



UNIVERSIDAD NACIONAL DE LA PATAGONIA  
SAN JUAN BOSCO

Facultad de Ciencias Naturales

**Bioprospección de secuencias de enzimas  
fucoidanasas en genomas y metagenomas  
microbianos**

Tesis presentada para optar al título de Licenciada en Ciencias Biológicas

Jessica Alejandra Gonzalez

Directora: Dra. Hebe Mónica Dionisi

Asesora: Prof. Isabel Albarracín

Co-Asesora: Ing. Qca. Ruth Salomón

# Agradecimientos

Y llegó el momento de escribir los agradecimientos.. Miro atrás y no me alcanzan las palabras para agradecer a todas las personas que me han acompañado durante esta etapa, voy a hacer mi mejor intento para expresarles mi gratitud. Los últimos dos meses resultaron ser muy difíciles y no hubiese podido llegar hasta acá sin la paciencia, la tolerancia y el cariño de todos ustedes. No podría estar más agradecida... La felicidad es también por haber llegado hasta aquí por y gracias a todos ustedes.. Nada en la vida tendría valor si no hay personas con quien compartirlo. A veces no puedo creer el estar rodeada de tan maravillosas personas.. Es por eso que me voy a tomar un tiempo y espacio de la tesis para agradecerles..

A mi directora de tesis, Dra. Hebe Dionisi por su confianza y por trasmitirme la pasión que pone en su trabajo. A los integrantes del laboratorio de microbiología ambiental: Mariana, Fernando, Matías, Luciano y Marina, por sus colaboraciones y por compartir sus experiencias conmigo.

A los jurados: Leandro Jones, Ruth Salomón y Marcela Cravero por haber aceptado evaluar el trabajo a pesar de que el tiempo era una gran limitante. Sus aportes han contribuido a mejorar la versión final de esta tesis.

A los grandes profesores de la FCN, aquellos con los que me crucé durante este recorrido y que ansío seguir encontrando. Gracias por sus ayudas y compartir sus experiencias de vida, gracias a ellas aprendí y crecí como futura profesional.

A la universidad pública. Estoy orgullosa y agradecida de vivir en un país donde la educación es gratuita y de calidad.

A los compañeros que tuve la suerte de encontrarme en esta carrera: Chimi, Belén V, Romi, Cris, Bren, Anto L, Fakiu, Jessi G, Pope, Ema, Sil S, Fer B, Bel M, Car, Barbi, David, Iña, Maxi, Lari y Jeny. Mi especial agradecimiento a Chimi y Belén V ("Primi") por demostrarme que siempre se puede y por haber estado constantemente preocupadas por mí para que yo pudiera lograr esta meta. Fueron un gran apoyo en el momento más duro que me tocó vivir.

A mis amigas de toda la vida: Ailu, Sofi, Lu L y Taty, que a pesar de vernos muy poco siempre están ahí acompañándome en todo.

A toda mi familia, por su eterna entrega y capacidad de mantenerme la ilusión de que esta meta la iba a alcanzar, a pesar de no entender muy bien qué es lo que estaba haciendo. Quiero

comenzar con mis padres (Irenesita y Alejandro), quienes me enseñaron que con paciencia y perseverancia se puede llegar muy lejos. La fortaleza que me transmitió mi mamá fue fundamental durante estos dos últimos meses para luchar día a día y alcanzar este objetivo. Si no hubiese sido así, este último tiempo hubiera sido imposible de sobrellevar. Papá, hoy me estas acompañando desde otro lado, desde mi corazón. Me cuesta muchísimo expresar el agradecimiento eterno que tengo hacia vos, no existen palabras.. y si las existieran no alcanzarían.. Recuerdo cuando te dije que quería viajar a un congreso en Mendoza y me dijiste: ¿Cuándo salimos?.. Gracias por caminar siempre a mi lado, hoy estas dentro de mí y así va a ser por siempre... Te amo y tuve mucha suerte de haber tenido un padre tan maravilloso como vos, esta tesis está dedicada a vos.

A mis hermanos: Martín, Pao y Geor. Gracias por preocuparse por mi trabajo, por interesarse en mi proyecto, por preguntarme cuánto falta, por acompañarme en esta etapa. Fueron un refugio emocional en las buenas y en las malas...sigamos caminando juntos para siempre.

A María mi cuñada, por esos abrazos que llenan el alma. Porque junto con mi hermano me dieron la oportunidad de ser tía de un pequeño terremoto llamado Nahuel! Espero poder seguir respondiendo sus preguntas sobre la naturaleza (reconozco que algunas me dejan sin palabras! Jaja).

A mis tías y tíos, en especial a mi madrina Ana y mi tía Norma por preocuparse e interesarse en cómo iba todo.. A mi abuela Hortencia, por los llamados inesperados que me han alentado y ayudado.

A todos ustedes...¡¡¡ GRACIAS !!!

**“Hay una fuerza motriz más poderosa que el vapor, la electricidad y la energía atómica: la voluntad” (Albert Einstein)**

# Índice

|  |           |
|--|-----------|
| <b>1. Introducción</b> .....   | <b>1</b>  |
| 1.1. Bioprospección microbiana.....  | 1         |
| 1.2. Genómica y metagenómica como herramienta para la bioprospección .....               | 1         |
| 1.2.1. Genómica .....  | 1         |
| 1.2.2. Metagenómica .....  | 2         |
| 1.2.3. Genómica de células únicas .....  | 3         |
| 1.2.4. Bases de datos genómicos y metagenómicos .....                                    | 4         |
| 1.3. Las macroalgas y el rol de los microorganismos en su descomposición .....           | 4         |
| 1.3.1. Las macroalgas.....   | 4         |
| 1.3.2. Degradación de las macroalgas .....   | 5         |
| 1.4. Fucooidanos .....   | 6         |
| 1.4.1. Algas pardas .....  | 6         |
| 1.4.1.1. Rol de los fucooidanos en las algas pardas .....                                | 7         |
| 1.4.1.2. Estructura química.....   | 8         |
| 1.5. Enzimas activas sobre carbohidratos.....  | 8         |
| 1.6. Enzimas que depolimerizan a los fucooidanos.....                                    | 9         |
| 1.6.1. $\alpha$ -L-fucosidasas y fucooidanasas.....                                      | 9         |
| 1.6.2. Sulfatasas.....   | 10        |
| 1.6.3. Enzimas fucooidanasas caracterizadas hasta el momento.....                        | 11        |
| 1.7. Propiedades bioactivas de los fucooidanos .....                                     | 13        |
| 1.7.1. Propiedades de los fucooidanos de alto peso molecular .....                       | 14        |
| 1.7.2. Propiedades de los fucooidanos de bajo peso molecular (fuco-oligosacáridos) ..... | 14        |
| 1.7.2.1. Ventajas de la hidrólisis enzimática .....                                      | 15        |
| 1.8. Relevancia regional del estudio de enzimas fucooidanasas .....                      | 15        |
| 1.9. Objetivos .....   | 16        |
| 1.9.1. Objetivo general .....  | 16        |
| 1.9.2. Objetivos específicos.....  | 16        |
| <b>2. Materiales y Métodos</b> .....   | <b>17</b> |
| 2.1. Identificación de secuencias homólogas a fucooidanasas .....                        | 18        |
| 2.1.1. Búsqueda en base de datos de genomas bacterianos .....                            | 18        |

|   |           |
|---|-----------|
| 2.1.2. Búsqueda en base de datos de metagenomas .....   | 19        |
| 2.2. Análisis de las secuencias identificadas .....   | 19        |
| 2.2.1. Alineamientos múltiples de secuencias identificadas en genomas bacterianos .....   | 19        |
| 2.2.2. Reconstrucción filogenética de secuencias genómicas .....  | 20        |
| 2.2.3. Análisis de evidencia funcional de las secuencias metagenómicas.....   | 20        |
| 2.2.4. Análisis de similitud de las secuencias metagenómicas con las depositadas en GenBank, e identificación de motivos/dominios conservados.....                    | 20        |
| 2.3. Análisis del contexto genómico y búsqueda de sintenia compartida .....   | 21        |
| 2.3.1. Organización génica en genomas bacterianos.....  | 21        |
| 2.3.2. Sintenia compartida en genomas bacterianos.....  | 21        |
| 2.4. Análisis del potencial degradativo de carbohidratos .....  | 21        |
| <b>3. Resultados .....</b>  | <b>22</b> |
| 3.1. En genomas bacterianos .....   | 23        |
| 3.1.1. Identificación de secuencias homólogas a enzimas fucoidanasas.....   | 23        |
| 3.1.1.1 Análisis filogenético de las secuencias identificadas a partir de las enzimas pertenecientes a la familia GH107. ....   | 24        |
| 3.1.1.2. Análisis filogenético de las secuencias identificadas a partir de los dominios catalíticos de las enzimas de la cepa <i>Flavobacterium</i> sp. SA-0082 ..... | 28        |
| 3.1.2. Análisis del contexto genómico y sintenia compartida .....   | 30        |
| 3.1.3. Análisis del potencial degradativo de carbohidratos .....  | 35        |
| 3.2. En metagenomas marinos.....  | 36        |
| 3.2.1. Identificación de secuencias homólogas .....   | 36        |
| 3.2.2. Evidencias funcionales en las secuencias metagenómicas.....  | 36        |
| 3.2.3. Relación con secuencias genómicas y búsqueda de dominios conservados .....   | 42        |
| <b>4. Discusión.....</b>  | <b>46</b> |
| <b>5. Conclusiones.....</b>   | <b>57</b> |
| <b>6. Bibliografía.....</b>   | <b>60</b> |

## Lista de Tablas

---

|  |    |
|--|----|
| <b>Tabla 1.1.</b> Sitios web con datos genómicos y metagenómicos útiles para la bioprospección microbiana. Extraído y modificado de Lozada y Dionisi (12). .....   | 4  |
| <b>Tabla 1.2.</b> Organismos en los cuales se han reportado hasta el momento presencia de enzimas fucoidanasas. ....   | 12 |
| <b>Tabla 1.3.</b> Matriz de porcentajes de identidad entre las secuencias de enzimas fucoidanasas que se encuentran en los organismos especificados.....   | 13 |
| <b>Tabla 1.4.</b> Efectos sobre la salud de fucoidanos aislados de diferentes especies de algas pardas. Extraído y modificado de Wijesinghe y colaboradores (53). ....   | 14 |
| <b>Tabla 3.1.</b> Linaje de las bacterias marinas identificadas que podrían presentar genes que codifiquen para enzimas fucoidanasas .....   | 23 |
| <b>Tabla 3.2.</b> Secuencias homólogas a las enzimas fucoidanasas de referencia identificadas en la base de datos de genomas microbianos IMG. Se indican los organismos en los cuales se identificaron las secuencias, el ambiente en el que fue aislado (marino: proveniente de la columna de agua, sedimento: sedimentos marinos y sd: sin determinar), el código del gen, el nombre del producto que codifica y por último el número de acceso a la base de datos IMG. .. | 25 |
| <b>Tabla 3.3.</b> Número asignado a las secuencias de los contextos genómicos incluidos en las Figuras 3.3. y 3.4. En todos los casos se indica el nombre del producto génico asignado por la anotación funcional en el sistema IMG. Se resaltan en negrita los nombres de las secuencias relacionadas con el metabolismo de carbohidratos. ....   | 34 |
| <b>Tabla 3.4.</b> Familias y subfamilias GH (el número se indica entre paréntesis) reportadas para los organismos identificados en genomas microbianos. ....   | 35 |
| <b>Tabla 3.5.</b> Estudios metagenómicos de la base de datos IMG/M que contenían secuencias homólogas a fucoidanasas. Se indica el número de set de datos totales y el número de metagenomas en los cuales se identificaron secuencias homólogas. ....   | 37 |
| <b>Tabla 3.6.</b> Secuencias homólogas a enzimas fucoidanasas identificadas en estudios metagenómicos reportados en la base de datos IMG/M. Se indica el nombre, código de identificación y nombre de la muestra de la cual se identificaron secuencias homólogas, el dominio catalítico utilizado, la identificación del gen y el scaffold.....   | 38 |
| <b>Tabla 3.7.</b> Evidencias funcionales identificadas en las secuencias metagenómicas.....  | 42 |
| <b>Tabla 3.8.</b> Dominios conservados (DC) detectados en las secuencias metagenómicas a partir del algoritmo DELTA-BLAST. Si no se identificaron dominios se indica con la abreviación “sd” (sin determinar). A la derecha, se indican las secuencias que resultaron más similares a las secuencias metagenómicas.....  | 44 |

## Lista de Figuras

- Figura 1.1.** Esquema simplificado de la secuenciación de un genoma. En primer lugar, se extrae el ADN del organismo, el cual es luego fragmentado y secuenciado utilizando plataformas de alto rendimiento. Las secuencias obtenidas son ensambladas, obteniéndose una o varias secuencias que representan el genoma del organismo analizado. Extraído y modificado de Ng y Kirkness (10)..... 2
- Figura 1.2.** Estrategia metagenómica de secuenciación al azar. En primer lugar, se extrae ADN de la comunidad microbiana de una muestra ambiental. Debido al gran tamaño de los genomas, éstos son escindidos en fragmentos más pequeños para posteriormente ser secuenciados. Las lecturas resultantes son ensambladas y anotadas por medio de programas computacionales. Extraído de “Teach the Microbiome” (17)..... 3
- Figura 1.3.a–i** Macroalgas marinas que producen polisacáridos de importancia medicinal: (a) *Enteromorpha linza* (Ulvales, Chlorophyta); (b) *Capsosiphon fulvescens* (Ulotrichales, Chlorophyta); (c) *Sargassum muticum* (Fuciales, Phaeophyceae); (d) *Ecklonia cava* (Laminariales, Phaeophyceae); (e) *Undaria pinnatifida* (Laminariales, Phaeophyceae); (f) *Laminaria japonica* (Laminariales, Phaeophyceae); (g) *Carpopeltis affinis* (Cryptonemiales, Rhodophyta); (h) *Chondrus ocellatus* (Gigartinales, Rhodophyta); (i) *Scinaia japonica* (Nemaliales, Rhodophyta). Extraído de Syntsyra y colaboradores (21)..... 5
- Figura 1.4.** Representación esquemática de la degradación de materia orgánica por un microorganismo que produce enzimas extracelulares endo-líticas (enzima superior, escisión en región medial) y exo-lítica (enzima inferior, escisión en región terminal) dando lugar a compuestos de bajo peso molecular. La zona oscura en la superficie celular representa un canal para la incorporación del producto resultante de la hidrólisis enzimática. Extraído de Arnosti (24)..... 6
- Figura 1.5.** Modelo de la pared celular de algas pardas pertenecientes al orden Fuciales. Las microfibrillas son escasas y con forma de cinta. Los alginatos y fucoidanos son los polímeros que constituyen la mayor parte de la pared celular. Los fenoles suelen encontrarse asociados a alginatos y proteínas. Grandes cantidades de yoduro también se encuentran en la pared. Extraído de Deniaud-Bouët y colaboradores (31)..... 7
- Figura 1.6.** Estructura del fucoidano de las algas pardas *Chorda filum* (Laminariales) y *Ascophyllum nodosum* (Fuciales). Extraído de Ale y Meyer (37)..... 8
- Figura 1.7.** Patrones de catálisis de los residuos de  $\alpha$ -L-fucosa de un fucoidano tipo. La actividad exolítica es característica de las  $\alpha$ -L-fucosidasas, mientras que la endo-hidrólisis es catalizada por enzimas fucoidanasas. Extraído de Ale y Meyer (37)..... 10
- Figura 1.8.** Estructura modular de la enzima fucoidanasa (FcnA) de *Mariniflexile fucanivorans* SW5..... 11
- Figura 1.9.** Genes parálogos producto de una duplicación del gen ancestral en el mismo organismo..... 12

**Figura 1.10.** Potenciales aplicaciones de los fucoidanos extraídos del alga parda *Undaria pinnatifida*. ..... 16

**Figura 3.1.** Árbol filogenético no enraizado obtenido por el método estadístico del vecino más cercano y el método de sustitución de aminoácidos de Jones-Taylor-Thornton (JTT) para las secuencias de referencia pertenecientes a la familia GH107 (señaladas con un cuadrado verde) y las secuencias genómicas identificadas a partir de ellas. Los círculos en los nodos hacen referencia a los valores de los soportes de “bootstrap” para 1000 réplicas (● = valores entre 80 y 100 %, ○ = valores entre 50 y 80 %). Las longitudes de las ramas son proporcionales al número de variaciones de aminoácidos. Los números corresponden al número de acceso a IMG y el nombre del organismo se indica entre corchetes. .... 28

**Figura 3.2.** Árbol filogenético no enraizado de las secuencias de referencia reportadas en la patente (cuadrado violeta) y las identificadas en la base de datos IMG. El árbol fue obtenido por el método del vecino más cercano. Los círculos en los nodos hacen referencia a los valores de los soportes de “bootstrap” para 1000 réplicas (● = valores entre 80 y 100%, ○ = valores entre 50 y 80%). Los números corresponden al número de acceso a IMG y el nombre del organismo se muestra entre corchetes a excepción de las dos secuencias identificadas en un microorganismo perteneciente al orden Planctomycetales..... 29

**Figura 3.3.** Esquema de genes próximos a las secuencias homólogas identificadas (señaladas con rojo) a partir de las enzimas pertenecientes a la familia GH107 (señaladas con un cuadrado verde). Los números sobre las flechas corresponden al nombre del producto génico indicados en la Tabla 3.3. Los número en negrita corresponden a elementos involucrados en el metabolismo de carbohidratos. Las secuencias homólogas a las enzimas fucoidanasas de la familia GH107 se señalan con un recuadro negro. .... 32

**Figura 3.4.** Esquema de genes próximos a las secuencias homólogas identificadas (señaladas con rojo) a partir de las secuencias pertenecientes al organismo *Flavobacterium* sp. SA-0082 (señaladas con un cuadrado violeta). Los números sobre las flechas corresponden al nombre del producto génico indicados en la Tabla 3.3. Los número en negrita corresponden a elementos involucrados en el metabolismo de carbohidratos. Las secuencias homólogas a las enzimas fucoidanasas se señalan con un recuadro negro. .... 33

---

# **1. Introducción**

### ***1.1. Bioprospección microbiana***

Los microorganismos, el grupo de organismos más diverso y abundante de la Tierra, ocupan todos los nichos posibles. Su participación es clave en los ciclos biogeoquímicos, regulan la composición de la atmósfera y eliminan compuestos contaminantes (1). La gran diversidad de las comunidades microbianas y sus funciones ecológicas y metabólicas se están explorando en una gran variedad de ambientes naturales, entre ellos en suelos (2), aire (3) y océanos (4).

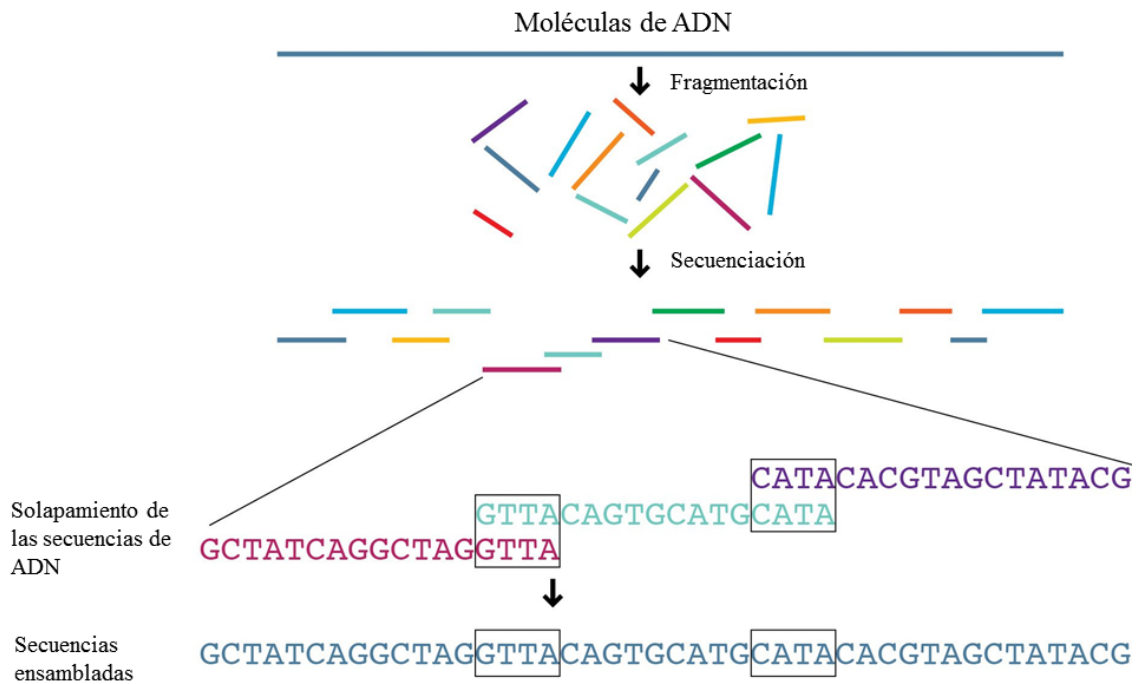
La bioprospección hace referencia a la búsqueda de compuestos químicos novedosos, genes, proteínas, microorganismos u otros productos que presenten potenciales aplicaciones comerciales, a partir de los ambientes naturales (5). En particular, los microorganismos marinos representan una reserva casi inagotable de diversidad genética y funcional mayormente inexplorada, acumulada durante millones de años de evolución de adaptación a diferentes presiones selectivas (6). A pesar de su papel vital en la sostenibilidad de la Tierra y sus aplicaciones biotecnológicas actuales y potenciales, gran parte de la diversidad y capacidades metabólicas de los microorganismos aún sigue sin explotar (7).

### ***1.2. Genómica y metagenómica como herramienta para la bioprospección***

#### **1.2.1. Genómica**

La genómica es el estudio del material genético completo de un organismo (genoma), e intenta comprender el contenido, la organización, la función y la evolución de la información genética contenida en el genoma. La genómica consiste en dos ramas complementarias: la estructural y la funcional. La genómica estructural determina la organización y la secuencia de la información genética contenidas dentro de un genoma. Por otro lado, la genómica funcional, caracteriza la función de las secuencias dilucidadas a partir de la genómica estructural (8).

En la actualidad, la secuencia completa de un genoma bacteriano se obtiene utilizando plataformas de secuenciación en gran escala. Este método comienza con la fragmentación al azar del ácido desoxirribonucleico (ADN), seguido por su secuenciación y su ensamblado por superposición mediante métodos computacionales (9). La secuencia del genoma es luego anotada a fin de obtener información sobre la ubicación de las secuencias codificantes, y su potencial función (Figura 1.1.).



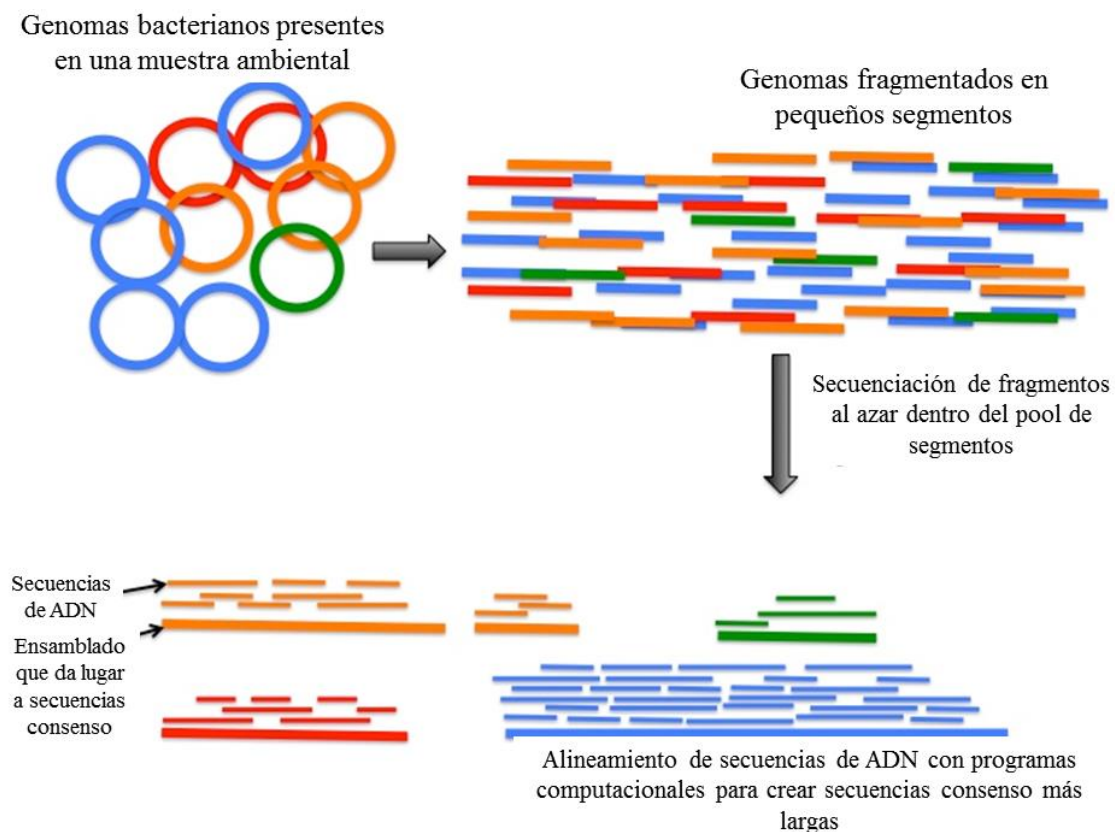
**Figura 1.1.** Esquema simplificado de la secuenciación de un genoma. En primer lugar, se extrae el ADN del organismo, el cual es luego fragmentado y secuenciado utilizando plataformas de alto rendimiento. Las secuencias obtenidas son ensambladas, obteniéndose una o varias secuencias que representan el genoma del organismo analizado. Extraído y modificado de Ng y Kirkness (10).

