



**Centro de Investigación y de Estudios Avanzados
del Instituto Politécnico Nacional
Unidad Mérida
Departamento de Recursos del Mar**

**Evaluación del potencial metabólico para la síntesis
microbiana de polihidroxicanoatos en el suroeste del Golfo
de México, mediante un análisis metagenómico**

Tesis que presenta:

LBG. Andrea Zambrano Arizpe

Para obtener el grado de:

Maestra en Ciencias

En la especialidad de:

Biología Marina

Director de la Tesis:

Dr. José Quinatzin García Maldonado

Mérida, Yucatán

Septiembre de 2021

1 RESUMEN

Los polihidroxicanoatos (PHAs) son polímeros lipídicos intracelulares con propiedades fisicoquímicas similares a los de origen petroquímico, que funcionan como reserva de energía en algunas bacterias y arqueas. Estudios previos relacionados con el metabolismo de PHAs han estado principalmente dirigidos hacia bacterias cultivables. Sin embargo, trabajos recientes han evidenciado una gran diversidad de metabolismos en microorganismos no cultivables. Por lo tanto, el objetivo de este trabajo fue explorar el potencial metabólico para la producción de polihidroxicanoatos en muestras ambientales marinas y en una comunidad enriquecida degradadora de hidrocarburos, aislada previamente del suroeste del Golfo de México, a través de herramientas bioinformáticas. Se logró identificar y ensamblar el 82.7% del genoma de una bacteria no cultivable perteneciente al Orden Dehalococcoidales, a partir del metagenoma de una comunidad bentónica, la cual presentó la enzima PHA sintasa tipo I, involucrada en la síntesis de poli(3-hidroxibutirato). Además, se encontró que los sustratos potenciales para la acumulación de este biopolímero pueden estar relacionados con el catabolismo de carbohidratos, y con el de hidrocarburos. Por otro lado, la predicción funcional basada en secuencias parciales del gen 16S ARNr de la comunidad microbiana enriquecida, evidenció la presencia de las enzimas requeridas para la síntesis de polihidroxicanoatos de cadena corta y, se infirió que la mayor síntesis de PHAs se lleva a cabo a las 72 horas de incubación en presencia de petróleo ligero. Los resultados de este trabajo exhiben el potencial metabólico que tienen algunos microorganismos bentónicos del suroeste del Golfo de México, para la producción de polihidroxicanoatos con propiedades fisicoquímicas de interés biotecnológico que tendrán que ser evaluadas posteriormente en experimentos de laboratorio.

2 ABSTRACT

Polyhydroxyalkanoates (PHAs) are intracellular lipid polymers with physicochemical properties similar to that of petrochemical origin, which work as an energy reserve in some bacteria and archaea. Previous studies related to the metabolism of PHAs have been mainly directed towards culturable bacteria. However, recent works have shown a great diversity of metabolisms in non-culturable microorganisms. Therefore, the objective of this work was to explore the metabolic potential to produce polyhydroxyalkanoates in marine environmental samples and an enriched hydrocarbon-degrading community, previously isolated from the southwestern Gulf of Mexico, through bioinformatics tools. It was possible to identify and assemble 82.7% of the genome of a non-culturable bacterium belonging to the Order Dehalococcoidales, from the metagenome of a benthic community, which presented the enzyme PHA synthase type I, involved in the synthesis of poly (3-hydroxybutyrate). In addition, it was found that the potential substrates for the accumulation of this biopolymer may be related to the catabolism of carbohydrates, and hydrocarbons. On the other hand, the functional prediction based on partial sequences of the 16S rRNA gene of the enriched microbial community evidenced the presence of the enzymes required for the synthesis of short-chain polyhydroxyalkanoates and, it was inferred that the highest synthesis of PHAs is carried out at 72 hours of incubation in the presence of light oil. The results of this work show the metabolic potential of some benthic microorganisms from the southwest of the Gulf of Mexico, for the production of polyhydroxyalkanoates with physicochemical properties of biotechnological interest that will have to be evaluated later in laboratory experiments.

« La vie n'est facile pour aucun de nous. Mais quoi, il faut avoir de la persévérance, et surtout de la confiance en soi. Il faut croire que l'on est doué pour quelque chose, et que, cette chose, il faut l'atteindre coûte que coûte »

- Marie Curie

3 AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología por la beca de manutención de la Convocatoria Becas Nacional (Tradicional) 2019 – 2 con número 1007358.

Esta investigación fue apoyada por el Fondo Sectorial CONACyT - SENER - Hidrocarburos, proyecto 201441.

Al Dr. Emanuel Hernández Núñez, por los recursos computacionales proporcionados para realizar los análisis bioinformáticos. Así mismo, por compartirme todos sus conocimientos.

Al Dr. Daniel Cerqueda García, por toda su dedicación y apoyo en los análisis bioinformáticos y por sus consejos en la redacción y revisión de la tesis. Pero, sobre todo, muchas gracias por toda su paciencia durante el desarrollo de este proyecto.

A Edgar Iván Morales Cruz, por alentarme a seguir mis sueños y nunca dejar que me rindiera, por apoyarme tanto en los buenos y malos momentos, por ser la persona que creyó y confió en mí ciegamente.

Al Dr. José García Maldonado, por darme la oportunidad de poder ingresar al Laboratorio de Microbiología y por sus consejos para el desarrollo de esta tesis.

A la Dra. Leopoldina Aguirre Macedo y Dr. Alejandro López Cortés por su ayuda en mi formación académica.

4 TABLA DE CONTENIDO

1	RESUMEN	II
2	ABSTRACT	III
3	AGRADECIMIENTOS	V
4	TABLA DE CONTENIDO.....	VI
5	LISTA DE FIGURAS	VIII
6	LISTA DE TABLAS.....	IX
7	DEFINICIONES DE ABREVIATURAS	X
8	LISTA DE UNIDADES.....	XIII
9	INTRODUCCION	1
10	MARCO TEÓRICO.....	3
10.1	POLIHIDROXIALCANOATOS	3
10.1.1	<i>Definición y función.....</i>	3
10.1.2	<i>Propiedades fisicoquímicas y aplicaciones de los PHA.....</i>	4
10.1.3	<i>Producción actual de bioplásticos.....</i>	5
10.1.4	<i>Rutas metabólicas</i>	6
10.2	PRODUCCIÓN DE POLIHIDROXIALCANOATOS A PARTIR DE HIDROCARBUROS	8
10.3	METAGENOMAS DE COMUNIDADES MICROBIANAS MARINAS	9
11	ANTECEDENTES	11
12	JUSTIFICACIÓN	15
13	PREGUNTAS DE INVESTIGACIÓN.....	16
14	OBJETIVO GENERAL.....	17
14.1	OBJETIVOS ESPECÍFICOS	17
15	MATERIALES Y MÉTODOS	18
15.1	ENSAMBLE DE LOS METAGENOMAS Y RECONSTRUCCIÓN DE LOS GENOMAS INDIVIDUALES.....	18
15.2	DETECCIÓN DE PHA SINTASAS EN LOS METAGENOMAS	19
15.3	IDENTIFICACIÓN TAXONÓMICA DE BACTERIAS PRODUCTORAS DE POLIHIDROXIALCANOATOS A PARTIR DE METAGENOMAS	20
15.4	RECONSTRUCCIÓN METABÓLICA DE PHAZGOM18.....	20

15.5	PREDICCIÓN DEL POTENCIAL METABÓLICO DE UN CONSORCIO MICROBIANO MARINO AISLADO DE SEDIMENTOS DEL SUROESTE DEL GOLFO DE MÉXICO	21
15.5.1	<i>PICRUS12</i>	21
16	RESULTADOS	23
16.1	EXPLORACIÓN DEL POTENCIAL METABÓLICO PARA LA SÍNTESIS DE POLIHIDROXIALCANOATOS A PARTIR DE METAGENOMAS DEL SUROESTE DEL GOLFO DE MÉXICO.....	23
16.2	POTENCIAL METABÓLICO PARA LA PRODUCCIÓN DE POLIHIDROXIALCANOATOS POR UN CONSORCIO MICROBIANO DEGRADADOR DE HIDROCARBUROS	29
17	DISCUSIÓN	36
17.1	EXPLORACIÓN DEL POTENCIAL METABÓLICO PARA LA SÍNTESIS DE POLIHIDROXIALCANOATOS A PARTIR DE METAGENOMAS DEL SUROESTE DEL GOLFO DE MÉXICO.....	36
17.2	POTENCIAL METABÓLICO PARA LA PRODUCCIÓN DE POLIHIDROXIALCANOATOS POR UN CONSORCIO MICROBIANO DEGRADADOR DE HIDROCARBUROS.....	39
18	CONCLUSIONES.....	43
19	LITERATURA CITADA.....	44
20	ANEXOS	56
20.1	MANUAL DE ENSAMBLE DE METAGENOMAS Y BÚSQUEDA DE ENZIMAS CLAVE.	56

5 LISTA DE FIGURAS

Figura 1: Ruta metabólica para la síntesis de polihidroxicanoatos de cadena corta. Tomado y adaptado de Sudesh <i>et al.</i> , (2000).	6
Figura 2: Ruta metabólica para la síntesis de polihidroxicanoatos de cadena mediana. Tomado y adaptado de (Kim <i>et al.</i> , 2007).	8
Figura 3: Degradación de alcanos hasta ácido carboxílico. Se muestra como ingresa el ácido carboxílico a la vía de β -oxidación para la obtención de precursores involucrados en la síntesis de polihidroxicanoatos. Tomado y adaptado de Hernández (2015).	9
Figura 4: Árbol filogenético que muestra las relaciones de parentesco del genoma ensamblado a partir del metagenoma de sedimento de la estación D18 del Golfo de México, llamado PHAZGoM18 y resaltado en color rojo. El análisis estuvo basado en el principio de máxima verosimilitud (Maximum likelihood). Los valores de soporte se calculan con la prueba Shimodaira-Hasegawa.	25
Figura 5: Visualización gráfica de genoma PHAZGoM18. Las flechas indican la posición de las enzimas PHA sintasas localizadas. En el interior se refleja la cadena antisentido, mientras que, en el exterior la cadena sentido.	26
Figura 6: PHA sintasas detectadas en el genoma de PHAZGoM18. A) se observa la ubicación del gen <i>phaC</i> . B) indica la posición de los genes <i>phaE_1</i> y <i>phaE_2</i>	27
Figura 7: Evaluación funcional de PHAZGoM18. Resumen de las funciones conocidas presentes en el genoma anotado con RASTtk de PHAZGoM18. Mapa de calor de abundancia relativa en una escala blanco-azul.	28
Figura 8: Reconstrucción metabólica involucrada en la producción de polihidroxicanoatos por PHAZGoM18.	29

6 LISTA DE TABLAS

Tabla 1: PHA sintasas localizadas en el metagenoma de sedimento del sitio de muestreo D18-SED del suroeste del Golfo de México.	23
Tabla 2: Principales rutas metabólicas predichas por PICRUST2 a partir secuencias del 16S ARNr del consorcio microbiano proveniente de sedimentos del sur de Golfo de México creciendo con petróleo ligero como única fuente de carbono. Los valores de abundancia se muestran en porcentaje.	32
Tabla 3: Principales enzimas predichas por PICRUST2 a partir de la secuenciación del 16S ARNr del consorcio microbiano proveniente de sedimentos del sur de Golfo de México creciendo con petróleo ligero como única fuente de carbono. Los valores de abundancia se muestran en porcentaje.	33
Tabla 4: Principales enzimas predichas por PICRUST2 a partir de la secuenciación del 16S ARNr del consorcio microbiano proveniente de sedimentos del sur del Golfo de México creciendo con petróleo ligero como única fuente de carbono. Los valores de abundancia se muestran en porcentaje.	34
Tabla 5: Enzimas involucradas en la producción de PHAs predichas por PICRUST2 a partir de la secuenciación del 16S ARNr del consorcio microbiano proveniente de sedimentos del sur del Golfo de México creciendo con petróleo ligero como púnica fuente de carbono. Los valores de abundancia se muestran en porcentaje.	35

7 DEFINICIONES DE ABREVIATURAS

¹H RMN	Resonancia Magnética Nuclear de Protón
3HD	3-hidroxicanoato
3HDD	3-hidroxicanoato
3HHD	3-hidroxiheptadecanoato
3HHx	3-hidroxiheptanoato
3HO	3-hidroxiheptanoato
3HTD	3-hidroxitetradecanoato
3HV	3-hidroxiheptanoato
4HB	4-hidroxiheptanoato
ACP	Proteína Portadora de Acilo
ARNr	Ácido ribonucleico ribosomal
ASV	Amplicon Sequence Variant
ATCC	American Type Culture Collection
BLAST	Basic Local Alignment Search Tool
BRENDA	Braunschweig Enzyme Database
BTEX	Benceno, Tolueno, Etilbenceno y Xileno
CoA	Coenzima A
COG	Grupos Ortólogos de Proteínas
Cys	Cisteína
DADA2	Divisive Amplicon Denoising Algorithm
DNA	Ácido Desoxirribonucleico

DSC	Calorimetría Diferencial de Barrido
E.C	Comisión enzimática
FTIR	Espectrometría Infrarroja con Transformada de Fourier
G+C (mol%)	Contenido, por ciento mol Guanina-Citosina
GC-MS	Cromatografía de Gases acoplada a Espectrometría de Masas
GPC	Cromatografía de Permeación en Gel
GTDB-Tk	Genome Taxonomy Database Toolkit
HDPE	Poliétileno de alta densidad
KEGG	Kyoto Encyclopedia of Genes and Genomes
LDPE	Poliétileno de bajad densidad
NCBI	National Center for Biotechnology Information
NGS	Next-Generation Sequencing
P3H4MV	Poli(3-hidroxi-4-metilvalerato)
P3HB	Poli(3-hidroxibutirato)
P3HHx	Poli(3-hidroxihexanoato)
P3HV	Poli(3-hidroxivalerato)
P4HB	Poli(4-hidroxibutirato)
P5HV	Poli(5-hidroxivalerato)
pb	pares de bases
PCR	Reacción en Cadena de la Polimerasa
Pfam	Protein family (Base de datos)
PHA	Polihidroxialcanoato
PHA mcl	Polihidroxialcanoatos de cadena mediana

PHA scl	Polihidroxicanoatos de cadena corta
PICRUSt2	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
PRODIGAL	Prokaryotic dynamic programming Gene-finding algorithm
PVC	Policloruro de vinilo
QIIME2	Quantitative Insights Into Microbial Ecology
RASTtk	Rapid Annotation using Subsystem Technology
RefSeq	Reference Sequence (Base de datos)
RMN	Resonancia Magnética Nuclear
rxn	Reacción
Tg	Temperatura de transición vítrea
TGA	Análisis termogravimétrico
Tm	Temperatura de fusión
UV	Ultravioleta
WGA DNA	Whole Genome Amplication
XRD	Difracción de rayos X

8 LISTA DE UNIDADES

g	Gramos
°C	Grados centígrados
%	Porcentaje
cm	Centímetro
Da	Dalton
kDa	Kilodalton
km	Kilometro
l	Litros
mg	Miligramo
µg	Microgramo
µm	Micrómetro

9 INTRODUCCION

La producción global de los plásticos sintéticos de origen petroquímico, se incrementó significativamente en las pasadas décadas de 1.7 millones de toneladas en 1950 (Auta *et al.*, 2017) a 348 millones de toneladas en el año 2017 (PlasticsEurope, 2018). Sin embargo, el gran aumento en su aplicación en la vida diaria no ha sido acompañada de un buen manejo para su disposición o degradación, estimando que el plástico representa cerca del 54% de los desechos antropogénicos en el ambiente (Horton *et al.*, 2017).

Los plásticos de origen petroquímico, son considerados como contaminantes recalcitrantes debido a la ausencia de grupos funcionales favorables, hidrofobicidad, cristalinidad, alto peso molecular, por lo tanto, no pueden ser asimilados por la mayoría de los microorganismos (Urbanek *et al.*, 2018). Como consecuencia, los plásticos están presentes en los ecosistemas terrestres, acuáticos, estuarios, costeros, marinos, zonas urbanizadas, incluso, se ha encontrado en sedimentos del fondo marino y encapsulado en el hielo marino del Ártico (Horton *et al.*, 2017). Aunque se han llevado a cabo estudios sobre la degradación de estos polímeros sintéticos a partir de bacterias, estos no han tenido mucho éxito. En un estudio realizado por Kumari *et al.*, (2019) indicaron que el porcentaje de degradación fue de solamente 0.26, 0.96 y 1% para PVC, LDPE y HDPE en 90 días de incubación con *Bacillus* sp. Por lo tanto, la búsqueda de alternativas a este problema y a los efectos tóxicos que causan los plásticos en el ambiente ha mostrado un gran interés en el desarrollo de plásticos biodegradables, como los polihidroxicanoatos.

Los polihidroxicanoatos son un grupo de poliésteres producidos por una amplia variedad de microorganismos. Al contrario de los plásticos sintéticos, son biodegradables y biocompatibles con el ambiente (Jiang *et al.*, 2016). Además, presentan propiedades mecánicas similares a los polímeros sintéticos, como el polipropileno, por lo que representan un alto potencial comercial con aplicación en la industria, medicina, farmacia y agricultura (Madkour *et al.*, 2013).

El conocimiento del metabolismo de bioplásticos se ha restringido a bacterias cultivables y productoras de polihidroxicanoatos, en donde se realiza una detección

fenotípica de los gránulos con Rojo Nilo, Negro Sudán y Azul Nilo (Vahlis, 2017). Este método clásico para la búsqueda de nuevas enzimas se basa principalmente en el aislamiento, cultivo y crecimiento de microorganismos bajo diferentes condiciones (Cheema *et al.*, 2011). Sin embargo, se estima que solamente el 1% de los microorganismos ambientales pueden ser cultivados (Alma'abadi *et al.*, 2015). El número total estimado de células microbianas en la Tierra es alrededor de 10^{30} . Por lo tanto, la diversidad de los procariotas es única y representa un reservorio genético no explorado que puede ser explotado para la búsqueda de nuevas enzimas para obtener mejores bioplásticos con mejores propiedades y con una amplia variedad de sustratos (Cheema *et al.*, 2011; Ngara & Zhang, 2018).

