



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**POSGRADO EN CIENCIAS BIOLÓGICAS**  
INSTITUTO DE ECOLOGÍA  
BIOLOGÍA EXPERIMENTAL

**RELEVANCIA DE LAS CAPACIDADES METABÓLICAS MICROBIANAS EN LA FORMACIÓN  
DE CONSORCIOS Y COMUNIDADES. SU ANÁLISIS MEDIANTE HERRAMIENTAS  
METAGENÓMICAS Y BIOLOGÍA COMPUTACIONAL**

**TESIS**

QUE PARA OPTAR POR EL GRADO DE:

**DOCTOR EN CIENCIAS**

PRESENTA:

**M. en C. DANIEL CERQUEDA GARCÍA**

TUTORA PRINCIPAL DE TESIS:

**DRA. LUISA ISAURA FALCÓN ÁLVAREZ**  
INSTITUTO DE ECOLOGIA, UNAM

COMITÉ TUTOR:

**DR. LEÓN PATRICIO MARTÍNEZ CASTILLA**  
FACULTAD DE QUIMICA, UNAM

**DR. LUIS JOSÉ DELAYE ARREDONDO**  
CINVESTAV IRAPUATO, IPN

**MÉXICO, CDMX.**

**Octubre, 2017.**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.





**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**POSGRADO EN CIENCIAS BIOLÓGICAS**  
INSTITUTO DE ECOLOGÍA  
BIOLOGÍA EXPERIMENTAL

**RELEVANCIA DE LAS CAPACIDADES METABÓLICAS MICROBIANAS EN LA FORMACIÓN  
DE CONSORCIOS Y COMUNIDADES. SU ANÁLISIS MEDIANTE HERRAMIENTAS  
METAGENÓMICAS Y BIOLOGÍA COMPUTACIONAL**

**TESIS**

QUE PARA OPTAR POR EL GRADO DE:

**DOCTOR EN CIENCIAS**

PRESENTA:

**M. en C. DANIEL CERQUEDA GARCÍA**

TUTORA PRINCIPAL DE TESIS:

**DRA. LUISA ISAURA FALCÓN ÁLVAREZ**  
INSTITUTO DE ECOLOGIA, UNAM

COMITÉ TUTOR:

**DR. LEÓN PATRICIO MARTÍNEZ CASTILLA**  
FACULTAD DE QUIMICA, UNAM

**DR. LUIS JOSÉ DELAYE ARREDONDO**  
CINVESTAV IRAPUATO, IPN

**MÉXICO, CDMX.**

**Octubre, 2017.**






Lic. Ivonne Ramírez Wence  
Directora General de Administración Escolar, UNAM  
Presente

Me permito informar a usted, que el Subcomité de Biología Evolutiva y Sistemática, en su sesión ordinaria del día 07 de agosto de 2017, aprobó el jurado para la presentación de su examen para obtener el grado de **DOCTOR EN CIENCIAS**, del Posgrado en Ciencias Biológicas, del alumno **CERQUEDA GARCÍA DANIEL** con número de cuenta **98251003**, con la tesis titulada: **"RELEVANCIA DE LAS CAPACIDADES METABÓLICAS MICROBIANAS EN LA FORMACIÓN DE CONSORCIOS Y COMUNIDADES. SU ANÁLISIS MEDIANTE HERRAMIENTAS METAGENÓMICAS Y BIOLOGÍA COMPUTACIONAL"**, bajo la dirección de la **DRA. LUISA ISAURA FALCÓN ÁLVAREZ**:

Presidente:	DR. ALFONSO VALIENTE BANUET
Vocal:	DR. ARTURO CARLOS II BECERRA BRACHO
Secretario:	DR. LEÓN PATRICIO MARTÍNEZ CASTILLA
Suplente:	DR. LUIS EDUARDO SERVIN GARCIDUEÑAS
Suplente	DR. LUIS DAVID ALCARAZ PERAZA

Sin otro particular, me es grato enviarle un cordial saludo.

**ATENTAMENTE**  
**"POR MI RAZA HABLARA EL ESPIRITU"**  
Cd. Universitaria, Cd. Mx., a 02 de octubre de 2017

  
**DR. ADOLFO GERARDO NAVARRO SIGÜENZA**  
**COORDINADOR DEL PROGRAMA**



## **Agradecimientos**

Al Posgrado en Ciencias Biológicas, UNAM por permitirme completar mi formación académica.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por haberme otorgado la beca para mis estudios de doctorado (No. 290817)

Al Programa de Apoyo a los Estudios de Posgrado (PAEP), UNAM por los apoyos recibidos para estancias y congresos.

A los miembros de mi comité tutor: Dra. Luisa Isaura Falcón Álvarez, Dr. Luis José Delaye Arredondo y Dr. León Patricio Martínez Castilla, por sus aportaciones a este proyecto.

## **Agradecimientos personales**

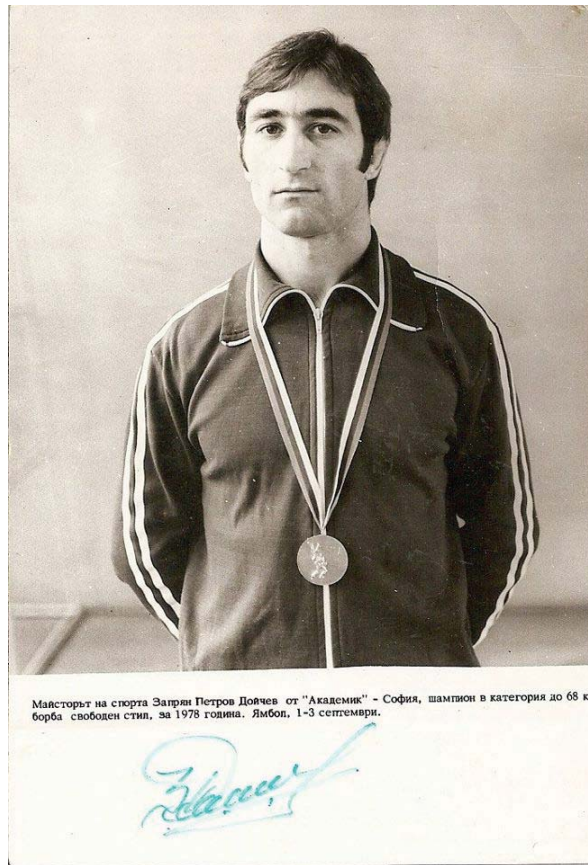
A mi madre y mis hermanos.

A mis amigos, esos que conocí en Ciencias y que ya son parte de mi vida.

A los miembros del laboratorio de Ecología Microbiana

A los miembros del jurado de esta tesis: Dr. Alfonso Valiente Banuet, Dr. Arturo Carlos II Becerra Bracho, Dr. Luis David Alcaraz Peraza y Dr. Luis Eduardo Servín Garcidueñas.

*IN MEMORIAM*



*Prof. Zaprian Petrov Doychev  
1952-2017*

*Quien fue el primer profesor que conocí al ingresar a la UNAM y murió al término de mis estudios de doctorado. Entrenador, guía, amigo y padre adoptivo de muchos. Gracias por mantenernos en el camino correcto y ser un ejemplo de vida, enseñándonos, a ser disciplinados, trabajar duro, amar lo que hacemos y nunca dejar de luchar.*



<b>Índice</b>	
<b>Resumen</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>Introducción</b>	
<i>Interacciones metabólicas microbianas en comunidades</i>	<b>3</b>
<i>Chlorochromatium aggregatum, modelo de consorcio entre dos especies</i>	<b>6</b>
<i>Tapetes microbianos y microbialitos como modelos de comunidades</i>	<b>9</b>
<i>Estudio de la importancia del metabolismo bacteriano en las comunidades</i>	<b>13</b>
<i>Anotación y reconstrucción de modelos metabólicos a partir de metagenomas</i>	<b>13</b>
<b>Objetivos</b>	<b>25</b>
<b>Capítulo 1</b>	
<i>La construcción del nicho y el concepto de holobionte, hacia la reestructuración de un paradigma</i>	<b>26</b>
<b>Capítulo 2</b>	
<i>Metabolic analysis of Chlorobium chlorochromatii CaD3 reveals clues of the symbiosis in ‘Chlorochromatium aggregatum’</i>	<b>30</b>
<b>Capítulo 3</b>	
<i>Metabolic potential of microbial mats and microbialites: Autotrophic capabilities described by an in silico stoichiometric approach from shared genomic resources</i>	<b>39</b>
<b>Discusión general</b>	<b>55</b>
<b>Conclusiones</b>	<b>57</b>
<b>Referencias</b>	<b>58</b>

## Índice de Figuras y Cuadros

**Cuadro 1.** Interacciones entre especies que determinan la estructura y composición de las comunidades.

**Figura 1.** Estructura de *Chlorochromatium aggregatum*

**Cuadro 2.** Actividades metabólicas de los principales grupos de bacterias que integran los tapetes microbianos.

**Figura 2.** Microbialitos de Pozas Azules, Cuatro Ciénegas, Coahuila, México.

**Figura 3.** Diagrama de flujo del proceso para la anotación funcional y reconstrucción de la red metabólica y el análisis de las capacidades metabólicas de una comunidad por medio de modelos estequiométricos.

**Figura 4.** Diagrama de una red metabólica hipotética con su matriz estequiométrica asociada.

**Figura 5.** Diagrama que muestra los modos elementales de flujo (EFM) de la red hipotética de la Figura 4.

**Cuadro 3.** Significado de términos utilizados.

**Figura 6.** Principales campos de aplicación para los MCSs (modificada de Klamt, 2006)

**Cuadro 4.** Metagenomas usados obtenidos de la base de datos del MG-RAST

## Resumen

En la naturaleza, las especies interactúan entre sí y con su medio abiótico, formando consorcios y comunidades, como los tapetes microbianos y microbialitos. De entre los tipos de interacciones entre los microorganismos, las interacciones metabólicas son de gran importancia en la evolución. Por ejemplo mediante su actividad metabólica, los microorganismos son capaces de conducir los ciclos biogeoquímicos de la Tierra debido a su diversidad metabólica. Con su metabolismo, los microorganismos pueden modificar el ambiente a escalas micro y planetarias, cambiando de esta manera las presiones de selección ambientales en otros organismos, como en el cambio *redox* de la atmósfera documentado en el registro fósil. Por ello es importante el estudio de las interacciones metabólicas a escala comunitaria.

Con el desarrollo de la secuenciación de alto rendimiento, surgió la metagenómica como un enfoque en el estudio de las comunidades, lo que ha generado una gran cantidad de información sin la necesidad del cultivo de especies. Esta información puede ser integrada con enfoques de la Biología de Sistemas, como el análisis de modos elementales de flujo (*elementary flux modes*-EFM), que es un análisis de la estructura de las redes metabólicas, con el cual se puede predecir tasas de consumo y producción de metabolitos. En este estudio utilizamos el enfoque de los EFMs, integrados con datos genómicos y metagenómicos. Primero utilizamos al consorcio *Chlorochromatium aggregatum*, como modelo de una asociación célula-célula entre dos especies, conformado por una betaproteobacteria y la bacteria verde del azufre *Chlorobium chlorochromatii*. Después utilizamos un conjunto de diez metagenomas provenientes de microbialitos y tapetes microbianos, para modelar su capacidad metabólica autotrófica y caracterizar el intercambio metabólico y sus capacidades metabólicas a nivel comunidad. En el primer caso predecimos que hay un intercambio de carbono-nitrógeno en la interacción célula-célula asociado al modo de vida simbiote. En el segundo caso encontramos redundancia en las rutas de incorporación de carbono inorgánico en las comunidades estudiadas.

## Abstract

In nature, the species interact with each other and with its abiotic environment, forming consortia and communities as microbial mats and microbialites. Of the types of interactions in the microorganisms, the metabolic interactions are the most relevant in the coevolution. For example, by its metabolic activity, the microorganism can drive the biogeochemical cycle from Earth, owing to its wide metabolic diversity. Therefore, with its metabolism, the microorganisms can modify the environment to a micro and planetary scale, changing the environmental selective pressures in other organisms, the redox shift of the atmosphere in the fossil record is an example of it. Thus the study of the metabolic interactions to a community scale is very relevant.

The development of the high-throughput sequencing made possible the metagenomic approaches in the study of communities. It has enabled the access to a plethora of information without the requirement of species cultivation. This information can be integrated by System Biology approaches, as the elementary flux modes analysis (EFM), which is an analysis of the structure of metabolic networks. The EFM analysis can predict the rates of uptakes and outputs of metabolites. In this study, we used the EFM approach, integrated with genomic and metagenomic data. First, we used the consortium *Chlorochromatium aggregatum* as model of a cell-cell interaction among two species; a betaproteobacteria and the green-sulfur bacterium *Chlorobium chlorochromatii*. Then we used a dataset of ten metagenomes from microbialites and microbial mats, with the aim of modeling its autotrophic metabolic capabilities and to describe the metabolic exchange and its metabolic capabilities at community level. In the first case study, we found an exchange of carbon-nitrogen in the cell-cell interaction associated to the symbiotic life style. In the second case study, we found redundancy in the pathways for the uptake of inorganic carbon in the communities studied.

## Introducción

### *Interacciones metabólicas microbianas en comunidades*

En el campo de estudio de la ecología, el concepto de comunidad se refiere a los ensambles de múltiples especies, en donde los organismos viven juntos en un ambiente continuo e interactúan entre sí. Este concepto surgió para el estudio de los organismos en los ecosistemas, donde se trata de responder cómo están estructurados los ensambles biológicos, cuáles son sus interacciones funcionales y cómo cambia la estructura de la comunidad en el tiempo y el espacio (Konopka, 2009). En las comunidades el componente biológico dominante es microbiano el cual es responsable de los procesos biogeoquímicos de los seis principales componentes de las macromoléculas (C, H, O, N, S y P) (Falkowski *et al.*, 2008). Esto pone en un papel principal a los microorganismos en los procesos de mantenimiento de las comunidades y los ecosistemas de la biosfera. Entre las interacciones en estas comunidades, resumidas en el Cuadro 1, las más importantes en los procesos biogeoquímicos, son las interacciones metabólicas.

Desde el origen de la vida, la actividad metabólica de los microorganismos ha modificado las condiciones del planeta. Un ejemplo de ello es el cambio de las condiciones *redox* de la atmósfera a partir de la generación de oxígeno libre, producto de la fotosíntesis oxigénica. El registro fósil indica un Gran Evento Oxidativo alrededor de 2,450 MA en donde mediante el rompimiento de la molécula de agua se cambió la composición de la atmósfera volviéndola oxidante. Además, desde este acontecimiento la actividad fotosintética ha proporcionado una fuente de oxígeno, ya que al ser una especie altamente reactiva, sin su recambio constante pasaría a ser parte de la corteza terrestre en alguna forma de óxido, alcanzando su equilibrio químico (Falkowski & Godfrey, 2008). El potencial metabólico microbiano es considerado como reservorio de los procesos metabólicos capaces de conducir los ciclos biogeoquímicos, ya que los microorganismos pueden sobrevivir a sucesos catastróficos como impactos de meteoritos o glaciaciones, contrariamente a las plantas y animales en los que hay mayor probabilidad de que se produzcan extinciones de especies.

Existe una redundancia metabólica en diferentes linajes microbianos, lo cual es uno de los mecanismos que ha asegurado el funcionamiento biogeoquímico de la Tierra a lo largo de la historia de la vida (Falkowski *et al.*, 2008). La comprensión y caracterización de los procesos metabólicos de los flujos de materia y energía entre los microorganismos, nos brinda conocimiento sobre la dinámica y los

principios que gobiernan y modifican los procesos biogeoquímicos. Por ello, las comunidades con componentes microbianos dominantes como los tapetes microbianos, que forman estructuras laminadas y microbialitos, son de un interés especial para la Ecología (Des Marais, 2003).



**Cuadro 1.** Interacciones entre especies que determinan la estructura y composición de las comunidades. Modificada de Konopka, 2009.

Categoría	Clase	Ejemplo	Referencia
Competencia por recursos.	De superficie celular: transportadores de nutrientes Enzimas extracelulares Sideróporos/lanzaderas de electrones		(Sowell <i>et al.</i> , 2009)  (Allison, 2005) (Miethke&Marahiel, 2007)
Interacciones metabólicas.	Sintrofia <i>Cross-feeding</i> Utilización secuencial de nutrientes Modificación química del ambiente	Acoplamiento de los ciclos biogeoquímicos Nitrificadores, catabolismo de xenobióticos Sucesión debida a la alteración del pH por ácidos orgánicos ' <i>Chlorochromatium aggregatum</i> '	(Dolfing, 2001) (Overmann & Gemerden 2000) (Villareal <i>et al.</i> , 1991)  (Wanner <i>et al.</i> , 2008)
Alelopatía	Metabolitos microbianos	Bacteriocinas Inhibición de la formación del quórum sensing	(Gillor <i>et al.</i> , 2008; Riley & Wertz 2002) (Gonzales & Keshavan, 2006)
Señalización	Microbiana con organismos pluricelulares	Formación de quórum interespecífico Nódulos radiculares	(Ryan & Dow, 2006) (Cooper, 2007)
Interacciones estructurales	Consortios Biocapas y tapetes microbianos	' <i>Chlorochromatium aggregatum</i> ' Microbialitos	(Wanner <i>et al.</i> , 2008) (Des Marais, 2003)
Interacciones nivel trófico	a Circuito microbiano		(Azam & Malfatti, 2007) (Suttle, 2007)
Transferencia horizontal de genes	de Parasitismo	Virus	(Juhás <i>et al.</i> , 2009)
Coevolución			(Moran <i>et al.</i> , 2008)

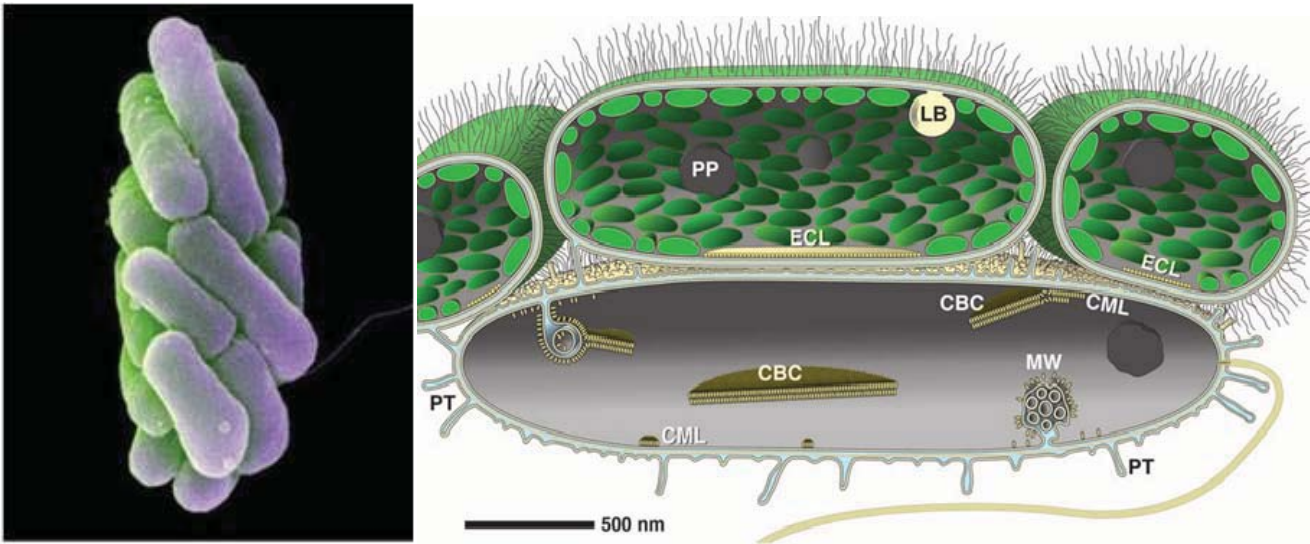
## *Chlorochromatium aggregatum, modelo de consorcio entre dos especies*

El consorcio denominado *Chlorochromatium aggregatum*, está conformado por dos especies bacterianas; una  $\beta$ -proteobacteria central recientemente denominada “*Candidatus Symbiobacter* (Ca. S.) *mobilis*”, rodeada por células espibiontes de *Chlorobium chlorochromatii*, bacteria verde del azufre del phylum chlorobia (Fröstl *et al.*, 1998; Liu *et al.*, 2013). *C. chlorochromatii* es una bacteria fotosintética, no motil, anaeróbica estricta, Gram <sup>(-)</sup>, fijadora de nitrógeno, en condiciones naturales solo se ha encontrado en simbiosis, pero puede vivir en vida libre en condiciones de laboratorio (Vogl *et al.*, 2006). ‘Ca. S. mobilis’ es incolora monoflagelada. No se conoce mucho de su fisiología, ya que es un simbionte obligado, pero mediante análisis de su gen 16S se sabe que está emparentada con la familia commamonadacea, cercana a *Variovorax paradoxus* (Glaeser *et al.*, 2003; Kanzler 2005). El consorcio es motil debido al flagelo de la  $\beta$ -proteobacteria y presenta movimiento escotofóbico, es decir, se mueve en respuesta a cambios de la intensidad de la luz y se aglomera a una longitud de onda de 749 nm (longitud de onda donde la bacterioclorofila *c* tiene la máxima absorción) (Fröstl *et al.*, 1998). Entre las dos especies existen mecanismos de comunicación química, ya que la división celular es coordinada y se forman dos consorcios completos. El periplasma de las dos especies esta interconectado por túbulos periplásmicos (Figura 1) (Wanner *et al.*, 2008).

Wenter y col. (2010) mostraron que en los estados de simbiosis y vida libre las células de *C. chlorochromatii* expresan 352 genes de manera diferencial entre las dos condiciones. Entre estos, los que codifican para las enzimas del metabolismo de los aminoácidos son los que muestran diferencias más significativas (Wenter *et al.*, 2010; Overmann, 2010). Así mismo detectaron genes que se expresan únicamente en simbiosis: un gen que codifica para la proteína PII reguladora del metabolismo del nitrógeno, un gen que codifica una glutamina sintetasa (GS) y un gen que codifica para un transportador tipo ABC de aminoácidos ramificados.

Las enzimas que participan en el metabolismo de los aminoácidos están reguladas por la disponibilidad de nitrógeno de las bacterias (Brown & Herbert, 1977; Magasanik, 1982; Rudnick *et al.*, 1997). Wenter y colaboradores (Wenter *et al.*, 2010) postulan que *C. chlorochromatii* en simbiosis tiene acceso limitado al nitrógeno. En simbiosis se realiza la síntesis de PII, GS y la transcripción de la nitrogenasa es elevada, lo que parece indicar que la célula necesita fijar N<sub>2</sub>, esta gran necesidad de fijar nitrógeno en simbiosis puede deberse a una sintrofia por el nitrógeno entre los simbiontes. Por todo lo anterior, es

importante analizar la red metabólica de *C. chlorochromatii* para saber qué rutas metabólicas son modificadas en la simbiosis y si existe sintrofia por algún recurso.



