



UNIVERSIDAD DE GRANADA  
FACULTAD DE CIENCIAS

**Diversidad taxonómica y funcional de la comunidad procariótica de la rizosfera de plantas xerófitas del altiplano central de México: Una aproximación metagenómica**

*Gloria Torres Cortés*

Tesis Doctoral  
2010

Editor: Editorial de la Universidad de Granada  
Autor: Gloria Torres Cortés  
D.L.: GR 3484-2010  
ISBN: 978-84-693-5235-9



# **Diversidad taxonómica y funcional de la comunidad procariótica de la rizosfera de plantas xerófitas del altiplano central de México: Una aproximación metagenómica**

Memoria que presenta la licenciada en Biología  
Dña. Gloria Torres Cortés  
como aspirante al grado de Doctor

Fdo: ***Gloria Torres Cortés***

VºBº de los Directores de la Tesis Doctoral

Fdo: Dr Francisco Martínez-  
Abarca Pastor  
Doctor en Ciencias Biológicas  
Investigador Científico del CSIC

Fdo.: Dr Nicolás Toro García  
Doctor en Ciencias Biológicas  
Profesor de Investigación del  
CSIC

Universidad de Granada  
2010



Este trabajo de Tesis Doctoral ha sido realizado en el Departamento de Microbiología y Sistemas Simbióticos (Grupo de investigación “Ecología Genética de la Rizosfera”) de la Estación Experimental del Zaidín (CSIC-Granada).

Para la realización del presente trabajo la Lda. Gloria Torres Cortés disfrutó de las siguientes formas de financiación:

I) Beca predoctoral asociada al proyecto BBVA BIOCON 04-084 Nov 2005.

II) Beca I3P-CSIC, Enero 2006- Diciembre 2009.

Becas concedidas por el CSIC para estancias breves en:

- Instituto de Ciencias del Mar- CSIC Barcelona. En el Laboratorio de Ecología Marina con la Dra Silvia González-Acinas. Periodo: Junio-Agosto 2008.
- Mikrobiologie & Biotechnologie, Biozentrum Klein Flottbek, Universität Hamburg. En el laboratorio del Profesor: Wolfgang Streit. Mayo-Agosto de 2009.

Los resultados mostrados en esta tesis doctoral se han presentado en comunicaciones en tres congresos Internacionales (una de ellas seleccionada para comunicación oral) y otros tres congresos Nacionales (dos de ellos como comunicación oral). Además están dando lugar a los siguientes trabajos en vías de ser publicados en Revistas Internacionales:

1) **Torres G.**, Millán V., Ramírez-Saad H., Toro N., Martínez-Abarca F. “Novel genes involved in Antibiotic Resistance found by functional metagenomics reveal significant differences between soil samples of distinct origin”. En proceso de revisión.

2) **Torres G.**, Millán V., Ramírez-Saad H., Toro N., Martínez-Abarca F. “A Novel Trimethoprim resistance gene from a metagenomic soil DNA library relates a Dihydrofolate reductase activity with oxoacyl-(acyl-carrier-protein) reductases protein family” En preparación.

3) Aguirre-Garrido, JF, **Torres, G.** Daniel Montiel-Lugo, D., Hernández-Rodríguez, C.H., Millán V., Martínez-Abarca, F., Toro, N. and Ramírez-Saad, H.C. “Bacterial community structure in the rhizosphere of cactus species growing in xerophytic highlands in Central Mexico”. En preparación.



*A mis padres y hermana*





<b>Introducción</b>	
<b>1.EL SUELO COMO ECOSISTEMA</b>	3
<b>2.PAPEL MICROORGANISMOS DEL SUELO</b>	5
<b>3.LA RIZOSFERA</b>	7
<b>4.ECOLOGÍA DE LOS MICROORGANISMOS EDÁFICOS</b>	9
<b>5.RESERVA DE la BIOSFERA DE TEHUACÁN-CUICATLÁN</b>	11
<b>6.NUEVAS TÉCNICAS DE BIOLOGÍA MOLECULAR</b>	15
6.1.Marcador filogenético universal: gen ribosomal 16S	16
6.2.Análisis de la diversidad microbiana mediante técnicas de “fingerprinting”.	18
6.3. Análisis de la diversidad microbiana mediante secuenciación	20
6.3.1. <i>Métodos de secuenciación masiva:tecnología 454</i>	
<i>pirosecuenciación</i>	21
6.3.2. <i>Bases de datos de genes 16S. Ribosomal Data base Project</i>	24
<b>7.LA METAGENÓMICA: UNA HERRAMIENTA PARA EL ESTUDIO DE LOS MICRRORGANISMOS PERESTES EN EL SUELO</b>	26
<b>8.ACCESO A LA “DIVERSIDAD ESCONDIDA” EN LA RIZOSFERA</b>	29
<b>8.1. Ecología del phylum <i>Acidobacteria</i> en el suelo</b>	32
<b>9. METAGENÓMICA COMO HERRAMIENTA PARA LA BÚSQUEDA DE NUEVAS ENZIMAS</b>	34
<b>9.1. Búsqueda de nuevas enzimas con actividad lipasa en genotecas de suelos</b>	34
9.1.2. Papel de las lipasas/esterasas en el suelo	34
9.1.2. Propiedades generales de las lipasas	36
9.1.3. Estructura y mecanismo de acción de las enzimas con carácter lipolítico	38
9.1.4. Distribución y clasificación de las carboxilesterasas y lipasas bacterianas	40
9.1.5. Aplicación industrial de las lipasas	43
<b>10.. EL SUELO COMO RESERVORIO DE NUEVOS GENES DE RESISTENCIA A ANTIBIÓTICOS</b>	44
10.1. Mecanismos de resistencia a antibióticos	46
10.2 Concepto de resistoma	48
<b>OBJETIVOS</b>	55

<b>MATERIAL Y MÉTODOS</b>	57
<b>1.TOMA DE MUESTRA</b>	59
<b>1.1.GEOGRAFÍA, OROGRAÍA Y TOPOLOGÍA</b>	59
<b>1.2.CLIMA</b>	60
<b>1.3.TIPO DE SUELO</b>	61
<b>1.4.VEGETACION</b>	61
<b>1.5.TOMA DE MUESTRAS EN EL CAMPO</b>	62
<b>1.6.NOMENCLATURA DE LAS MUESTRAS</b>	63
<b>2.ESPECIES Y CEPAS BACTERIANAS</b>	64
<b>3.PLÁSMIDOS Y VECTORES DE CLONAJE</b>	65
<b>4.OLIGONUCLEÓTIDOS</b>	67
<b>5.MEDIOS DE CULTIVO BACTERIANO Y ANTIBIÓTICOS</b>	68
5.1.MEDIOS DE CULTIVO Y ANTIBIÓTICOS	68
5.2.AISLAMIENTO DE CEPAS DE SUELO CON ACTIVIDAD LIPOLÍTICA	70
5.3.CONDICIONES DE CULTIVO	70
5.4.CONSERVACIÓN DE CULTIVOS BACTERIANOS	71
<b>6.AISLAMIENTO DE ÁCIDOS NUCLEICOS</b>	71
6.1.EXTRACCIÓN DE ADN DE MUESTRAS DE SUELO	71
6.2.EXTRACCIÓN DE ADN PLASMÍDICO	72
6.2.1. Extracción de ADN plasmídico por precipitación con sales de magnesio	72
6.2.2 Aislamiento de ADN plasmídico mediante kit comercial	73
6.4. CUANTIFICACIÓN DE LA CONCENTRACIÓ DE ADN	73
<b>7. MANIPULACIÓN Y AMPLIFICACIÓN DE ADN</b>	73
7.1 DIGESTIÓN CON ENZIMAS DE RESTRICCIÓN	73
7.2 REACCIÓN DE DESFOSFORILACIÓN	74
7.3 PCR.	74
7.3.1. PCR de colonia	75
<b>8. ELECTROFORESIS DE ÁCIDOS NUCLEICOS</b>	76
8.1. ELCETROFORESIS EN GELES DE AGAROSA NO DESNATURALIZANTES	76
8.2. ELECTROFORESIS EN GELES DE POLIACRILAMIDA NO DESNATURALIZANTES	76
8.3. MARCADORES DE PESO MOLECULAR	77
8.4. ELECTROFORESIS EN GEL EN GRADIENTE DE TEMPERATURA (TGGE)	77

<b>9. CLONAJE DE FRAGMENTOS</b>	78
9.1. AISLAMIENTO DE GEL DE AGAROSA	78
9.2. LIGACIÓN	79
9.2.1 <i>Clonaje de fragmentos amplificados por la reacción de PCR.</i>	79
9.3. PREPARACIÓN DE CÉLULAS COMPETENTES	80
9.3.1. <i>Células competentes por métodos químicos: RbCl</i>	80
9.3.2. <i>Células electrocompetentes</i>	81
9.4. TRANSFORMACIÓN DE CÉLULAS COMPETENTES.	81
9.4.1. Transformación por choque térmico	81
9.4.2. Electrotransformación	82
9.5. GENERACIÓN DE GENOTECAS 16S ARNr	82
9.5.1. <i>Análisis de las secuencias de los genes ribosomales 16S procedentes de la construcción de las librerías metagenómicas</i>	83
<b>10. PREPARACIÓN DE AMPLICONES PARA LA REACCIÓN DE PIROSECUENCIACIÓN</b>	84
10.1. Análisis de las secuencias procedentes de pirosecuenciación	86
<b>11. CONSTRUCCIÓN DE LIBRERÍAS METAGENÓMICAS EN FAGOS</b>	87
11.1. PREPARACIÓN DEL ADN MEDIOAMBIENTAL	87
11.2. CONSTRUCCIÓN DE LAS GENOTECAS	87
11.3. TITULACIÓN DE LAS GENOTECAS	90
11.4. AMPLIFICACIÓN DE LAS GENOTECAS	90
11.5. ESCISIÓN DE LAS GENOTECAS	91
11.5.1. <i>Escisión de una sola unidad formadora de calva</i>	91
11.5.2. <i>Escisión en masa</i>	93
<b>12. PREPARACIÓN Y ANÁLISIS DE PROTEÍNAS</b>	94
12.1 DETERMINACIÓN DEL CONTENIDO PROTEICO	94
12.2.ELECTROFORESIS DE PROTEÍNAS EN GEL EN CONDICIONES DESNATURALIZANTES (SDS-PAGE)	95
12.2.1. <i>Tinción de plata</i>	96
12.2.2. <i>Tinción con azul Coomassie</i>	97
12.2.3. <i>Zimogramas para la actividad esterasa</i>	97
12.2.4 <i>Secado de los geles de poliacrilamida</i>	98
<b>13. EXPRESIÓN Y PURIFICACIÓN DE PROTEÍNAS</b>	99
13.1. Expresión y purificación de proteínas en <i>E. coli</i> desde el vector pBCK-CMV	99
13.1.1 <i>Expresión en E. coli y preparación de las muestras de proteínas</i>	99

13.1.2. <i>Proceso de purificación</i>	100
1.3.2. Expresión y purificación de proteínas en <i>E. coli</i> desde el sistema Pet	101
13.2.1. <i>Expresión en E. coli y preparación de las muestras de proteínas</i>	101
13.2.2. <i>Proceso de purificación</i>	102
13.3. Cromatografía analítica de filtración en gel	103
<b>14. DETERMINACIÓN DE ACTIVIDADES ENZIMÁTICAS</b>	104
14.1. Ensayos enzimáticos relacionados con actividades lipolíticas	104
14.1.1. <i>Búsqueda de genes con actividad lipolítica en las genotecas metagenómicas</i>	104
14.1.1.1 <i>Ensayos enzimáticos de especificidad de sustrato</i>	105
14.1.1.2 <i>Ensayos enzimáticos para la determinación del pH óptimo, temperatura óptima y termoestabilidad de las enzimas esterasa</i>	106
14.1.1.3 <i>Efecto de detergentes no iónicos, de solventes orgánicos y de los iones sobre la actividad enzimática</i>	106
14.2. Ensayos enzimáticos relacionados con actividad dihidrofolato reductasa	107
<b>15. SISTEMA DE MUTAGÉNESIS DIRIGIDA</b>	107
<b>16. ANÁLISIS <i>in silico</i></b>	108
16.1. Secuenciación	108
16.2. Comparación e identificación de secuencias	108
16.3. Análisis filogenético	109
<b>17. DETERMINACIONES DE BIODIVERSIDAD</b>	109
17.1. Diversidad alfa	109
17.1A. Medición de la riqueza específica	109
17.1.A1. <i>Riqueza de especies (S)</i>	109
17.1.A2. <i>Curvas de rarefacción</i>	110
17.1.A3. <i>Índice de Chao</i>	110
17.1B. Medición de la estructura	111
17.1.B1. <i>Índice de Equidad de Shannon-Wiener (H')</i>	111
17.1.B2. <i>Índice de Pielou</i>	111
17.2. Diversidad beta	111
17.2.1 <i>Índice de Simpson</i>	111
17.2.2. <i>Coeficiente de similitud de Jaccard</i>	112
17.2.3. <i>Coeficiente de similitud de Sørensen</i>	112

<b>CAPÍTULO I</b>	113
<b>I1. CLASIFICACIÓN Y TOMA DE LAS MUESTRAS DE SUELO RIZOSFÉRICO.</b>	
I1.1. <i>Extracción de ADN de las muestras de suelo recogidas</i>	115
I1.2. Análisis de la diversidad procariota ( $\alpha$ -proteobacterias) mediante el perfil genético (adn fingerprint) obtenido con la técnica electroforesis en gel en gradiente de temperatura (TGGE)	117
<b>I2. Análisis de la diversidad procariota presente en la rizosfera de <i>Mammillaria carnea</i> y estudio de su estacionalidad</b>	122
I2.1. Análisis de la diversidad procariota en la rizosfera de <i>Mammillaria carnea</i> mediante la construcción de las librerías 16s	122
I2.1.1. Análisis de las secuencias	123
I2.1.2. Riqueza observada	125
I2.1.3. Índices de equidad y riqueza específica (índice de Chao)	126
I2.1.4. Diferencias entre la diversidad procariota presente en época seca y época lluviosa	127
<b>I3. Análisis de la diversidad procariota presente en la rizosfera de <i>Mammillaria carnea</i> en época lluviosa y seca mediante pirosecuenciación (tecnología 454)</b>	128
I3.1. Preparación de la muestra objeto de la pirosecuenciación	129
I3.2 Análisis de las secuencias y definición de Filotipos	129
I3.3. Análisis de la estructura de la comunidad microbiana	132
I3.4 Índices de equidad y riqueza específica	132
I3.5 Comparación de las dos zonas de muestreo	133
I3.5.1. <i>Número de OTUs comunes entre época seca y lluviosa</i>	133
I3.5.2. <i>Phyla diferenciados entre época seca y época lluviosa</i>	133
<b>I4. Análisis de la comunidad de acidobacterias presente en la rizosfera de <i>Mammillaria carnea</i> y en suelo suelto</b>	141
I4.1. Construcción de las genotecas específicas del phylum <i>Acidobacteria</i>	141
I4.2. Análisis de las secuencias	142
I4.3. Riqueza observada	142
I4.4 Índices de equidad y riqueza específica	143
I4.5 Comparativa de las librerías	143
<b>DISCUSIÓN CAPÍTULO I</b>	147
<b>ANEXO CAPÍTULO I</b>	162

<b>II BÚSQUEDA DE ENZIMAS DE INTERÉS INDUSTRIAL Y DE BIOMARCADORES EN LA RIZOSFERA DE MAMMILLARIA CARNEA</b>	163
II1. Construcción de genotecas metagenómicas en fagos	165
II2. El suelo, un reservorio potencial de genes de resistencia a antibióticos no estudiados	167
II2.1. <i>Búsqueda de genes de resistencia a antibióticos en las genotecas metagenómicas</i>	167
II2.2 <i>Medida de la concentración mínima inhibitoria de los genes de resistencia encontrados</i>	168
II2.3. Análisis de los genes de resistencia a antibióticos encontrados	169
II2.3.1. <i>Análisis de los clones de resistencia a ampicilina</i>	170
II2.3.2. <i>Análisis de los clones de resistencia a gentamicina</i>	173
II2.3.3. <i>Análisis de los clones de resistencia a cloranfenicol</i>	176
II2.3.4. <i>Análisis de los clones de resistencia a Trimetroprima</i>	179
II2.4. Análisis del clon de resistencia a trimetroprima: tm8.4. Caracterización de una nueva enzima con actividad dihidrofolato reductasa	182
II2.4.1 <i>Análisis de la secuencia del clon Tm8.4</i>	182
II2.4.2 <i>Purificación de la proteína Tm8.4</i>	185
II2.4.3 <i>Características bioquímicas de la proteína Tm8.4</i>	188
<b>DISCUSIÓN CAPÍTULO II2</b>	191
<b>ANEXO CAPÍTULO II2</b>	198
<b>II3 BÚSQUEDA DE ENZIMAS LIPOLÍTICAS DE INTERÉS INDUSTRIAL EN LAS GENOTECAS</b>	205
II3.1. <i>Estudio de bacterias cultivables con actividad lipolítica</i>	205
II3.2. Búsqueda de genes con actividad lipolítica en la genoteca metagenómica EEZ	208
II3.2.1. <i>Análisis de las secuencias de los genes de los clones con actividad lipolítica</i>	209
II3.2.2. <i>Relaciones filogenéticas de las enzimas aisladas</i>	212
II3.2.3. <i>Caracterización bioquímica de las enzimas encontradas</i>	216
II3.2.3.1 <i>Ensayos de temperatura y pH óptimos</i>	217
II3.2.3.2 <i>Determinación de las constantes cinéticas de las enzimas con carácter lipolítico para diferentes tipos de sustratos</i>	217
II3.2.3.3 <i>Estabilidad frente a solventes orgánicos y a detergentes</i>	219

II3.2.4. Purificación de enzimas esterases y caracterización bioquímica	222
II3.2.4.1. Características bioquímicas de las enzimas esterases purificadas	225
II3.2.5. Estudio de la resistencia a solventes de la enzima pE2 frente a la enzima pE1	227
II3.2.5.1 ESTUDIOS PRELIMINARES	227
II3.2.5.2 Análisis de la estructura primaria, secundaria y terciaria de las enzimas pE1 y pE2	228
II3.2.5.3 Construcción de las quimeras pE1Q y pE2Q	230
II3.2.5.4. Sobre-expresión de las quimeras	232
II3.2.5.5. Análisis de la actividad de las quimeras	232
II3.2.6 Caracterización de la enzima pE12	234
II3.2.6.1. Sobreexpresión y purificación de la enzima pE12	235
II3.2.6.2 Caracterización enzimática	236
II3.3 Búsqueda de genes con actividad lipolítica en las genotecas construidas a partir de muestras de ADN de la rizosfera de <i>mamillaria carnea</i>	239
II3.3.1. Subcloning y análisis de las secuencias	239
II3.3.2. Caracterización bioquímica preliminar	244
<b>DISCUSIÓN</b>	245
<b>ANEXO CAPÍTULO II3</b>	255
<b>II.4. ANÁLISIS DE UN BIOMARCADOR EN LA RIZOSFERA DE <i>Mammillaria carnea</i></b>	263
II4.1. Análisis de los clones 8.25 Y 6.2	263
II4.2. Uso de los clones 8.25 y 6.2 como posibles bioindicadores o biomarcadores	264
<b>DISCUSIÓN</b>	269
<b>ANEXO CAPÍTULO II4</b>	272
<b>CONCLUSIONES</b>	277
<b>BIBLIOGRAFÍA</b>	283





# **INTRODUCCIÓN**



## 1. EL SUELO COMO ECOSISTEMA

El suelo es el ecosistema terrestre que presenta mayor biodiversidad (Anderson, 1975; Usher *et al.*, 1979, Giller, 1996; Germida y Siciliano, 2001). En él conviven una gran cantidad de microorganismos, como bacterias, hongos, protozoos y algas (Zwolinski, 2007; Van Elsas *et al* 2007; tabla 1).

Pero entre los grupos de microorganismos presentes, el más abundante y diverso es el constituido por los procariotas, posiblemente por ser capaces de crecer más rápidamente y poder utilizar una gran variedad de compuestos como fuentes de carbono y de nitrógeno. Las bacterias se localizan en la superficie de las partículas del suelo, pudiendo interactuar con las raíces de las plantas, ya que en esa zona existe una alta concentración de nutrientes.

<b>Microorganismo</b>	Número estimado (por gramo de peso seco)
<b>Virus</b>	$10^{10}$
<b>Bacterias</b>	$10^6$ - $10^{10}$
<b>Actinomicetos</b>	$10^8$
<b>Hongos</b>	$10^6$
<b>Algas</b>	$10^6$
<b>Protozoos</b>	$10^5$
<b>Nematodos</b>	$10^2$

**Tabla 1.** Estimación del número de microorganismos (Maier *et al.* 2000; Metting 1993; Sylvia *et al.* 2005; Trevors and Cook 1992; VanElsas *et al.* 2007)

Una comunidad bacteriana edáfica, normalmente contiene entre  $10^7$  y  $10^9$  células viables y cultivables por gramo de suelo, incluyendo bacterias autotróficas y heterotróficas, que pueden agruparse en unos 40 fenotipos y 4000 genotipos (Torsvik *et al.*, 1990). Asumiendo que los individuos con una homología del ADN de más del 70% pertenecen a la misma *especie* (Wayne *et al.*, 1987), estos autores sugieren que puede haber entre 20.000 y 40.000 *especies* bacterianas por gramo de suelo.

En el suelo, se encuentra una mayor cantidad de géneros y especies de microorganismos que en otros hábitats microbianos. Algunas de estas especies se encuentran en poca cantidad, probablemente debido a que las

condiciones nutricionales y los factores fisico-químicos ambientales que requieren para su supervivencia y crecimiento están restringidos a determinados lugares. En otros casos las poblaciones bacterianas se pueden diversificar mediante el fenómeno descrito como variación de fase, por el cual se dan modificaciones fenotípicas y genotípicas dentro de una población de manera que se forman subpoblaciones. Este proceso se ha relacionado con la adaptación a condiciones ambientales diversas, así como a cambios bruscos en el ecosistema (Saunders, 1994; Davidson *et al* , 2004) que ofrecen a las diferentes subpoblaciones una ventaja para poder sobrevivir en mayor cantidad de hábitats y hacer frente a los posibles cambios (Dybvig, 1993).

A pesar de ello, atendiendo al número y a la heterogeneidad microbiana existente en el suelo, éste parece ser un ambiente apto para estos organismos que ya se encuentran adaptados a este medio austero. Este hecho es debido a que las bacterias poseen, quizás más que otros organismos, una gran adaptabilidad, tanto fisiológica como genética (transferencias horizontales y verticales, elevada tasa de mutación, variación de fase; Davidson *et al* 2004; Smets y Barkay, 2005; Smalla y Heuer, 2006; Huer *et al*, 2008), frente a la variación de las condiciones.

El suelo, por su enorme diversidad, parece un ecosistema ideal a la hora de estudiar ecología microbiana; sin embargo, es también, probablemente, el medio que más exigencias técnicas presenta a los microbiólogos. Esto se debe no solo al tamaño de su comunidad microbiana y a la enorme diversidad de especies presente, sino también a la naturaleza multifásica del medio (que incluye gases, agua y materiales sólidos) y a las cambiantes características químicas y biológicas del medio. Todos estos factores provocan una extrema heterogeneidad espacial y temporal en la microbiología del suelo.

El suelo difiere de otros hábitats microbianos en que la proporción sólido/líquido es muy alta, dominando la fase sólida, que está formada por partículas de diferentes tamaños rodeadas por fases acuosas y gaseosas, cuyo porcentaje y composición fluctúa en el tiempo y en el espacio.

La fase sólida es un sistema compuesto por minerales, residuos de plantas, animales y microorganismos en diferentes fases de descomposición, y la microbiota viva.

La fase acuosa, en la que se encuentran las sustancias orgánicas e inorgánicas que sirven como nutrientes o inhibidores de los microorganismos, está disuelta y se dispone de forma discontinua en el suelo, excepto cuando está saturado.

El espacio que no está ocupado por la fase acuosa, se *rellena* con aire u otros gases ( $O_2$ ,  $CO_2$ ,  $CH_4$ ,  $CO$ ,  $NO_x$ , ácidos grasos de cadena corta, alcoholes, aldehidos y compuestos aromáticos), siendo unos desplazados por otros y afectando de este modo a la composición y actividad de los microorganismos (por ejemplo, aerobios frente a aerobios facultativos, anaerobios o microaerófilos). De hecho, las concentraciones de  $O_2$  y  $CO_2$  varían de un suelo a otro, y de una porción de suelo a otra, dependiendo de la temperatura, humedad (Yamaguchi *et al.*, 1967), contenido en materia orgánica (Abrosimova y Revut, 1964), profundidad del suelo (Wood *et al.*, 1993) y tipo de cultivo (Buyanovsky y Wagner, 1983).

El suelo es, además, un medio en el que hay una baja disponibilidad de nutrientes, especialmente de fuentes de carbono y otros recursos energéticos, que llegan incluso a ser limitantes (o restrictivos) para determinados microorganismos edáficos.

Todas estas características hacen que en el suelo se formen **microhábitats**, ya que es sólo en determinados lugares en los que hay nutrientes y un continuo aporte hídrico, donde los microorganismos son metabólicamente más activos. Esta y es la cuestión clave de la diversidad de los mismos en el suelo (Numan *et al.*, 2002, 2003; Zhou *et al.*, 2002). La arquitectura del suelo, por tanto, define el hábitat a colonizar por los microorganismos, y las características de los espacios porosos determina las relaciones entre los organismos que habitan en ellos (Young an Rizt, 2000; Ranjard and Richaume, 2001).

## **2. EL PAPEL MICROORGANISMOS DEL SUELO**

Los microorganismos del suelo juegan un papel clave en los ciclos biogeoquímicos del carbono y del nitrógeno, en el mantenimiento de los ecosistemas naturales y en la producción agrícola.

Las superficies de los suelos están ocupadas por poblaciones indígenas de arqueas, bacterias, hongos, algas y protozoos. También son numerosos

otros microorganismos como los virus (tabla 1), aunque poco se conoce de este tipo de organismos.

Además de estas poblaciones indígenas, se pueden introducir microorganismos a través de la actividad humana, por ejemplo, la introducción deliberada de bacterias como agentes de control biológico o como agentes biodegradativos, o la introducción indirecta de las mismas como resultado de las prácticas en agricultura. Los animales también introducen microorganismos, como por ejemplo mediante los excrementos. Dependiendo de la fuente, la introducción de nuevos microorganismos a un suelo raramente supone un cambio en la abundancia y distribución de las poblaciones indígenas.

La población microbiana del suelo, por tanto, es crítica para el mantenimiento de la función del mismo, tanto en suelos naturales como agrícolas, por su implicación en procesos clave como la formación de la estructura del suelo, la descomposición de la materia orgánica, la eliminación de toxinas y los ciclos de carbono, nitrógeno, fósforo y azufre (van Elsas y Trevors, 1997, Canbolat *et al.*, 2006; Kohler *et al.*, 2006). Además, los microorganismos juegan un papel crucial en la defensa de enfermedades y en la promoción del crecimiento de las plantas, así como en los cambios en la vegetación (Doran *et al.*, 1996; Garbeva *et al.*, 2004).

Por tanto, los microorganismos del suelo tienen un papel muy importante en el mantenimiento de la calidad del suelo en sistemas agrícolas (Kennedy *et al.*, 1995). Las características microbianas del suelo pueden indicar cambios en la disponibilidad de recursos, la estructura del suelo o la contaminación y pueden representar una clave importante para entender el impacto de factores ambientales y antropogénicos (DeLong y Pace, 2001; Tiedje *et al.*, 1999; Pankhurst *et al.*, 1996).

La diversidad de los microorganismos del suelo puede representar la capacidad de un suelo para afrontar las perturbaciones (Bardgett, 2002; Johnsen *et al.*, 2001; Kennedy, 1999). La alta densidad y diversidad que caracterizan a la microbiota del suelo, hacen que se adapte rápidamente a cambios en las condiciones ambientales por medio de ajustes en i) las tasas de actividad, ii) biomasa y iii) estructura de la comunidad, tres parámetros a tener en consideración para evaluar la calidad de un suelo (Schloter *et al.*, 2003). Por tanto, los análisis de diversidad, actividad y estructura de la

comunidad de microorganismos del suelo parecen ser esenciales cuando se monitorizan factores ambientales que influyen en la calidad del suelo.

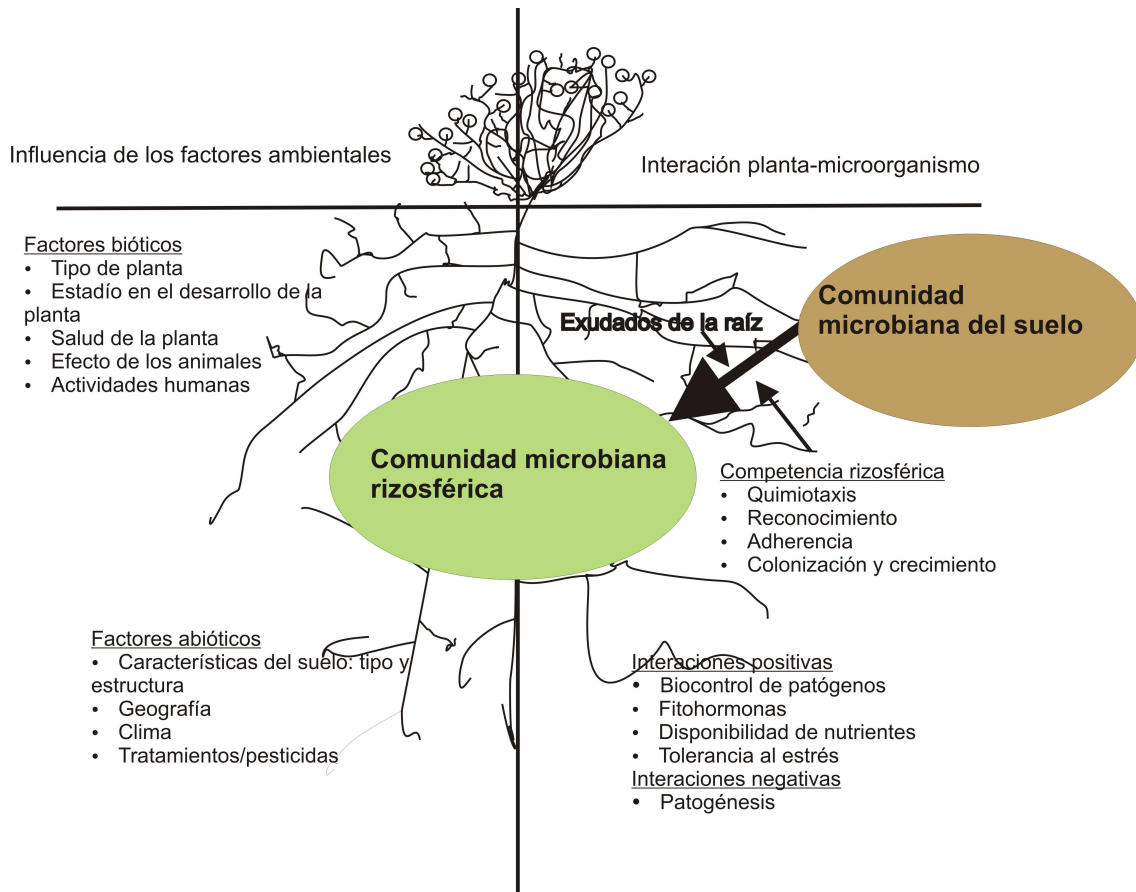
### **3. LA RIZOSFERA**

El término “rizosfera” fue acuñado por Hiltner en 1904 para describir la porción de suelo bajo la influencia de la raíz en la cual tienen lugar numerosos procesos mediados por microorganismos. El papel de la rizosfera es vital para la salud, calidad y nutrición de la planta. En comparación con el suelo suelto, en la rizosfera la biomasa y la actividad microbiológica son elevadas debido a los compuestos exudados por la raíz. (Sorensen, 1997; Raaijmakers *et al.*, 2009).

De este modo, el término “rizosfera” define el compartimiento edáfico directamente bajo la influencia de las raíces (Kent y Triplett, 2002). En este nicho, existen complejas comunidades microbianas que son mantenidas por los nutrientes exudados por la raíz (principalmente azúcares, ácidos orgánicos y aminoácidos; Smalla *et al.*, 2001; Kowalchuk *et al.*, 2002), mucílago y células radiculares en descomposición (Curl y Truelove, 1986). Así, la rizosfera induce un cambio característico en la distribución y actividad de los microorganismos asociados a las raíces en comparación con los del suelo (Berg *et al.*, 2002; 2005; Gomes *et al.*, 2003; ZaharHaichar *et al.*, 2008), denominado efecto rizosférico. La composición de las poblaciones rizosféricas depende de la composición del exudado radicular, así como de la especie vegetal, tipo de raíz, edad de la planta, tipo de suelo e historial del suelo (Anderson *et al.*, 1993, Gabriele Berg y Kornelia Smalla 2009). Sin embargo, por regla general, se suele observar una dominancia de bacilos gram negativos (como por ejemplo, cepas del género *Pseudomonas*).

El “efecto rizosférico” fue descrito por primera vez por Lorenz Hiltner (Hiltner, 1908), que observó un efecto “estimulante” de las raíces sobre las poblaciones microbianas del suelo. A su vez, los microorganismos rizosféricos tienen un enorme impacto sobre las plantas, ya que se han aislado una enorme variedad de microorganismos que poseen efectos estimulantes o inhibitorios sobre la salud de las mismas (Campbell, 1990). De hecho, la utilización de inóculos bacterianos promotores del crecimiento vegetal o su empleo en el biocontrol de enfermedades vegetales se ha convertido en un enorme campo de estudio.





**Figura 1.** Esquema de las interacciones que influyen a los microorganismos rizosféricos. En parte de la izquierda, se señalan los factores bióticos y abióticos que afectan a la población microbiana de la rizosfera (como el tipo de planta, estructura del suelo...). En la parte de la porción superior derecha, se presentan los elementos capaces de seleccionar a las poblaciones rizosféricas (exudados de la raíz y competencia) y en la parte inferior los beneficios y perjuicios que obtienen las plantas de esta interacción.

El uso de aproximaciones polifásicas que se han venido desarrollando en la última década y que combinan las técnicas moleculares con técnicas más tradicionales para el estudio de las poblaciones microbianas, han dado numerosos resultados que nos ayudan a tener un mayor conocimiento tanto de la estructura como de la función de los microorganismos en la rizosfera.

La diversidad de las comunidades bacterianas rizosféricas está influida por numerosos factores tanto bióticos como abióticos (figura 1), por ejemplo, el clima y las estaciones, los animales, el uso de pesticidas, el tipo y la estructura del suelo, el tipo de planta... (Lemanceau *et al*, 1995; Siciliano *et al*, 2001; Graner *et al*, 2003; Garveba *et al*, 2004; Jousset *et al*, 2006, Rasche *et al*, 2006).

Se asume que los distintos tipos de suelos poseen una comunidad microbiana específica de los mismos y, además, en los últimos tiempos se

están describiendo estudios a escala continental de este tipo de comunidades (Fierer y Jackson, 2006). En contraste con lo que conocemos sobre los macroorganismos, la biogeografía microbiana está controlada principalmente por variables edáficas, fundamentalmente por el pH (Fierer y Jackson, 2006). El efecto de las plantas sobre la diversidad de las comunidades microbianas rizosféricas también ha sido muy estudiado; cada planta segrega unos exudados radiculares concretos que seleccionan microorganismos específicos. Estos exudados no solo proveen nutrientes específicos para los microorganismos, sino que en algunas plantas también contienen metabolitos antimicrobianos únicos. La existencia de una elevada diversidad de especies de plantas, que está estimada entre 310000 hasta 422000 (Pitman y Jörgensen, 2002) y la correspondiente diversidad de metabolitos secundarios (Buchanan *et al.*, 2000) afecta de forma considerable a la diversidad microbiana.

#### **4. ECOLOGÍA DE LOS MICROORGANISMOS EDÁFICOS**

La mayoría de las comunidades naturales son altamente diversas y parece que las comunidades de microorganismos no son una excepción (Curtis *et al.*, 2002; Dykhuizen, 1998; Kent y Triplett, 2002; Pace, 1997; Rappé y Giovannoni, 2003; Torsvik *et al.*, 2002, Woese, 1987).

El empleo de técnicas moleculares para el estudio de las comunidades microbianas de muestras ambientales (Olsen *et al.*, 1986; Pace *et al.*, 1986) ha revelado la existencia de una increíble variedad de genotipos y especies en todos los hábitats conocidos, muchos de los cuales parecen ser ecológicamente equivalentes en términos de requerimientos de recursos (Rappé y Giovannoni, 2003; Schloter *et al.*, 2000).

La existencia de tal diversidad requiere una explicación, ya que está en desacuerdo con lo esperado en la teoría de que la selección natural elimina todo excepto los tipos mejor adaptados bajo un conjunto de condiciones ambientales dadas. Si esto es así, ¿cómo se pueden mantener esos niveles de diversidad en las comunidades microbianas naturales?

Los ecólogos y los genetistas poblacionales se han aferrado a lo largo de muchos años a asuntos relacionados con el mantenimiento de la diversidad, haciendo que gran parte de la motivación de la investigación en ecología,

genética de poblaciones y evolución haya sido impulsada por la necesidad de explicar la paradoja de la diversidad (Hutchinson, 1961; Tilman, 2000).

Pese a las diferentes aproximaciones, tanto los ecólogos interesados en la diversidad de las especies en las comunidades, como los genetistas poblacionales centrados en la variedad de genotipos dentro de las poblaciones, tanto unos como otros, han identificado principios similares que gobiernan los patrones de la diversidad en ambientes naturales. Éste es el principio de exclusión de nicho (Gause, 1934; Hardin, 1960), que afirma que un solo nicho sólo puede mantener a un tipo, bien sea un genotipo o una especie.

Un nicho es la variedad de factores en un ambiente que limitan el crecimiento de un tipo con relación a los otros. Un factor limitante podría ser un recurso o nutriente esencial, un conjunto de condiciones físicas (como temperatura, pH o salinidad) o la existencia de refugios a la predación. El ambiente puede ser interpretado como el número de nichos disponibles para el número de tipos que se mantienen en un ambiente determinado por el número de nichos (Levins, 1968, Tilman, 1982).

Como ya se ha comentado, el suelo representa un ambiente heterogéneo para la microbiota que lo habita. Los organismos que residen en el suelo están expuestos a condiciones abióticas y nutricionales que pueden variar incluso por encima de la escala micrométrica, existirían muchos nichos habitados por diferentes tipos de organismos. En un sistema estable se podría especular que cada microhábitat está ocupado por organismos que fueron mejores en la colonización del nicho y se establecieron en él. En el conjunto de estos organismos subyace la catálisis de los procesos bioquímicos del suelo.

En el suelo existe un amplio rango de factores que afectan a la vida de los microorganismos (van Veen *et al.*, 1997), pero aún quedan por responder muchas preguntas clave. ¿Cuáles son los factores principales que controlan la estructura y diversidad de las comunidades microbianas del suelo?

Las tres hipótesis principales son:

1. El tipo de planta es el principal determinante de la estructura de las comunidades de microorganismos en el suelo, ya que las plantas son los principales proveedores de fuentes de carbono y energía específicas (Grayston *et al.*, 1996; Somers *et al.*, 2004).

2. El tipo de suelo es el principal determinante de la estructura de las comunidades, ya que la combinación de textura, estructura del suelo, materia orgánica, estabilidad de microagregados, pH y la

presencia de nutrientes clave determina los nichos habitables en el suelo. (Lucas García *et al.* 2001; Fierer y Jackson, 2006).

3. El régimen de manejo agrícola (como la rotación de cultivos, labranza, aplicación de herbicidas y fertilizantes e irrigación) es el determinante clave de la estructura de las comunidades microbianas del suelo (Garbeva *et al.*, 2004).

No hay un consenso claro sobre cuál de estos factores, las propiedades del suelo y el tipo de plantas, tiene más influencia sobre la estructura y función de las comunidades microbianas. En la literatura hay numerosos estudios contradictorios que indican que el factor dominante es la planta o la estructura del suelo (Grayston *et al.*, 1998; Girvan *et al.*, 2003; Numan *et al.*, 2005), sin embargo, está ampliamente estudiado que los dos provocan efectos sobre las mismas dependiendo de las condiciones bióticas o abióticas (Gabriele Berg y Kornelia Smalla, 2009).

## **5. RESERVA DE LA BIOSFERA DE TEHUCÁN-CUICATLÁN**

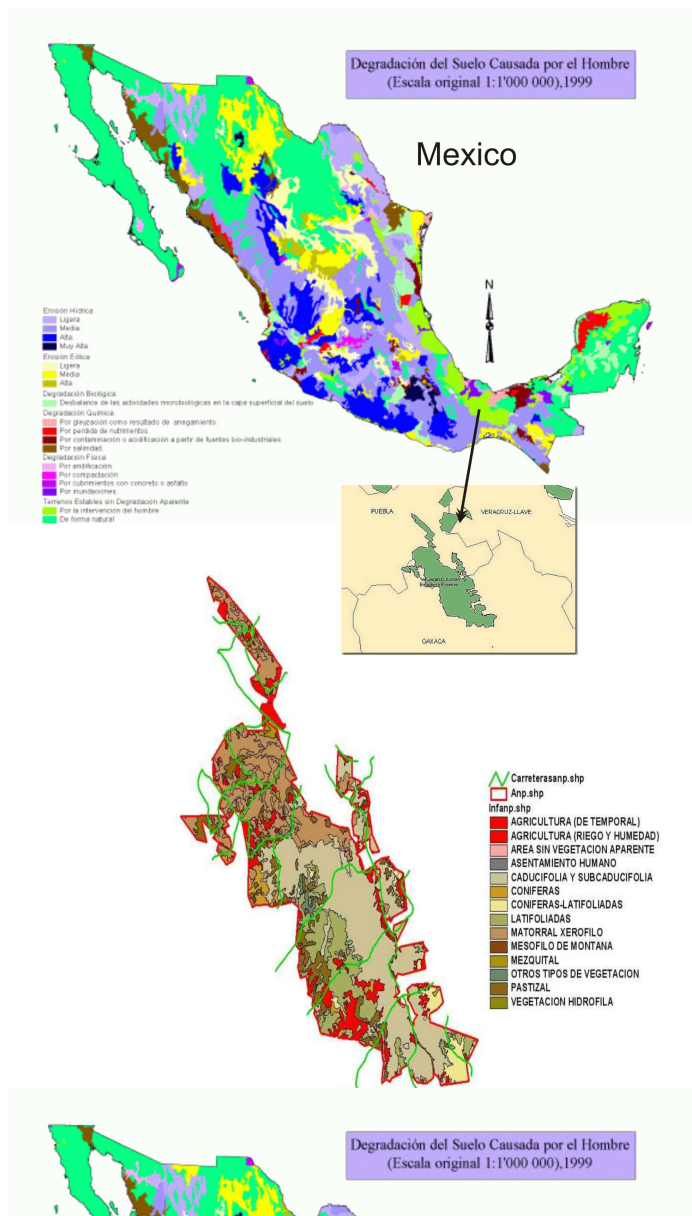
El estudio que se presenta en esta Tesis Doctoral se encuentra centrado en el análisis de la diversidad de los microorganismos edáficos presentes en la rizosfera de cactáceas situadas en la Reserva de la biosfera de Tehucán-Cuicatlán.

Esta provincia florística está ubicada en los estados de Puebla y Oaxaca (figura 2) y es considerada mundialmente una de las áreas de mayor biodiversidad. Como ejemplo, puede citarse que alberga más de 2700 especies de plantas vasculares, muchas de ellas consideradas endémicas. Por tal motivo, se proclamó Área Natural Protegida (ANP) con carácter de Reserva de la Biosfera una superficie de 490.186 ha, que abarca desde la Sierra de Tecamachalco, Puebla hasta Atatlahuaca en la Cañada Chica, Oaxaca.

Esta región es considerada una zona árida y semiárida por su baja precipitación pluvial, pero en realidad corresponde a una zona con clima semiárido (García, 1988). El régimen de lluvias es de verano con dos estaciones secas.

Desde el punto geológico pertenece a la Zona de la Cañada o de la Depresión de Tehuacán-Cuicatlán-Quiatepec. Esta zona es un largo y estrecho valle delimitado por fallas paralelas del tipo “riftoide”, rodeado por las sierras

Mixteca (a occidente), Zongolica y Juárez (a oriente) y de Ixtlán, angosta y dispuesta en sentido noroeste-sureste.



**Figura 2:** Localización geográfica de la Reserva de la Biosfera de Tehuacán-Cuicatlán. La Reserva de la Biosfera de Tehuacán-Cuicatlán fue decretada el 18 de septiembre de 1998, con una modificación el 23 de mayo de 1999. Abarca 4897 km<sup>2</sup>, en los cuales hay 51 municipios con 307 poblados y 35 223 personas. La reserva contempla la mayor parte de la provincia florística de Tehuacán-Cuicatlán, que se encuentra constituida principalmente por matorral xerófilo en donde predominan las cactáceas columnares (45 especies) que forman zonas de muy alta densidad (1100 ind./Ha).

A pesar de que los ecosistemas de zonas áridas cubren la mitad de la superficie total de México, su productividad biológica típicamente baja y su

poca densidad de individuos inducirían a pensar que su diversidad biológica es inferior a la mayoría de las demás zonas ecológicas de México.

Sin embargo, las zonas áridas mexicanas están compuestas por diferentes provincias florísticas que tienen sus taxa característicos y cada una contribuye con cerca de 6000 especies, lo que es equivalente a la selva subhúmeda (Rzedowski *et al*, 1991). Las zonas áridas y semiáridas no sólo son ricas en número de especies, sino que también se destacan por su alto grado de endemismo, calculado en cerca del 22% a nivel genérico en los pastizales y de hasta el 43% en matorrales xerófilos; a nivel específico es cercana al 60% (Golubov, 2001; Rzedowski *et al*, 1993).

Los recursos bióticos y abióticos en estos ambientes son escasos y están distribuidos de manera heterogénea durante la mayor parte del año, dependiendo en gran parte de los periodos muy variables de precipitación (Noy-Meir, J. 1973). La variación espacial y temporal de la precipitación unida a las temperaturas promedio elevadas y a las texturas de suelo dan lugar a una heterogeneidad ambiental muy marcada (Mandujano, M. 2002; Pianka, E 1967). Como las condiciones ambientales son altamente variables, la productividad primaria es baja (6000 a 40.000 kg/ha) y la distribución y abundancia de los niveles tróficos superiores es limitada (Huenneke, L 1996; Noy-Meir, J. 1973).

La Reserva de la Biosfera de Tehuacán-Cuicatlán fue decretada el 18 de septiembre de 1998, con una modificación el 23 de mayo de 1999, y abarca 4897 km<sup>2</sup> (figura 1). La reserva contempla la mayor parte de la provincia florística de Tehuacán-Cuicatlán, que se encuentra constituida principalmente por matorral xerófilo en donde predominan las cactáceas columnares (figura 3; 45 especies) que forman zonas de muy alta densidad (1100 ind./Ha).

Hasta 1990, en la zona se habían recolectado 2620 especies, pertenecientes a 911 géneros y 154 familias de plantas vasculares, y se espera encontrar un total de 2800 especies (Davila, P 1991). Dentro de esta flora extraordinariamente diversa, tres géneros son endémicos a la provincia y se calcula que, a nivel específico, hay un endemismo de cerca del 38%, por ejemplo las especies *Casimiroa calderoniae*, *Fouquieria purpusii*, *Solisia pectinata*, *Leucophyllum pringlei*, *Gypsacanthus nelsonii* y *Agave macroacantha*, algunas de las cuales están en peligro de extinción.

La familia *Cactaceae* ha experimentado una radiación secundaria en México, constituyéndose en el más importante de los dos mayores centros de

diversificación con cerca de 850 especies agrupadas en 52 géneros, el 84% de las cuales son endémicas (Arias-Montes, S. 1993).



**Figura 3.** Imagen panorámica de la Reserva de la biosfera de Tehuacán-Cuicatlán, constituida principalmente por material xerófito donde predominan las cactáceas columnares.

El estado de Oaxaca es un estado especialmente rico en especies de *Cactaceae* (112 especies en 28 géneros), que se encuentran distribuidas principalmente en dos zonas: la zona del istmo de Tehuantepec y la comprendida por el Valle de Tehuacán-Cuicatlán y la Mizteca Alta (Hernández, H. M 2004). De las especies identificadas en la zona de Tehuacán-Cuicatlán, encontramos que 24 de las 75 especies registradas en esta provincia son endémicas, por ej. *Mammillaria solisioides*, *M. pectinifera*, *M. hernandezii*, *M. kraehenbuehlii*, *M. huitzilopochtli*, *Neobuxbaumia tetezto*, *N. macrocephala* y *N. mezcalensis*, (Dávila, P., 1993).

Dada la alta diversidad específica de estas zonas, hay una creciente preocupación por la conversión de las zonas áridas, que puede llevar a procesos de desertificación y una consecuente pérdida de diversidad biológica (Huenneke, L 1996). En especial, las sabanas y zonas semiáridas de América del norte (geográficamente incluye a México) han sufrido un deterioro ambiental considerable en términos de composición de especies y estructura, debido primordialmente a la fragmentación y al uso agrícola y ganadero (Archer, S 1988). De ahí que resulte necesario determinar la importancia relativa de la interacción de las especies en estos ambientes para poder

cuantificar los costos ecológicos de la fragmentación y de la conversión del uso de suelo para proponer medidas adecuadas de manejo y conservación.

## 6. NUEVAS TÉCNICAS DE BIOLOGÍA MOLECULAR PARA EL ANÁLISIS DE LA DIVERSIDAD MICROBIOLÓGICA BASADAS EN EL ESTUDIO DE LA SUBUNIDAD PEQUEÑA DEL GEN RIBOSOMAL 16S

Numerosos estudios indican que, en ambientes naturales, más del 99% de los microorganismos no son realmente cultivables (tabla 2) y por tanto, no son accesibles para un análisis posterior (Amann *et al*, 1995). La principal limitación a la hora de describir nuevas especies en el ámbito de la microbiología radica en la necesidad del cultivo previo de las mismas. Aunque se está dedicando mucho esfuerzo al desarrollo de nuevas fórmulas de medios de cultivo que imiten en la mayor parte posible a las condiciones naturales, los avances en este sentido son muy limitados.

<b>Organismo</b>	<b>Especies estimadas</b>	<b>Accesibilidad (% del total)</b>
<b>Animales (mamíferos, aves peces)</b>	3,5.10 <sup>4</sup>	>90
<b>Artrópodos/invertebrados</b>	10 <sup>6</sup> -10 <sup>7</sup>	10
<b>Nematodos</b>	5.10 <sup>5</sup>	3
<b>Plantas superiores</b>	2,7.10 <sup>5</sup>	>90
<b>Algas</b>	10 <sup>4</sup> -10 <sup>5</sup>	70
<b>Hongos</b>	1,5. 10 <sup>6</sup>	5
<b>Bacterias</b>	10 <sup>4</sup> -10 <sup>5</sup>	1-10
<b>Arqueas</b>	10 <sup>5</sup> - 10 <sup>6</sup>	0,1-1
<b>Virus</b>	10 <sup>5</sup> - 10 <sup>6</sup>	4

**Tabla 2.** Número de especies estimadas en la naturaleza (Ferrer M, 2005).

El desarrollo de técnicas para el aislamiento y manipulación de ácidos nucleicos de un determinado nicho ecológico nos ha permitido acceder a esa diversidad “escondida”, permitiendo a su vez el estudio de cambios cuantificables en respuesta a fluctuaciones medioambientales tanto naturales (presencia de plantas, características del suelo, etc.) como causadas por el



hombre. De esta forma, la caracterización de esta diversidad microbiana adquiere un valor adicional, no sólo mediante la generación de colecciones de microorganismos cultivables, sino también por la creación de bancos genéticos que contienen el ADN medioambiental (**metagenoma**; Handelsman, 2004). Este tipo de enfoques puede tener un mayor interés sobre ecosistemas microbiológicos poco explorados como es el caso del ecosistema de la Reserva Natural de Tehucán -Cuicatlán.

Conseguir de forma colectiva todos los genomas microbianos presentes en un determinado hábitat es un objetivo de los programas de descubrimiento de nuevas enzimas y de los programas de ecología microbiana más básicos, encaminados a la utilización de estos recursos evitando el problema asociado al cultivo tradicional.

El uso de todas estas nuevas técnicas para el estudio de la distribución, diversidad y estructura de las poblaciones de microorganismos en diferentes ambientes, sin la limitación de los métodos tradicionales de cultivos, ha desembocado en el nacimiento de una nueva rama de la ecología, que es lo que actualmente denominada *Ecología Molecular*.

### **6.1. Marcador filogenético universal: el gen ribosomal 16s**

Gran parte de la estimación de esa diversidad “escondida” se debe al análisis de la secuencia de la subunidad pequeña del gen ribosomal (16S rARN; Woese et al, 1990), que permite obtener información acerca de los tipos de microorganismos presentes en un ambiente dado.

El estudio de este gen ha revolucionado la forma de trabajar en ecología microbiana y ha sido uno de los núcleos centrales en los trabajos efectuados en esta modalidad de los últimos 15 años; mucho de lo que conocemos hoy en día acerca de la diversidad y la estructura de la comunidad tanto de bacterias como de Archeas en diferentes ambientes, se lo debemos al estudio comparativo de la secuencia de esta subunidad pequeña.

La secuencia de estos genes ribosomales nos permite inferir relaciones evolutivas debido a que:

a) Son esenciales para la síntesis de proteínas, estando presente en todos los organismos, generalmente en varias copias.

b) Contiene regiones variables y otras altamente conservadas, tanto en su estructura primaria como secundaria, lo que permite el diseño de cebadores universales y específicos.

c) Parecen cambiar en su secuencia muy lentamente, porque son secuencias que se han mantenido muy uniformes a lo largo de la evolución (se pueden considerar como relojes moleculares).

Estas propiedades permiten establecer las diferencias filogenéticas, gracias a las cuales podemos establecer orígenes comunes y esquemas evolutivos. Las construcciones de estos árboles son conceptualmente simples. Pares de secuencias de rRNA 16s de organismos distintos se compararan, las diferencias se valoran y se toma como medida de la distancia evolutiva entre esos organismos.

Para obtener la secuencia del gen 16S rARN amplificada mediante la técnica de PCR (reacción en cadena de la polimerasa), se pueden usar diferentes cebadores (tabla 3). Gracias a las 9 regiones variables (V1-V9) que posee este gen, durante los últimos años se han ido diseñando una amplia batería de cebadores para la amplificación de estas zonas variables (DarSarma y Feleischmann, 1995; Elwood *et al*, 1985; Kolganova *et al*, 2002; Watanabe *et al*, 2001). Como ya se ha comentado, algunos de estos cebadores han sido diseñados específicamente para amplificar sólo algunos taxones, mientras que otros se han diseñado para amplificar la secuencia del gen 16S rARN de todos los procariotas, cebadores que normalmente se denominan como “universales” (tabla 3, Baker G.C. *et al*, 2003).

Nombre <sup>a</sup>	Secuencia 5'-3'	N <sup>b</sup>	Ta <sup>b</sup>	B <sup>c</sup>	A <sup>c</sup>	E <sup>c</sup>	Ref <sup>d</sup>
<b>E8F</b>	AGAGTTTGATCCTGGCTCAG	20	55	+++	+	-	1
<b>E9F</b>	GAGTTTGATCCTGGCTCAG	19	53	+++	+	++	2
<b>E334F</b>	CCAGACTCCTACGGGAGGCAGC	21	65	+++	-	-	3
<b>E341F</b>	CCTACGGGIGGCIGCA	16	51	+++	+	-	4
<b>E786F</b>	GATTAGATACCCTGGTAG	18	47	+++	+	-	5
<b>E533R</b>	TIACCGIIICTICTGGCAC	19	56	+++	+	++	6
<b>E926R</b>	CCGICIATHTTTIAGTTT	20	50	+++	+++	++	7
<b>E939R</b>	CTTGTGCGGGCCCCGTCAATTC	23	71	+++	-	-	8
<b>E1115R</b>	AGGGTTGCGCTCGTTG	16	47	+++	-	-	9
<b>E1541R</b>	AAGGAGGTGATCCANCCRCA	20	57	+++	+	-	10

**Tabla 3. Cebadores específicos para el gen 16s rARN bacteriano (Baker G.C et al, 2003), facilitados por la web del Ribosomal Database Project (<http://www.rdp.cme.msu.edu/>).**

<sup>a</sup> El número de los nombres está relacionado con la posición en el gen 16S en la bacteria *Escherichia coli* y la última letra con la dirección: F (Forward) y R (Reverse).

<sup>b</sup> Número de nucleótidos del cebador (N) y temperatura de anillamiento (Ta).

<sup>c</sup> El número total de correspondencias para Bacterias (B), Archeas (A) y Eucariotas (E) se define de la siguiente forma (., ninguna; +, <25; ++, 25-100; +++, <100).

<sup>d</sup> Referencias: **1.** Reysenbach and Pace, 1995; Martinez-Murcia et al., 1995; Reysenbach et al., 1994. **2.** McInnery et al., 1995; Hansen et al., 1998; Farelly et al., 1995. **3.** Rudi et al., 1997. **4.** Watanabe et al., 2001. **5.** Cologhoun, 1997. **6.** Watanabe et al., 2001. **7.** Watanabe et al., 2001. **8.** Rudi et al., 1997. **9.** Reysenbach and Pace, 1995. **10.** Suzuki and Giovanni, 1996.

Para determinar la estructura de una comunidad microbiana en muestras medioambientales basándonos en la amplificación mediante diferentes cebadores (tabla 3) del gen ribosomal 16S, se pueden usar varios tipos de aproximaciones, aunque en esta Tesis Doctoral nos hemos centrado en dos:

- i) Análisis del gen 16 rARN mediante diferentes técnicas indirectas, sin usar la secuenciación. En este tipo de aproximaciones se incluyen las técnicas, por ejemplo, del “fingerprinting” o de TGGE.
- ii) Aproximación basada en la secuenciación directa del gen 16S y comparación con otros genes presentes en las bases de datos.

En los capítulos siguientes desarrollaremos estos dos tipos de aproximaciones para el estudio de comunidades complejas.

## 6.2 Análisis de la diversidad microbiana mediante técnicas de “fingerprinting”

En el capítulo anterior, ya hemos descrito las características tan peculiares de la secuencia del gen ribosomal 16S para el estudio de la diversidad microbiana.

Además de la secuenciación directa de este gen, existen otra serie de técnicas indirectas que, sin usar la secuenciación, aprovechan las características intrínsecas de la secuencia 16S para comparar las especies microbianas presentes en distintas muestras complejas.

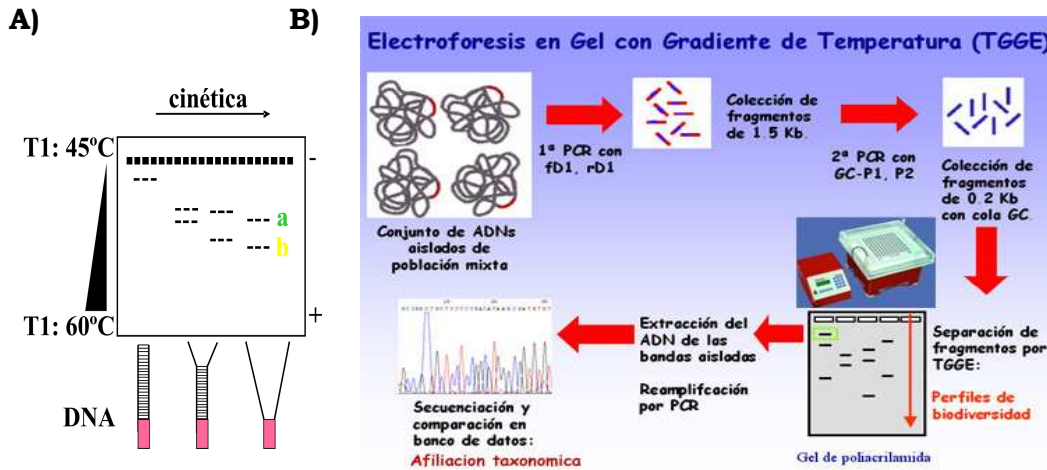
Una de las primeras aproximaciones basadas en fingerprinting que se adoptaron rápidamente en ecología microbiana fue la Electroforesis en gel con gradiente desnaturizante (DGGE; Muyzer et al, 1993). Esta técnica consiste en la separación de productos de PCR en base a su diferente movilidad electroforética en un gel con un gradiente de algún agente desnaturizante.

Una variante de este tipo de técnica es la Electroforesis en Gel con Gradiente de Temperatura (TGGE), que incorpora un gradiente de temperatura en el gel para separar las moléculas de ADN en función de su temperatura de fusión. El fundamento de la técnica se basa en dos hechos:

- La composición de las bases de cada molécula de ADN determina la temperatura a la cual éste se desnaturaliza.
- La migración electroforética de las moléculas bicatenarias de ADN es menor cuando se encuentran parcialmente desnaturalizadas.

Cuando la electroforesis se realiza bajo un gradiente lineal de temperaturas, las moléculas con distinta secuencia se desnaturalizan a distintas temperaturas, deteniéndose en su migración en distintas posiciones del gel. La separación completa de las hebras de ADN de una molécula se evita con una secuencia GC en el extremo que se denomina “grapa” (figura 4).

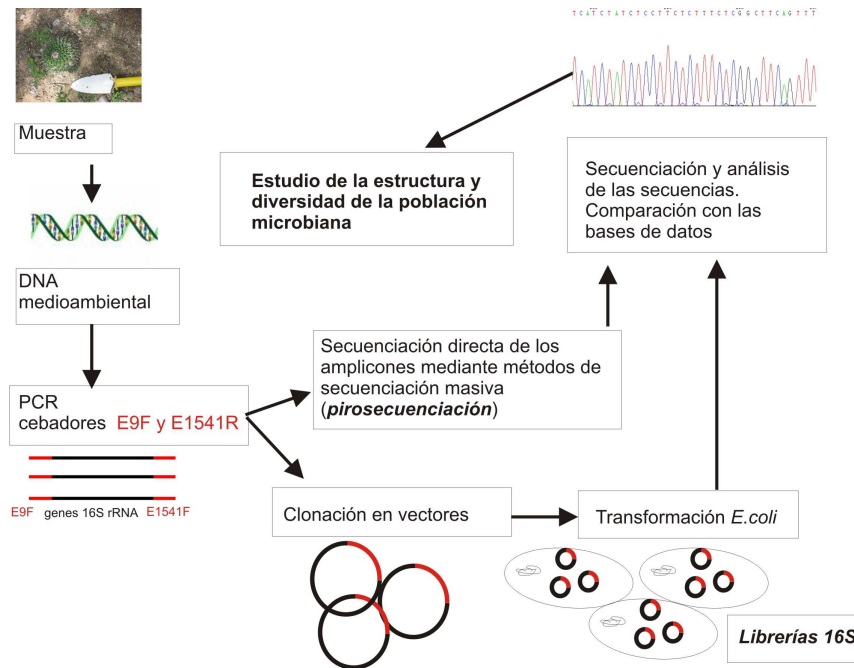
Para una comunidad microbiana compleja con diferentes microorganismos, el resultado serán varias bandas en el gel. Cada banda corresponde con un gen 16S diferente y, teóricamente, con una población diferente de la comunidad microbiana original. La intensidad relativa de cada banda es una medida cuantitativa de la abundancia de esa población. Además, las bandas pueden cortarse y secuenciarse para obtener más información sobre las relaciones filogenéticas de esa población.



**Figura 4.** Fundamento de TGGE. A) Se establece un gradiente de temperatura en el gel que permite que dos amplicones de dos bacterias diferentes (a y b) tengan una movilidad electroforética distinta. B) Generalmente se utilizan dos reacciones de PCR. En la segunda, se asocia al producto amplificado a una grapa (35 GC clamp) que permite que las hebras de ADN desnaturalizado no lleguen a separarse. La extracción y reamplificación de las bandas específicas permite la adscripción y abundancia de un determinado grupo taxonómico (OTUs) en un determinado ecosistema.

**6.3. Análisis de la diversidad microbiana mediante secuenciación del gen ribosomal 16s**

El protocolo que tradicionalmente se ha usado para el análisis de la diversidad microbiana presente en una muestra medioambiental mediante la secuencia del gen ribosomal 16S se esquematiza en la figura 5.



**Figura 5.** Análisis de la estructura y diversidad de la población microbiana de un determinado ecosistema mediante la construcción de librerías 16S o mediante la secuenciación directa (capítulo 6.3.1).

De esta forma, el primer paso es elegir los cebadores adecuados en función de la población que se quiera estudiar en el ecosistema (tabla 3); el segundo paso supone realizar con estos cebadores reacciones de PCR sobre las muestras, en las condiciones óptimas para que la enzima taq-polimerasa cometa los menos errores posibles; posteriormente se clonan estos productos de PCR que contendrán fragmentos del gen 16S en los vectores adecuados; se transforman células de *Escherichia coli* (*E.coli*) para tener un número elevado de plásmidos por célula (al conjunto de células de *E.coli* transformadas con los vectores que portan los fragmentos de las secuencias del gen 16S se lo denomina **librería o genoteca 16S**) y finalmente se manda a secuenciar el inserto que portan los vectores. Las secuencias así obtenidas se podrán comparar con otras depositadas en las bases de datos para establecer las relaciones filogenéticas.

Este tipo de técnica exige seleccionar los clones portadores de vector con inserto y mandarlos a secuenciar de manera aislada, usándose tradicionalmente para la secuenciación el método de Sanger (Sanger *et al*, 1977). El problema con esta aproximación radica en el número limitado de secuencias a analizar, entre 50 y 500, que hace que no se pueda abarcar al elevada diversidad que hay en el suelo.

En la actualidad, se están desarrollando otro tipo de aproximaciones para el estudio de la secuencia del gen ribosomal 16S en muestras medioambientales complejas. Todas ellas son métodos de secuenciación masiva centrados en secuenciar directamente los amplicones 16S (figura 5, capítulo 6.3.1) incrementando en 10 veces los datos obtenidos.

### **6.3.1. Métodos de secuenciación masiva. Tecnología 454**

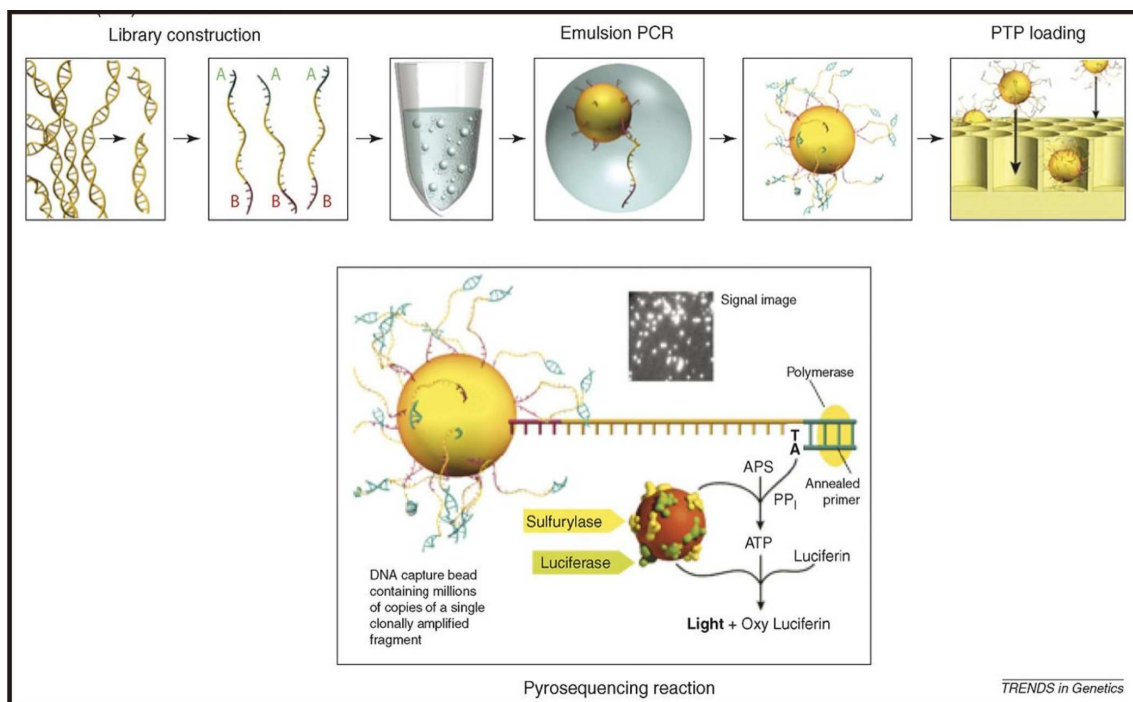
Este tipo de aproximaciones están basadas en la secuenciación masiva de los amplicones de PCR, gracias a la técnica denominada pirosecuenciación o tecnología 454.

La pirosecuenciación se desarrolló originariamente a mediados de los 90 (Ronaghi *et al*, 1996,1998) y se ha ido desarrollando y mejorando hasta su comercialización ([www.pyrosequencing.com](http://www.pyrosequencing.com); <http://www.roche.com/media/>).

Explicaremos este método en mayor detalle, por ser un método realmente novedoso que se ha aplicado a estos estudios.

A diferencia del método de Sanger, la obtención de la secuencia se basa en el proceso de polimerización “*in situ*”. Amplicones con una única molécula de ADN se amplifican en escala nanométrica en una placa de 1.6 millones de pocillos. En cada pocillo se dispone una mezcla de enzimas necesarias para la secuenciación, como la ADN polimerasa, la ATP sulforilasa y la luciferasa. Durante los distintos pasos de la secuenciación, los cuatro nucleótidos se van añadiendo en serie a la placa multipocillos. Si se añade el nucleótido complementario a la secuencia de ADN de un pocillo, la polimerasa extiende la hebra de ADN añadiendo el nucleótido. Al añadir el nucleótido, se libera pirofosfato, que inicia una reacción enzimática que genera una señal de luz que, a su vez, es recogida por una cámara optimizada para tal uso. El pico de la señal de luz es proporcional al número de nucleótidos que se han incorporado. Integrando las señales de luz, se va obteniendo la secuencia.

La eliminación de los vectores de clonación y con esto las limitaciones asociadas a la clonabilidad de determinados fragmentos de ADN son una de las principales ventajas de esta tecnología (Hyman, 1988; Ronaghi *et al*, 1996, 1998; Margulies *et al.*, 2005).



**Figura 6.** Esquema del proceso de pirosecuenciación: La muestra de ADN a pirosecuenciar lleva dos adaptadores, uno para unirse con las bolitas y otro para unirse a la taqpolimerasa. Después de unirse a las bolas, se realiza una PCR en emulsión y se cargan en una placa multipocillos donde tendrá lugar la reacción de pirosecuenciación.

Además, esta técnica permite leer estructuras secundarias y su rendimiento es muy elevado, ya que se puede leer una gran cantidad de secuencias por día (más de 100 Mb por corrida, <http://www.454.com/enabling-technology/index.asp>), y la estimación de esta capacidad de lectura está mejorando (tabla 4). Se infiere así la gran ventaja que resulta este tipo de tecnología para el estudio de ecosistemas complejos de gran diversidad como es el suelo, en el que se necesita un número muy elevado de lecturas para poder tener representada en la mayor proporción posible toda la comunidad microbiana.

Uno de los principales inconvenientes radica en que genera lecturas muy cortas, mucho más pequeñas que la secuenciación tradicional por Sanger. Sin embargo, la longitud de las secuencias generadas ha ido mejorando paulatinamente con el tiempo, desde la plataforma inicial GS20 (aproximadamente 100 bases), la plataforma GSFLX (unas 250 bases) y la última versión denominada “Titanium” ([http://www.roche.com/media/media\\_releases/med\\_dia\\_2008-10-01.htm](http://www.roche.com/media/media_releases/med_dia_2008-10-01.htm)), que genera secuencias de 400 bp y secuencia desde 400 a 600 millones de bp por corrida (tabla 4).

	<b>Rendimiento (millón bp/día)</b>	<b>Rendimiento (millón bp/corrida)</b>	<b>Rendimiento (lecturas/día)</b>	<b>Longitud lecturas (bp)</b>	<i>Coste aproximado 1.000 lecturas</i>
<b>3730xl- Sanger</b>	1-2	0.8	8000	>800	5.000 euros
<b>454 Titanium</b>	1000	500	2.500.000	400	12 euros
<b>SOLiD</b>	>300	3000-10000	>10.000.000	25-35	2 euros
<b>Solexa</b>	>200	1500-3000	>10.000.000	25-35	2 euros

**Tabla 4.** Comparación de los diferentes proyectos de secuenciación (Wen-tso liu y Janet Janson, 2010).

En la actualidad, cada vez hay más trabajos que reflejan el uso de este nuevo tipo de tecnologías para el análisis de la comunidad microbiana en diferentes ecosistemas (Rousk *et al*, 2010; Lauber *et al*, 2010; Schauer *et al*, 2010). Por razones obvias, mejora el rendimiento (se obtienen muchas más



secuencias que por el método tradicional de Sanger) y el tiempo y el trabajo manual se reducen.

Paralelamente al desarrollo de esta tecnología, se han desarrollado diferentes plataformas online (MegaRast, Ribosomal data base project) y distintos programas bioinformáticas (MOTUR) para poder trabajar con el volumen de secuencias (en ocasiones más de un millón de lecturas) que se obtienen de una muestra medioambiental al usar esta tecnología. De este modo, se puede calcular índices de diversidad, comparar con otras secuencias depositadas en las bases de datos, alinear, hacer análisis estadísticos, etc.

### **6.3.2. Bases de datos de genes 16S. Ribosomal Database Project**

Hoy en día existen numerosas bases de datos como el Ribosomal Database Project (RDP; Olsen *et al*, 1991; Cole *et al*, 2009), que pueden consultarse libremente para realizar la comparación estadística de las secuencias obtenidas de un aislamiento contra las que ya están publicadas.

La primera gran colección de secuencias de genes ribosomales 16S fue desarrollada por Carl Woese en 1992, contenía 708 secuencias y sería el núcleo principal de lo que es hoy el RDP que, a fecha de 31 de marzo de 2010, posee 1.396.793 secuencias de 16S rARNs.

Esta base de datos online contiene secuencias del gen 16S de muy diversa índole, incluyendo secuencias de alta calidad y de gran tamaño (286,680; dato de 2008, Wen-tso liu y Janet Janson, 2010), de muestras medioambientales (559.135), de aislados bacterianos cultivables (146.968), y de especies bacterianas tipo (5534). Las especies bacterianas tipo cubren aproximadamente el 60% de los nombres bacterianos adscritos hasta el momento y son muy importantes porque permiten unir la filogenia basada en el gen 16S con la taxonomía bacteriana. El RDP también proporciona una serie de herramientas que permiten al usuario comparar sus propias secuencias en el contexto de las secuencias depositadas en la base de datos. Entre ellas se incluyen herramientas para la clasificación taxonómica (RDP Classifier), para análisis filogenéticos (Tree Builder), comparación de muestras (Library Compare), visualización taxonómica (Taxomatic) y para el análisis de muestras procedentes de métodos de secuenciación masiva (Pyrosequencing Pipeline, capítulo 6.3.1).

Como muestra del potencial de este gen en estudios de ecología microbiana en diferentes ambientes, en la figura siguiente (figura 7), se muestra un árbol filogenético construido a partir de 3502 secuencias de 16S rADN procedentes de distintos ecosistemas (Tamamaes 2010).

El árbol que se muestra en el centro se ha creado tomando una secuencia representativa de cada familia. Las familias, a su vez, se han coloreado según su phyla correspondiente y sólo las familias con 10 observaciones o más han sido consideradas para la construcción del árbol. Las barras en el círculo exterior indican el número de veces que cada familia ha sido observada en una muestra de un ambiente en particular y las barras marcadas con una estrella han sido reducidas de su tamaño original para ajustarlas al esquema (Tamames et al, 2010) . En la figura, llama la atención lo cosmopolitas que son las distintas familias, ya que sólo 3 de las 211 familias son específicas de un ambiente en concreto: dos *Clostrida* (*Lachnospiracea* y *Oscillospiracea*), y la familia de gamma-proteobacteria *Succinivibrionacea*, todas ellas específicas del sistema gastrointestinal de animales.

Además, también se puede apreciar lo estudiado que se encuentran (en lo que a diversidad microbiana basada en el gen 16S ARNr se refiere) algunos ecosistemas como el suelo, las aguas marinas o el sistema digestivo de los animales.

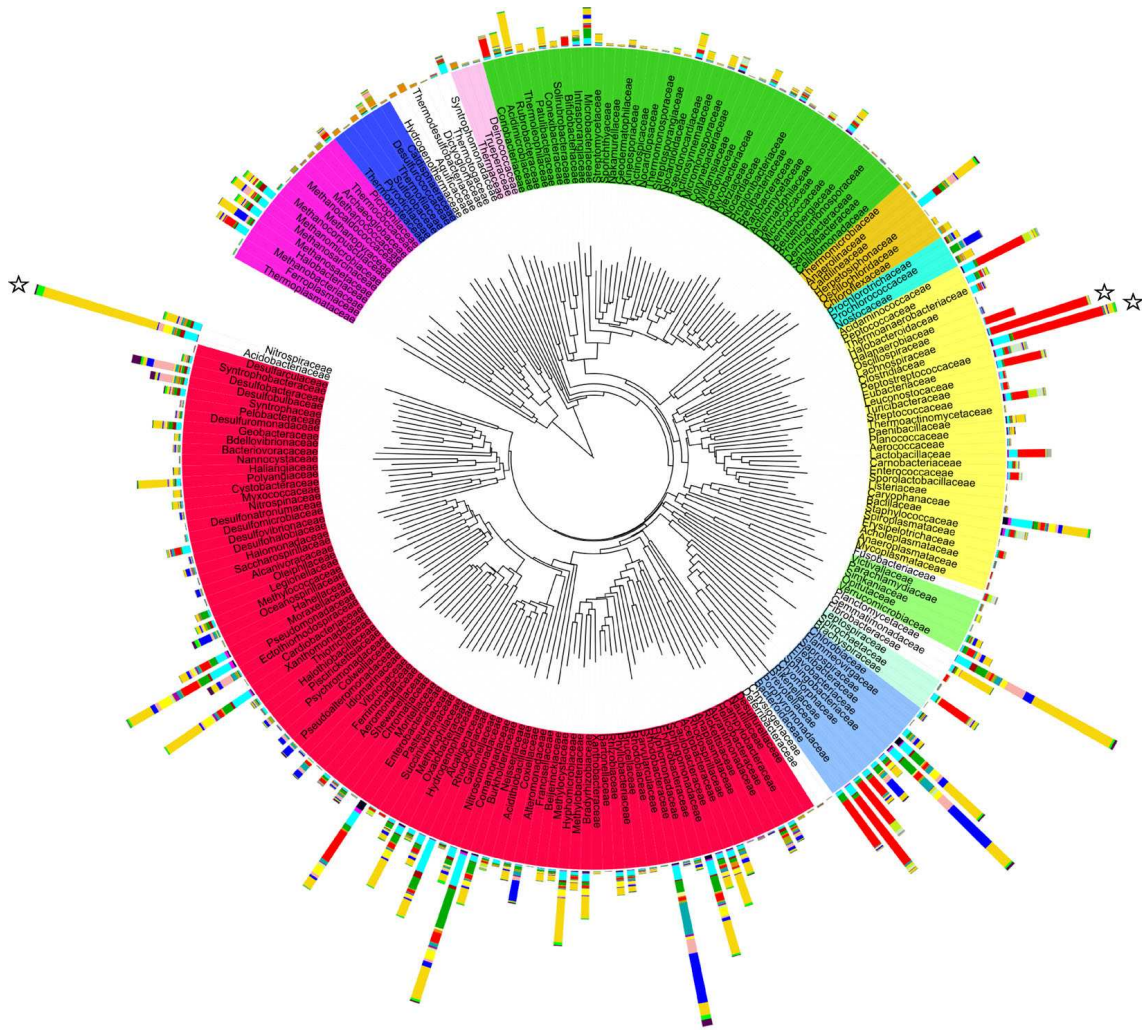
Una de las limitaciones del gen ribosomal 16S, sin embargo, es que no se recomienda su uso para identidades taxonómicas a nivel de especie. En este sentido se recomienda el uso de otros marcadores filogenéticos que pueden dar mayor resolución, como *recA*, *gyrB* y *rpoB*, por ejemplo, y que ya se usan con regularidad en los cultivos puros.

Además, debido a la gran dificultad para definir el propio concepto de especie bacteriana (Rossello-Mora y Amann, 2001; Stackebrandt et al, 2002), en este tipo de aproximaciones se suele hablar de Unidades Taxonómicas Operativas (OTUS) a diferentes niveles de significación.

El uso de aproximaciones moleculares de este tipo centradas en el ADN, obviamente, dependerá del éxito en la extracción de este ácido nucleico del ecosistema de interés. La extracción de ADN del suelo, históricamente, era una tarea ardua y no siempre exitosa por la propia composición del suelo, que portan sustancias que inhiben los procesos posteriores de biología molecular como el clonaje, ligación o amplificación. Sin embargo, en los últimos años,

con el uso de kits comerciales adecuados a este tipo de muestras, eso está cambiando.

Mientras, como ya hemos observado, el gen 16S rARN ha revolucionado la ecología y taxonomía de bacterias y arqueas, el gen 18S y las secuencias intergénicas ITS (“internally transcribed spacer”) están generando una propulsión parecida en el caso de hongos y otros microorganismos eucariotas (Horton and Bruns, 2001). Las bases de datos para estos genes o regiones, sin embargo, todavía no están tan desarrolladas como sus homónimas para el estudio de los genes 16S.



**Environments (Outer bars)**

● Aerial	● Freshwaters-Saline waters	● Oil	● Saline waters
● Animal host	● Geothermal	● Oral	● Soil
● Artificial	● Gut	● Other tissue	● Soil-Freshwaters interfase
● Freshwater sediment	● Hydrothermal	● Plants	● Soil-Saline waters interfase
● Freshwaters	● Marine host	● Saline sediment	● Vagina

**Phyla (Inner circle)**

● Crenarchaeota	● Actinobacteria	● Firmicutes	● Bacteroidetes-Chlorobi
● Euryarchaeota	● Chloroflexi	● Spirochaetes	● Proteobacteria
● Deinococcus-Thermus	● Cyanobacteria	● Chlamydia-Verrucomicrobia	

**Figura 7.** Distribución de las familias taxonómicas en diferentes ambientes. Figura extraída del trabajo de Tamames et al, 2010.

## 7. LA METAGENÓMICA: UNA HERRAMIENTA PARA EL ESTUDIO DE LOS MICROORGANISMOS PRESENTES EN EL SUELO

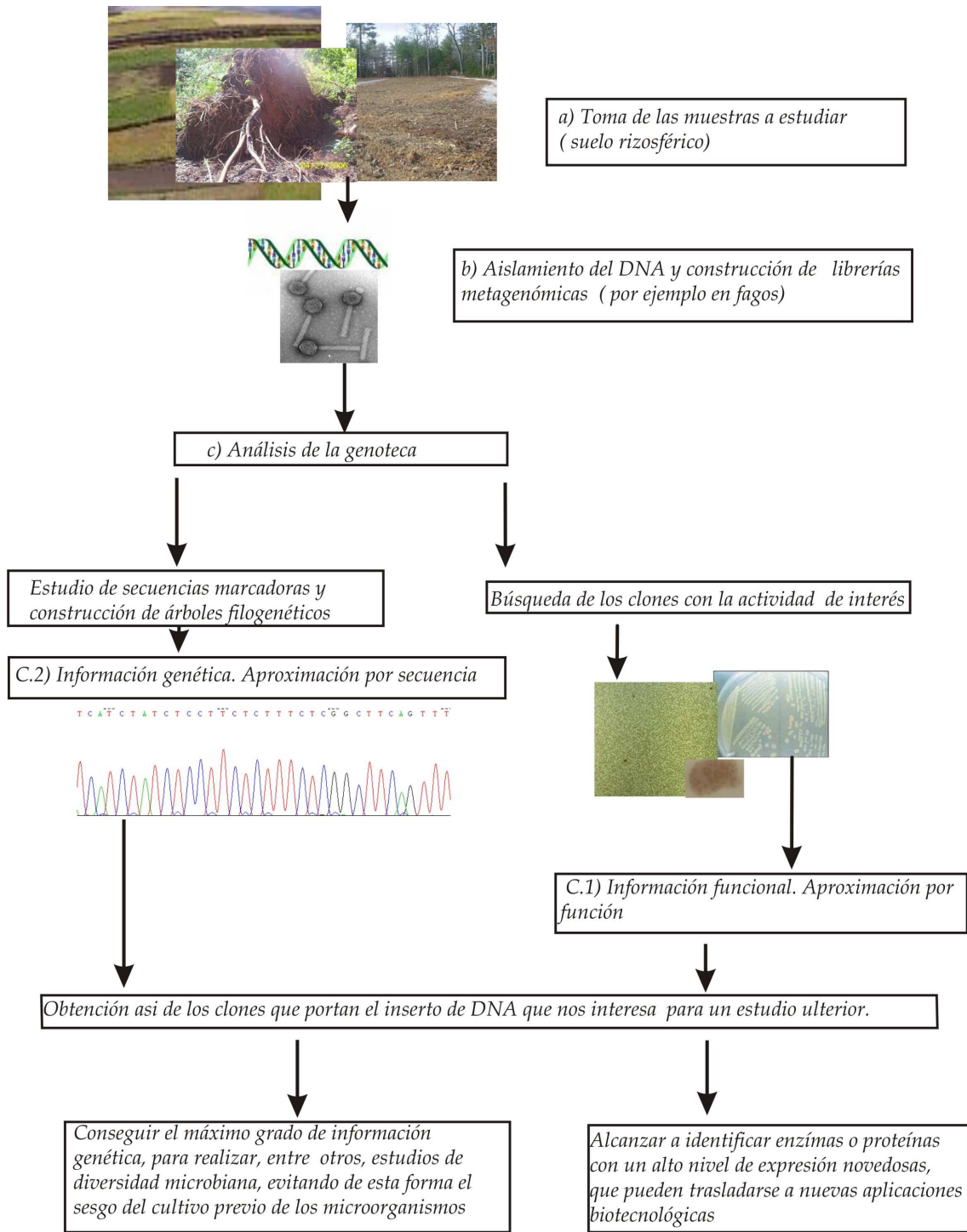
La comparación entre el porcentaje de bacterias cultivables en laboratorio y el recuento de células totales en diferentes hábitats naturales muestra enormes discrepancias (Amann *et al.*, 1995), fenómeno denominado

“*La gran anomalía de conteo en placa*” (Nichols, 2007). Se estima que el 95-99% de la comunidad microbiana presente en el ambiente no es realmente accesible por métodos tradicionales de cultivo.

El conocimiento actual sobre la diversidad microbiana y su papel en la naturaleza es, como vemos, pobre. Esto es debido, principalmente, a que los métodos tradicionales de estudio, como la microscopía, el cultivo en medios específicos o la caracterización bioquímica, sólo tienen un uso limitado en clasificación e identificación. Por otro lado, el hecho de que alrededor del 99% de todas las células bacterianas presentes en el suelo no sean actualmente cultivables en el laboratorio, da lugar a que la clasificación de los microorganismos en función de su fisiología o caracteres bioquímicos sea casi imposible.

Los enfoques de tipo metagenómico permiten así obtener información de dos tipos. Por un lado, se accede al gran reservorio genético de los microorganismos no cultivables, y por otro permite la identificación de genes y enzimas novedosas que pueden trasladarse a nuevas aplicaciones biotecnológicas (figura 8).

Si recordamos las estimaciones de la diversidad presente en el suelo (11.000-21.000 OTUs por gramo, que supera con creces las estimaciones basadas en microorganismos cultivables; Roesch *et al*, 2007), nos podemos hacer una idea de la magnitud de la información nueva que se está generando. En el campo de la microbiología de suelos, los estudios metagenómicos se han centrado en la detección de nuevos biocatalizadores y compuestos bioactivos, como por ejemplo enzimas líticos y nuevos antimicrobianos (Knietsch *et al*, 2009; Sharma *et al*, 2010; Cieśliński H, 2009); en la caracterización de genomas de microorganismos no cultivados como *Acidobacterium* o arqueas del suelo (Jones *et al*, 2009); o en la caracterización de genes de resistencia a metales en las raíces de plantas adaptadas a minas ácidas (Mirete *et al*, 2007). Además, estos estudios nos indican la gran diversidad metabólica existente en el suelo y en la rizosfera (Daniel, 2005).



**Figura 8. Construcción de librerías metagenómicas a partir de muestras de ADN medioambiental y análisis de las mismas.** En la construcción de este tipo de librerías, los pasos a seguir son: **a)** toma de muestras a estudiar, en nuestro caso, suelo rizosférico debido a la gran diversidad microbiológica existente en el mismo, **b)** extracción del ADN medioambiental y clonación en los vectores correspondientes, **c)** la creación de genotecas metagenómicas puede realizarse bien para c.2) obtener el máximo grado de información genética o bien para c.1) obtener enzimas con un alto grado de expresión.

A pesar de los enormes recursos derivados de la diversidad molecular presente en un ADN metagenómico, el número de actividades enzimáticas novedosas es todavía limitado, probablemente debido a problemas tecnológicos y logísticos derivados de la utilización de métodos de búsqueda adecuados y el éxito de la clonación de los fragmentos de ADN. Es por ello que el éxito de estos métodos depende en gran medida de la capacidad de clonar fragmentos de ADN suficientemente grandes en los vectores apropiados, de su capacidad para expresar la información genética en hospedadores heterólogos de forma eficiente y del desarrollo de protocolos adecuados de búsqueda (Kowalchuk *et al*, 2007).

Se espera que, en un futuro próximo, el número de genes nuevos identificados mediante este tipo de técnicas supere al número de genes que se han identificado mediante la secuenciación de microorganismos individuales (figura 4). Sin embargo, las investigaciones van encaminadas no sólo a la identificación de nuevos genes, sino también al entendimiento de la función de estos genes y proteínas dentro de los nichos microbiológicos correspondientes y su papel en los ciclos biogeoquímicos y en el reciclaje de materiales. Las posibilidades que genera la metagenómica y las demás técnicas moleculares mencionadas anteriormente son múltiples, por lo que podemos esperar importantes avances en el conocimiento del funcionamiento de las comunidades microbianas de suelo.

## **8. ACCESO A LA “DIVERSIDAD ESCONDIDA” EN LA RIZOSFERA. CARACTERÍSTICAS DEL PHYLUM ACIDOBACTERIA**

Como ya hemos comentado en el capítulo anterior, las comunidades microbianas rizosféricas difieren tanto cualitativa como cuantitativamente en su composición. La proporción relativa de cada uno de los miembros de la comunidad está sujeta a variaciones debidas a causas externas (factores ambientales; figura 1) o internas (competición interespecífica, etc.), por lo que organismos que son abundantes y cultivables bajo ciertas condiciones ambientales pueden convertirse en formas aletargadas y posiblemente no cultivables bajo otras condiciones ambientales.

Sin embargo, se ha observado una serie de phyla (da Rocha *et al*, 2009) dominantes en todas las comunidades rizosféricas. Estos phyla son los

siguientes: *alfa*-, *beta*- y *gamma*-*proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Verrumicrobia*, *Planctomycetes*, *Bacteroidetes* y *Firmicutes* (figura 9).

Los estudios de estos phyla basados en el gen 16S rARN, se desarrollan bien a partir del análisis de esta secuencia en aislados cultivables o bien a partir del estudio de la secuencia en microorganismos no cultivables.