La metagenómica ha sido eficazmente utilizada para identificar nuevas PHA sintasas y genes involucrados en el metabolismo de polihidroxialcanoatos, evitando las limitaciones causadas por los métodos de cultivo tradicionales. Estos enfoques facilitan el estudio tanto de bacterias cultivables y no cultivables a partir de una gran cantidad de nichos ecológicos (Vahlis, 2017).

10 MARCO TEÓRICO

10.1 Polihidroxicanoatos

10.1.1 Definición y función

Los polihidroxicanoatos son polímeros de ácidos hidroxicanoicos conectados por un enlace éster, razón por la cual también son conocidos como biopolíesteres (González *et al.*, 2013; Mannina *et al.*, 2020). Existe una amplia variedad de microorganismos de bacterias y arqueas que tienen la capacidad de sintetizar y acumular polihidroxicanoatos. Los microorganismos productores de PHAs pueden ser clasificados en tres grupos, de acuerdo con las condiciones de cultivo requeridas para su acumulación. En el primer grupo, la síntesis de polihidroxicanoatos se da como componente de reserva de energía intracelular bajo limitación de fósforo, nitrógeno, oxígeno y con un exceso de fuente de carbono (Babel & Steinbüchel, 2001; Ponnusamy *et al.*, 2019). Algunas de las bacterias que pertenecen a este grupo son *Ralstonia eutropha*, *Protomonas extorquens* y *Protomonas oleovorans*. Por otro lado, el segundo grupo no requiere de la limitación de nutrientes y puede ser acumulado durante el crecimiento. Este grupo incluye bacterias como *Alcaligenes latus*, *Azotobacter vinelandii* y *Escherichia coli* recombinante (Lee, 1996; Alavi *et al.*, 2014; Rudnik 2010; Kamravamanesh *et al.*, 2018; Koller, 2020). Por último, el tercer grupo requiere de factores de estrés ambiental como la temperatura, peróxido de hidrógeno, irradiación UV, etanol y presión osmótica para llevar a cabo la acumulación de polihidroxicanoatos (Bhat & Nambisan, 2014).

Los gránulos de PHAs son acumulados en el citoplasma rodeados por una monocapa de fosfolípidos, enzimas sintetasas, despolimerasas (González *et al.*, 2013) y fasininas (Mezzina & Pettinari, 2016). Estas últimas promueven el crecimiento bacteriano, la síntesis del polihidroxicanoato, afectan el número, tamaño y distribución de los gránulos (Mezzina & Pettinari, 2016). La forma de los gránulos de los polihidroxicanoatos es esférica y el tamaño depende del microorganismo, así mismo, llegan a representar del 30% al 80% del peso seco (Babel & Steinbüchel, 2001).

Se han reportado más de 150 tipos de monómeros intracelulares (Arumugam, 2019). Los polihidroxicanoatos son clasificados por la estructura de las unidades repetitivas en los polímeros. De esta forma, se definen los polihidroxicanoatos de cadena corta y en polihidroxicanoatos de cadena mediana (PHA scl y PHA mcl, por sus siglas en inglés). Los PHA scl están compuestos por monómeros con 3 a 5 átomos de carbono, mientras que, los PHA mcl presentan unidades monoméricas de 6 a 14 carbonos (Babel & Steinbuchel, 2001; González *et al.*, 2013).

Las principales enzimas implicadas en la biosíntesis de polihidroxicanoatos son PHA sintasas codificadas por el gen *phaC*. El sustrato de estas enzimas debe estar unido a una coenzima A. Existen cuatro clases de PHA sintasas basadas en el número de subunidades y sustratos. La PHA sintasa tipo I representada por la enzima de *Cupriavidus necator* y tiene una sola subunidad codificada por *phaC_{Re}*. La PHA sintasa tipo II es representada por la PHA sintasa 1 de *Pseudomonas oleovorans*, codificada por *phaC_{PO}*. La PHA sintasa tipo III representada por la enzima de *Allochromatium vinosum*, consiste en las subunidades *PhaC_{CV}* y *PhaE_{CV}* (Babel & Steinbuchel, 2001). Por último, PHA sintasa tipo IV es representada por *Bacillus* y está compuesta por dos unidades catalíticas, *PhaC* y *PhaR* (Martínez-Gutiérrez, 2018). Las PHA sintasas tipo I, III y IV prefieren sintetizar poliésteres con PHA scl, mientras que, la PHA sintasa tipo II tiene preferencia por poliésteres de PHA mcl (Babel & Steinbuchel, 2001; Martínez-Gutiérrez, 2018).

10.1.2 Propiedades fisicoquímicas y aplicaciones de los PHA

Los polihidroxicanoatos presentarán diferentes propiedades físicas en función a su composición monomérica. En general, el peso molecular de estos biopolímeros dependerá de la bacteria utilizada para su producción y diferirá de 2×10^5 a 3×10^6 g/mol (Babel & Steinbuchel, 2001; Sathiyarayanan *et al.*, 2016). Los PHA scl tienen un valor de T_m relativamente alto (180°C), un T_g entre -5 a 20°C y 60-80% de cristalinidad y son considerados como termoplásticos. Por otro lado, PHA mcl tienen una T_m de 42 - 58°C , un T_g entre -62 y 26°C y son altamente amorfos; siendo considerados como elastómeros

(Babel & Steinbuchel, 2001; González *et al.*, 2013). El poli(3-hidroxiбутirato) tiene una estabilidad térmica de 225°C (Hernández-Núñez *et al.*, 2019), mientras que, se ha reportado que los PHA mcl tienen una estabilidad térmica de alrededor de 180°C (Sin *et al.*, 2010).

Debido a estas propiedades, los polihidroxicanoatos tienen un alto potencial biotecnológico. Los PHA scl han sido utilizados para la generación de macetas biodegradables, tubos, liberación controlada de factores de crecimiento, medicamentos, pesticidas y herbicidas, fabricación de botellas desechables, pañales, servilletas, hilo para sutura y para estimular la formación del hueso. Por otro lado, los PHA mcl se utilizan generalmente como cubierta de quesos y hule biodegradable (Babel & Steinbuchel, 2001; González *et al.*, 2013), entre otras.

10.1.3 Producción actual de bioplásticos

A nivel global, existen 24 compañías enfocadas en la producción de polihidroxicanoatos. Entre los productos ya disponibles se encuentran: Biopol TM (copolímero de 3HB y 3HV), Biomer TM (P3HB) y Nodax TM (copolímero de 3HB y 3HV), producidos por cultivos axénicos (*E. coli* recombinante, *A. latus*, *R. eutropha* recombinante, respectivamente) (Arumugam, 2019; Serafim *et al.*, 2008). Sin embargo, la producción actual de bioplásticos es del 1% del volumen de producción total, y una de las principales limitantes en la producción de polihidroxicanoatos a nivel industrial es el costo de producción en relación con los plásticos convencionales. Entre estos, se encuentra el costo en la fermentación, incluye la fuente de carbono y la extracción del biopolímero. El primero, representa el 50% del costo total (Serafim *et al.*, 2008; Arumugam, 2019).

10.1.4 Rutas metabólicas

10.1.4.1 PHA scl

La ruta metabólica para la biosíntesis de PHA scl ha sido ampliamente estudiada en *Ralstonia eutropha*. La biosíntesis empieza con un sustrato de carbono, como carbohidratos (a través de glucólisis) o ácidos grasos mediante β -oxidación (Fig. 1). Por lo tanto, al término de la glucólisis se da la condensación de dos moléculas de acetil-CoA a acetoacetil-CoA mediante la β -cetotilasa (codificada por el gen *phaA*). Posteriormente, se reduce a (*R*)-3-hidroxiacetil-CoA por la enzima acetoacetil-CoA reductasa dependiente de NADPH (codificada por el gen *phaB*). Por último, la PHA sintasa (codificada por el gen *phaC*) cataliza la polimerización de 3-hidroxiacetil en poli(3-hidroxiacetato) (Sudesh *et al.*, 2000; Babel & Steinbuchel, 2001; González *et al.*, 2013; Martínez-Gutiérrez, 2018; Mannina *et al.*, 2020). Por otro lado, cuando se emplea la vía metabólica de degradación de ácidos grasos a través de una serie de β -oxidación, al llegar al intermediario butiril-CoA, la enzima butiril-CoA deshidrogenasa lo convierte a 2-butenoil-CoA. Entonces, la enzima enoil-CoA hidratasa convierte este producto a (*S*)-3-hidroxiacetil-CoA, donde a través de una reacción catalizada por 3-hidroxiacetil-CoA deshidrogenasa se obtiene acetoacetil-CoA. Este último entra a la ruta metabólica mediante una reducción por *phaB*.

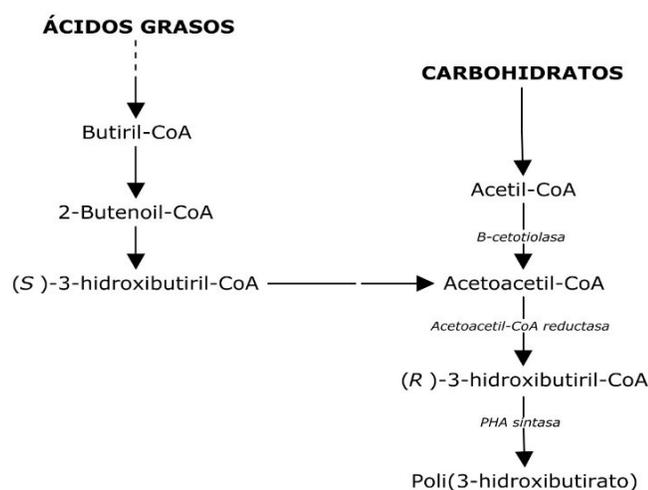


Figura 1. Ruta metabólica para la síntesis de polihidroxicanoatos de cadena corta. Tomado y adaptado de Sudesh *et al.*, (2000).

10.1.4.2 PHA mcl

La síntesis de PHA mcl se ha asociado a precursores intermediarios de tres diferentes rutas metabólicas: 1) biosíntesis de ácidos grasos de *novo* (Kim *et al.*, 2007), 2) β -oxidación (Sudesh *et al.*, 2000) y 3) elongación de cadenas de ácidos grasos, en el cual el acil-CoA se extiende con acetil-CoA. Estas rutas metabólicas producen precursores intermediarios de PHA mcl como (*R*)-3-hidroxiacil-ACP, (*S*)-3-hidroxiacil-CoA, 2-trans-Enoil-CoA y 3- cetoacil-CoA (Kim *et al.*, 2007) (Fig. 2).

Las reacciones para la conversión de estos intermediarios en (*R*)-3-hidroxiacil-CoA son las siguientes (Kim *et al.*, 2007):

1. (*S*)-3-hidroxiacil-CoA es convertido a (*R*)-3-hidroxiacil-CoA mediante la acción de 3-hidroxiacil-CoA epimerasa
2. 3- cetoacil-CoA es convertido a (*R*)-3-hidroxiacil-CoA con la ayuda de 3-cetoacil-CoA reductasa.
3. La enzima enoil-CoA hidratasa cataliza la hidratación de 2-trans-enoil-CoA a (*R*)-3-hidroxiacil-CoA.
4. La enzima (*R*)-3-hidroxiacil-ACP-CoA transferasa ha sido identificada como la conexión de la síntesis de ácidos grasos de *novo* con la síntesis de PHA mcl. Esta enzima cataliza la conversión de (*R*)-3-hidroxiacil-ACP a (*R*)-3-hidroxiacil-CoA

Finalmente, la PHA sintasa tipo II cataliza la polimerización de las moléculas de (*R*)-3-hidroxiacil-CoA en PHA mcl (Kim *et al.*, 2007).

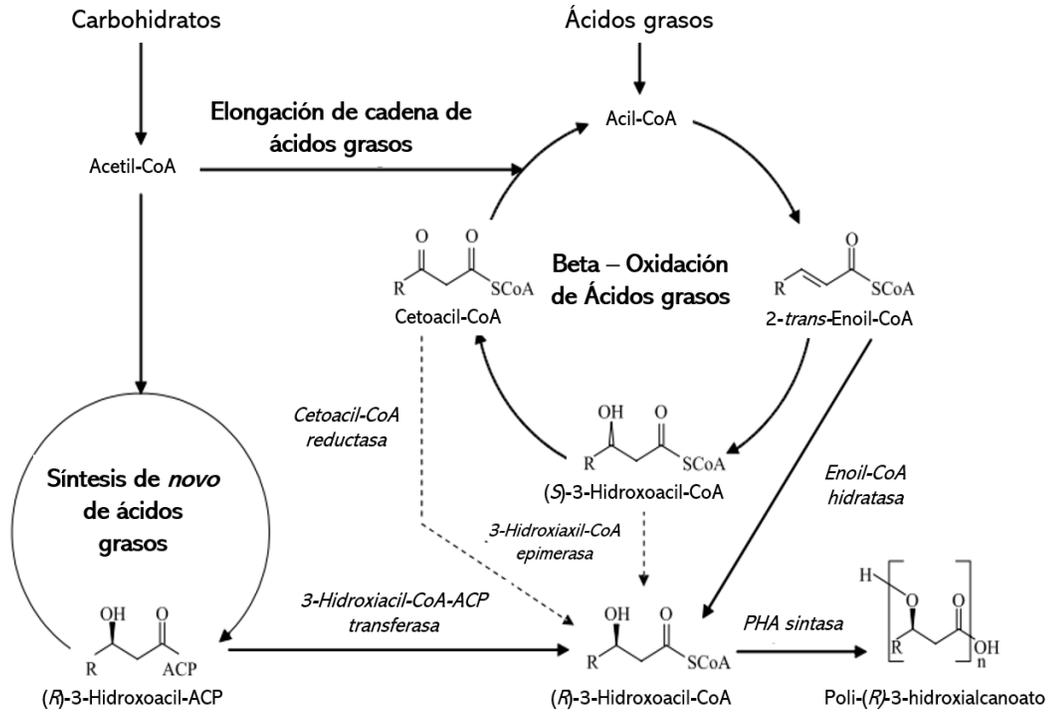


Figura 2. Ruta metabólica para la síntesis de polihidroxicanoatos de cadena mediana. Tomado y adaptado de (Kim *et al.*, 2007).

10.2 Producción de polihidroxicanoatos a partir de hidrocarburos

El uso de hidrocarburos como fuente de carbono y energía por microorganismos fue documentado desde hace más de 70 años por Zobell (1946). Las bacterias hidrocarbonoclasticas están representadas por un grupo de diversos microorganismos, no obstante, la mayoría pertenece a los géneros *Bacillus*, *Pseudomonas*, *Rhodococcus*, *Sphingomas*, y *Alcanivorax* (Sabirova, 2010).

El ecosistema marino es un ambiente oligotrófico, que puede tener entradas de hidrocarburos por fuentes antropogénicas que representa un incremento de carbono disponible. Con estas condiciones, las bacterias son capaces de desarrollar estrategias de adaptación y supervivencia, como la acumulación de polihidroxicanoatos (Manilla-Pérez *et al.*, 2010). Por lo tanto, la aplicación de estos microorganismos en biorremediación en combinación con la recuperación productos de valor como los biopolímeros provee una alternativa de plásticos sintéticos (Mwamburi *et al.*, 2019).

La biosíntesis de polihidroxicanoatos a partir de hidrocarburos está determinada por los precursores intermediarios de la vía de β -oxidación (Fig. 2) (Sabirova, 2010). En el caso de la degradación de alcanos (Fig. 3), los microorganismos utilizan la enzima monooxigenasa para la oxidación del último carbono de la molécula, obteniendo así un alcohol. Posteriormente, este grupo alcohol se oxida a un grupo aldehído y finalmente a un grupo carboxilo, el cual entrará a la vía de β -oxidación (Hernández, 2015).

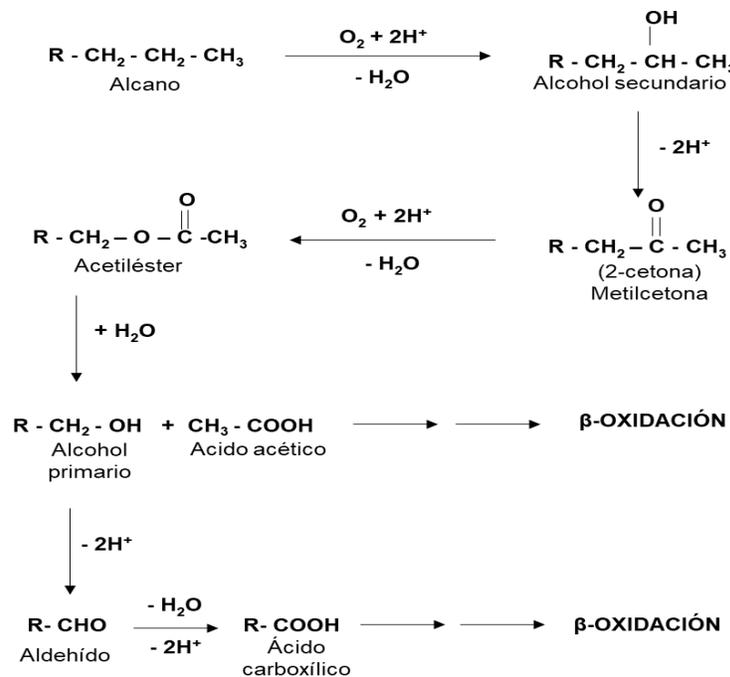


Figura 3. Degradación de alcanos hasta ácido carboxílico. Se muestra como ingresa el ácido carboxílico a la vía de β -oxidación para la obtención de precursores involucrados en la síntesis de polihidroxicanoatos. Tomado y adaptado de Hernández (2015).

