

La investigación reportada en esta tesis es parte de los programas de investigación del CICESE (Centro de Investigación Científica y de Educación Superior de Ensenada, B.C.).

La investigación fue financiada por el CONACYT (Consejo Nacional de Ciencia y Tecnología).

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México). El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo o titular de los Derechos Autor.

CICESE © 2022, Todos los Derechos Reservados, CICESE

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



Maestría en Ciencias en Ecología Marina

Análisis de diversidad bacteriana y funciones fototróficas en la zona superficial de ambientes oligotróficos y productivos

Tesis

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

Presenta:

Alberto Netzahualcoyotl Arroyo Cervantes

Ensenada, Baja California, México

2022

Tesis defendida por

Alberto Netzahualcoyotl Arroyo Cervantes

y aprobada por el siguiente Comité

Dra. Laura Gómez Consarnau

Codirectora de tesis

Dra. María Asunción Lago Lestón

Codirectora de tesis

Dr. Helmut Maske Rubach

Dr. Edgardo Alfredo Sepúlveda Sánchez Hidalgo



Dr. Rafael Andrés Cabral Tena
Coordinador del Posgrado en Ecología Marina

Dr. Pedro Negrete Regagnon
Director de Estudios de Posgrado

Resumen de la tesis que presenta Alberto Netzahualcoyotl Arroyo Cervantes como requisito parcial para la obtención del grado de Maestro en Ciencias en Ecología Marina.

Análisis de diversidad bacteriana y funciones fototróficas en la zona superficial de ambientes oligotróficos y productivos

Resumen aprobado por:

Dra. Laura Gómez Consarnau

Codirector de tesis

Dra. María Asunción Lago Lestón

Codirector de tesis

La fototrofia con proteorodopsinas es el tipo de fotoheterotrofia más ampliamente extendido en las comunidades bacterianas de la superficie del océano. Las tecnologías de secuenciación han permitido identificar a los grupos productores de proteorodopsinas dentro de las comunidades microbianas de manera cualitativa, mientras que las mediciones de concentraciones de retinal permiten cuantificar el nivel de fotoheterotrofia en un ambiente. Se ha observado que la fototrofia basada en retinal domina en regiones altamente oligotróficas siendo de gran importancia para el ecosistema, mientras que en sistemas productivos existe poca información al respecto. En el presente estudio se realiza un análisis comparativo de las comunidades microbianas marinas en regiones con distintos niveles de productividad usando secuencias del gen 16S-ARNr para analizar la distribución de grupos fotoheterótrofos conocidos (e.g. SAR11, Flavobacteriales, SAR86) en dichos ambientes y su relación con mediciones de pigmentos y otros factores ambientales. Los datos analizados provienen de muestras de las regiones del Mediterráneo (crucero HOTMIX), del golfo de México (campañas XIXIMI-5,6,7) y la corriente de California (punto de muestreo SPOT). Para el Mediterráneo y la corriente de California se utiliza el conteo de heterótrofos totales por citometría para obtener una estimación de la abundancia de grupos taxonómicos de interés a nivel de orden, con el fin de aportar más información a la distribución presentada por dichos grupos. Para comparar la distribución de grupos productores de proteorodopsinas en los distintos ambientes se realizaron análisis de beta diversidad, análisis multivariados con las mediciones de parámetros, análisis de correlaciones y comparaciones directas entre las secuencias de cada región. Los análisis de beta diversidad demostraron que las comunidades microbianas del oeste del Mediterráneo y de la corriente de California presentan poca distancia filogenética, además de que ambas regiones presentan una correlación positiva entre la abundancia de Flavobacteriales y el retinal, indicando que los grupos fotoheterótrofos pueden ser igual de abundantes que en una región oligotrófica. En cambio, las secuencias idénticas compartidas entre el golfo de México y el este del Mediterráneo permitieron observar similitudes entre ambientes altamente oligotróficos pero que presentan distintas condiciones de circulación.

Palabras clave: Proteorodopsina, Fotoheterotrofia, Bacterias marinas, amplicón 16S-ARNr

Abstract of the thesis presented by Alberto Netzahualcoyotl Arroyo Cervantes as a partial requirement to obtain the Master of Science degree in Marine Ecology.

Bacterial diversity and phototrophic function analysis in the superficial layer of productive and oligotrophic environments

Abstract approved by:

Dra. Laura Gómez Consarnau

Thesis Co-Director

Dra. María Asunción Lago Lestón

Thesis Co-Director

Proteorhodopsin phototrophy is the most widespread type of photoheterotrophy in marine bacteria communities from the surface ocean. High throughput sequencing technologies have allowed to qualitatively identify taxonomic groups that are proteorhodopsin producers in the microbial community, while retinal concentration measurements are a quantification of the photoheterotrophic level in a given marine environment. Retinal based phototrophy has been reported to dominate in highly oligotrophic environments being an important mechanism in the ecosystem, whilst there is little about its role in more productive ecosystems. In this study a comparative analysis was made between regions with different productivity levels using 16S-rRNA sequences in order to analyze the distribution of known photoheterotrophic groups (e.g. SAR11, Flavobacteriales, SAR86) and their relation to measurements of pigments and other environmental factors. The data comes from samplings made in the regions of the Mediterranean Sea (HOTMIX cruise), the Gulf of Mexico (XIXIMI-5,6,7) and the California Current (SPOT). In the samples from the Mediterranean and the California Current, total heterotrophic bacterial counts allowed to calculate an abundance estimate for relevant taxonomic groups at order level, this to contribute to better explain their distribution. To compare the proteorhodopsin producer groups between the different regions the methods used were beta diversity analysis, multivariate analysis, correlation analysis and exact sequence comparison between regions. Beta diversity analysis showed little phylogenetic distance between the samples from the West Mediterranean and the California Current, besides the positive correlation found between retinal and Flavobacteria abundance in both regions shows that photoheterotrophic bacteria can dominate in a productive environment. Identical sequence comparison between the Gulf of Mexico and the East Mediterranean Sea showed a higher similarity between two highly oligotrophic environments which have different circulation conditions.

Keywords: Proteorhodopsin, Photoheterotrophy, Marine bacteria, 16S-rRNA

Dedicatoria

A mi madre Maribel †, que me enseñó el valor de la educación y me inculcó su amor por la naturaleza.

Agradecimientos

Al Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California, por permitirme realizar mis estudios de maestría y darme la oportunidad relacionarme con personas que sienten pasión por la ciencia.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por brindarme el apoyo económico para realizar mis estudios de maestría (No. CVU: 991230).

Al posgrado de Ecología Marina por darme la oportunidad de conocer otros puntos de vista y por el apoyo brindado durante toda la maestría.

A mis directoras de tesis, la Dra. Laura Gómez-Consarnau y la Dra. María Asunción Lago Lestón por darme la oportunidad de trabajar en este proyecto y ayudarme a cada paso.

A mis profesores, incluidos mis sinodales.

Tabla de contenido

	Página
Resumen en español	ii
Resumen en inglés	iii
Dedicatoria	iv
Agradecimientos	v
Lista de figuras	viii
Lista de tablas	x
Capítulo 1. Introducción	
1.1. Importancia de las bacterias en el océano	1
1.2. Estrategias tróficas en bacterias marinas	2
1.3. Estudio de comunidades microbianas marinas	2
1.4. Fotoheterotrofia en bacterias marinas	4
1.5. Grupos productores de proteorodopsinas	5
1.6. Regiones de estudio	6
1.6.1. Mar Mediterráneo	6
1.6.2. Corriente de California	7
1.6.3. Golfo de México	7
1.7. Antecedentes	7
1.8. Justificación	9
1.9. Hipótesis	10
1.10. Objetivos	10
1.10.1. Objetivo general	10
1.10.2. Objetivos específicos	10
Capítulo 2. Metodología	
2.1. Descripción de los datos	11
2.1.1. Mar Mediterráneo	11
2.1.2. Corriente de California	12
2.1.3. Golfo de México	13
2.2. Caracterización de la comunidad microbiana en las distintas regiones	13
2.3. Análisis de diversidad	15
2.3.1. Diversidad alfa	15
2.3.2. Diversidad beta	15
2.4. Análisis de variables ambientales	16
2.5. Análisis de estimaciones de abundancia	17
2.6. Análisis de correlaciones	17
2.7. Comparación de secuencias entre regiones	17
Capítulo 3. Resultados	
3.1. Estructura de la comunidad	18
3.2. Análisis de diversidad	22

3.3.	Análisis de variables ambientales	25
3.4.	Análisis de estimaciones de abundancias	26
3.5.	Análisis de correlaciones	29
3.6.	Comparación de secuencias entre regiones	32
Capítulo 4.	Discusión	
4.1.	Bacterias fotoheterótrofas y su relación con el retinal	34
4.2.	Fotoheterotrofia en regiones con distintos niveles de productividad	36
4.3.	Comparaciones entre regiones	38
Capítulo 5.	Conclusiones	
Literatura citada	40
Anexo	45

Lista de figuras

Figura	Página
1. Red trófica adaptada de Fenchel (2008). Las flechas verdes resaltan el paso del carbono orgánico disuelto por el "bucle microbiano".	1
2. Diagrama del gen 16S indicando las regiones conservadas en blanco y variables en gris, mostrando también la localización del ARN ribosomal 16S en el ribosoma bacteriano. Imagen obtenida de Fukuda et al. (2016)	3
3. Metodología para el análisis de amplicón 16S.	4
4. Proteorodopsina funcional indicando la dirección del bombeo de protones y las longitudes de onda (azul 490nm y verde 530nm), se muestra la estructura del retinal. Imagen modificada de Pinhassi et al. (2016).	5
5. Distribución de pigmentos en el mar Mediterráneo para A: rodopsinas y B: clorofila. Imagen tomada de Gómez-Consarnau et al. (2019). Las estaciones de muestreo se pueden ver en la figura 6	9
6. Trayecto del crucero HOTMIX 2014 a través del mar Mediterráneo abarcando el este del Mediterráneo (1-13), el oeste del Mediterráneo (14-24) y parte del Atlántico (25-29). . .	12
7. Punto de muestreo SPOT (<i>San Pedro Ocean Time series</i>) en el canal de San Pedro en la corriente de California (latitud: -91.999817, longitud: 23.000383).	13
8. A. Localización de todas las estaciones de muestreo para el golfo de México. B. Estaciones de cada campaña oceanográfica, a la izquierda XIXIMI-5 (2016), al centro XIXIMI-6 (2017) y a la derecha XIXIMI-7 (2019).	14
9. Estructura de la comunidad del Mediterráneo a distintas profundidades. A: 3 metros, B: 50 metros, C: 60-130 metros. Las profundidades de cada muestra se indican en la parte superior de cada columna en caso de no ser 3 o 50 metros. Las líneas punteadas marcan los límites entre las regiones: océano Atlántico, oeste del Mediterráneo y este del Mediterráneo de izquierda a derecha.	19
10. Estructura de la comunidad del sitio SPOT. A: Muestras a 5 metros; B: máximo de clorofila, la profundidad de las muestras del máximo de clorofila viene indicada en cada columna. Los ejes verticales de la izquierda indican la concentración de clorofila, mientras que los que están a la derecha la de retinal.	20
11. Estructura de la comunidad del golfo de México, con los niveles de fluorescencia indicados (sin unidades). A: Muestras superficiales. B: Muestras del máximo de clorofila. La profundidad de cada muestra se observa en la parte superior de cada columna. Las líneas punteadas dividen las muestras por campañas de muestreo, indicadas entre las figuras. .	22
12. Diagrama de cajas por región y por nivel de profundidad para los estimadores Chao1 a la izquierda y Shannon a la derecha.	23
13. Análisis de coordenadas principales (PCoA) para las muestras del Mediterráneo usando distancias de UniFrac con peso a partir del árbol construido con fasttree. Se indican las regiones Atlántico (cuadrado), este (rombo) y oeste (círculo) del Mediterráneo; y los niveles de retinal (pM). En cada muestra se indica la estación seguida de la profundidad. Las figuras en color verde son muestras que no cuentan con medición de retinal.	24

Figura	Página
14. Análisis de coordenadas principales (PCoA) para las muestras del Mediterráneo y SPOT usando distancias de UniFrac con peso a partir del árbol construido con Fasttree. Se indican las regiones Atlántico (cuadrado), este (rombo) y oeste (triángulo) del Mediterráneo y SPOT (círculo). En cada muestra se indica la estación seguida de la profundidad. Las figuras en color verde son muestras que no cuentan con medición de retinal.	24
15. División de las muestras del golfo de México entre regiones este y oeste.	25
16. Análisis de coordenadas principales (PCoA) para las muestras del golfo de México usando distancias de UniFrac con peso a partir del árbol construido con Fasttree. Se indica el crucero de origen con el color y la región indicada en A. Para cada muestra graficada se indica el nombre de la estación y la profundidad.	25
17. PCA para las mediciones ambientales del Mediterráneo. A: Muestra los factores indicados por <i>bioenv</i> . B: Agregando retinal.	26
18. PCA mostrando el la distancia que producen las estimaciones de abundancia de heterótrofos entre muestras con diferente productividad. Las figuras en color verde son muestras que no cuentan con medición de retinal.	28
19. PCA indicando el nivel de retinal en el color de cada punto además de las distancias de las estimaciones de abundancia para los grupos más importantes, mostrando el nivel de retinal en el color. Se omitieron muestras si mediciones de retinal.	28
20. Matriz de correlaciones obtenida para los datos del Atlántico, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.	29
21. Matriz de correlaciones obtenida para los datos del oeste del Mediterráneo, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.	30
22. Matriz de correlaciones obtenida para los datos del este del Mediterráneo, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.	31
23. Matriz de correlaciones obtenida para los datos de la corriente de California, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.	31
24. Matriz de correlaciones obtenida para los datos del golfo de México, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.	32
25. Comparación entre regiones indicando el número de ASV idénticos que comparten. . . .	33

Lista de tablas

Tabla		Página
1.	Descripción de parámetros de cada región.	16
2.	Resultados de prueba Tukey posterior a ANOVA para revisar diferencias entre regiones del Mediterráneo.	27
3.	Resultados de la prueba ANOVA y Tukey para las estimaciones de abundancia de los órdenes SAR11, Flavobacteriales y SAR86.	27
4.	Clasificación taxonómica de los ASV compartidos entre el este del Mediterráneo y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.	45
5.	Clasificación taxonómica de los ASV compartidos entre el este del Mediterráneo y el oeste del Mediterráneo a nivel de clase y orden. Se muestran los primeros 15.	45
6.	Clasificación taxonómica de los ASV compartidos entre el Atlántico y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.	46
7.	Clasificación taxonómica de los ASV compartidos entre el este del Mediterráneo y el Atlántico a nivel de clase y orden. Se muestran los primeros 15.	46
8.	Número de ASV compartidos entre el oeste del Mediterráneo y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.	46
9.	Clasificación taxonómica de los ASV compartidos entre el oeste del Mediterráneo y el Atlántico a nivel de clase y orden. Se muestran los primeros 15.	47
10.	Clasificación taxonómica de los ASV compartidos entre SPOT y el Atlántico a nivel de clase y orden. Se muestran los primeros 15.	47
11.	Número de ASV compartidos entre SPOT y el este del Mediterráneo a nivel de clase y orden. Se muestran los primeros 15.	47
12.	Clasificación taxonómica de los ASV compartidos entre SPOT y el oeste del Mediterráneo a nivel de clase y orden. Se muestran los primeros 15.	48
13.	Número de ASV compartidos entre SPOT y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.	48

Capítulo 1. Introducción

1.1. Importancia de las bacterias en el océano

Las bacterias son los organismos más abundantes en el océano, tanto entre los heterótrofos (Whitman et al., 1998) como entre los fotosintéticos o autótrofos (Scanlan et al., 2009). Además, los procesos metabólicos de las bacterias y las arqueas controlan la biogeoquímica del océano (Falkowski et al., 2008). Los elementos imprescindibles para la vida como el C, N y P junto con otros micronutrientes son mayormente asimilados o transformados a sus formas más simples (Gasol y Kirchman, 2018) por el metabolismo de estos microorganismos. Dentro del ciclo del carbono, las bacterias heterótrofas asimilan carbono orgánico disuelto que difícilmente sería aprovechado por organismos de mayor tamaño, y al ser depredadas por protistas bacterívoros, permiten que este carbono sea transferido a niveles tróficos superiores a través del llamado "bucle microbiano" (Azam et al., 1983; Pernthaler, 2005). Este carbono orgánico disuelto representa la mayor parte del carbono fijado en la fotosíntesis, dando una idea de la importancia de este proceso (Ducklow et al., 1993). En la figura 1 se puede observar una simplificación de la red trófica en la columna de agua que resalta la asimilación y transporte de carbono orgánico disuelto por parte de las bacterias heterótrofas y su producción por parte de las autótrofas.

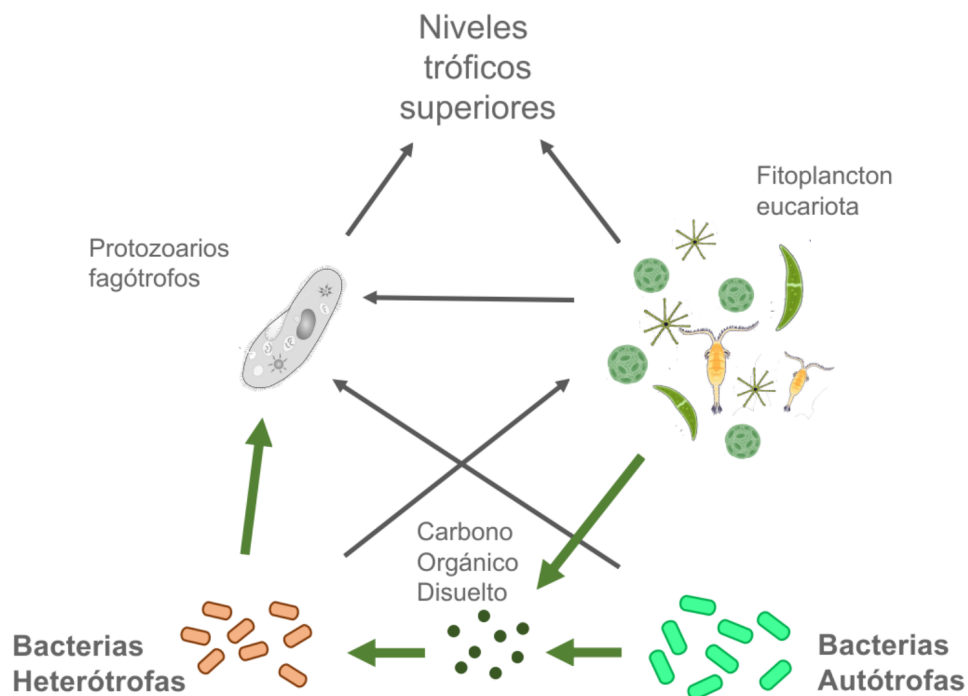


Figura 1. Red trófica adaptada de Fenchel (2008). Las flechas verdes resaltan el paso del carbono orgánico disuelto por el "bucle microbiano".

1.2. Estrategias tróficas en bacterias marinas

Las concentraciones de nutrientes disponibles en el agua de mar no se presentan de manera homogénea. Las regiones con bajas concentraciones de nutrientes se denominan oligotróficas, mientras que las regiones eutróficas tienen mayores concentraciones y generalmente muestran mayor productividad y biomasa. Dentro de las comunidades microbianas marinas, las bacterias oligotróficas presentan una baja tasa de crecimiento, genomas reducidos y un tamaño relativamente pequeño, lo que explica su adaptación a ambientes oligotróficos. En cambio, las bacterias copiotróficas presentan mayores tasas de crecimiento, genomas más grandes y por lo tanto una mayor cantidad de funciones para responder a cambios en el ambiente (Lauro et al., 2009).

En general, las regiones productivas se encuentran cerca de las costas, mientras que las regiones oceánicas alejadas de las costas son mayormente oligotróficas. Por lo tanto, la mayoría de las bacterias heterótrofas en el agua de mar están adaptadas a ambientes oligotróficos y en consecuencia son organismos que no crecen en medios ricos de cultivo (Carini et al., 2013).

