

**Centro de Investigación Científica y de Educación  
Superior de Ensenada, Baja California**



---

**Maestría en Ciencias  
en Ciencias de la Vida con orientación  
en Biotecnología Marina**

---

Análisis metagenómico de los microbiomas asociados a esponjas marinas del Parque Nacional Cabo Pulmo (BCS), dirigido a la búsqueda de genes involucrados en la síntesis de policétidos y péptidos no ribosomales con potencial actividad antimicrobiana.

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de  
Maestro en Ciencias

Presenta:

Fernanda Isabel Ariza Pérez

Ensenada, Baja California, México  
2017

Tesis defendida por

**Fernanda Isabel Ariza Pérez**

y aprobada por el siguiente Comité

---

**Dra. María Asunción Lago Lestón**  
Directora de tesis

**Dra. Johanna Bernáldez Sarabia**

**Dr. Carlos Alberto Brizuela Rodríguez**

**Dra. Alejandra Prieto Davó**



---

**Dra. Clara Elizabeth Galindo Sánchez**  
Coordinador del Posgrado en Ciencias de la Vida

---

**Dra. Rufina Hernández Martínez**  
Directora de Estudios de Posgrado

Fernanda Isabel Ariza Pérez © 2017

*Queda prohibida la reproducción parcial o total de esta obra sin el permiso formal y explícito del autor y director de la tesis.*

Resumen de la tesis que presenta **Fernanda Isabel Ariza Pérez**, como requisito parcial para la obtención del grado de Maestro en Ciencias de la Vida con orientación en Biotecnología Marina

Análisis metagenómico de los microbiomas asociados a esponjas marinas del Parque Nacional Cabo Pulmo (BCS), dirigido a la búsqueda de genes involucrados en la síntesis de policétidos y péptidos no ribosomales con potencial actividad antimicrobiana.

Resumen aprobado por:

---

**Dra. María Asunción Lago Lestón**

Directora de tesis

Las esponjas marinas son animales filtradores, muy antiguos que poseen relaciones simbióticas con distintas comunidades de microorganismos, las cuales pueden ser de alta (HMA) o baja (LMA) abundancia microbiana dependiendo de la especie de esponja. Una consecuencia interesante de esta asociación es la producción de metabolitos secundarios con potencial bioactivo, como los policétidos (PK) y péptidos no ribosomales (NRP). Estos compuestos constituyen el grupo más grande de productos naturales de origen microbiano con reportada actividad antibiótica y anticancerígena. La manera en que estos compuestos son biosintetizados ofrece grandes ventajas para la biotecnología, por lo que su búsqueda es común. Una forma innovadora de explorar variedades de estos compuestos, es a través del estudio metagenómico de esponjas, que consiste en secuenciar todo el ADN de sus comunidades microbianas y buscar *in silico*, grupos de genes biosintéticos (*BGC*, por sus siglas en inglés) que codifiquen para las enzimas (sintasas): PKS y NRPS. En el presente trabajo, se analizaron los microbiomas de 3 especies de esponja, agua y sedimento recolectadas de 4 sitios con profundidades distintas del arrecife del Parque Nacional de Cabo Pulmo, B.C.S, en el mes de abril 2016. Las esponjas fueron identificadas mediante el gen de la subunidad I de la citocromo oxidasa mitocondrial (COI), como: *Aplysina revillagigedi*, *Cliona celata* y *Prosuberites laughlini*. La caracterización de sus microbiomas se realizó por medio de la amplificación de la región V4 del 16S del ARNr, utilizando tecnología de secuenciación Illumina y la paquetería de análisis bioinformáticos de microbiomas *QIIME*. Se realizó un análisis funcional de los microbiomas de esponjas basado en inferencias filogenéticas utilizando, *PICRUSt*, como primer acercamiento para la identificación de los microbiomas con mayor contenido de genes involucrados en la producción de compuestos bioactivos, en particular PKs y NRPs. Basándonos en los resultados de este análisis, así como en la biodiversidad y la singularidad de su comunidad microbiana, se seleccionó a la especie *A. revillagigedi*, para secuenciar por *shotgun* su metagenoma y realizar una búsqueda de *BGCs* de PKS y NRPS en las siguientes bases de datos especializadas: *antiSMASH* y *NaPDoS*. Se detectaron 5 *BGCs* involucrados en la síntesis de PKS tipo 1 por *antiSMASH*; 3 de los cuales presentaron dominios asociados a la producción de 2 compuestos anticancerígenos: Epotilona y Fumonisina. A la par, se realizó un ensayo colorimétrico utilizando el extracto crudo de la esponja para evaluar su potencial antibiótico contra *Mycobacterium tuberculosis*. El experimento se llevó a cabo contra dos cepas: H37Rv, una cepa prototipo de esta bacteria, y 1820/16, una cepa multidrogo resistente (MDR), utilizando 2 concentraciones del extracto crudo: 400 µg/mL y 200 µg/mL. Ambas concentraciones tuvieron un efecto significativo sobre el crecimiento de la cepa H37Rv.

Palabras clave: Esponjas, Microbiomas, Metagenómica, PKS, NRPS, Actividad Antibiótica, Bioprospección

Abstract of the thesis presented by **Fernanda Isabel Ariza Pérez** as a partial requirement to obtain the Master Science degree in Life Sciences with orientation in Marine Biotechnology

**Metagenomic Analyses of Microbiomes Associated to Marine Sponges from Cabo Pulmo National Park, (BCS), in Search of Polyketides and Nonribosomal Peptides with Potential Antibiotic Activity**

Abstract approved by:

---

Dra. María Asunción Lago Lestón  
Directora de tesis

Marine sponges are ancient filter-feeder animals that possess symbiotic relationships with distinct communities of microorganisms, which can be of high or low microbial abundance depending on the sponge species. An interesting outcome of this association is the production of secondary metabolites with bioactive potential, such as polyketides (PK) and non-ribosomal peptides (NRP). These compounds constitute the largest group of natural products of microbial origin, with known antibiotic and anticancerous activity. They are highly sought after because their biosynthesis represents a biotechnological advantage. An innovative way to explore varieties of these compounds is through the metagenomic study of sponges, which consists of sequencing all the genetic material present in their microbiome, to search (by *in silico* means) for clusters of biosynthetic genes (BGC) that codify for the enzymes (synthases) that produce these compounds, namely: PKS and NRPS. In the present study, the microbiomes of three species of sponges, water and sediment were analyzed. These were collected from 4 sites with different depths in the coral reef of The National Park of Cabo Pulmo, BCS, in April of 2016. The sponges were identified through the mitochondrial cytochrome oxidase c subunit I gene (COI), as *Aplysina revillagigedi*, *Cliona celata* and *Prosuberites laughlini*. Their microbiomes were characterized through the amplification of the V4 region of the 16S rRNA, using Illumina sequencing technology and QIIME, a software used for the bioinformatic analysis of microbiomes. The sponge microbiomes were functionally analyzed based on phylogenetic inferences, through PICRUSt. This was done in order to identify the microbiome with the greatest quantity of genes involved in the production of bioactive compounds, such as PKs and NRP. Based on the results of this analysis, as well as the biodiversity and singularity of each sponges microbial community, the sponge *Aplysina revillagigedi* was selected to sequence its complete metagenome (shotgun metagenome) in search for BGC of PKS and NRPS in the following specialized databases: antiSMASH and NaPDoS. Five BGCs involved in the synthesis of type I PKS were detected through antiSMASH; 3 of which had domains associated to the production of the following two anticancerous compounds: Epotilone and Fumonisin. Parallel to the aforementioned studies, a colorimetric assay was performed on the crude extract of the sponge to assess its antibiotic potential against *Mycobacterium tuberculosis*. The experiment was developed against 2 strains of this bacterium: the prototype strain, H37Rv and a multidrug resistant strain, known as 1820/16. The crude extract was evaluated using the following concentrations: 400 µg/mL and 200 µg/mL. These had a significant effect on the growth of H37Rv.

Keywords: Sponges, Microbiomes, Metagenomics, PKS, NRPS, Antibiotic Activity, Bioprospection.

## **Dedicatoria**

A mi familia, por todo su apoyo y paciencia.

## Agradecimientos

Quisiera agradecerle al CONACYT por haberme otorgado los fondos necesarios para estudiar el posgrado en el CICESE. Así también a la dirección de posgrados y en particular a la Dra. Rufina Hernández Martínez, cuyo apoyo fue imprescindible para poder acudir a un congreso internacional de ciencia y tecnología en Noruega.

A la Dra. María Asunción Lago Lestón por enseñarme a cuestionar y ser más crítica de mi trabajo, por toda su comprensión a lo largo de este proceso, y por todas las oportunidades y el apoyo que me dio para viajar e ir a congresos para presentar este trabajo.

A la Dra. Jennifer Chong, postdoc en nuestro grupo de trabajo, por representar un ejemplo de tolerancia, organización y disciplina para mí. Gracias por todos tus consejos y apoyo en el lab.

A mis sinodales: Dra. Johanna Bernaldez, Dr. Carlos Alberto Brizuela y Dra. Alejandra Prieto-Davo, por toda la flexibilidad y comprensión que mostraron hacia conmigo, así como por su dedicación en las correcciones de esta tesis.

Al Dr. Dante Magdaleno, por que siempre estaba ocupado y aún así se tomaba el tiempo para aclararme las mil dudas que me surgían en el día o en la noche.

Al Lic. Maxime Barrault, por apoyarme con la identificación molecular de esponjas, y hacerme compañía en los largos días de extracciones de ADN y PCRs.

A la MC. Elizabeth Brassea, quién se convirtió en mi mayor fuente de motivación en la recta final, gracias por ser una excelente amiga, que se preocupa y me consiente cuando más lo necesito, te quiero mucho!

A la Lic. Arlette Pacheco, por su genuino compañerismo y amistad, gracias por tener siempre una actitud servicial hacia conmigo.

Al Lic. German el bioinformático introvertido del departamento de telemática, quien se convirtió en mi maestro y mi amigo después de una larga sesión de ensamblaje de un metagenoma. Gracias por todo lo que me enseñaste, no lo hubiera podido hacer sin ti!

A la Dra. Johanna Bernaldez, por tener siempre una disposición amable y gentil conmigo. Gracias por todos sus consejos y apoyo en el laboratorio de patógenos.

A el MC. Duahmet Ruiz Gueraca, por su compañerismo al momento de elaborar mis bioensayos, gracias por compartirme todo tus tips!

A mi esposo Luis, por encargarte de la casa cuando me ausentaba físicamente y mentalmente por estar trabajando. Gracias por tu compañía a distancia y en silencio, gracias por escucharme, y gracias por todos los abrazos, los apapachos y las palabras de aliento que me diste en los momentos más difíciles de crisis y estrés.

A la Lic. Vanessa Vazquez, porque hizo de la maestría un proceso más llevadero. Gracias por todas las carcajadas y buenos ratos.

Al Lic. Eduardo Alvarez (Ed), por ser un excelente compañero, con la mejor disposición para ayudar.

Al Lic. Victor Vargas, por ser un excelente amigo, y compañero, siempre dispuesto a escuchar y a dar buenos consejos.

## Tabla de contenido

	Página
Resumen en español.....	ii
Resumen en inglés.....	iii
Dedicatorias.....	iv
Agradecimientos.....	v
Tabla de contenido.....	vi
Lista de figuras.....	ix
Lista de tablas.....	xii
<b>Capítulo 1. Introducción.....</b>	<b>1</b>
1.1 Esponjas: Propiedades Generales.....	1
1.2 Esponjas como holobiontes.....	4
1.3 Origen microbiano de compuestos bioactivos en esponjas.....	7
1.4 Potencial antibiótico de PKS y NRPS.....	8
1.5 Síntesis biológica de PKS y NRPS.....	9
1.6 Análisis metagenómico de esponjas marinas.....	12
1.7 Área de estudio.....	14
1.8 Justificación .....	16
1.9 Objetivos.....	16
1.9.1 Objetivo general.....	16
1.9.2 Objetivos específicos.....	16
<b>Capítulo 2. Metodología.....</b>	<b>18</b>
2.1 Colecta de organismos.....	18
2.2 Identificación molecular de esponjas utilizando la subunidad I de la citocromo oxidasa mitocondrial (COI).....	19
2.2.1 Extracción de ADN de esponjas marinas.....	19
2.2.2 Identificación molecular.....	20
2.3 Preparación de librerías 16S del ARNr para secuenciar por Illumina MiSeq System.....	20
2.3.1 Amplificación de la región V4 del gen 16S del ARNr.....	20
2.3.2 Preparación de metalibrerías.....	21
2.4 Análisis bioinformáticos.....	22
2.4.1 QIIME: <i>Quantitative Insights into Microbial Ecology</i> .....	22

2.4.2 PICRUSt: <i>Phylogenetic Investigation of Communities by Reconstruction of Unobserved States</i> .....	24
2.4.3 MG-RAST: <i>metagenomics Rapid Annotation using Subsystem Technology</i> .....	24
2.4.4 antiSMASH: <i>antibiotics &amp; Secondary Metabolite Analysis Shell</i> .....	25
2.4.5 NaPDos: <i>Natural Product Domain Seeker</i> .....	26
2.5 Generación de librería shotgun para secuenciar el microbioma de <i>Aplysina revillagigedi</i> .	27
2.5.1 Separación celular.....	27
2.5.2 Extracción de ADN.....	28
2.5.3 Librería de secuenciación <i>shotgun</i> .....	28
2.6 Evaluación de la actividad antibiótica de extractos crudos de <i>Aplysina revillagigedi</i> contra <i>M. tuberculosis</i> .....	29
2.6.1 Preparación de extracto crudo.....	30
2.6.2 Ensayo colorimétrico.....	31
2.7 Análisis Estadísticos.....	33
2.7.1 Variabilidad entre estaciones.....	33
2.7.2 Variabilidad entre réplicas biológicas.....	33
2.7.3 Efecto del extracto crudo de <i>Aplysina revillagigedi</i> sobre <i>M. tuberculosis</i> .....	34
2.7.4 Análisis comparativo de microbiomas entre esponjas y su ambiente.....	34
<b>Capítulo 3. Resultados</b> .....	36
3.1 Colecta de esponjas.....	36
3.2 Identificación molecular de esponjas.....	37
3.3 Análisis de los microbiomas mediante amplificación del gel 16S del ARNr de muestras de sedimento, agua y esponjas.....	50
3.3.1 Amplicones de la SEQ1.....	50
3.3.2 Amplicones de la SEQ2.....	52
3.4 Resultados de secuenciación.....	54
3.5 Análisis ecológicos.....	56
3.6 Variabilidad entre réplicas.....	61
3.7 Variabilidad entre estaciones.....	63
3.8 OTUs especie-específicos de esponjas (diagramas de Venn) .....	64
3.9 Análisis funcional con PICRUSt.....	67
3.10 Separación celular y extracción de ADN del microbioma de <i>Aplysina revillagigedi</i> para secuenciación <i>shotgun</i> .....	69

3.11 Análisis metagenómico del microbioma de <i>Aplysina revillagigedi</i> por mgRAST.....	70
3.12 Resultados antiSMASH y NaPDoS.....	72
3.13 Bioensayos.....	74
<b>Capítulo 4 . Discusión.....</b>	<b>76</b>
4.1 Identificación molecular de esponjas marinas provenientes del Parque Nacional Cabo Pulmo (BCS).....	76
4.2 Caracterización y comparación del microbioma de 3 especies de esponjas abundantes en el PNCP.....	77
4.3 Análisis del metagenoma de <i>Aplysina revillagigedi</i> : búsqueda de PKS y NRPS.....	80
4.4 Bioensayos.....	82
<b>Capítulo 5 . Conclusiones.....</b>	<b>84</b>
5.1 Identificación de especies.....	84
5.2 Caracterización del microbioma de las tres especies mas abundantes del PNCP.....	84
5.3 Análisis del metagenoma de <i>Aplysina revillagigedi</i> .....	84
5.4 Bioensayos.....	85
<b>Literatura citada.....</b>	<b>86</b>
<b>Anexos.....</b>	<b>95</b>

## Lista de figuras

Figura		página
1.	Micrografía electrónica de un fósil de esponja.....	1
2	Diferentes formas de crecimiento en las esponjas marinas.....	3
3	Interacciones entre huésped-microbios.....	5
4	Evidencia de alto contenido de compuestos de origen microbiano en esponjas HMA.....	8
5	Organización estructural y bloques constructivos de PKSs y NRPSs.....	9
6	Mecanismo general PKS.....	11
7	Estructura modular de un NRPS.....	11
8	Mapa del Parque Nacional de Cabo Pulmo.....	15
9	Diseño y secuencia de los primers utilizados para amplificar la región V4 del gen 16S ARNr.....	21
10	Aspecto del extracto crudo de <i>Aplysina revillagigedi</i> posterior a su evaporación.....	30
11	Diseño de placa utilizada en bioensayos colorimétricos contra <i>M. tuberculosis</i> .....	32
12	Registro fotográfico de los ejemplares de esponja más abundantes encontradas en el PNCP...	36
13	Registro fotográfico de las especies de esponja marina menos abundantes del PNCP.....	37
14	Comparación fotográfica entre 3 especies de esponjas candidatas para la identificación taxonómica de del espécimen 306.....	39
15	Alineamiento de la secuencia del espécimen ID: 301, contra <i>Aplysina fistularis</i> .....	40
16	Alineamiento de la secuencia del espécimen ID:301 con <i>Aplysina gerardogreeni</i> .....	41
17	Alineamiento de la secuencia del espécimen ID:301 con <i>Aplysina revillagigedi</i> .....	42
18	Resultados de <i>Blast 2 sequences</i> entre espécimen ID:301 y <i>Aplysina revillagigedi</i> .....	43
19	Registro fotográfico de el espécimen ID: 306, <i>Aplysina revillagigedi</i> y su mapa de distribución en el mundo.....	44
20	Registro fotográfico de el espécimen ID: 306, <i>Prosuberites laughlini</i> y su mapa de distribución en el mundo.....	45
21	Registro fotográfico de el espécimen ID: 306, <i>Cliona celata</i> y su mapa de distribución en el mundo.....	46
22	Resultados de la extracción del ADN de las 4 especies y 3 géneros de esponja identificados....	48

23	Comparación entre homogenización manual contra homogenización mecánica del espécimen ID: 404, identificado como <i>Aplysina revillagigedi</i> .....	49
24	Gel de electroforesis del producto de amplificación del gen COI de especímenes de esponjas.	49
25	Gel de electroforesis con los productos PCR de esponjas secuenciadas en la SEQ1.....	51
26	Gel de electroforesis con los productos PCR del ambiente secuenciadas en la SEQ1.....	51
27	Gel de electroforesis con los productos PCR de las muestras de esponja secuenciadas en la SEQ2.....	53
28	Gel de electroforesis con los productos PCR de las muestras de ambiente secuenciadas en la SEQ2.....	54
29	Curvas de rarefacción de el número de OTUs observados para todas las muestras secuenciadas.....	55
30	Abundancia relativa a nivel filo ente esponjas y ambiente.....	57
31	Porcentaje de los 5 fijos más abundantes en cada especie de esponja.....	58
32	Comparación entre el índice de Chao1 y el número de OTUs observados en <i>Prosuberites laughlini</i> , <i>Aplysina revillagigedi</i> y <i>Cliona celata</i> .....	59
33	Índice de biodiversidad de Shannon para: <i>Prosuberites laughlini</i> , <i>Aplysina revillagigedi</i> , y <i>Cliona celata</i> .....	60
34	PCoA 2D de datos ponderados de beta diversidad.....	61
35	Variabilidad individual entre muestras.....	62
36	Variabilidad entre estaciones.....	64
37	Comparación entre el número de OTUs filtrado presentes en cada especie de esponja, contra el número de OTUs filtrado presentes en el ambiente.....	65
38	Número de OTUs especie-específicos.....	67
39	Potencial bioactivo de los microbioma de cada especie de esponja.....	68
40	Extracción enriquecida con ADN de procariota para secuenciación <i>shotgun</i> .....	70
41	Resultados MG-RAST de la distribución de categorías taxonómicas a nivel filo utilizando la base de datos REFseq.....	71
42	Número de lecturas anotadas por bases de datos utilizadas por mgRAST.....	71
43	Perfil de categorías funcionales según la base de datos: <i>Subsystems</i> .....	72
44	Muestra los resultados de la búsqueda de metabolitos secundarios en antiSMASH.....	73

45	Efecto inhibitorio del extracto crudo de <i>Aplysina revillagigedi</i> , sobre dos cepas de TB: (A) H37Rv y (B) 1820/16 MDR.....	75
46	Alineamiento blastn de la secuencia consenso de <i>Prosuberites laughlini</i> , muestra 303 de la estación Marios reef.....	95
47	Alineamiento blastn de la secuencia consenso de <i>Aplysina revillagigedi</i> , muestra 301 de la estación Marios Reef.....	96
48	Alineamiento por blastn de la secuencia consenso de <i>Timea</i> sp, muestra 302 de Marios Reef.....	97
49	Alineamiento por blastn de la secuencia consenso de la muestra 212, de la estación El Cantil.....	98
50	Alineamiento por blastn de la secuencia consenso de <i>Clathria prolifera</i> , de la muestra 407 de La Esperanza.....	99
51	Alineamiento por blastn de la secuencia consenso de <i>Phelloderma</i> sp, de la muestra 408, que proviene de La Esperanza.....	100
52	Alineamiento por blastn de la secuencia consenso de la esponja <i>Cliona celata</i> , proveniente de la estación Las Navajas (ID: 504). ....	101

## Lista de tablas

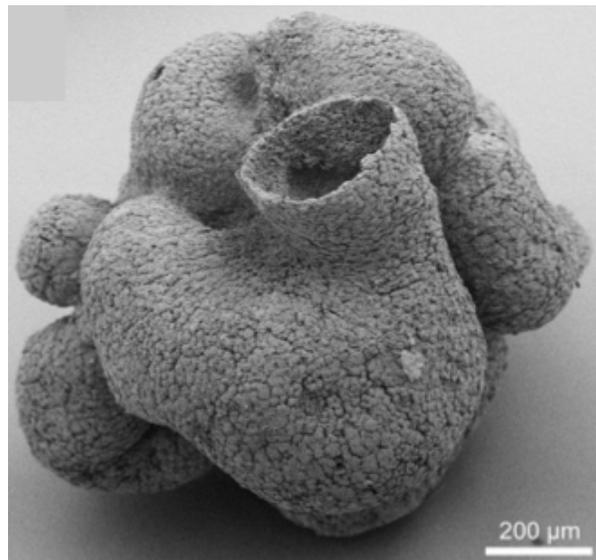
Tabla	página
1 Resultados de análisis BLAST de secuencias consenso de cada especie identificada de esponja.....	38
2 Información de secuenciación relevante de la SEQ2.....	55
3 Muestra el promedio de OTUs observados y su desviación estándar, de cada réplica de esponja ( <i>Aplysina revillagigedi</i> , <i>Cliona celata</i> y <i>Prosuberites laughlini</i> ), sedimento y agua.....	63
4 Número de OTUs totales vs filtrados para sedimento, agua y esponjas.....	65
5 Lista de los 10 OTUs exclusivos de cada esponja más dominantes (en orden descendente) de: <i>Aplysina revillagigedi</i> , <i>Cliona celata</i> y <i>Prosuberites laughlini</i> .....	66
6 Promedio de el enriquecimiento de genes para tipos de terpenoides y policétidos presentes en las réplicas de cada esponja.....	69
7 Muestra los resultados de la búsqueda por NaPDos, de los productos moleculares que se asocian a los dominios anotados que estuvieron presentes en los clusters BGCs involucrados en la síntesis de PKS encontrados por antiSMASH.....	74
8 Información relevante a la calidad de cada esfuerzo de secuenciación realizado en este trabajo.....	102
9 Resultados de secuenciones: SEQ1 y SEQ2 de las réplicas biológicas de esponja, sedimento y agua secuenciada, así como información relevante acerca de la profundidad de secuenciación. ....	102

## Capítulo 1. Introducción

---

### 1.1 Esponjas: Propiedades generales

Las esponjas (Filo *Porifera*) son metazoarios acuáticos, sésiles, muy antiguos, que datan desde hace 600 millones de años, según sugieren relojes moleculares calculados para la divergencia filogenética de eumetazoos, así como evidencia fósil (Fig.1) encontrada en un yacimiento de fosfatos al Sur de China (Formación Doushantuo) por Li *et al.*, (1998) (Yin *et al.*, 2015; Peterson *et al.*, 2005; Muscente *et al.*, 2015).



**Figura 1.** Micrografía electrónica de un fósil posiblemente porífero, encontrado en la Formación Doushantuo. Las dimensiones del espécimen fueron de 1.2mm de ancho por 1.1mm de largo, en la imagen se puede distinguir claramente una estructura tubular que asemeja el “ósculo” porífera, o apertura principal por donde sale el agua que es filtrada por la esponja (modificada de Yin *et al.*, 2015).

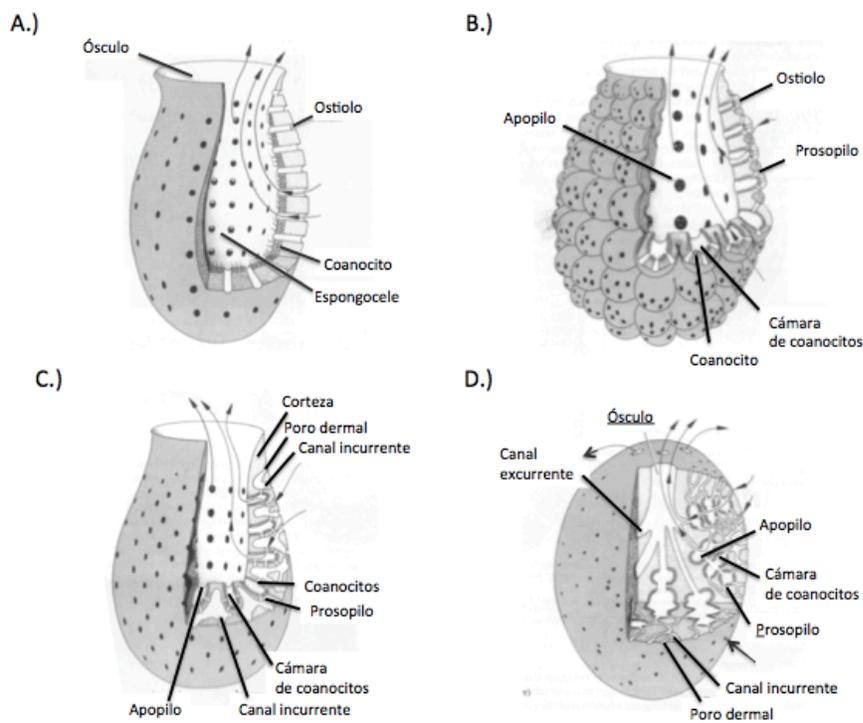
Su estatus primitivo, ha hecho que las esponjas sean animales versátiles para colonizar el bentos marino, pues se les puede encontrar viviendo a cualquier latitud o profundidad de ambientes de agua tanto salada como dulce. Aunque sus sitios de preferencia son las zonas tropicales, como los arrecifes caribeños y los litorales no contaminados (Brusca y Brusca, 2003). Ecológicamente, desempeñan un rol muy importante en los nichos que habitan, como: guarida y alimento para otros organismos bentónicos, agentes preventivos de la erosión de los arrecifes y animales participes en procesos biogeoquímicos. Por ejemplo: el ciclo del nitrógeno, carbono y sílice (Díaz *et al.*, 2001; Gatti *et al.*, 2002; Fiore *et al.*, 2013).

Fisiológicamente, las células de las esponjas marinas tienen la capacidad de re-agregarse para formar una esponja funcional, después de haber sido desasociadas. Esta particularidad, se descubrió desde principios del siglo XX, por H.V Wilson, un profesor de biología de la Universidad de North Carolina, en Chapel Hill, USA, y ha servido como base para el descubrimiento de moléculas específicas, involucradas en la adhesión celular (ej. cadherinas y N-CAM) (Harris, 2017; Costello, 1961). Otra peculiaridad de las esponjas es que poseen células totipotentes, que pueden cambiar de función según las necesidades de la esponja. Entre las funciones que pueden desempeñar están las siguientes: actividad fagocítica (realizada por arqueocitos), actividad reproductiva (oocitos y espermatozoides), porocitos (células en forma de canal), lofocitos (células encargadas de generar espongina o fibras de colágeno, material que constituye gran volumen de la esponja) y esclerocitos (células encargadas de generar espículas que conforman el esqueleto porifera) (Brusca y Brusca, 2003).

Anatómicamente, se describen a través de tres capas: el pinacodermo (capa externa), el mesohilo (capa media) y el coanodermo o espongocele (capa interna). El pinacodermo, se compone por células en forma de placa conocidas como pinacocitos. Esta capa presenta series de orificios pequeños, denominados "ostiolos", los cuales se pueden abrir o cerrar como válvulas para dar paso al flujo de agua que la esponja filtrará para alimentarse. Por otra parte, la capa media o mesohilo, es un espacio interno entre el pinacodermo y el coanodermo que está compuesto por espongina (fibras de colágeno) y espículas (estructuras en forma de aguja compuestas por sílice o carbonato de calcio), las cuales generan una matriz en la que se encuentran flotando células propias de la esponja, así como de microorganismos simbiotes de la misma. Por su parte, el coanodermo, está compuesto por células flageladas conocidas como coanocitos, las cuales le sirven a la esponja para bombear agua a su interior, y establecer un sistema de flujo (Taylor *et al.*, 2007; Brusca y Brusca, 2003). La complejidad de este sistema de flujo depende de la organización celular de la esponja (Brusca y Brusca, 2003).

Existen 3 formas de organización celular: asconoide, siconoide y leuconoide (Fig.2). Las formas asconoide y siconoide son las formas más simples de crecimiento. Se caracterizan por poseer un pinacodermo y coanodermo compuesto de una sola capa de células, las cuales forman la estructura de un tubo ya sea de manera continua (asconoide), o plegada (siconoide). Su mesohilo no abarca mucho volumen de tal manera que el flujo del agua es directo a través de los porocitos. Por el contrario, la forma leuconoide es más compleja, en este tipo de crecimiento el mesohilo es más grueso, y los coanocitos se organizan para formar compartimentos interconectados entre sí, por los cuales se hará pasar el agua y se capturarán partículas pequeñas para transferirlas al mesohilo donde se encuentran los arqueocitos, que son células encargadas de la digestión de la esponja, entre otras cosas (Taylor *et al.*,

2007; Brusca y Brusca, 2003). Dependiendo de la clase de esponja que se trate presentarán las distintas formas de crecimiento, y estas se clasifican según el tipo de esqueleto que presenten (Brusca y Brusca, 2003). Hay cuatro clases, estas son: Calcárea que son esponjas con un esqueleto compuesto por espículas de calcio, esta clase suele presentar las formas asconoides y siconoides; Hexactinélida, también conocidas como esponjas vítreas por poseer espículas síliceas, pueden presentar las formas siconoides o leuconoides; Demospongiae, que pueden presentar espículas de sílice o una mezcla de estas con espongina (fibras de colágeno) conformando una organización leuconoide; y Homoscleromorfa, esponjas con un esqueleto variable (con o sin espículas de sílice), de crecimiento leuconoide (Gazave *et al.*, 2011; CONABIO, 2012).



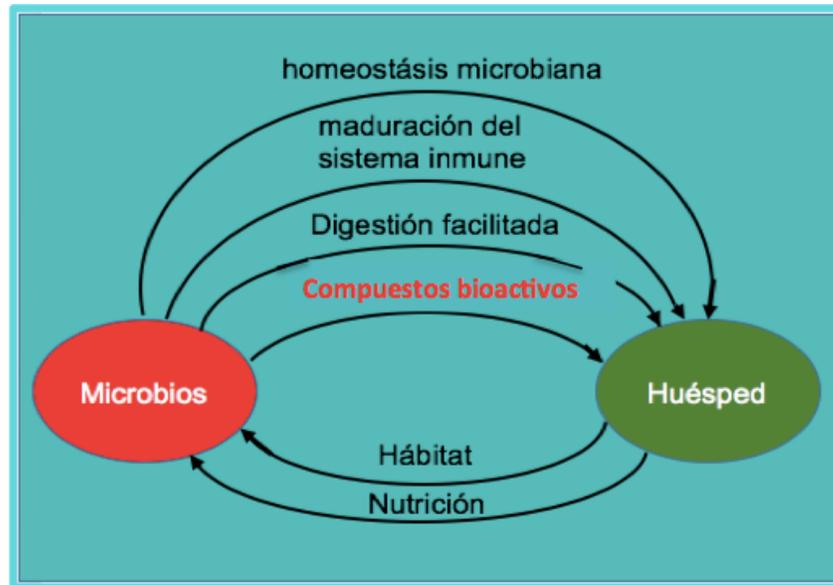
**Figura 2.** Diferentes formas de crecimiento en las esponjas, así como los tipos de sistemas de flujo de agua y componentes asociados a cada morfología. A.) asconoide; B.) siconoide; C.) variación de siconoide más compleja (presencia de corteza); D.) leuconoide (tomado de Brusca y Brusca, 2003).

A la fecha se reportan 8,845 especies de esponja, siendo la mayoría de ellas pertenecientes a la clase Demospongiae (Van Soest *et al.*, 2017). Aunque, se estima una diversidad real de 15,000 especies (Webster y Thomas, 2016). La identificación taxonómica tradicional de esponjas se basa en la forma, el tamaño y el número de espículas presentes en su tejido (Uriz, 1986). El problema con este método es

que no todas las esponjas poseen un esqueleto de espículas, ej. la clase Homoscleromorpha, que se identificó más recientemente por medio de métodos moleculares (Morrow & Cárdenas, 2015). Otra limitante de este método, es que las esponjas presentan una alta plasticidad ecológica, lo que significa que dos especies diferentes pueden parecer la misma, si se desarrollaron en el mismo ambiente. Por lo anterior, las pruebas moleculares, que utilizan regiones genéticas conservadas, con suficiente variabilidad como para distinguir esponjas a nivel especie, son cada vez más utilizadas en conjunto con el método tradicional (Shearer *et al.*, 2002). La región genética más utilizada es la subunidad I de la citocromo oxidasa mitocondrial (COI), para la cual existen una variedad de opciones de *primers* (oligonucleótidos utilizados para amplificar regiones específicas del ADN) reportados para amplificar esta región con diversos grados de efectividad dependiendo de la especie bajo estudio (Palaeontologie.geo.uni-muenchen.de, 2017). Los únicos factores que pueden dificultar el análisis adecuado de los amplicones generados son, la coextracción de ADN proveniente de macro o microorganismos viviendo en las esponjas, y la presencia de metabolitos secundarios que puedan inhibir las reacciones enzimáticas que se llevan a cabo al momento de amplificar, ya que las esponjas son conocidas por poseer gran cantidad de este tipo de compuestos. También está el limitante de las bases de datos utilizadas para la identificación filogenética, pues continúan en proceso de crecimiento y depuración. Por lo anterior, la combinación de técnicas moleculares en conjunto con el método clásico de clasificación, es la mejor forma de aproximarse a una identificación taxonómica de esponjas más confiable (Vargas *et al.*, 2012).

## 1.2 Esponjas como holobiontes

Las esponjas marinas son los organismos que albergan la mayor cantidad de microbios viviendo en simbiosis con ellas (sus microbiomas pueden ocupar hasta el 50% de su biomasa (Wang, 2006)), por lo que se les considera como un organismo holobionte. Un holobionte es una unidad funcional compuesta por el huésped (la esponja) y las comunidades microbianas que residen en el (Webster y Thomas, 2016). Los microorganismos que típicamente componen las comunidades simbiotes de esponjas son: bacterias, arqueas, microalgas, y hongos (Thomas *et al.*, 2010). Todos estos pueden estar influenciando a su huésped en diferentes aspectos fisiológicos (Fig. 3).



**Figura 3.** Interacciones entre huésped (esponjas) y microbios. En rojo se resalta la producción de compuestos bioactivos que las esponjas pueden utilizar para su defensa contra depredadores (Fraune y Bosch, 2010; Ryu *et al.*, 2016) (modificado de Fraune y Bosch 2010).

La carga microbiana que posean las esponjas se correlaciona con el tamaño y la eficiencia de su sistema de flujo, siendo aquellas de mayor tamaño y con un sistema de flujo pobre, las que tienden a presentar los números más altos de bacterias (Kennedy *et al.*, 2007; Wang, 2006). Las esponjas se denominan según su carga microbiana como esponjas con alta/baja abundancia microbiana o *High/Low Microbial Abundance* (HMA y LMA, respectivamente por sus siglas en inglés) (Ryu *et al.*, 2016). Otra manera para referirse a esponjas HMA es como bacterio-esponjas, estas esponjas pueden albergar entre  $10^8$ - $10^{10}$  células microbianas por gramo de peso húmedo de esponja (Hentschel *et al.*, 2006). La clase Demospongiae, posee la mayoría de las esponjas HMA. Por su parte, las especies LMA se caracterizan por presentar abundancias microbianas similares a las del agua circundante ( $10^5$ - $10^6$  células de microbios por gramo de peso húmedo de esponja). Esta similitud también se extiende en cuanto a la riqueza de especies microbianas, según lo sugieren los resultados de Schmitt *et al.*, 2012, quienes compararon esponjas LMA y HMA de la misma localidad, respecto al filo Chloroflexi, y se dieron cuenta que las especies LMA contenían especies de Chloroflexi presentes en el agua circundante, mientras que las esponjas HMA albergaban una comunidad de Chloroflexi diferente y menos variable que la primera. La distribución de los microorganismos en la esponja, varía dependiendo del tipo de metabolismo que estos posean. Así, por ejemplo, si son organismos fotolitoautótrofos, se encontrarán habitando el pinacodermo, donde pueden estar influenciando en aspectos como el color de la esponja (Rutzler, 1985;

Wilkinson, 1992). Por otro lado, los heterotróficos y otros autotróficos se concentran en el mesohilo o capa intermedia, donde pueden estar ayudando a la esponja en la producción de sustancias bioactivas como defensa y en la digestión (Hentschel *et al.*, 2003).

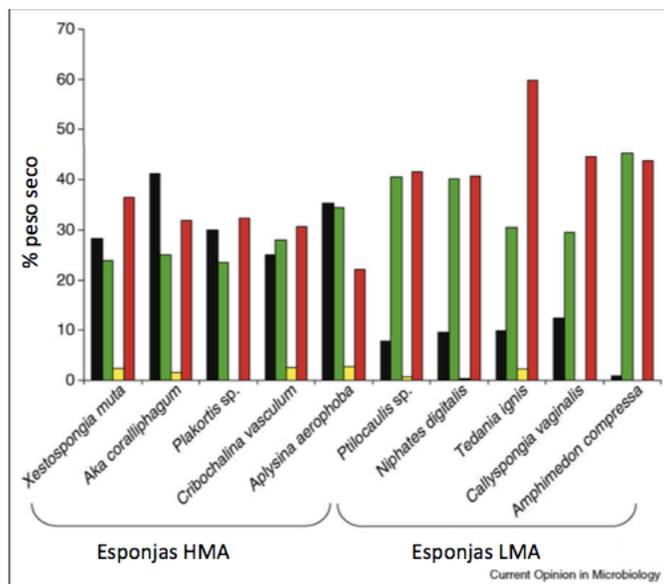
Otros factores que influyen en la composición y estructura de las comunidades simbiotes a través del tiempo y el espacio son los siguientes: nutrientes producidos por el huésped, su pH, y su respuesta inmune (Thomas *et al.*, 2016), sin embargo, los factores ecológicos y evolutivos que posiblemente moldean estas comunidades aún son grandemente desconocidos (Hacquard *et al.*, 2015). La manera en que las esponjas adquieren estos microbios se puede explicar en parte por la transmisión vertical dada por vía reproductiva (Schmitt *et al.*, 2007; Sharp *et al.*, 2007; Schmitt *et al.*, 2008; Lee *et al.*, 2009), así como transmisión horizontal, donde algunas bacterias que originalmente se encuentran en el medio ambiente logran escapar la digestión de la esponja y asentarse en ella (Wehrl *et al.*, 2007).

Se han reportado un total de 52 fila de organismos procariotas asociados a esponjas (Webster y Thomas, 2016). Este número incluye también a fila candidatos, los cuales forman parte de una comunidad nombrada “rara”, que se presenta en bajas abundancias en las esponjas (Reveillaud *et al.*, 2014). En un estudio reciente elaborado por Thomas *et al.*, (2016), sobre la composición de las comunidades microbianas asociadas a 81 especies de esponjas provenientes de 20 países que colindaban con los Océanos: Pacífico, Atlántico, e Índico, se encontró que el filo simbiote más dominante reportado para esponjas es el de las Proteobacterias (en particular, alfa y gamma); Chloroflexi, Cianobacteria y Crenarchaeota también pueden llegar a presentarse con abundancias relativas altas (10%). Los perfiles microbianos encontrados se caracterizaron por una combinación de especies de microorganismos generalistas y especialistas. Los generalistas se definen en el artículo como taxa cosmopolita que se encuentra en diferentes especies de esponjas, generalmente formando parte de lo que se conoce como el microbioma central o núcleo que puede variar de tamaño (entre 7-20 especies). Por su parte, los especialistas se definen como taxa que se encuentra en una o pocas especies de esponjas. Se encontró que muchos de estos simbiotes eran exclusivos a esponjas, por lo anterior se sugirió que los microbiomas de esponjas marinas son una fuente global significativa de diversidad microbiana única.

