

# UNIVERSIDAD DE GUADALAJARA

CENTRO UNIVERSITARIO DE CIENCIAS BIOLÓGICAS Y  
AGROPECUARIAS



## ANÁLISIS DE LA MICROBIOTA RELACIONADA CON EL CAMARÓN BLANCO (*Litopenaeus vannamei*): INTESTINO Y SEDIMENTO

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
**LICENCIADO EN BIOLOGÍA**

PRESENTA:  
**Luigui Michel Gallardo Becerra**

Las Agujas, Zapopan, Jalisco, Enero del 2018.

UNIVERSIDAD DE GUADALAJARA  
CENTRO UNIVERSITARIO DE CIENCIAS BIOLÓGICAS Y  
AGROPECUARIAS



ANÁLISIS DE LA MICROBIOTA RELACIONADA CON EL  
CAMARÓN BLANCO (*Litopenaeus vannamei*):  
INTESTINO Y SEDIMENTO

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
**LICENCIADO EN BIOLOGÍA**

PRESENTA:  
**Luogui Michel Gallardo Becerra**

DIRECTOR:  
Dr. Adrián Ochoa Leyva

ASESOR:  
Dr. Alejandro Muñoz Urias

Las Agujas, Zapopan, Jalisco, Enero del 2018.



UNIVERSIDAD DE GUADALAJARA  
Centro Universitario de Ciencias Biológicas y Agropecuarias  
Secretaría Académica  
Coordinación de la carrera de Biología

C. Luigui Michel Gallardo Becerra  
Presente

Manifestamos a usted que con esta fecha, ha sido aprobado su tema de titulación en la modalidad de Tesis, Tesina e Informes, opción Tesis con el trabajo intitulado: "ANÁLISIS DE LA MICROBIOTA RELACIONADA CON EL CAMARÓN BLANCO (*Litopenaeus vannamei*): INTESTINO Y SEDIMENTO", para obtener el título de licenciado en Biología.

Al mismo tiempo le informamos que ha sido aceptado como director de dicho trabajo al Dr. Adrián Ochoa Leyva, y como asesor al Dr. Alejandro Muñoz Urias.

Sin más por el momento, aprovechamos la ocasión para enviarle un cordial saludo.

A T E N T A M E N T E  
"PIENSA Y TRABAJA"

Las Agujas, Nextipac, Zapopan, Jal. 29 de septiembre de 2016

MTRA. CINTHYA ARACELI LÓPEZ LÓPEZ  
PRESIDENTE DEL COMITÉ DE TITULACIÓN

DRA. KARIN MARÍA EUGENIA BALZARETTI HEYM  
SECRETARIO DEL COMITÉ DE TITULACIÓN

**FORMATO 4**  
**Autorización impresión de tesis**

Las Agujas, Zapopan, Jalisco  
07 de noviembre de 2017

MTRA. CINTHYA ARACELI LÓPEZ LÓPEZ  
PRESIDENTE DEL COMITÉ DE TITULACIÓN  
LICENCIATURA EN BIOLOGÍA  
CENTRO UNIVERSITARIO DE CIENCIAS BIOLÓGICAS Y AGROPECUARIAS  
UNIVERSIDAD DE GUADALAJARA  
PRESENTE:

Por este conducto nos permitimos informar a Usted que habiendo revisado el trabajo de titulación "Análisis de la microbiota relacionada con el camarón blanco (*Litopenaeus vannamei*): intestino y sedimento" modalidad: **Tesis, Tesina e Informes**, opción específica: **Tesis**, del Reglamento General de Titulación de la Universidad de Guadalajara, que realizó el pasante: Luigui Michel Gallardo Becerra con código 208744963, consideramos que ha quedado debidamente concluido para su impresión.

Sin otro particular quedamos de usted con un cordial saludo.

Firma   
Nombre Dr. Adrián Ochoa Leyva  
Director/a del trabajo,

Firma Alejandro Henríquez Urias  
Nombre Alejandro Henríquez Urias  
Asesor(es) del trabajo

Grado y nombre completo del sinodal	Código	Firma	Fecha
Dra. Luz Patricia Castro Félix	8401494	Patricia Castro	9-11-2017
Dr. Leopoldo Díaz Pérez	2952002		9-11-2017
Dra. Josefina Casas Solís	9508775	Joséfa Casas S	9/Nov/17
Dr. Fernando Rodríguez Zúñiga Suplente	2739259	Fernando Rodríguez Zúñiga	13/Nov/17



---

Me gustaría agradecer a mis sinodales: Dr. Leopoldo Díaz Perez, Dra. Luz Patricia Castro Felix, Dra. Josefina Casas Solís y Dr. Fabián Alejandro Rodríguez Zaragoza por la corrección y evaluación de esta tesis; por su tiempo y compromiso muchas gracias.

Gracias al Dr. Alejandro Muñoz Urias que no solamente fue el asesor de esta tesis, sino que fue también uno de mis mejores profesores en la universidad y fue un ejemplo para seguir con mis estudios.

Dr. Adrián Ochoa Leyva que es el tutor de esta tesis y que me guía en mis estudios de posgrado en el Instituto de Biotecnología. Gracias por darme la oportunidad y confianza de realizar este proyecto y por enseñarme cada día que la investigación es una forma de vida apasionante donde puedes llegar tan lejos como te lo propongas. Gracias por todo.

Por último quiero agradecer a las personas que facilitaron la elaboración de esta tesis con su apoyo técnico. A nuestro técnico del laboratorio Filiberto Sánchez López por su enseñanza y apoyo en la parte de biología molecular de este proyecto.

De igual forma a Juan Manuel Hurtado Ramírez, Técnico Académico de la Unidad de Cómputo del Instituto de Biotecnología que administra y facilita el uso del servidor del laboratorio.

También me gustaría agradecer al proyecto PAPIIT-DGAPA UNAM (IA203516) 'Estudio de la dinámica del microbioma asociado al cultivo de dos cepas de camarón y su entorno a través del tiempo' que financió esta investigación.

---

Cualquier persona del nicho académico puede comprender en su totalidad esta tesis; sin embargo la parte personal, que va más allá de un título académico, más allá de los méritos escolares, del conocimiento en sí, es muy complicada de comprender. El haber acompañado a una persona en su búsqueda por superarse, por cumplir una de sus metas, por alcanzar sus sueños, eso es para mí muy difícil de poner en la piel de otro para que realmente lo entienda. Es por lo anterior que estas palabras valen la pena, para que esas personas sepan que la influencia que han dejado en ti han hecho que seas una mejor persona.

Quiero agradecer a mi familia. Cada uno de mis tíos y tías que son padres y madres para mí, que son parte de lo que soy ahora y que sin ellos simplemente yo no podría ser yo. A mi abuela que es la cabeza de nuestra gran familia. En especial quiero dedicar todo el trabajo y esfuerzo a mi madre por todo lo que ha hecho por mí. Desde las levantadas temprano para llevarme a la escuela desde que tengo uso de memoria, hasta ese último día que me llevó a la parada de autobús hacia la Universidad. Apoyándome en mis momentos difíciles, de indecisión acerca de mi futuro, en mis rebeldías y en mi falta de entusiasmo. Ahora visto desde lejos puedo observar que todo lo que hiciste por mí tuvo una razón de ser importante. Y como alguna vez le dije cuando había terminado el kinder: ¡lo logramos mamá!.

Me gustaría agradecer también a cada una de las personas que estuvieron conmigo en este viaje de formación, aprendizaje y porque no decirlo, un poco de locura y autodescubrimiento. Adrián, Daniel y Nigel, mis amigos de la preparatoria que con nuestras largas pláticas y discusiones formamos nuestro criterio desde los 15 años. A mis profesores que me cambiaron la forma de ver a la Biología, de su tiempo, enseñanza, profesionalismo y exigencia que formaron en mi hábitos para ser un buen profesionista. A cada uno de mis amigos de la Licenciatura en Biología: Christian, Miroslava, Alejandra, Polo, Paulina, Elizabeth, Julieta, Naye, Cuco, Ana y Sebastián. A los que me acompañaron en la parte final de mi carrera, en el momento que tuve que tomar las decisiones que repercuten fuertemente hoy en mi vida: Rosy, Pollo, Tachi y Victor.

Muchas gracias al Dr. Adrián Ochoa Leyva por aceptarme en el laboratorio del cual formo ahora parte en IBT; por su enseñanza, motivación y confianza que cada día me hace ser una persona más completa y apasionada por la ciencia. A mis colegas y amigos del laboratorio Andrea, Isabel, Abigail, Edgar, Omar, Miriam, Fernanda, Shirley, Sandra, Alfredo, Melani, Juan Pablo, Axel y Fili, que me hacen sentir todos los días en casa y que me apoyan en este nuevo camino en mi vida.

Gracias a todas esas personas que me dieron una oportunidad, que es lo único que se necesita.

---

*“... en la Tierra no vivimos en la Era del Hombre o de los humanos, vivimos hoy, y siempre, en la Era de la Bacterias.”*

*Stephen Jay Gould*



## Resumen

---

En este trabajo se busca conocer el impacto de la genética y el ambiente en la microbiota intestinal del camarón blanco del Pacífico (*Litopenaeus vannamei*) que es de gran importancia comercial a nivel mundial. Los especímenes considerados para la investigación provienen de tres estanques de una granja acuícola del norte del estado de Sinaloa, dos de los cuales poseen una genética similar por provenir de la misma fuente de desove. Para conocer los microorganismos presentes se utilizó la secuenciación del gen 16S ARNr en sus regiones hipervariables V3-V4 con la tecnología de MiSeq de Illumina. El análisis bioinformático se realizó en gran parte con el software de uso libre QIIME y la base de datos Greengenes. Nuestros resultados muestran el efecto de la genética y el ambiente en la microbiota intestinal de *L. vannamei*.

# Índice general

---

<b>Índice de figuras</b>	<b>vii</b>
<b>Índice de tablas</b>	<b>xvi</b>
<b>Introducción</b>	<b>1</b>
<b>Marco teórico</b>	<b>3</b>
Camarón blanco del Pacífico . . . . .	3
Cultivo de camarón . . . . .	4
Microbiota . . . . .	7
Herramientas para el análisis del gen 16S ARNr . . . . .	7
Características del gen 16S ARNr . . . . .	7
Regiones hipervariables . . . . .	8
Secuenciación por síntesis de Illumina® . . . . .	9
Control de calidad (QC) . . . . .	10
Estrategias de selección de OTU . . . . .	10
Análisis de Diversidad . . . . .	11
Cultivo de camarón . . . . .	14
<b>Antecedentes</b>	<b>15</b>
<b>Hipótesis</b>	<b>18</b>
<b>Objetivos</b>	<b>19</b>
Objetivo general . . . . .	19
Objetivos particulares . . . . .	19
<b>Materiales y métodos</b>	<b>20</b>
Colecta de las muestras . . . . .	21
Extracción del ADN con miniprep . . . . .	22
Amplificación de la región V3-V4 del gen 16S ARNr . . . . .	24

---

Programación del termociclador . . . . .	24
PCR <i>clean up</i> . . . . .	25
Secuenciación masiva . . . . .	26
Preprocesamiento de los datos . . . . .	26
Asignación taxonómica con QIIME . . . . .	27
Comparación entre bases de datos . . . . .	27
Asignación de taxonomía en muestras de intestino con Greengenes	28
Análisis LEfSe . . . . .	29
Análisis de diversidad $\alpha$ . . . . .	29
Análisis de diversidad $\beta$ . . . . .	30
Árbol filogenético . . . . .	30
PCA con abundancias de OTU . . . . .	31
Mapas de calor por ambiente . . . . .	31
<b>Resultados</b> . . . . .	<b>33</b>
Extracción de ADN y obtención de amplicones . . . . .	33
Pre-procesamiento . . . . .	36
FastQC . . . . .	36
Corte de los extremos de las lecturas con Trimmomatic . . . . .	39
Unión de lecturas R1 y R2 . . . . .	40
Limpieza por calidad en ventanas con Trimmomatic . . . . .	40
Resultados del pre-procesamiento . . . . .	42
Comparación de las bases de datos para secuencias de 16S ARNr . . . . .	44
Asignación de taxonomía en muestras de intestino con Greengenes . . . . .	45
Comparación a nivel phylum . . . . .	45
Comparación a nivel clase . . . . .	48
Comparación a nivel orden . . . . .	52
Comparación a nivel familia . . . . .	55
Comparación a nivel género . . . . .	58
Comparación a nivel especie . . . . .	61
Análisis LEfSe . . . . .	65
Pruebas de LEfSe por genética . . . . .	65
Pruebas LEfSe por ambiente . . . . .	73
Análisis de Diversidad $\alpha$ . . . . .	78
Comparación entre estanques . . . . .	78
Comparación entre ambientes . . . . .	82
Comparación entre genéticas . . . . .	84
Análisis de Diversidad $\beta$ . . . . .	87
Árbol filogenético . . . . .	90
Análisis de Componentes Principales . . . . .	91
Agrupamiento jerárquico con mapas de calor . . . . .	96

---

---

Mapas de calor por ambiente . . . . .	96
Mapas de calor por genética . . . . .	103
<b>Discusión</b>	<b>109</b>
<b>Conclusión</b>	<b>112</b>
<b>A. Apéndices</b>	<b>114</b>
A.1. Cebadores utilizados . . . . .	114
A.2. Código . . . . .	114
A.3. Figuras suplementarias . . . . .	119
Referencias . . . . .	129

# Índice de figuras

---

1.	Imagen que muestra el sistema digestivo típico de un camarón; tomada y modificada de . . . . .	4
2.	Esquema del cultivo de camarón; existen dos etapas importantes para la colonización microbiana:la muda de Nauplio a Zoea y la etapa de crecimiento de Post-larva a Sub-adulto que es en la cual generalmente se cosecha al camarón; tomado de (3). . . . .	5
3.	Producción mundial de camarón blanco del Pacífico ( <i>L. vannamei</i> ); tomada de (3). . . . .	6
4.	Producción de camarón blanco del Pacífico ( <i>L. vannamei</i> ); tomada y modificada de (4). . . . .	6
5.	Imagen del ARNr 16S, con sus regiones hipevariables (V1-V9) así como sus respectivos tamaños; tomada y modificada de (17). . . . .	8
6.	Esquema que muestra las regiones hipervariables del 16S ARNr (V1:V9); tomada y modificada de . . . . .	9
7.	Diagrama que muestra la metodología dividida en seis bloques principales. . . . .	20
8.	Mapa con los estanques y sitios de muestreo; Estanque Estrella (1), Estanque Mazatlán (2) y Estanque Lajitas (3). . . . .	21
9.	Gel de agarosa al 2 % con buffer TBE 1X, donde se observan los amplicones de ≈550 pb en las muestras de intestino; foto a los 10 min con 100 V. . . . .	35
10.	Gel de agarosa al 2 % con buffer TBE 1X, donde se observan los amplicones después de la limpieza con AMPure XP; foto a los 10 min con 100 V; las muestras con X, no son muestras de intestino. . . . .	35
11.	Imagen obtenida con FastQC de la muestra R1-IA9 antes del pre-procesamiento; muestra en contenido de N (nucleotidos no asignados). . . . .	36
12.	Imagen obtenida con FastQC de la muestra R2-IA9 antes del pre-procesamiento; muestra en contenido de N (nucleotidos no asignados)	37

---

13.	Imagen obtenida con FastQC de la muestra R1-IA9 antes del preprocesamiento; muestra la calidad por base basado en el <i>Phred Quality Score</i> . . . . .	37
14.	Imagen obtenida con FastQC de la muestra R2-IA9 antes del preprocesamiento; muestra la calidad por base basado en el <i>Phred Quality Score</i> . . . . .	38
15.	Imagen obtenida con FastQC de la muestra R1-IA9 después del corte de los primeros 5 nucleótidos. . . . .	39
16.	Imagen obtenida con FastQC de la muestra R2-IA9 después del corte de los primeros 6 nucleótidos. . . . .	39
17.	Gráfica de los resultados de la unión de R1 y R2 de la muestra IA9 a diferentes cortes (de 230 a 245); el corte con mayor número de lecturas juntas en esta muestra fue de 232 con 60354 secuencias totales. . . . .	40
18.	Imagen obtenida con FastQC de la muestra IA9 después de juntar R1 y R2; muestra la calidad por base basado en el <i>Phred Quality Score</i> ; se observa que hay calidades por debajo del Q20 con una mayor probabilidad de error al asignar una base. . . . .	41
19.	Imagen obtenida con FastQC después de la limpieza por ventanas; muestra la calidad por base basado en el <i>Phred Quality Score</i> ; se observa que no hay bases con calidad menor a Q20. . . . .	41
20.	Diagrama de Venn a nivel de género entre las tres bases de datos comparadas; (GG: Greengenes; RDP: Ribosomal Database Project; y SILVA). . . . .	44
21.	Gráfica que muestra los phyla con una abundancia mayor al 0.1 %. . . . .	46
22.	Diagrama de Venn que muestra los phyla compartidos entre los estanques Lajitas, Mazatlán y Estrella. . . . .	47
23.	Diagrama de Venn que muestra los phyla compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella). . . . .	48
24.	Diagrama de Venn que muestra los phyla compartidos entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas). . . . .	48
25.	Gráfica que muestra las clases con una abundancia mayor al 0.1 %. . . . .	49
26.	Diagrama de Venn que muestra las clases compartidas entre los estanques Lajitas, Mazatlán y Estrella. . . . .	51
27.	Diagrama de Venn que muestra las clases compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella). . . . .	51
28.	Diagrama de Venn que muestra las clases compartidas entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas). . . . .	52
29.	Gráfica que muestra los órdenes con una abundancia mayor al 0.1 %. . . . .	53
30.	Diagrama de Venn que muestra los órdenes compartidos entre los estanques Lajitas, Mazatlán y Estrella. . . . .	54

---

---

31.	Diagrama de Venn que muestra los órdenes compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).	54
32.	Diagrama de Venn que muestra los órdenes compartidos entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).	55
33.	Gráfica que muestra las familias con una abundancia mayor al 0.1 %.	56
34.	Diagrama de Venn que muestra las familias compartidas entre los estanques Lajitas, Mazatlán y Estrella. . . . .	57
35.	Diagrama de Venn que muestra las familias compartidas entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).	57
36.	Diagrama de Venn que muestra las familias compartidas entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).	58
37.	Gráfica que muestra los géneros con una abundancia mayor al 0.1 %.	59
38.	Diagrama de Venn que muestra los géneros compartidos entre los estanques Lajitas, Mazatlán y Estrella. . . . .	60
39.	Diagrama de Venn que muestra los géneros compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).	60
40.	Diagrama de Venn que muestra los géneros compartidos entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).	61
41.	Gráfica que muestra las especies con una abundancia mayor al 0.1 %.	62
42.	Diagrama de Venn que muestra las especies compartidas entre los estanques Lajitas, Mazatlán y Estrella. . . . .	63
43.	Diagrama de Venn que muestra las especies compartidas entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).	64
44.	Diagrama de Venn que muestra las especies compartidas entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).	64
45.	Análisis con LEfSe a nivel de phylum, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran los phyla que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde). . . . .	66
46.	Análisis con LEfSe a nivel de clase, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran las clases que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde). . . . .	66
47.	Análisis con LEfSe a nivel de orden, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran los órdenes que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde). . . . .	68

- 
48. Análisis con LEfSe a nivel de familia, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran las familias que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde). . . . . 69
49. Análisis con LEfSe a nivel de género, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran los géneros que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde). . . . . 70
50. Análisis con LEfSe a nivel de especie, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran las especies que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde). . . . . 71
51. Análisis con LEfSe a nivel de phylum, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran los phyla que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo). . . . . 73
52. Análisis con LEfSe a nivel de clase, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran las clases que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo). . . . . 74
53. Análisis con LEfSe a nivel de orden, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran los órdenes que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo). . . . . 75
54. Análisis con LEfSe a nivel de familia, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran las familias que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo). . . . . 76
55. Análisis con LEfSe a nivel de género, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran los géneros que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo). . . . . 77

---

56.	Análisis con LEfSe a nivel de especie, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran las especies que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo). . . . .	78
57.	Curva de rarefacción con el índice de Shannon, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y estanque: sedimento estanque Lajitas (morado), sedimento estanque Estrella (verde), sedimento estanque Mazatlán (amarillo), intestino estanque Lajitas (azul), intestino estanque Estrella (rojo) e intestino estanque Mazatlán (naranja). . . . .	79
58.	Curva de rarefacción con el índice de PD, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y estanque: sedimento estanque Lajitas (morado), sedimento estanque Estrella (verde), sedimento estanque Mazatlán (amarillo), intestino estanque Lajitas (azul), intestino estanque Estrella (rojo) e intestino estanque Mazatlán (naranja). . . . .	80
59.	Gráfica de caja y bigotes de la comparación de la diversidad $\alpha$ con el índice de Shannon entre los intestinos y sedimentos de los tres estanques. . . . .	80
60.	Curva de rarefacción con el índice Chao1, que estima la riqueza de especies; se encuentran agrupadas por tipo de muestra y estanque: sedimento estanque Lajitas (morado), sedimento estanque Estrella (verde), sedimento estanque Mazatlán (amarillo), intestino estanque Lajitas (azul), intestino estanque Estrella (rojo) e intestino estanque Mazatlán (naranja). . . . .	81
61.	Gráfica de caja y bigotes de la comparación de la diversidad $\alpha$ con el índice de Chao1 entre los intestinos y sedimentos de los tres estanques. . . . .	81
62.	Curva de rarefacción con el índice de Shannon, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y ambiente: sedimento ambiente 1 (amarillo), sedimento ambiente 2 (verde), intestino ambiente 1 (rojo) e intestino ambiente 2 (azul). . . . .	82
63.	Curva de rarefacción con el índice de PD, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y ambiente: sedimento ambiente 1 (amarillo), sedimento ambiente 2 (verde), intestino ambiente 1 (rojo) e intestino ambiente 2 (azul). . . . .	83
64.	Curva de rarefacción con el índice de Chao1, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y ambiente: sedimento ambiente 1 (amarillo), sedimento ambiente 2 (verde), intestino ambiente 1 (rojo) e intestino ambiente 2 (azul). . . . .	83

---

65.	Curva de rarefacción con el índice de Shannon, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y genética: sedimento genética 1 (amarillo), sedimento genética 2 (verde), intestino genética 1 (rojo) e intestino genética 2 (azul). . . . .	84
66.	Gráfica de caja y bigotes de la comparación de la diversidad $\alpha$ con el índice de Shannon entre los intestinos y sedimentos separados por genética. . . . .	85
67.	Curva de rarefacción con el índice Chao1, que estima la riqueza de especies; se encuentran agrupadas por tipo de muestra y genética: sedimento genética 1 (amarillo), sedimento genética 2 (verde), intestino genética 1 (rojo) e intestino genética 2 (azul). . . . .	85
68.	Curva de rarefacción con el índice PD, que estima la riqueza de especies; se encuentran agrupadas por tipo de muestra y genética: sedimento genética 1 (amarillo), sedimento genética 2 (verde), intestino genética 1 (rojo) e intestino genética 2 (azul). . . . .	86
69.	Gráfica de caja y bigotes de la comparación de la diversidad $\alpha$ con el índice de Chao1 entre los intestinos y sedimentos separados por genética . . . . .	86
70.	PCoA ponderado ( <i>weighted</i> con las muestras de intestino de camarón; izquierda: A, C y E) y no ponderado ( <i>unweighted</i> ; derecha: B, D y F) con los PC que explican el mayor porcentaje de variación. Muestras de intestino de camarón; las muestras están agrupadas por ambiente (A y B), genética (C y D) y estanque (E y F). . . . .	88
71.	PCoA ponderado ( <i>weighted</i> ) con muestras de intestino de camarón y sedimento; izquierda: A, C y E) y no ponderado ( <i>unweighted</i> ; derecha: B, D y F) con los PC que explican el mayor porcentaje de variación. Muestras de intestino de camarón y de sedimento; las muestras están agrupadas por ambiente (A y B), genética (C y D) y estanque (E y F). . . . .	89
72.	Árbol generado con Jackknife ponderado ( <i>weighted</i> ); cada nodo cuenta con su valor confianza. . . . .	90
73.	PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de phylum. . . . .	91
74.	PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de clase. . . . .	92
75.	PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de orden. . . . .	93

---

76. PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de familia. . . . .	94
77. PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de género. . . . .	95
78. PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de especie. . . . .	96
79. Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de phylum; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	97
80. Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de clase; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	98
81. Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de orden; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	99
82. Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de familia; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	100
83. Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de género; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	101
84. Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de especie; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	102
85. Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de phylum; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	103
86. Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de clase; se consideran solamente los OTU con una abundancia mayor al 0.1 %. . . . .	104

---

87. Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de orden; se consideran solamente los OTU con una abundancia mayor al 0.1 %.	105
88. Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de familia; se consideran solamente los OTU con una abundancia mayor al 0.1 %.	106
89. Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de género; se consideran solamente los OTU con una abundancia mayor al 0.1 %.	107
90. Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de especie; se consideran solamente los OTU con una abundancia mayor al 0.1 %.	108
A.1. Cebador <i>forward</i> utilizado para seleccionar la región V3-V4 del gen 16S ribosomal.	114
A.2. Cebador <i>reverse</i> utilizado para seleccionar la región V3-V4 del gen 16S ribosomal.	114
A.3. PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de phylum.	119
A.4. PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de clase.	119
A.5. PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de orden.	120
A.6. PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de familia.	120
A.7. PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de género.	121
A.8. PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de especie.	121
A.9. Árbol generado con Jackknife no ponderado ( <i>unweighted</i> ); cada nodo cuenta con su valor confianza.	122
A.10. Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel phylum.	123
A.11. Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel clase.	123
A.12. Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel orden.	124

---

A.13.Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel familia. . . . .	124
A.14.Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel género. . . . .	125
A.15.Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel especie. . . . .	125
A.16.Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel phylum. . . . .	126
A.17.Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel clase. . . . .	126
A.18.Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel orden. . . . .	127
A.19.Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel familia. . . . .	127
A.20.Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel género. . . . .	128
A.21.Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel especie. . . . .	128

# Índice de tablas

---

1.	Clasificación taxonómica de <i>L. vannamei</i> . . . . .	3
2.	Valores de Phred (Q); el valor ideal de calidad en secuenciación es de Q30; tomada y modificada de . . . . .	10
3.	Tabla con las muestras de intestino y sedimento para cada uno de los estanques; para cada estanque (Lajitas, Mazatlán y Estrella) se consideraron tres puntos de colecta y en cada uno de éstos tres muestras de intestino y una muestra de sedimento. . . . .	22
4.	Concentraciones de ADN en ng. . . . .	33
5.	Tabla con los volúmenes utilizados para la reacción de PCR. . . .	34
6.	Resultados del pre-procesamiento de muestras de intestino de camarón en porcentaje; se muestran el porcentaje resultantes en cada paso. . . . .	42
7.	Resultados del pre-procesamiento de muestras de intestino de camarón; se muestran el número de lecturas resultantes en cada paso. . . . .	43
8.	Comparación entre las asignaciones de QIIME con los valores por defecto para las tres bases de datos (GG: Greengenes; RDP: Ribosomal Database Project; y SILVA) . . . . .	44
9.	Comparación entre los índices de diversidad; se utilizó un tamaño de 150,110 secuencias para la comparación (GG: Greengenes; RDP: Ribosomal Database Project; y SILVA). . . . .	45
10.	Tabla con los 10 OTU más abundantes a nivel de phylum para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. . . . .	46
11.	Tabla con los 10 OTU más abundantes a nivel de clase para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. . . . .	49
12.	Tabla con los 10 OTU más abundantes a nivel de orden para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. . . . .	52

---

13.	Tabla con los 10 OTU más abundantes a nivel de familia para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. Los OTU que no obtuvieron una clasificación completa se colocan con '??' (orden: o? y familia: f?). . . . .	56
14.	Tabla con los 10 OTU más abundantes a nivel de género para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. Los OTU que no obtuvieron una clasificación completa se colocan con '??' (familia: f? y género: g?). . . . .	59
15.	Tabla con los 10 OTU más abundantes a nivel de especie para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. Los OTU que no obtuvieron una clasificación completa se colocan con '??' (familia: f?, género: g? y especie: s?). . . . .	61
16.	Tabla que muestra las diferencias encontradas con LEfSe en los distintos niveles taxonómicos, para las dos comparaciones: con todos los OTU y utilizando solamente los OTU con una abundancia mayor al 0.1%. . . . .	65

# Introducción

---

El camarón blanco del Pacífico (*Litopenaeus vannamei*) es una especie que tiene un gran potencial comercial (1), equivalente a más del 70 % de la producción mundial de camarón (2) con más de 3,5 millones de toneladas anuales (3). Es la especie de camarón mas producida en México con cerca de 120,000 toneladas anuales (4). Sin embargo existen dos eventos de gran relevancia que impactaron fuertemente la producción en nuestro país, en 2010 llega el virus de mancha blanca a Sonora, causando pérdidas de hasta el 50 % de la producción estatal. Posteriormente en 2013 *Vibrio parahaemolyticus* provoca un brote de EMS (Síndrome de Mortalidad Temprana, del inglés *Early Mortality Syndrome*) en todo México causando pérdidas del 60-90 %. Es por esto que investigaciones recientes buscan la relación de la microbiota del camarón con dicha propensión a dichas enfermedades (5).

La microbiota intestinal tiene un papel importante en la digestión, nutrición, el desarrollo del sistema inmune y en la resistencia a colonización de patógenos (1)(6)(5). Además de que se ha demostrado que promueve la salud en el organismo donde habita, la microbiota intestinal puede ser asociada a enfermedades del hospedero (7)(8)(9)(5). Una disbiosis, que es un desequilibrio en la composición bacteriana normal, puede comprometer la salud del hospedero ya que lo vuelve susceptible a la invasión de patógenos oportunistas que pueden llevar al crecimiento ineficiente e incluso la muerte del organismo lo cual tiene un impacto directo en su producción (9)(8). Por esta razón actualmente existe un interés especial en estudiar las relaciones, benéficas o patógenas, entre microorganismos (bacterias, arqueas y eucariotas unicelulares) y macroorganismos, en especial en los de importancia médica o comercial (7).

Debido a que el 99 % de los microorganismos de una muestra de microbiota no pueden ser cultivados en el laboratorio (11), para determinar la diversidad microbiana se utilizan técnicas basadas en el análisis múltiple de secuencias de ADN, esto es: el análisis metagenómico o perfil de secuencias de ADN marcadores. Uno de los ideales para estudiar la diversidad microbiana es el análisis del perfil del gen 16S ARNr (6); esto acompañado con análisis bionformático permite conocer

## INTRODUCCIÓN

---

cuáles son las especies presentes en una muestra determinada, en este caso la microbiota intestinal del camarón.

A pesar de su relevancia en la acuicultura mundial, existen pocas investigaciones para conocer el impacto de la genética y el ambiente en la diversidad microbiana intestinal de *L. vannamei*. Nuestros resultados permiten conocer cuál es la microbiota en muestras de intestino de *L. vannamei* de tres estanques diferentes en la misma granja camaronera, dos de los cuales comparten el mismo linaje por lo que cuentan con una genética similar, de igual forma nos permiten identificar las diferencias entre estanque (lo que definimos como ambiente). Esto podría brindarnos un punto de vista más amplio de cómo es la estructura de las comunidades microbianas en el intestino de esta especie de camarón, que podría servir como fuente de información para desarrollo de prebióticos y probióticos, así como para la prevención de enfermedades causadas por patógenos bacterianos.

# Marco teórico

---

## El camarón blanco del Pacífico

El organismo en el cual se enfoca nuestro interés es *Litopenaeus vannamei* (tabla 1) se encuentra dentro de la familia Penaeidae que incluye a otras especies de importancia comercial como el camarón blanco del Atlántico (*L. setiferus*) (3) y el camarón tigre (*Penaeus monodon*)(8).

**Tabla 1:** Clasificación taxonómica de *L. vannamei*.

<b>Phylum</b>	Arthropoda
<b>Clase</b>	Malacostraca
<b>Orden</b>	Decapoda
<b>Familia</b>	Penaeidae
<b>Género</b>	<i>Litopenaeus</i>
<b>Especie</b>	<i>Litopenaeus vannamei</i> (Boone, 1931)

## Sistema digestivo

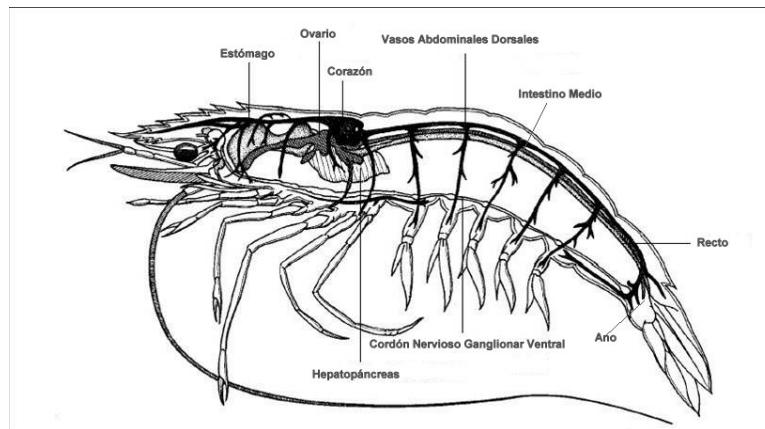
Las funciones del sistema digestivo son: ingestión de alimentos, transporte de la ingesta, digestión, absorción y transporte de nutrientes, así como finalmente la excreción de desechos. Se puede dividir en cinco partes principales: boca, esófago, estómago, hepatopáncreas e intestino (figura 1). Éste último a su vez se puede subdividir en anterior, medio y posterior (Ceccaldi, 1997).

## Ciclo de vida

El ciclo de vida del camarón (figura 2) inicia en la etapa de desove después de la fecundación externa, que se realiza en aguas profundas (fase marina). Pos-

## MARCO TEÓRICO

---



**Figura 1:** Imagen que muestra el sistema digestivo típico de un camarón; tomada y modificada de .

teriormente le sigue la fase incubación en la cual los huevos maduran y pasan a diferentes estadíos larvales (Nauplio, Zoea y Mysis) hasta llegar a la etapa de Post-larva que se asemeja a un camarón adulto pero de menor tamaño y sin madurez sexual. Durante el crecimiento de la etapa de Post-larva el camarón se encuentra en aguas poco profundas (fase estuarina), y en un lapso de 3 a 4 meses llega a la etapa de Sub-adulto.

Para nuestra investigación de microbiota intestinal son de importancia dos etapas principalmente. La primera es la muda de Nauplio a Zoea (ver figura 2) en la cual se da la colonización microbiana en el interior del camarón (12). La segunda es la etapa de crecimiento de Post-larva a Sub-adulto que es en la cual generalmente se cosecha al camarón (3) y es la etapa de los especímenes de nuestro estudio.

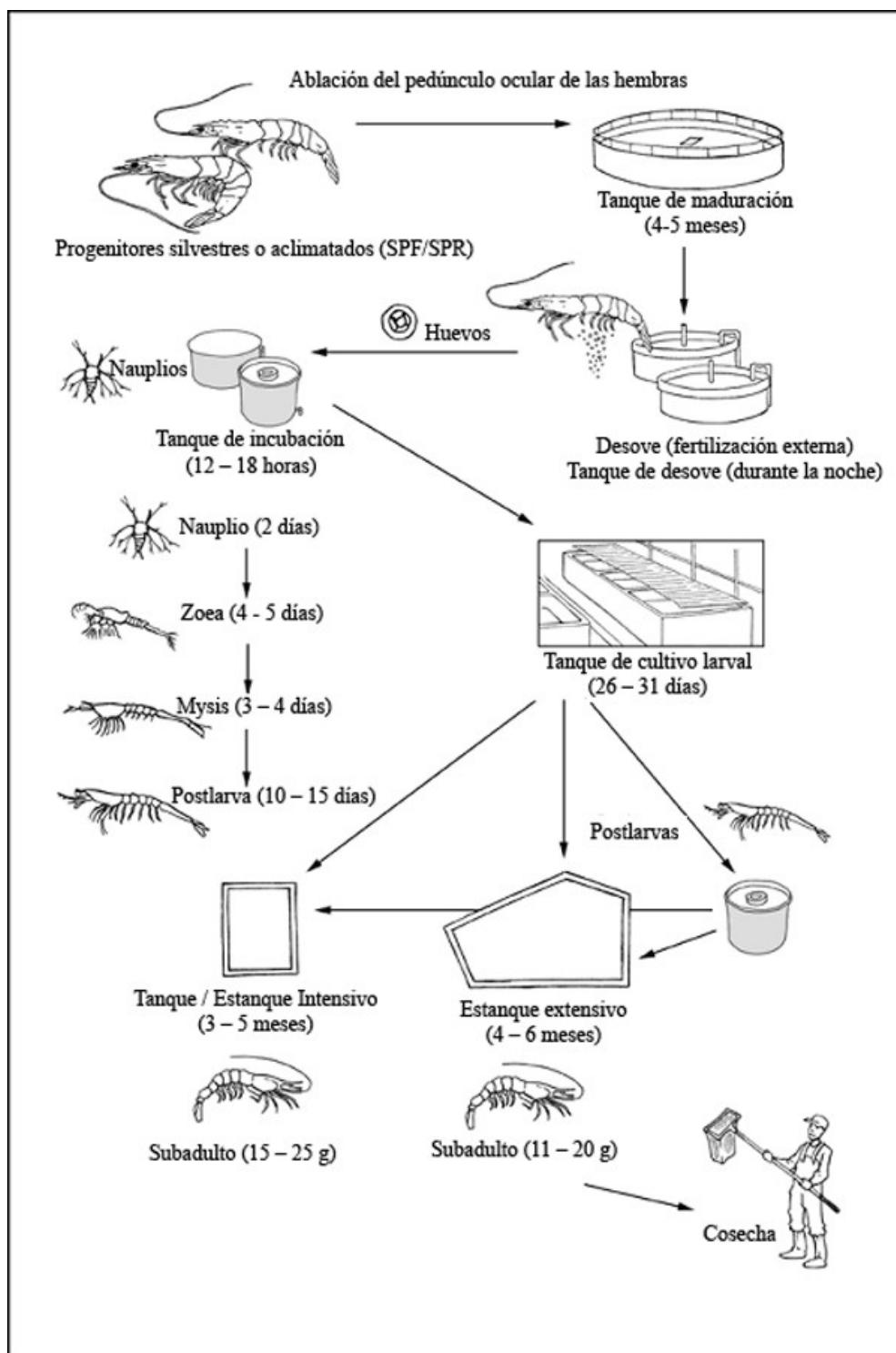
Mientras el camarón se encuentra en crecimiento se pueden presentar diversas enfermedades causadas por microorganismos como: camarón manchado (*Brown spot disease*), vibriosis sistémica, necrosis hepatopancreática aguda (AHPND o EMS), entre otras.

## Cultivo de camarón

Esta especie de camarón tiene un gran potencial comercial (1) y es equivalente a más del 70 % de la producción mundial de camarón (2) con más de 3,5 millones de toneladas anuales (figura 3)(3).

## MARCO TEÓRICO

---



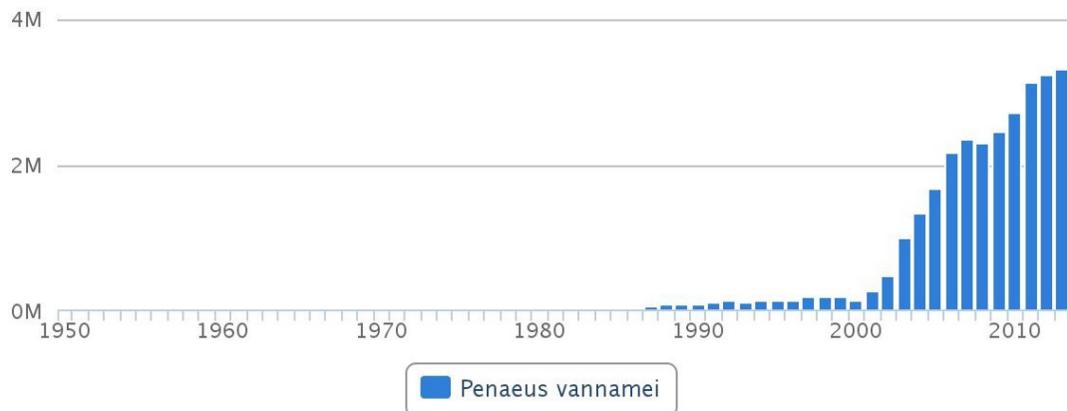
**Figura 2:** Esquema del cultivo de camarón; existen dos etapas importantes para la colonización microbiana: la muda de Nauplio a Zoea y la etapa de crecimiento de Post-larva a Sub-adulto que es en la cual generalmente se cosecha al camarón; tomado de (3).

## MARCO TEÓRICO

---

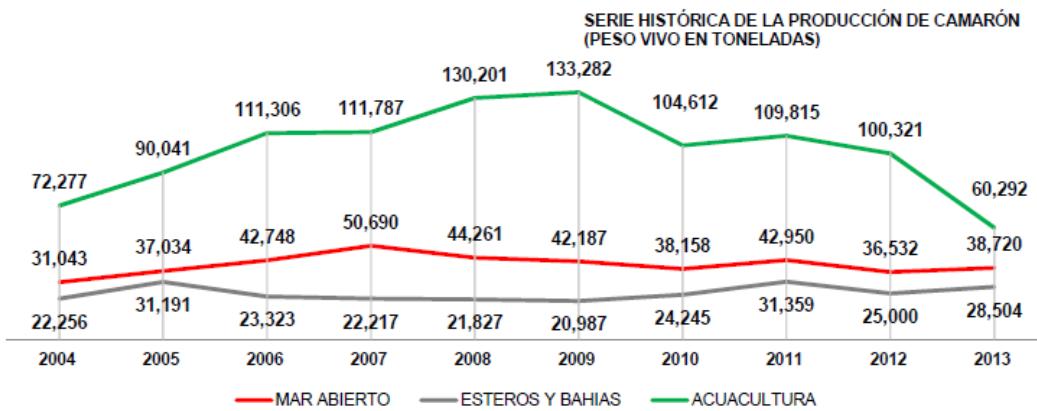
### Producción acuícola mundial (toneladas)

Fuente: FAO FishStat



**Figura 3:** Producción mundial de camarón blanco del Pacífico (*L. vannamei*); tomada de (3).

Entre los estados de la república con mayor producción de camarón se encuentran Sonora y Sinaloa. En la figura 4 se muestra la producción de camarón blanco en Sinaloa hasta el año 2013.