10.3 Metagenomas de comunidades microbianas marinas

Se estima que existen 3.67×10^{30} microorganismos en el ambiente marino, el cual representa el 71% de la superficie del planeta Tierra (Mineta & Gojobori, 2016). Debido a esta gran diversidad microbiana y a las condiciones extremas generalmente presentes en este ambiente, la comunidad microbiana marina ha ganado gran interés principalmente en el descubrimiento de nuevas enzimas, biomoléculas y polímeros (Kumar *et al.*, 2020). Incluso, se ha sugerido que la heterogeneidad genética de este

ambiente es tan grande que existe más genes desconocidos que secuencias estudiadas, por lo que podría destacar como una fuente potencial de microorganismos productores de polihidroxicarbohidratos (Heidelberg *et al.*, 2010; Kumar *et al.*, 2020).

El término de *metagenoma* fue utilizado por primera vez por Handelsman en 1998 para definir a los genomas de todos los microorganismos encontrados en la naturaleza (Mineta & Gojobori, 2016; Tseng & Tang, 2014). La metagenómica ha surgido como una excelente herramienta para estudiar la biodiversidad y las aplicaciones biotecnológicas en el ambiente marino (Alma'abadi *et al.*, 2015).

La obtención de metagenomas a partir de muestras ambientales ha tenido un rápido incremento debido a los avances en las tecnologías de secuenciación de alto rendimiento o de nueva generación (NGS). Dos diferentes enfoques son usados en los estudios metagenómicos, el primero basado en amplicones y el segundo, en Shotgun (Heidelberg *et al.*, 2010; Tseng & Tang, 2014; Mineta & Gojobori, 2016;). El enfoque basado en amplicones genera un gran número de lecturas del gen marcador 16S ARNr de la comunidad a bajo costo y de forma rápida (Mineta & Gojobori, 2016). Este enfoque permite hacer la asignación taxonómica de las lecturas a partir de alineamientos de con bases de datos de referencia, como SILVA (Heidelberg *et al.*, 2010). Por otro lado, un enfoque basado en Shotgun es más costoso, pero genera millones de secuencias cortas de diferentes regiones de genomas que posteriormente deberán ser ensambladas para posteriormente identificar secuencias homólogas contra bases de datos de referencia, para la búsqueda de genes específicos (Mineta & Gojobori, 2016).

La gran cantidad de secuencias generadas a partir del enfoque Shotgun requiere de un extenso análisis bioinformático para asegurar la correcta interpretación de las secuencias (Mineta & Gojobori, 2016). Para esto, se puede utilizar la herramienta BLAST para comparar lecturas de genes parciales obtenidas con todas las secuencias existentes en la base de datos, de forma rápida. Esto permite el entendimiento de la información de las secuencias metagenómicas a predicciones funcionales bioquímicas y fisiológicas (Heidelberg *et al.*, 2010).

11 ANTECEDENTES

A partir de un metagenoma de suelos contaminados con petróleo de la refinería Oil and Natural Gas Corporation, en Mahesana, Gujarat, India fueron identificadas nueve PHA sintasas. Estos resultados estimaron que cerca del 3.4% de los microorganismos presentes en este sitio contienen el gen *phaC*. Además, cinco de las nueve enzimas localizadas correspondieron a microorganismos no cultivados. Una de estas enzimas (PhaC-40N22 sintasa) fue expresada en *Pseudomonas putida* cepa KT2440 $\Delta phaC1$ utilizando diferentes fuentes de carbono (hexanoato, heptanoato, octanoato, decanoato, ácido oleico y glucosa). El análisis demostró que el uso de octanoato de sodio conlleva a la producción de un terpolímero que consiste en P(3HB-co-3HHx-co-3HO) con una temperatura de descomposición alrededor de 293-295°C. Estos resultados confirman que la metagenómica puede ser considerada como una herramienta clave para la búsqueda de enzimas involucradas en la producción polihidroxicanoatos con mejores propiedades fisicoquímicas y una amplia especificidad de sustratos (Cheema *et al.*, 2011).

Narancic *et al.*, (2012) emplearon la cepa TN301 de *Pseudomonas* sp., que fue aislada de sedimentos de un río cercano a una industria petroquímica. En este trabajo se evaluó la producción de polihidroxicanoatos a partir de compuestos monoaromáticos y poliaromáticos. Para confirmar la presencia de polihidroxicanoatos se utilizó cromatografía de gases acoplado a espectroscopia de masas. Los resultados mostraron que la mayor producción (25 % de PHA mcl) se llevó a cabo cuando se utilizó una mezcla de BTEX (benceno, tolueno, etilbenceno y xileno) y no los componentes individuales. Por otro lado, el uso de naftaleno, fenantreno y criseno llevó a la producción de 22.9%, 3.5% y 5% de PHA mcl, respectivamente.

Foong *et al.*, (2014) reportaron un estudio independiente de cultivo para identificar la diversidad del gen *phaC* en agua marina colectada a diferentes profundidades en Nankai Trough y la fosa de Japón. A través del análisis por WGA DNA se identificaron 15 nuevas o no caracterizadas enzimas PHA sintasa. Así mismo, una nueva enzima PHA

sintasa funcional de *Marinobacter* sp. fue descubierta con una producción de 60% de poli(3-hidroxi-butirato).

El uso de octadecano como sustrato para la producción de polihidroxicanoatos utilizando dos consorcios microbianos aislados de suelos contaminados fue evaluado por Ghavipankeh *et al.*, (2015). Los resultados revelaron que el consorcio ABN52 tuvo la capacidad de degradar el 48.2% de *n*-octadecano y la producción del 20.4% de polihidroxicanoato (peso seco celular). Así mismo, el consorcio ISN205 degradó el 22.9% del hidrocarburo y produjo 4.8% de polihidroxicanoato. El espectro de FTIR confirmó las bandas de absorción características en 1097 cm⁻¹ (-C-O-C), 2854 cm⁻¹, 2927 cm⁻¹, 2966 cm⁻¹ (C-H) y 3410 cm⁻¹ (O-H).

En el estudio realizado por Tai *et al.*, (2016), se aplicó un enfoque metagenómico para buscar la presencia del gen *phaC* (PHA sintasa tipo I y II) en suelo caliza. Un total de 87 clonas fueron identificadas como únicas, debido a que no mostraban una identidad cercana (<85%) a PHA sintasas ya conocidas. De estas, siete clonas fueron seleccionadas ya que se trataban de bacterias no reportadas previamente como productores de polihidroxicanoatos. Una de estas, presentó 81% de identidad a la PHA sintasa de *Chromobacterium violaceum* cepa ATCC 12472 y fue expresada exitosamente en *C. necator* P3HB-4 resultando en la acumulación del 55% de polihidroxicanoato en peso seco.

El reporte del primer análisis metagenómico de bacterias productoras de polihidroxicanoatos a través de un perfil taxonómico en rizosferas y suelo fue realizado por Vahlis (2017). En este trabajo, se detectaron fueron detectados 110,872 genes de PHA sintasa tipo I, 32,575 genes de PHA sintasa tipo II y 23,165 genes de PHA sintasa tipo III. En metagenomas de rizosfera, fueron detectados *Sphingomonadaceae* (18.3%), *Rhizobiaceae* (11.665), *Caulobacteraceae* (8.3%) y *Comamonadaceae* (7.1%), mientras que, en suelo *Bradyrhizobiaceae* (13.2%), *Rhodospirillaceae* (7.7%), *Burkholderiaceae* (7.4%) y *Rhodocyclaceae* (6.2%) fueron los phylum dominantes con PHA sintasa tipo I. En el caso de la PHA sintasa tipo II, *Pseudomonadaceae*, *Geodermatophilaceae* y *Nocardiaceae* fueron dominantes tanto en metagenomas de rizosfera y suelo. Por último, la PHA sintasa tipo III fue encontrada principalmente en *Xanthomonadaceae* (66.3%) en

rizosferas, mientras que, en suelo, *Bradyrhizobiaceae* (18.6%) y *Methylobacteriaceae* fueron abundantes. Este estudio reveló nuevos taxa productores de polihidroxicanoatos.

En otro estudio realizado por Amelia *et al.*, (2018) a partir de metagenomas de la esponja marina *Aaptos aaptos* se identificaron los genes *phaC* de las bacterias *Pseudomonas stutzeri* cepa 1317 (88% de identidad), bacteria no cultivada ARS-9d_16 (93% de identidad) y *Rhodocista pekingensis* (74% de identidad). Esto demuestra que aún existen ambientes potenciales no explorados para la identificación de nuevas enzimas PHA sintasa.

El uso de hidrocarburos monoaromáticos (Fenol y 4-clorofenol) e hidrocarburos poliaromáticos (Naftaleno y Fenantreno) también ha sido evaluado para la producción de polihidroxicanoatos utilizando un cultivo axénico de *Bacillus* sp. CYR1 (Venkateswar *et al.*, 2015). En este caso, se llevó a cabo la extracción química del polihidroxicanoato y se realizó su caracterización mediante FTIR y RMN. Además, se utilizaron cromatografía GPC, TGA, XRD, DSC, para determinar sus propiedades fisicoquímicas. Los resultados mostraron que el fenol fue el único utilizado para el crecimiento, con 93% de degradación y una producción de 51% de Poli(3-Hidroxi-butirato). En el resto de los sustratos no se registró un crecimiento inicial, por lo tanto, se agregó Tween-80 para favorecer la disponibilidad de la fuente de carbono. Esto incrementó el crecimiento con una producción de 32%, 42% y 18% de poli(3-hidroxi-butirato) con 4-clorofenol, naftaleno y fenantreno, respectivamente. Se determinó que el M_w fue de 709 kDa, estabilidad térmica de 100°C, T_m de 171°C y baja cristalinidad.

Los manglares, por otra parte, también contienen una alta diversidad microbiana que constantemente está expuesta a estrés abiótico tales como salinidad y condiciones anóxicas. Foong *et al.*, (2017) a partir de metagenomas de sedimentos de manglar, reportaron el descubrimiento de una nueva enzima PHA sintasa. Esta fue identificada como BP-M-CPG4, una PHA sintasa tipo I de *Fodinicurvata fenggangensis* (47% de identidad). La expresión de BP-M-CPF4 en *Cupriavidus necator* PHB⁻4, mostró que tiene una amplia variedad de sustratos (fructosa, aceite de palma, ácido isocaproico, valerato de sodio, 5-hidroxi-valerato de sodio, 4-hidroxi-butirato de sodio, gamma-butirolactona y

hexanoato de sodio) y que puede transformarlos a homopolímeros y copolímeros de polihidroxicanoatos. Los espectros de ^1H RMN confirmaron que se trataba de poli(3-hidroxicanoato), 4HB y 3HHx con un peso molecular entre 70,000 a 260,000 Da.

En otro estudio, se realizó el primer reporte del enriquecimiento y aislamiento de un consorcio bacteriano degradador de *n*-hexadecano con la capacidad de producir polihidroxicanoatos, provenientes de un tapete microbiano de Guerrero Negro, Baja California Sur, México. Para confirmar esto, se realizó la amplificación de los genes *alkB* y *phaC*. Los gránulos de polihidroxicanoatos fueron detectados mediante la técnica de Rojo Nilo (Martínez-Gutiérrez, 2018).

Por otro lado, a partir de un muestreo en un sitio contaminado con hidrocarburos aromáticos procíclicos (fenantreno, pireno y fluoranteno), se aislaron tres microorganismos con la capacidad de degradarlos y producir polihidroxicanoatos. Las bacterias aisladas fueron identificadas como *Pseudomonas aeruginosa*, *Pseudomonas* sp. y *Ralstonia* sp. Todas fueron capaces de degradar fenantreno, pireno y fluoranteno. Los resultados mostraron que el porcentaje de polihidroxicanoato fue mayor en cultivos axénicos que en cultivos mixtos, $38.20 \pm 1.9\%$ para *P. aeruginosa*, $33.23 \pm 5.09\%$ para *Ralstonia* sp, y *Pseudomonas* sp, presentó un $29 \pm 0.72\%$ del peso seco celular. Adicionalmente, la composición de los polihidroxicanoatos fue determinada usando GC-MS, e incluye, 3HO, 3HD, 3HDD, 3HTD, 3HHD. El estudio demuestra que la producción de polihidroxicanoato utilizando hidrocarburos aromáticos procíclicos tiene diversas ventajas sobre la biorremediación y producción de polihidroxicanoatos en la reducción de costos (Sangkharak *et al.*, 2020).

12 JUSTIFICACIÓN

Los ecosistemas marinos albergan una gran diversidad microbiana. Particularmente, el Golfo de México es una de las cuencas mejor estudiadas en el mundo por su extraordinaria biodiversidad y reserva de petróleo. Esto debido, principalmente, a las filtraciones naturales de combustibles fósiles y contaminación antropogénica. Por esa razón, el Golfo de México es un ecosistema de gran interés para el estudio de las comunidades microbianas y el descubrimiento de nuevas enzimas de origen microbiano de interés en biotecnología.

Actualmente, existe una amplia cantidad de literatura sobre la síntesis y extracción de polihidroxicarbohidratos en cultivos axénicos y consorcios microbianos, así como, su caracterización fisicoquímica. No obstante, se estima que solamente el 1% de las bacterias ambientales pueden ser cultivadas, por lo que hay un gran vacío en el conocimiento de las enzimas y microorganismos no cultivables que están implicados en el metabolismo de polihidroxicarbohidratos. Por lo tanto, el uso de aproximaciones independiente de cultivo, como la metagenómica, ha surgido como una buena herramienta para estudiar el potencial metabólico de los microorganismos en la producción de polihidroxicarbohidratos.

13 PREGUNTAS DE INVESTIGACIÓN

1. ¿Cuáles son los microorganismos que tienen el potencial de sintetizar polihidroxicanoatos en muestras de sedimentos y columna de agua del suroeste del Golfo de México?
2. ¿Un consorcio microbiano aislado del sedimento del suroeste del Golfo de México creciendo con petróleo ligero como única fuente de carbono, tendrá la capacidad de producir polihidroxicanoatos?

14 OBJETIVO GENERAL

Explorar, a través de herramientas metagenómicas y de predicción funcional basado en secuencias parciales del gen 16S ARNr, respectivamente, el potencial metabólico para la síntesis de polihidroxicanoatos en muestras ambientales del suroeste del Golfo de México y en una comunidad enriquecida con hidrocarburos como única fuente carbono.

14.1 OBJETIVOS ESPECÍFICOS

1. Detectar enzimas involucradas en la síntesis de PHA en metagenomas ensamblados de comunidades microbianas de sedimento y de columna de agua provenientes del suroeste del Golfo de México.
2. Identificar taxonómicamente a los microorganismos con la capacidad de producir de polihidroxicanoatos, detectados a partir de los metagenomas ensamblados.
3. Reconstruir el metabolismo involucrado en la síntesis de polihidroxicanoatos de los microorganismos potenciales productores de polihidroxicanoatos.
4. Predecir el potencial metabólico para la producción de polihidroxicanoatos por un consorcio microbiano degradador de hidrocarburos aislado de sedimentos del suroeste del Golfo de México.

15 MATERIALES Y MÉTODOS

15.1 Ensamble de los metagenomas y reconstrucción de los genomas individuales

Las secuencias nucleotídicas utilizadas para realizar los análisis bioinformáticos de este objetivo fueron descargadas del NCBI con número de acceso SRX791344 y SRX7913439, las cuales fueron publicados previamente por Raggi *et al.*, 2020. Se utilizaron los datos de la secuenciación Shotgun generados a partir de sedimentos del Golfo de México (19.93113 N 94.3407 W) en la estación denominada D18 (D18-SED), que tiene una profundidad de 1320 m. Además, se utilizaron las secuencias correspondientes a la columna de agua de esa misma estación D18, del estrato de máxima fluorescencia (D18-MAX), que fue localizado a los 76 m de profundidad.

La incisión del adaptador y el análisis de la calidad de las lecturas fue procesada mediante Trim Galore v0.4.4. Posteriormente, se utilizó Megahit v1.2.8 (Li *et al.*, 2016) para el ensamble del metagenoma, el cual se fundamenta en el gráfico de Brujin. La elección de largo de los kmers fue basada en el largo de las lecturas, por lo tanto, para la muestra de sedimento se trabajó con un kmer mínimo de 21, un kmer máximo de 141 y un incremento en el tamaño del kmer en cada interacción de 10 (k-step). Por otro lado, para las lecturas de la muestra de columna de agua fueron utilizados los siguientes parámetros: k-max igual a 21, k-min de 71 e igualmente, un k-step de 10. En este punto, se registraron 109897 contigs para D18-MAX con un tamaño promedio de 940 pb, mientras que, de D18-SED se obtuvieron 851389 contigs con 736 pb de tamaño promedio.

La recuperación de los de los genomas individuales a partir del metagenoma fue logrado utilizando el algoritmo MaxBin v2.2.7. Este es un software que se basa en el algoritmo “Expectation-Maximization” (Wu *et al.*, 2014).

15.2 Detección de PHA sintasas en los metagenomas

Para esto, se descargaron las secuencias aminoacídicas de las enzimas PHA sintasa de referencia en NCBI. Las enzimas fueron: PHA sintasa I (SMD46940.1), PHA sintasa II (PHA1 (AAK95382.1) y PHA2 (AAK95384.1)), PHA sintasa III (Subunidad C (ABK60193.1) y Subunidad E (ABK60192.1)) y PHA sintasa IV (Subunidad R (ABF29869.1) y Subunidad C (SCV19024.1)). La búsqueda de los marcos de lecturas en los metagenomas fue utilizado el software PRODIGAL v2.6.3 (Hyatt *et al.*, 2010), en donde se realizó la traducción de la secuencia nucleotídica a aminoacídica.