### 1.2.2. Metagenómica

Los microorganismos, como miembros de comunidades, son frecuentemente interdependientes metabólicamente, o bien utilizan recursos específicos que limitan su cultivabilidad. Los microorganismos marinos son particularmente resistentes al cultivo, dado que > 99,9 % de los mismos resultan imposibles de cultivar utilizando métodos tradicionales (11). Las estrategias metagenómicas permiten estudiar los genomas de (idealmente) todos los microorganismos presentes en una comunidad microbiana sin una etapa previa de cultivo, proporcionando información no solo sobre la estructura de la comunidad sino también de su potencial genético (12). Si bien la metagenómica presenta aún muchos desafíos, representa una herramienta clave para la bioprospección microbiana, ofreciendo la oportunidad de estudiar más ampliamente el potencial biotecnológico de las comunidades microbianas.

El análisis metagenómico comienza por lo general con la extracción y purificación del ADN de una muestra ambiental (Figura 1.2.). El material genético extraído es llamado “ADN metagenómico” (13). Este ADN puede utilizarse para la construcción de una biblioteca metagenómica (14) o, alternativamente, puede secuenciarse al azar a través de tecnologías de secuenciación en gran escala (15). A partir de la secuenciación al azar, se obtienen millones de

lecturas cuya longitud depende de la plataforma utilizada (16). Algunos ejemplos de plataformas son Illumina HiSeq, Roche454 y PacBio. A fin de recomponer fragmentos más largos de los genomas de los microorganismos que constituyen la comunidad, las lecturas obtenidas por secuenciación al azar son ensambladas utilizando herramientas bioinformáticas (13). El ensamblado se basa en el sobresolapamiento de las secuencias para reconstruir secciones completas y contiguas de ADN llamadas cóntigos (“contigs”), con la finalidad de obtener fragmentos más extensos. Los cóntigos se ensamblan nuevamente en fragmentos de mayor longitud, para formar secuencias consenso denominadas supercóntigos (“scaffolds”). Los supercóntigos contienen algunas regiones cuya secuencia no pudo ser determinada, o que presentan un bajo nivel de certeza. En el primer caso, cada base será indicada con una “N”, y en el segundo caso la base será indicada en minúscula.



**Figura 1.2.** Estrategia metagenómica de secuenciación al azar. En primer lugar, se extrae ADN de la comunidad microbiana de una muestra ambiental. Debido al gran tamaño de los genomas, éstos son escindidos en fragmentos más pequeños para posteriormente ser secuenciados. Las lecturas resultantes son ensambladas y anotadas por medio de programas computacionales. Extraído de “Teach the Microbiome” (17).

### 1.2.3. Genómica de células únicas

Además de la metagenómica, se ha desarrollado otra estrategia que permite determinar la información genética contenida en un genoma sin necesitar del cultivo del microorganismo, la

cual se denomina genómica de una única célula (SCG del inglés “single-cell genomics”). El análisis de células individuales implica el aislamiento de una célula individual a partir de una muestra ambiental, seguido de la purificación de su ADN, la amplificación al azar del mismo y su secuenciación (18). Esta estrategia presenta la ventaja de vincular el potencial metabólico de un microorganismo con su identidad taxonómica (19), lo cual es mucho más dificultoso en la metagenómica. Esta estrategia se utiliza a menudo de forma complementaria a la metagenómica.

#### 1.2.4. Bases de datos genómicos y metagenómicos

En la actualidad se cuenta con numerosos servidores públicos que permiten acceder a la información almacenada sobre datos genómicos y metagenómicos, como así también realizar análisis de secuencias a través de diversas herramientas y recursos bioinformáticos. En la Tabla 1.1. se especifican algunos de los sitios de Internet que pueden ser utilizados para la bioprospección de secuencias de interés biotecnológico.

**Tabla 1.1.** Sitios web con datos genómicos y metagenómicos útiles para la bioprospección microbiana. Extraído y modificado de Lozada y Dionisi (12).

| Sitio web   | Descripción   |
|---|---|
| Base de datos online de genomas (GOLD)<br><a href="http://www.genomesonline.org/cgi-bin/GOLD/index.cgi">www.genomesonline.org/cgi-bin/GOLD/index.cgi</a>  | Es un recurso que permite el acceso integral a la información sobre proyectos de secuenciación de genomas y metagenomas, y sus metadatos asociados. |
| Sistema integrado de genomas y metagenomas microbianos (IMG y IMG/M)<br><a href="http://img.jgi.doe.gov">http://img.jgi.doe.gov</a>   | Base de datos y herramientas de análisis de datos genómicos y metagenómicos, que permite integrar ambos tipos de información.                       |
| Recurso Genome del Centro Nacional de información biotecnológica (NCBI)<br><a href="http://www.ncbi.nlm.nih.gov">www.ncbi.nlm.nih.gov</a>   | El recurso "Genome" organiza información sobre genomas, incluyendo secuencias, mapas, cromosomas, ensamblajes y anotaciones.                        |
| Servicio de Análisis Genómicos (RAST, <a href="http://rast.nmpdr.org">http://rast.nmpdr.org</a> ) y Metagenómicos (MG-RAST; <a href="http://metagenomics.anl.gov">http://metagenomics.anl.gov</a> ) | Los servidores RAST y MG-RAST son plataformas de anotación automática y de análisis automatizado para genomas y metagenomas, respectivamente.       |

### 1.3. Las macroalgas y el rol de los microorganismos en su descomposición

#### 1.3.1. Las macroalgas

Las algas marinas son un grupo de organismos muy diversos que constituyen una fuente potencial de compuestos químicos con utilidad en biomedicina, como por ejemplo vitaminas o

ácidos grasos poliinsaturados (20). Entre los compuestos bioactivos presentes en las macroalgas, los polisacáridos son de especial interés debido a sus propiedades físicas y químicas (21). Debido al alto contenido de estos compuestos, varias especies de algas marinas son comúnmente utilizadas para fines alimenticios y medicinales (Figura 1.3.). Las macroalgas contienen una amplia variedad de polisacáridos, los cuales pueden ser estructurales formando parte de la pared celular, o bien de almacenamiento como reserva de energía (21).



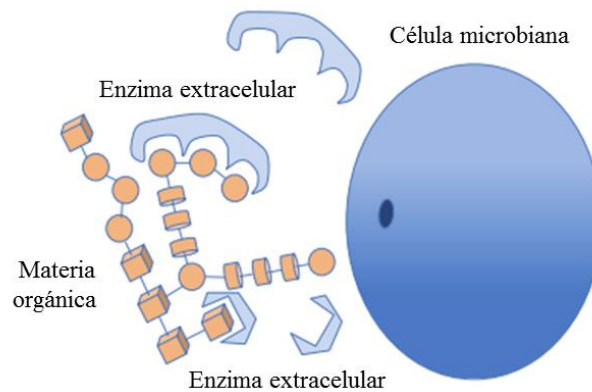
**Figura 1.3.** a-i Macroalgas marinas que producen polisacáridos de importancia medicinal: (a) *Enteromorpha linza* (Ulvales, Chlorophyta); (b) *Capsosiphon fulvescens* (Ulotrichales, Chlorophyta); (c) *Sargassum muticum* (Fuciales, Phaeophyceae); (d) *Ecklonia cava* (Laminariales, Phaeophyceae); (e) *Undaria pinnatifida* (Laminariales, Phaeophyceae); (f) *Laminaria japonica* (Laminariales, Phaeophyceae); (g) *Carpopeltis affinis* (Cryptonemiales, Rhodophyta); (h) *Chondrus ocellatus* (Gigartinales, Rhodophyta); (i) *Sciniaia japonica* (Nemaliales, Rhodophyta). Extraído de Synytsya y colaboradores (21).

### 1.3.2. Degradación de las macroalgas

Los polisacáridos de las macroalgas constituyen una fuente de carbono y energía importante para muchas bacterias marinas. La degradación microbiana de esta materia orgánica constituye un paso importante en el ciclo del carbono, dado que impide el secuestro a largo plazo del carbono en los ambientes costeros (22). Los microorganismos comienzan la degradación de los componentes de las paredes celulares de las algas marinas mediante la secreción de enzimas extracelulares específicas (por ejemplo glucósido-hidrolasas o polisacárido liasas, Figura 1.4.) capaces de metabolizar estos compuestos a tamaños lo suficientemente pequeños,

aproximadamente 700 Da (23), para que puedan ser absorbidos y utilizados para su crecimiento y metabolismo (24, 25). Estas enzimas presentan una alta especificidad con respecto a qué tipos de sustratos son capaces de depolimerizar (24). Los oligosacáridos resultantes de la actividad enzimática ingresan a la célula por medio de sistemas transportadores presentes tanto en la membrana interna como en la membrana externa, en el caso de las bacterias Gram negativas (24). El mecanismo por medio del cual los productos de degradación de bajo peso molecular se incorporan al organismo puede variar considerablemente entre ellos (26-28), así como también entre ambientes (29).

Debido a que estas enzimas depolimerizantes funcionan fuera de la célula, los productos de la hidrólisis pueden ser aprovechados por otros microorganismos que no necesariamente segregan sus propias enzimas (30). En general, la producción microbiana de enzimas extracelulares presenta mayores tasas de actividad en ambientes bentónicos en comparación con los pelágicos (30). Estas diferencias pueden interpretarse como resultado de una estrategia “costo-beneficio”, en donde el retorno de la inversión en la producción de las enzimas extracelulares es mayor en los sedimentos y por lo tanto, también es mayor el grado de producción de enzimas (24).



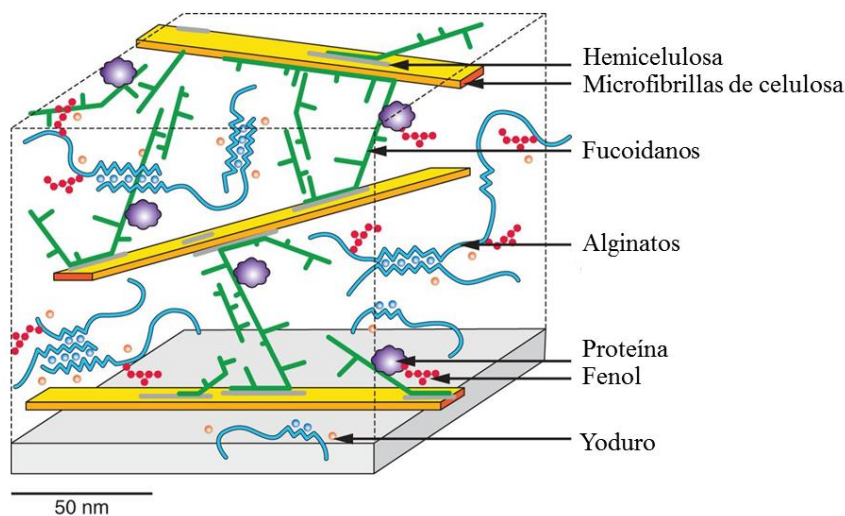
**Figura 1.4.** Representación esquemática de la degradación de materia orgánica por un microorganismo que produce enzimas extracelulares endo-líticas (enzima superior, escisión en región medial) y exo-lítica (enzima inferior, escisión en región terminal) dando lugar a compuestos de bajo peso molecular. La zona oscura en la superficie celular representa un canal para la incorporación del producto resultante de la hidrólisis enzimática. Extraído de Arnosti (24).

## 1.4. *Fuoidanos*

### 1.4.1. Algas pardas

Las algas pardas (Phaeophyceae) deben su nombre a su coloración predominante, determinada por los pigmentos clorofila a y c, beta-caroteno, violaxantina y fucoxantina. Presentan en la pared celular fibrillas de celulosa, y en la matriz dos tipos mayoritarios de polisacáridos: alginatos y fucooidanos (21). Deniaud-Bouët y colaboradores (31) determinaron la composición

de la pared celular de cinco especies de algas del orden de Fucales, por medio de fraccionamientos enzimáticos y análisis cromatográficos. Estos estudios revelaron que los fucoidanos y los alginatos abarcan la porción principal de la pared en las algas pardas (constituyendo hasta el 45 % de su peso seco), mientras que la celulosa sólo representa una pequeña fracción (1-8 % del peso seco). Una parte de la molécula de fucoidanos está estrechamente asociadas a microfibrillas de celulosa, sugiriendo que estos polisacáridos actúan como importantes componentes en la reticulación de las paredes celulares de algas pardas (Figura 1.5.).



**Figura 1.5.** Modelo de la pared celular de algas pardas pertenecientes al orden Fucales. Las microfibrillas son escasas y con forma de cinta. Los alginatos y fucoidanos son los polímeros que constituyen la mayor parte de la pared celular. Los fenoles suelen encontrarse asociados a alginatos y proteínas. Grandes cantidades de yoduro también se encuentran en la pared. Extraído de Deniaud-Bouët y colaboradores (31).

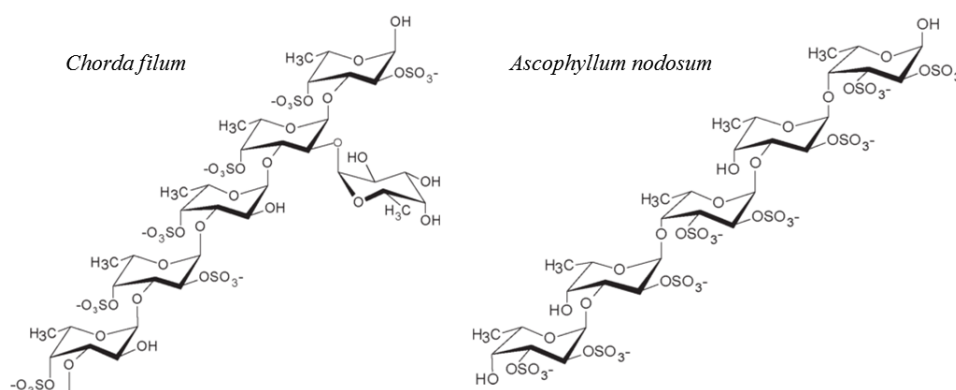
#### 1.4.1.1. Rol de los fucoidanos en las algas pardas

Aunque la función de los fucoidanos en las algas no se ha investigado a fondo, hay varias hipótesis sobre qué tipo de beneficio aportarían estos compuestos para el alga. Se ha reportado que el contenido de fucoidano en el alga difiere entre la zona intermareal (altas cantidades de fucoidanos) y la zona debajo de la línea de bajamar (menos cantidades de fucoidanos), por lo que podría tratarse de una estrategia frente a la deshidratación (32, 33). Otra hipótesis propuesta es que mejoraría la estabilidad de la pared celular (34), la cual podría ser apoyada por el descubrimiento de que el contenido de hidratos de carbono en las algas aumenta gradualmente en los meses con mayor exposición a la luz UV (35).

### 1.4.1.2. Estructura química

El término fucoidanos hace referencia a una serie de polisacáridos sulfatados ramificados ricos en fucosa (36). Estas moléculas constan de una cadena principal de residuos de  $\alpha$ -L-fucosa ( $\alpha$ -L-Fucp) unida por enlaces del tipo  $\alpha$ -(1 $\rightarrow$ 3),  $\alpha$ -(1 $\rightarrow$ 4) o ambos alternados. Los residuos de fucosa pueden presentar sustituciones sulfato y/o acetilo en el carbono 2 y/o 4, raramente en el 3 (37). Se han reportado también la presencia de otros monosacáridos tales como glucosa, galactosa, ácido úrico, xilosa y/o mannososa (38). A pesar de que los fucoidanos se conocen desde el año 1913, se han establecido pocas estructuras (37). La composición química y la estructura de los fucoidanos son complejas y variables, reflejando diferencias en la ruta biosintética entre las distintas especies de algas pardas.

Generalmente, en los órdenes Chordariales, Ectocarpales y Laminariales, los residuos de  $\alpha$ -L-fucosa presentan enlaces del tipo 1 $\rightarrow$ 3- $\alpha$ -L-Fucp, mientras que las algas pardas pertenecientes al orden Fucales se caracterizan porque el esqueleto de  $\alpha$ -L-Fucp contiene uniones alternadas del tipo 1 $\rightarrow$ 3 y 1 $\rightarrow$ 4 (39, 40). En la Figura 1.6. se ejemplifica la estructura del fucoidano del alga *Chorda filum* como representante del orden Laminariales y a *Ascophyllum nodosum* del orden Fucales. En el primer caso, los residuos de  $\alpha$ -L-Fucp presentan uniones del tipo 1 $\rightarrow$ 3 con sustituciones sulfato en el carbono de la posición 2 y 4 (C2 y C4 respectivamente) y ramificaciones de fucopiranososa en el C2. Por otro lado, la especie perteneciente al orden Fucales presenta un esqueleto de  $\alpha$ -L-Fucp con uniones alternas 1 $\rightarrow$ 3 y 1 $\rightarrow$ 4, con sulfataciones principalmente en el C2.



**Figura 1.6.** Estructura del fucoidano de las algas pardas *C. filum* (Laminariales) y *A. nodosum* (Fucales). Extraído de Ale y Meyer (37).

## 1.5. Enzimas activas sobre carbohidratos

La clasificación de enzimas que presentan como sustrato a los carbohidratos, basada en la similitud de sus secuencias, se encuentra recopilada en el servidor CAZy (Enzimas Activas

sobre Carbohidratos, [www.cazy.org](http://www.cazy.org)). La base de datos incluye un listado de módulos o dominios funcionales de enzimas que degradan, modifican o crean enlaces glucosídicos (41). Por lo tanto, este servidor incluye a un amplio espectro de enzimas involucradas en el metabolismo de carbohidratos, a menudo denominadas CAZymas. Dentro de los módulos que catalizan la ruptura, biosíntesis o modificación de carbohidratos se encuentran:

- Glucósido hidrolasas (GH): catalizan la hidrólisis de enlaces glucosídicos
- Glicosiltransferasas (GTs): catalizan la formación de enlaces glucosídicos
- Polisacárido Liasas (PL): producen la escisión no hidrolítica de enlaces glucosídicos
- Carbohidrato Esterasas (CES): producen hidrólisis de ésteres de hidratos de carbono
- Actividades Auxiliares (AAS): enzimas redox que actúan en conjunto con CAZymas

Las enzimas GH, también conocidas como glucosidasas, catalizan la hidrólisis del enlace glucosídico presente entre dos o más azúcares o entre un azúcar y otra parte de la molécula no glicosilada. Las GH están clasificadas con el código enzimático (EC) 3.2.1. Representan la clase más numerosa, compuesta por un total de 133 familias (familias GH). Cada familia contiene péptidos que son similares en cuanto a la secuencia de aminoácidos (estructura primaria) y en su estructura secundaria. Las familias GH5, GH13 y GH30 han sido clasificadas en subfamilias, que corresponden a subgrupos dentro de una familia que comparten un ancestro común más reciente y que, por lo general, resultan más uniformes en sus funciones moleculares (41). Algunas GH son enzimas multifuncionales debido a la presencia de uno o más módulos, que pueden pertenecer a varias familias. La clasificación de estas enzimas continúa evolucionando, ya que existen más de 4000 secuencias de GH que aún no han sido clasificadas en familias. Las enzimas PL, en cambio, catalizan la depolimerización de los polisacáridos que contienen ácidos urónicos por medio de un mecanismo de  $\beta$ -eliminación generando un residuo de ácido hexenurónico insaturado y un nuevo extremo reductor en el punto de clivaje (42). Esta clase de CAZymas es menos diversa, contando actualmente con sólo 23 familias.

### ***1.6. Enzimas que depolimerizan a los fucoidanos***

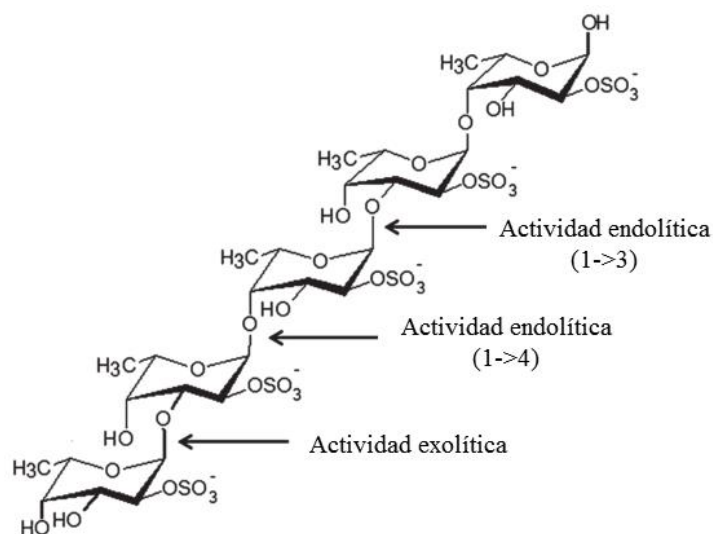
Si bien aún no se ha estudiado en detalle el proceso completo de depolimerización de las complejas y variables estructuras de los fucoidanos, se conocen algunas enzimas que catalizan la ruptura de estos compuestos, las cuales se describen a continuación.

#### **1.6.1. $\alpha$ -L-fucosidasas y fucoidanasas**

Las enzimas que catalizan la hidrólisis de enlaces entre residuos  $\alpha$ -L-Fucp pertenecen a la clase GH, las cuales se dividen en dos grandes grupos de acuerdo a su modo de acción: exo-

hidrolíticas y endo-hidrolíticas. Las enzimas con actividad exo-hidrolítica son capaces de escindir enlaces de residuos de fucosa no sulfatada o sulfatada en el extremo no reducido del fucoidano (36). Estas enzimas son llamadas  $\alpha$ -L-fucosidasas (EC: 3.2.1.51). En cambio, las enzimas que escinden enlaces glucosídicos internos de la molécula de fucoidano, dando lugar a fuco-oligosacáridos de bajo peso molecular, se denominan fucoidanasas (EC: 3.2.1.44) (37) (Figura 1.7.). Estas enzimas son muy diversas, y presentan una alta especificidad con respecto al tipo de enlace ( $1\rightarrow3$  ó  $1\rightarrow4$ ) y la presencia de grupos sulfato (43). Las enzimas fucoidanasas presentan un gran potencial biotecnológico, ya que su actividad endolítica permite obtener compuestos de bajo peso molecular, los cuales son biológicamente activos (36), como se describirá más adelante. Debido a estas propiedades, los fuco-oligosacáridos presentan potenciales aplicaciones en la industria farmacéutica, nutracéutica, cosmética y alimenticia.

Además de las enzimas que catalizan la endo-hidrólisis conocidas como fucoidanasas, también se han descrito endo-liasas que presentan como sustrato a los fucoidanos (43, 44). Sin embargo, éstas han sido descritas únicamente en patentes, y por lo tanto esta investigación no ha sido sometida al correspondiente proceso de revisión por pares.



**Figura 1.7.** Patrones de catálisis de los residuos de  $\alpha$ -L-fucosa de un fucoidano tipo. La actividad exolítica es característica de las  $\alpha$ -L-fucosidasas, mientras que la endo-hidrólisis es catalizada por enzimas fucoidanasas. Extraído de Ale y Meyer (37).

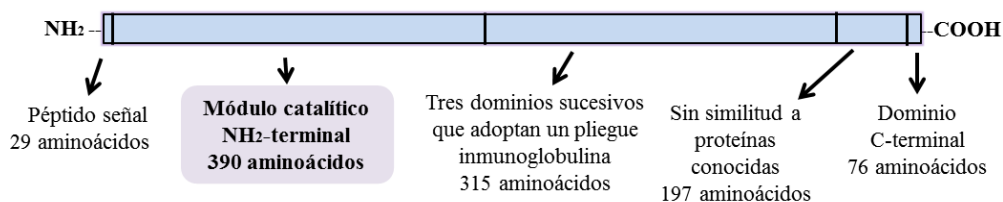
### 1.6.2. Sulfatasas

Otras enzimas que catalizan la escisión parcial de los fucoidanos de algas pardas son las sulfatasas bacterianas (EC 3.1.6.-). Estas enzimas eliminan grupos sulfato a partir de moléculas sulfatadas (39). La desulfatación se ha propuesto como potencial herramienta para la

investigación de las características estructurales y biológicas de los fucoidanos. Sin embargo, los reportes sobre enzimas sulfatasas específicas para estos compuestos son escasos (43).