**Figura 4:** Producción de camarón blanco del Pacífico (*L. vannamei*); tomada y modificada de (4).

## Microbiota

Microbiota, que es sinónimo de comunidad microbiana, puede ser definida como el conjunto de microorganismos que existen en un mismo espacio y tiempo en una muestra determinada (ej.: suelo, agua, fluidos corporales, órganos, etc.) (6)(13)(11). Incluye a las bacterias, arqueas, hongos unicelulares, protozoarios, así como sus virus (6). Debido a su complejidad se puede delimitar el campo de estudio a uno de estos grupos en particular. Por ejemplo cuando se realiza un análisis del gen 16S ARNr se considera solamente a los procariotes (arqueas y bacterias) presentes en la muestra, ya que este gen es exclusivo de dichos microorganismos (14).

## Herramientas para el análisis del gen 16S ARNr

Para poder estudiar a gran escala la microbiota presente en una muestra se utilizan herramientas bioinformáticas basadas en la metagenómica, en donde se incluyen a todos los genomas de una muestra (13) o bien el uso de los genes como marcadores moleculares para la clasificación de los seres vivos, como las secuencias de ADN de la subunidad ribosomal 16S (14).

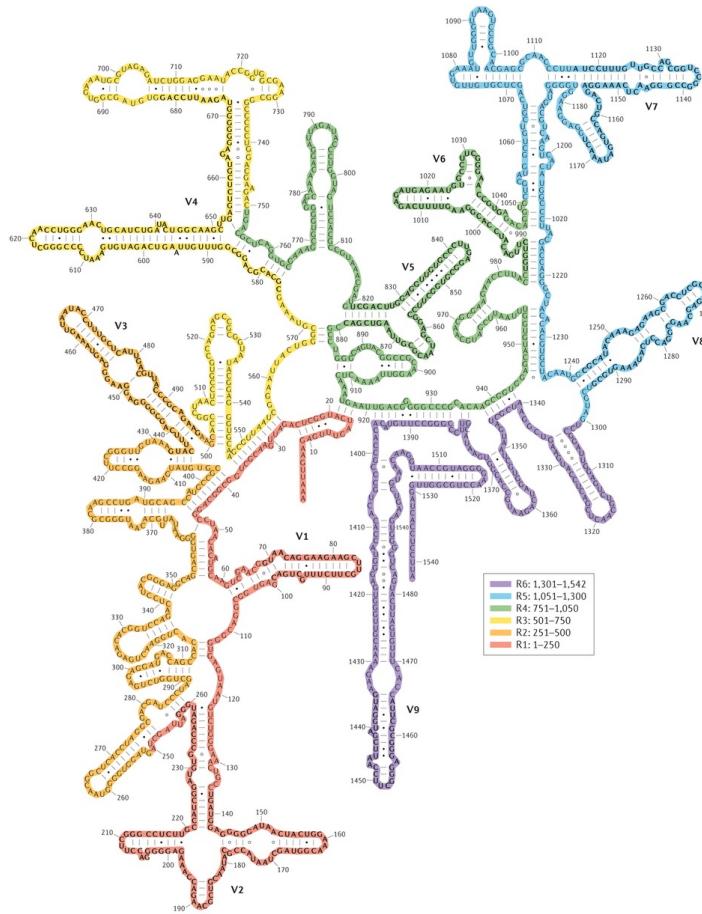
Debido a la disminución de los costos y del tiempo de procesamiento de las secuencias cortas de ADN en la última década muchos estudios para caracterizar comunidades microbianas se basan en la secuenciación del gen ARNr 16S que nos permite identificar la abundancia y diversidad de bacterias y arqueas presentes en una muestra sin necesidad de secuenciar sus genomas completos (15)(16). Ésto es conocido como análisis del perfil del gen 16S ARNr (6).

## Características del gen 16S ARNr

La maquinaria ribosomal en procariotas está compuesta por dos subunidades: la subunidad mayor que se conforma de los ARNr 23S y 5S junto con 36 proteínas, y a la subunidad menor que incluye al ARNr 16S (figura 5) junto con 24 proteínas.

## MARCO TEÓRICO

---



**Figura 5:** Imagen del ARNr 16S, con sus regiones hipevariables (V1-V9) así como sus respectivos tamaños; tomada y modificada de (17).

Existen varios motivos por los cuales se ha seleccionado al gen 16S ARNr como marcador. El primero es que es un gen universal para todos los organismos procariotes; esto permite la identificación, comparación y evaluación de todos los organismos en una muestra utilizando solamente este gen. El segundo es que el gen tiene evolución constante relativamente lenta a través del tiempo lo cual nos permite separar entre grandes grupos de organismos, por ejemplo separar a las arqueas de las bacterias, así como diferenciar entre los phyla de los reinos anteriores (11)(14)(17).

## Regiones hipervariables

El gen 16S ARNr está conformado por nueve regiones hipervariables distribuidas a lo largo de su secuencia, delimitadas cada una de ellas por regiones

## MARCO TEÓRICO

---

conservadas y tiene una longitud aproximada de 1500 pares de bases (figuras 5 y 6). Como se menciona anteriormente estas regiones hipervariables nos permiten diferenciar entre organismos de diferentes niveles taxonómicos (phylum, clase, orden, familia, género y especie); organismos de un mismo grupo taxonómico tendrán una menor variabilidad y esto nos permite agruparlos y determinarlos en sus diferentes niveles.

Para los estudios de diversidad basados en el 16S ARNr se amplifican una o dos regiones hipervariables, que nos facilitan el distinguir entre un organismo de otro y así poder asignarles taxonomía (6). Cuál región del 16S ARNr utilizar es un área de debate. La región que nos brinda más información depende del tipo de muestra, el diseño experimental y el objetivos (6). En investigaciones que buscan determinar la microbiota intestinal humana, como el consorcio *MetaHit*, se utiliza la región hipervariable V4, que además es el mismo propuesto por el *Earth Microbiome Project*. En el protocolo de MiSeq de Illumina® se incluye la preparación de las regiones V3 y V4, que son de las mejores determinando la diversidad de la muestras (18).



**Figura 6:** Esquema que muestra las regiones hipervariables del 16S ARNr (V1:V9); tomada y modificada de .

### Secuenciación por síntesis de Illumina®

La secuenciación del gen ribosomal 16S con tecnologías de secuenciación masiva nos permiten explorar la composición así como la dinámica que existe en una comunidad microbiana (19). Entre ellas se encuentra la secuenciación por síntesis que, debido a su costo por lectura y porcentaje de error menor que otras tecnologías, es una de las más utilizadas actualmente (6)(20)(21). Se basa en reacciones fluorescentes de los cuatro diferentes nucleótidos (adenina, timina, citosina y guanina) que son detectadas por una cámara y que determina el nucleótido añadido a la lectura (20).

Una de las tecnologías desarrolladas por Illumina® es MiSeq *paired-end* que nos permite secuenciar lecturas de hasta 600pb en dos fragmentos, lo cual facilita el estudio de la diversidad microbiana basada en el perfil del gen 16S ARNr, que comprende aproximadamente 460 pares de bases.

## Control de calidad (QC)

La métrica de calidad comúnmente más usada en las tecnologías de secuenciación masiva es el nivel de calidad de Phred (valor Q). Este valor muestra la probabilidad de que una base dada sea mal identificada por el secuenciador (22). Los valores Q son definidos por la ecuación:

$$q = -10 \times \log_{10} p \quad (1)$$

Donde  $p$  es la probabilidad estimada del error para la asignación de esa base y  $q$  es el valor Phred. (tabla con los valores de Q). Un valor alto de calidad corresponde a una probabilidad baja de error y viceversa.

**Tabla 2:** Valores de Phred (Q); el valor ideal de calidad en secuenciación es de Q30; tomada y modificada de .

Valor de calidad Phred	Probabilidad de error	Precisión de asignación
10	1 cada 10	90 %
20	1 cada 100	99 %
30	1 cada 1,000	99.9 %
40	1 cada 10,000	99.99 %
50	1 cada 100,000	99.999 %

## Estrategias de selección de OTU

Para la clasificación de los amplicones obtenidos es necesario juntarlos en unidades taxonómicamente funcionales (OTU por sus siglas en inglés *Operational Taxonomic Unit*) que es un conjunto de secuencias similares entre sí; por ejemplo, si un grupo de secuencias tienen una identidad mayor o igual al 97% se asume que se trata de la misma especie; esto se considera también a niveles taxonómicos más altos y nos permite organizar a las lecturas basándose en su similitud (6). La organización en OTU nos permite determinar la diversidad de la muestra aunque su clasificación taxonómica no se encuentre en las bases de datos para el 16S (6)(19).

Existen tres tipos principales de selección de OTU, así como múltiples algoritmos para cada uno de los tipos de selección:

- **De novo:** agrupa las secuencias basándose en la identidad de éstas, sin considerar bases de datos de referencia.

- **Con referencia cerrada (*closed-reference*):** considera bases de referencia para agrupar a las secuencias basándose en la similitud de lo ya conocido; si una secuencia falla al emparejarse con la referencia es desechada.
- **Con referencia abierta (*open-reference*):** al igual que en la referencia cerrada compara las secuencias con una base de datos, sin embargo si una secuencia no se ajusta a las existentes, esta se agrupa como un nuevo OTU al cual se le asigna un identificador.

La selección de OTU recomendada es la de referencia abierta (*open-reference*), porque nos permite descubrir diversidad que no se encuentra en las base de datos de referencia (6), sin embargo si lo que se busca es realizar una comparación entre diversidades también es útil la referencia cerrada (*closed-reference*).

## Análisis de Diversidad

### Diversidad $\alpha$

La diversidad  $\alpha$  es aquella que considera la diversidad, en este caso bacteriana, que existe en una muestra determinada (23). Las aproximaciones para medir esta diversidad son las curvas de acumulación de especies o curvas de rarefacción, así como diversos índices de diversidad que se basan en la riqueza o abundancia de especies.

#### Curva de rarefacción:

La rarefacción es una técnica desarrollada por Sanders (1968) y corregida posteriormente por Hulbert (1971), calcula el número de especies esperadas en el caso de que todas las muestras poseyeseen el mismo número de individuos capturados. Una *curva de acumulación de especies* es la gráfica del número de especies observadas como función de alguna medida del esfuerzo del muestreo requerido para observarlas (24).

Su aplicación a un colectivo taxonómico vendría dada por el siguiente algoritmo (24):

$$E(G) = \sum_i \left[ 1 - \frac{\left( \frac{N-N_i}{n} \right)}{\left( \frac{N}{n} \right)} \right] \quad (2)$$

## MARCO TEÓRICO

---

Donde  $E(G)$  es el número de géneros esperados para un número de especies  $n$  de un colectivo con  $N$  especies en total de las cuales  $N_i$  están en el género  $i$ .

Para los individuos de una misma muestra, la rarefacción clásica (*basada en individuos*) puede ser utilizada para producir una curva suave que estima el número de especies que se observaría para cualquier número menor de individuos, bajo el supuesto de mezcla aleatoria de individuos (24). Para conjuntos de muestras replicados (conjuntos de muestras), el número esperado de especies que sería observado para cualquier número menor de muestras se puede estimar mediante la rarefacción *basada en muestras*, bajo el supuesto de orden aleatorio de muestras (25).

- **Índice de Shannon-Wiener ( $H'$ ):** Es uno de los índices más utilizados para cuantificar la biodiversidad específica, derivado de la teoría de información como una medida de la entropía. El índice refleja la heterogeneidad de una comunidad sobre la base de dos factores: el número de especies presentes y su abundancia relativa. Conceptualmente es una medida del grado de incertidumbre asociada a la selección aleatoria de un individuo en la comunidad (26). El índice de Shannon-Wiener se define como:

$$H' = - \sum_{i=1}^S \pi_i \ln \pi_i \quad (3)$$

Donde  $\pi$  es igual a la abundancia relativa ( $n_i/N$ ) y  $S$  es igual a número total de especies en el muestreo. El índice tiene valores de 0 cuando hay una sola especie y un valor máximo de  $\ln S$ . Entre más alto sea su valor, se considera que es mayor la diversidad representada por este índice.

- **Índice de Simpson ( $D$ ):** El índice de Simpson es una medida de diversidad que tiene en cuenta tanto la riqueza de especies, y una uniformidad de la abundancia de las especies presentes. En esencia se mide la probabilidad de que dos individuos seleccionados al azar de un área pertenezcan a la misma especie (27).

La fórmula para el cálculo de  $D$  se presenta como:

$$D = \frac{\sum n_i (n_i - 1)}{N(N - 1)} \quad (4)$$

Donde  $n_i$  es el número total de organismos de cada especie individual y  $N$

## MARCO TEÓRICO

---

es el total de organismos de todas las especies.

El valor de  $D$  cambia de 0 a 1. Con este índice, 0 representa diversidad infinita y, 1 que se considera una sola especie, por lo cual no hay diversidad. Esto es, mientras más grande sea el valor, menor será la diversidad.

- **Chao 1:** El estimador Chao 1 calcula la estimación de la diversidad de especies verdaderas de una muestra por la ecuación:

$$S_1 = S_{obs} + \frac{F_1^2}{2F_2} \quad (5)$$

Donde  $S_{obs}$  es el número de especies en la muestra,  $F_1$  el número de *singletons* (por ejemplo, el número de especies con una sola aparición en la muestra). La idea detrás del estimador es que si una comunidad ha sido muestreada, y especies raras (*singletons*) aún están siendo descubiertas, es probable que especies aún más raras no sean descubierta tan pronto como todas las especies que se han recuperado al menos dos veces (*doubletons*). Varios análisis del estimador han demostrado que proporciona estimaciones razonables, al menos para los conjuntos de datos modernos. Por supuesto, como el valor es una estimación hay un grado de incertidumbre, y un método para calcular la varianza de los estimadores ha sido proporcionada por Chao, en 1987, en la forma de:

$$var(S_1) = F_2 \left[ \left( \frac{(F_1/F_2)}{4} \right)^4 + (F_1/F_2)^3 + \left( \frac{F_1/F_2}{2} \right)^2 \right] \quad (6)$$

### Diversidad $\beta$

La diversidad  $\beta$  considera las diferencias entre las composiciones bacterianas para diferentes muestras (23). Existen dos aproximaciones para medir la diversidad  $\beta$ : aquellas que consideran las diferencias evolutivas entre secuencias, conocidas como de diversidad  $\beta$  filogenética; y aquellas que no consideran la filogenia, que están basadas en taxonomía (23)(28)(29).

- **UniFrac:** Es una métrica que utiliza información filogenética para comparar muestras ambientales. UniFrac junto con técnicas estadísticas multivariadas como el análisis de componentes principales (PCoA por sus siglas en inglés *Principal Coordinates Analysis*), identifican diferencias entre comunidades microbianas.

## MARCO TEÓRICO

---

El cálculo de las diferencias entre dos muestras con UniFrac está determinado por la suma de su historia evolutiva. Esta historia evolutiva es calculada por la longitud de la rama en un árbol filogenético que tiene cada muestra, sin considerar las ramas que comparten las muestras comparadas. Existen dos formas de calcular UniFrac: la no balanceada (en inglés *unweighted*), que no considera las frecuencias relativas, y la balanceada (en inglés *weighted*), que además de la filogenia considera las diferencias entre las abundancias relativas.

## Antecedentes

---

Las técnicas independientes de cultivo para caracterizar a las comunidades microbianas permitieron dar los primeros indicios de la microbiota presente en diferentes ambientes incluyendo en el intestino de *L. vannamei*. En una investigación realizada por Johnson y colaboradores en el 2008 se realizó el primer análisis de la comunidad microbiana del intestino de *L. vannamei*, así como del agua de su tanque de cultivo. Utilizando electroforesis en gel con gradiente de desnaturallización (DGGE) del gen ARNr 16S determinaron que las especies sobresalientes en el intestino de camarón son *Vibrio* spp., *Mycobacterium* spp., *Desulfocapsa* spp., *Desulfobulbus* spp. y *Clorella* spp.; mientras que en el agua se encontró gran diversidad relacionadas al género *Vibrio*. Aunque no se determinan todas las especies existente en el intestino del camarón blanco del Pacífico por las técnicas que utilizan, cuentan con conclusiones importantes para el conocimiento de la microbiota de esta especie. Una de dichas conclusiones es que la diversidad del agua del estanque es mayor que la del intestino del camarón, lo cual sugiere que existe una selectividad de microorganismos por parte camarón aunque no se sugiere algún mecanismo para la selección (1).

Si bien con el avance de las tecnologías de secuenciación masiva se han abierto las posibilidades de caracterizar con mayor detalle a la microbiota de múltiples ambientes y nichos, la aplicación de estas tecnologías ha sido utilizada en pocas investigaciones para conocer la diversidad de especies microbianas relacionadas con la familia Penaeidae. Un trabajo importante es el realizado por Rungranssamme y colaboradores en el 2014, que fue el primer artículo publicado que caracterizó la microbiota intestinal del camarón tigre (*Penaeus monodon*) considerando especímenes de vida libre y cultivados. Dicha especie está cercanamente emparentada a *L. vannamei*. Se determinaron a los phyla Proteobacteria, Firmicutes, Bacteroidetes, Fusobacteria y Actinobacteria en todas las muestras de intestino del camarón tigre; estos grupos taxonómicos se conservan entre las muestras, aunque cambiando sus abundancias, lo cual habla también de selección de ciertas especies en el intestino del camarón tigre más allá del ambiente donde habita (de vida libre o cultivado) (30).

## ANTECEDENTES

---

Posteriormente Rungranssamme y colaboradores realizaron un trabajo comparativo entre las comunidades microbianas de dos especies de camarón *Penaeus monodon* y *L. vannamei* ante la invasión de *Vibrio harveyi* el cual es un patógeno que causa vibriosis en camarones y que puede llevar a la muerte del organismo y con esto mermar su producción. En dicha investigación se utilizó un equipo de pirosecuenciación Roche 454 GS FLX+ sobre las regiones V3-4 del gen ARNr 16S. Los cuatro phyla con mayor abundancia encontrados en ambas especies de camarón fueron Firmicutes, Actinobacteria, Bacteroidetes y Proteobacteria, siendo éste último el que contenía mayor número de secuencias ( 82 % en ambas especies de camarón); dentro del phylum Proteobacteria la clase Gammaproteobacteria fue la más abundante. En este artículo demuestran que *L. vannamei* es más resistente ante la invasión de patógenos que *P. monodon* lo cual puede ser causado por la microbiota benéfica relacionada con *L. vannamei* (8).

Para determinar la microbiota intestinal del camarón blanco del Pacífico (*L. vannamei*), Xiong y colaboradores en el 2015 realizaron una investigación para de especímenes sanos y enfermos así como sedimentos de sus hábitats, utilizando secuenciación masiva con MiSeq de Illumina sobre la región hipervariable V4. En este trabajo concluyen que se concluye en este trabajo que el ambiente es el factor dominante en proporcionar la variabilidad entre las comunidades microbianas del intestino de *L. vannamei*. Además determinaron que los phyla más dominantes son Proteobacteria (en su mayoría de las clases Alphaproteobacteria y Gammaproteobacteria), Firmicutes y Bacteroidetes (31).

Por último, en una investigación realizada por Cornejo-Granados y colaboradores en el 2017, se caracterizó la microbiota intestinal de *L. vannamei* utilizando la secuenciación de siete regiones hipervariables del gen ARNr 16S (V2, V3, V4, V5, V6-7, V8 y V9) con la tecnología de IonTorrent para caracterizar las comunidades microbianas de muestras de intestino, hepatopáncreas y sedimento, dentro de los cuales se consideraron muestras de individuos cultivados (sanos, enfermos) y tomados de vida libre en océano Pacífico. La importancia de conocer la microbiota del sedimento donde habita el camarón es por el hábito alimenticio del camarón el cual alimentarse ingiere también sedimento, lo cual tiene un impacto directo en el microbioma de éste. En este artículo se determinó que los phyla más abundantes en el intestino son Proteobacteria, Cyanobacteria, Bacteroidetes y Actinobacteria, con cambios entre las abundancias en individuos sanos, enfermos y de vida libre; mientras que en las muestras predominan los phyla Proteobacteria, Gemmatimonadetes, Actinobacteria, Bacteroidetes y Acidobacteria. Los géneros más abundantes en el intestino de camarón fueron *Vibrio*, *Photobacterium* y *Paracoccus*. Además se concluye que la diversidad de especies la microbiota intestinal es más similar a la del sedimento que a la del hepatopáncreas (5), lo cual concuerda con el hábito alimenticio del camarón y su ingesta directa de sedimento.

La selección de la microbiota por parte del camarón no tiene resultados con-

---

cluyentes que nos muestren si es la genética o el ambiente lo que modula la microbiota intestinal, sin embargo dichos trabajos nos permiten contrastar con los resultados de nuestros experimentos.

## Hipótesis

---

La microbiota del camarón blanco (*Litopenaeus vannamei*) está determinada principalmente por su genética y posteriormente por el ambiente, por lo cual los estanques que comparten la misma genética contarán con comunidades microbianas más similares que los que cuentan con genética diferente.

# Objetivos

---

## Objetivo general

Determinar en la microbiota intestinal del camarón blanco del Pacífico (*Lithopaeus vannamei*) la influencia de la genética y el ambiente.

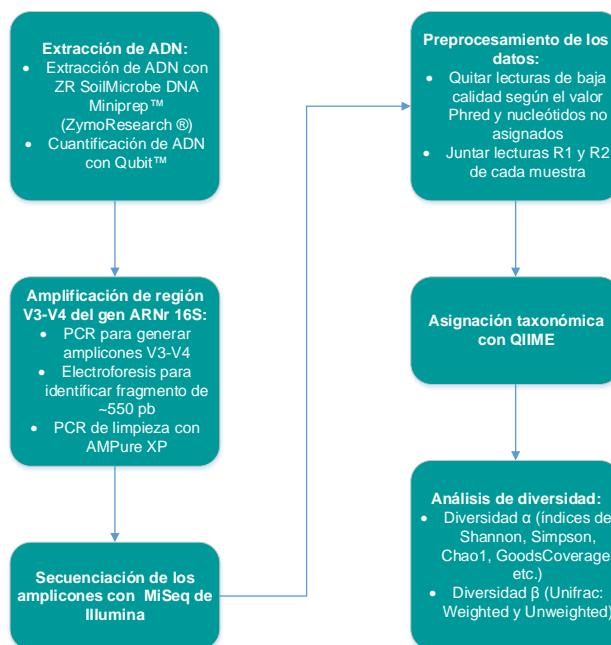
## Objetivos particulares

1. Estandarizar un protocolo de extracción de ADN para microbiota intestinal de camarón.
2. A partir del ADN de las muestras, obtener un conjunto de amplicones que abarquen a las regiones hipervariables V3 y V4 del gen 16S ARNr.
3. Análisis de los datos obtenidos de la secuenciación con el MiSeq de Illumina® de los amplicones obtenidos a partir de muestras de intestino de camarón y sedimento.
4. Comparaciones entre los datos obtenidos de la secuenciación entre la microbiota del intestino del camarón y del sedimento.

# Materiales y métodos

---

La metodología a seguir en este trabajo se agrupa en seis bloques principales (figura 7) divididos en dos partes: la primera parte que incluye las técnicas de biología molecular y que termina al momento de la secuenciación; y la segunda parte que abarca el análisis bionformático de la información producida tras la secuenciación con MiSeq de Illumina.



**Figura 7:** Diagrama que muestra la metodología dividida en seis bloques principales.

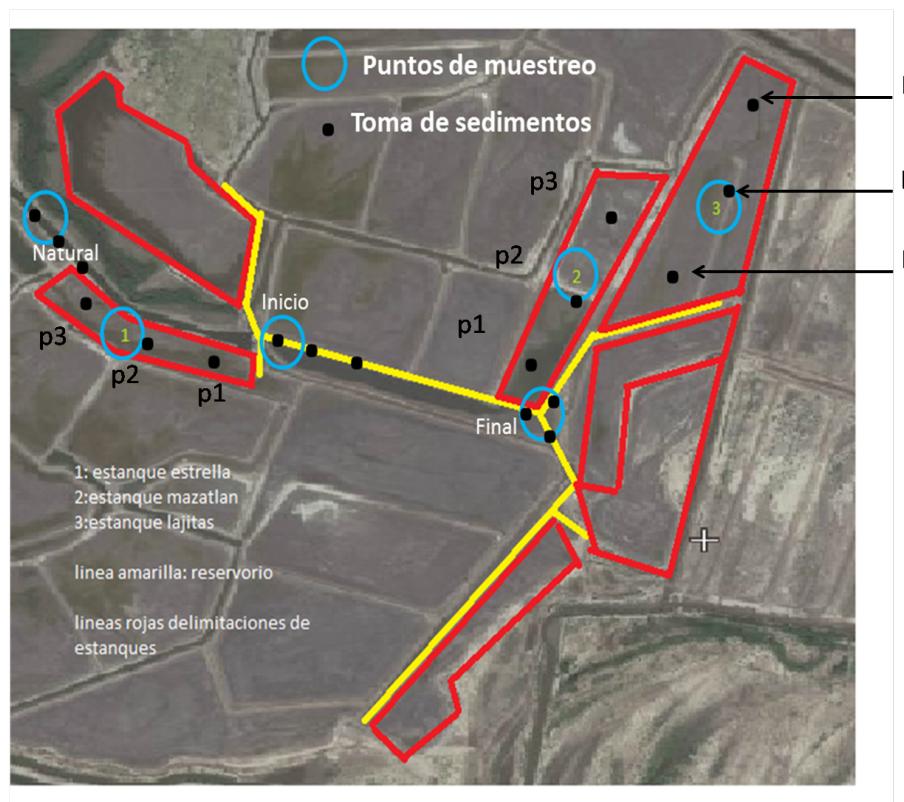
## MATERIALES Y MÉTODOS

---

### Colecta de las muestras

La colecta de los especímenes se realizó el 17 y el 18 de Junio del 2015 en tres estanques de cultivo de camarón en Sinaloa: Estanque Mazatlán (L), Estanque Lajitas (M) y Estanque Estrella (E). Dentro de cada estanque se determinaron tres puntos (P1, P2 y P3) en los cuales se tomaron los especímenes de camarón y la respectiva muestra de sedimento. Se tomaron nueve camarones de cada estanque, tres en cada uno de los puntos de muestreo del estanque en cuestión, a los cuales se les extrajo el intestino en el lugar de la colecta y se les colocó en *RNA later®*. Se colocaron en tubos independientes a los cuales se les agregó una etiqueta que los identificara (tabla 3 y figura 8).

Para la muestra de sedimento se utilizó un tubo falcón de 50 ml (o una bolsa de plástico) y además se tomó 1 g que se colocó en un tubo de 1.5 ml con *RNA later®* en cada uno de los puntos de colecta (tabla 3 y figura 8).



**Figura 8:** Mapa con los estanques y sitios de muestreo; Estanque Estrella (1), Estanque Mazatlán (2) y Estanque Lajitas (3).

## MATERIALES Y MÉTODOS

---

**Tabla 3:** Tabla con las muestras de intestino y sedimento para cada uno de los estanques; para cada estanque (Lajitas, Mazatlán y Estrella) se consideraron tres puntos de colecta y en cada uno de éstos tres muestras de intestino y una muestra de sedimento.

Estanque Lajitas		Estanque Mazatlán		Estanque Estrella	
Intestino	Sedimento	Intestino	Sedimento	Intestino	Sedimento
IA1	SA1	IL1	SL1	IE1	SE1
IA2		IL2		IE2	
IA3		IL3		IE3	
IA4	SA2	IL4	SL2	IE4	SE2
IA5		IL5		IE5	
IA6		IL6		IE6	
IA7	SA3	IL7	SL3	IE7	SE3
IA8		IL8		IE8	
IA9		IL9		IE9	

## Extracción del ADN con miniprep

Para la extracción del ADN tanto en las muestras de intestino se utilizó el *ZR SoilMicrobe DNA Miniprep™* de la marca *ZymoResearch®* (núm. de catálogo D6001). Se siguió el protocolo que se indica en el manual del producto con algunas variaciones para optimizar la extracción ADN bacteriano en nuestras muestras.

El procedimiento que se realizó para la extracción de ADN en cada una de las muestras fue el siguiente:

1. A un tubo *ZR BashingBead™ LysisTube* se le agregó la muestra:
  - a) Intestino: todo el que se obtuvo del camarón en cuestión (en las muestras 3, 6 y 9 de cada estanque solamente se tomó la mitad del intestino)
  - b) Sedimento: 250 mg.
2. Se añadieron 750 µL *LysisSolution* al *ZR BashingBead™ LysisTube*. Utilizar el vortex para homogeneizar la muestra con la *LysisSolution*.

## MATERIALES Y MÉTODOS

---

3. En un *VortexGenie<sup>TM</sup>* se colocó *ZR BashingBead<sup>TM</sup> LysisTube* para promover la lisis del tejido o sedimento a la máxima velocidad, con agitación constante durante 5 min.
4. Posteriormente se llevó el *ZR BashingBead<sup>TM</sup> LysisTube* a una microcentrífuga a 10,000 g durante 1 min para separar la parte acuosa.
5. Se tomaron 400 µL del sobrenadante y se llevaron a un *Zymo-Spin<sup>TM</sup> IV Spin Filter* el cual se colocó dentro de *CollectionTube* para obtener el producto del filtrado tras una centrifugación a 7,000 g durante 1 min.
6. Al *CollectionTube* del paso anterior se le agregaron 1200 µL de *Soil DNA Binding Buffer*. Se homogeneizó con la punta de micropipeta, subir y bajar 10 veces.
  - a) Como lo indica el manual de uso del *ZR SoilMicrobe DNA Miniprep<sup>TM</sup>* se agregó beta-mercaptoetanol al *Soil DNA Binding Buffer* con una concentración final de 0.5% (v/v) para un óptimo desempeño.
7. De los 1600 µL de la mezcla del paso anterior, se tomaron 800 µL y se colocaron en un *Zymo-Spin<sup>TM</sup> IIC Column* dentro de un *CollectionTube* y se centrifugó a 10,000 g durante 1 min. Pasando la centrifugación se tiró el filtrado del tubo colector.
8. Los 800 µL restantes de la mezcla del paso 6 se colocaron en el mismo *Zymo-Spin<sup>TM</sup> IIC Column* dentro del *CollectionTube* y se centrifugó a 10,000 g durante 1 min. El filtrado del tubo colector se desechará.
9. La *Zymo-Spin<sup>TM</sup> IIC Column* se traspasó a un nuevo *CollectionTube* y se le agregaron 200 µL de *DNA Pre-Wash Buffer* y se centrifugó a 10,000 g durante 1 min. Al finalizar se tiró el filtrado del tubo colector.
10. Se agregaron 500 µL de *Soil DNA Wash Buffer* a la *Zymo-Spin<sup>TM</sup> IIC Column* dentro del mismo *CollectionTube* y se centrifugó nuevamente a 10,000 g durante 1 min. Al finalizar se tiró el filtrado junto con el *CollectorTube*.
11. La *Zymo-Spin<sup>TM</sup> IIC Column* fue llevada a un tubo para microcentrífuga de 1.5 mL y se añadieron 300 µL 1 de agua libre de ADNasas y ARNasas directo a la matriz de la columna para las muestras 3, 6 y 9 de cada estanque se colocaron 30 µL. Se centrifugó a 10,000 g durante 30 s.
12. Se transfirió nuevamente la *Zymo-Spin<sup>TM</sup> IIC Column* a un nuevo tubo para microcentrífuga de 1.5 mL y se añadieron 30 µL, para una segunda elución, de agua libre de ADNasas y ARNasas (para las muestras 3, 6 y 9 de cada

## MATERIALES Y MÉTODOS

---

estanque se añadieron 100  $\mu$ L). Por último se centrifugó a 10,000 durante 30 s.

13. El ADN eluído del último paso se transfirió a un *Zymo-Spin<sup>TM</sup> IV-HRC Spin Filter* (previamente centrifugado a 8,000 g durante 1 min) dentro de un tubo para microcentrífuga de 1.5 mL y se realizó una última centrifugación a 8,000 g durante 1 min.

Posteriormente para determinar el contenido de ADN extraído se cuantificó con QUBIT<sup>TM</sup> siguiendo el protocolo del fabricante.

## Amplificación de la región V3-V4 del gen 16S ARNr

Los amplicones se generaron basándose en el protocolo propuesto por Illumina para la preparación de librerías para utilizarse en MiSeq . Las regiones hipervariables utilizadas fueron la V3-V4 contenidas en un solo amplicón de tamaño esperado de ~460 pb. Los cebadores utilizados fueron los siguientes:

Foward 5':

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGNGGC WG-CAG

Reverse 5':

GTCTCGTGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGG TAT- CTAATCC

Los cebadores fueron sintetizados en la Unidad de Síntesis de Instituto de Biotecnología de la UNAM (figuras A.1 y A.2). Para su uso fueron diluidos a una concentración de 2.5  $\mu$ g/ $\mu$ L. La polimerasa utilizada para la preparación de los amplicones fue la NEBNext® High-Fidelity.

## Programación del termociclador

Para las muestras de intestino se tomaron 300 ng del templado de ADN, ajustando el volumen para cada una de las muestras según su concentración. Se

## MATERIALES Y MÉTODOS

---

programó el termociclador de la siguiente manera:

- 95 °C durante 3 min
- 35 ciclos de:
  - 95 °C durante 30 s
  - 55 °C durante 30 s
  - 72 °C durante 30 s
- 72 °C durante 5 min
- Mantener a 4 °C

Para las muestras de sedimento se tomaron 12.5 ng de templado de ADN, ajustando el volumen de la muestra dependiendo de su concentración. Se programó el termociclador de la siguiente manera:

- 95 °C durante 3 min
- 25 ciclos de:
  - 95 °C durante 30 s
  - 55 °C durante 30 s
  - 72 °C durante 30 s
- 72 °C durante 5 min
- Mantener a 4 °C

## PCR *clean up*

Los productos de PCR se purificaron con AMPure XP, que son perlas magnéticas con afinidad al ADN las cuales nos permiten retirar a los contaminantes resultantes de la PCR, como cebadores, sales, dinucleótidos, etc. Se siguió el protocolo del fabricante y posteriormente se tomaron 5 µL y se corrieron en un gel de agarosa al 2% para determinar su integridad.

## Secuenciación masiva

La secuenciación de los amplicones ya purificados se llevó a cabo en el Instituto de Medicina Genómica (INMEGEN) con un equipo MiSeq de Illumina *paired-end*, que utiliza la secuenciación por síntesis. Para poder secuenciar las regiones hiper-variables V3-V4 (~460 pb). La longitud máxima de las lecturas de MiSeq es de 300 pb, sin embargo con la opción de *paired-end* podemos secuenciar fragmentos de ADN de hasta 600 pb divididos en dos lecturas: R1 y R2; por lo cual para cada una de las muestras se obtuvieron dos archivos *fastqc* (R1-muestra.fastqc y R2-muestra.fastqc).

## Preprocesamiento de los datos

Uno de los pasos más importantes para determinar la diversidad microbiana es el pre-procesamiento de los datos, ya que se puede modificar notablemente el éxito de la clasificación de las secuencias y con esto la explicación de la diversidad de la muestra (6)(21). Para cada muestra se debe de utilizar un pre-procesamiento de los datos dependiendo de cuáles sean los factores que se encuentren presentes en dicha muestra. estos factores pueden ser: la extracción del ADN, la elección de los cebadores, la variación del número de copias, la presencia de quimeras, errores de secuenciación y error de clasificación en OTUs (6)(19).

Existen diversas herramientas para visualizar las muestras obtenidas de la secuenciación, una de las más destacadas es FastQC (32) por las funciones que ofrece, entre las que destacan:

1. Estadísticas básicas de la muestra como número de lecturas que se obtuvieron del secuenciador así como su tamaño mínimo y máximo.
2. La calidad de la lectura por cada base y por la lectura completa.
3. Contenido de GC por base y por la lectura completa.
4. Contenido de bases sin asignar (N) en la lectura completa.
5. Distribución del tamaño de las lecturas en toda la secuenciación.
6. Niveles de duplicación de la lectura.
7. Secuencias sobre-representadas en cada lectura
8. Contenido de k-meros.

## MATERIALES Y MÉTODOS

---

Una de las características más útiles de esta herramienta es que destaca cuáles son los factores que pudiesen tener errores y que es necesario revisar. Una vez conociendo los resultados de calidad es necesario realizar curaciones a los datos y con esto disminuir los errores al momento de realizar la clasificación.

Debido a que se observó que existía un porcentaje elevado de nucleótidos no asignados (N) en todas las muestras se cortaron los extremos iniciales de todas las muestras. Basándonos en los resultados de FastQC para los archivos R1 se cortaron 5 nucleótidos y para los archivos R2 6 nucleótidos con Trimmomatic (figuras 15 y 16).

Posteriormente se juntaron ambas lecturas (R1 y R2) para generar el amplicón con las regiones hipervariables V3-V4. Este procedimiento fue realizado con el comando *fastq-join* del programa Bioconda. Se observó que existían variaciones entre el número de secuencias obtenidas al juntar los fragmentos R1 y R2, por lo cual se cortaron las lecturas de 230 a 245 nucleótidos y se comparó el número de secuencias finales obtenidas y se seleccionó para cada una de las muestras cuál era el corte con el cual se obtenían más lecturas finales.

Cuando ya se contaba con el amplicón junto se procedió a limpiarlo por calidad. La calidad mínima considerada fue de Q20, que según el *Phred Quality Score* se tiene un 99 % de exactitud al asignar una base. Se utilizó el filtrado por ventanas de 6 bases, esto es que se promediaba la calidad de la ventana y si era superior a Q20 se conservaba la lectura y continuaba el filtrado avanzando una base y nuevamente promediando hasta llegar al final de la secuencia. En el caso de que la ventana no cumpliera con la calidad mínima promedio esta se cortó, quitando la parte de baja calidad.

El último filtro utilizado fue el de nucleótidos degenerados (N) dentro de la secuencia. Como se menciona en el primer paso del pre-procesamiento se buscó la presencia de N en las secuencias, sin embargo FastQC detecta la presencia de N si se encuentran en más del 5 % en una posición, por lo cual si hay secuencias con N en baja proporción dicho programa no podría detectarlas. Otro de los filtros con en este paso fue el de eliminar las secuencias menores a 5 bases, que podrían ser solamente restos de oligos secuenciados.

## Asignación taxonómica con QIIME

### Comparación entre bases de datos

Para poder asignar taxonomía a un OTU es necesario realizar una comparación de su secuencia contra una secuencia ya conocida. Para dicho fin se han

## MATERIALES Y MÉTODOS

---

establecido bases de datos públicas de donde se puede descargar el conjunto de secuencias del 16S ARNr. Las tres más importantes y con mayor número de secuencias depositadas son: Greengenes (33), Ribosomal Database Project (RDP) (34) y SILVA (35).

Cada una de estas bases de datos son curadas por diferentes grupos o instituciones, por lo cual el contenido de información respecto a las secuencias y sus asignaciones taxonómicas varía. Para nuestra investigación es de interés conocer cuál es la base de datos que ofrece una cobertura mayor de la diversidad microbiana en las muestras de intestino de camarón por consiguiente se decidió realizar una comparación entre la asignación de taxonomía entre las tres bases de datos, así como una comparación entre los índices de diversidad resultantes.

Para realizar la comparación se tomó una muestra al azar, que ya había pasado por el pre-tratamiento antes mencionado. Se procesó con QIIME (36) con sus valores por defecto pero cambiando la base de datos en cuestión (Greengenes, RDP y SILVA).

La primera evaluación que se realizó fue cuantitativa en la cual se evaluaron el número de lecturas asignadas, número de OTU generados, porcentajes de asignación dentro de los niveles taxonómicos (8). Posteriormente se realizó un diagrama de Venn a nivel de género entre las tres bases de datos (figura 20).

Con la base de datos con mejor asignación (Greengenes) se llevó a cabo la selección de OTU con referencia cerrada. Se consideraron los parámetros de 97% de identidad así como predicción de la cadena complementaria para aumentar en número de lecturas por muestra.

## Asignación de taxonomía en muestras de intestino con Greengenes

Debido a que el objetivo es conocer si la microbiota intestinal está modulada por su genética o por su ambiente, se tomaron en cuenta estas dos aproximaciones para realizar diferentes análisis. Con el fin de determinar si la influencia de la microbiota se debía a su genética, se compararon las muestras del estanque Mazatlán (genética 1) y del estanque Lajitas (genética 2); se consideran estos dos grupos ya que comparten el mismo ambiente (figura 8). De igual forma se compararon las muestras de los dos estanques que poseían la misma genética (estanques Mazatlán y Estrella), esto como prueba para determinar si el ambiente era el que tenía una mayor influencia sobre la microbiota intestinal del camarón. Los resultados que se muestran incluyen a los 10 OTU con mayor abundancia para cada nivel taxonómico. De igual forma se filtraron los OTU con una abundancia mayor a 0.1% y se utilizaron diagramas de Venn para contrastar entre la diversidad de

## MATERIALES Y MÉTODOS

---

los estanques, genéticas y ambientes. El código utilizado en QIIME se muestra en el Apéndice de esta tesis.

