



**Benemérita Universidad Autónoma de Puebla**

Instituto de Investigación en Ciencias  
microbiológicas

Laboratorio de Interacción bacteria-planta

**“Análisis del gen *cheY* de *Azospirillum baldaniorum* Sp245”**

**Tesis**

para obtener el grado de:

**MAESTRO EN CIENCIAS (MICROBIOLOGÍA)**

**CON OPCIÓN EN: BIOQUÍMICA Y GENÉTICA MICROBIANA**

**Presenta**

QBP Uriel Cardona Baltazar

**Directora de tesis**

DC Beatriz Eugenia Baca

Puebla, Puebla.

Noviembre 2021





## **Agradecimientos**

A mis padres, por apoyarme en todo momento y por sus consejos. Siempre serán mi más grande inspiración.

A mi asesora de tesis, la Dra. Beatriz Eugenia Baca quien dirigió este trabajo con profesionalismo y paciencia.

A mis sinodales, la Dra. Sandra Reyes Carmona, al Dr. Alberto Ramírez Mata (Laboratorio de Interacción bacteria- planta, CICM, BUAP), al Dr. Miguel Castañeda Lucio (Laboratorio genética molecular microbiana, CICM, BUAP) y la Dra. Lucia Soto Urzua (Laboratorio de Fisiología Microbiana de la Interacción Microorganismo-Hospedero, CICM, BUAP), por sus valiosas aportaciones durante los seminarios y tutorales de tesis y por su contribución en la revisión del manuscrito.

A mis compañeros de laboratorio por hacer más ameno la estancia y en especial a la QFB. Ma. Luisa Xiqui Vázquez y MC. Ibelith Herrera por su gran apoyo en el desarrollo de algunos experimentos.

De igual forma, al Centro de Investigaciones en Ciencias Microbiológicas de la BUAP por haberme abierto sus puertas para continuar mis estudios de posgrado, a la Vicerrectoría de Investigación y Estudios de Posgrado (VIEP) por el apoyo económico otorgado para la conclusión de esta tesis dentro del programa IV. Investigación y posgrado. Apoyar a los programas de posgrado para lograr su incorporación al Padrón Nacional de Calidad. Indicador establecido en el Plan de Desarrollo Institucional 2017-2021. Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por otorgarme una beca de maestría.

## **Abreviaturas**

%GC: porcentaje guanina citocina

aa: aminoácidos

ABA: Ácido abscísico

ACC: enzima 1-aminociclopropano 1-carboxilato

ADN: Ácido desoxirribonucleico

AIA: ácido indol acético

ATP: Adenina trifosfato

CLE: proteínasectoras tipo CheY

CS: cinasa detectora

C-terminal: carboxilo terminal

CW: en dirección a las manecillas del reloj

CWC: en dirección contraria a las manecillas del reloj

DGC: diguanilato ciclasa

di-GMPc: monofosfato dimérico cíclico de la guanosina

DO: densidad óptica

DR: dominio receptor

DT: dominio trasmisor

FA: fragmento A

FAB: fragmento A y B

FB: fragmento B

GA3: Ácido giberélico

*gusA-Km<sup>R</sup>*: casete de resistencia a kanamicina-beta-glucoronidasa

HGT: transferencia horizontal de genes

HK: cinasa histidinica

HKH: cinasa histidinica hibrida

lacZ: gen de la betagalactosidasa

LPS: lipopolisacáridos

MCP: proteínas aceptoras de grupo metilo  
MPM: marcador de peso molecular  
NaCl: cloruro de sodio  
N-terminal: amino terminal  
PCR: reacción en cadena de la polimerasa  
PDB: base de datos de proteínas  
PDE: fosfodiesterasa  
PGPR: rizobacteria promotora del crecimiento vegetal  
PH: proteína hipotética  
RR: reguladores de respuesta  
SDC: sistemas de doble componente  
SD-RR: regulador de respuesta de dominio único  
UFC: unidades formadoras de colonia  
X-gal: 5-bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosido

## Índice de figuras

Titulo	Pagina
Figura 1. Quimiotaxis de <i>Escherichia coli</i> .	12
Figura 2. Esquema de las proteínas que forman parte de los sistemas de dos componentes.	13
Figura 3. Reguladores de respuesta	14
Figura 4. Modelo en el cual CheY y Cle median el control del flagelo de <i>Caulobacter crescentus</i> .	16
Figura 5. Curva de crecimiento de la cepa 245 (WT) y la cepa mutante 5A	20
Figura 6. Mapa del contexto genético	20
Figura 7. Curva de crecimiento de <i>A. baldaniorum</i> $\Delta hkkB$ y <i>A. baldaniorum</i> Sp245 en MM K-malato.	21
Figura 8. Ensayo de movilidad	22
Figura 9. Fotografía del gel de agarosa de los productos de PCR	31
Figura 10. Fotografía de gel de agarosa de los productos de PCR y digestión enzimática para verificar la construcción pGEMFAB	32
Figura 11. Comprobación de la presencia del vector pCDR-2.	33
Figura 12. fotografía de gel de agarosa de productos PCR utilizando como oligonucleótidos forward del fragmento A y reverse del fragmento B.	34
Figura 13. Obtención de la mutante <i>A. baldaniorum</i> cheY-like Sp245.	33
Figura 13. Árbol filogenético de las 45 proteínas de <i>A. baldaniorum</i> y los 6 CheY's de <i>R. sphaeroides</i>	43
Figura 14. Árbol filogenético realizado con secuencias de proteínas de los resultados del BlastP	44
Figura 16. Alineamiento de secuencias de las proteínas Cle de <i>C. crescentus</i> y secuencias de proteínas de longitud mayor de <i>A. baldaniorum</i> Sp245	45
Figura 17. Dominio y estructura secundaria de la proteína CheY-like de <i>A. baldaniorum</i> Sp245	46
Figura 18. Modelos tridimensionales de la proteína CheY-like.	47
Figura 18. Gráfica de Ramachandran	48
Figura 19. Empalme de la estructura if4v y el modelo de CheY-like.	49
Figura 20. Empalme de múltiples estructuras de dominio REC.	51

## Índice de tablas

<b>Título</b>	<b>Página</b>
Tabla 1. Tabla de iniciadores utilizados en este trabajo.	25
Tabla 2. Material biológico utilizado en este trabajo.	25
Tabla 3. Programa con las condiciones necesarias para amplificar FA.	27
Tabla 4. Programa con las condiciones necesarias para amplificar FB.	27
Tabla 4. Reguladores de respuesta de dominio único y su ubicación en <i>A. baldaniorum</i> Sp245.	36
Tabla 5. Tabla del contexto genético de los reguladores de respuesta de dominio único en <i>A. baldaniorum</i> Sp245, no se agregan los genes huérfanos.	39

## Índice de contenidos

<b>Resumen</b>	<b>8</b>
<b>Introducción</b>	<b>10</b>
<b>Rizosfera</b>	<b>10</b>
<b><i>Azospirillum</i></b>	<b>10</b>
<b>Quimiotaxis</b>	<b>11</b>
<b>Sistema de doble componente</b>	<b>11</b>
<b>Cinasa histidinica</b>	<b>13</b>
<b>Regulador de respuesta</b>	<b>14</b>
<b>Reguladores CheY</b>	<b>16</b>
<b>Antecedentes generales</b>	<b>16</b>
<b><i>Azospirillum baldaniorum</i> Sp245</b>	<b>16</b>
<b>Quimiotaxis en <i>Azospirillum</i></b>	<b>17</b>
<b>Antecedentes específicos</b>	<b>19</b>
<b>Justificación</b>	<b>23</b>
<b>Objetivos</b>	<b>23</b>
<b>General</b>	<b>23</b>
<b>Específicos</b>	<b>23</b>
<b>Metodología</b>	<b>24</b>
<b>Materiales y métodos</b>	<b>25</b>
<b>Cultivo celular de <i>A. baldaniorum</i> Sp245</b>	<b>26</b>
<b>PCR para el amplificado del FA y FB de <i>A. baldaniorum</i> Sp245</b>	<b>26</b>
<b>Clonación FA y FB</b>	<b>26</b>
<b>Análisis bioinformático de la proteína CheY-Like</b>	<b>27</b>
<b>Herramientas empleadas</b>	<b>27</b>
<b>BLAST</b>	<b>27</b>

<b>RAST</b>	27
<b>SMART</b>	28
<b>MIST3</b>	28
<b>MEGAX</b>	28
<b>SWISS-MODEL</b>	28
<b>I-TASSER</b>	24
<b>PyMOL y UCSF CHIMERA</b>	24
<b>Búsqueda de proteínas de dominio REC</b>	24
<b>Construcción del árbol filogenético de secuencias obtenidas del GenBank</b>	24
<b>Análisis comparativo</b>	24
<b>Resultados</b>	25
<b>Construcción del plásmido mutagénico pCDR-2 <math>\Delta</math>cheY.L</b>	25
<b>Obtención del vector pGEMT-easy FAB</b>	25
<b>Modificación del vector pCDR</b>	27
<b>Vector suicida pCDR-2FAB</b>	28
<b>Análisis bioinformático de la proteína CheY-L</b>	29
<b>Búsqueda de reguladores de respuesta de dominio único</b>	29
<b>Contexto genético de los RR de dominio único tipo CheY en una superfamilia de Sp245</b>	38
<b>Árbol filogenético de las 45 proteínas de dominio único REC en <i>Azospirillum baldaniorum</i> Sp245</b>	42
<b>Búsqueda de región rica en arginina en proteínas con mayor longitud</b>	45
<b>Análisis estructural de la proteína CheY-L</b>	45
<b>Estructura terciaria de CheY-L</b>	46
<b>Análisis del modelo CheY-like con la estructura CheY de <i>E. coli</i></b>	48
<b>Análisis con otras estructuras REC de las proteínas WspR, CheY, RstA y CheY-L</b>	50
<b>Discusión</b>	52
<b>Bibliografía</b>	58

## RESUMEN

Las rizobacterias utilizan la materia orgánica o la mineral como fuente de nutrimentos y energía. Estos organismos son atraídos por las rizodeposiciones, es decir, los nutrimentos, exudados y el mucigel liberados por las raíces de las plantas, lo cual les permite establecerse en la rizósfera o colonizar el tejido interno de la raíz causando efectos significativos en el crecimiento y nutrición de la planta hospedera. Estos microorganismos pueden establecer una relación de simbiosis, asociación mutualista o parasitismo con la planta. *Azospirillum baldaniorum* Sp245 es una de las bacterias más estudiadas por ser promotora del crecimiento vegetal, especialmente considerada por su potencial como inoculante de plantas.

Estas bacterias perciben y responden a condiciones ambientales cambiantes. La percepción y procesamiento intracelular de estos estímulos se lleva a cabo por circuitos moleculares que detectan y amplifican las señales, permitiendo la generación de respuestas específicas. En procariontes existe una amplia red de detectores y efectores llamados sistemas de dos componentes, que transmiten señales a través de la fosforilación sucesiva de dos proteínas, una cinasa detectora y otra proteína definida como regulador de la respuesta, con residuos conservados de histidina y aspartato, respectivamente. Estos procesos reguladores son realizados por varios reguladores de respuesta que comparten el fosfoceptor común (receptor REC).

Por la importancia de las proteínas que conforman los sistemas de doble componente en los mecanismos de regulación, se realizó una búsqueda con herramientas bioinformáticas de los reguladores de respuesta de dominio REC único en el genoma de *A. baldaniorum* Sp245, así como de su contexto genético. Fueron identificados 45 genes que de manera presuntiva codifican para proteínas CheY, las cuales se podrían clasificar como una superfamilia, localizando los genes homólogos descritos en *Azospirillum brasilense* Sp7, otros que difieren en su asignación funcional implicadas en quimiotaxis, i.e. tres que presentan homología con las proteínas Cle de *Caulobacter crescentus*, las cuales unen al di-GMPc, además de varios genes *cheY* ligados a genes que codifican a cinasas histídnicas, lo que apunta a una versatilidad de probables funciones. Se realizó un breve análisis de relaciones evolutivas entre los genes *cheY*'s secuenciados de diferentes bacterias, ubicando a la proteína CheY en una rama diferente al género *Azospirillum*, sugiriendo la transferencia del gen por vía horizontal. De manera adicional se realizó el análisis de la proteína CheY-like en estructura primaria, secundaria y terciaria. Los datos bioinformáticos obtenidos mostraron que los residuos de aminoácidos requeridos para la fosforilación de la proteína están altamente conservados. Sin embargo, estos análisis no nos permitieron proponer su interacción con el aparato flagelar.

Se construyó el vector mutagénico pCDR-2 $\Delta$ *cheY*-like que permitió obtener la cepa mutante *A. baldaniorum* Sp245 $\Delta$ *cheY*-like. Posteriores estudios fenotípicos aportarán conocimiento relevante de la función de este gen.

## INTRODUCCIÓN

### 1. Rizosfera

Las rizobacterias son el grupo de bacterias que se aíslan de la rizosfera. La rizosfera es un sistema complejo y altamente dinámico compuesto de microambientes afectados por plantas y actividades microbianas. Algunas son fitopatógenas, es decir, provocan daño a la planta en la que se hospedan, otras promueven el desarrollo de las plantas donde son aisladas de la rizósfera; éstas se conocen como bacterias promotoras del crecimiento vegetal (PGPR, acrónimo del inglés, Plant Growth Promoting Rhizobacteria). Colonizan la parte de la raíz y suelo llamado rizósfera (Singh, 2018).

La mayoría de los organismos del suelo utilizan la materia orgánica o la mineral como fuente de nutrimentos y energía (Cassán *et al.*, 2015.) Estos organismos son atraídos por las rizodeposiciones, es decir, los nutrientes, exudados y el mucigel liberados por las raíces de las plantas, lo cual les permite establecerse en la rizósfera o colonizar el tejido interno de la raíz causando grandes efectos sobre el crecimiento y nutrición de la planta huésped. Estos microorganismos pueden establecer una relación de simbiosis, asociación mutualista o parasitismo con la planta. Las asociaciones mutualistas se clasifican en dos tipos: protectores contra patógenos y proveedores de nutrientes como los hongos arbusculares micorrízicos y las bacterias que producen sustancias reguladoras del crecimiento vegetal. Las PGPR pueden inducir el crecimiento vegetal directa o indirectamente. La influencia directa incluye la producción de fitohormonas, como el ácido indolacético (AIA), del grupo de las auxinas; también de ácido giberélico (GA3), citoquininas y ácido abscísico (ABA); o bien la capacidad de producir la enzima 1-aminociclopropano 1-carboxilato (ACC) desaminasa, que reduce el nivel de etileno en las plantas (Glick, 2014).

Además, los mecanismos directos incluyen la solubilización de fosfatos y micronutrientes y la fijación biológica de nitrógeno. Los efectos indirectos se deben a la modificación del ambiente rizosférico y su ecología, actuando como agentes de biocontrol de fitopatógenos mediante la liberación de sustancias como sideróforos,  $\beta$ -1, 3- glucanasas, quitinasas, antibióticos, pigmentos fluorescentes y ácido cianhídrico. Dentro de los microorganismos PGPR más estudiados pertenecen los géneros *Rhizobium*, *Azotobacter*, *Klebsiella*, *Beijerinckia*, *Pseudomonas*, *Bacillus*, *Azoarcus*, *Arthrobacter*, *Enterobacter*, *Burkholderia*, *Serratia*, y *Azospirillum* (Singh, 2018).