### 1.3 Origen microbiano de compuestos bioactivos en esponjas marinas

El filo Porifera es considerado el taxon de organismos marinos que produce mayor cantidad de metabolitos secundarios con potencial farmacéutico y existe evidencia que muchos de estos componentes son producidos por los microbios simbiotes que viven en ellas (Thomas *et al.*, 2010; Hentschel *et al.*, 2012; Webster y Thomas, 2016). Por otra parte, las actinobacterias y ascomicetos son los productores dominantes de compuestos con potencial uso terapéutico (Thomas *et al.*, 2010). La evidencia de que muchos de los compuestos bioactivos aislados de esponjas marinas son de origen microbiano, se basa en un parentesco estructural de estos compuestos con productos típicos de vías metabólicas microbianas (Hentschel *et al.*, 2012). Un estudio reciente que mostró evidencia de lo anterior, es el de Della Sala *et al.* (2014), quienes trabajaron con la esponja caribeña *Plakortis halichondrioides* y encontraron que los dominios de sintasas de policétidos tipo I (PKSI - sintasas de metabolitos secundarios altamente bioactivos), provenían de genes de bacterias y protistas que viven en simbiosis con esta esponja.

Previo al estudio anterior, Shirmer *et al.* (2005) presentaron una de las evidencias más convincentes del origen microbiano de policetidos en esponjas. En este trabajo se secuenció el primer grupo de genes PKSI pertenecientes a un clado específico de esponjas, proveniente del metagenoma de la esponja *Discodermia dissoluta*. En sus resultados reportaron que el producto de expresión de éste cluster de genes, denominado *sup*, era un ácido graso metilado. Debido a que los ácidos grasos metilados forman gran parte de la capa lipídica protectora de bacterias como *Mycobacterium tuberculosis*, los autores concluyeron que la presencia de estos compuestos en la esponja se debía a las bacterias que vivían en ella. Esto último fue apoyado por la presencia de un mayor número de ácidos grasos metilados en bacterio-esponjas en comparación a esponjas libres de bacterias, en un estudio donde también se evaluaron otros ácidos grasos presentes en esponjas como: ácidos grasos no saturados, ácidos grasos con ciclopropano y ácidos grasos saturados y de cadena lineal (Fig. 4) (Grozdanov y Hentschel, 2007).

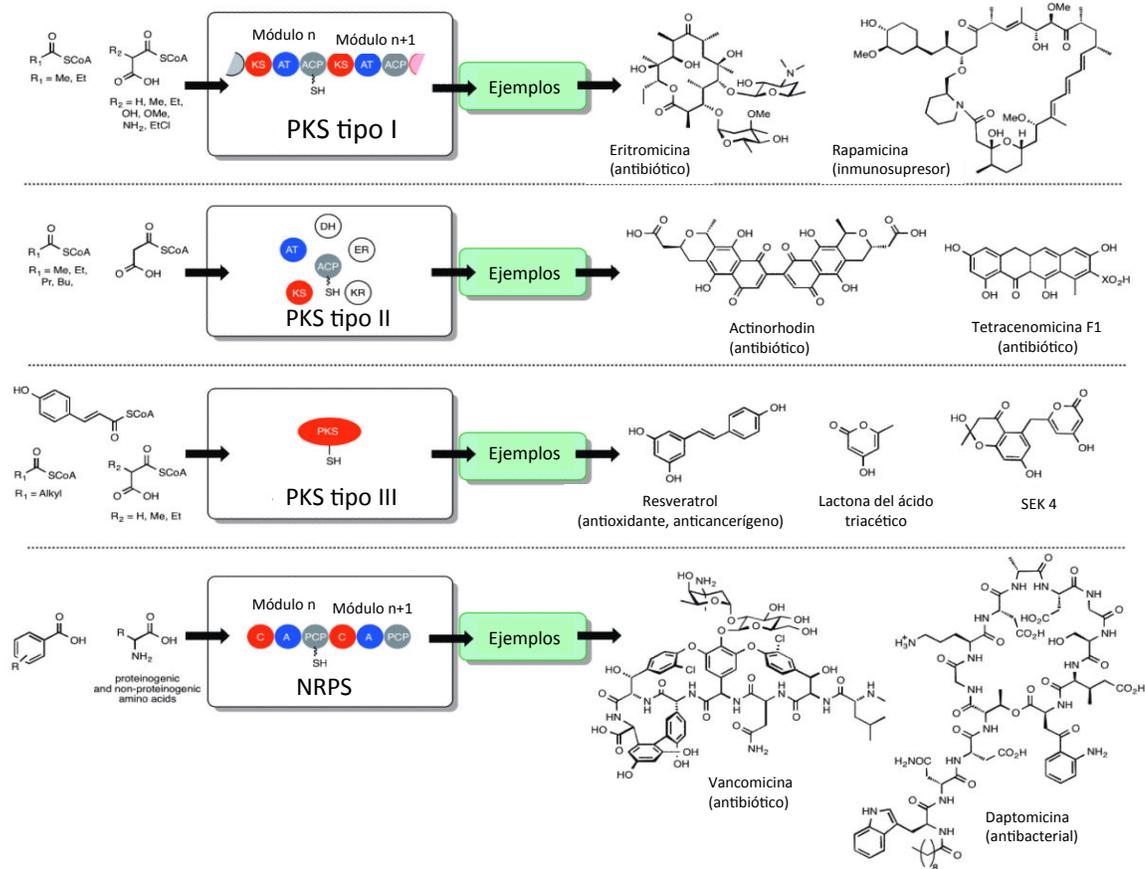


**Figura 4.** Evidencia del alto contenido de compuestos de origen microbiano en esponjas HMA. Las barras negras representan a los ácidos grasos metilados; las barras verdes son ácidos grasos no saturados; las barras amarillas son ácidos grasos con ciclopropano; y las barras rojas son ácidos grasos saturados y de cadena lineal (modificada de Grozdanov y Hentschel, 2007).

#### 1.4 Potencial antibiótico de PKS y NRPS

Actualmente la población mundial se enfrenta a una elevada tasa de resistencia bacteriana a antibióticos (World Health Organization, 2017) dada por la aparición de cepas multidrogoresistentes (MDR) que se definen como aquellas cepas resistentes a antibióticos de primera línea de defensa, tales como: isoniazida (INH, 1952) y rifampicina (RIF, 1966), entre otros (Bernáldez-Sarabia, 2013). La *Mycobacterium tuberculosis* que provoca la tuberculosis (TB) pulmonar es un ejemplo de una bacteria que se ha hecho MDR. En México, la TB pulmonar es un problema grave, especialmente para el Estado de Baja California, el cual presenta el mayor número de muertes por esta razón en el país (Sistema Nacional de Vigilancia Epidemiológica, 2012 en Bernáldez-Sarabia, 2013). Por lo anterior, hay una urgente necesidad de nuevas clases de antibióticos cuyo desarrollo pueda igualar el paso de la resistencia bacteriana (Laport *et al.*, 2009). Frente a la búsqueda de antibióticos alternativos, se encuentran los policétidos (PKs) y péptidos no ribosomales (NRPs), quienes representan dos clases de compuestos naturales complejos que se encuentran en microbios y plantas. Estos metabolitos poseen

diversas actividades farmacológicas, entre ellas: anticancerígena (ej. calicheamicina y bleomicina), inmunosupresora (ej. rapamicina) y antibacterial (ej. eritromicina y vancomicina) (Williams, 2013; Minowa *et al.*, 2007) (Fig.5).



**Figura 5.** Organización estructural y bloques constructivos de PKSs y NRPSs. Se muestra una selección típica de bloques constructivos que ocurren en la naturaleza (modificada de Williams, 2013).

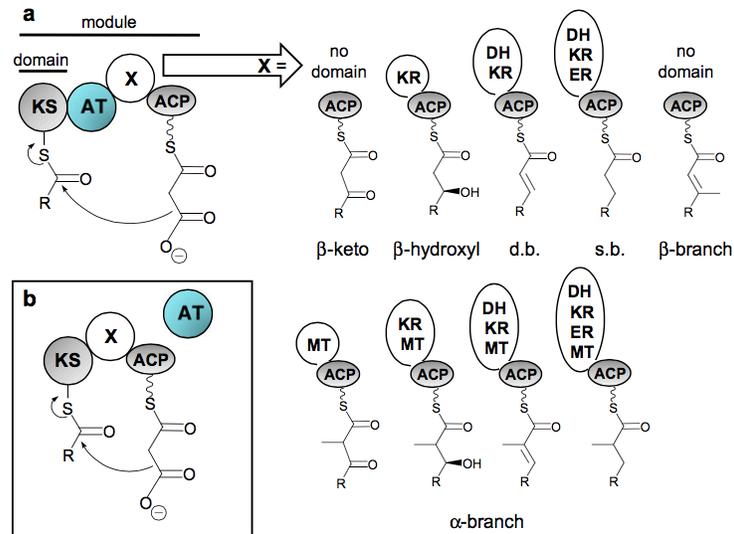
## 1.5 Síntesis biológica de PKS y NRPS

Los policétidos se producen de una manera similar a los ácidos grasos, a través de una condensación de Claisen seguida por una serie de reacciones químicas que dependen del dominio presente (Fig. 6). Pueden ser sintetizados por 3 tipos de sintasas de policétidos conocidas como: PKS I, PKS II y PKS III. Las PKS I se dividen en modulares o iterativas dependiendo del tipo de ensamblaje que

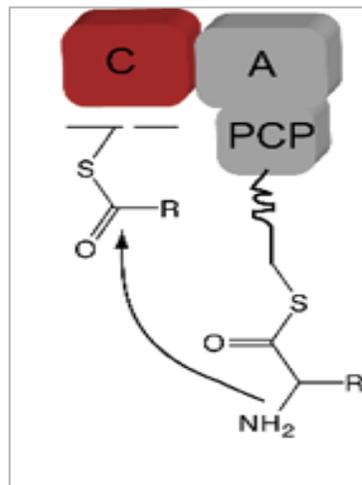
presenten. Las modulares presentan múltiples módulos por los cuales el policétido puede crecer en función a los dominios catalíticos de cada uno, extendiendo así varias cadenas a partir de un sólo policétido. Para las PKS I iterativas, todo el proceso se lleva a cabo en un sólo policétido que es utilizado repetidas veces. Los dominios más básicos presentes en estos módulos son: la proteína acarreadora de acilos (ACP), la cual está unida al policétido creciente e interactúa con la aciltransferasa, (AT). La AT proporciona los monómeros de acilo que constituyen la cadena. También está, el dominio de la cetosintasa (KS) que desempeña una reacción química con la que se une cada monómero con la cadena creciente (Hochmuth y Piel, 2009).

Por su parte, los péptidos no ribosomales, son sintetizados por sintasas de péptidos no ribosomales (NRPS), y tienen una síntesis en serie similar a los policétidos, a excepción de que estos son elongados a partir de un péptido que es sostenido por una proteína cargadora de péptidos (PCP). Esta proteína se une a amino ácidos transportados por el dominio de adenilación, por medio de un dominio de condensación que genera un enlace peptídico entre los primeros dos (Fig.7) (Schwarzer *et al.*, 2003). De todos los dominios catalíticos reportados, los dominios KS y C en policétidos y péptidos no ribosomales, respectivamente, son los más conservados y por ende los más estudiados (Yadav *et al.*, 2009).

Con base en la organización de dominios y modelos de sintasas de policétidos (PKS) y péptidos no ribosomales (NRPS), es posible predecir la estructura química de productos derivados de clusters de genes que codifican para PKS y NRPS (Fischbach y Walsh, 2006). Dado que muchos PKS y NRPs han sido aislados de esponjas marinas, y que existe fuerte evidencia de que son producidos por simbiontes microbios, la secuenciación de los metagenomas de esponja representan una fuente útil para la identificación de los microorganismos productores de estos compuestos naturales. Entre los simbiontes de esponjas descritos como productores de moléculas bioactivas estructuralmente complejas se encuentran los siguientes grupos bacterianos: Proteobacteria, Actinobacteria, Acidobacteria, Cyanobacteria y Enthothionella (perteneciente al filo Tectomicrobia (Nakashima *et al.*, 2016).



**Figura 6.** Mecanismo general de la síntesis de un PK. (a) tipos de dominios y reacciones encadenadas por un módulo PKS cis-AT. (b) módulo PKS trans. Cetosintasa (KS), aciltransferasa (AT), proteína acarreadora de acil (ACP), ketoreductasa (KR), deshidrogenasa (DH), enoilreductasa (ER), metiltransferasas (MT), doble enlace (d.b), enlace sencillo (s.b) (modificada de Hochmuth y Piel, 2009).



**Figura 7.** Estructura modular de un NRP: dominio A (adenilación) trae consigo los amino ácidos; PCP (proteína cargadora de péptidos) sostiene al péptido creciente como un tioéster; Dominio C (condensación) lleva a cabo el enlace peptídico, (Napdos.ucsd.edu, 2017).

## 1.6 Análisis Metagenómico de esponjas marinas

La caracterización de los microbiomas asociados a esponjas marinas es posible por medio de su estudio metagenómico, el cual consiste en secuenciar todo el ADN presente en una muestra de tejido de esponja, secuenciarlo y posteriormente analizarlo por medio de herramientas bioinformáticas (Oulas *et al.*, 2015). Hay dos maneras en las que se puede realizar un análisis metagenómico de microbiomas: secuenciación de regiones específicas (secuenciación de amplicones), y secuenciación del genoma completo (secuenciación *shotgun*).

La primera manera consiste en la amplificación de genes marcadores, tales como el gen que codifica para la subunidad menor (16S) del ARN ribosomal (ARNr) (Oulas *et al.*, 2015). Este es un gen muy conservado que se utiliza para identificar a bacterias y arqueas, utilizando alguna región hipervariable de las 9 que contiene el gen. Estas regiones, se suelen utilizar individualmente o en pares de acuerdo al diseño de los *primers* que se van a utilizar, hecho que está en función de los grupos de microorganismos que se quieren estudiar. Conocer el perfil taxonómico de un microbioma, es útil para cuestiones ecológicas, para conocer abundancia y diversidad de los mismos y como varían estas comunidades en ambientes específicos (mediante la estimación de la diversidad alfa y beta), así como para inferir funciones biológicas con base en sus relaciones filogenéticas. La diversidad alfa proporciona información sobre la variedad de especies presentes (de acuerdo a su dominancia), dentro de una muestra. Existen tres formas comúnmente utilizadas para calcular la alfa diversidad a partir de datos provenientes de secuencias: curvas de rarefacción, estimación de la riqueza de especies e índices de diversidad de comunidades. Las curvas de rarefacción son gráficas que se construyen a partir de conteos promediados de unidades operacionales taxonómicas (*OTUs* por sus siglas en inglés) observadas al momento de submuestrear diferentes intervalos de secuencias originadas para una muestra. Si la curva alcanza su asíntota, significa que la profundidad de la secuenciación de la muestra en cuestión fue suficiente, de lo contrario se considera una secuenciación incompleta. Por otro lado, para el estimado de la riqueza de especies, se calcula el índice de Chao1. Este índice suma el total de especies observadas más la razón que hay entre el número de *OTUs* que presentaron una sola secuencia asignada (*OTUs* referidos como raros) y el número de *OTUs* con más de una secuencia asignada. Si hay una muestra con un gran número de *OTUs* raros, el índice de Chao1 estimará una riqueza de especies alta para ella, bajo la premisa de que la presencia de muchos *OTUs* raros es indicativo de una mayor probabilidad de que existan más *OTUs* que no hayan sido detectados. En contraste, para estimar la diversidad de comunidades bacterianas el índice más utilizado es el de Shannon-Wiener. Este índice toma en consideración la abundancia relativa (el

número de individuos por especie) así como la riqueza (número de taxa) presente. Varía de 0 para comunidades con un sólo taxón a valores más altos representativos de comunidades con muchos taxa, cada uno con pocos individuos (Somes, 2017).

Por su parte, la diversidad beta compara la composición de diferentes comunidades o muestras entre sí. Hay varias maneras de visualizar y analizar la beta diversidad (Rammette, 2007). Primero se construye una matriz de distancia o disimilitudes. Unifrac es una métrica muy utilizado en estudios de metagenomas que compara árboles filogenéticos, y calcula las distancias de las fracciones de rama que son únicas a cada muestra. Si no se generan ramas únicas entre muestras, Unifrac agrupa las muestras que tienen distancias filogenéticas en común (Lozupone y Knight, 2005). Con esto se genera una matriz de distancia que posteriormente puede ser visualizada a través del uso de métricas multidimensionales como, “Análisis de Coordenadas Principales” (PCoA, por sus siglas en inglés) (Qiime.org, 2017a). Éste método calcula una matriz central que es descompuesta en sus componentes (eigenvalores y eigenvectores). Los eigenvectores, son estandarizadas al dividirlos por la raíz cuadrada de sus eigenvalores correspondientes (por cada eigenvalor hay un eigenvector, y hay tantos eigenvectores y eigenvalores como hay filas en la matriz de distancia) y representan las coordenadas principales de la gráfica; permiten ver los resultados de una matriz de distancia en un espacio de 2 o 3 dimensiones (Gower, 1966). Los análisis de diversidad alfa y beta pueden realizarse como parte de flujos de trabajos en programas bioinformáticos de libre acceso como el, “*Quantitative Insights into Microbial Ecology*” (QIIME, por sus siglas en inglés).

A partir de los datos taxonómicos generados por QIIME u otros programas, es posible inferir el contenido funcional del metagenoma en cuestión, a través del programa “*Phylogenetic Investigation of Communities by Reconstruction of Unobserved State’s* (PICRUSt, por sus siglas en inglés) (Languille *et al.*, 2013), una herramienta bioinformática que utiliza la ontología de genes para hacer este tipo de predicciones. Su única desventaja es que sus predicciones están limitadas a las familias de genes ya reportadas (en bases de datos como KEGG, por ejemplo), por lo que aquellos genes que no estén reportados, no serán predichos. Por otra parte, está la secuenciación *shotgun*, que como su nombre lo indica consiste en fragmentar el ADN metagenómico en pequeñas secuencias, para posteriormente secuenciarlo y volverlo a ensamblar. Este tipo de análisis metagenómico, proporciona una capacidad ilimitada para explorar la totalidad de genes (reportados y no reportados) involucrados en la producción de compuestos bioactivos (Oulas *et al.*, 2015).

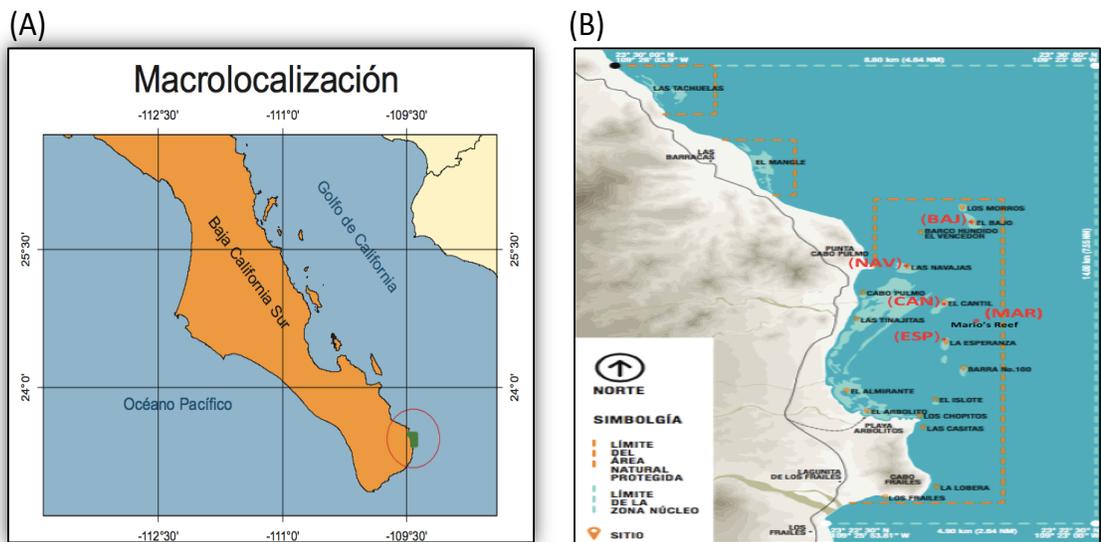
Ambas formas de generar datos metagenómicos son posible gracias a la secuenciación masiva, una tecnología que poseen secuenciadores de última generación o (NGS por sus siglas en inglés) entre los cuales figuran el *Illumina MiSeq/HiSeq* y *ROCHE-454 pyrosequencer*, recientemente, como los más populares para fines de estos estudios. Cada uno de estos secuenciadores tiene sus ventajas y desventajas en cuanto a tasas de error, tamaño de fragmento, tiempo de corrida y costo por corrida. Por ejemplo, aunque actualmente existe mucha información generada por el secuenciador *ROCHE-454* (lo cual facilita la comparación de información reciente con aquella más antigua), secuenciar por *Illumina* es más barato, y tiene rendimientos de hasta 60Gpb. Aunque las secuencias generadas por *Illumina* son más cortas (300pb) que las que puede generar el *ROCHE-454 pyrosequencer* (600-800pb), hecho que puede dificultar el ensamblaje de secuencias *de novo*. Sin embargo, cada vez se optimizan más los programas bioinformáticos utilizados más adelante en el análisis de datos, y la información generada por el secuenciador *Illumina* es cada vez más citada, por lo que la elección de la plataforma de secuenciación queda a las posibilidades e intereses de cada investigador (Thomas, Gilbert y Meyer., 2012).

El presente trabajo hizo uso del secuenciador *Illumina MiSeq*, para secuenciar el resultado de amplificar la región V4 del gen 16S ARNr, de 3 esponjas marinas recolectadas en diferentes puntos del Parque Nacional Cabo Pulmo, a fin de identificar la esponja con mayor diversidad microbiana. Se seleccionó esta región de amplificación puesto que ha sido citada por varios trabajos como aquella región que minimiza el sesgo entre la asignación taxonómica entre arqueas y bacterias (Caporaso *et al.*, 2010; Walters *et al.*, 2011). Además, el protocolo de *Illumina* recomienda el uso de esta región por haber sido la que tuvo mejor efecto en discriminar la ambigüedad taxonómica. Posteriormente, se secuenció por medio de *shotgun* el microbioma de la esponja con mayor biodiversidad, y se detectaron grupos de genes biosintéticos participes en la síntesis de PKS y NRPs, por medio del uso de las bases de datos especializadas: antiSMASH y NaPDoS. Paralelo a lo anterior, la esponja cuyo metagenoma fue secuenciado también fue sometida a pruebas *in vitro*, utilizando extractos crudos contra una cepa prototipo de tuberculosis (H37Rv), y otra MDR (1820/6). Con éste trabajo se espera contribuir a los esfuerzos realizados para la búsqueda de compuestos antibióticos alternativos, en respuesta a la elevada tasa de resistencia antimicrobiana a la que se enfrenta la población mundial.

## 1.7 Área de estudio

El Parque Nacional Cabo Pulmo está ubicado en el Municipio de Los Cabos, Estado de Baja California Sur

(entre las ciudades de La Paz y San José del Cabo), entre los 23° 22' 30" y los 23° 30' 00" latitud norte y los 109° 28' 03" y los 109° 23' 00" longitud oeste, en la costa del Golfo de California. Su superficie es de 7,111 (siete mil ciento once hectáreas) (Fig. 8). En 1995 fue decretado como área natural protegida por ser hogar del único arrecife coralino en el Golfo de California. Éste arrecife es uno de los más antiguos de los continentes americanos; data desde hace 20,000 años (CONANP, 2006). A la fecha no existe estudio alguno sobre la diversidad de las especies de esponjas y microbios en este parque o su potencial biotecnológico, a pesar del importante papel que juegan ambos organismos en el mantenimiento y aprovechamiento del arrecife (CONANP, 2006, Brusca y Brusca, 2003). Según Perdicaris *et al.*, (2016), los arrecifes coralinos son un buen lugar para encontrar esponjas HMA, debido a la alta competencia por alimento y espacio que caracteriza a este nicho. Por lo anterior, el Parque Nacional Cabo Pulmo representa un buen sitio para empezar investigaciones dirigidas a la búsqueda de compuestos bioactivos provenientes de estos organismos.



**Figura 8.** (A) Macrolocalización del Parque Nacional de Cabo Pulmo; (B) Mapa del Parque; en rojo se señalan los 5 sitios donde se recolectaron esponjas en el mes de abril 2016: El Bajo, El Cantil, Mario's Reef, Las Navajas y La Esperanza (abreviados en mapa como: BAJ, CAN, MAR, ESP y NAV, respectivamente). Las líneas punteadas indican el límite del Área Natural Protegida (modificada de Fertziger *et al.*, 2012).

## 1.8 Justificación

El Parque Nacional Cabo Pulmo, B.C.S, es un área natural protegida que alberga el único arrecife coralino del Golfo de California. A pesar de haber mucha información sobre los corales que lo habitan y la riqueza íctica que posee, poco o nada se sabe sobre las esponjas y su microbiota. Las esponjas son organismos sésiles que están bajo una constante presión por el espacio, así como expuestas a la depredación, por lo que su primera línea de defensa es la producción de compuestos citotóxicos. Esta producción está favorecida por la alta abundancia microbiana que vive íntimamente asociada a ellas. La exploración de estos microbiomas representa una fuente potencial de compuestos bioactivos, tales como PKs y NRPs, los cuales han mostrado una actividad antibiótica prometedora. La presencia de los genes involucrados en la síntesis de estos compuestos puede ser explorada de manera más amplia, rápida y económica gracias a la disponibilidad de herramientas para el estudio de metagenomas, tales como la secuenciación masiva y la bioinformática. La identificación *in silico* de las enzimas que los producen facilitaría su futura síntesis y su prueba en modelos *in vitro*. Además, contribuiría a la búsqueda de nuevos antibióticos que ayuden a combatir la elevada tasa de resistencia bacteriana a la cual se enfrenta la población mundial. Finalmente, esta tesis también aportará información relevante sobre aspectos ecológicos gracias a la caracterización y comparación de los microbiomas de esponjas, que permitirá un mejor entendimiento de la relación entre el simbiote y su huésped.

## 1.9. Objetivos Generales

Caracterizar el metagenoma de una esponja marina procedente del Parque Nacional Cabo Pulmo, a fin de identificar *in silico* a los genes que codifiquen para PKS y NRPS.

### 1.9.1 Objetivos Particulares

- 1.) Identificar molecularmente a las especies de esponjas marinas obtenidas del Parque Nacional Cabo Pulmo.
- 2.) Caracterizar y comparar los microbiomas de al menos dos de las esponjas marinas obtenidas, para seleccionar aquella que presente mayor diversidad microbiana

- 3.) Realizar una primera aproximación de la actividad antimicrobiana usando extractos crudos de la esponja con mayor biodiversidad microbiana contra *Mycobacterium tuberculosis*
- 4.) Secuenciar el metagenoma de la especie con mayor biodiversidad microbiana e identificar *in silico* aquellos genes que codifiquen para policétidos y péptidos no ribosomales con potencial actividad antimicrobiana, a través de su comparación con bases de datos propias de estos compuestos.

## Capítulo 2. Materiales y Métodos

---

### 2.1 Colecta de organismos

La colecta de esponjas se realizó vía buceo libre, el 27 y 28 de abril del 2016, en 5 puntos distintos del arrecife coralino del Parque Nacional Cabo Pulmo (PNCP), B.C.S, con la colaboración del equipo del Dr. Héctor Reyes de la Universidad Autónoma de Baja California Sur (UABCS). Los sitios de muestreo fueron seleccionados en base a las diferentes profundidades de los arrecifes y a información local sobre la distribución de las especies. En la figura 8 se enumeran los lugares de colecta en el orden en que fueron muestreados: 1. El Bajo (BAJ) (profundidad: 14.8 m); 2. El Cantil (CAN) (profundidad: 13.4 m); 3. Mario's Reef (MAR) (profundidad: 23.5 m); 4. La Esperanza (ESP) (profundidad: 16.1 m) y 5. Las Navajas (NAV) (profundidad: 5.2 m). La temperatura varió entre 23-25 °C, siendo 23 °C, la temperatura de las estaciones más profundas: MAR y ESP y las estaciones con menor profundidad presentaron la temperatura mas alta.

El proceso de colecta consistió en despegar cuidadosamente un pedazo de esponja de 4 x4 cm aproximadamente y colocarla en una bolsa de plástico hermética con un poco de agua de mar. Todas las bolsas se encontraban rotuladas con el número de estación de colecta (1,2,3,4 o 5), seguido por un código numérico individual. Se procuró tomar un registro fotográfico del aspecto submarino de cada espécimen dada la predisposición de algunas esponjas a cambiar de color al entrar en contacto con el aire. La estación 1 (BAJ), fue una estación en la que las condiciones ambientales no permitieron generar un registro fotográfico de calidad de las esponjas recolectadas en esta zona, por lo que la descartamos para este trabajo. Posterior a su colecta, los especímenes se volvieron a fotografiar y se eliminó de su superficie restos de sedimento o cualquier organismo adherido al pinacodermo, con un poco de agua destilada. Con un bisturí esterilizado, se tomó una muestra de la parte interna de la esponja, que es donde se encuentran habitando la mayor parte de los endosimbiontes. Las muestras se guardaron en criotubos y congelaron en nitrógeno líquido para los análisis moleculares. Una muestra de cada espécimen fue conservado en etanol al 70% para su identificación morfológica. Una vez en el laboratorio todas las muestras congeladas fueron mantenidas a -80 °C hasta su procesamiento.

## 2.2. Identificación molecular de esponjas utilizando la subunidad I de la citocromo oxidasa mitocondrial (COI).

### 2.2.1 Extracción de ADN de esponjas marinas

Se extrajo el ADN total de las muestras recolectadas utilizando el kit comercial: *PowerSoil* de *MOBIO*, cuyo protocolo fue modificado para incluir fenol: cloroformo: isoamil alcohol (25:24:1) en el paso de la homogenización del tejido, para promover una mejor lisis celular y purificación del ADN (Barnett and Larson, 2011). Esto se realizó añadiendo 200  $\mu$ L de fenol: cloroformo: isoamil alcohol, a un tubo *Power Soil* con perlas, al cual se le había añadido previamente 0.25 g de muestra de esponja y 60  $\mu$ L de solución C1 del kit. Posteriormente se incubó la muestra por 5 min, se le dió vortex por 5 min, y se centrifugó a 14 000 x g a temperatura ambiente, por 1 min. El sobrenadante de la muestra se transfirió a tubos limpios y se continuó haciendo la extracción de acuerdo a lo reportado por el protocolo de *PowerSoil* de *MOBIO*. La cuantificación del ADN extraído fue realizada con un espectrofotómetro Nano-Drop 8000 (*Thermo Scientific*, E.U.A). Para visualizar los resultados se realizaron geles de electroforesis con 0.8% de agarosa, aplicando 80 V por 30 min. Después se amplificó por medio de la reacción en cadena con polimerasa (PCR), al gen de la subunidad I de la citocromo C oxidasa mitocondrial (COI). Para ello se utilizaron los siguientes primers degenerados: dgLCO1490 y dgHCO2198 (Meyer *et al.*, 2003), mandados a elaborar a la empresa SIGMA-ALDRICH.

La reacción de amplificación fue de un volumen total de 20  $\mu$ L y consistió de los siguientes reactivos: 20 ng de ADN, solución tampón de PCR a 10X, 25 mM de  $MgCl_2$ , 10  $\mu$ M de *primer forward* ó sentido (*f*) y *reverse* ó antisentido (*r*), 10 mM de dNTPs, y (5 U/ $\mu$ L) de Taq (Biolasa de BIOLINE). Para realizar la amplificación se utilizó un Termociclador Veriti (Applied Biosystem) y el siguiente programa de amplificación: 1 min de desnaturalización inicial a 95 °C; 37 ciclos de: 40 s a 95 °C (desnaturalización de ADN), 40 s a 51 °C (anillamiento de primers) y 60 s a 72 °C (extensión de amplicon).

Los amplicones obtenidos de la reacción PCR, se visualizaron a través de un gel de electroforesis compuesto por 1.2% de agarosa, aplicando 100V por 40 min (Fig. 24). Posteriormente, los amplicones fueron purificados utilizando el kit "QIAquick PCR Purification Kit" de "QIAGEN" para eliminar residuos de la reacción PCR, y después fueron enviados a secuenciar por el método Sanger, a la empresa SeqXcel, Inc. en San Diego, CA.

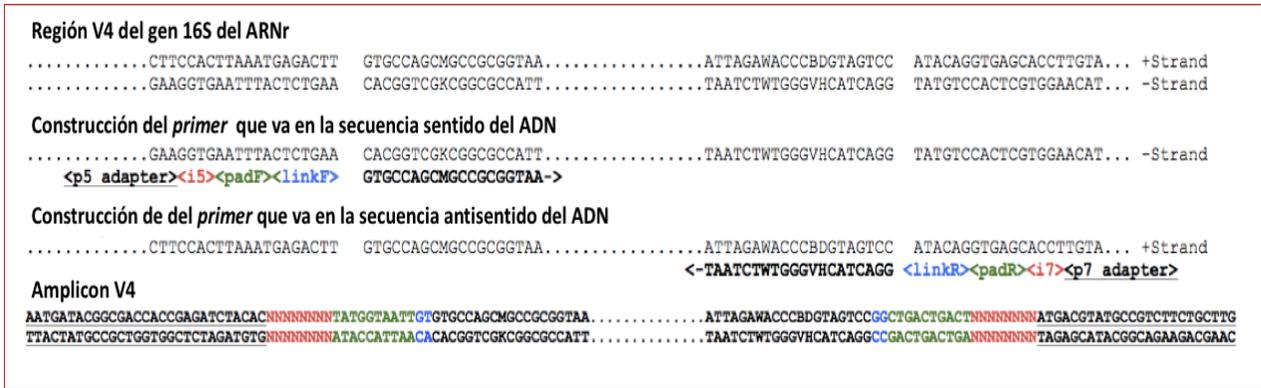
### 2.2.2 Identificación molecular

Se utilizó el programa MEGA7 (Kumar, Stetcher and Tamura, 2016), para revisar la calidad de las secuencias (*f* y *r*) y convertirlas a su archivo fasta. Las secuencias consenso se generaron mediante la unión de la secuencia *f* y *r* utilizando la opción “*align two or more sequences*” ofrecida por la plataforma del *National Center for Biotechnology Information* (NCBI por sus siglas en inglés): <<BLAST: blastn suite> (Blast.ncbi.nlm.nih.gov, 2017). La identificación molecular se realizó por medio de una serie de alineaciones por BLAST (Altschul, *et al.*, 1990) de las secuencias: *f*, *r* y consenso (alineación entre *f* y *r*) de cada muestra, contra la base de datos para nucleótidos no redundante (nt/nr). El criterio seguido para seleccionar los mejores hits, fue el reportado por Claverie y Notredame, (2007) que se basa en un porcentaje de identidad mayor a 70%, valores de E cercanos a 0.0, puntajes de mayor a 50 bits y porcentajes de cobertura próximas al 100%. En conjunto con este análisis, se llevó a cabo una búsqueda en las bases de datos: “*World Register of Marine Species: World Porifera data base*” (*Marinespecies.org*, 2017) y “*The Sponge Guide*” (*Spongeguide.org*, 2017), para verificar que la identificación taxonómica dada por BLAST coincidiera con la distribución esperada de la esponja bajo estudio.

## 2.3. Preparación de librerías 16S del ARNr para secuenciar por Illumina MiSeq System

### 2.3.1 Amplificación de la región V4 del gen 16S del ARNr y preparación de las librerías de secuenciación

La amplificación del locus del gen 16S del ARNr y la preparación de las librerías se realizaron en un sólo paso, siguiendo el protocolo reportado por Kozich *et al.*, (2013). En éste método se amplifica la región variable V4 del gen 16S procariota que codifica para el ARNr, al mismo tiempo que se hace el indexado de adaptadores e índices necesarios para su secuenciación. Los *primers* reportados por el autor (Fig.9), fueron sintetizados por la empresa SIGMA-ALDRICH®. Estos *primers* tienen la característica distintiva de que tanto el que se coloca en la secuencia sentido como el del antisentido poseen un índice distintivo que permite realizar combinaciones de estos para etiquetar un número mayor de muestras destinadas a secuenciar.



**Figura 9.** Diseño y secuencia de los *primers* sentido y antisentido utilizados para la estrategia 1 (imagen modificada de Kozich *et al.*, 2013). Los *primers* se componen de las siguientes secuencias: el adaptador para Illumina, un índice que los identifique al momento de secuenciarlos, una sección con 10 nucleótidos que previene formaciones secundarias no deseadas (padF), una secuencia de 2 nucleótidos no complementaria al gen 16S (linkF), y el primer que amplifica la región V4 del gen 16S que codifica para el ARNr.

Posteriormente se realizó la reacción PCR utilizando un volumen de 20  $\mu\text{L}$  (4  $\mu\text{L}$  de buffer marca “*Bioline-MyTaq*”, 2  $\mu\text{L}$  de primers, 0.125  $\mu\text{L}$  de enzima marca “*Bioline-MyTaq*”, 1  $\mu\text{L}$  de ADN 10 ng/ $\mu\text{L}$  y 12.875  $\mu\text{L}$  de  $\text{H}_2\text{O}$  desionizada). Las condiciones del termociclador fueron las siguientes: 95 °C por 3 min (desnaturalización); 25 ciclos a: 95 °C por 30 s, 55 °C por 30 s (hibridación de *primers*) y 72 °C por 30 s (elongación), y finalmente 72 °C por 5 min (elongación final). Según Kozich *et al.*, (2013), el tamaño esperado del amplicon para dicha reacción es de 318-322 pb.

### 2.3.2 Preparación de metalibrerías

Todos los productos PCR generados fueron visualizados en un gel de electroforesis elaborado con las condiciones previamente descritas (Fig. 25, 26, 27 y 28). Según la intensidad del amplicon se generaron distintos grupos o pools que después fueron purificados y normalizados utilizando el kit *SequalPrep Normalization Plate* de *Invitrogen*, siguiendo las indicaciones del fabricante. Los resultados de esta reacción fueron cuantificados utilizando un QUBIT HS (de la empresa *Invitrogen*<sup>TM</sup>). Para la normalización del metapool se igualaron todos los pools generados por el *Sequalprep* a la concentración más baja, considerando un tamaño de fragmento de 390 pb para hacer los cálculos correspondientes. Posteriormente se desnaturalizó el metapool, utilizando NaOH (200 mM) y se llevó a una concentración entre 5-6 pM (que corresponde a la concentración comúnmente utilizada en una placa de secuenciación

tamaño micro, la cual genera hasta 4 millones de secuencias) en 1 mL. Lo mismo se realizó con el PhiX (librería utilizada como control para la secuenciación). A continuación, se preparó la solución de carga con 1 mL del metapool generado y 5% de PhiX. Esta solución se incubó en un Termoblock (*Termostat Plus*) por 2 min a 96 °C, para asegurar la completa desnaturalización de los amplicones y después se dejó en hielo por 5 min. Finalmente, se colocó un volumen mínimo de 650 µL de la solución carga en la placa de secuenciación. La amplificación de la secuenciación fue *paired end*, lo que significa que se secuenció desde los dos extremos del amplicón (16S Metagenomic Sequencing Library Preparation: Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System, 2013).

Se realizaron 2 secuenciaciones (SEQ1 y SEQ2) siguiendo la metodología descrita arriba. En la primera secuenciación (SEQ1) se sometieron las 3 especies más abundantes del parque, estas fueron: *Prosuberites laughlini*, *Aplysina revillagigedi*, y *Cliona celata* (se secuenció un ejemplar por cada estación de proveniencia), y también se analizaron 4 muestras de agua, y 4 muestras de sedimento (una proveniente de cada estación). Por otra parte, en la segunda secuenciación (SEQ2) se secuenciaron 3 réplicas biológicas para cada especie de esponja proveniente de cada estación, y 3 submuestras de las muestras de sedimento y agua de cada estación. En anexos (1.4) se muestran los resultados del número de secuencias generadas para cada muestra en cada secuenciación, así como información relevante sobre la tasa de error de cada secuenciación realizada.

