



# UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS BIOLÓGICAS Y DE LA SALUD  
DEPARTAMENTO DE INVESTIGACIONES CIENTÍFICAS Y  
TECNOLÓGICAS

POSGRADO EN BIOCENCIAS

---

**ANÁLISIS TRANSCRIPTÓMICO DE LA DIATOMEA  
*Chaetoceros muelleri*, BAJO CONDICIONES  
LIMITANTES DE FÓSFORO Y SU RELACIÓN CON  
LA ACUMULACIÓN DE ACEITES**

**TESIS**

que para obtener el grado de:

**DOCTOR EN BIOCENCIAS**

presenta:

**JOSÉ PABLO LOVIO FRAGOSO**

Hermosillo, Sonora, México

8 de enero de 2021

# Universidad de Sonora

Repositorio Institucional UNISON



**"El saber de mis hijos  
hará mi grandeza"**



Excepto si se señala otra cosa, la licencia del ítem se describe como openAccess

Hermosillo, Sonora a 20 de febrero del 2021.

Asunto: Cesión de derechos

**UNIVERSIDAD DE SONORA  
P R E S E N T E.**

Por este conducto hago constar que soy autor y titular de la obra denominada "Análisis transcriptómico de la diatomea *Chaetoceros muelleri*, bajo condiciones limitantes de fósforo y su relación con la acumulación de aceites", en los sucesivo LA OBRA, realizada como trabajo terminal con el propósito de obtener el Grado de Doctor en Biociencias, en virtud de lo cual autorizo a la Universidad de Sonora (UNISON) para que efectúe la divulgación, publicación, comunicación pública, distribución, distribución pública, distribución electrónica y reproducción, así como la digitalización de la misma, con fines académicos o propios de la institución y se integren a los repositorios de la universidad, estatales, regionales, nacionales e internacionales.

La UNISON se compromete a respetar en todo momento mi autoría y a otorgarme el crédito correspondiente en todas las actividades mencionadas anteriormente.

De la misma manera, manifiesto que el contenido académico, literario, la edición y en general cualquier parte de LA OBRA son de mi entera responsabilidad, por lo que deslindo a la UNISON por cualquier violación a los derechos de autor y/o propiedad intelectual y/o cualquier responsabilidad relacionada con la OBRA que cometa el suscrito frente a terceros.

ATENTAMENTE



José Pablo Lovio Fragoso

---

**Nombre y Firma del Autor**

Hermosillo, Sonora, México

Julio, 2020.



LIC. GILBERTO LEÓN LEÓN  
Abogado General  
UNIVERSIDAD DE SONORA

ANÁLISIS TRANSCRIPTÓMICO DE LA DIATOMEA *Chaetoceros muelleri*, BAJO  
CONDICIONES LIMITANTES DE FÓSFORO Y SU RELACIÓN CON LA  
ACUMULACIÓN DE ACEITES

TESIS

Que para obtener el grado de

DOCTOR EN BIOCENCIAS

Presenta

M.C. JOSÉ PABLO LOVIO FRAGOSO

Hermosillo, Sonora, México

08 de enero de 2021

## APROBACIÓN

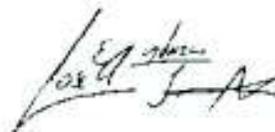
Los miembros del Comité designado para revisar la tesis intitulada “Análisis transcriptómico de la diatomea *Chaetoceros muelleri*, bajo condiciones limitantes de fósforo y su relación con la acumulación de aceites” presentada por M.C. José Pablo Lovio Fragoso, la han encontrado satisfactoria y recomiendan que sea aceptada como requisito parcial para obtener el grado de Doctor en Biociencias con Especialidad en Biotecnología de Recursos Naturales.



---

**Dra. Ángela Corina Hayano Kanashiro**

Directora y Presidenta



---

**Dr. José Antonio López Elías**

Co-Director y Secretario



---

**Dr. Luis Ángel Medina Juárez**

Sinodal interno



---

**Dr. Luis Rafael Martínez Córdova**

Sinodal interno



---

**Dr. Miguel Ángel Hernández Oñate**

Sinodal externo

## DEDICATORIA

A **Dios** y a la **Divina Providencia**, sin afán de blasfemar, gracias por mostrar tu predilección hacia mi persona. Si bien, no me exentas de los momentos de prueba, nunca me has soltado de la mano y siempre me alzas victorioso. Gracias infinitas por colmarme de amor y de bendiciones; por hacerme fuerte y mantener mi alma intacta a pesar de los tiempos adversos y de prueba. Ruego por siempre ser instrumento tuyo y testimonio vivo y en acción de tu existencia, así como a darte gloria con mi vivir en donde Tú quieras que esté.

A **mis padres y mi hermana Alexia**, quiero que sepan que lo que más valoro y lo que es más sagrado en mi vida ustedes me lo dieron. Cada uno a su forma y con todo el amor me lo han dado todo. No me alcanza la vida para agradecerles tanto, tanto. Los amo con el alma.

A **mi familia** de sangre y de espíritu, en especial a **Germán Martín Fragozo Robles**, quien antes de partir, hizo el bien y llenó de luz a cada uno de nosotros. En mi caso, dejaste una estrella inmortal que trajo su resplandor en medio de mucha oscuridad y que siempre alumbrará mi camino; eres el responsable de que mi vida sea mejor y más plena.

A **Ángela Corina Hayano Kanashiro, José Antonio López Elías, Ana Edilia Alvarado Rosas, Laura Elisa Gassós Ortega, Olga Lidia Tavares Sánchez y Leonardo Miguel Aguirre Gallardo** quiénes han contribuido en gran medida en mi formación académica, profesional, humana, espiritual y en mi despertar de conciencia. Gracias por su guía, enseñanzas y por día con día fortalecer mi corazón y espíritu.

A **Luis Gustavo Hodgers Robles y Víctor Espinoza Gómez** quiénes me enseñaron el valor del tiempo, de la vida y a no dejar para después el dar a quiénes amamos lo mejor de nosotros mismos. Un abrazo en donde quiera que estén.

A **todas y cada una de las personas** que han dejado una huella profunda en mi alma, a quienes me han ayudado a salir adelante en cada etapa de mi vida y me han hecho ser el hombre que hoy soy.

También dedico este trabajo a todas aquellas personas que me dijeron: “No puedes”, “No eres suficiente”, “No sirves para esto”; porque me di cuenta de que al único a quien tengo que hacer feliz es a mí mismo, a quien tengo que dar cuentas es a Dios y mi conciencia y que mi único rival está en el espejo. Al sumar cada avance, cada mejora en mi persona, cada victoria, cada “raya al tigre” y cada aprendizaje, el total da igual a una gloria personal que nada ni nadie me podrá quitar.

Y por último y no por eso menos importante, a *ti*, que me acompañaste en gran parte de este hermoso proyecto. Gracias por tu tiempo, tus sonrisas, tu confianza, comprensión, apoyo, cariño, ocurrencias, detalles e invitación a formar *parte de tu mundo*. Doy gracias a Dios y a la vida el haber coincidido contigo, por haber compartido un proyecto en común hasta donde ambos pudimos y espero que encuentres lo que andas buscando, que encuentres aquello que tu vida necesita y que llegue el amor y el éxito que, sin duda, mereces. Eternamente agradecido por todos los días y momentos que nos quisimos regalar.

## **AGRADECIMIENTOS**

A la **Universidad de Sonora** y al **Posgrado en Biociencias del Departamento de Investigaciones Científicas y Tecnológicas** por todo el apoyo, los espacios, facilidades y oportunidades brindadas durante el desarrollo y término de esta tesis de doctorado.

Al **Consejo Nacional de Ciencia y Tecnología (CONACyT)** por el apoyo en el otorgamiento de la beca de doctorado No. 292922.

A la **Dra. Ángela Corina Hayano Kanashiro** por todo su apoyo, ‘coaching’ y contribución a mi formación en toda la extensión de la palabra. Gracias por el tiempo compartido, su amistad y disponibilidad al 200 % para que este proyecto se llevara a cabo. Gracias por reforzar mi convicción de hacer las cosas bien hechas, de entrega al trabajo, el auxiliar a otros y enseñarme a anticiparse y prever situaciones para en caso de que surjan imprevistos, poder sortearlos lo mejor posible. Puede estar segura de que la higuera ha dado frutos y espero en un futuro seguir colaborando con usted.

Al **Dr. José Antonio López Elías**, por el apoyo, el abrirme las puertas en su laboratorio, el consejo, la disponibilidad y sus palabras de aliento en los momentos que más lo necesité. Gracias por introducirme en el mundo de las microalgas y contagiarme el gusto por su estudio y darle sentido a la formación profesional que he llevado hasta ahora. Espero en el futuro seguir colaborando y contando con usted. Gracias por todo.

A los **Dres. Luis Ángel Medina Juárez, Luis Rafael Martínez Córdova, Miguel Ángel Hernández Oñate** y **Luis Fernando García Ortega** por su acompañamiento, asesoría, apoyo, disposición, tiempo y atinadas observaciones para mi crecimiento y desarrollo de este proyecto de investigación. Completamente abierto y honrado en seguir colaborando con ustedes en un futuro.

A los **Dres. Alejandro Varela Romero y José Manuel Grijalva Chon** por su apoyo y las facilidades brindadas al permitirme hacer uso del espacio y equipos en su laboratorio de Ecología Molecular del DICTUS.

A los **Dres. Enrique de la Re-Vega, Marco Antonio López Torres, Nohemí Gámez Meza, Armando Tejeda Mansir, Kadiya del Carmen Calderón Alvarado, Maribel Ovando Martínez, Roberto Vázquez Euan, Eduardo Aguirre Hinojosa, María del Carmen Garza Aguirre, Luis Enrique Gutiérrez Millán, Fernando Enríquez Ocaña, María Guadalupe Burboa Zazueta, Mayra Lizett González Félix, Martín Pérez Velázquez, Perla Urquidez Bejarano y José Raul Romo León.** A la **M.C. Claudia Celeste Molina Domínguez, M.C. Erick José Ponce Manjarrez, M.C. Dolores Alicia Vásquez del Castillo, M.C. Edgar Sandoval Petris, L.A. María de los Ángeles Ibarra Madrigal** y técnicos **Álvaro Murguía López y José Luis Nieblas** por su apoyo, seguimiento y palabras de ánimo durante mi formación en el Posgrado en Biociencias de la Universidad de Sonora.

A **Lic. Ana Cecilia Piri Alcaraz** por sus finas atenciones, orientación y apoyo. Muchas gracias por siempre recibirme con una sonrisa, facilitar la fluidez de todos los trámites y sobre todo por su amistad y consejo cuando lo he necesitado.

Al **Dr. Juan Carlos Ramírez Suárez** y a las **M.C. María Gisela Carvallo Ruiz y M.C. María Elena Sánchez Lugo** por su apoyo, asistencia y asesoramiento durante la estancia realizada en el **Centro de Investigaciones en Alimentación y Desarrollo, A.C. (CIAD)** en el que se cumplió uno de los objetivos planteados para el desarrollo de este proyecto de investigación.

A mis compañeros de laboratorio, **Fernando Guadalupe Razo Mendivil, Karla Fabiola Yescas Romo, Marielena Clark Rivera, Génesis Valenzuela Quintero, Carlos Cota, Martín Fernando Encinas Yáñez, Víctor Andréé de la Torre Velázquez, Damaristelma de Jesús Campos, Dylann Córdova, Yeimi Vega Ruiz, Diana Fimbres Olivarría, Ana Lucía Gómez, Jesús Encinas Arzate, Ricardo Iván González, Norma García Lagunas, Carlos de la Rosa Ureña, María Luisa Juárez Gómez, Carlos Maldonado Othón y Emmanuel Villanueva** por todo el apoyo, las palabras de ánimo, la convivencia y momentos que atesoraré siempre.

A **Raymundo Favela López, Martha Cruz Flores, Lizeth Adriana Rodríguez Valdez y Martín Valenzuela Porchas:** Han sido varios los proveedores con los que he tenido contacto, sin embargo, ustedes han sido los más frecuentes y les agradezco siempre su tiempo, seguimiento, orientación y confianza.

A otras personas que de manera indirecta o directa han hecho esto posible. Gracias por su amistad, apoyo y accesibilidad: **Ismael, Irene, Don Gaby, Eli, Cynthia, Ana...** en fin, la lista es muy larga y no acabaría nunca de agradecer y si no aparecen aquí, lo seguro es que están en mi corazón.

## RESUMEN

El fósforo (P) es un macronutriente esencial para varios procesos biológicos y su limitación en el cultivo de microalgas es una estrategia eficiente para inducir la acumulación de lípidos. El objetivo de este estudio fue analizar el transcriptoma de *Chaetoceros muelleri* en respuesta a diferentes concentraciones limitantes de P y su relación con la acumulación de aceites. Se cultivó esta diatomea en medio F bajo diferentes concentraciones de P: control (72  $\mu\text{M}$ ), exceso (144  $\mu\text{M}$ ) y dos concentraciones limitantes (7  $\mu\text{M}$  y 18  $\mu\text{M}$ ). Se evaluó el efecto de la limitación y exceso de P en el crecimiento, producción de biomasa y composición bioquímica; además, se cuantificó el perfil de ácidos grasos siendo los más abundantes: mirístico, palmítico, palmitoleico y eicosapentaenoico. En el análisis de RNA-seq se identificaron 959, 405 y 34 transcritos expresados diferencialmente en *C. muelleri* bajo las concentraciones de P de 7, 18 y 144  $\mu\text{M}$  respectivamente. El análisis de enriquecimiento GO mostró que los transcritos que se inducen ante la limitación de P están relacionados con el metabolismo de fósforo, síntesis de sulfolípidos, biosíntesis de carbohidratos y metabolismo de lípidos, mientras que los transcritos que se encontraron reprimidos se encontraron relacionados con la traducción de proteínas y el metabolismo del nitrógeno principalmente. Bajo condiciones de exceso de P no se identificaron procesos biológicos enriquecidos. Estos resultados sugieren que ante la limitación de P esta diatomea implementa estrategias para satisfacer sus requerimientos vitales y asegurar su supervivencia. Finalmente, cuatro genes que codifican a proteínas transportadoras de P y enzimas involucradas en el metabolismo de ácidos grasos fueron analizadas *in silico*, encontrándose similitudes estructurales y funcionales con otras especies de diatomeas; estos resultados permitirán a futuro plantear estudios que amplíen la comprensión de la regulación de la biosíntesis de lípidos en *C. muelleri* bajo la limitación de nutrientes.

## ABSTRACT

Phosphorus (P) is an essential macronutrient for several biological processes and its limitation in microalgae culturing it is an efficient strategy to induce lipid accumulation. The aim of this study was to analyze the transcriptome of *Chaetoceros muelleri* in response to different P conditions and its relationship to lipid accumulation. This diatom was cultured in F medium under different P concentrations: control (72  $\mu\text{M}$ ), excess (144  $\mu\text{M}$ ) and two limiting concentrations (7  $\mu\text{M}$  and 18  $\mu\text{M}$ ). The effect of P limitation and excess on growth, biomass production and biochemical composition was evaluated. In addition, the fatty acid profile was quantified being the most abundant: myristic, palmitic, palmitoleic and eicosapentaenoic. RNA-seq analysis identified 959, 405 and 34 differentially expressed transcripts in *C. muelleri* under P concentrations of 7, 18 and 144  $\mu\text{M}$ , respectively. GO enrichment analysis showed that transcripts induced under P limitation were related to phosphorus metabolism, sulfolipid synthesis, carbohydrate biosynthesis and lipid metabolism, whereas transcripts that were found repressed were mainly related to translation and nitrogen metabolism. Under conditions of P excess, no enriched biological processes were identified. These results suggest that under P limitation this diatom implements strategies to satisfy its vital requirements and ensure its survival. Finally, four genes encoding P transporter proteins and enzymes involved in fatty acid metabolism were analyzed *in silico* finding structural and functional similarities with other diatom species; these results will allow future studies to improve the understanding of the regulation of lipid biosynthesis in *C. muelleri* under nutrient limitation.

# ÍNDICE GENERAL

	Página
<b>APROBACIÓN</b> .....	<i>i</i>
<b>DEDICATORIA</b> .....	<i>ii</i>
<b>AGRADECIMIENTOS</b> .....	<i>iv</i>
<b>RESUMEN</b> .....	<i>vii</i>
<b>ABSTRACT</b> .....	<i>viii</i>
<b>ÍNDICE DE FIGURAS</b> .....	<i>xii</i>
<b>ÍNDICE DE TABLAS</b> .....	<i>xvi</i>
<b>INTRODUCCIÓN</b> .....	1
<b>I. ANTECEDENTES</b> .....	4
1.1 Generalidades de las microalgas .....	4
1.1.1 Características adaptativas de las microalgas .....	5
1.1.2 Importancia económica y ecológica de las microalgas .....	6
1.2 Clasificación de las microalgas .....	7
1.2.1 Diatomeas .....	8
1.2.1.1 Género <i>Chaetoceros</i> .....	9
1.2.1.1.1 <i>Chaetoceros muelleri</i> .....	10
1.3 Aplicaciones de las microalgas .....	12
1.3.1 Biocombustibles .....	13
1.4 Producción de aceites a partir de microalgas .....	15
1.5 Producción de biodiesel a partir de aceites de microalgas .....	16
1.5.1 Perspectiva a futuro de la producción de biomasa a través del cultivo de microalgas .....	17
1.6 Cultivo de microalgas .....	19
1.6.1 Factores que influyen en el cultivo de microalgas .....	19
1.6.1.1 Efecto de la limitación de fósforo en microalgas .....	20
1.7 Transcriptómica .....	23
1.7.1 RNA-seq .....	23
1.7.1.1 Análisis de datos generados por RNA-seq .....	24
1.8 Análisis <i>in silico</i> de proteínas .....	27
<b>II. JUSTIFICACIÓN</b> .....	30
<b>III. HIPÓTESIS</b> .....	31
<b>IV. OBJETIVOS</b> .....	32
4.1 Objetivo general .....	32
4.2 Objetivos específicos .....	32

<b>V. MATERIALES Y MÉTODOS</b> .....	33
5.1 Material biológico .....	33
5.2 Condiciones de crecimiento .....	33
5.3 Determinación de la concentración celular .....	34
5.4 Determinación de biomasa y análisis bioquímicos .....	35
5.5 Perfil de ácidos grasos de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	36
5.5.1 Predicción de propiedades del biodiesel a partir de los perfiles de ácidos grasos de <i>C. muelleri</i> cultivada .....	37
5.6 Análisis transcriptómico de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	38
5.6.1 Extracción de ARN total .....	38
5.6.2 Cuantificación e integridad del ARN total.....	39
5.6.2.1 Purificación de ARN extraído de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	40
5.6.3 Construcción de librerías de ADNc y secuenciación .....	40
5.6.4 Filtrado de calidad secuencias, ensamblado <i>de novo</i> del transcriptoma y anotación funcional.....	41
5.6.5 Análisis de expresión diferencial y análisis de enriquecimiento.....	42
5.7 Análisis <i>in silico</i> de genes involucrados en el transporte de P y metabolismo de lípidos en <i>C. muelleri</i> .....	43
<b>VI. RESULTADOS Y DISCUSIÓN</b> .....	45
6.1 Crecimiento y determinación de la concentración celular de <i>C. muelleri</i> bajo diferentes concentraciones de P .....	45
6.2 Biomasa y análisis bioquímicos de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	47
6.3 Perfil de ácidos grasos de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	50
6.3.1 Predicción de propiedades del biodiesel a partir de perfiles de ácidos grasos de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P ..	55
6.4 Análisis transcriptómico de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	60
6.4.1 Integridad del ARN total .....	60
6.4.2 Análisis de calidad y filtrado de secuencias .....	61
6.4.3 Ensamblado del transcriptoma y alineamiento .....	64
6.4.4 Anotación funcional del transcriptoma de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	66
6.4.5 Análisis de los genes expresados diferencialmente en <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	69

6.4.6	Análisis de enriquecimiento GO de los genes expresados diferencialmente en <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P .....	75
6.4.7	Respuesta transcripcional de <i>C. muelleri</i> cultivada bajo concentraciones limitantes de P .....	77
6.4.8	Respuesta transcripcional de <i>C. muelleri</i> cultivada bajo exceso de P ..	87
6.4.9	Respuesta transcripcional de <i>C. muelleri</i> cultivada bajo concentraciones limitantes de P y su relación con la acumulación .....	89
6.5	Análisis <i>in silico</i> de proteínas transportadoras de P de <i>C. muelleri</i> .....	94
6.5.1	<i>NPT2B</i> : Proteína de transporte de fosfato dependiente de sodio 2B ....	96
6.5.2	<i>VTC4</i> : Chaperona de transporte vacuolar 4 .....	102
6.6	Análisis <i>in silico</i> de proteínas involucradas en el metabolismo de lípidos y ácidos grasos en <i>C. muelleri</i> .....	107
6.6.1	<i>ACC</i> : Acetil CoA carboxilasa .....	109
6.6.2	<i>FAD2</i> : Delta 12-desaturasa .....	115
<b>VII.</b>	<b>CONCLUSIONES</b> .....	122
<b>VIII.</b>	<b>RECOMENDACIONES</b> .....	124
<b>IX.</b>	<b>LITERATURA CITADA</b> .....	125
<b>X.</b>	<b>APÉNDICES</b> .....	151

## ÍNDICE DE FIGURAS

<b>Figura</b>		<b>Página</b>
1	Célula de <i>Chaetoceros muelleri</i> , vista cingular. Barra de escala = 10 $\mu\text{m}$ .	12
2	Principales países productores de biodiesel. Tomada de Ganesan <i>et al.</i> (2020)	17
3	Diseño experimental para evaluar el efecto de las diferentes concentraciones de P en la diatomea <i>C. muelleri</i> .	34
4	Proceso de estandarización llevado a cabo para la obtención de ARN total de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P	38
5	Crecimiento de <i>C. muelleri</i> durante 9 días en cuatro tratamientos experimentales con diferentes concentraciones de P: 7 $\mu\text{M}$ , 18 $\mu\text{M}$ , 72 $\mu\text{M}$ y 144 $\mu\text{M}$ . Las barras de error representan la desviación estándar de los conteos celulares de cuatro réplicas biológicas compuestas por tres réplicas técnicas cada una por tratamiento.	46
6	Perfil de ácidos grasos de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P. Los valores son la media $\pm$ desviación estándar de cuatro réplicas biológicas y cada una está compuesta por tres réplicas técnicas. Los datos se presentan como porcentaje de ácidos grasos. Los valores en cada barra con diferente superíndice (a-c) indican una diferencia significativa ( $p < 0.05$ ).	52
7	ARN total de <i>C. muelleri</i> extraído con reactivo TRIzol (Invitrogen) y purificado con el kit comercial RNeasy MinElute Clean Up Kit (QIAGEN) en gel de agarosa al 1 % teñido con GelRed (Biotium) en condiciones no desnaturalizantes. Carril del 1 al 4 corresponden a la primera réplica biológica (RB1) y del carril 5 al 8 corresponde a la segunda réplica biológica (RB2). Carril 1: 7 $\mu\text{M}$ _RB1. Carril 2: 18 $\mu\text{M}$ _RB1. Carril 3: 72 $\mu\text{M}$ _RB1. Carril 4: 144 $\mu\text{M}$ _RB1. Carril 5: 7 $\mu\text{M}$ _RB2. Carril 6: 18 $\mu\text{M}$ _RB2. Carril 7: 72 $\mu\text{M}$ _RB2. Carril 8: 144 $\mu\text{M}$ _RB2.	60
8	Conteo de datos crudos (lecturas) por librería y por tratamiento previo al filtrado de calidad.	62
9	Porcentaje de GC (citosina y guanina) en los datos crudos de secuenciación.	62
10	Conteo de lecturas por librería y por tratamiento posterior al filtrado de calidad con los softwares Trimmomatic y SortMeRNA.	63
11	Porcentaje de GC (citosina y guanina) de las lecturas posterior al filtrado de calidad.	63
12	Histogramas de clasificaciones de transcritos anotados de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P con la base de datos Gene Ontology (GO). A: Procesos biológicos. B: Componentes	68

	celulares. C: Funciones moleculares. D: Rutas metabólicas anotadas con KEGG.	
13	Histograma de clasificaciones de transcritos anotados de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P con la base de datos Clusters Orthologs Groups (COG).	69
14	Diagrama de Venn de los genes expresados diferencialmente de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P.	70
15	Diagramas de Venn de genes expresados diferencialmente en <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P. A: Número de DEG inducidos en <i>C. muelleri</i> . B: Número de DEG reprimidos en <i>C. muelleri</i>	71
16	Heatmap de análisis de correlación de Pearson tomando como referencia la matriz de conteos de isoformas de <i>C. muelleri</i> . Tratamiento 7 $\mu$ M (Limitante 1.1 y Limitante 1.2). Tratamiento 18 $\mu$ M (Limitante 2.1 y Limitante 2.2). Tratamiento 72 $\mu$ M (Control.1 y Control.2) y Tratamiento 144 $\mu$ M (Exceso.1 y Exceso.2).	71
17	Análisis de escalado multidimensional (MDS, por sus siglas en inglés) tomando como referencia la matriz de conteos de isoformas de <i>C. muelleri</i> . Tratamiento 7 $\mu$ M (Limitante 1.1 y Limitante 1.2) en color negro. Tratamiento 18 $\mu$ M (Limitante 2.1 y Limitante 2.2) en color rojo. Tratamiento 72 $\mu$ M (Control.1 y Control.2) en color verde y Tratamiento 144 $\mu$ M (Exceso.1 y Exceso.2) en color azul.	72
18	Heatmap de agrupamiento jerárquico de los genes expresados diferencialmente de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P. La escala de colores representa el Log2FC de los genes expresados diferencialmente (FDR < 0.01 y log2FC $\geq \pm 1$ ), oscilando de azul, para la mayoría de los genes reprimidos, a amarillo, para la mayoría de genes inducidos. El dendrograma indica la relación entre los perfiles genéticos utilizando el método de agrupamiento ward.D.	74
19	Heatmap del análisis de enriquecimiento de los principales grupos de Gene Ontology (GO) observados <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P. Cada columna contiene los términos GO reprimidos (azul) e inducidos (naranja). Sólo se muestran los términos GO más específicos (FDR < 0.05). La escala de colores representa el porcentaje de genes que pertenecen a cada término GO. Los asteriscos representan la importancia estadística del enriquecimiento (*FDR < 0.05, **FDR < 0.01 y ***FDR < 0.001 respectivamente).	76
20	Modelo hipotético de la respuesta transcripcional de <i>C. muelleri</i> ante la limitación por P. Cuadros en color rojo: Transcritos inducidos. Cuadros en color azul marino: Transcritos reprimidos. Creado en Biorender: <a href="http://www.biorender.com">www.biorender.com</a>	86

21	Modelo hipotético de la respuesta transcripcional de <i>C. muelleri</i> ante el exceso de P. Cuadros en color rojo: Transcritos inducidos. Cuadros en color azul marino: Transcritos reprimidos. Creado en Biorender: <a href="http://www.biorender.com">www.biorender.com</a>	89
22	Modelo hipotético de la acumulación de lípidos en <i>C. muelleri</i> ante la limitación por P. Valores en color rojo: log <sub>2</sub> FC en tratamiento con concentración de P de 7 μM. Valores en color azul: log <sub>2</sub> FC en tratamiento con concentración de P de 18 μM. +: Transcritos inducidos. - : Transcritos reprimidos. Creado en Biorender: <a href="http://www.biorender.com">www.biorender.com</a>	93
23	Composición aminoacídica de las secuencias de las proteínas transportadoras de P de <i>C. muelleri</i> . A: NPT2B. B: VTC4.	95
24	NPT2B: Proteína de transporte de fosfato dependiente de sodio 2B. Creada en Biorender: <a href="https://biorender.com/">https://biorender.com/</a>	97
25	Dominio conservado de la familia de los cotransportadores de sodio y P en la secuencia aminoacídica de NPT2B de <i>C. muelleri</i> .	98
26	Alineamiento de las secuencias de aminoácidos de NPT2B, mostrando 5 motivos significativamente conservados. La secuencia de NPT2B de <i>C. muelleri</i> está resaltada en rojo.	99
27	Representación gráfica de la predicción de la estructura secundaria de la proteína NPT2B de <i>C. muelleri</i> realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.	100
28	Modelo de la estructura terciaria teórica de la proteína cotransportadora de sodio y fósforo 2B (NPT2B) de <i>C. muelleri</i> obtenida a través del servidor SWISS MODEL.	101
29	VTC4: Chaperona de transporte vacuolar 4. Creada en Biorender: <a href="https://biorender.com/">https://biorender.com/</a>	103
30	Dominios conservados en la secuencia aminoacídica de VTC4 de <i>C. muelleri</i> .	104
31	Alineamiento de las secuencias de aminoácidos de VTC4, mostrando 4 motivos significativamente conservados. La secuencia de VTC4 de <i>C. muelleri</i> está resaltada en rojo.	105
32	Representación gráfica de la predicción de la estructura secundaria de la proteína VTC4 de <i>C. muelleri</i> realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.	106
33	Modelo de la estructura terciaria teórica de la proteína chaperona de transporte vacuolar 4 (VTC4) de <i>C. muelleri</i> obtenida a través del servidor SWISS MODEL.	107

34	Composición aminoacídica de las secuencias de las enzimas involucradas en la síntesis de ácidos grasos de <i>C. muelleri</i> . A: ACC. B: FAD2.	108
35	ACC: Acetil CoA carboxilasa. Creada en Biorender: <a href="https://biorender.com/">https://biorender.com/</a>	110
36	Dominios conservados en la secuencia aminoacídica de ACC de <i>C. muelleri</i> .	111
37	Alineamiento de las secuencias de aminoácidos de ACC, mostrando 3 motivos identificados. La secuencia de ACC de <i>C. muelleri</i> está resaltada en rojo.	113
38	Representación gráfica de la predicción de la estructura secundaria de la proteína ACC de <i>C. muelleri</i> realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.	114
39	Modelo de la estructura terciaria teórica de la enzima acetil-CoA carboxilasa (ACC) de <i>C. muelleri</i> obtenida a través del servidor SWISS MODEL.	114
40	FAD2: Delta 12-desaturasa. Creada en Biorender: <a href="https://biorender.com/">https://biorender.com/</a>	116
41	Dominio conservado de desaturasas de ácidos grasos en la secuencia aminoacídica de FAD2 de <i>C. muelleri</i> .	117
42	Alineamiento de las secuencias de aminoácidos de FAD2, mostrando 4 motivos significativamente conservados. La secuencia de FAD2 de <i>C. muelleri</i> está resaltada en rojo.	119
43	Representación gráfica de la predicción de la estructura secundaria de la enzima FAD2 de <i>C. muelleri</i> realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.	120
44	Modelo de la estructura terciaria teórica de la delta 12-desaturasa (FAD2) de <i>C. muelleri</i> obtenida a través del servidor SWISS MODEL.	121

## ÍNDICE DE TABLAS

Tabla		Página
1	Densidades celulares y tasas de crecimiento de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P.	47
2	Producción de biomasa y composición bioquímica de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P.	48
3	Perfil de ácidos grasos de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P.	51
4	Predicción de propiedades del biodiesel calculadas a partir del perfil de ácidos grasos de <i>C. muelleri</i> cultivada bajo diferentes concentraciones de P.	55
5	Secuencias de ARN ribosomal detectadas en las librerías de <i>C. muelleri</i> por el software SortMeRNA.	64
6	Métricas proporcionadas por rnaQUAST, BUSCO y Gene Mark S-T a partir de los transcriptomas ensamblados con Trinity, rnaSPAdes y SOAPdenovo.	65
7	Porcentaje de alineamiento de lecturas con los transcriptomas ensamblados con la herramienta Trinity.	66
8	Lista de transcritos inducidos en <i>C. muelleri</i> cultivada bajo limitación por P de acuerdo al análisis de RNA-seq.	83
9	Lista de transcritos reprimidos en <i>C. muelleri</i> cultivada bajo limitación por P de acuerdo al análisis de RNA-seq.	84
10	Propiedades fisicoquímicas de las secuencias aminoacídicas de las proteínas transportadoras de fósforo NPT2B y VTC4 de <i>C. muelleri</i> .	94
11	Comparación del porcentaje de identidad de la secuencia aminoacídica NPT2B de <i>C. muelleri</i> con las secuencias correspondientes a otras diatomeas.	97
12	Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de NPT2B de <i>C. muelleri</i> obtenidos por el servidor NPS@.	101
13	Comparación del porcentaje de identidad de la secuencia aminoacídica VTC4 de <i>C. muelleri</i> con las secuencias correspondientes a otras diatomeas.	103
14	Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de VTC4 de <i>C. muelleri</i> obtenidos por el servidor NPS@.	106
15	Propiedades fisicoquímicas de las secuencias aminoacídicas de las proteínas involucradas en el metabolismo de lípidos y ácidos grasos ACC y FAD2 de <i>C. muelleri</i> .	108

16	Comparación del porcentaje de identidad de la secuencia aminoacídica ACC de <i>C. muelleri</i> con las secuencias correspondientes a otras diatomeas.	111
17	Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de ACC de <i>C. muelleri</i> obtenidos por el servidor NPS@.	114
18	Comparación del porcentaje de identidad de la secuencia aminoacídica FAD2 de <i>C. muelleri</i> con las secuencias correspondientes a otras diatomeas.	117
19	Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de FAD2 de <i>C. muelleri</i> obtenidos por el servidor NPS@.	120

## INTRODUCCIÓN

Las diatomeas son el grupo más abundante de fitoplancton y el quinto grupo más diverso de células eucarióticas (Malviya *et al.*, 2016; de Vargas *et al.*, 2015). Se consideran un grupo filogenéticamente reciente, ya que se estima que aparecieron hace aproximadamente 190 millones de años a partir de una endosimbiosis seriada primero con algas verdes y posteriormente con alga rojas (Nakov *et al.*, 2018; Petersen *et al.*, 2014). Se estiman entre 8, 000 y 10, 000 especies descritas de diatomeas distribuidas en 1200 géneros y 45 órdenes, sin embargo, se cree que hay entre 20, 000 y 200, 000 especies en la naturaleza (D'ippolito *et al.*, 2015; Fields, 2014). Además, están involucradas en varios procesos biogeoquímicos en el océano, entre ellos el del carbono (C), nitrógeno (N) y silicio (Si) y, además, se les atribuye el 20 - 25 % de la productividad primaria mundial y se ha reportado que poseen una gran capacidad de adaptación a condiciones extremas (Jain, 2020; Armbrust, 2009).

Al igual que otras especies de microalgas, las diatomeas se han utilizado en varias aplicaciones biotecnológicas (como suplementos alimenticios, producción de biocombustibles, productos farmacéuticos, nutracéuticos y cosméticos), ya que se ha reportado que dentro de su composición bioquímica hay varios compuestos y metabolitos secundarios de interés y además poseen ventajas que otros recursos naturales como los cultivos terrestres no tienen como a) crecimiento rápido, b) medios de cultivo relativamente baratos para su producción, c) no necesitan tierras cultivables ni una estación específica del año para su cultivo, y d) pueden utilizarse para el tratamiento de aguas residuales (Khan *et al.*, 2018; Griffiths *et al.*, 2011). Dentro de la composición bioquímica de las diatomeas, el silicio ( $\text{SiO}_2$ ) es un componente clave en sus paredes celulares; también son ricas en compuestos de almacenamiento, entre ellos carbohidratos, como la crisolaminarina, y en cuanto a los lípidos, específicamente los ácidos grasos, los más abundantes son los ácidos mirístico, palmítico, palmitoleico y eicosapentaenoico (Jónasdóttir, 2019; Pérez-García *et al.*, 2011; Wilhelm *et al.*, 2006).

La composición bioquímica de las diatomeas, al igual que la de otras especies de microalgas, depende de las condiciones de cultivo y de la disponibilidad de nutrientes. Se ha reportado que en condiciones de estrés abiótico y de limitación de nutrientes, una de las

principales respuestas de las microalgas es la acumulación de compuestos de almacenamiento como los lípidos en forma de triacilglicéridos (TAG), lo que ha generado interés ya que pueden utilizarse en algunas aplicaciones como la producción de biodiésel (Khan *et al.*, 2018; Paliwal *et al.*, 2017). Las diatomeas son excelentes acumuladores de lípidos y tienden a almacenarlos en el interior de la célula en organelos llamados gotas lipídicas, las cuales representan un reservorio de energía durante periodos de desbalance en el flujo de electrones o ante la deficiencia de nutrientes y además poseen un papel importante en la homeostasis redox celular (Murphy, 2012); sin embargo, hasta el momento la comprensión del efecto de diferentes factores abióticos y la disponibilidad de nutrientes para diatomeas es escasa en comparación con otras microalgas como las clorofitas y las cianobacterias (Maeda *et al.*, 2017).

Una de las estrategias más eficaces para inducir la acumulación de lípidos en las microalgas es la limitación del fósforo (P) (Yang *et al.*, 2018). El P es un macronutriente vital para el cultivo de microalgas, ya que participa en el crecimiento celular y en procesos metabólicos como la transferencia de energía, la biosíntesis de macromoléculas (fosfolípidos y ácidos nucleicos) y la fotosíntesis. Dentro de las respuestas que han observado en microalgas expuestas a la limitación de P son: a) disminución del crecimiento y la proliferación celular, b) remodelación de la membrana celular mediante la sustitución de fosfolípidos por compuestos que no poseen P como los sulfolípidos o los lípidos de betaína, c) síntesis de proteínas inorgánicas de transporte de P, d) síntesis de hidrolasas como las fosfatasas alcalinas, que hidrolizan los enlaces de los ésteres a partir del fósforo orgánico disuelto (DOP) para facilitar su asimilación por las células, e) uso de polifosfatos (Poli-P) (como reservorio de P) que fueron sintetizados mientras que las concentraciones de P inorgánico (Pi) en el medio ambiente eran favorables para cubrir sus necesidades y, f) síntesis de compuestos de almacenamiento como los carbohidratos y los lípidos (Dyhrman *et al.*, 2016).

*Chaetoceros muelleri* es una diatomea central representativa del género *Chaetoceros* (uno de los géneros de diatomeas más abundantes) que es nativa del Noroeste de México. Uno de los principales usos de esta diatomea en la región es como alimento vivo para diferentes especies de peces y crustáceos en acuicultura. Se han realizado varios estudios para comprender las respuestas fisiológicas y bioquímicas a la limitación por P en *C. muelleri*; sin embargo, a nivel molecular esta información es escasa y no es clara (Wang *et al.*, 2014; López-Elías *et al.*, 2005).

Las ciencias ómicas han permitido una amplia comprensión de la biología de los organismos, su interacción con el medio ambiente y su respuesta a diferentes factores bióticos y abióticos (Ibraheem *et al.*, 2018). RNA-Seq es una poderosa herramienta empleada para el desarrollo de estudios transcriptómicos que ha permitido dilucidar los mecanismos que utilizan los distintos organismos para responder a diferentes factores, que tiene la ventaja de no requerir un genoma de referencia para realizar este tipo de estudios y obtener resultados fiables que pueden ser de relevancia para aplicaciones futuras (Lowe *et al.*, 2017; Pereira *et al.*, 2017).

Dentro de los retos a superar en la obtención de metabolitos de interés a partir de microalgas, se encuentra la producción baja de biomasa, así como la necesidad de mejorar el rendimiento de los subproductos derivados del efecto de uno o varios factores de estrés; de ahí que la integración generada por cada una de las ciencias ómicas, podría permitir una comprensión más amplia de la respuesta ante condiciones de estrés y mediante ingeniería genética se podría llevar a cabo el mejoramiento genético de cepas a fin de aprovechar todo su potencial (Lauritano *et al.*, 2019; Guarnieri y Pienkos, 2015; Wijffels, 2015). Por lo tanto, el objetivo de este estudio fue analizar el transcriptoma de *Chaetoceros muelleri* en respuesta a diferentes concentraciones limitantes de P y su relación con la acumulación de aceites.

## I. ANTECEDENTES

### 1.1 Generalidades de las microalgas

La palabra ‘plancton’, proveniente del griego y que fue acuñada por Victor Hensen en 1887, hace alusión a aquello que es ‘errante’ o que ‘se encuentra a la deriva’ y se utiliza para definir al conjunto de organismos que se encuentran suspendidos en la columna de agua (planctónicas) o habitando los sedimentos marinos o sobre diferentes sustratos (bentónicas), con o sin la capacidad de desplazarse por sí mismos. Estos organismos se clasifican en fitoplancton (organismos unicelulares vegetales) y zooplancton (organismos unicelulares animales) (Pachiappan *et al.*, 2019). El fitoplancton, o también conocidas como microalgas, son un conjunto de microorganismos unicelulares que poseen la capacidad de fijar el dióxido de carbono (CO<sub>2</sub>) por medio de la fotosíntesis y sintetizar biomoléculas. De acuerdo con un estimado, las microalgas pueden utilizar el 9 % de las radiaciones solares y así producir 280 toneladas de biomasa seca por hectárea y por año (Chen *et al.*, 2018; Dhillon y von Wuehlisch, 2013) al consumir alrededor de 513 toneladas de CO<sub>2</sub> y además producir cerca del 50 % de oxígeno (O<sub>2</sub>) del planeta (Lü *et al.*, 2011).

Las microalgas poseen una pared celular, membrana plasmática, citoplasma, núcleo, organelos y plástidos. El tamaño de las células microalgales puede variar de unos pocos µm a aproximadamente 200 µm. La biodiversidad de especies microalgales es considerada abundante ya que se estima la existencia de alrededor de 200, 000 y 800, 000 especies, aunque sólo se han descrito en la literatura entre 40,000 y 50,000 especies (Suganya *et al.*, 2016). Se caracterizan por crecer en diversos hábitats acuáticos, presentando una capacidad alta de adaptarse rápidamente a condiciones extremas, produciendo a su vez una gran variedad de metabolitos secundarios (biológicamente activos) conformados por estructuras particulares y actividades biológicas que generalmente no se encuentran en otros organismos (Anbuezhian *et al.*, 2015; Herrero *et al.*, 2013).

### 1.1.1 Características adaptativas de las microalgas

La capacidad de las microalgas para adaptarse a una amplia diversidad de ecosistemas se debe principalmente al proceso de evolución, el cual, ha sido moldeado por procesos tales como endosimbiosis, transferencia genética vertical y horizontal. Existe evidencia de que estos procesos han contribuido considerablemente al origen de genomas mosaico que se caracterizan por una mezcla y homología alta de genes que provienen de diferentes organismos fotosintéticos, eucariotas heterótrofos y bacterias (Hopes y Mock, 2015; Armbrust, 2009).

Las microalgas han desarrollado diversos mecanismos de adaptación en varios ecosistemas acuáticos para su crecimiento y supervivencia. En regiones costeras y estuarios, las especies de microalgas que predominan son diatomeas y prasinofitas (Rodrigues *et al.*, 2014); estas regiones se caracterizan por niveles altos y fluctuantes de nutrientes y condiciones de luz que pueden ser tanto bajas como variables. Las especies que habitan estas regiones han desarrollado varios mecanismos de adaptación entre ellos "estrategia de velocidad" para la adquisición de nutrientes cuando se encuentran en concentraciones altas en el ambiente (Kwon *et al.*, 2013, Litchman *et al.*, 2007). La elevada asimilación de nutrientes y las tasas menores de crecimiento permiten que algunas diatomeas almacenen nutrientes en sus vacuolas durante las temporadas de altas concentraciones (Litchman *et al.*, 2007), asegurando su supervivencia cuando los niveles son fluctuantes.

Otro mecanismo de adaptación aprovechado por diatomeas que habitan regiones costeras y estuarios para la adquisición de nutrientes es su morfología, ya que al girar sobre sí mismas, generan microturbulencia con sus setas permitiendo una mayor asimilación de nutrientes; por otro lado, la forma de la frústula en diatomeas, principalmente pennadas, puede auxiliar la asimilación y distribución de nutrientes, así como la captación de luz cuando esta se encuentra limitada (De Stefano *et al.*, 2007, Pahlow *et al.*, 1997).

En el mar abierto las microalgas están representadas por grupos tales como coccolitóforos, prasinofitos, dinoflagelados y pelagófitos; estas regiones se caracterizan por aguas estratificadas, bajo contenido de nutrientes y alta irradiación (Litchman *et al.*, 2007; Rodrigues *et al.*, 2014). Dentro de los nutrientes limitantes para dichas especies se encuentra el hierro, y entre los mecanismos que estas especies han desarrollado para hacer frente a dicha limitación

se ha observado la producción de varias enzimas, que utilizan zinc, cobre, manganeso y níquel en lugar de hierro, ayudando así a reducir las demandas de este nutriente, mientras que las proteínas transportadoras de hierro y aglutinantes optimizan su adquisición (Read *et al.*, 2013; Worden *et al.*, 2009).

### **1.1.2 Importancia económica y ecológica de las microalgas**

Los organismos marinos comprenden aproximadamente la mitad de la biodiversidad presente en la tierra y el ecosistema marino es un reservorio enorme de recursos potenciales para uso humano (Boeuf, 2011). Dentro de los organismos presentes en el océano, las microalgas son los microorganismos fotosintéticos responsables de alrededor de la mitad de la producción primaria global de materia orgánica mediante un proceso de conversión del CO<sub>2</sub> en carbono orgánico, del cual, un porcentaje es empleado en el ciclo del carbono (Hülse *et al.*, 2017). Además, las microalgas contribuyen ampliamente en la producción global de oxígeno (hasta un 50 %). Su participación en estos procesos, así como en el reciclaje de nutrientes las hace un cimiento importante para la mayoría de la vida en el planeta (Basu y Mackey, 2018). También poseen la capacidad de mitigar la contaminación del agua como del suelo provocada por desechos industriales tales como sustancias orgánicas e inorgánicas y metales pesados; por otro lado, la remoción del exceso de nitrógeno y carbono del agua es de utilidad para prevenir la eutroficación en los ambientes acuáticos (Abdel-Raouf *et al.*, 2012). Con lo que respecta al suelo, el rol que desempeñan las microalgas es el mejorar las condiciones físicas y químicas del mismo, sintetizando y secretando metabolitos promotores del crecimiento vegetal y con diversas actividades biológicas proveyendo así un balance ecológico (Barone *et al.*, 2018a).

Asimismo, las microalgas son consideradas excelentes indicadores de cambios en el ecosistema ya que son altamente sensibles a los cambios en la calidad de luz, temperatura, disponibilidad de nutrientes y contaminación; un ejemplo de lo anterior, son las diatomeas, las cuales son utilizadas ampliamente para monitorear el estado ecológico de ecosistemas lóticos ya que proporciona varias ventajas tales como que su uso para este objetivo es económico, los resultados generados son comparables, su implementación es rápida y precisa (Bere y Tundisi, 2009). Las diatomeas al poseer un ciclo de vida corto permiten detectar rápidamente los cambios

en la calidad de agua relacionados con actividades antropogénicas y en su forma fósil, las diatomeas también son buenas indicadores de las variaciones climáticas (Bigler y Hall, 2002).

Por otro lado, el crecimiento rápido de la población mundial ha resultado en un incremento continuo en la demanda de energía y alimentos sustentables (Banerjee y Ramaswamy, 2019). Asimismo, como resultado de la creciente demanda de recursos renovables adecuados para combatir el cambio climático global, surge la necesidad de extender la búsqueda de recursos sustentables para la producción de productos de alto valor agregado a escala comercial. Las microalgas son consideradas un recurso adecuado para la obtención de energía, alimentos y compuestos bioactivos (Laurens, 2017); poseen algunas ventajas en comparación de los cultivos terrestres, tales como productividad alta, no requieren de tierras de cultivo y pueden crecer en agua de baja calidad. De ahí que la evaluación precisa del costo de la obtención de biomasa microalgal posee una importancia económica crucial para su aprovechamiento en aplicaciones diversas (Banerjee y Ramaswamy, 2019).

El uso de microalgas como un recurso eficiente y asequible para la obtención de energía es una realidad que no se espera alcanzar a corto plazo ya que el costo producción y biorefinería actual es alto, por lo que no es competitivo ante el costo de producción de combustibles fósiles y el uso de recursos renovables tradicionales (García *et al.*, 2019); sin embargo, la biotecnología microalgal se está orientando hacia la obtención de productos de alto valor agregado cuyos precios de venta pueden garantizar la recuperación de la inversión (Barsanti y Gualtieri, 2018). La mayoría de los metabolitos encontrados en microalgas poseen importantes actividades antioxidantes, antiinflamatorias, antifúngicas, antibióticas, antivirales, inmuno-estimulantes y antitumorales. Adicionalmente, poseen designación GRAS que por sus siglas en inglés significa “generalmente reconocido como seguro”, lo que permite su uso para consumo humano y animal (FDA, 2016).

## **1.2 Clasificación de las microalgas**

Las microalgas pueden clasificarse en dos grandes grupos: procariotas y eucariotas (Harun *et al.*, 2010). Dentro del grupo de microalgas procariotas se encuentran las cianobacterias (algas verdeazuladas) las cuales se caracterizan principalmente por su crecimiento rápido,

almacenamiento de nutrientes y compuestos tales como fosfatos, cianoficina (un polímero compuesto por ácido aspártico y arginina) y un poliglicano ramificado; algunas de las especies pertenecientes a este grupo poseen la capacidad de sintetizar neurotoxinas y hepatotoxinas mientras que otros producen compuestos terapéuticos (Villarruel-López *et al.*, 2017).

Por su parte, en el grupo de microalgas eucariotas se pueden distinguir varios grupos: clorofitas (considerado uno de los grupos más numerosos), euglenofitas (organismos unicelulares y flagelados), rodofitas (se caracterizan por tener dentro de su composición bioquímica ficoeritrina que les da su color rojizo), criptofitas (las especies pertenecientes a este grupo se caracterizan por ser unicelulares y que poseen capacidad de moverse por sí mismas), dinoflagelados (son considerados un grupo de microalgas amplio, sin embargo, algunas especies no son utilizadas comercialmente por su producción de toxinas), haptofitas (este grupo de microalgas se caracterizan por la presencia de carbonato de calcio que cubre sus células y son responsables del 50 % de la precipitación de este compuesto en el océano), glaucofitas (especies de agua dulce cuyo organelo fotosintético principal recibe el nombre de cianela), las cloraracniofitas (un grupo pequeño de microalgas eucariotas flageladas) y por último las heterocontofitas (dentro de este grupo se encuentran varias especies de microalgas pertenecientes a las clases Phaeophyceae, Chrysophyceae, Xanthophyceae y Bacillariophyceae; en esta última se encuentra las diatomeas) (Duong *et al.*, 2012; Barsanti *et al.*, 2008).

### **1.2.1 Diatomeas**

Las diatomeas, pertenecientes a la clase Bacillariophyceae, son el grupo de microalgas más ubicuo y rico en especies representando alrededor del 40 % del total de organismos fitoplanctónicos, teniendo registrados hasta el momento alrededor de 500 géneros, de los cuales 350 (12,000 especies) aún existen y 150 (5,000 especies) se encuentran extintos (Jain, 2020; Williams y Kociolek, 2011; Hoek *et al.*, 2009); se estima que el número de especies de este grupo oscila aproximadamente de entre 200, 000 y 10 millones de especies; esta variación es resultado del conocimiento limitado de su diversidad, en particular de las especies que habitan en los trópicos (Kale y Karthick, 2015). Se estima que estas especies son responsables de entre un 20 y 25 % de la producción de oxígeno y fijación de carbono en el mundo, una cifra que es

mayor a la productividad de todos los bosques tropicales existentes (Jain, 2020; Kale y Karthick, 2015).

Previo a la secuenciación de genomas, estudios bioquímicos y ultraestructurales han generado evidencia convincente de que las diatomeas han evolucionado a partir de una endosimbiosis secundaria ocurrida en algas rojas que se estima tuvo lugar entre 700 y 1200 millones de años (Parker *et al.*, 2008; Gibbs, 1981). Más adelante, de acuerdo a estudios de metabarcoding realizados por la campaña de muestreo global de plancton de Tara Oceans analizando la región hipervariable V9 del gen 18S ADNr han señalado a las diatomeas como el grupo fotosintético más abundante y el quinto grupo marino más numeroso (Malviya *et al.*, 2016; de Vargas *et al.*, 2015).

Las diatomeas en función del hábitat en donde se desarrollan se clasifican en planctónicas y bentónicas. Las primeras son aquellas que se encuentran suspendidas en la columna de agua mientras que las últimas se encuentran en los sedimentos marinos adheridas a una superficie (Kale y Karthick, 2015). Otra de las características distintivas de las diatomeas se encuentra la composición de la pared celular, principalmente de silicio, que son conocidas como frústulas (Kale y Karthick, 2015). Por otro lado, se ha reportado que algunas diatomeas sintetizan varios compuestos bioactivos con actividad antibiótica, inhibición enzimática entre otras (Kale y Karthick, 2015). Dentro de los compuestos más comercializados de las diatomeas se encuentran los lípidos y se distinguen por ser uno de los principales productores de ácidos grasos poliinsaturados en la cadena trófica acuática y se sabe que bajo ciertas condiciones algunas especies son capaces de producir hasta un 70 % de su peso seco en lípidos, por lo que se consideraran como un recurso potencial para la producción de biocombustibles (Cuellar-Bermúdez *et al.*, 2015; Draaisma *et al.*, 2013).

#### 1.2.1.1 Género *Chaetoceros*

El género *Chaetoceros* es considerado el más abundante y diverso de las diatomeas y que se encuentra ampliamente distribuido alrededor del mundo, principalmente en ecosistemas costeros y regiones de surgencia (Li *et al.*, 2017). Estudios metagenómicos recientes han avalado que este género es efectivamente uno de los más numerosos de las especies de

diatomeas planctónicas (Piredda *et al.*, 2018; Malviya *et al.*, 2016). Los miembros de este grupo abundan primordialmente en ecosistemas marinos y sólo unas cuantas especies en agua dulce; la mayoría se encuentran en la zona nerítica (10 a 200 metros de profundidad) del ecosistema marino y sólo unas pocas se encuentran en la región oceánica (Cupp, 1943). Actualmente se estiman entre 400 y 529 especies, sin embargo, debido a las numerosas descripciones inadecuadas y al gran número existente de sinónimos, el número actual de taxones aceptados son 218 (Guiry y Guiry, 2016).

Las especies del género *Chaetoceros* son diatomeas centrales cuyo tamaño puede oscilar entre 2 – 85 µm de diámetro y de 2 - 45 µm de longitud y que pueden crecer bajo temperaturas que pueden ir entre los – 2 °C hasta los 30 °C y salinidades que varían desde 18 hasta 38 ups (OBIS, 2012). Además, se caracterizan porque sus miembros forman cadenas de varias células, las cuales se encuentran separadas por aberturas y estructuras largas que sobresalen de cada uno de los extremos de las células que reciben el nombre de setas; también en la naturaleza pueden encontrarse las células de forma solitaria; se reproducen de forma asexual por división mitótica; cada división deriva en la disminución del tamaño de las células. Una vez que las células ya no pueden dividirse por mitosis, la reproducción cambia a sexual por medio de auxosporas. Se ha reportado que gran parte de este grupo de diatomeas se distingue por la formación de esporas en condiciones desfavorables de crecimiento y además de que dichas esporas pueden tolerar condiciones extremas tales como la deficiencia de nutrientes (Itakura, 2000).

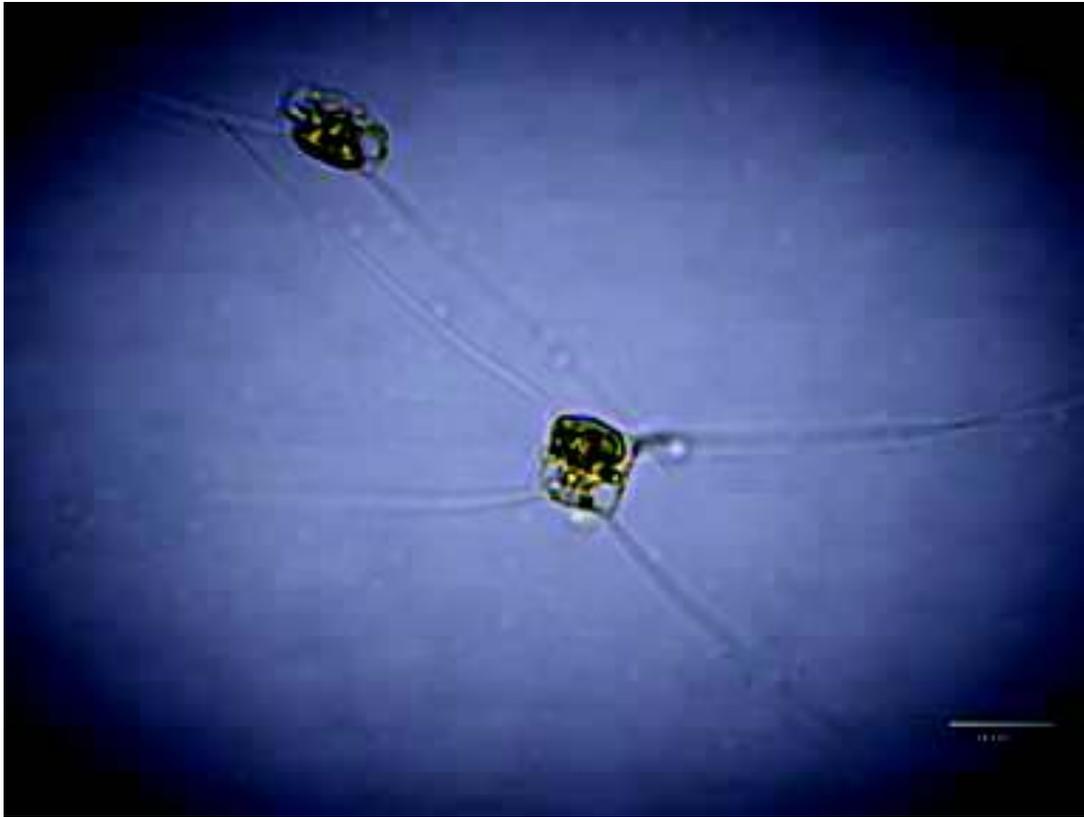
Las especies de este género también han despertado un interés comercial por su contenido lipídico, enfocándose algunas industrias en la obtención de ácidos grasos poliinsaturados para la producción de nutraceuticos o bien para su uso en acuicultura como alimento vivo para distintas especies de crustáceos y peces en su etapa larval (López-Elías *et al.*, 2003).