### 2. *Azospirillum*

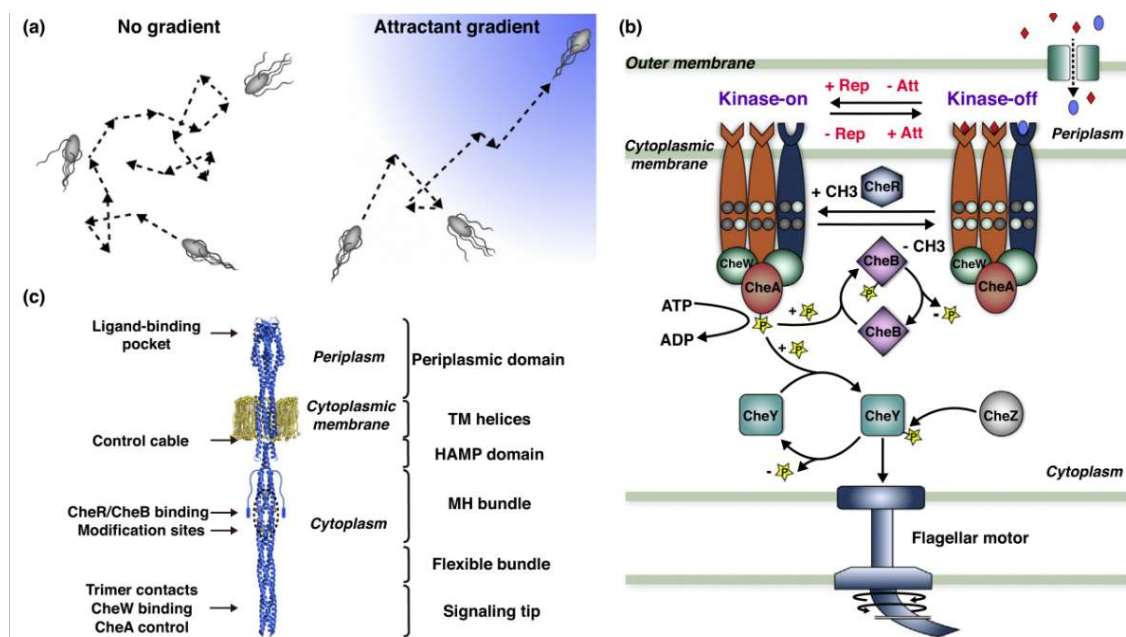
Uno de los géneros más estudiados es *Azospirillum* debido a su capacidad de mejorar significativamente el crecimiento y desarrollo, así como el rendimiento de las numerosas especies vegetales de interés agrícola (Bashan & de-Bashan, 2010). Las bacterias del género *Azospirillum* son rizobacterias Gram negativas, fijadoras

de nitrógeno de vida libre, de forma curva (vibroide), producen flagelos polares y peritricos; fija el nitrógeno molecular a amonio bajo condiciones microaeróbicas y en ausencia de nitrógeno combinado. Reduce el nitrato a nitritos, óxido nítrico y N<sub>2</sub>O y finalmente a N<sub>2</sub>; puede utilizar una amplia gama de azúcares, alcoholes y ácidos orgánicos como fuentes de carbono. Para la identificación de este género las características más importantes son la forma vibroide, pleomorfismo y movilidad en espiral, debida a la presencia de un flagelo polar responsable del movimiento natatorio y la presencia de flagelos peritricos responsables del movimiento en superficie. En la base de datos bioinformáticas del NCBI, se reconocen 24 especies del género *Azospirillum*: *A. agrícola*, *A. baldaniorum*, *A. brasilense*, *A. canadense*, *A. doebereineriae*, *A. fermentarium*, *A. formosense*, *A. halopraeferens*, *A. himalayense*, *A. humicireducens*, *A. largimobile*, *A. lipoferum*, *A. melinis*, *A. oleiclasticum*, *A. oryzae*, *A. palatum*, *A. palustre*, *A. picis*, *A. ramasamyi*, *A. rugosum*, *A. soli*, *A. thermophilum*, *A. thiophilum*, *A. zea*. (<https://lpsn.dsmz.de/genus/azospirillum>).

### 3. Quimiotaxis

La quimiotaxis bacteriana puede encontrar mejores entornos para el crecimiento de la bacteria, migrar hacia sustancias químicas favorables (atrayentes) y alejarse de sustancias químicas desfavorables (repelentes) (Porter, 2011). A pesar de características generales de la señalización quimiotáctica que parecen conservarse a través de bacterias, el núcleo conservado de proteínas de quimiotaxis canónicas comprende: proteínasceptoras de grupo metilo (MCP), las cuales al detectar algún tipo de gradiente químico activan a la proteína CheW que acopla a la cinasa histidinica CheA que, al sufrir el cambio conformacional se autofosforila y se convierte en una proteína fosfodonadora para el regulador de respuesta CheY que una vez fosforilado, tiene la capacidad de unir al motor del flagelo en la proteína FlIM, para así, cambiar la dirección del flagelo y redireccionar la natación de la bacteria (Figura 1A).

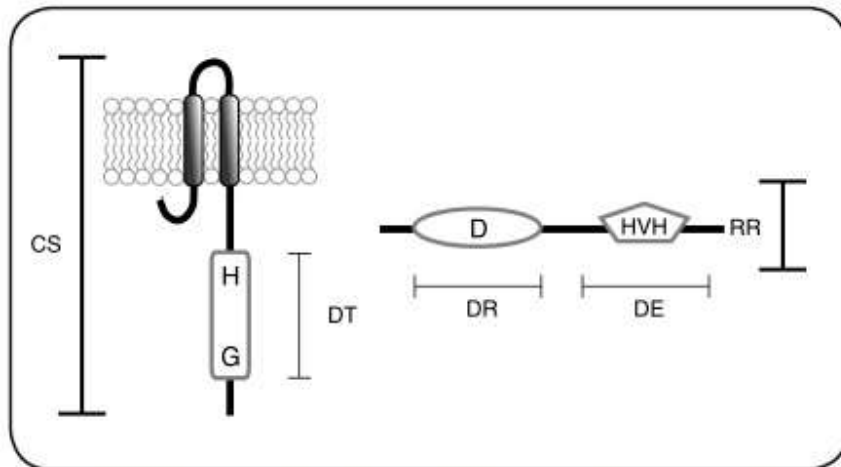
Además del núcleo de señalización conservado algunas bacterias poseen vías de señalización adicionales como el caso de las proteínas CheD, CheZ, CheC, CheX, etc. que se han observado juegan un papel importante en la modulación de la señalización del sistema de quimiotaxis canónico (Sourjik & Bi, 2018).



**Figura 1. Quimiotaxis de *Escherichia coli*.** **a)** Estrategia de quimiotaxis, en ausencia de un gradiente quimioefector, el movimiento de *E. coli* consiste en nados suaves interrumpido por vueltas transitorias. En un gradiente atractante la bacteria realiza una caminata aleatoria sesgada y periodos largos de corridas en dirección del gradiente. **b)** Vía de señalización. Atrayentes (Att; indicados de rombos y óvalos) promueven una inactividad (cinasa off) del complejo sensor, el cual gira en contra de las manecillas del reloj del motor flagelar y una natación suave. Repelentes (Rep) promueven un estado activo de la cinasa, girando en dirección a las manecillas del reloj la rotación del motor y direccionando. CheR y CheB modulan la metilación del receptor de acuerdo con el estado activo. **c)** estructura del receptor, representado por un diagrama de punteado con dominios y características específicas indicadas. Figura tomada de Sourjik & Bi, 2018.

#### 4. Sistemas de doble componente

Los organismos perciben y responden a condiciones ambientales cambiantes. La percepción y procesamiento intracelular de estos estímulos se lleva a cabo por circuitos moleculares que detectan y amplifican las señales, permitiendo la generación de respuestas específicas (Barba, 2014). En procariontes y algunos grupos de eucariontes, existe una amplia red de detectores y efectores llamados sistemas de dos componentes (SDC), que transmiten señales a través de la fosforilación sucesiva de dos proteínas, una cinasa detectora de histidina (CS) y otra conocida como regulador de la respuesta (RR), con residuos conservados de histidina y aspartato respectivamente (Figura 2) (De Phillips *et al.*, 2004).



**Figura 2. Esquema de las proteínas que forman parte de los sistemas de dos componentes.** Los sistemas de dos componentes están constituidos por una Cinasa detectora (CS) que posee un dominio transmisor (DT) que incluye el residuo Histidina conservada (H) y un dominio de unión a nucleótido importante para la unión de ATP conocido como caja G (G). Los reguladores de respuesta (RR) poseen un dominio receptor (DR) en donde se encuentra el residuo Asp (Tomado de Barba, 2014).

#### 4.1. Cinasas histidínicas

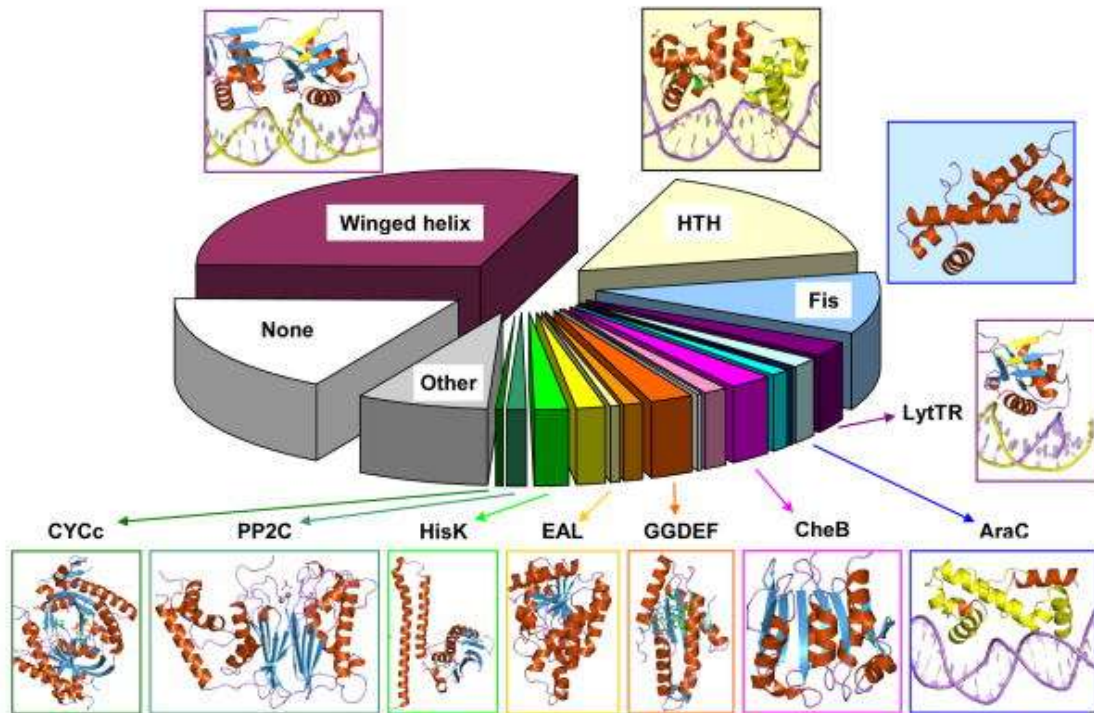
Las cinasas histidínicas son las primeras proteínas involucradas en la señalización en los SDC, y operan normalmente como dímeros los cuales se activan por medio de un proceso de trans-autofosforilación; y normalmente se encuentran asociados a la membrana, sin embargo, también pueden existir en el citoplasma bacteriano (Borland *et al.*, 2015) La histidina cinasa está constituida por un dominio sensor localizado en el extremo N-terminal y un dominio transmisor localizado en el extremo C-terminal de la proteína. El dominio sensor monitorea el ambiente en búsqueda de estímulos y el dominio transmisor incluye grupos de residuos esenciales para la activación y regulación de la actividad de la cinasa histidínica (Jung, 2012).

#### 4.2. Reguladores de respuesta

Los sistemas de dos componentes afectan procesos a nivel de genes y proteínas (cambios en la expresión génica, regulación alostérica de la actividad enzimática), a nivel de regulones globales de las células (motilidad celular, esporulación) y nivel multicelular (agregación celular, formación de biopelículas). Estos procesos reguladores son realizados por varios reguladores de respuesta que comparten el fosfoceptor común (receptor REC) pero difieren en sus dominios de salida (Figura 4). Por tanto, la asignación funcional de un RR se basa típicamente en el reconocimiento de su dominio de salida y su asignación funcional a la categoría de

proteína transportadora, enzima, de unión a ADN, de unión a ARN, de unión a ligando, de unión a proteínas, o transportadora (Galperin, 2011).

A



**Figura 3. Reguladores de respuesta.** Distribución de los reguladores de respuesta. Figura tomada de Galperin, 2011.

### 4.3. Reguladores CheY

Aunque la mayoría de los RR contienen el dominio REC con uno o más dominios de salida, el término "regulador de respuesta" fue acuñado en 1977 por Daniel Koshland para describir la proteína de quimiotaxis CheY que consiste únicamente en el dominio REC. Estas proteínas en un inicio se conocían por controlar la movilidad bacteriana y participar en la señalización de fosforrelievo (Galperin, 2011).

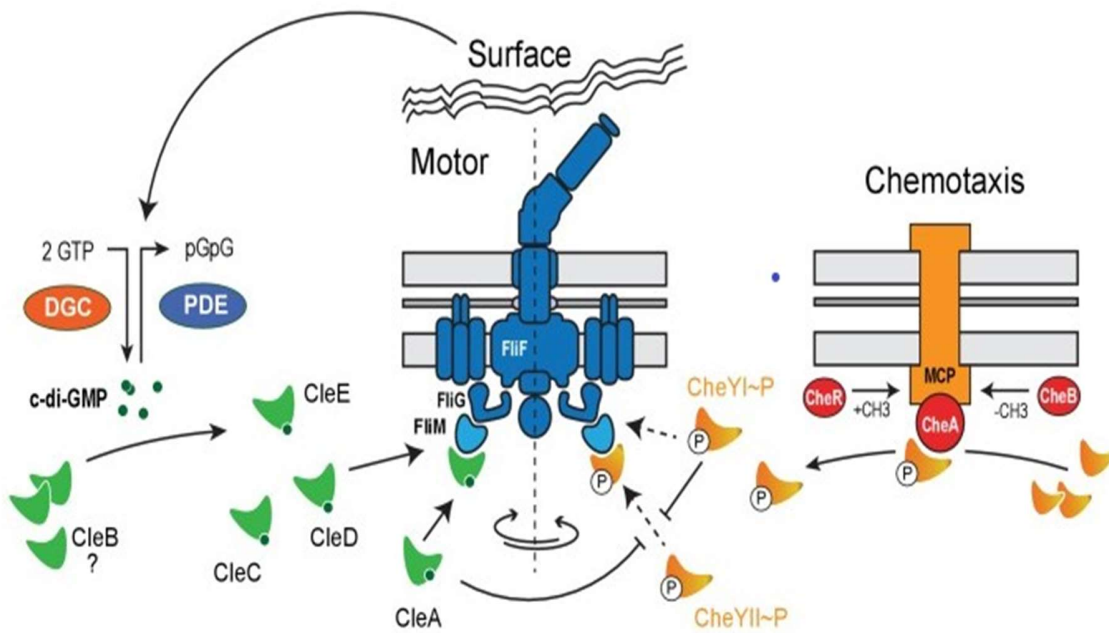
Un problema clave en el estudio de RR utilizando las técnicas estándar de análisis del genoma, es que la similitud de secuencia de RR no necesariamente indica identidad funcional. Por ejemplo, se ha descrito que las secuencias de las proteínas CheY de *A. brasilense* Sp7, las cuales regulan la movilidad (CheY1, CheY4, CheY6 y CheY7), comparten una identidad aproximada de 35% y tienen función diferente respecto a la movilidad de la bacteria (Mukherjee *et al.*, 2019). Además, también se

ha establecido la participación de reguladores de la quimiotaxis en otros procesos, como es el caso de CheY3 de *Shewanella oneidensis*, el cual se ha identificado como único regulador de la quimiotaxis, además participa en la formación de biopelícula (definiéndola como la fase de biopelícula adherida al fondo del recipiente) y película (El desarrollo de las células en la interfase aire/líquido), ya que al eliminar el gen *cheY3* que codifica esta proteína, se inhibe la biogénesis de biopelícula desde el inicio de la formación hasta la maduración de ésta, a través de la interacción física de CheY3 con al menos cuatro diguanilato ciclasas enzimas que participan en la formación del di-GMPc y de la biopelícula, lo que conduce a la inhibición de la actividad de estas enzimas y como consecuencia a la disminución de los niveles del segundo mensajero di-GMPc (Castelli *et al*, 2020).

Otra función descrita para las proteínas tipo CheY3 referida en el patógeno *Vibrio cholerae*, es que la proteína CheY3 participa en el mecanismo de dispersión de la biopelícula, la cual es una etapa del ciclo de la biopelícula donde las células pueden diseminarse y formar nuevas biopelículas y la participación de la proteína CheY3 es requerida para que las células puedan escapar por medio de los poros, que resultan de una degradación preliminar de la matriz de la biopelícula, sin embargo la mutación puntal del motivo (D13,D12 y D57) que debe ser fosforilado no afecta la función de dispersión de la biopelícula, lo que sugiere que la quimiotaxis no es necesaria (Bassier *et al*, 2020).

Por otra parte, se describió en *C. crescentus* el papel del segundo mensajero di-GMPc al modular la transición de la célula móvil a sedentaria a través de interferir con la función motora. Por medio de una nueva clase de proteínas similares a CheY denominadas proteínas Cle que se activan mediante la unión del segundo mensajero di-GMPc. Las proteínas Cle posee péptidos de mayor tamaño alrededor de 140 residuos de aminoácidos, que contienen regiones ricas en arginina y que comprenden sitios de unión de alta afinidad para di-GMPc, siendo activados por éste segundo mensajero y no por medio de fosforilación como los CheY canónicos (Nesper *et al.*, 2017). También se ha logrado determinar que las proteínas Cle unen a la proteína FliM del anillo C que conforma el motor flagelar, y se propone que funcionan como “freno” transitorio de la motilidad de la célula (Beeby *et al*, 2020) (Figura 4).

Los anteriores estudios nos indican una mayor versatilidad funcional de las proteínas CheY, que sugiere una adaptación a los nichos ecológicos de las bacterias.



