

03088  
3



UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO

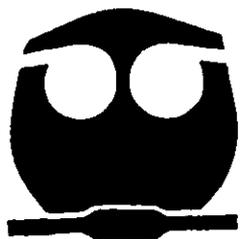
FACULTAD DE QUIMICA

"DETERMINACION DE LA DIVERSIDAD DE  
BACTERIAS LACTICAS DEL POZOL POR ANALISIS  
DE SECUENCIAS DE ADNr 16S"

T E S I S  
QUE PARA OBTENER EL GRADO DE:  
DOCTOR EN BIOTECNOLOGIA  
P R E S E N T A :

M. EN B. JOSE ADELFO ESCALANTE LOZADA

286747



MEXICO. D. F.

2000



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



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

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

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

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

**“DETERMINACIÓN DE LA DIVERSIDAD DE BACTERIAS LÁCTICAS DEL  
POZOL POR ANÁLISIS DE SECUENCIAS DE ADN<sub>r</sub> 16S”**

**TESIS QUE PARA OBTENER EL GRADO DE  
DOCTOR EN BIOTECNOLOGÍA  
PRESENTA:**

**M. EN B. JOSÉ ADELFO ESCALANTE LOZADA**

**ASESOR**

**DRA. AMELIA FARRÉS GONZÁLEZ-SARAVIA**

Departamento de Alimentos y Biotecnología

Facultad de Química, UNAM

**COMITÉ TUTORIAL ASIGANDO:**

**DRA. VALERIA SOUZA-SALDÍVAR**

Instituto de Ecología, UNAM

**DR. FERNANDO VALLE BAHEZA**

Instituto de Biotecnología, UNAM

**ESTE TRABAJO SE REALIZÓ EN EL  
LABORATORIO 312,  
DEPARTAMENTO DE ALIMENTOS Y BIOTECNOLOGÍA,  
CONJUNTO E, FACULTAD DE QUÍMICA, UNAM.**

**JURADO ASIGANDO PARA LA REVISIÓN DE ESTE TRABAJO:**

**DRA. ESPERANZA MARTÍNEZ**

**Centro de Investigaciones en Fijación de Nitrógeno, UNAM**

**DRA. VALERIA SOUZA SALDÍVAR**

**Instituto de Ecología, UNAM**

**DRA. MARÍA DEL CARMEN WACHER-RODARTE**

**Facultad de Química, UNAM**

**DR. CARLOS ESLAVA CAMPOS**

**Facultad de Medicina, UNAM**

**DR. JOSÉ FRANCISCO FERMÁNDEZ PERRINO**

**Depto. de Biotecnología, Universidad Autónoma Metropolitana**

**DR. ENRIQUE MERINO**

**Instituto de Biotecnología, UNAM**

**DR. GUILLERMO GOSSET**

**Instituto de Biotecnología, UNAM**

**DR. JOSÉ FRANCISCO FERMÁNDEZ PERRINO**

**Depto. de Biotecnología, Universidad Autónoma Metropolitana**

Para el desarrollo del presente trabajo se contó con los siguientes apoyos:

CONACYT 4688N

DGAPA 210194

PADEP 005353, 005357, 030515

BECA DE CONACYT

A mi querida esposa Martha y a mis dos pequeñas y adoradas hermanas Ximi y Vi,  
a quienes dedico este trabajo.

Gracias por todo su cariño y paciencia sin los cuales nunca hubiera podido  
concluir esta tesis. Todo lo que representa este trabajo es más para ustedes que  
para mí.

Amelia:

Sería muy difícil expresar en varias líneas todo lo que me han dejado todos estos años que pase en el grupo de los (las) Farrés. Tomando como pretexto las bacterias lácticas y empezando con el club de los Tobis, pasando por dos mudanzas y el estrenar un laboratorio, hasta éste momento por demás importante para mí. Creo que sería muy poco expresar con un simple "gracias Amelia" toda la confianza, oportunidades y toda tu ayuda y apoyo incondicional durante mi formación y sobre todo en algunos momentos muy difíciles para mí y mi joven familia.

Agradezco a cada uno de los miembros del Jurado asignado para la revisión de este documento, por todos sus amables comentarios y críticas que contribuyeron de forma muy importante para la presentación final de este trabajo.

A Ismael, Rodolfo, Rene y Noé. Como memoria de aquellos encuentros filosóficos entre el Doom y los -OH's. Gracias por haber hecho toda una experiencia nuestra estancia en el 312.

A papá, mamá, mis hermanos y mi familia Giles Gómez. Su apoyo y ayuda también hicieron posible este trabajo. Gracias.

## ÍNDICE

CONTENIDO	Página
<b>RESUMEN</b>	i
<b>ABSTRACT</b>	ii
<b>Capítulo 1. INTRODUCCIÓN</b>	1
<b>Capítulo 2. GENERALIDADES</b>	6
2.1 IMPORTANCIA DE LA DIVERSIDAD MICROBIANA.	6
2.2 APROXIMACIONES METODOLÓGICAS AL ESTUDIO DE LA DIVERSIDAD MICROBIANA DE UN AMBIENTE.	9
2.3 MÉTODOS MOLECULARES PARA LA DETERMINACIÓN DE LA ESTRUCTURA MICROBIANA DE UNA MUESTRA AMBIENTAL.	11
Extracción de ácidos nucleicos de una muestra ambiental.	15
Amplificación, clonación y secuenciación de genes específicos de una muestra ambiental.	16
2.4 EXTRACCIÓN DEL ADN MICROBIANO TOTAL DE UNA MUESTRA AMBIENTAL.	17
Extracción indirecta de ADN.	20
Extracción directa de ADN.	21
2.5 SECUENCIAS PARCIALES Y COMPLETAS DE ARNr 16S. SU USO E IMPLICACIONES EN ESTUDIOS DE ECOLOGÍA MICROBIANA MOLECULAR.	26
La búsqueda de un marcador óptimo para estudios filogenéticos.	28
Las especies de ARNr como marcadores filogenéticos.	29
El uso de secuencias parciales y la posibilidad de determinar líneas de principales descendencia.	31
Determinación de la posición filogenética de una clona a	

partir de una secuencia parcial de ADNr.	34
2.6 ANÁLISIS DE DATOS MOLECULARES: ALINEAMIENTO, CALCULO DE DISTANCIAS Y CONSTRUCCIÓN DE ÁRBOLES.	37
Alineamiento.	37
Tipos de datos: distancias <i>versus</i> caracteres.	38
Métodos de agrupamiento <i>versus</i> métodos de búsqueda.	39
Medidas de distancia y similitud.	40
Medidas de similitud.	43
Clasificación de los métodos filogenéticos más comunes para la construcción de árboles.	43
Algoritmos para la construcción de árboles de distancia.	44
Métodos discretos.	47
<b>Capítulo 3. HIPÓTESIS y OBJETIVOS.</b>	49
<b>Capítulo 4. MATERIAL Y MÉTODOS.</b>	51
4.1 EXTRACCIÓN DEL ADN TOTAL DE LA FRACCIÓN BACTERIANA DEL POZOL POR UN MÉTODO INDIRECTO.	51
Obtención de una muestra de pozol en la localidad de San Cristóbal de las Casas, Chiapas.	51
Separación de la fracción bacteriana del pozol.	51
Extracción y purificación del ADN bacteriano.	55
Lisis de la fracción bacteriana aislada.	55
Purificación del ADN bacteriano.	56
4.2 AMPLIFICACIÓN POR PCR DE LOS ADNr 16S PRESENTES, CLONACIÓN DEL PRODUCTO DE AMPLIFICACIÓN E IDENTIFICACIÓN DE LAS CLONAS CON INSERTO.	58
Amplificación por PCR de los ADNr 16S presentes.	58
Clonación del producto de amplificación e identificación de las clonas con inserto.	58
Reacción de ligación.	59
Transformación.	59
Selección del fenotipo blanco-azul.	60

Verificación de la presencia de in inserto clonado correspondiente al tamaño del ADNr 16S.	62
4.3 SECUENCIACIÓN PARCIAL DE LOS ADNr 16S CLONADOS.	63
4.4 IDENTIFICACIÓN Y ANÁLISIS DE LAS SECUENCIAS PARCIALES DE ADNr 16S CLONADOS.	66
Alineamiento de secuencias.	66
Construcción de un árbol filogenético.	66
Análisis de posibles secuencias quiméricas.	67
Titularidad de los programas de cómputo empleados para los análisis de secuencias.	71
<b>Capítulo 5. RESULTADOS</b>	72
5.1 EXTRACCIÓN DEL ADN TOTAL DE LA FRACCIÓN BACTERIANA DEL POZOL POR UN MÉTODO INDIRECTO.	72
Separación de la fracción bacteriana del pozol y lisis celular.	72
Extracción de ADN bacteriano del pozol.	72
5.2 AMPLIFICACIÓN POR PCR DE LOS ADNr 16S, CLONACIÓN DEL PRODUCTO DE AMPLIFICACIÓN E IDENTIFICACIÓN DE LAS CLONAS CON INSERTO.	75
Amplificación por PCR de los ADNr 16S.	75
Clonación del producto de amplificación y evaluación de la presencia del inserto	75
5.3 IDENTIFICACIÓN Y ANÁLISIS DE LA SECUENCIA PARCIAL DE LAS CLONAS Pz.	79
Identificación de clonas en la base de datos de secuencias nucleotídicas no redundantes del NCBI.	79
Análisis de posibles secuencias quiméricas.	80
Alineamiento de secuencias.	80
Construcción de un árbol filogenético y determinación de la filiación taxonómica de las clonas del pozol.	80
<b>Capítulo 6. DISCUSIÓN</b>	86
6.1 DIVERSIDAD DE BACTERIAS LÁCTICAS DEL POZOL	86

6.2	IMPLICACIONES BIOTECNOLÓGICAS DEL ESTUDIO DE LA DIVERSIDAD BACTERIANA DEL POZOL.	99
	<b>Capítulo 7. CONCLUSIONES y PERSPECTIVAS.</b>	103
	<b>Capítulo 8. REFERENCIAS</b>	106
	<b>Capítulo 9. APÉNDICE</b>	118
	RESULTADOS DEL ANÁLISIS DE LAS SECUENCIAS DE LAS CLONAS PZ EN EL PROGRAMA BLAST	118

## ÍNDICE DE FIGURAS

Figura 1. Representación general de varios componentes celulares y técnicas utilizadas para la identificación y clasificación bacteriana.	12
Figura 2. Resolución taxonómica de las metodologías comúnmente utilizadas para identificación bacteriana.	13
Figura 3. Estrategias basadas en el análisis de ADNr/ARNr para la caracterización de comunidades microbianas sin el cultivo bacteriano.	18
Figura 4. Sumario de las dos aproximaciones metodológicas para la extracción de ADN de muestras ambientales.	23
Figura 5. Comparación de los árboles filogenéticos construidos con diferente cantidad de información de secuencias del ADNr 16S: incapacidad de utilizar secuencias parciales para determinar la posición intergenérica de especies del género <i>Vibrio</i> .	32
Figura 6. Muestra de pozol tipo indígena colectada del mercado de San Cristóbal de Las Casas, Chiapas y empleado como sustrato para la extracción del ADN de la fracción bacteriana presente.	52
Figura 7. Proceso de extracción indirecta de ADN bacteriano del pozol.	53
Figura 8. Obtención de las secuencias parciales de los ADNr 16S obtenidos por PCR a partir del ADN bacteriano extraído del pozol y su posterior análisis.	57
Figura 9. Electroforesis en gel de agarosa 1% de ADN de la fracción bacteriana extraída del pozol.	74
Figura 10. Electroforesis en gel de agarosa 1.5% del los ADNr 16S amplificados por PCR a partir de la fracción de ADN bacteriano extraída del pozol, empleando el par de <i>primers</i> fD1+rD1.	76
Figura 11. Electroforesis en gel de agarosa 1% de digestiones	

enzimáticas con <i>Bam</i> HI – <i>Hind</i> III del plásmido pT7-Blue con el ADNr 16S amplificado por PCR clonado.	77
Figura 12. Electroforesis en gel de agarosa 1.5% del producto de amplificación por PCR empleando como templado ADN plasmídico obtenido de minipreparaciones alcalinas, y clonado en el plásmidos pT7Blue.	78
Figura 13. Alineamiento de secuencias de clonas Pz con secuencias de referencia en el programa CLUSTAL W(versión 1.60).	82
Figura 14. Árbol filogenético de secuencias parciales de ADNr 16S de clonas Pz y secuencias de microorganismos de referencia.	85
Figura 15. Árboles filogenéticos construidos a partir del análisis de 450 bases.	93
Figura 16. Árbol construido por el método de UPGMA a partir del análisis de las secuencias de las clonas PZ y secuencias de referencia obtenidas de la base de datos del GenBank.	95
Figura 17. Procesos microbianos presentes en el pozol con posibles implicaciones biotecnológicas a partir de microorganismos específicos.	102

## LISTA DE TABLAS

Tabla 1. Categorías y caracteres aplicados en sistemática bacteriana.	14
Tabla 2. <i>Primers</i> empleados para la amplificación por PCR de genes de ARNr 16S.	19
Tabla 3. Factores a considerar en cada etapa de extracción de ADN microbiano a partir de muestras ambientales.	22
Tabla 4. Comparación de los protocolos de extracción directo e indirecto de ADN.	24
Tabla 5. Grado de similitud entre diferentes regiones hipervariables de ADNr 16S.	33
Tabla 6. Valores de similitud de ADNr 16S obtenidos de secuencias de clonas ambientales y su parentesco cultivable más cercano.	36
Tabla 7. Características de los <i>primers</i> rD1 y fD1 empleados para la amplificación de genes de ARNr 16S.	60
Tabla 8. Eficiencia del proceso de extracción de la fracción bacteriana del pozol.	73
Tabla 9. Organismos identificados en la biblioteca de clonas Pz.	81
Tabla 10. Trabajos en los que se emplean secuencias parciales y completas de ADNr para establecer la filogenia y ubicación taxonómica de clonas ambientales en estudios de determinación de diversidad microbiana.	91

## RESUMEN

En el presente trabajo se determinó la diversidad de bacterias lácticas presente en una muestra de pozol tipo indígena con 8-10 horas de fermentación, por medio del análisis de secuencias parciales de ADNr 16S. Estos genes fueron amplificados por PCR empleando un par de *primers* universales hacia bacterias Gram (+) y relacionadas, a partir del ADN extraído de la fracción bacteriana total de la muestra de pozol por un método indirecto. Los ADNr 16S amplificados por PCR fueron clonados y se obtuvo una colección de 36 clones de los cuales se detectaron 16 que presentaron una secuencia única. Cada una de ellas fue analizada en la base de datos de secuencias nucleotídicas no redundantes del NCBI empleando el programa BLAST. Se determinó que 14 secuencias corresponden a ADNr 16S y dos a ADN no ribosomal de origen bacteriano y humano respectivamente. El análisis en el programa BLAST cada secuencia parcial (650 bases) permitió identificar las secuencias de ADNr 16S como bacterias lácticas de los géneros *Lactococcus lactis*, *Streptococcus bovis/suis*, *Lactobacillus plantarum*, *L. casei*, *L. alimentarium* y *L. Delbruekii* y una clona fue identificada como *Clostridium sp* Lo que permite señalar a estos microorganismos como la microbiota dominante en la muestra analizada. Las secuencias de ADNr 16S fueron alineadas con secuencias de referencia obtenidas del GENBAK, se construyó una matriz de distancia por medio del algoritmo de Jukes & Cantor y un árbol por *neighbour joining*. La posición de las clonas en el árbol respecto a las cepas de referencia validan su identidad. Se presenta una discusión sobre la presencia, abundancia y redundancia de las secuencias detectadas y de las implicaciones biotecnológicas de esta aproximación metodológica en el estudio de los alimentos fermentados tradicionales.

## ABSTRACT

This work reports the diversity of lactic acid bacteria present in a sample of a traditional Mexican Indian pozol with a fermentation time of 8-10 hours, by analysis of partial sequences of 16S rDNA. These genes were PCR amplified using a set of universal primers specific for Gram (+) bacteria and relatives using as template DNA extracted from the bacterial fraction isolated from pozol. Amplified genes lead to obtain a 36 clones library. 16 shown unique sequences. Unique partial sequences were analyzed using the BLAST program at the non-redundant nucleotidic database at NCBI homepage. This analysis lead to identify 14 clones as 16S rDNA and two non-ribosomal sequences. Identified organisms were *Lactococcus lactis*, *Streptococcus bovis/suis*, *Lactobacillus plantarum*, *L. casei*, *L. alimentarium*, *L. delbruekii* and one clone as *Clostridium sp.* All these clones were aligned, and a distance matrix was constructed using the Jukes-Cantor algorithm. A tree was obtained using the neighbour joining method. Position of each 16S rDNA clone sequences respect reference ones retrieved from GENBANK database validates its identity. The presence, relative abundance and redundancy of identified sequences and the biotechnological implications of this methodological approach in the study of traditional fermented foods are discuted.

Habiéndose acercado el tiempo de la creación, el *Ahau Tepeu* y el *Ahau Gucumatz* buscaron sustancia para hacer la carne del hombre.

Consultaron entre sí de que forma lo harían, porque los pasados hombres habían sido imperfectos, buscando cosa que pudiera servir para carne de aquél, se les manifestó en esta forma: Cuatro animales les manifestaron la existencia de las mazorcas de maíz blanco y de maíz amarillo. Estos animales fueron: *Yak*, el Gato de Monte; *Utiú*, el Coyote; *Quel*, la Cotorra, y *Hoh*, el Cuervo. La abuela *Ixmucamé* tomó del maíz blanco y del maíz amarillo e hizo comida y bebida, de las que salió la carne y la gordura del hombre, y de esta comida fueron hechos sus brazos y sus pies.

De esto formaron el Señor *Tepeu* y *Gucumatz* a nuestros primeros padres y madres.

*Popol Vuh.*

# 1

## INTRODUCCIÓN.

---

Los alimentos fermentados tradicionales presentan algunas características que los convierten en sistemas de estudio muy interesantes desde el punto de vista biotecnológico (Wacher 1993; Fleet 1999). Algunas de estas características son:

- a. Gran importancia regional
- b. Transformación de materias primas en productos por biocatálisis
- c. Mejoramiento de la calidad nutricional por fermentación
- d. Presencia de microorganismos con características de interés, tales como producción de enzimas hidrolíticas, producción de vitaminas o aminoácidos esenciales y producción de ácidos orgánicos
- e. Desarrollo de una microbiota predominante característica de cada alimento, con capacidades metabólicas especializadas
- f. Inhibición de microorganismos no deseables durante la fermentación, debido a la producción de sustancias antimicrobianas por la microbiota dominante.

Desde el punto de vista de la ecología microbiana, los alimentos fermentados tradicionales se pueden considerar como comunidades microbianas extremadamente complejas debido a la variedad de microorganismos presentes, a las interacciones existentes entre los microorganismos y entre éstos con el ambiente.

En países en desarrollo la elaboración de los alimentos fermentados tradicionales son procesos aún generalmente procesos artesanales que involucran en muchos casos técnicas ancestrales; la investigación que se realiza en estos alimentos en la mayoría de los casos está enfocada a la identificación de los microorganismos involucrados en el proceso, con la finalidad de seleccionar a aquéllos que pudieran ser empleados como iniciadores.

Es posible que el estudio de la microbiota presente en un determinado ambiente por métodos microbiológicos que se basan en el aislamiento y cultivo en forma pura de un microorganismo sólo ha permitido la caracterización parcial de dichos ambientes, lo que da lugar en muchos casos a interpretaciones incompletas de los fenómenos o a planteamiento procesos microbianos existentes. La principal explicación que se ha establecido para este problema es que este tipo de aproximaciones metodológicas no permitían detectar la presencia de microorganismos que no son susceptibles de crecer bajo las condiciones de laboratorio, o aquellos de los que no se conocen las condiciones de crecimiento que les resulten adecuadas (Akkermans *et al.* 1994; Boone *et al.* 1995; Pot 1998).

Los microorganismos presentes en un determinado ambiente juegan un papel esencial en la degradación de substratos orgánicos e inorgánicos, dando como resultado una modificación de las características de dicho ambiente. En el caso de los alimentos fermentados tradicionales, las actividades microbianas de los microorganismos que participan en la fermentación sobre los substratos empleados como materia prima para su elaboración definen las características sensoriales finales de dichos productos (Wacher 1993; Fleet 1999).

La introducción de algunas técnicas de biología molecular para el estudio de diferentes ambientes ha permitido la identificación y análisis de los miembros de una comunidad microbiana (Bull *et al.* 1992; Akkermans *et al.* 1994). Con el desarrollo de la metodología de amplificación de un gen por la reacción en cadena de la polimerasa (PCR), del uso de sondas de genes específicos y de protocolos de extracción y purificación del ADN y ARN bacteriano total de buena

calidad a partir de muestras ambientales, ha sido posible la detección de microorganismos específicos a partir de muestras extraídas de ambientes tales como suelos, sedimentos, muestras de agua, contenidos intestinales y en algunos ambientes extremos en los que se han encontrado microorganismos no detectados previamente por métodos microbiológicos tradicionales, permitiendo así una caracterización más completa de la estructura microbiana presente (Steffan & Atlas, 1991; Josephson *et al.* 1993; Barns *et al.* 1994; O'Sullivan 1999).

Estas aproximaciones han sido también empleadas exitosamente para la detección de microorganismos patógenos como *Aerococcus*, varias cepas de *Escherichia coli*, *Listeria monocytogenes*, *Pseudomonas*, *Salmonella*, *Shigella*, varias especies de *Vibrio*, *Yersinia enterocolitica*, etc., a partir del análisis de muestras de ADN bacteriano total extraído de alimentos (Notermans *et al.* 1989; Kim *et al.* 1991; Wesley & Ashton 1991; Howard *et al.* 1992; Tsai y Olson 1992; Wiedmann *et al.* 1992; Czajka *et al.* 1993; Datta *et al.* 1993; Le Roy *et al.* 1993; Cano *et al.* 1993; Dickinson *et al.* 1995; Feng 1997; Pot 1998).

A pesar de la aplicación de estas aproximaciones moleculares al análisis de alimentos, existen pocos trabajos sobre el estudio de la microbiota presente en alimentos fermentados tradicionales. Estos reportes están enfocados principalmente al seguimiento de algunos iniciadores durante la fermentación de la leche (Ahmad & Stewart 1991); detección y caracterización de especies de *Lactococcus* en ambientes lácteos (Klijn *et al.* 1991a); estudios taxonómicos de *Leuconostoc* en salsas fermentadas (Collins *et al.* 1993); determinación de la diversidad de especies de lactobacilos en vino (Cuoto & Hogg, 1994); detección de diferentes especies de *Lactobacillus* a partir de quesos (Ehrmann, *et al.* 1994; Ward & Timmins 1999); análisis de la diversidad genética de microorganismos aislados de masas ácidas de sorgo y de maíz (Hamad *et al.* 1997; Agati *et al.* 1998) y detección de especies de *Lactobacillus* en productos alimenticios para la salud (Zhong *et al.* 1998).

El pozol (del náhuatl *pozolli*, espumoso) se ha definido como “una masa de maíz fermentado en forma de bolas de varias formas y tamaños, envueltas en hojas de plátano”. Durante la fermentación de la masa ocurren dos cambios esenciales, que son el desarrollo de un sabor ácido y el de un aroma característico que imparten al pozol sus propiedades refrescantes. Este alimento es consumido diluyendo la masa fermentada en agua en el momento en que va a ser ingerido, de manera que se toma en forma de bebida, al mismo tiempo refrescante y nutritiva. Se elaboran y consumen diferentes variedades de pozol por comunidades indígenas y mestizas de los estados mexicanos de Chiapas, Tabasco, Campeche y Yucatán y en menor forma se utiliza también como alimento en algunas regiones de Veracruz y Oaxaca (Wacher *et al.* 1993).

La fermentación del pozol es un proceso que se desarrolla en una fase semisólida. En este proceso está involucrada, por una parte, una microbiota compleja compuesta por diferentes grupos bacterianos, hongos filamentosos y levaduras y, por otra parte, un substrato complejo compuesto de carbohidratos, lípidos, proteínas y material fibroso. El análisis microbiológico del pozol ha permitido identificar la presencia de tres grandes grupos bacterianos (bacterias lácticas, bacterias mesófilas no lácticas y enterobacterias), además de algunos hongos filamentosos y levaduras (Wacher *et al.* 2000).

Las bacterias del pozol han sido muy poco estudiadas. Se ha inferido que las bacterias lácticas del pozol son los microorganismos responsables de la producción de ácido láctico durante la fermentación, el cual imparte su sabor ácido a dicho alimento (Nuraida *et al.* 1995; Wacher *et al.* 1993; 2000). La presencia de los restantes grupos bacterianos es principalmente el resultado del manejo no aséptico de la materia prima y de los utensilios empleados para la elaboración de este alimento (Wacher *et al.* 1993; 2000)

El principal problema que se ha tenido para entender el papel de las bacterias en el desarrollo de la fermentación del pozol es que se requiere de un proceso exhaustivo de aislamiento y purificación de los microorganismos presentes en la muestra a analizar. Su identificación por medio de pruebas

bioquímicas convencionales requiere necesariamente de la disponibilidad de las cepas a identificar en forma pura, lo que requiere una gran cantidad de tiempo invertido. A pesar de ello, los resultados obtenidos han permitido la identificación únicamente de algunas bacterias a nivel de género en muestras de pozol procedente de algunas localidades del estado de Chiapas.

Con el interés de conocer la microbiota específica responsable del desarrollo de las características de alimentos fermentados de gran importancia regional, entre ellos el pozol, en esta tesis se propuso emplear diversas estrategias metodológicas que permitan la detección de los microorganismos presentes en una muestra de este alimento, a partir del análisis de secuencias de genes de ARNr 16S amplificadas por PCR y clonadas a partir del ADN bacteriano total extraído de una muestra de pozol tipo indígena.

# 2

## GENERALIDADES.

---

### 2.1 IMPORTANCIA DE LA DIVERSIDAD MICROBIANA.

La diversidad microbiana puede ser considerada como un importante recurso biológico. Sin embargo, el estudio de la misma ha representado tradicionalmente un problema debido a que las técnicas tradicionales de aislamiento favorecen únicamente el crecimiento de escasos grupos microbianos.

Una gran cantidad de microorganismos presentes en muestras ambientales no puedan ser observados o cultivados, por lo que existe un importante vacío en el conocimiento de la diversidad microbiana, particularmente aparente en el caso de las bacterias. Las evidencias existentes sugieren que puede haber de 300000 a 1 000000 de especies bacterianas en nuestro planeta, de las cuales de 3100 a 5000 han sido descritas formalmente, en contraste con medio millón de especies de insectos descritas (Boone *et al.* 1995; Pace 1996; 1997; Dykhuizen, 1998). Las estimaciones sugieren que una muestra de un gramo de un suelo típico contiene alrededor de 1 billón de bacterias, pero solamente el 1% puede ser cultivado. De la misma forma, sólo fracciones similares de microorganismos han sido cultivadas a partir de agua dulce y ambientes oceánicos (Marshall 1993; Torsvik *et al.* 1990; 1993).

La aplicación de nuevas tecnologías (particularmente en el análisis de ácidos nucleicos, en informática, química analítica y en las estrategias de muestreo y caracterización de ambientes) al estudio de la diversidad microbiana, la colocan como una disciplina de vanguardia de la que se desprenden algunos conceptos:

- a. Los microorganismos son de crítica importancia para la sustentabilidad de la vida en nuestro planeta;
- b. La diversidad microbiana no descubierta representa un recurso potencial de nuevos genes y microorganismos de gran valor en biotecnología;
- c. Los patrones de diversidad microbiana pueden ser empleados para monitorear y predecir cambios ambientales; y
- d. Los microorganismos juegan un papel importante en la conservación y restauración biológica de organismos superiores y ambientes degradados.

En 1995, 43 científicos expertos en el estudio de diferentes ambientes, grupos microbianos y metodologías se reunieron para discutir e identificar necesidades de investigación e infraestructura en el estudio de la diversidad microbiana. Este grupo de trabajo identificó cuatro áreas generales de importancia para tener un mejor conocimiento, manejo y utilización del vasto recurso microbiano (Boone *et al.* 1995):

#### I. Entender los orígenes y los patrones de la diversidad microbiana.

- a. Seleccionar hábitats de interés que podrían ser estudiados intensivamente;
- b. Obtener un mejor entendimiento de los patrones espaciales y temporales de la diversidad microbiana, y de como el ambiente determina estos patrones;  
y
- c. Reforzar nuestros conocimientos sobre las tasas y rangos de la dispersión, evolución y extinción de especies microbianas.

#### II. Identificar y caracterizar la diversidad microbiana.

- a. Refinar las unidades taxonómicas que definen la diversidad microbiana y desarrollar la metodología adecuada para medir estas unidades;
- b. Descubrir nuevas formas microbianas, bioquímicas, líneas evolutivas y hábitats;

- c. Mejorar la metodología para caracterizar, aislar e identificar microorganismos no cultivables y miembros raros de comunidades microbianas;
- d. Desarrollar instrumentación sofisticada y programas de cómputo que provean nuevos y más rápidos métodos para caracterizar aislados y comunidades enteras; y
- e. Fomentar la investigación en taxonomía polifásica, particularmente la integración de información fenotípica, genética y ecológica.

### III. Conservar la diversidad microbiana.

- a. Desarrollar investigaciones que permitan la conservación de comunidades mixtas;
- b. Mejorar estrategias de conservación de cultivos, tales como conservación automatizada, miniaturización y condiciones óptimas para grupos microbianos difíciles de preservar;
- c. Mantener una red sólida y coordinada de colecciones de cultivos accesibles a nivel mundial; y
- d. Conservación de hábitats con especies microbianas raras o amenazadas.

### IV. Establecer necesidades de organización y necesidades de infraestructura.

- a. Debido al gran esfuerzo requerido, desarrollar medios para coordinar la investigación sobre la diversidad microbiana a un nivel internacional;
- b. Desarrollar bases electrónicas integrales que incluyan información sobre hábitats, localización geográfica, características fenotípicas, genotípicas, morfológicas y de las vías de acceso a ellas;
- c. Involucrar investigadores de otros campos, especialmente en el área de informática, óptica, electrónica, química, detección remota, macroecólogos y taxónomos;
- d. Recomendar esfuerzos centralizados donde son requeridas necesidades especializadas y para el desarrollo de bases de datos;
- e. Entrenar a nuevos científicos con conocimientos modernos en diversidad microbiológica, fisiología y taxonomía; y

- f. Fortalecer la conciencia pública sobre el papel vital que juega la diversidad microbiana en sus vidas y los alcances que este campo ofrece.

---

## **2.2 APROXIMACIONES METODOLÓGICAS AL ESTUDIO DE LA DIVERSIDAD MICROBIANA DE UN AMBIENTE.**

Las características fenotípicas clásicas de las bacterias consideradas en los esquemas de clasificación tradicional comprenden caracteres morfológicos, fisiológicos y bioquímicos. Individualmente, muchas de estas características han resultado ser de poca relevancia, pero en conjunto permiten tener información muy valiosa para reconocer diferentes taxones. La morfología de una bacteria incluye características tanto celulares (forma, endosporas, flagelos, cuerpos de inclusión, tinción de Gram), como coloniales (color, dimensión, forma). Las características fisiológicas y bioquímicas incluyen datos sobre crecimiento en diferentes temperaturas, valores de pH, concentraciones de sales o condiciones atmosféricas, crecimiento en presencia de diferentes sustancias tales como agentes antimicrobianos, y datos sobre la presencia de actividad de diferentes enzimas, capacidad de metabolizar distintos substratos, etc. (Busse *et al.* 1996; Pot 1998).

Los datos fenotípicos fueron los primeros en ser analizados por comparación numérica asistida por computación. En la década de los 50's, la taxonomía numérica surgió de manera paralela al desarrollo de computadoras y esto permitió algunos años después la comparación de cantidades muy importantes de datos fenotípicos de una gran diversidad de cepas analizadas. Las matrices de datos obtenidas muestran el grado de similitud entre cada par de cepas y su análisis por agrupamiento genera un dendrograma que refleja la consistencia fenotípica de un grupo particular de microorganismos. Tan pronto como resultó posible comparar los resultados del análisis de los grupos obtenidos con otras aproximaciones taxonómicas, el análisis de tan grandes cantidades de características fenotípicas resultaron efectivamente relevantes desde el punto de vista taxonómico, dando como resultado que las pruebas fenotípicas clásicas o

tradicionales sean utilizadas en esquemas de clasificación de la mayoría de los laboratorios microbiológicos del mundo y constituyen la base de la descripción formal de los taxones, de especies a subespecies, hasta los niveles de género y familia (Busse *et al.* 1996; Vandamme *et al.* 1996).

Sin embargo, en muchos casos la escasez de características fenotípicas distintivas o de la instrumentación requerida para su determinación, limita la utilidad de los análisis fenotípicos. Esta claro que el fenotipo de endosimbiontes y de bacterias no cultivables se encuentra más allá de los alcances de las técnicas presentes. Por otro lado, los microorganismos considerados como referencia y disponibles en las distintas bases de datos y colecciones minimizan la posibilidad de determinar la diversidad microbiana real presente en un ambiente, mediante la comparación contra una referencia. En estos casos, es necesario la aplicación de técnicas fenotípico - genotípicas poder identificar satisfactoriamente a una cepa (Boone 1995; Pace 1996; 1997).

La información genotípica se deriva del análisis de los ácidos nucleicos (ADN y ARN) presentes en la célula, mientras que la información fenotípica se deriva de las proteínas y sus funciones, de diferentes marcadores quimiotaxonómicos y de una amplia gama de otras características expresadas (figura 1). El número de moléculas que han sido aplicadas en estudios taxonómicos es muy grande y sus aplicaciones como marcadores son múltiples. Algunos de los métodos hasta ahora usados, tales como la determinación del porcentaje contenido de moles de G+C y los estudios de hibridación ADN-ADN, se han convertido en aproximaciones clásicas y han sido aplicados en análisis taxonómicos de prácticamente cualquier bacteria, otras, tales como la secuencia de aminoácidos, han sido aplicadas en un número limitado de taxones, debido a que son laboriosos, técnicamente demandantes, o bien porque sólo son aplicables a un taxón en particular. La selección de un método se basa en el nivel de información taxonómica que puede proporcionar y la complejidad técnica para poder aplicarlo. El nivel de resolución taxonómica de algunas de estas técnicas se ilustra en la figura 2, mientras que las categorías y caracteres aplicados se muestran en la tabla 1.

Este tipo de nuevas metodologías han tenido un impacto importante no sólo en la descripción de la estructura microbiana de un ambiente, sino que ha tenido aplicaciones que van desde la detección de patógenos de importancia clínica difícilmente detectables por otros métodos y de microorganismos asociados a alimentos. Han sido útiles en la detección y monitoreo de microorganismos manipulados genéticamente liberados al ambiente, en la búsqueda de nuevos microorganismos específicos y actividades, productos y procesos de importancia biotecnológica. Con el uso de estas técnicas ha sido posible monitorear con un alto grado de sensibilidad y precisión la presencia de diferentes microorganismos en una gran variedad de ambientes (Olson & Tsai 1992; Akkermans *et al.* 1994; Hirsh & Sigmund 1995; Feng 1997; Jansson & Prosser 1997).

---

### **2.3 MÉTODOS MOLECULARES PARA LA DETERMINACIÓN DE LA ESTRUCTURA MICROBIANA DE UNA MUESTRA AMBIENTAL.**

La detección e identificación de un microorganismo específico a partir de una muestra ambiental y el análisis de la microbiota total presente han tenido tradicionalmente tres problemas principales:

- a. La complejidad de la muestra
- b. El número de microorganismos a analizar presente en la muestra
- c. El conocimiento de los requerimientos nutricionales para su cultivo.

La detección de bacterias en diferentes ambientes requiere de métodos que eviten la necesidad del cultivo microbiano, pero que permitan su identificación con una alta sensibilidad y, preferiblemente con la capacidad de poder ser cuantitativos.