## 2.4 Análisis bioinformáticos

### 2.4.1 QIIME: *Quantitative Insights into Microbial Ecology*

Para realizar el análisis de las secuencias generadas, primero se descargaron en formato *.fastq* de la aplicación en línea, *BaseSpace Sequence Hub*. Esta aplicación, entre otras cosas, permite monitorear y analizar los datos generados por el secuenciador en tiempo real. La calidad de las secuencias se analizó, mediante el programa FastQC. Este programa permite comprobar el tamaño de ambas secuencias generadas (secuencia sentido: R1 y antisentido: R2), así como la calidad de sus nucleótidos. Cada nucleótido tiene un valor de calidad ó Q SCORE (QS) asociado a la probabilidad que tuvo el secuenciador de cometer un error al momento de incorporarlo a la secuencia. Se procuró que todas las secuencias tuvieran el 50% de los nucleótidos con un QS igual ó mayor a 28. Un valor mayor a 28, significa que el secuenciador tuvo una probabilidad de cometer un error en 1 de cada 1000 bases incorporados a una secuencia y que los nucleótidos observados en las secuencias generadas tienen una

exactitud de aproximadamente 99.9% (*Quality Scores for Next Generation Sequencing: Assessing sequencing accuracy using Phred quality scoring*, 2011). A continuación, se utilizó el programa QIIME versión 1.9.1, para: el preprocesamiento de las secuencias, la identificación de las unidades taxonómicas operacionales (OTUs), asignación taxonómica, análisis de diversidad alfa y beta, así como los análisis estadísticos necesarios para evaluar significancia de datos (valores-P), siguiendo el siguiente flujo de trabajo (Qiime.org, 2017).

1. Unión de las secuencias (R1 con R2); comando utilizado: *join\_paired\_ends.py*
2. Filtración de puntajes de calidad menores a un *quality score (QS) < 28*; comando utilizado: *fastq\_quality\_filter.py*
3. Conversión del formato fastq a formato fasta; comando utilizado: *convert\_fastaqual\_fastq.py*.
4. Asignación de etiquetas para uso de QIIME; comando utilizado: *add\_qiime\_labels.py*

Enseguida el archivo obtenido del paso 4, se procesó para identificar y eliminar quimeras por referencia con el algoritmo USEARCH 5.2.236. La base de datos utilizada como referencia fue la “gold.fasta” descargada de la página de UCHIME.

5. Organización de secuencias en grupos de OTUs (límite de similitud: 97%); comando utilizado *pick\_otus.py*: método de agrupamiento: *uclust\_ref* (Edgar, 2010); base de datos de referencia utilizada: *SILVA123*.
6. Elaboración de tabla de OTUs; comando utilizado: *make\_otu\_table.py*.
7. Filtrado de cloroplastos y mitocondrias de la tabla de OTUs; comando utilizado: *filter\_taxa\_from\_otu\_table.py*.
8. Conversión de la tabla de OTUs de formato .biom a .txt; comando utilizado: *biom convert*
9. Generación de gráficos de barras de las frecuencias relativas de los OTUs; comando utilizado: *summarize\_taxa\_through\_plots.py*.
10. Cálculo de la diversidad alfa; comando utilizado: *alpha\_diversity.py*: índices calculados: Chao1 e índice de Shannon.
11. Cálculo de rarefacción; comando utilizado: *alpha\_rarefaction.py*.
12. Cálculo de beta diversidad (generación de matriz de distancia utilizando Unifrac) y análisis de coordenadas principales; comando utilizado: *beta\_diversity\_through\_plots.py*.

13. Normalización de la matriz de distancia por Jackknife al número máximo de secuencias en común en todas las muestras, y generación de gráficos de coordenadas principales; comando utilizado: *jackknifed\_beta\_diversity.py*.
14. Generación de gráficos en 2D de los resultados de la matriz de distancia generada en el paso 12.; comando utilizado: *make\_2d\_plots.py*.

#### **2.4.2 PICRUSt: *Phylogenetic Investigation of Communities by Reconstruction of Unobserved States***

Se realizó la predicción de las funciones del microbioma de tres de las especies de esponja identificadas, con base en la tabla de OTUs generada por QIIME, mediante el uso del programa: *PICRUSt* (Languille *et al.*, 2013). Este programa fue utilizado desde la página web: <http://huttenhower.sph.harvard.edu/galaxy/>. El PICRUSt, consiste en la ejecución de 3 comandos básicos. El primero es para normalizar la tabla de OTUs, para esto se divide la abundancia (número de lecturas) de cada OTU entre el posible número de copias que pudieran existir del gen 16S para cada OTU. Este número de copias se obtiene según la ubicación del OTU en un árbol filogenético construido con organismos procariotas cuyos genomas han sido secuenciados completamente. El segundo paso consiste en predecir el contenido de genes de cada OTU, de acuerdo al contenido de genes presentes en sus parientes evolutivos más cercanos, esto se realiza multiplicando el contenido de genes inferidos por el número de OTUs de cada muestra. El tercero, y último paso consiste en categorizar dichos genes por su función, utilizando el esquema funcional: “*KEGG Orthology*” (*KOs*, por sus siglas en inglés), (Languille *et al.*, 2013).

#### **2.4.3 MG-RAST: *metagenomics Rapid Annotation using Subsystem Technology***

El MG-RAST, es una plataforma que permite analizar la composición y la función de comunidades microbianas a partir de datos metagenómicos en crudo o ensamblados. El portal ofrece servicios como: control de calidad automatizado, anotación de genes basado en similitud, análisis comparativos y el archivamiento de información (Meyer, *et al.*, 2008; Metagenomics.anl.gov, 2017). El sistema MG-RAST se utilizó vía su interface web. Primero se le cargó un archivo de metadatos, con la descripción de las muestras y posteriormente se sometieron las secuencias en crudo (en archivo fasta) generadas por la secuenciación shotgun. Las opciones de control de calidad son de opción libre e incluyen varias opciones

como recorte de lecturas basado en: Q SCORE, longitud de secuencia, y número de bases ambiguas, la eliminación de artefactos de la secuenciación como lecturas duplicadas, y filtrado de ADN de organismos modelos (ej. humano). Todos estos parámetros fueron aplicados con los valores por defecto (*default*) de MG-RAST previamente a someter las secuencias al proceso de anotación. Al someter las secuencias, MG-RAST le asignó un número de acceso al trabajo, que dependiendo de su tamaño y su prioridad (estatus público tiene mayor prioridad que estatus privado) puede ejecutarse dentro de unas cuantas horas a una semana según la demanda computacional de cada trabajo. El trabajo aquí presente fue sometido de manera privada y tardó alrededor de 12 horas en realizarse. Durante este tiempo MG-RAST, ejecutó una serie de búsquedas de similitud, entre proteínas predichas y secuencias proteicas en bases de datos (para *shotgun*), y de ácidos nucleicos (para lecturas similares a secuencias 16S y 18S). Los resultados de la anotación anterior se muestran como parte de los resultados de este trabajo.

#### **2.4.4 antiSMASH: *antibiotics & Secondary Metabolite Analysis Shell***

Para ampliar las anotaciones obtenidas por MG-RAST, se ensambló *de novo* el metagenoma utilizando el ensamblador de MEGAHIT. Este ensamblador utiliza k-meros (secuencias de longitud k en el ADN) de menor a mayor tamaño para realizar múltiples gráficos De Bruijn comprimidos (sDBG, por sus siglas en inglés) a través de una serie de iteraciones. De esta manera los k-meros son ordenados y cuantificados en secuencias sobrelapadas o contigs. Entre mayor número de fragmentos de diferentes tamaños se tenga, el ensamblado del metagenoma es mejor, puesto que se pueden unir huecos entre secuencias con fragmentos pequeños, o detectar secuencias repetidas en el caso de fragmentos grandes (Li *et al.*, 2015). Posterior a la generación de contigs, el archivo final (fasta) fue separado en archivos con un máximo de 50,000 secuencias. Esto se logró utilizando el comando split de seqkit. Lo anterior se realizó debido a que antiSMASH tiene un límite de <30MB y <50,000 secuencias individuales para los archivos que les son subidos.

Cada uno de los archivos fue sometido a un análisis por la version web 3.0 de antiSMASH (Medema *et al.*, 2011; Weber *et al.*, 2015; Antismash.secondarymetabolites.org, 2017), una herramienta bioinformática automatizada para la identificación de grupos de genes involucrados en la síntesis de metabolitos secundarios. Este programa ofrece la posibilidad de ejecutar una serie de algoritmos con diferentes funciones, que son seleccionados de acuerdo a las necesidades del usuario. La base de los análisis que se llevan a cabo, parte del conocimiento de que los genes involucrados en la vía biosintética

de diferentes clases de metabolitos secundarios se organizan en grupos o clusters, conocidos como clusters de genes biosintéticos (BGCs, por sus siglas en inglés) (Weber *et al.*, 2015). Para fines de este trabajo se seleccionaron los siguientes 4 algoritmos al momento de ejecutar la búsqueda de PKS y NRPS por antiSMASH:

- ClusterFinder, algoritmo probabilístico que emplea un modelo oculto de Markov para detectar BGCs de clases de compuestos desconocidos. Se basa en la detección de dominios Pfam (genes que codifican para enzimas comunes a todas las vías biosintéticas de compuestos secundarios, ej. oxidoreductasas y metiltransferasas), para predecir regiones BGC “parecidas” ya que estos suelen encontrarse formando parte de BGC’s. Los BGCs “parecidos” son categorizados en los siguientes tipos: sacáridos, “Cf\_saccharide”, ácidos grasos, “Cf\_fatty acid” y putativos “Cf\_putative”
- ClusterBlast, herramienta con la cual se remueven réplicas de secuencias y se detectan homologías entre los BGC’s encontrados y aquellos presentes en una base de datos construida por antiSMASH, la cual contiene todos los BGC’s reportados para todo el árbol de la vida.
- Clusters de genes del metabolismo secundario de grupos ortólogos (smCOGs), este módulo es utilizado para predecir y categorizar las funciones de genes accesorios y para calcular árboles filogenéticos para cada gen con un alineamiento semilla de su familia de proteína smCOG.
- KnownClusterBlast, compara BGCs identificados específicamente con la siguiente base de datos: “*Minimum Information about a Biosynthetic Gene Cluster*”, que consiste de 1172 BGCs experimentalmente caracterizados a través de un proyecto comunitario.

#### **2.4.5 NaPDos: *Natural Product Domain Seeker***

También se utilizó la base de datos NaPDos (Ziemert *et al.*, 2012), dirigida a la identificación de dominios KS (ketosintasa) y C (carboxilo) característicos de PKS y NRPS, respectivamente. Esta base de datos, puede analizar datos provenientes de secuenciaciones por amplicones, así como datos metagenómicos de *shotgun*. Funciona a través de la identificación y comparación de secuencias pertenecientes a dominios candidatos de metabolitos secundarios (a partir de una secuencia de ADN o amino ácidos ingresada a la base de datos) contra secuencias de dominios de referencia, los cuales forman parte de una base de datos curada por sus autores (Ziemert *et al.*, 2012). A partir de esta comparación, crea agrupaciones filogenéticas de dominios-específicos, con los cuales hace predicciones de sus productos putativos y calcula su similitud con otras vías biosintéticas conocidas (Napdos.ucsd.edu,

2017). Las secuencias que se sometieron a NaPDos, fueron las de los dominios anotados como KS por antiSMASH que fueron identificados en los BGC's de PKS tipo I por la misma base de datos.

## **2.5 Generación de librería *shotgun* para secuenciar el microbioma de *Aplysina revillagigedi***

Para la extracción del ADN procariota, se realizó una separación celular del tejido de la esponja para minimizar la presencia del ADN eucariota, debido a que el análisis del metagenoma completo implica que se secuencia todo el ADN presente en la muestra, tanto el ADN del huésped como el del hospedero. Para la separación celular se utilizó el protocolo reportado por Takahiro *et al.*, (2012), y para la extracción de ADN, el de Thomas *et al.*, (2010), con algunas modificaciones en ambos casos. En el primer protocolo se modificó el número de fracciones generadas de 5 a 4, y en el segundo se modificaron los volúmenes utilizados, para ajustarlos al material utilizado. Por otro lado, la preparación de la librería se realizó con el kit *Nextera XT Library Preparation*, que es un protocolo utilizado para secuenciar genomas microbianos o que son menores a 5 Mb.

### **2.5.1 Separación celular**

Se lavaron 20g de esponja con solución salina sin calcio y magnesio (DPBS) de la marca SIGMA-ALDRICH 2 veces para remover macroorganismos adheridos. A continuación, se maceraron con mortero grupos de 5 g de esponja con buffer TEN a pH de 8.5 por 5-10 min cada uno y se filtró utilizando una malla de 20  $\mu\text{m}$ . El volumen del homogenizado se repartió en 4 tubos Falcon de 15 mL y se rotularon como fracción 1. Esta fracción fue centrifugada a 500 x g por 5 min a 4 °C (para precipitar a las células de esponja). Posteriormente, se vació el sobrenadante a un tubo rotulado como fracción 2 y este se centrifugó a 1000 x g por 15 min a 4 °C con una centrifuga refrigerada *Heraeus Multifuge 15-R*. Se decantó el nuevo sobrenadante sobre un tubo rotulado como fracción 3 y se centrifugó a 3000 x g por 15 min a 4 °C y el sobrenadante obtenido se volvió a decantar en un tubo rotulado como fracción 4 y se centrifugó a 8000 x g por 15 min a 4 °C. Posteriormente se recuperó el precipitado y se lavó con 15 mL de buffer TEN, para después centrifugarlo a 10,000 x g por 20 min a 4 °C , (este paso se repitió una segunda vez).

### 2.5.2 Extracción de ADN

Para la extracción de ADN se le se le añadió 30 mL de una solución con buffer TE y lisozima (10 mg/mL) (la solución TE+ lisozima contenía 13.8 mL de TE + 1.2 mL de lisozima (125 mg/mL)) al tubo falcon de 50 mL donde se juntó el pellet de las fracciones 4. Esta solución se incubó por 1 hr a 37 °C. y se añadió 1.875 mL de Proteinasa K (20 mg/mL) para obtener una concentración final de 2.5 mg/mL y 1.5 mL de SDS 20 %, para llevarlo al 2 % en un volumen final de 18.375 mL; se incubó por 2 hr a 50 °C. Después se añadió, un volumen 1:1 de fenol:cloroformo:isoamil alcohol, y se centrifugó a 2000 x g por 10 min 2 veces. Se recuperó la fase acuosa y se repitió este paso. Se precipitó el ADN añadiendo un volumen y medio de etanol frío (25.25 mL) y se dejó a -20 °C durante 16 hr. Posteriormente, se transfirió un volumen de 11.875 mL a 4 tubos falcon de 15 mL para centrifugarlos a 14,800 x g por 30 min a 4 °C. Después se secó el pellet a temperatura ambiente (TA), y se resuspendió en TE con RNasaA (10 µg/mL) e incubó por 1 hr a 37 °C. A continuación, se añadió un volumen igual de fenol:cloroformo:isoamil alcohol 25:25:1, y se precipitó de la fase acuosa con 0.1 volúmenes de 3 M acetato de Na (20 µL) y 2.5 volúmenes de etanol (500 µL). Se continuó a precipitar el ADN a 14,000 x g por 30 min a 4 °C, y se lavó con 300 µL de etanol al 70 %. Se pasó a secar por 17 min a 60 °C, y finalmente se resuspendió en buffer TE.

Para evaluar la integridad del ADN, se realizó un gel de electroforesis con 0.8 % de agarosa en TBE 0.5X. La cuantificación del ADN se realizó utilizando el sistema de ensayos: QUBIT BS (Invitrogen). Posteriormente se hizo una dilución de la muestra para llevarla a una concentración de 100 ng/µL, y a partir de ahí se hicieron dos diluciones seriadas; 50 ng/µL y 10 ng/µL. Cada una de las diluciones se cuantificaron en QUBIT BS, y de la dilución de 10 ng/µL se preparó la dilución final a 2 ng/µL.

Paralelo a esto se verificó el éxito de la separación celular mediante PCR de los genes COI, 18s y 16s.

### 2.5.3 Librería de secuenciación *shotgun*

El proceso de preparación de la librería de secuenciación *shotgun* consiste en 4 pasos principales: la tagmentación, la amplificación, purificación y la normalización. El proceso de tagmentación se realizó utilizando la tecnología transposoma de Nextera, que consiste en hacer una reacción enzimática (50 µL) para fragmentar el ADN y añadir etiquetas al ADN. La nota más enfatizada

por el protocolo de Nextera XT para asegurar el éxito de esta reacción, es asegurarse que la cantidad de ADN ingresada sea lo más exacta posible (1 ng). Por lo anterior, se utilizó el método de cuantificación fluorométrico QUBIT HS marca Invitrogen<sup>TM</sup>. Posteriormente se amplificaron por PCR los fragmentos creados, en esta reacción se añadieron los adaptadores necesarios para la secuenciación. La purificación del producto PCR, se realizó utilizando *SequalPrep*, siguiendo el protocolo del fabricante. En este paso, la reacción de 50 µL obtenida de la tagmentación se dividió en dos, dado que el protocolo inicia con 25 µL máximo. La primera reacción consiste en la unión del ADN a la placa *SequalPrep*, a continuación, se pasó el sobrenadante a otro pocillo para una segunda y una tercera reacción de unión. Al final se utilizó un sólo volumen de elución para limpiar todos los pocillos. Se cuantificaron las reacciones purificadas en QUBIT HS y se eligió la reacción con concentración más alta del *SequalPrep* para secuenciar. La librería se normalizó a 16 pM (utilizando un fragmento de 700 pb para hacer los cálculos correspondientes). La solución carga fue enriquecida con 2 % de PhiX. La amplificación en la secuenciación fue *single read* o de un sólo extremo, con un tamaño de 151 pb.

## **2.6 Evaluación de la actividad antibiótica de extractos crudos de *Aplysina revillagigedi* contra *M. tuberculosis***

El extracto crudo de la esponja *Aplysina revillagigedi*, se llevó a cabo siguiendo el protocolo reportado por Azevedo *et al.*, 2008 y Kazanjian *et al.*, 2006 quienes trabajaron con una esponja del mismo género *Aplysina caissara*. Posterior a la obtención del extracto crudo, este fue sometido a ensayos de citotoxicidad utilizando dos concentraciones: 400 µg/mL (Azevedo *et al.*, 2008) y 200 µg/mL, (concentración basal, comúnmente utilizada para evaluar otros compuestos de origen marino, (ej. conotoxinas) en ensayos antituberculosos (Bernáldez-Sarabia., 2013). Se evaluó la capacidad de inhibir el crecimiento de 2 cepas de *Mycobacterium tuberculosis*: una cepa prototipo denominada, “H37Rv” y una cepa MDR denominada, “1820/16”. Se utilizó un método colorimétrico para evaluar el efecto del extracto crudo sobre las diferentes cepas. El agente revelador utilizado fue un compuesto de sales de tetrazolio [3-(4,5-dimetiltiazol-2-yl)-5-(3-carboximetoxifenil)-2-(4-sulfofenil)-2H-tetrazolio (MTS), combinadas con un reactivo de acoplamiento de electrones, denominado fenazín etosulfato (PES); ambos son propios del “CellTiter 96 AQueous One Solution Cell Proliferation Assay” de Promega. La manera en que funciona el ensayo es la siguiente: cuando hay células viables en un pocillo de ensayo, el MTS es reducido a formazán (un compuesto color café oscuro), por una de las dos moléculas reductoras (NADPH ó NADH) producidas por las enzimas deshidrogenasas que están presentes en células metabólicamente

activas (Berridge *et al.*, 1993). La intensidad de este cambio de color se mide con un espectrofotómetro a 450nm de longitud de onda, y es directamente proporcional al número de células metabólicamente activas en el cultivo (Cory *et al.*, 1991; Riss *et al.*, 1992).

### 2.6.1 Preparación del extracto crudo

Se pesaron 20 g de muestra húmeda de esponja y se lavaron con agua destilada estéril. Posteriormente se cortó en fragmentos el tejido para eliminar la fauna endobiótica visible, y se procedió a adicionar metanol (0.3 g/mL) al 100%. El tejido fue macerado con un homogenizador mecánico hasta que la muestra quedara completamente desintegrada (15-20 s). Posteriormente se incubó en hielo por 1 hr, y después se centrifugó a 5000 x g por 5 min, de acuerdo a lo reportado por Marinho *et al.*, 2008. La extracción se repitió de la misma manera una segunda y tercera vez, a excepción de que en la tercera vez se dejó incubando por una noche según lo sugerido por Sacristan-Soriano *et al.*, 2010. Transcurrido el último período de incubación, se filtró la solución con un filtro Whatman de 0.2  $\mu\text{m}$  para esterilizar, ya que con este tamaño de poro se pueden eliminar bacterias. Posteriormente se evaporó lo más posible del volumen de metanol con una rotoevaporadora y el residuo de eso con nitrógeno gaseoso (Fig.10). El volumen restante de los dos procedimientos anteriores, se liofilizó por dos días. Posteriormente se pesó y se almacenó a -80 °C.



**Figura 10.** Muestra el aspecto del extracto crudo de *Aplysina revillagigedi* durante su evaporación con nitrógeno gaseoso. Este paso llevó aproximadamente 10 hr y se recuperó aproximadamente 4-5 mL de extracto crudo.

### 2.6.2 Ensayo colorimétrico

Para realizar el ensayo colorimétrico, el extracto crudo debe estar resuspendido en dimetilsulfóxido (DMSO) al 1%. Esta concentración de DMSO es la cantidad óptima para trabajar con *M. tuberculosis*, según un trabajo realizado previamente por la Dra. Johanna Bernáldez Sarabia en el laboratorio del Departamento de Innovación Biomédica del CICESE, quien generó una curva de calibración del DMSO contra la cepa prototipo H37Rv. A cada pocillo de la microplaca se le añade un volumen final de 200  $\mu\text{L}$ , de tal manera que para que la concentración final del DMSO en cada pocillo sea del 1% deben adicionarse 2  $\mu\text{L}$  del extracto crudo resuspendido en DMSO al 100%. Como se evaluaron dos concentraciones del extracto crudo (400  $\mu\text{g}/\text{mL}$  y 200  $\mu\text{g}/\text{mL}$ ), se generó una alícuota del extracto crudo con DMSO (100%), que tuviera una concentración final de 40  $\mu\text{g}/\mu\text{L}$ . Para esto se resuspendieron 10 mg del extracto crudo liofilizado en 250  $\mu\text{L}$  de DMSO al 100%. La solución generada se agitó con un *BenchMixer*<sup>TM</sup> Vortexer, 115V ajustable 200-3200 RPM por 30 s y se mantuvo a TA hasta que se disolviese completamente (aproxímadamente 3 horas). A partir de esta solución se realizaron las diluciones necesarias para obtener una concentración final por triplicado de 400  $\mu\text{g}/\text{mL}$  y 200  $\mu\text{g}/\text{mL}$  en cada pocillo prueba.

Para el montaje del experimento, se utilizó una placa de 96 pocillos (Fig.11) a la cual inicialmente, se le añadieron 200  $\mu\text{L}$  de agua en la periferia, con el fin de evitar la evaporación de los pocillos centrales, los cuales contienen el medio de cultivo con el extracto. El medio de cultivo utilizado fue el *MiddleBrook 7H9*, enriquecido con 10 % de ácido oleico, albúmina, dextrosa y catalasa (OADC), para favorecer el crecimiento de la cepa. Primero, se transfirieron 196  $\mu\text{L}$  de éste medio a 3 pozos de trabajo. A estos mismos pozos se les añadieron 4  $\mu\text{L}$  de la solución del extracto crudo con DMSO al 100%, y se homogenizó cada uno 5 veces (con punta distinta). Se transfirieron 100  $\mu\text{L}$  de cada uno de estos pocillos a pocillos adyacentes a los cuales se les había añadido 100  $\mu\text{L}$  de *Middlebrook* previamente. Se homogenizó cada uno de estos segundos pocillos con punta diferente, y posteriormente se descartaron 100  $\mu\text{L}$  de estos pocillos. Posteriormente se les agregó 100  $\mu\text{L}$  de una suspensión bacteriana ( $1.5 \times 10^7$  células/mL) de un cultivo de *M. tuberculosis* realizado en medio Lowenstein-Jensen a todos los pocillos prueba, de tal manera que los primeros pocillos quedarán con una concentración final de 400  $\mu\text{g}/\mu\text{L}$ , y los segundos con 200  $\mu\text{g}/\text{mL}$ . También se realizó un control para el DMSO (98  $\mu\text{L}$  de medio *MiddleBrook* con 2  $\mu\text{L}$  de DMSO); un control de esterilidad (CE) del medio *MiddleBrook* (200  $\mu\text{L}$ ); un control de crecimiento (CC) que consistió de 100  $\mu\text{L}$  de medio *MiddleBrook*, más 100  $\mu\text{L}$  de inóculo; 1 control positivo: etambutol, que es un fármaco antifímico de primera línea (concentración utilizada: 4  $\mu\text{g}/\text{mL}$ , Martin *et al.*, 2005 en Bernaldez-Sarabia., 2013). A todo los pocillos control (excepto el control de

esterilidad) se les adicionaron 100  $\mu\text{L}$  de suspensión bacteriana. La preparación de la suspensión bacteriana consistió en generar una turbidez similar al estándar 1.0 de la escala de McFarland ( $3 \times 10^8$  células/mL) y posteriormente diluir el inóculo en una proporción 1:20 con medio *Middlebrook*.

El ensayo se dejó incubando a 37 °C por 7 días, debido a que *M. tuberculosis* es un patógeno de crecimiento lento. Transcurrido este tiempo, se le adicionaron 20  $\mu\text{L}$  de MTS a uno de los pozos de CC. El cambio de color indica que el patógeno creció adecuadamente con las condiciones en que la placa fue montada (esto tarda aproximadamente 2-4 horas en observarse). Si se observa un cambio de color, se revela el resto de la placa con el reactivo, y se vuelve a dejar incubando por 2 -4 hrs a 37 °C antes de hacer la medición de absorbancia a 450 nm de la placa entera.

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B1	B2	B3	B4	B5	B6 400 $\mu\text{g/mL}$	B7 400 $\mu\text{g/mL}$	B8 400 $\mu\text{g/mL}$	B9 ETB	B10 DMSO	B11 CE	B12
C1	C2	C3	C4	C5	C6 200 $\mu\text{g/mL}$	C7 200 $\mu\text{g/mL}$	C8 200 $\mu\text{g/mL}$	C9 ETB	C10 DMSO	C11 CE	C12
D1	D2	D3	D4	D5	D6	D7	D8	D9 ETB	D10 DMSO	D11 CE	D12
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11 CC	E12
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11 CC	F12
G1	G2	G3	G4	G5	G6	G7	G8	G9	G10 CC	G11 CC	G12
H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

**Figura 11.** Diseño de placa utilizada en bioensayos colorimétricos contra *M. tuberculosis*. Todos los pocillos de la perifería (enmarcados en azul) se rellenan con agua para evitar la evaporación de los triplicados de los pocillos experimentales. Cada concentración del extracto crudo utilizada (400  $\mu\text{g/mL}$  y 200  $\mu\text{g/mL}$ ), así como los controles de esterilidad (CE), DMSO al 1% y etambutol (ETB, control positivo) se hacen por triplicado. El control de crecimiento (CC) se hace 4 veces, el cuarto pocillo es el primero que se revela después de 7 días. El cambio de color de este pocillo después de 2-4 hr, indica si se podrá revelar el resto de la placa o si se tiene que dejar incubando por más tiempo.

## 2.7 Análisis estadísticos

### 2.7.1 Variabilidad entre estaciones

Para conocer si existían diferencias significativas en la composición taxonómica del microbioma de muestras de esponja provenientes de diferentes estaciones de colecta (diversidad beta), se realizó un t-test no paramétrico, utilizando la simulación Monte Carlo, la cual crea una distribución estadística de los datos a través de la permutación de valores de abundancia de muestras (réplicas biológicas) de los diferentes grupos (estaciones de colecta) (Bustad *et al.*, 2006). Esta prueba permite entre otras cosas comparar grupos de muestras con tamaños de muestra distintos (Costello *et al.*, 2009) y obtener un valor de significancia ( $P < 0.05$ ). Su cálculo se basa en la fracción del tiempo que se tarda la simulación Monte Carlo en detectar un valor “t” estadístico más grande o igual a un valor “t” estadístico observado (Qiime.org, 2017). La simulación se ejecutó bajo parámetros basales (999 permutaciones) y se utilizó la mediana de los datos de abundancia taxonómica de las réplicas de la esponja *Aplysina revillagigedi*, la cual se encontró en todas las zonas muestradas. La mediana fue calculada con el comando de QIIME: *dissimilarity\_mtx\_Stats.py*, utilizando como entrada una matriz de distancia filogenética ponderada calculada por Unifrac, la cual fue previamente normalizada por jackknife a el número máximo de secuencias en común entre todas las muestras (5533). La gráfica se generó utilizando el comando de QIIME: *make\_distance\_boxplot.py*.

Así también se realizó otra prueba no paramétrica, ANOSIM (Cc.oulu.fi, 2017), utilizando como entrada la matriz de distancia ponderada generada por Unifrac, para el análisis de diversidad beta, para calcular un valor de P, que indica si hay diferencias significativas entre las diferentes estaciones de las que fue recolectada *Aplysina revillagigedi*. Este valor de P, fue calculado con 999 permutaciones, utilizando el comando de QIIME: *compare\_categories.py –method ANOSIM*.

### 2.7.2 Variabilidad entre réplicas biológicas

Para evaluar la variabilidad entre réplicas biológicas se utilizó el comando de QIIME: *compare\_alpha\_diversity.py*, con el cual se calculó un valor P, utilizando un t-test no paramétrico (como se describió arriba), 999 permutaciones, una profundidad de rarefacción de 5533 (el número mínimo de

secuencias en todas las muestras) y el método Bonferroni que corrige un error asociado a comparaciones múltiples (Qiime.org, 2017b).

### **2.7.3 Efecto del extracto crudo de *Aplysina revillagigedi* sobre *M. tuberculosis***

Los análisis estadísticos se realizaron utilizando el programa, *Graphpad Prism*. Primeramente, los datos de absorbancia fueron sometidos a una prueba de *Shapiro-Wilk* para evaluar la normalidad, y una prueba de *Levene* para evaluar la homoscedasticidad de los datos. Pasadas estas pruebas (indicado por un valor de  $P > 0.05$ ), se utilizó un ANOVA de una vía seguida por una prueba de comparaciones múltiples Tukey para evaluar si había diferencias significativas entre el control de crecimiento y los tratamientos. Los resultados se reportan de acuerdo a O'Neill *et al.*, (2014), quienes reportan la siguiente fórmula:

$$\% \text{ de inhibición de crecimiento} = (1 - \text{d.o extracto} / \text{d.o promedio del CC}) \times 100$$

donde d.o es la lectura de densidad óptica.

El cálculo anterior se realizó para el valor de absorbancia de cada tratamiento, posteriormente se calculó el promedio de cada uno, junto con la desviación estándar.

### **2.7.4 Análisis comparativo de microbiomas entre esponjas y su ambiente**

Con el fin de encontrar el número de OTUs especie-específicos asociados a cada muestra de esponja, se realizaron diagramas de Venn usando el programa Venny (Bioinfogp.cnb.csic.es, 2017). Se seleccionaron únicamente aquellos OTUs cuyas secuencias sumadas fueran al menos 50 secuencias o más, para hacer las comparaciones. Esto para asegurar que los OTUs considerados en este análisis fueran parte real del microbioma de cada esponja, y no OTUs que pudieran haberse excluido porque el esfuerzo de secuenciación no fuese todo lo exhaustivo posible. El valor de corte utilizado en consideración al margen de error asociado a los artefactos de la amplificación PCR (Wintzingerode, *et al.*, 1997) es un valor arbitrario que se elige a criterio del investigador (Mason *et al.*, 2014; Arat *et al.*, 2014). Aquí se eligieron 50 secuencias, porque este valor se consideró una cantidad suficiente para asumir que un microorganismo pueda estar presente de manera constante y específica en una esponja, y que no sea un

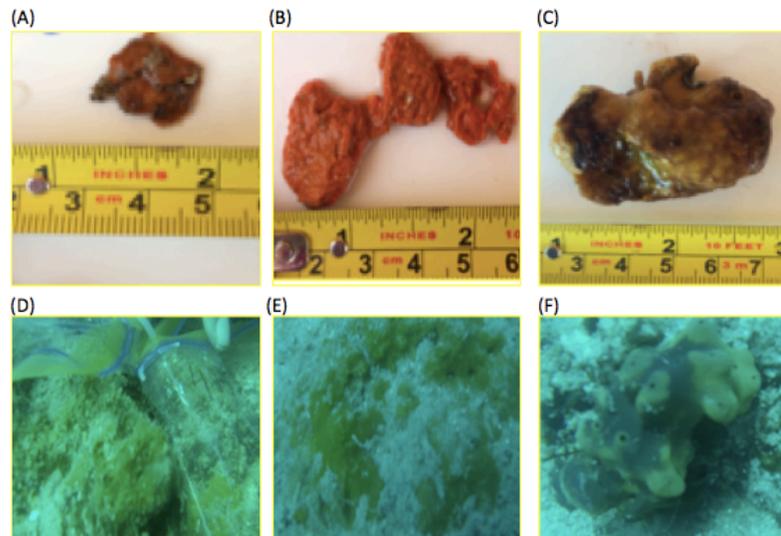
microorganismo que simplemente se encontraba transitando por su sistema de filtración al momento de su colecta (Kennedy *et al.*, 2014).

## Capítulo 3. Resultados

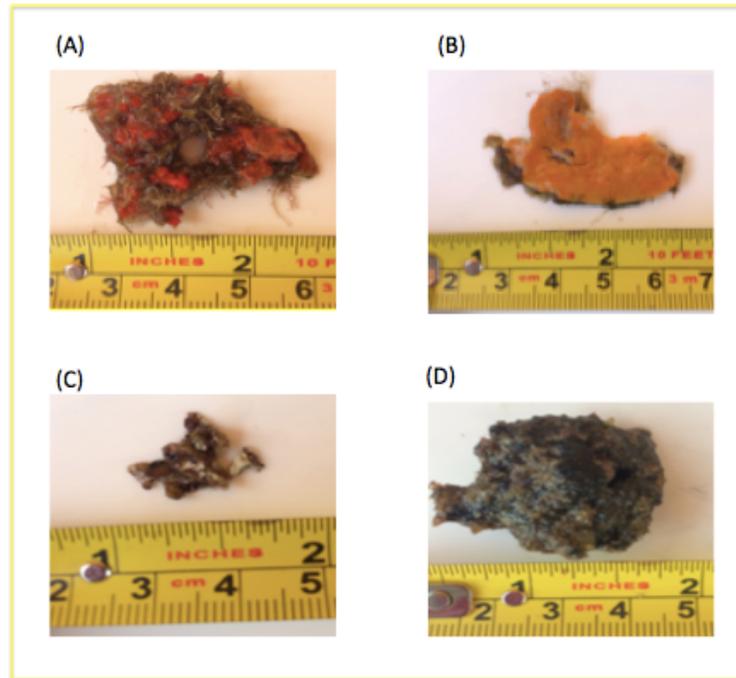
---

### 3.1 Colecta de esponjas

Se recolectaron un total de 61 muestras en el mes de abril del 2016, en 5 puntos distintos del Parque Nacional de Cabo Pulmo (PNCP): 5 de El Bajo (BAJ), 18 de El Cantil (CAN), 13 de Marios Reef (MAR), 12 de La Esperanza (ESP) y 13 de Las Navajas (NAV). Se tomó registro fotográfico de la apariencia de todas las esponjas colectadas dentro y fuera del agua que se encontraron en el PNCP. La figura 12 muestra las esponjas de las cuales se encontraron más ejemplares en el Parque, y la figura 13 aquellas esponjas de las cuales sólo se hallaron 4 o menos ejemplares.



**Figura 12.** Registro fotográfico de los ejemplares de esponja más abundantes encontradas en el PNCP. (A-C) Muestra su apariencia después de ser procesadas (E-G) Muestra su apariencia submarina.



**Figura 13.** Registro fotográfico de las especies de esponja marina menos abundantes del PNCP. De (A), se encontraron solamente 4 ejemplares, y de (B-D) sólo uno.

### 3.2 Identificación molecular de esponjas

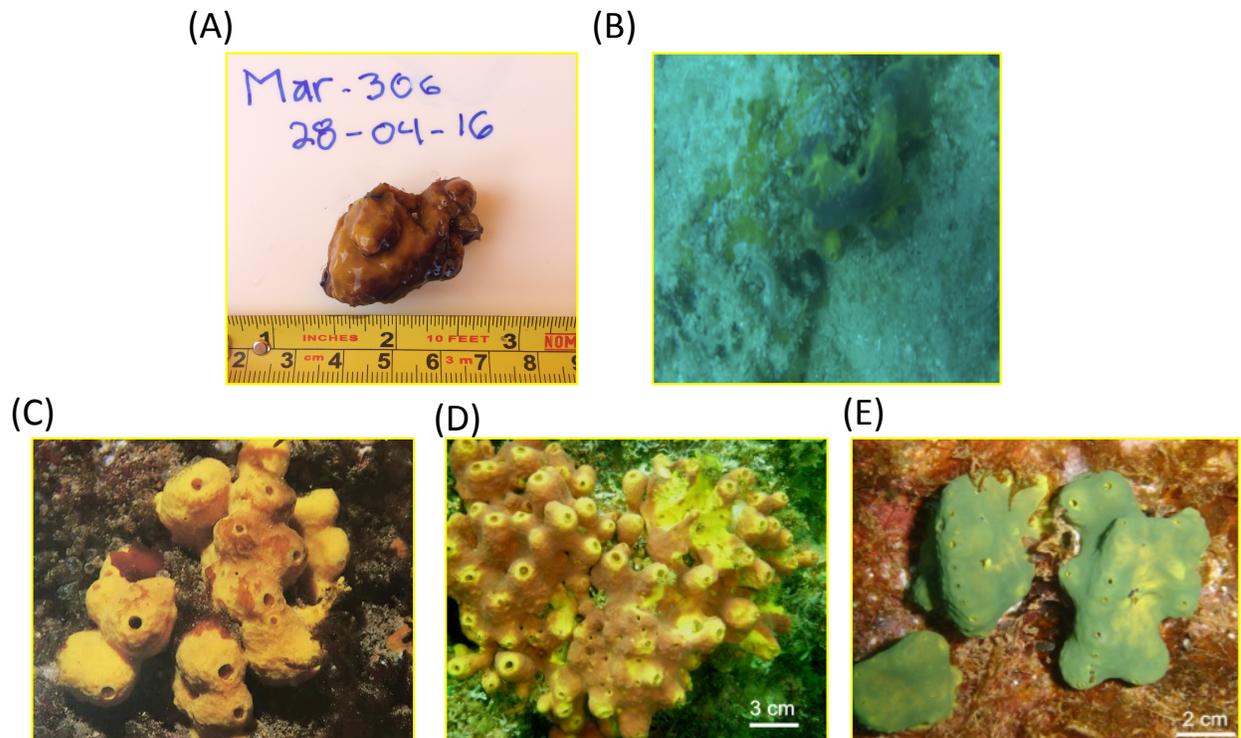
El ADN (Fig.22) obtenido de las 61 muestras recolectadas de esponjas se utilizó para amplificar su región COI (Fig.24) e identificarlas molecularmente a través del análisis comparativo (blastn) de sus secuencias de ADN, con secuencias depositadas en la base de datos nr/nt de NCBI. Con base en estos análisis, así como en una búsqueda bibliográfica y geográfica que reforzase los resultados generados por los blastn, se lograron identificar 7 especies de esponja: 4 identificadas a nivel de especie y 3 a nivel de género, con diferentes grados de confianza sobre su identificación molecular. Los parámetros utilizados para la identificación taxonómica de cada una de ellas se muestran en la Tabla 1, los resultados de sus alineamientos por blastn se muestran en Anexos (Fig.46-Fig.52).