1.3. Estudio de comunidades microbianas marinas

El desarrollo de las tecnologías de secuenciación ha permitido estudiar a las comunidades microbianas a partir del material genético obtenido de muestras ambientales. El genoma de un organismo consiste en el total de los genes que lo componen, en cambio el metagenoma se refiere a los genomas de los organismos que forman una comunidad. A grandes rasgos, la metagenómica consiste en el aislamiento del ADN de la muestra ambiental, seguido de un paso de rompimiento del ADN en fragmentos y la incorporación de secuencias cortas conocidas llamadas adaptadores a los extremos de los fragmentos, gracias a los adaptadores se puede realizar la amplificación del total del material genético, seguido de la secuenciación y la obtención de la información. Las secuencias obtenidas consisten en fragmentos que pueden usarse para reconstruir los genomas de las muestras usando herramientas computacionales (Gilbert y Dupont, 2011).

Del metagenoma, que consiste en los genomas de la comunidad, se pueden extraer genes de interés, ya sea para estudiar alguna función o para usarlo como marcador taxonómico.

Análisis de amplicón 16S-ARNr

El gen de la subunidad 16S del ARN ribosomal (16S-ARNr) es el marcador taxonómico más utilizado para describir la estructura de comunidades bacterianas. Como se observa en la figura 2, el ribosoma en

todas las bacterias y arqueas está formado por las subunidades 30S y 50S, donde el gen 16S es parte de la subunidad 30S del ribosoma. El ribosoma está involucrado en el proceso de síntesis de proteínas, una función altamente conservada entre bacterias y arqueas, por lo que la secuencia del gen 16S ribosomal es también altamente conservada. Sin embargo, presenta nueve regiones variables que permiten analizar la filogenia de los miembros de la comunidad a partir de sus secuencias. En la parte inferior de la figura 2 se observa un esquema con las regiones variables del gen 16S que van de V1 a V9 (Fukuda et al., 2016).

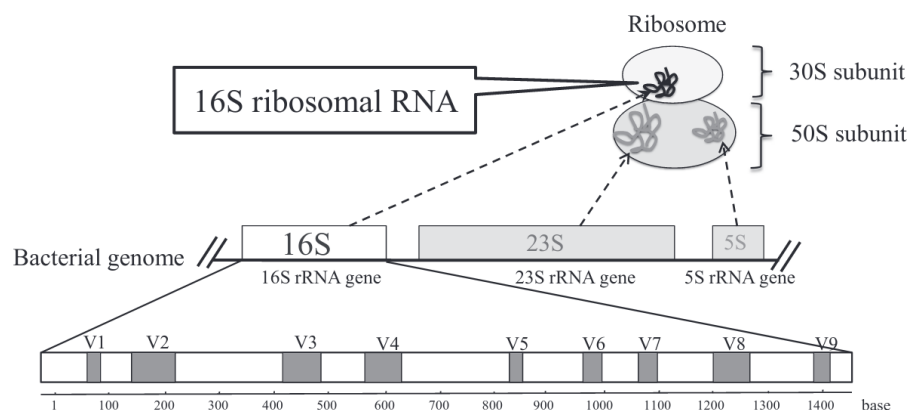


Figura 2. Diagrama del gen 16S indicando las regiones conservadas en blanco y variables en gris, mostrando también la localización del ARN ribosomal 16S en el ribosoma bacteriano. Imagen obtenida de Fukuda et al. (2016)

El gen 16S-ARNr puede extraerse del metagenoma para analizar la estructura de la comunidad de la muestra, pero también puede realizarse la secuenciación únicamente de los amplicones de alguna región de interés del gen 16S en específico. A diferencia de la metagenómica, en el análisis de amplicón no se realiza el rompimiento del material genético ni se agregan adaptadores, sino que se utilizan cebadores previamente diseñados que se alinean a regiones específicas del material genético permitiendo la amplificación de los genes de interés. El producto de esta amplificación se conoce como amplicón, y es lo que posteriormente pasa a ser secuenciado. Cuando el objetivo es analizar la estructura de la comunidad, la menor cantidad de información generada en la secuenciación para un solo gen representa una ventaja sobre la metagenómica. En cambio, el uso de algunos cebadores puede introducir sesgos en la amplificación para ciertos grupos taxonómicos (Parada et al., 2016). En la figura 3 se observa un ejemplo de la metodología para obtener la estructura de la comunidad a partir del amplicón 16S (Kembel et al., 2012).

Las secuencias generadas de esta metodología se denominan lecturas, y deben pasar por otra etapa de

procesamiento para obtener las secuencias representativas del gen 16S-ARNr que pueden ser identificadas y asignadas taxonómicamente usando bases de datos de secuencias conocidas. Tradicionalmente se utilizan OTU (Operational Taxonomic Unit) para la clasificación taxonómica, pero al basarse en técnicas de agrupamiento para la asignación no permiten la comparación directa entre diferentes muestras, además de generar más falsos positivos. Los ASV (Amplicon Sequence Variant) se obtienen a partir de la premisa de que las secuencias sin errores tienen más probabilidad de repetirse, realizando una inferencia estadística para cada secuencia. Esto permite identificar secuencias conocidas de muestras ambientales con un buen nivel de resolución con la posibilidad de comparar entre diferentes muestras (Callahan et al., 2017).

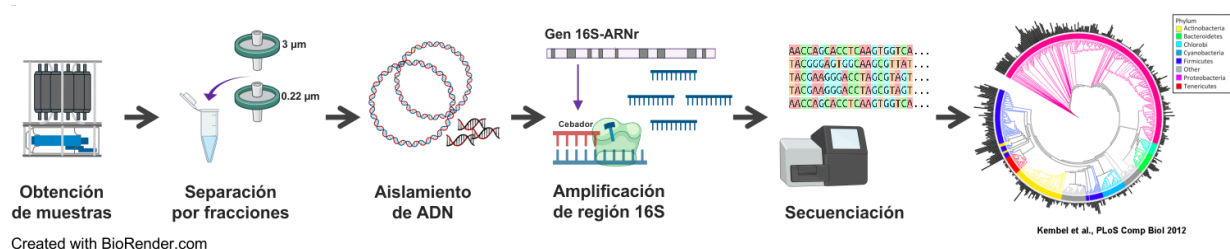


Figura 3. Metodología para el análisis de amplicón 16S.

1.4. Fotoheterotrofia en bacterias marinas

Las rodopsinas son fotosistemas que consisten en una estructura peptídica, la opsina, y un pigmento, el retinal (Figura 4) y se encuentran ampliamente distribuidas en todos los dominios de la vida con una gran variedad de funciones (Ernst et al., 2014; Nagata y Inoue, 2021). El análisis de metagenomas obtenidos de muestras de agua de mar llevó al descubrimiento de rodopsinas en bacterias heterótrofas y por lo tanto un nuevo tipo de fototrofia en el océano (Béjà et al., 2000). Posteriormente, su identificación en genomas de diferentes regiones oceánicas y sistemas acuáticos demostró que la fotoheterotrofia basada en retinal se encuentra ampliamente extendida en la zona fótica del océano (Béjà et al., 2001; Eiler, 2006; Finkel et al., 2013).

Las proteorodopsinas son un tipo de rodopsinas y son particularmente abundantes en las bacterias marinas (Finkel et al., 2013; Pinhassi et al., 2016). La parte peptídica de la proteorodopsina, la opsina, es una proteína transmembrana codificada por un único gen denominado *prd*, mientras que el retinal puede ser obtenido del rompimiento de β -caroteno, cuya reacción enzimática está codificada en el gen *blh*, además de genes relacionados al metabolismo de carotenoides *crtEIBY* (Martinez et al., 2007). Estos genes generalmente se encuentran dentro de un cassette génico que presenta diferentes variaciones entre grupos taxonómicos, incluyendo diferencias entre especies cercanas y similitudes entre grupos taxonómicos distantes, sugiriendo que esta función trófica es transferida horizontalmente (Pinhassi et al.,

2016).

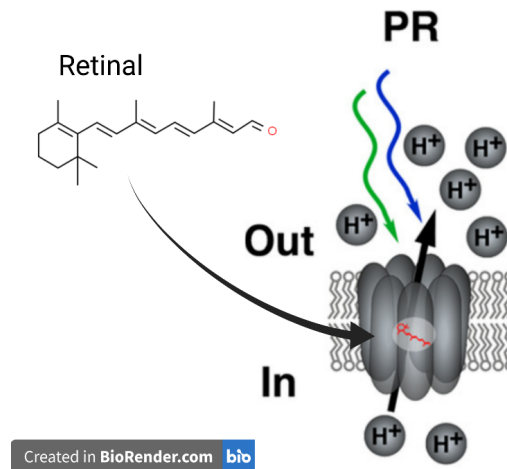


Figura 4. Proteorodopsina funcional indicando la dirección del bombeo de protones y las longitudes de onda (azul 490nm y verde 530nm), se muestra la estructura del retinal. Imagen modificada de Pinhassi et al. (2016).

El fotosistema completo de la proteorodopsina funciona como bomba de protones hacia el exterior de la célula en presencia de luz. El bombeo de protones tiene como efecto aumentar el potencial de membrana, lo cual puede acoplarse a la síntesis de ATP de manera similar a la fosforilación oxidativa (Kirchman y Hanson, 2013). La energía adicional obtenida de la luz reduce los requerimientos energéticos que de otra forma serían provistos por la oxidación de carbono orgánico (Gómez-Consarnau et al., 2007; Steindler et al., 2011).

1.5. Grupos productores de proteorodopsinas

Los datos obtenidos a partir de la secuenciación de metagenomas y metatranscriptomas han permitido identificar genes de proteorodopsinas en ambientes acuáticos de diferentes partes del mundo (Finkel et al., 2013) y evaluar su nivel de expresión en diferentes grupos taxonómicos microbianos (Sieradzki et al., 2018). Dentro de los órdenes taxonómicos que presentan una mayor expresión de proteorodopsinas y que además son organismos ubicuos en la superficie y la zona fótica del océano se encuentran los grupos SAR11 (Alfaproteobacterias), SAR86 (Gamaproteobacterias) y Flavobacteriales (Bacteroidetes). Estos organismos presentan diferentes estrategias tróficas, los organismos oligotróficos como SAR11 presentan una baja tasa de crecimiento y se adaptan a ambientes con bajas concentraciones de nutrientes, mientras que los organismos copiotróficos generalmente son oportunistas con tasas de crecimiento que se incrementan con la disponibilidad de nutrientes (Lauro et al., 2009).

SAR11

Las bacterias pertenecientes al grupo SAR11 (orden Pelagibacterales) son consideradas los organismos más abundantes en el océano (Morris et al., 2002; Giovannoni, 2017). Este grupo presenta el mayor nivel de expresión de genes de proteorodopsinas en comunidades microbianas marinas (Sieradzki et al., 2018). Estas bacterias se caracterizan por una baja tasa de crecimiento, un genoma reducido (Giovannoni et al., 2005b) y el uso de mecanismos de fotoheterotrofia ,i.e., proteorodopsinas (Giovannoni et al., 2005a). Su metabolismo está adaptado a asimilar moléculas de bajo peso molecular, lo que le proporciona ventajas selectivas en ambientes oligotróficos (o con limitaciones de nutrientes) (Giovannoni, 2017).

Flavobacteriales

El orden Flavobacteriales también es de los más abundantes en el océano. Si se comparan con organismos del grupo SAR11, los Flavobacteriales son heterótrofos que presentan mayores tasas de crecimiento y un comportamiento más copiotrófico. Su abundancia se relaciona con florecimientos de clorofila y están adaptados a utilizar materia orgánica disuelta de alto peso molecular (Kirchman, 2002; Gómez-Pereira et al., 2012; Williams et al., 2013).

SAR86

Las bacterias del grupo SAR86 son heterótrofas ubicuas en la superficie del océano (Hoarfrost et al., 2020). De acuerdo con información obtenida a través del estudio de metagenomas, presentan genomas reducidos y tienen importancia en los ciclos biogeoquímicos del carbono y el azufre (Dupont et al., 2012).

1.6. Regiones de estudio

Las regiones analizadas en el presente estudio se encuentran cerca de la zona subtropical norte (25° a 40°) pero presentan diferentes características geográficas y estados tróficos, es decir diferentes niveles de concentraciones de nutrientes y productividad.

1.6.1. Mar Mediterráneo

El mar Mediterráneo es una cuenca semi cerrada cuyas características geográficas limitan la circulación y el ingreso de nutrientes en el estrecho de Gibraltar entre el Atlántico y el oeste del Mediterráneo, mientras que el canal de Sicilia se encuentra entre el este y el oeste del Mediterráneo. Estas barreras

geográficas provocan diferencias muy marcadas en las concentraciones de pigmentos, nutrientes inorgánicos, productividad y temperatura entre estas regiones. Lo anterior causa que el Mediterráneo sea considerado una región altamente oligotrófica donde la disponibilidad de nutrientes limita la productividad (Longhurst, 2007).

1.6.2. Corriente de California

La corriente de California se encuentra en el este del océano Pacífico norte, es un sistema de surgencias de borde oriental por lo que es una región altamente productiva con surgencias a lo largo del año, presentando concentraciones de clorofila- $\alpha > 1\text{mg} \cdot \text{m}^{-3}$ durante florecimientos de fitoplancton (Longhurst, 2007).

1.6.3. Golfo de México

Esta región consiste en una cuenca oceánica en el oeste del Atlántico cuya circulación, al igual que en el Mediterráneo, se ve limitada por barreras geográficas como islas y penínsulas (al sur la de Yucatán y al norte la de Florida). Su circulación consiste en la continuación de la corriente del Caribe a través del canal de Yucatán, formando la corriente de Lazo en el este del golfo de México y terminando en la corriente del Golfo. El agua proveniente de la corriente del Caribe y que forma la corriente de Lazo es pobre en nutrientes, haciendo del golfo un ambiente poco productivo que presenta niveles de clorofila de $0.4\text{-}0.6 \text{mg} \cdot \text{m}^{-3}$ (Longhurst, 2007).

1.7. Antecedentes

Las tecnologías de secuenciación genómica de alto rendimiento y los muestreos oceanográficos a gran escala han hecho posible estudiar cómo los factores ambientales (por ejemplo, temperatura, incidencia de luz y disponibilidad de nutrientes) pueden influir en la estructura de la comunidad (DeLong, 2005; Sunagawa et al., 2015). Los principales grupos taxonómicos en las comunidades de la zona fótica del océano son cosmopolitas, pero existen variaciones estacionales y espaciales en la estructura de la comunidad que pueden deberse a sus interacciones con el medio y funciones específicas de estos grupos (Fuhrman, 2009).

Conocer la abundancia relativa de grupos taxonómicos de interés junto con mediciones de factores ambientales permite interpretar acerca del estilo de vida de dichos grupos. Dubinsky et al. (2017) reportaron el cambio en la diversidad bacteriana con relación a la concentración de clorofila desde una

zona cercana a la costa hacia el mar abierto en el este del Mediterráneo. Se encontró que la abundancia de flavobacterias (y cianobacterias) es mayor en regiones cercanas a la costa donde hay una mayor productividad (i.e., clorofila), mientras que la abundancia del grupo SAR11 fue mayor en regiones alejadas de la costa. La distribución geográfica observada para estos grupos taxonómicos se debe posiblemente a sus estrategias tróficas y adaptaciones para habitar ambientes con distinta disponibilidad de nutrientes.

Needham y Fuhrman (2016) analizaron las correlaciones presentes entre grupos taxonómicos de bacterias, fitoplancton y factores ambientales utilizando una serie de tiempo del 11 de marzo al 1 de abril del 2011, en la estación de muestreo del canal de San Pedro (SPOT: San Pedro Ocean Time series) cerca del puerto de Los Ángeles (California, USA), en la corriente de California. Encontraron que las comunidades microbianas se correlacionaban mejor entre ellas que con los factores ambientales, pero no reportaron correlación entre abundancias relativas y factores ambientales. Posteriormente, Sieradzki et al. (2018) identificaron la presencia y expresión de genes de funciones fototróficas bacterianas en la zona fótica en el mismo punto de muestreo, pero abarcando las distintas estaciones del año (de julio de 2012 a abril de 2013). Analizando los datos de metagenomas y metatranscriptomas de las muestras, fue posible identificar a los grupos taxonómicos responsables de la producción de proteorodopsina. A partir del metatranscriptoma se identificó a SAR11 como el principal productor de proteorodopsinas a lo largo del año (valor promedio, seguido por SAR86 y Flavobacteriales), además de que se encontró una correlación positiva entre la abundancia de genes de proteorodopsinas de SAR11 y los genes 16S-ARNr de SAR11, indicando que la contribución a la producción de rodopsina se incrementa junto con la abundancia relativa de este grupo. En cuanto a la abundancia relativa de genomas con rodopsinas se encontró un intercambio en la dominancia entre SAR11 y SAR86 atribuido a florecimientos de fitoplancton. El orden Flavobacteriales también se reportó como uno de los más abundantes en la estación de SPOT durante las fechas con mayor productividad (enero y abril).

Los enfoques ómicos dan mucha información acerca de la presencia de grupos microbianos, su abundancia relativa e incluso sus funciones potenciales, pero sigue siendo información cualitativa. Gómez-Consarnau et al. (2019) describieron la distribución de pigmentos en la zona fótica a través del mar Mediterráneo, reportando la distribución de retinal y clorofila desde el este del Mediterráneo hasta la entrada al Atlántico (Figura 5). La máxima concentración de retinal se encontró en la zona oligotrófica del este del Mediterráneo. En el oeste del Mediterráneo también se hallaron concentraciones altas de retinal en puntos con menores concentraciones de clorofila, sugiriendo una complementariedad entre los mecanismos de fototrofía en este sistema oceánico oligotrófico.

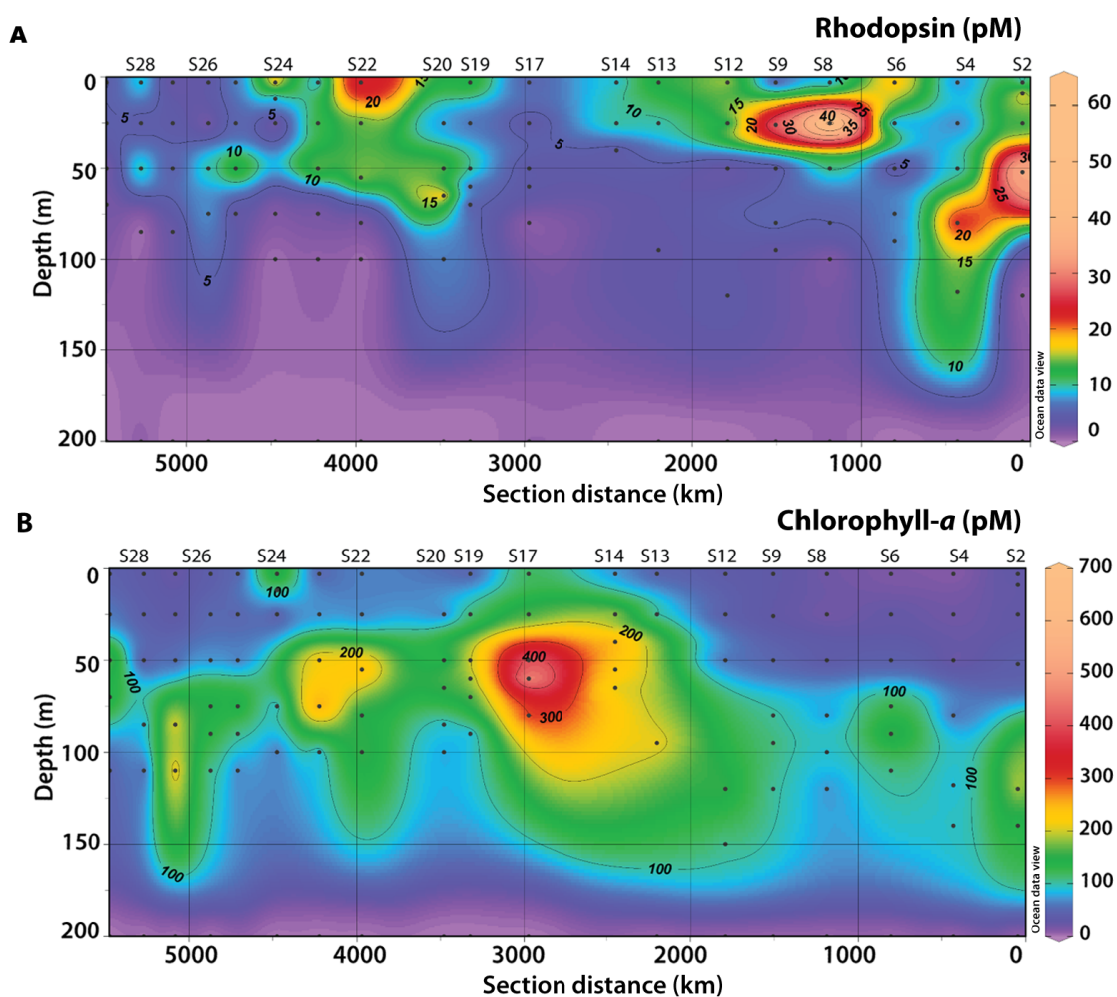


Figura 5. Distribución de pigmentos en el mar Mediterráneo para A: rodopsinas y B: clorofila. Imagen tomada de Gómez-Consarnau et al. (2019). Las estaciones de muestreo se pueden ver en la figura 6

1.8. Justificación

La fototrofia basada en retinal representa un mecanismo de adaptación importante y ampliamente extendido en el ecosistema marino, por lo que su estudio puede ayudar a comprender qué regiones favorecen el metabolismo fotoheterótrofo.