## INFORMACIÓN GENOTÍPICA

### ADN

#### ADN TOTAL

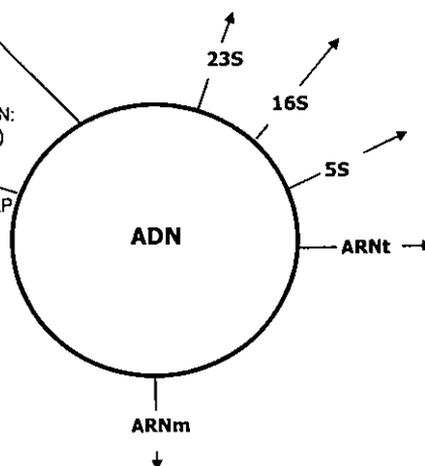
- % de moles de G + C
- Patrones de restricción (RFLP, PFGE)
- Tamaño del genoma
- Hibridaciones ADN:ADN

#### FRAGMENTOS DE ADN

- Análisis basados en PCR del ADN: (ribotipos, ARDRA, RAPD, AFLP)
- Sondas de ADN
- Secuenciación de ADN
- Patrones de restricción (RFLP, PFGE)

### ARN

- Secuenciación de bases del ADNc
- Perfiles de ARN de BPM



### PROTEÍNAS

- Patrones electroforéticos de proteínas totales o proteínas extracelulares (1D o 2D)
- Patrones enzimáticos (electroforesis de isoenzimas o enzimas multilocus)

### MARCADORES QUIMIOTAXONÓMICOS

- Ácidos grasos celulares
- Ácidos micólicos
- Lípidos polares
- Quinonas
- Poliaminas
- Componentes de la pared celular
- Exopolisacáridos

### CARACTERÍSTICAS EXPRESADAS

- Morfología
- Fisiología (Biolog, API)
- Enzimología (APIZYM)
- Serología (monoclonal, policlonal)

## INFORMACIÓN GENOTÍPICA

**Figura 1.** Representación general de varios componentes celulares y técnicas utilizadas para la identificación y clasificación bacteriana. ADN: ácido desoxirribonucleico; ARN: ácido ribonucleico; PCR: reacción en cadena de la polimerasa; RFLP: polimorfismo de fragmentos de restricción; PFGE: electroforesis de campo pulsado; ARDA: análisis de restricción de ADN amplificado; RAPD: amplificación aleatoria de ADN polimórfico; AFLP: polimorfismo de fragmentos amplificados; BPM: bajo peso molecular; 1D, 2D: una- y dos- dimensiones, respectivamente. Tomado de Vandame *et al.* 1996.

Técnica	Familia	Género	Especie	Cepa
Polimorfismo de fragmentos de restricción (RFLP)				
Análisis de fragmentos de restricción de baja frecuencia (LFRFA, PFGE)				
Ribotipos				
Amplificación de ADN (AFLP, AP-PCR, rep-PCR, DAF, RAPD, ARDRA)				
Tipificación de fagos y bacteriocinas				
Técnicas serológicas (monoclonal, policlonal)				
Zimogramas (polimorfismo de enzimas multilocus)				
Patrones electroforéticos de proteínas celulares totales				
Hibridación ADN-ADN				
ARNt - PCR				
Marcadores quimiotaxonómicos (poliaminas, quinonas)				
Tipificación de ácidos grasos celulares (FAME)				
Estructura de la pared celular				
Fenotipo (características morfológicas y coloniales, API, Biolog...)				
Secuenciación de ARNr				
Sondas de ADN				
Secuenciación de ADN				

**Figura 2.** Resolución taxonómica de las metodologías comúnmente utilizadas para identificación bacteriana. Las abreviaciones son definidas en la Figura 1. Tomado de Vandame *et al.* 1996 y Gürtler & Anker 1994.

**Tabla 1.**  
Categorías y caracteres aplicados en sistemática bacteriana.

Categorías	Ejemplos
Cultivables	Morfología colonial. Color de colonias Cuerpos fructíferos
Morfológicas	Micelio Morfología celular Tamaño celular Movilidad Tipo de flagelación Materiales de reserva Tinción de Gram Tinción ácida
Fisiológicas	Rango de temperatura Rango de pH Tolerancia a salinidad
Bioquímicas	Utilización de fuente de carbono Oxidación de carbohidratos Fermentación de carbohidratos Perfil de enzimas
Pruebas inhibitorias	Medios selectivos Antibióticos Colorantes
Serológicas	Aglutinación Inmunodifusión
Quimiotaxonómicas	Ácidos grasos Lípidos polares Ácidos micólicos Composición de lipopolisacáridos Electroforesis en gel de acrilamida de lipopolisacáridos Diaminoácidos de pared celular Azúcares celulares totales Pigmentos celulares Sistema de quinonas Contenido de poliaminas
Genotípicas	Electroforesis en gel de acrilamida de proteínas totales Proporción base del ADN (contenido de G + C) Amplificación aleatoria de ADN polimórfico (RAPD) Polimorfismo de fragmentos de restricción (RFLP) Electroforesis de campo pulsado (PFGE) de fragmentos de ADN
Filogenéticas	Sondas de ADN Hibridación ADN:ADN Hibridación ADN:ARN Secuencia de ARNr 16S Secuencia de ARNr 23S Secuencia de la subunidad $\beta$ de la ATP sintetasa Secuencia de GroEL (chaperonina) Secuenciación de otros genes

Tomado de Busse *et al.* 1996, Pot 1999.

El análisis de genes bacterianos específicos en muestras ambientales, por medio de métodos moleculares que no involucran el aislamiento de los microorganismos presentes en la muestra por técnicas microbiológicas tradicionales, ha permitido dar una nueva dimensión a nuestro entendimiento de la estructura de una población bacteriana. Estas metodologías han permitido, por primera vez, contar con una mejor aproximación para conocer la diversidad microbiana presente en un determinado ambiente.

Durante los pasados 20 años se han desarrollado una serie de estrategias metodológicas basadas en métodos moleculares para la detección e identificación rápida y directa de microorganismos en muestras ambientales, basadas en el análisis de secuencias específicas de ADN y ARN de diferentes microorganismos presentes en la muestra ambiental (figura 3). Estas estrategias típicamente están conformadas por las siguientes etapas:

- a. Extracción del ADN o ARN total de una muestra
- b. Amplificación *in vitro* de un fragmento particular del ADN/ARN por medio de PCR
- c. Clonación del producto de PCR
- d. Secuenciación y comparación de la secuencia obtenida en una base de datos, o bien, el desarrollo de una sonda específica y la hibridación de esta sonda contra los ácidos nucleicos extraídos de la muestra ambiental.

#### **Extracción de ácidos nucleicos de una muestra ambiental.**

Los métodos desarrollados para el análisis de los genes presentes en una muestra ambiental permiten detectar tanto secuencias específicas como la diversidad genética total presente. Estas técnicas toman como base metodológica la extracción del ADN o del ARN bacteriano a partir de muestras ambientales y fueron aplicadas en primer lugar para el análisis microbiano de suelos. En la actualidad existen una gran cantidad de variaciones a estas técnicas originales, de acuerdo al tipo de muestra y grupo microbiano a analizar.

La extracción y el análisis del ARN total presente en una muestra es una estrategia empleada en diferentes trabajos para detectar la actividad

microbiológica total presente, mientras que las técnicas que permiten detectar la expresión de ciertos genes son empleadas para medir la actividad de componentes específicos de la comunidad microbiana. Una de las ventajas del análisis del ARN total microbiano extraído de una muestra ambiental es su abundancia relativa. Existen tres clases principales de ARN: dos clases estables, el ARN de transferencia (ARNt), y el ARN ribosomal (ARNr), y una clase menos estable, el ARN mensajero (ARNm). En relación a la abundancia de los diferentes tipos de ARNs presentes en una célula, para *Escherichia coli* B con un tiempo de duplicación de 40 minutos, el ARN total representa alrededor del 21% del peso celular seco. De este ARN total, el ARNr constituye el 81%, el ARNt el 14% y el ARNm 5%. Dentro del ARN se distinguen tres tipos, el 5S, 16S y 23S, mientras que existen 60 tipos de ARNt y más de 4000 tipos de ARNm<sup>1</sup> (Pichard & Paul, 1994).

Sin embargo, a pesar de esta abundancia relativa, los métodos basados en la extracción del ARN total han sido menos utilizados que los métodos basados en el análisis del ADN, ya que el ARN es mucho menos estable que aquel y más fácilmente degradable debido a la presencia de ribonucleasas (RNasas). Aunque se han reportado metodologías que permiten extraer ARN exitosamente con muy buena calidad y rendimientos, una alternativa utilizada con mucha frecuencia cuando se desea analizar el ARNr (principalmente el 16S y 23S), es la amplificación por PCR de los genes correspondientes al ARNr con el uso de *primers* específicos (Wellington *et al.* 1994; Pichard & Paul 1996).

### **Amplificación, clonación y secuenciación de genes específicos de una muestra ambiental.**

La amplificación parcial o completa de genes por medio de PCR a partir del ADN microbiano total aislado de una muestra ambiental es una técnica empleada para la detección e identificación de bacterias en el ambiente. La reacción de amplificación de genes específicos por PCR puede ser realizada con el uso de *primers* específicos para un microorganismo particular presente en el ambiente y

---

<sup>1</sup> Para levaduras se estima un número de 4000 ARNm, mientras que para tejidos somáticos de eucariontes superiores se estima que el número de ARNm oscila entre 10,000 a 15,000 (Levin, 1997)

el análisis del producto de amplificación en un gel de agarosa permite determinar la presencia o ausencia del microorganismo de interés. Sin embargo, la eficiencia de esta técnica para determinar la diversidad total presente en una muestra ambiental está en función de la disponibilidad de *primers* para los microorganismos que se suponen pueden estar presentes en la muestra a analizar.

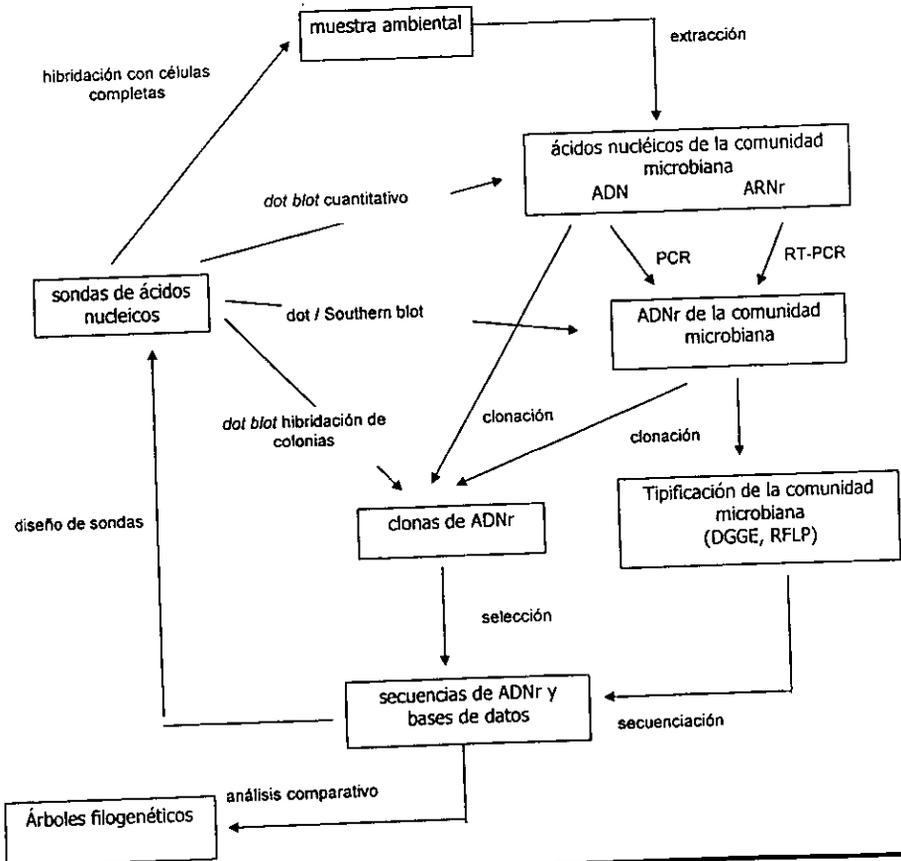
Desde hace más de 20 años, se ha reportado en la literatura el uso de *primers universales* o inespecíficos (tabla 2), para la amplificación por PCR de los genes de ARNr 16S presentes en una muestra de ADN extraída de un determinado ambiente (Weisburg *et al.* 1985; 1991; Devereux & Willis, 1995). El resultado de esta amplificación es una mezcla compleja de genes de ARN ribosomal. El producto resultante es entonces clonado para obtener entidades individuales de los genes amplificados para posteriormente secuenciar. El análisis de las secuencias en bases de datos de genes ribosomales, permite determinar la relación existente entre los microorganismos cuyos genes de ARNr 16S fueron amplificados, clonados y secuenciados y microorganismos reportados. Esta aproximación metodológica ha permitido no solo establecer la presencia de microorganismos, tanto cultivables como no cultivables, presentes en una muestra, sino que ha permitido determinar las relaciones filogenéticas entre los microorganismos presentes (Lane *et al.* 1985; Weisburg *et al.* 1985; 1991; Akkermans *et al.* 1994; Ludwig & Schleifer, 1994; Lee *et al.* 1996; Pace 1996).

---

## **2.4 EXTRACCIÓN DE ADN MICROBIANO TOTAL DE UNA MUESTRA AMBIENTAL.**

Los protocolos de extracción de ADN desarrollados para el análisis de sedimentos han sido modificados de acuerdo a los siguientes factores:

- a. Tipo y composición de la muestra ambiental
- b. Eficiencia en el proceso de lisis de la fracción bacteriana presente



**Figura 3.** Estrategias basadas en el análisis de ADNr/ARNr para la caracterización de comunidades microbianas sin el cultivo bacteriano. Las flechas indican la interconectividad existente en la utilidad de los métodos, materiales experimentales e información en el estudio de ecosistemas microbianos. RT-PCR, transcriptasa reversa para copiar ARN en ADN, seguido por PCR; DGGE, electroforesis desnaturizante en gel en gradiente; RFLP, polimorfismo de fragmentos de restricción. Tomado de Pace 1996.

Tabla 2. Primers empleados para la amplificación por PCR de genes de ARNr. 16S.

Primer <sup>a</sup>	Secuencia <sup>b</sup>	Sitio de unión <sup>c</sup>	Especificidad	Comentarios <sup>d</sup>
rD1	CcgaattctcgacaacAGAGTTTGATCTGGCTCAG	8-27	Mayoría de eubacterias	del, EcoRI, Sall
rD2	CcgaattctcgacaacAGAGTTTGATCTGGCTCAG	8-27	Enterobacterias y relacionadas	del, EcoRI, Sall
rD3	CcgaattctcgacaacAGAGTTTGATCTGGCTCAG	8-27	Borellia, espiroquetas	del, EcoRI, Sall
rD4	CcgaattctcgacaacAGAAATTTGATCTGGTTCCAG	9-27	Chlamydiae	del, BamHI
5'	GcggatccGAGTTTGATCTGGCTCAG	8-27	Eubacterias	del
PA	AGAGTTTGATCTGGCTCAG	1-21	Eucariontes	del, EcoRI, Sall
Primer A	CcgaattctcgacaAACCTGGTTGATCTGGCCAGT	8-27	Eubacterias	del, Sall, SacI
Primer A delantero	CCTCCGTTGATCTGCC	7-23	Empleado con metanógenos hipertermofílicos	Del
27d	AGAGTTTGATCTGGCTCAG	9-27	Mayoría de eubacterias	del, adaptado de un primer de secuenciación
rD1	CcggatccaagctAAGGAGGTGATCCAGCC	1541-1525	Mayoría de eubacterias	rev, HindIII, BamHI, XmaI
rP1	CcggatccaagctACGGTTACCTGTTACGACTT	1512-1492	Enterobacterias (y otras eubacterias)	rev, HindIII, BamHI, XmaI
rP2	CcggatccaagctACGGTACCTGTTACGACTT	1512-1492	Mayoría de eubacterias	rev, HindIII, BamHI, XmaI
rP3	CcggatccaagctACGGATACCTGTTACGACTT	1512-1492	Fusobacterias (y otras eubacterias)	rev, HindIII, BamHI, XmaI
3'	CcggatccAGAAAGGAGGTGATCCAGCC	1542-15254	Mayoría de eubacterias	del, BamHI
PH	AAGGAGGTGATCCAGCCGCA	1541-1522	Eubacterias	Rev
reverso	GGTACCTGTTACGACTT	1510-1492	Eubacterias	Rev
Primer B	CcggatccaagctTGATCTCTGCAGGTTACCTAC	1795-1772	Eucariontes	rev, SmaI, BamHI, HindIII, PstI
Primer B reverso	CcggatccaagctAAGGAGGTGATCCANCCRCA	1541-1518	Eubacterias	rev, SmaI, KpnI, HindIII
1492f	TACGGYATACCTTGTACGACTT	1539-1524	Empleado con metanógenos hipertermofílicos	rev
1525r	AAGGAGGTGWTCCARCC	1513-1492	Mayoría de eubacterias, Archaeobacterias	rev, primer de secuenciación adaptado para PCR
		1541-1525	Mayoría de eubacterias, Archaeobacterias	rev, primer de secuenciación adaptado para PCR

<sup>a</sup> Algunos primers no son designados de manera específica por los autores. Los nombres "Primer A" y "Primer B" son empleados por diferentes autores para referirse a primers diferentes

<sup>b</sup> Secuencias escritas en el sentido 5' → 3'. La secuencia en letras minúsculas indica una secuencia única que contiene sitios de reconocimiento de endonucleasas de restricción, las letras mayúsculas indican la región de hibridación con el ADNr. M= A y C; R= G y A; Y= C y T; W= A y T; N= C, A, T y G.

<sup>c</sup> Sitios de unión indicados con referencia en la numeración del ARNr. 16S de *Escherichia coli* para los primers para eubacterias y archaeobacterias, y en referencia a la numeración del ARNr. 16S de *Saccharomyces cerevisiae* para primers eucariontes.

<sup>d</sup> del = delantero, primer homólogo a la secuencia de ARNr con una extensión al extremo 3' del ARNr, rev = reverso, primer complementario al ARNr con extensión hacia el extremo 5' del ARNr. Las enzimas de restricción que corran en la región agregada se indican en cada caso. Para el primer B, el sitio de PstI está localizado en la región de hibridación con el ADNr.

Tomado de Devereux & Willis, 1995.

c. Integridad del ADN recuperado

d. Pureza del ADN

La importancia relativa de cada uno de estos aspectos varía de acuerdo al tipo de estudio a realizar y de la aplicación del ADN recuperado. En la tabla 3 se puntualizan los aspectos más importantes a considerar en cada una de las etapas que conforman un protocolo típico de extracción de ADN a partir de una muestra ambiental (Trevors 1992; Wheeler & Stahl 1996).

La extracción del ADN de muestras ambientales ha sido optimizada particularmente para el caso de bacterias y existen pocos estudios para eucariontes (Nazar *et al.*, 1996). El ADN total puede ser obtenido por dos estrategias principales (figura 4): (a) el método indirecto de extracción, que involucra la separación de las células presentes en la muestra con la posterior lisis celular y extracción y purificación del ADN; y (b) el método de extracción directa, que se basa en la lisis directa de las células en la muestra, con la posterior extracción y purificación del ADN. Los protocolos desarrollados para la extracción de ADN de diferentes muestras ambientales por métodos directos e indirectos han sido optimizados para obtener ADN de mejor calidad y con rendimientos cada vez mayores. En la tabla 4 se puntualizan las principales ventajas y desventajas de cada una de estas técnicas.

#### **Extracción indirecta del ADN de una muestra.**

Existen diferentes métodos que permiten la separación de células microbianas de una muestra ambiental. Tradicionalmente, la homogeneización de la muestra en una solución amortiguadora usando agitación mecánica ha sido la aproximación más común y se puede realizar tanto por homogeneización como por agitación rápida. La sonicación de la muestra ha sido empleada también con gran éxito para el estudio de suelos y sedimentos, pero su eficiencia depende del tipo de muestra a analizar, de la duración y de la amplitud de la sonicación. Sin embargo, la sonicación prolongada tiene el riesgo potencial de lisar las células y por ello las condiciones ideales de tratamiento varían de protocolo en protocolo y de acuerdo al tipo de muestra.

Una vez que las células han sido separadas, las partículas del substrato son sedimentadas por centrifugación a baja velocidad. El sobrenadante que contiene la

fracción celular separada es mantenido en hielo o en refrigeración, mientras que el precipitado es reprocesado algunas veces más para recuperar la mayor cantidad de células presentes antes de pasar a la etapa de lisis celular y purificación del ADN.

La eficiencia de extracción del ADN obtenido por este método está en función del éxito de la separación de las células bacterianas de la muestra. Aunque esta metodología no ha sido tan extensamente estudiada en comparación con la estrategia de extracción directa, ha sido aplicada para el análisis de grupos bacterianos específicos y en estudios de la diversidad microbiana total presente en muestras de suelos y sedimentos. Las eficiencias de recuperación de células para algunos de los trabajos que emplean esta metodología indirecta van de 7.3 a 23% incrementándose este valor al someter la muestra a extracciones repetidas y ciclos de lisis adicionales, lo que resulta en rendimientos variables de ADN obtenido (Holben *et al.* 1988; Steffan *et al.* 1988; Torsvik *et al.* 1990; Pillai *et al.* 1991; Trevors 1992; Breen *et al.* 1995).

#### **Extracción directa del ADN de una muestra.**

La extracción directa de ADN no requiere del aislamiento de las células bacterianas y se evita el proceso de homogeneización-centrifugación diferencial de la muestra. Los procedimientos para la extracción directa de ADN fueron utilizados en un principio para la obtención de ADN bacteriano y fúngico de suelos con la obtención de rendimientos variables (aunque estos valores dependen de los aspectos señalados en la tabla 3).

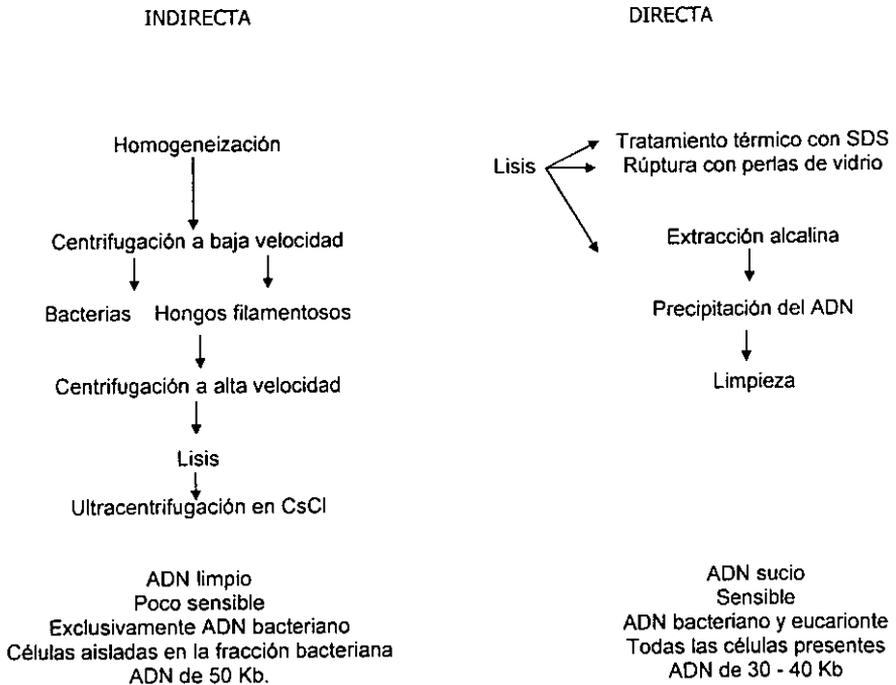
La extracción directa de ADN requiere la suspensión inicial de la muestra en una solución amortiguadora y el tratamiento físico, químico, enzimático o mecánico (o una combinación de ellos) para asegurar la lisis de los microorganismos presentes, con la posterior extracción y purificación del ADN. La eficiencia del proceso de lisis celular es particularmente importante, y se reportan valores que van del 20 al 90% (Porteus *et al.* 1994; Zhou *et al.* 1996). Sin embargo, la extracción directa en muchas ocasiones está acompañada de la contaminación del ADN por partículas presentes en la muestra, lo que obliga a tomar una serie de consideraciones durante el proceso de lisis y purificación del ADN (Trevors 1992).

Tabla 3.

**Factores a considerar en cada etapa de extracción de ADN microbiano a partir de muestras ambientales.**

1. Relación de la masa o volumen de la muestra: solución de extracción	Selección de solución de extracción y pH más adecuados de acuerdo al objetivo del estudio a realizar
2. Separación eficiente de células de las partículas de la muestra por homogeneización.	Considerar el tipo de agitación / homogeneización a emplear
3. Evaluación de la presencia de componentes de la muestra que pudieran contaminar el ADN liberado o reaccionar químicamente (método de extracción directo)	Adición de sustancias que reaccionen específicamente con estos componentes antes de realizar la lisis
3 a. Separación diferencial eficiente de células procariontes de eucariontes y partículas del sustrato (método de extracción indirecto)	Centrifugación diferencial de acuerdo a la densidad células/sustrato, que garantice la obtención de una fracción bacteriana libre de eucariontes y restos orgánicos de la muestra
4. Tratamiento lítico empleado	<b>Método de extracción directo:</b> Selección del método de lisis (enzimática, química, física) que garantice la degradación de la estructura celular de bacterias Gram(+) y microorganismos productores de polisacáridos extracelulares, considerando también la complejidad de la muestra. <b>Método de extracción indirecto:</b> Tratamiento lítico enzimático (lisozima/proteinasa K) en combinación con un tratamiento químico-térmico (uso de detergentes tales como SDS o sarcosil, con ciclos de calentamiento/enfriamiento) por períodos de incubación prolongados para asegurar la lisis de bacterias Gram (+) y microorganismos productores de polisacáridos extracelulares.
5. Centrifugación para remover restos celulares y partículas de la muestra. Re-extracción del ADN del pellet resultante (método de extracción directa)	Generalmente velocidad de centrifugación baja (5 000 × g durante 5 - 10 min)
6. Purificación del ADN: eliminación de ARN y proteínas	Extracción con fenol-cloroformo, tratamiento con ribonucleasas, ultracentrifugación en gradiente de CsCl <sub>2</sub> -EtBr, columnas de hidroxipatita.
7. Precipitación del ADN	Uso de 2 volúmenes de etanol e incubación a -20°C y lavado con etanol al 70% para eliminar sales
8. Resuspensión del ADN	Uso de agua destilada, bidestilada o solución amortiguadora de TRIS-EDTA de acuerdo al tipo de análisis a realizar.

Tomado de Trevors 1992; Wellington *et al.* 1994; Wheeler & Sthal 1996.



**Figura 4.**

Sumario de las dos aproximaciones metodológicas para la extracción de ADN de muestras ambientales. Tomado de Wellington *et al.* 1994.

**Tabla 4.**

Comparación de los protocolos de extracción directo e indirecto de ADN.

<b>Extracción indirecta (aislamiento de bacterias)</b>	
<b>Ventajas</b>	<b>Desventajas</b>
1. Asegura ADN de origen bacteriano	1. Las células bacterianas se pueden adherir estrechamente a las partículas de sedimentos y pueden no ser separadas
2. Evita el contacto directo del ADN con el sedimento	2. Procedimiento laborioso
3. No existe contaminación con sustancias o compuestos de la muestra	3. Bajos rendimientos en comparación con la extracción directa
<b>Extracción directa</b>	
<b>Ventajas</b>	<b>Desventajas</b>
1. Recuperación representativa	1. Recuperación de ADN no bacteriano y extracelular
2. Rápido, menos laborioso	2. El ADN entra en contacto directo con sedimentos
3. Altos rendimientos	3. Contaminación con diferentes sustancias presentes en la muestra
	4. ADN de bajo peso molecular

Tomado de Wheeler &amp; Stahl 1996.

A pesar de que la principal aplicación de las metodologías de extracción directa del ADN bacteriano de una muestra estuvo enfocada en un principio al análisis de suelos, estos procedimientos han sido empleados con éxito para el análisis de diferentes ambientes con eficiencias de recuperación de aproximadamente el 80% para comunidades microbianas autotróficas y heterotróficas (Paul 1996), muestras de agua dulce (Pickup *et al.* 1995), aguas residuales y lodos (Smalla 1996), películas microbianas (Bateson & Ward 1996), sedimentos acuáticos (Wheeler & Stahl, 1996), muestras de rizoplaneo (Bramwell *et al.* 1996a), microorganismos del filoplaneo (Bramwell *et al.* 1996b), nódulos actinorrízicos (Ramírez-Saad *et al.* 1996).

Antes de realizar un estudio basado en el análisis de los genes microbianos presentes en una muestra ambiental a partir de una muestra de ADN total, es importante considerar que los ácidos nucleicos presentes en la solución provienen tanto de los microorganismos viables presentes en ese ambiente como de microorganismos muertos cuyo ADN no está degradado. El origen del ADN libre ha sido investigado principalmente en ambientes tales como aguas oceánicas y continentales y en algunos sedimentos y suelos. En estos casos se ha determinado que el ADN es liberado al ambiente continuamente como parte de un proceso natural de descomposición de microorganismos muertos. Las células vivas pueden también liberar ADN durante su ciclo de vida, probablemente como parte de procesos de transferencia genética, tales como la transformación. La proporción de ADN libre en el ambiente puede llegar a ser alta, así por ejemplo, el uso de timina- $H^3$  para marcar bacterioplancton heterotrófico permitió detectar ADN disuelto en el ambiente tanto por lisis como por excreción. Su concentración fue aproximadamente un 4% de la concentración total presente en el ambiente por día, detectándose un promedio de 10 - 13  $\mu g$  de ADN disuelto / por litro de agua en un lapso de 6.5 horas. El origen del ADN libre en algunas muestras de suelo ha sido menos estudiado, pero algunos estudios han demostrado la presencia de valores de aproximadamente 5  $\mu g$  de ADN / g de suelo, con una tasa de aparición similar a la observada para ambientes acuáticos. Para otro tipo de muestras ambientales no existen estudios realizados que permitan estimar la cantidad de ADN libre presente (Wellington, *et al.* 1994).

Las metodologías de extracción de ADN de una muestra han sido aplicadas también al análisis del contenido microbiológico específico de muestras de alimentos. Ehrmann *et al.* (1994) emplearon un procedimiento de lisis directa para obtener ADN de bacterias lácticas susceptible de ser empleado en experimentos de hibridación, basado en la homogeneización exhaustiva con ultraturrax y posterior lisis por tratamiento con SDS y ciclos de calentamiento - enfriamiento. Dickinson, *et al.* (1995) describen dos procedimientos para la extracción directa de alimentos: un procedimiento de lisis rápida, empleando proteinasa K y precipitación de las partículas solubilizadas del alimento y del ADN con isopropanol y un segundo método, que asegura una lisis celular eficiente empleando tolueno y mutanolisina. Después de la remoción de las proteínas con solventes orgánicos, el ADN es precipitado con isopropanol. En este caso, ambos métodos permitieron obtener ADN susceptible de ser amplificado por PCR a partir de muestras de queso, col fermentada y carne de pollo y demostraron ser métodos eficientes para la detección específica de *Yersinia enterocolitica*, *Aerococcus viridans* y *Listeria monocytogenes*. Agersborg *et al.* (1997) desarrollaron un procedimiento para la extracción del ADN de *Listeria monocytogenes* inoculada en alimentos de origen marino, mediante un método basado en el tratamiento con lisozima, proteinasa K y Triton X-100. Este permitió obtener ADN de buena calidad, susceptible de ser amplificado por PCR. Sin embargo, no existen reportes que mencionen la aplicación de estas metodologías como un primer paso en el estudio de la determinación de la composición microbiana total presente en una muestra de alimento.

---

## **2.5 SECUENCIAS PARCIALES Y COMPLETAS DE ARN<sub>r</sub> 16S. SU USO E IMPLICACIONES EN ESTUDIOS DE ECOLOGÍA MICROBIANA MOLECULAR.**

Lo que empezó hace más de 20 años con la determinación de una serie de secuencias parciales de un grupo limitado de organismos procariontes y eucariontes llegó a revolucionar desde entonces nuestro entendimiento sobre el origen de la vida en la Tierra, permitiendo establecer un sistema de clasificación basado en la definición de los linajes Bacteria, Archaea y Eucariota a partir de todas las formas

de vida existentes, y nos ha permitido comprobar la existencia de una desconocida y abundante diversidad microbiana en el ambiente (Akkermans *et al* 1994; Pace 1996; Hattori *et al.* 1997).

Los resultados obtenidos del estudio de los ARNr 16S no solo fueron causa del resurgimiento del interés por la sistemática, sino que también generaron la colaboración entre diferentes disciplinas que tradicionalmente no habían sido consideradas en estudios microbiológicos y el fortalecimiento de otras materias, como la taxonomía y la ecología microbianas. Gracias a esto, los criterios empleados en sistemática y taxonomía (muchas veces subjetivos), con los que se juzgaban diferentes hechos y que prevalecieron durante décadas, fueron reemplazados por criterios sólidos y objetivos basados en pruebas experimentales. De este modo, las diferentes técnicas de biología molecular (la amplificación por PCR, clonación, secuenciación, hibridación y el uso de sondas, por ejemplo) fueron aceptadas como metodologías de gran importancia en el contexto de estas disciplinas.

La determinación de las relaciones filogenéticas de los microorganismos mediante el empleo de tales técnicas fue uno de los hechos más importantes en la microbiología y ecología microbiana a finales de la década de los 80's. Sin embargo, la gran mayoría de los organismos procariontes presentes en el ambiente y el papel que tienen en un ecosistema concreto permanecen aún sin identificar. Como consecuencia de esto, hasta que sea identificada la totalidad de la diversidad de Bacteria, Archaea, hongos filamentosos, levaduras, algas y protozoarios no se podrán establecer las bases de una discusión sobre la importancia que tienen los microorganismos en el contexto de la importancia de la diversidad microbiana, los aspectos de conservación de la biodiversidad, la preservación de pozas genéticas y el papel que tienen estos microorganismos con otros organismos más evolucionados para mantener la estabilidad de un ecosistema, incluyendo también a aquellos microorganismos no cultivables (Boone, *et al.* 1995; Akkermans, *et al.* 1994; Pace, 1996; Busse, *et al.* 1996; Vandamme *et al.* 1996; Hattori *et al* 1997).

### **La búsqueda de un marcador óptimo para estudios filogenéticos.**

El potencial del análisis comparativo de la estructura primaria de una macromolécula para análisis filogenéticos de especies procariontes, fue vislumbrada desde la década de los 60's, cuando se estableció la relación entre una tasa evolutiva y el cambio en la secuencia de una molécula. Basados solamente en la secuencia de unas cuantas proteínas, tales como el citocromo C y las hemoglobinas, se estableció que las moléculas biológicas se pueden ubicar en tres categorías de información. Las moléculas más importantes son las *moléculas semantofóricas* o *semántidas*, que son los genes o sus transcritos (moléculas primarias: ADN, secundarias: ARN y terciarias: proteínas). La información genotípica, como las secuencias de ADN y ARN, es el registro histórico de la evolución y la determinación de su estructura primaria proporciona una importante aproximación para poder medir las relaciones evolutivas, no solamente entre las moléculas secuenciadas sino en relación también con la evolución del organismo del cual fue generada la secuencia.

Las *moléculas episemánticas* son sintetizadas bajo el control de las semántidas terciarias, y en ellas se incluyen compuestos tales como el ATP, carotenoides y los marcadores quimiotaxonómicos (Stackebrandt & Rainey 1995). Estas moléculas no son consideradas como útiles para poder derivar a partir de ellas conclusiones evolutivas ya que enzimas con diferentes estructuras primarias pueden permitir la síntesis de moléculas episemánticas idénticas o similares en diferentes organismos en la proporción en que los sitios activos de estas enzimas sean similares. Las moléculas episemánticas no son de valor filogenético, pero sí son de gran valor en la caracterización de los taxones definidos por las semántidas. Las *moléculas asemántidas*, tales como las vitaminas, diferentes iones, etc., no expresan ningún tipo de información del organismo que las contiene, y consecuentemente no son de valor filogenético (aunque pueden eventualmente ser empleadas para la descripción general de una cepa).

