



**Centro de Investigación en  
Alimentación y Desarrollo, A.C.**

**DETECCIÓN DE UNIDADES TAXONÓMICAS  
OPERACIONALES EMPLEANDO CEBADORES  
DISEÑADOS A PARTIR DE FRAGMENTOS CON ALTO  
GRADO DE CONSERVACIÓN DEL ARNr 16S**

---

Por:

**Rocío Aracely Valenzuela Gutiérrez**

TESIS APROBADA POR LA

COORDINACIÓN DE ALIMENTOS DE ORIGEN ANIMAL

Como requisito para obtener el grado de

**MAESTRÍA EN CIENCIAS**

## APROBACIÓN

Los miembros del comité designado para la revisión de la tesis de Rocío Aracely Valenzuela Gutiérrez, la han encontrado satisfactoria y recomiendan que sea aceptada como requisito parcial para obtener el grado de Maestra en Ciencias.



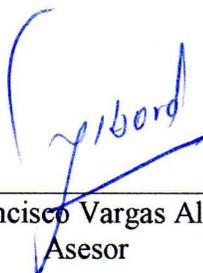
---

Dr. Marcel Martínez Porchas  
Director de Tesis



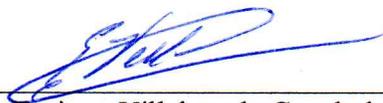
---

Dra. Teresa Gollas Galván  
Asesor



---

Dr. Francisco Vargas Albores  
Asesor



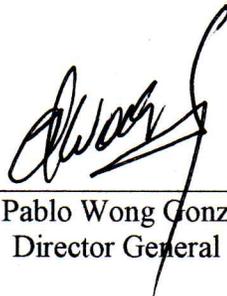
---

M.C. Enrique Villalpando Canchola  
Asesor

## DECLARACIÓN INSTITUCIONAL

La información generada en esta tesis es propiedad intelectual del Centro de Investigación en Alimentación y Desarrollo, A.C. (CIAD). Se permiten y agradecen las citas breves del material contenido en esta tesis sin permiso especial del autor, siempre y cuando se dé crédito correspondiente. Para la reproducción parcial o total de la tesis con fines académicos, se deberá contar con la autorización escrita del Director General del CIAD.

La publicación en comunicaciones científicas o de divulgación popular de los datos contenidos en esta tesis requiere la autorización escrita, del manuscrito en cuestión, del director o directora de tesis. En estos casos siempre se deberá dar los créditos al CIAD



---

Dr. Pablo Wong González  
Director General

## **AGRADECIMIENTOS**

Agradecimientos al Consejo Nacional de Ciencia y Tecnología (CONACYT), por el apoyo prestado durante mi posgrado.

Agradecimientos al Centro de Investigación en Alimentación y Desarrollo, A.C. (CIAD, A.C), por permitirme llevar a cabo mi tesis de Maestría.

Al equipo de trabajo del Laboratorio de Biología Experimental de la Coordinación de Tecnología en Alimentos de Origen Animal (CTAOA), por el apoyo y la confianza brindada a lo largo de mi maestría.

A mi director de tesis, el Dr. Marcel Martínez Porchas, por su confianza, apoyo y siempre alentarme a dar lo mejor de mí.

A mi comité de tesis, la Dra. Teresa Gollas Galván, el Dr. Francisco Vargas Albores y al M.C. Enrique Villalpando Canchola. Gracias por su paciencia, apoyo, consejos y esfuerzo en el desarrollo de mi tesis.

A mi familia, amigos y compañeros.

## **DEDICATORIA**

Primeramente, a Dios por darme fuerzas y ánimos para seguir adelante.

A mi familia querida, por todo su amor, apoyo, consejos y nunca dejarme sola, gracias a ustedes he llegado tan lejos y soy lo que soy ahora. Los amo, ¡Muchas gracias!

Para mi prometido, Tadeo Olivarria Parra. Sin ti esto no hubiera sido posible, gracias por tu amor, apoyo y siempre estar a mi lado. Te amo, ¡Muchas gracias!

Para mi otra familia, Olivarria Parra por todo su amor, apoyo y por considerarme parte de su familia. Los amo, ¡Muchas gracias!

A mis amigos y compañeros, en especial a Estefanía Garibay Valdez por demostrarme lo que es una amistad sincera, gracias por todo. Te quiero, ¡Muchas gracias!

## CONTENIDO

<b>APROBACIÓN</b> .....	2
<b>DECLARACIÓN INSTITUCIONAL</b> .....	3
<b>AGRADECIMIENTOS</b> .....	4
<b>DEDICATORIA</b> .....	5
<b>CONTENIDO</b> .....	6
<b>LISTA DE FIGURAS</b> .....	8
<b>LISTA DE TABLAS</b> .....	10
<b>RESUMEN</b> .....	11
<b>ABSTRACT</b> .....	12
<b>1. INTRODUCCIÓN</b> .....	13
<b>2. ANTECEDENTES</b> .....	15
2.1. Métodos Moleculares para la Detección e Identificación Bacteriana. ....	16
2.2. Identificación y Asignación Taxonómica Bacteriana Mediante Secuenciación del ARNr 16S.....	20
2.2.1. Estructura del ARNr 16S.....	20
2.2.1.1. Regiones conservadas.....	22
2.2.1.2. Regiones variables.....	24
2.3 .Análisis del ARNr 16S en Estudios de Diversidad Bacteriana.....	26
<b>3. HIPÓTESIS</b> .....	31
<b>4. OBJETIVOS</b> .....	32
4.1. General .....	32
4.2. Específicos .....	32
<b>5. MATERIALES Y MÉTODOS</b> .....	33
5.1. Diseño de cebadores.....	33
5.2. Evaluación <i>in silico</i> de los cebadores.....	34
5.3. Evaluación <i>in vitro</i> de los cebadores.....	35
5.3.1. Selección y Colección de Muestras Biológicas.....	35
5.3.2. Extracción y Cuantificación de ADN.....	35
5.3.3. Amplificación del ARNr16S (V3- V4 y V7-V9) .....	36
5.3.4. Preparación de Biblioteca de Amplicones del ARNr16S.....	36
5.3.5. Secuenciación de los Amplicones del ARNr 16S .....	38
5.3.6. Análisis de Datos.....	39
<b>6. RESULTADOS Y DISCUSIÓN</b> .....	40
6.1. Diseño y Evaluación <i>in silico</i> de los Cebadores .....	40
6.2. Evaluación <i>in silico</i> de los Pares de Cebadores .....	45

## CONTENIDO (Continuación)

6.3. Evaluación <i>in vitro</i> de los cebadores.....	60
6.3.1. Evaluación <i>in vitro</i> de los cebadores para la amplificación de las regiones variables 3- 4 y 7-9.....	61
<b>7. CONCLUSIONES.....</b>	<b>74</b>
<b>8. RECOMENDACIONES .....</b>	<b>75</b>
<b>9. REFERENCIAS.....</b>	<b>76</b>

## LISTA DE FIGURAS

<b>Figura</b>		<b>Página</b>
1	Modelo sistemático de técnicas moleculares no dependientes de cultivo, utilizadas en estudios de identificación y caracterización bacterianas en muestras de cualquier ambiente.....	17
2	Procesos de amplificación.....	18
3	Procesos de secuenciación.....	19
4	Estructura secundaria del ARNr 16S (Tomada de Neefs et al., 1990).....	20
5	Representación esquemática del ARNr 16S, indicando las regiones variables (V1-V9), las conservadas (C1-C10) (Tomada de Ortiz-Estrada, 2017).....	21
6	Secuencias detectadas y no detectadas por los cebadores diseñados a partir de fragmentos con alto grado de conservación (1, 2, 3, 4) y los cebadores propuestos por Klindworth y col, 2013 (5) para la amplificación de las regiones variables 3 y 4 del ARNr 16S.....	49
7	Filos detectados por cada uno de los cebadores diseñados a partir de fragmentos con alto grado de conservación (1, 2, 3 y 4) respecto a los cebadores propuestos por Klindworth y col, 2013 (5).....	53
8	Secuencias detectadas y no detectadas por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones variables 7-9 del ARNr 16S.....	54
9	Filos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V7-V9. 6) S-D-Bact-1048-a-S-15 S-D-Bact-1493-a-A-17; 7) S-D-Bact-1054-a-S-17  S-D-Bact-1493-a-A-17.....	55
10	Filos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V3-V4 (1-4) y V7-V9 (6 y 7) y los propuestos por Klindworth y col. 2013 (5).....	58
11	Ilustración de las diferentes regiones del ARNr 16S.....	59

## LISTA DE FIGURAS (Continuación)

<b>Figura</b>	<b>Página</b>	
12	Amplificaciones de ADN de muestras de saliva (S), Agua de mar (A) y suelo agrícola (Sa) empleando cebadores para las regiones variables 3 y 4 (1-5) y las regiones variables 7- 9 (6-7).....	60
13	Análisis de componentes principales de diversidad bacteriana en muestras de agua de mar, saliva y suelo agrícola utilizando asignaciones taxonómicas a nivel de filo.....	61
14	Análisis de diversidad bacteriana en agua de mar. a) Análisis de componentes principales utilizando asignaciones taxonómicas a nivel de filo; b) Alfa y beta diversidad entre los siete diferentes pares de cebadores; c) Perfil taxonómico detectado por cada par de cebador hasta nivel de filo.....	64
15	Análisis de diversidad bacteriana en saliva. a) Análisis de componentes principales utilizando asignaciones taxonómicas a nivel de filo; b) Alfa y beta diversidad entre los siete diferentes pares de cebadores; c) Perfil taxonómico detectado por cada par de cebador hasta nivel de filo.....	66
16	Análisis de diversidad bacteriana en suelo agrícola. a) Análisis de componentes principales utilizando asignaciones taxonómicas a nivel de filo; b) Alfa y beta diversidad entre los siete diferentes pares de cebadores; c) Perfil taxonómico detectado por cada par de cebador hasta nivel de filo.....	68
17	Mapa de calor. Filos detectados por los cebadores para la amplificación de las regiones V3-V4 (1-5) y las regiones V7-V9 (6-7) en tres tipos de muestra: Agua de mar, saliva y suelo agrícola.....	71

## LISTA DE TABLAS

<b>Figura</b>		<b>Página</b>
1	Lista de los 12-meros encontrados con más frecuencia en las regiones conservadas del ARNr 16S (Martínez-Porchas et al., 2017).....	24
2	Filos reportados para tres tipos de ambientes diferentes.....	29
3	Criterios para la selección de cebadores.....	33
4	Secuencias de adaptadores propuestos en el protocolo de Illumina.....	36
5	Propiedades y evaluación <i>in silico</i> de los cebadores sentido y anti-sentido diseñados a partir de las regiones conservadas 3 y 5 del ARNr 16S.....	43
6	Propiedades y evaluación <i>in silico</i> de los cebadores sentido y anti-sentido diseñados a partir de las regiones conservadas 7 y 10 del ARNr 16S.....	44
7	Evaluación <i>in silico</i> de las combinaciones cebadores diseñados a partir de las regiones conservadas 3 y 5 del ARNr 16S. Cebadores sentido: A) ARACTCCTACGGRNG; B) CARACTCCTACGGRNG.....	46
8	Evaluación <i>in silico</i> de las combinaciones cebadores diseñados a partir de las regiones conservadas 7 y 10 del ARNr 16S.....	47
9	Cebadores para la amplificación de las regiones variables 3-4 y 7-9.....	48
10	Porcentajes de cobertura de los fillos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V3-V4 del ARNr 16S.....	51
11	Porcentajes de cobertura de los fillos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V7-V9 del ARNr 16S.....	56

## RESUMEN

Las regiones conservadas que facilitan el diseño de cebadores para la amplificación del ARNr 16S, cuya secuencia es utilizada en la clasificación taxonómica de bacterias, presentan ciertas posiciones cuya variación es muy baja. La identificación y utilización de estas zonas para el diseño de nuevos cebadores, incrementaría la cobertura de secuencias bacterianas en estudios de diversidad bacteriana. El objetivo de este trabajo es detectar un mayor número de secuencias bacterianas empleando estos nuevos cebadores. Para ello, se diseñaron cebadores a partir de los fragmentos con alto grado de conservación para la amplificación de las regiones variables 3-4 y 7-9 del ARNr 16S y se evaluaron *in silico* utilizando la base de datos SILVA, e *in vitro* utilizándolos en secuenciación de muestras biológicas (agua de mar, saliva y suelo agrícola). Se obtuvieron en total seis pares de cebadores, cuatro para la amplificación de las regiones V3-V4 y dos para las regiones V7-V9. Presentando coberturas de secuencias entre 80% y 60%, respectivamente y especificidades del 100%. Todos los cebadores para las regiones V3-V4 detectaron en promedio 459,796 secuencias y entre 60 y 65 filis y los cebadores para las regiones V7-V9 detectaron 159,900 de secuencias y entre 74 y 75 filis de bacterias. En el análisis *in vitro*, los cebadores para la amplificación de las regiones V3-V4, registraron una detección de filis bacterianos similar entre ellos, así como con respecto a los cebadores reportados por Klindworth y col. Los cebadores para la amplificación de las regiones V7-V9, mostraron similitud en cuanto a comunidades bacterianas. En general los pares de cebadores detectaron 25 filis bacterianos. Los cebadores propuestos para la amplificación de las regiones V3-V4, mostraron una buena capacidad de detección de OTU's y además lograron detectar los filis *Acidobacteria*, *Verrucomicrobia* y *Plancomycetes*. Los cebadores propuestos para la amplificación de las regiones V7-V9 fueron los únicos en detectar filis poco estudiados como *Deferribacteres*, *Deinococcus-Thermus* y *Candidatus poribacteria*. Los cebadores propuestos demostraron ser fiables, por lo que sería recomendable, su utilización en estudios de diversidad bacteriana.

**Palabras clave:** Comunidades bacterianas, Regiones conservadas del ARNr 16S, cebadores, clasificación taxonómica, secuenciación masiva.

## ABSTRACT

The conserved regions that facilitate the design of primers for the amplification of the 16S ribosomal gene, whose sequence is used in the taxonomic classification of bacteria, have certain positions whose variation is very low. The identification and use of these zones for the design of new primers, would increase the coverage of bacterial sequences in studies of bacterial diversity. The objective of this work is to detect a greater number of bacterial sequences using these new primers. To do this, primers were designed from fragments with a high degree of preservation for the amplification of variable regions 3-4 and 7-9 of the 16S gene and were evaluated *in silico* using the SILVA database, and *in vitro* using them in sequencing of biological samples (seawater, saliva and soil). A total of six pairs of primers were obtained, four for the amplification of the V3-V4 regions and two for the V7-V9 regions. Presenting sequence coverage between 80% and 60%, respectively and 100% specificities. All the primers for the V3-V4 regions detected on average 459.796 sequences and between 60 and 65 phyla and the primers for the V7-V9 regions were detected 159.900 of sequences and between 74 and 75 phylum of bacteria. In the *in vitro* analysis, the primers for the amplification of the V3-V4 regions recorded a similar detection of bacterial phyla between them, as well as with respect to the primers reported by Klindworth et al. The primers for the amplification of the V7-V9 regions, showed similarity in terms of bacterial communities. In general, the pairs of primers detected 25 bacterial phyla. The primers proposed for the amplification of the V3-V4 regions showed a good detection capacity of OTUs and also managed to detect the *Acidobacteria*, *Verrucomicrobia* and *Plancomycetes*. The primers proposed for the amplification of the V7-V9 regions were the only ones to detect poorly studied phyla such as *Deferribacter*, *Deinococcus-Thermus* and *Candidatus poribacteria*. The proposed primers proved to be reliable, so it would be advisable to use them in studies of bacterial diversity.

**Key words:** Bacterial communities, conserved regions of 16S rRNA, primers, taxonomic classification, massive sequencing.

# 1 INTRODUCCIÓN

El estudio de la diversidad bacteriana a ha ido en aumento en los últimos años, convirtiéndose en una línea de investigación de alto impacto en ecología. Sin embargo, el estudio de las comunidades bacterianas no es sencillo, debido a que la gran mayoría de las bacterias existentes no son cultivables (Torsvik y Ovreas, 2002). Para contrarrestar esta problemática, se han desarrollado técnicas independientes de cultivo basadas en el análisis del material genético de bacterias contenidas en cualquier muestra biológica. Una de las más utilizadas actualmente es el análisis de biomarcadores taxonómicos, particularmente el ARNr 16S. La secuencia de este gen es utilizada como un código de barras para la identificación de taxones, a fin de establecer relaciones filogenéticas y obtener una estimación más precisa de la diversidad bacteriana (Yang et al., 2016).

En todos los organismos procariotas el ARNr 16S, es conocido como un marcador universal debido a su presencia ya que contiene regiones conservadas que permiten el diseño de cebadores que facilitan su análisis y regiones variables que contienen información para estudios de filogenia y taxonomía (Baker et al., 2003). No todas las regiones variables proporcionan la misma información taxonómica, por lo que existen diversos puntos de vista con respecto a cuál o cuáles son las regiones que proporcionan mayor información. Por lo general, se emplean las regiones centrales del gen; de ellas, las más analizadas son el segmento que incluye las regiones variables tres y cuatro (V3-V4), debido a que presentan una elevada variabilidad que proporciona una capacidad resolutive mayor para distinguir entre especies; además, existe en bases de datos, un mayor número de secuencias parciales correspondientes de estas regiones (Chakravorty et al., 2007; Kim y Morrison, 2011). Las regiones variables a analizar, así como la eficiencia de los cebadores seleccionados, tienen un efecto determinante en la estimación de la diversidad bacteriana en muestras de estudio (González et al., 2015). Por lo que, el primer seleccionado debe reaccionar con la mayoría de las secuencias; en caso contrario, una cobertura deficiente de secuencias del ARNr 16S, podría sesgar de manera significativa los resultados de una investigación.

Los cebadores más usados en estudios de diversidad bacteriana son los diseñados por Herlemann y colaboradores en 2011, los cuales amplifican las regiones variables V3-V4. Posteriormente fueron evaluados por Klindworth y colaboradores en 2013. Estos autores reportaron una mayor cobertura de secuencias (80%) en comparación con un listado de cebadores ya reportados en un análisis *in silico*. Actualmente, los cebadores reportados por Klindworth y colaboradores son empleados en plataformas de Illumina (MiSeq y HiSeq), entre otras que utilizan la tecnología de secuenciación por síntesis (Ju et al., 2006). Usando estos cebadores y la capacidad de la plataforma MiSeq, por ejemplo, es posible conseguir hasta 25 millones de lecturas; convirtiéndose en una de las tecnologías de secuenciación más utilizadas en estudios de diversidad bacteriana en muestras biológicas (López de Heredia, 2016).

Sin embargo, aún los cebadores que se consideran universales presentan limitaciones para detectar todas las secuencias de especies bacterianas presentes en una muestra medioambiental (Baker et al., 2003; Huws et al., 2007). Es decir, hay una porción de microorganismos que para poder amplificar su ARNr 16S, probablemente se necesitan otros cebadores. Esto se debe a que las secuencias de las regiones conservadas, que sirven de base para la creación de cebadores para la amplificación de regiones variables, no son totalmente conservadas (Martínez-Porchas et al., 2017). Esto resulta en una cobertura deficiente de secuencias de microorganismos en estudios de taxonomía y filogenia.

Considerando que el número de secuencias disponibles aumenta constantemente, una revisión de las regiones conservadas y las regiones con mayor grado de conservación, podría ser de gran utilidad para la ratificación de cebadores diseñados o proponer variaciones que presenten con una mayor cobertura. Por tanto, el objetivo de este estudio es la detección de regiones con alto grado de conservación dentro del gen del ARNr 16S para proponer nuevos cebadores con mayor cobertura *in silico* y, a partir de ahí, lograr la amplificación de un mayor número de secuencias bacterianas presentes en muestras biológicas.

## 2 ANTECEDENTES

Los microorganismos son responsables de una parte importante de los ciclos biogeoquímicos y, por lo tanto, influyen significativamente en la vida terrestre. Sin embargo, se estima que solo se conoce del 0.1 al 10% de las bacterias contenidas en el medioambiente (Torsvik y Ovreas, 2002). Por ello, el conocimiento de la vida microbiana, así como el papel que ésta juega en el ambiente, es todavía poco entendido. Más difícil aún es tratar de conocer la diversidad microbiana y predecir sus funciones; por ejemplo, se ha estimado que existen alrededor de  $3.6 \times 10^{28}$  microorganismos con una densidad promedio de  $5 \times 10^5$  células  $\text{mL}^{-1}$  en el ambiente marino (Cock et al., 2010), mientras que una muestra de suelo podría contener hasta  $4 \times 10^6$  diferentes taxas (Curtis et al., 2002). Los microorganismos tienen la capacidad de regular la disponibilidad de nutrientes en los distintos ambientes, favoreciendo la recirculación de estos a través de distintas cadenas tróficas; en el ambiente terrestre, por ejemplo, favorecen la fertilidad del suelo y, consecuentemente, el desarrollo de la vegetación que constituye la base de la cadena alimentaria (Guerrero y Berlanga, 2005). Por ello, el estudio de las comunidades bacterianas presentes en los diversos ambientes es relevante, ya que nos permite obtener conocer la diversidad biológica en la muestra y entender su participación en los diferentes ciclos biogeoquímicos.