#### 1.2.1.1.1. *Chaetoceros muelleri*

*Chaetoceros muelleri* es una diatomea marina planctónica considerada una de las especies más representativas del género *Chaetoceros* perteneciente a la clase Bacillariophyceae (Wang *et al.*, 2014). Se caracteriza por ser unicelular, de morfología central, su tamaño varía entre 3 y 10 µm de diámetro y presenta cuatro setas en sus extremos (Brotosudarmo *et al.*, 2019; Pacheco-

Vega y Sánchez-Saavedra, 2009) (Figura 1). Además, se caracteriza por poseer un complejo de proteínas de unión de fucoxantina y clorofila, siendo fucoxantina el pigmento más abundante involucrada en la asimilación de luz y responsable del 60 % de la transferencia de energía a la clorofila *a* en diatomeas (Brotosudarmo *et al.*, 2019; Wang *et al.*, 2018). Por otra parte, dentro de su composición química destaca la presencia de silicio en su pared celular y se caracteriza por su capacidad de acumular compuestos de reserva tales como carbohidratos (principalmente crisolaminarina) y lípidos al ser cultivada bajo ciertas condiciones de cultivo o limitación de nutrientes (Lin *et al.*, 2018; De Stefano *et al.*, 2009; Størseth *et al.*, 2005)

*C. muelleri* es una de las especies más importantes utilizadas en acuicultura en el Noroeste de México como alimento vivo por su crecimiento rápido, capacidad de adaptación a condiciones adversas, relativa facilidad de cultivo y su alto valor nutricional, particularmente por su contenido lipídico rico en ácidos grasos poliinsaturados (principalmente el ácido eicosapentanoico, EPA) (Martínez-Córdova *et al.*, 2012; Pacheco-Vega y Sánchez-Saavedra, 2009; López-Elías *et al.*, 2005). Se ha demostrado que su uso en acuicultura es adecuado por su fácil digestión e influencia directa en el desarrollo y supervivencia de poslarvas de camarón promoviendo el fortalecimiento de su sistema inmune (Lemus *et al.*, 2006; Gutiérrez, 2002). Por otro lado, ha sido considerada como una de las especies de microalgas más adecuadas para su uso comercial y como un recurso potencial para la producción de biodiesel por el perfil de ácidos grasos obtenido al ser cultivada bajo concentraciones limitantes de nutrientes (De Jesus-Campos *et al.*, 2020; Lin *et al.*, 2018).



**Figura 1.** Célula de *Chaetoceros muelleri*, vista cingular. Barra de escala = 10  $\mu\text{m}$ .

### 1.3 Aplicaciones de las microalgas

Desde un punto de vista comercial, el objetivo de la biotecnología microalgal es desarrollar productos rentables y para ello es necesario identificar algún metabolito de interés y a las especies de microalgas que lo produzcan y acumulen (Pachiappan *et al.*, 2019). El conocimiento de la composición de la biomasa microalgal es esencial para seleccionar las especies adecuadas para la producción de una fracción específica, identificar los componentes presentes con potencial para ser considerados subproductos y a su vez estos puedan ser integrados a un proceso de biorefinería y evaluar estrategias en función de estos metabolitos de interés que faciliten su obtención y procesamiento (Velazquez-Lucio *et al.*, 2018). Las especies de microalgas de los géneros *Chlorella*, *Arthrospira* y *Dunaliella* son las más utilizadas comercialmente debido su contenido proteico alto, su perfil aminoacídico, valor nutricional y por contar con protocolos estandarizados para la obtención de un rendimiento alto de biomasa. Actualmente, el mercado

global de microalgas está valuado en 40 millones de dólares al año (Pachiappan *et al.*, 2019; Ventura *et al.*, 2017; Vigani *et al.*, 2015).

Las microalgas han despertado interés en investigación ya que son un recurso rico en compuestos bioactivos, los cuales han manifestado actividad antibacterial, antifúngica, antiprotoso y antiviral (Sosa-Hernández *et al.*, 2019). Otra de las propiedades característica de las microalgas, especialmente observada en cianobacterias, es el potencial para la degradación de surfactantes y herbicidas así como la capacidad de metabolizar o biotransformar compuestos químicos tales como el petróleo, fluometurón y lindano, obteniendo resultados prometedores al mezclar diferentes especies de microalgas (Mansy *et al.*, 2002); por otro lado, el tratamiento de aguas residuales mediante el uso de microalgas para disminuir los niveles de contaminación ha permitido que este grupo de organismos sea considerado con potencial para su uso en biorremediación (Scherer *et al.*, 2017).

En los últimos 20 años se ha incrementado el número de publicaciones relacionadas con la biotecnología microalgal, siendo el periodo comprendido entre 2015 y 2016 donde se registró la mayor cantidad de estudios (Vingiani *et al.*, 2019). También se ha observado un incremento en la generación de estudios ómicos, los cuales brindarán información que permitirá mejorar las cepas para la producción de compuestos de alto valor agregado mediante la sobreexpresión de genes, la identificación y caracterización de las rutas metabólicas involucradas en la producción de compuestos bioactivos y la búsqueda de genes con aplicación biotecnológica directa en genomas y transcriptomas de microalgas y así expandir el potencial de las microalgas como un recurso de interés biotecnológico (Lauritano *et al.*, 2019; Vingiani *et al.*, 2019; Lauritano *et al.*, 2018).

### **1.3.1 Biocombustibles**

El término “biocombustible” se define como el combustible producido por organismos vivos y se clasifican en cuatro generaciones principales (Demirbas, 2011). La "primera generación de biocombustibles" se produce directamente a partir de cultivos comestibles como maíz, trigo y soja, sin embargo, a pesar de que se pueden obtener beneficios a partir de ellos, su principal objetivo que es el rubro alimenticio se ve desvirtuado (Olabi, 2013). La "segunda generación de

biocombustibles" comprende cultivos no comestibles o residuos orgánicos (biomasa lignocelulósica) (Gómez *et al.* 2008); sin embargo, el uso tierras de cultivo es un inconveniente para varios biocombustibles de segunda generación. Para dar solución a los inconvenientes anteriores surge la " tercera generación de biocombustibles" que se derivan principalmente de algas (Singh *et al.* 2011), mientras que una "cuarta generación" utiliza algas o microorganismos fotosintéticos modificados mediante ingeniería metabólica para producir biocombustibles mediante la mejora de la fotosíntesis (Ravindran *et al.*, 2017).

Debido a su capacidad de síntesis de biomoléculas tales como carbohidratos, proteínas y lípidos, las microalgas están siendo consideradas como un sistema renovable viable para la producción de biocombustibles (Musa *et al.*, 2019). La producción sostenible de bioproductos a partir de microalgas tiene un gran potencial para convertirse en un biocombustible de próxima generación debido a su mayor productividad, su capacidad para crecer en tierras no cultivables y su alta eficiencia en la conversión de la energía solar (Khan *et al.*, 2018). Recientemente, ha aumentado el interés por la explotación de microalgas y cianobacterias para la producción de precursores de biocombustibles, como los lípidos en forma de TAG y almidón, que pueden ser transformados en biodiesel y bioetanol, respectivamente (Shuba y Kifle, 2018).

El agotamiento de los suministros de combustibles fósiles y el impacto negativo en el clima mundial del uso excesivo de recursos no renovables exige la producción de energía limpia, renovable y sostenible y las microalgas pueden ser la respuesta a esta problemática ya que varias características de su fisiología les permite ser consideradas para su evaluación como fuentes potenciales de biocombustibles entre ellas (Dismukes *et al.*, 2008; Hu *et al.*, 2008; Chisti, 2007):

- a) La eficiencia fotosintética de las microalgas es mayor que la de plantas superiores, ya que éstas son intrínsecamente convertidores de energía solar más eficientes.
- b) A diferencia de los cultivos terrestres, la ausencia de biopolímeros intratables elimina la necesidad de tratamientos previos para descomponer los productos celulares.
- c) Su diversidad metabólica y ecológica permite la selección de taxones que se adaptan para el crecimiento en los acuíferos disponibles localmente o que tienen características morfológicas que permiten una cosecha rentable.

d) La manipulación de sus productos finales mediante el control biosintético de la composición química mediante la disponibilidad de nutrientes y estrés abiótico.

La investigación en la producción de biocombustibles a partir de microalgas es extensa y continúa incrementándose; sin embargo, las especies más estudiadas hasta el momento en relación a la obtención de biocombustibles son *Chlamydomonas reinhardtii* y *Chlorella vulgaris*. Por otro lado, la ingeniería genética y las técnicas de biotecnología molecular también se están aplicando en microalgas para aumentar la productividad lipídica (Shuba y Kifle, 2018).

#### **1.4 Producción de aceites a partir de microalgas**

Las microalgas son consideradas como organismos oleaginosos, ya que son capaces de acumular lípidos. El contenido de lípidos de microalgas está considerablemente influenciado por las condiciones ambientales del hábitat y se ha observado que oscila entre 5 y 70 % de su peso seco (Mata *et al.*, 2010). Las principales clases de lípidos en microalgas son los lípidos de membrana (glicosilglicéridos, fosfoglicéridos y lípidos de betaína) y los lípidos de almacenamiento (triacilgliceroles) (Li-Beisson *et al.*, 2016). Las microalgas también poseen pequeñas cantidades de otras clases de lípidos como terpenoides, esfingolípidos y esteroides están presentes en diferentes porcentajes dependiendo de la especie (Li-Beisson *et al.*, 2019).

Los TAG son los lípidos más comunes que se encuentran en las microalgas, formados a partir de ácidos grasos y glicerol. Estos lípidos son utilizados por las microalgas como un compuesto de almacenamiento y son similares a los lípidos que se encuentran en el aceite de otros vegetales. Además de los TAG, las microalgas producen otros tipos de lípidos como los fosfolípidos (lípidos polares, que están compuestos de dos ácidos grasos y un grupo fosfato unidos a una molécula de glicerol, componentes esenciales de la membrana celular), glicolípidos (compuestos de una molécula de glicerol combinado con dos ácidos grasos que se encuentran en las membranas de cloroplastos) y sulfolípidos (ésteres de sulfato de glicolípidos que también se encuentran en las membranas de cloroplastos) (Mondal *et al.*, 2017; Khozin-Goldberg y Cohen, 2011).

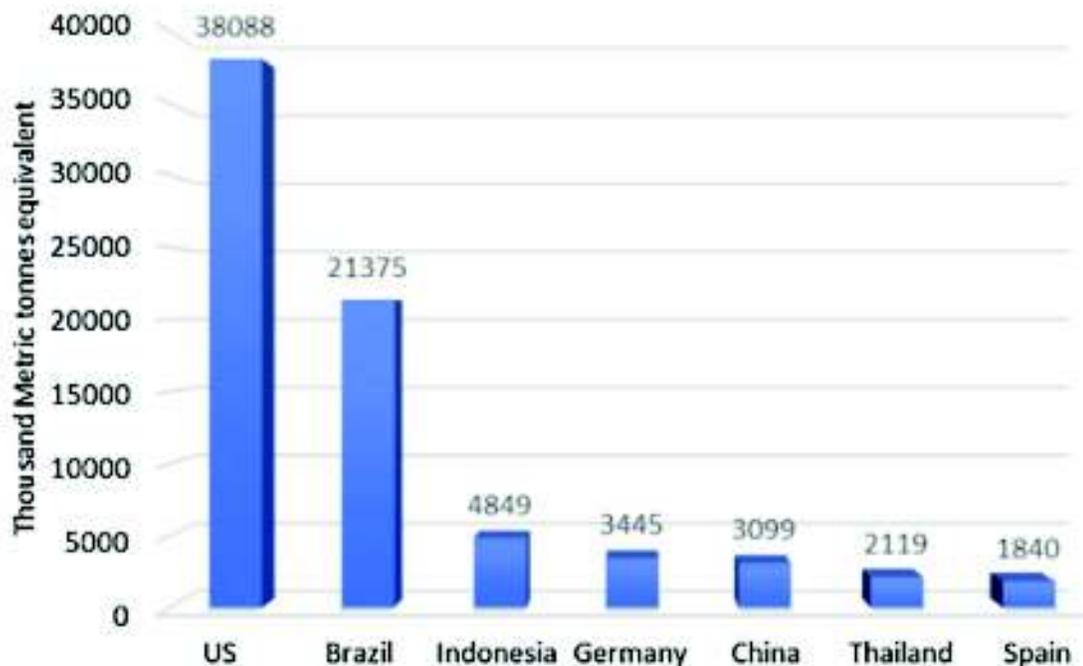
En plantas superiores, los ácidos grasos más comunes son aquellos cuya longitud de cadenas hidrocarbonadas constan de 16 a 18 átomos de carbono, entre saturados e insaturados

con hasta tres dobles enlaces presentes en sus cadenas. Por otro lado, en microalgas, los ácidos grasos más abundantes son aquellos que constan con cadenas hidrocarbonadas de 20 a 22 átomos de carbono. La producción de lípidos y su regulación en microalgas son específicas de la especie y están influenciadas por las condiciones ambientales (Khozin-Goldberg, 2016).

### **1.5 Producción de biodiesel a partir de aceites de microalgas**

Las microalgas ofrecen ventajas sobre los cultivos oleaginosos tradicionales para la producción de biocombustibles y compuestos bioactivos tales como productividad de lípidos alta, eficiencia fotosintética, capacidad para adaptación a diferentes condiciones ambientales, no compiten por tierra ni con cultivos comestibles, fijan con rapidez el carbono atmosférico y pueden ser cultivadas en aguas residuales y durante todo el año (Ahmad *et al.*, 2011). Existen diferentes formas de biocombustibles a partir de microalgas que incluyen combustión directa (biomasa), combustibles gaseosos (biogás, biometano, gas de síntesis y biohidrógeno) o combustibles líquidos (bioetanol, biobutanol y biodiesel) (Raheem *et al.*, 2018).

El biodiésel es una mezcla de ésteres metílicos o etílicos de ácidos grasos obtenidos por transesterificación del aceite extraído de microalgas, aceites vegetales o grasas animales. Estas materias primas oleosas están compuestas principalmente por TAG (90-98%) y pequeñas cantidades de mono y diglicéridos, ácidos grasos libres (1-5%) y cantidades trazas de otros subproductos como carotenos, tocoferoles, fosfátidos, fosfolípidos y compuestos de azufre y trazas de agua (Bozbas, 2008). El biodiésel es similar al diésel tradicional o "fósil" en sus características físicas y químicas y, por tanto, es considerado un combustible alternativo. El biodiesel posee varios beneficios tanto para los motores como para el medio ambiente, ya que se quema de manera más eficiente y da como resultado una reducción significativa de las emisiones de monóxido de carbono, hidrocarburos y material particulado como carbón y lodos (Chauhan y Shukla, 2011). Actualmente, los principales países productores de biodiésel en el mundo son Estados Unidos, Brasil, Indonesia, China, Tailandia y España (Ganesan *et al.*, 2020) (Figura 2).



**Figura 2.** Principales países productores de biodiesel. Tomada de Ganesan *et al.* (2020)

El uso de aceite de microalgas para la obtención de biodiésel requiere una alta productividad lipídica de las microalgas y de rápido crecimiento. Las condiciones de crecimiento óptimas conducen a la producción de grandes cantidades de biomasa, pero con contenidos de lípidos comparativamente bajos. La implementación de estrategias para favorecer el incremento de los lípidos en microalgas podría mejorar la economía del biodiésel (Sharma *et al.*, 2012). Las condiciones de crecimiento desfavorables de factores como la luz, la temperatura, los nutrientes (nitrógeno y fósforo), la salinidad y los metales pesados modulan las rutas biosintéticas de los lípidos en muchas especies de microalgas, lo que genera y acumula lípidos neutros, principalmente como TAG, apoyando a las microalgas para sobrevivir bajo tales condiciones adversas (Paliwal *et al.*, 2017; Ravindran *et al.*, 2017).

### **1.5.1 Perspectiva a futuro de la producción de biomasa a través del cultivo de microalgas.**

En México, todos los esfuerzos por producir biocombustible de microalgas han estado acotados a nivel de investigación. El desafío es por tanto el desarrollo de tecnologías de cultivo, cosecha

y transformación de microalgas en tiempos cortos. Las tecnologías para cultivo y cosecha de microalgas son a la fecha muy costosas para poder competir con los precios de los combustibles fósiles, y aunque en los últimos años se ha progresado mucho en estos aspectos, los cambios abruptos recientes en los precios del petróleo han hecho aún más difícil la entrada de las microalgas al mercado energético. Por lo anterior, los retos científicos y tecnológicos que se requieren lograr para disminuir los costos de producción de microalgas, así como de producción de biodiésel a costos competitivos son (Castillo *et al.*, 2017; Loera-Quezada y Olguín 2010):

- ✓ Aumentar la productividad de biomasa microalgal con el mayor contenido de lípidos del perfil adecuado.
- ✓ Lograr mayores rendimientos de biomasa microalgal en reactores que ofrezcan costos de inversión y de operación competitivos.
- ✓ Diseñar nuevos procesos de cosecha de biomasa por medio de biofloculación o uso de cepas con capacidades intrínsecas de auto floculación.
- ✓ Optimizar métodos de extracción de lípidos y de su transesterificación, que permitan mayores rendimientos a menores costos y que no afecten la calidad de la biomasa ni de los lípidos.
- ✓ Diseñar nuevos procesos para la generación de co-productos de alto valor agregado, proponiendo esquemas de biorrefinería para el aprovechamiento adecuado de todos los compuestos bioactivos que pueden estar presentes en la biomasa microalgal.
- ✓ Lograr la modificación genética de microalgas para incrementar la productividad de biomasa y lípidos (Castillo *et al.*, 2017; Loera-Quezada y Olguín 2010).

Finalmente, se requiere de una vinculación para que los empresarios inviertan en grupos de investigación de países en vías de desarrollo. El apoyo de las entidades gubernamentales más relacionadas a esta temática se hace fundamental para lograr que los grupos de investigación avancen en sus estudios y proyectos para que puedan llevar a cabo evaluaciones técnicas y económicas a una escala piloto o semiindustrial. Otro punto de gran importancia es que se debe promover la generación de nuevas políticas a nivel nacional e internacional para el fomento de

la investigación y desarrollo en la producción de biodiésel a partir de microalgas (Castillo *et al.*, 2017).

## **1.6 Cultivo de microalgas**

El cultivo de microalgas ha captado la atención en los últimos años, debido a sus aplicaciones tales como captura de CO<sub>2</sub>, biocombustibles, nutracéuticos, producción de biomoléculas y tratamiento de aguas residuales (Khan *et al.*, 2018). Al igual que la agricultura, el cultivo de microalgas cuenta con múltiples factores a considerar para que alcance una alta estabilidad y obtener una alta productividad para la generación de biomasa que pueda comercializarse a nivel mundial (Bracellar-Mendes y Namorato de Souza, 2018). Existen dos enfoques para el cultivo de microalgas a gran escala: sistema abierto y cerrado; la diferencia entre ambos reside en cómo operan (enfriamiento e intercambio de gases), influencia del medio ambiente (lluvia y presencia de especies ajenas al cultivo) y costos para la construcción del sistema (Jerney y Spilling, 2018). La elección de un tipo de sistema para el cultivo a gran escala de microalgas estará en función de la especie y el propósito del cultivo (Khan *et al.*, 2018).

### **1.6.1 Factores que influyen en el cultivo de microalgas**

Las microalgas, al igual que cualquier organismo, tienen requerimientos específicos para su crecimiento y supervivencia, incluyendo tanto factores físicos como químicos, tales como la luz, la temperatura, la concentración disponible de dióxido de carbono y la contribución de macronutrientes (como nitrógeno y fósforo) y oligoelementos. Además, se ha estudiado ampliamente, cómo estos factores no sólo afectan su crecimiento, sino también su composición bioquímica, observándose que las cantidades de estos productos metabólicos (proteínas, carbohidratos, lípidos, pigmentos) están estrechamente ligadas a las condiciones de crecimiento (Abo *et al.*, 2019, Juneja *et al.*, 2013).

#### 1.6.1.1. Efecto de la limitación de fósforo en microalgas

El fósforo (P) es un macronutriente importante que contribuye a varios procesos metabólicos, como vías de señalización, generación de energía y fotosíntesis; este elemento está presente en los fosfolípidos de la membrana celular como componente principal. También es un componente importante de ácidos nucleicos y juega un papel importante en la producción de energía química en forma de ATP y NADPH, que son moléculas indispensables durante el proceso de fotosíntesis (Díaz *et al.*, 2018; Brembu *et al.*, 2017; Lin *et al.*, 2016). Entre las respuestas que se han descrito en microalgas para hacer frente a la limitación de P se encuentran los cambios en la asignación de P, la síntesis de transportadores de fosfato, el uso de fuentes alternativas de P (DOP), la remodelación de la membrana celular y la modulación de procesos como la glucólisis y traducción de proteínas (Dyhrman *et al.*, 2012).

Actualmente, la comprensión de cómo las diatomeas detectan la disponibilidad de fósforo en el medio ambiente es escasa, sin embargo, recientemente se ha comenzado a describir los mecanismos por los que las diatomeas son capaces de detectar la presencia o ausencia de nutrientes. En *P. tricornutum* al ser cultivada en concentraciones bajas de fósforo (1.8  $\mu\text{M}$ ) por cuatro días se observó un incremento rápido y transitorio de los niveles de  $\text{Ca}^{2+}$  en el citosol al reabastecerse con medio de cultivo con una concentración de P de 36  $\mu\text{M}$ , dando indicios de que una ruta de señalización dependiente de  $\text{Ca}^{2+}$  puede estar involucrada la homeostasis celular de P (Helliwell *et al.*, 2020). Hasta el momento, la activación de la ruta de señalización por  $\text{Ca}^{2+}$  en diatomeas no se ha observado como una respuesta ante la deficiencia de otros nutrientes y metales traza, lo que pudiera indicar que es una respuesta característica de la limitación por P en estos organismos (Liu *et al.*, 2017).

Cuando las microalgas se enfrentan a una deficiencia de P, una estrategia común es la inducción de genes que codifican a proteínas transportadoras de alta afinidad de fosfatos, nucleotidasas y fosfatasas alcalinas, las cuales son responsables del transporte y la asimilación eficiente del  $\text{P}_i$ , con el objetivo de garantizar los niveles de este nutriente en las células para su correcto funcionamiento (Dell'Aquila y Maier, 2020; Alipanah *et al.*, 2018; Zhang *et al.*, 2016; Dyhrman *et al.*, 2012); en condiciones óptimas de crecimiento, las microalgas pueden almacenar el exceso de P en forma de Poli-P, los cuales pueden ser utilizados cuando el suministro externo se vuelve limitante (Solovchenko *et al.*, 2019a,b).

Por otra parte, la remodelación de la membrana celular mediante la sustitución de fosfolípidos por lípidos que no contienen P en su composición es otro de los mecanismos comunes en cianobacterias y algas eucariotas (Brembu *et al.*, 2017). Los fosfolípidos representan un reservorio importante de P celular que se moviliza en respuesta a bajas concentraciones de este nutriente y su degradación libera P para permitir que las células de microalgas se dividan incluso en condiciones adversas y los fosfolípidos son reemplazados por sulfolípidos, galactolípidos y lípidos de betaína (BL) (Shemi *et al.*, 2016; Van Mooy *et al.*, 2009)

Otro de los procesos que se ven afectados por la limitación de P es la fotosíntesis. En diatomeas cultivadas bajo la limitación de P, algunos efectos inmediatos son la inhibición de la síntesis de sustratos requeridos para el ciclo de Calvin y consecuentemente en la reducción de los niveles de fosforilación fotosintética, tasa fotosintética y la síntesis de moléculas como NADP y NADPH afectando el transporte de electrones a través del fotosistema II (PSII) (Barsanti y Gualtieri, 2014; Mittler, 2002). Por otro lado, genes que codifican a proteínas de unión de pigmentos tales como clorofila y fucoxantina se han visto reprimidos bajo limitación por P en diatomeas, observándose una disminución en el contenido de estos pigmentos (Alipanah *et al.*, 2018; Feng *et al.*, 2015). Aunque las moléculas de clorofila no contienen P, la disminución de los niveles de fotosíntesis como resultado de una menor disponibilidad de P puede conducir a una menor demanda de absorción de fotones por parte de las clorofilas y, en última instancia, a la supresión de la síntesis de clorofila (Alipanah *et al.*, 2018; Liu *et al.*, 2011).

Aunque la fotosíntesis, y en consecuencia la actividad del ciclo de Calvin, se reduce en las células microalgales limitadas en P, la producción de estructuras de carbono sigue superando las necesidades celulares; al disminuir la síntesis de proteínas y fosfolípidos, este exceso de carbono es redireccionado hacia la biosíntesis de compuestos de reserva en forma de TAG o excretado fuera de la célula como polisacáridos extracelulares (EPS, por sus siglas en inglés) (Brembu *et al.*, 2017; Alcoverro *et al.*, 2000). Con respecto a lípidos, se ha observado que bajo la limitación o deficiencia de P, el contenido total de ácidos grasos aumenta en algunas especies de microalgas; sin embargo, la expresión de algunos genes involucrados en la biosíntesis de los ácidos grasos tales como *ACC* que codifica a la acetil-CoA carboxilasa que es una enzima clave en el inicio de la síntesis de ácidos grasos y aquellos que codifican para desaturasas como *SAD* y *D9SD* permanece constante o disminuye y los autores explican que esta diferencia en los

niveles de expresión de estos genes puede atribuirse al grado de limitación de P, la especie evaluada y la duración del cultivo (Alipanah *et al.*, 2018; Shih *et al.*, 2015; Dyhrman *et al.*, 2012).

En relación con la acumulación de carbohidratos, dentro de los genes que se han visto inducidos bajo condiciones limitantes de P en diatomeas se han observado aquellos involucrados en la iniciación de la gluconeogénesis a partir del piruvato tales como los que codifican para piruvato carboxilasa (PC), fosfoenolpiruvato carboxicinasa (PEPCK) y fructosa-1,6-bifosfato (FBP) (Alipanah *et al.*, 2018; Yang *et al.*, 2014; Jitrapakdee y Wallace, 1999). Además, bajo deficiencia de P, se ha observado una mayor inducción de los genes que participan en la degradación de la crisolaminarina en comparación con su biosíntesis como es el caso de *GCK* que codifica a una glucocinasa, lo que sugiere que la degradación de este compuesto podría ser una estrategia para la síntesis de polisacáridos más complejos (Cruz de Carvalho *et al.*, 2016; Dyhrman *et al.*, 2012).

La comprensión amplia de la respuesta de la limitación por P en diatomeas se ha derivado a partir de la aparición de las ciencias ómicas y de las técnicas de secuenciación masiva, las cuales han permitido la disponibilidad de varios genomas y el desarrollo de análisis transcriptómicos, proteómicos y metabólicos (Dell'Aquila y Maier, 2020; Brembu *et al.*, 2017; Bowler *et al.*, 2008; Armbrust *et al.*, 2004). La mayoría de estudios ómicos realizados en diatomeas relacionados con su respuesta ante la limitación por P se han basado en el estudio de transcriptomas, ya sea mediante el uso de microarreglos o por la técnica de RNA-seq, siendo el primer estudio llevado a cabo mediante este enfoque en *T. pseudonana* (Dyhrman *et al.*, 2012).

Aunque los estudios ómicos realizados hasta el momento en diatomeas para evaluar la limitación de P difieren entre sí con respecto a las especies estudiadas y el diseño experimental, la visión general del conjunto de información generada es consistente ya que se ha corroborado que la limitación de este nutriente genera un impacto significativo en el metabolismo global de las diatomeas y de la implementación de estrategias moleculares que pueden ser clasificadas tanto como específicas de la limitación de P como de respuestas secundarias generadas del estrés (Dell'Aquila y Maier, 2020; Brembu *et al.*, 2017).

## 1.7 Transcriptómica

La transcriptómica es el conjunto de técnicas utilizadas para el estudio de los transcriptomas de organismos, los cuales se definen como la suma total de todos los transcritos presentes en una célula, tejido u órgano bajo condiciones fisiológicas o específicas en un momento determinado, comprendiendo tanto moléculas de ARN codificantes como no codificantes (Lowe *et al.*, 2017; Assis *et al.*, 2014). La determinación de los niveles de expresión de los genes de un organismo o tejido, en condiciones o momentos determinados provee información valiosa que permite inferir las funciones de los genes que no han sido anotados previamente (Lowe *et al.*, 2017). Actualmente, existen dos técnicas contemporáneas para el desarrollo de este tipo de estudios: los microarreglos, los cuales cuantifican un conjunto de secuencias predeterminado y RNA-seq la cual emplea las técnicas de secuenciación masiva para abarcar todas las secuencias de un organismo (Lowe *et al.*, 2017).

### 1.7.1 RNA-seq

RNA-seq hace referencia a la secuenciación de transcritos (ADNc) en los cuales su abundancia es definida por el número de conteos o reads alineados a cada transcrito. Esta técnica se vio influenciada por el desarrollo de las tecnologías de secuenciación masiva (Wang *et al.*, 2009). Desde su inicio en 2006 hasta la fecha, RNA-seq se ha usado para el estudio de múltiples transcriptomas, lo que ha permitido obtener importantes avances en la biología de organismos modelo y sobretodo en organismos no modelo (Lowe *et al.*, 2017). Esta herramienta posee ciertas ventajas evidentes sobre otros métodos empleados para desarrollar análisis transcriptómicos como la detección de transcritos conocidos y novedosos, así como de eventos de splicing alternativo y fusión de genes (Pereira *et al.*, 2017). Adicionalmente, RNA-seq permite detectar genes que se expresan a niveles bajos, además de polimorfismos de una sola base e isoformas, detección de transcritos no codificantes y cambios en las secuencias a nivel postranscripcional (mutaciones o edición) (Lowe *et al.*, 2017; Pereira *et al.*, 2017).

Con el fin de desarrollar un experimento de RNA-seq exitoso, es crítico un buen diseño experimental, el cual estará sujeto a factores extrínsecos (costo, tiempo, disponibilidad de muestras) e intrínsecos (variabilidad transcripcional entre tejidos, muestras y organismos). La

cantidad de recursos disponibles es generalmente, el factor principal que delimitará la decisión del investigador y de ahí será necesario definir el objetivo del experimento de RNA-seq para elegir el enfoque más apropiado (Pereira *et al.*, 2017).

#### 1.7.1.1 Análisis de datos generados por RNA-seq

Los experimentos de RNA-seq generan una gran cantidad de lecturas de secuenciación (secuencias/*reads*), las cuales requieren pasar por un procesamiento para la obtención de información relevante (Zhao *et al.*, 2016). El análisis de datos implica el uso de varias herramientas bioinformáticas las cuales variarán en función del diseño y objetivos del experimento (Sahraeian *et al.*, 2017); de ahí que no existe un protocolo adecuado y definido que aplique a todos los posibles enfoques que pueden ser abordados mediante la tecnología de RNA-seq (Conesa *et al.*, 2016). Sin embargo, un proceso de análisis transcriptómico por RNA-seq, generalmente puede desglosarse en los siguientes pasos: análisis y filtrado de calidad de secuencias, alineamiento, ensamblado, anotación funcional, cuantificación y expresión diferencial, enriquecimiento y validación del transcriptoma (Pereira *et al.*, 2017; Kukurba y Montgomery, 2015).

✓ Análisis y filtrado de calidad de secuencias: El análisis de RNA-seq comienza con los archivos que contienen los datos crudos obtenidos del proceso de secuenciación (lecturas/secuencias/*reads*). El análisis y filtrado de calidad de los datos crudos es un paso muy importante que facilita la detección de sesgos generados durante el procedimiento de secuenciación que, si no se eliminan correctamente, pueden generar problemas como un mapeo incorrecto durante la alineación y afectar los resultados finales (Conesa *et al.*, 2016). Entre las herramientas que se utilizan para llevar a cabo el análisis y filtrado de calidad de las secuencias se encuentra FASTQC (diseñado principalmente para la plataforma Illumina) la cual es una de las más populares y utiliza un formato estándar utilizado para almacenar las secuencias de nucleótidos junto con una puntuación de calidad por base en la escala de registro Phred que van de 0 a 40 (recomendándose valores de calidad superior a 25,  $Q \geq 25$ ) (Andrews, 2010).

✓ Alineamiento: Una vez que las lecturas han sido filtradas, el siguiente paso es el alineamiento de las secuencias a un genoma o transcriptoma de referencia. Durante este paso, las secuencias de lectura de los fragmentos de ADNc procedentes de la retrotranscripción de los transcritos de ARN se alinean con los genomas de referencia, sin embargo, si la especie estudiada carece todavía de un genoma de referencia, las lecturas pueden ser primero ensambladas *de novo* y posteriormente los reads pueden utilizarse ahora para alinearse al transcriptoma ensamblado y determinar el porcentaje de ellos que se encuentran representados; se recomienda que entre un 70 y 90 % de los reads se alineen al genoma de referencia o al transcriptoma ensamblado (Geniza *et al.*, 2017).

✓ Ensamblado: El proceso de ensamblado pudiera definirse como la reconstrucción del transcriptoma mediante la obtención de una colección de fragmentos, donde cada uno de ellos representa un transcrito de longitud completa (lo más precisa posible), a partir de lecturas cortas (Smith-Unna *et al.*, 2016; Martin *et al.*, 2010). El proceso de ensamblado es un paso fundamental en el análisis transcriptómico ya que, dependiendo de la calidad del mismo, se puede utilizar para el descubrimiento de transcritos novedosos y genes desconocidos, computar el nivel de expresión de los genes e incluso aquellos transcriptomas que sean correctamente anotados pueden ser de utilidad para mejorar la anotación de los genomas de los organismos de estudio en caso de que estén disponibles (Trapnell *et al.*, 2012). Existen dos tipos de enfoques para realizar el ensamblado de un transcriptoma: 1-) guiado por un genoma de referencia o 2-) '*de novo*'. La estrategia de ensamblado de transcriptoma guiado por un genoma de referencia hace alusión a alinear primero las lecturas de secuenciación a un genoma de referencia y luego ensamblar las alineaciones superpuestas en transcritos mientras que el enfoque de ensamblado '*de novo*', hace referencia a la reconstrucción directamente de las lecturas superpuestas en transcritos utilizando la redundancia de las propias lecturas de secuenciación; en el caso de los organismos sin genoma de referencia, sólo es posible el ensamblaje de su transcriptoma por esta vía (Lu *et al.*, 2013).

✓ Anotación funcional: En biología computacional, la anotación es el proceso de identificación de la ubicación y secuencia de los elementos genómicos y/o la asignación de una función biológica a los mismos. El término "funcional" tiene un significado diferente al de los biólogos evolutivos, interesados en la conservación, y los biólogos experimentales, interesados en las funciones bioquímicas; en este sentido, implica la asociación de una descripción funcional con un gen, después de identificar una secuencia similar utilizando herramientas bioinformáticas y bases de datos (Mudge y Harrow, 2016).

✓ Cuantificación y expresión diferencial: el objetivo principal de la mayoría de los proyectos de RNA-seq es cuantificar y comparar la expresión génica en diferentes condiciones e inferir la interpretación biológica a la expresión diferencial a nivel de genes o de transcritos (Pereira *et al.*, 2017). Para este fin se realiza el conteo de lecturas y normalización de las muestras mediante distintos programas bioinformáticos y métodos estadísticos (Pereira *et al.*, 2017; Conesa *et al.*, 2016). Otro aspecto importante a considerar durante un análisis de expresión diferencial es el definir los puntos de corte que indicarán qué genes o transcritos se encuentran expresados significativamente; dos de los parámetros utilizados para este propósito son el 'log<sub>2</sub>fold-change' o 'fold-change' que se define como el grado de cambio en número de veces en que la expresión de un gen se ve modificada con respecto a un control y la tasa de descubrimiento de falsos positivos (FDR, por sus siglas en inglés) que se define como la proporción de casos positivos que la prueba detecta como negativo (Schurch *et al.*, 2016; Tusher *et al.*, 2001).

✓ Análisis de enriquecimiento: la identificación de los genes que se encuentran expresados diferencialmente es sólo el punto de partida para obtener los conocimientos biológicos sobre sistemas experimentales, etapas de desarrollo, comprensión de enfermedades o mecanismos moleculares (Zhao *et al.*, 2016). El análisis y la interpretación de estos datos representan un desafío, por lo que, para comprender el contexto biológico de los genes expresados diferencialmente (DEG, por sus siglas en inglés), se realiza un análisis de enriquecimiento funcional, el cual es un método computacional capaz de determinar si un conjunto predefinido de genes muestra diferencias significativas entre las muestras (Pereira *et al.*, 2017). El análisis

de enriquecimiento transforma la información de los perfiles de expresión génica en un resumen de las rutas metabólicas que se ven influidas por el efecto de la variable de estudio (Zhao *et al.*, 2016). Las rutas metabólicas se someten a pruebas estadísticas para comprobar su representación en la lista de genes experimentales, las cuales toman en cuenta el número de genes detectados en el experimento, su clasificación relativa y el número de genes anotados en una ruta metabólica de interés (Reimand *et al.*, 2019).

✓ Validación del transcriptoma: aunque RNA-seq proporciona una estimación precisa y exacta de la abundancia de ARN, se recomienda que los resultados obtenidos mediante esta técnica se validen mediante la técnica de PCR cuantitativo, también conocida como RT-qPCR (cuando se refiere al proceso de retrotranscripción del ARN previo a la reacción) o PCR en tiempo real, ya que todavía se considera la técnica por excelencia para la cuantificación de la expresión génica (Pereira *et al.*, 2017; Bustin *et al.*, 2005). Sin embargo, la necesidad de validar los estudios de RNA-seq mediante RT-qPCR sigue siendo cuestionable ya que se ha observado en estudios previos una alta correlación entre los resultados de RNA-seq y de RT-qPCR (Wu *et al.*, 2013; Camarena *et al.*, 2010; Core *et al.*, 2008). Existe un debate en cuanto al número de réplicas biológicas del experimento de RNA-seq consideradas suficientes (a partir de 3 ó 6) para el considerar no realizar la validación por RT-qPCR, no obstante, sigue siendo recomendable precisamente por la alta correlación existente entre ambas técnicas ya que a través de RT-qPCR se pueden confirmar, mediante el análisis de réplicas biológicas diferentes a las que fueron secuenciadas, los resultados y conclusiones generadas a partir de RNA-seq (Pereira *et al.*, 2017; Fang y Cui, 2011).

## **1.8 Análisis *in silico* de proteínas**

Las proteínas son macromoléculas complejas que poseen funciones esenciales en la actividad celular y por ende en varios procesos o reacciones biológicas por lo que su estudio es fundamental para determinar su función y comprender los mecanismos en los que se encuentran involucradas (Singh *et al.*, 2016). La identificación de la función de una proteína *in vivo* sigue requiriendo un enfoque de genómica funcional que implica la detección de actividades

promotoras específicas de los tejidos, la localización subcelular, las posibles modificaciones posteriores a la traducción y el estudio de mutantes en donde las proteínas de interés sean sobreexpresadas o silenciadas (Free *et al.*, 2009). Este proceso suele ser costoso, en ocasiones inasequible y requiere de una inversión alta de tiempo (Mooney *et al.*, 2011). Por otra parte, los avances en el desarrollo de tecnologías de secuenciación de nueva generación en transcriptómica y proteómica basada en la espectrometría de masas han permitido la generación de una cantidad masiva de datos a nivel de ADN, ARN y proteínas que junto con la aparición y disponibilidad de un número considerable de herramientas bioinformáticas han representado una alternativa para la predicción de las características funcionales y estructurales de proteínas de organismos de interés a partir de sus secuencias aminoacídicas (Yavuz y Öztürk, 2017; Drozdetskiy *et al.*, 2015). Por lo tanto, la anotación *in silico* de una proteína implica principalmente la predicción de su función por medio de la homología, las posibles interacciones y las modificaciones postraduccionales, así como su localización subcelular (Yavuz y Öztürk, 2017; Moulton, 2008).

Mooney *et al.* (2011) mencionan que los métodos de predicción de características funcionales y estructurales de proteínas se dividen en dos: comparativos y ‘*de novo*’. Los métodos de predicción comparativos son aquellos que evalúan la homología de las secuencias del organismo de estudio con aquellas cuya estructura ya es conocida y se considera que este enfoque es más preciso que los métodos ‘*de novo*’ que únicamente utilizan softwares especializados para la predicción de la estructura secundaria de las proteínas de interés.

Dentro de los elementos conformacionales de las proteínas que son importantes considerar para su análisis mediante un enfoque comparativo son los dominios y motivos; los primeros se definen como unidades funcionales y/o estructurales conservadas de una proteína que generalmente son responsables de una función o interacción en particular que contribuye con el funcionamiento general de la proteína (Sangrador-Vegas y Mitchell, 2011). Por otra parte, los motivos son conocidos como patrones de secuencias cortas, generalmente de longitud fija, que pueden representar importantes características estructurales o funcionales en las secuencias de ácido nucleicos y proteínas, como sitios de unión de transcripción, uniones de empalme, sitios activos o interfaces de interacción (Mohamed *et al.*, 2016). La identificación de motivos y

dominios en las proteínas es un aspecto importante de la clasificación de las secuencias de proteínas y su anotación funcional (Xiong, 2006).

## II. JUSTIFICACIÓN

El P es un nutriente esencial para el crecimiento y desarrollo de las microalgas y su concentración tiene un efecto en su composición química y dicho efecto, ha despertado un interés específico sobre la síntesis de lípidos. Se han realizado estudios diversos bajo condiciones limitantes de P en varias diatomeas e incluso se han realizado análisis transcriptómicos, pero para el género *Chaetoceros* la información aún es limitada. Debido a la importancia que *Chaetoceros muelleri* representa en la acuicultura en el Noroeste de México por su crecimiento rápido, su calidad nutricia, capacidad de adaptación a las condiciones climatológicas de esta región y la acumulación de lípidos bajo concentraciones limitantes de nutrientes, es interesante conocer los procesos biológicos y moleculares involucrados en su respuesta ante la falta de P, ya que hasta el momento los estudios realizados en esta especie se han centrado básicamente en su composición bioquímica y producción de biomasa. Actualmente el uso de técnicas moleculares y el aporte de las nuevas áreas del conocimiento como la transcriptómica, proteómica y metabolómica facilitan la comprensión global de la biología, la interacción de los organismos y su respuesta ante los factores que les rodean en su medio ambiente. De ahí que la información generada a partir de estas áreas del conocimiento y de las tecnologías de nueva generación permitirá un mejor entendimiento de la regulación de las rutas metabólicas empleadas por la diatomea en respuesta a la limitación por P y desarrollar estudios posteriores. En virtud de lo anterior, se pretende relizar un análisis transcriptómico bajo distintas concentraciones limitantes de P y su relación con la síntesis de lípidos en *C. muelleri* por medio de RNA-seq.

### PLANTEAMIENTO DE LA PREGUNTA DE INVESTIGACIÓN

En base a los antecedentes ya mencionados se plantea la siguiente pregunta de investigación:

¿Cómo afectan las distintas concentraciones de fósforo en el nivel de expresión de los genes y la síntesis de lípidos en *Chaetoceros muelleri*?

### **III. HIPÓTESIS**

Habr  una mayor cantidad de genes reprimidos en condiciones limitantes de f sforo en comparaci n con las condiciones de exceso de este nutriente en *Chaetoceros muelleri* y existir  una relaci n en la expresi n diferencial con la acumulaci n de l pidos.

## IV. OBJETIVOS

### 4.1 Objetivo general

Analizar el transcriptoma de *Chaetoceros muelleri* en respuesta a diferentes concentraciones limitantes de P y su relación con la acumulación de aceites.

### 4.2 Objetivos específicos

- ✓ Determinar el efecto de la limitación y el exceso de fósforo en el crecimiento, producción de biomasa y composición proximal de *Chaetoceros muelleri*.
- ✓ Evaluar la producción de biomasa y la calidad bioquímica de los cultivos de *Chaetoceros muelleri* en los medios limitantes y con exceso de fósforo.
- ✓ Cuantificar el perfil de los ácidos grasos de *Chaetoceros muelleri* bajo diferentes condiciones de fósforo.
- ✓ Analizar el perfil transcriptómico de *Chaetoceros muelleri* cultivada bajo diferentes concentraciones de fósforo utilizando la tecnología RNA-seq.
- ✓ Predecir *in silico* las características de las secuencias aminoacídicas de las proteínas transportadoras de fósforo NPT2B y VTC4 y de las enzimas ACC y FAD2 involucradas en la biosíntesis de ácidos grasos en *Chaetoceros muelleri*.

## V. MATERIALES Y MÉTODOS

### 5.1 Material biológico

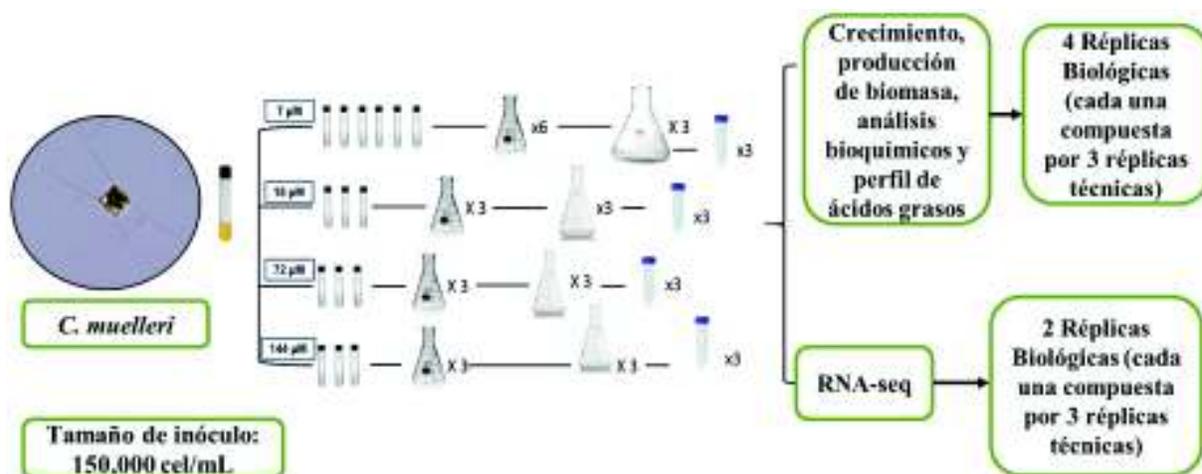
La cepa de *Chaetoceros muelleri* utilizada en este estudio fue aislada en Bahía de Kino (Coordenadas UTM: zona 12 N, X 4077, Y 3187302) (Sonora, México) y mantenida bajo condiciones de laboratorio en el cepario de microalgas del Departamento de Investigaciones Científicas y Tecnológicas (DICTUS) de la Universidad de Sonora.

### 5.2 Condiciones de crecimiento

La diatomea *Chaetoceros muelleri* fue cultivada en condiciones axénicas, creciendo en medio F (Guillard y Ryther, 1962), siendo modificadas las concentraciones de fosfatos. Se consideraron cuatro tratamientos con diferentes concentraciones de P, incluyendo un control (fosfatos: 10 g/L, 72  $\mu\text{M}$ ), un exceso (fosfatos: 20 g/L, 144  $\mu\text{M}$ ) y dos tratamientos con deficiencia en P que, por motivos prácticos, se denominaron limitante A (fosfatos 2.5 g/L, 18  $\mu\text{M}$ ) y limitante B (fosfatos: 1 g/L, 7  $\mu\text{M}$ ); las concentraciones del resto de nutrientes que componen el medio F permanecieron constantes. Los cultivos se mantuvieron bajo condiciones controladas de temperatura ( $22 \pm 1^\circ\text{C}$ ), iluminación constante ( $220 \mu\text{mol m}^{-2} \text{s}^{-1}$ , Phillips F96T8/TL865, 60 W), salinidad (30 psu) y sin aeración.

El efecto de las diferentes concentraciones de P en *C. muelleri* se evaluó primero mediante una fase de aclimatación de la diatomea en tubos de ensayo con 10 mL de medio de cultivo durante una semana para después realizar un escalamiento a matraces de 250 mL por 5 días y finalmente a matraces de 1 L por cuatro días. Previo a la inoculación en matraces de 1 L, se realizaron conteos para cada uno de los tratamientos y se definió como concentración de inóculo para cada tratamiento de 150,000 células por mililitro. Cada tratamiento fue realizado por triplicado (tanto en la fase de aclimatación como en el proceso de escalamiento). Para la evaluación del crecimiento, producción de biomasa, análisis bioquímicos y ácidos grasos se consideraron cuatro réplicas biológicas las cuales estuvieron compuestas por tres réplicas técnicas por cada tratamiento; por su parte, para el análisis de RNA-seq se consideraron dos

réplicas biológicas cada una de ellas compuestas por tres réplicas técnicas por tratamiento (Figura 3).



**Figura 3.** Diseño experimental para evaluar el efecto de las diferentes concentraciones de P en la diatomea *C. muelleri*.

### 5.3 Determinación de la concentración celular

Se llevó a cabo un escalamiento de cultivo de la diatomea *Chaetoceros muelleri*, creciendo en cada uno de los tratamientos mencionados anteriormente, iniciando en tubos de cultivo para después pasar a matraces Erlenmeyer de 250 mL, considerándose cuatro réplicas por tratamiento. La determinación de la concentración celular se obtuvo haciendo mediciones diarias por nueve días, de manera indirecta utilizando un espectrofotómetro (Spectronic 20D+ Thermo Scientific) haciendo lecturas a 550 nm y de forma directa mediante conteos celulares haciendo uso de un hematocitómetro de 0.1 mm de profundidad, utilizando un microscopio óptico y aplicando la siguiente fórmula:

$$\text{No. Células/mL} = (\text{No. total de células/No. de cuadros contados}) * 10,000.$$

Se estimaron también las tasas de replicación celular:  $\mu$  máxima,  $\mu$  acumulada y  $\mu$  promedio, siguiendo las siguientes fórmulas (Andersen, 2005):

Tasa de crecimiento máxima:  $\mu_{\max} = \text{Log}_2 B_n - \text{Log}_2 B_{n-1} / (t_n - t_{n-1})$

Tasa de crecimiento acumulada:  $\Sigma\mu$

Tasa de crecimiento promedio:  $\mu_{\text{prom}} = \text{Log}_2 B_n - \text{Log}_2 B_0 / (t_n - t_0)$

#### 5.4 Determinación de biomasa y análisis bioquímicos

Para la determinación de biomasa y análisis bioquímicos, la obtención de muestra se realizó al día 4 de iniciado el cultivo, ya que en este momento todos los tratamientos se encontraban en la fase logarítmica tardía de crecimiento.

Para la determinación de biomasa (peso seco), los cultivos fueron homogenizados y las muestras fueron obtenidas mediante filtración de 200 mL de medio de cultivo a través de filtros de microfibras de vidrio Whatman GF-C de 47 mm de diámetro previamente lavados, secados y pesados. Luego los filtros con las células filtradas fueron lavados con formiato de amonio (36 g L<sup>-1</sup>) para remover residuos de sales y se secaron en una estufa de convección a 60 °C por 8 h y se pesaron en una balanza analítica (Ohaus Explorer Analytical Balance Ohaus E011403). Posteriormente, estos mismos filtros fueron incinerados a 480 °C en una mufla por 12 h para determinar el contenido de materia inorgánica (cenizas) y se pesaron (Ohaus Explorer Analytical Balance Ohaus E011403). Finalmente, se determinó el contenido de materia orgánica mediante la diferencia entre biomasa y cenizas.

El contenido de clorofila *a* fue determinado mediante la obtención de 50 mL de cultivo previamente homogenizado y centrifugado a 2000 g por 5 min. Los pigmentos fotosintéticos fueron extraídos del pellet obtenido de la centrifugación, añadiéndose acetona al 90 % y después se sonicó por 5 min (Sonic 1510 R-DTH Branson Ultrasonics Corporation, CT, USA) y finalmente los extractos fueron almacenados en oscuridad por 24 h a 4 °C. El contenido de clorofila *a* se determinó espectrofotométricamente y haciendo uso de la ecuación propuesta por Jeffrey y Humphrey (1975):

$$\text{Clorofila } a = 11.47 E_{664} - 0.40 E_{630}$$

Para la determinación del contenido total de lípidos y proteínas, se tomaron muestras de cada tratamiento y se filtraron a través de filtros de microfibras de vidrio Whatman GF-C de 25 mm de diámetro; posteriormente, los filtros fueron lavados con formiato de amonio (36 g L<sup>-1</sup>) y

se almacenaron a  $-20^{\circ}\text{C}$  previo a los análisis bioquímicos. El contenido total de lípidos se determinó colorimétricamente siguiendo la técnica de Pande *et al.* (1963). Una vez extraídos los lípidos, se utilizó el método de Bligh y Dyer (1959) modificado por López Elías *et al.* (1993). El contenido total de proteínas se determinó también colorimétricamente por el método de Lowry *et al.* (1951) modificado por Malara y Charra (1972) y siguiendo las recomendaciones para su extracción de Clayton *et al.* (1988) y López Elías (1993). La cuantificación de lípidos y proteínas totales se realizó mediante el uso de espectrofotometría (CARY 100 BIO UV-Visible Spectrophotometer).

### **5.5 Perfil de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P.**

El análisis de perfil de ácidos grasos se llevó a cabo en el Laboratorio de Bioquímica y Calidad de Productos Pesqueros del Centro de Investigación en Alimentación y Desarrollo, A.C (CIAD), para lo cual se utilizaron muestras previamente liofilizadas de cada uno de los tratamientos de *C. muelleri*.

Los lípidos fueron extraídos siguiendo el método propuesto por Ryckebosch *et. al* (2012). Se pesaron 350 mg de muestra y se añadió una mezcla de cloroformo y metanol (1:1) y se homogenizó en vórtex por 30 seg. Los solventes se evaporaron a  $60^{\circ}\text{C}$  bajo una atmósfera de nitrógeno. Para la derivatización se utilizó una mezcla de ácido sulfúrico (al 1 % en metanol) y tolueno y se incubó por toda la noche a  $50^{\circ}\text{C}$ . Finalmente, se añadió hexano y una solución de NaCl (5%) para después dejarse reposar hasta que se observaron dos fases. La muestra recibió un flujo de nitrógeno y la fase superior fue transferida un vial y congelada a  $-80^{\circ}\text{C}$  hasta su análisis.

La determinación del perfil de ácidos grasos de todas las muestras se llevó a cabo mediante la inyección de 1  $\mu\text{L}$  de muestra en un cromatógrafo de gases (Agilent 6890 Series) equipado con una columna analítica de sílice modelo Agilent DB-23 (diámetro de 0.25 mm, grosor de 0.25  $\mu\text{m}$  y longitud de 60 m). El gas acarreador fue  $\text{N}_2$  y el modo de detección fue por ionización de flama (FID). La identificación de los ácidos grasos fue basada en una mezcla de ácidos grasos estandarizada como referencia (Supelco 37 Component FAME Mix by Sigma-Aldrich©).

### 5.5.1 Predicción de las propiedades del biodiesel a partir de los perfiles de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P.

Se calcularon varios parámetros de calidad del biodiesel a partir del perfil de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P, tomando en consideración el tamaño de la cadena hidrocarbonada, la presencia de dobles enlaces y el contenido de cada ácido graso identificado. El índice de saponificación (SV, por sus siglas en inglés), índice de yodo (IV, por sus siglas en inglés) y número de cetano (CN) se estimaron a partir del perfil de ácidos grasos haciendo uso de las siguientes fórmulas (1-3) (Krisnangkura, 1986):

$$SV = \sum (560 \times N_i) / M_i \quad (1)$$

$$IV = \sum (254 \times DN_i) / M_i \quad (2)$$

Donde  $N_i$  es el contenido de cada ácido graso en porcentaje,  $D$  es el número de dobles enlaces y  $M_i$  es el peso molecular de cada ácido graso. Posteriormente, el número de cetano (CN) fue calculado utilizando la siguiente fórmula:

$$CN = 46.3 + (5.458 / SV) - (0.225 \times IV) \quad (3)$$

Además, se estimaron otras propiedades del biodiesel tales como el grado de insaturación (DU, por sus siglas en inglés), punto de turbidez (CP, por sus siglas en inglés) y estabilidad oxidativa (OS, por sus siglas en inglés). El grado de insaturación (DU), se calculó en función del contenido de ácidos grasos monoinsaturados (MUFA) y poliinsaturados (PUFA). El CP fue calculado utilizando la ecuación propuesta por Sarin *et al.* (2009), mientras que la OS fue calculada por la ecuación propuesta por Wang *et al.* (2012).

$$DU = \sum MUFA + (2 \times PUFA) \quad (4)$$

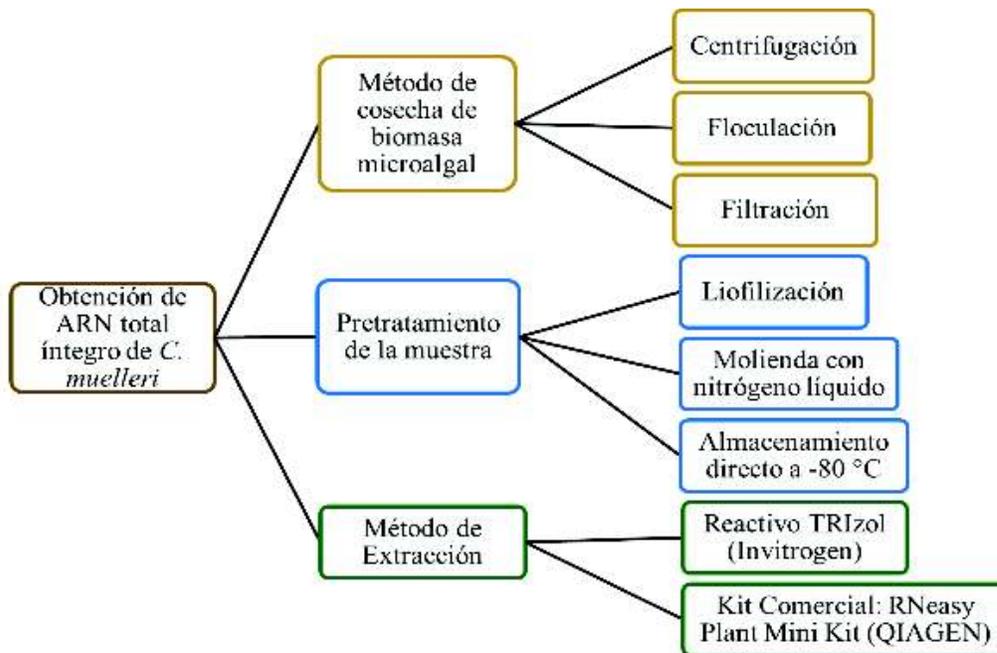
$$CP = (0.526 \times C_{16}) - 4.992 \quad (5)$$

$$OS = -0.0384 \times DU + 7.770 \quad (6)$$

## 5.6 Análisis transcriptómico de *C. muelleri* cultivada bajo diferentes concentraciones de P

### 5.6.1 Extracción de ARN total

Para llevar a cabo la extracción de ARN total de *C. muelleri* se llevó a cabo el siguiente proceso de estandarización (Figura 4) en el que se evaluaron diferentes métodos de cosecha, pretratamiento de la muestra, así como método de extracción de ARN total:



**Figura 4.** Proceso de estandarización llevado a cabo para la obtención de ARN total de *C. muelleri* cultivada bajo diferentes concentraciones de P.