**Figura 4. Modelo en el cual CheY y Cle median el control del flagelo de *Caulobacter crescentus*.** CheY (naranja) y proteínas Cle (verde) interactúan con la proteína FliM que actúa como interruptor flagelar. CheYII es el homólogo funcional de *E. coli* y es responsable de inducir la reversión del motor una vez fosforilado por el quimiorreceptor acoplado a la cinasa CheA (rojo) y subsecuentemente une a FliM en el motor flagelar. CleA (en su forma unida al di-GMPc) y CheYI-P compite con CheYII-P para unirse a FliM y de este modo reducir en general la tasa de reversiones del motor flagelar. CleA puede promover el nado suave adelante o hacia atrás a través de la competencia con CheYII en condiciones que aumentan los niveles de di-GMPc en *C. crescentus*. Las enzimas que regulan la concentración intracelular del di-GMPc (verde) y (DGC, en rojo; PDE, azul). CleC, CleE y posiblemente CleD interfiere con la función del motor para aumentar el rápido anclaje a la superficie. Estas proteínas pueden ser parte del mecanismo de retroalimentación a través del cual el flagelo detecta la superficie mediada por las enzimas DGC y/o PDE al aumentar la concentración del di-GMPc. Tomando de: Nesper, *et al.* (2017).

## ANTECEDENTES GENERALES

### 1. *Azospirillum baldaniorum*

En el año 2020 *Azospirillum brasilense* Sp245 recibe una reclasificación basada en su genoma como la cepa tipo de *Azospirillum baldaniorum*, nombrada así en honor a los doctores José Ivo Baldani y Vera Divan Baldani microbiólogos brasileños, por su contribución pionera a la investigación de *Azospirillum*. Por lo tanto, se adoptará la nueva notación de la bacteria en este trabajo. La cepa 245 es una de las bacterias más estudiadas por ser promotora del crecimiento vegetal, especialmente

considerada por su potencial como inoculante de plantas. Fue aislada de la superficie de la raíz desinfectada de trigo (*Triticum aestivum*) en Brasil. El contenido del DNA G+C basado en el genoma es del 68.4%. Las células son ligeramente curvas, de 1.6-2.1  $\mu\text{m}$  de largo y 0.5-0.7  $\mu\text{m}$  de ancho. Móvil con un flagelo polar largo y varios flagelos laterales de menor longitud. Fija nitrógeno en vida libre y las células pueden crecer en medio libre de nitrógeno. Crecen a 20 y 38 °C, reduce los nitratos a nitritos. Crecen en una variedad de fuentes de carbono como: L-Arabinosa, D-Arabitól, D-Fructosa, L-Fucosa,  $\alpha$ D-Galactosa, Gentioniosa, AD-Manitol, D-Manosa, Melibiosa, D-Picosa, L-Ramnosa, Acido fórmico, D-Galactónico ácido lactona, D-Acido glucosamínico,  $\alpha$ -ceto butírico,  $\alpha$ -Keto valérico, Glucoronamida, D-serina, Timidina, Fenitilamina, Glucosa-6-fosfato (dos Santos Ferreira *et al*, 2020).

## 2. Quimiotaxis en *Azospirillum*

La capacidad de *Azospirillum* de colonizar a la raíz depende de la movilidad y quimiotaxis, en respuesta a diversas moléculas químicas. *Azospirillum* dirige su nado hacia nichos que le benefician como compuestos orgánicos, azúcares y algunos aminoácidos. En tanto a los nichos que le son perjudiciales tiende a alejarse, como el exceso de oxígeno, aceptores de electrones, actividad redox y sustancias xenotóxicas. La bacteria quimiotáctica cuando no detecta ningún tipo de gradiente realiza un nado errático, que se caracteriza en nados de distancias cortas y en frecuentes cambios de dirección. Mientras que la célula que percibe algún tipo de attractante, se dirige hacia el gradiente de concentración con un nado en distancias mayores, así como de velocidad y disminuyen la frecuencia de giros en el nado (Figura 1A) (Rossmann *et al.* 2020).

El análisis de secuencias de genomas de varias especies de *Azospirillum* revelan que la respuesta quimiotáctica es coordinada por múltiples vías de quimiotaxis, codificadas por dos operones, Che1 y Che4, con receptores adicionales de quimiotaxis, cuyos genes codificantes se localizan en otros sitios del genoma (Mukherjee *et al.* 2016).

En *A. brasilense* Sp7 se han identificado 4 sistemas de quimiotaxis y múltiples proteínas homólogas de CheY que modulan la dirección de rotación del motor flagelar (Alexandre, 2015). En algunos casos los múltiples homólogos de CheY son codificados en diferentes regiones del genoma y no siempre ligados a un sistema de quimiotaxis. No está definido por qué la bacteria contiene múltiples reguladores de respuesta relacionados a la quimiotaxis (Mukherjee *et al.* 2019).

*A. brasilense* es móvil utilizando un solo flagelo polar que le permite nadar en medio líquido y cuando la viscosidad del medio incrementa las células producen múltiples flagelos laterales, estructuralmente distintos al flagelo polar, que permite el movimiento en superficies o "swarming" (Moens *et al.*, 1996). El flagelo polar de *A.*

*brasilense* Sp7 rota en ambas direcciones, es decir; en contra de las manecillas del reloj (CWC) y en dirección de las manecillas del reloj (CW), (Zhulin & Armitage, 1993) siendo la quimiotaxis quien controla estos cambios de dirección del flagelo (Mukherjee *et al*, 2016, 2019). La movilidad y la quimiotaxis son importantes para la colonización de la bacteria a la planta (Greer-Phillips *et al.*, 2004; O'Neal *et al.*, 2019, 2020; Cruz-Pérez *et al.*, 2021). Los datos experimentales obtenidos sugieren que el flagelo polar media la adsorción de las células de *A. brasilense* a las raíces de la planta en un proceso de dos etapas (Croes *et al.*, 1993). El flagelo polar media la formación de biopelícula *in vitro* (Viruega-Gongora *et al.* 2020). La flagelina polar de *A. brasilense* es glicosilada, contiene una cadena de tetrasacáridos de repetidos de ramanosa, fructosa, galactosa y N-acetil-glucosamina de estructura similar al antígeno O<sub>3</sub> del LPS (Belyakov *et al.*, 2012).

Dos diferentes sistemas regulan la quimiotaxis de *A. brasilense*, anotados como Che1 y Che4 y los RR CheYs (CheY6 y CheY7), no ligados a Che1 y Che4 (Bible *et al* 2012; Mukherjee *et al*, 2016, 2019). CheA1 y CheY1 se codifican en el operón Che1, estos regulan cambios transitorios de la velocidad del nado en la quimiotaxis. Los genes que codifican a CheA4 y CheY4 están ligados, estos controlan la probabilidad de cambios en la dirección de nado en la quimiotaxis. El fenotipo de las mutantes de *cheY6* y *cheY4* presentan reversión similar durante el movimiento natatorio, mientras que la mutante de *cheY7* no revierte. Todas las mutantes de *cheY* presentan alteración en el nado comparado con la cepa nativa en la ausencia de un gradiente. Evidencia genética y análisis fenotípicos indican que CheY6 es controlado por CheA1, y la actividad de CheY7 es controlado por CheA4. Las mutantes  $\Delta cheA4$ ,  $\Delta cheY4$  o  $\Delta cheY7$  no presentan quimiotaxis en gradiente espacial de quimioefectores, mientras que las mutantes que carecen de *cheY1* y *cheY6* son quimiotácticas bajo esas condiciones. La frecuencia de pausas en el movimiento natatorio se reduce en las mutantes *cheY1*, *cheY6* y *cheY7* pero ésta se incrementa en una cepa que carece *cheY4*. El mecanismo de estos efectos diferentes aún se desconoce. Los efectos de estos CheYs homólogos sobre el patrón de la movilidad natatoria de *A. brasilense* sería una estrategia para optimizar la navegación en el ambiente espacial y físicamente heterogéneo del suelo (Mukherjee *et al*, 2026, 2019).

Se describió que la rotación del flagelo que controla la movilidad natatoria estaría también participando en la transición de la célula móvil a sésil de diversas bacterias, quizá actuando el flagelo como un mecanodetector (Gordon & Wang, 2019). Así como la transición de móvil a sésil ocurre durante el "swarming" en superficies, en la formación de biopelícula y en la adherencia a las superficies (Guttenplan & Kearns, 2013).

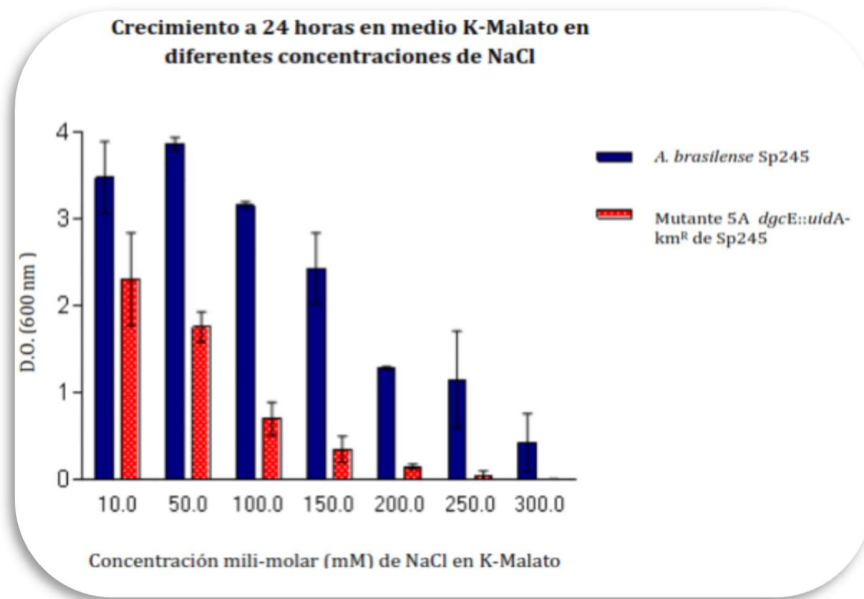
Se ha mostrado que la quimiotaxis y las mutantes de quimiotaxis intervienen en funciones sociales bacterianas: interacciones célula a célula o en la interacción con huésped eucarionte. En particular, los efectos de los CheY's de *A. brasilense* en la rotación del flagelo polar y quimiotaxis y su contribución en la transición de la célula

móvil a sésil, durante el “swarming”, adherencia a la superficie y formación de biopelícula (Ganusova *et al.*, 2021).

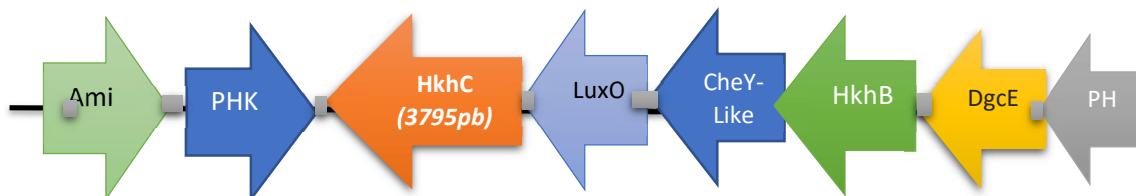
## ANTECEDENTES ESPECÍFICOS

Con el fin de analizar la participación de las proteínas con dominios GGDEF de *A. baldaniorum* Sp245 implicadas en la síntesis del segundo mensajero di-GMPc en funciones como el crecimiento y estrés, Gamboa en el 2014, realizó una mutación por inserción en el gen de la diguanilato ciclasa E, mediante un casete de resistencia a kanamicina-beta-glucuronidasa (*gusA-Km<sup>R</sup>*). El análisis fenotípico en cuanto a crecimiento de esta cepa mutante *A. baldaniorum* Sp245 5 A *cdgE::uidA-Km<sup>R</sup>*, fue realizado mediante la curva de crecimiento en base a las unidades formadoras de colonia (UFC) utilizando medio mínimo K-Malato. En cuanto al estrés, el crecimiento fue evaluado determinando densidad óptica, tras un periodo de incubación de 24 h, sometiendo a la bacteria a diferentes concentraciones de NaCl utilizando medio mínimo K-Malato. La cepa mutante *A. baldaniorum* 5 A registro una notable disminución del crecimiento, cuyo tiempo de generación de 190.77 minutos fue aproximadamente 3 veces mayor con respecto al tiempo de 65.1 minutos de la cepa silvestre *A. baldaniorum* Sp245, siendo posible observar directamente sobre la curva tal disminución a lo largo de las fases de crecimiento.

En relación a los datos obtenidos en condiciones de estrés osmótico, ambas cepas presentaron afección a la condición salina, siendo ésta más marcada entre mayor concentración de NaCl, además los datos presentados sugieren que la cepa mutante exhibió una significativa disminución del crecimiento cuando se compara con la cepa silvestre (Figura 5). Sin embargo, el fenotipo observado probablemente sea generado por una mutación polar, por la inserción del casete *Km<sup>R</sup>-GusA* en el gen *cdgE*, ocasionando la afectación de genes río abajo, tal como se observa en el contexto genético, donde se predice la posible presencia de un operón figura 6 (Gamboa 2014).



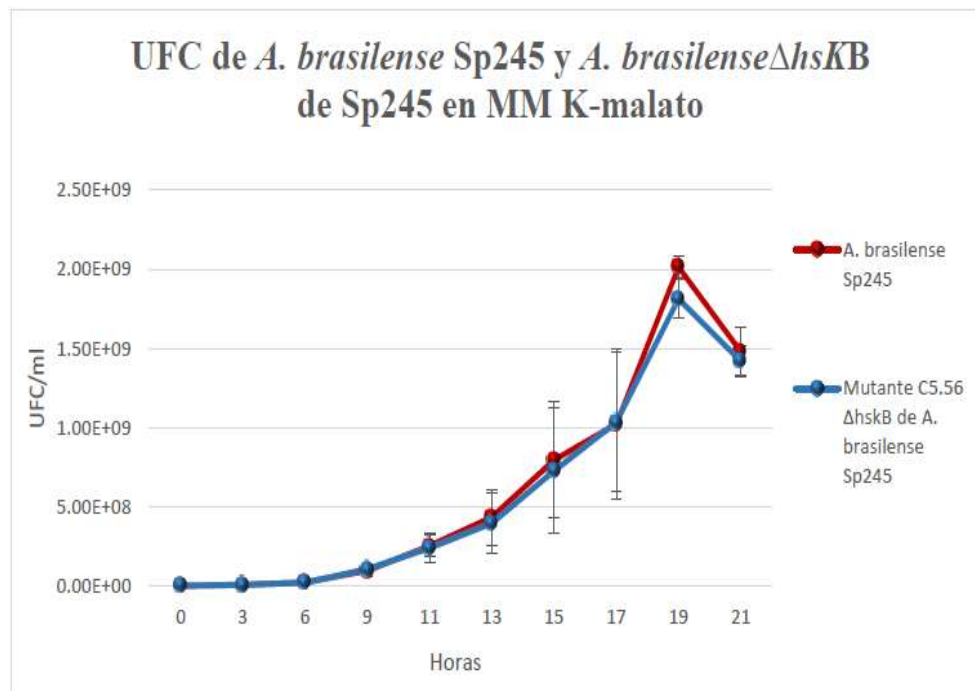
**Figura 5.** Curva de crecimiento de la cepa 245 (WT) y la cepa mutante *dgcE::uidA-km<sup>R</sup>* de Sp245 en distintas concentraciones de NaCl. Tomado de Gamboa, 2014.



**Figura 6. Mapa del contexto genético.** Ubicado en el plásmido 3 de *Azospirillum baldaniorum* Sp245 se encuentran en el mismo sentido de la transcripción, el regulador de respuesta CheY junto a una cinasa histidinica hibrida HkhB, un regulador de respuesta con función de factor de transcripción LuxO, una proteína cinasa histidinica hibrida (HyHKC; río abajo presente una diguanilato ciclasa *dgcE* y una proteína hibrida (PH). El mapa no está representado a escala de la longitud de los genes. Tomado de Gamboa 2014

Posteriormente, Avalos obtuvo la mutante *A. baldaniorum*  $\Delta cdgE$  7U3 (2016). La cual al analizar los fenotipos de crecimiento y de susceptibilidad al NaCl se observó que no existía una diferencia significativa de la mutante con respecto a la silvestre, como lo describió Gamboa (2014). Por lo cual la inserción de material genético *gusA-Km<sup>R</sup>* pudo generar una mutación polar, al afectar genes río abajo (Figura 6). Lo que llevó a estudiar el comportamiento de los demás genes, que se encuentran

en el contexto genético mediante mutaciones por eliminación del gen. Fue así como Cordero en el 2017 se dispuso a realizar la mutación por eliminación del gen *khkB*, el cual mostró que la mutante *A. baldaniorum*  $\Delta khkB$  C5.56 no presentó cambios en el crecimiento con respecto a la cepa silvestre *A. baldaniorum* Sp245. Las curvas de crecimiento obtenidas se muestran en DO 600nm (Figura 7).



**Figura 7. Curva de crecimiento de *A. baldaniorum*  $\Delta khkB$  y *A. baldaniorum* Sp245 en MM K-malato.** Se muestra la determinación mediante UFC (21 h). Los datos mostrados son el promedio de 4 ensayos independientes +/- desviación estándar. Tomado de Cordero, 2017.

Los datos obtenidos por Ávalos y Cordero sugerían la presencia de un operon constituido por cinco genes localizados en el contexto genético en secuencia 5'-3', con regiones intergénica cortas. De manera interesante al analizar la movilidad de las mutantes obtenidas se observó alteración en la movilidad realizada en medio mínimo K-malato, como se muestra en la Figura 8.