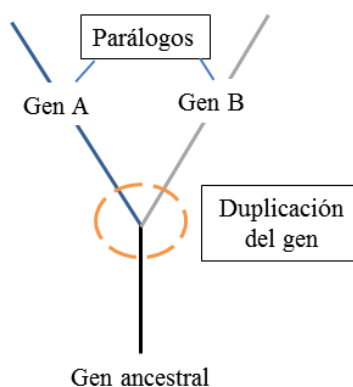
### 1.6.3. Enzimas fucoidanasas caracterizadas hasta el momento

Hasta el momento, sólo una secuencia que codifica para una fucoidanasa ha sido caracterizada en detalle, FncA (“sulfated fucan endo- $\alpha$ -1,4-L-fucanase”) de *Mariniflexile fucanivorans* SW5 (45). Esta enzima libera tetrasacáridos y hexasacáridos a partir de los fucoidanos del alga parda *Pelvetia canaliculata*. Análisis bioinformáticos realizados por Colin y colaboradores (45) determinaron la presencia de 5 dominios en la secuencia de aminoácidos. Un dominio ubicado cerca del extremo amino-terminal, de ~ 400 residuos, es responsable de su actividad catalítica (Figura 1.8.).



**Figura 1.8.** Estructura modular de la enzima fucoidanasa (FncA) de *Mariniflexile fucanivorans* SW5.

La enzima FncA definió, junto a otras tres secuencias menos caracterizadas, a la familia CAZy GH107 (41). Las tres fucoidanasas restantes se denominan SVI\_0379, Fda1 y Fda2. La primera fue identificada en la cepa *Shewanella violacea* DSS12 (Tabla 1.2.), mientras que las enzimas Fda1 y Fda2 fueron aisladas del organismo *Altermonas* sp. SN-1009. Dado sus altos porcentajes de identidad y por pertenecer al mismo organismo, las secuencias que codifican para las enzimas Fda1 y Fda2 son probablemente el producto de una duplicación génica. Los genes homólogos en el mismo organismo (que se originan por duplicación de un gen individual en el pasado evolutivo) se denominan parálogos (Figura 1.9.).



**Figura 1.9.** Genes parálogos producto de una duplicación del gen ancestral en el mismo organismo.

De forma similar, dos secuencias que codifican para enzimas fucoidanasas han sido identificadas en el organismo *Flavobacterium* sp. SA-0082 y reportadas en la patente EP 1826272/A2 (46). Estas enzimas no fueron nombradas y no están incluidas dentro de la familia GH107. En la base de datos GenBank del NCBI se reportan solamente como secuencias nucleotídicas del género *Flavobacterium* (Tabla 1.2.). Estas secuencias probablemente (al igual que para Fda1 y Fda2) se originaron a partir de una duplicación del gen ancestral correspondiendo a enzimas parálogas.

**Tabla 1.2.** Organismos en los cuales se han reportado hasta el momento presencia de enzimas fucoidanasas.

| GenBank <sup>(1)</sup> | Enzima   | Longitud | Organismo                             | Linaje   |
|------------------------|--|----------|---------------------------------------|--|
| DL231623               | sd   | 2091 nt  | <i>Flavobacterium</i> sp. SA-0082     | Bacteria; Bacteroidetes,<br>Flavobacteriia;<br>Flavobacteriales;<br>Flavobacteriaceae      |
| DL231622               | sd   | 2112 nt  | <i>Flavobacterium</i> sp. SA-0082     |  |
| CAI47003               | Sulfated fucan<br>endo- $\alpha$ -1,4-L-fucanase<br>(FcnA) | 1007 aá  | <i>Mariniflexile fucanivorans</i> SW5 |  |
| AAO00508               | Fucanase (Fda1)  | 814 aá   | <i>Alteromonas</i> sp. SN-1009        | Bacteria;<br>Proteobacteria;<br>Gammaproteobacteria;                                       |
| AAO00509               | Fucanase (Fda2)  | 881 aá   | <i>Alteromonas</i> sp. SN-1009        | Ateromonadales;<br>Alteromonadaceae  |
| BAJ00350               | SVI_0379   | 574 aá   | <i>Shewanella violacea</i> DSS12      | Bacteria;<br>Proteobacteria;<br>Gammaproteobacteria;<br>Alteromonadales;<br>Shewanellaceae |

(1) Numero de acceso a base de datos GenBank del NCBI (<http://www.ncbi.nlm.nih.gov/genbank/>). Las 2 primeras secuencias corresponden a secuencias de nucleótidos (nt), las 4 restantes a secuencias de aminoácidos (aá). sd, nombre de la enzima sin determinar.

Análisis comparativos entre las secuencias de enzimas fucoidanasas reportadas hasta el momento, mostraron que existe un alto grado de divergencia entre ellas (47). Se observaron porcentajes de identidad mayores al 30 % solamente entre los genes parálogos de la misma cepa (Tabla 1.3.). Dado que las enzimas identificadas en la patente no fueron nombradas, se asignó

con fines prácticos, “1” a aquella secuencia con el código de acceso a GenBank DL231623 y “2” a la que presenta el código DL231622 (Tabla 1.3.).

Por otro lado, resultados de análisis comparativos de dominios/motivos conservados entre las secuencias de enzimas fucoidanasas utilizando el recurso InterPro (48) indicaron la ausencia de un motivo común, a excepción de aquellas enzimas identificadas en el mismo microorganismo (47).

**Tabla 1.3.** Matriz de porcentajes de identidad entre las secuencias de enzimas fucoidanasas que se encuentran en los organismos especificados.

| Enzima (organismo)                            | % de identidad |           |           |           |          |      |
|---|----------------|-----------|-----------|-----------|----------|------|
|   | Fda1           | Fda2      | 1         | 2         | SVI_0379 | FcnA |
| Fda1 ( <i>Alteromonas</i> sp. SN-1009)        | 100            | <b>71</b> | 10        | 12        | 29       | 13   |
| Fda2 ( <i>Alteromonas</i> sp. SN-1009)        | <b>71</b>      | 100       | 10        | 12        | 28       | 13   |
| 1 - ( <i>Flavobacterium</i> sp. SA-0082)      | 10             | 10        | 100       | <b>60</b> | 7        | 11   |
| 2 - ( <i>Flavobacterium</i> sp SA-0082)       | 12             | 12        | <b>60</b> | 100       | 7        | 7    |
| SVI_0379 ( <i>Shewanella violacea</i> DSS12)  | 29             | 28        | 7         | 7         | 100      | 14   |
| FcnA ( <i>Mariniflexile fucanivorans</i> SW5) | 13             | 13        | 11        | 7         | 14       | 100  |

### 1.7. Propiedades bioactivas de los fucoidanos

Las propiedades terapéuticas de los fucoidanos han sido extensamente estudiadas. Dentro de las actividades biológicas conocidas, se incluyen actividad anticoagulante/antitrombótico, antiinflamatoria, inmunomoduladora, antiviral, antitumoral y antioxidante (21). Varios trabajos reportan el potencial antiviral de los fucoidanos extraídos de *U. pinnatifida* contra diferentes virus de la familia herpesviridae (HSV-1, HSV-2 y HCMV) (49), como así también sus propiedades antitumorales, antioxidantes y anticoagulantes (50, 51). Estas propiedades podrían ser aprovechadas para el desarrollo de fármacos o alimentos funcionales. Se ha reportado que los dos principales factores determinantes de las bioactividades son el peso molecular (52-54) y la posición y número de grupos éster-sulfato. Sin embargo, otras propiedades como la composición de monosacáridos y la tecnología de extracción, son de importancia para mantener la integridad estructural de los polisacáridos y para la conservación de sus propiedades biológicas (37, 55). A su vez, el tipo de actividad biológica que presenta el polisacárido sulfatado depende de la fuente de macroalga utilizada, confirmando claramente la biodiversidad de los fucoidanos en los diferentes órdenes taxonómicos de algas pardas (Tabla 1.4.). Fucoidanos extraídos de algas del mismo orden presentan diferente composición y por lo tanto, distintas bioactividades.

### 1.7.1. Propiedades de los fucoidanos de alto peso molecular

Se ha reportado que los polisacáridos sulfatados extraídos a partir de las algas pardas han dado resultados prometedores contra patógenos gastrointestinales en mamíferos mostrando un efecto anti-adhesivo (56). Estos compuestos son poco hidrolizables por las enzimas gástricas de los mamíferos. Sin embargo, son potencialmente fermentables por la microbiota intestinal autóctona, quienes lo utilizan como fuente de carbono.

### 1.7.2. Propiedades de los fucoidanos de bajo peso molecular (fuco-oligosacáridos)

El alto peso molecular de los fucoidanos es uno de los motivos por los cuales se han obstaculizado sus potenciales aplicaciones, especialmente como agentes terapéuticos (57). Es por ello, que la obtención de oligosacáridos de bajo peso molecular (fuco-oligosacáridos) es de especial interés por presentar una actividad biológica más específica. Se ha demostrado que los fucoidanos de bajo peso molecular resultaron ser antagónicos a una amplia gama de patógenos, inhibiendo la adhesión a las células del organismo (55). Resultados de análisis comparativos de fucoidanos de variado peso molecular indicaron que cuanto más pequeño es el peso que presentan, mayor es la actividad anticoagulante (58). Otras propiedades bioactivas que aumentan a medida que el peso molecular disminuye son: como agentes anticancerígenos (54) y como agentes antioxidantes (59, 60).

**Tabla 1.4.** Efectos sobre la salud de fucoidanos aislados de diferentes especies de algas pardas. Extraído y modificado de Wijesinghe y colaboradores (53).

| Efecto en la salud               | Fuente  | Referencia |
|----------------------------------|---|------------|
| Anticoagulante<br>Antitrombótico | <i>Ecklonia cava</i> (Laminariales)                 | (61)       |
|                                  | <i>Fucus evanescens</i> (Fucales)                   | (62)       |
|                                  | <i>Padina gymnospora</i> (Dyctiotales)              | (63)       |
|                                  | <i>Ascophyllum nodosum</i> (Fucales)                | (64)       |
|                                  | <i>Sargassum fulvellum</i> (Fucales)                | (65)       |
|                                  | <i>Laminaria cichorioides</i> (Laminariales)        | (66)       |
| Inmunomoduladora                 | <i>Fucus vesiculosus</i> (Fucales)                  | (51)       |
|                                  | <i>Undaria pinnatifida</i> (Laminariales)           | (67)       |
| Antiinflamatoria                 | <i>Laminaria japonica</i> (Laminariales)            | (68)       |
|                                  | <i>Ecklonia cava</i>                                | (69)       |
| Antitumoral                      | <i>Undaria pinnatifida</i>                          | (70)       |
|                                  | <i>Ecklonia cava</i>                                | (61)       |
|                                  | <i>Fucus evanescens</i>                             | (71)       |
| Antiviral                        | <i>Undaria pinnatifida</i>                          | (49)       |
| Protección de la mucosa gástrica | <i>Cladosiphon okamuranus</i> TOKIDA (Chordariales) | (72)       |

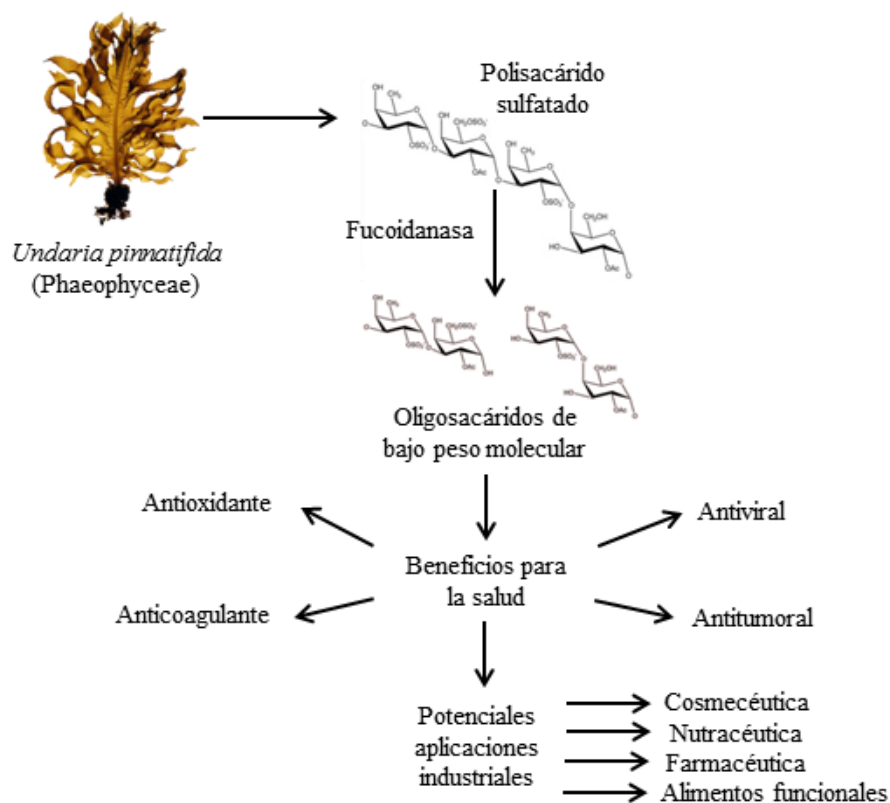
### 1.7.2.1. Ventajas de la hidrólisis enzimática

Los tratamientos convencionales para la degradación de fucoïdanos, ya sean químicos o térmicos, pueden producir pérdida de la integridad del polisacárido (incluyéndose los grupos sulfato) y por lo tanto de sus propiedades bioactivas (40). Es por ello que el tratamiento enzimático utilizando fucoïdanasas se considera la mejor opción para obtener oligosacáridos en donde se conserven los grupos sulfato y en consecuencia presenten una mayor actividad biológica (73).

## 1.8. Relevancia regional del estudio de enzimas fucoïdanasas

El alga parda *U. pinnatifida* (Laminariales, Phaeophyceae) es una especie nativa del noroeste de Asia (China, Japón y Corea del Sur), que fue detectada por primera vez en Patagonia en 1992 en Puerto Madryn, luego de ser transportada por vectores antrópicos (74). La invasión por *U. pinnatifida* ha generado efectos negativos sobre el ecosistema costero Patagónico, como por ejemplo la modificación de la estructura de las comunidades de algas locales, y en consecuencia de la fauna bentónica asociada (75). Produce además graves daños económicos, al afectar a embarcaciones, emprendimientos de acuicultura, actividades turísticas como el buceo o el uso recreacional de las playas. A partir del punto inicial de introducción, se ha extendido hacia el norte y hacia el sur, cubriendo actualmente las costas de la Patagonia desde 41,69° S hasta 47,75° S (76). Más aún, la invasión de *U. pinnatifida* puede potencialmente extenderse hacia el norte, pudiendo colonizar otras zonas costeras de Argentina, además de algunos puntos en la costa de Uruguay y Brasil (76). De hecho, recientemente se ha reportado un nuevo punto de introducción en la costa de Mar del Plata (77).

Considerando su amplia distribución actual, su alta fertilidad y la persistencia de sus estructuras reproductivas, la erradicación de *U. pinnatifida* de nuestras costas se considera imposible (76). Se puede reducir su impacto en áreas protegidas, emprendimientos de acuicultura o zonas de buceo por medio de su extracción manual (76). Como parte de una estrategia de control del impacto de la invasión, y a fin de generar nuevas oportunidades económicas en la región, sería deseable desarrollar aplicaciones comerciales a partir de la biomasa de esta especie. Una potencial aplicación es la producción de fuco-oligosacáridos de alta pureza, para el desarrollo de fármacos o suplementos alimentarios (Figura 1.10.). Para ello, resulta necesario disponer de enzimas capaces de depolimerizar estos polisacáridos.



**Figura 1.10.** Potenciales aplicaciones de los fucoidanos extraídos del alga parda *Undaria pinnatifida*.

## 1.9. Objetivos

### 1.9.1. Objetivo general

El objetivo general de esta tesis de grado consiste en incrementar nuestro conocimiento sobre los mecanismos que utilizan los microorganismos marinos para asimilar los polisacáridos de las algas pardas, con el fin de utilizar dicho conocimiento para el desarrollo de productos y procesos biotecnológicos.

### 1.9.2. Objetivos específicos

Los objetivos específicos definidos para este trabajo de tesis de grado son:

- Identificar secuencias homólogas a aquellas reportadas para enzimas fucoidanasas, a partir de genomas bacterianos y metagenomas depositados en bases de datos públicas.
- Analizar las relaciones evolutivas entre las secuencias de aminoácidos identificadas y las secuencias de fucoidanasas.
- Analizar el contexto genómico y su organización génica, para las secuencias más semejantes a las enzimas fucoidanasas.

---

## **2. Materiales y Métodos**

## **2.1. Identificación de secuencias homólogas a fucoidanasas**

A los efectos de facilitar la identificación de secuencias homólogas a enzimas fucoidanasas minimizando el riesgo de incluir secuencias que codifiquen otras actividades enzimáticas, se utilizó una estrategia de búsqueda basada en similitudes con el dominio esencial para la actividad catalítica de las 6 secuencias de fucoidanasas reportadas hasta el momento.

### **2.1.1. Búsqueda en base de datos de genomas bacterianos**

Se identificaron secuencias homólogas a las enzimas fucoidanasas de referencia a partir de genomas bacterianos depositados en la base de datos IMG (dependiente del Joint Genome Institute, Departamento de Energía, EEUU, <https://img.jgi.doe.gov>). Se analizaron todos los genomas bacterianos reportados en la base de datos IMG, los cuales se encuentran agrupados en 4 categorías de acuerdo al tipo de ambiente del que fue aislado el organismo: aire, acuático, terrestre y sin clasificar. Se utilizaron como secuencias problema (“query”) el módulo catalítico amino-terminal (aproximadamente de 400 aminoácidos) de las seis secuencias aminoacídicas indicadas en la Tabla 1.2.

La identificación de las secuencias se realizó utilizando la herramienta de búsqueda de genes del IMG de la pestaña “Find Genes”, basada en el algoritmo BLAST (78). El algoritmo BLAST (“Basic Local Alignment Search Tool”) es una herramienta de búsqueda básica de alineamiento local entre secuencias, la cual permite identificar secuencias similares a la utilizada como problema. Una de las variantes es el programa blastp, utilizado en este análisis. En esta búsqueda, se compararon las seis secuencias de aminoácidos problema con todas las secuencias en aminoácidos de los genomas depositados en el sistema IMG (79).

El punto de corte de significancia estadística (Expected value o E value) utilizado fue de  $1e^{-10}$ . La significancia estadística indica la probabilidad de que la similitud que se encuentre entre la secuencia problema y la identificada sea debida al azar. La matriz de intercambio de aminoácidos utilizada fue BLOSUM62 (“BLOcks of Amino Acid SUBstitution Matrix”). Se utilizaron los valores preestablecidos de los restantes parámetros [penalizaciones por desfase (gap penalties): presencia: 11, extensión: 1; umbral del alineamiento inicial (neighboring words threshold): 11]. Se seleccionaron aquellas secuencias que presentaron porcentajes de identidad mayor o igual al 30 %. Todas las secuencias seleccionadas fueron exportadas en formato FASTA para ser posteriormente analizadas.

### 2.1.2. Búsqueda en base de datos de metagenomas

La identificación de las secuencias se realizó utilizando las herramientas de búsqueda del Sistema Integrado de Manejo de Secuencias Metagenómicas (IMG/M), dependiente del Joint Genome Institute (Departamento de Energía, EEUU, <https://img.jgi.doe.gov>). Para analizar los metagenomas de interés de ambientes marinos, en primer lugar estos fueron identificados utilizando la herramienta “Genome Search” de la pestaña “Find Genome”. Posteriormente, se realizó una búsqueda por categorías de metadatos filtrado por tipo de ecosistema, utilizándose como palabra clave “marine”. Los metagenomas identificados fueron seleccionados por medio de la opción “Add Selected to Genome Cart”.

La búsqueda fue realizada sobre los metagenomas ensamblados seleccionados, utilizando como secuencias de entrada los dominios catalíticos de las secuencias de referencia, como se indicó en el punto anterior. Se utilizó la herramienta de búsqueda blastp de la pestaña “Find Genes”. El valor de significancia estadística establecido fue de  $1e^{-10}$  y se utilizaron demás valores preestablecidos [Matriz: Blosum 62; Penalizaciones por desfase (gap penalties): presencia: 11, extensión: 1; umbral del alineamiento inicial (“neighboring words threshold”): 11].

Las secuencias obtenidas con una longitud  $\geq 100$  aminoácidos y porcentajes de identidad  $\geq 30$  % fueron seleccionadas para su posterior análisis.

## 2.2. Análisis de las secuencias identificadas

### 2.2.1. Alineamientos múltiples de secuencias identificadas en genomas bacterianos

Las secuencias identificadas a partir de genomas bacterianos se alinearon con las secuencias de referencia. Dado que se desconoce si las secuencias pertenecientes a la familia GH107 son homólogas a las secuencias de fucoidasas no incluidas dentro de esta familia, las secuencias identificadas a partir del dominio catalítico de las enzimas pertenecientes a la familia GH107 fueron analizadas de forma independiente a aquellas identificadas con el dominio catalítico de las dos secuencias del género *Flavobacterium*.

El alineamiento se realizó en el programa Clustal X (80) con los valores establecidos por defecto. Clustal es un software que realiza alineamientos múltiples de proteínas y secuencias nucleotídicas a partir de secuencias que se encuentran en archivos multiFASTA. Se basa en la utilización de un algoritmo progresivo que agrupa secuencias y realiza alineamientos por pares para generar matrices de distancia, en los que se puede determinar el porcentaje de identidad entre cada par de secuencias (80).

### **2.2.2. Reconstrucción filogenética de secuencias genómicas**

El alineamiento múltiple fue exportado al programa MEGA6 (81), en el que se realizó un análisis de reconstrucción filogenética utilizando el algoritmo estadístico del vecino más cercano (“Neighbor-joining”) y el modelo de sustitución de aminoácidos de Jones-Taylor-Thornton (JTT). Las longitudes de rama fueron calculadas utilizando el método de corrección de Poisson (82). Los datos faltantes fueron tratados por el método de eliminación por pareja, y se realizaron pruebas para evaluar la robustez de la filogenia por medio del método de “bootstrap” con 1000 repeticiones. Dado el escaso conocimiento actual sobre las enzimas fucoidanasas, no fue posible utilizar grupos externos en los análisis.

### **2.2.3. Análisis de evidencia funcional de las secuencias metagenómicas**

Este análisis se basó en la información obtenida durante la anotación funcional de las secuencias sobre coincidencias con las bases de datos COG (Agrupamientos de Bloques Ortólogos) y Pfam (Familias de Proteínas). Dicha información proporciona evidencia sobre la posible función de las secuencias. Estos análisis son además utilizados por el sistema IMG como criterio para la asignación de nombres de las secuencias. Si la secuencia no sobrepasa un determinado umbral de longitud de alineamiento y de porcentaje de similitud en el análisis utilizando estas herramientas, la misma es anotada como proteína hipotética (83).

### **2.2.4. Análisis de similitud de las secuencias metagenómicas con las depositadas en GenBank, e identificación de motivos/dominios conservados**

A los efectos de evaluar el grado de similitud entre las secuencias identificadas en los metagenomas marinos con respecto a aquellas depositadas en la base de datos GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), se utilizó la herramienta de búsqueda blastp. Solo fueron consideradas aquellas secuencias cuyos alineamientos presentaron el mayor valor de porcentaje de identidad y cobertura con respecto a la secuencia metagenómica de entrada.

La identificación de dominios o motivos conservados se realizó utilizando la herramienta de búsqueda blastp del GenBank. El algoritmo aplicado en la búsqueda fue DELTA-BLAST (“Domain Enhanced Lookup Time Accelerated BLAST”). Este algoritmo realiza búsquedas de dominios conservados dentro de las secuencias aminoacídicas en un subconjunto de base de datos del NCBI (79).

### ***2.3. Análisis del contexto genómico y búsqueda de sintenia compartida***

#### **2.3.1. Organización génica en genomas bacterianos**

Este análisis se basó en la observación de que genes relacionados funcionalmente se encuentran a menudo presentes en la misma región del genoma (84), por lo que podría proveer evidencias adicionales sobre la posible función de las secuencias identificadas. De la inferencia filogenética obtenida, se seleccionaron aquellas secuencias que presentaron altos porcentajes de similitud con las secuencias de referencia, a fin de analizar sus organizaciones génicas y evaluar sus similitudes. El análisis se basó en la función predicha en la anotación funcional en el sistema IMG para los genes ubicados en cercanía de las secuencias identificadas.