Sin embargo, la contribución de la información procedente de los aislados cultivables es muy variable. En *Proteobacteria*, por ejemplo, la cantidad de información derivada de microorganismos cultivables es la misma que la derivada de microorganismos no cultivables. Sin embargo, en el phylum *Actinobacteria*, la información derivada de los aislados excede a la información proporcionada por los genes 16S rARN de microorganismos no cultivables. Para los otros cinco grupos bacterianos, la mayor parte de información filogenética tiene su origen en estudios basados en organismos no cultivables. La mayoría de los individuos presentes en estos grupos son por tanto, todavía no cultivables.

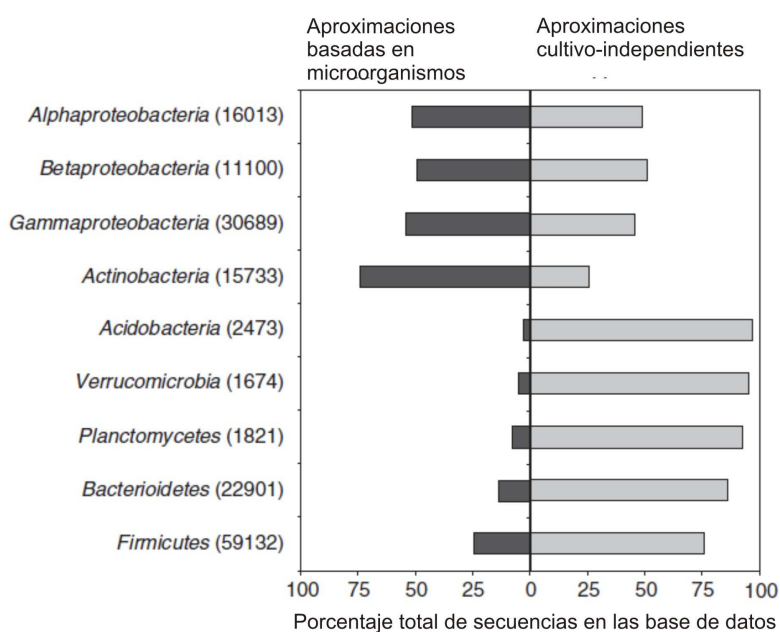
En el caso de los phyla *Acidobacteria*, *Verrumicrobia* y *Planctomycetes*, prácticamente toda la información que se conoce proviene de estudios básicos de microorganismos no cultivables. De este modo, nuestro conocimiento de estos taxa en la rizosfera es todavía muy limitado y está centrado principalmente en las fluctuaciones del gen 16S rARN en la rizosfera (Sanguin, *et al*, 2006; Zul *et al*, 2007).

De estos grupos, *Acidobacteria* y *Verrumicrobia* poseen un número muy bajo de especies recogidas en la literatura. El grupo *Verrumicrobia* tiene 10 especies definidas (Ward-Rainey *et al*, 1995; Garrity *et al*, 2005) y el phylum *Acidobacteria* sólo tres (Garrity *et al*, 2005, Bryant *et al*, 2007; Eichorst *et al*, 2007; Koch *et al*, 2008).

Aunque hay representantes de estos phyla que han sido cultivados a partir de muestras de suelo (Janssen *et al*, 2002; Sait *et al*, 2002; Stevenson *et al*, 2004; Davis *et al*, 2005), ninguno se ha cultivado a partir de muestras rizosféricas. Es necesario aislar directamente este tipo de microorganismos en la rizosfera para apoyar la idea de que son organismos relevantes en este medio, idea que subyace de los estudios basados en microorganismos no cultivables. Sin embargo, los miembros de *Acidobacteria* y *Verrumicrobia* que se han aislado del suelo son aeróbicos, heterótrofos y tienen un crecimiento muy lento (Janssen *et al*, 2002; Sait *et al*, 2002; Stevenson *et al*, 2004; Davis *et al*, 2005), características ampliamente distribuidas en otros aislados



rizosféricos, lo que podría indicar que podrían crecer bien en el ambiente de la rizosfera.



**Figura 9.** Porcentaje de secuencias del gen 16S rARN de los grupos mayoritarios en la rizosfera, obtenidas por vía dependiente de cultivo e independiente. Todas las secuencias (1200 pb) se obtuvieron del RDP, en paréntesis se indica el número de secuencias analizadas (figura tomada del trabajo de da Rocha *et al*, 2009).

De este modo, observamos que se ha descrito una elevada cantidad de bacterias no cultivables presentes en la rizosfera (figura 9; Kaiser *et al*, 2001; Kuske *et al*, 2002; Gremion *et al*, 2003; Schmalenberger y Tebbe, 2003; Graff y Conrad, 2005; Sharma *et al*, 2005; Stafford *et al*, 2005; Sanguin *et al*, 2006; Wang *et al*, 2007; Zhang *et al*, 2007; Van Elsas *et al*, 2008) basándose solo en información filogenética, pero nada se conoce acerca de su papel funcional. Se debe seguir enfatizando en las estrategias basadas en nuevos métodos de cultivo y en las técnicas moleculares para poder entender los posibles roles y funciones de toda esta microbiota no cultivable en la rizosfera.

En el siguiente capítulo nos vamos a centrar en el phylum *Acidobacteria*, del que ya hemos visto que los métodos moleculares indican que es muy abundante en la rizosfera por ser un phyla muy importante en los suelos objeto de análisis de esta Tesis Doctoral (capítulo I de *Resultados*).

### 8.1. Ecología del phylum *Acidobacteria* en el suelo

La primera especie identificada se aisló de un ambiente mineral ácido: *Acidobacterium capsulatum* en 1991 (Kishimoto *et al*, 1991), y fue el único representante del phylum hasta el año 1995 (Hiraishi *et al*, 1995).

Análisis moleculares de la microbiota del suelo posteriores confirmaron la existencia de acidobacterias en el suelo y observaron la increíble diversidad filogenética existente en este phylum (Kuske *et al*, 1997).

Desde entonces, numerosos estudios han revelado que este grupo bacteriano es uno de los phyla más abundantes ciertos suelos (Janssen, 2006). Algunas bacterias cultivables han indicado algún papel ecológico de este phylum. Sin embargo, dada la incapacidad de crecerlas en cultivo, se conoce muy poco de las características ecológicas de las acidobacterias.

Algunas técnicas de cultivo en laboratorio han conseguido cultivarlas con éxito, como por ejemplo: usando medios con un pH bajo (Sait *et al*, 2006), incrementando la concentración de CO<sub>2</sub> (Stevenson *et al*, 2004), seleccionando al azar en una amplia variedad de sustratos (Joseph *et al*, 2003), o extendiendo el tiempo de cultivo (Davis *et al*, 2005). A pesar de todos estos esfuerzos, los aislados cultivables no reflejan la diversidad real de las acidobacterias. De los 74 aislados, cerca del 75% (55/74) pertenece al subgrupo 1 (RDP v.10.4). Estos aislados cultivables nos indican que algunas acidobacterias tienen una tasa de crecimiento lenta y que son capaces de tolerar determinadas condiciones ambientales.

Alternativamente al uso de este tipo de técnicas, se han usado diferentes técnicas moleculares para el estudio de los patrones biogeográficos de las acidobacterias, que han revelado importantes datos acerca de la ecología, fisiología y modo de vida de este grupo bacteriano. En 1999, Barns y colaboradores diseñaron cebadores específicos para las acidobacterias, gracias a los cuales se ha podido cuantificar la abundancia relativa de las mismas en diferentes muestras (Felske *et al*, 2000; Fierer *et al*, 2005, 2007), se ha podido marcar el ADN genómico de las acidobacterias (Quaiser *et al*, 2003), se han desarrollado técnicas de DGGE específicas de este grupo (Born *et al*, 2002) y se han creado librerías 16S de acidobacterias (Barns *et al*, 2007; Bryant *et al*, 2008).

Un estudio reciente usando PCR cuantitativa ha analizado la abundancia relativa de seis grupos bacterianos en 71 muestras de suelo, y ha

determinado que la abundancia de acidobacterias está relacionada negativamente con la disponibilidad de carbono (Fierer *et al*, 2007). Además, otros estudios en suelos han revelado que las acidobacterias se encuentran en menor número en suelos jóvenes que en suelos maduros (Nemergut *et al*, 2007; Tarlera *et al*, 2008). También se han relacionado con el pH y se han encontrado en mayor abundancia en suelos con pH bajo (Fierer *et al*, 2007; Mannisto *et al*, 2007, Lauber *et al*, 2008). Si se tienen en cuenta todos estos datos, podemos llegar a la conclusión de que las acidobacterias normalmente son bacterias oligotróficas y su abundancia en la comunidad microbiana puede estar regulada por el pH.

Otros estudios también apuntan al papel regulador del pH en las comunidades de acidobacterias. Un análisis realizado con 87 suelos estudiados mediante pirosecuenciación y generación de librerías 16 S (Jones *et al*, 2009) mostró que efectivamente, en general, las acidobacterias representan una gran proporción de la comunidad microbiana en suelos con bajo pH. Dentro de los diferentes subgrupos de acidobacterias, los 1,2,3,4 y 6 son los más abundantes en suelos. Además, la abundancia relativa de estos subgrupos también está relacionada con el pH; los subgrupos 1,2,3,12 y 13 tienen una correlación negativa; mientras que la abundancia de los subgrupos 4,6,7,10,11,16,17,18,22 tienen una correlación positiva y abundan en suelos con pH neutro o ligeramente básico (entre 6-8,5).

También se ha observado la predilección de este grupo bacteriano por el suelo suelto en comparación con la rizosfera (Kielak, 2009, Marrilley y Aragno, 1999; Sanguin *et al*, 2006; Fierere *et al*, 2007). Su preferencia por el suelo no asociado a raíces parece que está relacionada con su lenta tasa de crecimiento y la disponibilidad de carbono en el medio. Como ya hemos comentado, suelen ser bacterias oligotróficas, por lo que la elevada cantidad de carbono en la rizosfera podría propiciar el rápido crecimiento de otros microorganismos que desplazarían por competencia a este phylum. Aunque también existen otros estudios que las asocian a suelo rizosférico pero no lo estudian comparativamente con suelo suelto (Chow *et al*, 2002; Stafford *et al*, 2005; Singh *et al*, 2007). También se ha corroborado que se encuentran metabólicamente activas, como en la rizosfera del castaño (Lee *et al*, 2008), donde también son numéricamente dominantes, o en la rizosfera de *Erica andavalensis*, en un ambiente ácido, donde la subdivisión 1 es metabólicamente activa (Mirete *et al*, 2007).

De este modo, observamos como este grupo microbiano es un claro ejemplo de phylum del que se conoce muy poco por métodos tradicionales de cultivo y cuya ecología, estilo de vida y fisiología se está intentando dilucidar mediante el uso de técnicas moleculares. Es necesario seguir haciendo hincapié en el desarrollo de métodos de cultivo adecuados para conocer más acerca de este tipo de microorganismos que, como se ha demostrado, es uno de los más importantes, en términos numéricos, en el suelo y en la rizosfera en particular.

## **9. METAGENÓMICA COMO HERRAMIENTA PARA LA BÚSQUEDA DE NUEVAS ENZIMAS DE INTERÉS INDUSTRIAL**

Como ya hemos comentado, los enfoques de tipo metagenómico permiten obtener información de dos tipos (figura 8); se puede analizar la estructura y diversidad de la comunidad microbiana mediante el estudio del gen 16S o de otros genes (capítulos 6,7 y 8), y además se puede usar para la búsqueda de nuevas enzimas con aplicaciones biotecnológicas como celulasas (Jian C, 2009; Berlemont R, 2009; Pang H, 2009; Kin SJ, 2008), amilasas (Sharma S, 2010; Yun J, 2004), lipasas (Ciésłinski, 2009; Yuhong X, 2009; Wei P, 2009; Hong ks, 2007; Kin YJ, 2006; Lee sw, 2004), etc.

Con este tipo de técnica se ha logrado encontrar enzimas con homologías muy bajas a las ya existentes, fundamentalmente porque no se necesita tener información a priori de secuencia (Danil R, 2004, Warnecke, F, 2007). Sin embargo, la principal limitación de este tipo de técnica es que el sistema de detección enzimática sea el adecuado para que la búsqueda culmine con éxito.

En los siguientes capítulos nos centraremos en la búsqueda de nuevas enzimas en librerías metagenómicas.

### **9.1. Búsqueda de nuevas enzimas con actividad lipasa en genotecas de suelos**

#### **9.1.1. Papel de las lipasas/esterasas en el suelo**

En esta Tesis Doctoral, se profundiza en la detección y aislamiento de nuevas enzimas estererasas de muestras de suelo. ¿Pero cuál es el papel de estas enzimas en la rizosfera?

Las plantas utilizan innumerables estrategias de defensa frente a los microorganismos patógenos. Una de las primeras estrategias y probablemente una de las más importantes es la barrera física que supone la presencia de pared celular. El tamaño de los poros de la pared celular es demasiado pequeño como para dejar pasas a virus y bacterias, por lo que los microorganismos penetran aprovechando pequeños huecos en la pared o bien utilizando diversas maquinarias enzimáticas para disolver la pared, que es lo que se conoce como “disolución enzimática” (Williamson G et al, 1998).

De los componentes de la pared celular, la lignina es el más difícil de degradar y se encuentra especialmente en tejidos vegetales maduros. Después de la lignina, en orden decreciente de resistencia, encontramos la celulosa, la hemicelulosa y las pectinas. Estas últimas son abundantes en frutos y por eso suelen ser más susceptibles al ataque microbiano. La lignina y los polisacáridos de la pared celular se asocian formando una estructura compleja y organizada que resiste el ataque enzimático gracias a su inaccesibilidad física y su estructura química. La estabilidad química aumenta gracias al establecimiento de cadenas laterales que, en muchos de los polisacáridos existentes, se establecen en todas sus unidades monoméricas. Hay dos tipos principales de cadenas laterales: las sustituciones para limitar el acceso de las enzimas al núcleo del polisacárido, o los “cross-linkages” entre los polisacáridos que aumentan la integridad de la pared celular. La hemicelulosa y las pectinas poseen una gran variedad de cadenas laterales adheridas al núcleo, como por ejemplo grupos acetil, metil y feruoil, cada uno de los cuales modifican las propiedades físicas del polisacárido (Williamson G et al, 1998).

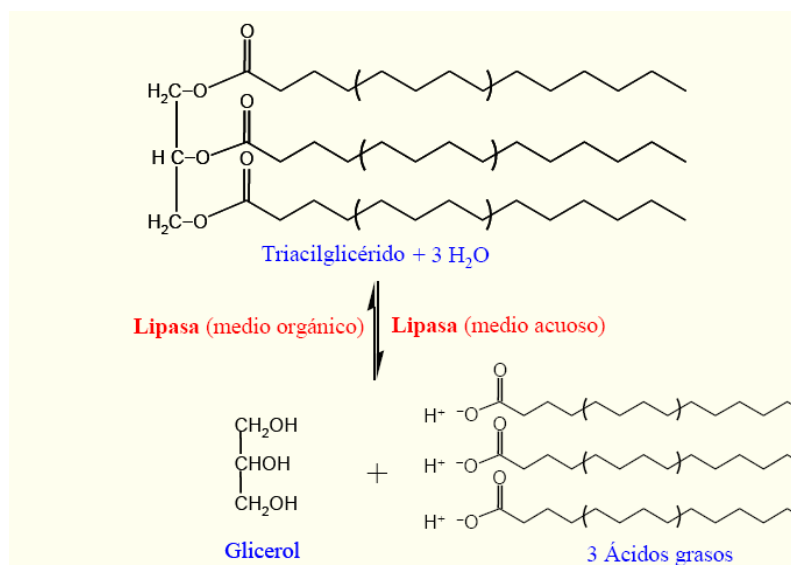
Los microorganismos utilizan enzimas, en particular **esterasas y lipasas** (tabla 5), para eliminar estas cadenas laterales y hacer más accesible el núcleo de la pared celular a las enzimas que degradan las cadenas principales.

Enzima	Subclase
<b>Carboxilesterasas (EC 3.1.1.1)</b>	Cinnamoil esterazas: FAEA/FAE-III (feruloil esteraza) Hidrolasas cinnamoil ester CinnAE (esteraza p-cumaroil/cinnamoil)
<b>Acetilesterasas (EC 3.1.1.6)</b>	Acetil-xilan esterazas Rhamnogalacturon-acetilesterasas Pectin acetilesterasas
<b>Pectin metilesterasas (EC 3.1.1.11)</b>	Pectin-metilesterasas

**Tabla 5.** Clasificación de las esterazas microbianas que participan en la degradación de la pared celular basada en la nomenclatura de la comisión internacional de enzimas (Jaeger KE *et al*, 2002).

### 9.1.2. Propiedades generales de las lipasas

Las lipasas son ésteres hidrolasas carboxílicas que rompen los enlaces éster de los acilglicéridos mediante la adición de una molécula de agua, dando lugar a ácidos grasos libres y glicerol (Figura 5).



**Figura 10.** Hidrólisis y síntesis de ácidos grasos mediante lipasas.

“Lipasa” o “enzima lipolítico” son términos genéricos que incluyen dos grupos principales de enzimas: (1) carboxilesterasas (EC 3.1.1.1) y (2) lipasas “auténticas” (EC 3.1.1.3), aunque a veces también se incluyen otras enzimas que actúan sobre acilglicéridos y otros lípidos similares como las cutinasas, las fosfolipasas y las lipoproteínas lipasas.

El principal criterio para diferenciar a las carboxilesterasas de las lipasas “auténticas” corresponde a la especificidad de sustrato. Las lipasas auténticas son específicas para acilglicérols y otros ésteres con ácidos grasos de cadena larga (que tengan al menos 10 átomos de carbono), siendo la trioleína su sustrato de referencia; mientras que las carboxilesterasas actúan específicamente sobre acilglicérols de cadena corta (menos de 10 átomos de carbono) y otros ésteres simples, siendo la tributirina su sustrato estándar (Arpigny y Jaeger, 1999). Las carboxilesterasas siguen una reacción enzimática que está regida por la ecuación de Michaelis-Menten (Brockhoff y Jensen, 1974; Sarda y Desnuelle, 1958), mientras que las reacciones cinéticas de las lipasas suelen estar representadas por curvas sigmoidales (Brockhoff y Jensen, 1974), porque muestran un efecto denominado “activación interfacial”. A bajas concentraciones de sustrato, la actividad enzimática es también baja, pero si la concentración de sustrato se incrementa hasta un punto en el que se forman micelas, la actividad enzimática aumenta rápidamente. La explicación a este fenómeno se dedujo de la estructura cristalina de la lipasa pancreática y de la lipasa del hongo *Rhizomucor miehei*. Esta estructura terciaria muestra una  $\alpha$ -hélice que cubre el sitio activo, en contacto con una interfase o un medio orgánico. Esta  $\alpha$ -hélice denominada tapadera, cambia su estructura y el sitio activo se vuelve ahora accesible para el sustrato (Brzozowski *et al*, 1991; van Tilbergh *et al*, 1993). Sin embargo, se han identificado numerosas lipasas que no muestran activación interfacial, como por ejemplo la lipasa LipA de *B. subtilis*, que no posee la tapadera, y la lipasa de *P. aeruginosa* LipA que no muestra activación interfacial aunque sí posee la estructura en  $\alpha$ -hélice denominada tapadera (Jaeger *et al*, 1993; Lesuisse *et al*, 1993; van Pouderooyen *et al*, 2001). De este modo, la presencia de esta tapadera no es algo distintivo de las lipasas y las carboxilesterasas.

Las lipasas “auténticas” muestran un mayor número de aminoácidos no polares pequeños localizados en las zonas expuestas al solvente y en la región del centro activo que las carboxilesterasas, así como un rango de sustratos más amplio, una mayor regio y estereoespecificidad, una mayor estabilidad en solventes orgánicos y un potencial electrostático en el centro activo más básico que las carboxilesterasas (Fojan P *et al*, 2000).

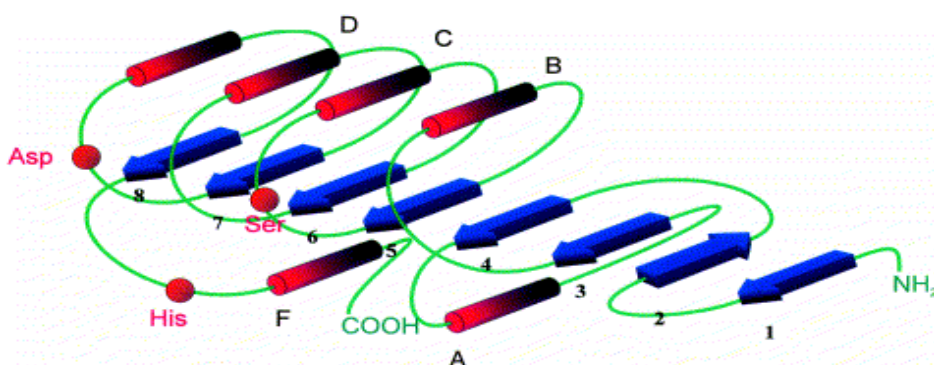
Una de las principales características de las lipasas consiste en que son enzimas muy versátiles: pueden catalizar reacciones de hidrólisis sobre un amplio rango de lípidos (en presencia de solventes orgánicos). Son también

capaces de catalizar reacciones de síntesis o de intercambio de grupos, transesterificación e interesterificación, entre diferentes moléculas, incluyendo lípidos, glúcidos y aminoácidos. Además, estas enzimas pueden realizar todas estas reacciones con una elevada especificidad de sustrato, así como de forma regio o enantioespecífica, lo que hace de ellas unas de las enzimas más interesantes en el campo de la biotecnología (Bornscheuer UT et al, 2002).

### 9.1.3. Estructura y mecanismo de acción de las enzimas con carácter lipolítico

Las lipasas se clasifican en 8 familias basadas en su secuencia de aminoácidos y en su función fisiológica (Aprigny y Jaeger, 1999; Jaeger y Eggert, 2002).

Con respecto a su estructura, todas las lipasas (al igual que otras enzimas como proteasas o amidasas, Nardini and Dijkstra, 1999) presentan el plegamiento típico de las  $\alpha/\beta$  hidrolasas (“ $\alpha/\beta$  hydrolase fold”), que consiste en una estructura central formada por 8 láminas  $\beta$  interconectadas por hélices  $\alpha$  (Nardini and Dijkstra, 1999; Ollis et al, 1992; figura 11).



**Figura 11.** Representación esquemática del plegamiento tipo  $\alpha/\beta$  hidrolasas.

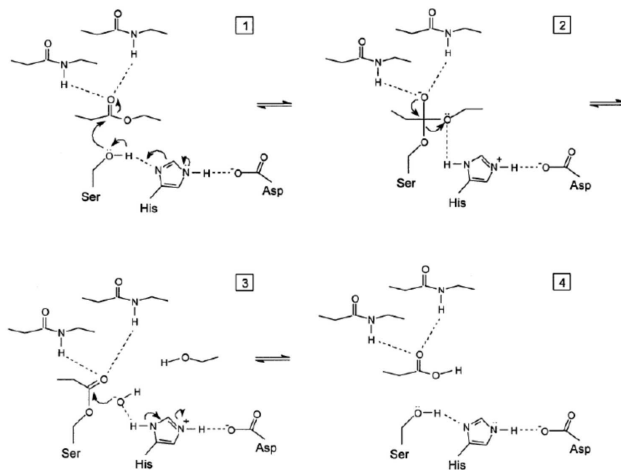
El centro activo de estas enzimas, tiene tres aminoácidos catalíticos, cuya posición dentro del plegamiento suele estar conservada: una **serina** nucleofílica, un grupo ácido (**aspártico o glutámico**) y una **histidina**. La serina catalítica está generalmente incluida en el pentapéptido conservado Gly(Ala)-Xaa-Ser-Xaa-Gly, el cual forma un giro entre la lámina  $\beta$ 5 y la siguiente hélice  $\alpha$ , su función es estabilizar y orientar la serina nucleofílica (Ollis *et al*, 1992) y es conocido como ‘codo nucleofílico’ (“nucleophilic elbow”). Además, el centro activo contiene otras estructuras que intervienen en la estabilización de los intermediarios producidos durante la catálisis (“oxyanion



hole”), o que se encargan de acomodar el ácido graso a escindir (“scissile fatty acid binding pocket”) u otras partes del sustrato, y cuyas características condicionan la especificidad de sustrato y la selectividad de estas enzimas.

La hidrólisis del enlace éster del sustrato por lipasas o carboxylesterasas (figura 12) sucede mediante un mecanismo similar al de las serín-proteasas con los siguientes pasos:

1. El sustrato se sitúa en el centro activo de la lipasa de tal manera que el carbono carbonílico toma contacto con el grupo -OH de la serina. Al final de este paso, se produce un intermediario tetraédrico ( $T_{d1}$ ) (Semeriva *et al*, 1974).
2. La carga negativa, en un principio situada en el  $O^-$  de la serina, sufre una translocación hacia el oxígeno de grupo carbonilo, originando un oxianión. Este oxianión encaja en un hueco, denominado “oxyanion hole”.
3. El intermediario tetraédrico se rompe, el alcohol se libera y seguidamente se forma el complejo acil-enzima.
4. Se produce un segundo ataque nucleofílico, esta vez de una molécula de agua sobre el carbono carbonílico del complejo acil-enzima. Se forma un nuevo intermedio tetraédrico ( $T_{d2}$ ) que, a continuación, se rompe, liberando un ácido graso y regenerando el -OH de la serina. (Jaeger y Reetz, 1998; Jaeger *et al*, 1999).



**Figura 12.** Mecanismo de acción de las lipasas. 1. Unión al centro activo y ataque nucleofílico de la Ser al grupo carbonílico. 2. El intermediario tetraédrico es estabilizado por dos residuos del “oxyanion hole” y los residuos de Asp e His del centro activo. 3. Se libera el alcohol. El complejo enzima-sustrato es atacado por una molécula de agua. 4. Formación del grupo carboxil y regeneración de la Serina del sitio activo.

#### **9.1.4. Distribución y clasificación de las carboxilesterasas y lipasas bacterianas**

Inicialmente, se clasificaron 53 lipasas y carboxilesterasas en 8 familias, basándose en homologías de secuencia y diferentes propiedades (Arpigny y Jaeger, 1999). Desde entonces, numerosas enzimas lipolíticas han sido identificadas y caracterizadas. Una excelente colección de lipasas se puede encontrar en la base de datos denominada BRENDA (Pharkya *et al*, 2003; Schomburg *et al*, 2004).

En la actualidad (abril, 2010), esta base de datos contiene 2813 secuencias de enzimas carboxilesterasas (de las cuales 625 pertenecen a bacterias) y 2509 secuencias de lipasas verdaderas (773 de bacterias).

Muchas de las nuevas enzimas descritas no se pueden clasificar en los grupos de familias existentes (de Pacale *et al*, 2008, Kakugawa *et al*, 2007). De este modo, el estudio de las enzimas lipolíticas representa un excitante campo de investigación donde se pueden encontrar muchas enzimas y actividades nuevas con nuevas aplicaciones.

A continuación, describiremos las familias en las que se clasifican este tipo de proteínas:

- **Triacilglicerol lipasas (familia I)**

La familia I contiene todas las lipasas verdaderas y está compuesta por 8 subgrupos.

Las enzimas de la subfamilia I.1 se diferencian de las de la subfamilia I.2 por su secuencia aminoacídica. Las de la subfamilia I.1 presentan una alta similitud con la lipasa de *Pseudomonas aeruginosa*. Las lipasas de la subfamilia I.2 poseen una gran homología con la lipasa de *Burkholderia glumae*. Ambas familias están formadas por lipasas de bacterias gram-negativas. Muchas lipasas interaccionan con chaperonas específicas que se denominan “proteínas Lifs” (Amand *et al*, 1994; Frenken *et al*, 1993; Honson *et al*, 1993; Rosenau *et al*, 2004).

En el caso de la subfamilia I.3, tienen un peso molecular de entre 50 y 65 Kda. Recientemente se ha estudiado la estructura terciaria de la enzima LipA de *Serratia marcescens* (Meier *et al*, 2007), que ha revelado mucho acerca de la configuración espacial de esta enzima. Las proteínas pertenecientes a esta subfamilia están muy relacionadas con las proteínas de procariotas.

La subfamilia I.4 contiene las lipasas más pequeñas identificadas, con una masa molecular de unos 20 KDa (Arpigny y Jaeger, 1999). En este grupo

se encuentran diversas lipasas de géneros gram positivos, como los *Bacillus*. Las lipasas de esta familia muestran un pH óptimo de 10-11 y son independientes de  $\text{Ca}^{+2}$ .

La subfamilia I.5 son lipasas de diferentes bacterias gram positivas como *Bacillus*, *Geobacillus* y *Clostridium*. El pH óptimo va desde 9.5 para la lipasa de *B.stearothermophilus* hasta 7.5 para la enzima de *B.Thermoleovorans* (Kim *et al*, 1998; Lee *et al*, 1999; Rua *et al*, 1997). El peso molecular de las proteínas presentes en esta subfamilia ronda los 46 KDa.

La subfamilia I.6 contiene numerosas lipasas de especies de *Staphylococcus*, que son factores de virulencia.. Poseen una amplia especificidad de sustratos (Rosentein y Goeth, 2000; Simons *et al*, 1999; Tyski *et al*, 1983) y son estables en un pH entre 4 y 9 (Sayari *et al*, 2001).

La subfamilia I.7 presenta gran homología con las enzimas pertenecientes a la subfamilia I.2. Se han identificado dos lipasas en este grupo de enzimas, originarias de *Streptomyces cinnamomeus* (Sommer *et al*, 1997) y *Propionibacterium acnes* (Miskin *et al*, 1997), con un peso molecular entre 30 y 36 Kda.

Finalmente, la subfamilia I.8 representa una nueva subfamilia de lipasas que fue propuesta recientemente, basada en la identificación y caracterización de la lipasa Lip1 de la bacteria psicrófila *Pseudoalteromonas haloplanktis* (de Pascale *et al*, 2008).

- ***Esterasas familia II-GDSL***

Esta familia comprende las enzimas llamadas GDSL/SGNH que pierden el motivo G-X-S-X-G y, en lugar de él, poseen un tetrapeptido conservado localizado en el N-terminal de la proteína G-D-S-(L) (Arpigny y Jaeger, 1999). Esta familia fue identificada en 1995 (Upon y Buckley, 1995) y contiene 5 bloques de aminoácidos conservados con el tetrapéptido G-D-S-(L) en el bloque I (Molgaard, 2002). En cada bloque poseen un aminoácido estrictamente conservado y que es esencial para la catálisis, normalmente Serina, Glicina, Asparagina e Histidina. Es por esto por lo que se suelen llamar hidrolasas SGNH.

- ***Esterasas familia III***

A esta familia pertenecen enzimas extracelulares de *Streptomyces* y *Moraxela* (Cruz *et al*, 1994; Feller *et al*, 1990; Perez *et al*, 1993). Este tipo de enzimas poseen una masa molecular de 32-35 KDa y muestran gran

homología con las isoformas intracelulares y plasmáticas de la acetilhidrolasa del factor de activación plaquetario humano (PAF-AH; Arpigny and Jaeger, 1999).

- ***Esterasas familia IV***

En esta familia encontramos esterases de procariotas muy diversas, incluyendo microorganismos psicrófilos, mesófilos y termófilos. Las enzimas pertenecientes a esta familia poseen una alta similitud con la familia de las lipasas sensibles a hormonas de mamíferos (HSL; Hemilae *et al*, 1994). En este caso, se identifican diferentes bloques con los residuos conservados, los bloques II y III contienen los residuos de la tríada catalítica (Arpigny and Jaeger, 1999). El bloque I contiene una secuencia conservada H-G-G-G que está relacionada con las interacciones en forma de puentes de hidrógeno para estabilizar el “oxyanion hole” y promover la catálisis.

Se han estudiado numerosas estructuras de enzimas de la clase IV (de Simone *et al*, 2000; de Simone *et al*, 2001; Wei *et al*, Zhy *et al*, 2003) y todas ellas muestran una estructura como un “tapón”, totalmente diferente a la estructura en forma de tapadera encontrada en las lipasas verdaderas (Wang *et al*, 2005), que está formada por dos regiones en forma de hélices muy separadas y recubre el sitio activo. Recientemente se ha identificado un inhibidor que se une específicamente a la serina del sitio activo. Esta inhibición es específica de este grupo, por lo que podría servir para discriminar este tipo de enzimas de otras de otra familia (Ben Ali *et al*, 2006).

- ***Esterasas familia V***

Las esterases de esta familia tienen como origen diferentes géneros bacterianos, tanto mesófilos como adaptados al frío o al calor, como por ejemplo *Pseudomonas*, *Haemophilus* y *Moraxella*. Los aminoácidos conservados se localizan en tres bloques, los residuos de la tríada catalítica están localizados: la Serina en el bloque II y el Asp y la His en el bloque III.

- ***Esterasas familia VI***

Las enzimas clasificadas dentro de la familia VI son enzimas pequeñas (23-26 KDa), con la excepción de una esterasa aislada de una librería metagenómica que tiene un peso molecular de 31.6 KDa (Arpigny y Jaeger 1999; Kim *et al*, 2006). Estas enzimas muestran alrededor de un 40% de homología con otras lisofosfolipasas de eucariotas y poseen tres bloques conservados, como ocurre con las proteínas de los grupos IV y V.

- ***Esterasas familia VII***

Este tipo de enzimas tiene un peso molecular aproximado de 55 KDa y muestra una gran similitud con acetilcolinas esterasas y con carboxilesteras del intestino o del hígado (Arpigny and Jaeger, 1999). Encontramos enzimas de grupos como *Bacillus*, *Peaenibacillus* y *Geobacillus*. Tienen cinco bloques con aminoácidos lo suficientemente conservados como para el diseño de cebadores específicos para la búsqueda de esterasas de esta familia en genomas de *Bacillus* o en muestras de suelo o de agua de minas (Nthangeni *et al*, 2005). El rango típico de hidrólisis de sustratos incluye ácidos grasos de cadena corta y triglicéridos, pero también puede hidrolizar diferentes compuestos de interés biotecnológico (Zock *et al*, 1994; Maqboll *et al*, 2006).

- **Esterasas familia VIII**

Al contrario que el resto de esterasas, la estructura de la familia VIII de esterasas difiere significativamente del plegamiento típico de las  $\alpha/\beta$ -hidrolasas y muestra una gran homología con las  $\beta$ -lactamasas. La serina catalítica está localizada al principio de una  $\alpha$ -hélice (Wagner *et al*, 2002).

Suelen poseer una masa molecular de 42 KDa aproximadamente y muchas de ellas son activas en forma multimérica (Elend *et al*, 2006; Nishizawa *et al* 1995; Schuette y Fetzner, 2007), mientras que otras actúan en forma monomérica (Kim *et al*, 1994; McKay *et al*, 1992; Rashmuse *et al*, 2007; Wagner *et al*, 2002). La alta tolerancia a solventes orgánicos de estas esterasas las hace tener un interés añadido en la industria (Elend *et al*, 2006; Orgino *et al*, 2004; Schuette y Fetzner, 2007). Como ya hemos comentado, este tipo de enzimas está estructuralmente relacionado con las  $\beta$ -lactamasas, pero sólo en casos puntuales se han descrito esterasas con actividad  $\beta$ -lactámica (Govardhan y Pratt, 1987; Jones y Page, 1991; Pratt y Govardhan, 1984).

### **9.1.5. Aplicación industrial de las lipasas**

Las lipasas microbianas son motivo de gran interés debido a su potencial biotecnológico. Entre otras aplicaciones, estas enzimas se utilizan en la industria alimentaria (producción de aromas, emulgentes, etc.), en química orgánica (síntesis de antibióticos, pesticidas, producción de compuestos enantiopuros, etc.), en la industria de la limpieza (aditivos en detergentes, producción de surfactantes para jabones y productos de limpieza, etc.), en la industria papelera (eliminación del “pitch”, de tintas, etc; Bajpai *et al*, 1998, Zollner *et al* 1998), así como en el tratamiento de productos residuales o

tóxicos, en biosensores, en la producción de biodiesel, etc. ( Reetz MT *et al*, 2002).

La importancia de estas aplicaciones ha llevado a la optimización de las reacciones catalizadas por las lipasas, a mejoras en la producción y purificación de estas enzimas y a la obtención de lipasas con nuevas propiedades catalíticas o una mayor estabilidad, lo que se ha conseguido mediante la modificación de enzimas ya existentes o mediante el aislamiento de nuevas lipasas (Jaeger KE *et al*, 2002).

Muchas de estas nuevas lipasas se han obtenido de microorganismos del suelo, donde se encuentran bacterias muy activas en el reciclaje de nutrientes como las pertenecientes al género *Bacillus* y otros géneros relacionados. Estos géneros son considerados como una de las mayores fuentes de enzimas con aplicación industrial.

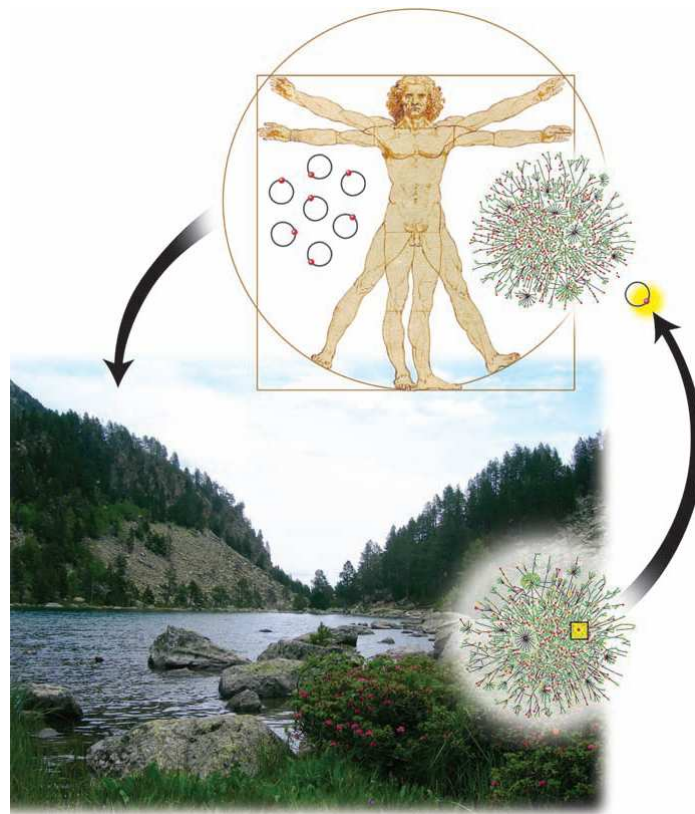
Estas enzimas son altamente estables en un amplio rango de temperaturas, pHs y solventes orgánicos, aunque la mayoría son mesófilas o prefieren pHs neutros o básicos. Son también estables frente a diferentes detergentes, iones y agentes químicos, y, en general, no requieren cofactores. Por estos motivos y debido al conocimiento existente sobre estas enzimas, a su disponibilidad y al hecho de que los procesos en los que intervienen son generalmente menos costosos y menos contaminantes, las lipasas bacterianas son las enzimas más versátiles y ampliamente utilizadas en biotecnología (Gupta R, 2004).

## **10. EL SUELO COMO RESERVORIO DE NUEVOS GENES DE RESISTENCIA A ANTIBIÓTICOS. CONCEPTO DE RESISTOMA**

A pesar de que los mecanismos de resistencia a antibióticos están ampliamente estudiados en clínica (Levy *et al*, 1997), no se conoce apenas nada sobre los reservorios ambientales de este tipo de genes y su contribución a la aparición de multiresistencias en la sanidad. La identificación de este tipo de reservorios y el estudio de su posible migración desde ecosistemas naturales hasta el ambiente humano puede ayudar a combatir la resistencia a antibióticos en patógenos (figura 13).

El éxito en el tratamiento de las infecciones bacterianas en clínica habitualmente se ve afectado por la aparición de cepas multi-resistentes a diferentes tipos de antibióticos (Levy and O'Brien, 2005). Además, el

incremento en la incidencia de enfermedades y en la longitud de vida, está provocando un enorme gasto en sanidad (Holmber *et al*, 1987; Cosgrove, 2006). Para poder hacer frente a este tipo de crisis, se necesita un amplio arsenal de sustancias antimicrobianas (Levy y Marshall, 2004) y profundizar en los mecanismos de resistencia a antibióticos, en su ecología, posible origen y movimiento (Aminov y Mackie, 2007).



**Figura 13.** Movimiento de los genes de resistencia a antibióticos desde ambientes naturales hasta la clínica. La mayoría de los genes de resistencia que se adquieren por transferencia horizontal provienen de microorganismos naturales. En sus hospedadores naturales, estos determinantes forman parte de rutas metabólicas. Después de su diseminación, su integración en nuevas rutas metabólicas puede ser difícil y su único papel puede estar relacionado con la resistencia. Esquema tomado de Martínez JL, 2008.

Dada la complejidad y diversidad del suelo, actualmente se está considerando este ecosistema como un potencial reservorio de nuevos mecanismos de resistencia (D´Costa *et al*, 2006).

### 10.1. Mecanismos de resistencia a antibióticos

Aproximadamente el 80% de los antibióticos usados en clínica provienen de bacterias del suelo (Keiner T, 2000). La clase *Actinomycetos* es la responsable de la síntesis de la mayoría de este tipo de compuestos. En la tabla 6 se muestran diferentes ejemplos de mecanismos de resistencia a antibióticos (asociados la mayoría a *Streptomyces spp.*) e incluye mecanismos de inhibición de la resistencia, bombas de membrana y modificaciones específicas en la propia molécula del antibiótico (De la Cruz F, 2002).

Grupo antimicrobiano	Mecanismo	Proteína natural relacionada	Reservorio natural
<b>Aminoglicósidos</b>	Acetilación/Fosforilación/metilación del gen 16S rARN	Histonas acetilhidrolasas/Proteínas quinasa/ metilasas	<i>Streptomyces</i>
<b>Tetraciclinas</b>	Bombas de extrusión	Superfamilia EF-TU, EF-G	<i>Streptomyces</i> , <i>Micromonospora</i>
<b>Cloroamfenicol</b>	Acetilación/Bombas de extrusión	Acetilasa/ EF-Tu	<i>Streptomyces</i>
<b>Macrólidos</b>	Mutación del antibiótico diana	50S subunidad del ribosoma	<i>Streptomyces</i>
<b>Fluoriquinonas*</b>	Protección de la topoisomerasa	Proteínas relacionadas con Qnr-A y Qnr-S	<i>Shevanella algae</i> , <i>Vibrio Splendidus</i>
<b>B-lactamasas*</b>	Hidrolasas	PBP (transpeptidasas)	<i>Kluvera spp.</i>

**Tabla 6.** Reservorios naturales de determinantes de resistencia que afectan a diversos agentes antimicrobianos. PBP: Proteínas de unión a Penicilina. (\*) Nuevos genes de resistencia asociados con bacterias no productoras de antibióticos. Tabla extraída del trabajo, Canton, R. 2009.

Los métodos tradicionales mediante los cuales los microorganismos adquieren resistencia a antibióticos se pueden esquematizar en tres tipos de mecanismos:

#### 1. *Inactivación enzimática*

El ejemplo más típico de este mecanismo está relacionado con la resistencia a antibióticos betalactámicos como la ampicilina. Dentro de este grupo, se incluye, por ejemplo, la desactivación enzimática de la penicilina G



en algunas bacterias resistentes a la penicilina mediante la producción de beta-lactamasas.

### *2. Alteración del punto de acción*

Existen diversas estrategias para alcanzar este objetivo. Destacan algunas como: modificaciones en el gen que codifica la propia diana del antibiótico; como por ejemplo las alteraciones en las PBP de *Streptococcus pneumoniae* que confiere resistencia a penicilina e incluso a ceftriaxona; la adquisición de genes que codifiquen para sustitutos de las dianas originales, como PBP2' en *Staphylococcus spp.* meticilinoresistentes o la dihidrofolato reductasa alternativa en las cepas resistentes a trimetoprima.

La trimetoprima es un antibiótico bacteriostático que inhibe a la enzima dihidrofolato reductasa. La dihidrofolato reductasa participa en la conversión del ácido dihidrofolato a su forma funcional, el ácido tetrahydrofolato. Su inhibición agota las reservas de folato, un cofactor esencial en la biosíntesis de purinas y, en último lugar, del ADN bacteriano, por lo que su inhibición impide la división del microorganismo. Los mecanismos de resistencia suelen estar relacionados con la síntesis de una enzima dihidrofolato reductasa resistente al antibiótico, o la sobre-expresión de esta enzima.

### *3. Alteraciones de la permeabilidad de membrana*

Hay diversos ejemplos:

- Alteraciones de las membranas bacterianas para evitar que penetre el antibiótico o alteraciones en la entrada de antibióticos dependiente de energía, como ocurre en la primera etapa de ingreso de los aminoglucósidos.

- Aumento de la salida de antibióticos. La resistencia por eflujo es un mecanismo inespecífico, que afecta a diferentes grupos de antibióticos como betalactámicos, quinolonas, tetraciclinas y cloranfenicol.

En los últimos años, se han identificado mecanismos de resistencia a antibióticos en bacterias de ambientes naturales que no producen antibióticos.

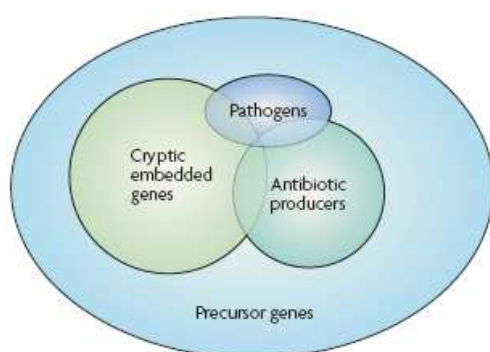
La mayoría de los mecanismos en las bacterias originales, son determinantes de resistencia crípticos que se expresan eficientemente en el nuevo hospedador. Un ejemplo son las  $\beta$ -lactamasas CTX-M de amplio espectro, que constituyen uno de los mecanismos de resistencia más importantes de las cefalosporinas de amplio espectro, o las proteínas

relacionadas con las topoisomerasas Qnr, que inhiben la actividad de las fluoroquinonas (Cantón, 2006; Robiscsek, 2006).

## 10.2 Concepto de resistoma

Las herramientas metagenómicas han identificado nuevos genes de resistencia a antibióticos de ambientes naturales (Riesenfeld *et al*, 2004; Mori *et al*, 2008; Allen *et al*, 2009 Kazimierczak, K.A *et al*, 2000). Estos genes forman parte de la maquinaria preexistente en los aislados bacterianos capaces de producir sustancias con actividad microbiana o, como últimamente se está demostrando, son moléculas antimicrobianas u otro tipo de moléculas con otro papel en su ambiente natural (Linares, 2006; Martínez J.L., 2008).

El término “**resistoma**” (figura 14) comprende el conjunto de todos los genes de resistencia a antibióticos de los microorganismos en un ambiente dado, incluyendo los de las bacterias patógenas y no patógenas. Este término está en relación con el concepto de un único reservorio de genes de resistencia, que incluye los genes de los productores de antibióticos y sus precursores que, bajo la apropiada presión selectiva, se pueden convertir en determinantes de resistencia. La mayoría son genes crípticos que normalmente no se expresan en los aislados naturales (Wright, GD. 2007).



**Figura 14.** Esquema del concepto de *resistoma*. El resistoma comprende todos los genes de resistencia a antibióticos. Figura tomada de Wright, GD 2007.

D’Costa y colaboradores (Da costa *et al*, 2006) fueron los primeros en usar el concepto de resistoma: construyeron una librería de más de 480 cepas procedentes de distintos ambientes de *Streptomyces sp.* y buscaron genes de resistencia frente a 21 antibióticos diferentes. Se identificaron determinantes de resistencia en todas las cepas y los mecanismos afectaban no sólo a los antimicrobianos naturales sino también a los sintéticos. Algunos de estos

mecanismos de resistencia no habían sido todavía caracterizados en los aislados clínicos.

Una de las claves del éxito de este tipo de aproximaciones está basada en el sistema de búsqueda de los determinantes de resistencia. Habitualmente se ha usado la PCR como técnica de búsqueda de nuevos genes de resistencia (de la Torre *et al.*, 2003; Henriques *et al.*, 2006; Demaneche *et al.*, 2008), pero mediante esta herramienta solamente se puede acceder a aquellos genes que son similares en secuencia a los ya existentes. De este modo, las aproximaciones metagenómicas te permiten: i) acceder al reservorio de los microorganismos no cultivables ii) no necesitas tener un conocimiento previo de la secuencia y iii) puedes obtener genes completos.

Algunas aproximaciones metagenómicas de este tipo ya han sido aplicadas con éxito en librerías construidas a partir de ADN extraído del suelo, y se han encontrado distintos genes de resistencia a antibióticos (Riesenfeld *et al.* 2004; D'Costa *et al.* 2006; Allen *et al.* 2008). En el trabajo de Allen *et al.*, se construyen librerías metagenómicas a partir de ADN extraído de suelos naturales de Alaska, encontrando diferentes agentes betalactámicos y la primera enzima bifuncional  $\beta$ -lactamasa descrita hasta el momento. Del mismo modo, en el estudio de Riesenfeld y colaboradores, se identifican nueve genes que expresan resistencia a aminoglicósidos y un gen de resistencia a tetraciclina.

Ahora bien, nos podríamos preguntar si estas moléculas tienen el mismo papel “in vivo” que “in vitro”. El éxito en la búsqueda de microorganismos productores de antibióticos en suelo puede llevarnos a pensar que el papel ecológico de este tipo de sustancias pueda ser inhibir el crecimiento de sus competidores (Waksman, 1940). Esto es cierto en parte.

Sin embargo, también es plausible que las moléculas seleccionadas por su capacidad antibiótica en concentraciones terapéuticas puedan tener otro papel en la naturaleza a concentraciones más bajas. Por ejemplo, moléculas implicadas en la comunicación entre microorganismos pueden tener actividad antimicrobiana (A.Fajardo, 2008). Esta dependencia de la concentración podría ser la clave para que determinadas moléculas tengan un papel u otro. Igual puede ocurrir con los determinantes de resistencia a antibióticos. Algunos pueden tener ese papel en la naturaleza, pero también es probable que en condiciones naturales tengan otra función y cuando se los somete a la presión selectiva de la presencia de algún antibiótico actúen como

determinantes de resistencia. Esta hipótesis abre muchas vías a la hora de buscar genes de resistencia a antibióticos y da idea de la enorme diversidad que puede existir en este tipo de moléculas.



# **OBJETIVOS**



La obtención del ADN de una comunidad microbiana en ambientes naturales de forma directa (metagenoma) nos permite acceder a distintos niveles de información:

- (i) A la diversidad genética global (la accesible y la hasta ahora inaccesible mediante técnicas clásicas de microbiología) de dicha comunidad.
- (ii) A la dinámica de dicha comunidad microbiana asociada a cambios medioambientales bióticos y abióticos (presencia de plantas, características del suelo, sequía, etc.).
- (iii) A su posible explotación biotecnológica.

Estos estudios aplicados a ecosistemas microbiológicos poco explorados (p.ej. sobre la rizosfera de especies de plantas endémicas) pueden contribuir a generar un valor añadido extra para su conservación y para un uso racional de esta diversidad biológica presente.

En el contexto de búsquedas de estrategias innovadoras para el estudio de un ecosistema complejo como es el suelo, se plantea en esta Tesis Doctoral el siguiente objetivo principal: **El análisis de la diversidad taxonómica y funcional de la comunidad procariótica presente en la rizosfera de plantas xerófitas del altiplano central de México mediante una estrategia metagenómica.**

Este objetivo general, está directamente ligado a la consecución de los siguientes objetivos específicos:

I Estudio de la influencia de la planta, tipo de suelo y la estacionalidad en la población microbiana mediante el análisis del gen marcador basado en la subunidad pequeña del ribosoma (16S ARNr). Identificación y caracterización de los phyla procarióticos mayoritarios.

II. Generación y análisis funcional de genotecas medioambientales de expresión mediante:

II.1. Análisis de genes marcadores de resistencia a antibióticos (*resistoma*) presentes en las muestras de estudio.

II.2. Caracterización y análisis de enzimas con carácter lipolítico (de posible interés biotecnológico).

II.3. Validación de posibles genes marcadores como indicadores de seguimiento de la microbiota característica de estos tipos de suelos.





# **MATERIAL Y MÉTODOS**



## **1. TOMA DE MUESTRAS**

Como se ha indicado en el Capítulo de Introducción de esta Tesis Doctoral, la zona de estudio se sitúa en el Parque de Tehuacán-Cuicatlán, espacio Natural Protegido de gran importancia ecológica. Esta provincia florística está ubicada entre los estados de Puebla y Oaxaca (ver figura 1 de la Introducción) y es considerada mundialmente una de las áreas de mayor biodiversidad en lo que a cactáceas se refiere.

### **1.1 Geografía, orografía y topología**

La región de Cuicatlán se ubica entre las montañas de la Sierra de Juárez, la Sierra de Huautla, la Sierra de Zongolica y la Sierra de Mixteca e Ixcateca, al noreste del Estado de Oaxaca.

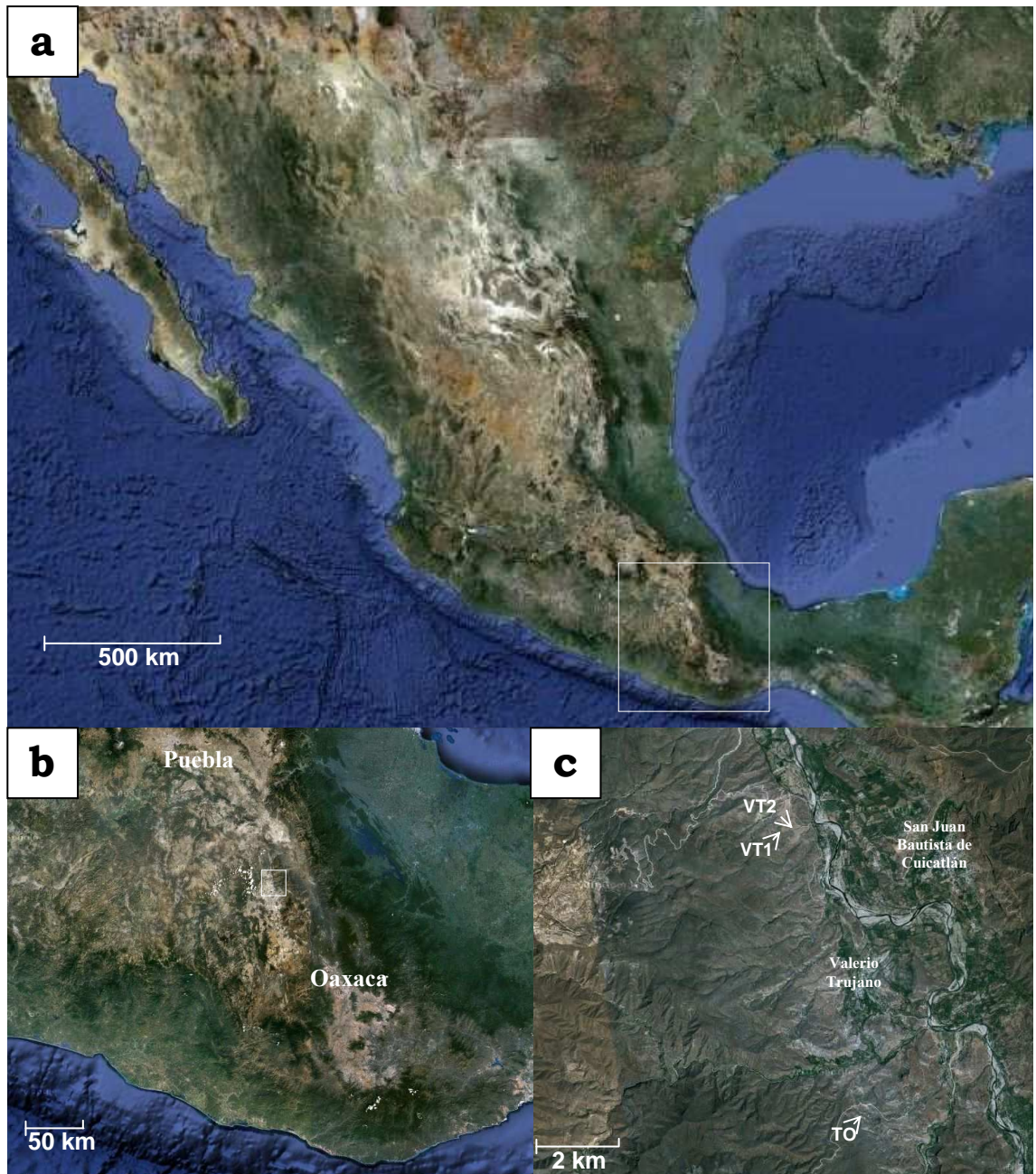
Presenta, dentro de sus límites, distintas zonas que se pueden distinguir en función del tipo de vegetación primaria, humedad relativa, topología y tipo de suelo, entre otros factores. En este sentido, se han seleccionado para nuestro estudio dos zonas diferenciadas entre sí en la parte sur del parque: Tomellín y Valerio Trujano. De esta última, se tomaron muestras de dos sitios diferentes (zona 1 y zona 2).

La primera zona de muestra, Valerio Trujano, Loma Larga (Zona VT; Geo Referencia: N: 17° 48.58', W: 96° 59.90') tiene una altitud de 589 m, la temperatura aproximada a mediodía es: bajo dosel 32°C y suelo abierto 39°C. Constituye una zona modificada por el hombre con claros signos de actividad humana, aunque quedan zonas próximas donde claramente se aprecia la influencia del dosel sobre la colonización de las cactáceas del terreno. Se observa la recuperación del paisaje silvestre tras la actuación humana. Como ya se ha comentado, se muestrearon dos áreas diferentes dentro de esta zona, una que parecía más manipulada por el hombre, denominada VT1, y otra algo más silvestre: VT2.

La tercera zona fue Tomellín (To; Geo Referencia: N: 17° 44.675', W: 96° 58.506', altitud: 755 m). La temperatura: a las 16:00 horas bajo dosel 34°C, abierto: 43°C.

Tomellín se encuentra camino de Sta. María Texcatitlán. Es una zona bien conservada, con gran cobertura vegetal de herbáceas y presencia de leñosas y cactáceas (Zona To: Como peculiaridad, en este caso parecía que la

cactácea globosa *Mamillaria carnea* era menos abundante que en la zona VT, donde realmente parecía la dominante).



**Figura 1.** Mapa de las zonas de muestreo. A) Mapa en relieve de México b) La reserva de Tehuacán-Cuicatlán está localizada entre los estados de Puebla y Oaxaca c) Parte Sur del parque de Tehuacán-Cuicatlán. Las zonas de muestreo se indican con flechas. Se observa como las zonas VT1 y VT2 están más próximas entre sí que la zona TO.

### 1.2. Clima

El clima es un clima seco en el que la evaporación excede a la precipitación. Los meses más fríos son diciembre y enero, con una

temperatura promedio mensual cercana a los 21°C, y el mes más cálido es mayo, con una temperatura media de 29,6°C. El régimen de lluvias es de verano con dos estaciones secas. El clima de la región de Cuicatlán está determinado por el efecto de sombra orográfica que generan las Sierras de Juárez y Huautla, las cuales provocan la retención de la humedad proveniente del golfo.

### **1.3. Tipo de suelo**

Estos suelos se caracterizan por ser del tipo feozen háplico, que se distinguen por tener un horizonte A melánico (de color pardo negruzco). Es el suelo característico de zonas semiáridas con una capa superficial oscura de materia orgánica.

### **1.4. Vegetación**

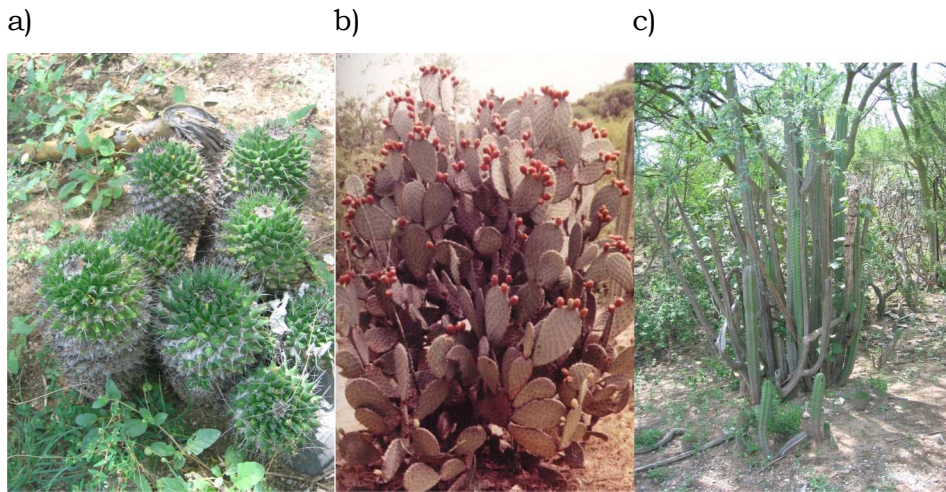
La vegetación de Cuicatlán representa la parte oaxaqueña de esta provincia florística, con elementos vegetales distintivos y únicos (Taktajan, 1986). Los tipos de vegetación están determinados principalmente por factores abióticos como tipo de suelo, altitud, orientación y temperatura.

En las zonas seleccionadas predomina el material xerófilo, que presenta una diversidad enorme en cuanto a componentes vegetales y fisionomía. Generalmente, se compone de arbustos y plantas suculentas, muchas de ellas endémicas.

A pesar de las diferencias en los factores mencionados en el capítulo 1.1 que se pueden establecer entre estas zonas, en todas se puede encontrar como parte de la vegetación natural el mezquite (*Prosopis juliflora*) un árbol espinoso perenne que puede actuar como planta nodriza y bajo cuya fronda se pueden encontrar retoños de cactáceas pertenecientes a diversos géneros y formas, entre los que destacan: *Mammillaria* y *Coryphantha* (cactáceas globosas), *Neobuxbaumia* (cactáceas columnares), *Opuntia* (cactáceas planas o nopales), *Cylindropuntia* (nopales cilíndricos). Dentro de estos géneros hay varias especies que se encuentran amenazadas o en peligro de extinción (*Mammillaria huitzilopochtli*, *M. crucigera*, *M. solilsioides*, *M. pectinifera*).

En este análisis se seleccionó como cactácea globosa de estudio la especie *Mammillaria carnea* (que en los muestreos señalamos mediante las

siglas Mc; figura 2) y como cactáceas columnares las especies *Stenocereus stellatus* (Sc), *Opuntia pilifera* (Op) y *Myrtillocactus geometrizans* (Mg).



**Figura 2.** Especies de plantas que se seleccionaron para este estudio: a) *Mammillaria carnea* b) *Opuntia pilifera* c) *Myrtillocactus geometrizans*.

### 1.5. Toma de muestras en el campo

Esta Tesis Doctoral gira entorno a la búsqueda de especies bacterianas asociadas a cactáceas, por lo que nos centramos en el análisis de la biodiversidad microbiológica presente en 2 ambientes de interés:

- Suelo bajo la influencia de la asociación planta leñosa-cactácea (Do).
- Suelo árido con poca cobertura vegetal, sin presencia de cactáceas o leñosas (Ab).

Este análisis se contrastó en dos épocas del año claramente diferenciadas: época seca y época lluviosa (figura 3). En ellas, el paisaje, las temperaturas y las condiciones ambientales cambian drásticamente.



**Figura 3.** Vista panorámica de la zona de Tomellín durante la época seca (imagen de la izquierda) y la época lluviosa (imagen de la derecha).

### 1.6. Nomenclatura de las muestras

En el muestreo tenemos diferentes variables:

- Dos zonas geográficas con tres grupos de muestras: To y VT (1 y 2).
- Dos ambientes: bajo dosel (Do) y abierto (Ab) (diferenciación original que posteriormente se desechó como menos relevante)

- Cuatro especies vegetales, más suelo no asociado a raíz:

*Mammillaria carnea* (Mc)

*Opuntia pilifera* (Op)

*Stenocereus stellatus* (Ss) y también *Myrtillocactus geometrizans* (Mg)

Muestras no rizosféricas: Bs (Bulk soil).



**Figura 3.** Detalles del muestreo durante la temporada seca (abril 2006).A) Muestreo del suelo rizosférico en la planta *Mammillaria carnea* en la zona de Valerio Trujano 1; B) Recogida del suelo rizosférico; C) Suelo íntimamente unido a raíz que en un primer momento se diferenció del suelo rizosférico.

La nomenclatura se generó siguiendo unas siglas y en función de las características de la muestras. De este modo, los nombres tendrán tres apartados. Como se puede observar en la tabla 1, uno indica la zona donde la hemos tomado, otro el ambiente y el último el tipo de planta. Por ejemplo, una muestra denominada To-Do-Mc será: muestra de suelo de Tomellín (To) bajo Dosel (Do) y rizosfera de *Mammillaria carnea* (Mc).

<b>Zona Geográfica</b>	<b>Ambiente</b>	<b>Especie de planta</b>	<b>Nombre de la muestra</b>
Tomellín (To)	Dosel (Do)	<i>Mammillaria carnea</i> (Mc)	To-Do-Mc
Valerio Trujano 1 (VT1)	Abierto (Ab)	<i>Opuntia pilifera</i> (Op)	VT1-Ab-Op
Valerio Trujano 2 (VT2)	Dosel (Do)	<i>Stenocereus stellatus</i> (Ss)	VT2-Do-Ss

**Tabla 1.** Ejemplo de nomenclatura de las muestras recogidas en función de sus características: zona geográfica, ambiente y tipo de planta.



Cada muestra de suelo, además, se subdividió en: suelo rizosférico y suelo íntimamente asociado a la raíz, esto es, el suelo que resultaba de agitar estrictamente las raíces (figura 3).

También se tomaron otro tipo de muestras no relacionadas con la Reserva y geográficamente muy distantes utilizadas como “suelo control” de una parcela agrícola ampliamente estudiada situada en la Estación Experimental del Zaidín, Granada (Geo Referencia: N: 37° 9.95', W: 3° 35.52'). Las características analíticas del suelo, determinadas según el método de Lachica et al., 1965 (La Chica et al, 1965), fueron: pH 8.1; P, 6.2mg/kg; N, 0.3mg/kg y K,132mg/kg. La composición del suelo fue: 35.8% arena, 43.6% limo, 20.5% arcilla y 1.8% materia orgánica.

## **2. ESPECIES Y CEPAS BACTERIANAS**

Todas las especies y cepas utilizadas en los ensayos de laboratorio se incluyen en la tabla 2.

La cepa de *E.Coli* XL1-Blue permite la correcta multiplicación de los vectores fagémidos en los que se construye la librería metagenómica (capítulo 11 de Material y Métodos) y, por tanto, la aparición de calvas de lisis; mientras que la cepa XL0LR es en la que se mantienen los plásmidos y se usa en la escisión de los vectores fagémidos a plásmidos.

La cepa DH5a se utiliza para el mantenimiento de los plásmidos, debido a que posee mejores cualidades de crecimiento que las anteriores y para el subclonado de los fragmentos.

Por su parte, las células *E. coli* Rosetta-gami B(DE3) se han usado para la correcta sobreexpresión y purificación de proteínas (capítulo 12), mientras que las células *E. coli* HB101 y *E. coli* ES1301 mutS se han utilizado para la construcción de mutantes (capítulo 15 del Material y Métodos).

CEPA BACTERIANA	CARACTERÍSTICAS	REFERENCIA
<i>Escherichia coli</i> DH5a	F <sup>-</sup> , ø80dlacZΔM15, Δ(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17(rK <sup>-</sup> , mK <sup>+</sup> ), phoA, supE44, λ <sup>-</sup> , thi-1, gyrA96, relA1	Bethesda Research Lab.
<i>Escherichia coli</i> XL1Blue	Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F <sup>-</sup> proAB lacIqZΔM15 Tn10 (Tetr)]	Stratagene
<i>Escherichia coli</i> XL0LR	Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 thi-1 recA1 gyrA96 relA1 lac [F <sup>-</sup> proAB lacIqZΔM15 Tn10 (Tetr)] Su <sup>-</sup> (nonsuppressing) λr (lambda resistant)	Stratagene
<i>E. coli</i> Rosetta-gami B(DE3)	F <sup>-</sup> recA1 hsdR(r <sub>K12</sub> <sup>-</sup> m <sub>K12</sub> <sup>+</sup> ) (DE3) (Rif <sup>R</sup> )	Novagen
<i>E. coli</i> HB101	supE44, Δ(mcrC-mrr), recA13, ara-14, proA2, lacY1, galK2, rpsL20, xyl-5, mtl-1, leuB6, thi-1	Promega
<i>E. coli</i> ES1301 mutS	lacZ53, mutS201::Tn5, thyA36, rha-5, metB1, deoC, IN(rrnD-rrnE)	Promega
<i>Sinorhizobium meliloti</i> GR4	Aislado de suelo de la EEZ; Nod <sup>+</sup> , Fix <sup>+</sup>	Casadesús y Olivares, 1979

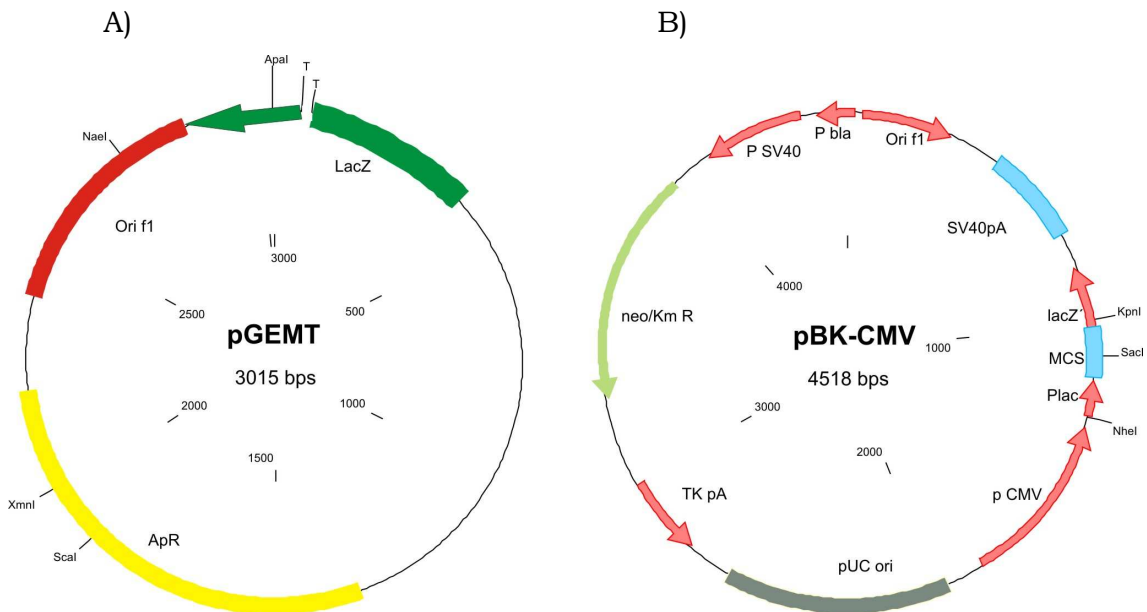
**Tabla 2.** Especies y cepas utilizadas en los diferentes ensayos.

### 3. PLÁSMIDOS Y VECTORES DE CLONAJE

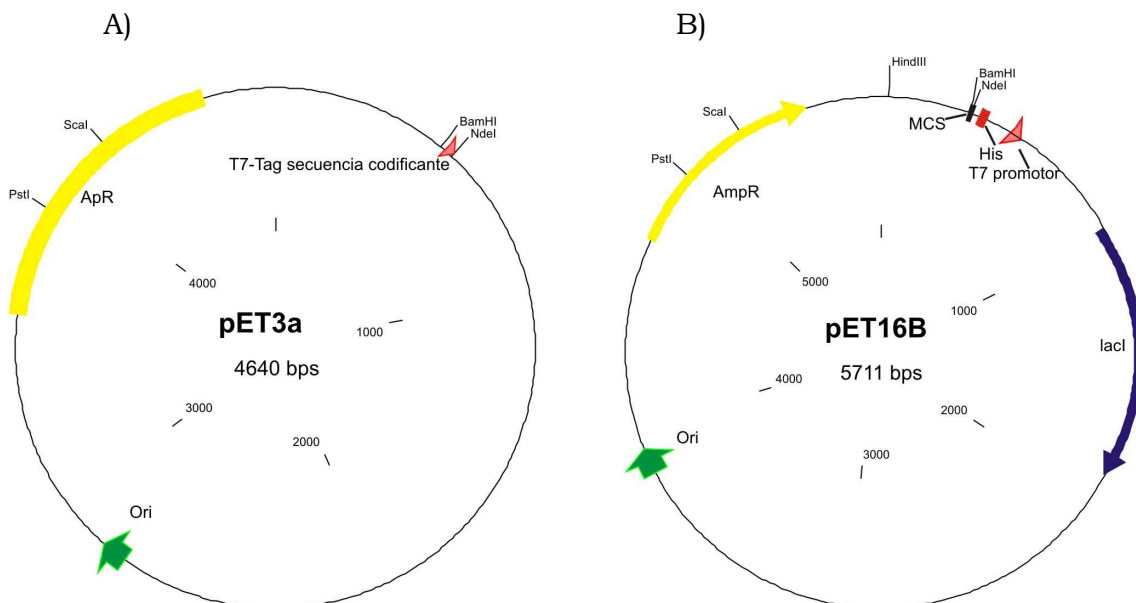
Se utilizaron una serie de vectores para el clonaje de fragmentos dependiendo de la estrategia diseñada en cada momento. Por lo general, se utilizaron vectores de alto número de copia y que permiten la selección de colonias blancas como pGEM-T (Promega) o pKSBluescript (Stratagene).

También se usaron vectores para la correcta purificación y sobreexpresión de las proteínas analizadas en este estudio (figura 5 y Tabla 3).

Además de vectores relacionados con genotecas de expresión construidas den fagos, vector pBCK-CMV.



**Figura 4.** Vectores de clonación. A) Vector pGEMT de alto número de copias utilizado rutinariamente para la clonación de productos de PCR. B) Vector pBK-CMV, este vector es el resultado de la escisión del genoma del fago lambda a plásmido (capítulo 11 de Material y Métodos).



**Figura 5.** Vectores que junto a las células de *E.coli* Rosetta-gama B (DE3) han sido utilizados para la correcta sobreexpresión y purificación de proteínas. A) Vector pET3a que no posee colas de histidinas. B) Vector pET16B con colas de histidinas.

Nombre	Uso	Referencia
pBlueScript KS/SK	Clonaje del fragmento genómico que contiene la secuencia 135F. Proporciona ApR	Stratagene
pGEM-T/pGEM-T easy	Clonaje de productos de PCR. ApR	Promega
Vector Zap Express <sup>R</sup>	Vector usado para la construcción de las librerías metagenómicas en fagos.	Stratagene
pBK-CMV	Plásmido resultante de la excisión de fagos en las librerías metagenómicas. KmR	Stratagene
pET-3a	Vector usado para la sobreexpresión de proteínas.	Novagen
PET-16b	Vector usado para la sobreexpresión de proteínas, con una cola de histidinas	Novagen

**Tabla 3.** Plásmidos usados en este trabajo.

#### 4. OLIGONUCLEÓTIDOS

Para la síntesis de oligonucleótidos se ha utilizado el servicio del Instituto de Parasitología y Biomedicina “López-Neyra” (C.S.I.C., Granada). Según indicaciones del jefe de servicio, los oligonucleótidos empleados en este trabajo se han sintetizado con *DNA Synthesizer*<sup>®</sup> 3400 (Applied Biosystem). También se han usado oligonucleótidos sintetizados en la casa comercial VWR.

En los anexos tras cada capítulo de resultados se presentan tablas que contienen los oligonucleótidos utilizados en la construcción de los mutantes descritos en el correspondiente capítulo. Aquellos oligonucleótidos específicos de cada técnica se detallan junto con la técnica en la que se utilizaron.

## 5. CULTIVOS CELULARES Y ANTIBIÓTICOS

### 5.1. Medios de cultivo bacterianos y antibióticos

#### **Medio LB**

Como medio de cultivo rutinario de las cepas de *Escherichia coli* se utilizó el medio de Luria-Bertani (LB) (Sambrook *et al.*, 1989), preparado con agua desionizada:

NaCl 5 g/l

Triptona (DIFCO) 10 g/l

Extracto de levadura 5 g/l

Para solidificar el medio se añadió agar (PANREAC) al 1,6%. El medio se esterilizó en autoclave a 120°C durante 20 minutos.

#### **Medio TY**

El medio completo utilizado para el crecimiento rutinario de *S. meliloti* fue TY (Beringer, 1974), preparado en agua desionizada:

$\text{CaCl}_2 \times 2\text{H}_2\text{O}$  0,9 g/l

Triptona (DIFCO) 5 g/l

Extracto de levadura (DIFCO) 3 g/l

Para la preparación del medio sólido se añadió agar (PANREAC) al 1,6%. La esterilización se realizó en autoclave durante 20 minutos a 120°C.

#### **Medio NZY**

Como medio de cultivo rutinario para la infección y propagación del fago lambda en las células de *E.coli* XL1-Blue se usó el medio NZY preparado con agua desionizada:

NaCl 5g/L

$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  2g/L

Extracto de levadura (DIFCO) 5g/L

Casaminoácidos o hidrolizado de caseína 10 g/L

En la solidificación del medio se le añadió agar (PANREAC) al 1,6%. El medio así preparado se esterilizó en el autoclave durante 20 minutos a 120 grados.

Para la preparación del top-agar que se usa en el mantenimiento del fago lambda, se mezclaron en igual proporción medio NZY líquido y medio NZY sólido fundido y se mantuvo a 45°C para evitar que gelificara antes de su uso.

### **Medio M9-RhodaminaB**

Como medio de cultivo discriminante de cepas con carácter lipolítico, se usó el Medio M9 con Rhodamina B. Si la bacteria produce alguna lipasa, al hidrolizar el lípido y liberar el ácido graso producirá una disminución del pH del medio, que en la placa se observará como un halo de color amarillo-naranja bajo la luz ultravioleta. Este medio se prepara de la siguiente forma con agua desionizada:

1°. Se prepara Medio M9 (MM).

Na<sub>2</sub>HPO<sub>4</sub> 60 g/L

KH<sub>2</sub>PO<sub>4</sub> 30 g/L

NaCl 5 g/L

NH<sub>4</sub>Cl 2 g/L

Se ajusta el pH a 7,4 y se autoclava durante 20 minutos a 120°C.

2°. Se le añaden en esterilidad los suplementos al medio M9.

- Glucosa al 2% preparada en agua desionizada y esterilizada por filtración.
- Prolina y leucina (20 µg/ml) preparadas en agua desionizada y esterilizadas por filtración.
- Tiamina (33 µg /ml)
- SO<sub>4</sub>Mg 0,01 M
- Cl<sub>2</sub>Ca 0,01 M

3°. Se prepara el medio mezclando en esterilidad todo lo demás de la siguiente forma:

200 ml de Medio Mínimo M9

2 ml Rhodamina B 1mg/ml

6.250 Aceite de oliva

2 ml de Cicloheximina 10 mg/ml

En la solidificación del medio se le añadió agar purificado y esterilizado (OXOID) al 1,6%.

### **Antibióticos**

La adición de antibióticos a los medios de cultivo se hizo a partir de soluciones 100 veces concentradas preparadas en agua desionizada (Km, Ap, Sm, Gm), en agua-etanol al 50% (Tc, Cm) o en metanol al 100% (Tm). Las soluciones preparadas en agua se esterizaron por filtración utilizando unidades Minisart® NML (Sartorius) de 0,2 µm de tamaño de poro.

<b>Antibiótico</b>	<b>Efecto</b>	<b>E.Coli</b>
Tetraciclina (Tc, Sigma)	Bactericida	10 mg/L
Kanamicina (Km, Roche)	Bactericida	200 mg/L
Ampicilina (Ap, Sigma)	Bacteriostático	50 mg/L
Streptomycin (Sm, Sigma)	Bactericida	325 mg/L
Gentamicina (Gm, Sigma)	Bactericida	100 mg/L
Cloroanfenicol (Cm, Sigma)	Bacteriostático	50 mg/L
Trimetoprima (Tm, Sigma)	Bactericida	50 mg/L
Cicloheximida (ClhX, Sigma)	Fungicida	10 mg/L

**Tabla 4.** Concentraciones finales de los antibióticos precisados en los diferentes protocolos para el crecimiento de la cepa *E.coli*.

### 5.2. Aislamiento de cepas de suelo con actividad lipolítica

Para ello se partió de 1 gr de suelo húmedo que se resuspendió en 25 ml de sarcosil 0,1% en tampón TE (previamente esterilizado por filtración). Tras 1 h en agitación a 229 rpm a temperatura ambiente se centrifugó a 3000 rpm 5 minutos y se sembraron diluciones seriadas en medio mínimo.

Posteriormente, se hicieron réplicas en placa que contenían sustratos específicos (como la Rhodamina B y Trioleína) para confirmar si estas bacterias del suelo eran o no lipasas positivas.

### 5.3. Condiciones de cultivo

#### *E.coli*

La cepa bacteriana *E.coli* se incubó a 37°C y su tiempo de generación en medio líquido fue de aproximadamente 30 min a 190 r.p.m. en agitador orbital.

Para la construcción de las genotecas en fagémidos y su posterior escisión a plásmido se utilizaron las cepas de *E.coli* XL1-Blue y XL0LR respectivamente. Ambas se crecieron a 37°C en medio Luria-Bertani (LB) suplementado con 0,2% maltosa y 10 mM de MgSO<sub>4</sub>.

La maltosa es necesaria para que se exprese en la membrana bacteriana el receptor correspondiente a este azúcar, que es por el que penetra el fago a la célula, y el MgSO<sub>4</sub> ayuda a equilibrar las cargas de la membrana para que el fago se pueda adherir mejor a la misma.

Para el mantenimiento de los plásmidos se usaron las cepas XL0LR y DH5 $\alpha$  crecidas a 37° en medio LB con el antibiótico correspondiente de cada plásmido.

### ***S.meliloti***

Las distintas cepas de *S.meliloti* se crecieron a 28°C. Su tiempo de generación en medio líquido fue de aproximadamente 2,5 h a 190 r.p.m. en agitador orbital.

## **5.4. Conservación de cultivos bacterianos**

Para la conservación prolongada de los cultivos bacterianos se utilizó la crio-conservación. Este método se basa en la paralización del metabolismo celular al disminuir la disponibilidad de agua. Para el mantenimiento de la viabilidad celular durante el periodo de conservación se utilizó glicerol 20% (v/v) como crioprotector. El glicerol estéril dispuesto en criotubos se mezcló con cultivos crecidos (DO 660>1). Los viales se congelaron rápidamente en Nitrógeno Líquido y se almacenaron en Ultracongeladores a -80°C.

## **6. AISLAMIENTO DE ÁCIDOS NUCLEICOS**

En este apartado vamos a describir cómo se obtuvieron la mayoría de las materias primas utilizadas en los diferentes ensayos.

### **6.1. Extracción de ADN de muestras de suelo**

La extracción de ADN total del suelo se hizo partiendo de 0,5 gramos de muestra y siguiendo el protocolo que describió Porteous et al, que se resume en varias etapas en las que se combinan diversas técnicas enzimáticas y de calor para lisar las células y extraer el ADN lo más limpio posible sin contaminación con compuestos fenólicos, ya que solo de esta forma será posible su uso posterior con éxito.

De este modo, se dividen y se pesan las muestras de suelo (0,5g) y se someten a varios tratamientos:

i) *Homogenización de la muestra.* Se usan detergentes y después se somete a sonicación suave (en un baño de sonicación de agua), cambios bruscos de temperatura y centrifugación a 4°C.



ii) *Concentrado del extracto*. Se concentra la muestra de ADN añadiendo acetato potásico, polietilenglicol y glicógeno y, se deja precipitar a  $-20^{\circ}\text{C}$ .

iii) *Lavado del precipitado*. En este último paso se procede a limpiar el ADN con distintas centrifugaciones y con dos precipitaciones: la primera con isopropanol y la última con etanol al 100%; iv) *Eliminación de sales*. Se resuspende la muestra en TAE y se vuelve a concentrar en un sistema de diálisis basado en filtros para microcentrífuga Microcon (Millipore YM100).

## **6.2. Extracción de ADN plasmídico**

Dependiendo de las necesidades de cada ensayo, se utilizó uno u otro protocolo. En líneas generales, el aislamiento mediante kit comercial se restringió a la preparación de plásmidos para secuenciación y en protocolos que precisasen de un ADN libre de ARNasas. El procedimiento de lisis seguida de precipitación con sales de magnesio es apropiado para extracción de plásmidos en en células de *E. coli*.

### **6.2.1. Extracción de ADN plasmídico por precipitación con sales de magnesio**

Esta técnica está basada en el método de Studier (1991), que permite el aislamiento rápido de plásmidos de *E.coli*. Recogemos 1,5 ml de cultivo bacteriano por centrifugación y, tras retirar el sobrenadante, se resuspenden las células en 100  $\mu\text{l}$  de agua desionizada. La lisis tiene lugar por la adición de 100  $\mu\text{l}$  de una solución compuesta por NaOH 0,1 M, EDTA 10 mM y 2% SDS. Se homogeneiza por agitación y se hierven durante 2 min. El ADN lineal, normalmente cromosómico, se precipita por la adición de 50  $\mu\text{l}$  de  $\text{MgCl}_2$  1 M. Tras homogeneizar completamente (aparición de precipitado blanco), se incuba en hielo de 2-5 min. Se centrifuga 1 min para sedimentar el precipitado en el fondo del tubo y añadimos 50  $\mu\text{l}$  de acetato potásico 5 M para eliminar proteínas. El tubo invertido se mezcla con vórtex evitando que se levante el precipitado blanco del fondo y se incuba 5 min en hielo. Se centrifuga durante 5 min y recogemos el sobrenadante a un tubo nuevo donde precipitaremos el ADN plasmídico con 0,6 ml de etanol 100% guardado a  $-20^{\circ}\text{C}$ . Se deja 5 min a temperatura ambiente y se centrifuga 5 min. Se lava el sedimento con 200  $\mu\text{l}$  de etanol al 70% y dejamos secar a vacío. Se resuspende en el volumen adecuado dependiendo de la concentración deseada o necesaria para las

diferentes aplicaciones: lo más habitual para un plásmido de alto número de copias es resuspender en 25  $\mu$ l de agua desionizada con ARNasa a una concentración de 10  $\mu$ g/ml.

### **6.2.2. Aislamiento de ADN plasmídico mediante kit comercial**

Hemos usado distintos kits comerciales para el aislamiento de ADN plasmídico. En todos los casos se han seguido las indicaciones de uso del proveedor: Wizard® Plus SV Minipreps DNA Purification Systems (Promega), QIAprep Spin Miniprep Kit (Qiagen) y PerfectPrep® plasmid mini (Eppendorf). El fundamento de estas preparaciones es la lisis celular seguida de la unión del ADN en disolución a una resina/membrana y elución con agua. La ventaja fundamental es la limpieza del ADN y la ausencia de ARNasa en los pasos finales.

### **6.3. Cuantificación de la concentración de ADN**

El cálculo de la concentración de ácidos nucleicos se estimó siguiendo el método espectrofotométrico descrito en Sambrook et al. (1989). Se midió la absorbancia de las muestras en solución acuosa a longitudes de onda 260 y 280 nm utilizando cubetas de cuarzo. La concentración se determinó en función de los valores estándar: 1 U a D.O.260nm se corresponde con 50  $\mu$ g/ml de ADN de cadena doble, 40  $\mu$ g/ml de ADN de cadena sencilla y ARN, y 33  $\mu$ g/ml de oligonucleótidos. La relación D.O.260nm / D.O.280nm se usó como estimador de la pureza de la preparación, considerándose valores por debajo de 1,8 como indicadores de contaminación por proteínas o fenol.

Generalmente, se ha usado un espectrofotómetro (NanoDrop® ND-1000) que permite trabajar con volúmenes de muestra 1-2  $\mu$ l; facilitando el proceso de análisis.

## **7. MANIPULACIÓN Y AMPLIFICACIÓN DE ADN**

### **7.1. Digestión con enzimas de restricción**

La digestión completa del ADN se realizó siguiendo las indicaciones de temperatura y tampón óptimos recomendadas por los proveedores (Roche y New England Biolabs). Lo más común es utilizar 1-5 U de enzima/ $\mu$ g ADN a

digerir (entre 0,1 y 0,2  $\mu$ l de enzima por 10-15  $\mu$ l de volumen final de la digestión). En el caso de ADN total, esta cantidad suele duplicarse así como el tiempo de incubación.

En el caso de digestiones con 2 enzimas de restricción simultáneamente hay que tener precaución con el tampón elegido. Debemos tratar de buscar un tampón compatible para ambas (instrucciones del proveedor) y, en caso de no ser posible, realizar primero la digestión con la enzima que precise el tampón con menor fuerza iónica. Si aún así no fuese factible, el ADN se limpiaría de sales y enzima tras la primera digestión, y se procedería con la segunda, así sucesivamente. Hay dos alternativas para eliminar las sales: la primera es el uso de kits comerciales como QIAquick PCR Purification Kit (Qiagen), pero también podemos optar por fenolizar y precipitar nuestro ADN.

### **7.2. Reacción de defosforilación**

En ocasiones es necesaria la eliminación de los grupos fosfato presentes en los extremos 5' de vectores linearizados para evitar que vuelvan a ligarse. Esto nos evita el análisis de muchas colonias que sean vector re-ligado en vez de contener el inserto que nos interesa. La defosforilación de 5  $\mu$ g de plásmido digerido precisa 1 U de fosfatasa alcalina (CIP, Calf Intestine Phosphatase; Roche) en su tampón correspondiente durante 1 h a 37°C. La reacción se para añadiendo 55  $\mu$ l de una solución compuesta por tampón STE 1x [1 mM Tris-HCl pH8, 100 mM NaCl y 1 mM EDTA] y SDS 1% e incubando 15 min a 68°C. Para obtener mejor rendimiento en la ligación, es conveniente fenolizar y precipitar.

### **7.3. PCR**

La reacción en cadena de la polimerasa (Polymerase Chain Reaction, PCR) consiste en la amplificación en sucesivos ciclos de un material de partida gracias al uso de una pareja de finalidades y en nuestro caso la hemos utilizado para amplificar insertos que después necesitamos clonar añadiendo determinados enzimas en los extremos, para la construcción de mutantes, para la comprobación rutinaria de la presencia de un determinado inserto o para la amplificación del gen ribosomal 16S y su posterior clonación en vectores como pGEMT o resolución mediante distintas técnicas como el TGGE.

Consiste en 3 pasos fundamentales repartidos en 3 ciclos: desnaturalización, anillamiento y extensión. La desnaturalización suele realizarse a 94°C 30 s aunque el tiempo depende del tipo de molde; al comienzo del programa de PCR suele incluirse un primer ciclo de desnaturalización de 2 min para disponer de todo el ADN como cadena simple. La temperatura de anillamiento depende de la composición de los cebadores pero suele rondar los 55°C y el tiempo estándar es 30 s. La extensión no es más que la temperatura óptima para la actividad de la enzima (generalmente 72-74°C); el tiempo de este ciclo suele estar en función del tamaño del amplificado y del tipo de enzima usado (por ejemplo, Taq polimerasa 1Kb/30 s). Además el programa de PCR suele terminarse con un ciclo adicional de extensión de 5-10 min para completar los fragmentos. Dependiendo de la finalidad se utilizaron distintas polimerasas: TripleMaster® (Eppendorf), High Fidelity® (Roche) o Pfu polimerasa (Promega), si era necesaria una baja tasa de errores en el producto final (como ocurre con la amplificación del gen ribosomal 16S donde se usó la enzima High Fidelity de Roche).

Las reacciones se llevaron a cabo generalmente en un volumen final de 25 µl conteniendo 0,1 a 100 ng del ADN molde, 25 pmol de cada uno de los cebadores específicos, 0,1 mM de deoxinucleósidos trifosfato, el tampón especificado para cada enzima y 2 unidades de ADN polimerasa. La amplificación se realizó en un termociclador Eppendorf Mastercycler® y el programa empleado varió dependiendo del objetivo.