**Figura 1.** Estructura de *Chlorochromatium aggregatum*. A la izquierda, una imagen de *C. aggregatum* hecha por microscopía electrónica de alta resolución (tomada de la portada del *Journal of Bacteriology* vol. 190 num. 10 año 2008) donde se observan a los epibiontes y el flagelo de la bacteria central. A la derecha, un esquema tomado de Wanner y col. (2008) donde se resumen las observaciones por microscopía electrónica de la ultra estructura del consorcio: LB, cuerpos lipídicos; PP, glóbulos de polifosfato; ECL, capa de contacto del epibionte (caracterizada por la ausencia de clorosomas); PT, túbulos periplásmicos (haciendo contacto entre las dos bacterias formando un espacio periplásmico común); MW, invaginación de la membrana citoplásmica; CML, capa membranas de la bacteria central; CBC, cristales de la bacteria central (estructuras parecidas a cristales formadas en la cara interna de la membrana citoplásmica por acumulación de estructuras globulares).

### *Tapetes microbianos y microbialitos como modelos de comunidades*

Estas comunidades son ecosistemas a escala milimétrica, ya que son autosuficientes debido a la gran diversidad metabólica que contienen en su estructura multilaminada (Des Marais, 2003; Dupraz *et al.*, 2004; Dupraz *et al.*, 2009). Las especies microbianas que las componen son clasificadas en gremios y ensamblajes dependiendo de sus características metabólicas y su afiliación filogenética. Estas se desarrollan en una gran diversidad de ambientes con gradientes fisicoquímicos (ej. de luz, azufre y oxígeno), en donde se pueden presentar desde una condición super saturada a una escasa (por ejemplo ir de un ambiente anóxico a oxigénico), resultando en distintos micro-ambientes (Vissher & Stolz, 2005). Dentro de su estructura, estas comunidades contienen grupos tróficos esenciales (productores primarios, secundarios y descomponedores), por ende, estos micro ecosistemas son dinámicos y sostienen una gran diversidad de especies con extensas habilidades metabólicas en un pequeño espacio (Vischer & Stolz, 2005).

En general en una comunidad de tapetes microbianos y/o de microbialitos, se encuentran los siguientes cuatro grupos funcionales:

1. Fotótrofos oxigénicos, principalmente cianobacterias, son los productores primarios y acoplan energía lumínica a la fijación de CO<sub>2</sub> y N<sub>2</sub>.
2. Bacterias aerobias heterótrofas, oxidan la mayor parte del carbono fijado, obteniendo energía por medio de la respiración utilizando oxígeno como aceptor final de electrones.
3. Bacterias reductoras de sulfato (SRB), son los organismos anaerobios dominantes y producen sulfuro en la respiración.
4. Bacterias oxidadoras de sulfuro (SOB), son organismos quimiolitótrofos que oxidan sulfuros utilizando oxígeno o nitrato como aceptores de electrones.

El Cuadro 2 puntualiza las diferencias de las actividades metabólicas de los diferentes grupos funcionales dependiendo del ciclo diario.

**Cuadro 2.** Actividades metabólicas de los principales grupos de bacterias que integran tapetes microbianos. Modificado de Visscher & Stolz, 2004.

Grupo funcional	Función metabólica diurna	Función metabólica nocturna
Cianobacterias	Fijación de CO <sub>2</sub> (fotosíntesis): CO <sub>2</sub> + H <sub>2</sub> O → CH <sub>2</sub> O + O <sub>2</sub>	Fermentación (incluyendo producción de H <sub>2</sub> ), Fijación de N <sub>2</sub> , Degradación de glucógeno
Heterótrofos aerobios	Oxidación de C (respiración): CH <sub>2</sub> O + O <sub>2</sub> → CO <sub>2</sub> + H <sub>2</sub> O	Fermentación, desnitrificación: 5CH <sub>2</sub> O + 2H <sub>2</sub> O → HCO <sub>3</sub> <sup>-</sup> + H <sup>+</sup> + 4CH <sub>3</sub> O y 5CH <sub>2</sub> O + 4NO <sub>3</sub> <sup>-</sup> → 5HCO <sub>3</sub> <sup>-</sup> + H <sup>+</sup> + 2N <sub>2</sub> + 2H <sub>2</sub> O
Oxidadores de sulfuros	Oxidación de sulfuros: H <sub>2</sub> S + 2O <sub>2</sub> → SO <sub>4</sub> <sup>2-</sup> + 2H <sup>+</sup> (puede estar acoplada a la fijación de CO <sub>2</sub> )	Desnitrificación, fermentación: 5HS <sup>-</sup> + 8NO <sub>3</sub> <sup>-</sup> → 5SO <sub>4</sub> <sup>2-</sup> + 4N <sub>2</sub> + H <sub>2</sub> O + 3OH <sup>-</sup>
Fotótrofos oxidadores de sulfuros	Fijación de CO <sub>2</sub> (fotosíntesis anóxica acoplada a la oxidación de sulfuro): 2CO <sub>2</sub> + H <sub>2</sub> S + 2H <sub>2</sub> O → 2CH <sub>2</sub> O + SO <sub>4</sub> <sup>2-</sup> + 2H <sup>+</sup>	Fermentación, síntesis de Bcl <sub>a</sub> , degradación de glucógeno
Heterótrofos anaeróbicos reductores de sulfatos	Oxidación de C (sulfato como aceptor de electrones): 2CH <sub>2</sub> O + SO <sub>4</sub> <sup>2-</sup> → 2HCO <sub>3</sub> <sup>-</sup> + 2H <sub>2</sub> S	Lo mismo que en el día
Heterótrofos anaeróbicos metanógenos	Respiración de carbonato: 4H <sub>2</sub> + CO <sub>2</sub> → CH <sub>4</sub> + 2H <sub>2</sub> O y 2CH <sub>2</sub> O → CH <sub>4</sub> + CO <sub>2</sub>	Lo mismo que en el día



En estas comunidades, el ciclaje de elementos necesarios para la vida es muy eficiente y funcionan como bioreactores, en donde prácticamente solo se requiere de luz para funcionar y por medio de la fotosíntesis, se produce el poder reductor para sostener la fijación de carbono y nitrógeno (Des Marais, 2003; Dupraz *et al.*, 2004; Dupraz *et al.*, 2009, Visscher & Stolz, 2005). Otra característica metabólica importante es la secreción extracelular de sustancias poliméricas (*extracellular polymeric secretions*, EPS), principalmente carbohidratos y aminoácidos (Decho *et al.*, 2005). La formación de una matriz de EPS influye en la arquitectura de la comunidad, ya que funciona como una trampa de sedimento y matriz de quelación de carbonato precipitado (Kawaguchi & Decho, 2002).

Existen diferencias entre las comunidades formadoras de tapetes microbianos y aquellas que tienen capacidad de litificación, lo cual conduce a la formación de microbialitos (Figura 2), que son estructuras órgano-sedimentarias formadas por la captura y unión de sedimentos, así como la precipitación de carbonatos asociadas a la actividad metabólica microbiana. Esta característica es importante, ya que mediante la formación de carbonato, estas comunidades pueden secuestrar carbono a largo plazo en forma de mineral (Dupraz *et al.*, 2004; Dupraz *et al.*, 2009).



**Figura 2.** Microbialitos de Pozas Azules, Cuatro Ciénegas, Coahuila, México. A la Izquierda una foto panorámica de los microbialitos en las pozas, a la derecha un microbialito con un corte transversal. Imágenes cortesía de la Dra. Patricia Valdespino Castillo.

## ***Estudio de la importancia del metabolismo bacteriano en las comunidades***

Con el avance en las tecnologías de secuenciación de ADN, se hizo posible la secuenciación masiva de muestras ambientales, surgiendo la metagenómica como herramienta en el estudio de las comunidades microbianas. En el enfoque metagenómico se puede obtener el material genético de la comunidad sin el problema que conlleva el cultivar a las especies microbianas que la estructuran, tomando un enfoque a nivel de sistema (Vieites *et al.*, 2010). Por ello se han generado millones de datos de secuencias obtenidas de muestras ambientales. En paralelo también se han desarrollado herramientas computacionales para el ensamblado de los fragmentos de secuencias y anotación de esta gran cantidad de información (Prakash & Taylor, 2012).

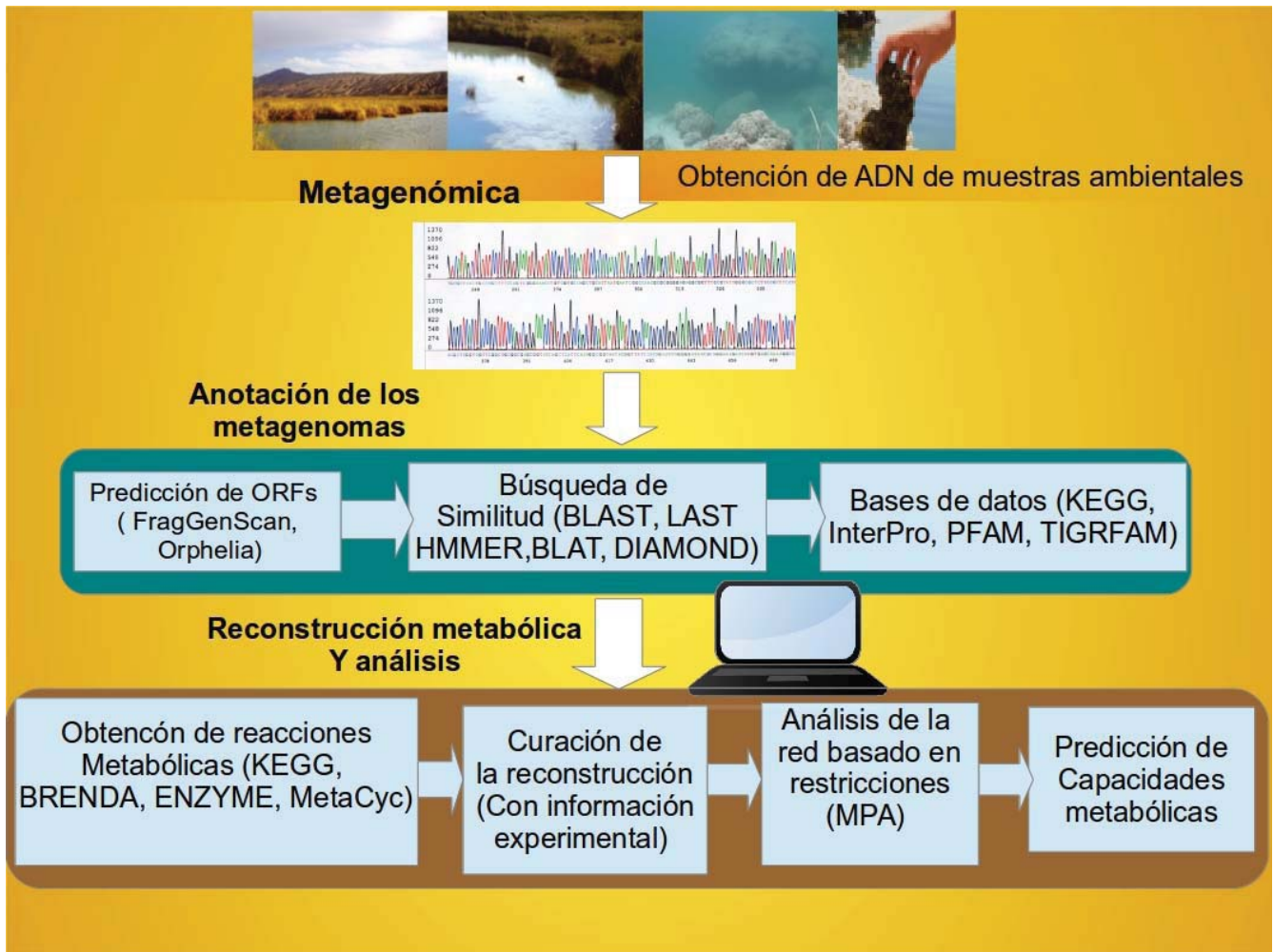
En estudios con fines de descripción de especies en las comunidades microbianas, se han usado marcadores, incluyendo secuencias altamente conservadas del gen 16S rDNA. De esta manera se ha obtenido un panorama de los tipos de bacterias que las habitan. Sin embargo, estos enfoques no han sido tan útiles para obtener conocimiento sobre las características metabólicas funcionales (Chistoserdova, 2014). Para un estudio con un enfoque funcional, se pueden utilizar los metagenomas completos de las comunidades y de aquí surge la importancia de tener una anotación lo más precisa posible de las funciones genéticas codificadas.

## ***Anotación y reconstrucción de modelos metabólicos a partir de metagenomas***

La anotación de las secuencias metagenómicas requiere de herramientas bioinformáticas que se basan en predicción de genes y búsqueda de homología en secuencias. Para esto existen *pipelines*, que consisten en una serie de pasos que parten de las lecturas (*reads*) obtenidas en la secuenciación, hasta la anotación funcional (como el mostrado en la Figura 1), donde se inicia con la predicción de los marcos de lectura de los genes (*Open Reading Frames-ORFs*). Para la predicción de los ORFs se utilizan programas como fragGenScan (Rho *et al.*, 2010), Prodigal (Hyatt *et al.*, 2010) u Orphelia (Hoff *et al.*, 2008). Una vez predichos los marcos de lectura, el siguiente paso es asignarles una función usando programas como BLAST (Altschul *et al.*, 1997), y HMMER (Eddy, 2011). Debido a la gran cantidad de datos, puede haber un cuello de botella en los análisis debido al requerimiento computacional, por ello se han desarrollado nuevos algoritmos como GHOSTX (Suzuki *et al.*, 2014), BLAT (Kent, 2002), LAST (Frith *et al.*, 2010) y DIAMOND (Buchfink *et al.*, 2015), los cuales aprovechan los sistemas con

múltiples procesadores para hacer un análisis más rápido.

Muchas veces mediante el análisis de homología de las secuencias, se pueden asignar funciones a los genes pertenecientes a ciertas funciones celulares que se conservan en muchos organismos. Por ejemplo, en el caso del metabolismo, muchas de las secuencias de los genes pertenecientes a rutas como glucólisis, gluconeogénesis, TCA, entre otros, son altamente conservadas. Cuando esto no es posible, las funciones de los genes se pueden asignar utilizando los motivos característicos de los sitios catalíticos de las enzimas, utilizando bases de datos como Pfam (Punta *et al.*, 2012), KEGG (Kanehisa *et al.*, 2012), TIGRFAM (Haft *et al.*, 2003), UniProt (Apweiler *et al.*, 2004) y COG (Tatusov *et al.*, 2003). El costo computacional que se requiere para la anotación adecuada de los genes va aumentando de acuerdo a la cantidad de las secuencias a analizar, por ello han surgido plataformas web como MG-RAST (Meyer *et al.*, 2008) y EBI Metagenomics (Mitchell *et al.*, 2015), que funcionan como repositorios y las secuencias son analizadas de una manera automática.

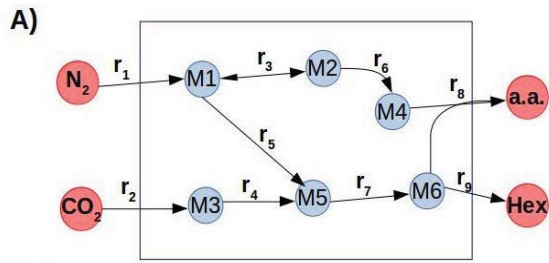


**Figura 3.** Diagrama de flujo del proceso para la anotación funcional y reconstrucción de la red metabólica y el análisis de las capacidades metabólicas de una comunidad por medio de modelos estequiométricos.

Una herramienta prometedora para el aprovechamiento de la información de los metagenomas, es la construcción de modelos matemáticos de los metabolismos, basados en la estequiometría de las reacciones enzimáticas que incluyen restricciones fisicoquímicas (Gombert & Nielsen 2000; Klitgord & Segre, 2011). Estos modelos son el resultado de nuevos paradigmas como la Biología de Sistemas y han sido útiles en aplicaciones de Genómica Funcional e Ingeniería Metabólica. Los modelos estequiométricos representan en un modelo matemático las propiedades metabólicas de los sistemas biológicos. A diferencia de los modelos mecanísticos, no requieren de parámetros cinéticos, lo que los hace atractivos ya que han permitido aplicar la visión de sistema en estudios metabólicos al incluir todas las reacciones involucradas en el metabolismo celular (Trinh *et al.*, 2009). Esto representa una gran ventaja en el caso del estudio de comunidades debido a la dificultad de cultivar a las especies que las componen y poder extraer los parámetros cinéticos.

El modelo parte de una reconstrucción de la red metabólica del sistema a estudiar (ver Figura 1 y 2). Ésta consiste en catalogar el total de las reacciones metabólicas que conforman al sistema bioquímico, en este caso la catalogación de las reacciones metabólicas presentes en la comunidad. Originalmente los modelos metabólicos a escala genómica se desarrollaron para microorganismos modelo como *E. coli* y *Saccharomyces cerevisiae*. En los modelos a escala genómica de estos organismos, la curación de la reconstrucción se hace con información experimental y en el caso de no tenerla, una forma de asignar reacciones bioquímicas es buscando con métodos bioinformáticos los sitios catalíticos de las enzimas, la presencia de los genes que codifiquen el total de las subunidades y los operones (Feist *et al.*, 2009; Palsson, 2004; Schellenberger *et al.*, 2011). En el caso de una reconstrucción a partir de un metagenoma, es difícil recurrir a esta información ya que es complicado ensamblar los genomas completos de los organismos y tener las secuencias de las subunidades completas de las enzimas, debido a que el ADN está presente en segmentos de secuencias de muchas especies (cientos o miles), que son secuenciadas al mismo tiempo. Sin embargo, para mejorar el modelo se puede recurrir a las mediciones experimentales de procesos metabólicos que se encuentren registradas en la literatura, como por ejemplo, las tasas de fijación de carbono, nitrógeno, nitrificación, etc., medidas en los sistemas de estudio, con ello se puede tener un punto de partida en el caso de haber incertidumbre en un cierto número de enzimas encontradas para un proceso metabólico.

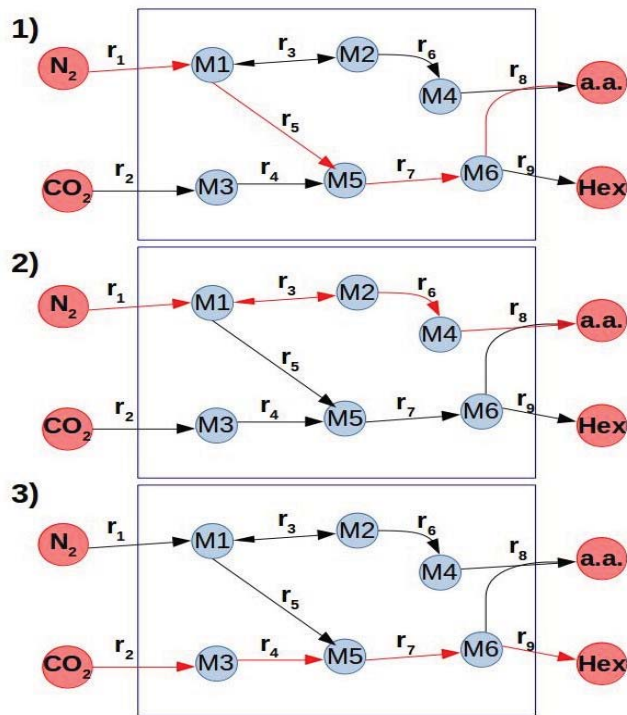




B)

	r <sub>1</sub>	r <sub>2</sub>	r <sub>3</sub>	r <sub>4</sub>	r <sub>5</sub>	r <sub>6</sub>	r <sub>7</sub>	r <sub>8</sub>	r <sub>9</sub>
M1	1	0	-1	0	-1	0	0	0	0
M2	0	0	1	0	0	-1	0	0	0
M3	0	1	0	-1	0	0	0	0	0
M4	0	0	0	0	0	1	0	-1	0
M5	0	0	0	1	1	0	-1	0	0
M6	0	0	0	0	0	0	1	-1	-1

**Figura 4.** Diagrama de una red metabólica hipotética con su matriz estequiométrica asociada. A) la red metabólica está compuesta por seis metabolitos internos (en color azul) y cuatro externos (en rojo), los cuales son transformados por 9 reacciones (flechas negras), los metabolitos externos son fuentes (en este caso hipotético es N<sub>2</sub> y CO<sub>2</sub>) y sumideros (en este caso se produce Hex = hexosa y a.a. = aminoácido) de materia en el sistema (entradas y salidas). B) la matriz estequiométrica asociada a la red de tamaño  $m \times n$ , el coeficiente 1 se asigna si una reacción  $r_n$  produce un metabolito  $M_x$ , el -1 si la reacción consume el metabolito y el 0 si el metabolito no participa en esa reacción.



**Figura 5.** Diagrama que muestra los modos elementales de flujo (EFM) de la red hipotética de la Figura 4. En flechas rojas se muestran los EFM, 1) y 2) producen a.a. a partir de N<sub>2</sub>, 3) produce Hex a partir de CO<sub>2</sub>. El análisis de EFM descompone a la red metabólica en grupos mínimos de enzimas que sostienen el estado estacionario del sistema.

Una vez que se ha reconstruido la red, se procede a hacer un análisis de vías metabólicas (*Metabolic Pathway Analysis*-MPA), el cual es un análisis de la estructura de la red (Trinh et al., 2009). Las ecuaciones matemáticas para el análisis se basan en el principio de conservación de masa de los metabolitos internos (ver el Cuadro 1) (Papin *et al.*, 2004; Palsson, 2006). La ecuación general que describe este principio es:

$$\frac{dC}{dt} = S * r - \mu * C \quad \text{Ecuación (1)}$$

En donde el término del lado izquierdo es el cambio de la concentración de los metabolitos internos  $C$  a través del tiempo  $t$ ;  $r$  es el vector de la tasa de las reacciones (flujo) que interconvierten a los metabolitos;  $S$  es la matriz estequiométrica de tamaño  $m \times n$ , cuyos elementos  $S_{ij}$  representan al coeficiente estequiométrico del metabolito  $i$  involucrado en la reacción  $j$  y  $\mu$  es la tasa de dilución específica con el cambio en el volumen del sistema.

A partir del balance general representado en la Ecuación 1, se construye un sistema de ecuaciones diferenciales ordinarias con un número de ecuaciones equivalentes a la cantidad de metabolitos internos. En un sistema bioquímico complejo sería una cantidad de ecuaciones quizás sin solución. Sin embargo, esto se puede simplificar tomando en cuenta que en un sistema biológico, el orden del segundo término, el cual involucra  $\mu$ , es menor, por lo que es despreciable (Villadsen *et al.*, 2011).

Una consideración importante en estos modelos es el estado pseudo-estacionario. Es decir, se considera que la replicación celular alcanza un estado estacionario, y que debido al orden de magnitud del tiempo que toma una reacción cinética es mucho menor al tiempo de replicación celular, lo que implica que la concentración de metabolitos se puede ver como si no cambiara en el tiempo. Haciendo a la concentración  $C$  constante e independiente del tiempo, nos encontramos en un estado estacionario en donde no hay acumulación de metabolitos internos:

$$\frac{dC}{dt} = 0$$

Por lo tanto, la ecuación general se transforma en un sistema de ecuaciones lineales:

$$S * r = 0 \quad \text{Ecuación (2)}$$

Debido a que algunas reacciones son reversibles, las reacciones  $r_i$  en  $r$  requieren de restricciones adicionales en valores positivos  $r_i \geq 0$ .