**Tabla 1.** Resultados de análisis BLAST de secuencias consenso de cada especie identificada de esponja.

Estación	Especies/Género	Query Cover %	E. Value	ID %	Score (bits)	ID espécimen
Mario's Reef (MAR)	<i>Prosuberites laughlini</i>	100	0	98	1098	303
Mario's Reef (MAR)	<i>Aplysina revillagigedi</i>	100	0	99	961	301
Las Navajas (NAV)	<i>Cliona celata</i>	97	2E-174	85	623	504
Mario's Reef (MAR)	<i>Timea sp</i>	98	0	99	1182	302
El Cantil (CAN)	<i>Spongia sp</i>	99	0	96	966	212
La Esperanza (ESP)	<i>Clathria prolifera</i>	99	0	93	915	407
La Esperanza (ESP)	<i>Phelloderma sp</i>	99	1E-159	89	573	408

En principio, *Aplysina revillagigedi* fue una esponja difícil de identificar, pues los resultados del blastn resultaron en tres posibles candidatos con una distribución geográfica cerca del área de estudio, estos fueron: *Aplysina fistularis*, *Aplysina gerardogreeni*, *Aplysina revillagigedi* (Fig.14). Al hacer una comparación fotográfica entre las diferentes especies de esponjas con la nuestra, se observó una mayor similitud morfológica con *A. revillagigedi* que con otra esponja (Fig.14). Sin embargo, los resultados de los alineamientos por blastn (Fig. 15, 16 y 17), mostraron el mayor porcentaje de identidad con *A. gerardogreeni* (100%), mientras que con *A. revillagigedi* y *A. fistularis* el porcentaje de identidad fue de 99%. Aún más, no se presentaron *gaps* en los alineamientos de *A. gerardogreeni* ó *A. revillagigedi*, mientras que *A. fistularis* si se observó 1 *gap* (recuadro rojo, Fig.15). Sin embargo, cabe mencionar que el porcentaje de cobertura fue menor para *A. gerardogreeni* y *A. revillagigedi* (ambas 86%) que para *A. fistularis* (100%). Para saber si el porcentaje de cobertura de *A. revillagigedi* (la esponja que se asemejó más a la nuestra), se debía a que la secuencia reportada para esta esponja era más corta que la nuestra, se descargó la secuencia reportada para ella y se realizó un *blast 2 sequences* entre esta secuencia y la nuestra (Fig.18). Al hacer esto el porcentaje de cobertura cambió de 86% a 100%, ya que en efecto el

porcentaje que blastn había dado inicialmente (86%) era porque la secuencia del COI reportada para *A. revillagigedi*, es menor (523 pb) que la nuestra (604 pb). Por todo lo anterior, se inclinó por *A. revillagigedi* como la especie presente en el PNCP.



**Figura 14.** Registro fotográfico de un ejemplar de las esponjas recolectadas en el PNCP (ID ESPECIMEN:306) (A) fuera y (B) dentro del agua, y las 3 especies de esponjas candidatas para su identificación taxónomica (C) *Aplysina fistularis*, (D) *Aplysina gerardogreeni*, y (E) *Aplysina revillagigedi*.

*Aplysina fistularis* voucher UCMPWC1067 cytochrome oxidase subunit I gene, partial cds; mitochondrial

Max score	Total score	Query cover	E value	Ident	Accession
1105	1105	100%	0.0	99%	<a href="#">AY561987.1</a>

***Aplysina fistularis* voucher UCMPWC1067 cytochrome oxidase subunit I gene, partial cds; mitochondrial**

Sequence ID: [AY561987.1](#) Length: 709 Number of Matches: 1

Range 1: 50 to 654 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1105 bits(598)	0.0	603/605(99%)	1/605(0%)	Plus/Plus
Query 1	TTTTGCTGGTATGATAGGAACAGCATT	TTAGTATGTTAATTAGACTGGAGTTATCTGCACC		60
Sbjct 50	TTTTGCTGGTATGATAGGAACAGCATT	TTAGTATGTTAATTAGACTGGAGTTATCTGCACC		109
Query 61	AGGCTCAATGTTAGGAGACGATCATT	TATATAATGTTATAGTAACGGCACATGCTTTTGT		120
Sbjct 110	AGGCTCAATGTTAGGAGACGATCATT	TATATAATGTTATAGTAACGGCACATGCTTTTGT		169
Query 121	TATGATATTCTTTTAGTAATGCCTGTT	TATGATAGGGGTTTTGGTAATTGATTAGTTCC		180
Sbjct 170	TATGATATTCTTTTAGTAATGCCTGTT	TATGATAGGGGTTTTGGTAATTGATTAGTTCC		229
Query 181	ATTATATATTGGAGCACCAGATATGGCT	TTCCCTCGGTTAAATAATATTAGTTTTTGATT		240
Sbjct 230	ATTATATATTGGAGCACCAGATATGGCT	TTCCCTCGGTTAAATAATATTAGTTTTTGATT		289
Query 241	ATTACCCCAGCACTAATTTTATATTGGGGT	CTGCTTTTATAGAACAAGGGCCGGTAC		300
Sbjct 290	ATTACCCCAGCACTAATTTTATATTGGGGT	CTGCTTTTATAGAACAAGGGCCGGTAC		349
Query 301	GGGTTGAACCGTCTATCCGCCATTATCAAGT	TACAAACACATTCTGGGGTTTCGGTAGA		360
Sbjct 350	GGGTTGAACCGTCTATCCGCCATTATCAAGT	TACAAACACATTCTGGGGTTTCGGTAGA		409
Query 361	TATGGCAATATTTAGTCTTCATTTGGCCGGAAT	TTCCCTCCATATTAAGCTCTATTAATTT		420
Sbjct 410	TATGGCAATATTTAGTCTTCATTTGGCCGGAAT	TTCCCTCCATATTAAGCTCTATTAATTT		469
Query 421	TATAACTACAATTTTAAATATGAGGGCGCCAGGGAT	GACTATGGACAGATTACCTTTATT		480
Sbjct 470	TATAACTACAATTTTAAATATGAGGGCGCCAGGGAT	GACAATGGACAGATTACCTTTATT		529
Query 481	CGTTTGATCTATTTTATTAACAACAATATTATTAGT	TATTAGCATTGCCCGTTTTAGCTGG		540
Sbjct 530	CGTTTGATCTATTTTATTAACAACAATATTATTAGT	TATTAGCATTGCCCGTTTTAGCTGG		589
Query 541	AGCAATTACTATGCTTTTAAACCGATAGAAATTTTAA	CACAACA-TTTTGGATCCAGCAGG		599
Sbjct 590	AGCAATTACTATGCTTTTAAACCGATAGAAATTTTAA	CACAACAATT-TTTTGGATCCAGCAGG		649
Query 600	CGGGG	604		
Sbjct 650	CGGGG	654		

**Figura 15.** Alineamiento de la secuencia del espécimen de esponja, ID: 301 con *Aplysina fistularis*. El recuadro rojo simboliza el único gap encontrado entre nuestra secuencia y la reportada en la base de datos nr/nt de NCBI.

Max score	Total score	Query cover	E value	Ident	Accession
-----------	-------------	-------------	---------	-------	-----------

[Aplysina gerardogreeni voucher LEB-ICML-UNAM-2042 cytochrome oxidase subunit I \(COI\) gene, partial cds; mitoch](#) 966 966 86% 0.0 100% [JQ437578.1](#)

Aplysina gerardogreeni voucher LEB-ICML-UNAM-2042 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial

Sequence ID: [JQ437578.1](#) Length: 523 Number of Matches: 1

[▶ See 1 more title\(s\)](#)

**Related Information**

Range 1: 1 to 523 [GenBank](#) [Graphics](#) [▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Gaps	Strand
966 bits(523)	0.0	523/523(100%)	0/523(0%)	Plus/Plus
Query 25	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA	84		
Sbjct 1	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA	60		
Query 85	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATTCFTTTTAGTAATGCC	144		
Sbjct 61	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATTCFTTTTAGTAATGCC	120		
Query 145	TGTTATGATAGGGGTTTGGTAATTGATTAGTTCATTATATATGGAGCACCAGATAT	204		
Sbjct 121	TGTTATGATAGGGGTTTGGTAATTGATTAGTTCATTATATATGGAGCACCAGATAT	180		
Query 205	GGCTTCCCTCGGTTAAATAATATAGTTTTGATTATTACCCCAGCACTAATTTTATT	264		
Sbjct 181	GGCTTCCCTCGGTTAAATAATATAGTTTTGATTATTACCCCAGCACTAATTTTATT	240		
Query 265	ATTGGGTCTGCTTTTATAGAACAAGGGCCGGTACGGTTGAACCGTCTATCCGCCATT	324		
Sbjct 241	ATTGGGTCTGCTTTTATAGAACAAGGGCCGGTACGGTTGAACCGTCTATCCGCCATT	300		
Query 325	ATCAAGTATACAAACACATTCGGGGTTCGGTAGATATGGCAATTTAGTCTTCATTT	384		
Sbjct 301	ATCAAGTATACAAACACATTCGGGGTTCGGTAGATATGGCAATTTAGTCTTCATTT	360		
Query 385	GGCCGGAATTCCTCCATATTAAGCTCTATTAATTTATAACTACAATTTAAATATGAG	444		
Sbjct 361	GGCCGGAATTCCTCCATATTAAGCTCTATTAATTTATAACTACAATTTAAATATGAG	420		
Query 445	GGCGCCAGGATGACTATGGACAGATTACCTTTATTCGTTGATCTATTTTATAACAAC	504		
Sbjct 421	GGCGCCAGGATGACTATGGACAGATTACCTTTATTCGTTGATCTATTTTATAACAAC	480		
Query 505	AATATTATTAGTATTAGCATTGCCGTTTGTAGCTGGAGCAATT	547		
Sbjct 481	AATATTATTAGTATTAGCATTGCCGTTTGTAGCTGGAGCAATT	523		

Figura 16. Alineamiento de la secuencia del espécimen de esponja, ID: 301 con *Aplysina gerardogreeni*.

Max score	Total score	Query cover	E value	Ident	Accession
-----------	-------------	-------------	---------	-------	-----------

[Aplysina revillagigedi voucher LEB-ICML-UNAM-1236 cytochrome oxidase subunit I \(COI\) gene, partial cds; mitochor](#) 961 961 86% 0.0 99% [JQ437580.1](#)

Aplysina revillagigedi voucher LEB-ICML-UNAM-1236 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial

Sequence ID: [JQ437580.1](#) Length: 523 Number of Matches: 1

[Related Information](#)

Range 1: 1 to 523 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
961 bits(520)	0.0	522/523(99%)	0/523(0%)	Plus/Plus
Query 25	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA			84
Sbjct 1	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA			60
Query 85	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATCTTTTGTAGTAATGCC			144
Sbjct 61	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATCTTTTGTAGTAATGCC			120
Query 145	TGTTATGATAGGGGTTTGGTAATTGATTAGTTCATTATATATTGGAGCACCAGATAT			204
Sbjct 121	TGTTATGATAGGGGTTTGGTAATTGATTAGTTCATTATATATTGGAGCACCAGATAT			180
Query 205	GGCTTCCCTCGGTTAAATAATATTAGTTTTGATTATTACCCAGCAGCTAATTTTATT			264
Sbjct 181	GGCTTCCCTCGGTTAAATAATATTAGTTTTGATTATTACCCAGCAGCTAATTTTATT			240
Query 265	ATTGGGTCGCTTTTATAGAACAAGGGCCGGTACGGTTGAACCGTCTATCCGCCATT			324
Sbjct 241	ATTGGGTCGCTTTTATAGAACAAGGGCCGGTACGGTTGAACCGTCTATCCGCCATT			300
Query 325	ATCAAGTATACAACACATCTGGGGTTCGGTAGATATGGCAATTTAGTCTTCATTT			384
Sbjct 301	ATCAAGTATACAACACACTCTGGGGTTCGGTAGATATGGCAATTTAGTCTTCATTT			360
Query 385	GGCCGGAATTCCTCCATATTAAGCTCTATTAATTTTATAACTACAATTTAAATATGAG			444
Sbjct 361	GGCCGGAATTCCTCCATATTAAGCTCTATTAATTTTATAACTACAATTTAAATATGAG			420
Query 445	GGCGCCAGGGATGACTATGGACAGATTACCTTTATTCGTTTGATCTATTTTATTAACAAC			504
Sbjct 421	GGCGCCAGGGATGACTATGGACAGATTACCTTTATTCGTTTGATCTATTTTATTAACAAC			480
Query 505	AATATTATAGTATTAGCATTGCCCGTTTGTAGCTGGAGCAATT			547
Sbjct 481	AATATTATAGTATTAGCATTGCCCGTTTGTAGCTGGAGCAATT			523

Figura 17. Alineamiento de la secuencia del espécimen de esponja, ID: 301 con *Aplysina revillagigedi*.

Max score	Total score	Query cover	E value	Ident	Accession
961	961	100%	0.0	99%	Query_44603

lcl|Query\_90133:54-657 SPG301-39-R

Sequence ID: Query\_44603 Length: 604 Number of Matches: 1

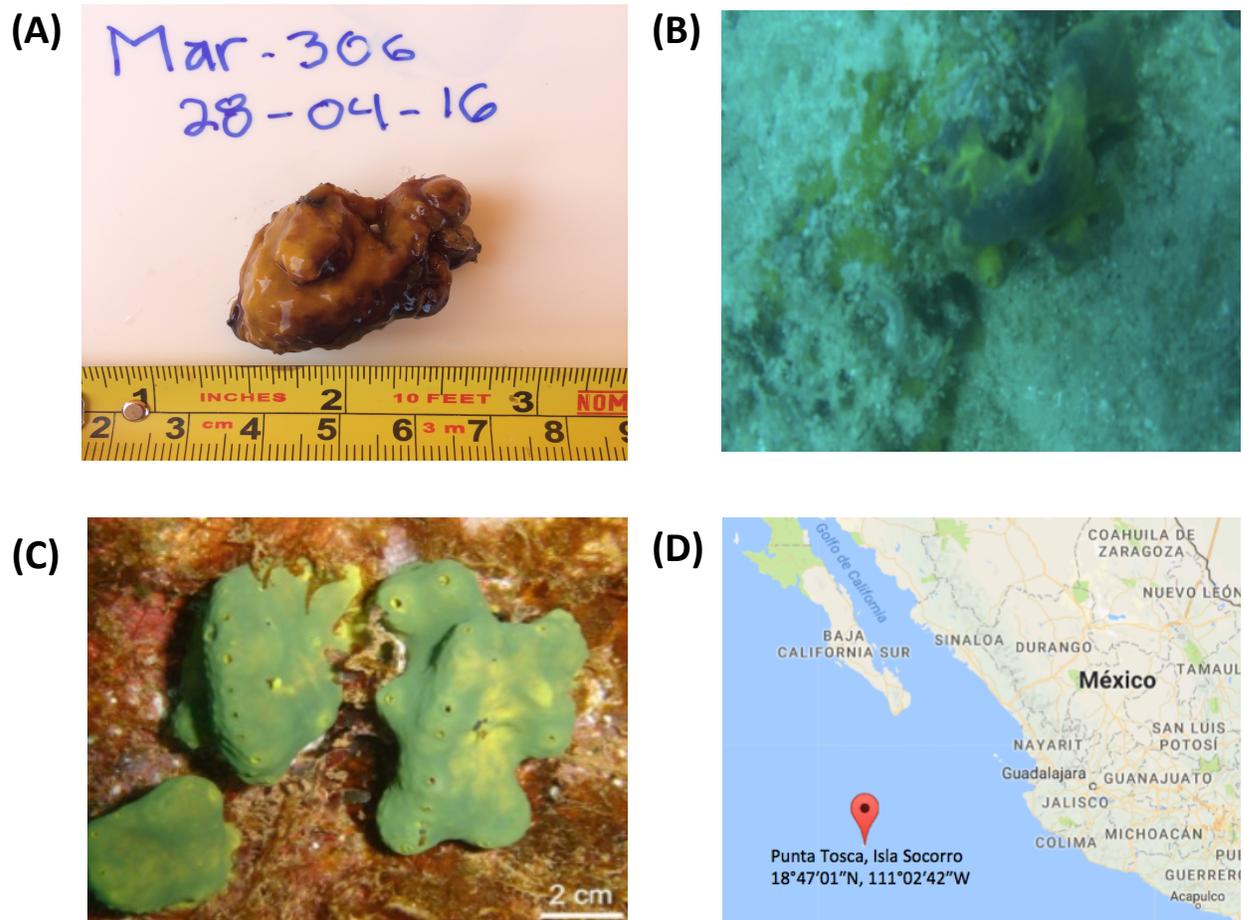
Range 1: 25 to 547 [Graphics](#)

▼ Next Match ▲ Previous Match

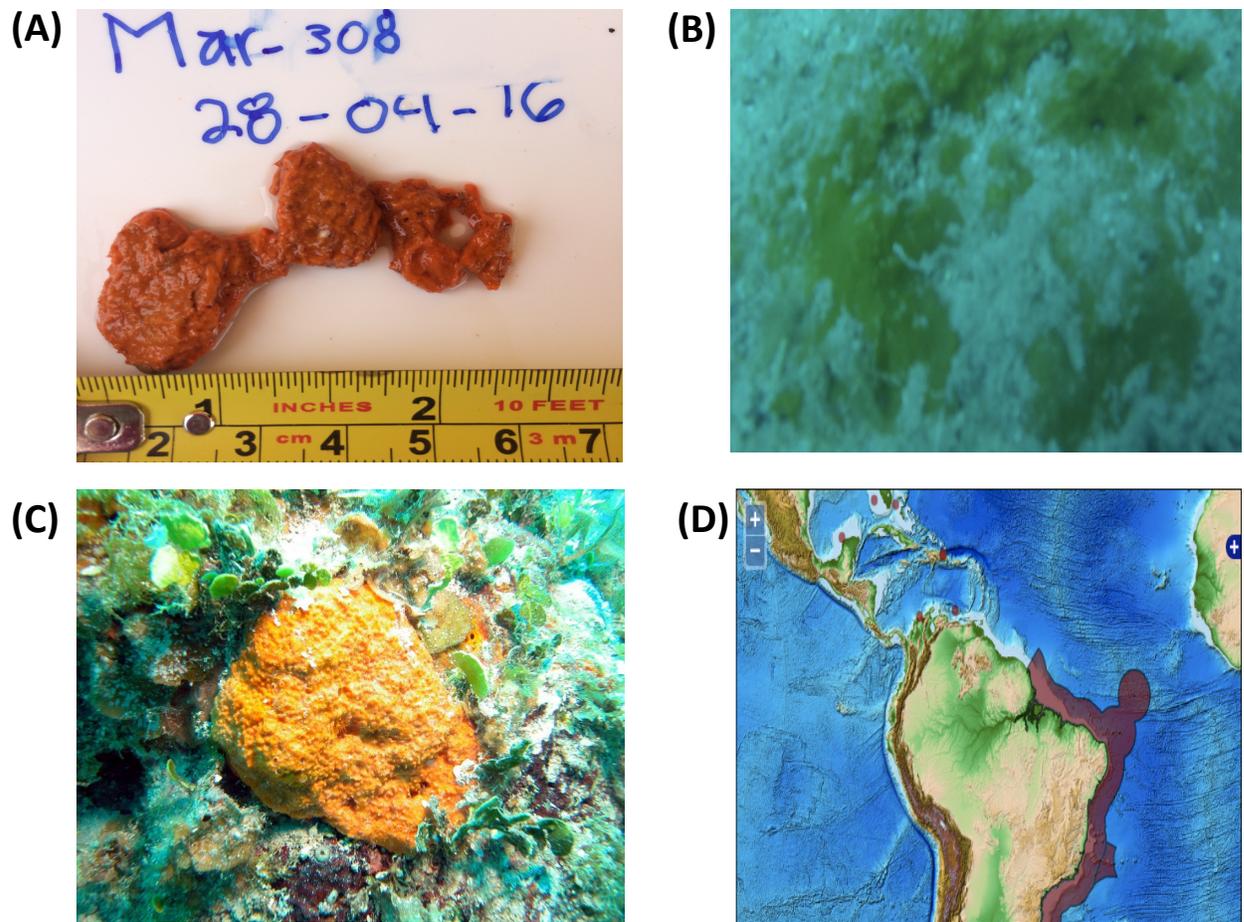
Score	Expect	Identities	Gaps	Strand
961 bits(520)	0.0	522/523(99%)	0/523(0%)	Plus/Plus
Query 1	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA	60		
Sbjct 25	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA	84		
Query 61	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATTCTTTTAGTAATGCC	120		
Sbjct 85	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATTCTTTTAGTAATGCC	144		
Query 121	TGTTATGATAGGGGGTTTGGTAATTGATTAGTTCCATTATATATTGGAGCACCAGATAT	180		
Sbjct 145	TGTTATGATAGGGGGTTTGGTAATTGATTAGTTCCATTATATATTGGAGCACCAGATAT	204		
Query 181	GGCTTTCCTCGGTAAATAATATAGTTTTGATTATTACCCCAGCACTAATTTTATT	240		
Sbjct 205	GGCTTTCCTCGGTAAATAATATAGTTTTGATTATTACCCCAGCACTAATTTTATT	264		
Query 241	ATTGGGGTCTGCTTTTATAGAACAAGGGCCGGTACGGGTGAACCGTCATCCGCCATT	300		
Sbjct 265	ATTGGGGTCTGCTTTTATAGAACAAGGGCCGGTACGGGTGAACCGTCATCCGCCATT	324		
Query 301	ATCAAGTATACAAACACACTCTGGGGTTCGGTAGATATGGCAATATTTAGTCTTCATTT	360		
Sbjct 325	ATCAAGTATACAAACACATTCTGGGGTTCGGTAGATATGGCAATATTTAGTCTTCATTT	384		
Query 361	GGCCGGAATTTCCATATTAAGCTCTATTAATTTATAACTACAATTTAAATATGAG	420		
Sbjct 385	GGCCGGAATTTCCATATTAAGCTCTATTAATTTATAACTACAATTTAAATATGAG	444		
Query 421	GGCGCCAGGGATGACTATGGACAGATTACCTTTATTGTTGATCTATTTTATTAACAAC	480		
Sbjct 445	GGCGCCAGGGATGACTATGGACAGATTACCTTTATTGTTGATCTATTTTATTAACAAC	504		
Query 481	AATATTATTAGTATTAGCATTGCCGTTTTAGCTGGAGCAATT	523		
Sbjct 505	AATATTATTAGTATTAGCATTGCCGTTTTAGCTGGAGCAATT	547		

**Figura 18.** Alineamiento obtenido de realizar un *Blast 2 sequences* de la secuencia de nucleótidos del espécimen, ID: 301 con la secuencia de *A. revillagiedi* (número de acceso: JQ437580.1).

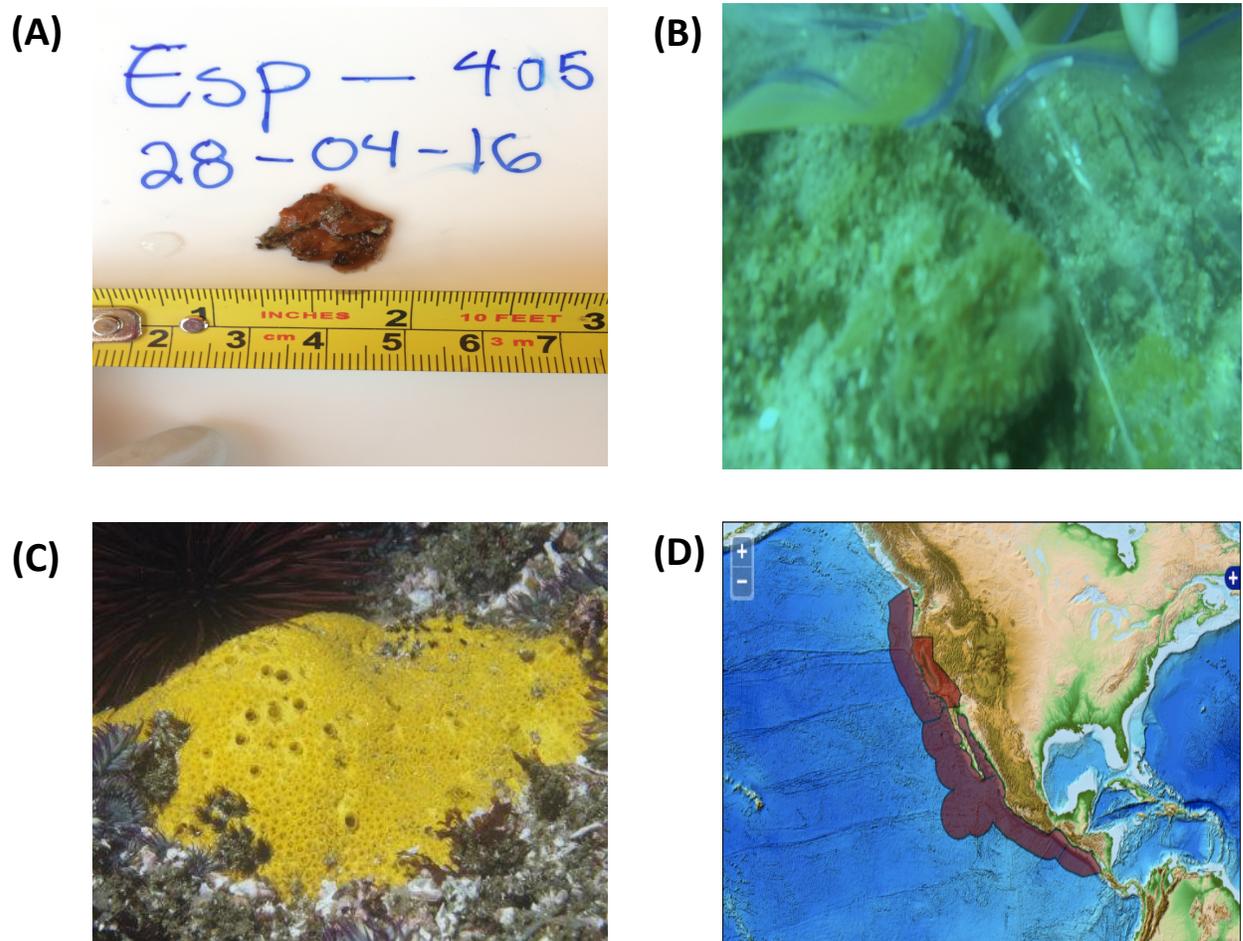
A continuación, se muestra un registro fotográfico para cada una de las 3 especies de esponja más abundantes del PNCP (fig.12), junto con la especie que se propone que es, con base en su registro fotográfico y la distribución geográfica de la especie sugerida (fig.19 –fig.21).



**Figura 19.** Registro fotográfico de la de la esponja identificada como *Aplysina revillagigedi* (A) muestra su aspecto fuera del agua; (B) dentro del agua; (C) muestra una fotografía de registro para la especie (recuperada de Cruz-Barraza *et al.*, 2012); (D) muestra la zona geográfica donde se ha reportado su presencia (Cruz-Barraza *et al.*, 2012).



**Figura 20.** Registro fotográfico de la de la esponja identificada como *Prosuberites laughlini* (A) muestra su aspecto fuera del agua; (B) muestra su aspecto dentro del agua; (C) muestra una fotografía de registro para la especie (recuperada de [spongeguide.org](http://spongeguide.org)); (D) distribución geográfica de la especie (recuperada de [marinespecies.org](http://marinespecies.org)).



**Figura 21.** Registro fotográfico de la de la esponja identificada como *Cliona celata* (también conocida como *Cliona californiana* (marinespecies.org) (A) muestra su aspecto fuera del agua; (B) dentro del agua; (C) muestra una fotografía de registro para la especie (fotografía recuperada de studyblue.com); (D) distribución geográfica de la especie (recuperada de marinespecies.org).

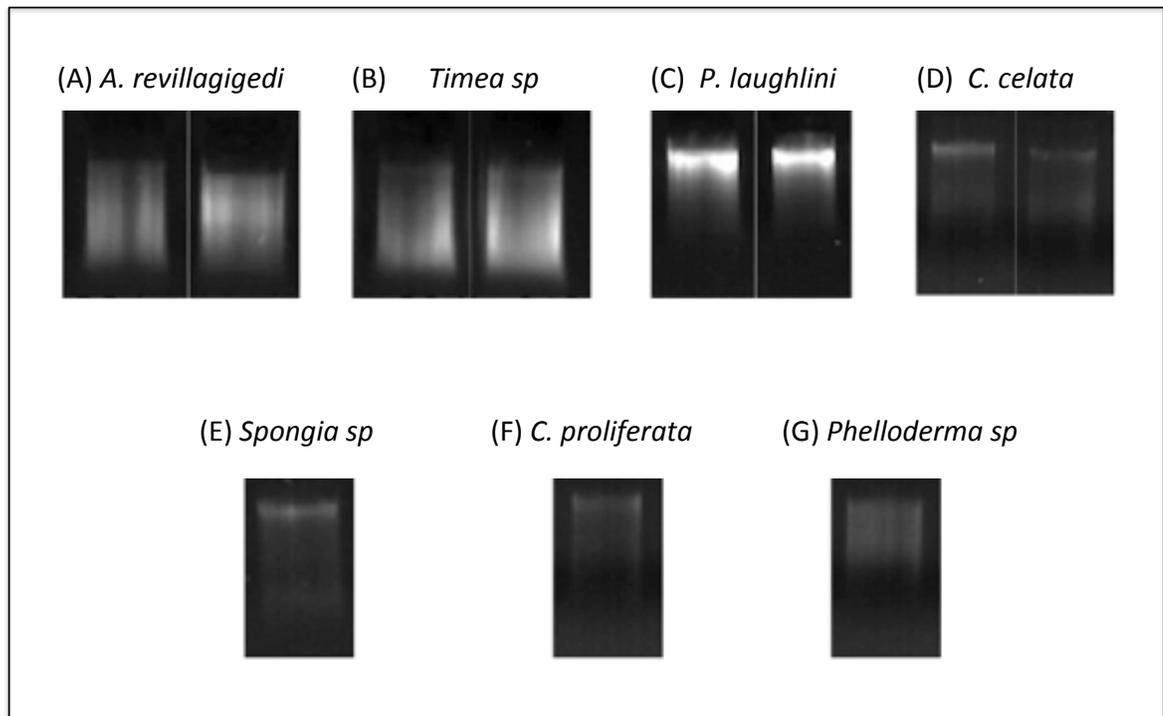
De entre las tres especies más abundantes, la esponja sugerida como *Aplysina revillagigedi* fue la más comúnmente encontrada (se recolectaron 33 ejemplares en las 4 estaciones muestreadas: CAN, MAR, ESP y NAV) a lo largo del parque. *Aplysina revillagigedi*, se describe en la literatura como una esponja de tipo masiva (entre 0.5 a 3cm de altura), que puede variar entre un color verde a amarillo limón, y que al ser preservada cambia a un color más oscuro (tal como se observó en nuestras muestras). Se encuentra cerca de la costa Este del Océano Pacífico Mexicano, específicamente se ha encontrado en las islas Isla Socorro, e Isla Clarion, así como en el archipiélago Revillagigedo, por el cual se le da su nombre. Habita aguas someras como profundas (4-30m), y suele encontrarse creciendo sobre sustratos rocosos expuestos a una alta agitación del agua (Cruz-Barraza *et al.*, 2012).

Le sigue *Cliona celata*, de la cual se se encontraron 12 ejemplares, pero exclusivamente en dos estaciones: CAN y MAR. Esta esponja se observa típicamente con una parte expuesta caracterizada por la presencia de papilas amarillas, mientras que el resto de la esponja permanece oculta dentro del substrato sobre el que crece. Su tamaño es de al menos 300mm en diámetro. Su distribución es cosmopolita, pues puede encontrarse desde Alaska hasta Baja California. Excava dentro de organismos marinos con concha, gracias a unas células que secretan material que fragmenta pequeños pedazos del substrato al que esta adherido la esponja. Puede encontrarsele habitando la zona intermareal hasta 120m de profundidad (Gotshall, 2005).

Finalmente, *Prosuberites laughlini*, de la cual se identificaron 7 ejemplares entre las muestras colectadas, únicamente en las estaciones: ESP y NAV. Esta esponja tiene una distribución en el Caribe sur, al igual que *Clathria prolifera*, otra esponja identificada en el parque.

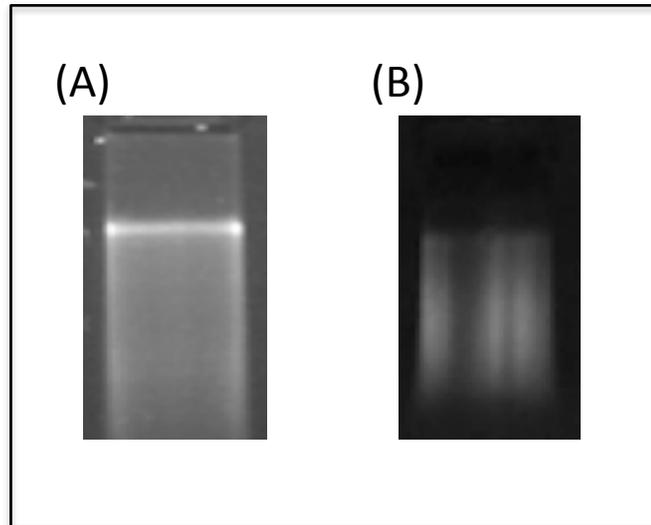
Los resultados del análisis de la distribución geográfica de las demás esponjas recolectadas, por las bases de datos: WoRMS, y *Sponge Guide* indican que presentan una distribución a lo largo de la costa Este del Océano Pacífico. Todas las especies identificadas pertenecen a la Clase Demospongiae. Esta clase de esponjas puede presentar diversas morfologías entre ellas: masiva (caracterizada por tener tamaño grande y forma protuberante, ej. en tubo), incrustante (que adoptan la forma del sustrato que sobre yacen) y excavadora (que erosionan el sustrato y crecen hacia dentro de él) (Silva, 2015). Se observaron los 3 tipos de morfologías en las esponjas aquí analizadas: 1 excavadora (*Cliona celata*); 1 masiva (*Aplysina revillagigedi*), y 5 incrustantes (*Clathria prolifera*, *Prosuberites laughlini*, *Phelloderma sp*, *Spongia sp* y *Timea sp*).

Por último, en cuestiones técnicas, se observó que la integridad y la cantidad del ADN extraído, tendía a variar dependiendo de la especie de esponja, el tipo de homogenización (mecánica vs tubo con perlas) y el uso o no de FCI. Las esponjas identificadas como: *Aplysina revillagigedi*, *Timea sp*, y *Phelloderma sp* fueron las que presentaron el ADN más fragmentado o barrido en los geles de electroforesis (Fig.22), a comparación de las demás esponjas identificadas, las cuales presentaron un ADN con un aspecto más íntegro, representado por una ligera banda que se puede apreciar en la parte superior del gel para algunas de estas muestras, estas esponjas fueron: *Prosuberites laughlini*, *Cliona celata*, *Spongia sp* y *Clathria prolifera*.



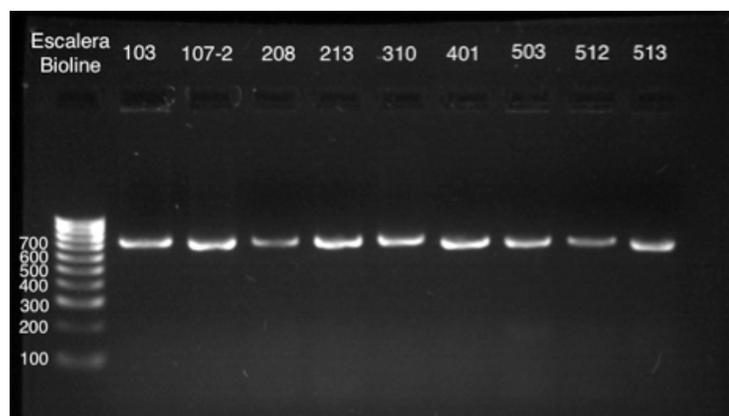
**Figura 22.** Resultados de la extracción del ADN de las 4 especies y 3 géneros de esponjas identificadas: (A) *Aplysina revillagigedi*; (B) *Timea sp*; (C) *Prosuberites laughlini*; (D) *Cliona celata*; (E) *Spongia sp*; (F) *Clathria proliferata*; (G) *Phelloderma sp*. Para las muestras A-D, se muestran las extracciones de las réplicas biológicas.

Por su parte, se realizó una breve prueba con un ejemplar de la esponja *Aplysina revillagigedi* (ID espécimen: 404) cuya región COI amplificaba muy poco, para ver si la utilización de un homogenizador mecánico para macerar la muestra generaba una mejor extracción de ADN, a comparación de homogenizar la muestra utilizando el tubo con perlas del kit de MOBIO. Se observó que la homogenización mecánica es una buena opción para mejorar la calidad del ADN (A260/A280: 2.03 vs A260/A280: 1.91), así como la cantidad (85.65 ng/ $\mu$ L vs 76.7 ng/ $\mu$ L) (Fig.23). Sin embargo, cabe mencionar que para la mayoría de las esponjas, la calidad obtenida de realizar las extracciones con el tubo con perlas fue suficiente para amplificar satisfactoriamente su gen COI y realizar su identificación molecular.



**Figura 23.** ADN del espécimen de *Aplysina revillagigedi* (ID espécimen: 404, y en gráficas del análisis de microbiomas se refiere como: SPG.APLY.ESP.3) que se utilizó para realizar una prueba comparativa entre la extracción de ADN con homogenización mecánica y fenol:cloroformo:isoamil alcohol (25:24:1) (A), en comparación a la extracción de ADN con homogenización en tubo con perlas y FCI (B). Condiciones de gel: 0.8% de agarosa, 80 V por 20 min. Se utilizó la muestra 404 con homogenización mecánica para la secuenciación.

A continuación, se muestra la calidad de los amplicones COI generados para las especies de esponja más abundantes del PNCP: *Cliona celata* (103 y 107.2), *Aplysina revillagigedi* (208, 213, 401, 503, 512, 513), *Prosuberites laughlini* (310).



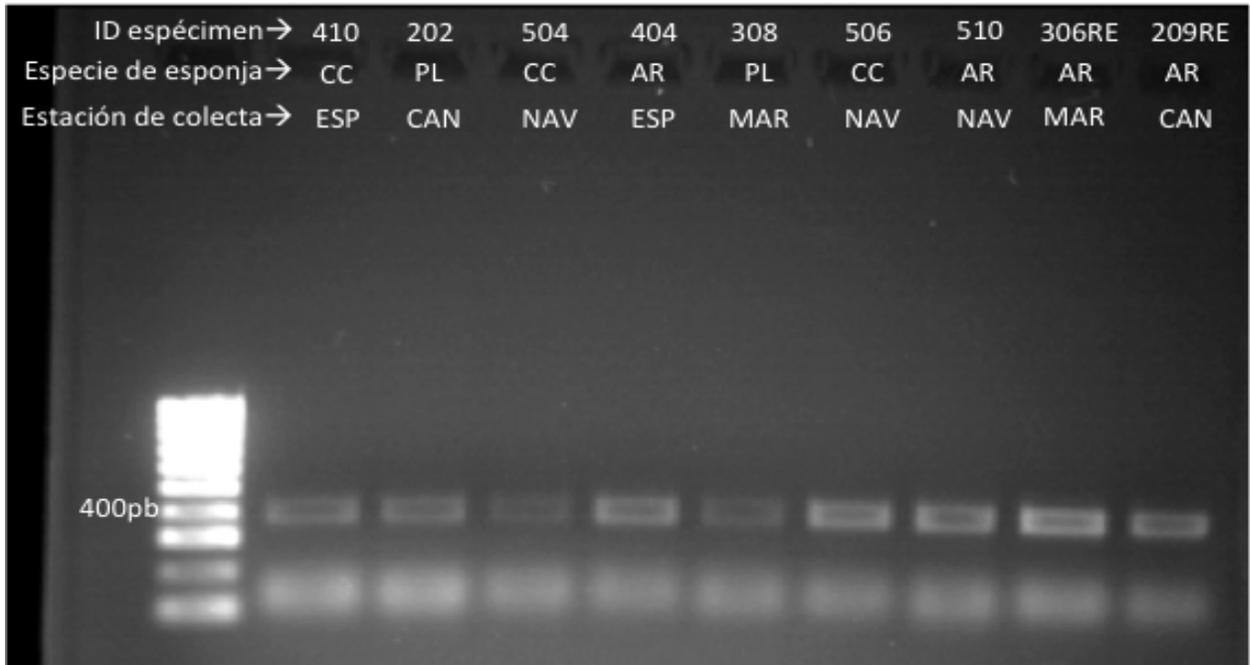
**Figura 24.** Amplicones del gen COI de ejemplares de las 3 esponjas más abundantes del PNCP: *Cliona celata* (103 y 107.2), *Aplysina revillagigedi* (208, 213, 401, 503, 512, 513), y *Prosuberites laughlini* (310). El tamaño de las bandas coincide con el esperado de 680 pb según Meyer *et al.*, (2003). Condiciones de gel: 1.2% de agarosa, 100V por 40 min.

### 3.3 Análisis de los microbiomas mediante amplificación del gen 16S del ARNr de muestras de sedimento, agua y esponjas.