### Análisis LEfSe

Con el fin de evaluar cuáles son los taxa que varían en los dos ambientes comparados, así como las dos genéticas, se utilizó el algoritmo LEfSe. Éste nos permite evaluar las diferencias significativas entre los grupos. El resultado en valor LDA (del análisis discriminante lineal que es el último paso de la prueba LEfSe) nos muestra que tanto varía un elemento, entre los grupos comparados (ambientes o genéticas). Entre mayor sea el valor de LDA, mayor será la diferencia de ese elemento en comparación a las otras condiciones. Se realizó la prueba a diferentes niveles taxonómicos (phylum, clase, orden, familia, género y especie) comparando a las muestras por estanque, genética y ambiente, así como dos diferentes niveles de abundancia de OTU: uno con todos los OTU y otro con solamente los que cuentan con más del 0.1% de la abundancia. Como se muestra en la tabla 16 cuando se utilizan los grupos con abundancias mayores a 0.1% se encuentran un menor número de taxones diferenciales por lo que se muestran dichos resultados. Además se realizó el análisis LEfSe en los diferentes niveles taxonómicos agrupando a las muestras por genéticas y por ambientes. Las muestras de microbiota intestinal del estanque Mazatlán se consideraron como ambiente 1 y las muestras del estanque Estrella como ambiente 2, esto como se menciona anteriormente se debe a que los camarones de los cuales se extrajo el intestino provienen de un linaje similar para estos estanques pero difieren en su ambiente.

### Análisis de diversidad $\alpha$

La diversidad  $\alpha$  es aquella que existe en una muestra determinada. Se separó a las muestras por estanque y por genética al igual que en la taxonomía. Además se añaden las muestras de sedimento para determinar qué es más diverso, si la microbiota intestinal o la que se encuentra en el sedimento.

Algunas de las muestras tenían una menor cantidad de lecturas por lo cual sus curvas de rarefacción tienen una profundidad menor. Para poder realizar una comparación más robusta se realizaron 10 000 iteraciones tomando subgrupos con el tamaño máximo de lecturas de la muestra más pequeña (con menos lecturas). Además se realizó un cuadro a una profundidad en la cual se pudieran apreciar todas las muestras, incluyendo a las de menor cantidad de lecturas.

## MATERIALES Y MÉTODOS

---

Para cada una de las curvas de rarefacción se realizó una comparación con un método no paramétrico (Monte-Carlo) para identificar si existen diferencias significativas entre las diversidades  $\alpha$  de los grupos considerados, tanto por estanque y por genética. La comparación entre ambientes se realizó agrupando a las muestras del estanque Mazatlán como ambiente 1 y las del estanque Estrella como ambiente 2. Estos dos estanques comparten la misma fuente de desove, esto es que pertenecen al mismo linaje y cuentan con una genética similar. La comparación se realizó al igual que con los estanques con los índices de diversidad Shannon, diversidad filogenética y Chao1.

### Análisis de diversidad $\beta$

La diversidad  $\beta$  es aquella que existe entre un conjunto de muestras y nos permite identificar diferencias entre éstas. Al igual que con el análisis de taxonomía y de diversidad  $\alpha$  se separaron las muestras por tipo de muestra (sedimento o intestino), estanque y genética. Se utilizó UniFrac como métrica para comparar las comunidades microbianas.

Se realizaron los PCoA (Análisis de Coordenadas Principales, en inglés *Principal Coordinates Analysis*) con 10 000 iteraciones considerando la abundancia (ponderado, en inglés *weighted*) y sin considerarla (no ponderado, en inglés *unweighted*). Debido a que algunas muestras tienen una menor cantidad de lecturas los subgrupos considerados para las iteraciones fueron al 75 % de la muestra más pequeña.

El análisis se realizó considerando las siguientes comparaciones: entre las muestras de los estanque Mazatlán y Estrella (ambiente), comparando las muestras de Mazatlán y Lajitas (genética) y por último considerando todos los estanques. De igual forma se realizaron los análisis considerando solamente las muestras de intestino, así como junto con las muestras de sedimento.

### Árbol filogenético

Para agrupar a las muestras en un árbol filogenético se utilizó el método UPG-MA junto con 1000 submuestreos con Jackknife. Se hicieron dos aproximaciones: considerando la abundancia de los OTU (en inglés *weighted*) o tomando en cuenta solamente la presencia o ausencia de éstos (en inglés *unweighted*).

## MATERIALES Y MÉTODOS

---

### PCA con abundancias de OTU

Por último se realizó un análisis de componentes principales PCA con el fin de evaluar si la diversidad entre estanques agrupaba a las muestras de intestino y sedimento por su genética o por su ambiente. En este análisis se consideran a todos los OTU, incluyendo a los de abundancias menores al 0.1 %.

### Mapas de calor por ambiente

Se utilizaron mapas de calor para organizar a las muestras de intestino a partir de sus frecuencias absolutas, que en este tipo de gráfico se representan como un gradiente de color. Se realizó un mapa de calor para cada nivel taxonómico y se conservaron solamente a los OTU que tenían una abundancia mayor al 0.1 %, contrastando primero entre ambientes y posteriormente entre genéticas.



## Resultados

---

### Extracción de ADN y obtención de amplicones

**Tabla 4:** Concentraciones de ADN en ng.

Muestra	Concentración de ADN
IA1	23.13
IA2	23.87
IA3	104
IA4	25
IA5	20.6
IA6	88.4
IA7	12.93
IA8	39.6
IA9	131.73
IL1	21.13
IL2	4.09
IL3	187.5
IL4	9.87
IL5	9.79
IL6	107.14
IL7	15.69
IL8	17.92
IL9	65.21
IE1	13.63
IE2	22.33
IE3	55.47
IE4	5.19
IE5	4.7
IE6	56.53
IE7	14.46
IE8	4.91
IE9	53.2

## RESULTADOS

---

La mezcla que se utilizó para la reacción de PCR se puede observar en la tabla 5. En las figuras 9 y 10 se observan ejemplos de los geles de agarosa que se consideraron para determinar la integridad de ADN, antes y después de su purificación, respectivamente.

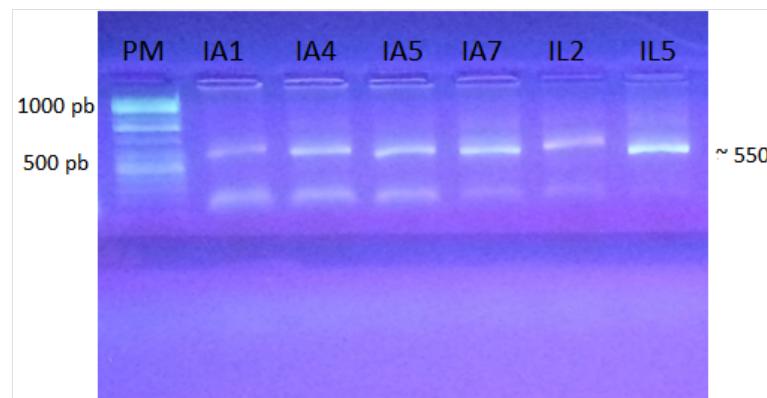
**Tabla 5:** Tabla con los volúmenes utilizados para la reacción de PCR.

Muestra	ADN	Rev	Fow	H2O	NEBNext® High-Fidelity
IA1	13 µL	2 µL	2 µL	3 µL	20 µL
IA2	13 µL	2 µL	2 µL	3 µL	20 µL
IA3	2.9 µL	2 µL	2 µL	5.6 µL	12.5 µL
IA4	12 µL	2 µL	2 µL	4 µL	20 µL
IA5	15 µL	2 µL	2 µL	1 µL	20 µL
IA6	4.6 µL	2 µL	2 µL	5.1 µL	12.5 µL
IA7	23 µL	2 µL	2 µL	0 µL	25 µL
IA8*	7.5 µL	5 µL	5 µL	2.5 µL	20 µL
IA9	2.3 µL	2 µL	2 µL	6.2 µL	12.5 µL
IL1	15 µL	2 µL	2 µL	1 µL	20 µL
IL2	36 µL	2 µL	2 µL	0 µL	40 µL
IL3	1.6 µL	2 µL	2 µL	6.9 µL	12.5 µL
IL4	31 µL	2 µL	2 µL	0 µL	30 µL
IL5	31 µL	2 µL	2 µL	0 µL	35 µL
IL6	2.8 µL	2 µL	2 µL	5.7 µL	12.5 µL
IL7	19 µL	2 µL	2 µL	2 µL	25 µL
IL8*	16.8 µL	5 µL	5 µL	0 µL	25 µL
IL9	9 µL	2 µL	2 µL	0 µL	12.5 µL
IE3	5.4 µL	2 µL	2 µL	3.1 µL	12.5 µL
IE5	30 µL	2 µL	2 µL	0 µL	35 µL
IE6	5.3 µL	2 µL	2 µL	3.2 µL	12.5 µL
IE9	5.7 µL	2 µL	2 µL	2.8 µL	12.5 µL

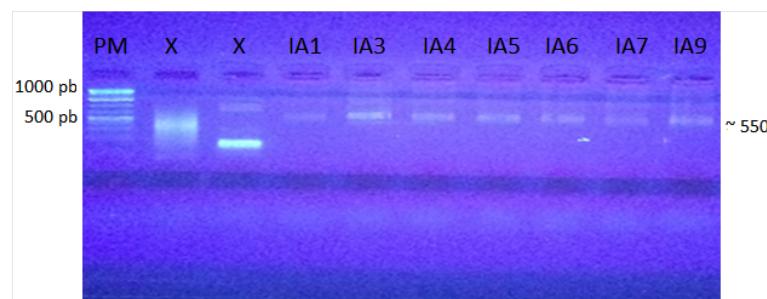
De las muestras IL5 (intestino del estanque Mazatlán 5) e IE4 (intestino del estanque Estrella 4) no se obtuvo el amplicón, por lo cual no se enviaron a secuenciar.

## RESULTADOS

---



**Figura 9:** Gel de agarosa al 2% con buffer TBE 1X, donde se observan los amplicones de  $\approx$ 550 pb en las muestras de intestino; foto a los 10 min con 100 V.



**Figura 10:** Gel de agarosa al 2% con buffer TBE 1X, donde se observan los amplicones después de la limpieza con AMPure XP; foto a los 10 min con 100 V; las muestras con X, no son muestras de intestino.

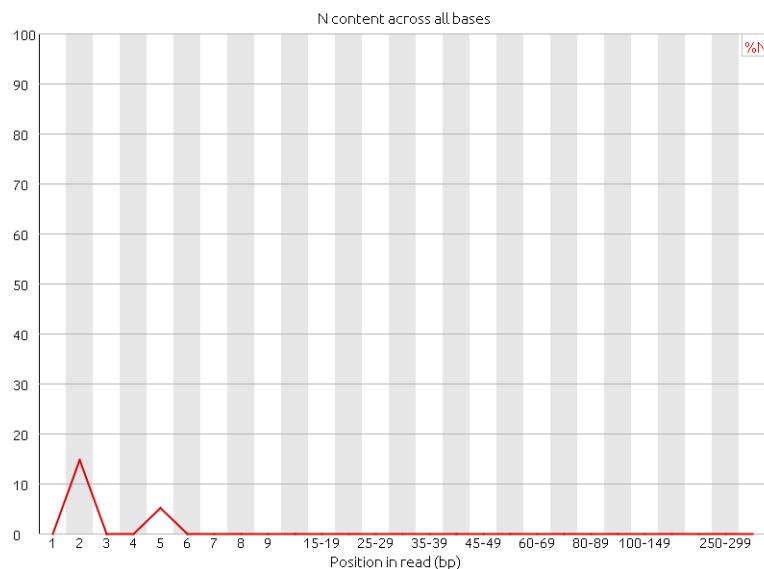
## RESULTADOS

---

### Pre-procesamiento

#### FastQC para visualizar las secuencias en crudo

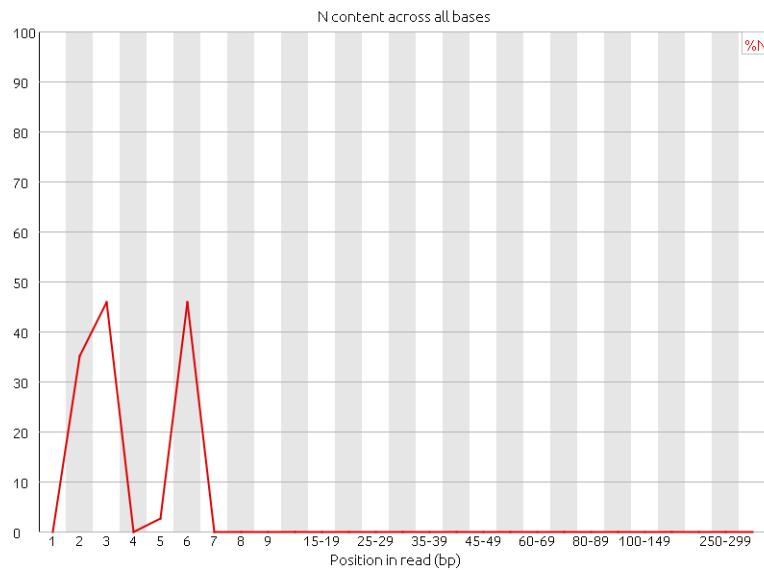
De las figuras 11 a la 14 se pueden observar el contenido de N (nucleótidos no asignados) y la calidad por base a lo largo de los amplicones.



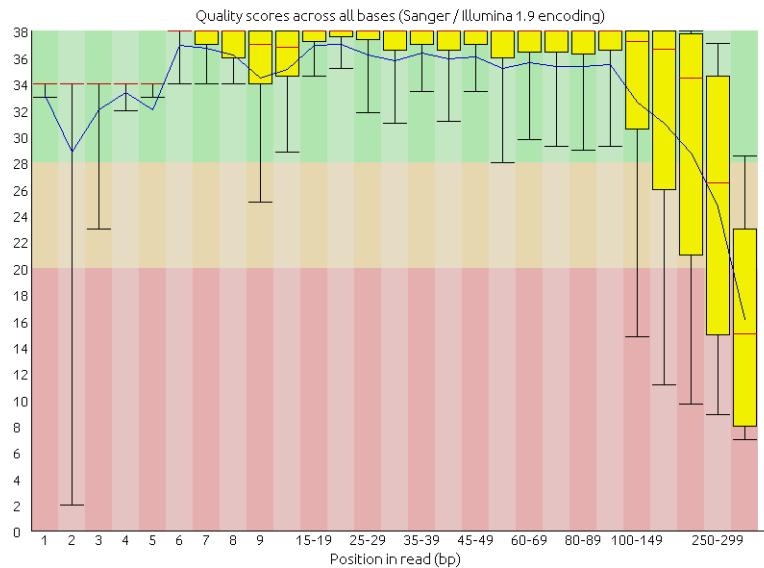
**Figura 11:** Imagen obtenida con FastQC de la muestra R1-IA9 antes del pre-procesamiento; muestra en contenido de N (nucleotidos no asignados).

## RESULTADOS

---



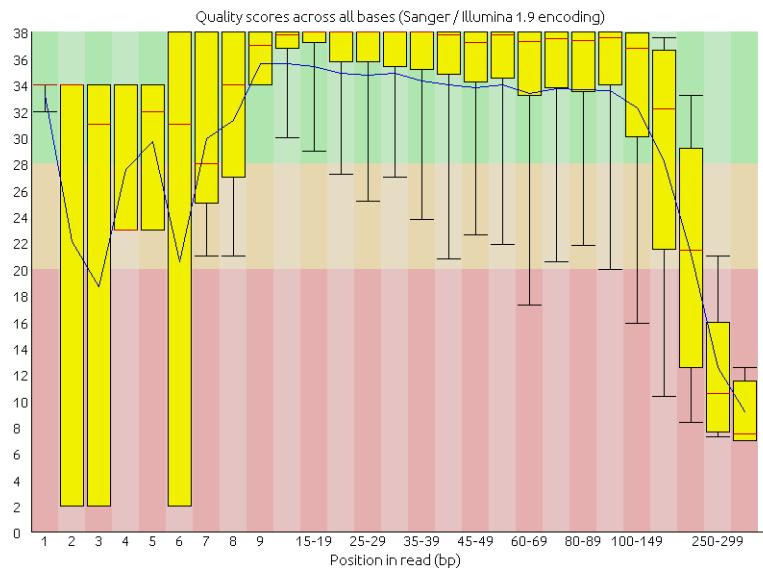
**Figura 12:** Imagen obtenida con FastQC de la muestra R2-IA9 antes del pre-procesamiento; muestra en contenido de N (nucleotidos no asignados)



**Figura 13:** Imagen obtenida con FastQC de la muestra R1-IA9 antes del pre-procesamiento; muestra la calidad por base basado en el *Phred Quality Score*.

## RESULTADOS

---

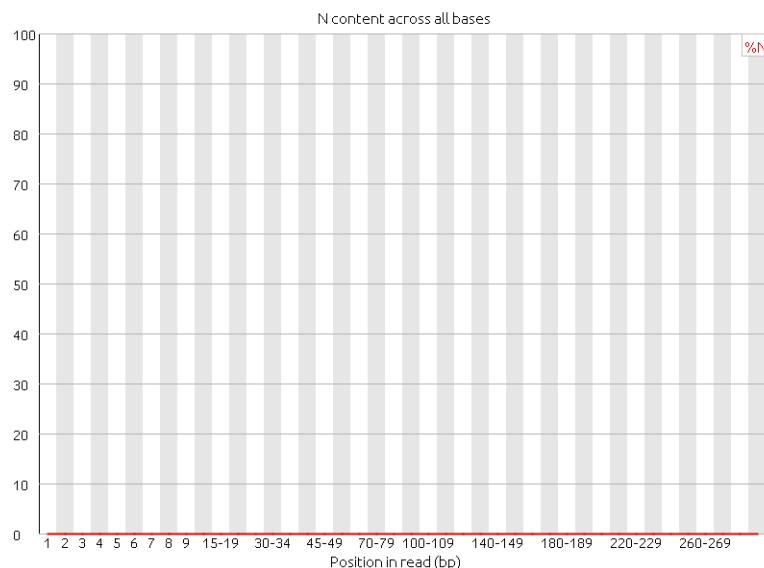


**Figura 14:** Imagen obtenida con FastQC de la muestra R2-IA9 antes del pre-procesamiento; muestra la calidad por base basado en el *Phred Quality Score*.

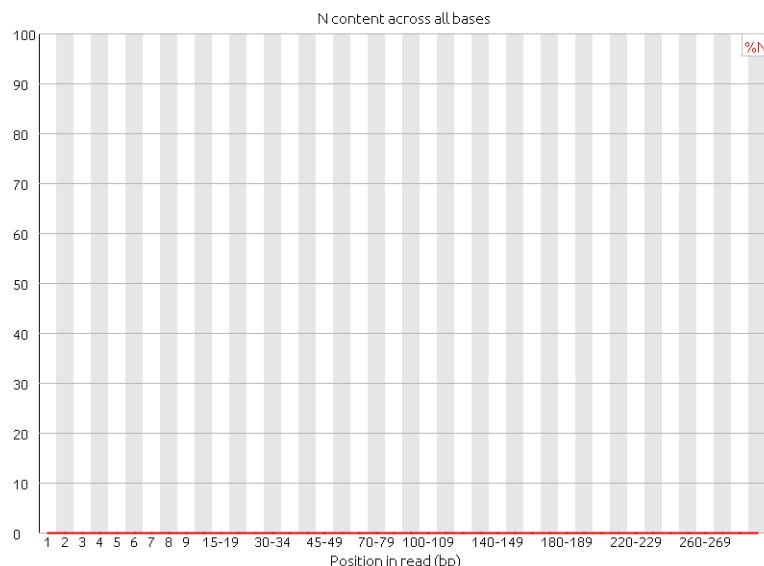
## RESULTADOS

---

### Corte de los extremos de las lecturas con Trimmomatic



**Figura 15:** Imagen obtenida con FastQC de la muestra R1-IA9 después del corte de los primeros 5 nucleótidos.

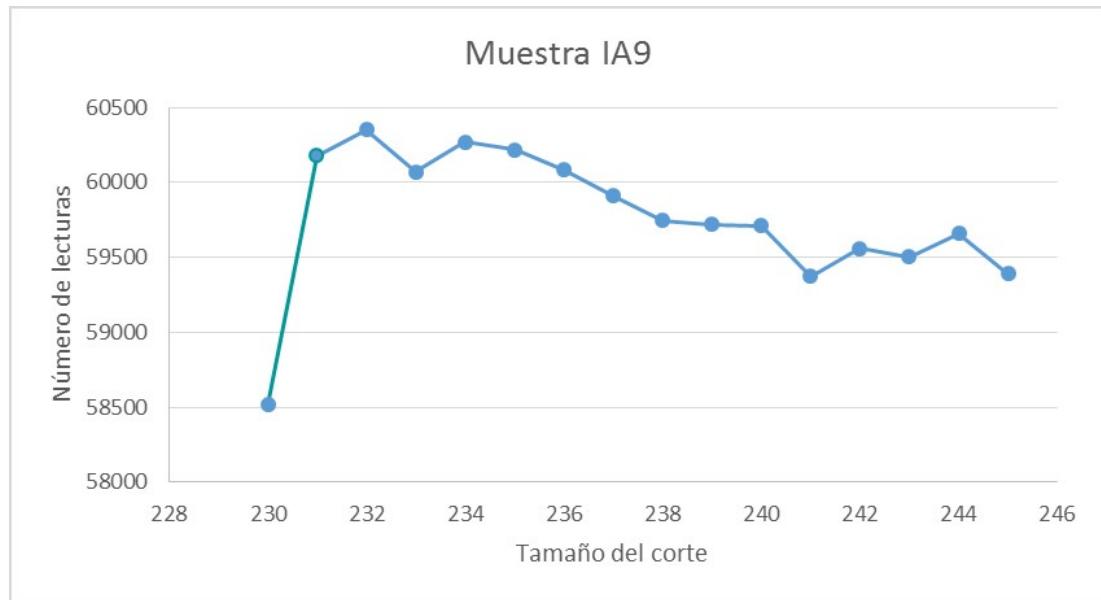


**Figura 16:** Imagen obtenida con FastQC de la muestra R2-IA9 después del corte de los primeros 6 nucleótidos.

## RESULTADOS

---

### Unión de lecturas R1 y R2 para obtener amplicón completo



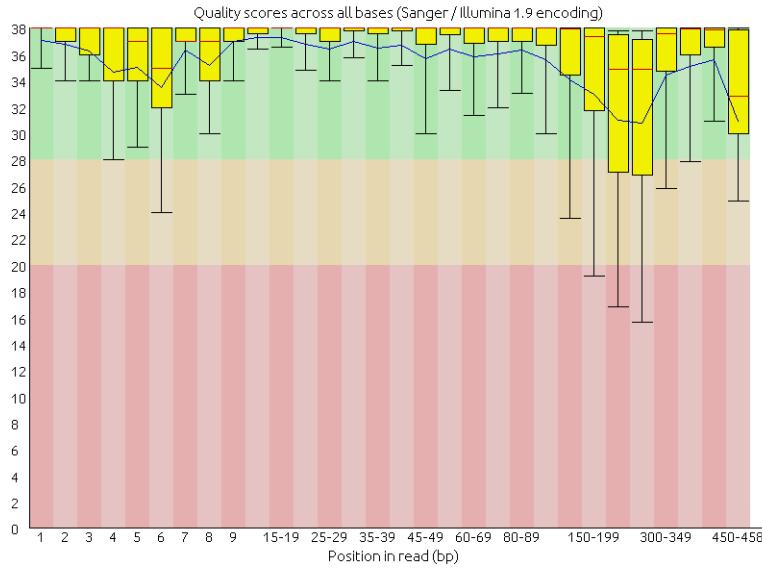
**Figura 17:** Gráfica de los resultados de la unión de R1 y R2 de la muestra IA9 a diferentes cortes (de 230 a 245); el corte con mayor número de lecturas juntas en esta muestra fue de 232 con 60354 secuencias totales.

### Limpieza por calidad en ventanas con Trimmomatic

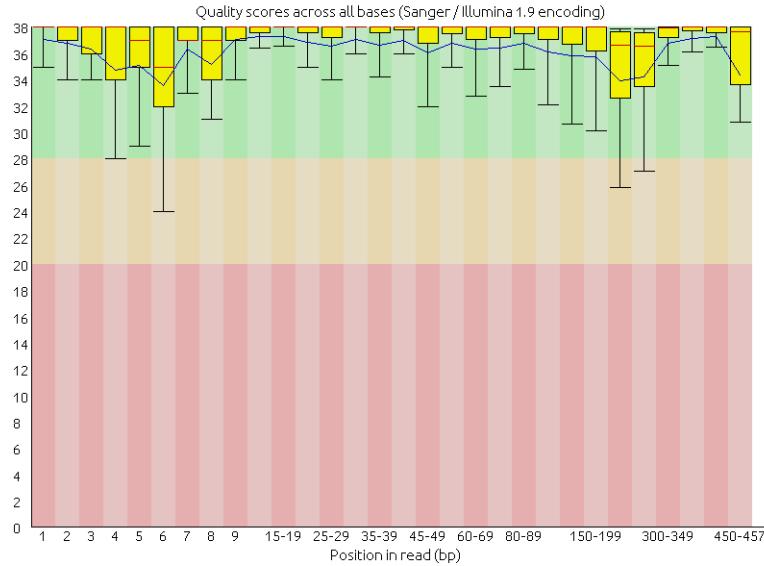
La calidad de las secuencias antes de la limpieza por ventanas (figura 18) si se ve mejorada al realizar la limpieza por ventanas (figura 19)

## RESULTADOS

---



**Figura 18:** Imagen obtenida con FastQC de la muestra IA9 después de juntar R1 y R2; muestra la calidad por base basado en el *Phred Quality Score*; se observa que hay calidades por debajo del Q20 con una mayor probabilidad de error al asignar una base.



**Figura 19:** Imagen obtenida con FastQC después de la limpieza por ventanas; muestra la calidad por base basado en el *Phred Quality Score*; se observa que no hay bases con calidad menor a Q20.

## RESULTADOS

---

### Resultados del pre-procesamiento

Algunas librerías conservan un mayor porcentaje de lecturas (tablas 6 y 7) al finalizar la limpieza, esto es independiente del pre-tratamiento.

**Tabla 6:** Resultados del pre-procesamiento de muestras de intestino de camarón en porcentaje; se muestran el porcentaje resultantes en cada paso.

Muestra	Crudo	Corte extremo	JuntasR1 y R2	Ventanas 6:Q20	Sin N
IA1	100 %	100 %	52.16 %	52.08 %	52.06 %
IA2	100 %	100 %	80.13 %	80.07 %	80.04 %
IA3	100 %	100 %	80.06 %	80.01 %	79.99 %
IA4	100 %	100 %	40.86 %	40.76 %	40.75 %
IA5	100 %	100 %	81.89 %	81.84 %	81.82 %
IA6	100 %	100 %	11.96 %	11.92 %	11.91 %
IA7	100 %	100 %	80.43 %	80.38 %	80.36 %
IA8	100 %	100 %	82.87 %	82.84 %	82.81 %
IA9	100 %	100 %	70.99 %	70.93 %	70.90 %
IL1	100 %	100 %	52.49 %	52.45 %	52.43 %
IL2	100 %	100 %	59.38 %	59.29 %	59.25 %
IL3	100 %	100 %	19.51 %	19.49 %	19.47 %
IL4	100 %	100 %	83.97 %	83.94 %	83.91 %
IL6	100 %	100 %	43.76 %	43.72 %	43.70 %
IL7	100 %	100 %	16.65 %	16.61 %	16.60 %
IL8	100 %	100 %	77.59 %	77.55 %	77.53 %
IL9	100 %	100 %	42.14 %	42.10 %	42.01 %
IE1	100 %	100 %	37.41 %	37.41 %	37.41 %
IE2	100 %	100 %	71.45 %	71.39 %	71.37 %
IE3	100 %	100 %	41.60 %	41.55 %	41.54 %
IE5	100 %	100 %	75.90 %	75.84 %	71.82 %
IE6	100 %	100 %	9.65 %	9.57 %	9.57 %
IE7	100 %	100 %	72.43 %	72.37 %	72.34 %
IE8	100 %	100 %	39.96 %	39.87 %	39.86 %
IE9	100 %	100 %	66.55 %	66.47 %	66.44 %

## RESULTADOS

---

**Tabla 7:** Resultados del pre-procesamiento de muestras de intestino de camarón; se muestran el número de lecturas resultantes en cada paso.

Muestra	Crudo	Corte extremo	JuntasR1 y R2	Ventanas 6:Q20	Sin N
IA1	16388	16388	8548	8536	8533
IA2	221426	221426	177447	177316	177242
IA3	150685	150685	120647	120571	120541
IA4	38466	38466	15719	15679	15676
IA5	102231	102231	83720	83669	83646
IA6	35297	35297	4223	4210	4206
IA7	627531	627531	504737	504465	504305
IA8	42542	42542	35257	35243	35231
IA9	85007	85007	60354	60303	60272
IL1	213688	213688	112169	112091	112051
IL2	15787	15787	9375	9361	9355
IL3	106896	106896	20864	20837	20815
IL4	219766	219766	184558	184476	184424
IL6	223954	223954	98005	97913	97881
IL7	10134	10134	1688	1684	1683
IL8	46133	46133	35796	35777	35767
IL9	2311	2311	974	973	971
IE1	1871	1871	700	700	700
IE2	606582	606582	433437	433074	432944
IE3	69709	69709	29000	28969	28963
IE5	134170	134170	101846	101767	101738
IE6	22465	22465	2168	2152	2151
IE7	121026	121026	87666	87594	87557
IE8	56184	56184	22454	22403	22395
IE9	63686	63686	42386	42336	42318

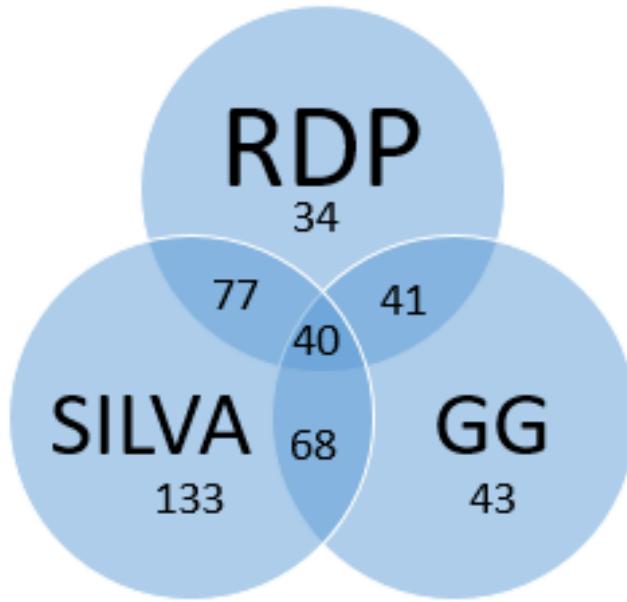
## RESULTADOS

---

### Comparación de las bases de datos para secuencias de 16S ARNr

**Tabla 8:** Comparación entre las asignaciones de QIIME con los valores por defecto para las tres bases de datos (GG: Greengenes; RDP: Ribosomal Database Project; y SILVA)

Base de datos	Greengenes	RDP	SILVA
Secs. asignadas	176356	214039	230707
Núm. OTU's	244	120	413
% Phylum	99.97	100	99.97
% Clase	99.94	100	99.90
% Orden	99.67	100	31.33
% Familia	52.42	99.71	30.84
% Género	21.07	30.35	30.45



**Figura 20:** Diagrama de Venn a nivel de género entre las tres bases de datos comparadas; (GG: Greengenes; RDP: Ribosomal Database Project; y SILVA).

Considerando los diferentes índices como la diversidad filogenética (PD), Chao1, Shannon y Simpson, se consideró que la base de datos que explicaba mejor la di-

## RESULTADOS

---

versidad de la muestra es Greengenes por lo cual se decidió considerarla para la asignación de taxonomía de todas las librerías (tabla 9).

**Tabla 9:** Comparación entre los índices de diversidad; se utilizó un tamaño de 150,110 secuencias para la comparación (GG: Greengenes; RDP: Ribosomal Database Project; y SILVA).

Base de datos	PD	Chao1	GC	OTU observados	Shannon	Simpson
Greengenes	44.967	1321.435	0.998	872.600	4.513	0.922
RDP	10.491	349.622	0.999	214.500	1.699	0.494
SILVA	46.693	1507.272	0.997	958.100	3.097	0.678

## Asignación de taxonomía en muestras de intenso con Greengenes

### Comparación a nivel phylum

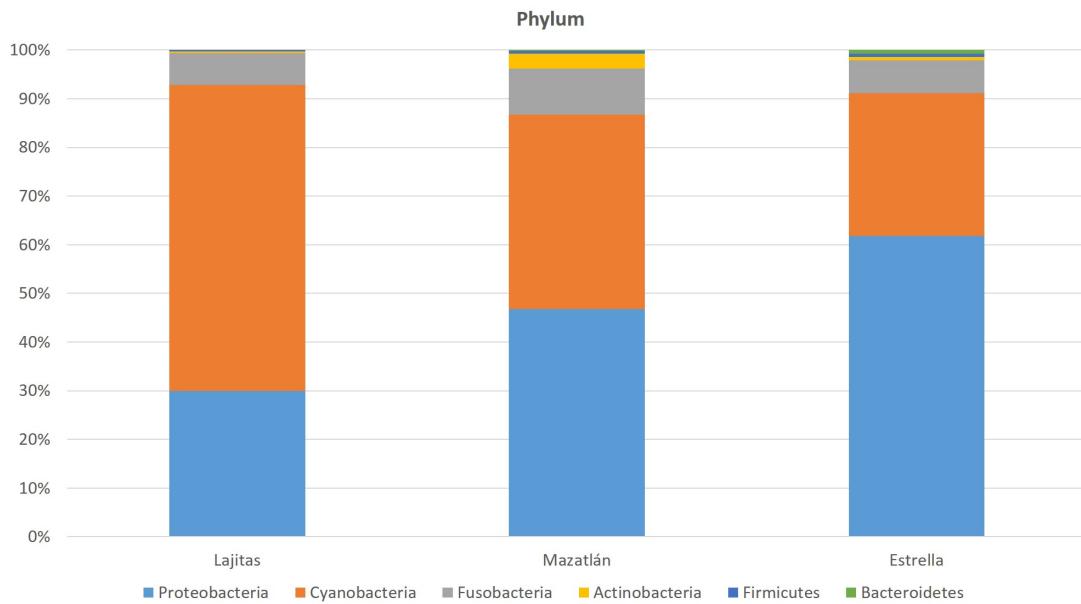
En la tabla 10 se pueden observar los phyla más abundantes para los tres estanques, entre ellos se encuentra Proteobacteria, Cyanobacteria, Fusobacteria, Actinobacteria y Firmicutes. Estos phyla conciden con los que se grafican con una abundancia mayor al 0.1 % (figura 21).

## RESULTADOS

---

**Tabla 10:** Tabla con los 10 OTU más abundantes a nivel de phylum para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque.

Phyla	Lajitas	Mazatlán	Estrella
Bacteria Proteobacteria	32.62	46.84	61.66
Bacteria Cyanobacteria	56.30	31.33	25.76
Bacteria Fusobacteria	6.22	7.07	6.07
Bacteria Actinobacteria	1.35	5.72	1.70
Bacteria Firmicutes	1.13	4.16	1.71
Bacteria Bacteroidetes	0.51	1.32	1.15
Bacteria Planctomycetes	0.55	1.40	0.82
Bacteria Chloroflexi	0.41	0.62	0.42
Bacteria Verrucomicrobia	0.18	0.41	0.28
Bacteria Acidobacteria	0.21	0.22	0.14



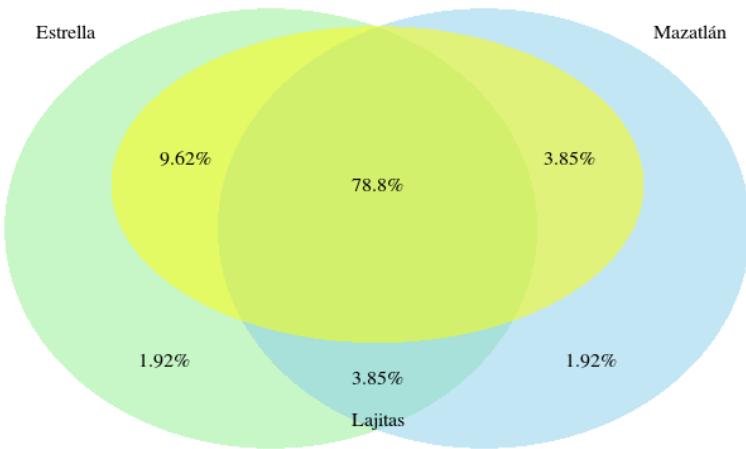
**Figura 21:** Gráfica que muestra los phyla con una abundancia mayor al 0.1 %.

Las muestras de microbiota intestinal del estanque Lajitas cuentan con una

## RESULTADOS

---

mayor proporción de Cyanobacteria (más del 50 %) y una menor cantidad de Cyanobacteria, en comparación con las muestras de los otros dos estanques. De forma contraria el estanque Estrella se puede observar una frecuencia relativa mayor de Proteobacteria, así como una disminución del phylum Cyanobacteria. Por último en el estanque Mazatlán se observa una proporción similar entre los phyla Cyanobacteria y Proteobacteria. De igual forma, se puede observar una mayor frecuencia relativa de Actinobacteria y Firmicutes únicamente en este estanque (tabla 10 y figura 21).

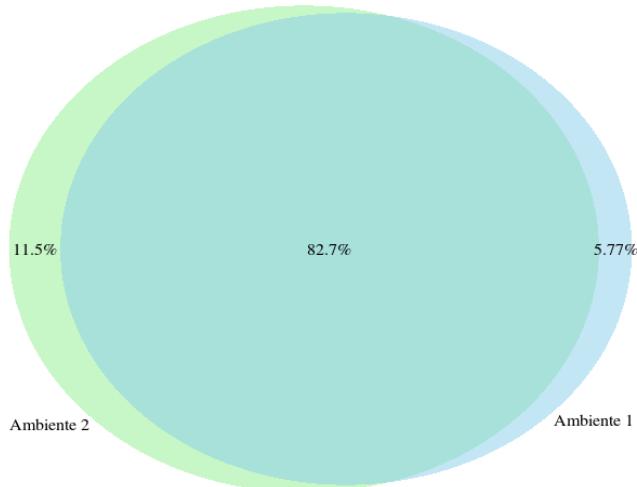


**Figura 22:** Diagrama de Venn que muestra los phyla compartidos entre los estanques Lajitas, Mazatlán y Estrella.

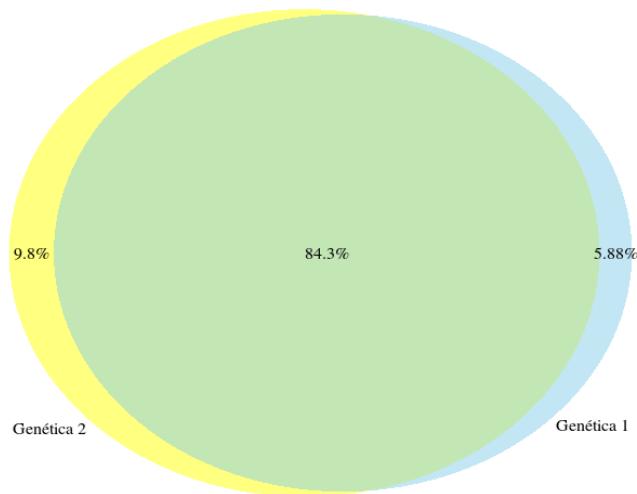
Se puede observar en las figuras 24 y 23, el número de phyla compartidos y no compartidos, así como su porcentaje son muy similares entre sí.

## RESULTADOS

---



**Figura 23:** Diagrama de Venn que muestra los phyla compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).



**Figura 24:** Diagrama de Venn que muestra los phyla compartidos entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).

## Comparación a nivel clase

En la tabla 11 se pueden observar los 10 OTU con mayor abundancia a nivel de clase para los tres estanques. En la figura 25 se observan las clases con una

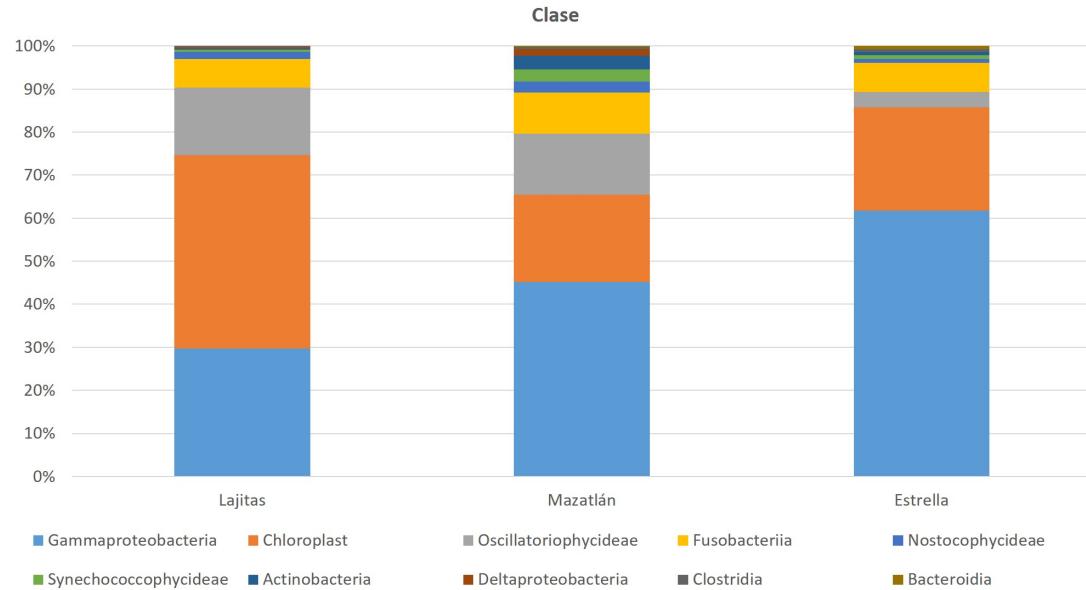
## RESULTADOS

---

abundancia mayor a 0.1 %.