El conjunto de todas las soluciones posibles para la Ecuación 2 se puede determinar mediante un

análisis convexo, en donde el espacio solución del flujo admisible está delimitado por los vectores base que forman un cono convexo (Palsson, 2006; Schuster *et al.*, 2007; Gombert & Nielsen, 2011). Para el análisis existe un algoritmo que parte de la búsqueda en el total de las reacciones y determina los modos elementales de flujo (*Elementary Flux Modes*, EFM), los cuales son las combinaciones lineales de los vectores base del cono convexo, por lo que determina todas las soluciones posibles (Orth *et al.*, 2010; Schuster *et al.*, 2007; Latendresse *et al.*, 2012). El algoritmo de los EFM es el más adecuado en el caso del estudio de las comunidades arriba descritas, ya que puede localizar modos de flujo (equivalentes a rutas metabólicas, ver Figura 3 y Cuadro 1) que no se conozcan y así se pueden encontrar rutas potenciales de C, H, O, N, P y S dentro de la comunidad (Taffs *et al.*, 2009).

Los modos elementales de flujo no se pueden calcular en redes muy grandes. Esto se debe a que se presenta una “explosión combinatoria” de posibles soluciones con el aumento del número de reacciones en la red (Stelling *et al.*, 2002), por lo que existe una limitante técnica por el uso excesivo de la memoria del equipo de cómputo. Sin embargo se han desarrollado paquetes de software con mejoras en este aspecto, como EFMTools, que está implementado en MATLAB©, y puede analizar redes utilizando multinúcleos. CellNetAnalyzer es una *toolbox* implementada en MATLAB© para analizar redes metabólicas, de señalización y regulación. EFMTools viene implementado dentro de CellNetAnalyzer (Klamt *et al.*, 2007).

**Cuadro 3.** Significado de términos utilizados.

---

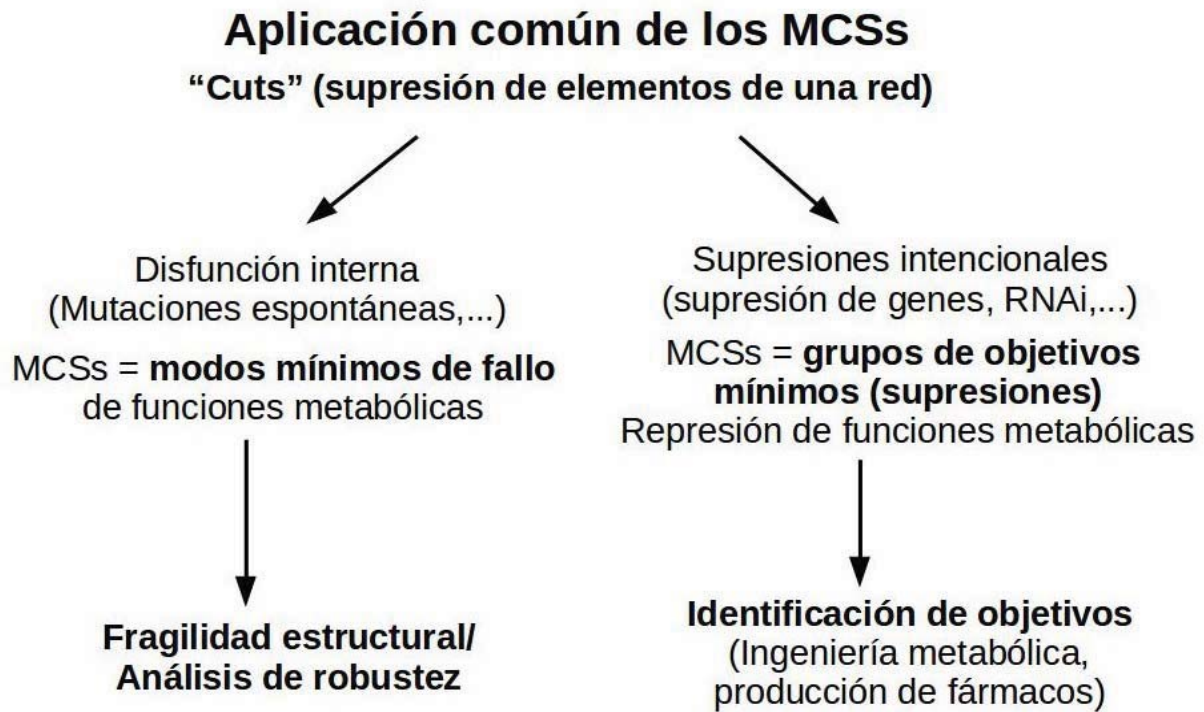
Reconstrucción y análisis metabólico basado en restricciones	Métodos de modelación en donde se hacen reconstrucciones de redes metabólicas que son curadas manualmente. Se incluyen restricciones termodinámicas (direccionalidad de las reacciones), de conservación de masa (de acuerdo a su estequiometría), restricciones detectadas experimentalmente y datos de regulación.
Metabolito interno	Metabolitos de reacciones que participan dentro del modelo. Debe haber un balance entre su formación y su consumo (asumiendo un estado estacionario)
Metabolito externo	Metabolitos que no necesariamente son balanceados debido a que se pueden acumular o están conectados a reservorios. Son considerados como recursos o sumideros (nutrientes o productos de desecho, productos almacenados o excretados y precursores para su utilización posterior).
Modo elemental de flujo (EFM)	Grupo mínimo de enzimas que funcionan en estado estacionario con todas las reacciones irreversibles procediendo en la dirección apropiada.
Topología de una red metabólica	Conjunto de enzimas y metabolitos, conectadas con sus reacciones estequiométricas en direccionalidad correcta.
Vector de flujo metabólico	Representación matemática del flujo metabólico en forma de vectores. Un vector de flujo corresponde a un vector en el espacio vectorial/solución de un análisis convexo.

---

Otro aspecto importante en los sistemas bioquímicos es su fragilidad y robustez. En el caso de las comunidades biológicas, es de suma importancia conocer qué tanto pueden resistir a condiciones cambiantes y perturbaciones. Predecir la importancia de un tipo de metabolismo (asociado a algún grupo de especies) en el mantenimiento de las comunidades es un objetivo clave en la ecología. Una aproximación se puede hacer por medio de los *minimal cut sets* (MCS), los cuales se determinan a partir de los EFMs calculados (Klamt & Gilles, 2004). Si los EFMs son unidades funcionales mínimas, los MCSs son los modos mínimos de fallo, que consisten en grupos mínimos de reacciones que en su ausencia, llevarían a la falla del sistema, rompiendo el estado estacionario y el flujo metabólico. Los MCSs se usan para determinar si la supresión de un cierto número de enzimas, interrumpe el flujo de un metabolito deseado (por ejemplo alguna toxina o la producción de biomasa) o si interrumpe todo el flujo del sistema (determinar la fragilidad total), con el fin de identificar las enzimas esenciales (o metabolismos esenciales) (Klamt, 2006). El algoritmo para calcular los MCSs viene integrado en CellNetAnalyzer y calcula la fragilidad total, dada por el coeficiente de fragilidad  $\mathbf{F}$ , el cual es el valor promedio de  $Fi$  entre todas las reacciones ( $q$ ):

$$\mathbf{F} = \frac{\sum_i Fi}{q}$$

En donde  $Fi$  es el coeficiente de fragilidad específico descrito por Klamt y Gilles (Klamt & Gilles, 2004) y corresponde al recíproco del tamaño promedio de todos los MCSs en los cuales la reacción  $i$  participa.  $\mathbf{F}$  toma valores de 0 a 1; el más alto valor se da cuando todas las reacciones son esenciales. Los MCSs de menor tamaño son más representativos, ya que nos dicen que se requiere de una menor perturbación (menor supresión de enzimas) para que el flujo metabólico se interrumpa.



**Figura 6.** Los dos principales campos de aplicación para los MCSs (modificada de Klamt, 2006)

En este proyecto usamos el enfoque de los EFM, primero en la reconstrucción metabólica de *C. chlorochromatii* a partir de su genoma (No. de acceso NC\_007514.1), para el análisis metabólico en un consorcio modelo de dos especies. Luego para el análisis a nivel comunidad, usamos un conjunto de diez metagenomas de microbialitos y tapetes microbianos (Cuadro 4), con el objetivo de predecir el intercambio de carbono y nitrógeno entre la comunidad y el ambiente, así como predecir el flujo metabólico a nivel comunitario.



**Cuadro 4.** Metagenomas usados obtenidos de la base de datos del MG-RAST

MG-RAST ID	Lugar	Referencia
4440067.3	Microbialito de Pozas Azules	(Breitbard <i>et al.</i> , 2009)
4440060.4	Microbialito de Rio Mezquites	(Breitbard <i>et al.</i> , 2009)
4440061.3	Microbialito de Highborne Cay, Bahamas	(Dinsdale <i>et al.</i> , 2008)
4441363.3	Tapete microbiano verde, Cuatro Ciénegas	(Peimbert <i>et al.</i> , 2012)
4442466.3	Tapete microbiano rojo, Cuatro Ciénegas	(Peimbert <i>et al.</i> , 2012)
4445126.3	Tapete microbiano en plataforma Markham en el ártico	(Varin <i>et al.</i> , 2012)
4445129.3	Tapete microbiano de la plataforma Ward Hunt en el ártico	(Varin <i>et al.</i> , 2012)
4460448.3	Tapete microbiano de Diamond Fork	(Gomez-Alvarez <i>et al.</i> , 2012)
4460449.3	Tapete microbiano de Diamond Fork	(Gomez-Alvarez <i>et al.</i> , 2012)
4454153.3	Tapete microbiano de la marisma de Sippewissett	(Wilbanks <i>et al.</i> , 2014)

## Objetivos

### *Objetivo general*

Reconstruir las redes metabólicas de comunidades microbianas modelo, a partir de datos genómicos y metagenómicos y analizarlas por medio de los EFMs para predecir el intercambio de materia y energía.

### *Objetivos específicos*

- Predecir las consecuencias metabólicas de la simbiosis en el consorcio *Chlorochromatium aggregatum*.
- Predecir el potencial metabólico autotrófico de una comunidad microbiana.

## Capítulo 1

- *Cerqueda-García, D., & Falcón, L. I. (2016). La construcción del nicho y el concepto de holobionte, hacia la reestructuración de un paradigma. Revista Mexicana de Biodiversidad, 87(1), 239-241.*



Nota de opinión

## La construcción del nicho y el concepto de holobionte, hacia la reestructuración de un paradigma

*The niche-construction and the holobiont concept, toward the restructuring of a paradigm*

Daniel Cerqueda-García y Luisa I. Falcón\*

Laboratorio de Bacteriana, Instituto de Ecología, Universidad Nacional Autónoma de México, Circuito exterior s/n, Cd. Universitaria, 04510, México D.F., México

En la concepción clásica de la adaptación, el ambiente gradualmente moldea los organismos para que estos se adecuen. Este enfoque implicaba que los organismos son un ente pasivo que habita el ambiente. Sin embargo, esto no es así, ya que los organismos no son agentes pasivos. Los organismos, por medio de su metabolismo y su comportamiento, modifican su medio, creando y destruyendo su ambiente desde una escala local hasta global. Algunos ejemplos de modificación del ambiente son la construcción de nidos, madrigueras, redes y formación de gradientes químicos. Estos procesos de modificación y selección ambiental se conocen con la metáfora de «construcción del nicho», considerando aquí el espacio «n» dimensional (según Hutchinson, 1957) «ocupado» por la especie, en el sentido de donde actúa —siendo un agente causal—, no solo en donde se localiza. La construcción del nicho no es exclusiva de algunas especies, sino que es un hecho de la vida, todos los organismos utilizan materia inorgánica para desarrollarse y sustentarse; además, todos secretan desechos (Day, Laland y Odling-Smee, 2003; Laland, Odling-Smee y Feldman, 2001; Odling-Smee, Laland y Feldman, 1996).

La construcción del nicho tiene un gran impacto en la evolución de los organismos y en la concepción de la adaptación, ya que modifica las presiones de selección natural al modificar el medio; según Day et al. (2003), en la naturaleza pueden existir 2 escenarios: 1) como resultado de la selección natural los organismos pueden desarrollar características que los hagan adecuarse a su ambiente; y 2) los organismos pueden modificar su ambiente para adecuarlo a sus características actuales; es decir, modificar factores ambientales de la estructura del nicho como sucede en la producción de microclimas, lo que modifica la temperatura y la humedad.

La construcción del nicho puede modificar las presiones de selección de los constructores, de las especies con las que cohabitan —el componente biótico del nicho— y de sus descendientes, ya que generalmente la modificación persiste por mayor tiempo que la vida de los constructores, por lo que la modulación del impacto ambiental continúa para las generaciones posteriores. Por lo tanto, los organismos no nacen y empiezan sus ciclos de vida en un ambiente *de novo*, sino que existe una herencia ecológica del nicho. La herencia ecológica se define como cualquier caso en el cual un organismo experimenta una relación funcional modificada entre él y su ambiente, como consecuencia de la construcción del nicho de sus ancestros genéticos o ecológicos (Laland et al., 2001). De aquí surge el concepto de «ingeniería de ecosistemas», donde se argumenta que los organismos no solo se adaptan a su ambiente, sino que también lo construyen, por lo que existe una coevolución entre el ambiente y el organismo, existiendo una retroalimentación. Un ejemplo claro de la construcción del nicho es la modificación de la atmósfera por parte de las bacterias fotosintéticas oxigénicas (Falkowski, 2011), que a una escala global modificaron las presiones de selección hacia metabolismos que toleran el oxígeno, siendo estas bacterias nuestros ancestros ecológicos.

Si bien hasta aquí se ha considerado la construcción del nicho refiriéndonos a la modificación del componente fisicoquímico —abiótico—, que a la vez afecta a las presiones de selección de otros organismos —indirectamente—, existen casos en donde este ambiente es conformado por otro organismo, como en los organismos pluricelulares, ya que surgieron en una «sopa microbiana» en donde rápidamente pudieron emerger interacciones patógenas y mutualistas, reteniendo selectivamente colonizadores que conferían un beneficio al huésped aumentando su adecuación (Yeoman et al., 2011). Toda planta o animal presenta relaciones simbióticas con microorganismos, de ahí surge el concepto de «holobionte», que se refiere a la planta o animal con todos sus microorganismos asociados —microbioma— (Zilber-Rosenberg y Rosenberg, 2008). Sabemos que 9 de cada

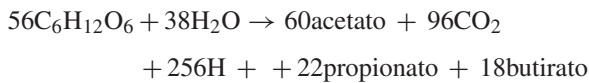
\* Autor para correspondencia.

Correo electrónico: [falcon@ecologia.unam.mx](mailto:falcon@ecologia.unam.mx) (L.I. Falcón).

La revisión por pares es responsabilidad de la Universidad Nacional Autónoma de México.

10 células de los seres humanos son bacterianas y, por lo tanto, un animal o planta puede ser considerado como un ecosistema, el cual posee numerosos nichos epiteliales. Dependiendo de la localización del nicho puede estar colonizado por bacterias, arqueas, bacteriófagos, hongos y/o protozoarios (Yeoman et al., 2011).

Los componentes del microbioma son un claro ejemplo de organismos constructores de nicho, como lo que ocurre en el conducto vaginal de las hembras humanas, que es colonizado por *Lactobacillus* sp., que sintetizan ácido láctico y, como consecuencia, el pH vaginal disminuye, atenuando la colonización de otros organismos patógenos (Menard et al., 2010); otro ejemplo es el que ocurre en el tracto digestivo de los mamíferos, en donde bacterias metanógenas y reductoras de sulfato subsisten gracias al hidrógeno producido por bacterias fermentadoras (Macfarlane y Gibson, 1997), ya que en la reacción general de la fermentación se produce hidrógeno:



Las bacterias metanógenas lo pueden utilizar para la producción de metano:



Y las bacterias reductoras del azufre para la reducción disimilatoria del azufre:



Este es un ejemplo donde al construirse y modificarse el ambiente se abren nuevos nichos —o condiciones favorables— para ser ocupados por otras bacterias. Por ello, los organismos multicelulares eucariontes son un ecosistema constantemente modificado por su microbioma, dando como consecuencia una intrincada red de interacciones simbióticas. La construcción del nicho puede ser el preámbulo para la facilitación presente en muchas especies, que principalmente se ha descrito en plantas (Valiente-Banuet y Verdú, 2007), pero que en las relaciones entre microorganismos formadores de microbioma queda por analizarse. Las interacciones facilitadoras son comunes en la complementación metabólica bacteriana, como ocurre en las comunidades formadoras de microbialitos, en donde debido a la presencia de diversos gremios bacterianos funcionan como una unidad en donde prácticamente solo se requiere luz y materia inorgánica para su desarrollo (Visscher y Stolz, 2005).

El holobionte es, por lo tanto, una comunidad biótica que se retroalimenta debido a las actividades y características de sus componentes, pero entonces ¿qué implicaciones tiene esto en las presiones de selección y la adaptación? Para responder a esto se ha considerado a toda la comunidad —al holobionte— como una unidad de selección, ya que la cooperación entre la microbiota y el hospedero generalmente producen un aumento en la adecuación. Asimismo, la diversidad genética del microbioma puede aumentar el intervalo de tolerancia a los ambientes en los cuales el holobionte puede vivir (Rosenberg, Sharon y Zilber-Rosenberg, 2009; Zilber-Rosenberg y Rosenberg, 2008).

La presencia de algunas especies está facilitando el desarrollo de otras por medio de la complementación metabólica, por ende la estructura comunitaria está funcionando como una unidad. Además, la diversidad genética del holobionte, conocido como «hologenoma», el cual es la suma de la diversidad genética del microbioma y el hospedero, varía entre holobiontes y también es transferida entre ellos. Esto último parece tener características lamarckianas, ya que se pueden perder y ganar habilidades —por medio de la adquisición o pérdida de componentes del microbioma— que modifiquen la adecuación en una misma generación (Rosenberg et al., 2009). Un claro ejemplo de lo anterior es la transferencia del microbioma materno en el nacimiento de los humanos, que en un parto por cesárea el microbioma materno no se transfiere y el infante presenta posteriormente una mayor propensión a enfermedades inmunes, tales como el asma, y una mayor propensión a alergias (Madan, Farzan, Hibberd y Karagas, 2012; Thanabalasuriar y Kubes, 2014). Igualmente, la alimentación de los recién nacidos con leche materna resulta en la colonización del tracto digestivo por un microbioma que digiere los oligosacáridos presentes en la leche, conformado por bacterias del género *Bifidobacterium* sp., las cuales ayudan a disminuir el pH del tracto digestivo y producen ácidos grasos de cadena corta que favorecen la función del sistema inmune (Hinde y Lewis, 2015). La estructura del holobionte nos lleva a considerar que la selección actúa a múltiples niveles, desde genes a genomas, a especies y a comunidades. Dentro del propio microbioma pueden presentarse todas las interacciones ecológicas —interacciones como depredador-presa y cooperación-facilitación—. Esto implica que la estructura del microbioma no solo está determinada por las características del huésped, como en el caso del microbioma del tracto digestivo de animales, en donde se ha pensado que únicamente la dieta del huésped modifica la estructura del microbioma, la cual es una perspectiva clásica en donde el medio moldea a los organismos. Desde la perspectiva de la construcción del nicho y el concepto de holobionte, las interacciones bióticas en el microbioma y las modificaciones emergentes de las presiones de selección pueden tener un papel importante en la determinación de la estructura del mismo.

En resumen, el descubrimiento del microbioma abre nuevas posibilidades para la investigación sobre los factores que influyen en la evolución de las especies. Si bien la idea de que la selección natural actúa a diferentes niveles de la jerarquía biológica no es nueva, y biólogos como Lewontin y Gould, desde los años 70 del siglo pasado, ya discutían sobre cuál es la unidad darwiniana de la selección y a qué niveles actúa (Gould y Lloyd, 1999; Lewontin, 1970; Levins y Lewontin, 1980; Vrba y Gould, 1986), en donde la unidad darwiniana debe cumplir principios básicos para ser un agente evolutivo como un punto distinguible de nacimiento, un punto distinguible de muerte y una estabilidad suficiente, además, debe dejar descendencia y esta descendencia debe ser más parecida a este «individuo darwiniano» que a los otros miembros de la generación parental. En el holobionte —la comunidad que lo conforma— se cumplen estos principios vernáculos para la unidad de selección. Aunque la conceptualización de la comunidad como unidad de selección ha sido criticada por el hecho de que puede violar el principio

de suficiente estabilidad (Gould, 2002), en la comunidad conformada por el holobionte este principio se cumple, ya que el tiempo de estabilidad es generalmente el tiempo generacional del organismo hospedero pluricelular.

## Referencias

- Day, R. L., Laland, K. N. y Odling-Smee, F. J. (2003). Rethinking adaptation: the niche-construction perspective. *Perspectives in Biology and Medicine*, 46, 80–95.
- Falkowski, P. G. (2011). The biological and geological contingencies for the rise of oxygen on Earth. *Photosynthesis Research*, 107, 7–10.
- Gould, S. J. (2002). *The structure of evolutionary theory*. Cambridge, Massachusetts and London: The Belknap Press Harvard University Press.
- Gould, S. J. y Lloyd, E. A. (1999). Individuality and adaptation across levels of selection: how shall we name and generalize the unit of Darwinism? *Proceedings of the National Academy of Science*, 96, 11904–11909.
- Hinde, K. y Lewis, Z. T. (2015). Mother's littlest helpers. *Science*, 348, 1427–1428.
- Hutchinson, G. E. (1957, January). Concluding remarks. *Cold spring harbor symposium on quantitative biology*, 22, 415–427.
- Laland, K. N., Odling-Smee, F. J. y Feldman, M. W. (2001). Niche construction, ecological inheritance, and cycles of contingency in evolution. En S. Oyama, P. E. Griffiths, y D. G. Russell (Eds.), *Cycles of contingency: developmental systems and evolution* (pp. 117–126). Cambridge, London: The MIT Press.
- Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics*, 1–18.
- Levins, R. y Lewontin, R. (1980). Dialectics and reductionism in ecology. *Synthese*, 43, 47–78.
- Macfarlane, G. T. y Gibson, G. R. (1997). Carbohydrate fermentation, energy transduction and gas metabolism in the human large intestine. In *Gastrointestinal microbiology*. pp. 269–318. New York: Springer.
- Madan, J. C., Farzan, S. F., Hibberd, P. L. y Karagas, M. R. (2012). Normal neonatal microbiome variation in relation to environmental factors, infection and allergy. *Current Opinion in Pediatrics*, 24, 753.
- Menard, J. P., Mazouni, C., Salem-Cherif, I., Fenollar, F., Raoult, D., Boubli, L., et al. (2010). High vaginal concentrations of *Atopobium vaginae* and *Gardnerella vaginalis* in women undergoing preterm labor. *Obstetrics and Gynecology*, 115, 134–140.
- Odling-Smee, F. J., Laland, K. N. y Feldman, M. W. (1996). Niche construction. *American Naturalist*, 147, 641–648.
- Rosenberg, E., Sharon, G. y Zilber-Rosenberg, I. (2009). The hologenome theory of evolution contains Lamarckian aspects within a Darwinian framework. *Environmental Microbiology*, 11, 2959–2962.
- Thanabalasuriar, A. y Kubes, P. (2014). Neonates, antibiotics and the microbiome. *Nature Medicine*, 20, 469–470.
- Valiente-Banuet, A. y Verdú, M. (2007). Facilitation can increase the phylogenetic diversity of plant communities. *Ecology Letters*, 10, 1029–1036.
- Visscher, P. T. y Stolz, J. F. (2005). Microbial mats as bioreactors: Populations, processes, and products. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 219, 87–100.
- Vrba, E. S. y Gould, S. J. (1986). The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. *Paleobiology*, 12, 217–228.
- Yeoman, C. J., Chia, N., Yildirim, S., Miller, M. E. B., Kent, A., Stumpf, R., et al. (2011). Towards an evolutionary model of animal-associated microbiomes. *Entropy*, 13, 570–594.
- Zilber-Rosenberg, I. y Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Reviews*, 32, 723–735.