**Figura 8. Ensayo de movilidad.** Se observan las cepas utilizadas *A. baldaniorum* Sp245 (WT 245), *A. baldaniorum*  $\Delta$ hkhB (C5.56), *A. baldaniorum*  $\Delta$ dgcE (7U-3) y *A. baldaniorum* dgcE:gusA:Km (5-A). Tomado de Cordero, 2017.

Estos resultados apuntan a que la mutante por inserción *A. baldaniorum* 5-A generó un efecto polar y que los genes que constituyen un presunto operón participan en movilidad. Es probable que también se haya originado un defecto en la utilización de la fuente de carbono malato, que es el mejor quimioattractante de *A. brasilense*. Por lo que, es relevante continuar con el análisis de los siguientes genes que constituyen esta región de señalización. De aquí que se proponga el estudio del gen *che* Y-like.

## JUSTIFICACIÓN

*A. baldaniorum* Sp245 es una bacteria de gran interés en el área de la agricultura ya que se ha utilizado en inoculantes debido a sus efectos de PGPR, en especial la cepa Sp245, que es la cepa tipo de la especie *baldaniorum*. En esta bacteria se ha descrito múltiples sistemas de señalización, así como proteínas que integran estos sistemas, tales como cinasas histidínicas y reguladores de respuesta, pero aún falta mucho por resolver y entender. Los reguladores de respuesta de dominio único son componentes de señalización de vías de fosforilación de dos componentes que albergan un dominio receptor de grupo fosfato, pero carecen de un dominio de salida especializado, por lo que resulta relevante entender su función. Si bien, en un inicio se le adjudicaba solamente la función de proteína con participación de sistema de quimiotaxis, se han descrito gran diversidad de funciones como el arresto de la movilidad en las proteínas CLE de *C. crescentus* y que además de manera interesante estas proteínas no van a estar moduladas por medio de fosfotransferencia sino por medio de la unión del segundo mensajero di-GMPc. También se ha observado participación independiente de los sistemas de quimiotaxis. Por estos motivos es importante e interesante estudiar de manera más detallada a los reguladores de respuesta de dominio único para identificar su función en la bacteria, así como su regulación dentro de los sistemas de señalización.

## OBJETIVO GENERAL

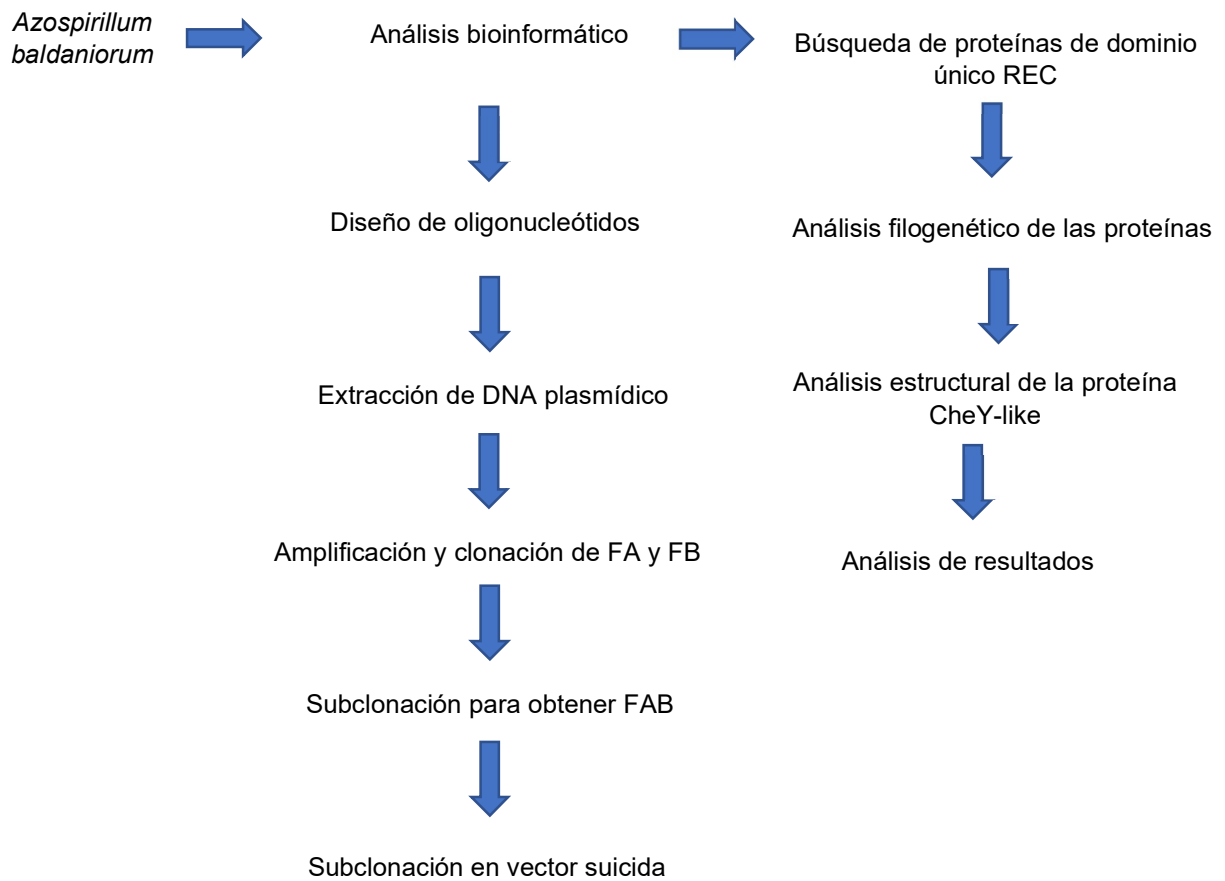
- Analizar del gen *cheY*-like de *A. baldaniorum* Sp245

## OBJETIVOS ESPECÍFICOS

- Construir el plásmido mutagénico pCDR-2, para la eliminación del gen *cheY*-like
- Analizar de manera bioinformática la proteína CheY

## METODOLOGÍA

### 1. Esquema del trabajo



## 2. Material y métodos

**Tabla 5.** Tabla de iniciadores utilizados en este trabajo.

OLIGONUCLEOTIDO	SECUENCIA DE DNA 5-3'	AMPLIFICADO
CheYSdaI-F	ATCCTGECAGGCGGTAATACAGATCCTCG	FA
CheYXbaI-R	CATCTAGAGGGTAGAGTCGCTTGGAG	FA
CheYXbaI-F	AGTCTAGACTCAAGAGTAGGCGGGTTCTGC	FB
CheYXmaI-R	ATCCCGGGCCCTCCATCAGGTTGATT	FB

Las bases subrayadas indican el sitio de restricción de la enzima. Los iniciadores se obtuvieron de *A. baldaniorum* Sp245

**Tabla 6.** Material biológico utilizado en este trabajo.

NOMBRE	DESCRIPCIÓN	REFERENCIA
<b>CEPAS</b>		
<i>E. coli</i> DH5 $\alpha$	<i>hds17 endA1 thi-1 gyrA96 relA1 recA1 supE44 D-lacU169 (<math>\Phi</math>80dlacZ <math>\Delta</math>M15)</i>	Gibco-BRL
<i>A. baldaniorum</i> Sp245	Cepa silvestre	Baldani et al. 1986
<i>E. coli</i> BL21	<i>E. coli. fhuA2 [lon] ompT gal [dcm] <math>\Delta</math>hdsS</i>	NEB
<b>VECTORES</b>		
PGEM®-Teasy	Vector de clonación. Amp <sup>R</sup>	PROMEGA
pCDR	Derivado de pJMS-pKm-sacB con fragmento XhoI/SnaBI de 3,075 pb de pGEM: lacZ traJ, ColE1, mob, rop, OriT, Amp <sup>S</sup> , Km <sup>S</sup> , Tc <sup>R</sup>	Cordero 2018
<b>PLÁSMIDOS</b>		
pGEM:FA	Derivado de pGEM Teasy con fragmento homólogo de 1038 pb. Amp <sup>R</sup>	Este trabajo
pGEM:FB	Derivado de pGEM Teasy con fragmento homólogo de 1086 pb. Amp <sup>R</sup>	Este trabajo
pGEM:FAB	Derivado de pGEM Teasy con fragmento de 2124 pb. Amp <sup>R</sup>	Este trabajo
pCDR-2	Derivado de pCDR sin el gen lacZ	Este trabajo
pCDR-2 $\Delta$ cheY-like	Derivado del vector pCDR-2 con fragmento de 2124 pb	Este trabajo

### 3. Cultivo celular

Se empleó *A. baldaniorum* Sp245 como la cepa silvestre. De un vial mantenido en crio-preservación se inocularon 50 µl en un tubo de 5 ml de medio mínimo K-malato (previamente esterilizado) y se incubó por un lapso de 18 h/30°C (pre-cultivo). Posteriormente, se realizaron cajas de medio rojo congo, las cuales se inocularon con el precultivo. Las clonas se aislaron por medio de estría cruzada y la caja de medio rojo congo se incubó 48 h/30°C.

### 4. PCR para el amplificado del FA y FB

La extracción de ADN genómico se realizó por la técnica de CTAB (Sambrook et al., 1989) Se partió de la mezcla de reacción para la PCR de acuerdo con las indicaciones del proveedor, variando de acuerdo al tipo de enzima utilizando la *Taq* DNA polimerasa recombinante (Invitrogen®) para comprobación; y las polimerasas de alta fidelidad *Taq* DNA polimerasa Platinum (Invitrogen®) con fines de secuenciación y clonación. Las condiciones de amplificación para los fragmentos A (1,038 pb) y B (1,086 pb) se muestran en la-Tabla y Tabla-4, respectivamente.

### 5. Clonación

Los productos de PCR, fragmento A (FA) y fragmento B (FB) fueron clonados en el vector pGEMT easy de manera independiente para ser liberados, posteriormente por medio de restricción enzimática utilizando las enzimas *SdaI-XbaI* para el FA y *XbaI-XmaI* para el FB para su posterior clonación sitio dirigida al vector suicida pCDR-2. Todas las clonaciones se transformaron en células quimio-competentes de la cepa *E. coli* DH5α, seleccionando las transformantes en placas de LB (100 µg/ml de Ap) para los plásmidos construidos en pGEMT Easy (pGEM:FA, pGEM:FB y pGEM:FAB), o bien en placas de LB (20 µg/ml de Km) para el plásmido modificado pCDR-2, de acuerdo a lo mencionado por Sambrook *et al.* (1989). Las restricciones enzimáticas y ligaciones fueron realizadas de acuerdo con las indicaciones del proveedor. La extracción de DNA y purificación del vector e inserto, se realizó de acuerdo con Sambrook *et al.* (1989). La comprobación de las construcciones fue realizada mediante PCR utilizando los iniciadores diseñados para el amplificado del FA y FB; y un perfil de restricción enzimático.

**Tabla 7.** Programa con las condiciones necesarias para amplificar FA.

Temperatura	Tiempo	Ciclos
94°C	5 min	1
94°C	30 seg	30
63°C	30 seg	
72°C	1 min	
72°C	5 min	1
4°C	∞	1

**Tabla 8.** Programa con las condiciones necesarias para amplificar FB.

Temperatura	Tiempo	Ciclos
94°C	5 min	1
94°C	30 seg	30
63°C	30 seg	
72°C	1 min	
72°C	5 min	1
4°C	∞	1

## ANÁLISIS BIOINFORMÁTICO

### 1. Herramientas bioinformáticas empleadas:

#### BLAST

Herramienta Básica de Búsqueda de alineación local, por sus siglas en inglés “BLAST”, la cual encuentra regiones de similitud entre secuencias biológicas. El programa compara secuencias de proteínas con las bases de datos de secuencias y calcula la significancia estadística. Herramienta en línea con dirección: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

#### RAST

El servidor RAST (Rapid Annotation using Subsystem Technology) proporciona anotaciones genómicas de alta calidad para procariontes en todo el árbol filogenético. Hace que una anotación de calidad SEED esté disponible como un servicio con un tiempo de respuesta de 48 horas. El entorno de SEED y las estructuras de datos SEED se utilizan para calcular las anotaciones automáticas. Una vez que se completa la anotación, los genomas pueden descargarse en una variedad de formatos para verse en línea. La anotación del genoma proporcionada incluye un

mapeo de genes a subsistemas y una reconstrucción metabólica. El servidor es en línea con la dirección: [RAST Server - RAST Annotation Server \(nmpdr.org\)](http://rast.nmpdr.org)

### SMART

Es una herramienta de investigación de arquitectura modular simple, permite la identificación y anotación de dominios genéticamente móviles y el análisis de arquitecturas de dominio. Estos dominios están ampliamente anotados con respecto a las distribuciones filéticas, la clase funcional, las estructuras terciarias y los residuos funcionalmente importantes. Cada dominio que se encuentra en una base de datos de proteínas no redundante, así como los parámetros de búsqueda y la información taxonómica, se almacenan en un sistema de bases de datos relacional. Las interfaces de usuarios de esta base de datos permiten búsquedas de proteínas que contienen combinaciones específicas de dominios en taxones definidos. dirección: <http://smart.embl-heidelberg.de/>

### MIST3

La base de datos MiST3.0. proporciona perfiles de transducción de señales de más de 125 000 genomas de bacterias y arqueas. Esta base de datos es el resultado de un escalamiento sustancial para acomodar datos genómicos microbianos en constante crecimiento. Aunque MiST3 se centra en la transducción de señales, se puede explorar utilizando funciones convenientes implementadas en Mist sin ejecutar software externo (Gumerov *et al*; 2020). Dirección: <https://mistdb.com/>

### CLUSTALOMEGA

Clustal Omega es un programa de alineación de secuencias múltiples que utiliza árboles guía sembrados y técnicas de perfil de perfil HMM para generar alineaciones entre tres o más secuencias.

### MEGAX

Es un software de computadora para realizar análisis estadísticos de la evolución molecular y para construir árboles filogenéticos. Incluye muchos métodos y herramientas sofisticados para filogenómica y filomedicina. Tiene diversas funciones como: construcción de alineación de secuencia, manejo de datos, visor de datos de secuencia, estimación basada en MCL(Probabilidad compuesta máxima) de patrones de sustitución de nucleótidos, prueba de homogeneidad del patrón de sustitución, método de estimación de distancias de nucleótido, pruebas de selección, entre otros (Masatoshi Nei *et al*; 2016).

### SWISS-MODEL

Es un servidor integrado basado en la web (dirección <https://swissmodel.expasy.org/interactive>) dedicado al modelado por homología de

estructuras de proteínas. Guía al usuario en la construcción de modelos de homología de proteínas en diferentes niveles de complejidad. Para la construcción de un modelo por homología, en este servidor se requiere: 1) identificación de plantillas estructurales, 2) alineación de la secuencia objetivo y estructura (s) de plantillas, 3) construcción de modelos y 4) evaluación de la calidad del modelo. Estos pasos requieren software especializado e integran bases de datos de secuencia y estructura de proteínas actualizadas. Cada uno de los pasos anteriores se puede repetir de forma interactiva hasta que se logre un resultado de modelado satisfactorio. Dirección: <https://swissmodel.expasy.org/>

#### I-TASSER

El servidor I-TASSER es una plataforma en línea que implementa los algoritmos basados en I-TASSER para la estructura de proteínas y predicciones de funciones. Permite a los usuarios generar automáticamente predicciones de modelos de alta calidad de la estructura 3D y la función biológica de las moléculas de proteínas a partir de sus secuencias de aminoácidos (Zhang *et al.*, 2010). Dirección: <https://zhanggroup.org/I-TASSER/>

#### PyMOL y USCF CHIMERA

PyMOL es un visor molecular de código abierto y auspiciado por usuarios creado por Warren Lyford Delano y comercializado por Delano Scientific LLC (Warren Lyford De Lano). Dirección: <https://pymol.org/2/>

USCF CHIMERA es un programa para la visualización interactiva y el análisis de estructuras moleculares y datos relacionados, incluidos mapas de densidad, trayectorias y alineaciones de secuencias. Dirección: <https://www.cgl.ucsf.edu/chimera/>

## 2. Búsqueda de proteínas de dominio único REC

Para realizar la búsqueda de las proteínas de dominio único REC en *A. baldaniorum* se utilizó el servidor RAST (Rapid Annotation using Subsystem Technology) versión 2.0, en donde se ingresó el genoma de *A. baldaniorum*, una vez ingresado, se examinó el genoma comparando con la secuencia de la proteína CheY-like analizada en este trabajo, se realizó el BLASTP y los resultados se filtraron para seleccionar y analizar únicamente las proteínas de dominio único REC. Para la selección de proteínas se utilizó la plataforma SMART (Simple Modular Architecture Research Tool) para la identificación y análisis de arquitectura de los dominios.

### 3. Construcción del árbol filogenético de secuencias obtenidas del GenBank

Las secuencias de los 45 reguladores de respuesta de dominio único y los seleccionados de *A. brasilense* Sp7 y *Rhodobacter sphaeroides* JPA1337 se ingresaron al software MEGA (Molecular Evolutionary Genetics Analysis) versión 10.2.2. para la construcción del árbol filogenético que se realizó por Maximum likelihood con un bootstrap de 1000.