**Tabla 11:** Tabla con los 10 OTU más abundantes a nivel de clase para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque.

Clases	Lajitas	Mazatlán	Estrella
Bacteria Proteobacteria Gammaproteobacteria	29.56	39.68	58.95
Bacteria Cyanobacteria Chloroplast	39.65	14.13	20.80
Bacteria Cyanobacteria Oscillatoriophycideae	13.73	12.17	3.07
Bacteria Fusobacteria Fusobacteriia	6.22	7.07	6.07
Bacteria Proteobacteria Alphaproteobacteria	2.24	4.61	2.08
Bacteria Actinobacteria Actinobacteria	0.58	3.99	1.03
Bacteria Cyanobacteria Synechococcophycideae	1.27	2.77	1.09
Bacteria Firmicutes Clostridia	0.85	3.08	1.19
Bacteria Cyanobacteria Nostocophycideae	1.53	2.10	0.71
Bacteria Proteobacteria Deltaproteobacteria	0.78	2.18	0.55



**Figura 25:** Gráfica que muestra las clases con una abundancia mayor al 0.1 %.

## RESULTADOS

---

Las clase compartida más abundante fue Gammaproteobacteria con un promedio de 42.73 % en su frecuencia absoluta (tabla 11). Esta clase es muy diversa e incluye a familias comunes como Enterobacteriaceae, Psudomonaceae, Vibronaceae, Aeromonaceae, entre otras. Más específicamente se puede observar que las muestras del estanque Lajitas, en comparación con los otros dos estanques, cuentan con un aumento en la proporción de una clase de cyanobacterias (cerca del 40 % de la frecuencia relativa), además de que se observa una disminución de Gammaproteobacteria (figura 25).

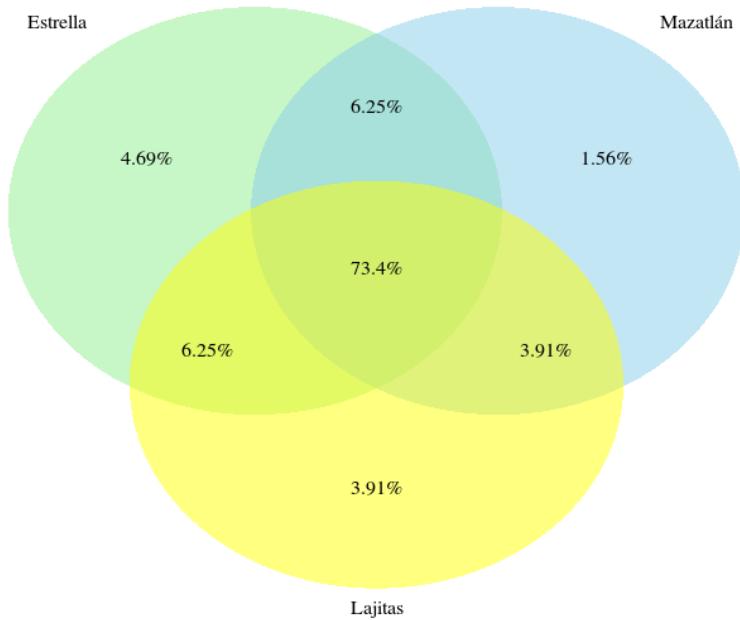
Se puede observar que hay un enriquecimiento de la clase Gammaproteobacteria en el estanque Estrella, que en comparación con el estanque Lajitas es casi del doble y que cuenta con una frecuencia relativa de cerca del 60 % (tabla 11). En este estanque se puede observar también una disminución en la proporción de la subclase Oscillatoriophycidaea (figura 25).

Por último en el estanque Mazatlán se puede observar frecuencias relativas más diversas y no tan dominantes. Cuenta con una proporción de un poco más de 40 % de Gammaproteobacteria, así como frecuencias similares entre sí de Chloroplast, Oscillatoriophycidaea (tabla 11). Se nota además un aumento en la proporción de Alphaproteobacteria y Actinobacteria en comparación a los otros dos grupos (figura 25).

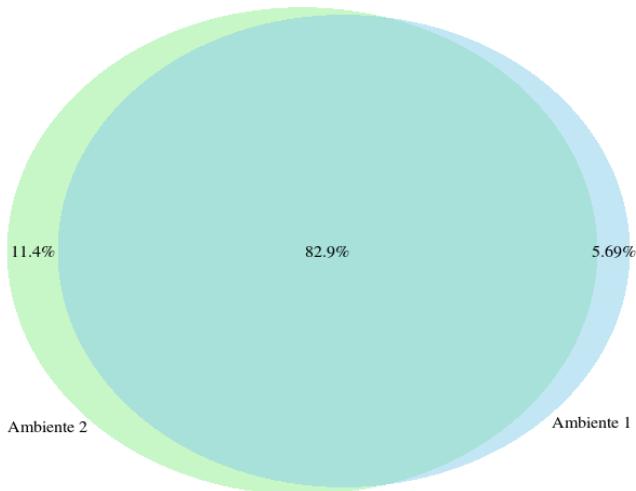
En la figura 26 se muestra el diagrama de Venn con las clases compartidas entre los estanques, que equivalen a 74.4 %. Al igual que a nivel de phylum, no se puede determinar una influencia por parte del ambiente o la genética considerando los gráficos de Venn que se realizaron a nivel de clase (figuras 27 y 28), ya que tanto el número de clases compartidas y no compartidas son similares entre sí.

## RESULTADOS

---



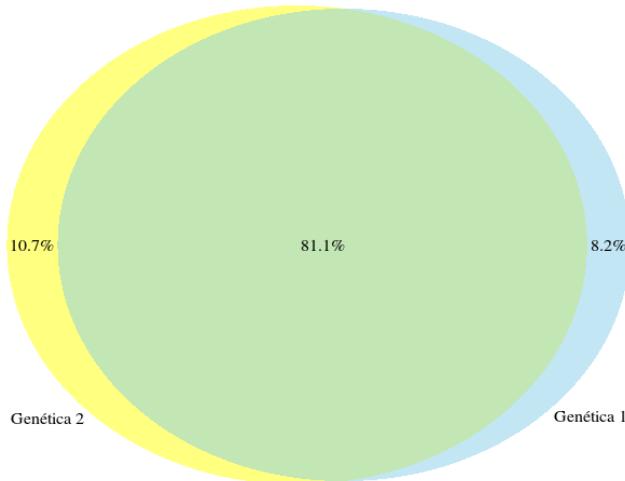
**Figura 26:** Diagrama de Venn que muestra las clases compartidas entre los estanques Lajitas, Mazatlán y Estrella.



**Figura 27:** Diagrama de Venn que muestra las clases compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).

## RESULTADOS

---



**Figura 28:** Diagrama de Venn que muestra las clases compartidas entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).

### Comparación a nivel orden

En la tabla 12 se muestran los OTU con mayor frecuencia absoluta para los tres estanques a nivel de orden. Se pueden observar también los órdenes con una abundancia mayor al 0.1 % en la figura 29.

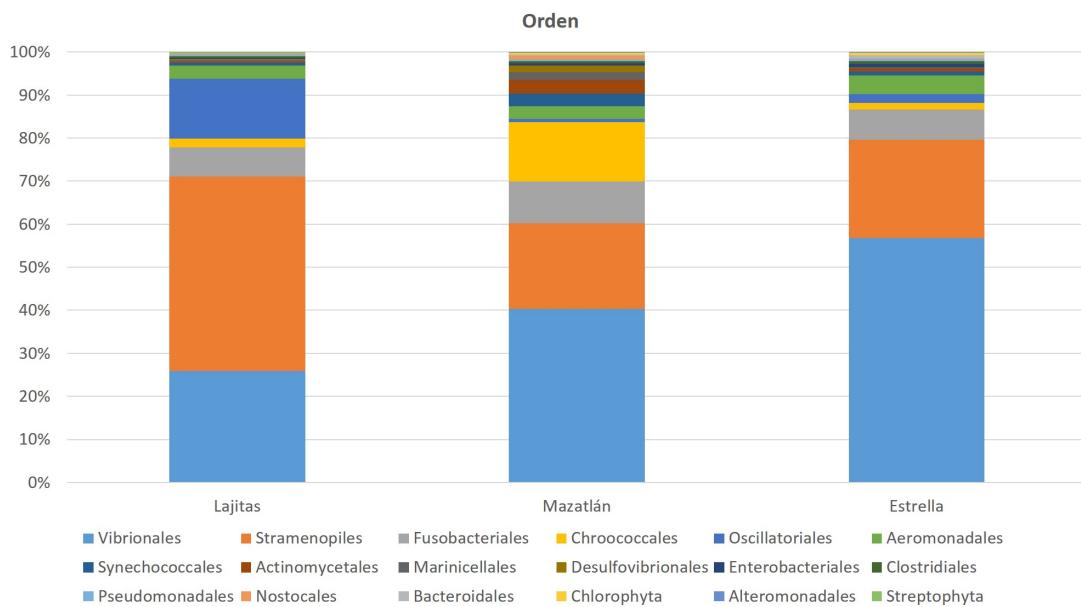
**Tabla 12:** Tabla con los 10 OTU más abundantes a nivel de orden para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque.

Órdenes	Lajitas	Mazatlán	Estrella
Bacteria Proteobacteria Gammaproteobacteria Vibrionales	24.06	31.43	50.72
Bacteria Cyanobacteria Chloroplast Stramenopiles	39.17	13.50	19.42
Bacteria Fusobacteria Fusobacteriia Fusobacteriales	6.22	7.07	6.07
Bacteria Cyanobacteria Oscillatoriophycideae Oscillatoriales	11.79	0.63	1.58
Bacteria Cyanobacteria Oscillatoriophycideae Chroococcales	1.94	11.55	1.49
Bacteria Proteobacteria Gammaproteobacteria Aeromonadales	2.80	2.48	3.88
Bacteria Proteobacteria Alphaproteobacteria Rhodobacterales	1.23	3.27	1.65
Bacteria Actinobacteria Actinobacteria Actinomycetales	0.58	3.98	1.03
Bacteria Firmicutes Clostridia Clostridiales	0.81	3.00	1.17
Bacteria Proteobacteria Gammaproteobacteria Alteromonadales	0.57	1.55	1.32

## RESULTADOS

---

En comparación con las gráficas de la asignación taxonómicas vistas en las secciones anteriores a nivel orden se puede observar de bandeo más variado a pesar de que se evalúan los mismos órdenes (figura 29) lo cual implica una mayor variabilidad entre la diversidad de los estanques a este nivel. En todos los estanques existe la presencia de dos órdenes con abundancia importante: Vibrionales y Stramenopiles (tabla 12 y figura 29).



**Figura 29:** Gráfica que muestra los órdenes con una abundancia mayor al 0.1 %.

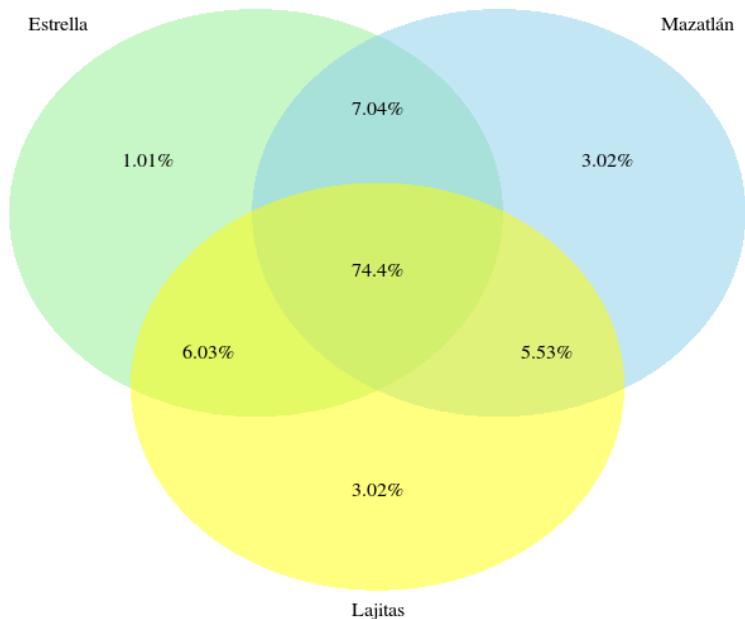
En el estanque Lajitas se puede observar un aumento en la proporción de Stramenopiles y Oscillatoriales en comparación a los otros dos estanques, así como una disminución en el orden Vibrionales (figura 29), lo cual concuerda con lo observado en los OTU más abundantes a este nivel (tabla 12).

Para las muestras del estanque Mazatlán se observa una composición más heterogénea en la cual sobresale el aumento Chroococcales, Actinomycetales, Clostridiales. Hay casi un 40 % de Vibrionales, que resulta más que la proporción del estanque Lajitas pero menos a la abundancia de las muestras del estanque Estrella (tabla 12 y figura 29).

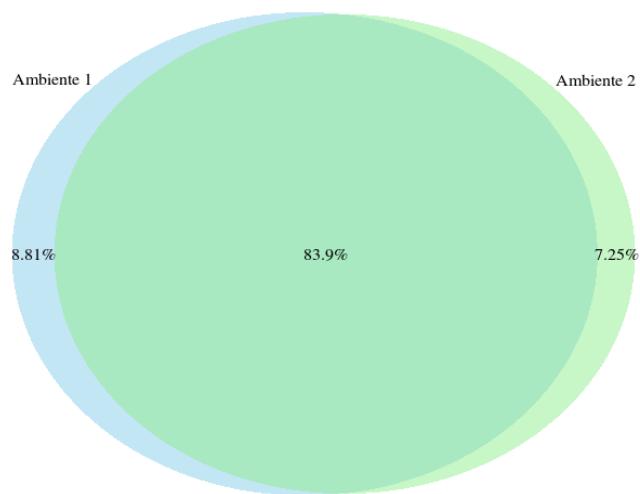
En la figura 30 se observan los órdenes compartidos entre los estanques Lajitas, Mazatlán y Estrella, de que comparten 248 taxones (74.4 %). Los diagramas de Venn difieren poco en número y porcentaje para las dos comparaciones realizadas (figuras 31 y 32) por lo cual se puede inferir poco acerca de la influencia de ambas variables con estos gráficos.

## RESULTADOS

---



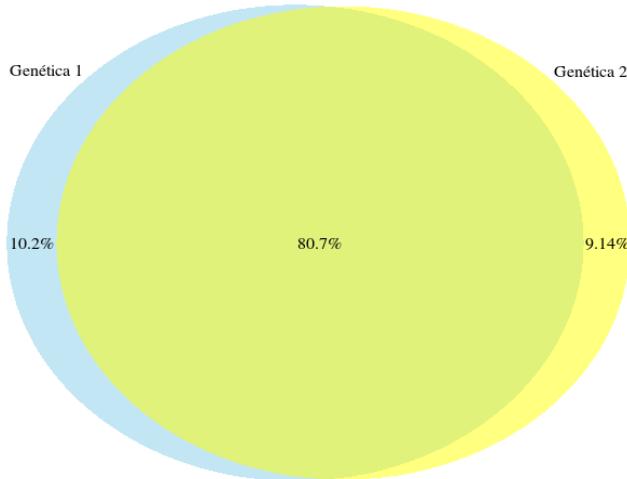
**Figura 30:** Diagrama de Venn que muestra los órdenes compartidos entre los estanques Lajitas, Mazatlán y Estrella.



**Figura 31:** Diagrama de Venn que muestra los órdenes compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).

## RESULTADOS

---



**Figura 32:** Diagrama de Venn que muestra los órdenes compartidos entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).

### Comparación a nivel familia

Se pueden observar los OTU a nivel familia con mayor frecuencia absoluta para los tres estanques en la tabla 13, sin embargo no todos los OTU fueron asignados completamente, por lo cual se coloca un signo de interrogación en el nivel faltante. Lo que podemos concluir de esta tabla es que una familia del orden Stramenopiles tiene el promedio más alto en abundancia para los tres estanques, seguida de las familias Pseudoalteromonadaceae y Vibrionaceae.

Las familias con una abundancia mayor a 0.1 % (que fueron clasificadas completamente) según el promedio de todos los grupos son: Pseudolateromonadace, Vibrionaceae, Fusobacteriaceae y Phormidiaceae (figura 33).

El estanque Lajitas está dominado por Pseudoalteromonadaceae seguido de Phormidiaceae que, a comparación con los otros estanques, cuenta con un enriquecimiento mayor. Posteriormente le siguen Vibrionaceae y Fusobacteriaceae. Para el estanque Mazatlán se pueden observar tres grupos con una frecuencia relativa similar: Pseudolateromonadace, Vibrionaceae y Cyanobacteriaceae; son relevantes también las familias Cyanobacteriaceae, Rhodobacteraceae Microbacteriaceae que aparece con una abundancia mayor a comparación con los otros estanques. Vibrionaceae y Pseudoalteromonadaceae son los grupos dominantes en el estanque Estrella, que juntos suman el 80 % de su frecuencia relativa. Es importante notar que en este estanque se observa una abundancia superior de Vibrionaceae (tabla 13 y figura 33).

## RESULTADOS

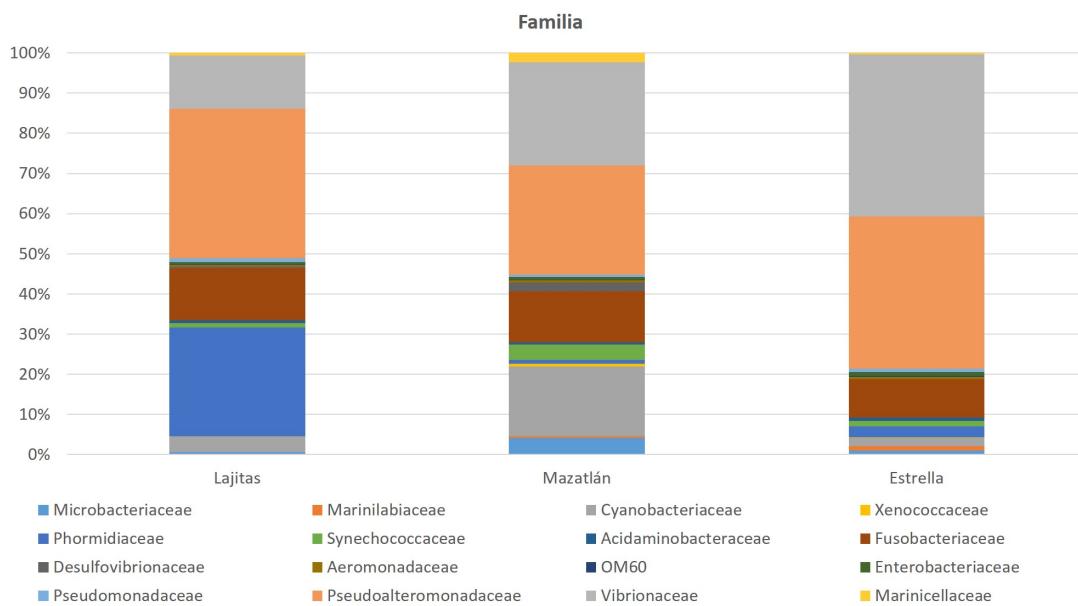
---

**Tabla 13:** Tabla con los 10 OTU más abundantes a nivel de familia para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. Los OTU que no obtuvieron una clasificación completa se colocan con '?' (orden: o? y familia: f?).

Familias	Lajitas	Mazatlán	Estrella
Bacteria   Cyanobacteria   Chloroplast   Stramenopiles   f?	39.17	13.50	19.42
Bacteria   Proteobacteria   Gammaproteobacteria   Vibrionales   Pseudoalteromonadaceae	17.54	17.25	23.98
Bacteria   Proteobacteria   Gammaproteobacteria   Vibrionales   Vibrionaceae	6.52	14.18	26.73
Bacteria   Fusobacteria   Fusobacteriia   Fusobacteriales   Fusobacteriaceae	6.16	6.86	5.79
Bacteria   Cyanobacteria   Oscillatoriophycideae   Oscillatoriales   Phormidiaceae	11.79	0.62	1.58
Bacteria   Cyanobacteria   Oscillatoriophycideae   Chroococcales   Cyanobacteriaceae	1.69	10.60	1.29
Bacteria   Proteobacteria   Gammaproteobacteria   Aeromonadales   f?	2.60	2.09	3.40
Bacteria   Proteobacteria   Alphaproteobacteria   Rhodobacterales   Rhodobacteraceae	1.23	3.24	1.64
Bacteria   Actinobacteria   Actinobacteria   Actinomycetales   Microbacteriaceae	0.33	3.18	0.83
Bacteria   Cyanobacteria   Nostocophycideae   o?   f?	1.22	1.23	0.58

m

**Figura 33:** Gráfica que muestra las familias con una abundancia mayor al 0.1 %.

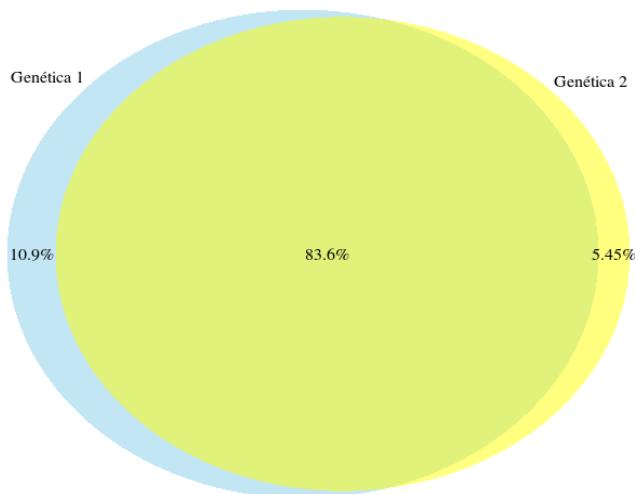
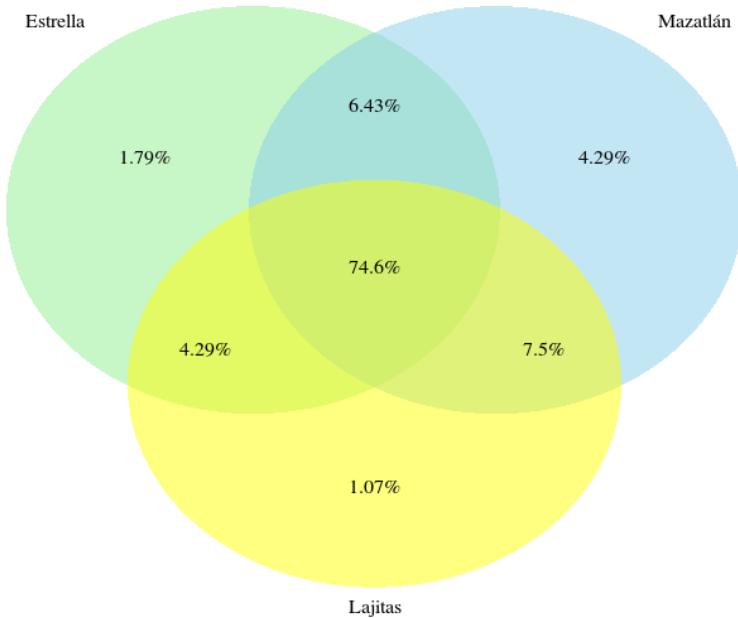


Entre los estanques se comparten 209 familias (el 74.6 %) y el estanque Mazatlán es el que cuenta con un mayor número de familias no compartidas (4.3%) (figura 34).

## RESULTADOS

---

**Figura 34:** Diagrama de Venn que muestra las familias compartidas entre los estanques Lajitas, Mazatlán y Estrella.



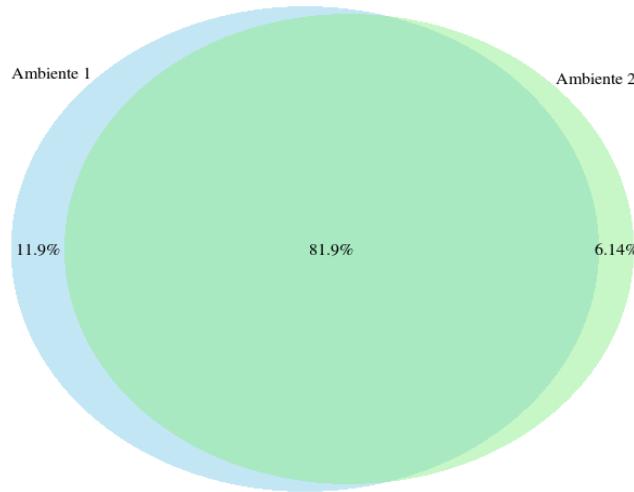
**Figura 35:** Diagrama de Venn que muestra las familias compartidas entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).

Se puede inferir poco acerca de la influencia del ambiente y de la genética con los diagramas de Venn realizados (figuras 36 y 35), ya que entre los análisis difiere

## RESULTADOS

---

poco en número de familias compartidas e individuales, así como su porcentaje.



**Figura 36:** Diagrama de Venn que muestra las familias compartidas entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).

## Comparación a nivel género

En la tabla 14 se muestran los OTU más abundantes a nivel de género. La mayoría de estos OTU no cuentan con una clasificación completa. El OTU con frecuencia promedio más alta es un género perteneciente a la orden Stramenopiles, posteriormente le sigue un género de la familia Pseudoalteromonadaceae. El OTU clasificado completamente con mayor abundancia a nivel de género es *Vibrio*, con frecuencias relativas del 4 % al 20 % que varía entre los estanques (tabla 14).

El género *Vibrio* es uno de los más abundantes para todos los grupos, en especial en el estanque Estrella donde alcanza más del 50 % de la frecuencia relativa a este nivel; en los estanques Lajitas y Mazatlán se encuentra cerca del 30 %. El siguiente más abundante es *Propionigenium*, que sobresale en el estanque Lajitas con más del 30 %. *Cyanobacterium* está enriquecido en el estanque Mazatlán, pero tiene el tercer lugar en abundancia para los tres estanques. Por último el género *Photobacterium* en el estanque Estrella muestra una mayor abundancia (figura 37).

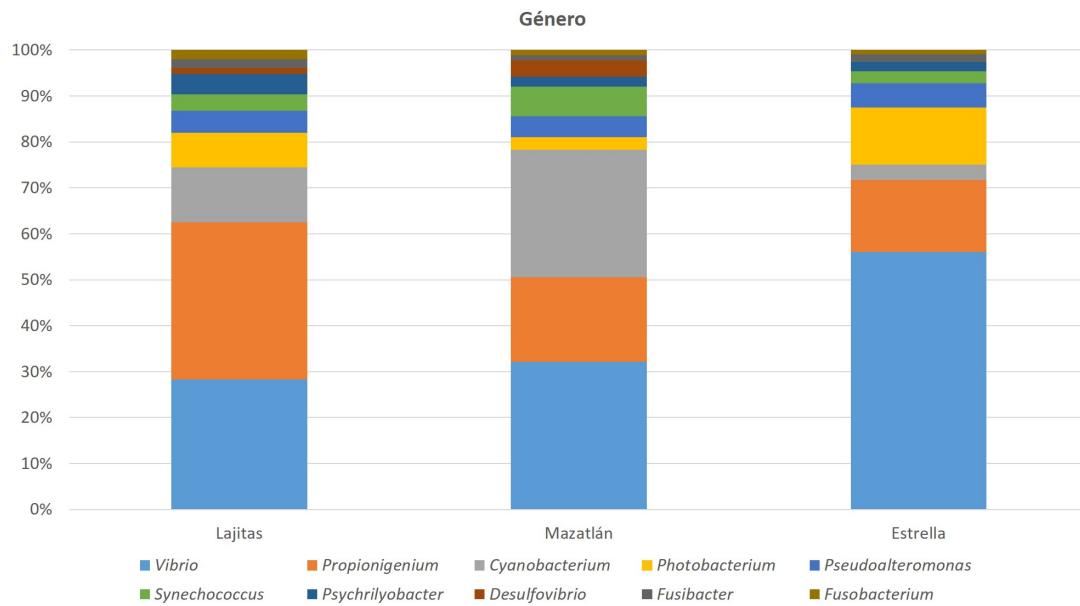
## RESULTADOS

---

**Tabla 14:** Tabla con los 10 OTU más abundantes a nivel de género para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. Los OTU que no obtuvieron una clasificación completa se colocan con '?' (familia: f? y género: g?).

Géneros	Lajitas	Mazatlán	Estrella
Bacteria Cyanobacteria Chloroplast Stramenopiles f? g?	39.17	13.50	19.42
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Pseudoalteromonadaceae g?	16.69	15.55	22.13
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Vibrionaceae Vibrio	4.60	10.61	19.59
Bacteria Cyanobacteria Oscillatoriophycideae Oscillariales Phormidiaceae g?	11.29	0.55	1.56
Bacteria Fusobacteria Fusobacteriia Fusobacteriales Fusobacteriaceae Propionigenium	5.21	5.82	4.88
Bacteria Cyanobacteria Oscillatoriophycideae Chroococcales Cyanobacteriaceae Cyanobacterium	1.59	9.84	1.08
Bacteria Proteobacteria Gammaproteobacteria Aeromonadales f? g?	2.60	2.09	3.40
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Vibrionaceae Photobacterium	1.09	0.85	4.26
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Vibrionaceae g?	0.79	2.67	2.87
Bacteria Proteobacteria Alphaproteobacteria Rhodobacterales Rhodobacteraceae g?	1.00	2.60	1.31

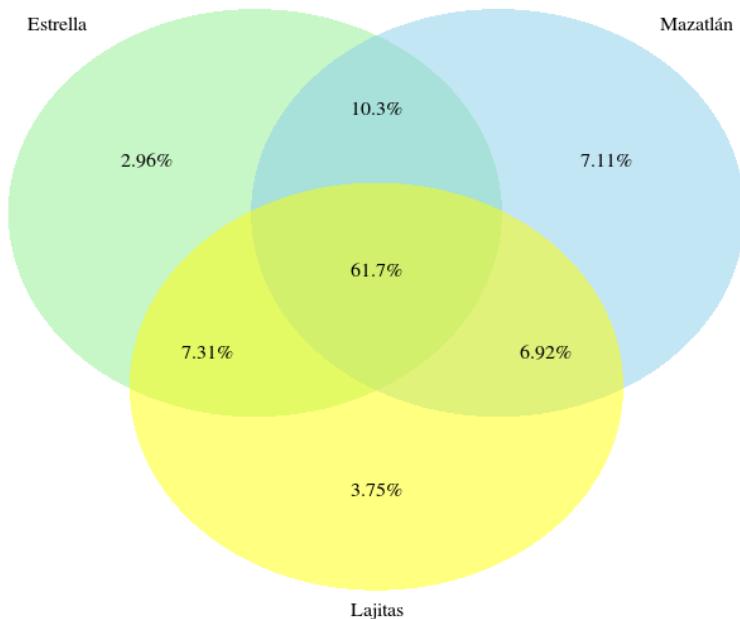
**Figura 37:** Gráfica que muestra los géneros con una abundancia mayor al 0.1 %.



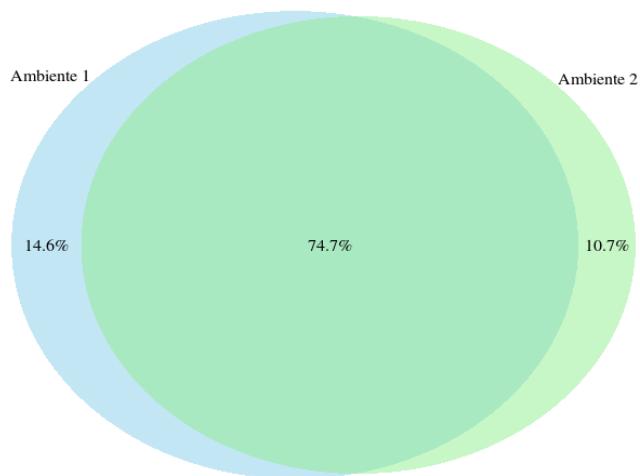
En la figura 38 se puede observar que a comparación con los niveles taxonómicos anteriores el porcentaje de géneros compartidos es menor, 61.7 %, que equivale a 312 géneros. Los diagramas de Venn comparando a la genética y el ambiente se muestran en las figuras 40 y 39, cuyos porcentajes de similitud son de 70.7 % y 74.6 %, respectivamente.

## RESULTADOS

---



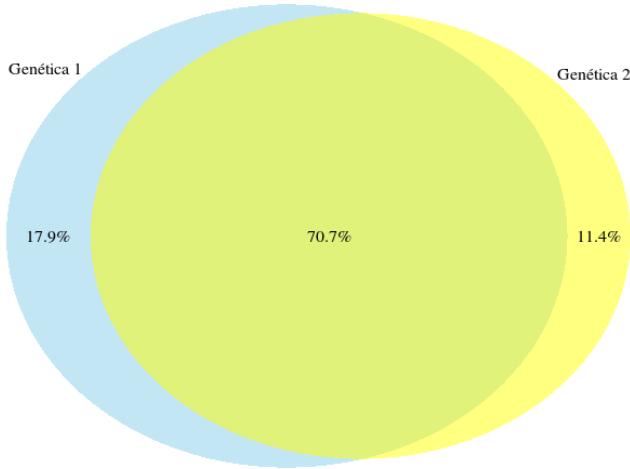
**Figura 38:** Diagrama de Venn que muestra los géneros compartidos entre los estanques Lajitas, Mazatlán y Estrella.



**Figura 39:** Diagrama de Venn que muestra los géneros compartidos entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).

## RESULTADOS

---



**Figura 40:** Diagrama de Venn que muestra los géneros compartidos entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).

### Comparación a nivel especie

Los 10 OTU con una abundancia mayor a nivel de especie se muestran en la tabla 15, sin embargo no todos los OTU cuentan con una asignación completa hasta el nombre de la especie.

**Tabla 15:** Tabla con los 10 OTU más abundantes a nivel de especie para los tres estanques; se muestra la frecuencia absoluta promedio (en porcentaje) por estanque. Los OTU que no obtuvieron una clasificación completa se colocan con '?' (familia: f?, género: g? y especie: s?).

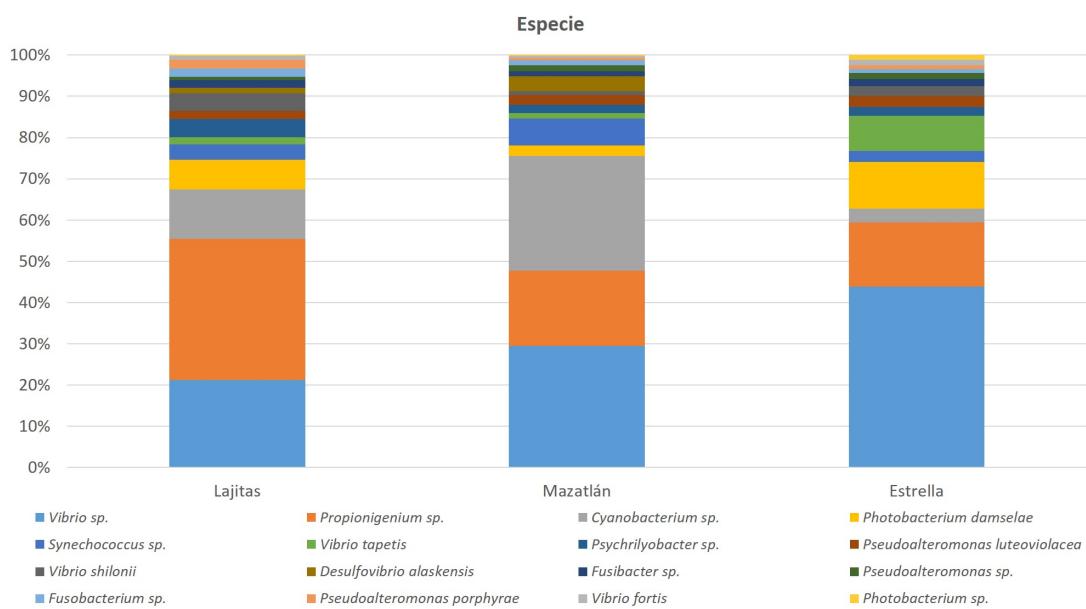
Species	Lajitas	Mazatlán	Estrella
Bacteria Cyanobacteria Chloroplast Stramenopiles f? g? s?	39.17	13.50	19.42
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Pseudoalteromonadaceae g? s?	16.69	15.55	22.13
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Vibrionaceae Vibrio s?	3.68	9.72	15.59
Bacteria Cyanobacteria Oscillatoriophycideae Oscillatoriiales Phormidiaceae g? s?	11.29	0.55	1.56
Bacteria Fusobacteria Fusobacteriia Fusobacteriales Fusobacteriaceae Propionigenium s?	5.21	5.82	4.88
Bacteria Cyanobacteria Oscillatoriophycideae Chroococcales Cyanobacteriaceae Cyanobacterium s?	1.59	9.84	1.08
Bacteria Proteobacteria Gammaproteobacteria Aeromonadales f? g? s?	2.60	2.09	3.40
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Vibrionaceae g? s?	0.79	2.67	2.87
Bacteria Proteobacteria Gammaproteobacteria Vibrionales Vibrionaceae Photobacterium damselae	1.03	0.72	3.74
Bacteria Proteobacteria Alphaproteobacteria Rhodobacterales Rhodobacteraceae g? s?	1.00	2.60	1.31

Como se puede observar en la figura 41 las frecuencias relativas de las especies encontradas en los diferentes estanques son muy variables. Aunque no se

## RESULTADOS

---

asigna completamente el nombre de la especie, se puede saber que una especie del género *Vibrio* es una de las que cuenta con más del 40 % de la frecuencia relativa del estanque Estrella, que al contrario del estanque Lajitas se ve disminuida a cerca del 20 %. Esto se observa a la inversa con *Propionigenium sp.* que se muestra abundante en el estanque Lajitas pero no el estanque Estrella. Para las muestras del estanque Mazatlán se obtienen proporciones similares de *Vibrio sp.* y de *Propionigenium sp.*, pero además se muestra un aumento significativo en *Cyanobacterium sp.* (figura 41).



**Figura 41:** Gráfica que muestra las especies con una abundancia mayor al 0.1 %.

En promedio la especie más abundante es *Photobacterium damsela* (que también se muestra en la tabla 15) seguida de *Vibrio tapetis*, sin embargo estas abundancias varían mucho entre los estanques (figura 41).

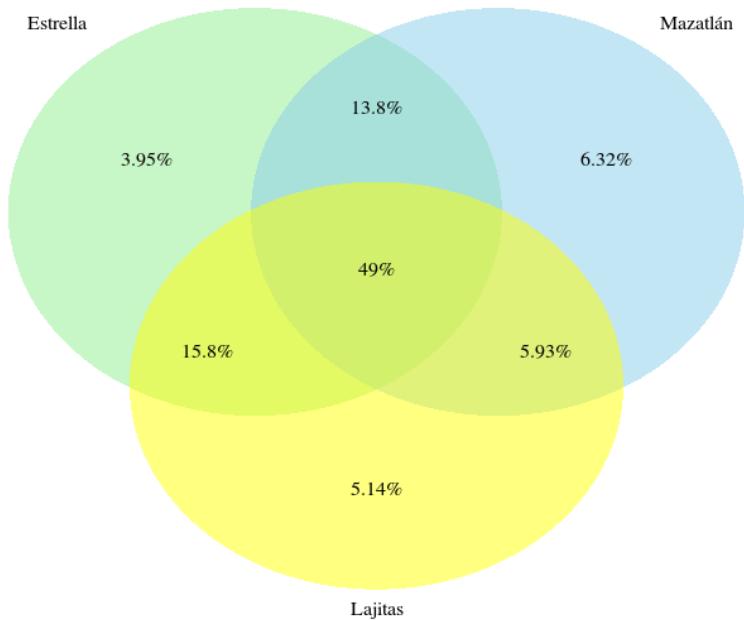
Para el estanque Lajitas tenemos que las especies más abundantes son *Photobacterium damsela* y *Vibrio shilonii*. En el estanque Mazatlán, que resultó el más variado, tenemos que la especie más frecuente es *Desulfovibrio alaskensis*, seguida de *Pseudoalteromonas luteoviolacea* y *Photobacterium damsela* con porcentajes similares; además de un Enriquecimiento notorio de *Clostridium perfringens*. El estanque Estrella tiene proporciones cercanas entre *Photobacterium damsela* y *Vibrio tapetis*, que suman aproximadamente el 70 % de la abundancia.

Es importante resaltar que la frecuencia absoluta de las lecturas asignadas a nivel especie para cada estanque son muy pocas (cerca del 5 % en promedio), pero

## RESULTADOS

---

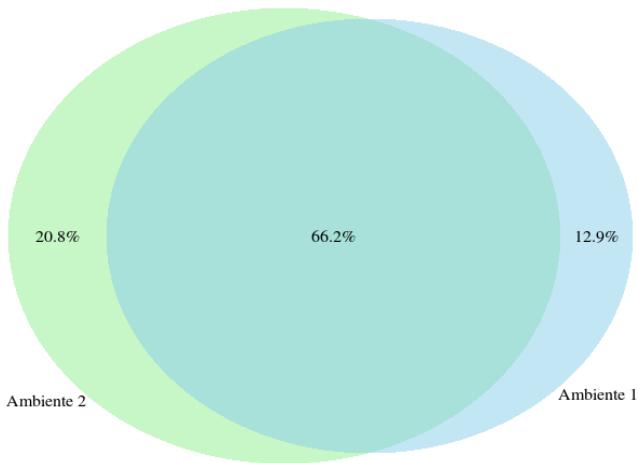
que pueden ser significativas al momento de interpretar la diversidad microbiana. En la figura 42 se muestra el diagrama de Venn entre las especies compartidas entre los estanques, 124 especies que equivalen al 49 %. Las comparaciones del ambiente y la genética se muestran en las figuras 43 y 44.