Parte de la remineralización de materia orgánica disuelta que ocurre en la zona fótica del océano es llevada a cabo por bacterias fotoheterótrofas que pueden utilizar la energía de la luz, por lo que estudiar las comunidades microbianas de ambientes con distinta productividad puede ayudar a entender la importancia de los metabolismos fotoheterótrofos en procesos biogeoquímicos de remineralización como el bucle microbiano

1.9. Hipótesis

La abundancia relativa de bacterias productoras de proteorodopsinas es mayor en regiones oligotróficas que en regiones productivas.

Los patrones de distribución de bacterias productoras de proteorodopsinas es similar entre ambientes oligotróficos.

1.10. Objetivos

1.10.1. Objetivo general

Analizar los patrones de abundancia de bacterias fotoheterótrofas con respecto a la distribución de pigmentos y otros parámetros ambientales en regiones marinas con distintos niveles de productividad.

1.10.2. Objetivos específicos

- Analizar la estructura de la comunidad de la zona fótica de tres regiones oceánicas con distintas condiciones ambientales.
- Analizar la distribución de grupos productores de proteorodopsinas y su relación con pigmentos como el retinal y la clorofila.
- Analizar las estimaciones de abundancia de grupos taxonómicos a nivel de orden usando los conteos de heterótrofos totales.

Capítulo 2. Metodología

2.1. Descripción de los datos

En el presente estudio se realizó la comparación de comunidades microbianas utilizando el gen 16S-ARNr como marcador taxonómico, mediciones de retinal, clorofila y otros parámetros ambientales como temperatura, salinidad, conteo de bacterias heterótrofas en regiones con diferentes niveles de productividad para hallar diferencias o similitudes entre los grupos bacterianos productores de proteorodopsinas. El criterio para considerar una región como oligotrófica o productiva es la concentración de clorofila y las características oceanográficas. Para SPOT, además de ser una zona de surgencias de borde oriental, se reportan concentraciones de clorofila $> 1\text{mg} \cdot \text{m}^{-3}$ ($\approx 900 \text{ pM}$), por lo que es la única región considerada productiva. La región del oeste del Mediterráneo presenta las máximas mediciones de clorofila en este trayecto ($\approx 700 \text{ pM}$), con lo que puede considerarse una región ligeramente más productiva que el resto de las regiones. Para el Atlántico, a pesar de presentar un estado trófico similar al oeste del Mediterráneo, no se observan mediciones importantes de pigmentos. Finalmente, el golfo de México y el este del Mediterráneo se consideran las regiones más oligotróficas del estudio (Longhurst, 2007).

2.1.1. Mar Mediterráneo

Las muestras provienen del crucero HOTMIX 2014 de un trayecto realizado del 30 de abril al 28 de mayo de 2014 abarcando las regiones del este del Mediterráneo, oeste del Mediterráneo y parte del océano Atlántico en el barco RV Sarmiento de Gamboa en 29 estaciones de muestreo (Figura 6). Las profundidades de muestreo abarcan desde la superficie hasta la región batipelágica, pero en el presente proyecto se analizan únicamente muestras de la zona fótica (3 metros, 50 metros y el máximo de clorofila: 70-100 m) excluyendo también las muestras a 200 metros. A diferencia del presente estudio, Sebastián et al. (2021) realizaron el análisis de la comunidad y diversidad para los mismos puntos de muestreo usando profundidades de toda la columna de agua.

Las mediciones de temperatura, salinidad y fluorescencia se obtuvieron directamente de la sonda, a la par que se recolectó agua en botellas Niskin de 12 litros. Las mediciones de los nutrientes inorgánicos nitrato, fosfato y silicato se realizaron a bordo usando análisis de flujo segmentado con detección colorimétrica. La concentración de clorofila- α se obtuvo mediante fluorimetría usando un equipo Turner Designs 10-AU. Para cada medición de retinal se filtraron 8 litros de agua de mar para almacenar la fracción microbiana 0.2-3 μm , realizar el proceso de extracción y finalmente la medición por LC/MS/MS

(Gómez-Consarnau et al., 2019). Las abundancias de *Prochlorococcus*, *Synechococcus*, piceocariotas, células con clorofila y bacterias heterótrofas se obtuvieron mediante citometría de flujo (Suffridge et al., 2018).

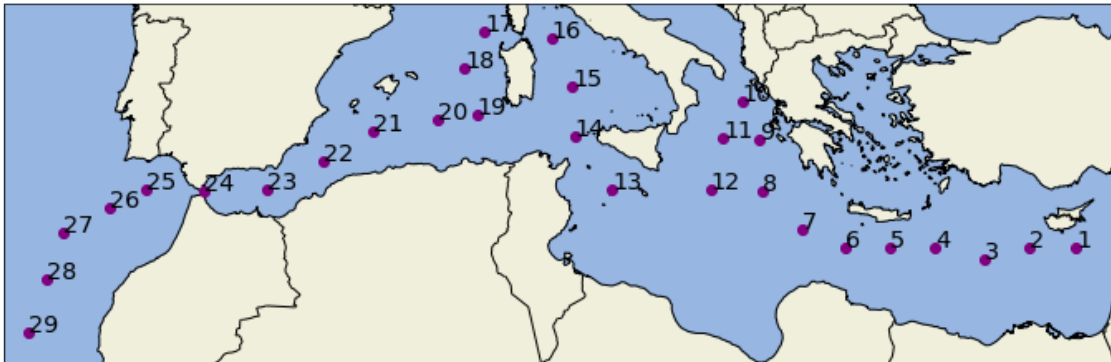


Figura 6. Trayecto del crucero HOTMIX 2014 a través del mar Mediterráneo abarcando el este del Mediterráneo (1-13), el oeste del Mediterráneo (14-24) y parte del Atlántico (25-29).

El trayecto fue dividido en tres secciones para su análisis: la región del Atlántico, entre la estación 29 y la 25; el oeste del Mediterráneo, entre las estaciones 24 y 14; y el este del Mediterráneo, abarcando las estaciones restantes (13 a 1). Las profundidades de las muestras también se dividieron en tres niveles para su análisis: superficial (3 metros), intermedio (50 metros) y profundo (60-130 metros). Las muestras de ADN se obtuvieron de la fracción 0.2-3 μm . La amplificación se realizó con los cebadores 515F/926R para la región V4-V5 del gen ARNr-16S, que presentan una buena resolución para bacterias marinas heterótrofas de vida libre (Parada et al., 2016). La secuenciación se realizó en la plataforma Illumina MiSeq (2x250 bp).

2.1.2. Corriente de California

Las muestras de la corriente de California fueron obtenidas del sitio de series de tiempo del canal de San Pedro (SPOT, latitud: -91.999817, longitud: 23.000383, Figura 7) cerca del puerto de la ciudad de Los Ángeles en las fechas de noviembre de 2016 y en 2017 los meses de enero, abril, mayo, junio y noviembre. En cada fecha se tomó una muestra superficial a 5 metros y una muestra en el máximo de clorofila, cerca de los 20 metros de profundidad, dando como resultado un total de 12 muestras (Hassanzadeh et al. *en preparación*). Se analizó la fracción 0.2-3 μm de las muestras (Villarreal Peralta, 2021). La amplificación del gen 16S-ARNr se realizó con los cebadores 515F/926R (Parada et al., 2016). La secuenciación se realizó en la plataforma Illumina (2x250 bp). Las mediciones de clorofila y retinal se obtuvieron con la misma metodología que en la región del mar Mediterráneo (Gómez-Consarnau et al.,

2019).

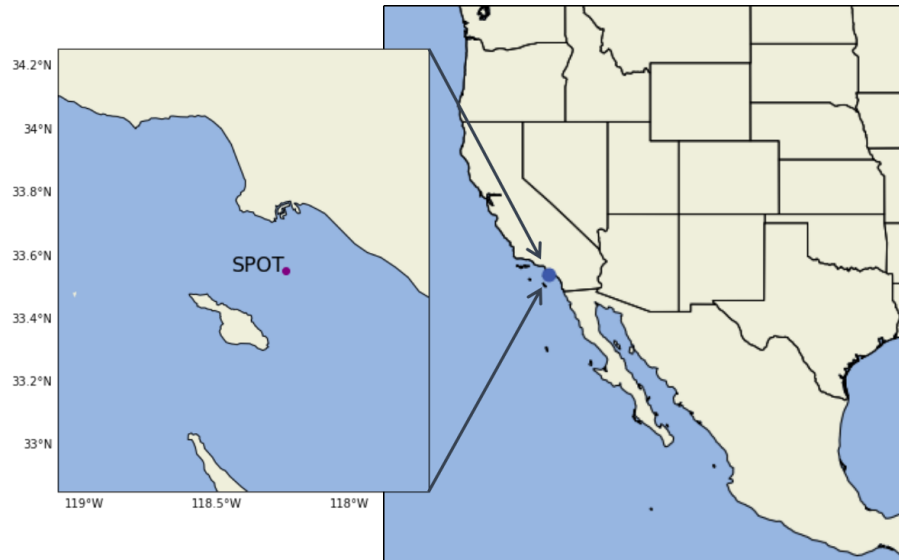


Figura 7. Punto de muestreo SPOT (*San Pedro Ocean Time series*) en el canal de San Pedro en la corriente de California (latitud: -91.999817, longitud: 23.000383).

2.1.3. Golfo de México

Los datos del golfo de México provienen de los cruceros XIXIMI-5 (junio 2016), XIXIMI-6 (agosto-septiembre 2017) y XIXIMI-7 (mayo-junio 2019). En la figura 8 se indican los puntos de muestreo para cada campaña. Consisten en la fracción 0.22-0.8 μm a dos profundidades: 5 metros en la profundidad superficial y en el máximo de clorofila, a una profundidad variable entre 70 y 110 metros. Las secuencias amplificadas para el gen 16S-ARNr consisten en la región V4 con el par de cebadores 515F/806R (Caporaso et al., 2011). La secuenciación se realizó en la plataforma Illumina MiSeq (2x150 bp). Para las muestras con profundidades entre 70 y 110 metros fue posible obtener mediciones de temperatura, salinidad, fluorescencia y oxígeno disuelto por medio del CTD.

2.2. Caracterización de la comunidad microbiana en las distintas regiones

Se utilizó la plataforma de Rstudio a través de una imagen de Docker creada durante el transcurso del proyecto con las librerías necesarias instaladas (R-4.0.5) (Boettiger y Eddelbuettel, 2017). Para la caracterización de la comunidad se utilizó el flujo de trabajo de DADA2, que presenta una mayor sensibilidad y resolución en comparación a otras herramientas (Prodan et al., 2020).

Mar Mediterráneo

Para las muestras del trayecto del Mediterráneo, ya se había realizado tanto preprocesamiento como

asignación taxonómica. Sebastián et al. (2021) describen el preprocesamiento para la obtención de ASV a través de la plataforma de DADA2 v1.8 y usando Cutadapt para eliminar los cebadores de las secuencias (Martin, 2011). Las secuencias de los ASVs habían sido asignados taxonómicamente con la base de datos de SILVA v.132 utilizando la función *assignTaxonomy* de DADA2. Por lo tanto, fueron reclasificadas usando la versión actualizada de la base de datos SILVA v.138 (Quast et al., 2013) y usando el clasificador *IdTaxa* de la paquetería DECIPHER, reportado como el que produce un menor número de falsos positivos en la clasificación (Murali et al., 2018). El filtrado de secuencias asignadas como cloroplasto, mitocondria y las secuencias sin clasificar a nivel de dominio y filo (*phylum*) se realizó con el paquete Phyloseq (McMurdie y Holmes, 2013). Usando la base de datos SILVA v.138.1 para DADA2 (McLaren y Callahan, 2021) con el clasificador nativo de DADA2 (Wang et al., 2007) se obtuvieron secuencias que no fueron asignadas por la función *IdTaxa*.

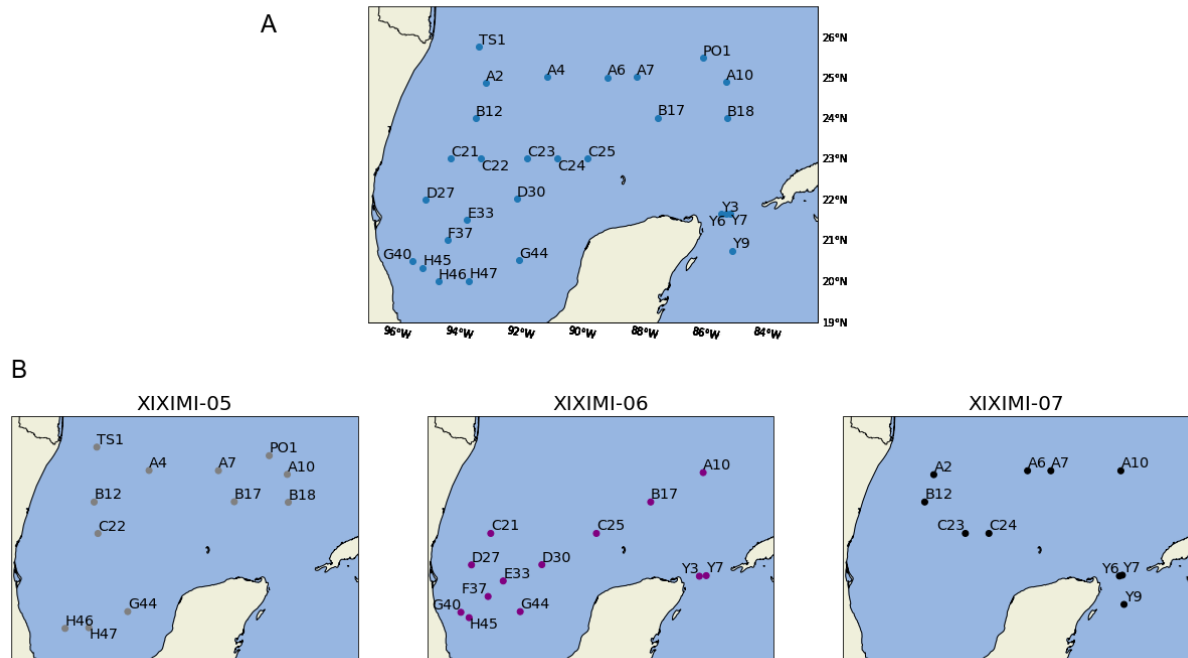


Figura 8. A. Localización de todas las estaciones de muestreo para el golfo de México. B. Estaciones de cada campaña oceanográfica, a la izquierda XIXIMI-5 (2016), al centro XIXIMI-6 (2017) y a la derecha XIXIMI-7 (2019).

Corriente de California

Para las secuencias de 16S ARNr de SPOT, la primera parte del preprocesamiento consistió en eliminar los cebadores con el programa Cutadapt (Martin, 2011) a través de R. El parámetro error esperado en el filtrado se fijó a 10 para retener un mayor número de secuencias. La eliminación de secuencias quiméricas se realizó con la función *removeBimeraDenovo* incluida en DADA2. La asignación taxonómica siguió la misma metodología descrita para las secuencias del Mediterráneo, combinando el resultado de ambos

clasificadores. Se eliminaron las secuencias clasificadas como cloroplasto, mitocondria y las secuencias sin clasificar a nivel de dominio y filo (*phylum*).

Golfo de México

El preprocesamiento fue realizado por parte del equipo de ómica del departamento de Innovación Biomédica de CICESE. Las lecturas del golfo de México fueron procesadas con el flujo de trabajo de DADA2 v1.8.0. Las secuencias se filtraron por longitud >245 y <260 con base en la distribución de las longitudes. La asignación taxonómica de nuevo siguió la misma metodología descrita en la región del Mediterráneo, combinando el resultado de los dos clasificadores. Se eliminaron los ASV asignados como cloroplasto, mitocondria, además de secuencias sin asignar a nivel de dominio y filo (*phylum*) y secuencias únicas.

2.3. Análisis de diversidad

2.3.1. Diversidad alfa

La diversidad alfa permite describir la diversidad de forma local o en una muestra mediante un índice numérico que se relaciona a la presencia de especies. El índice de Shannon expresa la cantidad de especies en una muestra, puede tomar valores de cero (una especie) y su valor aumenta con el número de especies (Kim et al., 2017). Chao1 es un estimador de riqueza que le da mayor peso a las especies poco abundantes (Chao, 1984). Para obtener los estimadores Chao1 y Shannon se utilizó la función *plot_richness* de Phyloseq, que calcula los estimadores para cada muestra.

2.3.2. Diversidad beta

Se construyeron árboles filogenéticos para los ASV de cada región utilizando Fasttree (Price et al., 2010) a través de qiime2:2019.1. El árbol filogenético generado fue importado directamente en el ambiente de Rstudio y agregado al objeto de Phyloseq. Con esta información fue posible calcular distancias UniFrac y realizar análisis de coordenadas principales (PCoA) para comparar las muestras de cada región y entre algunas regiones. Las muestras de SPOT y el Mediterráneo se compararon entre sí, tomando en cuenta que las primeras corresponden a una serie temporal de todo un año y las otras a una distribución espacial posterior a un florecimiento. Las muestras del mar Mediterráneo provienen de una metodología similar a la de SPOT, se utilizó la misma fracción (0.22-3 μm) y el mismo par de cebadores para la amplificación 515F/926R (Parada et al., 2016). En cambio, las muestras del golfo de México consisten

en la fracción 0.22-0.8 μm y fueron amplificadas con los cebadores 515F/806R (Caporaso et al., 2011). Se decidió utilizar distancias UniFrac con peso, que toma en cuenta también las abundancias relativas, y para este análisis no se utilizó ningún método de normalización, ya que de esta manera se explicaba la mayor variabilidad de los datos (McKnight et al., 2019; McMurdie y Holmes, 2013).

2.4. Análisis de variables ambientales

El análisis de variables ambientales se realizó solo para las muestras del Mediterráneo, ya que es la región que tenía más mediciones (Tabla 1). Para identificar qué variables ambientales expresan mejor la variabilidad en las muestras se utilizó el comando *bioenv* (Clarke y Ainsworth, 1993), que identifica las variables ambientales que mejor se correlacionan con una matriz de disimilitud de distancias de Bray-Curtis obtenidas de las tablas de conteo.

Tabla 1. Descripción de parámetros de cada región.