## Capítulo 2

- *Cerqueda-García, D., Martínez-Castilla, L. P., Falcón, L. I., & Delaye, L. (2014). Metabolic analysis of Chlorobium chlorochromatii CaD3 reveals clues of the symbiosis in 'Chlorochromatium aggregatum'. The ISME journal, 8(5), 991-998.*



## ORIGINAL ARTICLE

# Metabolic analysis of *Chlorobium chlorochromatii* CaD3 reveals clues of the symbiosis in ‘*Chlorochromatium aggregatum*’.

Daniel Cerqueda-García<sup>1,2</sup>, León P Martínez-Castilla<sup>1</sup>, Luisa I Falcón<sup>2</sup> and Luis Delaye<sup>3</sup>

<sup>1</sup>Departamento de Bioquímica, Facultad de Química, Universidad Nacional Autónoma de México, México city, México; <sup>2</sup>Instituto de Ecología, Universidad Nacional Autónoma de México. Circuito ext. sn, Cd. Universitaria, México city, México and <sup>3</sup>Departamento de Ingeniería Genética, Centro de Investigación y de Estudios Avanzados (CINVESTAV), Irapuato, México

**A symbiotic association occurs in ‘*Chlorochromatium aggregatum*’, a phototrophic consortium integrated by two species of phylogenetically distant bacteria composed by the green-sulfur *Chlorobium chlorochromatii* CaD3 epibiont that surrounds a central  $\beta$ -proteobacterium. The non-motile *chlorobia* can perform nitrogen and carbon fixation, using sulfide as electron donors for anoxygenic photosynthesis. The consortium can move due to the flagella present in the central  $\beta$ -proteobacterium. Although *Chl. chlorochromatii* CaD3 is never found as free-living bacteria in nature, previous transcriptomic and proteomic studies have revealed that there are differential transcription patterns between the symbiotic and free-living status of *Chl. chlorochromatii* CaD3 when grown in laboratory conditions. The differences occur mainly in genes encoding the enzymatic reactions involved in nitrogen and amino acid metabolism. We performed a metabolic reconstruction of *Chl. chlorochromatii* CaD3 and an *in silico* analysis of its amino acid metabolism using an elementary flux modes approach (EFM). Our study suggests that in symbiosis, *Chl. chlorochromatii* CaD3 is under limited nitrogen conditions where the GS/GOGAT (glutamine synthetase/glutamate synthetase) pathway is actively assimilating ammonia obtained via  $N_2$  fixation. In contrast, when free-living, *Chl. chlorochromatii* CaD3 is in a condition of nitrogen excess and ammonia is assimilated by the alanine dehydrogenase (AlaDH) pathway. We postulate that ‘*Chlorochromatium aggregatum*’ originated from a parasitic interaction where the  $N_2$  fixation capacity of the *chlorobia* would be enhanced by injection of 2-oxoglutarate from the  $\beta$ -proteobacterium via the periplasm. This consortium would have the advantage of motility, which is fundamental to a phototrophic bacterium, and the syntrophy of nitrogen and carbon sources.**

*The ISME Journal* (2014) 8, 991–998; doi:10.1038/ismej.2013.207; published online 28 November 2013

**Subject Category:** Microbe-microbe and microbe-host interactions

**Keywords:** elementary flux modes; syntrophy; *chlorobia*; nitrogen metabolism

## Introduction

Symbiosis is widespread in the microbial world (Moya *et al.*, 2008). Symbiotic interactions can fluctuate between parasitism and mutualism through time (Toft and Andersson, 2010). These interactions are a source of evolutionary innovation through genetic rearrangements that give rise to metabolic capabilities and emergence of syntrophy to exploit resources, stabilizing in a mutualistic relationship (Margulis and Fester, 1991; Moya *et al.*, 2008; Moya and Peretó, 2011).

The phototrophic consortium ‘*Chlorochromatium aggregatum*’ is a model of symbiotic interactions involving prokaryotes. It was isolated from Lake Dagow in Germany (Fröstl and Overmann, 1998) and is formed by green-sulfur bacteria *Chlorobium chlorochromatii* (epibionts) that surround a central, colorless and motile  $\beta$ -proteobacterium, phylogenetically related to Comamonadaceae (Kanzler *et al.*, 2005). *Chl. chlorochromatii* CaD3 belong to the *Chlorobia* phylum. They are Gram-negative, non-motile, anaerobic photosynthetic and diazotrophic bacteria. These bacteria perform anoxic photosynthesis in the chlorosoma, which contains a photosystem I (PSI) reaction center. Their major photosynthetic pigment is bacteriochlorophyll c, and sulfide is used as electron donor.  $CO_2$  is fixed via the reverse tricarboxylic acid cycle (rTCA) (Vogl *et al.*, 2006). In natural conditions they are exclusively found in symbiosis although they can be

Correspondence: L Delaye, Departamento de Ingeniería Genética, Centro de Investigación y de Estudios Avanzados (CINVESTAV), Km. 9.6 Libramiento Norte Carretera Irapuato-León, Irapuato, Guanajuato 36821, México.

E-mail: ldelaye@ira.cinvestav.mx

Received 24 June 2013; revised 25 September 2013; accepted 7 October 2013; published online 28 November 2013

cultured without their symbiont, which is not the case for the obligate-symbiotic  $\beta$ -proteobacterium (Müller and Overmann, 2011). Previous studies have identified genes that code for virulence factors in *Chl. chlorochromatii* CaD3 that seem to be relevant in the symbiosis, one of them seems to have been acquired by horizontal gene transfer (Vogl *et al.*, 2008).

A recent study compared the proteome and gene expression of symbiotic and free-living states in *Chl. chlorochromatii* CaD3 (Wenter *et al.*, 2010). It showed that ~350 genes of *Chl. chlorochromatii* CaD3 have differential expression patterns between both states (Wenter *et al.*, 2010). A significant fraction of these correspond to genes involved in nitrogen and amino acid metabolism. Among these there were genes coding for a nitrogen regulator protein PII, a gene coding a glutamine synthetase (GS) enzyme and the *nifH* gene coding for nitrogenase reductase. These genes seem to be having a regulatory role in nitrogen assimilation between symbiotic and free-living states (Brown and Herbert, 1977; Magasanik, 1982; Forchhamer, 2004; Texeira *et al.*, 2010). In this study, we performed the metabolic network reconstruction of *Chl. chlorochromatii* CaD3 and an elementary flux mode (EFM) analysis (Schuster *et al.*, 1999; Jevremovic *et al.*, 2011) of the reconstructed amino-acid metabolism sub-network, to better understand the metabolic shift between symbiotic and free-living states.

## Materials and methods

### Genomes used

The following genomes were downloaded from the Genbank database: *Chlorobium chlorochromatii* CaD3 (NC\_007514.1), *Chlorobium tepidum* TLS (NC\_002932.3), *Chlorobium phaeobacteroides* BS1 (NC\_010831.1), *Chlorobium phaeobacteroides* DSM 266 (NC\_008639.1), *Chlorobium luteolum* DSM 273 (NC\_007512.1), *Chlorobium phaeovibrioides* DSM 265 (NC\_009337.1), *Chlorobium limicola* DSM 245 (NC\_010803.1), *Pelodictyon phaeoclathratiforme* BU-1 (NC\_011060.1), *Prosthecochloris aestuarii* DSM 271 (NC\_011059.1), *Chloroherpeton thalassium* ATCC 35110 (NC\_011026.1), *Chlorobaculum parvum* NCIB 8327 (NC\_011027.1) and *Escherichia coli* K-12 (MG1655) (NC\_000913.2).

### Reconstruction of the metabolic capabilities of *Chl. chlorochromatii* CaD3

We first reconstructed the metabolism of *Chl. chlorochromatii* CaD3 with Pathway-Tools software (Karp *et al.*, 2010). The Pathway-Tools software inferred all metabolic pathways directly from the Genbank file. Thereby, we obtained a Pathway/Genome Database (PGDB) to be manually curated. As *C. tepidum* is the most studied species of *chlorobia*, we also performed the automatic reconstruction of its metabolism with Pathway-Tools

software from its annotated genome downloaded from Genbank database (Eisen *et al.*, 2002). The automatic reconstruction was curated with overall experimental information, and metabolic function was assigned only if it could be confirmed in the literature. Then, the metabolic reconstruction of *C. tepidum* was used to cure the metabolic reconstruction of *Chl. chlorochromatii* CaD3 (see below).

To improve and curate the initial metabolic reconstruction we performed a search for orthologous proteins between *Chl. chlorochromatii* CaD3, *C. tepidum* and the other species of *Chlorobia* listed under 'Genomes used'. We also searched for orthologous proteins between the *Chlorobia* and *E. coli* and used BranchClust software to find orthologs between those genomes (Poptsova and Gogarten, 2007). We used the *E. coli* genome assuming that it is the best curated bacterial genome and is available through EcoCyc database (Keseler *et al.*, 2011). The genomes of the others species of *Chlorobia* were added to the analysis to improve the performance of BranchClust. For the 12 genomes included in this study, we asked BranchClust to consider as orthologs all monophyletic groups of sequences having at least eight different species represented. For those peptide sequences of *Chl. chlorochromatii* CaD3 that had no orthologs, we looked for conserved protein domains against the Pfam database (Finn *et al.*, 2010) with HMMER3 software (Johnson *et al.*, 2010) by using an *E*-value cutoff of 0.01. Based on the orthologous search with BranchClust and domain identification with Pfam, we assigned 382 enzymatic reactions to the metabolism of *Chl. chlorochromatii* CaD3 (Supplementary Material 1). These reactions were obtained from KEGG (Kanehisa *et al.*, 2012) and BRENDA databases (in their WWW available versions in 2011–2012) (Scheer *et al.*, 2011). Finally the reactions were integrated within YANA software (Schwarz *et al.*, 2005).

### EFM analysis of symbiotic and free-living states

We performed an EFM analysis of the amino-acid metabolism in the symbiotic and free-living states of *Chl. chlorochromatii* CaD3. The EFM algorithm decomposes the metabolic network in EFMs, where each EFM is a minimal set of enzymes operating in steady state with all reactions proceeding in the direction dictated by its thermodynamic restrictions (for theoretical details, see Schuster *et al.*, 1999 and Wagner, 2004) being equivalent to a metabolic pathway that supports the network structure in a specific growth condition. The inferred metabolic network of the amino-acid metabolism consists of 99 enzymatic reactions from the whole reconstruction. These reactions were identified following the metabolic pathway classification of KEGG database. Then, based on the expression profiles published by Wenter *et al.* (2010), we removed with YANA software 5 and 10 reactions not transcribed in symbiotic and free-living states respectively, thus

generating two sub-networks: one with 94 reactions corresponding to the symbiotic state and the other with 89 reactions corresponding to the free-living state (see Supplementary Material 1, 2 and 3).

To identify the EFM of the symbiotic and free-living states, the two sub-networks in METATOOL file format (Schuster *et al.*, 1999; Supplementary Material 1) were analyzed with ElMo-comp-1.0.4 software (Jevremovic *et al.*, 2011). For each sub-network there were two output files: one containing the EFMs found, and the other containing the reactions of each EFM. These output files were analyzed with scripts in python v2.7 to see whether the EFMs in both states (free-living and symbiosis) were the same. We further analyzed the EFMs with ACoM-c software (Pérès *et al.*, 2011). This software groups the EFMs in motifs, which are clusters of EFMs that share at least three reactions. Motifs were visualized with CellDesigner (Funahashi *et al.*, 2007) (Supplementary Material 2). Finally, to prove that the obtained EFM reconstructions were independent of chance, we performed 100 random simulations of the symbiotic and free-living networks as follows: from the whole network of amino-acid metabolism containing 99 reactions, we randomly removed 5 and 10 reactions to simulate symbiotic and free-living states with 94 and 89 reactions each. We repeated this procedure 100 times for each simulated state. The 200 resulting networks were analyzed with ElMo-comp-1.0.4 and the EFMs obtained were compared against each other by a python script. All scripts are available upon request.

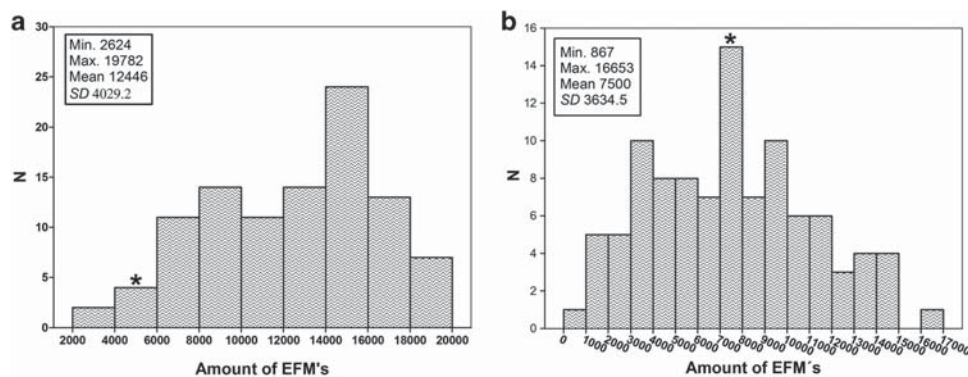
## Results and discussion

### *EFM analysis of amino-acid metabolism in free-living and symbiotic states*

The automatic metabolic reconstruction with Pathway-Tools software assigned approximately 1000 enzymatic reactions to *Chl. chlorochromatii* CaD3. From this initial reconstruction, only 382 reactions

remained after manual curation (Supplementary Material 1). As mentioned above, several enzymes related to amino acid and nitrogen metabolism change their expression profile between free-living and symbiotic states (Wenter *et al.*, 2010). To better understand the metabolic changes related to both states, we selected those reactions belonging to amino acid metabolism for further EFM analysis. In brief, EFM analysis identifies all unique and non-decomposable steady state flux distributions in a metabolic pathway, which represents the metabolic potential of the network in a specific condition (Klamt and Stelling, 2002). The amino-acid metabolic networks of the free-living and symbiotic states contained 89 and 94 reactions, respectively. These networks overlap by 84 reactions.

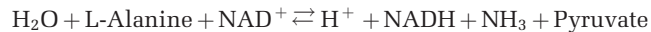
Although in EFM analysis, the amount of EFMs normally increases according to the number of reactions (Klamt and Stelling, 2002), in this study the EFM analysis of amino-acid metabolism showed that the free-living state network containing 89 reactions has more EFMs (7069 EFMs), than the network in symbiotic state with 94 reactions (5075 EFMs). This is easily explainable because five of the reactions of the symbiotic state are consecutive in two different pathways (three in the biosynthesis of isoleucine and two in the biosynthesis of leucine). In contrast, in the simulated networks, the number of EFMs increased with the number of reactions. The random networks simulating the number of enzymes in the free-living state had a mean of 7500 EFMs, which approximated the 7069 EFMs obtained for the free-living state (Figure 1). In contrast, the random sampling simulating the symbiotic state had a mean of 12 446 EFMs, approximately two standard deviations ( $\sigma=4029.2$ ) higher than the EFMs obtained in the symbiotic network (5075 EFMs). These results suggest that there is a low probability of obtaining this number of EFMs in a random configuration of 94 reactions and that the symbiotic state is clearly constraining the structure of the metabolic network of amino acids in *Chl. Chlorochromatii* CaD3.



**Figure 1** Histograms showing the distribution of EFMs in the randomly resampled networks of the symbiotic (a) and free-living (b) states of *Chl. chlorochromatii* CaD3. Asterisks indicate the bin on which the amount of EFMs obtained in symbiosis (a) and free-living (b) states would be located.



Experimental data showed that GS was present only in symbiosis (Wenter *et al.*, 2010). Accordingly, the motifs obtained with ACoM-c software showed that the EFMs were clearly different in hotspots of  $\text{NH}_3$  assimilation between symbiotic and free-living states (Figure 2). In free-living state 5805 EFMs (82%) contained a reaction performed by alanine dehydrogenase (AlaDH) EC:1.4.1.1:



Whereas in symbiosis, 3438 EFMs (68%) contained a reaction performed by GS EC:6.3.1.2:

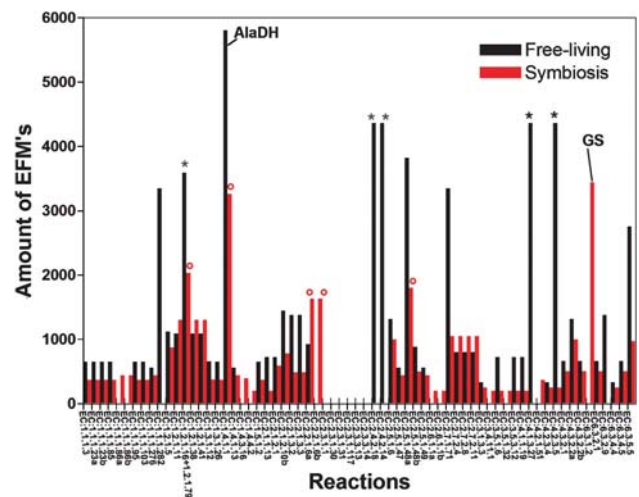


The reaction performed by AlaDH in the symbiotic state was only present in 3258 EFMs (64%). Random simulations of the free-living and symbiotic states gave very different percentages of the number of EFMs containing the reaction performed by AlaDH (Table 1).

#### Is nitrogen assimilated via AlaDH in free-living state?

Previous work has shown that  $\text{N}_2$  fixed to ammonia by the green-sulfur bacteria is assimilated via GS/GOGAT and the GDH (glutamine dehydrogenase) pathways (Brown and Herbert, 1977; Wahlund and Madigan, 1993). Both mechanisms are regulated by nitrogen availability. The GS/GOGAT pathway is a high-affinity mechanism that is active when there is low availability of nitrogen, and in the case of diazotrophic bacteria, when they are fixing nitrogen (Rudnick *et al.*, 1997; Schulz *et al.*, 2001). On the contrary, the assimilation of ammonia by GDH is a mechanism of low affinity that is active when there is high availability of nitrogen (Brown and Herbert, 1977; Rudnick *et al.*, 1997).

Unlike the other green-sulfur bacteria species, *Chl. chlorochromatii* CaD3 lacks GDH, and expresses GS only when growing in symbiotic state. How does *Chl. chlorochromatii* CaD3 assimilate nitrogen when growing in a free-living condition? One possibility is by the use of alanine dehydrogenase AlaDH (gene Cag\_1878). The alanine dehydrogenase pathway is another N assimilation mechanism. This enzyme has aminase-deaminase activity and is active in conditions of high nitrogen availability. Although this role for AlaDH has not been identified in green-sulfur bacteria, it has been observed in diazotrophic-photosynthetic bacteria that lack GDH such as *Rhodospseudomonas capsulata* (Johansson and Gest, 1976), *R. acidophila* (Herbert *et al.*, 1978), *Anabaena cylindrica* (Rowell and Stewart, 1975), and in non-photosynthetic diazotrophic bacteria such as *Bradyrhizobium sp.* (Allaway *et al.*, 2000). The increase of EFMs containing AlaDH in the metabolic network corresponding to the free-living state supports this possibility.



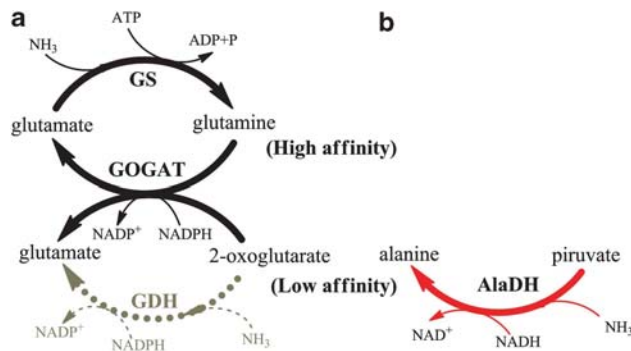
**Figure 2** Distribution of EFMs in the metabolism of amino acids of *Chl. chlorochromatii* CaD3 in free-living and symbiotic states. Reactions that contain 1-0 EFMs were not included in this plot. The reactions represented by bars with circles are involved in EFMs of the GS and by bars with stars are involved in EFMs of the AlaDH pathways.

**Table 1** Number of times the reactions performed by GS and AlaDH were found in free-living and symbiotic states

Enzyme	Symbiosis	Symbiosis (random) <sup>a</sup>	Free-living	Free-living (random) <sup>a</sup>
GS	3438 (68%)	42%	—	50%
AlaDH	3258 (64%)	90%	5805 (82%)	64%

<sup>a</sup>The percentage of the mean number of times the reaction was found in 100 random simulations of free-living and symbiotic states.

The EFMs analysis suggests that in symbiosis, ammonia is assimilated via the GS pathway (Figure 2). It is known that GS is a hotspot in nitrogen metabolism, distributing nitrogen in glutamine form to the whole metabolic network (Tyler, 1978; Reitzer, 2003). Likewise, there is a significant ratio of EFMs that goes through the AlaDH pathway. However, in free-living condition, GS is not present and AlaDH is present in a larger number of EFMs (Figure 2). These shifts in the mechanisms of nitrogen assimilation suggest that the availability of nitrogen in *Chl. chlorochromatii* CaD3 changes between symbiosis and free-living states. Wenter *et al.* (2010) and Overmann (2010) postulate that this bacterium in symbiosis is in a condition of low nitrogen availability, where the assimilation of ammonia becomes fundamental. The shift in the metabolic networks observed in our analysis supports this hypothesis. In symbiosis, the *chlorobia* seem to be in a condition of low nitrogen availability and the high-affinity mechanism is active; while in free-living state, where ammonia is available in high concentrations, the *chlorobia* use the low-affinity mechanism (Figure 3).



**Figure 3** Mechanism of nitrogen assimilation of *Chl. chlorochromatii* CaD3. The GS/GOGAT and GDH pathways (modified from the study by Yan (2007)) characteristic of the green-sulfur bacteria are shown in (a); GS/GOGAT (black solid arrow) is assimilating nitrogen in symbiosis, however a gene coding for GDH (gray dashed arrow) is not present in *Chl. Chlorochromatii* CaD3, instead in (b) the AlaDH (red solid arrow) is assimilating nitrogen in free-living state.