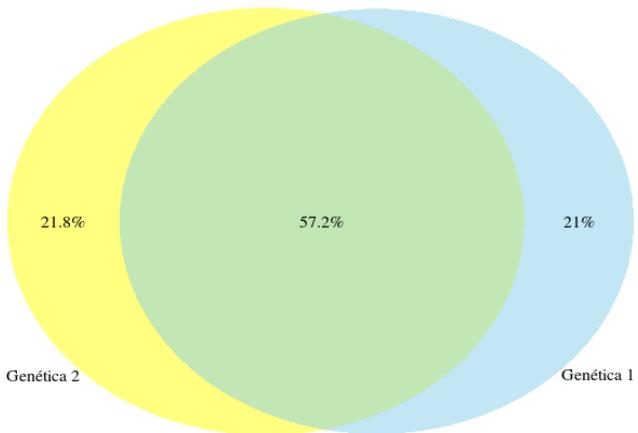
**Figura 42:** Diagrama de Venn que muestra las especies compartidas entre los estanques Lajitas, Mazatlán y Estrella.

## RESULTADOS

---



**Figura 43:** Diagrama de Venn que muestra las especies compartidas entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella).



**Figura 44:** Diagrama de Venn que muestra las especies compartidas entre la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas).

## RESULTADOS

---

### Análisis LEfSe

En la tabla 16 se muestran el número de taxones con diferencias significativas con LEfSe, donde se puede observar que cuando se utilizan los grupos con abundancias mayores a 0.1 % se encuentran un menor número de taxones diferenciales que con utilizando todos los OTU, es por esto que a continuación se muestran solamente estos resultados.

**Tabla 16:** Tabla que muestra las diferencias encontradas con LEfSe en los distintos niveles taxonómicos, para las dos comparaciones: con todos los OTU y utilizando solamente los OTU con una abundancia mayor al 0.1 %.

	Todos los OTU		
	Genética (M vs L)	Genética 2 (E vs L)	Ambiente (M vs E)
Phylum	4	3	2
Clase	12	7	5
Orden	20	12	8
Familia	14	21	13
Género	26	21	13
Especie	32	30	12

Primero se muestra el análisis realizado comparando la genética 1 y 2, estanques Mazatlán y Lajitas, respectivamente. Cada estanque obtiene a las larvas de camarón de una fuente de desove diferente, por lo cual se considera que pertenecen a linajes diferentes.

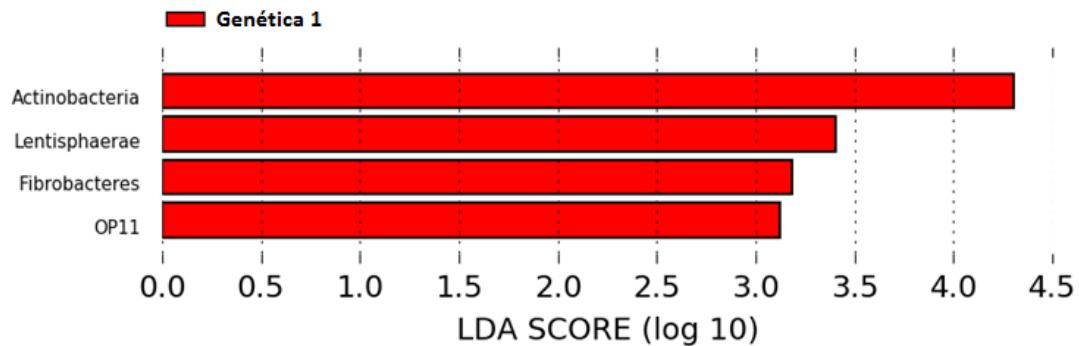
### Pruebas de LEfSe por genética

#### Análisis con LEfSe a nivel phylum

En la prueba realizada a nivel phylum se obtuvieron cuatro phyla con diferencias significativas para la genética 1 (Actinobacteria, Lentisphaerae, Fibrobacteres y OP11), sin embargo no se encontraron diferencias significativas para los phyla de la genética 2 (figura 45).

## RESULTADOS

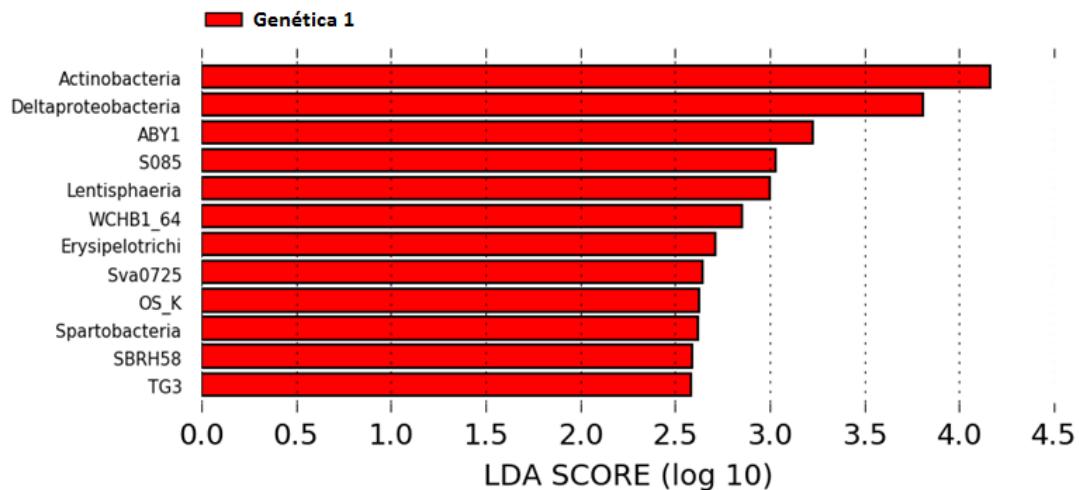
---



**Figura 45:** Análisis con LEfSe a nivel de phylum, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran los phyla que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde).

### Análisis con LEfSe a nivel clase

Al igual que el análisis realizado a nivel phylum, solamente se encontraron clases con diferencias significativas en la genética 1. En total son 12 clases: Actinobacteria, Deltaproteobacteria, ABY1, S085, Lentisphaeria, WCHB1\_64, Erysipelotrichi, Sva0725, OS\_K, Spartobacteria, SBRH58 y TG3 (figura 46).



**Figura 46:** Análisis con LEfSe a nivel de clase, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran las clases que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde).

## RESULTADOS

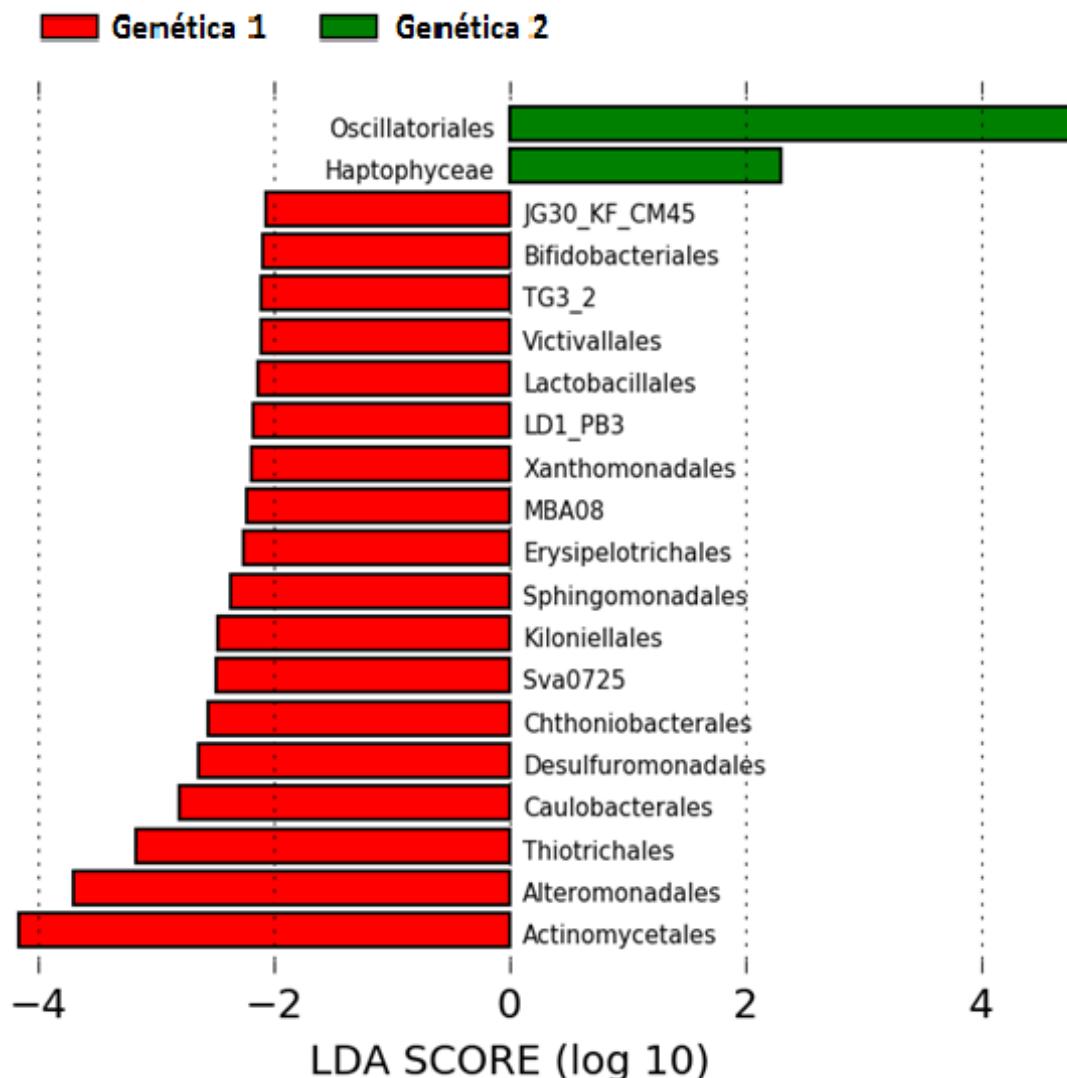
---

### Análisis con LEfSe a nivel orden

Para la genética 1 se encontraron 18 órdenes con diferencias significativas en comparación la genética 2; éstos son: JG30\_KF\_CM45, Bifidobacteriales, TG3\_32, Victivallales, Lactobacillales, LD1\_PB3, Xanthomonadales, MBA08, Erysipelotrichales, Sphingomonadales, Kiloniellales, Sva0725, Chthoniobacteriales, Desulfovimonadales, Caulobacteriales, Thiotrichales, Alteromonadales y Actinomycetales. En la genética 2 se encontraron dos órdenes con abundancias más altas y diferencialmente significativas, Oscillatoriales y Haptophyceae (fig 47).

## RESULTADOS

---



**Figura 47:** Análisis con LEfSe a nivel de orden, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran los órdenes que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde).

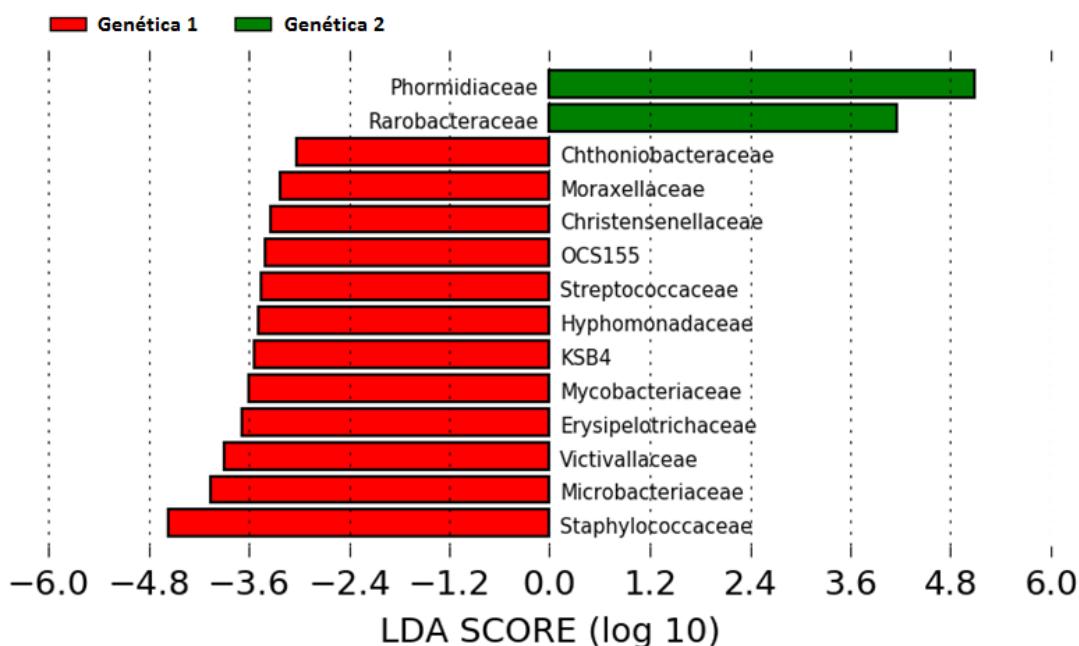
### Análisis con LEfSe a nivel familia

El análisis con LEfSe a nivel de familia obtuvo 12 familias con diferencias significativas (Chthoniobacteraceae, Moraxellaceae, Christensenellaceae, OCS155, Streptococcaceae, Hyphomonadaceae, KSB4, Mycobacteriaceae, Erysipelotrichaceae, Victivallaceae, Microbacteriaceae y Staphylococcaceae), mientras que para

## RESULTADOS

---

la genética 2 se obtuvieron solamente dos (Phormidiaceae y Rarobacteraceae).



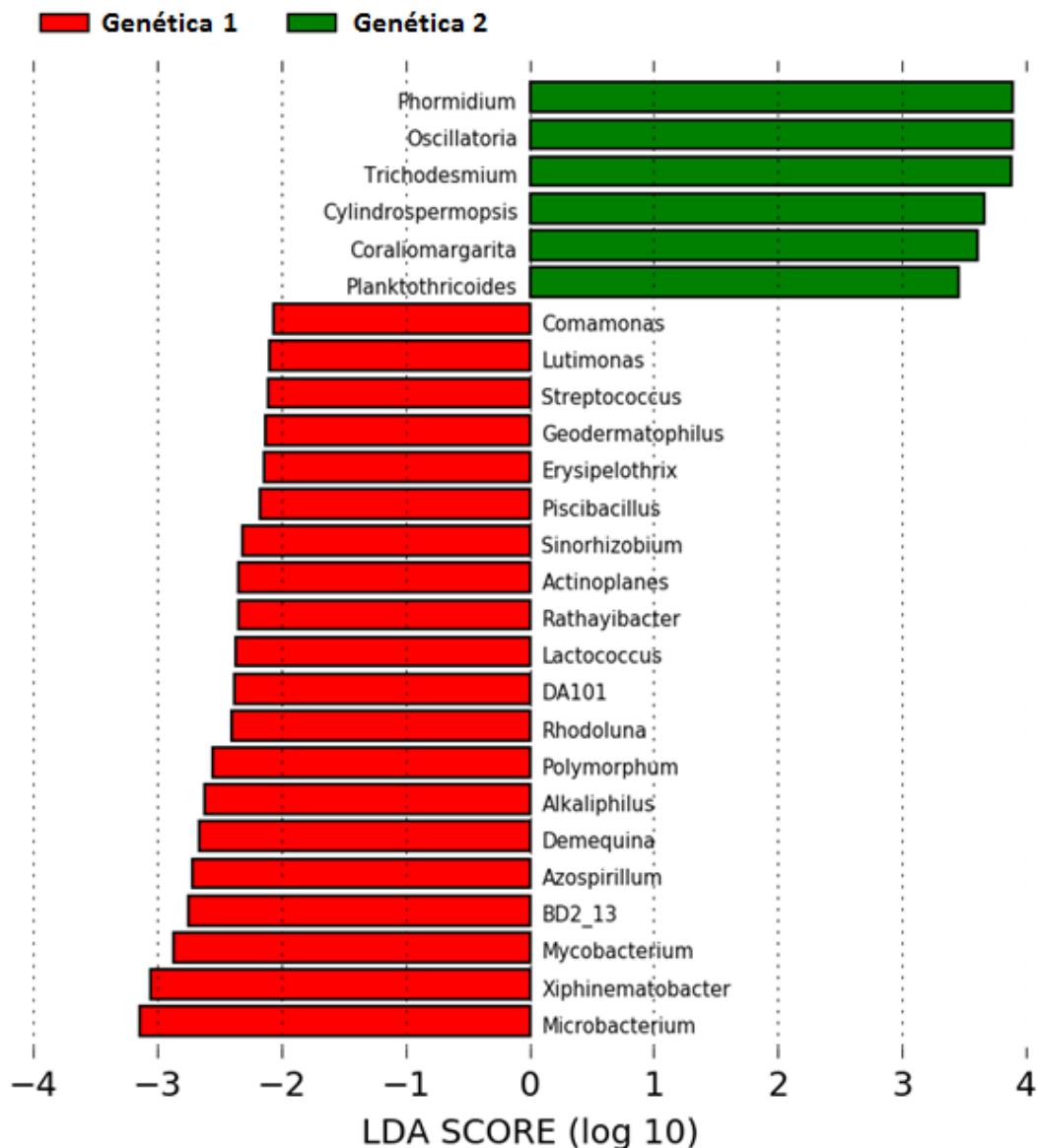
**Figura 48:** Análisis con LEfSe a nivel de familia, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran las familias que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde).

### Análisis con LEfSe a nivel género

Para la genética 1 se obtuvieron 20 géneros con diferencias significativas, éstos son: *Comamonas*, *Lutimonas*, *Streptococcus*, *Geodermatophilus*, *Erysipelothrix*, *Piscibacillus*, *Sinorhizobium*, *Actinoplanes*, *Rathayibacter*, *Lactococcus*, DA101, *Rhodoluna*, *Polymorphum*, *Alkaliphilus*, *Demequina*, *Azospirillum*, BD2\_13, *Mycobacterium*, *Xiphinema* y *Microbacterium*. En la genética 2 se encontraron seis géneros con abundancias diferencialmente significativas, éstos son: *Phormidium*, *Oscillatoria*, *Trichodesmium*, *Cylindrospermopsis*, *Coraliomargariata* y *Planktothricoides*.

## RESULTADOS

---



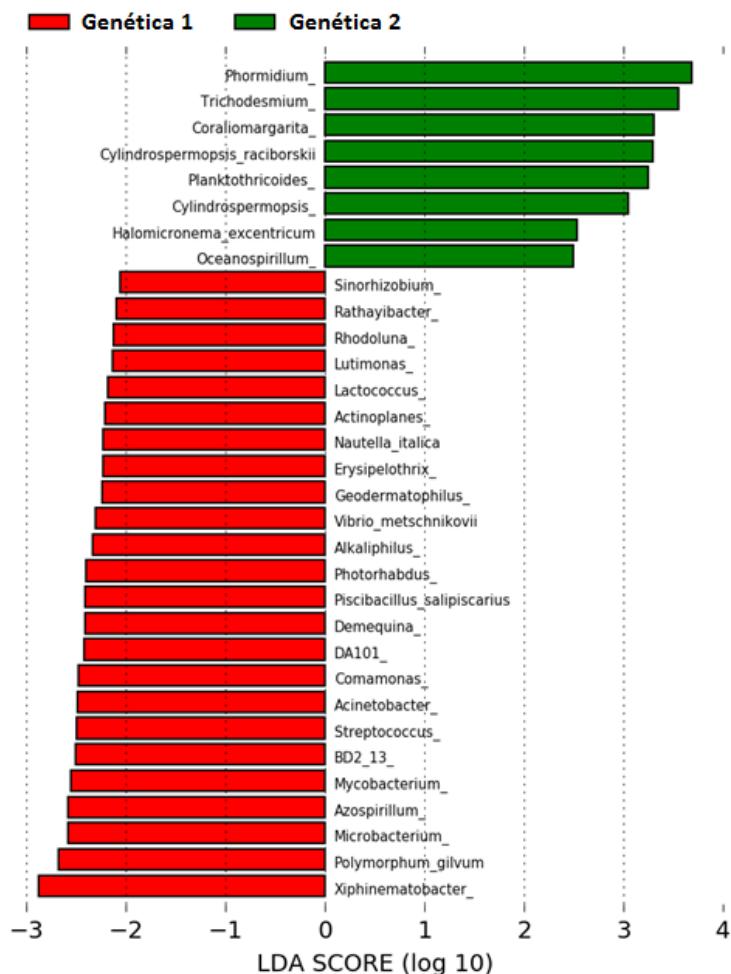
**Figura 49:** Análisis con LEfSe a nivel de género, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran los géneros que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde).

## RESULTADOS

---

### Análisis con LEfSe a nivel especie

Para la genética 1 se encontraron 21 especies con diferencias significativas en comparación la genética 2 (figura 50).



**Figura 50:** Análisis con LEfSe a nivel de especie, comparando las muestras de intestino de camarón agrupadas por genética; solamente se muestran las especies que obtuvieron diferencias significativas entre los grupos comparados. Genética 1: estanque Mazatlán (rojo); y genética 2: estanque Lajitas (verde).

Para la genética 1 éstas son: *Azospirillum sp.*, *Vibriometschnikovii sp.*, *Geodermatophilus sp.*, *Lactococcus sp.*, *Erysipelothrix sp.*, *Photorhabdus sp.*, *Demequina sp.*, *Sinorhizobium sp.*, *Actinoplanes sp.*, *Mycobacterium sp.*, *Microbacterium sp.*, *DA101*, *Alkaliphilus sp.*, *Streptococcus sp.*, *Rathayibacter sp.*, *Rhodo-*

## RESULTADOS

---

*luna* sp., *Xiphinematobacter* sp., *Piscibacillus salipiscarius*, *Polymorphumgilvum* sp., *BD213*, *Acinetobacter* sp., *Lutimonas* sp., *Comamonas* sp. y *Nautellaitalica* sp. En la genética 2 se encontraron 8 especies con abundancias más altas y diferencialmente significativas: *Coraliomargarita* sp., *Halomicronema excentricum*, *Cylindrospermopsis* sp., *Trichodesmium* sp., *Cylindrospermopsis raciborskii*, *Oceanospirillum* sp., *Planktothricoides* sp. y *Phormidium* sp.

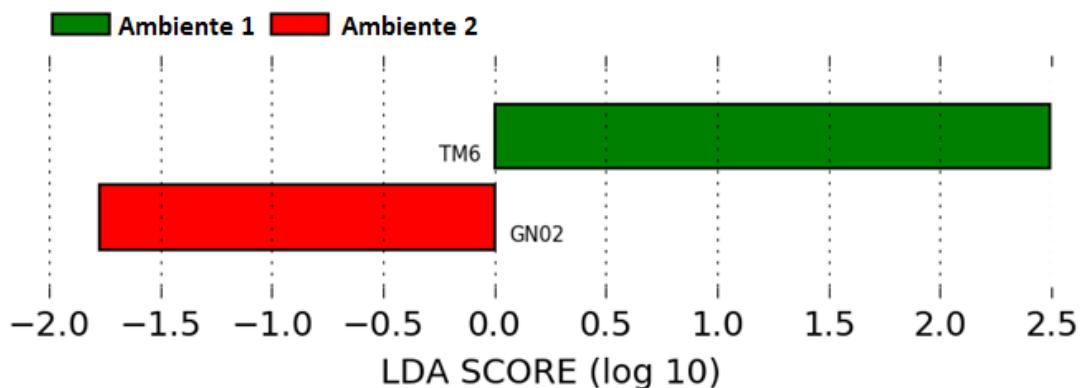
## RESULTADOS

---

### Pruebas LEfSe por ambiente

#### Análisis con LEfSe a nivel phylum

Para la comparación a nivel phylum se obtuvieron dos phyla con diferencias significativas, el ambiente 1 con TM6 y para el ambiente 2 GN02. Ninguno de estos dos phyla son cultivables (figura 51).



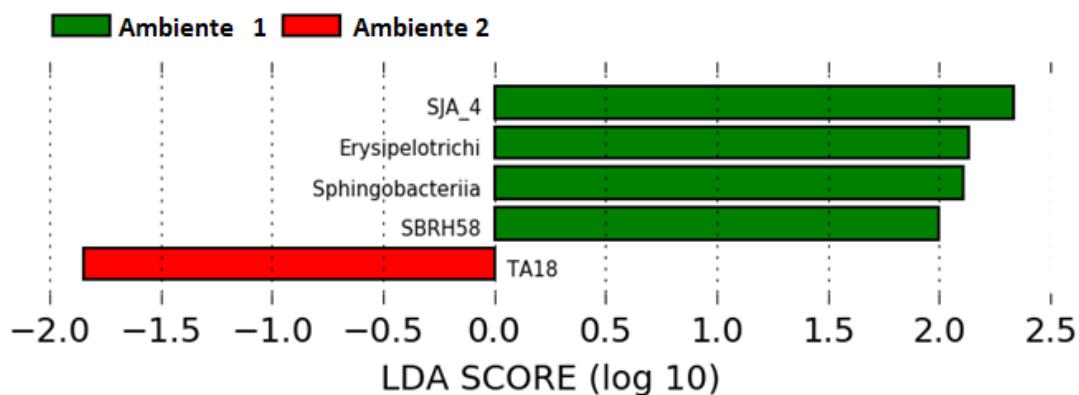
**Figura 51:** Análisis con LEfSe a nivel de phylum, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran los phyla que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo).

#### Análisis con LEfSe a nivel clase

En el análisis a nivel de clase se obtuvieron para el ambiente 1 cuatro clases con diferencias significativas, éstos son: SJA\_4, Erysipelotrichi, Sphingobacteriia y SBRH58. De igual forma para el ambiente 2 se obtuvo una sola clase con abundancia sobresaliente, TA18 (figura 52).

## RESULTADOS

---



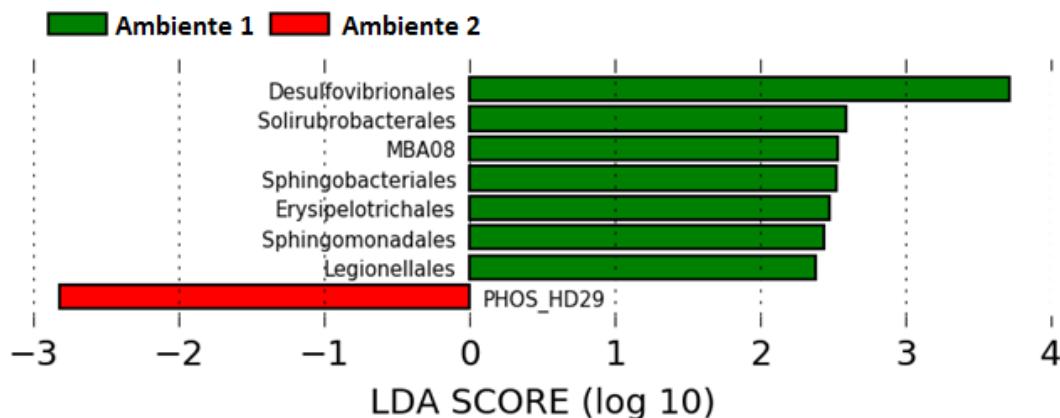
**Figura 52:** Análisis con LEfSe a nivel de clase, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran las clases que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo).

### Análisis con LEfSe a nivel orden

A nivel de orden el análisis LEfSe mostró 8 órdenes con diferencias significativas entre los ambientes; para el ambiente 1 se obtuvieron 7 órdenes (Desulfovibrionales, Solirubrobacteriales, MBA08, Sphingobacteriales, Erysipelotrichales, Sphingomonadales y Legionellales) mientra que para el ambiente 2 solamente se obtuvo un orden (PHOA\_HD29)(figura 53).

## RESULTADOS

---



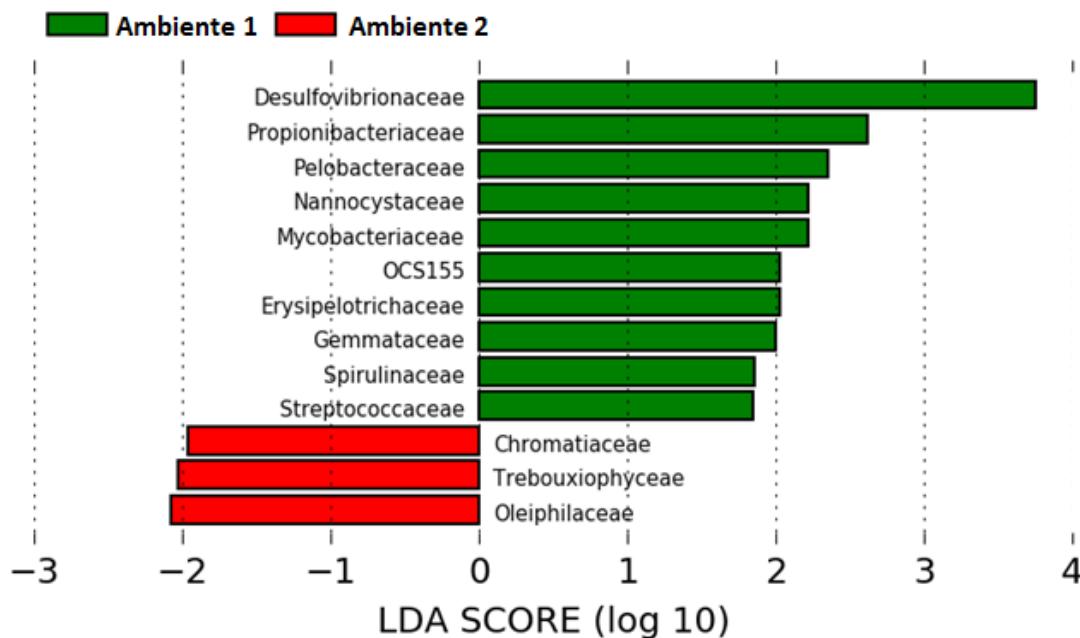
**Figura 53:** Análisis con LEfSe a nivel de orden, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran los órdenes que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo).

### Análisis con LEfSe a nivel familia

Para el ambiente 1 se obtuvieron 10 familias con diferencias significativas, éstas son: Desulfovibrionaceae, Propionibacteriaceae, Pelobacteraceae, Nannocysteaceae, Mycobacteriaceae, OCS155, Erysipelotrichaceae, Gemmataceae, Spirulinaeae y Streptococcaceae. Mientras que para el ambiente 2 se encontraron tres familias: Chromatiaceae, Trebouxiophyceae y Oleophilaceae.

## RESULTADOS

---



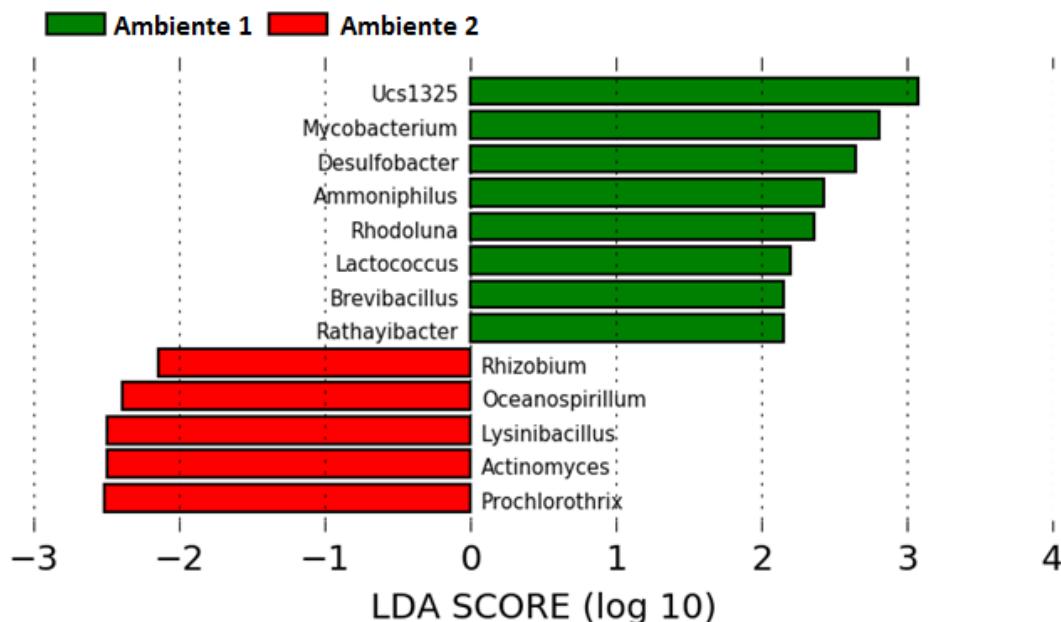
**Figura 54:** Análisis con LEfSe a nivel de familia, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran las familias que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo).

### Análisis con LEfSe a nivel género

En la comparación a nivel de género se obtuvieron en total 13 géneros con diferencias significativas, 8 géneros para el ambiente 1 y 5 para el ambiente 2. Los géneros del ambiente 1 son: *Ucs1325*, *Mycobacterium*, *Desulfobacter*, *Ammoniphilus*, *Rhodoluna*, *Lactococcus*, *Brevibacillus* y *Rathayibacter*. Mientras que para el ambiente 2 fueron: *Rhizobium*, *Oceanospirillum*, *Lysinibacillus*, *Actinomyces* y *Prochlorothrix*.

## RESULTADOS

---



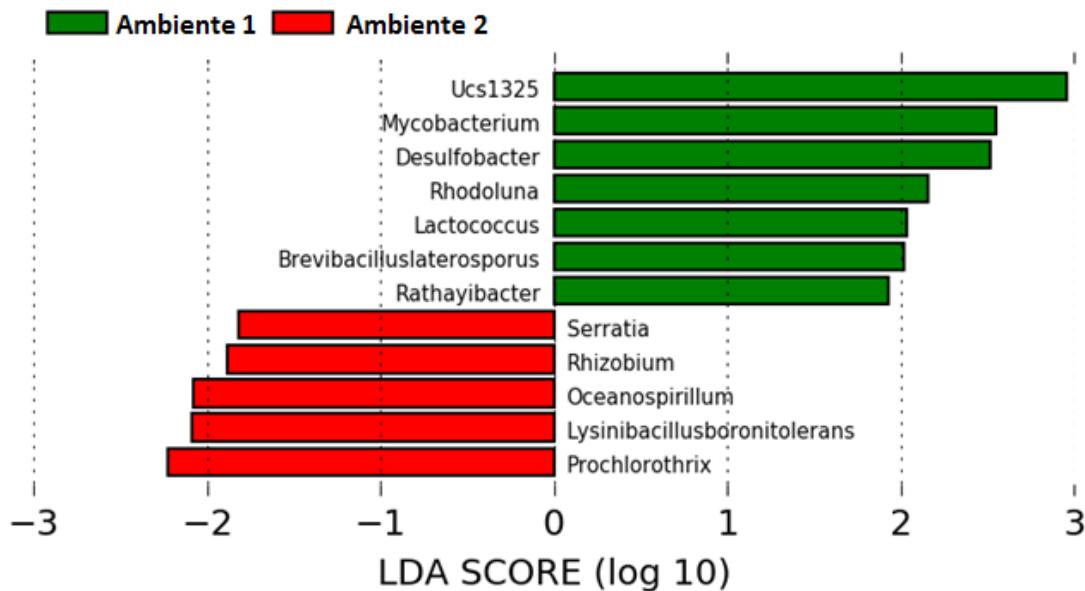
**Figura 55:** Análisis con LEfSe a nivel de género, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran los géneros que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo).

### Análisis con LEfSe a nivel especie

En la prueba realizada a nivel phylum se obtuvieron cuatro phyla con diferencias significativas para la ambiente 1 (*Ucs1325*, *Lactococcus sp.*, *Rhodoluna sp.*, *Rathayibacter sp.*, *Mycobacterium sp.*, *Brevibacilluslaterosporus sp.* y *Desulfobacter* ), mientras que para el ambiente 2 se encontraron 5 especies diferenciales: *Prochlorothrix sp.*, *Rhizobium sp.*, *Serratia sp.*, *Oceanospirillum sp.* y *Lysinibacillus boronitolerans* (figura 56).

## RESULTADOS

---



**Figura 56:** Análisis con LEfSe a nivel de especie, comparando las muestras de intestino de camarón agrupadas por ambiente; solamente se muestran las especies que obtuvieron diferencias significativas entre los grupos comparados. Ambiente 1: estanque Mazatlán (verde); y ambiente 2: estanque Estrella (rojo).

## Análisis de Diversidad $\alpha$

### Comparación entre estanques

En las curvas de rarefacción agrupadas por estanque (figuras 57 y 60) se puede observar, con los diferentes índices, que la diversidad de los sedimentos es mayor a la observada en las muestras de intestino. Además las diferencias entre tipo de muestra, sedimento o intestino, no son significativamente diferentes (figuras 59 y 61).

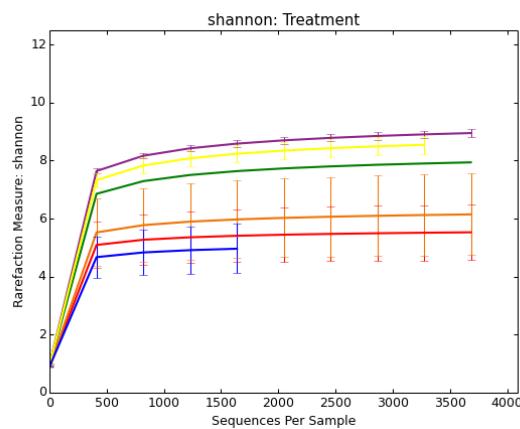
Según el índice de Shannon, que mide la diversidad de especies, el sedimento más diverso es el del estanque Lajitas, mientras que el menos diverso es el estanque Mazatlán. Para las muestras de intestino el estanque más diverso fue el del estanque Mazatlán y el menos diverso fue el estanque Lajitas (figura 57). Se considera que cuando la curva se approxima a la asíntota se tiene una buena aproximación a la diversidad de la muestra y que con mayor profundidad de secuenciación se descubriría cada vez menos diversidad. En todas los estan-

## RESULTADOS

---

ques según esta consideración, se tiene una buena estimación de la diversidad de especies.

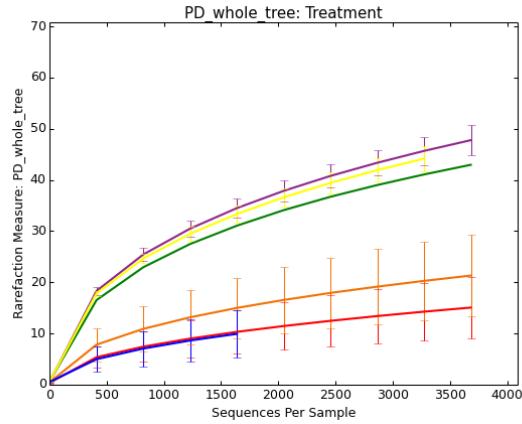
Se consideró al índice Chao1 como estimador de la riqueza de especies. Se puede observar en la figura 60 que la muestra de sedimento con mayor riqueza de especies es la del estanque Lajitas y la menor la del estanque Mazatlán. En las muestras de intestino se encuentra una mayor riqueza de especies en el estanque Mazatlán y una riqueza similar entre los estanques Lajitas y Estrella. Tomando en cuenta la proximidad de las curvas a la asíntota, según este índice no se alcanza la riqueza de especies de ninguna muestra, aunque comienza a aproximarse a una profundidad de secuenciación de 3500 lecturas.



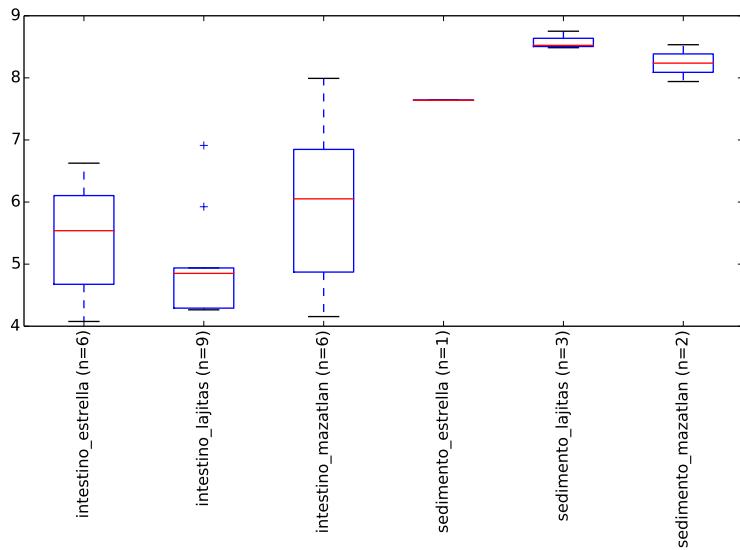
**Figura 57:** Curva de rarefacción con el índice de Shannon, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y estanque: sedimento estanque Lajitas (morado), sedimento estanque Estrella (verde), sedimento estanque Mazatlán (amarillo), intestino estanque Lajitas (azul), intestino estanque Estrella (rojo) e intestino estanque Mazatlán (naranja).

## RESULTADOS

---



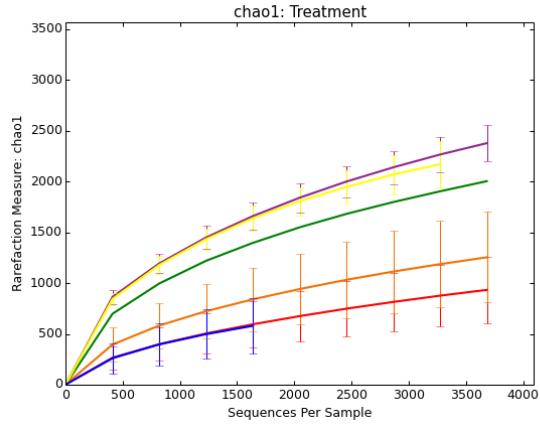
**Figura 58:** Curva de rarefacción con el índice de PD, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y estanque: sedimento estanque Lajitas (morado), sedimento estanque Estrella (verde), sedimento estanque Mazatlán (amarillo), intestino estanque Lajitas (azul), intestino estanque Estrella (rojo) e intestino estanque Mazatlán (naranja).