En este momento resulta imposible emplear toda la secuencia primaria completa del ADN de un organismo con fines comparativos, pero sí es posible seleccionar de entre las secuencias de genes de proteínas, secuencias de aminoácidos y secuencias

de las diferentes especies de ARN's. El número y diferencias en los genes que codifican para el ARNr y las proteínas reflejan sus filogenias y, consecuentemente, permiten el reconocimiento de pares o grupos de moléculas que se han derivado de ancestros comunes, así como la posibilidad de ordenamiento del organismo secuenciado de acuerdo a su historia evolutiva. Para tener la posibilidad de medir exactamente el tiempo transcurrido desde la divergencia del ancestro más reciente, los cambios deben de ser selectivamente neutrales e independientes del fenotipo y de la presión selectiva dictada por esta función. De lo anterior se puede determinar que el cronómetro molecular más confiable puede ser representado por una parte con menor grado de compromiso funcional del genoma que evoluciona aleatoriamente.

Considerando que hay un número limitado de posibilidades para una posición particular de una secuencia (los cuatro nucleótidos), los cambios neutrales se fijan rápidamente y solo es posible medir eventos evolutivos recientes. La realidad, sin embargo, es diferente, ya que todos los genes secuenciados (y especies de ARN), constituyen una estructura *quimérica* de regiones con un grado variable de conservación. La tasa y la naturaleza de la sustitución de bases difiere en posiciones individuales de las secuencias, dependiendo de su importancia para la función de la molécula o del correspondiente producto del gen, pero no son sujetas a cambios neutrales en grandes cambios revolucionarios. En su lugar, son acumulados múltiples cambios en regiones variables; éstas partes no reflejan el número acumulado de eventos mutacionales y pueden oscurecer la historia de la molécula simulando identidades falsas. Las moléculas están compuestas por regiones con mayor o menor grado de compromiso funcional, por lo que el grado de diferencia no puede ser directamente asociado con el tiempo acumulado (Ludwig & Schleifer 1994; Schleifer & Ludwig 1994; Stackebrandt & Rainey 1995; Pace 1996).

### **Las especies de ARNr como marcadores filogenéticos.**

En la actualidad, las especies de ARNr pueden ser consideradas como los cronómetros moleculares más útiles, ya que reúnen todas las características que definen un marcador filogenético: se encuentran ampliamente distribuidas en la naturaleza, por lo que tienen la posibilidad de cubrir diferentes taxones y son

homólogas, al derivar de un ancestro común. Su función ha permanecido constante durante millones de años, y como la función es determinada por la secuencia, un modo comparable de evolución de los organismos más lejanamente relacionados puede ser esperado solamente de moléculas funcionalmente equivalentes.

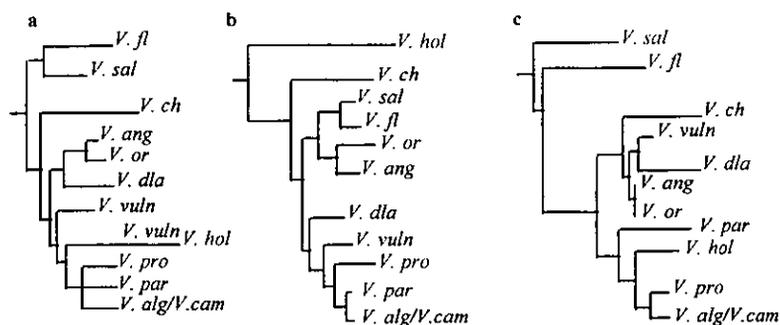
Los genes de ARN ribosomales parecen ser genéticamente estables. Las especies ribosomales 16S y 23S son moléculas razonablemente grandes y este factor es importante en el momento de analizar un número elevado de regiones o posiciones cambiando independientemente de un número cada vez más elevado de niveles filogenéticos. Las estructuras primarias de estos marcadores moleculares contienen dominios con evolución independiente, permitiendo determinar la relevancia de la identidad o diferencia de regiones de secuencias locales para obtener conclusiones filogenéticas. Por otro lado, diferentes regiones de estos ARNr varían a diferentes tasas: sobre un rango de cambios lentos la región permanece constante en su composición en relación a la extensión del tiempo evolutivo; y sobre un rango de cambios rápidos pueden ser generalmente vistas como diferencias en composición de una determinada región entre dos cepas de la misma especie (Ludwig & Schleifer 1994; Schleifer & Ludwig 1994; Stackebrandt & Goebel 1994; Stackebrandt & Rainey 1995, Busse *et al.* 1996).

Algunas proteínas que han sido empleadas como marcadores para estudios filogenéticos (tales como los citocromos, las ferredoxinas y las zurinas, ATPasas, factores *tuf* y proteínas de choque térmico) son también semántidas adecuadas, pero no son tan útiles como los son los ARNr. Las razones de esto son: (i) la degeneración del código genético previene la formación de regiones conservadas necesarias para el desarrollo de *primers* para la amplificación por PCR que podrían cubrir el rango total de la diversidad de procariontes; (ii) muchas de estas proteínas que fueron empleadas en el pasado no se encuentran ampliamente distribuidas, lo que restringe el análisis al establecimiento de árboles filogenéticos parciales, y (iii) la existencia de familias de proteínas, que muchas veces hacen difícil el decidir si una proteína es verdaderamente ortóloga o paróloga (Schleifer & Ludwig 1994; Stackebrandt & Rainey 1995).

### **El uso de secuencias parciales y la posibilidad de determinar líneas principales de descendencia.**

Resulta obvio que la posición filogenética de un taxón se debe basar en el análisis de secuencias completas de los genes de ARNr (ADNr). Las bases de datos existentes en la actualidad contienen un gran número de secuencias completas ya que el análisis de secuencias de ADNr ha llegado a convertirse casi en una técnica de rutina, debido al incremento en el número de secuenciadores automáticos en diferentes laboratorios y también debido a que el costo y el tiempo involucrados en el análisis de una secuencia completa es en algunos casos mucho mayor que el empleado para un análisis parcial.

El uso de secuencias parciales de ADNr presenta una serie de problemas cuando se pretenden realizar estudios filogenéticos. El siguiente ejemplo, sobre el género *Vibrio*, puede demostrar algunos de estos problemas (figura 5). El análisis basado en secuencias completas permite observar que *V. fischeri* y *V. salmonicida* representan el par de microorganismos con una ramificación más profunda, mientras que *V. hollisae* ramifica entre *V. proteolyticus* y sus relacionados y de *V. vulnificus*. Cuando el análisis se restringe al uso de 900 nucleótidos terminales del extremo 3', *V. hollisae* aparece como la ramificación más profunda, mientras que *V. fischeri* y *V. salmonicida* forman un grupo junto con *V. angillarum* y *V. ordalii*. Es posible observar otra topología cuando se emplean para el análisis las secuencias de 450 nucleótidos, en donde a pesar de que se observa un patrón de ramificación más parecido al obtenido en el análisis de secuencias completas, se puede observar una diferencia significativa en los puntos de ramificación de todas las especies analizadas. La variación en las topologías de los árboles filogenéticos se debe a las diferencias significativas en los valores de similitud determinados para ciertas regiones de secuencias de un grupo de microorganismos relacionados entre sí (tabla 5). La posición de las regiones hipervariables en la estructura primaria de los ADNr difiere de taxón en taxón y necesita ser determinada individualmente para poder ser empleada en el establecimiento de relaciones.



**Figura 5.**

Comparación de los dendrogramas filogenéticos construidos a partir de diferente contenido de información de secuencias del ADNr 16S: incapacidad de secuencias parciales para determinar la posición intergenérica de especies del género *Vibrio*. a. Análisis basado en la secuencia completa. b. Análisis basado en una región de la posición 501 a 1450. c. Análisis basado en la utilización de 500 nucleótidos del extremo 5'. *V. alg* = *V. algillarum*, *V. pro.* = *V. proteolyticus*, *V. fl* = *V. fisheri*, *V. sal* = *V. salmonicida*, *V. hol* = *V. hollisae*, *V. vuln* = *V. vulnificus*. El resto de las abreviaturas no son definidas por los autores. Tomado de Stackebrandt & Rainey 1995.

Tabla 5.

Grado de similitud entre diferentes regiones hipervariables de ADNr 16S.

a. *Streptomyces ambofaciens* y representativos de diferentes líneas principales de descendencia del dominio Bacteria.

Organismos	Región				total	contenido de G+C de ADNr. 16S
	61-108	170-227	433-490	1009-1044		
<i>S. ambofaciens/S. lividans</i>	100	83.3	100	100	93.9	59.1/58.8
<i>S. ambofaciens/A. globiformis</i>	68.4	62.1	97.8	94.4	53.1	59.1/57.0
<i>S. ambofaciens/B. subtilis</i>	55.6	50	72.3	64.7	56.2	59.1/55.1
<i>S. ambofaciens/E. coli</i>	60.5	54.5	76.1	69.7	50.0	59.1/54.5
<i>S. ambofaciens/Bt. vulgaris</i>	43.2	49.2	67.4	58.3	31.2	59.151.2
<i>S. ambofaciens/P. staleyii</i>	40.5	45.5	60.0	48.5	34.4	59.1/54.4
<i>S. ambofaciens/T. maritima</i>	65.7	52.6	76.6	60.0	37.5	59.1/63.6

b. *Escherichia coli* y representativos de diferentes líneas principales de descendencia del dominio Bacteria.

Organismos	Región				total	contenido de G+C de ADNr. 16S
	61-108	170-227	433-490	1009-1044		
<i>E. coli/S. lividans</i>	60.5	54.5	76.1	50.0	69.7	78.4
<i>E. coli/A. globiformis</i>	60.5	54.5	76.1	56.9	69.7	78.4
<i>E. coli/B. subtilis</i>	62.5	63.6	73.9	52.8	70.7	76.9
<i>E. coli/E. coli</i>	59.5	57.7	53.8	5.3	62.5	77.8
<i>E. coli/Bt. vulgaris</i>	60.9	50.0	54.2	34.4	62.5	71.3
<i>E. coli/P. staleyii</i>	43.2	57.1	54.0	44.4	39.4	72.3
<i>E. coli/T. maritima</i>	48.6	60.0	51.6	50.0	57.6	74.2

S. - *Streptomyces*; A. - *Arthobacter*; B. - *Bacillus*; E. - *Escherichia*; Bt. - *Bacteroides*; P. - *Pirellula*; T. - *Thermotoga*La secuencia de *Methanosarcina barkeri*, representante del dominio Archaea, fue empleada como referencia externa.

Las columnas en negritas indican la mayor variación.

Tomado de Stackebrandt &amp; Rainey 1995.

A pesar de esto, existen algunos trabajos que han empleado secuencias parciales de ADNr para establecer filogenias (Lane *et al.* 1985; Böttger *et al.* 1989; Arahai *et al.* 1996; Rheims *et al.* 1996; Felske *et al.* 1997; Zhou *et al.* 1997; Burrel *et al.* 1998; Felske *et al.* 1998; Nüsslein & Tiedje 1998). El uso de una secuencia parcial de al menos 500 bases de una clona de ADNr procedente de muestras ambientales puede permitir la asignación de la línea principal de descendencia. Esta aproximación resulta útil en la determinación del grado de diversidad en una biblioteca de clonas y en el diseño de sondas oligonucleotídicas para determinar la abundancia de clonas positivas en una biblioteca genómica. Sin embargo, en el contexto del análisis de una biodiversidad global, es motivo de discusión sobre una estrategia que involucre el análisis de secuencias de la misma región del ADNr 16S y el mínimo de nucleótidos requeridos para tal fin. Debido a que muchos investigadores han incluido en su análisis el extremo 5' de esta molécula, se recomienda la secuenciación de 300 a 500 nucleótidos del extremo 5' (Stackebrandt & Rainey 1995).

#### **Determinación de la posición filogenética de una clona a partir de una secuencia parcial de ADNr.**

Una vez obtenida la secuencia parcial de una clona, el siguiente paso es la determinación de su identidad o del origen de este fragmento clonado. La identificación del microorganismo a partir del cual esta clona se origina, permite asumir que ese microorganismo estuvo presente en el ambiente del cual se analiza la diversidad microbiana.

El primer paso en el proceso de identificación de la clona es alinear la secuencia contra un grupo de secuencias, en el que generalmente se incluyen secuencias representativas de las líneas principales de descendencia entre bacterias. Cuando éste se realiza por un procedimiento manual, es posible obtener una idea de la identidad de la clona considerando la línea principal con la que presente el mejor alineamiento.

Una vez que las secuencias han sido alineadas, se puede construir una matriz de similitud. Esto permite observar un valor de similitud entre la clona y el grupo

filogenético principal. La tabla 6, muestra los valores de similitud de secuencias parciales de clonas ambientales desconocidas con secuencias conocidas y consideradas como referencia para determinar su identidad. La construcción de un árbol filogenético a partir ésta matriz permite a su vez ubicar la posición filogenética de la clona ambiental con referencia a las secuencias conocidas empleadas en el análisis. Sin embargo, se debe de hacer énfasis en que cuando el punto de ramificación es cercano a la raíz de un *phylum*, la posición puede cambiar de un *phylum* a otro cuando se tiene la secuencia completa y/o se emplea una selección de organismos de referencia distinta.

Cuando se determina el grupo principal al que la secuencia desconocida pertenece, ésta puede ser entonces comparada con aquéllas disponibles en dicho grupo. Esto permite determinar la posición exacta de la secuencia de la clona en uno de varios niveles taxonómicos que van desde el nivel familia al nivel de especie. Es aquí en donde se requiere el conocimiento detallado de las características del grupo al que pertenece la clona identificada, ya que el no incluir las secuencias de todos los representantes de este grupo puede causar el establecimiento de conclusiones erróneas como el que la clona representa un taxón no secuenciado previamente o no conocido.

La carencia de secuencias disponibles para muchos géneros y especies bacterianos representa un obstáculo importante para la determinación adecuada de la identidad de una clona secuenciada. En la actualidad existen una gran cantidad de secuencias completas y parciales que se encuentran disponibles para el análisis de comparación, por lo que la mayoría de las especies pueden ser ubicadas filogenéticamente en cada uno de los *phyla* y reinos existentes. Por ejemplo, en la Base de Datos del Proyecto Ribosomal (RDP por sus siglas en inglés, <http://www.cme.msu.edu/RDP/html/index.html>) se encuentran depositadas 16,277<sup>2</sup> secuencias de ARNr 16S (Maidak *et al.* 2000). Sin embargo, para algunos géneros no existen datos de secuencias disponibles, por lo que el descubrimiento de nuevas líneas de descendencia del grupo principal debe ser visto con ciertas precauciones.

---

<sup>2</sup> Número de secuencias reportado para el mes de Octubre de 2000 en la base de datos RDP.

**Tabla 6.**

Valores de similitud de ADNr 16S obtenidos de secuencias de clonas ambientales y su parentesco cultivable más cercano. Los análisis se basaron en los 500 nucleótidos terminales del extremo 5'.

	CLONAS				
	H7	H101	H3	H79	Mc100
<i>Agrobacterium tumefaciens</i>	<b>78.5</b>	70.2	76.8	74.0	60.8
<i>Rhodopseudomonas palustris</i>	77.9	71.3	74.6	70.2	61.7
<i>Rhodocyclus palustris</i>	75.9	<b>85.6</b>	73.5	73.5	61.9
<i>Alcaligenes faecalis</i>	67.4	81.2	76.8	69.6	61.9
<i>Escherichia coli</i>	71.3	71.3	<b>82.9</b>	69.6	56.9
<i>Vibrio parahaemolyticus</i>	73.5	75.1	76.2	66.9	62.4
<i>Desulfobacter postgatii</i>	73.5	65.7	74.0	75.7	60.2
<i>Pirellula staleyi</i>	67.4	68.0	72.4	66.9	<b>74.6</b>
<i>Arthrobacter globiformis</i>	71.3	68.5	70.2	<b>86.7</b>	62.4
<i>Bacillus subtilis</i>	70.2	67.4	70.7	75.1	65.2
<i>Clostridium oceanicum</i>	74.0	71.8	68.5	71.8	63.5
<i>Bacteroides fragilis</i>	63.0	64.1	65.7	72.4	59.1
<i>Cloroflexus aurantiacus</i>	66.3	69.6	65.2	74.6	61.9
<i>Thermus ruber</i>	73.5	72.4	70.7	72.4	64.6

Origen: H, tracto intestinal de holoturino, Great Barrier Reef; Mc, muestra de suelo de Mount Coo-tha, Brisbane, Australia.

Tomado de Stackebrandt & Rainey 1995.

Cuando una secuencia analizada es ubicada dentro de un grupo con organismos que son considerados como representativos de un género particular, pero no se puede determinar una identidad específica al compararla con las secuencias de todas las especies empleadas para el análisis, el investigador debe de considerar en primer lugar la existencia de otras especies del mismo género cuyas secuencias no han sido determinadas aún, antes de llegar a la conclusión de que la clona secuenciada representa una especie nueva, que puede ser incluso considerada como no cultivable.

Por otro lado, la presencia de clonas con secuencias quiméricas, resultantes durante la reacción de amplificación por PCR en presencia de una gran fracción de ADN degradado, debe ser un hecho a considerar cuidadosamente. La presencia no detectada de este tipo de secuencias puede dar la impresión errónea de la presencia de especies procariontes únicas en un determinado ambiente. Para evitar un análisis por demás tedioso para determinar la presencia de secuencias quiméricas para cada una de las clonas secuenciadas en la base de datos RDP, ofrece una opción automatizada para verificar la presencia de quimeras (programa CHECK\_CHIMERA) (Stackebrandt & Rainey 1995).

---

## **2.6 ANÁLISIS DE DATOS MOLECULARES: ALINEAMIENTO, CÁLCULO DE DISTANCIAS Y CONTRUCCIÓN DE ÁRBOLES**

### **Alineamiento.**

El alineamiento de secuencias se realiza con la finalidad de comparar un grupo de secuencias para maximizar la similitud existente. El alineamiento de un grupo de secuencias se realiza de tal forma que es posible obtener una medida de la similitud o distancia existente entre las secuencias. Estas medidas son posteriormente empleadas por un algoritmo para la construcción de un árbol.

Los métodos de alineamiento más comunes se basan en una aproximación de programación dinámica desarrollado por Needleman & Wunsch en 1970 (Russek-Cohen & Colwell, 1996) en el cual, se le asigna un valor positivo a una igualdad, un

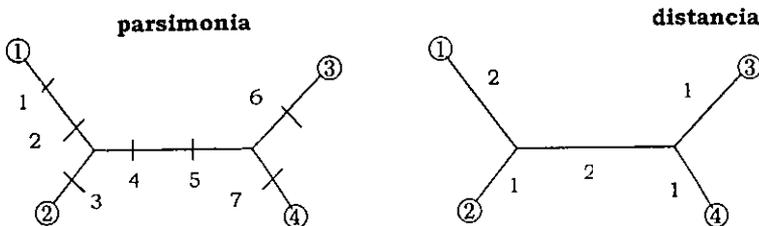
valor negativo a un *gap* y un valor de cero a un mal emparejamiento. Muchos de los algoritmos que existen en la actualidad difieren en la velocidad y en los valores asignados (penalizaciones), sin embargo existen aproximaciones recientes que son considerablemente más rápidos.

### Tipos de datos: distancias *versus* caracteres discretos.

Esta división se basa en la forma en que los datos son tratados. Los métodos de distancia primero convierten las secuencias alineadas en una matriz de distancia para posteriormente introducir esa matriz en un método para construir un árbol, mientras que los métodos discretos consideran cada sitio de un nucleótido (o alguna función de ese sitio) directamente. Como un ejemplo, se pueden considerar las siguientes secuencias y la correspondiente matriz de distancia:

secuencias		sitios							distancias				
		1	2	3	4	5	6	7					
secuencias	1	T	T	A	T	T	A	A					
	2	A	A	T	T	T	A	A	secuencias	2	3		
	3	A	A	A	A	A	T	A	3	5	4		
	4	A	A	A	A	A	A	T	4	5	4	2	
									1	2	3		

Los árboles obtenidos por parsimonia (un método discreto) o un método de mínima evolución (un método de distancia) presentan una topología y longitud de ramas idénticas:



El análisis de parsimonia permite identificar siete sustituciones y los sitios en los que se presentan en las cinco ramas del árbol, mientras que el árbol de distancia

aporta únicamente la distancia entre las secuencias sobre las ramas del árbol. Se puede observar también que ambos métodos proporcionan las mismas estimaciones de las longitudes para cada rama. Bajo parsimonia, cada uno de los siete sitios requiere un cambio, para un total de siete cambios. Si se suma la longitud de cada rama en el árbol de distancia, se obtiene el mismo valor:  $2+1+2+1+1=7$ . Sin embargo, es importante notar que el árbol de parsimonia proporciona información adicional de que sitio contribuye a la longitud de cada rama y una vez que son convertidas las secuencias a distancias, se pierde esa información. Además, los métodos discretos permiten realizar inferencias sobre los atributos de los ancestros extintos, en este caso, secuencias de nucleótidos extintas. En éste caso, estos ancestros reconstruidos pueden aportar datos sobre la evolución molecular.

#### **Métodos de agrupamiento versus métodos de búsqueda.**

Otra forma de dividir los métodos de construcción de árboles es la forma en la que éstos los construyen. De acuerdo a esto, se tiene a los **métodos de agrupamiento**, los cuales siguen una serie de pasos (algoritmo) para llegar a un árbol. Por ejemplo, si se tienen cinco secuencias se podría empezar con tres de ellas y decidir donde colocar la cuarta. Una vez que se tiene el árbol con la cuarta secuencia, entonces se decide en que lugar se adiciona la quinta y última secuencia al árbol.

Los métodos de agrupamiento tienen la ventaja de ser fáciles de implementar, resultando en programas de computación rápidos. Además casi siempre producen un solo árbol. Ésta combinación de velocidad y una aparente respuesta no ambigua hace que éstos métodos hayan ganado una gran popularidad. Sin embargo, poseen algunas limitaciones como herramientas analíticas.

El resultado obtenido de un algoritmo de agrupamiento depende con frecuencia del orden en el que son agregadas las secuencias al árbol en construcción, con lo que se podría llegar a un árbol distinto. Pero la limitación más importante es que estos métodos no nos permiten evaluar hipótesis, sino que solamente producen un árbol. Es posible que dos diferentes árboles puedan explicar una serie de datos de la misma forma, o muy parecidamente, por lo que si no se tiene una forma de medir la correspondencia entre los datos y el árbol, no será posible hacerlo.

Los métodos de agrupamiento de la segunda clase emplean un **criterio de optimización** para escoger de entre un grupo de todos los posibles árboles. Este criterio es empleado para asignar a cada árbol un valor o rango el cual es una función de las relaciones ente los árboles y datos (p.e. máxima parsimonia y máxima posibilidad, *maximum likelihood*).

Estos métodos tienen la gran ventaja de que permiten evaluar la calidad de cualquier árbol, por lo que es posible comparar que tan bien correlaciona una hipótesis sobre relaciones evolutivas con el grupo de datos analizados. La gran desventaja de éstos métodos es que son muy costosos computacionalmente. Un método óptimo posee dos problemas que deben de ser resueltos: primero, para un determinado grupo de datos y un árbol, ¿cuál es el valor del criterio óptimo para ese árbol?, por ejemplo, ¿cuál es el número mínimo de eventos evolutivos requeridos para explicar un grupo de datos observados?. Y en segundo lugar, ¿cuál de todos los posibles árboles posee el valor máximo de este criterio?. El primer problema es generalmente resuelto de forma sencilla, sin embargo, el segundo es mucho más complejo y pertenece a una clase de problemas llamados por los matemáticos como problemas *NP-complete* (*NP = non-deterministic polynominal*), para los cuales no existen algoritmos eficientes para su solución (Page & Holmes, 1999).

Históricamente los métodos de construcción de árboles que se basan en el agrupamiento han sido ampliamente utilizados en microbiología (Russek-Cohen & Colwell, 1996).

### **Medidas de distancia y similitud.**

La medida de la distancia permite definir que tan apartados están dos taxones, p.e. que tan diferentes son fenotípicamente, genotípicamente y por supuesto, filogenéticamente.

La medida empleada se selecciona con base en el tipo de caracteres registrados. Mientras que las distancias Euclidiana o Manhattan son frecuentemente empleadas para datos continuos, otras medidas tales como la de Jukes & Cantor, Kimura y Fitch-Margoliash, son empleadas para comparar datos de secuencias.

Jukes & Cantor (JC): Esta comparación emplea un método de un solo parámetro en el cual se asume que todas las sustituciones son igualmente probables. Siendo  $S$  la fracción de posiciones en común entre un par de secuencias y  $D = 1-S$ , entonces la distancia de Jukes & Cantor es:

$$d = -\frac{3}{4} \left[ \ln - (1 - 4D/3) \right]$$

El modelo original de JC no considera el efecto de *gaps* o deleciones. Sin embargo, la ecuación original se modificó para que el cálculo de la distancia considere *gaps*, deleciones e inserciones, dando una ecuación alternativa:

$$d_{AB} = -\frac{3}{4} \left[ 1 - \frac{4}{3} \left( \frac{S}{I+S} \right) \right] \left[ 1 - \frac{G}{N} \right] + \frac{G}{N}$$

Donde  $I$  es el número de bases idénticas,  $S$  es el número de sustituciones y  $G$  es el número de *gaps* en una secuencia con respecto a la otra. Se define a  $N$  como la suma de  $I+S+G$ .

De Wachter et. al. en 1985 (Russek-Cohen & Colwell, 1996) contabilizan cada posición por separado, los autores recomiendan que los *gaps* adyacentes sean considerados como uno solo:  $k_{nuc}$

En este procedimiento, a las tranversiones (pirimidina - purina) se le asigna un mayor peso que a una transición (purina - purina o pirimidina - pirimidina). En este caso, la medición de distancia se realiza por:

$$d = \frac{1}{2} \ln \left[ 1 - (2P - Q) / \sqrt{1 - 2Q} \right]$$

donde  $P = U_P/N$ ,  $Q = U_Q/N$  y  $N = M + U_P + U_Q$ .  $P$  es la fracción de posiciones de las secuencias diferentes por una transición,  $Q$  es la fracción de posiciones diferentes por una tranversión,  $N$  es el número de posiciones de secuencias comparadas en las cuales ambas secuencias contiene un nucleótido.  $M$  es el número de posiciones

en las que las dos secuencias son iguales. Este método falla al considerar un *gap* equivalente a una transversión (Russek-Cohen & Colwell, 1996).

Fitch-Margoliash: Estos dos autores propusieron una medida de distancia para comparar secuencias de ADN. En el procedimiento, ellos examinan las proteínas que son codificadas por estas secuencias y plantean la pregunta de ¿cuántas mutaciones en el ADN serían requeridas para codificar cada proteína?, comparando una proteína con otra (Russek-Cohen & Colwell, 1996).

Distancia Euclidiana-Manhattan: Una medida común de la distancia entre dos cepas cuando se emplean datos continuos es la distancia euclidiana. Esta aproximación es aplicable para el análisis de datos de ácidos grasos y otros atributos. Si  $(X_1, X_2, \dots, X_N)$  son el grupo o conjunto de variables cuantitativas registradas para la primera cepa ( $X$ ) y  $(Y_1, Y_2, \dots, Y_N)$  son el grupo correspondiente de variables cuantitativas para la segunda cepa ( $Y$ ), la distancia entre  $X$  y  $Y$  es:

$$d = \sqrt{\sum_1^N (X_i - Y_i)^2}$$

Si en su lugar se emplea:  $d = \sum_1^N |X_i - Y_i|$

Entonces se obtiene la distancia Manhattan o distancia de bloque (*city block distance*). Con ambas distancias (Euclidiana y Manhattan), si los atributos registrados tienen diferentes grados de variabilidad de diferentes unidades alternativas, las variables son frecuentemente estandarizadas antes del análisis:

$$\frac{X_i - (\text{Promedio de } X_i \text{'s})}{\text{Desviación estándar de } X_i \text{'s}}$$

### **Medidas de similitud.**

Una medida de similitud es una cantidad numérica registrada para cada par de taxones. Esta medida esta dada entre 0 y 1, si se considera a dos taxones idénticos. Dos de las más comunes medidas de similitud en taxonomía bacteriana es el coeficiente simple de identidad (*ssm*, *simple matching coefficient*) y el coeficiente de similitud de Jaccard ( $S_J$ ). El *ssm* es la proporción de caracteres que comparten dos cepas, p.e. características en común. El  $S_J$  es similar, pero ignora dobles negativos, p.e. si ambas cepas carecen de una característica dada, dentro del número total de caracteres. Estas medidas son empleadas con frecuencia cuando se usan datos fenotípicos binarios, pero la medida del *ssm* es incluso utilizada cuando se analizan datos de secuencias alineadas. La distancia obtenida por el método de JC es solo una transformación simple del *ssm* (Russek-Cohen & Colwell, 1996).

Medida de Sorensen. Para el análisis de ácidos grasos y cromatografía de gases (CG) se emplea la similitud de Sorensen. Esta es la relación de la suma de la mínima cantidad medida para cada prueba, p.e. compuestos orgánicos en el análisis de (CG) relativo al total de material para todas las pruebas para el par de cepas (Russek-Cohen & Colwell, 1996).

### **Clasificación de los métodos filogenéticos más comunes empleados para la construcción de árboles.**

Basados en la distinción de los métodos que usan distancias *versus* aquellos que emplean caracteres discretos y los métodos que emplean un algoritmo de agrupamiento y aquellos que se basan en un criterio de optimización, se pueden clasificar a estos métodos de la siguiente manera (Page & Holmes, 1999):

		Tipo de datos	
		Distancias	Sitios de nucleótidos
Métodos de construcción de árboles	Algoritmo de agrupamiento	UPGMA Neighbour joining	
	Criterio de optimización	Evolución mínima	Máxima parsimonia Maxima similitud ( <i>maximum likelihood</i> )

#### Algoritmos para la construcción de árboles de distancia.

Existen dos elecciones que se hacen frecuentemente para la construcción de árboles basados en distancias. Estas son el **método de grupos no emparejados con medias aritméticas** (*Unweighted pair group method with arithmetic means* o UPGMA) y el de vecinos unidos (*neighbour joining*, NJ) de Saitou & Nei (1987). Ambos métodos pueden ser alimentados con una matriz de similitud.

Estos métodos difieren en la forma en que los taxones son agrupados en cada etapa del análisis. Aunque históricamente el UPGMA ha sido el método de elección, existe evidencia de que el NJ puede ser más efectivo en discernir el mejor árbol a escoger, particularmente cuando se analizan datos moleculares. Sin embargo hay diversos autores que no están de acuerdo con esto. En este contexto, Huelsenbenck & Hillis en 1993 (Russek-Cohen & Colwell, 1996), indican que el algoritmo que probablemente sea más efectivo es aquel que es dependiente de las suposiciones que gobiernan las tasas de evolución. Por lo anterior, no es probable que haya un solo método correcto por el que se logre conseguir este hecho. Esto sugiere entonces, que uno debe de experimentar con dos o más procedimientos y esperar una correlación entre ellos (Russek-Cohen & Colwell, 1996).

UPGMA. Éste método construye un árbol ultramérico, en el cual todas las puntas son equidistantes de la raíz del árbol, lo que es equivalente a asumir un reloj molecular. En este método, cada cepa o UTO, es considerada a ser una de  $n$  clusters separados en un estado 1. Empleando una distancia métrica, los dos taxones más cercanos son combinados. Esto implica que en este estado se tienen  $n-1$  clusters. Los dos clusters más cercanos son entonces combinados para formar  $n-2$  clusters. La distancia entre ellos es  $(d_{ij})$ , donde  $i$  es elemento del cluster  $j$ .

El proceso de combinar los clusters más cercanos se repite secuencialmente hasta que se define un solo cluster enramado a todas ( $n$ ) las cepas. El resultado de este análisis típicamente es presentado como un árbol o dendograma. El eje debajo de él representa la distancia en la cual un determinado taxón une un cluster (Russek-Cohen & Colwell, 1996).

Neighbor joining (NJ). Este método es ampliamente utilizado para la construcción de árboles ya que combina la velocidad computacional con la obtención de un solo resultado (un solo árbol). Estos dos atributos lo hacen muy atractivo. NJ es un método de agrupamiento y presenta la limitación de que no optimiza un criterio de correlación entre los datos y el árbol obtenido. Sin embargo, es un buen método heurístico para la estimación de un árbol evolutivo mínimo.

Este algoritmo asume que existen cambios independientes en todos los sitios con igual probabilidad. Si una base cambia es independiente de su identidad y cuando cambia hay una misma probabilidad de terminar en cada una de las otras tres bases. Por lo que la matriz de transición para un periodo de tiempo corto o  $dt$  es:

		Hacia	A	G	C	T
		:				
De:	A		1-3a	a	a	A
	G		a	1-3a	a	A
	C		a	a	1-3a	a
	T		a	a	a	1-3a

donde  $a$  es  $u dt$ , el producto de la tasa de substitución por unidad de tiempo  $u$  y la duración  $dt$  del intervalo de tiempo. Para periodos de tiempo largos esto implica que la probabilidad de que dos secuencias difieran en un sitio dado es:

$$p = \frac{3}{4} \left(1 - e^{-\frac{4u}{3}}\right)$$

y por lo que si observamos  $p$ , se puede computar una estimación de la longitud de la rama  $ut$  invirtiendo la ecuación anterior para obtener:

$$ut = -\frac{3}{4} \log_e \left(1 - \frac{4}{3} p\right)$$

En ésta aproximación, el vecino más cercano es seleccionado de la siguiente forma: un par de vecinos es un par de UTO's conectados por medio de un solo nodo interior en un árbol bifurcado no enraizado.

El número de pares de vecinos es  $\frac{n}{2}$  cuando  $n$  es par y  $\frac{(n-1)}{2}$  si es impar. Inicialmente se asume una topología en forma de red o de estrella del árbol a construir. Se unen dos taxones resultando la suma de la longitud de las ramas. Esto se realiza al examinar el número posible de árboles  $\frac{n(n-1)}{2}$  que pueden ser computados cuando un par de UTO's son combinados en una unidad taxonómica hipotética o UTH.

Una vez que los taxones más cercanos (#1 y #2 por conveniencia) son combinados en un solo UTH, el resultado es de  $n-1$  UT's. La distancia de la UTH combinada a la UTO <sub>$j$</sub>  se define como:

$$D_{(1-2j)} = (D_{1j} + D_{2j})/2 \quad 3 \leq j \leq n$$

Esto resulta en una matriz revisada. El proceso se repite buscando nuevos vecinos. Este proceso se repite con todos las UTO's restantes y se genera un árbol no enraizado (Felsenstein, 1989; Russek-Cohen & Colwell, 1996; Page & Holmes, 1999).

### **Métodos discretos.**

En contraste con los métodos de distancia, los métodos discretos operan directamente en las secuencias, o en las funciones derivadas de las secuencias, en lugar de actuar sobre las distancias. Por lo que evitan la pérdida de información que se presenta cuando una secuencia es convertida a distancias. Los dos métodos discretos principales son la máxima parsimonia y la máxima posibilidad (*maximum likelihood*). La parsimonia máxima escoge el árbol (o árboles) que explica(n) un menor número de cambios evolutivos. La máxima posibilidad escoge el árbol (o árboles) que a partir de todos los árboles, es el único más similar al que han generado los datos observados.

Máxima parsimonia. En cladística, un árbol es construido bajo el principio de máxima parsimonia o el principio de evolución mínima. Esta aproximación es usada más exactamente en áreas de la biología como son la entomología y la paleontología más que en microbiología (Russek-Cohen & Colwell, 1996). Los árboles obtenidos asumen la representación de las relaciones evolutivas entre los taxones analizados.

La clasificación basada en datos fenotípicos, tales como variables binarias tienden a ser completamente funcionales en estudios cladísticos, un hecho crítico es la direccionalidad de los caracteres. En cladística, se asume que uno puede determinar que el estado de carácter binario (p.e. 0 ó 1) se encuentra en un estado primitivo o avanzado. Es claro que no es posible hacer que este tipo de asignaciones fácilmente, por lo que en el pasado, los microbiólogos evitaron emplear estos métodos. Sin embargo, con el uso de datos moleculares es posible considerar las relaciones filogenéticas entre microorganismos.