A continuación, el metagenoma ensamblado con los marcos de lecturas identificados fue convertido en una base de datos utilizando makeblastdb v2.2.22+ (Camacho *et al.*, 2019). Posteriormente, se utilizó Blastp v2.7.1 (Camacho *et al.*, 2009) para la identificación de las enzimas PHA sintasas en los metagenomas. En R v3.5.3 (2019-03-11) (R Core Team, 2019), se hizo la filtración de las secuencias de PHA sintasas localizadas con una identidad igual o mayor a 40%, ya que es muy probable que las secuencias homólogas que comparten más de este porcentaje de identidad compartan funciones según la Comisión de Enzimas (Pearson, 2013). A continuación, se recuperaron el ID de las presuntas PHA sintasas presentes en los metagenomas.

Las secuencias aminoacídicas de las presuntas PHA sintasas fueron descargadas en formato FASTA utilizando filter_fasta.py v1.9.1. Con el fin de demostrar y confirmar la función de estas enzimas, se realizó la búsqueda del dominio proteico conservado (PhaC_N) en la base de datos Pfam v33.1 (Mistry *et al.*, 2021) utilizando un E-value de 1.00. Solamente cinco PHA sintasas en el metagenoma de sedimento fueron confirmadas y sus secuencias aminoacídicas fueron separadas en un archivo formato FASTA.

La metodología detallada se encuentra en el apartado de Anexos.

15.3 Identificación taxonómica de bacterias productoras de polihidroxicanoatos a partir de metagenomas

Las PHA sintasas fueron localizadas en los genomas individuales, esto con el fin de identificar a los microorganismos potenciales en la producción de polihidroxicanoatos. Se buscó de forma individual y manual cada uno de los ID de las PHA sintasas confirmadas con el dominio PhaC_N en todos los genomas individuales. Por lo tanto, el ensamblaje fue descargado e ingresado a la plataforma de KBase. Posteriormente, se realizó la anotación del genoma con RASTtk v1.073. A continuación, se determinó el porcentaje del genoma recuperado mediante el programa CheckM v1.0.18 (Parks *et al.*, 2015). Después, se realizó la clasificación taxonómica de este utilizando programa GTDB-Tk classify v0.14, el cual utiliza como criterio la divergencia evolutiva relativa para establecer rangos taxonómicos (Chaumeil *et al.*, 2020).

Se realizó un árbol filogenético basado en la máxima probabilidad del Bin 91 y un conjunto de genomas estrechamente relacionados seleccionados de la importación de genomas públicos KBase de RefSeq ($1 \leq 100 \leq 200$) utilizando el programa SpeciesTree v2.2.0. La construcción del árbol parte de un conjunto de 49 genes centrales definidos por familias de genes COG. Para esto, fue necesario convertir el genoma en un GenomeSet utilizando Build GenomeSet v.1.0.1. Para la interpretación y visualización se utilizó el programa FigTree v1.4.4. Es este punto, debido a los resultados preliminares, se decidió nombrar al genoma del Bin 91 como PHAZGoM18. Finalmente, la visualización del genoma circular de PHAZGoM18 se realizó en CGView Server (<http://cgview.ca/>). Para esto, se ingresó el ensamblaje completo y se realizó la anotación con Prokka utilizando como código genético *Bacterial and Plant Plastid*. Así mismo, se agregó el contenido G+C (mol%).

15.4 Reconstrucción metabólica de PHAZGoM18

Se construyó un resumen de la evaluación funcional, a través de los genes con funciones conocidas presentes en PHAZGoM18, utilizando la asignación de genes COG. Para ello, se utilizó el programa View Function Profile for Genomes - v1.4.0, con un

GenomeSet anotado con RASTtk v1.073. Así mismo, se realizó la reconstrucción metabólica involucrada en la producción de polihidroxicanoatos de PHAZGoM. A través de KBase, se utilizó el programa Build Metabolic Model el cual usa el genoma anotado con RASTtk v1.073. Como criterio para construir la ruta metabólica, se basó del hecho que la PHA sintasa localizada en el genoma de PHAZGoM18 fue una PHA sintasa tipo I que llevaba a la producción de poli(3-hidroxibutirato). Las reacciones fueron conectadas y fue utilizado la base de datos ModelSEED, para el complementar con las reacciones ausentes.

15.5 Predicción del potencial metabólico de un consorcio microbiano marino aislado de sedimentos del suroeste del Golfo de México

Las secuencias del gen 16S ARNr utilizadas para este trabajo fueron generadas previamente a partir de un consorcio microbiano proveniente de sedimentos del suroeste del Golfo de México, en experimentos de degradación de petróleo ligero en condiciones de laboratorio. El ensayo consistió en 30 días de incubación; donde se tomó una muestra para secuenciación del gen 16S ARNr en el tiempo cero y cada tercer día (Uribe-Flores *et al.*, 2019). Las secuencias utilizadas para este análisis fueron obtenidas del NCBI con número de acceso PRJNA494696.

15.5.1 PICRUST2

Las secuencias demultiplexadas en formato FASTQ fueron importadas a QIIME2 pipeline (2017.11). Se realizó la eliminación de las secuencias de los primers y la corrección de los errores de secuenciación con el plugin DADA2 para resolver los ASVs. Posteriormente, las secuencias de los ASVs en formato FASTA fueron sometidas al software bioinformático PICRUST2 (Douglas *et al.*, 2020), con los parámetros default el cual está diseñado para predecir el contenido funcional del metagenoma a partir de genes marcadores y genomas completos (Douglas *et al.*, 2020). Se obtuvieron tres

documentos: 1) Rutas metabólicas, 2) Enzimas según KEGG data base y 3) las enzimas en un esquema de clasificación de la Comisión de Enzimas.

Los tres archivos fueron analizados en Lenguaje de programación R version 3.5.3. Para la identificación y confirmación de las rutas metabólicas y las enzimas se utilizaron las bases de datos: KEGG database (Kanehisa & Goto, 2000) y BRENDA (Jeske *et al.*, 2019).

16 RESULTADOS

16.1 Exploración del potencial metabólico para la síntesis de polihidroxicanoatos a partir de metagenomas del suroeste del Golfo de México

En este trabajo, dos metagenomas provenientes del suroeste del Golfo de México fueron ensamblados y evaluados para determinar la síntesis potencial de polihidroxicanoatos. El metagenoma de sedimento tuvo un total de 162961050 lecturas con un tamaño de 150 pb cada una y un contenido G+C (%mol) de 53.11%. Por otra parte, se obtuvieron 77396766 lecturas en el metagenoma de columna de agua con un tamaño de 75 pb y 49.59% de contenido G+C (%mol).

Una vez reconstruidos los genomas individuales, se localizaron 34 Bins en el metagenoma de columna de agua, mientras que, en el metagenoma de sedimento se reconocieron 151 Bins. Además, se realizó la búsqueda de las enzimas PHA sintasas de referencia en los metagenomas de sedimento y columna de agua, debido a que es la enzima clave en la síntesis de estos gránulos de reserva. En el metagenoma de sedimento, fueron predichas 16 enzimas, no obstante, solamente cinco enzimas PHA presentaron el dominio proteico PhaC_N (Tabla 1). Por otra parte, en el metagenoma de columna de agua, se identificaron cinco secuencias de enzimas PHA sintasas, sin embargo, no presentaron el dominio proteico clave que conlleva a la función de la enzima.

Tabla 1. PHA sintasas localizadas en el metagenoma de sedimento del sitio de muestreo D18-SED del suroeste del Golfo de México.

REFERENCIA	ENZIMA	% IDENTIDAD	DESCRIPCIÓN
SMD46940.1	k141_122519_1	47.500	PHA sintasa I
ABK60192.1	k141_685128_2	40.751	PHA sintasa III (Subunidad E)
ABK60192.1	k141_931424_1	60.976	PHA sintasa III (Subunidad E)
ABK60192.1	k141_931424_2	64.085	PHA sintasa III (Subunidad E)
ABK60192.1	k141_1563817_1	42.553	PHA sintasa III (Subunidad E)

En el metagenoma de sedimento, las enzimas k141_122519_1, k141_931424_1 y k141_931424_2 fueron encontradas dentro del genoma del Bin 91, mientras que, k141_685128_2 y k141_1563817_1 no fueron asociadas a otro Bin. Debido a que no se localizó la subunidad PhaC de la PHA sintasa tipo III, se propuso que el Bin 91 solamente tiene la enzima PHA sintasa tipo I que lleva a la producción de polihidroxicanoatos de cadena corta.

Para indagar la taxonomía del Bin 91, en la Fig. 4 se observa que el genoma de este microorganismo se encuentra relacionado filogenéticamente con *Ardenticatena marítima*, del Phylum Chloroflexi, Orden Ardenticatenales. Sin embargo, el análisis realizado en el programa GTDB-Tk classify v0.14 demostró que, por divergencia evolutiva, el Bin 91 puede estar más relacionado al Orden Dehalococoidales del Phylum Chloroflexi. Debido al alto valor del soporte (100%) se indica que hay una fuerte evidencia que las secuencias de estas bacterias se agrupan con exclusión de cualquier otro.

En la Fig. 5 se observa que se logró reconstruir el 82.7% del genoma del Bin 91, el cual tiene un tamaño de 6013805 pb y un contenido G+C (mol%) del 60.9%. En vista de la ausencia de una clasificación completa y de la falta de reportes de este microorganismo, en este trabajo se propuso nombrar al genoma de la bacteria como PHAZGoM18.

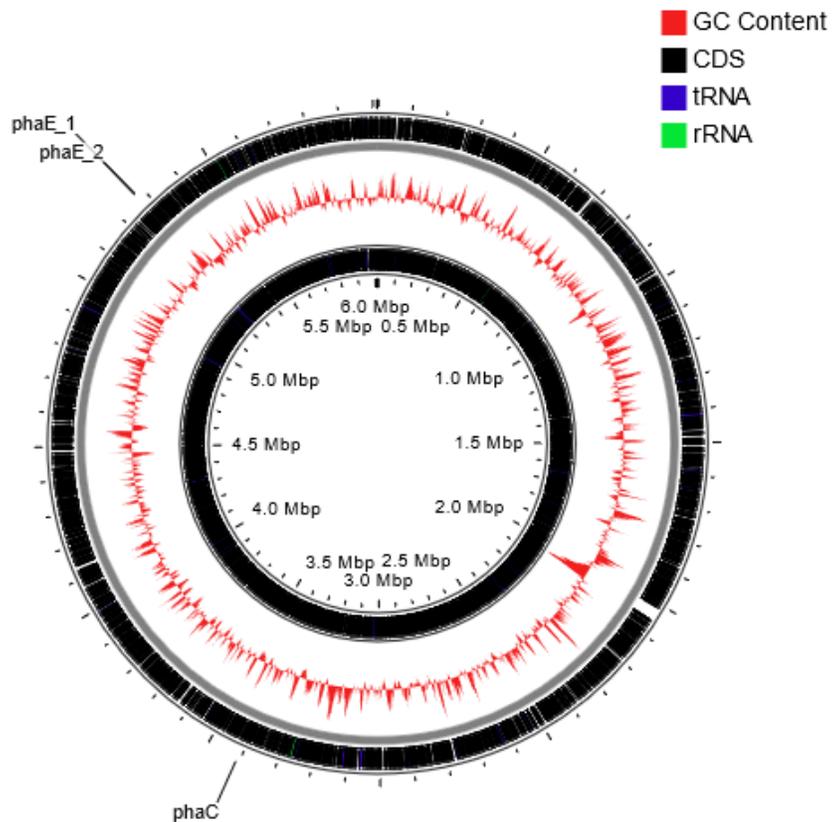


Figura 5. Visualización gráfica de genoma PHAZGoM18. Las flechas indican la posición de las enzimas PHA sintasas localizadas. En el interior se refleja la cadena antisentido, mientras que, en el exterior la cadena sentido.

En la Fig. 6 se observa que los genes *phaC*, *phaE_1* y *phaE_2* se encuentran en la cadena sentido del genoma. El gen *phaC* comienza en la posición 3409284 pb y termina en 3409481 pb, con un tamaño de 198 pb. Sin embargo, no se pudieron anotar correctamente los genes cercanos, a excepción de *gyrB_3*, que codifica para una ADN

girasa (Fig. 6A). Por otro lado, los genes *phaE_1* y *phaE_2* se encuentran de forma adyacente (Fig. 6B). El primero inicia en 5273999 pb y termina en 5274253 con un tamaño de 255 pb, mientras que *phaE_2*, se localiza entre 5274257 a 5274769 pb, teniendo un tamaño de 513 pb. Así mismo, río abajo fueron anotados los genes fenilacetato-CoA ligasa, *sufT_1* y *lutC*, en tanto río arriba, *ribD_2* y *ribE_2*. El resto de los genes involucrados en el metabolismo de polihidroxicanoatos de cadena corta no fueron encontrados cercanos al gen PHA sintasa.

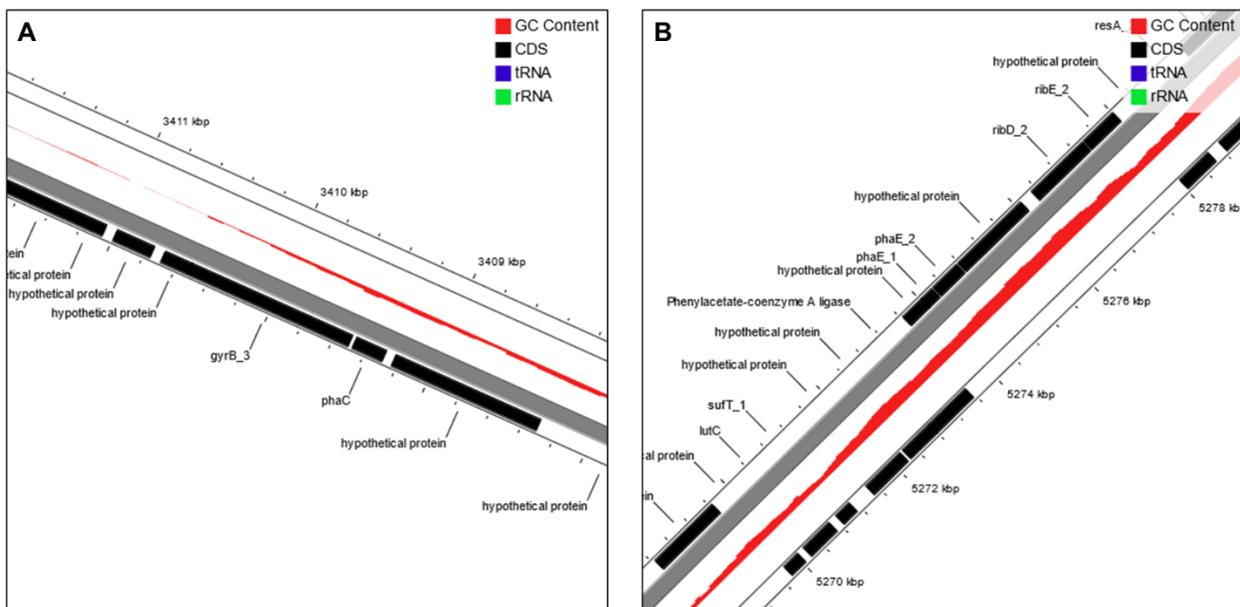


Figura 6. PHA sintasas detectadas en el genoma de PHAZGoM18. A) se observa la ubicación del gen *phaC*. B) indica la posición de los genes *phaE_1* y *phaE_2*.

La evaluación funcional de PHAZGoM18 mostró que las rutas metabólicas más abundantes, estuvieron agrupadas en mecanismos de traducción de señales (T), traducción, estructura y biogénesis ribosomal (J), producción y conversión de energía (C), metabolismo y transporte de aminoácidos (E) y lípidos (I) (Fig. 7). En cambio, la biosíntesis de metabolitos secundarios (Q), como lo son los polihidroxicanoatos, se encuentra en baja abundancia.



KEY

CELLULAR PROCESSES AND SIGNALING	D	Cell cycle control, cell division, chromosome partitioning
	M	Cell wall/membrane/envelope biogenesis
	N	Cell motility
	O	Post-translational modification, protein turnover, and chaperones
	T	Signal transduction mechanisms
	U	Intracellular trafficking, secretion, and vesicular transport
	V	Defense mechanisms
INFORMATION STORAGE AND PROCESSING	W	Extracellular structures
	Y	Nuclear structure
	Z	Cytoskeleton
	A	RNA processing and modification
METABOLISM	B	Chromatin structure and dynamics
	J	Translation, ribosomal structure and biogenesis
	K	Transcription
	L	Replication, recombination and repair
POOR CHAR	C	Energy production and conversion
	E	Amino acid transport and metabolism
	F	Nucleotide transport and metabolism
	G	Carbohydrate transport and metabolism
	H	Coenzyme transport and metabolism
	I	Lipid transport and metabolism
	P	Inorganic ion transport and metabolism
	Q	Secondary metabolites biosynthesis, transport, and catabolism
R	General function prediction only	
S	Function unknown	

Figura 7. Evaluación funcional de PHAZGoM18. Resumen de las funciones conocidas presentes en el genoma anotado con RASTtk de PHAZGoM18. Mapa de calor de abundancia relativa en una escala blanco-azul.

La reconstrucción de la ruta metabólica involucrada en la síntesis de polihidroxicanoatos indicó que el microorganismo con el genoma PHAZGoM18 tiene la capacidad de producir poli(3-hidroxiбутirato), a partir del metabolismo de carbohidratos o hidrocarburos (Fig. 8).