La estrategia por la cual se obtuvieron mejores resultados en la extracción de ARN total de *C. muelleri* fue mediante la cosecha de la biomasa microalgal a través de centrifugación (condiciones: 2500 g por 10 minutos a 4 °C) para posteriormente almacenar la muestra directamente a – 80 °C hasta su uso. Para la extracción de ARN total, las muestras permanecieron en nitrógeno líquido hasta que se vertieron dentro de un tubo Eppendorf de 1.5

mL con 1 mL de TRIzol® Reagent (Thermo Fischer Scientific) frío, siguiéndose las instrucciones del fabricante: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/trizol\\_reagent.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/trizol_reagent.pdf)

### **5.6.2 Cuantificación e integridad del ARN total**

Para la cuantificación del ARN total se tomaron 2 µL de cada muestra y se cuantificó en el espectrofotómetro NanoDrop 1000 (Thermo Scientific). Se consideró que la relación de absorbancia 260/280 nm fuera entre el rango >1.8 a 2. La integridad del ARN total se analizó por medio de electroforesis en un gel de agarosa al 1 % preparándose de la siguiente manera:

- ✓ Se pesó 0.5 g de agarosa y se depositó en un matraz Erlen-Meyer de 250 mL con 50 mL de buffer TAE 1X estéril.
- ✓ Posteriormente se calentó la mezcla de agarosa con TAE 1X haciendo uso de un horno microondas por un minuto y medio, verificando su disolución cada 30 seg. Ya que la mezcla no presentó grumos y se tornó transparente, se dejó reposar a temperatura ambiente hasta que la temperatura del matraz fuera soportable al dorso de la mano.
- ✓ Se añadió 1 µL de GelRed™ (Biotium) y se agitó vigorosamente para después vaciarse lentamente en la placa de electroforesis previamente preparada con los sellos acrílicos, esto para evitar la formación de burbujas. Se colocaron los peines de separación correspondiente y se dejó reposar hasta que la mezcla se solidificara (15 a 20 min).
- ✓ Una vez formado el gel, se depositó en el tanque de la cámara de electroforesis. Se vertió la cantidad apropiada de buffer TAE 1X estéril hasta que cubriera el gel y se retiraron los peines con cuidado, evitando que se dañara el gel (Para este punto, la cámara de electroforesis debe estar en total asepsia, siendo previamente lavada con Benzal al 4 %).
- ✓ Se cargaron 2 µL de muestra de ARN total mezclado con 2 µL de buffer de carga y se depositaron en cada uno de los pozos.
- ✓ Se llevó a cabo la electroforesis conectando la cámara de electroforesis a una fuente de poder a 80 Volts por 35 min aproximadamente.

✓ Una vez terminado el proceso, se continuó con la visualización del gel de electroforesis en el fotodocumentador Bio-Imaging System MiniBis pro (Bio America Inc), guardándose la imagen para su análisis posterior.

#### 5.6.2.1 Purificación de ARN extraído de *C. muelleri* cultivada bajo diferentes concentraciones de P

Con el fin de eliminar posibles impurezas remanentes del proceso de extracción de ARN total de *C. muelleri*, se procedió con su purificación mediante el uso del kit RNeasy MinElute Cleanup Kit de Qiagen, siguiéndose las instrucciones del fabricante: <https://www.qiagen.com/mx/resources/resourcedetail?id=bb039a0a-b168-478f-88a7-ebcada67f7dd&lang=en>. Las muestras de ARN se volvieron a analizar por NanoDrop 1000 (Thermo Scientific), se corrieron en gel de agarosa para después visualizarse en el fotodocumentador y finalmente almacenarse a  $-80\text{ }^{\circ}\text{C}$  hasta su uso. Se obtuvieron dos réplicas biológicas (cada una compuesta por tres réplicas técnicas) de ARN total por cada tratamiento, teniendo un total de 8 muestras, las cuales fueron enviadas al Laboratorio de Servicios Genómicos de LANGEBIO-CINVESTAV Unidad Irapuato, Guanajuato, México.

#### 5.6.3 Construcción de librerías de ADNc y secuenciación

La construcción y secuenciación de las genotecas fue realizada en la Unidad de Genómica Avanzada del Laboratorio de Genómica para la Biodiversidad del Centro de Investigaciones Avanzadas del Instituto Politécnico Nacional (LANGEBIO-CINVESTAV) en la Unidad de Irapuato, Guanajuato. Para cada muestra, se determinó el RIN, que por sus siglas en inglés significa número de integridad del ARN en un 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Las genotecas fueron preparadas usando  $3\text{ }\mu\text{g}$  de ARN total (a concentración de al menos  $50\text{ ng mL}^{-1}$ ), un  $\text{RIN} \geq 6$  y TruSeq RNA Sample Prep v2 LT de Illumina; la secuenciación de las 8 genotecas (dos réplicas biológicas por tratamiento) se realizó en modo paired-end (2 x 150 bp) MID en la plataforma NextSeq® de Illumina (San Diego, CA, USA)

#### 5.6.4 Filtrado de calidad secuencias, ensamblado *de novo* del transcriptoma y anotación funcional.

Posterior a la secuenciación, se recibió un enlace para descargar toda la información de las lecturas crudas generadas por la plataforma Illumina NextSeq en archivos con formato FastQ, los cuales fueron analizadas usando FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) (Andrews, 2010). Posteriormente, con la finalidad de obtener lecturas fiables para el ensamblado y análisis del transcriptoma, se utilizaron las herramientas Trimmomatic v0.36 (Bolger *et al.*, 2014) la cual permite la eliminación de adaptadores, lecturas de baja calidad ( $Q < 25$ ) y lecturas ambiguas (“N”) y SortMeRNA (Kopylova *et al.*, 2013) la cual permite la detección y eliminación de secuencias de ARN ribosomal. Para corroborar que las lecturas conservadas fueron de alta calidad, se realizó un segundo análisis utilizando el programa FastQC. Una vez que se confirmó la calidad de las lecturas se continuó con el ensamble del transcriptoma de *C. muelleri*.

En vista de que hasta el momento no se cuenta con el genoma de *C. muelleri* secuenciado, para el ensamblado *de novo* del transcriptoma se probaron tres softwares: Trinity en su versión default y haciendo uso de la opción ‘jaccard clip’ (Grabherr *et al.*, 2011), rnaSPAdes (Bushmanova *et al.*, 2019) y SOAPdenovo (Luo *et al.*, 2012). Para definir la calidad del ensamblado, se utilizó la herramienta rnaQuast (Bushmanova *et al.*, 2016), la cual hace sinergia de otras herramientas como BUSCO (Simão *et al.*, 2015) y GeneMarkS-T (Tang *et al.*, 2015) para proveer información estadística como el total de los unigenes, su distribución, longitud y estadístico N50, que permitirán la selección de la versión del transcriptoma más adecuada para los análisis posteriores. Además, se utilizó la herramienta HISAT2 (Kim *et al.*, 2015) para verificar la cobertura promedio de las lecturas con el transcriptoma ensamblado.

La anotación funcional del transcriptoma se realizó siguiendo el protocolo Trinotate (<https://github.com/Trinotate/Trinotate.github.io/wiki/TrinotateWeb>) (Bryant *et al.*, 2017). Trinotate utiliza varios métodos diferentes y bien referenciados para la anotación funcional, incluida la búsqueda de homología con datos de secuencia conocidos como Swiss-Prot utilizando BLASTx y BLASTp (valor de  $E < 1e^{-3}$ ), identificación de dominios de proteínas (HMMER / PFAM), péptido señal de proteínas y predicción de dominios transmembranales (signalP / tmHMM) así como también usa la información de otras bases de datos como eggNOG

(Powell *et al.*, 2011) y GO (Ashburner *et al.*, 2000) y KEGG (Kanehisa *et al.*, 2012). Todos los datos de anotación funcional derivados de este análisis fueron integrados en una base de datos SQLite permitiendo la búsqueda rápida y eficiente de términos con cualidades específicas para finalmente crear un reporte de anotación completo para un transcriptoma. Por otro lado, con la finalidad de robustecer el análisis, se realizó un BLASTx de las secuencias con proteomas de otras microalgas (*Phaeodactylum tricornutum*, *Thalassiosira pseudonana* y *Chlamydomonas reinhardtii*) depositados en la base de datos de Swiss-Prot. Después de seguir el protocolo de Trinotate y recopilar toda la información depositada en diferentes fuentes externas y bases de datos, todos los resultados obtenidos se tabularon en un archivo Excel.

### **5.6.5 Análisis de expresión diferencial y análisis de enriquecimiento**

La estimación de abundancia de transcritos fue determinada utilizando la herramienta RSEM, la cual a su vez utiliza Bowtie2 para llevar a cabo el alineamiento utilizando los parámetros por defecto; esto es que se permitió como máximo 2 bases mal apareadas en la secuencia semilla a la cual se le asignó un tamaño de 25 bases, se determinó un tamaño de inserto entre lecturas pareadas de 1 -1000 (ambos valores por defecto en RSEM) y para el cálculo de la abundancia de transcritos se reportaron todos los alineamientos encontrados para cada fragmento, excluyendo todos aquellos fragmentos con más de 200 alineamientos (todos estos parámetros son por defecto de RSEM) (Li y Dewey, 2011; Langmead, 2010). Se utilizaron los conteos crudos de cada transcrito de cada librería para generar una matriz con todas las condiciones experimentales y sus réplicas. Los conteos crudos entre las muestras y las réplicas fueron normalizados por la media ajustada de los valores M (TMM) usando el paquete estadístico edgeR de Bioconductor (Robinson *et al.*, 2010).

La correlación y el coeficiente de variación biológica entre las muestras y las réplicas biológicas se determinaron con un análisis de correlación y de escalamiento multidimensional (MDS) de Pearson, utilizando los datos de expresión normalizados por TMM. Para evaluar la respuesta transcripcional de *C. muelleri* a las diferentes concentraciones de P, se compararon los genes de *C. muelleri* cultivada bajo una concentración de P de 72  $\mu\text{M}$  (control) frente a los tratamientos en los que se cultivó esta diatomea bajo concentraciones limitantes (7 y 18  $\mu\text{M}$ ) y

de exceso (144  $\mu\text{M}$ ) de P, utilizando el paquete edgeR (Robinson *et al.* 2010), utilizando un método estadístico basado en comparaciones por pares y una dispersión común calculada como sugiere el manual de EdgeR. Los transcritos que mostraron una tasa de descubrimiento de falsos positivos (FDR, por sus siglas en inglés) de 0.01 y un valor de cambio ( $\log_2$  fold-change)  $\log_2\text{FC} \geq \pm 1$  fueron considerados como genes expresados diferencialmente. El agrupamiento jerárquico de los genes expresados diferencialmente en *C. muelleri* se mostró en un heatmap que se realizó usando las funciones ‘hclust’ y ‘heatmap.2’ con el método de agrupamiento de ward.D y la biblioteca de ‘gplots’ de R (<http://cran.r-project.org/web/packages/gplots/index.html>).

Para el conjunto de transcritos expresados diferencialmente a lo largo de las comparaciones evaluadas se realizó un análisis de enriquecimiento funcional utilizando el paquete de R, GOSEQ v1.26.0 (FDR < 0.05) (Young *et al.* 2010). En este análisis se utilizaron las categorías de Ontología Génica (GO, por sus siglas en inglés) obtenidas en las anotaciones funcionales llevadas a cabo por Trinotate para caracterizar las respuestas transcripcionales considerando los procesos biológicos, componentes celulares y funciones moleculares asociados a los genes.

### **5.7 Análisis *in silico* de genes involucrados en el transporte de P y metabolismo de lípidos en *C. muelleri*.**

Se seleccionaron y extrajeron del transcriptoma de *C. muelleri*, genes codificantes para proteínas transportadoras de P (NPT2B y VTC4) y genes involucrados en el metabolismo de lípidos (ACC y FAD2) para el análisis *in silico* de sus secuencias aminoacídicas.

Se utilizó la herramienta ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) para traducir las secuencias nucleotídicas de *C. muelleri*, encontrar marcos de lectura abiertos (ORF’s por sus siglas en inglés) y seleccionar el de mayor tamaño para después realizar un Blastp (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) y comparar dichas secuencias con otras depositadas en la base de datos del NCBI.

Por otro lado, se identificaron los dominios y motivos conservados más probables empleando las herramientas Conserved Domain Search y PROSITE de ExPASy (<https://prosite.expasy.org/>) - NIH de NCBI (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Además, se utilizó el software Seq2Logo 2.0 (Thomsen y Nielsen, 2012) de DTU Health Tech para la elaboración de logos de secuencia, los cuales son una representación gráfica del contenido de información almacenado en un alineamiento múltiple de secuencias (MSA, por sus siglas en inglés) y proporcionan una representación compacta y altamente intuitiva de la composición de aminoácidos específica en secuencias biológicas (Thomsen y Nielsen, 2012). Por otra parte, se realizaron alineamientos múltiples con secuencias aminoacídicas homólogas pertenecientes a otras especies de microalgas y plantas a través del software Jalview ver. 2.11.1.0 (Waterhouse *et al.*, 2009) donde fueron visualizados y editados.

Asimismo, se determinó el peso molecular (kDa) y el punto isoeléctrico de las secuencias aminoacídicas mediante el servidor ProtParam de ExPASy (<https://web.expasy.org/protparam/>). Además, se empleó el servidor NPS@ ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_seccons.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html)) para efectuar la predicción de la estructura secundaria y determinar los porcentajes de alfa hélices y láminas betas. Finalmente, se empleó el software SWISS MODEL para buscar modelos viables para elucidar la estructura de las proteínas codificadas por las secuencias identificadas en *C. muelleri*, a través de homología a partir de la utilización de modelos tridimensionales reales de proteínas similares.

## VI. RESULTADOS Y DISCUSIÓN

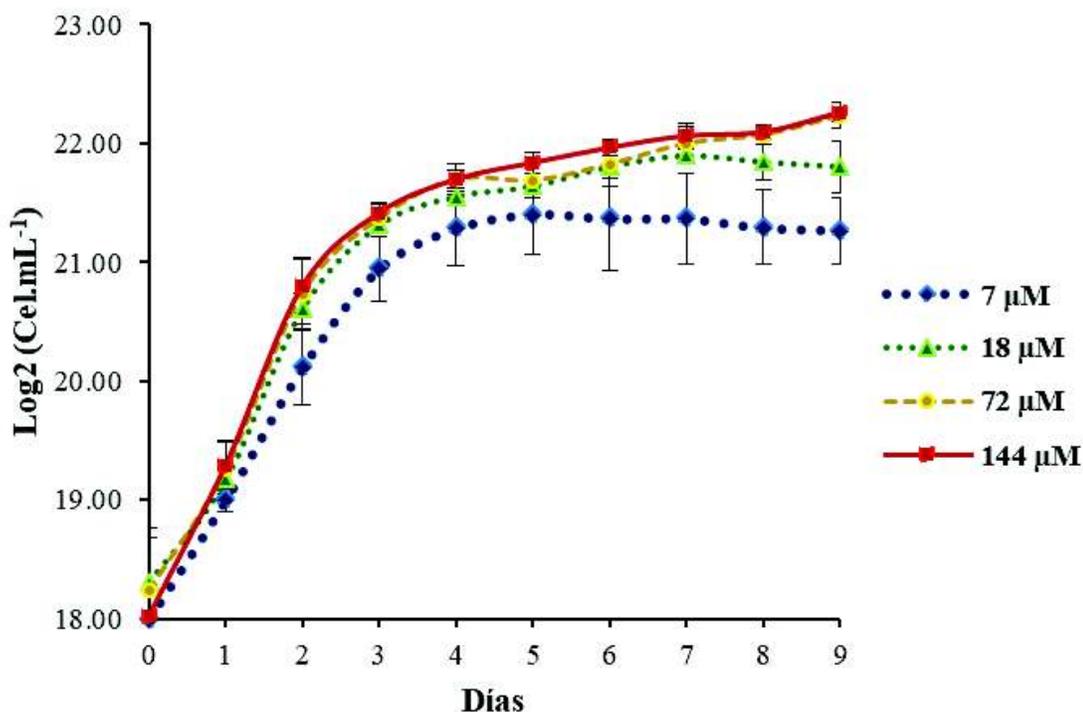
### 6.1 Crecimiento y determinación de la concentración celular de *C. muelleri* bajo diferentes concentraciones de P.

La Figura 5 muestra el crecimiento de *C. muelleri* hasta los nueve días de cultivo y se puede observar en términos generales que el crecimiento de los cultivos fue ascendente hasta el día cuatro en donde se alcanzó la fase logarítmica tardía, observándose una menor densidad celular en el tratamiento cuya concentración de P fue de 7  $\mu\text{M}$  la cual fue de  $21.28 \pm 0.30$  log base 2 ( $2.60 \times 10^6$  células  $\text{mL}^{-1}$ ) en comparación a los tratamientos control y exceso los cuales presentaron densidades celulares de  $21.69 \pm 0.07$  log base 2 ( $3.40 \times 10^6$  células  $\text{mL}^{-1}$ ) y  $21.70 \pm 0.12$  log base 2 ( $3.42 \times 10^6$  células  $\text{mL}^{-1}$ ) no encontrándose diferencias significativas entre ambos durante todo el periodo de cultivo ( $p < 0.05$ ). Por su parte, el tratamiento cuya concentración de P fue de 18  $\mu\text{M}$  presentó una densidad celular de  $21.55 \pm 0.04$  log base 2 ( $2.94 \times 10^6$  células  $\text{mL}^{-1}$ ), no encontrándose diferencias significativas con el resto de los tratamientos ( $p < 0.05$ ).

Al final de los nueve días de cultivo, se mantuvo una menor densidad celular en el tratamiento cuya concentración de fósforo fue de 7  $\mu\text{M}$  la cual fue de  $21.27 \pm 0.27$  log base 2 ( $2.56 \times 10^6$  células  $\text{mL}^{-1}$ ) en comparación con los tratamientos control y exceso donde se observaron densidades celulares de  $22.23 \pm 0.11$  log base 2 ( $4.94 \times 10^6$  células  $\text{mL}^{-1}$ ) y  $22.26 \pm 0.08$  log base 2 ( $5.04 \times 10^6$  células  $\text{mL}^{-1}$ ) respectivamente ( $p < 0.05$ ). Por su parte, el tratamiento cuya concentración fue de 18  $\mu\text{M}$  presentó una densidad celular de  $21.81 \pm 0.22$  log base 2 ( $3.52 \times 10^6$  células  $\text{mL}^{-1}$ ) ( $p < 0.05$ ). Estos resultados corroboran el rol importante del P como un nutriente esencial para el crecimiento y división celular en las microalgas, por lo que bajas concentraciones y el tiempo de duración del cultivo influirán en la densidad celular y en las tasas de crecimiento (Brembu *et al.*, 2017; Dyrman, 2016).

Por otra parte, el valor más alto de tasa máxima de crecimiento fue observado en el tratamiento cuya concentración de P fue de 144  $\mu\text{M}$  ( $1.51 \pm 0.21$ ) y el menor en el tratamiento cuya concentración de P fue 7  $\mu\text{M}$  ( $1.11 \pm 0.32$ ); las tasas de crecimiento media y acumulada fueron independientes de las concentraciones de P en el medio de cultivo al no encontrarse

diferencias significativas entre los tratamientos, aunque se mostró una tendencia a aumentar a medida que la concentración de P aumentaba (Tabla 1). Estos resultados pueden ser explicados por lo propuesto por Finenko y Krupatkina-Akinina (1974) en donde sugieren que tanto las concentraciones presentes en el medio de cultivo como las concentraciones intracelulares de P influyen en las tasas de crecimiento, sin embargo, para ser concluyentes se requerirán estudios posteriores.



**Figura 5.** Crecimiento de *C. muelleri* durante 9 días en cuatro tratamientos experimentales con diferentes concentraciones de P: 7  $\mu\text{M}$ , 18  $\mu\text{M}$ , 72  $\mu\text{M}$  y 144  $\mu\text{M}$ . Las barras de error representan la desviación estándar de los conteos celulares de cuatro réplicas biológicas compuestas por tres réplicas técnicas cada una por tratamiento.

**Tabla 1.** Densidades celulares y tasas de crecimiento de *C. muelleri* cultivada bajo diferentes concentraciones de P.

Concentración de P	Densidad celular máxima (x 10 <sup>6</sup> mL <sup>-1</sup> )	Tasa máxima (μ) div/día	Tasa acumulada (μ)	Tasa promedio(μ) div/día
7 μM	2.56 ± 0.46 <sup>a</sup>	1.11 ± 0.32 <sup>a</sup>	3.26 ± 0.50 <sup>a</sup>	0.36 ± 0.05 <sup>a</sup>
18 μM	3.52 ± 0.54 <sup>b</sup>	1.43 ± 0.01 <sup>b</sup>	3.49 ± 0.26 <sup>a</sup>	0.39 ± 0.03 <sup>a</sup>
72 μM	4.94 ± 0.36 <sup>c</sup>	1.48 ± 0.15 <sup>b</sup>	4.00 ± 0.34 <sup>a</sup>	0.44 ± 0.04 <sup>a</sup>
144 μM	5.04 ± 0.27 <sup>c</sup>	1.51 ± 0.21 <sup>b</sup>	4.24 ± 0.80 <sup>a</sup>	0.47 ± 0.09 <sup>a</sup>

Los valores son la media ± desviación estándar de cuatro réplicas biológicas y cada una está compuesta por tres réplicas técnicas. Los valores en cada columna con diferente superíndice (a-c) indican una diferencia significativa (p<0.05).

## 6.2 Biomasa y análisis bioquímicos de *C. muelleri* cultivada bajo diferentes concentraciones de P.

La producción más alta de biomasa fue detectada en el tratamiento control (72 μM) con 311.42 ± 12.3 mg L<sup>-1</sup>; no se encontraron diferencias entre los tratamientos cuya concentración fue de 18 μM (250.26 ± 11.2 mg L<sup>-1</sup>) y 144 μM (248.07 ± 11.2 mg L<sup>-1</sup>). La producción más baja de biomasa se observó en el tratamiento cuya concentración de P fue de 7 μM (199.92 ± 2.7 mg L<sup>-1</sup>) (Tabla 2). Dentro de los efectos inmediatos observados por limitación en P en microalgas es la disminución de la síntesis y regeneración de sustratos del ciclo Calvin-Benson y como consecuencia una disminución en el uso de la energía lumínica para la fijación de carbono resultando en una producción de biomasa y tasas de crecimiento bajas (Belotti *et al.*, 2013; Dyrhman, 2016). Por su parte en el exceso de P se observó una menor producción de biomasa seca con respecto al control, esto puede explicarse al observarse una menor cantidad de cenizas en el exceso de P con respecto al control; se ha reportado para diatomeas que del total de su biomasa seca, la mayor porción la constituyen las cenizas lo que sugiere que a pesar de no encontrarse diferencias en el contenido de materia orgánica ni en la densidad celular entre control y exceso de P, el tamaño de las células fue menor en el tratamiento de exceso de P con respecto al control (Lee *et al.*, 2009; Cordero-Esquivel *et al.*, 1993).

El contenido más alto de clorofila *a* fue observado en el tratamiento control (0.5064 ± 0.02 mg L<sup>-1</sup>), mientras que menores cantidades fueron encontradas en los tratamientos de exceso (144 μM) y limitantes (18 y 7 μM) de P, encontrándose diferencias significativas entre todos los tratamientos (p < 0.05) (Tabla 2). Los niveles bajos de clorofila *a* no pueden ser atribuidos

directamente a la limitación de P ya que dentro de la composición de su molécula este elemento se encuentra ausente. Sin embargo, Roopnarain *et al.* (2014) sugieren que la disminución en el contenido de clorofila *a* ocurre debido a la incapacidad de las células de producir ATP y NADPH, las cuales son moléculas involucradas en su síntesis, afectando por ende todo el metabolismo y como una estrategia de la célula para darle prioridad a la síntesis de compuestos de reserva. Por otra parte, en *C. muelleri* cultivada bajo una concentración de P de 144  $\mu\text{M}$ , se observó que los niveles de clorofila *a* disminuyeron; si bien el efecto de concentraciones excesivas de P en los niveles de clorofila *a* en diatomeas es escaso, en plantas como frijol y soja se ha observado este patrón y dichos estudios sugieren que esto se debe a un antagonismo con los iones  $\text{NO}_3^-$  ya que tanto la capacidad fotosintética y en consecuencia los niveles de clorofila están directamente relacionados con el contenido de N y P en las plantas (Novoa *et al.*, 2018; Quintana-Blanco *et al.*, 2017; Singh y Reddy, 2015).

**Tabla 2. Producción de biomasa y composición bioquímica de *Chaetoceros muelleri* cultivada bajo diferentes concentraciones de P.**

Concentración de P	Biomasa seca (mg L <sup>-1</sup> )	Materia orgánica (mg L <sup>-1</sup> )	Cenizas (mg L <sup>-1</sup> )	Clorofila <i>a</i> ( $\mu\text{g mL}^{-1}$ )	Lípidos totales (%)	Proteínas totales (%)
7 $\mu\text{M}$	199.92 $\pm$ 2.7 <sup>a</sup>	91.78 $\pm$ 18.07 <sup>a</sup>	107.80 $\pm$ 14.8 <sup>a</sup>	0.1078 $\pm$ 0.02 <sup>a</sup>	4.42 $\pm$ 0.44 <sup>a</sup>	4.29 $\pm$ 0.50 <sup>a</sup>
18 $\mu\text{M}$	250.26 $\pm$ 11.2 <sup>b</sup>	113.28 $\pm$ 12.79 <sup>b</sup>	115.03 $\pm$ 17 <sup>ab</sup>	0.1759 $\pm$ 0.01 <sup>b</sup>	3.73 $\pm$ 0.64 <sup>b</sup>	3.80 $\pm$ 0.52 <sup>b</sup>
72 $\mu\text{M}$	311.42 $\pm$ 12.3 <sup>c</sup>	137.92 $\pm$ 12.08 <sup>c</sup>	158.60 $\pm$ 11 <sup>c</sup>	0.5064 $\pm$ 0.02 <sup>c</sup>	3.27 $\pm$ 0.56 <sup>c</sup>	3.81 $\pm$ 0.57 <sup>ab</sup>
144 $\mu\text{M}$	248.07 $\pm$ 11.7 <sup>b</sup>	133.95 $\pm$ 18.43 <sup>c</sup>	123.51 $\pm$ 11 <sup>b</sup>	0.4375 $\pm$ 0.03 <sup>d</sup>	3.18 $\pm$ 0.46 <sup>c</sup>	4.22 $\pm$ 0.63 <sup>ab</sup>

Los valores son la media  $\pm$  desviación estándar de cuatro réplicas biológicas y cada una está compuesta por tres réplicas técnicas. Los valores en cada columna con diferente superíndice (a-d) indican una diferencia significativa ( $p < 0.05$ ).

El contenido más alto de lípidos totales fue registrado en el tratamiento cuya concentración de P fue de 7  $\mu\text{M}$  y se observó además que su contenido se veía reducido a medida que aumentaba la concentración de P (Tabla 2). Se ha reportado ampliamente en varias especies de microalgas que, bajo una deficiencia de P, una respuesta común es la acumulación de lípidos neutrales y una reducción de los niveles de fosfolípidos, siendo estos últimos reemplazados por lípidos cuya composición carece de P (Van Mooy *et al.*, 2009; Cruz de Carvalho *et al.*, 2016). Recientemente, Lin *et al.*, (2018) evaluaron la privación de N, Si y P en el medio de cultivo en *C. muelleri*, reportando que bajo la privación de P, el contenido de lípidos totales se incrementó (26.74  $\pm$  1.00 % de su peso seco) en comparación a las condiciones establecidas como control

( $22.23 \pm 1.44\%$  de su peso seco; concentración de P de  $36 \mu\text{M}$ ); esta respuesta fue observada también en este estudio. Los valores reportados por Lin *et al.*, (2018) son más altos comparados a los de este estudio, lo cual pudiera ser atribuido a la cepa, a una menor intensidad de luz empleada ( $150 \mu\text{mol m}^{-2} \text{s}^{-1}$ ) y el tiempo de duración del experimento ya que los autores cosecharon las muestras para los diferentes análisis bioquímicos en el día seis de cultivo.

Por otra parte, se encontraron diferencias significativas en el contenido total de proteínas entre los tratamientos cuya concentración de P fueron  $7$  y  $18 \mu\text{M}$ ; sin embargo, entre los tratamientos limitantes de P y los tratamientos control ( $72 \mu\text{M}$ ) y exceso ( $144 \mu\text{M}$ ) de P, no se encontraron diferencias significativas (Tabla 2). Por su parte, Berdalet *et al.* (1994) observaron el efecto de la deficiencia de P y nitrógeno en el contenido total de proteínas y ácidos nucleicos de *Heterocapsa* sp., reportando que bajo deficiencia de P la síntesis de proteínas permanecerá constante siempre y cuando en el medio de cultivo exista niveles suficientes de nitrógeno, al menos en un periodo inicial de cultivo. Liang *et al.* (2013) observaron también en *Chlorella* sp. cultivada bajo limitación de P que el contenido total de proteínas no se vio modificado en comparación con condiciones definidas de suficiencia de nutrientes y limitación por nitrógeno, siendo en esta última donde se apreció una disminución del contenido total de proteínas; además los autores hacen hincapié en que la respuesta a la limitación de diferentes nutrientes dependerá en función de las especies estudiadas, coincidiendo así con otros estudios (Zhao *et al.*, 2009; Berdalet *et al.* 1994).

De acuerdo con lo anterior, Velasco *et al.* (2016) reportaron un contenido total de proteínas de  $59 \pm 0.9\%$  y un contenido total de lípidos de  $31 \pm 1.3\%$  (basándose en el contenido de materia orgánica) para una cepa de *C. muelleri* nativa del mar Caribe; las condiciones en las que el crecimiento de esta cepa se llevó a cabo fue en medio F/2 bajo una temperatura de  $24 \pm 0.5 \text{ }^\circ\text{C}$ , salinidad de 35 ups y suministro de aire e iluminación constante de  $45 \mu\text{mol m}^{-2} \text{s}^{-1}$ . Estas condiciones difieren con las establecidas en este estudio, ya que Velasco *et al.*, (2016) emplearon una cepa aislada del mar Caribe, la salinidad fue mayor, la intensidad de luz fue menor y además se suministró  $\text{CO}_2$  de manera constante y se ha reportado que aunado a la limitación de nutrientes, estos factores pueden influir en la composición bioquímica en general y en especial en la acumulación de lípidos (Rahmadi *et al.*, 2020; Jensen *et al.*, 2019; Zulu *et al.*, 2018). Por otra parte, los estudios relacionados a evaluar el contenido total de proteínas en

microalgas y en específico en diatomeas y bajo condiciones de exceso de P son escasos, sin embargo, se sugiere que, a mayor disponibilidad de P, la síntesis de ATP se verá favorecida y por ende la síntesis de proteínas, no obstante, es necesario realizar estudios posteriores.

### **6.3 Perfil de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de fósforo**

Se identificaron 14 ácidos grasos en el perfil de *C. muelleri* cultivada bajo diferentes concentraciones de P siendo los más abundantes palmitoleico (34.61 – 39.15 %), palmítico (18.62 – 20.62 %), eicosapentaenoico (9.46 – 20.36 %) y mirístico (13.31 – 15.11 %) (Tabla 3). Además, se encontraron cantidades traza del ácido graso miristoleico únicamente en el tratamiento cuya concentración de P fue de 144  $\mu$ M. El ácido graso behénico sólo fue detectado en los tratamientos control (72  $\mu$ M) y exceso (144  $\mu$ M) de P, mientras que el ácido lignocérico no se detectó en el tratamiento control (Figura 6). Estos resultados coinciden con lo reportado en la literatura en que un rasgo característico de las diatomeas es que dentro de su perfil de ácidos grasos, los ácidos mirístico, palmítico, palmitoleico y eicosapentaenoico son predominantes mientras que los ácidos grasos de 18 átomos de carbono generalmente se encuentran en cantidades traza; por el contrario, especies de microalgas pertenecientes al grupo de las clorofitas y cianofitas los ácidos grasos más abundantes son aquellos cuya cadenas hidrocarbonadas oscilan entre los 16 y 18 átomos de carbono (Vuppaladadiyam *et al.*, 2018; Zulu *et al.*, 2018).

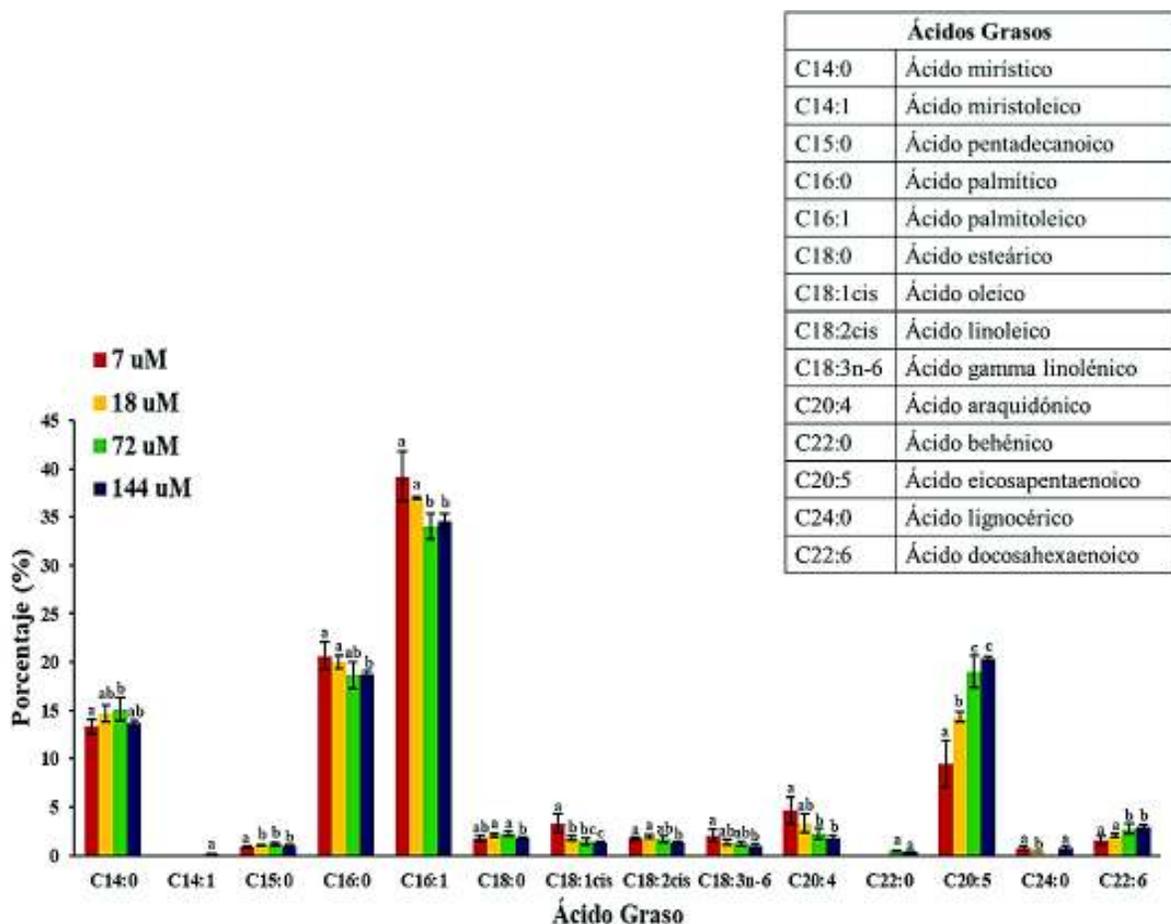
En este estudio se observó en los tratamientos con limitación de P (7 y 18  $\mu$ M) un incremento significativo en el contenido del ácido graso palmitoleico en comparación de los tratamientos control (72  $\mu$ M) y exceso (144  $\mu$ M) de P; respecto al contenido de ácido palmítico, se encontraron diferencias significativas entre los tratamientos limitantes y el exceso de P, siendo en los primeros donde se encontró la mayor cantidad de este ácido graso. Por otra parte, el contenido del ácido eicosapentaenoico se vio afectado por la limitación de P, encontrándose la menor cantidad en el tratamiento cuya concentración fue de 7  $\mu$ M ( $p < 0.05$ ). Por último, el contenido más alto del ácido mirístico se observó en el tratamiento control (72  $\mu$ M) y la menor cantidad en el tratamiento con menor concentración de P (7  $\mu$ M); no se encontraron diferencias

significativas entre los tratamientos cuyas concentraciones de P fueron de 18 y 144  $\mu\text{M}$  para este ácido graso.

**Tabla 3. Perfil de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P.**

Ácido Graso	7 $\mu\text{M}$	18 $\mu\text{M}$	72 $\mu\text{M}$	144 $\mu\text{M}$
Ácido mirístico	13.31 $\pm$ 0.72 <sup>a</sup>	14.67 $\pm$ 0.88 <sup>ab</sup>	15.11 $\pm$ 1.19 <sup>b</sup>	13.74 $\pm$ 0.17 <sup>ab</sup>
Ácido miristoleico	N.D.	N.D.	N.D.	0.14 $\pm$ 0.03 <sup>a</sup>
Ácido pentadecanoico	0.83 $\pm$ 0.10 <sup>a</sup>	1.09 $\pm$ 0.06 <sup>b</sup>	1.17 $\pm$ 0.18 <sup>b</sup>	1.03 $\pm$ 0.04 <sup>b</sup>
Ácido palmítico	20.62 $\pm$ 1.41 <sup>a</sup>	19.97 $\pm$ 0.70 <sup>a</sup>	18.62 $\pm$ 1.41 <sup>ab</sup>	18.77 $\pm$ 0.30 <sup>b</sup>
Ácido palmitoleico	39.15 $\pm$ 2.58 <sup>a</sup>	36.96 $\pm$ 0.13 <sup>a</sup>	34.04 $\pm$ 1.34 <sup>b</sup>	34.61 $\pm$ 0.70 <sup>b</sup>
Ácido estearico	1.73 $\pm$ 0.34 <sup>ab</sup>	2.09 $\pm$ 0.19 <sup>a</sup>	2.21 $\pm$ 0.22 <sup>a</sup>	1.77 $\pm$ 0.08 <sup>b</sup>
Ácido oleico	3.29 $\pm$ 0.97 <sup>a</sup>	1.78 $\pm$ 0.24 <sup>b</sup>	1.45 $\pm$ 0.31 <sup>bc</sup>	1.29 $\pm$ 0.10 <sup>c</sup>
Ácido linoleico	1.78 $\pm$ 0.17 <sup>a</sup>	1.93 $\pm$ 0.17 <sup>a</sup>	1.67 $\pm$ 0.29 <sup>ab</sup>	1.35 $\pm$ 0.06 <sup>b</sup>
Ácido $\gamma$ -linolénico	2.05 $\pm$ 0.61 <sup>a</sup>	1.38 $\pm$ 0.26 <sup>ab</sup>	1.22 $\pm$ 0.24 <sup>ab</sup>	0.99 $\pm$ 0.22 <sup>b</sup>
Ácido araquidónico	4.69 $\pm$ 1.36 <sup>a</sup>	3.30 $\pm$ 0.95 <sup>ab</sup>	2.26 $\pm$ 0.55 <sup>b</sup>	1.76 $\pm$ 0.21 <sup>b</sup>
Ácido behénico	N.D.	N.D.	0.47 $\pm$ 0.04 <sup>a</sup>	0.40 $\pm$ 0.03 <sup>a</sup>
Ácido eicosapentaenoico	9.46 $\pm$ 2.37 <sup>a</sup>	14.31 $\pm$ 0.52 <sup>b</sup>	18.99 $\pm$ 1.66 <sup>c</sup>	20.36 $\pm$ 0.21 <sup>c</sup>
Ácido lignocérico	0.80 $\pm$ 0.17 <sup>a</sup>	0.43 $\pm$ 0.01 <sup>b</sup>	N.D.	0.77 $\pm$ 0.17 <sup>a</sup>
Ácido docosahexaenoico	1.57 $\pm$ 0.45 <sup>a</sup>	2.03 $\pm$ 0.16 <sup>a</sup>	2.73 $\pm$ 0.49 <sup>b</sup>	2.96 $\pm$ 0.19 <sup>b</sup>
% Total de AGS	37.30 $\pm$ 2.01 <sup>a</sup>	38.27 $\pm$ 1.81 <sup>a</sup>	37.59 $\pm$ 0.56 <sup>a</sup>	36.50 $\pm$ 0.34 <sup>a</sup>
% Total de AGMI	42.44 $\pm$ 2.24 <sup>a</sup>	38.75 $\pm$ 0.32 <sup>b</sup>	35.50 $\pm$ 1.18 <sup>c</sup>	36.05 $\pm$ 0.75 <sup>c</sup>
% Total de AGPI	19.56 $\pm$ 2.60 <sup>a</sup>	22.97 $\pm$ 1.49 <sup>a</sup>	26.89 $\pm$ 1.54 <sup>b</sup>	27.44 $\pm$ 0.46 <sup>b</sup>

Los valores son la media  $\pm$  desviación estándar de cuatro réplicas biológicas y cada una está compuesta por tres réplicas técnicas. Los datos se presentan como porcentaje de ácidos grasos. Los valores en cada fila con diferente superíndice (a-c) indican una diferencia significativa ( $p < 0.05$ ). AGS: Ácidos grasos saturados. AGI: Ácidos grasos insaturados. AGMI: Ácidos grasos monoinsaturados. AGPI: Ácidos grasos poliinsaturados. N.D: No detectado.



**Figura 6.** Perfil de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P. Los valores son la media  $\pm$  desviación estándar de cuatro réplicas biológicas y cada una está compuesta por tres réplicas técnicas. Los datos se presentan como porcentaje de ácidos grasos. Los valores en cada barra con diferente superíndice (a-c) indican una diferencia significativa ( $p < 0.05$ ).

Siron *et al.* (1989) sugieren que el incremento de los ácidos grasos palmítico y palmitoleico en la diatomea *Phaeodactylum tricornutum* bajo condiciones limitantes de P es una respuesta relacionada con la acumulación de TAG y disminución del contenido de fosfolípidos y que ambos ácidos grasos pueden representar una fuente de energía de reserva. Por el contrario, los niveles del ácido eicosapentaenoico se vieron reducidos drásticamente bajo deficiencia de P. Los autores sugieren que la limitación de P afecta el grado de insaturación de los lípidos microalgales principalmente la síntesis de los AGPI, los cuales presentan mejores condiciones para su síntesis mientras el cultivo se encuentra en su fase exponencial en donde la actividad celular es intensa y no se presenta una deficiencia de P. Por otro lado, Cañavate *et al.* (2017)

corroboraron este mismo patrón en *P. tricornutum* cultivándola en medio F/2 y privación de P observando que se duplicó el contenido del ácido graso palmitoleico y se redujo a la mitad el contenido del ácido eicosapentaenoico bajo privación de P comparado con el tratamiento control (100  $\mu$ M de P).

En relación a las especies del género *Chaetoceros*, han sido escasos los estudios que han evaluado en específico la limitación o privación de P y su efecto en el perfil de ácidos grasos; por su parte, Cañavate *et al.* (2017) evaluaron el efecto de la privación de P en el perfil de ácidos grasos en varias especies de microalgas, entre ellas *Chaetoceros gracilis*, en la cual se observó que el contenido del ácido palmítico aumentó considerablemente en comparación al contenido del ácido palmitoleico bajo privación de P y al igual que en otros estudios realizados en diatomeas, el contenido del ácido eicosapentaenoico se vio reducido. Lin *et al.* (2018) realizaron el primer estudio, hasta donde se conoce, de la privación de P en *C. muelleri* y su efecto en el perfil de ácidos grasos y compararon los resultados con otra diatomea, *Thalassiosira weissflogii*; los autores observaron en *C. muelleri* una respuesta similar a lo ya reportado en *C. gracilis*, sin embargo, en *T. weissflogii* el contenido del ácido palmitoleico aumentó considerablemente a diferencia del ácido palmítico ante la ausencia de P en el medio.

Por lo que se refiere al estudio del efecto del exceso de nutrientes en microalgas y en particular de P, este ha sido poco explorado y recientemente han comenzado a generarse estudios en diatomeas. Lopes *et al.* (2019) evaluaron tres concentraciones de P (76, 114 y 228  $\mu$ M) y dos fases de cultivo (exponencial y estacionaria) y su efecto en el perfil de ácidos grasos en *P. tricornutum* cultivada en medio F/2; al igual que en estudios anteriores en diatomeas, los ácidos mirístico, palmítico, palmitoleico y eicosapentaenoico comprometieron en mayor parte el contenido total de ácidos grasos (~ 80 %), sin embargo, no se detectaron diferencias significativas en el contenido de los ácidos grasos entre las diferentes concentraciones de P pero sí se encontraron diferencias significativas al contrastar la fase exponencial con la estacionaria, viéndose en esta última una disminución en el contenido del ácido eicosapentaenoico y un aumento de los ácidos grasos palmítico y palmitoleico.

De acuerdo a la saturación e insaturación global de los ácidos grasos en este estudio, el contenido total de ácidos grasos saturados no se vio modificado entre los tratamientos ( $p < 0.05$ ). En cambio, el contenido total de ácidos grasos monoinsaturados (AGMI) se incrementó a

medida que la concentración de P disminuía, siendo el tratamiento de 7  $\mu\text{M}$  donde se encontró un mayor contenido de AGMI, seguido del tratamiento cuya concentración fue 18  $\mu\text{M}$  encontrándose diferencias significativas entre ambos. Con respecto a los tratamientos cuya concentración de P fue 72 y 144  $\mu\text{M}$ , no presentaron diferencias significativas entre ellos en el contenido de AGMI totales, pero sí al compararse con los tratamientos con bajas concentraciones de P (7 y 18  $\mu\text{M}$ ) (Tabla 4). Del mismo modo, el contenido total de ácidos grasos poliinsaturados (AGPI) se vio afectado por la concentración de P, el cual aumentó a medida que las concentraciones de P aumentaban encontrándose los valores más altos en los tratamientos control (72  $\mu\text{M}$ ) y exceso (144  $\mu\text{M}$ ) sin encontrarse diferencias significativas entre ellos, pero si con los tratamientos con limitación de P. El menor contenido de AGPI se registró en los tratamientos limitantes de P (7 y 18  $\mu\text{M}$ ), no encontrándose diferencias significativas entre ambos tratamientos (Tabla 4).

Los resultados anteriores presentan un patrón similar a lo reportado para otras especies de diatomeas y microalgas en las que se evaluó la limitación por P, observándose que el incremento en el contenido de AGMI se ve favorecido principalmente ya que este tipo de ácidos grasos junto con los AGS se encuentran relacionados con la acumulación de TAG como compuestos de reserva (Anne-Marie *et al.*, 2020; Lopes *et al.*, 2019; Cañavate *et al.*, 2017). Algunos estudios muestran un mayor contenido total de AGS que de AGMI ante la limitación de P, sin embargo, esto puede explicarse en función de la especie estudiada, su perfil de ácidos grasos, el grado de limitación de P y la duración del cultivo (Lopes *et al.*, 2019; Lin *et al.*, 2018; Cañavate *et al.*, 2017). Por otro lado, como se mencionó anteriormente, el contenido total de AGPI está relacionado con los lípidos de membrana y su contenido se ve favorecido cuando la actividad celular es intensa por lo que a concentraciones bajas de P se puede verse reducido su contenido como un mecanismo de respuesta para generar compuestos de reserva y asegurar la supervivencia de las células algales en lugar de dividirse (Cañavate *et al.*, 2017; Benning *et al.*, 1995; Siron *et al.*, 1989).

### 6.3.1 Predicción de propiedades del biodiesel a partir de perfiles de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P

Se calcularon algunas propiedades de calidad del biodiesel a partir del perfil de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P tomando en consideración diversos métodos de estimación (Tabla 4).

**Tabla 4. Predicción de propiedades del biodiesel calculadas a partir del perfil de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P.**

Parámetro	Concentración de P			
	7 $\mu$ M	18 $\mu$ M	72 $\mu$ M	144 $\mu$ M
Índice de saponificación	202.48 $\pm$ 2.88 <sup>a</sup>	203.49 $\pm$ 0.72 <sup>a</sup>	202.30 $\pm$ 0.85 <sup>a</sup>	201.42 $\pm$ 0.25 <sup>a</sup>
Índice de yodo	108.22 $\pm$ 9.79 <sup>a</sup>	120.45 $\pm$ 5.54 <sup>a</sup>	135.13 $\pm$ 5.73 <sup>b</sup>	139.44 $\pm$ 0.85 <sup>b</sup>
Número de cetano	48.91 $\pm$ 2.45 <sup>a</sup>	46.01 $\pm$ 1.15 <sup>a</sup>	42.87 $\pm$ 1.18 <sup>b</sup>	42.02 $\pm$ 0.16 <sup>b</sup>
Grado de insaturación	81.58 $\pm$ 5.17 <sup>a</sup>	84.69 $\pm$ 3.30 <sup>ab</sup>	89.29 $\pm$ 2.00 <sup>bc</sup>	90.93 $\pm$ 0.29 <sup>c</sup>
Punto de turbidez	5.85 $\pm$ 0.74 <sup>a</sup>	5.51 $\pm$ 0.37 <sup>a</sup>	4.80 $\pm$ 0.74 <sup>ab</sup>	4.88 $\pm$ 0.16 <sup>b</sup>
Viscosidad cinemática	2.73 $\pm$ 0.27 <sup>a</sup>	2.61 $\pm$ 0.00 <sup>b</sup>	2.42 $\pm$ 0.00 <sup>c</sup>	5.94 $\pm$ 0.00 <sup>d</sup>
Estabilidad oxidativa	4.63 $\pm$ 0.19 <sup>a</sup>	4.51 $\pm$ 0.12 <sup>ab</sup>	4.34 $\pm$ 0.07 <sup>bc</sup>	4.27 $\pm$ 0.01 <sup>c</sup>
Ácido $\gamma$ -linolénico	2.05 $\pm$ 0.61 <sup>a</sup>	1.38 $\pm$ 0.26 <sup>ab</sup>	1.22 $\pm$ 0.24 <sup>ab</sup>	0.99 $\pm$ 0.22 <sup>b</sup>
Dobles enlaces $\geq$ 4	15.73 $\pm$ 1.99 <sup>a</sup>	19.65 $\pm$ 1.59 <sup>b</sup>	23.99 $\pm$ 1.60 <sup>c</sup>	25.09 $\pm$ 0.38 <sup>c</sup>

Los valores son la media  $\pm$  desviación estándar de cuatro réplicas biológicas y cada una está compuesta por tres réplicas técnicas. Los datos se presentan como porcentaje de ácidos grasos. Los valores en cada fila con diferente superíndice (a-d) indican una diferencia significativa ( $p < 0.05$ ).

El índice de saponificación (SV, por sus siglas en inglés) es un dato específico de cada grasa y aceite que está relacionado con el peso promedio, el contenido en porcentaje y longitud de las cadenas hidrocarbonadas de todos los ácidos grasos que forman parte de un aceite o grasa y se expresa en mg de hidróxido de potasio necesario para saponificar 1 g de aceite (Patterson, 2011). En este estudio, se encontraron valores de índice de saponificación entre 201 y 203 en *C. muelleri* y no se encontraron diferencias significativas entre los tratamientos.

Los valores de SV de *C. muelleri* en este estudio fueron similares a lo que se ha reportado para la diatomea *P. tricornutum*, varias especies de clorofitas tales como *Nannochloropsis oculata*, *Chlorella vulgaris*, *Ankistrodesmus fusiformis*, *Ankistrodesmus falcatus* y *Scenedesmus obliquus* e incluso a plantas oleaginosas como olivo (196 mg KOH/g de aceite), soja (195 mg KOH/g de aceite) y algodón (192 mg KOH/g de aceite), las cuales ya han sido utilizadas en ensayos para la obtención de biodiesel (Islam *et al.*, 2013; Talebi *et al.*, 2013; Ramos *et al.*, 2009; Dermibas, 2003). Por lo tanto, los resultados obtenidos en este parámetro para *C. muelleri* sugieren que su perfil de ácidos grasos es adecuado para la producción de biodiesel, ya que entre menor sea este valor a los ya reportados para las especies anteriormente mencionadas, es un indicador de un contenido alto de ácidos grasos poliinsaturados afectando la estabilidad del aceite y entre mayor sea (>240) indica un contenido alto de ácidos grasos saturados siendo más factible la elaboración de jabón ya que dichos valores son característicos de una grasa sólida (Bart *et al.*, 2010).

Otro parámetro calculado fue el índice de yodo (IV, por sus siglas en inglés), el cual se define como una medida que indica la cantidad de yodo ( $I_2$ ) en g que reacciona con 100 g de muestra y está relacionado directamente con el grado de insaturación total de ácidos grasos en un aceite o grasa (Barabás y Todorut, 2011). La definición de un valor adecuado de índice de yodo ha generado controversia en distintas legislaciones al considerar que el IV es simplemente un indicador del grado de insaturación mientras que el contenido de dobles enlaces influye contundentemente en la estabilidad oxidativa del aceite (Barabás y Todorut, 2011). En este estudio, se obtuvieron valores de IV de 108 y 139 para los perfiles de ácidos grasos de *C. muelleri*. Los lineamientos ASTM D6751 de Estados Unidos para la producción de biodiesel no establecen un valor adecuado para este parámetro; sin embargo, otras normativas recomiendan valores que no excedan entre 120 y 130 de IV, ya que entre mayor sea este valor, indica un mayor contenido de ácidos grasos poliinsaturados y aumenta la inestabilidad del aceite (Barabás y Todorut, 2011; European Committee for Standardization, 2008). Lo anterior sugiere que el perfil de ácidos grasos de *C. muelleri* que cumple mejor con las recomendaciones de estándares internacionales, fue el que se obtuvo al ser cultivada esta diatomea bajo una concentración de P de 7  $\mu$ M ( $108.22 \pm 9.79$  g de  $I_2$  / 100 g de muestra).

El número de cetano (CN, por sus siglas en inglés) es un parámetro que mide la calidad de ignición de un combustible y permite relacionar su idoneidad para su uso en motores de encendido por compresión. Es considerado uno de los indicadores principales para definir la calidad de un combustible y se expresa de forma adimensional (Barabás y Todorut, 2011). Teóricamente, el número de cetano puede alcanzar valores que van de 15 a 100 y se ha reportado que el rango adecuado de número de cetano es entre 41 – 56 (Barabás y Todorut, 2011). Si el número de cetano es bajo, el encendido del motor será difícil especialmente a bajas temperaturas y el motor funcionará de manera irregular y ruidosa con ciclos sin combustión, se demorará en calentar, las combustiones serán incompletas y aumentará la contaminación con gases de efecto invernadero (Knothe, 2008). Por el contrario, números de cetano altos conducen a una ignición prematura sin una mezcla apropiada con el aire, resultando en una combustión incompleta y aumento de las emisiones de humo; además el combustible se encendería cerca del inyector causando un sobrecalentamiento donde partículas del combustible que no lleguen a consumirse puedan tapan las boquillas del inyector (Barabás y Todorut, 2011).

Se estimaron valores de número de cetano promedio de 42 – 48 a partir de los perfiles de ácidos grasos de *C. muelleri* cultivada bajo diferentes concentraciones de P. Valores similares de este parámetro han sido observados en la diatomea *P. tricornutum* y *Picochlorum* sp (CN = 47.3) (Islam *et al.*, 2013); por su parte reportes de Stansell *et al.* (2012) establecieron un número de cetano promedio para biodiesel obtenido para de diatomeas de 48.5. Los estándares para producción de biodiesel de Brasil y Estados Unidos ASTM establecen valores mínimos de número de cetano de 45 y 47 respectivamente; por su parte las especificaciones en el continente europeo son aún más estrictas al establecer un valor mínimo aceptable de número de cetano de 51 (Barabás y Todorut, 2011). En vista de lo anterior, sólo el número de cetano estimado para el perfil de ácidos grasos de *C. muelleri* cultivada en una concentración de P de 7  $\mu\text{M}$  ( $48.91 \pm 2.45$ ) pudiera cumplir con lo propuesto por los estándares internacionales.