### 4. Análisis comparativo

Las proteínas CheY1, CheY4, CheY6 y CheY7 descritas en *A. brasilense* Sp7 con función en quimiotaxis (Mukherjee *et al.*, 2016; Mukherjee *et al.*, 2019) y *Rhodobacter sphaeroides* (Porter *et al.*, 2006), se emplearon para analizar las proteínas homólogas de *A. baldaniorum*, realizando los alineamientos de secuencias con el programa Clustal Omega versión 2.0.12. Este mismo programa se utilizó para realizar el análisis de residuos de aminoácidos conservados implicados en la función quimiotáctica tomando como modelo las secuencias de las proteínas CheY descritas de manera funcional.

Para el análisis estructural de la proteína CheY-like se ingresó la secuencia de aminoácidos en el servidor I-TASSER (Iterative Threading ASSEmby Refinement) para la predicción de la estructura secundaria, así como tridimensional. Utilizando el modelo 1 ya que presentaba mejores parámetros de calidad de predicción. Se utilizó también para la estructura secundaria la herramienta PHYRE2 (Protein Fold Recognition Server).

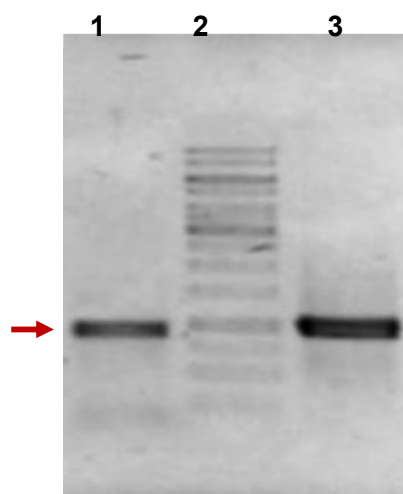
En tanto al modelado por homología se utilizó el servidor SWISS-MODEL para un análisis estructural de las proteínas utilizando como plantillas el cristal CheY de *E. coli* (Lee *et al.*, 2001) apoyado del servidor PyMOL y CHIMERA 1.15 para la visualización de las moléculas.

## RESULTADOS

### OBTENCIÓN DE LA MUTANTE *A. baldaniourm* $\Delta$ *cheY-L*

#### 1. Construcción del plásmido mutagénico

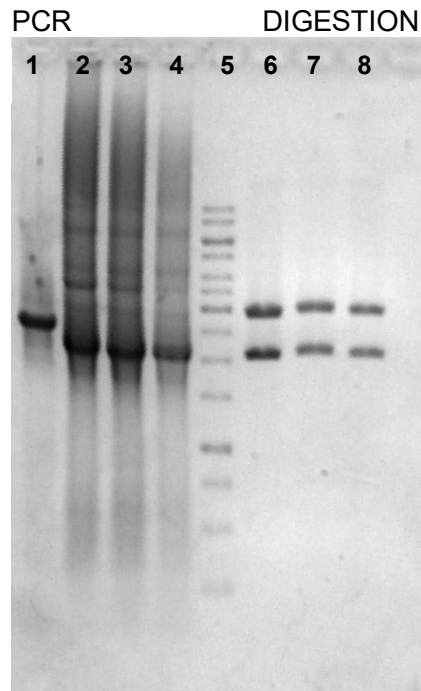
La construcción del plásmido mutagénico se inició con la amplificación de los fragmentos A y B, fragmentos que se encuentran aledaños al gen *cheY-like*. Ambos fragmentos se amplificaron con los oligonucleótidos descritos en tabla 1. Se obtuvieron las bandas correspondientes al peso molecular de ambos fragmentos Figura 9.



**Figura 9.** Electroforetograma de los productos de PCR, utilizando los oligonucleótidos diseñados para amplificar fragmento A y B que se encuentran río arriba y río abajo del gen *cheY-like*. Carril 1) Fragmento A, carril 2) MPM 1 kb y carril 3) Fragmento B. Flecha indica banda de 1000 pb.

#### 2. Obtención del vector pGEMT-easy FAB

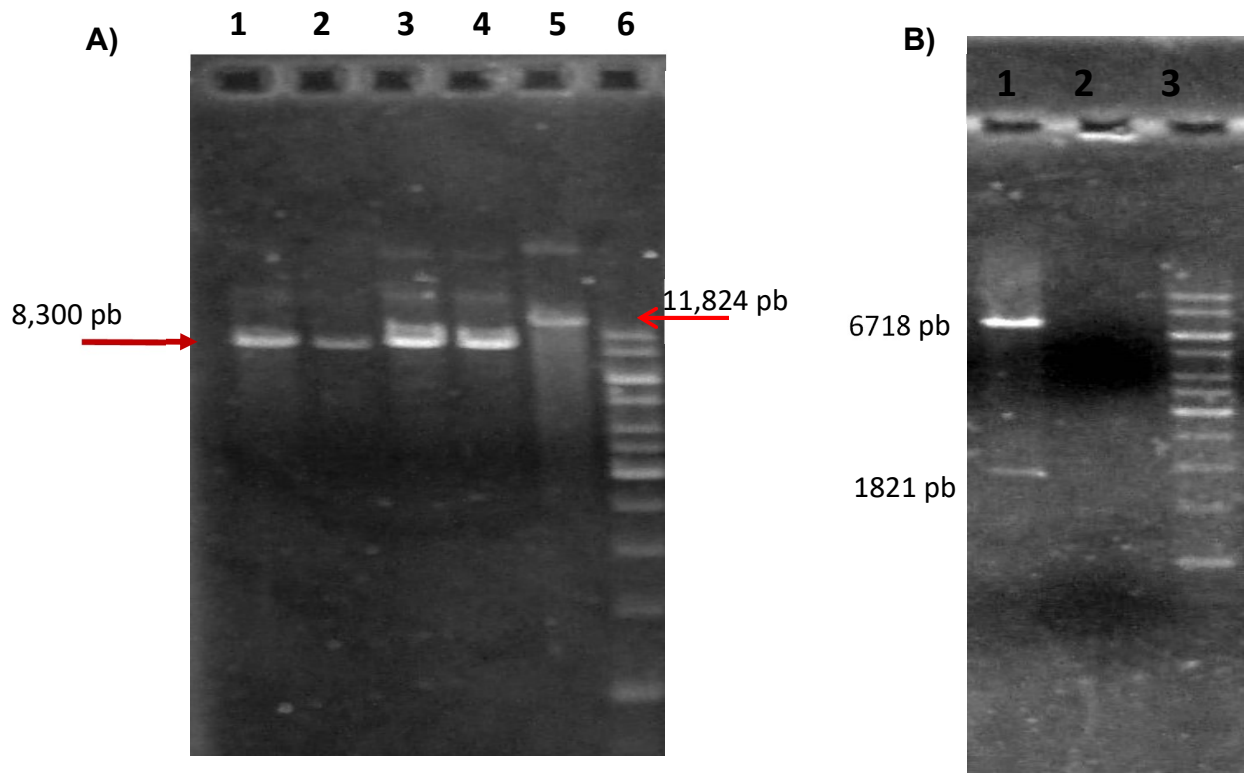
Posteriormente, a partir de los amplificados de los fragmentos A y B, se realizaron subclonaciones de cada uno en el vector de mantenimiento pGEMT-easy, lo que permitió realizar una clonación de ambos fragmentos en el mismo vector (pGEMT-easy FAB). Una vez obtenido pGEMFAB, se realizó la liberación del FAB por digestión con las enzimas *SdaI* y *XmaI*. El FAB presenta un tamaño 2035 pb. La construcción se verificó por medio de PCR y restricciones por digestión con las enzimas *SdaI* y *XmaI* (Figura 10).



**Figura 10.** Electroforetograma de los productos de PCR y digestión enzimática para verificar la construcción pGEMFAB. Carril 1) producto de PCR del fragmento A y B en la construcción, carril 2, 3 y 4 amplificados del plásmido de *A. baldaniorum* Sp245, carril 6,7 y 8 productos digestión enzimática.

### 3. Modificación del vector pCDR

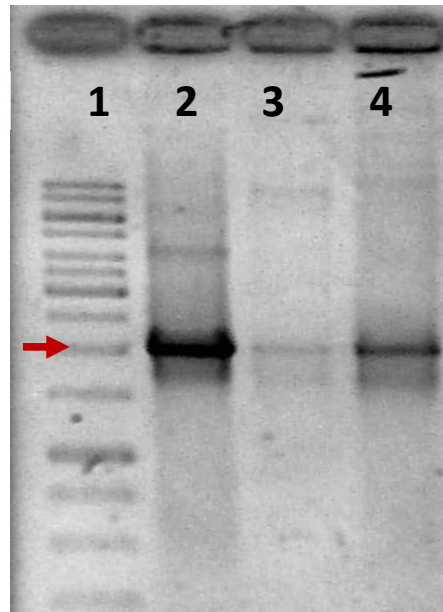
Con el fin de propiciar la recombinación del vector suicida pCDR y optimizar la conjugación, se eliminó el gen de la beta galactosidasa (*lacZ*) con la enzima *NodI*. Para corroborar la construcción se probó el crecimiento con el compuesto X-gal. Se observaron las clonas que contenían el plásmido pCDR intacto (con el gen *lacZ*) la en color azul, mientras que las clonas que contienen el plásmido pCDR-2 (sin *lacZ*) las colonias se observaron blancas (Datos no mostrados). Una vez seleccionadas las clonas se realizó su comprobación con enzimas de restricción que permitieron observar la disminución del peso molecular del plásmido (Figura 11).



**Figura 11. Comprobación de la presencia del vector pCDR-2.** **A)** Electroforetograma de productos de digestión enzimática del vector pCDR-2 con enzima *XhoI*. Carril 1, 2, 3, 4 digestión del vector modificado, carril 5 digestión del vector sin modificar y carril 6 marcador de peso molecular 1 kb. **B)** fotografía de gel de agarosa de productos de digestión enzimática con la enzima *PstI*. carril 1 clona 1, carril 2 vacío y carril 3 marcador de peso molecular.

#### 4. Vector suicida pCDR-2FAB

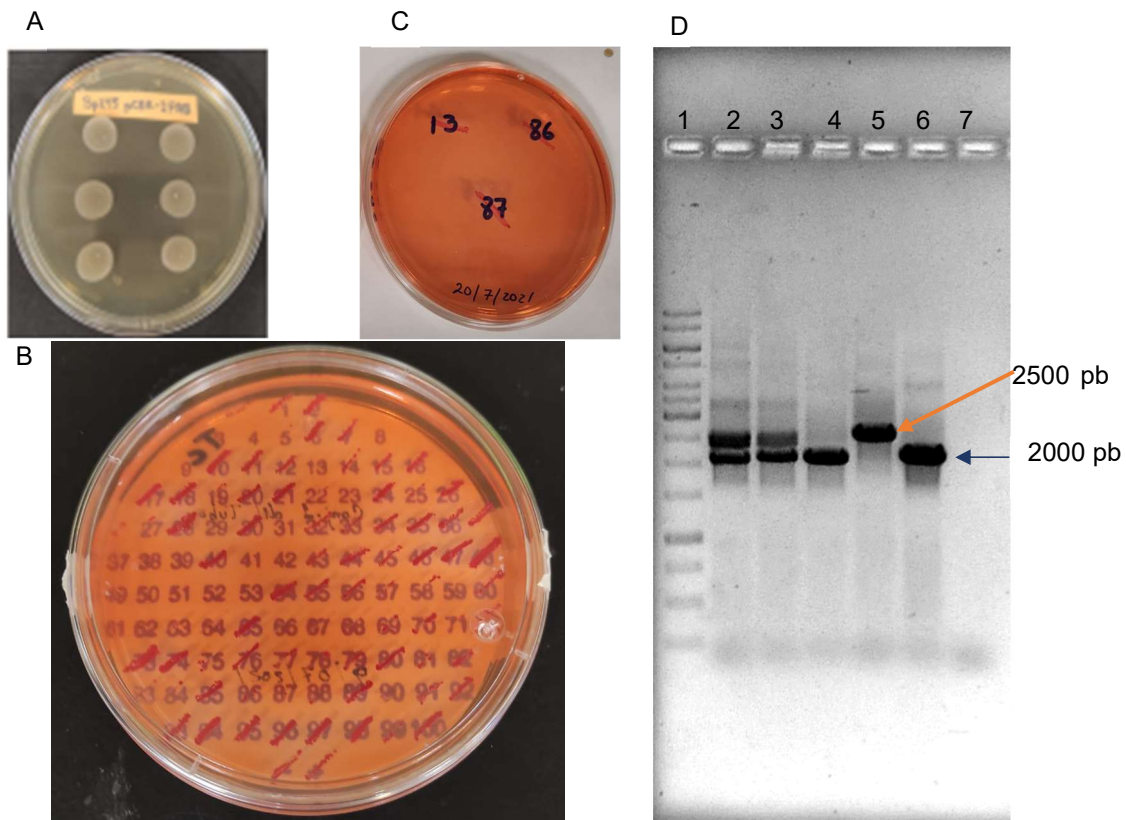
Una vez obtenido el vector pCDR-2 se realizó la subclonación de FAB, primeramente, liberado de la construcción pGEM-easyFAB. Para ello, se linealizó el vector suicida con la enzima *XmaI*, se realizó la ligación y posteriormente se transformó a la cepa *E.coli* S17.1. Para corroborar la construcción del vector mutagénico pCDR-2FAB se realizó un análisis de PCR. En donde se obtuvo un amplificado positivo para los iniciadores “Foward” fragmento A y “Reverse” fragmento B (Figura12).



**Figura 12.** Electroforetograma de productos PCR utilizando como oligonucleótidos forward del fragmento A y reverse del fragmento B. Carril 1 marcador de peso molecular, carril 2 clona 17, carril 3, clona 36 y carril 3 clona 37. La flecha indica la banda con peso molecular de 2000 pb.

### **5. Obtención de la mutante *A. baldaniorum* $\Delta$ *cheY-like* de Sp245 utilizando el vector pCDR-2FAB**

El plásmido pCDR-2FAB fue incorporado en células quimiocompetentes *E. coli* S17.1 por medio de transformación, la cual fue utilizada como cepa donadora en los ensayos de conjugación biparental realizados con la cepa receptora *A. baldaniorum* Sp245. La selección positiva, se hizo en placas de medio mínimo de K-malato. Posteriormente, la selección negativa, se hizo en placas de medio mínimo k-malato adicionado con sacarosa al 15%, siendo posibles candidatas las clonas sensibles a Tetraciclina. Como comprobación de las posibles mutantes se realizó un amplificado de la región afectada por medio de los iniciadores Foward fragmento A y Reverse fragmento B (Figura13).



**Figura 13. Obtención de la mutante *A. baldaniorum cheY-like Sp245*.** A) conjugación biparental entre *A. baldaniorum* Sp245 y *E. coli* S17.1 pCDR-2FAB. B) medio mínimo adicionado con k-malato rojo Congo+Tc, se seleccionan las clonas sensibles a tetraciclina, clonas 13, 86 y 87. C) Medio K-malato+ RC+15 % sacarosa. D) Electroforetograma de PCR utilizando Foward FA y reverse FB. Carril 1 MPM 1 kb, carril 2 clona13, carril3 clona 86, carril 4 clona 87, carril 5 *A. baldaniorum*, carril 6 *E. coli* S17.1 pCDR-2FAB y carril 7 blanco de reactivos.







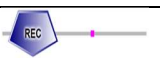





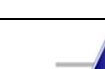
## ANÁLISIS BIOINFORMÁTICO














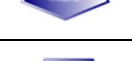




### 1. Búsqueda de reguladores de respuesta de dominio único









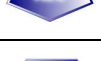





Se realizó la búsqueda de proteínas que solamente contuvieran el dominio REC en su arquitectura modular dentro de los genomas de *A. baldaniorum* Sp245 (Tabla) y *A. brasilense* Sp7. Así mismo, se realizó un análisis sistemático y una comparación de la distribución para identificar los productos putativos de reguladores de respuesta de dominio único (REC, SD-RR) que están involucrados en sistemas de

señalización. Este análisis llevó a la identificación de 45 reguladores de respuesta de dominio único en *A. baldaniorum* Sp245 y 41 en *A. brasilense* Sp7. Se observó una gran variabilidad en la longitud de secuencia de estas proteínas identificadas, que van desde 119 hasta los 629 aminoácidos. Así mismo se identificó la distribución en el genoma de Sp245 observando los RR de dominio único, en el cromosoma y los crómidos 1, 2, y 3, y el plásmido 4, Tabla 4.