A continuación, se muestran los resultados de amplificar el gen 16S del ARNr con 25 ciclos y 20 ng/ $\mu$ L de ADN proveniente de las 3 especies de esponja más abundantes del PNCP, estas fueron: *Aplysina revillagigedi*, *Cliona celata* y *Prosuberites laughlini*. Por cada sitio de colecta se amplificaron 3 réplicas biológicas de cada especie. Así también se amplificaron muestras de agua y sedimento (n=3) proveniente de cada estación de colecta. Los amplicones generados fueron secuenciados en dos eventos de secuenciación: SEQ1 y SEQ2. La información relevante a la calidad de cada secuenciación se puede encontrar en la Tabla 8 de Anexos.

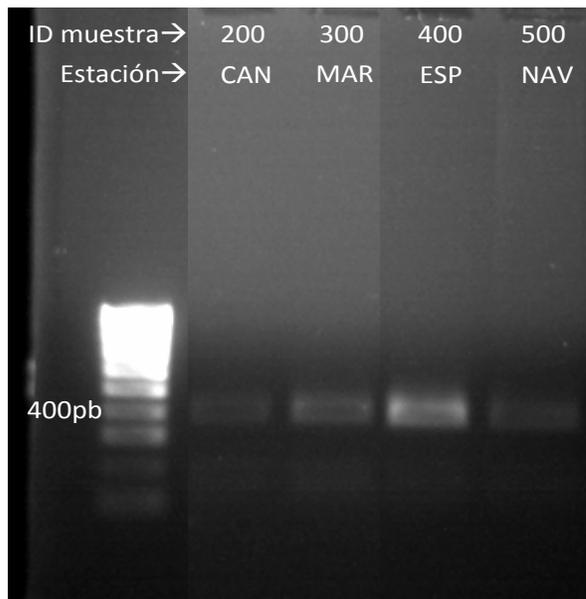
#### 3.3.1 Amplicones de la SEQ1

Las muestras secuenciadas en SEQ1 incluyeron: 4 muestras de *Cliona celata* (2 muestras provenientes de cada estación de la cual se colectó la esponja), estas fueron: 410, 405, 506 y 504; 4 muestras de *Aplysina revillagigedi* (una de cada estación de la cual se colectó): 209RE, 306RE (RE: re-extracción de la original), 404 y 510; y 2 de *Prosuberites laughlini*: 202 y 308 (Fig.25). Así también se amplificaron 4 muestras de sedimento (SD), estas fueron: CAN (ID: 200), MAR (ID: 300), ESP (ID: 400) y NAV (ID: 500), y 4 muestras de agua: CAN (ID: 3 y 5), MAR (ID: 2), ESP (ID: 4), provenientes de cada una de las estaciones donde se recolectaron esponjas (Fig.26). La calidad de sus amplificaciones se muestra a continuación. Todas las muestras amplificaron bien, sin inespecificidades, sólo hubo diferencias en relación a la intensidad de los productos obtenidos, y estas diferencias fueron independientes de la naturaleza de la muestra (esponja, sedimento o agua) o especie de esponja (abreviadas en las figuras como: AR (*Aplysina revillagigedi*), PL (*Prosuberites laughlini*), y CC (*Cliona celata*). El tamaño del amplicon generado (entre 400pb -300pb) corresponde con el esperado (380pb).

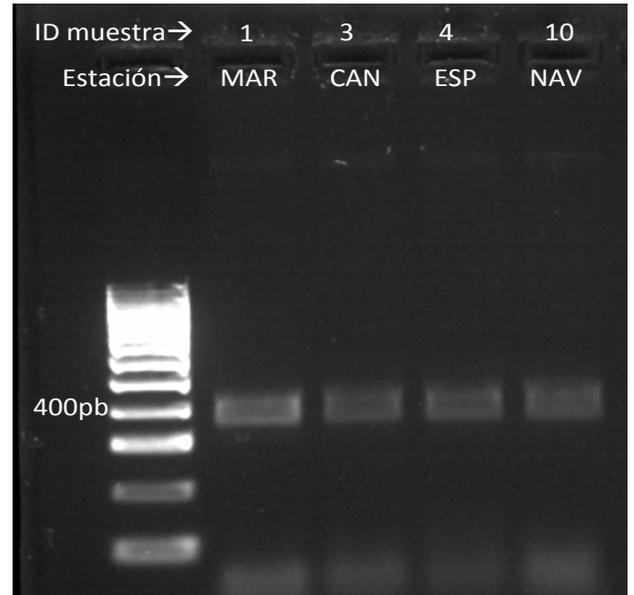


**Figura 25.** Amplicones de diferentes especies de esponja (CC: *Cliona celata*, PL: *Prosuberites laughlini*, AR: *Aplysina revillagigedi*) provenientes de diferentes estaciones (CAN, MAR, ESP y NAV), secuenciados en la SEQ1. La reacción PCR fue realizada a 25 ciclos con 20 ng/μL de concentración de ADN para amplificar la región V4 del gen 16S del ARNr. Se observó un amplicón de tamaño entre 400-300 pb, el fragmento esperado era de 380 pb.

(A)



(B)

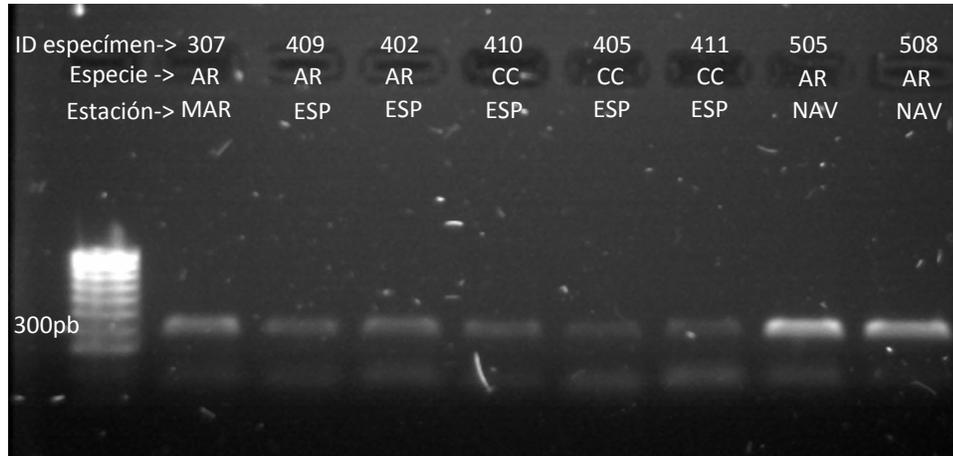


**Figura 26.** (A) Amplicones de muestras de sedimento y (B) agua, provenientes de las siguientes estaciones: MAR, CAN, ESP y NAV, secuenciados en la SEQ1. La reacción PCR fue realizada a 25 ciclos con 20 ng/μL de concentración de ADN para amplificar la región V4 del gen 16S del ARNr; se observó un amplicón de tamaño entre 400-300 pb, el fragmento esperado era de 380 pb.

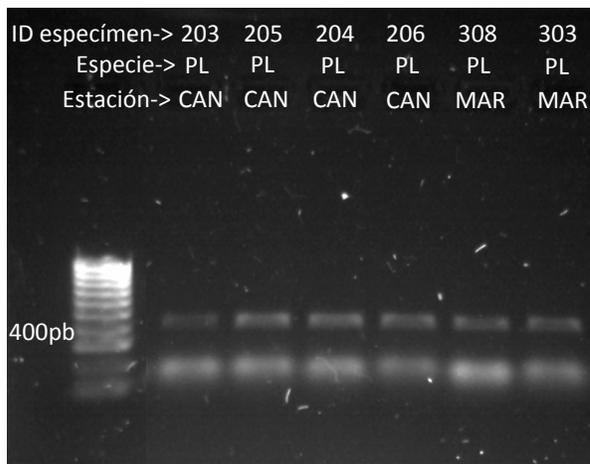
### 3.3.2 Amplicones de SEQ2

Las muestras secuenciadas en SEQ2 fueron dos especímenes de cada especie de esponja provenientes de cada estación de colecta, que en conjunto con la SEQ 1, sumaron un total de 3 réplicas biológicas por cada estación de la cual se recolectaron esponjas de cada especie. Las muestras de esponja que fueron secuenciadas en SEQ2 fueron las siguientes: *Cliona celata*: 410, 405, 411, 506 y 502, *Prosuberites laughlini*: 202, 203, 205, 308, 303 y 311; *Aplysina revillagigedi*: 204, 206, 301, 307, 409, 402, 505, y 508, la calidad de sus amplicones se muestra en la Fig.27. Así también se muestran las calidades de los amplicones generados para las submuestras de sedimento y agua que fueron secuenciadas en la SEQ2 (Fig.28). Cabe mencionar que, en algunos casos de las muestras de esponja, fue necesario aumentar la concentración de ADN (de 20 ng/μL a 30 ng/μL) para generar suficiente cantidad del amplicon (inciso C, fig. 27).

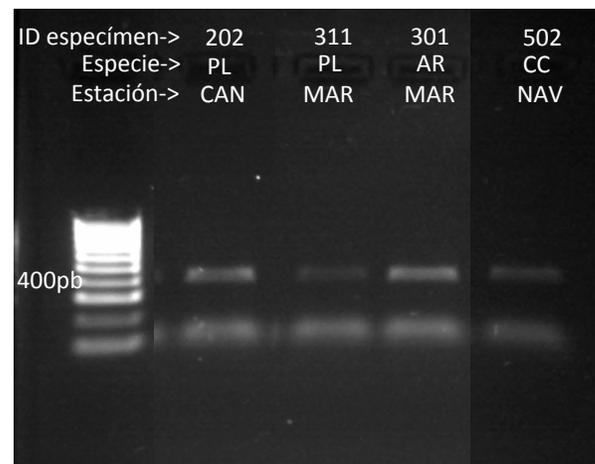
(A)



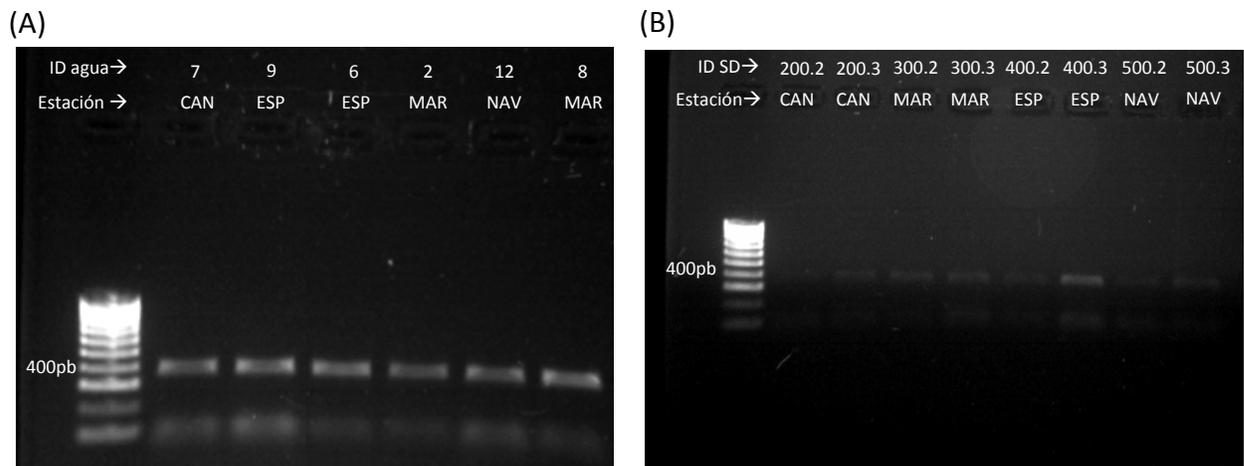
(B)



(C)



**Figura 27.** Resultados de la reacción PCR realizada para amplificar la región V4 del gen 16S del ARNr. El PCR se realizó a 25 ciclos con 20 ng/μL de concentración de ADN para los amplicones de las especies de esponja: AR: *Aplysina revillagigedi* y CC: *Cliona celata*, provenientes de las estaciones: MAR, ESP y NAV (A) y los amplicones de especímenes de la esponja, PL: *Prosuberites laughlini*, provenientes de las estaciones: CAN y MAR (B). Los amplicones de especímenes cuya concentración de ADN se incrementó a 30 ng/μL por motivos técnicos se muestran en (C). En todos los geles se observó un amplicón de tamaño entre 400-300 pb, que corresponde con el fragmento esperado de 380 pb.



**Figura 28.** Amplicones de las muestras de agua (A) y sedimento (B) provenientes de las estaciones: CAN, ESP MAR y NAV, que fueron secuenciadas en SEQ2. Las muestras de agua fueron amplificadas utilizando una concentración de ADN de 10 ng/μL, y las de sedimento de 20 ng/μL de ADN. Los PCRs realizados fueron a 25 ciclos; se observó un amplicón de tamaño entre 400-300 pb (el fragmento esperado era de 380pb).

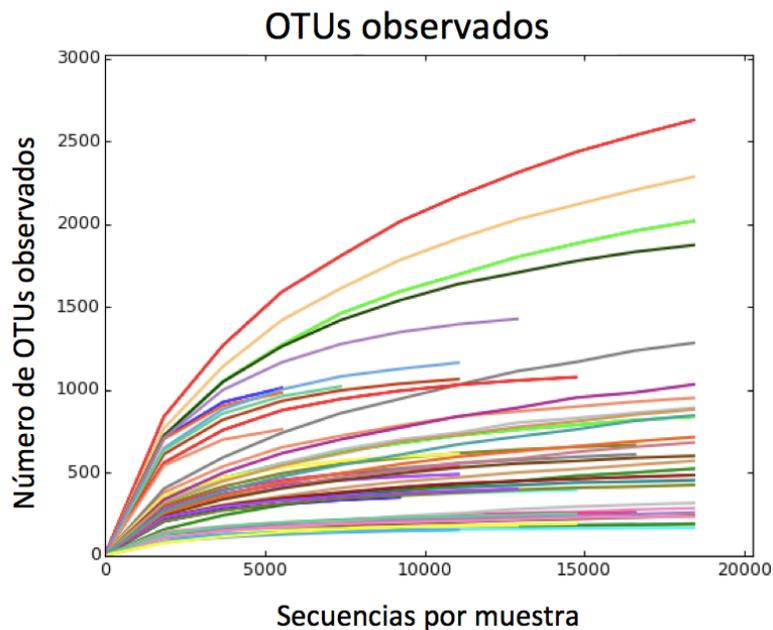
### 3.4 Resultados de secuenciación

Los datos procedentes de las dos secuenciaciones realizadas (SEQ1 y SEQ2) se juntaron en un único archivo de metadatos y se analizaron en conjunto. En total se secuenciaron 53 muestras: 12 de sedimento, 12 de agua, 12 de la esponja *Aplysina revillagigedi*, 9 de la esponja *Cliona celata* y 8 de la esponja *Prosuberites laughlini*. Se obtuvieron un total de 1' 654,824 secuencias, por el esfuerzo de secuenciación de 2 secuenciaciones, estas secuencias fueron asignadas a 8,754 OTUs: 3730 OTUs en los microbiomas de esponja, y el resto fueron exclusivos de las muestras de sedimento y agua. En la tabla 2 se detallan los estadísticos básicos de la secuenciación por cada tipo de muestra, estos incluyen: el número total de secuencias asignadas por muestra; el número total de OTUs detectados a un 97% de similitud con la base de datos SILVA (contando "singletons" u OTUs con una sola secuencia asignada) y el OTU con mayor número de secuencias asignadas. Este último dato se obtuvo de sumar las secuencias presentadas por cada réplica de cada muestra. En orden descendente las esponjas que tuvieron mayor número de OTUs observados fueron: *C. celata* (3081 OTUs), le sigue *P. laughlini* (1209) y finalmente *A. revillagigedi* (962). El OTU más representado para *C. celata*, fue una Cianobacteria, para *P. laughlini* fue una Gammaproteobacteria, y para *A. revillagigedi* fue una Chloroflexi.

**Tabla 2.** Conteo de secuencias, número de OTUs totales y el OTU más representado en cada tipo de muestra: sedimento, agua, *A. revillagigedi*, *C. celata* y *P. laughlini*.

Muestras	# total de secuencias	#OTUs totales	OTU más representado
Sedimento	250530	6203	BD7-8 Grupo marino (Gammaproteobacteria)
Agua	518374	3602	Prochlorococcus (Cianobacteria)
<i>A. revillagigedi</i>	371107	962	Caldilineae (Cloroflexi)
<i>C. celata</i>	296368	3081	Sinecoccus (Cianobacteria)
<i>P. laughlini</i>	218445	1209	E01-9C-26 marine group (Gammaproteobacteria)

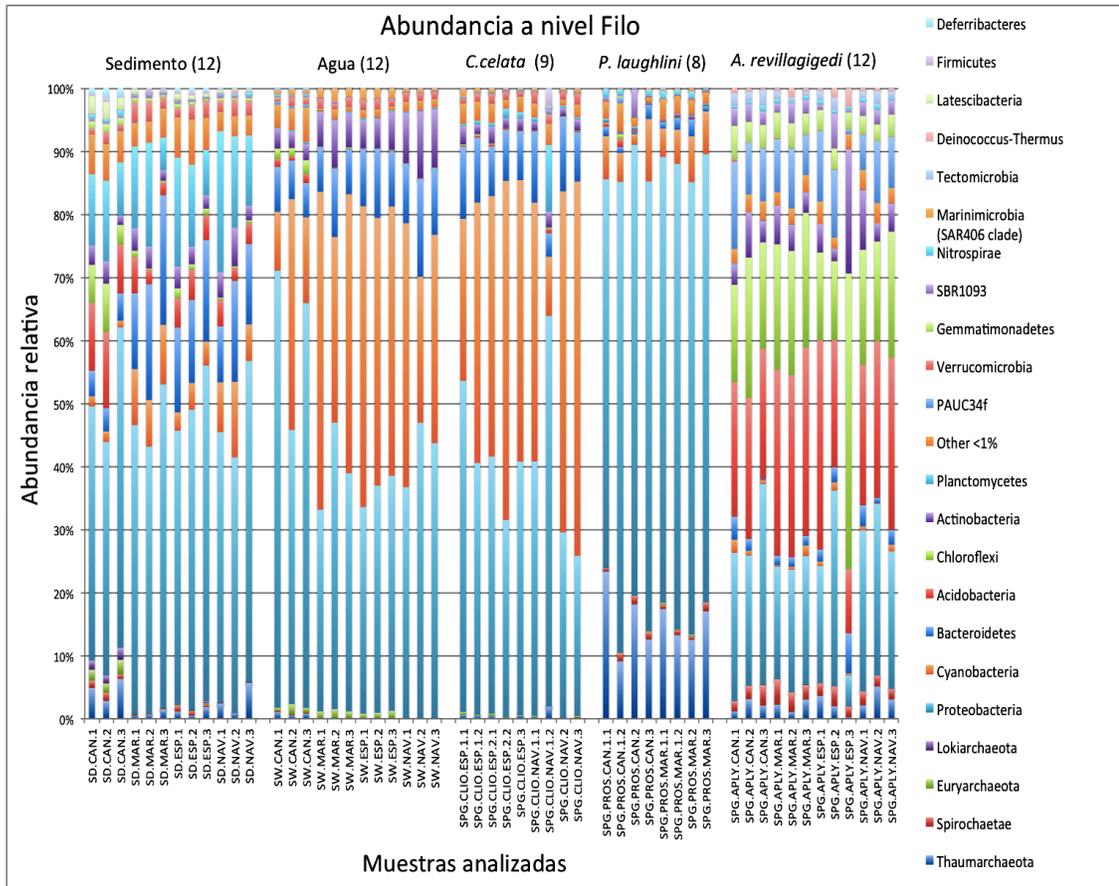
El esfuerzo de secuenciación fue satisfactorio para el análisis de las muestras, lo cual se refleja en la gráfica de rarefacción del número de especies observadas (Fig.29). La mayoría de las curvas alcanzan su asíntota o están cerca de alcanzarla, lo que implica que se extrajo prácticamente toda la información presente en la muestra.



**Figura 29.** Curvas de rarefacción del número de OTUs observados en las 53 muestras secuenciadas que indica que la totalidad de la diversidad de cada muestra está cerca de llegar a la asíntota, lo que implica que el esfuerzo de secuenciación ha sido satisfactorio.

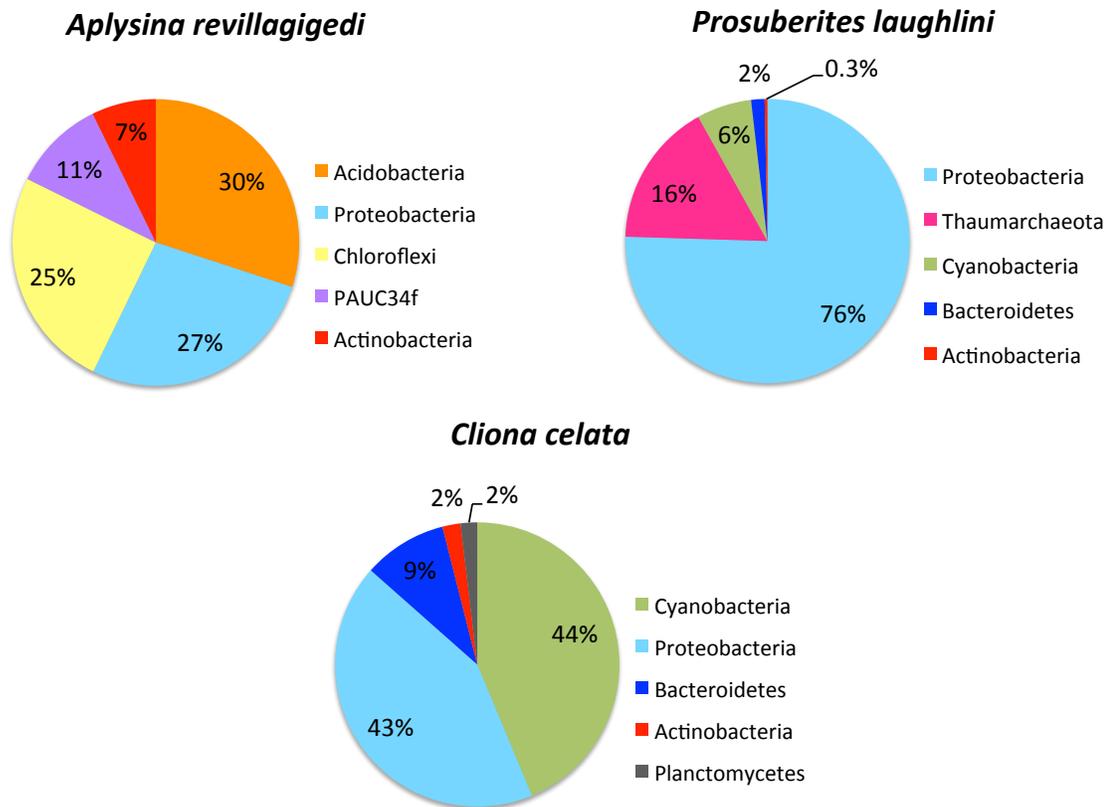
### 3.5 Análisis ecológicos

Se identificaron un total de 63 filo, de los cuales 30 estaban presentes en las esponjas, los demás fueron exclusivos del ambiente. En la figura 30 se muestran las abundancias relativas a nivel filo (para aquellos grupos presentes con una abundancia mayor al 1%), donde se pueden observar los perfiles microbianos de las 3 especies de esponjas (con sus respectivas réplicas, n=3 por estación), así como los del ambiente (muestras de agua y sedimento provenientes de cada zona de colecta). Dos cosas resaltan inmediatamente de esta gráfica: la primera de ellas, es que la estación de colecta no cambia la composición taxonómica del perfil microbiano en ninguna de las especies de esponjas, pues se pueden diferenciar claramente 3 perfiles distintos independiente del origen o el número de réplicas de cada esponja. La segunda es que el perfil microbiano de la esponja *Cliona celata* es muy similar al del agua. Esta esponja presentó la mayor riqueza de filo (35), a comparación de *Prosuberites laughlini* (30) y *Aplysina revillagigedi* (28), pero muy pocos de ellos presentaron abundancias mayores al 1%. Aún más, se observó que el filo Proteobacteria, fue el que presentó la mayor riqueza de especies en los tres microbiomas (*Cliona celata*: 418 especies distintas de Proteobacteria; *Prosuberites laughlini*: 211 y *Aplysina revillagigedi*: 181), estas especies pertenecieron mayoritariamente a las Clases Gammaproteobacteria y Alphaproteobacteria, siendo la especie *Endozoicomonas* perteneciente a la clase Gammaproteobacteria las más abundante en *Cliona celata*; la Gammaproteobacteria: E01-9c-26 en *Prosuberites laughlini*, y una especie no cultivada de Gammaproteobacteria en *Aplysina revillagigedi*. El filo Bacteroidetes, le siguió a las Proteobacterias, con 128 diferentes tipos de especies para *Cliona celata*, 64 para *Aplysina revillagigedi* y 54 para *Prosuberites laughlini*. Los filo Acidobacteria, Planctomicetes, Cianobacteria y Actinobacteria, presentaron un promedio de 27-19 especies distintas, respectivamente en las tres especies de esponja.



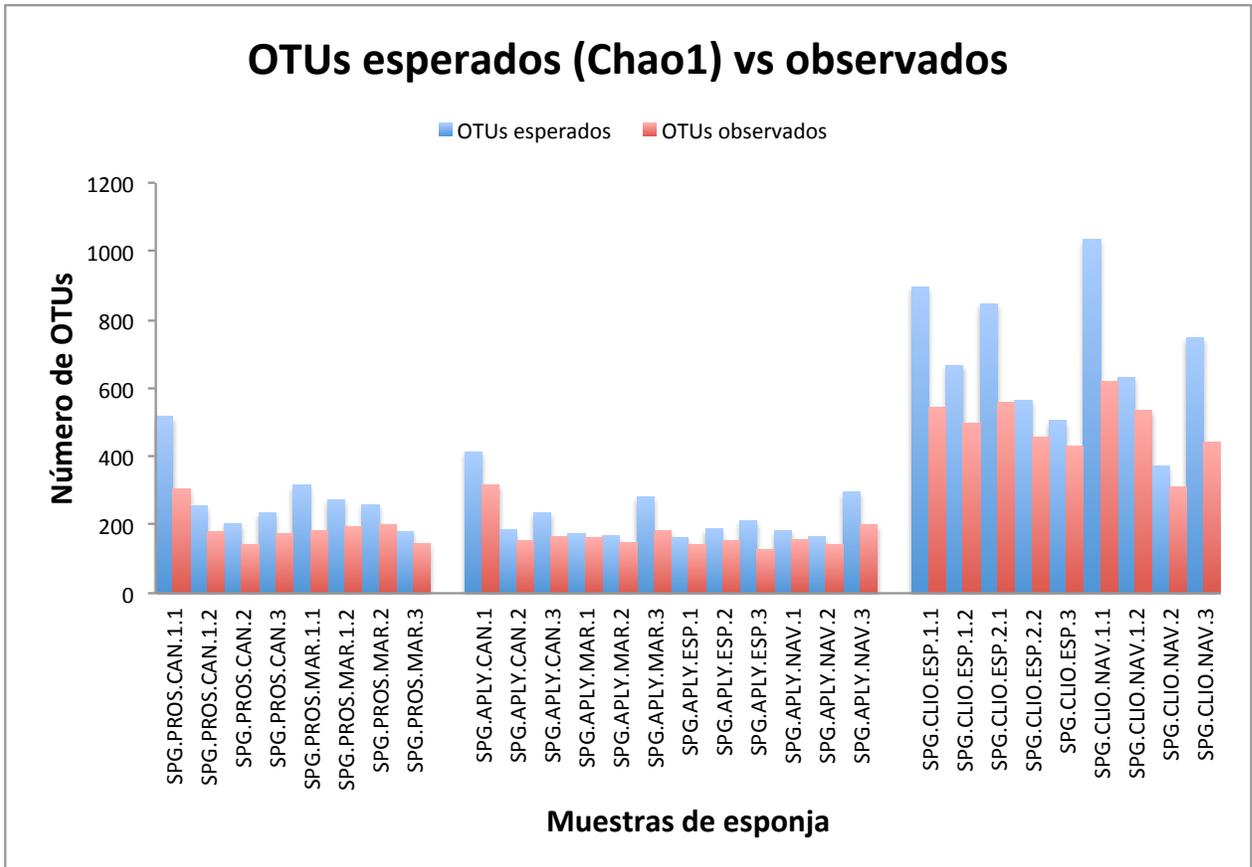
**Figura 30.** Abundancia relativa a nivel filo para todos aquellos filo con abundancias mayores al 1%. En la abscisa se muestran los nombres de las réplicas (de izquierda a derecha) de: sedimento (SD), agua (SW) y esponjas (SPG): *C. celata*, *P. laughlini* y *A. revillagigedi*; CAN (El cantil), MAR (Mario's Reef), ESP (La Esperanza) y NAV (Las Navajas), son las estaciones de colecta. La leyenda a la derecha (vista de abajo hacia arriba), ordena de manera descendente primero a las arqueas, y después a las bacterias.

En la figura 31, se muestran los 5 filo más abundantes de cada especie de esponja, con sus respectivos porcentajes, los cuales fueron calculados como el promedio de las abundancias a nivel filo del conjunto de réplicas de cada especie de esponja. Se observa que a pesar de que la composición taxonómica de cada microbioma de esponja se compone en términos generales de los mismos filo, estos no se presentan en las mismas abundancias en cada una de ellas. Para *Aplysina revillagigedi*, el filo más abundante fue el de las Acidobacterias. En contraste, para *Prosuberites laughlini* fueron las Proteobacterias, mientras que para *Cliona celata*, el filo de las Cianobacterias fue el más abundante.



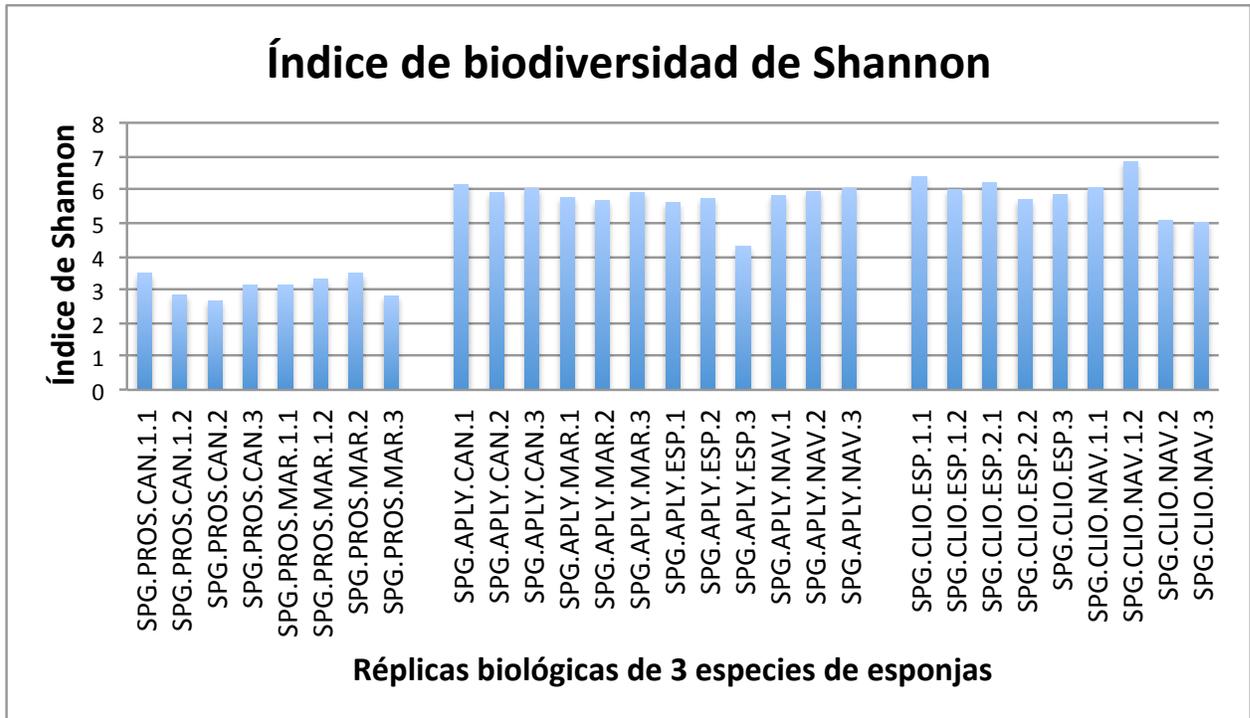
**Figura 31.** Los 5 fila más abundantes en cada especie de esponja. El porcentaje fue calculado a partir del promedio de las abundancias a nivel filo de las réplicas de cada especie de esponja.

En cuanto al cálculo de los índices de diversidad alfa (Chao1 y Shannon), se observó que el mayor número de OTUs esperados (Chao1) y observados en 5533 secuencias (valor utilizado para normalizar los valores comparados), lo presentó la esponja *Cliona celata* (Fig. 32). Esto significa que a comparación de *P. laughlini* y *A. revillagigedi*, *C. celata* probablemente tiene una mayor riqueza de especies de microorganismos asociados, los cuales no se lograron observar debido a que están presentes en muy bajas abundancias en su microbioma.



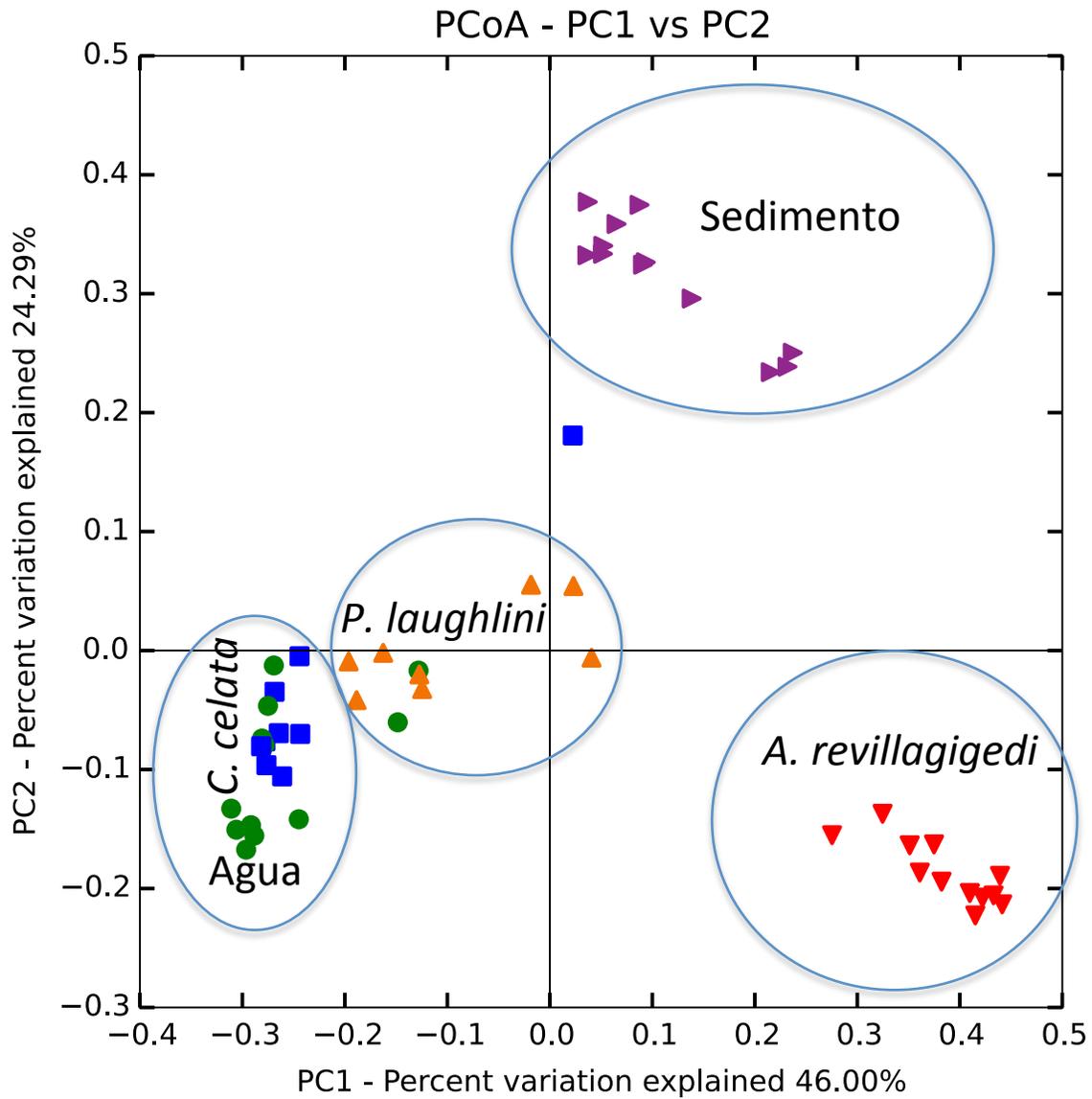
**Figura 32.** Comparación entre el número esperado de OTUs observados (Chao1, azul), y el número de OTUs observados (rojo) al mostrar 5533 secuencias (mínimo número de secuencias observados en una muestra). De izquierda a derecha se muestran los valores para *Prosuberites laughlini*, *Aplysina revillagiedgi* y *Cliona celata*.

Por otra parte, la esponja *Cliona celata* y *Aplysina revillagiedgi*, presentaron índices de biodiversidad de Shannon similares, estos fueron en promedio: 5.91 y 5.74, respectivamente. Ambos fueron mayores que el índice de Shannon de *Prosuberites laughlini* (3). Los valores mostrados del índice de Shannon en la figura 33, son representativos de 5533 secuencias.



**Figura 33.** Muestra el índice de Shannon para cada réplica de esponja, tomando como valor de normalización 5533 secuencias; de izquierda a derecha se muestran las réplicas de: *Prosuberites laughlini*, *Aplysina revillagigedi* y *Cliona celata*. El número de réplicas varía según el número de estaciones en que se encontró cada especie de esponja.

Como parte del análisis de diversidad beta realizado por QIIME, se realizó un análisis de coordenadas principales (PCoA por sus siglas en inglés), utilizando una matriz de distancia filogenética ponderada que se calculó utilizando el métrico Unifrac (Lozupone and Knight, 2005). Los resultados de este análisis muestran 4 agrupaciones (Fig. 34). Cada agrupación representa un microbioma único o filogenéticamente distinto. En general, se observa que las réplicas de la esponja de *C. celata* se agrupan casi en su totalidad con las del agua, lo cual indica que el microbioma de esta esponja es muy similar al encontrado en el agua (como se observó en la Fig. 29). Las réplicas de *P. laughlini*, también forma su propio grupo muy cerca de las de *C. celata* y el agua. Las agrupaciones filogenéticas observadas son dadas más que nada por la especie de la esponja, ya que se observa otra grande agrupación compuesta por las réplicas biológicas de *A. revillagigedi*. Finalmente, observamos que el microbioma de sedimento también es filogenéticamente distinto a las demás muestras.

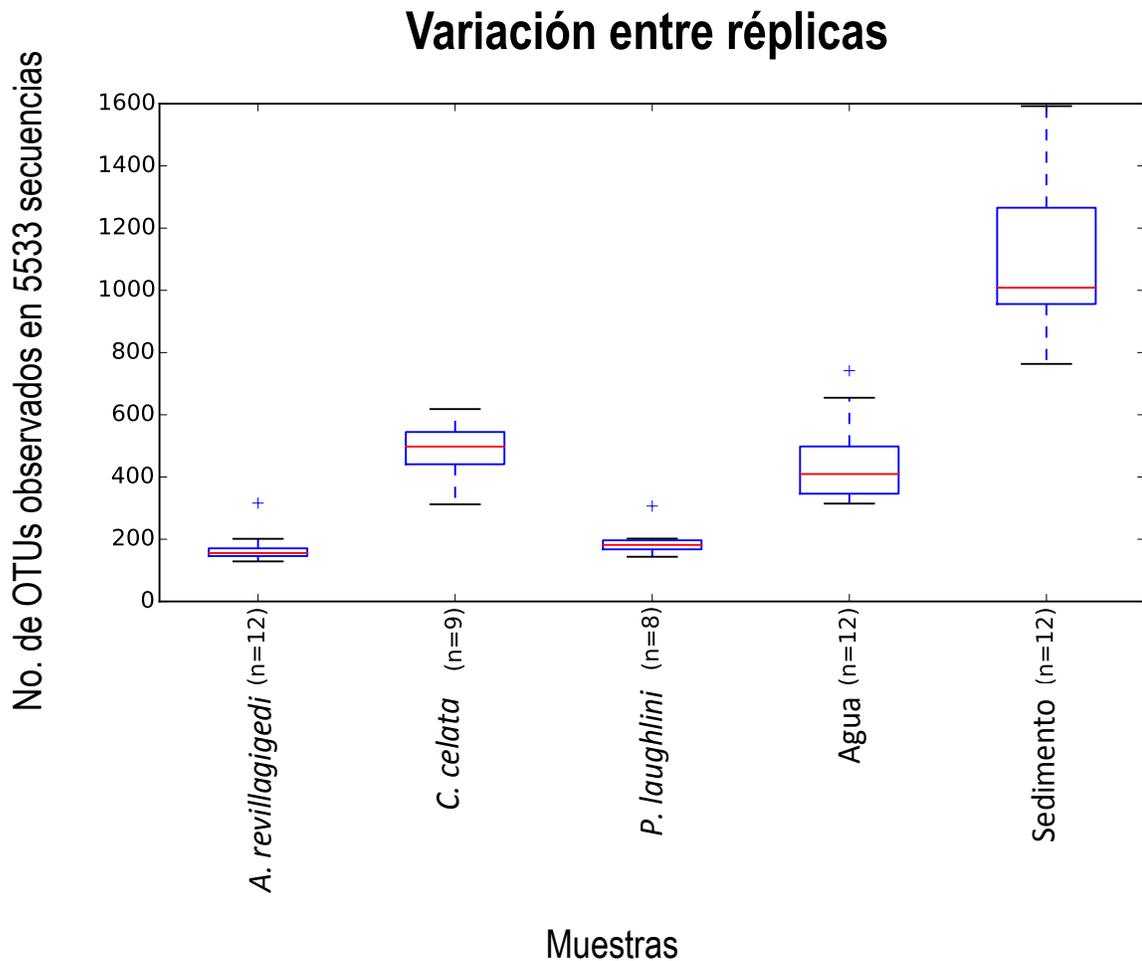


**Figura 34.** PCoA 2D de datos ponderados de beta diversidad. (A) se observan 4 agrupaciones dadas por el tipo de muestras, rojo: *A. revillagigedi*; azul: *C. celata*; naranja: *P. laughlini*; verde: agua; morado: sedimento.