Por otro lado, el conocimiento de las comunidades bacterianas permite determinar las capacidades bioquímicas de las diferentes poblaciones que la conforman (Huber et al. 2004, Domínguez et al. 2006, Aly et al. 2008), facilitando el uso de cepas con actividades funcionales aprovechables por el hombre. El estudio de la diversidad bacteriana va en aumento, convirtiéndose en una de las líneas de investigación más relevantes en ecología; sin embargo, se estima que la mayoría de las bacterias son no cultivables o difíciles de cultivar, probablemente debido al desconocimiento de sus requerimientos nutricionales y fisiológicos para su crecimiento. La incapacidad de cultivar a los microorganismos representa un serio problema, debido a que son de gran importancia para la seguridad alimenticia y la salud pública. Muchas especies potencialmente peligrosas sobreviven y persisten en alimentos procesados, leche pasteurizada, agua potable y en el ambiente en

general (Sardessai, 2005). Para contrarrestar esta limitante, se ha recurrido a métodos complementarios a las técnicas tradicionales de aislamiento e identificación bacteriana (identificación fenotípica, pruebas bioquímicas), tales como las técnicas moleculares. Las técnicas moleculares permiten el análisis de secuencias de genes o genomas completos de bacterias cultivables y no cultivables de diferentes ambientes (Torsvik y Ovreas 2002, Croci et al. 2007).

## 2.1 Métodos Moleculares para la Detección e Identificación Bacteriana.

La dificultad de estudiar bacterias dependientes de cultivo, fue contrarrestada por la capacidad de aislar y secuenciar el material genético de una comunidad bacteriana presente en cualquier ambiente y su posterior análisis mediante técnicas moleculares y análisis filogenéticos (Figura 1). La reacción en cadena de la polimerasa (PCR) desarrollada por Kary Mullis en 1987, es una de las técnicas moleculares más importantes que contribuyeron de manera significativa al estudio de los ácidos nucleicos, revolucionando la biología molecular. Esta técnica permite llevar a cabo amplificaciones sucesivas de ADN o ARN en una muestra para su posterior análisis.

El desarrollo de la PCR y el incremento en el estudio de las comunidades bacterianas en distintos nichos, aumentó el desarrollo y uso de algunas herramientas moleculares; entre ellas los métodos de Fingerprinting (huella génica): Polimorfismo del ADN amplificado al azar (RADP), utilizado para el aislamiento y caracterización de bacterias ácido lácticas en queso (Ramos et al., 2009), Análisis de fragmentos de restricción del gen ribosomal 16S (ADRA), reportado en estudios de diversidad en sedimentos de (Ribeiro et al., 2013) y en la identificación de bacterias ácido lácticas en mosto de uva y vino (Rodas et al., 2003), Polimorfismo en el tamaño de los fragmentos de restricción (T-RFLP), en estudios de diversidad bacteriana en la producción de mezcal (Damián et al., 2012), Electroforesis en gel de gradiente desnaturalizante o temperatura (DGGE, TGGE) utilizado en estudios de comunidades bacterianas en ecología marina (Lyautey et al., 2004) y estudios de



tiempo y trabajo (Ortiz et al., 2017). Así el siguiente paso en el estudio de ácidos nucleicos fue el desarrollo de las plataformas de secuenciación de alto rendimiento también conocidas como secuenciación de nueva generación (NGS: Next-Generation Sequencing).

Entre las NGS se encuentran Roche 454 (Hoffmann-La Roche), Illumina/Solexa (Illumina Inc.), SOLiD (Applied Biosystems), Ion Proton (Thermo Fisher Scientific), Helicos (Helicos Biosciences), PacBio (Pacific Biosciences), Nanopore (Oxford Nanopore Technologies), FRET (Thermo Fisher Scientific). Las últimas cuatro son catalogadas como tecnologías de tercera generación y, característicamente, tienen una elevada capacidad de lecturas por corrida y producen lecturas de mayor tamaño; sin embargo, en la actualidad algunas se encuentran en desarrollo y/o no se comercializan (López de Heredia, 2016). La primera tecnología NGS que se desarrolló fue la de pirosecuenciación 454 (Roche, 2005), la cual emplea PCR por emulsión. Posteriormente, salió al mercado la técnica basada en secuenciación tipo puente o secuenciación por síntesis (Illumina, 2006)

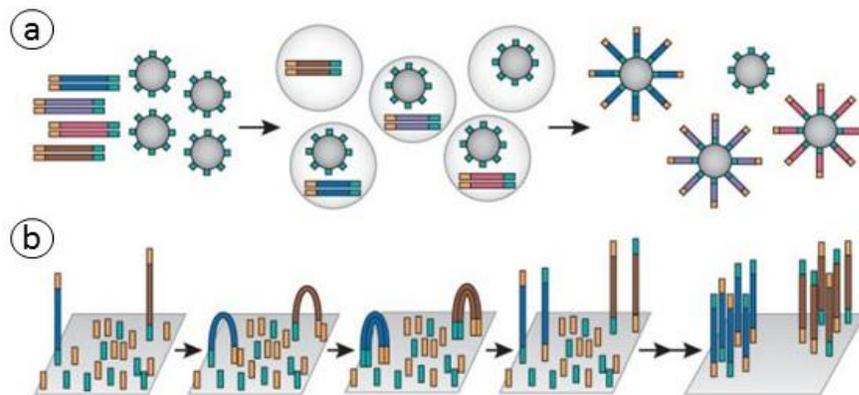


Figura 2. Procesos de amplificación. (a) PCR en emulsión, la reacción consiste en una emulsión de aceite-agua que encapsula complejos de ADN y nanoesferas dentro de pequeñas gotas de agua. Posteriormente se lleva a cabo la amplificación, de tal modo que cada nanoesfera queda cubierta por miles de copias de la misma secuencia molde (454, SOLiD, Ion Proton). (b) PCR tipo puente, consiste en la ligación del ADN a una superficie de cristal con adaptadores, posteriormente da inicio la amplificación tipo puente y se repite la operación hasta conseguir muchas localizaciones aisladas que contienen múltiples copias idénticas de un fragmento de ADN (Illumina /Solexa). Tomada de la Revista Médica Genética, 2017. NGS: Secuenciación de Segunda Generación. <https://revistageneticamedica.com/blog/ngs-secuenciacion/>.

(Figura 2 y 3). Ambas son catalogadas como técnicas de alto rendimiento, debido a que permiten secuenciar millones de fragmentos de ADN o ARN, además de tener un menor costo de operación (Schendure y Ji, 2008).

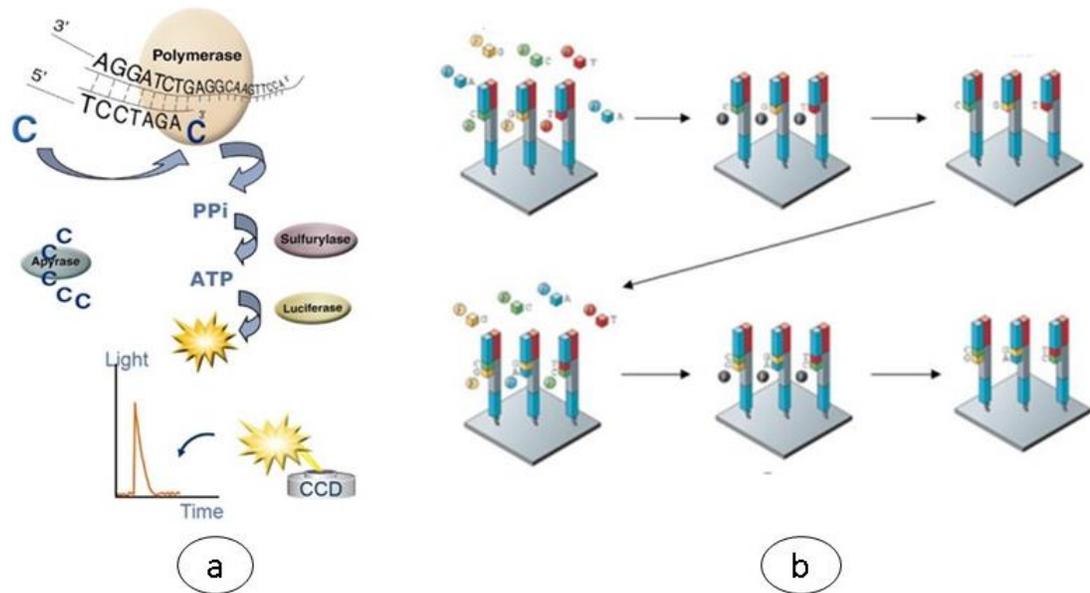


Figura 3. Procesos de secuenciación. (a) Pirosecuenciación, el cebador se une con la cadena y da paso a la reacción en cadena de la polimerasa. Se une uno de los cuatro desoxirribonucleótidos trifosfato, si resulta ser el complementario a la cadena molde se liberará la molécula pirofosfato (PPi), que a su vez reaccionará con la enzima luciferasa y emitirá una luz fluorescente que será detectada por una cámara (Tomada de Biosonda S.A. 2016. <http://www.biosonda.cl/producto/pyromark>). (b) Secuenciación por síntesis, consiste en la incorporación de cuatro nucleótidos terminadores cada uno marcado con un fluoróforo diferente de forma que se detecta cada nucleótido cuando se produce la unión a la cadena molde (Tomada de Metzker, M.L., 2010).

Existen dos aplicaciones en la secuenciación de nueva generación: Secuenciación todo el material genético (Shotgun) y secuenciación dirigida (Genes). La secuenciación dirigida se basa en el análisis de genes diana o regiones del genoma que presentan una variabilidad y una divergencia genética significativa a nivel de especie, además de poseer sitios conservados adyacentes a estas regiones variables que permiten el diseño de cebadores universales para su amplificación por PCR secuenciación de forma fácil, reproducible y

precisa (Kress y Erickson, 2012). En estudios de taxonomía bacteriana, el análisis del gen que codifica para el ARN de la subunidad ribosómica 16S (ARNr 16S) es la herramienta más utilizada. El análisis de la secuencia del ARNr 16S, ha permitido el establecimiento de firmas moleculares a varios niveles taxonómicos, utilizadas como la base para la identificación bacteriana por comparación filogenética (Theron y Cloete, 2000).

## 2.2 Identificación y Asignación Taxonómica Bacteriana Mediante Secuenciación del ARNr 16S

### 2.2.1 Estructura del ARNr 16S

El ARNr 16S es un polirribonucleótido (aprox. 1500 nucleótidos) codificado por el gen *rrs*, es de cadena sencilla y adquiere una estructura secundaria caracterizada por presentar segmentos de doble cadena que permiten la formación de asas y hélices (Neefs et al., 1990) (Figura 4).

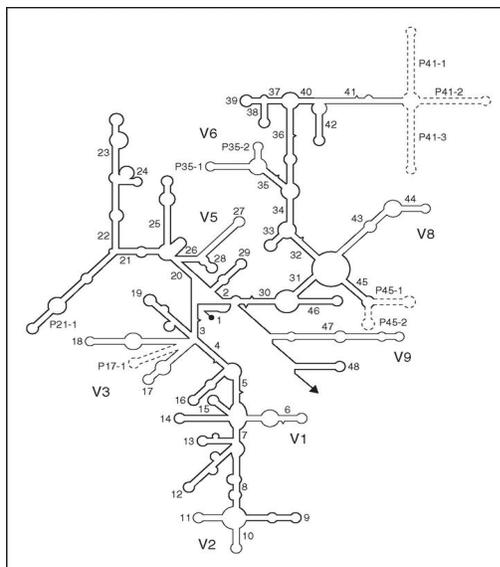


Figura 4. Estructura secundaria del ARNr 16S (Tomada de Neefs et al., 1990).

Así mismo, contiene nueve regiones menos conservadas o variables (Baker et al. 2003), que aportan información útil para estudios de filogenética y taxonomía. Toda la variabilidad genética se concentra en estas regiones, mas no en la misma proporción; algunas de estas presentan mayor variabilidad, permitiendo obtener asignaciones taxonómicas más precisas (Liu et al., 2007; Youssef et al., 2009). La mayoría de los estudios basados en el ARNr 16S, se concentran en las regiones más centrales ya que aportan la mayor cantidad de información discriminatoria (Wang et al., 2016). La diferencia en la variabilidad se debe a que cada región variable posee su propia tasa de mutación y, por ende, tienen diferente grado de utilidad en la identificación de ciertos grupos de microorganismos (Claesson et al., 2010; Kunin et al., 2010; Mizrahi-Man et al., 2013). También contiene diez regiones conservadas utilizadas como base para el diseño de cebadores o iniciadores universales. Estos permiten la amplificación de las secuencias de diversas regiones variables de los ARNr16S de la mayoría de microorganismos presentes en una comunidad (Figura 5).



Figura 5. Representación esquemática del ARNr 16S, indicando las regiones variables (V1-V9), las conservadas (C1-C10) (Tomada de Ortiz-Estrada, 2017).

El RNAr 16S se considera como un poderoso marcador universal, presente en todos los organismos procariotas conocidos. Además, su estructura y su función están altamente conservados. Sin embargo, los cambios aleatorios en su secuencia generan suficiente variabilidad para diferenciar no sólo los organismos más alejados, sino también para poder diferenciar especies, cepas o variedades (Woese, 1987).

2.2.1.1 Regiones conservadas. En el diseño y creación de oligonucleótidos iniciadores o cebadores se utilizan las regiones conservadas del gen, ya que, al flanquear las regiones variables, es posible su amplificación. La selección de los cebadores es de suma importancia para lograr la mayor cobertura de secuencias en cualquier estudio. Para la secuenciación completa del ARNr 16S, se utilizan cebadores diseñados para las regiones conservadas extremos 5' y 3' del gen que originan amplicones alrededor de 1,500 pb. Pese a la gran capacidad y velocidad de las tecnologías de NGS para la generación de millones de secuencias bacterianas simultáneamente, solamente llegan a amplificar fragmentos de hasta 600 pares de bases (Rodríguez- Armengol, 2012). En estas circunstancias, el gen completo no puede ser secuenciado en una sola reacción y, en forma alternativa, se utilizan cebadores que permiten la amplificación de fragmentos de menor tamaño (Rodicio y Mendoza, 2004; Cadena et al., 2016).

Se han utilizado diferentes cebadores dirigidos a regiones variables, garantizando una amplia cobertura y buena amplificación (Hiergeist, et al., 2016, Yang et al., 2016); sin embargo, ningún cebador es verdaderamente universal, es decir, es probable que solo una porción de la diversidad de especies que se encuentran en cualquier ambiente, sea cubierta por alguna combinación de cebadores (Martínez-Porchas et al., 2017). Además, la cobertura se puede ver alterada por factores intrínsecos, como el diseño de los cebadores, el número de ambigüedades, reactivos químicos usados, condiciones de amplificación y otros sesgos de PCR; o bien por factores extrínsecos, como: tipos de las muestras, composición bacteriana y ambiente, e inhibidores de la PCR transportados en el proceso de muestreo (Albertsen et al., 2015). Por lo que, la información proporcionada acerca de la cobertura y condiciones de cualquier par de cebadores, representa información relevante.

Estudios sobre la eficacia de cebadores que se enfocan al estudio de muestras biológicas, han sido útiles para aumentar el panorama de la investigación de comunidades bacterianas (Cruaud et al, 2014); sin embargo, los factores intrínsecos y extrínsecos que afectan en el desempeño de estos, no permiten concluir si tienen la mejor cobertura posible, ya que solo permiten inferir si un par de cebadores es mejor que otros, pero no aportan información concluyente acerca de la cobertura. Otra limitante para la obtención de una mayor

cobertura taxonómica por parte de los cebadores en estudios de diversidad bacteriana, es que fragmentos de las regiones conservadas utilizadas para la creación de estos presentan variaciones, por lo que no son tan conservadas como se espera. Para contrarrestar esta limitación se ha propuesto el uso de cebadores degenerados, para la amplificación de ADN, cubriendo un mayor número de genes de procariotas inespecíficos. Los cebadores degenerados se diseñaron inicialmente insertando degeneraciones después de varios alineamientos mediante programas sofisticados (Linhart y Shamir, 2002; Najafabadi et al., 2008; Qu et al., 2012). Sin embargo, aún falta información para comprender las variaciones que poseen en las regiones conservadas del ARNr 16S (Martínez-Porchas et al., 2017).

Para medir el grado de conservación e inferir que posiciones en la secuencia se conservan verdaderamente, Martínez-Porchas y colaboradores (2017) realizaron un análisis con todos los cebadores diseñados reportados para cada región conservada del ARNr 16S. Considerando toda la información contenida en bases de datos, realizaron un alineamiento de todos estos; posteriormente, se llevó a cabo un escaneo de secuencias (posición por posición), incluyendo degeneraciones y combinaciones de regiones conservadas, mediante barrido con fragmentos de 12 nucleótidos (12-meros). Como resultado, los autores reportaron que aun en las regiones conservadas existen variaciones; sin embargo, también registraron la existencia de fragmentos con un alto grado de conservación. Estas zonas altamente conservadas se mostraron con mayor frecuencia en cada conjunto de secuencias registradas por regiones conservadas analizadas, particularmente en las regiones C3, C4, C5 y C7, que mostraron porcentajes arriba del 95%. Sin embargo, el análisis también mostro que las regiones extremas en el gen, registraron frecuencias más bajas, entre 30 y 40% de las más de 513,000 secuencias reportadas en la base de datos SILVA SSU Ref NR 99 (Versión 123) (Tabla 1).

Debido a la poca cantidad de secuencias pertenecientes a las regiones extremas del gen los estudios basados en el ARNr 16S, se concentran en las regiones más centrales (Yarza et al., 2014). Pese a la falta de información de las regiones más externas y a la presencia de variaciones en todas las regiones conservadas, las zonas altamente conservadas

posibilitan el diseño de cebadores más fiables que permita profundizar y ampliar los estudios taxonómicos en comunidades bacterianas.

Tabla 1. Lista de los 12-meros encontrados con más frecuencia en las regiones conservadas del ARNr 16S (Martínez-Porchas et al., 2017)

Región conservada	Secuencia (12-mero)	Número	Porcentaje
01	ATYMTGGCTCAG	195,901	38.16%
02	SYGGCGNACGGG	405,570	79.01%
03	GGRNGGCNGCAG	500,253	97.46%
04	CVGCNGCYGCGG	496,412	96.71%
5a	GMGGTGAAATKC	382,156	74.45%
5b	TAGAWACCCNNG	493,348	96.11%
6a	RAATWGRCGGGG	501,792	97.76%
6b	GTGGTTTAATTC	389,530	75.89%
6c	GARGAACCTTAC	393,614	76.68%
7a	GYYGYCGTCAGC	499,976	97.40%
7b	CGAGCGCAACCC	489,290	95.32%
8a	AGGYGGGGAYGA	454,807	88.60%
8b	GCKACACACGYG	382,857	74.59%
09	GYACWCWCCGCC	388,911	75.77%
10	AGTCRTAACAAAG	172,918	33.69%

2.2.1.2 Regiones variables. Las regiones variables permiten establecer relaciones filogenéticas entre microorganismos, por lo que son ampliamente utilizadas en estudios de diversidad bacteriana (Liu et al., 2007). Cada región muestra diferentes grados de variabilidad genética (Youssef et al., 2009) por lo que su utilidad es diferente y, en algunos casos, para diferentes filos bacterianos. Por otro lado, debido a que las nuevas tecnologías de secuenciación producen fragmentos cortos (600 pb), para estudios taxonómicos se recomienda el uso una o dos regiones variables (Baker et al., 2003). Existen diversos puntos de vista con respecto a cuál o cuáles son las regiones más eficientes para realizar las asignaciones taxonómicas más precisas. Varios autores han estudiado distintas regiones variables y sus combinaciones, con el fin de reconocer aquellas que permitan obtener mayor información taxonómica (Liu et al., 2008; Kim y Morrison, 2011; Shakya et al., 2013).

Se ha reportado que para la clasificación de la microbiota intestinal, usar el segmento que abarca las regiones V4-V5 del ARNr 16S, es más eficiente que el segmento comprendido por las regiones V7-V8 para la clasificación de la microbiota intestinal (Liu et al., 2008). Shakya y colaboradores (2013), reportaron que el amplicón de la región V3 fue eficiente para la identificación de linajes de *Euryarchaeota*, pero no detectaron especies del filo *Pyrobaculum*; además, realizaron combinaciones de cebadores que amplifican la región V4-V8, que tendieron a subestimar *Crenarchaeota* y *Euryarchaeota*. Sin embargo, se reportó que los segmentos que comprenden las regiones V1-V3 y V1-V4, son la más utilizadas para clasificación microbiana en muestras ambientales debido a que presentan ventajas tales como su elevada variabilidad y la existencia de un mayor número de secuencias parciales de estas regiones en las bases de datos (Chakravorty et al., 2007; Kim y Morrison, 2011). Martínez-Porchas y colaboradores (2016) realizaron un análisis *in silico* para determinar la eficiencia de regiones variables del gen ARNr 16S para propósitos de clasificación taxonómica, concluyendo que el fragmento interno V3-V5 puede ser considerado un candidato por su alto rendimiento para la secuenciación masiva. Al utilizar el fragmento V3-V5 se obtuvo que el 55.6% de las secuencias obtenidas recibieron la misma clasificación (subespecie), que al haber utilizado la secuencia completa del 16S (V1-V9).