La búsqueda de un árbol único con la máxima parsimonia es una tarea que involucra la demanda de mucho tiempo, y puede ser difícil o hasta imposible. Un problema es el gran número de árboles que se necesitan examinar para encontrar el

árbol más parsimonioso. Por ejemplo, con  $t$  UTO's es posible encontrar  $(2t-3)!/2^{t-2}(t-2)!$  árboles, o cuando  $t = 10$ , esto significa que se obtendrán 34,459,425 árboles. Como consecuencia, son requeridos procedimientos que limiten el número de árboles. Si la resolución de los datos no es adecuada, se pueden encontrar varios árboles parsimoniosos, para construir posteriormente un árbol consenso, que describa las características de todos o la mayoría. Un paquete computacional que ofrece esta alternativa es el paquete PAUP (Russek-Cohen & Colwell, 1996; Page & Holmes, 1999).

Maximum likelihood (máxima posibilidad). Felsenstein en 1981 (Russek-Cohen & Colwell, 1996) propuso una aproximación de conjunto de procedimientos estadísticos conocidos como de máxima posibilidad (*maximum likelihood*, ML). La técnica se basa en la estimación de la probabilidad de diferentes tipos de sustitución (p.e. la probabilidad de que A pueda cambiar a C).

El método asume que todas las bases en una secuencia se comportan independientemente y asume también que los cambios son verosímiles aunque sean poco probables generando un número mínimo que explique todas las diferencias.

Este método requiere de una gran cantidad de tiempo para poder obtener un árbol y tiene el riesgo de que los resultados obtenidos no necesariamente concuerden con lo existente, por lo que podría explicar hechos biológicos falsos o erróneos (Russek-Cohen & Colwell, 1996).

# 3

## HIPÓTESIS y OBJETIVOS

---

### HIPÓTESIS DEL PROYECTO.

“Es posible determinar la diversidad de bacterias lácticas en una muestra de pozol tipo indígena, por medio del análisis de secuencias parciales de ADNr 16S obtenidas a partir de la fracción de ADN bacteriano extraído de este alimento”

---

### OBJETIVOS.

#### Objetivo general.

Conocer la diversidad de bacterias lácticas presente en una muestra de pozol tipo indígena, por medio del análisis de la secuencia de ADNr 16S obtenidas a partir de una muestra de ADN bacteriano total extraído de este alimento.

#### Objetivos particulares.

1. Desarrollo una metodología de extracción de ADN de la fracción bacteriana total presente en una muestra de pozol.
2. Amplificación por PCR de los ADNr 16S presentes en la muestra de ADN obtenido, empleando un par de *primers* universales para Eubacterias.
3. Clonación de los ADNr 16S amplificados por PCR y caracterización de las clonas con inserto.
4. Secuenciación de los ADNr 16S clonados.

5. Comparación de las secuencias obtenidas contra aquellas depositadas en una base de datos para conocer su identidad y análisis, para determinar así la diversidad bacteriana presente en la muestra de pozol.

# 4

## MATERIAL Y MÉTODOS

---

### 4.1 EXTRACCIÓN DEL ADN TOTAL DE LA FRACCIÓN BACTERIANA DEL POZOL POR UN MÉTODO INDIRECTO.

#### **Obtención de una muestra de pozol procedente de la localidad de San Cristóbal de las Casas, Chiapas.**

Se obtuvo una muestra de pozol tipo indígena elaborado con maíz amarillo y envuelto en hojas de plátano del mercado de la ciudad de San Cristóbal de las Casas, Estado de Chiapas, México (figura 6). La muestra de pozol, con un tiempo de fermentación aproximado de 8 - 10 horas en el momento de su obtención, fue colocada en una bolsa de plástico estéril (STOMACHER, SEWARD MEDICAL) y mantenida en congelación en una mezcla de hielo con NaCl comercial durante su traslado a la Ciudad de México. La muestra fue mantenida en congelación a -20°C hasta el momento de su análisis.

#### **Separación de la fracción bacteriana del pozol.**

Una muestra de 50 g de pozol tipo indígena se utilizó como sustrato para la extracción indirecta del ADN de la fracción bacteriana presente, empleando una modificación del protocolo descrito por Holben *et al.* (1988). La muestra analizada contenía porciones del centro, parte media y superficie de la bola de pozol para asegurar una representatividad de cada uno de los posibles microambientes presentes en el alimento. El procedimiento de extracción de ADN empleado se muestra en el diagrama de la figura 7 y estuvo conformado por la separación



**Figura 6.**

Muestra de pozol tipo indígena que se consiguió en el mercado de San Cristóbal de Las Casas, Chiapas y empleado como sustrato para la extracción del ADN de la fracción bacteriana presente (0.3x).

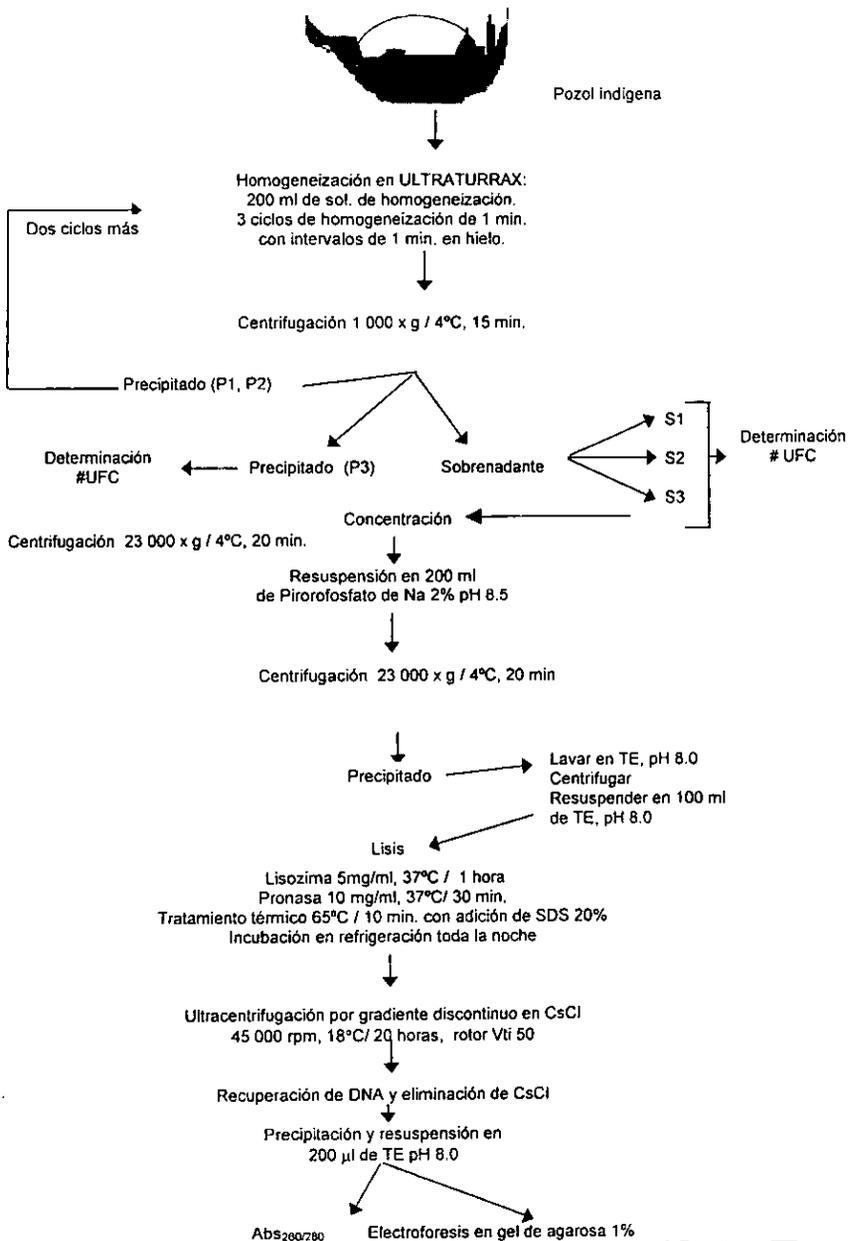


Figura 7.  
Proceso de extracción indirecta del ADN bacteriano del pozo.

de la fracción bacteriana por un proceso de homogeneización/centrifugación, lisis celular y purificación del ADN.

La muestra de 50 g de pozol fue colocada con ayuda de una espátula en una botella de centrifuga de 500 ml, adicionándose 200 ml de una solución modificada de Winogradski ( $K_2HPO_4$  1.43 mM (BAKER),  $MgSO_4 \cdot 7H_2O$  1.11mM (BAKER), NaCl 2.14 mM (BAKER),  $Fe(SO_4)_3$  4.75  $\mu$ M (BAKER),  $MnSO_4$  14.8  $\mu$ M (BAKER), pH 7.0), diluida 1:20 (v/v) en  $H_2O$  destilada justo antes de ser utilizada. Para asegurar una separación eficiente de la fracción bacteriana del sustrato, la muestra de pozol fue homogeneizada en un mezclador Ultraturrax (ULTRATURRAX T25, IKA LABOR TECHNIK) a una velocidad media, en tres ciclos de homogeneización de 1 min. con intervalos de incubación en hielo de 1 min. (para evitar el calentamiento de la muestra). El homogeneizado fue centrifugado a 1000 x g durante 15 min. a 4°C, para separar células fúngicas y partículas del alimento. El sobrenadante (S1) se transfirió a otra botella de centrifuga de 250 ml y fue mantenido en hielo. Para realizar una extracción exhaustiva de la fracción bacteriana, el *precipitado* resultante fue homogeneizado en Ultraturrax dos veces más como se describió anteriormente, obteniéndose los sobrenadantes S2 y S3 respectivamente. Los sobrenadantes S1, S2 y S3 fueron recentrifugados a 23000 x g durante 20 min a 4°C para colectar la fracción bacteriana. Para evaluar la eficiencia de la extracción bacteriana del alimento, se plaquearon alícuotas de los tres sobrenadantes y del precipitado final del proceso de homogeneización/centrifugación (P3) en agar-APT (DIFCO) y se calculó la eficiencia de recuperación celular.

Los precipitados bacterianos pS1, pS2 y pS3 fueron resuspendidos y concentrados en 200 ml de una solución de pirofosfato de sodio 2%, pH 8.5 (SIGMA). La muestra se centrifugó a 23000 x g durante 20 min. a 4°C. Este paso es reportado como necesario para eliminar material orgánico no celular del alimento, remanente en la fracción celular separada (Holben *et al.* 1988). El precipitado se lavó dos veces con 200 ml de una solución de TE (TRIS-HCl 33mM (SIGMA), EDTA 1 mM (BAKER), pH 8.0) y se centrifugó en las condiciones anteriormente descritas. El precipitado se resuspendió finalmente en 100 ml de TE. Esta solución fue utilizada para iniciar el proceso de lisis celular.

El manejo de la muestra de pozol a analizar y el proceso de extracción celular fueron realizados bajo condiciones ambientales asépticas y con material y soluciones estériles para evitar la posible contaminación con microorganismos no asociados al alimento.

### **Extracción y purificación del ADN bacteriano.**

**Lisis de la fracción bacteriana aislada.** A la suspensión bacteriana en TE se le adicionaron 25 ml de una solución de NaCl 5M. La muestra fue incubada a temperatura ambiente (t.a.) durante 10 min. y las células fueron centrifugadas a  $23000 \times g$  durante 20 min. a  $4^{\circ}\text{C}$ . Se desechó el sobrenadante y el *precipitado* se resuspendió en 10 ml de una solución TS (TRIS-HCl 50 mM, NaCl 50mM, pH 8.0) y se transfirió a otro tubo de centrifuga de 50 ml al que se le añadieron 50  $\mu\text{l}$  de una solución de SDS al 20% (SIGMA). Se incubó la muestra a t.a. durante 10 min. y se centrifugó como se describió anteriormente. El *precipitado* fue recuperado y resuspendido en 3.5 ml de una solución de TSE (TRIS-HCl 50mM, Sacarosa 0.75 M (BAKER), EDTA 10 mM, pH 8.0). Se adicionó lizozima cristalina (SIGMA) hasta una concentración final de 5 mg/ml y se incubó a  $37^{\circ}\text{C}$  durante 1 hora, con agitación ocasional. Al término de la incubación con lizozima se adicionaron 0.5 ml de una solución de 10 mg/ml de Pronasa predigerida (BOEHRINGER MANHEIM) (en una solución de TS). La muestra se homogeneizó en vórtex a máxima velocidad y se incubó a  $37^{\circ}\text{C}$  durante 1 h. Al término de la incubación, la muestra se calentó a  $65^{\circ}\text{C}$  en un baño de agua y se adicionaron 0.25 ml de una solución de SDS 20%, incubando 10 min. a  $65^{\circ}\text{C}$ . Después de este tratamiento la muestra se mantuvo en refrigeración toda la noche.

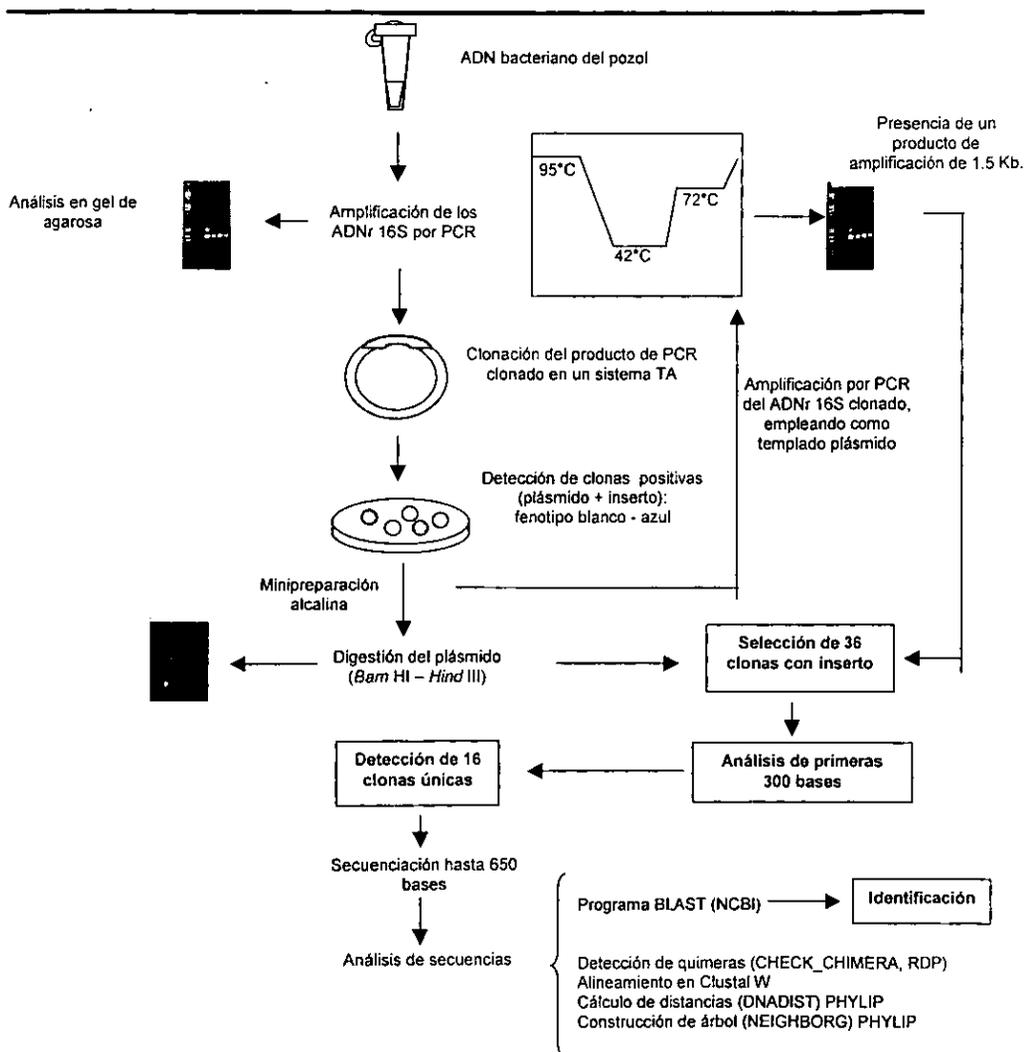
Para evaluar la posibilidad de interferencia de la masa en el proceso de lisis celular, se realizaron los siguientes controles: Control 1. Una muestra de 100 g de masa de maíz obtenida en una tortillería fue disuelta en un cultivo de 100 ml de *Lactobacillus amylophilus* incubado en caldo APT (DIFCO) y procesado de acuerdo a la metodología descrita para la extracción de células y lisis. Control 2, un cultivo de 100 ml en caldo APT de *L. amylophilus* fue sometido al proceso de lisis descrito. En ambos casos, los cultivos de *L. amylophilus* fueron incubados toda la noche. Para determinar la eficiencia del proceso de lisis en ambos controles, se plaqueó una alícuota de las mismas en agar-APT antes de iniciar el

proceso de lisis, después del tratamiento con lisozima-pronasa y después del tratamiento con SDS.

**Purificación del ADN bacteriano.** La suspensión lisada en refrigeración se clarificó por centrifugación a 40 000 x *g* durante 1 hora a 4°C y alícuotas de 5 ml del sobrenadante se transfirieron a tubos de centrifuga de 50 ml previamente pesados. Las muestras fueron preparadas para ser purificadas por ultracentrifugación en un gradiente discontinuo de cloruro de cesio (CsCl<sub>2</sub>)/bromuro de etidio (EtBr) en una centrifuga XL-90 (BECKMAN), de acuerdo a la metodología descrita por Sambrook *et al.* (1988).

La banda de ADN cromosomal obtenida del proceso de ultracentrifugación, se visualizó bajo luz ultravioleta de onda larga y se recuperó con ayuda de una aguja en un tubo de microcentrifuga de 1.5 ml. El ADN fue purificado eliminando el EtBr por lavados sucesivos con 1 volumen de isopropanol (BAKER) y centrifugando a 200 x *g* durante 2 min. a t.a., hasta obtener la fase acuosa libre de EtBr. El CsCl<sub>2</sub> se eliminó lavando la fase acuosa recuperada y concentrada de cada tubo, con 2 volúmenes de agua destilada y 6 volúmenes de etanol absoluto frío (BAKER), mezclando manualmente e incubando exactamente 30 min. a 4°C. Este tratamiento se realizó en un tubo de centrifuga de 50 ml estéril. El ADN se recuperó por centrifugación de la muestra incubada a 15000 x *g* durante 30 min. a 4°C. El sobrenadante fue descartado y el *precipitado* secado al vacío y resuspendido en 400 µl de agua destilada. La integridad y calidad del ADN obtenido se evaluó por electroforesis en un gel de agarosa al 1% (SIGMA) y su pureza y concentración se determinó espectrofotométricamente.

A partir del ADN obtenido de la fracción bacteriana del pozol, se obtuvieron las secuencias parciales de los ADN<sub>r</sub> 16S amplificados por PCR, las cuales fueron identificadas posteriormente. El procedimiento empleado se muestra en el diagrama de la figura 8.



**Figura 8.**

Obtención de las secuencias parciales de los ADNr 16S obtenidos por amplificación por PCR a partir del ADN bacteriano extraído del pozo y su posterior análisis

---

#### **4.2 AMPLIFICACIÓN POR PCR DE LOS ADNr 16S PRESENTES, CLONACIÓN DEL PRODUCTO DE AMPLIFICACIÓN E IDENTIFICACIÓN DE LAS CLONAS CON INSERTO.**

##### **Amplificación por PCR de los ADNr 16S presentes.**

Los ADNr 16S se amplificaron por PCR de acuerdo a las condiciones descritas por Weisburg, *et al.* (1991). Se empleó un par de *primers* denominados fD1 5'd(AGAGTTTGACCTCCTCCTGGCTCAG)3' y rD1 5'd(AAGGAGGTGATCCAGCC)3', reportados para la amplificación de ADNr 16S de los diferentes grupos bacterianos mostrados en la tabla 7. La reacción de PCR se realizó en tubos de microcentrífuga de 0.5 ml con 100 µl de la siguiente mezcla de reacción: PCR reaction buffer 1X (TRIS-HCl 200mM, pH 8.0, KCl 500mM) (GIBCO-BRL), MgCl<sub>2</sub> 1.5 mM, pH 8.0 (GIBCO-BRL), 200µM de cada dNTP (dATP, dCTP, dGTP, dTTP, AMERSHAM), AmpliTaq DNA Polimerasa recombinante (2.5 U/100µl) (GIBCO-BRL), *primer* rD1 (1.0 µM), *primer* fD1 (1.0 µM) y ADN templado, 14 ng/µl (concentración final) y llevados a 100 µl con agua destilada estéril. La mezcla de reacción se cubrió con 50 µl de aceite mineral estéril (BAKER). Las condiciones de reacción fueron: 1 ciclo de desnaturalización inicial a 95°C durante 5 min., y 35 ciclos de amplificación: 95°C durante 2 min., 42°C durante 0.5 min. y 72°C durante 4 min. La reacción de amplificación se realizó en un termociclador multibloques (Robocycler 40 STRATAGENE). Al término de la reacción de amplificación una muestra de 10 µl de reacción fue analizada en electroforesis en gel de agarosa al 1.5%.

##### **Clonación del producto de amplificación e identificación de las clonas con inserto.**

El producto de amplificación de 5 reacciones independientes fue concentrado y el aceite mineral remanente se eliminó por extracción con 1 volumen de cloroformo (BAKER). La fase acuosa fue recuperada y precipitada con 2 volúmenes de etanol absoluto frío (BAKER) e incubado 30 min. a -20°C, se centrifugó en las condiciones descritas, se lavó con 1 ml de etanol al 70%, se secó al vacío y se resuspendió en 50 µl de H<sub>2</sub>O destilada estéril. Una alícuota del producto de PCR

fue clonada en el *kit* para clonación de productos de amplificación por PCR pT7Blue T-Vector Kit (NOVAGEN cat. no. 69836-1). Este tipo de vectores, denominados vectores de clonación TA, están diseñados para la clonación rápida de productos de PCR, aprovechando la característica de muchas ADN polimerasas termoestables, tales como la *Taq* ADN Polimerasa, de dejar una cadena sencilla de adeninas (A) en el extremo 3' de sus productos. Estas moléculas pueden ser ligadas directamente a un vector compatible que contiene una cadena sencilla de timinas (T) (Ausubel *et al.* 1992).

**Reacción de ligación.** La reacción de ligación y transformación se realizó de acuerdo al protocolo estándar descrito por el proveedor: para una reacción de ligación estándar se ligaron 50 ng (0.03 pmol) del plásmido pT7Blue T-Vector con 0.2 pmol del producto de amplificación (proporción molar aproximada de inserto a vector de 5:1 bajo estas condiciones) en un tubo de 1.5 ml. Cada uno de los componentes empleados en la reacción de ligación, son incluidos en el *kit* mencionadas anteriormente:

1.0 µl	Buffer de ligasa (TRIS-HCl 200mM pH 7.6, MgCl <sub>2</sub> 50mM)
0.5 µl	DTT 100 mM
0.5 µl	ATP 10 mM
1.0 µl	pT7Blue T-Vector 50 ng/µl
2.0 µl	Producto de amplificación (0.2 pmol aproximadamente)
14.5 µl	H <sub>2</sub> O
0.5 µl	ADNLigasa T4 (2-3 Unidades Weiss) (Adicionada al final).

La mezcla de reacción se homogeneizó por pipeteo suave y se incubó 2 h a 16°C.

**Transformación.** Una alícuota de 20 µl de células competentes de *Escherichia coli* NovaBlue (incluidas en el *kit* de clonación), fue descongelada en hielo y se transfirió a un tubo de microcentrifuga de 1.5 ml en hielo. De la reacción de ligación se pipeteó 1 µl directamente a las células, se mezcló por agitación suave. El tubo con las células fue incubado 30 min. en hielo y calentado exactamente 40 seg. a 42°C en un baño de agua sin agitar y transferido inmediatamente a un baño de hielo e incubado por 2 min. (Sambrook *et al.*, 1988). Al término de la incubación se agregaron 80 µl de medio SOC a t.a. (incluido en el *kit*), se incubó en agitación a 250 rpm durante 1 hora a 37°C.

Tabla 7.

Características de los *primers* rD1 y fD1 empleados para la amplificación de genes de ARNr 16S.

---

**fD1** 5'- ccgaattgctccacaacAGAGTTTGACCTCCTCCTGGCTCAG -3'  
 longitud: 37 nucleótidos  
 % G + C: 51.53  
 Tm: 69.1° C  
 Sitios de restricción para *EcoRI* y *Sall* (región en minúsculas)  
 Sitio de unión al gen de ADNr 16S de *E. coli* en las posiciones 8-27

Amplificación de genes de ARNr 16S de:

Bacterias Gram (+) y relacionadas,  
 Bacterias púrpuras y relacionadas (proteobacteria),  
*Cyanobacteria*,  
*Bacteroides/Flavobacteria*,  
*Deinococcus* y relacionados,  
 Espiroquetas,  
*Plachtiomyces* y relacionados,  
*Thermotaga*.

**rD1** 5'- cccgggtccaagcttAAGGAGGTGATCCAGCC - 3'  
 longitud: 33 nucleótidos  
 % G + C: 60.61  
 Tm: 70.68° C  
 Sitios de restricción para *BamHI*, *HindIII* y *XmaI* (región en minúsculas)  
 Sitio de unión al gen de ADNr 16S de *E. coli* en las posiciones 1541-1525

Amplificación de genes de ARNr 16S de:

Bacterias púrpuras y relacionadas,  
 Bacterias Gram (+) y relacionadas,  
*Cyanobacteria*,  
 Espiroquetas,  
*Plachtiomyces*,  
*Chlorobium*,  
*Thermotaga*.

La combinación fD1+rD1 genera un producto de amplificación de aproximadamente 1500 kb.

---

Tomado de Weisburg, *et al.* 1991.

Después de esta incubación, se plaqueó una alícuota 50  $\mu$ l de mezcla de transformación se plaquearon en placas de agar-Luria-Bertani preparado a partir de sus ingredientes (Sambrook, *et al.*, 1988), suplementadas con ampicilina 50  $\mu$ g/ml (SIGMA) (amp) y tetraciclina 15  $\mu$ g/ml (SIGMA) (tet). La tet asegura la selección del fenotipo F' que contiene el gen marcador *lacZAM15*, eliminando la interferencia de colonias blancas no recombinantes que han perdido el fenotipo F'. Para la selección de las colonias recombinantes blancas/azules, es necesario incluir IPTG 40 mM (BOHERINGER-MANHEIM) y X-gal 80  $\mu$ g/ml (BOHERINGER-MANHEIM) a las placas con agar-LB. Estos compuestos son adicionados bajo condiciones asépticas una vez que el medio agar-LB ha sido esterilizado y con una temperatura aproximada de 40°C (Ausubel *et al.* 1992). Las cajas se incubaron toda la noche a 37°C.

**Selección del fenotipo blanco/azul.** El vector pT7Blue permite la selección de recombinantes por selección de colonias blancas/azules. El plásmido codifica una copia funcional del  $\alpha$ -péptido *lacZ* que complementa el fragmento  $\omega$ -*lacZ* expresado por la cepa hospedera (NovaBlue). Como resultado de esta complementación, la enzima activa  $\beta$ -galactosidasa puede hidrolizar el sustrato cromogénico X-gal para generar colonias con un color azul. Los insertos que son clonados en un sitio de corte para la enzima *EcoRV* modificado localizado en la región que codifica al  $\alpha$ -péptido interfieren con su función, por lo que las colonias recombinantes aparecen blancas cuando son plaqueadas en cajas suplementadas con IPTG/X-gal.

Sin embargo, adicionalmente al fenotipo azul y blanco de las colonias que han crecido en las placas, se pueden observar colonias azul claro resultantes de una autoligación (poco eficiente, pero detectable) de los extremos de dT por la ligasa T4 y colonias blancas con centro azul como resultado de la síntesis de pequeñas cantidades del  $\alpha$ -péptido funcional. Para evitar la posibilidad de aparición de falso positivos, las colonias blancas resultantes se transfirieron a una caja nueva de agar-LB (complementada como se describió previamente) para corroborar el fenotipo *lac* (colonias blancas). Las cajas se incubaron a 37°C toda la noche y las colonias blancas resultantes fueron transferidas asépticamente a

tubos con 3 ml de caldo LB suplementado con amp (50 µg/ml) y tet (12.5 µg/ml), incubándose a 275 rpm y 37°C toda la noche.

De los cultivos resultantes, 1 ml fue mantenido en congelación a -70°C en glicerol al 30% (BAKER) y 1.5 ml fueron procesados para extraer el plásmido, utilizando el protocolo de minipreparaciones por el método de lisis alcalina (Ausubel *et al.* 1992).

**Verificación de la presencia de un inserto clonado correspondiente al tamaño del ADNr 16S.** Para confirmar la presencia de un inserto de aproximadamente 1.5 kb en la colección de clonas se realizaron las siguientes pruebas:

- (a) Digestión enzimática del plásmido+inserto: una alícuota de cada minipreparación de plásmido fue digerida simultáneamente con las enzimas *Sal I* y *Hind III* (GIBCO-BRL) a 37°C toda la noche en reacciones de 20 µl. La digestión fue analizada por electroforesis en gel de agarosa 1.5% evaluando la presencia de una banda correspondiente al tamaño del plásmido y de una banda de aproximadamente 1.5 kb correspondiente al tamaño promedio del ADNr 16S amplificado por PCR (Weisburg *et al.* 1991).
- (b) Amplificación por PCR: una alícuota de cada minipreparación de plásmido fue utilizada como templado para una desarrollar una reacción de PCR para la amplificación del ADNr 16S clonado en el plásmido. Para este experimento, se emplearon las condiciones de amplificación y concentraciones de reactivos ya empleadas y descritas previamente.

Como resultado de estos análisis se detectaron 36 clonas positivas.

---

#### 4.3. SECUENCIACIÓN PARCIAL DE LOS ADN<sub>r</sub> 16S CLONADOS.

La colección de 36 clonas recombinantes con el plásmido+inserto fue secuenciada hasta 300 bases para determinar la presencia de clonas únicas. A partir de cada una de las 36 clonas, se obtuvo una minipreparación de plásmidos recombinantes para ser empleado como templado en una reacción de secuenciación a partir de ADN de doble cadena. Se realizaron cultivos de 3 ml de caldo LB suplementado con amp 50 µg/ml y tet 12.5 µg/ml a 37°C toda la noche. Una alícuota de 1.5 ml de cada cultivo fue procesada de acuerdo al protocolo de preparación de templado de doble cadena descrito por Kovalenko *et al.* (1994). Este protocolo permite la obtención de plásmidos libres de ARN y con un rendimiento promedio de hasta 5 µg de ADN por 1.5 ml de un cultivo en saturación de *E. coli*:

- (a) Un cultivo de 1.5 ml de cada clona se centrifugó a 3000 rpm durante 3-5 min a t.a.
- (b) El *precipitado* fue resuspendido en 0.1 ml de TE (TRIS-HCl 10 mM, pH 8.0, EDTA 1 mM) y se adicionó 0.1 ml de fenol:cloroformo:isoamílico (25:25:1 v/v) (BAKER). La mezcla se agitó en vórtex a máxima velocidad durante 20 seg y se centrifugó a 5000 rpm durante 5 min a t.a.
- (c) La fase no acuosa se transfirió a un segundo tubo de 1.5 ml y se adicionó 0.1 ml de TNEM (TRIS-HCl 10 mM, pH 8.0, NaCl 0.1 M, EDTA 1 mM, MgCl<sub>2</sub> 0.1 M (BAKER)) y 100 µg/ml de Ribonucleasa A (libre de DNAsas, 10 mg/ml stock, congelado a -20°C (SIGMA)). La mezcla fue centrifugada a 5000 - 7000 rpm durante 5 min. a t.a.
- (d) El sobrenadante se transfirió a un tercer tubo y el plásmido se precipitó con la adición de 0.1 ml de acetato de amonio 10 M (BAKER) y 0.8 ml de etanol absoluto frío, incubándose en hielo durante 15 min y colectándose por centrifugación a 5000 rpm durante 5 min. a t.a.
- (e) El etanol residual se decantó y el *precipitado* fue lavado con 1 ml de etanol 70% y secado al vacío.

Para la reacción de secuenciación se empleó el *kit* de Thermo-Sequenase radiolabeled terminator cycle sequencing (Amersham US79750) bajo las siguientes condiciones de reacción:

- (a) Preparación de reacciones de terminación. Para cada secuencia a realizar se preparo una mezcla de reacción empleando 2  $\mu\text{l}$  de la mezcla de terminación (termination master mix) y 0.5  $\mu\text{l}$  de ddNTP's marcados, de acuerdo al siguiente cuadro:

Mezcla de terminación ( $\mu\text{l}$ )	1 secuencia			
	G	A	T	C
$[\alpha\text{-P}^{33}]\text{ddGTP}$	0.5			
$[\alpha\text{-P}^{33}]\text{ddATP}$		0.5		
$[\alpha\text{-P}^{33}]\text{ddTTP}$			0.5	
$[\alpha\text{-P}^{33}]\text{ddCTP}$				0.5
Total	2.5	2.5	2.5	2.5

- (b) Una vez preparada esta mezcla se repartieron 2.5  $\mu\text{l}$  de cada reacción de terminación en tubos rotulados respectivamente como G, A, T, C para cada secuencia. Cada tubo se mantiene en incubación en hielo hasta su uso.
- (c) Para el desarrollo de la reacción se preparó la siguiente mezcla en tubos eppendorf de 0.5 ml:

Buffer de reacción	2 $\mu\text{l}$
Templado	<sup>A</sup> $\mu\text{l}$
Primer <sup>C</sup>	1 $\mu\text{l}$ [2 pmol/ $\mu\text{l}$ ]
H <sub>2</sub> O	<sup>B</sup> $\mu\text{l}$
ThermoSecuenasa	2 $\mu\text{l}$

<sup>A</sup> Dependiendo de la concentración, máximo hasta 10  $\mu\text{l}$ , <sup>B</sup> Llevar a un volumen final de 20  $\mu\text{l}$ , <sup>C</sup> Primer incluido en el *kit* de Termosecuenaasa (M13 - 40 forward).

Una vez preparada la mezcla de reacción, se colocaron 4.5  $\mu\text{l}$  de cada mezcla de reacción a cada tubo de terminación (G, A, T, C ) y se agregaron 20  $\mu\text{l}$  de aceite mineral.

- (d) La reacción de terminación fue desarrollada bajo las siguientes condiciones: 35 ciclos a 95°C por 30 seg, 55°C por 30 seg, 72°C por 2 min, con un ciclo inicial de desnaturalización a 95°C por 5 min. Esta reacción se realizó en un termociclador multibloques (Robocycler 40 STRATAGENE). Al término de la reacción se adicionaron 4  $\mu\text{l}$  de *stop solution* a cada tubo. Finalmente las muestras se calentaron 10 min a 72°C justo antes de ser analizadas en un gel para secuenciación de urea-acrilamida (MERCK).
- (e) Una vez cargadas las muestras en el gel, la electroforesis se desarrollo a 1500 volts y 50 watts. Al término de ésta, el gel fue separado de los vidrios, pegado a una hoja de papel filtro, secado por 1 hora a 80°C y expuesto a una película fotográfica (KODAK) toda la noche a -20°C. Al término del período de exposición, la película fue revelada y analizada.

Con el análisis de las primeras 300 bases de cada clona, se determinó la presencia de 16 clonas únicas. Para esto, una muestra de 15  $\mu\text{l}$  de cada preparación de templado con una concentración de 200 ng/ $\mu\text{l}$  fue enviada para su análisis al servicio de secuenciación automatizado de la Escuela de Ciencias Biológicas de la Universidad de Auckland, Nueva Zelanda (Private Bag 92019, Auckland, New Zealand).

La reacción de secuenciación realizada se basa en el proceso *Dye Terminator Chemistry*, en el cual las reacciones de extensión de ADN son amplificadas linealmente y el colorante marcador es incorporado al producto de reacción por la *AmpliTaq* ADN polimerasa. En este caso son empleados cuatro distintos colorantes fluorescentes como marcadores, en donde un color equivale a cada una de las cuatro bases que forman al ADN. El producto purificado de la reacción de extensión es analizado en un equipo AB 373A Strectch DNA Sequencer, el cual hace uso de una separación electroforética que permite la detección de los productos de ADN marcados por fluorescencia como resultado de

la reacción de secuenciación. Para todas las reacciones de secuenciación se empleó el primer universal M13/pUC (-40) 5'd(GTTTCCAGTCACGAC)3' (AMERSHAM). Como resultado de este análisis se obtuvo una secuencia parcial de 650 bases para cada clona analizada.