La limpieza de este tipo de productos (eliminación de cebadores y sales) se realizó mediante columnas MicroSpin™ S-300HR (GE Healthcare) o con el uso del QIAquick PCR Purification Kit (Qiagen).

### **7.3.1. PCR de colonia**

Es una estrategia muy rápida para la comprobación de la presencia de determinados insertos a partir de las colonias en placa. Se trata de repartir los 25 µl de la mezcla de reactivos de PCR y, añadir el ADN molde aplicando directamente una colonia de la placa mediante un palillo y agitando dentro de la mezcla.

## **8. ELECTROFORESIS DE ÁCIDOS NUCLEICOS**

La electroforesis es una técnica separativa basada en la migración diferencial en sentido y velocidad de partículas cargadas en el seno de un campo eléctrico, que fue desarrollada por Tiselius en 1937. Consiste en la separación de macromoléculas en función de su tamaño y carga. Los ácidos nucleicos, con carga negativa, migran hacia el polo positivo.

### **8.1. Electroforesis en geles de agarosa no desnaturalizantes**

El análisis de productos de PCR, de ADN total o plasmídico o de fragmentos de restricción se llevó a cabo mediante electroforesis horizontal en geles de agarosa (agarosa SeaKem® LE, Cambrex/Iberlabo) al 1% en TAE 1x: [TrisHCl 40 mM, EDTA Na<sub>2</sub> 2 mM, ácido acético glacial 0,1142%(v/v)]. El voltaje habitual de trabajo fue de 120 V aunque determinados procedimientos requirieron velocidades inferiores (Cubetas de BioRad). Como tampón de carga se utilizó una solución 6x de 0,25% (p/v) de azul de bromofenol y glicerol 45% (v/v).

El revelado de los geles de agarosa se realizó por inmersión de los mismos en una solución de bromuro de etidio 1 mg/ml en agua durante 15 a 20 min. Este intercalante (cancerígeno) permite visualizar ADN o ARN en un transiluminador que emite luz UV de 365 nm de longitud de onda. Las fotografías de los geles se realizaron con Gelprinter y Gelstation de TDI para grabar y procesar las imágenes.

### **8.2. Electroforesis en geles de poliacrilamida no desnaturalizantes**

Para la resolución de fragmentos inferiores a 400 bp, se recurre a electroforesis vertical en geles de poliacrilamida.

En este caso el tampón de electroforesis suele ser TBE 1X [Tris-HCl 0,089 M pH 8, ácido bórico 0,089 M y EDTA 0,002 M]. Hemos trabajado con disoluciones 40% (29:1) de acrilamida:bis-acrilamida (Bio-Rad). El porcentaje y el grosor del gel varía según las necesidades y con éstos también cambian las proporciones relativas del agente polimerizante (TEMED, Sigma) y el catalizador (persulfato amónico, APS; Sigma).

Para el revelado de estos geles se utilizó la tinción con nitrato de plata (Heuer et al., 1997). Se fija con una solución compuesta por etanol 10% (v/v) y ácido acético glacial 0,5% (v/v) (2 veces durante 3 min). A continuación, se incubaba en agitación durante 30 min con una solución de nitrato de plata 0,2% en agua. Se lava 4 veces con agua para eliminar el exceso de plata y se revela con una mezcla compuesta por tetrahidrobórato sódico 10% (p/v)/formaldehído 0,4% (v/v)/NaOH 1,5% (p/v), controlando la aparición de las bandas. Se lava con agua destilada y el revelado se para por la incubación con carbonato sódico 0,75%. Antes de secarlo, es conveniente mantener el gel toda la noche en una solución de etanol 25% y glicerol 10%.

### **8.3. Marcadores de peso molecular**

Los marcadores de peso molecular fueron en todos los casos comerciales:

- Marcador II: ADN del fago  $\lambda$  digerido con la enzima HindIII. Está compuesto por 8 fragmentos que abarcan desde los cientos de nucleótidos hasta más de 20Kbp (Roche; Universidad Autónoma de Madrid). También se dispone marcado con digoxigenina que marcan el peso molecular en las hibridaciones ADN-ADN.

- Marcador III: ADN del fago  $\lambda$  digerido con las enzimas HindIII y EcoRI. En este caso, el patrón de digestión comprende 13 fragmentos que abarcan desde los cientos de nucleótidos hasta más de 20Kbp (Roche). También marcado con digoxigenina.

- Marcador pGEM: ADN del plásmido pGEM-T digerido con las enzimas HinfI y EcoRI generando 15 fragmentos que se encuentran entre las pocas nt hasta las ~2500 bp (Promega).

- Marcador  $\Phi$ 29: ADN del fago con el mismo nombre digerido con HindIII. El patrón de digestión se compone de 14 fragmentos: 72, 156, 273, 453, 579, 611, 759, 1150, 1331, 1933, 2201, 2498, 2899 y 4370 bp.

### **8.4. Electroforesis en gel en gradiente de temperatura (TGGE)**

Para la resolución de los amplificadores del gen ribosomal 16S en función de su secuencia, se empleó un sistema de electroforesis basado en geles desnaturizantes con gradiente de temperatura (TGGE Maxi, BIOMETRA),

que fue utilizado según las especificaciones del fabricante. Se utilizaron geles con un 6% de acrilamida/bisacrilamida, 1% TAE, 8 M Urea, 20% Formamida desionizada y 2% Glicerol. La polimerización fue iniciada con 100µl de TEMED y 200µl de Persulfato Amónico al 10%. Posteriormente, se cargaron entre 2-8 de producto de PCR al cual se le añadió 1/5 de tampón de carga (5xTAE, 0,1% Tritón X-100 y 0,01 de azul de bromofenol). La electroforesis se realizó durante 16h a 130 V, bajo un gradiente térmico de 44,5°C a 56°C para los fragmentos de 16S ARNr.

La presencia de las diferentes bandas del gel fue evidenciada mediante tinción con nitrato de plata (Heuer et al, 1997). Todos los pasos se realizaron en agitación: dos lavados de 3 minutos en solución de fijación, incubación durante 15 minutos en solución de tinción, cuatro lavados (aproximadamente 4 segundos cada vez) con agua destilada, inmersión en solución de revelado hasta que se observen todas las bandas, lavado con agua destilada e inmersión en la solución de parada durante 10 minutos. Finalmente, se dejó reposar en solución de conservación durante 7 minutos.

<b>Solución</b>	Composición
Fijación	10% Etanol, 0,5% Ácido Acético
Tinción	0,2% Nitrato de plata
Revelado	1,5% NaOH, 0,01 Borohidruro Sódico, 0,15% Formaldehído
Paro	0,75% p/v Carbonato sódico
Conservación	25% w/w etanol, 10% v/v glicerol

**Tabla 5.** Composición de las soluciones empleadas en la tinción con plata.

## **9. CLONAJE DE FRAGMENTOS**

En este apartado se incluyen distintos protocolos utilizados habitualmente para el clonaje de insertos de interés en vectores.

### **9.1. Aislamiento de fragmentos de ADN de gel de agarosa**

Para el aislamiento de fragmentos de restricción separados por electroforesis horizontal en gel de agarosa se utilizaron kits comerciales de acuerdo con las directrices proporcionadas por los proveedores: PerfectPrep

Gel Cleanup® (Eppendorf) y/o QIAquick Gel Extraction Kit (Qiagen). El paso previo para el uso del kit precisa cortar la banda de interés, para lo que nos servimos del transiluminador UV, una vez se ha teñido el gel con bromuro de etidio (aunque se procura minimizar la exposición del ADN a la luz UV).

## **9.2. Ligación de fragmentos de restricción**

La unión covalente de dos moléculas de ADN lineal se lleva a cabo mediante la adición de 1 U T4 ADN ligasa (Roche) en su tampón adecuado. La proporción molar más habitual vector:inserto suele ser 1:3. Las reacciones se llevaron a cabo en un volumen final de 10-15 µl a 14-16°C durante toda la noche.

### **9.2.1 Clonaje de fragmentos amplificados por la reacción de PCR**

En ocasiones el producto que se precisa ligar no es un fragmento de restricción sino un producto de PCR. En el mercado se encuentran disponibles vectores que permiten la clonación de amplificados con T protuberantes (extremos que deja la Taq polimerasa) como por ejemplo el vector pGEMTeasy (PROMEGA, figura 4), pero algunas polimerasas con actividad exonucleasa 3'-5' dejan extremos romos. Por eso es un procedimiento habitual la adición de adeninas en los extremos de estos fragmentos; este procedimiento está catalizado por la Taq polimerasa (5 U) en presencia de 0,2 mM de dATP (70°C durante 30 min). Este producto puede ser utilizado directamente en la reacción de ligación.

Estos amplicones se ponen a ligar en una mezcla que contiene: vector pGEMT y el amplicón (en una proporción molar de 1:3), ligasa (3u/µl) y en el tampón de ligación suministrado. La reacción de ligación se deja toda la noche a 16°C.

Posteriormente se transforman células competentes (ver más adelante) con esta reacción de ligación y se seleccionan las colonias blancas que son las que llevarán los fragmentos de diferentes genes ribosomales 16S para su posterior estudio.

Este método adquiere un valor adicional en la generación de las genotecas génicas conteniendo el gen ribosomal 16.



### **9.3. Preparación de células competentes**

Una vez que tenemos el ADN circularizado necesitamos que se replique en la bacteria. Con este fin se utilizan un tipo de células que está preparado para captar ADN exógeno, las células competentes. Su preparación requiere de métodos químicos o físicos que alteran la pared celular de la bacteria. En nuestro caso, sólo se han preparado células competentes de distintas cepas de *E.coli* y observamos una mayor eficiencia de transformación en células electrocompetentes.

#### **9.3.1. Células competentes por métodos químicos: RbCl**

En general, este tipo de protocolos deben llevarse a cabo en condiciones de esterilidad. Es aconsejable que el proceso sea rápido y en frío para que el metabolismo celular se ralentice en el mayor grado posible y consigamos el mayor número de células viables para la transformación.

Se cultivan las células de *E coli* en 100 ml de medio LB hasta que se alcanza una A600 de 0,4 (fase exponencial de crecimiento). Entonces, se detiene el crecimiento incubando el cultivo en hielo durante 15 min. Transcurrido este tiempo, las bacterias se sedimentan por centrifugación (6000 r.p.m., 10 minutos, 4 °C). Con cuidado, se resuspende el sedimento con 32 ml de solución RF1 estéril y preenfriada a 4 °C [RF1 (por 100 ml): 1,2 g de RbCl; 0,99 g de MnCl<sub>2</sub> x 4 H<sub>2</sub>O; 0,294 g de acetato potásico; 0,15 g de CaCl<sub>2</sub> x H<sub>2</sub>O; 11,9 ml de glicerol; agua destilada hasta 100 ml. Ajustar el pH a 5,8 con una solución de ácido acético 0,2 M. Esterilizar por filtración y conservar a 4°C]. Se incubaron en hielo durante 15 minutos y después se centrifugaron 6000 r.p.m. durante 10 minutos. Se retiró el sobrenadante y las células se resuspendieron en 4 ml de solución RF2 preenfriada a 4°C [RF2 (por 50 ml): 0,1046 g de ácido morfolino propanosulfónico (MOPS); 0,06 g de RbCl; 0,55 g de CaCl<sub>2</sub> x H<sub>2</sub>O; 5,95 ml de glicerol, como agente crioprotector ; agua destilada hasta 50 ml. Ajustar el pH a 6,8 con una solución diluida de NaOH. Esterilizar por filtración y conservar a 4°C]. Se repartieron en alícuotas de 100 µl en tubos preenfriados a 4°C y se congelaron inmediatamente con nitrógeno líquido. Se conservaron por un periodo limitado a -80°C. El rango de eficiencia de transformación se encuentra entre 1-5 x 10<sup>5</sup> células/µg ADN.

### **9.3.2. Células electrocompetentes**

Para este protocolo se precisa la limpieza completa de las sales del medio de cultivo de manera que no interfieran con la electroporación, el verdadero causante de la desorganización de la pared celular. A su vez, se añade un agente crioprotector (glicerol) que va a permitir la viabilidad de las células conservadas a  $-80^{\circ}\text{C}$ . Consiste en sucesivos lavados con volúmenes cada vez más reducidos de glicerol 10%; partimos de 500 ml de cultivo a una A600 de 0,5, cuyo metabolismo se retrasa por incubación en hielo durante 20 min. Se recogen a 6000 r.p.m. durante 15 minutos y  $4^{\circ}\text{C}$  y se resuspenden en 100 ml de glicerol 10% frío, volviéndose luego a centrifugar. En el siguiente paso, se resuspenden en 20 ml de glicerol para acabar, tras este segundo lavado, con las bacterias resuspendidas en 2 ml de glicerol 10%. Se reparten en alícuotas de 50  $\mu\text{l}$  que inmediatamente se congelan en  $\text{N}_2$  líquido. Se conservan a  $-80^{\circ}\text{C}$  durante al menos 4-6 meses. El rango de eficiencia de transformación ronda las  $10^6$  células/ $\mu\text{g}$  ADN.

### **9.4. Transformación de células competentes**

Se pueden usar diferentes maneras de inducir la desestabilización de las membranas celulares. En células cuya pared se encuentra dañada, se suele usar un choque térmico para abrir poros en la membrana de la pared. En cambio, las células electrocompetentes deben mantener intacta su pared celular, por lo que nos ayudamos de una descarga eléctrica para cambiar el potencial de membrana y permitir la captación del ADN.

#### **9.4.1. Transformación por choque térmico**

Las células competentes preparadas por métodos químicos se sacaron a hielo hasta su descongelación total (~20 min). Se añadió en condiciones de esterilidad el ADN, procedente de ligación o no, y se dejó incubando con las células durante otros 20-25 min. Las células son sometidas a un choque térmico de  $42^{\circ}\text{C}$  durante 90 seg, para seguidamente devolverlas a hielo durante otros 5 min, en los que se permitirá la captación del ADN. Ahora debemos permitir la recuperación de las células por la adición de 900  $\mu\text{l}$  de LB fresco durante 1 h a  $37^{\circ}\text{C}$  (en agitación a bajas revoluciones). Transcurrido este tiempo, se siembra parte o todo el cultivo (dependiendo de la eficiencia de transformación esperada) en medio LB sólido con el antibiótico

correspondiente (ver tabla 4). Si el sistema de selección de clones con inserto lo permite se añadió además a las placas 40 µg/ml de X-Gal (bromo-chloro-indolyl-galactopyranosido) que permite seleccionar solo las colonias blancas como positivas.

#### **9.4.2. Electrotransformación**

En este protocolo también debimos descongelar las células en hielo, tras lo cual añadimos el ADN con el que deseábamos transformar. Transferimos la suspensión celular a una cubeta de electroporación que se sometió a un pulso eléctrico según las recomendaciones hechas por el fabricante. En nuestro caso, usamos un electroporador Eppendorf 2510 que recomienda un pulso de 1800 V durante 3-5 milisegundos. Como en el caso anterior, se restauró la estabilidad celular por la adición de 950 µl de LB fresco. A continuación, se procedió como en el apartado anterior. Es vital que el ADN con el que queremos transformar esté libre de sales. Para eliminar las sales presentes tras una reacción de digestión o ligación, se dializan las muestras frente a agua utilizando filtros de nitrocelulosa VSWP de 0,025 mm (MILLIPORE).

#### **9.5. Generación de genotecas 16S ARNr**

Para tal fin, se realiza un PCR con los cebadores U1510R (5'-GGTTACCTTGTTACGACTT-3'; Reysenbach y Pace, 1995) y E9F (5'-GAGTTTGATCCTGGCTCAG-3'; McInnery y., 1995; Hansen *et al.*, 1998; Farelly *et al.*, 1995), en el caso de las librerías donde se amplificó el gen 16S ARNr de bacterias en general; o con los cebadores E9F y 31F (5'-GATCCTGGCTCAGAATC-3; Barns SM et al, 1999) en el caso de las librerías donde se amplificó el gen 16S ARNr específico del grupo acidobacterias.

Las condiciones de PCR son las siguientes: 4 minutos a 94°C para desnaturalizar el ADN, seguidos de 30 ciclos, en el caso de muestras realizadas con los cebadores generales para el gen 16S ARNr y 25 ciclos en el caso de muestras realizadas con los cebadores específicos del phylum acidobacterias. Cada ciclo conlleva 1 minuto a 94°C, 1 minuto a 52°C y 1,5 minutos a 72°C. Finalmente, hay un paso de extensión a 72°C durante 10 minutos. La PCR se llevó a cabo con la enzima High Fidelity® (Roche), usando 0,5 µl de ADN molde (a 30 ng/µl de concentración inicial), 0,25 µl de cada

cebador (a 100 pmo/ $\mu$ l de concentración inicial) y agua MilliQ. El resto de reactivos se añadió siguiendo las recomendaciones del fabricante en un volumen final de 25  $\mu$ l. El tamaño de los amplicones fue próximo a los 1500 pares de bases.

Estos amplicones se ponen a ligar en una mezcla que contiene: vector pGEMT y el amplicón (en una proporción normalmente 1:3), ligasa (3u/ $\mu$ l) y tampón de ligación. La reacción de ligación se deja toda la noche a 16°C.

Posteriormente, se transforman células competentes de ClRb con esta reacción de ligación y se seleccionan las colonias blancas, que son las que llevarán los fragmentos de diferentes genes ribosomales 16S para su posterior estudio. Al conjunto de colonias resultantes de este proceso es lo que se denomina librería génica.

#### ***9.5.1. Análisis de las secuencias de los genes ribosomales 16S procedentes de la construcción de las librerías metagenómicas***

Un primer análisis de la diversidad de los clones procedentes de las distintas genotecas realizadas consistió en electroforesis del perfil de restricción con HaeIII de los amplicones correspondientes a 200 clones de cada una de las genotecas realizadas.

Cada perfil distinto fue secuenciado con el cebador universal 1541r con objeto de obtener de cada clon unos 900 nt de lectura mediante secuenciación tipo Sanger. Las lecturas devueltas por el servicio de secuenciación fueron visualizadas con el programa informático Sequence Scanner versión 1.0. Las tareas de edición de secuencias se realizaron con el programa informático Clone Manager Professional Suite versión 6.00. El chequeo de quimeras (secuencias que pertenecen a más de un individuo que son amplificadas como un solo producto de PCR, dando lugar a un amplicón de individuos diferentes) se realizó en el servidor online Bellerophon Server. El alineamiento de secuencias se llevó a cabo en el servidor MAFFT version 6 (Kato, K et al, 2008). Para la clasificación filogenética de las secuencias se utilizaron en paralelo dos estrategias:

- I) Uso del Blastn con cada una de las secuencias.
- II) Introducción de las secuencias en RDP Release 10. (Wang, Q et al, 2007).

Para la inferencia de árboles filogenéticos se utilizó el programa MEGA (Kumar, S et al, 2008).

Para estudiar la biodiversidad, la riqueza y comparar las librerías se utilizaron las herramientas estadísticas DOTUR, SONS y [-LIBSHUFF. Actualmente, estas herramientas se ofrecen como un único paquete informático llamado MOTHUR.

## **10. PREPARACIÓN DE AMPLICONES PARA LA REACCIÓN DE PIROSECUENCIACIÓN**

Además del análisis de los genes 16 ARNr mediante la construcción de librerías génicas, estos genes se pueden analizar secuenciando los productos de PCR directamente mediante algún método de secuenciación masiva, como por ejemplo la pirosecuenciación mediante el sistema titanium GS FLX, que permite obtener lecturas de 400 pb.

La pirosecuenciación de una placa resultaría teóricamente en un millón de lecturas. Para parte de los resultados expuestos en esta Tesis Doctoral, se enviaron dos muestras (correspondientes a muestras rizosféricas en época seca y lluviosa) a pirosecuenciar mediante el sistema titanium GS FLX a la empresa Life Secuenciencing.

Estas dos muestras se enviaron junto a otras 9 más de otros estudios que están siendo desarrollados en el grupo de investigación, en total 11 muestras que se pirosecuenciaron en un cuarto de placa (corresponderían a 250.000 secuencias totales, teóricamente 22.727 lecturas por muestra).

La preparación de las muestras para pirosecuenciar se realizó de la siguiente manera.

Se realiza una PCR usando como molde el ADN metagenómico de estudio y utilizando unos cebadores universales para amplificar así el gen ribosomal 16S, tanto de bacterias como de arqueas. Estos cebadores tienen añadida, además, una secuencia para poder unirse a las bolitas del pirosecuenciador y para que se pueda unir la enzima taq polimerasa (tabla X), así como otra secuencia que nos permite diferenciar entre muestras (código de barras, habrá 11 códigos de barras distintos, cada muestra tendrá unos cebadores concretos con un código de barras que nos permitirá diferenciarla del resto).

De este modo, los cebadores que nosotros usamos fueron el U926R (CCGTCAATTCCTTTRAGTTT; Baker, G.C *et al*, 2003) y el U519F (CAGCMGCCGCGTAATWC; Baker, G.C *et al*, 2003), junto a la secuencia de

unión a las bolitas del pirosecuenciador y a la taq polimerasa y el código de barras.

Nombre	Secuencia 5´-3´
G1F (muestra época seca)	CGTATCGCCTCCCTCGCGCCATCAGAAGCCAGT CAGCMGCCGCGGTAATWCT
G1R (muestra época seca)	CTATGCGCCTTGCCAGCCCGCTCAGAAGCCAGT CCGTCAATTCCCTTTRAGTTT
G2F (muestra época lluviosa)	CGTATCGCCTCCCTCGCGCCATCAGTTGAACTG CAGCMGCCGCGGTAATWC
G2R (muestra época lluviosa)	CTATGCGCCTTGCCAGCCCGCTCAGTTGAACTG CCGTCAATTCCCTTTRAGTTT

**Tabla 6.** Secuencia de los cebadores usados para la pirosecuenciación. En color verde se señala la secuencia de unión a las bolitas y a la taq polimerasa, en color negro la secuencia correspondiente al código de barras (una secuencia ‘alien’ que permitirá diferenciar los amplicones procedentes de nuestra muestra en la placa) y en color rojo la secuencia que anilla al gen 16S ARNr.

La PCR con los primers así diseñados se realizó en las siguientes condiciones: 4 minutos a 94°C para desnaturalizar el ADN, seguidos de 25 ciclos. Cada ciclo conlleva 15 segundos a 94°C, 45 segundos a 55°C y 1 minuto a 72°C. Finalmente, hay un paso de extensión a 72°C durante 10 minutos. La PCR se llevó a cabo con la enzima Master Taq (5 Prime suministrada por VWR), usando 10 ng de ADN Molde, 0,25 µl de cada primer (a 100 pg/µl de concentración inicial) y agua MilliQ. El resto de reactivos se añadió siguiendo las recomendaciones del fabricante en un volumen final de 25 µl. El tamaño de los amplicones fue próximo a los 1500 pares de bases.

La empresa para pirosecuenciar te pide que mezcles todas las muestras y mandes 5 µg de ADN total a una concentración de 300 ng/µl; como mandamos a secuenciar 11 muestras, tenemos que tener cada muestra a una concentración de 30 ng/ µl.

Es muy importante que las muestras en la mezcla se encuentren a una concentración equimolar, porque si no, se pueden pirosecuenciar preferencialmente unas muestras sobre otras.

De igual forma, también es muy importante que no queden restos de cebadores en las muestras que se envían, porque si esto ocurre se secuenciarán estos cebadores y no los productos de PCR. Para solventar este problema, los amplificadores se limpiaron con unas columnas basadas en filtros

para microcentrífuga Microcon YM-100 (Millipore), que te permitían eliminar los cebadores y al mismo tiempo concentrar la muestra hasta el volumen adecuado.

### **10.1. Análisis de las secuencias procedentes de pirosecuenciación**

Debido al elevado número de secuencias que proporciona la pirosecuenciación, su manejo se hace dificultoso. Para el manejo de estas secuencias, se usó la interfaz específica para datos de pirosecuenciación que presenta el RDP Release 10 (<http://pyro.cme.msu.edu/>).

En esta página, se pueden introducir los archivos recibidos tras la pirosecuenciación en el apartado de “pipeline inicial process” y el programa se encarga de eliminar las secuencias pertenecientes a los cebadores dentro de los amplicones, de orientarlas todas de la misma manera y también de limpiar las secuencias que tengan alguna base nucleotídica mal leída (que se manifiesta como una “n”). De este modo, se pueden tener todas las secuencias limpias en muy poco tiempo.

En esta página hay diferentes herramientas que permiten trabajar con datos, se pueden alinear todas las secuencias (aligner), o agruparlas en Unidades Taxónomicas Operativas (OTUs) al nivel de significación que se desee (complete linkage clustering). También se pueden calcular índices de diversidad (capítulo 15 material y métodos; Shannon and Chao index) y curvas de rarefacción (rarefaction), observar las diferencias entre dos muestras mediante los índices de Jaccard y Sørensen, y comparar taxonómicamente dos grupos de secuencias mediante la herramienta “Lib compare”.

Del mismo modo, se puede usar la herramienta “classifier” para ver la clasificación taxonómica de todas tus secuencias, observando los porcentajes de los diferentes phyla, géneros u órdenes.

El RDP permite además poder hacer subselecciones dentro de las muestras. Por ejemplo, si se tienen 15000 secuencias de una muestra y se quieren tener aisladas las 200 secuencias que pertenecen al phylum *Cyanobacterias* que se engloban dentro de tu muestra, se puede hacer usando la opción de “fasta sequence selection”. Además, permite generar archivos que se pueden usar en paquetes bioinformáticos un poco más complejos, como el paquete informático R, el MOTHUR o el SPADE.

## 11. CONSTRUCCIÓN DE LIBRERÍAS METAGENÓMICAS EN FAGOS

En este apartado se describe la construcción de librerías génicas en fagos conteniendo fragmentos de ADN medioambiental (figura 7)

### 11.1. Preparación del ADN medioambiental

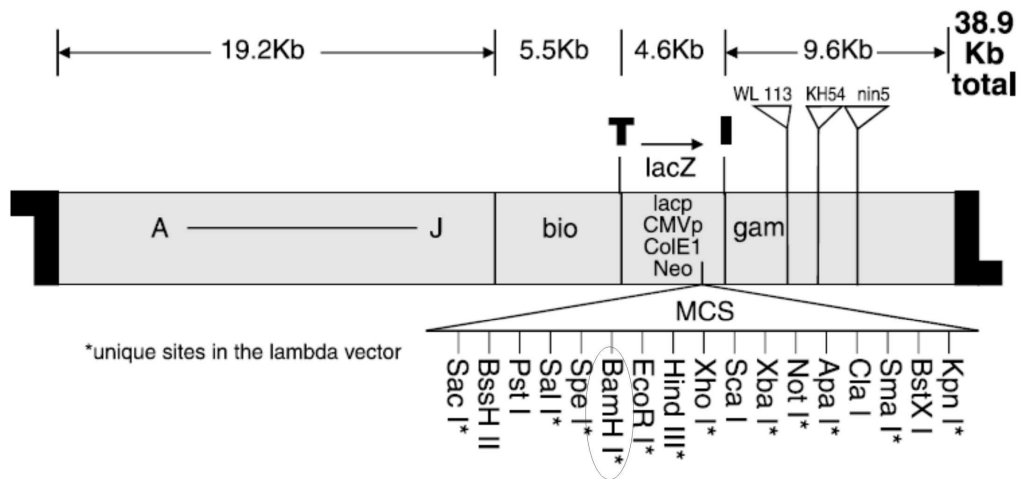
El ADN medioambiental (obtenido según se indica en el capítulo 6.1) se digirió parcialmente con la enzima de restricción *Sau3A*, resolviéndose después en un gel de agarosa de bajo punto de fusión a un voltaje constante toda la noche a 4°C. De este gel se corta la banda que corresponda a un rango de ADN de entre 9-10 kb. Se trata con gelasa según las recomendaciones del proveedor (Epicentre Tech) y se vuelve a concentrar en un sistema de diálisis basado en filtros para microcentrífuga Microcon YM-100 (Millipore).

### 11.2. Construcción de las genotecas

Para la construcción de las genotecas se utilizará el sistema “ZAP Express® Predigested Vector Kit” que proporciona el genoma parcial del fago lambda digerido con *BamH1* (figura X).

En este genoma (de 38,9 Kb en total), encontramos genes necesarios para la propagación del fago, como los genes de la cápsida, de lisis, etc. además de un sitio de clonaje con 12 dianas para enzimas de restricción únicas. Los insertos clonados en este vector pueden ser de hasta 12 Kb y se pueden escindir fuera del fago en forma de un vector fagémido resistente a kanamicina.





**Figura 6.** Mapa del vector ZAP Express<sup>R</sup>. En negrita, en cada extremo, se señalan los sitios cos que permiten la formación de concatémicos (ver más abajo).

El kit comercial te proporciona este vector digerido por la enzima *BamHI* y desfosforilado para evitar la autoligación. Además, como ya hemos comentado en el apartado anterior, también poseemos el ADN de la muestra medioambiental digerido parcialmente con la enzima de restricción *Sau3A*.

La enzima *Sau3A* genera extremos protuberantes que pueden ser ligados con los extremos protuberantes generados por *BamHI* (figura X). De este modo, se liga el ADN medioambiental digerido por *Sau3A* con el vector ZAP Express en el sitio *BamHI*.

Para la ligación, se realiza una mezcla que contiene 0,2 µl de ligasa T4 (20U/ µl), 1 µl de vector Zap Express (1 µg/ µl), 2 µl de ADN medioambiental (a 60 ng/ µl), 0,5 µl de tampón 10x y se completa con agua hasta 5 µl. Esta mezcla se incuba durante toda la noche a 4°C.

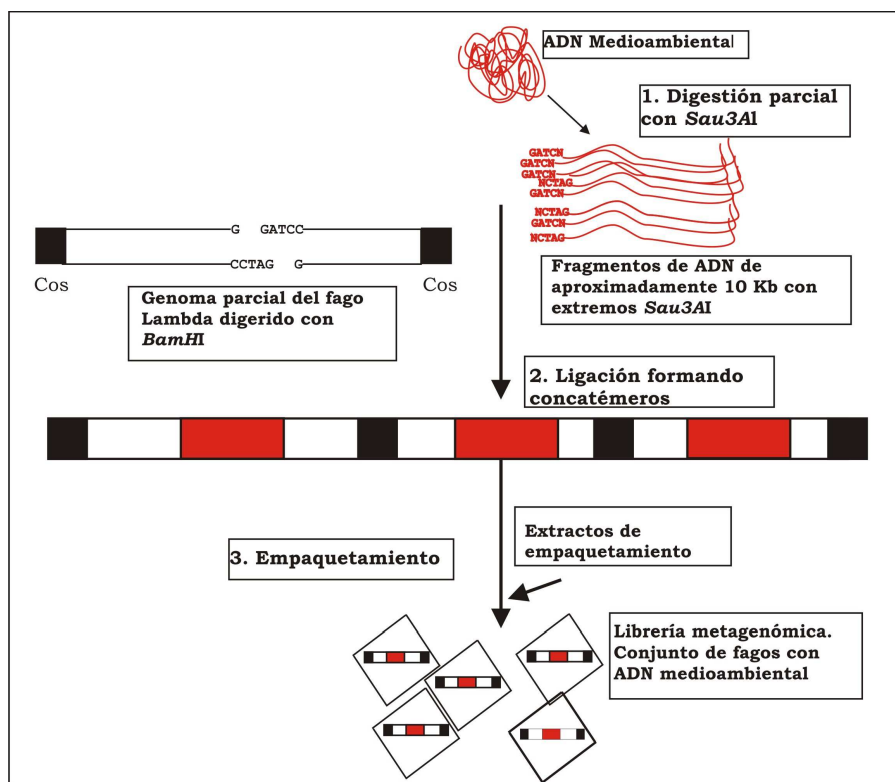
Después de ligar el inserto en el genoma del fago lambda, se empaqueta en partículas víricas utilizando extractos de empaquetamiento. Estos extractos te los proporciona la casa comercial y se deben descongelar muy lentamente. Se les añade 4 µl de la ligación y se dejan a temperatura ambiente durante 2 horas. Posteriormente, en esterilidad, se les añade 500 µl de tampón SM (NaCl, MgSO<sub>4</sub> · 7H<sub>2</sub>O 1 M Tris-HCl pH 7,5 y 5 ml 2% gelatina en un litro) y 20 µl de cloroformo. El fago así empaquetado se puede guardar a cuatro grados durante meses.

Es importante para que el empaquetamiento se produzca de un modo correcto, que el ADN medioambiental se encuentre libre de contaminantes que podrían inhibir este proceso. Por ejemplo, el polietilenglicol que se encuentra en algunos tampones de ligación inhibe el empaquetamiento.

Además, para que la eficiencia del empaquetamiento sea alta, también se necesita que las moléculas de ADN se unan formando concatémoros gracias a las secuencias cos y no se unan formando moléculas circulares (figura 7). Para que esto ocurra así, se recomienda usar durante la ligación concentraciones de ADN altas.

Los vectores que no tengan inserto teóricamente no serán empaquetados, porque se necesita de un tamaño mínimo para el empaquetamiento. De igual forma, los vectores que tengan un tamaño de inserto por encima de las 12 Kb tampoco se empaquetarán porque excederán el tamaño máximo. De este modo, se tiene construida una librería génica con insertos que se estiman entre 0 y 12 Kb.

Una vez construida la genoteca se procede a su titulación para conocer por cuántos fagos está compuesta.



**Figura 7.** Esquema de la construcción de librerías metagenómicas con el sistema Zap Express. Se digiere el ADN medioambiental con la enzima Sau3A1, se liga con el genoma parcial del fago lambda en el sitio *HindIII* y se empaqueta en partículas virales.

### **11.3. Titulación de las genotecas**

La titulación se realiza para conocer cuántas partículas virales posee nuestra librería metagenómica.

Para ello, se inocula un cultivo de células de *E.coli* XL1Blue en medio LB con maltosa al 0,2% (p/v) y 10 mM de MgSO<sub>4</sub> hasta una densidad óptica de 0,5. El MgSO<sub>4</sub> mejora la adhesión del fago a la célula y la maltosa actúa induciendo el operón de la maltosa que sintetiza un receptor de membrana, que es la molécula que usa el fago para penetrar en la célula.

Un volumen de 200 µl de células así preparadas se infectan con diferentes diluciones de los fagos procedentes de las librerías. Se incuban durante 15 minutos a 37°C y se mezclan en un tubo con 6,5 ml de medio top agar NZY precalentado a 45°C (preparado con volúmenes iguales de medio NZY líquido y medio sólido fundido). Esta mezcla se vierte en una placa de 40 ml medio sólido NZY y, tras dejar que gelifique, se incuba a 37°C.

Las calvas de lisis empiezan a ser visibles a las 8 o 9 horas de la incubación. De este modo, contando las distintas calvas de lisis que aparecen en las diferentes diluciones, se puede estimar el número de fagos que posee nuestra librería metagenómica.

### **11.4. Amplificación de las genotecas**

Una vez que la genoteca está titulada y conocemos su número, se procede a su amplificación. De este modo, podremos tener muchos más fagos multiplicando su número, para poder realizar diferentes búsquedas representativas de actividades específicas o de determinados genes de interés.

Al igual que en el caso anterior, se inocula un cultivo de células de *E.coli* XL1Blue en medio LB con maltosa al 0,2% (p/v) y 10 mM de MgSO<sub>4</sub> hasta una densidad óptica de 0,5.

Posteriormente se mezclan 600 µl de las células así preparadas con una alícuota de fagos de la librería metagenómica que contenga  $5 \cdot 10^4$  unidades formadoras de calvas de lisis (pfu “plates formation units”). Esta mezcla se incuba durante 15 minutos a 37 grados y, posteriormente, se vierte en sobre capa, como ya se ha especificado en el apartado anterior, junto a top agar en una placa de medio NZY.

Se incuba la placa a 37°C durante 6-8 horas. Trascurrido este tiempo, empiezan a aparecer las calvas de lisis. Se añaden entonces 9 ml de tampón SM y se incuban a 4°C toda la noche para permitir al fago que difunda desde el medio al tampón.

Al día siguiente se recoge la suspensión de fagos en tampón SM y se le añade un 5% (v/v) de cloroformo para lisar las posibles bacterias infectadas por fagos que no se hayan lisado todavía. Se agita bien y se incuba durante quince minutos a temperatura ambiente. Posteriormente se centrifuga 10 minutos a 500x y el sobrenadante se pasa a un tubo estéril. Si el sobrenadante continua turbio, se puede volver a añadir cloroformo y volver a centrifugar.

A este sobrenadante se le añade un 0,3% (v/v) de cloroformo y se guarda a 4°C. Es recomendable guardar también una alícuota de la librería metagenómica amplificada a -80°C añadiendo un 7% (v/v) de dimetilsulfóxido.

Una vez que ya tenemos la genoteca amplificada, se procede a su titulación para poder conocer el número de unidades formadoras de calvas que tenemos. En este punto se asume que tras este paso podríamos tener de  $10^9$  a  $10^{11}$  pfu/ml.

### **11.5. Escisión de las genotecas**

Como ya hemos comentado anteriormente, el vector Zap Express permite la escisión de los fragmentos clonados a un fagémido que se puede seleccionar en células de *E.coli* gracias a que presenta resistencia a kanamicina.

Este fagémido se denomina pBK-CMV (figura 4) y permite poder tener los fragmentos clonados en células de *E.coli* en un número elevado, así como poder sobreexpresarlos mediante la adicción de IPTG o comprobar los clones con inserto mediante una selección de colonias blancas y azules.

#### **11.5.1. Escisión de una sola unidad formadora de calva**

Para la escisión de una unidad formadora de calva individual se procede de la siguiente manera. Se pica en la calva de interés con una punta estéril y se resuspende en un tubo con 500 µl de tampón SM y 20 µl de cloroformo para lisar las células. Se agita y se incuba durante una o dos horas

a temperatura ambiente o a 4°C durante toda la noche (el fago, de esta forma, es estable durante más de 6 meses).

Se inoculan dos cultivos independientes de células XL0LR y XL1Blue de *E. coli* y con 0,2% de maltosa y 10 mM de MgSO<sub>4</sub> hasta una densidad óptica de 1 (que correspondería a 8 10<sup>8</sup> células/ml).

Posteriormente, se mezclan 100 µl de células XL1Blue de densidad óptica 1, con 125 µl de la mezcla del fago de interés en SM (con cuidado de no tomar cloroformo) y 1 µl del fago helperEx assist. Se incuba esta mezcla durante 15 minutos a 37°C para que el fago ataque a las células.

A continuación, se le añade a esta mezcla 1,5 ml de medio LB con maltosa al 0,2% y 10 mM de MgSO<sub>4</sub>, y se incuba a 37 grados durante 2,5 o 3 horas. Como la representación clonal no es relevante, la reacción de escisión puede dejarse también toda la noche.

Transcurrido este tiempo, se calienta la mezcla a 65-70°C durante 20 minutos para lisar las partículas del fago lambda y las células. Después, se centrifuga a 1000g durante 15 minutos para lograr la sedimentación de las células.

El sobrenadante resultante de esta centrifugación contendrá el fagémido pBK-CMV escindido, empaquetado como una partícula de fago filamentoso y podrá ser guardado de esta forma durante 1-2 meses a 4°C.

En el siguiente paso, se infectan células XL0LR a densidad óptica 1 (crecidas con maltosa y sulfato magnésico) con la mezcla que porta el fagémido escindido en forma de fago filamentoso. De este modo, se mezclan 200 µl de células XL0LR con 100 µl del sobrenadante que contiene el fagémido. Se incuba la mezcla a 37°C durante 15 minutos para que el fago infecte a la célula y, posteriormente, se le añade medio líquido NZY, 300 µl, y se incuba durante 1 hora a 37 grados para permitir la correcta expresión del gen de resistencia a kanamicina que porta el fagémido, antes de la siembra en el medio con el antibiótico.

Finalmente, se plaquean 200 µl de esta mezcla en placas de medio LB con kanamicina (50 mg/L) y se incuban a 37°C toda la noche.

Las colonias que aparecerán en la placa poseerán el vector fagémido pBK-CMV procedente de la escisión del vector Zap Express.

El fago helper junto con la cepa XL0LR están diseñados para la correcta escisión del vector fagémido pBK-CMV del vector ZAP Express. El fago helper tiene una mutación ambar que evita su replicación cuando su hospedador

corresponde con cepas de *E.coli* no supresivas, como las XLOR. Estas cepas simplemente permiten la replicación del fagémido escindido, eliminando de este modo la posibilidad de una coinfección con el fago helper. Además, las células XLOR son resistentes a una infección con el fago lambda, evitando de esta forma la posible contaminación del ADN después de la escisión.

### **11.5.2. Escisión en masa**

En algunos casos, en lugar de la escisión de una calva de interés, lo que interesa es la escisión en masa de toda la librería metagenómica para tenerla en forma de fagémidos dentro de las células XLOR de *E.coli* y no en forma de fagos.

Para ello, y al igual que en el método anterior, se inoculan cultivos independientes de células XLOR y XL1Blue de *E. coli* con sulfato magnésico y maltosa a las concentraciones que ya se han indicado anteriormente hasta una densidad óptica de 1.

Posteriormente, se combinan en un tubo cónico de 50 ml una proporción de la librería metagenómica amplificada en el fago lambda con las células XL1Blue y con el fago helper en proporción 1:10:100.

Se recomienda que la cantidad de fago lambda de la genoteca metagenómica que se use, sea 10 o 100 veces mayor que el de la genoteca metagenómica original sin amplificar para asegurar la representatividad estadística en los clones escendidos. Del mismo modo, se añade 100 veces cantidad de fago helper que de células para asegurar que todas las células son coinfectadas con el fago lambda y con el fago helper.

Por ejemplo, si el título de nuestra librería metagenómica fuera de  $10^5$  unidades formadoras de calvas, se recomendaría preparar una mezcla con  $10^7$  pfu procedentes de la librería metagenómica (100 veces la original),  $10^8$  células XL1Blue y  $10^9$  pfu de fago helper.

Se incuba la mezcla así preparada a 37°C durante 15 minutos para permitir al fago atacar a las células y, posteriormente, se le añaden 50 ml de medio LB con maltosa y sulfato magnésico. A continuación, se deja incubando en un tubo cónico durante 3 horas (también se puede incubar durante toda la noche) a 37°C en agitación.

Pasado este tiempo, se calienta la muestra para lisar las células y las partículas de fago lambda y se centrifugan a 1000x g durante 10 minutos para decantar las células.

El sobrenadante contendrá un conjunto de vectores fagémidos escindidos procedentes de los fagos lambda de la genoteca inicial. Para titularlos, se combina 1  $\mu$ l de este sobrenadante con 200  $\mu$ l de células XL0LR de densidad óptica 1. Se incuba la mezcla a 37°C durante 15 minutos y después se le añaden 40  $\mu$ l de medio NZY concentrado 5 veces (para que la concentración final sea 1X).

Posteriormente, se deja a 37°C durante una hora para permitir que se exprese la resistencia a kanamicina y se plaquea en placas LB Km (50 mg/L).

De este modo, podremos pasar de tener una librería metagenómica en fagos a poseer vectores fagémidos de alto número de copias en células de *E.coli*, sobre las que se podrán realizar diferentes tipos de búsquedas (por ejemplo, búsqueda de genes de resistencia a antibióticos).

## **12. PREPARACIÓN Y ANÁLISIS DE PROTEÍNAS**

En los siguientes apartados trataremos todos los aspectos relacionados con la manipulación, análisis y preparación de muestras de proteínas.

### **12.1 Determinación del contenido proteico**

La concentración proteica de las diferentes muestras ensayadas en el estudio se determinó siguiendo el método colorimétrico de Bradford (1976), el cual se basa en la observación del cambio del máximo de absorbancia, de 465 a 595nm, de una solución ácida del azul Coomassie R-250, al unirse a proteínas. Se cuantifica la unión del azul Coomassie a la proteína problema, y se compara esta unión con la de cantidades diferentes de una proteína de referencia, que en este caso fue albúmina sérica bovina, a dicho colorante.

Las muestras se prepararon en un volumen total de 800  $\mu$ l, completando con agua destilada. Se añadió a cada muestra 200  $\mu$ l de la solución colorante, solución de Azul Coomassie R-250 en ácido fosfórico y metanol (BioRad) y se agitó.

Posteriormente se realizó la medida a 595 nm en un espectrofotómetro.

Los resultados obtenidos se extrapolaron en una curva patrón utilizando como estándar albúmina sérica bovina (50  $\mu$ g/ml).

Ocasionalmente las proteínas (en fracciones superiores a 50-100ng/ $\mu$ l) fueron medidas en espectrofotómetro Nano Drop ND-1000 que realiza las determinaciones a una  $\lambda$  de 280 nm.

## **12.2. Electroforesis de proteínas en gel en condiciones desnaturalizantes (SDS-PAGE)**

El análisis de las proteínas se llevó a cabo según el método descrito por Laemmli (1970), mediante una electroforesis vertical en geles de poliacrilamida usando el sistema Mini Proten II (BioRad) y siguiendo las recomendaciones de la casa comercial.

La concentración de acrilamida/bisacrilamida para el gel separador fue del 12% y para el gel concentrador del 5 %. Los geles polimerizan tras la adición de un 0.01 % de persulfato amónico (p/v) y 50 % de TEMED (v/v). El resto de componentes de los geles y sus proporciones se observan en la tabla 7.

Como tampón de carga se usó una solución dos veces concentrada de Tris-Hcl pH6.8 50 mM, 2 % SDS, 10 % glicerol, azul de bromofenol al 0.006 % (p/v) y 10 mM de DTT. En el caso de electroforesis en condiciones no reductoras el DTT se suprimió del tampón de carga. Las muestras preparadas con el tampón de carga se calientan a 95°C durante dos minutos para desnaturalizar las proteínas.

Para el marcador de peso molecular, se usó un marcador estándar de bajo rango, sin teñir suministrado por BioRad, en el que se observa las bandas correspondientes a las proteínas: fosforilasa b de 97.4 KDa, Seroalbúmina bovina de 66.2 KDa, Ovoalbúmina de 45 KDa, anhidrasa carbónica de 31 KDa, Tripsinógeno de 21.5 kDa y lisozima de 14.1 KDa.

Los geles se corren usando como tampón de electroforesis 0.25 M Tris pH 8.3, 1.9 de Glicina, 1 % SDS, durante 45 minutos a un voltaje constante de 200 V.



	<b>Gel separador</b>	<i>Gel concentrador</i>
<b>40 % Poliacrilamida</b>	3 ml	248 $\mu$ l
<b>15 mM Tris-Hcl pH 6.8</b>	/	250 $\mu$ l
<b>15 mM Tris-Hcl pH 8.8</b>	2.5 ml	/
<b>SDS 10 %</b>	100 $\mu$ l	20 $\mu$ l
<b>Agua</b>	4.3 ml	1.485

**Tabla 7.** Composición de los geles de poliacrimida usados en este estudio.

Las proteínas después de la electroforesis se visualizaron por tinción con azul coomassie o por tinción de plata según la cantidad de proteína que hubiera o la calidad de pureza que se quisiera observar.

### **12.2.1. Tinción de plata**

Se hizo según una modificación del procedimiento descrito por Blum *et al.* (1987), que nos permite detectar concentraciones de proteína de 10-100 ng. El protocolo seguido fue el siguiente:

En primer lugar se somete el gel al proceso de fijación: se sumerge durante 30 minutos (al menos aunque se puede dejar toda la noche) en 100 ml de una solución de etanol absoluto 40 % (v/v), ácido acético 10% (v/v) y agua al 50 % (v/v) con agitación.

Posteriormente el gel se incuba en una solución de sensibilización durante 30 minutos exactos, compuesta por etanol absoluto al 30 % (v/v), tiosulfato sódico al 0.2 % (p/v), acetato sódico al 6.8 % y agua.

Después del proceso de sensibilización, se realizan 3 lavados de cinco minutos cada uno, con agua destilada.

El siguiente paso es el de la tinción con plata. Se prepara una solución de nitrato de plata al 0.25 % (p/v), y el gel se incuba en esta solución durante 20 minutos. Es necesario proteger la tinción de plata de la luz, por lo que el recipiente que contiene el gel se suele cubrir con un material opaco durante esta fase.

Antes de proceder al revelado, el gel se debe lavar con agua durante 1 minuto, dos veces.

Posteriormente se le añade una solución de revelado que posee carbonato sódico al 2.5 % y 0.004 % (v/v) de formaldehído al 37 %. El gel se incuba en esta solución hasta cuando aparecen las bandas, en este momento

la solución se cambia por la solución de parada compuesta por EDTA sódico al 1.46 % (p/v). Se incuba en esta solución durante 10 minutos y posteriormente se lava durante tres veces cinco minutos con agua destilada.

### **12.2.2. Tinción con azul Coomassie**

Con esta técnica se detectan concentraciones de proteína del orden de 0.1-1 µg.

El gel se sumerge en una solución de tinción compuesta por ácido acético glacial al 9.2% (v/v), metanol 45% (v/v) y Azul Coomassie R-250 al 0.25% (p/v) durante 15-30 minutos con agitación. La solución de tinción se eliminó entonces y se sustituyó por la decolorante que contenía ácido acético glacial al 15% (v/v) y metanol al 5% (v/v). Se dejó agitando de nuevo y se hicieron frecuentes cambios de solución para eliminar el colorante extraído (Garfin, 1990). Esto se hizo durante el tiempo necesario para que el fondo quedase claro y las bandas de proteínas patentes.

### **12.2.3. Zimogramas para la actividad esterasa**

Una parte de los resultados expuestos en esta tesis están relacionados con enzimas con actividad esterasa. Para el trabajo rutinario con este tipo de enzimas, sobre todo en el caso de extractos proteicos no purificados, resulta muy útil poseer un ensayo rápido usando sustratos cromogénicos o fluorescentes.

En este apartado se describe una técnica para revelar la actividad esterasa en geles SDS-PAGE (Díaz P, et al, 1999), de forma que podamos conocer con precisión si la enzima causante de la actividad se encuentra en el extracto proteico por un lado y por otro su peso molecular.

Para ello, se realiza una electroforesis de proteínas SDS-PAGE, tal y como se describe en el apartado 12.2, con la salvedad de que la muestra de proteínas con el tampón de carga se somete a diferentes tiempos de hervido, cargando en el gel muestras desde sin hervir, hasta hervidas 10 minutos, porque este proceso afectará a la renaturalización de las proteínas posterior.

Después de que la electroforesis haya terminado y por tanto se haya producido la separación de las proteínas, se elimina el SDS del gel incubándolo en agitación durante 30 minutos a temperatura ambiente con una solución de Tritón-X-100 al 2.5 %. Posteriormente, el gel se lava durante 2 minutos con 50 mM de tampón fosfato pH7.

Una vez que el gel está lavado, se le añade el sustrato que va a permitir observar las bandas con actividad esterasa; se trata del sustrato 4-metilumberiferil acetato (4-MUF-acetato; sigma), se añade una solución 100  $\mu\text{M}$  de este sustrato en tampón tris HCl pH7 y se incuba con el gel durante un minuto. Pasado ese tiempo las bandas con actividad se observan iluminando con luz ultravioleta.

Al contrario que otros sistemas basados en el revelado de actividad esterasa en geles SDS PAGE (Yadav et al, 1998), la sensibilidad de este protocolo es muy alta, permitiendo la detección de  $1.5 \cdot 10^{-7}$  U de lipasa pancreática (Díaz P, et al, 1999).

Una vez que se ha revelado la actividad esterasa, el gel se puede teñir normalmente, bien cono Coomassie o bien con la tinción de plata. La yuxtaposición de las imágenes resultantes de la tinción con 4-MUF-Acetato y de la tinción con Coomassie o con plata nos permite identificar la banda que correspondería con la enzima de carácter lipolítico.

Una variante de este tipo de ensayos, muy útil para testar la actividad esterasa en extractos proteicos durante la purificación (por ejemplo para testarla en diferentes fracciones cromatográficas) consiste en añadir en un papel de filtro una gota de 5  $\mu\text{l}$  de 4-MUF-Acetato 200  $\mu\text{M}$  y sobre ésta, 5  $\mu\text{l}$  o 10  $\mu\text{l}$  del extracto proteico (dependiendo de su concentración). Al iluminar la gota en el ultravioleta, podremos observar si el extracto en cuestión posee actividad lipolítica o no.

#### **12.2.4 Secado de los geles de poliacrilamida**

Tras la tinción de los geles, éstos se secaron entre dos papeles de celofán, para su conservación. A tal fin, se lavaron con agua y se sumergieron en una solución de glicerol al 3% (v/v) y metanol 30% (v/v), durante 30 minutos, con agitación constante.

Al cabo de ese tiempo se colocaron sobre un papel de celofán, empapado en ácido acético al 3% (v/v), que a su vez iba sobre una bastidor y un marco de metacrilato. Se colocó encima de los geles un segundo celofán, empapado de igual modo, con la precaución de no dejar ninguna burbuja de aire. Por último se colocó el segundo marco de metacrilato, y se sujetó un marco al otro con pinzas metálicas. Se dejó secar todo el sistema al aire.

### **13. EXPRESIÓN Y PURIFICACIÓN DE PROTEÍNAS**

La metodología seguida para la expresión y purificación de proteínas depende de la proteína y del vector en el que se esté expresando. En un primer momento, se procedió a la purificación de las proteínas directamente desde el vector pBCK-CMV (figura 4), ya que este vector permite su inducción con IPTG. Con este método se purificaron las proteínas pE1, pE2 y pE4.

Para otras proteínas cuya expresión desde este vector no resultaba tan buena, se clonaron en vectores específicos (figura X) de expresión y purificación de proteínas como son los vectores pET3a y el pET16b (figura X) que además posee una cola histidinas que puede ser usada para purificar la proteína mediante cromatografía de afinidad. Un ejemplo de las proteínas que se purificaron siguiendo este procedimiento son la proteína Tm8.4 y la proteína pE12; por su parte también se expresaron aunque no purificaron las proteínas quiméricas pE1Q y pE2Q.

#### **13.1. Expresión y purificación de proteínas en *E. coli* desde el vector pBCK-CMV**

##### **13.1.1 Expresión en *E. coli* y preparación de las muestras de proteínas**

Para ello, se parte de 500 ml de cultivo de la cepa XL0LR de *E. coli* que contiene el plásmido pBK-CMV con el gen esterasa expresado desde el promotor lac, el cultivo se incubaba a 37°C en medio LB con 50µg/ml de kanamicina.

Cuando la densidad óptica a 600 nm varía entre 0,5-0,6 se le añade IPTG hasta una concentración final de 1 mM para inducir la expresión.

Las células se recogen 3 horas después de haber inducido y se lavan con 25 ml de tampón Tris-HCl pH 7, 50 mM (tampón A), guardándose el pellet a -80°C.

El botón de células se resuspende en 5 ml de tampón A con inhibidor de proteasas y 0,5 mg/ml de DNasa, después se procede a su lisis por sonicación (4-5 pulsos de 30 segundos, seguidos de pausa de 1 minuto; Sonicador Bransson) a 4°C. La fracción soluble se separa de la insoluble por centrifugación a 10.000 g durante 30 minutos a 4°C, se dializa toda la noche frente a 2 L de tampón A (3 cambios) y se concentra mediante ultrafiltración

en una membrana Centricon YM-10 (Amicon, Millipore) hasta un volumen final de 5 ml.

Antes de proceder a la cromatografía para el proceso de purificación, se le añaden, a los 5 ml de muestra de proteínas, sulfato amónico hasta una concentración de 1M, se centrifuga a 4°C, 5000 rpm, 15 minutos. Después de recoger el sobrenadante se corrobora que éste mantiene la actividad esterasa. A continuación, se filtra la muestra con filtros de 0,45 µm.

### **13.1.2. Proceso de purificación**

El proceso de purificación de las proteínas recombinantes se llevó a cabo utilizando un sistema FPLC de Amersham Biosciences en el que se integraron distintas columnas cromatográficas, el equipo se encuentra provisto de un detector óptico a 280 nm que permite la visualización y análisis del cromatograma. Se controló el proceso de purificación por análisis electroforético y medida de actividad enzimática de las fracciones obtenidas de las columnas cromatográficas después de cada paso. Se empleó el método de Bradford para estimar el contenido proteico de cada etapa de la purificación.

De este modo, el primer paso cromatográfico para la purificación está basado en una cromatografía de hidrofobicidad. Se carga la muestra en una columna de Phenyl-Sefarosa HR 10/10 de 5 ml (Pharmacia biotech), previamente equilibrada con tampón B (50 mM Tris-HCl, pH 7, 1 M  $(\text{NH}_4)_2\text{SO}_4$ ). Las proteínas eluyen siguiendo un gradiente lineal decreciente (40 ml) de  $(\text{NH}_4)_2\text{SO}_4$ , desde 1 M hasta 0. El flujo fue de 0,15 ml/min y se toman fracciones de 1 ml. Siguiendo esta técnica cromatográfica, las proteínas eluyen en función de sus propiedades hidrofóbicas, primero eluirán las menos hidrófobas y posteriormente las más hidrófobas.

Una vez recogidas todas las fracciones, se hace un ensayo para determinar la actividad esterasa y se cargan las muestras en geles de proteínas SDS-PAGE al 12%, viendo de este modo el perfil de bandas y la identificación de la posible proteína en función de su peso molecular.

De este modo, se unen todas las fracciones que tenían actividad positiva y se pasa esta muestra por una columna PD-10 (Amersham) para eliminar las sales, después se concentra con una columna Centricon YM-10 (Amicon, Millipore) hasta 1,5 ml.

La muestra concentrada se filtra (0,20 µm) y se purifica con otro tipo de columna, Mono-Q HR 5/5 de 1 ml. Las proteínas se eluyen siguiendo un

gradiente lineal de concentración (25 ml) de NaCl (de cero hasta 1 M), con un flujo de 0,5 ml/min y se recogen fracciones de 1 ml. Se testa la actividad esterasa en cada fracción y luego se corroboran los datos con geles SDS-PAGE, viendo en algunos casos la banda correspondiente a nuestra proteína purificada y en otros, se procedió a hacer una purificación más usando como técnica de cromatografía la filtración en gel.

### **13.2. Expresión y purificación de proteínas en *E. coli* desde el sistema pET**

#### **13.2.1. Expresión en *E. coli* y preparación de las muestras de proteínas**

El sistema de expresión pET está basado en la transcripción de genes por parte de la enzima T7-RNA polimerasa (Studier and Moffatt, 1986). La secuencia correspondiente al gen que codifica la proteína de interés se clona a continuación de la secuencia promotora de la enzima T7-RNA polimerasa y la construcción se transforma en cepas de *E. coli* que contienen en su genoma una copia del gen que codifica esta enzima. La transcripción del gen de la T7-RNA polimerasa en estas células se encuentra reprimida por medio de la secuencia reguladora lac, pero es inducible por IPTG. La adición de IPTG permite la transcripción y posterior expresión de la enzima T7-RNA polimerasa, la cual dispara a su vez la transcripción del gen de interés. Se alcanzan de esta forma unos elevados niveles de expresión de proteína recombinante.

Con cada una de las construcciones pET3a/pET16b con los insertes de interés, se transforman células de *E. coli* frescas de la cepa BL21 (DE3), las cuales contienen el gen que codifica la T7-RNA polimerasa. Como ventaja adicional, estas células están modificadas en genes que codifican proteasas y peptidasas de *E. coli*, de forma que permiten la sobreexpresión de proteínas recombinantes.

En la expresión se parte de un cultivo de 100 ml de células de *E. coli* BL21 (DE3) transformadas con el plásmido de interés que se incubaba en agitación (190 rpm) a 37°C hasta una densidad óptica a 600 nm de 0.5. En este momento se le añade el IPTG para inducir la expresión proteica, añadimos IPTG hasta una concentración final de 1 mM en el caso de los vectores pET16b y hasta una concentración final de 0.4 mM en el caso de los

vectores IPTG. La expresión de las proteínas recombinantes también se estudió sin la adición de IPTG. Los cultivos así inducidos se dejan toda la noche a 20°C, que fue la temperatura que observamos como óptima para la expresión de las proteínas.

Posteriormente, se recogen las células y tras varios lavados con tampón A, se resuspenden en un volumen de tampón de lisis dependiente de la densidad óptica de los cultivos, normalmente el volumen suele ser de 2 ml. El tampón de lisis contiene 50 mM de tris-HCl pH 7.5, 1 mM EDTA pH 7.5, 5 % de glicerol, 1 mM DTT y 1 M NaCl.

Tras resuspender con el tampón de lisis, las muestras se lisan por sonicación (pulsos de 30 segundos, seguidos de pausa de 1 minuto) a 4°C. La fracción soluble se separa de la insoluble por centrifugación a 12.000 g durante 30 minutos a 4°C, y en el caso en el que sea necesario se concentra hasta 2 ml usando la columna Centricon YM-10.

### **13.2.2. Proceso de purificación**

Al igual que en el punto 13.1, el proceso de purificación de las proteínas recombinantes se llevó a cabo utilizando un sistema FPLC de Amersham Biosciences en el que se integraron distintas columnas cromatográficas.

En el caso de las proteínas clonadas en el vector sin colas de histidinas, el pET3a, la elevada expresión de proteína recombinante, nos permite tener que realizar solo dos pasos cromatográficos para su purificación.

La muestra concentrada del paso anterior, se filtra (0,20 µm) y pasa por una columna de intercambio iónico Mono-Q HR 5/5 de 1 ml. Las proteínas eluyen siguiendo un gradiente lineal de concentración (25 ml) de NaCl (de cero hasta 1 M), con un flujo de 0,5 ml/min y se recogen fracciones de 1 ml.

Al igual que en el caso anterior, se testa la actividad en cada fracción y luego se corroboran los datos con geles SDS-PAGE. Después de unir las fracciones en que las eluye la proteína de interés, se somete a un sistema de diálisis para eliminar las sales y se concentra hasta 1 ml.

Con la muestra así preparada se somete a otro paso cromatográfico basado en la separación de las moléculas en función de su peso molecular, denominado filtración en gel. Tras este último paso cromatográfico, se corrobora la actividad y la pureza de la muestra mediante geles de SDS-PAGE.

Para la proteína Tm8.4p aislada mediante esta metodología, finalmente se conservó en un tampón con 10 mM de Tris-HCL pH8, 0.5 mM DTT, 5 $\mu$ M NADPH, inhibidor de proteasas (PMSF 1 mM) y 50 % de glicerol

En el caso de las proteínas recombinantes expresadas desde el vector pET16b, éstas poseen colas de histidina bien en su extremo N-terminal o bien en el C-terminal que facilitan su purificación mediante un solo paso cromatográfico en una columna de afinidad (HisTRap FF 1 ml, BioRad) equilibrada previamente con níquel.

Las histidinas tienen afinidad con el níquel, de forma que la proteína recombinante con la cola de histidina se pegará a la columna. Posteriormente, la columna se somete a un gradiente de imidazol desde cero hasta 0.5 M (25 ml) que compite por el níquel con las histidinas, eluyendo así la proteína recombinante en fracciones de 0.5 ml, en una concentración de imidazol concreta. Posteriormente, se testa la actividad y se corrobora mediante geles de SDS-PAGE que efectivamente la muestra proteica está pura.

### **13.3. Cromatografía analítica de filtración en gel**

En todos los procesos de purificación se procedió a una cromatografía analítica mediante filtración en gel, para calcular el peso molecular y el radio de las proteínas, además de para poder corroborar si se encuentran en forma monoméricas o multiméricas.

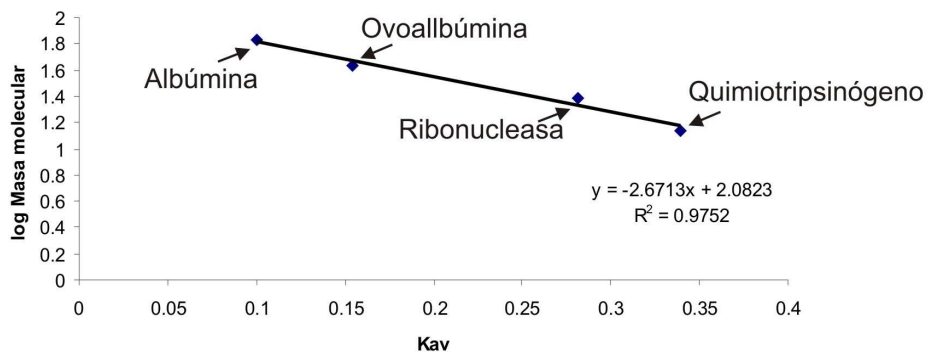
Para ello, se usó la columna cromatográfica una de Superdex 75 HR integradas en el sistema FPLC de Amersham Biosciences. La calibración de la columna se realizó utilizando proteínas marcadoras estándar (Amersham Biosciences) de masa molecular y radio de Stokes conocidos: En el caso de la columna Superdex HR-75 se usaron los marcadores ribonucleasa (13.7 KDa, 16.4 Å), quimotripsinógeno (25 KDa, 20.9 Å), ovoalbúmina (43 KDa, 30.5 Å), y albúmina de suero bovina (BSA) (67 KDa, 35.5 Å).

Para cada proteína,  $K_{av}$  (la fracción del volumen de gel estacionario que está disponible para la difusión de un soluto dado, este valor permite definir la conducta cromatográfica del soluto independientemente de la longitud del lecho y del grado de empaquetamiento) se calculó a partir de la ecuación 1.

$$K_{av} = (V_e - V_o)/(V_t - V_o) \quad (\text{Eq. 1})$$



Donde  $V_e$  es el volumen de elución para cada proteína,  $V_t$  es el volumen total de la columna (24.4 ml) y  $V_o$  el volumen de exclusión de la columna (7.98 ml) para que se determinó inyectando azul de dextrano. La representación del log de la masa molecular frente a  $K_{av}$  de las proteínas marcadoras (figuras x) corresponde a una regresión lineal que permite determinar la masa molecular de una proteína conociendo su  $K_{av}$ .



**Figura 8.** Calibración de la columna Superdex HR-75 mediante el uso de proteínas marcadoras.

## 14. DETERMINACIÓN DE ACTIVIDADES ENZIMÁTICAS

### 14.1. Ensayos enzimáticos relacionados con actividades lipolíticas

#### 14.1.1. Búsqueda de genes con actividad lipolítica en las genotecas metagenómicas

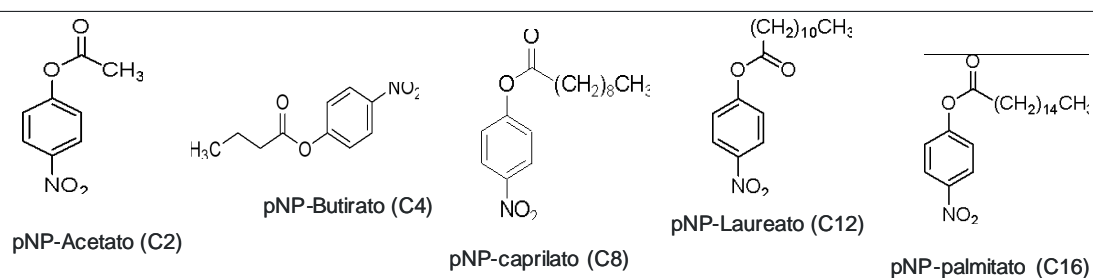
La búsqueda de enzimas de carácter lipolítico en las genotecas metagenómicas construidas se realizó directamente sobre las calvas de lisis producidas por los fagos en un primer momento, aunque también se pueden testar colonias bacterianas. De este modo, se realizó un ensayo en placa donde se puso de manifiesto la actividad esterasa mediante la adición naftil-acetato y fast-blue (SIGMA). En este ensayo se prepara una solución de naftil-acetato de 0,2 mg/ml en acetato y otra de fase-blue a 0,8 mg/ml en dimetilsulfóxido, estos dos componentes se mezclan en agarosa (800  $\mu$ l de sustratos por cada 45 ml de agarosa) al 0,5% y se vierte en sobrecapa sobre la colonia o colonias cuya actividad se quiere testar. Al hidrolizarse el naftil-acetato reacciona con el fast-blue dando lugar a un precipitado de color marrón que indica la colonia o la calva (en el caso de que sea un fago), donde se está produciendo actividad esterasa.

### 14.1.2. Actividades enzimáticas

Las actividades enzimáticas de las enzimas esterasa se determinaron usando ésteres de para-nitrofenil (pNps; figura X; Sigma-Aldrich). De este modo, se midió la variación de absorbancia a 405 nm resultante de la liberación de p-nitrofeniles por la enzima esterasa. Si no se especifica, la actividad enzimática se determina usando un ensayo estándar con 0,8 mM de p-nitrofenil-acetato en tampón 50mM Tris-HCL pH7, 15% acetonitrilo y 0,038 mM de Triton-X-100 durante 15 minutos. La reacción se detiene añadiendo 0,023 M de carbonato sódico. Todos los ensayos si hicieron por triplicado y corrigiendo con la muestra en blanco para evitar así el error proveniente de la autohidrólisis del sustrato. Una unidad de enzima se define como la actividad enzimática capaz de liberar, a partir de un éster de para-nitrofenil, 1  $\mu$ mol de p-nitrofenil por minuto.

#### 14.1.2.1 Ensayos enzimáticos de especificidad de sustrato

Para hallar la especificidad de sustrato de las distintas enzimas encontradas, se usaron los ésteres p-nitrofenil, en concreto se usó el p-nitrofenil-acetato (C2), el p-nitrofenil-butirato (C4), el p-nitrofenil-caprilato (C8), el p-nitrofenil-laureato (C12) y el p-nitrofenil-palmitato (C16) (ver tabla 3). Los p-nitrofeniles se disolvieron en isopropanol hasta una concentración final de 16 mM, en las reacciones para determinar la especificidad de sustrato, la concentración final del mismo fue de 8 mM. En la determinación de la ecuación de Michaelis-Menten se monitorizó la reacción en función de distintas concentraciones de sustrato.



**Figura 9.** Derivados p-nitrofenilos (pNP) utilizados en los ensayos colorimétricos para la determinación de la actividad esterasa.

Con la enzima pE12 purificada también se evaluó la especificidad de sustrato frente a otras sustancias de interés farmacéutico o industrial como son: 2-(4-isobutylphenyl) propanoate, 2-phenyl propanoate, 3-phenyl propanoate, cyclohexanoate, 2-(3-benzoylphenyl) propanoate, 2-naphtoate, 1-

napthoate, adamantanoate, 2-(6-methoxynaphtalen-2-propanoate). La enzima se incubó con 475 µl de tampón fosfato pH7 y con 25 µl (5 mg/ml) de cada sustrato durante 15 minutos.

*14.1.2.2. Ensayos enzimáticos para la determinación del pH óptimo, temperatura óptima y termoestabilidad de las enzimas esterasa*

Para la determinación del pH óptimo, se midió la actividad esterasa en intervalos de pH de entre 6,8-12,6, no se usaron valores de pH más bajos porque el ambiente ácido influye en la autohidrólisis del sustrato, produciéndose así una hidrólisis masiva de sustrato sin intervención de la enzima. El tampón usado fue Tris-Hcl 50 mM a distintos pH (6,8; 7; 7,6; 8,3; 9,5; 12,6) y se monitorizó la liberación de p-nitrofenil acetato a 405 nm, a la temperatura óptima de cada enzima.

La temperatura óptima se determinó usando un rango de temperaturas de 0-42°C bajo las condiciones estándar del ensayo enzimático ya descritas. Se ensayó la termoestabilidad midiendo la actividad residual después de incubar las enzimas durante intervalos determinados de tiempo a distintas temperaturas.

*14.1.2.3. Efecto de detergentes no iónicos, de solventes orgánicos y de los iones sobre la actividad enzimática*

El posible efecto inhibitorio o inductor de este tipo de sustancias se halló en el caso de los detergentes no iónicos Triton-X-100 y SDS 10%, midiendo la actividad enzimática sobre p-nitrofenil-acetato en un tampón con 10 % de volumen de cada uno de los detergentes.

En el caso de los solventes orgánicos se determinó la actividad residual de las enzimas después de su incubación durante 6 horas a temperatura ambiente en concentraciones de acetonitrilo y dimetilsulfóxido (DMSO) que variaban entre el 0 y el 98%. Para el cálculo de la actividad residual se siguieron las condiciones de los ensayos estándar de actividad enzimática.

Con respecto al efecto de los iones sobre la actividad lipolítica, las enzimas se incubaron en distintas concentraciones de iones como calcio o Manganeso. Los ensayos con extractos crudos no pueden ser evaluables y solo se realizaron este tipo de ensayos con las proteínas puras. Tras el tiempo de incubación, se centrifugaban durante 2 minutos a 4°C a 1000 x g para

eliminar los posibles precipitados que pudieran interferir en los valores de absorbancia.

#### **14.2. Ensayos enzimáticos relacionados con actividad dihidrofolato reductasa**

Para una de las enzimas aisladas en esta Tesis Doctoral, la proteína Tm8.4p, se realizaron diferentes ensayos con el fin de poder corroborar si poseía actividad dihidrofolato reductasa.

Los ensayos para la actividad dihidrofolato reductasasa se realizaron en un tampón con 100 mM de ClNa, 25 mM de Tris-HCL pH7, 25 mM de MES pH7, 75 mM de  $\beta$ -mercaptoetanol y 1 mg/ml de seroalbúmina bovina. A este tampón se le añadía 100  $\mu$ M de NADPH (Sigma) y la enzima Tm8.4p, la reacción se dispara tras la adicción de 100 100  $\mu$ M de dihidrofolato (SIGMA). Se mide de esta forma la variación de absorbancia a 340 nm debida al cambio del NADPH a NADP. El coeficiente de extinción molar del NADPH es de 6.22 mM<sup>-1</sup>. cm<sup>-1</sup>. Todos los ensayos si hicieron por triplicado y corrigiendo con la muestra en blanco. Una unidad de enzima se define como la actividad enzimática capaz de transformar a 1  $\mu$ mol de NADPH por minuto a una temperatura dada (generalmente 20-25 °C).

#### **15. SISTEMA DE MUTAGÉNESIS DIRIGIDA**

El sistema que se usó para la construcción de las quimeras pE1Q y pE2Q fué la mutagénesis puntual mediante el kit comercial Altered Sites® II in vitro Mutagenesis Systems (Promega).

El fundamento del protocolo es la introducción de mutaciones mediante hibridación con oligonucleótidos que contienen cambios respecto a la secuencia original. Deberemos tener clonada nuestra secuencia a mutar en el vector pALTER-1® (Promega). Posteriormente, se anillan tres oligonucleótidos a la vez: uno de ellos fue nuestro oligonucleótido mutagénico y los otros dos van a servir para cambiar la resistencia que el plásmido confiere a la bacteria (además de darnos una idea de la probabilidad de que se haya introducido nuestra mutación).

Originalmente el plásmido es resistente a tetraciclina pero sensible a ampicilina; si ambos oligonucleótidos anillan en su lugar correspondiente, tras

la síntesis y reparación de las muescas, deberemos obtener un plásmido ApR pero TcR.

Una vez que hemos comprobado por secuenciación que la mutación es correcta los insertos se introducen por digestión en el vector definitivo. Así, los correspondientes mutantes se transfirieron a los vectores pET3a y pET16b.

## **16. ANÁLISIS *in silico***

### **16.1. Secuenciación**

Para la secuenciación de las construcciones presentadas en este trabajo hemos recurrido al servicio de secuenciación de ADN/Genómica del Instituto de Parasitología y Biomedicina “López-Neyra”. Este servicio amplifica las muestras con un termociclador PE-9600 (Perkin-Elmer) y las secuencia mediante el secuenciador ABI 373 XL Stretch (Perkin-Elmer) usando el kit ABI PRISM Big Dye™ Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer).

La visualización de los picos de secuencia se realizó mediante el programa Chromas Lite 2.01. El trabajo rutinario en las secuencias (búsqueda de dianas de restricción, diseño de cebadores, comparativa de secuencias, búsqueda de ORFs, etc.) se llevó a cabo con el programa Clone Manager Professional Suite®.

### **16.2. Comparación e identificación de secuencias**

Para la comparativa e identificación de las secuencias de las diferentes proteínas identificadas en las genotecas metagenómicas, así como para la búsqueda de nuevas secuencias en las bases de datos, se utilizaron distintos algoritmos BLAST disponibles en la página web del NCBI (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>) (Pearson y Lipman, 1988; Altschul et al., 1990). Para el alineamiento múltiple de secuencia se han usado dos versiones del programa Clustal disponibles en red (<http://www.bioweb.pasteur.fr> y <http://clustalw.genome.ad.jp/>) (Thompson et al., 1994; Higgins et al., 1996; Thompson et al., 1997).

### **16.3. Análisis filogenético**

Las distancias genéticas entre las secuencias analizadas se analizó alineándolas usando el programa MAFT versión 6 (Kato, K et al, 2008), este programa ofrece diversas estrategias de alineamientos, nosotros usamos el método Clustal W. Los análisis filogenéticos posteriores se llevaron a cabo con el método *Neighbour-Joining* usando el paquete informático *MEGA* (Kumar, S et al, 2008) y estableciendo árboles consenso tras 500 repeticiones.

## **17. DETERMINACIONES DE BIODIVERSIDAD**

Gracias a las herramientas que se enumeran en los capítulos 9.5 y 10.1, se han calculado diferentes índices de diversidad en base a los cambios nucleótidos dentro de la secuencia del gen 16S.

La gran mayoría de los métodos propuestos para evaluar la diversidad de especies se refieren a la diversidad dentro de las comunidades (alfa). Para diferenciar los distintos métodos en función de las variables biológicas que miden, los dividimos en dos grandes grupos: i) Métodos basados en la cuantificación del número de especies presentes (riqueza específica) ii) Métodos basados en la estructura de la comunidad.

La diversidad beta o diversidad entre hábitats es el grado de reemplazamiento de especies o cambio biótico a través de gradientes ambientales (Whittaker, 1972).

A continuación se indican los parámetros utilizados para estimar la diversidad de especies y la estructura dentro de una población microbiana.

### **17.1. Diversidad alfa**

#### **17.1A. Medición de la riqueza específica**

##### *17.1.A1. Riqueza de especies (S)*

También conocido como Índice de Riqueza Específica, se define de manera clásica como el número total de especies encontradas por censo de una comunidad.

En nuestro caso, al ser el concepto de especie procariota tan controvertido, no se habla de especies sino de OTUs a un nivel de significación dado. Serían por tanto, grupos donde se englobarían secuencias parecidas que

se diferencian del resto de grupos en función del nivel de significación que se quiera utilizar.

*17.1.A2. Curvas de rarefacción*

Las curvas de rarefacción o esfuerzo muestran el cambio en el valor esperado de riqueza de especies de acuerdo al tamaño de la muestras. Calcula el número esperado de especies de cada muestra si todas las muestras fueran reducidas a un tamaño estándar, es decir, si la muestra fuera considerada de  $n$  individuos ( $n < N$ ), ¿cuántas especies se habrían registrado?

$$E(S) = \sum 1 - [(N - N_i) \cdot n / (N \cdot n)]$$

Donde:

$E(S)$ : número esperado de especies

$N$ : Número total de individuos en la muestra

$N_i$ : Número de individuos de la  $i$ ésima especie

$n$ : Tamaño de la muestra estandarizado.

Normalmente vienen representadas por un gráfico, en el que en el eje de la “X” se representa el número de secuencias y en el eje de la “Y” el número de OTUs (unidades taxónomicas operativas) al nivel de significación que se desee. Te proporciona la información de si has podido acceder a todos los OTUs presentes en tu muestra con las secuencias estudiadas, o deberías de poseer muchas más secuencias para abarcar a la población.

*17.1.A3. Índice de Chao*

Es un estimador del número de especies en una comunidad basado en el número de especies raras en la muestra (Chao, 1984; Chao y Lee, 1992; Smith y van Belle, 1984).

$$\text{Chao } 1 = S + a^2/2b$$

“S” es el número de especies en una muestra, “a” es el número de especies que están representadas solamente por un único individuo en esa muestra (número de “singletons”) y b es el número de especies representadas por exactamente dos individuos en la muestra (número de “doubletons”, Colwell, 1997; Colwell y Coddington, 1994).

### **17.1B. Medición de la estructura**

#### *17.1.B1. Índice de Equidad de Shannon-Wiener ( $H'$ )*

Este índice de abundancia proporcional relaciona la cantidad de especies presentes con el número de individuos de cada una. Expresa, por tanto, el grado promedio de incertidumbre en la predicción de la especie de un individuo escogido al azar desde una comunidad (Magurran, 1988). Su expresión es:

$$H' = -\sum p_i \ln(p_i)$$

donde  $p_i$  es la abundancia proporcional de una especie, morfotipo u OTU.

#### *17.1.B2. Índice de Pielou*

Mide la proporción de la diversidad observada con relación a la máxima diversidad esperada. Su valor va de 0 a 1 que corresponde a situaciones donde todas las especies son igualmente abundantes (Magurran, 1988).

$$J = H' / H'_{\max}$$

Este índice es muy práctico cuando se quiere comparar el índice de Shannon entre dos muestras que no tienen el mismo número de especies, ya que ayuda a relativizarlo y que no dependa del número de especies analizadas.

### **17.2. Diversidad beta**

#### **17.2.1 Índice de Simpson**

$$\Lambda = \sum p_i^2$$

Donde  $p_i$  = abundancia proporcional de la especie  $i$ , es decir, el número de individuos de la especie  $i$  dividido entre el número total de individuos de la muestra.

Manifiesta la probabilidad de que dos individuos tomados al azar de una muestra sean de la misma especie. Está fuertemente influido por la importancia de las especies más dominantes (Magurran, 1988; Peet, 1974). Como su valor es inverso a la equidad, la diversidad puede calcularse como  $1/\Lambda$  (Lande, 1996).



**17.2.2. Coeficiente de similitud de Jaccard**

$$I_J = c/a+b-c$$

Donde

a= número de especies presentes en el sitio A

b= número de especies presentes en el sitio B

c= número de especies presentes en ambos sitios A y B.

El intervalo de valores para este índice va de 0 cuando no hay especies compartidos entre ambos sitios, hasta 1 cuando los dos sitios tienen la misma composición de especies.

**17.2.3. Coeficiente de similitud de Sørensen**

$$I_s = 2c/ a +b$$

Relaciona el número de especies común con la media aritmética de las especies en ambos sitios (Magurran, 1988).

Todos estos índices, se han calculado usando la herramienta bioinformática DOTUR para un número de secuencias menor a 200 y la interfaz del RDP dedicada a pirosecuenciación cuando se tiene mayor número de secuencias del gen ribosomal 16S.

**I. ANÁLISIS DE LA  
DIVERSIDAD PROCARIOTA EN  
LA RIZOSFERA DE cactaceas  
del altiplano central de  
México;  
*Mammillaria carnea***

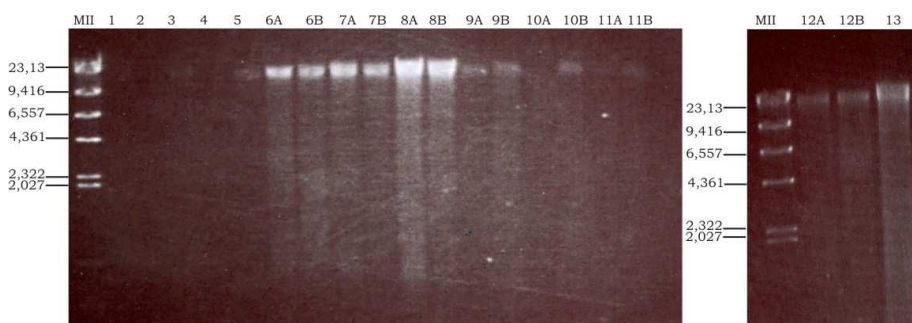


## I.1. CLASIFICACIÓN Y TOMA DE LAS MUESTRAS DE SUELO RIZOSFÉRICO.

### 11.1. Extracción de ADN de las muestras de suelo recogidas

La extracción total de ADN del suelo se hizo partiendo de 0,5 gramos de muestra tamizada por un filtro de 0,2 mm de grosor, siguiendo el protocolo que describió Porteous et al, que se resume en el material y métodos de esta tesis doctoral.

De este modo, se tomaron 50 muestras de suelo, 24 en época seca y 26 en época lluviosa. Se extrajo ADN de suelo de cada muestra por triplicado y la media de estas réplicas puede observarse en la tabla I.1 como microgramos de ADN por gramo de suelo. Normalmente, el tamaño medio del ADN extraído era aproximadamente de 20 Kb (ver figura 3), aunque se observa una degradación hacia bandas de un rango inferior.



**Figura I.1.** Geles de agarosa al 0,8% con ADN extraído de las distintas muestras. Carriles: **1:** VTDoBs; **2:** VTabBs; **3:** VTDoBs; **4:** ToDoBs; **5:** ToAbBs; **6A/6B:** VTDoMc1; **7A/7B:** VTDoMc2; **8A/8B:** VTDoMc3; **9A/9B:** VTabMc1; **10A/10B:** VT2DoMc1; **11A/11B:** VT2AbMc1; **12A/12B:** ToDoMc1; **13:** ToAbMc1

En la figura anterior, observamos que en los carriles del 1 al 5 hay menos ADN que en el resto de muestras. Estos corresponden a las extracciones relacionadas con suelo suelto, mientras que las otras son muestras rizosféricas en las que sí se observa más cantidad de ADN.

Época lluviosa				Época seca			
Zona	Ambiente	Planta	µgADN/gsuelo	Zona	Ambiente	Planta	µgADN/gsuelo
VT1	Do	Bs	0,509	VT1	Do	Bs	0,214
VT1	Ab	Bs	0,165	VT1	Ab	Bs	0,203
VT1	Do	Mc1	3,021	VT1	Do	Mc1	4,267
VT1	Do	Mc2	2,805	VT1	Ab	Mc1	1,488
VT1	Do	Mc3	6,337	VT1	Do	Op1	0,684
VT1	Ab	Mc1	1,635	VT1	Ab	Op1	0,108
VT1	Do	Op1	3,291	VT1	Do	Ss1	1,431
VT1	Ab	Op1	1,608	VT1	Ab	Mg	0,645
VT1	Do	Mg1	2,172	VT2	Do	Bs	0,741
VT1	Do	Ss1	0,384	VT2	Ab	Bs	0,648
VT2	Do	Bs	0,413	VT2	Do	Mc1	0,075
VT2	Do	Mc1	0,808	VT2	Ab	Mc1	0,285
VT2	Ab	Mc1	0,451	VT2	Do	Op1	3,237
VT2	Do	Op1	0,393	VT2	Ab	Op1	0,349
VT2	Ab	Op1	1,569	VT2	Do	Ss1	1,344
VT2	Do	Mg1	0,663	VT2	Ab	Ss1	4,803
VT2	Ab	Mg1	1,113	To	Do	Bs	0,457
VT2	Do	Ss1	0,375	To	Ab	Bs	0,132
To	Do	Bs1	0,115	To	Do	Mc1	0,093
To	Ab	Bs1	0,274	To	Ab	Mc1	1,716
To	Do	Mc1	4,756	To	Do	Op1	0,705
To	Do	Op1	0,579	To	Ab	Op1	0,369
To	Ab	Op1	0,408	To	Do	Ss1	0,591
To	Do	Ss1	0,411	To	Ab	Ss1	2,991
To	Ab	Ss1	0,498	/	/	/	/
To	Do	Mg1	0,400	/	/	/	/

**Tabla I.1.** Rendimiento de extracción de ADN en las diferentes zonas, en época lluviosa y en época seca .

Una vez extraído el ADN, se planteó el estudio estadístico de las medias de los rendimientos en la extracción para ver si podrían existir diferencias significativas entre las distintas muestras como medida indirecta de la actividad microbiana.

De este modo, se analizaron las medias mediante un análisis de varianza (ANOVA de una vía; tabla I.2) y observamos que no existen diferencias estadísticamente significativas entre el ADN extraído en época seca y lluviosa, ya que se obtienen valores muy similares en la extracción en ambos casos (1,15 y 1,35 µgADN/g de suelo de media respectivamente). Tampoco existen diferencias estadísticamente significativas entre las medias de las extracciones de ADN en las muestras bajo dosel y abierto, ni en las diferentes zonas, Valerio Trujano 1 y 2 y Tomellín, aunque en las muestras de Tomellín y Valerio Trujano 2 el rendimiento fue algo más bajo que en la zona Valerio Trujano 1 (0,9 y 1,08 frente a 1,72 respectivamente).

En el caso de rizosfera y suelo suelto, sí se observan diferencias estadísticamente significativas en los rendimientos de extracción de ADN (tabla I.2). En rizosfera se obtiene unas cuatro veces más ADN que en suelo suelto, lo que puede deberse al efecto rizosférico (Berg et al, 2002), es decir, al aumento del número de microorganismos que se produce en los ambientes rizosféricos, debido a los productos secretados por las raíces de las plantas (azúcares, aminoácidos, nutrientes, antimicrobianos, etc.), que permiten la proliferación de grupos específicos de microorganismos.

Variable <sup>a</sup>	Nºmuestra <sup>b</sup>	Media <sup>c</sup> (µgADN/g)	Error estándar <sup>d</sup>	P-value <sup>e</sup>
<b>Seco</b>	24	1,15	0,27	0,625
<b>Lluvioso</b>	26	1,35	0,3	
<b>Total</b>	50			
<b>Dosel</b>	29	1,42	0,3	
<b>Abierto</b>	21	1,02	0,25	0,337
<b>Total</b>	50			
<b>Valerio Trujano1</b>	18	1,72	0,4	0,221
<b>Valerio Trujano2</b>	16	1,08	0,31	
<b>Tomellín</b>	16	0,9	0,31	
<b>Total</b>	50			
<b>Rizosfera</b>	41	1,44	0,2	<b>0,024*</b>
<b>Suelo suelto</b>	9	0,39	0,07	
<b>Total</b>	50			

**Tabla I.2.** Medias y estadísticas de los rendimientos en las extracciones de ADN de suelo. Con un asterístico se muestra el p-valor significativo menor de 0,05

<sup>a</sup> Variables que se están comparando.

<sup>b</sup> Número de muestras en cada comparación.

<sup>c</sup> Media aritmética del rendimiento de ADN obtenido en cada variable, expresado como microgramos de ADN por gramo de suelo.

<sup>d</sup> Error estándar de las medias calculadas.

<sup>e</sup> Grado de significación al 95%.

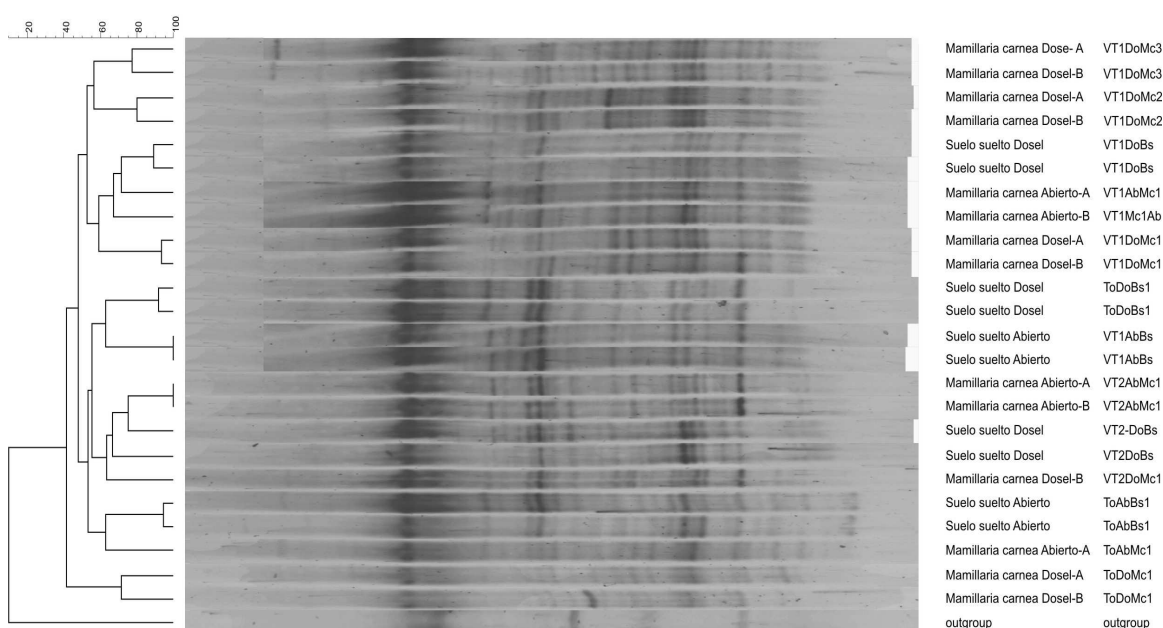
### **11.2. Análisis de la diversidad procariota ( $\alpha$ -proteobacterias) mediante el perfil genético (ADN fingerprint) obtenido con la técnica electroforesis en gel en gradiente de temperatura (TGGE)**

Para observar las diferencias en la estructura y diversidad de las comunidades microbianas presentes en las diferentes zonas, se generaron perfiles genéticos mediante la amplificación por PCR de los genes 16S ADNr presentes en el ADN total extraído de cada muestra, seguido de la separación en base a la secuencia de cada amplicón por medio de TGGE.