Sin embargo, ninguna región parcial podría mejorar la utilidad de la secuencia completa para estimar con precisión la riqueza de especies (Kim et al., 2011). Por lo que, los resultados en estudios de diversidad con fragmentos parciales, probablemente conduzcan a conclusiones parciales (Guo et al., 2013). Por otro lado, de la selección adecuada de los cebadores, es de suma importancia seleccionar la plataforma informática para el análisis de los datos, con el fin de obtener la mayor información taxonómica posible. La amplificación de las regiones variables 3 y 4 es la más utilizada en estudios de diversidad bacteriana (Huse et al., 2008). Para ello, los cebadores reportados por Klindworth y colaboradores en 2013 se han convertido en un estándar y actualmente, estos son empleados en las plataformas con tecnologías de secuenciación más utilizadas (Miseq o HiSeq XTen de Illumina, entre otras).

### 2.3 Análisis del ARNr 16S en Estudios de Diversidad Bacteriana

El análisis de la secuencia del ARN ribosómico (ARNr) 16S, es una de las herramientas moleculares más utilizadas en estudios de detección y clasificación filogenética de bacterias (Pace et al., 1986; Woese, 1987). Esto consiste primeramente en la amplificación del gen, gracias a la reacción en cadena de la polimerasa (PCR) empleando cebadores diseñados a partir de regiones conservadas como ya se mencionó. La selección de las regiones a amplificar y el tamaño de los amplicones que serán secuenciados dependerán del interés del estudio. Posterior a la secuenciación, se realiza el procesamiento de los datos, el cual está dividido en tres fases.

La primera fase consiste en el análisis de la calidad de la secuencia, la detección y eliminación de quimeras (lecturas creadas a partir de distintos ARNr16S) y posteriormente la lectura de la longitud después de la eliminación de las bases de baja calidad (Weinstock, 2012). La segunda fase consiste en la generación de datos, es decir, la creación de tablas de taxones que se comparan con secuencias del ARNr 16S contenidas en bases de datos (SILVA: Bases de datos de ARN ribosomal de alta calidad, Ribosomal Database Project (RDP), EzTaxon) o mediante el uso de programas sofisticados que agrupan las lecturas en OTUs (Unidades taxonómicas operacionales). Estas permiten diferenciar los datos obtenidos mediante similitud para la clasificación de microorganismos (Weinstock, 2012). Por último, la tercera fase consiste en utilizar los datos para producir árboles filogenéticos u otras representaciones de la similitud entre las comunidades, abundancia, diversidad, y otro descriptor de la estructura de una comunidad bacteriana (Abubucker et al., 2012; Caporasso et al., 2016).

En estudios de diversidad se emplean diferentes medidas para la estimación de especies como: diversidad alfa ( $\alpha$ ), beta ( $\beta$ ) y gamma ( $\gamma$ ). La diversidad alfa se refiere a la diversidad de especies a lo largo de todas las subunidades locales relevantes (hábitat) y abarca dos variables: la riqueza de especies y la abundancia relativa de especies (Carmona y Carmona, 2013). Existen diferentes índices de diversidad alfa, que emplean las variables de riqueza y abundancia relativa; por ejemplo, los índices de Fisher, Simpson y Shannon-

Wiener (Halffer et al., 2001). La diversidad beta, se comprende como el diferencial o cambio en la composición de especies de una comunidad determinada (diversidad alfa) y la diversidad total de un conjunto de comunidades, es decir, representa la diversidad de especies entre comunidades que no está compartida y refleja la respuesta de las especies a la heterogeneidad espacial (Whittaker, 1960; Villarreal et al., 2006) y se define mediante:

$$\text{Diversidad } \beta = (S_1 - C) + (S_2 - C)$$

*Dónde: S= Riqueza de especies en dicha comunidad y C= Riqueza de especies compartidas entre comunidades.*

Por último, la diversidad gamma se define como la “diversidad total de especies en una región, donde se incluyen varias comunidades (diversidad alfa  $\alpha$ ) o el recambio entre regiones (diversidad beta  $\beta$ )” (Whittaker, 1960; Halffer et al., 2001) y puede ser calculada por:

$$\text{Diversidad } \gamma = \alpha * \beta$$

Otra herramienta complementaria utilizada para el estudio de diversidad, es el análisis estadístico de componentes principales (ACP). Esta técnica permite estudiar las relaciones entre variables cuantitativas (Palm, 1998), teniendo como objetivo la reducción de éstas perdiendo la menor cantidad de información posible. Este método ha sido aplicado en un variado número de problemas en la ciencia como en tecnología de alimentos y desarrollo de productos, en agronomía, evaluación de suelos productivos, entre otros (Ruiz, 2012).

El análisis de la secuencia del ARNr 16S, ha sido una herramienta útil para estudios de filogenia y taxonomía bacteriana, complementado por técnicas estadísticas como las anteriormente mencionadas y programas bioinformáticos. Además, su estudio ha permitido inferir la estructura funcional de una comunidad bacteriana, es decir, sus características metabólicas consecuencia de la diversidad genética taxo-específica (Zak y col., 1994), permitiendo, además la estimación de la diversidad bacteriana, la

determinación del rol de las comunidades bacterianas en el ambiente en el que se desarrollan. La relevancia de la utilización de estas medidas en el estudio de comunidades bacterianas, radica en aumentar en el conocimiento de la diversidad biológica global, así como el entendimiento de los ciclos biogeoquímicos de los que son partícipes. La combinación de técnicas tradicionales de cultivo y las técnicas moleculares en conjunto con el uso de técnicas analíticas, se ha vuelto una aproximación fundamental para abordar, de una manera más completa, el estudio de las poblaciones bacterianas en el medioambiente (López y Zaballos 2005).

Existen numerosos estudios de diversidad bacteriana mediante el análisis del ARNr 16S de diversos ambientes. En general, casi cualquier hábitat contiene gran cantidad y diversidad de vida microbiana (Youssef et al., 2015). Algunos de los microbiomas más estudiados, son los provenientes de muestras de agua (lagos, arroyos, aguas oceánicas, agua residual), suelos (agrícolas, desérticos, sedimentos residuales); así como de procedencia animal (intestinos, variedades de queso, desechos fecales) incluyendo la humana (intestinos, cavidad bucal) (Referencias en tabla 2). Cada uno distinguido por poseer una amplia diversidad de filos bacterianos, relativamente diferente.

Tabla 2. Filos reportados en diferentes tipos de ambientes.

Ambiente	Filos	Referencia
Suelos áridos	<i>Acidobacteria</i> <i>Cyanobacteria</i> <i>Chloroflexi</i> <i>Firmicutes</i> <i>Verrucomicrobia</i> <i>Thermomicrobia</i> <i>Actinobacteria</i> <i>Proteobacteria</i> <i>Gemmatimonadetes</i>	Fierer et al., 2007 Bachar et al., 2010 Bachar et al., 2012
Agua de Mar	<i>Acidobacteria</i> <i>Actinobacteria</i> <i>Proteobacteria</i> <i>Bacteroidetes</i> <i>Cyanobacteria</i> <i>Verrucomicrobia</i> <i>Planctomycetes</i> <i>Firmicutes</i>	Campbell et al., 2011
Cavidad bucal (Saliva)	<i>Firmicutes</i> <i>Actinobacteria</i> <i>Proteobacteria</i> <i>Bacteroidetes</i> <i>Fusobacteria</i> <i>Synergistetes</i> <i>Spirochaetes</i>	Aas et al., 2005 Cameron et al., 2015 Acharya et al., 2017

El ARNr 16S es la macromolécula más ampliamente utilizada en estudios de diversidad taxonómica de bacterias, debido a que presentan regiones variables que son utilizadas como base para la clasificación de microorganismos y regiones conservadas usadas para el diseño de cebadores que permiten la amplificación de estas regiones variables. Se ha reportado que las regiones conservadas poseen variaciones en su secuencia, esto podría resultar en una cobertura deficiente de secuencias de microorganismos por parte de los

cebadores y tener un efecto negativo en la estimación de la diversidad. Sin embargo, las regiones conservadas también poseen fragmentos cuya variación es muy baja, por lo que la identificación y utilización de estos fragmentos para el diseño de nuevos cebadores podría incrementar la cobertura de secuencias bacterianas, permitiendo la obtención de más información en estudios de filogenia y taxonomía.

### **3 HIPÓTESIS**

La utilización de cebadores diseñados a partir de las regiones más conservadas del gen codificante para el ARNr 16S tendrá mayor cobertura y, consecuentemente, permitirá la detección de más unidades taxonómicas operacionales en el análisis de muestras ambientales.

## 4 OBJETIVOS

### 4.1 General

Detectar un mayor número de secuencias del gen codificante para el ARNr 16S bacteriano presentes en muestras biológicas empleando cebadores diseñados a partir de regiones con mayor grado de conservación del ARNr 16S.

### 4.2 Específicos

1. Diseñar cebadores a partir de las secuencias con mayor grado de conservación de las regiones C3, C5, C7 y C10 del ARNr 16S.
2. Evaluar *in silico* la cobertura de los cebadores diseñados, utilizando la base de datos SILVA SSU NR Ref 99 (Versión 128).
3. Evaluar *in vitro* la capacidad de los mejores cebadores diseñados, amplificando y secuenciando muestras biológicas.
4. Comparar el perfil bacteriano obtenido de la secuenciación masiva empleando los diversos cebadores diseñados.

## 5 MATERIALES Y MÉTODOS

### 5.1 Diseño de Cebadores

Los cebadores se diseñaron a partir de fragmentos altamente conservados en secuencias consenso de los cebadores reportados para las regiones conservadas 3, 5, 7, 10 del ARNr 16S, previamente reportados por (Martínez-Porchas y col. 2017). Se incluyeron cebadores de diferentes tamaños, desde 13 hasta 20 bases de longitud, para las regiones que contienen los fragmentos V3-V4 y V7-V9 del ARNr 16S. Posteriormente, éstos se analizaron con el programa Primer analyzer, para elegir los que cumplieran con los criterios establecidos para el diseño de cebadores (Tabla 3) (Sharrocks y Shaw, 1991; Ayyadevara et al., 2000).

Tamaño	20-25 nucleótidos de longitud
Base en el extremo 3´	Debe ser G o C
Temperatura de fusión (Tm)	50-80°C
Contenido GC	40-60%
Auto-complementariedad	Debe ser evitada. Para evitar la formación de estructuras secundarias y dimerizaciones en el/los cebadores
Similitud	Debe tener un 100% de apareamiento con el molde

Los cebadores fueron nombrados de acuerdo a la terminología publicada por Alm y col, (1996). Cada nombre está formado por siete componentes que lo describen:

Ejemplo: ‘S-D-Bact-0338-a-A-18’

1. Gen objetivo (S= Small subunit RNAr)
2. Indicación del grupo taxonómico más grande (D= dominio)
3. Descripción abreviada de primeras tres o cinco letras del grupo taxonómico o filogenético de interés (Bact: Bacteria)

4. Un número de cuatro dígitos que indica la posición de la hebra. (Por ejemplo, '0338' representa la posición inicial 338 en el sistema de nomenclatura de *Escherichia coli*)
5. Una sola letra minúscula representa la versión de prueba del cebador (a= primera versión).
6. Una sola letra mayúscula indica si la secuencia de la sonda es idéntica a la hebra de sentido del ADN (S) o a la hebra anti-sentido (A).
7. El número indica el tamaño del cebador.

## 5.2 Evaluación *in silico* de los Cebadores

Se empleó la base de datos SILVA SSU Ref NR 99 (Versión 128) que contiene 645,151 secuencias de tamaños entre 900 y 1200 bases. Se evaluó la cobertura (porcentaje de secuencias bacterianas obtenidas) y especificidad (porcentaje de secuencias pertenecientes al filo bacteria) de los cebadores mediante las herramientas Test probe y Test prime de SILVA (con un nivel de astringencia de 0 mismatch), para evaluar a los cebadores en particular, así como la combinación de éstos (Quast et al., 2013).

Posteriormente, se determinó el tamaño del fragmento a amplificar (amplicon) por cada pareja de cebadores, así como la temperatura de fusión entre ellos, empleando la herramienta Primer Blast de la base de datos del Centro Nacional de Información Biotecnológica (NCBI, por sus siglas en inglés). Solo los pares de cebadores que presentaron una cobertura general y especificidad definidas, un tamaño de amplicon entre 400 y 550 pares de bases y temperaturas de fusión menores o iguales a 5°C entre ellos, fueron considerados como posibles candidatos para ser propuestos para análisis *in vitro*.

Finalmente, se seleccionaron los pares de cebadores que presentaron coberturas mayores a 80% y especificidades de 100% para el dominio bacteria, para posteriormente continuar con el análisis *in vitro*.

### 5.3 Evaluación *in vitro* de los Cebadores

#### 5.3.1 Selección y Colección de Muestras Biológicas

Se colectaron muestras de saliva, agua de mar (zona costera) y suelo agrícola, las cuales se colocaron en tubos estériles de 15 mL y se almacenaron a -20°C. Las muestras de agua de mar y saliva, se filtraron por vacío a través de membranas de celulosa de 0.45 µm. Posteriormente la membrana se dividió en dos y cada parte se colocó en tubos de lisis.

#### 5.3.2 Extracción y Cuantificación de ADN

Para la extracción de ADN de la muestra de suelo, se utilizó el kit de aislamiento de ADN FastDNA™ Spin kit for soil (MP BIO Laboratories, Solana Beech, CA, USA), utilizando 500 mg de muestra. Para las muestras de agua de mar y saliva se utilizó el kit de extracción de ADN/ARN taco™ (GeneReach Biotechnology Corporation), utilizando 200 µL de muestra. Las extracciones se realizaron siguiendo las indicaciones del fabricante.

Posteriormente, se determinó la concentración del ADN de cada muestra mediante espectrometría de fluorescencia, utilizando el kit dsDNA Quant-iT™ PicoGreen® (Life Technologies) y el lector de microplacas SpectraMax® (Molecular Devices). Una vez confirmada la presencia de ADN en las muestras, así como su cantidad y calidad, se procedió al proceso de amplificación.

### 5.3.3 Amplificación del ARNr16S (V3- V4 y V7-V9)

Se amplificaron los segmentos que comprenden las regiones V3-V4 y V7-V9 del ARNr 16S con los cebadores diseñados en la etapa anterior. Se preparó la reacción de PCR de las muestras de ADN obtenidas; para ello, se utilizaron 2.5 µL de muestra de ADN (5 ng/µL), 5 µL de cebador 16S fw, 5 µL del cebador 16S Rv y 12.5 µL del reactivo 2x KAPA HiFi HotStart Ready Mix (KAPA Biosystems, Woburn, MA, USA) para un volumen total de 25 µL. Las condiciones de PCR fueron las siguientes: 95°C por 3 minutos, 25 ciclos de 95°C por 30 segundos, 55°C por 30 segundos y 72°C por 30 segundos, 72°C por 5 minutos.

Los adaptadores utilizados para cada cebador se muestran en la Tabla 4.

Tabla 4. Secuencias de adaptadores propuestos en el protocolo Illumina	
Adaptador	
Forward (Fw)	5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG 3'
Reverse (Rv)	5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG 3'

Los adaptadores de los cebadores se utilizaron para indexar las muestras y poder identificarlas. Posteriormente, se realizó la preparación de la biblioteca de los productos de PCR obtenidos, según la guía de secuenciación del ARNr 16S publicada por Illumina para el sistema Illumina MiSeq.

### 5.3.4 Preparación de Biblioteca de Amplicones del ARNr16S

Según la guía de Illumina (Illumina, USA), los productos de PCR deben cumplir con ciertos parámetros relacionados con la pureza de los amplicones para que pueda prepararse una biblioteca y posteriormente llevar a cabo su secuenciación.

Primeramente, se llevó a cabo la purificación de amplicones de PCR mayores o iguales a 100 pares de bases, utilizando el reactivo AMPure Beads (Beckman Coulter, USA), el cual fue incubado previamente a temperatura ambiente (TA), y homogeneizado por 30 segundos utilizando un vortex. Posteriormente, se agregaron 20  $\mu\text{L}$  del reactivo en una microplaca con los productos de PCR, se agitó la mezcla en 10 veces y se incubó a temperatura ambiente por 5 minutos. Se colocó la microplaca con la mezcla de AMPure Beads y productos de PCR sobre la placa magnética para desechar el sobrenadante con una pipeta. Se realizó un lavado con etanol al 80% (dos veces) y se descartó el sobrenadante, dejando secar al aire, por 10 minutos. Se agregaron 52.2  $\mu\text{L}$  de Tris (tris (hydroxymethyl) aminomethane) 10 mM con pH 8.5 para la separación de los productos de PCR de las perlas magnéticas del reactivo AMPure Beads. Después, se colocó la microplaca en la placa magnética durante 2 minutos para que el sobrenadante se aclarara. Finalmente, se transfirieron 50  $\mu\text{L}$  del sobrenadante a una microplaca nueva.

Para la diferenciación de las muestras se realizó una PCR de indexado o etiquetado, que consiste en la colocación de cadenas de nucleótidos conocidas mediante enzimas, en los extremos de cada amplicon generado, permitiendo la agrupación en bibliotecas. Para ello, se utilizó la enzima del kit Nextera XT Index. Primeramente, se transfirieron 5  $\mu\text{L}$  de los productos de PCR purificados con el reactivo AMPure Beads a una microplaca nueva. Los 45  $\mu\text{L}$  restantes se almacenaron a  $-80\text{ }^{\circ}\text{C}$ . Posteriormente, se agregaron dos índices por cada muestra. Se tomaron 5  $\mu\text{L}$  de cada índice de manera horizontal (índices 2) y verticalmente (índices 1), para realizar una combinación diferente por cada muestra. Posteriormente, se llevó a cabo una reacción de PCR para que los índices pudieran unirse a los adaptadores que componen a la biblioteca. Se agregaron 25  $\mu\text{L}$  de 2x KAPA HiFi Hot Start Ready Mix y 10  $\mu\text{L}$  de agua para PCR, teniendo un volumen final de 50  $\mu\text{L}$  por cada muestra y se agitó la mezcla 10 veces. Las condiciones de PCR fueron las siguientes:  $95^{\circ}\text{C}$  por 3 minutos, 8 ciclos de  $95^{\circ}\text{C}$  por 30 segundos,  $55^{\circ}\text{C}$  por 30 segundos y  $72^{\circ}\text{C}$  por 30 segundos,  $72^{\circ}\text{C}$  por 5 minutos y se mantuvo a  $4^{\circ}\text{C}$ .

Para eliminar los contaminantes después de la segunda reacción de PCR, se llevó a cabo una segunda limpieza de productos de PCR. Se realizaron los mismos pasos empleados en la primera limpieza de productos de PCR, con excepción de que se agregaron 56  $\mu\text{L}$

del reactivo a cada muestra de productos de PCR con los índices; se lavó con 200  $\mu\text{L}$  de etanol al 80% y 27.5  $\mu\text{L}$  de TRIS 10 mM con pH 8.5 y se tomaron 25  $\mu\text{L}$  de sobrenadante para transferirlos a una microplaca nueva. Una vez limpios los productos de PCR, se realizó una medición de cantidad y concentración de ADN.

Se llevó a cabo una conversión de la concentración de las bibliotecas a nM. Para ello, la cantidad de ADN obtenida en ng/ $\mu\text{L}$  obtenida con el instrumento SpectraMax®, se utilizó la siguiente fórmula:

$$\frac{\text{Concentración en ng}/\mu\text{l}}{(660 \text{ g/mol} \times \text{tamaño promedio de la librería})} \times 10^6 = \text{concentración en nM}$$

Después de obtener la concentración de las muestras en nM, se realizó un ajuste de las bibliotecas.

### **5.3.5 Secuenciación de los Amplicones del ARNr 16S**

Las bibliotecas se ajustaron a una concentración 4 nM utilizando Tris 10 mM (pH 8.5) como diluyente. Posteriormente, de cada biblioteca se tomaron 5  $\mu\text{L}$  y se realizó una agrupación de bibliotecas (pool). Posteriormente, la biblioteca se desnaturalizó con 5  $\mu\text{L}$  de NaOH 0.2 N. Se utilizó una biblioteca PhiX 10 nM como control interno para verificar el adecuado funcionamiento del proceso de secuenciación. De igual forma, la biblioteca control PhiX se llevó a una concentración de 4 nM con Tris 10 mM (pH 8.5), se tomaron 5  $\mu\text{L}$  de PhiX diluido y se desnaturalizó con 5  $\mu\text{L}$  de NaOH 0.2 N, teniendo una concentración final de 2 nM. Se agregaron 990  $\mu\text{L}$  de buffer de hibridación para ambas muestras y se obtuvo una concentración final de 20 pM. Ambas bibliotecas se combinaron con 570  $\mu\text{L}$  de la biblioteca de ADN y 30  $\mu\text{L}$  de la biblioteca de PhiX.