#### **2.3.2. Sintenia compartida en genomas bacterianos**

Se realizó un análisis comparativo del orden de los genes ubicados en cercanía de las secuencias homólogas a fucoidanasas identificadas, a los efectos de evaluar si existía sintenia compartida entre los fragmentos genómicos que contenían dichas secuencias. Este análisis se llevó a cabo sólo con aquellas secuencias que presentaron los mayores porcentajes de identidad con respecto a las seis secuencias de referencia, dado la gran incertidumbre para definir la función de una secuencia en base a su similitud con secuencias de proteínas conocidas.

### ***2.4. Análisis del potencial degradativo de carbohidratos***

Los microorganismos capaces de utilizar polisacáridos como fuentes de carbono generalmente producen múltiples enzimas extracelulares con distintos rangos de sustratos. A fin de conocer el potencial general para degradar carbohidratos de las cepas bacterianas que contienen secuencias homólogas a fucoidanasas, se evaluó el rango completo de enzimas GH (familias y subfamilias) codificadas en sus genomas. Para realizar esta búsqueda se utilizó la base de datos CAZy (<http://www.cazy.org/>). En el buscador inicial se filtró por nombre de la cepa seleccionándose aquellos identificados por homología en la base de datos IMG.



## **3. Resultados**

### 3.1. En genomas bacterianos

#### 3.1.1. Identificación de secuencias homólogas a enzimas fucoidanasas

En el momento de realizar la búsqueda de secuencias homólogas a fucoidanasas, existían 33.804 genomas bacterianos depositados en la base de datos IMG. Para la búsqueda de secuencias putativas de fucoidanasas, como secuencias problema se utilizaron los potenciales dominios catalíticos de secuencias de enzimas fucoidanasas reportadas hasta el momento (Tabla 1.2.). La búsqueda utilizando el programa blastp permitió la identificación de 50 secuencias que se ajustaban a las condiciones preestablecidas, a partir de cinco de los seis dominios catalíticos evaluados. Solamente el dominio catalítico de la enzima fucoidanasa (FcnA) de *M. fucanivorans* SW5 no permitió la identificación de secuencias. Las secuencias putativas de fucoidanasas se identificaron en cepas bacterianas pertenecientes a cuatro phyla: Bacteroidetes, Proteobacteria, Planctomycetes y Verrucomicrobia, que incluían un total de 15 géneros (Tabla 3.1.). Por lo tanto, sólo una pequeña proporción de los genomas de la base de datos IMG contenía secuencias que cumplían con los criterios de búsqueda elegidos, y estos genomas contenían de 1 a 4 secuencias.

**Tabla 3.1.** Linaje de las bacterias marinas identificadas que podrían presentar genes que codifiquen para enzimas fucoidanasas

| PHYLUM          | CLASE               | ORDEN               | FAMILIA             | GÉNERO                  |
|-----------------|---------------------|---------------------|---------------------|-------------------------|
| Bacteroidetes   | Flavobacteriia      | Flavobacteriales    | Flavobacteriaceae   | <i>Aquimarina</i>       |
|                 |                     |                     |                     | <i>Cellulophaga</i>     |
|                 | Cytophagia          | Cytophagales        | Cyclobacteriaceae   | <i>Echinicola</i>       |
|                 |                     |                     |                     | Cytophagaceae           |
| Sphingobacteria | Sphingobacteriales  | Sphingobacteriaceae | Flammeovirgaceae    | <i>Flexithrix</i>       |
|                 |                     |                     | <i>Pedobacter</i>   |                         |
| Bacteroidia     | Bacteroidales       | Porphyromonadaceae  | <i>Articibacter</i> |                         |
| Proteobacteria  | Gammaproteobacteria | Alteromonadales     | Alteromonadaceae    | <i>Dysgonomonas</i>     |
|                 |                     |                     | Colwelliaceae       | <i>Microbulbifer</i>    |
|                 |                     | Cellvibrionales     | Cellvibrionaceae    | <i>Colwellia</i>        |
| Planctomycetes  | Planctomycetia      | Planctomycetales    | Planctomycetaceae   | <i>Saccharophagus</i>   |
| Verrucomicrobia | Verrucomicrobiae    | Verrucomicrobiales  | Verrucomicrobiaceae | <i>Rhodopirellula</i>   |
|                 |                     |                     |                     | <i>Verrucomicrobium</i> |

Un total de 47 secuencias fueron identificadas a partir del dominio catalítico de las dos secuencias pertenecientes al género *Flavobacterium*, con las que compartían 30 a 65 % de identidad a nivel de aminoácidos.

En la búsqueda de secuencias homólogas a las enzimas pertenecientes a la familia GH107, en cambio, solo se identificaron tres secuencias pertenecientes a *Colwellia* sp. SCGC AC281-N11 (genoma obtenido a partir de una célula individual), *Rhodopirellula* sp. SWK7 y *Rhodopirellula sallentina* SM41 (Tabla 3.2.). Estas secuencias compartieron 20 a 70 % de identidad con el dominio catalítico de las enzimas de referencia.

El tipo de ambiente en el que fueron aislados los microorganismos que contenían secuencias homólogas a fucoidanasas fue variado, entre ellos: marino (provenientes de la columna de agua y de sedimentos), suelo y ambientes que no habían sido determinados (Tabla 3.2.).

### **3.1.1.1 Análisis filogenético de las secuencias identificadas a partir de las enzimas pertenecientes a la familia GH107.**

El árbol filogenético obtenido a partir de los módulos catalíticos de las secuencias de referencia pertenecientes a la familia GH107 y las secuencias homólogas identificadas a partir de ellos se muestran en la Figura 3.1. Los módulos catalíticos que presentaron el mayor porcentaje de identidad fueron aquellos que pertenecían al organismo *Alteromonas* sp. SN-1009. Estas dos secuencias conformaron un grupo (clúster) cuyo grupo hermano fue la secuencia putativa identificada en el organismo *Colwellia* sp. SCGC AC281-N11. La fucoidanasa putativa de *S. violacea* DSS12, por otro lado, presentó un porcentaje de identidad a nivel de aminoácidos de aproximadamente 40 % con respecto a las secuencias de referencia de *Alteromonas* sp. SN-1009 y *Colwellia* sp. SCGC AC281-N11. Si bien la secuencia identificada en este estudio compartía porcentajes de identidad moderados con las secuencias de referencia de este clúster, el alto valor de “bootstrap” que presenta sugiere que esta secuencia pertenecería a la familia GH107. Los tres géneros incluidos en este clúster están cercanamente relacionados, dado que pertenecen al mismo orden (Alteromonadales) dentro de las Gammaproteobacterias.

Por otro lado, las secuencias putativas identificadas en cepas pertenecientes al género *Rhodopirellula* se agruparon, compartiendo un 48 % de identidad. Debido al bajo porcentaje de identidad que presentaron con secuencias de referencia de fucoidanasas (20 – 30 % de identidad) y su posición en el árbol filogenético, la potencial pertenencia de estas secuencias a la familia GH107 es incierta.

La secuencia de la enzima fucoidanasa de *M. fucanivorans* SW5, a partir de la cual no se identificaron secuencias en la base de datos de genomas bacterianos, fue altamente divergente del resto las secuencias analizadas (porcentajes de identidad menores al 25 %).

**Tabla 3.2.** Secuencias homólogas a las enzimas fucoidanasas de referencia identificadas en la base de datos de genomas microbianos IMG. Se indican los organismos en los cuales se identificaron las secuencias, el ambiente en el que fue aislado (marino: proveniente de la columna de agua, sedimento: sedimentos marinos y sd: sin determinar), el código del gen, el nombre del producto que codifica y por último el número de acceso a la base de datos IMG.

| Organismo  | Ambiente en el que fue aislado | ID Gen               | Nombre del producto génico   | IMG        |
|--|--------------------------------|----------------------|--|------------|
| <i>Aquimarina latercula</i> DSM-2041               | Marino                         | H526DRAFT_01831      | “Hypothetical protein”   | 2523974136 |
| <i>Aquimarina latercula</i> DSM-2041               | Marino                         | H526DRAFT_01830      | “Hypothetical protein”   | 2523974135 |
| <i>Aquimarina</i> sp. SW150                        | Sd                             | NoneDRAFT_00547      | “Por secretion system C-terminal sorting domain”                   | 2580986083 |
| <i>Aquimarina</i> sp. SW150                        | Sd                             | NoneDRAFT_04297      | “Por secretion system C-terminal sorting domain”                   | 2580989826 |
| <i>Aquimarina</i> sp. SW150                        | Sd                             | NoneDRAFT_04296_PA14 | “Domain/Carbohydrate binding module (family 6)/f5/8 type C domain” | 2580989825 |
| <i>Aquimarina</i> sp. SW150                        | Sd                             | NoneDRAFT_01542      | “Carbohydrate binding module (family 6)”                           | 2580987074 |
| <i>Arcticibacter svalbardensis</i> MN12-7          | Sd                             | ADIARSV_4252         | “Esterase”   | 2545498006 |
| <i>Arcticibacter svalbardensis</i> MN12-7          | Sd                             | ADIARSV_2216         | “Putative protein_transmembrana prediction”                        | 2545495967 |
| <i>Arcticibacter svalbardensis</i> MN12-7          | Suelo                          | ADRIARSV_2179        | “Hypothetical protein”   | 2545495930 |
| <i>Arcticibacter svalbardensis</i> MN12-7          | Suelo                          | ADIARSV_2919         | “Putative protein-transmembrane prediction”                        | 2545496670 |
| <i>Cellulophaga baltica</i> NN016038               | Sd                             | M667DRAFT_00713      | “Hypothetical protein”   | 2581797417 |
| <i>Colwellia</i> sp. SCGC AC281-N11 <sup>(1)</sup> | Sd                             | Ga0063149_01860      | “Concavalin A-like lectin/glucanases”                              | 2599809154 |
| <i>Cytophaga fermentans</i> IAM-14302DSM-9555      | Sd                             | CytfeDRAFT_1496      | “Carbohydrate binding module (family 6)”                           | 2509836672 |
| <i>Cytophaga fermentans</i> IAM-14302DSM-9555      | Sd                             | CytfeDRAFT_1524      | “Hypothetical protein”   | 2509836700 |
| <i>Cytophaga fermentans</i> IAM-14302DSM-9555      | Sd                             | CytfeDRAFT_0040      | “Hypothetical protein”   | 2509835216 |
| <i>Cytophaga fermentans</i> IAM-14302DSM-9555      | Sd                             | CytfeDRAFT_1093      | “Hypothetical protein”   | 2509836269 |
| <i>Cytophaga fermentans</i> JCM-21142              | Sd                             | JCM21142_1601        | “Endo-1,4-beta-xylanase A precursor”                               | 2579598062 |
| <i>Cytophaga fermentans</i> JCM-21142              | Sd                             | JCM21142_1564        | “Hypothetical protein”   | 2579598026 |
| <i>Cytophaga fermentans</i> JCM-21142              | Sd                             | JCM21142_93803       | “Hypothetical protein”   | 2579601207 |

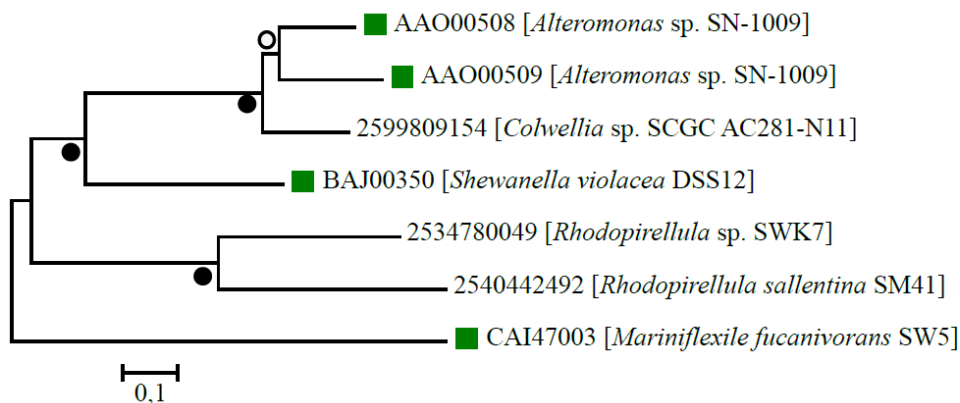
Tabla 3.2. Continuación.

| Organismo                                   | Ambiente en el que fue aislado | ID Gen            | Nombre del producto génico                                  | IMG <sup>(2)</sup> |
|---|--------------------------------|-------------------|---|--------------------|
| <i>Cytophaga fermentans</i> JCM-21142       | Sd                             | JCM21142_93373    | “Carbohydrate binding module”                               | 2579600783         |
| <i>Dysgomonas gradei</i> ATCCBAA-286        | Sd                             | HMPREF9455_02394  | “Hypothetical protein”                                      | 651387290          |
| <i>Echinicola pacifica</i> DSM 19836        | Sd                             | B050DRAFT_02714   | “Right handed beta helix region”                            | 2515475267         |
| <i>Flexithrix dorotheae</i> DSM-6795        | Sedimento                      | A3EMDRAFT_00125   | “Hypothetical protein”                                      | 2515802484         |
| <i>Formosa agariphila</i> KMM-39001         | Marino                         | BN863_18640       | “Conserved hypothetical protein”                            | 2586114395         |
| <i>Formosa agariphila</i> KMM-39001         | Marino                         | BN863_10080       | “Pectin lyase fold protein”                                 | 2586113537         |
| <i>Formosa agariphila</i> KMM-3901          | Sd                             | BN863_100700      | “Protein lyase fold protein”                                | 2586113536         |
| <i>Formosa agariphila</i> KMM-3901          | Suelo                          | BN863_10050       | “Pectin lyase fold protein”                                 | 2586113534         |
| <i>Formosa agariphila</i> KMM-3901          | Marino                         | BN863_18600       | “GDSL-like lipase/Acylhydrolase family protein”             | 2586114391         |
| <i>Formosa agariphila</i> KMM-3901          | Marino                         | BN863_10380       | “Por secretion system C-terminal sorting domain”            | 2586113567         |
| <i>Maribacter forsetti</i> DSM-18668        | Marino                         | P177DRAFT_02776   | “Hypothetical protein”                                      | 2574427858         |
| <i>Maribacter forsetti</i> DSM-18668        | Marino                         | P177DRAFT_02788   | “Hypothetical protein”                                      | 2574427870         |
| <i>Maribacter</i> sp. Hel_I_7               | Suelo                          | P178DRAFT_3736    | “Hypothetical protein”                                      | 2558433713         |
| <i>Microbulbifer thermaoleans</i> DSM-19189 | Marino                         | Ga0056041_01120   | “Carbohydrate binding module (family 6) F5/8 type C domain” | 2599618235         |
| <i>Microbulbifer thermaoleans</i> DSM-19189 | Sedimento                      | Ga0056041_02869   | “F5/8 type C domain”  | 2599619982         |
| <i>Pedobacter saltans</i> DSM-12145         | Suelo                          | Pedsa_3763        | “Hypothetical protein”                                      | 649984677          |
| Planctomycetes                              | Sd                             | pue10645_00011640 | “Autotransporter-associated beta repeat”                    | 2509277163         |
| Planctomycetes                              | Sd                             | pue06694_00024590 | “Hypothetical protein”                                      | 2509285542         |
| <i>Rhodopirellula baltica</i> SH1           | Marino                         | Ga0065370_115465  | “Right handed beta helix region”                            | 2607882865         |
| <i>Rhodopirellula baltica</i> SH1           | Sd                             | RB1470            | “Hypothetical protein-transmembrana prediction”             | 637433777          |
| <i>Rhodopirellula baltica</i> SH28          | Marino                         | RBSH_03031        | “Right handed beta helix region”                            | 2537187736         |
| <i>Rhodopirellula baltica</i> SWK14         | Sd                             | RBSWK_00449       | “Right handed beta helix region”                            | 2563978961         |
| <i>Rhodopirellula baltica</i> WH47          | Sedimento                      | RBWH47_04797      | “Right handed beta helix region”                            | 2536129258         |
| <i>Rhodopirellula europaea</i> 6C           | Marino                         | RE6C_05531        | “Putative secreted protein”                                 | 2534016826         |
| <i>Rhodopirellula europaea</i> SH398        | Sedimento                      | RESH_01934        | “Right handed beta helix region”                            | 2534019429         |

Tabla 3.2. Continuación.

| Organismo  | Ambiente en el que fue aislado | ID Gen             | Nombre del producto génico                       | IMG <sup>(2)</sup> |
|--|--------------------------------|--------------------|--|--------------------|
| <i>Rhodopirellula sallentina</i> SM41 <sup>(1)</sup> | Sedimento                      | RSSM_06681         | “Putative secreted protein”                      | 2540442492         |
| <i>Rhodopirellula</i> sp. SWK 7                      | Sedimento                      | RRSWK_03901        | “Dystroglycan-type cadherin-like domain protein” | 2534779559         |
| <i>Rhodopirellula</i> sp. SWK7 <sup>(1)</sup>        | Sedimento                      | RRSWK_04389        | “Putative secreted protein”                      | 2534780049         |
| <i>Saccharophagus degradans</i> 2-40                 | Marino                         | Sde_3709           | “Hypothetical protein”                           | 6379922806         |
| <i>Saccharophagus degradans</i> 2-40                 | Marino                         | Ga0065403_113879   | “F5/8 type C domain”                             | 2606911028         |
| <i>Verrucomicrobium spinosum</i> DSM-4136            | Sd                             | VspiD_010100000905 | “Hypothetical protein”                           | 642329229          |

(1) Secuencias identificadas a partir de las enzimas pertenecientes a la familia GH107.

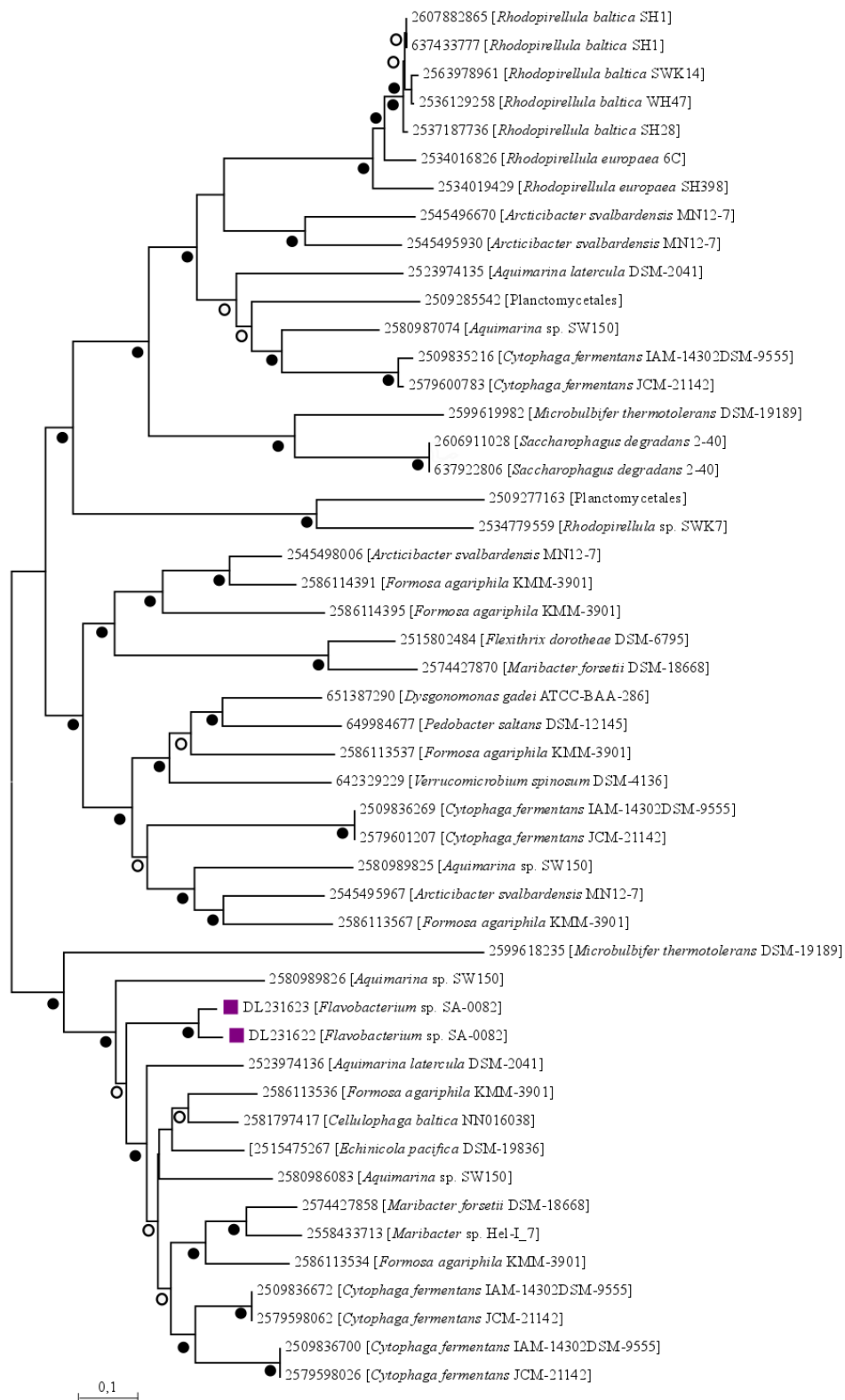


**Figura 3.1.** Árbol filogenético no enraizado obtenido por el método estadístico del vecino más cercano y el método de sustitución de aminoácidos de Jones-Taylor-Thornton (JTT) para las secuencias de referencia pertenecientes a la familia GH107 (señaladas con un cuadrado verde) y las secuencias genómicas identificadas a partir de ellas. Los círculos en los nodos hacen referencia a los valores de los soportes de “bootstrap” para 1000 réplicas (● = valores entre 80 y 100 %, ○ = valores entre 50 y 80 %). Las longitudes de las ramas son proporcionales al número de variaciones de aminoácidos. Los números corresponden al número de acceso a IMG y el nombre del organismo se indica entre corchetes.

### 3.1.1.2. Análisis filogenético de las secuencias identificadas a partir de los dominios catalíticos de las enzimas de la cepa *Flavobacterium* sp. SA-0082

Las relaciones filogenéticas entre las dos secuencias de fucoidanasas reportadas en una patente y las identificadas a partir de ellas en la base de datos IMG se muestran en la Figura 3.2. Se observaron tres grandes grupos, soportados por altos valores de “bootstrap”. Solo uno de estos clústers incluyó a las dos secuencias de referencia pertenecientes a *Flavobacterium* sp. SA-0082 (señaladas con un cuadrado violeta, Figura 3.2.). En algunos casos, las secuencias identificadas a partir de distintas cepas de la misma especie se agruparon en distintos grupos con secuencias que pertenecían a otras especies. Este resultado refleja una baja similaridad en la secuencia aminoacídica incluso en organismos de una misma especie.

Todas las secuencias de organismos pertenecientes al orden Planctomycetales se ubicaron en el clúster superior del árbol filogenético. Estas secuencias fueron altamente divergentes de las secuencias de fucoidanasas de *Flavobacterium* sp. SA-0082 (15 - 25 % de identidad), al igual que los otros miembros de este clúster (< 30 %). Por el contrario, los porcentajes de identidad entre las secuencias de referencia y las 14 secuencias identificadas incluidas en el clúster inferior del árbol filogenético fueron menos divergentes (> 55 %). Este grupo de secuencias es investigado con mayor detalle en el siguiente punto.



**Figura 3.2.** Árbol filogenético no enraizado de las secuencias de referencia reportadas en la patente (cuadrado violeta) y las identificadas en la base de datos IMG. El árbol fue obtenido por el método del vecino más cercano. Los círculos en los nodos hacen referencia a los valores de los soportes de “bootstrap” para 1000 réplicas (● = valores entre 80 y 100%, ○ = valores entre 50 y 80%). Los números corresponden al número de acceso a IMG y el nombre del organismo se muestra entre corchetes a excepción de las dos secuencias identificadas en un microorganismo perteneciente al orden Planctomycetales.