---

#### **4.4 IDENTIFICACIÓN Y ANÁLISIS DE LAS SECUENCIAS PARCIALES DE ADN<sub>r</sub> 16S CLONADOS.**

Las secuencias de 650 nucleótidos de cada clona Pz fueron analizadas en la base de datos de secuencias nucleotídicas no redundantes del GenBank empleando el programa BLAST (Altschul, *et al.* 1997), en la página en Internet del National Center for Biotechnological Information (NCBI, <http://www.ncbi.nlm.nih.gov>), para corroborar su naturaleza como secuencia de ADN<sub>r</sub> 16S y su identidad.

**Alineamiento de secuencias.** Todas estas secuencias fueron alineadas con secuencias de referencia, obtenidas de la base de datos GenBank, correspondientes a los organismos identificados en el análisis previo (incluyendo la secuencia de *Sulfolobus acidocaldarius* D14053), como grupo externo. Se realizó un alineamiento empleando 650 nucleótidos de las secuencias de las clonas del pozol obtenidas junto con secuencias referencia en el programa CLUSTAL W versión 1.6 (Thompson *et al.* 1994). Los parámetros de alineamiento empleados fueron los presentados por el programa para un análisis de secuencias estándar.

**Construcción de un árbol filogenético.** Con el resultado del alineamiento se obtuvo una matriz de distancia de la tasa de sustitución de nucleótidos con el algoritmo de Jukes & Cantor (1969) empleando el programa DNADIST del paquete computacional para análisis filogenéticos PHYLIP versión 3.5c (Phylogenetic Inference Package, Felsenstein 1989). El árbol filogenético fue

construido por el método de *neighbor-joining* (Saitou & Nei, 1987) empleando el programa NEIGHBOR del paquete PHYLIP versión 3.5c (Felsenstein, 1989). El programa NEIGHBOR construye un árbol por agrupamiento (*clustering*) sucesivo de linajes, colocando la longitud de las ramas como la unión de linajes. El árbol obtenido por este método no asume un reloj evolutivo y es un árbol no enraizado. El análisis por este programa resulta mucho más rápido que otros del mismo paquete (FITCH o KITSCH) y resulta particularmente útil para el estudio de un gran número de OTU's (Unidad taxonómica operacional). Se aplicó el método de *bootstrapping* para calcular como un criterio de confianza de la topología del árbol obtenida del método de *neighbor joining* ya mencionado (100 replicas), empleando el programa BOOTSTRAP incluido en el paquete Phylip versión 3.5c. Para evitar cualquier tendencia potencial introducida como resultado de la adición de secuencias, la adición de OTU's fue aleatoria en el análisis de *neighbor joining*. El árbol obtenido fue ilustrado con el programa TREE VIEW (Page 1996) y enraizado a *Sulfolobus acidocaldarius*.

**Análisis de posibles secuencias quiméricas.** Todas las secuencias únicas fueron analizadas en el programa CHECK\_CHIMERA de la base de datos RDP (<http://www.cme.msu.edu/RDP/html/index.html>, Maidak *et al.* 2000), para detectar la presencia de posibles artefactos quiméricos resultantes de la amplificación por PCR.

El programa CHECK\_CHIMERA puede ayudar a determinar si la secuencia a analizar por el usuario es una quimera, es decir, si está compuesta de dos fragmentos que poseen una similitud con secuencias distintas a la correspondiente secuencia enviada por el usuario del programa y que están contenidas en la base de datos.

La detección de quimeras intenta identificar tales fragmentos (el método ignora la posibilidad de que una secuencia pueda estar compuesta por más de

dos fragmentos). El programa emplea secuencias no alineadas y se basa en la comparación de oligómeros (7 oligos para secuencias de ADNr 16S y 8 para 23S). El resultado del análisis asiste al usuario para tomar sus propias decisiones y no intenta darle una respuesta definitiva o cuantitativa.

Descripción del método (Maidak *et al.* 2000). Primero se calcula el número de oligos único en común entre la secuencia a analizar y aquella más similar en la base de datos. Este número es designado como **N-Total**. Entonces, por cada décima posición en la secuencia a analizar (llamada **Pos**), se calculan dos números:

- (a) **NI**. El número de oligómeros únicos en común entre la subsecuencia de la posición 1 a **Pos** (Fragmento 1) y las secuencias más similares al Fragmento 1 en la base de datos.
- (b) **N2**. El número de oligómeros únicos en común entre la subsecuencia de la posición **Pos + 1** hasta el final (Fragmento 2) y las secuencias más similares al Fragmento 2 en la base de datos.

Si la suma de **N1 + N2** es mayor a **N total**, entonces uno o ambos fragmentos son similares a algo en la base de datos en lugar que a la secuencia completa. Mientras más grande sea la diferencia es mayor la indicación de que los dos fragmentos –en una posición dada- no son del mismo organismo.

Como resultado del análisis se muestra un histograma, el cual es la suma del valor **N1 + N2 - N** por cada décimas posición de la secuencia. Uno pudiera esperar un verdadero “punto de cambio” (*break point*) cerca de la posición correspondiente al más alto valor, mientras que los valores bajos se esperan que estén lejanos a este punto para una posible secuencia quimérica. Idealmente, uno debería de ver un perfil del histograma con caídas muy suaves a partir del punto de cambio a ambos lados, dando la forma de una curva en forma de campana, tal y como se muestra en el siguiente ejemplo:



**FRAGMENT 1, approximate sequence positions 1 to 120:**

Short-ID	S_ab	Olis	Full name
Cln7,III	-		120Cln7,III 979 bp rRNA RNA 15-JAN-1989, 979 bCln7,III
<u>Dsm.sp.2</u>	.950		530Desulfuromonas sp.
<u>AF019933</u>	.683		1225Geobacter sp. SDB-1 16S ribosomal RNA gene, partial sequence.
<u>AF019929</u>	.683		1249Geobacter sp. Ala-6 16S ribosomal RNA gene, partial sequence.
<u>Gbc.metred</u>	.667		1453Geobacter metallireducens str. GS-15 ATCC 53774 (T).
<u>Dbb.elonga</u>	.658		1383Desulfobulbus elongatus str. FP DSM 2908 (T).
<u>DE16SFP</u>	.658		1383D.elongatus 16S ribosomal RNA.
<u>Drp.vacuo1</u>	.650		1382Desulforhopalus vacuolatus str. Itk10.
<u>Gbc.slfred</u>	.650		1328Geobacter sulfurreducens str. PCA ATCC 51573 (T).
<u>DSU85473</u>	.650		1386Desulfobulbus sp. BG25 16S ribosomal RNA gene, partial sequence.
<u>AF030774</u>	.642		1402Marine snow associated bacterium AdriaticR16 16S ribosomal RNA

**FRAGMENT 2, approximate sequence positions 121 to 400:**

Short-ID	S_ab	Olis	Full name
Cln7,III	-		278Cln7,III 979 bp rRNA RNA 15-JAN-1989, 979 bCln7,III
<u>Mcl.phphov</u>	.734		1360Microlunatus phosphovorius str. NM-2.
<u>Mcl.phpho3</u>	.734		1360Microlunatus phosphovorius str. NM-1.
<u>Mcl.phpho2</u>	.734		1381Microlunatus phosphovorius DSM 10555 (T).
<u>Mcl.sp.Y73</u>	.730		1357Microlunatus sp. str. Y-73.
<u>PFY10819</u>	.694		1416P. freudenreichii subsp. shermanii partial 16S rRNA, 23S rRNA and
<u>Prp.freud4</u>	.694		1416Propionibacterium freudenreichii subsp. shermanii str. E11.1 DSM
<u>Lut.japon3</u>	.683		1365Luteococcus japonicus DSM 10546 (T).
<u>LJZ78208</u>	.683		1365L.japonicus 16S rRNA gene.
<u>D85487</u>	.683		1409Luteococcus japonicus DNA for 16S rRNA.
<u>Lut.japon2</u>	.683		1375Luteococcus japonicus IFO 12422 (T).

**FULL LENGTH SEQUENCE, from position 1 to 400:**

Short-ID	S_ab	Olis	Full name
Cln7,III	-		398Cln7,III 979 bp rRNA RNA 15-JAN-1989, 979 bCln7,III
<u>Tsc.bndign</u>	.575		1391Tessaracoccus bendigoensis str. Ben106 ACM 5119 (T).
<u>PFY10819</u>	.575		1416P. freudenreichii subsp. shermanii partial 16S rRNA, 23S rRNA and
<u>Prp.freud4</u>	.575		1416Propionibacterium freudenreichii subsp. shermanii str. E11.1 DSM
<u>Mcl.sp.Y73</u>	.573		1357Microlunatus sp. str. Y-73.
<u>Mcl.phpho3</u>	.570		1360Microlunatus phosphovorius str. NM-1.
<u>Mcl.phpho2</u>	.570		1381Microlunatus phosphovorius DSM 10555 (T).
<u>Mcl.phphov</u>	.570		1360Microlunatus phosphovorius str. NM-2.
<u>Prf.innocu</u>	.565		1340Propioniferax innocua NCTC 11082 (T).
<u>Prp.freud1</u>	.563		1353Propionibacterium freudenreichii subsp. freudenreichii DSM 20271
<u>Prp.freud3</u>	.560		1403Propionibacterium freudenreichii subsp. freudenreichii DSM 20271
<u>Frd.antarc</u>	.558		1381Friedmanniella antarctica str. AA-1042 DSM 11053 (T).
<u>Lut.japon3</u>	.553		1365Luteococcus japonicus DSM 10546 (T).
<u>LJZ78208</u>	.553		1365L.japonicus 16S rRNA gene.
<u>Prp.freude</u>	.550		1348Propionibacterium freudenreichii subsp. freudenreichii DSM 20271
<u>Lut.japon2</u>	.550		1375Luteococcus japonicus IFO 12422 (T).
<u>D85487</u>	.550		1409Luteococcus japonicus DNA for 16S rRNA.
<u>Prp.avidu2</u>	.545		1397Propionibacterium avidum DSM 4901 (T).
<u>Prp.ropi2</u>	.545		1401Propionibacterium propionicus DSM 43307 (T).
<u>PPAJ3058</u>	.545		1401Propionibacterium propionicus DSM 43307 16S rRNA gene.
<u>PAAJ3055</u>	.545		1397Propionibacterium avidum DSM 4901 16S rRNA gene.

Junto al histograma se muestran también como parte de los resultados resultado tres tablas con la lista de las secuencias más similares en la base de datos. La primera es para el Fragmento 1 (de la posición 1 al punto de cambio hipotético y que están marcadas en color café), la segunda es para el Fragmento 2 (del punto de cambio hipotético al final de la secuencia y marcado en color verde). Y la tercera, para la secuencia completa.

Cada línea en estas tablas contiene cuatro elementos, que de izquierda a derecha son:

- (a) Un identificador usado para una secuencia en la base de datos RDP.
- (b) Un valor de similitud codificado (S\_ab). Este valor es el número de oligos únicos compartidos entre la secuencia analizada y una secuencia dada de la base de datos RDP dividido entre el valor más bajo de oligos únicos en cualquiera de los dos.
- (c) El número de oligos únicos presentes en una secuencia dada (Olig). Si los mismos oligos están presentes más de una vez en una secuencia, se contabiliza solo una vez, por lo que el valor para Olig refleja la longitud de la secuencia. Al contabilizar los oligos únicos, la tendencia de la composición de bases es corregida (por ejemplo, en los insertos que tienden a ser ricos en GC, es muy probable que el mismo oligo rico en GC este presente varias veces. Al contabilizarlo solo una vez, este artefacto es de menor importancia)
- (d) El nombre completo. La línea de definición de la base de datos RDP con formato de archivo del GENBANK.

Resultados del análisis y recomendaciones. El ejemplo señalado anteriormente es ideal ya que en ocasiones existen casos en los que los valores de salida no dan el suficiente soporte para tomar una decisión absoluta. El método puede ser mejorado por varias formas, incluyendo el uso de secuencias alineadas y análisis de su estructura secundaria. En su estado actual, deben de ser considerados los siguientes puntos sobre el programa:

- (a) Si los valores del histograma tienden de forma consistente hacia un máximo y entonces bajan, entonces la secuencia en cuestión es una

- quimera, En la experiencia, el indicador más confiable ha sido esta tendencia hacia valores cada vez mayores, con una posterior caída.
- (b) Si el valor de  $S_{ab}$  para la secuencia completa es significativamente menor que aquellos valores para ambos fragmentos, se tiene entonces, otro indicador positivo.
  - (c) Una secuencia quimérica de una región del ARNr altamente conservada puede presentar pequeños pero consistentes incrementos en los valores, ya que las secuencias no difieren mucho. Se esperarían valores altos si contiene una región variable. Si la secuencia incluye tanto regiones variables como conservadas, el histograma puede presentar un perfil asimétrico y desviarse del ejemplo ideal.
  - (d) No se ha estimado aún un valor real de “línea basal” de valores a partir de los cuales se puede considerar a una secuencia como quimera con una cierta seguridad. Una razón de esto, es que los valores tomados de las secuencias conocidas consideradas como no quimeras varían de forma considerable, y el método “no sabe “ que parte de la estructura del ARNr está cubierta por la secuencia a analizar.

**Titularidad de los programas de cómputo empleados para los análisis de secuencias.** Los análisis de las secuencias parciales de los clones Pz fueron realizados *en línea* a partir de las siguientes direcciones: CLUSTAL W versión 1.6, <http://bioweb.pasteur.fr/seqanal/interfaces/clustalw-simple.html>; PHYLP versión 3.5c <http://sdmc.krdl.org.sg:8080/~lxzhang/phylip/>. El programa TREEVIEW, fue obtenido como *freeware* a partir de la siguiente dirección: [http://bcdnt1.weizmann.ac.il/Services/treeview.html#\\_Toc356614772](http://bcdnt1.weizmann.ac.il/Services/treeview.html#_Toc356614772).

# 5

## RESULTADOS

---

### 5.1 EXTRACCIÓN DEL ADN TOTAL DE LA FRACCIÓN BACTERIANA DEL POZOL POR UN MÉTODO INDIRECTO.

#### **Separación de la fracción bacteriana del pozol y lisis celular.**

La eficiencia del proceso de extracción de la fracción bacteriana de la muestra de pozol tipo indígena analizada fue del 23% (tabla 8). Los controles realizados para evaluar la posible interferencia de algún componente de la masa de maíz mostraron que en una suspensión de *Lactobacillus amylophilus* la eficiencia de lisis celular fue del 92% (control 1), comparado con una lisis del 100% empleando las mismas condiciones con un cultivo de *L. amylophilus* sin masa (control 2).

#### **Extracción del ADN bacteriano del pozol.**

El protocolo de extracción de ADN empleado permitió la obtención de una sola banda de ADN cromosomal la cual después de ser analizada en un gel de agarosa, presentó una migración al nivel del marcador de peso molecular de 21 Kb (figura 9). La muestra se observó libre de plásmidos de bajo peso molecular y de ARN. El ADN obtenido fue incoloro y con la suficiente pureza para ser empleado como templado para amplificación de los ADNr 16S por PCR.

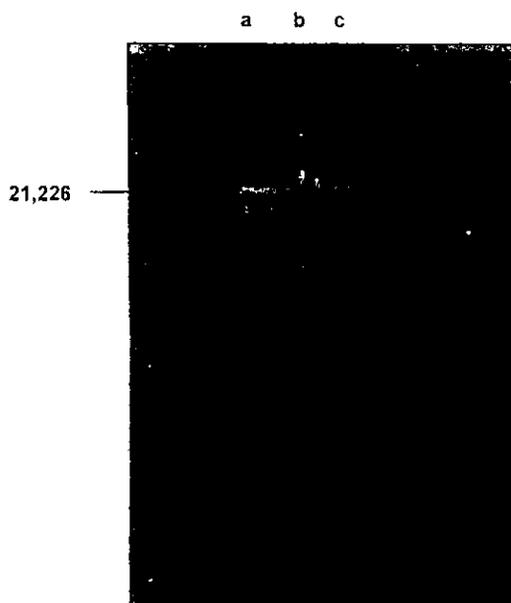
Tabla 8.

Eficiencia del proceso de extracción de células bacterianas del pozol.

Muestra	#UFC/ml
Sobrenadante 1 <sup>a</sup> (S1)	$11 \times 10^6$
Sobrenadante 2 <sup>a</sup> (S2)	$8 \times 10^6$
Sobrenadante 3 <sup>a</sup> (S3)	$2 \times 10^6$
Precipitado 3 <sup>b</sup> (P3)	$7 \times 10^7$
Total recuperado (S1 + S2 + S3) = $21 \times 10^6$ UFC/ml	
No. de microorganismos detectados en la muestra residual (P3) = $7 \times 10^7$ UFC/ml	
$\% \text{ de extracción} = \frac{S1 + S2 + S3}{S1 + S2 + S3 + P3} = 23.07 \%$	

<sup>a</sup> Sobrenadantes recuperados después de cada ciclo de homogeneización/centrifugación.

<sup>b</sup> Precipitado final obtenido después del tercer ciclo de homogeneización/centrifugación.



---

**Figura 9.**

Electroforesis en gel de agarosa 1% de ADN de la fracción bacteriana extraída del pozol. a. Marcadores de PM,  $\lambda$  digerido con *Eco* RI- *Hind* III (en pb); b. ADN de  $\lambda$  sin digerir; c. ADN de la fracción bacteriana del pozol.

---

## **5.2 AMPLIFICACIÓN POR PCR DE LOS ADNr 16S PRESENTES, CLONACIÓN DEL PRODUCTO DE AMPLIFICACIÓN Y DETECCIÓN DE LAS CLONAS CON INSERTO.**

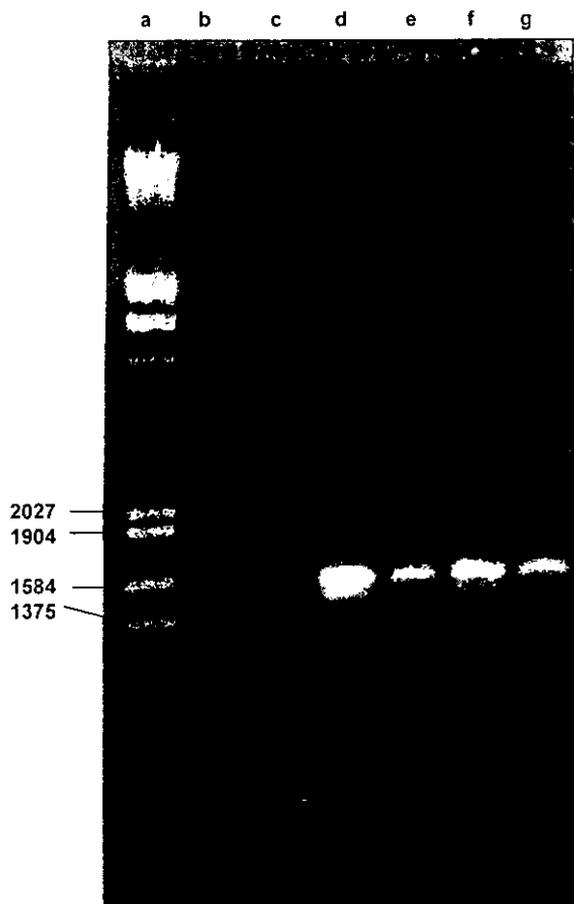
### **Amplificación por PCR de los ADNr 16S.**

La amplificación por PCR de los ADNr 16S presentes en la muestra de ADN bacteriano aislado del pozol bajo las condiciones de reacción empleadas y con el par de *primers* fD1+rD1, resultó en un producto de aproximadamente 1.5 Kb. De acuerdo a lo reportado por Weisburg, *et al.* (1991), corresponde al tamaño promedio de los ADNr's 16S (figura 10)

### **Clonación del producto de amplificación y evaluación de la presencia del inserto.**

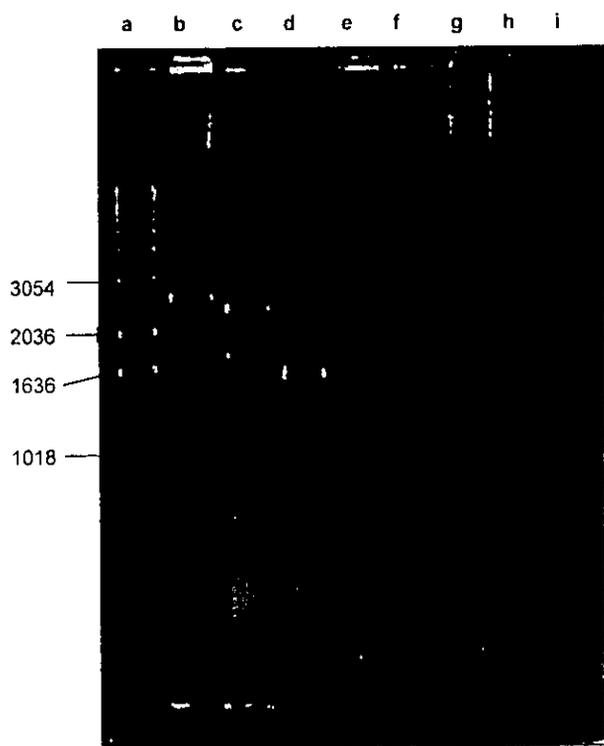
La figura 11 muestra las digestiones enzimáticas de algunas de las minipreparaciones de plásmidos de clonas Pz positivas. Se observa una banda correspondiente al tamaño del plásmido pT7Blue (aproximadamente 2.9 Kb) y una banda correspondiente al tamaño promedio del producto de amplificación de 1.5 Kb aproximadamente.

La figura 12 muestra la amplificación por PCR de los ADNr 16S empleando como templado minipreparaciones de plásmidos con inserto clonado. Se observa una banda correspondiente al tamaño promedio de los insertos clonados (1.5 Kb). Este resultado confirma la clonación del producto de amplificación por PCR y permitió la detección de 36 clonas con inserto.



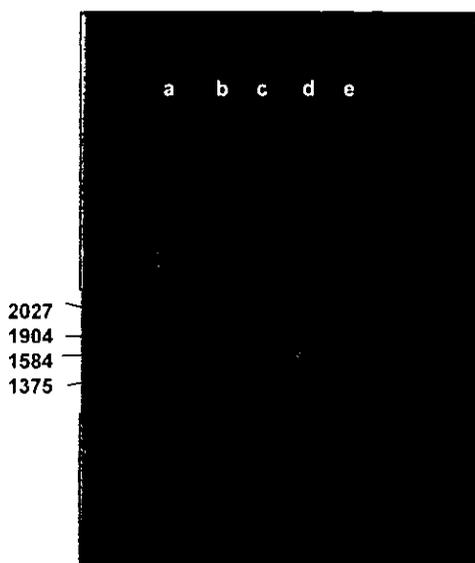
**Figura 10.**

Electroforesis en gel de agarosa 1.5% del los ADNr 16S amplificados por PCR a partir de la fracción de ADN bacteriano extraída del pozol, empleando el par de *primers* fd1+rD1. a. Marcadores de PM,  $\lambda$  digerido con *Eco* RI- *Hind* III (en pb); b. Reacción de PCR sin templado (control negativo); c. Reacción de PCR con ADN cromosomal de *E. coli* como templado (control negativo); d-g. Reacción de PCR empleando ADN de la fracción bacteriana del pozol como templado.



**Figura 11.**

Electroforesis en gel de agarosa 1% de digestiones enzimáticas con *Bam* HI - *Hind* III del plásmido pT7-Blue con el ADNr 16S amplificado por PCR clonado. a, h. Marcador de peso molecular 1Kb; b. Plásmido pT7-Blue; c. Clona Pz314; d. Clona Pz406; e. Clona Pz525; f. Clona Pz322; g. Clona Pz547; i. Clona Pz619. Solo se muestra la digestión de algunas clonas con fines ilustrativos.



**Figura 12.**

Electroforesis en gel de agarosa 1.5% del producto de amplificación por PCR del inserto clonado en el plásmido pT7Blue, empleando como templado ADN plasmídico obtenido de minipreparaciones alcalinas. a. Marcadores de PM,  $\lambda$  digerido con *Eco* RI- *Hind* III (en pb); b. Clona Pz322; c. Clona Pz518; d. Clona Pz619; e. Clona Pz547. Solo se muestra algunas muestras con fines ilustrativos.

---

### 5.3 IDENTIFICACIÓN Y ANÁLISIS DE LA SECUENCIA PARCIAL DE LAS CLONAS Pz.

#### **Identificación de clonas en la base de datos de secuencias nucleotídicas no redundantes del NCBI.**

Como resultado del análisis de las primeras 300 bases de cada una de las 36 clonas con inserto detectadas, se determinó la presencia de 16 clonas únicas. Las 650 bases secuenciadas para cada una de ellas se analizaron en la base de datos de secuencias nucleotídicas no redundantes del NCBI (<http://www.ncbi.nlm.nih.gov>), empleando el programa BLAST. Como resultado de este análisis se determinó que de las 16 clonas únicas analizadas, dos de ellas correspondieron a genes no ribosomales. La tabla 9 muestra la identidad determinada como resultado de este análisis para cada clona única.

De los genes de ADNr 16S clonados, 13 secuencias únicas analizadas corresponden a bacterias lácticas y una a *Clostridium* sp. De las secuencias identificadas como bacterias lácticas, 13 de ellas fueron identificadas como *Lactococcus lactis*, siendo este el grupo bacteriano más abundante en la colección de clonas. Las secuencias identificadas como *Lactobacillus plantarum* y *L. alimentarius*, corresponden al segundo grupo más abundante de clonas presente. *L. delbrueckii* y *Lb. casei* fueron también detectados y así como un gen no ribosomal de *Alcaligenes branchisepticus* y otro de humano. La identidad de la clona Pz507 se consideró como cuestionable debido a que los resultados del análisis en el programa BLAST (Apéndice) mostraron que esta secuencia posee el mismo porcentaje de identidad (de acuerdo al valor de salida del análisis) con las secuencias de ADNr 16S de *Clostridium sordellii* (93%), de una eubacteria no cultivada (93%). Se observó el mismo caso para las secuencias identificadas como *Streptococcus bovis*, las cuales presentan un porcentaje de identidad similar al de *S. suis*.

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

**Análisis de posibles secuencias quiméricas.**

Las secuencias únicas de ADNr 16S identificadas fueron analizadas en el programa CHECK\_CHIMERA de la base de datos RDP (Maidak, *et al.* 2000). Los resultados de este análisis mostraron que las dos únicas secuencias consideradas como posibles quimeras fueron las clonas designadas como Pz522 y Pz612. Este resultado corrobora la naturaleza de estas secuencias como genes no ribosomales.

**Alineamiento de secuencias.**

La figura 13 muestra el resultado del alineamiento hecho en el programa Clustal W versión 1.6 (Thompson, *et al.* 1994) de las clonas Pz. En este análisis se incluyeron secuencias de referencia en la base de datos del GenBank para cada uno de los microorganismos identificados en el análisis en la base de datos nucleotídicos no redundantes del NCBI.

**Construcción de un árbol filogenético y determinación de la filiación taxonómica de las clonas del pozol.**

Los 650 nucleótidos terminales del extremo (5'-) de cada clona única se usaron para realizar un análisis filogenético. La identidad de las clonas Pz obtenidas del análisis con el programa BLAST correlaciona con la posición filogenética de los organismos conocidos y considerados como referencia (figura 14). La relación específica entre las clonas Pz 314 y Pz 634 con relación a las secuencias de referencia de *L. delbruekii* y *L. alimentarius* no fue resuelta en este análisis.

Tabla 9.

Organismos identificados en la biblioteca de clonas Pz.

Organismo identificado <sup>a</sup>	No. (%) de organismos identificados <sup>b</sup>	Designación de clonas únicas <sup>c</sup>	Porcentaje de identidad <sup>d</sup>	Número de acceso al GENBANK
<i>Lactococcus lactis</i>	13 (36.1)	Pz 518 Pz 608 Pz 638 Pz 513 Pz 547 Pz 505 Pz 504 Pz 406 Pz 525 Pz 619 Pz 322 Pz 634 Pz 314 Pz 507 Pz 522	85 96 98 96 97 99 94 95 96 96 95 96 98 93 RNM-	AF095555 AF095544 AF095560 AF095538 AF095542 AF095551 AF095550 AF095537 AF095557 AF095564 AF095561 AF095562 AF095563 AF095565 NS
<i>Streptococcus bovis/suis</i>	3 (8.3)		94	AF095550
<i>Lactobacillus plantarum</i>	6 (16.6)		96	AF095564
<i>Lactobacillus casei</i>	2 (5.6)		95	AF095561
<i>Lactobacillus alimentarius</i>	8 (22.2)		96	AF095562
<i>Lactobacillus delbruekii</i>	1 (2.8)		98	AF095563
<i>Clostridium</i> sp.	1 (2.8)		93	AF095565
Gen no ribosomal de <i>Alcaligenes bronchisepticus</i> (arilmalonato descarboxilasa)	1 (2.8)		RNM-	NS
ADN de origen humano	1 (2.8)	Pz 612	RNM	NS

<sup>a</sup> Organismo que muestra el porcentaje de identidad más alto en los resultados del análisis en la base de datos nucleotídicos no redundantes del NCBI por medio del programa BLAST; <sup>b</sup> Número de secuencias con la misma identidad (porcentaje de secuencias que fueron identificadas como un microorganismo dado en la base de datos) presente en la colección de clonas Pz; <sup>c</sup> Designación de clonas únicas; <sup>d</sup> Valor de identidad de acuerdo al resultado del análisis en el programa BLAST; NS. Secuencia no introducida en la base de datos GENBANK; RNM. Resultado no mostrado.





Pz608 -----  
 Pz505 -----  
 Pz639 -----  
 Pz247 TGCGCGCG -----  
 Streptolacti -----  
 Pz519 -----  
 Pz513 -----  
 Streptobovis GTGAAHTDC -----  
 Pz496 -----  
 Pz406 -----  
 Lbplantaru -----  
 Pz619 -----  
 Lbplimenta -----  
 Pz314 -----  
 Ldelbruec -----  
 Lbcasei -----  
 Pz332 -----  
 Clostridiu GTGAAHTDCGTGAGCTTACGAGAA -----  
 Sulfolobus G-----

Lactococcu, *Lactococcus lactis* X54260; Streplacti, *Streptococcus lactis* M58837; Strepbovis, *Streptococcus bovis* AB002482; Lbplantaru,  
*Lactobacillus plantarum* M58827; Lbalimenta, *Lactobacillus alimentarius* M58827; Ldelbruec, *Lactobacillus delbruekii* X52654; Lbcasei,  
*Lactobacillus casei* D16548; Clostridiu, *Clostridium* sp. X98395; Sulfolobus, *Sulfolobus acidocaldarius* D14053.

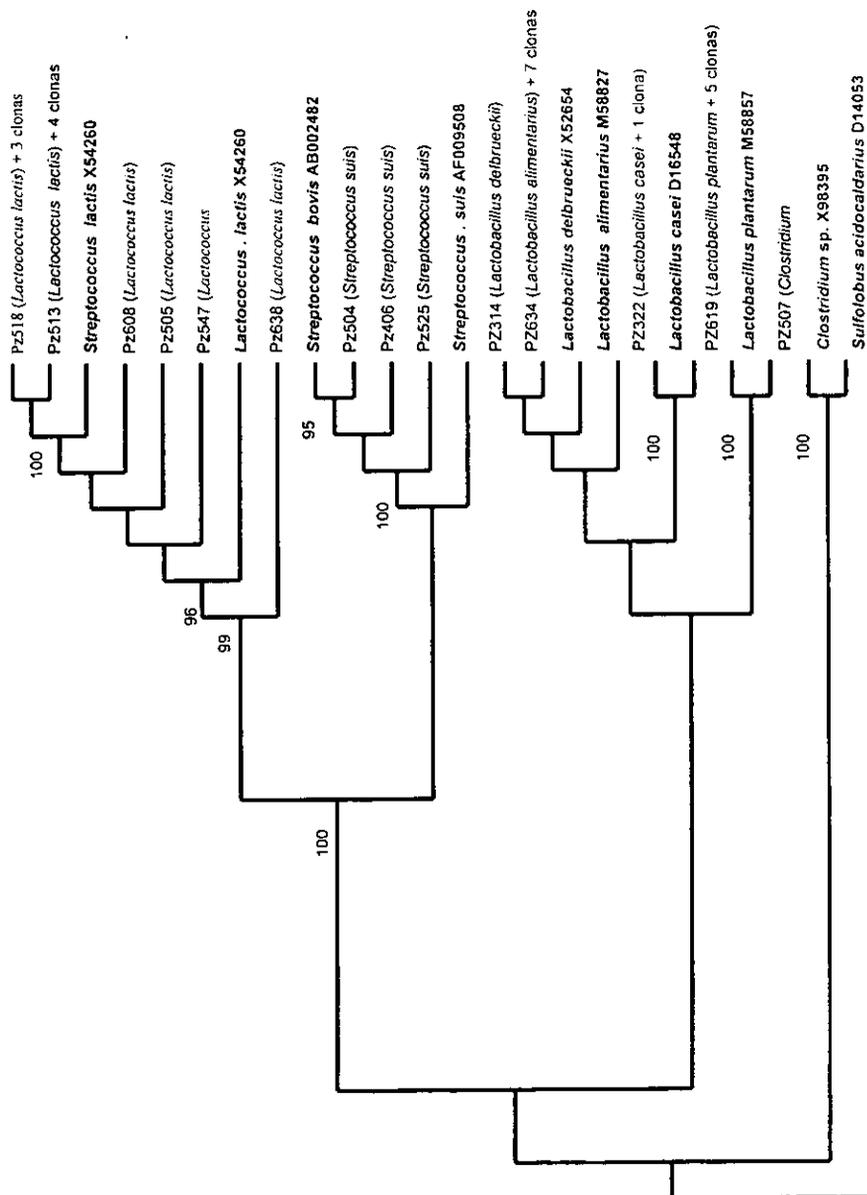


Figura 14.

Árbol filogenético de secuencias parciales de ADNr 16S de las clonas Pz y secuencias de bacterias conocidas y consideradas como referencia. Entre paréntesis se muestra la identidad obtenida. Se señala en negritas el número de accesión de las secuencias de organismos de referencia. El porcentaje resultante de 100 muestreos de *bootstrapping* que soporta cada elemento topológico en el análisis de *neighbour-joining* se indican en la figura. No se muestran valores para grupos con un valor de *bootstrapping* menor de 90%.

# 6

## DISCUSIÓN

---

### 6.1 DIVERSIDAD DE BACTERIAS LÁCTICAS DEL POZOL.

En nuestro grupo de trabajo, la diversidad de las bacterias lácticas del pozol ha sido ampliamente estudiada mediante el uso de técnicas microbiológicas tradicionales y recientemente con el empleo de técnicas moleculares. Algunos de los resultados previos obtenidos muestran la presencia de bacterias lácticas del género *Lactococcus* (originalmente identificadas como *Streptococcus*), *Leuconostoc* y *Lactobacillus* en las primeras horas de fermentación de masas de pozol tipo indígena del Estado de Chiapas (Nuraida, et al. 1995). Posteriormente se identificaron las bacterias lácticas *Lactobacillus plantarum*, *Lactobacillus pentosus*, *Leuconostoc mesenteroides*, *Leuconostoc citreum*, *Lactococcus lactis* y *Lactococcus raffinolactis*, como los microorganismos más abundantes en muestras de pozol, en una colección de 92 cepas procedentes de la ciudad de San Cristóbal de Las Casas y Tapachula, Chiapas (Flores, 1996).

La aproximación metodológica empleada en este trabajo permitió determinar la presencia de *Clostridium* sp. y algunas especies de bacterias lácticas pertenecientes a los géneros *Lactobacillus*, *Streptococcus* y *Lactococcus* en la muestra de pozol tipo indígena estudiada. Los microorganismos identificados correlacionan con las características del alimento: desarrollo de una fermentación láctica asociada a bacterias lácticas (Wacher et al. 1993; 2000) y

condiciones microaerófilas y posiblemente anaerobias presentes en las porciones internas y medias de la masa respectivamente.