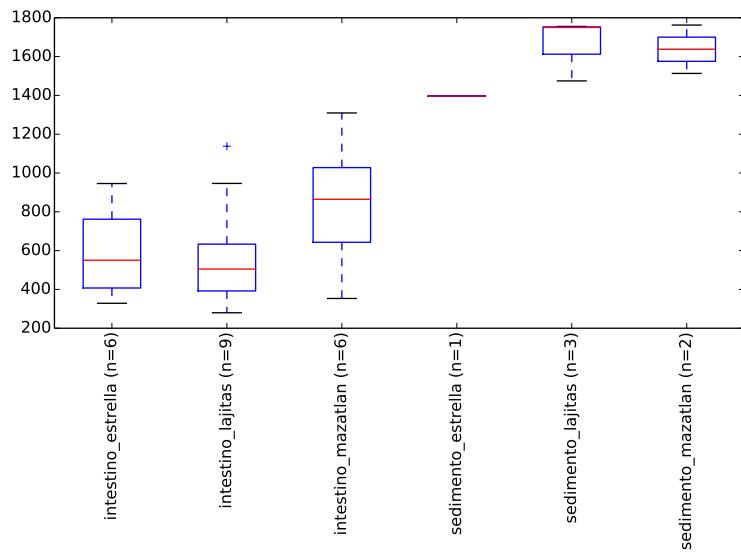
**Figura 59:** Gráfica de caja y bigotes de la comparación de la diversidad  $\alpha$  con el índice de Shannon entre los intestinos y sedimentos de los tres estanques.

## RESULTADOS

---



**Figura 60:** Curva de rarefacción con el índice Chao1, que estima la riqueza de especies; se encuentran agrupadas por tipo de muestra y estanque: sedimento estanque Lajitas (morado), sedimento estanque Estrella (verde), sedimento estanque Mazatlán (amarillo), intestino estanque Lajitas (azul), intestino estanque Estrella (rojo) e intestino estanque Mazatlán (naranja).



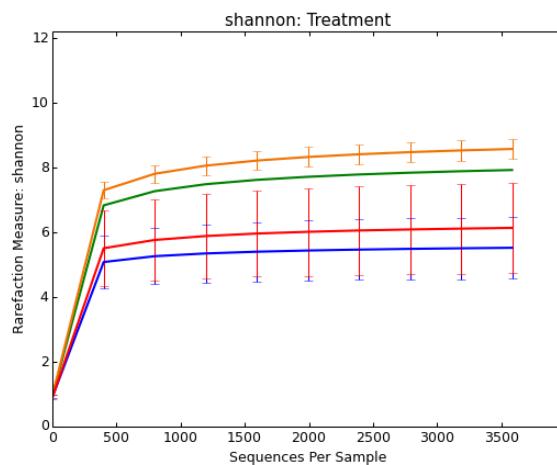
**Figura 61:** Gráfica de caja y bigotes de la comparación de la diversidad  $\alpha$  con el índice de Chao1 entre los intestinos y sedimentos de los tres estanques.

## RESULTADOS

---

### Comparación entre ambientes

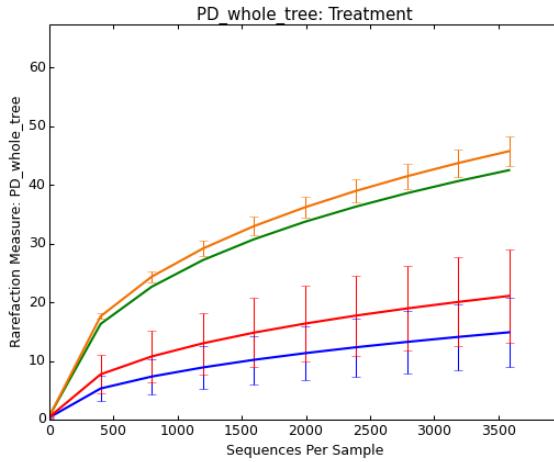
Como se puede observar en las figuras 62, 63 y 64 las muestras de sedimento tienen una mayor de diversidad que las de las muestras de intestino. Esto concuerda con lo observado en las curvas de rarefacción comparando a los ambientes. Se puede distinguir además que , tanto el sedimento como los intestinos del ambiente 1, cuentan con una mayor diversidad de especies coincidiendo todos los índices de diversidad considerados.



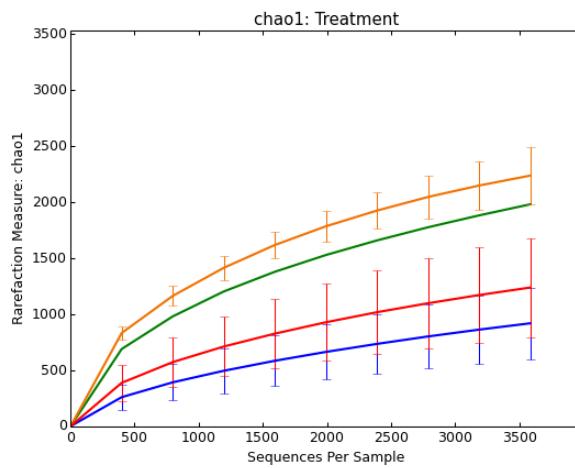
**Figura 62:** Curva de rarefacción con el índice de Shannon, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y ambiente: sedimento ambiente 1 (amarillo), sedimento ambiente 2 (verde), intestino ambiente 1 (rojo) e intestino ambiente 2 (azul).

## RESULTADOS

---



**Figura 63:** Curva de rarefacción con el índice de PD, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y ambiente: sedimento ambiente 1 (amarillo), sedimento ambiente 2 (verde), intestino ambiente 1 (rojo) e intestino ambiente 2 (azul).



**Figura 64:** Curva de rarefacción con el índice de Chao1, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y ambiente: sedimento ambiente 1 (amarillo), sedimento ambiente 2 (verde), intestino ambiente 1 (rojo) e intestino ambiente 2 (azul).

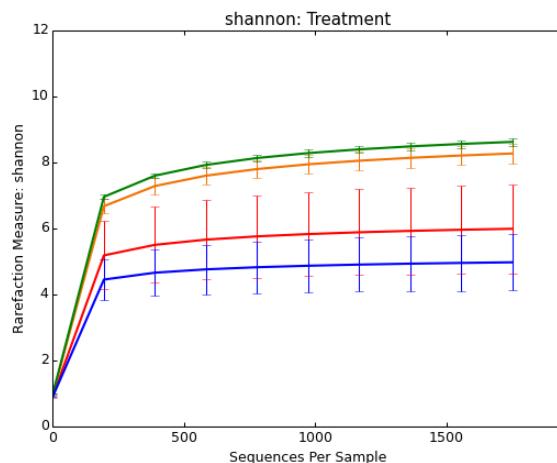
## RESULTADOS

---

### Comparación entre genéticas

Comparando las curvas de rarefacción con el índice de Shannon no existen diferencias significativas entre las muestras de sedimento o intestino [66](#). Sin embargo se puede observar en la figura [65](#) que el sedimento de la genética 2 tiene ligeramente una mayor diversidad y que las muestras de intestino de la genética 1 tienen una mayor diversidad.

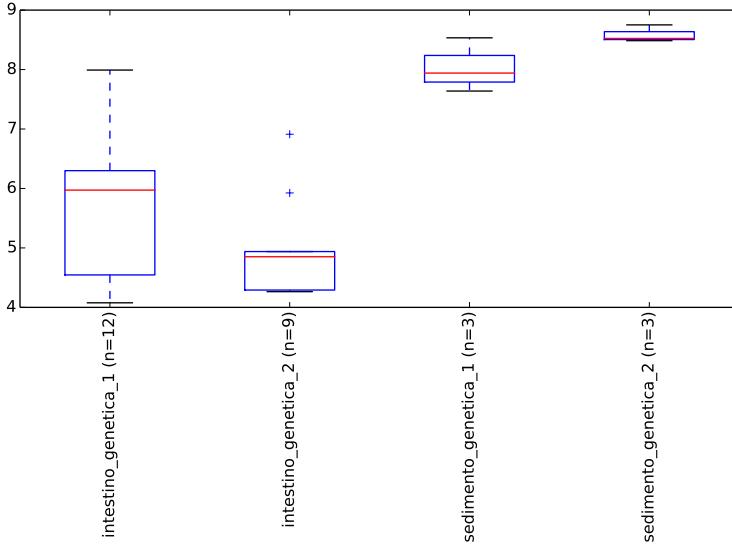
Considerando la riqueza de especies según el índice de Chao1 tampoco se encontraron diferencias significativas. El sedimento de la genética 2 tiene una riqueza ligeramente superior al igual que las muestras de intestino de la genética 1 (figura [67](#)).



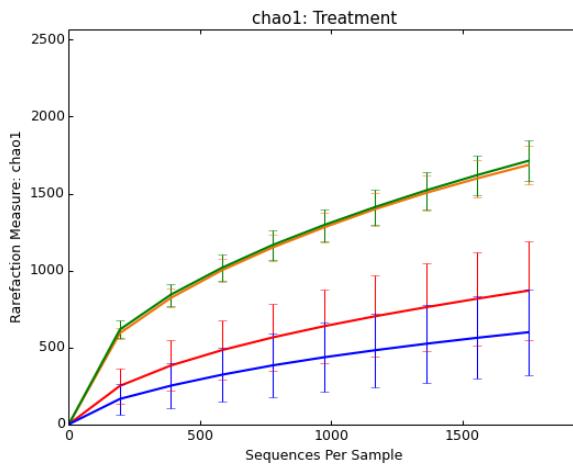
**Figura 65:** Curva de rarefacción con el índice de Shannon, que estima la diversidad de especies; se encuentran agrupadas por tipo de muestra y genética: sedimento genética 1 (amarillo), sedimento genética 2 (verde), intestino genética 1 (rojo) e intestino genética 2 (azul).

## RESULTADOS

---



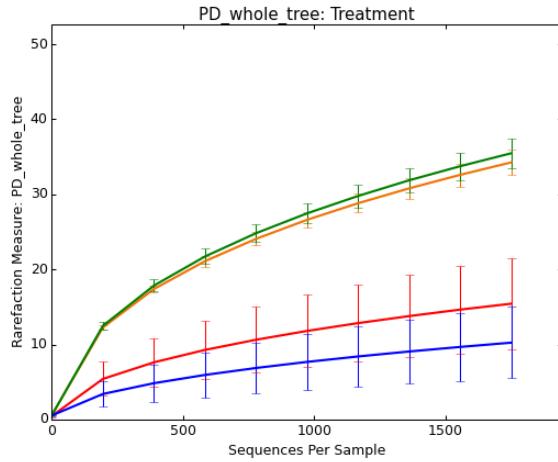
**Figura 66:** Gráfica de caja y bigotes de la comparación de la diversidad  $\alpha$  con el índice de Shannon entre los intestinos y sedimentos separados por genética.



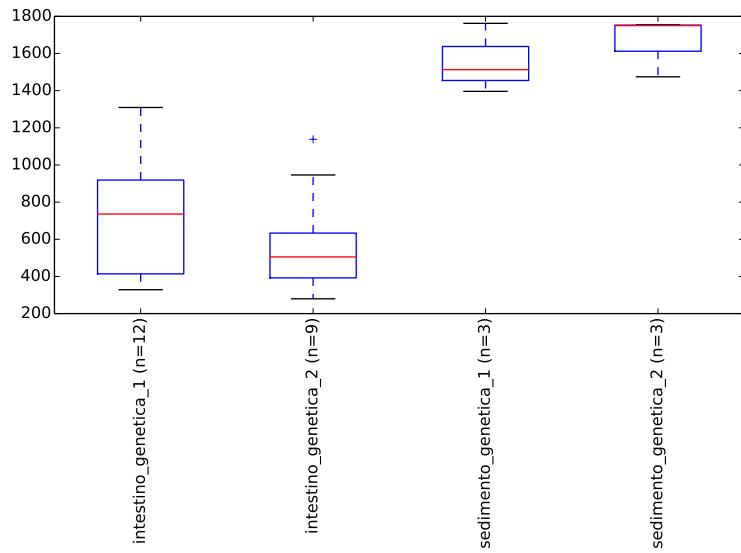
**Figura 67:** Curva de rarefacción con el índice Chao1, que estima la riqueza de especies; se encuentran agrupadas por tipo de muestra y genética: sedimento genética 1 (amarillo), sedimento genética 2 (verde), intestino genética 1 (rojo) e intestino genética 2 (azul).

## RESULTADOS

---



**Figura 68:** Curva de rarefacción con el índice PD, que estima la riqueza de especies; se encuentran agrupadas por tipo de muestra y genética: sedimento genética 1 (amarillo), sedimento genética 2 (verde), intestino genética 1 (rojo) e intestino genética 2 (azul).



**Figura 69:** Gráfica de caja y bigotes de la comparación de la diversidad  $\alpha$  con el índice de Chao1 entre los intestinos y sedimentos separados por genética

## RESULTADOS

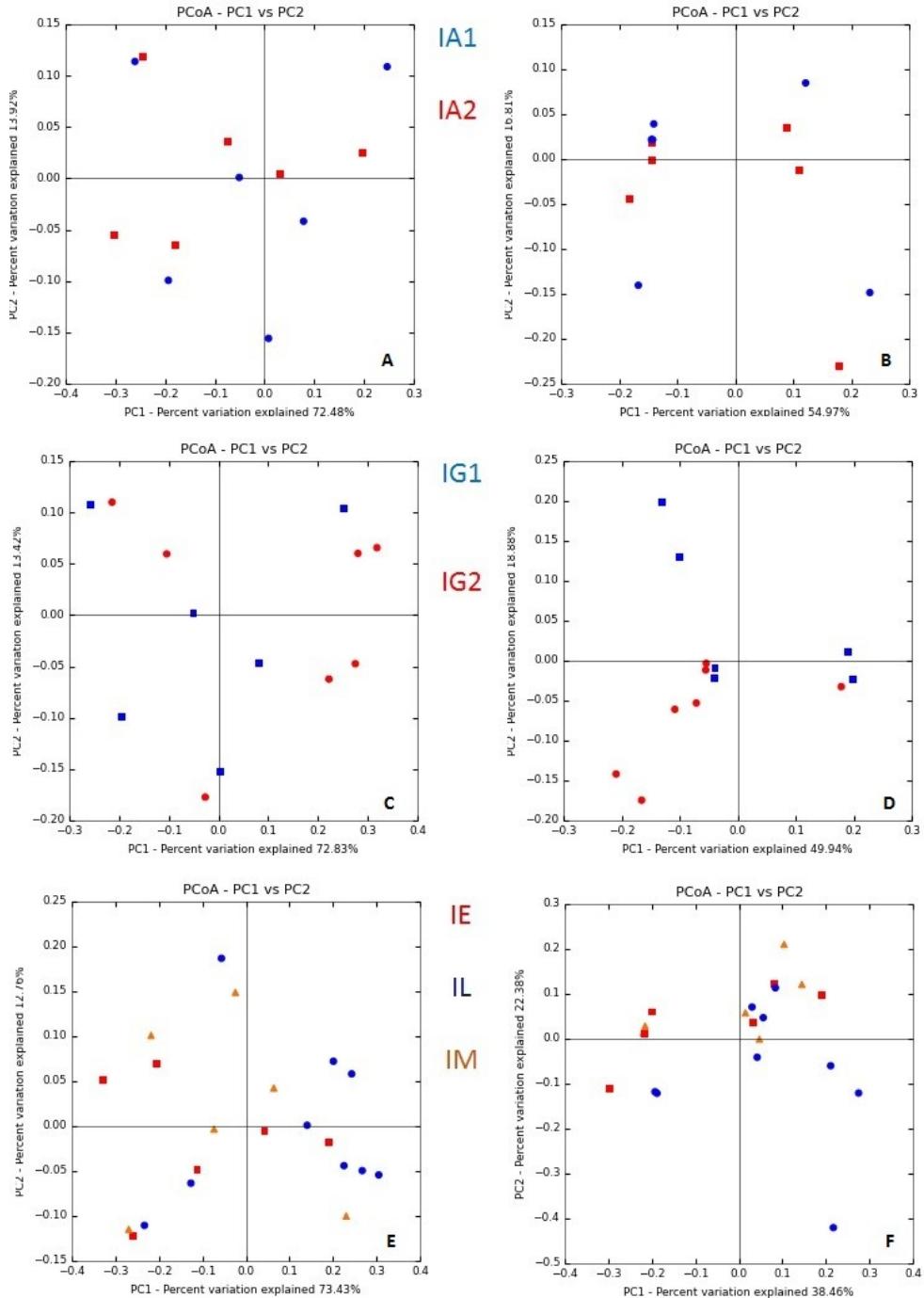
---

### Análisis de Diversidad $\beta$

En la figura 70 se muestra el PCoA sin considerar las muestras de sedimento, mientras que en el gráfico 71 se agrupan las muestras considerando también los sedimentos de su respectivo estanque.

## RESULTADOS

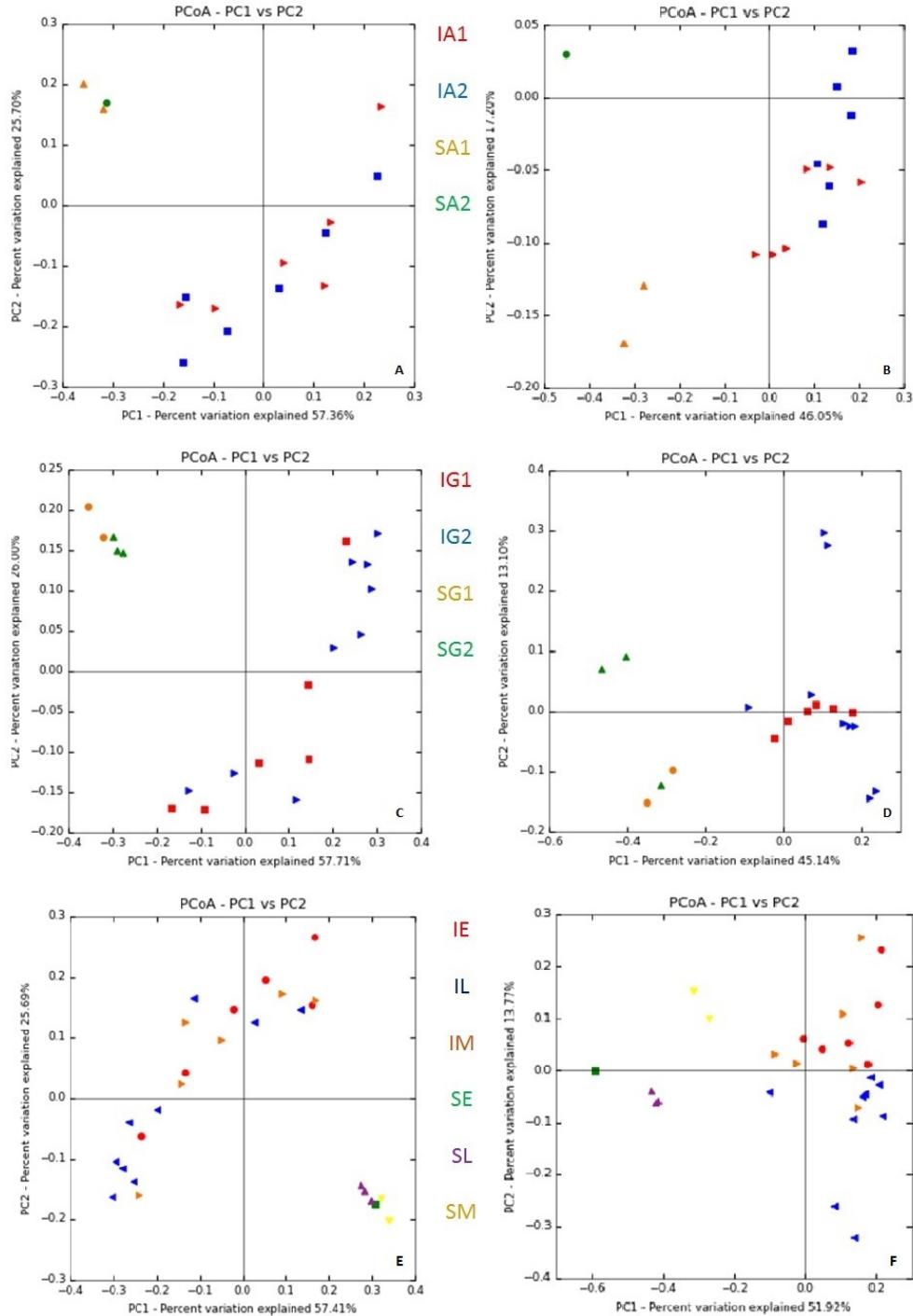
---



**Figura 70:** PCoA ponderado (*weighted*) con las muestras de intestino de camarón; izquierda: A, C y E) y no ponderado (*unweighted*; derecha: B, D y F) con los PC que explican el mayor porcentaje de variación. Muestras de intestino de camarón; las muestras están agrupadas por ambiente (A y B), genética (C y D) y estanque (E y F).

## RESULTADOS

---



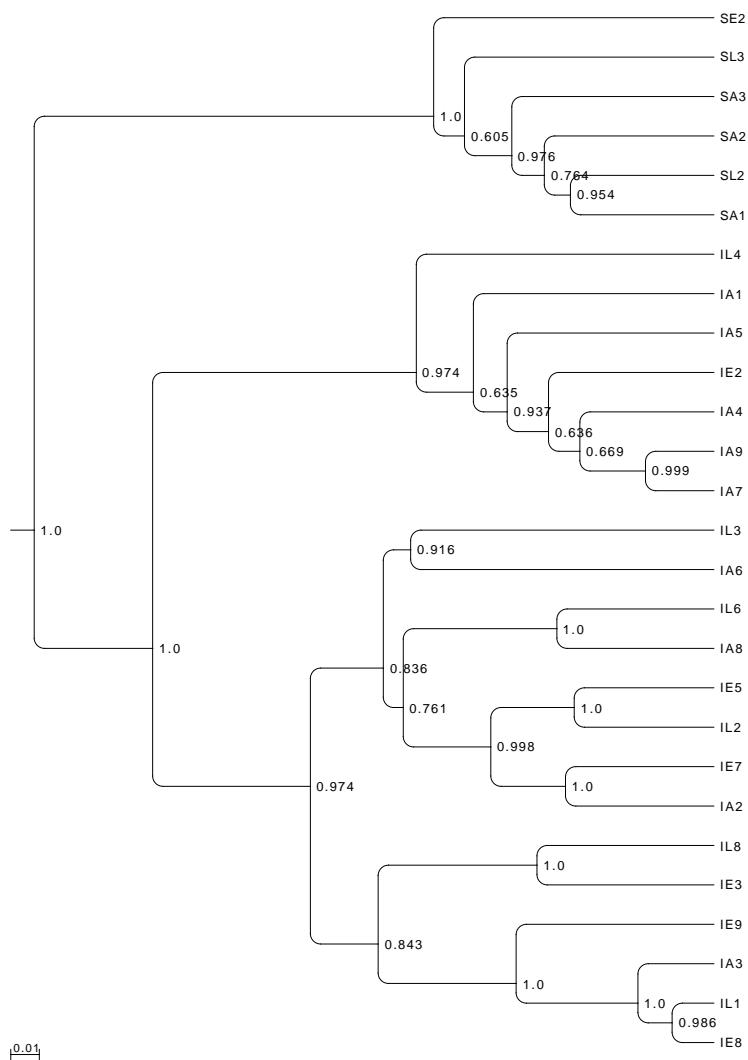
**Figura 71:** PCoA ponderado (*weighted*) con muestras de intestino de camarón y sedimento; izquierda: A, C y E) y no ponderado (*unweighted*; de recha: B, D y F) con los PC que explican el mayor porcentaje de variación. Muestras de intestino de camarón y de sedimento; las muestras están agrupadas por ambiente (A y B), genética (C y D) y estanque (E y F).

## RESULTADOS

---

### Árbol filogenético

En la figura 72 (árbol ponderado) se puede observar que las muestras de sedimento se separan completamente de las de intestino y que todas las muestras de intestino forman un gran grupo pero sin separarlas por genética o estanque. De igual forma se observa que el árbol no ponderado (figura A.9) mantiene a los sedimentos en una rama y a la mayoría de las muestras de intestino con excepción de la muestra IE9. Ésto concuerda con los PCoA anteriores.



**Figura 72:** Árbol generado con Jackknife ponderado (*weighted*); cada nodo cuenta con su valor confianza.

## RESULTADOS

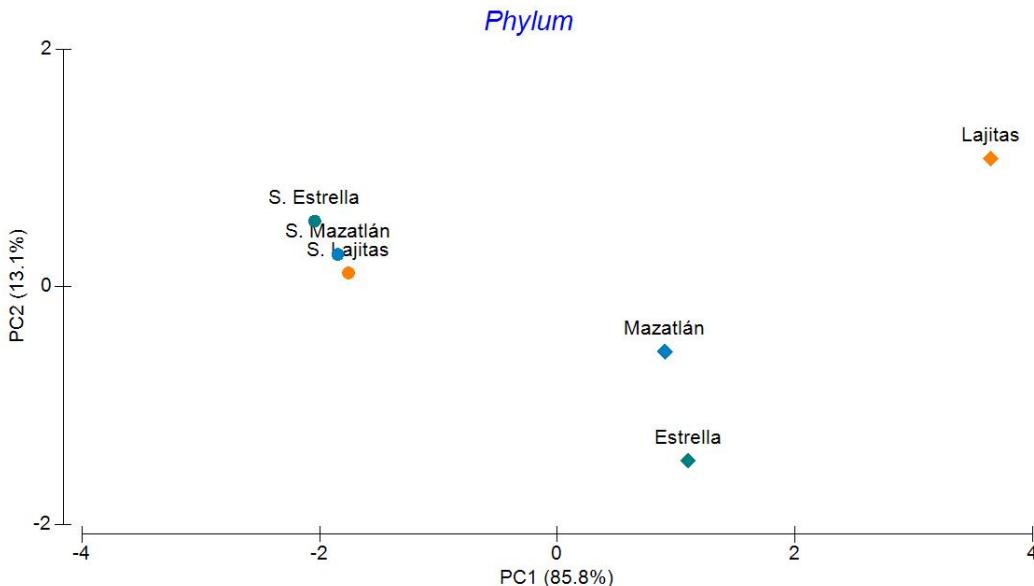
---

### Análisis de Componentes Principales

En este análisis se consideran a todos los OTU, incluyendo a los de abundancias menores al 0.1 %.

#### PCA a nivel de phylum

En la gráfica a nivel de phylum (figura 73) se muestran los PC que explican el mayor porcentaje de variación (98.9 %). Se puede observar que las muestras de sedimentos se agrupan, mientras que las muestras de intestino tienden agruparse por su genética (Mazatlán y Estrella).



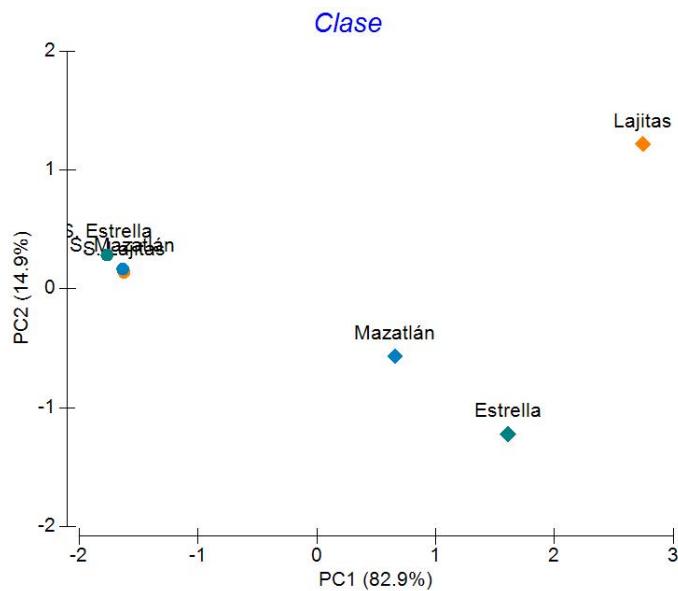
**Figura 73:** PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de phylum.

#### PCA a nivel de clase

El PCA con las muestras agrupadas por clase explica el 97.8 % de la variabilidad entre las muestras. Se puede observar que los sedimentos de los tres estanques se agrupan casi completamente, mientras que las muestras de intestino se agrupan más por su genética (estanques Mazatlán y Estrella) que por su ambiente (figura 74).

## RESULTADOS

---



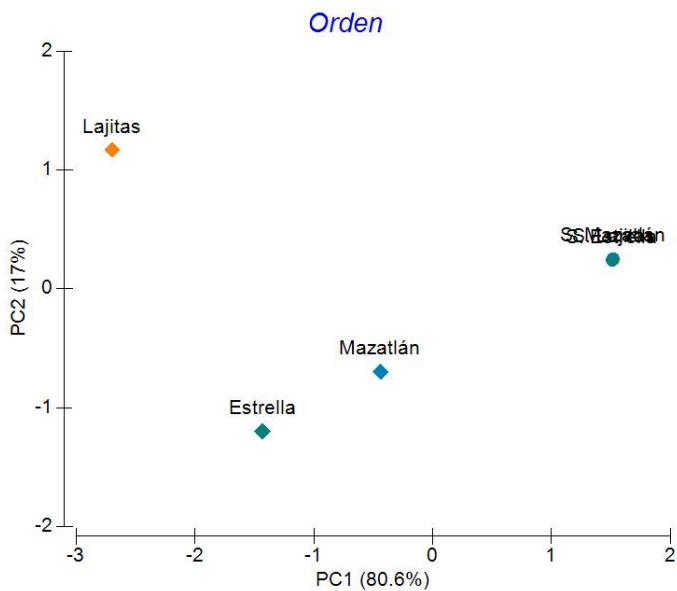
**Figura 74:** PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de clase.

### PCA a nivel de orden

En la figura 75 se muestra el PCA a nivel de orden, que explica el 97.6 % de la variación a este nivel. Se muestra en la figura que las muestras de sedimento se agrupan totalmente y que, al igual que en los niveles anteriores, las muestras de intestino tienden a agruparse por su genética más no por su ambiente.

## RESULTADOS

---



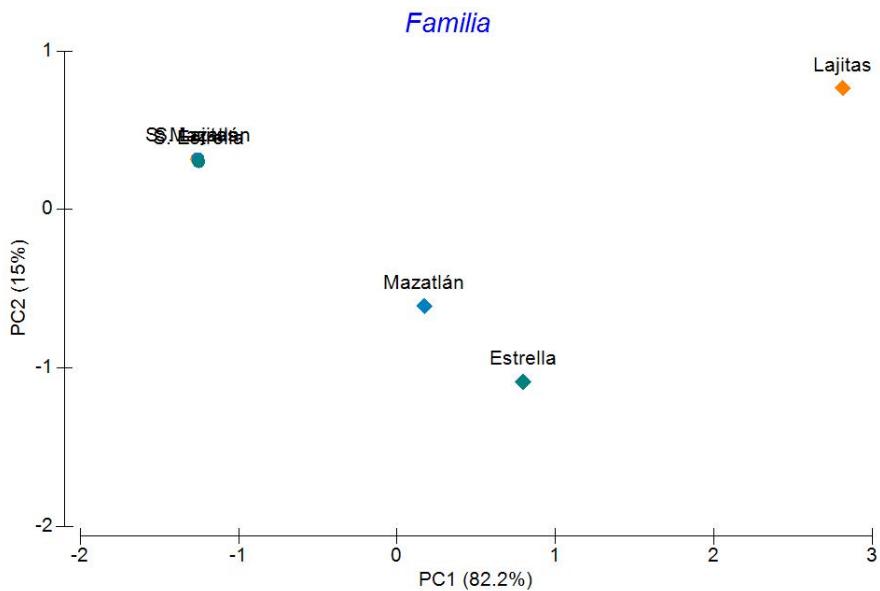
**Figura 75:** PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de orden.

### PCA a nivel de familia

El PCA con las muestras agrupadas a nivel de familia explica el 97.2 % de la variabilidad entre las muestras. En la figura se puede observar que los sedimentos de los tres estanques se agrupan casi completamente, mientras que las muestras de intestino tienden a agruparse más por su genética (estanques Mazatlán y Estrella) que por su ambiente (figura 76).

## RESULTADOS

---



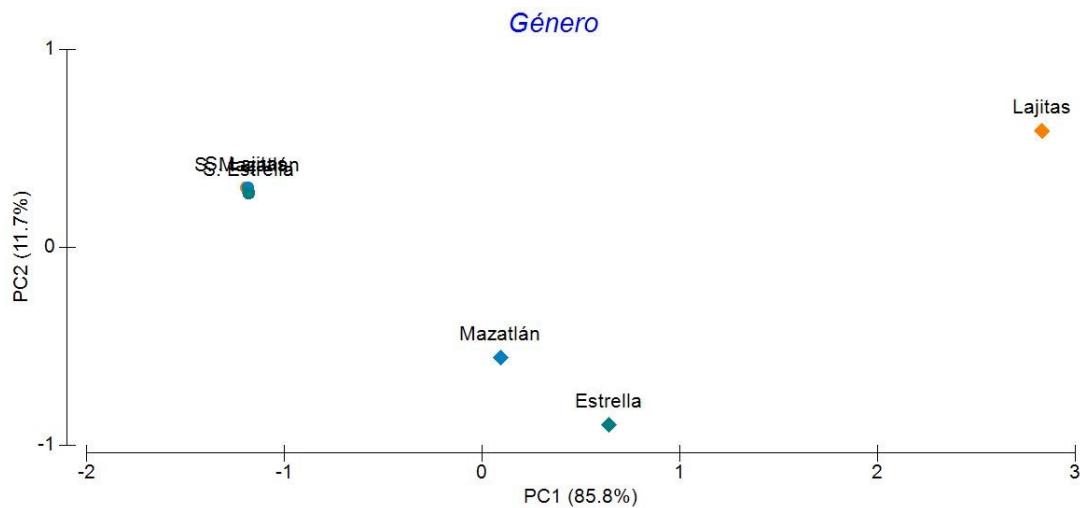
**Figura 76:** PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de familia.

### PCA a nivel de género

En la gráfica a nivel de género (figura 77) se muestran los PC que explican el mayor porcentaje de variación (97.5 %). Se puede observar que las muestras de sedimentos se agrupan, mientras que las muestras de intestino tienden agruparse por su genética (los estanques Mazatlán y Estrella provienen del mismo linaje) más que por su ambiente (los estanques Estrella y Lajitas comparten un ambiente similar).

## RESULTADOS

---



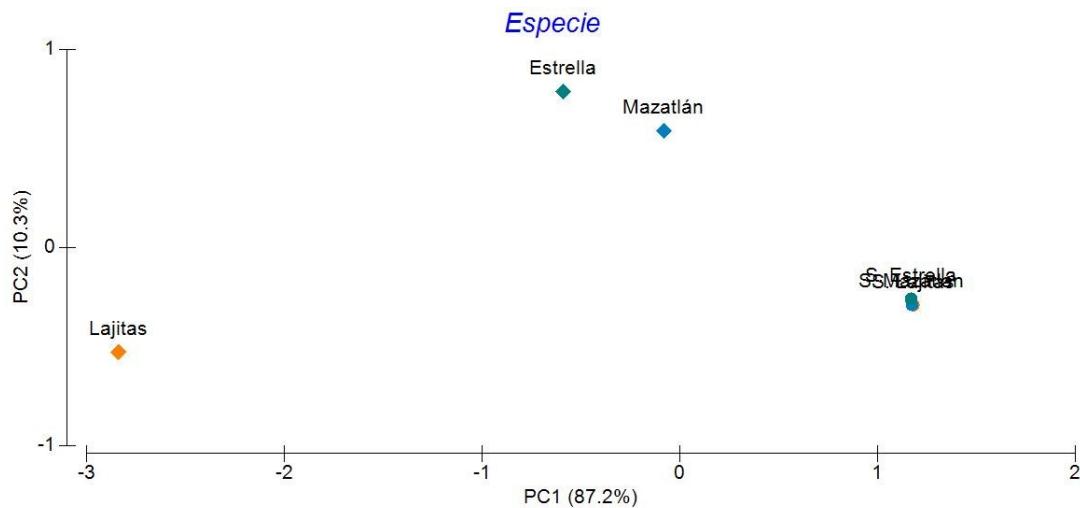
**Figura 77:** PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de género.

### PCA a nivel de especie

Por último, en el PCA con los OTU a nivel de especie se puede observar que las muestras de sedimento se agrupan totalmente, mientras que las muestras de intestino de los estanques se agrupan por su genética compartida, de los estanques Mazatlán y Estrella, y que solamente las muestras del estanque Lajitas se orientan en otra dirección. Este PCA nos explica el 97.5 % de la variación a este nivel, con los PC1 y PC2 (figura 78).

## RESULTADOS

---



**Figura 78:** PCA con las muestras agrupadas por estanques (Lajitas, Mazatlán y Estrella) así como tipo de muestras (intestino y sedimento) con los OTU a nivel de especie.

## Agrupamiento jerárquico con mapas de calor

### Mapas de calor por ambiente

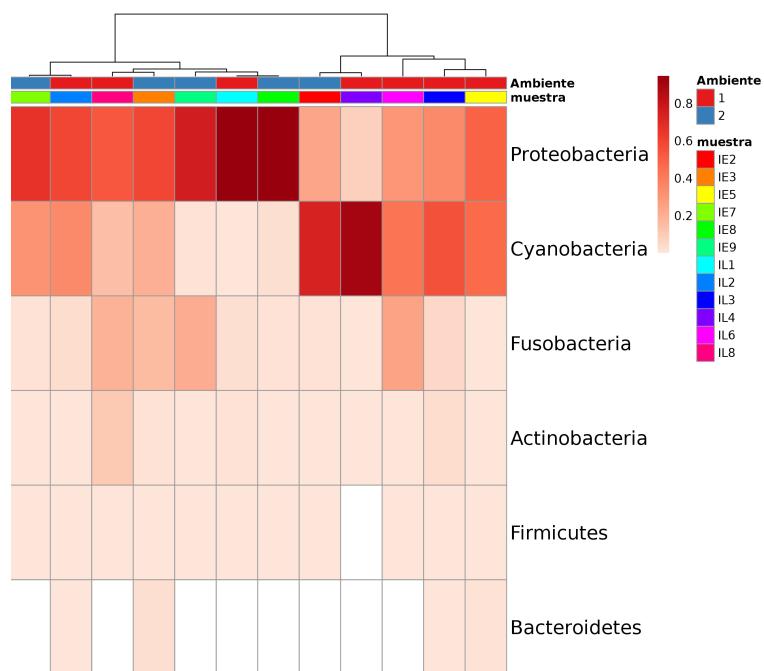
#### Mapa de calor a nivel phylum entre ambientes

En la figura 79 se puede observar el mapa de calor a nivel de phyla entre el ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella). Se obtuvo una variedad en la tonalidad de rojo, en la cual los espacios sin colorear equivalen a valores menores a 0.01 % de la frecuencia relativa y las tonalidades más fuertes hasta el 80 % de esta misma.

Se pueden observar dos grandes ramas en las cuales se agrupan las muestras, sin embargo no se separan claramente las muestras entre ambientes (figura 79).

## RESULTADOS

---



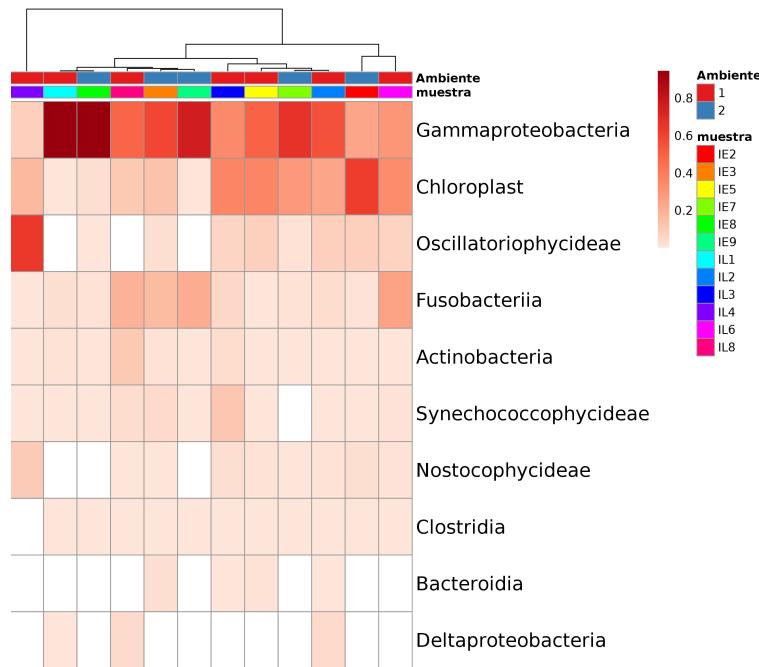
**Figura 79:** Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de phylum; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel clase entre ambientes

En el mapa de calor entre los ambientes que considera a los OTU con abundancia mayor a 0.1 % se puede observar una gran rama que abarca prácticamente a todas las muestras, excepto a la muestra IL4 (perteneciente al ambiente 1). No se puede distinguir una jerarquía entre las muestras de los dos ambientes con este mapa de calor (figura 80).