### Does 2-oxoglutarate mediate $N_2$ fixation in ‘*C. aggregatum*’?

Several observations indicate that highly specific regulation mechanisms exist between *Chl. chlorochromatii* CaD3 and the central  $\beta$ -proteobacterium in ‘*C. aggregatum*’ (Müller and Overmann, 2011). For instance, in ‘*Pelochromatium roseum*’, a closely related consortium, incorporation of 2-oxoglutarate by the central  $\beta$ -proteobacterium seems to be controlled by the physiological state of the *Chlorobia* (Glaeser and Overmann, 2003b). This is because incorporation of 2-oxoglutarate is strictly dependent on the simultaneous presence of light and sulfide, which is used by the epibiont as an electron donor in photosynthesis. Besides, addition of 2-oxoglutarate was necessary for the isolation of ‘*C. aggregatum*’, and this compound is pivotal to culture maintenance where the entire consortium ‘*C. aggregatum*’ seems to be unable to grow photolithoautotrophically or chemotrophically (Fröstl and Overmann, 1998). However, the nature of the molecules responsible for communication between both bacterial species is still unknown.

The PII protein, a regulatory protein of nitrogen metabolism, is exclusively expressed in symbiosis in *Chl. chlorochromatii* CaD3 (Wenter et al., 2010). All PII-signaling proteins known so far, can bind up to three molecules of ATP/ADP and 2-oxoglutarate, thereby sensing the intracellular energy and carbon status (Ninfa and Jiang, 2005; Forchhammer 2008, 2010). PII proteins are also central in regulating processes related to nitrogen assimilation (Forchhammer 2010). In *E. coli*, PII signal protein GlnB is uridylylated (GlnB-UMP) by GlnD. This modification allows the sensing of the cellular nitrogen status directly by GlnD, assessed as the glutamine concentration (Forchhammer, 2007). Also in *E. coli*, PII protein GlnB controls the expression of nitrogen-regulated genes through the histidine kinase nitrogen regulation protein B (NtrB) and the

activity of GS through GS adenylyltransferase GlnE (Forchhammer 2008). The result is that high concentration of 2-oxoglutarate stimulates ammonia assimilation through GS (Ninfa and Jiang, 2005).

In cyanobacteria and methanogenic *Archaea* that have an incomplete TCA, because of the lack of 2-oxoglutarate dehydrogenase, the concentration of 2-oxoglutarate is the signal for PII proteins to sense the carbon/nitrogen status of the cell (Forchhammer 2008; Zhang and Zhao, 2008). This is because the cellular concentration of 2-oxoglutarate depends only on the rate of its formation and its consumption for ammonia assimilation, in most cases through the GS-GOGAT cycle. By contrast, in species with a complete TCA cycle, 2-oxoglutarate is an intermediate and its cellular concentration depends only partially on nitrogen assimilation. Therefore, measuring of 2-oxoglutarate concentration by PII indicates the level of available carbon skeletons for ammonia assimilation, and glutamine is used as a primary nitrogen status signal (Forchhammer 2007, 2008). *Chl. chlorochromatii* CaD3 has a complete reverse rTCA and 2-oxoglutarate is also an intermediate, and its concentration depends partially on the GS-GOGAT cycle.

As mentioned above, PII regulates GS in bacteria by a post-transcriptional modification via adenylation and deadenylation according to nitrogen availability (Magasanik, 1982; Rudnick et al., 1997; Zhang et al., 2005). In cyanobacteria, PII senses the concentration of 2-oxoglutarate, determining its modified state (phosphorylated) in response to nitrogen availability (Forchhammer et al., 1999; Zhang and Zhao, 2008). In a similar way, in *Rhodospirillum rubrum*, another phototrophic bacterium, the GS pathway reaches its higher ratio of deadenylation when 2-oxoglutarate is present, as the PII protein and 2-oxoglutarate form a complex to allow the deadenylation of GS. Further, 2-oxoglutarate is detected by the PII protein and stimulates ammonia assimilation and  $N_2$  fixation (Forchhammer 2004; Jonsson et al., 2007). In *Azotobacter vinelandii*, 2-oxoglutarate stimulates the transcription of *nifA*, which is the master regulator of the *nif* operon, containing all genes involved in  $N_2$  fixation (Dixon and Kahn, 2004).

As reviewed above, the sensing by PII of a high concentration of 2-oxoglutarate stimulates  $N_2$  fixation and assimilation through the GS-GOGAT cycle. Could it be possible for the  $\beta$ -proteobacterium to use 2-oxoglutarate to stimulate nitrogen assimilation by the epibiont? This hypothesis is attractive because it links previous knowledge on the mechanisms of regulation of  $N_2$  fixation and assimilation, with the observation that proteins involved in nitrogen metabolism, like PII and GS, are only expressed in symbiotic state. It also links the chemotactic behavior of ‘*C. aggregatum*’ towards 2-oxoglutarate with  $N_2$  fixation and assimilation. However, this hypothesis has to overcome some observations:

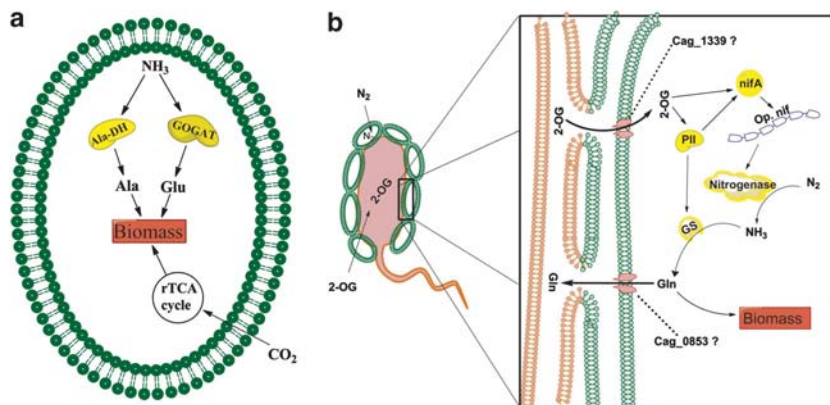
First, the membranes of phototrophic bacteria are impermeable to 2-oxoglutarate, this allows a fine

regulation of this important regulator and effector in the assimilation of ammonia and  $N_2$  fixation (Vásquez-Bermúdez *et al.*, 2000; Teixeira *et al.*, 2010). In addition, none of the known species of green sulfur bacteria utilize 2-oxoglutarate for growth (Glaeser and Overmann, 2003b). *Chl. chlorochromatii* CaD3 is not the exception; it cannot assimilate 2-oxoglutarate when free-living (Vogl *et al.*, 2006). This limitation could be overcome if there is a transporter protein for 2-oxoglutarate that only expresses in the symbiotic state. *Chl. Chlorochromatii* CaD3 codes indeed for a homolog of the 2-oxoglutarate transporter in *E. coli*. The protein coded by the gene Cag\_1339 is expressed only in the symbiotic state (Wenter *et al.*, 2010) and is homologous to five transporter proteins in *E. coli*, among these the 2-oxoglutarate transporter (gene *kgtP*). However, its best reciprocal hit in the *E. coli* genome is another predicted transporter coded by the gene *yhjE*. Therefore, it is not clear if Cag\_1339 could function as a 2-oxoglutarate transporter in *Chl. chlorochromatii* CaD3.

Secondly, 2-oxoglutarate has to travel from the  $\beta$ -proteobacterium to the epibiont. Ultrastructural characterizations of '*C. aggregatum*' suggest that there is a common periplasmic space between both bacteria (Wanner *et al.*, 2008). This would facilitate the diffusion of 2-oxoglutarate from the central bacterium to the *chlorobia*. However, recent experiments suggest that the proposed periplasmic space is not a place where molecules can freely diffuse (Müller and Overmann, 2011). Anyway, it has been shown that *Chl. chlorochromatii* CaD3 when free-living, excretes a variety of amino acids and sugars (Pfannes, 2007), and according to the proposed role of the epibiont in '*C. aggregatum*', these molecules have to reach the  $\beta$ -proteobacterium, likely through the periplasmic space.

Although *Chl. chlorochromatii* CaD3 does not assimilate 2-oxoglutarate when free-living, '*C. aggregatum*' shows chemotactic behavior towards 2-oxoglutarate (Fröstl and Overmann, 1998). Several lines of evidence suggest that the  $\beta$ -proteobacterium and not the *chlorobia*, assimilates this metabolite from the medium (Fröstl and Overmann, 1998; Overmann *et al.*, 1998; Glaeser and Overmann, 2003b; Vogl *et al.*, 2006). However, it is not known whether the central bacterium transfers organic carbon to the epibiont (Müller and Overmann, 2011). Previous studies analyzed the  $\Delta^{13}C$  ratio (that is, the difference between the  $\delta^{13}C$  values of biomarkers and the  $\delta^{13}C$  of ambient) to elucidate the carbon source used by the epibionts in '*P. roseum*' (Glaeser and Overmann, 2003a). It was found that the  $\Delta^{13}C$  ratio of the *chlorobia* in '*P. roseum*' is consistent with photoautotrophic growth (Glaeser and Overmann, 2003a). As *Chl. chlorochromatii* CaD3 has a complete rTCA, 2-oxoglutarate can be used to build pyruvate in addition to glutamine (Madigan *et al.*, 2009). Therefore, if 2-oxoglutarate enters the *chlorobia* from the  $\beta$ -proteobacterium it would end up forming biomass, contradicting the  $\Delta^{13}C$  ratio observations in '*P. roseum*', but in accordance with the fact that the entire consortium requires of 2-oxoglutarate supplementation to grow (Fröstl and Overmann, 1998).

If the theoretical difficulties described above can be overcome, we propose that the  $\beta$ -proteobacterium diffuses 2-oxoglutarate to the *Chlorobia* via their shared periplasm (Wanner *et al.*, 2008). Figure 4 summarizes what this study proposes, where in free-living state *Chl. chlorochromatii* CaD3 can assimilate ammonia via the AlaDH, where ammonia is a low-affinity substrate, and perhaps also by the GOGAT pathway because this enzyme can use ammonia instead of glutamine as a low-affinity



**Figure 4** Diagram showing the proposed mode of nitrogen assimilation in symbiosis and free-living states. (a) *Chl. chlorochromatii* CaD3 in free-living state is assimilating ammonia by alanine dehydrogenase (AlaDH) and likely by glutamate synthetase (GOGAT) pathways. (b) *Chl. chlorochromatii* CaD3 in symbiotic state where 2-oxoglutarate (2-OG) is assimilated by the  $\beta$ -proteobacterium and diffused to *Chl. chlorochromatii* CaD3 through the shared periplasm that is, periplasmic tubules (Wanner *et al.*, 2008). 2-oxoglutarate interacts with the PII and NifA proteins, stimulating the expression of the *nif* operon (Op. *nif*) for  $N_2$  fixation, and transformation of  $NH_3$  to glutamine (Gln) by GS, a portion of which is transported to the  $\beta$ -proteobacterium. Cag\_1339 is shown as the probable gene coding for a 2-oxoglutarate permease and Cag\_0853 as probable gene coding for a glutamine transporter, both only expressed in symbiosis (Wenter *et al.*, 2010).



substrate (Mäntsälä and Zalkin, 1976; Adachi and Suzuki, 1977; Matsuoka and Kimura, 1986). In symbiosis, the increase in the intracellular concentration of 2-oxoglutarate produces a regulatory shift in nitrogen assimilation, stimulating the PII and NifA proteins and increasing GS activity. *Chl. chlorochromatii* CaD3 in symbiosis has high rates of N<sub>2</sub> fixation and GS activity (Wenter et al., 2010). We postulate that this regulatory shift is induced by the 'injection' of 2-oxoglutarate by the  $\beta$ -proteobacterium, and part of the glutamine synthesized by *Chl. chlorochromatii* CaD3 is diffused to the  $\beta$ -proteobacterium through the periplasmic tubules. The above is supported by experimental evidence that shows the expression of an ABC transporter of amino acids by *Chl. chlorochromatii* CaD3 only in symbiosis coded by gene Cag\_0853, and directly influenced by the  $\beta$ -proteobacterium (Wenter et al., 2010).

Last but not the least, *Chl. chlorochromatii* CaD3 like cyanobacteria lack GlnD homologs (Arcondéguy et al., 2001). However, differing from cyanobacteria that have only GlnB, *Chl. chlorochromatii* CaD3 has three PII proteins. One of them (Cag\_1998) is coded next to an ammonium transporter, similar to GlnK in *E. coli*. The other two PII proteins are coded next to each other in an operon containing the *nifH* gene that codes for nitrogenase reductase. *Chl. chlorochromatii* CaD3 also lacks the gene *pphA* (sll1771 in *Synechocystis* sp. PCC 6803) responsible for the phosphorylation of PII in cyanobacteria (Forchhammer, 2004). Therefore, it is likely that regulation of nitrogen assimilation in the epibiont differs in the details from what has been described so far.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgements

This study was funded by projects UNAM, Facultad de Química awarded to LPMC (Project PAIP 4290-07); CONACYT-México Ciencia Básica to LD (project: CB-157220) and PASPA-UNAM and CONACYT grants for sabbatical leave (LIF). DCG was awarded a graduate fellowship from CONACYT, and acknowledges Posgrado en Ciencias Biológicas, UNAM for training and support to complete his doctoral degree.

## References

Adachi K, Suzuki I. (1977). Purification and properties of glutamate synthase from *Thiobacillus thio-parus*. *J Bacteriol* **129**: 1173–1182.

Allaway D, Ludwig EM, Crompton LA, Wood M, Parsons R, Wheeler TR et al. (2000). Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids. *Mol Microbiol* **36**: 508–515.

Arcondéguy T, Jack R, Merrick M. (2001). P(II) signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiol Mol Biol Rev* **65**: 80–105.

Brown CM, Herbert RA. (1977). Ammonia assimilation in purple and green sulphur bacteria. *FEMS Lett* **1**: 39–42.

Dixon R, Kahn D. (2004). Genetic regulation of biological nitrogen fixation. *Nat Rev Microbiol* **2**: 621–631.

Eisen JA, Nelson KE, Paulsen IT, Heidelberg JF, Wu M, Dodson RJ et al. (2002). The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc Natl Acad Sci USA* **99**: 9509–9514.

Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE et al. (2010). The Pfam protein families database. *Nucleic Acids Res* **38**: D211–D222.

Forchhammer K. (2004). Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. *FEMS Microbiol Rev* **28**: 319–333.

Forchhammer K. (2007). Glutamine signaling in bacteria. *Front Biosci* **12**: 358–370.

Forchhammer K. (2008). PII signal transducers: novel functional and structural insights. *Trends Microbiol* **16**: 65–72.

Forchhammer K. (2010). The network of P(II) signalling protein interactions in unicellular cyanobacteria. *Adv Exp Med Biol* **675**: 71–90.

Forchhammer K, Hedler A, Strobel H, Weiss V. (1999). Heterotrimerization of PII-like signalling proteins: implications for PII-mediated signal transduction systems. *Mol Microbiol* **33**: 338–349.

Frösl MJ, Overmann J. (1998). Physiology and tactic response of the phototrophic consortium 'chlorochromatium aggregatum'. *Arch Microbiol* **169**: 129–135.

Funahashi A, Jouraku A, Matsuoka Y, Kitano H. (2007). Integration of CellDesigner and SABIO-RK. *In Silico Biol* **7**: S81–S90.

Glaeser J, Overmann J. (2003a). Characterization and in situ carbon metabolism of phototrophic consortia. *Appl Environ Microbiol* **69**: 3739–3750.

Glaeser J, Overmann J. (2003b). The significance of organic carbon compounds for in situ metabolism and chemotaxis of phototrophic consortia. *Environ Microbiol* **5**: 1053–1063.

Herbert RA, Siefert E, Pfennig N. (1978). Nitrogen assimilation in *Rhodospseudomonas acidophila*. *Arch Microbiol* **119**: 1–5.

Jevremovic D, Trinh CT, Srienc F, Sosa CP, Boley D. (2011). Parallelization of Nullspace Algorithm for the computation of metabolic pathways. *Parallel Computing* **37**: 261–278.

Johansson BC, Gest H. (1976). Inorganic nitrogen assimilation by the photosynthetic bacterium *Rhodospseudomonas capsulata*. *J Bacteriol* **128**: 683–688.

Johnson LS, Eddy SR, Portugaly E. (2010). Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* **11**: 431.

Jonsson A, Teixeira PF, Nordlund S. (2007). The activity of adenylyltransferase in *Rhodospirillum rubrum* is only affected by alpha-ketoglutarate and unmodified PII proteins, but not by glutamine, in vitro. *FEBS J* **274**: 2449–2460.

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**: D109–D114.

Kanzler BE, Pfanes KR, Vogl K, Overmann J. (2005). Molecular characterization of the nonphotosynthetic partner bacterium in the consortium 'Chlorochromatium aggregatum'. *Appl Environ Microbiol* **71**: 7434–7441.



- Karp PD, Paley SM, Krummenacker M, Latendresse M, Dale JM, Lee TJ *et al.* (2010). Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Brief Bioinform* **11**: 40–79.
- Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil MP, Gama-Castro S, Muniz Rascado L *et al.* (2011). EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Res* **39**: D583–D590.
- Klamt S, Stelling J. (2002). Combinatorial complexity of pathway analysis in metabolic networks. *Mol Biol Rep* **29**: 233–236.
- Madigan HT, Martinko JM, Dunlap PV, Clark DP. (2009). *Brock. Biología de los microorganismos*, 12 ed. Pearson Educación: Madrid, España.
- Magasanik B. (1982). Genetic control of nitrogen assimilation in bacteria. *Ann Rev Genet* **16**: 135–168.
- Margulis L, Fester R. (1991). *Symbiosis as a source of evolutionary innovation: speciation and morphogenesis*. MIT Press: Cambridge, MA, USA.
- Matsuoka K, Kimura K. (1986). Glutamate synthase from *Bacillus subtilis* PCI 219. *J Biochem* **99**: 1087–1100.
- Moya A, Peretó J. (2011). *Simbiosis seres que evolucionan juntos*. Editorial síntesis S.A: España.
- Moya A, Peretó J, Gil R, Latorre A. (2008). Learning how to live together: genomic insights into prokaryote–animal symbioses. *Nat Rev Genet* **9**: 218–229.
- Mäntsälä P, Zalkin H. (1976). Glutamate synthase. Properties of the glutamine-dependent activity. *J Biol Chem* **251**: 3294–3299.
- Müller J, Overmann J. (2011). Close interspecies interactions between prokaryotes from sulfurous environments. *Front Microbiol* **2**: 146.
- Ninfa A, Jiang P. (2005). PII signal transduction proteins: sensors of  $\alpha$ -ketoglutarate that regulate nitrogen metabolism. *Curr Opin Microbiol* **8**: 168–173.
- Overmann J. (2010). The phototrophic consortium ‘Chlorochromatium aggregatum’—a model for bacterial heterologous multicellularity. *Adv Exp Med Biol* **675**: 15–29.
- Overmann J, Tuschak C, Fröstl J, Sass H. (1998). The ecological niche of the consortium ‘Pelochromatium roseum’. *Arch Microbiol* **169**: 120–128.
- Pfannes KR. (2007). *Characterization of the Symbiotic Bacterial Partners in Phototrophic Consortia* Dissertation, University of Munich, 180pp.
- Poptsova MS, Gogarten JP. (2007). BranchClust: a phylogenetic algorithm for selecting gene families. *BMC Bioinformatics* **8**: 120.
- Péres S, Vallée F, Beurton-Aimar M, Mazat JP. (2011). ACoM: a classification method for elementary flux modes based on motif finding. *Biosystems* **103**: 410–419.
- Reitzer L. (2003). Nitrogen assimilation and global regulation in *Escherichia coli*. *Annu Rev Microbiol* **57**: 155–176.
- Rowell P, Stewart WD. (1975). Alanine dehydrogenase of the N<sub>2</sub>-fixing blue-green alga, *Anabaena cylindrica*. *Arch Microbiol* **107**: 115–124.
- Rudnick P, Meletzus D, Green A, He L, Kennedy C. (1997). Regulation of nitrogen fixation by ammonium in diazotrophic species of proteobacteria. *Soil Biol Biochem* **29**: 831–841.
- Scheer M, Grote A, Chang A, Schomburg I, Munnaretto C, Rother M *et al.* (2011). BRENDA, the enzyme information system in 2011. *Nucleic Acids Res* **39**: D670–D676.
- Schulz AA, Collett HJ, Reid SJ. (2001). Nitrogen and carbon regulation of glutamine synthetase and glutamate synthase in *Corynebacterium glutamicum* ATCC 13032. *FEMS Microbiol Lett* **205**: 361–367.
- Schuster S, Dandekar T, Fell DA. (1999). Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol* **17**: 53–60.
- Schwarz R, Musch P, von Kamp A, Engels B, Schirmer H, Schuster S *et al.* (2005). YANA—a software tool for analyzing flux modes, gene-expression and enzyme activities. *BMC Bioinformatics* **6**: 135.
- Teixeira PF, Selao TT, Henriksson V, Wang H, Norén A, Nordlund S. (2010). Diazotrophic growth of *Rhodospirillum rubrum* with 2-oxoglutarate as sole carbon source affects regulation of nitrogen metabolism as well as the soluble proteome. *Res Microbiol* **161**: 651–659.
- Toft C, Andersson SG. (2010). Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat Rev Genet* **11**: 465–475.
- Tyler B. (1978). Regulation of the assimilation of nitrogen compounds. *Annu Rev Biochem* **47**: 1127–1162.
- Vogl K, Glaeser J, Pfannes KR, Wanner G, Overmann J. (2006). *Chlorobium chlorochromatii* sp. nov., a symbiotic green sulfur bacterium isolated from the phototrophic consortium ‘Chlorochromatium aggregatum’. *Arch Microbiol* **185**: 363–372.
- Vogl K, Wenter R, Dressen M, Schlickerrieder M, Plösch M, Eichacker L *et al.* (2008). Identification and analysis of four candidate symbiosis genes from ‘Chlorochromatium aggregatum’, a highly developed bacterial symbiosis. *Environ Microbiol* **10**: 2842–2856.
- Vásquez-Bermúdez MF, Herrero A, Flores E. (2000). Uptake of 2-oxoglutarate in *Synechococcus* strains transformed with the *Escherichia coli* kgtP gene. *J Bacteriol* **182**: 211–215.
- Wagner C. (2004). Nullspace approach to determine the elementary modes of chemical reaction systems. *J Phys Chem B* **108**: 2425–2431.
- Wahlund TM, Madigan MT. (1993). Nitrogen fixation by the thermophilic green sulfur bacterium *Chlorobium tepidum*. *J Bacteriol* **175**: 474–478.
- Wanner G, Vogl K, Overmann J. (2008). Ultrastructural characterization of the prokaryotic symbiosis in ‘Chlorochromatium aggregatum’. *J Bacteriol* **190**: 3721–3730.
- Wenter R, Hütz K, Dibbern D, Li T, Reisinger V, Plösch M *et al.* (2010). Expression-based identification of genetic determinants of the bacterial symbiosis ‘Chlorochromatium aggregatum’. *Environ Microbiol* **12**: 2259–2276.
- Yan D. (2007). Protection of the glutamate pool concentration in enteric bacteria. *Proc Natl Acad Sci USA* **104**: 9475–9480.
- Zhang Y, Pohlmann EL, Roberts GP. (2005). GlnD is essential for NifA activation, NtrB/NtrC-regulated gene expression, and posttranslational regulation of nitrogenase activity in the photosynthetic, nitrogen-fixing bacterium *Rhodospirillum rubrum*. *J Bacteriol* **187**: 1254–1265.
- Zhang Y, Zhao JD. (2008). PII, they key regulator of nitrogen metabolism in the cyanobacteria. *Sci China Ser C-Life Sci* **51**: 1056–1065.