### 3.1.2. Análisis del contexto genómico y sintenia compartida

En general, se observó un bajo nivel de conservación en la vecindad de las secuencias genómicas identificadas (señaladas con una flecha roja, Figura 3.3. y Figura 3.4.). En la Figura 3.3. se observa el análisis del contexto genómico para las secuencias homólogas a las enzimas de la familia GH107. Con respecto a las secuencias de referencia, solo el genoma de la cepa *S. violacea* DSS12 se encontraba depositado en la base de datos IMG. El entorno genómico de las secuencias identificadas en este estudio pertenecientes a los organismos *Colwellia* sp. SCGC AC281-N11 y *R. sallentina* SM41 contenía secuencias que podrían codificar para enzimas arilsulfatasas (representadas con el número 2 en la Figura 3.3., Tabla 3.3.). Esta enzima es requerida durante la depolimerización de polisacáridos sulfatados para catalizar la hidrólisis de los grupos sulfato. Otro elemento identificado en el genoma del organismo *R. sallentina* SM41 que podría estar relacionado con la degradación de fucoidanos es la enzima  $\alpha$ -L-fucosidasa (EC: 3.2.1.51), indicada con el número 11. Por otro lado, se identificaron otros elementos que potencialmente podrían estar involucrados en el transporte de oligómeros al interior celular y en su metabolismo (números 3 y 4). El componente 3 corresponde al receptor dependiente de TonB y el 4 a la proteína Concavalina A unido a lectina/familia glucanasa. Por lo tanto, este análisis permitió identificar varios elementos en el contexto genómico de las secuencias identificadas que posiblemente estén involucrados en la asimilación de fucoidanos. Este resultado provee evidencia adicional sobre la posible función de las secuencias putativas de fucoidanasas identificadas a partir de genomas bacterianos.

En la Figura 3.4. se muestra el contexto genómico de secuencias identificadas a partir del dominio catalítico de las fucoidanasas pertenecientes al organismo *Flavobacterium* sp. SA-0082. En este análisis sólo se incluyeron las secuencias más cercanas a las fucoidanasas, las cuales pertenecían a organismos del phylum Bacteroidetes. El genoma de *Flavobacterium* sp. SA-0082 no se encontraba disponible en la base de datos, por lo que no se muestra en la Figura 3.4.

La organización génica de los fragmentos genómicos presentaba un bajo grado de conservación, aún en organismos pertenecientes al mismo género. Se pudo identificar, sin embargo, un elemento que acompañó a todas las secuencias putativas de fucoidanasas, a excepción de aquellas de cepas pertenecientes al género *Aquimarina*. Este gen podría codificar para la enzima arilsulfatasa (señalada con el número 2, Figura 3.4.) y resultó relativamente abundante en el genoma de algunos organismos tales como aquellos pertenecientes a los géneros *Maribacter* y *Cellulophaga*.

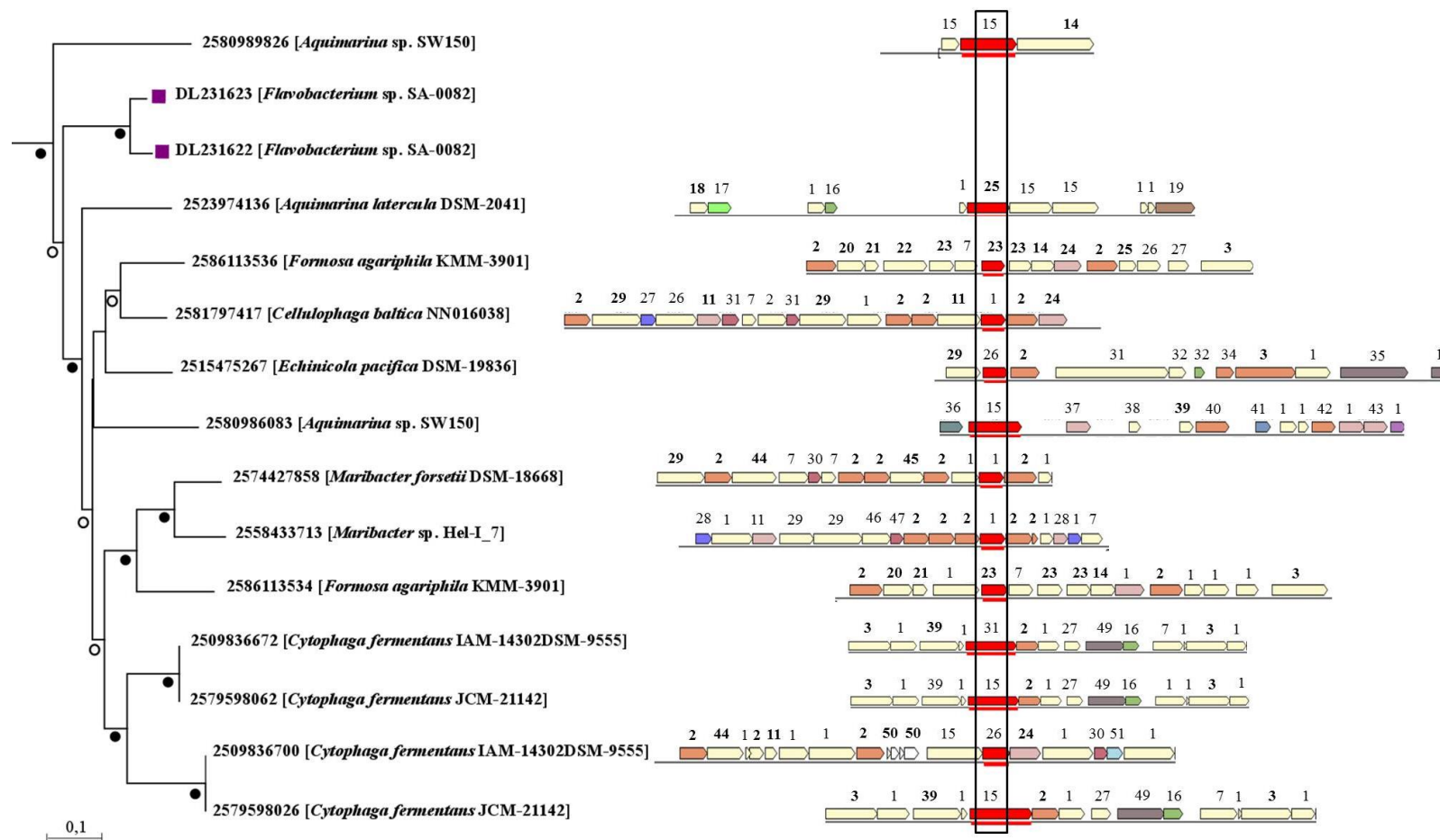
Genes que codifican a la enzima  $\alpha$ -L-fucosidasa también se identificaron en los organismos *Cellulophaga báltica* NN016038, *Maribacter* sp. He-I\_7 y *Cytophaga fermentans* IAM-14302

---

DSM-9555. Otras secuencias de los contextos genómicos analizados podrían codificar para acetil-esterasas (EC 3.1.1.6), enzimas que también podrían estar implicadas en el metabolismo de fucoidanos (número 20 en la Figura 3.4.). Secuencias codificantes para esta enzima, que cataliza la hidrólisis de grupos acetil-éster, se encontraron en el genoma del organismo *Formosa agariphila* KMM-3901.

Las secuencias de fucoidanasas putativas pertenecientes al organismo *F. agariphila* KMM-3901 presentaron una correspondencia parcial en cuanto a su entorno genómico, a pesar de que formaron parte de distintos clústers. Situación similar ocurrió con las secuencias pertenecientes a la cepa *C. fermentans* JCM-21142 y una de las secuencias de la cepa *C. fermentans* IAM-14302DSM-9555.





**Figura 3.4.** Esquema de genes próximos a las secuencias homólogas identificadas (señaladas con rojo) a partir de las secuencias pertenecientes al organismo *Flavobacterium* sp. SA-0082 (señaladas con un cuadrado violeta). Los números sobre las flechas corresponden al nombre del producto génico indicado en la Tabla 3.3. Los números en negrita corresponden a elementos que podrían estar involucrados en el metabolismo de carbohidratos. Las secuencias homólogas a las enzimas fucoidanasas se señalan con un recuadro negro.

**Tabla 3.3.** Número asignado a las secuencias de los contextos genómicos incluidos en las Figuras 3.3. y 3.4. En todos los casos se indica el nombre del producto génico asignado por la anotación funcional en el sistema IMG. Se resaltan en negrita los nombres de las secuencias relacionadas con el metabolismo de carbohidratos.

| Número asignado en el gráfico | Nombre dado por anotación funcional                                  |
|-------------------------------|--|
| 1                             | "Hypothetical protein"   |
| 2                             | <b>"Arylsulfatase A"</b>   |
| 3                             | <b>"Ton-B dependent receptor"</b>                                    |
| 4                             | <b>"Concavalina A-like lectin/glucanases superfamily protein"</b>    |
| 5                             | "NAD <sup>+</sup> synthase (glutamine-hydrolysing)"                  |
| 6                             | "Membrane or secreted protein"                                       |
| 7                             | "Protein of unknown function"  |
| 8                             | "putative secreted protein"  |
| 9                             | "Outer membrane efflux protein"                                      |
| 10                            | "tRNA-Leu"   |
| 11                            | <b>"Alpha-L-fucosidase"</b>  |
| 12                            | "Planctomycete cytochrome C"   |
| 13                            | "SLA1 homology domain 1"   |
| 14                            | <b>"Carbohydrate binding module"</b>                                 |
| 15                            | "Por secretion system C-terminal sorting domain-cointaining protein" |
| 16                            | "Component transcriptional regulator"                                |
| 17                            | "Imidazolonepropionase"  |
| 18                            | <b>"Beta-lactamase"</b>  |
| 19                            | "Outer membrane protein TolC"  |
| 20                            | <b>"Acetyl esterase"</b>   |
| 21                            | <b>"Alpha/beta hydrolase family"</b>                                 |
| 22                            | <b>"Alpha D-xyloside xylohydrolase"</b>                              |
| 23                            | <b>"Pectin lyase fold protein"</b>                                   |
| 24                            | <b>"Xylan 1,4-beta-xylosidase"</b>                                   |
| 25                            | <b>"Pectate lyase superfamily protein"</b>                           |
| 26                            | "Right handed beta helix region"                                     |
| 27                            | "Unsaturated chondroitin disaccharide hydrolase"                     |
| 28                            | "Acetyl esterase/lipase"   |
| 29                            | <b>"Heparinase II/III -like protein"</b>                             |
| 30                            | "Lysophospholipase L1"   |
| 31                            | "Gliding motility-associated C-terminal domain-containing protein"   |
| 32                            | "Type IX secretion system membrane protein, PorP/SprF family"        |
| 33                            | "RNA polymerase sigma factor, sigma-70 family"                       |
| 34                            | "FecR family protein"  |
| 35                            | "PAS/PAC sensor hybrid histidine kinase (EC 2.7.13.3)"               |
| 36                            | "Phenylacetate-CoA ligase"   |
| 37                            | "Phosphomannomutase"   |
| 38                            | "Rhomboid family protein"  |
| 39                            | <b>"Alginate lyase"</b>  |

| Número asignado en el gráfico | Nombre dado por anotación funcional               |
|-------------------------------|---|
| 40                            | "Sulfate permease, SulP family"                   |
| 41                            | "EamA domain-containing membrane protein RarD"    |
| 42                            | "Sodium/proton antiporter, CPA1 family"           |
| 43                            | "Glucose-1-phosphate adenylyltransferase"         |
| 44                            | <b>"Alpha-glucosidase"</b>                        |
| 45                            | <b>"Glucosyl hydrolases family 28"</b>            |
| 46                            | "FAD dependent oxidoreductase"                    |
| 47                            | "Acyl-CoA thioesterase"                           |
| 48                            | "Gluconate 5-dehydrogenase"                       |
| 49                            | "Ligand-binding sensor domain-containing protein" |
| 50                            | <b>"Beta-fructosidases (levanase/invertase)"</b>  |
| 51                            | "Formylglutathione hydrolase FrmB"                |

### 3.1.3. Análisis del potencial degradativo de carbohidratos

El rango de enzimas glucósido hidrolasa (GH) reportados en la base de datos CAZy de algunos de los microorganismos identificados en este estudio se detallan en la Tabla 3.4. Dado que las familias GH5, GH13 y GH30 han sido clasificadas en subfamilias, se indica además el número de las mismas entre paréntesis. Las familias GH solo son reportadas para 5 organismos, dado que los restantes organismos detallados en la Tabla 3.2. no registraron entradas sobre el potencial general degradativo de hidratos de carbono en la base de datos CAZy. Todos los genomas contenían una gran cantidad de genes pertenecientes a la clase GH, y la diversidad de familias es indicativo del potencial que presentan estas cepas para degradar una gran diversidad de carbohidratos.

**Tabla 3.4.** Familias y subfamilias GH (el número se indica entre paréntesis) reportadas para los organismos identificados en genomas microbianos.

| Organismo                            | Familias GH  |
|--------------------------------------|--|
| <i>Rhodopirellula baltica</i> SH1    | GH2, GH5, GH10, GH13 (2), GH16, GH20, GH29, GH32, GH33, GH43, GH57, GH77, GH78, GH86, GH95, GH97, GH117, GH127, GH130.   |
| <i>Cellulophaga baltica</i> NN016038 | GH1, GH2, GH3, GH5 (1), GH13, GH15, GH16, GH20, GH23, GH28, GH29, GH29, GH31, GH32, GH37, GH39, GH63, GH65, GH73, GH82, GH88, GH95, GH104, GH113, GH114, GH115, GH117, GH127, GH130.                       |
| <i>Formosa agariphila</i> KMM-3901]  | GH1, GH2, GH3, GH5 (1), GH13 (2), GH15, GH16, GH17, GH20, GH23, GH25, GH26, GH28, GH29, GH30, GH31, GH39, GH43, GH65, GH73, GH78, GH86, GH88, GH95, GH97, GH105, GH106, GH109, GH113, GH115, GH117, GH130. |

| Organismo                            | Familias GH   |
|--------------------------------------|---|
| <i>Pedobacter saltans</i> DSM-12145  | GH1, GH2, GH3, GH5 (1), GH10, GH11, GH13, GH16, GH18, GH20, GH23, GH24, GH25, GH26, GH27, GH28, GH29, GH30 (2), GH31, GH33, GH36, GH43, GH51, GH53, GH67, GH73, GH78, GH88, GH89, GH92, GH93, GH97, GH105, GH106, GH109, GH115, GH117, GH127, GH128, GH130.               |
| <i>Saccharophagus degradans</i> 2-40 | GH2, GH3, GH5 (5), GH6, GH9, GH10, GH11, GH13 (5), GH15, GH16, GH18, GH20, GH23, GH24, GH26, GH27, GH28, GH30 (1), GH31, GH35, GH43, GH50, GH51, GH53, GH55, GH63, GH67, GH73, GH77, GH79, GH81, GH86, GH94, GH95, GH97, GH103, GH105, GH108, GH115, GH117, GH128, GH130. |

### 3.2. En metagenomas marinos

#### 3.2.1. Identificación de secuencias homólogas

En la base de datos IMG/M existían 50 estudios metagenómicos de ambientes marinos al momento de realizar la búsqueda de secuencias putativas de fucoidanasas. En 11 de estos estudios se identificaron secuencias que cumplían con los parámetros fijados para este análisis, si bien en sólo algunas de las muestras de cada estudio (Tabla 3.5.). En total se identificaron 31 secuencias homólogas a las enzimas de referencia en 22 metagenomas, cuyos códigos de acceso a la base de datos IMG/M se indican en la Tabla 3.6. (columna Gen ID). Todos los productos de las secuencias identificadas correspondieron a proteínas hipotéticas, a excepción de la secuencia “TDF\_OR\_ARG05\_123mDRAFT\_10138301”, cuyo nombre otorgado por la anotación funcional del sistema IMG/M fue “*alginate lyase*”.

El rango de longitud de los scaffolds identificados (señalado en la columna “scaffold” de la Tabla 3.4.) fue desde un mínimo de 335 nucleótidos correspondiendo a aproximadamente 111 aminoácidos, hasta un máximo de 2870 nucleótidos que corresponde a aproximadamente 956 aminoácidos.

#### 3.2.2. Evidencias funcionales en las secuencias metagenómicas

Solo 7 secuencias metagenómicas presentaron evidencias funcionales relacionadas con las bases de datos COG y Pfam (Tabla 3.7.). El dominio Pfam “Beta hélix” fue identificado en la gran mayoría de estas secuencias. Este dominio determina la configuración de la estructura secundaria de las proteínas que lo presentan, adoptando una estructura formada por asociaciones de láminas beta paralelas dispuestas en un patrón helicoidal de dos o tres caras. Otro dominio Pfam identificado fue “*Alginate lyase*”, implicado en la depolimerización de alginatos. Con respecto al análisis basado en la base de datos de grupos de ortólogos COG, permitió inferir la

función de sólo una secuencia, correspondiente a una “galacturonasa”, implicada en el metabolismo y transporte de carbohidratos.

**Tabla 3.5.** Estudios metagenómicos de la base de datos IMG/M que contenían secuencias homólogas a fucoidanasas. Se indica el número de set de datos totales y el número de metagenomas en los cuales se identificaron secuencias homólogas.

| Nombre del estudio   | N° de set de datos | N° de set de datos con resultados |
|--|--------------------|-----------------------------------|
| Marine microbial communities from Global Ocean Sampling (GOS).   | 56                 | 1                                 |
| Deep ocean microbial communities from the Global Malaspina Expedition.                                       | 77                 | 1                                 |
| Saline water and sediment microbial community from Etoliko Lagoon, Greece.                                   | 6                  | 4                                 |
| Marine sediment microbial communities from the Peru Margin, Gulf of Mexico and South Pacific Gyre.           | 1                  | 1                                 |
| Marine microbial communities from Delaware Coast.  | 7                  | 1                                 |
| Marine and estuarine microbial communities from Columbia River Coastal Margin.                               | 12                 | 4                                 |
| Marine archaeal communities from Monterey Bay, CA, that are ammonia-oxidizing.                               | 33                 | 5                                 |
| Marine sediment archaeal communities from Santa Barbara Basin, CA, that are anaerobic and methane-oxidizing. | 7                  | 2                                 |
| Hypersaline microbial communities from Lake Vida, Antarctica.  | 5                  | 2                                 |
| Fossil microbial community from Whale Fall, Santa Cruz Basin of the Pacific Ocean.                           | 3                  | 1                                 |
| Marine microbial communities from chronically polluted sediments in four geographic locations.               | 23                 | 8                                 |

**Tabla 3.6.** Secuencias homólogas a enzimas fucoidanasas identificadas en estudios metagenómicos reportados en la base de datos IMG/M. Se indica el nombre, código de identificación y nombre de la muestra de la cual se identificaron secuencias homólogas, el dominio catalítico utilizado, la identificación del gen y el scaffold.

| Nombre del estudio  | ID del estudio | Dominio catalítico utilizado         | Nombre de la muestra   | Gen ID                     | Scaffold  |
|---|----------------|--------------------------------------|--|----------------------------|---|
| Marine microbial communities from Global Ocean Sampling (GOS)                                     | Gs0055716      | FcnA - SVI_0379 - Fda1 - Fda2        | Mangrove swamp microbial communities from Isabella Island, Ecuador - GS032   | GOS2247_10085102           | GOS2247_1008510 (1256 nt)   |
| Deep ocean microbial communities from the Global Malaspina Expedition                             | Gs0053074      | SVI_0379 - Fda2                      | Marine viral communities from the Deep Pacific Ocean - MSP-118 (Metagenome Std Draft M1678 Malaspina viral metagenome MSP-118, ASSEMBLY_DATE=20130718) | JGI24657J20077_10709771    | JGI24657J20077_1070977 (357nt)  |
| Saline water and sediment microbial community from Etoliko Lagoon, Greece                         | Gp0092397      | 1 y 2<br><i>Flavobacterium</i><br>sp | Saline sediment microbial community from Etoliko Lagoon, Greece (combined Illumina-454-Sanger assembly)  | 2154732346                 | Sediment microbial community from Etoliko Lagoon, Greece (combined Illumina-454-Sanger assembly) : Etoli__NODE21NE_1345959_923_ (943nt) |
|   |                |                                      |  | 2154482561                 | Sediment microbial community from Etoliko Lagoon, Greece (combined Illumina-454-Sanger assembly) : Etoli__1st61650 (413nt)              |
|   |                |                                      | Saline sediment microbial community from Etoliko Lagoon, Greece Reannotation (re-annotation)   | Ga0074242_120217421        | Ga0074242_12021742 (943nt)  |
|   |                |                                      | Ga0074242_115443361  | Ga0074242_11544336 (413nt) |   |
| Marine sediment microbial communities from the Peru Margin, Gulf of Mexico and South Pacific Gyre | Gs0063499      | Fda1                                 | Marine sediment microbial communities from Arctic Ocean, off the coast from Alaska - sample from low methane PC12-247-20cm                             | ASLM20b_07669100           | ASLM20b_GJMWM0N02G3 HIM (335nt)   |

Tabla 3.6. Continuación

| Nombre del estudio  | ID del estudio | Dominio catalítico utilizado          | Nombre de la muestra   | Gen ID                        | Scaffold                          |
|---|----------------|---------------------------------------|--|-------------------------------|-----------------------------------|
| Marine microbial communities from Delaware Coast                              | Gs0063440      | SVI_0379 - Fda2                       | Marine microbial communities from Delaware Coast, sample from Delaware MO Late spring/early summer         | DelMOSpr2010DRAFT_101531<br>1 | DelMOSpr2010DRAFT_101531 (1133nt) |
| Marine and estuarine microbial communities from Columbia River Coastal Margin | Gs0047387      | SVI_0379 - Fda1 - Fda2                | Estuarine microbial communities from Columbia River, sample from South Channel ETM site, GS313-3LG-ETM-15m | none_1117511                  | none_p111751 (496nt)              |
|   |                | FcnA                                  | Marine microbial communities from Columbia River, CM, sample from Newport Hydroline, GS310-0p8-Hyp-75m     | none_0228941                  | none_p0228941 (496nt)             |
|   |                | 1 y 2<br><i>Flavobacterium</i><br>sp. |  | none_0518291                  | none_p0518291 (442nt)             |
| Marine archaeal communities from Monterey Bay, CA, that are ammonia-oxidizing | Gs0099541      | 1 y 2<br><i>Flavobacterium</i><br>sp. | Ammonia-oxidizing marine microbial communities from Monterey Bay, California, USA - CAN11_08_M0_20         | JGI26112J46591_10812521       | JGI26112J46591_1081252 (369nt)    |
|   |                |                                       | Ammonia-oxidizing marine microbial communities from Monterey Bay, California, USA - CAN11_07_M0_10         | JGI26115J46592_10573961       | JGI26115J46592_1057396 (449nt)    |
|   |                |                                       | Ammonia-oxidizing marine microbial communities from Monterey Bay, California, USA - CAN11_18_M0_10         | JGI26114J46594_10474891       | JGI26114J46594_1047489 (646nt)    |
|   |                |                                       | Ammonia-oxidizing marine microbial communities from Monterey Bay, California, USA - CAN11_66_BW_10         | JGI26118J46590_10025313       | JGI26118J46590_1002531 (2870nt)   |
|   |                | 2<br><i>Flavobacterium</i><br>sp.     | Ammonia-oxidizing marine microbial communities from Monterey Bay, California, USA - CAN11_54_BW_10         | JGI26119J46589_10639041       | JGI26119J46589_1063904 (440nt)    |

Tabla 3.6. Continuación

| Nombre del estudio  | ID del estudio | Dominio catalítico utilizado       | Nombre de la muestra   | Gen ID                           | Scaffold                                 |
|---|----------------|------------------------------------|--|----------------------------------|--|
| Marine sediment archaeal communities from Santa Barbara Basin, CA, that are anaerobic and methane-oxidizing | Gs0046212      | SVI_0379 y Fda2                    | Marine sediment archaeal communities from Santa Barbara Basin, CA, that are methane-oxidizing, sample 0-3 cm           | ASA120_00228020                  | ASA120_GJFD58A02GM4XR (418nt)            |
|   |                | 2<br><i>Flavobacterium</i> sp.     |  | ASA120_02388680                  | ASA120_GJFD58A02G3UQB (497nt)            |
| Hypersaline microbial communities from Lake Vida, Antarctica  | Gs0067859      | SVI_0379                           | Hypersaline microbial communities from Lake Vida, Antarctica - Brine Hole Two >0.2 micron                              | JGI1221J13441_100292331          | JGI1221J13441_10029233 (1574nt)          |
|   |                |                                    |  | JGI1221J11331_10716531           | JGI1221J11331_1071653 (612nt)            |
| Fossil microbial community from Whale Fall, Santa Cruz Basin of the Pacific Ocean                           | Gs0055718      | 1 y 2<br><i>Flavobacterium</i> sp. | Fossil microbial communities from Whale Fall, Santa Cruz Basin of the Pacific Ocean – Bone                             | 2001497724                       | 2001402879 (1417nt)                      |
| Marine microbial communities from chronically polluted sediments in four geographic locations               | Gs0063447      | 1<br><i>Flavobacterium</i> sp.     | Marine microbial communities from chronically polluted sediments in the Baltic Sea-site KBA sample SWE21_205m          | BS_KBA_SWE21_205mDRAFT_100399101 | BS_KBA_SWE21_205mDRAFT_10039910 (1149nt) |
|   |                |                                    | Microbial communities from chronically polluted sediments in Antarctica-King George Island site 52 sample ANT05_23.45m | KGI_S2_ANT05_2345mDRAFT_10666891 | KGI_S2_ANT05_2345mDRAFT_c1066689 (572nt) |
|   |                | 1 y 2<br><i>Flavobacterium</i> sp. | Microbial communities from chronically polluted sediments in Antarctica-King George Island site 52 sample ANT05_23.45m | KGI_S2_ANT05_2345mDRAFT_10328402 | KGI_S2_ANT05_2345mDRAFT_c1032840 (919nt) |
|   |                |                                    | Microbial communities from chronically polluted sediments in Antarctica-King George Island site S1 sample ANT03_9.5m   | KGI_S1_ANT03_95mDRAFT_10035671   | KGI_S1_ANT03_95mDRAFT_c1003567 (2151nt)  |
|   |                |                                    | Microbial communities from chronically polluted sediments in Antarctica-King George Island site S1 sample ANT03_9.5m   | KGI_S1_ANT03_95mDRAFT_10236371   | KGI_S1_ANT03_95mDRAFT_c1023637 (759nt)   |