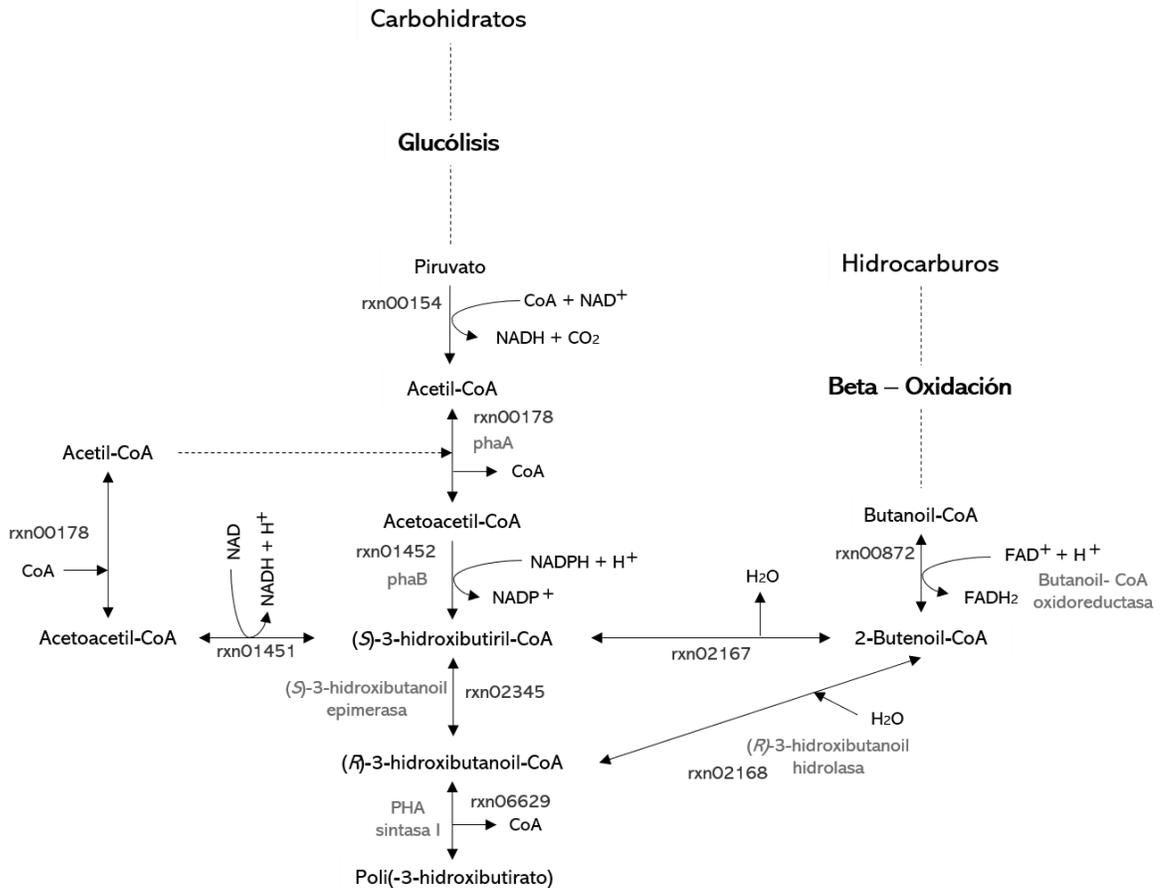


Figura 8. Reconstrucción metabólica involucrada en la producción de polihidroxicanoatos por PHAZGoM18.

16.2 Potencial metabólico para la producción de polihidroxicanoatos por un consorcio microbiano degradador de hidrocarburos

En los datos analizados, se identificaron rutas metabólicas que proveen de intermediarios para la síntesis de PHA, como la β -oxidación y elongación de ácidos grasos (Tabla 2). La β -oxidación la cual provee de tres intermediarios para la síntesis de (*R*)-3-hidroxiacil-CoA ((*S*)-3-hidroxiacil-CoA, 3- cetoacil-CoA y 2-trans-enoil-CoA), es la tercera ruta metabólica más abundante en el consorcio microbiano. Los resultados indicaron una variación en la abundancia, donde se registró un aumento al tercer día de incubación con 1.395% y a continuación se observó un descenso hasta llegar al día 30 (0.946%). Así mismo, la elongación de ácidos grasos fue localizada; al día cero, presentó

una abundancia del 0.830%, aumentando al tercer día a 0.899%. Posteriormente, tiene una abundancia oscilando entre 0.885-0.0911%. Por otro lado, la oxidación de octano es la cuarta ruta más abundante presente en el consorcio microbiano. Al inicio el ensayo, presenta una abundancia de 0.779% que aumenta al llegar al día tres a 1.264%. Al igual que la β -oxidación, a partir del este día, se registra una disminución hasta llegar a 0.803% al término del ensayo.

En la Tabla 3 se observa la presencia de acetil-CoA acetiltransferasa (EC 2.3.1.9) con una abundancia de 0.622% a las 72 horas de incubación, la enzima 3-oxoacil-[ACP]-reductasa (EC 1.1.1.100, 0.462-0.523%) y enoil-hidratasa (EC 4.2.1.17, 0.465-0.491%), enzimas relacionadas con la producción de polihidroxicanoatos en bacterias y arqueas. Por otra parte, la Tabla 4 muestra las 25 enzimas más abundantes considerando KEGG database. De igual forma se observa la presencia de acetil-CoA acetiltransferasa (K00626, 0.288-0.307%) y 3-oxoacil-[ACP] reductasa (K00059, 0.228-0.273%).

La Tabla 5 muestra las enzimas implicadas en la producción de PHA encontradas en menor abundancia. Las enzimas phaA, phaC, phaD, phaE, phaF y phaG (K05559-K05564) presentan la misma abundancia entre ellas durante los diez tiempos de toma de muestra para la secuenciación. Así mismo, se identificó que la enzima acetoacetyl-CoA reductasa (P3HBB, EC 3.1.1.35, K00023) al inicio del ensayo presentó una abundancia del 0.0254% que al llegar al tercer día disminuye en 0.0168% y posteriormente se mantiene oscilando entre 0.0172 a 0.0223%. Al inicio del ensayo (T0), la enzima poli(3-hidroxi-butirato) despolimerasa (phaZ, E.C 3.1.1.75, K05973) presenta una abundancia de 0.0249%. No obstante, a las 72 horas de incubación se registra una disminución a 0.0167%. Más adelante, se observaron fluctuaciones en la abundancia hasta llegar a 0.022% al final del ensayo.

La enzima clave en la síntesis de la producción de los biopolímeros, PHA sintasa (EC 2.3.1.-, K03821), mostró una mayor abundancia al tercer día (0.0939%), mientras que, los días siguientes del ensayo, mostró una variación entre 0.0770-0.0869%. Por último, se observó la presencia de dos enzimas enoil-CoA hidratasa en menor abundancia. Al igual que en la tabla 2, se localizó la enoil-CoA hidratasa (EC 4.2.1.17, K01692) pero con una abundancia menor (0.0933-0.1524%) registrándose con mayor

presencia al día 30 de incubación y con menor abundancia el día 3. Así mismo, se localizó la enoil-CoA hidratasa 2 (EC 4.2.1.119) con una abundancia entre 0.014% a 0.055%.

Tabla 2. Principales rutas metabólicas predichas por PICRUSt2 a partir secuencias del 16S ARNr del consorcio microbiano proveniente de sedimentos del suroeste de Golfo de México creciendo con petróleo ligero como única fuente de carbono. Los valores de abundancia se muestran en porcentaje.

DESCRIPCIÓN	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
Ácidos grasos (Salvage)	0.940	1.453	1.099	1.243	1.117	1.137	1.096	1.027	1.022	1.013	0.999
Respiración aeróbica I	1.480	1.448	1.506	1.497	1.512	1.517	1.522	1.524	1.528	1.528	1.527
Beta oxidación	0.894	1.395	1.051	1.184	1.060	1.080	1.042	0.975	0.969	0.961	0.946
Oxidación de octano	0.779	1.264	0.970	1.062	0.944	0.947	0.912	0.851	0.830	0.829	0.803
Fermentación de piruvato a isobutanol	1.104	1.249	1.116	1.148	1.103	1.105	1.093	1.072	1.058	1.061	1.058
Biosíntesis L-isoleucina II	0.987	1.015	1.023	1.027	1.025	1.028	1.030	1.018	1.013	1.015	1.015
Biosíntesis L-isoleucina I	0.964	0.996	1.001	1.005	1.001	1.003	1.004	0.991	0.985	0.987	0.987
Biosíntesis L-valina	0.964	0.996	1.001	1.005	1.001	1.003	1.004	0.991	0.985	0.987	0.987
Biosíntesis de góndrato	0.955	0.957	1.028	1.020	1.044	1.052	1.061	1.057	1.071	1.070	1.078
Elongación de ácidos grasos	0.830	0.899	0.902	0.887	0.885	0.898	0.905	0.887	0.897	0.904	0.911
Biosíntesis de micolato	0.838	0.864	0.846	0.856	0.838	0.843	0.850	0.828	0.831	0.823	0.812
Supervía de biosíntesis de aminoácidos ramificados	0.836	0.863	0.885	0.882	0.888	0.891	0.894	0.888	0.887	0.886	0.886
Biosíntesis de oleato IV (anaeróbico)	0.803	0.826	0.813	0.822	0.804	0.809	0.817	0.795	0.799	0.790	0.779
Biosíntesis de estearato	0.785	0.826	0.787	0.805	0.782	0.789	0.795	0.770	0.773	0.767	0.756
Biosíntesis de cis-vaccenato	0.867	0.825	0.959	0.927	0.957	0.954	0.966	0.977	0.978	0.961	0.963
Biosíntesis de (5Z) -dodec-5-enoato	0.769	0.795	0.776	0.786	0.764	0.769	0.776	0.754	0.756	0.747	0.734
Biosíntesis de palmitoleato I (de (5Z) -dodec-5-enoato)	0.772	0.795	0.773	0.784	0.764	0.770	0.777	0.754	0.758	0.750	0.739
Biosíntesis de CDP-diacilglicerol I	0.739	0.775	0.759	0.763	0.749	0.751	0.750	0.745	0.743	0.743	0.740
Biosíntesis de CDP-diacilglicerol II	0.739	0.775	0.759	0.763	0.749	0.751	0.750	0.745	0.743	0.743	0.740
Biosíntesis III de L-isoleucina	0.744	0.763	0.776	0.775	0.777	0.779	0.781	0.778	0.777	0.776	0.775
Ciclo TCA	0.675	0.753	0.740	0.759	0.756	0.757	0.752	0.750	0.749	0.746	0.740
Supervía de biosíntesis de L-isoleucina I	0.696	0.745	0.725	0.735	0.723	0.724	0.726	0.722	0.723	0.719	0.716
Supervía de biosíntesis de fosfolípidos I	0.676	0.699	0.690	0.698	0.690	0.693	0.691	0.687	0.687	0.684	0.682
Supervía de nucleobases de pirimidina (Salvage)	0.673	0.694	0.699	0.702	0.700	0.702	0.701	0.701	0.700	0.693	0.688
Supervía de biosíntesis de L-treonina	0.630	0.692	0.654	0.670	0.651	0.652	0.653	0.649	0.652	0.646	0.643

Tabla 3. Principales enzimas predichas por PICRUST2 a partir de la secuenciación del 16S ARNr del consorcio microbiano proveniente de sedimentos del suroeste de Golfo de México creciendo con petróleo ligero como única fuente de carbono. Los valores de abundancia se muestran en porcentaje.

ENZIMAS	DESCRIPCIÓN	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
EC:2.7.7.7	DNA polimerasa	1.155	1.149	1.198	1.189	1.197	1.198	1.203	1.205	1.208	1.199	1.190
EC:2.7.13.3	Histidina quinasa	0.989	1.081	1.007	1.013	0.974	0.985	0.978	0.963	0.963	0.980	0.982
EC:3.6.4.12	DNA helicasa	1.053	1.064	1.071	1.068	1.058	1.058	1.056	1.056	1.053	1.044	1.036
EC:5.2.1.8	Peptidil-propil isomerasa	0.819	0.888	0.825	0.848	0.795	0.804	0.800	0.782	0.788	0.792	0.788
EC:2.3.1.9	Acetil-CoA acetiltransferasa	0.570	0.622	0.551	0.584	0.563	0.567	0.557	0.548	0.547	0.557	0.565
EC:1.3.8.7	Acil-CoA deshidrogenasa de cadena mediana	0.425	0.604	0.487	0.552	0.517	0.527	0.512	0.486	0.493	0.502	0.512
EC:3.6.4.13	RNA helicasa	0.473	0.574	0.489	0.501	0.455	0.461	0.455	0.440	0.440	0.447	0.445
EC:2.5.1.18	Glutati3n transferasa	0.434	0.563	0.491	0.515	0.480	0.496	0.495	0.466	0.478	0.501	0.515
EC:1.6.5.3	Ubiquinona reductasa	0.760	0.513	0.684	0.632	0.714	0.701	0.718	0.751	0.741	0.755	0.764
EC:2.7.7.6	RNA polimerasa	0.478	0.474	0.508	0.497	0.505	0.505	0.506	0.510	0.509	0.505	0.503
EC:1.1.1.100	3-oxoacil- [acil-proteína transportadora] reductasa	0.494	0.462	0.480	0.474	0.502	0.505	0.507	0.507	0.507	0.511	0.523
EC:4.2.1.17	Enoil-CoA hidratasa	0.469	0.461	0.435	0.470	0.469	0.473	0.466	0.463	0.468	0.478	0.491
EC:6.3.5.6	Asparaginil-tRNA sintasa	0.383	0.459	0.422	0.430	0.413	0.416	0.410	0.402	0.396	0.398	0.396
EC:6.3.5.7	Glutaminil-tRNA sintasa	0.383	0.459	0.422	0.430	0.413	0.416	0.410	0.402	0.396	0.398	0.396
EC:1.1.1.1	Alcohol deshidrogenasa	0.341	0.458	0.366	0.403	0.369	0.369	0.355	0.345	0.339	0.337	0.334
EC:1.9.3.1	Citocromo C-oxidasa	0.440	0.451	0.443	0.447	0.441	0.445	0.447	0.442	0.450	0.448	0.447
EC:1.1.1.35	3-hidroxiacil-CoA deshidrogenasa	0.261	0.407	0.316	0.345	0.311	0.317	0.310	0.290	0.290	0.292	0.290
EC:2.2.1.6	Acetolactato sintasa	0.427	0.385	0.385	0.376	0.375	0.374	0.376	0.374	0.371	0.373	0.376
EC:2.7.11.1	Serina/treonina proteína quinasa	0.409	0.376	0.389	0.384	0.391	0.379	0.371	0.390	0.376	0.362	0.358
EC:6.4.1.2	Acetil-CoA carboxilasa	0.374	0.360	0.385	0.383	0.394	0.393	0.394	0.399	0.399	0.394	0.391
EC:6.2.1.3	Ácido graso de cadena larga-CoA ligasa	0.278	0.333	0.309	0.326	0.330	0.335	0.332	0.326	0.329	0.333	0.339
EC:1.11.1.15	Peroxiredoxina	0.360	0.329	0.329	0.331	0.323	0.322	0.322	0.322	0.323	0.317	0.315
EC:3.1.3.18	Fosfoglicolato fosfatasa	0.281	0.314	0.281	0.282	0.262	0.267	0.266	0.258	0.257	0.265	0.265
EC:1.7.1.15	Nitrito reductasa	0.218	0.312	0.231	0.253	0.215	0.214	0.204	0.196	0.192	0.193	0.189
EC:1.13.12.16	Nitronato monooxigenasa	0.188	0.306	0.232	0.262	0.231	0.236	0.230	0.211	0.214	0.213	0.212

Tabla 4. Principales enzimas predichas por PICRUST2 a partir de la secuenciación del 16S ARNr del consorcio microbiano proveniente de sedimentos del suroeste del Golfo de México creciendo con petróleo ligero como única fuente de carbono. Los valores de abundancia se muestran en porcentaje.