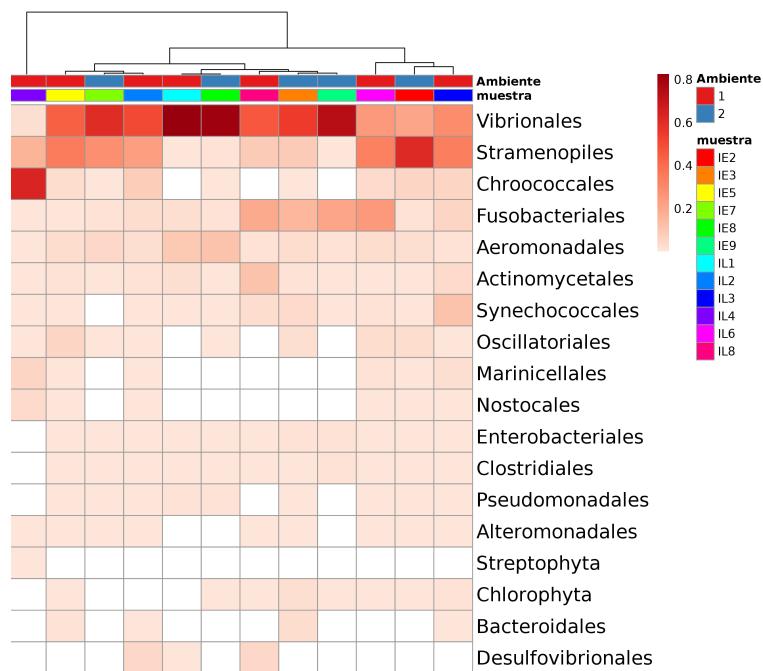
**Figura 80:** Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de clase; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel orden entre ambientes

En la figura 81 se puede observar el mapa de calor obtenido a nivel orden con las muestras de intestino agrupadas por ambiente. Como se muestra en dicha figura, se obtiene una gran rama que incluye a prácticamente a todas las muestras con excepción de IL4 del ambiente 1. Esto se observa también en el mapa de calor anterior (figura 81). Dentro de la rama que abarca al resto de las muestras no se agrupan las muestras por ambiente, por lo menos a este nivel taxonómico (figura 81).



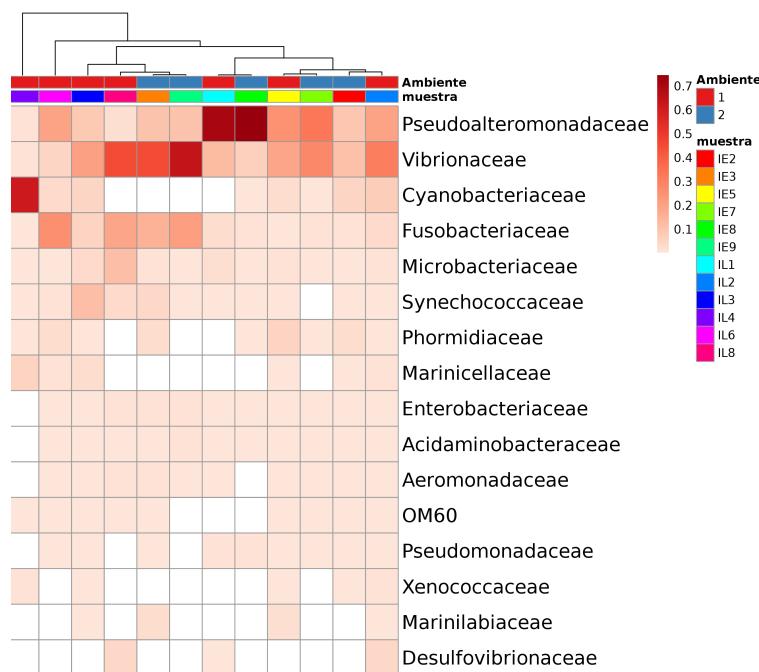
**Figura 81:** Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de orden; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel familia entre ambientes

Al igual que en las figuras 80 y 81 en el mapa de calor obtenido considerando a los OTU a nivel de familia contrastando entre ambiente (figura 82) no se obtiene una clasificación clara que nos permita separar entre muestras de intestino de un ambiente y otro. De igual forma se sale una muestra de la agrupación que abarca prácticamente a todas las muestras, IL4 del ambiente 1, aunque no se puede inferir alguna otra clasificación.



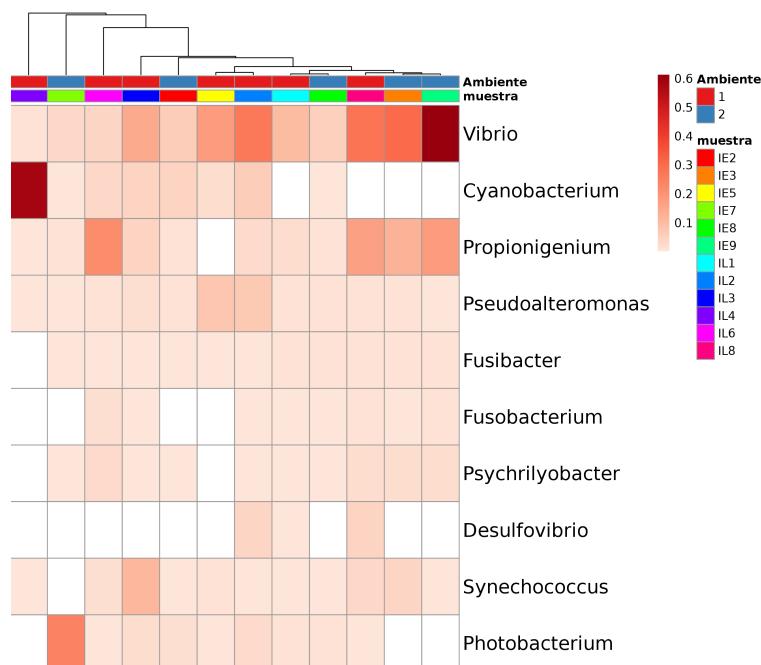
**Figura 82:** Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de familia; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel género entre ambientes

El mapa de calor a nivel de género se muestra en la figura 83. Al contrario de los mapas de calor obtenidos en los niveles anteriores, todas las muestras se agrupan en una gran rama. Sin embargo, no se puede discernir alguna jerarquía entre las muestras de los dos ambientes considerados. Se puede observar fácilmente en la figura que hay muestras que sobre-expresan algún género en específico, lo cual la saca de la clasificación.



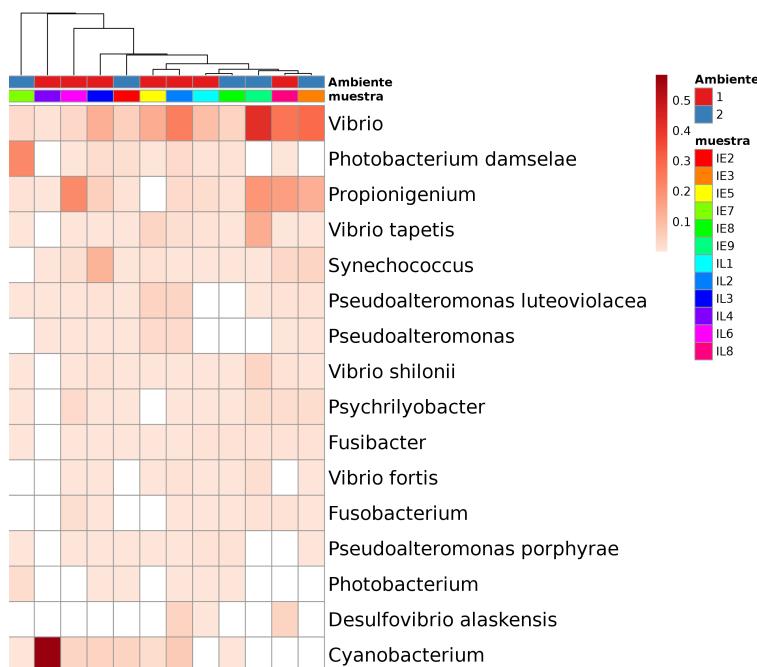
**Figura 83:** Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de género; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel especie entre ambientes

El mapa de calor entre las muestras de los dos ambientes se puede observar en la figura 84, el cual es un mapa similar al de la comparación a nivel de género (figura 83). No se puede inferir mucho acerca de dichos mapas ya que las muestras no tienen una organización jerárquica aparente, sin embargo, se pueden observar fácilmente las especies que se tienen una abundancia sobresaliente entre las muestras. Un ejemplo claro es el de *Cyanobacterium sp.*, lo cual explicaría porque la muestra IL4 sale de las clasificaciones jerárquicas a niveles más altos (figuras 81-83).



**Figura 84:** Mapa de calor entre muestras del ambiente 1 en rojo (estanque Mazatlán) y el ambiente 2 en azul (estanque Estrella) a nivel de especie; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

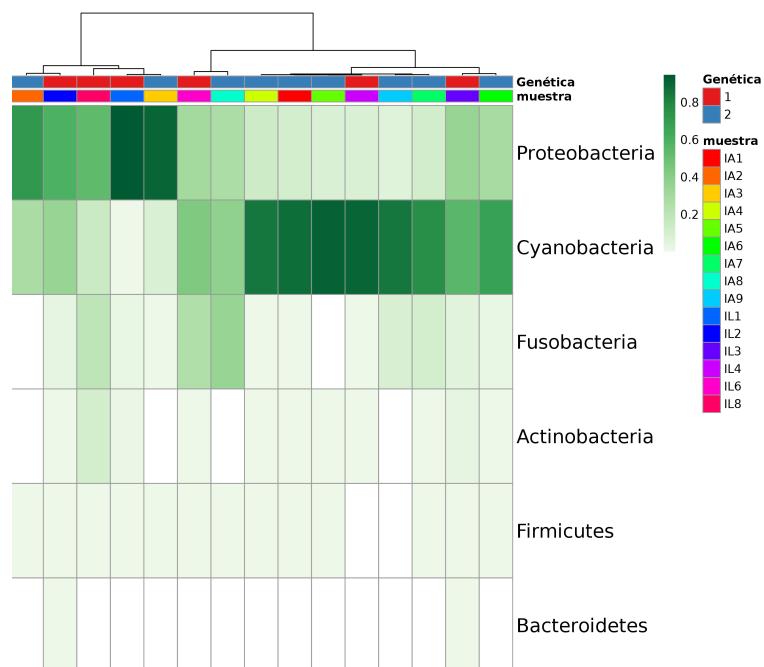
---

### Mapas de calor por genética

En la sección anterior se muestran los mapas de calor obtenidos para clasificar a las muestras entre ambientes. En esta sección se muestran dichos mapas de calor para clasificar entre las muestras de la genética 1 (estanque Mazatlán) y la genética 2 (estanque Lajitas) en los diferentes niveles taxonómicos.

#### Mapa de calor a nivel phylum entre genéticas

En la figura 85 se muestra el mapa de calor para clasificar a las muestras de las dos genéticas. Se pueden observar dos grandes ramas, las cuales se separan por su proporción de Cyanobacteria o Proteobacteria. Se distingue una jerarquía entre las muestras de las dos genéticas, las cuales se agrupan en dos ramas principales.



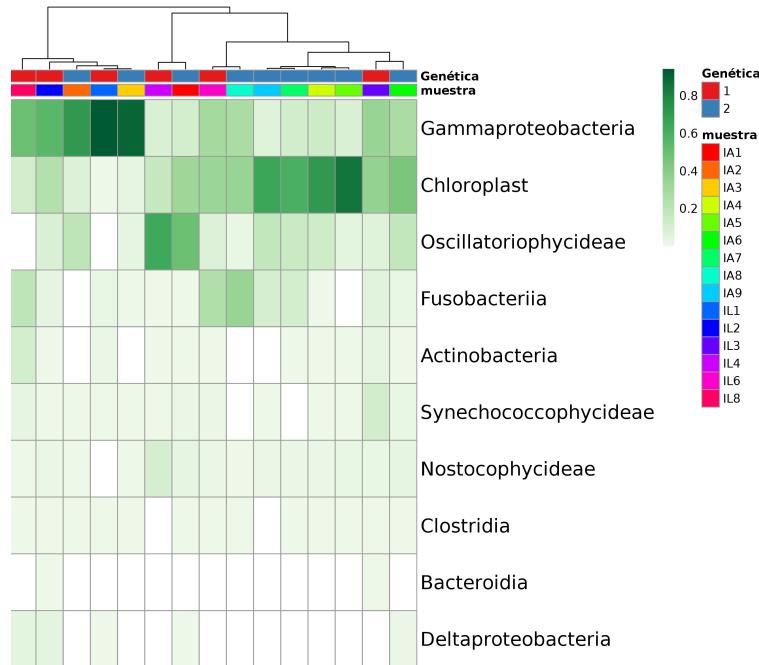
**Figura 85:** Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de phylum; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel clase entre genéticas

El mapa de calor entre las muestras de las dos genéticas se puede observar en la figura 86. Al igual que en mapa de calor a nivel de phylum (figura 85) se pueden observar dos grandes ramas que separan a las muestras, a nivel de clase, por la presencia de Gammaproteobacteria y Chloroplast. Se puede distinguir también que las muestras tienen a separarse por genéticas (1 o 2).



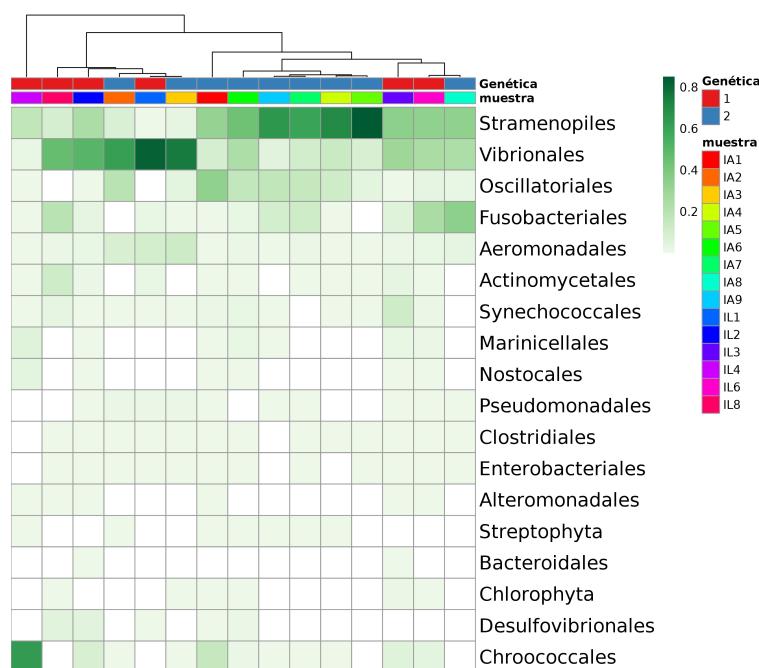
**Figura 86:** Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de clase; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel orden entre genéticas

En la clasificación con mapa de calor a nivel de orden se pueden distinguir dos grandes ramas que abarcan a la mayoría de las muestras, y una muestra (IL4) que se sale de la clasificación formando su propia rama. Se puede distinguir también que hay una tendencia más clara a separar a las muestras según su genética y que la clasificación se ve fuertemente influenciada por los dos órdenes más abundantes: Stramenopiles y Vibrionales (87).



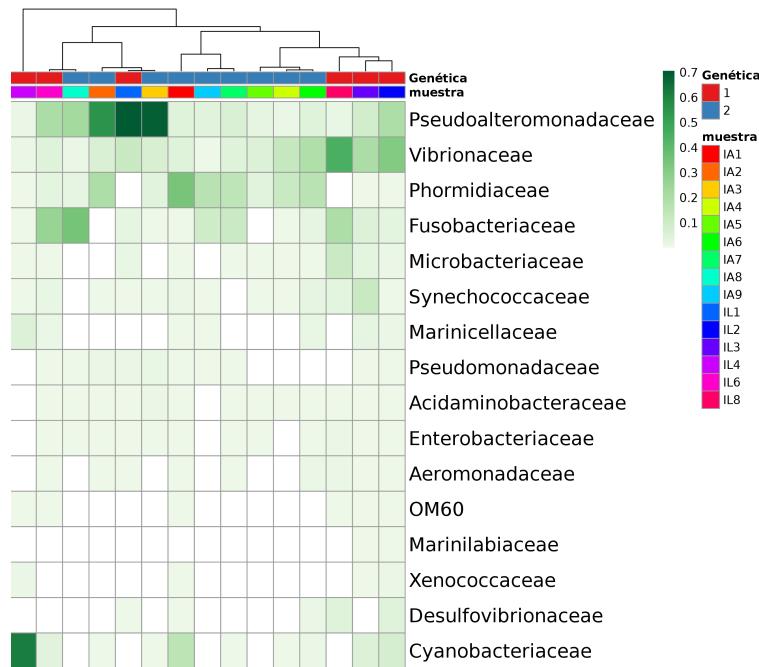
**Figura 87:** Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de orden; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel familia entre genéticas

Para el mapa de calor a nivel de familia se pueden observar una similitud entre la clasificación de las muestras, y de igual forma, la muestra IL4 de la genética 1 se sale de la clasificación. Para este nivel Pseudoalteromonadaceae, Vibrionaceae, Phormidiaceae y Fusobacteriaceae son las familias que determinan la clasificación jerárquica (88).



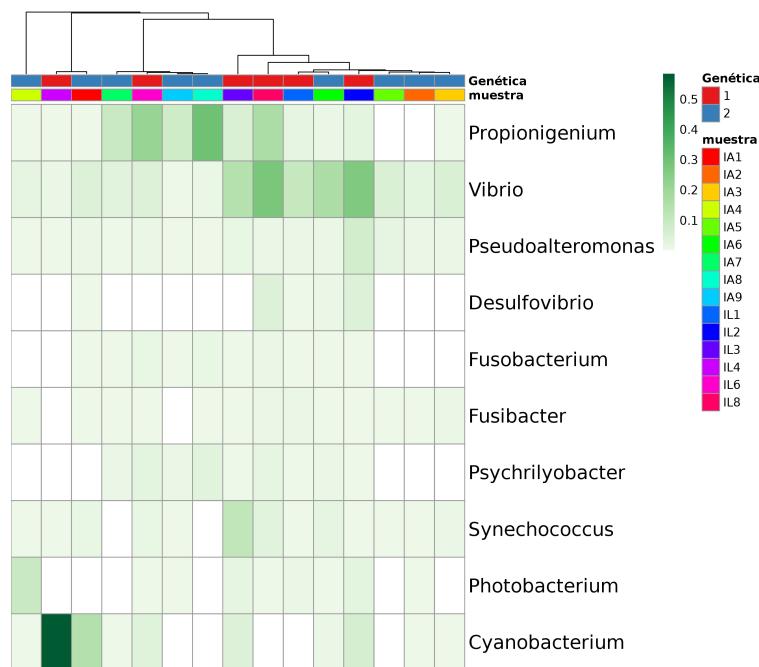
**Figura 88:** Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de familia; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel género entre genéticas

La clasificación jerárquica a nivel de género se muestra en la figura 89. Este mapa de calor no clasifica a las muestras por su genética, y se ve influenciado por géneros como *Propionigenium*, *Vibrio* y *Cyanobacterium*. En contraste con las figuras anteriores se obtiene que cada una de las muestras cuenta con género diferentes, por lo cual la clasificación con el mapa de calor es más complicada.



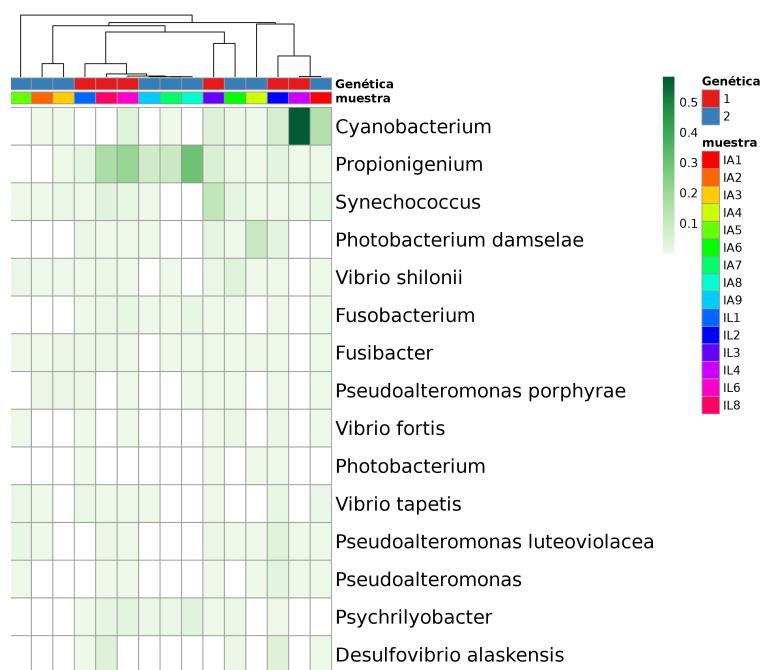
**Figura 89:** Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de género; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## RESULTADOS

---

### Mapa de calor a nivel especie entre genéticas

El mapa de calor a nivel de especie se muestra en la figura 90. Al igual que en la clasificación entre genéticas a nivel género (figura 89) se obtiene una gran rama que abarca a todas las muestras con excepción de la muestra IA5, que forma su propia rama saliendo de la clasificación. El mapa de calor a este nivel tiene influencia de especies como *Cyanobacterium sp.*, *Propionigenium sp.* y *Synechococcus* (90).



**Figura 90:** Mapa de calor entre muestras de la genética 1 en rojo (estanque Mazatlán) y la genética 2 en azul (estanque Lajitas) a nivel de especie; se consideran solamente los OTU con una abundancia mayor al 0.1 %.

## Discusión

---

En este trabajo se busca determinar si la microbiota intestinal del camarón está modulada por la genética o por el ambiente en que se encuentra. Las investigaciones realizadas bajo el enfoque de la secuenciación masiva de regiones de hipervariables del gen ARNr 16S son muy escasas en camarones y en especial en esta especie (*Litopenaeus vannamei*), por lo cual una comparación que busque determinar cuáles son los factores que influencian la microbiota es de gran importancia.

Existen diferentes bases de datos con las secuencias del gen ARNr 16S. Cada una de éstas cuenta con secuencias diferentes por lo que su asignación taxonómica varía. En este trabajo se compararon las bases de datos Greengenes (33), RDP (34) y SILVA (35), y se concluyó que la base de datos que caracterizaba mejor la diversidad fue Greengenes.

La asignación taxonómica con QIIME (36) y la base de datos de 16S Greengenes (33) nos permitió determinar que el intestino del camarón está dominado por dos phyla en especial Proteobacteria y Cyanobacteria, con un promedio de 45.60 % y 39.65 %, respectivamente. Esto no coincide totalmente con lo reportado anteriormente en la literatura. El dominio de Proteobacteria es comúnmente reportado en camarones (37)(38)(30)(8)(10)(2)(5) sin embargo la presencia de un porcentaje alto de Cyanobacteria no se menciona con una abundancia relevante. Estos resultados pueden ser consecuencia de las diferencias en técnicas para extracción, amplificación y purificación del ADN microbiano, del pre-tratamiento aplicado e incluso de los parámetros utilizados para la asignación taxonómica (6).

Al nivel de phylum existen variaciones en abundancia relativa de los taxa más abundantes. Por ejemplo en la figura 21 se ve un aumento de Proteobacteria notable en el estanque Estrella a comparación con el estanque Lajitas. De igual forma en el estanque Lajitas hay una proporción mayor de Cyanobacteria que no se repite en los otros dos estanques. Esto podría ser explicado por que estos estanques pertenecen a linajes, o genética, diferentes (genética 1 el estanque Estrella y genética 2 el estanque Lajitas) ya que cuando se observan las abundancias relativas en las dos genéticas se observan patrones de bandeo diferentes.

## DISCUSIÓN

---

Las especies más abundantes pertenecen a los géneros *Vibrio*, *Propionigenium* y *Cyanobacterium*, pero se encuentran en proporciones diferentes en los estanques (figura 41), lo cuál se observa frecuentemente en otros estudios del perfil del 16S a niveles taxonómicos más bajos. Una de las especies con mayor relevancia es *Photobacterium damselae* que es una bacteria gram-negativa que causa enfermedades en los animales acuáticos (39), mientras que las especies pertenecientes a *Vibrio* pueden estar relacionadas con enfermedades en camarones, pero se encuentran normalmente en la microbiota del camarón (1)(40).

Con las pruebas de LEfSe, que miden las diferencias significativas entre condiciones dadas, se observa que hay variaciones a todos los niveles cuando se considera a todos los OTU (figura 16), mas no cuando se consideran a las abundancias mayores a 0.1 %. Esto se debe a que las variantes que existen entre los estanques, ambientes y genéticas tienen una abundancia baja pero que son importantes al buscar cuál es el principal modulador de la microbiota.

Considerando el análisis de LEfSe se puede observar que las variaciones entre los diferentes taxones son notorios al comparar las genéticas (figuras 45-50), y son mayores que considerando los ambientes (figuras 51-56). El estanque Estrella y el estanque Mazatlán pertenecen a la misma fuente de desove, por lo cual tienen una genética similar y deberían de existir pocas diferencias al comparar los grupos con LEfSe, sin embargo sí se determinaron diferencias significativas entre los estanques. Considerando las genéticas se obtuvo un mayor porcentaje de variación entre las muestras, lo cual nos infiere que la genética puede tener un impacto más significativo sobre la modulación de la microbiota (16). Una de las diferencias más importantes entre los estanques a nivel género es la presencia de *Lactococcus* en el estanque Mazatlán. Mientras que en el estanque Lajitas sobresalen géneros pertenecientes a Cyanophyta. Estos géneros se encuentran en una frecuencia baja, sin embargo pueden tener gran impacto modulando la microbiota como en el caso de *Lactococcus* que es conocido en la promoción de microbiota sana (41).

Con los resultados de la diversidad  $\alpha$  se puede determinar que existe una mayor diversidad bacteriana en el sedimento donde habitan los camarones que en su intestino. Esto anteriormente se había reportado por Johnson y colaboradores en el 2009, los cuales mencionan que la diversidad del intestino es selectiva, por lo cual no todos los microorganismos presentes en el ambiente pueden encontrarse en el intestino a pesar de que los hábitos alimenticios del camarón incluyan el ingerir sustrato (1). Las curvas de rarefacción obtenidas (figuras 57-67) nos permiten concluir que se estima en buena medida la diversidad y riqueza de especies con la profundidad de secuenciación utilizada. Se puede observar que cuando se comparan las diversidades entre genéticas, hay una cambio entre la genética del intestino y de los sedimentos, esto es que no existe una relación directa entre la microbiota que pueda existir en el ambiente y la microbiota intestinal de *L.*

## DISCUSIÓN

---

*vannamei* (65-67).

Para el análisis de diversidad  $\beta$  se utilizó un análisis de coordenadas principales (figuras 70 y 71) en las cuales se puede observar que los sedimentos se agrupan en todos los PCoA, mientras que las muestras de intestino se agrupan cuando no se considera la abundancia (*unweighted*, figuras 71 (D) y 70 (D)). Al comparar los ambientes 1 y 2 (*unweighted*, figuras 71 (A y B) y 70 (A y B)) no se observa que se las muestras se agrupen con el PCoA. Para el agrupamiento de las muestras se utilizaron árboles de filogenia (figuras 72 y A.9). En el árbol que agrupa a las muestras considerando las abundancias de los OTU (*weighted*, figura 72) se puede observar que las muestras de sedimento e intestino se agrupan en ramas diferentes. Sin embargo, no se puede determinar cuál factor, si el ambiente o la genética, tiene un mayor impacto sobre la microbiota intestinal. Resultados similares se observa también en árbol que no toma en cuenta la abundancia de los OTU (*unweighted*, figura A.9), sin embargo la distancia entre los nodos, así como su nivel de confianza, son menores; además de que una muestra de intestino sale de la agrupación (IE9) por lo cual podemos inferir poco con la información que nos brinda este árbol.

Considerando los PCA realizados, se puede notar que a todos los niveles los sedimentos se agrupan totalmente, además de que los estanques Mazatlán y Estrella tienden a agruparse, los cuales cuentan con la misma genética (figuras 73-78), por lo cual con este análisis también se puede coincidir en que la genética tiene una mayor influencia sobre la microbiota que el ambiente. Por último la agrupación jerárquica en mapas de calor nos permite organizar a las muestras utilizando a sus frecuencias absolutas así como un gradiente de color. Se puede observar en los mapas de calor resultantes (figuras 85-90) que existe una agrupación parcial por genética 1 y 2, mas no se observa cuando se agrupan las muestras por ambientes (figuras 79-84).

## Conclusión

---

Uno de los objetivos de este trabajo fue la estandarización de un método para la extracción de ADN, así como la amplificación de la región V3-V4 del gen ARNr 16S, lo cual se logra exitosamente para las muestras de intestino de camarón. Además de que se demuestra que la base de datos que mejor caracteriza la microbiota intestinal de ésta especie es Greengenes.

La microbiota intestinal del camarón blanco del Pacífico (*Litopenaeus vannamei*) está dominada por los phyla Proteobacteria y Cyanobacteria; los géneros *Vibrio*, *Propionigenium* y *Cyanobacterium*. Con los resultados obtenidos se puede determinar que la diversidad y la riqueza de especies, calculadas con curvas de rarefacción con diferentes índices, son similares entre muestras de intestino y de igual forma entre muestras de sedimento (de los tres diferentes estanques), pero son diferentes entre tipo de muestra, siendo más diversa la microbiota del sedimento que la del intestino de *L. vannamei*. Esto concuerda con la selección en la microbiota intestinal previamente reportada (1).

Los PCoA realizados muestran que existe una tendencia a separar a la microbiota intestinal según su genética (figuras 71 (B) y 70 (B)). Se puede concluir de forma similar considerando el árbol de filogenia (figura A.9). De igual forma el mayor número de diferencias significativas con LEfSe se obtiene cuando se comparan a las genética 1 y 2 (figuras 16). Los PCA realizados (figuras 73-78) agrupan a las muestras por su genética, lo cual también nos permite concluir que existe un mayor peso sobre la microbiota intestinal; esto también se observa en los mapas de calor obtenidos al comparar las muestras por genética y ambiente (figuras 85-90).

Determinar el factor que modula a la diversidad microbiana intestinal del camarón blanco del Pacífico, el ambiente o la genética, no se logra completamente y se necesitarían más muestras para obtener un respuesta concluyente con más valores estadísticos que la respalden. Sin embargo con los análisis realizados se puede concluir que la microbiota intestinal en *L. vannamei* tiene una mayor influencia por su genética, el linaje de donde proviene, que por su ambiente.

Debido a la importancia comercial de esta especie de camarón a nivel mundial

## CONCLUSIÓN

es necesario realizar más estudios que busquen determinar su microbiota intestinal y con esto completar un esquema más amplio de su diversidad microbiana.

---

## Apéndice A

# Apéndices

---

### A.1. Cebadores utilizados

Reporte de síntesis					
OLIGO: 248					
Solicita: ADRIAN OCHOA		Grupo: AO		Dependencia: IBT	
Clave: V3-V4_fwd_MiSeq	Longitud: 50	Secuencia (5'→3'): TCGTCGGAGCGTCAGATGTATAAGAGACAGCCTACGGNGGCWGCAG			
Coef.: 559.75	NanoDrop A(260): 82.612	Volumen (μl): 500	OD's totales: 19.70	Conc.(ug/μl): 1.30	
C+G (%): 59.66	Tm(1): 79.88	Tm(2): 159.66	p.m.: 16374.30		picomol/μl: 79.39
<b>NOTA:</b> Para preparar 100 μL de una solución <b>10 uM</b> , tome 12.6 microlitros del OLIGO y agregue 87.4 microlitros de AGUA. $10\mu M = 10pmol/\mu L$					
Observaciones: OLIGO CRUDO. EXTRAIDO DE SOLUCION AMONIACAL c/ nBUTANOL. SECADO EN SAVANT y RESUSPENDIDO EN BUFFER DE ELUCION . UNA BANDA POR UV-SHADOWING EN GEL DE ACRILAMIDA 20% 7M UREA .					
EN LA UNIDAD DE SÍNTESIS Y SECUENCIACIÓN HACEMOS UN GRAN EFUERZO POR ENTREGAR TUS OLIGOS Y SECUENCIAS CON ALTA CALIDAD Y EFICIENCIA. CONSIDERALO EN TUS PUBLICACIONES. Atte. E. LOPEZ, S. BECERRA, J. YAÑEZ, P. GAYTAN					

**Figura A.1:** Cebador *forward* utilizado para seleccionar la región V3-V4 del gen 16S ribosomal.

Reporte de síntesis					
OLIGO: 249					
Solicita: ADRIAN OCHOA		Grupo: AO		Dependencia: IBT	
Clave: V3-V4_rev_MiSeq	Longitud: 55	Secuencia (5'→3'): GTCTCGGGCTGGAGATGTATAAGAGACAGGACTACHVGGGTATCTAACCC			
Coef.: 616.6	NanoDrop A(260): 78.416	Volumen (μl): 500	OD's totales: 17.34	Conc.(ug/μl): 1.14	
C+G (%): 50.31	Tm(1): 77.42	Tm(2): 165.34	p.m.: 17983.30		picomol/μl: 63.62
<b>NOTA:</b> Para preparar 100 μL de una solución <b>10 uM</b> , tome 15.7 microlitros del OLIGO y agregue 84.3 microlitros de AGUA. $10\mu M = 10pmol/\mu L$					
Observaciones: OLIGO CRUDO. EXTRAIDO DE SOLUCION AMONIACAL c/ nBUTANOL. SECADO EN SAVANT y RESUSPENDIDO EN BUFFER DE ELUCION . UNA BANDA POR UV-SHADOWING EN GEL DE ACRILAMIDA 20% 7M UREA .					
EN LA UNIDAD DE SÍNTESIS Y SECUENCIACIÓN HACEMOS UN GRAN EFUERZO POR ENTREGAR TUS OLIGOS Y SECUENCIAS CON ALTA CALIDAD Y EFICIENCIA. CONSIDERALO EN TUS PUBLICACIONES. Atte. E. LOPEZ, S. BECERRA, J. YAÑEZ, P. GAYTAN					

**Figura A.2:** Cebador *reverse* utilizado para seleccionar la región V3-V4 del gen 16S ribosomal.

### A.2. Código

Código utilizado para la asignación de taxonomía con QIIME (versión 1.9.0+dfsg-0biolinux5) en Biolinaux 8.

---

```
> qiime
> pick_closed_reference_otus.py -p $PWD/parameters.txt -i $PWD/
    seqs.fna -o $PWD/pick_closed_greengenes -a -O 60
> pick_rep_set.py -i $PWD/pick_closed_greengenes/
    uclust_ref_picked_otus/seqs.txt -f $PWD/seqs_LIMPIO.fna -o
    $PWD/pick_closed_greengenes/seqs_rep_set.fna -l $PWD/
    pick_closed_greengenes/uclust_ref_picked_otus/seqs.log
> align_seqs.py -i $PWD/pick_closed_greengenes/seqs_rep_set.fna -
    o $PWD/pick_closed_greengenes/pynast_aligned_seqs/
> filter_alignment.py -i $PWD/pick_closed_greengenes/
    pynast_aligned_seqs/seqs_rep_set_aligned.fasta -o $PWD/
    pick_closed_greengenes/pynast_aligned_seqs/filtered_alignment/
> make_phylogeny.py -i $PWD/pick_closed_greengenes/
    pynast_aligned_seqs/seqs_rep_set_aligned.fasta -o $PWD/
    pick_closed_greengenes/rep_phylo.tre
> filter_otus_from_otu_table.py -i $PWD/pick_closed_greengenes/
    otu_table.biom -o $PWD/pick_closed_greengenes/
    otu_table_no_singletons.biom -n 2
> biom summarize-table -i $PWD/pick_closed_greengenes/
    otu_table_no_singletons.biom/otu_table_no_singletons.biom -o
    $PWD/pick_closed_greengenes/summary_biom
> make_otu_network.py -m $PWD/datos_pretratados/map_beta_.txt -i
    $PWD/pick_closed_greengenes/otu_table_no_singletons.biom/
    otu_table_no_singletons.biom -o $PWD/pick_closed_greengenes/
    otu_network
> summarize_taxa.py -i $PWD/pick_closed_greengenes/
    otu_table_no_singletons.biom -o $PWD/pick_closed_greengenes/
    taxa_sp -L 2,3,4,5,6,7
> cd pick_closed_greengenes/taxa_sp
> plot_taxa_summary.py -i otu_table_no_singletons_L2.txt,
    otu_table_no_singletons_L3.txt,otu_table_no_singletons_L4.txt,
    otu_table_no_singletons_L5.txt,otu_table_no_singletons_L6.txt,
```

---

```
otu_table_no_singletons_L7.txt -o ../taxa_sp_plot
```

Código utilizado para el análisis de diversidad  $\alpha$ .

```
> qiime  
> alpha_rarefaction.py -a -O 60 -i $PWD/pick_closed_greengenes/  
    otu_table_no_singletons.biom -t $PWD/pick_closed_greengenes/  
    rep_phylo.tre -m $PWD/map.txt -o $PWD/pick_closed_greengenes/  
    alpha_rarefaction -p $PWD/alpha_params.txt  
> compare_alpha_diversity.py -i $PWD/alpha_rarefaction/  
    alpha_div_collated/chao1.txt -m $PWD/map.txt -c Treatment,  
    Description -o $PWD/alpha_rarefaction/compare_alpha_chao1 -d  
    1638  
> compare_alpha_diversity.py -i $PWD/alpha_rarefaction/  
    alpha_div_collated/shannon.txt -m $PWD/map.txt -c Treatment,  
    Description -o $PWD/alpha_rarefaction/compare_alpha_shannon -d  
    1638  
> compare_alpha_diversity.py -i $PWD/alpha_rarefaction/  
    alpha_div_collated/simpson.txt -m $PWD/map.txt -c Treatment,  
    Description -o $PWD/alpha_rarefaction/compare_alpha_simpson -d  
    1638  
> compare_alpha_diversity.py -i $PWD/alpha_rarefaction/  
    alpha_div_collated/observed_otus.txt -m $PWD/map.txt -c  
    Treatment,Description -o $PWD/alpha_rarefaction/  
    compare_alpha_observed_otus -d 1638  
> compare_alpha_diversity.py -i $PWD/alpha_rarefaction/  
    alpha_div_collated/goods_coverage.txt -m $PWD/map.txt -c  
    Treatment,Description -o $PWD/alpha_rarefaction/  
    compare_alpha_goods_coverage -d 1638  
> compare_alpha_diversity.py -i $PWD/alpha_rarefaction/  
    alpha_div_collated/PD_whole_tree.txt -m $PWD/map.txt -c  
    Treatment,Description -o $PWD/alpha_rarefaction/  
    compare_alpha_PD_whole_tree -d 1638
```

Código utilizado para el análisis de diversidad  $\beta$ .