Una vez combinadas las bibliotecas, se realizó la carga de muestras. Las muestras se cargaron en el cartucho con reactivo MiSeq V3 (Illumina). El cartucho se colocó en el instrumento de secuenciación MiSeq (Illumina), el cual tiene una celda de flujo con capacidad de 25 millones de lecturas y se llevó a cabo una secuenciación por ambos lados (602 ciclos: 2 x 301).

### 5.3.6 Análisis de Datos

Finalmente, se llevó a cabo una comparación entre los resultados obtenidos con los cebadores diseñados y los cebadores propuestos por Illumina, los cuales son ampliamente utilizados. Para llevar a cabo la comparación, en ambos casos se midieron parámetros de diversidad tales como alfa diversidad empleando el índice de Shannon-Wiener (Ecuación 1), beta diversidad (Ecuación 2), así como un análisis de componentes principales con el objetivo de establecer si el uso de un par u otro de cebadores arroja resultados distintos en cuanto al perfil bacteriano. Lo anterior, se llevó a cabo utilizando la plataforma MG-RAST: Metagenomics RAST Server (Meyer et al., 2008) y el programa STAMP: Statistical analysis of metagenomics profiles (Parks et al., 2014).

$$\text{Índice de Shannon- Wiener (Ecuación 1): } H' = \sum_{i=1}^S p_i \times \log_2 (p_i)$$

Donde,  $S$  es el número de especies (riqueza de especies),  $p_i$  representa la abundancia relativa de especies.

$$\text{Beta- diversidad } (\beta) \text{ (Ecuación 2): } \beta = (S_1 - C) + (S_2 - C)$$

Donde,  $S$  representa la riqueza de especies en dicha comunidad y  $C$  representa la riqueza de especies compartidas entre comunidades.

## 6 RESULTADOS Y DISCUSIÓN

### 6.1 Diseño y Evaluación *in silico* de los Cebadores

Para la amplificación del fragmento que contiene las regiones V3-V4, se diseñaron 70 cebadores sentido (forward) a partir de la región C3 y 67 cebadores anti-sentido (reverse) a partir de la región C5, del ARN 16S. Al evaluar las propiedades cada uno de ellos mediante el programa Primer Analyzer, solo 39 cebadores de tamaño entre 15 y 20 bases, cumplieron con los criterios establecidos. Los cebadores con tamaños de 13 y 14 bases no fueron seleccionados para el análisis *in silico*, debido a que no cumplieron completamente con los criterios, al presentar temperaturas de fusión menores a 50 °C o porcentajes de GC mayores del 60% (Tabla 5).

Por otro lado, para la amplificación del fragmento que contiene las regiones V7-V9, se diseñaron 98 cebadores sentido (forward) correspondientes a la región C7, y 31 cebadores anti-sentido (reverse) usando la región C10. De los 129 cebadores diseñados, solamente 22 cumplieron con los criterios requeridos (Tabla 6). En su mayoría cebadores de tamaño entre 17 y 20 bases; sin embargo, al contrario de los cebadores diseñados para la amplificación de las regiones variables 3-4, algunos cuyo tamaño osciló entre 13 y 15 bases, sí cumplieron con los criterios establecidos.

Para una adecuada optimización de la reacción de PCR, es importante considerar la utilización de cebadores de una longitud mínima que garantice temperaturas de fusión de al menos 54° C y un porcentaje de GC- de 40-60 %; esto provee condiciones para obtener una especificidad y eficiencia adecuadas (Dieffenbach et al., 1993; Chenna et al., 2003). El tamaño del cebador es proporcional a la eficiencia de la hibridación con el fragmento del ADN que se quiere amplificar. Si la secuencia del cebador es mayor a 24 b menor a 10 b puede no hibridar, o unirse a regiones que no son de interés generando productos inespecíficos, lo que se establece como una PCR ineficiente.

Por lo general, los cebadores utilizados para amplificar genes o fragmentos de genes considerados como biomarcadores taxonómicos, suelen tener un tamaño entre 16 y 20 pares de bases, longitud suficiente para lograr una adecuada especificidad, la cual debe ser sumada a una óptima temperatura de hibridación (Capel-Salinas y Yuste Lisboa, 2016).

Posteriormente, a partir del proceso de selección de cebadores idóneos se obtuvieron coberturas y especificidades variables. Por ejemplo, los cebadores diseñados a partir de las regiones C3 y C5, presentaron coberturas entre 80-100% y especificidades de entre 70 y 100%. Un total de 39 cebadores (28 cebadores sentido y 11 cebadores anti-sentido), presentaron coberturas mayores al 90% y fueron específicos para el dominio Bacteria (Tabla 5). Estos últimos cebadores fueron pre-seleccionados para continuar con el análisis *in silico*; los cuales contaban con un tamaño de 15 a 20 bases en su mayoría, en conformidad con lo reportado con Capel-Salinas y Yuste-Lisbona (2016), quienes sugieren la utilización de cebadores de 16 a 20 bases. Con excepción de 11 que eran de 19 a 20b y que se probaron porque presentaron porcentajes de cobertura y especificidad idóneos.

Una cobertura alta indica que un alto porcentaje o proporción de secuencias bacterianas obtenidas pueden hibridar con el cebador; mientras que una alta especificidad indica que tal cebador solamente reacciona con secuencias del grupo de interés, en este caso el dominio Bacteria (Wang, 2007). Entre más largo sea el cebador, menor será la probabilidad de repetirse en la hebra molde, aunque también disminuye la probabilidad de considerar aquellas secuencias que presentan una o más variaciones en posiciones específicas, incrementando así el nivel de astringencia (Morales y Holben, 2009). Una de las estrategias para resolver esta limitante, es el uso de degeneraciones o la reducción del tamaño de los cebadores. Una longitud de entre 15 y 16 bases, mostró ser adecuada, considerando sus altas coberturas y especificidades. Existen reportes de cebadores eficientes y ampliamente utilizados en análisis de diversidad bacteriana, cuya longitud se posiciona en estos tamaños (Lane et al., 1985; Muyzer et al., 1993; Hansen et al., 1998 y Claesson y O'Toole, 2010).

En cuanto a los 22 cebadores diseñados a partir de las regiones C7 y C10, éstos presentaron coberturas (48-96%) y especificidades (18 hasta 100%) con mayor variabilidad (Tabla 6). Para los cebadores sentido, se obtuvieron coberturas mayores al 80% y especificidades mayores de 90%, excluyendo un cebador de 18 bases que presentó una especificidad de 18%, el cual no fue considerado para continuar con el análisis *in silico*. Los cebadores anti-sentido presentaron coberturas generales bajas entre 45 y 69%, pero especificidades elevadas ( $\geq 90\%$ ). Los 22 cebadores (13 cebadores sentido y 9 cebadores anti-sentido), fueron específicos para el dominio bacteria; sin embargo, registraron coberturas bajas en comparación con los cebadores diseñados de las regiones conservadas 3 y 5. Esto probablemente se deba al reducido número de secuencias registradas en la base de datos SILVA (versión SSU 128 NR Ref 99) que contienen el fragmento de las regiones variables 7-9; en este caso, se evaluaron *in silico* todos los cebadores.

Aunque la información contenida en las bases de datos especializadas en el ARNr 16S ha incrementado casi exponencialmente gracias al uso de tecnologías de secuenciación masiva, es también cierto que el número de secuencias incompletas ha ido en aumento, debido a la limitada capacidad de los equipos actuales para producir secuencias completas (Rodríguez-Armengol, 2012).

Una vez terminada esta etapa se procedió a la formación y evaluación *in silico* de pares de cebadores con todas las secuencias que presentaron mayores porcentajes de cobertura y especificidad para cada región conservada.

Tabla 5. Propiedades y evaluación *in silico* de los cebadores sentido y anti-sentido diseñados a partir de las regiones conservadas 3 y 5 del ARNr 16S.

Cebadores Sentido	Bases	Tm	%G-C	Cobertura
CTGAGAYACGGYCCARACTC	20	63.5	57.5	92%
TGAGAYACGGYCCARACTCC	20	66.1	57.5	92%
GAYACGGYCCARACTCCTAC	20	62.1	57.5	92%
AYACGGYCCARACTCCTACG	20	64.8	57.5	93%
YACGGYCCARACTCCTACG	19	64.5	60.5	92%
GAGAYACGGYCCARACTCC	19	63.3	60.5	93%
ACTGAGAYACGGYCCARAC	19	61.4	55.3	93%
AYACGGYCCARACTCCTAC	19	60.2	55.3	92%
TGAGAYACGGYCCARACTC	19	62.6	55.3	93%
YACGGYCCARACTCCTAC	18	59.6	58.3	93%
CTGAGAYACGGYCCARAC	18	60.5	58.3	93%
AGAYACGGYCCARACTCC	18	61.4	58.3	92%
ACGGYCCARACTCCTACG	18	63.2	60.1	93%
TGAGAYACGGYCCARAC	17	59.4	55.9	93%
AGAYACGGYCCARACTC	17	57.2	55.9	92%
ACGGYCCARACTCCTAC	17	58	58.8	94%
GAGAYACGGYCCARAC	16	55.8	59.4	93%
GAYACGGYCCARACTC	16	55.8	59.4	93%
AYACGGYCCARACTCC	16	58.1	59.4	94%
CARACTCCTACGGRNG	16	55.4	59.4	96%
ARACTCCTACGGRNGC	16	56.7	59.4	88%
ACTGAGAYACGGYCC	15	53.2	60	96%
AGAYACGGYCCARAC	15	53.2	56.7	93%
AYACGGYCCARACTC	15	53.2	56.7	94%
GGYCCARACTCCTAC	15	50.5	60	94%
GYCCARACTCCTACG	15	52	60	95%
YCCARACTCCTACGG	15	53.8	60	95%
ARACTCCTACGGRNG	15	51.7	56.7	97%
Cebadores Reversos				
GAWACCCNNGTAGTCCACGC	80	65.4	60	86%
AGAWACCCNNGTAGTCCACG	20	62.8	55	86%
TAGAWACCCNNGTAGTCCAC	20	58	50	87%
GAWACCCNNGTAGTCCACG	19	61.9	57.9	86%
AGAWACCCNNGTAGTCCAC	19	57.9	52.6	88%
AACRGGATTAGAWACCC	17	52.8	44.1	88%
WACCCNNGTAGTCCACG	17	59	58.8	86%
AWACCCNNGTAGTCCAC	17	54.4	52.9	88%
ACRGGATTAGAWACCC	16	50.7	46.9	88%
AACRGGATTAGAWACC	16	47.9	40.6	89%
CRGGATTAGAWACCC	15	49.2	50	90%

En ningún caso se determinó la formación de dímeros, Especificidades de 100%

Tabla 6. Propiedades y evaluación *in silico* de los cebadores sentido y anti-sentido diseñados a partir de las regiones conservadas 7 y 10 del ARNr 16S

Sentido	Bases	Tm	%G-C	Cobertura	Especificidad
ATGGYYGYCGTCAGCTCGTG	20	71.8	62.5	94%	100%
TGCATGGYYGYCGTCAGCTC	20	73	62.5	94%	93%
TGNTGCATGGYYGYCGTCAG	20	72.7	60	94%	97%
NTGCATGGYYGYCGTCAG	18	67.9	61.1	95%	89%
TGNTGCATGGYYGYCGTC	18	69.6	61.1	94%	97%
AGGTGNTGCATGGYYGYC	18	67.3	61.1	94%	95%
TCAGCTCGTGYCGTGAG	17	63.3	61.8	87%	100%
ATGGYYGYCGTCAGCTC	17	63.9	61.8	95%	93%
TGCATGGYYGYCGTCAG	17	67.3	61.8	95%	89%
TGNTGCATGGYYGYC	15	61.5	60	95%	94%
ATGGYYGYCGTCAG	14	55.5	60.7	95%	89%
GTGNTGCATGGYYG	14	55.7	60.7	95%	18%
NTGCATGGYYGYC	13	53.8	61.5	96%	82%
Anti-sentido					
GAAGTCRTAACAAGGTANCC	20	55.8	45	49%	97%
TGAAGTCRTAACAAGGTANC	20	54.8	40	46%	100%
AAGTCRTAACAAGGTANCC	19	53.7	42.1	66%	90%
GAAGTCRTAACAAGGTANC	19	51.6	42.1	51%	96%
GGGTGAAGTCRTAACAAGG	19	58.8	50	47%	100%
AGTCRTAACAAGGTANCC	18	51.8	44.4	68%	90%
GGTGAAGTCRTAACAAGG	18	54.6	47.2	48%	96%
GGGTGAAGTCRTAACAAG	18	54.6	47.2	48%	100%
GTCRTAACAAGGTANCC	17	50.2	47.1	69%	90%

En ningún caso se determinó la formación de dímeros.

## 6.2 Evaluación *in silico* de los Pares de Cebadores

En base a las secuencias conservadas de la región C3 y C5 se diseñaron 28 cebadores Fw y 11 Rv, respectivamente. Con ellos fue posible hacer 392 combinaciones, de las cuales solamente 22 pares presentaron coberturas mayores al 80% y fueron específicas para el dominio bacteria (Tabla 7), por lo que fueron seleccionados para análisis posteriores. Estos cebadores seleccionados contaban con longitudes de entre 15 y 20 bases; a este respecto, se han reportado anteriormente cebadores eficientes de entre 15 y 19 bases (Capel-Salinas y Yuste-Lisbona, 2016), los cuales han sido ampliamente utilizados en numerosos estudios, posiblemente un tamaño menor que el promedio sugerido podría contribuir mejorar el nivel de las coberturas (Jaric, 2013). Los pares de cebadores seleccionados presentaron un diferencial de Tm alrededor de los 5°C y generaron amplicones de tamaños entre 460-472 pb; estos son resultados prometedores para su potencial uso *in vitro*, ya que además de demostrar que se está amplificando el fragmento esperado, el tamaño resulta compatible con la capacidad de la celda de flujo, el kit de secuenciación y del sistema Miseq (600 pb como máximo) (Rodríguez-Armengol, 2012).

Para los cebadores diseñados a partir de las regiones conservadas C7 y C10, se obtuvieron 117 combinaciones, de las cuales 25 presentaron coberturas mayores al 60% y fueron específicas para el dominio bacteria. Estas bajas coberturas posiblemente se deben a la proporción considerable de secuencias incompletas que hay en la base de datos SILVA versión 128 (Martínez-Porchas et al., 2017). El tamaño de los pares de cebadores que fueron los mejores candidatos, osciló en el rango de 15 a 20 bases, presentaron un Tm menor o igual a 5° C, así como, tamaños de amplicones entre 455 y 465 pb. Un total de 14 pares de cebadores fueron seleccionados para el análisis *in vitro* (Tabla 8)

Tabla 7. Evaluación *in silico* de las combinaciones de cebadores diseñados a partir de las regiones conservadas 3 y 5 del ARNr 16S.  
Cebadores sentido: A) ARACTCCTACGGRNG; B) CARACTCCTACGGRNG

	Anti-sentido (Reverse)	Cobertura	Amplicon (pb)	$\Delta T_m$ cebadores
A	GAWACCCNNGTAGTCCACGC	80%	472	>5
B		80%	473	>5
A	AACRGGATTAGAWACC	80%*	462	<5
B		80%*	463	5
A	GAWACCCNNGTAGTCAACGC	83%	471	>5
B		83%	472	>5
A	TAGAWACCCNNGTAGTCCAC	83%	470	>5
B		83%	470	>5
A	GAWACCCNNGTAGTCCACG	80%	468	>5
B		80%	471	>5
A	AGAWACCCNNGTAGTCCAC	80%	471	<5
B		84%	472	5
A	WACCCNNGTAGTCCACG	80%	469	<5
B		84%	470	>5
A	AWACCCNNGTAGTCCAC	83%	465	>5
B		83%	466	5
A	ARGGATTAGAWACCC	80%*	462	<5
B		80%*	463	<5
A	ACRGGATTAGAWACCC	80%	462	<5
B		83%	472	5
A	AACRGGATTAGAWACCC	80%	462	<5
B		83%	471	5

\*Cebadores seleccionados para el análisis *in vitro*, Especificidades del 100%.

Tabla 8. Evaluación <i>in silico</i> de las combinaciones de cebadores diseñados a partir de las regiones conservadas 7 y 10 del ARNr 16S.				
Sentido (Forward)	Anti-sentido (Reverse)	Cobertura	Amplicón (pb)	$\Delta T_m$ cebadores
atggyygycgtcag	agtcrtacaaggtancc	65%	461	<5
	gtcrtacaaggtancc	66%	461	<5
gtgntgcatggyyg	agtcrtacaaggtancc	65%	461	<5
	gtcrtacaaggtancc	66%	461	<5
tgntgcatggytgyc	agtcrtacaaggtancc	65%	462	<5
	gtcrtacaaggtancc	65%*	462	<5
ntgcatggyygyc	agtcrtacaaggtancc	65%	460	<5
	gtcrtacaaggtancc	66%	460	<5
tcagctcgtgycgtgag	agtcrtacaaggtancc	59%	456	5
	gtcrtacaaggtancc	60%	456	5
atggytgycgtcagctc	agtcrtacaaggtancc	65%	456	5
	gtcrtacaaggtancc	64%*	456	5
tgcattggyygycgtcag	agtcrtacaaggtancc	65%	456	5
	gtcrtacaaggtancc	66%	456	5

\*Cebadores seleccionados para el análisis *in vitro*, Especificidades entre 94-100%.

Los pares de cebadores que obtuvieron las coberturas y especificidades más altas fueron seleccionados para continuar con el análisis *in silico*. Se eligieron entonces seis pares de cebadores, de los cuales cuatro fueron para la amplificación de las regiones variables 3 y 4, que obtuvieron una cobertura máxima de 80% y especificidades de 100%, y dos para las regiones variables 7-9 que obtuvieron coberturas mayores a 60% y especificidades de 100%. A continuación, en la tabla 9 se detallan las secuencias de cada par de cebadores.

Tabla 9. Cebadores para la amplificación de las regiones variables 3-4 y 7-9.		
	Nombre	Secuencia (5'-3')
1	S-D-Bact-0329-a-S-15	ARACTCCTACGGRNG
	S-D-Bact-0776-a-A-15	GGGTWTCTAATCCYG
2	S-D-Bact-0329-a-S-15	ARACTCCTACGGRNG
	S-D-Bact-0774-a-A-16	GGTWTCTAATCCYGTT
3	S-D-Bact-0328-b-S-16	CARACTCCTACGGRNG
	S-D-Bact-0776-a-A-15	GGGTWTCTAATCCYG
4	S-D-Bact-0328-b-S-16	CARACTCCTACGGRNG
	S-D-Bact-0775-a-A-16	GGGTWTCTAATCCYGT
5	S-D-Bact-1048-a-S-15	TGNTGCATGGYYGYC
	S-D-Bact-1493-a-A-17	GGNTACCTTGTTAYGAC
6	S-D-Bact-1054-a-S-17	ATGGYYGYCGTCAGCTC
	S-D-Bact-1493-a-A-17	GGNTACCTTGTTAYGAC

Se compararon los porcentajes de cobertura y especificidad de los pares de cebadores candidatos para la amplificación de las regiones variables 3 y 4, con los cebadores propuestos por Klindworth y col. Todos los pares de cebadores presentaron coberturas mayores o iguales a 80% y especificidades de 100%. Si bien el porcentaje de cobertura de los cebadores propuestos por Klindworth y col. es del 86% (5), se considera que al menos numéricamente no existe una diferencia significativa con respecto a las combinaciones de cebadores propuestas. Estos resultados son favorables, debido a que indican una alta capacidad de detección de secuencias por parte de los nuevos cebadores y similares a los cebadores reportados por Klindworth y col. 2013. Los cebadores seleccionados detectaron en promedio 459,796 de las 575,268 secuencias registradas para las regiones variables 3 y 4 contenidas en la base de datos SILVA. En la figura 6, se observan las secuencias detectadas y no detectadas por cada par de cebadores.