**Tabla 4.** Reguladores de respuesta de dominio único y su ubicación en *A. baldaniorum* Sp245.

#	MIST3 id	Número de acceso GenBank	Localización	SMART	Longitud Residuos de aa
1	GCF_000237365.1-AZOBR_RS14225	WP_014241666.1	Cromosoma		125
2	GCF_000237365.1-AZOBR_RS14200	WP_014241661.1	Cromosoma		138
3	GCF_000237365.1-AZOBR_RS12415	WP_014241240.1	Cromosoma		121
4	GCF_000237365.1-AZOBR_RS12660	WP_014241309.1	Cromosoma		129
5	GCF_000237365.1-AZOBR_RS09890	WP_014240697.1	Cromosoma		137
6	GCF_000237365.1-AZOBR_RS09475	WP_014240604.1	Cromosoma		350
7	GCF_000237365.1-AZOBR_RS10740	WP_014240885.1	Cromosoma		364
8	GCF_000237365.1-AZOBR_RS35795	WP_051657866.1	Cromosoma		629
9	GCF_000237365.1-AZOBR_RS07695	WP_082188147.1	Cromosoma		179
10	GCF_000237365.1-AZOBR_RS06010	WP_014239878.1	Cromosoma		134
11	GCF_000237365.1-AZOBR_RS06395	WP_014239960.1	Cromosoma		343
12	GCF_000237365.1-AZOBR_RS04260	WP_014239494.1	Cromosoma		137
13	GCF_000237365.1-AZOBR_RS07605	WP_014240212.1	Cromosoma		168

14	GCF_000237365.1-AZOBR_RS01020	WP_014238757.1	Cromosoma		132
15	GCF_000237365.1-AZOBR_RS00885	WP_051658041.1	Cromosoma		119
16	GCF_000237365.1-AZOBR_RS02220	WP_014239024.1	Cromosoma		121
17	GCF_000237365.1-AZOBR_RS21690	WP_014198656.1	Crómido 1		264
18	GCF_000237365.1-AZOBR_RS20910	WP_014198488.1	Crómido 1		162
19	GCF_000237365.1-AZOBR_RS20090	WP_014198312.1	Crómido 1		121
20	GCF_000237365.1-AZOBR_RS21935	WP_014198710.1	Crómido 1		125
21	GCF_000237365.1-AZOBR_RS21825	WP_014198687.1	Crómido 1		124
22	GCF_000237365.1-AZOBR_RS19295	WP_014198146.1	Crómido 1		138
23	GCF_000237365.1-AZOBR_RS20765	WP_014198447.1	Crómido 1		184
24	GCF_000237365.1-AZOBR_RS20560	WP_014198411.1	Crómido 1		127
25	GCF_000237365.1-AZOBR_RS17435	WP_014197714.1	Crómido 1		159
26	GCF_000237365.1-AZOBR_RS17205	WP_014197668.1	Crómido 1		136
27	GCF_000237365.1-AZOBR_RS16990	WP_014197623.1	Crómido 1		144
28	GCF_000237365.1-AZOBR_RS16625	WP_014197547.1	Crómido 1		161
29	GCF_000237365.1-AZOBR_RS16185	WP_014197449.1	Crómido 1		137
30	GCF_000237365.1-AZOBR_RS14720	WP_014197107.1	Crómido 1		127
31	GCF_000237365.1-AZOBR_RS14590	WP_014197081.1	Crómido 1		135

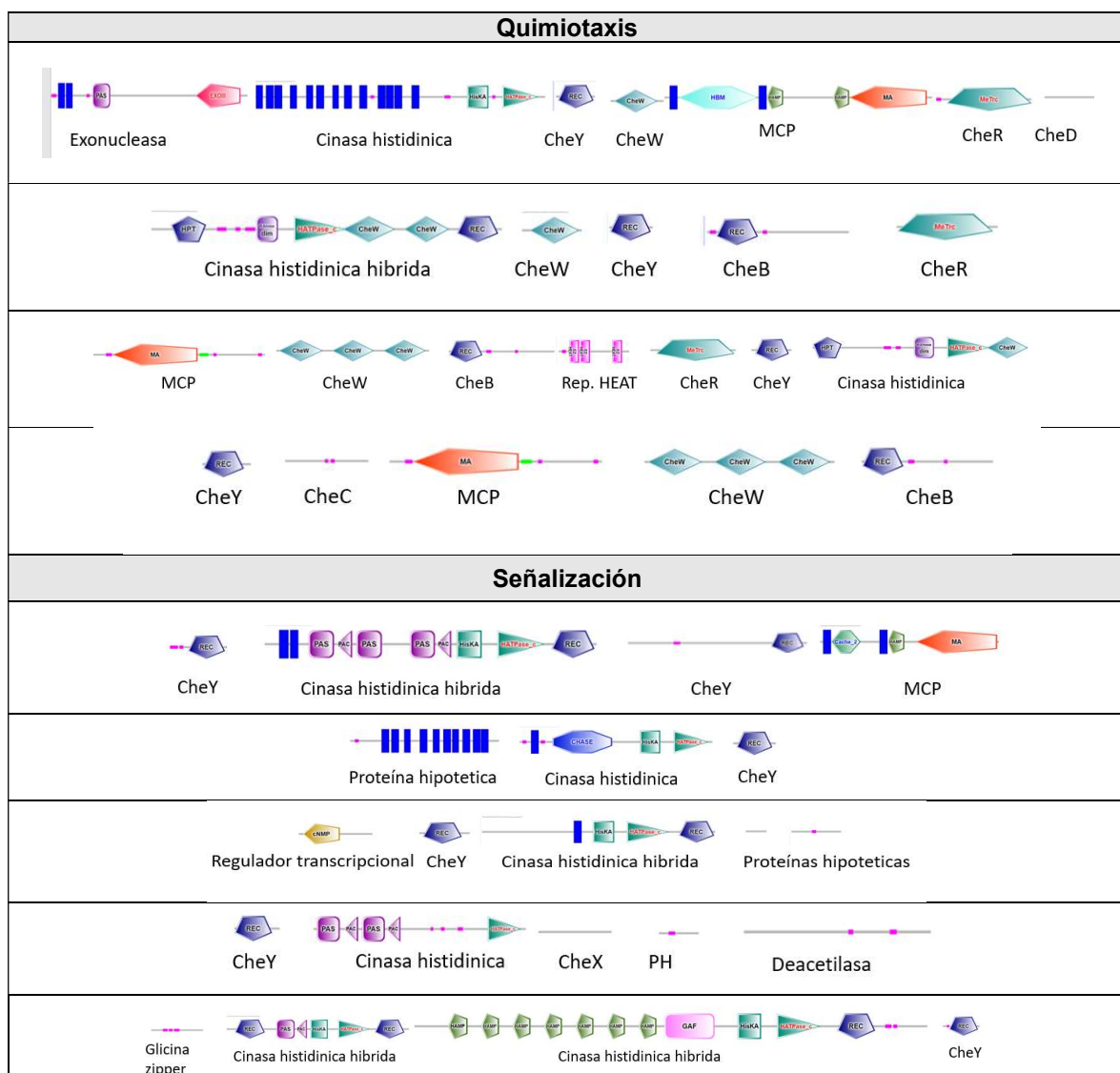
32	GCF_000237365.1-AZOBR_RS25995	WP_014242548.1	Crómido 2		132
33	GCF_000237365.1-AZOBR_RS25980	WP_041813258.1	Crómido 2		137
34	GCF_000237365.1-AZOBR_RS25940	WP_014242539.1	Crómido 2		135
35	GCF_000237365.1-AZOBR_RS25765	WP_014242497.1	Crómido 2		144
36	GCF_000237365.1-AZOBR_RS24210	WP_014242147.1	Crómido 2		127
37	GCF_000237365.1-AZOBR_RS24025	WP_014242094.1	Crómido 2		124
38	GCF_000237365.1-AZOBR_RS23095	WP_014241880.1	Crómido 2		124
39	GCF_000237365.1-AZOBR_RS22685	WP_014241789.1	Crómido 2		135
40	GCF_000237365.1-AZOBR_RS29510	WP_014199555.1	Crómido 3		139
41	GCF_000237365.1-AZOBR_RS26955	WP_014198983.1	Crómido 3		128
42	GCF_000237365.1-AZOBR_RS26875	WP_014198966.1	Crómido 3		149
43	GCF_000237365.1-AZOBR_RS26135	WP_014198787.1	Crómido 3		144
44	GCF_000237365.1-AZOBR_RS31795	WP_014200077.1	Plásmido 4		120
45	GCF_000237365.1-AZOBR_RS30855	WP_014199861.1	Plásmido 4		160

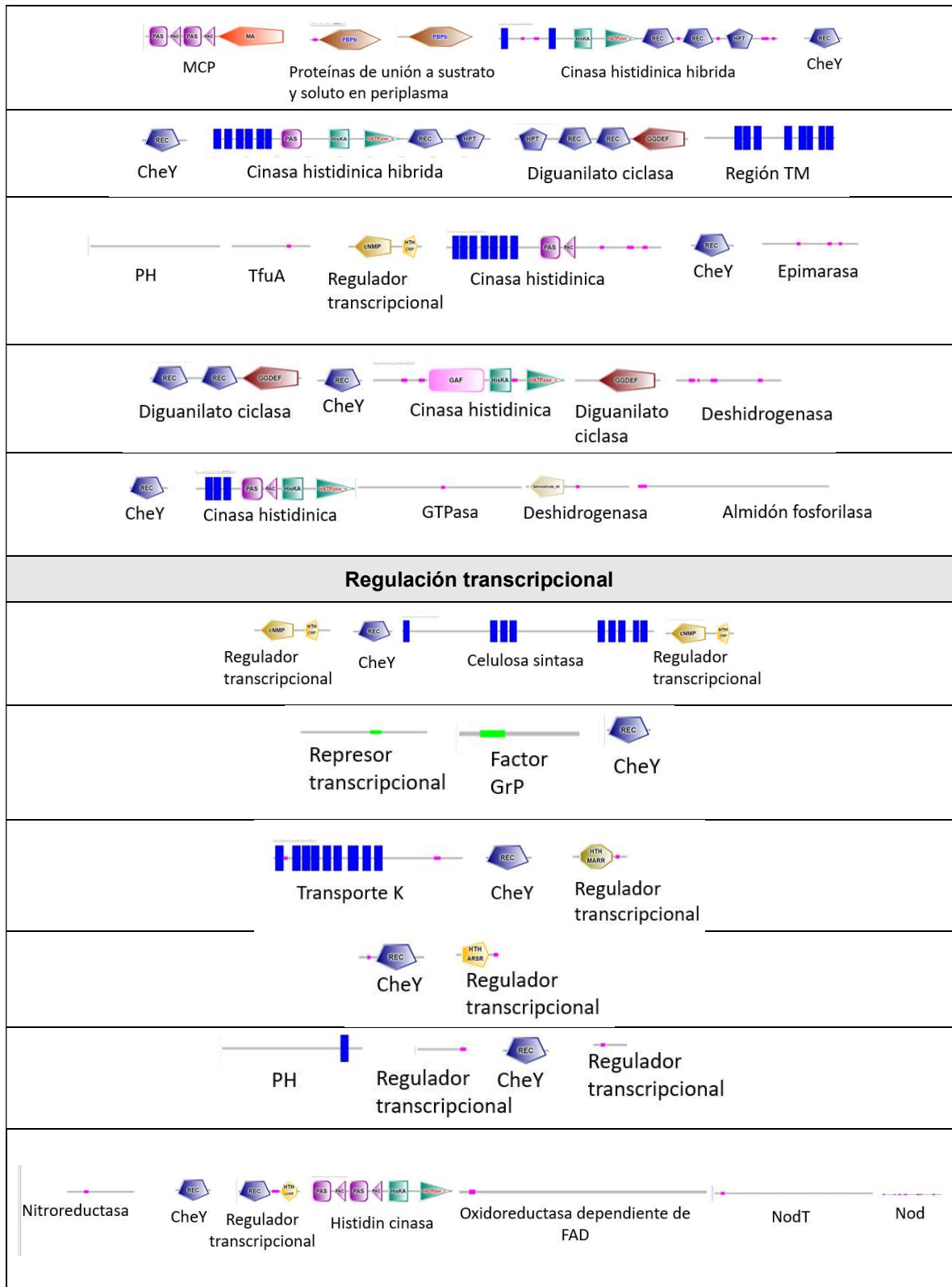
## 2. El Contexto genético de los RR de dominio único tipo CheY en una superfamilia de Sp245

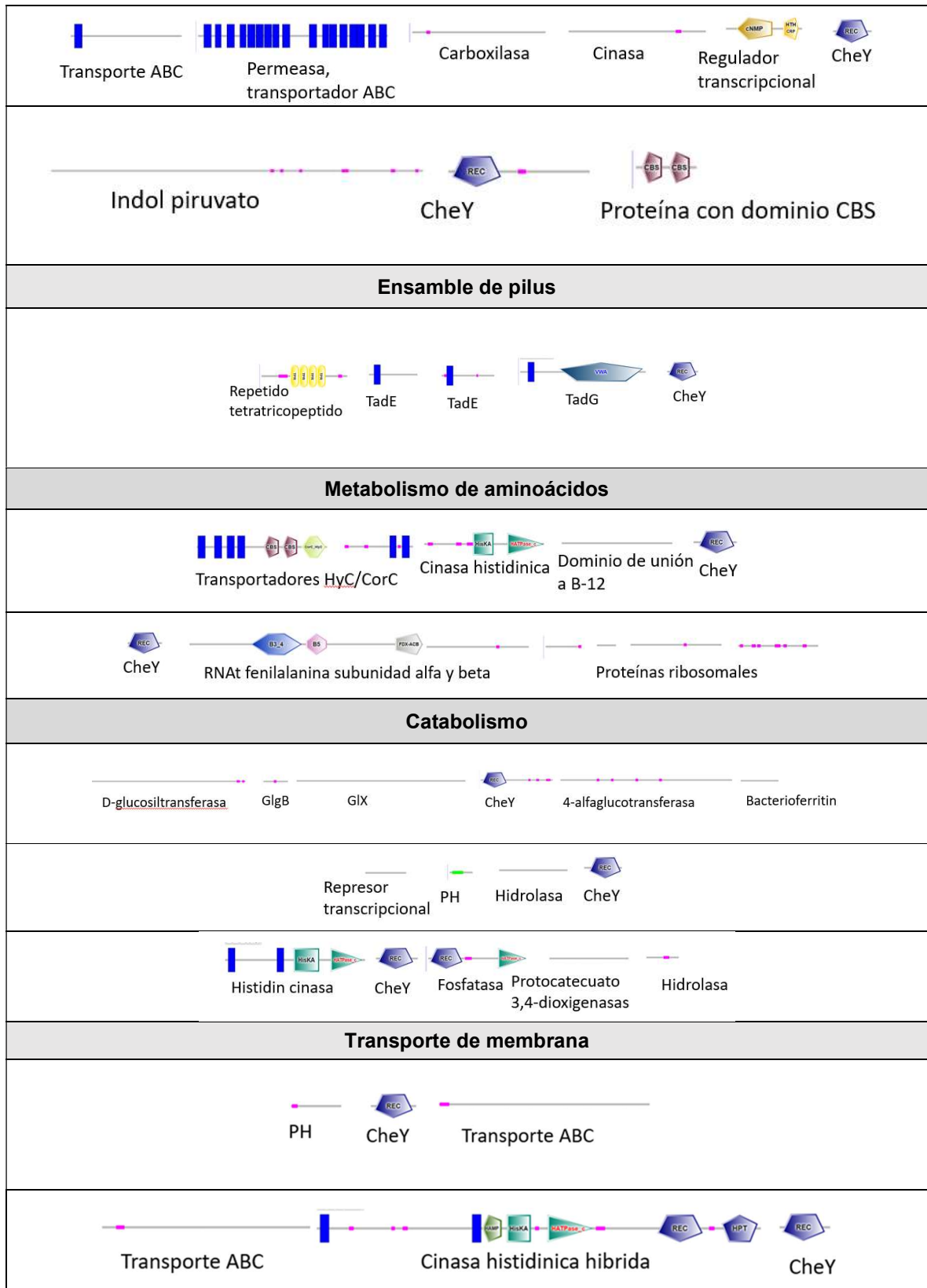
Se continuó con el análisis de reguladores de respuesta de dominio único, con el fin de establecer su distribución en los replicones y probable función, para lo cual se realizó el análisis del contexto genético de cada una de las proteínas putativas identificadas en *A. baldaniorum* Sp245. Destacando la diversidad con base al contexto genético, ya que se observó no solamente en sistemas de quimiotaxis canónicos, si no también, adyacentes a diversas proteínas involucradas en sistemas

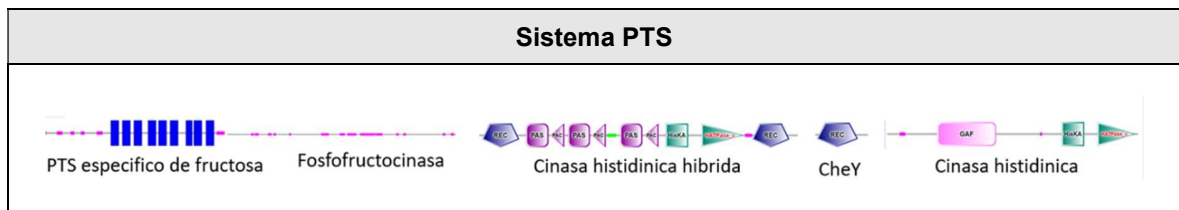
de señalización o de doble componente, pero también a proteínas de regulación transcripcional, transporte de aminoácidos, catabolismo o en sistema PTS, e incluso varios genes que se localizaron huérfanos, sorprendiendo la gran distribución de acuerdo al contexto genético ubicados en cromosoma, crómidos y plásmidos. De igual manera se localizaron los dos operones de quimiotaxis canónicos similares a *A. brasilense* Sp7 y otros dos operones asociados a sistemas de quimiotaxis de funciones celulares diversas (Tabla 5).