Parámetros	Sitios de muestreo		
	Mediterráneo	Corriente de California	Golfo de México
Pigmentos	Clorofila, Retinal	Clorofila, Retinal	Fluorescencia
Mediciones ambientales	Temperatura, Salinidad, Nitratos, Fosfatos, Silicatos, Producción bacteriana, Carbono orgánico total	Temperatura	Temperatura, Salinidad, Oxígeno disuelto
Citometría	Bacterias heterótrofas totales , <i>Synechococcus</i> , <i>Prochlorococcus</i>	Bacterias heterótrofas totales	-
Gen marcador taxonómico	16S-ARNr V4-V5 (515F-926R)	16S-ARNr V4-V5 (515F-926R)	16S-ARNr V4 (515F-806R)

Se utilizaron muestras que incluyeran mediciones de retinal y clorofila, omitiendo las que no tuvieran mediciones de estos pigmentos con lo que quedaron 43 muestras del Mediterráneo para el análisis con *bioenv*. La profundidad no se agregó ya que en los análisis posteriores se utilizó como variable categórica. No se tuvo en cuenta la salinidad debido a su relación con la temperatura, considerando que podría dar información redundante. Las mediciones de temperatura tenían una distribución cuya probabilidad de normalidad de acuerdo con la prueba de Shapiro-Wilks era $p = 0.024$, por lo que se asume normalidad para esta variable. Se utilizó el paquete *bestNormalize* para obtener una transformación de normalización para el resto de las variables ambientales. El paquete sugiere un tipo de transformación que permita normalizar los datos. La producción bacteriana, la clorofila, nitrato y fosfato se normalizaron con una transformación logarítmica. La normalización del retinal se realizó usando una transformación de raíz cuadrada. Los datos de carbono orgánico total y el conteo total de heterótrofos ya tenían una distribución normal. El fosfato no fue considerado debido a la poca variabilidad que presenta en la zona fótica de todo el trayecto. Se realizó un PCA con los resultados de *bioenv* para explorar la variabilidad de las muestras.

2.5. Análisis de estimaciones de abundancia

Debido a que la abundancia de bacterias heterótrofas totales ya tenía una distribución normal, se realizaron análisis de varianza (ANOVA) para buscar diferencias entre regiones y entre niveles de profundidad respectivamente. A partir de la estructura de la comunidad obtenida de la secuenciación del gen 16S-ARNr se obtuvo el porcentaje de ciertos grupos taxonómicos de interés a nivel de orden (SAR11, Flavobacteriales, SAR86). Al eliminar las secuencias asignadas a grupos autótrofos, se obtuvo la estructura de la comunidad con los grupos únicamente heterótrofos a nivel de orden. Al multiplicar este porcentaje por la abundancia bacterias heterótrofas totales en la muestra proporcionó una estimación de la abundancia de estos grupos taxonómicos. Después de normalizar los datos de estimación de abundancia, se realizaron análisis de varianza con respecto a la región para estos grupos taxonómicos. Para normalizar las estimaciones de abundancias se utilizó el paquete BestNormalize, que sugirió transformaciones logarítmicas para las estimaciones de Flavobacteriales, SAR86, Rhodobacterales y Actinomarinales ($\log_b a$); mientras que para SAR11 y Rhodospirillales sugirió transformaciones con exponentes ($x^{a/b}$). Se utilizó la prueba de Tukey para analizar entre qué grupos se encontraban las diferencias.

2.6. Análisis de correlaciones

Se realizó un análisis de correlaciones con el paquete RcmdrMisc usando el coeficiente de correlación de Spearman. Usando las estimaciones de abundancia de grupos taxonómicos a nivel de orden, se realizaron correlaciones con las mediciones ambientales en las regiones de estudio del Mediterráneo y en la corriente de California. Para el golfo de México las correlaciones se realizaron usando las abundancias relativas.

2.7. Comparación de secuencias entre regiones

Debido al sesgo presentado por ciertos grupos taxonómicos en los cebadores utilizados en las muestras del golfo de México (Parada et al., 2016), se realizó una comparación directa entre las secuencias del golfo de México con el resto de las regiones, buscando asignaciones taxonómicas y secuencias iguales entre las regiones, y posteriormente revisando las secuencias de organismos fotoheterótrofos presentes en las distintas regiones. Debido a que la subestimación no afecta la identificación de organismos en las muestras, es posible encontrar coincidencias entre los ASV obtenidos de las distintas regiones.

Capítulo 3. Resultados

3.1. Estructura de la comunidad

Mar Mediterráneo

La estructura de la comunidad para las muestras del Mediterráneo se visualiza en la figura 9 como abundancias relativas a nivel taxonómico de orden. A tres metros de profundidad (Figura 9A), las estaciones del Atlántico presentan una dominancia del orden Synechococcales (34-52 %), seguido por el grupo SAR11 (22-28 %). En cambio, en las estaciones del este y el oeste del Mediterráneo el grupo SAR11 domina la mayoría de las muestras de superficie (26-46 %). Los órdenes Flavobacteriales (3-16 %), SAR86 (4-11 %), Puniceispirillales (7-15 %), Rhodobacterales (3-10 %) y Synechococcales (5-20 %) también presentan una abundancia relativa importante en las muestras del Mediterráneo. También se puede observar una mayor abundancia relativa del orden de arqueas Marine Group II en la estación 24 (9 %), correspondiente al estrecho de Gibraltar, que a la vez es la única muestra a 3 metros donde domina el orden Flavobacteriales con 35 % de la comunidad.

A los 50 metros de profundidad (Figura 9B) el orden Punicespirituales presenta una menor abundancia relativa que a 3 metros (4-9 % a 50m vs. 7-15 % a 3m), y lo mismo ocurre con los Rhodobacterales (2-5 % a 50m vs. 3-10 % a 3m) y Opitutales (1-3 % a 50m vs. 2-5 % a 3m). Las muestras del Atlántico a 50 metros presentan una estructura similar a la superficie dominada por fotoautótrofos. En el este del Mediterráneo los Synechococcales se encuentran en mayor abundancia relativa (16-40 %) que a 3 metros (10-20 %). El grupo SAR86 presenta una abundancia relativa mayor en el este (5-9 %) que en el oeste del Mediterráneo (1-7 %), mientras que a 3 metros no se observaron diferencias importantes. En la estación 14 indicada a 40 metros de profundidad, correspondiente a la parte norte del estrecho de Sicilia, se presenta una dominancia del orden de arqueas Nitrosopumilales (32 %).

La figura9C presenta la estructura de la comunidad microbiana en las muestras a mayor profundidad, entre 60 y 130 metros. En general, se puede observar que el orden Actinomarinales presenta mayor abundancia relativa (1-7 %) con respecto a las muestras más superficiales, donde no es mayor al 5 %. Se observa que el orden de arqueas Nitrosopumilales presenta mayor abundancia relativa a mayor profundidad (hasta 36 %). El orden Opitutales presenta una menor abundancia relativa en comparación a las muestras a menor profundidad. Las muestras del Atlántico presentan cambios en las proporciones de los grupos dominantes (grupo SAR11, Synechococcales y Nitrosopumilales) al acercarse al estrecho de Gibraltar. Las muestras del oeste del Mediterráneo a este nivel de profundidad presentan una mayor

abundancia del orden Flavobacteriales (10-22%) que corresponde con el aumento de retinal en las estaciones 19 a 23 (Figura 5), faltando mediciones de este pigmento que correspondan a las estaciones 15 y 16 (Figura 9). En el este del Mediterráneo Synechococcales es el orden dominante abarcando el 30-65% de la comunidad, correspondiendo con las mediciones de máximas de clorofila cerca de esta profundidad. También se observa una menor abundancia relativa de los grupos descritos en las muestras más superficiales como Puniceispirillales, Opituales y Rhodobacterales. Para los Flavobacteriales no se observan cambios importantes en su abundancia relativa, que es ligeramente menor que las muestras superficiales, pero sigue siendo de los grupos más abundantes en la comunidad.

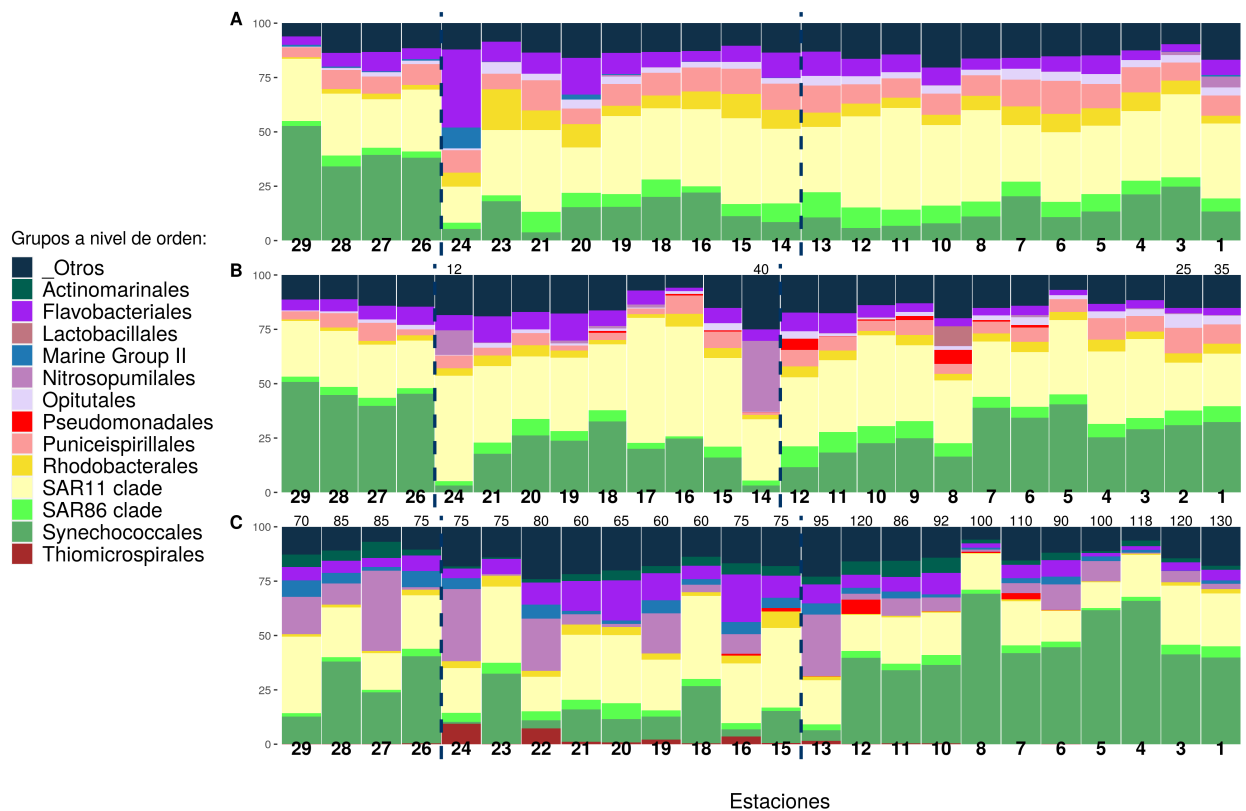


Figura 9. Estructura de la comunidad del Mediterráneo a distintas profundidades. A: 3 metros, B: 50 metros, C: 60-130 metros. Las profundidades de cada muestra se indican en la parte superior de cada columna en caso de no ser 3 o 50 metros. Las líneas punteadas marcan los límites entre las regiones: océano Atlántico, oeste del Mediterráneo y este del Mediterráneo de izquierda a derecha.

Corriente de California

En la figura 10 se observa la estructura de la comunidad microbiana en la corriente de California, una región altamente productiva a lo largo del año (Longhurst, 2007). Se puede observar un dominio de grupos fotoheterótrofos (e.g., SAR11, Flavobacteriales, SAR86, Rhodobacterales) durante todo el año, tanto a los 5 metros como en el máximo de clorofila, que se encuentra entre los 20 y 30 metros.

La abundancia relativa del orden Synechococcales es menor en comparación con el resto de los grupos microbianos dominantes, especialmente a la profundidad el máximo de clorofila y en fechas con mayores concentraciones de este pigmento (11 % a 5m vs 8 % en el máximo de clorofila). Los grupos dominantes en la superficie fueron SAR11 y Flavobacteriales a lo largo de todo el año con una abundancia media de 36 % y 20 %, respectivamente. Los órdenes Rhodobacterales (6 %), SAR86 (5 %), Cellvibrionales (3 %) también se encontraron de manera constante. La máxima abundancia del grupo Flavobacteriales a los 5 metros se dio en el mes de abril (30 %), coincidiendo con el máximo de retinal a esta profundidad a lo largo del año.

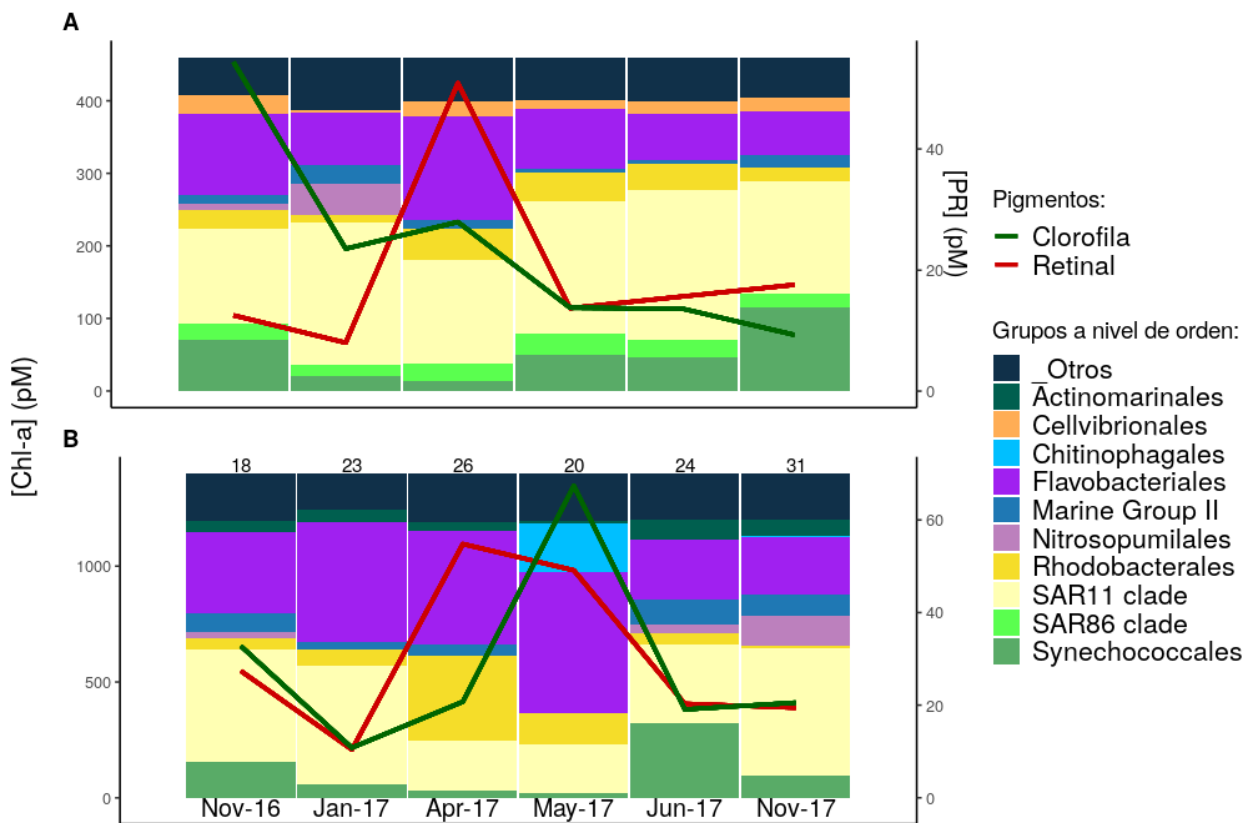


Figura 10. Estructura de la comunidad del sitio SPOT. A: Muestras a 5 metros; B: máximo de clorofila, la profundidad de las muestras del máximo de clorofila viene indicada en cada columna. Los ejes verticales de la izquierda indican la concentración de clorofila, mientras que los que están a la derecha la de retinal.

En la profundidad del máximo de clorofila se puede observar una menor abundancia relativa del orden SAR11 (27 %) y un aumento en la abundancia relativa de Flavobacteriales (30 %), principalmente en fechas de mayor concentración de pigmentos. La muestra del máximo de clorofila de abril (a 26 m de profundidad) presenta una dominancia de Flavobacteriales (35 %) y Rhodobacterales (26 %). La dominancia de Flavobacteriales (37 %) se mantiene en mayo, mes donde se obtuvo el máximo de retinal

(60 pM). Otra diferencia entre las profundidades es la mayor abundancia del orden Actinomarinales a mayor profundidad, al igual que en las muestras del Mediterráneo. También se observa una menor abundancia relativa de SAR11 en los meses de abril y mayo para este nivel de profundidad (15 %) que en la superficie (30 y 40 % respectivamente). Se puede observar la diferencia en la profundidad del máximo de clorofila, ya que en las muestras del Mediterráneo los máximos de clorofila se encuentran a partir de los 50 metros (Figura 5).

Se observa que la diversidad de grupos taxonómicos a nivel de orden en esta región es menor que en el Mediterráneo (Figura 9). El orden Cellvibrionales (3.4 %) se encuentra en mayor abundancia relativa, en cambio Puniceispirillales (1 %) y Opitutales (<1 %) presentan menores abundancias relativas, mientras que Lactobacillales, y Pseudomonadales no se encuentran en SPOT. La presencia de estos grupos taxonómicos relacionados a la superficie que aparentemente no son cosmopolita podría relacionarse al nivel trófico de estas aguas productivas.

Golfo de México

La estructura de la comunidad para las muestras del golfo de México se muestra en la figura 11 dividida en nivel de profundidad y año. La abundancia relativa de SAR11 en las muestras de superficie es 4 % en promedio. Para las muestras de superficie (Figura 11A) se observa una abundancia relativa importante para los órdenes SAR86 (13 %), Alteromonadales (5 %), Oceanospirillales (2 %), Bacteriovorales (2 %), todas las gamaproteobacterias, que tienen una abundancia relativa menor en las muestras del Mediterráneo y en SPOT. En general, se observa una dominancia del orden Synechococcales en todas las muestras de superficie (30 %). Los órdenes Flavobacteriales (4-11 %), Actinomarinales (2-7 %) y Puniceispirillales (3-7 %) se presentan en todas las muestras con una abundancia relativa similar. El orden Rhodospirillales también presentó una abundancia relativa importante en las muestras de superficie (3-17 %).

Las muestras del máximo de clorofila presentaron una estructura de la comunidad similar a las de superficie. El orden SAR11 se encontró en la misma abundancia relativa promedio (4 %). El orden Synechococcales presentó una abundancia relativa ligeramente menor (27 %), mismo patrón que se observa para la mayoría de los grupos taxonómicos descritos para superficie. El orden Rhodobacterales presenta una abundancia relativa baja, siendo visible únicamente en las muestras de mayor profundidad (1 %). Los órdenes de arqueas Nitrosopumilales (1-15 %) y Marine Group II (5-20 %) presentaron una mayor abundancia relativa en esta profundidad.

En la figura 11B se observan los niveles de fluorescencia para cada muestra, como aproximación a los

niveles de clorofila de las muestras, habiendo correlaciones entre los mayores niveles de fluorescencia, que es un indicador de la clorofila, y una mayor abundancia relativa de ciertos órdenes taxonómicos, como los Flavobacteriales. Para las muestras de 2019 se observa en general un menor nivel de fluorescencia en las muestras que se encuentran más cercanas al interior del golfo de México (C23, C24), mientras que en las muestras de 2017 y 2016 se observan picos de fluorescencia en muestras del interior del golfo (G44, H46). No se observan cambios importantes en la estructura de la comunidad entre las expediciones XIXIMI-5, XIXIMI-6 y XIXIMI-7.