ENZIMAS	DESCRIPCIÓN	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
K03088	Factor de RNA polimerasa sigma-70	0.424	0.368	0.445	0.438	0.467	0.460	0.464	0.476	0.477	0.461	0.459
K00626	Acetil-CoA acetiltransferasa	0.291	0.307	0.283	0.299	0.294	0.296	0.291	0.288	0.288	0.292	0.297
K02014	Proteína receptora de la membrana externa (Complejo Fe)	0.266	0.299	0.276	0.307	0.285	0.293	0.292	0.279	0.294	0.291	0.289
K00249	Acil-CoA deshidrogenasa	0.217	0.298	0.250	0.282	0.270	0.274	0.267	0.256	0.259	0.264	0.269
K07044	Proteína no caracterizada	0.108	0.285	0.182	0.220	0.173	0.175	0.163	0.143	0.140	0.134	0.124
K00799	Glutación-S transferasa	0.222	0.278	0.252	0.264	0.251	0.259	0.259	0.245	0.251	0.263	0.270
K00936	Sensor de histidina quinasa	0.127	0.275	0.219	0.250	0.225	0.227	0.221	0.206	0.210	0.204	0.196
K07107	Acil-CoA tioéster hidrolasa	0.205	0.271	0.229	0.246	0.225	0.226	0.222	0.212	0.211	0.206	0.200
K00257	Acil-ACP deshidrogenasa	0.187	0.245	0.200	0.222	0.212	0.212	0.205	0.200	0.197	0.199	0.200
K03704	cspA	0.218	0.241	0.217	0.222	0.210	0.212	0.210	0.205	0.204	0.205	0.206
K00059	3-oxoacil-[ACP] reductasa	0.250	0.228	0.246	0.241	0.260	0.261	0.263	0.264	0.265	0.266	0.273
K07090	Proteína no caracterizada	0.213	0.219	0.220	0.217	0.218	0.219	0.220	0.219	0.219	0.219	0.219
K01990	Proteína de unión a ATP del sistema ABC-2	0.258	0.203	0.248	0.237	0.259	0.253	0.254	0.268	0.266	0.257	0.256
K00100	Butanol deshidrogenasa	0.080	0.190	0.136	0.163	0.136	0.139	0.132	0.120	0.122	0.121	0.118
K03496	Proteína de partición cromosómica.	0.168	0.175	0.180	0.176	0.180	0.181	0.182	0.182	0.179	0.183	0.183
K07001	Proteína de la familia NTE	0.145	0.173	0.144	0.152	0.139	0.140	0.138	0.131	0.131	0.131	0.130
K01992	Proteína permeasa del sistema de transporte tipo ABC-2	0.229	0.167	0.200	0.190	0.209	0.205	0.205	0.217	0.215	0.212	0.214
K01997	Proteína permeasa del sistema de transporte de aminoácidos de cadena ramificada	0.138	0.166	0.146	0.143	0.143	0.146	0.148	0.137	0.134	0.138	0.140
K01897	Acil-CoA sintetasa de cadena larga	0.142	0.164	0.159	0.167	0.172	0.174	0.173	0.171	0.173	0.175	0.178
K00666	Acil-graso-CoA sintasa	0.202	0.163	0.163	0.169	0.179	0.179	0.177	0.180	0.180	0.186	0.194
K01091	Fosfoglicolato fosfatasa	0.143	0.155	0.144	0.144	0.137	0.139	0.139	0.136	0.135	0.139	0.139
K00001	Alcohol deshidrogenasa	0.066	0.153	0.107	0.123	0.101	0.102	0.098	0.087	0.086	0.083	0.079
K03559	Proteína de transporte de biopolímero ExbD	0.146	0.151	0.160	0.161	0.161	0.162	0.163	0.162	0.163	0.162	0.162
K00459	Nitronato monooxigenasa	0.096	0.151	0.119	0.134	0.121	0.123	0.120	0.111	0.113	0.112	0.111
K06911	Proteína no caracterizada	0.117	0.150	0.126	0.141	0.131	0.133	0.129	0.124	0.124	0.123	0.120

Tabla 5. Enzimas involucradas en la producción de PHAs predichas por PICRUST2 a partir de la secuenciación del 16S ARNr del consorcio microbiano proveniente de sedimentos del suroeste del Golfo de México creciendo con petróleo ligero como púnica fuente de carbono. Los valores de abundancia se muestran en porcentaje.

EC	KEGG	ABREVIACIÓN	DESCRIPCIÓN	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
-	K05564	phaG	-	0.029	0.042	0.033	0.035	0.030	0.031	0.030	0.028	0.028	0.029	0.028
-	K05563	phaF	-	0.029	0.042	0.033	0.035	0.030	0.031	0.030	0.028	0.028	0.029	0.028
-	K05562	phaE	-	0.029	0.042	0.033	0.035	0.030	0.031	0.030	0.028	0.028	0.029	0.028
-	K05561	phaD	-	0.029	0.042	0.033	0.035	0.030	0.031	0.030	0.028	0.028	0.029	0.028
-	K05560	phaC	-	0.029	0.042	0.033	0.035	0.030	0.031	0.030	0.028	0.028	0.029	0.028
-	K05559	phaA	-	0.029	0.042	0.033	0.035	0.030	0.031	0.030	0.028	0.028	0.029	0.028
EC:1.1.1.36	K00023	P3HBB	Acetoacetyl-CoA reductasa	0.025	0.017	0.019	0.017	0.019	0.019	0.019	0.019	0.020	0.021	0.022
EC:3.1.1.75	K05973	phaZ	Poli (3)-hidroxibutirato despolimerasa	0.025	0.017	0.019	0.017	0.018	0.018	0.019	0.019	0.019	0.021	0.022
EC:2.3.1.-	K03821	P3HBC, phaC	Polihidroxiálcanoato sintasa	0.083	0.094	0.080	0.087	0.080	0.081	0.080	0.076	0.077	0.079	0.080
EC 4.2.1.17	K01692	-	Enoil-CoA hidratasa	0.146	0.093	0.105	0.121	0.132	0.133	0.130	0.134	0.138	0.143	0.152

17 DISCUSIÓN

17.1 Exploración del potencial metabólico para la síntesis de polihidroxicanoatos a partir de metagenomas del suroeste del Golfo de México

Durante el desarrollo de este trabajo, se lograron reconstruir 34 Bins provenientes del metagenoma de una muestra de columna de agua y 151 Bins a partir del metagenoma de la muestra de sedimento, reportados previamente por Raggi *et al.*, (2020). Los autores de este artículo reportaron que la diversidad de microorganismos, estimado a través del Índice de Shannon, para la muestra de sedimento fue mayor que la detectada para la muestra de agua. Con ello, ya se esperaba que para este análisis se hubieran ensamblado más genomas para la muestra de sedimentos que en la muestra de columna de agua. No obstante, no fue posible adjudicar la menor presencia de Bins en el metagenoma de la columna de agua en términos de diversidad, ya que se debe considerar que el tamaño de las lecturas reportadas (77396766) de la secuenciación de esta muestra fue de solamente 75 pb.

Los valores de Bins reportados en este trabajo pueden ser atribuidos a la baja cobertura en la secuenciación. Autores como López, (2016) y Jo *et al.*, (2020) han indicado que la cobertura, tamaño de la lectura y la tasa de error afectan el desempeño en el análisis de la comunidad microbiana. Incluso, Covarrubias, (2020) demuestra que existe una correlación positiva entre el número de lecturas y el número de OTUs obtenidos por muestra. Por lo que, el aumento en la profundidad de la secuenciación es esencial si se espera detectar un mayor número de OTUs. Además, de que esto facilita el realizar ensambles de *novο* de especies donde no se dispone de un genoma de referencia.

Continuando con la búsqueda de las enzimas PHA sintasas en ambos metagenomas, cinco enzimas PHA sintasa fueron confirmadas en la muestra de sedimento (Tabla 1). En la actualidad, no existen reportes de la presencia del gen *phaC* en el Golfo de México. No obstante, solamente se reporta un estudio realizado

en México por Martínez-Gutiérrez *et al.*, (2018) en la región de Guerrero Negro, Baja California Sur. En dicho trabajo, se logró la identificación del gen *phaC* en tapetes microbianos. Por otro lado, si se ha reportado la presencia de PHA sintasas en sedimento de manglar de Balik Pulau, Malasia a través de un análisis metagenómico. Foong *et al.*, (2017) encontraron un total de 33,632 fragmentos de posibles PHA sintasa, sin embargo, solamente se reportó una enzima funcional PHA sintasa que llevaba a la producción de P3HB, P3HV, P4HB, P3H4MV, P5HV y P3HHx a partir de una amplia variedad de sustratos. De esta forma, en este trabajo se presenta el primer reporte de la presencia de PHA sintasas en el suroeste del Golfo de México, lo que representa un sitio potencial para el estudio de polihidroxicanoatos.

Las PHA sintasas localizadas fueron clasificadas como tipo I (k141_122519_1) y la subunidad E de la PHA sintasa tipo III (k141_931424_1 y k141_931424_2). En el caso de las PHA sintasas que presentan una sola subunidad, esta misma presenta los dos grupos tioles a partir de la Cys-319 lo cual permite su actividad catalítica. En el caso de las PHA sintasas con dos subunidades (Clase III y IV), se especula que además de la cisteína presente en la subunidad C, se requiere un segundo grupo tiol proveniente de una cisteína localizada en la subunidad PhaE o PhaR (según sea el caso) (Sudesh *et al.*, 2000). Por esta razón y debido a que no se localizó la subunidad C de la PHA sintasa tipo III, se puede deducir que no se podría llevar a cabo la síntesis de polihidroxicanoatos. Por lo tanto, se espera que la PHA sintasa tipo I encontrada en el metagenoma de sedimento lleve a la producción de polihidroxicanoatos de cadena corta.

La enzima clave para la síntesis de polihidroxicanoatos fue localizada en el Bin 91, del cual se logró reconstruir el 82.7% del genoma (6013805 pb) (Fig. 5). Debido a la falta de reportes de esta bacteria en particular, su genoma y metabolismo enfocado a la producción de polihidroxicanoatos, en este trabajo se le nombro como PHAZGoM18. Así mismo, a partir del análisis por GTDB-Tk classify v0.14 se demostró que pertenece al orden Dehalococcoidales, el cual ha sido asociado con la degradación de hidrocarburos poliaromáticos (Oni *et al.*, 2015; Yan *et al.*, 2019)

y compuestos orgánicos halogenados (Löffler *et al.*, 2013; Matturro *et al.*, 2017) en sedimentos.

A pesar de existir pocos reportes sobre los miembros de este orden, se conoce que son bacterias anaerobias con un metabolismo estricto de respiración de organohaluros, no móviles, en forma de disco con hendiduras en los lados planos opuestos de la célula, con un tamaño de 0.3 -1 μm de diámetro y de 0.1 a 0.2 μm de espesor (Löffler *et al.*, 2013). Aunado a esto, en este trabajo se presenta el primer reporte del metabolismo asociado a la producción de polihidroxicanoatos en este orden. Una vez realizado la reconstrucción metabólica, se confirmó que el genoma PHAZGoM18 presenta las enzimas intermediarias y la PHA sintasa tipo I que conlleva a la producción de poli(3-hidroxiбутирато).

La reconstrucción metabólica presentada en la Fig. 8, indica que el genoma PHAZGoM18 puede utilizar tanto sustratos de hidrocarburos como de carbohidratos para la síntesis de poli(3-hidroxiбутирато). En la región suroeste del Golfo de México se ha reportado la presencia de hidrocarburos en un rango de 56.07-125.44 $\mu\text{g}/\text{kg}$ para hidrocarburos aromáticos y 0.47-6363.72 $\mu\text{g}/\text{kg}$ para hidrocarburos alifáticos (Godoy-Lozano *et al.*, 2018). Así mismo se ha estimado que, el 25% de las manchas de petróleo naturales totales en Golfo de México han sido observadas en esta región (MacDonald *et al.*, 2015). No obstante, la biosíntesis y catabolismo de metabolitos secundarios es baja (Fig. 7). Para confirmar si existe un uso preferencial de este tipo de sustrato o bien, si esta concentración es suficiente para favorecer la producción de polihidroxicanoatos, es necesario realizar posteriores ensayos de laboratorio con diferentes sustratos y determinaciones fisicoquímicas del biopolímero.

En este trabajo no fue posible identificar a otros genes relacionados con la síntesis de poli(3-hidroxiбутирато) cercanos a las PHA sintasa (Tipo I y subunidad E de la PHA sintasa tipo III) en forma de clúster (Fig. 6). Esto puede ser debido a que es la primera vez que se reporta el genoma de esta bacteria y no existe información de sus genes y enzimas en las bases de referencia. En un estudio realizado por Tan *et al.*, (2020) en *Pseudomonas* sp. fueron identificados el clúster tipo I con los genes

phaC, *phaR*, *phaP*, *P3HBR*, *phaB* y *phaA* que conllevan a la síntesis de polihidroxicanoatos de cadena corta; y el clúster tipo II con los genes *phaZ*, *phaD*, *phaF* y *phal*, *phaC1* y *phaC2* que llevan a la síntesis de polihidroxicanoatos de cadena mediana. Además, se identificó el clúster I para *Janthinobacterium* sp. Por otro lado, en un estudio en organismos de la clase *Betaproteobacteria* de igual forma, fueron identificados dos clústers. El primero incluye genes asociados a *phaC1*, mientras que, el segundo incluye genes relacionados con *phaC2* y *phaC3* (Kutralam-Muniasamy *et al.*, 2018).

17.2 Potencial metabólico para la producción de polihidroxicanoatos por un consorcio microbiano degradador de hidrocarburos

El software PICRUSt2 no mostró la predicción de una ruta metabólica específica para la producción de polihidroxicanoatos. No obstante, se identificaron rutas metabólicas que proveen de intermediarios para la PHA sintasa. Entre ellas se encuentra la β -oxidación y la elongación de ácidos grasos, implicadas en la biosíntesis de polihidroxicanoatos de cadena mediana en bacterias (Kim *et al.*, 2007). En el caso de la β -oxidación, la mayor abundancia se observó a las 72 horas con un 1.395% (Tabla 1). Este dato coincide con la mayor abundancia de la oxidación del octano con un 1.264% que también ocurrió a las 72 horas. En ambos casos, a partir de este día, se observa una disminución en su abundancia hasta el día 30. Por otro lado, la abundancia de la elongación de los ácidos grasos se mantuvo entre 0.830-0.902% durante los 30 días del ensayo. Estos datos, sugieren que la degradación de petróleo ligero se da principalmente al tercer día de incubación, sin embargo, este trabajo no cuenta con datos experimentales necesarios para confirmar esta predicción.

En este trabajo, se identificaron dos enzimas enoil-CoA hidratasa (EC 4.2.1.17, EC 4.2.1.119), la cual cataliza la reacción de *trans*-2-enoil-CoA a (*R*)-3-hidroxiacil-CoA (Kim *et al.*, 2007; Kihara *et al.*, 2017). Esto sugiere que como parte de la degradación de los hidrocarburos por β -oxidación, se pueden obtener el

intermediario (*R*)-3-hidroxiacil-CoA. A pesar de tener dos de las enzimas involucradas en la síntesis de polihidroxicanoatos de cadena mediana, no se puede confirmar la producción de este tipo de bioplástico, es necesario de datos experimentales. Esto debido a que las predicciones cuantitativas dependen de la abundancia de miembros de la comunidad, que puede verse afectado por la variación del número de copias del gen 16S ARNr, además, puede existir un sesgo y baja resolución ya que PICRUST2 se basa en genomas de referencia. Así mismo, se ha reportado que la exactitud de las funciones relacionadas con la producción de metabolitos secundarios (como los polihidroxicanoatos) es menor. Moskowitz & Merrick (1969), de igual forma, reportaron la presencia de dos de estas enzimas en la bacteria *Rhodospirillum rubrum* que conllevan a la síntesis de β -hidroxibutiril-CoA. Así mismo, diferentes tipos de esta enzima específica de cadena corta (C4-C6) y cadena mediana (C8-C12) han sido reportados Kihara *et al.*, (2017). En un estudio realizado por Hisano *et al.*, (2001) mostraron que la enzima EC 4.2.1.119 muestra actividad hacia sustratos de cadena corta en la bacteria *Aeromonas caviae*, sin embargo, no existen reportes sobre la especificidad de EC 4.2.1.17.

La enzima (*R*)-3-hidroxiacil-ACP-CoA transferasa (PhaG), presentó la mayor abundancia al tercer día con 0.042% (Tabla 4). Esta enzima ha sido reportada en el género *Pseudomonas*, en el cual es la conexión de la biosíntesis de grasos de *novo* con la síntesis de polihidroxicanoatos de cadena mediana a partir de fuentes de carbono como glucosa, gluconato o etanol (Kim *et al.*, 2007; Gutiérrez-Gómez *et al.*, 2019; Mozejko-Ciesielska *et al.*, 2019). En este trabajo, los datos provienen de un consorcio microbiano creciendo con petróleo ligero como única fuente de carbono, por lo tanto, estos resultados sugieren que para la biosíntesis de ácidos grasos se está utilizando el acetyl-CoA liberado a través de cada ciclo de β -oxidación (Babel & Steinbuchel, 2001). Además, la enzima enoil-CoA hidratasa (reportada por duplicado) y enoil-CoA hidratasa 2, se reportaron con una abundancia promedio de 0.1298%, 0.0291% y 0.4676%, mientras que, solamente se reportó una abundancia promedio de 0.0313% para (*R*)-3-hidroxiacil-ACP-CoA transferasa. Esto indica que la principal ruta involucrada en la síntesis del intermediario (*R*)-3-hidroxiacil-CoA es

la β -oxidación, siendo las 72 horas donde se podría generar la mayor producción, de acuerdo con las abundancias reportadas.

La PHA sintasa tipo II, es la enzima que puede llevar a la síntesis de polihidroxicanoatos de cadena mediana. En la Tabla 2, se observó la presencia de la enzima E.C 2.3.1.-. No obstante, la predicción por el software PICRUST2, no determinó el tipo de PHA sintasa encontrada. Por lo tanto, a pesar de tener dos de las enzimas involucradas en la síntesis de polihidroxicanoatos de cadena mediana, no se puede confirmar la producción de este tipo de bioplástico, es necesario de datos experimentales.

Por otro lado, se realizó la búsqueda de las enzimas involucradas en la síntesis de polihidroxicanoatos de cadena corta. En las Tablas 2 y 3, se reporta la enzima acetil-CoA acetiltransferasa (K00626, E.C 2.3.1.9, sinónimo beta-cetotiolasa, *phaA*), con la mayor abundancia registrada al tercer día de incubación con 0.622% y 0.307%, respectivamente. Además, la enzima acetoacetil-CoA reductasa (E.C 1.1.1.36, K00023, *phaB*) presentó la mayor abundancia al tiempo cero (0.025%). Por otra parte, en la tabla 4, se observa la presencia de la subunidad *phaC* (K05560) y *phaE* (K05562). Diversos autores han indicado que los genes típicos involucrados en la biosíntesis de polihidroxicanoatos de cadena corta son *phaA*, *phaB* y *phaC*. Así mismo, la presencia de estas dos subunidades da lugar a la PHA sintasa tipo III, la cual tiene preferencia por utilizar monómeros de tres a cinco átomos de carbono (Babel & Steinbuchel, 2001; Revelo-Romo *et al.*, 2014; Kutralam-Muniasamy *et al.*, 2017; Thomas *et al.*, 2019; Andler *et al.*, 2019; Tan *et al.*, 2020). Con esto, se sugiere que el consorcio microbiano marino aislado de sedimentos del Golfo de México puede tener la capacidad de producir polihidroxicanoatos de cadena corta.