**Tabla 5.** Tabla del contexto genético de los reguladores de respuesta de dominio único en *A. baldaniorum* Sp245, no se agregan los genes huérfanos.





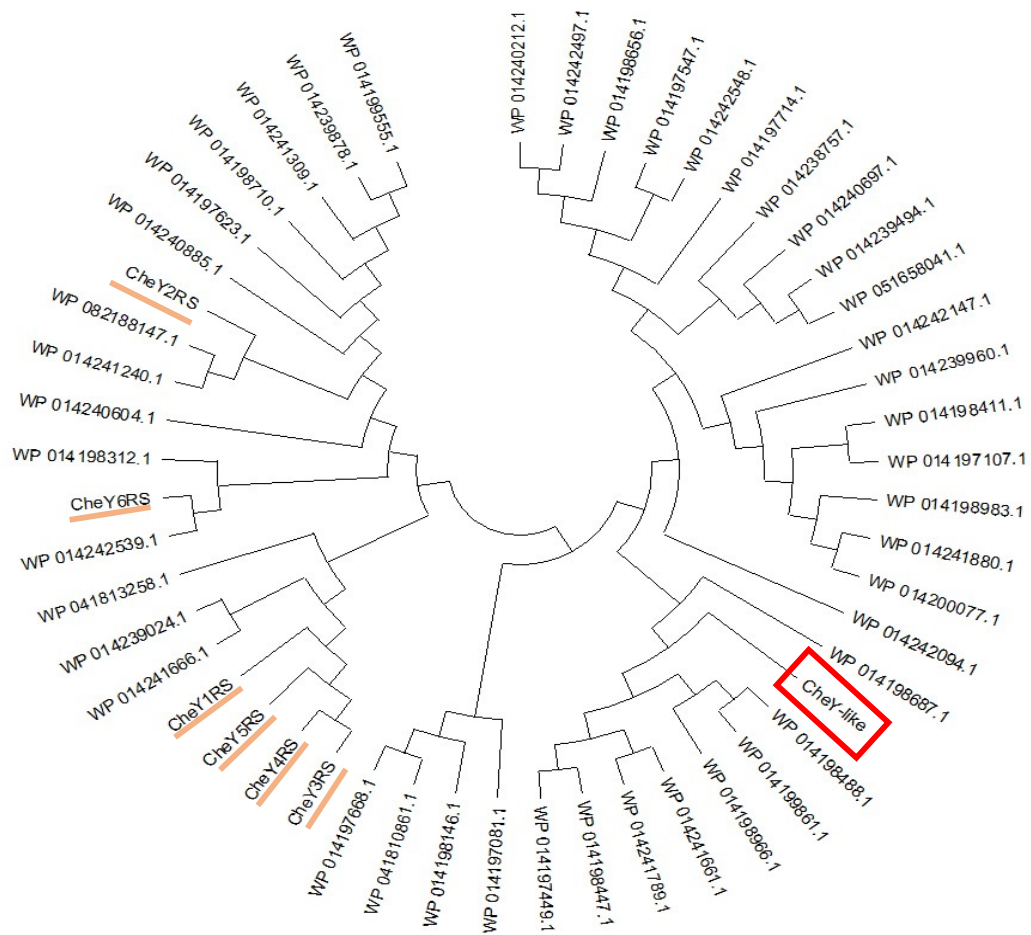




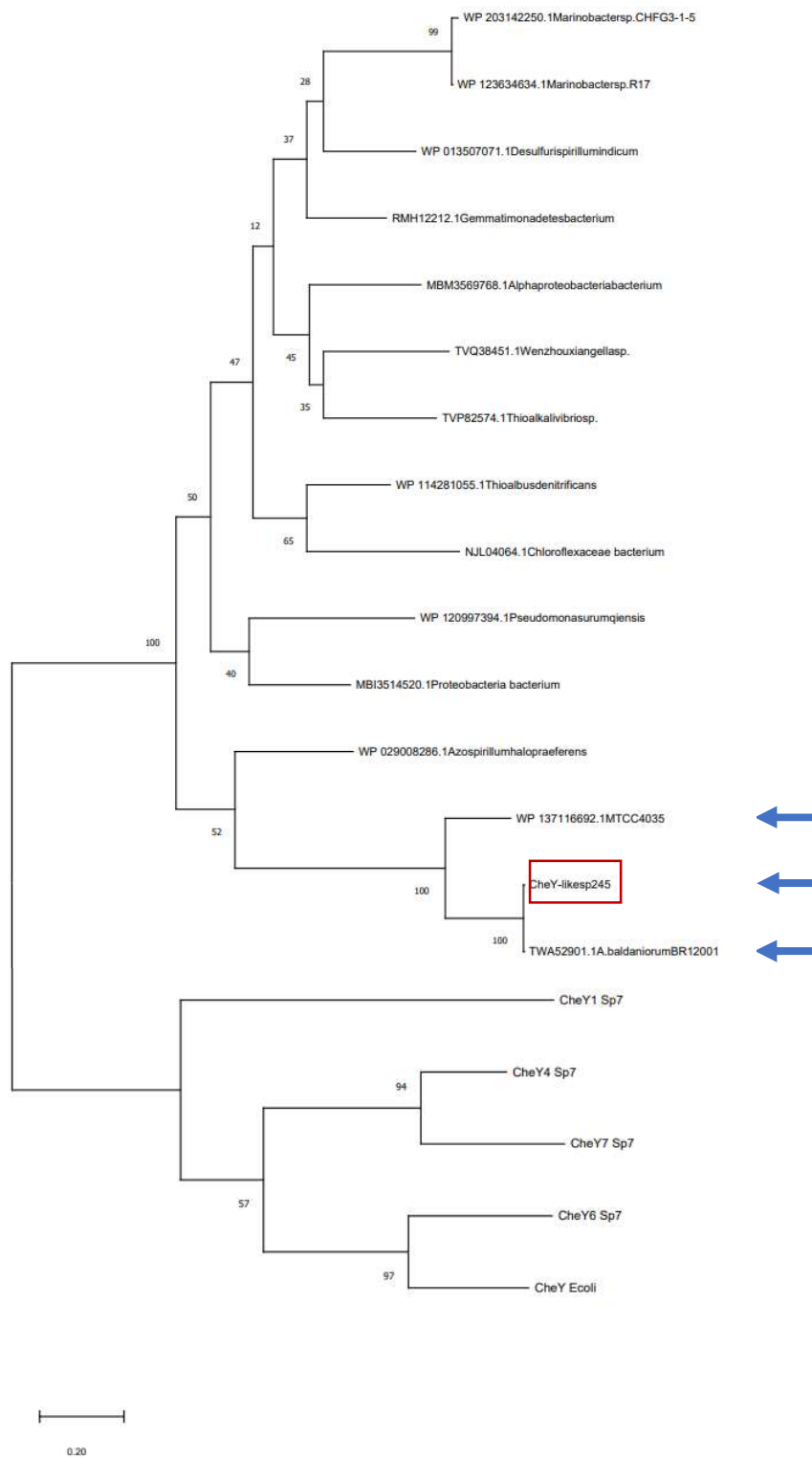
### 3. Árbol filogenético de las 45 proteínas de dominio único REC en *A. baldaniorum* Sp245

Con el fin de analizar las relaciones evolutivas entre el total de las 45 proteínas encontradas en el genoma de *A. baldaniorum* Sp245, éste se inició realizando un árbol filogenético de las 45 proteínas putativas identificadas (Figura 13) y además se incluyen las secuencias de los 6 CheY's de *Rhodobacter sphaeroides*, bacteria que habita en el ambiente acuático, relacionada filogenéticamente con el género *Azospirillum*. Al analizar el árbol, las secuencias de las proteínas se distribuyeron en diversas ramas. Se observó que las proteínas homologas de CheY1, CheY4, CheY6 y CheY7 de las que se han descrito su participación en quimiotaxis en *A. brasilense* Sp7 se encuentran en una misma rama, lo que indicaría que comparten un ancestro en común. Mientras que la proteína CheY-like que se analiza en este proyecto se ubicó en una rama diferente respecto a las proteínas CheYs de *R. sphaeroides*.

Al advertir la distancia genética entre las proteínas de *R. sphaeroides*, nos dimos a la tarea de construir un árbol filogenético con las secuencias de los resultados de la búsqueda de BlastP en NCBI, incluyendo aquellas que corresponden al > 40% de identidad y cobertura >80%. Se consideraron también las secuencias de las proteínas CheY1, CheY4, CheY6 CheY7 de *A. brasilense* Sp7 y la secuencia de la proteína CheY de *E. coli* con fines de comparación (Figura 14). Se observó en el árbol como las secuencias se distribuyen en diversas ramas. Las proteínas CheY1, CheY4, CheY6 y CheY7 de *A. brasilense* Sp7 y la proteína CheY de *E. coli* se encuentran compartiendo rama del árbol, mientras que la proteína que se analiza en este trabajo se ubica compartiendo rama con otras dos proteínas, estas proteínas corresponden a las cepas 12001 (aislada en Brasil) y MTCC4035 (aislada en la India), cepas en las que se han analizado e identificado el clúster de 5 genes que codifican presuntamente al operón de señalización conservado Figura 14.



**Figura 13. Árbol filogenético de las 45 proteínas de *A. baldaniorum* y los 6 CheY's de *R. sphaeroides*.** En cuadro rojo la proteína putativa que se estudia en el proyecto subrayadas las proteínas CheY de *R. sphaeroides*. Árbol obtenido en MEGA-X



**Figura 14. Árbol filogenético realizado con secuencias de proteínas de los resultados del BlastP. Las flechas indican las proteínas que se encuentran en el mismo arreglo de clúster de 5 genes. Árbol obtenido en MEGA-X.**

#### 4. Búsqueda de región rica en arginina en proteínas con mayor longitud en *A. baldaniorum* Sp245

Con el fin de identificar la región rica en arginina que se ha descrito en las proteínas CLE, región que une a di-GMPc para activar y unir al motor flagelar en *C. crescentus*, se realizó un alineamiento de secuencias con las proteínas que encontramos en la búsqueda inicial en *A. baldaniorum* Sp245. Se identificaron tres secuencias (WP\_014210604.1, WP\_014240885.1 y WP\_014239960.1) que poseen en su C-terminal dicha región rica en arginina.



Figura 16. Alineamiento de secuencias de las proteínas Cle de *C. crescentus* y secuencias de proteínas de longitud mayor de *A. baldaniorum* Sp245.

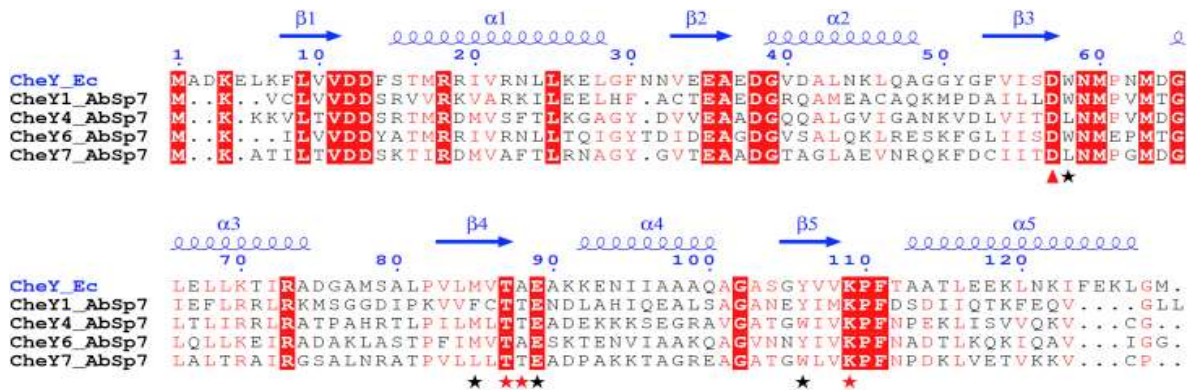
## ANÁLISIS ESTRUCTURAL

### 1. Análisis de proteína CheY-like de *A. baldaniorum* Sp245

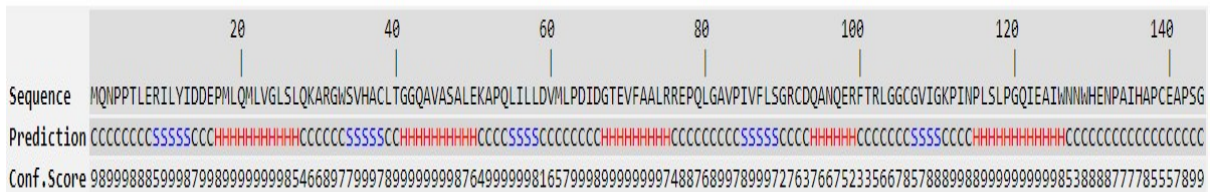
Tomando en consideración los datos obtenidos, procedimos a determinar las características estructurales de la proteína CheY-like de *A. baldaniorum* Sp245. En primera instancia se realizó un alineamiento de la secuencia proteica comparándola con proteínas funcionales de *A. brasilense* Sp7 y *E. coli*. Los resultados nos permitió constatar que en CheY-like ocurren los residuos de aminoácidos que intervienen en la recepción del grupo fosfato Figura 17 A. Así como la predicción de la estructura secundaria apoya el dato, como se muestra en la figura 17B los elementos característicos de la estructura secundaria de las proteínas CheY, están presentes como cinco alfa hélices y cinco beta-plegadas ( $\alpha_1 \beta_1 \alpha_2 \beta_2 \alpha_3 \beta_3 \alpha_4 \beta_4 \alpha_5 \beta_5$ ),

además se observa que al terminar la beta plegada 3, se encuentra el aminoácido aspartato (D) que es característica de las proteínas CheY y que se ha descrito como el aminoácido donde se lleva la fosforilación (Mukherjee et al., 2019).

**A**



**B**



**Figura 17. Dominio y estructura secundaria de la proteína CheY-like de *A. baldaniorum* Sp245. A)** Predicción de dominio de la secuencia de la proteína CheY-like en la plataforma SMART. **B)** Predicción de estructura secundaria de la secuencia de CheY-like en la plataforma I-TASSER.

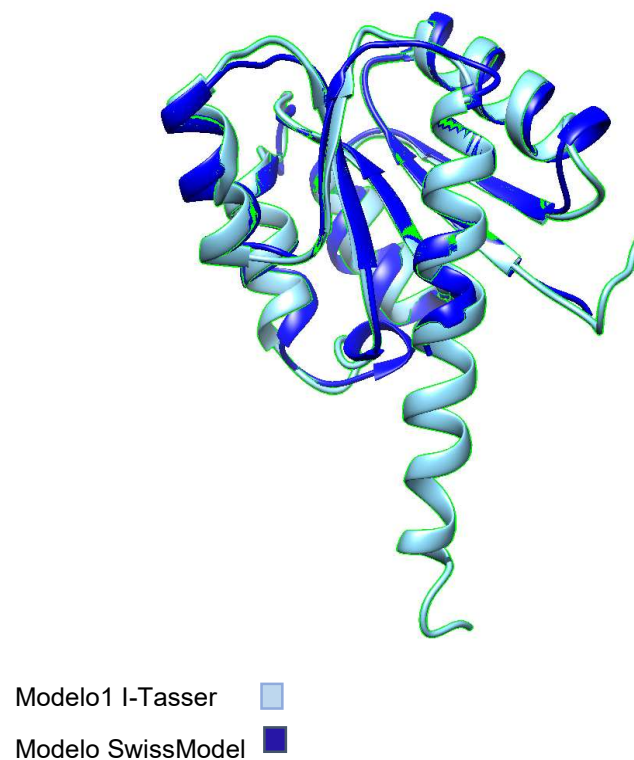
## 2. Estructura terciaria

La secuencia de la proteína CheY-like se ingresó en las plataformas I-Tasser y Swiss Model se obtuvo la predicción de su estructura tridimensional.

Empleamos la plataforma Swiss Model utilizando como plantilla para generar el modelo por homología el cristal 5e3j de *Acinetobacter baumannii* (PDB DOI: 10.2210/pdb5E3J), el dominio REC del regulador transcripcional de la proteína RstA, observando la coincidencia del análisis previo de la estructura secundaria, es decir, que el modelo conserva las cinco alfa hélices y las cinco beta-plegadas, así como la secuencia:  $\alpha_1 \beta_1 \alpha_2 \beta_2 \alpha_3 \beta_3 \alpha_4 \beta_4 \alpha_5 \beta_5$  figura 17.

En I-Tasser se obtiene el modelo, el cual muestra el plegamiento de las cinco alfas hélices y de las cinco beta-plegadas que concuerda con el análisis previo de estructura secundaria, así como el modelo obtenido en Swiss Model.