A pesar de que en los tratamientos en los que *C. muelleri* fue cultivada bajo concentraciones de P de 72 y 144  $\mu\text{M}$  ( $42.87 \pm 1.18$  y  $42.02 \pm 0.16$  respectivamente) se encontraron valores de número de cetano dentro de lo teóricamente recomendado, éstos se encuentran por debajo de los estándares internacionales y además poseen un contenido considerable de AGPI en comparación al ser cultivada bajo concentraciones limitantes de P por

lo que probablemente, bajo estas condiciones, convendría su uso en otras aplicaciones como alimento en acuicultura, suplementos alimenticios o farmacéutica. Sin embargo, una de las estrategias sugeridas para mejorar o aumentar números de cetano bajos son mezclar estos aceites con ácidos grasos saturados o implementar el uso de aditivos que previamente ya han sido utilizados para mejorar este parámetro como nitrato de 2-etilhexilo, nitrato de ciclohexilo y éter 2-metoxietilo, observándose un incremento de este valor de 45.5 a 63.5 (Li *et al.*, 2014).

Otro parámetro que se calculó en función de los perfiles de ácidos grasos de *C. muelleri* fue el grado de insaturación, el cual es un indicador del número de dobles enlaces presentes en los ácidos grasos que conforman el aceite; por lo que, una cantidad alta de dobles enlaces indicará un alto grado de insaturación. Asimismo, se ha reportado que otros indicadores de calidad del biodiesel como el índice de yodo y viscosidad cinemática están relacionados con este parámetro (Giakoumis y Sarakatsanis, 2019). Se encontraron diferencias significativas entre los tratamientos cuya concentración de P fue de 7  $\mu\text{M}$  ( $81.58 \pm 5.17$ ) y 144  $\mu\text{M}$  ( $90.93 \pm 0.29$ ), observando que el grado de insaturación aumenta con la concentración de P lo que también es congruente con el contenido de ácidos grasos, principalmente poliinsaturados, que también se incrementa (Tabla 4).

Otro de los indicadores importantes para evaluar la calidad del biodiesel es el punto de turbidez (CP, por sus siglas en inglés), el cual se define como la temperatura a la que el combustible comienza a manifestar cristales aglomerados, lo que conduce a la formación de un gel conforme baja la temperatura, haciendo que el combustible ya no fluya adecuadamente por el sistema del motor y tape los filtros (Barabás y Todorut, 2011). En este estudio, se estimaron valores promedios de CP de 4.88 y 5.85 °C para los perfiles de ácidos de *C. muelleri*. En *P. tricornutum* se ha reportado valores de CP de 4.47 °C y en biodiesel en general se ha reportado una temperatura promedio de 5 °C, los cuales son similares a lo calculado en este estudio (Babu y Anand, 2019). Cabe señalar que este parámetro se vuelve relevante cuando el combustible es utilizado en regiones geográficas donde predominan climas fríos por lo que se recomienda que, en estos casos, el biodiesel se mezcle con diésel (ya que posee un CP menor al biodiesel, CP = -28 °C) y aditivos para reducir el punto de turbidez e inhibir la aparición de cristales (Pérez y Lisinger, 2016).

Otro de los parámetros calculados a partir del perfil de ácidos grasos de *C. muelleri* fue la viscosidad cinemática (KV, por sus siglas de inglés). La viscosidad se define como la resistencia de un líquido a fluir debido a la fricción interna de una parte del fluido desplazándose sobre otro. Este indicador es importante ya que muestra la capacidad del aceite a fluir en los motores. La viscosidad alta puede causar tamaños de gotas grandes, una vaporización pobre, un ángulo de inyección más estrecho y una mayor penetración en el cilindro de la pulverización de combustible (Hasimoglu *et al.*, 2008; Knothe *et al.*, 2005). Los resultados obtenidos en este estudio para este indicador oscilaron entre 2.42 – 5.94 mm<sup>2</sup>/s. Las especificaciones para la producción de biodiesel ASTM D6751 de Estados Unidos mencionan como valores aceptables de viscosidad de 1.9 – 6.0 mm<sup>2</sup>/s, lo cual es satisfecho en todos los tratamientos planteados en este estudio (ASTM, 2008).

Dentro de las propiedades de importancia a evaluar del biodiesel se encuentra la estabilidad oxidativa (OS, por sus siglas en inglés), la cual está definida por la composición de ácidos grasos presentes en el aceite, especialmente por el contenido de AGPI, el cual es inversamente proporcional al valor que se obtenga en este parámetro (Knothe *et al.*, 2005). Otros factores que influyen en la estabilidad oxidativa son el almacenamiento, temperatura, tiempo, presencia de oxígeno, metales traza y luz (Almeida *et al.*, 2019; Hoekman *et al.*, 2012). Los valores obtenidos de OS en este estudio oscilaron entre 4.27 – 4.63 h, cumpliendo con el valor mínimo establecido por los estándares de calidad para producción de biodiesel ASTM D6751 de Estados Unidos (OS = 3h) (ASTM, 2008).

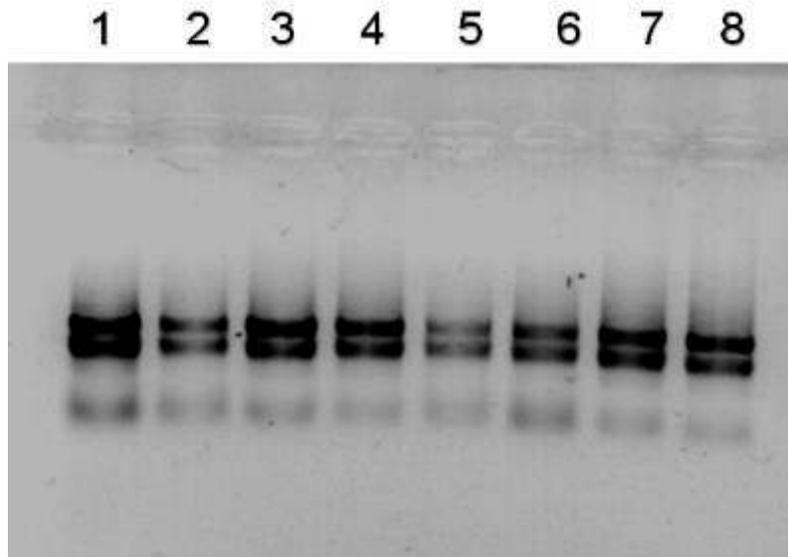
Otras recomendaciones establecidas para considerar biomasa potencial para la producción de biodiesel de acuerdo a los estándares internacionales, es que éstas posean cantidades bajas de ácido linolénico (< 12 %) y de ácidos grasos con más de cuatro dobles enlaces (<1%); la primera de las especificaciones es satisfecha en este estudio en todos los tratamientos mientras que la segunda no es cumplida en ningún caso (European Committee for Standardization, 2008). Por su parte, para los estándares ASTM D6751 de Estados Unidos, estas dos características (contenido de ácido linolénico y contenido de ácidos grasos con más de cuatro dobles enlaces) no son relevantes (ASTM, 2008). En conclusión, los resultados obtenidos de la estimación de parámetros de calidad de biodiesel a partir de los perfiles ácidos grasos de *C. muelleri* de este estudio, indican que aquel que fue obtenido en el tratamiento con una

concentración de P de 7  $\mu\text{M}$  es un recurso potencial para la producción de biodiesel al cumplir con los requerimientos de la norma ASTM D6751 de Estados Unidos para la producción de biodiesel.

## 6.4 Análisis transcriptómico de *C. muelleri* cultivada bajo diferentes concentraciones de P

### 6.4.1 Integridad del ARN total

Se obtuvieron dos réplicas biológicas (cada una de ellas estuvo compuesta por tres réplicas técnicas) de ARN total de *C. muelleri* por cada tratamiento con la calidad, concentración y cantidad requerida para llevar a cabo el proceso de secuenciación (Figura 7). Posteriormente, las muestras fueron enviadas al Laboratorio de Servicios Genómicos de LANGEBIO-CINVESTAV Unidad Irapuato, Guanajuato, México.



**Figura 7.** ARN total de *C. muelleri* extraído con reactivo TRIzol (Invitrogen) y purificado con el kit comercial RNeasy MinElute Clean Up Kit (QIAGEN) en gel de agarosa al 1 % teñido con GelRed (Biotium) en condiciones no desnaturizantes. Carril del 1 al 4 corresponden a la primera réplica biológica (RB1) y del carril 5 al 8 corresponde a la segunda réplica biológica (RB2). Carril 1: 7  $\mu\text{M}$ \_RB1. Carril 2: 18  $\mu\text{M}$ \_RB1. Carril 3: 72  $\mu\text{M}$ \_RB1. Carril 4: 144  $\mu\text{M}$ \_RB1. Carril 5: 7  $\mu\text{M}$ \_RB2. Carril 6: 18  $\mu\text{M}$ \_RB2. Carril 7: 72  $\mu\text{M}$ \_RB2. Carril 8: 144  $\mu\text{M}$ \_RB2.

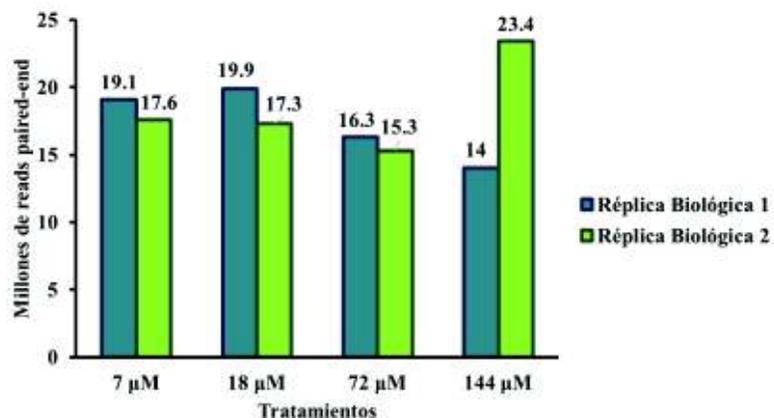
Una vez que se las muestras de ARN total de *C. muelleri* fueron recibidas en el Laboratorio de Servicios Genómicos de LANGEBIO-CINVESTAV Unidad Irapuato, Guanajuato, México, nuevamente la concentración e integridad del ARN total fue corroborada espectrofotométricamente en un Agilent 2100 Bioanalyzer (Agilent Technologies) y se determinó el número de integridad del ARN (RIN, por sus siglas en inglés), encontrándose con la calidad requerida para dar continuidad con el proceso de secuenciación.

#### **6.4.2 Análisis y filtrado de calidad de secuencias**

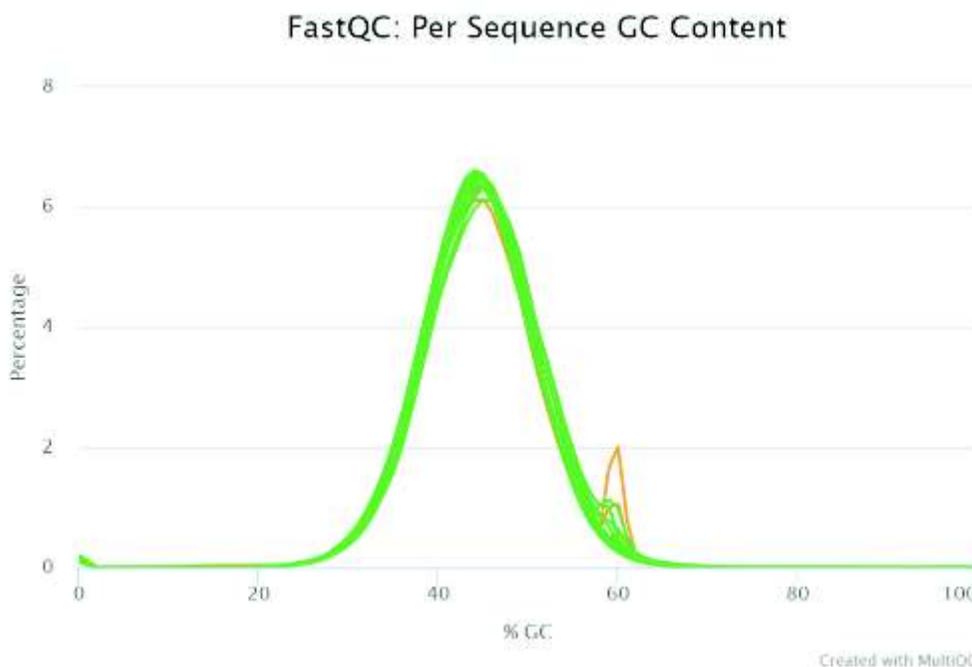
Los datos crudos de secuenciación fueron descargados de la base de datos de los Laboratorios de Servicios Genómicos de LANGEBIO-CINVESTAV y se obtuvieron en promedio 142.9 millones de lecturas pareadas (Figura 8). Posteriormente, se realizó un análisis de calidad de cada una de las librerías utilizando el programa FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) (Andrews, 2010) para determinar la calidad de las secuencias de cada genoteca. Se observó que el porcentaje de GC entre las muestras se mantuvo entre 43 y 45 %, sin embargo, en el gráfico correspondiente a este parámetro se pudo observar una curva adicional de menor tamaño que pudiera indicar algún tipo de contaminación (adaptadores, secuencias de ARNr, entre otros) (Figura 9). Por otra parte, el contenido de pares de bases ambiguas ('N') se encontró por debajo del 1 %.

Después del análisis de calidad de los datos crudos de secuenciación, se procedió a realizar un filtrado de las lecturas mediante el uso del software Trimmomatic v0.36 (Bolger *et al.*, 2014) con el fin de remover lecturas de baja calidad ( $Q_{\text{phred}} < 25$ ), adaptadores, bases ambiguas ('N') y lecturas con una longitud menor a 50 pb. Además, se utilizó el software SortMeRNA (Kopylova *et al.*, 2013) para eliminar la posible presencia de secuencias de ARNr. Luego de este procesamiento de los datos, nuevamente se hizo uso del software FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) (Andrews, 2010) para verificar si los inconvenientes observados en el análisis de calidad de los datos crudos fueron corregidos. Una vez finalizado el filtrado de calidad de las secuencias (Figura 10), la calidad promedio en la escala de Phred tanto de las bases como las secuencias permaneció superior a 30 ( $Q_{\text{phred}} > 30$ ), el contenido de GC osciló entre 43 y 44 % (Figura 11), se logró la eliminación de adaptadores,

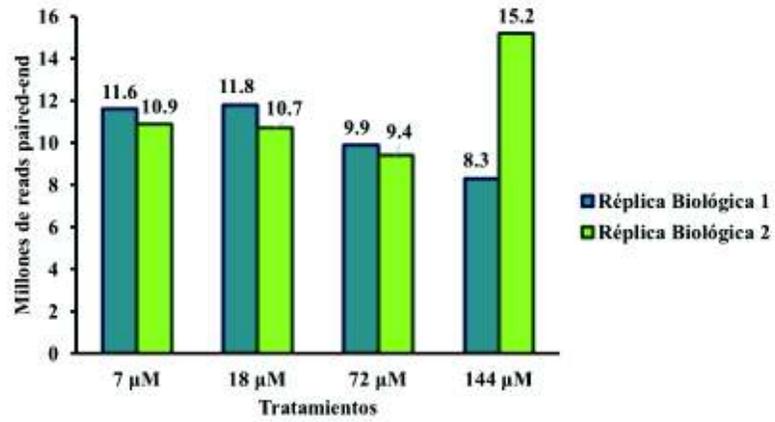
bases de baja calidad y ambiguas así como de secuencias de ARNr (Tabla 5), obteniéndose en total 87.8 millones de lecturas pareadas, las cuales fueron utilizadas para el ensamblado *de novo* del transcriptoma (Figura 10).



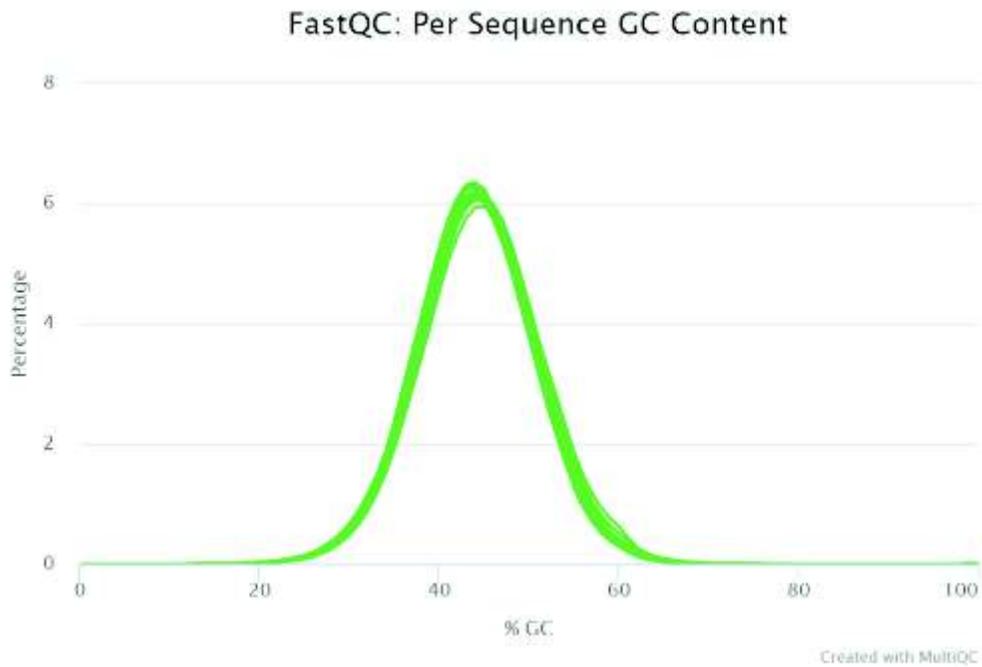
**Figura 8.** Conteo de datos crudos (lecturas) por librería y por tratamiento previo al filtrado de calidad.



**Figura 9.** Porcentaje de GC (citosina y guanina) en los datos crudos de secuenciación.



**Figura 10.** Conteo de lecturas por librería y por tratamiento posterior al filtrado de calidad con los softwares Trimmomatic y SortMeRNA.



**Figura 11.** Porcentaje de GC (citosina y guanina) de las lecturas posterior al filtrado de calidad.

**Tabla 5. Secuencias de ARN ribosomal detectadas en las librerías de *C. muelleri* por el software SortMeRNA.**

<b>Tratamiento</b>	<b>Réplica</b>	<b>Hit a ARNr 18S</b>
7 $\mu$ M	1	916,024
18 $\mu$ M	1	445,440
72 $\mu$ M	1	1,624,866
144 $\mu$ M	1	798,676
7 $\mu$ M	2	442,202
18 $\mu$ M	2	426,226
72 $\mu$ M	2	250,726
144 $\mu$ M	2	359,272

### 6.4.3 Ensamblado del transcriptoma y alineamiento

Una vez que las lecturas fueron filtradas se continuó con el ensamblado *de novo* del transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P. Para este fin se consideraron tres ensambladores: Trinity (en condiciones por defecto y utilizando la opción ‘jaccard clip’), rnaSPAdes y SOAPdenovo. Posteriormente, la calidad de cada uno de los transcriptomas ensamblados se evaluó a través de las herramientas rnaQUAST, BUSCO con la base de datos de Stramenopiles y Gene Mark S-T. Tras realizar un análisis de las métricas proporcionadas por cada una de las herramientas se definió que los transcriptomas ensamblados con Trinity presentaron los resultados más adecuados para seguir con el análisis (Tabla 6).

**Tabla 6. Métricas proporcionadas por rnaQUAST, BUSCO y Gene Mark S-T a partir de los transcriptomas ensamblados con Trinity, rnaSPAdes y SOAPdenovo.**

	Ensamblado Trinity default	Ensamblado Trinity jaccard clip	Ensamblado rnaSPAdes	Ensamblado SOAPdenovo	
Transcritos	33466	35272	26725	68215	
Transcritos > 500 pb	26715	28282	22593	1886	
Transcritos > 1000 pb	21603	21941	16592	335	
Tamaño promedio de los transcritos ensamblados (pb)	1956.179	1735.062	1709.95	175.262	
Transcrito de mayor tamaño (pb)	17911	17909	17778	8987	
N50	1719	1856	2467	125	
<b>BUSCO</b>					
Porcentaje de genes completos recuperados (%)	Con una sola copia	52	59	56	0
	Duplicados	46	39	37	0
	Total	98	98	93	0
Porcentaje de genes fragmentados recuperados (%)		1	1	4	2
Porcentaje de genes ausentes (%)		1	1	3	98
<b>GeneMarkS-T</b>					
Número de genes predichos	22105	24500	18778	1070	

Una vez seleccionados los transcriptomas candidatos para continuar con el análisis, se realizó un alineamiento de las lecturas filtradas contra los transcriptomas ensamblados con Trinity utilizando la herramienta HISAT2 con el fin de determinar el nivel de cobertura y representatividad de dichas lecturas en el transcriptoma y se recomienda que entre un 70 y 90 % de dichas lecturas deben de alinearse al transcriptoma ensamblado (Geniza y Jaiswal, 2017). El porcentaje de alineamiento de las lecturas con los transcriptomas ensamblados varió entre 79.62 y 81.63 % (Tabla 7), considerándose adecuados ambos ensamblados, sin embargo, al verificar nuevamente las métricas obtenidas para ambos ensamblados, se escogió el transcriptoma ensamblado con Trinity utilizando la opción ‘jaccard clip’ como el más adecuado para dar continuidad a los análisis ya que de acuerdo a la evaluación realizada por la herramienta BUSCO a pesar de que ambos transcriptomas mostraron el mismo porcentaje total de transcritos completos recuperados, el transcriptoma ensamblado con Trinity utilizando la opción ‘jaccard

clip’, mostró un mayor porcentaje de transcritos completos recuperados únicos (59 %) en comparación con el ensamblado en sus condiciones por default (52 %).

**Tabla 7. Porcentajes totales de alineamiento de lecturas con los transcriptomas ensamblados con la herramienta Trinity.**

<b>Tratamiento</b>	<b>Réplica</b>	<b>Ensamblado default (%)</b>	<b>Ensamblado jaccard clip (%)</b>
7 µM	1	81.48	81.23
18 µM	1	80.26	80.28
72 µM	1	80.64	80.54
144 µM	1	79.62	79.97
7 µM	2	80.78	80.64
18 µM	2	80.67	80.71
72 µM	2	80.27	80.58
144 µM	2	81.47	81.63

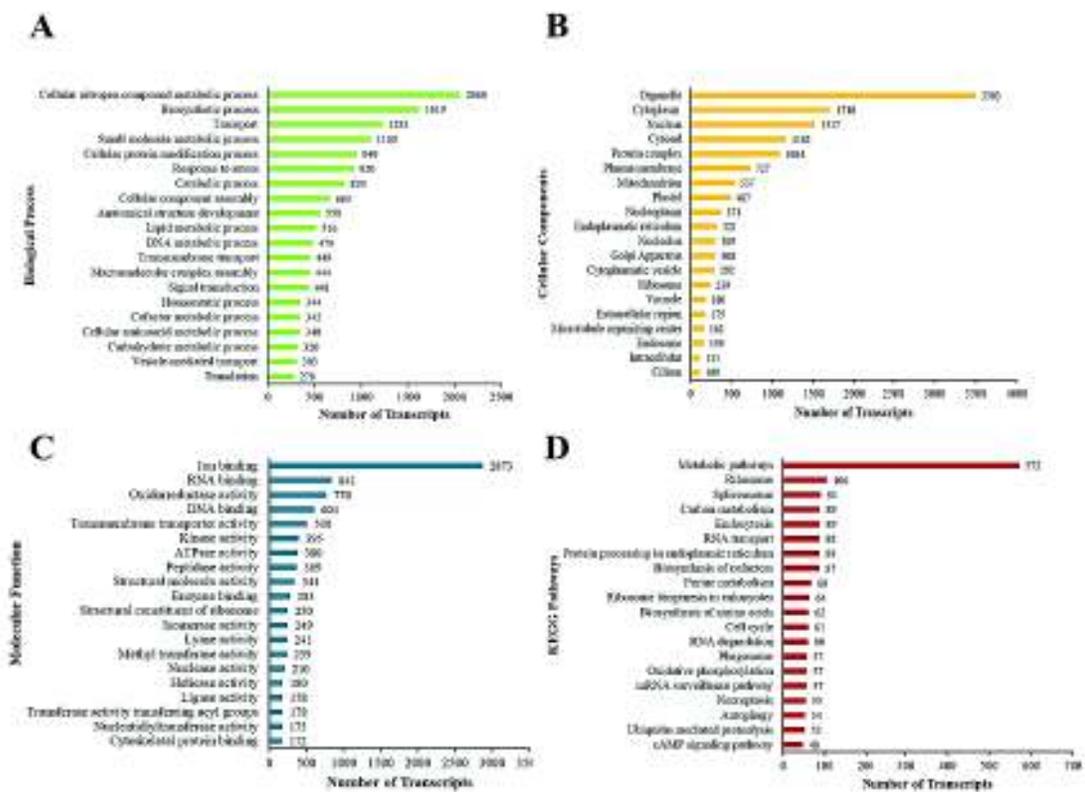
#### **6.4.4 Anotación funcional del transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P**

De los 35,272 transcritos ensamblados por Trinity utilizando la opción ‘jaccard clip’, se predijeron 54,951 marcos de lectura abiertos (ORFs, por sus siglas en inglés) distribuidos en 28,220 transcritos mediante la herramienta Transdecoder, la cual posee la capacidad de identificar dentro de los transcritos posibles regiones codificantes. Se realizó un blast del transcriptoma de *C. muelleri* utilizando ‘blastx’ (valor de  $E < 1e^{-3}$ ) contra la base de datos de UniProtKB/Swiss-Prot dando hit a 15,236 secuencias de proteínas; también se realizó un ‘blastp’ utilizando los ORF predichos con Transdecoder igualmente contra la base de datos UniProtKB/Swiss-Prot dando hit a 9,030 secuencias de proteínas. Además, con la finalidad de robustecer el análisis se realizó un ‘blastx’ del transcriptoma de *C. muelleri* contra proteomas de otras microalgas depositados en la base de datos de de UniProtKB/Swiss-Prot; de los 35,272

transcritos, 17,705 dieron hit a secuencias de proteínas en *Chlamydomonas reinhardtii*, 28,702 para *Thalassiosira pseudonana* y 30,279 a *Phaeodactylum tricornutum*.

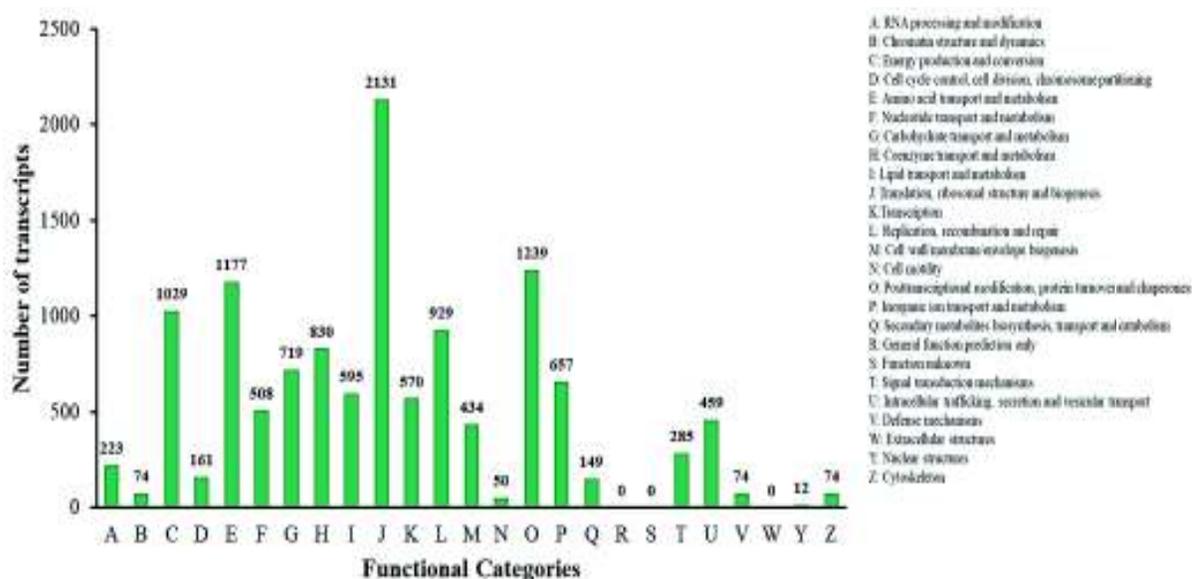
De los 35,272 transcritos identificados en el transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P, 15,839 transcritos fueron anotados por Gene Ontology (GO), los cuales fueron clasificados en tres categorías GO principales: Procesos biológicos (42.3 %), componentes celulares (31.9 %) y funciones moleculares (25.8 %) (Figura 12). Estas categorías a su vez se subdividen en otras categorías facilitando así la descripción de la identidad biológica de aquellos transcritos que pudieron ser anotados. Dentro de la categoría de procesos biológicos, las categorías más representativas fueron: Proceso metabólico del compuesto de nitrógeno celular (2060 transcritos), proceso biosintético (1619 transcritos) y transporte (1231 transcritos) (Figura 12A). Con respecto a la categoría de componentes celulares, las categorías con mayor número de transcritos fueron: Organelo (3503 transcritos), citoplasma (1710 transcritos) y núcleo (1517 transcritos) (Figura 12B). Finalmente, en relación a la categoría de funciones moleculares, las subcategorías más abundantes fueron: Unión de iones (2873 transcritos), unión de ARN (841 transcritos) y actividad oxidorreductasa (604 transcritos) (Figura 12C).

Para el caso de la anotación funcional generada con la base de datos de KEGG, se observó un total de 5,721 transcritos anotados, los cuales estuvieron involucrados principalmente en rutas metabólicas (572 transcritos), ribosoma (106 transcritos), espliceosoma (91 transcritos) y metabolismo del carbono (89 transcritos) (Figura 12D).



**Figura 12.** Histogramas de clasificaciones de transcritos anotados de *C. muelleri* cultivada bajo diferentes concentraciones de P con la base de datos Gene Ontology (GO). A: Procesos biológicos. B: Componentes celulares. C: Funciones moleculares. D: Rutas metabólicas anotadas con KEGG.

Por otro lado, 12,380 transcritos fueron anotados con la base de datos de Grupos de Ortólogos (COG, por sus siglas en inglés) y clasificados dentro de 22 categorías funcionales de las 25 que integran esta base de datos (Figura 13). La categoría que presentó un mayor número de transcritos anotados correspondió a “Traducción, biogénesis y estructural ribosomal” (2131 transcritos), seguido de “Modificaciones postranscripcionales, recambio proteico y chaperonas” (1239 transcritos) y “Transporte y metabolismo de aminoácidos” (1177 transcritos); por el contrario, dentro las categorías funcionales “Predicción de funciones generales”, “Función desconocida” y “Estructuras extracelulares” no se identificaron transcritos anotados (Figura 13).

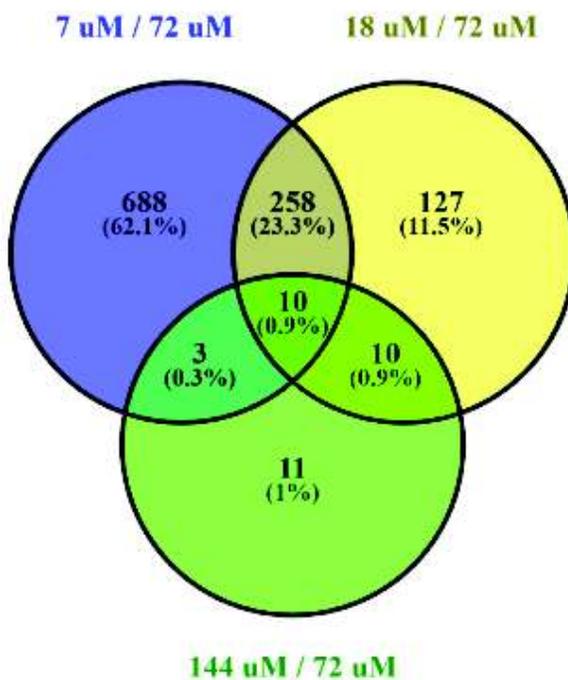


**Figura 13.** Histograma de clasificaciones de transcritos anotados de *C. muelleri* cultivada bajo diferentes concentraciones de P con la base de datos Clusters Orthologs Groups (COG).

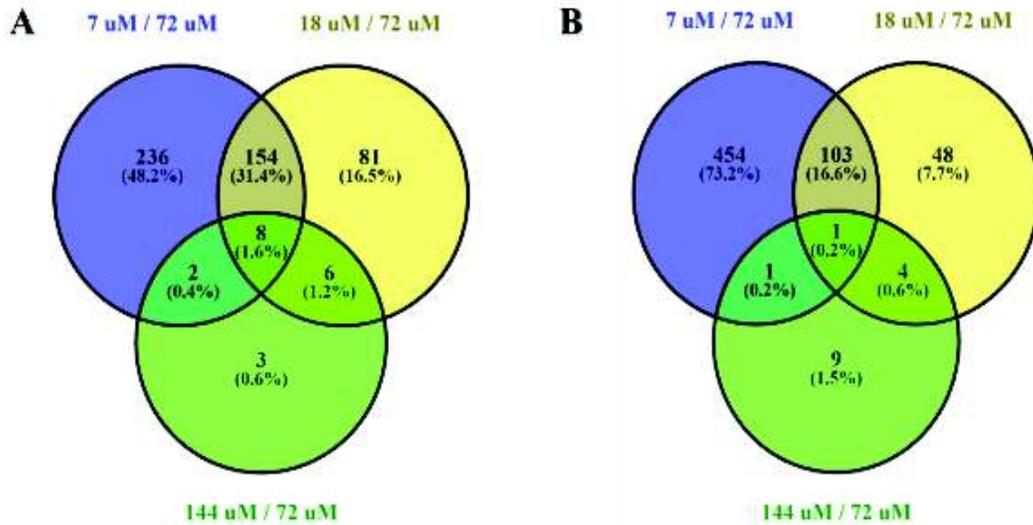
#### 6.4.5 Análisis de los genes expresados diferencialmente en *C. muelleri* cultivada bajo diferentes concentraciones de P

La identificación de los transcritos expresados diferencialmente en *C. muelleri* cultivada bajo diferentes concentraciones de P fue realizado haciendo uso del paquete EdgeR. Los transcritos que presentaron una tasa de descubrimiento de falsos positivos (FDR, por sus siglas en inglés) <0.01 y un valor de cambio (log2fold-change)  $\geq \pm 1$  fueron considerados como expresados diferencialmente. Los transcritos expresados diferencialmente en *C. muelleri* cultivada bajo

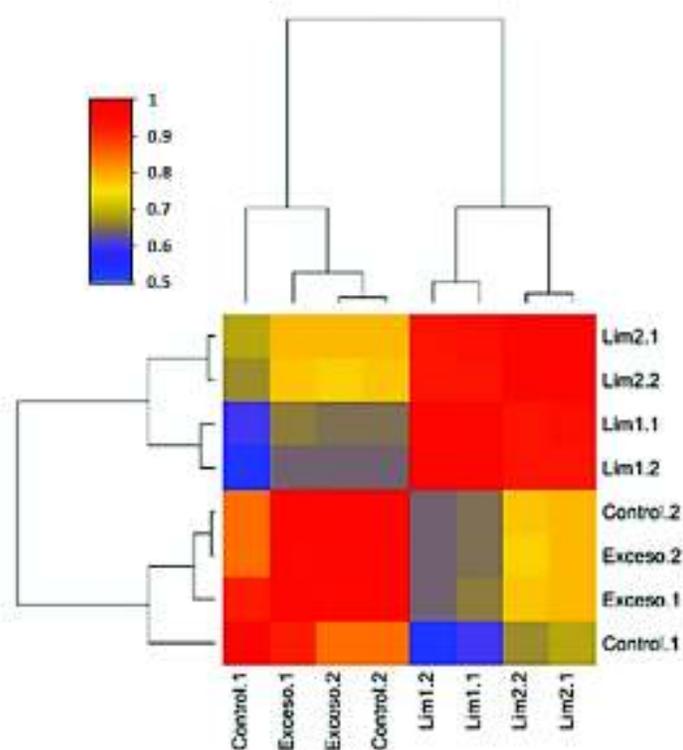
concentraciones limitantes (7 y 18  $\mu\text{M}$ ) y exceso de P (144  $\mu\text{M}$ ) fueron determinados en comparación contra el tratamiento considerado como control (72  $\mu\text{M}$ ). El mayor número de transcritos expresados diferencialmente se observó en el tratamiento cuya concentración de P fue de 7  $\mu\text{M}$  (959; 400 inducidos y 559 reprimidos), seguido del tratamiento con concentración de P de 18  $\mu\text{M}$  (405; 249 inducidos y 156 reprimidos) y por último el tratamiento con exceso de P (34; 19 inducidos y 15 reprimidos) (Figuras 14 y 15). Se puede observar que el número de genes expresados diferencialmente incrementó a medida que la concentración de P fue menor; por otro lado, en base a los resultados obtenidos de los análisis de correlación de Pearson y de escalado multivariado (Figuras 16 y 17), se pudo apreciar que tanto el tratamiento control como exceso son muy parecidos por lo que eso explicaría en parte el escaso número de transcritos expresados diferencialmente.



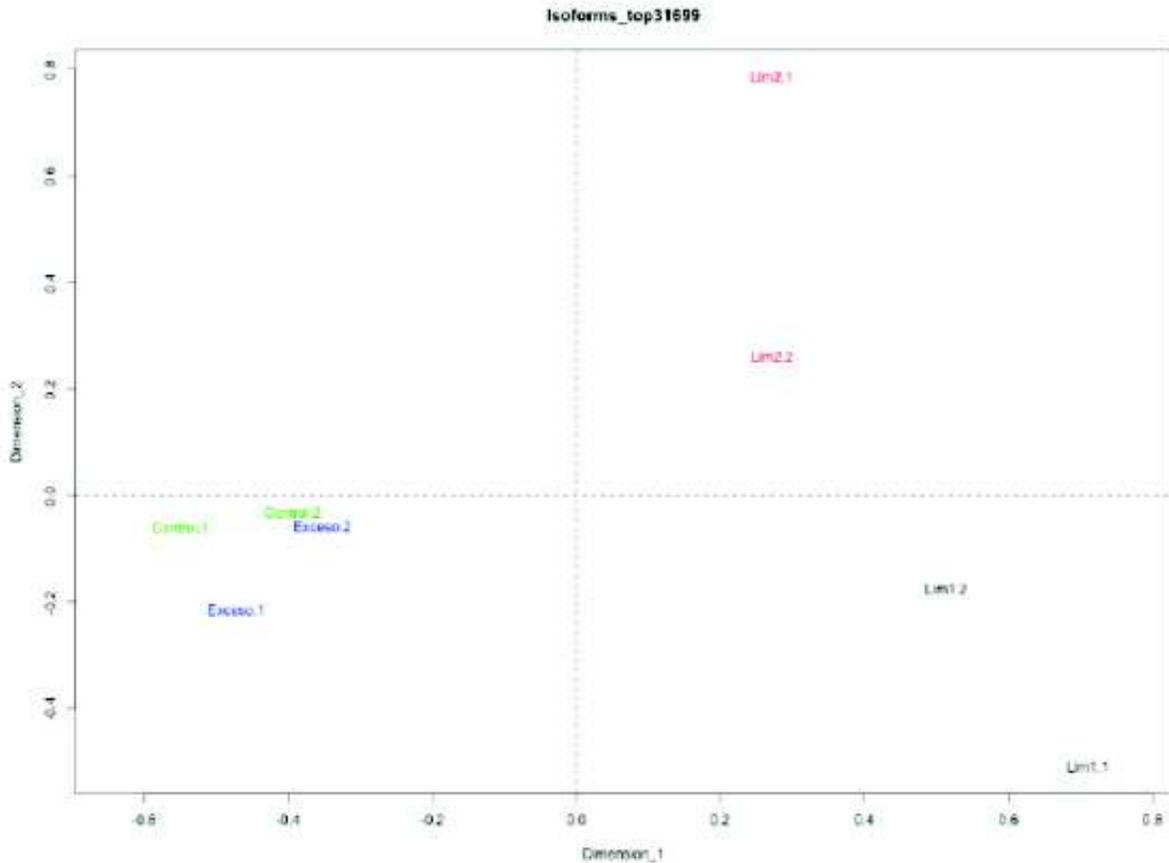
**Figura 14.** Diagrama de Venn de los genes expresados diferencialmente de *C. muelleri* cultivada bajo diferentes concentraciones de P.



**Figura 15.** Diagramas de Venn de genes expresados diferencialmente en *C. muelleri* cultivada bajo diferentes concentraciones de P. A: Número de DEG inducidos en *C. muelleri*. B: Número de DEG reprimidos en *C. muelleri*.



**Figura 16.** Heatmap de análisis de correlación de Pearson tomando como referencia la matriz de conteos de isoformas de *C. muelleri*. Tratamiento 7  $\mu$ M (Limitante 1.1 y Limitante 1.2). Tratamiento 18  $\mu$ M (Limitante 2.1 y Limitante 2.2). Tratamiento 72  $\mu$ M (Control.1 y Control.2) y Tratamiento 144  $\mu$ M (Exceso.1 y Exceso.2).

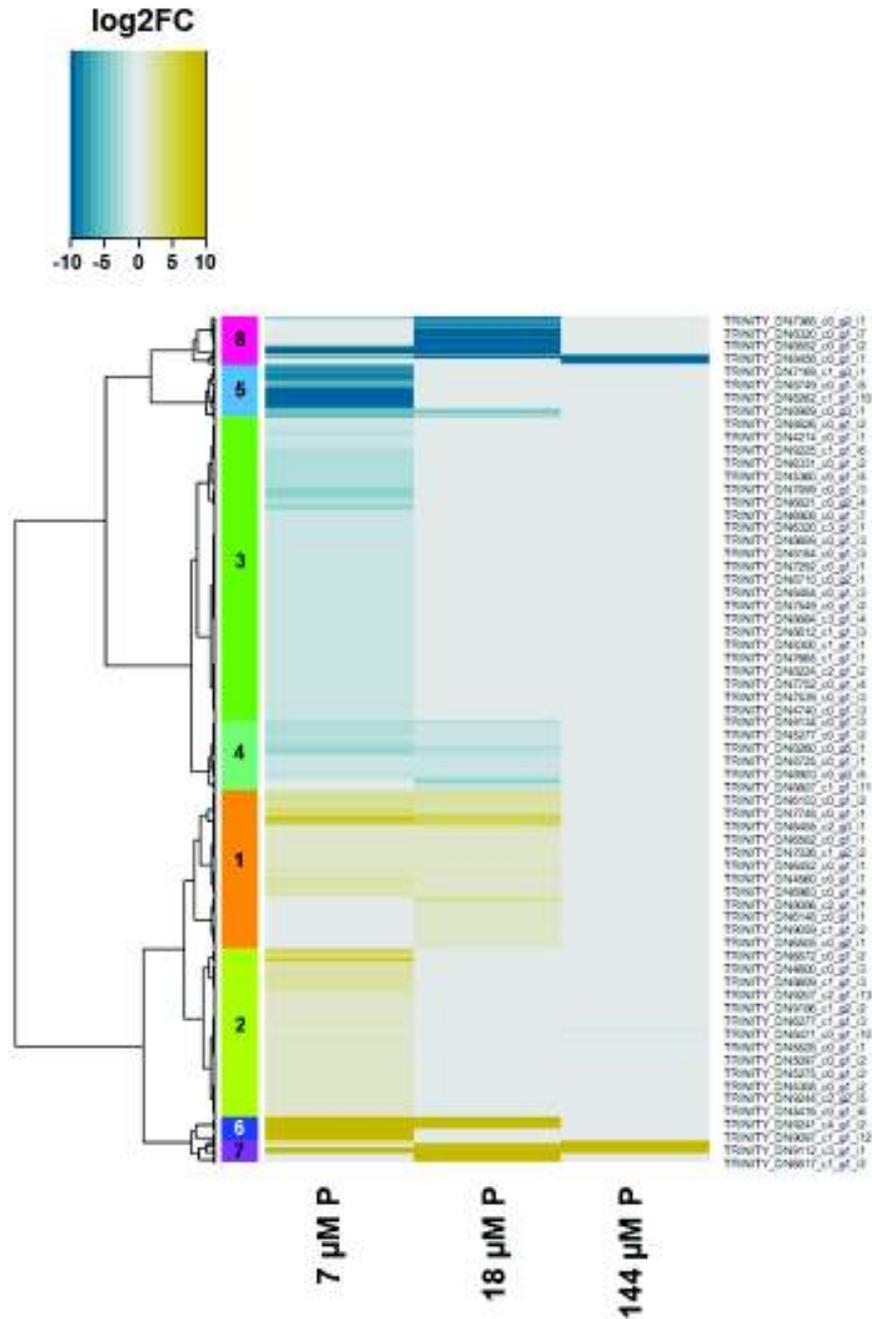


**Figura 17.** Análisis de escalado multidimensional (MDS, por sus siglas en inglés) tomando como referencia la matriz de conteos de isoformas de *C. muelleri*. Tratamiento 7  $\mu\text{M}$  (Limitante 1.1 y Limitante 1.2) en color negro. Tratamiento 18  $\mu\text{M}$  (Limitante 2.1 y Limitante 2.2) en color rojo. Tratamiento 72  $\mu\text{M}$  (Control.1 y Control.2) en color verde y Tratamiento 144  $\mu\text{M}$  (Exceso.1 y Exceso.2) en color azul.

Posteriormente, se realizó un agrupamiento jerárquico de los genes expresados diferencialmente en *C. muelleri* de acuerdo a su perfil y a la concentración de P que fue cultivada (Figura 18). Se observaron 8 clusters, de los cuales el cluster 3 presentó la mayor cantidad de transcritos (397) expresados diferencialmente, específicamente reprimidos en la concentración de 7  $\mu\text{M}$  de P, relacionados principalmente con procesos biológicos tales como traducción de proteínas, fotosíntesis y asimilación de la luz y ensamblado de la subunidad pequeña ribosomal, así como también actividades de unión de clorofila y ARN. El segundo cluster de mayor tamaño fue el número 2 y comprendió de 221 transcritos principalmente inducidos en el tratamiento de 7  $\mu\text{M}$  de P involucrados en el metabolismo de glucosa y homeostasis celular redox, así como actividades hidrolasa y de unión de fosfolípidos dependiente de calcio. Por su parte, el cluster 1

comprendió de 208 genes expresados diferencialmente que se encontraron inducidos en ambas concentraciones limitantes de P (7 y 18  $\mu\text{M}$ ), los cuales se encuentran principalmente relacionados con autofagia, catabolismo de lípidos y gluconeogénesis; también en este cluster se observaron transcritos con actividad de unión de ADN y transporte por antiporte o contransporte antiparalelo.

Por otro lado, el cluster 4 comprendió de 91 transcritos reprimidos en ambas concentraciones limitantes de P (7, y 18  $\mu\text{M}$ ) y dentro del cual se identificaron transcritos involucrados en procesos biológicos tales como traducción de proteínas y transporte transmembranal de amonio, así como relacionados con funciones moleculares como componente estructural del ribosoma y unión de ARN. Por lo que se refiere al cluster 5, se observó que está comprendido por 68 transcritos fuertemente reprimidos correspondientes al tratamiento con mayor limitación de P (7  $\mu\text{M}$ ) y que están relacionados con actividad de transporte simporte ó paralelo y unión de iones metálicos. Por lo que se refiere al cluster 8, está formado por 63 transcritos fuertemente reprimidos en las tres concentraciones de P (7, 18 y 144  $\mu\text{M}$ ) que fueron comparadas contra el control (72  $\mu\text{M}$ ) y se encontraron algunos genes que presentan actividad ATPasa y de unión de ATP. El cluster 6 se compuso por 30 transcritos altamente inducidos en ambas concentraciones limitantes (7, y 18  $\mu\text{M}$ ) y se encontraron genes con actividad de unión de ATP, mientras que el cluster 7 se conformó por 29 transcritos altamente inducidos en las concentraciones de 7, 18 y 144  $\mu\text{M}$  en donde también se identificaron genes relacionados con actividad de unión de ATP.

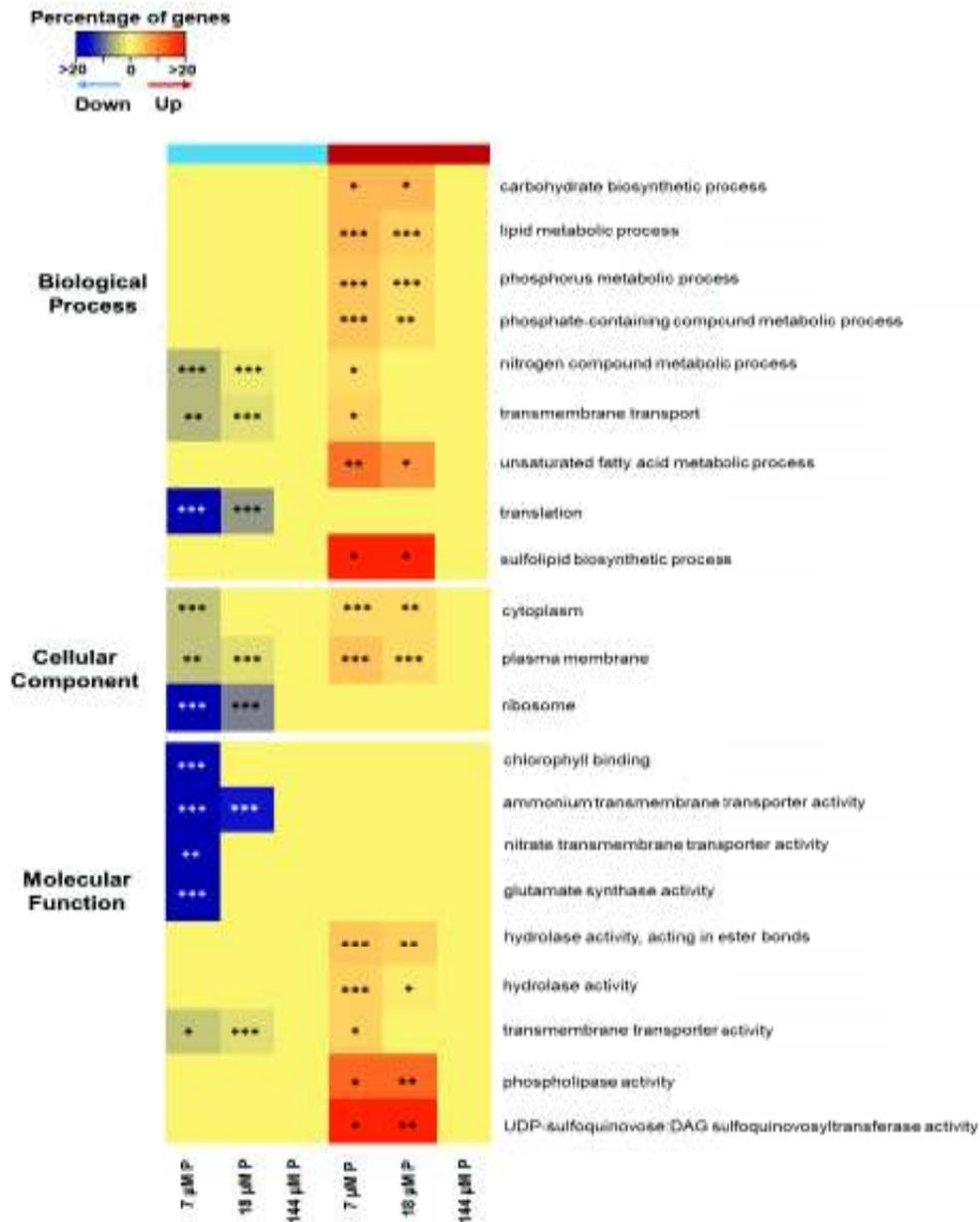


**Figura 18.** Heatmap de agrupamiento jerárquico de los genes expresados diferencialmente de *C. muelleri* cultivada bajo diferentes concentraciones de P. La escala de colores representa el Log<sub>2</sub>FC de los genes expresados diferencialmente (FDR < 0.01 y log<sub>2</sub>FC ≥ ± 1), oscilando de azul, para la mayoría de los genes reprimidos, a amarillo, para la mayoría de genes inducidos. El dendrograma indica la relación entre los perfiles genéticos utilizando el método de agrupamiento ward.D.

#### 6.4.6 Análisis de enriquecimiento GO de los genes expresados diferencialmente en *C. muelleri* cultivada bajo diferentes concentraciones de P

El análisis de enriquecimiento para la clasificación funcional en procesos biológicos, componentes celulares y funciones moleculares que se regularon diferencialmente se realizó mediante la asignación de términos GO por Trinotate y haciendo uso del paquete de R GOSEQ v1.26.0 considerando un FDR < 0.05 (Figura 19). En condiciones limitantes (7 y 18  $\mu\text{M}$ ) de P en *C. muelleri* se pudo observar la inducción de genes involucrados en procesos biológicos relacionados con el metabolismo del P y de compuestos en cuya composición se encuentre presente, metabolismo de ácidos grasos insaturados, transporte transmembranal, procesos metabólicos de compuestos de nitrógeno, metabolismo de lípidos y biosíntesis de carbohidratos y sulfolípidos. Por otro lado, genes involucrados en procesos biológicos tales como traducción, transporte transmembranal y procesos metabólicos de compuestos de nitrógeno se vieron reprimidos (Figura 19). Con respecto al análisis de términos GO más específicos se observó que la respuesta transcripcional de *C. muelleri* cultivada bajo concentraciones limitantes de P (7 y 18  $\mu\text{M}$ ) implicó principalmente la inducción de genes relacionados con actividad hidrolasa en general y actuando sobre enlaces ésteres, actividad de transporte transmembranal, actividad fosfolipasa y actividad UDP-sulfoquinovosa: DAG sulfoquinovosiltransferasa; por el contrario, en el caso de genes reprimidos, se observaron aquellos involucrados en actividad de transporte transmembranal, unión de clorofila, actividad de transporte transmembranal de amonio y nitratos así como de actividad de glutamato sintasa. Por otra parte, en el caso del tratamiento de exceso de P (144  $\mu\text{M}$ ) en *C. muelleri*, bajo estos parámetros, no se identificaron categorías GO enriquecidas significativamente (Figura 19).

Estos resultados son congruentes con lo observado en estudios previos en diatomeas en donde dentro de las respuestas que se han observado para lidiar ante la limitación o ausencia de P se encuentran la inducción de genes que codifican a enzimas relacionadas con el transporte, la asimilación, la removilización y el reciclaje de este nutriente, así como de aquellos involucrados en la síntesis de compuestos de reserva como carbohidratos y lípidos. Además, la inhibición de la fotosíntesis, del metabolismo de nitrógeno y síntesis de proteínas está relacionada con la limitación de P de diatomeas (Alipanah *et al.*, 2018; Zhang *et al.*, 2016; Feng *et al.*, 2015; Shih *et al.*, 2015; Dyhrman *et al.*, 2012).



**Figura 19.** Heatmap del análisis de enriquecimiento de los principales grupos de Gene Ontology (GO) observados *C. muelleri* cultivada bajo diferentes concentraciones de P. Cada columna contiene los términos GO reprimidos (azul) e inducidos (naranja). Sólo se muestran los términos GO más específicos (FDR<0.05). La escala de colores representa el porcentaje de genes que pertenecen a cada término GO. Los asteriscos representan la importancia estadística del enriquecimiento (\*FDR<0.05, \*\*FDR<0.01 y \*\*\*FDR<0.001 respectivamente).

#### 6.4.7 Respuesta transcripcional de *C. muelleri* cultivada bajo concentraciones limitantes de P

Como resultado de la presente investigación doctoral, se reporta por primera vez el transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P. Al igual que en otros estudios transcriptómicos realizados en otras diatomeas que fueron cultivadas bajo limitación por P (Alipanah *et al.*, 2018; Zhang *et al.*, 2016; Cruz de Carvalho *et al.*, 2016; Shih *et al.*, 2015; Yang *et al.*, 2014; Dyhrman *et al.*, 2012), en el transcriptoma de *C. muelleri* se identificaron respuestas tanto específicas a la disponibilidad y asimilación de P como respuestas generales derivadas de la limitación de este nutriente (Figura 20).

Cuando los niveles de Pi en el medio de cultivo o en el ambiente disminuyen, se ha reportado para diatomeas la inducción de transcritos que codifican a proteínas transportadoras de P e hidrolasas ésteres (tales como fosfatasas alcalinas y nucleotidasas) involucradas en la asimilación, transporte, distribución y aprovechamiento eficiente de este nutriente (Alipanah *et al.*, 2018; Zhang *et al.*, 2016; Shih *et al.*, 2015; Dyhrman *et al.*, 2012). De acuerdo al análisis de RNA-seq realizado en este estudio, dentro de las estrategias implementadas por *C. muelleri* relacionadas directamente con la respuesta a la limitación por P, se observó la inducción de un transcrito que codifica a una proteína cotransportadora de fosfatos dependiente de sodio (NPT2B) a la cual se le ha atribuido un rol importante en la asimilación de P en diatomeas (Shih *et al.*, 2015; Dyhrman *et al.*, 2012). Por otra parte, se identificaron en *C. muelleri* transcritos inducidos que codifican a proteínas transportadoras de Pi pertenecientes a las familias de genes PHT1 y PHT5, las cuales han sido identificadas tanto en levadura como en plantas y están relacionadas con la asimilación, translocación, movilización y homeostasis celular del Pi (Chang *et al.*, 2019; Lira-Morales *et al.*, 2019; Yang *et al.*, 2017) (Tabla 8).

Por otra parte, otras de las estrategias empleadas por diatomeas ante la limitación de P es la síntesis de fosfatasas y otras enzimas hidrolíticas con la finalidad de que, en caso de que el Pi sea escaso en el medio ambiente, pueda hacer uso de otros recursos como el DOP u otras moléculas que posean P en su composición bioquímica para su asimilación (Lin *et al.*, 2016). En *C. muelleri* se identificaron dos transcritos inducidos que codificaron a una fosfatasa alcalina (APase) y una NAD 5'-nucleotidasa (5-NAD), lo que sugiere que bajo estas condiciones, esta diatomea además de transportar el escaso Pi que puede estar presente en el medio de cultivo,

también desarrolla estrategias para hidrolizar biomoléculas (fosfolípidos, ácidos nucleicos y fosfonatos) que posean P en su forma orgánica de tal manera que se vuelva biodisponible para su aprovechamiento (Tabla 8).