Estudios previos en el grupo, han demostrado una mayor discriminación en el perfil genético utilizando esta técnica, cuando el análisis de la diversidad se focaliza mediante cebadores exclusivos para el sub-phylum de alfa-proteobacterias (Villadas et al, 2007).

De este modo, se realizó una primera PCR con cebadores específicos del grupo microbiano  $\alpha$ -proteobacterias (primers  $\alpha$ U175 y L492) y una segunda PCR (sobre ésta, PCR anidada) con cebadores específicos del TGGE que contenían la cola GC (primers R1378 y U968GC) y que permiten la correcta separación de los amplificados en un gradiente de temperatura (capítulo 8.4 Material y Métodos).

Como se observa en la figura I.2, la separación electroforética mediante TGGE del producto de la PCR anidada, obtenido a partir del ADN aislado de las muestras de suelo, permitió obtener un patrón de bandas bien diferenciado que pudo ser objeto de un posterior análisis con el programa *fingerprinting*, que permite el cálculo de las relaciones filogenéticas a partir del análisis de las bandas.



**Figura I.2** Análisis del Perfil genético mediante TGGE de amplicones parciales del 16S ADN<sub>r</sub> de distintas muestras de suelo recogidas durante la **estación húmeda**. A la izquierda se muestra un análisis de similitud (Jaccard con un índice de tolerancia del 1%) en base a los perfiles de bandas mostrado. La nomenclatura de las muestras se especifica en la sección de Material y Métodos 1.6

En la figura I.2, se muestran los perfiles resultantes de amplificar con los primers del grupo bacteriano  $\alpha$ -proteobacterias, los distintos ADNs extraídos en las muestras procedentes de ejemplares de *M. carnea* de época lluviosa. Con las letras A y B, especificamos el tipo de muestra; A es suelo íntimamente adherido a la raíz (se extrajo moviendo bruscamente las raíces y se recogiendo el suelo que se desprendía; Material y métodos capítulo 1) y B es el suelo que rodea a la raíz. Como se observa en la figura I.2, en ninguno de los casos existen diferencias entre el perfil de bandas en A y B, por lo que después de esta evidencia las muestras de ADN de A y B de cada planta se unieron en una sola que se denominó como rizosférica.

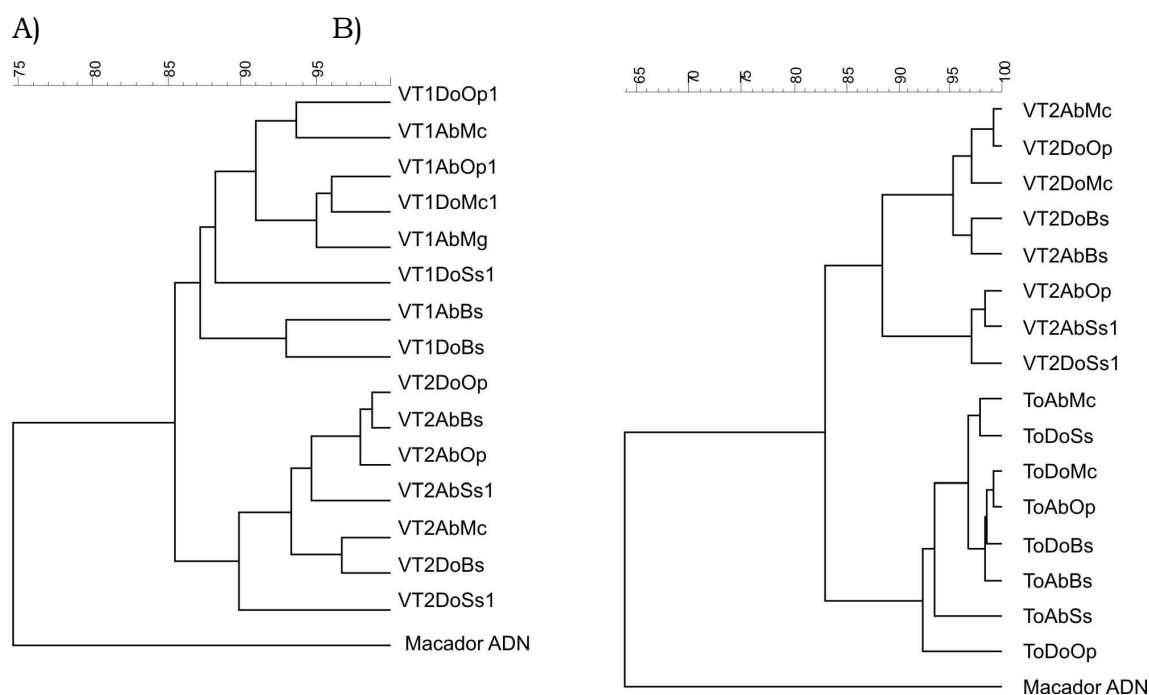
Sí que existen diferencias entre las muestras de diferente origen, permitiéndonos esta técnica agrupar a las muestras por localización geográfica (Valerio Trujano1 y 2 y Tomellín), lo que nos está indicando que las poblaciones microbianas que habitan en cada región, en cada tipo de suelo, son diferentes. Además, también existen diferencias claras entre la comunidad de  $\alpha$ -proteobacterias presente en suelo suelto y rizosfera, separándolas dentro de cada tipo de suelo, en subgrupos diferentes. Sin embargo el efecto dosel producido por la planta nodriza no parece tener una influencia clara, puesto que no existe un agrupamiento claro de las muestras obtenidas en la rizosfera de plantas bajo la influencia de la planta nodriza (muestras Dosel) frente a las de zonas sin esa influencia (Abierto).

De manera similar, también se realizaron perfiles genéticos usando los mismos cebadores tomando de molde los ADNs extraídos de las muestras de época seca, en este caso de las distintas plantas estudiadas: *M. carnea*, *O. pillifera* y *S. stellatus* (figura I.3 ).

Como se puede observar en los árboles de similitud, en este caso incluso la agrupación por localización geográfica es aún más patente que en época húmeda; quedando por un lado las zonas de VT1 y VT2 separadas a un 85% y a una mayor distancia (a un 65% de la VT2) la de Tomellín (I.3 A y B).

La diferenciación de las muestras derivada del tipo de planta parece no tener una influencia determinante, de ahí que los agrupamientos de estos árboles no parecen correlacionarse con el tipo de planta analizada. De igual manera a lo que sucedía en las muestras de época húmeda, tampoco parece haber una influencia clara del efecto Dosel en las muestras, puesto que no parece haber un agrupamiento claro de las muestras obtenidas en la rizosfera de plantas bajo la influencia de la planta nodriza.

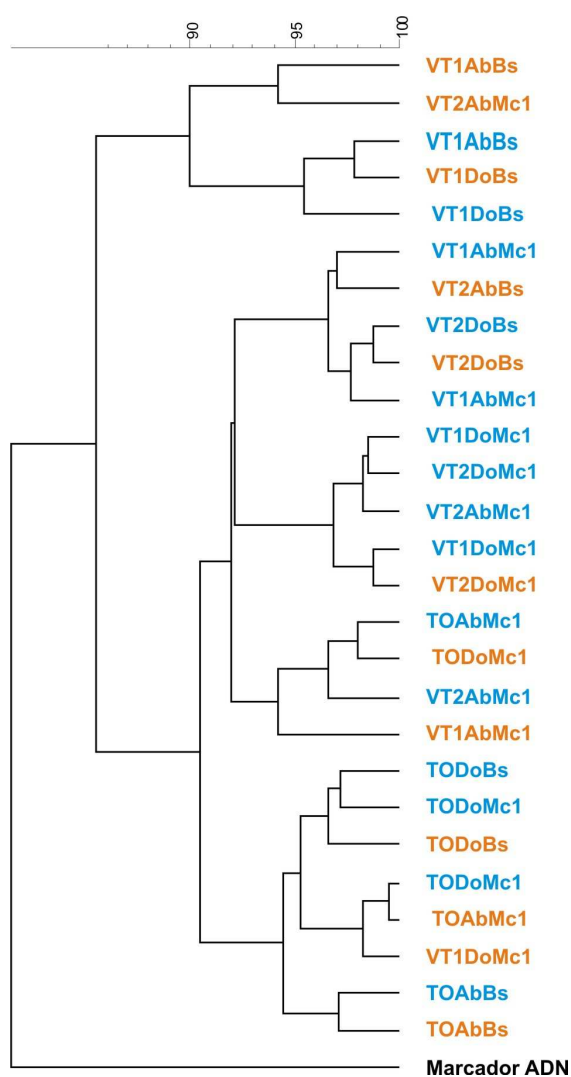




**Figura I.3.** Análisis de similitud (como en la figura I2) en base a los perfiles de bandas mostrado por el TGE de amplicones parciales de 16S ADNr de la comunidad de  $\alpha$ -proteobacterias presentes en la rizosfera de plantas recogidas en la estación seca (A) de la zona de Valerio Trujano VT1 y VT2 y (B) VT2 y Tomellín. La nomenclatura de las muestras se especifica en la sección de Material y Métodos 1.6

Los datos físico-climáticos de la toma de muestras, así como las condiciones y diferencias en la estacionalidad respecto a la recogida de las muestras, apuntan a que deben de existir cambios cíclicos (anuales) en la microbiota en relación a fluctuaciones de temperatura, grado de hidratación, ... . En tanto las diferencias existentes entre las distintas plantas no parecían ser muy acusadas en la temporada seca, decidimos establecer la comparación estacional restringiéndonos a muestras rizosféricas de una única planta, en este caso *M. carnea*, en las distintas temporadas (Fig I4).

Sorprendentemente, del análisis de similitud no parece apreciarse una diferencia clara entre el perfil de proteobacterias de las muestras debido a la estacionalidad; las muestras se agrupan indistintamente de su procedencia estacional. Sin embargo, la división por el suelo de procedencia se sigue manteniendo como ya se había observado en las anteriores comparaciones (Figuras I2-I3) las muestras procedentes de Valerio Trujano (1 y 2) están más relacionadas entre sí y diferencialmente a las de Tomellín.



**Figura I.4.** Análisis de similitud (como en la figura I2) en base a los perfiles de bandas mostrado por el TGGE de amplicones parciales de 16S ADNr de la comunidad de  $\alpha$ -proteobacterias presentes en la rizosfera de *M. carnea* obtenidas en época seca (en naranja) y lluviosa (en azul). La nomenclatura de las muestras se especifica en la sección de Material y Métodos 1.6.

Estos estudios sobre las diferencias en los rendimientos en las extracciones de ADN y los diferentes análisis en los perfiles electroforéticos de la técnica TGEE nos han servido como aproximación preliminar para estimar la diversidad procariota presente en todas las muestras.

Dada la complejidad del diseño experimental y basándonos en estos resultados, para continuar con el estudio nos centramos en la diversidad microbiana presente en la cactácea globosa endémica de la zona, denominada *Mammillaria carnea*, por varias razones: está bien representada en las tres zonas de muestreo; presenta elevados rendimientos en la extracción de ADN en época lluviosa en contraposición con la época seca (tabla I.2); y además en los estudios de TGGE se vio que podrían existir diferencias entre las poblaciones de bacterias en época lluviosa y época seca en la rizosfera de *Mammillaria carnea* en las zonas VT1 y VT2 (figura I.4).

Se pensó, de este modo, que sería interesante conocer como fluctúa la comunidad bacteriana en estas dos épocas y estimar la diversidad y

estructura de la comunidad procariota existente asociada a la raíz de esta cactácea.

## **I2. ANÁLISIS DE LA DIVERSIDAD PROCARIOTA PRESENTE EN LA RIZOSFERA DE *Mammillaria carnea* Y ESTUDIO DE SU ESTACIONALIDAD**

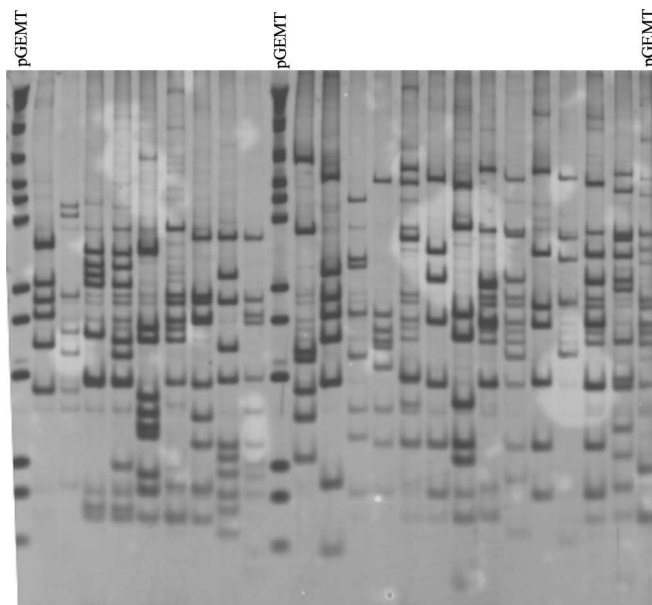
Por los motivos expuestos anteriormente, en este capítulo se va a analizar la diversidad procariota presente en la rizosfera de *Mammillaria carnea*. Se estudiará, en primer lugar, la población microbiana en época seca y lluviosa y, finalmente, nos centraremos en un grupo mayoritario y se analizará su composición en suelo suelto y en rizosfera.

### ***I2.1. Análisis de la diversidad procariota en la rizosfera de *mammillaria carnea* mediante la construcción de las librerías 16s***

Para la construcción de las genotecas 16S partimos de las muestras de ADN de la rizosfera de la cactácea globosa *Mammillaria carnea*.

De este modo, se amplificó el gen 16S ARNr con los primers universales E1591R/E9F y se obtuvo así un amplificado de aproximadamente 1500 pb que corresponde casi a la totalidad de este gen. Una vez amplificada la región, el producto de PCR se clonó en el vector pGEMT y se transformaron células de *Escherichia coli* tal y como se detalla en el material y métodos de esta Tesis Doctoral.

Se seleccionaron 210 clones con inserto que se analizaron mediante *fingerprint* utilizando la enzima de restricción *HaeIII* (figura I.5). Después de analizar los perfiles de digestión, se eliminaron las secuencias pertenecientes al gen ribosomal de *E.coli* e identificamos las que poseían un mismo patrón de digestión, mandando de esta forma a secuenciar 103 clones diferentes que contienen amplicones del gen 16S, 47 pertenecientes a época seca y 56 a época lluviosa. La secuenciación se hizo con el cebador U1510R y se obtuvo en este caso secuencias correspondientes a la región variable V8-V9.



**Figura I.5.** Digestión con *HaeIII* de varios de los amplicones 16S de las librerías construidas. Los carriles nombrados como pGEMT, son el vector digerido por la enzima de restricción *HaeIII*, los otros carriles son las muestras de estudio.

### ***12.1.1 Análisis de las secuencias***

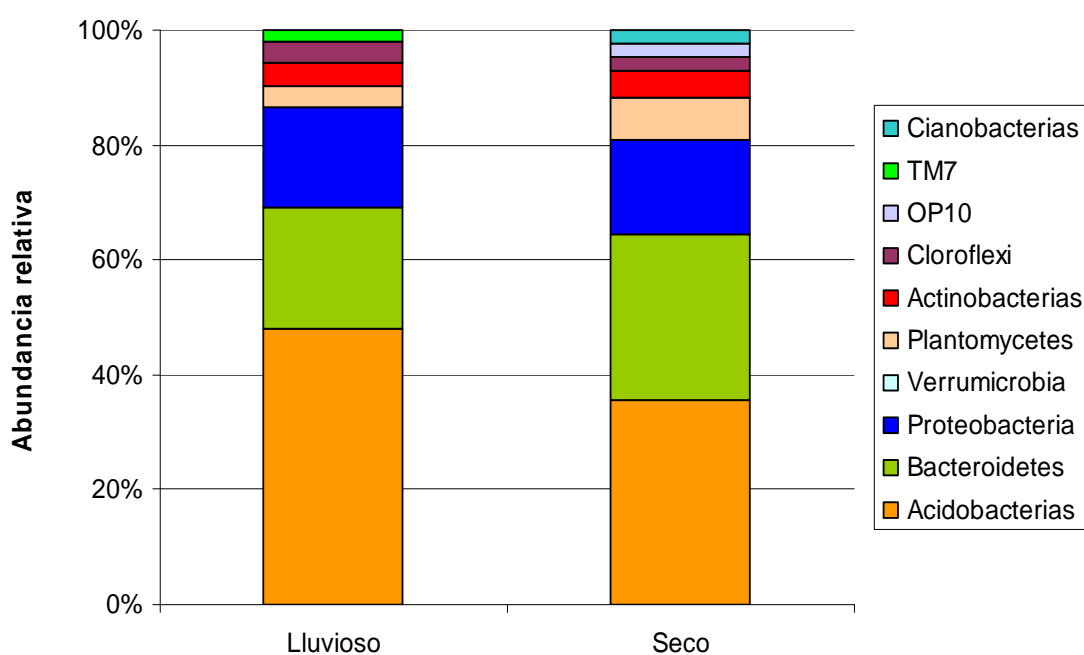
Se utilizaron las 103 secuencias para inferir un árbol filogenético, usando la herramienta correspondiente de la página web del *ribosomal data base* (<http://rdp.cme.msu.edu/>).

La clasificación taxonómica se llevó a cabo mediante el método Naïve Bayesian, introduciendo las 103 secuencias en la página web del RDP y enfrentándola a su base de datos específica de genes ribosomales 16S. Un resumen de la clasificación taxonómica se muestra en el siguiente gráfico, que distingue entre época seca y lluviosa.

De este modo se observa como la población microbiana mayoritaria en este tipo de muestras son las acidobacterias (aproximadamente un 40% de las secuencias encontradas en las librerías pertenecen a acidobacterias) y parece que predominan más en época lluviosa (44,6 frente a un 31,9 % en época seca).

*Acidobacteria*, *Verrumicrobia* y *Plantomycetes* (figura I.6; da Rocha et al, 2009) son tres phyla muy poco representados dentro de los microorganismos cultivables. La información acerca de estos taxa es muy limitada y está restringida prácticamente a estudios de genes 16S rRNA. Dentro de estos grupos, *Acidobacteria* y *Verrucomicrobia* tienen un número muy bajo de especies recogidas en la literatura. Sin embargo, son de los grupos más abundantes en numerosos suelos ((Lauber *et al*, 2008; Jones *et al*, 2009; Quaiser *et al*, 2003).

Tradicionalmente, las acidobacterias se han considerado habitantes de ambientes ácidos. De hecho, el primer individuo que se aisló en 1991, *Acidobacterium capsulatum*, proviene de una mina ácida (Kishimoto et al, 1991). Sin embargo, este phylum bacteriano está compuesto por diferentes subgrupos (1-26) y además, como se ha observado posteriormente, son bacterias muy ubicuas que se pueden encontrar en diferentes tipos de suelos, sedimentos y aguas marinas. La relativa abundancia de la mayoría de los subgrupos de Acidobacterias sí está fuertemente relacionada con el pH del suelo. El subgrupo GP4, por ejemplo, que es el más abundante en nuestras genotecas (tabla I.4), está relacionado con pHs neutros o altos (Jones et al, 2009).



**Figura I.6.** Representación del porcentaje de secuencias dentro de los 10 phyla encontrados en las librerías 16S.

	Época Seca	Época lluviosa
<b>Acidobacteria</b>	Gp6 (2)	Gp3 (1)
	Gp3 (2)	Gp7 (2)
	Gp4 (10)	Gp6 (3)
	unclassified_"Acidobacteria" (1)	Gp4 (19)
<b>Bacteroidetes</b>	Sporocytophaga (1)	Flavisolibacter (4)
	Flavisolibacter (4)	unclassified_"Chitinophagaceae" (2)
	Ferruginibacter (1)	unclassified_"Sphingobacteriales" (3)
	Sphingobacteriales (2)	unclassified_"Bacteroidetes" (2)
	unclassified_"Chitinophagaceae" (4)	
<b>Proteobacteria</b>		unclassified_Sphingomonadaceae (2)
	Steroidobacter (1)	unclassified_Alphaproteobacteria (1)
	Escherichia/Shigella (1)	unclassified_Nannocystineae (2)
	unclassified_Sphingomonadaceae (2)	Thiobacter (2)
	unclassified_Comamonadaceae (1)	unclassified_"Proteobacteria"
	unclassified_"Proteobacteria" (1)	
<b>Verrumicrobia</b>	Subdivision3_genera_incertae_sedis (4)	Subdivision3_genera_incertae_sedis (2)
		Spartobacteria_genera_incertae_sedis (2)
<b>Plantomycetes</b>	Zavarzinella (1)	unclassified_Planctomycetaceae (2)
	unclassified_Planctomycetaceae (2)	
<b>Actinobacterias</b>	Rubrobacter (1)	Rubrobacter (1)
	Arthrobacter (1)	Actinoplanes (1)
<b>Cloroflexi</b>	unclassified_"Chloroflexi" (1)	Herpetosiphon (2)
<b>OP10</b>	OP10_genera_incertae_sedis (1)	
<b>Cianobacterias</b>	unclassified_Cyanobacteria (1)	
<b>TM7</b>		TM7_genera_incertae_sedis

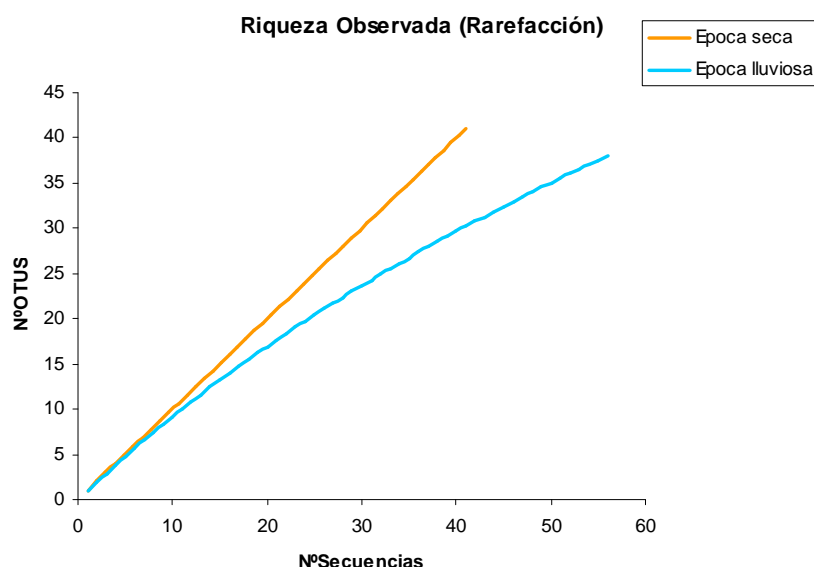
**Tabla I.3.** OTUs representados en las diferentes genotecas 16S con un nivel de confianza >80%. Los números entre paréntesis a la derecha de algunos géneros indican el número de secuencias que se clasifican dentro de ese género

### 12.1.2 Riqueza observada

Una vez definidos los diferentes tipos de secuencias o filotipos y antes de proceder a realizar los análisis de diversidad, se determinó si el número de clones analizados permitía estimar adecuadamente la diversidad procariota presente en las raíces de *Mammillaria carnea* en las dos épocas muestreadas. Para ello, se realizaron las curvas de esfuerzo o rarefacción para cada una de las épocas a una distancia molecular del 0,03%.

La tendencia lineal de esta gráfica nos indica que el número de secuencias analizadas no sería suficiente para estimar la diversidad procariota. Probablemente, esto sea debido a la gran heterogeneidad del ecosistema (el suelo en este caso). Para poder hacer este tipo de estudios,

posiblemente se tengan que abordar otro tipo de metodologías, como la pirosecuenciación o el estudio de un phylum específico.



**Figura I.7.** Curva de rarefacción. El número de secuencias analizadas en época lluviosa es de 56 y en época seca es de 47. El número de OTUs obtenido es de 41 en época seca y 38 en época lluviosa. La distancia considerada entre OTUs fue de 0,03 (3%).

Se observa, en este caso (figura I.7), que la diversidad observada es ligeramente mayor en las muestras recogidas en época seca que en las muestras procedentes de la época lluviosa.

**12.1.3. Índices de equidad y riqueza específica (índice de Chao).**

Para analizar la diversidad presente en las muestras, se calcularon los índices de diversidad de Shannon-Wiener y de riqueza específica. Gracias a la riqueza específica, podemos saber el número de OTUs estimados que existirían en nuestras muestras a un grado de significación determinado.

En ambas muestras los índices son altos y muy similares, un poco más altos en época seca, al igual que ocurría con las curvas de rarefacción (tabla I.3).

	Riqueza específica	Índice de Shannon-Wiener (H')
Época seca	207	3,64
Época lluviosa	77	3,47

**Tabla I.4.** Índice de diversidad de Shannon-Wiener y riqueza específica para cada una de las épocas muestreadas. La distancia considerada entre OTUs fue de 0,03(3%).

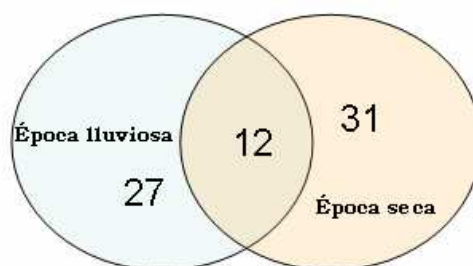
**12.1.4. Diferencias entre la diversidad procariota presente en época seca y época lluviosa**

Gracias a la herramienta informática  $\zeta$ -Libshuff, comparamos las secuencias de las librerías construidas en época seca y lluviosa. Para ello, se analizaron las diferencias a nivel de secuencia usando los datos que nos proporcionó una matriz de distancias. Como se observa en la tabla I.5, en la que se representan los p-valores obtenidos mediante este test, no existen diferencias estadísticamente significativas entre las librerías construidas en época seca y época lluviosa, ya que los p-valores son mayores a 0,05.

	Época seca	Época lluviosa
Época seca	0	0,8589
Época lluviosa	0,6030	0

**Tabla I.5.** p-valores obtenidos con la herramienta informática  $\zeta$ -Libshuff tras hacer la comparativa de ambas genotecas.

La homología entre ambas comunidades se determinó usando el paquete bioinformático SONS, que calcula el número de OTUs comunes entre parejas de comunidades usando distintos índices. Los datos de los OTUs se transformaron en tablas binarias (los OTUs se cuentan como presencia de secuencias de las dos genotecas o ausencia) en una hoja de Excel para calcular el número de OTUs compartidos entre los suelos.

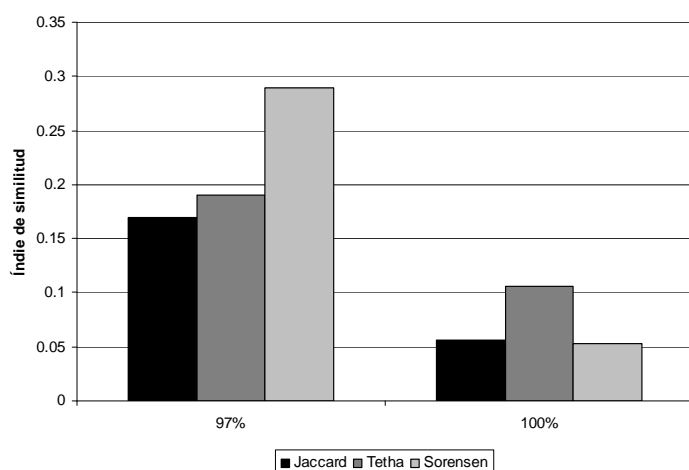


**Figura I.8.** Diagrama de Venn mostrando los OTUs compartidos entre época seca y lluviosa. Los OTUs se definen con un porcentaje de similitud del 97%.

Este tipo de gráficos (figura I.8) nos muestra que de los 70 OTUs que se forman a una distancia del 0,03%, habría 31 OTUs endémicos de época seca y 27 que sólo aparecen en época lluviosa, mientras que 12 OTUs constituirían un núcleo común entre época seca y lluviosa.



En este trabajo presentamos también el índice de Jaccard (figura I.9), que se calcula como el número de OTUs compartidos en comparación el número total de OTUs en ambas muestras, y es una simple medida de las especies compartidas sin involucrar a la abundancia. En el índice  $\theta$  sí que se relaciona la abundancia.



**Figura I.9.** Índices de similitud de Jaccard, Sørensen y Tetha entre época seca y lluviosa a distintas distancias moleculares 100% y 97%.

Reforzando lo que nos muestran los diagramas de Venn, el índice de Jaccard es relativamente bajo a un nivel de significación del 97 % (0,16) lo que nos indica que el número de especies que se comparten entre época seca y lluviosa no es excesivamente alto.

El índice de Sørensen es similar al Jaccard, se trata de un índice de riqueza, mientras que el índice Theta ya no sólo tiene en cuenta los OTUs que se comparten sino también la abundancia relativa de estos. Ambos índices, Theta y Sørensen, son también bajos, por debajo del 0,3 de similitud.

### **13. ANÁLISIS DE LA DIVERSIDAD PROCARIOTA PRESENTE EN LA RIZOSFERA DE *Mamillaria carnea* EN ÉPOCA LLUVIOSA Y SECA MEDIANTE PIROSECUENCIACIÓN**

Los ensayos descritos en el capítulo I.2 anterior demuestran la enorme diversidad microbiana presente en la rizosfera de la cactácea *Mamillaria carnea* tanto en época seca como en época lluviosa (índices de Shannon de 3,64 en época seca 3,47 lluviosa respectivamente). Además, las curvas de rarefacción (figura I.7) nos permiten inferir que el número de clones analizados en las genotecas 16S (56 y 47 que corresponden a 41 y 38 OTUs al 0,03 de significación) es inferior al número de clones necesarios para abarcar toda la diversidad presente en las muestras (índice de Chao, 207 y 77).

De este modo, se consideró que el siguiente paso de esta investigación debería ser el estudio en profundidad mediante pirosecuenciación de la dinámica temporal en la composición de la comunidad microbiana que coloniza la planta objeto de estudio.

Averiguar esa dinámica mediante otro tipo de abordaje es precisamente el objetivo de las investigaciones que se presentan en este capítulo.

### **I3.1. Preparación de la muestra objeto de la pirosecuenciación**

Para preparar la muestra que se mandó a pirosecuenciar, se amplificó la región del gen 16S V4-V5 con los cebadores G1R/G1F y G2R/G2F obteniéndose un amplicón de unos 408 pb, tal y como se detalla en el Material y Métodos de esta tesis doctoral.

Tras las amplificaciones y concentraciones en las columnas Microcon YM-100 (Millipore), poseíamos un volumen de 200  $\mu\text{l}$  de cada muestra (de época seca y lluviosa), a una concentración de 25 ng/ $\mu\text{l}$ . Estas dos muestras se unieron con otras 9 procedentes de otros estudios que están siendo desarrollados en el grupo de investigación en cantidad equimolar y se concentraron hasta 500 ng/ $\mu\text{l}$  mandando a la empresa encargada de la pirosecuenciación 20  $\mu\text{l}$  de esta muestra.

### **I3.2. Análisis de las secuencias y definición de Filotipos.**

Como se simplifica en la tabla I.6, en el archivo madre resultante del proceso de pirosecuenciación tenemos 15410 secuencias del gen 16s en la muestra de época seca y 13083 secuencias de época lluviosa con una media de 409 nucleótidos de longitud en ambos casos. Incrementamos en dos ordenes los datos obtenidos mediante la generación de genotecas clásicas de 16S ADNr

Tras someter estas secuencias al proceso de limpieza (eliminación de secuencias cortas y con “n”; Material y Métodos sección 10 ) en el pipeline del *rdp* (<http://rdp.cme.msu.edu/>), nos quedamos con 14410 secuencias en época seca y 12075 secuencias para época lluviosa.

	<b>Secuencias totales</b>	<b>Secuencias limpiadas</b>	<b>Longitud media de las secuencias</b>
<b>Época lluviosa</b>	13083	12075	409
<b>Época seca</b>	15410	14410	409

**Tabla I.6.** Número de secuencias de amplicones del gen ribosomal 16S obtenidos en la pirosecuenciación: a) Número de secuencias totales b) Número de secuencias tras eliminar aquellas que tuvieran menos de 150 nucleótidos y cuya lectura no fuera buena.

Una vez que nuestras secuencias se encontraban libres de cebadores, y que tenían la calidad adecuada pasamos a realizar la clasificación taxonómica de las mismas mediante la herramienta RDP-CLASSIFIER presente en esta página web (<http://rdp.cme.msu.edu/>) (tabla I.7 y Figura I.10).

De todas las secuencias 16S analizadas, un 1,3% en época seca y un 1,4% en época lluviosa, pertenecen a **Archaeas** (tabla I.7). Dentro de este grupo, en ambos casos, más del 90% de las secuencias pertenecen al phylum Chrenarcheota y a su vez a la clase Termoprotei. En esta clase se encuentran microorganismos principalmente autótrofos y extremófilos, capaces de vivir y adaptarse a altas temperaturas. Hasta la fecha no existen evidencias que asocien individuos de esta clase a muestras de suelos rizosféricos o a las características microclimáticas que se observan en este tipo de suelos.

Con respecto a las secuencias de **bacterias**, hay un gran porcentaje de secuencias sin clasificar por el RDP database al nivel del 80%, en torno a un quinto de las secuencias obtenidas

Centrándonos en los grupos microbianos que sí pudimos clasificar, encontramos 13 phyla con un número de secuencias superior a 10 (para facilitar la interpretación de los datos los phyla con menos de 10 secuencias, menos del 0.1% del total, se descartaron del análisis).

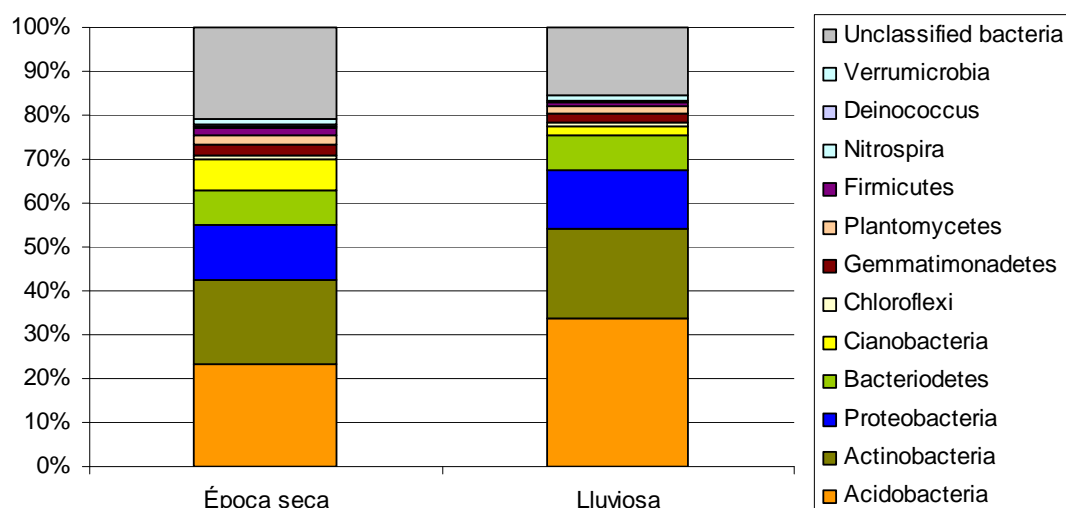
El phylum mayoritario con un porcentaje aproximado del 30% de las secuencias son las acidobacterias, en ambos casos, tanto en época lluviosa como en época seca; aunque parece que el número aumenta en época lluviosa (32.9% frente a un 23.2).

Los siguientes phyla mayoritarios son las actinobacetas (20%), proteobacterias (12%) y bacteroidetes (8%). A estos phyla ya les siguen otros minoritarios como el de las gemmatimonadetes, plantomycetes y verrumicrobia que son tres phyla con un número muy escaso de

representantes en lo que a microorganismos cultivables se refiere. Cabe destacar también al phylum cianobacteria cuyo porcentaje aumenta unas 3,5 veces en época seca.

Bacteria	Época seca	Época lluviosa
	Secuencias <sup>a</sup> (%) <sup>b</sup>	Secuencias <sup>a</sup> (%) <sup>b</sup>
<b>Acidobacteria</b>	3291 (23,2)	3912 (32,9)
<b>Unclassified bacteria</b>	2971 (20,9)	2340 (19,7)
<b>Actinobacteria</b>	2719 (19,1)	1798 (15,1)
<b>Proteobacteria</b>	1769 (12,5)	1543 (13)
<b>Bacteriodetes</b>	1178 (8,3)	921 (7,7)
<b>Cianobacteria</b>	956 (6,7)	229 (1,9)
<b>Gemmatimonadetes</b>	334 (2,4)	205 (4,2)
<b>Plantomycetes</b>	313 (2,2)	194 (1,6)
<b>Firmicutes</b>	232 (1,6)	126 (1,1)
<b>Verrumicrobia</b>	191 (1,3)	119 (1)
<b>Chloroflexi</b>	153 (1,1)	137 (1,2)
<b>Nitrospira</b>	76 (0,5)	48 (0,4)
<b>Deinococcus</b>	10 (0,1)	/
Archeas		
<b>Euryarcheota</b>	7 (3,9)	1,2 (2)
<b>Chrenarcheota</b>	166 (91,7)	163 (98,2)
<b>Unclassified</b>	8 (4,4)	1(1)

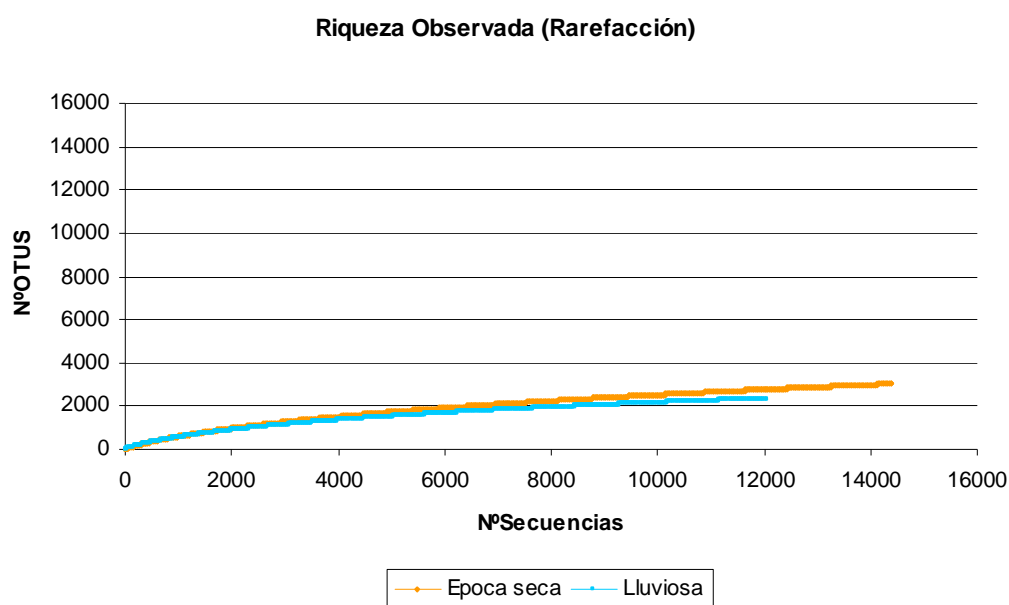
**Tabla I.7.** OTUs representados en las muestras objeto de estudio con un nivel de confianza >80%. a) Número neto de secuencias b) Porcentaje frente al total de secuencias de esa muestra.



**Figura I.10.** Abundancia relativa de OTUs al 0,03% en las diferentes muestras de estudio.

### I3.3 Análisis de la estructura de la comunidad microbiana

Como se muestra en las curvas de esfuerzo presentadas en la figura I.12, el análisis de 14410 y 12075 secuencias en época seca y lluviosa respectivamente fue suficiente para cubrir la diversidad procariota presente en la raíz de *Mammillaria carnea* en los tiempos ensayados, al contrario que con los datos que procedían de las genotecas 16S (figura I.7). Además, se observa que la diversidad en la época seca es ligeramente superior a la diversidad en la época lluviosa.



**Figura I.11.** Curvas de rarefacción al 0,03% en época seca y lluviosa.

### I3.4. Índices de equidad y riqueza específica

Al igual que en el capítulo anterior, para analizar la diversidad presente en las muestras se calcularon los índices de diversidad de Shannon-Wiener y de riqueza específica para cada una de las épocas muestreadas.

No existen diferencias significativas entre ambas épocas en cuanto a diversidad se refiere.

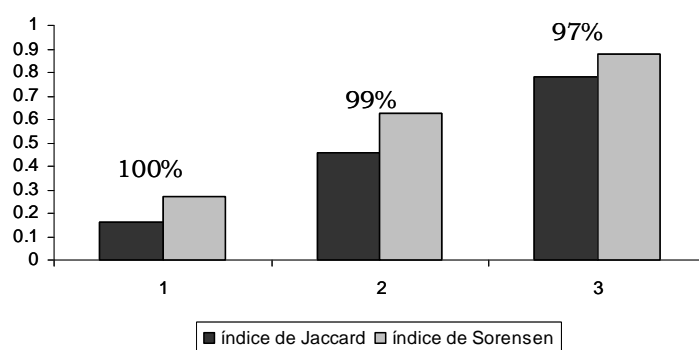
	<b>Riqueza específica (índice de Chao)</b>	<b>Riqueza observada</b>	<b>(H')</b>	<b>H' máxima</b>	<b>J'</b>
Época seca	5277,66	3033,97	6,92	8,02	0,86
Época lluviosa	3688	2347,04	6,65	7,76	0,85

**Tabla I.8.** Índices de riqueza y diversidad en época seca y lluviosa con un nivel de significación del 0,03%.

### 13.5. Comparación de las dos zonas de muestreo

#### 13.5.1 Número de OTUs comunes entre época seca y lluviosa

Para poder estimar que secuencias se compartían y cuales eran únicas en época seca y lluviosa, se usó el índice de Jaccard (figura I.12), que compara el número de OTUs presentes en los dos sitios frente al número de especies comunes.



**Figura I.12.** Índices de similitud de Jaccard y Sørensen entre época seca y lluviosa a diferentes distancias moleculares, las primeras columnas son OTU con un 97% de similitud, las segundas con un 99% y las terceras con un 100% de similitud.

El elevado índice de Jaccard al 0,03 (0,782, figura I.13) nos indica que el número de OTUs comunes que tienen las dos muestras a esta distancia molecular es muy elevado.

#### 13.5.2. Phyla diferenciados entre época seca y época lluviosa

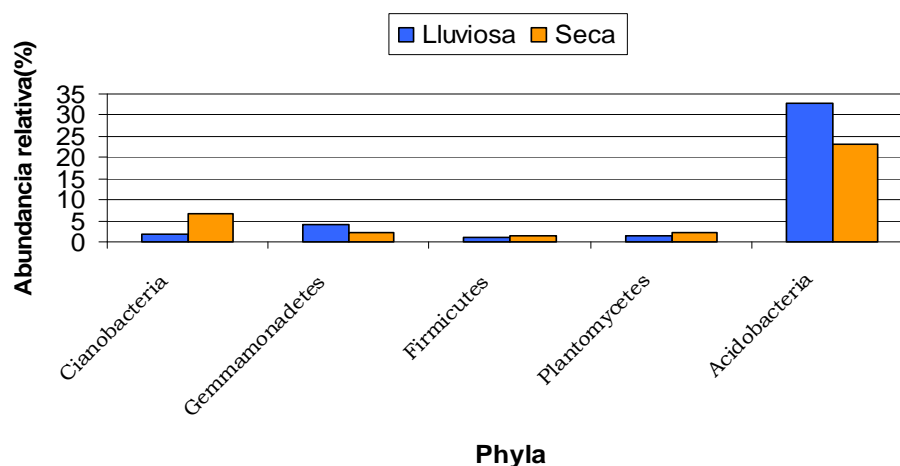
Además de este estudio, para poder determinar si existían diferencias estadísticamente significativas entre los diferentes phyla de ambas épocas, se usó la herramienta *lib compare* del *rdp*. (<http://rdp.cme.msu.edu/comparison/comp.jsp>)

Phyla	Número secuencias (lluviosa/seco)	p-valor
<b>Cyanobacteria</b>	228/955	6,00E-14
<b>Gemmamonadetes</b>	492/337	6,00E-14
<b>Firmicutes</b>	124/224	6,00E-04
<b>Plantomycetes</b>	191/315	3,40E-04
<b>Acidobacteria</b>	3907/3286	6,00E-14

**Tabla I.9.** p-valor resultante de la comparación de los diferentes phyla en época seca y lluviosa.

De este modo, encontramos diferencias estadísticamente significativas (con un nivel de significación por debajo del 0,01) en los phyla acidobacterias,

cianobacterias, gemmatimonadetes, plantomycetes y firmicutes (tabla I.9; figura I.13) entre época seca y lluviosa. Las acidobacterias y gemmatimonadetes aumentan en número en la época lluviosa, mientras que en el caso de los plantomycetes, firmicutes y cianobacterias son más numerosos en época seca.



**Figura I.13.** Abundancia relativa de los phyla cianobacterias, gemmatimonadetes, firmicutes, plantomycetes y acidobacteria en época seca y lluviosa.

Observando estos resultados, se planteó el estudio pormenorizado de cada uno de estos phyla (tabla I.10) para observar las posibles diferencias en el caso de riqueza específica (índice de Chao) y diversidad (índice de Shannon e índice de Piloni). Asimismo, además de estos phyla en los que existían diferencias significativas, se amplió el estudio a los otros phyla mayoritarios, es decir proteobacterias, bacteroidetes y actinobacterias.

### ***Proteobacterias y Actinobacterias***

Si nos fijamos en detalle en los grupos del phylum *Proteobacteria* (tabla I.11), observamos que la clase mayoritaria es la de las alfa-proteobacterias y que la abundancia relativa de ésta es mayor en época seca que en época lluviosa. En el caso de las Betaproteobacterias, su abundancia relativa es mayor en época lluviosa que en época seca.

	0,03%	Total	<i>Proteobacteria</i>	<i>Acidobacteria</i>	<i>Bacteriodes</i>	<i>Cyanobacteria</i>	<i>Gemmamonades</i>	<i>Plantomicetes</i>	<i>Actinobacteria</i>
<b>Seqs</b>	Seca	14410	1775	3289	1176	956	331	314	2716
	Lluviosa	12075	1545	3912	925	228	501	191	2342
<b>Chao</b>	Seca	5277.66	802.8	378	466	136	125.95	297.53	767.62
	Lluviosa	3688	745.51	295	347	53	143.85	336.07	467
<b>R-Observada</b>	Seca	3033.9	484.86	366	251	103	94	158	354
	Lluviosa	2347	434.87	238	227	40	116	111	443
<b>H'</b>	Seca	6.97	5.45	4.13	4.71	3.57	4.12	4.74	5.17
	Lluviosa	6.65	5.47	4.1	4.76	1.86	4.16	4.47	5.05
<b>H' máx</b>	Seca	8.02	6.18	5.90	5.53	4.63	4.54	5.06	6.64
	Lluviosa	7.76	6.08	5.47	5.42	3.69	4.75	4.71	6.14
<b>J'</b>	Seca	0.87	0.88	0.70	0.85	0.77	0.91	0.94	0.77
	Lluviosa	0.86	0.90	0.75	0.88	0.50	0.88	0.95	0.82

**Tabla I.10.** Tabla resumen en la que se indican los índices de diversidad y riqueza al 0.03% de los phya mayoritarios (acidobacterias, proteobacterias y bacteriodes) y de los que existen diferencias significativas entre época seca y lluviosa (cianobacterias, gemmamonades, plantomicetes y actinobacterias). En la tabla se especifican el índice de Shanon (H), índice de Shanon máximo (H') e índice de Pilou (J'), también se muestran valores de riqueza como el índice de Chao o la riqueza observada. Las secuencias de las que se parte este estudio vienen indicadas como "Seqs"



A rasgos generales, podríamos decir que el número global de proteobacterias es muy parecido en ambas épocas (1775/1545 en época seca/época lluviosa), así como que los índices de riqueza y diversidad también son muy similares, por lo que parece que es un grupo que permanece estable a lo largo del año, sin variaciones importantes.

	Seco	Lluvioso
<b>Gammaproteobacteria</b>	4.5 %	7.2 %
<b>Betaproteobacteria</b>	13.6 %	27.9 %
<b>Deltaproteobacteria</b>	16.3 %	13.7 %
<b>Alphaproteobacteria</b>	61.4 %	48.7 %
<b>unclassified_ "Proteobacteria"</b>	4.2 %	2.5 %

**Tabla I.11** Abundancia relativa de las secuencias del phylum Proteobacterias.

Algo parecido a lo anterior también ocurre con el *phylum Actinobacteria*. Parece un grupo muy robusto y estable sin apenas variaciones entre época lluviosa y seca en lo que a riqueza y diversidad se refiere.

### ***Bacterioidetes***

Los bacterioidetes, por su parte, aumentan en número en época seca y la diversidad, sin embargo, es mayor en época lluviosa (0,88 frente a 0,85 en el índice de Pilou). Casi todos los bacterioidetes presentes en las dos épocas pertenecen a la clase Sphingobacteria.

	Seca	Lluviosa
<b>Bacteroidia</b>	0.20%	/
<b>Flavobacteria</b>	1%	0.30%
<b>Sphingobacteria</b>	81.70%	87.80%
<b>unclassified_ "Bacterioidetes"</b>	17.10%	11.90%

**Tabla I.12.** Abundancia relativa de las secuencias del phylum bacterioidetes.

### ***Gemmamonadetes, Plantomycetes***

En el caso de los gemmamonadetes y plantomycetes, ambos aumentan en número en época lluviosa y la diversidad es mayor en época seca.

En cuanto a la clasificación taxonómica de las secuencias, todos los Gemmamonadetes pertenecen al género *Gemmatimonas*. En el caso de los Plantomycetes, pertenecen a la familia *Plantomycetaceae*.

**Firmicutes**

Los individuos pertenecientes al phylum *Firmicutes* aumentan en época seca y el 50% de las secuencias se clasifican dentro del grupo *Clostridia*.

**Cianobacterias**

Las diferencias observadas sobre el phylum de las cianobacterias son destacables. Este grupo aumenta cuatro veces en número durante la época seca, pasando de 229 a 956 lecturas.

Fijándonos en los índices de diversidad (tabla I.10), es, además, mucho más diverso en esta época, en la seca, que en época lluviosa (con un índice de pilou de 0.77 frente a 0.5 respectivamente). Esta diversidad menor en época lluviosa se explica tras analizar los subgrupos presentes en ambas muestras (tabla I.13). Existe un número de secuencias muy elevado sin clasificar (780 lecturas) en la muestra de época seca (un 80 % de los datos de cianobacterias obtenidos).

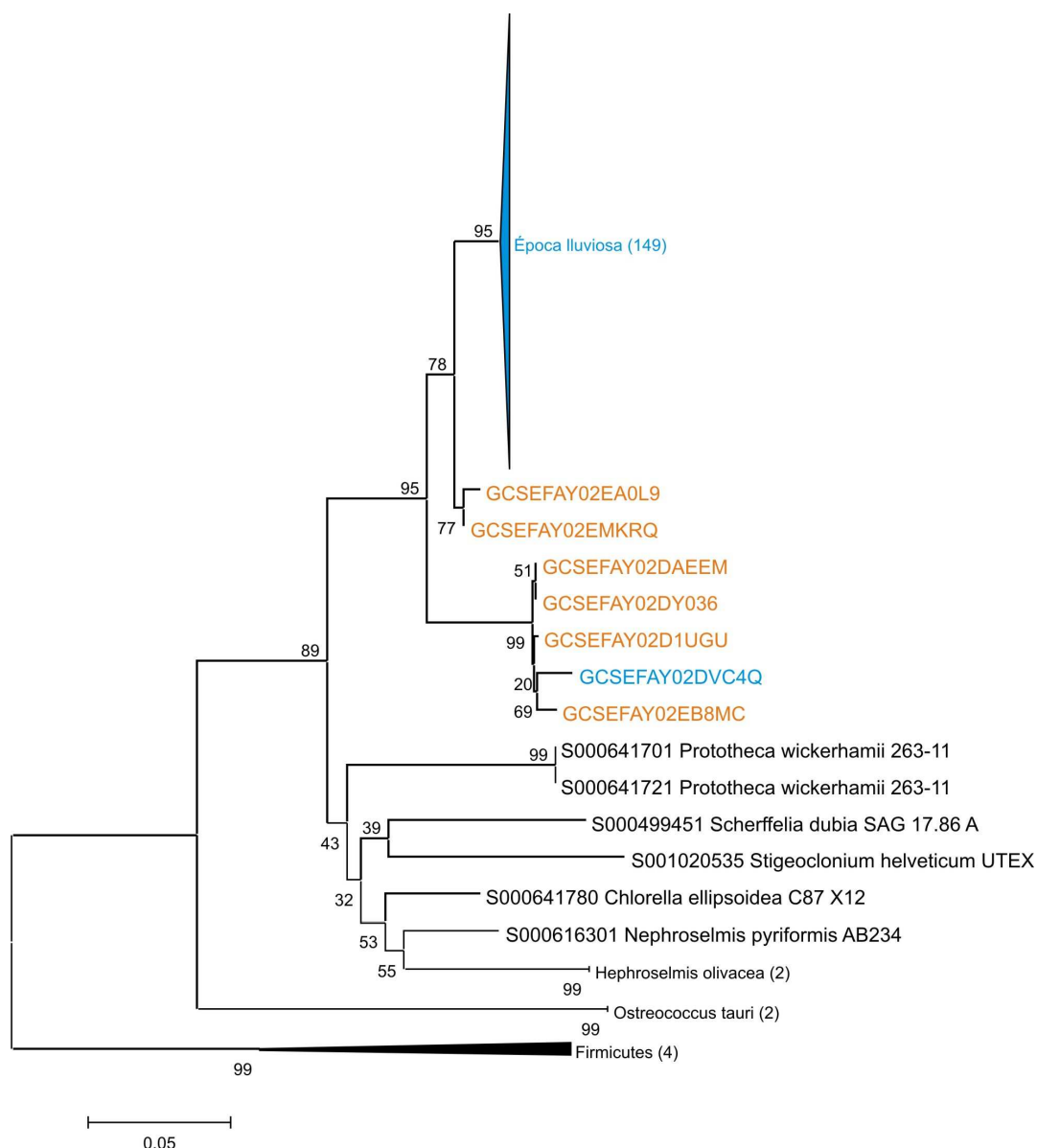
Otra diferencia entre las muestras observable en relación a este phyla es el referente al grupo 16S ADNr relacionado con plastidios de plantas (Familia Chloroplast; tabla I.14). Más del 50 % de las secuencias encontradas en la muestra de época húmeda se corresponden a secuencias pertenecientes a este grupo taxonómico; sin embargo, de este grupo sólo encontramos 6 representantes, (un 0,63%) en la época seca.

La explicación de este resultado podría estar relacionada con un aumento clonal de un grupo específico de cianobacterias que disminuiría la diversidad del ecosistema en cuanto a este phylum se refiere en comparación con la diversidad existente en época seca

	Lluviosa	Seca
	Secuencias (%)	
<b>Familia XIII</b>	7 (3.07)	30 (3.14)
<b>Familia I</b>	8 (3.51)	135 (14.14)
<b>Familia Chloroplast</b>	150 (65.79)	6 (0.63)
<b>Familia V</b>	/	2 (0.21)
<b>Familia IX</b>	/	2 (0.21)
<b>Unclassified</b>	63 (27.63)	780 (81.68)
<b>Total</b>	228 (100.00)	955 (100.00)

**Tabla I.13.** Abundancia relativa y número de secuencias de las diferentes familias dentro del phylum *Cianobacteria*.

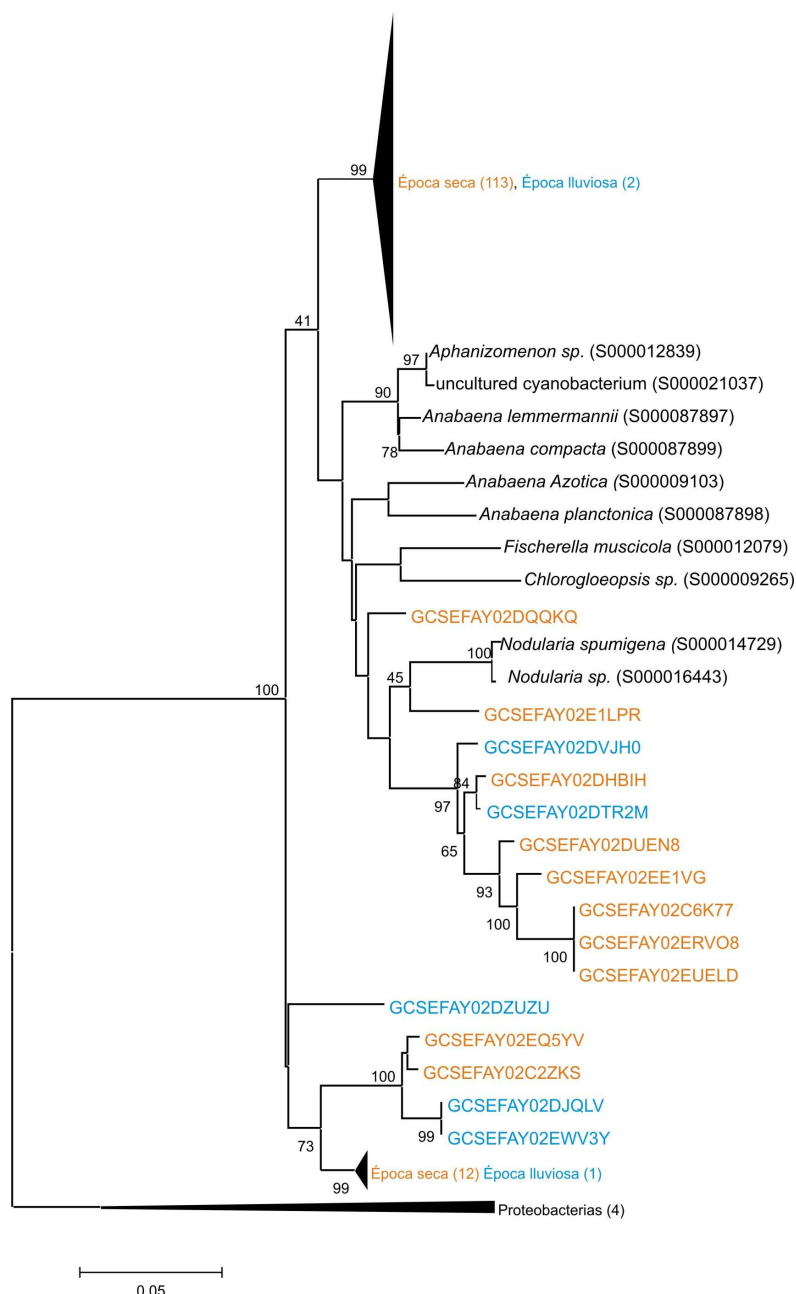
Esta hipótesis se refuerza mediante el análisis árbol filogenético con las 150 secuencias de época lluviosa y las 6 secuencias de época seca pertenecientes a la familia Chloroplast (Figura I.14; familia constituida por secuencias ribosomales relacionadas con las que se encuentran en plastidios de plantas y secuencias de microorganismos no cultivables). En el mismo, se observa como 149 de las secuencias presentes a época lluviosa están muy relacionadas entre sí formando un grupo monofilético específico muy concreto que estaría englobado por la diversidad presente en lecturas de la época seca.



**Figura I.14.** Árbol filogenético donde se observan las distancias filogenéticas de las 150 secuencias encontradas en época lluviosa (subrayadas en azul), 6 secuencias encontradas en época seca (en naranja) correspondientes al “grupo” Familia Chloroplast (Tabla I.13) y 10 secuencias de genes 16S bajados de la base de datos del rdp correspondientes a plastidios de las plantas indicadas. El grupo *Firmicutes* se usó como *outgroup*.

Con respecto a la familia I perteneciente al phylum cianobacterias, vemos que también difiere entre época seca y lluviosa, aumentando en número en época seca al contrario que ocurría con el género *Streptophyta*.

Si se realiza un árbol filogenético con las 143 secuencias aisladas en nuestro trabajo y otras 10 obtenidas de las bases de datos, observamos como en época seca habría una explosión clonal de un tipo de microorganismos que harían que la población microbiana cambie



**Figura I.15.** Árbol filogenético donde se observan las distancias filogenéticas de las 135 secuencias encontradas en época seca (en naranja), 8 secuencias encontradas en época lluviosa (en azul) y 10 secuencias de genes 16S obtenidos de la base de datos del rdp correspondientes al “grupo” de la Familia I (grupo donde se engloban estas secuencias, tabla I.13) 4 secuencias del phylum *Protobacteria* se utilizaron como *outgroup*.

De este modo, hemos observado dos ejemplos de cómo fluctúa la población microbiana de cianobacterias en función de las condiciones ambientales, época seca o lluviosa, predominando un tipo de cianobacterias u otras.

**Acidobacterias**

Finalmente, en el caso de las acidobacterias, constituyen el phylum mayoritario en las dos épocas en este tipo de suelos, con una abundancia relativa de un 23,2% de las secuencias totales en época seca y un 32,9% de las secuencias totales en época lluviosa. Su número aumenta en época lluviosa y su diversidad disminuye.

Un porcentaje elevado (alrededor del 80% en ambos casos) de las acidobacterias presentes en época seca y en época lluviosa pertenece al grupo GP4. Este grupo es un grupo característico de suelos con pH neutro (Jones *et al*, 2009). También aparecen otros grupos minoritarios como el GP6, el GP7 y el GP3 (ver tabla I.14)

	<b>Seca</b>	<b>Lluviosa</b>
<b>Gp4</b>	85.6	88.5
<b>Gp6</b>	5.7	5.9
<b>Gp7</b>	3.9	2.7
<b>Gp3</b>	3.9	2
<b>Gp10</b>	0.8	0.5
<b>Gp17</b>	0.1	0.1
<b>Gp1</b>	0	0
<b>Gp5</b>	0	0
<b>Unclassified</b>	0.5	0.2

**Tabla I.14.** Porcentaje de abundancia relativa de los diferentes grupos de acidobacterias.

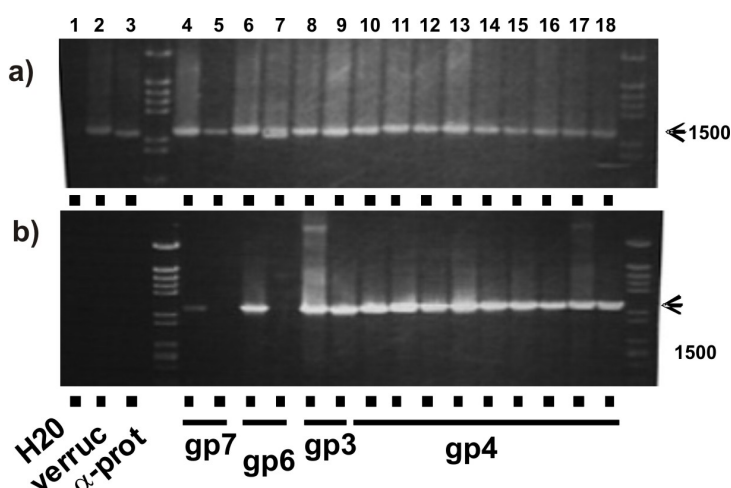
## 14. ANÁLISIS DE LA COMUNIDAD DE ACIDOBACTERIAS PRESENTE EN LA RIZOSFERA DE *Mamilaria carnea* Y EN SUELO SUELTO

Como ya se ha descrito en el apartado anterior, la población microbiana predominante en este tipo de suelos pertenece al phylum *Acidobacteria* y, además, se encuentra en un elevado número tanto en época lluviosa como en época seca, aunque predomina en época lluviosa y existen diferencias estadísticamente significativas entre estas dos épocas.

Con el fin de profundizar en la ecología de este phylum bacteriano tan desconocido y además poder observar si la población de acidobacterias varía entre rizosfera y suelo suelto como ya se ha descrito en otros trabajos (Lauber *et al*, 2008; Jones *et al*, 2009; Quaiser *et al*, 2003), se planteó el diseño de genotecas específicas de acidobacterias usando como molde el ADN extraído de la rizosfera de la cactácea globosa *Mamilaria carnea* y el ADN procedente de una muestra de suelo no rizosférico, ambas muestras tomadas durante la época lluviosa.

### 14.1. CONSTRUCCIÓN DE LAS GENOTECAS ESPECÍFICAS DEL PHYLUM ACIDOBACTERIA

Los cebadores escogidos fueron el primer universal 1492r y el 31f descrito en la literatura como específico para acidobacterias (Burns *et al*, 1999). Antes de construir las genotecas y como estudios preliminares para observar si esta pareja de primers podía amplificar específicamente el gen 16s del phylum *Acidobacteria*, se realizaron PCRs con ADN de diferentes clones secuenciados obtenidos en las genotecas 16S de los que ya se conocía su adscripción filogenética (figura I.17).

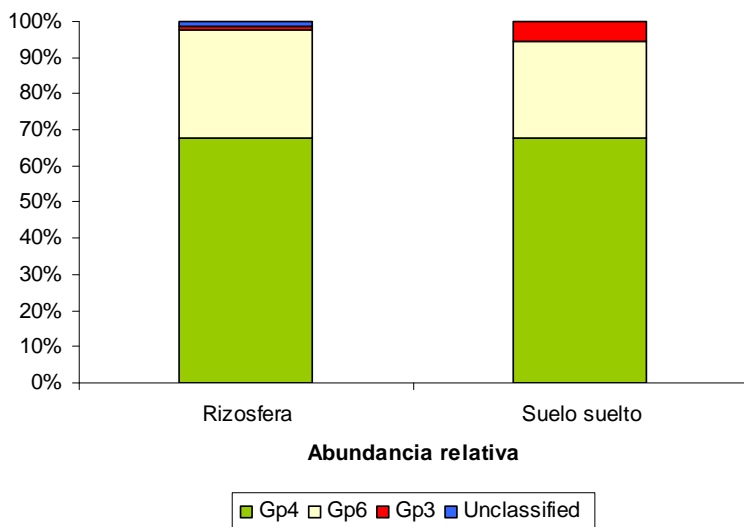


**Figura I.16.** a) Productos de PCR resultantes de la amplificación con los primers 9f y 1492r, que amplifican un amplio espectro de phyla bacterianos. b) Productos de PCR resultantes de la amplificación con los primers 1492r y 31f específicos del phylum *Acidobacteria*.

Una vez comprobado las condiciones de PCR en las cuales los cebadores eran específicos, se procedió a la construcción de las genotecas según se detalla en el material y métodos de esta Tesis Doctoral. De estas genotecas se secuenciaron 153 clones, 71 pertenecientes a suelo suelto y 82 procedentes a la rizosfera.

#### 14.2. ANÁLISIS DE LAS SECUENCIAS

Al igual que en el caso anterior en la genoteca construida con los primers generales, la clasificación taxonómica se llevó a cabo mediante el método Naïve Bayesian, introduciendo las secuencias en la página web del *Ribosomal Data Base* y enfrentándola a su base de datos específica de genes ribosomales 16S. Los resultados de la adscripción filogenética se observan en el siguiente gráfico, en el que predomina el género GP4 tanto en suelo suelto como en rizosfera. Se aprecia, además, un número ligeramente más elevado del grupo GP3 en rizosfera que en suelo suelto.

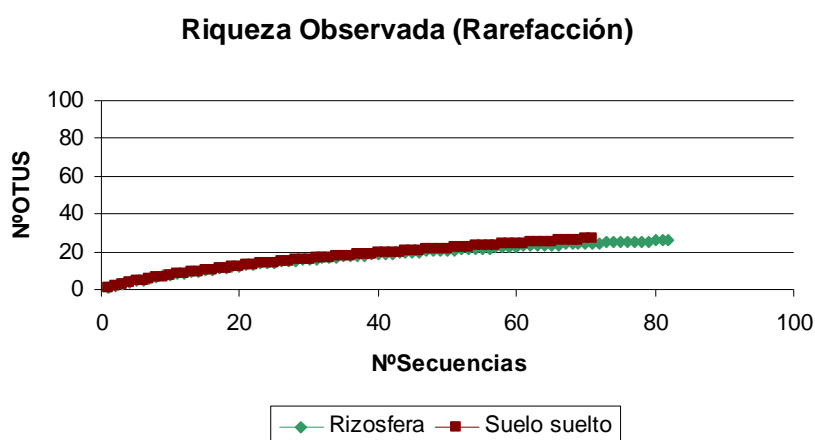


**Figura I.17.** Grupos de acidobacterias encontrados en rizosfera y suelo suelto.

#### 14.3. RIQUEZA OBSERVADA

Para estimar la riqueza observada, se realizaron las curvas de esfuerzo o rarefacción al 0,03% de distancia para cada una de las muestras analizadas en suelo suelto y rizosfera. De este modo, con 71 secuencias en el caso del

suelo suelto, establece 26 OTUs, y en la rizosfera, con 81 secuencias, establece 26 OTUs, lo cual indicaría que la rizosfera es ligeramente menos rica que el suelo suelto en cuanto a acidobacterias se refiere.



**Figura I.18.** Curvas de rarefacción al para el phylum *Acidobacteria* en rizosfera y suelo suelto 0,03%.

#### 14.4 ÍNDICES DE EQUIDAD Y RIQUEZA ESPECÍFICA

Con respecto a la riqueza específica (índice de Chao), se estima mayor número de OTUs al 0,03% en suelo suelto que en rizosfera (como ya se apuntaba en las curvas de rarefacción). Sin embargo, la diversidad no varía entre ambos tipos de muestras (índice de Piloni,  $J'$ , es 0,87 en ambos casos)

	Riqueza Observada	Riqueza Específica	Índice de Shannon-Wiener ( $H'$ )	$H'_{max}$	$J'$
<b>Suelo suelto</b>	27	56	2.85	3.29	0.87
<b>Rizosfera</b>	26	37	2.86	3.26	0.87

**Tabla I.15** Índices de riqueza y diversidad calculados a partir de las secuencias de acidobacterias obtenidas en rizosfera y suelo suelto.

#### 14. 4 COMPARATIVA DE LAS LIBRERÍAS

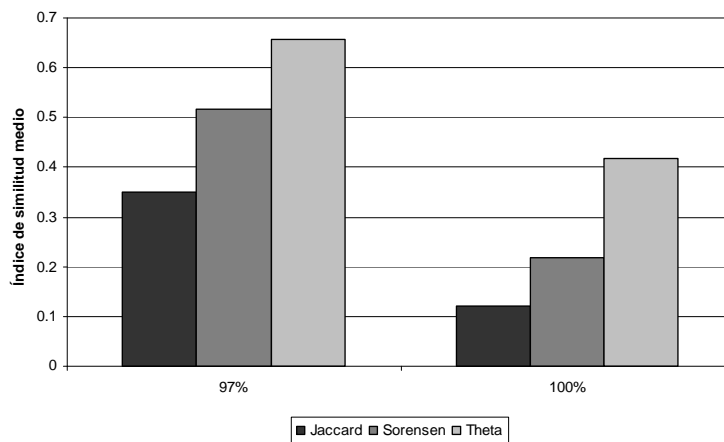
Se usó la herramienta del rdp *Lib compare* para comparar ambas librerías y no obtuvimos diferencias estadísticamente significativas en ninguno de los casos. Además, también se usó la herramienta informática  $\zeta$ -*Libshuff* y tampoco se encontraron diferencias estadísticamente significativas.



	Rizosfera	Suelo suelto
Rizosfera	0	0,5993
Suelo suelto	0,4286	0

**Tabla I.16.** p-valor calculado usando la herramienta estadística  $\zeta$ -Libshuff para ambas genotecas.

También se determinaron los índices de Jaccard, Theta y Sørensen. La rizosfera y el suelo suelto, a un 97%, comparten un 0,35 (índice de Jaccard) de los OTUs y, si tenemos en cuenta la abundancia relativa y no solo los OTUs diferentes, observamos como la similitud aumenta (índice Theta 0,65).



**Figura I.19.** Índices de Jaccard, Sørensen y Theta para el phylum *Acidobacteria* en rizosfera a diferentes distancias moleculares 0,03 y 0%.

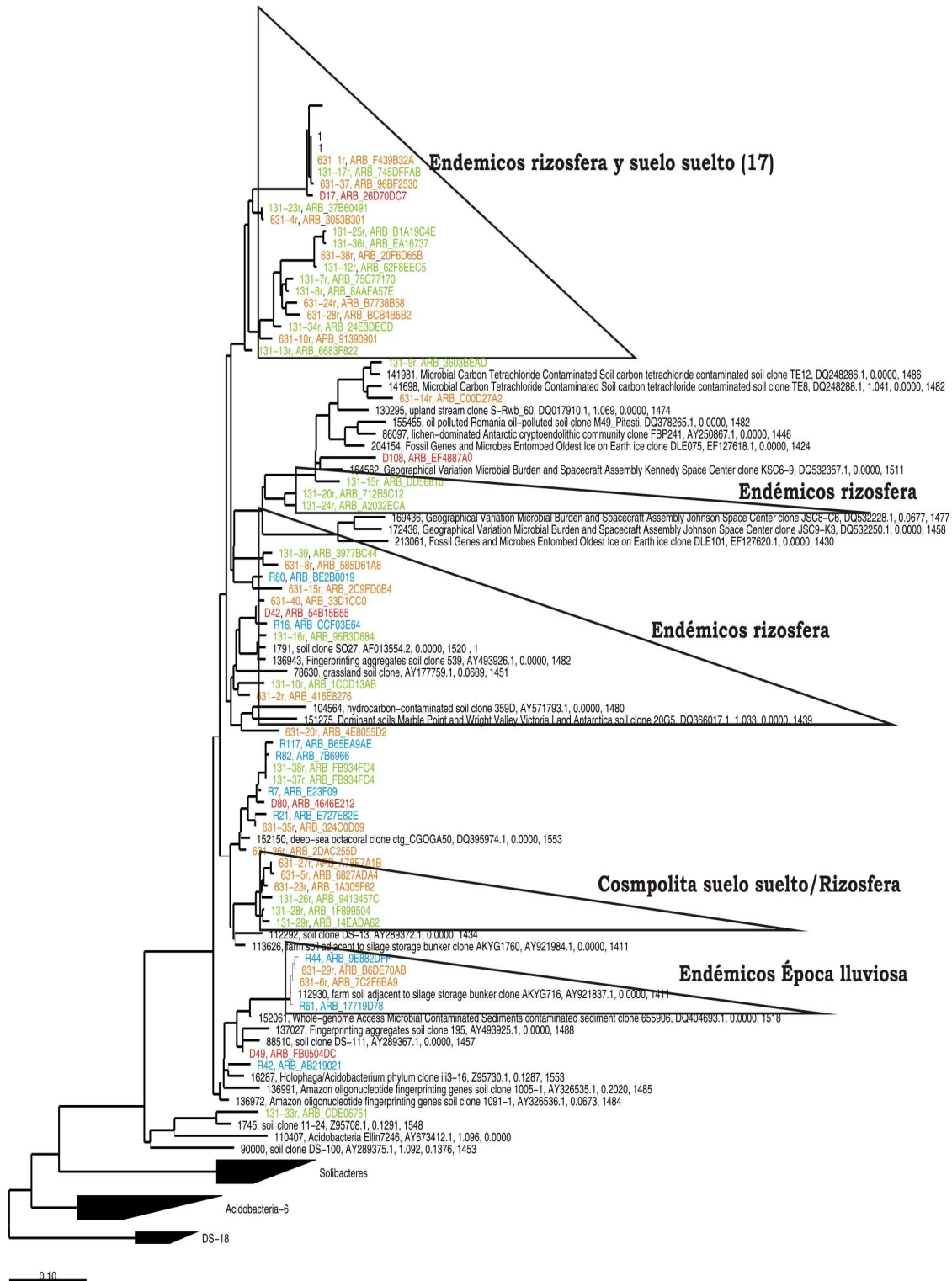
En el siguiente gráfico se muestran los diagramas de Venn con los OTUs compartidos al 0,03% entre rizosfera y suelo suelto. Ambos ecosistemas comparten 15 OTUs y tienen como endémicos 14 OTUs que solo existen o en rizosfera o en suelo suelto.



**Figura I.20.** Diagrama de Venn con OTUs al 0,03% de distancia entre rizosfera y suelo suelto.

De este modo, observamos cómo la población microbiana mayoritaria de estos suelos son las acidobacterias, pero no encontramos diferencias estadísticamente significativas entre la población de acidobacterias presente en suelo suelto y la población presente en rizosfera. En ambos casos abundan los grupos GP4 y GP6, aunque en suelo suelto también aparece el grupo GP3 y es más diverso que la rizosfera.

Para observar las relaciones filogenéticas entre todas las acidobacterias de este estudio, se construyó un árbol filogenético con el programa *ARB* (Ludwig et al, 2004,), en el que se enfrentaron las secuencias de acidobacterias procedentes de las genotecas 16S de época seca y lluviosa y de rizosfera y suelo suelto a una base de datos con 150400 secuencias en las que se seleccionaron específicamente después del alineamiento las secuencias de acidobacterias. De este modo, se observan como las secuencias de acidobacterias pertenecientes a este estudio se agrupan juntas, lo que indica que son microorganismos realmente novedosos y no descritos anteriormente. Además, también se forman grupos endémicos de suelo suelto y rizosfera.



**Figura I.21.** Árbol filogenético en el que se enfrentan las secuencias de acidobacterias procedentes de este estudio con las depositadas en las bases de datos del ARB (<http://www.arb-home.de/>). En rojo se señala las secuencias procedentes de época seca y en azul las procedentes de época lluviosa. De la misma forma, en verde se señala las secuencias procedentes de la rizosfera y en naranja las muestras procedentes de suelo suelto.

## DISCUSIÓN

Los estudios preliminares desarrollados en este capítulo, han permitido focalizar nuestro estudio en las muestras rizosféricas pertenecientes a la cactácea *M. carnea*.

De este modo, se han llevado a cabo diferentes análisis para determinar la diversidad procariota existente en la rizosfera de esta cactácea, en época seca y lluviosa. Hemos observado que el phylum mayoritario en las muestras es el de acidobacterias y que la población microbiana de este grupo varía en función de la época; fluctuación que se observa también, por ejemplo en los phyla *Cyanobacteria*, *Firmicutes* y *Plantomycetos* que son más abundantes en época seca, o en el phylum *Gemmamonadetes* que por el contrario, aumenta en época lluviosa.

También se ha estudiado en profundidad el phylum *Acidobacteria* en rizosfera y suelo suelto y se ha observado, que la población mayoritaria es la del grupo GP4 en ambos tipos de muestras, aunque en suelo suelto aparece también el grupo GP3 y es más diverso en general que el suelo asociado a raíz.

**Clasificación preliminar del muestreo mediante sus perfiles genéticos (TGGE).** De las aproximaciones preliminares para el estudio de la diversidad procariota (capítulo I.1), observamos que se extrae más material biológico, en forma de ADN, en las muestras de suelo rizosférico que en muestras de suelo suelto.

Estos resultados coinciden en parte con los estudios llevados a cabo en la Universidad de México por el licenciado Felix Aguirre sobre microorganismos cultivables en estas mismas muestras, en las que se observan diferencias estadísticamente significativas en el número de células viables en rizosfera y suelo suelto en época lluviosa. Y están en concordancia con otros estudios que plantean el efecto de la rizosfera sobre las poblaciones microbianas, análisis que apuntan a que en la rizosfera la biomasa y la actividad microbiológica es elevada debido a los compuestos exudados por la raíz (Raaijmakers *et al*, 2009). De este modo, la rizosfera induciría un cambio en la población microbiana del suelo, seleccionando unos individuos y aumentando la biomasa microbiana, aumento que se puede observar de forma indirecta mediante las extracciones de ADN de suelo.

Los estudios desarrollados mediante el análisis poblacional del 16S ADNr de  $\alpha$ -proteobacterias usando los perfiles genéticos obtenidos gracias a la

técnica de TGGE que se describen en este capítulo, permitieron demostrar, entre otros datos, que no había diferencias entre las muestras de suelo íntimamente pegado a la raíz y suelo rizosférico. Este resultado además permitió concluir que el perfil poblacional de las  $\alpha$ -proteobacterias era estable en distintas extracciones (capítulo 1 de Material y Métodos).

Además estos estudios, nos permitieron enfatizar en las diferencias entre las poblaciones microbianas de la rizosfera y del suelo suelto ya que los análisis de similitud (UPGMA) diferenciaban ambas comunidades; análisis que apoyan lo que se ha observado de manera indirecta mediante las extracciones de ADN de suelo: la comunidad microbiana parecía cambiar en número bajo la influencia de la raíz de la planta.

Los estudios de TGGE sobre esta población ( $\alpha$ -proteobacterias) también nos permitieron observar las diferencias entre las diferentes zonas, agrupando de forma diferencial a las tres zonas de muestro: Valerio Trujano 1 y 2 y Tomellín. Estas tres zonas tienen características edafológicas claramente diferenciadas, y como ya hemos comentado en la Introducción de esta Tesis Doctoral, el suelo es uno de los determinantes principales en la estructura de las comunidades debido a los micronichos que se crean en él (Fierer et al, 2006). En este caso, parece ser incluso un factor determinante más decisivo que la rizosfera ;en tanto que el uso de esta técnica no permitió diferenciar las poblaciones microbianas de diferentes plantas en una misma región, o diferenciar claramente las poblaciones microbianas en época seca y lluviosa.

Sin embargo, a pesar de que la población de  $\alpha$ -proteobacterias no pareciera regirse por una planta o estación de muestreo particular, decidimos abordar un estudio poblacional más amplio mediante la generación de genotecas globales comparando la rizosfera de una de estas plantas: *M. carnea* en dos épocas extremas (seca y lluviosa) y determinar la influencia de la presencia de la planta en la modulación de las poblaciones procarióticas.

## **2. LA DIVERSIDAD PROCARIÓTICA EN LA RIZOSFERA DE *Mammillaria carnea*.**

Para el análisis de las fluctuaciones estacionales en la diversidad y la estructura de la comunidad microbiana en la rizosfera de *M. carnea*, se determinó la secuencia de amplicones utilizando cebadores universales para la subunidad pequeña del ribosoma.

Se analizó la misma muestra mediante dos aproximaciones experimentales complementarias:

I) Por un lado, se construyeron genotecas construidas en *E. coli* del 16S ARNr utilizando los cebadores 9f y 1541r (Reysenbach y Pace 1995; McInnerty *et al.*, 1995), y se analizaron 103 fragmentos de 700 bs aproximadamente. El tamaño de los fragmentos nos permitía tener una adscripción filogenética con mejor nivel de significación, pero el reducido número de secuencias (103), nos restringía el entendimiento de la diversidad y estructura de la comunidad. Además, las regiones del gen ribosomal 16S que se estudiaron comprendían las zonas variables V8-V9.

II) Las mismas muestras se sometieron a un protocolo de secuenciación masiva (454 FLX Titanium) de amplicones del gen 16S ARNr. En este caso los cebadores utilizados se correspondieron con G1R/G1F y G2R/G2F (tabla AI.1) que comprenden las regiones V4 y V5 para la asignación filogenética. El tamaño de lectura medio consistió en 400 nt de información que limitaba la identificación de los microorganismos de referencia en comparación con el método anterior. Sin embargo, el número de secuencias obtenidas se incrementó en más de dos órdenes de magnitud, permitiendo poder analizar la estructura y diversidad de nuestro ecosistema de estudio.

### **- Análisis de las secuencias definición de filotipos**

El número de secuencias que puedes estudiar mediante la construcción de genotecas 16S y su posterior secuenciación con el método Sanger, limita el número de datos entre 50 y 500

Con la aparición de los métodos de secuenciación masiva, el rendimiento y número de datos es muy superior (al menos dos ordenes de magnitud; tabla I.17) Aunque el tamaño de las secuencias sea más pequeño (400 nt en la actualidad mediante 454 Titanium de Roche), en muchos abordajes es suficiente para establecer adscripciones filogenéticas que

determinen diferencias en las muestras (Jones T.R et al, 2009; Schauer et al, 2010.).

	<b>Número Secuencias</b>		<b>Tamaño medio de lecturas (nt)</b>	
	<b>Época seca</b>	<b>Época lluviosa</b>	<b>Época seca</b>	<b>Época lluviosa</b>
<b>454</b>	15410	13083	409	409
<b>Sanger</b>	47	56	650	650

**Tabla I.17.** Secuencias y tamaño de las mismas obtenidas a partir de un proceso tradicional de generación de genotecas 16S seguido de secuenciación tipo Sanger y a partir de un proceso de pirosecuenciación.

El ecosistema que estudiamos, el suelo, es muy diverso y como nos indican las curvas de esfuerzo o rarefacción, difícil de cubrir mediante el número de secuencias tan limitado que se obtiene con una aproximación Sanger. El Abordaje mediante 454, sí que te permite poder tener un número de secuencias más idóneo para cubrir la diversidad procariota presente en las muestras de suelo (figura I.11, Castro H.F, 2009).

Comparando ambos tipos de aproximaciones, podemos constatar que el grupo mayoritario presente en la rizosfera de *M.carnea* es el phylum *Acidobacteria* y el número de microorganismos en este grupo aumenta en época lluviosa en las dos aproximaciones (tabla I.18). Las proteobacterias, que son otro de los phyla mayoritarios, también están representadas en ambos tipos de abordaje.

Sin embargo, la metodología 454 te permite observar cambios en grupos minoritarios que con las secuencias procedentes de Sanger son muy difíciles de observar. Como ejemplo, se pueden citar las cianobacterias, en época lluviosa mediante el método Sanger no encontramos ninguna secuencia y con la tecnología 454 tuvimos secuencias suficientes (250) como para hacer un estudio más pormenorizado (figura I.14).

	454		Sanger	
	Época seca Número (%)	Época lluviosa Número (%)	Época seca Número (%)	Época lluviosa Número (%)
<b>Acidobacteria</b>	3291(23.2)	3912(32.9)	15(31.9)	25(44.6)
<b>Unclassified bacteria*</b>	2971(20.9)	2340(19.7)	/	/
<b>Actinobacteria</b>	2719(19.1)	1798(15.1)	2(4.3)	2(3.6)
<b>Proteobacteria</b>	1769(12.5)	1543(13)	7(14.9)	9(46.1)
<b>Bacteroidetes</b>	1178(8.3)	921(7.7)	12(25.5)	11(19.6)
<b>Cianobacteria</b>	956(6.7)	229(1.9)	1(2.1)	
<b>Gemmatimonadetes</b>	334(2.4)	205(4.2)	/	/
<b>Plantomycetes</b>	313(2.2)	194(1.6)	3(6.4)	2(3.6)
<b>Firmicutes</b>	232(1.6)	126(1.1)	/	
<b>Verrumicrobia</b>	191(1.3)	119(1)	5(10.6)	4(7.1)
<b>Chloroflexi</b>	153(1.1)	137(1.2)	1(2.1)	2(3.6)
<b>Nitrospira</b>	76(0.5)	48(0.4)	/	/
<b>Deinococcus</b>	10(0.1)	/	/	/
<b>OP10</b>	/	/	1(2.1)	
<b>TM7</b>	/	/		1(1.8)

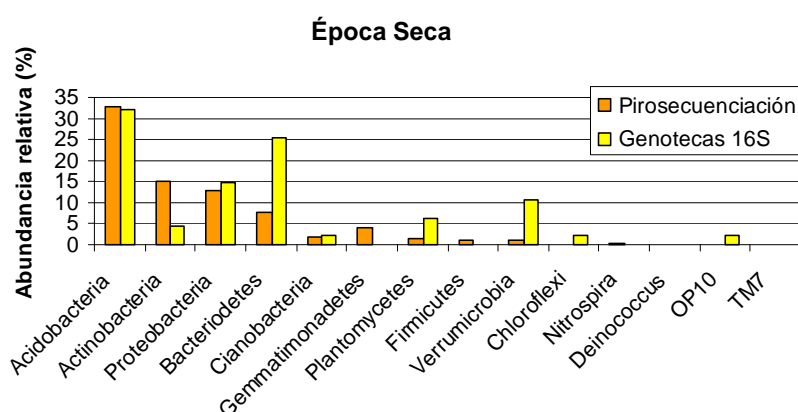
**Tabla I.18.** Filotipos descritos en las secuencias procedentes de la tecnología 454 y de Sanger. La clasificación se realizó en la base de datos del RDP a un nivel de significación del 80%.

Gracias a los datos de pirosecuenciación el grupo actinobacterias afloró como el segundo más importante en cuanto a número se refiere, estando así subrepresentado en las secuencias obtenidas mediante el abordaje tipo Sanger. El caso contrario ocurre con los bacteroidetes y verrumicrobia que están sobrerrepresentados con este último abordaje tipo Sanger (figura I.22 A y B).

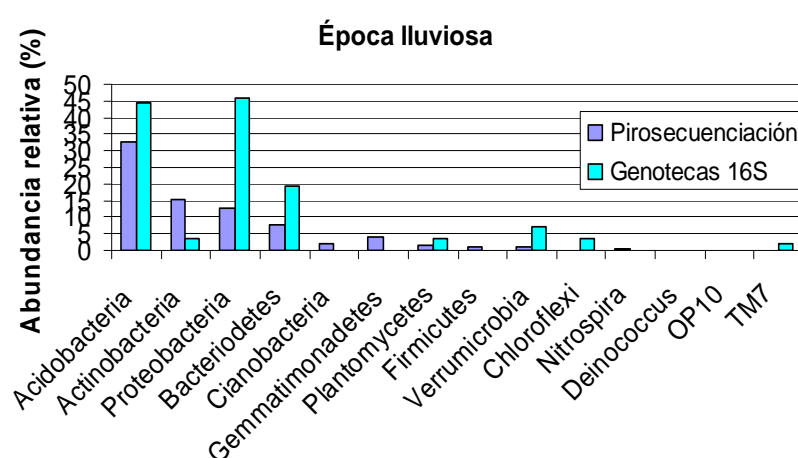
Salvando estas diferencias, si observamos la figura I.22, nos daremos cuenta que el perfil de los diferentes phyla que hay en nuestras muestras de suelo rizosférico no varía mucho usando ambos tipos de metodologías, si bien, la pirosecuenciación te permite hacer un estudio más detallado con respecto a los phyla minoritarios



A)



B)



**Figura I.22.** Clasificación taxonómica en phyla de las secuencias obtenidas en época seca (A) y en época lluviosa (B) mediante pirosecuenciación y el método tradicional de secuenciación tipo Sanger

**- Diversidad y Riqueza.**

Si comparamos los resultados de los índices de diversidad ( $H'$  y  $J'$ ) usando las secuencias obtenidas a partir de las dos aproximaciones (tabla I.19), nos damos cuenta que en el caso de la aproximación tipo Sanger se sobreestima la diversidad. La explicación a este hecho es simple, se debe al error de muestreo que representa tener tan pocas secuencias.