### 3.6 Variabilidad entre réplicas

Los resultados de la prueba t-test no paramétrica calculada mediante una simulación Monte Carlo, con el fin de evaluar la variabilidad entre las réplicas biológicas de cada tipo de muestra, indican

que las diferencias entre réplicas biológicas de esponja no son significativas (Fig. 35). La mayor variabilidad la presentaron las muestras ambientales. *Cliona celata* fue la esponja que tuvo mayor variabilidad entre las 3, con una desviación estándar de 84.96 (en comparación a 47.49 y 47.75 dada por *Aplysina revillagigedi* y *Prosuberites laughlini*, respectivamente (Tabla 3).



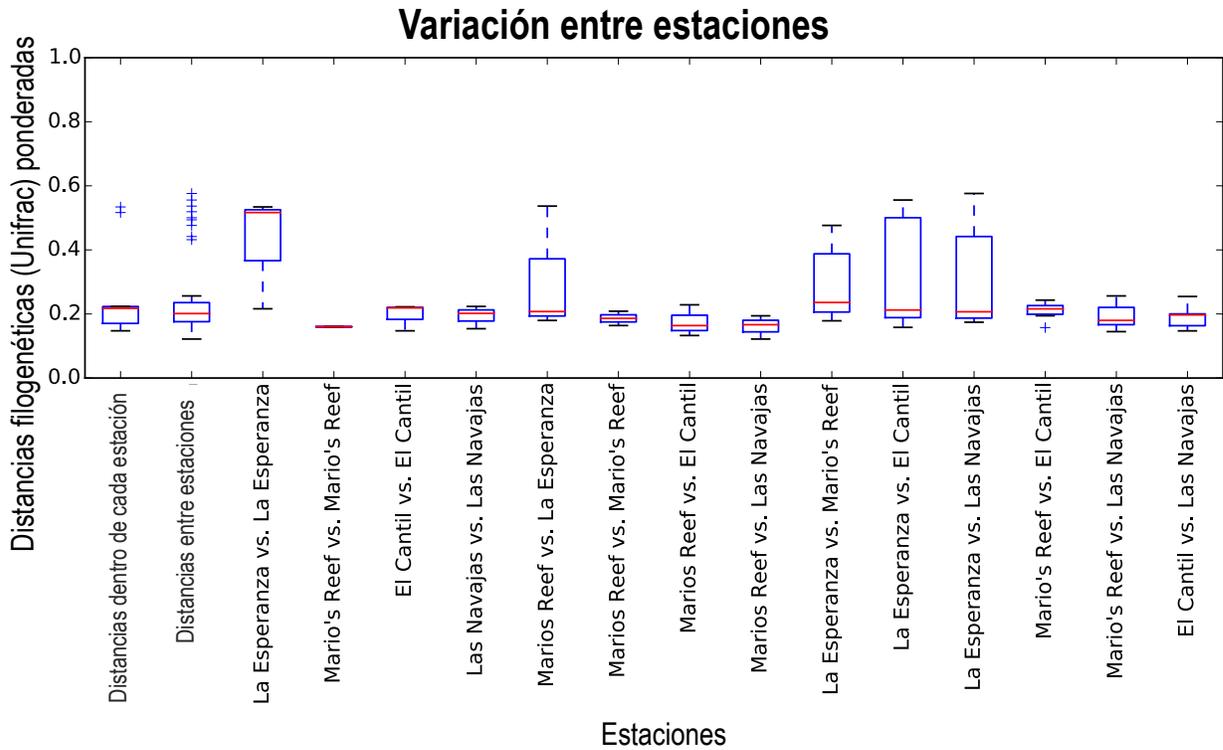
**Figura 35.** Resultados de prueba t-test no paramétrica, que muestra la dispersión del número de OTUs observados (normalizados a 5533 secuencias) representada por el conjunto de réplicas biológicas de *A. revillagigedi* (n=12), *C. celata* (n=9), *P. laughlini* (n=8), agua (n=12) y sedimento (n=12).

**Tabla 3.** Muestra el promedio de OTUs observados al muestrear 5533 secuencias, y su desviación estándar para cada muestra: *A. revillagigedi*, *C. celata* y *P. laughlini*, sedimento y agua. El número en paréntesis indica el número de réplicas de cada muestra.

Estadísticos	<i>A. revillagigedi</i> (12)	<i>C. celata</i> (9)	<i>P. laughlini</i> (9)	SD (12)	SW (12)
Promedio	171.13	488.1	191.31	1104.85	449.33
Desviación estándar	47.49	84.96	47.75	231.90	130.58

### 3.7 Variabilidad entre estaciones

Los resultados de la prueba t-test no paramétrica, indican que el sitio de colecta no es un factor de variación que esté influenciando la composición taxonómica del microbioma de *A. revillagigedi*, pues las réplicas no son significativamente diferentes, ni aquellas que proceden de la misma estación ni de las que proceden de otras estaciones, ( $P > 0.05$ ), (Fig.36).



**Figura 36.** Resultados de prueba t-test no paramétrica, sobre la mediana de las distancias filogenéticas en cada muestra de *Aplysina revillagigedi* proveniente de diferentes estaciones, en una matriz de distancia ponderada, calculada por Unifrac. No hubo valores de  $p < 0.05$ , lo que indica que no hay diferencias significativas entre las muestras provenientes de diferentes estaciones.

Los resultados obtenidos del análisis realizado por ANOSIM sobre la matriz de distancia ponderada generada por Unifrac, mostraron un valor de P igual a 0.635, lo que también indica que las estaciones no son significativamente diferentes ( $P > 0.05$ ).

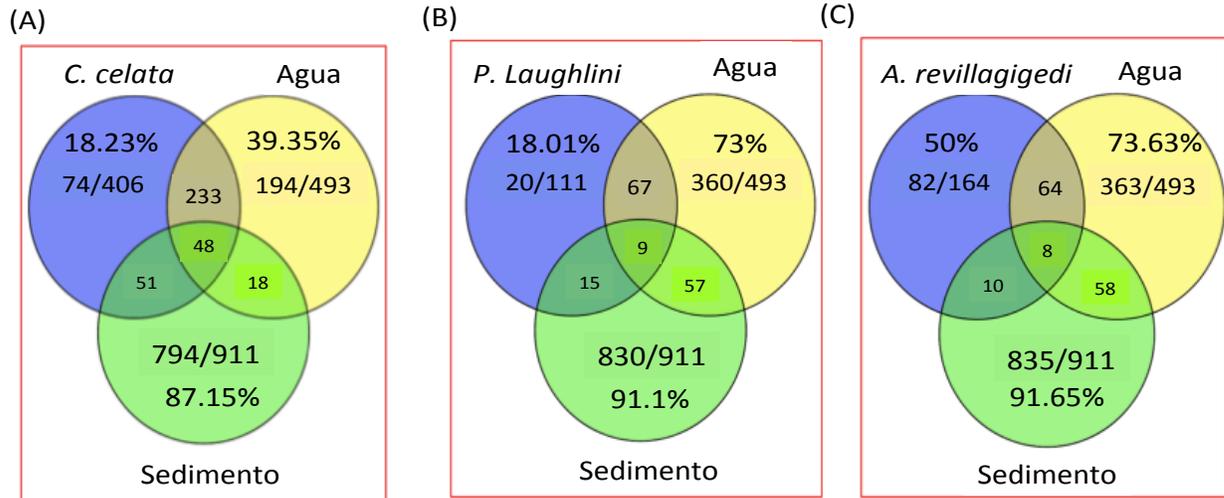
### 3.8 OTUs especie-específicos de esponjas (diagramas de Venn)

Los resultados de la comparación de los OTUs de las distintas muestras, usando aquellos con al menos 50 secuencias por OTU presentes en cada especie de esponja, contra el número de OTUs presentes en el medio circundante indican que el microbioma de *Cliona celata* y *Prosuberites laughlini* tienen un 18% de OTUs específicos (que no se encuentran ni en el agua ni en el sedimento), lo que significa que más del 80% de los OTUs que componen mayoritariamente a su microbioma, también son encontrados en el medio, particularmente en el agua, y siempre refiriéndonos a aquellos

microorganismos detectados con más de 50 secuencias por OTU (# de OTUs filtrados, Tabla 4). Por su parte, el microbioma de *Aplysina revillagigedi*, tiene un 50% de especies en común con el ambiente, y otro 50% de especies exclusivamente presentes en sus réplicas (Fig.37).

**Tabla 4.** Número de OTUs detectados con un 97% de similitud en la base de datos: *SILVA*, en cada tipo de muestra: sedimento, agua y esponjas. El número total de OTUs incluye OTUs con una sola secuencia asignada o *singletons*, y el número de OTUs filtrados es el conteo de los OTUs que fueron asignados al menos 50 secuencias o más.

Muestras	#OTUs totales	#OTUs filtrados
Sedimento	6203	437
Agua	3602	301
<i>A. revillagigedi</i>	962	137
<i>C. celata</i>	3081	249
<i>P. laughlini</i>	1209	62



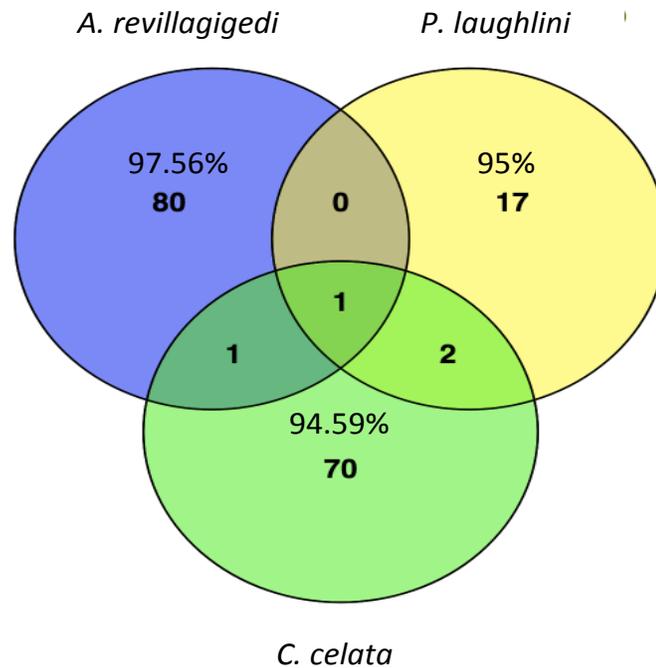
**Figura 37.** Comparación del número de OTUs de alta abundancia entre cada especie de esponja y su ambiente. La fracción en cada círculo indica el número de OTUs exclusivos a cada dominio entre el total, y sus respectivos porcentajes se indican por encima. *C. celata* (A) y *P. laughlini* (B) compartieron más del 80% de sus OTUs con el ambiente, mientras que *A. revillagigedi* (C) sólo compartió el 50%.

A continuación, se muestra una lista de los 10 OTUs exclusivos de cada especie que presentan una mayor abundancia (Tabla 5). El filo más abundante en las esponjas *Cliona celata* y *Prosuberites laughlini* fue el de las Proteobacterias, siendo una beta-proteobacteria el más dominante en *Cliona celata* y una gamma-proteobacteria en *Prosuberites laughlini*. Por su parte, *Aplysina revillagigedi* presentó mayor número de diferentes fila entre las tres esponjas, entre los cuales Chloroflexi, y Actinobacteria fueron los más dominantes.

**Tabla 5.** Lista de los 10 OTUs exclusivos y dominantes en orden descendente, de las esponjas: *Aplysina revillagigedi*, *Cliona celata* y *Prosuberites laughlini*.

<i>Aplysina revillagigedi</i>	<i>Cliona celata</i>	<i>Prosuberites laughlini</i>
TK10 (Chloroflexi)	Nitrosomonadales (Betaproteobacteria)	Salinisphaerales (Gammaproteobacteria)
Acidimicrobiia (Actinobacteria)	Oceanospirillales (Gammaproteobacteria)	Spirochaetales (Spirochaetae)
Marine Group I (Thaumarchaeota)	Opiutae (Verrucomicrobia)	Ambiguous taxa (SBR1093)
Deinococci (Deinococcus-Thermus)	Oceanospirillales (Gammaproteobacteria)	Acaryochloris (Cyanobacteria)
Caldilineae (Chloroflexi)	Flavobacteriia (Bacteroidetes)	Rhodospirillales (Alphaproteobacteria)
BD2-11 terrestrial group (Gemmatimonadetes)	Oceanospirillales (Gammaproteobacteria)	E01-9C-26 marine group (gammaproteobacteria)
Incertae Sedis (Bacteroidetes)	Planctomycetacia (Planctomycetes)	Flavobacteriia (Bacteroidetes)
Acidimicrobiia (Actinobacteria)	Alteromonadales (Gammaproteobacteria)	Rhodospirillales (Alphaproteobacteria)
Incertae Sedis (Bacteroidetes)	Flavobacteriia (Bacteroidetes)	Rhodospirillales (Alphaproteobacteria)
S085 (Chloroflexi)	Planctomycetacia (Planctomycetes)	HOC36 (gammaproteobacteria)

Además de la comparación de cada esponja con su ambiente, se compararon los OTUs que resultaron específicos de cada esponja (Fig.37), para hacer una comparación exclusivamente entre esponjas, con el fin de ver cuántos de estos OTUs eran comunes a las tres especies, y cuántos eran especie-específicos. Los resultados de esta comparación (Fig.38) mostraron que la mayoría de los OTUs son específicos de cada especie y solo tres están presentes en las 3 especies de esponjas. Estos fueron: una bacteria no cultivada (perteneciente al filo Bacteroidetes) y una Deltaproteobacteria (S110(24)-1), presente únicamente en las esponjas: *Aplysina revillagigedi* y *Cliona celata*; una Alfabroteobacteria no cultivada y *Spongiibacteria pacificum* (especie de Bacteroidetes), presentes exclusivamente en *Cliona celata* y *Prosuberites laughlini* (Fig.38).



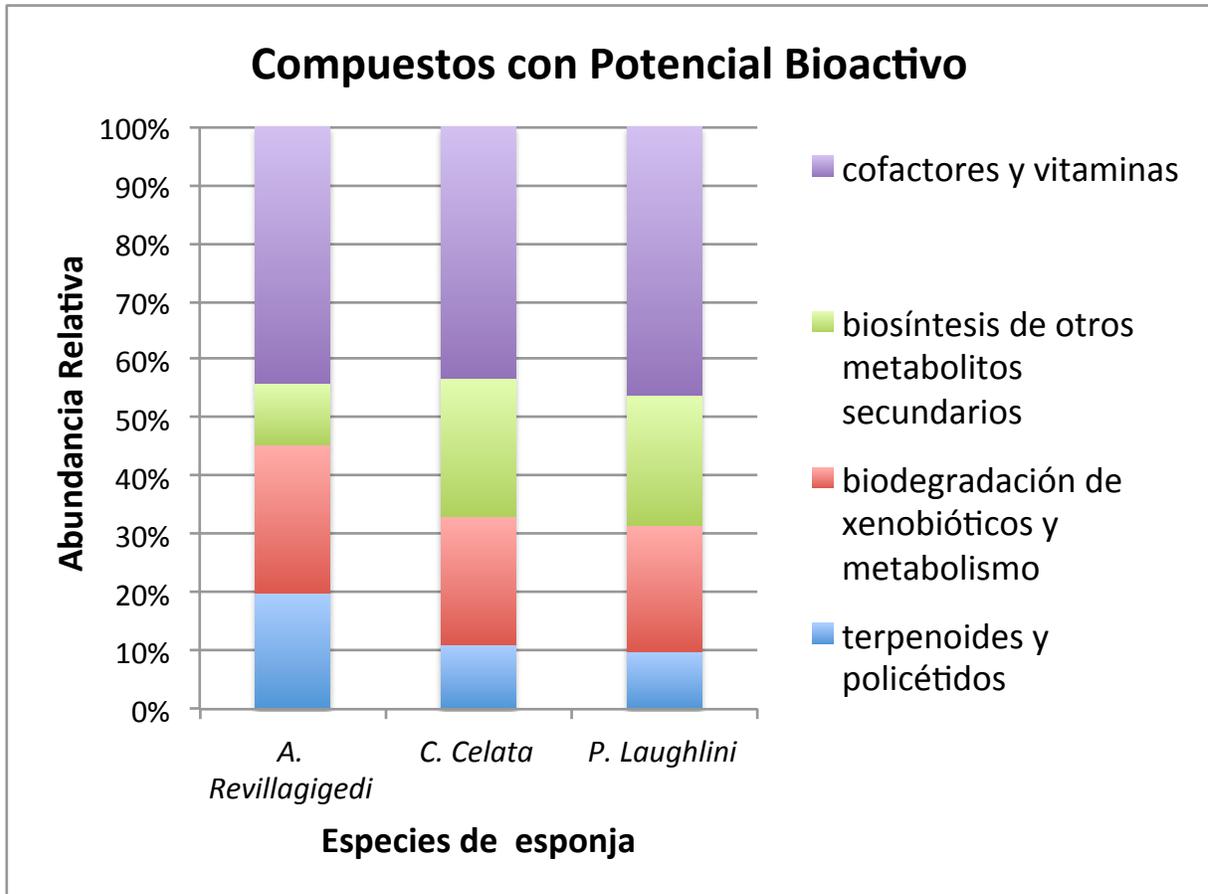
**Figura 38.** Diagrama de Venn que compara el número de OTUs especie-específicos de cada esponja. Se observa que sólo hay 1 ó 2 OTUs compartidos entre esponjas.

### 3.9 Análisis funcional con PICRUSt.

Este análisis se enfocó a analizar únicamente la cantidad de genes involucrados en la vía de producción de compuestos con potencial bioactivo predichos en cada esponja, dado que el objetivo de este trabajo fue la búsqueda de PKS y NRPS. Por esta razón se omitieron todas las categorías funcionales (KO) que formaban parte de procesos celulares, procesamiento de información genética, procesamiento de información ambiental, y metabolismo de macromoléculas, por lo que los porcentajes que se muestran a continuación no reflejan la misma proporción en los resultados totales del análisis funcional.

En la figura 39 que muestra el porcentaje de abundancia de categorías funcionales como: cofactores y vitaminas, biosíntesis de otros metabolitos secundarios, biodegradación de xenobióticos y metabolismo, y terpenoides y policetidos, se observa que el microbioma de las tres esponjas presentó una gran abundancia de genes enriquecidos para la producción de cofactores y vitaminas (40-45%) seguido por la biodegradación de xenobióticos y metabolismo (30%). Las esponjas *C. celata* y *P. laughlini* presentaron un mayor porcentaje de genes involucrados en la biosíntesis de otros metabolitos

secundarios (20%), que *A. revillagigedi*, quien sólo presentó un 10% de sus genes involucrados en esta categoría, sin embargo esta última esponja presentó la mayor cantidad de terpenoides y policétidos (20%), que son los compuestos de interés para este trabajo.



**Figura 39.** Abundancia relativa de genes con potencial para biosintetizar compuestos bioactivos en el microbioma de *A. revillagigedi*, *C. celata*, y *P. laughlini*. Los valores de abundancia relativa fueron calculados a partir del promedio de los genes presentes en las réplicas de cada una de las esponjas.

En la Tabla 6, se muestra el número de genes predecidos en el microbioma de cada esponja, involucrado en la producción de terpenoides y policétidos. Se observa que *Aplysina revillagigedi* es la esponja con mayor cantidad de genes involucrados específicamente en la síntesis de PKS tipo II (4413), a comparación de *Cliona celata* que sólo tuvo 8, y *Prosuberites laughlini* con 12. Por otra parte en la síntesis de NRPS, *Cliona celata*, superó a las otras dos esponjas con 17076 genes involucrados en su síntesis, en comparación a 12755 de *Aplysina revillagigedi*, y 6529 de *Prosuberites laughlini*. En general, estos resultados indican que *Prosuberites laughlini* fue la esponja con menos genes enriquecidos para la biosíntesis de este tipo de compuestos, mientras que las esponjas *Aplysina revillagigedi* y *Cliona celata*,

representan una fuente con mucho más potencial para su búsqueda, en particular *Aplysina revillagigedi* quien presentó el enriquecimiento de genes más alto para los compuestos de interés.

**Tabla 6.** Promedio de el enriquecimiento de genes para tipos de terpenoides y policétidos presentes en las réplicas de cada esponja.

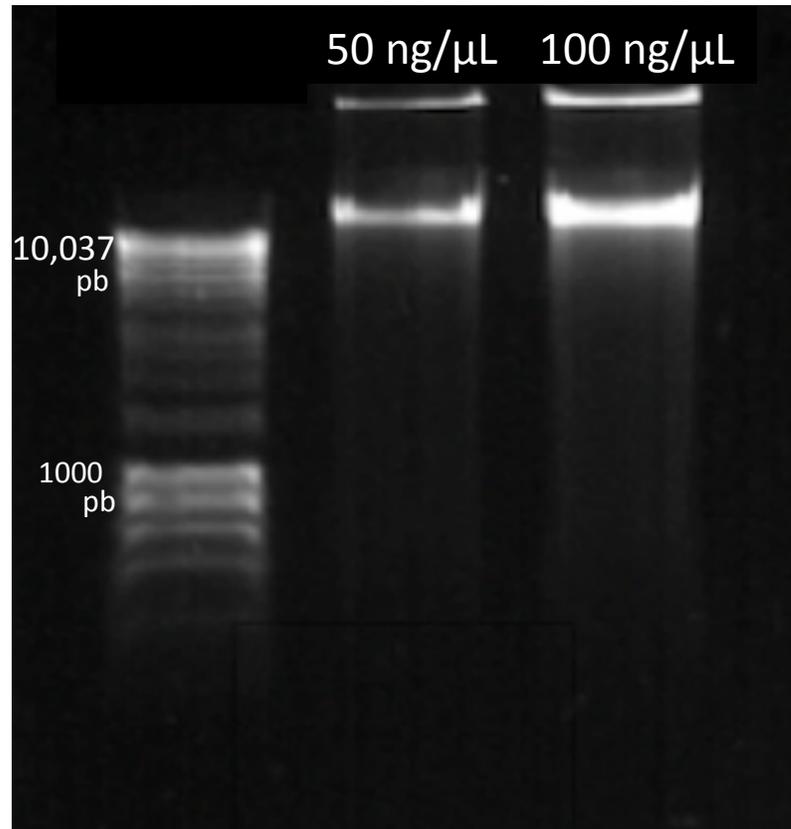
<b>Terpenoides y Policétidos</b>	<b><i>A. revillagigedi</i></b>	<b><i>P. laughlini</i></b>	<b><i>C. celata</i></b>
Terpenoid backbone biosynthesis	207652	123012	246015
Prenyltransferases	151058	88489	201752
Limonene and pinene degradation	104508	72511	108103
Geraniol degradation	107127	71120	81634
Polyketide sugar unit biosynthesis	78031	35131	70942
Tetracycline biosynthesis	57785	38581	87228
Carotenoid biosynthesis	33019	12388	80996
Biosynthesis of vancomycin group antibiotics	27081	11019	30108
Biosynthesis of ansamycins	23706	13262	25109
Zeatin biosynthesis	14294	6696	17507
Biosynthesis of siderophore group nonribosomal peptides	12755	6529	17076
Biosynthesis of type II polyketide products	4413	12	8
Biosynthesis of 12-, 14- and 16-membered macrolides	3	52	30

### **3.10 Separación celular y extracción de ADN del microbiomas de *Aplysina revillagigedi* para secuenciación *shotgun***

Con base en la caracterización taxonómica y la inferencia funcional obtenida del análisis de la comunidad microbiana, se seleccionó a *Aplysina revillagigedi* para realizar una secuenciación *shotgun* de su microbioma.

El resultado de la extracción del ADN que se pretendió estuviese enriquecido con ADN procariota mediante la separación de las células de esponja de las bacterianas (Fig.40) fue exitosa y se logró obtener un ADN íntegro y de alto peso molecular con el protocolo seguido; dos atributos decisivos para proseguir con el proceso de secuenciación por *shotgun*, pues entre más completo esté el ADN

mejor tamaño de fragmentos se pueden generar, los cuales facilitarán grandemente el proceso de ensamblaje que se realiza posterior a su secuenciación.

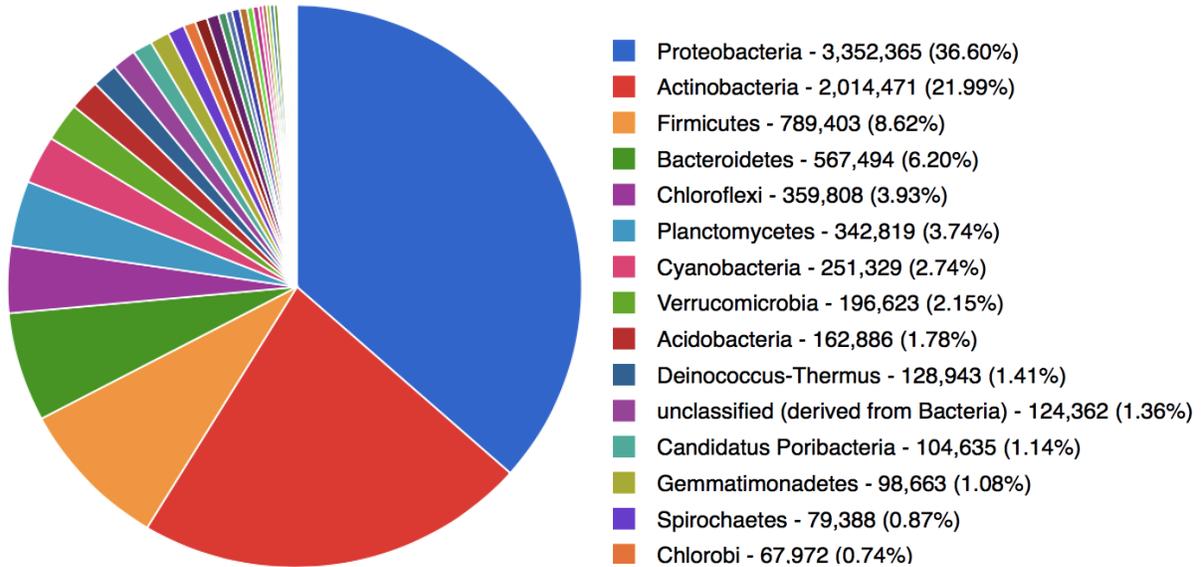


**Figura 40.** Calidad de la extracción de ADN procariota evaluada a 50 ng/μL y 100 ng/μL para fines comparativos, donde se observa que se obtuvo ADN de alto peso molecular según una escalera molecular Bioline de 1 kb, que genera 14 bandas, y para facilitar su identificación las más intensas son 10,037 pb y 1000 pb, respectivamente.

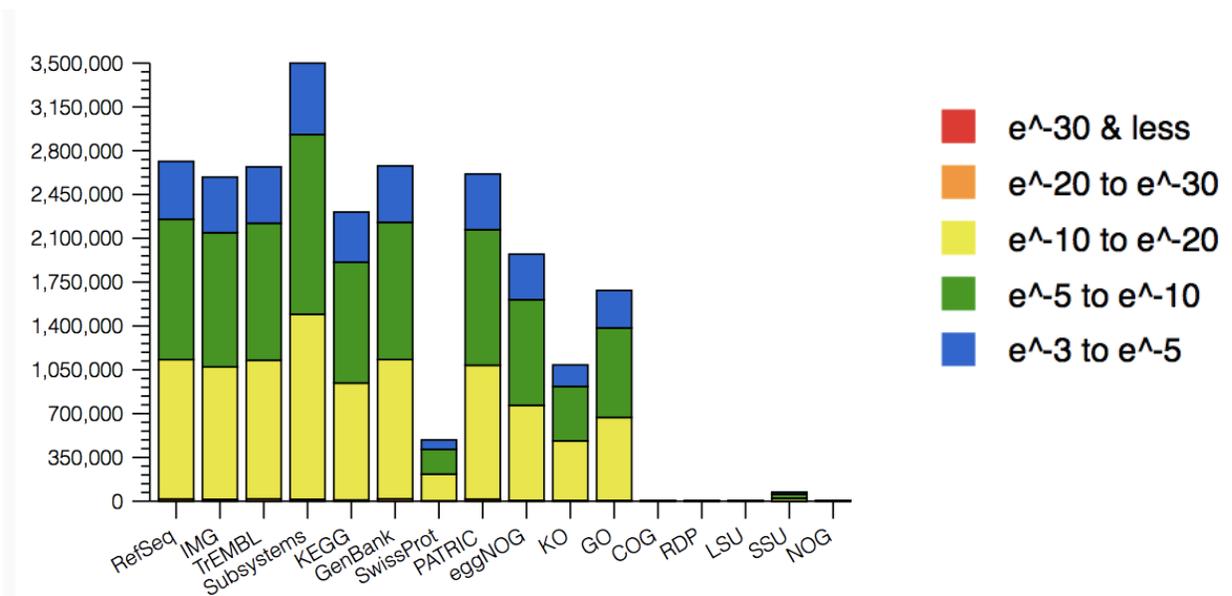
### 3.11 Análisis metagenómico del microbioma de *Aplysina revillagigedi* por mgRAST

Las secuencias crudas generadas por la secuenciación *shotgun* del microbioma de *Aplysina revillagigedi* se analizaron directamente en la plataforma MG-RAST. Los resultados de este análisis indican que se obtuvieron 13,634,651 lecturas (single end) con una longitud promedio de 137 pb, de las cuales 8,846,975 (96.59%) pertenecieron a bacterias, y sólo 169,316 (1.85%) a eucariotas, lo que indica que el enriquecimiento de ADN procariota realizado fue muy exitoso. Los resultados taxónomicos (Fig.

41) muestran 5 fila más abundantes para esta especie según la base de datos REFseq fueron: Proteobacteria, Actinobacteria, Firmicutes, Bacteroidetes y Chloroflexi.



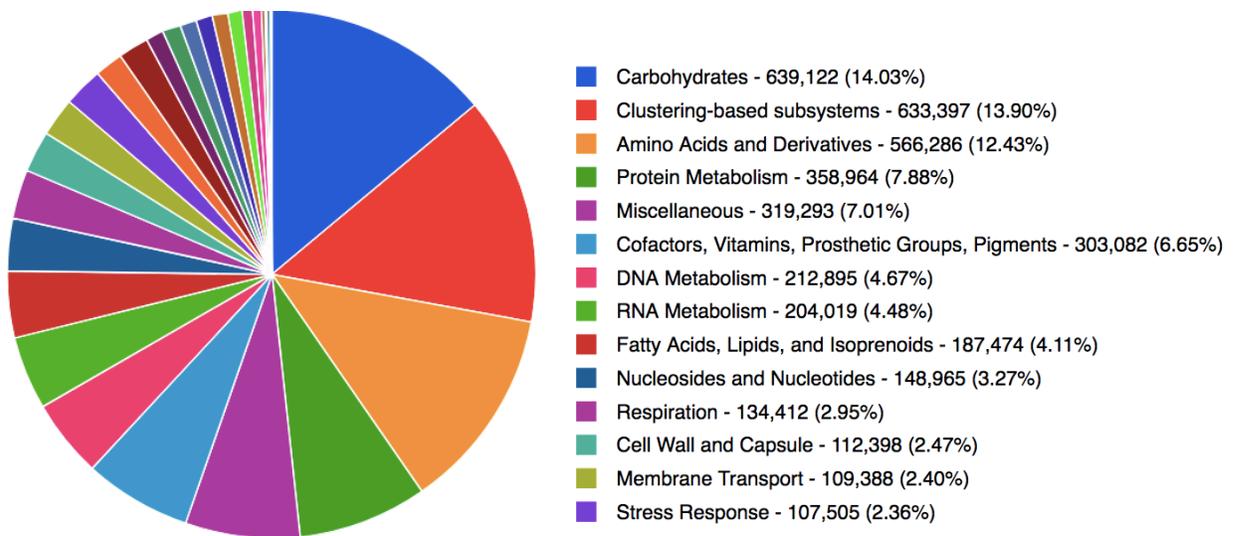
**Figura 41.** Resultados MG-RAST de la distribución de categorías taxonómicas a nivel filo utilizando la base de datos REFseq.



**Figura 42.** Cantidad de lecturas (barras) anotadas por distintas bases de datos que incluyen bases de datos de: proteínas, proteínas con información jerárquica funcional, y ARNr. Las barras están coloreadas en función al intervalo del valor e que presenten las secuencias anotadas en cada base de datos.

En cuanto a su análisis funcional se observó que la mayoría de las lecturas generadas se identificaron por medio de la base de datos *subsystems* (Fig.42), por lo que el resultado de la distribución de categorías funcionales a partir de esta base de datos se muestra a continuación.

Se observa que las dos funciones principales del microbioma de *Aplysina revillagiedi* (Fig.43) son la síntesis de carbohidratos (14.03%), y la síntesis de un grupo de función desconocida, denominado *clustering-based subsystems* (13.90%) que incluye a todos aquellos genes para los cuales existe evidencia funcional de un acoplamiento en una vía biosintética, que sugiere una agrupación física de ellos en *clusters* (Nmpdr.org, 2017). También es notorio el bajo porcentaje de genes involucrados en respuestas antiestrés (2.36%), transporte de membrana (2.40%) y la síntesis de paredes celulares y cápsulas de protección (2.47%).

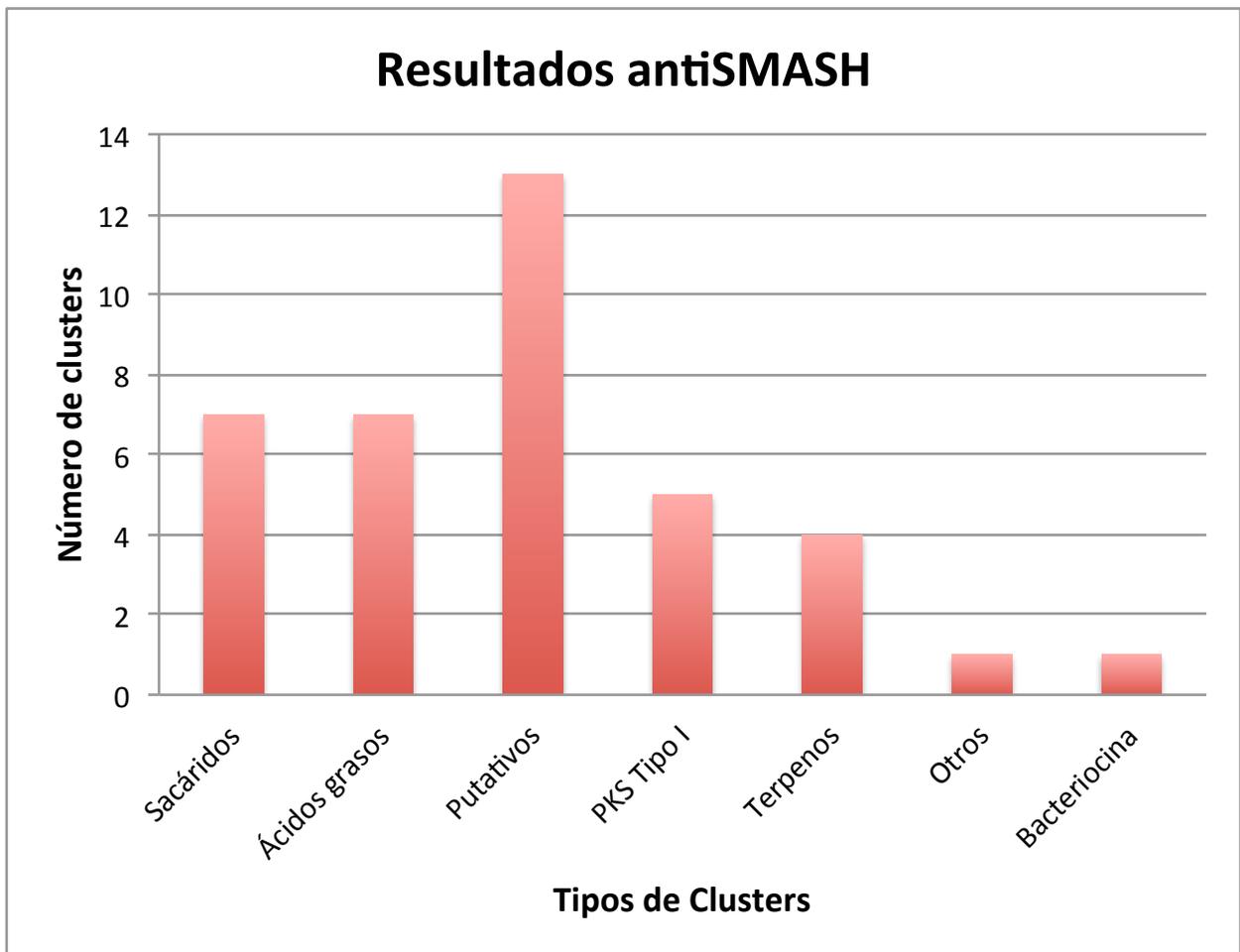


**Figura 43.** Perfil de categorías funcionales de Subsistemas. Cada parte de la gráfica representa el porcentaje de lecturas, anotadas a la categoría en cuestión. El microbioma presenta una diversidad de genes enriquecidos para la síntesis de macromoléculas principales.

### 3.12 Resultados antiSMASH y NaPDoS

Se identificaron un total de 38 clusters (Fig.44) de genes biosintéticos involucrados en la síntesis

de metabolitos secundarios a partir de contigs ensamblados con MEGAHIT originarios de la secuenciación *shotgun* del microbioma de *Aplysina revillagigedi* por medio de la base de datos de antiSMASH. Cinco de estos pertenecieron a PKS tipo 1, cuatro a terpenos, y uno a bacteriocina. El resto de los BGC, estuvieron involucrados en la síntesis de sacáridos (7), ácidos grasos (7), y más del 30% en clusters identificados como, “BGCs putativos”. Esta última categoría incluye metabolitos secundarios que forman parte de clusters biosintéticos desconocidos (que no están presentes en las bases de datos que utilizaba antiSMASH), sin embargo, se denominan putativos porque incluyen genes que codifican para enzimas que típicamente se encuentran en este tipo de clusters. No se encontró ningún cluster involucrado en la síntesis de NRPS.