Una vez que el Pi es absorbido y transportado al interior de la célula, éste puede ser empleado para satisfacer los requerimientos celulares o ser almacenado en la vacuola mediante la acción del complejo de transporte vacuolar (VTC, por sus siglas en inglés) (Gerasimaité *et al.*, 2014; Hothorn *et al.*, 2009). En *C. muelleri* se identificaron dos transcritos inducidos que codifican a las subunidades VTC1 y VTC4 del complejo de transporte vacuolar, el cual se encuentra involucrado en la síntesis de Poli-P y en la translocación, distribución y homeostasis del Pi en el interior de la célula (Tabla 8) (Yang *et al.*, 2017; Uttenweiler *et al.*, 2007; Müller *et al.*, 2003). Se ha reportado que la inducción de los genes que codifican a las subunidades VTC1 y VTC4 en otros estudios transcriptómicos sugiere que bajo la limitación por P, en lugar de llevarse a cabo la síntesis de Poli-P, la actividad del complejo vacuolar se encontrará relacionada con la regulación de los niveles de Poli-P presentes en la vacuola con el fin de satisfacer las necesidades celulares por lo que se sugiere que los reservorios de Poli-P presentes en *C. muelleri* están siendo utilizados con el objetivo de mitigar la limitación de P (Zhang *et al.* 2016; Dyhrman *et al.*, 2012).

La disponibilidad de P es un factor importante que regula varios procesos entre ellos el crecimiento y la división celular (Zhang *et al.*, 2016; Shih *et al.*, 2015; Kavanová *et al.*, 2006). Se identificaron varios transcritos inducidos involucrados en la regulación del ciclo celular en el transcriptoma de *C. muelleri* que codifican a ciclinas del tipo B, las cuales son subunidades reguladoras positivas de las cinasas dependientes de las ciclinas (CDK) y, por lo tanto, desempeñan un papel esencial en el control del ciclo celular, en particular mediante su destrucción durante la división celular y su presencia puede ser necesaria para retrasar la citocinesis (Tabla 8) (Lim y Kaldis, 2013; Morgan, 2007). En el transcriptoma de la diatomea *S. costatum* bajo limitación por P varios genes que codificaban a ciclinas B y CDK se encontraron inducidos, pero una vez que las concentraciones de P fueron favorables, estos genes se encontraron reprimidos; los autores sugieren que la inducción de estos genes mantuvo detenido el ciclo celular y, por lo tanto, la división celular cesó (Zhang *et al.*, 2016); este

antecedente genera evidencia de que esto mismo pudiera estar sucediendo en *C. muelleri* ante la limitación de P de acuerdo al análisis de RNA-seq.

Otro de los procesos que se vio afectado por la limitación de P fue la fotosíntesis en donde la mayoría de los genes relacionados con la síntesis de pigmentos como clorofila *a* y fucoxantina se encontraron reprimidos, lo que conduce a una disminución en el contenido de pigmentos provocando fotoinhibición, daño del fotosistema II y aumento en los niveles de radicales libres derivando en estrés oxidativo (Tabla 9) (Alipanah *et al.*, 2018; Moseley *et al.*, 2006). Sin embargo, según el análisis de RNA-seq para evitar daño al aparato fotosintético en *C. muelleri*, dentro de las estrategias que se implementan se encuentra la inducción de genes que codifican a enzimas denominadas “válvulas de electrones”, las cuales dentro de sus funciones se encuentra la depuración del exceso de energía de la cadena de transporte de electrones fotosintética (Moseley *et al.*, 2006; Noriega *et al.*, 2004); entre los transcritos que se encontraron inducidos que codificaron a este tipo de enzimas se encontraron la ATP sintasa (ATPB) y la ubiquinol oxidasa 4 (AOX4), las cuales se han visto tanto en plantas como *Arabidopsis* y sojay en microalgas que protegen al aparato fotosintético del daño fotooxidativo (Wang *et al.*, 2019; Moseley *et al.*, 2006; Kanazawa y Kramer, 2002).

En relación con lo anterior, de acuerdo al análisis de RNA-seq, otra de las estrategias que *C. muelleri* podría estar llevando a cabo para contrarrestar el estrés oxidativo en los cloroplastos derivado de la limitación por P es la inducción de transcritos relacionados con el ciclo de las xantófilas, el cual se ha visto también en plantas que es uno de los mecanismos que se sabe están involucrados en la protección celular del estrés fotooxidativo mediante la disipación del exceso de energía y regulación de los niveles de especies reactivas de oxígeno (ROS, por sus siglas en inglés) (Hernández y Bosch, 2015; Marijuan y Munné-Bosch, 2014) (Tabla 8). Entre los transcritos inducidos encontrados en *C. muelleri* se observó uno que codifica a zeaxantina epoxidasa (ZEP), la cual se encarga de llevar a cabo la conversión de zeaxantina en anteraxantina y posteriormente a violaxantina la cual posee actividad antioxidante (Fu *et al.*, 2011).

Por otro lado, la inducción de ZEP también se encuentra relacionada con el metabolismo del ácido abscísico (ABA), el cual ha sido poco estudiado en diatomeas en comparación a plantas en cuanto a su rol en la respuesta ante la limitación por P (Jarosz *et al.*, 2020; Yoshida *et al.*, 2019; Stigter y Plaxton, 2015). En plantas, ABA es la fitohormona más importante que proporciona resistencia a condiciones de estrés mediante el incremento de su contenido y promoviendo la inducción de genes involucrados en la adaptación a dichas condiciones adversas (Yoshida *et al.*, 2019; Sah *et al.*, 2016; Yoshida *et al.*, 2014). Otros transcritos que se encontraron expresados diferencialmente en *C. muelleri* relacionados con la señalización por ABA fueron aquellos que codificaron a la subunidad 2 del complejo de elongación de la ARN polimerasa II (ELP2), syntaxina-61 (SYP61) y una glicotransferasa (EPC1) (Tablas 8 y 9). Además, se identificó un transcrito reprimido que codifica a una proteína transportadora ABC (ABCG25) la cual se encarga de secretar ABA hacia el exterior de la célula como un mecanismo de regulación de los niveles de esta hormona; esta respuesta pudiera indicar que bajo limitación por P, *C. muelleri* favorece la concentración de ABA en su interior para inducir la señalización mediada por esta hormona y así responder a las condiciones de estrés (Park *et al.*, 2016; Romanenko *et al.*, 2015; Cramer *et al.*, 2011).

Otros de los transcritos que se encontraron inducidos como respuesta al estrés oxidativo derivado de la limitación por P en *C. muelleri* fueron aquellos que codificaron a peroxirredoxinas e hidrolasas NUDIX. La síntesis de peroxirredoxinas es vital para la protección celular contra el estrés oxidativo ya que poseen la capacidad de eliminar rápidamente ROS en diferentes organelos como cloroplastos y mitocondrias (Rhee *et al.*, 2016) (Tabla 8). Por otra parte, los transcritos que codificaron a hidrolasas NUDIX en *C. muelleri* están involucradas en la protección de los ácidos nucleicos del daño oxidativo promoviendo la degradación de nucleósidos difosfato, la reducción de ROS y la eliminación de mutaciones de ADN y errores transcripcionales (Zhang *et al.*, 2016; Yoshimura *et al.*, 2015; Fonseca y Dong, 2014).

En relación con el metabolismo de ácidos nucleicos, dentro de los transcritos que se encontraron expresados en *C. muelleri* fueron aquellos que codificaron para adenilosuccinato sintetasa (AdSS) y una proteína de unión de poliadenilato 1 (PABP-1) (Tablas 8 y 9). AdSS es una enzima importante en la ruta *de novo* y en la ruta de recuperación de la biosíntesis de purinas que cataliza el primer paso en la biosíntesis de AMP a partir de inosina monofosfato (IMP, por

sus siglas en inglés) (Stayton *et al.*, 1983). El transcrito que codificó a AdSS se encontró inducido y en función de los resultados que se han observado hasta el momento, se sugiere que la biosíntesis de purinas en *C. muelleri* bajo estas condiciones se da a través de la ruta de recuperación ya que esta se centra en el reciclamiento de bases nitrogenadas y nucleósidos para mantener un aporte equilibrado en células que se encuentran en reposo (Baynes y Dominiczak, 2009). Por otro lado, PABP-1 está involucrado en la estabilidad del ARNm mediante la adición de cola poli-A y el transcrito que codificó a esta enzima en *C. muelleri* se encontró reprimido por lo que indicaría que el proceso de maduración de ARNm se está viendo afectado y que puede influir más adelante en el proceso de traducción de proteínas (Luong y Contig, 2019; Dyhrman *et al.*, 2012; Van Mooy *et al.*, 2009).

Por otra parte, de acuerdo al análisis de RNA-seq, dentro de los procesos más afectados por la limitación por P en *C. muelleri* fue el de traducción, en donde además se encontró el mayor número de transcritos expresados diferencialmente, los cuales estuvieron relacionados con la síntesis, estructura y función de las subunidades ribosomales y en su mayoría se encontraron reprimidos (Tabla 9). Esto puede explicarse dado que bajo condiciones limitantes de P, el proceso de traducción se ralentiza debido a la disminución en la síntesis de ribosomas ya que implica una inversión energética para la célula. Estos organelos son reservorios importantes de ARN que como se sabe son moléculas ricas en P, por lo que bajo estas condiciones se pretende preservar el escaso P que pueda estar disponible para satisfacer las necesidades básicas de supervivencia (Alipanah *et al.*, 2018; Li *et al.*, 2018; Dyhrman *et al.*, 2012).

Al disminuir la síntesis de ribosomas y por ende la síntesis de proteínas a través del proceso de traducción, el metabolismo del nitrógeno (N) y el metabolismo de aminoácidos también resultaron afectados por la limitación por P en *C. muelleri* y se identificaron en el análisis de RNA-seq varios transcritos que codificaron a proteínas transportadoras de N y relacionados con la síntesis de aminoácidos reprimidos (Tabla 9). En plantas se ha observado que ante la limitación por P, se incrementan los niveles de aminoácidos libres derivados de un proceso catabólico de proteínas generando “señales” de un supuesto exceso de N por la acumulación de aminoácidos y por tanto inhibiendo el transporte de N (Ahmad *et al.*, 2018; Huang *et al.*, 2008; Misson *et al.*, 2005). Lo anterior coincide en lo observado en microalgas

expuestas a bajas concentraciones de P, que al disminuir los requerimientos necesarios para la biosíntesis de ribosomas y proteínas, el transporte de N disminuyó y se ha observado en estudios transcriptómicos la inducción de transcritos relacionados con el catabolismo de aminoácidos (Alipanah *et al.*, 2018; Wurch *et al.*, 2011), por lo que probablemente esto mismo pudiera estar ocurriendo en *C. muelleri* bajo estas condiciones.

A pesar de que la división celular, la fotosíntesis y la síntesis de proteínas disminuyen por efecto de la limitación de P, se ha reportado en diatomeas que la síntesis de estructuras de carbono sigue superando las necesidades celulares y que dicho exceso de carbono es redireccionado hacia la biosíntesis de compuestos de reserva en forma de TAG o excretado fuera de la célula como EPS (Brembu *et al.*, 2017; Zhang *et al.*, 2016; Shih *et al.*, 2015; Dyhrman *et al.*, 2012; Alcoverro *et al.*, 2000). En *C. muelleri* se identificaron varios transcritos expresados diferencialmente que codificaron a enzimas que se encuentran involucradas tanto en el proceso de glucólisis como gluconeogénesis, sin embargo, cabe destacar que se observó en *C. muelleri* la inducción de un transcrito que codifica a piruvato carboxilasa (PC), una enzima clave en gluconeogénesis, lo que sugiere que bajo estas condiciones este proceso se está llevando a cabo en lugar de la glucólisis, ya que su inducción marca el inicio de la obtención de glucosa a partir de piruvato u oxaloacetato como una estrategia para la obtención de energía (Tabla 8) (Yang *et al.*, 2014; Garrett *et al.*, 2008; Jitrapakdee y Wallace, 1999).

Por otra parte, la inducción de transcritos relacionados con la degradación de crisolaminarina y transporte de azúcares hacia el aparato de Golgi en *C. muelleri* indica que se están generando precursores para la síntesis de sulfolípidos o polisacáridos que pueden acumularse o ser excretados fuera de la célula (EPS), siendo esto último una respuesta que se ha reportado en otras diatomeas ante la limitación por P como un mecanismo para reducir el exceso de carbono en el interior de las células y así además proveer de ciertos beneficios a las diatomeas como protección de depredadores, retención de agua e inclusive dichos EPS representan una fuente de carbono para bacterias heterotróficas, las cuales pueden crear una relación simbiótica con las diatomeas proveyéndolas de Pi a medida de que éstas consumen estos polisacáridos (Alipana *et al.*, 2018; Brembu *et al.*, 2017; Guerrini *et al.*, 2000).

**Tabla 8. Lista de transcritos inducidos en *C. muelleri* cultivada bajo limitación por P de acuerdo al análisis de RNA-seq.**

<b>Nombre del gen</b>	<b>Nombre de proteína</b>	<b>Proceso</b>
<i>SLC34A2</i>	Proteína transportadora de fósforo dependiente de sodio 2B (NPT2B)	Transporte de Pi
<i>PHT1</i>	Proteína putativa transportadora de Pi	Transporte de Pi
<i>PHT5</i>	Proteína transportadora vacuolar de fosfatos 1	Homeostasis Pi
<i>phoA</i>	Fosfatasa alcalina	Hidrólisis DOP
<i>5NTD</i>	NAD 5'-nucleotidasa	Hidrólisis DOP
<i>VTC1</i>	Chaperona de transporte vacuolar 1	Homeostasis celular Pi
<i>VTC4</i>	Chaperona de transporte vacuolar 4	Homeostasis celular Pi
<i>CycB3</i>	Ciclina B3	Ciclo celular
<i>CYCB1-3</i>	Ciclina B1-3	Ciclo celular
<i>CCNB2</i>	Ciclina-2 específica de G2/M (ciclina tipo B)	Ciclo celular
<i>atpB</i>	ATP sintasa subunidad B	Fotosíntesis
<i>AOX4</i>	Ubiquinol oxidasa 4	Fotosíntesis
<i>ZEP</i>	Zeaxantina epoxidasa	Ciclo xantófilas y señalización ABA
<i>SYP61</i>	Sintaxina-61	Señalización ABA
<i>EPCI</i>	Proteína C4 de la familia de las glicosiltransferasas	Señalización ABA
<i>sIII1621</i>	Peroxirredoxina sIII1621	Respuesta a estrés
<i>PRXIII-2</i>	Peroxirredoxina-2E	Respuesta a estrés
<i>PRXIIF</i>	Peroxirredoxina-2F	Respuesta a estrés
<i>NUDX1</i>	Hidrolasa Nudix 1	Protección ácidos nucleicos
<i>NUDX8</i>	Hidrolasa Nudix 8	Protección ácidos nucleicos
<i>PHATRRAFT_26256</i>	Adenilosuccinato sintetasa (AdSS)	Síntesis purinas

**Tabla 8. Lista de transcritos inducidos en *C. muelleri* cultivada bajo limitación por P de acuerdo al análisis de RNA-seq (continuación).**

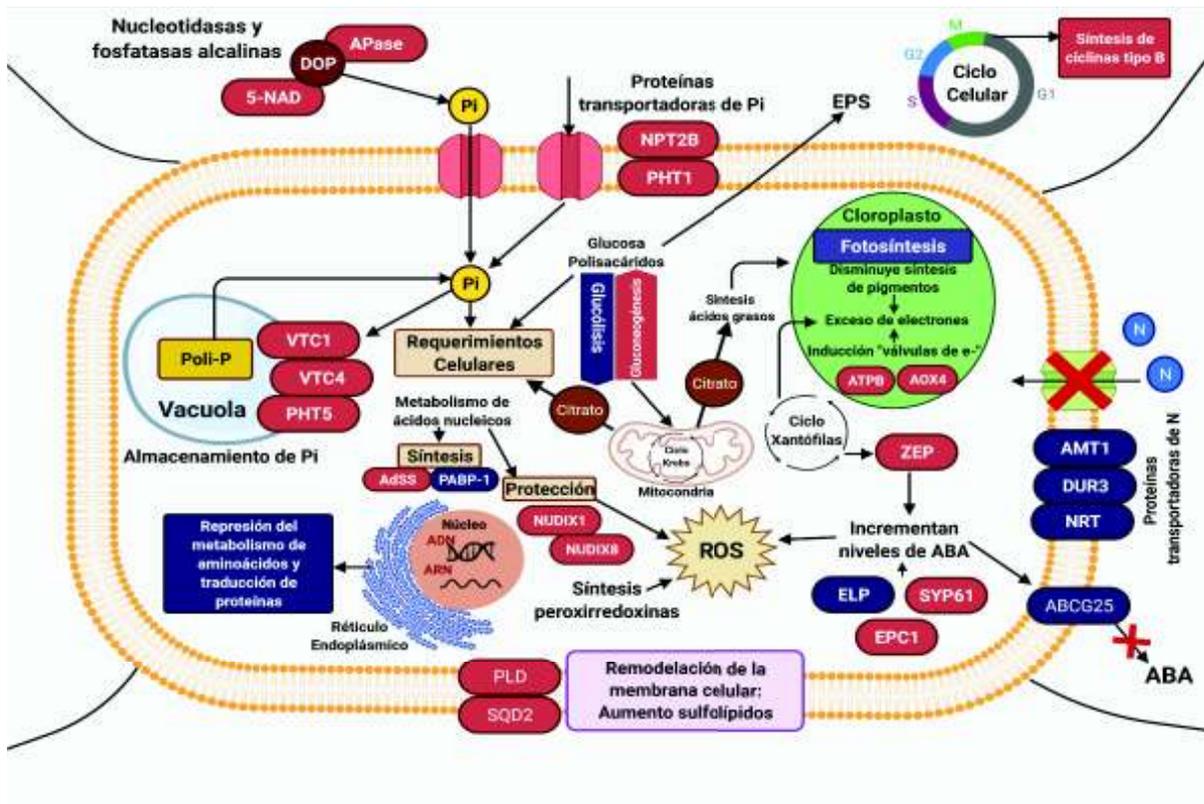
Nombre del gen	Nombre de proteína	Proceso
<i>pyc</i>	Piruvato carboxilasa	Gluconeogénesis
<i>gap3</i>	Gliceraldehído-3-fosfato deshidrogenasa 3	Glucólisis / Gluconeogénesis
<i>galE</i>	UDP-glucosa-4 epimerasa	Metabolismo de carbohidratos
<i>UTR4</i>	Proteína transportadora 4 de UDP glucosa y UDP galactosa.	Transporte a aparato de Golgi
<i>gpmA</i>	Fosfogliceromutasa	Glucólisis / Gluconeogénesis
<i>PGMI</i>	Fosfoglucomutasa	Degradación crisolaminarina / Síntesis de glucosa
<i>glk</i>	Glucocinasa	Síntesis de glucosa

**Tabla 9. Lista de transcritos reprimidos en *C. muelleri* cultivada bajo limitación por P de acuerdo al análisis de RNA-seq.**

Nombre del gen	Nombre de proteína	Proceso
<i>FCPE</i>	Proteína de unión de fucoxantina y clorofila a-c, E	Fotosíntesis
<i>FCPA</i>	Proteína de unión de fucoxantina y clorofila a-c	Fotosíntesis
<i>FCPF</i>	Proteína de unión de fucoxantina y clorofila a-c, F	Fotosíntesis
<i>chlD</i>	Magnesio quelatasa subunidad chlD	Síntesis de clorofila
<i>PSBO</i>	Proteína extrínseca del fotosistema II	Síntesis de clorofila
<i>ABCG25</i>	Proteína transportadora ABC tipo G	Transporte de ABA
<i>ELP2</i>	Subunidad 2 del complejo de elongación de la ARN polimerasa II	Señalización ABA
<i>Pabpc1</i>	Proteína de unión de poliadenilato 1 (PABP-1).	Maduración ARNm
<i>RPL4A</i>	Proteína 60S ribosomal	Traducción
<i>rps24</i>	Proteína 40S ribosomal S24	Traducción
<i>Rps7</i>	Proteína 40S ribosomal S7	Maduración ARNr

**Tabla 9. Lista de transcritos reprimidos en *C. muelleri* cultivada bajo limitación por P de acuerdo al análisis de RNA-seq (continuación).**

<b>Nombre del gen</b>	<b>Nombre de proteína</b>	<b>Proceso</b>
<i>NRT2.5</i>	Proteína transportadora de nitratos	Transporte de N
<i>AMT1-3</i>	Proteína transportadora de amonio	Transporte de N
<i>DUR3</i>	Proteína transportadora de urea	Transporte de N
<i>Os01g0681900</i>	Glutamato sintasa 1	Biosíntesis de aminoácidos
<i>gltS</i>	Glutamato sintasa 2 dependiente de ferredoxina	Biosíntesis de aminoácidos
<i>NIA</i>	Nitrato reductasa	Transporte de N
<i>GLNA</i>	Glutamina sintetasa	Biosíntesis de glutamina
<i>gcs-1</i>	Glutamato cisteína ligasa	Biosíntesis de glutatión



**Figura 20.** Modelo hipotético de la respuesta transcripcional de *C. muelleri* ante la limitación por P. Cuadros en color rojo: Transcritos inducidos. Cuadros en color azul marino: Transcritos reprimidos. Creado en Biorender: [www.biorender.com](http://www.biorender.com)

#### 6.4.8 Respuesta transcripcional de *C. muelleri* cultivada bajo exceso de P

Por lo que respecta al efecto de la concentración de exceso de P en *C. muelleri*, éste solo pudo ser percibido mediante el análisis de RNA-seq al identificarse la expresión diferencial de algunos transcritos puntuales ya que, de acuerdo a los análisis bioquímicos realizados en el presente estudio de manera general, no se encontraron diferencias significativas con respecto al tratamiento control. En referencia al análisis de RNA-seq, a pesar de que no se identificaron términos GO enriquecidos, se identificaron transcritos expresados diferencialmente relacionados con el ciclo celular y metabolismos tales como el de ácidos nucleicos, hormonas, lípidos y glucólisis (Figura 21).

Se identificó un transcrito reprimido relacionado con ciclo celular que codificó a una ciclina de tipo D (CYCD3), las cuales están relacionadas con la transición de la fase G1 (interfase) y S (replicación de ADN) (O'Connor *et al.*, 2010; Alvarado-Moreno y Mayani, 2006). Como ya se mencionó anteriormente, la síntesis y degradación de ciclinas es importante para regular el proceso de división celular y se sabe que al encontrarse presentes interactúan con complejos cinasas dependientes de ciclinas (CDK, por sus siglas en inglés), los cuales requieren de estas enzimas para activarse y garantizar el correcto desarrollo del ciclo celular; una vez que cumplen su función, las ciclinas se disocian de los CDK permitiendo la progresión del ciclo celular (Yang, 2018; Alvarado-Moreno y Mayani, 2006). Por lo tanto, se sugiere que en *C. muelleri* bajo estas condiciones el proceso de división celular se está efectuando, ya que al encontrarse *CYCD3* reprimida, indica que no se encuentra presente y por tanto no se une a su CDK correspondiente permitiendo la progresión de la fase G1 a S1 del ciclo celular.

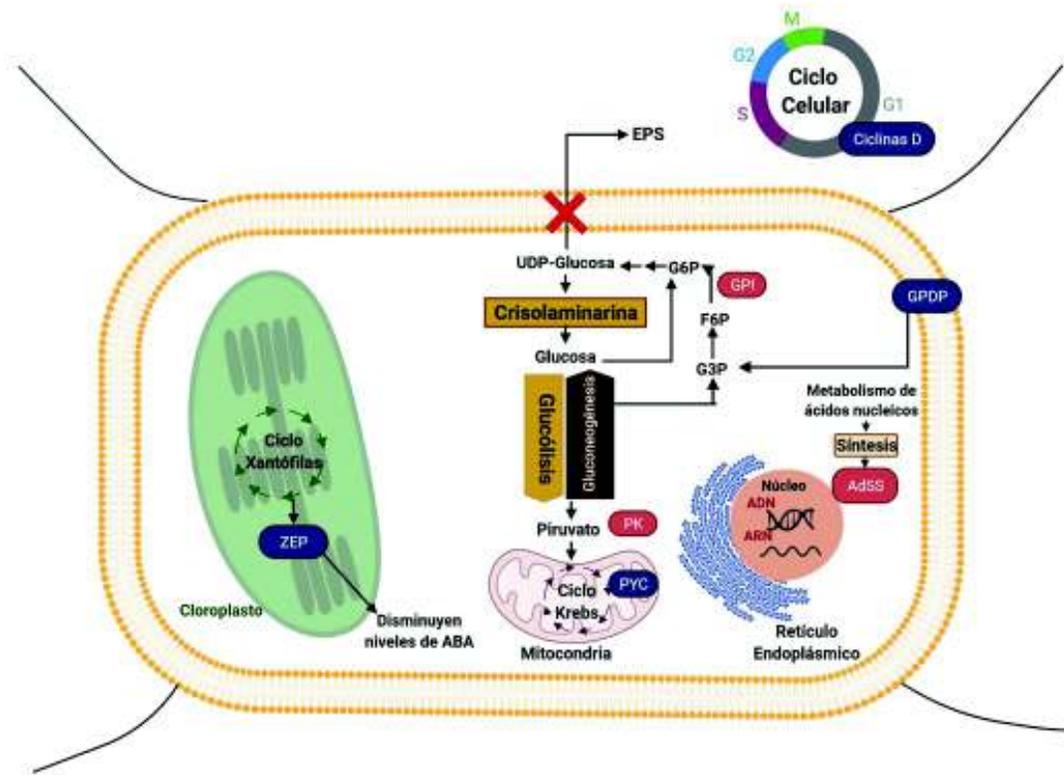
Por otra parte, al igual que bajo limitación por P, se observó bajo condiciones de exceso la inducción de adenilosuccinato sintetasa (AdSS) la cual se encuentra involucrada en la síntesis de purinas (Baynes y Dominiczak, 2009; Stayton *et al.*, 1983). Como se mencionó anteriormente, la síntesis de purinas es alta en células que se preparan para dividirse, la diferencia con las condiciones limitantes y exceso de P en *C. muelleri* pudiera estar relacionada en que en las primeras se está llevando a cabo como consecuencia de un proceso de reciclamiento de bases nitrogenadas y nucleósidos para mantener un aporte equilibrado en células mientras que en el exceso es el resultado de la activación de la ruta 'de novo' de síntesis

de purinas que se lleva a cabo a partir de metabolitos básicos y para dar cumplimiento a los requerimientos celulares (Baynes y Dominiczak, 2009; Stayton *et al.*, 1983).

Otro transcrito que se encontró reprimido en *C. muelleri* bajo exceso de P codificó a zeaxantina epoxidasa (ZEP), la cual como ya se mencionó, esta relacionada con el ciclo de las xantófilas y la síntesis de carotenoides y la acumulación de ABA (Murillo *et al.*, 2019; Sah *et al.*, 2016). Además, la actividad de ZEP está relacionada con el grado de daño fotooxidativo en los cloroplastos derivado también de la disminución de pigmentos como la clorofila y fucoxantina derivados del estrés (Reinhold *et al.*, 2018; Kuczynska *et al.*, 2015). Por lo tanto, se sugiere que bajo estas condiciones en *C. muelleri*, los niveles de pigmentos son suficientes para la protección del aparato fotosintético y no hay aparentemente una condición de estrés que favorezca la señalización por ABA.

En lo que respecta al metabolismo de lípidos, se observó un transcrito reprimido que codifica a glicerolfosfodiéster fosfodiesterasa (GPDP) la cual está involucrada en la degradación de fosfolípidos en el proceso de remodelación de la membrana celular e interviene en la homeostasis celular de P (Alipanah *et al.*, 2018; Mehra *et al.*, 2018; Cheng *et al.*, 2011). En *C. muelleri* bajo estas condiciones al no ser el P un factor limitante, no requiere de la actividad de GPDP para obtener P biodisponible ni ácidos grasos libres para que puedan ser utilizados como precursores para la síntesis de TAG o de energía.

Por otro lado, a diferencia de las condiciones limitantes de P en *C. muelleri*, en exceso de P se observó la inducción de un transcrito que codifica a piruvato cinasa (PK) que se sabe posee un rol clave durante la glucólisis catalizando la conversión de fosfoenolpiruvato a piruvato mientras que el transcrito que codificó a piruvato carboxilasa (PC) que cataliza la conversión de oxaloacetato a piruvato se encontró reprimido y que es clave en el proceso de gluconeogénesis (Garrett *et al.*, 2008; Raven, 2005). La actividad de ambas enzimas es clave ya que definirán el proceso (glucólisis o gluconeogénesis) que se llevará a cabo en la célula dependiendo de su estatus energético (Alipanah *et al.*, 2018; Jitrapakdee y Wallace, 1999). Esto indica que en *C. muelleri* bajo estas condiciones se está llevando a cabo el proceso de glucólisis a diferencia de lo que ocurre en condiciones limitantes en donde es necesaria la gluconeogénesis como estrategia para la obtención de glucosa ante la deficiencia energética.



**Figura 21.** Modelo hipotético de la respuesta transcripcional de *C. muelleri* ante el exceso de P. Cuadros en color rojo: Transcritos inducidos. Cuadros en color azul marino: Transcritos reprimidos. Creado en Biorender: [www.biorender.com](http://www.biorender.com)

#### 6.4.9 Respuesta transcripcional de *C. muelleri* cultivada bajo concentraciones limitantes de P y su relación con la acumulación de lípidos

A pesar de que las respuestas a la limitación de P en las diatomeas son bastante similares entre las especies estudiadas hasta ahora, varios estudios transcriptómicos han demostrado que la expresión de genes clave implicados en el metabolismo de los lípidos dependerá de la especie, las concentraciones de P, las condiciones de cultivo y la duración del experimento (Alipanah *et al.*, 2018; Feng *et al.*, 2015; Shih *et al.*, 2015; Dyhrman *et al.*, 2012).

En relación al metabolismo de lípidos en condiciones limitantes de P en *C. muelleri* se observó la inducción de transcritos involucrados en la remodelación de la membrana celular, entre ellos aquellos que codifican a fosfolipasas tales como fosfolipasa D (PLD), fosfolipasa A1

(PLA1) y fosfolipasa A2 (PLA2) las cuales dentro de sus funciones se encuentran la degradación de fosfolípidos con el fin de proveer Pi a la célula para satisfacer sus requerimientos y proporcionar precursores para la síntesis de triacilgéridos tales como ácido fosfatídico y ácidos grasos libres; esta respuesta ha sido observada en distintas especies tanto de microalgas como de plantas ante la limitación por P (Xu y Miao, 2020; Alipanah *et al.*, 2018; Cruz de Carvalho *et al.*, 2016; Cruz-Ramírez *et al.*, 2006). Otro transcrito inducido en *C. muelleri* que se encuentra involucrado en la remodelación de la membrana celular es el que codifica para sulfoquinovosil transferasa 2 (SQD2) la cual está relacionada con la síntesis de sulfolípidos, los cuales sustituirán bajo estas condiciones a los fosfolípidos que han sido degradados (Van Mooy *et al.*, 2009); esta respuesta es considerada característica de la limitación por P en microalgas y que se ha observado tanto en análisis bioquímicos, transcriptómicos y proteómicos (Alipanah *et al.*, 2018; Shih *et al.*, 2015; Dyhrman *et al.*, 2012; Van Mooy *et al.*, 2009) (Figura 22).

Por otra parte, una de las enzimas consideradas claves en el flujo de carbono es fosfoenolpiruvato carboxicinas (PEPCK), la cual cataliza la conversión reversible de oxaloacetato a fosfoenolpiruvato (PEP) y participa en el conjunto de reacciones que conforman el ciclo de Krebs o ciclo del ácido tricarbóxico (TCA, por sus siglas en inglés) (Liu *et al.*, 2018; Wu *et al.*, 2016). El ciclo del TCA es una ruta metabólica clave que unifica el metabolismo de los carbohidratos, los lípidos y las proteínas, y se ha reportado que la acumulación de TAG bajo estrés por nutrientes es una consecuencia de la reasignación de carbono, principalmente de los intermediarios de esta ruta metabólica y se considera a PEPCK como un punto clave en el proceso de acumulación de lípidos ya que se ha reportado que al reprimirse el gen que codifica para esta enzima, deriva a una deficiencia de fosfoenolpiruvato y acelera la biosíntesis y transporte de citrato de la mitocondria al citosol, dirigiendo el flujo de carbono hacia la síntesis de piruvato, acetil-CoA y ácidos grasos (Luo *et al.*, 2017; Wu *et al.*, 2016; Yang *et al.*, 2016). En *C. muelleri* se identificó un transcrito reprimido que codifica a PEPCK, por lo que el flujo de carbono está dirigiéndose hacia la obtención de precursores necesarios (como acetil Co-A y ácidos grasos) para la síntesis de lípidos (Figura 22).

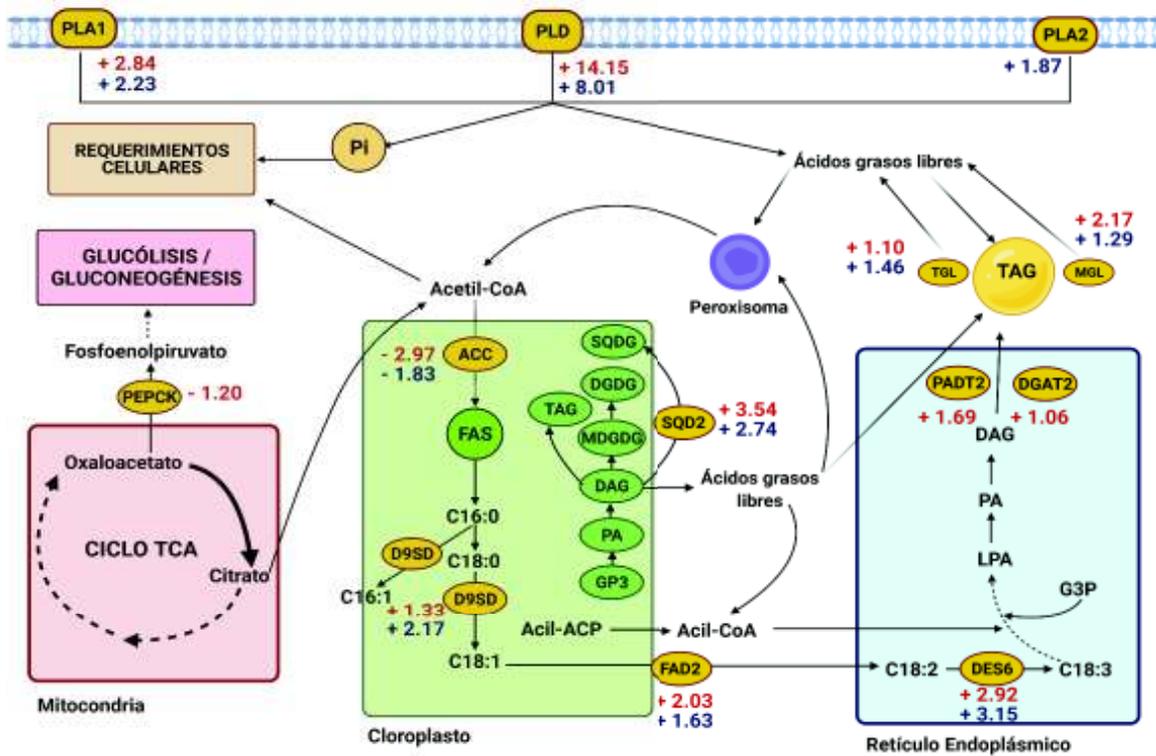
En lo que respecta a la síntesis de ácidos grasos, la mayoría de los transcritos que codificaron a enzimas involucradas en este proceso se encontraron reprimidos o no se expresaron diferencialmente en *C. muelleri*. Tal es el caso de acetil-CoA carboxilasa (ACC),

considerada uno de los puntos clave en la síntesis de ácidos grasos, en el que se encontraron transcritos reprimidos que codificaron para esta enzima (Sasaki y Nagano, 2004; Page *et al.*, 1994). Esto pudiera deberse a que se ha reportado que la actividad de ACC es regulada por AMPK, un complejo enzimático considerado un regulador maestro del metabolismo, el cual puede activar o inactivar esta enzima en función de los niveles de citrato, palmitoil-CoA y la relación ATP/AMP en la célula. Cuando los niveles de ATP son bajos en la célula, AMPK fosforila a ACC inhibiéndola hasta que el contenido energético vuelva a ser favorable (Berg *et al.*, 2012). De tal manera que se sugiere en función de estos resultados que para *C. muelleri* bajo estas condiciones, la acumulación de lípidos no puede ser atribuida directamente a la biosíntesis de ácidos grasos sino que estos se están derivando de otros procesos como la degradación de fosfolípidos (Figura 22).

Por otro lado, se identificaron transcritos inducidos que codificaron a varias desaturasas, entre ellas, delta 9-estearoil desaturasa (D9SD), delta 12-desaturasa (FAD2) y delta 6 desaturasa (DES6) (Figura 22). D9SD cataliza la desaturación del ácido estearico al ácido oleico e interviene también en la desaturación del ácido palmítico al palmitoleico, FAD2 es responsable de la desaturación del ácido oleico a linoleico y DES6 está involucrada en la síntesis de los ácidos grasos alfa linolénico (ALA) y gamma linolénico (GLA) y es considerado un punto clave que definirá la síntesis de ácidos grasos omega 3 y 6 en diatomeas y la producción de EPA (Dar *et al.*, 2017). En *C. muelleri*, la inducción de estos transcritos coincide moderadamente con lo observado en el perfil de ácidos grasos obtenido en este estudio ya que en primer lugar la síntesis de ácidos grasos monoinsaturados se ve favorecida sobre la síntesis de ácidos grasos poliinsaturados en la limitación por P y estas desaturasas están involucradas principalmente en la síntesis de ácidos grasos de 18 carbonos con más de una insaturación, los cuales en diatomeas en general se encuentran en cantidades ínfimas. Estos resultados pudieran sugerir en primer lugar que aún bajo estas condiciones, no ha cesado la síntesis de ácidos grasos insaturados, que el tiempo de vida media de estos transcritos puede ser corta, que pueden presentarse modificaciones postranscripcionales o postransduccionales que sean reguladas por la limitación de P y además se ha reportado que algunos de los genes que codifican para estas desaturasas también están involucrados en la respuesta a factores abióticos (Anne-Marie *et al.*, 2020; Vuppaladadiyam *et al.*, 2018; Yang *et al.*, 2018).

Finalmente, en relación a la síntesis de TAG, se encontró en *C. muelleri* dos transcritos inducidos que codificaron a las enzimas fosfolípido:diacilglicerol aciltransferasa 2 (PDAT2) y diacilglicerol aciltransferasa (DGAT2) las cuales están involucradas en el paso final del ensamblaje de TAG y la inducción de ambos sugiere que bajo la limitación por P se está efectuando la acumulación de lípidos (Figura 22) (Yang *et al.*, 2018; Mülhroth *et al.*, 2017; Cruz Carvalho *et al.*, 2016). Cabe destacar que entre especies de diatomeas cercanas filogenéticamente, esta respuesta puede variar ante la limitación de P ya que se ha observado en otros estudios que estos genes pueden encontrarse reprimidos o sin mostrar cambios en su expresión; además, también se ha reportado que la inducción de PDAT2 y DGAT2 no es exclusiva de la limitación de este nutriente lo que indica que entre especies hay distintos mecanismos de respuesta a un mismo factor (Alipanah *et al.*, 2015; Levitan *et al.*, 2015; Shih *et al.*, 2015).

Otro aspecto interesante que se encontró en *C. muelleri* ante la limitación de P y su relación con la acumulación de lípidos, fue la inducción de dos transcritos que codificaron a monoacilglicerol lipasa (MGL) y a una TAG lipasa (TGL), las cuales están involucradas en la degradación de TAG (Figura 22). En estudios transcriptómicos en plantas como *Arabidopsis* y benjuí, se observó la inducción de *MGL* como una estrategia de estos organismos ante la deficiencia de carbono y energía (Wu *et al.*, 2020; Kim *et al.*, 2016). Por otro lado, en *C. affinis* bajo limitación por P, *TGL* no mostró cambios en su nivel de expresión (Shih *et al.*, 2015); esto genera evidencia de que a pesar que en dicho estudio se consideró una concentración menor de P (3  $\mu$ M) a la del presente estudio, las respuestas entre especies cercanas filogenéticamente pueden variar. Se ha reportado que la movilización y degradación de los TAG en condiciones de estrés es necesaria para la liberación de ácidos grasos y glicerol para que estos puedan a su vez ser degradados en los peroxisomas para la producción de energía o utilizados como sustratos para síntesis de lípidos de membrana (Warakanont *et al.*, 2019). De ahí que se sugiere que en *C. muelleri* bajo limitación por P, los TAG que se están produciendo a su vez son conducidos a un proceso catabólico con el fin de proveer precursores (glicerol, ácidos grasos, acetil Co-A) para otros procesos y la producción de energía para el correcto funcionamiento de la célula y garantizar su supervivencia.



**Figura 22.** Modelo hipotético de la acumulación de lípidos en *C. muelleri* ante la limitación por P. Valores en color rojo: log<sub>2</sub>FC en tratamiento con concentración de P de 7 μM. Valores en color azul: log<sub>2</sub>FC en tratamiento con concentración de P de 18 μM. +: Transcritos inducidos. —: Transcritos reprimidos. Creado en Biorender: [www.biorender.com](http://www.biorender.com)

### 6.5 Análisis *in silico* de proteínas transportadoras de P de *C. muelleri*.

El transporte de Pi a través de la membrana celular es esencial para el funcionamiento celular. La inducción de genes que codifican a proteínas transportadoras de P es una respuesta común ante la limitación de P en varios organismos, desde bacterias hasta seres humanos. En estudios transcriptómicos en diatomeas cultivadas bajo limitación y ausencia de este nutriente, se ha observado que la inducción de los genes que codifican para este tipo de proteínas aumenta considerablemente (Shih *et al.*, 2015; Yang *et al.*, 2014; Dyhrman *et al.*, 2012). Se seleccionaron dos genes que codifican a proteínas transportadoras de P que se vieron inducidos, de acuerdo al análisis de RNA-seq, en el transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P: NPT2B (proteína de transporte de fosfato dependiente de sodio 2B) y VTC4 (proteína chaperona de transporte vacuolar 4). Las secuencias nucleotídicas de estos genes fueron identificadas y extraídas del transcriptoma para después ser traducidas a secuencias aminoacídicas para su posterior análisis. En la Tabla 10 se muestran varios parámetros fisicoquímicos estimados por la herramienta ProtParam de ExPASy para cada una de las secuencias aminoacídicas y en la Figura 23 la composición de aminoácidos correspondiente a cada secuencia.

**Tabla 10. Propiedades fisicoquímicas de las secuencias aminoacídicas de las proteínas transportadoras de P NPT2B y VTC4 de *C. muelleri***

	NPT2B	VTC4
Número de aminoácidos	474	735
Peso Molecular (Daltons)	50778.50	83513.9
Punto isoeléctrico (pI)	6.33	6.54
Índice de inestabilidad	28.03	38.09
Índice alifático	112.68	86.87
Hidropatía GRAVY	0.522	-0.386

<b>A</b> Amino acid composition:			<b>B</b> Amino acid composition:		
Ala (A)	36	7.6%	Ala (A)	47	6.4%
Arg (R)	17	3.6%	Arg (R)	36	4.9%
Asn (N)	13	2.7%	Asn (N)	29	3.9%
Asp (D)	17	3.6%	Asp (D)	53	7.2%
Cys (C)	9	1.9%	Cys (C)	8	1.1%
Gln (Q)	11	2.3%	Gln (Q)	23	3.1%
Glu (E)	21	4.4%	Glu (E)	44	6.0%
Gly (G)	47	9.9%	Gly (G)	37	5.0%
His (H)	6	1.3%	His (H)	22	3.0%
Ile (I)	41	8.6%	Ile (I)	47	6.4%
Leu (L)	54	11.4%	Leu (L)	66	9.0%
Lys (K)	19	4.0%	Lys (K)	56	7.6%
Met (M)	13	2.7%	Met (M)	12	1.6%
Phe (F)	23	4.9%	Phe (F)	27	3.7%
Pro (P)	17	3.6%	Pro (P)	36	4.9%
Ser (S)	34	7.2%	Ser (S)	52	7.1%
Thr (T)	37	7.8%	Thr (T)	46	6.3%
Trp (W)	5	1.1%	Trp (W)	15	2.0%
Tyr (Y)	10	2.1%	Tyr (Y)	27	3.7%
Val (V)	44	9.3%	Val (V)	52	7.1%
Pyl (O)	0	0.0%	Pyl (O)	0	0.0%
Sec (U)	0	0.0%	Sec (U)	0	0.0%
(B)	0	0.0%	(B)	0	0.0%
(Z)	0	0.0%	(Z)	0	0.0%
(X)	0	0.0%	(X)	0	0.0%

**Figura 23.** Composición aminoacídica de las secuencias de las proteínas transportadoras de P de *C. muelleri*. A: NPT2B. B: VTC4.

Los resultados del análisis realizado por la herramienta ProtParam de ExPASy pueden brindar información de la naturaleza de estas secuencias aminoacídicas; se puede observar que, por el tamaño como el peso molecular de ambas secuencias aminoacídicas analizadas de *C. muelleri* correspondientes a NPT2B y VTC4 pueden clasificarse como polipéptidos (>100 aa; PM  $\geq$  5000 Daltons) (Berthold, 2014) (Tabla 10).

Un parámetro importante estimado por la herramienta ProtParam, es la hidropatía GRAVY (Grand Averigüe of Hidropatía) la cual determina la hidrofiliidad o hidrofobicidad de una proteína y se calcula mediante la suma de los valores de hidropatía de todos los aminoácidos presentes en la secuencia aminoacídica y se divide entre la longitud de dicha secuencia; sus valores varían entre -2 y 2 para la mayoría de las proteínas en donde los valores positivos están

relacionados con comportamientos hidrofóbicos (Walker *et al.*, 2005). NPT2B mostró valores positivos de hidropatía GRAVY (0.522) por lo que pudiera sugerir que es una proteína con cierta naturaleza hidrófoba mientras que VTC4 (-0.386) pudiera considerarse como hidrofílica al obtenerse valores negativos en el parámetro de hidropatía de GRAVY (Tabla 10).

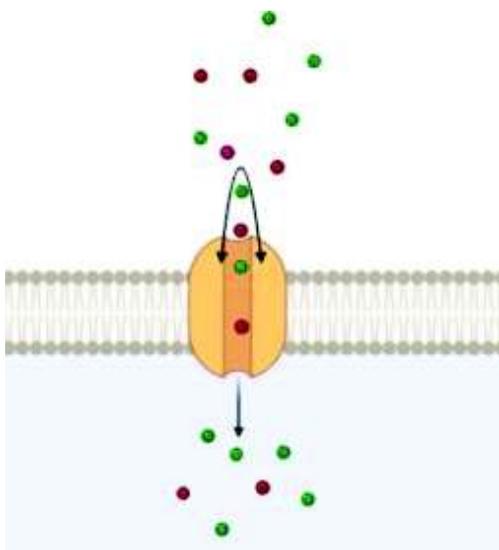
### 6.5.1 NPT2B: Proteína de transporte de fosfato dependiente de sodio 2B

Desde bacterias hasta mamíferos se han identificado proteínas transportadoras de P dependientes de sodio pertenecientes a las familias SLC20 y SLC34 (Forster *et al.*, 2013; Li *et al.*, 2006). Este tipo de proteínas son simportadoras, es decir, transfieren P y sodio a través de la membrana celular hacia su interior (Lin *et al.*, 2016). Por su parte, la familia de proteínas SLC34 es una clase única de proteínas de transporte de membrana que no comparte una homología obvia con otras familias de transportadores de solutos, ni siquiera a nivel bacteriano; consiste en tres miembros, NaPi-IIIa (SLC34A1), NaPi-IIIb (SLC34A2) y NaPi-IIIc (SLC34A3), que median la translocación del fosfato inorgánico divalente ( $\text{HPO}_4^{2-}$ ) junto con dos (NaPi-IIIc) o tres iones de sodio (NaPi-IIIa y NaPi-IIIb), respectivamente (Forster *et al.*, 2013).

Una respuesta común ante la deficiencia de P es la síntesis de proteínas transportadoras de P (Dyhrman *et al.*, 2016; Lin *et al.*, 2016); se han identificado en estudios transcriptómicos de varias microalgas con limitación y ausencia de P, genes homólogos que codifican a proteínas transportadoras de P dependientes de sodio de la familia SLC20 y SLC34, siendo de ésta última la isoforma NaPi-IIIb (SLC34A2) o NPT2B, a la que se le ha atribuido una función importante en la asimilación de P en diatomeas bajo la limitación de este nutriente (Shih *et al.*, 2015; Yang *et al.*, 2014; Dyhrman *et al.*, 2012) (Figura 24).

Se realizó un BLASTp con la secuencia aminoacídica de NPT2B de *C. muelleri* contra la base de datos del NCBI dando como resultado una similitud alta con proteínas co-transportadoras de sodio y fósforo pertenecientes a la familia SLC34 de varias especies de diatomeas. Las especies que presentaron una mayor homología con la secuencia aminoacídica de NPT2B de *C. muelleri* fueron *C. affinis* (62.42 %), *P. tricorutum* (57.20 %) y *T. pseudonana* (56.62%). (Tabla 11). En la Tabla 11, además del porcentaje de homología con otras especies de diatomeas, se incluye el valor de e y el número de acceso de secuencias aminoacídicas para

cada especie. El valor de e (o E-value, en inglés) hace referencia al número esperado de hits aleatorios para un puntaje de alineación dado; entre menor sea su valor o más cercano a cero, mayor es la significancia del resultado; este parámetro toma en consideración el tamaño de la secuencia proporcionada, ya que a menor tamaño de secuencia aumenta la posibilidad de encontrar coincidencias con las secuencias analizadas disminuyendo la significancia del resultado. En este estudio el valor de e fue bajo lo que da certeza de la significancia del análisis.



**Figura 24.** NPT2B: Proteína de transporte de fosfato dependiente de sodio 2B. Creada en Biorender: <https://biorender.com/>

**Tabla 11. Comparación del porcentaje de identidad de la secuencia aminoacídica NPT2B de *C. muelleri* con las secuencias correspondientes a otras diatomeas.**

Organismo	Valor de e	Porcentaje de identidad (%)	Número de acceso
<i>Chaetoceros affinis</i>	4e-159	62.42	AKI88062.1
<i>Phaeodactylum tricornutum</i>	1e-166	57.20	XP_002181914.1
<i>Thalassiosira pseudonana</i>	6e-140	56.62	XP_002292964.1
<i>Fragilariopsis cylindrus</i>	1e-134	56.56	OEU13556.1
<i>Thalassiosira oceanica</i>	9e-136	56.26	EJK69182.1
<i>Fistulifera solaris</i>	9e-137	53.50	GAX18419.1

La herramienta de BLAST se encuentra relacionada con otra herramienta conocida como CD search (Conserved Domain search) mediante la cual se identificó en la secuencia aminoacídica de NPT2B de *C. muelleri*, el dominio 2a58 que hace referencia a la familia de proteínas de transporte y unión de fósforo y sodio (Figura 25).

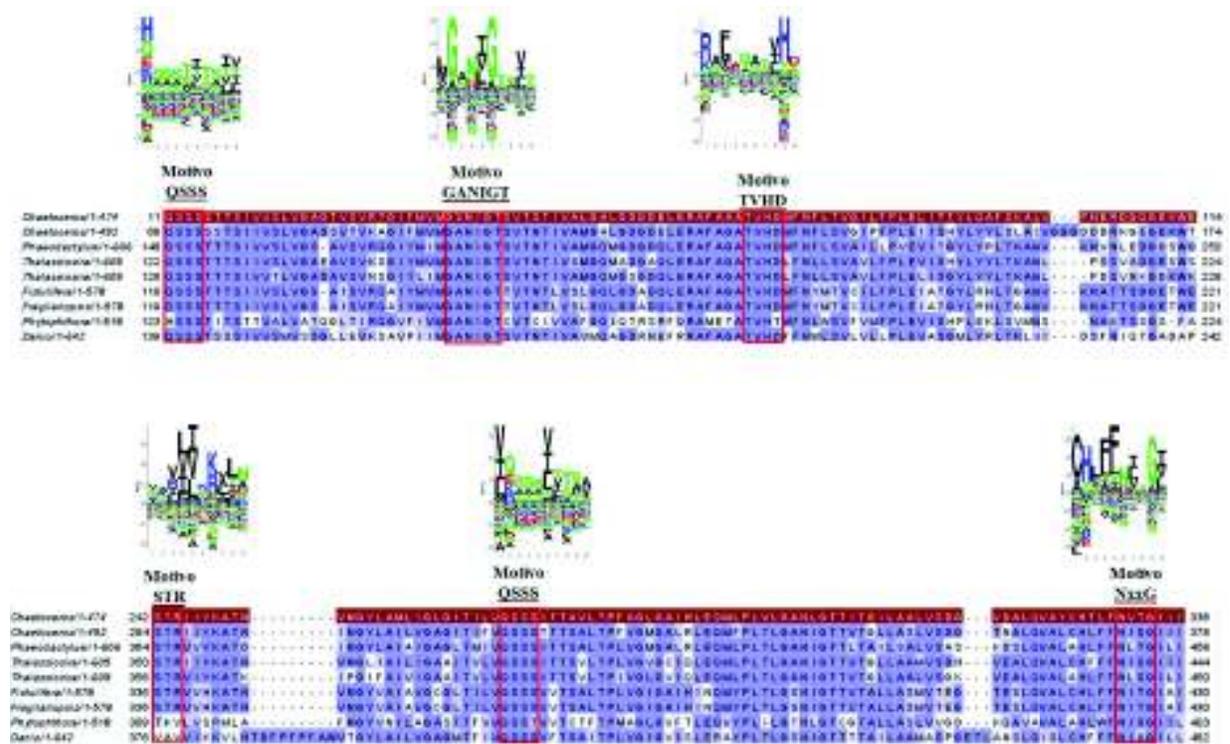


**Figura 25.** Dominio conservado de la familia de los cotransportadores de sodio y fósforo en la secuencia aminoacídica de NPT2B de *C. muelleri*.

Por otro lado, la secuencia aminoacídica de NPT2B de *C. muelleri* fue contrastada con secuencias similares pertenecientes a diferentes especies en un alineamiento múltiple de secuencias y se identificaron algunos motivos mediante la herramienta PROSITE de Expasy (Figura 26). De acuerdo a Fenollar-Ferrer y Forrest (2019), el motivo “QSSS” es característico de la familia de proteínas SLC34, se presenta dos veces en la secuencia aminoacídica de estas proteínas y están involucradas en la unión de sodio (Na<sub>2</sub> y Na<sub>3</sub>) y P, lo cual puede observarse en todas las secuencias consideradas en el análisis lo que sugiere la certeza de la función de estas proteínas en el transporte simporte tanto de P como de sodio.

Por otra parte, se detectó el motivo GANIGT, el cual representa un sitio de miristoilación en el extremo N-terminal, altamente conservado en todas las secuencias a excepción de la secuencia aminoacídica de *C. muelleri* que presentó una sustitución de alanina por serina; esta sustitución aparentemente no representa una modificación sustancial en la función de la proteína ya que puede darse el caso de que se puedan presentar sustituciones en la secuencia aminoacídica siempre y cuando los residuos que se presenten no se encuentren cargados y no se presenten residuos hidrofóbicos de gran tamaño (Udenwobele *et al.*, 2017; McIlhinney, 1998).

Otros motivos identificados en la secuencia aminoacídica de *C. muelleri* y que se observó un cierto grado de conservación con el resto de secuencias con las que se comparó fueron TVHD, STR y NxxG. El motivo TVHD representa un sitio de fosforilación de caseína cinasa II (T<sup>71</sup>); la caseína cinasa II (CKII) es una proteína serina-treonina cinasa ubicua y altamente conservada en todos los eucariotas y participa en la regulación de procesos celulares esenciales (Ahmed *et al.*, 2002). Existe evidencia de que CKII posee un rol clave en la regulación de la asimilación y homeostasis de P mediante su interacción con proteínas transportadoras de este nutriente (Chen *et al.*, 2015).

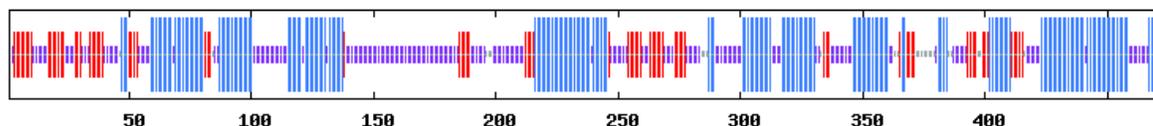


**Figura 26.** Alineamiento de las secuencias de aminoácidos de NPT2B, mostrando 5 motivos significativamente conservados. La secuencia de NPT2B de *C. muelleri* está resaltada en rojo.

En lo que se refiere al motivo STR, hace alusión a un sitio de fosforilación por proteínas cinasas de tipo C (PKC), las cuales suelen fosforilar residuos de serina o treonina en secuencias básicas; se ha reportado que la activación de PKC inhibe la actividad de enzimas transportadoras de sodio y fósforo (Wagner *et al.*, 1996; Hayes *et al.*, 1995). Finalmente, el motivo NxxG

representa un sitio de N-glicosilación, en donde N es un residuo de asparagina, en donde se pueden suscitar modificaciones postraduccionales mediante la unión de un oligosacárido de complejidad variable (Helenius, 1994). La incorporación de oligosacáridos en proteínas cotransportadoras de sodio y fósforo del tipo II es de vital importancia para su maduración y se ha observado en un estudio en donde se evitó la glicosilación de esta proteína que, aunque no se afectó su función, sí disminuyó la tasa de absorción y los autores sugirieron que esto podía deberse en parte, a una distribución reducida de la proteína en la superficie celular (Hayes *et al.*, 1994).