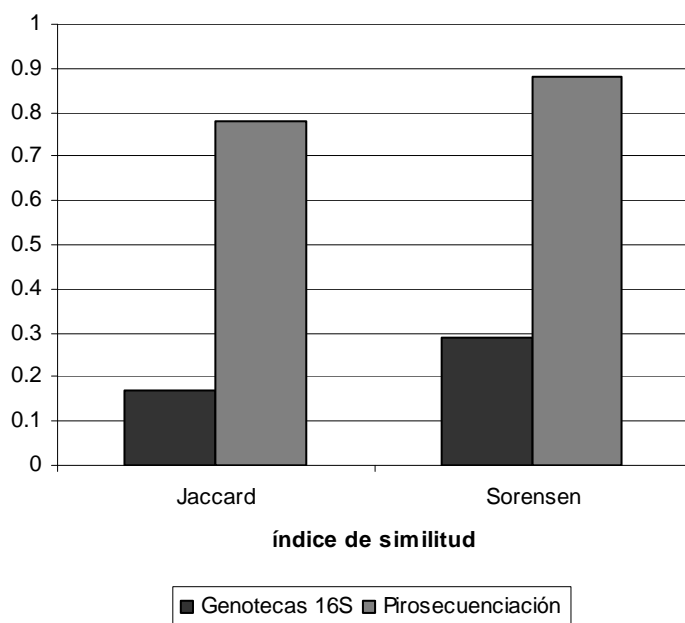
Algo similar ocurre con la riqueza específica que se sobreestima debido al número reducido de secuencias obtenido mediante el abordaje tipo Sanger (de 41 OTUs en época seca por ejemplo el índice de Chao estima 207 OTUs, cinco veces más, y en el caso de los datos del 454 de 3033.97 se estima 5277.66, menos del doble)

		Riqueza específica	Riqueza Observad.	H'	H' max	J'
<b>Sanger</b>	Época seca	207	41	3.64	3.71	0.98
	Época lluviosa	77	38	3.47	3.64	0.95
<b>454</b>	Época seca	5277.66	3033.97	6.92	8.02	0.86
	Época lluviosa	3688	2347.04	6.65	7.76	0.85

**Tabla I.19.** Índices de diversidad y riqueza específica al 0.03%, en las secuencias obtenidas a partir de aproximación Sanger y 454.

**- Índices de similitud.**

Si nos fijamos en los índices de similitud de Jaccard y Sørensen al 0.03% (figura I.25), podemos inferir que en el caso de el abordaje tipo Sanger, estos índices están subestimados (valores de 0.17 y 0.29 para Jaccard y Sørensen respectivamente), lo que nos indicaría que las muestras seleccionadas tendrían poca similitud y muchos OTUs endémicos de cada zona; mientras que si observamos los índices que se obtienen a partir de las muestras obtenidas mediante la pirosecuenciación a este mismo nivel de significación, vemos como la similitud aumenta considerablemente (0.78 y 0.88 para Jaccard y Sørensen respectivamente), concluyendo así que un número reducido de secuencias (<100) en estudios de suelo genera un error de muestreo (al igual que ocurría con los índices de diversidad) que provoca una subestimación del índice de similitud.



**Figura I.23.** Índices de similitud de Jaccard y Sørensen al 0.03% calculado a partir de las secuencias obtenidas mediante 454 y Sanger.

Para poder encontrar diferencias significativas entre unos ecosistemas y otros usando la herramienta  $\zeta$ -*Libshuff* o bien el lib *compare* del *rdp*, entre unos tiempo de muestreo y otros, es fundamental tener un número de secuencias lo suficientemente alto que te permita hacer comparaciones.

En nuestro caso, con las 103 secuencias el abordaje tipo Sanger, no encontrábamos diferencias estadísticamente significativas entre época seca y lluviosa, mientras que con las 28493 secuencias de la aproximación 454, si podemos establecer diferencias estadísticamente significativas entre ambas épocas y además se pueden señalar tanto los phyla bacterianos causantes de estas diferencias como los que permanecen invariables en el tiempo, los que no fluctúan.

De este modo, gracias a la aproximación 454 se puede tener un número muchísimo más elevado de secuencias del gen 16S, que te permita hacer estudios de diversidad microbiana sin sobreestimarla, estudiar los phyla minoritarios y hacer comparaciones entre diferentes épocas sin cometer un error de muestreo.

La principal limitación de este tipo de aproximación frente a la tradicional de la elaboración de las genotecas 16S ya no es el coste económico (que ya está casi solventado con el abaratamiento de los precios) es determinar si el tamaño de las secuencias, que es de 400 pb (en la actualidad) nos permite un nivel de discriminación suficiente para entender las fluctuaciones en el ecosistema.

#### **-Phyla mayoritarios en la rizosfera de *M. carnea*.**

Una de las consecuencia del aumento en número de lecturas, es que te permite hacer un estudio pormenorizado de los distintos phyla presentes: además, de las *Acidobacterias* los otros phyla mayoritarios en la rizosfera de *M. carnea* son: *Actinobacterias*, *Proteobacterias* y *Bacteroidetes*; phyla que ya se han descrito como abundantes en suelos de otro tipo de estudios (Janssen et al, 2006; Da rocha et al , 2009). Todos estos phyla, además, se manifiestan como mayoritarios en los dos tipos de abordajes metodológicos usados, salvo el caso de las *Actinobacterias* que en los estudios basados en las genotecas 16S aparecían muy poco representadas.

Al contrario que ocurría con las acidobacterias, éstos phyla no sufren variaciones estadísticamente significativas relacionadas con las épocas seca y lluviosa.

En el caso de los *Bacteroidetes*, existen evidencias en otros trabajos que apuntan a que pueden estar poco representados en abordajes de este tipo por el sesgo de la amplificación por PCR (Cottrell et al, 2000; Eilers, H et al, 2000); aún así es uno de los grupos mayoritarios en nuestras muestras, y dentro de éste predomina la clase *Sphingobacteria*. Algunos miembros de este grupo son aerobios, mientras que otros son anaerobios o aerobios facultativos, y la composición de especies de esta clase en un suelo depende en parte de los niveles de oxígeno en el mismo (Janssen et al, 2006).

#### **-Phyla minoritarios en la rizosfera de *M. carnea*.**

Además de las acidobacterias, existen otros phyla, aunque ya no mayoritarios, en los que sí se observan diferencias estadísticamente significativas entre época seca y lluviosa (tabla I.11). Estos son *Cyanobacterias*, *Firmicutes* y *Plantomycetes* que se observa que aumentan en época seca en los dos tipos de metodologías usadas (con la salvedad de los firmicutes que en las genotecas 16S no aparecen) y *Gemmamonadetes* que aumentan en época lluviosa.

Dentro del phylum *Firmicutes*, tanto en época seca como en lluviosa predomina el género *Clostridia*, compuesto por bacterias anaerobias formadoras de esporas. Es posible de esta forma apuntar que, durante la época seca los clostridios se encuentren en forma de esporas para poder sobrevivir a las condiciones ambientales. Aunque en otros trabajos (Castro et al, 2010), describen que la abundancia relativa tanto de firmicutes como de *Plantomicetos* aumenta con la humedad, al contrario que lo que observa en nuestras muestras.

Con respecto a los *Plantomycetes*, la mayoría de los aislados provienen de ecosistemas acuáticos y no se consideran buenos modelos ni genética ni fisiológicamente para extrapolar a los plantomicetos procedentes del suelo. Los aislados que se han conseguido de este grupo en suelo, incluyen algunos miembros del linaje WPS-1 (Davis et al, 2005; Josehp et al, 2003; Wang et al, 2002), pero no son suficientes para poder cubrir toda la diversidad que

sugieren los estudios de las secuencias del gen ribosomal 16S en suelos, ni para poder inferir su posible papel en el suelo.

El phylum *Cyanobacteria* por su parte, es casi cinco veces más numeroso durante la época seca. Si observamos la distribución de las secuencias de este grupo (tabla I.13), podremos observar que la mayoría de ellas no se pueden clasificar al nivel del 80 %. Si consideramos otro nivel de significación más bajo, como del 50 % por ejemplo, hay algunas secuencias que si las clasifica en otras familias como la familia V, IX ó XIII; pero las familias Streptophyta y familia I que es donde se encuentran las mayores diferencias entre épocas apenas varían.

La familia Streptophyta está compuesta principalmente por genes 16S rARN de cloroplastos de plantas y de microorganismos no cultivables. En este trabajo, se observa como las secuencias que guardan homología con esta familia aumentan 100 veces en época lluviosa. Si se elabora un árbol filogenético con todas estas secuencias (figura I.14), observamos como la mayoría de las secuencias de la época lluviosa son muy parecidas entre sí y a su vez se diferencian de las secuencias de época seca, indicando que existiría una explosión clonal en época lluviosa de un tipo de microorganismo que sería el que dominaría esta clase de Cianobacterias durante esta época.

En el caso de la familia I, pasa algo parecido (figura I.15). Su número aumenta considerablemente en época seca. En esta familia ya sí podemos encontrar algunos microorganismos cultivables como *Anabaena* o *Nodularia*. Y al igual que en el caso anterior, en el árbol filogenético vemos como en época seca estaría ocurriendo una explosión clonal de un determinado tipo de microorganismo que sería el que dominaría esta clase de cianobacterias hasta la época lluviosa.

En las cianobacterias estaríamos observando dos ejemplos claros de modulación por el ambiente, en el que se produce una proliferación de los microorganismos adaptados a cada ambiente. El hecho de que se produzcan en direcciones opuestas, con indiferencia respecto al número global de cada grupo, da validez a los resultados obtenidos.

Para finalizar, el grupo de las gimmatimonadetas (*Gemmatimonadetes*) aumenta en los suelos de estudio, en época lluviosa.

Se han descrito que constituyen un porcentaje del 2 % de las comunidades microbianas de suelo (Janssen et al, 2006). Contiene una sola especie, *Gemmatimonas aurantiacus* (Zhang et al, 2003), que es gram-

negativa, heterótrofa y aeróbica; fue aislada de un tratamiento de aguas residuales y pertenece a la subdivisión 1, también conocida como la clase *Gemmatimonadetes*.

Del suelo se han obtenido cuatro aislados (Davis et al, 2005; Josehp et al, 2003), que pertenecen también a la subdivisión 1 y presentan un fenotipo heterótrofo aeróbico. En nuestro estudio, este phylum aumenta en época lluviosa, este efecto se ha descrito en otro tipo de trabajos (Castro et al, 2009), pero dado lo poco que se conoce de la ecología de este grupo bacteriano, no se puede dilucidar qué papel podría estar teniendo las gimmatimonadetas en el suelo durante la época lluviosa.

### **3. ANÁLISIS DE LA DIVERSIDAD Y ESTRUCTURA DEL Phylum ACIDOBACTERIA EN LA RIZOSFERA DE *Mammillaria carnea***

Como se ha comentado en el apartado anterior, los datos aportados mediante sanger y 454 coinciden en que el phylum mayoritario en la rizosfera de *M carnea* en ambas estaciones (seca y húmeda) es el **phylum *Acidobacteria***. Hasta la fecha, ningún estudio había relacionado este grupo microbiano con la rizosfera de cactáceas en un suelo con pH neutro; los análisis que existen las relacionan fundamentalmente con pHs ácidos, describiendo este phylum como mayoritario en suelos con estas características (Lauber et al, 2009).

Se conoce muy poco de la ecología de este phylum bacteriano debido a que es muy difícil de cultivar. Casi todos los aspectos recogidos en la literatura respecto a él, provienen de datos de estudios sobre el gen 16S rARN (Jones et al, 2009; Sang-Hoon Lee et al, 2008; Quaiser et al, 2003; Lauber et al, 2009; Navarro-Noya et al, 2009).

Sin embargo, es un un phylum importante, abundante y diverso en ecosistemas de suelos; si consideramos su diversidad filogenética y patrón de distribución ecológica, el grupo Acidobacteria es un grupo comparable genética y metabólicamente al phylum, ampliamente estudiado, de las *Proteobacterias* (Hugenholtz et al, 1998; Barns te tal, 1999). Ya se ha observado mayoritario en diferentes tipos de suelos, como por ejemplo en suelos contaminados con metales pesados en México, en la rizosfera de plantas colonizadoras (Navarro-Noya et al, 2010).

Hay estudios que apuntan a que es un grupo de bacterias oligotróficas (Fierer et al, 2007), que se encuentran en mayor número en suelos maduros

(Tarlera et al, 2008) y que cuya abundancia puede estar regulada por el pH (Fierer et al). Hay pocos trabajos sin embargo, que relacionen la estructura y diversidad de la población de acidobacterias con la disponibilidad de agua.

En un estudio (Castro et al, 2009) donde se analiza la respuesta de la comunidad microbiana frente a diferentes cambios ambientales relacionados con el cambio climático (cambios en las condiciones de dióxido de carbono, de temperatura, de humedad...), se observa que la abundancia de la población de acidobacterias es mayor en ambientes secos que en húmedos. Los autores justifican esta observación, con el hecho de que este grupo microbiano parece que prefiere habitar nichos oligotróficos.

Nuestros resultados, sin embargo, apuntan justo lo contrario. Los dos conjuntos de datos, indican que la abundancia de acidobacterias es mayor en época lluviosa que en época seca. Indicando en este caso, una posible correlación positiva de la abundancia de acidobacterias con la humedad.

Además, el grupo mayoritario dentro de las acidobacterias, es el **grupo GP4**. Este grupo predomina independientemente de las condiciones que se estén dando (época seca o lluviosa) y es un grupo asociado con suelos con un pH neutro, como son los suelos objetos de estudio. Además, también hay estudios que relacionan los suelos con pH neutros con una elevada diversidad microbiana (Lauber et al, 2009).

Esta abundancia en rizosfera, nos sugirió la posibilidad de una relación específica con *M. carnea*, por lo que se estudio la composición de este phylum en particular en una nueva comparativa: suelo suelto y rizosfera, puesto que en otros estudios se ha observado que existen diferencias significativas, en la distribución, abundancia y composición en las comunidades rizosféricas y de suelo suelto de acidobacterias.

Existen datos contradictorios, algunos estudios apuntan a que predominan en suelo suelto (Kielah, 2009; Fierer et al, 2007; Sanguin et al, 2006) y otros las asocian claramente a suelo rizosférico; como por ejemplo, se ha descrito como phylum dominante en la rizosfera del castaño en suelos ácidos (Lee et al, 2008) o en la rizosfera del pino (Chow et al, 2002) Stafford et al, 2005; Singh et al, 2007).

Además, la variabilidad dentro del phylum *Acidobacterias* es muy elevada, mucho más que otros phylum procariontas (Lauber et al, 2009). En nuestro estudio, partíamos de la hipótesis de que la población microbiana de acidobacterias era abundante y dominante tanto en suelo suelto como en

rizosfera y queríamos observar si existían patrones de microdiversidad dentro de este grupo en función del ambiente; esto es, si la variabilidad dentro del grupo acidobacterias varía en función del suelo suelto o rizosfera. Para ello se diseñaron primers específicos de este grupo bacteriano y se generaron genotecas 16S.

De este modo, se observó que el grupo **GP4** era el grupo mayoritario tanto en rizosfera como en suelo suelto; sin embargo, la diversidad aumentó ligeramente en el caso del suelo suelto, aumentando el grupo GP3 que es un grupo que está presente tanto en ambientes ácidos como neutros (Jones et al, 2009).

También se concluye que de los 43 OTUs totales a un nivel de significación del 0.03 % que existen en estas genotecas, solo se comparten 15 OTUs, estableciéndose de este modo una gran proporción de OTUs endémicos en suelo suelto y rizosfera (figura I.20-I.21), aunque debido a la escasa información filogenética que existe del phylum acidobacterias no se puede establecer que tipo de OTUs serían endémicos de rizosfera y cuales serían endémicos de suelo suelto.





## **ANEXO CAPÍTULO I**

## Apéndice Capítulo 1

Nombre	Secuencia	Aplicación	Ref.
<b>aU175</b>	5' -CCGCATACGCCCTACGGGGGAAAGATTTAT-3'	TGGE, $\alpha$ -proteob.	(Gomes et al, 2001)
<b>1512Ur</b>	5' -CTACGGHTACCTTGTACGAC-3'	TGGE $\alpha$ -proteob.	(Smalla et al, 2001)
<b>F984GC</b>	5' -CGCCCGGGCGCGCCCCGGGCGGGGCGGGGCGGGGACGCGAAGAACCTTAC-3'	TGGE, cola GC	(Heuer et al, 2007)
<b>R1378</b>	5' -CGGTGTGTACAAGGCCCGGAACG-3'	TGGE	(Heuer et al, 2007)
<b>U1510R</b>	5' -GGTTACCTTGTACGACTT-3'	Universal, genotecas 16S	(Reysenbach and Pace 1995)
<b>E9F</b>	5' -GAGTTTGATCCTGGCTCAG-3'	Universal, genotecas 16S	McInnerty et al, 1995
<b>31F</b>	5' - GATCCTGGCTCAGAATC-3'	Acidobacterias	Barns SM et al 1999
<b>G1F</b>	5' -CGTATCGCCTCCCTCGCGCCATCAGAAGCCAGTCAGCMGCCGCGTAATWCT-3'	Pirosecuenciación (seca)	En este trabajo (Baker et al, 2003)
<b>G1R</b>	5' -CTATGCGCCTTGCCAGCCCGCTCAGAAGCCAGTCCGTCAATTCCTTTRAGTTT-3'	Pirosecuenciación (seca)	En este trabajo Baker et al, 2003)
<b>G2F</b>	5' -CGTATCGCCTCCCTCGCGCCATCAGTTGAACTGCAGCMGCCGCGTAATWC-3'	Pirosecuenciación (lluviosa)	En este trabajo Baker et al, 2003)
G2R	5' - CTATGCGCCTTGCCAGCCCGCTCAGTTGAACTGCCGTCAATTCCTTTRAGTTT-3'	Pirosecuenciación (lluviosa)	En este trabajo Baker et al, 2003)

Tabla IA.1. Cebadores usados en este capítulo.

**II METAGENOMICA FUNCIONAL APLICADA A LA  
RIZOSFERA DE LA CACTÁCEA GLOBOSA**

*Mammillaria carnea*



## II BÚSQUEDA DE ENZIMAS DE INTERÉS INDUSTRIAL Y DE BIOMARCADORES EN LA RIZOSFERA DE *MAMMILLARIA CARNEA*

La diversidad procariota presente en nuestras muestras que revela el capítulo anterior indica que estos suelos son potencialmente un reservorio de nuevos genes, así como de proteínas inéditas con un posible interés industrial o biomarcadores.

Con este fin, se construyeron genotecas medioambientales en fagos con un doble enfoque: por un lado, poder caracterizar el resistoma en las muestras objeto de análisis y por otro, realizar una búsqueda de nuevas enzimas con carácter lipolítico que pudieran tener una potencial aplicación en la industria.

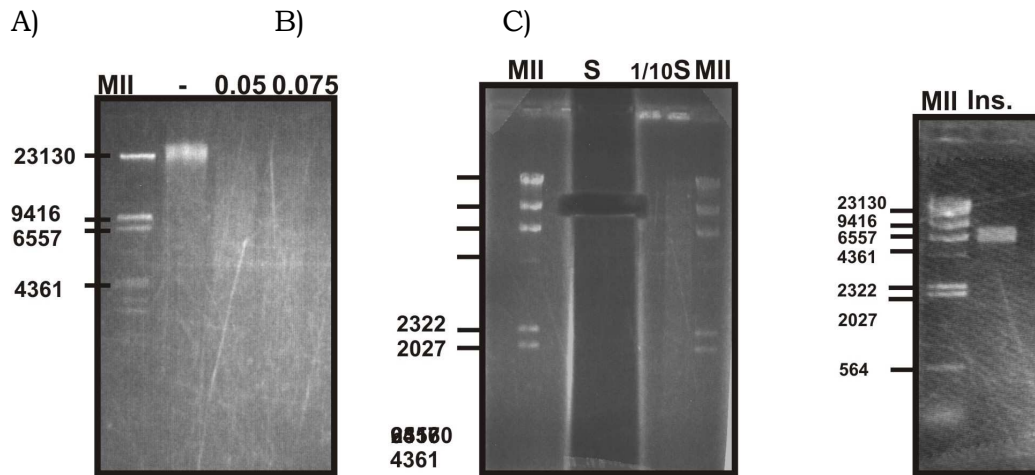
### III.1. CONSTRUCCIÓN DE GENOTECAS METAGENÓMICAS EN FAGOS

Se construyeron tres genotecas, dos pertenecientes a la rizosfera de *Mammillaria carnea* en época seca y lluviosa (denominadas genotecas lluviosa y seca) y otra a la que nosotros llamamos “muestra control” (denominada genoteca EEZ), perteneciente a un suelo de una parcela próxima a nuestro centro de trabajo y que a diferencia de las dos anteriores presenta una fuerte influencia antropogénica (historial de cultivos, ...).

Así, para la construcción de las genotecas metagenómicas, el ADN medioambiental de estos tres ecosistemas se digirió parcialmente con la enzima de restricción *Sau3A*. De este modo, se aisló un fragmento del gel de entre 9,5-10 kb de ADN a una concentración de unos 60 ng/ $\mu$ l. El siguiente paso implicó la ligación con el genoma del fago lambda a 4°C y el posterior empaquetamiento (introducción del ADN en una partícula viral), tal y como se detalla en el capítulo 11 del Material y Métodos de esta Tesis Doctoral.

Obtuvimos, de esta forma, una genoteca de  $6,45 \cdot 10^5$  pfu en el caso de la llamada EEZ (ver tabla II.1.1). El tamaño medio de inserto de la genoteca fue de unas 6,5 kb.

En el caso de las otras dos genotecas pertenecientes a la rizosfera de *Mammillaria carenea* en época seca y lluviosa, el rendimiento en la construcción de las genotecas fue mayor y se obtuvieron en ambos casos aproximadamente 250.000 clones.



**Figura II.1.1.** Construcción de las genotecas de expresión en fagos. A) Gel de agarosa al 0,8%, donde se observan las digestiones parciales con la enzima Sau3A, carril con ADN sin digerir (-), digerido con 0,05 U de enzima (0,05) y con 0,075 U (0,075) B) Gel de agarosa de bajo punto de fusión en el que se ha cortado una banda de entre 9 y 10 Kb. S es el carril donde se encontraba la muestra y 1/10S una dilución 1/10 de la muestra C) Inserto de aproximadamente 9 Kb que se obtiene tras cortar la banda y que se pondrá a ligar con el genoma del fago lambda.

Nombre de la genoteca	Nº clones	Tamaño medio de inserto (Kb)	Cantidad de ADN clonado (Gigabases)	Mb/Clon de resistencia
EEZ	65000	6,5	0,42	52,5
Lluviosa-mex	250000	7	1,7	850
Seca-mex	254000	7	1,5	1500

**Tabla II.1.1.** Características de las genotecas medioambientales construidas.

## II2. EL SUELO, UN RESERVORIO POTENCIAL DE GENES DE RESISTENCIA A ANTIBIÓTICOS NO ESTUDIADOS

### II2.1. Búsqueda de genes de resistencia a antibióticos en las genotecas metagenómicas

Con las genotecas metagenómicas así construidas, se realizó una escisión en masa (capítulo 11.5.2 de Material y Métodos) y se obtuvieron colonias de células XLOR de *Escherichia coli*. Cada una portaba el fagémido pCK-CMV (figura 4 de Material y Métodos,) con un inserto de ADN metagenómico.

La escisión de fago a fagémido de nuestras genotecas metagenómicas nos permitió hacer una búsqueda de genes de resistencia usando diferentes sustancias que inhibían el crecimiento bacteriano. De este modo, empleamos varios tipos de antibióticos con distintas dianas en la célula, como por ejemplo los antibióticos estreptomicina y gentamicina, que son aminoglicósidos: eritromicina, que afecta a la traducción de proteínas; ampicilina, que afecta a la pared celular, y trimetroprima y cloranfenicol, que inhiben rutas metabólicas esenciales para la supervivencia de la célula.

Se partió de una alícuota de aproximadamente  $1.10^6$  células de cada genoteca, que se extendió en placas de medio LB con kanamicina (que es el antibiótico que permite seleccionar el fagémido pBK-CMV; figura 4 de Material y Métodos). A cada placa se le añadieron diferentes concentraciones letales de cada antibiótico: de ampicilina, estreptomicina, cloroanfenicol y trimetroprima se usó una concentración de 50 mg/L, en el caso de la gentamicina se usaron 20 mg/L, y por último usamos eritromicina a una concentración de 60 mg/L.

No se encontró ninguna célula resistente a estreptomicina. En el caso de la eritromicina, en una primera búsqueda encontramos numerosas células resistentes a este antibiótico que, tras diferentes análisis posteriores, resultaron ser falsos positivos.

Con respecto al resto, encontramos, de forma global, 11 genes de resistencia a los antibióticos usados (tabla II.2). De estos, 7 pertenecen a la genoteca construida a partir de ADN de la Estación Experimental del Zaidín, que es mucho más pequeña en número que las genotecas construidas a partir de suelo del parque natural de México (tabla II.2). La alta incidencia de estos genes en esta genoteca, podría indicar una relación entre este tipo de marcadores y suelos más o menos perturbados por el hombre.



Nombre	Genoteca	Tamaño inserto (Kb)	Antibiótico usado para la búsqueda
<b>Ap6-8</b>	Mex Seca	6,9	Ampicilina
<b>Gm8-3</b>	Mex Lluviosa	3	Gentamicina
<b>Tm8-4</b>	Mex Lluviosa	7.5	Trimetroprima
<b>GmE1</b>	EEZ	7.5	Gentamicina
<b>ApE1</b>	EEZ	7	Ampicilina
<b>ApE14</b>	EEZ	10	Ampicilina
<b>CmE1</b>	EEZ	9	Cloranfenicol
<b>CmE6</b>	EEZ	9	Cloranfenicol
<b>TmE1</b>	EEZ	4.2	Trimetroprima
<b>TmE5</b>	EEZ	6,5	Trimetroprima
<b>TmE25</b>	EEZ	5,85	Trimetroprima

**Tabla II.2.** Clones de resistencia a antibióticos encontrados en las genotecas metagenómicas construidas.

### II2.2 Medida de la concentración mínima inhibitoria de los genes de resistencia encontrados

Se midió la concentración mínima inhibitoria (MIC) de los clones encontrados para los antibióticos usados en la selección (tabla II2.2).

En general, los clones encontrados en las genotecas construidas a partir de ADN de suelo del parque natural de Tehuacán-Cuicatlán presentan más resistencia que los clones aislados de la genoteca de la Estación Experimental del Zaidín.

Por ejemplo, si nos fijamos en los genes de resistencia a ampicilina, Ap6-8 es el doble de veces de resistente que los genes ApE1 y ApE14 encontrados en la genoteca de la EEZ. Algo parecido ocurre, asimismo, con los genes de resistencia a gentamicina: El clon Gm8-3 es 2,5 veces más resistente a gentamicina que el clon GmE1.

Clon	Genoteca	Concentración mínima inhibitoria (µg/ml)			
		Ampicilina	Gentamicina	Trimetroprima	Cloranfenicol
pBK-CMV <sup>b</sup>		<10	2,5	12,5	<10
Ap6-8	Seco-MEX	100 (10x) <sup>c</sup>			
Gm8-3	Lluviosa-MEX		<50 (20x) <sup>c</sup>		
Tm8.4	Lluviosa-MEX			>150 (12x) <sup>c</sup>	
GmE1	EEZ		>20 (8x) <sup>c</sup>		
ApE1	EEZ	>50(5x) <sup>c</sup>			
ApE14	EEZ	>50(5x) <sup>c</sup>			
CmE1	EEZ				>50(5x) <sup>c</sup>
CmE6	EEZ				>50(5x) <sup>c</sup>
TmE1	EEZ			>50(4x) <sup>c</sup>	
TmE5	EEZ			>50(5x) <sup>c</sup>	
TmE25	EEZ			>50(5x) <sup>c</sup>	

**Tabla II.3.** Valores de MIC (*Minimum Inhibitory Concentration*) de los clones de resistencia frente a los antibióticos de estudio

<sup>a</sup> Las medidas se hicieron en la cepa hospedadora procedente de la genoteca metagenómica (*E.coli* cepa XLOLR) crecida en medio LB con los antibióticos indicados. La tabla muestra la media de tres experimentos independientes.

<sup>b</sup> Vector desprovisto de inserto.

<sup>c</sup> El número en paréntesis indica el número de veces que el clon es más resistente que el vector control sin inserto.

### II2.3.Análisis de los genes de resistencia a antibióticos encontrados

El rango del tamaño de los insertos que portaban los vectores, cuyo fenotipo resultaba en resistencia a antibióticos, varía entre 4,2 a 10 kb (tabla II.2).

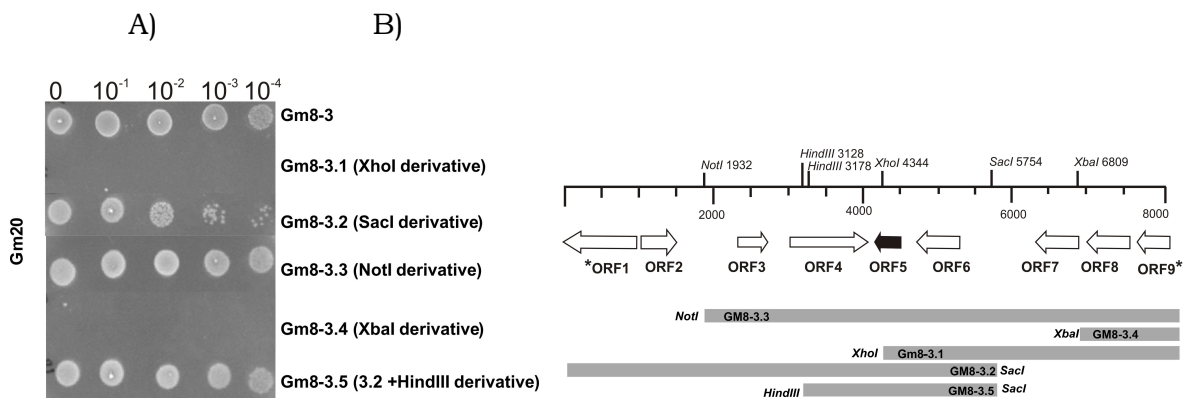
Para identificar dentro de los diferentes insertos los marcos abiertos de lectura que estaban causando el carácter de la resistencia, se generaron diferentes subclones mediante digestiones con distintas enzimas de restricción y se testaba la resistencia o no al antibiótico, reduciendo así la información genética a secuenciar necesaria para tener el marco abierto de lectura del gen de resistencia.

Un ejemplo de la metodología usada para la secuenciación de los marcos abiertos de lectura de los genes de resistencia a antibióticos se observa en la figura II.2 (para el clon de resistencia a gentamicina Gm8-3 (8 kb)).

Generando diferentes subclones mediante enzimas de restricción, observamos que el fenotipo se mantiene en los subclones Gm8-3.2, Gm8-3.3 y Gm8-3.5. Este último subclon posee un tamaño considerablemente menor al

original, 3 kbs, pudiendo obtener la secuencia fácilmente mediante la técnica de “primer walking”.

De todas formas, en este caso concreto del clon Gm8-3, finalmente se secuenciaron las 8 kb que portaba el inserto y se observó que, efectivamente, el marco abierto de lectura que contenía el gen de resistencia se encontraba en el subclon Gm8-3.5 y se correspondía con el denominado ORF 5 (marco abierto de lectura, de las siglas en inglés: *Open Reading Frame*), que codifica para una aminoglicósido 3´N-acetiltransferasa.



**Figura II.2.** Características del clon de resistencia a gentamicina Gm8-3. A) Ensayos de resistencia para el clon Gm8-3 y los diferentes subclones en una placa con medio LB y 20 mg/L de gentamicina. Para el ensayo se usaron diferentes gotas con diluciones serias desde 0 a 10<sup>-4</sup>. B) Organización esquemática de los diferentes ORFs identificados en el inserto del clon Gm8-3. Las flechas indican la localización y la dirección de la transcripción en los ORFs. El ORF implicado en la resistencia a gentamicina se marca en negro. Las barras indican la longitud de los subclones, generados a partir de digestiones con las enzimas de restricción que se indican en el dibujo.

En los siguientes apartados, desarrollaremos las características de cada uno de los tipos de resistencia encontrados frente a cada antibiótico:

*II2.3.1. Análisis de los clones de resistencia a ampicilina*

Los tres clones que confieren resistencia a ampicilina contienen insertos de tamaño de 6,5, 6,9 y 10 kb. Como ya se ha comentado, se realizaron diferentes subclones para determinar qué ORFs eran los que determinaban el patrón de resistencia y se observó cómo, en los tres casos, los ORFs que conferían resistencia a ampicilina codifican para enzimas beta-lactamasas.

El estudio de la secuencia de estos genes reveló que las beta-lactamasas correspondientes a los clones ApE6-8 y ApE1 pertenecían a la familia A de beta-lactamasas, mientras que en el caso del clon ApE14, codifica una enzima que pertenece a la familia C (figura II.3). La clasificación molecular de las familias beta-

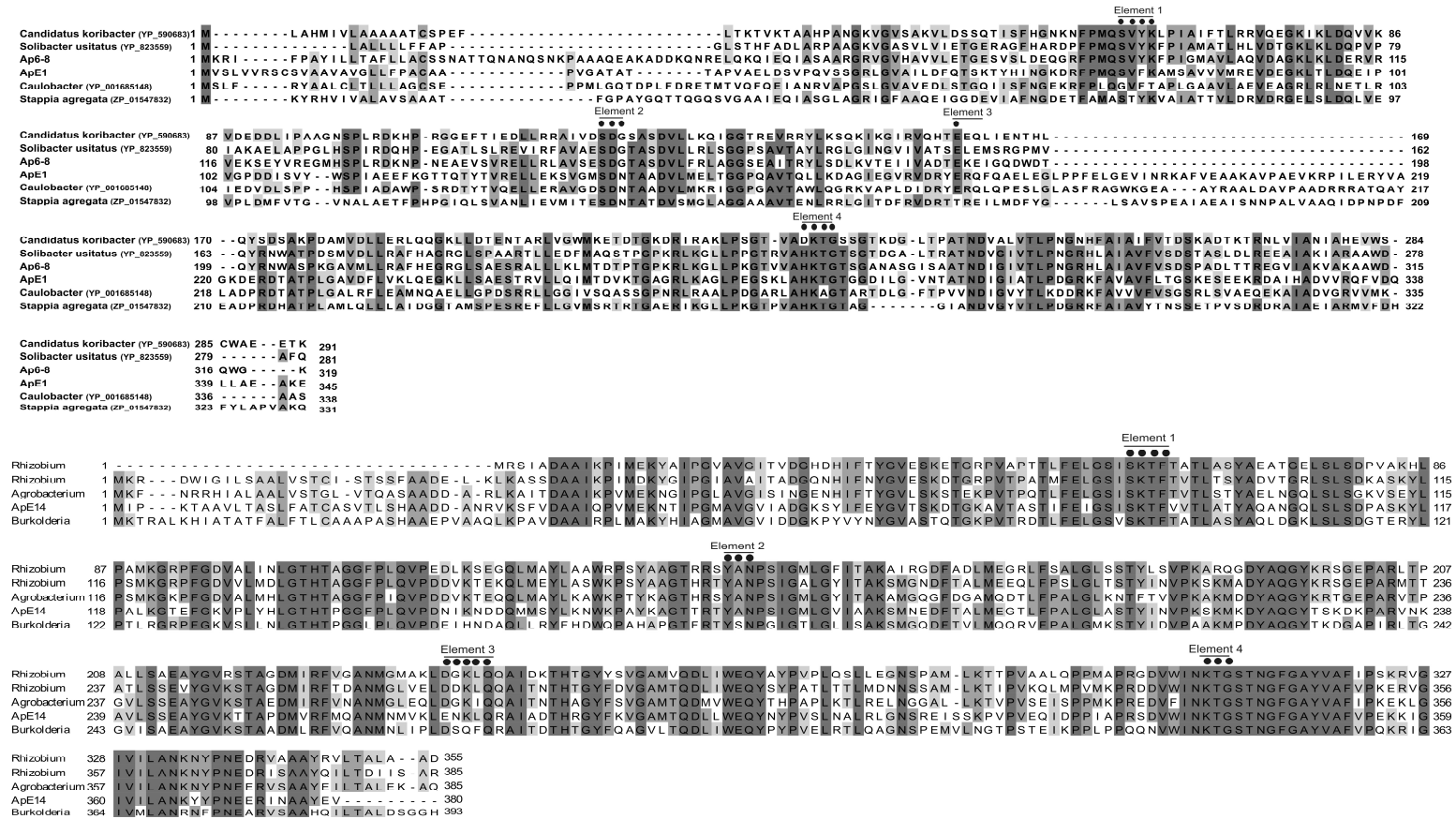
lactámicas se realiza en función de la secuencia aminoacídica y nucleotídica y teniendo en cuenta diferentes características fenotípicas (Ambler, R.P. 1980).

Las clases A, C y D actúan con un mecanismo basado en una serina, mientras que las enzimas de la clase B, denominadas metalo-beta-lactamasas, necesitan del ión zinc para su activación.

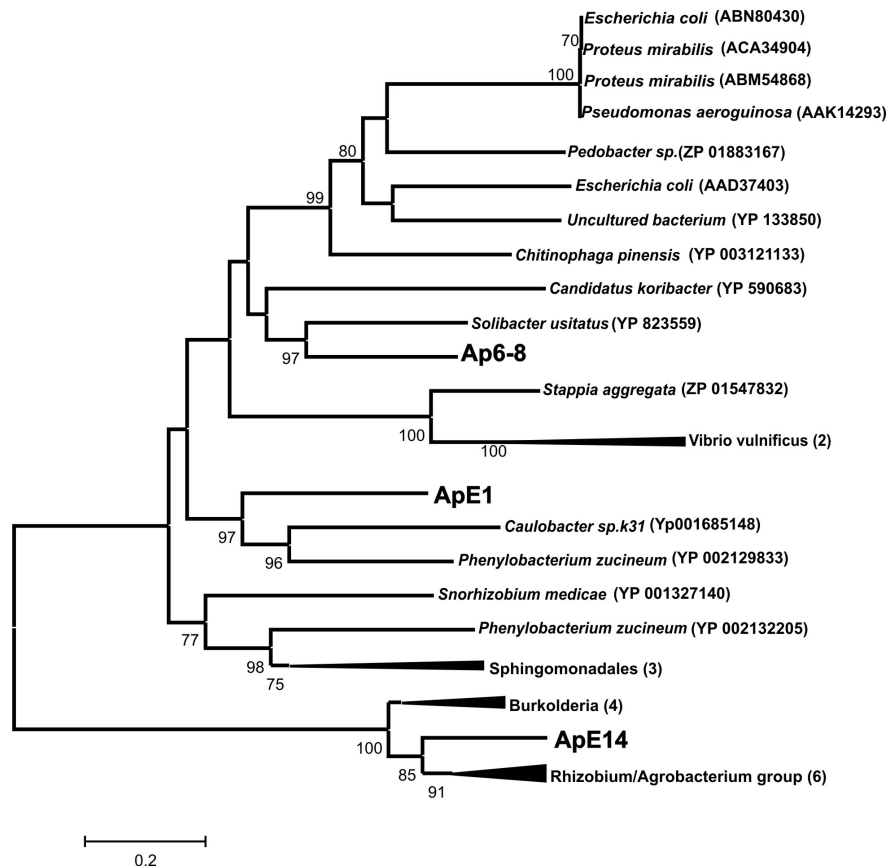
Las tres enzimas difieren significativamente de otras beta-lactamasas descritas en las bases de datos (el porcentaje de identidad es del 57, 45 y 63 para Ap6-8, ApE1 y ApE14 respectivamente). Los análisis filogenéticos (figura II.4) revelan que las enzimas beta-lactamasas de los clones encontrados en la genoteca EEZ (ApE1 y ApE14) están muy relacionadas con otros genes de beta-lactamasas encontrados en el cromosoma de diferentes estirpes de los géneros *Phenylobacterium* y *Rhizobium* respectivamente.

Sin embargo, en el caso del clon perteneciente a la genoteca de suelo de México (Ap6-8), se encuentra relacionado con beta-lactamasas encontradas en bacterias del género *Solibacter sp.* (tabla AII.1 y figura II.3). Este género, *Solibacter* pertenece al phylum *acidobacteria*, el cual ya hemos visto a través de otro tipo de técnicas que es el más abundante en este tipo de suelos (capítulo I de Resultados).

## Capítulo II-Resistoma



**Figura II.3. Secuencias, alineamientos e identificación de los residuos conservados en las beta-lactamasas ApE1, ApE12 y Ap6-9 encontradas en las genotecas metagenómicas.** A) Alineamiento múltiple de las secuencias de los clones Ap6-8 y ApE1 con otras secuencias aminoacídicas íntimamente relacionadas identificadas en la base de datos del NCBI usando el servidor BLAST. La mayor parte de estas proteínas forman parte de la clase A de las beta-lactamasas **B)** Alineamiento múltiple de la secuencia del clon ApE14 con otras beta-lactamasas de la clase C (los elementos del 1 al 4, marcados con negrita, son los residuos conservados descritos para las beta-lactamasas; [Joris B et al, 1991](#)). Los alineamientos de la secuencias de proteínas se realizaron con la herramienta online: <http://align.bmr.kyushu-u.ac.jp/mafft/software/>.

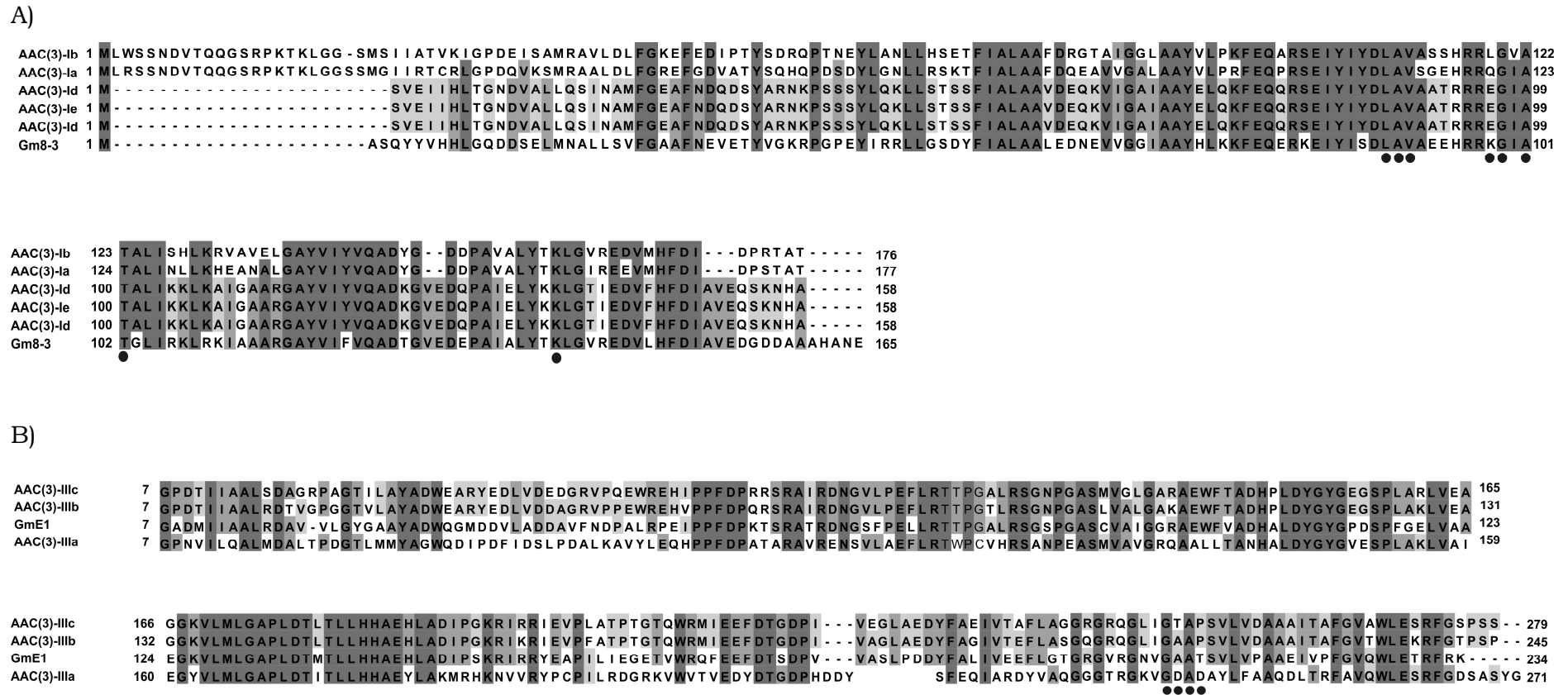


**Figura II.4. Árbol construido con el método del vecino más próximo usando 1000 iteraciones, que muestra las relaciones filogenéticas de las proteínas de resistencia a ampicilina.** En este árbol consenso, en los nodos se representan los valores de porcentaje superiores al 70%. Los nombres de las diferentes especies van seguidos de su número de acceso en el NCBI. Algunas de las ramas del árbol han sido colapsadas (entre paréntesis se indica el número de secuencias agrupadas) para mejorar la visibilidad del árbol. Las secuencias de este estudio se indican en negrita.

### II.2.3.2. Análisis de los clones de resistencia a gentamicina

Se encontraron dos genes de resistencia a gentamicina, uno en la genoteca de la EEZ (GmE1) y otro en la genoteca construida a partir de ADN del parque natural de México en época lluviosa (Gm8-3).

Tras estudiar el mapa de restricción (figura II.2), se observó que en ambos clones la resistencia a gentamicina se daba debido a una enzima 3´N-acetiltransferasa (AAC(3); Azucena E. *et al*, 2001) codificada a partir del denominado ORF5 en el caso del clon Gm8-3 (figura II.2) y ORF2 en el caso del clon GmE1 (tabla anexo AII.2.1).

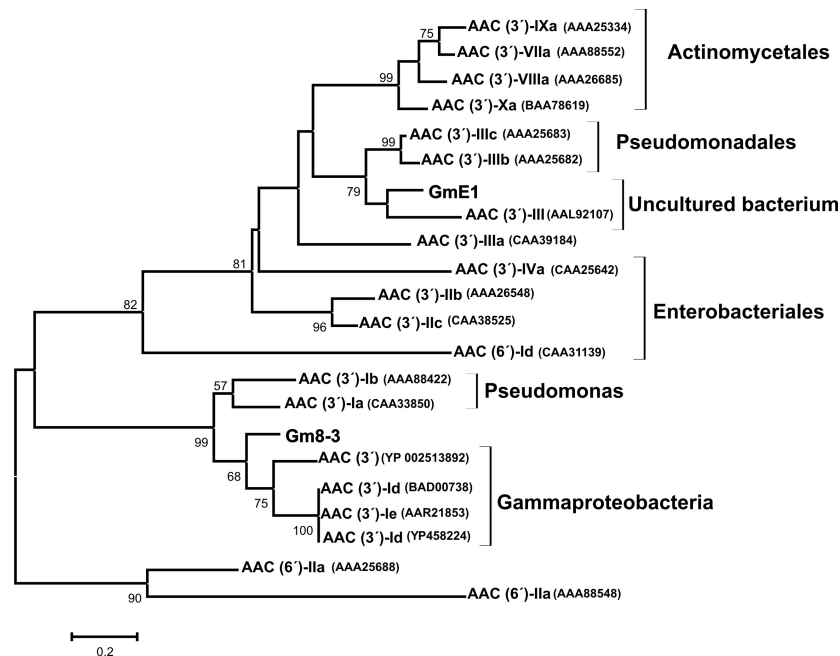


**Figura II.5. Alineamiento múltiple de las dos enzimas aisladas de las genotecas metagenómicas que confieren resistencia a gentamicina. A)** La enzima Gm8.3 está alineada con proteínas de la familia AAC(3)- I [AAC(3)-Ia, AAC(3)-Ib, AAC(3)-Id y AAC(3)-Ie] (Vakulenko *et al*, 2003) **B)** La enzima GmE1 se encuentra alineada con proteínas de la familia AAC(3)-III. Los alineamientos se han realizado en la página web: <http://align.bmr.kyushu-u.ac.jp/mafft/software/>. La figura muestra con un círculo, en el caso de Gm8-3, los aminoácidos que catalizan el traspaso de un grupo acetil desde el acetil-coenzima A hasta el aminoglicósido (Dyda F *et al*, 2005). En el caso del GmE1, no parece contener estos dominios.

Ambas enzimas contienen el motivo “GNAT” característico de las enzimas relacionadas con GCN5 (Dyda F *et al*, 2000; figura II.5). Los aminoácidos implicados en transferir el grupo acetil desde el acetil-coenzima A hasta la amina primaria del antibiótico aminoglicósido están claramente identificados en el caso de la enzima del clon Gm8-3, pero no se han podido identificar en el caso de la enzima del clon GmE1.

Estas enzimas tienen menos de un 69% de homología con las secuencias aminoacídicas de otras enzimas presentes en las base de datos (tabla AII.2.1).

Como se observa en las figura II.5-II.6, la enzima GmE1 está relacionada con las enzimas AAC(3)-III. Se encuentra próxima a otras enzimas procedentes de diferentes especies de *Pseudomonas* que muestran similar perfil de resistencia. Del mismo modo, se puede apreciar que está claramente diferenciada del grupo de genes codificados por *Actionomicetos*. Los análisis filogenéticos (figura II.6) sugieren que podría ser un nuevo grupo dentro de la familia de las enzimas AAC(3)-III. Presenta, además, una elevada homología con otro clon aislado de una genoteca metagenómica de suelo que confiere resistencia a kanamicina (tabla anexo AII.2.1; Shaw *et al*, 1993). La proteína Gm8-3 está relacionada, sin embargo, con otro grupo diferente de acetiltransferasas, como son las del grupo I (AAC(3)-I; Riccio, M.L *et al*, 2003; Ahmed, A.M *et al*, 2004).

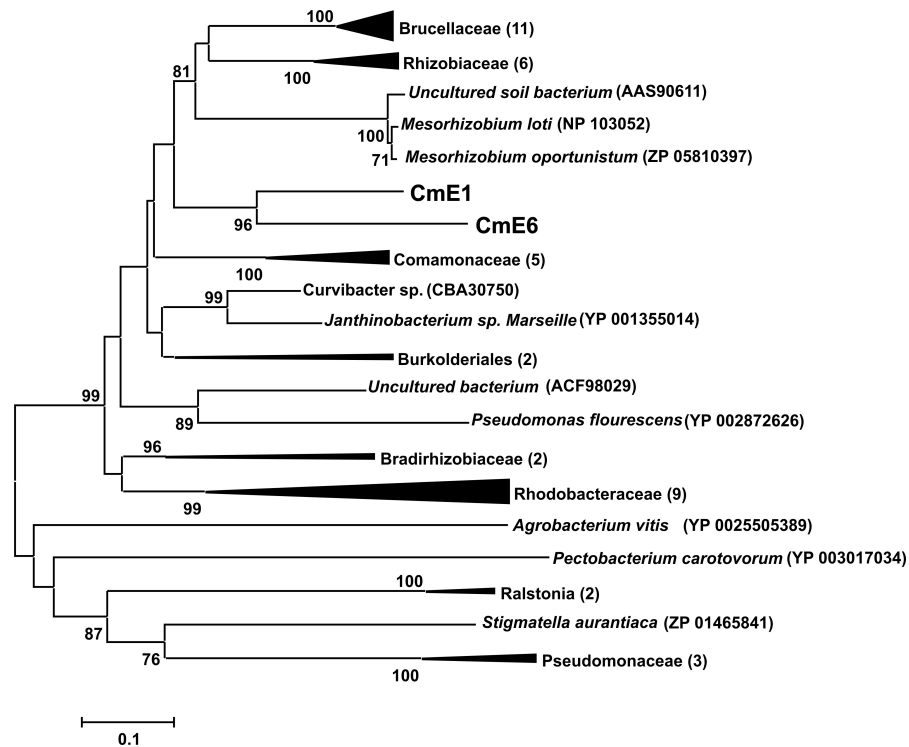


**Figura II.6. Árbol filogenético en el que se infiere las posibles afiliaciones filogenéticas de los genes de resistencia a gentamicina encontrados en este estudio.** Las abreviaturas de los genes y las definiciones de los diferentes clados se muestran acorde con el trabajo de Shaw *et al*, 1993. Se presenta un árbol consenso, tras 1000 iteraciones, en cuyos nodos sólo se muestran los valores superiores al 70%. Los genes de la familia AAC6' se han usado como grupo diferencial. Los nuevos genes aislados de nuestro estudio se presentan en negrita.



II2.3.3. Análisis de los clones de resistencia a cloranfenicol

Los dos genes que confieren resistencia a cloranfenicol han sido aislados únicamente de la genoteca construida a partir de ADN de la Estación Experimental del Zaidín (CmE1 y CmE6).



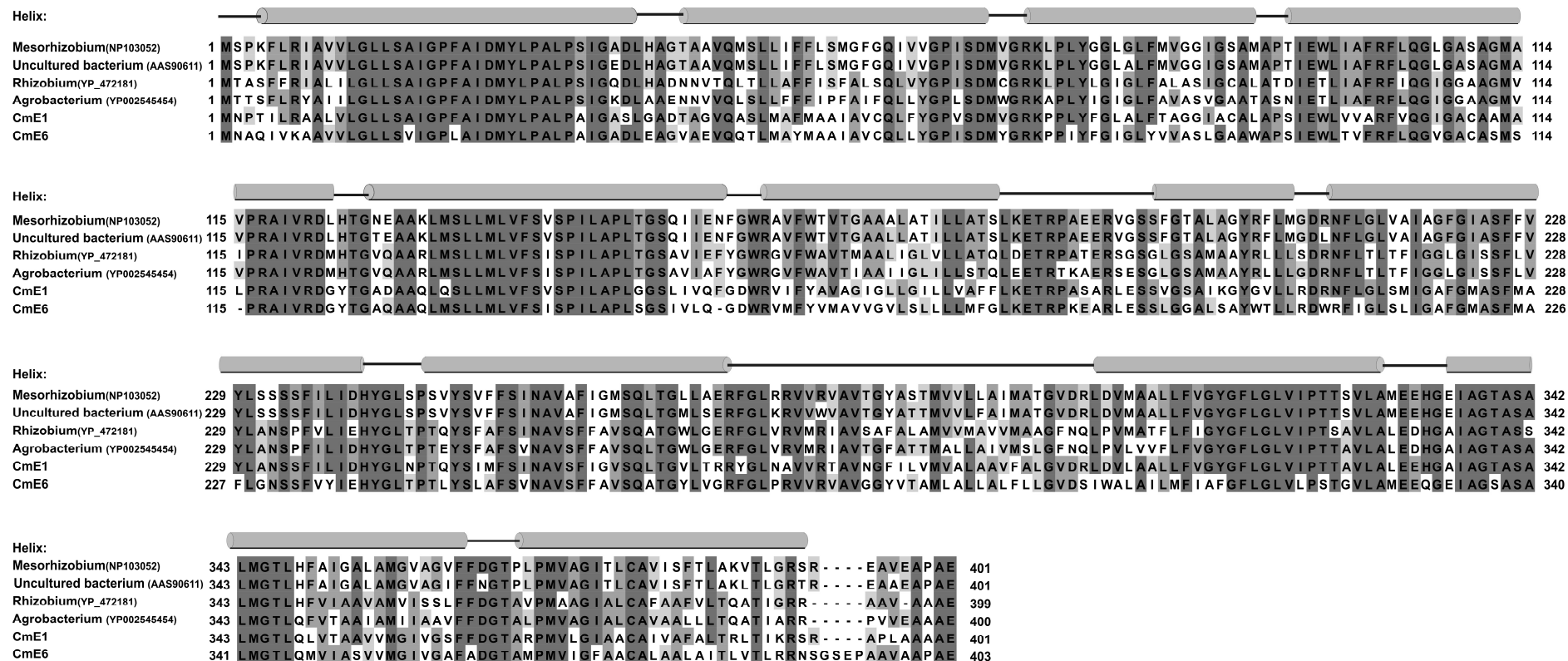
**Figura II.9.** Árbol filogenético construido siguiendo el método del vecino más próximo en el que se infieren las relaciones filogenéticas de los genes de resistencia a cloranfenicol, CmE1 y CmE6.

El árbol filogenético construido a partir de las secuencias de estas proteínas y de otras encontradas en la base de datos del NCBI (figura II.9), muestra como las proteínas CmE1 y CmE6 están muy relacionadas entre ellas (con un 65% de identidad a nivel de aminoácido) y también con otras bombas de extrusión presentes en proteobacterias del suelo.

Algunas de estas bombas de extrusión relacionadas con la resistencia a cloranfenicol están involucradas también en resistencia a tetraciclina (Liu, Z,Q *et al*, 2008). Además, tanto la proteína CmE1 como la CmE6 presentan una gran homología con CR4 (Figura II.9, número de acceso AAS90611; Riesenfeld *et al*, 2004), que es una proteína aislada también mediante un enfoque metagenómico y que confiere resistencia a tetraciclina. Se testó, de este modo, la resistencia a tetraciclina para ambos casos, y los dos clones, CmE1 y CmE6, confirieron resistencia a este antibiótico en células de *Escherichia coli*.

Fijándonos en su estructura, ambos genes presentan homología con otros transportadores de membrana que confieren resistencia a determinadas drogas, en concreto con la subfamilia de transportadores Bcr/CflA (tabla anexo AII.2.11), aunque son considerablemente diferentes a las proteínas descritas dentro de esta subfamilia (presentan menos de un 57% de identidad con otras proteínas presentes en las bases de datos). La predicción de la estructura de las proteínas CmE1 y CmE6 (figura II.10) muestra que tienen 12 regiones alfa-hélice transmembrana, al igual que ocurre con los miembros de la subfamilia Bcr/CflA.

## Capítulo II-Resistoma



**Figura II.10. Alineamiento múltiple de las proteínas que confieren resistencia a cloranfenicol y tetraciclina, CmE1 y CmE6, con otras pertenecientes a la familia Bcr/CflA.** La figura muestra la estructura secundaria de estas proteínas con doce dominios transmembrana (realizado con la herramienta online: <http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>).

#### *II.2.3.4. Análisis de los clones de resistencia a Trimetoprima*

De la genoteca EEZ se aislaron tres genes que confieren resistencia a trimetoprima (TmE1, TmE5, TmE25), mientras que en la genoteca México lluviosa, se aisló un único clon (Tm8.4) del que por sus peculiares características se tratará en un capítulo posterior (capítulo II.2.4).

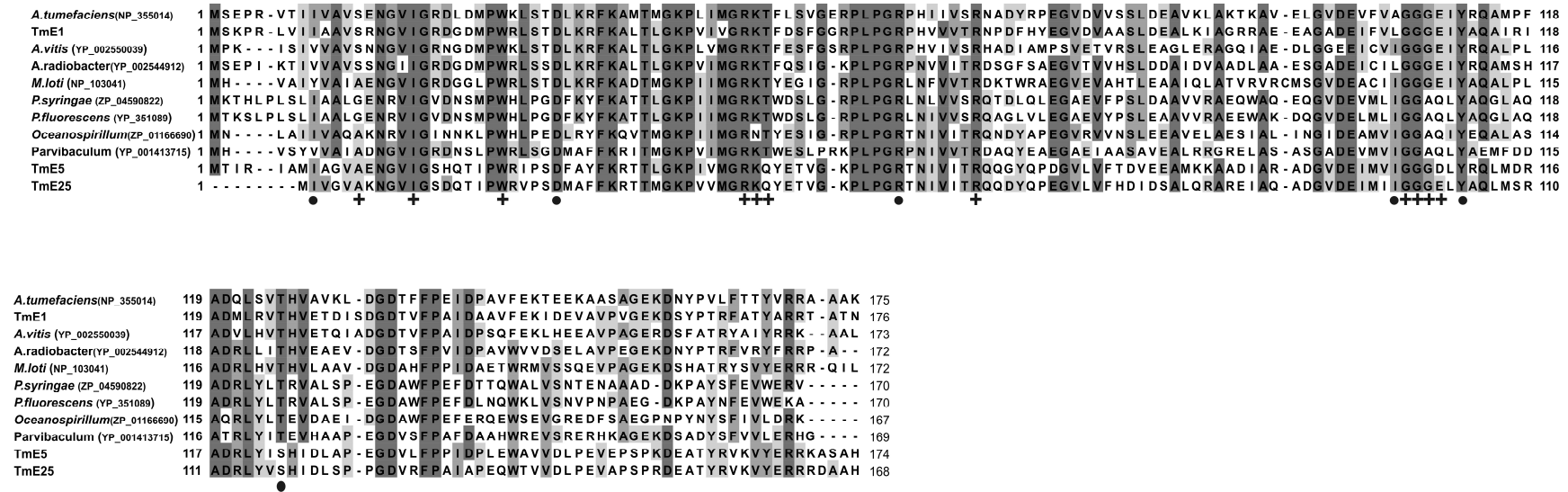
Con respecto a estos tres genes que confieren resistencia a trimetoprima, todos codifican para dihidrofolato reductasas de tipo I, que son enzimas codificadas por genes situados en el cromosoma bacteriano (*dhfr*; tabla anexo AII.2.1, figura II.7; Howell, E; 2005). Se pueden identificar, claramente, los residuos del sitio activo que interaccionan con el folato y con el NADP.

Los análisis filogenéticos (figura II.8) revelan que la enzima TmE1 está más relacionada con otras *dhfr* encontradas en bacterias del orden *Rhizobiales* y que se encuentra más alejada de otras relacionadas con las beta y gamma proteobacterias.

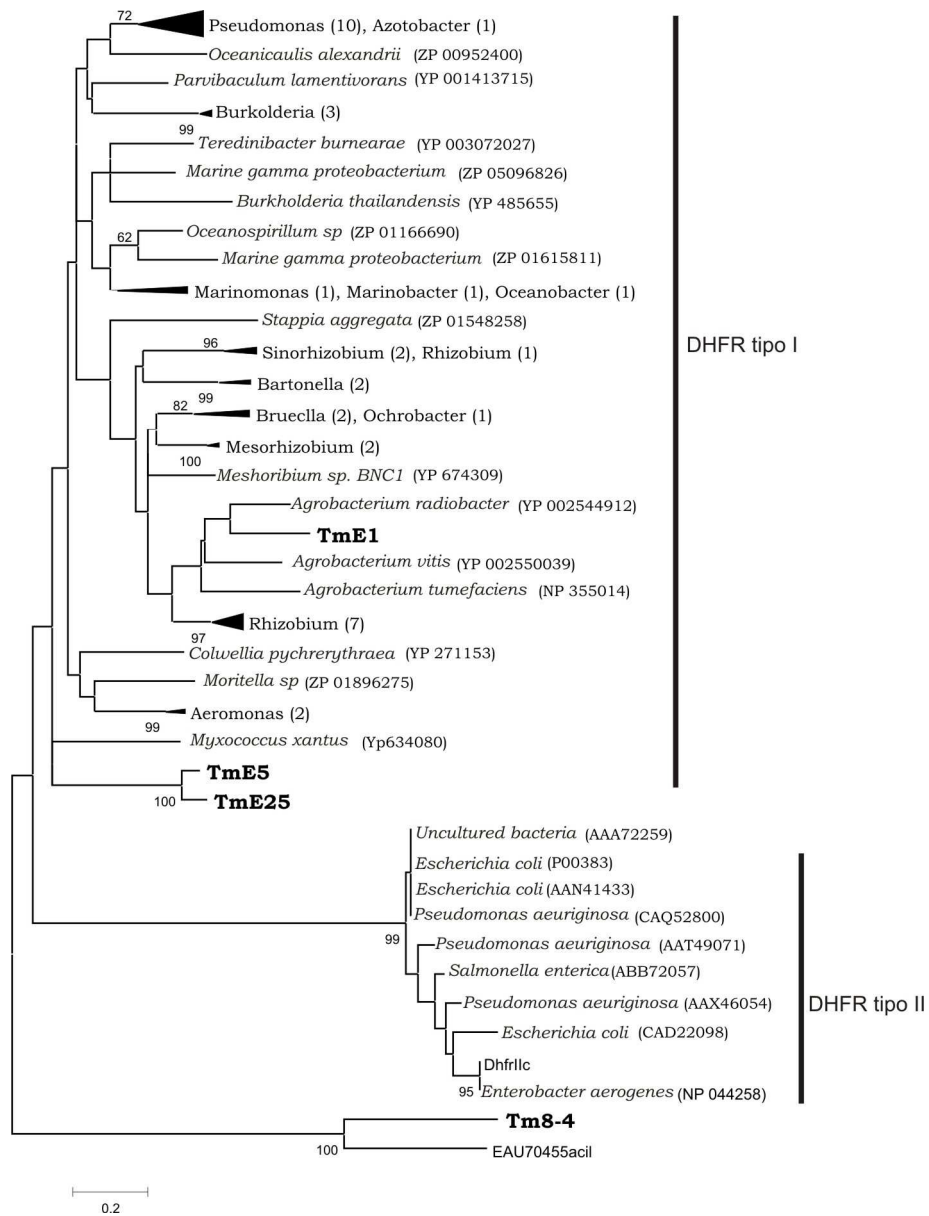
Las proteínas TmE5 y TmE25, sin embargo, están muy relacionadas entre ellas y se sitúan muy distantes de otros grupos de proteínas, siendo las proteínas a las que más se le parecen codificadas por la bacteria *Mysococcus xantus*. Todas estas enzimas están claramente diferenciadas del otro grupo descrito de dihidrofolato reductasas (el tipo II), que engloba principalmente a proteínas codificadas por elementos móviles, como integrones (Alonso, H. *et al*, 2006).

Por su parte, la proteína Tm8.4 no se engloba ni en el grupo de las dihidrofolato reductasas de la clase I ni en el grupo de las dihidrofolato reductasas de la clase II. Puede tratarse, por tanto, de un nuevo tipo de dihidrofolato reductasa, lo cual se tratará más adelante.

## Capítulo II-Resistoma



**Figura II.7. Alineamiento de las proteínas TmE1, TmE5 y TmE25 con otras pertenecientes a la familia I de dihidrofolato reductasas e identificación de los motivos conservados.** El alineamiento se ha realizado con la herramienta online *Clustal W* (<http://align.bmr.kyushu-u.ac.jp/mafft/software/>). En la figura se muestran con un círculo los residuos aminoacídicos que interaccionan con el foltato y con una cruz los residuos que interaccionan con el NADP (Howel, E et al, 2005).



**Figura II.8.** Árbol filogenético construido siguiendo el método del vecino más próximo en el que se infieren las relaciones filogenéticas de los genes de resistencia a trimetoprima aislados a partir de la genoteca de la Estación Experimental del Zaidín (TmE1, TmE5 y TmE25).

## **II2.4. ANÁLISIS DEL CLON DE RESISTENCIA A TRIMETROPRIMA: Tm8.4. CARACTERIZACIÓN DE UNA NUEVA ENZIMA CON ACTIVIDAD DIHIDROFOLATO REDUCTASA**

### **II2.4.1 Análisis de la secuencia del clon Tm8.4.**

Como ya hemos comentado en el capítulo II.2.3.3, de la genoteca México lluviosa se aisló un clon (Tm8.4), que confería resistencia a trimetoprima en células de *E.coli*.

Como en los casos anteriores, se hicieron diferentes subclones usando distintas dianas para enzimas de restricción, para poder delimitar de este modo el fragmento del clon que contenía el marco abierto de lectura que era capaz de codificar la proteína que confiere resistencia a este antibiótico (figura II.11.a). Así, observamos que el subclon Tm8.4 $Xho$ I de 1,5 Kb seguía confiriendo resistencia a trimetoprima.

Tras secuenciar estas 1,5 Kb encontramos dos marcos abiertos de lectura (Figura II.11a, tabla anexo AII.2.1). Uno codificaba para una proteína con un 42 % de homología a nivel de aminoácido con una proteína reguladora de respuesta de dos componentes (ORF1), y el otro marco abierto de lectura (ORF2) codificaba para una proteína con un porcentaje de homología de un 41% con una 3-oxoacil (acilcarrier) reductasa, que es una enzima implicada en el metabolismo de ácidos grasos.

La trimetoprima es un antibiótico bacteriostático que inhibe a la enzima dihidrofolato reductasa, encargada de convertir el dihidrofolato en tetrahidrofolato, un grupo metilo requerido para la síntesis de novo de purinas, timidina monofosfato y ciertos aminoácidos.

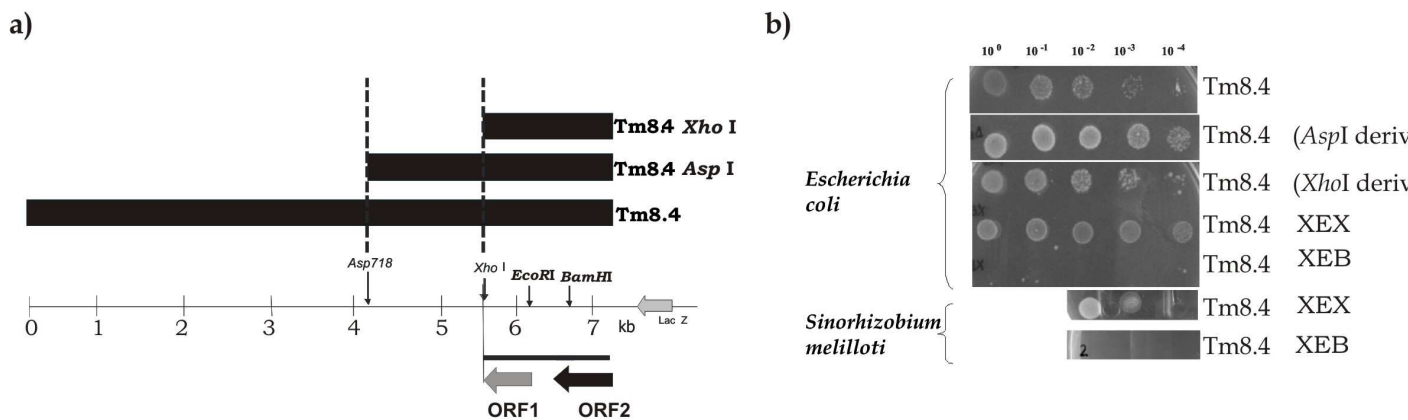
Los mecanismos descritos hasta el momento para la resistencia a trimetoprima están relacionados con una reducción en la permeabilidad celular, la sobreproducción de la enzima dihidrofolato reductasa o la producción de una reductasa de conformación alterada con disminución de la afinidad por el antibiótico. Dichas formas de resistencia pueden aparecer por mutación o, más frecuentemente, por la intervención de reductasas resistentes codificadas por plásmidos.

A priori, ninguno de los dos ORFs encontrados en el subclon Tm8.4 $Xho$ I explicaba una posible actividad que confiriera resistencia a trimetoprima. Tan sólo podría apuntar a ser candidato el ORF2 por codificar una proteína con una posible actividad reductasa.

Para poder determinar cual de los marcos abiertos de lectura identificados estaban causando la resistencia, se realizaron mutaciones puntuales de rupturas del marco abierto de lectura basados en la estrategia de matar las dianas de restricción *Bam*HI (XEB; que producía un cambio en la pauta de lectura del ORF1) y *Eco*RI (XEX; que produce un cambio en la pauta de lectura del gen ORF2) mediante el uso de la polimerasa Klenow.

Como se observa en la figura II.11b, la ruptura de la fase de lectura del ORF1 no producía ningún cambio en el perfil de resistencia, el clon mantenía la capacidad de conferir resistencia a trimetoprima, mientras que la eliminación del ORF2 hacía que el clon perdiera esta capacidad y que las células de *Escherichia coli* que lo portaban fueran sensibles a trimetoprima.

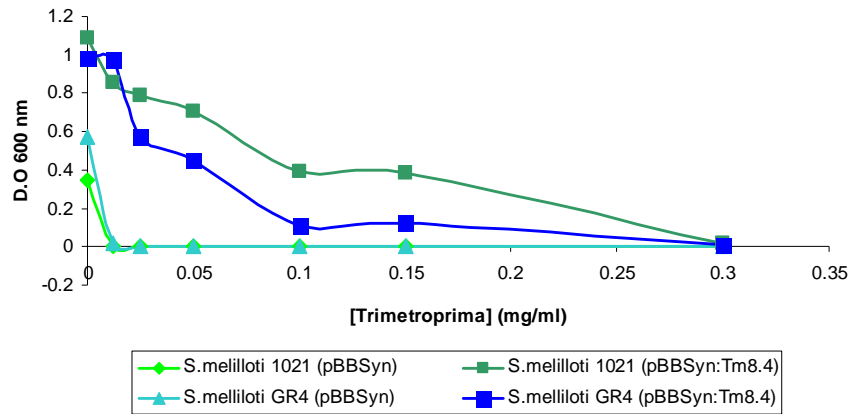
Este experimento demuestra, por tanto, que el gen que está causando la resistencia a trimetoprima es el ORF2.



**Figura II.11.** a) Mapa físico del clon de resistencia a trimetoprima Tm8.4 y de sus derivados. b) Los subclones derivados de las digestiones *Xho*I y *Asp*7418 siguen manteniendo la resistencia a trimetoprima. Por otro lado, las construcciones resultantes de un cambio en la pauta del ORF1 mantienen la resistencia, mientras que las resultantes de un cambio en el ORF 2 la pierden. Por tanto, la resistencia se debe al gen señalado en negro, que guarda homología con 3-oxoacyl-[acyl-carrier protein] reductasas, confiriendo resistencia tanto en células de *Escherichia coli* como en células de *Sinorhizobium meliloti*.

Para comprobar si este gen además de conferir resistencia a trimetoprima en células de *E. coli* podía conferir también resistencia en otros hospedadores distintos del sub-phyllum  $\gamma$ -proteobacteriase, se ensayó el gen del ORF2 en *Sinorhizobium meliloti*, ( $\alpha$ -proteobacteria). Se diseñaron cebadores específicos y se clonó de forma aislada el gen correspondiente al ORF2 en un replicón que funciona adecuadamente en células de *S. melliloti* (pBBSym) bajo la expresión de un promotor constitutivo. De este modo vimos (Figura II.12 y II.11b), que este gen era efectivamente, capaz de conferir resistencia en diferentes estirpes de *S. melliloti*.

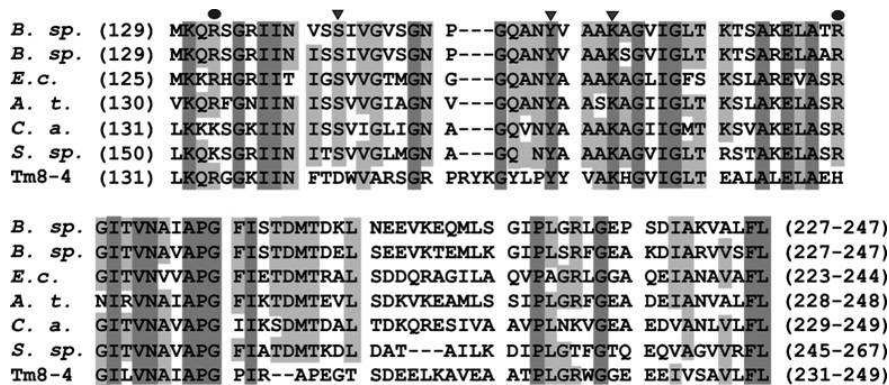




**Figura II.12.** Valores de densidad óptica de diferentes cepas de *S. meliloti* crecidas en medio Ty durante 10 horas con distintas concentraciones de trimetoprima.

Analizando la secuencia de la proteína codificada por el ORF2 (proteína Tm8.4), vimos que, al contrario de lo que ocurría con las otras proteínas que conferían resistencia a trimetoprima descritas en nuestro estudio (TmE1, TmE5 y TmE25; figura II.7), era complicado identificar los residuos del sitio activo implicados en la unión al folato y al NADP.

Sí pudimos encontrar, sin embargo, los residuos conservados descritos en otras oxoacyl-[acyl-carrier protein] reductasas para la unión del acetoacetyl coenzima A (figura II.13; Howell *et al*, 2005).



**Figura II.13.** Alineamiento parcial del clon Tm8-4 con 3-OxoAcyl-ACP reductasas (OARs). *Bacillus sp* OARs (*B. sp.* nos de acc: ZP\_01171570 y ZP\_01861660); *E. coli* (*E.c.* no de acc: P25716); *A. thermophilum* (YP\_002573030), *C. aetobutylicum* (*C.a.* NP\_350157) y *Synechococcus* (*S. sp.* ZP01470074). Los residuos que forman parte del centro activo conservados entre OARs de bacterias y plantas están señalados en triángulos. Los círculos señalan dos residuos de Arginina claves en la unión al acetoacetyl coenzima A para algunas OAR procarióticas y eucarióticas (Howell *et al*, 2005).

Además, la proteína Tm8.4 no posee ninguna homología a nivel de secuencia aminoacídica con las dihidrofolato reductasas descritas hasta el momento. Si observamos el árbol de la figura II.8, veremos como la asocia a otro grupo totalmente aparte, junto con otra oxoacyl-[acyl-carrier protein] reductasa.

De este modo, en este capítulo se ha demostrado que, poseemos un gen que codifica para una proteína con una homología muy baja en secuencia aminoacídica con una enzima relacionada, quizás, con el metabolismo de lípidos (3-oxoacil (acilcarried) reductasa), que confiere resistencia a trimetroprima tanto en *E.coli* como en *S.melilloti* y que en principio no guarda correlación ninguna con las dihidrofolato reductasas descritas hasta el momento.

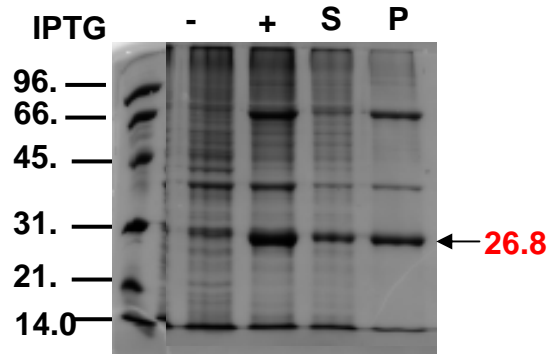
Se nos plantean varios interrogantes en este punto: ¿Cual es el mecanismo por el que esta proteína genera la resistencia a trimetroprima?, ¿Tiene esta enzima acaso actividad dihidrofolato reductasa, cuales son sus constantes bioquímicas?, ¿posee actividad 3-oxoacil (acilcarried) reductasa como apuntan los datos de homología obtenidos *en silico*?

#### **II2.4.2 Purificación de la proteína Tm8.4**

Se planteó una estrategia de purificación con el fin de poder caracterizar a la proteína Tm8.4.

De este modo, se diseñaron cebadores que permitieran el clonaje del gen de tal forma que se pudiera expresar de forma nativa en los sistemas pET (Studier *et al*). El amplicón obtenido se digirió con los enzimas *NdeI* y *BamHI* y se clonó en el vector de expresión pET3a. Con el vector así preparado, se transformaron células *Rosetta* de *E.coli*.

Ensayando diferentes condiciones, se demostró que el máximo nivel de expresión (figura II.14) se lograba creciendo las células *Rosetta* de *E.coli* que contenían el plásmido pET3a:Tm8.4p hasta la fase exponencial de crecimiento (hasta una densidad óptica de 0,6 aproximadamente) y añadiendo posteriormente 0,4 mM de IPTG e incubando los cultivos a 20°C. Las células se recogían a las 12 horas de la inducción.



**Figura II.14.** Gel SDS PAGE al 12% con muestras de la sobreexpresión de Tm8.4p. Carril 1: marcador; carriles 2-5: sobreexpresión de la proteína Tm8.4 (plásmido pEt3a:Tm8.4p), en el gel se muestran los extractos proteicos sin inducir (-); tras inducir (+); de sobrenadante (S) y precipitado (P). Se observa como la proteína Tm8.4p se acumula a una altura de unas 26,8 Kb.

Se comprobó si el extracto proteico (pET3a:Tm8.4p) así recogido poseía actividad dihidrofolato reductasa (tabla II.4), para lo cual se usó NADPH y dihidrofolato. Si tiene actividad dihidrofolato reductasa, la enzima reduciría el dihidrofolato a tetrahidrofolato, oxidando el NADPH a NADP. El cambio de NADPH a NADP se puede seguir espectrofotométricamente a 350nm, por lo que se comprobó frente a un extracto control con el vector sin inserto (pET3a) que efectivamente la enzima Tm8.4p era capaz de reducir el NADPH.

Para poder analizar en detalle la proteína Tm8.4, se siguieron diferentes pasos de purificación que se resumen en la tabla II.4 y en la figura II.15.

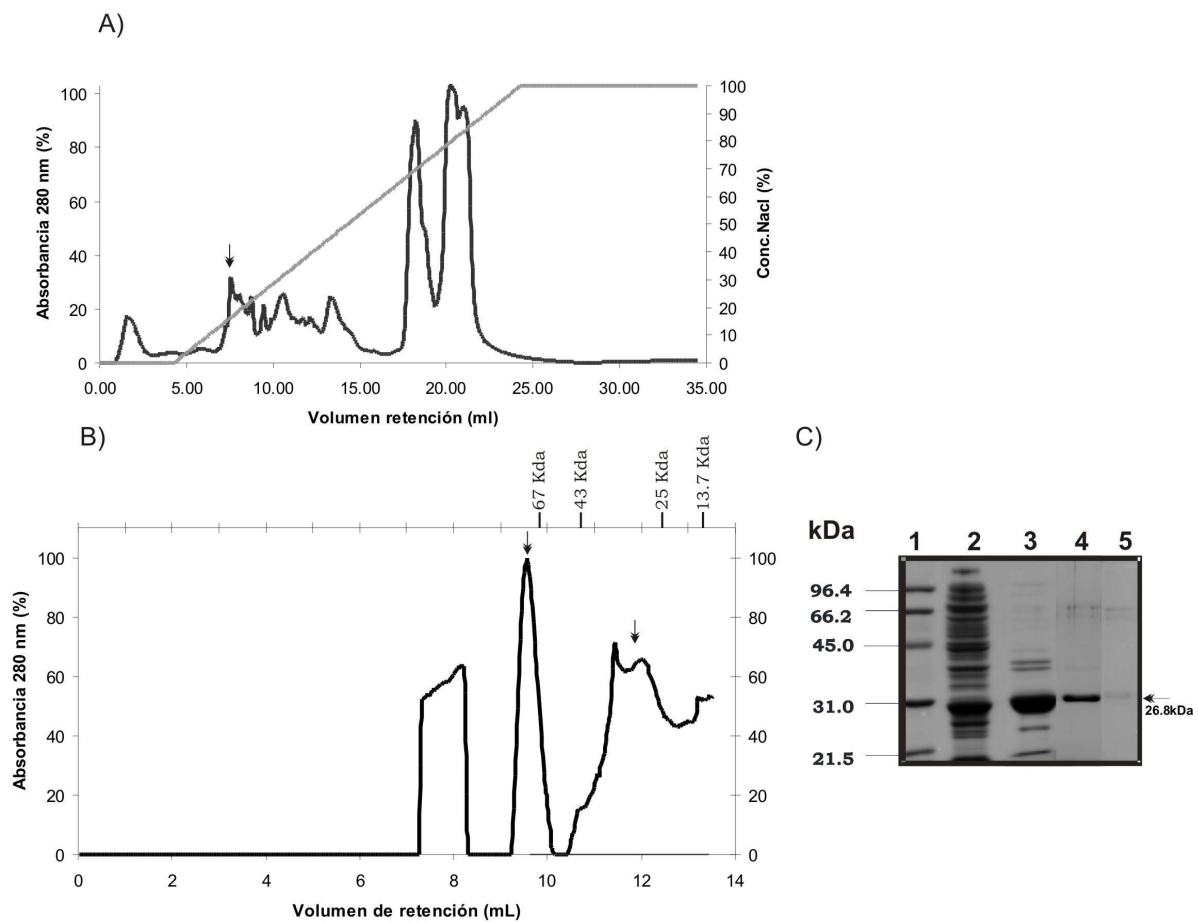
En primer lugar, se pasó el extracto crudo con una concentración de 22,1 mg/ml por una columna de intercambio iónico (mono Q), estableciendo un gradiente de NaCl de 0 a 1 M, y se observó que la proteína eluía entre 150 mM y 200 mM de NaCl a una concentración de 0,94 mg/ml. Se comprobó que esta fracción poseía actividad dihidrofolato reductasa.

Posteriormente, se pasó la muestra por una columna de filtración en gel, eluyendo la proteína pura en un 98% en dos fracciones diferentes: Una con un volumen de elución de 12ml que coincidiría con una forma monomérica (25,93 KDa) y otra con un volumen de elución de 9,5 ml y una forma dimérica (la mayoritaria; 67,8 KDa).

Con las fracciones obtenidas a partir de la cromatografía de intercambio iónico y la de filtración en gel también se realizaron diferentes mediciones enzimáticas para corroborar que la enzima seguía manteniendo su actividad (tabla II.4).

Paso Purificación	Proteína	Volumen	Total Proteína	Actividad DHFR		Fold	Total actividad	Rendimiento
	mg/ml			mg	u/ml			
<b>Extracto Crudo</b>	22,1	2,5	55,25	4,47	0,2	1	11,18 u	100%
<b>Intercambio Iónico</b>	0,94	1	0,94	3,38	3,6	18,0	3,38 u	75,7%
<b>Hidrofobicidad</b>	B6 0,09	0,5	0,06	0,3	5,0	25	0,15 u	28,0%

**Tabla II.15.** Resumen de la purificación de la proteína Tm8.4. En la tabla se muestra la concentración de proteína y volumen de partida en cada paso de purificación, además de la actividad dihidrofolato reductasa.



**Figura II.15.** Esquema de purificación de la proteína Tm8.4p. A) Cromatografía de intercambio iónico con un gradiente de 0 a 1 M de NaCl. B) Cromatografía de filtración en gel, se indica en la parte superior del gráfico el volumen de retención de diferentes compuestos de peso molecular conocido; la flecha en ambas gráficas indica en las fracciones en las que eluye la proteína. C) Gel en el que se observa la purificación, Carril 1: Marcador, Carril 2: Extracto crudo; Carril 3: Extracto tras pasar por la columna de intercambio iónico; Carriles 4 y 5: Extractos tras pasar por la columna de filtración en gel, en forma dimérica (carril 4) y en forma monomérica (carril 5).

Como hemos comentado, la proteína eluye en forma monomérica (25,93 Kda) y dimérica (67,8 Kda), aunque en estas condiciones principalmente eluye en forma de dímero. Para poder observar si los monómeros se unían por puentes disulfuro

para formar el dímero, se corrieron geles SDS PAGE al 12% sin ningún agente reductor como el DTT, observándose en todos los casos una única banda de 25,93 Kda y no viendo una banda mayor correspondiente al dímero. Este experimento indica que deben de existir otro tipo de interacciones (como interacciones iónicas, hidrofóbicas, puentes de hidrógeno) que estén interviniendo en la unión de los monómeros.

Con la enzima así purificada, se pudo realizar su caracterización bioquímica completa, la cual se detalla en los siguientes apartados.

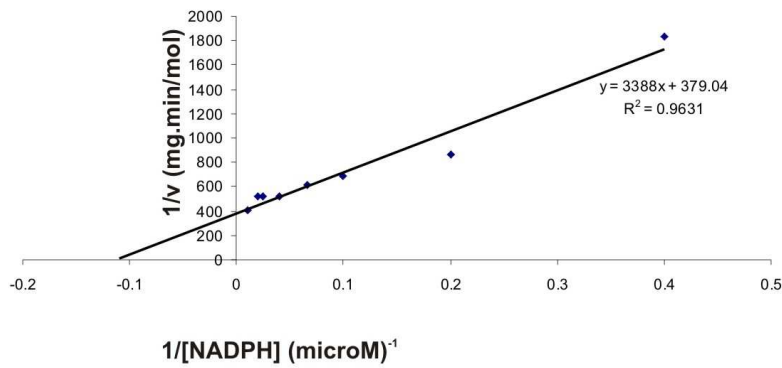
#### **II2.4.3 Características bioquímicas de la proteína Tm8.4**

Para determinar como de efectiva es esta enzima para los sustratos acetil-CoA y dihidrofolato se calcularon sus constantes cinéticas ( $K_m$  y  $V_m$ ). Considerando que en las fases iniciales de la reacción enzimática la concentración de sustrato en relación con el tiempo es casi lineal, se pueden obtener datos exactos de la velocidad en función de la concentración de sustrato si se conoce la concentración inicial de éste.

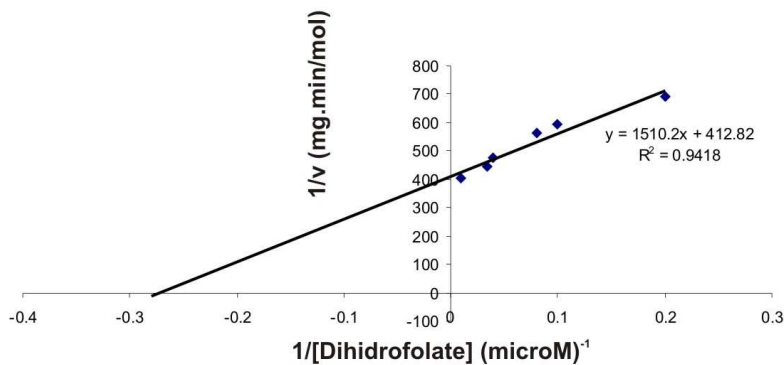
Con este fin, se fijó una concentración de extracto con la enzima constante y se varió la concentración de sustratos, de folato y NADP entre 2,5-100  $\mu\text{M}$ , midiendo en cada caso la velocidad inicial de la reacción a una temperatura de 25°C y pH de 7,4. Se calculó la representación de Lineweaver-Burk para cada sustrato: en esta representación la extrapolación lineal de los datos nos proporciona tanto  $V_{\text{max}}$  como  $K_m$ .

De este modo (figura II.16), los valores de  $K_m$  para el dihidrofolato y para el NADPH son de 3,7 y 9  $\mu\text{M}$  respectivamente (presentando de esta forma más especificidad por el NADPH que por el dihidrofolato). Asimismo, los valores de velocidad máxima son de 0,0024 y 0,0026 para el dihidrofolato y el NADPH respectivamente.

A)



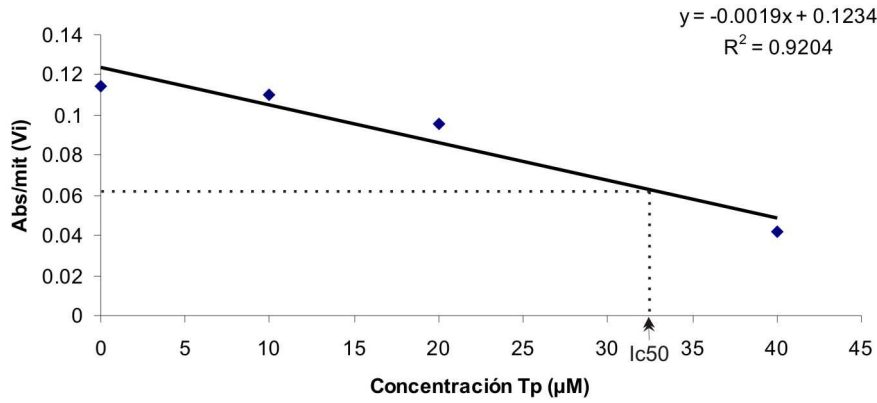
B)



**Figura II.16.** Representación de la actividad de la enzima Tm8.4 mediante la ecuación de Lineweaver-Burk. En esta gráfica doble inversa, se representa  $1/V_i$  frente a  $1/[S]$ . A) Representación del sustrato NADPH. B) Representación del sustrato dihidrofolato. Los valores que se representan son la media de tres experimentos independientes.