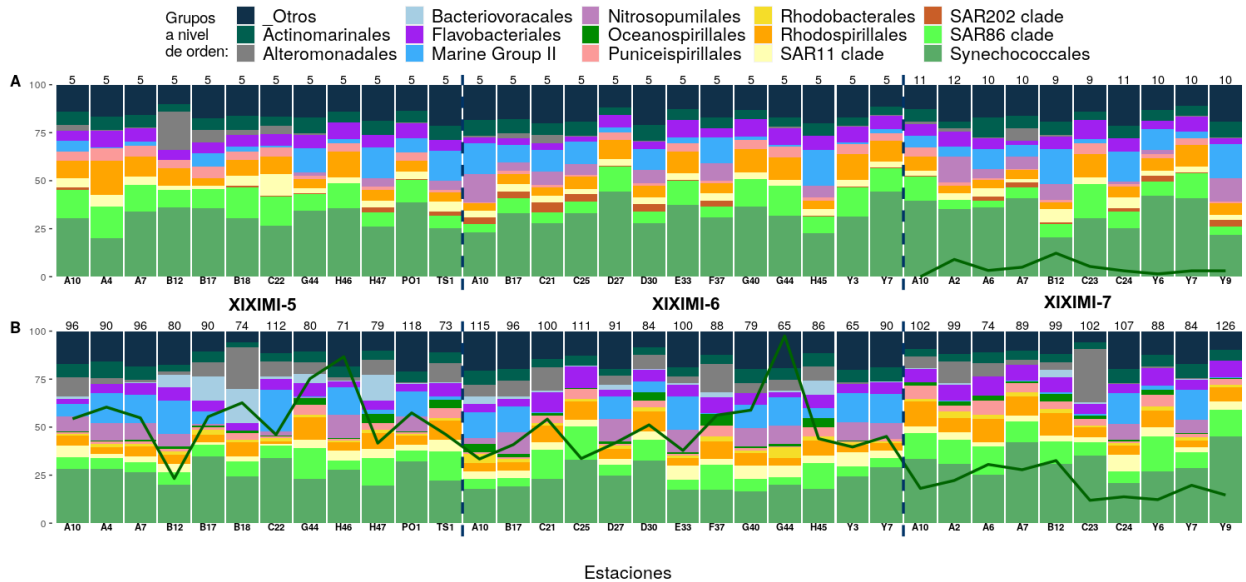


Figura 11. Estructura de la comunidad del golfo de México, con los niveles de fluorescencia indicados (sin unidades). A: Muestras superficiales. B: Muestras del máximo de clorofila. La profundidad de cada muestra se observa en la parte superior de cada columna. Las líneas punteadas dividen las muestras por campañas de muestreo, indicadas entre las figuras.

3.2. Análisis de diversidad

Diversidad alfa

La figura 12 muestra diagramas de caja para los estimadores de Chao1 y Shannon para cada región y profundidad. En general, se observa poca dispersión para los valores en muestras cercanas a la superficie. Las muestras de SPOT presentan valores menores a 100 para Chao1, indicando una baja cantidad de especies poco abundantes. Las muestras del golfo de México entre 60 y 130 metros de profundidad presentan valores mayores para ambos estimadores que las regiones del Mediterráneo y SPOT.

Diversidad beta

Se utilizó la distancia de UniFrac con peso ya que toma en cuenta la distancia filogenética y abundancias relativas en las muestras. Para el Mediterráneo (Figura 13) se observa un agrupamiento de las

muestras del este del Mediterráneo en la parte inferior de la figura. Las muestras del oeste se acomodan a la derecha de la figura, mientras que las del Atlántico se distribuyen desde el centro hacia la izquierda de la figura. Se puede observar un agrupamiento de las muestras de mayor profundidad en el este del Mediterráneo con las muestras del Atlántico, coincidiendo con la observación de presentar comunidades similares a pesar de ser regiones diferentes. Se puede observar como las muestras del estrecho de Gibraltar a 12 metros se aleja del ordenamiento (muestra 24-12 en la figura 13).

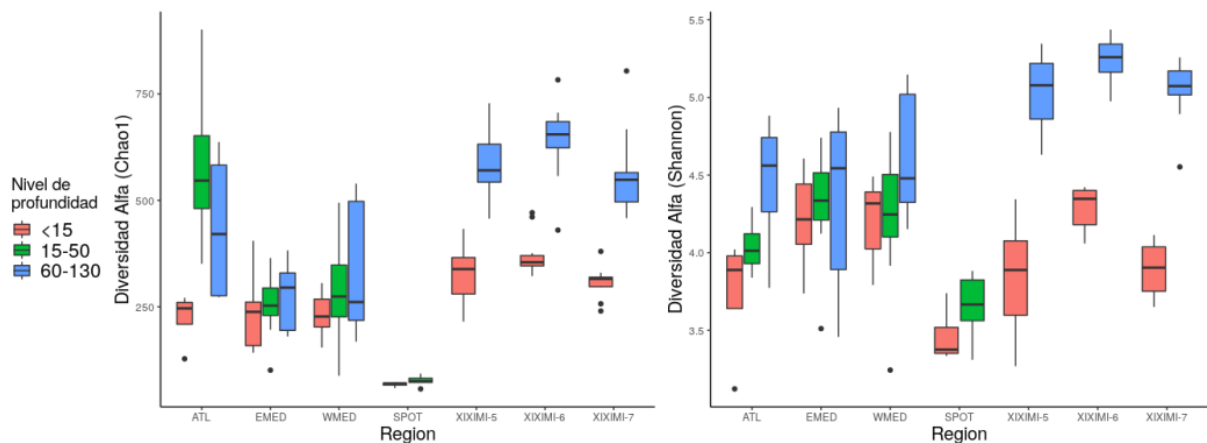


Figura 12. Diagrama de cajas por región y por nivel de profundidad para los estimadores Chao1 a la izquierda y Shannon a la derecha.

Debido a la similitud en su metodología de procesamiento y la baja cantidad de muestras, el análisis de diversidad beta para la corriente de California se realizó en conjunto con las muestras del Mediterráneo (Figura 14). Se puede observar un patrón similar de agrupamiento al descrito en la figura 13, con la diferencia de que las muestras del Atlántico se encuentran agrupadas al centro de la figura, mientras que las muestras de la corriente de California se agrupan principalmente con las muestras del oeste del Mediterráneo. Únicamente la muestra superficial de enero de SPOT (Jan-17 en la figura 14) se agrupó cerca de las muestras del este del Mediterráneo, mientras que las muestras del mes de noviembre, con niveles altos de retinal, se agrupan con las muestras del Mediterráneo de mayor profundidad.

Para el golfo de México, las muestras se dividieron en dos regiones: este y oeste (Figura 15). El análisis de diversidad beta para el golfo de México en la figura 16 se observa un agrupamiento de las muestras de superficie en la parte derecha del gráfico. Las muestras se agrupan independientemente del año de la campaña de muestreo. Las muestras del máximo de clorofila que se encuentran del lado izquierdo del gráfico presentan un patrón que se divide en dos grupos: el grupo de la parte inferior que incluye muestras más alejadas del golfo; mientras que las muestras de la parte superior de la gráfica provienen principalmente del interior del golfo (Figura 16).

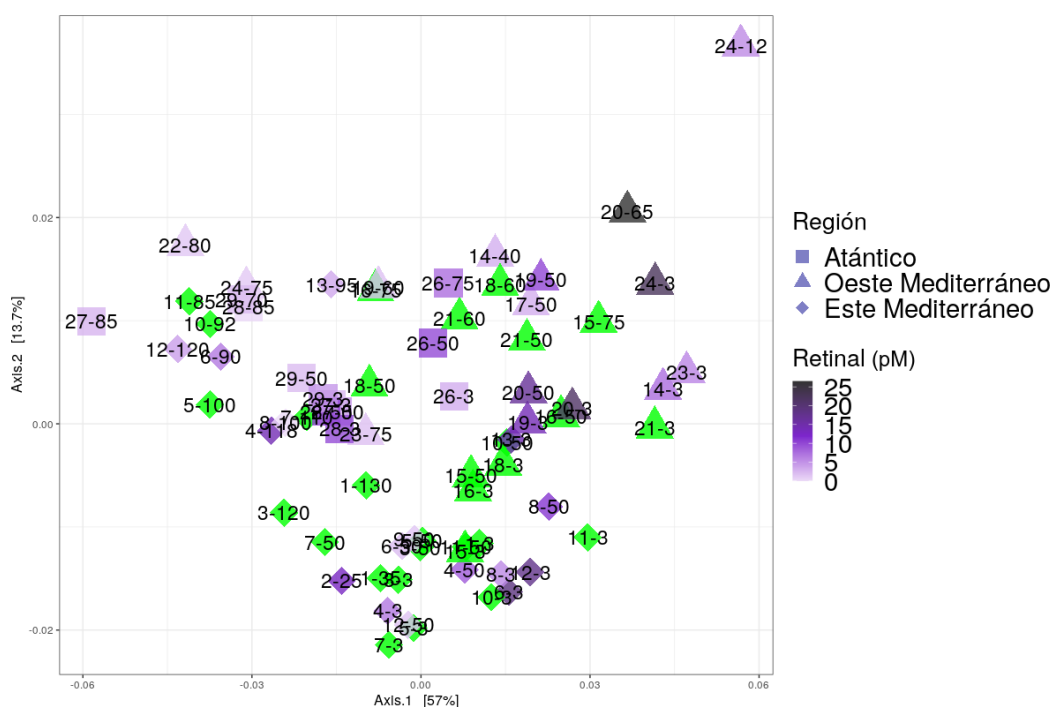


Figura 13. Análisis de coordenadas principales (PCoA) para las muestras del Mediterráneo usando distancias de UniFrac con peso a partir del árbol construido con fasttree. Se indican las regiones Atlántico (cuadrado), este (rombo) y oeste (círculo) del Mediterráneo; y los niveles de retinal (pM). En cada muestra se indica la estación seguida de la profundidad. Las figuras en color verde son muestras que no cuentan con medición de retinal.

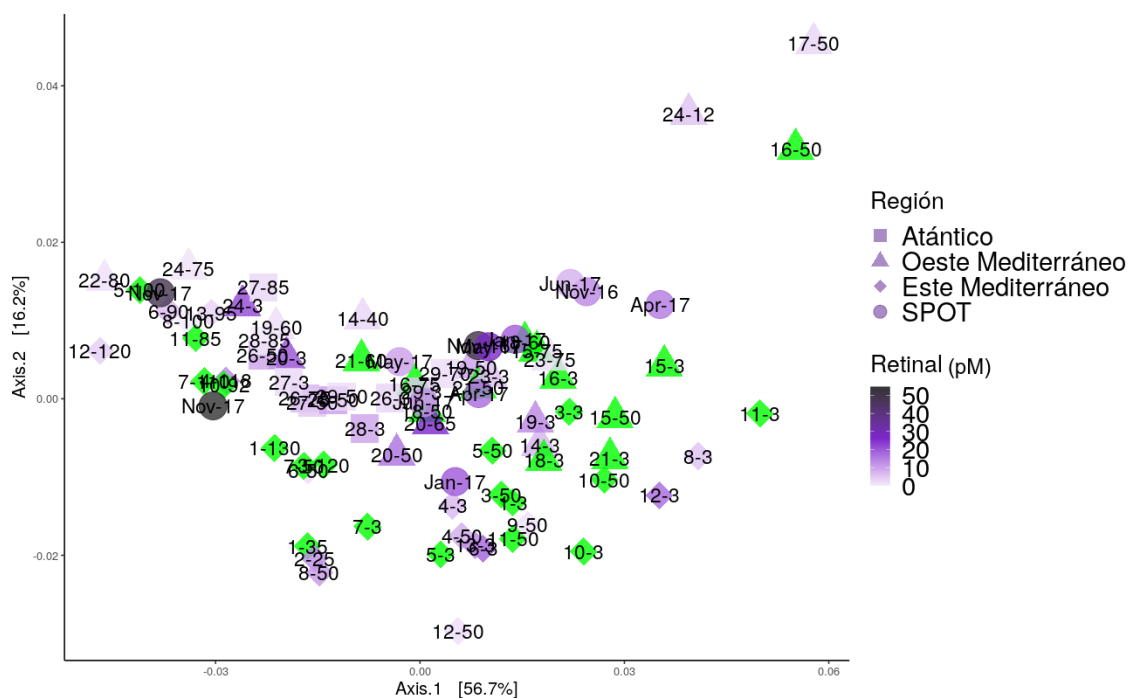


Figura 14. Análisis de coordenadas principales (PCoA) para las muestras del Mediterráneo y SPOT usando distancias de UniFrac con peso a partir del árbol construido con Fasttree. Se indican las regiones Atlántico (cuadrado), este (rombo) y oeste (triángulo) del Mediterráneo y SPOT (círculo). En cada muestra se indica la estación seguida de la profundidad. Las figuras en color verde son muestras que no cuentan con medición de retinal.

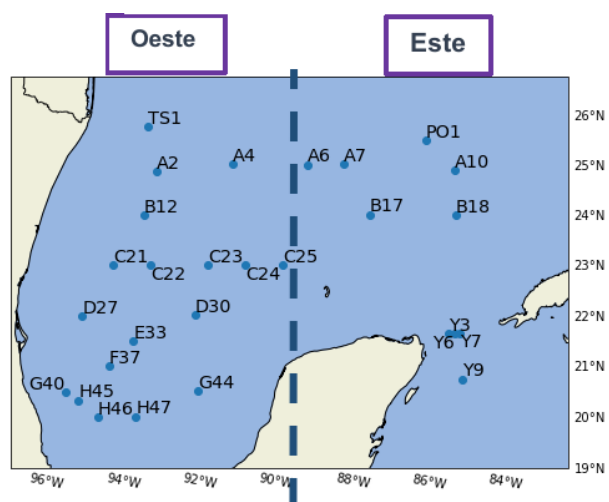


Figura 15. División de las muestras del golfo de México entre regiones este y oeste.

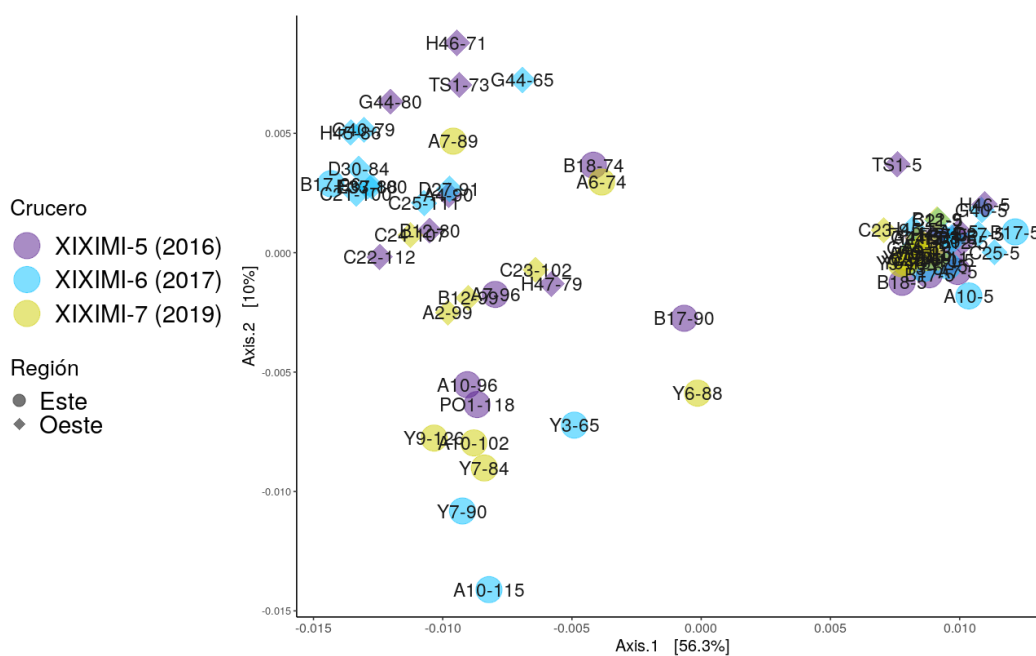


Figura 16. Análisis de coordenadas principales (PCoA) para las muestras del golfo de México usando distancias de UniFrac con peso a partir del árbol construido con Fasttree. Se indica el crucero de origen con el color y la región indicada en A. Para cada muestra graficada se indica el nombre de la estación y la profundidad.

3.3. Análisis de variables ambientales

Debido a la cantidad de parámetros en cada región, este análisis se realizó solamente para las muestras del Mediterráneo (Tabla 1). Tomando en cuenta que la cantidad de parámetros en el golfo de México y SPOT, la variabilidad de la comunidad se estudió en el análisis de correlaciones. El comando *bioenv* determinó que las variables ambientales que mejor describen las distancias entre las muestras son

las concentraciones de fosfato, nitrato, clorofila y abundancia de bacterias heterótrofas totales (indicado como bacteria en la figura 17). La figura 17A muestra un PCA realizado con los factores sugeridos por el comando *bioenv*. Se puede observar una separación entre las muestras del este del Mediterráneo con cierto traslape con las otras regiones. La división entre el oeste del Mediterráneo y el Atlántico es menos aparente.

Al agregar la concentraciones de retinal en la figura 17B se puede observar que esta variable aporta poca información acerca de la separación de las muestras, indicando además una correlación negativa con el fosfato. En ambas figuras se observa una mejor separación entre el este del Mediterráneo y las otras dos regiones marcada principalmente por la abundancia de bacterias heterótrofas totales y en menor medida la concentración de clorofila.

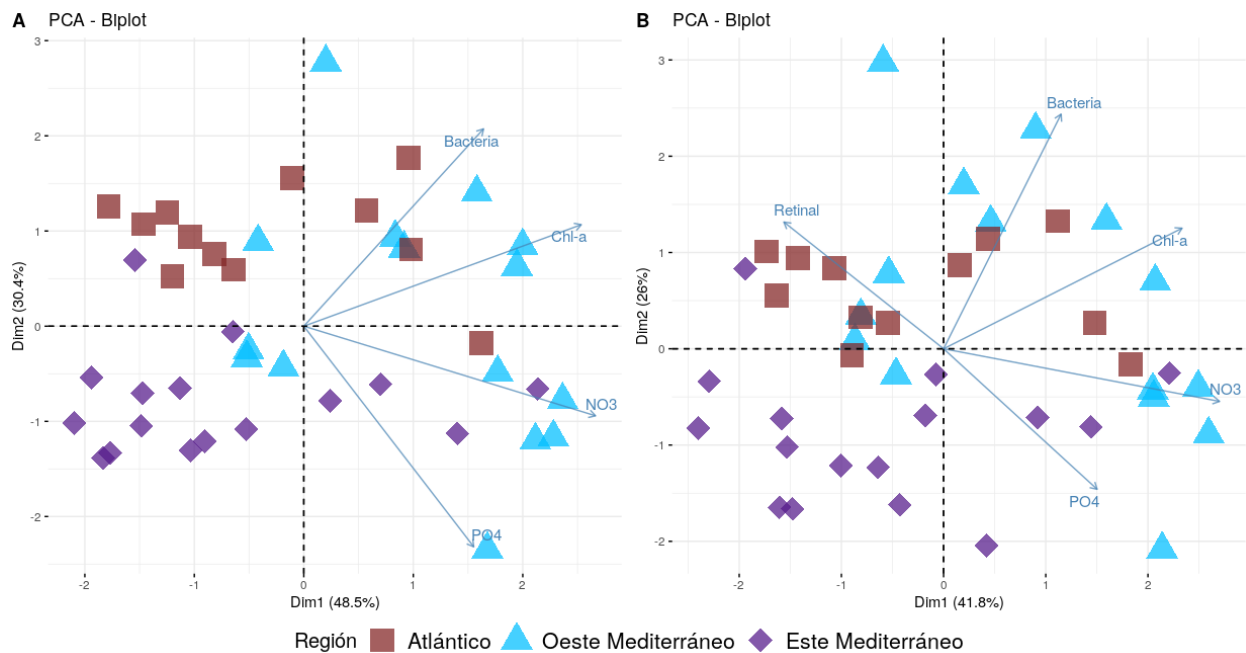


Figura 17. PCA para las mediciones ambientales del Mediterráneo. **A:** Muestra los factores indicados por *bioenv*. **B:** Agregando retinal.

3.4. Análisis de estimaciones de abundancias

Usando ANOVA de una vía se encontró que no existen diferencias significativas en las abundancias de bacterias heterótrofas totales a las distintas profundidades ($p = 0.45$) en las muestras del Mediterráneo. En cambio, sí existen diferencias significativas entre las regiones ($p < 0.001$), específicamente entre las muestras del este del Mediterráneo con las regiones del oeste del Mediterráneo y el océano Atlántico (Tukey, $p < 0.001$; Tabla 2).

Para las estimaciones de abundancia de grupos taxonómicos, también se encontraron diferencias

Tabla 2. Resultados de prueba Tukey posterior a ANOVA para revisar diferencias entre regiones del Mediterráneo.

Prueba Tukey			
Regiones	Este-Atlántico	Oeste-Atlántico	Este-Oeste
Resultado	$H_1 : \mu_E < \mu_A$ p<0.001	$H_0 : \mu_W = \mu_A$ p=0.53	$H_1 : \mu_E < \mu_W$ p<0.001

significativas para los grupos SAR11 y Flavobacteriales entre el este y las otras dos regiones, coincidiendo con los resultados de la abundancia de bacterias heterótrofas totales (Tabla 3). Para los órdenes SAR86, Rhodospirillales y Opituales no se encontraron diferencias significativas entre regiones ($p = 0.29$, $p = 0.20$, $p = 0.60$, Tabla 3). Las estimaciones de abundancia de los Rhodobacterales mostraron diferencias significativas entre regiones ($p<0.01$), presentando diferencias entre el oeste del Mediterráneo y las otras dos regiones (Tukey, $p<0.05$).