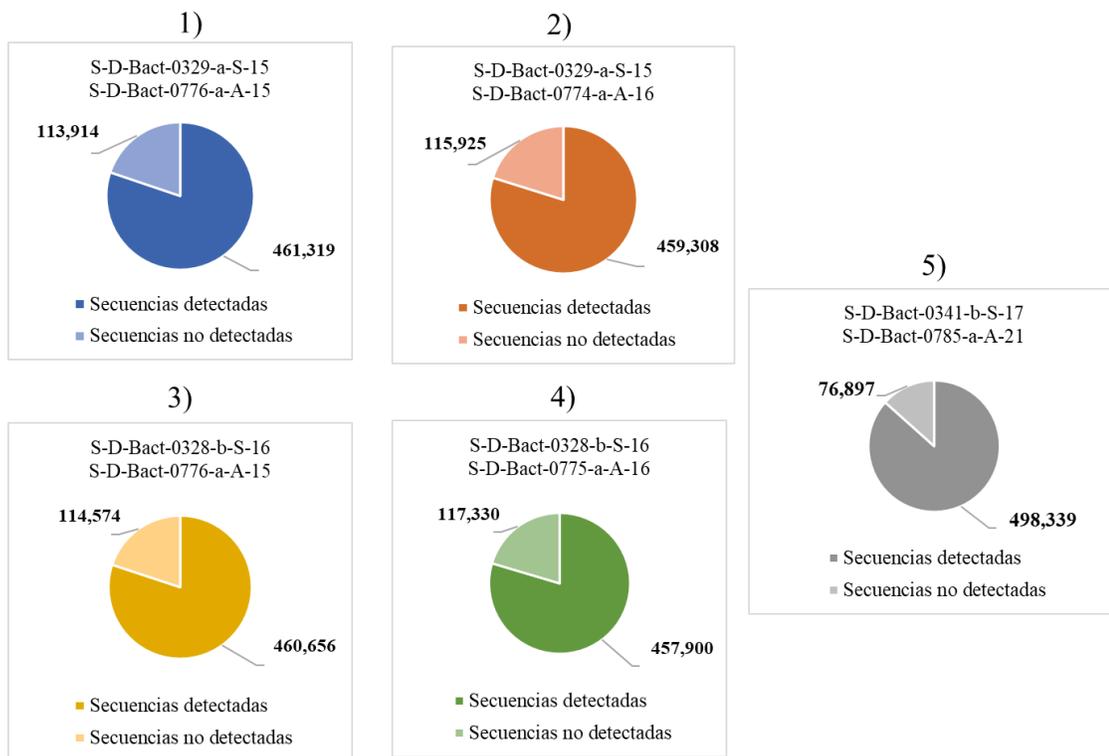


Figura 6. Secuencias detectadas y no detectadas por los cebadores diseñados a partir de fragmentos con alto grado de conservación (1, 2, 3, 4) y los cebadores propuestos por Klindworth y col, 2013 (5) para la amplificación de las regiones variables 3 y 4 del ARNr 16S.

Se obtuvo una alta detección de secuencias por parte de los cebadores propuestos y la cantidad de estas son similares entre ellos. Los cebadores reportados por Klindworth y col. detectaron en promedio 30,000 secuencias más. Pese a esto, se confirmó que los cebadores diseñados a partir de fragmentos con alto grado de conservación, poseen una cobertura de secuencias elevada y muy similar a la reportada para los cebadores reportados por Klindworth y col. Sin embargo, esto no representa información suficiente para concluir si poseen una mayor cobertura taxonómica, es decir, una mayor cantidad de grupos detectados. Por ejemplo, en una base de datos, un solo filo puede tener hasta medio millón de secuencias registradas (ej *Proteobacteria*), mientras que otro puede tener solo una secuencia registrada (ej. *Thermosulfidibacteraeota*).

Por lo tanto, un par de cebadores pudiera presentar una elevada cobertura de secuencias, pero si estas pertenecen a solo unos cuantos fillos que posean gran cantidad de secuencias registradas, en realidad no se estaría demostrando una buena cobertura taxonómica. Debido a esto, se llevó a cabo un análisis del perfil taxonómico bacteriano obtenido por cada par de cebadores y se encontró que, para los cuatro pares de cebadores propuestos, se detectaron entre 60 y 65 fillos (Tabla 10), mientras que para los reportados por Klindworth y col. se detectaron un total de 74 fillos. Los cuatro pares de cebadores propuestos presentaron coberturas de secuencias del 100% para los fillos *Chrysiogenetes*, *Dictyoglomi*, *Firestonebacteria*, *GBS-1*, *WS4*. Sin embargo, como se mencionó anteriormente, la base de datos SILVA contiene un número de registro reducido de algunos fillos (entre 3 y 10 secuencias por filo).

En cuanto a los fillos con un número mayor de secuencias registradas en la base de datos como: *Actinobacteria*, *Bacteroidetes*, *Proteobacteria*, *Acidobacteria*, *Firmicutes*, entre otros, los cebadores propuestos cubrieron entre 80 y 95% de secuencias. No obstante, para el filo *Cyanobacteria* solo se cubrió el 1% (176 de 13,468) de las secuencias registradas, un porcentaje bajo comparado con los cebadores propuestos por Klindworth y col. Con los cuales se detectó 76%. Estos fillos de bacterias poseen una mayor cantidad de secuencias, debido a que son más abundantes en la naturaleza, pero también han sido ampliamente

estudiados y reportados en estudios metagenómicos de microbiota intestinal, cavidad bucal, suelos y agua (Hugenholtz et al., 1998).

Tabla 10. Porcentajes de cobertura de los filos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V3-V4 del ARNr 16S.

<i>Chrysiogenetes</i>	100%	<i>Spirochaetes</i>	76%
<i>Dictyoglomi</i>	100%	<i>Kiritimatiellaeota</i>	73%
<i>Firestonebacteria</i>	100%	<i>WS1</i>	72%
<i>GBS-1</i>	100%	<i>Nitrospinae</i>	68%
<i>WS4</i>	100%	<i>Marinimicrobia</i>	61%
<i>GAL15</i>	98%	<i>Elusimicrobia</i>	60%
<i>Schekmanbacteria</i>	98%	<i>Desantisbacteria</i>	50%
<i>Epsilonbacteraeota</i>	95%	<i>Armatimonadetes</i>	49%
<i>Nitrospirae</i>	94%	<i>Omnitrophicaeota</i>	47%
<i>BHI80-139</i>	93%	<i>Tenericutes</i>	46%
<i>Acidobacteria</i>	91%	<i>Rokubacteria</i>	28%
<i>Margulisbacteria</i>	91%	<i>Chloroflexi</i>	26%
<i>Zixibacteria</i>	91%	<i>WOR-1</i>	26%
<i>CK-2C2-2</i>	90%	<i>Planctomycetes</i>	22%
<i>Entotheonellaeota</i>	90%	<i>Patescibacteria</i>	17%
<i>Gemmatimonadetes</i>	90%	<i>TA06</i>	15%
<i>WPS-2</i>	90%	<i>Coprothermobacteraeota</i>	7%
<i>Calditrichaeota</i>	88%	<i>Thermotogae</i>	5%
<i>Dependentiae</i>	88%	<i>AncK6</i>	5%
<i>FBP</i>	88%	<i>Aerophobetes</i>	5%
<i>Firmicutes</i>	88%	<i>Hydrogenedentes</i>	3%
<i>Modulibacteria</i>	88%	<i>GN01</i>	3%
<i>Latescibacteria</i>	87%	<i>Verrucomicrobia</i>	2%
<i>Proteobacteria</i>	87%	<i>Cyanobacteria</i>	1%
<i>FCPU426</i>	86%	<i>Fusobacteria</i>	1%
<i>Poribacteria</i>	86%	<i>PAUC34f</i>	1%
<i>Deferribacteres</i>	85%	<i>Deinococcus-Thermus</i>	0%
<i>Atribacteria</i>	84%	<i>Acetothermia</i>	0%
<i>WS2</i>	84%	<i>Aegiribacteria</i>	0%
<i>Actinobacteria</i>	83%	<i>Aquificae</i>	0%
<i>Cloacimonetes</i>	83%	<i>Calescamantes</i>	0%
<i>LCP-89</i>	83%	<i>Dadabacteria</i>	0%
<i>Chlamydiae</i>	82%	<i>Edwardsbacteria</i>	0%
<i>Fibrobacteres</i>	82%	<i>Fervidibacteria</i>	0%
<i>Halanaerobiaeota</i>	82%	<i>Hydrothermae</i>	0%
<i>Bacteroidetes</i>	81%	<i>Lindowbacteria</i>	0%
<i>Lentisphaerae</i>	81%	<i>MAT-CR-M4-B07</i>	0%
<i>BRC1</i>	79%	<i>RsaHF231</i>	0%
<i>Synergistetes</i>	78%	<i>Thermosulfidibacteraeota</i>	0%
<i>Caldiserica</i>	76%	uncultured	0%

Aproximadamente 19 filos fueron cubiertos en porcentajes menores del 80% tanto por los cebadores propuestos y como los reportados por Klindworth y col. Además, 10 filos fueron detectados en porcentajes menores de 10% por los cebadores propuestos; mientras que un total de 14 filos no pudieron ser detectados, en contraste con los cebadores propuestos por Klindworth y col., los cuales detectaron solamente 5 filos con una cobertura por debajo del 10% y otros 4 que no fueron detectados por el par de cebadores.

En general un total de 64 filos tuvieron en común todos los cebadores, entre ellos filos como: *Actinobacteria*, *Bacteroidetes*, *Proteobacteria*, *Acidobacteria*, *Firmicutes*, *Cyanobacteria*, *Spirochaetes*, *Kiritimatiellaeota*, *Marinimicrobia*, *Armatimonadetes*, *Omnitrophicaeota*, *Tenericutes*, *Chloroflexi*, *Planctomycetes*, *Patescibacteria* y *Nitrospinae* (Figura 7); lo que sugiere la posible obtención de perfiles taxonómicos similares para una misma comunidad bacteriana. Por último, solamente 5 filos no pudieron ser detectados por algún par de cebadores: *AncK6*, *Aerophobetes*, *GN01*, *Aegiribacteria* y *Calescamantes* debido a que obtuvieron porcentajes de cero en el análisis. Solamente se registró un filo que los cebadores propuestos por Klindworth y col. no pudieron detectar, pero los cebadores propuestos sí; este es el filo *Zixibacteria*.

Estos filos no detectados resaltan debido a su mínima cantidad de secuencias pertenecientes a las regiones variables 3 y 4 registradas en la base de datos SILVA. Los cebadores no pueden llegar a ser universales, por lo que su diseño contempla como característica principal, la obtención de la mayor proporción de secuencias contenidas en las bases de datos, dejando fuera casos como estos.

Una posible explicación de la poca cantidad de secuencias registradas de los filos mencionados anteriormente, pudiera ser a que representan grupos de bacterias de las cuales no se posee aún suficiente información debido a que algunas han sido descubiertas recientemente (ej. *Aerophobetes*), o que habitan ambientes hostiles (ej. *Kiritimatiellaota*, *Marinimirobia*), lo que dificulta su estudio (Spring et al., 2016; Bertagnolli et al, 2017). Se ha reportado que las secuencias del ARNr 16S pertenecientes a bacterias que habitan ambientes hostiles, o que poseen alguna ventaja metabólica, poseen variaciones que las hacen ser resistentes a estos ambientes y son esas mismas variaciones las que ocasionan discordancias entre los cebadores utilizados, imposibilitando la correcta secuenciación del genoma de estas bacterias e incluso de fragmentos de interés (Bertagnolli et al, 2017). No obstante, hasta este punto la eficiencia obtenida a partir del uso *in silico* de los cuatro pares de cebadores propuestos para la detección de taxones bacterianos, parece ser inferior a la observada con los cebadores de Klindworth y col. 2013.

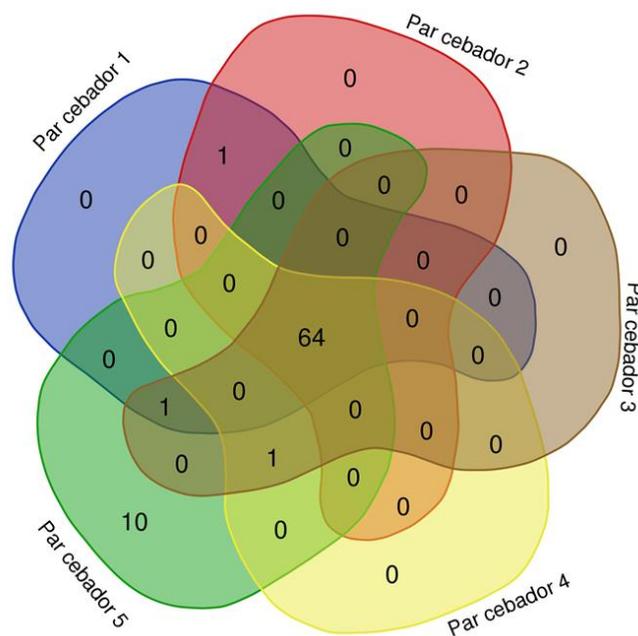


Figura 7. Filos detectados por cada uno de los cebadores diseñados a partir de fragmentos con alto grado de conservación (1, 2, 3 y 4) respecto a los cebadores propuestos por Klindworth y col, 2013 (5).

Por otro lado, también se compararon los porcentajes de cobertura y especificidad de los pares de cebadores candidatos para la amplificación de las regiones variables 7-9. Los dos pares de cebadores presentaron coberturas similares alrededor del 60% y especificidades del 100%. Para el primer par de cebadores, la cobertura general fue de 65%, mientras que para el segundo par fue de 64%. Si bien los porcentajes de cobertura de estos cebadores propuestos son muy bajos comparados con los cebadores que amplifican las regiones variables 3 y 4 presentados anteriormente, esto es de esperarse, debido a la menor cantidad de secuencias registradas en la base de datos SILVA para el segmento que comprende las regiones variables 7-9. Se detectaron en promedio 159,900 de 240,889 secuencias registradas para las regiones variables 7 y 9 contenidas en la base de datos SILVA. En la figura 8, se observan las secuencias detectadas y no detectadas por par de cebadores.

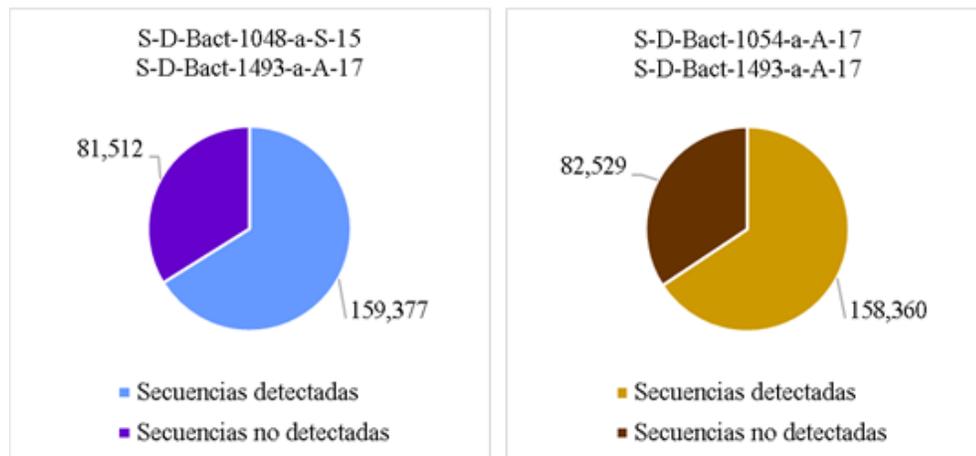


Figura 8. Secuencias detectadas y no detectadas por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones variables 7-9 del ARNr 16S.

La cantidad de secuencias detectadas por ambos pares son similares (Figura 10). Sin embargo, el par de cebadores S-D-Bact-1048-a-S-15| S-D-Bact-1493-a-A-17 (1) detectó alrededor de 1,000 secuencias más que el segundo par. Aunado a esto, la cobertura de secuencias bacterianas por parte de ambos pares es aceptable en comparación con cebadores utilizados para la amplificación de las regiones más extremas del ARNr 16S, utilizados por la base de datos SILVA (S-D-Bact-0008-a-S-16|S-D-Bact-1492-a-A-16) que

poseen coberturas de secuencias bacterianas de 77% (Klindworth y col. 2013). Cabe destacar que los nuevos cebadores solo se dirigen a la amplificación de tres zonas variables (7–9 de aprox. de 445 pb); en cambio, el par de cebadores que emplea la base de datos SILVA amplifica la totalidad del gen (aprox. 1,500 pb). Esto debido a que las regiones V7-V9 son menos estudiadas, por lo que existe una menor cantidad de fragmentos pertenecientes a estas regiones en las bases de datos (Quast et al., 2013).

Se encontró que para los dos pares de cebadores nuevos se detectaron 76 y 77 fillos (Tabla 11) y se obtuvieron un total de 76 fillos en común (Figura 9). Estos presentaron además coberturas de secuencias del 100% para 17 fillos. El par de cebadores S-D-Bact-1048-a-S-15|S-D-Bact-1493-a-A-17 detectó 9 fillos, uno más que el par Bact-1054-a-S-17|S-D-Bact-1493-a-A-17, el fillo *Desantisbacteria*. Sin embargo, como se mencionó anteriormente, los fillos detectados en un 100% cuentan con una baja cantidad de secuencias para las regiones variables 7 - 9 en la base de datos SILVA (entre 1 y 34 secuencias por fillo). Un total de 9 fillos fueron detectados entre 80 y 90%, entre ellos fillos como *Rokubacteria*, *Marinimicrobia*, *Halanaerobiaeota*, que son fillos que representan microorganismos no cultivables, debido que todavía se desconocen sus características metabólicas, cuya información permitiría el potencial desarrollo de medios de cultivo (Becraft et al., 2017).

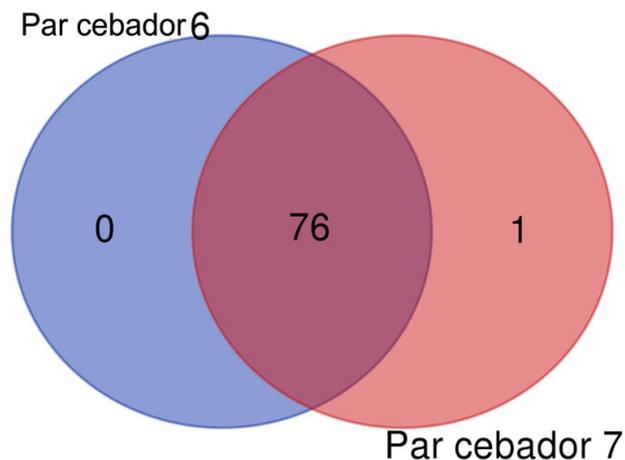


Figura 9. Fillos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V7-V9. 6) S-D-Bact-1048-a-S-15|S-D-Bact-1493-a-A-17; 7) S-D-Bact-1054-a-S-17|S-D-Bact-1493-a-A-17.

Además, aproximadamente 56 filos fueron cubiertos en porcentajes menores del 80% por ambos pares de cebadores. Incluyendo filos con una gran cantidad de secuencias registradas y de amplio estudio como: *Actinobacteria*, *Bacteroidetes*, *Proteobacteria*, *Acidobacteria*, *Firmicutes*, *Cyanobacteria*, entre otros. Estos filos representan entre el 20 y 60% de la población bacteriana total en ambientes como suelo participando en procesos como la descomposición de materia orgánica y el ciclo del carbono (Usha et al., 2011). Solo cinco filos presentaron porcentajes menores a 10%. El par de cebadores S-D-Bact-1048-a-S-15| S-D-Bact-1493-a-A-17 no pudo detectar cuatro filos: *Calescamantes*, *Fervidibacteria*, *MAT-CR-M4-B07* y *Zixibacteria*. En cambio, el par de cebadores S-D-Bact-1054-a-S-17| S-D-Bact-1493-a-A-17, si pudo detectar todas las secuencias registradas (100%) del filo *Calescamantes*.