Supplementary Information accompanies this paper on The ISME Journal website (<http://www.nature.com/ismej>)

### Capítulo 3

- *Cerqueda-García, D., & Falcón, L. I. (2016). Metabolic potential of microbial mats and microbialites: Autotrophic capabilities described by an in silico stoichiometric approach from shared genomic resources. Journal of Bioinformatics and Computational Biology, 14(04), 1650020.*

## Metabolic potential of microbial mats and microbialites: Autotrophic capabilities described by an *in silico* stoichiometric approach from shared genomic resources

Daniel Cerqueda-García\* and Luisa I. Falcón†

Universidad Nacional Autónoma de México  
Instituto de Ecología, Circuito Exterior  
Ciudad Universitaria, Distrito Federal 04510, Mexico

\*dacegabiol@ciencias.unam.mx

†falcon@ecologia.unam.mx

Received 24 February 2016

Revised 12 April 2016

Accepted 20 April 2016

Published 21 June 2016

Microbialites and microbial mats are complex communities with high phylogenetic diversity. These communities are mostly composed of bacteria and archaea, which are the earliest living forms on Earth and relevant to biogeochemical evolution. In this study, we identified the shared metabolic pathways for uptake of inorganic C and N in microbial mats and microbialites based on metagenomic data sets. An *in silico* analysis for autotrophic pathways was used to trace the paths of C and N to the system, following an elementary flux modes (EFM) approach, resulting in a stoichiometric model. The fragility was analyzed by the minimal cut sets method. We found four relevant pathways for the incorporation of CO<sub>2</sub> (Calvin cycle, reverse tricarboxylic acid cycle, reductive acetyl-CoA pathway, and dicarboxylate/4-hydroxybutyrate cycle), some of them present only in archaea, while nitrogen fixation was the most important source of N to the system. The metabolic potential to incorporate nitrate to biomass was also relevant. The fragility of the network was low, suggesting a high redundancy of the autotrophic pathways due to their broad metabolic diversity, and highlighting the relevance of reducing power source. This analysis suggests that microbial mats and microbialites are “metabolic pumps” for the incorporation of inorganic gases and formation of organic matter.

*Keywords:* Network analysis; elementary flux modes; fragility; minimal cut sets; eco-physiology.

### 1. Introduction

Bacteria and Archaea are fundamental components of biological communities. Fossil record shows that these organisms are the earliest living forms, evidenced by the appearance of stromatolites ~3500 Mya, which are lithified structures formed by the interaction of metabolic pathways coupled to geochemical process.<sup>1</sup> Stromatolites also represent the oldest evidence of communities and exemplify the importance of biological interactions that influence community composition and function.<sup>2,3</sup>

Modern examples forming microbial mats and microbialites<sup>4,5</sup> are self-sufficient communities, considered ecosystems at a millimetric scale due to their high phylogenetic diversity and elemental cycling capabilities. Biological diversity in microbial mats and microbialites includes phototrophs, sulfate reducing bacteria, sulfur oxidizing bacteria, methanogens, chemolithotrophs, and fermenters. The wide range of metabolic capabilities found in microbialites<sup>6,7</sup> is codified in the pan-genome, which is the total genetic resource of the organisms structuring the community (the core genome of the community).<sup>8</sup> These systems have been described as a metabolic unit forming a consortium that only requires inorganic gases and light to grow and produce available organic matter.<sup>9–11</sup> Lithified microbialites are organo-sedimentary structures formed by precipitation and trapping of silicates, carbonates, and sulfates in an extracellular polymeric substances (EPS) matrix, which is also formed by high rates of organic matter synthesis in the system, including carbohydrates and amino acids.<sup>12–16</sup>

With the emergence of “omics” tools, a plethora of sequencing data are available, and features of the phylogenetic composition and diversity of microbial mats and microbialites have been elucidated.<sup>17</sup> However, there are gaps in the functional understanding of these systems.<sup>18</sup> In this study, we analyzed to a community level, the pathways of C and N incorporation to biomass in microbialites and microbial mats from different environments, to formulate predictions of their metabolic autotrophic capabilities. We used available metagenomic data sets to make a metabolic network reconstruction of an archetypal community, which is based on stoichiometric constraints. The stoichiometric model characterizes their metabolic capabilities, capturing in mathematical form their biochemical transformations, describing the molecular balance of the biochemical reactions at a systems level.<sup>19</sup> Then we performed a metabolic pathway analysis (MPA) by an elementary flux modes (EFM) approach<sup>20</sup> which decomposes the metabolic network to minimal sets of enzymes operating in steady state with all reactions proceeding in the direction dictated by their thermodynamic restrictions.<sup>21,22</sup> Further, we analyzed the fragility of the network by minimal cut sets identification<sup>23</sup> to predict the key metabolic pathways involved in autotrophic biomass synthesis sustaining the microbial community.

## 2. Methods

### 2.1. Metagenome selection

Protein sequences and functional annotations of microbialites and microbial mat metagenomes were retrieved from the Application Programming Interface (API) of MG-RAST server.<sup>24,25</sup> The dataset included the metagenomes of microbial mats and microbialites from pristine ecosystems that had the largest amount of sequences, which to the date of this search consisted of at least  $\sim 200,000$  reads. The final dataset included lithifying microbialites from Pozas Azules and Rio Mesquites, Cuatro Ciénegas, Coahuila, Mexico (4,440,067.3 and 4,440,060.4<sup>26</sup>); Highborne Cay

Island, Bahamas (4,440,061.3<sup>27</sup>), and microbial mats from Cuatro Ciénegas, Coahuila, Mexico (red, 4,441,363.3 and green, 4,442,466.3<sup>28</sup>); the Markham Ice shelf and Ward Hunt Ice shelf in the Arctic (4,445,126.3 and 4,445,129.3,<sup>29</sup>); Diamond Fork (4,460,448.3 and 4,460,449.3<sup>30</sup>); and the Sippewissett salt marsh (4,454,153.3<sup>31</sup>).

## 2.2. Metabolic reconstruction

Each metagenomic annotation file downloaded from the MG-RAST pipeline, using a python script, was searched for annotated enzymes with enzyme code (EC) number assignments. We obtained a list of EC codes for each file, which was filtered to consider each EC number once. All files were compared to obtain the consensus of enzymes present in all metagenomes, hence, the enzymes annotated and shared in all metagenomes. With this enzymatic consensus, we retrieved the enzymatic reactions with KEGGtranslator<sup>32</sup> from the KEGG database<sup>33</sup> and YANA software<sup>34</sup> in SBML format, generating a draft metabolic reconstruction of 640 enzymatic reactions representing the shared metabolism for all metagenomes analyzed.

We focused in describing the relevance of the autotrophic metabolic pathways associated to C and N assimilation, with aims to formulate predictions of the capabilities of these communities in cycling of both constitutive elements. Therefore, from the draft reconstruction we selected the enzymatic reactions involved in the metabolic pathways that uptake CO<sub>2</sub> and N<sub>2</sub>, using the metabolic classification from KEGG PATHWAY as reference.<sup>35</sup> The autotrophic metabolic network was set by mapping the enzymes from the draft reconstruction in the pathways of energy metabolism (Ko00190, Ko00195, Ko00710, Ko00720, Ko00910, and Ko00929).

The gaps in the reconstructions were completed with experimental information from the literature supporting the incorporation of missing enzymes. We included the metabolic modules present in the KEGG MODULE database<sup>36</sup> to complete the missing steps for transformation of the metabolites and to remove dead-ends. The completion of missing steps was made if the pivotal enzymes of the pathways were present in the draft reconstruction. The curated reconstruction consisted of 104 metabolic reactions (see Supplementary Material). Biomass precursors were set to glutamate, pyruvate, 2-oxoglutarate, oxaloacetate, phosphoenolpyruvate, D-ribulose 5-phosphate, and D-erythrose 4-phosphate. Glutamate enters the GS-GOGAT cycle for amino acid formation in cells.<sup>37</sup> All other metabolites are precursors of carbohydrates that represent an organic carbon source for biomass formation and are also relevant in the EPS matrix, which supports structurally these microbial communities.<sup>38</sup> The carbon and nitrogen sources to the model were N<sub>2</sub>, NO<sub>3</sub>, and CO<sub>2</sub>.

## 2.3. EFM and minimal cut sets analysis of metabolic network

The metabolic networks were decomposed in EFM, which are the minimal set of enzymes operating in steady state with all reactions proceeding in the direction dictated by thermodynamic constraints. EFMs are equivalent to feasible metabolic pathways interconnecting the inputs and outputs of metabolites from a metabolic

system.<sup>21</sup> We performed the analysis in the CellNetAnalyzer (version 2015.1) software package using EFMtool algorithm.<sup>39,40</sup> Then we explored the fragility of the autotrophic network aiming to predict the key pathways to hold the organic matter production, by a minimal cut sets analysis to the precalculated EFMs. The minimal cut sets algorithms identify the structural failure modes, constituting the minimal sets of enzymes which if removed lead to failure of the metabolic flux. The analysis was run in MATLAB<sup>®</sup> (R2013a) in a GNU/Linux platform server with 32 cores and 64 GB of memory.

#### 2.4. *Phylogenetic identification in metagenomes*

Phylogenetic annotation to phylum and genus levels was obtained from MG-RAST. Phyla were retrieved by determination with the lowest common ancestor (LCA) method. Genera were retrieved by determination to the best hit with a cutoff of 60% similarity and  $e$ -value of  $1e^{-5}$ . Genera from phyla not included in the LCA determination were pruned. The genera that perform key metabolic pathways predicted in the model were identified and their abundance was agglomerated. The phylogenetic composition for each pathway was normalized from 0 to 1, plotted and clustered by complete-linkage with Manhattan distance in a heatmap.

### 3. Results and Discussion

#### 3.1. *Metabolic reconstruction and modeling*

There are few published metagenomic studies of microbialites ( $\sim 5$ ), and although microbial mats are better represented in metagenomic data sets, few have over 200,000 reads. Here, we included 10 metagenomic dataset compiling a total of 3,722,274 reads, three datasets representing microbialites and seven microbial mats. First we found 640 enzymes shared between all metagenomes, which were analyzed to identify the subset of autotrophic pathways. The network initially had gaps and dead-end of metabolites due to the nature of metagenomic sequencing, which produced disconnected genetic elements. It is well documented that the architecture of pathways is modular, and functional modules are conserved among species.<sup>41,42</sup> The KEGG MODULE database was built under this principle,<sup>36</sup> therefore we used the information of this database plus information from the literature to fill the gaps and curate the reconstruction. The KEGG modules used are listed in the Supplementary Material.

All communities analyzed shared the incorporation of nitrogen to biomass via  $N_2$  fixation and nitrate uptake, denitrification and anammox (see Fig. 1 and Supplementary Material). The metabolic reconstruction revealed that microbial mats and microbialites have a high diversity of metabolic pathways involved in  $CO_2$  fixation, including Calvin cycle, reverse tricarboxylic acid cycle, reductive acetyl-CoA pathway and dicarboxylate/4-hydroxybutyrate cycle. To gain insights on the phylogenetic implications of these pathways, we retrieved the phylogenetic affiliations of the

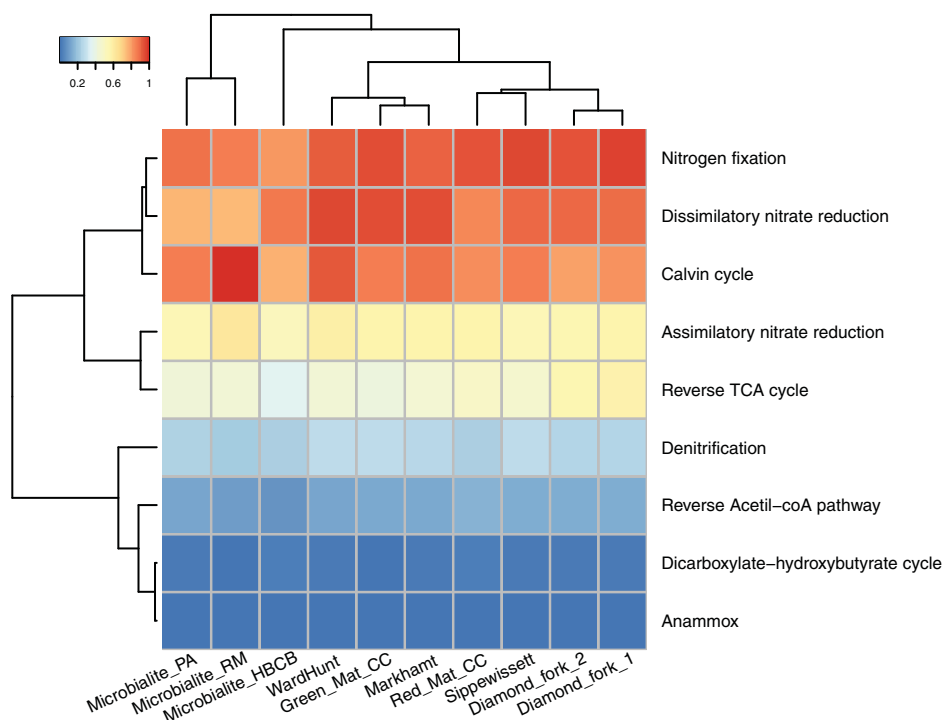


Fig. 1. Proportion of key metabolic pathways for N and C cycling at the community level predicted for microbial mats and microbialites from phylogenetic composition. The phylogenetic annotation was retrieved from MG-RAST to genus level assigned with best-hits at a cutoff of 60% of similarity and  $e$ -value of  $1e^{-5}$ . The standardized proportions of genera were agglomerated and transformed to 0–1 for each metabolic pathway. Dendrograms were clustered by complete-linkage with Manhattan distance.

metagenomic sequences from the MG-RAST server. An estimate of the abundance of the predicted pathways associated to the uptake of inorganic nitrogen and carbon was done from the abundance of the genera where these pathways are present (proportion of each genus shown in Supplementary Material). The genera performing Calvin cycle and  $N_2$  fixation were more abundant than those performing denitrification, anammox and the alternative  $CO_2$  fixing cycles (Fig. 1).

### 3.2. EFM analysis of metabolic network

The analysis of the autotrophic network obtained 3,722,274 EFMs. Although all EFMs have mathematical significance, this does not mean they all have biological significance. However, their abundance with respect to a pathway can give insights to the regulation and robustness of the networks. This can then be used in the prediction of the importance of the pathways in a metabolic process.<sup>23,43</sup> The calculated EFMs are not constrained by experimental flux data, but represent the flux with nonzero values to the system, thus the potential or possible fluxes. We grouped



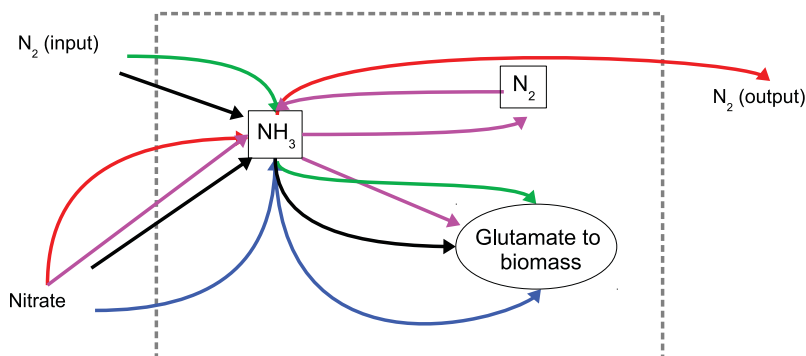


Fig. 2. Potential pathways of nitrogen utilization in microbial mats and microbialites (green arrow) 133,210 EFMs,  $N_2$  fixed by nitrogenase; (red arrow) 144,041 EFMs,  $N_2$  output, 76,561 EFMs from anammox and 37,845 EFMs from denitrification; (blue arrow) 1,389,422 EFMs, N acquired as nitrate; (black arrow) 80,057 EFMs, N acquired both as  $N_2$  and nitrate to produce biomass; (magenta arrow) 1,096,916 EFMs, N acquired as nitrate, internal production and uptake of  $N_2$  to produce biomass. From anammox: 495,121 EFMs. From denitrification: 601,795 EFMs.

the EFMs involved in the inputs of inorganic nitrogen and carbon to the system, as depicted in Figs. 2 and 3. This analysis shows the metabolic potential of all metagenomes analyzed to incorporate nitrogen via  $N_2$  fixation, and assimilation of nitrate. Most inputs of nitrogen go to biomass production. Based on the topology of the network, the EFMs predict that although there are potential pathways for  $N_2$  outgassing from denitrification and anammox (Fig. 2, red arrow), a major proportion of this  $N_2$  is re-assimilated by nitrogenase (Fig. 2, magenta arrow) and is not released from the system.

Carbon capture was attributed to four  $CO_2$  fixation pathways (Fig. 3), Calvin cycle, reverse TCA cycle, reductive acetyl CoA pathway, and dicarboxylate-hydroxybutyrate cycle. This prediction becomes relevant in light of the described

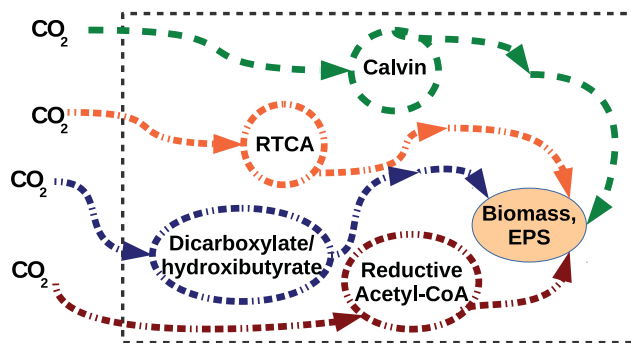


Fig. 3. Potential pathways of  $CO_2$  fixation, 780,849 EFMs from Calvin cycle (green arrows); 999,236 EFMs from reverse TCA cycle, (orange arrows); 336,028 EFMs from dicarboxylate/hydroxybutyrate cycle (blue arrows); 11,715 EFMs from reductive acetyl-CoA pathway (wine arrows).

importance of the Calvin cycle in primary productivity estimations,<sup>44–50</sup> but the other CO<sub>2</sub> fixation pathways remain understudied.<sup>51–54</sup> The EFM analysis is constrained by stoichiometry and does not violate the conservation of mass law, therefore it becomes useful to estimate the potential role of relevant metabolic pathways, including the alternative CO<sub>2</sub> fixation pathways, coupling metagenomic datasets, and ecophysiological rates to limit the solution flux space of the EFMs and make predictions for the rates of CO<sub>2</sub> fixation pathways involved.

The most studied metabolic process in microbial mats and microbialites is N<sub>2</sub> fixation.<sup>55–59</sup> Nitrogen is a limiting element in aquatic environments<sup>60,61</sup> and it is available to biological systems either through fixation (N<sub>2</sub>), and/or assimilation (NH<sub>4</sub> or NO<sub>3</sub>). In order to couple a measured rate for the pathways identified in the EFMs analysis, we used the ecophysiological measurements of N<sub>2</sub> fixation rates estimated by acetylene reduction assays of microbialites from Pozas Azules II in Cuatro Ciénegas, Mexico (published by Beltrán *et al.*<sup>59</sup>) Using the maximum rate of N<sub>2</sub> fixation (0.1 μmol N<sub>2</sub> gC<sup>-1</sup> h<sup>-1</sup>) and the availability of dissolved nitrate (1.5 μM) and phosphate (1.3 μM) in this system as inhomogeneous equalities to determine the elementary flux vectors<sup>62,63</sup> in the model. Then we obtained and estimated the rates of CO<sub>2</sub> fixation for each pathway under this constrained condition (see Fig. 4). Under these conditions, no flux towards degassing of inorganic N (either by denitrification or anammox) was predicted.

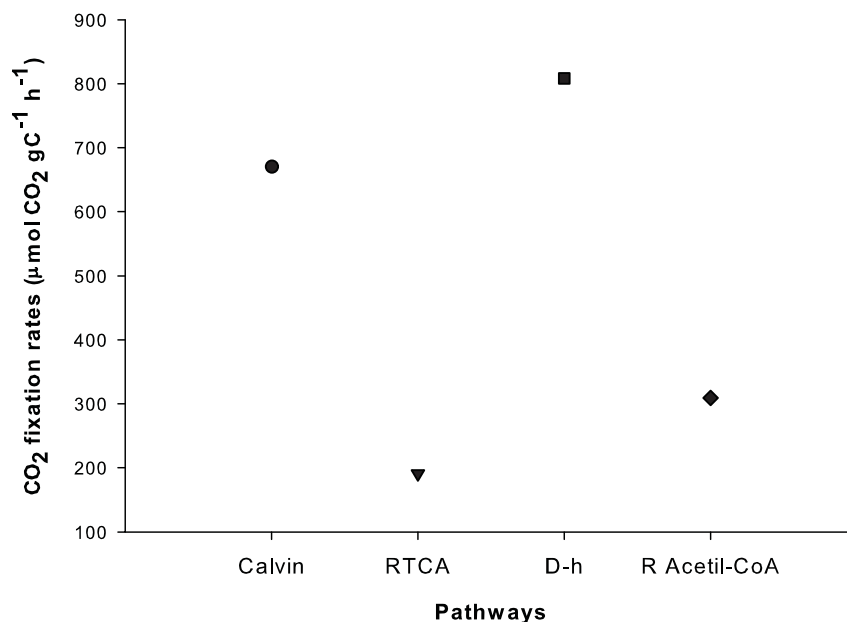


Fig. 4. Theoretical rates of CO<sub>2</sub> fixation calculated with the model, using published N<sub>2</sub> fixation rates and the availability of dissolved nitrate and phosphate in microbialites of PozasAzules II, Cuatro Ciénegas.<sup>59</sup> Reductive tricarboxylic acid cycle (RTCA); dicarboxylate–hydroxybutyrate cycle (D–h); reductive acetyl-CoA pathway (R Acetyl-CoA).

### 3.3. Minimal cut sets analysis and the fragility of the autotrophic network

A total of 6495 minimal cut sets were obtained from the autotrophic EFMs with an overall average fragility coefficient of 0.14582, with a minimal cut set size of five reactions and an average cut set size of 6.7885. The fragility coefficient takes values from 0 to 1, where 1 means that all metabolic reactions are essential.<sup>23</sup> The coefficient value depends on the size of the cut sets, since a cut set of one reaction shows that only one reaction is needed at a time to be removed (by a perturbation) to destroy EFMs and interrupt the metabolic flux. The significant cut sets should not have *pseudo* reactions, which are those that are not catalyzed by enzymes but are necessary for the model. Once the *pseudo* reactions are discarded, the enzymatic reactions involved in the cut sets and their metabolic modules are shown (Table 1).