Tabla 3. Resultados de la prueba ANOVA y Tukey para las estimaciones de abundancia de los órdenes SAR11, Flavobacteriales y SAR86.

Orden	ANOVA	Prueba Tukey		
		Este-Atlántico	Oeste-Atlántico	Este-Oeste
Flavobacteriales	$H_1 : \exists \mu_i het$ p<0.001	$H_1 : \mu_E < \mu_A$ p<0.001	$H_0 : \mu_W = \mu_A$ p=0.53	$H_1 : \mu_E < \mu_W$ p<0.001
SAR11	$H_1 : \exists \mu_i het$ p<0.001	$H_1 : \mu_E < \mu_A$ p<0.001	$H_0 : \mu_W = \mu_A$ p=0.58	$H_1 : \mu_E < \mu_W$ p<0.001
SAR86	$H_0 : \mu_A = \mu_E = \mu_W$ p=0.29	$H_1 : \mu_E = \mu_A$	$H_0 : \mu_W = \mu_A$	$H_1 : \mu_E = \mu_W$
Rhodospirillales	$H_0 : \mu_A = \mu_E = \mu_W$ p=0.20	$H_1 : \mu_E = \mu_A$	$H_0 : \mu_W = \mu_A$	$H_1 : \mu_E = \mu_W$
Opituales	$H_0 : \mu_A = \mu_E = \mu_W$ p=0.60	$H_1 : \mu_E = \mu_A$	$H_0 : \mu_W = \mu_A$	$H_1 : \mu_E = \mu_W$

Debido al mayor número de células heterótrofas en las muestras de la corriente de California, todos los grupos taxonómicos resultaron significativamente más abundantes en esta región productiva que en el mar Mediterráneo y océano Atlántico, indicando una limitación para utilizar la estimación de abundancia de grupos taxonómicos de manera individual.

El PCA permitió visualizar las similitudes entre las muestras a partir de sus estimaciones de abundancias entre las regiones del Mediterráneo y las muestras de la corriente de California (Figura 18). En la figura se observa que las muestras de cada región se ordenan de izquierda (oligotrófico) a derecha (productivo), con las muestras del este del Mediterráneo a la izquierda, seguidas del océano Atlántico, el oeste del Mediterráneo y las muestras de SPOT a la derecha. Se observa que grupos de vida libre relacionados a la superficie (Opituales, Puniceispirillales) se asocian a regiones más oligotróficas, además de relacionarse negativamente con la presencia de arqueas (Nitrosopumilales, Marine Group II), lo que también puede deberse a que son grupos relacionados a mayor profundidad.

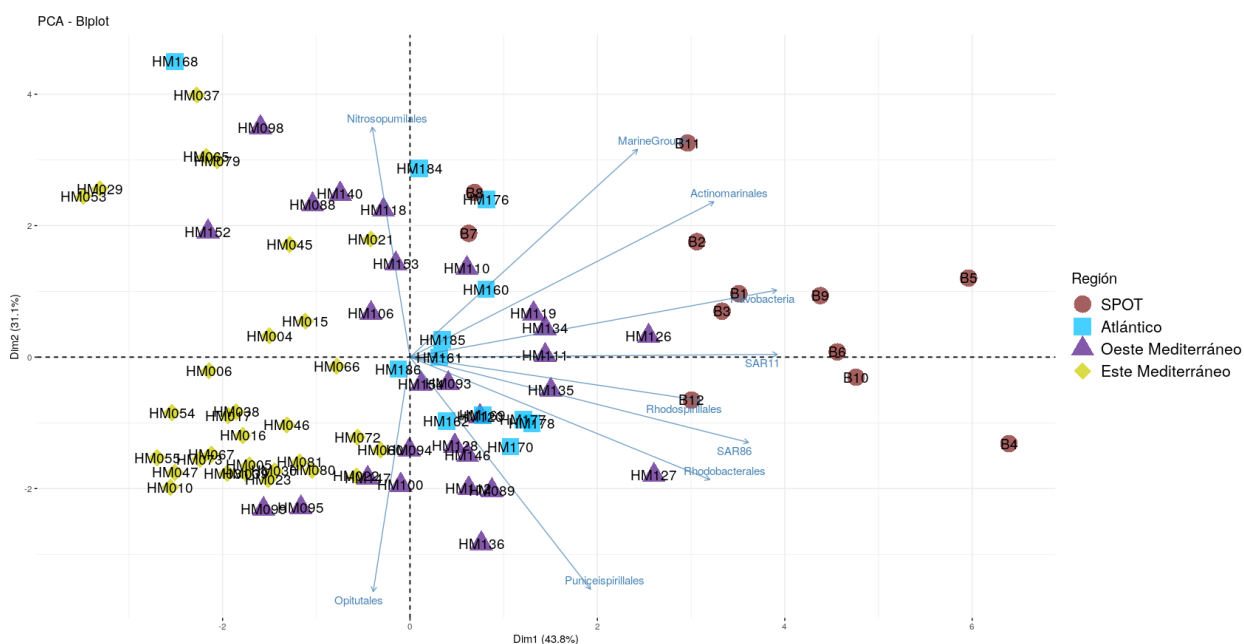


Figura 18. PCA mostrando el la distancia que producen las estimaciones de abundancia de heterótrofos entre muestras con diferente productividad. Las figuras en color verde son muestras que no cuentan con medición de retinal.

Al retirar las muestras que no presentan mediciones de retinal y repetir el PCA (Figura 19), se obtiene un patrón que relaciona la abundancia de grupos heterótrofos conocidos (SAR86, SAR11 y Flavobacteriales) en regiones productivas con retinal, y para regiones oligotrófica relaciona más la abundancia del orden Opitutales y en menor medida Puniceispirillales. La mayoría de las muestras con menores concentraciones de retinal se encuentran en el gradiente de mayor abundancia de arqueas.

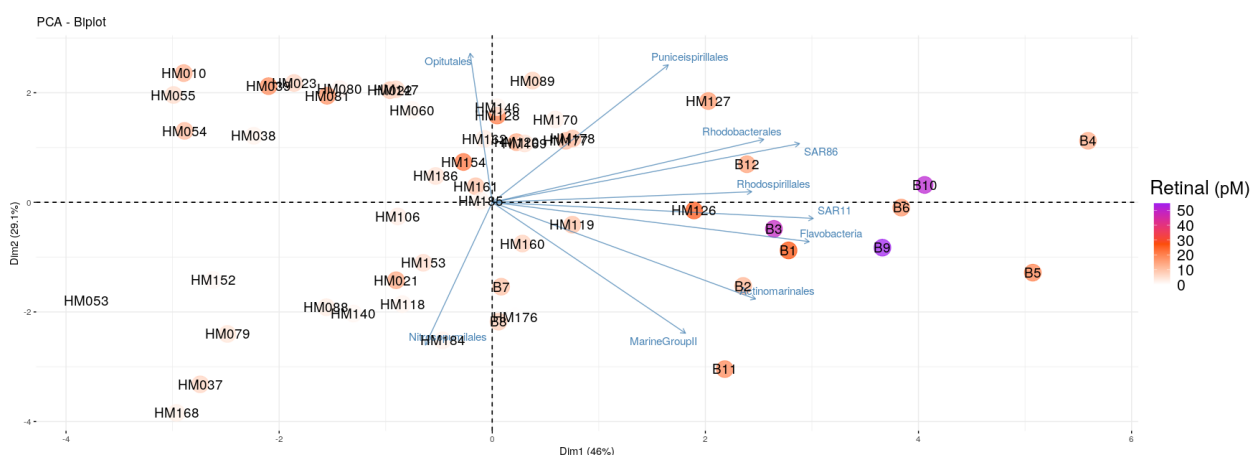


Figura 19. PCA indicando el nivel de retinal en el color de cada punto además de las distancias de las estimaciones de abundancia para los grupos más importantes, mostrando el nivel de retinal en el color. Se omitieron muestras si mediciones de retinal.

3.5. Análisis de correlaciones

Los análisis de correlaciones se realizaron por separado en cada región, mostrándose en las figuras 20 a 24. En los parámetros utilizados los nombres taxonómicos a nivel de orden corresponden a estimaciones de abundancia, mientras que los géneros *Synochococcus* y *Prochlorococcus* corresponden directamente a conteos por citometría. La figura 20 muestra las correlaciones entre las variables de los datos del Atlántico, se observan correlaciones altas entre el retinal y los órdenes SAR86 y Cellvibrionales. La clorofila se correlaciona positivamente con los conteos de *Synochococcus* y *Prochlorococcus* obtenidos por citometría, con los órdenes de arqueas Nitrosopumilales y Marine Group II y con el orden Actinomarinales. Además se correlaciona negativamente con órdenes taxonómicos presentes en la superficie (Puniceispirillales, Opitutales y Cellvibrionales). Se observa que los grupos que se correlacionan positivamente con la clorofila (*Synochococcus*, *Prochlorococcus*, Nitrosopumilales, Marine Group II) y los que se correlacionan negativamente (Puniceispirillales, Opitutales, Cellvibrionales) se correlacionan positivamente entre sí, y negativamente con los otros grupos.

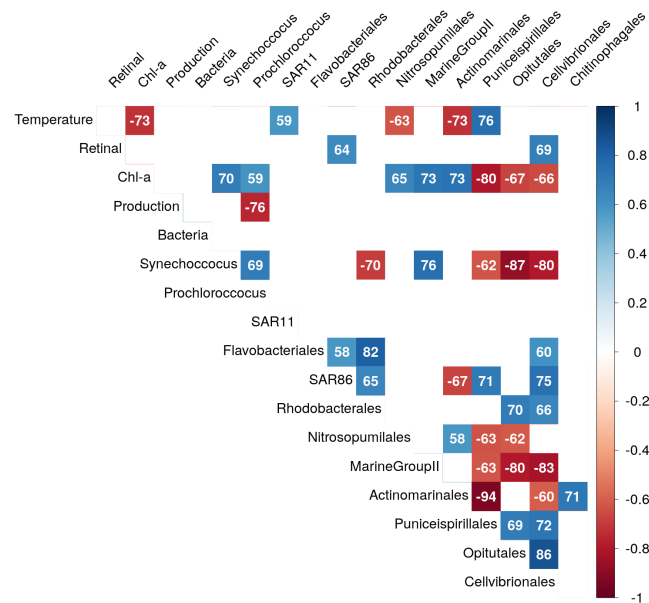


Figura 20. Matriz de correlaciones obtenida para los datos del Atlántico, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.

Para la región del oeste del Mediterráneo (Figura 21), el retinal presenta correlación positiva con la mayoría de los grupos taxonómicos heterótrofos (Flavobacteriales, SAR86, Rhodobacteriales, Puniceispirillales, Opitutales, Cellvibrionales y Chitinophagales). A pesar de esto no se presenta correlación con el grupo SAR11, considerado el principal responsable de expresar genes de proteorodopsinas (Sieradz-

ki et al., 2018). La producción bacteriana también presentó correlaciones positivas (Rhodobacterales y Opitutales) y negativas (*Prochlorococcus*, Nitrosopumilales y Actinomarinales) con grupos taxonómicos específicos en la región del oeste del Mediterráneo.

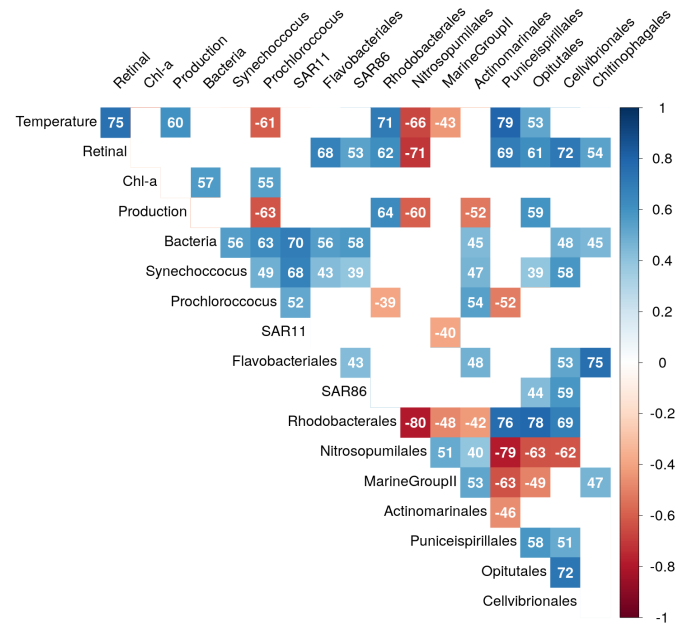


Figura 21. Matriz de correlaciones obtenida para los datos del oeste del Mediterráneo, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.

En el este del Mediterráneo (Figura 22) no se encuentran correlaciones del retinal con ningún grupo taxonómico, a pesar de que en esta región se obtuvieron los valores máximos. La clorofila presenta correlaciones positivas con *Prochlorococcus*, Nitrosopumilales, Marine Group II y Actinomarinales; y correlaciones negativas con Puniceispirillales, Opitutales y Cellvibrionales, al igual que en el Atlántico. Se vuelve a observar el mismo patrón en el que los órdenes taxonómicos de arqueas y los Actinomarinales se correlacionan positivamente entre sí y negativamente con los órdenes heterótrofos de vida libre.

Las mediciones de retinal de la corriente de California (Figura 23) presentan correlación con los grupos Flavobacteriales y Rhodobacterales, además de correlacionarse con la abundancia de bacterias heterótrofas totales (Bacteria). La clorofila se correlaciona positivamente con los órdenes Flavobacteriales y Chitinophagales. La mayoría de grupos taxonómicos heterótrofos se correlacionan positivamente entre sí, con excepción del orden de arqueas Nitrosopumilales, que presenta correlación negativa con Rhodobacterales y Cellvibrionales.

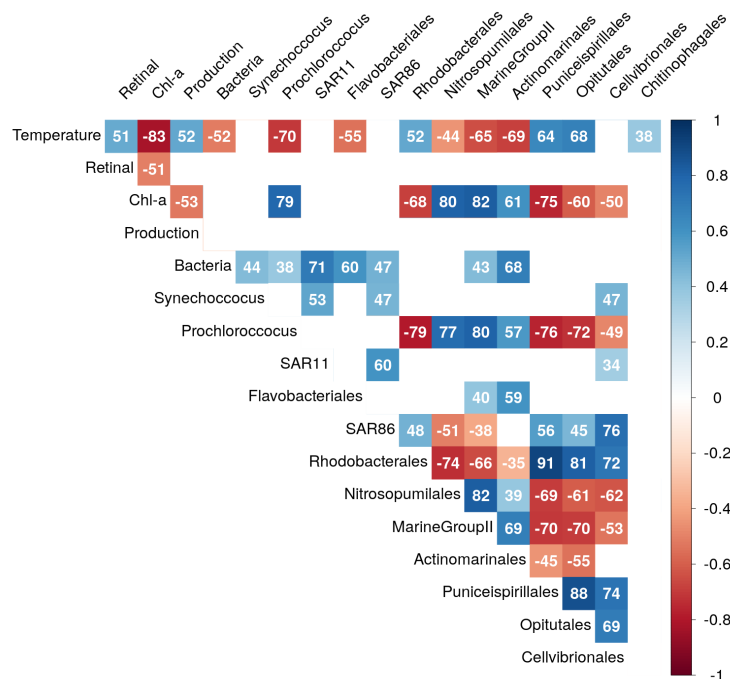


Figura 22. Matriz de correlaciones obtenida para los datos del este del Mediterráneo, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.

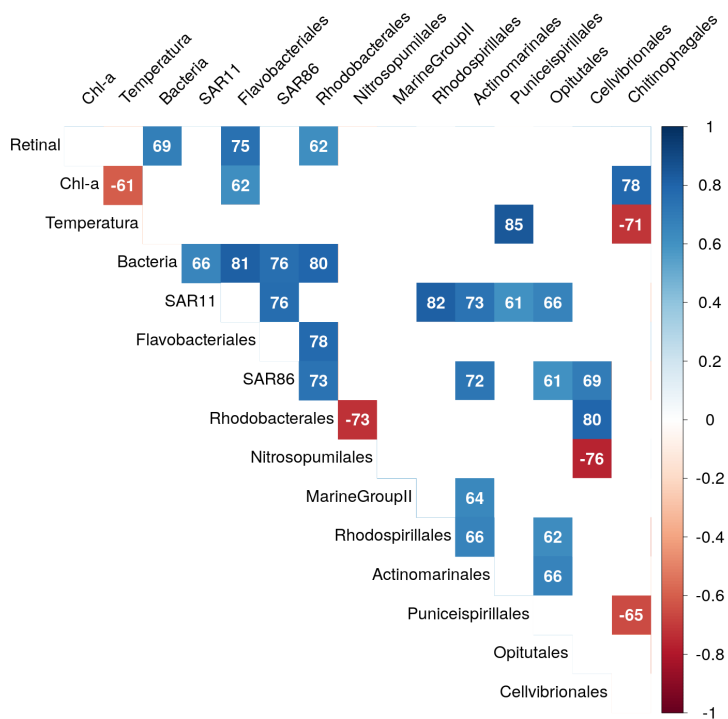


Figura 23. Matriz de correlaciones obtenida para los datos de la corriente de California, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.

Para los datos del golfo de México (Figura 24), los grupos taxonómicos son representados con abundancias relativas, debido a que no se contaba con la abundancia de bacterias heterótrofas totales. No se encontraron correlaciones entre las mediciones de fluorescencia con las abundancias relativas de los grupos a nivel de orden, únicamente se encontró una correlación negativa entre la fluorescencia y la abundancia relativa del orden Synechococales. En cambio, se encontró que las abundancias relativas de Flavobacteriales, Rhodospirillales, Puniceispirillales y Rhodobacterales se correlacionan entre sí. De igual manera se encontró que los grupos de arqueas Nitrosopumilales y Marine Group II se correlacionan entre sí. Entre los dos grupos descritos se hallan correlaciones negativas, lo que coincide con lo descrito en las regiones del Mediterráneo y Atlántico, que es menos aparente en la corriente de California.

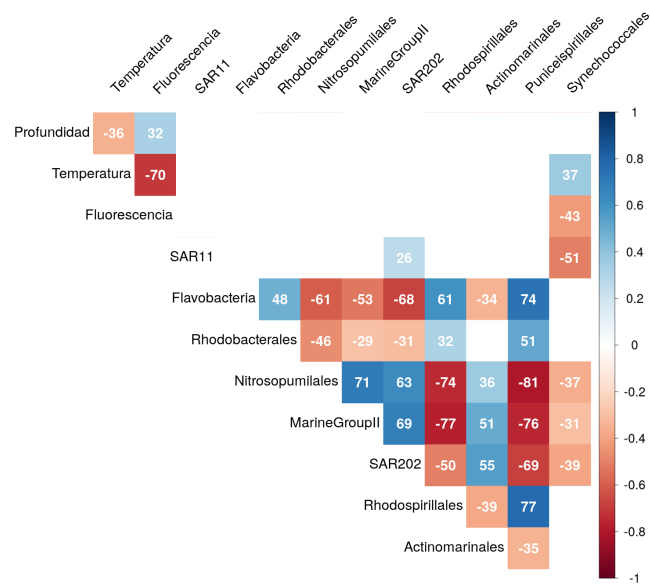


Figura 24. Matriz de correlaciones obtenida para los datos del golfo de México, los coeficientes de correlación se muestran en forma de porcentaje, indicando correlaciones positivas en azul y negativas en rojo.

3.6. Comparación de secuencias entre regiones

La figura 25 muestra la cantidad de secuencias idénticas compartidas por las diferentes regiones independientemente de la estructura de la comunidad. Se observa que la mayor cantidad de ASV compartidos entre regiones que no son contiguas es entre el este del Mediterráneo y el golfo de México. Se observa que las regiones del Mediterráneo comparten una gran cantidad de ASV, lo cual se esperaba al ser regiones conectadas geográficamente.