Tabla 3.6. Continuación

| Nombre del estudio  | ID del estudio | Dominio catalítico utilizado          | Nombre de la muestra  | Gen ID                          | Scaffold                                |
|---|----------------|---------------------------------------|---|---------------------------------|---|
| Marine microbial communities from chronically polluted sediments in four geographic locations | Gs0063447      | 1 y 2<br><i>Flavobacterium</i><br>sp. | Marine microbial communities from chronically polluted sediments in the Baltic Sea-site KBA sample SWE12_21m      | BS_KBA_SWE12_21mDRAFT_101822181 | BS_KBA_SWE12_21mDRAFT_c10182218 (501nt) |
|   |                |                                       | Marine microbial communities from chronically polluted sediments in Tierra del Fuego_site92 sample ARG04_12.3m    | TDF_OR_ARG04_123mDRAFT_10107671 | TDF_OR_ARG04_123mDRAFT_c1010767 (863nt) |
|   |                | 2<br><i>Flavobacterium</i><br>sp.     | Marine microbial communities from chronically polluted sediments in Tierra del Fuego, site 012 sample ARG05_12.3m | TDF_OR_ARG05_123mDRAFT_10138301 | TDF_OR_ARG05_123mDRAFT_1013830 (2049nt) |
|   |                |                                       | Marine microbial communities from chronically polluted sediments in the Baltic Sea-site KBA sample SWE02_21m      | BS_KBA_SWE02_21mDRAFT_103946921 | BS_KBA_SWE02_21mDRAFT_10394692 (343nt)  |

**Tabla 3.7.** Evidencias funcionales identificadas en las secuencias metagenómicas.

| Gen ID                          | Evidencia funcional                                 |
|---------------------------------|---|
| Ga0074242_120217421             | Pfam 13229 Beta hélix                               |
| Ga0074242_115443361             | Pfam 13229 Beta hélix                               |
| JGI26118J46590_10025313         | Pfam 13229 Beta hélix                               |
| JGI26119J46589_10639041         | Pfam 13229 Beta helix<br>COG 5434 Polygalacturonase |
| ASA120_02388680                 | Pfam 13229 Beta hélix                               |
| TDF_OR_ARG04_123mDRAFT_10107671 | Pfam 13229 Beta hélix                               |
| TDF_OR_ARG05_123mDRAFT_10138301 | Pfam 13229 Beta helix<br>Pfam 05426 Alginate lyase  |

### 3.2.3. Relación con secuencias genómicas y búsqueda de dominios conservados

A fin de conocer el grado de divergencia entre las secuencias identificadas en el sistema Integrado de Manejo de Secuencias Metagenómicas (IMG/M) y las secuencias depositadas en GenBank, como así también evaluar la posible presencia de dominios o motivos conservados en las mismas, se utilizó el algoritmo DELTA-BLAST. En la Tabla 3.8. se especifican las secuencias que presentaron el mayor porcentaje de identidad y cobertura con respecto a las secuencias metagenómicas. Solo una secuencia no obtuvo resultados en este análisis.

Las secuencias más cercanas de la base de datos correspondían en su mayor parte a proteínas hipotéticas de cepas del género *Rhodopirellula*, con porcentajes de cobertura e identidad de 30 - 100 % y 31 - 84 %, respectivamente. En menor proporción correspondieron a secuencias pertenecientes a los géneros *Saccharicrinis*, *Polaribacter*, *Paenibacillus*, *Tamlana* y *Algibacter*, los cuales no habían sido identificados con anterioridad en la identificación de secuencias homólogas en genomas bacterianos. El género *Paenibacillus* se encuentra dentro del phylum Firmicutes, el cual no había sido identificado en análisis previos. Los 4 géneros restantes pertenecen al phylum Bacteroidetes. Por otro lado, en 3 ocasiones la secuencia más cercana correspondió a la secuencia de aminoácidos de las enzimas fucoidanasas pertenecientes a cepas utilizadas como referencia en este trabajo: FcnA de *M. fucanivorans* SW5 y SVI\_0379 de *S. violacea* DSS12.

Gran parte de las secuencias genómicas que presentaron altos porcentajes de identidad y cobertura con respecto a las secuencias metagenómicas, pertenecieron a organismos que ya habían sido identificados en análisis previos. Entre ellas se encuentran: *Rhodopirellula* sp. SWK7, *R. europea* 6C, *F. agariphila* KMM-3901, *C. baltica* y *E. pacifica*. En este análisis no solo la cepa fue coincidente sino también la secuencia de aminoácidos con respecto a secuencias genómicas que habían sido identificadas en la base de datos de genomas microbianos (IMG).

En la gran mayoría de las secuencias metagenómicas no se pudieron identificar dominios conservados (sd, sin determinar), y los que se detectaron fueron diversos (Tabla 3.8.). El dominio conservado “AmyAc\_family superfamily (Alpha-L-fucosidase)” fue identificado en una de las secuencias. La abreviación AmyAc proviene del inglés “Alpha amylase catalytic domain family”, que comprende una de las mayores familias dentro de la clase GH, actuando sobre el almidón, el glucógeno y polisacáridos relacionados. Dentro de los miembros de la familia se incluye la enzima alfa-L-fucosidasa. El dominio conservado “Beta\_helix” fue detectado en 6 secuencias metagenómicas. En solo 2 secuencias metagenómicas se identificó el dominio catalítico “NosD\_copper\_fam”. La sigla NosD hace referencia a una proteína periplasmática requerida para la maduración de la enzima óxido nitroso reductasa. El dominio “Nodulin-21\_like\_1 (CC1\_like superfamily)” fue identificado en una sola secuencia metagenómica. Este dominio incluye a una familia de proteínas estrechamente relacionadas con la proteína Nodulin-21, una proteína que pareciera estar involucrada en la fijación simbiótica de nitrógeno. Este dominio también se relaciona con la proteína CCC1, que consta de una proteína transmembrana que funciona como un transportador de hierro y manganeso. Por último, en una única secuencia fue identificado el dominio conservado “Alginate\_lyase (AlgLyase superfamily)”, que incluye a enzimas que depolimerizan alginatos.

Los resultados obtenidos en este análisis son consistentes con aquellos descritos en el apartado 3.2.2. (Tabla 3.7), a excepción de las secuencias con ID “Ga00742\_120217421” y “TDF\_OR\_ARG05\_123mDRAFT\_10138301”, para los cuales el dominio conservado reportado por anotación funcional en el sistema IMG/M no fue detectado en el análisis de dominios conservados por medio de la herramienta DELTA-BLAST en GenBank.

**Tabla 3.8.** Dominios conservados (DC) detectados en las secuencias metagenómicas a partir del algoritmo DELTA-BLAST. Si no se identificaron dominios se indica con la abreviación “sd” (sin determinar). A la derecha, se indican las secuencias que resultaron más similares a las secuencias metagenómicas.

| Gen ID                    | DC  | Nombre  | Secuencia identificada |             |             |
|---------------------------|---|---|------------------------|-------------|-------------|
|                           |   |   | GenBank                | % cobertura | % identidad |
| GOS2247_10085102          | AmyAc_family superfamily (Alpha-L-fucosidase) | Putative secreted protein [ <i>Rhodopirellula</i> sp. SWK7]                     | EMI43088.1             | 30%         | 84%         |
| JGI24657J20077_10709771   | Sd  | Hypothetical protein [ <i>Rhodopirellula</i> sp. SWK7]                          | WP_044255141.1         | 100%        | 42%         |
| 2154732346                | Sd  | T9SS C-terminal target domain-cointaing protein [ <i>Aquimarina latercula</i> ] | WP_035084141.1         | 97%         | 60%         |
| 2154482561                | Beta_helix                                    | Hypothetical protein [ <i>Rhodopirellula europaea</i> ]                         | WP_008665660.1         | 100%        | 81%         |
| Ga0074242_120217421       | Sd  | Hypothetical protein [ <i>Saccharicrinis fermentans</i> ]                       | WP_044213879.1         | 100%        | 61%         |
| Ga0074242_115443361       | Beta_helix                                    | Putative secreted protein [ <i>Rhodopirellula europaea</i> 6C]                  | EMB13743.1             | 100%        | 82%         |
| ASLM20b_07669100          | Sd  | Putative secreted protein [ <i>Rhodopirellula</i> sp. SWK7]                     | EMI43088.1             | 87%         | 43%         |
| DelMOSpr2010DRAFT_1015311 | Sd  | Hypothetical protein [ <i>Shewanella violacea</i> ]                             | WP_013049664.1         | 98%         | 81%         |
| none_1117511              | Sd  | Sd  | sd                     |             |             |
| none_0228941              | Sd  | Hypothetical protein [ <i>Shewanella violacea</i> ]                             | WP_013049664.1         | 98%         | 31%         |
| none_0518291              | Sd  | Alpha-1,4-endofucoidanase precursor [ <i>Mariniflexile fucanivorans</i> ]       | CAI47003.1             | 99%         | 40%         |
| none_022141               | Sd  | Hypothetical protein [ <i>Polaribacter</i> sp. Hel1_85]                         | WP_036824722.1         | 99%         | 52%         |
| JGI26112J46591_10812521   | Sd  | Hypothetical protein [ <i>Rhodopirellula europaea</i> ]                         | WP_008665660.1         | 97%         | 80%         |
| JGI26115J46592_10573961   | Sd  | Hypothetical protein [ <i>Rhodopirellula europaea</i> ]                         | WP_008665660.1         | 100%        | 76%         |
| JGI26114J46594_10474891   | Sd  | Hypothetical protein [ <i>Rhodopirellula europaea</i> ]                         | WP_008665660.1         | 100%        | 68%         |
| JGI26118J46590_10025313   | Beta_helix                                    | Hypothetical protein [ <i>Rhodopirellula europaea</i> ]                         | WP_008665660.1         | 100%        | 66%         |
| JGI26119J46589_10639041   | Beta_helix                                    | Hypothetical protein [ <i>Rhodopirellula europaea</i> ]                         | WP_008665660.1         | 100%        | 83%         |

Tabla 3.8. Continuación.

| Gen ID                           | DC  | Nombre   | Secuencia identificada |             |             |
|----------------------------------|---|--|------------------------|-------------|-------------|
|                                  |   |  | GenBank                | % cobertura | % identidad |
| ASA120_00228020                  | Sd  | Putative secreted protein [ <i>Rhodopirellula</i> sp. SWK7]                      | EM143088.1             | 88%         | 36%         |
| ASA120_02388680                  | NosD_copper_fam y Beta_helix  | Hypothetical protein [ <i>Saccharicrinis fermentans</i> ]                        | WP_044213879.1         | 95%         | 66%         |
| JGI1221J13441_100292331          | Sd  | Putative secreted protein [ <i>Rhodopirellula</i> sp. SWK7]                      | EM143088.1             | 88%         | 36%         |
| JGI1221J11331_10716531           | Sd  | Hypothetical protein [ <i>Paenibacillus</i> sp. UNC451MF]                        | WP_051621184.1         | 100%        | 52%         |
| 2001497724                       | Sd  | Pectin lyase fold protein [ <i>Formosa agariphila</i> KMM 3901]                  | CDF78719.1             | 99%         | 84%         |
| BS_KBA_SWE21_20mDRAFT_100399101  | Sd  | Hypothetical protein [ <i>Cellulophaga baltica</i> ]                             | WP_029445482.1         | 59%         | 41%         |
| KGI_S2_ANT05_2345mDRAFT_10666891 | Sd  | Hypothetical protein [ <i>Echinicola pacifica</i> ]                              | WP_018475002.1         | 100%        | 82%         |
| KGI_S2_ANT05_2345mDRAFT_10328402 | CC1_like superfamily  | Hypothetical protein [ <i>Formosa algae</i> ]                                    | WP_057779154.1         | 93%         | 75%         |
| KGI_S1_ANT03_95mDRAFT_10035671   | Sd  | Hypothetical protein [ <i>Polaribacter</i> sp. Hel1_85]                          | WP_036824725.1         | 99%         | 76%         |
| KGI_S1_ANT03_95mDRAFT_10236371   | Sd  | Hypothetical protein [ <i>Polaribacter</i> sp. Hel1_85]                          | WP_036824725.1         | 99%         | 71%         |
| BS_KBA_SWE12_2mDRAFT_101822181   | Sd  | Hypothetical protein [ <i>Saccharicrinis fermentans</i> ]                        | WP_027471338.1         | 83%         | 48%         |
| TDF_OR_ARG04_123mDRAFT_10107671  | NosD_copper_fam   | Hypothetical protein [ <i>Algibacter</i> sp. HZ22]                               | WP_054726015.1         | 100%        | 76%         |
| TDF_OR_ARG05_123mDRAFT_10138301  | Alginate_lyase (AlgLyase superfamily), NosD_copper_fam y Beta_helix | Hypothetical protein [ <i>Algibacter</i> sp. HZ22]                               | WP_054726015.1         | 48%         | 72%         |
| BS_KBA_SWE02_21mDRAFT_103946921  | Sd  | T9SS C-terminal target domain-containing protein [ <i>Tamlana sedimentorum</i> ] | WP_044631250.1         | 78%         | 54%         |

---

## **4. Discusión**

El ecosistema marino es considerado hoy en día uno de los recursos más importantes para la biodiversidad y, al mismo tiempo, es un ambiente poco estudiado o explotado en lo que respecta a sus genes, proteínas y enzimas (85). Es por ello, que este tipo de ambiente resulta adecuado para la bioprospección de biomoléculas con potenciales aplicaciones biotecnológicas.

Durante más de un siglo, la identificación y caracterización de microorganismos se ha llevado a cabo exclusivamente a través de técnicas de cultivo tradicionales. Estos estudios revelaron el papel clave que desempeñan los microorganismos en los ciclos geoquímicos y biorremediación, y como representantes de una fuente inagotable de nuevos compuestos bioactivos (86–88). Hoy en día, las colecciones de cultivos proporcionan cepas puras disponibles para la experimentación reproducible, como así también, sirven como depósito de microorganismos útil para comparaciones con otros organismos descritos recientemente (89). La información obtenida a partir de cepas aisladas en medios de cultivo tales como la caracterización de su fisiología, sigue siendo considerada como una herramienta importante para la bioprospección (86). A pesar de proporcionar conocimientos considerables sobre el mundo microbiano, las técnicas dependientes del cultivo tienen ciertas limitaciones, pudiéndose aplicar solo a organismos que pueden crecer en condiciones de laboratorio (90). Afortunadamente, nuevos métodos de cultivo se han desarrollado como alternativas para superar estas limitaciones (91). Estos métodos se basan fundamentalmente en los avances en cuanto al conocimiento biológico y ecológico de los organismos con el fin de representar de forma adecuada el entorno natural.

Otro gran avance para el estudio de los microorganismos son las nuevas tecnologías de secuenciación (12). Con el desarrollo de las tecnologías de secuenciación de alto rendimiento, el costo de la secuenciación de los genomas bacterianos se ha reducido enormemente. Los resultados de los genomas se obtienen en un lapso menor de tiempo para luego ser sometidos a un proceso de anotación automática (92, 93). Sin embargo, durante el proceso de anotación automática se pueden introducir errores o anotaciones genómicas de baja confiabilidad, por lo que sus resultados a menudo requieren una curación manual (92). La mayoría de los sistemas de anotación usan métodos de similitud entre la secuencia incógnita con respecto a genomas de referencia. Este mecanismo puede introducir y propagar errores. Sin embargo, existen bases de datos curadas con la finalidad de detectar y eliminar estos errores, aunque muchas veces, dada la gran cantidad de información generada, no resulta factible la cura manual de las anotaciones genómicas (92).

La caracterización de la biodiversidad microbiana marina ha sido especialmente difícil debido a la complejidad y la naturaleza dinámica de estos ambientes (86). Como consecuencia, los grupos bacterianos marinos están poco representados en las bases de datos y a su vez, existen desviaciones en cuanto a los grupos bacterianos estudiados. Iniciativas internacionales han

desarrollado estrategias para mejorar la representatividad de los genomas bacterianos, entre ellas se encuentra la base de datos de genomas en línea (“Genomes Online Database, GOLD”, <https://gold.jgi.doe.gov/>). GOLD constituye un sistema de gestión de datos para la catalogación y control continuo de proyectos de secuenciación de todo el mundo (94). Esta base de datos tiene como objetivo proporcionar el repositorio más completo de información sobre metadatos asociados a los proyectos de secuenciación, como así también proveer información sobre los organismos y sus ambientes (95).

A pesar de los esfuerzos realizados, gran parte de los microorganismos marinos con potencial interés biotecnológico pueden permanecer sin estudiar debido al hecho de que los medios de cultivo utilizados no resultan adecuados para su crecimiento (12). Por lo tanto, una de las principales limitaciones para conocer, y por lo tanto acceder a la gran diversidad microbiana son la dependencia de medios artificiales de cultivo y la falta de conocimiento de las interacciones microbianas, lo cual imposibilita el cultivo de muchos de los microorganismos presentes en una muestra ambiental. Se estima que sólo un 1 % de las especies microbianas que habitan el planeta se pueden cultivar y aislar en laboratorio utilizando métodos tradicionales (11). Esta pequeña fracción representa un mínimo de la diversidad genómica presente en un entorno determinado, y por lo tanto, del espectro potencial de genes novedosos con potenciales aplicaciones biotecnológicas (96). En el caso de los microorganismos marinos, el porcentaje de microorganismos cultivables es aún menor, dado que los ambientes marinos son especialmente difíciles de replicar en el laboratorio debido a sus características fisicoquímicas (97, 98). Como resultado de la baja proporción de microorganismos cultivables, se han desarrollado métodos alternativos para estudiar el potencial genético de las comunidades microbianas (99, 100). Un ejemplo son las estrategias metagenómicas, que representan una oportunidad para dilucidar más ampliamente el potencial biotecnológico inmerso en las comunidades microbianas (101). Sin embargo, las comunidades microbianas de ambientes marinos son extremadamente diversas, lo cual dificulta obtener información de los genomas de los microorganismos que integran la comunidad con una cobertura adecuada debido a la desigualdad en la distribución y riqueza de las especies microbianas (102). En consecuencia, a menudo la información obtenida es altamente fragmentada. Por lo tanto, ambas estrategias, genómica y metagenómica, brindan información complementaria sobre el potencial genético de las comunidades microbianas de ambientes marinos. Estas estrategias presentan ventajas y desventajas para la identificación de genes de importancia biotecnológica, las cuales deben ser consideradas al momento de utilizar la información depositada en las bases de datos de acceso público.

Las algas pardas son importantes productores primarios en los ecosistemas marinos, ya que juegan un papel ecológico análogo al de las plantas en los ambientes terrestres (102). La degradación de su biomasa depende en gran parte de los microorganismos marinos capaces de

utilizar componentes de la misma como fuente de carbono y energía (103). Los polisacáridos son importantes constituyentes de la materia orgánica producida por las algas en los ambientes marinos, constituyendo más de la mitad de su peso seco (104). Para su degradación, los microorganismos marinos producen enzimas extracelulares que depolimerizan distintos polisacáridos, por ejemplo hidrolasas o liasas específicas (104). Los productos de degradación son luego transportados al interior de la célula, donde enzimas citoplasmáticas intervienen en su procesamiento para su incorporación al metabolismo central.

Los fucoidanos y los fuco-oligosacáridos de algas pardas han sido objeto de numerosos estudios científicos debido a sus diversas funciones biológicas y potenciales propiedades terapéuticas. Si bien las investigaciones sobre los fucoidanos y sus posibles aplicaciones son extensas, la comprensión de los mecanismos y de las características estructurales particulares responsables de las diversas actividades biológicas es aún incompleta (36). Esto es debido principalmente a la heterogeneidad estructural que presentan estos polisacáridos y a la falta de conocimiento actual sobre enzimas que los catalicen (43). A pesar de que las bacterias capaces de hidrolizar moléculas de fucoidanos fueron descritas por primera vez en 1959 (105), en la actualidad se han descrito solo unas pocas fuentes de enzimas con actividad fucoidanasa. Los datos sobre la especificidad de fucoidanasas, el tipo de enlace glucosídico escindido, la relación entre la actividad catalítica y el grado de sulfatación del sustrato, es aún escaso (43). Se conocen cuatro secuencias de aminoácidos de enzimas fucoidanasas bacterianas, agrupadas en la familia CAZy GH107 [FcnA de *M. fucanivorans* SW5 (45), SVI\_0379 de *S. violacea* DSS12 y Fda1 y Fda2 del organismo *Alteromonas* sp. SN-1009 antes llamado "*Fucanobacter lyticus* SN-1009" (46)]. Por otro lado, las secuencias de aminoácidos de las enzimas pertenecientes al organismo *Flavobacterium* sp. SA-0082 (o también llamado *Fucobacter marina* SA-0082) se encuentran solo reportadas en una patente (46).

No se han identificado dominios o motivos conservados específicos de las secuencias de fucoidanasas (48). Además, las CAZymas son particularmente difíciles de anotar correctamente debido a que comparten módulos con otros tipos de enzimas (45, 106). Es probable que las enzimas fucoidanasas estén anotadas con otro nombre o como "proteína hipotética", como es el caso de muchas secuencias genómicas o metagenómicas (107-111). Por lo tanto, en este estudio se utilizó el dominio catalítico de las enzimas reportadas hasta el momento para la identificación de secuencias homólogas de fucoidanasas en genomas y metagenomas marinos utilizando el algoritmo blastp.

Las secuencias de aminoácidos de las enzimas fucoidanasas reportadas hasta el momento y utilizadas como referencia en este trabajo, resultaron ser muy diferentes entre sí. Las enzimas Fda1, Fda2 y SVI\_0379, que han sido identificadas en organismos que pertenecen al mismo

orden taxonómico, presentaron bajos porcentajes de identidad a nivel de aminoácidos. Situación similar ocurre con las enzimas FcnA y aquellas identificadas en el organismo *Flavobacterium* sp. SA-0082, estos organismos a pesar de pertenecer a la familia Flavobacteriaceae presentan porcentajes de identidad relativamente bajos (Tabla 1.3.). Este hecho podría deberse a las distintas especificidades que presentan las enzimas en cuanto al sustrato (43). La enzima FcnA escinde enlaces entre los residuos de  $\alpha$ -L-fucosa del tipo  $\alpha$ -(1 $\rightarrow$ 4), mientras que las enzimas Fda1, Fda2 y las dos enzimas del género *Flavobacterium* catalizan la ruptura de enlaces glucosídicos del tipo  $\alpha$ -(1 $\rightarrow$ 3) (43).

A pesar de que las enzimas pertenecientes al organismo *Flavobacterium* sp. SA-0082 no forman parte de la familia GH107, fue a partir del dominio catalítico de estas enzimas que se identificó el mayor número de secuencias homólogas, tanto en genomas bacterianos y como en metagenomas marinos. Sakai y colaboradores (112, 43) reportaron que las enzimas de *Flavobacterium* sp. SA-0082 presentan una actividad endo-fuoidanólítica liasa, siendo capaz de catalizar la escisión de distintos fuoidanos extraídos de algas pardas de la familia Laminarinaceae: *Kjellmaniella crassifolia*, *U. pinnatifida* y *Lessonia nigrescens*. Los autores determinaron que la enzima es altamente específica, catalizando únicamente la ruptura de las moléculas de fuoidanos y no de otros polisacáridos tales como alginatos, hialuronato, sulfatos de condroitina A, B y C y heparina.