**Figura 44.** Número de BGCs encontrados para cada categoría de metabolitos secundarios reportadas por antiSMASH.

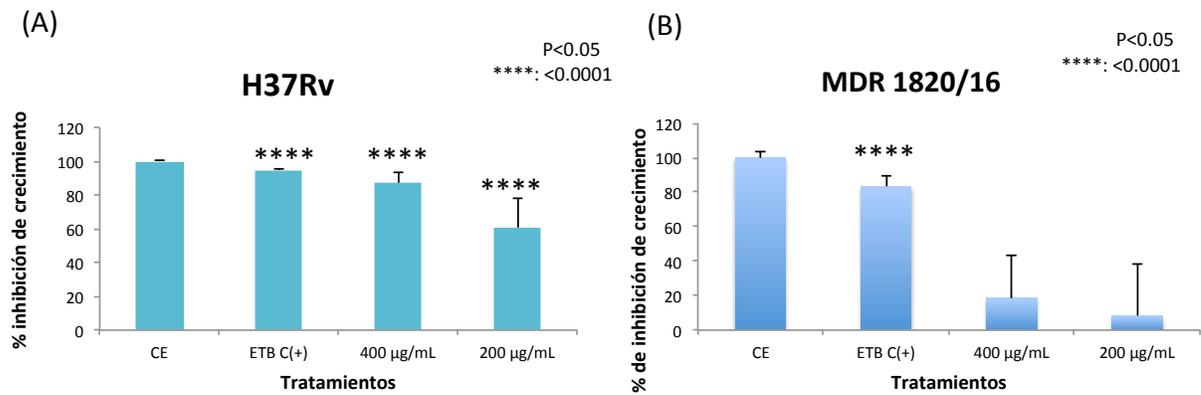
Aún más, de los 5 BGCs involucrados en la síntesis de PKS tipo 1 que fueron encontrados por antiSMASH, se lograron identificar 3 dominios KS por medio de la base de datos NaPDoS. Los resultados muestran que estos dominios podrían estar involucrados en la vía biosintética de los siguientes dos compuestos: Epotilona y Fumonisina. En la tabla 7, se muestra el porcentaje de identidad, la longitud de alineamiento y el valor E de su alineación con KS presentes en la base de datos NaPDoS. Se observa que la mayoría tuvo porcentajes mayores al 25% de identidad a lo largo de más de 100 amino ácidos, y valores-E menores a 0.001, que son los parámetros típicamente reportados para considerar homología entre secuencias de amino ácidos (Claverie y Notredame, 2007).

**Tabla 7.** Resultados de la búsqueda en la base de datos NaPDoS de los dominios KS que fueron anotados para los clusters BGCs involucrados en la síntesis de PKS tipo 1 por antiSMASH.

Tipo de dominio	Producto vía	Valor-E	Longitud de alineamiento	%ID
KS-modular	Epotilona	3e -109	406	53%
KS-modular	Epotilona	3e -71	248	55%
KS-iterativo	Fumonisina	5e -62	439	37%

### 3.12 Bioensayos

Los resultados de probar dos concentraciones del extracto crudo de la esponja *Aplysina revillagigedi* (400 µg/mL y 200 µg/mL, respectivamente), contra dos cepas de *M. tuberculosis*: H37Rv y 1820/16, se muestran en la figura 45 como el promedio del porcentaje de inhibición entre los triplicados, y su desviación estándar. El análisis estadístico (ANOVA de una vía) muestra que la esponja *Aplysina revillagigedi* inhibió significativamente el crecimiento de la cepa prototipo H37Rv de *M. tuberculosis* a ambas concentraciones probadas, por el contrario no se observó un efecto significativo del extracto crudo de esta esponja sobre la cepa multidrogorresistente: 1820/16 (Fig. 45).



**Figura 45.** Efecto inhibitorio del extracto crudo de *A. revillagigedi*, sobre dos cepas de TB: (A) H37Rv y (B) 1820/16 MDR. De izquierda a derecha se observa el porcentaje de inhibición de: el control de esterilidad (CE) (100%, pues no contiene inóculo); Etambutol (antibiótico que inhibe el crecimiento en ambas cepas, concentración utilizada: 4 µg/mL); y los tratamientos a: 400 µg/mL y 200 µg/mL. Los asteriscos representan actividad inhibitoria significativa según un valor  $P < 0.05$ , calculado por medio de un ANOVA de una vía seguido por una prueba de comparaciones múltiples de Tukey.

## Capítulo 4. Discusión

---

### 4.1 Identificación Molecular de Esponjas Marinas provenientes del Parque Nacional Cabo Pulmo, (BCS).

Se logró amplificar satisfactoriamente el gen COI de todas las muestras de esponja para su subsecuente identificación molecular, a pesar de que la calidad y la concentración de las extracciones de ADN tendía a variar dependiendo de la especie de esponja con la que se trabajaba. Según la literatura, uno de los posibles riesgos a la hora de extraer ADN de esponjas para su identificación molecular, es coextraer metabolitos secundarios que podrían interferir en las reacciones enzimáticas que se llevan a cabo al momento de amplificar el ADN (Shearer *et al.*, 2002; Vargas *et al.*, 2012). Otro factor que podría causar problemas es la presencia de compuestos húmicos, que son sustancias inhibitorias que están presentes en el sedimento que se encuentra en proximidad a la esponja. El uso de lavados previos a la extracción de ADN, podría ser una buena opción para evitar este tipo de riesgos. También, es importante trabajar lo más rápido posible en la extracción de ADN para evitar su degradación.

Sin embargo, el mayor problema que se presentó no fue por las extracciones si no la identificación molecular de las secuencias obtenidas, ya que daban resultados iguales que identificaban tres especies distintas. Para evitar este tipo de situaciones sería recomendable reforzar los resultados aquí reportados, con el análisis de las muestras de esponja por medio de un análisis morfológico (como el método tradicional del conteo de espículas), o a través de la utilización de otros marcadores moleculares, como ADNr de ribosomas nucleares, tales como: ITS1-5.8-ITS2. Estos marcadores junto con el COI, permiten hacer comparaciones entre especies más precisas, ya que se ha reportado que para algunas especies de esponja el gen COI no representa suficiente variabilidad como para hacer una identificación a nivel especie (Cruz-Barraza *et al.*, 2012). El uso exclusivo del marcador COI también pudo haber influenciado en que no se pudieran identificar algunas esponjas de este trabajo a nivel especie. Por lo anterior, fue crucial corroborar los resultados obtenidos del BLAST, por medio de búsquedas geográficas de las especies de esponja identificadas. La mayoría de ellas sí tenían registros de distribución a lo largo del Pacífico Este, como era de esperarse por la zona de estudio. Sin embargo, hubo excepciones a esto, como *Prosuberites laughlini* y *Clathria prolifera*, las cuales están reportadas con distribución en el Caribe sur y aguas del Atlántico sudamericano, respectivamente. Aunque pudo haber imprecisión o baja resolución en los resultados obtenidos de sus secuencias COI, dando como resultado una identificación errónea, también es posible que estas esponjas representen nuevos registros en esta zona, pues su presencia podría deberse a su introducción de manera accidental (como especies

invasoras), a través de su incrustación en barcos que cruzan el canal de Panamá, ya que ambas son de morfología incrustante (Bastida-Zavala *et al.*, 2014). Para el caso particular de *Clathria prolifera*, existen registros de su presencia en el Pacífico Este, particularmente en las costas de San Francisco, CA (1940); Washington, USA (1967); y Humboldt Bay, CA (1989), donde existe evidencia de que arribó por medio de su incrustación en cascos de barcos que transportaban ostiones a esas bahías (Cohen, 2016).

La distribución de las diferentes especies de esponjas a lo largo del PNCP llama la atención porque no todas fueron encontradas en todos los sitios muestreados, a excepción *Aplysina revillagigedi*. Se sabe que algunos animales que depredan a esponjas marinas como el pez angel, escáridos, nudibranquios y tortugas marinas (Loh & Pawlik, 2014) anidan y se alimentan en el PNCP periódicamente. Por lo que la amplia distribución de la esponja *Aplysina revillagigedi*, a lo largo del parque incita a pensar que esta propagación exitosa de la esponja se deba a que posee una armadura química más fuerte, que permite su supervivencia y consecuente dispersión en el parque. Más aún, se ha reportado que ejemplares del género *Aplysina sp.* producen una gran cantidad de compuestos bioactivos con propiedades antipredatorias, antifouling y antimicrobianas, como metabolitos bromados (Betancourt-Lozano, *et al.*, 1998; Azevedo *et al.*, 2008; Kazanjian y Fariñas 2006). Por otra parte, no se observó que los cambios de temperatura de los sitios de colecta influenciarán la distribución de las esponjas, porque las estaciones en las que se encontraron las otras dos especies de esponjas (*P. laughlini* y *C. celata*), se caracterizaban por presentar el mínimo y el máximo del rango de temperatura registrada (23°C - 25°C). Sin embargo *C. celata* sólo se encontró en las estaciones más someras (NAV: 5.2m, y ESP: 16.1m), por lo que para esta esponja quizás la profundidad si sea un factor que afecte a su distribución vertical. Esta última esponja es de tipo excavadora y su presencia en el PNCP puede llegar a ser perjudicial para el arrecife. En el trabajo de Carballo *et al.*, 2012, se menciona que las esponjas excavadoras suelen ser abundantes en arrecifes afectados por blanqueamiento coralino, y que juegan un rol significativo en la destrucción física de su estructura al debilitar la unión entre las colonias vivas y el marco de carbonato de calcio que componen al arrecife coralino.

## **4.2 Caracterización y comparación del microbioma de tres especies de esponjas abundantes en el PNCP.**

Se compararon los microbiomas de tres especies de esponjas abundantes en distintos puntos del PNCP, estas esponjas se identificaron como: *Cliona celata*, *Aplysina revillagigedi* y *Prosuberites laughlini*. Los esfuerzos de secuenciación realizados en este trabajo lograron identificar alrededor del 71.74% de

OTUs estimados por el índice Chao1 (Fig.32); el 30% restante de OTUs estimados que no se lograron observar, puede explicarse porque, o están presentes en cantidades muy bajas en el microbioma de cada esponja lo que hace muy difícil su detección, o porque debido a los tamaños de muestra tan grandes (decenas de miles de secuencias) el estadístico sobrestima esta abundancia, lo que parece la hipótesis más probable basándonos en las curvas de rarefacción. De todas formas siempre hay que tener en cuenta, que por la manera en que se lleva a cabo la secuenciación, es frecuente que algunos OTUs de baja abundancia no se logren detectar debido a la sobreamplificación de OTUs con altas abundancias. Este es un problema inherente a la secuenciación masiva de amplicones que se debe continuar mejorando (Caporaso *et al.*, 2011). En este contexto, cabe mencionar que todos los OTUs considerados para hacer la búsqueda de OTUs exclusivos de cada esponja y que se discutirán a continuación, tenían 50 o más lecturas por OTU para estar seguros de que los grupos taxonómicos muy probablemente conformaban una parte real del microbioma de cada especie de esponja analizada y no estaban presentes solo por azar. Lo anterior en consideración a la inevitable cuestión de los sesgos presentados por la amplificación por PCR, la region variable del gen del 16S del ARNr seleccionada para secuenciar, así como el número diferente de copias de este gen que pueden presentar en algunas especies bacterianas que puede afectar a la abundancia relativa de algunos grupos (Wintzingerode *et al.*, 1997).

Los resultados de los análisis ecológicos, indican que *Aplysina revillagigedi* es la esponja con el microbioma más diverso de entre las tres especies comparadas, pues presenta varios grupos que comparten abundancias parecidas. En contraste, con los microbiomas de *Cliona celata* y *Prosuberites laughlini* presentaron 2 ó 3 fila altamente abundantes. Además de lo anterior, al menos el 50% de los OTUs encontrados en *Aplysina revillagigedi*, resultaron ser especie-específicos de esta esponja, lo que sugiere una transmisión vertical de estos grupos por medio de la reproducción de la especie de esponja en cuestión. Hay varios estudios que han demostrado este mecanismo de transferencia entre esponjas de una misma especie (Schmitt *et al.*, 2007; Sharp *et al.*, 2007; Schmitt *et al.*, 2008; Lee *et al.*, 2009). Entre los OTUs exclusivos y más abundantes de *Aplysina revillagigedi* estuvieron especies representativas del filo Actinobacteria, el cual es responsable de más de la mitad de todos los antibióticos descubiertos a la fecha (Udwary, *et al.*, 2007). Lo anterior aunado a los resultados del análisis de inferencia funcional realizado por PICRUSt, que indicaron que esta especie poseía la mayor cantidad de genes involucrados en la síntesis de compuestos bioactivos (Fig. 39 y Tabla 6, respectivamente) comparada con las otras dos, haciendo de ella la elegida para continuar con el estudio de su metagenoma completo y así buscar novedades farmacológicas como las PKs y los NRPs que se había propuesto este trabajo. Aunque no se encontraron estudios metagenómicos realizados específicamente con esta especie en la literatura, sí hay estudios con esponjas del mismo género (Slaby *et al.*, 2017;

Thomas *et al.*, 2016; Hentschel *et al.*, 2002; Schmitt *et al.*, 2012) que dan soporte a los fila identificados en este trabajo para *A. revillagigedi* y sus respectivos porcentajes de abundancia. Entre estos estudios resalta el realizado por Thomas *et al.*, (2016), que incluyó el análisis del microbioma de 3 especies del género *Aplysina sp* (estas fueron: *Aplysina cavernícola*, *Aplysina aerophoba* y *Aplysina fulva*), donde se reportó que la especie de bacteria más abundante pertenece a una clase no cultivada de gammaproteobacterias (tal como se observó en los resultados de este trabajo), y que entre los fila bacterianos más dominantes (con abundancias mayores al 10%) se encuentran los 5 fila más abundantes reportados por los resultados de este trabajo (Fig.31).

Por otra parte, aunque los resultados del análisis de alfa diversidad mostraron que la esponja *Cliona celata* presentaba una alta biodiversidad microbiana, lo cual coincide con lo reportado previamente (Thomas *et al.*, 2016; Jeong *et al.*, 2015), se descartó para la búsqueda de PKS y NRPs en este trabajo, porque el análisis de beta diversidad reveló que el microbioma es muy similar al encontrado en el agua. Los resultados de beta diversidad también parecen acercarse a *Prosuberites laughlini* al del agua, aunque esta similitud es menor que la de *Cliona celata*. Estas similitudes entre el microbioma de estas especies y el agua pueden deberse a la textura blanda que las caracteriza y que dificultó la obtención de una muestra adecuada del mesohilo de ambas esponjas, lo que incita a pensar que el resultado observado por estos análisis puede estar grandemente influenciado por la presencia de bacterias presentes en la capa más externa de la esponja (la cual está en contacto con el agua), las cuales podrían haber sido accidentalmente incluidas al momento de procesar a las esponjas en el campo. La alta presencia de cianobacterias encontradas en el microbioma de *C. celata* soporta esta idea, puesto que las cianobacterias son un grupo de bacterias fotoautótrofas que típicamente se encuentra en la capa externa de las esponjas por su requerimiento de luz (Rutzler, 1985; Wilkinson, 1992). Su presencia particularmente abundante en el microbioma de *C. celata*, podría deberse a la distribución de esta esponja en el Parque. *C. celata* se encontró únicamente en zonas más someras.

Por su lado, el alto porcentaje de OTUs que compartieron los microbiomas de cada esponja con el ambiente (Fig. 37), sugiere que una buena parte de los OTUs que componen los microbiomas de estas esponjas son adquiridos por transferencia horizontal. La transmisión horizontal implica una fuerte dependencia de los siguientes factores: la abundancia de ciertos grupos simbiotes en el ambiente (con lo cual se pueden estar constantemente renovando ó que representen OTUs de los cuales se alimente la esponja), la eficiencia del sistema de filtración que posea la esponja, otros simbiotes ya establecidos en el microbioma de cada esponja ya sea por transmisión vertical u horizontal que estén controlando la colonización de nuevos grupos, por medio de la generación de compuestos bioactivos con propiedades

citotóxicas, que inhiban la colonización de otras especies competitivas de microbios presentes en el agua, y la presencia de características genómicas que faciliten la adaptación de ciertas especies de bacterias a un existencia simbiote como lo son: los elementos transposables, mecanismos de defensa, y proteínas similares a las producidas por eucariontes (Thomas *et al.*, 2010; Fan *et al.*, 2012; Hentshel *et al.*, 2012; Horn *et al.*, 2016). Aún más la presencia de pocos grupos “esponja-específicos” observados (Fig.38), sugiere que solo unos cuantos OTUs son verdaderamente cosmopolitas (que se encuentran exclusivamente en esponjas de diferentes especies), y que la gran mayoría de los OTUs exclusivos a cada esponja tienen un estatus especialista (que se encuentra en una o pocas especies de esponja), mientras que los resultados observados para la variación individual entre réplicas biológicas de cada especie de esponja (desviación estándar de la Tabla 4) sugiere la presencia de otros cuantos OTUs oportunistas (Torsten *et al.*, 2016).

### 4.3 Análisis del metagenoma de *Aplysina revillagigedi*: búsqueda de PKS y NRPS

Utilizando los protocolos reportados por Takahiro *et al.*, 2012 y Torsten *et al.*, 2010, fue posible enriquecer el ADN bacteriano de una muestra de tejido de esponja, y extraer ADN con alto peso molecular para continuar con la secuenciación de fragmentos o shotgun (Fig. 40). Los resultados de esta secuenciación analizados por medio de MG-RAST, (Fig.30), muestran una composición taxonómica diferente del microbioma de esta esponja, en comparación a la observada como resultado de la secuenciación de la amplificación del gen 16S del ARNr (Fig.31). Una posible causa para explicar dichas diferencias, podría ser la manera en que se extrajo el ADN. En el caso de la secuenciación shotgun el ADN se extrajo utilizando un método clásico como se describe en la metodología, mientras que para la secuenciación por amplicones se utilizó un kit de extracción de ADN para sedimento que fue modificado. Hay estudios (Glassing *et al.*, 2016; Salter, *et al.*, 2014) que reportan la presencia de contaminantes presentes en este tipo de kits de extracción de ADN y otros reactivos de laboratorio, que ocasionan el enriquecimiento de ciertas especies de bacterias en microbiomas de baja abundancia que incluyen a representantes del filo Acidobacteria, el cual curiosamente se presentó como el filo de mayor abundancia en esta esponja en el análisis de secuenciación por amplicones (Fig.31).

Por otra parte, en el perfil de categorías funcionales (Fig.43) se observa un grupo denominado *clustering-based subsystems* que incluye a todos aquellos genes que forman parte de una vía biosintética, para los cuales existe evidencia funcional de un acoplamiento de genes que indica su agrupación física en *clusters* (Nmpdr.org, 2017). Estos *clustering-based subsystems* representan el 13.9%

de las secuencias identificadas en el metagenoma de *Aplysina revillagigedi*, y se posiciona como el segundo porcentaje más abundante después de los genes involucrados en la producción de Carbohidratos (14.03%). Cabe mencionar que la síntesis de metabolitos secundarios, se hace a través de clusters de genes biosintéticos (BGC), por lo que esta categoría bien podría considerarse representativa de ellos, pues al ensamblar las secuencias generadas y someterlas a la base de datos de antiSMASH, para una búsqueda más detallada de este tipo de BGCs, se encontró que, en efecto, el metagenoma de esta especie de esponja está enriquecido con clusters de genes biosintéticos, pues se encontraron 38 BGCs, 5 de los cuales se encontraron asociados a PKS, con dominios KS modulares e iterativos (Weber *et al.*, 2015).

Entre los BGCs de PKS, se detectaron dos dominios KS, los resultados de la búsqueda de estos dominios por la base de datos NaPDoS (tabla 6), indicaron que estos dominios tienen alrededor de un 53%-55% de similitud con un dominio KS involucrado en la vía de síntesis del Epotilona, que forma parte de una clase de fármacos con potencial anticancerígeno que interfiere con la tubulina (Napdos.ucsd.edu, 2017). Aún más, el análisis por Blastp, de estos dominios mostró una identidad del 93%, con un valor de E igual a cero y un 100% de cobertura, con un PKS putativo de una bacteria no cultivada llamada 2063G, que está asociado con la esponja marina *Pseudoceratina clavata* (Kim y Fuerst, 2006). Sin embargo, el bajo porcentaje de similitud reportado por la base de datos NaPDoS, indica que quizás se trate de un par de PKS novedosos, al igual que se estipula para los otros 3 BGCs de PKS.

Por otra parte, se encontraron 7 BGC's involucrados en la síntesis de sacáridos. Se ha reportado que sacáridos asociados a componentes membranales tales como los lipopolisacáridos, polisacáridos capsulares, y polisacárido A, desempeñan un rol importante en las interacciones existentes entre huésped-microbio o microbio-microbio (Park *et al.*, 2009; Kadioglu *et al.*, 2008; Mazmanian *et al.*, 2005; Mazmanian *et al.*, 2008). También existen estudios que le atribuyen una serie de actividades antibacterianas a sacáridos fácilmente difusibles por la membrana (Flatt y Mahmud, 2007; Weitnauer *et al.*, 2001). La mayoría de los genes biosintéticos centrales de estos BGCs mostraron hits con la enzima glicosil transferasa tipo 1 (según los resultados de la búsqueda por smCOG y pHMM), que esta involucrada en la transferencia de monosacáridos a lípidos, proteínas, ADN, o la parte glucosídica de glucoproteínas y glucolípidos (Qasba y Ramakrishnan, 2007). Sin embargo, el análisis por Blastp de éstos resultados mostró porcentajes de similitud menores al 70% lo que sugiere que la mayoría de éstos clusters podrían tratarse de sacaridos novedosos.

En contraste, dado el alto porcentaje de BGCs del tipo putativo, se realizó una búsqueda de uno

de estos clusters por la base de datos Blastp, con el fin de dar una idea del tipo de proteínas asociadas a este tipo de clusters. Sin embargo, como era de esperarse se observaron únicamente *hits* con bajos porcentajes de identidad (35-44%), que sugerían proteínas hipotéticas de las bacterias *Candidatus Rokubacteria* (bacteria presente en sedimento acuífero, Hug *et al.*, 2016) y una alfa/beta hidrolasa de *Bacillus xerothermodurans* (una bacteria del sedimento) (Bond y Favero, 1977).

#### 4.4 Bioensayos

Según los resultados del análisis estadístico (ANOVA) y cálculo del porcentaje de inhibición del cultivo de *M. tuberculosis*, el extracto crudo de la esponja *Aplysina revillagigedi* inhibió significativamente el crecimiento de una de las dos cepas evaluadas de *M. tuberculosis*, esta fue la cepa prototipo H37Rv, indicando la presencia de compuestos con potencial antibiótico en esta esponja marina.

Hay algunos puntos de la metodología que son necesarios mejorar; particularmente, en el proceso de elaboración del extracto crudo. Al trabajar con esta esponja, se descubrió que posee un compuesto (aún desconocido) que oscurece al extracto con el paso de los días y forma un precipitado. Este precipitado interfiere grandemente con la medición de la absorbancia realizada al momento de revelar los resultados del ensayo colorimétrico. Algunas soluciones para este emergente problema, incluyen las siguientes: realizar el extracto crudo probando solventes distintos al utilizado, esto con el fin de descartar la posibilidad de una interacción entre el solvente y los compuestos extraídos del tejido de la esponja que resultó en la formación de precipitados. Otras opciones de solventes incluyen: el etanol, el EtOAc, el acetato de etilo, el PBS, el cloroformo y el acetonitrilo (Palermo, *et al.*, 1998; Kazanjian y Fariñas 2006; Oliveira *et al.*, 2006; Tasdemir, *et al.*, 2007; Millan-Aguiñaga, 2009). También, existe una posibilidad de que el extracto de esta esponja en particular sea fotosensible y que por tanto necesite manipularse de acuerdo a ello (evitando que entre en contacto con la luz), pues se sabe que otras especies de este género poseen compuestos sulfatados que hacen que la esponja cambie de color al entrar en contacto con el oxígeno del aire, como es el caso de *Aplysina aerophoba* (Sacristan *et al.*, 2011). También se podría realizar una separación celular previa a la extracción del tejido, de tal manera que si el compuesto responsable del cambio de color del extracto, está siendo producido por las células de la esponja, estas se eliminen y solo se extraigan los compuestos del microbioma o las bacterias asociadas a la esponjas, las cuales según la literatura son responsables de la gran mayoría de los

compuestos bioactivos presentes en las esponjas marinas (Sacristan *et al.*, 2011; Thomas *et al.*, 2010; Hentschel *et al.*, 2012; Webster y Thomas, 2016). Otras perspectivas para este trabajo incluyen la evaluación de la concentración mínima inhibitoria del extracto crudo contra la cepa protipo H37rv, así como la evidente necesidad de repetir el ensayo realizado, sin el ruido del precipitado formado, para corroborar la significancia de los resultados aquí presentados.

## Capítulo 5 Conclusiones

---

### 5.1 Identificación de especies

- 7 especies fueron identificadas molecularmente a través de la amplificación de la subunidad I de la citocromo oxidasa mitocondrial (COI), las más abundantes fueron: *Aplysina revillagigedi*, *Cliona celata* y *Prosuberites laughlini*.

### 5.2 Caracterización del microbioma de las tres especies más abundantes del PNCP

- Los microbiomas de las tres especies de esponjas estudiadas son significativamente diferentes entre ellas ( $P \leq 0.06$ ).
- Las esponjas *Cliona celata* y *Prosuberites laughlini* comparten más del 80% de su microbioma con el agua.
- *Aplysina revillagigedi* fue la esponja con mayor porcentaje de OTUs exclusivos (50%). Presentó alta biodiversidad de especies, entre las cuales figuran las Actinobacterias, un grupo conocido por su capacidad de producir moléculas bioactivas. Esto último también se ve reflejado por el enriquecimiento de genes con potencial bioactivo que arroja el análisis PICRUSt para esta esponja.
- No hubo diferencias significativas dadas por la estación de proveniencia, ni entre réplicas biológicas.

### 5.3 Análisis del metagenoma de *Aplysina revillagigedi*

- Los 5 filo más abundantes en el microbioma de esta esponja fueron: Proteobacteria, Actinobacteria, Firmicutes, Bacteroidetes y Chloroflexi.
- El microbioma de esta esponja presenta genes enriquecidos para la producción de carbohidratos, y subsistemas basados en clusters.
- Se detectaron un total de 38 clusters biosintéticos, la mayor parte de ellos putativos, 5 PKS tipo 1 y ningún NRPS.
- Todos los clusters biosintéticos presentaron un porcentaje de similitud con otras proteínas en las

bases de datos, menor al 70%, lo que sugiere que estos muy probablemente representan compuestos novedosos de PKS y sacáridos.

## 5.4 Bioensayos

- El extracto crudo de *Aplysina revillagigedi*, presentó un efecto antituberculoso significativo contra la cepa prototipo H37rv de *M. tuberculosis*.

## Literatura citada

---

- Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), pp.403-410.
- Antimicrobial resistance (2015). World Health Organization. Fact sheet Number 194. Recuperado el 27 de Junio de 2016 de: <http://www.who.int/mediacentre/factsheets/fs194/en/>.
- Antismash.secondarymetabolites.org. (2017). *antiSMASH bacterial version*. Recuperado el 16 de Octubre de 2017 de: <https://antismash.secondarymetabolites.org/#!/start>
- Azevedo, L., Muccillo-Baisch, A., Filgueira, D., Boyle, R., Ramos, D., Soares, A., Lerner, C., Silva, P. and Trindade, G. (2008). Comparative cytotoxic and anti-tuberculosis activity of *Aplysina caissara* marine sponge crude extracts. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, 147(1), pp.36-42.
- Bai, R.L., Paull, K.D., Herald, C.L., Malspeis, L., Petit, G.R., Hamel, E. (1991). Halichondrin B and homohalichondrin B, marine natural products binding in the vinca domain of tubulin. Discovery of tubulin-based mechanism of action by analysis of differential cytotoxicity data en *J Biol Chem*. 266 (24) 15882-9.
- Barnett, R. and Larson, G. (2011). A Phenol–Chloroform Protocol for Extracting DNA from Ancient Samples. *Methods in Molecular Biology*, pp.13-19.
- Bastida-Zavala, R., De Leon-González, J.A., Carballo-Cenizo, J.L., Moreno-Dávila, B., (2014). Invertebrados bénticos exóticos: esponjas, poliquetos y ascidias, en R. Mendoza y P. Koleff (coords.), *Especies acuáticas invasoras en México*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México, pp 317-336.
- Bayer, T., Neave, M.J., Alsheikh-Hussain, A., Aranda M., Yum, L.K., Mincer, T., Hughen, K., Apprill, A., Voolstra, C., (2013). The Microbiome of the Red Sea Coral *Stylophora pistillata* is Dominated by Tissue-Associated Endozoicomonas Bacteria en *Applied and Environmental Microbiology*. 79 (15) 4759-4762.
- Bernáldez-Sarabia, J., (2013). Aislamiento y caracterización de toxinas de caracoles marinos del género *Conus* con actividad en *Mycobacterium tuberculosis*. Tesis de Doctorado. Centro de Investigación Científica y Educación Superior de Ensenada, Baja California, México.pp: 18-19.
- Blast.ncbi.nlm.nih.gov. (2017). *Nucleotide BLAST: Search nucleotide databases using a nucleotide query*. Recuperado el 16 octubre de 2017 de: [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE\\_TYPE=BlastSearch](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch)
- Blunt, J.W.,Copp, B.R., Munro, M.H., Northcote, P.T., Prinsep, M.R., (2006). Marine natural products en *Nat Prod Rep*. 23: 26-78.
- Bond, W. and Favero, M. (1977). *Bacillus xerothermodurans* sp. nov., a Species Forming Endospores Extremely Resistant to Dry Heat. *International Journal of Systematic Bacteriology*, 27(2), pp.157-160.

- Brusca, R.C. & Brusca, G.J., (2003). Invertebrates. (2<sup>nd</sup>. ed). Sinauer. Sunderland, Massachusetts.
- Bustad, A., Terziivanov, D., Leary, R., Port, R., Schumitzky, A. and Jelliffe, R. (2006). Parametric and Nonparametric Population Methods. *Clinical Pharmacokinetics*, 45(4), pp.365-383.
- Caporaso, J., Lauber, C., Walters, W., Berg-Lyons, D., Lozupone, C., Turnbaugh, P., Fierer, N. and Knight, R. (2010). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(Supplement\_1), pp.4516-4522.
- Carballo, J., Bautista, E., Nava, H., Cruz-Barraza, J. and Chávez, J. (2013). Boring sponges, an increasing threat for coral reefs affected by bleaching events. *Ecology and Evolution*, 3(4), pp.872-886.
- Cimermancic, P., Medema, M., Claesen, J., Kurita, K., Wieland Brown, L., Mavrommatis, K., Pati, A., Godfrey, P., Koehrsen, M., Clardy, J., Birren, B., Takano, E., Sali, A., Lington, R. and Fischbach, M. (2014). Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell*, 158(2), pp.412-421.
- Claverie, J. and Notredame, C. (2007). 2nd ed. [ebook] Indianapolis, Indiana: Wiley Publishing, Inc., pp. 214. Recuperado el 13 de octubre de 2017 de: [https://is.muni.cz/el/1431/podzim2015/Bi9060/um/Claverie\\_Bioinformatics\\_for\\_Dummies\\_2nd\\_Edition\\_2007.pdf](https://is.muni.cz/el/1431/podzim2015/Bi9060/um/Claverie_Bioinformatics_for_Dummies_2nd_Edition_2007.pdf)
- Comisión Nacional de Áreas Naturales Protegidas (CONANP), (2006). Programa de conservación y manejo parque nacional Cabo Pulmo, México. Recuperado en 2017 de: [http://www.conanp.gob.mx/que\\_hacemos/pdf/programas\\_manejo/cabo.pdf](http://www.conanp.gob.mx/que_hacemos/pdf/programas_manejo/cabo.pdf)
- Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, (2012). Esponjas. Recuperado el 04 de agosto de 2016 de: [http://www.biodiversidad.gob.mx/especies/gran\\_familia/animales/Esponjas/esponjas.html](http://www.biodiversidad.gob.mx/especies/gran_familia/animales/Esponjas/esponjas.html)
- Cory, A.H., Owen, T.C., Bartltrop, J.A., Cory, J.G. (1991). Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. 3 (7). pp: 207-212.
- Costello, D. (1961). [ebook] Washington, D.C.: National Academy of Sciences. Recuperado el 13 de octubre de 2017 de: <http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/wilson-henry.pdf>
- Costello, E., Lauber, C., Hamady, M., Fierer, N., Gordon, J. and Knight, R. (2009). Bacterial Community Variation in Human Body Habitats Across Space and Time. *Science*, 326(5960), 1694-1697.
- Cruz-Barraza, J., Carballo, J., Rocha-Olivares, A., Ehrlich, H. and Hog, M. (2012). Integrative Taxonomy and Molecular Phylogeny of Genus *Aplysina* (Demospongiae: Verongida) from Mexican Pacific. *PLoS ONE*, 7(8), 42049.
- Della Sala, G., Hochmuth T., Teta R., Constatino, V., Mangoni, A., (2014). Polyketide synthases in the Microbiome of the Marine Sponge *Plakortis halichondrioides*: A metagenomic Update en *Marine Drugs*. 12: 5425-5440. DOI: 10.3390/md12115425.

- Diaz, C.M., Rützler, K. (2001). Sponges: an essential component of Caribbean coral reefs en *Bull Mar Sci.* 69: 535–546
- Edgar, R. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), pp.2460-2461.
- Fertziger, I., García, F., Godínez, G., Godoy, P., Vizzuett, C., De la Garza M., Pesenti, C. (2012). Plan estratégico para el desarrollo turístico en Cabo Pulmo. Amigos para la Conservación de Cabo Pulmo (ACCP). Recuperado en 2017 de: <http://cabopulmoamigos.org/assets/plan-estrat%C3%A9tico-accp-2012.pdf>
- Fiore, C.L., Baker, D.C., Lesser, M.P. (2013) Nitrogen Biogeochemistry in the Caribbean Sponge, *Xestopongia muta*: A Source or Sink of Dissolved Inorganic Nitrogen? en *PloS One*. 8(8): e72961: doi: [10.1371/journal.pone.0072961](https://doi.org/10.1371/journal.pone.0072961).
- Fischbach, M. and Walsh, C. (2006). Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *ChemInform*, 37(44).
- Fraune, S. & Bosch, T.C.G., (2010). Why bacteria matter in animal development and evolution en *Bioessays*.31: 571-580.
- Friedrich, A.B., Fischer, I., Proksch, P., Hacker, J., Hentschel, U., (2001). Temporal variation of the microbial community associated with the Mediterranean sponge *Aplysina aerophoba* en *FEMS Microbiol Ecol*.38: 105-113.
- Gatti, S. (2002). The Role of Sponges in High-Antarctic Carbon and Silicon Cycling – a Modelling Approach en *Bericht zur Polar und Meeresforschung*. 434: 1–134.
- Gazave, E., Lapébie, P., Ereskovsky, A.V., Vacelet, J., Renard, E., Cárdenas, P., Borchiellini, C., (2011). No longer Demospongiae: *Homoscleromorpha* formal nomination as a fourth class of Porifera en *Hydrobiologia*. 687(1): 3-10. DOI: 10.1007/s10750-011-0842-x.
- Glassing, A., Dowd, S., Galandiuk, S., Davis, B. and Chiodini, R. (2016). Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathogens*, 8(1).
- Gower, J.C. (1966) Some distance properties of latent root and vector methods used in multivariate analysis en *Biometrika* 53: 325-338.
- Grozdanov, L., & Hentschel, U., (2007). An environmental genomics perspective on the diversity and function of marine sponge-associated microbiota en *Current Opinion in Microbiology*. 10:215-220. DOI: 10.1016/j.mib.2007.05.012
- Hacquard, S., Garrido-Oter, R., Gonzalez, A., Spaepen, S., Ackermann, G., Lebeis, S., McHardy, A.C., Dangl, J.L., Knight, R., Ley, R., Schulze-Lefert, P. (2015). Microbiota and host nutrition across plant and animal kingdoms en *Cell Host Microbe*. 17: 603-616.
- Hammer, O. (2017) PAST-diversity. Recuperado el 05 de agosto de 2017 de: <http://folk.uio.no/ohammer/past/diversity.html>.

- Hardoim, C.C.P., Esteves A.I.S., Pires, F.R., Goncalves, J.M.S., Cox, C.J., Xavier, J.R., Costa, R. (2012). Phylogenetically and Spatially Close Marine Sponges Harbour Divergent Bacterial Communities en PLoS ONE. 7(12). DOI: 10.1371/journal.pone.0053029.
- Harris, A. (2017). *Sponge Cells*. [online] Labs.bio.unc.edu. Recuperado el 13 de Oct. de 2017: <http://labs.bio.unc.edu/Harris/Research/sponges.htm>.
- Hentschel, U., Fieseler, L., Wehrl, A., Gernert, C., Steinert, M., Hacker, J., Horn, M. (2003). Microbial diversity of marine sponges en Müller WEG (ed) *Sponges (Porifera)*. Springer, Berlin Heidelberg New York, pp 59–88.
- Hentschel, U., Piel, J., Degnan S.M., Taylor, M.W. (2012). Genomic insights into the marine sponge microbiome en *Nature Reviews Microbiology*. DOI: 10.1038/nrmicro2839.
- Hochmuth, T. y Piel, J. (2009). Polyketide synthases of bacterial symbionts in sponges- Evolution- based applications in natural products research en *Phytochemistry*.70: 1841-1849.
- Huttenhower.sph.harvard.edu. (2017). *Galaxy*. Recuperado el 16 octubre de 2017 de: [http://huttenhower.sph.harvard.edu/galaxy/root?tool\\_id=PICRUST\\_normalize](http://huttenhower.sph.harvard.edu/galaxy/root?tool_id=PICRUST_normalize)
- Hug, L., Thomas, B., Sharon, I., Brown, C., Sharma, R., Hettich, R., Wilkins, M., Williams, K., Singh, A. and Banfield, J. (2015). Critical biogeochemical functions in the subsurface are associated with bacteria from new phyla and little studied lineages. *Environmental Microbiology*, 18(1), pp.159-173.
- Jeong, J.B, Kim, K.H and Park, J.S. (2015). Sponge-Specific Unknown Bacterial Groups Detected in Marine Sponges Collected from Korea Through Barcoded Pyrosequencing en *J. Microbiol. Biotechnol.* 25: 1–10.
- Kazanjian, A. & Fariñas, M. (2006). Actividades biológicas del extracto acuoso de la esponja *Aplysina lacunosa* (Porifera: Aplysinidae). *Rev.Biol.Trop.* 54(3). pp:189-200.
- Kennedy, J. Marchesi, J.R., Dobson, A.D.W. (2007). Metagenomic approaches to exploit the biotechnological potential of the microbial consortio of marine sponges en *Appl Micrbiol Biotechnol.* 75: 11-20. DOI: 10.1007/S00253-007-0875-2.
- Kim, T. and Fuerst, J. (2006). Diversity of polyketide synthase genes from bacteria associated with the marine sponge *Pseudoceratina clavata*: culture-dependent and culture-independent approaches. *Environmental Microbiology*, 8(8), pp.1460-1470.
- Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., Schloss, P.D. (2013). Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform en *Applied and Environmental Microbiology*. 79 (17): 5112-5120. DOI: 10.1128/AEM.01043-13.
- Kumar, S., Stecher, G. and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), pp.1870-1874.
- Langille, M. G.I.\*; Zaneveld, J.\*; Caporaso, J. G.; McDonald, D.; Knights, D.; a Reyes, J.; Clemente, J. C.; Burkpile, D. E.; Vega Thurber, R. L.; Knight, R.; Beiko, R. G.; and Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences en *Nature Biotechnology*.8.pp: 1-10.