Tabla 11. Porcentajes de cobertura de los filos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V7-V9 del ARNr 16S

AncK6	100%	WS1	68%
Calescamantes	100%	FCPU426	67%
Desantisbacteria	100%	Planctomycetes	67%
Dictyoglomi	100%	Bacteroidetes	66%
Firestonebacteria	100%	Cloacimonetes	66%
GBS-1	100%	Verrucomicrobia	66%
Lindowbacteria	100%	Actinobacteria	64%
Thermosulfidibacteraeota	100%	Firmicutes	64%
uncultured	100%	Latescibacteria	64%
WOR-1	100%	Acidobacteria	62%
WS4	90%	Fibrobacteres	62%
Rokubacteria	84%	Aegiribacteria	61%
Marinimicrobia	83%	Lentisphaerae	60%
Chlamydiae	82%	Omnitrophicaeota	60%
Deferribacteres	82%	Fusobacteria	58%

Entotheonellaeota	82%	BHI80-139	57%
FBP	80%	Calditrichaeota	57%
Halanaerobiaeota	80%	Dependentiae	57%
Synergistetes	80%	Hydrogenedentes	56%
Margulisbacteria	79%	Chloroflexi	54%
Tenericutes	78%	Patescibacteria	51%
WPS-2	78%	Edwardsbacteria	50%
Gemmatimonadetes	77%	Acetothermia	49%
WS2	76%	TA06	49%
CK-2C2-2	75%	Spirochaetes	48%
Cyanobacteria	75%	Caldiserica	46%
GAL15	75%	BRC1	43%
Armatimonadetes	74%	Chrysiogenetes	40%
Elusimicrobia	74%	RsaHF231	40%
Aquificae	73%	Coprothermobacteraeota	38%
LCP-89	73%	Schekmanbacteria	37%
Nitrospinae	71%	Deinococcus-Thermus	34%
PAUC34f	71%	Modulibacteria	33%
Kiritimatiellaeota	70%	Aerophobetes	32%
Proteobacteria	70%	GN01	28%
Nitrospirae	69%	Hydrothermae	14%
Atribacteria	68%	Epsilonbacteraeota	3%
Dadabacteria	68%	Thermotogae	3%
Poribacteria	68%	Fervidibacteria	0%
		MAT-CR-M4-B07	0%
		Zixibacteria	0%

Los cebadores empleados para la amplificación de las regiones V3-V4, detectaron un mayor número de secuencias, en su mayoría secuencias pertenecientes a filos que han sido ampliamente reportados en diferentes estudios de diversidad bacteriana. En cambio, los cebadores que amplifican las regiones V7-V9 detectaron mayor número de filos. Por ejemplo, los filos *AncK6*, *Calescamantes*, *Desantisbacteria*, *Dictyoglomi*, *Firestonebacteria*, *GBS-1*, *Lindowbacteria*, *Thermosulfidibacteraeota*, *WOR-1*, solo pudieron ser detectados por los cebadores diseñados para estudiar las regiones V7-V9, los cuales, además, que presentaron coberturas del 100% para estos filos, los cuales en su mayoría se consideran como bacterias no cultivables. Solo hubo dos filos que no pudieron ser cubiertos por alguno de los cebadores: *Fervidibacteria* y *MAT-CR-M4-B07*, debido

quizá, a que son bacterias no cultivables descubiertas recientemente; por lo que aún no se cuenta con suficiente información en las bases de datos (Becraft et al., 2017). Considerando las diferentes especificidades entre los cebadores, a veces altamente específicos para determinados grupos de bacterias, algunos autores recomiendan el uso combinado de estos (Tingre y Hugenholtz, 2008; Hong et al., 2009; Wang et al., 2007).

En general, todos los cebadores diseñados a partir de fragmentos con alto grado de conservación dentro de las regiones conservadas el ARNr 16S presentaron buenas coberturas, especificidades y una detección considerable de filos (Figura 10); similares a los obtenidos por los cebadores propuestos por Klindworth y col. Por lo que, en base al análisis *in silico* pudiera concluirse parcialmente que estos cebadores resultaron fiables para utilizarse en análisis de diversidad bacteriana, aunque no resultaron ser mejores a los recomendados por Klindworth y col. 2013.

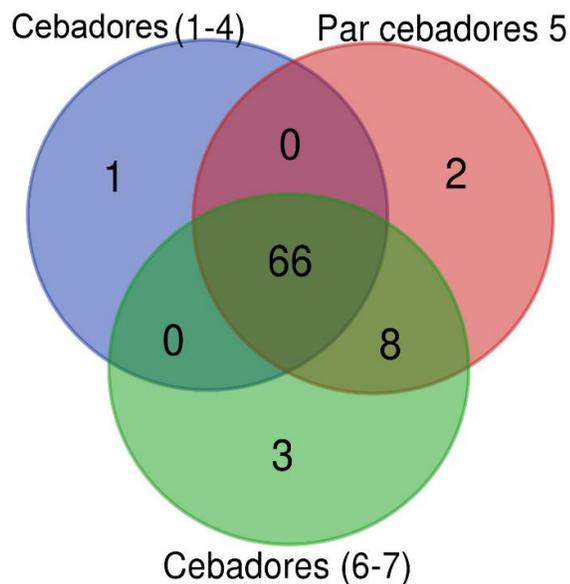


Figura 10. Filos detectados por los cebadores diseñados a partir de fragmentos con alto grado de conservación para la amplificación de las regiones V3-V4 (1-4) y V7-V9 (6 y 7) respecto a los propuestos por Klindworth y col. 2013 (5).

La siguiente figura muestra donde se encuentran posicionados los cebadores seleccionados para la amplificación de las regiones V3-V4 y V7-V9 y los cebadores propuestos por Klindworth y colaboradores (color rojo), en el ARNr 16S. Se observa que los cebadores seleccionados para las regiones V3-V4, están posicionados muy cerca de los cebadores de Klindworth y col. por lo que, pudiera explicarse los porcentajes de cobertura tan similares entre ellos y la cantidad de filos detectados en común. En cuanto a los cebadores seleccionados para la amplificación de las regiones V7-V9, se observa que estos están muy próximos entre ellos en la secuencia del ARNr 16S, por lo que de igual modo obtuvieron coberturas similares y la misma diversidad bacteriana detectada.

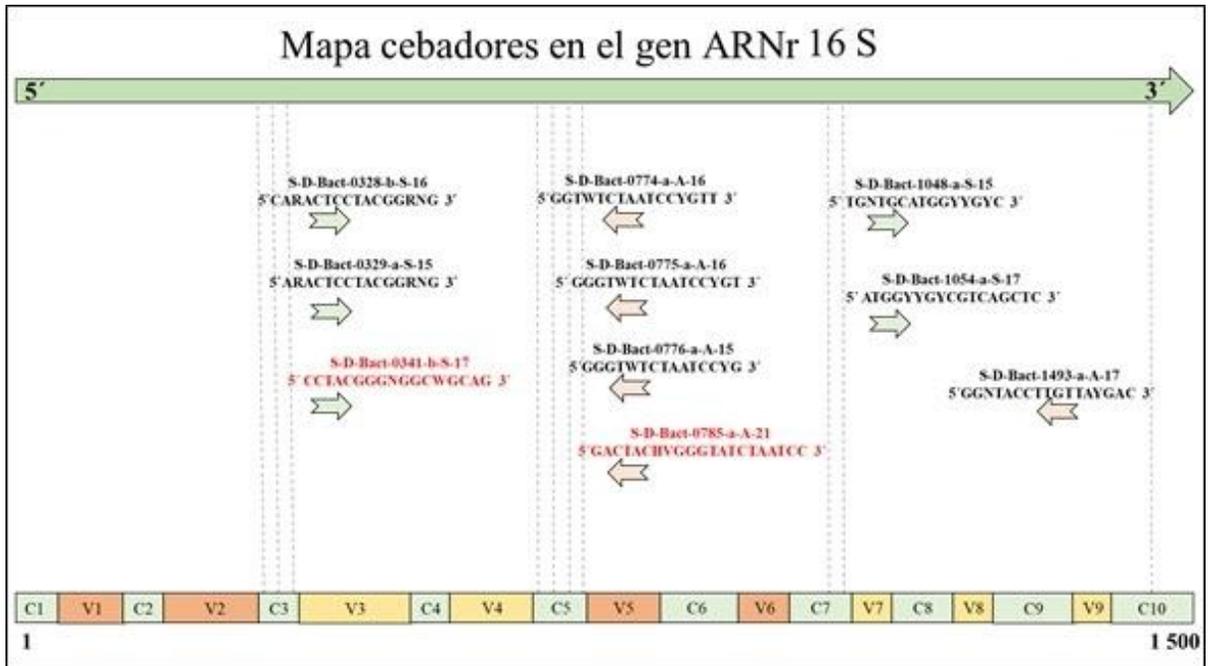


Figura 11. Ilustración de las diferentes regiones del ARNr 16S. Los recuadros en color amarillo representan a las regiones variables a amplificar. Letras en color negro representan a los pares de cebadores diseñados a partir de fragmentos altamente conservados, letras en color rojo representa el par de cebadores propuestos por Klindworth y col. 2013. Flecha color verde indican a los cebadores sentido o forward y flechas color rosa indican a los cebadores anti-sentido o reverse.

### 6.3 Evaluación *in vitro* de los Cebadores

Como primera etapa del análisis *in vitro* de los cebadores obtenidos se procedió a probar su capacidad de amplificación y, además, determinar la temperatura de alineamiento adecuada para cada cebador. Para ello, se llevó a cabo una PCR convencional de gradiente empleando ADN extraído de muestra de suelo. Se utilizaron 11 temperaturas siendo la mínima 46°C y la máxima 60°C.

La selección de las temperaturas se llevó a cabo tomando 5°C abajo y 5°C arriba de la temperatura de alineamiento de los cebadores (entre 50 y 55°C). Posteriormente, se realizó una electroforesis capilar utilizando cintas de alta sensibilidad (High Sensitivity D1000 ScreenTape Assay) que se colocaron en el instrumento TapeStation 2200 de la marca Agilent Technologies. Todos los cebadores tanto para la amplificación de las regiones V3-V4 y V7-V9 obtuvieron amplicones de un tamaño entre 450 y 500 pb y además se observó que las amplificaciones fueron más adecuadas entre los 46 y 57°C (Figura 12).

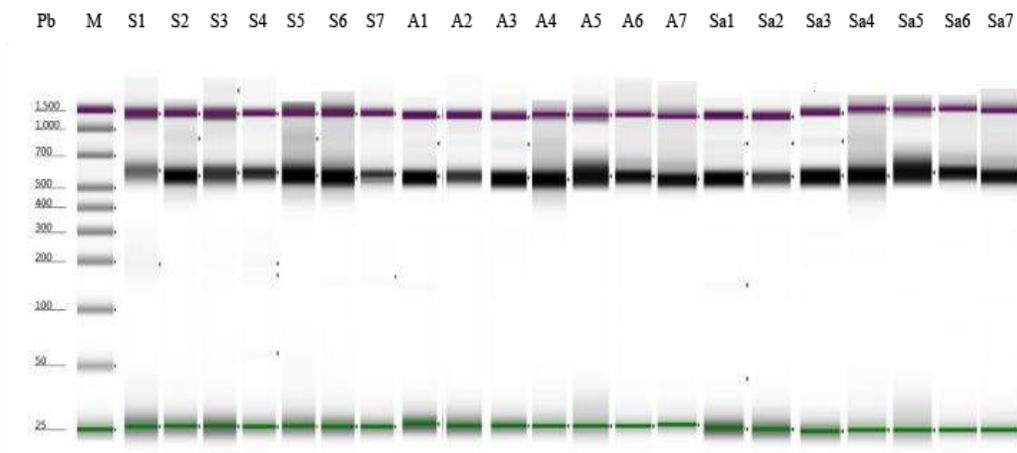


Figura 12. Amplificaciones de ADN de muestras de saliva (S), Agua de mar (A) y suelo agrícola (Sa) empleando cebadores para las regiones variables 3 - 4 (1-5) y las regiones variables 7-9 (6-7).

### 6.3.1 Evaluación *in vitro* de los Cebadores para la Amplificación de las Regiones Variables 3- 4 y 7-9

Una vez probados los cebadores propuestos y habiendo determinado las condiciones óptimas para su utilización, se procedió a realizar la secuenciación masiva del ARNr 16 S de las bacterias contenidas en muestras biológicas. Obtenidos los datos se procedió a su análisis con la finalidad de establecer diferencias en cuanto al perfil bacteriano y se observó que, independientemente del par de cebador empleado, ya sea los propuestos para las regiones V3-V4 y V7-V9 o los de Klindworth y col., la diversidad bacteriana detectada por estos en cada muestra fue similar en la mayoría de los casos (Figura 13). El análisis de componentes principales, mostró un agrupamiento por muestra de perfiles taxonómicos. Sin embargo, se observó que, tanto para las muestras de suelo agrícola como para las de saliva, los perfiles taxonómicos generados a partir de los cebadores que amplificaban las regiones V7-V9, se alejaban un poco de los obtenidos para V3-V4.

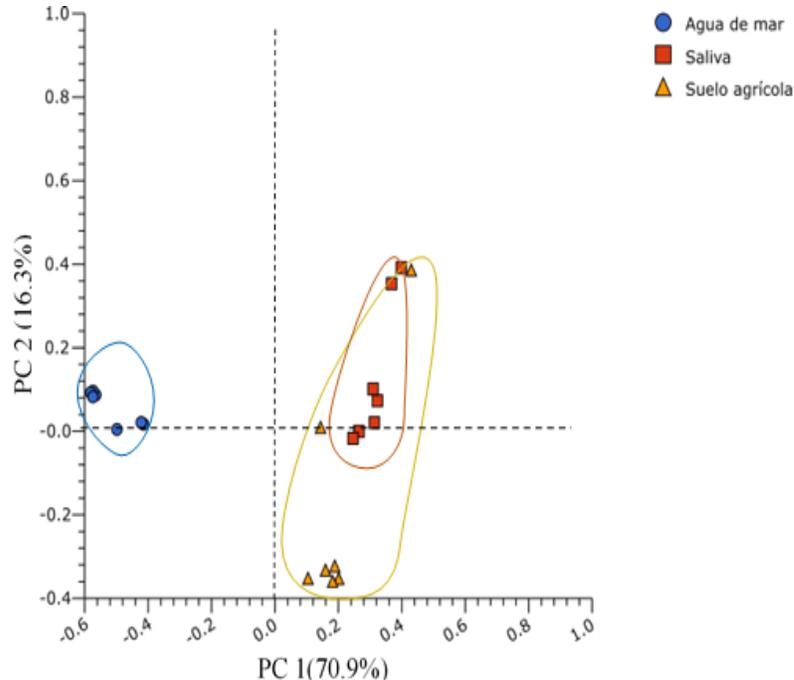


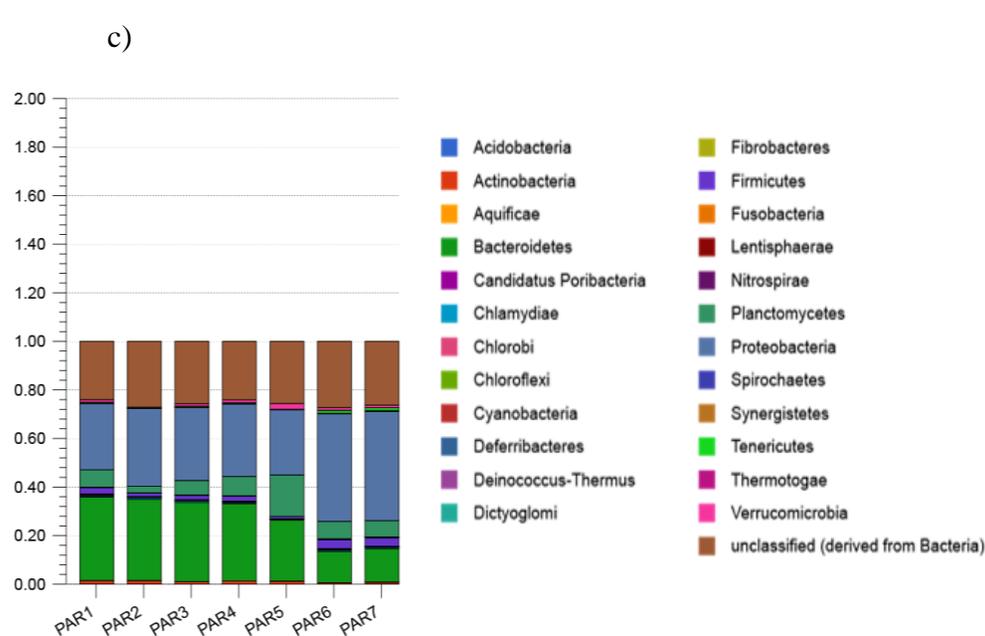
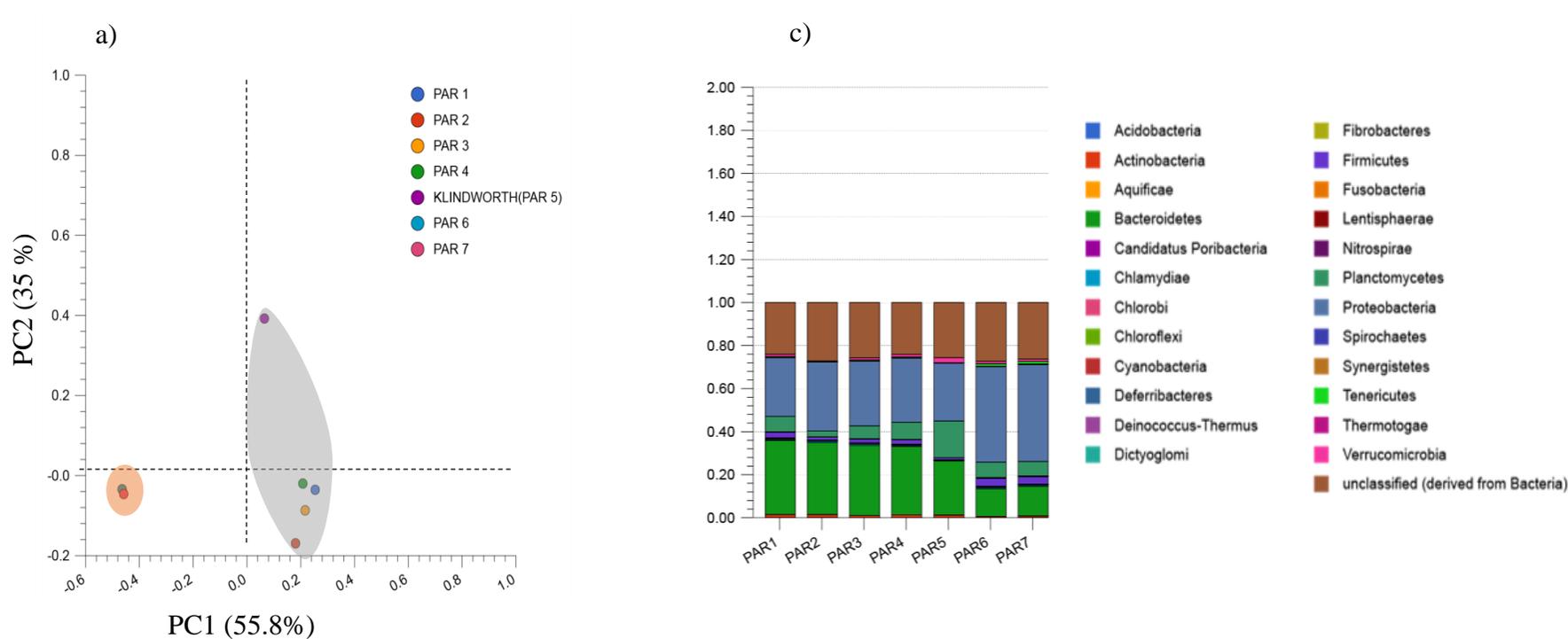
Figura 13. Análisis de componentes principales de diversidad bacteriana en muestras de agua de mar, saliva y suelo utilizando asignaciones taxonómicas a nivel de filo

Considerando cada muestra por separado, también se observaron ciertas diferencias con respecto a lo revelado por las diferentes regiones estudiadas. Por ejemplo, respecto a la muestra de agua de mar es posible apreciar dos diferentes conjuntos (Figura 16a). Los cebadores propuestos (1-4) para la amplificación de las regiones V3-V4, se encuentran muy próximos entre sí, aunque distantes de los perfiles obtenidos por los cebadores propuestos para la amplificación de las regiones V7-V9 (6 y 7) y los de Klindworth y col. Lo cual significa, que la cantidad de unidades taxonómicas operacionales (OTU'S) detectados por estos son muy similares. Esto se observa en las comparaciones pareadas de  $\beta$ -diversidad (Figura 16b) (Halffer et al., 2001) en donde las diferencias entre los cebadores 1 al 4, son menores (0.00-0.02), pero se acentúan cuando se comparan con los cebadores de Klindworth y col. (0.03-0.07) y con los diseñados para V7-V9 (0.06-0.10). Además, los resultados demuestran que la mayor riqueza taxonómica se obtuvo con los cebadores propuestos por Klindworth et al., 2013, seguido del par 1, 4, 3, 2 (V3-V4), 7 y 6 (V7-V9), respectivamente. Resultados similares se obtuvieron en el análisis *in silico* donde los cebadores propuestos por Klindworth presentaron un porcentaje ligeramente mayor (84%) a los cebadores propuestos (80%). Por lo que los resultados obtenidos *in silico* terminan siendo validados por los resultados del análisis *in vitro* para la muestra de agua de mar.

Evidentemente, las diferencias entre cebadores se deben a la mayor cantidad y diversidad de OTUs obtenidas por cada uno de los cebadores. Un contraste entre los resultados obtenidos por el análisis *in silico* e *in vitro*, es la baja riqueza taxonómica encontrada en agua marina (y en el resto de las muestras), mientras que los resultados *in silico* mostraron una mayor detección de filos. Esto puede deberse a que, aunque los ecosistemas son ricos en diversidad taxonómica de microorganismos, esta diversidad está influenciada por las variables medioambientales, que favorece a ciertos tipos de microorganismos ante otros, mientras que el análisis *in silico* considera el universo de las secuencias bacterianas registradas en las bases de datos.