Los resultados de la predicción de la estructura secundaria de la proteína NPT2B de *C. muelleri*, así como el porcentaje de alfa hélices y láminas beta se muestran en la Figura 27 y la Tabla 12 respectivamente. Se obtuvo un valor de 186 aminoácidos con plegamientos alfa hélice que representa el 39.24 % de la secuencia aminoacídica, seguido de un 16.67 % correspondiente a láminas beta (79 residuos), 185 residuos (39.03 %) para enrollamientos aleatorios y un 5.06 % de ambigüedades (24 residuos) del total de 474 aminoácidos que conforman la secuencia aminoacídica de NPT2B de *C. muelleri*. Se observó que NPT2B en su mayoría está compuesta por enrollamientos aleatorios, esto significa que posee plegamientos irregulares y por tanto no poseen una estructura secundaria definida, esto se debe a que en esta conformación los residuos que la componen interactúan con otros sin llegar a interacciones estables. Por otro lado, los enrollamientos aleatorios son considerados como “puentes de unión” entre los elementos principales de la estructura secundaria (hélices alfa y láminas beta) (Khrustalev *et al.*, 2013; Dill y Shortle, 1991).



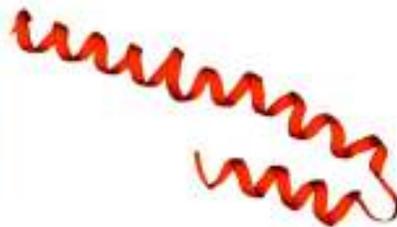
**Figura 27.** Representación gráfica de la predicción de la estructura secundaria de la proteína NPT2B de *C. muelleri* realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.

**Tabla 12.** Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de NPT2B de *C. muelleri* obtenidos por el servidor NPS@.

Tipo	Número de aminoácidos	Porcentaje (%)
Alpha helix (Hh)	186	39.24
310 helix (Gg)	0	0.00
Pi helix (Ii)	0	0.00
Beta bridge (Bb)	0	0.00
Extended strand (Ee)	79	16.67
Beta turn (Tt)	0	0.00
Bend region (Ss)	0	0.00
Random coil (Cc)	185	39.03
Ambiguous states (?)	24	5.06
Other states	0	0.00

Finalmente, la Figura 28 muestra el modelo propuesto de la estructura terciaria de la proteína NPT2B de *C. muelleri*. Se emplearon plantillas procedentes de modelos de proteínas de genes ortólogos de la base de datos de SWISS MODEL, encontrándose para esta proteína un porcentaje bajo de identidad (15.38 %) con el modelo de referencia y no concordó con la secuencia identificada. Esto puede deberse a que la predicción de la estructura terciaria, al basarse en algoritmos de homología, requiere de proteínas cuyas estructuras ya estén experimentalmente determinadas y que tengan un alto porcentaje de homología a la proteína analizada, por lo que en este caso pudiera verse limitado al ser escasa la información para diatomeas.

Global Quality Estimate		Local Quality Estimate		Comparison
Template	Seq Identity	Coverage	Description	
4v3p:12.A	15.38%	<input type="text"/>	40S ribosomal protein S13-1 The molecular structure of the left-handed supra-molecular helix of eukaryotic polyribosomes	

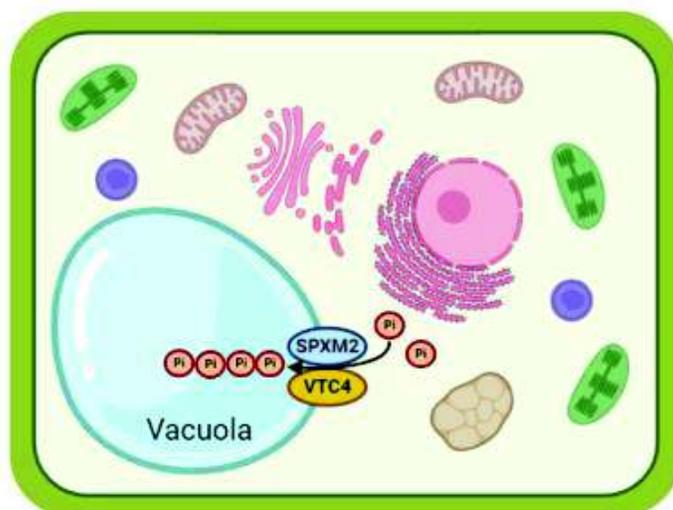


**Figura 28.** Modelo de la estructura terciaria teórica de la proteína cotransportadora de sodio y fósforo 2B (NPT2B) de *C. muelleri* obtenida a través del servidor SWISS MODEL.

#### 6.5.2 VTC4: Chaperona de transporte vacuolar 4

En procariotas, hongos y microalgas, los Poli-P, una cadena lineal de tres a miles de residuos de Pi unidos por enlaces de fosfoanhídrido de alta energía, son la principal forma de almacenamiento de P dentro de las vacuolas y se estima que, ante condiciones favorables de Pi en el medio ambiente, las vacuolas pueden almacenar entre un 70 y un 95 % del Pi intracelular y su concentración dentro de estos organelos fluctuará de acuerdo a la disponibilidad de P extracelular (Yang *et al.*, 2017; Bielecki, 1973). La síntesis de Poli-P ha sido ampliamente caracterizada en *S. cerevisiae* y se ha reportado que está regulada por el complejo chaperonas de transporte vacuolar (VTC), el cual está integrado por cinco subunidades (VTC1, VTC2, VTC3, VTC4 Y VTC5) y en un principio dentro de sus funciones principales se definieron su contribución en la fusión de las membranas de las vacuolas, la estabilidad de la V-ATPasa, el tráfico de membranas y la microautofagia; tiempo después se confirmó su rol en la síntesis de Poli-P (Uttenweiler *et al.*, 2007; Müller *et al.*, 2003).

Después de realizarse estudios de cristalografía de rayos X, se descubrió que la subunidad del complejo VTC responsable de catalizar la síntesis de poli-P utilizando ATP como sustrato es VTC4, mientras que el resto de subunidades fungen como accesorios involucradas en la regulación de la síntesis y transferencia de poliP (Gerasimaité *et al.*, 2014; Hothorn *et al.*, 2009). Cabe destacar que las proteínas homólogas del complejo VTC se encuentran sólo en protistas, hongos y microalgas y, están ausentes en plantas y animales (Secco *et al.*, 2012). En este estudio y de acuerdo al análisis de RNA-seq, un transcrito que codificó a VTC4 se encontró inducido en condiciones limitantes de P; esto es congruente con lo observado en otros estudios transcriptómicos en diatomeas y *S. cerevisiae* bajo limitación por P en donde el gen que codificó para VTC4 se vio inducido (Zhang *et al.*, 2016; Dyhrman *et al.*, 2012; Ogawa *et al.*, 2000). Dyhrman *et al.* (2012) sugieren que la inducción de este gen en estas condiciones puede estar relacionada a la distribución y uso de P disponible para la formación de Poli-P (Figura 29).



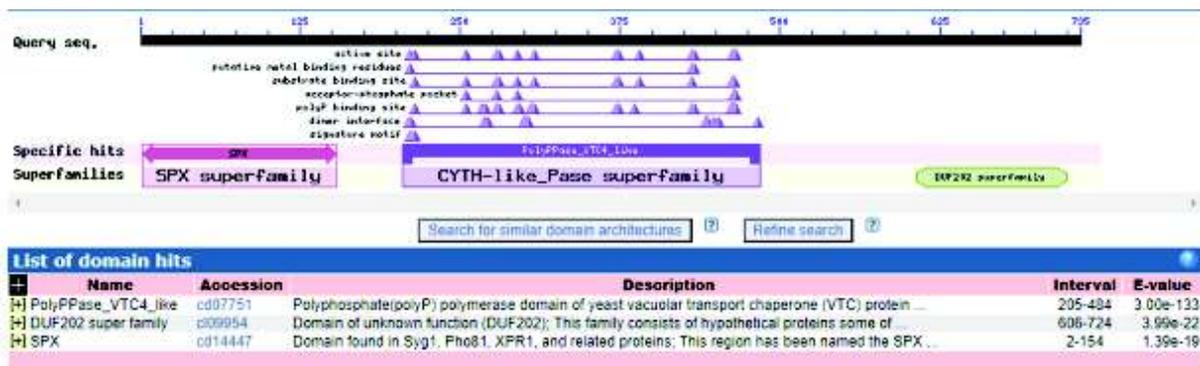
**Figura 29.** VTC4: Chaperona de transporte vacuolar 4. Creada en Biorender: <https://biorender.com/>

La secuencia nucleotídica de VTC4 de *C. muelleri* fue extraída del transcriptoma y se tradujo para obtener su secuencia aminoacídica para después realizar un BLASTp contra la base de datos del NCBI dando como resultado una similitud alta con proteínas VTC4 de varias especies de diatomeas. Las especies que presentaron una mayor homología con la secuencia aminoacídica de VTC4 de *C. muelleri* fueron *T. pseudonana* (69.08 %), *T. oceanica* (63.62 %) y *P. tricorutum* (63.45 %) (Tabla 13). En la Tabla 13, además del porcentaje de homología con otras especies de diatomeas, se incluye el valor de e y el número de acceso de secuencias aminoacídicas para cada especie.

**Tabla 13.** Comparación del porcentaje de identidad de la secuencia aminoacídica VTC4 de *C. muelleri* con las secuencias correspondientes a otras diatomeas.

Organismo	Valor de e	Porcentaje de identidad (%)	Número de acceso
<i>Thalassiosira pseudonana</i>	0.0	69.08	XP_002295322.1
<i>Thalassiosira oceanica</i>	0.0	63.62	EJK75212.1
<i>Phaeodactylum tricorutum</i>	0.0	63.45	XP_002184903.1
<i>Fistulifera solaris</i>	0.0	63.16	GAX11298.1
<i>Fragilariopsis cylindrus</i>	0.0	62.70	OEU19032.1
<i>Pseudo-nitzschia multistriata</i>	0.0	60.87	VEU38640.1

Después, mediante la herramienta CD search (Conserved Domain search) se identificaron tres dominios en la secuencia aminoacídica de VTC4 de *C. muelleri* correspondientes a las superfamilias de SPX, CYTH-like\_Pase y DUF202 (Figura 25). Las proteínas que contienen el dominio SPX están involucradas en la homeostasis, transporte y adaptación a condiciones limitantes de P (Jung *et al.*, 2018). Por su parte, el dominio CYTH-like\_Pase es característico de proteínas que hidrolizan sustratos que contienen trifosfato y requieren de cationes metálicos como cofactores; esta superfamilia de dominios también contiene ARN trifosfatasa, polimerasas de Poli-P asociadas a la membrana, tripolifosfatasa, nucleósidos trifosfatasa, nucleósidos tetrafosfatasa y otras proteínas con funciones desconocidas (Lu *et al.*, 2020). Finalmente, para el dominio DUF202 no es del todo clara su función, sin embargo, este dominio contiene regiones transmembranales predichas, y muchos miembros de la familia son supuestas proteínas de membrana (Rooney *et al.*, 2011). Estos resultados brindan certeza sobre las presuntas características funcionales de una proteína VTC4 en la secuencia aminoacídica de *C. muelleri*.

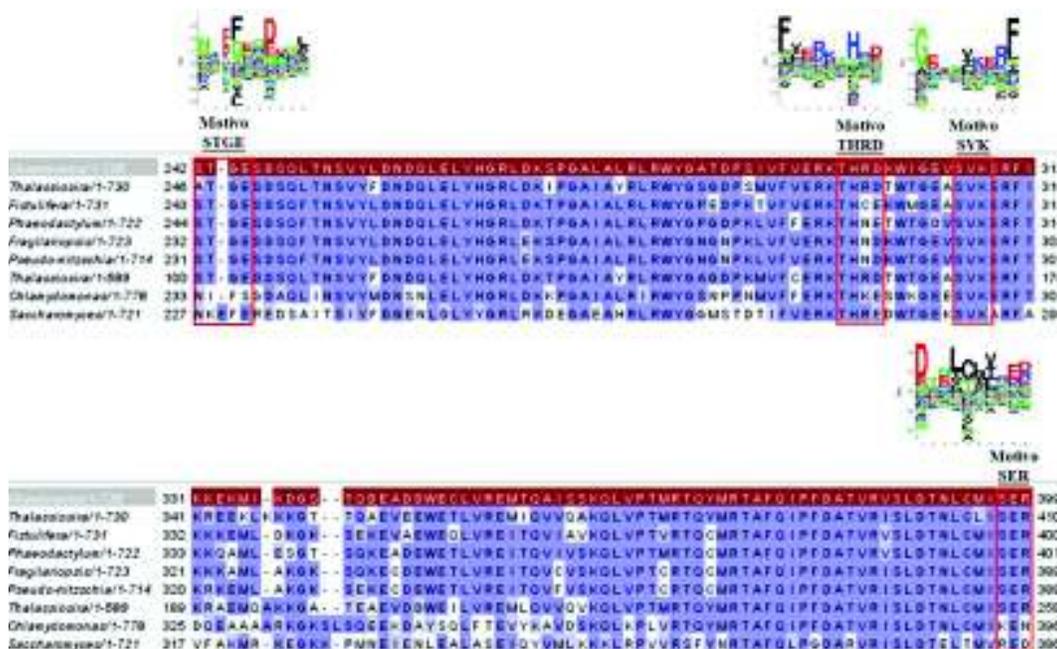


**Figura 30.** Dominios conservados en la secuencia aminoacídica de VTC4 de *C. muelleri*.

Posteriormente, la secuencia aminoacídica de VTC4 de *C. muelleri* fue contrastada con secuencias similares pertenecientes a diferentes especies de diatomeas y con una secuencia aminoacídica de *S. cerevisiae* en un alineamiento múltiple de secuencias y se identificaron algunos motivos mediante la herramienta PROSITE de ExPASy (Figura 31). Se detectaron dos motivos que representan sitios de fosforilación por caseína cinasa II (CKII) ('STGE' y 'THRD')

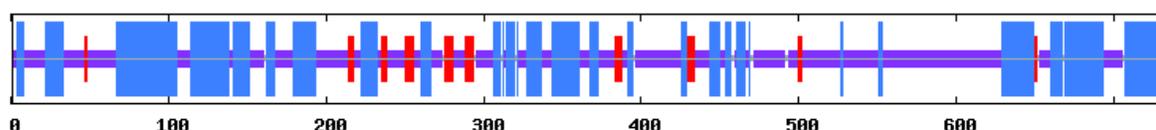
y dos sitios de fosforilación por proteína cinsana C (PKC) ('SVK' y 'SER'). Cabe mencionar que el motivo 'SVK' se encontró altamente conservado en todas las secuencias consideradas para el análisis, mientras que los motivos 'STGE' y 'SER' sólo se encontraron conservados entre las especies de diatomeas.

Por lo que se refiere al motivo 'THRD', a pesar de que se observaron algunas sustituciones de residuos entre las secuencias consideradas en el análisis, esto no afectó en ninguno de los casos la identificación del sitio de fosforilación por caseína cinsana II, ya que su aminoácido diana es treonina y en todos los casos se observó conservado. Sin embargo, la presencia de lisina en el extremo N-terminal del sitio de fosforilación por CKII puede reducir la tasa de fosforilación al ser un aminoácido básico (Pinna, 1990). Hasta donde se sabe, no existen estudios específicos sobre el efecto de fosforilación o ausencia de ella en estos sitios de fosforilación en VTC4, sin embargo, se sabe que CKII está involucrada en la regulación de la asimilación y homeostasis de P mediante su interacción con proteínas transportadoras de este nutriente, mientras que PKC dentro de sus funciones se encuentra la modulación de eventos de la estructura de membrana (Chen *et al.*, 2015; Newton, 1995).



**Figura 31.** Alineamiento de las secuencias de aminoácidos de VTC4, mostrando 4 motivos significativamente conservados. La secuencia de VTC4 de *C. muelleri* está resaltada en rojo.

Los resultados de la predicción de la estructura secundaria de la proteína VTC4 de *C. muelleri*, así como el porcentaje de alfa hélices y láminas beta se muestran en la Figura 32 y la Tabla 14 respectivamente. Se obtuvo un valor de 287 aminoácidos con plegamientos alfa hélice que representa el 39.05 % de la secuencia aminoacídica, seguido de un 5.85 % correspondiente a láminas beta (43 residuos), 390 residuos (53.06 %) para enrollamientos aleatorios y un 2.04 % de ambigüedades (15 residuos) del total de 735 aminoácidos que conforman la secuencia aminoacídica de VTC4 de *C. muelleri*. Al igual que NPT2B, VTC4 mostró un porcentaje alto de enrollamientos aleatorios, esto sugiere que en su mayoría posee plegamientos irregulares (Dill y Shortle, 1991).



**Figura 32.** Representación gráfica de la predicción de la estructura secundaria de la proteína VTC4 de *C. muelleri* realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.

**Tabla 14.** Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de VTC4 de *C. muelleri* obtenidos por el servidor NPS@.

Tipo	Número de aminoácidos	Porcentaje (%)
Alpha helix (Hh)	287	39.05
310 helix (Gg)	0	0.00
Pi helix (Ii)	0	0.00
Beta bridge (Bb)	0	0.00
Extended strand (Ee)	43	5.85
Beta turn (Tt)	0	0.00
Bend region (Ss)	0	0.00
Random coil (Cc)	390	53.06
Ambiguous states (?)	15	2.04
Other states	0	0.00

Por último, en la Figura 33 se muestra el modelo propuesto de la estructura terciaria de la proteína VTC4 de *C. muelleri*. Se emplearon plantillas procedentes de modelos de proteínas de genes ortólogos de la base de datos de SWISS MODEL, encontrándose para esta proteína un porcentaje de identidad de 35.59 % con el modelo de referencia y concordó con la secuencia

identificada, sugiriendo que hay un cierto grado de conservación y homología con organismos ampliamente estudiados.



**Figura 33.** Modelo de la estructura terciaria teórica de la proteína chaperona de transporte vacuolar 4 (VTC4) de *C. muelleri* obtenida a través del servidor SWISS MODEL.

## 6.6 Análisis *in silico* de proteínas involucradas en el metabolismo de lípidos y ácidos grasos en *C. muelleri*.

La identificación de las enzimas críticas para la síntesis de los ácidos grasos y TAG y el conocimiento profundo de su regulación son la base para mejorar su producción en las microalgas (Ran *et al.*, 2019). A pesar de que los conocimientos actuales son escasos, las ciencias ómicas tales como la transcriptómica, la proteómica y la metabolómica, han permitido comprender los cambios globales en el metabolismo durante la acumulación de TAG bajo la limitación de nutrientes en diversas especies de microalgas, lo que podría facilitar el desarrollo de cepas mediante ingeniería genética (Kong *et al.*, 2018).

Se seleccionaron dos genes que codifican a enzimas clave en la síntesis de ácidos grasos que se vieron expresados diferencialmente, de acuerdo al análisis de RNA-seq, en el transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P: ACC (Acetil CoA Carboxilasa) y FAD2 (Delta 12-desaturasa). Las secuencias nucleotídicas de estos genes fueron identificadas y extraídas del transcriptoma de *C. muelleri* para después ser traducidas a secuencias aminoacídicas para su posterior análisis. En la Tabla 15 se muestran varios parámetros fisicoquímicos estimados por la herramienta ProtParam de Expasy para cada una de las secuencias aminoacídicas y en la Figura 34 la composición de aminoácidos correspondiente a cada secuencia.

**Tabla 15. Propiedades fisicoquímicas de las secuencias aminoacídicas de las proteínas involucradas en el metabolismo de lípidos y ácidos grasos ACC y FAD2 de *C. muelleri*.**

	ACC	FAD2
Número de aminoácidos	2027	425
Peso Molecular (Daltons)	223750.1	48520.43
Punto isoeléctrico (pI)	5.11	6.85
Índice de inestabilidad	34	32.95
Índice alifático	90.63	78.02
Hidropatía GRAVY	-0.259	-0.116

<b>A</b> Amino acid composition:			<b>B</b> Amino acid composition:		
Ala (A)	165	8.1%	Ala (A)	30	7.1%
Arg (R)	108	5.3%	Arg (R)	12	2.8%
Asn (N)	79	3.9%	Asn (N)	17	4.0%
Asp (D)	132	6.5%	Asp (D)	21	4.9%
Cys (C)	12	0.6%	Cys (C)	4	0.9%
Gln (Q)	83	4.1%	Gln (Q)	7	1.6%
Glu (E)	158	7.8%	Glu (E)	16	3.8%
Gly (G)	161	7.9%	Gly (G)	34	8.0%
His (H)	32	1.6%	His (H)	20	4.7%
Ile (I)	134	6.6%	Ile (I)	15	3.5%
Leu (L)	175	8.6%	Leu (L)	43	10.1%
Lys (K)	115	5.7%	Lys (K)	23	5.4%
Met (M)	46	2.3%	Met (M)	10	2.4%
Phe (F)	75	3.7%	Phe (F)	30	7.1%
Pro (P)	94	4.6%	Pro (P)	23	5.4%
Ser (S)	131	6.5%	Ser (S)	28	6.6%
Thr (T)	91	4.5%	Thr (T)	26	6.1%
Trp (W)	16	0.8%	Trp (W)	19	4.5%
Tyr (Y)	59	2.9%	Tyr (Y)	21	4.9%
Val (V)	161	7.9%	Val (V)	26	6.1%
Pyl (O)	0	0.0%	Pyl (O)	0	0.0%
Sec (U)	0	0.0%	Sec (U)	0	0.0%
(B)	0	0.0%	(B)	0	0.0%
(Z)	0	0.0%	(Z)	0	0.0%
(X)	0	0.0%	(X)	0	0.0%

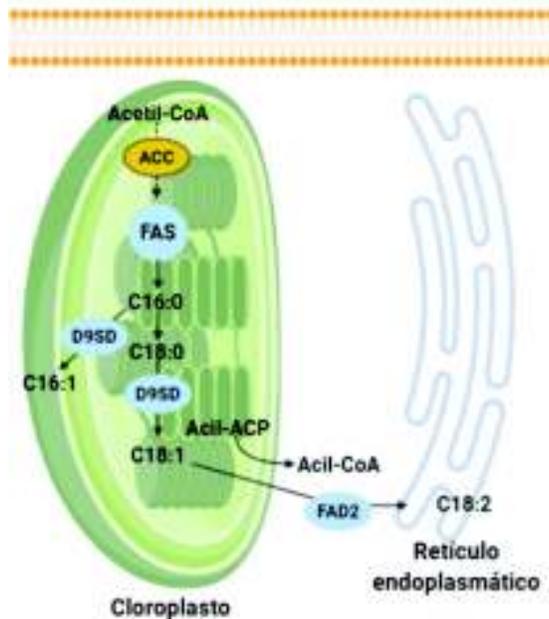
**Figura 34.** Composición aminoacídica de las secuencias de las enzimas involucradas en la síntesis de ácidos grasos de *C. muelleri*. A: ACC. B: FAD2.

Los resultados del análisis realizado por la herramienta ProtParam de ExPASy pueden brindar información de la naturaleza de estas secuencias aminoacídicas; se puede observar que, tanto por el tamaño como el peso molecular, las secuencias aminoacídicas analizadas de *C. muelleri* correspondientes ACC y FAD2 pueden clasificarse como polipéptidos ( $>100$  aa;  $PM \geq 5000$  Daltons) (Berthold, 2014) (Tabla 15).

Un parámetro importante estimado por la herramienta ProtParam, es la hidropatía GRAVY (Grand Averigüe of Hidropatía) determina la hidrofiliidad o hidrofobicidad de una proteína y se calcula mediante la suma de los valores de hidropatía de todos los aminoácidos presentes en la secuencia aminoacídica y se divide entre la longitud de dicha secuencia; sus valores varían entre -2 y 2 para la mayoría de las proteínas en donde los valores positivos están relacionados con comportamientos hidrofóbicos (Walker *et al.*, 2005). Las dos secuencias aminoacídicas consideradas en este estudio presentaron valores negativos, por lo que se sugiere cierta naturaleza hidrofílicas en estas dos enzimas (Tabla 15).

### **6.6.1 ACC: Acetil CoA carboxilasa**

El primer paso clave en la biosíntesis *de novo* de ácidos grasos es la carboxilación dependiente de ATP de acetil-CoA para producir malonil-CoA; una reacción de dos pasos catalizada por la enzima acetil CoA carboxilasa (ACC) que contiene biotina (Figura 35) (Sasaki y Nagano, 2004). En la naturaleza, la ACC se presenta en dos formas: una es un complejo heteromérico de múltiples unidades que contiene cuatro polipéptidos diferentes, entre ellos la biotina carboxilasa (BC), la proteína portadora de la biotina carboxilo (BCCP) y la  $\alpha$ - y  $\beta$ - carboxiltransferasa ( $\alpha$ - y  $\beta$ -CT); mientras que la otra es una forma homomérica en la que cada uno de los cuatro componentes antes mencionados se fusionan en tándem en un solo polipéptido grande (Sasaki y Nagano, 2004). La ubicación subcelular de la ACC varía entre los organismos (Li-Beisson *et al.*, 2019). La ACC animal, que es de la forma homomérica, está presente tanto en el citoplasma como en las mitocondrias, mientras que la ACC vegetal y algal, que es de la forma homomérica y heteromérica, se encuentra en el citoplasma y en los cloroplastos (Hu *et al.*, 2008; Abu-Elheiga *et al.*, 2000).



**Figura 35.** ACC: Acetil CoA carboxilasa. Creada en Biorender: <https://biorender.com/>

La ACC desempeña un papel importante en la regulación del flujo de carbono en la biosíntesis de ácidos grasos (Page *et al.*, 1994), incluso en microalgas, en las que se ha observado una correlación positiva entre la actividad de esta enzima y la cantidad de ácidos grasos en especies tales como *Chlorella vulgaris* (Leyva *et al.*, 2014). Varios estudios de ingeniería genética han sido llevados a cabo en esta enzima, siendo uno de sus principales objetivos incrementar el contenido de ácidos grasos tanto en plantas como en microalgas mediante su sobreexpresión; sin embargo, entre las especies que se han estudiado hasta el momento, los resultados han sido variables, desde incrementos significativos hasta no percibirse cambio alguno en el contenido de ácidos grasos; no obstante, no deja de ser una enzima atractiva para su estudio (Chen *et al.*, 2019; Roesler *et al.*, 1997; Dunahay *et al.*, 1996).

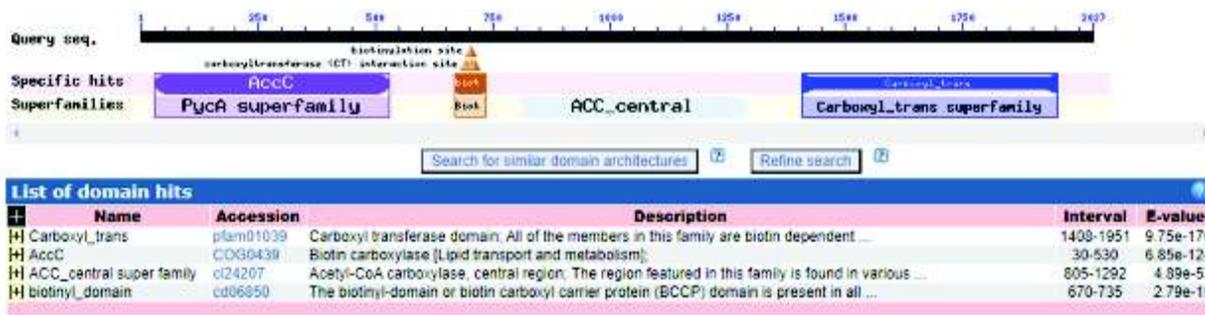
La secuencia nucleotídica de ACC de *C. muelleri* fue extraída del transcriptoma y se tradujo para obtener su secuencia aminoacídica para después realizar un BLASTp contra la base de datos del NCBI dando como resultado una similitud alta con proteínas ACC de varias especies de diatomeas. Las especies que presentaron una mayor homología con la secuencia aminoacídica de ACC de *C. muelleri* fueron *T. pseudonana* (72.81 %), *C. cryptica* (71.42%) y

*P. tricornutum* (71.19 %) (Tabla 16). En la Tabla 16, además del porcentaje de homología con otras especies de diatomeas, se incluye el valor de e y el número de acceso de secuencias aminoacídicas para cada especie.

**Tabla 16. Comparación del porcentaje de identidad de la secuencia aminoacídica ACC de *C. muelleri* con las secuencias correspondientes a otras diatomeas.**

Organismo	Valor de e	Porcentaje de identidad (%)	Número de acceso
<i>Thalassiosira pseudonana</i>	0.0	72.81	XP_002296083.1
<i>Cyclotella cryptica</i>	0.0	71.42	AAA81471.1
<i>Phaeodactylum tricornutum</i>	0.0	71.19	XP_002183067.1
<i>Skeletonema marinoi</i>	0.0	70.34	QKG32051.1
<i>Fistulifera solaris</i>	0.0	70.09	GAX09673.1
<i>Pseudo-nitzschia multistriata</i>	0.0	68.76	VEU44842.1

Asimismo, a través de la herramienta CD search (Conserved Domain search) se identificaron cuatro dominios conservados característicos de ACC en la secuencia aminoacídica de *C. muelleri*: carboxil transferasa, biotina carboxilasa, la región central de acetil coA-carboxilasa y biotina carboxilo portadora de proteínas (Figura 36).

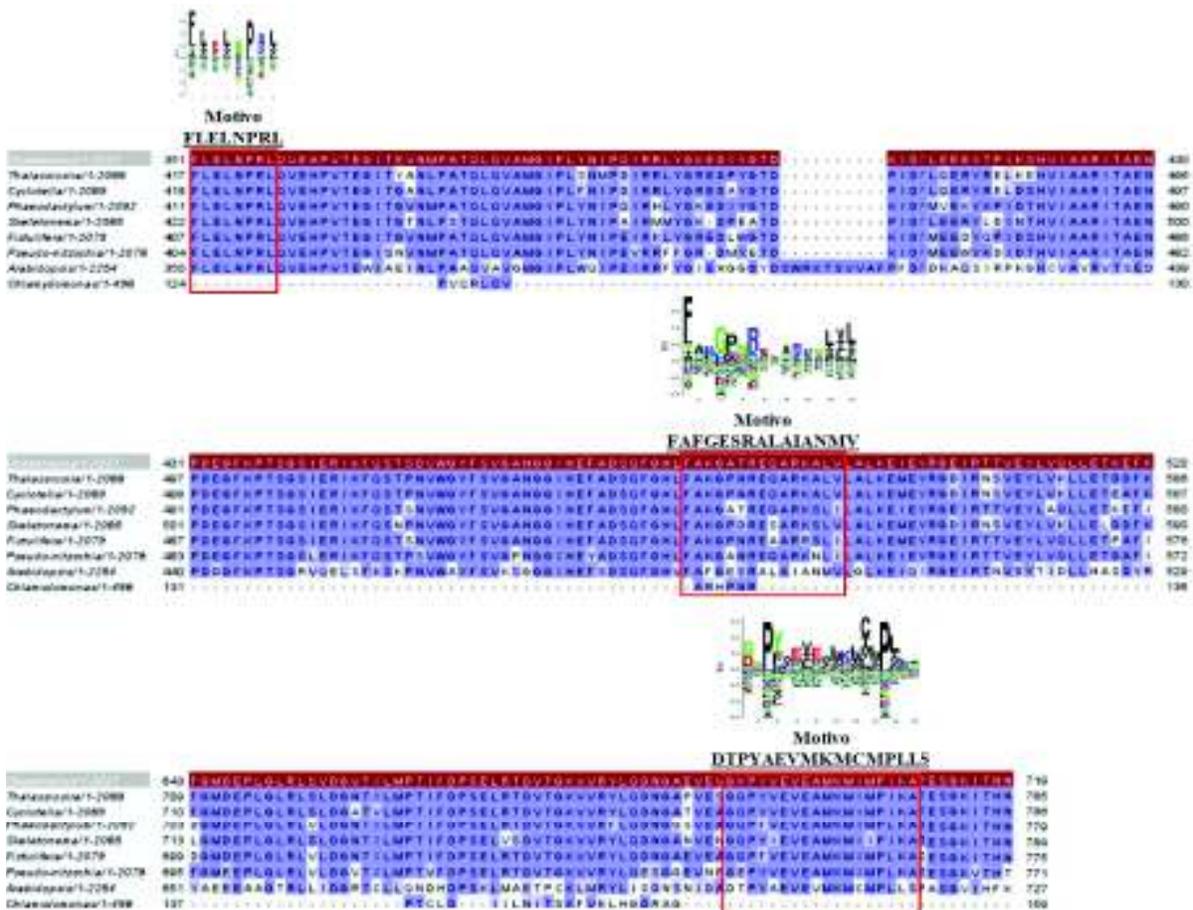


**Figura 36.** Dominios conservados en la secuencia aminoacídica de ACC de *C. muelleri*.

Por lo que se refiere a la predicción de motivos conservados en la secuencia aminoacídica de ACC en *C. muelleri*, se identificaron tres motivos mediante el análisis de las secuencias por la herramienta PROSITE de ExPASy y mediante la revisión de la literatura disponible para esta enzima (Figura 37). En primer lugar, el dominio ‘FLELNPRL’, que hace referencia a un subdominio de carbamoil-fosfato sintasa (CPS), se encontró altamente conservado en todas las secuencias aminoacídicas de las especies consideradas en el análisis de ACC a excepción de

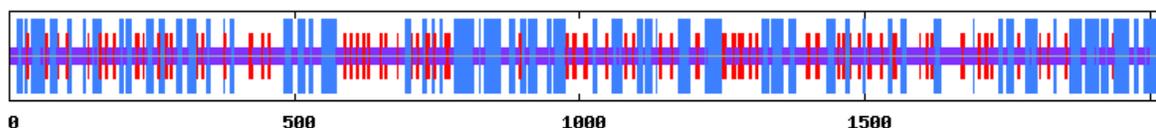
*Chlamydomonas reinhardtii*. CPS cataliza la formación de carbamoilfosfato a partir de MgATP, glutamina y bicarbonato a través de un intermediario de carboxifosfato; dicho intermediario es ampliamente aceptado que es formado por carboxilasas de biotina (Jitrapakdee y Wallace, 2003).

Por otra parte, el segundo motivo identificado fue 'FAFGESRALAIANMV' el cual de acuerdo al análisis realizado en PROSITE, hace alusión a la superfamilia de proteínas transportadoras ABC; este tipo de proteínas se caracterizan por hidrolizar ATP para energizar diversos sistemas biológicos (Andreoletti *et al.*, 2017). En diatomeas, se observó un ligero grado de conservación con la secuencia aminoacídica de *A. thaliana* pero aun entre ellas, se observaron diferencias por eventos de sustitución. Por último, otro motivo identificado por la herramienta PROSITE fue 'DTPYAEVMKMCMPLLS' y según el análisis hace referencia a un sitio de acoplamiento de las enzimas que requieren biotina; de acuerdo a Jitrapakdee y Wallace, (2003) la conservación de los residuos 'MKM' es crítica sobre todo de lisina (K) para su reconocimiento por la biotina ligasa, ya que mover este residuo en una posición hacia el extremo terminal N o C de su ubicación natural reduce la eficiencia de la reacción de biotinilación (Reche *et al.*, 1998). Esto fue congruente con todas las secuencias analizadas en este estudio a excepción de *Chlamydomonas reinhardtii*.



**Figura 37.** Alineamiento de las secuencias de aminoácidos de ACC, mostrando 3 motivos identificados. La secuencia de ACC de *C. muelleri* está resaltada en rojo.

En cuanto a la predicción de la estructura secundaria de ACC de *C. muelleri*, los resultados proporcionados por la herramienta NPS@ se muestran en la Figura 38 y la Tabla 16. Se obtuvo un valor de 689 aminoácidos con plegamientos alfa hélice que representa el 33.99 % de la secuencia aminoacídica, seguido de un 14.60 % correspondiente a láminas beta (296 residuos) y 1042 residuos (51.41%) para enrollamientos aleatorios del total de 2027 aminoácidos que conforman la secuencia aminoacídica de ACC de *C. muelleri*. Como puede apreciarse, poco más de la mitad está conformada por enrollamientos aleatorios, sugiriendo que en su mayoría posee plegamientos irregulares (Dill y Shortle, 1991).



**Figura 38.** Representación gráfica de la predicción de la estructura secundaria de la proteína ACC de *C. muelleri* realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.

**Tabla 17. Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de ACC de *C. muelleri* obtenidos por el servidor NPS@.**

Tipo	Número de aminoácidos	Porcentaje (%)
Alpha helix (Hh)	689	33.99
310 helix (Gg)	0	0.00
Pi helix (Ii)	0	0.00
Beta bridge (Bb)	0	0.00
Extended strand (Ee)	296	14.60
Beta turn (Tt)	0	0.00
Bend region (Ss)	0	0.00
Random coil (Cc)	1042	51.41
Ambiguous states (?)	0	0.00
Other states	0	0.00

Finalmente, la Figura 39 muestra el modelo propuesto de la estructura terciaria de la proteína ACC de *C. muelleri*. Se emplearon plantillas procedentes de modelos de proteínas de genes ortólogos de la base de datos de SWISS MODEL, encontrándose para esta enzima un porcentaje moderado de identidad (42.19 %) con el modelo de referencia; esto sugiere que hay un cierto grado de conservación y homología con organismos ampliamente estudiados.

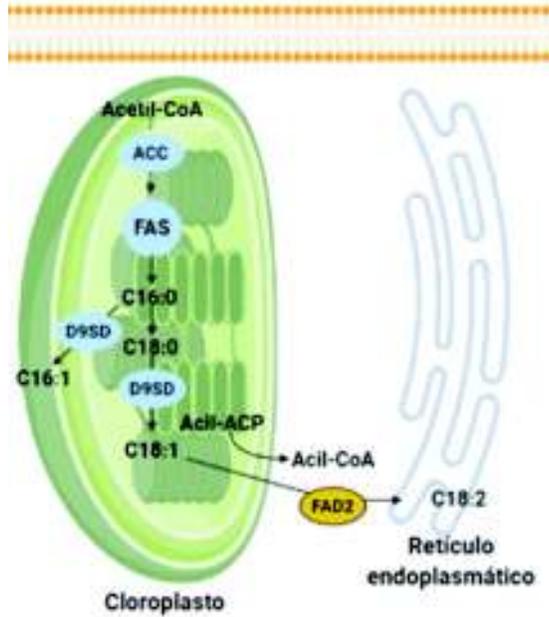


**Figura 39.** Modelo de la estructura terciaria teórica de la enzima acetil-CoA carboxilasa (ACC) de *C. muelleri* obtenida a través del servidor SWISS MODEL.

### 6.6.2 FAD2: Delta 12-desaturasa.

FAD2, también conocida como desaturasa de ácidos grasos omega-6, es la enzima clave responsable de la biosíntesis de ácidos grasos poliinsaturados (AGPI) y se encarga de catalizar la conversión del ácido oleico (C18:1- $\Delta$ 9) a linoleico (Figura 40) (Dar *et al.*, 2017). Además, desempeña un papel esencial en los sistemas de membranas biológicas, la señalización, el almacenamiento de energía, la adaptación a cambios de temperatura y resistencia a estrés biótico y abiótico (Cao *et al.*, 2013). El gen *FAD2* se identificó por primera vez en *Arabidopsis thaliana* con una sola copia, que se expresa constitutivamente (Okuley *et al.*, 1994); en estudios posteriores se ha identificado más de un gen *FAD2* en diversos cultivos terrestres (Kargiotidou *et al.*, 2008; Rolletschek *et al.*, 2007; Li *et al.*, 2007). Por otro lado, diversos estudios en ingeniería genética en diversos organismos se ha observado que la sobreexpresión de *FAD2* favorece la síntesis de ácido linoleico mientras que su silenciamiento lo disminuye y se incrementan los niveles de ácido oleico (Lamers *et al.*, 2019; Wu *et al.*, 2013).

El ácido linoleico es un ácido graso poliinsaturado esencial para la salud y la nutrición ya que no puede ser sintetizado por los seres humanos y deben ser suministrados a través de la dieta (Guan *et al.*, 2012). A pesar de los beneficios para la salud de los AGPI, su abundancia en el aceite hace que este sea más vulnerable al enranciamiento, afectando su sabor y reduciendo su vida útil (Pandey *et al.*, 2014). La estabilidad oxidativa y el valor nutritivo del aceite comestible dependen del contenido de ácidos grasos del aceite, especialmente de los ácidos oleico y linoleico (Cao *et al.*, 2013). Se comprobó que el ácido oleico tiene una mayor estabilidad oxidativa que el ácido linoleico, lo que da lugar a la prolongación de su vida útil (Ge *et al.*, 2015). Por lo tanto, hay una gran demanda de aceite de calidad superior rico en ácidos grasos monoinsaturados y bajo contenido en AGPI. Esos aceites son más deseables, tanto desde el punto de vista nutricional como comercial (Guan *et al.*, 2012).



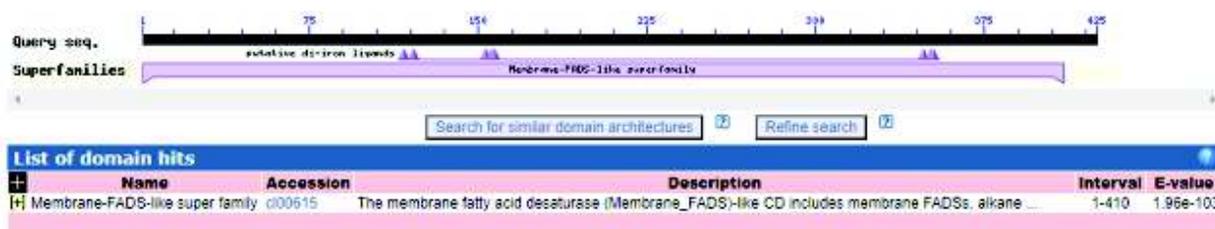
**Figura 40.** FAD2: Delta 12-desaturasa. Creada en Biorender: <https://biorender.com/>

La secuencia nucleotídica de FAD2 de *C. muelleri* fue extraída del transcriptoma y se tradujo para obtener su secuencia aminoacídica para posteriormente realizar un BLASTp contra la base de datos del NCBI dando como resultado una similitud alta con proteínas FAD2 de varias especies de diatomeas. Las especies que presentaron una mayor homología con la secuencia aminoacídica de FAD2 de *C. muelleri* fueron *T. oceanica* (68.59 %), *T. pseudonana* (68.28 %), y *F. solaris* (65.73 %) (Tabla 18). En la Tabla 18, además del porcentaje de homología con otras especies de diatomeas, se incluye el valor de e y el número de acceso de secuencias aminoacídicas para cada especie.

Por otra parte, se identificó un dominio correspondiente a la superfamilia de desaturasas de ácidos grasos de membrana (“Membrane-FADS-like superfamily”) en la secuencia aminoacídica de *C. muelleri* a través de la herramienta CD search (Conserved Domain search). Los miembros pertenecientes a esta superfamilia de proteínas se caracterizan por estar involucrados en la introducción regioselectiva de dobles enlaces en las cadenas alifáticas de ácidos grasos y por desempeñar un papel importante en el mantenimiento de la estructura y funcionamiento adecuado de las membranas biológicas (Lu *et al.*, 2020) (Figura 41).

**Tabla 18. Comparación del porcentaje de identidad de la secuencia aminoacídica FAD2 de *C. muelleri* con las secuencias correspondientes a otras diatomeas.**

Organismo	Valor de e	Porcentaje de identidad (%)	Número de acceso
<i>Thalassiosira oceanica</i>	0.0	68.59	EJK71021.1
<i>Thalassiosira pseudonana</i>	0.0	68.28	XP_002292071.1
<i>Fistulifera solaris</i>	0.0	65.73	BA027791.1
<i>Phaeodactylum tricorutum</i>	0.0	64.73	XP_002186139.1
<i>Fragilariopsis cylindrus</i>	0.0	62.19	OEU07512.1
<i>Pseudo-nitzschia multistriata</i>	0.0	60.97	VEU44279.1



**Figura 41.** Dominio conservado de desaturasas de ácidos grasos en la secuencia aminoacídica de FAD2 de *C. muelleri*.

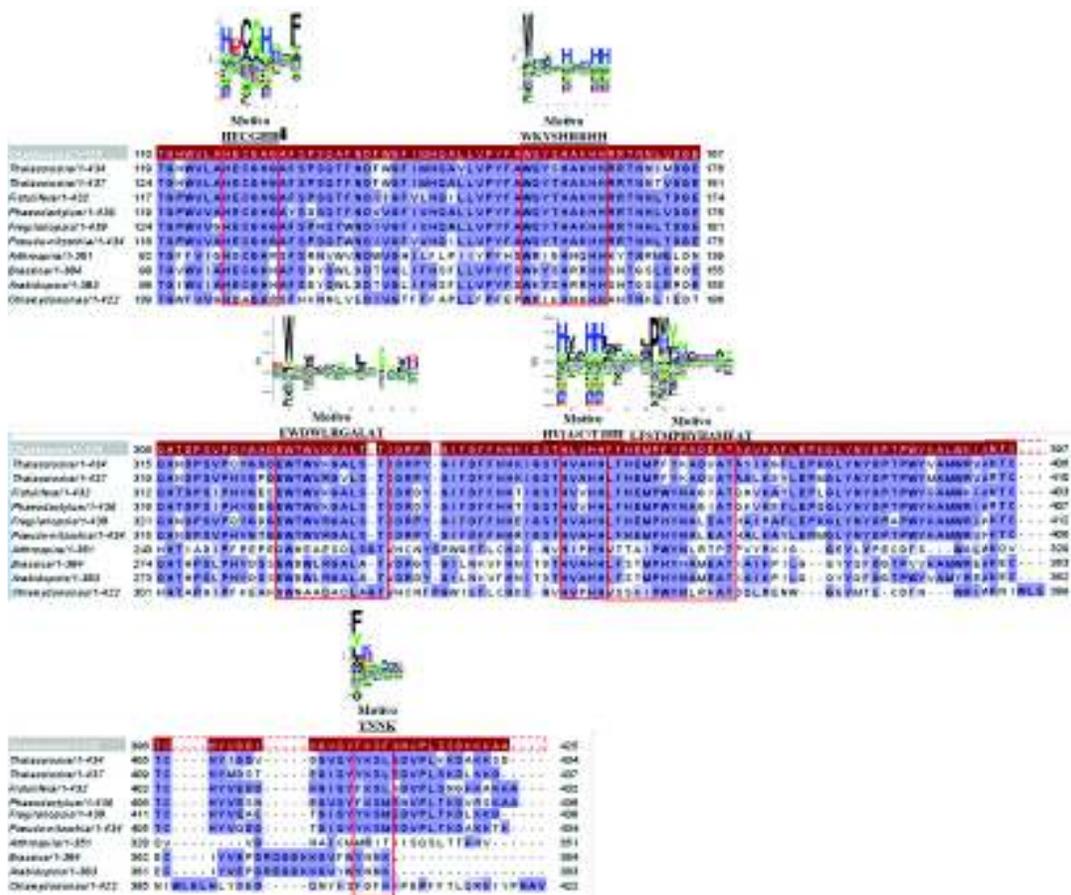
En relación con la identificación de motivos conservados en la secuencia aminoacídica de *C. muelleri*, se comparó en un alineamiento múltiple de secuencias con secuencias similares pertenecientes a diferentes especies de diatomeas, dos secuencias aminoacídicas de FAD2 de plantas (*Arabidopsis thaliana* y *Brassica napus*) y otras dos correspondientes a las microalgas *Chlamydomonas reinhardtii* y *Arthrospira platensis* (Figura 42). De acuerdo a un estudio *in silico* de FAD2 en 16 especies de plantas oleaginosas realizado por Nayeri y Yarizade (2014), encontraron que el primer motivo de histidina en todas las secuencias de FAD2 fue ‘HECGHH’, el cual se encontró altamente conservado en todos los organismos considerados en este estudio a excepción de *Chlamydomonas reinhardtii* y *Arthrospira platensis* que presentaron un grado ligero de sustitución de residuos. Además, Nayeri y Yarizade (2014) mencionaron que en todas las secuencias que analizaron, el residuo alanina fue el primero que se ubicó inmediatamente después de esta caja de histidina, lo cual fue consistente también en este estudio con las especies analizadas, salvo en las mismas especies en que se presentaron sustituciones en este motivo observándose la sustitución de alanina por una serina en ambas secuencias.

Otro motivo de histidina que está altamente conservado en las secuencias de FAD2 es ‘HRRHH’, el cual se encuentra albergado en la segunda caja de histidina ‘WKYSHRRHH’ (Shanklin *et al.*, 1994). En el caso particular de las diatomeas, todas las secuencias mostraron la misma sustitución de este motivo en donde en lugar de contar con dos residuos de arginina (R), se encuentran una alanina (A) y una lisina (K). Por otro lado, con respecto a la segunda caja de histidina (‘WKYSHRRHH’) en la que se encuentra este motivo, también observaron otras dos sustituciones de residuos en las secuencias aminoacídicas de diatomeas, en donde en lugar de una lisina (K), se encontró glutamina (Q) y en lugar de serina (S) se observó una treonina; sólo las secuencias aminoacídicas de *C. muelleri* y *T. pseudonana* conservaron el residuo serina.

Otro motivo identificado en las secuencias aminoacídicas analizadas en este estudio y que también se observó en el estudio *in silico* de Nayeri y Yarizade (2014), fue ‘HV[A/C/T]HH’ el cual se ubica hacia el extremo carboxílico. Nuevamente, las diatomeas mostraron una mayor homología con respecto a las secuencias aminoacídicas de plantas, sin embargo, la secuencia de *C. muelleri*, presentó en lugar de valina (V) una lisina y la alanina (A) fue sustituida por una valina; las secuencias aminoacídicas de *P. tricornutum* y *F. solaris* manifestaron la misma sustitución del residuo alanina por una valina.

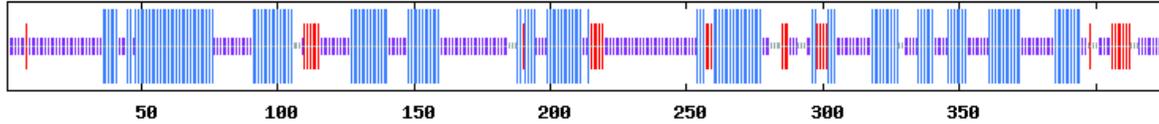
Nayeri y Yarizade (2014) señalaron también que entre esta última caja de histidina (‘HV[A/C/T]HH’), los motivos ‘EWDWLRGALAT’ y ‘LFSTMPHYHAMEAT’ son los más altamente conservados antes y después de la tercera caja de histidina respectivamente en secuencias aminoacídicas de plantas oleaginosas; estos mismos motivos fueron identificados en las secuencias aminoacídicas de FAD2 analizadas en este estudio, encontrándose un cierto grado de conservación con respecto a las especies de diatomeas.

Finalmente, otro motivo conservado en secuencias aminoacídicas de FAD2 de plantas oleaginosas es ‘YNNK’ [también puede encontrarse como Y(K/N)NKF o YRNKI], el cual se ubica en el extremo C-terminal. Se detectó la presencia de este motivo en las secuencias aminoacídicas de *Arabidopsis thaliana* y *Brassica napus*; sin embargo, se encontraron diferencias con el resto de especies analizadas en estudio e inclusive entre las especies de diatomeas. De acuerdo a Nayeri y Yarizade (2014), la disminución de la conservación de este motivo pudiera indicar la pérdida de la señal de recuperación del retículo endoplásmico, dando lugar a la disociación de esta enzima del retículo endoplásmico (McCartney *et al.*, 2004).



**Figura 42.** Alineamiento de las secuencias de aminoácidos de FAD2, mostrando 4 motivos significativamente conservados. La secuencia de FAD2 de *C. muelleri* está resaltada en rojo.

Por otra parte, se predijo la estructura secundaria de la proteína FAD2 de *C. muelleri* y se calcularon los porcentajes de alfa hélices y láminas beta los cuales se muestran en la Figura 43 y la Tabla 19 respectivamente. Se obtuvo un valor de 169 aminoácidos con plegamientos alfa hélice que representa el 39.76 % de la secuencia aminoacídica, seguido de un 7.29 % correspondiente a láminas beta (31 residuos), 204 residuos (48.00 %) para enrollamientos aleatorios y un 4.94 % de ambigüedades (21 residuos) del total de 425 aminoácidos que conforman la secuencia aminoacídica de FAD2 de *C. muelleri*. Al igual que el resto de secuencias aminoacídicas analizadas en este estudio, mostró un porcentaje alto de enrollamientos aleatorios, lo que indicaría que en su mayoría está compuesta por regiones plegadas irregularmente (Dill y Shortle, 1991).



**Figura 43.** Representación gráfica de la predicción de la estructura secundaria de la enzima FAD2 de *C. muelleri* realizada en el servidor NPS@. Las secciones que se encuentran en color azul representan las alfa hélices, las secciones en color rojo a las láminas beta y las secciones en morado a los enrollamientos aleatorios.

**Tabla 19.** Porcentajes de alfa hélices, láminas beta y enrollamientos aleatorios en la estructura secundaria de FAD2 de *C. muelleri* obtenidos por el servidor NPS@.

Tipo	Número de aminoácidos	Porcentaje (%)
Alpha helix (Hh)	169	39.76
310 helix (Gg)	0	0.00
Pi helix (Ii)	0	0.00
Beta bridge (Bb)	0	0.00
Extended strand (Ee)	31	7.29
Beta turn (Tt)	0	0.00
Bend region (Ss)	0	0.00
Random coil (Cc)	204	48.00
Ambiguous states (?)	21	4.94
Other states	0	0.00

Finalmente, la Figura 44 muestra el modelo propuesto de la estructura terciaria de la desaturasa FAD2 de *C. muelleri*. Se emplearon plantillas procedentes de modelos de proteínas de genes ortólogos de la base de datos de SWISS MODEL, encontrándose para esta enzima un porcentaje bajo de identidad (18.75 %) con el modelo de referencia. Esto puede deberse a que la predicción de la estructura terciaria, al basarse en algoritmos de homología, requiere de proteínas cuyas estructuras ya estén experimentalmente determinadas y que tengan un alto porcentaje de homología a la proteína analizada, por lo que en este caso pudiera verse limitado al ser escasa la información para diatomeas.

Template	Seq Identity	Coverage	Description
4zyo_1.A	16.75%	<input type="checkbox"/>	Acyl-CoA desaturase Crystal Structure of Human Integral Membrane Stearoyl-CoA Desaturase with Substrate



**Figura 44.** Modelo de la estructura terciaria teórica de la delta 12-desaturasa (FAD2) de *C. muelleri* obtenida a través del servidor SWISS MODEL.

## VII. CONCLUSIONES

- ✓ El fósforo (P) posee un efecto importante en el crecimiento, densidad celular y composición bioquímica en *C. muelleri* ya que, al disminuir la concentración de este nutriente, se incrementó el contenido de lípidos totales pero disminuyó la densidad celular. Por el contrario, el exceso de fósforo no mostró un efecto sobre la densidad celular ni en el contenido de lípidos totales con respecto al control.
- ✓ Las concentraciones de P y las condiciones de cultivo establecidas en este estudio no poseen un efecto en el contenido total de ácidos grasos saturados de *C. muelleri* al no encontrarse diferencias significativas entre los tratamientos.
- ✓ La síntesis de ácidos grasos monoinsaturados se ve favorecida bajo condiciones limitantes de P (7 y 18  $\mu\text{M}$  P) en *C. muelleri*, lo cual está relacionado con la síntesis de compuestos de reserva. En cambio, la síntesis de ácidos grasos poliinsaturados es inducida bajo concentraciones de P consideradas como control (72  $\mu\text{M}$  P) y exceso (144  $\mu\text{M}$  P), los cuales se encuentran relacionados con lípidos de membrana.
- ✓ Los resultados obtenidos del cálculo de las propiedades de biodiesel a partir de los perfiles de ácidos grasos, sugieren que *C. muelleri* cultivada bajo una concentración de P de 7  $\mu\text{M}$  es un recurso potencial para la producción de biodiesel de acuerdo a la norma ASTM D 6751 de Estados Unidos de América.
- ✓ La inducción de genes involucrados en procesos biológicos tales como metabolismo de P, transporte transmembranal, biosíntesis de sulfolípidos, metabolismo de lípidos y biosíntesis de carbohidratos sugieren que *C. muelleri* bajo condiciones limitantes de P implementa estrategias para incrementar la asimilación, transporte y uso eficiente de este nutriente y sintetizar compuestos de reserva para satisfacer sus requerimientos celulares.
- ✓ La represión de genes involucrados en el metabolismo de nitrógeno y traducción de proteínas pudiera indicar que se está viendo afectada la síntesis de proteínas y ribosomas ya que ello implica un gasto energético para la célula por lo que el P disponible pudiera estarlo empleando para su supervivencia. Por lo que se sugiere que el crecimiento y división celular están viéndose comprometidos.