---

```

> qiime
> beta_diversity_through_plots.py -i $PWD/pick_closed_greengenes/
    otu_table_no_singletons.biom -m $PWD/map.txt -t $PWD/
    pick_closed_greengenes/rep_phylo.tre -e 1638 -o $PWD/
    pick_closed_greengenes/beta_div
> make_2d_plots.py -i $PWD/pick_closed_greengenes/beta_div/
    weighted_unifrac_pc.txt -m $PWD/map.txt -o $PWD/
    pick_closed_greengenes/beta_div/weighted_unifrac_2d
> make_2d_plots.py -i $PWD/pick_closed_greengenes/beta_div/
    unweighted_unifrac_pc.txt -m $PWD/map.txt -o $PWD/
    pick_closed_greengenes/beta_div/unweighted_unifrac_2d

```

Código utilizado para generar los árboles filogenéticos.

```

> beta_diversity.py -i $PWD/pick_closed_reference/
    otu_table_no_singletons.biom -o $PWD/pick_closed_reference/
    jack_manual/beta_div_otu_table/ -t $PWD/pick_closed_reference/
    rep_phylo.tre
> upgma_cluster.py -i $PWD/pick_closed_reference/jack_manual/
    beta_div_otu_table/ -o $PWD/pick_closed_reference/jack_manual/
    beta_div_otu_table_results/
> multiple_rarefactions_even_depth.py -i $PWD/
    pick_closed_reference/otu_table_no_singletons.biom -o $PWD/
    pick_closed_reference/jack_manual/rarefied_otu_tables_1000/ -d
    1638 -n 1000
> beta_diversity.py -i $PWD/pick_closed_reference/jack_manual/
    rarefied_otu_tables_1000/ -m $PWD/pick_closed_reference/
    weighted_unifrac -o $PWD/pick_closed_reference/jack_manual/
    beta_div_weighted/ -t $PWD/pick_closed_reference/rep_phylo.tre
> upgma_cluster.py -i $PWD/pick_closed_reference/jack_manual/
    beta_div_weighted/ -o $PWD/pick_closed_reference/jack_manual/
    beta_div_weighted_results/
> tree_compare.py -m $PWD/pick_closed_reference/jack_manual/
    beta_div_otu_table_results/

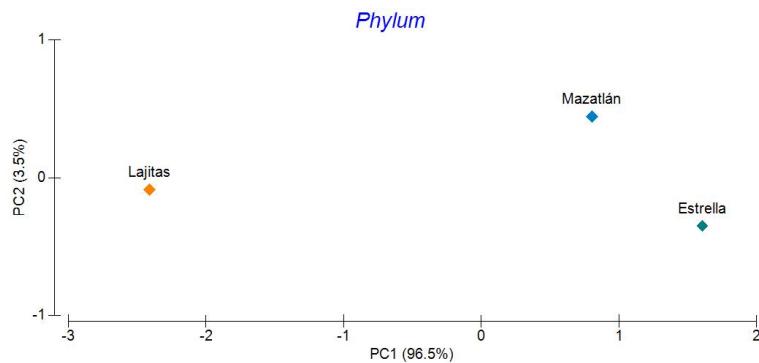
```

---

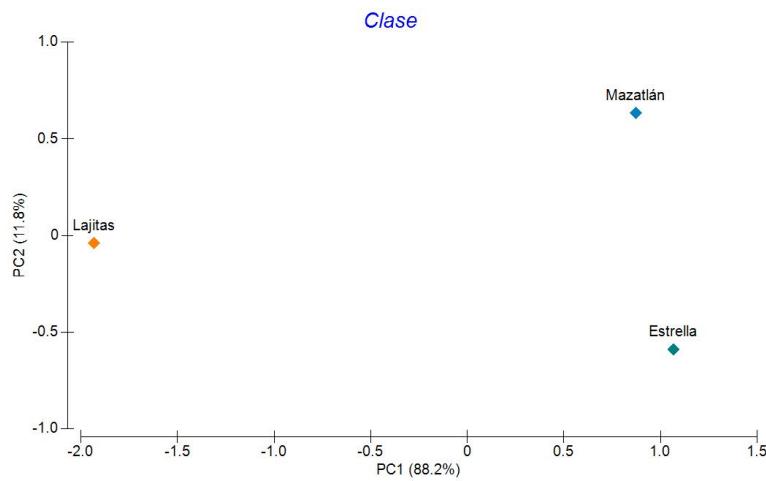
```
upgma_weighted_unifrac_otu_table_no_singletons.tre -s $PWD/
pick_closed_reference/jack_manual/beta_div_weighted_results/ -
o $PWD/pick_closed_reference/jack_manual/
beta_div_weighted_tree_compared/
> make_bootstrapped_tree.py -m $PWD/pick_closed_reference/
jack_manual/beta_div_weighted_tree_compared/master_tree.tre -s
$PWD/pick_closed_reference/jack_manual/
beta_div_weighted_tree_compared/jackknife_support.txt -o $PWD/
pick_closed_reference/jack_manual/
beta_div_weighted_tree_compared/jackknife_weighted_named_nodes.
pdf
> beta_diversity.py -i $PWD/pick_closed_reference/jack_manual/
rarefied_otu_tables_1000/ -m $PWD/pick_closed_reference/
unweighted_unifrac -o $PWD/pick_closed_reference/jack_manual/
beta_div_unweighted/ -t $PWD/pick_closed_reference/rep_phylo.
tre
> upgma_cluster.py -i $PWD/pick_closed_reference/jack_manual/
beta_div_unweighted/ -o $PWD/pick_closed_reference/jack_manual
/beta_div_unweighted_results/
> tree_compare.py -m $PWD/pick_closed_reference/jack_manual/
beta_div_otu_table_results/
upgma_unweighted_unifrac_otu_table_no_singletons.tre -s $PWD/
pick_closed_reference/jack_manual/beta_div_unweighted_results/ -
o $PWD/pick_closed_reference/jack_manual/
beta_div_unweighted_tree_compared/
> make_bootstrapped_tree.py -m $PWD/pick_closed_reference/
jack_manual/beta_div_unweighted_tree_compared/master_tree.tre
-s $PWD/pick_closed_reference/jack_manual/
beta_div_unweighted_tree_compared/jackknife_support.txt -o
$PWD/pick_closed_reference/jack_manual/
beta_div_unweighted_tree_compared/
jackknife_unweighted_named_nodes.pdf
```

---

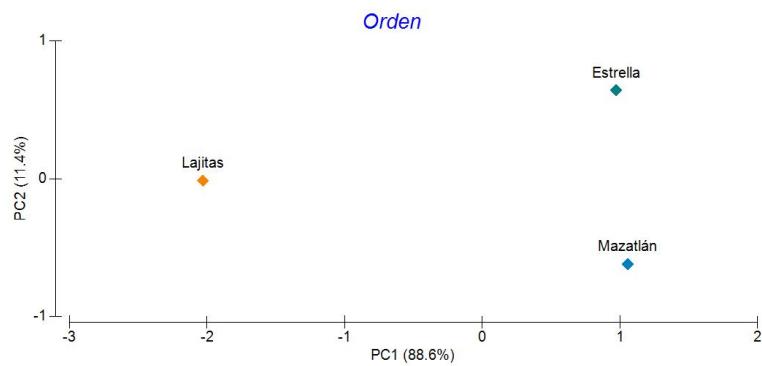
### A.3. Figuras supplementarias



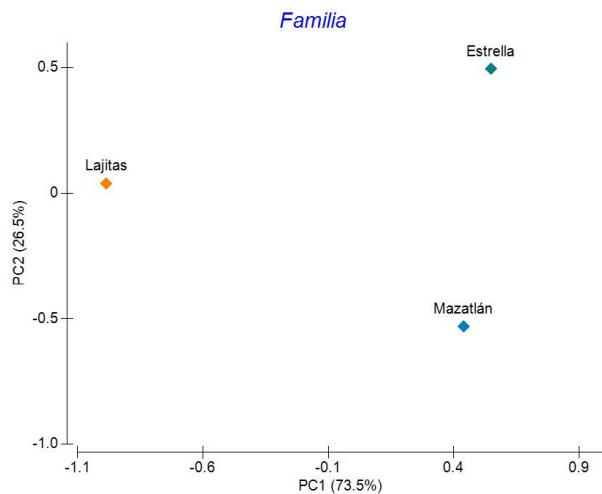
**Figura A.3:** PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de phylum.



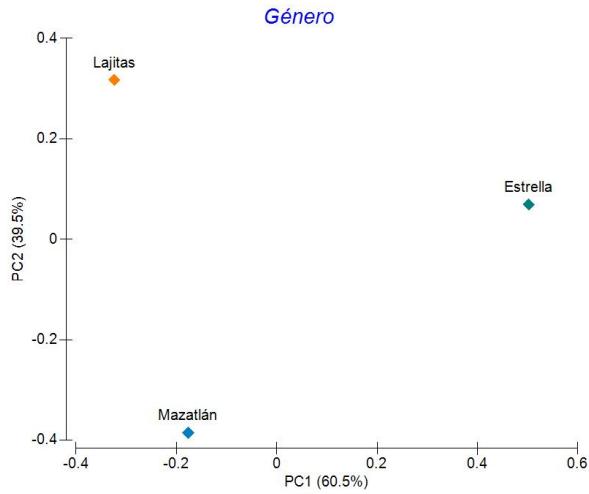
**Figura A.4:** PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de clase.



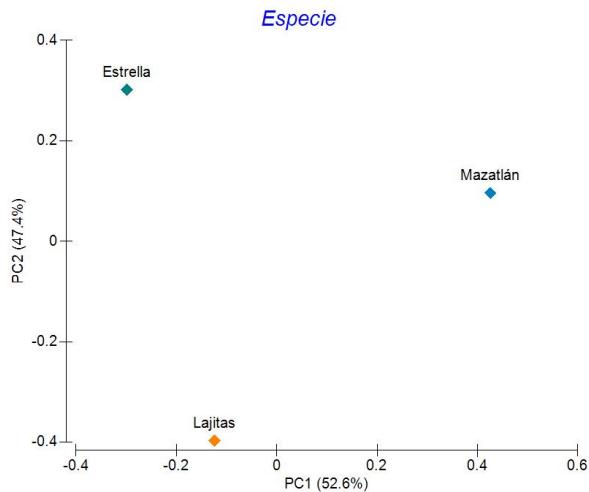
**Figura A.5:** PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de orden.



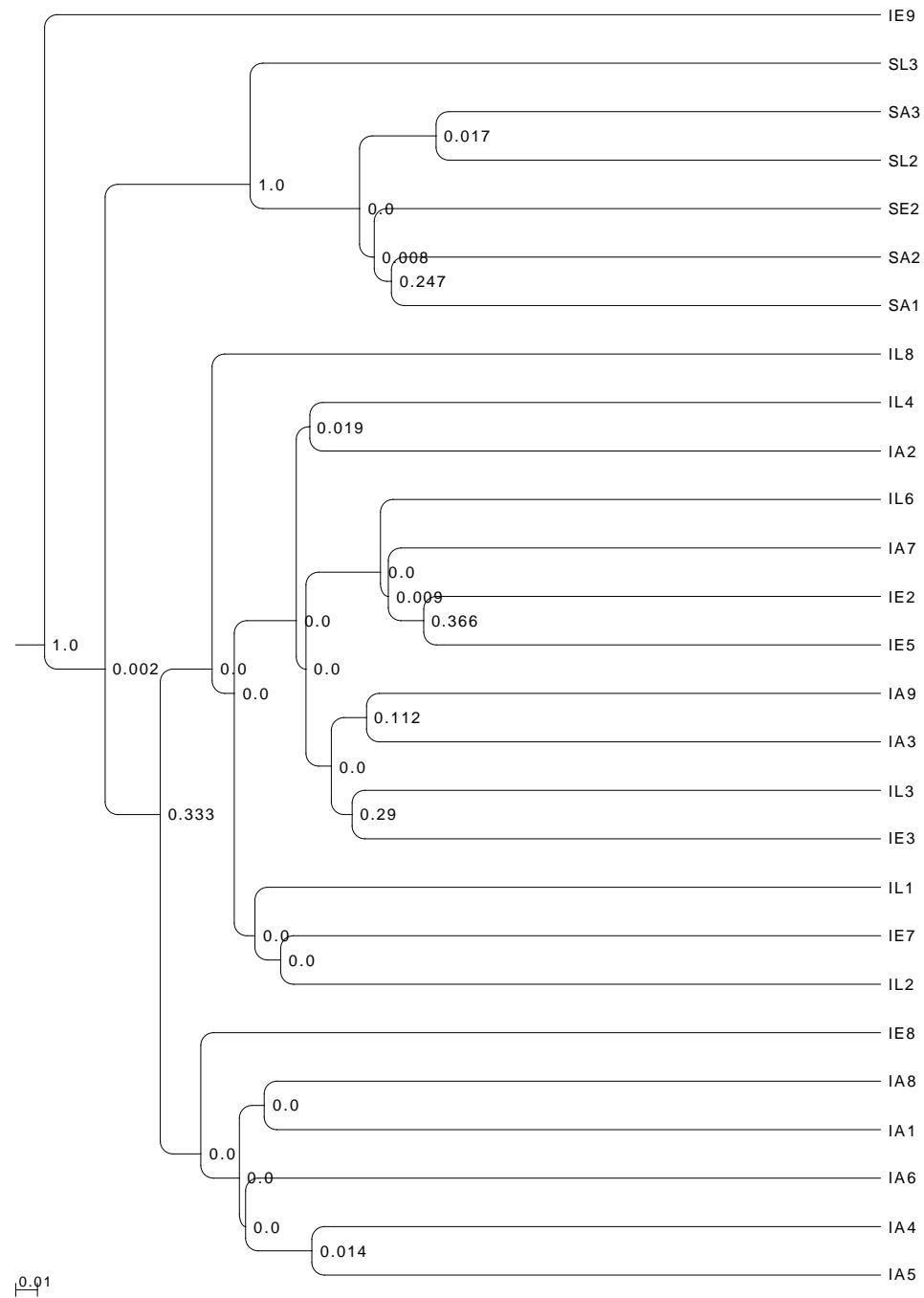
**Figura A.6:** PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de familia.



**Figura A.7:** PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de género.

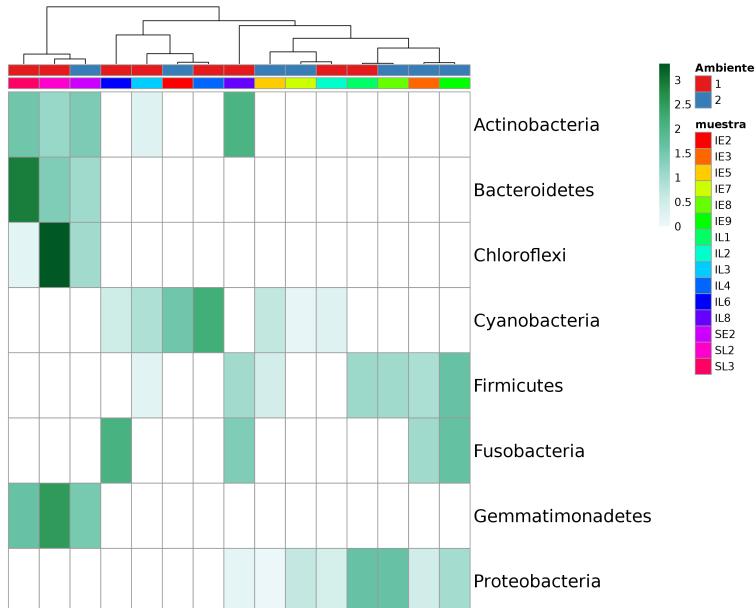


**Figura A.8:** PCoA con las muestras agrupadas por estanques (Estrella, Lajitas y Mazatlán) con los OTU a nivel de especie.

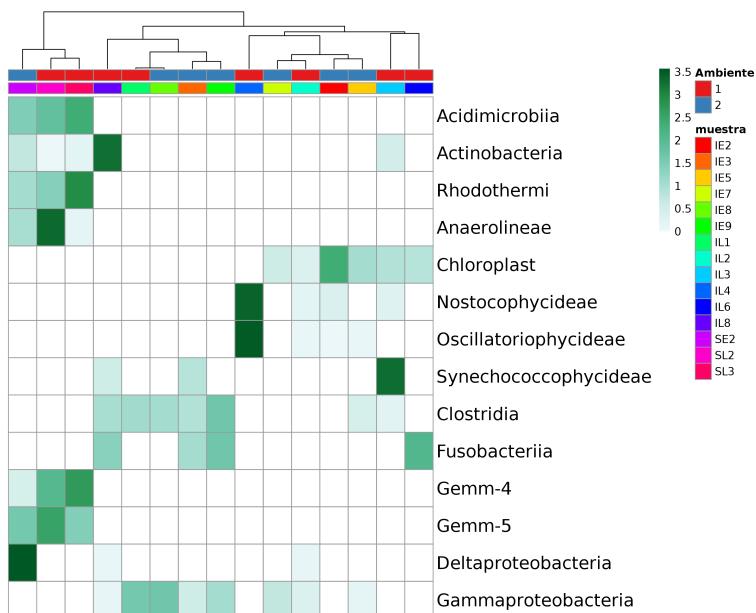


**Figura A.9:** Árbol generado con Jackknife no ponderado (*unweighted*); cada nodo cuenta con su valor confianza.

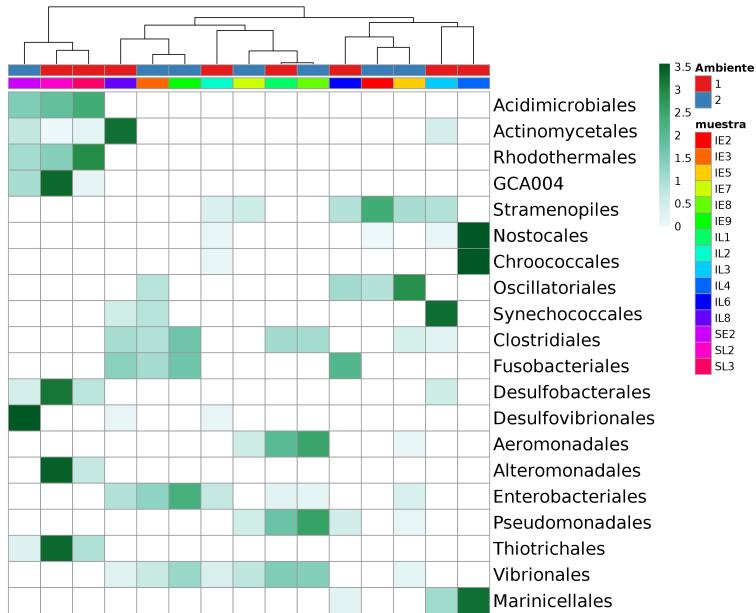
## Clustering jerárquico ambiente



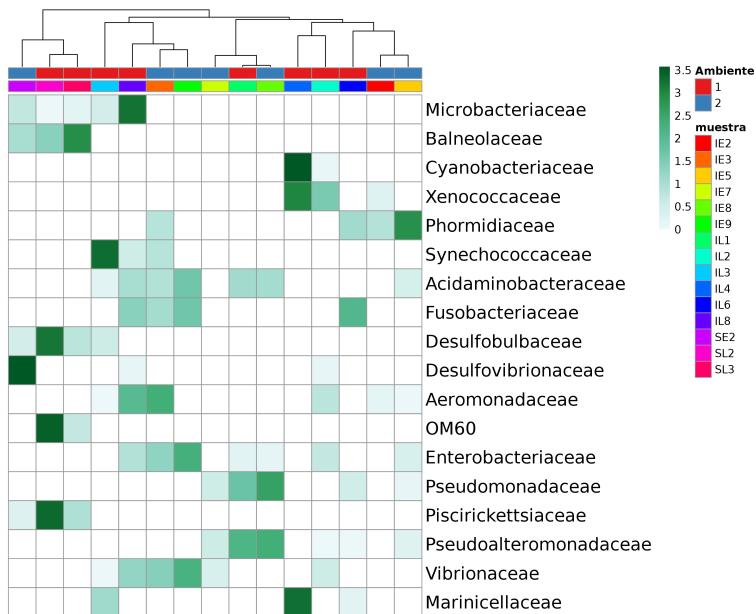
**Figura A.10:** Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel phylum.



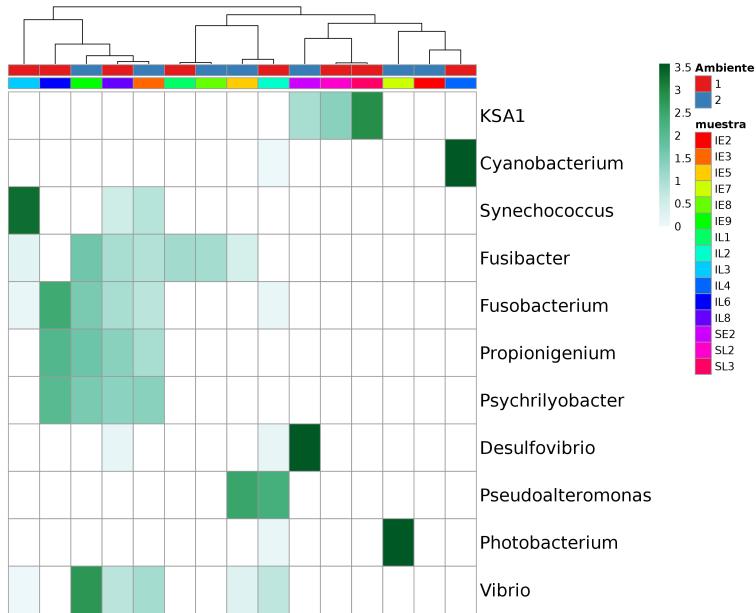
**Figura A.11:** Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel clase.



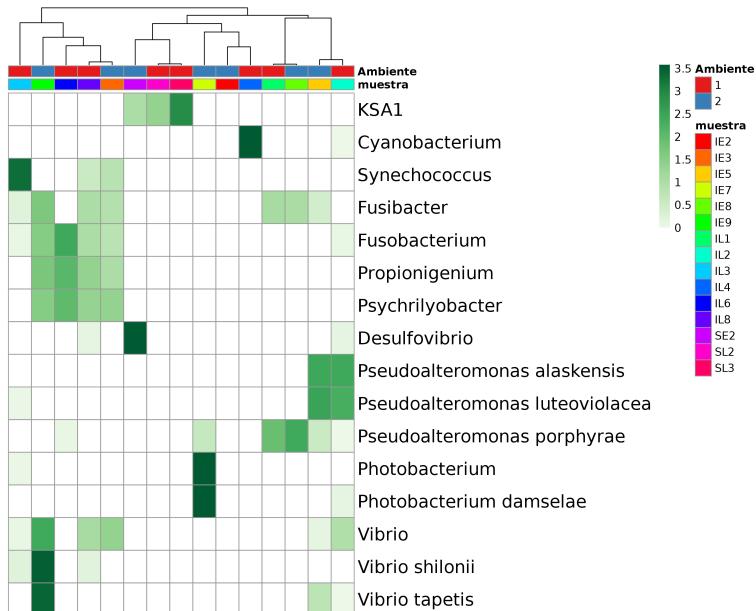
**Figura A.12:** Mapa de calor entre muestras del ambiente 1 (estanque Maatlán) y el ambiente 2 (estanque Estrella) a nivel orden.



**Figura A.13:** Mapa de calor entre muestras del ambiente 1 (estanque Maatlán) y el ambiente 2 (estanque Estrella) a nivel familia.

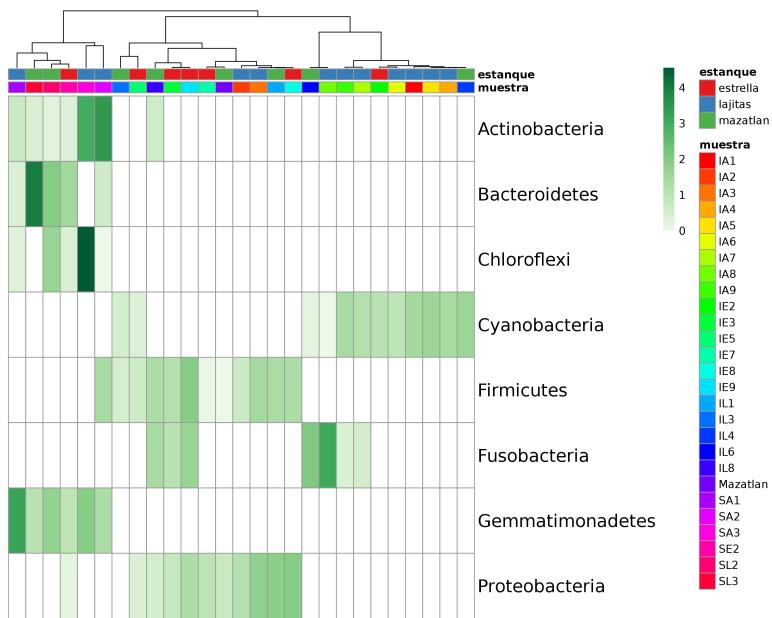


**Figura A.14:** Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel género.

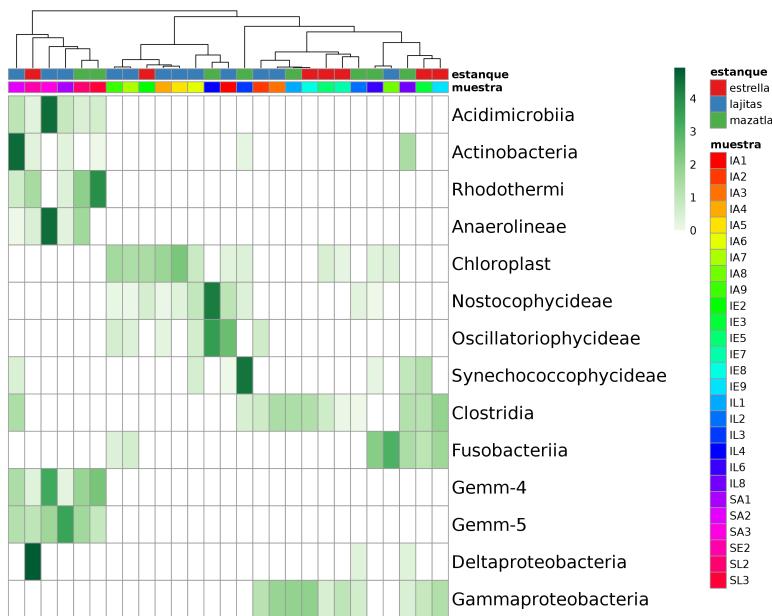


**Figura A.15:** Mapa de calor entre muestras del ambiente 1 (estanque Mazatlán) y el ambiente 2 (estanque Estrella) a nivel especie.

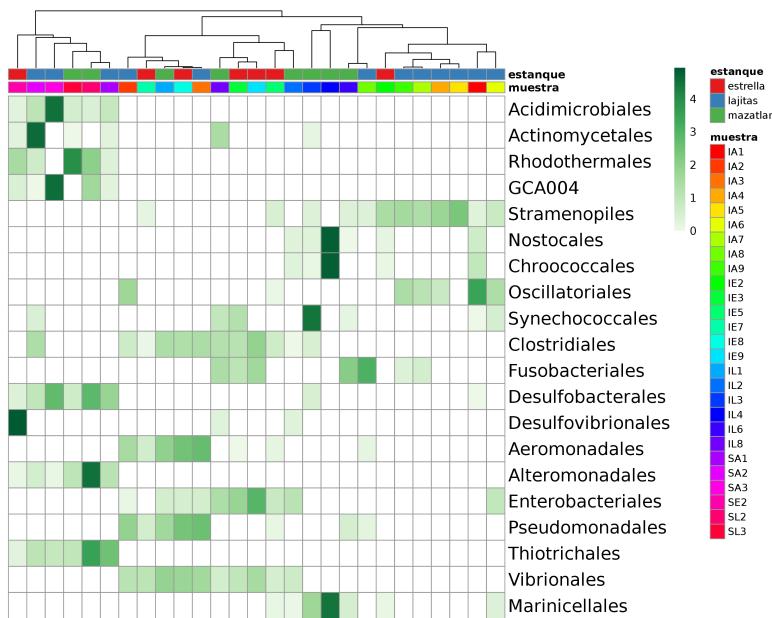
Clustering jerárquico por estanque



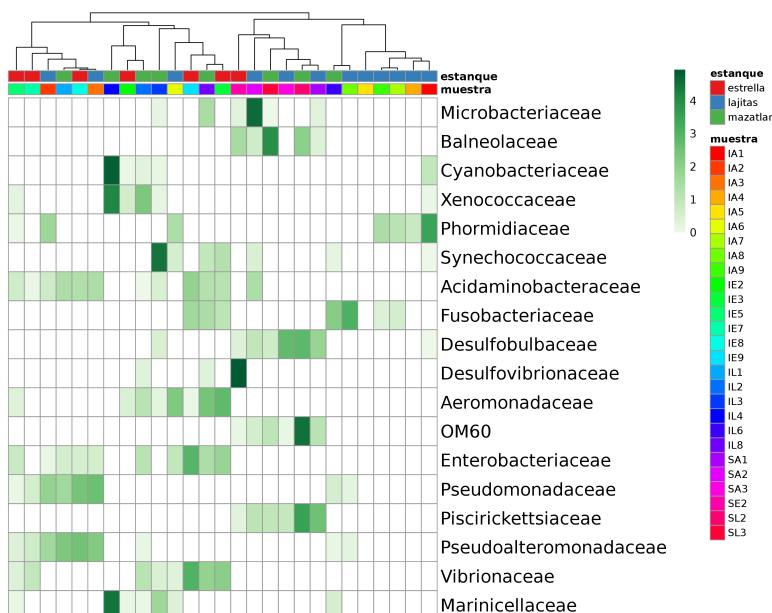
**Figura A.16:** Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel phylum.



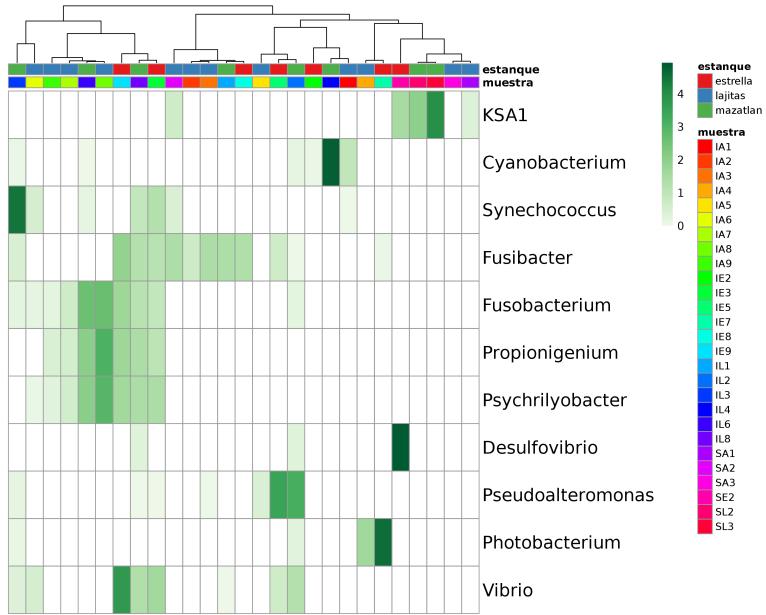
**Figura A.17:** Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel clase.



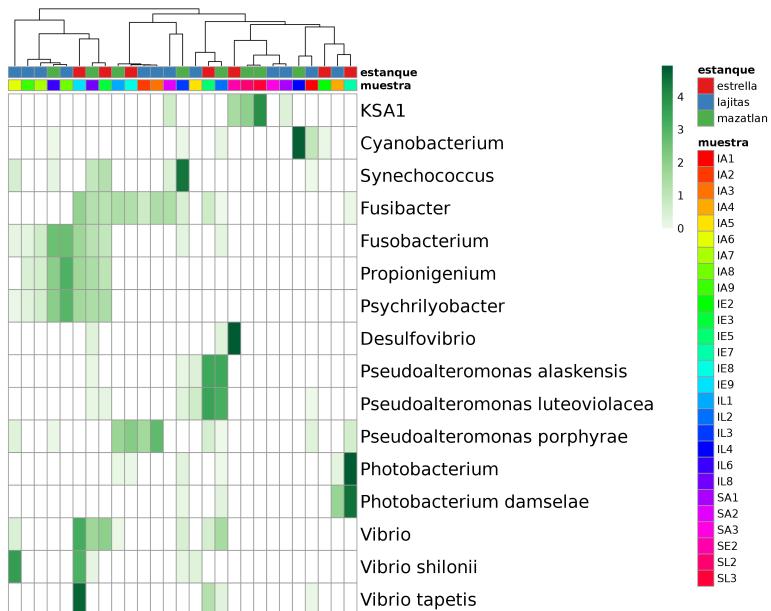
**Figura A.18:** Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel orden.



**Figura A.19:** Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel familia.



**Figura A.20:** Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel género.



**Figura A.21:** Mapa de calor entre muestras de la genética 1 (estanque Mazatlán) y el genética 2 (estanque Lajitas) a nivel especie.

---

## Referencias

- [1] Crystal N. Johnson, Sue Barnes, John Ogle, D. Jay Grimes, Yun Juan Chang, Aaron D. Peacock, and Liz Kline. Microbial community analysis of water, foregut, and hindgut during growth of pacific white shrimp, *Litopenaeus vannamei*, in closed-system aquaculture. *Journal of the World Aquaculture Society*, 39(2):251–258, 2008. [1](#), [4](#), [15](#), [110](#), [112](#)
- [2] Hao Zhang, Zhenli Sun, Bo Liu, Yinming Xuan, Ming Jiang, Yunsheng Pan, Yiming Zhang, Yeping Gong, Xiaoping Lu, Deshan Yu, Dhiraj Kumar, Xiaolong Hu, Guangli Cao, Renyu Xue, and Chengliang Gong. Dynamic changes of microbial communities in *Litopenaeus vannamei* cultures and the effects of environmental factors. *Aquaculture*, 455:97–108, 2016. [1](#), [4](#), [109](#)
- [3] Fisheries and Aquaculture Department. Fact sheets: *Litopenaeus vannamei*. Technical report, Food and Agriculture Organization of the United Nations, 2014. [1](#), [3](#), [4](#)
- [4] SAGARPA. Anuario Estadístico de Acuacultura y pesca 2013. *Anuario estadístico de Acuacultura y Pesca 2013*, 53(9):1689–1699, nov 2013. [1](#)
- [5] Fernanda Cornejo-Granados, Alonso A. Lopez-Zavala, Luigui Gallardo-Becerra, Alfredo Mendoza-Vargas, Filiberto Sánchez, Rodrigo Vichido, Luis G. Brieba, Maria Teresa Viana, Rogerio R. Sotelo-Mundo, and Adrián Ochoa-Leyva. Microbiome of Pacific Whiteleg shrimp reveals differential bacterial community composition between Wild, Aquacultured and AHPN-D/EMS outbreak conditions. *Scientific Reports*, 7(1):11783, dec 2017. [1](#), [16](#), [109](#)
- [6] Shirley Bikel, Alejandra Valdez-Lara, Fernanda Cornejo-Granados, Karina Rico, Samuel Canizales-Quinteros, Xavier Soberón, Luis Del Pozo-Yauner, and Adrián Ochoa-Leyva. Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: Towards a systems-level understanding of human microbiome. *Computational and Structural Biotechnology Journal*, 13:390–401, 2015. [1](#), [7](#), [9](#), [10](#), [11](#), [26](#), [109](#)
- [7] Man Kit Cheung, Ho Yin Yip, Wenyan Nong, Patrick Tik Wan Law, Ka Hou Chu, Hoi Shan Kwan, and Jerome Ho Lam Hui. Rapid Change of Microbiota Diversity in the Gut but Not the Hepatopancreas During Gonadal Development of the New Shrimp Model *Neocaridina denticulata*. *Marine Biotechnology*, 17(6):811–819, 2015. [1](#)

- 
- [8] Wanilada Rungrassamee, Amornpan Klanchui, Sawarot Maibunkaew, and Nitsara Karoonuthaisiri. Bacterial dynamics in intestines of the black tiger shrimp and the Pacific white shrimp during *Vibrio harveyi* exposure. *Journal of Invertebrate Pathology*, 133:12–19, 2015. [1](#), [3](#), [16](#), [109](#)
  - [9] Alexander V. Vlassov. The Human Microbiota: Composition, Functions, and Therapeutic Potential. *Medical Science Review*, 2:92–103, 2015. [1](#)
  - [10] Meiling Zhang, Yuhong Sun, Ke Chen, Na Yu, Zhigang Zhou, Liqiao Chen, Zhenyu Du, and Erchao Li. Characterization of the intestinal microbiota in Pacific white shrimp, *Litopenaeus vannamei*, fed diets with different lipid sources. *Aquaculture*, 434:449–455, 2014. [109](#)
  - [11] Jian Xu. *Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches*, chapter Microbial Ecology in the Age of Metagenomics. Wiley-Blackwell, Hoboken, New Jersey, 2011. [1](#), [7](#), [8](#)
  - [12] Nuno Simões, David Jones, Sonia Soto-rodríguez, and Ana Roque. Las Bacterias en el Inicio de la Alimentación Exógena en Larvas de Camarones Penídeos : Efectos de la Calidad del Agua , Tasas de Ingestión y Rutas de Colonización del Tracto Digestivo. *Memorias del VI Simposium International de Nutrición Acuícola*, pages 243–276, 2002. [4](#)
  - [13] J Handelsman, J Tiedje, National Research Council (US) Committee on Metagenomics: Challenges and Functional, , and National Research Council (Us) Committee on Metagenomics: Challenges Functional. *THE NEW SCIENCE OF METAGENOMICS: Revealing the Secrets of Our Microbial Planet.* 2007. [7](#)
  - [14] C R Woese and G E Fox. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, 74(11):5088–5090, 1977. [7](#), [8](#)
  - [15] P. C Y Woo, S. K P Lau, J. L L Teng, H. Tse, and K. Y. Yuen. Then and now: Use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infection*, 14(10):908–934, 2008. [7](#)
  - [16] M G Langille, J Zaneveld, J G Caporaso, D McDonald, D Knights, J A Reyes, J C Clemente, D E Burkepile, R L Vega Thurber, R Knight, R G Beiko, and C Huttenhower. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol*, 31(9):814–821, 2013. [7](#)

- 
- [17] Pablo Yarza, Pelin Yilmaz, Elmar Pruesse, Frank Oliver Glöckner, Wolfgang Ludwig, Karl-Heinz Schleifer, William B. Whitman, Jean Euzéby, Rudolf Amann, and Ramon Rosselló-Móra. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews. Microbiology*, 12(9):635–645, 2014. [8](#)
  - [18] Jennifer J. Barb, Andrew J. Oler, Hyung-Suk Kim, Natalia Chalmers, Gwenyth R. Wallen, Ann Cashion, Peter J. Munson, and Nancy J. Ames. Development of an Analysis Pipeline Characterizing Multiple Hypervariable Regions of 16S rRNA Using Mock Samples. *PLOS ONE*, 11(2):e0148047, feb 2016. [9](#)
  - [19] Ali May, Sanne Abeln, Wim Crielaard, Jaap Heringa, and Bernd W. Brandt. Unraveling the outcome of 16S rDNA-based taxonomy analysis through mock data and simulations. *Bioinformatics*, 30(11):1530–1538, 2014. [9](#), [10](#), [26](#)
  - [20] J Gregory Caporaso, Christian L Lauber, William a Walters, Donna Berg-Lyons, James Huntley, Noah Fierer, Sarah M Owens, Jason Betley, Louise Fraser, Markus Bauer, Niall Gormley, Jack a Gilbert, Geoff Smith, and Rob Knight. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, 6(8):1621–1624, 2012. [9](#)
  - [21] Alejandra Escobar-Zepeda, Arturo Vera-Ponce de León, and Alejandro Sanchez-Flores. The Road to Metagenomics: From Microbiology to DNA Sequencing Technologies and Bioinformatics. *Frontiers in Genetics*, 6(December):348, 2015. [9](#), [26](#)
  - [22] B Ewing, L D Hillier, and M C Wendl. Base-Calling of Automated Sequencer Traces UsingPhred. II. Error Probabilities. *Genome Research*, 8(3):186–194, 1998. [10](#)
  - [23] Juan Jovel. Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Evolutionary and Genomic Microbiology*, 7(April):459, 2016. [11](#), [13](#)
  - [24] Stuart H Hurlbert. the Nonconcept of Species Diversity: a Critique and Alternative Parameters. *Ecology*, 52(4):577–&, 1971. [11](#), [12](#)
  - [25] N J Gotelli and R K Colwell. Quantifyinf biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4 Abstract(4):379–391, 2001. [12](#)

- 
- [26] C. E. Shannon and W. Weaver. The Mathematical Theory of Communication. *The mathematical theory of communication*, 27(4):117, 1949. [12](#)
  - [27] E. H. Simpson. Measurement of Diversity, 1949. [12](#)
  - [28] Catherine Lozupone, Manuel E Lladser, Dan Knights, Jesse Stombaugh, and Rob Knight. UniFrac: an effective distance metric for microbial community comparison. *The ISME journal*, 5(2):169–172, 2011. [13](#)
  - [29] Catherine Lozupone and Rob Knight. UniFrac : a New Phylogenetic Method for Comparing Microbial Communities UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. *Applied and environmental microbiology*, 71(12):8228–8235, 2005. [13](#)
  - [30] Wanilada Rungrassamee, Amornpan Klanchui, Sawarot Maibunkaew, and Sage Chaiyapechara. Characterization of Intestinal Bacteria in Wild and Domesticated Adult Black Tiger Shrimp ( *Penaeus monodon* ). 9(3), 2014. [15, 109](#)
  - [31] Jinbo Xiong, Kai Wang, Jinfeng Wu, Linglin Qiuqian, Kunjie Yang, Yunxia Qian, and Demin Zhang. Changes in intestinal bacterial communities are closely associated with shrimp disease severity. *Applied Microbiology and Biotechnology*, 99(16):6911–6919, 2015. [16](#)
  - [32] Simon Andrews. Fastqc: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>, 2010. [26](#)
  - [33] T. Z. DeSantis, P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen. Greengenes, a chimerachecked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, 72(7):5069–5072, 2006. [28, 109](#)
  - [34] James R. Cole, Qiong Wang, Jordan A. Fish, Benli Chai, Donna M. McGarrell, Yanni Sun, C. Titus Brown, Andrea Porras-Alfaro, Cheryl R. Kuske, and James M. Tiedje. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, 42(D1):633–642, 2014. [28, 109](#)
  - [35] Christian Quast, Elmar Pruesse, Pelin Yilmaz, Jan Gerken, Timmy Schweer, Pablo Yarza, Jörg Peplies, and Frank Oliver Glöckner. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1):590–596, 2013. [28, 109](#)

- 
- [36] J Gregory Caporaso, Justin Kuczynski, Jesse Stombaugh, Kyle Bittinger, Frederic D Bushman, Elizabeth K Costello, Noah Fierer, Antonio Gonzalez Peña, Julia K Goodrich, Jeffrey I Gordon, Gavin a Huttley, Scott T Kelley, Dan Knights, Jeremy E Koenig, Ruth E Ley, Catherine a Lozupone, Daniel McDonald, Brian D Muegge, Meg Pirrung, Jens Reeder, Joel R Sevinsky, Peter J Turnbaugh, William a Walters, Jeremy Widmann, Tanya Yatsunenko, Jesse Zaneveld, and Rob Knight. correspondence QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nature Publishing Group*, 7(5):335–336, 2010. [28](#), [109](#)
- [37] Wanilada Rungrassamee, Amornpan Klanchui, Sage Chaiyapechara, Sawarat Maibunkaew, Sithichoke Tangphatsornruang, Pikul Jiravanichpaisal, and Nitsara Karoonuthaisiri. Bacterial Population in Intestines of the Black Tiger Shrimp ( *Penaeus monodon* ) under Different Growth Stages. 8(4), 2013. [109](#)
- [38] Wanilada Rungrassamee, Amornpan Klanchui, Sage Chaiyapechara, Sawarat Maibunkaew, Sithichoke Tangphatsornruang, Pikul Jiravanichpaisal, and Nitsara Karoonuthaisiri. Bacterial Population in Intestines of the Black Tiger Shrimp ( *Penaeus monodon* ) under Different Growth Stages. *PLoS ONE*, 8(4):e60802, 2013. [109](#)
- [39] H.M. Essam, G.S. Abdellrazeq, S.I. Tayel, H.A. Torky, and A.H. Fadel. Pathogenesis of *Photobacterium damsela* subspecies infections in sea bass and sea bream. *Microbial Pathogenesis*, 99:41–50, oct 2016. [110](#)
- [40] Jacqueline Tzuc, Diana Escalante, Rafael Rojas Herrera, Gabriela Gaxiola Cortés, and Maria Ortiz. Microbiota from *Litopenaeus vannamei*: digestive tract microbial community of Pacific white shrimp ( *Litopenaeus vannamei* ). *SpringerPlus*, 3(1):280, 2014. [110](#)
- [41] Julian R Marchesi, David H Adams, Francesca Fava, Gerben D a Hermes, Gideon M Hirschfield, Georgina Hold, Mohammed Nabil Quraishi, James Kinross, Hauke Smidt, Kieran M Tuohy, Linda V Thomas, Erwin G Zoetendal, and Ailsa Hart. The gut microbiota and host health: a new clinical frontier. *Gut*, 65(2):330–339, feb 2016. [110](#)