En general, se detectaron en promedio 22 filos. Aquellos que obtuvieron mayor porcentaje de detección fueron los filos *Bacteroidetes*, *Proteobacteria*, *Planctomycetes*, *Actinobacteria*, *Verrucomicrobia*, *Chlamydiae* y *Firmicutes* (Figura 16c), lo cual

concuerta con los resultados obtenidos en el análisis *in silico*, donde todos los pares de cebadores pudieron detectar a todos estos filos. Además, concuerda con los filos detectados en estudios de diversidad bacteriana en ambiente marino reportado por Campell y col. (2011).



b)

Índice de Shannon	Cebador 1	Cebador 2	Cebador 3	Cebador 4	Klindworth 5	Cebador 6	Cebador 7
$\alpha$ -Diversidad							
1.26	1.39	1.23	1.29	1.36	1.41	1.08	1.04
Cebador 1	0.00	0.02	0.01	0.01	0.03	0.08	0.08
Cebador 2	0.02	0.00	0.02	0.02	0.07	0.07	0.06
Cebador 3	0.01	0.02	0.00	0.00	0.03	0.06	0.06
Cebador 4	0.01	0.02	0.00	0.00	0.02	0.06	0.06
Klindworth 5	0.03	0.07	0.03	0.02	0.00	0.08	0.10
Cebador 6	0.08	0.07	0.06	0.06	0.08	0.00	0.03
Cebador 7	0.08	0.06	0.06	0.06	0.10	0.03	0.00

Figura 14. Análisis de diversidad bacteriana en agua de mar. a) Análisis de componentes principales utilizando asignaciones taxonómicas a nivel de filo; b) Alfa y beta diversidad entre los siete diferentes pares de cebadores; c) Perfil taxonómico detectado por cada par de cebador hasta nivel de filo.

Con respecto a la muestra de saliva es posible apreciar dos diferentes conjuntos (Figura 17a). Los cebadores propuestos (1-4) para la amplificación de las regiones V3-V4, se encuentran muy próximos entre sí, aunque distantes de los perfiles obtenidos por los cebadores propuestos para la amplificación de las regiones V7-V9 (6 y 7). Los pares de cebadores propuestos uno y dos, se encuentran más próximos a los cebadores propuestos por Klindworth y col. Lo cual indica que la cantidad de OTU'S detectados por estos son muy similares.

En las comparaciones pareadas de  $\beta$ -diversidad (Figura 17b) se observa que las diferencias entre los cebadores 1 al 4, son menores (0.00-0.01), de igual manera cuando se comparan con los cebadores de Klindworth y col. (0.00-0.02) y aumenta con los diseñados para V7-V9 (0.02-0.04) sin embargo, las diferencias siguen siendo mínimas. Además, los resultados demuestran que la mayor riqueza taxonómica se obtuvo con los cebadores propuestos por Klindworth et al., seguido del par 3, 2, 1, 4 (V3-V4), 7 y 6 (V7-V9), respectivamente. Estos resultados nos indican que, los cebadores propuestos para la amplificación de las regiones V3-V4 fueron muy similares en cuanto a la detección de filos bacterianos en comparación con los cebadores reportados por Klindworth y col. Por lo que, al menos en muestras de saliva es fiable la utilización de estos cebadores.

En general, se detectaron en promedio 21 filos. Aquellos que obtuvieron mayor porcentaje de detección fueron los filos *Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, *Firmicutes* (Figura 17c), lo cual concuerda con los resultados obtenidos en el análisis *in silico*, donde todos los pares de cebadores pudieron detectar a todos estos filos. Además, concuerda con los filos detectados en estudios de diversidad bacteriana en cavidad bucal en pacientes sanos reportado por Aas y col. (2005). También es importante destacar que la diversidad bacteriana en este tipo de ambiente es muy susceptible a cambios potenciados por enfermedades (Cameron et al., 2015).

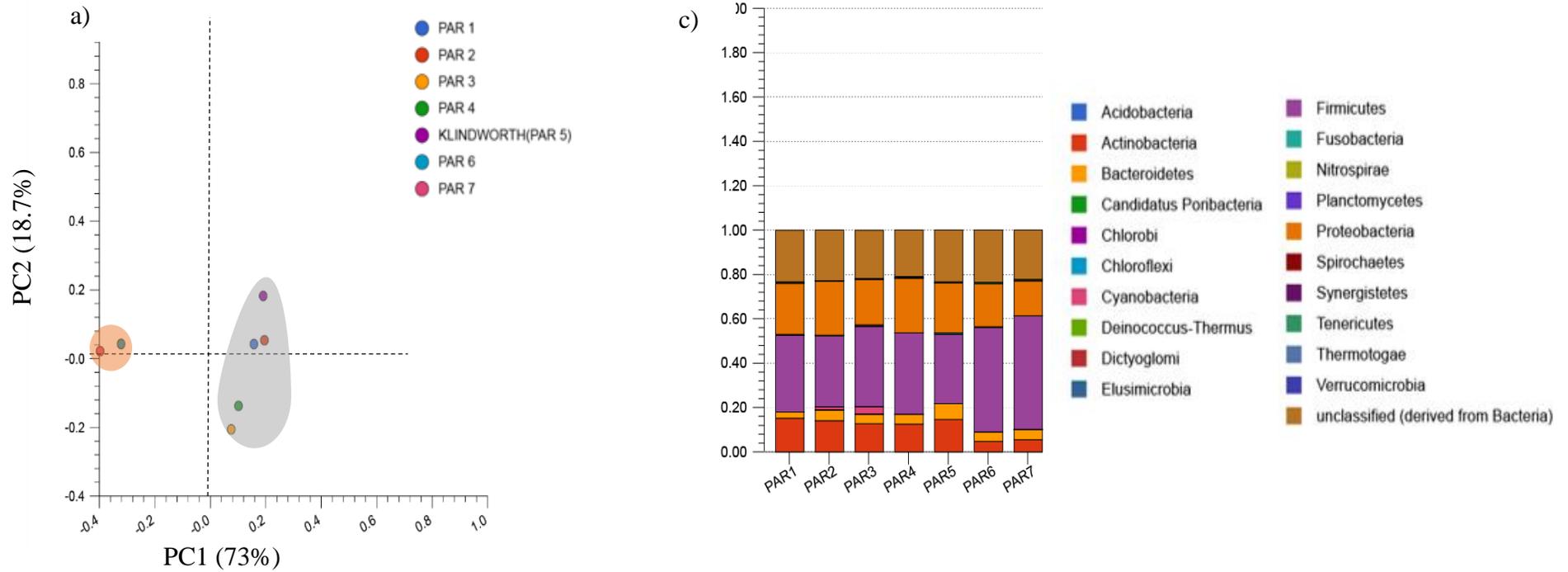


Figura 15. Análisis de diversidad bacteriana en saliva. a) Análisis de componentes principales utilizando asignaciones taxonómicas a nivel de filo; b) Alfa y beta diversidad entre los siete diferentes pares de cebadores; c) Perfil taxonómico detectado por cada par de cebador hasta nivel de filo.

De igual manera en la muestra de suelo se observó que los cebadores propuestos (1-4) para la amplificación de las regiones V3-V4, se encuentran muy próximos entre sí, aunque distantes de los perfiles obtenidos por los cebadores propuestos para la amplificación de las regiones V7-V9 (6 y 7) y los de Klindworth y col. (Figura 18a). Además, los dos pares de cebadores propuestos para la amplificación de las regiones V7-V9 (6 y 7), se mostraron muy alejados uno del otro. Lo que podría indicar, que la cantidad de secuencias o OTU'S que presentan pudieran no ser tan similares.

En el análisis de  $\beta$ -diversidad (Figura 18b) se observa que los valores entre los cebadores 1 al 4, son menores (0.00-0.04), estos aumentan levemente cuando se comparan con los cebadores de Klindworth y col. (0.01-0.07) y con los diseñados para V7-V9 (0.04-0.18). Además, los resultados demuestran que la mayor riqueza taxonómica se obtuvo con los cebadores propuestos por Klindworth et al., seguido por el par 3, 1, 4, 2 (V3-V4), 6 y 7 (V7-V9), respectivamente. El par siete para la amplificación de las regiones V7-V9, mostro menor riqueza taxonómica, lo que quiere decir que tuvo una detección de OTU'S menor al resto de los cebadores. Resultados diferentes a las muestras de agua de mar y saliva en donde este detecto más OTU'S.

En la muestra de suelo fue donde se detectaron más filos con un total de 28. Aquellos que obtuvieron mayor porcentaje de detección fueron los filos *Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, *Firmicutes*, *Planctomycetes*, *Chloroflexi*, *Cyanobacteria*, *Dictyoglomi*, *Nitrospirae*. (Figura 18c), lo cual concuerda con los resultados obtenidos en el análisis *in silico*, donde todos los pares de cebadores pudieron detectar a todos estos filos. Además, estos filos se asocian con el tipo de suelo árido tal y como lo reporta Bachar y col. 2012. Esta diferencia en la diversidad bacteriana en este tipo de suelo es debido a las características climáticas como la temperatura, la radiación solar (Kidron, 2009) así como la escasa vegetación contribuyen a los cambios en la diversidad bacteriana, ya que estos factores alteran la composición de estas. (Norris et al., 2002; Rainey et al., 2005).

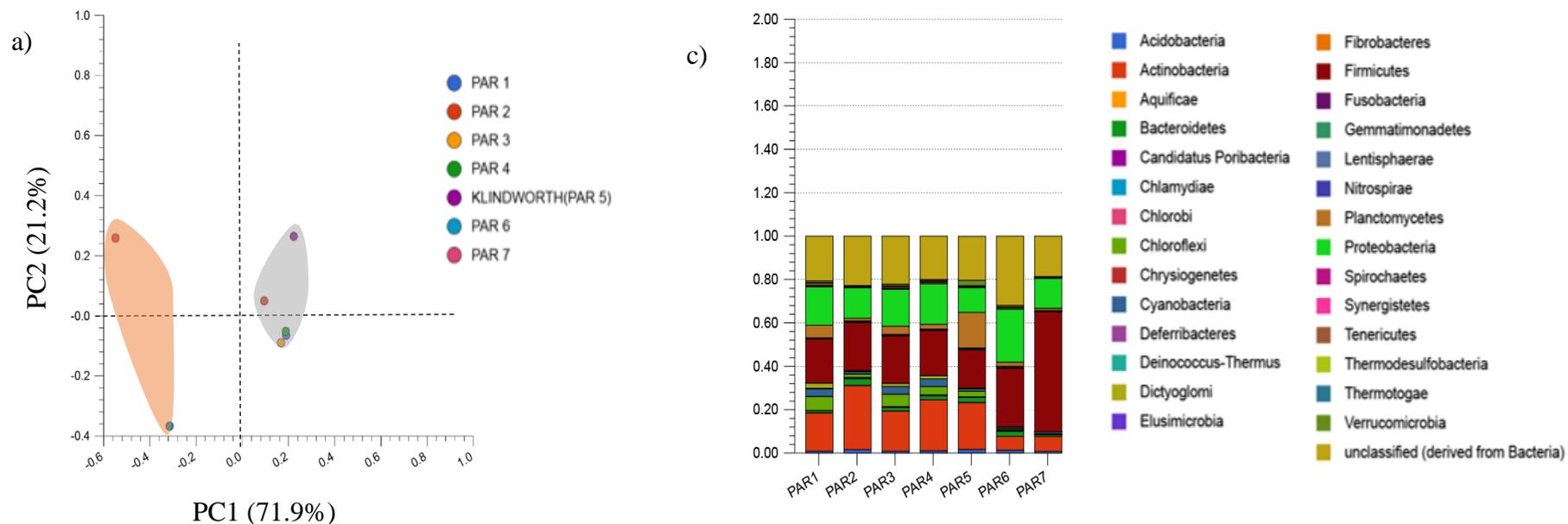


Figura 16. Análisis de diversidad bacteriana en suelo agrícola. a) Análisis de componentes principales utilizando asignaciones taxonómicas a nivel de filo; b) Alfa y beta diversidad entre los siete diferentes pares de cebadores; c) Perfil taxonómico detectado por cada par de cebador hasta nivel de filo.

En la figura 17, se presenta un mapa de calor que indica la abundancia de filios y un dendrograma que muestra la similitud entre los cebadores empleados en cada muestra. Se observa que todos los pares de cebadores fueron muy similares entre si dentro de cada muestra. Excepto los cebadores propuestos para la amplificación de las regiones V7-V9, empleados en la muestra de saliva, que presenta una mayor similitud en la diversidad bacteriana detectada por estos con respecto a la detectada por los mismos cebadores, pero empleados en la muestra de suelo agrícola.

En general, los filios que fueron detectados por los cebadores y presentaron mayor abundancia en las distintas muestras fueron *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, *Firmicutes*, *Verrucomicrobia*, *Acidobacteria* y *Planctomycetes*. Sin embargo, hubo pares de cebadores que sobreestimaron, es decir, presentaron mayor abundancia para ciertos filios o, por el contrario, no pudieron detectarlos dentro de una misma muestra. Por ejemplo, para el filo *Acidobacteria* solo el par tres de cebadores propuestos para la amplificación de las regiones V3-V4 pudo detectarlo en la muestra de saliva. Por otro lado, en la muestra de agua de mar solo el par de cebadores propuestos por Klindworth y col. 2013 no pudo detectar este filo. De nuevo en la muestra de saliva, solo cuatro pares pudieron detectar el filo *Planctomycetes*, los pares propuestos uno y dos para las regiones V3-V4 y los dos pares propuestos para las regiones V7-V9. Los cebadores propuestos por Klindworth y col. no pudieron detectar este filo en esta misma muestra. En cambio, solo dos pares de cebadores pudieron detectar al filo *Verrucomicrobia*, el par propuesto número tres y los cebadores propuestos por Klindworth y col. 2013.

Se observó una buena estimación de filios reportados para ese tipo de muestras por la mayoría de los cebadores propuestos. Los cebadores propuestos para la amplificación de las regiones variables 3 y 4, lograron detectar algunos filios no tan estudiados como *Chloroflexi*, *Chlorobi*, *Synergistetes*, *Spirochaetes*, *Lentisphaerae*, *Aquificae*, entre otros. Por otro lado, los pares de cebadores propuestos para la amplificación de las regiones extremas 7-9 detectaron los filios bacterianos reportados comúnmente para estos ambientes y filios que los cebadores que amplifican las regiones centrales del ARNr 16S no pudieron. Filios de bacterias no cultivables como *Deferribacteres*, *Deinococcus-Thermus* y

*Candidatus poribacteria*, que se desarrollan en ambientes hostiles haciendo difícil su estudio (Becraft et al., 2017).

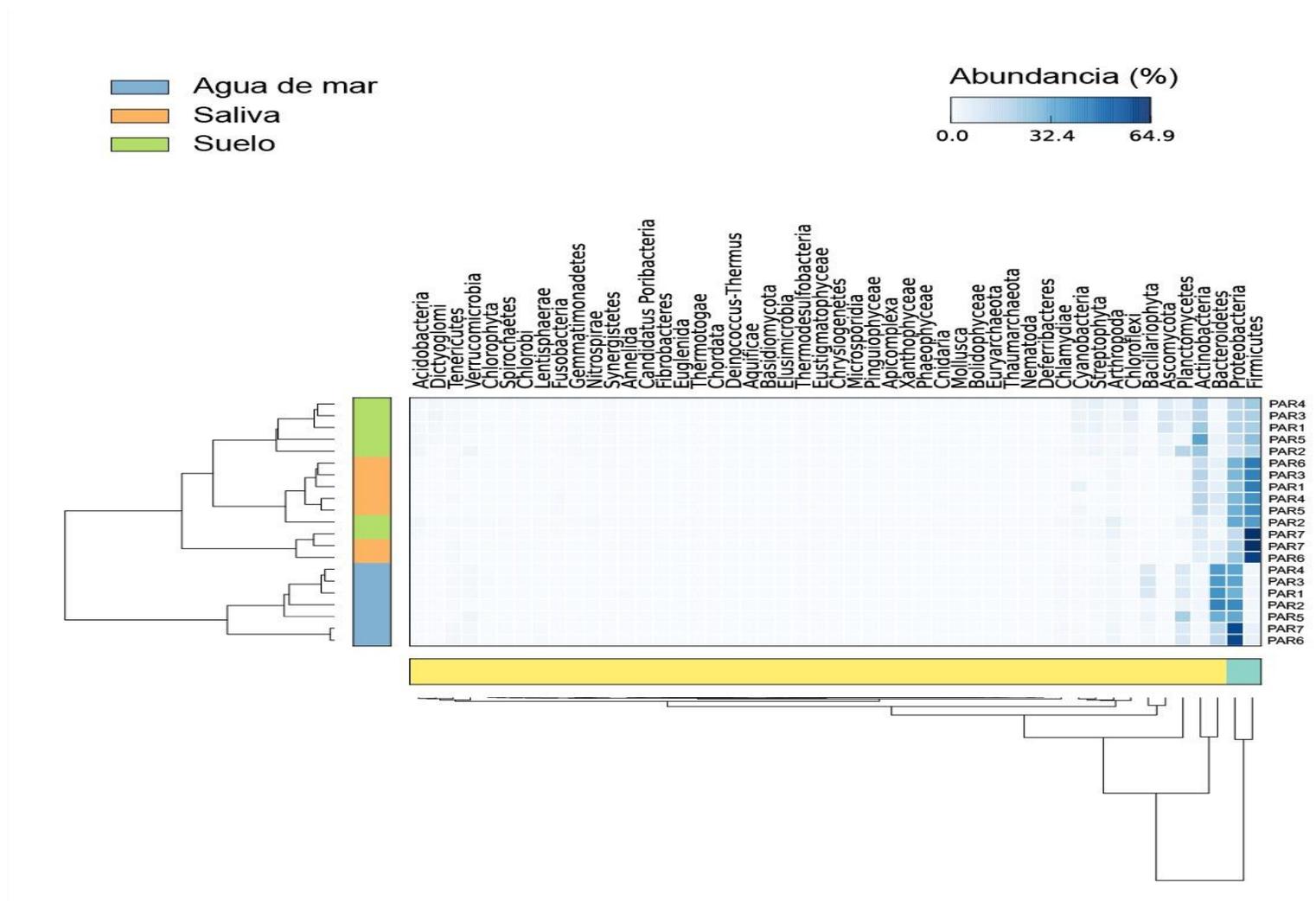


Figura 17. Mapa de calor. Filos detectados por los cebadores para la amplificación de las regiones V3-V4 (1-5) y las regiones V7-V9 (6-7) en tres tipos de muestra: Agua de mar, saliva y suelo agrícola.

Por otro lado, la diferencias en las secuencias sentido y anti-sentido de los pares de cebadores no fue una condicionante para diferenciar un par de otro. Por ejemplo, los pares de cebadores 1 y 2 poseen la misma secuencia sentido (5'ARACTCCTACGGRNG 3') pero las secuencias anti-sentido de ambos son diferentes ya que en varían en tamaño (15 y 16 nucleótidos, respectivamente) y conformación: el par 2 posee tres guaninas al extremo 5' y dos timinas al extremo 3'. (5'GGGTWTCTAATCCYG-GGTWTCTAATCCYGT 3'). De igual manera, las secuencias anti-sentido de los pares 3 y 4 (5'GGGTWTCTAATCCYG- GGTWTCTAATCCYGT 3') solo que la secuencia anti-sentido del par 4 posee una timina más al extremo 3' que el par 3. Por otra parte, los cebadores 6 y 7 diseñados para la amplificación de las regiones V7-V9 difieren en la secuencia sentido, las cuales varían en tamaño y conformación (5'TGNTGCATGGYYGY-ATGGYYGYCGTCAGCTC 3').

Uno de los criterios principales para una adecuada selección y/o diseño de cebadores es la posición 3' terminal del cebador, el cual es importante para controlar el apareamiento incorrecto (Ayyadevara et al., 2000). Se considera la colocación de una guanina o una citosina, debido a que ayudan a la correcta hibridación del extremo 3', gracias a los enlaces triples de Hidrogeno que se dan entre estas dos bases. Además, mejoran la eficiencia de la reacción de PCR reduciendo cualquier deshibridación que pudiera ocurrir. (Huang et al., 1992). El par de cebadores 2 posee en su extremo 3' dos timinas, sin embargo, esto no pareciera ser una limitante en la eficiencia en la detección de filos por parte de este par, dado que mostro resultados similares al resto de los cebadores.

Respecto a los tamaños, el porcentaje de guaninas y citosinas, y las terminaciones en los extremos 5' y 3' y contenido de degeneraciones son muy similares a las propiedades reportadas por los cebadores propuestos por Klindworth y col. 2013. Además, todos los cebadores que permiten la amplificación de las regiones variables 3-4 se encuentran ubicados casi en la misma posición o muy próximos entre sí en la secuencia del ARNr 16S; al igual que los cebadores que amplifican las regiones variables 7-9. Estas similitudes entre los cebadores dieron como resultado coberturas y especificidades y una detección de OTU's elevadas y similares entre sí, además de valores de  $\beta$ -diversidad bajos. Por lo cual

no se detectan diferencias entre estos en cuanto a diversidad bacteriana entre los pares que presentaban diferentes secuencias reverse (1-2 y 3-4) y los cebadores que presentaban diferentes secuencias forward (6-7). El tamaño del cebador y la diferencia en su conformación nucleotídica, no fueron limitantes en la eficiencia de los pares de cebadores individualmente como entre ellos. Por lo que, pudiera considerarse el uso de cualquiera de ellos en estudios de diversidad bacteriana.