El estudio de la diversidad bacteriana presente en un determinado ambiente por medio del empleo de técnicas moleculares que no dependen del cultivo microbiano, se ve influenciada por la aproximación metodológica empleada para este fin. A continuación se señalan algunos aspectos del presente trabajo que pudieron haber influenciado los resultados obtenidos y las conclusiones generadas a partir de su análisis.

La eficiencia de separación de la fracción bacteriana del pozol fue de un 23%, contra un 33% reportado previamente para una muestra de suelo con el empleo de esta técnica de extracción (Holben *et al.* 1988). Es importante señalar que no existen reportes de la aplicación de esta técnica de extracción de la fracción bacteriana en muestras de alimentos, por lo que el reporte de estos autores representa la única referencia para la comparación de la eficiencia del proceso en la muestra de pozol analizada.

Durante el proceso de extracción de la fracción bacteriana del pozol, se separó un 12, 9 y 2% del número de microorganismos considerado como total y determinado por cuenta en placa a partir de cada uno de los sobrenadantes S1, S2 y S3, respectivamente. Aunque el tercer ciclo de extracción permitió la separación de tan solo un 2% del número considerado como total, la cuenta en placa del precipitado 3 (P3) fue de  $7 \times 10^7$  UFC/ml, lo que indica que la cantidad bacteriana residual es aún importante en el P3.

Este resultado puede ser la consecuencia de una co-sedimentación de las células bacterianas junto con las partículas del alimento durante la centrifugación, lo que supone que el proceso de separación diferencial de células bacterianas no permite una extracción exhaustiva de la fracción bacteriana presente, considerando a este resultado como una desventaja del empleo de una metodología indirecta para la extracción del ADN bacteriano (Tabla 4, Wheeler & Sthal, 1996). Sin embargo, Bakken, 1985 y Holben, *et al.* (1988), han sugerido

que con base en la distribución del tamaño celular bacteriano, la técnica empleada podría liberar a todos los tipos bacterianos de una muestra analizada a lo largo de ocho ciclos de homogeneización/centrifugación con igual eficiencia, por lo que la fracción bacteriana liberada después de un solo ciclo es aparentemente representativa de la población bacteriana total, tanto como lo que sería aquella detectada después de tres ciclos de extracción. Holben, *et al.* (1988) señalan también que de acuerdo a la cuenta en placa total detectada (p.e.  $10^6$  UFC por gramo de muestra), el resultado obtenido se puede considerar como satisfactorio considerando que el organismo o grupo bacteriano de interés se encuentra en niveles mínimos de  $10^4$  UFC o más por gramo de muestra.

De acuerdo a lo anterior, en este trabajo se consideró que el empleo de tres ciclos de extracción de células bacterianas fue satisfactorio, ya que en la muestra analizada se detectó una cuenta microbiana en placa de  $10^6$  UFC/g de muestra para bacterias lácticas y mesófilos aerobios, mientras que para enterobacterias se reportó una cuenta en placa de  $10^4$  UFC/g de muestra al inicio de la fermentación (Wacher, *et al.* 2000). Sin embargo, es importante señalar que debido a que no es posible separar la totalidad de la fracción bacteriana presente en la muestra de pozol tipo indígena analizada mediante el empleo de esta técnica extracción / separación diferencial, la diversidad bacteriana detectada en el presente trabajo está determinada por la abundancia relativa de los grupos bacterianos presentes.

La incapacidad de lisar de manera eficiente la fracción bacteriana aislada de una muestra ambiental puede llevar a conclusiones erróneas sobre la estructura microbiana presente. De este modo Zhou *et al.* (1997), mostraron inicialmente que en una biblioteca de clonas ambientales de suelo utilizando ADN extraído con el empleo de un protocolo de lisis con SDS y un tratamiento térmico, únicamente 11% de las clonas identificadas correspondían al grupo de bacterias Gram (+), microorganismos que se han reportado como presentes en gran número para este tipo de ambientes (Felske, *et al.* 1997; 1998). Posteriormente la modificación del protocolo de lisis celular empleado permitió observar un mayor número de microorganismos de este grupo bacteriano.

El protocolo de lisis empleado en este trabajo pretendió asegurar la ruptura de todas las células bacterianas aisladas de la muestra de pozol. Entre los microorganismos identificados en muestras de este alimento por métodos microbiológicos tradicionales, se han detectado a *Bacillus* y bacterias lácticas de los géneros *Lactobacillus*, *Lactococcus* y *Leuconostoc* (Nuraida *et al.* 1995; Flores, 1996; Wacher *et al.* 1993; 2000), organismos cuya pared celular y presencia de exopolisacáridos en el caso de los últimos, hace difícil el procedimiento de lisis celular. Para el grupo de bacterias lácticas, algunos protocolos de lisis emplean tiempos largos de incubación con enzimas líticas, incluso hasta más de 8 horas (LeBourgeois *et al.* 1989; Klijn *et al.* 1991b), o bien el empleo de condiciones desnaturalizantes de la estructura celular tales como ebullición de una suspensión de células en baño de agua y enfriamiento en hielo (Erhman *et al.* 1994).

La evaluación de la eficiencia del protocolo empleado para lisar la fracción bacteriana aislada del pozol se realizó lisando un cultivo de *L. amilophylus* con una densidad celular conocida (controles 1 y 2), obteniéndose eficiencias de lisis de 92% y 100% respectivamente. Estos resultados permitieron asegurar la lisis de microorganismos con características estructurales similares a las del control (bacterias lácticas, Gram (+)), tal y como se observó en el tipo de microorganismos identificados en la biblioteca de clonas Pz.

El uso de secuencias de 650 bases de las clonas Pz, permitió determinar la identidad de cada una de las clonas analizadas tomando como marco de comparación las secuencias reportadas y depositadas en la base de datos de secuencias nucleotídicas no redundantes del NCBI. Este primer análisis permitió conocer la diversidad de bacterias lácticas presente en la muestra de pozol analizada. Sin embargo, el análisis de secuencias parciales de ADNr 16S puede generar una topología diferente al momento de construir una filogenia respecto a aquella obtenida con una mayor cantidad de información. El tratar de explicar la presencia de un microorganismo desconocido o considerado como no cultivable, puede llevar a conclusiones erróneas. Así de este modo, cuando se emplean únicamente partes de la molécula, existe la posibilidad de que la información

contenida genere cambios regionales en un determinado grupo -*cluster*-filogenético que puede influenciar de forma importante la topología global del árbol, mientras que la información filogenéticamente valiosa almacenada en otra parte de la molécula no se encuentra disponible para el análisis (Schleifer & Ludwig, 1994; Pace 1996; Hilali *et al.* 1997).

Diversos autores han empleado secuencias parciales y/o completas para la construcción de filogenias en el análisis de la estructura microbiana de distintos ambientes (tabla 10), logrando determinar la presencia de secuencias identificadas como pertenecientes a microorganismos no cultivados previamente o como no cultivables. En estos casos, después de obtener la colección de datos (secuencias), éstas son alineadas con secuencias consideradas de referencia y disponibles en diversas bases de datos, lo que hace que muchos errores del proceso de secuenciación, tales como compresiones o rearrreglos puedan ser detectados. Después del alineamiento, el siguiente paso es la estimación de la distancia existente entre las secuencias alineadas empleando diversos algoritmos y finalmente la construcción del árbol filogenético. Es importante considerar que este tipo de estudios únicamente permiten ubicar rápidamente clonas ambientales desconocidas en taxones superiores, por lo que el uso de las primeras 500 bases es recomendado para asignar la posición a una línea principal de descendencia (Devereux & Willis, 1995; Stackenbrandt & Rainey, 1995).

Para evaluar el impacto del empleo de secuencias parciales y completas de ADNr 16S en la topología de un árbol filogenético obtenido en este trabajo, se construyeron cuatro árboles empleando respectivamente 450 bases, 650 bases, 1000 bases y la secuencia completa ( $\cong$  1500 bases) a partir de secuencias conocidas y depositadas en la base de datos del GENBANK. Para la obtención del alineamiento de las secuencias, las matrices de distancia y la construcción de los respectivos árboles, se realizaron los mismos análisis computacionales descritos ya para el análisis de las secuencias de las clonas Pz. Los resultados obtenidos se muestran en la figura 15.

Tabla 10.

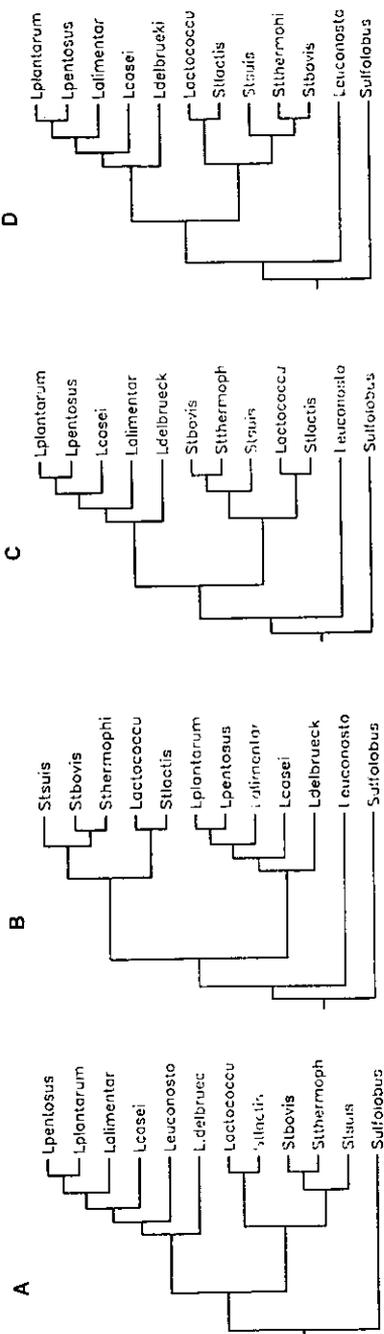
Trabajos en los que se emplean secuencias parciales y completas de ADNr para establecer la filogenia y ubicación taxonómica de clonas ambientales en estudios de determinación de diversidad microbiana.

Grupo bacteriano estudiado	Origen de la muestra analizada	Longitud del ADNr analizada (nt)	Referencia.
Archaea	Cepas tipo	350 y 950	Lane, <i>et al.</i> 1985
Enterobacterias	Cepas tipo	650 - 880	Böttger, <i>et al.</i> 1989
Archaea	Agua salada	1420	Arahal, <i>et al.</i> 1996
Proteobacteria, Grupo <i>Cytophaga-Flexibacter-Bacteroides</i>	Suelo	239 - 254	Bomeman, <i>et al.</i> 1996.
Gram (+)	Turba	1329 - 1373	Rheims, <i>et al.</i> 1996
Actinobacterias	Suelo	1100	Felske, <i>et al.</i> 1997
Gram (+). Proteobacterias, y Grupo de <i>Fibrobacter</i> , <i>Planctomyces</i> , <i>Spirochaetes</i> y <i>Bacteroides</i>	Suelo	214 - 475	Zhou, <i>et al.</i> 1997
Gram (+). Proteobacterias y <i>Nitrospora</i> y <i>Nitrobacter</i> .	Licor de un bioreactor oxidante de nitritos	1030	Burrell, <i>et al.</i> 1998.
Gram (+). Proteobacterias	Suelo	1500	Felske, <i>et al.</i> 1998
Gram (+)	Suelo	400	Nüsslein, <i>et al.</i> 1998

Este análisis permitió determinar que la topología de cada árbol cambia de acuerdo a la longitud de las secuencias empleadas. La secuencia que menos información proporciona es la de 450 bases, mientras con 650 bases la topología del árbol presenta variaciones pero mantiene con menores cambios la posición de las secuencias analizadas de acuerdo a lo que se observa en los árboles construidos con 1000 bases y la secuencia completa.

En el árbol obtenido empleando el algoritmo de *neighbour joining*, las clonas Pz505, Pz513, Pz518, Pz547, Pz608 y Pz638 que fueron identificadas como *L. lactis* se ubican en el mismo grupo (*cluster*) de las secuencias de referencia *Streptococcus lactis* y *Lactococcus lactis* X54260 (*S. lactis* ha sido reclasificado como *L. Lactis*, Vandamme, *et al.* 1995, aunque aún aparece con este nombre en la base de datos del GENBANK). Las clonas Pz406, Pz504 y Pz525, identificadas como *Streptococcus bovis/suis* se ubicaron en el mismo *cluster* de la secuencia de referencia de estos microorganismos. La clona Pz507 se ubicó en un mismo *cluster* junto a la secuencia de referencia de *Clostridium sp.* X98395, aunque la identidad de esta clona se puede considerar como ambigua, ya que de acuerdo a los resultados del análisis en la base de datos nucleotídicos no redundantes del NCBI, ésta clona presentó el mismo porcentaje de identidad (93%) que una eubacteria no cultivable.

Para el caso de las especies del género *Lactobacillus* se observa que el análisis de las secuencias realizado agrupa en el árbol obtenido por *neighbour joining* en un mismo *cluster* a *L. delbruekii* con *L. alimentarius*. Este resultado difiere del esquema de clasificación establecido para este género en el que se ubica en el grupo filogenético A (homofermentativos obligados) a *L. delbruekii*, mientras que *L. alimentarius*, *L. casei* y *L. plantarum* se ubican en este esquema de clasificación en el grupo filogenético B (heterofermentativos facultativos). Este resultado cuestiona que la topología del árbol obtenido por NJ refleje la posición taxonómica real de este grupo de microorganismos.

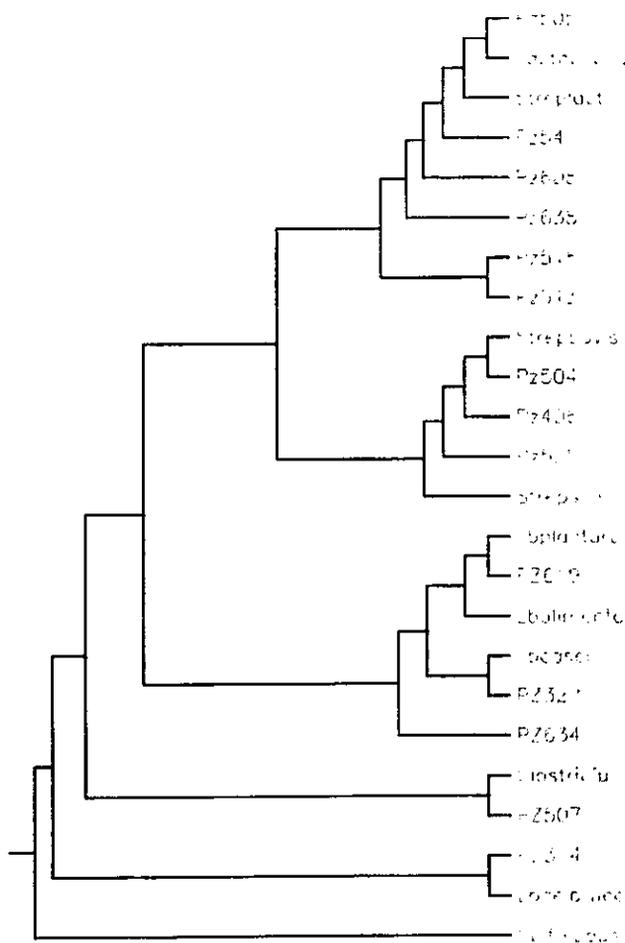


**Figura 15**  
 Árboles filogenéticos contruidos a partir del análisis de 450 bases (A); 650 bases (B); 1000 bases (C) y a partir de la secuencia completa (D). Todas las secuencias fueron obtenidas a partir de la base de datos del GENBANK y sometidas al mismo análisis computacional que las clonas Pz. Sulfolobus, *Sulfolobus acidocaldarius* D14053; Lactococcus; *Lactococcus lactis* X54260; Sthbovis, *Streptococcus bovis* AB002482; Ssthisuis, *Streptococcus suis* AF009508; L.delbruec, *Lactobacillus delbrueckii* X52654; Lplantarum, *Lactobacillus plantarum* M58827; Ssthermoni, *Lactobacillus alimentarius* M58804; Sthlactis, *Streptococcus lactis* M58837; Spentosus *Lactobacillus pentosus* D79211; Leuconosto, *Leuconostoc mesenteroides* AB023243; Lcasei, *Lactobacillus casei* D16553; Ssthermoph, *Streptococcus thermophilus* X68418.

Para evaluar si la topología observada en el árbol obtenido por NJ y la ubicación taxonómica de las clonas Pz se ve influenciada por el algoritmo para la construcción del árbol (NJ), se obtuvo un árbol empleando la misma matriz de distancia (JK) por el método de UPGMA. El resultado que se muestra en la figura 16, permite observar que el cluster de *Lactococcus lactis* se distribuyen todas clonas Pz identificadas como *Lactococcus* aunque la topología de este grupo varía respecto a lo que se observa en el árbol obtenido por NJ. El cluster de *Streptococcus* mantiene la misma topología que en el árbol de NJ y también permite separar a las especies de lactobacilos en clusters que correlacionan con el esquema de clasificación de grupos filogenéticos/patrón de fermentación. Este resultado permite concluir que la topología del árbol se ve influenciada por el algoritmo utilizado (NJ y UPGMA en este trabajo), y señala la necesidad de contar con información adicional sobre los microorganismos identificados en la colección de clonas Pz para corroborar la información obtenida del análisis de secuencias parciales de ADNr 16S (Page & Homes, 1999). Esta situación se puede considerar como crítica ya que puede llevar a conclusiones y suposiciones erróneas.

La presencia de genes no ribosomales de *Alcaligenes branchisepticus* y de humano como producto de amplificación por PCR, muestra que el par de primers empleados en este trabajo pueden generar una amplificación inespecífica hacia otros genes distintos a los ANDr 16S de los grupos bacterianos reportados como blanco (Weisburg *et al.* 1991). Esta inespecificidad ha sido observada previamente en análisis de diversidad microbiana de suelos (Borneman *et al.* 1996), en donde se detectó la presencia de ADNr de cloroplasto, de genes mitocondriales de insecto y de genes no ribosomales de humano con el empleo de primers con una especificidad teórica hacia genes ribosomales 16S de microorganismos Gram (+).

Estos autores reportaron también que, a pesar de la amplificación de estos genes no ribosomales, no se detectó la presencia de otros microorganismos como *Pseudomonas*, mientras que el análisis de la muestra analizada por cuenta en placa mostró la presencia de este microorganismo en una proporción del 3 al 15% del total observado. Con el par de primers fD1+rD1 empleado en este trabajo



**Figura 16**

Arbol construido por UPGMA a partir del análisis las secuencias de las clonas Pz y secuencias de referencia obtenidas de la base de datos del GenBank. *Sulfolobus*, *Sulfolobus acidocaldarius* D14053; *Lactococcus*; *Lactococcus lactis* X54260; *Streptococcus*, *Streptococcus bovis* AB002482; *Strepsuis*, *Streptococcus suis* AF009508; *L.delbuec*, *Lactobacillus delbrueckii* X52654; *L.plantaru*, *Lactobacillus plantarum* M58827; *L.alimenta*, *Lactobacillus alimentarius* M58804; *Stlactis*, *Streptococcus lactis* M58837; *L.casei*, *Lactobacillus casei* D16553; *Clostridriu*, *Clostridium* sp. X98395

no se detectó la presencia de genes (tanto ribosomales como no ribosomales) de maíz o cloroplasto, considerando la naturaleza de la materia prima empleada para la elaboración del pozol.

Se ha reportado en estudios de diversidad microbiana de muestras ambientales realizados mediante el análisis de secuencias de ADNr 16S, que la selección de secuencias amplificadas es un proceso aleatorio durante los procedimientos de ligación y transformación, por lo que aquellas secuencias consideradas como *raras* y pertenecientes a microorganismos representados en una baja proporción respecto a la microbiota dominante, tienen una posibilidad real baja de aparecer como un inserto clonado en la biblioteca de clones obtenida (Felske *et al.* 1997).

De acuerdo a lo anterior, la diversidad bacteriana detectada en este trabajo se puede considerar como representativa de los microorganismos más abundantes presentes en la muestra de pozol analizada. Si se considera que se detectó una cuenta total en placa de  $10^6$  a partir de los sobrenadantes S1, S2 y S3 en el proceso de extracción de la fracción bacteriana, es probable que en dichos sobrenadantes estuvieran presentes de forma abundante los microorganismos identificados. Si se considera también que las secuencias obtenidas fueron recuperadas aleatoriamente de una poza de ADNr 16S es posible suponer entonces que a partir de la muestra de ADN, aquellos microorganismos más abundantes son los que tendrían una mayor probabilidad de ser amplificados por PCR y posteriormente clonados.

Algunos autores han reportado que existe una correlación entre las clones únicas prevalentes y la estructura microbiana de una muestra ambiental (Borneman *et al.* 1996; Felske *et al.* 1997). Si se asume que tanto los microorganismos cultivables como los no cultivables son liberados con la misma eficiencia a partir de la muestra a analizar, la presencia de microorganismos considerados como no cultivables o no cultivados es nula en el pozol, de acuerdo a los resultados obtenidos en este trabajo, y a diferencia de lo observado en

análisis de suelos, agua o ambientes extremos (Barns *et al.* 1994; Borneman *et al.* 1996; Zhou *et al.* 1997; Felske *et al.* 1998).

La aproximación molecular empleada en este trabajo permitió observar que las secuencias identificadas como *Lc. lactis* (13 clonas con 6 filotipos únicos) corresponden al microorganismo más abundante, seguido de *L. alimentarius* y *L. plantarum* (8 y 5 clonas, respectivamente, representadas por un filotipo único). *Strep. bovis/suis* (3 filotipos únicos) fue el tercer organismo más abundante. La diversidad de bacterias lácticas detectada permite señalar que estos microorganismos juegan un importante papel en la fermentación del pozol. Sin embargo, y a pesar de que *Lactococcus lactis* fue el microorganismo más abundante, existen algunos reportes que señalan a otros microorganismos como la microbiota dominante en diferentes muestras de pozol.

Nuraida *et al.* (1995) y Flores (1996) determinaron que *Lactococcus lactis*, *Lactobacillus plantarum*, *L. pentosus* y *Leuconostoc mesenteroides* fueron los microorganismos más abundantes aislados de diferentes muestras de pozol de la localidad de San Cristóbal de las Casas, Chiapas, mientras que Ampe *et al.* (1999) y Ben Omar & Ampe (2000) señalan que *S. bovis* comprende del 25 - 75% de la población bacteriana total activa de una muestra de pozol procedente del Estado de Tabasco.

En el presente trabajo no se detectó a *Leuconostoc mesenteroides* ni a *Lactobacillus pentosus*, identificados como dos de los microorganismos más abundantes en un estudio previo (Flores, 1996). Se ha determinado que *Leuconostoc mesenteroides* es el microorganismo más abundante al inicio de la fermentación y que esta bacteria no tolera la acidez que alcanza la masa durante el desarrollo de la fermentación láctica (Wacher, *et al.* 2000). Si se considera que la muestra de pozol analizada contaba con un tiempo fermentación de 8 - 10 horas, es probable que la acidez generada haya determinado la desaparición de este grupo bacteriano al momento de realizar la extracción del ADN de la fracción

bacteriana. La ausencia de *Lactobacillus pentosus* en la colección de clonas Pz, y el hecho de que *L. plantarum* es el cuarto microorganismo más abundante detectado en la biblioteca de clonas Pz, permite sugerir que *L. pentosus* no constituye un grupo abundante en la muestra de pozol tipo indígena analizada.

*Lactococcus lactis*, comprende el 36% de los microorganismos identificados a partir de la biblioteca de clonas Pz. A diferencia de los resultados de Ampe *et al.* (1999) y Ben Omar & Ampe (2000) en los que señalan que *Lc. lactis* solo comprende menos del 5% de la microbiota activa de la muestra de pozol analizada. Ampe *et al.* (1999) y Ben Omar & Ampe (2000) observan también que *L. delbruekii*, *L. casei* y *L. plantarum* solo están presentes en la muestra de pozol después de 48 a 72 horas de fermentación, mientras que en el presente trabajo, se detectó a estos microorganismos en la muestra de pozol analizada con 8-10 horas de fermentación.

La biblioteca de clonas Pz comprende 36 clonas que representan 12 diferentes tipos de ADNr 16S únicos de bacterias lácticas, 2 genes no ribosomales y 20 secuencias redundantes. Otros reportes sobre el análisis de bibliotecas de clonas de ADNr 16S ambientales han encontrado una redundancia de secuencias que va de grande a pequeña o nula (Barns *et al.* 1994; Borneman *et al.* 1996; Felske *et al.* 1997; 1998). Estos resultados han sido interpretados como un indicador de la diversidad bacteriana relativa (Felske *et al.* 1998). La biblioteca de clonas Pz contiene una redundancia de secuencias alta, lo que puede indicar una diversidad bacteriana limitada, causada por influencias selectivas del ambiente (Felske *et al.* 1998). El claro dominio de *Lc. lactis*, *L. plantarum* y *L. alimentarius*, así como el limitado número de taxones es evidente en los resultados obtenidos. Si se comparan estos resultados con los obtenidos por Nuraida, *et al.* (1995) y Flores (1996), por métodos microbiológicos tradicionales, se observa que las bacterias lácticas son responsables de la fermentación del pozol independientemente de la región geográfica de la cual se colectó la muestra y del tipo de elaboración empleado, pero que no es constante el tipo de bacterias lácticas presente en dichas muestras.

Se ha reportado una gran redundancia de clonas de ADNr 16S en ambientes ácidos (suelos y turbas) (Borneman *et al.* 1996; Rheims *et al.* 1996), y se ha propuesto como explicación en estos casos que las condiciones ácidas de tales ambientes son causantes de una presión selectiva que tienen como resultado la redundancia de clonas observadas. Para el caso del pozol, durante la fermentación de la masa de maíz el pH varía significativamente de un valor cercano a 7 al inicio de la fermentación a un pH de 3.0 al final, como consecuencia del metabolismo ácido láctico. Esta situación podría explicar la gran redundancia observada en la biblioteca de clonas Pz.

---

## **6.2 IMPLICACIONES BIOTECNOLÓGICAS DEL ESTUDIO DE LA DIVERSIDAD BACTERIANA DEL POZOL POR MÉTODOS MOLECULARES.**

La biotecnología ha sido descrita como la aplicación de principios científicos e ingenieriles al procesamiento de materias primas para la provisión de bienes y servicios, mediante el uso de sistemas y agentes biológicos. En un sentido estricto, la biotecnología se originó a partir de las fermentaciones tradicionales de alimentos. Con el paso del tiempo, esta práctica pionera ha sido expandida y mejorada de tal forma que los microorganismos y otros agentes biológicos han encontrado aplicaciones muy importantes en otras áreas. Los avances recientes en la genética, enzimología, tecnología del ADN recombinante y tecnología de fermentaciones han permitido que los avances en biotecnología estén mucho más allá del contexto tradicional original.

En la actualidad, en muchos países en desarrollo se emplean aún las viejas técnicas artesanales para el procesamiento de alimentos. Estos países no pueden permanecer ajenos a los beneficios de la biotecnología moderna. Debido a la importancia potencial de muchos alimentos fermentados en diferentes contextos, se establecieron las siguientes prioridades de investigación en torno a

ellos: (1) identificar los microorganismos responsables de la fermentación y microbiota indeseable; (2) mejorar la comprensión de los procesos de fermentación; (3) simplificar y estandarizar el proceso y (4) desarrollar capacidades tecnológicas locales. Se debe dar mayor énfasis a aquellos alimentos fermentados que sirven como fuente principal de nutrimentos a grandes poblaciones o grupos étnicos de importancia (Gaden, *et al.* 1992).

En este contexto, el estado actual del estudio del pozol se puede ubicar en la primera prioridad de investigación señalada. El interés por parte de nuestro grupo de trabajo es el estudio de la microbiología y fisiología de las bacterias lácticas del pozol, un alimento fermentado tradicional de importancia regional en el cual se desarrolla una comunidad microbiana. Los resultados obtenidos en este trabajo pretenden complementar los estudios de la diversidad microbiana presente en muestras de pozol, mediante una aproximación polifásica que incluye también el uso de técnicas microbiológicas tradicionales, que han permitido identificar la microbiota presente y también permiten considerar a ésta como constante independientemente del origen de la muestra.

Las características del pozol como un ambiente microbiano complejo (diversidad de microorganismos y microambientes), hacen que este alimento sea un sistema a partir del cual sea posible describir diferentes procesos microbianos de potencial interés en biotecnología (figura 17).

El presente trabajo es el primer ejemplo de la aplicación del análisis de secuencias parciales de ADNr 16S obtenidas a partir de la fracción de ADN bacteriano extraída de una muestra de alimento fermentado tradicional, como parte de una aproximación al estudio de la estructura microbiana del pozol para determinar la presencia de microorganismos con características genéticas, fisiológicas y ecológicas de potencial importancia biotecnológica. La diversidad de bacterias lácticas presente en una muestra de pozol tipo indígena y descrita en este trabajo, muestra algunas especies de este grupo que no habían sido detectadas previamente por métodos microbiológicos tradicionales. Por otro lado,

no se identificaron microorganismos considerados como no cultivables o desconocidos en la muestra analizada.

A pesar de las ventajas que ofrecen este tipo de metodologías para el estudio de la estructura microbiana de un alimento fermentado tradicional, el punto de partida obligado para iniciar la investigación sobre la importancia de cada uno de los microorganismos presentes y la posibilidad de realizar una explicación de su papel en el desarrollo de las características de este alimento, es el aislamiento y el cultivo puro de la cepa o cepas a analizar. Lo que para el caso del pozol resulta de gran importancia, ya que una primera meta en el estudio de la diversidad bacteriana de este alimento es el establecimiento de la microbiota responsable del desarrollo de las características del alimento; otra es identificar a los microorganismos indeseables con la finalidad de optimizar su proceso de elaboración.

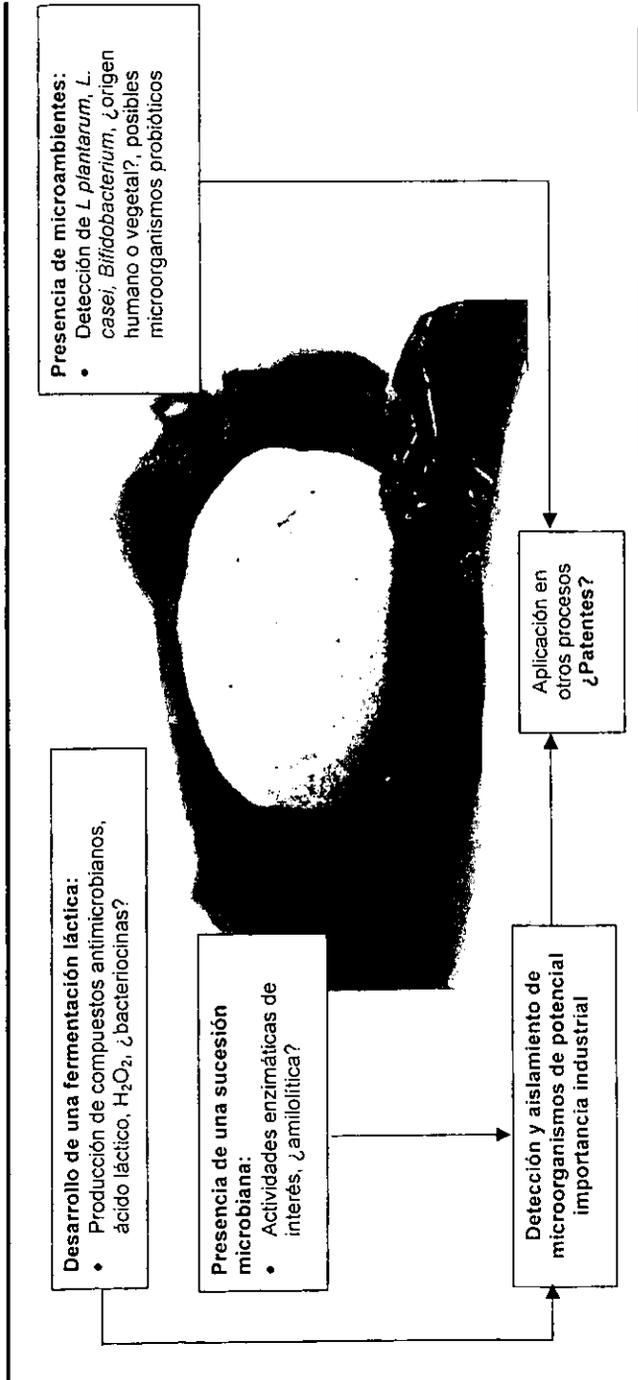


Figura 17. Procesos microbianos presentes en el queso con posibles implicaciones biotecnológicas a partir de microorganismos específicos.

# 7

## CONCLUSIONES y PERSPECTIVAS

---

El presente trabajo se engloba en una línea de investigación principal que pretende el estudio de la diversidad bacteriana del pozol. La aproximación molecular empleada en esta investigación pretende implementar el uso de técnicas no dependientes del cultivo microbiano para el análisis de la diversidad bacteriana presente en un alimento, tomando como modelo para evaluar la factibilidad del uso de éstas técnicas una muestra de pozol tipo indígena con un tiempo de fermentación de 8 – 10 horas.

La información obtenida mediante la aplicación de estas técnicas para el análisis de este alimento tiene la finalidad de ampliar el conocimiento sobre la diversidad bacteriana del pozol que ha sido obtenida por métodos microbiológicos tradicionales. De este modo, las principales conclusiones que se generan de este trabajo y que aportan una serie de conocimientos, no sólo sobre la diversidad de bacterias lácticas de una muestra de pozol tipo indígena sino también en forma de una serie de experiencias metodológicas a nuestro grupo de trabajo, son las siguientes:

1. En este trabajo se reporta por primera vez el análisis de la diversidad de bacterias lácticas de un alimento fermentado tradicional, en este caso una muestra de pozol tipo indígena con un tiempo de fermentación de 8-10 horas,

mediante el análisis de secuencias parciales de ADNr 16S amplificadas por PCR a partir del ADN bacteriano extraída de una muestra de este alimento.

2. Las secuencias de ADNr 16S analizadas fueron identificadas como bacterias Gram(+). Este resultado confirma la especificidad del par de *primers* fD1 + rD1 hacia bacterias Gram(+) y relacionadas, aunque se identificaron también dos genes no ribosomales.
3. Las secuencias obtenidas permitieron detectar la presencia de las bacterias lácticas *Lactobacillus alimentarius*, *L. delbruekii*, *L. plantarum*, *L. casei*, *Lactococcus lactis*, *Streptococcus bovis/suis*, y *Clostridium* sp.
4. Las especies *Lb. alimentarius*, *Lb. delbruekii* y *S. Bovis/suis* son reportadas por primera vez en una muestra de pozol tipo indígena.
5. La abundancia de clonas únicas detectada permite señalar a las especies *Lactococcus lactis*, *L. alimentarius* y *L. plantarum* como los microorganismos con una mayor abundancia relativa en la muestra de pozol tipo indígena analizada con 8-10 horas de fermentación
6. No se detectaron secuencias de ADNr 16S de microorganismos considerados como no cultivables o desconocidos en la muestra de pozol.

A partir de estas conclusiones se pueden señalar las siguientes recomendaciones que podrán permitir un mejor análisis de la estructura microbiana del pozol:

1. Emplear una menor cantidad de muestra de pozol a analizar, que sin embargo sea representativa de cada uno de los microambientes del pozol (zona interna, media y externa), evaluando el efecto del tamaño de muestra sobre la eficiencia de extracción celular y del ADN.