Las muestras de la corriente de California presentan la mayor cantidad de secuencias compartidas con la región del oeste del Mediterráneo, la única región que también presenta mayores niveles de productividad (Tabla 12). Las muestras del golfo de México comparten el mayor número de secuencias

con la región del este del Mediterráneo, que podría considerarse la más alejada geográficamente, pero la única que comparte similares condiciones de oligotrofia (Tabla 4).

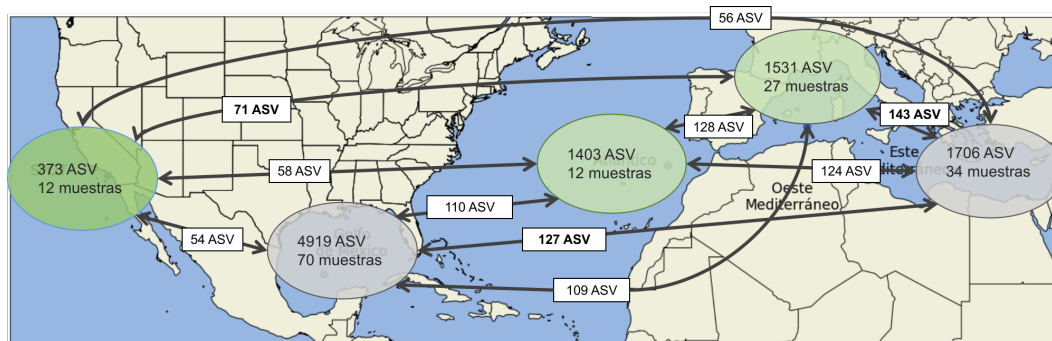


Figura 25. Comparación entre regiones indicando el número de ASV idénticos que comparten.

Capítulo 4. Discusión

4.1. Bacterias fotoheterótrofas y su relación con el retinal

El presente estudio permitió realizar la comparación de comunidades microbianas marinas en regiones con distintos niveles de productividad, analizando los grupos taxonómicos a nivel de orden y su relación con el potencial fotoheterotrófico (medido a través del retinal, pigmento que capta la luz en las proteorodopsinas) y otros parámetros ambientales. Estudiar la estructura de la comunidad de una muestra permite identificar a los grupos más abundantes en el agua, ayudando a entender los factores del medio que determinan su nicho ecológico y que tienen mayor influencia en su dominio en ciertos ambientes. Dentro de los grupos taxonómicos mayormente fotoheterótrofos, el clado SAR11 presenta una abundancia relativa importante, normalmente entre 20 y 45 %, independientemente de la concentración de retinal. A excepción de las muestras del océano Atlántico, bacterias del grupo SAR11 dominan la estructura de la comunidad en muestras superficiales a lo largo de todo el trayecto del Mediterráneo. En las muestras superficiales de la corriente de California (5 metros de profundidad), SAR11 también resultó ser el grupo más abundante. Únicamente en las muestras del golfo de México, el grupo SAR11 no domina la estructura de la comunidad. Esta observación probablemente se explique por la subestimación de los cebadores utilizados en las muestras del golfo de México, pero de igual manera se identifican secuencias ASV clasificadas como SAR11 en todas las muestras. A pesar de ser el grupo taxonómico generalmente más abundante en el océano (Giovannoni, 2017) y el principal responsable de la expresión de proteorodopsina (Sieradzki et al., 2018), no fue posible encontrar correlaciones entre SAR11 y retinal u otros grupos taxonómicos en ninguna de las regiones. En el oeste del Mediterráneo, la estimación de abundancia del grupo SAR11 únicamente presentó correlaciones positivas con el total de bacterias heterótrofas y con las abundancias de *Synochococcus* y *Prochlorococcus*. Lo primero probablemente debido a que la estimación de abundancia es una variable derivada del conteo de heterótrofos, mientras que lo segundo indica una disminución simultánea con las bacterias autótrofas en una región ligeramente productiva (oeste del Mediterráneo).

El orden Flavobacteriales resultó ser el grupo más abundante en las muestras de la corriente de California con máximos niveles de retinal (abril y mayo). Este patrón se observó también en las muestras del oeste del Mediterráneo, además de presentarse correlaciones positivas altas entre el retinal y la estimación de Flavobacteriales en ambas regiones en los análisis de correlaciones (Figuras 21 y 23). En la corriente de California también se encontró una correlación positiva entre Flavobacteriales y los niveles de clorofila. El PCA realizado en las estimaciones de abundancia también relaciona los sitios más

productivos y con mayores concentraciones de retinal con una mayor estimación de abundancia de Flavobacteriales (Figuras 19 y 18). En la región del este del Mediterráneo, a pesar de ser la región con mayores concentraciones de retinal (61 pM), mayores que las de la corriente de California (54 pM), la abundancia relativa de los Flavobacteriales fue menor que en el oeste. Las abundancias relativas presentadas en el golfo de México fueron similares a las del este del Mediterráneo, indicando que los Flavobacteriales no son un grupo dominante en ambientes oligotróficos. Lo anterior confirma la importancia de este orden taxonómico durante los florecimientos de fitoplancton para la degradación de materia orgánica compleja en ambientes productivos (Williams et al., 2013; Bauer et al., 2006).

El grupo SAR86 presentó una mayor abundancia relativa en las muestras del este del Mediterráneo que en el oeste (Sebastián et al., 2021). En cambio, el análisis de varianzas para las estimaciones de abundancias de SAR86 demostró que no se presentan cambios significativos de abundancia a través de todo el trayecto, a pesar de que el total de bacterias heterótrofas disminuye significativamente al entrar al este del Mediterráneo. Lo anterior indica un papel importante en regiones oligotróficas para este orden taxonómico. A pesar de que las muestras del golfo de México también presentan una mayor abundancia relativa para SAR86, esto puede deberse a la sobreestimación de gamaproteobacterias presentada por los cebadores usados como se reporta en (Parada et al., 2016). La abundancia de SAR86 presenta una correlación positiva con el retinal en las regiones del Atlántico y el oeste del Mediterráneo, pero dicha correlación deja de ser significativa en la corriente de California, donde también presenta la menor abundancia relativa para SAR86 de todas las regiones. A pesar de esta aparente preferencia por ambientes que no son altamente productivos, el PCA para estimaciones de abundancia asocia la abundancia de SAR86 a muestras más productivas y con mayores concentraciones de retinal (Figuras 19 y 18).

Los órdenes Opitutales y Puniceispirillales se hallaron en mayor abundancia en muestras de la superficie en las regiones del Mediterráneo. Presentaron una correlación positiva con el retinal en la región del oeste del Mediterráneo, mientras que en las muestras de la corriente de California ya no forma parte de los grupos más abundantes. Esto coincide con lo descrito para SAR86, además de que en el análisis de varianzas para las estimaciones de abundancias, tampoco se presentaron diferencias significativas a través del trayecto del Mediterráneo. En los PCA también se relaciona la mayor abundancia de estos dos órdenes taxonómicos con las muestras del este del Mediterráneo con concentraciones de retinal (Figuras 19 y 18).

Las correlaciones de órdenes taxonómicos con el retinal no necesariamente se deben a la producción de proteorodopsinas, sino más bien a la naturaleza de estos grupos taxonómicos que presentan ventajas selectivas en ambientes donde también se ve favorecido el metabolismo fotoheterótrofo de la proteo-

rodopsina. Los grupos taxonómicos mencionados contienen genomas que han reportado la presencia de genes de proteorodopsinas, pero no necesariamente sugiere que la mayoría de los miembros de este grupo sean fotoheterótrofos como el caso de SAR11 (Sieradzki et al., 2018). En cambio se observa una mayor cantidad de correlaciones entre los grupos taxonómicos que con variables ambientales, como fue reportado por Needham y Fuhrman (2016). Se puede observar una división entre los órdenes taxonómicos analizados en los análisis de correlaciones, donde dos grupos se correlacionan positivamente entre sí (SAR11, Flavobacteriales, SAR86, Puniceispirillales, Opitutales, Rhodospirillales) y negativamente con los miembros del otro grupo (Nitrosopumilales, Marine Group II, Actinomarinales, Synechococcales). Este patrón se puede observar también en los PCA de los análisis de estimaciones de abundancias (Figuras 19 y 18), donde estos grupos taxonómicos se relacionan con muestras de regiones con bajas concentraciones de retinal y muestras de mayor profundidad. Lo anterior sugiere que de los grupos taxonómicos fotoheterótrofos aumentan su abundancia de manera simultánea, mientras que su correlación negativa con el otro grupo indica desventajas selectivas que pueden deberse a condiciones del medio, o a la mayor profundidad donde los miembros del segundo grupo dominan.

4.2. Fotoheterotrofia en regiones con distintos niveles de productividad

Los sitios de estudio presentan diferentes niveles de productividad que van desde regiones altamente oligotróficas (este del Mediterráneo, golfo de México), regiones oligotróficas ligeramente productivas (oeste del Mediterráneo), regiones con nutrientes, pero sin productividad (Atlántico) y regiones altamente productivas (corriente de California). Dichas regiones describen un gradiente de productividad que se puede observar en la figura 18 en el análisis de estimaciones de abundancia de los grupos bacterianos a nivel de orden más abundantes, observándose un agrupamiento de las regiones oligotróficas a la izquierda (este Mediterráneo) y las más productivas a la derecha (corriente de California), mientras que las muestras de las regiones del Atlántico y el oeste del Mediterráneo quedan en el centro. Al dejar únicamente las muestras con datos disponibles de retinal en la figura 19, se observa una separación entre tres grupos, indicando posiblemente tres tipos de ambientes. Las muestras de ambientes productivos y con mayor número de células hacia la izquierda y al centro, las muestras de ambientes oligotróficos con niveles de retinal y finalmente las muestras con pocas concentraciones de retinal y con mayor abundancia relativa de Nitrosopumilales.

La agrupación de las muestras con bajas concentraciones de retinal y abundancia de Nitrosopumilales puede ser debida a la distribución de este orden taxonómico que es más abundante a mayores profundidades, mientras que los niveles de retinal disminuyen naturalmente al aumentar la profundidad. En

cambio, la división presentada entre las muestras con alta productividad y las oligotróficas sugiere que la mayor abundancia relativa de los órdenes taxonómicos Opituales y Puniceispirillales está relacionada a ambientes mayormente oligotróficos. A diferencia del orden SAR86, que a pesar de presentar una mayor abundancia relativa en el este del Mediterráneo que en el oeste y el Atlántico, también se observa en las muestras superficiales de SPOT.

La abundancia de bacterias heterótrofas totales en las muestras de la corriente de California se encontraban un orden de magnitud arriba de los reportados para el Mediterráneo, por lo que analizar las estimaciones de abundancia sólo indicó diferencias significativas para todos los grupos taxonómicos, siendo mayor en la región productiva que en cualquier región del trayecto del Mediterráneo. A pesar de esto, en las gráficas de abundancia relativa (Figuras 9A y 10A) se pudo observar que los fotoheterótrofos son los grupos dominantes en la superficie en las regiones del este y oeste del Mediterráneo y en la corriente de California.

Las mediciones de retinal son un buen indicador del nivel de fotoheterotrofia en cada región (Gómez-Consarnau et al., 2019), lo que permite observar los mayores niveles de fotoheterotrofia en la corriente de California y en el este del Mediterráneo. En el primer caso, una región altamente productiva, para las muestras que presentan los máximos niveles de pigmentos (abril y mayo) se observa que la estructura de la comunidad es dominada por los grupos Flavobacteriales, SAR86, Rhodobacterales. Mientras que en el este del Mediterráneo se observa una dominancia del orden SAR11, y una menor abundancia relativa de algunos grupos fotoheterótrofos (Flavobacteriales, Rhodobacterales), además de una mayor abundancia relativa de SAR86, Puniceispirillales y Opituales. Aunque se ha considerado a SAR11, Flavobacteriales y SAR86 como grupos taxonómicos mayormente fotoheterótrofos, la baja abundancia relativa presentada por Flavobacteriales en el este del Mediterráneo, donde se encontraban los valores máximos de retinal, sugieren que la fotoheterotrofia está distribuida entre la mayoría de grupos heterótrofos de vida libre, con diferentes condiciones óptimas para la expresión de proteorodopsinas en diferentes grupos taxonómicos (Pinhassi et al., 2016).

A pesar de que el análisis de beta diversidad (Figura 14) y el análisis de estimación de abundancia (Figura 18) indican una similitud entre las regiones del oeste del Mediterráneo y la corriente de California, las distribuciones de pigmentos se presentan de manera distinta. Mientras que en el oeste del Mediterráneo se observa una división entre las dominancias de retinal y clorofila (Figura 5), en las muestras de SPOT se observa un aumento simultáneo entre las concentraciones de retinal y clorofila. Al ser una región altamente productiva, la alta concentración de retinal puede tener importancia en el procesamiento de materia orgánica producida durante la fotosíntesis (González et al., 2011). En cambio,

las altas concentraciones de retinal en regiones oligotróficas como el este del Mediterráneo tienen un potencial energético similar a la clorofila (Gómez-Consarnau et al., 2019), que podría ser más relevante tomando en cuenta el número de células significativamente menor en comparación con el oeste, que se encuentra adyacente. Las comunidades microbianas del golfo de México, a pesar de los sesgos presentados por los cebadores, presentan similitud con las muestras del Mediterráneo. La estructura de la comunidad incluye a grupos como Flavobacteriales, Puniceispirillales y Rhodobacterales, además de una abundancia relativa importante del orden Synechococales. A pesar de la subestimación, SAR11 sigue presentando una abundancia relativa mayor al 5 %, lo que podría sugerir una abundancia relativa varias veces mayor en las muestras (Parada et al., 2016). Debido a que se trata de una región oligotrófica la fotoheterotrofia debe tener importancia para la comunidad microbiana, probablemente encontrándose en algún punto entre el Atlántico y el este del Mediterráneo.

4.3. Comparaciones entre regiones

El análisis de ASV idénticos compartidos entre regiones también sugiere que la similitud de condiciones ambientales implica una similitud entre comunidades microbianas. Los ASV compartidos entre sitios de estudio alejados confirman la naturaleza cosmopolita de las bacterias relacionadas a la superficie del océano.

Se esperaba que al ser regiones contiguas, la mayoría de los ASV entre las regiones del Mediterráneo presentaran pocas diferencias, sin embargo, el número de ASV compartidos entre el este y el oeste es mayor que con el Atlántico. El estrecho de Gibraltar forma una barrera geográfica que limita las aguas del Mediterráneo a una menor circulación y disponibilidad de nutrientes (Suffridge et al., 2018).

Al analizar la cantidad de ASV idénticos compartidos entre regiones se encontró que las regiones del este del Mediterráneo y del golfo de México son las que comparten una mayor cantidad de secuencias ASV idénticas (127) entre sitios no contiguos. El sesgo presentado por los cebadores durante la amplificación no impiden la detección y asignación de secuencias en las muestras, haciendo posible la detección de secuencias amplificadas de genes 16S-ARNr en sitios distintos de manera simultánea. La naturaleza de los ASV asignados compartidos entre regiones permite dar una idea de sus similitudes.

Capítulo 5. Conclusiones

Los grupos taxonómicos productores de proteorodopsinas presentan diferentes abundancias relativas en ambientes con distinta productividad de acuerdo con sus ventajas selectivas y naturaleza trófica, es decir si es un organismo más adaptado a un estilo de vida más oligotrófico o si responde a la disponibilidad de nutrientes y materia orgánica aumentando su tasa de crecimiento.

Los grupos taxonómicos fotoheterótrofos productores de proteorodopsina SAR11 y SAR86 presentan una mayor abundancia relativa en ambientes oligotróficos que en ambientes productivos. Mientras el orden Flavobacteriales presenta una dominancia en muestras de regiones productivas en fechas cercanas a crecimientos de fitoplancton, además de aumentar su abundancia relativa en el oeste del Mediterráneo en regiones con mayor productividad.

La producción de proteorodopsina en ambientes oligotróficos o productivos difícilmente puede asociarse a un grupo taxonómico específico, ya que la fotoheterotrofia por retinal puede estar dispersa entre la comunidad bacteriana y cada grupo puede responder a distintos estímulos para su expresión.

La presencia de ASV idénticos entre las regiones del este del Mediterráneo y el golfo de México indican una mayor similitud entre los ambientes oligotróficos que con regiones productivas.

Literatura citada

- Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., y Thingstad, F.. 1983. The Ecological Role of Water-Column Microbes in the Sea. *Marine Ecology Progress Series*, 10. doi: 10.3354/meps010257.
- Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., Würdemann, C. A., Quast, C., Kuhl, H., Knaust, F., Woebken, D., Bischof, K., Mussmann, M., Choudhuri, J. V., Meyer, F., Reinhardt, R., Amann, R. I., y Glöckner, F. O.. 2006. Whole genome analysis of the marine Bacteroidetes 'Gramella forsetii' reveals adaptations to degradation of polymeric organic matter. *Environmental Microbiology*, 8(12). doi: 10.1111/j.1462-2920.2006.01152.x.
- Béjà, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., Jovanovich, S. B., Gates, C. M., Feldman, R. A., Spudich, J. L., Spudich, E. N., y DeLong, E. F.. 2000. Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science*, 289(5486), pp. 1902–1906. doi: 10.1126/science.289.5486.1902.
- Béjà, O., Spudich, E. N., Spudich, J. L., Leclerc, M., y DeLong, E. F.. 2001. Proteorhodopsin phototrophy in the ocean. *Nature*, 411(6839), pp. 786–789. doi: 10.1038/35081051.
- Boettiger, C. y Eddelbuettel, D.. 2017. An introduction to Rocker: Docker containers for R. *R Journal*, 9(2). doi: 10.32614/rj-2017-065.
- Callahan, B. J., McMurdie, P. J., y Holmes, S. P.. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, 11(12). doi: 10.1038/ismej.2017.119.
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., y Knight, R.. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America*, 108(SUPPL. 1). doi: 10.1073/pnas.1000080107.
- Carini, P., Steindler, L., Beszteri, S., y Giovannoni, S. J.. 2013. Nutrient requirements for growth of the extreme oligotroph 'Candidatus Pelagibacter ubique' HTCC1062 on a defined medium. *ISME Journal*, 7(3). doi: 10.1038/ismej.2012.122.
- Chao, A.. 1984. Non-parametric estimation of the classes in a population. *Scandinavian Journal of Statistics*, 11(4).
- Clarke, K. R. y Ainsworth, M.. 1993. A method of linking multivariate community structure to environmental variables. *Marine Ecology Progress Series*, 92(3). doi: 10.3354/meps092205.
- DeLong, E. F.. 2005. Microbial community genomics in the ocean. *Nature Reviews Microbiology*, 3(6). doi: 10.1038/nrmicro1158.
- Dubinsky, V., Haber, M., Burgsdorf, I., Saurav, K., Lehahn, Y., Malik, A., Sher, D., Aharonovich, D., y Steindler, L.. 2017. Metagenomic analysis reveals unusually high incidence of proteorhodopsin genes in the ultraoligotrophic Eastern Mediterranean Sea. *Environmental Microbiology*, 19(3), pp. 1077–1090. doi: 10.1111/1462-2920.13624.
- Ducklow, H. W., Kirchman, D. L., Quinby, H. L., Carlson, C. A., y Dam, H. G.. 1993. Stocks and dynamics of bacterioplankton carbon during the spring bloom in the eastern North Atlantic Ocean. *Deep-Sea Research Part II*, 40(1-2). doi: 10.1016/0967-0645(93)90016-G.
- Dupont, C. L., Rusch, D. B., Yooseph, S., Lombardo, M. J., Alexander Richter, R., Valas, R., Novotny, M., Yee-Greenbaum, J., Selengut, J. D., Haft, D. H., Halpern, A. L., Lasken, R. S., Nealson, K., Friedman, R., y Craig Venter, J.. 2012. Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME Journal*, 6(6). doi: 10.1038/ismej.2011.189.