El escaso número de secuencias identificadas en este estudio a partir de los módulos catalíticos de las enzimas de la familia GH107 indica que la misma no es actualmente abundante en las bases de datos de genomas y metagenomas. Si bien existen desvíos en estas bases de datos, la baja abundancia de estas secuencias sugiere que las enzimas pertenecientes a esta familia no serían las que predominarían en los procesos de asimilación de fuoidanos de algas pardas. Es importante aclarar que se utilizaron solamente los metagenomas ensamblados para la identificación de secuencias homólogas de fuoidanasas, y que a partir de los datos obtenidos se seleccionaron solamente las secuencias más largas (mayor a 100 aminoácidos). Por lo tanto, existe la posibilidad de que estas secuencias estén presentes en los ambientes estudiados, pero en poblaciones poco abundantes o poco homogéneas genéticamente, lo que podría resultar en bajos porcentajes de ensamblado de las lecturas obtenidas a partir de sus genomas (113). Dado que en muchos casos existen dentro de las CAZymas múltiples familias de enzimas que comparten un mismo sustrato, es posible que existan familias aún no identificadas de enzimas fuoidanasas (106). Por otra parte, también podrían participar en el proceso de depolimerización enzimas exo-hidrolíticas y sulfatasas específicas.

Dentro de las cepas bacterianas que contenían en sus genomas secuencias homólogas a la familia GH107 se incluyen dos cepas pertenecientes al género *Rhodopirellula* (ambas aisladas a

partir de sedimentos marinos) y una perteneciente al género *Colwellia*, cuyo genoma fue obtenido a partir de la estrategia “Genómica a partir de una célula individual” (Tabla 3.2.) aislada de las aguas costeras del Golfo de México. Como se mencionó anteriormente, otras enzimas que podrían estar implicadas en la depolimerización de los fucoidanos son las  $\alpha$ -L-fucosidasas (EC: 3.2.1.55) y las sulfatasas (EC: 3.1.6.-). Se ha reportado para miembros del phylum Planctomycetes la presencia de un alto contenido de genes que codifican sulfatasas dentro de su genoma, sugiriendo que podrían desempeñar un papel importante en la degradación de polisacáridos sulfatados (114). En particular, para la cepa *R. baltica* SH1 se han reportado 110 genes que codifican sulfatasas y un gen que codifica para la enzima  $\alpha$ -L-fucosidasa (114, 115). En el genoma de esta cepa se identificó una secuencia homóloga al dominio catalítico de las enzimas del organismo *Flavobacterium* sp. SA-0082. En el análisis del entorno genómico próximo a la secuencia homóloga identificada, se identificaron numerosos elementos que codifican para la enzima arilsulfatasa. Este resultado no se muestra en la sección de resultados ya que este análisis solo fue realizado para aquellos organismos cuyas secuencias homólogas presentaron los mayores porcentajes de identidad con respecto a las enzimas de referencia de acuerdo a la inferencia filogenética realizada en el apartado 3.1.1.2. Por otro lado, los resultados del análisis de contexto genómico en las cepas pertenecientes al mismo género (Figura 3.3.) indicaron la presencia de genes que se encontraban próximos a la secuencia homóloga que podrían codificar para la enzima  $\alpha$ -L-fucosidasa (señalada con el número 11), no ocurriendo así para genes que codifican para enzimas sulfatasas.

En general, el tipo de ecosistema más frecuente del que fueron aislados los organismos en los cuales se identificaron secuencias homólogas a fucoidanasas fue el ambiente marino, tanto en la columna de agua como en sedimentos (Tabla 3.2.). Este resultado pone de manifiesto la existencia de organismos bacterianos con posible capacidad fucoidanólítica en estos ambientes. Se han reportado diferencias en cuanto a las capacidades enzimáticas entre estos dos ambientes marinos que posiblemente se deban a diferencias en la composición de las comunidades microbianas (116). Otros estudios han reportado que en los sedimentos marinos los niveles de actividad degradativa de fucoidanos es mayor que en la columna de agua (117). Por otro lado, las enzimas que se encuentran adheridas a superficies (por ejemplo, en sedimentos) son menos activas pero se encuentran más protegidas a la degradación que aquellas enzimas presentes en la columna de agua (118). Es por esto que los sedimentos marinos son los entornos que a menudo se suelen estudiar para bioprospectar enzimas con nuevas aplicaciones biotecnológicas, ya que son reconocidos como depósitos de una gran diversidad microbiana y en su mayoría sin explorar (119). Este ambiente a su vez, es filogenéticamente más diverso que cualquier otro tipo de ambiente, incluyendo el suelo (120).

De acuerdo a la inferencia filogenética realizada (Figura 3.2.), todos los organismos en los cuales se identificaron secuencias homólogas con altos valores de identidad, pertenecen al phylum Bacteroidetes. Resultados publicados sobre proyectos de secuenciación de genomas de bacterias marinas pertenecientes a este phylum determinaron que la gran mayoría de estas bacterias presentan una maquinaria especializada para la degradación de polisacáridos y otros sustratos de alto peso molecular (121). Estos reportes son consistentes con los resultados del análisis de contexto genómico, en el cual se identificó un gran número de genes que posiblemente estén involucrados en el transporte, asimilación y metabolismo de polisacáridos de algas (Figura 3.4.), tales como sulfatasas,  $\alpha$ -L-fucosidasas, receptores dependientes de TonB, beta-lactamasa, acetil-esterasas,  $\alpha/\beta$  hidrolasa,  $\alpha$ -D-xilósido xilohidrolasa, plegamiento pectin liasa, xilan 1,4- $\beta$ -xilosidasa, pectato liasa, heparinasa II/III, alginato liasa,  $\alpha$ -glucosidasa y  $\beta$ -fructosidasa. Un ejemplo son las enzimas acetil-esterasas (EC: 3.1.1.6), que exhiben una organización estructural secundaria del tipo  $\alpha/\beta$  hidrolasa (122). Tanto los genes que codifican para enzimas acetil-esterasa (cataliza la hidrólisis de grupos acetil-éster) como aquellos relacionados con la determinación de la estructura de estas enzimas, fueron identificados muy próximos entre ellos y a la secuencia homóloga en el contexto genómico de la cepa *F. agariphila* KMM-3901 (Figura 3.4.).

Otro elemento identificado en las secuencias es el módulo de unión a carbohidratos (“Carbohydrate binding module”). Este módulo forma parte de enzimas que están involucradas en el metabolismo de carbohidratos tales como las enzimas glucósido hidrolasas (123). A su vez, otros reportes sugieren que estos motivos podrían estar relacionados con la degradación de polisacáridos, tal como se conoce en otras bacterias celulolíticas, ya que permiten que las enzimas activas sobre carbohidratos se unan de manera específica a sus sustratos (123). Los fragmentos genómicos que contenían este motivo se identificaron en un reducido número de organismos, entre ellos, *R. sallentina* SM41 identificada a partir del dominio catalítico de enzimas de la familia GH107, y en los organismos *Aquimarina* sp. SW150 y *F. agariphila* KMM-3901 identificadas a partir del dominio catalítico de las enzimas reportadas en la patente.

El análisis de sintenia compartida de los contextos genómicos de las secuencias más cercanamente relacionadas con la fucoidanasa del organismo *Flavobacterium* sp. SA-002 mostró un bajo grado de conservación, incluso entre organismos pertenecientes al mismo género (por ejemplo, *Aquimarina*). Este resultado sugiere la ocurrencia de rearrreglos genómicos tales como inserciones y deleciones. Sin embargo, el orden de estos genes parece estar más conservado en las especies *F. agariphila* y *C. fermentans* pertenecientes al phylum Bacteroidetes.

El amplio espectro de enzimas GH identificadas en las cepas *C. baltica* NN016038, *F. agariphila* KMM3901 y *P. saltans* DSM-12145 en la base de datos CAZy (Tabla 3.6.) sugiere que estos organismos presentan una alta especialización en la utilización de carbohidratos como fuente de carbono, lo cual es coincidente con los antecedentes reportados para organismos pertenecientes al phylum Bacteroidetes (121).

La secuencia homóloga identificada en el organismo *S. degradans* 2-40 (phylum Proteobacteria) a partir del dominio catalítico de las enzimas del género *Flavobacterium*, a pesar de no incluirse dentro del análisis del contexto genómico por presentar un porcentaje de identidad bajo, resultados sobre el potencial degradativo de carbohidratos reportados en la base de datos CAZy determinó la presencia de un amplio espectro de enzimas GH (Tabla 3.6.), sugiriendo una alta especialización en el uso de carbohidratos como fuente de carbono y energía. Si los fucoidanos representan parte de estos carbohidratos es incierto, debido al bajo porcentaje de similaridad que presentaba la secuencia homóloga identificada con aquellas utilizadas como referencia.

El número de secuencias identificadas en los metagenomas ensamblados obtenidos a partir de ambientes marinos fue mucho menor que el número identificado a partir de genomas bacterianos. Como se mencionó anteriormente, este resultado podría estar subestimado, dado que sólo se consideraron las secuencias identificadas con longitudes mayores a 100 aminoácidos de los metagenomas ensamblados. A excepción de una secuencia, todas las secuencias metagenómicas identificadas presentaron nombres de proteínas hipotéticas. Al igual que otras CAZymas, las enzimas fucoidanasas conocidas hasta el momento presentan una estructura modular. Esta característica, como así también la gran diversidad entre las secuencias de aminoácidos de las enzimas conocidas, hacen que la anotación automática para este grupo de enzimas sea un proceso dificultoso (124). De hecho, la clasificación en familias y subfamilias que utiliza la base de datos CAZy, basada en homología de secuencias, no es automática, sino que requiere de una curación manual (41). En el caso particular de las enzimas fucoidanasas, resulta imposible una asignación de nombre durante la anotación automática, dado que no existen COG o Pfam específicos para este tipo de enzimas. Durante la anotación automática en el sistema IMG/M, el nombre de una secuencia metagenómica se basa en búsquedas en estas dos bases de datos. El nombre es asignado cuando sobrepasa un valor umbral de longitud de alineamiento y de porcentaje de similaridad en búsquedas basadas en las bases de datos COG y Pfam (83).

Durante la anotación automática pueden ocurrir errores (110), por lo que es siempre importante estudiar en detalle las secuencias detectadas, aunque éstas presenten otro nombre. La estrategia utilizada para analizar las secuencias metagenómicas fue por medio del análisis DELTA-BLAST. A partir de este análisis se evaluó la presencia de motivos conservados en las

secuencias metagenómicas identificadas. El dominio conservado más frecuente en las búsquedas fue el “Beta hélix” (Pfam 13229). Aunque un número creciente de funciones de este motivo están siendo vinculadas con un tipo de plegamiento helicoidal (125), este dominio se asocia particularmente a proteínas que se unen a hidratos de carbono (126). Sin embargo, se ha reportado solo para las familias de GH 28, 49, 55, 82 y 87 la presencia de pliegues Beta-helicoidal en su estructura secundaria (127-129). Otro motivo conservado identificado para una de las secuencias fue “AmyAc\_family superfamily (Alpha-L-fucosidase)”. Se ha reportado que la enzima  $\alpha$ -L-fucosidasa, también llamada  $\alpha$ -fucosidasa (Pfam 01120), se clasifica en las familias GH29 y 95 (130). Otros dominios conservados identificados en este análisis, como así también en la información proporcionada por evidencia funcional en el sistema IMG/M, parecieron no tener precedentes en la literatura publicada hasta la fecha en relación a familias GH ó en cuanto a la participación en el metabolismo de carbohidratos, entre ellos: “NosD\_copper fam” y “CC1\_like superfamily”. Por último, se identificó el motivo conservado “Alginate lyase (AlgLyase superfamily)” que presenta el código Pfam PF05426 (Tabla 3.7. y Tabla 3.8.). Esta familia incluye enzimas que catalizan la depolimerización de alginatos, que al igual que los fucoidanos, son importantes constituyentes de la pared celular de muchas especies de algas pardas (131).

Dado que los productos génicos de todas las secuencias metagenómicas identificadas correspondieron a proteínas hipotéticas (a excepción de una de ellas que codificaba para la enzima alginato liasa), se realizó un análisis de similitud con respecto a las secuencias depositadas en GenBank. Este análisis no aportó mayor información, dado que la gran mayoría de las secuencias que presentaron los mayores porcentajes de cobertura e identidad también se trataban de proteínas hipotéticas (Tabla 3.8.). Además, el 40 % de las secuencias metagenómicas compartían menos del 60 % de identidad con las secuencias más cercanas de GenBank. Las secuencias que presentaron porcentajes de identidad superiores al 60 % correspondieron en su mayoría a cepas pertenecientes al género *Rhodopirellula*, apoyando la idea de que estos organismos son actores claves en el metabolismo de hidratos de carbono en los ambientes marinos, donde los polisacáridos representan los componentes dominantes de la biomasa generada (132). Varias cepas que se encontraban dentro de este grupo pertenecían al phylum Bacteroidetes: *Algibacter* sp. HZ22, *Saccharicrinis fermentans*, *Polaribacter* sp. Hel1\_85, *E. pacifica* y *Formosa algae*. Varias de estas cepas presentaron reportes sobre la presencia de genes codificando numerosas enzimas GH, coincidiendo con lo reportado para los miembros del phylum Bacteroidetes (41, 85). Otros organismos que contenían secuencias cercanamente relacionadas con secuencias metagenómicas son *S. violacea* (phylum Proteobacteria) y *M. fucanivorans* (phylum Bacteroidetes). Estas secuencias corresponden a las enzimas FcnA y SVI\_0379 pertenecientes a la familia GH107 y utilizadas como referencia. Sin

embargo, la secuencia identificada en el organismo *S. violacea* presentó un número de acceso a GenBank (WP\_013049664) distinto a aquel que presenta la enzima SVI\_0379 utilizada como referencia (BAJ00350) a pesar de corresponder al mismo genoma. La secuencia identificada a partir del análisis de DELTA-BLAST podría haberse originado a partir de una duplicación génica en el mismo organismo (133) o bien, podría tratarse de un error en la anotación génica en el sistema GenBank, en donde la misma secuencia de aminoácidos presenta dos entradas a la base de datos.

Al evaluar el grado de divergencia entre las secuencias metagenómicas identificadas en la base de datos IMG/M y aquellas secuencias depositadas en GenBank que presentaron los mayores porcentajes de identidad y cobertura con respecto a las primeras, se observó que la gran mayoría de las secuencias de GenBank correspondieron a secuencias que ya habían sido identificadas con anterioridad en la base de datos de genomas microbianos (IMG). El sistema IMG incluye todos los genomas microbianos depositados en GenBank, pero dado que los vuelve a anotar para conseguir un criterio más homogéneo, las secuencias toman otro nombre. Por lo tanto, las siguientes secuencias son las mismas: *Rhodopirellula* sp. SWK7 (número de acceso a GenBank EMI43088) y la secuencia genómica con número de acceso a IMG 2534780049, *A. latercula* (WP\_035084141) y la cepa *A. latercula* DSM-2041 (2523974135), *C. baltica* (WP\_029445482) y *C. baltica* NN016038 (2581797417), *E. pacifica* (WP\_018475002) y *E. pacifica* DSM-19836 (2515475267), *F. agariphila* KMM-3901 (CDF78719) y la secuencia de la misma cepa con número de acceso a IMG 2586113536, *R. europaea* (EMB13743) y la cepa *R. europaea* 6C (2534016826), y *R. europaea* (WP\_008665660) y *R. europaea* SH398 (2534019429). Por lo tanto, las secuencias metagenómicas y genómicas identificadas a partir de las secuencias de fucoidanasas conocidas están relacionadas.

### Consideraciones futuras

Este estudio ha revelado un número importante de secuencias homólogas a enzimas fucoidanasas en genomas de bacterias aisladas de diversos ambientes (marinos y suelo) y en metagenomas marinos, las cuales presentaron una alta diversidad. Los resultados de estos análisis sugieren que al menos parte de las secuencias homólogas identificadas podrían tratarse de genes que codifiquen para enzimas fucoidanasas. Algunas de estas enzimas podrían formar parte de la familia GH107, mientras que otras podrían constituir una nueva familia aún no definida. Dado el escaso conocimiento sobre la diversidad de las enzimas fucoidanasas, los análisis basados en similitud de secuencias sólo podrán identificar parte de las secuencias de interés. Otras estrategias, basadas en la detección de actividad enzimática, permitiría el descubrimiento de nuevas familias de fucoidanasas. Por ejemplo, se podrían aislar microorganismos en base a su capacidad para depolimerizar fucoidanos, y a partir de ellos

---

identificar las secuencias que codifican para estas enzimas utilizando métodos tradicionales o genómica. Alternativamente, se podría utilizar la estrategia de metagenómica funcional, por medio del análisis funcional de clones de una biblioteca metagenómica. Por otra parte, resulta necesario la expresión, purificación y caracterización enzimática en laboratorio a los efectos de comprobar la actividad fucoidanólítica de las secuencias identificadas, como así también determinar el tipo de enlace glucosídico que catalizan y el rango de sustratos sobre el cual actúan. Estos resultados proporcionarán información relevante sobre bacterias marinas capaces de catalizar la endo-hidrólisis de fucoidanos. Este conocimiento podría aplicarse para la producción de fuco-oligosacáridos, que podrán utilizarse para el desarrollo de productos farmacéuticos, nutracéuticos y/o cosméticos. Esta información a su vez, servirá de base para el desarrollo de nuevas aplicaciones comerciales de la biomasa generada por la macroalga invasora *U. pinnatifida* en nuestras costas.



## **5. Conclusiones**

A partir de los resultados de este estudio, se pueden concluir los siguientes puntos:

1. El algoritmo de búsqueda de alineamiento local (BLAST) resultó eficiente para la identificación de secuencias homólogas a aquellas que codifican para enzimas fucoidanasas en bases de datos públicas de genomas y metagenomas.
2. La mayor cantidad de secuencias homólogas a las enzimas fucoidanasas de referencia en genomas y metagenomas fueron identificadas a partir de los potenciales dominios catalíticos de las enzimas endo-fucoidanasa liasas identificadas en el organismo *Flavobacterium* sp. SA-0082.
3. El escaso número de secuencias identificadas a partir de los módulos catalíticos de las enzimas de la familia GH107 sugiere que la misma no sería abundante en los microorganismos de ambientes marinos. Las bacterias marinas fucoidanólíticas podrían estar utilizando mayormente otras enzimas en sus procesos de depolimerización de fucoidanos, tales como otras familias no homólogas de fucoidanasas,  $\alpha$ -L-fucosidasas, sulfatasas y/o acetil-esterasas.
4. El tipo de ecosistema más frecuente del que fueron aislados los organismos en los cuales se identificaron secuencias homólogas fue el ambiente marino tanto pelágico como en sedimentos, poniendo de manifiesto la existencia de poblaciones bacterianas con posible capacidad fucoidanólítica en estos ambientes.
5. El análisis de contexto genómico demostró un bajo nivel de conservación en las regiones adyacentes de las secuencias genómicas identificadas, incluso entre cepas de la misma especie, sugiriendo la ocurrencia de rearrreglos genómicos tales como inserciones y deleciones. En la gran mayoría de los contextos genómicos, las secuencias adyacentes a la fucoidanasa putativa se encontraban relacionadas con el ingreso y/o metabolismo de hidratos de carbono. Esto provee evidencia adicional sobre el posible rol de las secuencias identificadas en la degradación de polisacáridos.
6. El número de secuencias codificantes identificadas en metagenomas de ambientes marinos fue mucho menor que el número identificado a partir de genomas bacterianos. En estas secuencias se identificaron motivos que sugieren que las mismas podrían estar involucradas en la degradación de hidratos de carbono.
7. El grado de divergencia entre las secuencias metagenómicas y las más cercanas depositadas en la base de datos GenBank fue variable, con porcentajes de identidad medio a elevado para más de la mitad de las secuencias. Los organismos que contenían secuencias similares a las metagenómicas correspondieron en su mayoría a cepas del género *Rhodopirellula*. Cepas de este género también habían mostrado contener secuencias homólogas a fucoidanasas en sus genomas, apoyando la idea de que estos

---

microorganismos poco estudiados jugarían un rol importante en los procesos de asimilación de polisacáridos en los ambientes marinos.

8. La gran mayoría de las secuencias identificadas a partir del análisis DELTA-BLAST pertenecían a organismos cuyas secuencias ya habían sido identificadas en la base de datos de genomas microbianos, lo cual muestra la alta cobertura del análisis de este estudio, y refuerza la posible relación de las secuencias identificadas con la depolimerización de fucoidanos.
9. Las estrategias de búsqueda en genomas y metagenomas brindó información complementaria sobre el potencial de los microorganismos marinos para asimilar polisacáridos. Dado el escaso conocimiento actual existente sobre las enzimas fucoidanasas y los microorganismos que las producen, este estudio provee información novedosa que servirá de base para la expresión y caracterización de estas enzimas. El conocimiento generado se podrá aplicar para el desarrollo de nuevas aplicaciones comerciales de la biomasa de *Undaria pinnatifida* en nuestra región.

---

## **6. Bibliografía**

1. Huse, S. M., Dethlefsen, L., Huber, J. A., Welch, D., Relman, D. A. & Sogin, M. L. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLOS Genet* **4**, e1000255.
2. Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., Robeson, M., Edwards, R. A., Felts, B. & other authors. (2007). Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Appl Environ Microbiol* **73**, 7059–7066..
3. Brodie, E. L., DeSantis, T. Z., Parker, J. P. M., Zubietta, I. X., Piceno, Y. M. & Andersen, G. L. (2007). Urban aerosols harbor diverse and dynamic bacterial populations. *Proc Natl Acad Sci* **104**, 299–304.
4. Alonso-Saez, L. & Gasol, J. M. (2007). Seasonal variations in the contributions of different bacterial groups to the uptake of low-molecular-weight compounds in northwestern Mediterranean coastal waters. *Appl Environ Microbiol* **73**, 3528–3535.
5. De Pascale, D., De Santi, C., Fu, J. & Landfald, B. (2012). The microbial diversity of Polar environments is a fertile ground for bioprospecting. *Mar Genomics* **8**, 15–22.
6. Warnecke, F., Hess, M. (2009). A perspective: metatranscriptomics as a tool for the discovery of novel biocatalysts. *Biotechnol. J* **142**, 91-95.
7. Singh, B. K. (2010). Exploring microbial diversity for biotechnology: the way forward. *Trends Biotechnol* **28**, 111–6.
8. Rodríguez Martínez, J. M. (2004). Secuenciación de genomas. *Arbor* **177**, 285–310.
9. Pierce, B. (2012). *Genetics: A conceptual approach.*, 3rd ed. Editorial Médica Panamericana.
10. Ng, C. P. & Kirkness, E. F. (2010). Whole Genome Sequencing. In *Genet Var*, pp. 215–226. Edited by M. R. Barnes & G. Breen. Humana Press.
11. Handelsman, J., Liles, M., Mann, D., Riesenfeld, C. & Goodman, R. (2002). Cloning the metagenome: culture-independent access to the diversity and functions of the uncultivated microbial world. *J Microbiol Methods* **33**, 241–255.
12. Lozada, M. & Dionisi, H. (2015). Microbial bioprospecting in marine environments. In *Springer Handb Mar Biotechnol*, pp. 307–326. Edited by S.-K. Kim. Springer Berlin Heidelberg.
13. Gilbert, J. A. & Dupont, C. L. (2011). Microbial metagenomics : beyond the genome. *Ann Rev Mar Sci* **3**, 347-371.
14. Streit, W. & Daniel, R. (2010). *Metagenomics: Methods and protocols. Methods Mol Biol*, 668th edn. New York.
15. Shokralla, S., Spall, J., Gibson, J. & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* **21**, 1794–1805.