Since the shortest cut sets of the network have a size of five reactions, to destroy the metabolic flux it would be necessary to remove a minimal combination of five reactions belonging to the metabolic modules listed in Table 1. This suggests a high redundancy in the network towards biomass and EPS production. The metabolic flux is destroyed if the metabolisms that produce reducing power are removed (originated from autotrophic pathways including photosynthesis, sulfur metabolisms, and nitrate reduction). The high diversity of metabolic pathways involved in CO<sub>2</sub> fixation make these systems highly resilient to perturbation, since only the removal of a high set of metabolisms involved in the carbon flux would be necessary to produce failure in the metabolic network. The N<sub>2</sub> fixation (nitrogenase) module seems to be an obvious cut set, considering that fixed N<sub>2</sub> is the most important source of reduced N to the system. Nonetheless, reducing power was more critical, which is biologically meaningful due to the high energetic cost of N<sub>2</sub> fixation and the extensive distribution of nitrogenase in distant phylogenetic groups.<sup>64</sup>

Table 1. Reactions involved in the cut sets with their corresponding metabolic module. All significant cut sets are listed in Supplementary Material.

Metabolic module	Reaction ID
Photosynthesis (light-dependent reactions)	R03817 R01195 R09542 R09503 R00081 R02161 R00086
Reductive pentose phosphate cycle (Calvin cycle)	R01061 R01063 R00216
Reductive citrate cycle (Arnon–Buchanan cycle)	R00405 R01196 R01082 R02164
Assimilatory sulfate reduction	R02021 R00509 R00858 R00859
Dicarboxylate–hydroxybutyrate cycle	R00238 R09279 R00342 R01975 R09281 R10782
Reductive acetyl-CoA pathway (Wood–Ljungdahl pathway)	R00943 R01220 R00134 R02289 R01655 R07168
Assimilatory nitrate reduction	R00791 R00796 R00794 R00798
Thiosulfate oxidation by SOX complex	R10151
Dissimilatory sulfate reduction	R00861

### 3.4. Validity and importance of the stoichiometric modeling to a community level

In the post genomic era the development of sequencing technologies has enabled metagenomic methods (including 16S rDNA amplicons) as fundamental sources for the analysis of phylogenetic and functional diversity at both the community and ecosystem levels.<sup>17,65-69</sup> However, to date there is a lack of sufficient direct measurements of relevant fluxes, and only a few ecophysiological data of rates of nitrogen and carbon uptake are available for microbial mat and microbialite systems (see Fig. 5). Further, in the field of systems biology, stoichiometric models have been developed which can integrate eco-physiological data to metagenomic analysis to predict metabolic rates and fluxes to a systems level.<sup>70,71</sup>

In this study we built a stoichiometric model of an archetypal microbial community, considering this as a compartment of pooled enzymes that interchange fundamental elements for biomass formation, including C and N with the environment (as described by Taff *et al.*<sup>72</sup>). This approach is at the supra-organismal level and over simplifies the system. It does not consider individual cells in the community, which is an advantage since there can be an incommensurable number of individuals, making the modeling nonviable or incomputable. Further, not all metabolic pathways in cells are important in the maintenance of the community, for example

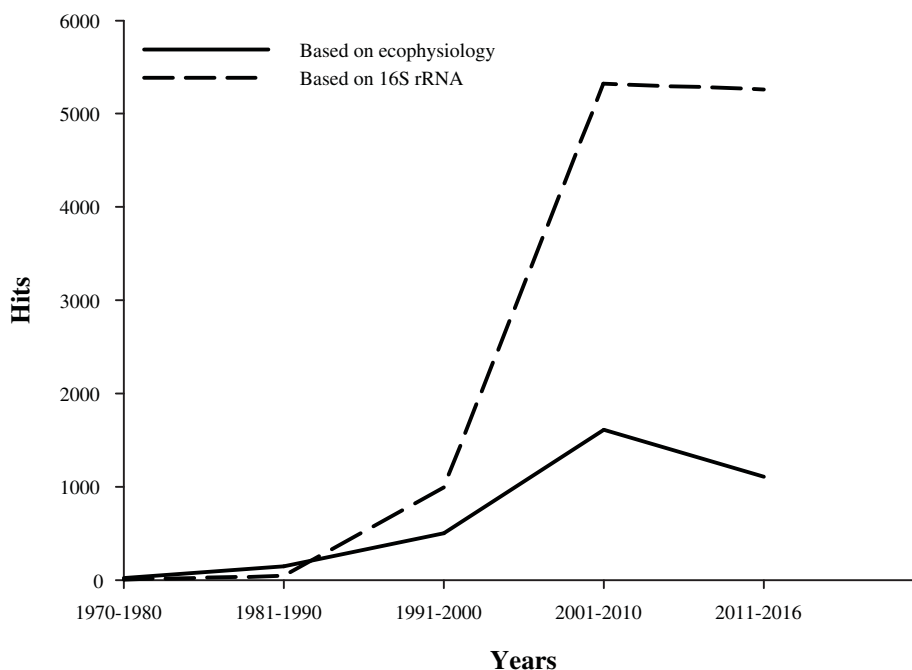


Fig. 5. Comparison of positive hits in Google Scholar since 1970 on the topic of microbial community ecology.

housekeeping metabolisms are inherent in the maintenance of the cell but are not relevant at the community level.

### 3.5. *Implication of alternative CO<sub>2</sub> fixation pathways in the microbial communities*

It is fairly well discussed that cyanobacteria have an important role in N<sub>2</sub>, CO<sub>2</sub> fixation and EPS matrix formation in microbial mats and microbialites.<sup>2,10,12,16</sup> However in this theoretical study, besides the Calvin cycle, we found three other carbon fixation cycles; reverse TCA cycle (present in bacteria); reductive acetyl-CoA cycle (present in bacteria and archaea), and the dicarboxylate–hydroxybutyrate cycle (present only in archaea).<sup>52</sup> The occurrences of these “alternative” cycles could have relevant ecological implications for microbial mats and microbialites as CO<sub>2</sub> sinks coupled to their capacity to capture carbon as carbonates. Carbon fixation pathways, including reductive acetyl-CoA pathway and rTCA cycle are amongst the most ancient and appeared during the early evolution of life.<sup>53,73–75</sup> The CO<sub>2</sub> fixation pathways associated exclusively to archaea have been very recently elucidated.<sup>76</sup> The stoichiometric model based on metagenomic analysis and eco-physiological rates here proposed, hypothesizes that the ecological role of archaea has been undermined, because diversity estimates based on 16S rRNA have biased our knowledge of genetic composition towards bacteria. An instance of this issue is shown in the study of Mobberley *et al.*<sup>77</sup> where both, bacterial and archaeal directed 16S rDNA sequencing were performed, and in the second case a considerable diversity of euryarchaeota and crenarchaeota were found, where the dicarboxylate–hydroxybutyrate and reductive acetyl-CoA cycles are present.<sup>53</sup>

## 4. Conclusions

Microbialites and microbial mats are relevant models for the study of communities, elemental cycling, and evolution. To date, next generation sequencing approaches provide a wealth of information regarding biological systems, including microbial mats and microbialites. This study identified alternative CO<sub>2</sub> fixation pathways that are relevant to biomass formation, yet had not been identified with ecophysiological methods. Further, the role of N<sub>2</sub> fixation as a fundamental metabolic pathway in these communities is confirmed. This study suggests an approach to exploit the growing metagenomic datasets with systems biology methods, and integrates ecophysiological measurements of nitrogenase activity rates to gain insights to the pathways that are relevant for elemental cycling in biological communities.

## Acknowledgments

This manuscript is part of the degree requirements of the Posgrado en Ciencias Biológicas, UNAM to D Cerqueda-García, who is recipient of a graduate studies fellowship from CONACyT, Mexico. Funding was provided by SEP-CONACyT

151796 and UNAM-PAPIIT 202016-1 (LIF). Authors appreciate the critical remarks from Dr. Roberto Olivares-Hernández who helped with the EFM analysis. Technical assistance is acknowledged to O Gaona.

## References

1. Allwood AC, Walter MR, Kamber BS, Marshall CP, Burch IW, Stromatolite reef from the Early Archaean era of Australia, *Nature* **441**:714–718, 2006.
2. Stal LJ, Physiological ecology of cyanobacteria in microbial mats and other communities, *New Phytol* **131**:1–32, 1995.
3. Konopka A, What is microbial community ecology? *Isme J* **3**(11):1223–1230, 2009.
4. Golubic S, Modern stromatolites: A review, in *Calcareous Algae and Stromatolites*, Springer, Berlin, pp. 541–561, 1991.
5. Foster JS, Green SJ, Microbial diversity in modern stromatolites, in *Stromatolites: Interaction of Microbes with Sediments*, Vol. 18, Springer, Netherlands, pp. 385–405, 2011.
6. Des Marais DJ, Biogeochemistry of hypersaline microbial mats illustrates the dynamics of modern microbial ecosystems and the early evolution of the biosphere, *Biol Bull* **204**:160–167, 2003.
7. Decker KLM, Potter CS, Bebout BM, Des Marais DJ, Carpenter S, Discipulo M, Hoehler TM, Miller SR, Thamdrup B, Turk KA, Visscher PT, Mathematical simulation of the diel O, S, and C biogeochemistry of a hypersaline microbial mat, *Fems Microbiol Ecol* **52**:377–395, 2005.
8. Fullmer MS, Soucy SM, Gogarten JP, The pan-genome as a shared genomic resource: Mutual cheating, cooperation and the Black Queen Hypothesis, *Front Microbiol* **6**:728, 2015.
9. Paerl HW, Pinckney JL, A mini-review of microbial consortia: Their roles in aquatic production and biogeochemical cycling, *Microb Ecol* **31**:225–247, 1996.
10. Paerl HW, Pinckney JL, Steppe TF, Cyanobacterial-bacterial mat consortia: Examining the functional unit of microbial survival and growth in extreme environments, *Environ Microbiol* **2**:11–26, 2000.
11. Visscher PT, Stolz JF, Microbial mats as bioreactors: Populations, processes, and products, *Palaeogeogr Palaeoclimatol Palaeoecol* **219**:87–100, 2005.
12. Kawaguchi T, Sayegh HA, Decho AW, Development of an indirect competitive enzyme-linked immunosorbent assay to detect extracellular polymeric substances (EPS) secreted by the marine stromatolite-forming Cyanobacteria, *Schizothrix* sp., *J Immunoassay Immunochem* **24**:29–39, 2003.
13. Dupraz C, Visscher PT, Baumgartner LK, Reid RP, Microbe-mineral interactions: Early carbonate precipitation in a hypersaline lake (Eleuthera Island, Bahamas), *Sedimentology* **51**:745–765, 2004.
14. Dupraz C, Visscher PT, Microbial lithification in marine stromatolites and hypersaline mats, *Trends Microbiol* **13**:429–438, 2005.
15. Braissant O, Cailleau G, Dupraz C, Verrecchia AP, Bacterially induced mineralization of calcium carbonate in terrestrial environments: The role of exopolysaccharides and amino acids, *J Sed Res* **73**:485–490, 2003.
16. Dupraz C, Reid RP, Braissant O, Decho AW, Norman RS, Visscher PT, Processes of carbonate precipitation in modern microbial mats, *Earth-Sci Rev* **96**:141–162, 2009.
17. Kowalchuk GA, Speksnijder AGCL, Zhang K, Goodman RM, van Veen JA, Finding the needles in the metagenome haystack, *Microb Ecol* **53**:475–485, 2007.

18. Chistoserdova L, Is metagenomics resolving identification of functions in microbial communities? *Microb Biotechnol* **7**:1–4, 2014.
19. Villadsen J, Nielsen J, Liden G, *Bioreaction Engineering Principles*, 3rd Edition, Springer, US, pp. 1–561, 2011.
20. Trinh CT, Wlaschin A, Sreenc F, Elementary mode analysis: A useful metabolic pathway analysis tool for characterizing cellular metabolism, *Appl Microbiol Biotechnol* **81**(5):813–826, 2009.
21. Schuster S, Dandekar T, Fell DA, Detection of elementary flux modes in biochemical networks: A promising tool for pathway analysis and metabolic engineering, *Trends Biotechnol* **17**:53–60, 1999.
22. Klamt S, Stelling J, Two approaches for metabolic pathway analysis? *Trends Biotechnol* **21**:64–69, 2003.
23. Klamt S, Generalized concept of minimal cut sets in biochemical networks, *Biosystems* **83**:233–247, 2006.
24. Meyer F, Paarmann D, D’Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA, The metagenomics RAST server — a public resource for the automatic phylogenetic and functional analysis of metagenomes, *BMC Bioinformatics* **9**(1):386, 2008.
25. Wilke A, Glass EM, Bischof J, Braithwaite D, D’Souza M, Gerlach W, Harrison T, Keegan K, Matthews H, Paczian T, MG-RAST Technical report and manual for version 3.3. 6–rev. 1, 2013.
26. Breitbart M, Hoare A, Nitti A, Siefert J, Haynes M, Dinsdale E, Edwards R, Souza V, Rohwer F, Hollander D, Metagenomic and stable isotopic analyses of modern freshwater microbialites in CuatroCiEnegas, Mexico, *Environ Microbiol* **11**:16–34, 2009.
27. Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, Moran MA, Nelson KE, Nilsson C, Olson R, Paul J, Brito B R, Ruan Y, Swan BK, Stevens R, Valentine DL, Thurber RV, Wegley L, White BA, Rohwer F, Functional metagenomic profiling of nine biomes, *Nature* **452**:629–632, 2008.
28. Peimbert M, Alcaraz LD, Bonilla-Rosso G, Olmedo-Alvarez G, Garcia-Oliva F, Segovia L, Eguiarte LE, Souza V, Comparative metagenomics of two microbial mats at CuatroCienegas basin I: Ancient lessons on how to cope with an environment under severe nutrient stress, *Astrobiology* **12**:648–658, 2012.
29. Varin T, Lovejoy C, Jungblut AD, Vincent WF, Corbeil J, Metagenomic analysis of stress genes in microbial mat communities from antarctica and the high arctic, *Appl Environ Microbiol* **78**:549–559, 2012.
30. Gomez-Alvarez V, Revetta RP, Domingo JWS, Metagenomic analyses of drinking water receiving different disinfection treatments, *Appl Environ Microbiol* **78**:6095–6102, 2012.
31. Wilbanks EG, Jaekel U, Salman V, Humphrey PT, Eisen JA, Facciotti MT, Buckley DH, Zinder SH, Druschel GK, Fike DA, Orphan VJ, Microscale sulfur cycling in the phototrophic pink berry consortia of the Sippewissett Salt Marsh, *Environ Microbiol* **16**:3398–3415, 2014.
32. Wrzodek C, Buechel F, Ruff M, Draeger A, Zell A, Precise generation of systems biology models from KEGG pathways, *BMC Syst Biol* **7**(1):1, 2013.
33. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M, Data, information, knowledge and principle: Back to metabolism in KEGG, *Nucl Acids Res* **42**:D199–D205, 2014.
34. Schwarz R, Musch P, von Kamp A, Engels B, Schirmer H, Schuster S, Dandekar T, YANA - A software tool for analyzing flux modes, gene-expression and enzyme activities, *BMC Bioinformatics* **6**(1):135, 2005.



35. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M, KEGG for integration and interpretation of large-scale molecular data sets, *Nucl Acids Res* **40**:D109–D114, 2012.
36. Muto A, Kotera M, Tokimatsu T, Nakagawa Z, Goto S, Kanehisa M, Modular architecture of metabolic pathways revealed by conserved sequences of reactions, *J Chem Inf Model* **53**:613–622, 2013.
37. Magasanik B, Genetic control of nitrogen assimilation in bacteria, *Annu Rev Genet* **16**:135–68, 1982.
38. Decho AW, Visscher PT, Reid RP, Production and cycling of natural microbial exopolymers (EPS) within a marine stromatolite, *Palaeogeogr Palaeoclimatol Palaeoecol* **219**(1):71–86, 2005.
39. Klamt S, Saez-Rodriguez J, Gilles ED, Structural and functional analysis of cellular networks with CellNetAnalyzer, *BMC Syst Biol* **1**(1):2, 2007.
40. Terzer M, Stelling J, Large-scale computation of elementary flux modes with bit pattern trees, *Bioinformatics* **24**:2229–2235, 2008.
41. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL, The large-scale organization of metabolic networks, *Nature* **407**:651–654, 2000.
42. Wagner A, Fell DA, The small world inside large metabolic networks, *Proc R Soc B Biol Sci* **268**:1803–1810, 2001.
43. Wilhelm T, Behre J, Schuster S, Analysis of structural robustness of metabolic networks, *Syst Biol* **1**(1):14–120, 2004.
44. Schidlowski M, Aharon P, Carbon cycle and carbon isotope record: Geochemical impact of life over 3.8 Ga of Earth history, in *Early Organic Evolution*, SpringerBerlin Heidelberg, pp. 147–175, 1992.
45. Knoll AH, The geological consequences of evolution. *Geobiology* **1**(1):3–14, 2003.
46. Tabita FR, Hanson TE, Li H, Satagopan S, Singh J, Chan S, Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs, *Microbiol Mol Biol Rev* **71**(4):576–599, 2007.
47. Falkowski P, Scholes RJ, Boyle EEA, Canadell J, Canfield D, Elser J *et al.*, The global carbon cycle: A test of our knowledge of earth as a system, *Science* **290**(5490):291–296, 2000.
48. Falkowski PG, Fenchel T, Delong EF, The microbial engines that drive Earth's biogeochemical cycles, *Science* **320**(5879):1034–1039, 2008.
49. Casamayor EO, García-Cantizano J, Pedrós-Alió C, Carbon dioxide fixation in the dark by photosynthetic bacteria in sulfide-rich stratified lakes with oxic-anoxic interfaces, *Limnol Oceanogr* **53**(4):1193–1203, 2008.
50. Raven JA, MacIntyre H, Berman T, Frank IB, Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments, *Aquat Microb Ecol* **56**(2–3):177–192, 2009.
51. Callieri C, Corno G, Caravati E, Rasconi S, Contesini M, Bertoni R, Bacteria, Archaea, and Crenarchaeota in the epilimnion and hypolimnion of a deep holo-oligomictic lake, *Appl Environ Microbiol* **75**(22):7298–7300, 2009.
52. Huegler M, Sievert SM, Beyond the Calvin cycle: Autotrophic carbon fixation in the ocean, *Annu Rev Mar Sci* **3**:261–289, 2011.
53. Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Huegler M, Alber BE, Fuchs G, Autotrophic carbon fixation in archaea, *Nat Rev Microbiol* **8**:447–460, 2010.
54. Bar-Even A, Noor E, Milo R, A survey of carbon fixation pathways through a quantitative lens, *J Exp Bot* **63**(6):2325–2342, 2012.
55. Stal LJ, Cyanobacterial mats and stromatolites, in *Ecology of Cyanobacteria II*, Springer, Netherlands, pp. 65–125, 2012.



56. Omoregie EO, Crumbliss LL, Bebout BM, Zehr JP, Determination of nitrogen-fixing phylotypes in *Lyngbya* sp. and *Microcoleus chthonoplastes* cyanobacterial mats from Guerrero Negro, Baja California, Mexico, *Appl Environ Microbiol* **70**(4):2119–2128, 2004.
57. Falcón LI, Carpenter EJ, Cipriano F, Bergman B, Capone DG, N<sub>2</sub> fixation by unicellular bacterioplankton from the Atlantic and Pacific Oceans: Phylogeny and in situ rates, *Appl Environ Microbiol* **70**(2):765–770, 2004.
58. Lesser MP, Falcón LI, Rodríguez-Román A, Enríquez S, Hoegh-Guldberg O, Iglesias-Prieto R, Nitrogen fixation by symbiotic cyanobacteria provides a source of nitrogen for the scleractinian coral *Montastraea cavernosa*, *Mar Ecol Prog Ser* **346**:143–152, 2007.
59. Beltrán Y, Centeno CM, García-Oliva F, Legendre P, Falcón LI, N<sub>2</sub> fixation rates and associated diversity (nifH) of micro-bialite and mat-forming consortia from different aquatic environments in Mexico, *Aquat Microb Ecol* **67**(1):15, 2012.
60. Falkowski PG, Evolution of the nitrogen cycle and its influence on the biological sequestration of CO<sub>2</sub> in the ocean, *Nature* **387**:272–275, 1997.
61. Canfield DE, Glazer AN, Falkowski PG, The evolution and future of Earth's nitrogen cycle, *Science* **330**:192–196, 2010.
62. Urbanczik R, Enumerating constrained elementary flux vectors of metabolic networks, *IET Syst Biol* **1**(5):274–279, 2007.
63. Klamt S, Mahadevan R, On the feasibility of growth-coupled product synthesis in microbial strains, *Metab Eng* **30**:166–178, 2015.
64. Dos Santos PC, Fang Z, Mason SW, Setubal JC, Dixon R, Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes, *BMC Genom* **13**:162, 2012.
65. Leveau JH, The magic and menace of metagenomics: Prospects for the study of plant growth-promoting rhizobacteria, *Eur J Plant Pathol* **119**(3):279–300, 2007.
66. Chistoserdova L, Recent progress and new challenges in metagenomics for biotechnology, *Biotechnol Lett* **32**(10):1351–1359, 2010.
67. Morales SE, Holben WE, Linking bacterial identities and ecosystem processes: Can 'omic'analyses be more than the sum of their parts? *FEMS Microbiol Ecol* **75**(1):2–16, 2011.
68. Prakash T, Taylor TD, Functional assignment of metagenomic data: Challenges and applications, *Brief Bioinformatics* **13**(6):711–727, 2012.
69. Knight R, Jansson J, Field D, Fierer N, Desai N, Fuhrman JA, Unlocking the potential of metagenomics through replicated experimental design, *Nat Biotechnol* **30**(6):513–520, 2012.
70. Zengler K, Palsson BO, A road map for the development of community systems (CoSy) biology, *Nat Rev Microbiol* **10**(5):366–372, 2012.
71. Song HS, Cannon WR, Beliaev AS, Konopka A, Mathematical modeling of microbial community dynamics: A methodological review, *Processes* **2**:711–752, 2014.
72. Taffs R, Aston JE, Brileya K, Jay Z, Klatt CG, McGlynn S, Mallette N, Montross S, Gerlach R, Inskeep WP, Ward DM, Carlson RP, In silico approaches to study mass and energy flows in microbial consortia: A syntrophic case study, *BMC Syst Biol* **3**(1):114, 2009.
73. Pereto JG, Velasco AM, Becerra A, Lazcano A, Comparative biochemistry of CO<sub>2</sub> fixation and the evolution of autotrophy, *Int Microbiol* **2**(1):3–10, 1999.
74. Russell MJ, Martin W, The rocky roots of the acetyl-CoA pathway, *Trends Biochem Sci* **29**:358–363, 2004.
75. Braakman R, Smith E, The emergence and early evolution of biological carbon-fixation, *PLoS Comput Biol* **8**(4):e1002455, 2012.

76. Huber H, Gallenberger M, Jahn U, Eylert E, Berg IA, Kockelkorn D, Eisenreich W, Fuchs G, A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum *Ignicoccus hospitalis*, *Proc Natl Acad Sci USA* **105**:7851–7856, 2008.
77. Mobberley JM, Ortega MC, Foster JS, Comparative microbial diversity analyses of modern marine thrombolitic mats by barcoded pyrosequencing, *Environ Microbiol* **14**:82–100, 2012.