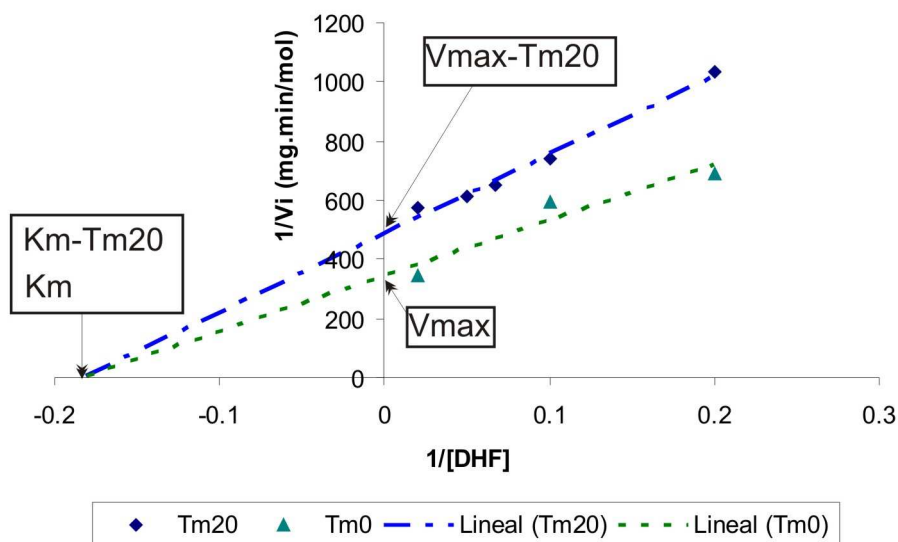
Además, con la enzima purificada también se testó la actividad frente al sustrato aceto-acetil-Coenzima A, sustrato de la familia de proteínas con la que guardaba homología (3-oxoacil-ACP reductasa). No se encontró actividad frente aceto-acetil-Coenzima A, comprobando de esta forma que la enzima Tm8.4p tiene, efectivamente, actividad dihidrofolato reductasa y no tiene actividad 3-oxoacil-carried reductasa.

Asimismo, se calculó la concentración de trimetoprima que produce un grado de inhibición del 50% (IC50) en determinadas condiciones (usamos concentraciones de sustratos saturantes: 20  $\mu\text{M}$  de dihidrofolato, 100  $\mu\text{M}$  de NADPH; y condiciones de actividad óptima: 25°C y pH 7,4). Se observó, de este modo, que la concentración de trimetoprima a la cual la velocidad de la proteína Tm8.4p es la mitad de la máxima es de 32,47  $\mu\text{M}$  (figura II.17).



**Figura II.17.** Cálculo de la concentración de trimetoprima que produce un grado de inhibición del 50%.

Para poder determinar qué tipo de inhibición causa la trimetoprima en la enzima Tm8.4p (si es inhibición reversible competitiva, no competitiva o mixta), se incubó la enzima en 20 μM de trimetoprima y se calcularon los diferentes valores de Km para el dihidrofolato en cada situación (figura II.18). De este modo, se observa cómo al aumentar la concentración de antibiótico la Vmax disminuye, mientras que la Km se mantiene inalterada, datos que sugieren que la trimetoprima estaría actuando en este caso como un inhibidor no competitivo, donde éste no afecta a la unión del sustrato en la enzima, es decir no afecta a la Km, pero obstaculiza la catálisis (disminuye la Vmax).



**Figura II.18.** Representación de la actividad de la enzima Tm8.4 mediante la ecuación de Lineweaver-Burk sin trimetoprima (Tm0) y con 20 μM de Trimetoprima (Tm20). En esta representación doble inversa, se representa  $1/V_i$  frente a  $1/[S]$ . La  $V_{max}$ , al incubar la enzima con el inhibidor, disminuye, mientras que la  $K_m$  permanece inalterada.

**Discusión Capítulo II.1. Caracterización del Resistoma**

En este capítulo se ha desarrollado un análisis funcional para acceder al conjunto de genes de resistencia a antibióticos (el resistoma; Wright *et al*, 2007) de diferentes tipos de suelo: dos suelos procedentes del Parque Natural de Tehuacán-Cuicatlán (en época seca y lluviosa; genotecas mex lluvioso y mex seca) y un suelo manipulado por el hombre perteneciente a una parcela experimental situada cerca de nuestro sitio de trabajo (EEZ).

Para llevar a cabo este estudio, se han construido genotecas de expresión en el fago lambda mediante el sistema lambda ZAP-Phagemid Expresión (Stratagene). Entre las ventajas de clonar el ADN medioambiental en este tipo de vectores, está el de conseguir un buen rendimiento en lo que respecta a ADN clonado y la posibilidad de trabajar mediante escisión en fagémido, no ya con fagos sino con plásmidos. De esta forma, se puede trabajar con estrategias de muestreo de selección sobre colonias de *E. coli*.

En nuestro estudio se han construido tres genotecas denominadas mex-lluvioso, mex-seco y EEZ, cuyas características se pueden observar en la tabla II.1. La genoteca EEZ es la genoteca más pequeña, con 65000 ufc y un tamaño de inserto de unas 7 Kb. Si asumimos que el tamaño medio de un genoma bacteriano es de unas 40000 kb, en esta genoteca tendríamos clonado el equivalente a 35 genomas similares al de *E.coli*. En el caso de las genotecas mex-lluviosa y mex-seca, poseen un tamaño superior, 250000 ufc, con un tamaño de inserto también en torno a las 7 kb. Esto supone que tendríamos clonado en las genotecas aproximadamente 1,5 gibases lo que equivaldría a aproximadamente 250 genomas bacterianos del tipo *E.coli*. La información genética que se puede tener clonada en este tipo de vectores es, por tanto, muy elevada. Unido eso, además, a la capacidad de expresar proteínas heterólogas, hace que este sistema tenga mayor éxito que otro tipos de vectores (BACs, plásmidos, etc.) o que una gran colección de clones (Riesenfeld *et al*, 2004; Allen *et al*, 2009).

Las principales limitaciones en este tipo de estrategias radican en tener un buen sistema de búsqueda de la actividad a testar, la expresión de la proteína heteróloga en *E.coli* y el mecanismo de escisión a fagémido.

De todos modos, este tipo de metodologías, asociadas a la búsqueda de nuevos genes de resistencia a antibióticos, presentan diversas ventajas sobre los métodos tradicionales de cultivo de células resistentes o de búsqueda de genes de resistencia mediante PCR: i) te permiten acceder a la enorme



diversidad de microorganismos no cultivables de un ecosistema, ii) no requieren conocimiento previo de las secuencias y iii) se puede acceder a la secuencia génica completa.

En este trabajo hemos identificado a partir de genotecas metagenómicas 11 clones que producen resistencia a diferentes tipos de antibióticos en *E.coli*. Estas proteínas tienen un rango de identidad a nivel de aa de entre el 40-69% con otros genes depositados en las bases de datos, lo cual parece indicar que los microorganismos del suelo poseen genes capaces de generar mecanismos de resistencia a antibióticos con mayor diversidad genética de lo descrito en otros trabajos. De estos 11 clones, 8 pertenecen a la genoteca denominada EEZ, que es mucho más pequeña (en número de unidades formadoras de calvas) que las genotecas mex-lluviosa y mex-seca, en las cuales se han aislado tan sólo 2 y 1 clon de resistencia respectivamente.

Además de la baja homología de estas secuencias con otras depositadas en las bases de datos, normalmente las proteínas más cercanas suelen ser proteínas de resistencia que proceden también de aproximaciones metagenómicas. Por ejemplo, los genes de resistencia a cloranfenicol encontrados en nuestras genotecas están muy relacionados con el gen *tcaB* del clon CR4 (Accession Number: AY566822) encontrado en otra genoteca medioambiental de suelo (Resenfeld *et al*, 2004), y el gen de resistencia a gentamicina GmE1 presenta una gran homología con otro clon aislado mediante este tipo de aproximación que confiere resistencia a kanamicina. Este hecho, está relacionado probablemente con similares estrategias de búsqueda, sobre todo el hecho de haber usado *E. coli* como hospedador para la búsqueda de genes de resistencia.

Para la identificación del posible origen de todos estos genes se usó el programa online PhyloPythia (tabla AII.2.1; McHardy *et al*, 2009) y TETRA. Este programa manifiesta un predominio del grupo *alfa-Proteobacteria* para los fragmentos de ADN clonados en las genotecas (ApE14, GmE1, CmE1, CmE6, TmE1 y TmE25). Esta asignación filogenética, además, se puede correlacionar con la identificación de sintenia a nivel de gen en algunos genomas de referencia.

Por ejemplo, los clones CmE6 y CmE1 poseen, aguas arriba del gen que codifica para el transportador de membrana Bcr/CflA, un regulador transcripcional que posee sintonía con el genoma de *Mesorhizobium loti* (tabla AII.2.1). De igual forma, aguas debajo de los genes *dhfr* encontrados en los

clones TmE1, TmE25 y TmE5, encontramos un gen que codifica para la subunidad de la proteasa HflK, gen relacionado normalmente con los genes dhfr. El origen probable de estos fragmentos de ADN es el orden Rhizobiales. Sin embargo, en el caso del clon Gm8.3, no se pudo establecer una sintenia clara con los genomas existentes en las bases de datos. Aguas arriba del gen de resistencia a gentamicina se han encontrado numerosos genes relacionados con distintas resistencias a metales.

En cualquier caso, no se encontraron en ninguno de los clones genes relacionados con transferencia horizontal o movilidad, como elementos repetidos, secuencias de inserción, trasposones... (tabla AII.2.1).

Las proteínas de resistencia a ampicilina encontradas en la genoteca EEZ son enzimas beta-lactamasas. Estas proteínas guardan gran homología con otras pertenecientes a bacterias típicas de suelo como *Rhizobium*. Sin embargo, la enzima beta-lactamasa aislada de la genoteca de mex-seca guarda una gran homología con otras enzimas procedentes de *Solibacter*, que pertenece al phylum *Acidobacteria*, phylum que según los datos obtenidos en las genotecas 16S (Capítulo I) es especialmente abundante en este tipo de suelos.

En el caso de las proteínas que confieren resistencia a gentamicina, se corresponden con enzimas N acetil transferasas. La proteína Gm8.3 está muy relacionada con las proteínas pertenecientes a la familia AAC (3')-I de gammaproteobacterias (Riccio, M.L *et al*, 2003; Ahmed, A.M *et al*, 2004) y la proteína GmE1 está relacionada con la familia AAC (3')-III, si bien por lo que se intuye de los árboles filogenéticos podría constituir una nueva subclase de enzimas dentro de esta familia. Además, estudios clínicos previos han indicado que existe una correlación entre la presión selectiva de los antibióticos y el patrón de resistencia a aminoglicósidos (Vakulenko *et al*, 2003). Nuestros resultados apoyan estos trabajos respecto a la alta frecuencia de genes del tipo AAC3 encontrados cuando se usa la gentamicina como antibiótico.

Todos los clones aislados en las genotecas ambientales que han conferido resistencia a cloranfenicol constituyen bombas de extrusión. No se ha encontrado ningún clon que porte algún tipo de enzima, como por ejemplo cloranfenicol-acetil-transferasasa. Este hecho podría estar reflejando que este tipo de determinantes de resistencia pueden estar muy poco representados en la población de estudio o, simplemente, ser producto de una limitación en el

sistema de búsqueda en *E.coli*, o ambos casos. Los dos clones aislados que confieren resistencia a cloranfenicol son muy parecidos entre sí y, a su vez, muy parecidos a otras bacterias del suelo como *Mesorhizobium*. Además, también confieren resistencia a tetraciclina, como se ha observado en otros casos (Liu, Z,Q et al, 2008).

En lo que respecta a los clones de resistencia a trimetoprima, codifican enzimas dihidrofolato reductasas que, o son capaces de sustituir la actividad de la dihidrofolato reductasa de *E.coli* sensible al antibiótico, o bien se están superproduciendo a partir del fagémido multicopia pBK-CMV. Las proteínas TmE1, TmE5 y TmE25 son dihidrofolato reductasas de tipo I, generalmente codificadas en el cromosoma bacteriano (Hoewel et al, 2005), a diferencia de las de tipo II cuyo origen es más bien plasmídico, y muy relacionadas con otras enzimas similares encontradas en bacterias del suelo (figura II.8)

La enzima Tm8.4 tiene unas características especiales. Guarda una homología muy baja (40%) con enzimas relacionadas con el metabolismo lipídico (3-oxoacil-acilcarried-reductasas) y su secuencia nucleotídica y aminoacídica no se parece en nada a las dihidrofolato reductasas descritas hasta el momento (con un 11% de homología con la DHFR de *E.coli*; Fierke et al, 1987; Lee et al, 2010; Hoewel et al, 2005), a las pertenecientes a la familia de dihidrofolato reductasas cromosómicas (familia *dhfrA*), ni a la familia de las *dhfrB* plasmídicas (White et al, 2001). No se pueden encontrar los residuos conservados de unión al tetrahidrofolato y al NADPH.

La purificación y posterior caracterización bioquímica de esta proteína ha demostrado que tiene actividad dihidrofolato reductasa y que no posee actividad 3-oxoacil-acilcarried-reductasa. . Lo cual, apunta en muchos casos a la necesidad de pruebas bioquímicas, además de los datos in silico, para la correcta asignación y papel de los genes encontrados mediante estrategias de secuenciación masiva, entre otras.

Con respecto a sus características bioquímicas (tabla II.16), esta enzima presenta actividades cinéticas únicas no descritas hasta el momento. Posee menor especificidad por el sustrato dihidrofolato que las dihidrofolato reductasas previamente descritas en *E.coli* (Fierke et al, 1987) o en *Streptococcus pneumoniae* (*S. pneumoniae*; Lee et al, 2010). La especificidad por el sustrato NADPH es muy parecida a la descrita en las otras enzimas.

Lo que realmente llama la atención de la enzima Tm8.4 es que es muy resistente a la inhibición por trimetoprima, teniendo una IC<sub>50</sub> mucho más

elevada que el resto de dihidrofolato reductasas descritas (100 veces mayor; tabla II.16).

Además, la estructura de la proteína Tm8.4 parece que también tiene ciertas peculiaridades. Las enzimas DHFR cromosómicas suelen ser monoméricas, mientras que las plasmídicas suelen ser tetraméricas (Howel et al, 2005). La proteína Tm8.4 descrita en este trabajo es dimérica.

El estudio de nuevas enzimas que confieran resistencias, como la Tm8.4, podría ayudar a diseñar nuevas moléculas antimicrobianas frente a las bacterias patógenas con DHFRs resistentes a drogas.

De este modo, en este trabajo se ha descrito una enzima que por secuencia y estructura no se parece en nada a las dihidrofolato descritas hasta el momento y que tiene menor afinidad por el sustrato tetrahydrofolato, por lo que podría ser una enzima que en su ambiente natural estuviera jugando otro papel diferente pero que, bajo la presión selectiva de la trimetroprima, aflora como una dihidrofolato reductasa.

<b>Enzima</b>	<b>K<sub>M(NADPH)</sub> (μm)</b>	<b>K<sub>M(H2Folato)</sub> (μm)</b>	<b>IC<sub>50</sub> (μm)</b>	<b>Estructura</b>
<b><i>S.pneumoniae</i></b>	15,7 <sup>a</sup>	4,4 <sup>a</sup>	0,48 <sup>a</sup>	/
<b><i>E.coli</i></b>	4,8 <sup>a</sup>	0,7 <sup>a</sup>	0,13 <sup>a</sup>	Monómero, 18.000 Da <sup>b</sup>
<b>R67 DHFR</b>	3 <sup>b</sup>	5,8 <sup>b</sup>	/	Tetrámero, 34.000 Da <sup>b</sup>
<b>Tm8.4</b>	<b>3,7</b>	<b>9</b>	<b>20</b>	<b>Dimero, 26.000 Da</b>

**Tabla II.16.** Tabla comparativa entre la enzima dihidrofolato Tm8.4 descrita en este estudio y otras dihidrofolato descritas en la literatura.

<sup>a</sup> Datos tomados de Lee et al, 2010

<sup>b</sup> Howel et al, 2005

En su conjunto, todos los datos presentados en este trabajo pueden dar una idea del posible origen de los diferentes mecanismos de resistencia a antibióticos y de los nuevos determinantes de resistencia que podrían surgir en la clínica. La baja homología encontrada entre las proteínas que confieren resistencia a antibióticos descritas en este trabajo y las depositadas en las bases de datos apoyan las hipótesis desarrolladas en los últimos años que

indican que estas proteínas podrían estar jugando otro papel muy diferente en condiciones naturales (Martínez J.L *et al*, 2008).

Cualquier compuesto químico puede ser tóxico a altas concentraciones. De esta forma, es factible que las moléculas que actualmente se están usando por sus propiedades antibióticas estén jugando un papel muy diferente en condiciones naturales, a otras concentraciones. Por ejemplo, podrían actuar como moléculas de señalización (Fajardo *et al*, 2009). Algo similar puede estar ocurriendo con los genes de resistencia a antibióticos. Por ejemplo, se conoce que las enzimas que inactivan los antibióticos aminoglicósidos pueden tener un papel metabólico en la célula (Miller *et al*, 1967) o estar involucradas en la acetilación del peptidoglicano (Macinga *et al*, 1999). Una estructura similar entre el antibiótico y el sustrato de la enzima podría conseguir que estas enzimas confieran resistencia, aunque normalmente tengan en otro papel muy diferente. En una situación de este tipo se podría encontrar la enzima descrita que produce resistencia a trimetroprima: Tm8.4.

## **Anexo Capítulo II2.**

Apéndice Capítulo II.2.

**Tabla AII.2.1** Secuencias codificantes predichas en los clones de resistencia a antibióticos aislados en las librerías metagenómicas de estudio.

Plásmido (tamaño de inserto y pb secuenciados)		G+C <sup>a</sup> (%)	Numero Acceso	ORF <sup>b</sup>	Tamaño proteína (aa)	Proteína relacionada (número aa),	Organismo	E-value <sup>c</sup>	Identidad Aa /similitud % <sup>c</sup>	Clasificación filogenética de los insertos Dominio/Phylum /Clase <sup>d</sup>
<i>Clones Resistencia a Ap</i>										
Ap6-8 (6.9 Kbp / 2,951)	MEX-Seca	56.1	FN640464	1 (343.. 1518 direct)	391	lipid-A-disaccharido synthasa (383), YP_001951013	[ <i>Geobacter lovleyi</i> SZ]	2e-65	40/57	Bacteria/Proteobacteria
				<b>2 (1523...2482 reverse)</b>	<b>319</b>	<b>β-lactamasa clase A (281), YP_823559</b>	<b><i>Solibacter usitatus Ellin6076</i></b>	<b>1e-81</b>	<b>57/70</b>	
ApE1 (6,5 Kbp/ 3,033)	EEZ	56.8	FN640466	1 (>1...612 reverse)	>170	N-acetylmuramyl-L-alanine amidasa, regulador negative de AmpC, AmpD (320), ZP_03631130.1	[ <i>bacterium Ellin514</i> ]	6e-40	59/74	Bacteria
				<b>2 (727...1764 reverse)</b>	<b>345</b>	<b>beta-lactamasa clase A (352), YP_002129833.1</b>	<b>[<i>Phenylobacterium zucineum HLK1</i>]</b>	<b>4e-65</b>	<b>45/58</b>	
ApE14 (10 Kbp/5,579)	EEZ	53.1	FN640467	1 (>2...186 reverse)	>62	Proteína hipotética (455), YP_674672.1	[ <i>Mesorhizobium sp. BNC1</i> ]	2e-21	79/87	Bacteria/Proteobacteria/Alphaproteobacteria
				2 (378...944 direct)	188	Proteína hipotética (177), YP_673334.1	<i>Mesorhizobium sp. BNC1</i>	5e-62	65/76	
				<b>3 (922...2064 reverse)</b>	<b>380</b>	<b>Beta-lactamasa (385), YP_002976497.1</b>	<b><i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> WSM1325</b>	<b>7e-139</b>	<b>63/79</b>	
				4 (2175-... 3044 direct)	289	Regulador trascricional familia LysR(292), YP_002282440.1	[ <i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> WSM2304]	8e-131	80/87	
				5 (3356....4339 reverse)	327	dihydrodipicolinato synthasa (321), YP_001370746.1	<i>Ochrobactrum anthropi</i> ATCC 49188	2e-79	55/71	
				6 (>4794...5579 direct)	>262	adenylate/guanylato ciclase (468), YP_002972998.1	[ <i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> ]	1e-140	90/95	

Anexo al capítulo II-Resistoma

							WSM1325]			
Clones de Resistencia a Gm										
Gm8-3 (8 Kbp/ 8,046)	MEX-LLuviosa	52.7	FN640465	1 (>1...978 reverse)	>242	Proteína hipotética (457), YP_001942216.1	[[ <i>Chlorobium limicola</i> DSM 245]	9e-23	32/48	Bacteria/Proteobacteria
				2 (1051...1482 direct)	143	Proteína hipotética TPR repeat protein (SEL1 SUPERFAMILY (327), YP_002149794.1]	[ <i>Proteus mirabilis</i> HI4320]	5e-13	47/69	
				3 (2361...2723 direct)	120	No conocida				
				4 (3105...3902 direct)	265	Proteína dominio MOSC (263), EDX86104	<i>Synechococcus</i> sp. PCC 7335	6e-65	44/65	
				<b>5 (4029...4526 reverse)</b>	<b>165</b>	<b>Gentamicin 3'-N-acetyltransferasa (157) , YP_002513892.1</b>	<b><i>Thioalkalivibrio</i> sp. HL-EbGR7]</b>	<b>5e-52</b>	<b>69/78</b>	
				6 (4704...5132 reverse)	142	GCN5-related N-acetyltransferasa (140), YP_563491	<i>Shewanella denitrificans</i> OS217	4e-40	55/76	
				7 (6454 ...6831 reverse)	125	Proteinadioxygenasa de resistencia a Glyoxalasa/bleomicina (140), ZP_01108256	<i>Flavobacteriales bacterium</i> HTCC2170	2e-22	43/65	
				8 (6963...7553 reverse)	196	Proteína de Resistencia aTellurito TehB (197), YP_392289	<i>Thiomicrospira crunogena</i> XCL-2	1e-67	61/76	
				9 (>7767...8046 reverse)	>92	Proteinadioxygenasa de resistencia a Glyoxalasa/bleomicina (133), YP_001869281	<i>Nostoc punctiforme</i> PCC 73102	2e-12	46/59	
GmE1 (7 Kbp/ 4,5 ) 1775 nt	EEZ	66.6	FN640468	1 (>1-856 direct)	>286	Proteína familia glycosyl transferasa(413), YP_771367	<i>Rhizobium leguminosarum</i> bv. <i>viciae</i> 3841	5e-77	54/65	Bacteria/Proteobacteria/Betaproteobacteria
				<b>2 (1071-1775 reverse)</b>	<b>&gt;234</b>	<b>aminoglicosido 3'-N-acetyltransferasa 262), AF486581_1</b>	<b>[uncultured bacterium]</b>	<b>8e-81</b>	<b>69/76</b>	
Clonres de resistencia a Cm										
CME1 (9 Kbp/ 4,428)	EEZ	62.1	FN640469	1(1...222 reverse)	>74	probable reductasa (194), ZP_01004299.1	<i>Loktanella vestfoldensis</i> SKA53	2e-18	58/73	Bacteria/Proteobacteria/Alphaproteobacteria
				<b>2(501...1706 reverse)</b>	<b>401</b>	<b>Transportadores de la subfamilia Bcr/CfIA</b>	<b><i>Mesorhizobium loti</i> MAFF303099]</b>	<b>1e-135</b>	<b>63/76</b>	



Anexo al capítulo II-Resistoma

				3 (1891...2538)	215	<b>(401), NP_103052.1</b> probable regulador transcripcional (229), TetR NP_103053	<i>Mesorhizobium loti</i> MAFF303099]	6e-51	50/64	
				4 (2686-3429 reverse)	247	Proteína relacionada con la biosíntesis del Pyridoxal phosphato PdxJ (247), ZP_02166111.1]	<i>Hoeflea phototrophica</i> DFL-43	5e-87	65/76	
				5 (3507-4061 reverse)	184	Subunidad reguladora de acetolactato sintetasa 3(190), NP_386219.1]	<i>Sinorhizobium meliloti</i> 1021	4e-73	71/82	
				6 (4243-4428 reverse)	>61	acetolactato sintetasa, subunidad grande (592), YP_002500595.1	<i>Methylobacterium nodulans</i> ORS 2060	7e-16	62/77	
CME6 (4 Kbp/ 4003)	EEZ	63.5	FN640470	1(>1-...606 reverse)	>202	probable regulador transcripcional (229), TetR NP_103053	[ <i>Mesorhizobium loti</i> MAFF303099]	3e-53	55/70	Bacteria/Proteobacteria/Alphaproteobacteria
				<b>2 (796...2013 direct)</b>	<b>405</b>	<b>Transportadores de la subfamilia Bcr/CfIA (399), YP_001355014.1</b>	<b><i>Mesorhizobium loti</i></b>	<b>7e-116</b>	<b>57/74</b>	
				3(2187...2651 direct)	154	No conocido				
				4 (2651...2878 direct)	75	Regulador transcripcional familia XRE (71), YP_001685144.1	<i>Caulobacter</i> sp. K31	6e-23	75/85	
				5 (2936...3886 direct)	316	Dominio de unión al Zn Alcohol dehidrogenasa (327), YP_002482726.1	<i>Cyanothece</i> sp. PCC 7425	2e-88	52/68	
<i>Clones de Resistencia a Tm</i>										
TmE1 (4,5 Kbp/ 2336)	EEZ	61.3	FN640471	1(>2...232 direct)	>76	UDP-glucose 4-epimerasa (298), YP_002979022.1]	<i>Rhizobium leguminosarum</i> bv. trifolii WSM1325	5e-27	76/85	Bacteria/Proteobacteria/Alphaproteobacteria
				2(332...1126 direct)	264	thymidylate synthasa (264) NP_386233.1	[ <i>Sinorhizobium meliloti</i> 1021]	1e-136	86/91	
				<b>3(1123...1653 direct)</b>	<b>176</b>	<b>Dhfr dihydrofolate reductasa (175), NP_355014.1]</b>	<b><i>Agrobacterium tumefaciens</i> str. [C58]</b>	<b>2e-60</b>	<b>65/76</b>	
				4(1794...2336 direct)	>181	HflK hidrolasa serin proteasa Proteína K subunidad de transmembrana (361), YP_001979067.1	<i>Rhizobium etli</i> [CIAT 652]	4e-59	67/82	

Anexo al capítulo II-Resistoma

TmE5 (6,5 Kbp/ 1844)	EEZ	59.8	FN640472	1 (124-1239reverse)	371	HflK hidrolasa serin proteasa Proteína K subunidad de transmembrana (376), YP_674308.1]	[ <i>Mesorhizobium sp.</i> <i>BNC1</i> ]	1e-89	52/65	Bacteria/Proteobacteria/
				<b>2 (1300...1824 reverse)</b>	<b>174</b>	<b>Dhfr dihydrofolato reductasa (246), YP_001174482.1</b>	<b><i>Pseudomonas stutzeri A1501</i></b>	<b>5e-41</b>	<b>46/66</b>	
TmE25 (6,8 Kbp/ 3396)	EEZ	61.1	FN640473	1 (2...358 reverse)	>119	Subunidad proteasa hflK (371), NP_103040.1	<i>Mesorhizobium loti</i> <i>MAFF303099</i>	5e-19	51/63	Bacteria/Proteobacteria/Alphaproteobacteria
				<b>2 (420...926-reverse)</b>	<b>168</b>	<b>Dhfr dihydrofolate reductasa (170), ZP_04590822.1</b>	<b><i>Pseudomonas syringae pv. oryzae str</i></b>	<b>3e-41</b>	<b>47/69</b>	
				3 (944...1432 reverse)	163	Acetyltransferasa, familia GNAT (160), YP_002131720.1	<i>Phenyllobacterium zucineum</i> <i>HLK1</i>	3e-50	63/72	
				4(1542...2543 reverse)	334	Proteína hipotética (350), YP_001525236.1	<i>Azorhizobium caulinodans</i> <i>ORS 571</i> ]	1e-74	44/60	
Tm8-4 (7,5 Kbp) 1490 nt sequence	65,64	Orf 1 (>1-612 reverse)	>204	1e-12	42/57	Proteína reguladora de respuesta de dos componentes	YP_002545888.1	[ <i>Agrobacterium radiobacter</i> <i>K84</i> ]		
		Orf 2 (678-1427 reverse)	249	3e-40	41/58	<b>3-oxoacyl-(acyl-carrier-protein) reductasa</b>	ZP_01470074.1	[ <i>Synechococcus sp.</i> <i>BL107</i> ]		

<sup>a</sup> Calculado en la página web: [http://www.genomicsplace.com/cgi-bin/gc\\_calculator.pl](http://www.genomicsplace.com/cgi-bin/gc_calculator.pl).

<sup>b</sup> Los ORFs correspondientes a los genes de resistencia al antibiótico usados para la selección se indican en negrita.

<sup>c</sup> Datos obtenidos de la base de datos del NCBI (Febrero 2010).

<sup>d</sup> Clasificación filogenética de cada fragmento obtenida con el programa phylopythia (29).

**Tabla AII.2.2.** Cebadores utilizados en este estudio.

Nombre	Secuencia	Aplicación	Referencia
T3	5'-d(ATTAACCTCACTAAAG)-3'	Primer secuenciar pBCK-CMV	Stratage
T7	5'-d(TAATACGACTCACTATAGGG)-3'	Primer secuenciar pBCK-CMV	Stratagene
SP6	5'-d(CATACGATTTAGGTGACTATAG)-3'	Primer secuenciar pBCK-CMV	Stratagene
Gme1a	5'-CGAAGTCACAAGCAACGGG-3'	Secuenciar clon GmE1	En este trabajo
GmE1b	5'-AAGCTGCCGTTGTGCGGG-3'	Secuenciar clon GmE1.	En este trabajo
GM83 559r	5' AGAGAAAGTAGGCTGTATTCC -3'	Secuenciar clon Gm8.3.	En este trabajo
GM83 597f	5'- TGAGCCAATCATTGGGGTC-3'	Secuenciar clon Gm8.3.	En este trabajo
Gm83cf	5'- GCTGTATTGGTAGAAGGTCTG -3'	Secuenciar clon Gm8.3.	En este trabajo
GM83 b	5'- CTGTTGCTAACGGAAGCAGG -3'	Secuenciar clon Gm8.3.	En este trabajo
GM83 x	5'- TGGCGACCGCTGCTACGC -3'	Secuenciar clon Gm8.3.	En este trabajo
GM83y	5'- CCGGCCGCTTTCCAACAT -3'	Secuenciar clon Gm8.3.	En este trabajo
GM83z	5'- GCGTGGCGGTGCGCAAAG -3'	Secuenciar clon Gm8.3.	En este trabajo
Ap68b 636f	5'- ATTGCGGGCGAGCGCGTT -3'	Secuenciar clon Ap6.8	En este trabajo
Ap68c 85r	5'- GGCCGATTTCTATGCAGAG -3'	Secuenciar clon Ap6.8	En este trabajo
ApE1a	5'- CCCGATGTCGTTTCGTCGCT -3'	Secuenciar clon ApE1	En este trabajo
ApE14z	5'- CCACTCGCCCGGGCGAC -3'	Secuenciar clon ApE14	En este trabajo
ApE14a	5'- CTCATTTGCGATGGAAAACAGG-3'	Secuenciar	En este

		clon ApE14	trabajo
ApE14x	5´- ATCCAGTCCTTCCGCTGCGG-3´	Secuenciar clon ApE14	En este trabajo
ApE14c	5´- ATTATTTGGAGCCTAGCTTGGT-3´	Secuenciar clon ApE14	En este trabajo
ApE14y	5´- GGTGATGGCTTCATCTCCGTT-3´	Secuenciar clon ApE14	En este trabajo
CmE1z	5´-TCTGGTATATGGCAGGGATAG-3	Secuenciar clon CmE1	En este trabajo
CmE1b	5´-CAGCACTGCAGTGGTCGG-3´	Secuenciar clon CmE1	En este trabajo
CmE1y	5´-AGCGTGCCACGACCTCGT´-3´	Secuenciar clon CmE6	En este trabajo
Cm6a:	5´-GGCTCACTTCCATAACGAGG-3´	Secuenciar clon CmE6	En este trabajo
Cm6z	5´-TTACCGACGGCCTCGACC-3´	Secuenciar clon CmE6	En este trabajo
CmE6y	5´- GCCGACGAATGCTGTTTGA-3´	Secuenciar clon CmE6	En este trabajo
TmE1a	5´-TGAAGGGCGATACGAACGTC-3´	Secuenciar clon TmE1	En este trabajo
TmE1z	5´-CGGGCATTACGCCACGCA-3´	Secuenciar clon TmE1	En este trabajo
TmE25z	5´-TTCGGGAGGATGGGATGAG-3´	Secuenciar clon TmE25	En este trabajo
TmE25x	5´-GCAAGGCCTTGCTGGTCAA-3´	Secuenciar clon TmE25	En este trabajo
TmE25y	5´-CAGCAAGTTCCGGCTCGTGG-3´	Secuenciar clon TmE25	En este trabajo
1427Tm8 4r	5´- GCGCATATGATGACGACCACCTTGGCGG-3´	Secuenciar clon Tm8.4	En este trabajo
678Tm84 f	5´- GCGCATATGATGACGACCACCTTGGCGG-3´	Secuenciar clon Tm8.4	En este trabajo
1427 Tm8-4 r	5´-GCGCATATGATGACGACCACCTTGGCGG-3´	Clonar tm8.4pET3a	En este trabajo
<b>678</b> <b>Tm8-4 f</b>	5´-GCGGATCCTCATCATCTGAGGTGCCGCCAC- 3´	Secuenciar clon Tm8.4 en pET3a	En este trabajo



### II3 BÚSQUEDA DE ENZIMAS LIPOLÍTICAS DE INTERÉS INDUSTRIAL EN LA GENOTECA EEZ

Con las genotecas medioambientales construidas, además del estudio del resistoma en los suelos, también se realizó una búsqueda de nuevas enzimas con carácter lipolítico.

#### II3.1. Estudio de bacterias cultivables con actividad lipolítica

Antes de efectuar la búsqueda de enzimas lipasas/esterasas en las tres genotecas construidas (mexico seca y lluviosa y la genoteca EEZ), se llevó a cabo una aproximación preliminar a partir de bacterias cultivables aisladas del suelo de la Estación Experimental del Zaidín que poseían actividad esterasa/lipasa. De esta manera, en este suelo pudimos comparar el distinto tipo de enzimas presentes dependiendo del enfoque (cultivables o metagenómico).

Se partió de 1 gr de suelo húmedo que se resuspendió en sarcosil y se agitó durante una hora a temperatura ambiente, para luego sembrarlo en diferentes medios de cultivo y contar las unidades formadoras de colonias en diluciones seriadas:

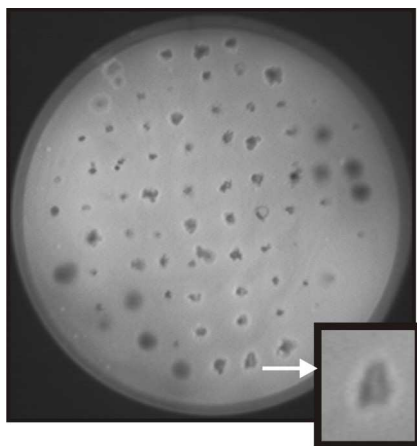
Medio de cultivo TY:  $5,05 \cdot 10^5$ /gr suelo.

Medio de cultivo LB:  $3,0 \cdot 10^5$ /gr de suelo.

Medio de cultivo mínimo MM:  $3,14 \cdot 10^5$  /gr de suelo.

De todos ellos, se escogió el MM, ya que nos iba a permitir posteriormente el análisis de los que presentaran fenotipo lipasa positivo.

A continuación se hizo réplica en placa que contenía medios para probar si eran lipasa positivas (figura II.3.1). Los análisis posteriores revelaron una gran homogeneidad en los tipos de colonias que eran lipasa positivas, la cual fue después confirmada mediante análisis del gen ribosomal 16S.



**Figura II.3.1.** Placa de medio mínimo que contiene trioleina y rodamina B observada bajo luz ultravioleta. Las colonias lipasa positivas se distinguen por un halo de color naranja brillante.

Partiendo del suelo de la EEZ, se aislaron 228 colonias de placas con medio mínimo, de las cuales el 33 % eran lipasa positivas (lip+). Las cepas positivas y algunas negativas aisladas del suelo se sometieron a un proceso de identificación basado en la amplificación de un fragmento de 475 bp del 16SrADN. Los amplicones obtenidos se sometieron a un proceso de TGGE que permitió la identificación mayoritaria de dos tipos de amplicones, ambos pertenecientes al género *Pseudomonas* dentro del phylum  $\gamma$ proteobacteria (Tabla II.3.1).

Grupo	Taxón	Aislado	Presencia	Fenotipo Lip
$\alpha$ -Proteobacteria	<i>Pseudomonas</i> K49.23	sp EEZ1-5-type	4% (9/228)	+
	<i>Pseudomonas</i> sp.AZ22L5	EEZ1-10-type	29% (67/228)	+/-

**Tabla II.3.1.** Aislados del suelo de la Estación Experimental del Zaidín.

El análisis de la actividad *in vitro* presente en los extractos mostró que no todos los aislados pertenecientes a un taxón determinado eran positivos en la actividad lipasa (Tabla II.3.3).

Microorganismo	$\mu\text{molmin}^{-1}\text{mg}^{-1}$ Triacetin	pH <sub>opt</sub>	T <sub>opt</sub> (°C)
<b>ESTERASA/LIPASA</b>			
EEZ1-5 <i>Pseudomonas</i> sp AZ18R1	1039 ± 101	8,0-8,5	40
EEZ1-10 <i>Pseudomonas</i> sp AZ22L5	939 ± 104	8,5	45
EEZ1-14 <i>Pseudomonas</i> sp AZ22L5	160 ± 14	8,5	40
EEZ2-40 <i>Pseudomonas</i> sp AZ22L5	190 ± 11	8,6	50
<b>Datos publicados</b>			
<i>Pseudomonas</i> sp. Strain KB700A	150	8,0-8,5	35
<i>Candida cylindracea</i>	<b>4570</b>	7,4	45
<i>Bacillus pumillus</i> B62	138	8,5	35
<i>Staphylococcus warnei</i>	25	7,0	25
<i>Aspergillu niger</i>	208	5,5	35-55
<i>Pseudomonas cepacia</i>	<b>4180</b>	7,5	60
<i>Penicillium cyclopium</i>	193	7,0	40
<i>Bacillus thermoleovorans</i> ID-1	116	7,5	70

**Tabla II.3.3.** Comparación de la actividad específica de esterasas/lipasas de bacterias del suelo de la EEZ frente a otras obtenidas de distintas fuentes.

1 La actividad específica se presenta en  $\mu\text{molmin}^{-1}\text{g}^{-1}$  de proteína total o en  $\mu\text{molmin}^{-1}\text{g}^{-1}$  de células liofilizadas expresando la esterasa o lipasa. La actividad hidrolítica se determinó mediante la titulación de los ácidos grasos libres tras la hidrólisis de triglicéridos (desde triacetina hasta trioleina) usando como método el pH-stato. La hidrólisis de los sustratos (35 mM) se ensayó a 40°C en 100 mM Tris HCl a pH 8.0 usando una solución de NaOH 0.1 M para titular el ácido graso liberado. La mezcla de reacción (20 ml) contenía el sustrato, 0,15 M de NaCl y 0,09% de acetonitrilo 0,09% (v/v). Una unidad de enzima se define como la cantidad que libera 1  $\mu\text{mol}$  de ácido graso libre por minuto.

De todos los aislados lipasa positivos, el que mayor actividad presentó fue el aislado EEZ1-5, y fue el que se utilizó para estudios posteriores de estereo-especificidad (Tabla II.3.2b).

Substrate	Positive samples
Solketal acetate	–
Geranyl acetate	EEZ1-5 (R)
Neomenthyl acetate	–
Methyl mandelate	–
Menthyl acetate	EEZ1-5 (S),
Methyl(R)-3-bromo-2-methylpropionate	–
N-benzyl-proline ethyl ester	EEZ1-5 (L)
Tryptofan methyl ester	–
Glycidyl-4-nitrobenzoate	–
Pantolactone	–
Dihydro-5-hydroxymethyl-2(3H)-furanone	–

**Tabla II.3.4:** Perfil de sustrato de las esterasas aisladas de EE1-5. Actividad frente a ésteres enantioméricamente puros. Sólo las muestras con exceso de actividad sobre el enantiómero superior al 90% y un ratio mayor de 20 se muestran como positivas.

De este modo, vemos como la esterasa aislada de EEZ1-5 presenta propiedades útiles para la resolución de posibles compuestos enantioméricos en procesos de síntesis, en concreto, para los sustratos: Geranyl acetato, Menthyl acetato y N-benzyl-proline ethyl ester (Tabla II.3.2b).

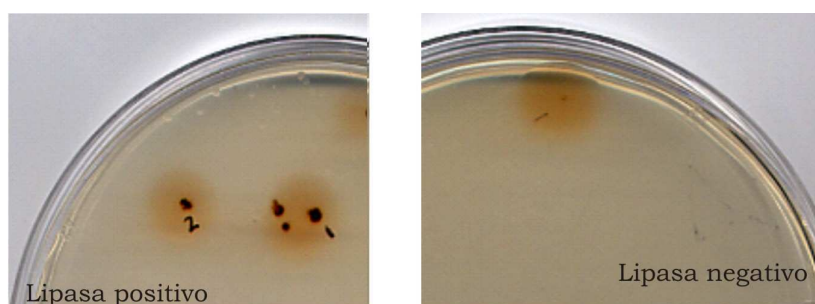


### II3.2. Búsqueda de genes con actividad lipolítica en la genoteca metagenómica EEZ

Con las genotecas metagenómicas construidas realizamos una búsqueda de genes de carácter lipolítico que se estuvieran expresando desde el ADN clonado en el vector del fago lambda. De este modo, se hizo un ensayo de actividad esterasa en placa, directamente sobre las calvas de lisis provocadas por el fago, tal y como se detalla en el Material y Métodos de esta tesis doctoral.

En la genoteca construida a partir del suelo de la Estación Experimental del Zaidín de 65.000 pfus de tamaño (ver tabla II.1.1), realizamos diferentes búsquedas; se hizo un primer test de  $3 \cdot 10^5$  pfu/placa de petri (la mitad del tamaño de la genoteca) que testaba 10 placas (aproximadamente 5 veces la genoteca inicial). En este primer test se identificaron unos 28 posibles fagos que contenían insertos que revelaban un fenotipo esterasa positivo. Tras el segundo muestreo (después de la amplificación de cada una de las calvas), sólo 12 se confirmaron como tales.

Posteriormente, se procedió a la escisión de los fagos y al ensayo esterasa en colonia para corroborar los resultados de las calvas de lisis ( figura II.3.2). Se pudieron escindir a plásmido todos los fagos excepto uno, que tras repetir los ensayos varias veces fue imposible de escindir. Probablemente, el fragmento de ADN clonado en él contenga alguna secuencia que sea letal para la célula de *E.coli*.



**Figura II.3.2.** Confirmación del ensayo esterasa en fagémidos escindidos a plásmido. La placa de la izquierda contiene colonias resultado de la escisión a plásmido del clon Est1. La placa de la derecha contiene colonias esterasa negativas. El fenotipo esterasa positivo se resuelve mediante un precipitado oscuro.

### **II.3.2.1. Análisis de las secuencias de los genes de los clones con actividad lipolítica**

Los fagos con actividad esterasa fueron aislados de la placa e incubados en medio SM y cloroformo (para lisis de las posibles células que hubiéramos tomado). A continuación, se sembraron individualmente para confirmar esta actividad esterasa. Cada uno de los clones positivos aislados se escindió a plásmido siguiendo las instrucciones del kit (plásmido pBK-CMV de 4.5 kb; figura 4 de Material y Métodos) y se elaboró un primer mapa físico de los fragmentos clonados mediante digestiones con enzimas de restricción. De este modo, además, se comprobó que de los 12 fagos sólo existían 8 diferentes, por lo que finalmente nos quedamos con 8 clones con actividad esterasa positiva procedentes de esta genoteca.

Una vez establecidos los mapas físicos (ver figura anexo AII.3.1), se generaron subclones, disminuyendo así el tamaño de inserto y analizando si conservan o no la actividad esterasa positiva.

La secuenciación de cada uno de los clones y subclones mediante los cebadores universales correspondientes al sitio de clonaje del plásmido pBCK (Stratagene), así como el uso de cebadores internos mediante “*primer walking*”, permite obtener la secuencia de cada uno de los clones del gen causante de la actividad esterasa.

Las secuencias de estos genes corresponden a enzimas diversas que pertenecen a distintos grupos o familias de esterases (Arpigni and Jaeger, 1999; tabla II.3.5). Los genes adyacentes encontrados no están conservados en ninguno de los casos. Esto indica la poca sintenia que presentan estas muestras y es reflejo de la diversidad presente en el suelo y en particular en nuestras genotecas.

La adscripción funcional de algunos de los clones en base a su homología de secuencia puede reflejar, por un lado, la comunidad microbiana presente y el posible papel que pueden estar cumpliendo estas enzimas en el suelo (figura anexo AII.3.1). En todos los clones se ha determinado la secuencia del gen esterasa y de al menos 1 a 2 kb aproximadamente del entorno de dicho gen.

En el caso del contexto del clon 1, aguas abajo encontramos de manera divergente un nuevo marco abierto de lectura (tabla II.3.5) y, entre ambos, una secuencia nucleotídica en forma de horquilla de unos 70 nt (probablemente un

sitio de terminación de la transcripción que además complicó su secuenciación).

El clon 6 posee varios genes juntos relacionados funcionalmente, los cuales forman un cluster. Guardan homología con las estructuras génicas descritas en la degradación del protocatecol y del catecol, aunque no tienen el mismo orden en el cluster (Iwagami *et al*, 2000).

En el caso del clon 7, éste posee varios genes esterasa en prácticamente 3 kb de inserto y podrían ser un posible cluster de esterasas. Además, dos de los tres genes que porta el clon (debido a que el tercero está truncado) son funcionales. Esto se ha comprobado mediante la construcción de subclones que sólo tenían uno de los genes, tanto el subclon 7.1 como el 7.2 son esterasa positivos (figura 10-Anexo).

De este modo, como se observa en la tabla II.3.5, de la genoteca EEZ se aislaron 8 clones con actividad esterasa y se han identificado 9 enzimas diferentes.

De estas, tres pertenecen a la familia IV-HSL (hormone sensitive family), que son enzimas que tienen una alta homología con la familia de las lipasas sensibles a hormonas de mamíferos; tres a la familia II (G-D-S-L); dos a la familia VIII y una no se pudo clasificar de forma clara en ninguna de las familias que existen hasta el momento. Estas enzimas poseen unos pesos moleculares que están en el rango de entre 25 y 50 Kda; un punto isoeléctrico variable, desde neutro hasta ácido en algunos casos (como la enzima pE4 con un punto isoeléctrico de 4.3), o claramente básico (por ejemplo la enzima pE1 con un punto isoeléctrico de 9.5) y bajos niveles de homología con las secuencias de las enzimas depositadas en las bases de datos (con rangos de identidad por debajo del 60%).

Clon	Tamaño	Nombre	Enzima esterasa							Microorganismo
			Número aa	Pm(kDa)	Pi <sup>a</sup>	enz.más homología <sup>b</sup>	Familia lipasas	Identidad <sup>c</sup>	E-valor <sup>d</sup>	
1	6 kb	pE1	238	26	9,535	Mucin-desulfatasa,multidominio	II (GDSL)	45%	7.e-49	<i>Blastopirellula marina</i>
2	8 kb	pE2	248	43,9	8,74	lypolitic enzyme	II (GDSL)	58%	2.e-118	<i>Caulobacter crescentus</i>
3	10 kb	pE3	473	51,8	9,76	Sialato-O-acetilesterasa	Sin clasificar	54%	2.e-148	<i>Flavobacterium john.</i>
4	8 kb	pE4	308	32,8	4,29	Esterasa/lipasa/thioester.	IV (HSL)	52%	3.e-68	<i>Ralstonia eutropha</i>
6	3 kb	pE6	379	41,35	4,62	Lactona hidrolasa	VIII	61%	8.e-129	<i>Rhodococcus ruber</i>
7	3 kb	pE7.1	341	36,63	6,84	Carboxylesterasa	IV (HSL)	37%	2.e-45	<i>Bacillus halodurans</i>
	N-term	pE7.2	420	44,57	7,05	Esterasa	VIII	42%	5.e-53	<i>Streptomyces averm.</i>
	parcial		185	/	/	Esterasa/lipasa/thioester	IV (HSL)	42%	1.e-30	<i>Ralstonia eutropha</i>
8	7 kb	pE8	418	43,99	8,74	Lypolytic enzyme G-D-S.L	II (GDSL)	46%	7.e-47	<i>Caulobacter crescentus</i>
12	7 kb	pEe12	309	34,96	6,01	Esterasa/lipasa	IV (HSL)	43%	1.e-57	<i>uncultured bacterium</i>

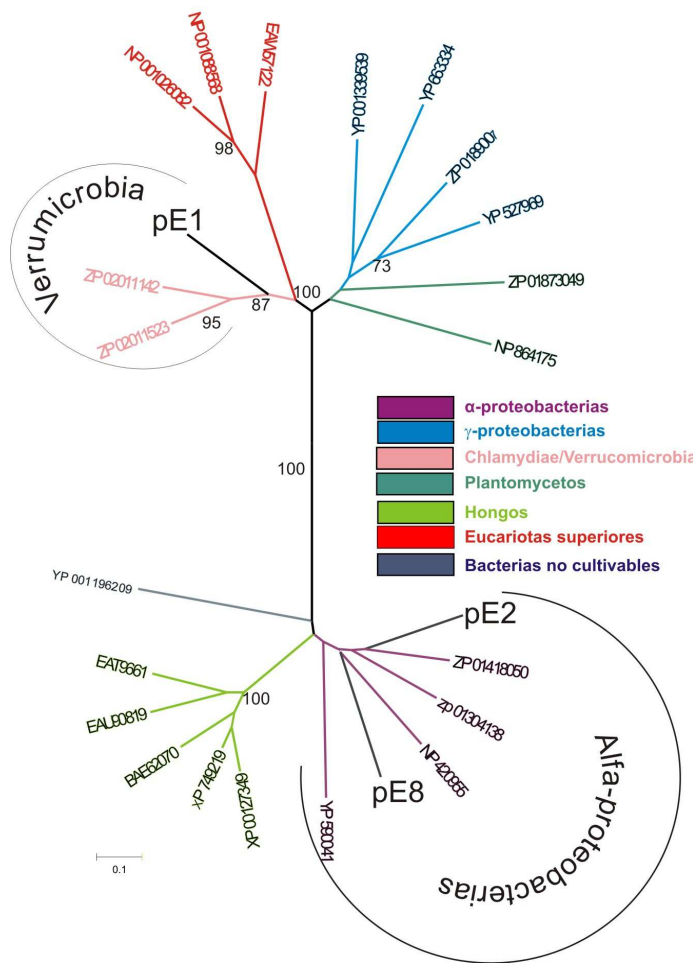
**Tabla II.3.5.** Características de los clones de la genoteca metagenómica que portan los genes con actividad esterasa.

- a) Punto isoeléctrico ([http://scansite.mit.edu/calc\\_mw\\_pi.html](http://scansite.mit.edu/calc_mw_pi.html)).
- b) Enzima con la que guarda más parecido en los alineamientos (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).
- c) Porcentaje de identidad de aa. Es un valor numérico que indica cuanto se parecen entre sí.
- d) Representa el número de alineamientos con puntuación igual o equivalente a la que podría haber sido encontrada sólo por azar.

**II3.2.2. Relaciones filogenéticas de las enzimas aisladas**

Una vez conocida la secuencia de cada uno de los clones y estudiadas sus características, se procedió al análisis de las secuencias de aminoácidos para la construcción de árboles filogenéticos que nos ayudaran a definir, por un lado, su origen en relación a familias génicas y por otro, predecir el posible microorganismo o taxón del cual provienen (figura II.3.3, 5 y 7).

**- Familia II.**

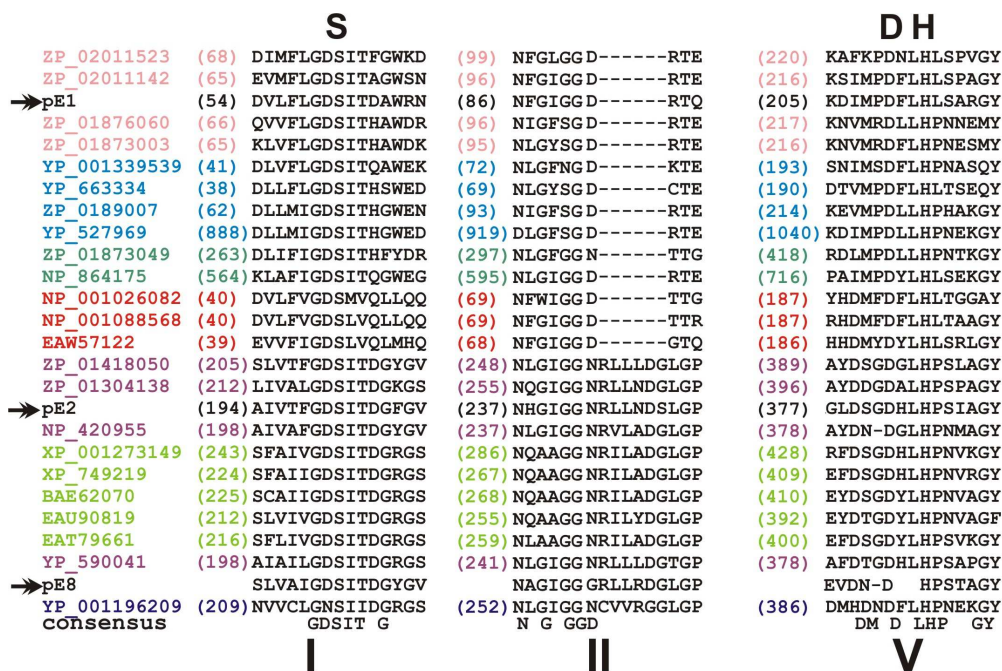


**Figura II.3.3** Relaciones filogenéticas entre las secuencias de las proteínas esterasa pE1, pE2 y pE8 obtenidas en este estudio y secuencias de referencia de la clase de lipasas II-GDSL procedentes de la base de datos. El árbol se construyó considerando las secuencias representativas con un valor de e inferior a e-38 resultantes del alineamiento en la base de datos y usando el método *Neibor Joining*. Los números sobre las ramas del árbol indican los valores de *bootstrap* cuando se realizan 100 repeticiones. La escala situada abajo a la izquierda es proporcional a la longitud de las ramas del árbol. En código de color posible adscripción a distintos phyla.

En la figura anterior, se observa un árbol con las enzimas aisladas de la genoteca metagenómica que pertenecen a la familia II (pE1, pE2 y pE8) y otras que guardan homología con éstas de la misma familia. De este modo, vemos como la pE2 y pE8 están muy relacionadas entre ellas y a su vez con algunas enzimas de esta familia encontradas en  $\alpha$ -proteobacterias, mientras que la

pE1, aunque posee parecido con varias proteínas de eucariotas superiores, se encuentra muy próxima a proteínas similares en procariotas, concretamente a las del grupo *Verrumicrobia*, que es un grupo muy poco caracterizado en cuanto a bacterias cultivables se refiere.

Este tipo de enzimas, las esterasas de la clase II, se denominan GDSL/SGNH enzimas porque no tienen el sitio activo clásico (G-X-S-X-G) y en su lugar poseen el tetrapéptido G-D-S(L), localizado en la parte N-terminal de la proteína (Arpigny and Jaeger, 1999), y cinco bloques conservados. En las enzimas pE1, pE2 y pE8 podemos señalar tanto los cinco bloques conservados como los residuos conservados del sitio activo (figura II.3.4).

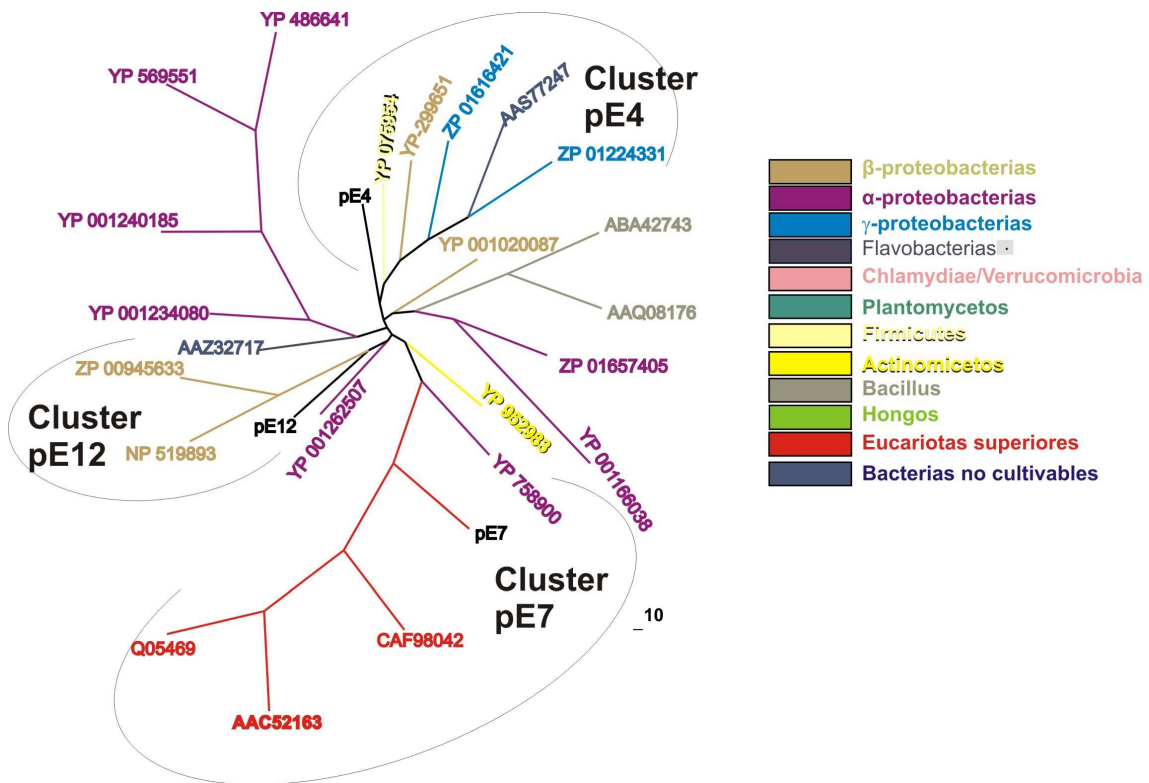


**Figura II.3.4.** Alineamiento de las proteínas pE1 y pE2 con otras enzimas pertenecientes a la familia GDSL de la base de datos e identificación de los residuos conservados. El alineamiento se realizó usando la herramienta ClustalW (<http://mafft.cbrc.jp/alignment/software/>). El código de color similar al utilizado en la figura II.3.3.

**- Familia IV**

En el caso del árbol filogenético con las proteínas pertenecientes a la clase IV (pE4, pE7.1 y pE12; figura II.3.5), observamos cómo estas tres proteínas, aunque pertenecen a una misma clase, se sitúan en clusters totalmente diferentes. La pE7.1 está relacionada con proteínas similares a algunas encontradas en eucariotas superiores, mientras que en el caso de la pE4 y pE12 no existe una relación clara entre este tipo de proteínas y otras de algún phylum en concreto.

La familia IV (HSL) de esterazas posee tres bloques con motivos conservados (Arpigny and Jaeger, 1999). Los bloques II y III son los que contienen los residuos de la tríada catalítica (figura III.3.6). El bloque I contiene una secuencia conservada H-G-G-G que está relacionada en las interacciones mediante puentes de hidrógeno que promueven la catálisis.



**Figura II.3.5.** Relaciones filogenéticas entre las secuencias de las proteínas esterasa pertenecientes a la familia IV (pE4, pE12 y pE7.1) y secuencias de referencia depositadas en las bases de datos. Para la construcción de los árboles, sólo se consideraron secuencias representativas con un valor de  $e$  inferior a  $e^{-58}$  de homología procedentes de la base de datos. El árbol se construyó usando el método *Neighbor Joining*. Los números sobre las ramas del árbol indican los valores de *bootstrap* cuando se realizan 100 repeticiones. La escala situada abajo a la izquierda es proporcional a la longitud de las ramas del árbol.

			S	D	H				
YP-299651	(79)	CHGGGWVVDG	(149)	AVGGDSAGGN (236)	RGLPPAFVVT	AEFDPLRDEG	EAYARRLAEA	GTFVECKRYD	GTIHGFC
YP_075854	(80)	FHGGGWVLGS	(150)	AVAGDSAGGN (235)	SGLPPAFVIT	AEYDPLRDEG	EAYARKLIEA	GSAAVTVKRYL	GQIHGFC
ZP_01224331	(93)	FHGGGWVIGD	(163)	SVGGDSAGGN (251)	STLPPALMMT	AEFDPLRDEG	EAYAQRLKAA	GVEVTVRRFD	GLVHGFF
AA577247	(81)	FHGGGWVIGN	(151)	AVGGDSAGGN (251)	AGLPPAFVLT	AEFDPLRDEG	EAYGQELRKA	GVAVTAKRYD	GVCHGFV
ZP_01616421	(76)	FHGGGWVIGD	(145)	AVGGDSAGGN (239)	TNLPATIIIT	AEFDPLRDEG	EAYAARLVAA	GNTVTVKRFD	GVVHGFF
→ pE4	(80)	LHGGGWVIGG	(150)	AVGGDSAGGN (237)	SGLAPATVIT	AEYDPLRDEG	DAYAAKLTDA	GVDVEHLRYA	GQIHGFF
ABA42743	(79)	YHGGGWVLGS	(148)	AVGGDSAGGN (237)	SSLPSATIIIT	AQYDPLRDSG	KDYADALKNH	GVPVTYENYE	TMIHGFL
AAQ08176	(79)	YHGGGWVLGS	(148)	AVGGDSAGGN (237)	SSLPSATIIIT	AQYDPLRDSG	KDYADALKNH	GVPVTYENYE	TMIHGFL
YP_001166038	(80)	YHGGGWVIGS	(149)	AVAGDSAGGN (209)	AGLPPATVIV	AQYDPLRDEG	LAYAEALKAA	GVPVTEELAP	GMIHGFF
ZP_01657405	(83)	YHGGGWVIGD	(153)	AVAGDSAGGN (214)	TGLPPAYVVT	AGFDVLRDEG	KAYAEALKKA	GVEVEYVNYE	GMIHGFF
YP_001020087	(81)	FHGGGFVFGN	(150)	AVAGDSAGGN (237)	KGLPAAATVIT	AGYDPLRDEG	DAYAKALAKA	GVAVDHVPNP	TMIHGFF
YP_486641	(88)	FHGGGWVIGN	(157)	SVGGDSAGGN (245)	AGLPPAFVIT	AGADPLRDEG	DEYARRLADA	GVPVTHRTYP	GQFHGFF
YP_569551	(88)	FHGGGWVIGN	(157)	SVGGDSAGGN (245)	AGLPPAFVMT	AGADPLRDEG	DEYARRLADA	GVPVTHRTYP	GQFHGFF
YP_001240185	(87)	FHGGGWVIGN	(157)	SVGGDSAGGN (245)	AGLPPAFVIT	AGADPLRDEG	DDYARRLREA	GVPVTHRTYP	GQFHGFF
YP_001234080	(83)	FHGGGWVIGD	(153)	AVGGDSAGGN (241)	DGLPPALVIT	AGFDPLRDEG	EAYARRLAEER	GVAVTLRRFP	GQIHGFL
AAZ32717	(83)	YHGGGWVIGD	(153)	AVGGDSAGGN (241)	ANLPPALVIT	AGFDPLRDEG	LHYAQKLESEA	GNRATHISFE	RQIHGFI
NP_519893	(108)	FHGGGFVIGS	(178)	AFGGDSAGGT (269)	RGVCPAWISV	AGFDPIRDAG	IGYADKLRAA	GVPVTLKMYE	GMIHDFD
ZP_00945633	(119)	FHGGGFVIGS	(189)	AFGGDSAGGT (280)	RGVCPAWIAV	AGFDPIRDAG	IDYADKLRAA	GAPVALKVYE	GMIHDFD
YP_001262507	(104)	FHGGGFVYCD	(174)	AIGGDSAGGN (257)	VGLPPAFMLT	AGYDPLRDEA	IDYARRLNEA	GVPVRHSEHR	GQIHGFA
→ pE12	(81)	FHGGGFVIGD	(151)	AVGGDSAGGN (238)	SGAPACLVT	AGFDPLKDEG	KAYAEKLTAA	GVPARHVEHA	GFVHDFY
YP_952983	(138)	YHGGGHVIGS	(208)	AVGGDSAGGN (297)	SGLSPALVVT	AGFDPLRDEG	RHYADALRAA	GTPVDYREYG	SLVHGFA
YP_758900	(84)	FHGGGFVLGD	(154)	AVAGDSAGGN (245)	KGLPPAHIIIV	CGWDPLHDEG	LAYAACKLRAM	GIAVSEKEYP	SMVHGFL
→ pE7	(105)	FHGGGFVVLGN	(175)	SVGGDSAGGN (264)	AELPSIFIVS	AKNDILHSDA	VAMETKLRAL	GVDVGHLLGVE	GAGHGFL
Q05469	(650)	FHGGGFVAQT	(719)	CLAGDSAGGN (891)	KSLPPVHIVA	CALDPMLDDS	VMLARRLRNL	GQPVTLRVVE	DLPHGFL
AAC52163	(347)	IHGGGFVAQT	(417)	CLAGDSAGGN (681)	KTLPPVHLVA	CALDPMLDDS	VMFARRLRDL	GQPVTLKVVE	DLPHGFL
CAF98042	(352)	FHGGGFVAQT	(422)	CLVGDSDAGGN (691)	KGLPPVHLVA	CALDPMLDDS	VMFAKRLRSV	DQPVTLCVVE	DLPHGFL
Consensus		FHGGFWVIG		AVGGDSAGGN	GLPPA	AG DPLRDEG	AYA L A	GVPVT Y G	IHGFC

**Figura II.3.6.** Alineamiento de las proteínas pE1 y pE2 con otras enzimas pertenecientes a la familia IV de la base de datos e identificación de los residuos conservados. El alineamiento se realizó usando la herramienta ClustalW (<http://mafft.cbrc.jp/alignment/software/>).

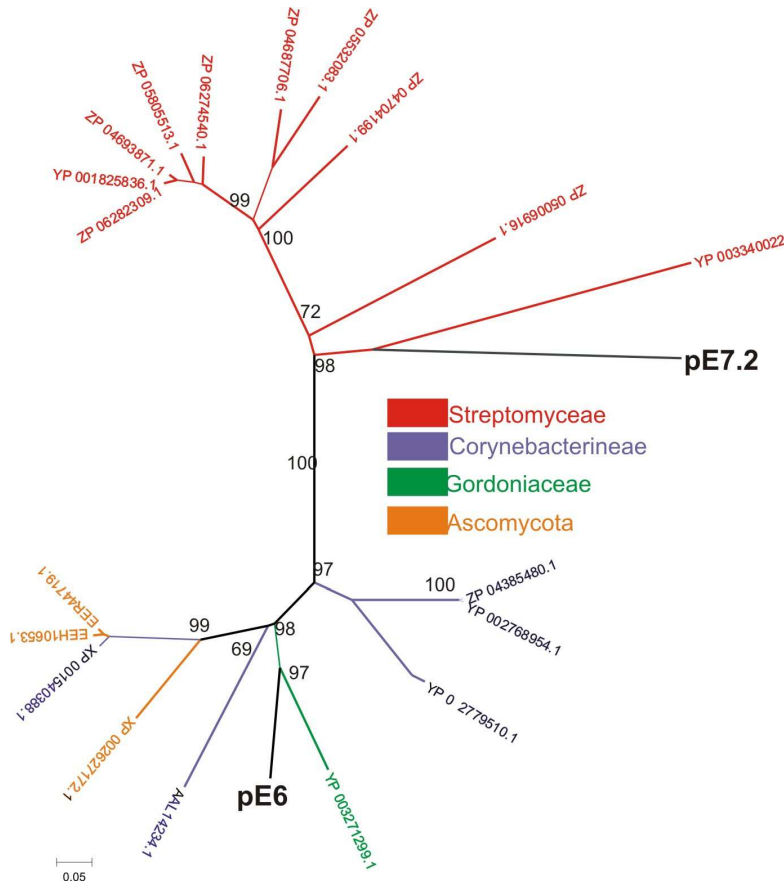
**- Familia VIII**

Finalmente, con respecto a las enzimas pE6 y pE7.2, se puede decir que ambas pertenecen a la familia VIII de lipasas. Ésta es una familia que se caracteriza por diferenciarse significativamente del plegamiento típico de las hidrolasas y poseer mucha homología con las β-lactamasas, ya que poseen un dominio típico de este tipo de proteínas. Se testó su actividad β-lactamasa y ninguno de los clones confiere resistencia a ampicilina en células de *E.coli* (datos no mostrados).

Ambas enzimas se agrupan (figura III.3.7) con otras proteínas encontradas en la clase actinobacteria. La esterasa pE7.2 es diferencialmente distinta a la pE6 y guarda gran homología con otras β-lactamasas pertenecientes al género *Streptomyces*.

En cuanto a la pE6, se agrupa con otras enzimas como β-lactamasas y lactona hidrolasas pertenecientes a los subórdenes de actinobacterias: *Gordoniaceae* y *Corynebacterineae*.





**Figura II.3.7.** Relaciones filogenéticas entre las secuencias de las proteínas esterasa pertenecientes a la familia VIII (pE7.2 y pE6) y secuencias de referencia de la base de datos. El árbol se construyó usando el método *Neighbor Joining*. Los números sobre las ramas del árbol indican los valores de *bootstrap* cuando se realizan 100 repeticiones. La escala situada abajo a la izquierda es proporcional a la longitud de las ramas del árbol.

Una vez identificados los genes causantes de la actividad lipolítica dentro de los clones y analizadas sus secuencias, el siguiente objetivo se centró en el estudio de las características bioquímicas de estas enzimas con el fin de encontrar algunas características potencialmente aplicables en la industria biotecnológica.

### II.3.2.3. Caracterización bioquímica de las enzimas encontradas

Para ello, se partió del plásmido pBK-CMV con el inserto de ADN metagenómico resultante de la escisión del fago cuya actividad lipasa era positiva, y en algún caso con alguno de los subclones.

Este vector contiene el promotor lac, por lo que la expresión de algunos genes contenidos en el inserto puede ser inducida añadiendo IPTG a las células. Esto ha resultado en 7 de los 8 clones obtenidos, ya que se encuentran orientados bajo este promotor. De este modo y como se detalla en el Material y Métodos de esta Tesis Doctoral (sección 13) cada uno de los clones se crecieron en el medio correspondiente hasta llegar a fase exponencial en la cual se le añadió IPTG para conseguir la correcta expresión de los genes.

A las 3 horas de la inducción con IPTG, se recogieron las células, se lisaron y se resuspendieron en el tampón necesario para hacer los posteriores análisis bioquímicos.

*II.3.2.3.1 Ensayos de temperatura y pH óptimos*

Las temperaturas óptimas varían entre las distintas proteínas, desde 20 grados hasta 37 grados. En general, son temperaturas óptimas relativamente bajas para lo que se describe para otro tipo de enzimas obtenidas con estrategias similares (Couto et al, 2010), que suele rondar en torno a los 40°C.

El clon pE7 que portaba dos enzimas (pE7.1 y pE7.2) no se incluyó en el análisis por la dificultad de obtener extracto crudo con las dos enzimas diferenciadas.

<b>EEZ, enzima</b>	<b>Temperatura óptima (°C)</b>	<b>pH óptimo</b>
<b>pE1</b>	27	7
<b>pE2</b>	37	7
<b>pE3</b>	20	8
<b>pE4</b>	15	8,5
<b>pE6</b>	37	7
<b>pE8</b>	10	8
<b>pE12</b>	20	7

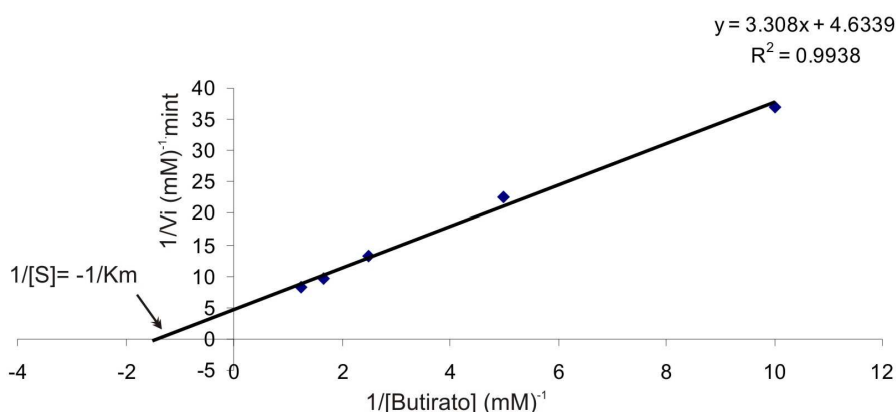
**Tabla II.3.6.** Temperatura y pH óptimos de los extractos crudos de los distintos clones aislados de la genoteca metagenómica EEZ. La temperatura óptima se calculó incubando la enzima en el tampón de actividad enzimática con 50 mM de pNp acetato a diferentes temperaturas durante 15 minutos. El pH óptimo se calculó incubando los extractos en tampones con diferentes pHs. Los ensayos se hicieron por triplicado.

En el ensayo para testar la actividad máxima a diferentes pHs, no se usaron pHs por debajo de 6,8, porque en ambientes ácidos se auto-hidrolizan los derivados p-nitrofenilos y no se pueden determinar bien los valores de absorbancia a 450 nm. En cualquier caso (tabla II.3.6), la mayoría de las enzimas mostraron un pH óptimo cercano al neutro.

*II.3.2.3.2 Determinación de las constantes cinéticas de las enzimas con carácter lipolítico para diferentes tipos de sustratos*

Para la determinación de las constantes cinéticas, se usaron los sustratos para-nitrofenilos cuya hidrólisis podemos medir con un espectrofotómetro a 405 nm, tal y como se detalla en el Material y Métodos de esta tesis.

Para el cálculo de los valores de  $K_m$  y  $V_m$ , se fijó una concentración de extracto con la enzima recombinante constante y se varió la concentración de sustratos entre 1,6 mM y  $4 \cdot 10^{-3}$  mM, midiendo en cada caso la velocidad inicial de la reacción. Se calculó la representación de Lineweaver-Burk para cada enzima y la extrapolación lineal de los datos nos proporcionó tanto  $V_{max}$  como  $K_m$ . Un ejemplo de este tipo de gráfica se observa en la figura II.3.8 en el caso de la enzima pE12 para el sustrato pNp-butirato.



**Figura II.3.8.** Representación de la actividad de la enzima con carácter lipolítico pE12 mediante la ecuación de Lineweaver-Burk. En esta representación doble inversa, se representa  $1/V_i$  frente a  $1/[S]$ . Los valores son la media de experimentos duplicados.

En la tabla II.3.7, se observan los valores de  $K_m$  y  $V_{max}$  para la mayoría de los clones aislados en la genoteca EEZ.

La constante  $K_m$  es un indicador de la especificidad de sustrato. Cuanto mas pequeño sea el valor de  $K_m$ , mayor especificidad sobre un sustrato determinado. Se define de esta manera el tipo de enzima, si es esterasa o lipasa verdadera.

De este modo, observamos cómo las enzimas pE1, pE2, pE4 y pE8 son esterasas verdaderas y tienen una gran especificidad por los ésteres de bajo número de carbonos, fundamentalmente acetato (C2) y butirato (C4). En el caso de los clones pE6 y pE12, sí que se podrían considerar lipasas verdaderas, ya que poseen unos valores de  $K_m$  muy bajos (son muy específicos) para ésteres de alto número de carbonos, como por ejemplo laureato (C12). Remarcable es lo específica que es la enzima pE12 para el sustrato caprilato (C8) (tabla II.3.7).

Km (mM)						
	pE1	pE2	pE4	pE6	pE8	pE12
Acetato	0,4371	0,079	0,2018	0,785	0,12	0,673
Butirato	nd	0,86	0,516	0,341	nd	0,714
Caprilato	nd	nd	nd	0,017	nd	0,00196
Laureato	nd	nd	nd	0,043	nd	0,03
Palmitato	nd	nd	nd	nd	nd	nd

**Tabla II.3.7.** Valores de Km para los extractos de los clones de la genoteca metagenómica EEZ. Combinarla y ponerla conjuntamente la II.3.7 yII.3.8.

Con respecto a la velocidad máxima, esta varía entre la hidrólisis de 11.11 mmol/min de acetato por minuto en el caso del pE1, hasta 0,07 mil moles de laureato por minuto en el caso de la pE12, que es la Vmáxima inferior en el rango de enzimas. En general, las esteras pE1, pE2, pE4 y pE8 tienen una Vmáx muy elevada para el acetato (sobre todo en el caso de la pE1) y las enzimas pE6 y pE12 tienen unas velocidades máximas con valores inferiores, presentando el máximo de hidrólisis con el sustrato caprilato.

Vmax (mmol/min)						
	pE1	pE2	pE4	pE6	pE8	pE12
Acetato	11,11	2,00	0,83	0,04	1,00	0,04
Butirato	nd	0,09	0,02	0,09	Nd	0,22
Caprilato	nd	nd	nd	0,11	Nd	0,44
Laureato	nd	nd	nd	0,05	Nd	0,07
Palmitato	nd	nd	nd	nd	Nd	nd

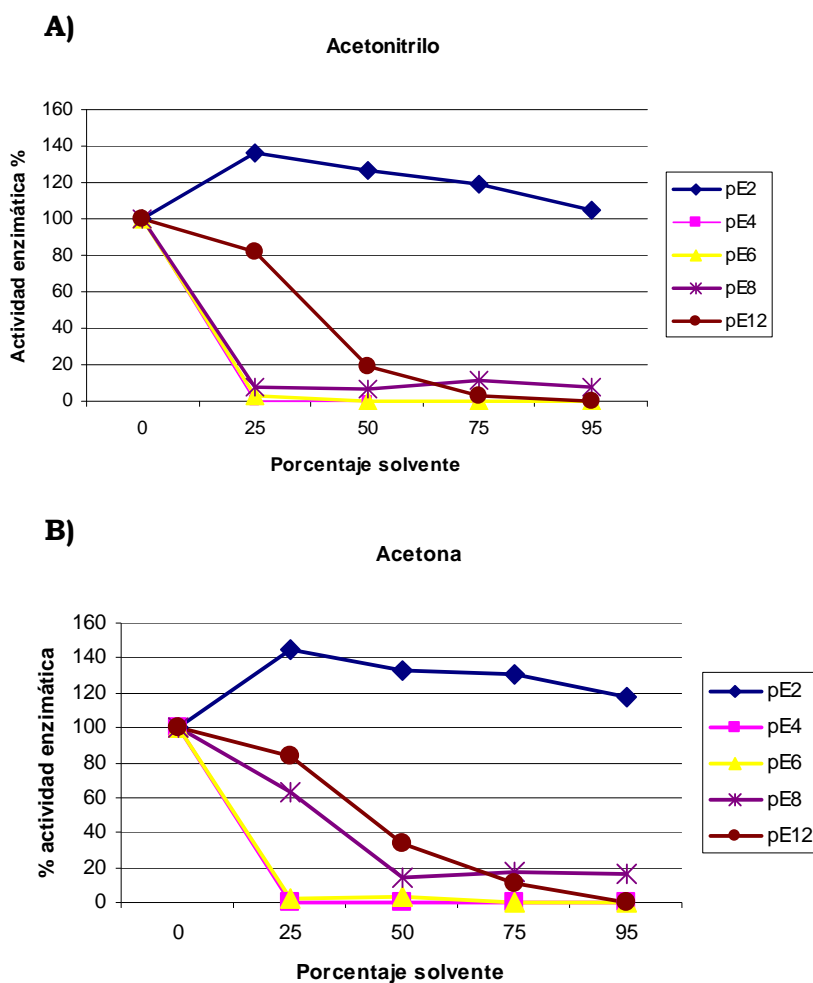
**Tabla II.3.8.** Valores de Vmáx. para las enzimas aisladas de la genoteca metagenómica EEZ.

### II.3.2.3.3. Estabilidad frente a solventes orgánicos y a detergentes

#### **Efecto de los solventes orgánicos en la actividad enzimática de los clones lipolíticos encontrados**

Para llevar a cabo una caracterización bioquímica completa de las enzimas y poder evaluar sus características en una posible aplicación en biotecnología, se ensayó su estabilidad frente a distintas condiciones.

En el caso de la resistencia a solventes orgánicos, se evaluó añadiendo acetona y acetonitrilo en diferentes concentraciones al extracto proteico con cada una de las enzimas a analizar (figura II.3.9).



**Figura II.3.9.** Estabilidad frente a solventes de las enzimas aisladas en la genoteca EEZ **A)** Actividad enzimática residual tras incubar los extractos a diferentes concentraciones de acetonitrilo. **B)** Actividad enzimática residual tras incubar los extractos a diferentes concentraciones de acetona. Los ensayos se realizaron por triplicado. Se observa en la gráfica la media de estas tres réplicas. Para ello, se incubaron los extractos durante 6 horas con diferentes concentraciones de solventes y se calculó posteriormente la actividad enzimática residual frente a un control sin solvente.

De este modo, las enzimas pE4, pE6, pE8 y pE12 resultaron ser muy poco estables frente a acetonitrilo. En el caso de la pE4 y la pE6, una concentración del 25% inhibiría su actividad por completo y, en el de la pE8 y la pE12, a partir del 25% va disminuyendo hasta hacerse prácticamente nula al 75%.

Todo lo contrario ocurre con la enzima pE2, cuya actividad relativa no sólo se mantiene, sino que aumenta incubándola con diferentes solventes. Conserva esta actividad incluso incubándola al 99% con los solventes ensayados.

El efecto de este tipo de sustancias en las reacciones enzimáticas está relacionado con los residuos hidrofóbicos de las proteínas. Las enzimas mantienen su estructura secundaria gracias a interacciones intramoleculares, hidrofóbicas, entre los aminoácidos.

En una solución acuosa tienen los dominios hidrofílicos en contacto con el agua y los hidrofóbicos plegados hacia el interior de la molécula. Cuando la polaridad del medio es baja porque hay un solvente orgánico, los dominios hidrofóbicos se dispersan, desplegándose así la molécula (Klibanov *et al*, 2001). Otro tipo de efecto está relacionado con la disponibilidad del sustrato, que en medio hidrofóbico estaría más soluble.

Estos resultados indican que la configuración nativa de la pE2 no varía en solución acuosa o en una solución con polaridad más baja, ya que su actividad enzimática tampoco lo hace, en contraposición con lo que ocurre con el resto de enzimas.

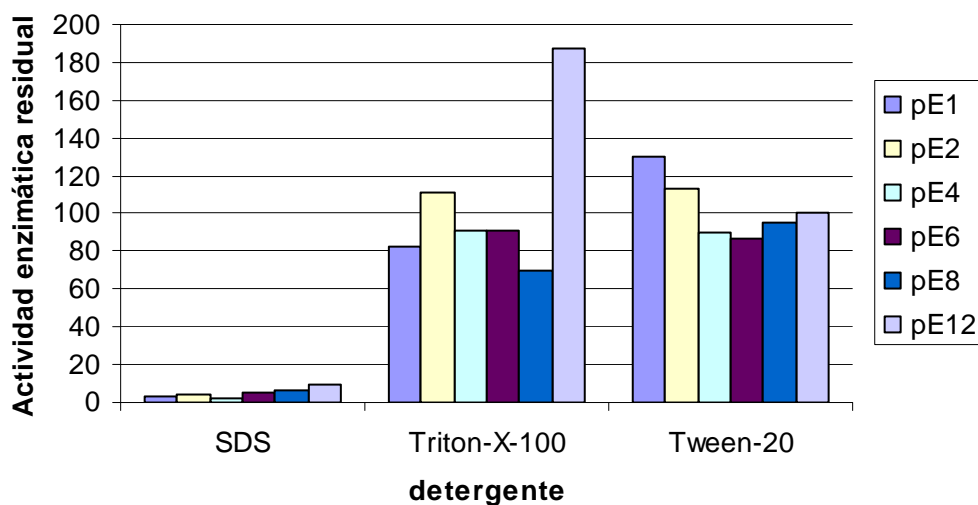
Las enzimas pE2, pE8 y pE1 pertenecen a la misma familia de esterasas, la familia II GDSL (figura X). En estructura primaria (secuencia de aminoácidos) la enzima pE8 es muy parecida a la pE2 y, sin embargo, tiene características en lo que a resistencia a solventes orgánicos se refiere muy diferentes, lo cual nos indica que en estructura secundaria probablemente sean muy diferentes. La pE1 difiere bastante de la pE2 y la pE8 en cuanto a estructura primaria se refiere y su perfil de resistencia a solventes orgánicos se presenta en el subapartado siguiente en comparación con la enzima pE2.

### ***Efecto de los detergentes en la actividad enzimática de los clones lipolíticos encontrados***

En lo que respecta al efecto de los detergentes, se ensayaron dos tipos: los iónicos (SDS) y los no iónicos (Tritón X-100)

En la figura siguiente, se observa cómo el detergente iónico SDS prácticamente inhibe la actividad enzimática de todos los clones. El pE12 es el que parece ser un poco más estable a la incubación con este tipo de detergente, manteniendo su actividad en el 5%.

Con respecto a los detergentes no iónicos, Tween-20 y Triton-X-100, las enzimas pE1, pE2, pE4, pE6 y pE8 mantienen su actividad más o menos constante al incubarlas con el detergente, mientras que la enzima pE12 se ve fuertemente activada al incubarla con Triton X-100



**Figura II.3.10.** Efecto de los detergentes iónicos y no iónicos en la actividad enzimática de los clones con actividad lipolítica aislados de la genoteca EEZ. La resistencia a estos dos tipos de detergentes se ensayó incubando los extractos con un 10% (p/v o v/v) de los mismos una hora a 25 grados. Después de esta hora de incubación, se calculó la actividad enzimática residual expresándola como porcentaje respecto a un control sin detergente. Los ensayos se hicieron por triplicado y en la gráfica se muestra la media de estas tres réplicas.

Con todas estas caracterizaciones bioquímicas preliminares, podemos llegar a la conclusión de que la enzima pE2 es una enzima muy estable a solventes orgánicos, lo que la podría hacer deseable en lo que concierne a una posible aplicación biotecnológica. Del mismo modo, también cabe resaltar la enzima pE12, que se activa por detergentes no iónicos, como el Tritón-x-100, y que es capaz de hidrolizar ésteres de alto número de carbonos.

### **II3.2.4.Purificación de enzimas esterases y caracterización bioquímica**

Para poder llevar a cabo estudios bioquímicos más precisos, y además poder corroborar los datos de los estudios preliminares descritos en los apartados anteriores, se sobreexpresaron todos los clones que portaban los insertos con genes esterasa siguiendo el protocolo de purificación que se describe en el material y métodos (esquematizado en la figura II.3.11).

De estos clones, siguiendo esta estrategia, se consiguió la purificación total (mayor al 90%) para las proteínas pE1, pE2 y pE4 (figura II.3.11) y se

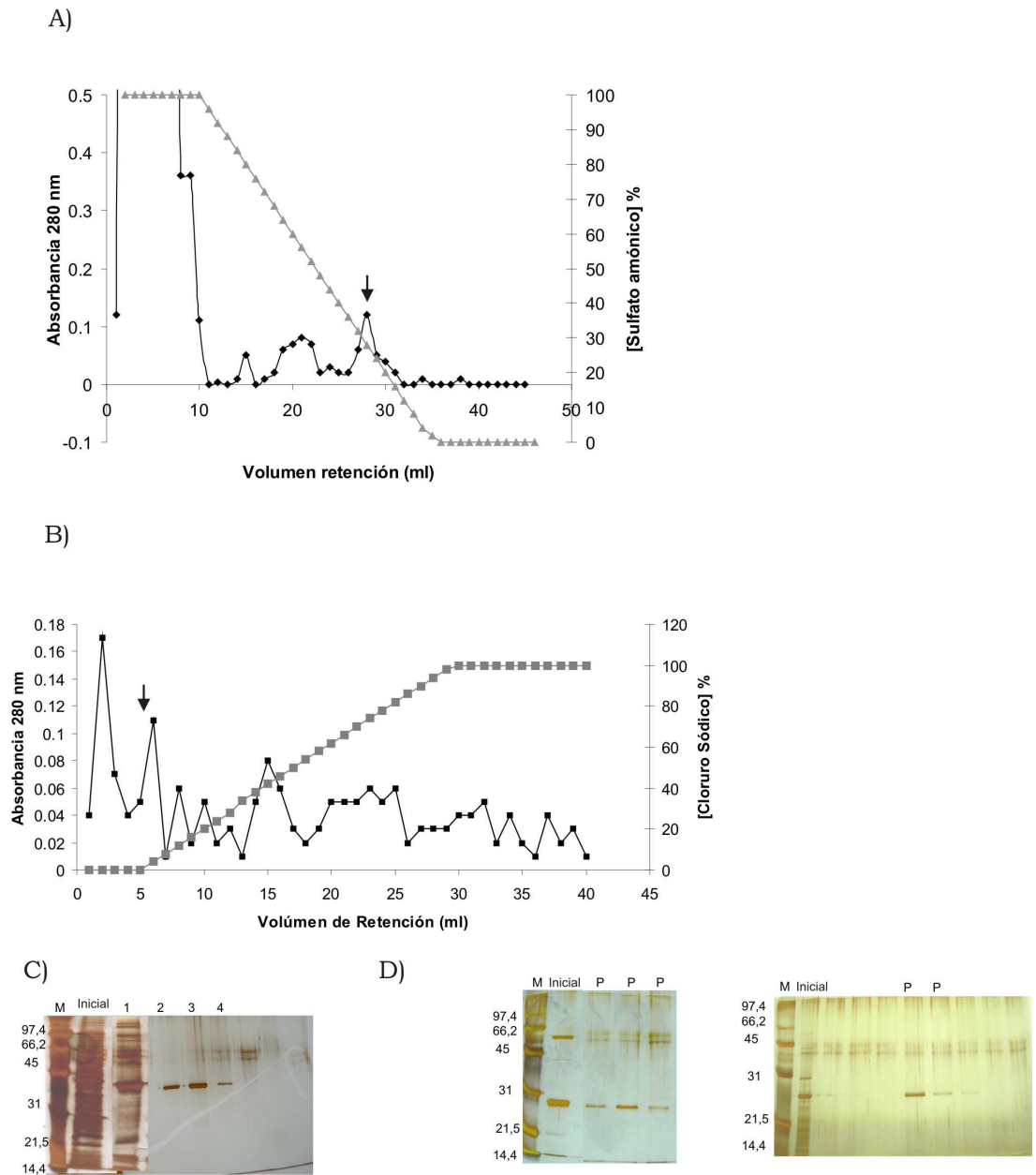
obtuvo una concentración de enzima de 0,18 mg/ml, 0,46 mg/ml, 0,03 mg/ml respectivamente.

La variabilidad en el rendimiento es explicable por las características particulares de cada uno de los clones, así como de las diferentes propiedades fisicoquímicas de cada esterasa.

De este modo, para el extracto proteico del clon 1, se pasó en un primer momento por una columna de intercambio catiónico, eluyendo a una concentración de NaCl de 200 mM. Posteriormente, se pasó por una columna de filtración molecular, filtración en gel, y es dónde ya se obtuvo pura. En el caso del clon 2 (figura II.3.11), se pasó el extracto por la columna de hidrofobicidad, eluyendo a una concentración de sulfato amónico de 0,45 mM. El siguiente paso cromatográfico fue una columna de intercambio iónico en la que eluyó en las primeras fracciones antes de incrementarse la concentración de cloruro sódico (eluyó a 0M de NaCl).