- Eiler, A.. 2006. Evidence for the ubiquity of mixotrophic bacteria in the upper ocean: Implications and consequences. *Applied and Environmental Microbiology*, 72(12). doi: 10.1128/AEM.01559-06.
- Ernst, O. P., Lodowski, D. T., Elstner, M., Hegemann, P., Brown, L. S., y Kandori, H.. 2014. Microbial and animal rhodopsins: Structures, functions, and molecular mechanisms. *Chemical Reviews*, 114(1), pp. 126–163. doi: 10.1021/cr4003769.
- Falkowski, P. G., Fenchel, T., y Delong, E. F.. 2008. The microbial engines that drive earth's biogeochemical cycles. *Science*, 320(5879). doi: 10.1126/science.1153213.
- Fenchel, T.. 2008. The microbial loop - 25 years later. *Journal of Experimental Marine Biology and Ecology*, 366(1-2). doi: 10.1016/j.jembe.2008.07.013.
- Finkel, O. M., Bèjà, O., y Belkin, S.. 2013. Global abundance of microbial rhodopsins. *ISME Journal*, 7, pp. 448–451. doi: 10.1038/ismej.2012.112.
- Fuhrman, J. A.. 2009. Microbial community structure and its functional implications. *Nature*, 459(7244). doi: 10.1038/nature08058.
- Fukuda, K., Ogawa, M., Taniguchi, H., y Saito, M.. 2016. Molecular approaches to studying microbial communities: Targeting the 16S ribosomal RNA gene. *Journal of UOEH*, 38(3). doi: 10.7888/juoeh.38.223.
- Gasol, J. M. y Kirchman, D. L.. 2018. *Microbial Ecology of the Oceans*. 3a ed., Wiley-Blackwell.
- Gilbert, J. A. y Dupont, C. L.. 2011. Microbial metagenomics: Beyond the genome. *Annual Review of Marine Science*, 3, pp. 347–71. doi: 10.1146/annurev-marine-120709-142811.
- Giovannoni, S. J.. 2017. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. *Annual Review of Marine Science*, 9(1). doi: 10.1146/annurev-marine-010814-015934.
- Giovannoni, S. J., Bibbs, L., Cho, J. C., Stapels, M. D., Desiderio, R., Vergin, K. L., Rappé, M. S., Laney, S., Wilhelm, L. J., Tripp, H. J., Mathur, E. J., y Barofsky, D. F.. 2005a. Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature*, 438(7064). doi: 10.1038/nature04032.
- Giovannoni, S. J., Tripp, H. J., Givan, S., Podar, M., Vergin, K. L., Baptista, D., Bibbs, L., Eads, J., Richardson, T. H., Noordewier, M., Rappé, M. S., Short, J. M., Carrington, J. C., y Mathur, E. J.. 2005b. Genetics: Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, 309(5738). doi: 10.1126/science.1114057.
- Gómez-Consarnau, L., González, J. M., Coll-Lladó, M., Gourdon, P., Pascher, T., Neutze, R., Pedrós-Alió, C., y Pinhassi, J.. 2007. Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature*, 445(7124), pp. 210–213. doi: 10.1038/nature05381.
- Gómez-Consarnau, L., Raven, J. A., Levine, N. M., Cutter, L. S., Wang, D., Seegers, B., Arístegui, J., Fuhrman, J. A., Gasol, J. M., y Sañudo-Wilhelmy, S. A.. 2019. Microbial rhodopsins are major contributors to the solar energy captured in the sea. *Science Advances*, 5(8). doi: 10.1126/sciadv.aaw8855.
- Gómez-Pereira, P. R., Schüler, M., Fuchs, B. M., Bennke, C., Teeling, H., Waldmann, J., Richter, M., Barbe, V., Bataille, E., Glöckner, F. O., y Amann, R.. 2012. Genomic content of uncultured Bacteroidetes from contrasting oceanic provinces in the North Atlantic Ocean. *Environmental Microbiology*, 14(1). doi: 10.1111/j.1462-2920.2011.02555.x.

- González, J. M., Pinhassi, J., Fernández-Gómez, B., Coll-Lladó, M., González-Velázquez, M., Puigbò, P., Jaenicke, S., Gómez-Consarnau, L., Fernández-Guerra, A., Goesmann, A., y Pedrós-Alió, C.. 2011. Genomics of the proteorhodopsin-containing marine flavobacterium *Dokdonia* sp. strain MED134. *Applied and Environmental Microbiology*, 77(24). doi: 10.1128/AEM.06152-11.
- Hoarfrost, A., Nayfach, S., Ladau, J., Yooseph, S., Arnosti, C., Dupont, C. L., y Pollard, K. S.. 2020. Global ecotypes in the ubiquitous marine clade SAR86. *ISME Journal*, 14(1). doi: 10.1038/s41396-019-0516-7.
- Kembel, S. W., Wu, M., Eisen, J. A., y Green, J. L.. 2012. Incorporating 16S Gene Copy Number Information Improves Estimates of Microbial Diversity and Abundance. *PLoS Computational Biology*, 8(10). doi: 10.1371/journal.pcbi.1002743.
- Kim, B. R., Shin, J., Guevarra, R. B., Lee, J. H., Kim, D. W., Seol, K. H., Lee, J. H., Kim, H. B., y Isaacson, R. E.. 2017. Deciphering diversity indices for a better understanding of microbial communities. *Journal of Microbiology and Biotechnology*, 27(12). doi: 10.4014/jmb.1709.09027.
- Kirchman, D. L.. 2002. The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiology Ecology*, 39(2). doi: 10.1016/S0168-6496(01)00206-9.
- Kirchman, D. L. y Hanson, T. E.. 2013. Bioenergetics of photoheterotrophic bacteria in the oceans. *Environmental Microbiology Reports*, 5(2). doi: 10.1111/j.1758-2229.2012.00367.x.
- Lauro, F. M., McDougald, D., Thomas, T., Williams, T. J., Egan, S., Rice, S., DeMaere, M. Z., Ting, L., Ertan, H., Johnson, J., Ferreira, S., Lapidus, A., Anderson, I., Kyrpides, N., Munkf, A. C., Detterg, C., Hang, C. S., Brown, M. V., Robb, F. T., Kjelleberg, S., y Cavicchioli, R.. 2009. The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37). doi: 10.1073/pnas.0903507106.
- Longhurst, A. R.. 2007. *Ecological Geography of the Sea*. 2a ed., Elsevier Inc.
- Martin, M.. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, 17(1). doi: 10.14806/ej.17.1.200.
- Martinez, A., Bradley, A. S., Waldbauer, J. R., Summons, R. E., y DeLong, E. F.. 2007. Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proceedings of the National Academy of Sciences of the United States of America*, 104(13). doi: 10.1073/pnas.0611470104.
- McKnight, D. T., Huerlimann, R., Bower, D. S., Schwarzkopf, L., Alford, R. A., y Zenger, K. R.. 2019. Methods for normalizing microbiome data: An ecological perspective. *Methods in Ecology and Evolution*, 10(3). doi: 10.1111/2041-210X.13115.
- McLaren, M. R. y Callahan, B. J.. 2021. Silva 138.1 prokaryotic SSU taxonomic training data formatted for DADA2. doi: 10.5281/ZENODO.4587955.
- McMurdie, P. J. y Holmes, S.. 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4). doi: 10.1371/journal.pone.0061217.
- Morris, R. M., Rappé, M. S., Connon, S. A., Vergin, K. L., Siebold, W. A., Carlson, C. A., y Giovannoni, S. J.. 2002. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*, 420(6917). doi: 10.1038/nature01240.
- Murali, A., Bhargava, A., y Wright, E. S.. 2018. IDTAXA: A novel approach for accurate taxonomic classification of microbiome sequences. *Microbiome*, 6(1). doi: 10.1186/s40168-018-0521-5.

- Nagata, T. y Inoue, K.. 2021. Rhodopsins at a glance. *Journal of Cell Science*, 134(22). doi: 10.1242/jcs.258989.
- Needham, D. M. y Fuhrman, J. A.. 2016. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nature microbiology*, 1. doi: 10.1038/nmicrobiol.2016.5.
- Parada, A. E., Needham, D. M., y Fuhrman, J. A.. 2016. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, 18(5). doi: 10.1111/1462-2920.13023.
- Pernthaler, J.. 2005. Predation on prokaryotes in the water column and its ecological implications. *Nature Reviews Microbiology*, 3(7). doi: 10.1038/nrmicro1180.
- Pinhassi, J., DeLong, E. F., Bèjà, O., González, J. M., y Pedrós-Alió, C.. 2016. Marine Bacterial and Archaeal Ion-Pumping Rhodopsins: Genetic Diversity, Physiology, and Ecology. *Microbiology and Molecular Biology Reviews*. doi: 10.1128/mubr.00003-16.
- Price, M. N., Dehal, P. S., y Arkin, A. P.. 2010. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, 5(3). doi: 10.1371/journal.pone.0009490.
- Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A. H., Nieuwdorp, M., y Levin, E.. 2020. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS ONE*, 15(1). doi: 10.1371/journal.pone.0227434.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., y Glöckner, F. O.. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1). doi: 10.1093/nar/gks1219.
- Scanlan, D. J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W. R., Post, A. F., Hagemann, M., Paulsen, I., y Partensky, F.. 2009. Ecological Genomics of Marine Picocyanobacteria. *Microbiology and Molecular Biology Reviews*, 73(2). doi: 10.1128/mubr.00035-08.
- Sebastián, M., Ortega-Retuerta, E., Gómez-Consarnau, L., Zamanillo, M., Álvarez, M., Arístegui, J., y Gasol, J. M.. 2021. Environmental gradients and physical barriers drive the basin-wide spatial structuring of Mediterranean Sea and adjacent eastern Atlantic Ocean prokaryotic communities. *Limnology and Oceanography*, 66(12). doi: 10.1002/lno.11944.
- Sieradzki, E. T., Fuhrman, J. A., Rivero-Calle, S., y Gómez-Consarnau, L.. 2018. Proteorhodopsins dominate the expression of phototrophic mechanisms in seasonal and dynamic marine picoplankton communities. *PeerJ*. doi: 10.7717/peerj.5798.
- Steindler, L., Schwalbach, M. S., Smith, D. P., Chan, F., y Giovannoni, S. J.. 2011. Energy starved candidatus pelagibacter ubique substitutes light-mediated ATP production for endogenous carbon respiration. *PLoS ONE*, 6(5). doi: 10.1371/journal.pone.0019725.
- Suffridge, C. P., Gómez-Consarnau, L., Monteverde, D. R., Cutter, L., Arístegui, J., Alvarez-Salgado, X. A., Gasol, J. M., y Sañudo-Wilhelmy, S. A.. 2018. B Vitamins and Their Congeners as Potential Drivers of Microbial Community Composition in an Oligotrophic Marine Ecosystem. *Journal of Geophysical Research: Biogeosciences*, 123(9). doi: 10.1029/2018JG004554.
- Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R., Alberti, A., Cornejo-Castillo, F. M., Costea, P. I., Cruaud, C., D'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J. M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B. T., Royo-Llonch, M., Sarmiento, H., Vieira-Silva, S., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Boss, E., Follows, M., Karp-Boss, L., Krzic, U.,

- Reynaud, E. G., Sardet, C., Sieracki, M., Velayoudon, D., Bowler, C., De Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M. B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S. G., y Bork, P. 2015. Structure and function of the global ocean microbiome. *Science*, 348(6237). doi: 10.1126/science.1261359.
- Villarreal Peralta, E.. 2021. Efecto de la disponibilidad de nutrientes sobre la regulación fototrófica del fitoplancton y bacterias con proteorodopsinas (PRs). Tesis de maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California.
- Wang, Q., Garrity, G. M., Tiedje, J. M., y Cole, J. R.. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73(16). doi: 10.1128/AEM.00062-07.
- Whitman, W. B., Coleman, D. C., y Wiebe, W. J.. 1998. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12). doi: 10.1073/pnas.95.12.6578.
- Williams, T. J., Wilkins, D., Long, E., Evans, F., Demaere, M. Z., Raftery, M. J., y Cavicchioli, R.. 2013. The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology*, 15(5). doi: 10.1111/1462-2920.12017.

Anexo

Tabla 4. Clasificación taxonómica de los ASV compartidos entre el este del Mediterráneo y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Alphaproteobacteria;SAR11 clade	116
Bacteroidia;Flavobacteriales	99
Gammaproteobacteria;SAR86 clade	59
Alphaproteobacteria;Rhodospirillales	49
Cyanobacteria;Synechococcales	46
Thermoplasmata;Marine Group II	33
Bdellovibrionia;Bdellovibrionales	27
Alphaproteobacteria;Puniceispirillales	23
Nitrososphaeria;Nitrosopumilales	12
Verrucomicrobiae;Opitutales	12
Acidimicrobiia;Actinomarinales	11
Alphaproteobacteria;Rickettsiales	11
Dehalococcoidia;SAR202 clade	11
Gammaproteobacteria;Coxiellales	11
Gammaproteobacteria;Oceanospirillales	11

Tabla 5. Clasificación taxonómica de los ASV compartidos entre el este del Mediterráneo y el oeste del Mediterráneo a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Alphaproteobacteria;SAR11 clade	136
Bacteroidia;Flavobacteriales	119
Gammaproteobacteria;SAR86 clade	64
Cyanobacteria;Synechococcales	41
Alphaproteobacteria;Puniceispirillales	39
Gammaproteobacteria;Cellvibrionales	27
Nitrososphaeria;Nitrosopumilales	27
Alphaproteobacteria;Rhodospirillales	25
Thermoplasmata;Marine Group II	24
Alphaproteobacteria;Rickettsiales	23
Alphaproteobacteria;Rhodobacterales	20
Dehalococcoidia;SAR202 clade	19
Verrucomicrobiae;Opitutales	17
Acidimicrobiia;Microtrichales	12
Acidimicrobiia;Actinomarinales	11

Tabla 6. Clasificación taxonómica de los ASV compartidos entre el Atlántico y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Alphaproteobacteria;SAR11 clade	129
Bacteroidia;Flavobacteriales	88
Alphaproteobacteria;Rhodospirillales	42
Alphaproteobacteria;Puniceispirillales	41
Thermoplasmata;Marine Group II	41
Gammaproteobacteria;SAR86 clade	37
Cyanobacteria;Synechococcales	32
Bdellovibrionia;Bdellovibrionales	23
Nitrososphaeria;Nitrosopumilales	15
Alphaproteobacteria;Rickettsiales	13
Acidimicrobiia;Actinomarinales	12
Alphaproteobacteria;Defluviococcales	11
Gammaproteobacteria;Cellvibrionales	9
Acidimicrobiia;Microtrichales	8
Bacteroidia;Cytophagales	8

Tabla 7. Clasificación taxonómica de los ASV compartidos entre el este del Mediterráneo y el Atlántico a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Alphaproteobacteria;SAR11 clade	148
Bacteroidia;Flavobacteriales	82
Cyanobacteria;Synechococcales	43
Gammaproteobacteria;SAR86 clade	42
Alphaproteobacteria;Puniceispirillales	34
Alphaproteobacteria;Rhodospirillales	34
Thermoplasmata;Marine Group II	23
Alphaproteobacteria;Rhodobacterales	17
Alphaproteobacteria;Rickettsiales	17
Bdellovibrionia;Bdellovibrionales	15
Acidimicrobiia;Actinomarinales	14
Gammaproteobacteria;Cellvibrionales	14
Alphaproteobacteria;Defluviococcales	13
Nitrososphaeria;Nitrosopumilales	13
Verrucomicrobiae;Opitutales	12

Tabla 8. Número de ASV compartidos entre el oeste del Mediterráneo y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Bacteroidia;Flavobacteriales	104
Alphaproteobacteria;SAR11 clade	85
Gammaproteobacteria;SAR86 clade	49
Thermoplasmata;Marine Group II	33
Alphaproteobacteria;Puniceispirillales	23
Alphaproteobacteria;Rhodospirillales	20
Cyanobacteria;Synechococcales	20
Nitrososphaeria;Nitrosopumilales	15
Gammaproteobacteria;Cellvibrionales	11
Verrucomicrobiae;Opitutales	11
Bdellovibrionia;Bdellovibrionales	10
Dehalococcoidia;SAR202 clade	10
Acidimicrobiia;Actinomarinales	9
Acidimicrobiia;Microtrichales	8
Alphaproteobacteria;Rickettsiales	7

Tabla 9. Clasificación taxonómica de los ASV compartidos entre el oeste del Mediterráneo y el Atlántico a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Alphaproteobacteria;SAR11 clade	119
Bacteroidia;Flavobacteriales	117
Gammaproteobacteria;SAR86 clade	39
Alphaproteobacteria;Puniceispirillales	36
Cyanobacteria;Synechococcales	31
Thermoplasmata;Marine Group II	29
Alphaproteobacteria;Rickettsiales	21
Alphaproteobacteria;Rhodobacterales	19
Alphaproteobacteria;Rhodospirillales	19
Gammaproteobacteria;Cellvibrionales	18
Verrucomicrobiae;Opitutales	13
Acidimicrobiia;Actinomarinales	11
Nitrososphaeria;Nitrosopumilales	11
Alphaproteobacteria;Defluviococcales	10
Acidimicrobiia;Microtrichales	9

Tabla 10. Clasificación taxonómica de los ASV compartidos entre SPOT y el Atlántico a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Bacteroidia;Flavobacteriales	29
Alphaproteobacteria;SAR11 clade	11
Gammaproteobacteria;SAR86 clade	10
Thermoplasmata;Marine Group II	9
Alphaproteobacteria;Parvibaculales	5
Alphaproteobacteria;Puniceispirillales	5
Alphaproteobacteria;Rhodobacterales	5
Alphaproteobacteria;Rhodospirillales	4
Cyanobacteria;Synechococcales	4
Acidimicrobiia;Actinomarinales	3
Gammaproteobacteria;Cellvibrionales	3
Gammaproteobacteria;Ectothiorhodospirales	3
Verrucomicrobiae;Verrucomicrobiales	3
Alphaproteobacteria;Defluviococcales	2
Alphaproteobacteria;Rickettsiales	2

Tabla 11. Número de ASV compartidos entre SPOT y el este del Mediterráneo a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Bacteroidia;Flavobacteriales	28
Alphaproteobacteria;SAR11 clade	11
Gammaproteobacteria;SAR86 clade	9
Thermoplasmata;Marine Group II	7
Alphaproteobacteria;Parvibaculales	5
Alphaproteobacteria;Puniceispirillales	5
Alphaproteobacteria;Rhodobacterales	5
Alphaproteobacteria;Rhodospirillales	4
Cyanobacteria;Synechococcales	4
Gammaproteobacteria;Cellvibrionales	4
Acidimicrobiia;Actinomarinales	3
Alphaproteobacteria;Defluviococcales	3
Gammaproteobacteria;Burkholderiales	3
Rhodothermia;Balneolales	3
Nitrososphaeria;Nitrosopumilales	2

Tabla 12. Clasificación taxonómica de los ASV compartidos entre SPOT y el oeste del Mediterráneo a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Bacteroidia;Flavobacteriales	49
Gammaproteobacteria;SAR86 clade	13
Alphaproteobacteria;SAR11 clade	12
Thermoplasmata;Marine Group II	10
Gammaproteobacteria;Cellvibrionales	9
Alphaproteobacteria;Rhodobacterales	8
Alphaproteobacteria;Parvibaculales	5
Alphaproteobacteria;Puniceispirillales	5
Bacteroidia;Chitinophagales	5
Verrucomicrobiae;Verrucomicrobiales	5
Alphaproteobacteria;Rhodospirillales	4
Cyanobacteriia;Synechococcales	4
Gammaproteobacteria;Burkholderiales	4
Acidimicrobiia;Actinomarinales	3
Alphaproteobacteria;Defluviococcales	3

Tabla 13. Número de ASV compartidos entre SPOT y el golfo de México a nivel de clase y orden. Se muestran los primeros 15.

Clase;Orden	Total
Bacteroidia;Flavobacteriales	29
Alphaproteobacteria;SAR11 clade	11
Thermoplasmata;Marine Group II	11
Gammaproteobacteria;SAR86 clade	10
Alphaproteobacteria;Parvibaculales	4
Alphaproteobacteria;Puniceispirillales	4
Cyanobacteriia;Synechococcales	4
Acidimicrobiia;Actinomarinales	3
Alphaproteobacteria;Rhodobacterales	3
Alphaproteobacteria;Rhodospirillales	3
Planctomycetes;Pirellulales	3
Bacteroidia;Sphingobacteriales	2
Gammaproteobacteria;Cellvibrionales	2
Nitrososphaeria;Nitrosopumilales	2
Rhodothermia;Balneolales	2