16. **Guazzaroni, M.-E., Beloqui, A., Golyshin, P. & Ferrer, M. (2009).** Metagenomics as a new technological tool to gain scientific knowledge. *World J Microbiol Biotechnol* **25**, 945–954.
17. **Teach the Microbiome.** <http://teachthemicrobiome.weebly.com/sequencing-the-microbiome.html>.
18. **Heidelberg, K., Gilbert, J. & Joint, I. (2010).** Marine genomics: at the interface of marine microbial ecology and biodiscovery. *Microb Biotechnol* **3**, 531–543.
19. **Stepanauskas, R. (2012).** Single cell genomics: an individual look at microbes. *Curr Opin Microbiol* **15**, 613–620.
20. **Robledo, D. (1998)** The seaweed resources of Mexico. In *Seaweed Resources of the World* (eds A. Critchley & M. Ohno), pp. 331–342. Japon.
21. **Synnytsya, A., Čopíková, J., Kim, W. J. & Park, Y. II. (2015).** Cell Wall Polysaccharides of Marine Algae. In *Mar Algal Biotechnol*, pp. 543–590. Edited by S. Kim. Springer Berlin Heidelberg.
22. **Arnosti, C., Bell, C., Moorhead, D. L., Sinsabaugh, R. L., Steen, a. D., Stromberger, M., Wallenstein, M. & Weintraub, M. N. (2013a).** Extracellular enzymes in terrestrial, freshwater, and marine environments: Perspectives on system variability and common research needs. *Biogeochemistry* **117**, 5–21.
23. **Weiss, M., Abele, U., Weckesser, J., Welte, W., Schiltz, E. & Schulz, G. (1991).** Molecular architecture and electrostatic properties of a bacterial porin. *Science* **254**, 1627–1630.
24. **Arnosti, C. (2014).** Patterns of microbially driven carbon cycling in the ocean: links between extracellular enzymes and microbial communities. *Adv Oceanogr* **2014**, 1–12.
25. **Michel, G., Nyval-Collen, P., Barbeyron, T., Czjzek, M. & Helbert, W. (2006).** Bioconversion of red seaweed galactans: a focus on bacterial agarases and carrageenases. *Appl Microbiol Biotechnol* **71**, 23–33.
26. **Elifantz, H., Malmstrom, R., Cottrell, M. & Kirchman, D. (2005).** Assimilation of polysaccharides and glucose by major bacterial groups in the Delaware Estuary. *Appl Environ Microbiol* **71**, 7799–7805.
27. **Elifantz, H., Dittel, A., Cottrell, M. & Kirchman, D. (2007).** Dissolved organicmatter assimilation by heterotrophic bacterial groups in the western Arctic Ocean. *Aquat Microb Ecol* **50**, 39–49.
28. **Kirchman, D., Moran, X. & Ducklow, H. (2009).** Microbial growth in the polar oceans—role of temperature and potential impact of climate change. *Nat Rev Microbiol* **7**, 451–459.
29. **Arnosti, C. & Steen, A. (2013b).** Patterns of extracellular enzyme activities and microbial metabolism in an Arctic fjord of Svalbard and in the northern Gulf of Mexico: contrasts in processing and fate of carbohydrates. *Front Microbiol* **4**, 318.
30. **Allison, S. D. (2005).** Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments. *Ecol Lett* **8**, 626–635.

31. Deniaud-Bouët, E., Kervarec, N., Michel, G., Tonon, T., Kloareg, B. & Hervé, C. (2014). Chemical and enzymatic fractionation of cell walls from Fucales: insights into the structure of the extracellular matrix of brown algae. *Ann Bot* **114**, 1203–16.
32. Black, W., Dewar, E. & Woodward, F. (1952). Manufacture of algal chemicals. IV. Laboratory-scale isolation of fucoidin from brown marine algae. *J Sci Food Agric* **3**, 122–9.
33. Holtkamp, von A.D. & Ruhr, M. an Der. (2009) *Isolation, Characterisation, Modification and Application of Fucoidan from Fucus Vesiculosus*. Von der Fakultät für Lebenswissenschaften der Technischen Universität Carolo Wilhelmina zu Braunschweig zur Erlangung des Grades einer. PhD Thesis.
34. Mabeau, S., Kloareg, B. & Joseleau, J. (1990). Fractionation and analysis of fucans from brown algae. *Phytochemistry* **29**, 2441–2445.
35. Honya, M., Mori, H., Anzai, M., Araki, Y. & Nisizawa, K. (1999). Monthly changes in the content of fucans, their constituent sugars and sulphate in cultured *Laminaria japonica*. *Hydrobiologia* **398**, 411–416.
36. Ale, M. T., Mikkelsen, J. D. & Meyer, A. S. (2011). Important determinants for fucoidan bioactivity: A critical review of structure-function relations and extraction methods for fucose-containing sulfated polysaccharides from brown seaweeds. *Mar Drugs* **9**, 2106–2130.
37. Ale, M. T. & Meyer, A. S. (2013). Fucoidans from brown seaweeds: an update on structures, extraction techniques and use of enzymes as tools for structural elucidation. *Rsc Adv* **3**, 8131–8141.
38. Descamps, V., Colin, S., Lahaye, M., Jam, M., Richard, C., Potin, P., Barbeyron, T., Yvin, J. C. & Kloareg, B. (2006). Isolation and culture of a marine bacterium degrading the sulfated fucans from marine brown algae. *Mar Biotechnol* **8**, 27–39.
39. Kusaykin, M., Bakunina, I., Sovo, V., Ermakova, S., Kuznetsova, T., Besednova, N., Zaporozhets, T. & Zvyagintseva, T. (2008). Structure, biological activity, and enzymatic transformation of fucoidans from the brown seaweeds. *Biotechnol J* **3**, 904–915.
40. Holtkamp, A. D., Kelly, S., Ulber, R. & Lang, S. (2009). Fucoidans and fucoidanases—focus on techniques for molecular structure elucidation and modification of marine polysaccharides. *Appl Microbiol Biotechnol* **82**, 1–11.
41. Lombard, V., Golaconda, R. H., Drula, E., Coutinho, P. M. & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acid Res* **42**, 490–495.
42. Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P. & Henrissat, B. (2010). A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem J* **432**, 437–444.
43. Kusaykin, M. I., Silchenko, A. S., Zakharenko, A. M. & Zvyagintseva, T. N. (2016). Fucoidanases. *Glycobiology*, **26**, 3-12.
44. Sakai, T., Ishizuka, K. & Kato, I. (2003a) Isolation and characterization of a fucoidan-degrading marine bacterium. *Mar Biotechnol*, **5**, 409–416.

- 
45. Colin, S., Deniaud, E., Jam, M., Descamps, V., Chevolot, Y., Kervarec, N., Yvin, J. C., Barbeyron, T., Michel, G. & Kloareg, B. (2006). Cloning and biochemical characterization of the fucanase FcnA: Definition of a novel glycoside hydrolase family specific for sulfated fucans. *Glycobiology* **16**, 1021–1032.
46. Takayama, M., Koyama, N., Sakai, T. & Kato, I. (2007). Enzymes capable of degrading a sulfated-fucose-containing polysaccharide and their encoding genes. *EP 1826272 A2*.
47. Gonzalez, J. A. & Dionisi, H. M. (2015). Análisis de secuencias de enzimas que hidrolizan fucoidanos, polisacáridos presentes en las algas pardas. *Taller Campo y Laboratorio. Univ Nac la Patagon San Juan Bosco*, Trelew.
48. Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G. & other authors. (2015). The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res* **43**, D213–D221.
49. Hemmingson, J. a., Falshaw, R., Furneaux, R. H. & Thompson, K. (2006). Structure and antiviral activity of the galactofucan sulfates extracted from *Undaria pinnatifida* (Phaeophyta). *J Appl Phycol* **18**, 185–193.
50. Maruyama, H., Tamauchi, H., Iizuka, M. & Nakano, T. (2006). The role of NK cells in antitumor activity of dietary fucoidan from *Undaria pinnatifida* sporophylls (Mekabu). *Planta Med* **72**, 1415–1417.
51. Kim, W. J., Kim, S. M., Lee, Y. H., Kim, H. G., Kim, H. K., Moon, S. H., Suh, H. H., Jang, K. H. & Park, Y. II. (2008). Isolation and characterization of marine bacterial strain degrading fucoidan from Korean *Undaria pinnatifida* sporophylls. *J Microbiol Biotechnol* **18**, 616–623.
52. Durand, E., Helley, D., Al Haj Zen, A., Dujols, C., Bruneval, P., Collic- Jouault, S., Fischer, A. M. & Lafont, A. (2008). Effect of low molecular weight fucoidan and low molecular weight heparin in a rabbit model of arterial thrombosis. *J Vasc Res* **45**, 529–537.
53. Wijesinghe, W. A. P. & Jeon, Y. J. (2012). Biological activities and potential industrial applications of fucose rich sulfated polysaccharides and fucoidans isolated from brown seaweeds: A review. *Carbohydr Polym* **88**, 13–20.
54. Yang, C., Chung, D., Shin, I. S., Lee, H., Kim, J., Lee, Y. & You, S. (2008). Effects of molecular weight and hydrolysis conditions on anticancer activity of fucoidans from sporophyll of *Undaria pinnatifida*. *Int J Biol Macromol* **43**, 433–437.
55. Berteau, O. & Mulloy, B. (2003). Sulfated fucans, fresh perspectives: Structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. *Glycobiology* **13**, 29–40.
56. Rodriguez Salvador Ojeda, M. (2008). *Efecto de algas marinas sobre la microbiota intestinal y el sistema inmune de ratones BALB/c*. Centro de investigaciones biológicas del noroeste, s.c. Tesis Doctoral.
57. Morya, V. K., Kim, J. & Kim, E. K. (2012). Algal fucoidan: Structural and size-dependent bioactivities and their perspectives. *Appl Microbiol Biotechnol* **93**, 71–82

58. Kim, W. J., Koo, Y.-K., Jung, M.-K., Moon, H. R., Kim, S. M., Synytsya, A., Yun-Choi, H. S., Kim, Y. S., Park, J. K. & Park, Y. II. (2010). Anticoagulating activities of low-molecular weight fuco-oligosaccharides prepared by enzymatic digestion of fucoidan from the sporophyll of Korean *Undaria pinnatifida*. *Arch Pharm Res* **33**, 125–31.
59. Wang, J., Wang, F., Zhang, Q., Zhang, Z., Shi, X. & Li, P. (2009). Synthesized different derivatives of low molecular fucoidan extracted from *Laminaria japonica* and their potential antioxidant activity in vitro. *Int J Biol Macromol* **44**, 379–384.
60. Choi, J., Kim, H. J., Kim, J. H., Byun, M., SooChun, B., HyunAhn, D., Hwang, Y., Kim, D., Kim, G. & Lee, J. (2009). Application of gamma irradiation for the enhanced physiological properties of polysaccharides from seaweeds. *Appl Radiat Isot* **67**, 1277–1281.
61. Athukorala, Y., Jung, W., Vasanthan, T. & Jeon, Y. (2006). An anticoagulative polysaccharide from an enzymatic hydrolysate of *Ecklonia cava*. *Carbohydr Polym* **66**, 184–191.
62. Kuznetsova, T., Besednova, N., Mamaev, A., Momot, A., Shevchenko, N. & Zvyagintseva, T. (2003). Anticoagulant activity of fucoidan from brown algae *Fucus evanescens* of the Okhotsk sea. *Bull Exp Biol Med* **136**, 471–473.
63. Silva, T., Alves, L., Queiroz, K., Santos, M., Marques, C., Chavante, S., Rocha, H. & Leite, E. (2005). Partial characterization and anticoagulant activity of a heterofucan from the brown seaweed *Padina gymnospora*. *Brazilian J Med Biol Res* 523–533.
64. Chevolot, L., Foucault, A. & Chauber, F. (1999). Further data on the structure of brown seaweed fucans: Relationships with anticoagulant activity. *Carbohydr Res* **319**, 154–165.
65. De Zoysa, M., Nikapitiya, C., Jeon, Y., Jee, Y. & Lee, J. (2008). Anticoagulant activity of sulfated polysaccharide isolated from brown seaweed *Sargassum fulvellum*. *J Appl Phycol* **20**, 67–74.
66. Yoon, S., Pyun, Y., Hwang, J. & Mourão, P. (2007). A sulfated fucan from the brown alga *Laminaria cichorioides* has mainly heparin cofactor II-dependent anticoagulant activity. *Carbohydr Res* **342**, 2326–2330.
67. Yoo, Y., Kim, W., Kim, S., Kim, S., Chung, M., Park, J., Suh, H., Lee, K. & Park, Y. (2007). Immunomodulating activity of a fucoidan isolated from Korean *Undaria pinnatifida* sporophyll. *Algae* **22**, 333–338.
68. Li, B., Lu, F., Wei, X. & Zhao, R. (2008). Fucoidan: Structure and bioactivity. *Molecules* **13**, 1671–1695.
69. Kang, S., Kim, K., Lee, S., Ahn, G., Cha, S., Kim, A., Yang, X.-D., Kang, M.-C. & Jeon, Y. (2011). Anti-inflammatory activity of polysaccharide purified from AMG-assistant extract of *Ecklonia cava* in LPS-stimulated RAW 264.7 macrophages. *Carbohydr Polym* **85**, 80–85.
70. Synytsya, A., Kim, W.-J., Kim, S.-M., Pohl, R., Synytsya, A., Kvasnička, F., Čopíková, J. & Park, Y. II. (2010a). Structure and antitumor activity of fucoidan isolated from sporophyll of Korean brown seaweed *Undaria pinnatifida*. *Carbohydr Polym* **81**, 41–48.

71. Alekseyenko, T., Zhanayeva, S., Venediktova, A., Zvyagintseva, T., Kuznetsova, T., Besednova, N. & Korolenko, T. (2007). Antitumor and antimetastatic activity of fucoidan, a sulfated polysaccharide isolated from the Okhotsk sea *Fucus evanescens* brown alga. *Bull Exp Biol Med* **143**, 730–732.
72. Shibata, H., Takagi, I., Nagaoka, M., Hashimoto, S., Aiyama, R., Iha, M., Ueyama, S. & Yokokura, T. (2000). Properties of fucoidan from *Cladosiphon okamuranus* TOKIDA in gastric mucosal protection. *BioFactors* **11**, 235–245.
73. Kim, W., Park, J., Park, J., Choi, D. & Park, Y. (2015). Purification and characterization of a fucoidanase (FNase S) from a Marine Bacterium *Sphingomonas paucimobilis* PF-1. *Mar Drugs* **13**, 4398–4417.
74. Casas, G., Scrosati, R. & Luz Piriz, M. (2004). The invasive kelp *Undaria pinnatifida* (Phaeophyceae, Laminariales) reduces native seaweed diversity in Nuevo Gulf (Patagonia, Argentina). *Biol Invasions* **6**, 411–416.
75. Irigoyen, A. J., Eyra, C. & Parma, A. M. (2010). Alien algae *Undaria pinnatifida* causes habitat loss for rocky reef fishes in north Patagonia. *Biol Invasions* **13**, 17–24.
76. Dellatorre, F. G., Amoroso, R., Saravia, J. & Orensanz, J. M. L. (2014). Rapid expansion and potential range of the invasive kelp *Undaria pinnatifida* in the Southwest Atlantic. *Aquat Invasions* **9**, 467–478..
77. Meretta, P., Matula, C. & Casas, G. (2012). Occurrence of the alien kelp *Undaria pinnatifida* (Laminariales, Phaeophyceae) in Mar del Plata, Argentina. *BioInvasions Rec* **1**, 59–63.
78. Madden, T. (2013). *The BLAST sequence analysis tool*, 2nd edn. Bethesda: The NCBI Handbook [Internet]. National Center for Biotechnology Information (US).
79. Boratyn, G. M., Schäffer, A. a, Agarwala, R., Altschul, S. F., Lipman, D. J. & Madden, T. L. (2012). Domain enhanced lookup time accelerated BLAST. *Biol Direct* **7**, 12.
80. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A. & other authors. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.
81. Tamura, K., Stecher, G., Peterson, D., Filipksi, A. & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
82. Zuckerkandl, E. & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In *Evol Genes Proteins*, pp. 97–166. Edited by V. Bryson & H. Vogel. New York: Academic Press.
83. Mavromatis, K., Chu, K., Ivanova, N., Hooper, S. D., Markowitz, V. M. & Kyrpides, N. C. (2009). Gene context analysis in the Integrated Microbial Genomes (IMG) data management system. *PLoS One* **4**.
84. Korb, O. E., Jensen, L., Von Mering, C. & Bork, P. (2004). Analysis of genomic context: prediction of functional associations from conserved bidirectionally transcribed gene pairs. *Nat Biotechnol* **22**, 911 – 917.

- 
85. Michel, G. & Czjzek, M. (2013). Polysaccharide -degrading enzymes from marine bacteria. In *Mar Enzym Biocatal sources, Biocatal Charact bioprocesses Mar Enzym*. Edited by A. Trincone. Elsevier.
86. Joint, I., Muhling, M. & Querellou, J. (2010). Culturing marine bacteria - an essential prerequisite for biodiscovery. *Microb Biotechnol* **3**, 564–575.
87. Van Hamme, J., Singh, A. & Ward, O. (2003). Recent advances in petroleum microbiology. *Microbiol Mol Biol Rev* **67**, 503–549.
88. Widada, J., Nojiri, H. & Omori, T. (2002). Recent developments in molecular techniques for identification and monitoring of xenobiotic-degrading bacteria and their catabolic genes in bioremediation. *Appl Environ Microbiol* **60**, 45–59.
89. Uruburu, F. (2003). History and services of culture collections. *Int Microbiol* **6**, 101–103.
90. Akondi, K. B. & Lakshmi, V. V. (2013). Emerging trends in genomic approaches for microbial cioprospecting. *Omi A J Integr Biol* **17**, 61–70.
91. Lewis, S. K., Epstein, A. D. & Ling, L. L. (2010). Un- cultured microorganisms as a source of secondary metabolites. *J Antibiot (Tokyo)* **63**, 468–476.
92. Richardson, E. J. & Watson, M. (2013). The automatic annotation of bacterial genomes. *Brief Bioinform* **14**, 1–12.
93. Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nat Rev Genet* **11**, 31–46.
94. Reddy, T. B. K., Thomas, A. D., Stamatis, D., Bertsch, J., Isbandi, M., Jansson, J., Mallajosyula, J., Pagani, I., Lobos, E. a. & Kyrpides, N. C. (2015). The Genomes OnLine Database (GOLD) v.5: A metadata management system based on a four level (meta)genome project classification. *Nucleic Acids Res* **43**, D1099–D1106.
95. Pagani, I., Liolios, K., Jansson, J., Chen, I. M. a, Smirnova, T., Nosrat, B., Markowitz, V. M. & Kyrpides, N. C. (2012). The Genomes OnLine Database (GOLD) v.4: Status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* **40**, 571–579.
96. Lee, H. S., Kwon, K. K., Kang, S. G., Cha, S. S., Kim, S. J. & Lee, J. H. (2010). Approaches for novel enzyme discovery from marine environments. *Curr Opin Biotechnol* **21**, 353–357.
97. Molinski, T., Dalisay, D., Lievens, S. & Saludes, J. (2009). Drug development from marine natural products. *Nat Rev Drug Discov* **8**, 69–85.
98. Krause, J. & Tobin, G. (2013). Discovery, development, and regulation of natural products. In *Using Old Solut to New Probl Drug Discov*, 21st edn., pp. 3–35. Edited by M. Kulka. Rijeka, Croatia: Tech.
99. Beloqui, A., de Maria, P., Golyshin, P. & Ferrer, M. (2008). Recent trends in industrial microbiology. *Curr Opin Microbiol* **11**, 240–248.

100. Gabor, E., de Vries, E. & Janssen, D. (2004). Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases. *Environ Microbiol* **6**, 948–958.
101. Steward, G. & Rappe, M. (2007). What's the 'meta' with metagenomics? *Isme J* **1**, 100–102.
102. Edwards, J. L., Smith, D. L., Connolly, J., McDonald, J. E., Cox, M. J., Joint, I., Edwards, C. & McCarthy, A. J. (2010). Identification of carbohydrate metabolism genes in the metagenome of a marine biofilm community shown to be dominated by Gammaproteobacteria and Bacteroidetes. *Genes (Basel)* **1**, 371–384
103. Warren, R. (1996). Microbial hydrolysis of polysaccharides. *Annu Rev Microbiol* **50**, 183–212.
104. Benner, R., Pakulski, J., Mccarthy, M., Hedges, J. & Hatcher, P. (1992). Bulk chemical characteristics of dissolved organic-matter in the ocean. *Science* **255**, 1561–1564.
105. Yaphe, W. & Morgan, K. (1959). Enzymic hydrolysis of fucoidin by *Pseudomonas atlantica* and *Pseudomonas carrageenovora*. *Nature* **183**, 761–762.
106. Cantarel, B. I., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res* **37**, 233–238.
107. Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
108. Gomez-Alvarez, V., Teal, T. & Schmidt, T. (2009). Systematic artifacts in metagenomes from complex microbial communities. *Isme J* **3**, 1314–1317.
109. Raes, J., Foerstner, K. & Bork, P. (2007). Get the most out of your metagenome: computational analysis of environmental sequence data. *Curr Opin Microbiol* **10**, 490–498.
110. Roberts, R. (2004). Identifying protein function a call for community action. *PLoS Biol* **2**, E42.
111. Sleator, R. & Walsh, P. (2010). An overview of in silico protein function prediction. *Arch Microbiol* **192**, 151–155
112. Sakai, T., Kimura, H., Kojima, K., Shimanaka, K., Ikai, K. & Kato, I. (2003b) Marine bacterial sulfated fucoglucuronomannan (SFGM) lyase digests brown algal SFGM into trisaccharides. *Mar Biotechnol*, **5**, 70–78.
113. Nagarajan, N. & Pop, M. (2013). Sequence assembly demystified. *Nat Rev Genet* **14**, 157–167.
114. Wegner, C. E., Richter-Heitmann, T., Klindworth, A., Klockow, C., Richter, M., Achstetter, T., Glöckner, F. O. & Harder, J. (2013). Expression of sulfatases in

- Rhodopirellula baltica* and the diversity of sulfatases in the genus *Rhodopirellula*. *Mar Genomics* **9**, 51–61. Elsevier B.V.
115. Glöckner, F. O., Kube, M., Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Beck, a, Borzym, K. & other authors. (2003). Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc Natl Acad Sci U S A* **100**, 8298–8303.
116. Arnosti, C. (2003). Microbial extracellular enzymes and their role in dissolved organic matter cycling. In *Aquat Ecosyst interactivity dissolved Org matter*, pp. 316–342. Edited by S. Findlay & R. Sinsabaugh. San Diego: Aquatic Ecology Series. Academic Press.
117. Arnosti, C. (2008). Functional differences between Arctic seawater and sedimentary microbial communities: Contrasts in microbial hydrolysis of complex substrates. *FEMS Microbiol Ecol* **66**, 343–351.
118. Wetzel, R. (1993). Humic compounds from wetlands: complexation, inactivation and reactivation of surface-bound and extracellular enzymes. *Verh Int Verein fur Liminol* **25**, 122–128.
119. Dionisi, H. M., Lozada, M. & Olivera, N. L. (2012). Bioprospection of marine microorganisms: biotechnological applications and methods. *Rev Argent Microbiol* **44**, 49–60.
120. Lozupone, C. A. & Knight, R. (2007). Global patterns in bacterial diversity. In *Proc Natl Acad Sci*, pp. 104:11436–40.
121. Gómez Pereira, P. (2010). *Marine Bacteroidetes: distribution patterns and role in the degradation of organic matter*. Doctoral dissertation.
122. Ollis, D., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S., Harel, M., Remington, S., Silman, I. & other authors. (1992). The alpha/beta-hydrolase fold. *Protein Eng* **5**, 197–211
123. Aurilia, V., Parracino, A. & D'Auria, S. (2008). Microbial carbohydrate esterases in cold adapted environments. *Gene* **410**, 234–240.
124. Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F. & Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* **40**, W445–W451.
125. Leinälä, E., Davies, P. & Jia, Z. (2002). Crystal structure of beta-helical antifreeze protein points to a general ice binding model. *Structure* **10**, 619–627.
126. Jenkins, J., Mayans, O. & Pickersgill, R. (1998). Structure and evolution of parallel beta-helix proteins. *J Struct Biol* **122**, 236–246.
127. Petersen, T., Christgau, S., Kofod, L., Kauppinen, S., Dalbøge, H., Johnson, A. & Larsen, S. (1997). Crystallization and preliminary X-ray studies of rhamnogalacturonase A from *Aspergillus aculeatus*. *Acta Crystallogr Sect D Biol Crystallogr* **53**, 105–107.
128. Michel, G., Chantalat, L., Fanchon, E., Henrissat, B., Kloareg, B. & Dideberg, O. (2001). The ι-Carrageenase of *Alteromonas fortis* A β-helix fold-containing enzyme for the degradation of a highly polyanionic polysaccharide. *J Biol Chem* **276**, 40202–40209.

- 
- 129. Rigden, D. J. & Franco, O. L. (2002).**  $\beta$ -Helical catalytic domains in glycoside hydrolase families 49, 55 and 87: Domain architecture, modelling and assignment of catalytic residues. *FEBS Lett* **530**, 225–232.
- 130. Sakurama, H., Tsutsumi, E., Ashida, H., Katayama, T., Yamamoto, K. & Kumagai, H. (2012).** Differences in the substrate specificities and active-site structures of two L-Fucosidases (Glycoside Hydrolase Family 29) from *Bacteroides thetaiotaomicron*. *Biosc Biotech Biochem* **76**, 1022–1024.
- 131. Rehm, B. (2009).** *Alginates: biology and applications*, 13th edn. Springer.
- 132. Gade, D., Gobom, J. & Rabus, R. (2005).** Proteomic analysis of carbohydrate catabolism and regulation in the marine bacterium *Rhodopirellula baltica*. *Proteomics* **5**, 3672–3683.
- 133. Gevers, D., Vandepoele, K., Simillion, C., Van de Peer, Y. (2004).** Gene duplication and biased functional retention of paralogs in bacterial genomes. *Trends in Microbiol* **12**, 148-154.