Como se indica en la tabla II.3.5, la proteína pE2 tiene un punto isoeléctrico de 8,74, por lo que a pH 7, que es el que se usa para este tipo de purificaciones, la proteína estaría cargada positivamente, lo que explica que eluya en las primeras fracciones en la cromatografía de intercambio iónico.





**Figura II.3.11.** Esquema del protocolo cromatográfico utilizado en la proteína pE2.

**A) Cromatografía de hidrofobicidad:** Se representa en la gráfica la absorbancia a 280

nm de cada una de las fracciones y con una flecha en el pico donde eluyó. **B)**

**Cromatografía de intercambio iónico:** Al igual que en el caso anterior, se muestra el

cromatograma. Se hizo un ensayo esterasa observando que la actividad esterasa ahora

se concentra en las primeras fracciones. **C)** Tras correr un gel SDS-PAGE, se corroboró

que efectivamente habíamos purificado la proteína. M: Marcador, Inicial: Muestra

inicial; 1: Muestra tras pasar por la cromatografía de hidrofobicidad, 2-4: Muestra

purificada. **D)** Geles SDS-PAGE al 12 % teñidos con plata. A la derecha se muestra la

purificación pE1 y la izquierda la de la purificación pE4. En todos los geles, el orden

es el siguiente:

L1: marcador de peso molecular.

L2: muestra inicial.

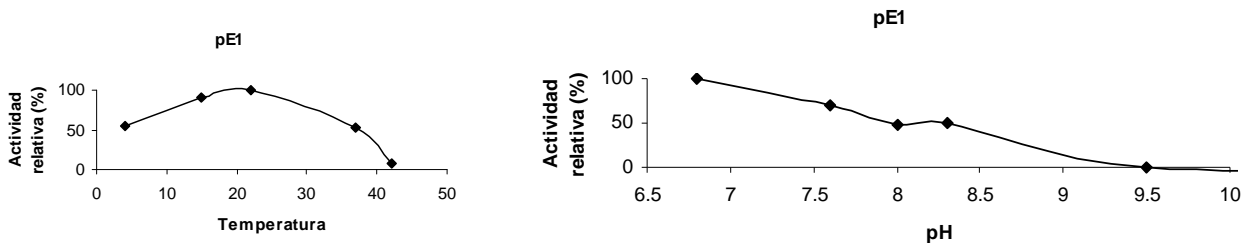
P: Muestra purificada.

Finalmente, para la pE4 se siguió el mismo esquema: Un primer paso en una columna de hidrofobicidad en la que eluyó a una concentración de sulfato amónico de 900 mM y un segundo paso en una columna de intercambio iónico en la que se consiguió purificar la proteína mediante su elución a una concentración de 360 mM de NaCl.

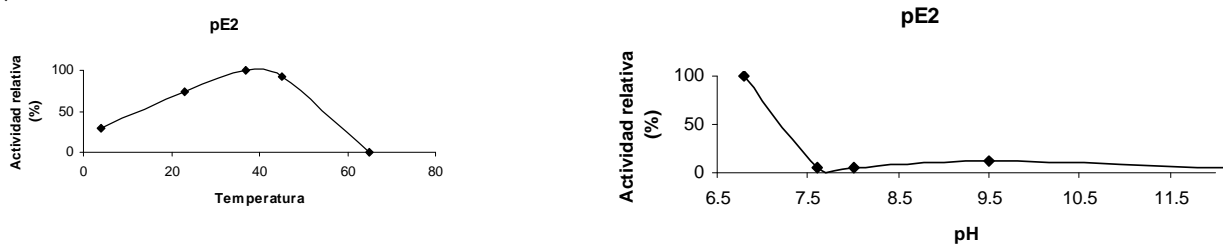
**II.3.2.4.1. Características bioquímicas de las enzimas esterasas purificadas**

Se caracterizaron las enzimas así purificadas determinando su temperatura y pH óptimo (figura II.3.12) y se vio que se obtenían datos similares a los que ya poseíamos de los extractos crudos.

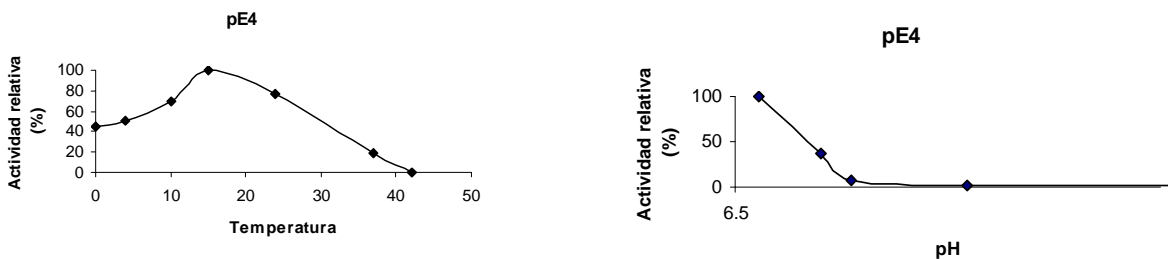
A)



B)



C)



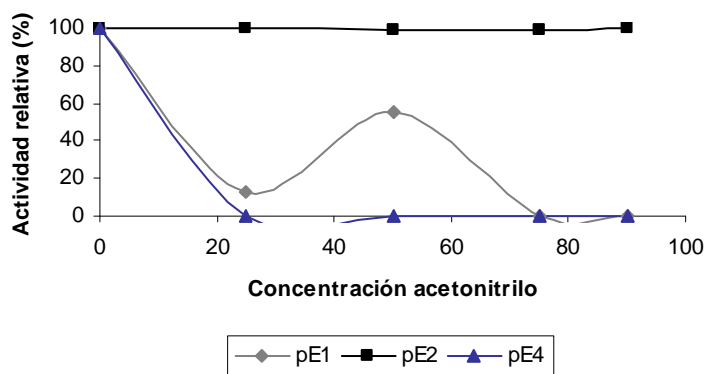
**Figura II.3.12.** Curvas de temperatura y de pH de las enzimas purificadas pE1 (A), pE2 (B) y pE4 (C).

Asimismo, también se calculó la actividad enzimática de cada una de las enzimas con los sustratos pNp de diferente longitud de átomos de carbono (Tabla II.3.9).

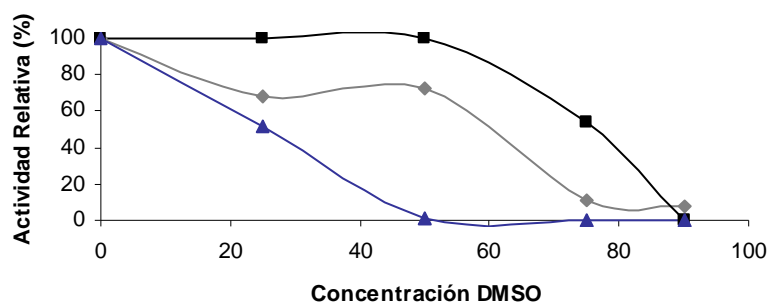
	AEE (nmol/min.mg pro <sup>-1</sup> )		
	pE1	pE2	pE4
acetato (C2)	11419,75	11555,55	87.962,90
butirato (C4)	12,34	6888,89	52037
caprilato (C8)	0	0	0
laureato (C12)	43,2	0	0
palmitato (C16)9	0	0	0

**Tabla II.3.9.** Especificidad de sustrato de las enzimas esterazas purificadas.

También se ensayó la resistencia a diferentes solventes, como acetonitrilo y dimetilsulfóxido.



**Figura II.3.13.** Efecto de los solventes orgánicos sobre la actividad enzimática. En la gráfica de la parte superior se observa el efecto del solvente acetonitrilo y en la parte inferior el solvente DMSO.



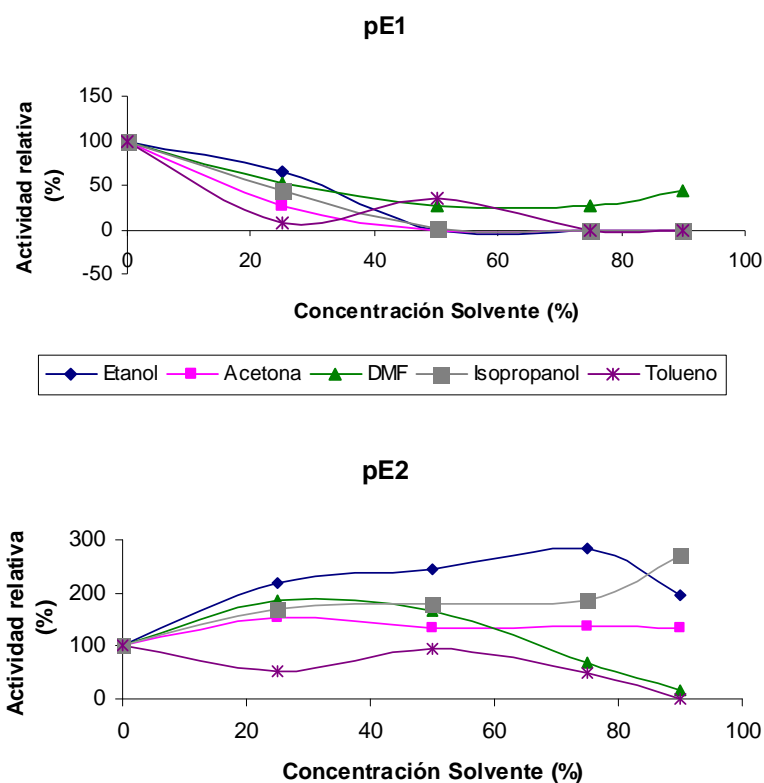
De este modo, se corroboran los datos obtenidos mediante extractos crudos, la enzima pE2 es muy estable en solventes orgánicos y, además, la pE4 aparece como una enzima muy activa frente a ésteres de cadena corta como pNp acetato.

### II3.2.5. Estudio de la resistencia a solventes de la enzima pE2 frente a la enzima pE1

#### II3.2.51. Estudios preliminares

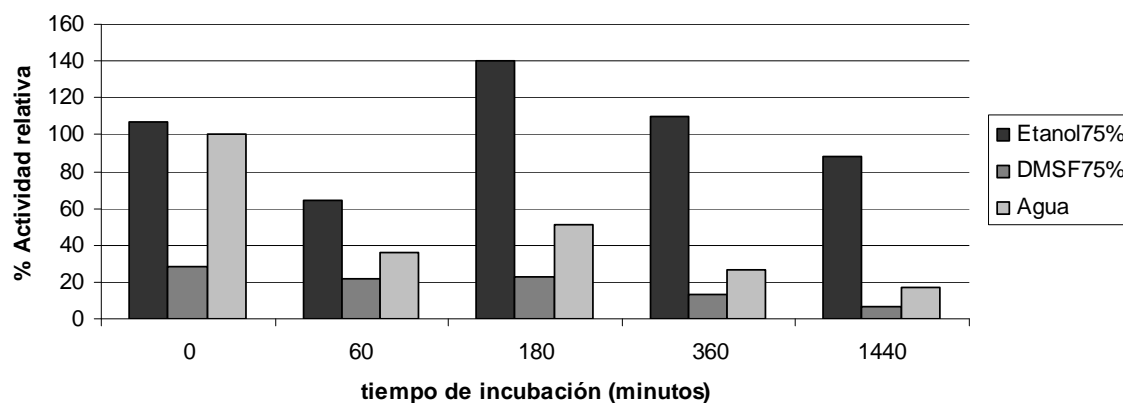
Como ya hemos descrito en los capítulos anteriores, las enzimas pE1 y pE2 pertenecen a la misma familia de esterasas (familia II, GDSL, tabla II.3.5). Sin embargo, tienen características bien diferenciadas. Como ya hemos comentado en el subapartado anterior, la enzima denominada pE2 es muy estable en solventes orgánicos, mientras que la pE1 es muy sensible.

Como estudios preliminares al análisis de la resistencia a solventes orgánicos por parte de la enzima pE2, se realizaron más experimentos usando diferentes solventes (no sólo acetona, acetonitrilo) y efectivamente, se observó que la proteína pE2 es muy estable a solventes frente a la pE1, sobre todo en lo que incubación con alcoholes se refiere (figura II.3.14.). En el caso del solvente tolueno no se pudo obtener ninguna conclusión clara debido a su incapacidad de mezclarse con una solución acuosa.



**Figura II.3.14.** Resistencia a solventes orgánicos de las enzimas pE1 y pE2. Los ensayos se hicieron por triplicado incubando las enzimas en diferentes concentraciones de solventes durante 6 horas a temperatura ambiente y calculando la actividad enzimática residual.

Asimismo, también se realizó otro experimento incubando la enzima pE2 a temperatura ambiente en diferentes tiempos (hasta un máximo de 24 horas) en una concentración constante tanto de etanol, como de DMSF, como de agua. De este modo, se observa que la enzima es mucho más estable en un solvente orgánico que en un solvente polar, como el agua.



**Figura II.3.15.** Actividad relativa de la enzima pE2 tras su incubación durante diferentes tiempos con etanol, dimetilsulfóxido y agua al 75%. Los ensayos se hicieron por triplicado y la actividad enzimática residual se calculó frente al control en el que no se ha incubado la enzima durante ningún tiempo en agua.

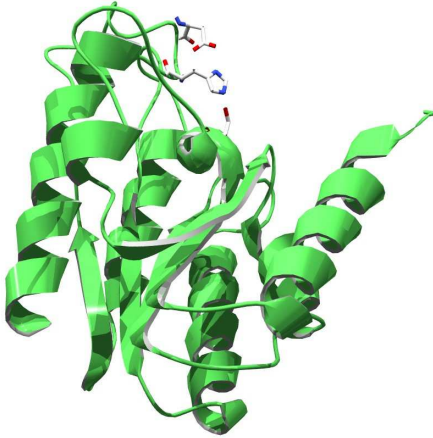
Una vez corroboradas las diferencias en estabilidad entre ambas enzimas de la misma familia, pE1 y pE2. Nos planteamos la posibilidad de determinar los dominios de la enzima pE2 que pudieran explicar este comportamiento.

### **II.3.2.5.2. Análisis de la estructura primaria, secundaria y terciaria de las enzimas PE1 Y PE2**

En el capítulo II.3.2.4. ya describimos las diferentes características en lo que a secuencia de aminoácidos se refiere de las proteínas pE2 y pE1. La proteína pE1 es mucho más pequeña que la proteína pE2 (238 aminoácidos frente a 405 respectivamente) y, además, la serina del sitio activo en el caso de la pE1 está mucho más cerca del extremo N-terminal (en el aminoácido 62), que en el caso de la pE2 que se encuentra en el aminoácido 202.

Para poder estudiar las posibles diferencias en cuanto a estructura terciaria de estas proteínas, procedimos a realizar una modelización virtual de la estructura a partir de la secuencia de aminoácidos. La proteína pE1 fue

fácil de modelar debido a su parecido con otras enzimas de la misma familia (figura II.3.16). Sin embargo, de la enzima pE2 no pudimos obtener un modelo claro por el poco parecido que presenta con otras enzimas hidrolíticas.



**Figura II.3.16.**

Estructura tridimensional de la proteína pE1. Explicar algo sobre los aa del centro activo!!. Se ha hecho en base a ...

Aunque no pudimos modelizar la proteína pE2, partimos de la hipótesis de que quizás la parte N-terminal de esta enzima de 200 aminoácidos podría estar involucrada en las características tan peculiares de la proteína con respecto a la resistencia a solventes.

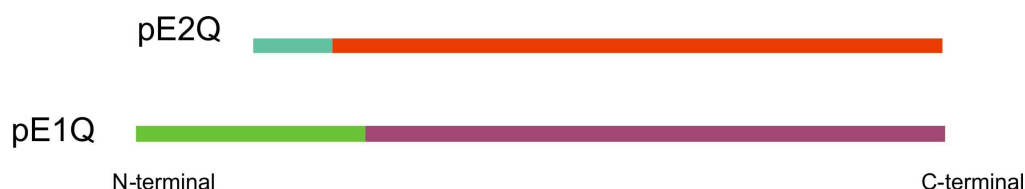
De este modo, se planteó diseñar una serie de experimentos para construir quimeras con el extremo N-terminal de la pE2 (representado de color verde en la figura II.3.17).

Concretamente, se propuso la construcción de dos tipos de quimeras: la denominada pE1Q, que es una enzima con el extremo N-terminal de la pE2 (color verde) pero con el resto la secuencia, con el sitio activo, de la pE1 (de color morado) y la quimera pE2Q que tiene el extremo N-terminal de la enzima pE1 (de color azul) pero con el resto de la secuencia de la pE2 (de color rojo, figura II.3.17).

**A) Enzimas pE1 y pE2 originales.**



**B) Quimeras construidas pE1Q y pE2Q.**



**Figura II.3.17.** Representación gráfica de la secuencia aminoacídica de las proteínas pE1 y pE2 (A) y las quimeras resultantes del intercambio del módulo N-terminal (B).

**II3.2.5.3. Construcción de las quimeras pE1Q Y pE2Q**

Para la construcción de las quimeras, en primer lugar estudiamos la secuencia nucleotídica de cada una de las enzimas.

De este modo, observamos que podríamos introducir mutaciones puntuales en la secuencia para generar una diana para la enzima de restricción *Sac* II. Con la diana así generada, podríamos digerir e intercambiar el C-terminal de la enzima pE1 por el C-terminal de la enzima pE2.

Para generar la diana *Sac* II en la secuencia de la proteína pE1, necesitábamos introducir una mutación de 3 nucleótidos (nucleótidos 66, 67, 68; figura II.3.18). Concretamente, esta mutación produce un cambio de un aminoácido neutro a otro con carga positiva, alanina por arginina, que se quedaría al límite del extremo C-terminal de la proteína pE1.

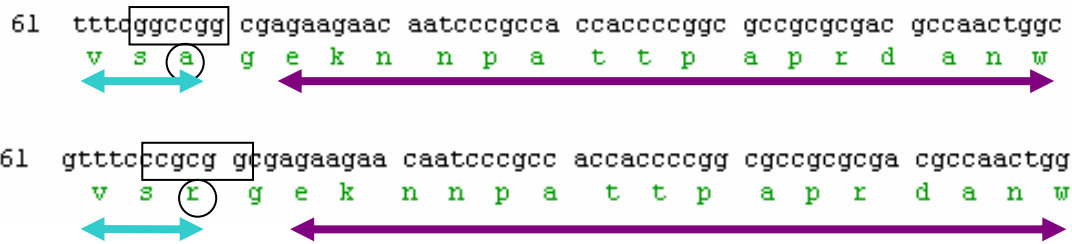
En el caso de la proteína pE2, para generar la diana *Sac* II en su secuencia nucleotídica, necesitábamos introducir dos mutaciones que no producen cambios en la secuencia aminoacídica en la pauta de lectura (figura II.3.18).

Los mutantes se construyeron siguiendo las indicaciones del sistema pALTER, tal y como se describe en el Material y Métodos de esta tesis doctoral

A)



B)



C)



**Figura II.3.18.** Estrategia seguida en la introducción de la diana para la enzima de restricción *Sac* II en la secuencia nucleotídica de las proteínas pE1 y pE2. **A)** En el caso de la pE1, se introducen 3 mutaciones en la secuencia, concretamente en los nucleótidos 66, 67 y 68, cambiando la secuencia “ggc” por otra “ccg”. Se indica en azul la parte correspondiente al N-terminal de la proteína pE1 y en morado la parte del C-terminal de la proteína pE1. **B)** Cambio que se produce en el aminoácido alanina por arginina en la secuencia de la proteína pE1. Se indica en azul la parte correspondiente al N-terminal de la proteína y en morado la parte del C-terminal de la proteína. **C)** Estrategia seguida para la introducción de la diana para la enzima de restricción *Sac* II en la secuencia nucleotídica de la proteína pE2. Se introducen 2 mutaciones en la secuencia, cambiando la secuencia “cagg” por otra “ccgc”. Se indica en rojo la parte correspondiente al N-terminal de la proteína pE3 y en verde la parte del C-terminal de la proteína pE2.

Una vez construidos los mutantes, se digirieron con la enzima de restricción *Sac* II y se intercambiaron los módulos: el extremo N-terminal de la enzima pE1 se puso a ligar con el extremo C-terminal de la enzima pE2, y el



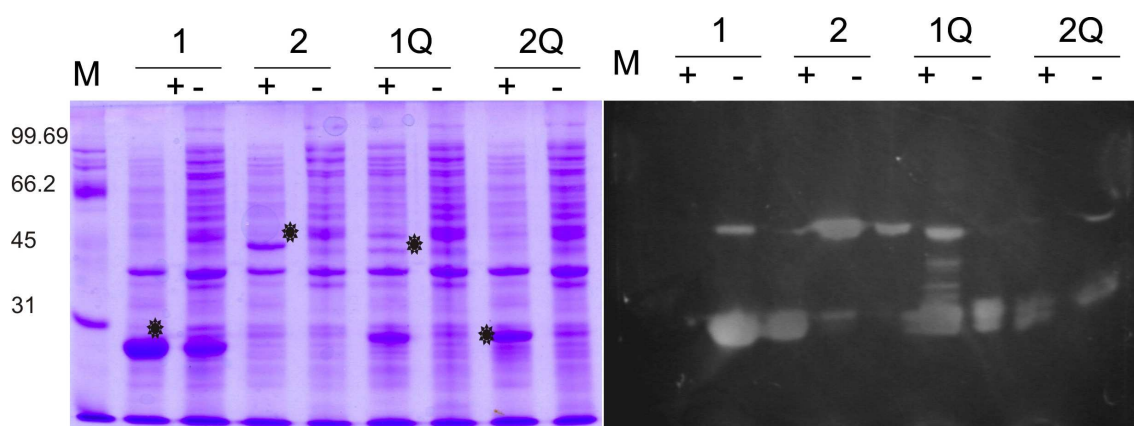
extremo N-terminal de la enzima pE2 se puso a ligar con el extremo C-terminal de la enzima pE1.

Las quimeras así construidas se clonaron en el sistema de expresión pET en dos plásmidos diferentes, uno que poseía colas de histidinas (plásmido pET16b) y otro sin colas (plásmido pET3a).

#### II3.2.5.4. Sobre-expresión de las quimeras

Los plásmidos que contenían las secuencias quiméricas y las secuencias originales se sobre-expresaron en células *E.coli*, induciendo y sin inducir con IPTG, tal y como se indica en la figura II.3.19. Los tamaños moleculares teóricos de las proteínas son 26,06, 43,18, 40,08 y 28,48 KDa para las enzimas pE1, pE2, pE1Q y pE2Q respectivamente. Estos tamaños se ven reflejados en los geles de la figura II.3.19.

Con los extractos así sobreexpresados, se hicieron diferentes ensayos de actividad y resistencia a solventes orgánicos.



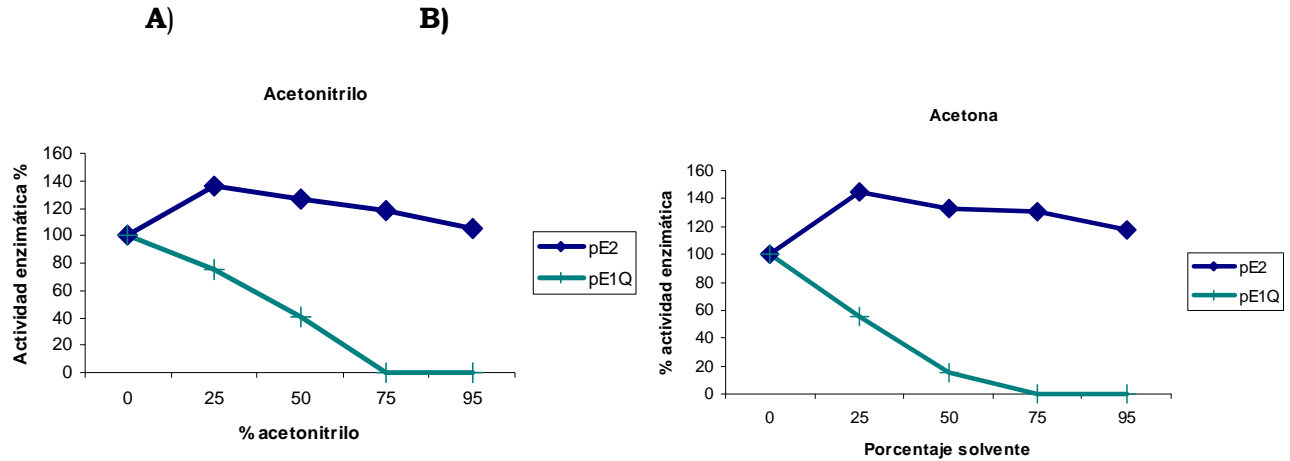
**Figura II.3.19.** Geles SDS-PAGE al 12% con las muestras obtenidas tras la sobreexpresión de las enzimas pE1, pE2, pE1Q y pE2Q. En los geles se muestran los extractos inducidos (+) y los sin inducir (-). El gel de la izquierda corresponde a un gel teñido con Coomassie blue, mientras que el de la derecha es un zimograma que muestra la actividad enzimática.

De este modo, se observó que la enzima pE2Q era inactiva. Esto significa que la proteína pE2 pierde la actividad sin su extremo N-terminal. Este módulo por tanto, sería de vital importancia para la actividad enzimática.

#### II3.2.5.6. Análisis de la actividad de las quimeras

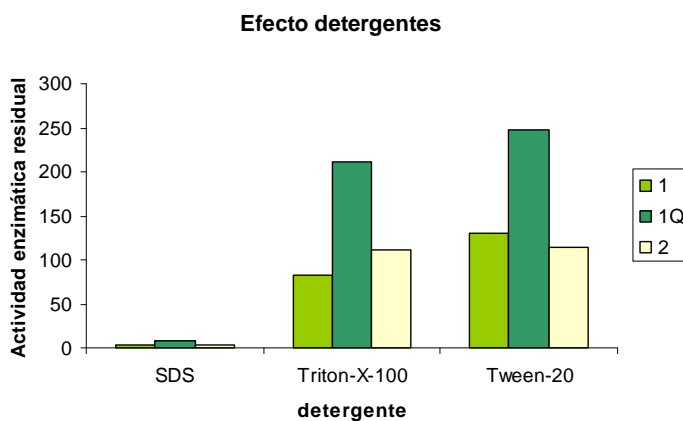
Con respecto a la resistencia a solventes orgánicos, se hicieron diferentes experimentos incubando las enzimas durante 6 horas a distintas concentraciones de solventes y calculando la actividad enzimática residual.

De este modo, vimos que la proteína pE1Q es sensible a la incubación en este tipo de sustancias y que no mantiene el fenotipo de la proteína pE2, por lo que la propiedad de estabilidad en medios no acuosos no residiría en este dominio de la proteína o no sería transferible.



**Figura II.3.20.** Gráfica de resistencia a solventes orgánicos con los extractos proteicos que contienen a las enzimas pE2 y pE1Q. Los ensayos se hicieron por triplicado y se observa la media de estas tres réplicas.

Además de la resistencia a solventes orgánicos, también se testaron otro tipo de propiedades, como la estabilidad en detergentes. Se observó, de este modo, que si bien la quimera pE1Q no había ganado en cuanto a estabilidad a solventes orgánicos, sí lo había hecho con respecto a la estabilidad en diferentes detergentes, siendo mucho más estable que las enzimas pE1 y pE2.



**Figura II.3.21.** Estabilidad frente a diferentes detergentes. Las enzimas se incubaron con un 10% de cada detergente durante una hora para calcular posteriormente la actividad enzimática residual.

Además, también calculamos la constante  $K_m$  para este tipo de enzimas. De este modo, vimos que la parte N-terminal de la pE2 hacía que la enzima pE1Q fuera el doble de específica para el sustrato pNp-Acetato que la enzima silvestre pE1, aunque no hasta el punto de la enzima pE2. Se calcularon unos valores de  $K_m$  de 0,4371, 0,079 y 0,231 para las enzimas pE1, pE2 y pE1Q respectivamente.

Con este tipo de aproximación, hemos visto que la parte N-terminal de la enzima pE2 resulta esencial para la actividad de la proteína, aunque en este “módulo” no se encuentre ningún residuo del sitio activo G-D-S-L. Este hecho no se repite en el caso de la proteína pE1, que es funcional sin esta parte N-terminal.

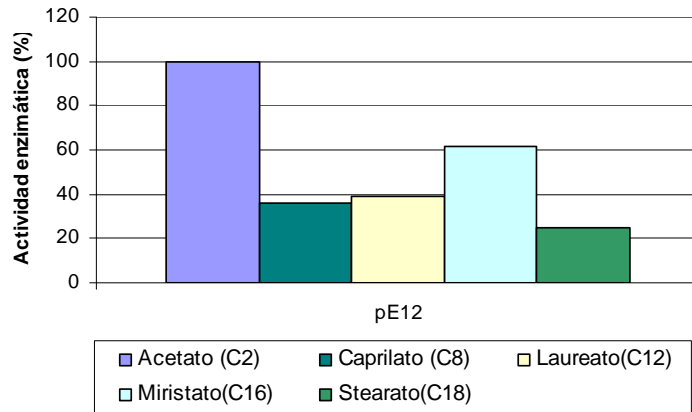
Además, si se transfiere este módulo (el N-terminal) a otra enzima que no presente estabilidad a solventes orgánicos, esta propiedad no se transfiere, aunque se consigue mejorar la especificidad por el sustrato y la resistencia en la incubación con diferentes detergentes. Para seguir profundizando en las características moleculares de la resistencia a solventes orgánicos por parte de la enzima pE2, son necesarios otros experimentos de mutagénesis dirigida y un sistema de búsqueda adecuado en el que se puedan discriminar los mutantes que pierdan el fenotipo. De este modo, se podría inferir qué residuos están implicados en esta propiedad.

### ***II.3.2.6. Caracterización de la enzima pE12***

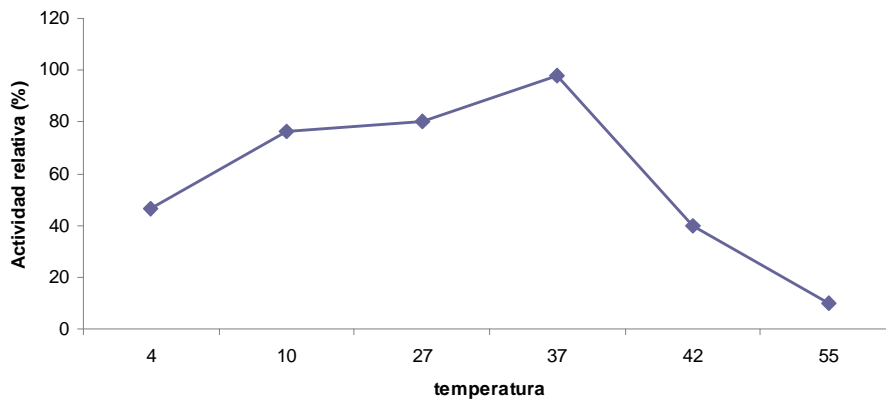
La enzima pE12 es una enzima de 309 aminoácidos perteneciente a la familia IV de esterasas, denominada HSL por su homología a enzimas sensibles a hormonas de mamíferos. Esta enzima tiene un porcentaje de identidad bajo (43%) con otras enzimas conocidas, y la proteína a la que más se parece es una perteneciente a una bacteria no cultivable. El pH óptimo es neutro y la temperatura óptima es de 37°C.

La peculiaridad de esta proteína radica en que tiene actividad en un margen muy amplio de temperaturas (figura II.3.22), se activa tras su incubación con tritón-X-100 y además es una lipasa verdadera que hidroliza ésteres de un alto número de carbonos como miristato o estearato (figura II.3.23), por lo que podría ser interesante para el uso en diferentes tipos de industria (papelera, farmacéutica, etc). Con el fin de profundizar en sus

peculiaridades, se desarrolló un protocolo para su purificación y posterior caracterización enzimática.



**Figura II.3.22.** Rango de hidrólisis de sustratos de la enzima pE12.



**Figura II.3.23.** Rango de actividad de la enzima pE12 a diferentes temperaturas.

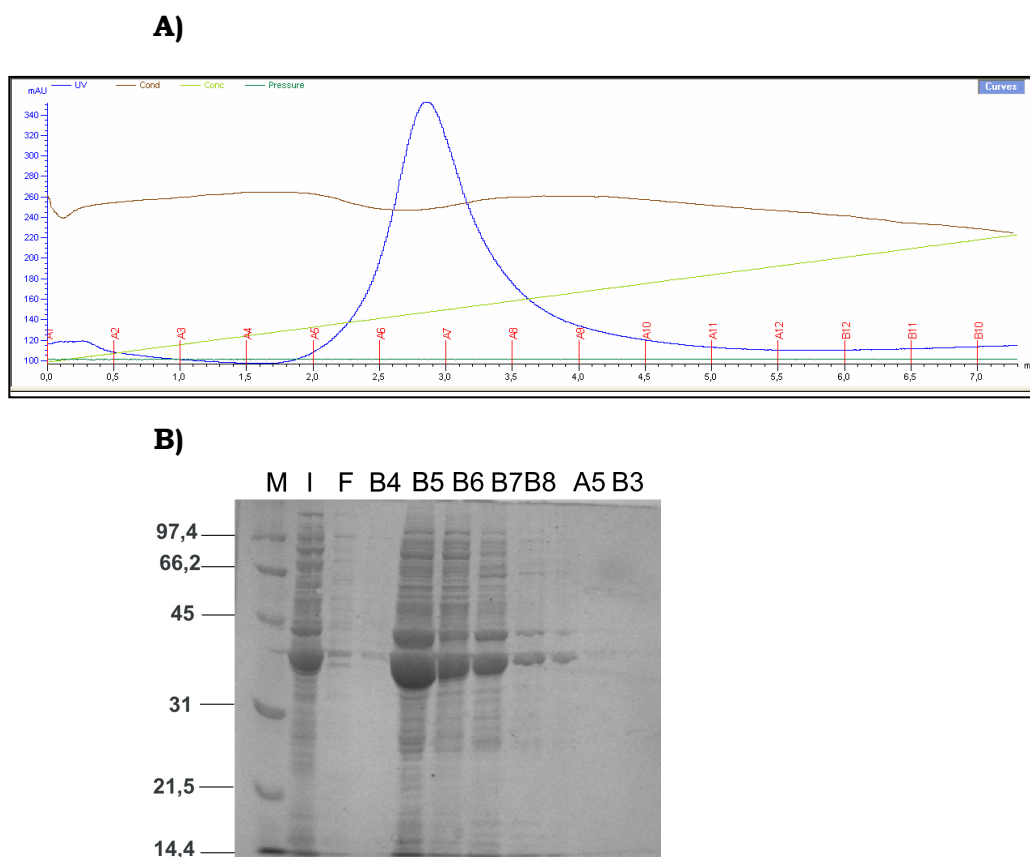
### II.3.2.6..1. Sobre-expresión y purificación de la enzima pE12

Para la sobreexpresión de la proteína pE12, se clonó la enzima en el vector de expresión pET24c, que te permite sobreexpresarla con una cola de histidinas. La clonamos de forma que la cola de histidinas estuviera en el extremo C-terminal.

Después del clonaje en este vector de expresión, se procedió a la sobreexpresión en células de *Rosetta* de *E.coli* (ver el capítulo 13 de Material y Métodos) para la purificación, usando una columna de afinidad equilibrada previamente coné níquel (figura II.3.23).

De este modo, partimos de 86 mg de proteína total, que se pasó por la columna previamente equilibrada con níquel y, tras establecer un gradiente de

imidazol de cero a 500 mM, la proteína eluyó en las fracciones que correspondían con 100 mM de imidazol (figura II.3.23).



**Figura II.3.23.** Proceso de purificación de la enzima pE12. A) Cromatograma resultante de pasar el extracto crudo por una columna de níquel. La línea azul corresponde a la absorbancia a 280 nm y la verde a la concentración de imidazol. Se observa el pico al que eluye la enzima pE12 en las fracciones desde la A5 hasta la A8. B) Gel con las diferentes muestras del proceso de purificación. M corresponde a marcador; I es la muestra inicial; F es la muestra que eluye al cargar la columna y A4-B3 son las diferentes fracciones.

#### II3.2.6.1. Caracterización enzimática

Con la enzima así purificada, se realizaron diferentes ensayos enzimáticos.

##### - *Sustratos*

Se calculó la actividad enzimática específica frente a diferentes sustratos, tanto sustratos pNp con ácidos grasos de diferente número de átomos de carbono, como sustratos enantioméricos (tabla x). Destaca la actividad que la enzima pE12 tiene sobre sustratos de interés farmacéutico como el nanoproxeno o sobre el compuesto químico naphthoato.

Además, se usó el pHstato para testar la actividad de nuestra enzima frente al sustrato poliglicerol-laureato, pero no se detectó actividad.

Substrate	Sp act (U/mg)		
	EstCE1	EstA3	pE12
<b>pNp-acylesters</b>			
Acetato (C2)	nd	nd	549
Butirato (C4)	31,1	513,6	201
Caprilato (C8)	1	3,6	89,32
Myristato (C14)			114
2-(4-isobutylphenyl)propanoate <i>ibuprofeno</i>	nd	7,2	nd
2-phenyl propanoato	nd	21	15,28
3-phenyl butanoato	1	31,6	5,85
cyclohexanoato	16,8	50	13,04
2-(3-benzoylphenyl)propanoato	0,8	26,7	10,5
2-naphtoato	nd	nd	3,69
1-naphtoato	nd	nd	10,79
adamantanoato	nd	nd	3,25
2-(6-methoxynaphtalen-2-ylpropanoat0) <i>nanoproxeno</i>	nd	26,7	10

**Tabla II.3.10.** Actividad enzimática específica frente a distintos sustratos. En la tabla se muestra la enzima aislada de la genoteca medioambiental de la Estación Experimental del Zaidín (pE12) y otras dos enzimas publicadas en la literatura (EstCE1 y EstA3; Elend C, *et al* 2006).

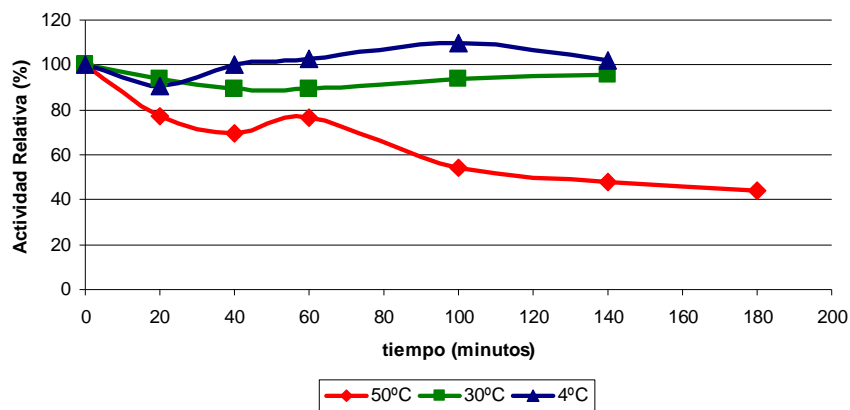
- *Temperatura óptima*

Se repitieron los experimentos de temperatura óptima y se corroboró que, efectivamente, la enzima, una vez purificada, es activa en un amplio rango de temperaturas que van desde los 4°C, donde mantiene aproximadamente el 30% de su actividad, hasta los 55°C, donde es activa en un 50%.

- *Estabilidad frente a temperaturas*

También se ensayó su estabilidad incubando la enzima a diferentes temperaturas durante distintos tiempos. De este modo, se observó cómo la enzima es muy estable a la incubación a temperaturas de 50 grados, puesto

que mantuvo aproximadamente el 40% de la actividad residual tras incubarla a esta temperatura durante 3 horas.



**Figura II.3.24.** Estabilidad de la enzima pE12 tras la incubación a diferentes temperaturas.

- *Efecto de los iones en la actividad enzimática*

Se ensayaron los efectos del ión potásico, manganesico y calcio sobre la actividad enzimática (tabla II.3.11). El ión potásico tiene un efecto positivo sobre la actividad, la estimula, mientras que el ión calcio la inhibe en parte. En el caso del ión manganeso no produce ningún efecto sobre la actividad enzimática.

	Concentración (mM)		
	1	5	10
<b>KCl</b>	100	93	160
<b>MnCl<sub>2</sub></b>	100	100	100
<b>CaCl<sub>2</sub></b>	100	50	50

**Tabla II.3.11.** Efecto de los iones potásico, manganeso y calcio sobre la actividad enzimática. El ensayo se realizó por triplicado añadiendo al tampón enzimático distinta concentración del compuesto (1,5 y 10 Mm).

### **II3.3. BÚSQUEDA DE GENES CON ACTIVIDAD LIPOLÍTICA EN LAS GENOTECAS CONSTRUIDAS A PARTIR DE MUESTRAS DE ADN DE LA RIZOSFERA DE *Mamillaria carnea***

Al igual que en el caso anterior con la genoteca de la Estación Experimental del Zaidín, se testaron las genotecas construidas a partir de ADN extraído en época seca y lluviosa de la rizosfera de *Mamillaria carnea* frente a un ensayo de actividad esterasa.

De este modo, se realizó un primer test con aproximadamente 37500 pfu/placa de petri usando 20 placas de petris (que supondrían 750000 pfu, es decir, unas 3 veces el tamaño de nuestras genotecas originales).

En este primer test se identificaron 4 calvas con actividad esterasa en el caso de la genoteca construida a partir de la muestra en época seca y 10 calvas con actividad esterasa en el caso de la genoteca de época lluviosa. Ensayos posteriores revelaron que realmente sólo poseíamos 2 clones con actividad esterasa para la genoteca de época seca y 5 clones con actividad esterasa en el caso de la genoteca de época lluviosa. Estas unidades formadoras de calvas se escindieron a plásmido para poder realizar los siguientes ensayos.

De los 7 clones encontrados, sólo pudimos realizar estudios preliminares basados en la secuencia y en las características bioquímicas más básicas, como pH y temperatura óptima, a 4 de ellos. Estudios posteriores a esta Tesis Doctoral revelarán las características bioquímicas más peculiares de estas enzimas, así como su posible papel en la industria.

#### **II3.3.1. subcloning y análisis de las secuencias**

Se generaron subclones disminuyendo así el tamaño de inserto para identificar el gen causante de la actividad esterasa. Las secuencias de estos genes corresponden a enzimas diversas que pertenecen a distintos grupos o familias de esterazas, aunque algunas no se han podido identificar con exactitud (Arpigny and Jaeger, 1999; tabla II.3.12).

En esta Tesis Doctoral no se presentan las características de los clones procedentes de la genoteca de época lluviosa 8.2 y 8.14, ni para el clon procedente de la genoteca de época seca 6.4.



De entre los clones de los que sí poseemos la secuencia del gen que codifica para la enzima esterasa, llaman especialmente la atención el clon 6.2 obtenido en la genoteca de época seca y el clon 8.25 obtenido en la genoteca de la época lluviosa. Ambos se solapan en unas 3 kbs, siendo dos fragmentos de ADN procedentes del mismo microorganismo que codifican para las mismas enzimas. Estas enzimas guardan gran homología con otras pertenecientes al phylum bacteriano *Acidobacteria*, phylum que ya hemos visto que es especialmente abundante en estos suelos (capítulo I).

Los clones 6.2 y 8.25 codifican para dos esterases, denominadas 6.2.1 y 6.2.2 (la 6.2.2 es una enzima multifunción muy grande en tamaño; tabla II.3.12), que no se han podido clasificar en ninguna de las familias de lipasas conocidas hasta el momento.

Las proteínas 6.2.1 y 6.2.2 forman un grupo independiente (figura II.3.25) con otras lipasas de *Bacteroidetes*. En su secuencia, la serina de la tríada catalítica podría estar en un motivo conservado (GLSMGG) muy parecido al motivo que contiene la serina catalítica en la familia V de esterases (GNSMGG). Sin embargo, este motivo está muy cerca del extremo C-terminal de la proteína y, además, no se identifican el resto de motivos que contienen la histidina y el aspártico del centro activo en la familia V de esterases.

Si se pudieron observar los residuos conservados de la proteína 8.1, que pertenece a la familia VIII de esterases, familia que guarda una gran homología con las enzimas beta-lactamasas

Clon	Tamaño	Genoteca	Nombre	Enzima esterasa							Microorganismo
				Número aa	Pm(kDa)	Pi <sup>a</sup>	enz.más homología <sup>b</sup>	Familia lipasas	Identidad <sup>c</sup>	E-valor <sup>d</sup>	
6.2*	7 kb	Seca	6.2.1	409	44.15	9.1	Esterasa putativa	Sin identificar	58%	3e-117	<i>Solibacter usitatus</i> <i>Ellin6076</i>
	7 kb	Seca	6.2.2	648	74.68	8.57	Esterasa putativa	Proteína multifunción	46%	1e-128	<i>Flavobacterium</i> <i>Johnsoniaen</i>
6.4	8 kb	Seca	6.4	//	/	/	/	/	/	/	/
8.1	8 kb	Lluviosa	8.1	370	39.53	7.63	Acetilesterasa putativa	Sin identificar	40%	8e-67	<i>Chthoniobacter</i> <i>flavus Ellin428</i>
8.2	11 kb	Lluviosa	8.2	/	/	/	/	/	/	/	/
8.25*	7 Kb	Lluviosa	6.2.1	409	44.15	9.1	Esterasa putativa	Sin identificar	58%	3e-117	<i>Solibacter usitatus</i> <i>Ellin6076</i>
			6.2.2	648	74.68	8.57	Esterasa putativa	Proteína multifunción	46%	1e-128	<i>Flavobacterium</i> <i>Johnsoniaen</i>
8.14	5 Kb	Lluviosa	8.14	/	/	/	/	/	/	/	/
8.17	7	Lluviosa	8.17	336	35.79	9.86	Enzima lipolítica familia G-D-S-L	Familia II-GDSL	44%	8e-42	<i>Chthoniobacter</i> <i>flavus Ellin428</i>
8.26	8,5 Kb	Lluviosa	8.26	575	59.8	9.37	Lipasa/esterasa	Familia VII	47%	5e-137	<i>Bacteria no</i> <i>cultivable</i>

**Tabla II.3.12.** Características de los clones de la genoteca metagenómica que portan los genes con actividad esterasa.

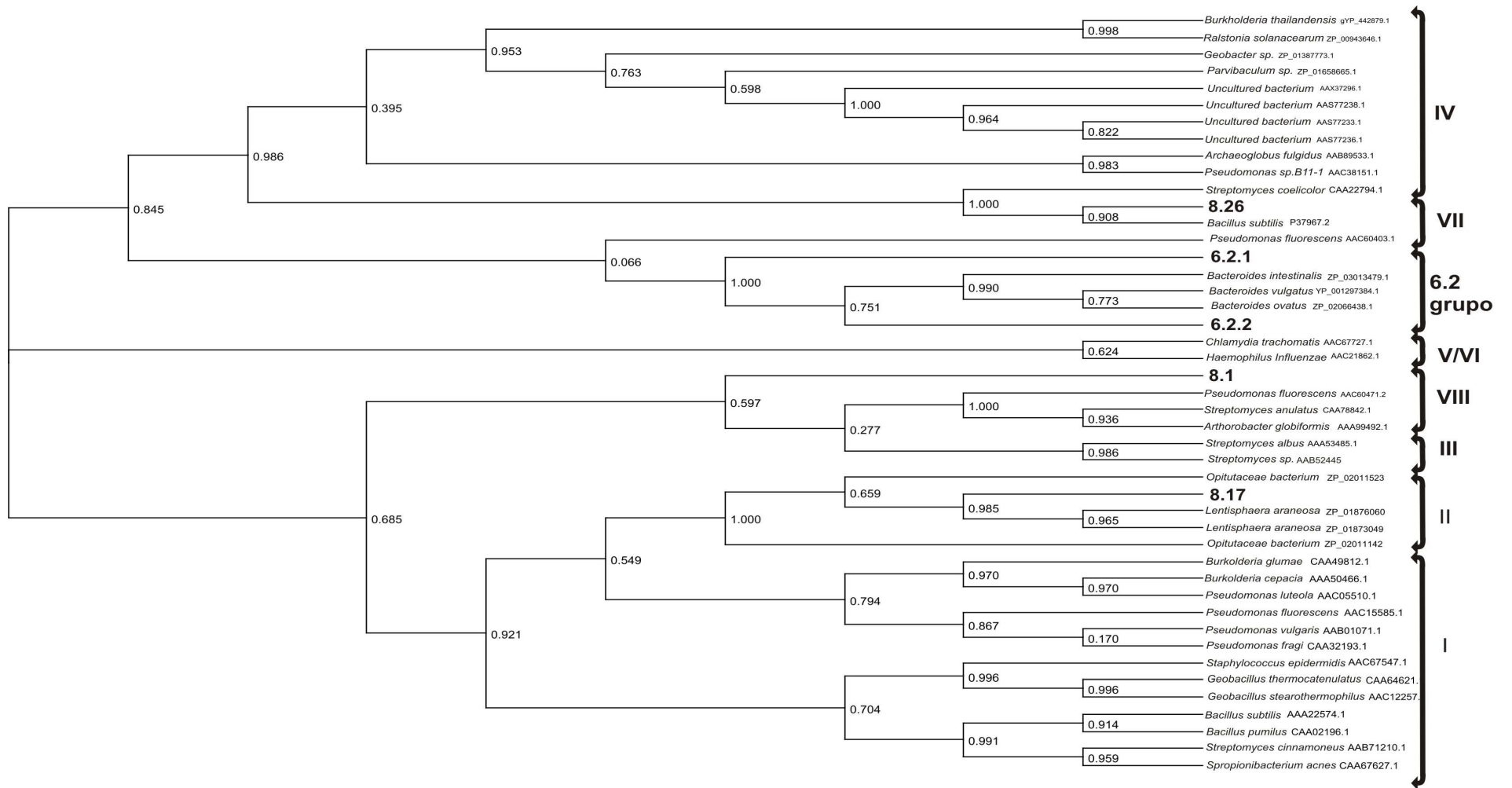
a) Punto isoeléctrico ([http://scansite.mit.edu/calc\\_mw\\_pi.html](http://scansite.mit.edu/calc_mw_pi.html)).

b) Enzima con la que guarda más parecido en los alineamientos (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

c) Es un valor numérico que indica cuánto se parecen entre sí.

d) Representa el número de alineamientos con puntuación igual o equivalente a la que podría haber sido encontrada sólo por azar.

\* El clon 6.2 y el 8.25 son clones que portan las mismas enzimas.



**Figura II.3.25.** Árbol filogenético con las proteínas aisladas en las genotecas metagenómicas mex-lluvioso y méx-seco y otras esterazas de las bases de datos. Se indica el microorganismo de procedencia de cada enzima, así como el número de acceso. Por su parte, el número romano que se representa a la derecha se corresponde con la familia de lipasas/esterasas en la que se agrupan las proteínas.

Con respecto a la enzima codificada por el clon 8.17, pertenece a la familia de lipasas II GDSL. Pueden identificarse los residuos conservados GDSIT, donde se sitúa la serina que participa en el sitio activo (figura II.3.26).

Por su parte, la enzima codificada por el clon 8.26 pertenece a la familia de lipasas tipo VII, familia de la que no se conocen muchos miembros.

Las proteínas de esta familia poseen gran homología con acetilcolinesterasas de eucariotas. Suelen ser enzimas con un gran interés en la industria, como la esterasa de *Arthrobacter oxydans*, enzima muy interesante biotecnológicamente porque posee actividad frente a determinados herbicidas (Pohlenz, H.D et al, 1992) o la esterasa de *B. subtilis*, que posee una actividad alta en la hidrólisis de ésteres de para-nitrobenzil.

Se han identificado, en este caso, cuatro bloques de aminoácidos conservados (Nthangeni et al, 2005) en función de los alineamientos de la secuencia de la proteína 8.26 con otras proteínas depositadas en las bases de datos (figura II.3.26).

A)

		S		D H
8.17	(159)	LLHGDSTIDWWLQ (187)	NFAVAGDTTQG (306)	ESFRPDNLHPQAKGY
<i>O.bacterium</i> ZP_02011523	(68)	IMFLGDSITFGWKD (99)	NFGLGGDRTEN (210)	KAFKPDNLHLSVGY
<i>O.bacterium</i> ZP_02011142	(65)	VMFLGDSITAGWSN (96)	NFGIGGDRTEH (216)	KSIMPDFLHLSVAGY
<i>L.araneosa</i> ZP_01876060	(66)	VVFLGDSITHAWDR (96)	NIGFSGDRTEH (217)	KNVMRDLHPNNEMY
<i>L.araneosa</i> ZP_01873049	(65)	LIFIGDSITHFYDR (95)	NLGFGGNTTCK (216)	RDLMPDLLHPNTKGY
Consenso		GDSIT	N G	D LH

I

II

V

B)

		I		II
8.26	(73)	WEGVFKANAFGPSC (163)	PVMWIIYGGGFTGGSGGQAWYDG	
<i>B.subtilis</i> P37967.2	(48)	WEDVLDTAYGSIC (98)	PVMWIIHGGAFYLGAGSEPLYDG	
<i>S.coleicolor</i> CAA22794.1	(52)	WDGVRDAGSFGPTA (104)	PVLVWLHGGALTRGSSAVPVYDG	
<i>A.oxydans</i> Q01470.1	(49)	WTGVRDASAYGPSA (101)	PVLVWIIHGGLLTGSGNLPNYAT	
Consenso		W A G	PV VMIHGG G	

8.26	(243)	AAFGGDPNNVTIAG (456)	KGYTYFFTRV-PLRNGQPQNGASHTAEISYAFNH-
<i>B.subtilis</i> P37967.2	(174)	SAFGGDPDNVTVFG (379)	PVWMYRFDW----HPKPPYKAFHALELPFVFGNL
<i>S.coleicolor</i> CAA22794.1	(176)	AAFGGDPDRVTVAG (389)	HV--YEFWPSNLPD----LGACHALELGFVFDG
<i>A.oxydans</i> Q01470.1	(173)	AAFGGDPNRITLVG (381)	PVRAYQFDVVSPLSDGA---LGAVHCIEMPTFANL
Consenso		AAFGGDP T G	Y F A H E F

III

IV

**Figura II.3.26.** Alineamiento de las proteínas 8.17 y 8.26 con otras depositadas en las bases de datos. **A)** Alineamiento de la proteína 8.17 con otras pertenecientes a la familia II e identificación de los residuos conservados. **B)** Alineamiento de la proteína 8.26 con proteínas pertenecientes a la familia VII e identificación de los bloques conservados.

### II.3.3.2. Caracterización bioquímica preliminar

Una vez identificados los genes causantes de la actividad lipolítica dentro de los clones y analizadas sus secuencias, el siguiente objetivo se centró en el estudio de las características bioquímicas de estas enzimas, al igual que se hizo en el caso de las enzimas encontradas en la genoteca de la Estación Experimental del Zaidín (capítulo X).

Para ello, se partió del plásmido pBK-CMV con el inserto de ADN metagenómico resultante de la escisión del fago cuya actividad lipasa era positiva.

Se han elaborado diferentes ensayos de pH y temperatura óptima que se resumen en la tabla siguiente. La temperatura óptima se encuentra entre los 37 y 10°C. Al igual que en el caso de las esterasas aisladas de la genoteca EEZ, estas temperaturas son relativamente bajas si se comparan con las características de otras enzimas aisladas de forma similar. Con respecto al pH óptimo, poseen pHs ligeramente básicos, lo que podría ser interesante para su posible aplicación en la industria.

Enzima	pH óptimo	T <sup>a</sup> óptima (°C)
6.2	8,25	37
8.1	8	22
8.25	8,25	37
8.17	9,1	10
8.26	8	37

**Tabla II.3.13.** pH y temperatura óptima de las enzimas aisladas de las genotecas mex-lluviosa y mex-seca. Los ensayos se realizaron por triplicado y en la tabla se representa la media de estas tres réplicas.

De este modo, en las genotecas mex-lluvioso y mex-seco se han aislado 7 enzimas con carácter lipolítico. De estas 7, hasta el momento sólo se han podido describir y someter a ensayos preliminares a 5 de ellas, que apuntan importantes diferencias bioquímicas respecto a las encontradas en el suelo de la EEZ.

## DISCUSIÓN

Los trabajos realizados en este capítulo han permitido identificar dieciséis esterasas/lipasas de origen microbiano no descritas con anterioridad y con elevada actividad sobre ésteres de ácidos grasos de cadena corta (algunas de ellas con actividad también sobre ésteres de ácidos grasos de cadena larga), después de realizar una búsqueda funcional para enzimas que hidrolizan este tipo de compuestos en tres genotecas construidas a partir de ADN procedente de tres ambientes totalmente diferentes (figura II.1).

### **I. METAGENÓMICA COMO FUENTE DE NUEVAS ENZIMAS CON CARÁCTER LIPOLÍTICO**

En la actualidad, numerosos programas de búsqueda de nuevas actividades están enfocados al análisis de la diversidad enzimática natural, usando para ello como herramienta la metagenómica.

De este modo, se pueden encontrar enzimas con nuevas actividades y muy poco relacionadas con las enzimas ya descritas aisladas mediante métodos tradicionales de cultivo o mediante PCR. Estudios de este tipo se realizan con éxito en diferentes ecosistemas de suelo (Elend et al, 2006, 2009; Sharma et al 2010; Cieslinki et al, 2009). Basándonos en la diversidad enzimática natural, en este capítulo de esta Tesis Doctoral se ha desarrollado una aproximación metagenómica para la búsqueda de enzimas esterasas en suelos.

Hemos observado diferencias significativas entre la aparición de clones positivos con actividad esterasa y el tamaño de las genotecas construidas. La genoteca EEZ es la más pequeña (65000 ufc) y, sin embargo, en ella es donde se han encontrado más proteínas con carácter lipolítico (9; tabla II.3.5).

Las dos genotecas de México tienen un tamaño muy superior (250 000 ufc) y en ellas se han encontrado sólo dos clones con actividad lipolítica en la de mex-seco y 5 clones en la de mex-lluvioso (tabla II.3.12).

Las diferencias en los rendimientos no se pueden explicar remitiéndose al sistema de búsqueda ni a la construcción de las genotecas, ya que éstos son iguales en los tres casos. Sin embargo, sí se pueden relacionar con las diferencias intrínsecas en los ecosistemas de estudio.

El suelo de la EEZ es un suelo muy manipulado y contaminado por el hombre, por lo que puede ser que se haya enriquecido en microorganismos

que poseen este tipo de enzimas con el paso del tiempo, debido a la contaminación con diversas sustancias. Los suelos procedentes de México son, por el contrario, de un parque natural conservado y protegido, muy poco manipulado con el hombre. Además, en estos suelos encontramos muchas menos actividades lipolíticas (2 frente a 5) en época seca frente a época lluviosa, lo cual podría indicar la baja actividad metabólica que pudiera existir en época seca por las condiciones tan extremas de esta estación.

Además de encontrarse diferencias en el hallazgo de genes esterasa en las tres genotecas, también existen discrepancias en el tipo de genes encontrados.

Si bien la identidad a nivel de proteína es más o menos similar en las tres genotecas (tablas II.3.5-II.3.12; todas las proteínas se parecen menos de un 58% en el caso de las genotecas de México, y de un 61 % en el caso de las genotecas de la EEZ a las secuencias depositadas en las bases de datos), el tipo de proteínas esterasa encontradas sí que difiere.

En el caso de la genoteca de la EEZ, la determinación de la estructura primaria de estos genes permite clasificar estas enzimas en algunas de las subclases ya descritas (3 proteínas de la familia II de esterasas, 3 de la familia IV, 2 de la familia VIII; Arpigny and Jaeger, 1999), y algunas no se engloban en ninguna de las clases descritas hasta el momento (como la pE3).

Los genes esterasa aislados de las genotecas de México codifican para proteínas que se engloban en familias representadas por muy pocos miembros (por ejemplo la 8.26 es miembro de la familia VII) o incluso codifican para proteínas en las que no se identifican los residuos conservados de ninguna familia de esterasas, formando en los árboles filogenéticos grupos específicos junto con otras proteínas (proteínas 8.1, 6.2.1 y 6.2.2; figura II.3.25). Esto podría indicar que forman parte de una nueva clase de enzimas.

Los microorganismos de los que podrían proceder estas enzimas (tabla II.3.5), en el caso de la genoteca de la EEZ, son fundamentalmente *Proteobacteria*, *Plantomycetes*, *Firmicutes* y *Actinobacteria*. Llama especialmente la atención que, en principio, las enzimas obtenidas no provengan de bacterias del género *Pseudomona*, siendo este tipo de bacterias las que presentaron un perfil positivo para esta actividad en el análisis preliminar basado en microorganismos aislados en placa (II.3.1). Ello indica que la ventana a la que hemos tenido acceso mediante el enfoque

metagenómico es, en todo punto, distinta a un enfoque basado en cultivables para esta misma muestra de suelo.

En el caso de las enzimas procedentes de las genotecas de México, las proteínas con las que guardan más homología provienen de bacterias con muy pocos representantes en cuanto a individuos cultivables se refiere, como *Acidobacteria*, *Bacterioidetes* y *Verrumicrobia*, pero son grupos mayoritarios en este tipo de suelos (capítulo I).

Estos resultados, en su conjunto, demuestran que la diversidad microbiana es una fuente importante de productos y procesos biotecnológicos y además auguran la existencia de nuevas enzimas en distintos ecosistemas.

En los apartados siguientes nos centramos en las características de las enzimas con carácter lipolítico aisladas de la genoteca EEZ, por ser éstas las mejor caracterizadas en esta Tesis Doctoral. Estudios posteriores a este trabajo nos revelarán las características más peculiares de las enzimas aisladas de las genotecas mex-seco y mex-lluvioso.

### **III. CARACTERÍSTICAS DE LAS ENZIMAS LIPOLÍTICAS AISLADAS EN LA GENOTECA EEZ.**

Como ya hemos comentado en el apartado anterior, de la genoteca EEZ se aislaron 9 enzimas con carácter lipolítico.

Así, las enzimas pE1, pE2 y pE8 pertenecen a la clase II de esterazas (Akoh et al, 2004), ya que se pueden identificar los residuos conservados del centro activo (figura II.3.4). La clase de esterazas GDSL es una clase relativamente reciente de enzimas que se caracterizan por poseer los residuos conservados Ser-Gly-Asn-His en cuatro bloques diferentes; la serina y la histidina, en este caso, forman parte del centro activo. Nuestras proteínas también presentan estos cuatro residuos estrictamente conservados.

Las secuencias pE2 y pE8 parecen estar relacionadas entre ellas y a su vez con otras enzimas pertenecientes a las alfa-proteobacterias (figura II.3.3). La pE1 está más relacionada con las esterazas pertenecientes a procariotas y con las del género *Verrumicrobia*.

Enzimas de esta clase también han sido aisladas mediante técnicas metagenómicas en muestras de suelos procedentes de un bosque de Corea (Kwang et al, 2007) o en genotecas construidas a partir de ADN de microorganismos que habitan en el rumen (Ferrer et al, 2005).



Esta familia posee enzimas con características muy interesantes, como una amplia especificidad de sustrato y regioselectividad, que se debe a que el tiene un centro activo flexible que cambia de conformación con la unión de diferentes sustratos (Akoh et al, 2004). En nuestro caso, las enzimas pE1, pE2 y pE8 no poseen amplia especificidad con los sustratos ensayados (tabla II.3.7), aunque sí poseen características significativas, como es el caso de la pE2, que es una enzima muy estable.

Las proteínas pE4, pE7.1 y pE12 pertenecen a la clase IV, denominada “The hormone sensitive lipase”. Se llaman así porque las lipasas presentes en esta familia poseen una secuencia aminoacídica que tiene gran similitud con las lipasas sensibles a hormonas de los mamíferos (estas lipasas participan en el metabolismo lipídico regulado por una señal hormonal en mamíferos). De hecho, los residuos del sitio activo se infirieron de los modelos en 3D de las lipasas de mamíferos. La proteína pE7 está más relacionada con las proteínas esterasas de procariotas, mientras que la pE12 está más relacionada con otras procedentes de las beta-proteobacterias, y la pE4 no se adscribe a ningún grupo claro en el árbol filogenético.

Estas proteínas suelen tener una alta actividad a bajas temperaturas, por debajo de los 15°C (Yeaman et al 2004; Rhee J et al, 2005, Manco G et al, 1998). Dicha característica se corrobora con nuestros datos, sobre todo en el caso de la enzima pE12 (figura II.3.23). Estos grupos de lipasas poseen enzimas con importantes aplicaciones biotecnológicas potenciales (Jaeger et al, 1998). Numerosos ejemplos de este tipo de proteínas se han aislado mediante aproximaciones metagenómicas (Jeon et al, 2009; Yuhong et al, 2009; Kwang et al, 2009).

Finalmente, las proteínas pE7.2 y pE6 pertenecen a la familia VIII de lipasas, que son enzimas muy relacionadas con beta-lacatamasas, aunque ninguno de los dos clones confiere resistencia a ampicilina, al igual que se ha descrito en otros clones de esta misma familia aislados de suelo y de agua (Elend et al, 2006).

### **III.1. Características bioquímicas de las enzimas aisladas**

Las especificidades de sustrato de las lipasas clonadas, así como sus características cinéticas, indican claramente que pertenecen al grupo de las esterasas (EC 3.1.1.1). De estas 8, 6 hidrolizan ésteres de ácidos grasos de cadena corta y 2 ésteres de ácidos grasos de cadena larga (tabla II.3.8). La

poca proporción de enzimas lipasas verdaderas encontradas en nuestros análisis puede deberse a que el sistema de búsqueda con el sustrato naftil acetato seleccione fundamentalmente enzimas esterazas frente a lipasas (aunque el sustrato puede ser degradado por ambos tipos de enzimas).

La temperatura óptima de las enzimas (tabla II.3.6) es, en general, ligeramente más baja (entre 10-37°C) que la de otras enzimas aisladas mediante estrategias similares (Elend et al, 2006; Jeon et al, 2009) o que la que se ha descrito como habitual para las lipasas: entre 30-60°C (Gupta R et al, 2004). Esto podría deberse a las peculiaridades del ecosistema, del suelo. Con respecto al pH óptimo, ronda la neutralidad o es ligeramente básico, siguiendo la norma de la mayoría de las proteínas esterasa (Gupta R et al, 2004). En este sentido, es importante mencionar que las enzimas del tipo HSL (como la pE4) se caracterizan por tener una temperatura óptima baja.

Todas las enzimas se sobre-expresaron en *E. coli* y se purificaron siguiendo un esquema cromatográfico en el que se han usado columnas de intercambio iónico, de hidrofobicidad y de filtración molecular. A pesar de la sobreexpresión de todas las proteínas, sólo se consiguieron purificar mediante este método las pE1, pE2 y pE4. Esta purificación nos permitió calcular la actividad enzimática específica, la cual es alta para el sustrato pNpAcetato, y corroborar los datos bioquímicos que poseíamos de los extractos, observando de este modo que la enzima pE2 era una proteína muy estable.

### **III.2. Resistencia a solventes orgánicos y detergentes de las enzimas lipasas**

Los ensayos de resistencia a detergentes nos indicaron que la enzima más estable en este tipo de medios es la enzima pE12. La activación por detergentes de tipo tritón-x-100 es una característica bien estudiada de las enzimas lipasas verdaderas, como parece que es el caso de esta proteína.

En el caso de los ensayos relacionados con la resistencia a disolventes orgánicos, cabe destacar la actividad de la proteína pE2, que es muy estable en los disolventes orgánicos ensayados al 90%: acetona, etanol, isopropanol, dimetilsulfóxido (DMSO), acetonitrilo y tolueno, al contrario que el resto de las enzimas ensayadas. Además, mantiene su actividad al 80% tras su incubación con etanol al 75% durante 24 horas.

La estabilidad en solventes orgánicos es una propiedad que adquiere vital importancia en la síntesis de distintos compuestos.

Habitualmente, las lipasas verdaderas (no está descrito en las esterasas) son estables en solventes orgánicos, con unas pocas excepciones de inhibición o de activación. El perfil de hidrólisis de sustratos de la pE2 nos indica claramente que es una esterasa, no una lipasa verdadera, pero que posee esta propiedad típica de las lipasas: es estable en solventes orgánicos (figura II.3.9). En el caso de la pE12, que es una lipasa verdadera, es la siguiente enzima más estable en este tipo de sustancias, sobre todo a concentraciones no muy altas de solvente (del 20 al 30%; figura II.3.9).

Hay diferentes ejemplos de los efectos de solventes orgánicos sobre lipasas de bacterias cultivables. Por ejemplo, la acetona, el etanol y el metanol al 30% activan la actividad lipasa de *B. thermocatenuatus* (Schmitdt-Dannert *et al*, 1994). Sin embargo, si la incubación se continúa durante una hora y media, la actividad disminuye.

En el caso de la lipasa de *P. aerugiona* YS-7, la acetona inhibe su actividad (Sugihara *et al*, 1991). También presentan una gran estabilidad frente a solventes orgánicos las lipasas de las bacterias *Thermoanaerobater thermohydrosulfuricus* SOL1 y *Caldanaerobacter subterraneus* (Royter *et al*, 2009), manteniendo la actividad a un 90% de solvente durante una hora y media.

No obstante, la enzima pE2 es más estable que estos ejemplos, debido a que mantiene la actividad tras la incubación durante 6 horas con los solventes a una concentración del 90%.

Con respecto a las características relativas a la resistencia en solventes orgánicos de esterasas de origen metagenómico, la enzima O.16 aislada de lagos hipersalinos de Ucrania (Ferrer M, *et al*, 2005) es una enzima lipasa y posee una elevada estabilidad frente a los solventes etanol y DMSO al 90%. Sin embargo, su actividad se inhibe por acetonitrilo y acetona (al contrario que ocurre con la enzima pE2, que es estable en todos los solventes enumerados).

Hay una gran demanda de este tipo de enzimas resistentes a solventes en la industria farmacéutica para la síntesis asimétrica, con el fin de poder tener compuestos farmacológicos enantiopuros. Muchos de estos compuestos son insolubles e inestables (Obino H *et al*, 2001) en agua, en ocasiones el equilibrio termodinámico de la reacción no es bueno en soluciones acuosas, por lo que obtener el producto deseado es muy difícil.

Estas características hacen que sea interesante disponer de enzimas estables en medios apolares. En este caso, la pE2 parece ser una buena

candidata para este tipo de aplicaciones. Para este ensayo, en un futuro, tendríamos que realizar además ensayos específicos de enantioselectividad con sustancias quirales. Además, otra posible aplicación es en la síntesis de ésteres, ya que en este tipo de medio se favorece la síntesis de ésteres de ácidos grasos a partir de ácidos y alcohol (Zaks et al, 1995).

Se ha observado que el extremo N-terminal de esta proteína (hasta el aminoácido 166) es indispensable para la actividad de la misma, aunque en él no se encuentra ninguno de los 4 bloques conservados identificados para este tipo de enzimas (Akoh et al, 2004). Además, si se pasa este módulo de la proteína a otra proteína de características similares (como la pE1), no se transfiere el carácter de resistencia a solventes orgánicos, aunque se mejora la actividad enzimática y la estabilidad en detergentes (caso de la pE1Q; figura II.3.21).

### **III.3. Lipasa verdadera pE12**

En este trabajo también se ha descrito, purificado y caracterizado la lipasa verdadera pE12.

Esta enzima es especialmente atractiva, porque posee actividad sobre ácidos grasos de cadena larga (tabla III.3.22) y es muy estable en un amplio rango de temperaturas. Esta última característica puede estar relacionada con el ecosistema de origen, debido a que en el suelo normalmente se registran temperaturas extremas entre el día y la noche, y puede ser que el microorganismo del cual procede esté adaptado a estas condiciones.

Además, presenta activación por el catión  $K^+$  e inactivación por el catión  $Ca^{+2}$ . Los cofactores generalmente no se requieren para la actividad lipolítica, pero los cationes divalentes como el calcio normalmente estimulan la catálisis enzimática. Se pueden enumerar numerosos ejemplos de lipasas activadas por el calcio, como por ejemplo las lipasas de *B. subtilis* 168 (Lesuisse et al, 1993), *B. thermoleovorans* ID-1 (Lee et al, 1999), *P. aeruginosa* EF2 (Gilbert et al, 1991), *S. aureus* 226 (Muraoka et al, 1982), *S. hyicus* (Van Oort et al, 1989), *C. viscosum* (Sigiura et al, 1974) y *Acinetobacter sp.* RAG-1 (Snellman et al, 2002); pero también hay ejemplos de lipasas que, como la pE12, se inactivan con el calcio: la lipasa de *P. aeruginosa* 10145 (Finkelstein et al, 1970).

En el caso de lipasas de origen metagenómico, hay datos tanto de activación por calcio como de inactivación. En este último caso, encontramos por ejemplo la enzima EstA3 derivada de una genoteca construida a partir de

muestras de agua (Elend C. et al, 2006) o las lipasas EstAT1 y EstAT11, derivadas de sedimentos del ártico (Jeon et al, 2009).

Esta enzima, la pE12, presenta además características interesantes por ser capaz de hidrolizar diferentes sustratos de importancia industrial o farmacéutica (tabla II.3.10).

En trabajos posteriores se deberá profundizar en la estructura de estas proteínas para obtener más información de sus mecanismos de catálisis, su estabilidad frente a solventes orgánicos (en el caso de la pE2) o su estabilidad a distintas temperaturas (pE12).

Además, el estudio de estas enzimas puede combinarse con técnicas de evolución dirigida, herramienta que se ha mostrado muy eficaz para el desarrollo de nuevas actividades partiendo de enzimas ya conocidas (Alcalde et al, 2003). Si esta herramienta se combina con la diversidad enzimática aún sin explorar, nos podemos encontrar en una situación ventajosa que implicaría un avance sustancial en la Biotecnología y Microbiología actuales.

Proyectos conjuntos que contemplen el descubrimiento de nuevas enzimas a partir de metagenomas y su posterior mejora mediante técnicas de evolución dirigida o ingeniería de la reacción y del biocatalizador, por ejemplo técnicas de inmovilización, son por tanto proyectos ambiciosos y de gran interés, tanto desde un punto de vista básico, conocimiento de nuevas estructuras y actividades, como desde un punto de vista aplicado, nuevas actividades que pueden trasladarse a una industria química verde libre de subproductos y respetuosa con el medioambiente.

## **Anexo Capítulo II3**



**Figura AII.3.1.** Mapa físico y posición de los genes identificados por secuencia en cada uno de los clones asilados. Los genes correspondientes a la actividad esterasa/lipasa se destacan en negro.

Plásmido	Mapa físico	Subclon/Clon	Actividad esterasa
a)pE1		pE1.1	+
		pE1.2	+
		pE1	+
b)pE2		pE2.1	+
		pE2	+
c)pE3		pE3.1	-
		pE3	+
d)pE4		pE4.1	-
		pE4	+
e)pE6		pE6	+
f)pE7		pE7	+



**Figura AII.3.1.** (Continuación)

Plásmido	Mapa Físico	Subclon/Clon	Actividad esterasa
g) Clon 8		pE8.1	-
		pE8.2	-
		pE8	+
h) Clon 12		pE12.1	-
		pE12.2	-
		pE12	+

**Tabla AII.3.2.** Tabla con los diferentes marcos abiertos de lectura identificados en los clones con carácter lipolítico aislados de la genoteca EEZ

- a) Datos obtenidos del *BLASTp* en la base de datos del NCBI  
 b) Web para el cálculo del contenido en g+c: <http://tubic.tju.edu.cn/GC-Profile/>  
 b) Definidos por Arpigny and Jaeger et al 1999.  
 e) \* El codon de iniciación es “gtg” en lugar de “atg”.

Plásmido (tamaño inserto)	Secuenciado nt/	G+C (%) <sup>b</sup>	Hipotetical CDS	Tamaño proteína (aa)	Pi	E-value <sup>a</sup>	Aa identidad/similitud % <sup>a</sup>	Familia esterasa/lipasa <sup>c</sup>	Dominios conservados/residuos catalíticos <sup>c</sup>	Proteína homóloga (Número de acceso) <sup>a</sup>	Posible microorganismo de origen	Características principales del microorganismo de origen
<b>pE1 (6Kbp)</b>	1951/XXXXXX	64,48	<b>Orf 1 325-1041 (reverse)</b>	<b>238</b>	<b>9,53</b>	<b>1<sup>-68</sup></b>	<b>58/70</b>	<b>II /GDSL</b>	<b>Posible lipasa_GDS(L) tríada catalítica: S<sub>62</sub> D<sub>211</sub> H<sub>214</sub></b>	<b>Proteína G-D-S-L (ZP_02011142)</b>	<i>Opiritaceae bacterium TAV2</i>	<b>Grupo de Clamydia y verrucomicrobia. Un aislado de suelo de arroz.</b>
			Orf 2 >1301-1951 (direct)	>217	--	3 <sup>-14</sup>	33/50	--	--	Proteína repetida Tetratricopeptido	<i>Acidobacteria bacterium Ellin345</i>	Subdivisión 1 del phylum <i>Acidobacteria</i> aislado de suelo de Australia .
<b>pE2 (8Kbp)</b>	3541/xxxx	71.48	Orf 1 >543-1(reverse)	>181	--	2e-67	68/81	--	--	Ribonucleasa (NP_101927 )	<i>Mesorhizobium loti MAFF303099</i>	Grupo alpha Proteobacteria. Bacterias Gram negativas que crecen en ambientes acuáticos .
			Orf 2 593-1479* (reverse)	294	--	9e-74	54/64	--	--	Polisacarido deacetylasa	<i>Caulobacter sp. K31</i>	
			<b>Orf 3 2693-1476 (reverse)</b>	<b>405</b>	<b>9,25</b>	<b>9e-127</b>	<b>58/73</b>	<b>II/GDSL</b>	<b>Posible lipasa_GDS(L) tríada catalítica: S<sub>202</sub> D<sub>382</sub> H<sub>386</sub></b>	<b>Lipolytic enzyme, G-D-S-L (ZP_01418050)</b>	<i>Caulobacter sp. K31</i>	
			Orf 4 >256 (direct)	256	--	3e-65	66/76	--	--	Proteína hipotética conservada (ZP_01420908)	<i>Caulobacter sp. K31</i>	

pE3 (10Kbp)	2443 / XXXX	40.48	Orf 1 692-2265 (direct)	473	9,76	7e-154	54/72	--	--	Sialato O- acetilsterasea(YP_001 196427)	<i>Flavobacterium johnsoniae UW101</i>	Grupo Bacteriodetes .
	921/XXX		Orf2 > 1- 625 (direct)	>207	--	2e-38	50/68	--	--	Transportador de Na+ (ZP_00831497)	<i>Yersinia frederiksenii ATCC 33641</i>	Las gamma proteobacterias son abundantes en el suelo y el agua de ambientes tropicales y subtropicales
			Orf 3 399-1 (reverse)	>133	--	8e-20	43/63	--	--	Proteína hipotética PBAL39_09221 (ZP_01883200)	<i>Pedobacter sp. BAL39</i>	Grupo bacteriodetes. Estos microorganismos se caracterizan por la capacidad de utilizar múltiples compuestos naturales
			Orf 4 516-921 (direct)	>135	--	4e-34	54/73	--	--	Regulador transcripcional familia BadM/Rrf2 (YP_001194757)	<i>Flavobacterium johnsoniae UW101</i>	
pE4 (8 kbp)	1726/xxx	73.04	Orf 1 1320-394 (reverse)	308	4,29	8e-72	52/64	IV/HSL	Esterasa/lipasa /thioesterasa.T riada catalítica: S <sub>156</sub> D <sub>250</sub> H <sub>280</sub>	Esterasa/lipasa/thioest erasa (YP_299651)	<i>Ralstonia eutropha JMPI34</i>	Grupo de las beta proteobacterias. Orden <i>Burkholderiales</i> . Este grupo se ha encontrado ampliamente distribuido en agua y suelos y tiene un gran potencial en bioremediación.
pE6 (3 kbp)	2981/XXX	66.69	Orf 1 147-758 (direct)	201	5,34	3e-44	67/55	--	--	Carboximuconolactona decarboxylasea(YP_371 718)	<i>Burkholderia sp. 383</i>	Grupo beta proteobacteria.
			Orf 2. 803-1097 (direct)	97	--	2e-26	75/62	--	--	Proteína hipotética BAL199_03004 (ZP_02187417)	<i>Alpha proteobacterium BAL199</i>	Microorganismo marino de alpha proteobacteria no clasificada.
			Orf 3 2254- 1103 (reverse)	397	4.62	1e-131	73/61	VIII	Desconocido	Lactona hidrolasa (AAL14234)	<i>Rhodococcus ruber</i>	Clase <i>Actinobacteria</i> . Microorganismos capaces de degradar una gran variedad de polychlorados
			Orf 4 2834- 2267 (reverse)	189	--	2e-16	53/33	--	--	Proteína hipotética Maqu_0265 (YP_957558 )	<i>Marinobacter aquaeolei VT8</i>	Grupo gamma proteobacteria. Bacteria degradadora de hidrocarburos ligeramente halófila.

pE7 (3 kbp)	3121/XXX	54.50	Orf 1 1049-42 (reverse)	341	6.70	1e-44	58/38	IV/HSL	Carboxilestera sa. Tríada catalítica: S <sub>185</sub> D <sub>281</sub> H <sub>311</sub>	carboxilestera A (YP_001485721)	<i>Bacillus pumilus</i> SAFR-032	Phylum <i>Firmicutes</i> .
			Orf 2 1468-2514 (direct)	348	5.12	1e-54	55/39	VIII		esterasa (YP_001624002)	<i>Renibacterium</i> <i>salmoninarum</i> ATCC 33209	Acidobacteria.
			Orf 3 2573- 312	>183	--	7e-48	53/70	IV/HSL		Enzima lipolítica (ABU55418)	<i>Sphingobium</i> <i>chungbukense</i>	Grupo alpha proteobacteria.
pE8 7 (Kpb)	2460/xxx	64.98	Orf1 1 1440-1 (reverse)	>183	--	7e-48	70/53	--	--	Enzima lipolítica (ABU55418)	<i>Sphingobium</i> <i>chungbukense</i>	Grupo alpha proteobacteria.
			Orf 2. 1500-... (reverse)	>481		1e-162	74/59	--	-	Proteína hipotética Sde_1755 (YP_527227)	<i>Saccharophagus</i> <i>degradans</i> 2-40	Gamma proteobacteria que degrada carbohidratos poco solubles.
pE12 6(Kpb)	2327/xxx	68.34	Orf 1 2327-1746 (reverse)	>181	--	7e-79	86/77	--	--	Sintetasa y ligasa dependientes de AMP (ZP_01422009)	<i>Caulobacter sp. K31</i>	Grupo alpha proteobacteria
			Orf 2 1737-806 (reverse)	309	5.51	3e-60	59/43	--	--	lipasa/esterasa (AAZ32717)	<i>Uncultured bacterium</i>	

**Tabla AII.3.2.** Oligonucleótidos usados en este estudio.

<b>Nombre</b>	<b>Secuencia</b>	<b>Aplicación</b>	<b>Ref.</b>
T3	5' -ATTAACCCCTCACTAAAG-3'	Cebador secuenciar pBCK-CMV	En este trabajo
T7	5' -TAATACGACTCACTATAGGG-3'	Cebador secuenciar pBCK-CMV	En este trabajo
SP6	5' -CATACGATTTAGGTGACACTATAG-3'	Cebador secuenciar pBCK-CMV	En este trabajo
Nde1- pE1	5' -GGCCATATGTCCGAGCGCAGTTTGA-3'	Construcc. pE1 con diana <i>NdeI</i>	En este trabajo
BamH1- pE1	5' - GGCGGATCCTCATCACTTCGCGCTGAGCATC- 3'	Construcc. pE1 con diana <i>BamHI</i>	En este trabajo
Nde1- pE2	5' - GGCCATATGCGGATCCTGTCTGGT-3'	Construcc. pE2 con diana <i>NdeI</i>	En este trabajo
BglII-pE2	5' - GGCAGATCTTCATCACTTGAACAGCGACAG-3'	Construcc. pE2 con diana <i>BglII</i>	En este trabajo
S-71pE1f	5' - TCGGGGACGTTTCCC GCGGCGAGAAGAA- 3'	Mutante <i>SacII</i> pe1	En este trabajo
S- 747pE2f	5' - GGCCACGTGGTTCGCCGCGGACGTAGAAGGTG- 3'	Mutante <i>SacII</i> pe2	En este trabajo
pE2.1 <sub>1317t</sub> 7	5' -CGTTCCTCAACGCCTGGA-3'	Secuenciar clon pE2	En este trabajo
E2.611t7	5' -TCGACGAGCGTCTTCATGGC-3'	Secuenciar clon pE2	En este trabajo
pE2 <sub>2011T7</sub>	5' -TCGGCCTGCAGGCGTTCGGC-3'	Secuenciar clon pE2	En este trabajo
pE3 <sub>1291T7</sub>	5' -ACAGCGGTTTCATTTGCAGAA-3'	Secuenciar clon pE3	En este trabajo
pE4 <sub>511</sub>	5' -GATGATGACCACGATGGTGC-3'	Secuenciar clon pE4	En este trabajo
pE4 <sub>668</sub>	5' -ATCGGCGGCATCGACACC-3'	Secuenciar clon pE4	En este trabajo
pE6 <sub>661t3</sub>	5' -GTCGCCGAGACGTACATCTA-3'	Secuenciar clon pE6	En este trabaj o
pE9.1, <sub>1371</sub>	5' - TTCAAGGTCAGTATGCCGAT-3'	Secuenciar clon pE7	En este trabaj o
pE7.822t 7	5' -CCTCCATGGAAGTTGAAGTG- 3'	Secuenciar clon pE7	En este trabajo

pE9.681T7	5´-CCTCATGACACTGGTACACG-3´	Secuenciar clon pE7	En este trabajo
pE9.731T3	5´-CTTCGTACGCTATCCGAAGG-3´	Secuenciar clon pE7	En este trabajo
pE10.711T7	5´-GAGCAGCAAGCTATCAATGC-3´	Secuenciar clon pE7	En este trabajo
pE7.519T3	5´-TACCCAGTCGACAGGCGTCT-3´	Secuenciar clon pE7	En este trabajo
pE8 <sub>1502,T7</sub>	5´-ATCAATGCTTGGCCACGCT-3´	Secuenciar clon pE8	En este trabajo
pE8731	5´-AACTTGCTGTTGTAGTTGCC-3´	Secuenciar clon pE8	En este trabajo
pE8 <sub>1502,T7</sub>	5´-ATCAATGCTTGGCCACGCT-3´	Secuenciar clon pE8	En este trabajo
pE12 <sub>1401T3</sub>	5´-ACGGATCCGTACACTGACGG-3´	Secuenciar clon pE12	En este trabajo
pE121281	5´-CCGGGCAATGTCGAGAACCT-3´	Secuenciar clon pE12	En este trabajo
pE12 <sub>1281T7</sub>	5´-CCGGGCAATGTCGAGAACCT-3´	Secuenciar clon pE12	En este trabajo
pE12.631t3	5´-CCAAGATCGAGATGCCGCCG-3´	Secuenciar clon pE12	En este trabajo
12 f	5´-GCCCATATGCTCGACCCGCAC-3´	Construcc.pE 12 diana <i>Nde</i> I	En este trabajo
12 r	5´-GCCGGATCCTCATCAGGCGAGCCCGT-3´	Construcc.pE 12 diana <i>Nde</i> I	En este trabajo
814c:	5´-GCGCGCCCTCTCAAACGA -3´	Secuenciar clon 8.14	En este trabajo
814x:	5´-AAGTCTTTATCGCCGGACACA-3´	Secuenciar clon 8-14	En este trabajo
8.14b	5´-CATCCGCCTTCTGGTATCTG-3´	Secuenciar clon 8.14	En este trabajo
8.14y	5´- CCTTGGTGGATTCTGTCAC-3´	Secuenciar clon 8.14	En este trabajo
6.2y	5´-CCGAGTCGCCCATTTCC-3´	Secuenciar clon 6.2	En este trabajo
8.25e	5´- AGTCACACACCGATTTGAGTG-3´	Secuenciar clon 8.25	En este trabajo
8.25y	5´- AGCTGGACGCGATTGCGG-3´	Secuenciar clon 8.25	En este trabajo
6.4c	5´- GGGTTGTTGATCTGCACATCC-3´	Secuenciar clon 6.4	En este trabajo
6.4y	5´-TCGATGAAGTCGGCGCGG-3´	Secuenciar clon 6.4	En este trabajo
8.1b	5´-CTGCCGCGGACAGTGA-3´	Secuenciar clon 8.1	En este trabajo
8.1y	5´-TCGCCCAGGCGAGGAAGA-3´	Secuenciar clon .8.1	En este trabajo
8.17y	5´- AGCAGGGGCGTGCGTTAC-3´	Secuenciar clon 8.17	En este trabajo
8.17d	5´-GATCGCCGGGACGCTG-3´	Secuenciar clon 8.17	En este trabajo
8.26a	5´-CCGCCGGTGAACCGCC-3´	Secuenciar clon 8.26	En este trabajo
P6.2 832r	5´-CAAACGCCTCGCCGACCA-3´	Secuenciar clon 6.2	En este trabajo
P6.2 52r	5´-CATGCCAACTCGCGGGA-3´	Secuenciar clon 6.2	En este trabajo
6.2b	5´-ACTTCCATGGGCGGGATG-3´	Secuenciar clon 6.2	En este trabajo

6.2z	5´-TTCATGTCCCGGGACGAGTA-3´	Secuenciar clon 8.25	En este trabajo
8.25z	5´-CGCTTGTACTCATCGCGCT-3´	Secuenciar clon 8	En este trabajo
8.25d	5´-GTAAGGTTTGCAGTTGAGTCATG-3´	Secuenciar clon 8.26	En este trabajo
6.4a:	5´-CATGACCGGCACGCGGAA-3´	Secuenciar clon 6.4	En este trabajo
6.4z	5´-GCAAGACCTACACGATCTCAG-3´	Secuenciar clon 6.4	En este trabajo
8.1a	5´-GAATTCGCTTCGGCACC-3´	Secuenciar clon 8.1	En este trabajo
8.1z:	5´-TCACGACCTCAACGGCTC-3´	Secuenciar clon 8.2	En este trabajo
8.2a	5´-GCCGGTGGCGCCCTTCA-3´	Secuenciar clon 8.2	En este trabajo
8.2z	5´-GGCGCACCATCGAGCTG-3´	Secuenciar clon 8.2	En este trabajo
8.14z	5´-ACCATGTCAACGGCGGTT-3´	Secuenciar clon 8.14	En este trabajo
8.17c	5´-CCACATCACGATGGTTGAG-3´	Secuenciar clon 8.17	En este trabajo
8.17z	5´-CATCACGATGGTTGAGGCGA-3´	Secuenciar clon 8.17	En este trabajo
8.26z	5´-GGGCCTGCAGATGGGAC-3´	Secuenciar clon 8-26	En este trabajo
6.2y	5´-CCGAGTCGCCCATTTCC-3´	Secuenciar clon 6.2	En este trabajo
8.25 e	5´-AGTCACACACCGATTTGAGTG-3´	Secuenciar clon 8.25	En este trabajo
8.25y	5´-AGCTGGACGCGATTGCGG-3´	Secuenciar clon 8.25	En este trabajo
6.4c	5´-GGGTTGTTGATCTGCACATCC-3´	Secuenciar clon 6.4	En este trabajo
6.4y	5´-TCGATGAAGTCGGCGCGG-3´	Secuenciar clon 6.4	En este trabajo
8.1b	5´-CTGCCGCGGACAGTGA-3´	Secuenciar clon 8.1	En este trabajo
81y	5´-TCGCCCAGGCGAGGAAGA-3´	Secuenciar clon 8.1	En este trabajo
8.14b	5´-CATCCGCTTCTGGTATCTG-3´	Secuenciar clon 8.14	En este trabajo
8.14y	5´-CCTTGGTGGATTTCGTCCAC-3´	Secuenciar clon 8.14	En este trabajo
8.17d	5´-GATCGCCGGGACGCTG-3´	Secuenciar clon 8.17	En este trabajo
8.17y	5´-AGCAGGGGCGTGC GTTAC-3´	Secuenciar clon 8.17	En este trabajo
8.26a	5´-CCGCCGGTGAAACCGCC-3´	Secuenciar clon 8.26	En este trabajo

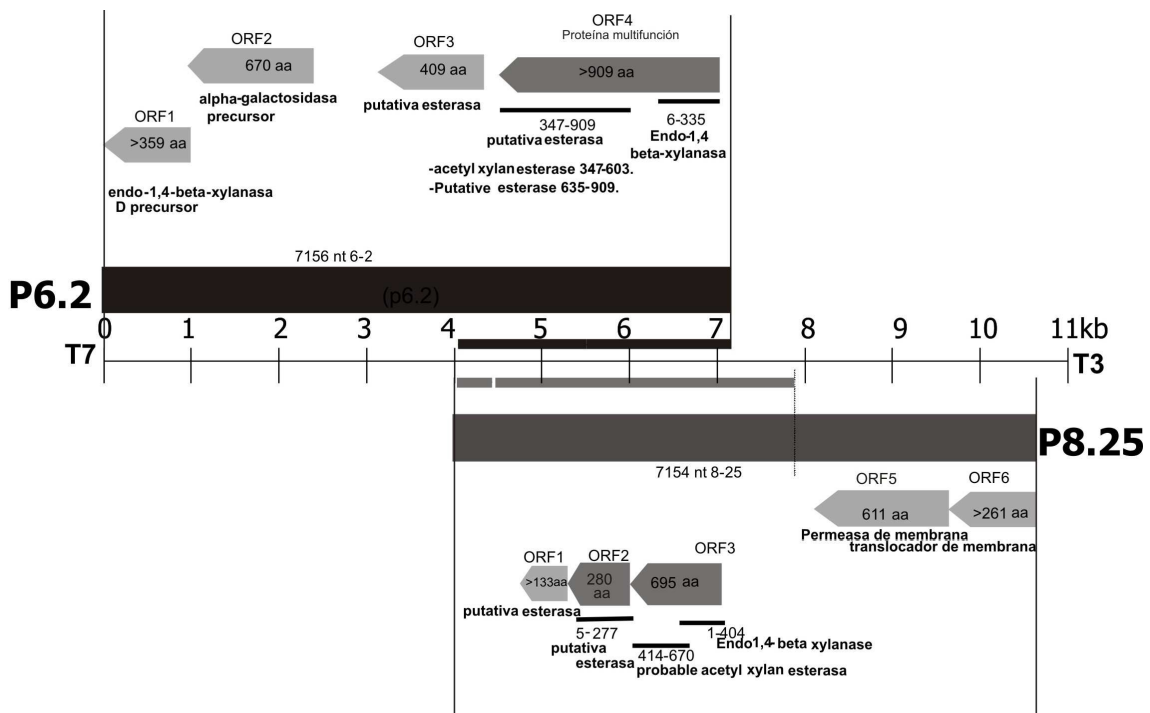
**II.4. ANÁLISIS DE UN BIOMARCADOR EN LA RIZOSFERA DE *Mammillaria carnea***

En la búsqueda de genes que codifican enzimas con carácter lipolítico en las genotecas construidas a partir de ADN de la rizosfera de *Mammillaria carnea* (capítulo II.3), dos de los clones encontrados, 6.2 y 8.5 tenían un alto grado de homología a nivel de nucleótido en unas 3 kb (Figura II.4.1).