**Daniel Cerqueda-García** received his MS degree from UNAM, Mexico in 2014. He is a PhD student in UNAM's Graduate Program in Biological Sciences.

**Luisa I Falcón** received her PhD degree from Stony Brook University, USA, in 2003. She is the P.I. of the Bacterial Ecology Laboratory at the Institute of Ecology, UNAM, Mexico.

## Discusión General

Los análisis metabólicos han surgido de un enfoque reduccionista, en donde se estudian proteínas, células y cultivos mono-específicos o monoclonales, por medio de la experimentación y modelación matemática. Si bien, estos estudios como la purificación y cristalización de proteínas son esenciales para obtener conocimiento sobre la función de las mismas, pueden no ser adecuados en cuanto a la predicción o descripción de su funcionamiento a nivel de sistema. Esto debido a que los componentes del metabolismo no son partes aisladas y tienen interacciones con otros componentes. Por ejemplo, las interacciones microbio-microbio y microbio-ambiente, que producen propiedades emergentes no aditivas y que moldean las dinámicas y evolución de sus ecosistemas.

La integración de los datos ómicos en modelos predictivos, es de gran importancia en la ecología microbiana. En este estudio, integramos información genómica y metagenómica para predecir el intercambio de materia y energía entre las especies formadoras de consorcios. Primero a partir de información genómica de *Chlorobium chlorochromatii*, en donde hicimos una predicción del intercambio de nitrógeno y carbono en el consorcio que forma este epibionte y la bacteria central '*Ca. S. mobilis*'. Esta simbiosis se presenta debido a un cambio en la síntesis de aminoácidos entre su modo de vida libre y en simbiosis, en donde la asimilación de nitrógeno por el sistema GS-GOGAT, presente en vida libre, cambia a ser asimilado por la enzima alanina deshidrogenasa (alaDH) en simbiosis. Aunque esta vía de asimilación de nitrógeno no había sido descrita para las bacterias verdes del azufre, fue confirmada posteriormente en *Chlorobium tepidum* (Levy *et al.*, 2016), una especie cercana a *C. chlorochromatii*.

Al término del análisis de la red metabólica de *C. chlorochromatii* en la presente tesis, Liu y col. (2013) publicaron el genoma completo de '*Ca. S. mobilis*', el cual tiene un tamaño reducido (~2.9 Mbp) en comparación a las especies de la familia commamonadaceae (4.8-6.8 Mbp), por lo que ha pasado por un proceso de reducción de su genoma, perdiendo genes involucrados en el metabolismo, su genoma evidencia el hecho de que no pueda vivir sin su epibionte y dependa de él para la obtención de metabolitos esenciales, ya que carece de rutas para la fijación de CO<sub>2</sub> y requiere de una fuente exógena de carbono. Además, no puede fijar N<sub>2</sub> ni asimilar nitrato o nitrito, pero tiene permeasas para amonio y transportadores de aminoácidos y azúcares, por lo que a partir del genoma, Lui y col también discuten que la fuente de nitrógeno puede estar siendo suplementada por *C. chlorochromatii*, lo que concuerda

con los resultados de el presente estudio.

*C. aggregatum* es un ejemplo de integración de especies en la naturaleza, es un hecho que ningún organismo vive aislado y las condiciones de aislamiento solo se consiguen mediante cuidados rigurosos en un laboratorio (Poole et al., 2015). Esta asociación que ahora es obligada para '*Ca. S. mobilis*', puede no ser un caso excepcional y por el contrario puede ser una regla para la vida, ya que todo organismo necesita suplementación externa para sobrevivir, incluso organismos fotótrofos y diazótrofos como las cianobacterias, sin una fuente de recambio de CO<sub>2</sub>, lo transformarían en carbohidratos hasta depletarlo y no podrían seguir sintetizando sus nutrientes (Los organismos fotosintéticos son organismos autótrofos obligados y no pueden asimilar carbono orgánico) (Vázquez-Bermúdez et al., 2000), por ello dependen de la actividad heterótrofa respiratoria. Entonces, los microorganismos que pueden ser relativamente simples, con sus interacciones surgen propiedades emergentes que los benefician.

Para el estudio a nivel comunidad, integramos secuencias metagenómicas en una reconstrucción metabólica de los metabolismos autótrofos presentes en comunidades de tapetes microbianos y microbialitos. Mediante el análisis de EFMs, hicimos la predicción del intercambio de nitrógeno y carbono entre una comunidad modelo y el ambiente. La predicción muestra que los metabolismos se conservan aunque cambien las especies. Además, existe una redundancia en las rutas de fijación de carbono, prediciendo la importancia de rutas de fijación que no han sido descritas profundamente en el funcionamiento de las comunidades.

Este estudio es un ejemplo de un enfoque ecosistémico del metabolismo, enfoque que puede ayudar a dar esa integración o salto conceptual en el estudio del metabolismo a nivel comunitario, para describir el acoplamiento de la comunidad a nivel biogeoquímico y poder predecir su comportamiento, más allá de la descripción de su composición filogenética. Modelando el metabolismo comunitario también podemos catalogar su potencial genético total (su pangenoma) (Fullmer et al., 2015), para hacer una predicción de las capacidades autotróficas y heterotróficas de las comunidades, es decir, enfrentar cuestiones cómo: si una comunidad está compuesta por fijadores de carbono pero también de heterótrofos, en su totalidad ¿esta comunidad está produciendo carbono inorgánico o lo está consumiendo? Estas cuestiones se han vuelto muy importantes a nivel mundial debido a los esfuerzos para la mitigación de gases de efecto invernadero en el calentamiento global.

El tipo de abstracción de los componentes de la comunidad en la construcción de un modelo, como la elección de las “unidades de interacción” (células, especies o la comunidad total) depende del objetivo del modelo y del tipo de sistema estudiado. En nuestro caso, tratamos a la comunidad de los tapetes microbianos y microbialitos como una unidad metabólica, con el objetivo de predecir su capacidad de usar el carbono y nitrógeno inorgánicos, así como las rutas metabólicas involucradas. Se sabe que estos sistemas son sésiles y sus límites espaciales son claramente diferenciables, funcionando como unidades semi-independientes de sus entornos, utilizando luz y materia inorgánica para funcionar, asemejándose a un bioreactor. Con este enfoque se simplifica la compartimentalización existente entre las células de la comunidad, sin embargo, también representa una ventaja técnica al permitir el cómputo de los modos elementales, al reducirse la complejidad de la red metabólica requerida. En una comunidad como la encontrada en una columna de agua quizás su delimitación espacial no sería tan clara o sus límites espaciales serían mucho más grandes.

## **Conclusiones**

Con la metagenómica se ha hecho posible la disposición de una gran cantidad de información correspondiente a las secuencias genómicas encontradas en muestras ambientales. Toda esta información puede ser aprovechada mediante el uso de modelos estequiométricos para caracterizar el flujo de materia y energía dentro de las comunidades. En el análisis de las comunidades se trata con un mayor número de componentes (a diferencia de los estudios del metabolismo de una sola especie), por lo que en este sentido aumenta la complejidad de los modelos y la mayor necesidad de poder de cómputo. Sin embargo, los modelos estequiométricos pueden construirse por módulos y con ello, predecir procesos metabólicos de importancia en el mantenimiento de las comunidades.

## Referencias

1. Allison, S. D. (2005). Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments. *Ecology Letters*, 8(6), 626-635.
2. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389-3402.
3. Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., ... & Martin, M. J. (2004). UniProt: the universal protein knowledgebase. *Nucleic acids research*, 32(suppl 1), D115-D119.
4. Azam, F., & Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology*, 5(10), 782-791.
5. Breitbart, M., Hoare, A., Nitti, A., Siefert, J., Haynes, M., Dinsdale, E., ... & Hollander, D. (2009). Metagenomic and stable isotopic analyses of modern freshwater microbialites in Cuatro Ciénegas, Mexico. *Environmental microbiology*, 11(1), 16-34.
6. Brown CM. & Herbert RA. (1977) Ammonia assimilation in purple and green sulphur bacteria. *FEMS Letters* 1:39-42.
7. Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature methods*, 12(1), 59-60.
8. Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, 7(5), 335-336.
9. Centeno, C. M., Legendre, P., Beltrán, Y., Alcántara-Hernández, R. J., Lidström, U. E., Ashby, M. N., & Falcón, L. I. (2012). Microbialite genetic diversity and composition relate to environmental variables. *FEMS microbiology ecology*, 82(3), 724-735.
10. Chistoserdova, L. (2014). Is metagenomics resolving identification of functions in microbial communities?. *Microbial biotechnology*, 7(1), 1-4.
11. Cooper, J. E. (2007). Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. *Journal of Applied Microbiology*, 103(5), 1355-1365.
12. Decho, A. W., Visscher, P. T., & Reid, R. P. (2005). Production and cycling of natural microbial exopolymers (EPS) within a marine stromatolite. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 219(1), 71-86.

13. Des Marais, D. J. (2003). Biogeochemistry of hypersaline microbial mats illustrates the dynamics of modern microbial ecosystems and the early evolution of the biosphere. *The Biological Bulletin*, 204(2), 160-167.
14. Dinsdale, E. A., Edwards, R. A., Hall, D., Angly, F., Breitbart, M., Brulc, J. M., ... & McDaniel, L. (2008). Functional metagenomic profiling of nine biomes. *Nature*, 452(7187), 629.
15. Dolfing, J. (2001). The microbial logic behind the prevalence of incomplete oxidation of organic compounds by acetogenic bacteria in methanogenic environments. *Microbial ecology*, 41(2), 83-89.
16. Dupraz, C., Reid, R. P., Braissant, O., Decho, A. W., Norman, R. S., & Visscher, P. T. (2009). Processes of carbonate precipitation in modern microbial mats. *Earth-Science Reviews*, 96(3), 141-162.
17. Dupraz, C., Visscher, P. T., Baumgartner, L. K., & Reid, R. P. (2004). Microbe–mineral interactions: early carbonate precipitation in a hypersaline lake (Eleuthera Island, Bahamas). *Sedimentology*, 51(4), 745-765.
18. Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS computational biology*, 7(10), e1002195.
19. Falkowski, P. G., & Godfrey, L. V. (2008). Electrons, life and the evolution of Earth's oxygen cycle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1504), 2705-2716.
20. Falkowski, P. G., Fenchel, T., & Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science*, 320(5879), 1034-1039.
21. Falkowski, P. G., Fenchel, T., & Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science*, 320(5879), 1034-1039.
22. Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., & Palsson, B. Ø. (2009). Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, 7(2), 129-143.
23. Frith, M. C., Wan, R., & Horton, P. (2010). Incorporating sequence quality data into alignment improves DNA read mapping. *Nucleic acids research*, 38(7), e100-e100.
24. Fröstl MJ. & Overmann J. (1998) Physiology and tactic response of the phototrophic consortium “*chlorochromatium aggregatum*” *Arch. microbiol* (1998) 169:129–135.
25. FU, Limin, et al. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 2012, vol. 28, no 23, p. 3150-3152.
26. Fullmer, M. S., Soucy, S. M., & Gogarten, J. P. (2015). The pan-genome as a shared genomic



- resource: mutual cheating, cooperation and the black queen hypothesis. *Frontiers in microbiology*, 6.
27. Gillor, O., Etzion, A., & Riley, M. A. (2008). The dual role of bacteriocins as anti-and probiotics. *Applied microbiology and biotechnology*, 81(4), 591-606.
  28. Glaeser J. & Overmann J. (2003) Characterization and in situ carbon metabolism of phototrophic consortia. *Appl. Environ. Microbiol.* 69(7):3739-50.
  29. Gombert, A. K., & Nielsen, J. (2000). Mathematical modelling of metabolism. *Current Opinion in Biotechnology*, 11(2), 180-186.
  30. Gombert, A. K., & Nielsen, J. (2000). Mathematical modelling of metabolism. *Current Opinion in Biotechnology*, 11(2), 180-186.
  31. Gomez-Alvarez, V., Revetta, R. P., & Santo Domingo, J. W. (2012). Metagenomic analyses of drinking water receiving different disinfection treatments. *Applied and environmental microbiology*, 78(17), 6095-6102.
  32. González, J. E., & Keshavan, N. D. (2006). Messing with bacterial quorum sensing. *Microbiology and Molecular Biology Reviews*, 70(4), 859-875.
  33. Haft, D. H., Selengut, J. D., & White, O. (2003). The TIGRFAMs database of protein families. *Nucleic acids research*, 31(1), 371-373.
  34. Hoff, K. J., Tech, M., Lingner, T., Daniel, R., Morgenstern, B., & Meinicke, P. (2008). Gene prediction in metagenomic fragments: a large scale machine learning approach. *BMC bioinformatics*, 9(1), 217.
  35. Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, 11(1), 119.
  36. Jevremović, D., Trinh, C. T., Srienc, F., Sosa, C. P., & Boley, D. (2011). Parallelization of nullspace algorithm for the computation of metabolic pathways. *Parallel computing*, 37(6), 261-278.
  37. Juhas, M., Der Meer, V., Roelof, J., Gaillard, M., Harding, R. M., Hood, D. W., & Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS microbiology reviews*, 33(2), 376-393.
  38. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic acids research*, 40(D1), D109-D114.

39. Kanzler BE., Pfanes KR., Vogl K. & Overmann J. (2005) Molecular characterization of the nonphotosynthetic partner bacterium in the consortium "*Chlorochromatium aggregatum*" Appl. Environ. Microbiol. 71(11):7434-41.
40. Kawaguchi, T., & Decho, A. W. (2002). Isolation and biochemical characterization of extracellular polymeric secretions (EPS) from modern soft marine stromatolites (Bahamas) and its inhibitory effect on CaCO<sub>3</sub> precipitation. Preparative Biochemistry and Biotechnology, 32(1), 51-63.
41. Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. Genome research, 12(4), 656-664.
42. Klamt, S. (2006). Generalized concept of minimal cut sets in biochemical networks. Biosystems, 83(2), 233-247.
43. Klamt, S., & Gilles, E. D. (2004). Minimal cut sets in biochemical reaction networks. Bioinformatics, 20(2), 226-234.
44. Klamt, S., Saez-Rodriguez, J., & Gilles, E. D. (2007). Structural and functional analysis of cellular networks with CellNetAnalyzer. BMC systems biology, 1(1), 2.
45. Klitgord, N., & Segre, D. (2011). Ecosystems biology of microbial metabolism. Current opinion in biotechnology, 22(4), 541-546.
46. Konopka, A. (2009). What is microbial community ecology. The ISME journal, 3(11), 1223-1230.
47. Latendresse, M., Krummenacker, M., Trupp, M., & Karp, P. D. (2012). Construction and completion of flux balance models from pathway databases. Bioinformatics, 28(3), 388-396.
48. Levy, A. T., Lee, K. H., & Hanson, T. E. (2016). *Chlorobaculum tepidum* modulates amino acid composition in response to energy availability, as revealed by a systematic exploration of the energy landscape of phototrophic sulfur oxidation. Applied and Environmental Microbiology, 82(21), 6431-6439.
49. Li, W. (2009). Analysis and comparison of very large metagenomes with fast clustering and functional annotation. BMC bioinformatics, 10(1), 359.
50. Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., ... & Law, M. (2012). Comparison of next-generation sequencing systems. BioMed Research International, 2012.
51. Liu, Z., Müller, J., Li, T., Alvey, R. M., Vogl, K., Frigaard, N. U., ... & Henke, P. (2013). Genomic analysis reveals key aspects of prokaryotic symbiosis in the phototrophic consortium "*Chlorochromatium aggregatum*". Genome biology, 14(11), R127.
52. Magasanik B. (1977) Regulation of bacterial nitrogen assimilation by glutamine synthetase.

- Trends Biochem. Sci. 2(1):9-12.
53. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., ... & Edwards, R. A. (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9(1), 386.
  54. Miethke, M., &Marahiel, M. A. (2007). Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Reviews*, 71(3), 413-451.
  55. Mitchell, A., Bucchini, F., Cochrane, G., Denise, H., ten Hoopen, P., Fraser, M., ... & Finn, R. D. (2015). EBI metagenomics in 2016—an expanding and evolving resource for the analysis and archiving of metagenomic data. *Nucleic acids research*, gkv1195.
  56. Moran, N. A., McCutcheon, J. P., &Nakabachi, A. (2008). Genomics and evolution of heritable bacterial symbionts. *Annual review of genetics*, 42, 165-190.
  57. Orth, J. D., Thiele, I., &Palsson, B. Ø. (2010). What is flux balance analysis?.*Nature biotechnology*, 28(3), 245-248.
  58. Overmann J. (2010) The phototrophic consortium "*Chlorochromatium aggregatum*" - a model for bacterial heterologous multicellularity. *Adv Exp Med Biol*. 675:15-29.
  59. Overmann, J., &Gemerden, H. (2000). Microbial interactions involving sulfur bacteria: implications for the ecology and evolution of bacterial communities. *FEMS microbiology reviews*, 24(5), 591-599.
  60. Padmanabhan, R., Mishra, A. K., Raoult, D., & Fournier, P. E. (2013). Genomics and metagenomics in medical microbiology. *Journal of microbiological methods*, 95(3), 415-424.
  61. Palsson, B. (2004). Two-dimensional annotation of genomes. *Nature biotechnology*, 22(10), 1218-1219.
  62. Palsson, B. Ø. (2006). *Systems biology: properties of reconstructed networks*. Chicago
  63. Papin, J. A., Stelling, J., Price, N. D., Klamt, S., Schuster, S., & Palsson, B. O. (2004). Comparison of network-based pathway analysis methods. *Trends in biotechnology*, 22(8), 400-405.
  64. Peimbert, M., Alcaraz, L. D., Bonilla-Rosso, G., Olmedo-Alvarez, G., García-Oliva, F., Segovia, L., ... & Souza, V. (2012). Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin I: ancient lessons on how to cope with an environment under severe nutrient stress. *Astrobiology*, 12(7), 648-658.
  65. Pfeiffer, T., Nu, J. C., Montero, F., & Schuster, S. (1999). METATOOL: for studying metabolic networks. *Bioinformatics*, 15(3), 251-257.

66. Poole, A. M., Lundin, D., & Rytönen, K. T. (2015). The evolution of early cellular systems viewed through the lens of biological interactions. *Frontiers in microbiology*, 6.
67. Prakash, T., & Taylor, T. D. (2012). Functional assignment of metagenomic data: challenges and applications. *Briefings in bioinformatics*, 13(6), 711-727.
68. Punta, M., Coghill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., ... & Finn, R. D. (2012). The Pfam protein families database. *Nucleic acids research*, 40(D1), D290-D301.
69. Reid, R. P., Visscher, P. T., Decho, A. W., Stolz, J. F., Bebout, B. M., Dupraz, C., ... & DesMarais, D. J. (2000). The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. *Nature*, 406(6799), 989-992.
70. Rho, M., Tang, H., & Ye, Y. (2010). FragGeneScan: predicting genes in short and error-prone reads. *Nucleic acids research*, 38(20), e191-e191.
71. Riley, M. A., & Wertz, J. E. (2002). Bacteriocins: evolution, ecology, and application. *Annual Reviews in Microbiology*, 56(1), 117-137.
72. Rudnick P., Meletzus D., Green A., He L. & Kennedy C. (1997) Regulation of nitrogen fixation by ammonium in diazotrophic species of proteobacteria. *Soil Biol. Biochem.* 29(5-6):831-841.
73. Ryan, R. P., & Dow, J. M. (2008). Diffusible signals and interspecies communication in bacteria. *Microbiology*, 154(7), 1845-1858.
74. Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., ... & Palsson, B. Ø. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature protocols*, 6(9), 1290-1307.
75. Schuster, S., von Kamp, A., & Pachkov, M. (2007). Understanding the roadmap of metabolism by pathway analysis. In *Metabolomics* (pp. 199-226). Humana Press.
76. Sowell, S. M., Wilhelm, L. J., Norbeck, A. D., Lipton, M. S., Nicora, C. D., Barofsky, D. F., ... & Giovanonni, S. J. (2009). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *The ISME journal*, 3(1), 93-105.
77. Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., & Gilles, E. D. (2002). Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420(6912), 190-193.
78. Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nature Reviews Microbiology*, 5(10), 801-812.
79. Suzuki, S., Kakuta, M., Ishida, T., & Akiyama, Y. (2014). GHOSTX: an improved sequence homology search algorithm using a query suffix array and a database suffix array. *PloS one*, 9(8), e103833.

80. Taffs, R., Aston, J. E., Brileya, K., Jay, Z., Klatt, C. G., McGlynn, S., ... & Carlson, R. P. (2009). In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC systems biology*, 3(1), 114.
81. Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., ... & Natale, D. A. (2003). The COG database: an updated version includes eukaryotes. *BMC bioinformatics*, 4(1), 41.
82. Terzer, M., & Stelling, J. (2008). Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics*, 24(19), 2229-2235.
83. Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics-a guide from sampling to data analysis. *Microb Inform Exp*, 2(3).
84. Trinh, C. T., Wlaschin, A., & Sreenc, F. (2009). Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Applied microbiology and biotechnology*, 81(5), 813-826.
85. Varin, T., Lovejoy, C., Jungblut, A. D., Vincent, W. F., & Corbeil, J. (2012). Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. *Applied and environmental microbiology*, 78(2), 549-559.
86. Vieites, J. M., Guazzaroni, M. E., Beloqui, A., Golyshin, P. N., & Ferrer, M. (2010). Molecular methods to study complex microbial communities. In *Metagenomics* (pp. 1-37). Humana Press.
87. Villadsen, J., Nielsen, J., & Lidén, G. (2011). *Bioreaction engineering principles*. Springer.
88. Villarreal, D. T., Turco, R. F., & Konopka, A. L. L. A. N. (1991). Propachlor degradation by a soil bacterial community. *Applied and environmental microbiology*, 57(8), 2135-2140.
89. Visscher, P. T., & Stolz, J. F. (2005). Microbial mats as bioreactors: populations, processes, and products. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 219(1), 87-100.
90. Wanner G., Vogl K. & Overmann J. (2008) Ultrastructural characterization of the prokaryotic symbiosis in "*Chlorochromatium aggregatum*". *J Bacteriol.* 190(10):3721-30.
91. Wanner, G., Vogl, K., & Overmann, J. (2008). Ultrastructural characterization of the prokaryotic symbiosis in "*Chlorochromatium aggregatum*". *Journal of bacteriology*, 190(10), 3721-3730.
92. Wenter R., Hütz K., Dibbern D., Li T., Reisinger V., Plösch M., Eichacker L., Eddie B., Hanson T., Bryant DA. & Overmann J. (2010) Expression-based identification of genetic determinants of the bacterial symbiosis "*Chlorochromatium aggregatum*". *Environ. Microbiol.* 12(8):2259-76.
93. Wilbanks, E. G., Jaekel, U., Salman, V., Humphrey, P. T., Eisen, J. A., Facciotti, M. T., ... &

Orphan, V. J. (2014). Microscale sulfur cycling in the phototrophic pink berry consortia of the Sippewissett Salt Marsh. *Environmental microbiology*, 16(11), 3398-3415.