- ✓ La represión de genes involucrados en la fotosíntesis y síntesis de pigmentos conduce a la disminución de la actividad del aparato fotosintético y de los niveles de estos pigmentos derivando en estrés oxidativo en *C. muelleri* bajo concentraciones limitantes de P.
- ✓ La modulación del metabolismo del ácido absícico (ABA) bajo condiciones limitantes de P ha sido escasamente estudiado en diatomeas, siendo este el primer estudio hasta nuestro conocimiento, que reporta como estrategia una posible acumulación de esta hormona dentro de la célula ante la limitación por P en *C. muelleri* para mitigar el daño derivado del estrés.
- ✓ A nivel molecular la acumulación de lípidos en *C. muelleri* no puede ser atribuida directamente a la biosíntesis de ácidos grasos ya que varios genes que codifican a enzimas clave como la acetil-CoA carboxilasa se encontraron reprimidos o no se expresaron diferencialmente bajo las condiciones establecidas en este estudio.
- ✓ El proceso de remodelación de la membrana celular posiblemente sea el mecanismo responsable de proveer ácidos grasos mediante la degradación de fosfolípidos para la síntesis de triacilglicéridos en *C. muelleri* cultivada bajo concentraciones limitantes de P.
- ✓ En base a los resultados obtenidos, se corrobora la hipótesis planteada ya que en condiciones de exceso de P, *C. muelleri* no requiere restringir su actividad celular ni implementar estrategias para incrementar la asimilación, reciclaje y aprovechamiento eficiente de este nutriente a diferencia de lo que ocurre bajo condiciones limitantes.
- ✓ El ensamblaje y el análisis *in silico* de secuencias de genes completos que codifican para proteínas transportadoras de fósforo y enzimas clave involucradas en la síntesis de ácidos grasos permitirá plantear estudios con mayor profundidad que nos ayuden a demostrar su participación y entender la regulación de la biosíntesis de lípidos en *C. muelleri* bajo limitación de nutrientes.

## VIII. RECOMENDACIONES

- ✓ Validar los resultados de RNA-seq generados en este estudio mediante la técnica de RT-qPCR.
- ✓ Determinar los parámetros de calidad del biodiesel experimentalmente a partir del aceite extraído de *C. muelleri* con el fin de corroborar la predicción realizada en este estudio.
- ✓ Profundizar en el estudio de transcritos no codificantes en el transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P.
- ✓ Seleccionar genes de interés identificados en el análisis de RNA-seq que codifican a enzimas claves involucradas en el metabolismo de lípidos tales como ACC, PDAT2, DGAT2, MGL y TAG lipasa para desarrollar estudios de ingeniería genética en *C. muelleri* lo que permitiría ampliar el conocimiento sobre su función en el proceso de acumulación de lípidos.

## IX. LITERATURA CITADA

- Abd El-Baky, H. H., F. K. Ri-Baz y G. S. El Baroty. 2010. Enhancing antioxidant availability in wheat grains from plants grown under seawater stress in response to microalgae extract treatments. *Journal of the Science of Food and Agriculture* 90: 299–303. doi: 10.1002/jsfa.3815
- Abdel-Raouf, N., A. A. Al-Homaidan, y I. B. M. Ibraheem. 2012. Microalgae and wastewater treatment. *Saudi Journal of Biological Sciences* 19(3): 257-275. doi: 10.1016/j.sjbs.2012.04.005
- Abo, B. O., E. A. Odey, M. Bakayoko y L. Kalakodio. 2019. Microalgae to biofuels production: a review on cultivation, application and renewable energy. *Reviews on Environmental Health*, 34(1): 91-99. doi: 10.1515/reveh-2018-0052
- Abu-Elheiga, L., W. R. Brinkley, L. Zhong, S. S. Chirala, G. Woldegiorgis y S. J. Wakil. 2000. The subcellular localization of acetyl-CoA carboxylase 2. *Proceedings of the National Academy of Sciences* 97(4): 1444-1449. doi: 10.1073/pnas.97.4.1444
- Ahmad, Z., F. Nadeem, R. Wang, X. Diao, Y. Han, X. Wang y X. Li. 2018. A larger root system is coupled with contrasting expression patterns of phosphate and nitrate transporters in foxtail millet [*Setaria italica* (L.) Beauv.] under phosphate limitation. *Frontiers in Plant Science* 9: 1367. doi: 10.3389/fpls.2018.01367
- Ahmad, A.L., N.M. Yasin, C.J.C. Derek y J.K. Lim. 2011. Microalgae as a sustainable energy source for biodiesel production: A review. *Renewable and Sustainable Energy Reviews* 15(1): 584-593. doi: 10.1016/j.rser.2010.09.018
- Ahmed, K., D. A. Gerber y C. Cochet. 2002. Joining the cell survival squad: an emerging role for protein kinase CK2. *Trends in Cell Biology* 12(5): 226-230. doi: 10.1016/S0962-8924(02)02279-1
- Alcoverro, T., E. Conte y L. Mazella. 2000. Production of mucilage by the adriatic epipellic diatom. *Journal of Phycology* 36: 1087–1095. doi: 10.1046/j.1529-8817.2000.99193.x
- Alipanah, L., P. Winge, J. Rohloff, J. Najafi, T. Brembu y A. M. Bones. 2018. Molecular adaptations to phosphorus deprivation and comparison with nitrogen deprivation responses in the diatom *Phaeodactylum tricornutum*. *PloS one* 13(2): e0193335. doi: 10.1371/journal.pone.0193335
- Alipanah, L., J. Rohloff, P. Winge, A. M. Bones y T. Brembu. 2015. Whole-cell response to nitrogen deprivation in the diatom *Phaeodactylum tricornutum*. *Journal of Experimental Botany* 66(20): 6281-6296. doi: 10.1093/jxb/erv340
- Almeida, D. T. D., T. V. Viana, M. M. Costa, C. D. S. Silva y S. Feitosa. 2019. Effects of different storage conditions on the oxidative stability of crude and refined palm oil, olein and stearin (*Elaeis guineensis*). *Food Science and Technology* 39: 211-217. doi: 10.1590/fst.43317
- Alvarado-Moreno, J. A. y H. Mayani. 2007. El ciclo celular y su papel en la biología de las células progenitoras hematopoyéticas. *Gaceta médica de México* 143(2): 149-161.
- Anbuechzian, R., V. Karuppiah y Z. Li. 2015. Prospect of marine algae for production of industrially important chemicals. 195-217 p. En: Das, D (Ed.), *Algal Biorefinery: An Integrated Approach*. Springer, Cham.
- Andreoletti, P., Q. Raas, C. Gondcaille, M. Cherkaoui-Malki, D. Trompier y S. Savary. 2017. Predictive structure and topology of peroxisomal ATP-binding cassette (ABC) transporters. *International Journal of Molecular Sciences* 18(7): 1593. doi: 10.3390/ijms18071593

- Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. *Braham Bioinformatics*.
- Anne-Marie, K., W. Yee, S. H. Loh, A. Aziz y T. San Cha. 2020. Effects of excess and limited phosphate on biomass, lipid and fatty acid contents and the expression of four fatty acid desaturase genes in the tropical Selenastraceae *Messastrum gracile* SE-MC4. *Applied Biochemistry and Biotechnology* 190(4): 1438-1456. doi: 10.1007/s12010-019-03182-z
- Armbrust, E.V. 2009. The life of diatoms in the world's oceans. *Nature* 459(7244): 185–192. doi:10.1038/nature08057.
- Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., S. Zhou, A.E. Allen, K.E. Apt, et al. 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306(5693): 79-86. doi: 10.1126/science.1101156
- Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, et al. 2000. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nature Genetics* 25: 25 – 29. doi: doi.org/10.1038/75556
- Assis, A. F., E. H. Oliveira, P. B. Donate, S. Giuliatti, C. Nguyen y G. A. Passos. 2014. What is the transcriptome and how it is evaluated? 3-48 p. En: Passos, G.A. (Ed.), *Transcriptomics in Health and Disease* Springer, Cham.
- ASTM. 2008. Standard specification for biodiesel fuel blend stock (B100) for middle distillate fuels. Report no. D6751-08.
- Athenstaedt, K. y G. Daum. 1999. Phosphatidic acid, a key intermediate in lipid metabolism. *European Journal of Biochemistry* 266 (1): 1–16. doi: 10.1046/j.1432-1327.1999.00822.x
- Babu y Anad. 2019. Influence of fuel injection timing and nozzle opening pressure on a CRDI-assisted diesel engine fueled with biodiesel-dieselalcohol fuel. 353 – 390 p. En: Azad, K (Ed.), *Advances in Ecology Fuels for a Sustainable Environment*. Woodhead Publishing. doi: 10.1016/B978-0-08-102728-8.00013-9
- Banerjee, S., y S. Ramaswamy. 2019. Comparison of productivity and economic analysis of microalgae cultivation in open raceways and flat panel photobioreactor. *Bioresource Technology Reports* 8: 100328. doi: 10.1016/j.biteb.2019.100328.
- Barabás, I. y I. A. Todoruț. 2011. Biodiesel quality, standards and properties. 3-28 p. En; Montero, G., Stoytcheva, M., (Eds.), *Biodiesel-quality, emissions and by-products*. InTech. Rijeka, Croacia.
- Barone, V., A. Baglieri, P. Stevanato, C. Broccanello, G. Bertoldo, M. Bertaggia, M. Cagnin, D. Pizzeghello, V. M. C. Moliterni, G. Mandolino, et al. 2018. Root morphological and molecular responses induced by microalgae extracts in sugar beet (*Beta vulgaris* L.). *Journal of Applied Phycology* 30(2): 1061-1071. doi: 10.1007/s10811-017-1283-3
- Barsanti, L. y P. Gualtieri. 2018. Is exploitation of microalgae economically and energetically sustainable? *Algal Research* 31: 107-115. doi: 10.1016/j.algal.2018.02.001
- Barsanti, L y P. Gualtieri. 2014. *Algae: Anatomy, Biochemistry, and Biotechnology*. CRC press: Pisa, Italy. 202 p. ISBN 978-1-4398-6732-7.
- Barsanti L., P. Coltelli, V. Evangelista, A.M. Frassanito, V. Passarelli, N. Vesentini y P. Gualtieri. 2008. The World of Algae. 281–304 p. En: Evangelista, V., Barsanti, L., Frassanito, A.M. Passarelli, V., Gualtieri, P. (Eds.), *Algal Toxins: Nature, Occurrence, Effect and Detection* Springer Science & Business Media B.V. doi: 10.1007/978-1-4020-8480-5\_1.

- Bart, J. C. J., N. Palmeri y S. Cavallaro. 2010. 6 - Emerging new energy crops for biodiesel production. 226-284 p. En: Bart, J. C. J., Palmeri, N., Cavallaro, S. (Eds.), Biodiesel Science and Technology. Woodhead Publishing. doi: 10.1533/9781845697761.226
- Basu, S., A.S. Roy, K. Mohanty y A.K. Ghoshal. 2013. Enhanced CO<sub>2</sub> sequestration by a novel microalga: *Scenedesmus obliquus* SA1 isolated from bio-diversity hotspot region of Assam, India. *Bioresource Technology* 143:369–377. doi: 10.1016/j.biortech.2013.06.010
- Bateman, J.M. y S. Purton. 2000. Tools for chloroplast transformation in *Chlamydomonas*: Expression vectors and a new dominant selectable marker. *Molecular and General Genetics* 263: 404–410. doi: 10.1007/s004380051184
- Baynes, J. y M. H. Dominiczak. 2009. *Medical biochemistry*. Elsevier Health Sciences, Reino Unido.
- Becker-André, M. y K. Hahlbrock. 1989. Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a PCR aided transcript titration assay (PATTY). *Nucleic Acids Research* 17: 9437 – 9446.
- Belotti, G., M. Bravi, B. de Caprariis, P. de Filippis y M. Scarsella. 2013. Effect of nitrogen and phosphorus starvations on *Chlorella vulgaris* lipids productivity and quality under different trophic regimens for biodiesel production. *American Journal of Plant Sciences* 4(12): 44-51. doi.org/10.4236/ajps.2013.412A2006.
- Benning, C., Z.H. Huang y D.A. Gage. 1995. Accumulation of a novel glycolipid and a betaine lipid in cells of *Rhodobacter sphaeroides* grown under phosphate limitation. *Archives of Biochemistry and Biophysics* 317: 103-111. doi: 10.1006/abbi.1995.1141
- Berdalet, E., M. Latasa y M. Estrada. 1994. Effects of nitrogen and phosphorus starvation on nucleic acid and protein content of *Heterocapsa* sp. *Journal of Plankton Research*, 16(4): 303-316. doi.org/10.1093/plankt/16.4.303.
- Bere, T. y J. G. Tundisi. 2010. Biological monitoring of lotic ecosystems: the role of diatoms. *Brazilian Journal of Biology* 70(3): 493-502. doi: 10.1590/S1519-69842010005000009.
- Berg, J., J. Tymoczko y L. Stryer. 2012. *Biochemistry*. Freeman and Company. New York, Estados Unidos. pp. 600–601.
- Bertholf, R. L. 2014. Proteins and albumin. *Laboratory Medicine* 45(1): e25-e41. doi: 10.1309/LMKRNRGW5J03APZQ
- Bialeski, R. L. 1973. Phosphate pools, phosphate transport, and phosphate availability. *Annual review of Plant Physiology* 24(1): 225-252. doi: 10.1146/annurev.pp.24.060173.001301
- Bigler, C. y R. I. Hall. 2002. Diatoms as indicators of climatic and limnological change in Swedish Lapland: a 100-lake calibration set and its validation for paleoecological reconstructions. *Journal of Paleolimnology* 27(1): 97-115. doi: 10.1023/A:1013562325326
- Bligh, E.G. y W.J. Dyer. 1959. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* 37(8): 911-917. doi: 10.1139/o59-099.
- Boeuf, G. 2011. Marine biodiversity characteristics. *Comptes Rendus Biologies* 334(5-6): 435-440. doi: 10.1016/j.crv.2011.02.009
- Bolger, A. M., M. Lohse y B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15): 2114-2120. doi: 10.1093/bioinformatics/btu170

- Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., U. Maheswari, C. Martens, F. Maumus, R. P. Otilar, et al. 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456(7219): 239-244. doi: 10.1038/nature07410
- Bozbas, K. 2008. Biodiesel as an alternative motor fuel: Production and policies in the European Union. *Renewable and Sustainable Energy Reviews* 12(2): 542-552. doi: 10.1016/j.rser.2005.06.001
- Bracellar-Mendes, L. y Namorato de Souza, H. 2018. New strains and the future of microalgae production. *Journal of Aquatic Science and Marine Biology* 1(4): 7 – 9.
- Brembu, T., A. Mühlroth L. Alipanah y A.M. Bones. 2017. The effects of phosphorus limitation on carbon metabolism in diatoms. *Philosophical Transactions of the Royal Society B: Biological Sciences* 372(1728): 20160406. doi: 10.1098/rstb.2016.0406
- Brotosudarmo, T. H., A. A. Wibowo y M. A. Adhiwibawa. 2019. Single cells diatom *Chaetoceros muelleri* investigated by homebuilt confocal fluorescence spectro-microscopy. En: *Third International Seminar on Photonics, Optics, and Its Applications (ISPhOA 2018)* (Vol. 11044, p. 110440E). International Society for Optics and Photonics.
- Bryant, D. M., K. Johnson, T. DiTommaso, T. Tickle, M. B. Couger, D. Payzin-Dogru, T.J. Lee, N.D. Leigh, T. H. Kuo, F. G. Davis. 2017. A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Reports* 18(3): 762-776. doi: 10.1016/j.celrep.2016.12.063
- Bullard, J. H., E. Purdom, K. D. Hansen y S. Dudoit. 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* 11(1): 1-13. doi: 10.1186/1471-2105-11-94
- Bushmanova, E., D. Antipov, A. Lapidus y A. D. Prjibelski. 2019. rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience*, 8(9), giz100.
- Bushmanova, E., D. Antipov, A. Lapidus, V. Suvorov y A. D. Prjibelski. 2016. rnaQUAST: a quality assessment tool for de novo transcriptome assemblies. *Bioinformatics* 32(14): 2210-2212. doi: 10.1093/bioinformatics/btw218
- Bustin, S.A., V. Benes, T. Nolan y M.W. Pfaffl. 2005. Quantitative real-time RT-PCR- a perspective. *Journal of Molecular Endocrinology* 34(3): 597 – 601. doi: 10.1677/jme.1.01755
- Camarena, L., V. Bruno, G. Euskirchen, S. Poggio y M. Snyder. 2010. Molecular mechanisms of ethanol-induced pathogenesis revealed by RNA-sequencing. *PLoS Pathog* 6(4): e1000834. doi: 10.1371/journal.ppat.1000834
- Cañavate, J. P.; I. Armada e I. Hachero-Cruzado. 2017. Common and species-specific effects of phosphate on marine microalgae fatty acids shape their function in phytoplankton trophic ecology. *Microbial Ecology* 74: 623-639. doi: 10.1007/s00248-017-0983-1
- Cao, S., X. R. Zhou, C. C. Wood, A. G. Green, S. P. Singh, L. Liu y Q. Liu. 2013. A large and functionally diverse family of *Fad2* genes in safflower (*Carthamus tinctorius* L.). *BMC Plant Biology* 13(1): 5. doi: 10.1186/1471-2229-13-5
- Castillo, O. S., S. G. Torres-Badajoz, C. A. Núñez-Colín, V. Peña-Caballero, C. H. Herrera Méndez y J. R. Rodríguez-Núñez. 2017. Producción de biodiésel a partir de microalgas: avances y perspectivas biotecnológicas. *Hidrobiológica* 27(3): 337-352.
- Cazzonelli, C.I., M. Vanstraelen, S. Simon, K. Yin, A. Carron-Arthur, N. Nisar, G. Tarle, A.J. Cuttriss, I.R. Searle, E. Benkova, et al., 2013. Role of the *Arabidopsis* PIN6 auxin transporter in auxin

- homeostasis and auxin-mediated development. *PLoS One* 8: e70069. doi: 10.1371/journal.pone.0070069
- Chai, M.K., Y.H. Tan y L.S. Wong. 2020. Macronutrient effect on biomass of Microalgae in biofuel production: A review. *Indian Journal of Science and Technology* 13(22): 2245-2263. doi: 10.17485/IJST/v13i22.176
- Chang, M. X., M. Gu, Y.W. Xia, X.L. Dai, C.R. Dai, J. Zhang, S.C. Wang, H.Y. Qu, N. Yamaji, J.F. Ma. 2019. OsPHT1; 3 mediates uptake, translocation, and remobilization of phosphate under extremely low phosphate regimes. *Plant Physiology* 179(2): 656-670. doi: 10.1104/pp.18.01097
- Chauhan, S. K. y A. Shukla. 2011. Environmental impacts of production of biodiesel and its use in transportation sector. Croatia: Environmental impacts of biofuels. InTech 1-18. doi: 10.5772/20923
- Chen, D., X. Yuan, L. Liang, K. Liu, H. Ye, Z. Liu, Y. Liu, L. Huang, W. He, Y. Chen et al. 2019. Overexpression of acetyl-CoA carboxylase increases fatty acid production in the green alga *Chlamydomonas reinhardtii*. *Biotechnology Letters* 41(10): 1133-1145. doi: 10.1007/s10529-019-02715-0
- Chen, Y., C. Xu y S. Vaidyanathan. 2018. Microalgae: a robust “green bio-bridge” between energy and environment. *Critical Reviews in Biotechnology* 38(3): 351-368. doi: 10.1080/07388551.2017.1355774
- Chen, J., Y. Wang, F. Wang, J. Yang, M. Gao, C. Li, Y. Liu, Y. Liu, N. Yamaji, J.F. Ma, et al. 2015. The rice CK2 kinase regulates trafficking of phosphate transporters in response to phosphate levels. *The Plant Cell* 27(3): 711-723. doi: 10.1105/tpc.114.135335
- Chen, Z., H. Zhang, D. Jablonowski, X. Zhou, X. Ren, X. Hong, R. Schaffrath, J.K. Zhu y Z. Gong. 2006. Mutations in ABO1/ELO2, a subunit of holo-Elongator, increase abscisic acid sensitivity and drought tolerance in *Arabidopsis thaliana*. *Molecular and Cellular Biology* 26(18): 6902-6912. doi: 10.1128/MCB.00433-06
- Cheng, Y., W. Zhou, N. I. El Sheery, C. Peters, M. Li, X. Wang y J. Huang. 2011. Characterization of the *Arabidopsis* glycerophosphodiester phosphodiesterase (GDPD) family reveals a role of the plastid- localized AtGDPD1 in maintaining cellular phosphate homeostasis under phosphate starvation. *The Plant Journal* 66(5): 781-795. doi: 10.1111/j.1365-313X.2011.04538.x
- Chisti, Y. 2007. Biodiesel from microalgae. *Biotechnology Advances* 25(3): 294-306. doi: 10.1016/j.biotechadv.2007.02.001
- Clayton, J.R., Q. Dortch, S.S. Thoresen y S.I. Ahmed. 1988. Evaluation of methods for the separation and analysis of proteins and free amino acids in phytoplankton samples. *Journal of Plankton Research*, 10(3): 341-358. doi.org/10.1093/plankt/10.3.341.
- Conesa, A., P. Madrigal, S. Tarazona, D. Gomez-Cabrero, A. Cervera, A. McPherson, M. W. Szczesniak, D.J. Gaffney, L. L. Elo, X. Zhang. 2016. A survey of best practices for RNA-seq data analysis. *Genome Biology* 17(1): 13. doi: 10.1186/s13059-016-0881-8
- Cordero Esquivel, B., D. Voltolina Lobina y F. Correa Sandoval. 1993. Biochemical composition of two diatoms after different preservation techniques. *Comparative Biochemistry and Physiology. B: Comparative biochemistry*. 105B (2): 369–373.
- Córdova-Izquierdo, A., C. G. Ruiz-Lang, J. A. Córdova-Jiménez, M. S. Córdova-Jiménez, J. E. Guerra-Liera, B. E. Rodríguez-Denis y K. Arancibia-Salinas. 2009. Oxidative stress and antioxidants in the spermatic conservation. *Revista Complutense de Ciencias Veterinarias* 3(1): 01-38.

- Core, L. J., J. J. Waterfall y J. T. Lis. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322(5909) 1845-1848. doi: 10.1126/science.1162228
- Cotteau, P. 1996. Micro-algae. 7 – 47 p. En: Lavens, P., y Sorgeloos, P. (Eds.), *Manual on the production and use of live food for aquaculture*. FAO, Ghent, Bélgica.
- Cramer, G. R., K. Urano, S. Delrot, M. Pezzotti y K. Shinozaki. 2011. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biology* 11(1): 1-14. doi: 10.1186/1471-2229-11-163
- Cruz de Carvalho, M. H., H. X. Sun, C. Bowler, N. H. Chua. 2016. Noncoding and coding transcriptome responses of a marine diatom to phosphate fluctuations. *New Phytologist* 210: 497-510. doi: 10.1111/nph.13787
- Cruz-Ramírez, A., A. Oropeza-Aburto, F. Razo-Hernández, E. Ramírez-Chávez y L. Herrera-Estrella. 2006. Phospholipase DZ2 plays an important role in extraplastidic galactolipid biosynthesis and phosphate recycling in *Arabidopsis* roots. *Proceedings of the National Academy of Sciences* 103(17): 6765-6770. doi: 10.1073/pnas.0600863103
- Cuellar- Bermúdez, S. P., I. Aguilar- Hernandez, D. L. Cardenas- Chavez, N. Ornelas- Soto, M. A. Romero- Ogawa y R. Parra- Saldivar. 2015. Extraction and purification of high- value metabolites from microalgae: essential lipids, astaxanthin and phycobiliproteins. *Microbial Biotechnology* 8(2): 190-209. doi: 10.1111/1751-7915.12167
- Cultek. 2006. Extracción y purificación de ácidos nucleicos. [En línea]. [Fecha de consulta: 19 de septiembre de 2015]. Disponible en: [http://www.divulgacion.ccg.unam.mx/files/pdfs/como/Purificacion\\_extraccion\\_DNA-RNA.pdf](http://www.divulgacion.ccg.unam.mx/files/pdfs/como/Purificacion_extraccion_DNA-RNA.pdf)
- Cupp, E. E. 1943. *Marine Plankton Diatoms of the West Coast of North America*. University of California Press. Berkeley, California. 238 p.
- Dar, A. A., A. R. Choudhury, P. K. Kancharla y N. Arumugam. 2017. The FAD2 gene in plants: occurrence, regulation, and role. *Frontiers in Plant Science* 8: 1789. doi: 10.3389/fpls.2017.01789
- De Jesús-Campos, D., J. A. López-Elías, L. Á. Medina-Juarez, G. Carvallo-Ruiz, D. Fimbres-Olivarria y C. Hayano-Kanashiro. 2020. Chemical composition, fatty acid profile and molecular changes derived from nitrogen stress in the diatom *Chaetoceros muelleri*. *Aquaculture Reports* 16: 100281. doi: 10.1016/j.aqrep.2020.100281
- Dell' Aquila, G. y U. G. Maier. 2020. Specific acclimations to phosphorus limitation in the marine diatom *Phaeodactylum tricornutum*. *Biological Chemistry* 401(12): 1495-1501. doi: 10.1515/hsz-2020-0197
- Demirbas, A. 2011. Competitive liquid biofuels from biomass. *Applied Energy* 88(1): 17-28. doi:10.1016/j.apenergy.2010.07.016
- Demirbas, A. 2003. Biodiesel fuels from vegetable oils via catalytic and non-catalytic supercritical alcohol transesterifications and other methods: a survey. *Energy Conversion and Management* 44:2093–2109. doi: 10.1016/S0196-8904(02)00234-0
- De Stefano, M., L. De Stefano y R. Congestri. 2009. Functional morphology of micro- and nanostructures in two distinct diatom frustules. *Superlattices and Microstructures* 46(1-2), 64-68. doi: 10.1016/j.spmi.2008.12.007
- De Stefano, L., I. Rea, I. Rendina, M. De Stefano y L. Moretti. 2007. Lensless light focusing with the centric marine diatom *Coscinodiscus walesii*. *Optics Express* 15 (26): 18082–18088. doi: 10.1364/OE.15.018082

- de Vargas, C., S. Audic, N. Henry, J. Decelle, F. Mahé, R. Logares, E. Lara, C. Berney, N. L. Bescot, I. Probert et al. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348: 1261605. doi:10.1126/science.1261605.
- Dhillon, R.S. y G. von Wuehlich. 2013. Mitigation of global warming through renewable biomass. *Biomass Bioenergy* 48:75 – 89. doi: 10.1016/j.biombioe.2012.11.005
- Diaz, J. M., A. Holland, J. G. Sanders, K. Bulski, D. Mollett, C. W. Chou, D. Phillips, Y. Tang y S. Duhamel. 2018. Dissolved organic phosphorus utilization by phytoplankton reveals preferential degradation of polyphosphates over phosphomonoesters. *Frontiers in Marine Science* 5:380. doi: 10.3389/fmars.2018.00380
- Dill, K. A. y D. Shortle. 1991. Denatured states of proteins. *Annual Review of Biochemistry* 60(1): 795-825. doi: 10.1146/annurev.bi.60.070191.004051
- d'Ippolito, G., A. Sardo, D. Paris, F. M. Vella, M.G. Adelfi, P. Botte, C. Gallo y A. Fontana. 2015. Potential of lipid metabolism in marine diatoms for biofuel production. *Biotechnology for Biofuels* 8(1): 28. doi: 10.1186/s13068-015-0212-4
- Dismukes, G. C., D. Carrieri, N. Bennette, G. M. Ananyev y M. C. Posewitz. 2008. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Current Opinion in Biotechnology* 19(3): 235-240. doi: 10.1016/j.copbio.2008.05.007
- Duong, V. T., Y. Li, E. Nowak y P. M. Schenk. 2012. Microalgae isolation and selection for prospective biodiesel production. *Energies* 5(6): 1835-1849. doi: 10.3390/en5061835
- Drozdetskiy, A., C. Cole, J. Proctery G. J. Barton. 2015. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research* 43(W1): W389–W394. doi: 10.1093/nar/gkv332
- Dunahay, T. G., E. E. Jarvis, S. S. Dais y P. G. Roessler. 1996. Manipulation of microalgal lipid production using genetic engineering. *Applied Biochemistry and Biotechnology* 57(1): 223. doi: 10.1007/978-3-319-24945-2
- Dyhrman, S.T. 2016. Nutrients and their acquisition: phosphorus physiology in microalgae. En: Borowitzka, M.A., Beardall, J. y Raven J.A. (Eds.), *The Physiology of Microalgae*. Springer Dordrecht. doi: 10.1007/978-3-319-24945-2
- Dyhrman, S. T., B. D. Jenkins, T. A. Ryneerson, M. A. Saito, M. L. Mercier, H. Alexander, L.P. Whitney, A. Drzewianowski, V. Bulygin, E.M. Bertrand. 2012. The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PloS one* 7(3): e33768. doi: 10.1371/journal.pone.0033768
- European Committee for Standardization. 2008. EN 14214: automotive fuels – fatty acid methyl esters (FAME) for diesel engines – requirements and test methods. Report no. EN 14214:2008. Management Centre.
- Evans, P.R., G.W. Farrants, P.J. Hudson y H.G. Britton. 1981. Phosphofructokinase: structure and control. *Philosophical Transactions of the Royal Society, B: Biological Sciences* 293: 53–62. doi: 10.1098/rstb.1981.0059
- FDA. 2016. FDA, Summary: Substance Generally Regarded as Safe, U.S. Food and Drug Administration (HHS). [En línea]. [Fecha de consulta: 17 de mayo de 2020]. Disponible en: <https://www.fda.gov/about-fda>
- Feng, T. Y., Z. K. Yang, J. W. Zheng, Y. Xie, D. W. Li, S. B. Murugan, W.D. Yang, J.S. Liu y H. Y. Li. 2015. Examination of metabolic responses to phosphorus limitation via proteomic analyses in the marine diatom *Phaeodactylum tricornutum*. *Scientific Reports* 5: 10373. doi: 10.1038/srep10373

- Fenollar-Ferrer, C. y L. R. Forrest. 2019. Structural models of the NaPi-II sodium-phosphate cotransporters. *Pflügers Archiv-European Journal of Physiology* 471(1): 43-52. doi: 10.1007/s00424-018-2197-x
- Fields, F. J. 2014. An Evolutionary Perspective on Selecting High-Lipid-Accumulating Diatoms (Bacillariophyta): Literature Review, New Data, and Future Prospects Tesis de Doctorado. University of Colorado at Boulder. Boulder, Colorado, Estados Unidos.
- Finenko, Z. Z. y D. K. Krupatkina-Akinina. 1974. Effect of inorganic phosphorus on the growth rate of diatoms. *Marine Biology* 26(3): 193-201.
- Fonseca, J. P. y X. Dong. 2014. Functional characterization of a Nudix hydrolase AtNUDX8 upon pathogen attack indicates a positive role in plant immune responses. *PloS one* 9(12): e114119. doi: 10.1371/journal.pone.0114119
- Forster, I. C., N. Hernando, J. Biber y H. Murer. 2013. Phosphate transporters of the SLC20 and SLC34 families. *Molecular Aspects of Medicine* 34(2-3): 386-395. doi: 10.1016/j.mam.2012.07.007
- Free, R. B., L. A. Hazelwood y D. R. Sibley. 2009. Identifying novel protein- protein interactions using co- immunoprecipitation and mass spectroscopy. *Current Protocols in Neuroscience* 46(1): 5-28. doi: 10.1002/0471142301.ns0528s46
- Fu, H., B. Xie, S. Ma, X. Zhu, G. Fan y S. Pan. 2011. Evaluation of antioxidant activities of principal carotenoids available in water spinach (*Ipomoea aquatica*). *Journal of Food Composition and Analysis* 24(2): 288-297. doi: 10.1016/j.jfca.2010.08.007
- Ganesan, R., S. Manigandan, M. S. Samuel, R. Shanmuganathan, K. Brindhadevi, N. T. L. Chi, P.A. Duc y A. Pugazhendhi. 2020. A review on prospective production of biofuel from microalgae. *Biotechnology Reports: e00509*. doi: 10.1016/j.btre.2020.e00509
- Garcia, E. S., G. Olivieri, L. Sijtsma, M.H.Vermuë, M. Barbosa, J. H. Reith, C. van den Berg, M.H.M. Eppink, R.H. Wijffels. 2019. Integrated Biorefineries for Algal Biomolecules. 293-317 p. En: Hallman, A; Rampelotto, P. (Eds), *Grand Challenges in Algae Biotechnology*. Springer, Cham.
- García-Martín, E. 2017. Relaciones estructura -función del sensor de  $Ca^{2+}$  sos3 de *Arabidopsis thaliana*. Tesis Doctoral. Universidad de Sevilla. Departamento de Bioquímica Vegetal y Biología Molecular. Sevilla, España.
- Garrett, R.H. y C. M. Grisham. 2008. *Biochemistry*. Boston: Twayne Publishers.
- Ge, Y., Y. Chang, W. L. Xu, C. S. Cui y S. P. Qu. 2015. Sequence variations in the FAD2 gene in seeded pumpkins. *GMR, Genetics and Molecular Research* 14: 17482-17488. doi: 10.4238/2015.December.21.19
- Geniza, M. y P. Jaiswal. 2017. Tools for building de novo transcriptome assembly. *Current Plant Biology* 11: 41-45. doi: 10.1016/j.cpb.2017.12.004
- Georgianna, D.R., M.J. Hannon, M.W.S. Marcuschi, K. Botsch, A.J. Lewis, J. Hyun, M. Mendez y S.P. Mayfield. 2013. Production of recombinant enzymes in the marine alga *Dunaliella tertiolecta*. *Algal Research* 2(1): 2-9. doi: 10.1016/j.algal.2012.10.004
- Gerasimaitė, R., S. Sharma, Y. Desfougères, A. Schmidt y A. Mayer. 2014. Coupled synthesis and translocation restrains polyphosphate to acidocalcisome-like vacuoles and prevents its toxicity. *Journal of Cell Science* 127(23): 5093-5104. doi: 10.1242/jcs.159772
- Giakoumis, E. G. y C. K. Sarakatsanis. 2019. A comparative assessment of biodiesel cetane number predictive correlations based on fatty acid composition. *Energies* 12(3): 422. doi: 10.3390/en12030422

- Gibbs, S.P. 1981. The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. *Annals of the New York Academy of Sciences* 361: 193 – 208. doi:10.1111/j.1749-6632.1981.tb54365.x
- Gomez, L. D., C. G. Steele- King y S. J. McQueen- Mason. 2008. Sustainable liquid biofuels from biomass: the writing's on the walls. *New Phytologist* 178(3) 473-485. doi: 10.1111/j.1469-8137.2008.02422.x
- Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29(7): 644-652. doi:10.1038/nbt.1883
- Griffiths, M. J., R. G. Dicks, C. Richardson y S. T. Harrison. 2011. Advantages and challenges of microalgae as a source of oil for biodiesel. *Biodiesel-feedstocks and processing technologies*. 177-200 p. En: Stoytcheva, M., Montero, J., (Eds.), *Biodiesel - Feedstocks and Processing Technologies*, Rijeka: In Tech.
- Guan, L. L., Y. B. Wang, H. Shen, K. Hou, Y. W. Xu y W. Wu. 2012. Molecular cloning and expression analysis of genes encoding two microsomal oleate desaturases (FAD2) from safflower (*Carthamus tinctorius* L.). *Plant Molecular Biology Reporter* 30(1): 139-148. doi: 10.1007/s11105-011-0322-5
- Guarnieri, M. T. y P. T. Pienkos. 2015. Algal omics: unlocking bioproduct diversity in algae cell factories. *Photosynthesis Research* 123(3): 255-263. doi: 10.1007/s11120-014-9989-4
- Guerrini, F., M. Cangini, L. Boni, P. Trost y R. Pistocchi. 2000. Metabolic responses of the diatom *Achnanthes brevipes* (Bacillariophyceae) to nutrient limitation. *Journal of Phycology* 36(5): 882-890. doi: 10.1046/j.1529-8817.2000.99070.x
- Guiry, M.D. y G.M. Guiry. 2016. *AlgaeBase*. [En línea]. [Fecha de consulta: 16 de septiembre de 2020]. Disponible en: <http://www.algaebase.org>
- Guo, S.L., X.Q. Zhao, Y. Tang, C. Wan, M.A. Alam, S.H. Ho, F.W. Bai y J.S. Chang. 2013. Establishment of an efficient genetic transformation system in *Scenedesmus obliquus*. *Journal of Biotechnology* 163(1): 61-68. doi: 10.1016/j.jbiotec.2012.10.020
- Gutiérrez, I. P. 2002. Estimación de los requerimientos alimenticios para el crecimiento de briópodo *Artemia franciscana* alimentado con la diatomea *Chaetoceros muelleri* orientada a la producción masica. Tesis. Universidad de Colima. Manzanillo, Colima, Méjico.
- Harun, R., M. Singh, G. M. Forde y M. K. Danquah. 2010. Bioprocess engineering of microalgae to produce a variety of consumer products. *Renewable and Sustainable Energy Reviews* 14(3): 1037-1047. doi: 10.1016/j.rser.2009.11.004
- Haşimoğlu, C., M. Ciniviz, İ. Özsert, Y. İçingür, A. Parlak y M. S. Salman. 2008. Performance characteristics of a low heat rejection diesel engine operating with biodiesel. *Renewable energy*, 33(7): 1709-1715. doi: 10.1016/j.renene.2007.08.002
- Hayes, G., A. E. Busch, F. Lang, J. Biber y H. Murer. 1995. Protein kinase C consensus sites and the regulation of renal Na/P i-cotransport (NaPi-2) expressed in *Xenopus laevis* oocytes. *Pflügers Archiv* 430(5): 819-824. doi: 10.1007/BF00386181
- Hayes, G., A. Busch, M. Lötscher, S. Waldegger, F. Lang, F. Verrey, J. Biber y Murer, H. 1994. Role of N-linked glycosylation in rat renal Na/Pi-cotransport. *Journal of Biological Chemistry* 269(39): 24143-24149. doi: 10.1016/S0021-9258(19)51060-2

- Helenius, A. 1994. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Molecular Biology of the Cell* 5(3): 253-265. doi: 10.1091/mbc.5.3.253
- Helliwell, K.E., E. Harrison, J. Christie-Oleza, A. Rees, J. Downe, M. Aguilo-Ferretjans, L. Al-Moosawi, C. Brownlee y G. Wheeler. 2020. A novel Ca<sup>2+</sup> signalling pathway co-ordinates environmental phosphorus sensing and nitrogen metabolism in marine diatoms. *bioRxiv*. doi: 10.1101/2020.05.13.090282.
- Herrero, M., J. A. Mendiola, M. Plaza y E. Ibañez. 2013. Screening for bioactive compounds from algae. 833-872 p. En: Lee, J (Ed), *Advanced Biofuels and Bioproducts*. Springer. New York, Estados Unidos de América.
- Hoek, C., D.G. Mann, y H.M. Jahns. 2009 *Algae. An introduction to phycology*. Cambridge University Press, Daryaganj.
- Hoekman, S. K., A. Broch, C. Robbins, E. Cenicerros y M. Natarajan. 2012. Review of biodiesel composition, properties, and specifications. *Renewable and Sustainable Energy Reviews* 16(1): 143-169. doi: 10.1016/j.rser.2011.07.143
- Hopes, A. y T. Mock. 2015. Evolution of microalgae and their adaptations in different marine ecosystems. *eLS*: 1-9. doi: 10.1002/9780470015902.a0023744
- Hothorn, M., Neumann, H., Lenherr, E. D., Wehner, M., Rybin, V., Hassa, P. O., A. Uttenweiler, M. Reinhardt, A. Schmidt, J. Seiler, et al. 2009. Catalytic core of a membrane-associated eukaryotic polyphosphate polymerase. *Science* 324(5926): 513-516. doi: 10.1126/science.1168120
- Hu, Q., M. Sommerfeld, E. Jarvis, M. Ghirardi, M. Posewitz, M. Seibert y A. Darzins. 2008. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal* 54(4): 621-639. doi: 10.1111/j.1365-313X.2008.03492.x
- Huang, C. Y., U. Roessner, I. Eickmeier, Y. Genc, D. L. Callahan, N. Shirley, P. Langridge y A. Bacic. 2008. Metabolite profiling reveals distinct changes in carbon and nitrogen metabolism in phosphate-deficient barley plants (*Hordeum vulgare* L.). *Plant and Cell Physiology* 49(5): 691-703. doi: 10.1093/pcp/pcn044
- Hülse, D., S. Arndt, J. D. Wilson, G. Munhoven y A. Ridgwell. 2017. Understanding the causes and consequences of past marine carbon cycling variability through models. *Earth-science Reviews*, 171: 349-382. doi: 10.1016/j.earscirev.2017.06.004
- Hwangbo, K., J. W. Ahn, J. M. Lim, Y. I. Park, J. R. Liu y W. J. Jeong. 2014. Overexpression of stearyl-ACP desaturase enhances accumulations of oleic acid in the green alga *Chlamydomonas reinhardtii*. *Plant Biotechnology Reports* 8(2): 135-142. doi: 10.1007/s11816-013-0302-3
- Ibraheem, O., R. O. Adigun y I. T. Olatunji. 2018. Omics Technologies in Unraveling Plant Stress Responses; Using Sorghum as a Model Crop, How Far Have We Gone. *Vegetos* 31: 2. doi: 10.4172/2229, 4473.
- Islam, M. A., M. Magnusson, R. J. Brown, G. A. Ayoko, M. Nabi y K. Heimann. 2013. Microalgal species selection for biodiesel production based on fuel properties derived from fatty acid profiles. *Energies* 6(11): 5676-5702. doi:10.3390/en6115676
- Itakura, S. 2000. Physiological ecology of the resting stage cells of coastal planktonic diatoms. *Bulletin of Fisheries and Environment of Inland Sea (Japan)*: 67-130.
- Jain, S. 2020. Marine Diatoms. 131-142 p. En: Jain, J (Ed), *Fundamentals of Invertebrate Palaeontology*. Springer, New Delhi.

- Jarosz, M., M. Van Lijsebettens y M. Woloszynska. 2020. Plant Elongator—Protein Complex of Diverse Activities Regulates Growth, Development, and Immune Responses. *International Journal of Molecular Sciences* 21(18): 6912. doi: 10.3390/ijms21186912
- Jeffrey, S. T. y G. F. Humphrey. 1975. New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae ad natural phytoplankton. *Biochimie und physiologie der pflanzen*, 167(2): 191-194. doi: 10.1016/S0015-3796(17)30778-3
- Jensen, E. L., K. Yangüez, F. Carrière y B. Gontero. 2020. Storage compound accumulation in diatoms as response to elevated CO2 concentration. *Biology* 9(1): 5. doi: 10.3390/biology9010005
- Jitrapakdee, S. y J. C. Wallace. 2003. The biotin enzyme family: conserved structural motifs and domain rearrangements. *Current Protein and Peptide Science* 4(3): 217-229. doi: 10.2174/1389203033487199
- Jitrapakdee, S. y J. C. Wallace. 1999. Structure, function and regulation of pyruvate carboxylase. *Biochemical Journal* 340(1): 1-16. doi: 10.1042/bj3400001
- Jónasdóttir, S. H. 2019. Fatty acid profiles and production in marine phytoplankton. *Marine Drugs* 17(3): 151. doi:10.3390/md17030151
- Juneja, A., R. M. Ceballos y G. S. Murthy. 2013. Effects of environmental factors and nutrient availability on the biochemical composition of algae for biofuels production: a review. *Energies*, 6(9): 4607-4638. doi: 10.3390/en6094607
- Jung, J. Y., M. K. Ried, M. Hothorn y Y. Poirier. 2018. Control of plant phosphate homeostasis by inositol pyrophosphates and the SPX domain. *Current Opinion in Biotechnology* 49: 156-162. doi: 10.1016/j.copbio.2017.08.012
- Kale, A., y B. Karthick. 2015. The diatoms. *Resonance* 20(10): 919-930. doi: 10.1007/s12045-015-0256-6.
- Kanehisa, M., S. Goto, Y. Sato, M. Furumichi y M. Tanabe. 2012. KEGG for integration and interpretation of largescale molecular data sets. *Nucleic Acids Research* 40: 109 – 114. doi: 10.1093/nar/gkr988
- Kargiotidou, A., D. Deli, D. Galanopoulou, A. Tsaftaris y T. Farmaki. 2008. Low temperature and light regulate delta 12 fatty acid desaturases (FAD2) at a transcriptional level in cotton (*Gossypium hirsutum*). *Journal of Experimental Botany* 59(8): 2043-2056. doi: 10.1093/jxb/ern065
- Kavanová, M., F. A. Lattanzi, A. A. Grimoldi y H. Schnyder. 2006. Phosphorus deficiency decreases cell division and elongation in grass leaves. *Plant Physiology* 141(2): 766-775. doi: 10.1104/pp.106.079699
- Khan, A. R., M. T. Pervez, M. E. Babar, N. Naveed y M. Shoaib. 2018. A comprehensive study of de novo genome assemblers: current challenges and future prospective. *Evolutionary Bioinformatics* 14: 1176934318758650. doi: 10.1177/1176934318758650
- Khazin-Goldberg I. 2016. Lipid Metabolism in Microalgae. En: Borowitzka, M.A., Beardall, J. y Raven, J.A. (Eds), *The Physiology of Microalgae*. Springer Cham. doi: 10.1007/978-3-319-24945-2
- Khazin-Goldberg, I. y Z. Cohen. 2011. Unraveling algal lipid metabolism: recent advances in gene identification. *Biochimie* 93(1): 91-100. doi: 10.1016/j.biochi.2010.07.020
- Khrustalev, V. V., T. A. Khrustaleva y E. V. Barkovsky. 2013. Random coil structures in bacterial proteins. Relationships of their amino acid compositions to flanking structures and corresponding genic base compositions. *Biochimie*, 95(9), 1745-1754.

- Kim, R. J., H. J. Kim, D. Shim y M. C. Suh. 2016. Molecular and biochemical characterizations of the monoacylglycerol lipase gene family of *Arabidopsis thaliana*. *The Plant Journal* 85(6): 758-771. doi: 10.1111/tpj.13146
- Kim, D., B. Langmead y S. L. Salzberg. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 12(4): 357-360. doi: 10.1038/nmeth.3317
- Knothe G. 2005. Oxidative stability of biodiesel: literature overview. 122–126 p. En: Knothe, G., Van Gerpen, J., Krahl, J. (Eds.), *The biodiesel handbook*. AOCS Press. Urbana, IL, Estados Unidos.
- Kong, F., I. T. Romero, J. Warakanont y Y. Li- Beisson. 2018. Lipid catabolism in microalgae. *New Phytologist* 218(4): 1340-1348. doi: 10.1111/nph.15047
- Kopylova, E., L. Noé y H. Touzet. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28(24): 3211-3217.
- Krisnangkura, K. 1986. A simple method for estimation of cetane index of vegetable oil methyl esters. *Journal of the American Oil Chemists' Society* 63(4): 552-553. doi: 10.1007/BF02645752
- Kuczynska, P., M. Jemiola-Rzeminska y K. Strzalka. 2015. Photosynthetic pigments in diatoms. *Marine Drugs* 13(9): 5847-5881. doi:10.3390/md13095847
- Kuzminov, A. 2011. Homologous Recombination—Experimental Systems, Analysis and Significance. *EcoSal Plus* 4(2): 1 – 73. doi: 10.1128/ecosalplus.7.2.6
- Kwon, H.K., S.J. Oh y H.S. Yang. 2013. Growth and uptake kinetics of nitrate and phosphate by benthic microalgae for phytoremediation of eutrophic coastal sediments. *Bioresource Technology* 129:387–395. doi: 10.1016/j.biortech.2012.11.078
- Lamers, D., B. Visscher, R. A. Weusthuis, C. Francke, R. H. Wijffels y C. Lokman. 2019. Overexpression of delta-12 desaturase in the yeast *Schwanniomyces occidentalis* enhances the production of linoleic acid. *Bioresource Technology* 289: 121672. doi: 10.1016/j.biortech.2019.121672
- Laurens, L. M. L. 2017. State of Technology Review—Algae Bioenergy. *Golden: IEA Bioenergy*. 1 – 158. doi: 10.13140/RG.2.2.11770.90560.
- Lauritano, C., M.I. Ferrante y A. Rogato. 2019. Marine Natural Products from Microalgae: An -Omics Overview. *Marine Drugs* 17: 269. doi: 10.3390/md17050269
- Lauritano, C. y A. Ianora. 2018. Grand Challenges in Marine Biotechnology: Overview of Recent EU-Funded Projects. 425–449 p. En: Rampellotto, P.H., y Trincone, A. (Eds.), *Grand Challenges in Marine Biotechnology*. Springer, Heidelberg, Alemania.
- Lee, S. H., R. Karawita, A. Affan, J. B. Lee, K. W. Lee, B. J. Lee, D.W. Kim y Y. J. Jeon. 2009. Potential of Benthic Diatoms *Achnanthes longipes*, *Amphora coffeaformis* and *Navicula* sp. (Bacillariophyceae) as Antioxidant Sources. *Algae* 24(1): 47-55. doi: 10.4490/ALGAE.2009.24.1.047
- Lemus, N., T. Urbano, B. Arredondo-Vega, M. Guevara, A. Vásquez, L. Carreón-Palau y N. Vallejo. 2006. Crecimiento y perfil bioquímico de *Chaetoceros muelleri* cultivada en sistemas discontinuos y semicontinuos. *Ciencias marinas* 32(3): 597-603.
- Levitan, O., J. Dinamarca, E. Zelzion, D. S. Lun, L. T. Guerra, M. K. Kim, J. Kim, B. A. S. Van Mooy, D. Bhattacharya y P. G. Falkowski. 2015. Remodeling of intermediate metabolism in the diatom *Phaeodactylum tricorutum* under nitrogen stress. *Proceedings of the National Academy of Sciences* 112(2): 412-417. doi: 10.1073/pnas.1419818112

- Leyva, L. A., Y. Bashan, A. Mendoza y L. E. de-Bashan. 2014. Accumulation fatty acids of in *Chlorella vulgaris* under heterotrophic conditions in relation to activity of acetyl-CoA carboxylase, temperature, and co-immobilization with *Azospirillum brasilense*. *Naturwissenschaften*, 101(10): 819-830. doi: 10.1007/s00114-014-1223-x
- Li, Y., A. Boonprakob, C. C. Gaonkar, W. H. C. F. Kooistra, C. B. Lange, D. Hernández-Becerril, Z. Chen, Ø. Moestrup, N. Lundholm. 2017. Diversity in the globally distributed diatom genus *Chaetoceros* (Bacillariophyceae): three new species from warm-temperate waters. *Plos One* 12(1): e0168887. doi:10.1371/journal.pone.0168887
- Li, J., D. Han, D. Wang, K. Ning, J. Jia, L. Wei, X. Jing, S. Huang, J. Chen, Y. Li et al. 2014. Choreography of transcriptomes and lipidomes of *Nannochloropsis* reveals the mechanisms of oil synthesis in microalgae. *The Plant Cell* 26(4): 1645-1665. doi: 10.1105/tpc.113.121418
- Li, M., Z. Zhang, J. Yuan, Y. Zhang y X. Jin. 2014. Altered glutamate cysteine ligase expression and activity in renal cell carcinoma. *Biomedical Reports* 2(6): 831-834. doi: 10.3892/br.2014.359
- Li, R., Z. Wang, P. Ni, Y. Zhao, M. Li y L. Li. 2014. Effects of cetane number improvers on the performance of diesel engine fuelled with methanol/biodiesel blend. *Fuel* 128: 180-187. doi: 10.1016/j.fuel.2014.03.011
- Li, L., X. Wang, J. Gai y D. Yu. 2007. Molecular cloning and characterization of a novel microsomal oleate desaturase gene from soybean. *Journal of Plant Physiology* 164(11): 1516-1526. doi: 10.1016/j.jplph.2006.08.007
- Li, Q., X. Gao, Y. Sun, Q. Zhang, R. Song y Z. Xu. 2006. Isolation and characterization of a sodium-dependent phosphate transporter gene in *Dunaliella viridis*. *Biochemical and Biophysical Research Communications* 340(1): 95-104. doi: 10.1016/j.bbrc.2005.11.144
- Liang, K., Q. Zhang, M Gu y W. Cong. 2013. Effect of phosphorus on lipid accumulation in freshwater microalga *Chlorella* sp. *Journal of Applied Phycology* 25: 311-318. doi.org/ 10.1007/s10811-012-9865-6.
- Li-Beisson, Y., J. J. Thelen, E. Fedosejevs y J. L. Harwood. 2019. The lipid biochemistry of eukaryotic algae. *Progress in Lipid Research* 74: 31-68. doi: 10.1016/j.plipres.2019.01.003
- Li-Beisson, Y., Y. Nakamura y J. Harwood. 2016. Lipids: from chemical structures, biosynthesis, and analyses to industrial applications. 1–18 p. En: Nakamura, Y., Li-Beisson, Y. (Eds.), *Lipids in plant and algae development*; Springer International Publishing, Cham.
- Lim, S. y P. Kaldis. 2013. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development* 140(15): 3079-3093.
- Lin, Q., W. H. Zhuo, X. W. Wang, C. P. Chen, Y. H. Gao y J. R. Liang. 2018. Effects of fundamental nutrient stresses on the lipid accumulation profiles in two diatom species *Thalassiosira weissflogii* and *Chaetoceros muelleri*. *Bioprocess and Biosystems Engineering* 41(8): 1213-1224. doi: 10.1007/s00449-018-1950-z
- Lin, S., R. W. Litaker y W. G. Sunda. 2016. Phosphorus physiological ecology and molecular mechanisms in marine phytoplankton. *Journal of Phycology* 52(1): 10-36. doi: 10.1111/jpy.12365
- Lira-Morales, J.D., M. Valenzuela-López, M.A. Islas-Osuna, T. Osuna-Enciso, J.A. López-Valenzuela, J.A. Sañudo-Barajas. 2019. Phosphorus transporter proteins from the PHT1 family and their potential use in modern agriculture. *Revista Mexicana Ciencias Agrícolas* 10(5): 1111-1123.

- Litchman E, C.A. Klausmeier, O.M. Schofield y P.G. Falkowski. 2007. The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. *Ecology Letters* 10 (12): 1170–1181. doi: 10.1111/j.1461-0248.2007.01117.x
- Liu, Z. H., T. Li, Q. Y. He, Z. Sun y Y. Jiang. 2018. Role of Mitochondria in Regulating Lutein and Chlorophyll Biosynthesis in *Chlorella pyrenoidosa* under Heterotrophic Conditions. *Marine Drugs* 16(10): 354. doi:10.3390/md16100354
- Liu K, Niu Y., Konishi M., Wu Y., Du H., Sun Chung H., Li L., Boudsocq M., McCormack M., Maekawa S., et al. 2017. Discovery of 633 nitrate–CPK–NLP signalling in central nutrient–growth networks. *Nature* 545:311–316. 634. doi:10.1038/nature22077
- Liu, S., Guo Z.L., T. Li, H. Huang y S.J. Lin. 2011. Photosynthetic efficiency, cell volume, and elemental stoichiometric ratios in *Thalassiosira weissflogii* under phosphorus limitation. *Chinese Journal of Oceanology and Limnology* 29:1048 - 1056. doi: 10.1007/s00343-011-0224-2
- Loera-Quezada, M.M. y E.J. Olguín. 2010. Las microalgas oleaginosas como fuente de biodiesel: retos y oportunidades. *Revista Latinoamericana de Biotecnología Ambiental y Algal* 1: 91-116.
- Lopes, R. G., H. Cella, J. J. Mattos, M. R. F. Marques, A. T. Soares, N. R. Antoniosi Filho, R.B. Derner y L. R. Rörig. 2019. Effect of phosphorus and growth phases on the transcription levels of EPA biosynthesis genes in the diatom *Phaeodactylum tricornutum*. *Revista Brasileira de Botânica* 42:13-22. doi: 10.1007/s40415-019-00515-4
- López-Elías, J.A., D. Voltolina, F. Enríquez-Ocaña y G. Gallegos-Simental. 2005. Indoor and outdoor mass production of the diatom *Chaetoceros muelleri* in a Mexican commercial hatchery. *Aquacultural Engineering* 33: 181-191. doi.org/10.1016/j.aquaeng. 2005.01.001.
- López Elías, J.A., D. Voltolina, C.O. Chavira-Ortega, B.B. Rodríguez-Rodríguez, L.M. Sáenz-Gaxiola, B. Cordero-Esquivel, M. Nieves. 2003. Mass production of microalgae in six commercial shrimp hatcheries of the Mexican northwest. *Aquacultural Engineering* 29(3-4): 155-164. doi: 10.1016/S0144-8609(03)00081-5
- López-Elías, J.A y D. Voltolina. 1993. Semicontinuous cultures of four microalgae with a nonconventional medium. *Ciencias Marinas* 19(2): 169-180. doi.org/ 10.7773/cm.v19i2.930.
- Lowe, R., N. Shirley, M. Bleackley, S. Dolan y T. Shafee. 2017. Transcriptomics technologies. *PLoS Computational Biology* 13(5): e1005457. doi: 10.1371/journal.pcbi.1005457
- Lowry, O.H., N.J. Rosebrough, A.L. Farr y R.J. Randall. 1951. Protein measurement with the Folin phenol reagent. *Journal of Biological Chemistry*, 193(1): 265-275.
- Lu, B., Z. Zeng y T. Shi. 2013. Comparative study of de novo assembly and genome-guided assembly strategies for transcriptome reconstruction based on RNA-Seq. *Science China Life Sciences* 56(2): 143-155. doi: 10.1007/s11427-013-4442-z
- Lü, J., C. Sheahan y P. Fu. 2011. Metabolic engineering of algae for fourth generation biofuels production. *Energy & Environmental Science* 4(7): 2451 - 2466. doi: 10.1039/C0EE00593B
- Lu, S., J. Wang, F. Chitsaz, M. K. Derbyshire, R. C. Geer, N. R. Gonzales, M. Gwadz, D.I. Hurwitz, G.H. Marchler, J.S. Song, et al. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Research* 48(D1): D265-D268. doi: 10.1093/nar/gkz991
- Luo, S., Y. Li, R. Ma, J. Liu, P. Xu, H. Zhang, K. Tang, J. Ma, N. Liu, Y. Zang et al. 2017. Downregulation of PCK2 remodels tricarboxylic acid cycle in tumor-repopulating cells of melanoma. *Oncogene* 36(25): 3609-3617. doi:10.1038/onc.2016.520

- Luo, R., B. Liu, Y. Xie, Z. Li, W. Huang, J. Yuan, G. He, Y. Chen, Q. Pan, Y. Liu et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1(1): 2047-217X. doi: 10.1186/2047-217X-1-18
- Luong, X.G. y M. Conti. 2019. RNA Binding Protein Networks and Translational Regulation in Oocytes. 193–220 p. En: Leung, P.C.K., Qiao, J (Ed.), *Human Reproductive and Prenatal Genetics*; Academic Press/Elsevier: London, UK; San Diego, CA, USA; Cambridge, MA, USA; Oxford, UK. doi: 10.1016/B978-0-12-813570-9.00009-7
- Maeda, Y., D. Nojima, T. Yoshino y T. Tanaka. 2017. Structure and properties of oil bodies in diatoms. *Philosophical Transactions of the Royal Society B: Biological Sciences* 372(1728): 20160408. doi: 10.1098/rstb.2016.0408
- Malara, G. y R. Charra. 1972. Dosage des glucides particuliers de phytoplancton selon la méthode de Dubois. Université de Paris, Station Zoologique de Villefranche-sur-Mer, Notes de Travail, 6: 1-12.
- Malviya, S. E. Scalco, S. Audic, F. Vincent, A. Veluchamy, J. Poulain, P. Wincker, D. Iudicone, C. de Vargas, L. Bittner, A. Zingone et al. 2016. Insights into global diatom distribution and diversity in the world's ocean. *Proceedings of the National Academy of Sciences USA* 113: E1516–E1525. doi:10.1073/pnas.1509523113.
- Mansy, A. E. R. y E. El-Bestawy. 2002. Toxicity and biodegradation of fluometuron by selected cyanobacterial species. *World Journal of Microbiology and Biotechnology* 18(2): 125-131. doi: 10.1023/A:1014490811121
- Martin, J., V. M. Bruno, Z. Fang, X. Meng, M. Blow, T. Zhang, G. Sherlock, M. Snyder y Z. Wang. (2010). Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics* 11(1): 1-8. doi: 10.1186/1471-2164-11-663
- Martínez-Córdova, L. R., A. Campaña-Torres, M. Martinez-Porchas, J. A. López-Elías y C. O. García-Sifuentes. 2012. Efecto de medios alternativos sobre la producción y composición proximal de la microalga *Chaetoceros muelleri* como alimento en cultivo del copépodo *Acartia* sp. *Latin American Journal of Aquatic Research* 40(1): 169-176. doi: 10.3856/vol40-issue1-fulltext-16
- Matallana-Ramírez, L. P., M. Rauf, S. Farage-Barhom, H. Dortay, G. P. Xue, W. Dröge-Laser, A. Lers, S. Balazadeh y B. Mueller-Roeber. 2013. NAC transcription factor ORE1 and senescence-induced BIFUNCTIONAL NUCLEASE1 (BFN1) constitute a regulatory cascade in *Arabidopsis*. *Molecular Plant* 6(5): 1438-1452. doi: 10.1093/mp/sst012
- Mata, T.M., A.A. Martins y N.S. Caetano. 2010. Microalgae for biodiesel production and other applications: A review. *Renewable and Sustainable Energy Reviews* 14:217–232. doi: 10.1016/j.rser.2009.07.020
- McCartney, A. W., J. M. Dyer, P. K. Dhanoa, P. K. Kim, D. W. Andrews, J. A. McNew y R.T. Mullen. 2004. Membrane- bound fatty acid desaturases are inserted co- translationally into the ER and contain different ER retrieval motifs at their carboxy termini. *The Plant Journal* 37(2): 156-173. doi: 10.1111/j.1365-313X.2004.01949.x
- McIlhinney, R. J. 1998. Membrane targeting via protein N-myristoylation. 211-225 p. En: Clegg, R.A. (Ed.), *Protein Targeting Protocols. Methods in Molecular Biology*. Humana Press. doi: 10.1385/0-89603-487-9:211
- McLean, T. I. 2013. “Eco-omics”: a review of the application of genomics, transcriptomics, and proteomics for the study of the ecology of harmful algae. *Microbial Ecology* 65(4): 901-915. doi: 10.1007/s00248-013-0220-5.