2. Desarrollar y evaluar el efecto del empleo de una estrategia de extracción directa de ADN del pozol sobre el estudio de la diversidad microbiana en este alimento, comparando con lo obtenido en este trabajo.
3. Determinar la diversidad de bacterias Gram(-) en el pozol, empleando una combinación de *primers* que incluya alguno de los empleados en este trabajo (fD1 o rD1), con otro específico para bacterias Gram(-) y en particular enterobacterias, segundo grupo bacteriano más abundante en el pozol después de las bacterias lácticas.
4. Determinar las diferencias en la estructura microbiana entre muestras de pozol de diferentes tipos (tipo indígena o tipo mestizo basado en la utilización de diferentes tipos de maíz) y diferentes regiones, considerando las observaciones señaladas en los puntos 1-3.
5. Realizar el seguimiento de la fermentación del pozol mediante el empleo de la aproximación metodológica utilizada en este trabajo y complementar esta información con otras técnicas la hibridación fluorescente *in situ* (FISH, por sus siglas en inglés) y complementar esta información con los datos obtenidos por métodos microbiológicos tradicionales. Todo esto, para evaluar bajo un criterio polifásico y en términos cuantitativos la diversidad microbiana del pozol y su variación durante el desarrollo de una fermentación.

La continuación de esta línea de investigación, bajo las recomendaciones señaladas, permitirá tener un panorama completo de la diversidad bacteriana del pozol mediante el empleo de métodos moleculares. Esta información permitirá establecer el tipo de microorganismos dominantes en muestras de pozol de diferentes tipos y tanto con el aislamiento selectivo de estos microorganismos por técnicas microbiológicas tradicionales permitirá la posibilidad de controlar las condiciones de fermentación de este alimento.

# 8

## REFERENCIAS

- 
- Agati V., Morlon-Guyot J.P., Talamond P. & Hounhouigan D.J. (1998) Isolation and characterization of new amylolytic strains of *Lactobacillus fermentum* from fermented maize doughs (mawe and oig) from Benin. *Journal of Applied Microbiology*. **85**: 512-520.
- Agersborg A., Dahl R. & Martinez I. (1997) Sample preparation and DNA extraction procedures for polymerase chain reaction identification of *Listeria monocytogenes* in seafoods. *International Journal of Food Microbiology*. **3**:275-280.
- Ahmad K. A. & Stewart G. S. A. (1991) The production of bioluminescent lactic acid bacteria suitable for the rapid assessment of starter culture activity in milk. *Journal of Applied Bacteriology*. **70**: 113-120.
- Akkermans D. L., Mirza M. S., Harmsen H. J. M., Blok H. K., Herron P. R., Sessitch A. & Akkermans W. M. (1994) Molecular ecology of microbes: a review of promises, pitfalls and true progress. *FEMS Microbiology Reviews*. **15**: 185-194.
- Altschul S. F., Madden T. L., Schäffer A. A., Zhang J., Zhang Z., Miller W. & Lipman D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Resesarch*. **25**: 3389-3402.
- Ampe F., Ben Omar N., Moizan C., Wachter C. & Guyot, JP. (1999). Polyphasic study of the spatial distribution of microorganisms in Mexican pozol, a fermented maize dough, demonstrates the need for cultivation-independent

- methods to investigate traditional fermentations. *Applied and Environmental Microbiology*. **12**: 5464-5473.
- Arahal D. R., Dewhirst F. E., Paster B. J., Volcani B. E. & Ventosa A. (1996) Phylogenetic analysis of some extremely halophilic Archaea isolated from Dead Sea water, determined on basis of their 16S rRNA sequences. *Applied and Environmental Microbiology*. **62**: 3779-3786.
- Ausubel F., Brent R., Kingston R., Moore D., Seidamn J. G., Smith J. A. & Struhl K. (Eds.) (1992) Short protocols in molecular biology. 2nd. Ed. John Wiley & Sons. New York, USA.
- Bakken L. R. (1985) Separation and purification of bacteria from soil. *Applied and Environmental Microbiology*. **49**: 1482-1487.
- Barns S. M., Fundyga R. E., Jeffries M. W. & Pace N. R. (1994) Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proceedings of the National Academy of Science USA*. **91**: 1609-1613.
- Bateson M. M. & Ward D. M. (1996) Methods for extracting DNA from microbial mats and cultivated micro-organisms: high molecular weight DNA from French press lysis. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.1.4**:1-7.
- Ben Omar N. & Ampe, F. (200) Microbial community dynamics during production of the Mexican fermented maize dough pozol. *Applied and Environmental Microbiology*. **66**: 3664-3673.
- Boone D.; Castenholz R.; Colwell R.; Dazzo F.; Devereaux R.; Forney L.; Frederick R.; Gherna R.; Ghiorse W.; Gosz J.; Grimes J.; Gross K.; Holben W.; Holt J.; Huang L. H.; Hunter-Cevera J. C.; Kirsop B.; Klopateck C.; Klug M.; Krichevsky M.; Lenski R.; Moore L.; Murray R.G.E.; Pace N.; Poindexter J.; Prieur D.; Riley M.; Roskoski J.; Schmidt T.; Sneath P.; Staley J.; Sugawara H.; Swings J.; Tiedje J.; Torsvik V.; Truper H.; Vidaver A.; White D. & Zavarzin G. (1995) Priorities for microbial biodiversity research. Summary and recommendations. Report of a workshop organized by the Center of Microbial Ecology at Michigan State University. pp. 1-8.

- Bormeman J., Skroch P. W., O'Sullivan K. M., Palus J. A., Rumjanek N. G., Jansen J. L., Nienhuis J. & Triplett E. W. (1996) Molecular microbial diversity of an agricultural soil in Wisconsin. *Applied and Environmental Microbiology*. **62**: 1935-1943.
- Böttger E. C. (1989) Rapid determination of bacterial ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA. *FEMS Microbiology Letters*. **65**: 171-176.
- Bramwell P. A., Barallon R. V., Rogers H. J. & Bailey M. J. (1996a) Extraction and PCR amplification of DNA from the rhizoplane. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.4.3**: 1-20.
- Bramwell P. A., Barallon R. V., Rogers H. J. & Bailey M. J. (1996b) Extraction of microbial DNA from phylloplane. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed.). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.4.3**:1-21.
- Breen A., Rope A. F., Taylor D., Loper J. C. & Sferra P. R. (1995) Application of DNA amplification fingerprinting (DAF) to mixed cultured bioreactors. *Journal of Industrial Microbiology*. **14**:10-16.
- Bull A. T., Goodfellow M & Slater J. H. (1992) Biodiversity as a source of innovation in Biotechnology. *Annual Review in Microbiology*. **46**: 219-252.
- Burrell, P. C., Keller J. & Blackall L. (1998) Microbiology of a nitrite-oxidizing bioreactor. *Applied and Environmental Microbiology*. **64**: 1878-1883.
- Busse H-J., Denner E. B. M. & Lubitz W. (1996) Classification and identification of bacteria: current approaches to an old problem. Overview of methods used in bacterial systematics. *Journal of Biotechnology*. **47**:3-38.
- Cano R. J., Rasmussen S. R., Sánchez-Fraga G. & Palomares J. C. (1993) Fluorescent detection-polymerase chain reaction (FD-PCR) assay on microwell plates as a screening test for salmonellas in foods. *Journal of Applied Bacteriology*. **75**: 247-253.
- Collins M. D., Samelis J., Metaxopoulos J. & Wallbanks S. (1993) Taxonomic studies on some leuconostoc-like organisms from fermented sausages: description of a new genus *Weisella* for the *Leuconostoc paramesenterioides* group of species. *Journal of Applied Bacteriology*. **75**: 595-603.

- Cuoto J. A. & Hogg T. A. (1994) Diversity of ethanol-tolerant lactobacilli isolated from Duoro fortified wine: clustering and identification by numerical analysis of electrophoretic protein profiles. *Journal of Applied Bacteriology*. **76**: 487-491.
- Czajka J., Bsat N., Piani M., Russ W., Sultana K., Wiedmann M., Whitaker R. & Batt C. A. (1993) Differentiation of *Listeria monocytogenes* and *Listeria innocua* by 16S rRNA genes and intraspecies discrimination of *Listeria monocytogenes* strains by random amplified polymorphic DNA polymorphisms. *Applied and Environmental Microbiology*. **59**: 304-308.
- Datta A. R., Moore M. A., Wentz B. A. & Lane J. (1993) Identification and enumeration of *Listeria monocytogenes* by non-radioactive DNA probe colony hybridization. *Applied and Environmental Microbiology*. **59**: 144-149.
- Devereux R. & Willis S. G. (1995) Amplification of ribosomal RNA sequences. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed.). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **3.3.1**:1-11.
- Dickinson J. H., Kroll R. G. & Grant K. A. (1995) The direct application of the polymerase chain reaction to DNA extracted from foods. *Letters in Applied Microbiology*. **4**:212-216.
- Dykhizen, D. E. (1998) Santa Rosalia revisited: why are so many species bacterias of bacteria?. *Antonie van Leeuwenhoek*. **73**:25-33.
- Ehrmann M., Ludwig W. & Schleifer K. H. (1994) Reverse dot blot hybridization: a useful method for the direct identification of lactic acid bacteria in fermented food. *FEMS Microbiology Letters*. **117**:143-150.
- Felsenstein J. (1989) PHYLIP—Phylogeny Inference Package (Version 3.2). *Cladistics*. **5**:164-166.
- Felske A., Rheims H., Wolterink A., Stackenbrandt E. & Akkermans A. D. L. (1997) Ribosome analysis reveals prominent activity of an uncultured member of the class Actinobacteria in grassland soils. *Microbiology*. **143**: 2983 - 2989.
- Felske A., Wolterink A., Van Lis R. & Akkermans A. D. L. (1998) Phylogeny of the main bacterial 16S rRNA sequences in Drentse a grassland soils (The Netherlands). *Applied and Environmental Microbiology*. **64**: 871-879.

- Feng, P. (1997) Impact of molecular biology on the detection of food borne pathogens. *Molecular Biotechnology*. **7**:267-278.
- Fleet G. H. (1999) Microorganisms in food ecosystems. *International Journal of Food Microbiology*. **50**:101-117.
- Flores T. (1996) Caracterización fisiológica de las bacterias lácticas aisladas del pozol. Tesis de Licenciatura. Facultad de Ciencias, UNAM. México. 71pp.
- Gaden E. L., Bokanga M., Harlander S., Hesseltine C. W. & Steinkraus K. H. (1992) Research priorities in traditional fermented foods. In *Applications of Biotechnology to traditional Fermented Foods*. Report of an had hoc Panel of the Board on Science and Technology for International Development. Office of International Affairs National Research Council. National Academy Press, Washington, D.C. pp. 3-7.
- Gürtler H. & Anker L. (1994) Industries requirements with regard to identification of bacteria. In: Priest F. G., Ramos-Comenzala A. & Tindall B. J. (ed). *Bacterial diversity and systematics*. FEMS Symposium No. 75. Plenum Press, New York. pp. 297-307.
- Hamad S. H., Dieng M. C., Ehrmann M. A. & Vogel R. F. (1997) Characterization of the bacterial flora of Sudanese sorghum flour and sorghum sourdough. *Journal of Applied Microbiology*. **83**:764-770.
- Hattori T., Mitsui H., Haga H., Wakao N., Shikano S., Grolach K., Kasahara Y., El-Beltagy A. & Hattori R. (1997) Advances in soil microbial ecology and the biodiversity. *Antonie van Leeuwenhoek*. **72**: 21-28.
- Hilali F., Saulnier P., Chachaty E. & Andremont A. (1997) Decontamination of polymerase chain reaction reagents for detection of low concentrations of 16S rRNA genes. *Molecular Biotechnology*. **7**: 207-216.
- Hirsch C. F. & Sigmund J. M. (1995) Use of polymerase chain reaction (PCR) fingerprinting to differentiate bacteria for microbial products screening. *Journal of Industrial Microbiology*. **15**:85-93.
- Holben W., Jansson J., Chelm B. & Tiedje J. (1988) DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Applied and Environmental Microbiology*. **54**:703-711.

- Howard P. J., Horsonp K. D. & Luchansky J. B. (1992) Differentiation of *Listeria innocua*, *Listeria ivanovii*, and *Listeria seeligeri* by pulse-field gel electrophoresis. *Applied and Environmental Microbiology*. **58**: 709-712.
- Jansoon J. K. & Prosser J. Y. (1997) Quantification of the presence and activity of specific microorganisms in nature. *Molecular Biotechnology*. **7**:103-120.
- Josephson K. L., Gerba C. P. & Pepper I. L. (1993) Polymerase chain reaction detection of nonviable bacterial pathogens. *Applied and Environmental Microbiology*. **59**: 3513-3515.
- Jukes T. H. & Cantor C. R. (1969) Evolution of protein molecules. In Munro H. N. (ed) Mammalian protein metabolism, vol. 3 Academic Press, Inc. New York, pp 21-132.
- Kim C., Swaminathan B., Cassaday P. K., Mayer L. W. & Holloway B. P. (1991) Rapid confirmation of *Listeria monocytogenes* isolated from foods by a colony blot assay using a digoxigenin-labeled synthetic oligonucleotide probe. *Applied and Environmental Microbiology*. **57**: 1609-1614.
- Klijn N., Weerkamp A. & de Vos W. (1991a) Detection and characterization of lactose-utilizing *Lactococcus* spp. in natural ecosystems. *Applied and Environmental Microbiology*. **61**:788-792.
- Klijn N., Weerkamp A. & de Vos W. (1991b) Identification of mesophilic lactic acid bacteria by using polymerase chain reaction - amplified variable regions of 16S rRNA and specific DNA probes. *Applied and Environmental Microbiology*. **57**:3390-3393.
- Kovalenko S. A., Tanaka M. & Ozawa T. (1994) Simple method for preparation of plasmid DNA yielding long and accurate sequence data. *Nucleic Acids Research*. **22**:5771-5772
- Lane D. J., Pace B., Olsen G. J., Stahl D. A., Sogin M. L. & Pace N. R. (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. *Proceedings of the National Academy of Sciences USA*. **82**:6955-6959.
- Le Bourgeois P., Mata M. & Ritzenthaler P. (1989) Genome comparison of *Lactococcus* strains by pulsed-field gel electrophoresis. *FEMS Microbiology Letters*. **59**:65-70.

- Le Roy, D., Bouchet, A., Saulnier, P., Pecquet, S. & Andremont A. (1993) Comparison of chemical assay, and dot blot hybridization for detection of aerobactin in members of the family *Enterobacteriaceae*. *Applied and Environmental Microbiology*. **59**: 942-944.
- Lee S., Bollinger J., Bezdicek D. & Ogram A. (1996) Estimation of the abundance of an uncultured soil bacteria by a competitive quantitative PCR method. *Applied and Environmental Microbiology*. **62**:3787-3793.
- Lewin B. (1997) Genes VI. Oxford University Press. Oxford. pp. 658.
- Ludwig W. & Schleifer K. H. (1994) Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiological Reviews*. **15**:155-173.
- Maidak B. L., Cole J. R., Lilburn T. G., Parker Jr C. T., Saxman P. R., Stredwick J. M., Garrity G. M., Li B., Olsen G. J., Pramanik S., Schmidt T. M. & Tiedje J. M. (2000) The RDP (Ribosomal Database Project) continues. *Nucleic Acids Research*. **28**:173-174.
- Marshall K. C. (1993) Microbial ecology: Whither goes thought?. In: Guerrero R. and Pedrós-Alió C. (ed.). *Trends in Microbial Ecology*. Proceedings of the Sixth International Symposium on Microbial Ecology. Spanish Society of Microbiology. pp. 5-8
- Nazar R. N., Robb E. J. & Volossioouk T. (1996) Direct extraction of fungal DNA from soil. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.3.6**:1-8.
- Notermans S., Chakraborty T., Leimmeister-Wächter M., Dufrenne J., Heuvelman K. J., Jansen W, Wernars K & Guinee P. (1989) Specific gene probe for detection of biotyped and serotyped *Listeria* strains. *Applied and Environmental Microbiology*. **55**: 902-906.
- Nuraida L., Wachter C. & Owens J. D. (1995) Microbiology of pozol, a Mexican fermented maize dough. *World Journal of Microbiology and Biotechnology* **11**:567-571.
- Nüsslein K. & Tiedje J. (1998) Characterization of the dominant and rare members of a young Hawaiian soil bacterial community with small-subunit ribosomal DNA amplified from DNA fractionated on the basis of its guanine

- and cytosine composition. *Applied and Environmental Microbiology*. **64**: 1283-1289.
- Oison B. H. & Tsai Y. (1992) Molecular approaches to environmental management. *Environmental Microbiology*. pp. 239-263.
- O'Sullivan D. J. (1999) Methods for analysis of the intestinal microflora. In: Tannock G. (Ed.). *Probiotics. A critical Review*. Horizon Scientific Press. Norfolk. pp. 23-44.
- Pace N. (1996) New perspectives on the natural microbial world: Molecular Microbial Ecology. *American Society of Microbiology News*. **66**:463-470.
- Pace N. (1997) A molecular view of microbial diversity and the biosphere. *Science*. **276**:734-740.
- Page R. D. M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applied Bioscience*. **12**:357-358
- Page R. D. M. & Holmes E. C. (1999) Molecular Evolution. A phylogenetic approach. Blackwell Science Ltd. Oxford: 172-227.
- Paul J. H. (1996) Extraction of microbial DNA from aquatic sources: marine environments. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.1.1**:1-13
- Pichard S. L. & Paul J. H. (1996) Extraction of microbial RNA from aquatic sources: marine environments. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.2.1**:1-13.
- Pickup R. W., Rhodes G. & Saunders J. R. (1995) Extraction of microbial DNA from aquatic sources: freshwater. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.1.2**:1-11.
- Pillai S., Josephson K., Bailey R., Gerba C. & Pepper I. (1991) Rapid method for processing soil samples for polymerase chain reaction amplification of specific gene sequences. *Applied and Environmental Microbiology*. **57**:2283-2286.
- Porteus L. A., Armstrong J. L., Seidler R. J. & Watrud L. S. (1994) An effective method to extract DNA from environmental samples for polymerase chain

- reaction amplification and DNA fingerprint analysis. *Current Microbiology*. **5**:301-307.
- Pot B. (1998) Future trends in identification of microorganisms – molecular techniques, protein profiling, immunological methods, etc. In: Dairy Science and Technology. Proceedings of the 25<sup>th</sup> International Dairy Congress: 293-307.
- Ramirez-Saad H., Akkermans W. M. & Akkermans D. L. (1996) DNA extraction from actinorhizal nodules. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.4.4**:1-11.
- Rheims H., Spröer C., Rainey F. A. & Stakebrandt E. (1996) Molecular biological evidence for the occurrence of uncultured members of the actinomycete line of descent in different environments and geographical locations. *Microbiology*. **142**: 2863-2870.
- Russek-Cohen E. & Colwell R R. (1996) Numerical classification of microorganisms. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **3.2.1.1**:1-21.
- Saitou N. & Nei M. (1987) The neighbor-joining method: a new method for reconstruction of phylogenetic trees. *Molecular Biological Evolution*. **4**:406-425
- Sambrook J., Fritsch E.F. & Maniatis T. (1989) Molecular cloning: a laboratory manual. 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y.
- Schleifer K.-H. & Ludwig W. (1994) Molecular taxonomy: classification and identification. In: Priest F. G., Ramos-Cormenzana A. & Tindall B. J. (ed.). *Bacterial diversity and systematics*. FEMS Symposium No. 75. New York: Plenum Press. pp. 1-15.
- Smalla K. (1996) Extraction of microbial DNA from sewage and manure slurries. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **1.1.3.1**:1-10.

- Stackebrandt E. & Goebel B. M. (1994) Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic Bacteriology*. **44**:864-849.
- Stackebrandt E. & Rainey F. A. (1995) Partial and complete 16S rDNA sequences, their use in generation of 16S rDNA phylogenetic trees and their implications in molecular ecological studies. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht. **3.1.1**:1-17.
- Steffan R. J. & Atlas R. M. (1991) Polymerase chain reaction: applications in environmental microbiology. *Annual Review in Microbiology*. **45**: 137-161.
- Steffan R., Goksoyr J., Bej A. & Atlas R. (1988) Recovery of DNA from soils and sediments. *Applied and Environmental Microbiology*. **54**:2908-2915.
- Thompson J. D., Higgins D. G., Gibson T. J. (1994) CLUSTAL W: improving the sensibility of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. **22**:4673-4680
- Torsvik V., Goksoyr J. & Daae F. L. (1990) High diversity in DNA of soil bacteria. *Applied and Environmental Microbiology*. **56**:782-787.
- Torsvik V., Goksoyr J., Daae F. L., Sorheim R., Michalsen J. & Salte K. (1993) Diversity of microbial communities determined by DNA reassociation technique In: Guerrero R. & Pedrós-Alió C. (ed.). *Trends in Microbial Ecology*. Proceedings of the Sixth International Symposium on Microbial Ecology Spanish Society of Microbiology. pp. 375-378.
- Trevors J. T. (1992) DNA extraction from soil. *Microbial Releases*. **1**:3-9.
- Tsai, Y. & Olson B. (1992) Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. *Applied and Environmental Microbiology*. **58**:2292-2295.
- Vandamme P., Pot B., Gillis M., de Vos P., Kersters K. & Swings J. (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological Reviews*. **60**: 407-438.

- Wacher C. (1993) Alimentos y bebidas fermentadas tradicionales. En: García-garibay M., Quintero-Ramírez R. y López Munguía A. (eds). *Biología alimentaria*. LIMUSA. México D.F. pp. 313-315.
- Wacher C., Cañas A., Cook P.E., Barzana E. & Owens J.D. (1993) Sources of microorganisms in pozol; a Mexican traditional fermented maize dough. *World Journal of Microbiology and Biotechnology*. **9**: 269-274.
- Wacher C., Cañas A., Barzana E., Lappe P., Ulloa M. & Owens J.D. (2000) Microbiology of Indian and Mestizo pozol fermentations. *Food Microbiology*. **17**: 251-256.
- Ward L. J. H. & Timmins M. J. (1999) Differentiation of *Lactobacillus casei*, *Lactobacillus paracasei* and *Lactobacillus rhamnosus* by polymerase chain reaction. *Letters in Applied Microbiology*. **29**:90-92.
- Weisburg W. G., Oyaizu Y., Oyaizu Y. & Woese R. (1985) Natural relationships between *Bacteroides* and *Flavobacteria*. *Journal of Bacteriology*. **164**:230-236.
- Weisburg W. G., Barns S. M., Pelletier D. A. & Lane D. J. (1991) 16S Ribosomal amplification for phylogenetic study. *Journal of Bacteriology*. **173**:697-703.
- Wellington E. M. H., Huddleston A. S. & Marsh P. (1994) The use of molecular markers for the detection and typing of bacteria in soil. In: Priest F. G., Ramos-Cormenzana A. & Tindall B. J. (ed.). *Bacterial diversity and systematics*. FEMS Symposium No. 75. New York: Plenum Press. pp. 137-152.
- Wesley I. V & Ashton F. (1991) Restriction enzyme analysis of *Listeria monocytogenes* strains associated with food-borne epidemics. *Applied and Environmental Microbiology*. **57**:969-975.
- Wheeler E. & Stahl D. A. (1996) Extraction of microbial DNA from aquatic sediments. In: Akkermans D. L., van Elsas J. D. & de Bruijn F. J. (ed). *Molecular microbial ecology manual*. Kluwer Academic Publishers. Dordrecht:**1.1.5**:1-29.
- Wiedmann M., Czajka J., Barany F. & Batt C. A. (1992) Discrimination of *Listeria monocytogenes* from other *Listeria* species by ligase chain reaction. *Applied and Environmental Microbiology*. **58**:3443-3447.

- Zhong W., Millsap K., Briakowska-Hobrzanska H. & Reid G. (1998) Differentiation of *Lactobacillus* species by molecular typing. *Applied and Environmental Microbiology*. **64**:2418-2423.
- Zhou J., Bruns M. A. & Tiedje J. M. (1996). DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology*. **2**:316-322.
- Zhou J., Davey M. E., Figueras J. B., Rivkina E., Gilichinsky D. & Tiedje J. (1997) Phylogenetic diversity of a bacterial community determined from Siberian tundra soil DNA. *Microbiology*. **143**: 3913-3919.

# 9

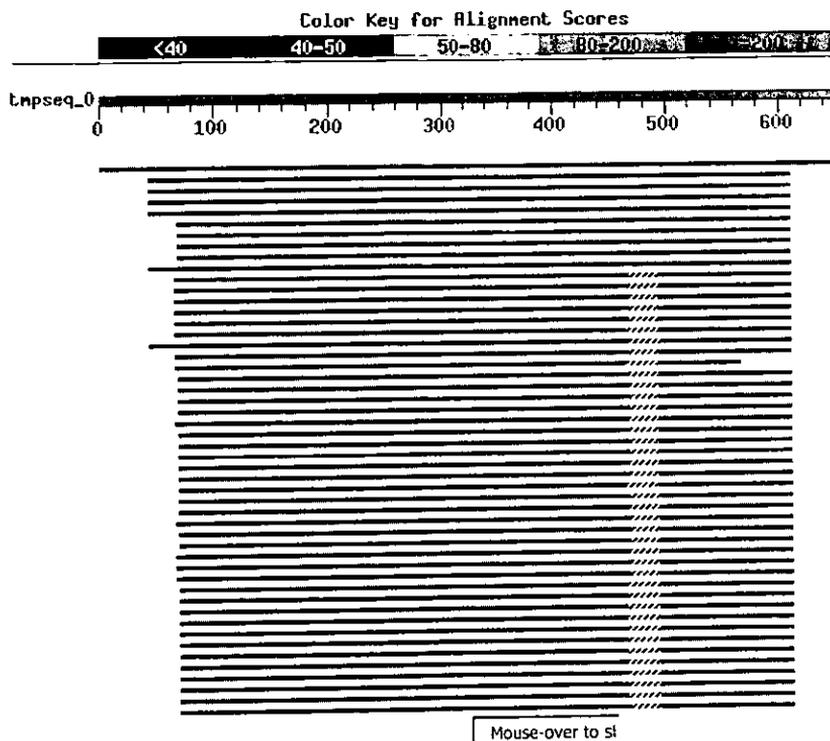
## APÉNDICE

---

**RESULTADOS DEL ANÁLISIS DE LAS SECUENCIAS DE LAS CLONAS PZ EN EL PROGRAMA BLAST.**

Resultados del análisis de la secuencia de la clona Pz608 en el programa BLAST.

### Distribution of 218 Blast Hits on the Query Sequence

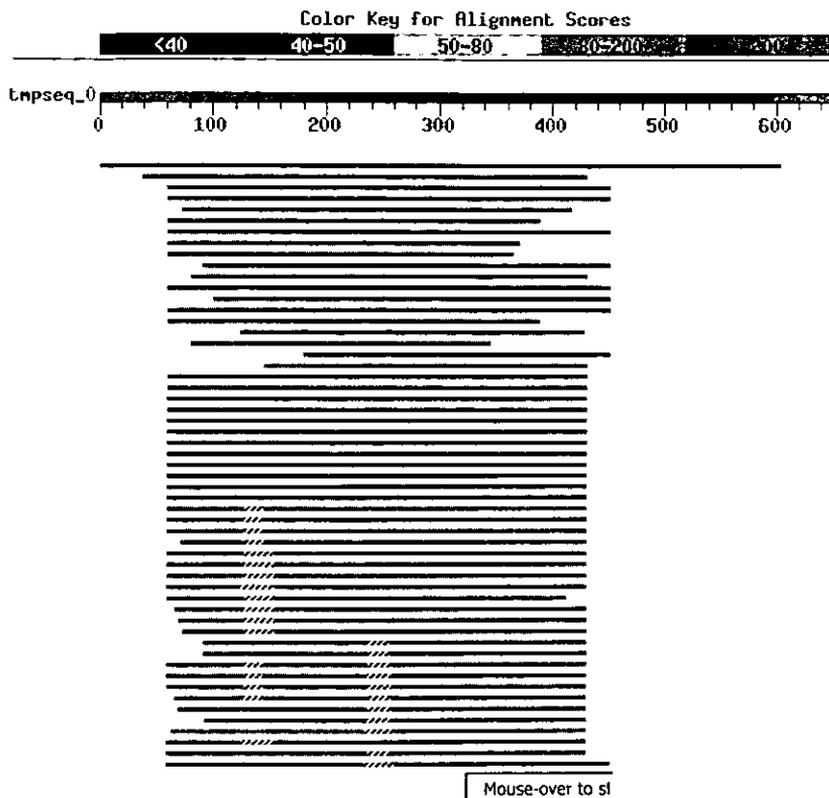


Sequences producing significant alignments:	Score (bits)	E Value	Identities* (%)
<u>gb AF095544.1 AF095544</u> Uncultured bacterium clone Pz608 16S...	1199	0.0	100
<u>emb X64887.1 LLRRNAOP</u> L.lactis ribosomal RNA operon encodin...	961	0.0	96
<u>emb AJ271851.1 LLA271851</u> Lactococcus lactis partial 16S rRN...	950	0.0	95
<u>gb N58837.1 STRRR16SD</u> Streptococcus lactis 16S ribosomal RNA	944	0.0	95
<u>gb N58836.1 STRRR16SC</u> Streptococcus cremoris 16S ribosomal RNA	938	0.0	95
<u>gb AF061005.1 AF061005</u> Enterococcus seriolicida 16S ribosom...	825	0.0	94
<u>dbj AB026843.1 AB026843</u> Enterococcus seriolicida gene for 1...	825	0.0	94
<u>gb L32813.1 ENERRD</u> Enterococcus seriolicida 16S ribosomal R...	801	0.0	93
<u>dbj AB012306.1 AB012306</u> Lactococcus garvieae gene for 16S r...	785	0.0	93
<u>emb X54260.1 LLA16S</u> Lactococcus lactis 16S ribosomal RNA	767	0.0	91

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz547 en el programa BLAST.

### Distribution of 201 Blast Hits on the Query Sequence

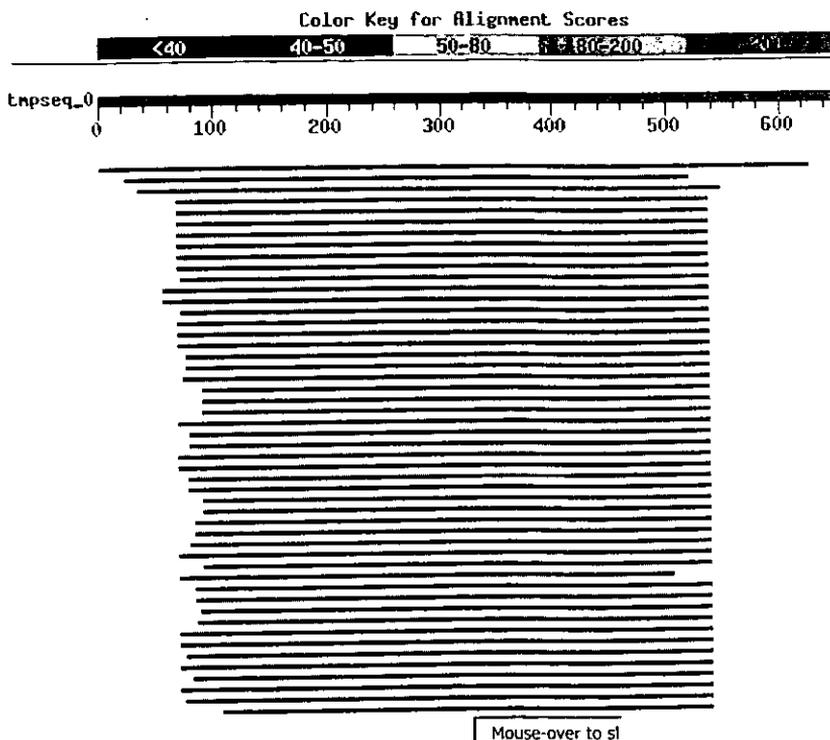


Sequences producing significant alignments:	Score (bits)	E Value	Identities (%)
<u>gb AF095542.1 AF095542</u> Lactococcus lactis clone Pz547 16S r...	882	0.0	98
<u>gb AF095538.1 AF095538</u> Lactococcus lactis clone Pz513 16S r...	632	e-179	94
<u>emb X64887.1 LLRRNAOP</u> L.lactis ribosomal RNA operon encodin...	620	e-175	93
<u>gb M58837.1 STRRR16SD</u> Streptococcus lactis 16S ribosomal RNA	620	e-175	93
<u>gb AY006859.1 </u> Uncultured rumen bacterium 16S ribosomal RNA...	571	e-160	95
<u>gb MS4920.1 LACRR16S1</u> Lactococcus lactis 16S ribosomal RNA,...	569	e-160	96
<u>gb AF107514.1 </u> Uncultured bacterium SY2-77 16S ribosomal RN...	565	e-159	92
<u>gb M55156.1 LACRR16S4</u> Lactococcus lactis 16S ribosomal RNA,...	561	e-157	97
<u>gb M54921.1 LACRR16S2</u> Lactococcus lactis 16S ribosomal RNA,...	555	e-156	97
<u>gb AF270792.1 AF270792</u> Lactococcus lactis N4 16S ribosomal ...	551	e-154	93

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz504 en el programa BLAST.

### Distribution of 101 Blast Hits on the Query Sequence

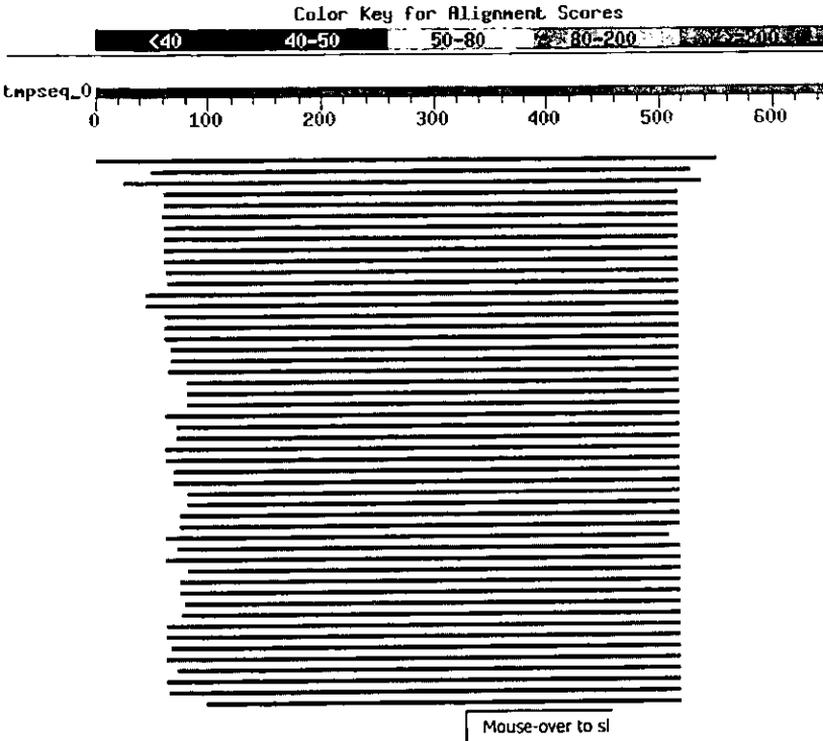


Sequences producing significant alignments:			Score	E	Identities*
			(bits)	Value	(%)
<u>gb AF095550.1 AF095550</u>	Uncultured bacterium clone Pz504 16S...	<u>1148</u>	0.0	100	
<u>gb AF095557.1 AF095557</u>	Uncultured bacterium clone Pz525 16S...	<u>858</u>	0.0	97	
<u>gb AF095537.1 AF095537</u>	Uncultured bacterium clone Pz406 16S...	<u>819</u>	0.0	95	
<u>gb AF104114.1 AF104114</u>	Streptococcus bovis strain ATCC43143...	<u>696</u>	0.0	94	
<u>gb AF104109.1 AF104109</u>	Streptococcus bovis strain JBI 16S r...	<u>696</u>	0.0	94	
<u>gb AF009482.1 AF009482</u>	Streptococcus suis 16S ribosomal RNA...	<u>690</u>	0.0	94	
<u>gb AY005044.1 </u>	Streptococcus sp. oral clone CH016 16S ribos...	<u>688</u>	0.0	94	
<u>gb AF009488.1 AF009488</u>	Streptococcus suis 16S ribosomal RNA...	<u>688</u>	0.0	94	
<u>gb AF009476.1 AF009476</u>	Streptococcus suis 16S ribosomal RNA...	<u>688</u>	0.0	94	
<u>gb AF082731.1 SB16SRRNA2</u>	Streptococcus bovis strain FM 16S ...	<u>684</u>	0.0	94	

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz406 en el programa BLAST.