## 7 CONCLUSIONES

Los cebadores diseñados a partir de zonas con altamente conservadas del ARNr 16S, demostraron ser óptimos al obtener de una amplia cobertura de secuencias bacterianas y una detección de filos amplia (entre 60 y 77 filos) en el análisis *in silico*.

Los filos *Proteobacteria*, *Bacteroidetes*, *Firmicutes*, *Actinobacteria*, fueron los más abundantes en las muestras de agua, saliva y suelo empleando los cebadores diseñados para la amplificación de las regiones variables 3-4 y 7-9 en el análisis *in vitro*.

Los cebadores diseñados para la amplificación de las regiones variables 3-4, mostraron la una buena capacidad de detección de OTU's en el análisis *in vitro*.

El par 3 propuesto para la amplificación de las regiones V3-V4, es eficiente para la detección de los filos *Acidobacteria* y *Verrucomicrobia* en muestras de saliva. Del mismo modo los pares 1 y 2 propuestos para la amplificación de las regiones V3-V4 y los pares 6 y 7 propuestos para la amplificación de las regiones V7-V9 son eficientes en la detección del filo *Planctomycetes*

Los cebadores propuestos para la amplificación de las regiones V7-V9, fueron los únicos en detectar filos poco estudiados como *Deferribacteres*, *Deinococcus-Thermus* y *Candidatus poribacteria*.

Los cebadores propuestos para la amplificación de las regiones V3-V4 y V7-V9 demostraron ser fiables para su utilización en estudios de diversidad bacteriana mediante secuenciación masiva en muestras de agua de mar, saliva y suelo.

## **8 RECOMENDACIONES**

Los cebadores propuestos para la amplificación de las regiones variables 3-4 y 7-9 del ARNr 16S, solo fueron evaluados en muestras de agua de mar, saliva y suelo. Por lo que, sería recomendable evaluar en muestra de diferente índole.

## 9 REFERENCIAS

- Aas, J., Paster, B., Stokes, L., Olsen, I. and Dewhirst, F. (2005). Defining the normal bacterial flora of the oral cavity. *Journal of Clinical Microbiology*.43: 5721-5732.
- Abubucker, S., Segata, N., Goll, J., Schubert, A., Izard, J., Cantarel, B., Rodriguez-Mueller, B., Zucker, J., Thiagarajan, M., Henrissat, B., White, O., Kelley, S., Methé, B., Schloss, P., Gevers, D., Mitreva, M. and Huttenhower, C.(2012). Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Computational Biology*. 8: e1002358.
- Acharya, A., Chan, Y., Kheur, S., Jin, L, Watt, R. and Mattheos, N. (2017). Salivary microbiome in non-oral disease: A summary of evidence and commentary. *Archives of Oral Biology*. 83: 169-173.
- Acinas, S.G. (2007). Diversidad y estructura de una comunidad microbiana. *Sociedad Española de Microbiología*. 44: 24-29.
- Albertsen, M., Karst, S., Ziegler, A., Kirkegaard, R and Nielsen, P. (2015). Back to basics the influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. *PLOS ONE*. 10: e0132783.
- Aly, S.M., Abd-El-Rahman, A.M., John, G. and Mohamed, M.F. (2008). Characterization of some bacterial isolated from *Oreochromis niloticus* and their potential use as probiotics. *Aquaculture*. 277: 1-6.
- Ayyadevara, S., Thaden, J. and Shmookler Reis, J. (2000). Discrimination of primer 3'-nucleotide mismatch by Taq DNA Polymerase during polymerase chain reaction. *Analytical Biochemistry*. 284:11-18.
- Bachar, A., Al-Ashhab, A., Soares, M., Sklarz, M., Angel, R., Ungar, E. and Gillor, O. (2010). Soil microbial abundance and diversity along a low precipitation gradient. *Microbial Ecology*. 60: 453-461.
- Bachar, A., Soares, M. and Gillor, O. (2012). The effect of resource islands on abundance and diversity of bacteria in arid soils. *Microbial Ecology*. 63: 694-700.
- Baker, G., Smith, J. and Cowan, D. (2003). Review and re-analysis of domain-specific 16S primers. *Journal Microbiology Methods*.55: 541-555.

- Becraft, E., Woyke, T., Jarett, J., Ivanova, N., Godoy-Vitorino, F., Poulton, N., Brown, J., Brown, J., Lau, M., Onstott, T., Eisen, J., Moser, D. and Stepanauskas, R. (2017). *Rokubacteria*: Genomic Giants among the Uncultured Bacterial Phyla. *Frontiers in Microbiology*. 8: 22-64.
- Bertagnolli, A. D., Padilla, C. C., Glass, J. B., Thamdrup, B. and Stewart, F. J. (2017), Metabolic potential and in situ activity of marine *Marinimicrobia* bacteria in an anoxic water column. *Environmental Microbiology*. 19: 4392-4416.
- Biosonda S.A. Reacción de pirosecuenciación (2016). Disponible en: <http://www.biosonda.cl/producto/pyromark>
- Cadena, J., Martínez, M., Guzmán, L. y Arteaga, L. (2016). Aplicación de secuenciación masiva para el estudio y exploración de diversidad microbiana y su aprovechamiento biotecnológico. *Agroproductividad*. 9: 70-83.
- Cameron, S., Huws, S., Hegarty, M., Smith, D. and Mur, L. (2015). The human salivary microbiome exhibits temporal stability in bacterial diversity. *FEMS Microbiology Ecology*. 91: 91.
- Campbell, B., Yu, L., Heidelberg, J. and Kirchman, D. (2011). Activity of abundant and rare bacteria in a coastal ocean. *Proceedings of the National Academy of Sciences of the United States of America*. 108: 12776-12781.
- Capel-Salinas, J. y Yuste- Lisbona, F. (2016). *Manual de Prácticas en Bioinformática*. Editorial Universidad de Almería. 5:112.
- Carmona, V. y Carmona, T. (2013). La Diversidad de los análisis de diversidad. *Bioma*. 14: 20-28.
- Chakravorty, S., Helb, D., Burday, M., Connell, N. and Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods*. 69: 330-339.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T., Higgins, D. and Thompson, J. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research*. 31: 3497–3500.
- Chun, J. and Rainey, F. (2014). Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *International Journal of Systematic and Evolutionary Microbiology*. 64: 316-324.

- Claesson, M., Wang, Q., O'Sullivan, O., Greene-Diniz, R., Cole, J., Ross, R. and O'Toole, P. (2010). Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Research*. 38: e200.
- Cruaud, P., Vigneron, A., Luchetti-Miganeh, C., Ciron, P., Godfroy, A. and Cambon-Bonavita, M. (2014). Influence of DNA extraction method, 16S rRNA targeted hypervariable regions, and simple origin on microbial diversity detected by 454 pyrosequencing in marine chemosynthetic ecosystems. *Applied and Environmental Microbiology*. 80:4626-4639.
- Croci, L., Suredini, E., Cozzi, L., Toti, L., Ottaviani, D., Pruzzo, C., Serratore, P., Fischetti, R., Goredó, E., Loredó, G. and Mioni, R. (2007). Comparison of different biochemical and molecular methods for the identification of *Vibrio parahaemolyticus*. *Journal of Applied Microbiology*. 102: 229-237.
- Cock, J., Tessmar-Raible, K., Boyen, C. and Viard, F. (2010). Introduction to Marine Genomics, *Advances in Marine Genomics*. Springer. Primera edición. Dordrecht, 399. pp.
- Curtis, T., Sloan, W. and Scannell, J. (2002). Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences*. 99: 10494-10499.
- Díaz, R. and Wachter, R. (2003). Métodos para el estudio de comunidades microbianas en alimentos fermentados. *Revista Latinoamericana de Microbiología*. 45: 30-40.
- Dieffenbach, C., Lowe, T. and Dveksler, G. (1993). General concepts for PCR primer design. *PCR methods applied*. 3: 30-37.
- Domínguez, M., Garibay, O., Poggi-Varaldo, H. y García, M. (2006). Caracterización de la diversidad de comunidades microbianas útiles en biorremediación y producción de probióticos por su huella genética. *Revista Latinoamericana Microbiología*. 48: 211-225.
- Ercolini, D., Moschetti, G., Blaiotta, G. and Coppola S. (2001). The potential of polyphasic PCR-DGGE approach in evaluating microbial diversity of natural whey cultures for wáter-buffalo Mozzarella cheese production: bias of culture dependent and culture independent approaches. *Systematic and Applied Microbiology*. 24: 610-617.

- Fierer, N., Bradford, M. and Jackson, R. (2007). Toward an ecological classification of soil bacteria. *Ecology*. 88:1354-1364.
- Guo, F., Ju, F., Cai, L. and Zhang, T., (2013). Taxonomic precision of different hypervariable regions of 16s rRNA gene and annotation methods for functional bacterial groups in biological wastewater treatment. *PLoS ONE*. 8: e76185.
- Guerrero, R. y Berlanga, M. (2005). Microbios en la niebla: descubriendo el papel de los microbios en la biosfera. *Ecosistemas*. 14: 3-10.
- Herlemann, D., Labrenz, M., Juergens, K., Bertilsson, S., Waniek, J., Andersson, A. (2011). Transition in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J*. 5:1571–1579.
- Huang, M., Arnheim, N., and Goodman, M.F. (1992). Extension of base mispairs by Taq DNA polymerase: Implications for single nucleotide discrimination in PCR. *Nucleic Acids Research*. 20: 4567-4573.
- Huber, I., Spanggaard, B., Appel, K.F., Rossen, L., Nielsen, T. and Gram, L. (2004). Phylogenetic analysis and in situ identification of the intestinal microbial community of rainbow trout *Oncorhynchus mykiss* (Walbaum). *Journal of Applied Microbiology*. 96: 117-132.
- Handelsman, J. (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews*. 68: 669-685.
- Hugenholtz, P., Goebel, B. M. and Pace, N. R. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology*. 180: 4765-4774.
- Huse, S., Dethlefsen, L., Huber, J., Welch, D., Relman, D. and Sogin, M. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable Tag sequencing. *PLOS Genet*. 4: e1000255.
- Huws, S., Edwards, J., Kim, E. and Scollan, N. (2007). Specificity and sensitivity of eubacterial primers utilized for molecular profiling of bacteria within complex microbial ecosystems. *Journal Microbiology Methods*. 70: 565-569.
- Ju, J., Kim, D., Bi, L., Meng, Q., Bai, X., Li, Z., Li, X., Marma, M., Shi, S., Wu, J., Edwards, J., Romu, A. and Turro, N. (2006). Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proceedings*

- of the National Academy of Sciences of the United States of America. 103: 19635-19640.
- Kim, D.H., Brunt, J. and Austin, B. (2007). Microbial diversity of intestinal contents and mucus in rainbow trout *Oncorhynchus mykiss*. *Journal of Applied Microbiology*. 102: 1654-1664.
- Kim, M., Morrison, M. and Yu Z. (2011). Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *Journal Microbiology Methods*. 84: 81-87.
- Kidron, G.J. (2009) The effect of shrub canopy upon Surface temperatures and evaporation in the Negev Desert. *Earth Surf Process Landforms*. 34:123-132.
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. and Glöckner, F.O. (2012) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. [Opens external link in new window](#) *Nucleic Acids Research*. 41: e1.
- Kunin, V, Engelbrektson, A., Ochman, H., and Hugenholtz, P. (2010). Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*. 12:118-123.
- Kress, W. and Erickson, D. (2012). *DNA Barcodes: Methods and Protocols*. *Methods in Molecular Biology*. Humana Press/ Springer. 1: 470.
- Lane, D., Pace, B., Olsen, G., Stahl, D., Sogin, M. and Pace, N. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences of the United States of America*. 82:6955-6959.
- Linhart, C. and Shamir, R. (2002). The degenerate primer design problem. *Bioinformatics*. 18:172-180.
- Liu, Z., Lozupone, C., Hamady, M., Bushman, f. and Knight, R. (2007). Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Research*. 35: e12.
- Liu, Z., De Santis, T., Andersen, G. and Knight, R., (2008). Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Research*. 36: e120.

- López de Heredia, U. (2016). Las técnicas de secuenciación masiva en el estudio de la diversidad biológica. *Munibe, Ciencias Naturales*. 64: 7-31.
- López, L. y Zaballos, M. (2005). Diversidad y actividad procariótica en ecosistemas marinos. *Ecosistemas*. 14: 30-40.
- Lyautey, E., Lacoste, B., Ten-Hage, L., Rols, J. and Garabetian, F. (2004). Analysis of bacterial diversity in rivers biofilms using 16S rDNA PCR-DGGE: methodological settings and fingerprints interpretation. *Water Research*. 39: 380-388.
- Martínez, M., Villalpando, E. and Vargas, F. (2016). Significant loss of sensitivity and specificity in the taxonomic classification occurs when short 16S rRNA gene sequences are used. *Heliyon*. 2: e00170.
- Martínez-Porchas, M., Villalpando-Canchola, E., Ortiz, L. and Vargas-Albores, F. (2017). How conserved are the conserved 16S rRNA regions? *PeerJ*. 5: e3036.
- Metzker, M.L. (2010). Sequencing technologies - the next generation. *Nature*. 11:31- 46.
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E., Kubal, M., Paczian, T., Rodríguez, A., Stevens, A., Wilke, A., Wilkening, J. and Edwards, R. (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*. 9: 386.
- Mizrahi-Man, O., Davenport, E. and Gilad, Y. (2013). Taxonomic classification of bacterial 16S rRNA genes using short sequencing reads: Evaluation of effective study designs. *PLoS ONE*. 8: e53608.
- Morales, S. and Holben, W. (2009). Empirical testing of 16S rRNA gene PCR primer pairs reveals variance in target specificity and efficacy not suggested by *in silico* analysis. *Applied Environmental Microbiology*. 75:2677–2683.
- Najafabadi, H., Torabi, N. and Chamankhah, M. (2008). Designing multiple degenerate primers via consecutive pairwise alignments. *BMC Bioinformatics*. 9:55.
- Neefs, J., Van de Peer, Y., Hendriks, L. and De Wachter, R. (1990). Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Research*. 18:2237-2330.
- Norris, T., Wraith, J., Castenholz, R. and McDermott, T. (2002). Soil microbial community structure across a thermal gradient following a geothermal heating event. *Applied Environmental Microbiology*. 68:6300–6309.

- Ortiz, M., Martínez, M. y Vargas, F. (2017). La secuencia completa del gen ARN ribosomal 16S, una promesa para mejorar la precisión en la asignación taxonómica microbiana. En: Arenas, M. y Chiappa, X. (eds.). Microbiología Ambiental en México. Diagnóstico, tendencias en investigación y áreas de oportunidad. D.R. Universidad Autónoma de México, Mérida, 50-62 pp.
- Pace, N., Stahl, D., Lane, D. and Olsen G. (1986). The analysis of natural microbial populations by ribosomal RNA sequences. In: Marshall KC (ed.), Advances in Microbial Ecology. Springer. 9: 1–55.
- Palm, R. (1998). L'analyse en composantes principales: principe et application. Notes de statistique et d'informatique. Gembloux, Belgique, 98/2: 31 p.
- Parks, D., Tyson, G., Hugenholtz, P. and Beiko, R. (2014). STAMP: Statistical analysis of taxonomic and functional profiles. Bioinformatics. 30: 3123-3124.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research. 41: 590-596.
- Qu, W., Zhou, Y., Zhang, Y., Lu, Y., Wang, X., Zhao, D., Yang, Y. and Zhang, C. (2012). MFEprimer- 2.0: a fast thermodynamics-based program for checking PCR primer specificity. Nucleic Acids Research. 40:205–208.
- Rainey, F., Ray, K., Ferreira, M., Gatz, B., Nobre, M., Bagaley, D., Rash, B., Park, M., Earl, A., Shank, N., Small, A., Henk, M., Battista, J., Kampfer, P. and da Costa, M. (2005). Extensive diversity of ionizing-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* obtained from a single soil sample. Applied Environmental Microbiology. 71:5225–5235.
- Ramos, B., Bucio, A., Bautista, C., Aranda, E. e Izquierdo, F. (2009). Aislamiento, identificación y caracterización de bacterias ácido lácticas para la elaboración de queso crema tropical. Universidad y ciencia. 25:159-171.
- Randazzo, C., Torriani, S., Akkermans, A., de Vos, W. and Vaughan, E. (2002). Diversity, dynamics, and activity of bacterial communities during production of an artisanal sicilian cheese as evaluated by 16S rRNA analysis. Applied Environmental Microbiology. 68:1882-1892.

- Revista Médica Genética. (2017). NGS: Secuenciación de Segunda Generación. Medigene Press, S.L. Parc Científic Universitat de València. <https://revistageneticamedica.com/blog/ngs-secuenciacion/>.
- Ribeiro, C., Steffens, M. Etto, R., Mazer, R., Galvão, C., Martins, C., Pedrosa, F., and Kolm, H. (2013). ARDRA profiles of bacteria and archaea in mangrove sediments with different levels of contamination in the estuarine complex of Paranaguá, Brazil. *Brazilian Archives of Biology and Technology*. 56: 275-281.
- Rodas, A., Ferrer, S. and Pardo, I. (2003). 16S-ARDRA, a Tool for Identification of Lactic Acid Bacteria Isolated from Grape Must and Wine. *ENOLAB Laboratorio de Microbiología Enológica. Departament de Microbiologia i Ecologia, Facultat de Ciències Biològiques, Universitat de València, Burjassot, València, Spain*. 26: 412-422.
- Rodicio, M. y Mendoza, M. (2004). Identificación bacteriana mediante secuenciación del ARNr 16S: Fundamento, metodología y aplicaciones en microbiología clínica. *Enfermedades Infecciosas y Microbiología Clínica*. 22: 238-24.
- Rodríguez, B. y Armengol, L. (2012). Tecnologías de secuenciación de nueva generación en diagnóstico genético pre- y postnatal. *Diagnóstico prenatal*. 23:56-66.
- Ruiz, C. (2012). Aplicación del análisis de componentes principales como técnica para obtener índices sintéticos de calidad ambiental. *UCV – Scientia*. 4: 145-153.
- Sardesai Y. N. (2005). Viable but non-culturable bacteria: their impact on public health. *Current Science*. 89: 1650.
- Shakya, M., Quince, C., Campbell, J., Yang, Z., Schadt, W. and Podar, M. (2013). Comparative metagenomic and Rna microbial diversity characterization using Archeal and Bacterial synthetic communities. *Environmental Microbiology*. 15: 1882-1899.
- Spring, S., Bunk, B., Spröer, C., Schumann, P., Rohde, M., Tindall, B. and Klenk, H. (2016). Characterization of the first cultured representative of *Verrucomicrobia* subdivision 5 indicates the proposal of a novel phylum. *The ISME Journal*. 10:2801-2816.

- Theron, J., and Cloete, T. (2000). Molecular techniques for determining microbial diversity and community structure in natural environments. *Critical Reviews in Microbiology*. 26: 37-57.
- Torsvik, V. and Ovreas, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion Microbiology*. 5: 240-245.
- Usha, R., Hema, S. and Kanchana, D. (2011). Antagonistic activity of actinomycetes isolates against human pathogen. *Journal Microbiology Biotechnology Research*. 1: 74-79.
- Villarreal, H., Álvarez, M., Córdoba, S., Escobar, F., Fagua, G., Gast, H., Mendoza, F., Ospina, M. y Umaña, A. (2006). *El Manual de métodos para el desarrollo de inventarios de biodiversidad*. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt. Segunda edición. Bogotá, 185-226 pp.
- Wang, Q., Garrity, G., Tiedje, J. and Cole, J. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied Environmental Microbiology*. 73: 5261-5267.
- Wang, S., Sun, B., Tu, J. and Lu, Z. (2016). Improving the microbial community reconstruction at the genus level by multiple 16S rRNA regions. *Journal of Theoretical Biology*. 398:1-8.
- Whittaker, R.H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*. 30: 279-338.
- Weinstock, G. M. (2012). Genomic approaches to studying the human microbiota. *Nature*. 489: 250-256.
- Woese, C. (1987). Bacterial evolution. *Microbiological Reviews*. 51: 221-271.
- Yang, B., Wang, Y. and Qian, P. (2016). Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics*. 17: 135.
- Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F., Ludwig, W., Schleifer, K., Whitman, W., Euzéby, J., Amann, R. and Rossello-Mora, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*. 12:635-645.
- Youssef, N., Sheik, C., Krumholz, L., Najar, F., Roe, B. and Elshahed, M. (2009). Comparison of species richness estimates obtained using nearly complete fragments

and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Applied Environmental Microbiology*. 75: 5227-5236.

Youssef, N., Couger, M., McCully, A., Guerrero, A. and Mostafa, S. (2015). Assessing the global phylum level diversity within the bacterial domain: A review. *Journal of Advanced Research*. 6: 269-282.

Zak, J., Willing, M., Moorhead, D. and Wildman, H. (1994). Functional diversity of microbial communities: A quantitative approach. *Soil Biology and Biochemistry*. 26:1101-1108.