- Mehra, P., B. K. Pandey, L. Verma y J. Giri. 2019. A novel glycerophosphodiester phosphodiesterase improves phosphate deficiency tolerance in rice. *Plant, Cell & Environment* 42(4): 1167-1179. doi: 10.1111/pce.13459
- Milner, H.W. 1951. Possibilities in photosynthetic methods for production of oils and proteins. *JAOCs* 28:363–367
- Misson, J., K. G. Raghothama, A. Jain, J. Jouhet, M. A. Block, R. Bligny, P. Ortet, A. Creff, S. Somerville, N. Rolland et al. 2005. A genome-wide transcriptional analysis using Arabidopsis thaliana Affymetrix gene chips determined plant responses to phosphate deprivation. *Proceedings of the National Academy of Sciences* 102(33): 11934-11939. doi: 10.1073/pnas.0505266102
- Mittler, R. 2002. Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science* 7: 405-410. [https://doi.org/10.1016/S1360-1385\(02\)02312-9](https://doi.org/10.1016/S1360-1385(02)02312-9)
- Mohamed, S. A. E. H., M. Elloumi y J. D. Thompson, 2016. Motif Discovery in Protein Sequences. En: Ramakrishnan, S. (Ed.), *Pattern Recognition-Analysis and Applications*. IntechOpen. doi: 10.5772/65441
- Mondal, M., S. Goswami, A. Ghosh, G. Oinam, O.N. Tiwari, P. Das, K. Gayen, M.K. Mandal y G.N. Halder. 2017. Production of biodiesel from microalgae through biological carbon capture: A review. *3 Biotech* 7(2): 99. doi: 10.1007/s13205-017-0727-4
- Mooney, C., N. Davey, A. J. Martin, I. Walsh, D. C. Shields y G. Pollastri. 2011. In silico protein motif discovery and structural analysis. 341-353p. En: Yu B. y M. Hinchcliffe (Eds). *In Silico Tools for Gene Discovery* (pp. 341-353). Humana Press.
- Morgan, D. O. 2007. *The cell cycle: principles of control*. New Science Press Ltd.
- Moseley, J.L., C.W. Chang y A.R. Grossman. 2006. Genome-Based Approaches to Understanding Phosphorus Deprivation Responses and PSR1 Control in *Chlamydomonas reinhardtii*. *Eukaryotic Cell* 5: 26–44. doi: 10.1128/EC.5.1.26-44.2006
- Moult, J. 2008. Comparative modeling in structural genomics. *Structure* 16(1):14-16. doi: 10.1016/j.str.2007.12.001
- Mudge, J. M. y J. Harrow. 2016. The state of play in higher eukaryote gene annotation. *Nature Reviews Genetics* 17(12): 758. doi: 10.1038/nrg.2016.119
- Müller, O., H. Neumann, M. J. Bayer y A. Mayer. 2003. Role of the Vtc proteins in V-ATPase stability and membrane trafficking. *Journal of Cell Science* 116(6): 1107-1115. doi: 10.1242/jcs.00328
- Murillo, A. G., S. Hu y M. L. Fernandez. 2019. Zeaxanthin: metabolism, properties, and antioxidant protection of eyes, heart, liver, and skin. *Antioxidants* 8(9): 390. doi: 10.3390/antiox8090390
- Murphy, D. J. 2012. The dynamic roles of intracellular lipid droplets: from archaea to mammals. *Protoplasma* 249(3): 541-585. doi: 10.1007/s00709-011-0329-7
- Musa, M., G. A. Ayoko, A. Ward, C. Rösch, R. J. Brown y T. J. Rainey. 2019. Factors affecting microalgae production for biofuels and the potentials of chemometric methods in assessing and optimizing productivity. *Cells* 8(8): 851. doi:10.3390/cells8080851
- Mussnug, J.H. 2017. Nuclear Transformation and Toolbox Development. En: Hippler, M. (Ed.). *Chlamydomonas: Molecular Genetics and Physiology*. Springer Cham. doi: 10.1007/978-3-319-66365-4\_2

- Nakov, T., J. M. Beaulieu y A. J. Alverson. 2018. Insights into global planktonic diatom diversity: The importance of comparisons between phylogenetically equivalent units that account for time. *The ISME Journal* 12(11): 2807-2810. doi: 10.1038/s41396-018-0221-y
- Nayeri, F. D. y K. Yarizade. 2014. Bioinformatics study of delta-12 fatty acid desaturase 2 (FAD2) gene in oilseeds. *Molecular Biology Reports* 41(8): 5077-5087. doi: 10.1007/s11033-014-3373-5
- Newton, A. C. 1995. Protein kinase C: structure, function, and regulation. *Journal of Biological Chemistry* 270(48): 28495-28498. doi: 10.1074/jbc.270.48.28495
- Nguyen, V. 2016. Developing tools to genetically engineer the microalga *Nannochloropsis* Tesis de Maestría. Norwegian University of Science and Technology. Trondheim, Ålesund, Gjøvik, Noruega.
- Noriega, G. O., K. B. Balestrasse, A. Batlle y M. L. Tomaro. 2004. Heme oxygenase exerts a protective role against oxidative stress in soybean leaves. *Biochemical and Biophysical Research Communications* 323(3): 1003-1008. doi: 10.1016/j.bbrc.2004.08.199
- Novoa, M. A., D. Miranda y L. Melgarejo. 2018. Efecto de las deficiencias y excesos de fósforo, potasio y boro en la fisiología y el crecimiento de plantas de aguacate (*Persea americana*, cv. Hass). *Revista Colombiana de Ciencias Hortícolas* 12(2): 293-307. doi: 10.17584/rcch.2018v12i2.8092
- OBIS: Ocean Biogeographic Information System. 2012. *Chaetoceros*. [En línea]. [Fecha de consulta: 24 de marzo de 2020]. Disponible en: [http://www.iobis.org/mapper/?taxon\\_id=611649](http://www.iobis.org/mapper/?taxon_id=611649).
- O'Connor, C. M., J. U., Adams y J. Fairman. 2010. *Essentials of Cell Biology*. Cambridge, MA: NPG Education 1: 54.
- Ogawa, N., J. DeRisi y P. O. Brown. 2000. New components of a system for phosphate accumulation and polyphosphate metabolism in *Saccharomyces cerevisiae* revealed by genomic expression analysis. *Molecular Biology of the Cell* 11(12): 4309-4321. doi: 10.1091/mbc.11.12.4309
- Okuley, J., J. Lightner, K. Feldmann, N. Yadav y E. Lark. 1994. *Arabidopsis* FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *The Plant Cell* 6(1): 147-158. doi: 10.1105/tpc.6.1.147
- Olabi, A. G. 2013. State of the art on renewable and sustainable energy. *Energy* 61: 2-5. doi:10.1016/j.energy.2013.10.013
- Olguín, E. J. 2012. Dual purpose microalgae–bacteria-based systems that treat wastewater and produce biodiesel and chemical products within a Biorefinery. *Biotechnology Advances* 30(5): 1031-1046. doi: 10.1016/j.biotechadv.2012.05.001
- Oswald, M., M. Fischer, N. Dirninger y F. Karst. 2007. Monoterpenoid biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Research* 7(3): 413-421. doi: 10.1111/j.1567-1364.2006.00172.x
- Pacheco-Vega, J. M., M. A. Cadena-Roa, M. P. Sánchez-Saavedra, D. Tovar-Ramírez y C. Rangel-Dávalos. 2010. Effect of culture medium and nutrient concentration on fatty acid content of *Chaetoceros muelleri*. *Revista Latinoamericana de Biotecnología Ambiental y Algal* 1: 6-15.
- Pacheco-Vega, J. M., y M. D. P. Sánchez-Saavedra. 2009. The biochemical composition of *Chaetoceros muelleri* (Lemmermann Grown) with an agricultural fertilizer. *Journal of the World Aquaculture Society* 40(4): 556-560. doi: 10.1111/j.1749-7345.2009.00276.x
- Pachiappan, P., P. Santhanam, A. Begum y B. B. Prasath. 2019. An introduction to plankton. 1-24 p. En: Santhanam, P., A. Begum y P. Pachiappan (Eds.), *Basic and Applied Phytoplankton Biology*. Springer, Singapore.

- Page, R.A., S. Okada y J.L. Harwood. 1994. Acetyl-CoA carboxylase exerts strong flux control over lipid synthesis in plants. *Biochimica et Biophysica Acta—Lipids and Lipid Metabolism* 1210(3): 369–372. doi: 10.1016/0005-2760(94)90242-9
- Pahlow, M., U. Riebesell y D.A. Wolf-Gladrow. 1997. Impact of cell shape and chain formation on nutrient acquisition by marine diatoms. *Limnology and Oceanography* 42 (8): 1660–1672. doi: 10.4319/lo.1997.42.8.1660
- Paliwal, C., M. Mitra, K. Bhayani, S. V. Bharadwaj, T. Ghosh, S. Dubey y S. Mishra. 2017. Abiotic stresses as tools for metabolites in microalgae. *Bioresource Technology* 244: 1216-1226. doi: 10.1016/j.biortech.2017.05.058
- Pande, S.V., R.P. Khan y T.A. Venkatasubramanian. 1963. Microdetermination of lipids and serum total fatty acids. *Analytical Biochemistry* 6(5): 415-423. doi.org/10.1016/0003-2697(63)90094-0.
- Pandey, M. K., M. L. Wang, L. Qiao, S. Feng, P. Khera, H. Wang, B. Tonnis, N.A. Barkley, J. Wang, C.C. Holbrook et al. 2014. Identification of QTLs associated with oil content and mapping FAD2 genes and their relative contribution to oil quality in peanut (*Arachis hypogaea*L.). *BMC Genetics* 15(1): 133. doi: 10.1186/s12863-014-0133-4
- Park, Y., Z. Y. Xu, S. Y. Kim, J. Lee, B. Choi, J. Lee, H. Kim, H.J. Sim y Hwang, I. 2016. Spatial regulation of ABCG25, an ABA exporter, is an important component of the mechanism controlling cellular ABA levels. *The Plant Cell* 28(10): 2528-2544. doi: 10.1105/tpc.16.00359
- Parker, M. S., T. Mock y E. V. Armbrust. 2008. Genomic insights into marine microalgae. *Annual Review of Genetics* 42: 619-645. doi: 10.1146/annurev.genet.42.110807.091417
- Patterson, H. B. W. 2011. Quality and Control. 329-350 p. En: List, G. R. y King J. W. (Eds.), *Hydrogenation of Fats and Oils*. AOCS Press. doi: 10.1016/B978-1-893997-93-6.50018-X
- Pawłowicz, I., A. Kosmala y M. Rapacz. 2012. Expression pattern of the psbO gene and its involvement in acclimation of the photosynthetic apparatus during abiotic stresses in *Festuca arundinacea* and *F. pratensis*. *Acta physiologiae plantarum* 34(5): 1915-1924. doi: 10.1007/s11738-012-0992-0
- Pereira, M. A., E. L. Imada y R. L. M. Guedes. 2017. RNA- seq: Applications and Best Practices. 3-36 p. En: Marchi, F.A., P.D.R Cirillo y E.C. Mateo (Eds.), *Applications of RNA-Seq and Omics Strategies-From Microorganisms to Human Health*. InTech, Croatia. doi: 10.5772/intechopen.69250
- Perez, A. y T. P. J. Linsinger. 2016. CERTIFICATION REPORT The certification of the cold filter plugging point (CFPP) and the cloud point (CP). *European Reference Materials*. doi: 10.2787/434305
- Pérez-García, O., F.M. Escalante, L.E. de-Bashan y Y. Bashan. 2011. Heterotrophic cultures of microalgae: metabolism and potential products. *Water Research* 45(1): 11-36. doi: 10.1016/j.watres.2010.08.037
- Petersen, J., A. K. Ludewig, V. Michael, B. Bunk, M. Jarek, D. Baurain y H. Brinkmann. 2014. *Chromera velia*, endosymbioses and the rhodoplex hypothesis—plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). *Genome Biology and Evolution* 6(3): 666-684. doi: 10.1093/gbe/evu043
- Piredda, R., J.M Claverie, J. Decelle, C. de Vargas, M. Dunthorn, B. Edvardsen, W. Eikrem, D. Forster, W. H. C. F. Kooistra, R. Logares et al. 2018. Diatom diversity through HTS-metabarcoding in coastal European seas. *Scientific Reports* 8(1): 1-12. doi: 10.1038/s41598-018-36345-9

- Pirt, S. J., Y. K. Lee, A. Richmond y M. W. Pirt. 1980. The photosynthetic efficiency of *Chlorella* biomass growth with reference to solar energy utilisation. *Journal of Chemical Technology and Biotechnology* 30(1): 25-34. doi: 10.1002/jctb.503300105
- Powell, S., D. Szklarczyk, K. Trachana, A. Roth, M. Kuhn, J. Muller, R. Arnold, T. Rattei, I. Letunic y T. Doerks. 2011. eggNOG v3. 0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Research* 40: D284–D289. doi: 10.1093/nar/gkr1060
- Pramanik, K., P. K. Ghosh, S. Ray, A. Sarkar, S. Mitra y T. K. Maiti. 2017. An in silico structural, functional and phylogenetic analysis with three-dimensional protein modeling of alkaline phosphatase enzyme of *Pseudomonas aeruginosa*. *Journal of Genetic Engineering and Biotechnology* 15(2): 527-537. doi: 10.1016/j.jgeb.2017.05.003.
- Qi, W., S. A. Baldwin, S. P. Muench, A. Baker. 2016. Pi sensing and signalling: from prokaryotic to eukaryotic cells. *Biochemical Society Transactions* 44(3): 766-773. doi: 10.1042/BST20160026
- Quintana-Blanco, W. A., E. H. Pinzón Sandoval y D. F. Torres. 2017. Efecto de un fosfato térmico sobre el crecimiento y producción de frijol (*Phaseolus vulgaris* L.) CV ICA CERINZA. *Revista UDCA Actualidad & Divulgación Científica* 20(1): 51-59.
- Raheem, A., P. Prinsen, A. K. Vuppaladadiyam, M. Zhao y R. Luque. 2018. A review on sustainable microalgae based biofuel and bioenergy production: Recent developments. *Journal of Cleaner Production* 181: 42-59. doi: 10.1016/j.jclepro.2018.01.125
- Rahmadi, A., Y. Mulyani y M. W. Lewaru. 2020. Effect of salinity difference on lipid content from *Chaetoceros muelleri* on continuous reactors. *Advanced Journal of Graduate Research* 7(1): 3-10. doi: 10.21467/ajgr.7.1.3-10
- Raja, R., S. Hemaiswarya, N. A. Kumar, S. Sridhar y R. Rengasamy. 2008. A perspective on the biotechnological potential of microalgae. *Critical Reviews in Microbiology* 34(2): 77-88. doi: 10.1080/10408410802086783
- Ramanan, R., B. H. Kim, D. H. Cho, H. M. Oh y H. S. Kim. 2016. Algae–bacteria interactions: evolution, ecology and emerging applications. *Biotechnology Advances* 34(1): 14-29. doi: 10.1016/j.biotechadv.2015.12.003
- Ramos, M.J., C.M. Fernández, A. Casas, L. Rodríguez y A. Pérez. 2009. Influence of fatty acids composition of raw materials on biodiesel properties. *Bioresource Technology* 100:261–268. doi: 10.1016/j.biortech.2008.06.039
- Ran, W., H. Wang, Y. Liu, M. Qi, Q. Xiang, C. Yao, Y. Zhang y X. Lan. 2019. Storage of starch and lipids in microalgae: Biosynthesis and manipulation by nutrients. *Bioresource Technology* 291: 121894. doi: 10.1016/j.biortech.2019.121894
- Raven, P. 2005. *Biology*. Boston: Twayne Publishers.
- Ravindran, B., M.B. Kurade, A.N. Kabra, B.H. Jeon y S.K. Gupta. 2017. Recent Advances and Future Prospects of Microalgal Lipid Biotechnology. En: Gupta, S.K., Malik A y Bux F(eds). *Algal Biofuels*. Springer Cham. doi: 10.1007/978-3-319-51010-1
- Read, B. A., J. Kegel, M. J. Klute, A. Kuo, S. C. Lefebvre, F. Maumus, C. Mayer, J. Miller, A. Monier, A.Salamov. 2013. Pan genome of the phytoplankton *Emiliana underpins* its global distribution. *Nature* 499(7457): 209-213. doi: 10.1038/nature12221
- Reche, P., Y. L. Li, C. Fuller, K. Eichhorn y N. R. Perham. 1998. Selectivity of post-translational modification in biotinylated proteins: the carboxy carrier protein of the acetyl-CoA carboxylase of *Escherichia coli*. *Biochemical Journal* 329(3): 589-596. doi: 10.1042/bj3290589

- Reimand, J., R. Isserlin, V. Voisin, M. Kucera, C. Tannus-Lopes, A. Rostamianfar, L. Wadi, M. Meyer, J. Wong, C. Xu, et al. 2019. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols* 14(2): 482-517. doi: 10.1038/s41596-018-0103-9
- Reinhold, C., S. Niczyporuk, K. C. Beran y P. Jahns. 2008. Short-term down-regulation of zeaxanthin epoxidation in *Arabidopsis thaliana* in response to photo-oxidative stress conditions. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1777(5): 462-469. doi: [10.1016/j.bbabi.2008.03.002](https://doi.org/10.1016/j.bbabi.2008.03.002)
- Rhee, S. G. 2016. Overview on peroxiredoxin. *Molecules and Cells* 39(1): 1. doi: 10.14348/molcells.2016.2368
- Robinson, M. D., D. J. McCarthy y G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1): 139-140. doi: 10.1093/bioinformatics/btp616
- Rodrigues, S. V., M.M. Marinob, C.C. Cubas Jonck, E.S. Gonçalves, V. F. Brant R. Paranhos, M. P. Curbelo A. P. Falcão. 2014. Phytoplankton community structures in shelf and oceanic waters off southeast Brazil (20–25 S), as determined by pigment signatures. *Deep Sea Research Part I: Oceanographic Research Papers* 88: 47-62. doi: 10.1016/j.dsr.2014.03.006
- Roesler, K., D. Shintani, L. Savage, S. Boddupalli y J. Ohlrogge. 1997. Targeting of the *Arabidopsis* homomeric acetyl-coenzyme A carboxylase to plastids of rapeseeds. *Plant Physiology* 113(1): 75-81. doi: 10.1104/pp.113.1.75
- Rolletschek, H., Borisjuk, L., A. Sánchez-García, C. Gotor, L. C. Romero, J. M. Martínez-Rivas y M. Mancha. 2007. Temperature-dependent endogenous oxygen concentration regulates microsomal oleate desaturase in developing sunflower seeds. *Journal of Experimental Botany* 58(12): 3171-3181. doi: 10.1093/jxb/erm154
- Romanenko, E. A., I. V. Kosakovskaya y P. O. Romanenko. 2015. Phytohormones of microalgae: biological role and involvement in the regulation of physiological processes. Pt I. auxins, abscisic acid, ethylene. *International Journal on Algae* 17(3). doi: 10.1615/InterJAlgae.v17.i3.80
- Rooney, P. J., L. Ayong, C. M. Tobin, S. N. Moreno y L. J. Knoll. 2011. TgVTC2 is involved in polyphosphate accumulation in *Toxoplasma gondii*. *Molecular and Biochemical Parasitology* 176(2): 121-126. doi: 10.1016/j.molbiopara.2010.12.012
- Roopnarain, A., V. M. Gray y S. D. Sym. 2014. Phosphorus limitation and starvation effects on cell growth and lipid accumulation in *Isochrysis galbana* U4 for biodiesel production. *Bioresource Technology* 156: 408-411.
- Roselet, F., D. Vandamme, K. Muylaert y P. C. Abreu. 2019. Harvesting of microalgae for biomass production. 211-243 p. En: Alam, M.A., y Wang Z. (Eds.), *Microalgae Biotechnology for Development of Biofuel and Wastewater Treatment*. Springer, Singapore. doi: 10.1007/978-981-13-2264-8\_10.
- Ryckebosch, E., K. Muylaert y I. Foubert. 2012. Optimization of an analytical procedure for extraction of lipids from microalgae. *Journal of the American Oil Chemists' Society* 89(2): 189-198. doi: 10.1007/s11746-011-1903-z
- Sah, S. K., K. R. Reddy y J. Li. 2016. Abscisic acid and abiotic stress tolerance in crop plants. *Frontiers in Plant Science* 7: 571. doi: 10.3389/fpls.2016.00571
- Sahraeian, S. M. E., M. Mohiyuddin, R. Sebra, H. Tilgner, P. T. Afshar y K. F. Au. 2017. Gaining comprehensive biological insight into the transcriptome by performing a broad-spectrum RNA-seq analysis. *Nature Communications* 8 (1): 59. doi: 10.1038/s41467-017-00050-4

- Sangrador-Vegas, A. y A. Mitchell. 2011. Protein classification. An introduction to EMBL-EBI resources. European Bioinformatics Institute (EMBL-EBI). doi: 10.6019/tol.prc.2011.00001.1
- Sarin, A., R. Arora, N. P. Singh, R. Sarin, R. K. Malhotra y K. Kundu. 2009. Effect of blends of Palm-Jatropha-Pongamia biodiesels on cloud point and pour point. *Energy* 34(11): 2016-2021. doi: 10.1016/j.energy.2009.08.017
- Sasaki, Y. y Y. Nagano. 2004. Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Bioscience, Biotechnology, and Biochemistry* 68(6): 1175-1184. doi: 10.1271/bbb.68.1175.
- Scherer, M. D., A. C. de Oliveira, F. J. C. Magalhães Filho, C. M. L. Ugaya, A. B. Mariano y J. V. C. Vargas. 2017. Environmental study of producing microalgal biomass and bioremediation of cattle manure effluents by microalgae cultivation. *Clean Technologies and Environmental Policy* 19(6): 1745-1759. doi: 10.1007/s10098-017-1361-x
- Schroeder, A., O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow y T. Ragg. 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology* 7(1): 1-14. doi: 10.1186/1471-2199-7-3
- Schurch, N. J., P. Schofield, M. Gierliński, C. Cole, A. Sherstnev, V. Singh, N. Wrobel, K. Gharbi, G.G. Simpson, T. Owen-Hughes et al. 2016. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *Rna* 22(6): 839-851. doi: 10.1261/rna.053959.115
- Secco, D., C. Wang, H. Shou y J. Whelan. 2012. Phosphate homeostasis in the yeast *Saccharomyces cerevisiae*, the key role of the SPX domain-containing proteins. *FEBS letters* 586(4): 289-295. doi: 10.1016/j.febslet.2012.01.036
- Shanklin, J., E. Whittle y B. G. Fox. 1994. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* 33(43): 12787-12794. doi: 10.1021/bi00209a009
- Sharma, U., S. K. Singh, D. Pal, R. Khajuria, A. K. Mandal y R. Prasad. 2012. Implication of BBM lipid composition and fluidity in mitigated alkaline phosphatase activity in renal cell carcinoma. *Molecular and Cellular Biochemistry* 369(1-2): 287-293. doi: 10.1007/s11010-012-1391-y
- Shemi, A., D. Schatz, H. F. Fredricks, B. A. Van Mooy, Z. Porat y A. Vardi. 2016. Phosphorus starvation induces membrane remodeling and recycling in *Emiliania huxleyi*. *New Phytologist* 211: 886-898. doi: 10.1111/nph.13940
- Shih, C. Y., L. K. Kang y J. Chang. 2015. Transcriptional responses to phosphorus stress in the marine diatom, *Chaetoceros affinis*, reveal characteristic genes and expression patterns in phosphorus uptake and intracellular recycling. *Journal of Experimental Marine Biology and Ecology* 470: 43-54. doi: 10.1016/j.jembe.2015.05.001
- Shuba, E. S., y D. Kifle. 2018. Microalgae to biofuels: 'Promising' alternative and renewable energy, review. *Renewable and Sustainable Energy Reviews* 81: 743-755. doi: 10.1016/j.rser.2017.08.042
- Sierra, E., F. G. Ación, J. M. Fernández, J. L. García, C. González y E. Molina. 2008. Characterization of a flat plate photobioreactor for the production of microalgae. *Chemical Engineering Journal* 138(1-3): 136-147. doi: 10.1016/j.cej.2007.06.004
- Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva y E. M. Zdobnov. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19): 3210-3212. doi: 10.1093/bioinformatics/btv351

- Singh, N., S. Upadhyay, A. Jaiswar y N. Mishra. 2016. In silico analysis of protein. *JSM Bioinformatics, Genomics and Proteomics* 1(2): 1007.
- Singh, S. K., G. Badgujar, V. R. Reddy, D. H. Fleisher y J. A. Bunce. 2013. Carbon dioxide diffusion across stomata and mesophyll and photo-biochemical processes as affected by growth CO<sub>2</sub> and phosphorus nutrition in cotton. *Journal of Plant Physiology* 170(9): 801-813. doi: 10.1016/j.jplph.2013.01.001
- Singh, A., P. S. Nigam y J. D. Murphy. 2011. Renewable fuels from algae: an answer to debatable land based fuels. *Bioresource Technology* 102(1): 10-16. doi: 10.1016/j.biortech.2010.06.032
- Siron, R., G. Giusti y B. Berland. 1989. Changes in the fatty acid composition of *Phaeodactylum tricornutum* and *Dunaliella tertiolecta* during growth and under phosphorus deficiency. *Marine Ecology: Progress Series* 55: 95-100.
- Smith-Unna, R., C. Bournnell, R. Patro, J. M. Hibberd y S. Kelly. 2016. TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Research* 26(8): 1134-1144. doi: 10.1101/gr.196469.115
- Solovchenko, A. E., T. T. Ismagulova, A. A. Lukyanov, S. G. Vasilieva, I. V. Konyukhov, S. I. Pogosyan, E.S. Lobakova, O. A. Gorelova. 2019a. Luxury phosphorus uptake in microalgae. *Journal of Applied Phycology* 31: 2755-2770. <https://doi.org/10.1007/s10811-019-01831-8>
- Solovchenko, A., I. Khozin-Goldberg, I. Selyakh, L. Semenova, T. Ismagulova, A. Lukyanov, I. Mamedov, E. Vinogradova, O. Karpova, I. Konyukhov, et al. 2019b. Phosphorus starvation and luxury uptake in green microalgae revisited. *Algal Research* 43: 101651. doi: 10.1016/j.algal.2019.101651
- Sosa-Hernández, J.E., K.D. Romero-Castillo, L. Parra-Arroyo, M.A. Aguilar-Aguila-Isaías, I.E. García-Reyes, I. Ahmed, R. Parra-Saldivar, M. Bilal, H. Iqbal. 2019. Mexican Microalgae Biodiversity and State-Of-The-Art Extraction Strategies to Meet Sustainable Circular Economy Challenges: High-Value Compounds and Their Applied Perspectives. *Marine Drugs* 17: 174. doi:10.3390/md17030174
- Stansell, G. R., V. M. Gray y S. D. Sym. 2012. Microalgal fatty acid composition: implications for biodiesel quality. *Journal of Applied Phycology* 24(4): 791-801. doi: 10.1007/s10811-011-9696-x
- Stayton, M. M., F. B. Rudolph y H. J. Fromm. 1983. Regulation, genetics, and properties of adenylosuccinate synthetase: a review. 103-141 p. En: Horecker, B. L., Stadtman, E. R. (Eds.), *Current topics in cellular regulation*. Academic Press. doi: 10.1016/B978-0-12-152822-5.50008-7.
- Stigter, K. A. y W. C. Plaxton. 2015. Molecular mechanisms of phosphorus metabolism and transport during leaf senescence. *Plants* 4(4): 773-798. doi: 10.3390/plants4040773
- Størseth, T. R., K. Hansen, K. I. Reitan, y J. Skjermo. 2005. Structural characterization of  $\beta$ -D-(1 $\rightarrow$ 3)-glucans from different growth phases of the marine diatoms *Chaetoceros mülleri* and *Thalassiosira weissflogii*. *Carbohydrate research*, 340(6), 1159-1164. doi: 10.1016/j.carres.2004.12.036
- Suganya, T., M. Varman, H. H. Masjuki y S. Renganathan. 2016. Macroalgae and microalgae as a potential source for commercial applications along with biofuels production: a biorefinery approach. *Renewable and Sustainable Energy Reviews* 55: 909 - 941. doi: 10.1016/j.rser.2015.11.026
- Talebi, A. F., S. K. Mohtashami, M. Tabatabaei, M. Tohidfar, A. Bagheri, M. Zeinalabedini, H.H. Mirzaei, M. Mirzajanzadeh, S. M. Shafaroudi. 2013. Fatty acids profiling: a selective criterion for

- screening microalgae strains for biodiesel production. *Algal Research* 2(3): 258-267. doi: 10.1016/j.algal.2013.04.003
- Tang, S., A. Lomsadze y M. Borodovsky. 2015. Identification of protein coding regions in RNA transcripts. *Nucleic Acids Research* 43(12): e78-e78. doi: 10.1093/nar/gkv227
- Thomsen, M. C. F. y M. Nielsen. 2012. Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Research* 40(1):281-287. doi: 10.1093/nar/gks469
- Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn y L. Pachter. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 7(3): 562-578. doi: 10.1038/nprot.2012.016
- Tusher, V. G., R. Tibshirani y G. Chu. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences* 98(9): 5116-5121. doi: 10.1073/pnas.091062498
- Udenwobebe, D. I., R. C. Su, S. V. Good, T. B. Ball, S. Varma Shrivastav y A. Shrivastav. 2017. Myristoylation: An important protein modification in the immune response. *Frontiers in Immunology* 8: 751. doi: 10.3389/fimmu.2017.00751
- Uttenweiler, A., H. Schwarz, H. Neumann y A. Mayer. 2007. The vacuolar transporter chaperone (VTC) complex is required for microautophagy. *Molecular Biology of the Cell* 18(1): 166-175. doi: 10.1091/mbc.e06-08-0664
- Van Mooy, B.A.S., G. Rocap, H.F. Fredricks, C.T. Evans y A.H. Devol. 2009. Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proceedings of the National Academy of Sciences of the United States of America* 103: 8607–8612. doi: 10.1073/pnas.0600540103
- Velasco, L. A., S. Carrera y J. Barros. 2016. Isolation, culture and evaluation of *Chaetoceros muelleri* from the Caribbean as food for the native scallops, *Argopecten nucleus* and *Nodipecten nodosus*. *Latin American Journal of Aquatic Research* 44(3): 557-568. doi: 10.3856/vol44-issue3-fulltext-14
- Velazquez-Lucio, J., R.M. Rodríguez-Jasso, L.M. Colla, A. Sáenz-Galindo, D.E. Cervantes-Cisneros, C.N. Aguilar, B.D. Fernandes, H.A. Ruiz. 2018. Microalgal biomass pretreatment for bioethanol production: a review. *Biofuel Research Journal* 5: 780–791. doi:10.18331/BRJ2018.5.1.5.
- Ventura, S.P.M., B.P. Nobre, F. Ertekin, M. Hayes, M. García-Vaquero, F. Vieira, M. Koc, L. Gouveia, M.R. Aires-Barros y A.M.F. Palavra. 2017. Extraction of value-added compounds from microalgae. 461–483 p. En: Gonzalez-Fernandez, C., Muñoz, R. (Eds.), *Microalgae-Based Biofuels and Bioproducts*; Elsevier: Reino Unido.
- Vigani, M., C. Parisi, E. Rodríguez-Cerezo, M.J. Barbosa, L. Sijtsma, M. Ploeg, C. Enzing. 2015. Food and feed products from micro-algae: Market opportunities and challenges for the EU. *Trends in Food Science and Technology* 42: 81–92. doi: 10.1016/j.tifs.2014.12.004
- Villarruel-López, A., F. Ascencio y K. Nuño. 2017. Microalgae, a potential natural functional food source—a review. *Polish Journal of Food and Nutrition Sciences* 67(4): 251-264. doi: 10.1515/pjfn-2017-0017
- Vingiani, G. M., P. De Luca, A. Ianora, A. D. Dobson y C. Lauritano. 2019. Microalgal enzymes with biotechnological applications. *Marine Drugs* 17(8): 459. doi:10.3390/md17080459

- Vuppaladadiyam, A. K., P. Prinsen, A. Raheem, R. Luque y M. Zhao. 2018. Microalgae cultivation and metabolites production: a comprehensive review. *Biofuels, Bioproducts and Biorefining* 12(2): 304-324. doi: 10.1002/bbb.1864
- Wagner, C. A., G. Raber, S. Waldegger, H. Oddwald, J. Biber, H. Murer, A. E. Busch y F. Lang. 1996. Regulation of the human brush border Na<sup>+</sup>/phosphate cotransporter (NaPi- 3) expressed in *Xenopus* oocytes by intracellular calcium and protein kinase C. *Cell Physiology and Biochemistry* 6: 105–111. doi: 10.1159/000154799
- Walker, J.M. 2005. Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*. Humana Press. 607p.
- Wang, C., Ying, S., Huang, H., Li, K., Wu, P., & Shou, H. (2009). Involvement of OsSPX1 in phosphate homeostasis in rice. *The Plant Journal*, 57(5), 895-904.
- Wang, F., M. Deng, J. Xu, X. Zhu y C. Mao. 2018. Molecular mechanism of phosphate signaling in plants. *Seminars in Cell & Developmental Biology* 74: 114–122. doi: 10.1016/j.semcdb.2017.06.013
- Wang, J., Q. Qin, J. Pan, L. Sun, Y. Sun, Y. Xue y K. Song. 2019. Transcriptome analysis in roots and leaves of wheat seedlings in response to low-phosphorus stress. *Scientific Reports* 9(1): 1-12. doi: 10.1038/s41598-019-56451-6
- Wang, L.J., Y. Fan, R.L. Parsons, G.R. Hu, P.Y. Zhang y F.L. Li. 2018. A rapid method for the determination of fucoxanthin in diatom. *Marine Drugs* 16(33): 1-13. doi:10.3390/md16010033
- Wang, X.W, J.R. Liang, C.S. Luo, C.P. Chen y Y.H. Gao. 2014. Biomass, total lipid production, and fatty acid composition of the marine diatom *Chaetoceros muelleri* in response to different CO<sub>2</sub> levels. *Bioresource Technology*, 161: 124-130. doi.org/10.1016/j.biortech.2014.03.012
- Wang, L. B., H. Y. Yu, X. H. He y R. Y. Liu. 2012. Influence of fatty acid composition of woody biodiesel plants on the fuel properties. *Journal of Fuel Chemistry and Technology* 40(4): 397-404. doi: 10.1016/S1872-5813(12)60018-8
- Wang, C., S. Ying, H. Huang, K. Li, P. Wu y H. Shou. 2009. Involvement of OsSPX1 in phosphate homeostasis in rice. *The Plant Journal* 57(5): 895-904. doi: 10.1111/j.1365-3113X.2008.03734.x
- Wang, Z., M. Gerstein y M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10:57 – 63. doi: 10.1038/nrg2484
- Warakanont, J., Y. Li-Beisson y C. Benning. 2019. LIP4 is involved in triacylglycerol degradation in *Chlamydomonas reinhardtii*. *Plant and Cell Physiology* 60(6): 1250-1259. doi: 10.1093/pcp/pcz037
- Waterhouse, A.M., J.B. Procter, D.M.A Martin, M. Clamp y G. J Barton. 2009. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. *Bioinformatics* 25(9):1189-1191. doi: 10.1093/bioinformatics/btp033
- Weeks, D.P. 2011. Homologous recombination in *Nannochloropsis*: A powerful tool in an industrially relevant alga. *Proceedings of the National Academy of Sciences* 108(52): 20859-20860. doi: 10.1073/pnas.1118670109
- Wijffels, R. H. 2015. The need and risks of using transgenic microalgae for the production of food, feed, chemicals and fuels. *Biosafety and the environmental uses of micro-organisms*. Netherlands 6071.
- Wilhelm, C., C. Büchel, J. Fisahn, R. Goss, T. Jakob, J. LaRoche, J. Lavaud, M. Lohr, U. Riebesell, K. Stehfest. 2006. The regulation of carbon and nutrient assimilation in diatoms is significantly different from green algae. doi: 10.1016/j.protis.2006.02.003

- Williams, D.M. y J.P. Kocielek. 2011. An overview of diatom classification with some prospects for the future. 47–91 p. En: Seckbach, J. y Kocielek, J.P. (Eds), Cellular origin, life in extreme habitats and astrobiology volume 19: the diatom world. Springer Publishing House, London.
- Wood, A. M., R. C. Everroad y L. M. Wingard. 2005. Measuring growth rates in microalgal cultures. 269-288 p. En: Andersen, R.A. (Ed.), Algal Culturing Techniques. Academic Press, China.
- Worden, A. Z., J. H. Lee, T. Mock, P. Rouzé, M. P. Simmons, A. L. Aerts, A. E. Allen, M. L. Cuvelier, E. Derelle, M. V. Everet et al. 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324(5924): 268-272. doi: 10.1126/science.1167222
- Wu, Q., Y. Cao, C. Chen, Z. Gao, F. Yu y R. D. Guy. 2020. Transcriptome analysis of metabolic pathways associated with oil accumulation in developing seed kernels of *Styrax tonkinensis*, a woody biodiesel species. *BMC Plant Biology* 20(1): 1-17. doi: 10.1186/s12870-020-2327-4
- Wu, P., S. Zhang, L. Zhang, Y. Chen, M. Li, H. Jiang y G. Wu. 2013. Functional characterization of two microsomal fatty acid desaturases from *Jatropha curcas* L. *Journal of Plant Physiology* 170(15): 1360-1366. doi: 10.1016/j.jplph.2013.04.019
- Wu, X., L. Jiang, M. Yu, X. An, R. Ma y Z. Yu. 2016. Proteomic analysis of changes in mitochondrial protein expression during peach fruit ripening and senescence. *Journal of Proteomics* 147: 197-211. doi: 10.1016/j.jprot.2016.06.005
- Wu, A. R., N. F. Neff, T. Kalisky, P. Dalerba, B. Treutlein, M. E. Rothenberg, F.M. Mburu, G.L. Mantalas, S. Sim, M.F. Clarke et al. 2014. Quantitative assessment of single-cell RNA-sequencing methods. *Nature Methods* 11(1): 41. doi: 10.1038/nmeth.2694
- Wurch, L. L., S. T. Haley, E. D. Orchard, C. J. Gobler y S. T. Dyhrman. 2011. Nutrient- regulated transcriptional responses in the brown tide- forming alga *Aureococcus anophagefferens*. *Environmental Microbiology* 13(2): 468-481. doi: 10.1111/j.1462-2920.2010.02351.x
- Xiong, J. 2006. Protein Motifs and Domain Prediction. 85-94 p. En: *Essential Bioinformatics*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511806087.008
- Xu, X. y X. Miao. 2020. Glyceroglycolipid Metabolism Regulations under Phosphate Starvation Revealed by Transcriptome Analysis in *Synechococcus elongatus* PCC 7942. *Marine Drugs* 18(7): 360. doi: 10.3390/md18070360
- Yang, F., W. Xiang, T. Li y L. Long. 2018. Transcriptome analysis for phosphorus starvation-induced lipid accumulation in *Scenedesmus* sp. *Scientific Reports* 8(1): 1-11. doi: 10.1038/s41598-018-34650-x
- Yang, V. W. 2018. The cell cycle. *Physiology of the Gastrointestinal Tract*. 197-219.
- Yang, S. Y., T. K. Huang, H. F. Kuo y T. J. Chiou. 2017. Role of vacuoles in phosphorus storage and remobilization. *Journal of Experimental Botany* 68(12): 3045-3055. doi: 10.1093/jxb/erw481
- Yang, X., K. Teng, J. Zhang, F. Wang, T. Zhang, G. Ai, P. Han, F. Bai y Zhong, J. 2017. Transcriptome responses of *Lactobacillus acetotolerans* F28 to a short and long term ethanol stress. *Scientific Reports* 7(1): 1-11. doi: 10.1038/s41598-017-02975-8
- Yang, J., Y. Pan, C. Bowler, L. Zhang y H. Hu. 2016. Knockdown of phosphoenolpyruvate carboxykinase increases carbon flux to lipid synthesis in *Phaeodactylum tricorutum*. *Algal Research* 15: 50-58. doi: 10.1016/j.algal.2016.02.004

- Yang, X., J. Nian, Q. Xie, J. Feng, F. Zhang, H. Jing, J. Zhang, G. Dong, Y. Liang, J. Peng, et al. 2016. Rice ferredoxin-dependent glutamate synthase regulates nitrogen–carbon metabolomes and is genetically differentiated between japonica and indica subspecies. *Molecular Plant* 9(11): 1520-1534.
- Yang, Z. K., J. W. Zheng, Y. F. Niu, W. D. Yang, J. S. Liu, H. Y. Li. 2014. Systems- level analysis of the metabolic responses of the diatom *Phaeodactylum tricornutum* to phosphorus stress. *Environmental Microbiology* 16: 1793-1807. doi: 10.1111/1462-2920.12411
- Yavuz, C. y Z. N. Öztürk. 2017. Working with Proteins in silico: A Review of Online Available Tools for Basic Identification of Proteins. *Turkish Journal of Agriculture-Food Science and Technology* 5(1): 65-70. doi: 10.24925/turjaf.v5i1.65-70.926
- Yoshida, T., T. Obata, R. Feil, J. E. Lunn, Y. Fujita, K. Yamaguchi-Shinozaki y A. R. Fernie. 2019. The role of abscisic acid signaling in maintaining the metabolic balance required for Arabidopsis growth under nonstress conditions. *The Plant Cell* 31(1): 84-105. doi: 10.1105/tpc.18.00766
- Yoshida, T., J. Mogami y K. Yamaguchi-Shinozaki. 2014. ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. *Current Opinion in Plant Biology* 21: 133-139. doi: 10.1016/j.pbi.2014.07.009
- Yoshimura, K. y S. Shigeoka. 2015. Versatile physiological functions of the Nudix hydrolase family in Arabidopsis. *Bioscience, Biotechnology, and Biochemistry* 79(3): 354-366. doi: 10.1080/09168451.2014.987207
- Young, M. D., M. J. Wakefield, G. K. Smyth y A. Oshlack. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* 11(2): R14. doi: 10.1186/gb-2010-11-2-r14
- Zhang, S. F., C. J. Yuan, Y. Chen, X. H. Chen, D. X. Li, J. L. Liu, L. Lin y D. Z. Wang. 2016. Comparative transcriptomic analysis reveals novel insights into the adaptive response of *Skeletonema costatum* to changing ambient phosphorus. *Frontiers in Microbiology* 7: 1476. doi: 10.3389/fmicb.2016.01476
- Zhao, S., B. Zhang, Y. Zhang, W. Gordon, S. Du, T. Paradis, M. Vincent y D. von Schack. 2016. Bioinformatics for RNA-seq data analysis. *Bioinformatics—Updated Features and Applications: InTech*, 125-49. doi: 10.5772/63267
- Zhao, Y.F., Z. Yu, X.X. Song y X.H. Cao. 2009. Biochemical compositions of two dominant bloom-forming species isolated from the Yangtze River Estuary in response to different nutrient conditions. *Journal of Experimental Marine Biology and Ecology* 368: 30-36. doi: org/10.1016/j.jembe.2008.09.023.
- Zhu, L.D., Z.H. Li y E. Hiltunen. 2016. Strategies for lipid production improvement in microalgae as a biodiesel feedstock. *BioMed Research International* 2016: 1 – 8. doi: 10.1155/2016/8792548
- Zorin, B., P. Hegemann y I. Sizova. 2005. Nuclear-gene targeting by using single-stranded DNA avoids illegitimate DNA integration in *Chlamydomonas reinhardtii*. *Eukaryot Cell* 4: 1264–1272. doi: 10.1128/EC.4.7.1264-1272.2005
- Zulu, N.N., K. Zienkiewicz, K. Vollheyde y I. Feussner. 2018. Current trends to comprehend lipid metabolism in diatoms. *Progress in Lipid Research* 70: 1-16. doi.org/10.1016/j.plipres.2018.03.001.

## X. APENDICES

**Nota:** Debido a la contingencia por COVID-19, el proceso de validación del transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P se detuvo; sin embargo, a continuación, se muestran los avances que se han logrado obtener hasta el momento.

### 10.1 Estandarización de oligos de genes expresados diferencialmente en el transcriptoma de *C. muelleri* seleccionados para la validación por la técnica de RT-qPCR.

#### 10.1.1 Extracción de ADN genómico de *C. muelleri*

Las células algales de cada tratamiento fueron colectadas mediante centrifugación a 2500 g x 10 min. El sobrenadante fue eliminado y los pellets resultantes fueron inmediatamente almacenados en tubos Falcon de 50 mL a – 80 °C previo a su procesamiento. Para la extracción de ADN genómico de *C. muelleri*, la muestra permaneció en nitrógeno líquido hasta que se vertió dentro de un tubo Eppendorf de 1.5 mL con 600 µl de buffer de lisis (Tris-HCl 100mM a pH 8.0, NaCl 20mM, EDTA 20mM, y N- Lauril sarcosina al 1%) y se dejó reposar por 5 min. Posteriormente, se siguieron los siguientes pasos para continuar con la extracción de ADN genómico:

- ✓ Se añadieron 2 µl de RNAsa y se dejó reposar por 15 min.
- ✓ Se agregaron 600µl de fenol:cloroformo:alcohol-isoamílico (25:24:1) y se centrifugó a 12 000 rpm por 20 min a 4°C.
- ✓ Se tomó el sobrenadante y se llevó a un tubo nuevo y se adicionó 600 µL de isopropanol al 100 %, dejándose incubar por 10 min a temperatura ambiente. Posteriormente, se centrifugó a 12 000 rpm por 10 min y se decantó el sobrenadante dejando el pellet precipitado.
- ✓ Se lavó el pellet con 1 mL de etanol al 70 % y se centrifugó a 7 500 rpm por 5 min. Luego, se eliminó el etanol dejando secar el pellet (ARN precipitado) durante 10 min y se resuspendió en agua libre de ARNasas estéril (50 µL - 100 µL).

Una vez realizado el procedimiento anterior, se verificó la integridad del ADN<sub>g</sub> resultante en electroforesis en un gel de agarosa al 1 % a 110 V por 35 min aproximadamente (Figura S1).

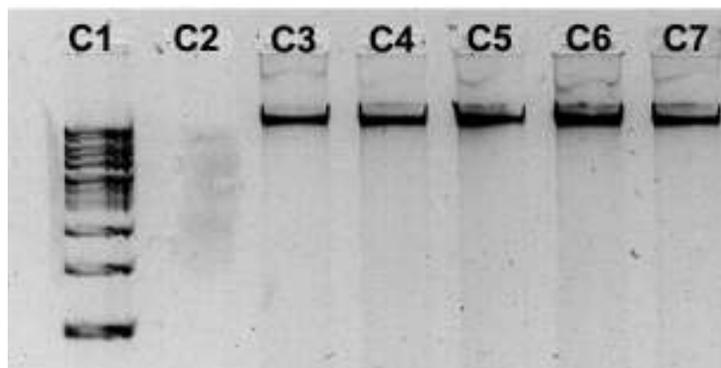


Figura S1. ADN genómico de *C. muelleri* en gel de agarosa al 1% teñido con GelRed™ (Biotium). A) ARN extraído con el reactivo TRIzol® B) ARN tratado con el kit RNeasy MinElute Cleanup Kit. Carril 1: Marcador de peso molecular (1 Kb Plus DNA Ladder, Invitrogen). Carril 3 – 7: Muestras de ADN genómico de *C. muelleri*.

### 10.1.2 Selección de genes expresados diferencialmente en el transcriptoma de *C. muelleri* cultivada bajo diferentes concentraciones de P.

Una vez que se llevó a cabo el análisis de RNA-seq y para su posterior validación, se seleccionaron 8 genes expresados diferencialmente en el transcriptoma de *C. muelleri* y dos genes constitutivos cuya estabilidad ya ha sido evaluada en diatomeas. Los oligos fueron diseñados usando el software Primer3 y luego analizados con la herramienta OligoEvaluator de Sigma-Aldrich. En la Tabla S1 se muestran los 10 pares de oligos que se estandarizaron mediante la técnica de PCR punto final, los cuales fueron sintetizados por la compañía T4 OLIGO ubicada en la ciudad de Irapuato, Guanajuato, México.

<b>Tabla S1.</b> Lista de oligos estandarizados en este estudio.		
<b>Nombre</b>	<b>Abreviación</b>	<b>Inicial (F)/ Reverso (R)</b>
Alkaline phosphatase	APA	F: GCAGGAGCGTAAACAGTGAT R: GATTGGATACGACCCCTCT
Fucoxanthin-chlorophyll a-c binding protein A	FCPA	F: CTCGGACCCTCTGGATCTAA R: AACCATCCTGCGACAAAGAA
Sodium-dependent phosphate transport protein 2B	NPT2B	F: CAAATGCTTCCTCTTGTGCT R: TGAAAGGGACAGGATACCAA
Vacuolar transporter chaperone 4	VTC4	F: TATCCACCTTGTCATTCCG R: GGCCATATATTCTCGGAGG
Putative inorganic phosphate transporter	PIPT	F: ACCAATCAACACCCAACAGT R: CTCCTTGGTATCGGTTGTGG
SPX domain-containing membrane protein	SPXM2	F: GCTCTGTTGCAATTCTCTCC R: GTCGTCTTTCGGGAACAGTA
Phospholipase A1-Igama3	PLA	F: ACTGCCGCAACTGTATCATT R: CGTACCCAAACCTGTGACTT
Ammonium transporter 1	AMT1	F: ATTTGCTGCTCCGAGACATAC R: TACACGGAACATACCAAGC
TATA-box-binding protein	TBP	F: TAAGAGAACCACGAGCAACG R: AACCCAACATCAATCGTCCC
Hypoxanthine-guanine phosphoribosyltransferase	HPRT	F: CCTTCCAATTATCGCGGTCT R: CACCGTGGGGATAGTTTGAA

### 10.1.3 Reacción de PCR

Se utilizó el GoTaq Colorless Master Mix (PROMEGA), siguiendo las instrucciones del fabricante, se añadieron los siguientes componentes a un microtubo estéril para un volumen final de reacción de 12.5  $\mu$ L:

- ✓ 6.3  $\mu$ L de GoTaq Colorless Master Mix
- ✓ 1  $\mu$ L de primer reverso y 1  $\mu$ L de primer forward (Dependiendo del gen).
- ✓ 3.2  $\mu$ L de H<sub>2</sub>O estéril.
- ✓ 1  $\mu$ L de ADN genómico de *C. muelleri*.

#### **10.1.4 Determinación de los parámetros de temperaturas y ciclos.**

La temperatura de alineamiento y los ciclos se definieron llevando a cabo un PCR de gradiente tanto para definir la temperatura de alineamiento como el número de ciclos necesarios para la visualización de los productos de PCR de la siguiente manera en un termociclador SimpliAmp™ de Applied Biosystems:

- Un ciclo de desnaturalización por 5 min a 94 °C.
- Para definir la temperatura de alineamiento se tomó un rango entre 50 °C y 65 °C. Las condiciones para cada ciclo de PCR fueron las siguientes: Temperatura de desnaturalización de 94 °C por 1 min, la temperatura de alineamiento por 30 seg y la temperatura de extensión fue 72 °C por 40 seg.
- Para determinar la cantidad de ciclos apropiada para cada gen se tomó un rango de entre 25 y 35 ciclos, hasta que la amplificación de los productos de PCR para cada gen, fuera visible en el gel de agarosa.

#### **10.1.5 Análisis de los productos de PCR**

Se realizó un gel de agarosa al 1 % para observar la integridad de los productos de PCR de la siguiente forma:

- ✓ Se pesaron 0.75 g de agarosa y se disolvieron en un matraz Erlen-Meyer con 75 mL de buffer TAE 1X estéril.
- ✓ Se calentó la mezcla de agarosa y TAE 1X en un horno microondas por un minuto y medio, verificando su disolución cada 30 seg. Hasta que la mezcla no presentó grumos y se tornó transparente, se dejó reposar hasta que adquiriera una temperatura soportable al dorso de la mano.
- ✓ Se agregó 1 µL de GelRed™ (Biotium) y se agitó vigorosamente para después vaciarse lentamente en la placa de electroforesis previamente preparada con los sellos acrílicos, esto para

evitar la formación de burbujas. Se colocaron los peines de separación correspondiente y se dejó reposar hasta que la mezcla se solidificara (15 a 20 minutos).

✓ Una vez formado el gel, se retiraron los sellos y se colocó la placa dentro del tanque de electroforesis. Se vertió la cantidad apropiada de buffer TAE 1X estéril hasta que cubriera el gel y se retiraron los peines con cuidado, evitando que se dañara el gel.

✓ Se cargaron los productos de PCR con volumen de 2  $\mu$ L de muestra y un 1  $\mu$ L de Marcador de peso molecular (1 Kb Plus DNA Ladder, Invitrogen) para determinar el tamaño de cada uno de los fragmentos, con 2  $\mu$ L de buffer de carga 1X y se depositaron en los pozos correspondientes.

✓ Se llevó a cabo el proceso de electroforesis conectando la cámara de electroforesis a una fuente de poder a 100 Volts por 35 min aproximadamente.

✓ Una vez terminado el proceso, se visualizó el gel de electroforesis en el fotodocumentador Bio-Imaging System MiniBis pro (Bio America Inc).

✓ Por último, se enjuagó con agua estéril todos los componentes de electroforesis al terminar la corrida, ya que con el paso del tiempo la acumulación de restos de sales puede dañar el equipo.

#### **10.1.6 Estandarización de condiciones para PCR**

Se definió para cada gen seleccionado del transcriptoma de *C. muelleri* para su posterior validación mediante la técnica de RT-qPCR, las condiciones adecuadas de PCR mediante gradientes de ciclos y de temperaturas. En la Tabla S2 se presentan las condiciones a las cuales fueron analizados cada uno de los genes. Para la estandarización de dichas condiciones se empleó ADN genómico de *C. muelleri*.

<b>Tabla S2. Condiciones definidas para PCR</b>		
<b>Gen</b>	<b>Número de ciclos</b>	<b>Temperatura de Alineamiento (°C)</b>
APA	27	53
FCPA	27	51
NPT2B	27	53
VTC4	27	57
PIPT	27	51
SPXM2	27	59
PLA	27	51
AMT1	27	57
TBP	27	59
HPRT	27	59