Para finalizar, la enzima poli(3)-hidroxibutirato depolimerasa (EC:3.1.1.75, K05973) presenta una abundancia de 0.017% al tercer día de incubación que aumenta hasta llegar a 0.022% al final del ensayo. Ong *et al.*, (2017) menciona que la actividad de esta enzima aumenta después del decaimiento en la fuente de carbono. Por lo tanto, se sugiere que al inicio del ensayo de la síntesis de polihidroxicanoato

(principalmente a las 72 horas) después, al ya no existir fuente de carbono disponible para las bacterias, se da la hidrolisis de los gránulos para obtener carbono y energía.

18 CONCLUSIONES

- En este trabajo se logró identificar y ensamblar el genoma de una bacteria no cultivable perteneciente al Orden Dehalococcoidales (nombrada como PHAZGoM18), a partir de los datos del metagenoma de la muestra de sedimento de la estación D18 en el suroeste del Golfo de México. Además, se encontró que esta bacteria presenta la enzima PHA sintasa tipo I y el resto de las enzimas necesarias para llevar a cabo la síntesis de poli(3-hidroxi-butirato). La reconstrucción metabólica permitió reconocer que la bacteria PHAZGoM18 puede ser capaz de utilizar carbohidratos y ácidos grasos como sustratos para la acumulación de este bioplástico.
- En los datos del metagenoma de la muestra de columna de agua no se lograron identificar microorganismos productores de polihidroxi-alcanoatos. Sin embargo, esto puede estar relacionado con la baja cobertura de secuenciación obtenida. Por lo tanto, estudios posteriores con mayor esfuerzo de secuenciación serán necesarios para explorar la diversidad de microorganismos productores de bioplástico en muestras de columna de agua.
- Este trabajo demostró que el consorcio microbiano degradador de petróleo ligero, aislado del Golfo de México, presenta las enzimas necesarias para la síntesis de polihidroxi-alcanoatos de cadena corta, y se infirió que la mayor síntesis de PHAs se lleva a cabo a las 72 horas de incubación en presencia de petróleo ligero. Sin embargo, son necesarios experimentos de laboratorio posteriores para validar estos resultados y así como para conocer las propiedades fisicoquímicas del biopolímero.
- Los resultados de este trabajo exhiben el potencial metabólico que tienen algunos microorganismos bentónicos del suroeste del Golfo de México, para la producción de polihidroxi-alcanoatos con propiedades fisicoquímicas de interés biotecnológico que tendrán que ser evaluadas posteriormente en experimentos de laboratorio.

19 LITERATURA CITADA

- Alma'abadi, A. D., Gojobori, T., y Mineta, K. 2015. Marine metagenome as a resource for novel enzymes. *Genomics, Proteomics and Bioinformatics*, 13(5), 290–295. <https://doi.org/10.1016/j.gpb.2015.10.001>
- Amelia, T. S. M., Amirul, A. A. A., y Bhubalan, K. 2018. Data on partial polyhydroxyalkanoate synthase genes (*phaC*) mined from *Aaptos aaptos* marine sponge-associated bacteria metagenome. *Data in Brief*, 16, 75–80. <https://doi.org/10.1016/j.dib.2017.11.011>
- Andler, R., Vivod, R., y Steinbüchel, A. 2019. Synthesis of polyhydroxyalkanoates through the biodegradation of poly(cis-1,4-isoprene) rubber. *Journal of Bioscience and Bioengineering*, 127(3), 360–365. <https://doi.org/10.1016/j.jbiosc.2018.08.015>
- Arumugam, A. 2019. Polyhydroxyalkanoates (PHA) production. In *Encyclopedia of Renewable and Sustainable Materials* (Elsevier I, pp. 236–252). <https://doi.org/10.1016/b978-0-12-803581-8.10571-5>
- Auta, H. S., Emenike, C. U., y Fauziah, S. H. 2017. Distribution and importance of microplastics in the marine environment. A review of the sources, fate, effects, and potential solutions. *Environment International*, 102, 165–176. <https://doi.org/10.1016/j.envint.2017.02.013>
- Babel, W., y Steinbüchel, A. 2001. Biopolyesters. *Advances in Biochemical Engineering/Biotechnology*. Spain: Springer-Verlag Berlin Heidelberg. <https://doi.org/10.1017/CBO9781107415324.004>
- Bhat, S., y Nambisan, P. 2014. Bacterial polyhydroxyalkanoates production and its applications. Retrieved from http://www.researchgate.net/publication/283568901_Bacterial_Polyhydroxyalkanoates_production_and_its_Applications
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., y Madden, T. L. 2009. BLAST+: Architecture and applications. *BMC*

Bioinformatics, 10, 1–9. <https://doi.org/10.1186/1471-2105-10-421>

Camacho, C., Madden, T., Tao, T., Agarwala, R., y Morgulis, A. 2019. BLAST command line applications user manual. *Bethesda (MD): National Center for Biotechnology Information (US)*, (MD), 1–28. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK279690/>

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., y Parks, D. H. 2020. GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*, 36(6), 1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>

Cheema, S., Bassas-Galia, M., Sarma, P. M., Lal, B., y Arias, S. 2011. Exploiting metagenomic diversity for novel polyhydroxyalkanoate synthases: Production of a terpolymer poly(3-hydroxybutyrate-co-3-hydroxyhexanoate-co-3-hydroxyoctanoate) with a recombinant *Pseudomonas putida* strain. *Bioresource Technology*, 103(1), 322–328. <https://doi.org/10.1016/j.biortech.2011.09.098>

Covarrubias, M. 2020. Caracterización de las comunidades microbianas en sedimentos del Golfo de México, mediante análisis metagenómico. Tesis de Maestría. Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California. 82 pp.

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., ... Langille, M. G. I. 2020. PICRUSt2 for prediction of metagenome functions. *Nature Biotechnology*, 38(6), 685–688. <https://doi.org/10.1038/s41587-020-0548-6>

Foong, C. P., Lakshmanan, M., Abe, H., Taylor, T. D., Foong, S. Y., y Sudesh, K. 2017. A novel and wide substrate specific polyhydroxyalkanoate (PHA) synthase from unculturable bacteria found in mangrove soil. *Journal of Polymer Research*, 25(1). <https://doi.org/10.1007/s10965-017-1403-4>

Foong, C. P., Lau, N. S., Deguchi, S., Toyofuku, T., Taylor, T. D., Sudesh, K., y Matsui, M. 2014. Whole genome amplification approach reveals novel polyhydroxyalkanoate synthases (PhaCs) from Japan Trench and Nankai Trough seawater. *BMC Microbiology*, 14(1), 1–11.

<https://doi.org/10.1186/s12866-014-0318-z>

- García-Cruz, N. U., Canul-Puc, R., Sanchez-Avila, J., Gold-Bouchot, G., y Aguirre-Macedo, M. L. 2014. Potential crude oil degradation by a microbial consortium isolated from the Gulf of Mexico. *Environmental Biotechnology and Engineering*, (January 2016), 684–692. Retrieved from https://www.researchgate.net/publication/272151561_Potential_crude_oil_degradation_by_a_microbial_consortium_isolated_from_the_Gulf_of_Mexico
- Ghavipankeh, F., Pazouki, M., Ziaei Rad, Z., y Hosseinnia, A. 2015. Biological conversion of normal chain octadecane by native microbial consortia. *Iranian Journal of Chemical Engineering*, 12(2), 50–58.
- Godoy-Lozano, E. E., Escobar-Zepeda, A., Raggi, L., Merino, E., Gutierrez-Rios, R. M., Juarez, K., ... Pardo-Lopez, L. 2018. Bacterial diversity and the geochemical landscape in the southwestern Gulf of Mexico. *Frontiers in Microbiology*, 9(OCT), 1–15. <https://doi.org/10.3389/fmicb.2018.02528>
- González, Y., Meza, J., González, O., y Córdova, J. 2013. Síntesis y biodegradación de polihidroxicanoatos plásticos de origen microbiano. *Rev. Int. Contam. Ambie.*, 29(1), 77–115.
- Gutiérrez-Gómez, U., Servín-González, L., y Soberón-Chávez, G. 2019. Role of β -oxidation and de novo fatty acid synthesis in the production of rhamnolipids and polyhydroxyalkanoates by *Pseudomonas aeruginosa*. *Applied Microbiology and Biotechnology*, (L). <https://doi.org/10.1007/s00253-019-09734-x>
- Heidelberg, K. B., Gilbert, J. A., y Joint, I. 2010. Marine genomics: At the interface of marine microbial ecology and biodiscovery: Minireview. *Microbial Biotechnology*, 3(5), 531–543. <https://doi.org/10.1111/j.1751-7915.2010.00193.x>
- Hernández-Núñez, E., Martínez-Gutiérrez, C. A., López-Cortés, A., Aguirre-Macedo, M. L., Tabasco-Novelo, C., González-Díaz, M. O., y García-Maldonado, J. Q. 2019. Physico-chemical characterization of poly(3-Hydroxybutyrate) produced by *Halomonas salina*, isolated from a hypersaline microbial mat. *Journal of*

Polymers and the Environment, 27(5), 1105–1111.
<https://doi.org/10.1007/s10924-019-01417-y>

Hernández, J. 2015. Identificación de cepas nativas de sustancias húmicas y suelos altamente impactados por hidrocarburos. Tesis de Licenciatura. Unidad Profesional Interdisciplinaria de Biotecnología. Instituto Politécnico Nacional. 57 pp.

Himmel, M., y Bomble, Y. 2020. Metabolic pathway engineering. *Methods in Molecular Biology* (Vol. 2096). Golden, CO, USA. https://doi.org/10.1007/978-1-0716-0195-2_13

Hisano, T., Fukui, T., Iwata, T., y Doi, Y. 2001. Crystallization and preliminary X-ray analysis of (R)-specific enoyl-CoA hydratase from *Aeromonas caviae* involved in polyhydroxyalkanoate biosynthesis. *Acta Crystallographica Section D: Biological Crystallography*, 57(1), 145–147.
<https://doi.org/10.1107/S0907444900014062>

Horton, A. A., Walton, A., Spurgeon, D. J., Lahive, E., y Svendsen, C. 2017. Microplastics in freshwater and terrestrial environments: Evaluating the current understanding to identify the knowledge gaps and future research priorities. *Science of the Total Environment*, 586, 127–141.
<https://doi.org/10.1016/j.scitotenv.2017.01.190>

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., y Hauser, L. J. 2010. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11. <https://doi.org/10.1186/1471-2105-11-119>

Jeske, L., Placzek, S., Schomburg, I., Chang, A., y Schomburg, D. 2019. BRENDA in 2019: A European ELIXIR core data resource. *Nucleic Acids Research*, 47(D1), D542–D549. <https://doi.org/10.1093/nar/gky1048>

Jiang, Hill, D. J., Kowalczyk, M., Johnston, B., Adamus, G., Irorere, V., y Radecka, I. 2016. Carbon sources for polyhydroxyalkanoates and an integrated

- Biorefinery. *International Journal of Molecular Sciences*, (17), 1157–1177. <https://doi.org/10.3390/ijms17071157>
- Jo, J., Oh, J., y Park, C. 2020. Microbial community analysis using high-throughput sequencing technology: a beginner's guide for microbiologists. *Journal of Microbiology*, 58(3), 176–192. <https://doi.org/10.1007/s12275-020-9525-5>
- Kamravamanesh, D., Lackner, M., y Herwig, C. 2018. Bioprocess engineering aspects of sustainable polyhydroxyalkanoate production in cyanobacteria. *Bioengineering*, 5(4), 1–18. <https://doi.org/10.3390/BIOENGINEERING5040111>
- Kanehisa, M., y Goto, S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(4), 27–30. <https://doi.org/10.3892/ol.2020.11439>
- Kihara, T., Hiroe, A., Ishii-Hyakutake, M., Mizuno, K., y Tsume, T. 2017. *Bacillus cereus*-type polyhydroxyalkanoate biosynthetic gene cluster contains R-specific enoyl-CoA hydratase gene. *Bioscience, Biotechnology and Biochemistry*, 81(8), 1627–1635. <https://doi.org/10.1080/09168451.2017.1325314>
- Kim, D. Y., Hyung, W. K., Moon, G. C., y Young, H. R. 2007. Biosynthesis, modification, and biodegradation of bacterial medium-chain-length polyhydroxyalkanoates. *Journal of Microbiology*, 45(2), 87–97.
- Kumar, V., Kumar, S., y Singh, D. 2020. Microbial polyhydroxyalkanoates from extreme niches: Bioprospection status, opportunities and challenges. *International Journal of Biological Macromolecules*, 147, 1255–1267. <https://doi.org/10.1016/j.ijbiomac.2019.09.253>
- Kumari, A., Chaudhary, D. R., y Jha, B. 2019. Destabilization of polyethylene and polyvinylchloride structure by marine bacterial strain. *Environmental Science and Pollution Research*, 26(2), 1507–1516. <https://doi.org/10.1007/s11356-018-3465-1>
- Kutralam-Muniasamy, G., Corona-Hernandez, J., Narayanasamy, R. K., Marsch, R., y Pérez-Guevara, F. 2017. Phylogenetic diversification and developmental

- implications of poly-(R)-3-hydroxyalkanoate gene cluster assembly in prokaryotes. *FEMS Microbiology Letters*, 364(13), 1–28. <https://doi.org/10.1093/femsle/fnx135>
- Kutralam-Muniasamy, G., Marsch, R., y Pérez-Guevara, F. 2018. Investigation on the evolutionary relation of diverse polyhydroxyalkanoate gene clusters in Betaproteobacteria. *Journal of Molecular Evolution*, 86(7), 470–483. <https://doi.org/10.1007/s00239-018-9859-3>
- Lee, S. Y. 1996. Plastic bacteria? Progress and prospects for polyhydroxyalkanoate production in bacteria. *Trends in Biotechnology*, 14(11), 431–438. [https://doi.org/10.1016/0167-7799\(96\)10061-5](https://doi.org/10.1016/0167-7799(96)10061-5)
- Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., ... Lam, T. W. 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, 102, 3–11. <https://doi.org/10.1016/j.ymeth.2016.02.020>
- Löffler, F. E., Yan, J., Ritalahti, K. M., Adrian, L., Edwards, E. A., Konstantinidis, K. T., ... Spormann, A. M. 2013. *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, Dehalococcoidia classis nov., order Dehalococcoidales ord. nov. and famil. *International Journal of Systematic and Evolutionary Microbiology*, 63(PART2), 625–635. <https://doi.org/10.1099/ijs.0.034926-0>
- López, U. 2016. Las técnicas de secuenciación masiva en el estudio de la diversidad biológica. *Munibe Ciencias Naturales*, 64(January). <https://doi.org/10.21630/mcn.2016.64.07>
- MacDonald, I., Garcia-Pineda, O., Beet, A., Daneshgar, S., Feng, L., Graettinger, G., ... Swayze, G. 2015. Natural and unnatural oil slicks in the Gulf of Mexico. *Journal of Geophysical Research: Oceans*, 120, 8364–8380. <https://doi.org/10.1002/2015JC011062>.Received
- Madkour, M. H., Heinrich, D., Alghamdi, A. M., Shabbaj, I. I., y Steinbuchel, A. 2013.