**Distribution of 101 Blast Hits on the Query Sequence**

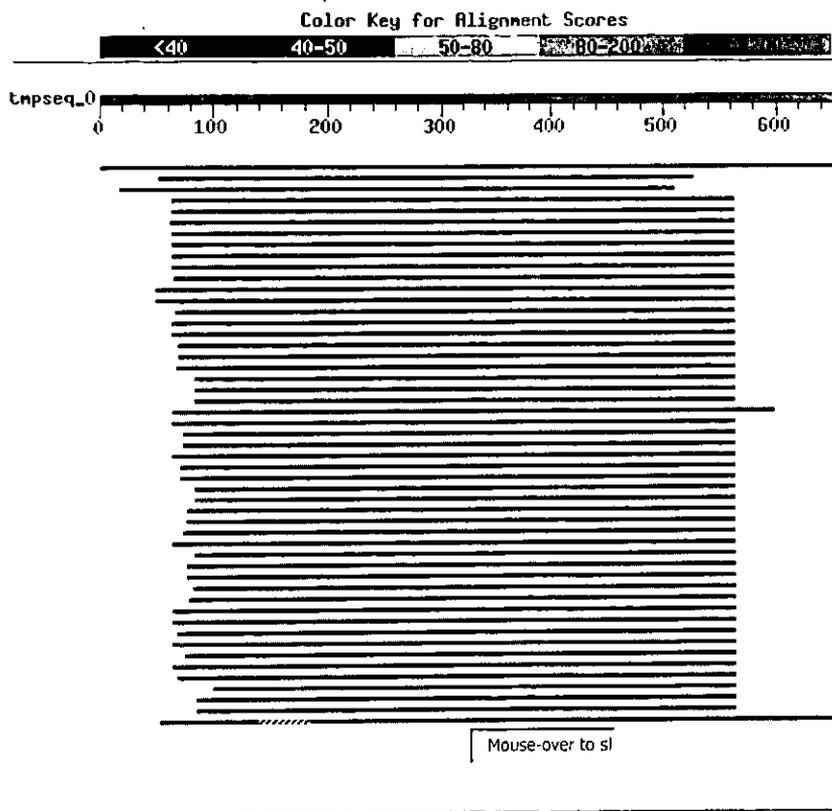


Sequences producing significant alignments:	Score (bits)	E Value	Identities* (%)
gb AF095537.1 AF095537	1045	0.0	100
gb AF095557.1 AF095557	870	0.0	98
gb AF095550.1 AF095550	819	0.0	95
gb AF104114.1 AF104114	728	0.0	95
gb AF104109.1 AF104109	728	0.0	94
gb AF009482.1 AF009482	722	0.0	94
gb AY005044.1	720	0.0	94
gb AF009488.1 AF009488	720	0.0	94
gb AF009476.1 AF009476	720	0.0	94
gb AF082731.1 SB16SRRNA2	716	0.0	94

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz525 en el programa BLAST.

### Distribution of 102 Blast Hits on the Query Sequence

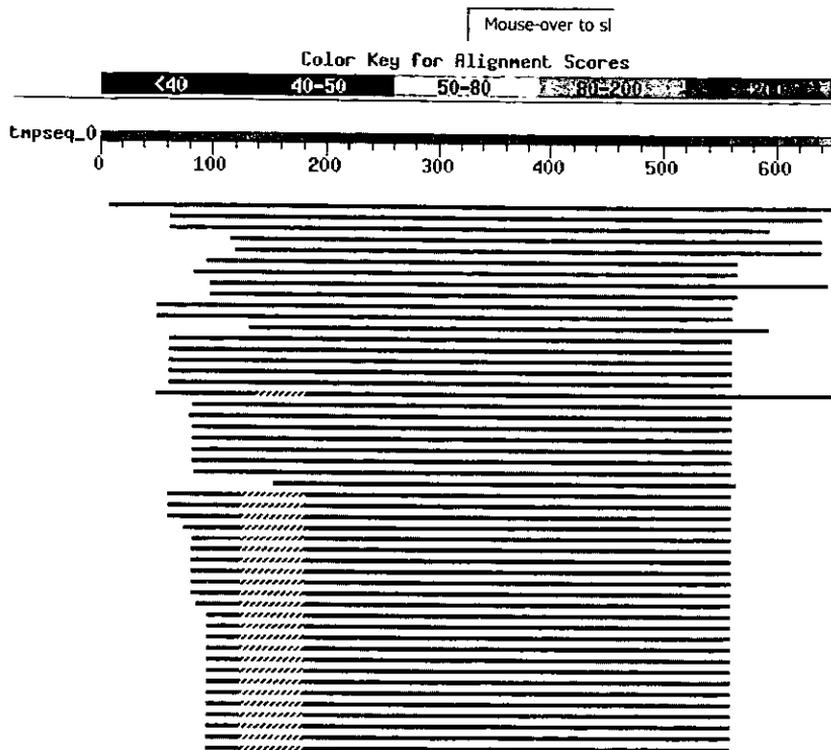


Sequences producing significant alignments:	Score (bits)	E Value	Identities* (%)
<u>gb AF095537.1 AF095537</u> Uncultured bacterium clone Pz525 16S...	1235	0.0	100
<u>gb AF095537.1 AF095537</u> Uncultured bacterium clone Pz406 16S...	870	0.0	98
<u>gb AF095550.1 AF095550</u> Uncultured bacterium clone Pz504 16S...	858	0.0	97
<u>gb AF104114.1 AF104114</u> Streptococcus bovis strain ATCC43143...	842	0.0	96
<u>gb AF104115.1 AF104109</u> Streptococcus bovis strain JBI 16S r...	842	0.0	96
<u>gb AF009482.1 AF009482</u> Streptococcus suis 16S ribosomal RNA...	837	0.0	96
<u>gb AY005344.1 </u> Streptococcus sp. oral clone CH016 16S ribos...	835	0.0	96
<u>gb AF009488.1 AF009488</u> Streptococcus suis 16S ribosomal RNA...	835	0.0	96
<u>gb AF009476.1 AF009476</u> Streptococcus suis 16S ribosomal RNA...	835	0.0	96
<u>gb AF082731.1 SB16SRRNA2</u> Streptococcus bovis strain FM 16S ...	831	0.0	96

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz638 en el programa BLAST.

### Distribution of 175 Blast Hits on the Query Sequence

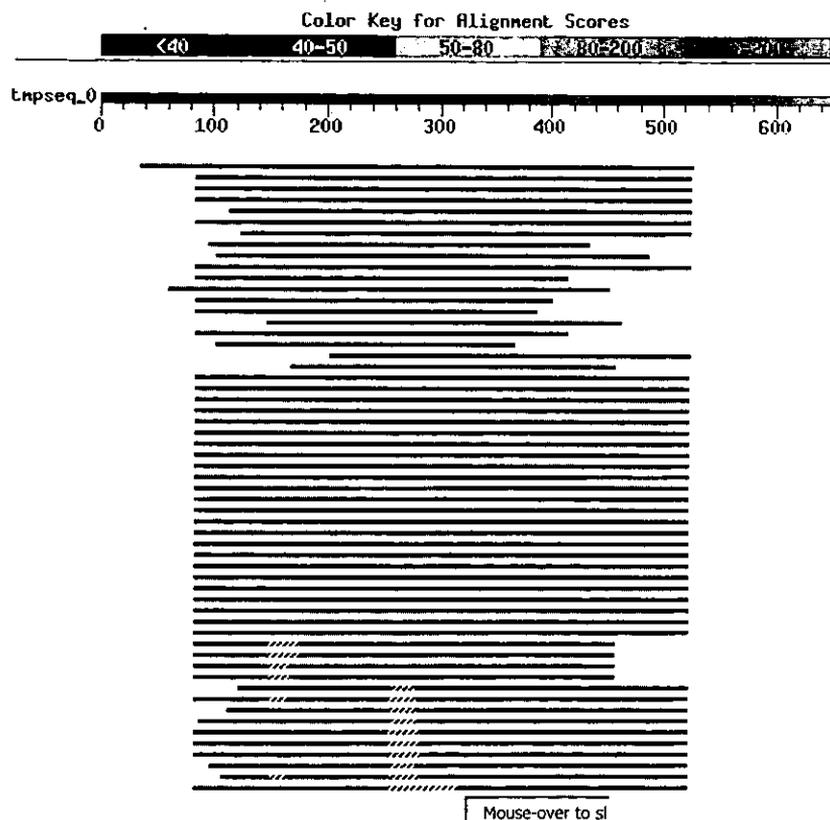


Sequences producing significant alignments:	Score (bits)	E Value	Identities* (%)
gb AF095560.1 AF095560 Uncultured bacterium clone Pz638 16S...	1223	0.0	100
emb X64887.1 LLRNAOP L.lactis ribosomal RNA operon encodin...	993	0.0	98
emb AJ271851.1 LLA271851 Lactococcus lactis partial 16S rRN...	989	0.0	99
gb M58836.1 STRRR16SC Streptococcus cremoris 16S ribosomal RNA	990	0.0	97
gb M58837.1 STRRR16SD Streptococcus lactis 16S ribosomal RNA	880	0.0	97
gb AF061005.1 AF061005 Enterococcus seriolicida 16S ribosom...	767	0.0	95
gb L32813.1 ENERRD Enterococcus seriolicida 16S ribosomal R...	765	0.0	95
dbj AB018211.1 AB018211 Lactococcus garvieae gene for 16S r...	720	0.0	94
dbj AB012306.1 AB012306 Lactococcus garvieae gene for 16S r...	720	0.0	93
emb AJ295853.1 SMI295853 Streptococcus mitis partial 16S rR...	716	0.0	92

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz513 en el programa BLAST.

### Distribution of 185 Blast Hits on the Query Sequence

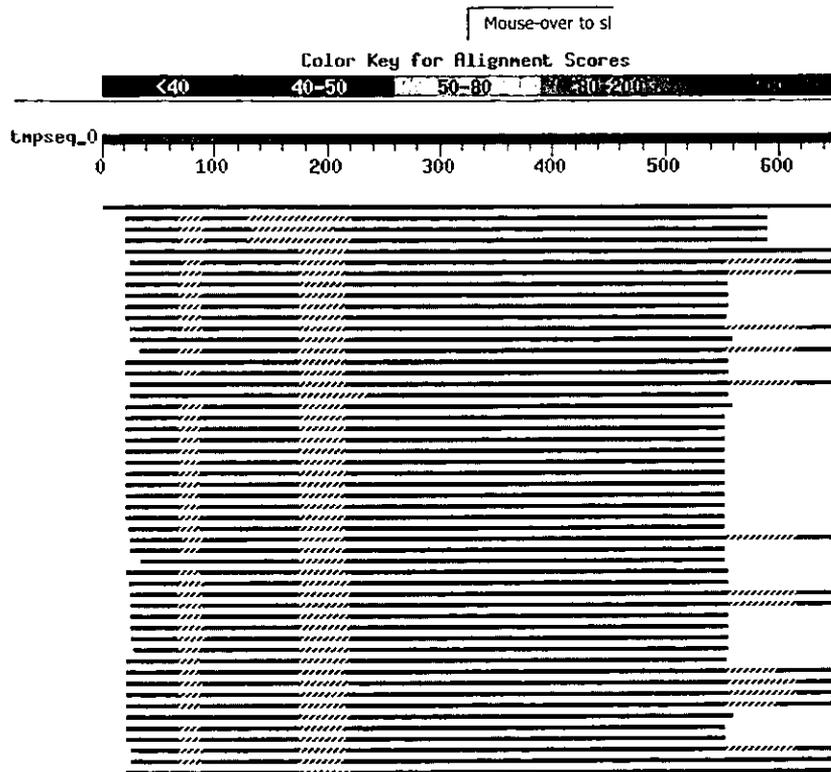


Sequences producing significant alignments:	Score (bits)	E Value	Identities (%)
<a href="#">gb AF095538.1 AF095538</a> Lactococcus lactis clone Pz513 16S r...	900	0.0	100
<a href="#">emb X64887.1 LLRRNAOP</a> L.lactis ribosomal RNA operon encodin...	753	0.0	96
<a href="#">gb M58837.1 STRRR16SD</a> Streptococcus lactis 16S ribosomal RNA...	753	0.0	96
<a href="#">gb AF107514.1 </a> Uncultured bacterium SY2-77 16S ribosomal RN...	698	0.0	94
<a href="#">gb AF270792.1 AF270792</a> Lactococcus lactis N4 16S ribosomal ...	692	0.0	95
<a href="#">gb M58836.1 STRRR16SC</a> Streptococcus cremoris 16S ribosomal RNA	670	0.0	93
<a href="#">gb AF270791.1 AF270791</a> Lactococcus lactis N1 16S ribosomal ...	664	0.0	95
<a href="#">gb AY006859.1 </a> Uncultured rumen bacterium 16S ribosomal RNA...	654	0.0	99
<a href="#">dbj AB030918.1 AB030918</a> Lactococcus lactis gene for 16S rRN...	654	0.0	96
<a href="#">emb X54260.1 LLA16S</a> Lactococcus lactis 16S ribosomal RNA	652	0.0	92

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz507 en el programa BLAST.

### Distribution of 314 Blast Hits on the Query Sequence

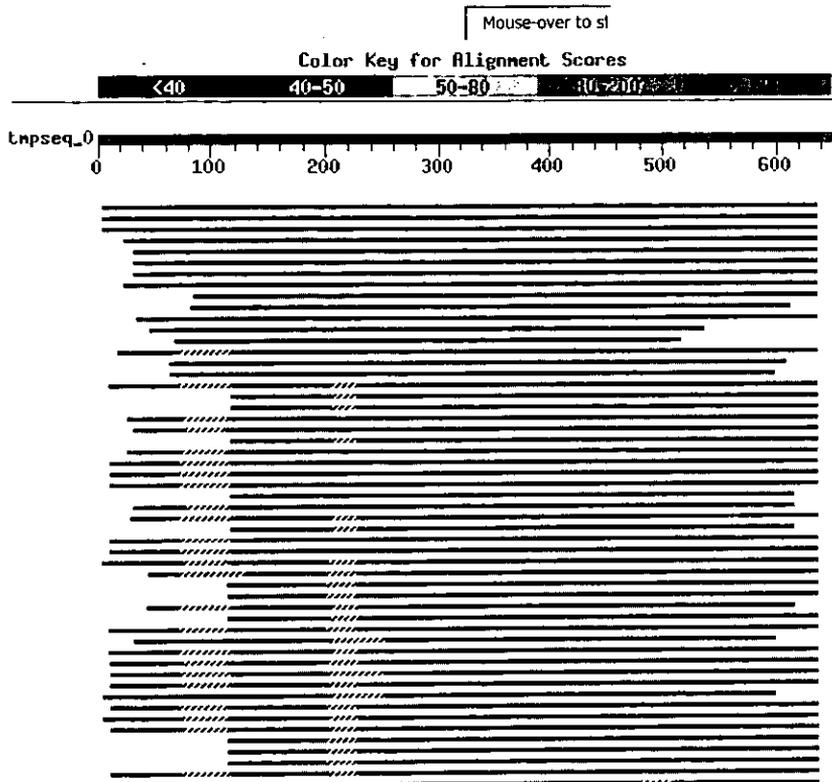


Sequences producing significant alignments:	Score (bits)	E Value	Identities* (%)
<u>gb AF095565.1 AF095565</u> Clostridium sp. clone Pz507 16S ribo...	<u>1164</u>	0.0	100
<u>gb AF018038.1 AF018038</u> Uncultured bacterium O8SE 16S riboso...	<u>541</u>	e-151	93
<u>emb X73447.1 CI16SRNA</u> C. irregularis 16S rRNA	<u>539</u>	e-151	92
<u>gb M59105.1 CLORR16SE</u> Clostridium sordellii 16S ribosomal RNA	<u>517</u>	e-144	93
<u>emb AJ009473.1 UEAJ9473</u> uncultured bacterium SJA-65 16S rRN...	<u>500</u>	e-139	89
<u>emb AJ229218.1 UEA229218</u> Unidentified eubacterium from anox...	<u>482</u>	e-133	92
<u>gb AF127023.1 AF127023</u> Clostridium algidicarnis 16S ribosom...	<u>480</u>	e-133	92
<u>gb AF127024.1 AF127024</u> Clostridium putrefaciens 16S ribosom...	<u>480</u>	e-133	92
<u>emb Y18177.1 CPU18177</u> Clostridium putrefaciens 16S rRNA gen...	<u>480</u>	e-133	92
<u>emb Y18189.1 APO18189</u> Anaerobacter polyendosporus 16S rRNA ...	<u>480</u>	e-133	92

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz619 en el programa BLAST.

### Distribution of 248 Blast Hits on the Query Sequence

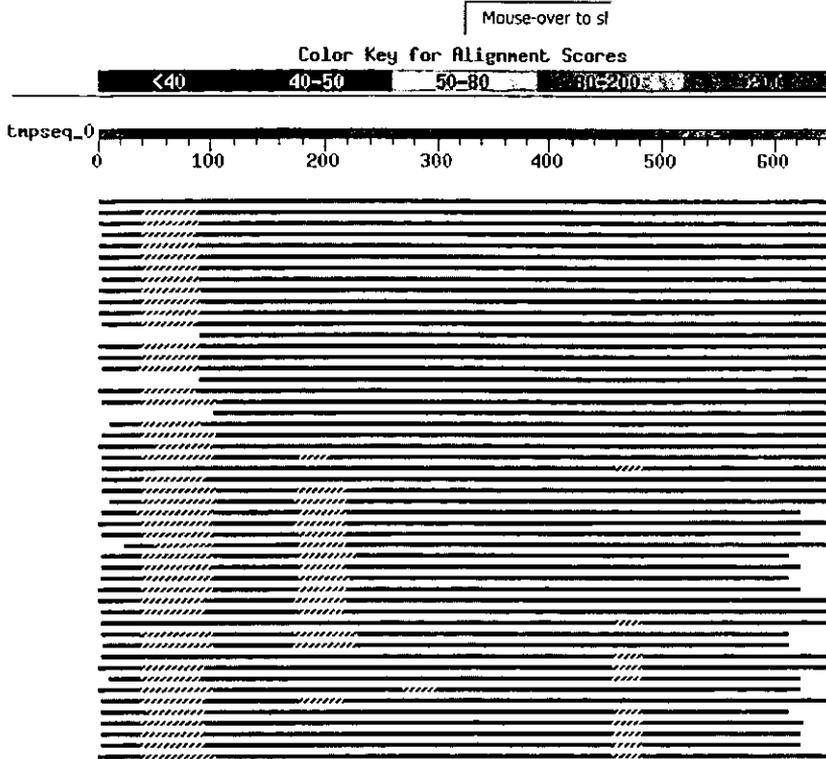


Sequences producing significant alignments:	Score (bits)	E Value	Identities* (%)
<u>gb</u>  AF095564.1 AF095564 Lactobacillus plantarum clone Pz619 ...	<u>1126</u>	0.0	100
<u>gb</u>  M58827.1 LBARR16SY Lactobacillus plantarum 16S ribosomal...	<u>1110</u>	0.0	96
<u>emb</u>  X52653.1 LP16SRRNA L.plantarum 16S rRNA	<u>1082</u>	0.0	97
<u>emb</u>  AJ272031.1 LPL272031 Lactobacillus plantarum 16S rRNA g...	<u>1068</u>	0.0	95
<u>dbj</u>  D79210.1 LBA16SRRN1 Lactobacillus plantarum gene for 16...	<u>1057</u>	0.0	96
<u>dbj</u>  D79211.1 LBA16SRRN2 Lactobacillus pentosus gene for 16S...	<u>1053</u>	0.0	96
<u>dbj</u>  D79213.1 LBA16SRRN3 Lactobacillus sp. gene for 16S rRNA...	<u>1029</u>	0.0	95
<u>emb</u>  AJ271852.1 LPL271852 Lactobacillus plantarum partial 16...	<u>965</u>	0.0	94
<u>gb</u>  AF296646.1 AF296646 Lactobacillus sp. 3F11 16S small sub...	<u>957</u>	0.0	96
<u>dbj</u>  AB025972.1 AB025972 Lactobacillus plantarum gene for 16...	<u>920</u>	0.0	96

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz322 en el programa BLAST.

### Distribution of 313 Blast Hits on the Query Sequence



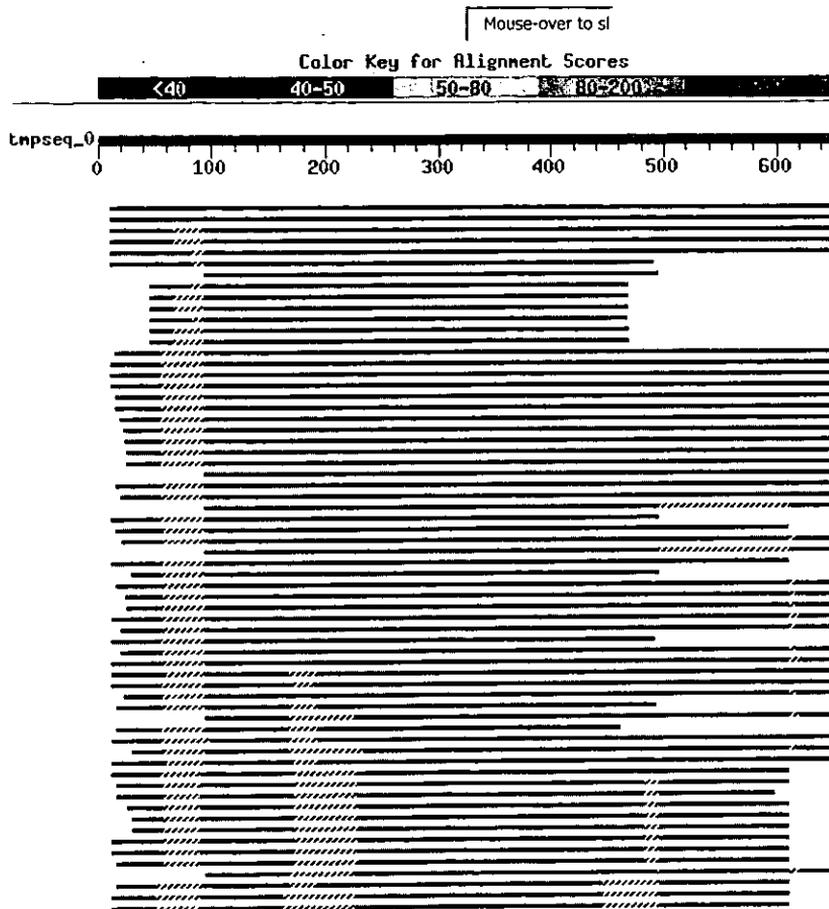
Sequences producing significant alignments:

	Score (bits)	E Value	Identities (%)
<u>gb AF095561.1 AF095561</u> Lactobacillus casei clone Pz322 16S ...	<u>1158</u>	0.0	100
<u>dbj D16552.1 D16552</u> Lactobacillus casei DNA for 16S rRNA	<u>932</u>	0.0	95
<u>dbj D16550.1 D16550</u> Lactobacillus casei DNA for 16S rRNA	<u>932</u>	0.0	95
<u>dbj D16548.1 D16548</u> Lactobacillus casei DNA for 16S rRNA	<u>932</u>	0.0	95
<u>dbj D16549.1 D16549</u> Lactobacillus casei DNA for 16S rRNA	<u>926</u>	0.0	95
<u>gb AF243147.1 AF243147</u> Lactobacillus paracasei subsp. parac...	<u>918</u>	0.0	95
<u>dbj D79212.1 LBA16SRRNA</u> Lactobacillus paracasei gene for 16...	<u>918</u>	0.0	95
<u>dbj D86517.1 D86517</u> Lactobacillus casei DNA for 16S ribosom...	<u>918</u>	0.0	95
<u>gb AF243146.1 AF243146</u> Lactobacillus rhamnosus strain Fl1 l...	<u>916</u>	0.0	95
<u>dbj D86518.1 D86518</u> Lactobacillus casei DNA for 16S ribosom...	<u>916</u>	0.0	95

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz314 en el programa BLAST.

### Distribution of 368 Blast Hits on the Query Sequence

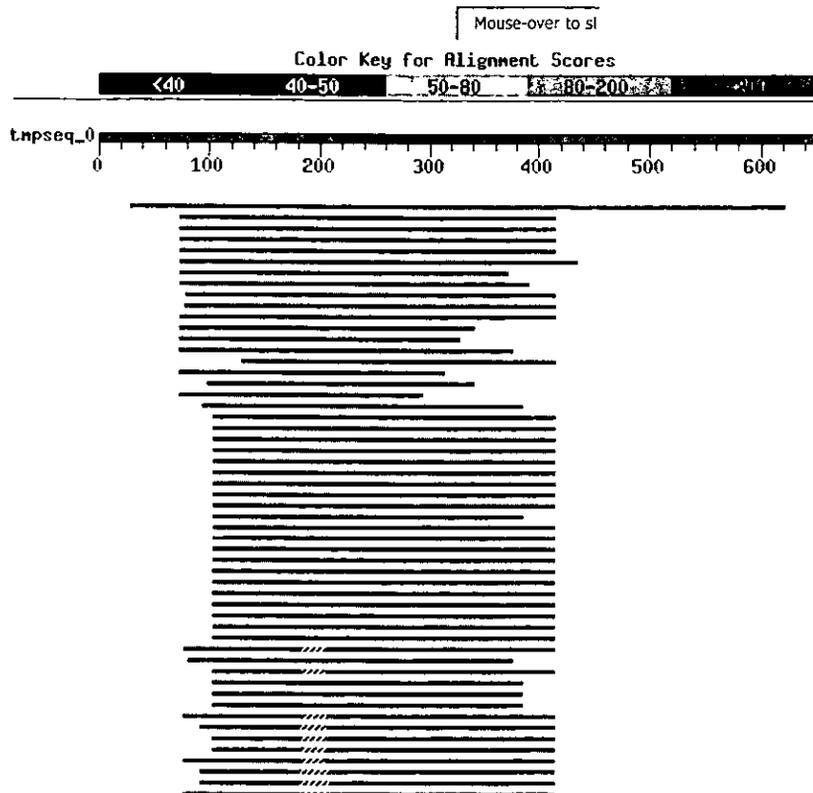


Sequences producing significant alignments:	Score (bits)	E Value	Identities (%)
<u>gb AF095563.1 AF095563</u> Lactobacillus delbrueckii clone Pz31...	1219	0.0	100
<u>emb X52654.1 LD16SRRNA</u> L.delbrueckii 16S rRNA	904	0.0	98
<u>gb M58823.1 LBARR16SU</u> Lactobacillus lactis 16S ribosomal RNA	739	0.0	98
<u>gb M58814.1 LBARR16SM</u> Lactobacillus delbrueckii 16S ribosom...	728	0.0	97
<u>dbj AB007908.1 AB007908</u> Lactobacillus delbrueckii gene for ...	728	0.0	97
<u>gb AF213703.1 AF213703</u> Lactobacillus delbrueckii subsp. lac...	720	0.0	97
<u>gb AF296649.1 AF296649</u> Lactobacillus delbrueckii 16S small ...	702	0.0	97
<u>emb AJ010834.1 LDE010834</u> Lactobacillus delbrueckii 16S rRNA...	692	0.0	98
<u>emb AJ010833.1 LDE010833</u> Lactobacillus delbrueckii 16S rRNA...	692	0.0	98
<u>emb AJ010832.1 LDE010832</u> Lactobacillus delbrueckii ssp. lac...	692	0.0	98

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz505 en el programa BLAST.

### Distribution of 156 Blast Hits on the Query Sequence

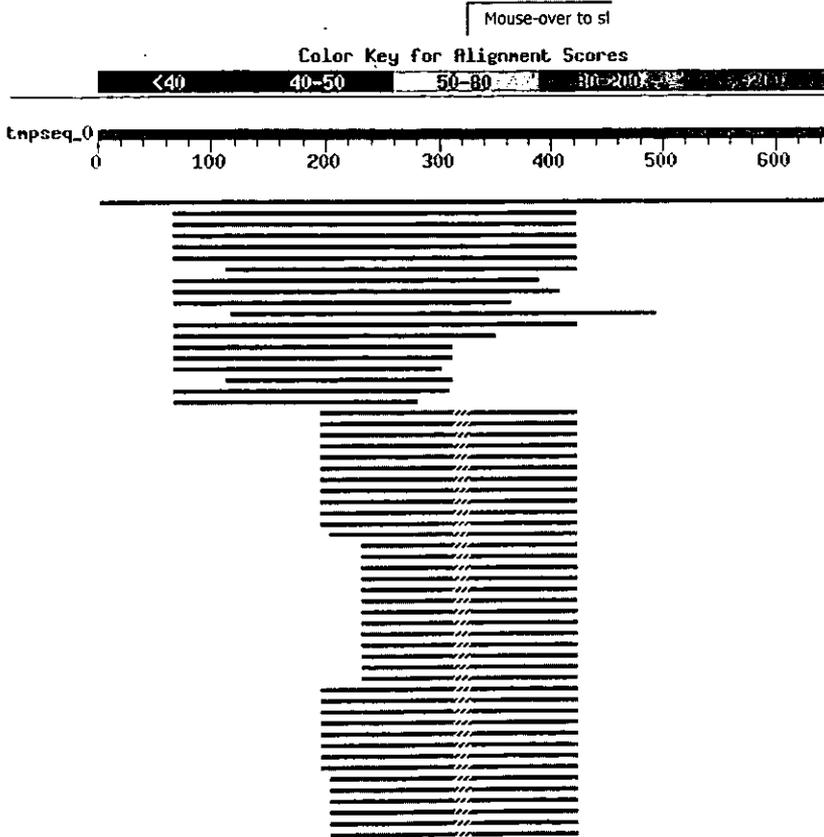


Sequences producing significant alignments:	Score (bits)	E Value	Identities (%)
<a href="#">gb AF095551.1 AF095551</a> Lactococcus lactis clone Pz505 16S r...	<u>1003</u>	0.0	97
<a href="#">gb AF270792.1 AF270792</a> Lactococcus lactis N4 16S ribosomal ...	<u>632</u>	e-179	99
<a href="#">gb AF270791.1 AF270791</a> Lactococcus lactis N1 16S ribosomal ...	<u>632</u>	e-179	99
<a href="#">emb X64887.1 LLRRNAOP</a> L.lactis ribosomal RNA operon encodin...	<u>632</u>	e-179	99
<a href="#">gb M58837.1 STRRR16SD</a> Streptococcus lactis 16S ribosomal RNA	<u>632</u>	e-179	99
<a href="#">dbj AB030918.1 AB030918</a> Lactococcus lactis gene for 16S rRN...	<u>609</u>	e-172	97
<a href="#">gb AY006859.1</a> Uncultured rumen bacterium 16S ribosomal RNA...	<u>587</u>	e-165	99
<a href="#">gb AF107514.1</a> Uncultured bacterium SY2-77 16S ribosomal RN...	<u>581</u>	e-163	97
<a href="#">gb AF095538.1 AF095538</a> Lactococcus lactis clone Pz513 16S r...	<u>581</u>	e-163	97
<a href="#">gb M58836.1 STRRR16SC</a> Streptococcus cremoris 16S ribosomal RNA	<u>557</u>	e-156	96

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz518 en el programa BLAST.

### Distribution of 182 Blast Hits on the Query Sequence

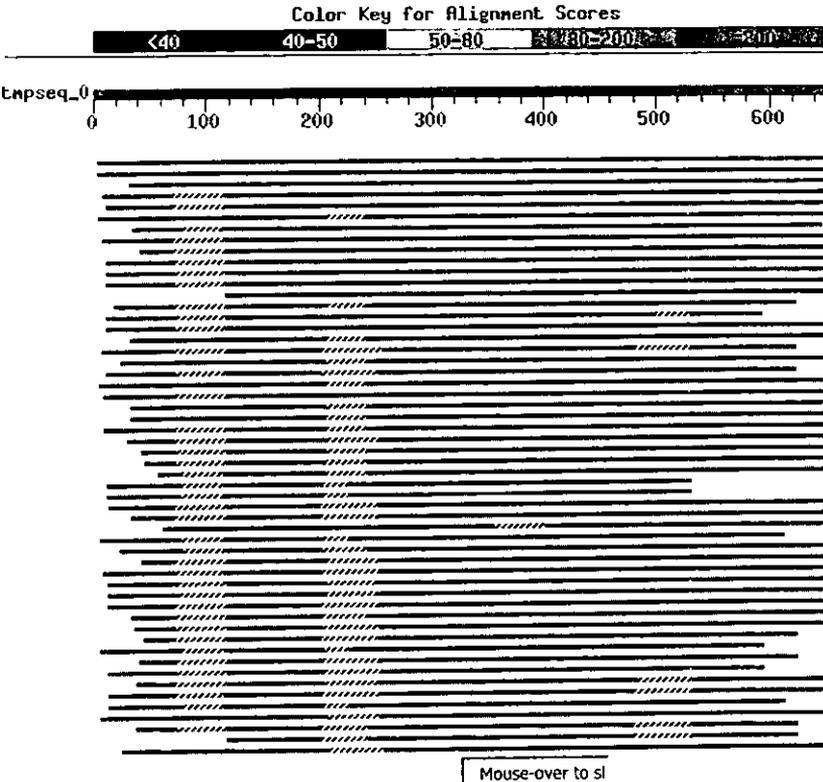


Sequences producing significant alignments:	Score (bits)	E Value	Identities* (%)
<u>gb AF095555.1 AF095555</u> Lactococcus lactis clone Pz518 16S r...	749	0.0	100
<u>gb AF270792.1 AF270792</u> Lactococcus lactis N4 16S ribosomal ...	325	2e-86	84
<u>gb AF270791.1 AF270791</u> Lactococcus lactis N1 16S ribosomal ...	325	2e-86	84
<u>emb X64897.1 LLRRNAQP</u> L.lactis ribosomal RNA operon encodin...	325	2e-86	84
<u>gb M50837.1 STRRR16SD</u> Streptococcus lactis 16S ribosomal RNA	325	2e-86	84
<u>gb AF107514.1 </u> Uncultured bacterium SY2-77 16S ribosomal RN...	293	5e-77	83
<u>gb M58836.1 STRRR16SC</u> Streptococcus cremoris 16S ribosomal RNA	289	8e-76	83
<u>gb AF095538.1 AF095538</u> Lactococcus lactis clone Pz513 16S r...	287	3e-75	85
<u>dbj AB030918.1 AB030918</u> Lactococcus lactis gene for 16S rRN...	283	5e-74	84
<u>gb AF095551.1 AF095551</u> Lactococcus lactis clone Pz505 16S r...	274	5e-71	85

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.

Resultados del análisis de la secuencia de la clona Pz634 en el programa BLAST.

### Distribution of 368 Blast Hits on the Query Sequence



Sequences producing significant alignments:	Score (bits)	E Value	Identities (%)
gb AF095562.1 AF095562 Lactobacillus alimentarius clone Pz6...	<u>1102</u>	0.0	100
gb M58804.1 LBARR16SC Lactobacillus alimentarius 16S riboso...	<u>882</u>	0.0	96
gb AF183558.1 AF183558 Lactobacillus kimchii 16S ribosomal ...	<u>710</u>	0.0	91
gb M59295.1 LBARR16SAH Lactobacillus vermiforme 16S ribosom...	<u>430</u>	e-118	87
gb M58824.1 LBARR16SV Lactobacillus mali 16S ribosomal RNA	<u>424</u>	e-116	86
gb M58817.1 LBARR16SP Lactobacillus farciminis 16S ribosoma...	<u>416</u>	e-114	92
dbj D87678.1 D87678 Pediococcus damnosus DNA for 16S rRNA, ...	<u>408</u>	e-111	86
gb M58821.1 LBARR16SQ Lactobacillus hilgardii 16S ribosomal...	<u>393</u>	e-107	85
dbj AB024300.1 AB024300 Lactobacillus kefir gene for 16S r...	<u>387</u>	e-105	85
emb X95423.1 LLRRL40 L.lindneri 16S rRNA gene, strain LTH 2...	<u>379</u>	e-102	85

\*Esta columna no aparece en el resultado del análisis en el programa BLAST. Los valores que se ilustran corresponden al porcentaje de identidad que presenta la clona analizada con las 10 secuencias con las que produce alineamientos significativos.