El hecho de haberlos encontrado en dos genotecas distintas, 6.2 en la época seca y 8.25 en la lluviosa, es indicativo de que el grupo microbiano al que pertenecen puede ser especialmente abundante en esta rizosfera.

**II.4.1. Análisis de los clones 8.25 Y 6.2**

Para poder responder a la pregunta de qué tipo de microorganismo es el que posee este fragmento genómico y qué genes posee, secuenciamos los fragmentos completos de los dos clones (8.25 y 6.2). En total, 14,310 kba de información genética que nos proporcionó pistas acerca de la posible adscripción filogenética de estos fragmentos.



**Figura II.4.1.** Mapas físicos de los clones 6.2 y 8.25. Se representan los ORFs mediante una flecha. Estos contienen el número de aminoácidos de las proteínas para las que codifican. En algunos casos también se especifican los nucleótidos. Se especifica el nombre de las proteínas con las que guardan similitud en las bases de datos.



De este modo, observamos (figura II.4.1, tabla anexo AII.4.1) que los genes que portaban estos dos fragmentos genómicos correspondían principalmente a secuencias codificantes para enzimas esterases o xilanasas, además de traslocadores de membrana. Los estudios de las secuencias revelan que, teóricamente, el fragmento 6.2 posee una proteína multifunción con varios dominios, esterasa y xilanasas, mientras que esta misma región, en el 8.25, codifica para proteínas independientes.

Para conocer algo sobre el microorganismo de origen de estos fragmentos metagenómicos, usamos una herramienta online :[http://insilico.ehu.es/oligoweb/my\\_distance/](http://insilico.ehu.es/oligoweb/my_distance/), que nos puso de manifiesto que estos fragmentos podrían proceder de una bacteria perteneciente al phylum *Acidobacteria*, phylum mayoritario, además, en este tipo de suelos (ver capítulo I de Resultados).

Estos datos están en concordancia con los datos proporcionados por la página web del NCBI (<http://blast.ncbi.nlm.nih.gov/>) cuando se realiza una búsqueda en la que se comparan nuestras secuencias proteicas con las depositadas en las bases de datos (tabla AII.4.1).

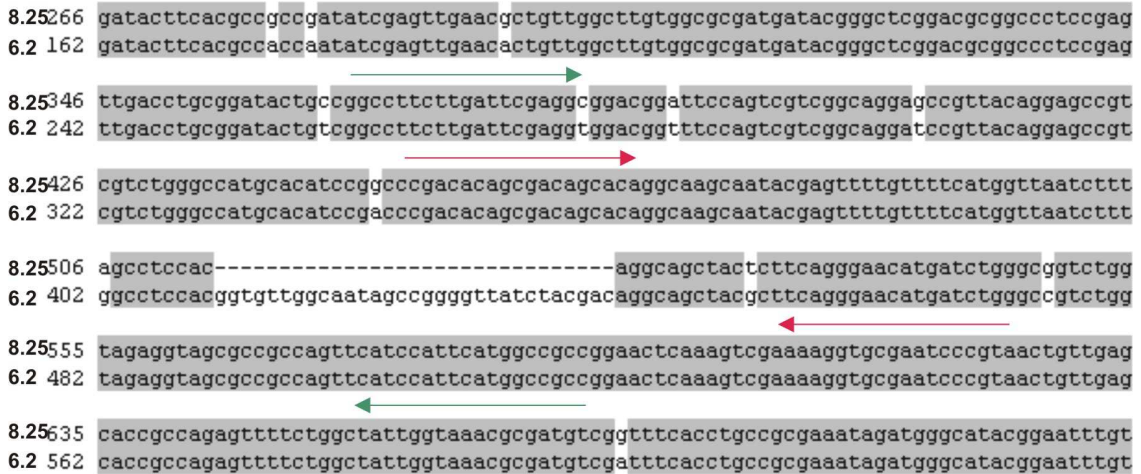
#### **II.4.2. Uso de los clones 8.25 y 6.2 como posibles bioindicadores o biomarcadores**

El análisis de la secuencia de estos clones en detalle permite observar cambios puntuales en la secuencia “común” a ambos fragmentos metagenómicos. Así, el clon 6.2 posee 30 nucleótidos extra en la región intergénica entre el ORF3 y el ORF4, que se encuentran ausentes en el clon 8.25 (en este caso entre el ORF1 y ORF3 (figuras II.4.1 y II.4.2 ))

Esta diferencia nos permitiría avanzar en si se podría utilizar esta región como bioindicador de cambios poblacionales en base a que:

- Ambos fragmentos parecen provenir de uno de los grupos mayoritarios: *Acidobacteria*.
- El haber aislado ambos fragmentos de dos genotecas distintas puede ser indicativo de variaciones en las subpoblaciones dentro de este phyla, género, especie, en definitiva, OTU.
- Una estrategia en base a cebadores anidados puede ser utilizada para tal fin.

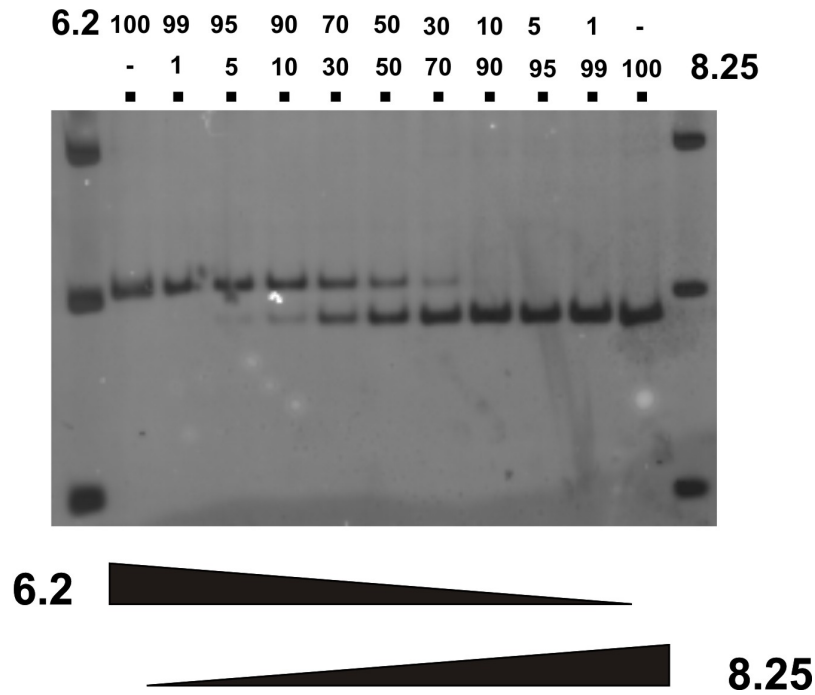
Con el fin de poder observar si el microorganismo al que pertenece el fragmento metagenómico 8.25 o 6.2 es más abundante en unas muestras que en otras, se diseñaron dos parejas de cebadores que amplifican esta región (figura II.4.2).



**Figura II.4.2.** Alienamiento de la secuencia intergénica del clon 8.25 con la del clon 6.2. Las zonas escogidas para el diseño de los cebadores están indicadas mediante flechas: en verde para los cebadores externos y en rojo para los cebadores internos. Dependiendo de si el amplicón procede de un clon u otro, se diferencian en 305 pb para el amplicón procedente del 6.2 y 8.25 pb para el procedente del 8.25.

Una vez diseñados los cebadores, se utilizaron para determinar si se podían utilizar de forma semicuantitativa, por lo que se usaron sobre ADN molde en concentraciones cambiantes de cada uno de los clones 8.25 y 6.2 (Figura II.4.3.). Los amplicones obtenidos se resolvieron en geles de acrilamida al 6 %.

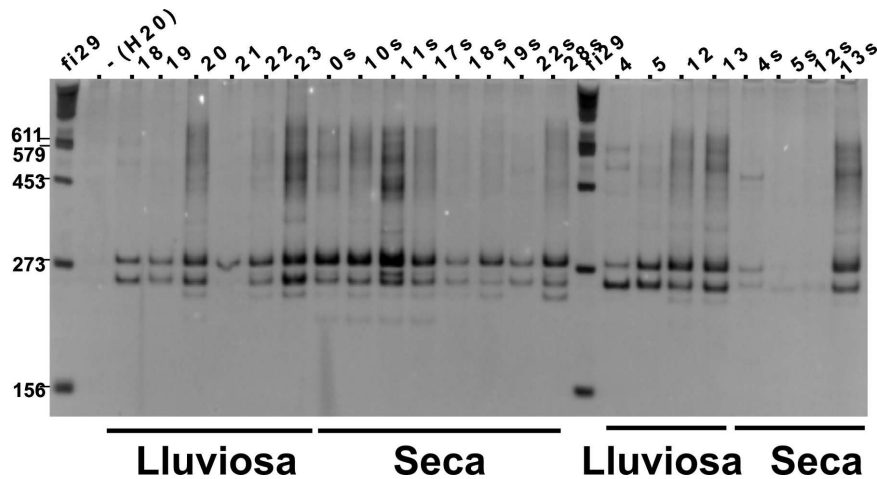
Una vez demostrado que los cebadores funcionaban y que eran capaces de detectar el fragmento 6.2 o el 8.25 a diferentes concentraciones, se hicieron diversas PCRs anidadas a partir del ADN metagenómico de todas las muestras que poseíamos del parque Natural de Tehuacán-Cuicatlán. Gracias a la diferencia de tamaños entre los dos amplicones, pudimos observar si predominaba más un tipo de secuencias que otras en determinadas zonas o si uno de los tipos es específico de unas muestras (figura II.4.3).



**Figura II.4.3.** Gel de poliacrilamida al 6% en el que se observan los fragmentos resultantes de la amplificación de mezclas a distintas concentraciones de los clones 6.2 y 8.25 mediante una PCR anidada con los cebadores externos e internos. En el primer y último carril se observa el marcador de peso molecular. Desde el segundo carril hasta el undécimo se encuentran diferentes mezclas de plásmidos que tienen clonado el gen que codifica para ambas enzimas: 8.25 y 6.2. En el segundo carril, sólo existe el plásmido con el fragmento 6.2 que produce un amplificado de 305 pb, y en el carril undécimo se observa el fragmento correspondiente al plásmido 8.25 de 274 pb

Siguiendo esta estrategia, se realizaron PCRs anidadas a partir del ADN de todas las muestras, se corrieron geles de poliacrilamida al 6% (figura II.4.4) y la intensidad de cada banda se calculó mediante un analizador de imagen (Quanti-One). Los resultados nos permitieron responder a los siguientes interrogantes:

¿Predomina algún morfotipo en general en estos suelos? ¿Existen diferencias entre suelo suelto y rizosfera? ¿Predomina algún morfotipo en la rizosfera de determinadas plantas? ¿Hay diferencias entre época seca y lluviosa? ¿Predomina en diferentes zonas?



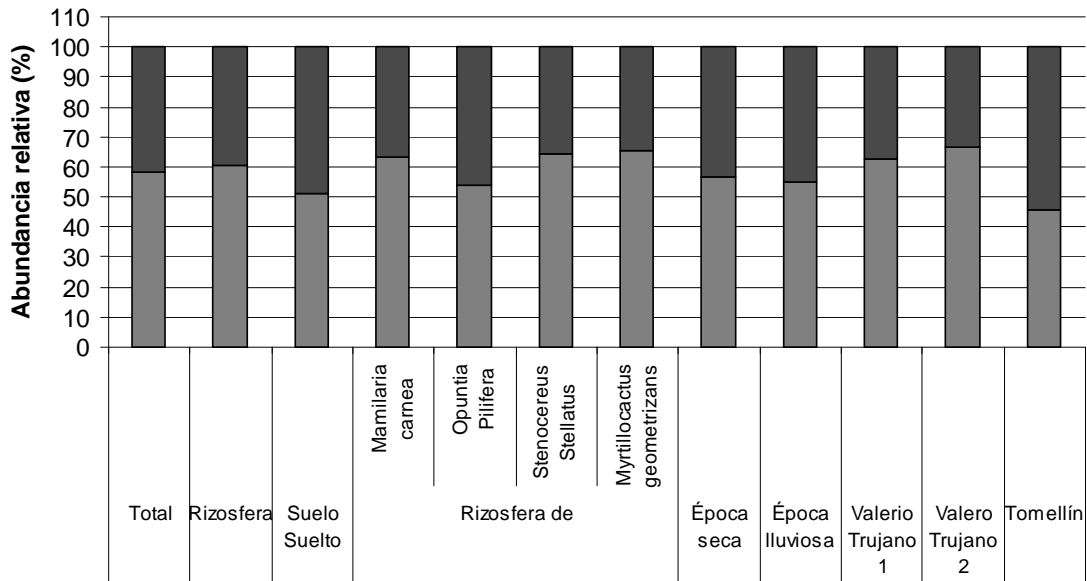
**Figura II.4.4.** Ejemplo de gel de poliacrilamida al 6% con los amplicones de las PCRs anidadas realizadas sobre ADN metagenómico de la zona Valerio Trujano 1. Nótese que en la mayoría de los casos aparecen las dos bandas, correspondientes a los dos morfotipos, aunque las intensidades varían. En alguna muestra también aparece una sola banda (por ejemplo, en la muestra 21).

De este modo, como se observa en la figura II.4.5, vimos que considerando todas las muestras en total, la abundancia del morfotipo 6.2 es relativamente superior a la del 8. 25 (58% frente a 41% respectivamente).

Lo mismo ocurre en la rizosfera de todas las plantas donde predomina el gemovar 6,2. En época seca y lluviosa, prácticamente aparecen ambos morfotipos por igual, sin que predomine uno en ninguno de los casos. Sin embargo, si observamos diferencias en la distribución de los morfotipos por zonas, en todas las zonas el morfotipo que predomina es el 6.2 menos en la zona de Tomellín, donde el morfotipo con más presencia es el 8.25.

En resumen, se puede decir que, en general, en nuestras muestras predomina el morfotipo 6.2, sobre todo en las muestras rizosféricas, y que existe también una distribución heterogénea de este morfotipo por zonas, ya que predomina en las regiones Valerio Trujano 1 y 2 (que son de características similares) y es más minoritario en la zona de Tomellín, donde predomina por tanto el morfotipo 8.25.

Este tipo de estrategia nos ha permitido usar uno de los clones encontrados mediante una aproximación funcional como posible biomarcador o bioindicador de las muestras objeto de estudio. Podemos así concluir que el morfotipo 6.2 es un biomarcador de rizosfera y de las zonas Valerio Trujano 1 y 2, mientras que el morfotipo 8.25 es un bioindicador de los suelos de la zona de Tomellín.



**Figura II.4.5.** Abundancia relativa de las genomovariedades 6.2 (en gris) y 8.25 (representado por el color negro) en las diferentes muestras.

## **DISCUSIÓN**

En este capítulo, se ha desarrollado una herramienta molecular para poder detectar e identificar la abundancia relativa de dos fragmentos de ADN en todas nuestras muestras de suelo. Ambos fragmentos se obtuvieron en las genotecas metagenómicas construidas a partir de las muestras rizosféricas de *Mamillaria carnea*.

Estos fragmentos de ADN deben de proceder de microorganismos muy abundantes en las muestras de procedencia y pensamos en el desarrollo de una herramienta que nos permitiera poder identificar qué tipo de microorganismo predomina en unas muestras u otras: el que poseería el fragmento 6.2, denominado “genomovar 6.2”, o el que poseería el fragmento 8.25, llamado “genomovar 8.25”.

Además, el análisis mediante herramientas informáticas de estos fragmentos de ADN nos ha revelado que podrían proceder de un microorganismo asociado al phylum *Acidobacteria*, el cual es muy abundante en nuestros suelos (capítulo I).

De este modo, gracias a esta herramienta, hemos observado que el morfotipo 6.2 parece ser un bioindicador de muestras rizosféricas y de muestras de la zona de Valerio Trujano, mientras que el morfotipo 8.25 es un bioindicador de la zona de Tomellín. Estos datos estarían correlacionados con los datos de los perfiles de TGGE desarrollados en el capítulo I de esta Tesis Doctoral, en los que se demuestra que la población microbiana de las alfa-proteobacterias es muy diferente entre las zonas Valerio Trujano 1 y 2 y

Así, se ha desarrollado una herramienta molecular sencilla para la detección de este bioindicador, diferente de los estudios tradicionales basados en la subunidad pequeña de gen 16S (Woese et al, 1999) o de otros genes marcadores como el gen RecA, GroEL, etc.



## **ANEXO CAPÍTULO II4**



**Tabla AII.4.1** Características de los ORFs de los clones 6.2 y 8.25.

- a) Contenido en G+C calculado en la página: <http://tubic.tju.edu.cn/GC-Profile/>.  
 b) Posibles ORFs.  
 c) Peso molecular y punto isoelectrico calculados en <http://www.scripps.edu/~cdputnam/protcalc.html>.  
 d) Características de las secuencias aminoacídicas obtenidas a partir del NCBI <http://blast.ncbi.nlm.nih.gov/>.

Plásmido (tamaño)	G+C <sup>a</sup> (%)	CDS hipotética <sup>b</sup>	Tamaño proteína (aa) <sup>c</sup>	Pi c <sup>b</sup>	E-value <sup>dc</sup>	Aa identidad/similitud % <sup>d</sup>	Tipo de proteína	Proteína homóloga (Nº de acceso) <sup>d</sup>	Posible microorganismo de procedencia <sup>d</sup>	Características principales del microorganismo de procedencia	
p6.2 (7156 bp)	59.8	Orf 1 (>1-1078 reverse)	>359	--	1e-110	58/73	glycoside hidrolasa, family 43	Proteína hipotética conservada ZP_03013524	<i>Bacteroides intestinalis</i> DSM 17393	Aislado de heces humanas. Grupo <i>Bacteroidetes</i> .	
		Orf 2 (1099-3138 reverse)	679		0.0	63/78	precursor Alpha-galactosidasa	Proteína Hipotética conservada ZP_03013525	<i>Bacteroides intestinalis</i> DSM 17393	Aislado de heces humanas. Grupo <i>Bacteroidetes</i> .	
		Orf 3 (3140-4369 reverse)	409	8.93	2e-117	58/70	<b>Esterasa putativa</b>	C-terminal de una putative esterasa(350-689 aa) YP_823378	<i>Solibacter usitatus</i> Ellin6076	Acidobacteria gp3.	
		Orf 4 (>4427-7156 reverse) Proteína multifuncional	>909	8.28							
			648-906		1e-40	38/53	<b>Putative esterase</b>	Proteína hipotética conservada EDV06009	<i>Bacteroides intestinalis</i> DSM 17393	Aislado de heces humanas. Grupo <i>Bacteroidetes</i> .	
			346-605		4e-107	69/81	<b>Acetyl xylan esterase</b>	Proteína hipotética conservada ZP_03016604	<i>Bacteroides intestinalis</i> DSM 17393	Aislado de heces humanas Grupo <i>Bacteroidetes</i>	
			6-335		3e-77	47/59	Endo-1,4-beta-xylanasa	Endo-1,4-beta-xylanasa YP_662224	<i>Pseudoalteromonas atlantica</i> T6c	Bacteria gram negativa, formadora de biofilms quimiorganotrofa que se ha encontrado en océanos.	
P8.25 (7117 bp)	59.02	Orf 1 = N terminal Orf3 >1-400 (reverse)	>133	--	1e-36	58/66	--	C-terminal de una putative esterasa (350-689 aa) YP_823378	<i>Solibacter usitatus</i> Ellin6076	Acidobacteria gp3.	
		Orf 2 = Orf4 C terminal 427-1269 (reverse)	280 <sub>6</sub> ?	--	2e-40		<b>Esterasa putativa</b>	Proteína hipotética conservada EDV06009	<i>Bacteroides intestinalis</i> DSM 17393	Aislado de heces humanas. Grupo <i>Bacteroidetes</i> .	
		Orf 3 = Orf4 N terminal 1270-3357	695 415-672		1e-106	69/81	Acetyl xylan esterasa	Proteína hipotética conservada ZP_03016604	<i>Bacteroides intestinalis</i> DSM 17393	Aislado de heces humanas. Grupo <i>Bacteroidetes</i> .	

	(reverse) Proteína multifunción	8-406	--	4e-81	43/57	Endo-1,4-beta-xylanasa	glycosido hidrolasa familia 10	<i>Pseudoalteromonas atlantica</i> T6c	Bacteria gram negativa, formadora de biofilms quimiorganotrofa que se ha encontrado en océanos.
	Orf 4 3470-5305 (reverse)	611		2e-88	32/50	Alpha-glucosidasa	Proteína hipotética conservada Acid 3433 YP_824691	<i>Solibacter usitatus</i> Ellin6076	Acidobacteria.
	Orf 5 5343-6308 (reverse)	321		5e-79	50/71	Transportador ABC	Transportador ABC NP_407350	<i>Yersinia pestis</i> CO92	Gamma proteobacteria causante de plagas.
	Orf 6 >6332-final (reverse)	>261		1e-77	59/73	Translocador	Translocador de membrana interna ZP_02966503	<i>bacterium Ellin514</i>	Verrucomicrobia.

**Tabla AII.4.2.** Cebadores usados en este capítulo.

<b>Nombre</b>	<b>Secuencia</b>	<b>Aplicación</b>	<b>Ref.</b>
6.2out1	5´ - CCCATCCTCGCCTTTAGTCAG - 3´	Para diferenciar genomovar 6.2/8.25 (cebador externo)	En este trabajo
6.2out2	5´ - CTTTCAGGAGAATGCCAGAGCC - 3´	Para diferenciar genomovar 6.2/8.25 (cebador externo)	En este trabajo
6.2in1	5´ - CGTCTGGGCCATGCACATC - 3´	Para diferenciar genomovar 6.2/8.25 (cebador interno)	En este trabajo
6.2in2	- 5´ CATCTATTTTCGCGGCAGGTGA - 3´	Para diferenciar genomovar 6.2/8.25 (cebador interno)	En este trabajo

## **CONCLUSIONES**



1. Los análisis de similitud de los perfiles genéticos de  $\alpha$ -proteobacterias de muestras de suelo del Altiplano Central de México revelan que el determinante principal en la distribución de esta población microbiana es el tipo de suelo. Influyendo este factor más que la estacionalidad, y que el tipo de planta (cactácea en este caso).
2. Estudios taxonómicos basados en la subunidad pequeña del ribosoma (16SADNr) indican que la población procariota mayoritaria en la rizosfera de *M.carnea* (80% de la población) se corresponde con cuatro phyla: *Acidobacterias* (20-30%), *Actinobacterias* (15-20%), *Proteobacterias* (12%) y *Bacteriodetes* (8%). En el caso de las Arqueas la población está dominada por el phylum *Chrenarcheota*. Existen diferencias estadísticamente significativas (con un nivel de significación por debajo del 0.01) en los phyla acidobacterias, cianobacterias, gemmatimonadetes, plantomycetes y firmicutes entre época seca y lluviosa. Las acidobacterias y gemmatimonadetes aumentan en número en la época lluviosa, mientras que en el caso de los plantomycetes, firmicutes y cianobacterias son más numerosos en época seca
3. El análisis mediante secuenciación de alto rendimiento -454 FLX Titanium- de las muestras rizosféricas de *M.carnea* nos permitió observar fluctuaciones estacionales en los grupos minoritarios, como en el phylum *Cyanobacterias*, en el que unidades taxonómicas relacionadas con plastidios de plantas (familia Chloroplast) aumentan 20 veces en época lluviosa; mientras que otras relacionadas con la familia I aumentan casi 17 veces en época seca. Por primer vez se observa de este modo, una explosión clonal de un determinado grupo de microorganismos en esta rizosfera.
4. El grupo dominante de *Acidobacterias* en la población rizosférica de *M.carnea*, es el taxón GP4, tanto en época seca como en época lluviosa. El segundo grupo mayoritario es el GP6, y en suelo suelto además aparece el grupo GP3. La micro-diversidad dentro del phylum *Acidobacterias* es mayor en suelo suelto que en suelo rizosférico.
5. La generación de genotecas en fagémidos de expresión es un método con un alto rendimiento para la búsqueda de nuevas actividades y

enzimas de interés biotecnológico, usando como material de partida ADN de muestras de suelo.

- 6.** Se ha descrito el resistoma de las muestras de estudio, encontrando 11 nuevos genes que codifican para proteínas que confieren resistencia a diferentes antibióticos, implicando distintos mecanismos de acción como inactivación (genes de resistencia frente a Ampicilina y Gentamicina), bombas de extrusión (genes de resistencia a Cloranfenicol) y enzimas resistentes al antibiótico (genes de resistencia a trimetoprima).
- 7.** Se ha analizado, purificado y caracterizado bioquímicamente una nueva proteína con actividad dihidrofolato reductasa con una elevada resistencia a trimetoprima: cuya secuencia, y características bioquímicas la señalan como representante de una nueva familia dentro de este grupo de enzimas.
- 8.** Se han encontrado 16 nuevas enzimas esterases de origen microbiano no descritas con anterioridad, que 12, constituyen nuevos miembros de familias de esterases conocidas (familia II: 4 enzimas; familia IV: 3 proteínas ; familia VIII:2 enzimas, familia VII: 1 enzima) y 4 podrían suponer un nuevo tipo de genes esterasa.
- 9.** Se ha purificado y caracterizado bioquímicamente una enzima esterasa, pE2, con una gran estabilidad a solventes orgánicos. Su actividad se mantiene intacta o se potencia tras la incubación durante 6 horas con los solventes al 90%: acetona, acetonitrilo, dimetilsulfóxido, etanol e isopropanol. Mantiene su actividad tras la incubación en etanol al 75% durante 24 horas. El extremo N-terminal de esta proteína (del aminoácido 1 al 166) resulta indispensable para su actividad, aunque en él no se encuentra ningún residuo de la tríada catalítica, y su paso a otra enzima sensible a solventes orgánicos no confiere estabilidad a este tipo de sustancias, aunque mejora su estabilidad en detergentes como el Triton-X-100.
- 10.** Se ha purificado y caracterizado bioquímicamente la lipasa verdadera pE12, que mantiene su actividad en un amplio rango de

temperaturas (a 4°C y a 42°C mantiene su actividad aproximadamente al 50%, siendo su óptimo 37°C). Esta lipasa además, mantiene su actividad al 40% tras incubarla a 50°C durante 3 horas, se activa mediante el cation K<sup>+</sup> y se inactiva con el catión Ca<sup>+2</sup>. Posee también actividad sobre diferentes compuestos de interés farmacéutico como el nanoproxeno o el compuesto químico naphtoato. Todas estas, cualidades de interés para su posible uso en la industria.

- 11.** La presencia de dos fragmentos (6.2 y 8.25) conteniendo genes codificantes para esterases relacionadas (nivel de homología 98%) en distintas genotecas de la rizosfera de *M. carnea* –época seca y lluviosa– nos ha permitido el desarrollo de herramientas moleculares (mediante PCR) para la distinción de subtipos mayoritarios dentro de las muestras de estudio. Encontrando que el genomovar 6.2 es un biomarcador de rizosfera y de las zonas Valerio Trujano 1 y 2; mientras que el genomovar 8.25 es un bioindicador de los suelos de la zona de Tomellín.





# **BIBLIOGRAFÍA**



## BIBLIOGRAFÍA

**Aamand JL, Hobson AH, Buckley CM, Joergensen ST, Diderichsen B, McConnell DJ. 1994.** Chaperone-mediated activation in vivo of a *Pseudomonas cepacia* lipase. *Mol Gen Genet* 245: 556-564

**Abrosimova, L.M. y Revut, I.B. 1964.** Biological activity and the composition of soil air in the low layer. *Soviet Soil Science*, 7:682-691

**Alcalde M.2003.** Evolución molecular dirigida. *Investigación y ciencia* 11:69-77

**Allen, H.K., Moe, L.A., Rodbumrer, J., Gaarder, A. and J. Handelsman. 2009.** Functional metagenomics reveals diverse  $\beta$ -lactamases in a remote Alaskan soil. *ISME J.* 3: 243-251.

**Amann RI, LudwigW, Scheleifer KH. 1995.** Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev*, 59:143-169

**Amann, R.I., Ludwig, W., & Schleifer, K.H. 1995** Phylogenetic Identification and *in-Situ* Detection of Individual Microbial-Cells without Cultivation. *Microbiological Reviews* 59: 143-169.

**Anderson, T.A., Guthrie, E.A., & Walton, B.T. 1993** Bioremediation in the Rhizosphere. *Environmental Science & Technology* 27: 2630-2636.

**Andeson, J. M. 1975.** The enigma of soil animal species diversity. En: *Progress In Soil Zoology*, (J. Vanek, Ed.) Proc. 5° Int. Coll. Soil Zool., 51-58 pp. *Appl Environ Microbiol* 70: 4748-4755.

**Arpigny JI, Jaeger KE 1999.** Bacterial lipolytic enzymes: classification and properties. *Biochem J* 343: 177-183  
Assignment of hitherto unidentified 165 rDNA species to a main line of descent within the domain Bacteria. *Microbiology* 141: 3247-3250  
bacterial communities. *PNAS*.vol. 103 no. 3 626-631

**Bajpai, P.; Bajpai, P.K.1998.** Deinking with enzymes: a review. *Tappi J.*,vol 81, 12,111-117

**Baker G.C., Smith J.J., Cowan D.A 2003.** Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* 55 541- 555

**Bardgett, R.D. 2002.** Causes and consequences of biological diversity in soil. *Zoology*, 105: 367-374

**Barns SM, Cain EC, Sommerville L, Kuske CR. 2007.** Acidobacteria phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum. *Appl Environ Microbiol* 73: 3113-3116.

**Barns SM, Takala SL; Kuske CR . 1999 .** Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. *Appl Environ Microbiol* 65: 1731-1737.

**Ben Ali Y, Chahinian H, Petry S, Muller G, Carriere F, Verger R, Abousalham A 2004.** Might the kinetic behaviour of hormone-sensitive lipase reflect the absence of the lid domain? *Biochemistry* 43: 9298-9306

**Berg G, Krechel A, Ditz M, Sikora A., Ulrich A., Hallmann J. 2005.** Endophytic and ectophytic potato-associated bacterial communities differ in structure and antagonistic function against plant pathogenic fungi, *FEMS Microbiology Ecology*. 5:1215–229

**Berg G, Roskot N, Steidle A, Eberl L, Zock A, and Smalla.2002.** K.Plant-Dependent Genotypic and Phenotypic Diversity of Antagonistic Rhizobacteria Isolated from Different *Verticillium* Host Plants *Appl and Environm Microbiol*. P3328-3338, V1 68, No7

**Berg G, Smalla K. 2009** Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere

**Berlemont R, Delsaute M, Pipers D, D'Amico S, Feller G, Galleni M, Power P. 2009** Insights into bacterial cellulose biosynthesis by functional metagenomics on Antarctic soil samples. *ISME J. Sep*;3(9):1070-81.

**Boeufgras J-M, Alabouvette C. 1995.** Effect of two plant species Flax (*Linum usitatissimum* L.) and tomato (*Lycopersicon esculentum* Mill.), on the diversity of soilborne populations of fluorescent pseudomonads. *Appl Environ Microbiol* 61:1004–1012

**Boon N, De Windt W, Verstraete W, Top EM. 2002.** Evaluation of nested PCR-DGGE (denaturing gradient gel electrophoresis) with group-specific 16S rRNA primers for the analysis of bacterial communities from different wastewater treatment plants. *FEMS Microbiol Ecol* 39: 101–112.

**Bornscheuer UT. 2002** Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiol Rev. Mar*;26(1):73-81.

**Brockerhoff J, Jensen RG 1974.** *Lipolytic Enzymes*. New York: Academic

**Bryant DA, Garcia Costas AM, Maresca JA et al. 2007.** *Candidatus Chloracidobacterium thermophilum*: an aerobic phototrophic acidobacterium. *Science* 317: 523–526.

**Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL. 2008.** Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proc Natl Acad Sci USA* 105: 11505–11511

**Brzozowski AM, Derewenda U, Drewenda ZS, Dodson GG, Lawson DM, Turkenburg JP, Bjorkling F, Huge-Jensen B, Patkar SA, Thim L 1991.** A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. *Nature* 351: 491-194

**Buyanovsky, G.A. y Wagner, G.H. 1983.** Annual cycles of carbon dioxide level in soil air. *Soil Science Society of America Journal*, 47: 1139-1145.

**Campbell, R., & Greaves, M.P. 1990.** *Anatomy and community structure of the rhizosphere*

- Cantón R, Coque TM 2006.** The CTX-M beta-lactamase pandemic. *Curr Opin Microbiol* ;9: 446-475
- Canton, R. 2009.** Antibiotic resistance genes from environment: a perspective through newly identified antibiotic resistance mechanisms in the clinical setting. *Clinic.Microbiol. Infect.* 15 (Suppl. 1): 20–
- Casimir C. Akoh a,b,1, Guan-Chiun Lee b, Yen-Chywan Liaw c, Tai-Huang Huang d, Jei-Fu Shaw. 2004** GDSL family of serine esterases/lipases. *Progress in Lipid Research* 43) 534–552.
- Castro H.F; Classen A.T; Austin E.E; Norby R.J; Schadt C.W. 2009.** Soil Microbial Community Responses to Multiple Experimental Climate Change Drivers. *Appl. Environ. Microbiol.* Doi: 10.1128/AEM-02874-09.
- Chow ML, Radomski CC, McDermott JM, Davies J & Axelrood PE 2002** Molecular characterization of bacterial diversity in Lodgepole pine (*Pinus contorta*) rhizosphere soils from British Columbia forest soils differing in disturbance and geographic source. *FEMS Microbiol Ecol* 42: 347–357
- Chung, E.J., Lim, H.K., Kim, J.C., Choi, G.J., Park, E.J., Lee, M.H., Chung, Y.R. and S.W. Lee 2008.** Forest soil metagenome gene cluster involved in antifungal activity expression in *Escherichia coli*. *Appl Environ Microbiol.* 74: 723-30.
- Cieśliński H, Białkowska A, Tkaczuk K, Długołęcka A, Kur J, Turkiewicz M.2009.** Identification and molecular modeling of a novel lipase from an Antarctic soil metagenomic library. *Pol J Microbiol.* 2009;58(3):199-204
- Cole J.R, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulman-Syed-Mohideen, A.S; McGarrel,D.M; Marsh T, Garrity GM et al. 2009.** The Ribosomal Data Base Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.*37,D141-D145; doi: 10.193/nar/gkn1879
- Coloqhoun, J.A., 1997.** Discovery of deep-sea actinomycetes. PhD Dissertation. Research School of Biosciences. University of Kent, Canterbury.
- Cottrell M.T; Kirchman D.L. 2000.** Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence in situ hybridization. *Appl Environ Microbiol* 66: 5116-5122.
- Couto GH, Glogauer A, Faoro H, Chubatsu LS, Souza EM, Pedrosa FO. 2010.** Isolation of a novel lipase from a metagenomic library derived from mangrove sediment from the south Brazilian coast. *Genet Mol Res.* 9(1):514-523
- Cruz H, Perez C, Wellington E, Castro C, Servin-Gonzalez L 1994.** Sequence of the *Streptomyces Albus G* lipase-encoding gene reveals the presence of a prokaryotic lipase family. *Gene:* 144: 141-142
- Curl, E., & Truelove, B. (1986)** *The Rhizosphere.* Springer-Verlag. Berlin, Germany: 288.

- Curtis, T.P., Sloan, W.T. y Scannell, J.W. 2002.** Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America*, 99:10494–99.
- Da Costa VM, McGrann KM, Hughes DW, Wright GD 2006.** Sampling the antibiotic resistome. *Science*; 311:374-377
- Da Costa, V.M., McGrann, K.M., Hughes, D.W. and G.D. Wright. 2006.** Sampling the antibiotic resistome. *Science*. 311: 374-377.
- Da Rocha, U.N; Overbeek L.V, van Elsas J.D. 2009.** Exploration of hitherto-uncultured bacteria from the rhizosphere *FEMS Microbiol Ecol* 69 (2009) 313–328
- Daniel, R.2004.**The soil metagenome – A rich resource for the discovery of novel natural products. *Curr. Opin. Biotechnol.*, 15, 199–204.
- DasSarma, S., Fleischmann, E.F., 1995.** Archaea: A Laboratory Manual—Halophiles. Cold Spring Harbour Laboratory Press, New York, pp. 269– 272
- Davidson CJ, Surette MG. 2008.** Individuality in bacteria *Annu Rev Genet*. 42:253-68. Review.
- Davis K.E.R.; Joseph S.J, Janssen H.P. 2005.** Effects of growth medium, inoculum size and incubation time on the culturability and isolation of soil bacteria. *Appl. Environ. Microbiol*. 71: 826-834
- De la Cruz F, Garcia Lobo jm, Davies J 2002.** Antibiotic resistance: how bacterial populations respond to a simple evolutionary force. In: Lewis KA, Salyers AA, Taber HW, Wax RG, eds *Bacterial resistance to antimicrobial*. New York: Marcel Dekker Inc; 19-36
- de Pascale D, Cusano AM, Autore F, Parrilli E, di Prisco G, Marino G, Tutino ML 2008.** The cold-active Lip1 lipase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125 is a member of a new bacterial lipolytic enzyme family. *Extremophiles* 12: 311-323
- de Simone G, Galdiero S, Manco G, Lang D, Rossi M, Pedone C 2008.** The cold –active Lip1 lipase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125 is a member of a new bacterial lipolytic enzyme family. *Extremophiles* 12: 311-323
- de Simone G, Menchise V, Manco G, Mandrich L, Sorrentino N, Lang D, Rossi M, Pedone C 2001.** The crystal structure of a hyper-thermophilic carboxylesterase from the archeon *Archaeoglobus fulgidus*. *J Mol Biol* 314: 507-518
- DeLong, E. F., y Pace, N. R. 2001.** Environmental diversity of bacteria and archaea. *Systematic Biology*, 50:470–478.
- Diaz P, Prim N and Pastor J. 1999.** Direct Fluorescence-Based Lipase Activity Assay. *BioTechniques* 27: 697-700.
- Doran, J.W., Sarrantonio, M. y Liebig, M.A. 1996.** Soil health and sustainability. *Advances in Agronomy*, 56: 2–54.

**Dybvig, K. 1993.** DNA rearrangements and phenotypic switching in prokaryotes. *Molecular Microbiology*, 10: 465-471.

**Dykhuizen, D.E. 1998.** Santa Rosalia revisited: Why are there so many species of bacteria? *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology*, 73:25-33.

**Eichorst SA, Breznak JA & Schmidt TM 2007.** Isolation and characterization of soil bacteria that define *Terriglobus* gen. nov., in the phylum Acidobacteria. *Appl Environ Microb* 73: 2708-2717.

**Eilers H.; Pernthaler J; Glöckner F.O, Amann R. 2000.** Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* 66: 3044-3051

**el Zahar Haichar, F., Marol C, Berge O, Rangel-Castro J.I, Prosser J.I, Balesdente J, Heulin T, Achouak, W 2008.** Plant host habitat and root exudates shape soil bacteria community structure. *ISME j.* 2, 1221-1230.

**Elend C, Schmeisser C, Leggewie C, Babiak P, Carballeira JD, Steele HL, Reymond JL, Jaeger KE, Streit WR 2006.** Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl Environ Microbio* 72; 3637-3645

**Elend C, Schmeisser C, Leggewie C, Babiak P, Carballeira JD, Steele HL, Reymond JL, Jaeger KE, Streit WR.2006** Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl Environ Microbiol.*;72(5):3637-45

**Elwood, H.J., Olsen, G.J., Sogin, M.L., 1985.** The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. *Mol. Biol. Evol.* 2, 399.

**Fajardo A, Martínez JL. 2008.** The role of antibiotics as signalling molecules. *Curr. Opin. Microbiol.* 11, 161

**Farely, V., Rainey, F.A., Stackebrandt, E., 1995.** Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Appl. Environ. Microbiol.* 61, 2798- 2801.

**Feller G, Thiry M, Gerday C 1990.** Sequence of a lipase gene from the antarctic psychrotroph *Moraxella* TA 144. *Nucleic Acids Res* 18:6431

**Felske A, de Vos WM, Akkermans ADL. 2000.** Spatial distribution of 16S rRNA levels from uncultured Acidobacteria in soil. *Lett Appl Microbiol* 31: 118-122

**Ferrer M, Golyshina O, Chernikova T, Khachane A.N, Mantins dos Santos V, Yakimov M, Timmis K, Golyshin P. 2005.** Microbial Enzymes Mined from the Urania Deep-Sea Hypersaline Anoxic Basin. *Chemistry and Biology*, Vol 12, 895-904



**Ferrer M, Golyshina OV, Chernikova TN, Khachane AN, Reyes-Duarte D, Santos VA, Strompl C, Elborough K, Jarvis G, Neef A, Yakimov MM, Timmis KN, Golyshin PN. 2005.** Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environ Microbiol*;7(12):1996-2010.

**Ferrer M. 2005.** Metagenoma: acceso a los recursos potencialmente ilimitados de microorganismos no cultivables. *Actualidad de la SEM*. 38 11-17

**Fierer N, Bradford M, Jackson R. 2007 aç.** Toward an ecological classification of soil bacteria. *Ecology* 88: 1354–1364.

**Fierer N, Jackson JA, Vilgalys R, Jackson RB. 2005.** Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl Environ Microbiol* 71: 4117–4120

**Fierer N, Jackson R.B. 2006.** The diversity and biogeography of soil bacterial communities. *PNAS*. Vol 103.nº3. 626-531

**Fojan P, Jonson PH, Petersen MT, Petersen SB 2000.** What distinguishes an esterase from a lipase: a novel structural approach. *Biochimie*. Nov;82(11):1033-41

**Frenken LG, de Groot A, Tommansen J, Verrips CT 1993.** Role of the lipB gene product in the folding of the secreted lipase of *Pseudomonas glumae*. *Mol Microbiol* 9:591-599

**Garbeva P, van Veen JA & van Elsas JD. 2004.** Microbial Diversity in soil: selection of microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* 42: 243–270.

**Garbeva, P., van Veen, J.A. y van Elsas, J.D. 2004.** Microbial diversity in soil: selection of microbial populations by plant and soil type and implications for disease suppressiveness. *Annual Review of Phytopathology*, 42:243–70.

**Garrity GM, Bell JA & Lilburn TG .2005.** Taxonomic outline of the prokaryotes. *Bergey's Manual of Systematic Bacteriology*, Second Edition. Release 5.0 May 2004 (Garrity GM, editor-in-chief), pp. 1–399. Springer-Verlag, NY.

**Gause, G.F. 1934.** The struggle for existence. Baltimore: Williams & Wilkins geographic regions. *Appl Environ Microbiol* 63: 3614–3621.

**George. A. Kowalchuk G.A, Buma D.S., de Boer W, Klinkhamer P.G.L & Germida, J.J. y Siciliano, S.D. 2001.** Taxonomic diversity of bacteria associated with the roots modern, recent and ancient wheat cultivars. *Biology and Fertility of Soils*, 33: 410-415.

**Gilbert EJ, Cornish A, Jones CW .1991.** Purification and properties of extracellular lipase from *Pseudomonas aeruginosa* EF2. *J.Gen. Microbiol* 137: 2233-2229

**Giller, P.S., 1996.** The diversity of soil communities, the poor man's tropical rainforest. *Biodiversity and Conservation*, 5: 135-168.

**Girvan, M.S., Bullimore, J., Pretty, J.N., Osborn, A.M. y Ball, A.S. 2003.** Soil type is the primary determinant of the composition of the total and active bacterial communities in arable soils. *Applied and Environmental Microbiology*, 69 (3), 1800-1809

**Godtfredsen SE. 1990.** Microbial lipases. In: Fogarty WM, Kelly ET (eds). *Microbial enzymes and biotechnology*. Elsevier, Amsterdam, 255-274

**Gomes N.C.M, Heuer H, Schönfeld J., Costa R, Mendonça-Hagler L, Smalla K. 2001.** Bacterial diversity of the rhizosphere of maize (*Zea mays*) grown in tropical soil studied by temperature gradient gel electrophoresis. *Plant Soil* 232: 167-180

**Gomes, N.C.M., Fagbola O., Costa R, Rumjanek, N.G, Buchner A., Mendona-Hagler, L. and Smalla, K. 2003.** Dynamics of fungal communities in bulk and maize rhizosphere soil in the tropics. *Appl. Environ. Microbiol* 69, 3758-3766

**Good, I.J. 1953.** The population frequencies of species and the estimation of population parameters. *Biometrika* 40: 237-264

**Govardhan CP, Pratt RF 1987.** Kinetics and mechanism of the serine beta-lactamase catalyzed hydrolysis of depsipeptides. *Biochemistry* 26: 3385-3395

**Graff A & Conrad R 2005** Impact of flooding on soil bacterial communities associated with poplar (*Populus* sp.) trees. *FEMS Microbiol Ecol* 53: 401-415.

**Graner, G., Persson, P., Meijer, J., and Alstrøm, S. 2003.** A study on microbial diversity in different cultivars of *Brassica napus* in relation to its wilt pathogen, *Verticillium longisporum*. *FEMS Microbiol Lett* 29: 269-276.

**Grayston, S.J., Vaughan, D. y Jones, D. 1996.** Rhizosphere carbon flow in trees, in comparison with annual plants: the importance of root exudation and its impact on microbial activity and nutrient availability. *Applied Soil Ecology*, 5: 29-56.

**Grayston, S.J., Wang, S., Campbell, C.D. y Edwards, A.C., 1998.** Selective influence of plant species on microbial diversity in the rhizosphere. *Soil Biology & Biochemistry*, 30: 369- 378.

**Gremion F, Chatzinotas A & Harms H 2003** Comparative 16S rDNA and 16S rRNA sequence analysis indicates that Actinobacteria might be a dominant part of the metabolically active bacteria in heavy metal-contaminated bulk and

**Gupta R, Gupta N, Rathi P. 2004.** Bacterial lipases: an overview of production, purification and biochemical properties. *Appl Microbiol Biotechnol*. 2004 Jun;64(6):763-81. Review

**Gupta R, Gupta N, Rathi P. 2004.** Bacterial lipases: an overview of production, purification and biochemical properties *Appl Microbiol Biotechnol* 64: 763-781

- Handelsman J. 2004.** Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and Molecular Biology Reviews*, p. 669-685, Vol. 68, No. 4
- Hansen, M.C., Tolker-Neilson, T., Givskov, M., Molin, S., 1998.** Biased 16S rDNA PCR amplification caused by interference from DNA flanking template region. *FEMS Microbiol. Ecol.* 26,141– 149.
- Hardeman F, Sjoling S. 2007** Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. *FEMS Microbiol Ecol* 59524-534
- Hardin, G. 1960.** The competitive exclusion principle. *Science*, 131:1292–97.
- Hemilae H, Koivula TT, Palva I 1994.** Hormone sensitive lipase is closely related to several bacterial proteins, and distantly related to acetylcholinesterase and lipoprotein lipase: identification of a superfamily of esterases and lipases. *Biochim Biophys Acta* 1210:249-253
- Heuer H, Abdo Z, Smalla K. 2008.** Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. *FEMS Microbiol Ecol.*65(3):361-71
- Heuer, H., Engelen, B. y Smalla, K. 1997.** A Modern Soil Microbiology (van Elsas, J. D., Wellington, F. M. H., y Trevors, J. T., eds.), pp. 353–373, Dekker, New York. 276
- Hiltner, L. 1904.** Uber neuere erfahrungen und probleme auf dem gebeit der bodenbakteriologie und unter besonderer berucksichtigung der grundungung und brache. *Arb. Dsch. Landwirt. Ges*, 98: 59-78.
- Hiraishi A, Kishimoto N, Kosako Y, Wakao N, Tano T. 1995.** Phylogenetic position of the menaquinonecontaining acidophilic chemo-organotroph *Acidobacterium capsulatum*. *FEMS Microbiol Lett* 132: 91–94.
- Hobson AH, Buckley CM, Aamand JL, Jorgensen ST, Didereichsen B, McConnel DJ 1993.** Activation of a bacterial lipase by its chaperone. *Proc Natl Acad Sci USA* 90: 5682-5686
- Hong KS, Lim HK, Chung EJ, Park EJ, Lee MH, Kim JC, Choi GJ, Cho KY, Lee SW. 2007.** Selection and characterization of forest soil metagenome genes encoding lipolytic enzymes. *J Microbiol Biotechnol.* Oct;17(10):1655-60
- Horton TR and Burns TD 2001.** The molecular revolution in ectomycorrhizal ecology: peeking into the black-box. *Molecular Ecology*, 10, 1855-1871.
- Hugenholtz P, Goebel BM; Pace NR 1998.** Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* 180: 4765-4774.
- Hutchinson, G.E. 1961.** The paradox of the plankton. *American Naturalist*, 95:137–45.

**Hyman BC, Beck JL, Weiss KC 1988.** Sequence amplification and gene rearrangement in parasitic nematode mitochondrial DNA. Genetics. Nov;120(3):707-12.

**Iwagami SG, Yang K, Davies J. 2000** Characterization of the protocatechuic acid catabolic gene cluster from *Streptomyces* sp. strain 2065. *Appl Environ Microbiol* Apr;66(4):1499-508

**J. T. Trevors Antonie van Leeuwenhoek.2010.** One gram of soil: a microbial biochemical gene library. 97:99–106. DOI 10.1007/s10482-009-9397-5

**Jaatinen K, Tuittila ES, Laine J, Yrjala K, Fritze H 2005.** Methane-oxidizing bacteria in a Finnish raised mire complex: effects of site fertility and drainage. *Microb Ecol.* Oct;50(3):429-39.

**Jaeger KE, Dijkstra BW, Reetz MT 1999.** Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Annu Rev Microbiol* 53: 315-351

**Jaeger KE, Eggert T 2002.** Lipases for biotechnology. *Curr Opin Biotechnol.* Aug;13(4):390-7. Review

**Jaeger KE, Ransac S, Koch HB, Ferrato F, Dijkstra BW 1993.** Topological Characterization and Modeling of the 3D Structure of Lipase form *Pseudomonas aeruginosa*. *FEBS Lett* 332:143-149

**Jaeger, K.E and Reetz, M.T. 1998** Microbial lipases form versatile tools for biotechnology, *Trends Biochem. Biotech* Vol.16, 396-403

**Janssen P.H. 2006.** Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16s RNA genes. *Appl Environ Microbio.* Minireviews. P 1719-1728

**Janssen PH, Yates PS, Grinton BE, Taylor PM & Sait M 2002** Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl Environ Microb* 68: 2391–2396.

**Janssen PH. 2006.** Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol* 72: 1719–1728.

**Jean Louis Arpigny and Karl-Erich Jaeger. 1999.** Bacterial lipolytic enzymes: classification and properties. *Biochem. J.* 343, 177-183

**Jeong Ho Jeon & Jun-Tae Kim & Sung Gyun Kang & Jung-Hyun Lee & Sang-Jin Kim. 2009.** Characterization and its Potential Application of Two Esterases Derived from the Arctic Sediment Metagenome. *Mar Biotechnol* 11:307–316

**Jiang C, Ma G, Li S, Hu T, Che Z, Shen P, Yan B, Wu B.2009** Characterization of a novel beta-glucosidase-like activity from a soil metagenome. *J Microbiol.* 2009 Oct;47(5):542-8

- Johnsen, K., Jacobsen, C.S., Torsvik, V. y Sorensen S. 2001.** Pesticide effects on bacterial diversity in agricultural soils. A review. *Biology and Fertility of Soils*, 33 (6): 443-453.
- Jones M, Page MI 1991.** An esterase with beta-lactamase activity. *J Chem Soc Chem Commun* 5: 316-317
- Jones RT, Michael SR, Lauber CL, Hamady M, Knight R, Fierer N. 2009.** A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *The ISME Journal* 3,442-453
- Jones T.R; Robeson M.S; Lauber CL; Hamady M; Knight R; Fierer N. 2009.** A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *The ISME Journal*. 3, 442-453
- Joseph SJ, Hugenholtz P, Sangwan P, Osborne CA, Janssen PH. 2003.** Laboratory cultivation of widespread and previously uncultured soil bacteria. *Appl Environ Microbiol* 69: 7210-7215.
- Jousset A., Lara E., Wall L.G. & Valverde C. 2006.** Secondary metabolites help biocontrol strain *Pseudomonas fluorescens* CHA0 to escape predation. *Appl. Environ. Microbiol.*, 72, 7083-7090.
- Kaiser O, Puhler A & Selbitschka W 2001** Phylogenetic analysis of microbial diversity in the rhizoplane of oilseed rape (*Brassica napus* cv. Westar) employing cultivation-dependent and cultivation-independent approaches. *Microb Ecol* 42: 136-149.
- Kakugawa S, Fushinobu S, Wakagi T, Shoun H 2007.** Characterization of a thermostable carbosylesterase from the hyperthermophilic bacterium *thermotoga maritima*. *Appl Microbial Biotechnol* 74:585-591
- Katoh, K. and H. Toh. 2008.** Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 9: 286-298.
- Kazimierczak, K.A., Scott, K.P., Kelly, D. and R.I. Aminov. 2009.** Tetracycline resistome of the organic pig gut. *Appl Environ Microbiol*. 75: 1717-1722.
- Kennedy, A.C., y Smith, K. L. 1995.** Soil microbial diversity and the sustainability of agricultural soils. *Plant and Soil*, 170:75-86.
- Kent, A.D. y Triplett, E.W. 2002.** Microbial communities and their interactions in soil and rhizosphere ecosystems. *Annual Review of Microbiology*, 56:211-36.
- Kent, A.D., & Triplett, E.W. 2002** Microbial communities and their interactions in soil and rhizosphere ecosystems. *Annual Review of Microbiology* 56: 211-236
- Kielak A, Pijl AS, Van Veen JA & Kowalchuk GA 2008** Differences in vegetation composition and plant species identity lead to only minor changes in soil-borne microbial communities in a former arable field. *FEMS Microbiol Ecol* 63: 372-382.

**Kiesser T, Bibb MJ, Buttner MJ, Chater KF, 2000** Hopwood DA: Practical Streptomyces Genetics. 1 edn. Norwich: The Jones Innes Foundation;

**Kim HK, Park SY, Lee JK, Oh Tk 1998.** Gene cloning and characterization of thermostable lipase form *Bacillus stearothermophilus* L1. *Biosci Biotechnol Biochem* 62: 66-71

**Kim SJ, Lee CM, Han BR, Kim MY, Yeo YS, Yoon SH, Koo BS, Jun HK.2008.** Characterization of a gene encoding cellulase from uncultured soil bacteria. *FEMS Microbiol Lett.* 2008 May;282(1):44-51

**Kim YJ, Choi GS, Kim SB, Yoon GS, Kim YS, Ryu YW 2006.** Screening and characterization of a novel esterase form a metagenomic library. *Protein Exp Purif* 45: 315-323

**Kim YS, Lee HB, Choi KD, Park S, Yoo OJ 1994.** Cloning of *Pseudomonas fluorescens* carboxylesterase gene and characterization of its product expressed in *Escherichia coli*. *Biosci Biotechnol Biochem* 58: 11-116

**Kishimoto N, Kosako Y, Tano T. 1991.** *Acidobacterium capsulatum* Gen-Nov, Sp-Nov—an acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. *Curr Microbiol* 22:1-7.

**Klibanov, A.M. .2001.** Improving enzymes by using them in organic solvents. *Nature*. Vol409.11.

**Knietsch A, Waschowitz T, Bowien S, Henne A, Daniel R.2003.** Metagenomes of complex microbial consortia derived from different soils as sources for novel genes conferring formation of carbonyls from short-chain polyols on *Escherichia coli*. *J Mol Microbiol Biotechnol.*;5(1):46-56.

**Koch IH, Gich F, Dunfield PF & Overmann J 2008** *Edaphobacter modestus* gen. nov., sp. nov., and *Edaphobacter aggregans* sp. nov., acidobacteria isolated from alpine and forest soils. *Int J Syst Evol Micr* 58: 1114-1122.

**Kolganova, T.V., Kuznetsov, B.B., Tourova, T.P., 2002.** Designing and testing oligonucleotide primers for amplification and sequencing of Archaeal 16S rRNA genes. *Mikrobiologiya* 71,283- 286.

**Kumar, S., Dudley, J., Nei, M. and K. Tamura.** 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* **9**: 299-306.

**Kuske CR, Barns SM, Busch JD. 1997.** Diverse uncultivated bacterial groups from soils of the arid southwestern United States that are present in many geographic regions *Appl. Environ. Microbiol.*,3614-3621, Vol 63, No. 9

**Kuske CR, Ticknor LO, Miller ME, Dunbar JM, Davis JA, Barns SM & Belnap J 2002** Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. *Appl Environ Microb* 68: 1854-1863

**Kwang Yun Cho, Seon-Woo Lee, Hong, Kyung Sik, He Kyoung Lim, Eu Jin Chung, Eun Jin Park, Myung Hwan Lee, Jin-Cheol Kim, Gyung Ja Choi.** 2007. Selection and Characterization of Forest Soil Metagenome Genes Encoding Lipolytic Enzymes. *J. Microbiol. Biotechnol*; 17(10): 1655-1660

**Lachica, M., Recalde, L., Esteban, E.** 1965. Métodos analíticos utilizados en la Estación Experimental del Zaidín. *Anal. Edafol. Agrobiol.* 24, 589-610

**Lauber C.L, Hamady M, Knight R, Fierer N.** 2009. Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. *Applied and Environmental Microbiology.* 5111-5120.

**Lauber CL, Strickland MS, Bradford MA, Fierer N.** 2008. The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol Biochem* 40: 2407-2415

**Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N.** 2010. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *FEMS Microbiol Lett.*

**Lee O-W, Koh Y-S, Kim K-J, Kim B-C, Choi H-J, Kim D-S, Suhartono MT, Pyun Y-R.** 1999. Isolation and characterization of a thermophilic lipase from *Bacillus thermoleovorans* ID-1. *FEMS Microbiol Lett* 179:393-400

**Lee S.H; Ka J.O; Cho J.C.** 2008. Members of the phylum Acidobacteria are dominant and metabolically active in rhizosphere soil. *FEMS Microbiol Lett* 285. 263-269

**Lee SW, Won K, Lim HK, Kim JC, Choi GJ, Cho KY.** 2004. Screening for novel lipolytic enzymes from uncultured soil microorganisms. *Appl Microbiol Biotechnol.* Nov;65(6):720-6

**Lemanceau P, Corberand T, Gardan L, Latour X, Laguerre G, Lesuisse E, Schack K, Colson C** 1993. Purification and preliminary characterization of the extracellular lipase of *Bacillus subtilis* 168, an extremely basic pH-tolerant enzyme. *Eur J Biochem* 216: 155-160

**Lesuisse E, Schanck K, Colson C.** 1993. Purification and preliminary characterization of the extracellular lipase of *Bacillus subtilis* 168, an extremely basic pH-tolerant enzyme. *Eur J Biochem* 216: 155-160

**Levins, R.** 1968. *Evolution in Changing Environments.* Princeton, NJ: Princeton Univ. Press.

**Linares JF, Gustafsson I, Baquero F, Martínez JL** 2006. Antibiotic as intermicrobial signaling agents instead of weapons. *Proc Natl Acad Sci USA*; 103: 19484-19489.

**Lucas García, J.A., Barbas, C., Probanza, A., Barrientos, M.L. y Gutiérrez Mañero, F.J. 2001.** Low molecular weight organic acids (LOAs) and fatty acids in root exudates of two *Lupinus* cultivars at flowering and fruiting stages. *Phytochemical Analysis*, 12: 305-311.

**MacKay DB, Jennings MP, Godfrey EA, MacRae IC, Rogers PJ, Geacham IR 1992.** Molecular analysis of an esterase-encoding gene from a lipolytic psychrotrophic pseudomonad. *J Gen Microbiol* 138: 701-708

**MacLean, Daniel; Jones Jonatha D.G; Studholme Davi J.2009.** Application of “next-generation”sequencing technologies to microbial genetics. *Nature reviews*. Volum7

**Macrae AR, Hammond RC. 1985.** Present and future applications of lipases. *Biotech Genet Eng Rev* 3: 193-217

**Manco G; Adinolfi E; Pisani F.M; Ottolina G; Carrea G; Rossi M. 1998.** Overexpression and properties of a new thermophilic and thermostable esterase from *Bacillus acidocaldarius* with sequence similarity to hormone-sensitive lipase subfamily. *Biochem J*.332,203-212.

**Mannisto MK, Tirola M, Haggblom MM. 2007.** Bacterial communities in Arctic fields of Finnish Lapland are stable but highly pH-dependent. *FEMS Microbiol Ecol* 59: 452-465.

**Maqbool QA, Johri S, Verma L, Riyaz-ul-Hassan S, Verma V, Koul S, Taneja SC, Parshad R, Qazi GN 2002.** Purification and characterization of a novel enantioselective hydrolase from *Bacillus subtilis*. *Biotechnol Appl Biochem* 36: 227-234

**Martinez-Murcia, A.J., Acinas, S.G., Rodriguez-Valera, F., 1995.** Evaluation of prokaryotic diversity by restriction digestion of 16S rDNA directly amplified from hypersaline environments. *FEMS Microbiol. Ecol.* 17, 247-256.

**Maxam AM, Gilbert W. 1977.** A new method for sequencing DNA *Feb*;74(2):560-4.

**McInnery, J.O., Wilkinson, M., Patching, J.W., Embley, T.M., Powell, R., 1995.** Recovery and phylogenetic analysis of novel archaeal rRNA sequences from deep-sea deposit feeder. *Appl. Environ. Microbiol.* 61, 1646-1648

**Mirete S, De Figueras CG & Gonzalez-Pastor JE 2007** Novel nickel resistance genes from the rhizosphere metagenome of plants adapted to acid mine drainage. *Appl Environ Microbiol* 73: 6001-601

**Miskin JE, Farrel AM, Cunliffe WJ, Holland KT 1977.** *Propionibacterium acnes*, a resident of lipid-rich human skin, produces a 33 kDa extracellular lipase encoded by *gehA*. *Microbioloty* 143(Pt 5):1745-1755

**Molgaard A 2002** Rhamnogalacturonan acetyltransferase, a member of the SGNH-hydrolase family. In *Advances in Pectin and Pectinase Research*. F Voragen, H Schols, RGF Visser (eds). Dordrecht: Kluwer, pp.299-313



**Mori, T., Mizuta, S., Suenaga, H. and K. Miyazaki. 2008.** Metagenomic Screening for Bleomycin Resistance Genes. *Applied and Environmental Microbiol.* 74: 6803-6805

**Muraoka T, Ando T, Okuda H. 1982.** Purification and properties of a novel lipase from *Staphylococcus aureus* 226. *J. Biochem* 92: 1933-1939.

**Muyzer G, de Wall, E.C, and Uitterlinden A.G. 1993.** Profiling of complex microbial populations by denaturing gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16 rRNA. *Appl. Environ. Microbiol.* 59,695-700.

**Navarro-Noya Y.E; Roblero J.J; González-Chavez M.; Hernández-Gama R; Hernández-Rodríguez C. 2010** Bacterial communities associated with the rhizosphere of Pioneer plants (*Bahia xyloperda* and *Viguiera linearis*) growing on heavy metals-contaminated soil. *Antonie van Leeuwenhoek*

**Nemergut DR, Anderson SP, Cleveland CC, Martin AP, Miller AE, Seimon A et al. 2007.** Microbial community succession in an unvegetated, recently deglaciated soil. *Microb Ecol* 53: 110-122

**Nichols D, 2007.** Cultivation gives context to the microbial ecologist. *FEMS microbiol Ecol* 60(3):351-357.

**Nishizawa M, Shimizu M, Ohkawa H, Kanaoka M 1995.** Stereoselective production of (+)-transchrysanthermic acid by microbial esterase: cloning, nucleotide sequence, and overexpression of the esterase gene of *Arthrobacter globiformis* in *Escherichia coli*. *Appl Environ Microbiol* 61: 3208-3215

**Nthangeni M.B, Ramagoma F, Matsobane G, Litthauer D. 2005.** Development of a versatile cassette for directional genome walking using cassette ligation-mediated PCR and its application in the cloning of complete lipolytic genes from *Bacillus* species. *Journal of Microbiological Methods* Volume 61, Issue 2 225-234

**Nthangeni MB, Patterton H, van Tondre A, Vergerr WP, Litthauer D 2001.** Over-expression and properties of a purified recombinant *Bacillus licheniformis* lipase: a comparative report on *Bacillus* lipases. *Enzyme Microb Technol* 28: 705-712

**Nunan N, Daniell TJ, Singh BK, Papert A, Mc Nicol JW & Prosser JI. 2005.** Links between plant and rhizosphere bacterial communities in grassland soils, characterized using molecular techniques. *Appl Environ Microbiol.* 71(11):6784-92.

**Obino H; Ishikawa H. 2001.** Enzymes which are stable in the presence of organic solvents. *Journal of bioscience and bioengineering.* Vol.91,No2,109-116.2001

**Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Herl M, Remington SJ, Silman I, Schrag J 1992.** The alpha/beta hydrolase fold. *Protein Eng* 5:197-211

**Olsen, G.J; Larsen N, and Woese, CR . 1991.** The ribosomal RNA database project. *Nucleic Acids Res.* 19,2917-2021

**Orgino H, Mimitsuka T, Muto T, Matsumura M, Yasuda M, Ishimi K, Ishikawa H 2004.** Cloning, expression, and characterization of a lipolytic enzyme gene (lip8) from *Pseudomonas aeruginosa* LST-03 *J Mol Microbiol Biotechnol* 7: 212-223

**P. Garbeva, J. A. van Veen & J. D. van Elsas. 2004.** Assessment of the diversity, and antagonism towards *Rhizoctonia solani* AG3, of *Pseudomonas* species in soil from different agricultural regimes *FEMS Microbial ecology*, 47. 51-64

**Pace, N.R. 1997.** A molecular view of microbial diversity and the biosphere. *Science*, 276:734-40

**Pang H, Zhang P, Duan CJ, Mo XC, Tang JL, Feng JX. 2009.** Identification of cellulase genes from the metagenomes of compost soils and functional characterization of one novel endoglucanase. *Curr Microbiol* Apr;58(4):404-8.

**Pankhurst, C.E., Ophel-Keller, K., Doube, B.M., y Gupta, V. V. S. R. 1996.** Biodiversity of soil microbial communities in agricultural systems. *Biodiversity and Conservation*, 5: 197-209.

**Perez C, Juarez K, García-Castelles E, Soberon G, Servin-Gonzalez L 1993.** Cloning, characterization, and expresión in *Streptomyces lividans* 66 of an extracellular lipase-encoding gene from *Streptomyces* sp M11. *Genes* 123: 109-114

**Pharkya P, Nikolaev EV, Maranas CD 2003.** Review of the BRENDA database. *Metab Eng* 5:71-73

**Phlenz H.D, Boidol W, Schuttke I, Steber WR .1992.** Purification and properties of an *Arthobacter oxydans* P52 carbamate hydrolase specific for the herbicide phenmedipham and nucleotide sequence of corresponding gene. *J Bacteriol* 174:6600-6607

**Pitman N.C.A, Jørgensen P.M. 2002.** Estimating the Size of the World's Threatened Flora . *Science*. Vol. 298. no 5595, p.989

**Porteous LA, Armstrong JL, Seidler RJ, Watrud LS. 1994.** An effective method to extract DNA from environmental samples for polymerase chain reaction amplification and DNA fingerprint analysis. *Curr Microbiol.* Nov;29(5):301-7.

**Pratt RF, Govarhan CP 1984.** Beta-Lactamase-catalyzed hydrolysis of acyclic depsipeptides and acyl transfer to specific amino acid acceptors. *Proc Nat Acad Sci USA* 81:1302-1306

- Quaiser A; Ochsenreiter T; Lanz C; Schuster S.C; Treusch A.H; Eck J.; Schleper C. 2003.** Acidobacteria form a coherent but highly diverse group within the bacterial domain: evidence from environmental genomics. *Molecular Microbiology*. 50 (2). 563-575.
- R. Croteau, T. M. Kutchan, N. G. Lewis.** 2000. "Natural Products (Secondary Metabolites)". En: Buchanan, Gruissem, Jones (editores). *Biochemistry and Molecular Biology of Plants*. American Society of Plant Physiologists. Rockville, Maryland, Estados Unidos. Capítulo 24.
- Raaijmakers J.M, Timothy P.C, Steinberg C., Alabouvette C., & Yvan Moëgne-Loccoz. 2009.** The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant Soil* 321:341–361
- Rappé, M.S. y Giovannoni, S.J. 2003.** The uncultured microbial majority. *Annual Review of Microbiology*, 57:369–94.
- Rasche F, Velvis H, Zachow C, Berg G, Van Elsas JD, Sessitsch A. 2006.** Impact of transgenic potatoes expressing anti-bacterial agents on bacterial endophytes is comparable with the effects of plant genotype, soil type and pathogen infection. *J Appl Ecol* 43:555–566.
- Rashmuse KJ, Burton SG, Stafford WH, Cowan DA 2007.** Molecular characterization of a novel family VIII esterase from *Burkholderia multivorans* UWC10. *J Mol Microbiol Biotechnol* 13: 181-188
- Reetz MT. 2002** Lipases as practical biocatalysts. *Curr Opin Chem Biol*. Apr;6(2):145-50. Review
- Reysenbach, A.-L., Pace, N.R., 1995.** Archaea: A Laboratory Manual—Thermophiles In: Robb, F.T., Place, A.R. (Eds.),. Cold Spring Harbour Laboratory Press, New York, pp. 101– 107
- Reysenbach, A.-L., Wickham, G.S., Pace, N.R., 1994.** Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 60, 2113–2119
- Rhee J; Dae-Gyun, A; Yeon-Gu K.; Won-Jong, O. 2005.** New thermophilic and thermostable Esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a Metagenomic library. *Applied and environmental microbiology*.p. 817-825
- Riesenfeld, C.S., Goodman R.M. and J. Hanselman. 2004.** Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Env Microbiol.* 6: 981-989
- Robicsek A, Jacoby BA, Hooper DC 2006.** The worldwide emergence of plasmid-mediated quinolona resistance. *Lancet Infect Dis*; 6: 629-640
- Roesch, I.f.w., Fulthorpe R.R., Riva, A., Casella G, Hadwin AKM, Kent AD, Daroub SH, Camargo, FAO, Farmerie, WG, Triplett E.W . 2007.** Pyrosequencing enumerates and contrast soil microbial diversity. *The ISME Journal* 1, 283-290.

- Ronaghi M, Karamohamed S, Pettersson B, Uhlén M, Nyrén P 1996** Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem*;242(1):84-9
- Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman.2000.** Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol.* Jun;66(6):2541-7.
- Rosch C, Mergel A, Bothe H. 2002.** Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. *Appl Environ Microbiol* Aug;68(8):3818-29
- Rosenau F, Tommassen J, Jaeger KE 2004.** Lipase-specific foldases. *Chembiochem* 5: 152-161
- Rosentein R, Goetz F 2000.** Staphylococcal lipases: biochemical and molecular characterization. *Biochimie* 82: 1005-1014
- Rossello-Mora R, Amann R .2001.** The species concept for prokaryotes. *FEMS Microbial Rev* 25,36-67.
- Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R, Fierer N. 2010.** Soil bacterial and fungal communities across a pH gradient in an arable soil *ISME J*
- Royter M, Schmidt M, Elend C, Höbenreich H, Schäfer T, Bornscheuer U.T, Antranikian G. 2009** Thermoestable lipases from the extreme thermophilic anaerobic bacteria *Thermoanaerobacter thermohydrosulfuricus* SOL1 and *Caldanaerobacter subterraneus* subsp. *Tengacongensis*. *Extremophiles*, 13: 769-783
- Rudi, K., Skulberg, O.M., Larsen, F., Jacoksen, K.S., 1997.** Strain classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from variable regions V6, V7 and V8. *Appl. Environ. Microbiol.* 63, 2593– 2599.
- Sait M, Davis KER, Janssen PH. 2006.** Effect of pH on isolation and distribution of members of subdivision 1 of the phylum Acidobacteria occurring in soil. *Appl Environ Microbiol* 72: 1852–1857.
- Sait M, Hugenholtz P & Janssen PH 2002** Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation independent surveys. *Environ Microbiol* 4: 654–666.
- Sambrook, J., Fritsch, E.F. y Maniatis, T. 1989.** Molecular cloning. A laboratory Manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.
- Sanger F, Nicklen S, Coulson AR. 1977** DNA sequencing with chain-terminating inhibitors. *Dec*;74(12):5463-7.

- Sanguin H, Remenant B, Dechesne A et al. .2006.** Potential of a 16S rRNA-based taxonomic microarray for analyzing the rhizosphere effects of maize on *Agrobacterium* spp. And bacterial communities. *Appl Environ Microb* 72: 4302–4312
- Sarda L, Desnuelle P 1958.** Actions of pancreatic lipase on esters in emulsions. *Biochim Biophys Acta* 30: 513-521
- Saunders, J.R. 1994.** Population genetics of phase variable antigens. En: *Population Genetics of Bacteria*, Cambridge University Press, 247-268 pp.
- Sayari A, Agrebi N, Jaoua S, Gargorui Y 2001.** Biochemical and molecular characterization of *Staphylococcus simulans* lipase. *Biochimie* 83: 863-871
- Schauer R, Bienhol C, Ramette A, Harder J. 2010** Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. *The ISME Journal* 4, 159-170.
- Schloss, P.D, Larget B.R, Handelsman J. 2004.** Integration of Microbial Ecology and Statistics: a Test to Compare Gene Libraries. *Appl Environ. Microbiol.* 70: 5485-5492.
- Schlöter, M., Dilly, O. y Munch, J.C. 2003.** Indicators for evaluating soil quality. *Agriculture, Ecosystems and Environment*, 98: 255–262.
- Schmalenberger A &Tebbe CC 2003** Bacterial diversity in maize rhizospheres: conclusions on the use of genetic profiles based on PCR-amplified partial small subunit rRNA genes in ecological studies. *Mol Ecol* 12: 251–261
- Schmidt-Dannert C, Rfia M.L, Atomi H., Schmid RD.1996.** Thermoalkalophilic lipase of *Bacillus thermocatenuatus*. I. . Molecular cloning, nucleotide sequence, purification and some properties. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* Volume 1301, Issues 1-2, 105-114
- Schomburg I, Chang A, Ebeling C, Bremse M, Hedlt C, Huhn G, Schomburg D 2004.** BRENDA, the enzyme database: updates and major new developments. *Nucleic Acids Res* 32: D431-433
- Schuette M, Fetzner S 2007.** EstA from *Arthrobacter nitroguajacolicus* Ru61a, a thermo- and solvent-tolerant carboxylesterase related to class C beta-lactamases. *Curr Microbiol* 54: 230-236
- Semeriva M, Chapus C, Bouvier-Lapierre C, Desnuelle P 1974.** On the transient formation of an acetyl enzyme intermediate during the hydrolysis of p-nitrophenyl acetate by pancreatic lipase. *Biochem Biophys Res Commun* 58: 808-813
- Sharma S, Aneja MK, Mayer J, Munch JC & Schlöter M 2005** Characterization of bacterial community structure in rhizosphere soil of grain legumes. *Microb Ecol* 49: 407–415.

**Sharma S, Khan FG, Qazi GN. 2010.** Molecular cloning and characterization of amylase from soil metagenomic library derived from Northwestern Himalayas *Appl Microbiol Biotechnol.*;86(6):1821-8

**Siciliano, S. D., N. Fortin, A. Mihoc, G. Wisse, S. Labelle, D. Beaumier, D. Ouellette, R. Roy, L. G. Whyte, M. K. Banks, P. Schwab, K. Lee y C. W. Greer. 2001.** Selection of specific endophytic bacterial genotypes by plants in response to soil contamination. *Appl. Environ. Microbiol.* 67: 2469- 2475.

**Simon C, Daniel R. 2009** Achievements and new knowledge unraveled by metagenomic approaches. *Appl Microbiol Biotechnol.*;85(2):265-76. Epub 2009 Sep 16.

**Simons JW, van Kampen MD, Riel S, Gotz F, Egmond WR, Verheij HM 1998.** Cloning, purification and characterisation of the lipase from *Staphylococcus epidermidis*-comparison of the substrate selectivity with those of other microbial lipases. *Eur J Biochem* 253: 675-683

**Singh BK, Munro S, Potts JM & Millard P 2007** Influence of grass species and soil type on rhizosphere microbial community structure in grassland soils. *Appl Soil Ecol* 36: 147-155.

**Smalla K, Haines AS, Jones K, Krögerrecklenfort E, Heuer H, Schloter M, Thomas CM. 2006.** Increased abundance of IncP-1beta plasmids and mercury resistance genes in mercury-polluted river sediments: first discovery of IncP-1beta plasmids with a complex mer transposon as the sole accessory element. *Appl Environ Microbiol.*;72(11):7253-9. Epub 2006 Sep 15.

**Smalla K, Wieland G, Buchenr A, Zock A, Parzy J, Kaiser S, Roskot N, Heuer H, Ber G. 2001.** Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Appl. Environ. Microbiol.* 67: 4742:4751.

**Smets BF, Barkay T. 2005.** Horizontal gene transfer: perspectives at a crossroads of scientific disciplines. *Nat Rev Microbiol.* 3(9):675-8

**Snellman EA, Sullivan ER, Colwell RR. 2002.** Purification and properties of the extracellular lipase, Lip A, of *Acinetobacter* sp. RAG-1- *Eur J. Biochem* 269: 5771-5779

**Somers E, Vanderleyden J & Srinivisam M. 2004.** Rhizosphere bacterial signaling: a love parade beneath our feet. *Crit Rev Microbiol* 30: 205-240

**Sommer P, Bormann C, Gotz F 1997.** Genetic and biochemical characterization of a new extracellular lipase from *Streptomyces cinnamomeus*. *Appl Environ Microbiol* 63: 35553-3560

**Sørensen, J. 1997.** The rhizosphere as a habitat for soil microorganisms. *In:* JD van Elsas; JT Trevors & EMH Wellington. (eds). *Modern Soil Microbiology*. Marcel Dekker, New York, pp 21-45

**Stackebrandt, E, Frederiksen W, Garrity GM, Grimont P.A.D., Kampfer P, Maiden M.C.J, Nesme, S, Rossello-Mora R, Swings J, and Truper, H.G et al .2002.** Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int.J.Syst.Evol.Microbiol.* 52,1043-1047.

**Stafford WHL, Baker GC, Brown SA, Burton SG & Cowan DA 2005** Bacterial diversity in the rhizosphere of Proteaceae species. *Environ Microbiol* 7: 1755–1768.

**Stevenson BS, Eichorst SA, Wertz JT, Schmidt TM & Breznak JA 2004** New strategies for cultivation and detection of previously uncultured microbes. *Appl Environ Microb* 70: 4748–4755

**Studier, F.W.** 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol Biol.* 219: 37-44.

**Sugiura J, Isobe M, Muroya N, Yamaguchi T. 1974.** Purification and properties of a Chromobacterium lipase with a high molecular weight. *Agric Biol Chem* 38: 947-952.

**Suzuki, M.T., Giovanni, S.J., 1996.** Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* 62, 625–630

**Tamames J, Abellán JJ, Pignatelli J, Camacho A, Moya A.2010.** Environmental distribution of prokaryotic taxa. *BMC Microbiology*,10:85

**Tarlera S, Jangid K, Ivester AH,Whitman WB,Williams MA. 2008.** Microbial community succession and bacterial diversity in soils during 77,000 years of ecosystem development. *FEMS Microbiol Ecol* 64: 129–140

**Tiedje, J.M., Asuming-Brempong, S., Nusslein, K., Marsh, T.L., y Flynn S.J. 1999.** Opening the black box of soil microbial diversity. *Applied Soil Ecology*, 13: 109–122.

**Tilman, D. 1982.** Resource Competition and Community Structure. Princeton, NJ: Princeton Univ. Press

**Tilman, D. 2000.** Causes, consequences and ethics of biodiversity. *Nature*, 405: 208–11

**Torsvik, V., Goksoyr, J.Y. y Daae, F.L. 1990.** High diversity in DNA of soil bacteria. *Applied and Environmental Microbiology*, 56: 782-787.

**Torsvik, V., Ovreas, L. y Thingstad, T.F. 2002.** Prokaryotic diversity—magnitude, dynamics, and controlling factors. *Science*, 296:1064–66.

**Tyski S, Hryniewicz W, Jeljaszewicz J 1983.** Purification and some properties of the staphylococcal extracellular lipase. *Biochim Biophys Acta* 749: 312-317

**van Elsas JD, Jansson JK, Trevors JT (eds). 2007.** Modern soil microbiology II. CRC Press-Taylor and Francis, Boca Raton, p 646

- Van Elsas JD, Speksnijder AJ & van Overbeek LS 2008** A procedure for the metagenomics exploration of diseasesuppressive soils. *J Microbiol Meth* 75: 515-522.
- Van Elsas, J.D., Trevors, J.T. y Wellington, E.M.H. 1997.** *Modern soil microbiology*. Marcel Dekker, New York.
- Van Oort MG, Deever AMTJ, Dijkman R, Tjeen ML, Verheij HM, Haas GH de, Wenzig E, Gotz F. 1989.** Purification and substrate specificity of *Staphylococcus hyicus* lipase. *Biochemistry* 28: 9278-9285.
- Van Pouderooyen G, Eggert T, Jaeger KE, Dijkstra BW 2001** The cristal struture of *Baccillus subtilis* lipase: a minimal alpha/beta hydrolase fold enzyme. *J Mol Biol* 309: 215-226
- Van Tillbeurgh H, Egloff MP, Martinez C, Rugani N, Verger R, Cambillau C 1993.** Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography. *Nature* 362: 814-820
- Van Veen JA. 2002.** Effects of above-ground plant species composition and diversity on the diversity of soil-borne microorganisms *Antonie van Leeuwenhoek* 81: 509-520.
- Van Veen, J.A., van Overbeek, L.S., van Elsas, J.D. 1997.** Fate and activity of microorganisms introduced into soil. *Microbiology and Molecular Biology Reviews*, 61:121-35.
- Villadas PJ, Fernández-López M, Ramírez-Saad H, Toro N. 2007.** Rhizosphere-bacterial community in *Eperua falcata* (Caesalpinaceae) a putative nitrogen-fixing tree from French Guiana rainforest. *Microb Ecol.* 53:317-27.
- Waksman S.A, Woodruff H.B. 1940** The soil as a source of microorganisms antagonistic to disease-producing bacteria.*J. Bacteriol.* 40, 581
- Wang J, Jenkins C, Webb R.I, Fuerst J.A. 2002.** Isolation of Gemmata-like and Isosphaera-like planctomycete bacteria from soil and freshwater. *Appl. Environ. Microbiol.* 68: 417-422.
- Wang J, Shen WJ, Pate S, Harada K, Kraemer FB 2005.** Mutational analysis of the “regulatory module” of hormone-sensitive liapse. *Biochemistry* 44:1953-1959
- Wang M, Chen JK & Li B 2007** Characterization of bacterial community structure and diversity in rhizosphere soils of three plants in rapidly changing salt marshes using 16S rDNA. *Pedosphere* 17: 545-556
- Wang, Q, G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007.** Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol.* 73(16): 5261-7.



- Ward-Rainey N, Rainey FA, Schiesner H & Stackebrandt E .1995.**  
**Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M. et al.2007** Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, 450, 560–565.
- Watanabe, K., Kodama, Y., Harayama, S., 2001.** Design and evaluation of PCR primers to amplify 16S ribosomal DNA fragments used for community fingerprinting. *J. Microbiol. Methods*44, 253–262
- Wayne, J. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. y Truper, H. G. 1987.** Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Bacteriology*, 37: 463-464.
- Wei P, Bai L, Song W, Hao G. 2009.**Characterization of two soil metagenome-derived lipases with high specificity for p-nitrophenyl palmitate. *Arch Microbiol.* Mar;191(3):233-40. Epub 2008 Nov 29
- Wei Y, Swanson L, Castro C, Derewenda U, Minor W, Arai H, Aoki J, Inoue K, Servin-Gonzalez L, Derewenda ZS 1998.** Structure of a microbial homologue of mammalian platelet-activating factor acetylhydrolases: *Streptomyces exfoliatus* lipase at 1.9 Å resolution. *Structure* 6: 511-519
- Wen-tso liu y Janet Janson, 2010.** Environmental molecular microbiology. Libro. +232 pp. ISBN 978-1-904455-52-3
- Whitman, W.B., Coleman, D.C. & Wiebe, W.J.1998.** Prokaryotes: the unseen majority *Proc. Natl. Acad. Sci. USA* 95: 6578-6583
- Williamson G, Paul A., Faulds C 1998.**Hairy plant polysaccharides: a close shave with microbial esterases. *Microbiology*, 144, 2011-2023
- Woese CR, Kandler O, Wheelis ML. 1990** Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya ;87(12):4576-9. *Proc Natl Acad Sci U S A*.
- Woese, C.R. 1987.** Bacterial evolution. *Microbiological Reviews*, 51: 221–71
- Wolfgang Ludwig, Oliver Strunk, Ralf Westram, Lothar Richter, Harald Meier, Yadhukumar, Arno Buchner, Tina Lai, Susanne Steppi, Gangolf Jobb, Wolfram Förster, Igor Brettske, Stefan Gerber, Anton W. Ginhart, Oliver Gross, Silke Grumann, Stefan Hermann, Ralf Jost, Andreas König, Thomas Liss, Ralph Lüssemann, Michael May, Björn Nonhoff, Boris Reichel, Robert Strehlow, Alexandros Stamatakis, Norbert Stuckmann, Alexander Vilbig, Michael Lenke, Thomas Ludwig, Arndt Bode and Karl-Heinz Schleifer.2004.** ARB: a software environment for sequence data. *Nucleic Acids Research* 32(4):1363-1371.
- Wood, R.D., Keller, C.K. Y Johnstone, D.L. 1993.** In situ measurement of microbial activity and controls on microbial CO<sub>2</sub> production in the unsaturated zone. *Water Resources Research*, 29: 647-659.

- Wright GD 2007.** The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol*; 5: 175-186.
- Yadav, RP, Saxena RK, Gupta R, Davidson WR. 1998.** Rapid zymogram for lipase. *Bio Techniques* 24: 754-756.
- Yamaguchi, M., Flocker, W.J. Y Howard, F.D. 1967.** Soil atmosphere as influenced by temperature and moisture. *Soil Science Society of America Proceedings*, 31: 164-167.
- Yeaman, S.J. 2004** Hormone-sensitive lipase-new roles for an old enzyme. *Biochem J.* 379, 11-22
- Yuhong Z, Shi P, Liu W, Meng K, Bai Y, Wang G, Zhan Z, Yao B. 2009.** Lipase diversity in glacier soil based on analysis of metagenomic DNA fragments and cell culture. *J Microbiol Biotechnol.*;19(9):888-97
- Yuhong Z, Shi P, Liu W, Meng K, Bai Y, Wang G, Zhan Z, Yao B. 2009.** Lipase diversity in glacier soil based on analysis of metagenomic DNA fragments and cell culture. *J Microbiol Biotechnol.*;19(9):888-97.
- Yun J, Kang S, Park S, Yoon H, Kim MJ, Heu S, Ryu S.2004** Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library. *Appl Environ Microbiol.* Dec;70(12):7229-35
- Zaks, A; Klibanow,A.M. 1985** Enzyme-catalyzed processes in organic solvents. *Proc. Natl.Acad.Sci. USA* 82, 3192-3196
- Zhang HB, Shi W, Yang MX, Sha T & Zhao ZW 2007** Bacterial diversity at different depths in lead-zinc mine tailings as revealed by 16S rRNA gene libraries. *J Microbiol* 45: 479-484.
- Zhang, H. Sekiguchi Y., Hanada S, Hugenholtz P, Kim H, Kamagata Y, Nakamura K. 2003.** *Gemmatimonas aurantiaca* gen. nov., sp. nov., a gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Gemmatimonadetes phyl.nov. *Int. J. Syst. Evol. Microbiol.* 53: 1155-1163
- Zhu X, Larsen NA, Basran A, Bruce NC, Wilson IA 2003.** Observation of an arsenic adduct in an acetyl esterase crystal structure. *J Biol Chem* 278: 2008-2014
- Zock J, Cantwell C Swartling J, Hodges R, Phol T, Sutton K, Rosteck P Jr, McGilvray D, Queener S 1994.** The *Bacillus subtilis* pnbA gene encoding p-nitrobenzyl esterase: cloning, sequence and high-level expression in *Escherichia coli*. *Gene* 151: 37-47
- Zollner, H.K; Schroeder, L.R. 1998.** Enzymatic deinking of nonimpact printed white office paper with  $\alpha$ -amylase. *Tappi J.*,vol.81,3,166-170

**Zul D, Denzel S, Kotz A & Overmann J .2007.** Effects of plant biomass, plant diversity, and water content on bacterial communities in soil lysimeters: implications for the determinants of bacterial diversity. *Appl Environ Microb* 73: 6916–6929.

**Zwolinski MD. 2007.** DNA sequencing: strategies for soil microbiology. *SSSAJ* 71:592–600