- PHA Recovery from Biomass. *Biomacromolecules*, 14(9), 2963-2972. [dx.doi.org/10.1021/bm4010244](https://doi.org/10.1021/bm4010244)
- Manilla-Pérez, E., Lange, A. B., Hetzler, S., y Steinbüchel, A. 2010. Occurrence, production, and export of lipophilic compounds by hydrocarbonoclastic marine bacteria and their potential use to produce bulk chemicals from hydrocarbons. *Applied Microbiology and Biotechnology*, 86(6), 1693–1706. <https://doi.org/10.1007/s00253-010-2515-5>
- Mannina, G., Presti, D., Montiel-Jarillo, G., Carrera, J., y Suárez-Ojeda, M. E. 2020. Recovery of polyhydroxyalkanoates (PHAs) from wastewater: A review. *Bioresource Technology*, 297, 122478. <https://doi.org/10.1016/j.biortech.2019.122478>
- Martínez-Gutiérrez. 2018. Aislamiento y análisis independiente de cultivo de comunidades productoras de polihidroxicanoatos y degradadoras de n-Hexadecano en tapetes microbianos de Guerrero Negro, México. Tesis de maestría. Centro de Investigaciones Biológicas del Noroeste. 89 pp.
- Martínez-Gutiérrez, C. A., Latisnere-Barragán, H., García-Maldonado, J. Q., y López-Cortés, A. 2018. Screening of polyhydroxyalkanoate-producing bacteria and PhaC-encoding genes in two hypersaline microbial mats from Guerrero Negro, Baja California Sur, Mexico. *PeerJ*, 2018(5), 1–19. <https://doi.org/10.7717/peerj.4780>
- Matturo, B., Frascadore, E., y Rossetti, S. 2017. High-throughput sequencing revealed novel Dehalococcoidia in dechlorinating microbial enrichments from PCB-contaminated marine sediments. *FEMS Microbiology Ecology*, 93(11), 1–10. <https://doi.org/10.1093/femsec/fix134>
- Mezzina, M. P., y Pettinari, M. J. 2016. Phasins, multifaceted polyhydroxyalkanoate granule-associated proteins. *Applied and Environmental Microbiology*, 82(17), 5060–5067. <https://doi.org/10.1128/AEM.01161-16>
- Mineta, K., y Gojobori, T. 2016. Databases of the marine metagenomics. *Gene*, 576(2), 724–728. <https://doi.org/10.1016/j.gene.2015.10.035>

- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., ... Bateman, A. 2021. Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1), D412–D419. <https://doi.org/10.1093/nar/gkaa913>
- Moskowitz, G. J., y Merrick, J. M. 1969. Metabolism of poly- β -hydroxybutyrate. II. Enzymatic synthesis of D-(-)- β -Hydroxybutyryl coenzyme A by an enoyl Hydrase from *Rhodospirillum rubrum*. *Biochemistry*, 8(7), 2748–2755. <https://doi.org/10.1021/bi00835a009>
- Mozejko-Ciesielska, J., Szacherska, K., y Marciniak, P. 2019. *Pseudomonas* species as producers of eco-friendly polyhydroxyalkanoates. *Journal of Polymers and the Environment*. <https://doi.org/10.1007/s10924-019-01422-1>
- Mwamburi, S. M., Mbatia, B. N., Remmy, K., Kirwa, E. M., & Noah, N. M. 2019. Production of polyhydroxyalkanoates by hydrocarbonoclastic bacteria. *African Journal of Biotechnology*, 18(17), 352–364. <https://doi.org/10.5897/ajb2019.16763>
- Narancic, T., Kenny, S. T., Djokic, L., Vasiljevic, B., O'Connor, K. E., y Nikodinovic-Runic, J. 2012. Medium-chain-length polyhydroxyalkanoate production by newly isolated *Pseudomonas* sp. TN301 from a wide range of polyaromatic and monoaromatic hydrocarbons. *Journal of Applied Microbiology*, 113(3), 508–520. <https://doi.org/10.1111/j.1365-2672.2012.05353.x>
- Ngara, T. R., y Zhang, H. 2018. Recent advances in function-based metagenomic screening. *Genomics, Proteomics and Bioinformatics*, 16(6), 405–415. <https://doi.org/10.1016/j.gpb.2018.01.002>
- Ong, S. Y., Chee, J. Y., y Sudesh, K. 2017. Degradation of polyhydroxyalkanoate (PHA): a review. *Journal of Siberian Federal University. Biology*, 10(2), 21–225. <https://doi.org/10.17516/1997-1389-0024>
- Oni, O. E., Schmidt, F., Miyatake, T., Kasten, S., Witt, M., Hinrichs, K. U., y Friedrich, M. W. 2015. Microbial communities and organic matter composition in surface and subsurface sediments of the Helgoland mud area, North Sea. *Frontiers in*

- Microbiology*, 6(NOV), 1–16. <https://doi.org/10.3389/fmicb.2015.01290>
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., y Tyson, G. W. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, 25(7), 1043–1055. <https://doi.org/10.1101/gr.186072.114>
- Pearson, W. 2013. An Introduction to Sequence Similarity (“Homology”) Searching. *Curr Protoc Bioinformatics.*, 1(10), 1286–1292. <https://doi.org/10.1002/0471250953.bi0301s42.An>
- PlasticsEurope. 2018. *The facts 2018. PlasticsEurope.* Europe. <https://doi.org/10.1016/j.marpolbul.2013.01.015>
- Ponnusamy, S., Viswanathan, S., Periyasamy, A., y Rajaiah, S. 2019. Production and characterization of PHB-HV copolymer by *Bacillus thuringiensis* isolated from *Eisenia foetida*. *Biotechnology and Applied Biochemistry*, 66(3), 340–352. <https://doi.org/10.1002/bab.1730>
- Raggi, L., García-Guevara, F., Godoy-Lozano, E. E., Martínez-Santana, A., Escobar-Zepeda, A., Gutierrez-Rios, R. M., ... Juarez, K. 2020. Metagenomic profiling and microbial metabolic potential of Perdido Fold Belt (NW) and Campeche Knolls (SE) in the Gulf of Mexico. *Frontiers in Microbiology*, 11(August), 1–18. <https://doi.org/10.3389/fmicb.2020.01825>
- Revelo-Romo, D. M., Obando, M. C., Gómez-Perdomo, M. I., Fernández Izquierdo, P. 2014. PHA synthase genes in oil-degrading marine bacteria from the Colombian Pacific Coast. *International Journal of Advanced Biotechnology and Research*, (February), 976–2612. Retrieved from <http://www.bipublication.com>
- Sabirova, J. 2010. Handbook of Hydrocarbon and Lipid Microbiology. Handbook of hydrocarbon and lipid microbiology. Berlin: Springer-Verlag. <https://doi.org/10.1007/978-3-540-77587-4>
- Sangkharak, K., Choonut, A., Rakkan, T., y Prasertsan, P. 2020. The degradation of Phenanthrene, Pyrene, and Fluoranthene and its conversion into medium-

chain-Length polyhydroxyalkanoate by novel polycyclic aromatic hydrocarbon-Degrading Bacteria. *Current Microbiology*, (0123456789).
<https://doi.org/10.1007/s00284-020-01883-x>

Sathiyarayanan, G., Saibaba, G., Kiran, G. S., Yang, Y. H., y Selvin, J. 2016. Marine sponge-associated bacteria as a potential source for polyhydroxyalkanoates. *Critical Reviews in Microbiology*, 43(3), 294–312.
<https://doi.org/10.1080/1040841X.2016.1206060>

Serafim, L. S., Lemos, P. C., Torres, C., Reis, M. A. M., y Ramos, A. M. 2008. The influence of process parameters on the characteristics of polyhydroxyalkanoates produced by mixed cultures. *Macromolecular Bioscience*, 8(4), 355–366. <https://doi.org/10.1002/mabi.200700200>

Sin, M. C., Gan, S. N., Annuar, M. S. M., y Tan, I. K. P. 2010. Thermodegradation of medium-chain-length poly(3-hydroxyalkanoates) produced by *Pseudomonas putida* from oleic acid. *Polymer Degradation and Stability*, 95(12), 2334–2342.
<https://doi.org/10.1016/j.polymdegradstab.2010.08.027>

Sudesh, K., Abe, H., y Doi, Y. 2000. Synthesis, structure and properties of polyhydroxyalkanoates: Biological polyesters. *Progress in Polymer Science (Oxford)*, 25(10), 1503–1555. [https://doi.org/10.1016/S0079-6700\(00\)00035-6](https://doi.org/10.1016/S0079-6700(00)00035-6)

Sun, S., Jones, R. B., y Fodor, A. A. 2020. Inference-based accuracy of metagenome prediction tools varies across sample types and functional categories. *Microbiome*, 8(1), 1–9. <https://doi.org/10.1186/s40168-020-00815-y>

Tai, Y. T., Foong, C. P., Najimudin, N., y Sudesh, K. 2016. Discovery of a new polyhydroxyalkanoate synthase from limestone soil through metagenomic approach. *Journal of Bioscience and Bioengineering*, 121(4), 355–364.
<https://doi.org/10.1016/j.jbiosc.2015.08.008>

Tan, I. K. P., Foong, C. P., Tan, H. T., Lim, H., Zain, N. A. A., Tan, Y. C., ... Sudesh, K. 2020. Polyhydroxyalkanoate (PHA) synthase genes and PHA-associated gene clusters in *Pseudomonas* spp. and *Janthinobacterium* spp. isolated from

- Antarctica. *Journal of Biotechnology*, 313(March), 18–28. <https://doi.org/10.1016/j.jbiotec.2020.03.006>
- Thomas, T., Elain, A., Bazire, A., y Bruzard, S. 2019. Complete genome sequence of the halophilic PHA-producing bacterium *Halomonas* sp. SF2003: insights into its biotechnological potential. *World Journal of Microbiology and Biotechnology*, 35(3), 0. <https://doi.org/10.1007/s11274-019-2627-8>
- Tseng, C. H., y Tang, S. L. 2014. Marine microbial metagenomics: From individual to the environment. *International Journal of Molecular Sciences*, 15(5), 8878–8892. <https://doi.org/10.3390/ijms15058878>
- Urbanek, A. K., Rymowicz, W., y Mirończuk, A. M. 2018. Degradation of plastics and plastic-degrading bacteria in cold marine habitats. *Applied Microbiology and Biotechnology*, 102(18), 7669–7678. <https://doi.org/10.1007/s00253-018-9195-y>
- Uribe-Flores, M. M., Cerqueda-García, D., Hernández-Nuñez, E., Cadena, S., García-Cruz, N. U., Trejo-Hernández, M. R., ... García-Maldonado, J. Q. 2019. Bacterial succession and co-occurrence patterns of an enriched marine microbial community during light crude oil degradation in a batch reactor. *Journal of Applied Microbiology*, 127(2), 495–507. <https://doi.org/10.1111/jam.14307>
- Vahlis, P. T. A. 2017. Comparative metagenomics of PHA synthase genes in soil. Tesis de Maestría. University of Waterloo. 78 pp.
- Venkateswar, M., Mawatari, Y., Yajima, Y., Seki, C., Hoshino, T., y Chang, Y. C. 2015. Poly-3-hydroxybutyrate (PHB) production from alkylphenols, mono and poly-aromatic hydrocarbons using *Bacillus* sp. CYR1: A new strategy for wealth from waste. *Bioresource Technology*, 192, 711–717. <https://doi.org/10.1016/j.biortech.2015.06.043>
- Wu, Y.W., Tang, Y.H., Tringe, S. G., Simmons, B. A., y Singer, S. W. 2014. MaxBin: an automated binning method to recover individual genomes from metagenomes using. *Microbiome*, 2(26), 4904–4909. Retrieved from

<https://microbiomejournal.biomedcentral.com/articles/10.1186/2049-2618-2-26>

Yan, Z., Hao, Z., Wu, H., Jiang, H., Yang, M., y Wang, C. 2019. Co-occurrence patterns of the microbial community in polycyclic aromatic hydrocarbon-contaminated riverine sediments. *Journal of Hazardous Materials*, 367, 99–108. <https://doi.org/10.1016/j.jhazmat.2018.12.071>

Zobell, C. E. 1946. Action of microorganism on hydrocarbons. *Bacteriol Rev.*, 10(295), 1–49.

20 ANEXOS

20.1 Manual de ensamble de metagenomas y búsqueda de enzimas clave.

PASO 1: INCISIÓN DEL ADAPTADOR Y ANÁLISIS DE LA CALIDAD CON TRIM_GALORE

Trim_galore un script para controlar y automatizar la calidad y el recorte del adaptador. Para ello, utiliza los primeros 13 pb de los adaptadores estándar de Illumina de forma predeterminada. Permite especificar la calidad Phred, la cual es una medida de calidad en la identificación de las bases obtenidas de la secuenciación, y la extracción del adaptador de forma individual para eliminar lecturas más cortas que lo especificado (Himmel & Bomble, 2020).

```
nohup trim_galore --illumina --path_to_cutadapt /home/alumnos/local/bin/cutadapt  
--paired -q 30 --length 50 --max_n 1 --retain_unpaired --trim-n  
SRR11308316_1.fastq.gz SRR11308316_2.fastq.gz &
```

```
nohup trim_galore --illumina --path_to_cutadapt /home/alumnos/local/bin/cutadapt  
--paired -q 30 --length 50 --max_n 1 --retain_unpaired --trim-n  
SRR11308320_1.fastq.gz SRR11308320_2.fastq.gz &
```

PASO 2: ENSAMBLE DE LOS METAGENOMAS CON MEGAHIT

MEGAHIT es un ensamblador NGS de *nov*o que puede ensamblar conjuntos de datos metagenómicos grandes y complejos de forma eficiente (tiempo y memoria). Hace uso de gráficos sucintos de Brujin (Li *et al.*, 2015, 2016).

```
nohup megahit -1 SRR11308316_1_val_1.fq.gz -2 SRR11308316_2_val_2.fq.gz -r  
SRR11308316_1_unpaired_1.fq.gz,SRR11308316_2_unpaired_2.fq.gz -o
```

```
ensam_SRR11308316 -t 22 --out-prefix SRR11308316 --k-min 21 --k-max 141 --k-  
step 10 -m 0.95 > mhit2.log &
```

```
nohup megahit -1 SRR11308320_1_val_1.fq.gz -2 SRR11308320_2_val_2.fq.gz -r  
SRR11308320_1_unpaired_1.fq.gz,SRR11308320_2_unpaired_2.fq.gz -o  
ensam_SRR11308320 -t 22 --out-prefix SRR11308320 --k-min 21 --k-max 71 --k-  
step 10 -m 0.95 > mhit2.log &
```

PASO 3: BINNING CON MAXBIN 2.2.7

MaxBin automatiza el agrupamiento de andamios metagenómicos ensamblados utilizando un algoritmo “Expectation-maximization”. En este enfoque, las frecuencias tetranucleóticos y las coberturas de andamios se utilizan para organizar las secuencias metagenómicas en Bins individuales, las cuales pueden ser predichas a partir de la identificación inicial de genes marcadores en secuencias ensambladas (Wu *et al.*, 2014).

Como el archivo de los contigs está dentro de una carpeta. Crear un archivo afuera de la misma.

```
In -s /scratch/ANDREA/ensam_SRR11308316/SRR11308316.contigs.fa .
```

```
In -s /scratch/ANDREA/ensam_SRR11308320/SRR11308320.contigs.fa .
```

A continuación:

```
nohup /scratch/ANDREA/MaxBin-2.2.7/run_MaxBin.pl -contig  
SRR11308316.contigs.fa -reads SRR11308316_1_unpaired_1.fq.gz -reads2  
SRR11308316_2_unpaired_2.fq.gz -reads3 SRR11308316_1_val_1.fq.gz -reads4  
SRR11308316_2_val_2.fq.gz -thread 20 -out summary -min_contig_length 200 &
```

```
nohup /scratch/ANDREA/MaxBin-2.2.7/run_MaxBin.pl -contig  
SRR11308320.contigs.fa -reads SRR11308320_1_unpaired_1.fq.gz -reads2
```

```
SRR11308320_2_unpaired_2.fq.gz -reads3 SRR11308320_1_val_1.fq.gz -reads4  
SRR11308320_2_val_2.fq.gz -thread 20 -out binning20 -min_contig_length 200 &
```

PASO 4: BÚSQUEDA DE LOS MARCOS DE LECTURA CON PRODIGAL

Es un algoritmo utilizado para predecir de genes faltantes, el cual utiliza una combinación de Critica, Glimmer y BLAST. Así mismo, corrige errores, reconoce el sitio de inicio de la traducción y reduce los falsos positivos (Hyatt *et al.*, 2010).

```
nohup prodigal -i SRR11308316.contigs. fa -a protein_16.faa -c -p meta &
```

```
nohup prodigal -i SRR11308320.contigs. fa -a protein_20.faa -c -p meta &
```

PASO 5: DESCARGAR SECUENCIAS DE REFERENCIA DEL NCBI

Descargar las secuencias aminoacídicas de referencia de las PHA sintasa del NCBI y copiarlas todas en un archivo formato FASTA utilizando gedit y subirlas a cada una de las carpetas (ensam_SRR11308316 y ensam_SRR11308320) utilizando nano en el servidor.

NOTA: No debe existir líneas en blanco entre las secuencias.

PASO 6: CONVERTIR LOS METAGENOMAS A BASE DE DATOS CON MAKEBLASTDB

```
makeblastdb -in protein_16.faa -out dbmet16 -dbtype prot -parse_seqids
```

```
makeblastdb -in protein_20.faa -out dbmet20 -dbtype prot -parse_seqids
```

PASO 7: BÚSQUEDA DE LAS ENZIMAS DE REFENCIA EN LOS METAGENOMAS CON BLASTP

BLAST es un programa de búsqueda de similitud de secuencias para buscar rápidamente en una base de datos (Camacho *et al.*, 2019).

```
nohup blastp -db dbmet16 -query pha_sintases -out prediccion16 -num_threads 20  
-outfmt 6 -evalue 0.003 &
```

```
nohup blastp -db dbmet20 -query pha_sintases -out prediccion20 -num_threads 20  
-outfmt 6 -evalue 0.003 &
```

PASO 8: FILTRAR SECUENCIAS EN LENGUAJE R.

Descargar el archivo con las PHA sintases predichas y en Lenguaje R, eliminar aquellas con un porcentaje de similitud menor al 40 %. Obtener los IDs de cada una de ellas y subirlas en un archivo a cada una de las carpetas de los ensamblajes utilizando nano (un ID por línea).

```
setwd("C:/Users/Andrea Zambrano/Documents/MobaXterm/home")  
a<- read.table("prediccion16")  
attach(a)  
(b <- a[V3>=40.000,])  
summary(b) #21 secuencias  
(id16 <- b[,2])  
save (id16, file="id16.RData")  
  
c <- read.table("prediccion20")  
attach (c)  
(d <- c[V3>=40.000,])  
summary (d)  
(id20 <- d[,2])
```

```
save (id20, file="id20.RData")
```

PASO 9: FILTRAR LAS SECUENCIAS DE LAS PHA SINTASAS PREDICHAS EN FORMATO FASTA CON FITLER_FASTA.PY

```
filter_fasta.py -f protein_16.faa -o phafiltradas16.fasta -s id16
```

```
filter_fasta.py -f protein_20.faa -o phafiltradas20.fasta -s id20
```

PASO 10: BÚSQUEDA DE DOMINIOS EN HHMER/Pfam

Es una base de datos de familia de proteínas y dominios que se utiliza para analizar genomas y metagenomas. Se basa en un alineamiento “seed”, es decir, consiste en encontrar coincidencias exactas. Un modelo de Markov se construye automáticamente a partir del alineamiento seed y se busca en una base de datos utilizando un software HMMER (Mistry *et al.*, 2021).

Descargar los archivos phafiltradas16.fasta y phafiltradas20.fasta. Ingresar a la siguiente página: <http://pfam.xfam.org/search#tabview=tab1>

- En este caso, solamente cinco enzimas del metagenoma de sedimento presentaron los dominios de una PHA sintasa. Por lo tanto, se copiaron y pegaron en un archivo gedit, manualmente.

PASO 11: BUSCAR LAS ENZIMAS EN LOS BINS

Las PHA sintasas confirmadas fueron buscadas en los Bins utilizando el siguiente comando: `grep -w "nombre"*.fasta` En donde, nombre corresponde al ID de la enzima PHA sintasa.