- Laport, M., Santos, O. and Muricy, G. (2009). Marine Sponges: Potential Sources of New Antimicrobial Drugs. *Current Pharmaceutical Biotechnology*, 10(1), pp.86-105.
- Lee OO, Chui PY, Wong YH, Pawlik JR, Qian PY. (2009). Evidence for vertical transmission of bacterial symbionts from adult to embryo in the Caribbean sponge *Svenzea zeai*. *Appl Environ Microbiol* 75: 6147–6156.
- Li, C.W., Chen, J.Y., Hua, T.E. (1998). Precambrian Sponges with Cellular Structures. *Science*, 279(5352), pp.879-882.
- Li, D., Liu, C., Luo, R., Sadakane, K. and Lam, T. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10), pp.1674-1676.
- Li, K., Bihan, M., Yooseph, S., Methé, B.A. (2012). Analyses of the Microbial Diversity across the Human Microbiome en *Plos One*. 7(6). DOI: 10.1371/journal.pone.0032118.
- Loh, T.L & Pawlik, J.R. (2014). Chemical defenses and resource trade-offs structure sponge communities on Caribbean coral reefs en *Proc Natl Sci U.S.A.* 111 (11): 4151-4156.
- Longford, S.R., Tujula, N.A., Crocetti, G.R., Holmes, A.J., Holmstrom, C., Kjelleberg, S., Steinberg, P.D., Taylor, M.W. (2007). Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes en *Aquatic Microbial Ecology*. 48: 217-229.
- Lozupone, C. and Knight, R. (2005). UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology*, 71(12), pp.8228-8235.
- Marinespecies.org. (2017). *World Porifera Database*. Recuperado el 13 de Oct de 2017 de: <http://www.marinespecies.org/porifera/>
- Martin, A., Moncillo, N., Lemus, D., Montoro, E., Telles, M. A., Simboli, N., Pontino, M., Porras, T., León, C., Velasco, M., Chacon, L., Barrera, L., Ritacco, V., Portales, F., Palomino, J. C. (2005). Multicenter study on MTT and resazurin assays for susceptibility testing to first-line antituberculosis drugs en *International Journal of Tuberculosis and Lung Diseases*. 9: 901-906.
- Medema, M., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M., Weber, T., Takano, E. and Breitling, R. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research*, 39(suppl\_2), pp.W339-W346.
- Metagenomics.anl.gov. (2017). *MG-RAST*. Recuperado el 13 octubre de 2017 de: <http://metagenomics.anl.gov/>
- Meyer, C.P. (2003). Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics en *Biological Journal of the Linnean Society*. 79: 401-459.
- Minowa, Y., Araki, M. and Kanehisa, M. (2007). Comprehensive Analysis of Distinctive Polyketide and Nonribosomal Peptide Structural Motifs Encoded in Microbial Genomes. *Journal of Molecular Biology*, 368(5), 1500-1517

- Morrow, C & Cárdenas, P. (2015). Proposal for a revised classification of the Demospongiae (Porifera) en Front. Zool.12:1-27.
- Muscente, A.D, Marc, M.F., Dale, J.G, Xiao, S. (2015). Assessing the veracity of Precambrian ‘sponge’ fossils using in situ nanoscale analytical techniques en Precambrian Res. 263: 142–156.
- Napdos.ucsd.edu. (2017). *NaPDoS | Home*. Recuperado el 13 de octubre de 2017 de: <http://napdos.ucsd.edu>.
- Nakashima, Y., Egami, Y., Kimura, M., Wakimoto, T. and Abe, I. (2016). Metagenomic Analysis of the Sponge Discodermia Reveals the Production of the Cyanobacterial Natural Product Kasumigamide by ‘Entotheonella’. *PLOS ONE*, 11(10), p.e0164468.
- Nmpdr.org. (2017). *ClusteringBasedSubsystem < FIG < NMPDR*. Recuperado el 16 octubre de 2017 de: <http://www.nmpdr.org/FIG/wiki/view.cgi/FIG/ClusteringBasedSubsystem>
- Oulas, A., Pavloudi, C., Polymenakou, P., Pavlopoulos, G., Papanikolaou, N., Kotoulas, G., Arvanitidis, C. and Iliopoulos, I. (2015). Metagenomics: Tools and Insights for Analyzing Next-Generation Sequencing Data Derived from Biodiversity Studies. *Bioinformatics and Biology Insights*, 9, p.BBI.S12462.
- Österlund, T., Nerman, O., Jonsson, E.C., *et al.* (2016). Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics en BMC Genomics. 17:78 DOI 10.1186/s12864-016-2386-y.
- Palaeontologie.geo.uni-muenchen.de. (2017). *Welcome to the Sponge Barcoding Project!* Recuperado el 13 de octubre de 2017 de: <http://www.palaeontologie.geo.uni-muenchen.de/SBP/>
- Peterson, K.J and Butterfield, N.J. (2005). Origin of the Eumetazoa: testing ecological predictions of molecular clocks against the Proterozoic fossil record en *Proc. Natl. Acad. Sci. U. S. A.* 102: 9547–9552.
- Qasba, P. and Ramakrishnan, B. (2007). X-Ray Crystal Structures of Glycosyltransferases. *Comprehensive Glycoscience*, 2, pp.251-281.
- Qiime.org. (2017a). *454 Overview Tutorial: de novo OTU picking and diversity analyses using 454 data — Homepage*. Recuperado el 13 de octubre de 2017 de: <http://qiime.org/tutorials/tutorial.html>
- Qiime.org. (2017b). *group\_significance.py – Compare OTU frequencies across sample groups — Homepage*. Recuperado el 16 octubre de 2017: [http://qiime.org/scripts/group\\_significance.html](http://qiime.org/scripts/group_significance.html)
- Quality Scores for Next Generation Sequencing: Assessing sequencing accuracy using Phred quality scoring. (2011). [ebook] San Diego, CA: Illumina Inc., p.1. Recuperado el 16 octubre. de 2017 de: [https://www.illumina.com/documents/products/technotes/technote\\_Q-Scores.pdf](https://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf).
- Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology*, 62(2),142-160.

- Reveillaud, J., Maignien, L., Eren, A.M., Huber, J.A., Apprill, A., Sogin, L.M., Vanreusel, A. (2014). Host-specificity among abundant and rare taxa in the sponge microbiome en *ISME J.* 8: 1198-1209.
- Rutzler K. (1985) Associations between Caribbean sponges and photosynthetic organisms en Rutzler K (ed). *New perspectives in sponge biology*. Smithsonian Institution Press, Washington, pp 455–466
- Ryu, T., Seridi, L., Moitinho-Silva, L., Oates, M., Jin Liew, Y., Mavromatis, C., Wang, X., Haywood, A., Lafi, F.F., Kupresanin, M., Sougrat R., Alzahrani, M.A., Giles, E., Ghosheh, Y., Schunter, C., Baumgarten, S., Berumen, M.L., Gao, X., Aranda, M., Foret, S., Gough, J., Voolstra C.R., Hentschel, U., Ravasi, T. (2016). Hologenome analysis of two marine sponges with different microbiomes en *BMC Genomics*. 14-158. DOI: 10.1186/s12864-016-2501-0.
- Salter, S., Cox, M., Turek, E., Calus, S., Cookson, W., Moffatt, M., Turner, P., Parkhill, J., Loman, N. and Walker, A. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, 12(1).
- Schmitt, S., Weisz, J.B., Lindquist, N., and Hentschel, U. (2007) Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix* en *Appl Environ Microbiol.* 73: 2067-2078.
- Schmitt, S., Angermeier, H., Schiller, R., Lindquist, N., and Hentschel, U. (2008) Molecular microbial diversity survey of sponge reproductive stages and mechanistic insights into vertical transmission of microbial symbionts en *Appl Environ Microbiol.* 74: 7694-7708.
- Schmitt, S., Deines, P., Behnam, F., Wagner, M., and Taylor, M.W. (2012) *Chloroflexi* bacteria are more diverse, abundant, and similar in high than in low microbial abundance sponges en *FEMS Microbiol Ecol* 78: 497-510.
- Schwarzer, D., Finking, R. and Marahiel, M. (2003). Nonribosomal Peptides: From Genes to Products. *ChemInform*, 34(38).
- Shearer, T.L, Van Oppen, M.J.H, Romano, S.L, Worheide, G. (2002). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria) en *Molecular Ecology*. 11: 2475–2487.
- Shirmer, A., Gadkari, R., Reeves, C.D., Ibrahim, F., DeLong, E.F., Hutchinson, C.R. (2005). Metagenomic analysis reveals diverse polyketide synthase gene Clusters in microorganisms associated with the marine sponge *Discodermia dissoluta* en *Appl Environ Microbiol.* 71: 4840-4849.
- Sharp, K.H., Eam, B., John Faulkner, D., and Haygood, M.G. (2007) Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp en *Appl Environ Microbiol.* 73: 622-629.
- Somes, R. (2017). SDR-IV Help. Recuperado el 05 de agosto de 2017 de: <http://www.piscesconservation.com/sdrhelp/index.html>.
- Spongeguide.org. (2017). *The Sponge Guide*. Recuperado el 13 de Oct. de 2017 de: <http://www.spongeguide.org/>
- Takahiro, A.B.E., Pinar-Sahin, F., Akiyama, K., Naito, T., Kishigami, M., Miyamoto, K., Sakakibara, Y., Uemura, D. (2012). Construction of a metagenomic Library for the Marine Sponge *Halichondria okadai* en *Biosci. Biotechnol. Biochem.* 76(4): 633-639. DOI: 10.1271/bbb.110533.

- Taylor, M.W., Radax, R., Steger, D. Wagner, M. (2007). Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential en *Microbiology and Molecular Biology Reviews*. 295-347. DOI: 10.1128/MMBR.00040-06.
- Thomas, T., Gilbert, J. and Meyer, F. (2012). Metagenomics - a guide from sampling to data analysis. *Microbial Informatics and Experimentation*, 2(1), p.3.
- Thomas, T., Kavlar, D., Lokabharathi, P. (2010). Marine Drugs from Sponge-Microbe Association en *Marine Drugs*. pp. 1417-1468.doi: 10.3390/md8041417.
- Thomas, T., Rusch, D., DeMaere, M., Yung, P., Lewis, M., Halpern, A., Heidelberg, K., Egan, S., Steinberg, P. and Kjelleberg, S. (2010). Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *The ISME Journal*, 4(12), pp.1557-1567
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J., Easson, C., Astudillo-García, C., Olson, J., Erwin, P., López-Legentil, S., Luter, H., Chaves-Fonnegra, A., Costa, R., Schupp, P., Steindler, L., Erpenbeck, D., Gilbert, J., Knight, R., Ackermann, G., Victor Lopez, J., Taylor, M., Thacker, R., Montoya, J., Hentschel, U. and Webster, N. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nature Communications*, 7, p.11870.
- Uriz, M.J. (1986). Clave de identificación de las esponjas más frecuentes de la península Ibérica en *Misc Zool*. 10: 7-22.
- Van Soest R.W.M, Boury-Esnault N., Hooper J.N.A., Rützler K., de Voogd N.J., Alvarez de Glasby B., Hajdu E., Pisera A.B., Manconi R., Schoenberg C., Klautau M., Picton B., Kelly M., Vacelet J., Dohrmann M., Díaz M.-C., Cárdenas P. (2017). WoRMS Porifera: World Porifera database. Consultado el día 5 de Agosto de 2017 en : [www.catalogueoflife.org/col](http://www.catalogueoflife.org/col).
- Vargas, S., Schuster, A., Sacher, K., Büttner, G., Schätzle, S., Lächli, B., Hall, K., Hooper, J., Erpenbeck, D. and Wörheide, G. (2012). Barcoding Sponges: An Overview Based on Comprehensive Sampling. *PLoS ONE*, 7(7), p.e39345.
- Walters, W., Caporaso, J., Lauber, C., Berg-Lyons, D., Fierer, N. and Knight, R. (2011). PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics*, 27(8), 1159-1161.
- Wang, G. (2006). Diversity and biotechnological potential of the sponge-associated microbial consortia en *Jind Microbiol Biotechnol*. 33: 545-551.
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H., Brucoleri, R., Lee, S., Fischbach, M., Müller, R., Wohlleben, W., Breitling, R., Takano, E. and Medema, M. (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research*, 43(W1), W237-W243.
- Webster, N. and Thomas, T. (2016). The Sponge Hologenome. *mBio*, 7(2), pp.e00135-16.
- Wehrl, M., Steinert, M., and Hentschel, U. (2007) Bacterial uptake by the marine sponge *Aplysina aerophoba* en *Microb Ecol*. 53: 355-365.
- Williams, G. (2013). Engineering polyketide synthases and nonribosomal peptide synthetases. *Current Opinion in Structural Biology*, 23(4), 603-612.

- Wilkinson, C.R. (1992). Symbiotic interactions between marine sponges and algae en Reisser W (ed) *Algae and symbioses: plants, animals, fungi, viruses, interactions explored*. Biopress, Bristol, pp.113–151.
- World Health Organization. (2017). *World Health Statistics 2015*. [online] Available at: [http://www.who.int/gho/publications/world\\_health\\_statistics/2015/en/](http://www.who.int/gho/publications/world_health_statistics/2015/en/) [Accessed 13 Oct. 2017].
- Yadav, G., Gokhale, R.S., Mohanty, D. (2009). Towards Prediction of Metabolic Products of Polyketide Synthases: An In Silico Analysis en *PLOS Computational Biology*. 5(4).
- Yin, Z., Zhu, M., Davidson, E., Bottjer, D., Zhao, F. and Tafforeau, P. (2015). Sponge grade body fossil with cellular resolution dating 60 Myr before the Cambrian. *Proceedings of the National Academy of Sciences*, p.201414577.
- 16S Metagenomic Sequencing Library Preparation: Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System. (2013). [ebook] San Diego, CA: Illumina Inc., pp.1-28. Recuperado el 16 Oct. de 2017 de: [https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry\\_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf](https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf).

# Anexo 1

## Alineamientos por BLASTn de las 7 especies de esponja identificadas

Max score	Total score	Query cover	E value	Ident	Accession
1098	1098	100%	0.0	98%	<a href="#">AY561960.1</a>

'*Prosuberites*' *laughlini* voucher UCMPWC875 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial  
 Sequence ID: [AY561960.1](#) Length: 709 Number of Matches: 1

Range 1: 42 to 666 [GenBank](#) [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1098 bits(594)	0.0	613/625(98%)	1/625(0%)	Plus/Plus
Query 1	TTGGGAGCATTAGTGGGATGATAGGAACAGCCTTTAGTGTGCTTATTAGATTAGAATTA			60
Sbjct 42	TTGGGAGCATTAGTGGGATKATAGGAACAGCCTTTAGTGTGCTTATTAGATTAGAATTA			101
Query 61	TCCGCTCCAGGTTCAATGTTAGGTGATGATCAGTTGTATAATGTTATAGTAACGCTCAT			120
Sbjct 102	TCCGCTCCAGGTTCAATGTTAGGTGATGATCAGTTGTATAATGNTATAGTAACGCTCAT			161
Query 121	GCCTTTGTAATGATAATCTTTCTGGTTATGCCGGTGTAAATCGGAGGGTTTGGAACTGA			180
Sbjct 162	GCCTTTGTAAATGATAATCTTTCTGGTTATGCCGGTGTAAATCGGAGGGKTTGGAACTGA			221
Query 181	TTCGTACCCCTTATACATAGGGGCTCCAGATATGGCGTTTCCAAGGCTAAATAATATCAGT			240
Sbjct 222	TTCGNACCCCTTATACATAGGGGCTCCAGATATGGCGTTTCCAAGGCTAAATAATATCAGT			281
Query 241	TTTTGACTATTACCCCATCATTAACCTTTATTACTAGGTCTGCTTTTGTAGAGCAGGGG			300
Sbjct 282	TTTTGACTATTACCCCATCATTAACCTTTATTACTAGGTCTGCTTTTGTAGAGCAGGGG			341
Query 301	GCCGGTACAGGGTGGACAGTTTACCCGCCGTTAGCAAGTATACAAGCAGCATTCCGGGGGA			360
Sbjct 342	GCCGGTACAGGGTGGACAGTCTACCCGCCGTTAGCAAGTATACAAGCAGCATTCCGGGGGA			401
Query 361	TCAGTAGACATGGTTATTTTAGTTTTCATTTAGCTGGTATTTCTTCAATATTAGGGGCG			420
Sbjct 402	TCAGTAGACATGGTTATTTTAGTCTTCATTTAGCTGGTATTTCTTCAATATTAGGGGCG			461
Query 421	ATGAATTTTATCACAACATCTTTAATATGAGGGGCCAGGGATAACAATGGATAAAATG			480
Sbjct 462	ATGAATTTTATCACAACATCTTTAATATGAGGGGCCAGGGATAACAATGGATAGAATG			521
Query 481	CCATTATTTGTGTGATCTATTTTAGTAACAGTATTTTATTGTTATTGTCTCTGCCAGTT			540
Sbjct 522	CCATTATTTGTGTGATCTATTTTAGTAACAGTATTTTATTGTTATTGWCTCTGCCAGTT			581
Query 541	TTAGCCGGGGCTATAACTATGCTTCTTACAGATAGAACTTTAACACTACC-TTTTGGAC			599
Sbjct 582	TTAGCCGGGGCTATAACTATGCTTCTTACAGATAGAACTTTAACACTACC-TTTTGGAC			641
Query 600	CCAGCAGGGGGGGTGGACCCCATAT 624			
Sbjct 642	CCAGCAGGGGGGGTGGACCCCATAT 666			

**Figura 46.** Alineamiento blastn de la secuencia consenso de *Prosuberites laughlini*, muestra 303 de la estación Marios reef.

Max score	Total score	Query cover	E value	Ident	Accession
961	961	86%	0.0	99%	<a href="#">JQ437580.1</a>

[Aplysina revillagiedi voucher LEB-ICML-UNAM-1236 cytochrome oxidase subunit I \(COI\) gene, partial cds; mitochondrial](#)

**Aplysina revillagiedi voucher LEB-ICML-UNAM-1236 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial**  
 Sequence ID: [JQ437580.1](#) Length: 523 Number of Matches: 1

Range 1: 1 to 523 [GenBank](#) [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
961 bits(520)	0.0	522/523(99%)	0/523(0%)	Plus/Plus
Query 25	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA			84
Sbjct 1	ATTTAGTATGTTAATTAGACTGGAGTTATCTGCACCAGGCTCAATGTTAGGAGACGATCA			60
Query 85	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATCTTTTTAGTAATGCC			144
Sbjct 61	TTTATATAATGTTATAGTAACGGCACATGCTTTTGTATGATATCTTTTTAGTAATGCC			120
Query 145	TGTTATGATAGGGGTTTTGGTAATTGATTAGTTCATTATATATGGAGCACCAGATAT			204
Sbjct 121	TGTTATGATAGGGGTTTTGGTAATTGATTAGTTCATTATATATGGAGCACCAGATAT			180
Query 205	GGCTTTCCTCGGTTAAATAATATTAGTTTTGATTATTACCCCAGCACTAATTTTATT			264
Sbjct 181	GGCTTTCCTCGGTTAAATAATATTAGTTTTGATTATTACCCCAGCACTAATTTTATT			240
Query 265	ATTGGGGTCTGCTTTTATAGAACAAGGGCCGGTACGGGTTGAACCGTCTATCCGCCATT			324
Sbjct 241	ATTGGGGTCTGCTTTTATAGAACAAGGGCCGGTACGGGTTGAACCGTCTATCCGCCATT			300
Query 325	ATCAAGTATACAAACACATCTGGGGTTCGGTAGATATGGCAATATTTAGTCTTCATTT			384
Sbjct 301	ATCAAGTATACAAACACACTCTGGGGTTCGGTAGATATGGCAATATTTAGTCTTCATTT			360
Query 385	GGCCGGAATTCCTCCATATTAAGCTCTATTAATTTTATAACTACAATTTTAAATATGAG			444
Sbjct 361	GGCCGGAATTCCTCCATATTAAGCTCTATTAATTTTATAACTACAATTTTAAATATGAG			420
Query 445	GGCGCCAGGGATGACTATGGACAGATTACCTTTATTCGTTTGATCTATTTTATTAACAAC			504
Sbjct 421	GGCGCCAGGGATGACTATGGACAGATTACCTTTATTCGTTTGATCTATTTTATTAACAAC			480
Query 505	AATATTATTAGTATTAGCATTGCCCGTTTTAGCTGGAGCAATT		547	
Sbjct 481	AATATTATTAGTATTAGCATTGCCCGTTTTAGCTGGAGCAATT		523	

**Figura 47.** Alineamiento blastn de la secuencia consenso de *Aplysina revillagiedi*, muestra 301 de la estación Marios Reef.

[Timea sp. G303973 cytochrome oxidase subunit I \(COI\) gene, partial cds; mitochondrial](#)

Max score	Total score	Query cover	E value	Ident	Accession
1116	1116	100%	0.0	99%	<a href="#">AY561968.1</a>

**Timea sp. G303973 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial**  
**Sequence ID: [AY561968.1](#) Length: 709 Number of Matches: 1**

Range 1: 42 to 653 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1116 bits(604)	0.0	610/613(99%)	1/613(0%)	Plus/Plus
Query 1	TTTGGAGCTTTTGCAGGAATGATCGGGACGGCCTTTAGTATGTTAATACGGCTGGAAC TA			60
Sbjct 42	TTTGGAGCTTTTGCAGGAATGATCGGGACGGCCTTTAGYATGTTAATACGGCTGGAAC TA			101
Query 61	TCTTCGCCAGGATCAATGCTAGGGGACGATCATTATATAACGTTCTTGTAAC TGC GCAT			120
Sbjct 102	TCTTCGCCAGGATCAATGCTAGGGGACGATCATTATATAACGTTCTTGTAAC TGC GCAT			161
Query 121	GCTTTTGTAATGATA tttttttt AGTTATGCCGTAATGATTGGAGGGTTTGGTAAC TGA			180
Sbjct 162	GCTTTTGTAATGATA tttttttt TAGTTATGCCGTAATGATTGGAGGGTTTGGTAAC TGA			221
Query 181	CTGGTGCCCTTGATATATGGAGCGCCTGATATGGCTTTTCCTCGATTAAATAATATAAG T			240
Sbjct 222	CTGGTGCCCTTGATATATGGAGCGCCTGATATGGCTTTTCCTCGATTAAATAATATAAG T			281
Query 241	TTTTGATTATTGCCCCGCCTTTAAGTTTATTGTTGGCGTCTGCTTTCGTTGGAGCAAGGG			300
Sbjct 282	TTTTGATTATTGCCCCGCCTTTAAGTTTATTGTTGGCGTCTGCTTTCGTTGGAGCAAGGG			341
Query 301	GCAGGAACAGGATGAACTGTTTATCCCCCTCTCTCGGTAATCAGGCTCATTCTGGCGGG			360
Sbjct 342	GCAGGAACAGGATGAACTGTTTATCCCCCTCTCTCGGTAATCAGGCTCATTCTGGCGGG			401
Query 361	TCAGTTGATTAGTAATCTTTAGCCTACATTTAGCCGGAATTTCTTCCATTTTAGGCGCT			420
Sbjct 402	TCAGTTGATTAGTAATCTTTAGCCTACATTTAGCCGGAATTTCTTCCATTTTAGGCGCT			461
Query 421	ATGAATTTTATTACTACAATAGTAAACATGCGGATAACGGGGATGACTTTTGATAGGTTG			480
Sbjct 462	ATGAATTTTATTACTACAATAGTAAACATGCGGATAACGGGGATGACTTTTGATAGGTTG			521
Query 481	CCTTTATTTGTTTGATCAATTTTAGTTACCGCGGTTTTATTGTTATTGTCCTTACCTGTA			540
Sbjct 522	CCTTTATTTGTTTGATCAATTTTAGTTACCGCGGTTTTATTGTTATTGTCCTTACCTGTA			581
Query 541	TTAGCGGGGGCTATTACCATGTTGTTAACCGACAGGAATTTAACTACTGC tttttttt GA			600
Sbjct 582	TTAGCGGGGGCTATTACCATGTTGTTAACCGACAGGAATTTAAC -ACTGCTTTTTTTGA			640
Query 601	TCCAGCGGGGGGA 613			
Sbjct 641	TCCAGCGGGGGGA 653			

**Figura 48.** Alineamiento por blastn de la secuencia consenso de *Timea sp*, muestra 302 de Marios Reef.

[Spongia sp. Alg15\\_240 cytochrome oxidase subunit I \(COI\) gene, partial cds; mitochondrial](#)

Max score	Total score	Query cover	E value	Ident	Accession
966	966	99%	0.0	96%	<a href="#">KX574853.1</a>

**Spongia sp. Alg14\_230 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial**  
 Sequence ID: [KX574847.1](#) Length: 636 Number of Matches: 1

Range 1: 1 to 587 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
966 bits(523)	0.0	569/591(96%)	4/591(0%)	Plus/Plus
Query 3	TGGGTTTATCTTTAAGTTTGATAATGCGGTTGGAGTTGTCTTCCCCGGGAACCTTTATTGG			62
Sbjct 1	TGGGTTTATCTTTAAGTCTGATAATGCGGTTTCGAGTTGTCTTCCCCGGGAACCTTTATTGG			60
Query 63	GGGATGATAATATATATAATGTTGTTGTAACAGCACATGGTTTGATAATGGTCtttttttt			122
Sbjct 61	GGGATGATACTATATATAATGTTGTTGTAACAGCACATGGTTTGATAATGGTCTTTTTTTT			120
Query 123	tGGTAATGCCTGTATTGATCGGGGGGTTTGGTAATTGGTTTGTTCCCCTGTATATAGGGT			182
Sbjct 121	TGGTAATGCCTGTATTGATCGGGGGGTTTGGTAATTGGTTTGTTCCCCTGTATATCGGGT			180
Query 183	CTCCGGATATGGCGTCCCAAGCTAAATAATTTAAGTTTTTGGTTATTACCGCCTGCAT			242
Sbjct 181	CGCCCGATATGGCGTCCCACGGCTAAATAATTTAAGTTTTTGGTTATTGCCCGGCAT			240
Query 243	TTATTATGTTAGTTGGGTCGTTTTTGTGGAAGAGGGGGCCGGACCGGCTGGACTTTGT			302
Sbjct 241	TTATTATGTTAGTTGGGTCGTTTTTGTGGAAGAGGGGGCCGGAACCGGCTGGACTTTGT			300
Query 303	ACCCTCCGCTATCTGGCGGGATGGGGCATTCCGGGCCATCTGTGGATATGGCTATTTTTA			362
Sbjct 301	ACCCCCGCTATCTGGCGGCATGGGGCATTCCGGGCCATCTGTGGATATGGCTATTTTTA			360
Query 363	GTTTGCATTTGGCGGGTATATCTTCTATCTTGGGGGCGATTAATTTTATTAGTACTATTT			422
Sbjct 361	GTTTGCATTTGGCGGGTATATCTTCTATCTTGGGGGCGATTAATTTTATTAGTACTATTT			420
Query 423	TTAATATGCGGGCGTCTGGAGTGGGTTGAACTCGAATCCCTTTGTTTGTATGATCTATTT			482
Sbjct 421	TTAATATGCGGGCGTCCGGAGTAGGTTGAACTCGAATGCCTTTGTTTGTATGATCTATTT			480
Query 483	TAGTTCCAGTGTTTTTATTATTATTGTCATTACCAGTATTGGCCGGGGGAATAACTATGT			542
Sbjct 481	TAGTTACAGTGTTTTTATTATTATTGTCATTACCAGTATTGGCCGGGGGAATAACTATGT			540
Query 543	TATTAACGATCGGAATTTTTAATATCTTCTTTCTTTGATCCCGGGGGA			593
Sbjct 541	TATTAACGATCGGAATTTT-AATA-CTT-CTTCTTTGATCC-GGTGGGA			587

**Figura 49.** Alineamiento por blastn de la secuencia consenso de *Spongia sp.*, muestras 212 de la estación El Cantil.

[Clathria prolifera voucher SERCINVERT0452 cytochrome oxidase subunit 1 \(COI\) gene, partial cds; mitochondrial](#)

Max score	Total score	Query cover	E value	Ident	Accession
915	915	99%	0.0	93%	<a href="#">KU905733.1</a>

**Clathria prolifera voucher SERCINVERT0452 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial**

Sequence ID: [KU905733.1](#) Length: 657 Number of Matches: 1

[▶ See 1 more title\(s\)](#)

Range 1: 14 to 635 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
915 bits(495)	0.0	581/623(93%)	3/623(0%)	Plus/Plus
Query 1	GTTTTGGAATTTTAGCAGGAATGATTGGTACTGCTTTTAGTATGTTAATTAGGTTAGAGT	60		
Sbjct 14	GTTTTGGGATTTTAGCAGGAATGATTGGTACTGCTTTCAGTATGTTAATTAGATTAGAGT	73		
Query 61	TATCGGGACCGGGATCAATGTTGGGGGATGATCATTATATAATGTTATTGTAAGTCTC	120		
Sbjct 74	TATCGGGACCGGGATCAATGTTAGGGGATGATCATTATACAATGTTATTGTTACTGCTC	133		
Query 121	ATGCTTTTGTATGAttttttttttAGTTATGCCTGTCATGATTGGGGGATTTGGTAATT	180		
Sbjct 134	ATGCTTTTGTATGATTTTTTTTtagtaATGCCTGTGATGATTGGGGGTTTGGGAATT	193		
Query 181	GGTTAGTGCCATTATATAATGGTGCCTCCAGATATGGCTTTCCCTCGATTAAATAATATAA	240		
Sbjct 194	GGTTAGTACCATTATATAATGGTGCCTCCAGATATGGCATTCCCTCGATTAAATAATATAA	253		
Query 241	GTTTTTGATTATTACCTCCGCCTTTAAGTTTATTATTAGTTTCTGCTTTTGTGGAGCAAG	300		
Sbjct 254	GTTTTTGATTATTACCGCCGGCTTTAAGTTTATTATTAGCTTCTGCTTTTGTGGAGCAAG	313		
Query 301	GGCCCGGAACAGGATGAACAGTTTATCCTCCTTTATCAGGGATTCAGGTTTCATTCCGGGG	360		
Sbjct 314	GGCCCGGAACAGGATGAACAGTTTATCCGCCTTTATCAGGGATTCAGGTCATTCTGGGG	373		
Query 361	G TTCAGTAGATTTAGTTATTTTTAGTTTACATTTGGCGGGAATTCCTCTATTTTAGGTG	420		
Sbjct 374	G TTCAGTAGATTTAGTTATTTTTAGTTTACATTTGGCGGGAATTCATCTATATTAGGTG	433		
Query 421	CTATGAATTTTATTACAACAATAATAATATGAGAGCCCCGGGGATAACTTTTGATAAAA	480		
Sbjct 434	CAATGAATTTTATTACAACAATAATAATATGAGAGCGCCGGGAATAACTTTTGATAGAA	493		
Query 481	TGCCTTTATTTGTTTGATCAATTTTAGTAAGTCTGTTTTATTATTATTATCTTTACCAG	540		
Sbjct 494	TGCCATTATTTGTTTGATCTATTTTAGTAAGTCAATTTTATTATTATTGTCCTTTACCGG	553		
Query 541	TATTAGCTGGGGCTATAACTATGTTATTAACAGATAGAAATTTAATACCAGCATTCTT-	599		
Sbjct 554	TATTAGCGGGGCTATAACTATGTTATTAACAGATAGAAATTTAATACT-GCATTTTTT	612		
Query 600	GAT-CTGCAGGAGGAGCAGACCC 621			
Sbjct 613	GACCCCTGCAGGGGAGGAGACCC 635			

**Figura 50.** Alineamiento por blastn de la secuencia consenso de *Clathria prolifera*, de la muestra 407 de La Esperanza.

Max score	Total score	Query cover	E value	Ident	Accession
-----------	-------------	-------------	---------	-------	-----------

[Phelloderma oxychaetoides mitochondrial partial COI gene for cytochrome oxidase subunit 1, specimen voucher SMF<DEU> 11726](#) 573 573 99% 2e-159 89% [HE611626.1](#)

**Phelloderma oxychaetoides mitochondrial partial COI gene for cytochrome oxidase subunit 1, specimen voucher SMF 11726**

Sequence ID: [HE611626.1](#) Length: 629 Number of Matches: 1

**Related**

Range 1: 1 to 472 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
573 bits(310)	2e-159	418/472(89%)	0/472(0%)	Plus/Plus
Query 2	CGGGGATGATAGGCAC	TGCCTTGAGCATGTTAATTCGATTGGAGTTATCTGCGCCAGGGT		61
Sbjct 1	CGGGGATGATGGGTACCTCCTTTAGTATGTTAATTCGATTGGAGTTATCTGCGCCAGGGT			60
Query 62	CAATGTTGGGAGACGACCATCTATATAATGTTATGTAAGTCTCCCGCTTTTGTCATGA			121
Sbjct 61	CAATGTTGGGAGATGATCATTTATATAATGTTATGTAAGTCTCATGCTTTTGTCATGA			120
Query 122	TCtttttttttATTTATGCCGTAACGATgggggggTTTGGTAATTGGATTGTGCCTTTAT			181
Sbjct 121	TCTTTTTTTTAGTTATGCCGTAATGATGGGGGATTTGGTAATTGGTTGTACCTTTAT			180
Query 182	ATATGGGGCCCCCATATGGCCTTCCGCGATTAAATAATATAAGTTTTGATTATGGC			241
Sbjct 181	ATATTGGTGCCCTGATATGGCTTCCCTCGATTAAATAATATAAGTTTTGATTATTC			240
Query 242	CCCCGCCTTCAAGTTTATTATTGGCGTCTGCTTTTGTAGAGCAAGGGCGGGAACAGGAT			301
Sbjct 241	CTCCGGCTTAAAGTTTATTATTGGCGTCTGCTTTTGTAGAGCAAGGGCGGGAACAGGAT			300
Query 302	GGACAGTTTATCCTCCTTATCCAGCATTTCAGGCCCATTCGGGGGTTTCAGTCGATTTAG			361
Sbjct 301	GGACAGTTTATCCCCCTTATCTGGAATTCAGGCCCATTCGGGGGTTCAATCGACTTAG			360
Query 362	TCATATTTAGTTTACATTTAGCAGGGATTTTTCTATATTGGCGCTATGAATTTTATTA			421
Sbjct 361	TAATATTTAGTTTACATTTAGCGGGTATTTCTTCTATATTGGCGCTATGAATTTTATTA			420
Query 422	CTACAATAATCAATATGAGACCATGGGGTAGGACTTTTGATAAAAGCCATT			473
Sbjct 421	CTACAATAATAAATATGAGACAAGGGGATGACTTTGGATAAAATGCCATT			472

**Figura 51.** Alineamiento por blastn de la secuencia consenso de *Phelloderma sp.*, de la muestra 408, que proviene de La Esperanza.

[Cliona celata voucher Mc4026 cytochrome oxidase subunit I \(cox1\) gene, partial cds; mitochondrial](#)

Max score	Total score	Query cover	E value	Ident	Accession
623	623	97%	2e-174	85%	<a href="#">HM999039.1</a>

**Cliona celata voucher Mc4026 cytochrome oxidase subunit I (cox1) gene, partial cds; mitochondrial**  
**Sequence ID: [HM999039.1](#) Length: 842 Number of Matches: 1**

Range 1: 1 to 621 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
623 bits(337)	2e-174	528/623(85%)	2/623(0%)	Plus/Minus
Query 3	ATAATATGTGAGATCCCCACCACCAGCAGGATCAAAAAATGTTGTATTTAAAAATTTCTGTC			62
Sbjct 621	ATAAAAT-TG-GATCTCCTCCACCTGCAGGATCAAAAAACGCAGTATTTAAAAATTTCTATC			564
Query 63	TGTAAGAAGCATTGTTATTGCCCTGCTAACACTGGTAAAGATAATAACAATAAAAACAGC			122
Sbjct 563	TGTAAGAAGCATTGTAATCGCCCCAGCTAACACTGGTAAAGACAATAATAATAAAAAAGC			504
Query 123	TGTGACTAAAAATAGACCACACAAATAATGGCATTCTGTCCATTGTAATCCCTGGTGCCT			182
Sbjct 503	AGTAACTAAAAATAGACCACACAAATAACGGCATTTCGATCCATTGTTATGCCCGAGCCCT			444
Query 183	CATATTTAAAAATGTTGTAATAAAATTCATTGCTCCTAATATTGAAGATATTCGGCTAA			242
Sbjct 443	CATATTTAAATATTGTAGTAATAAAATTCATTGCACCAATATAGAAGAAATCCCGCTAA			384
Query 243	ATGAAAATAAATATAGCCAGGTC AACAGAGCCGCCAGAATGTGTTTGGATGCTAGATAA			302
Sbjct 383	ATGCAAACTAAATATTGCCATATCAACTGATCCACCTGAATGCGCCTGTATACTTGCCAA			324
Query 303	GGGGGATAAACGGTTCATCCCGTCCCTGCCCTTGTCTACAAAGGCAGAACTTAATAA			362
Sbjct 323	AGGAGGATATACAGTTCATCCTGTGCCAGCCCCTTGTCTACAAACGCCGAACCTAATAA			264
Query 363	TAAAGTTAAAGAAGGAGGTAATAATCAGAACTAATATTATTTAATCTTGAAAAAGCCAT			422
Sbjct 263	TAAAGTTAAAGAAGGTGGTAATAATCAAAAACCTTATATTATTTAATCTTGAAAAAGCCAT			204
Query 423	ATCTGGCGCTCCGATATATAAAGGCACTAGTCAATTACCAAATCCTCCTATCATTACTGG			482
Sbjct 203	ATCCGGAGAACCGATATATAGCGGAACTAGTCAATTTCAAACCCCTATCATCACTGG			144
Query 483	CATTACTAGAAAAAATATCATTACAAAAGCATGAGCTGTTACTATAACATTGTATAAATG			542
Sbjct 143	CATAACTAAAAAATATCATTAAAAAAGCATGAGCAGTAACTATAACATTATATAAATG			84
Query 543	ATCATCCCCTAACATTGTTCTGGGGCTGATAATTCCTAATCTTATAAGCATGCTAAAGGC			602
Sbjct 83	ATCATCCCCAGCATTGAGCCCGGAGCTGATAATTCCTAACC GAATAAGCATACTAAACCC			24
Query 603	TGTTCTATCATCCCCGGAATG			625
Sbjct 23	AGTTCTATCATACCTGCAAATG			1

**Figura 52.** Alineamiento por blastn de la secuencia consenso de la esponja *Cliona celata*, proveniente de la estación Las Navajas (ID: 504).

## Anexo 2. Información de secuenciaciones 1 y 2

**Tabla 8.** Información relevante a la calidad de cada esfuerzo de secuenciación realizado en este trabajo: Número de secuenciación, número de muestras, sumatoria de secuencias asignadas a OTUs en cada secuencia, sumatoria de OTUs observados en cada secuencia, % de alineamiento para la secuencia de ADN sentido (R1), % de alineamiento para la secuencia de ADN antisentido (R2), % de error de R1 y R2.

Secuenciación	# de muestras	secuencias asignadas	OTUs_obs	% Alin_R1	% Alin_R2	% error (R1)	% error (R2)
1	18	1124842	23571	4.66 ± 0.17	4.58 ± 0.10	0.50 ± 0.14	0.50 ± 0.06
2	35	529982	19260	3.60 ± 0.08	3.52 ± 0.08	0.65 ± 0.04	0.56 ± 0.04

**Tabla 9.** Resultados de las secuenciaciones: SEQ1 y SEQ2 de las réplicas biológicas de esponja, sedimento y agua secuenciada, así como información relevante acerca de la profundidad de secuenciación.

ID Muestra	Secuenciación	# de secuencias asignadas a OTUs	# total de OTUs observados
SPG.APLY.CAN.3	SEQ1	47975	299
SPG.APLY.MAR.3	SEQ1	73922	365
SPG.APLY.ESP.3	SEQ1	74158	290
SPG.APLY.NAV.3	SEQ1	56903	364
SPG.CLIO.ESP.1.1	SEQ1	59126	1229
SPG.CLIO.ESP.2.1	SEQ1	52907	1144
SPG.CLIO.NAV.1.1	SEQ1	66351	1471
SPG.CLIO.NAV.3	SEQ1	53742	949
SPG.PROS.CAN.1.1	SEQ1	37849	634
SPG.PROS.MAR.1.1	SEQ1	79911	486
SD.CAN.3	SEQ1	64200	3643
SD.MAR.3	SEQ1	41144	2464
SD.ESP.3	SEQ1	39707	2731
SD.NAV.3	SEQ1	27635	2029
SW.CAN.3	SEQ1	104873	2198
SW.MAR.3	SEQ1	75058	812
SW.ESP.3	SEQ1	83859	1075
SW.NAV.3	SEQ1	85522	1388

SW.CAN.1	SEQ2	23748	991
SW.CAN.2	SEQ2	29698	918
SW.MAR.1	SEQ2	20172	434
SW.MAR.2	SEQ2	14605	410
SW.ESP.1	SEQ2	21301	497
SW.ESP.2	SEQ2	18422	456
SW.NAV.1	SEQ2	23325	628
SW.NAV.2	SEQ2	17791	620
SD.CAN.1	SEQ2	5561	1013
SD.CAN.2	SEQ2	5834	990
SD.MAR.1	SEQ2	11391	1170
SD.MAR.2	SEQ2	6361	782
SD.ESP.1	SEQ2	7831	1032
SD.ESP.2	SEQ2	13593	1440
SD.NAV.1	SEQ2	11153	1067
SD.NAV.2	SEQ2	16120	1090
SPG.PROS.CAN.1.2	SEQ2	16343	239
SPG.PROS.CAN.2	SEQ2	16171	198
SPG.PROS.CAN.3	SEQ2	19372	240
SPG.PROS.MAR.1.2	SEQ2	18419	267
SPG.PROS.MAR.2	SEQ2	16511	260
SPG.PROS.MAR.3	SEQ2	13869	187
SPG.APLY.CAN.1	SEQ2	14755	398
SPG.APLY.CAN.2	SEQ2	17572	189
SPG.APLY.MAR.1	SEQ2	9694	175
SPG.APLY.MAR.2	SEQ2	13599	171
SPG.APLY.ESP.1	SEQ2	12047	160
SPG.APLY.ESP.2	SEQ2	13570	184
SPG.APLY.NAV.1	SEQ2	18460	191
SPG.APLY.NAV.2	SEQ2	18452	169

SPG.CLIO.ESP.1.2	SEQ2	18062	677
SPG.CLIO.ESP.2.2	SEQ2	12298	544
SPG.CLIO.ESP.3	SEQ2	11446	494
SPG.CLIO.NAV.1.2	SEQ2	11487	618
SPG.CLIO.NAV.2	SEQ2	10949	361