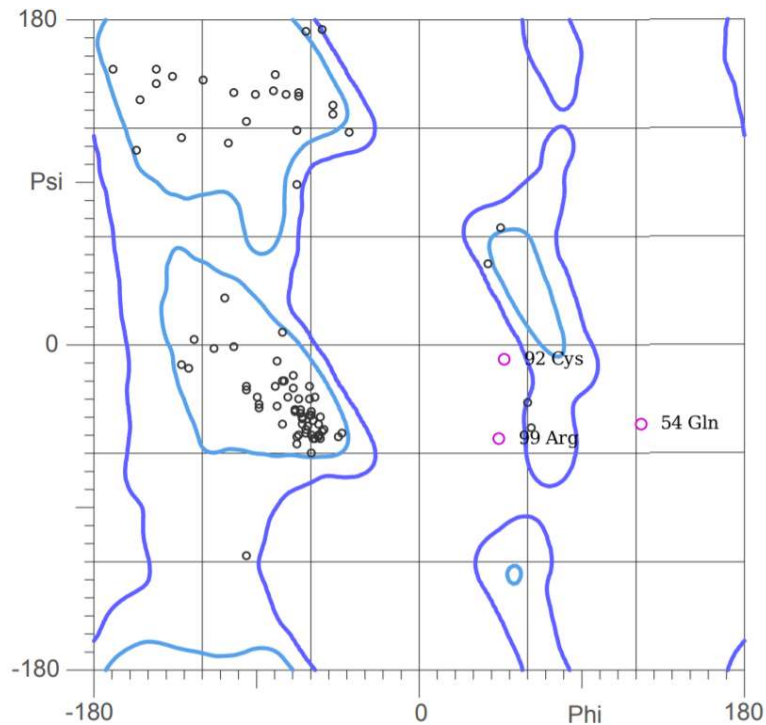
Sin embargo, al construir el empalme de ambos modelos (Fig.18) se observan diferencias. Una de ellas la más notable correspondía a la alfa hélice 5 de CheY-like de *A. baldaniorum* ya que es mayor al modelo de proteína de *Acinetobacter baumannii* (5e3j). Por lo cual seleccionamos el modelo de I-Tasser para análisis posteriores, por los parámetros obtenidos, los mismos que son más cercanos a la referencia.



**Figura 18. Modelos tridimensionales de la proteína CheY-like.** En azul claro el modelo 1 generado en I-Tasser teniendo valores de C-score: 0.19; TM-SCORE:0.74 +-0.11. En azul marino el modelo de Swiss Model con valores de Q MEAN-2.13; % de identidad del 27.7 con la plantilla.

Para evaluar el modelo 1 obtenido de I-Tasser, con base en la geometría de los aminoácidos que ocurren en la proteína, se realizó el gráfico de Ramachandran en la plataforma Molprobity, la cual utiliza como referencia los aminoácidos de proteínas ya reportadas y que se encuentran en la plataforma PDB. En la figura 18

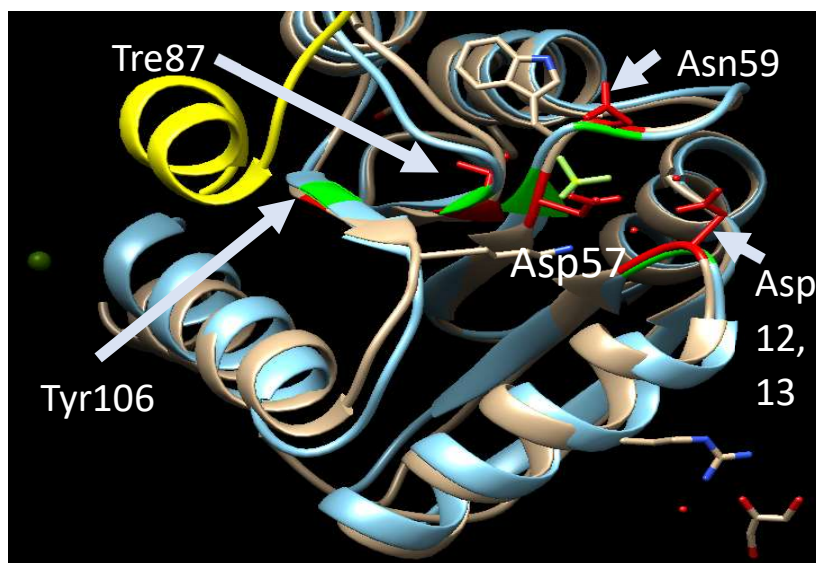
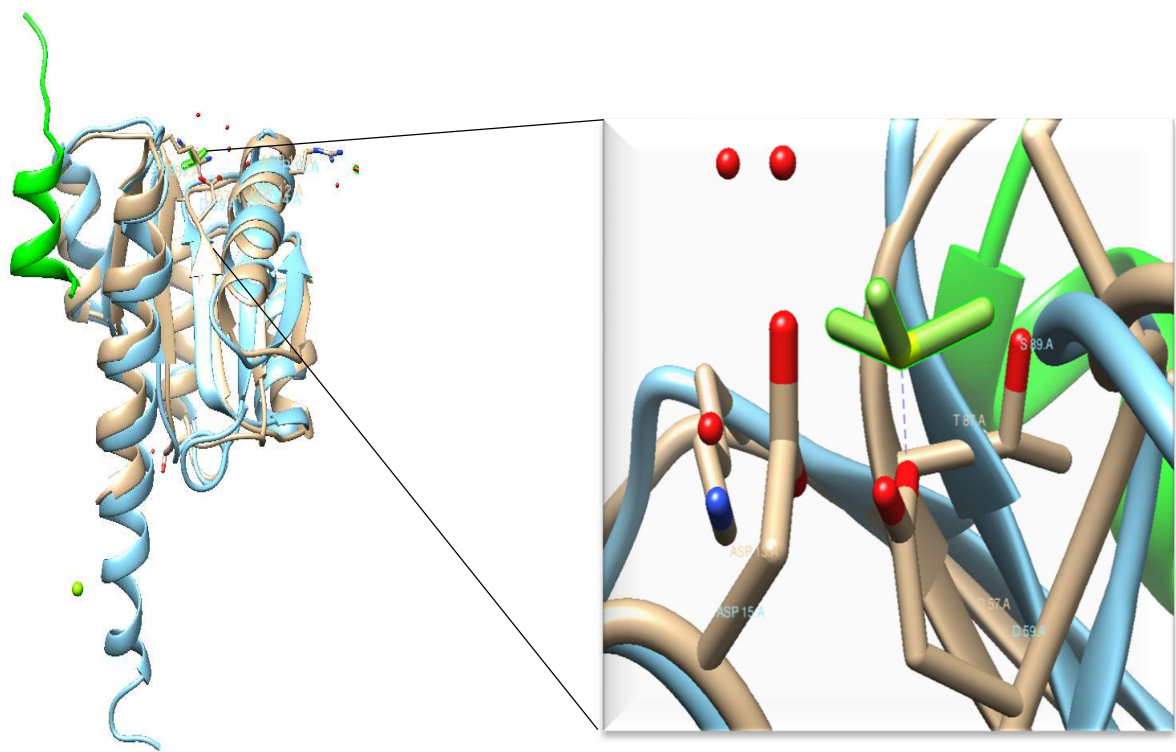
se observa que el 85% de los aminoácidos se ubican dentro de los parámetros indicados por el programa bioinformático, en donde se indica que la mayoría de los residuos de aminoácidos se localizan de manera correcta en la región favorable.



**Figura 18.** Gráfica de Ramachandran obteniendo un 85% de índice de Ramachandran. Obtenida de la plataforma Molprobit.

### 3. Análisis del modelo de CheY-like con la estructura de CheY de *E. coli*

De los reguladores de respuesta de dominio único cuyo papel está establecido de manera experimental por técnicas bioquímicas y de biofísica, es la proteína CheY de *E. coli*. Estudios que apoyan la función de modular la movilidad y quimiotaxis al interaccionar con el flagelo en *E. coli* (Hazelbauer *et al.*, 2008). La proteína se cristalizó con el análogo del grupo fosfato BeF3, y ligado al péptido de 16 aminoácidos de la proteína FliM (Lee *et al.*, 2001), cuyo código en el programa PDB es 1f4v. En el empalme con la estructura 1f4v (Figura 19) se observó que los aminoácidos que participan en la fosforilación de la proteína CheY de *E. coli* se encuentran altamente conservados en la proteína CheY-like, como se constató en el alineamiento inicial descrito en la Fig.17. Sin embargo, no se localizaron en el alineamiento los aminoácidos que participan en el cambio conformacional estructural, que intervienen en la unión con el análogo del grupo fosfato, ni en la interacción con FliM de CheY *E. coli* (Figura 19).



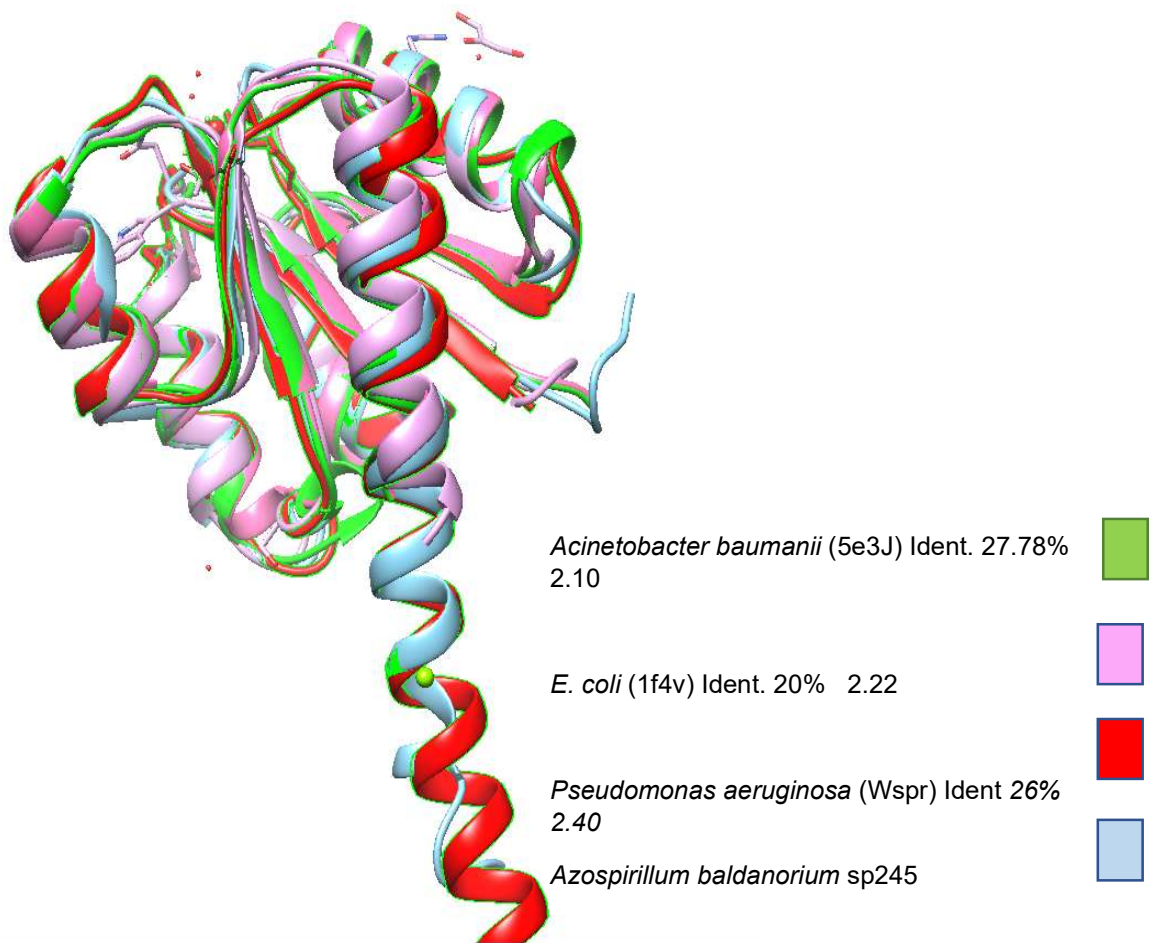
CheY <i>E. coli</i>	CheY-like <i>A. baldaniorum</i> Sp245
Asp57	Asp59
Asp12	Asp14
Asp13	Asp15
Tre87	Ser89
Asn59	Met61
Tyr106	Val108

**Figura 19. Empalme de la estructura if4v y el modelo de CheY-like.** Alineamiento de las secuencias de las proteínas CheY *E. coli* (Café claro) y CheY-like *A. baldaniorum* Sp245 (Azul). Dentro la tabla en la columna color rojo los aminoácidos que participan en la fosforilación Aspartato (D) posición 12, 13 y 59. Aminoácidos que participan en la modificación estructural de CheY *E. coli* treonina 87 (T) y tirosina 106 (Y). La estructura tridimensional visualizada y modificada en Chimera y Pymol. Alineamiento realizado en Clustal Omega.

#### 4. Análisis con otras estructuras de dominio REC

Tomando en consideración los datos obtenidos a nivel estructurales en donde se constató discrepancias entre los modelos a nivel de estructura tridimensional obtenido de las proteínas CheY de *E. coli* y CheY-like de *A. baumannii* consideramos pertinente buscar en el programa PDB cristales de diversos dominios REC, ya sea de proteínas con reguladores de respuesta de dominio único, de reguladores transcripcionales o de reguladores de respuesta con actividad enzimática como la diguanilato ciclasa WspR de *P. aeruginosa* (PDB número de acceso 3BRE, De *et al.*, 2008), *A. baumannii* 5e3j (PDB DOI: 10.2210/pdb5E3J), *E. coli* 1f4v (PDB DOI: 10.2210/pdb1F4V), con los cuales se realizó el empalme con dichas estructuras (Fig. 20).

Se destaca en este análisis el plegamiento característico de las cinco alfas hélices y cinco beta- plegadas, aún con la falta de identidad significativa de la secuencia peptídica, entre las diferentes estructuras primarias de las proteínas. De manera notable se reveló en el empalme la alfa hélice cinco de la proteína CheY-like de acuerdo con la estructura tridimensional, que ésta es similar a la alfa hélice de la proteína diguanilato ciclasa WspR de *P. aeruginosa* figura 20. En general el dominio REC presentó similitud a pesar de que éste está en las proteínas asociado a otros dominios, además de que las bacterias pertenecen a la familia de las gamma proteobacterias y géneros diferentes.



**Figura 20.** Empalme de múltiples estructuras de dominio REC.

## DISCUSION

La mutagénesis sitio dirigida es una herramienta fundamental para la correlación de las funciones celulares con regiones específicas de DNA y verificación de funciones predichas por análisis bioinformático (Wen *et al.*, 2014). Se realizan mutantes sin afectar el marco de lectura, sin la inserción disruptiva por un casete de resistencia. Sin embargo, la detección de eventos de doble entrecruzamiento se considera tardíos y laboriosos debido a la falta de métodos eficaces de contra-selección, por lo que el uso de un marcador contra-seleccionable, puede mejorar significativamente esta metodología (Sant'Anna *et al.*, 2011). Debido a los datos anteriores, se logró obtener un plásmido mutagénico que funciona para realizar la mutante por delección de una manera eficaz al poder utilizar diversos marcadores tanto de selección como de contra-selección. En este trabajo sólo se analizó la selección negativa por crecimiento en sacarosa, ya que se obtienen merodiploides fácilmente por la selección de la resistencia a Tc. Cordero en el 2017 nos sugiere que la presencia de TraJ (regulador de la expresión de los genes tra), mejora la transferencia por movilización del plásmido construido en este trabajo, en relación a vectores, que no lo contienen.

En este trabajo se identificaron de manera sorpresiva un total de 45 reguladores de respuesta de dominio único en *A. baldaniorum* Sp245 y 41 en *A. brasilense* Sp 7 (Tabla 4). Estos reguladores de respuesta de dominio único (SD-RR) ocurren en los tres dominios de la vida: bacterias, arqueas y eucariotas. Entre las bacterias, están presentes en los genomas secuenciados de todos los filos bacterianos excepto *Chlamydia*, *Fusobacteria* y *Tenericutes* (Mollicutes) (Jenal & Galperin 2011). Como se ha descrito en otros géneros bacterianos, el número de SD-RR se correlacionó con el tamaño del genoma, ya que el genoma de *A. baldaniorum* contiene 7.86 Mpb por lo tanto, la regla de que a mayor tamaño, mayor número total de SD-RR codificados, se confirmó en nuestro estudio. Mientras que el genoma de *R. sphaeroides* bacteria que pertenece también a la familia *Rhodospirillaceae* de hábitat acuático, solo contiene un total de 14 SD-RR; de esos, 6 han sido designados CheY's basados en su nivel de similitud a CheY de *E. coli*, su proximidad a otros *loci* de quimiotaxis y definidos experimentalmente como involucrados en movilidad y quimiotaxis (Porter *et al.*, 2006). Curiosamente de manera similar a *Myxococcus xanthus*, organismo cuyo estilo de vida es complejo con función de agregación y diferenciación celular basada en la motilidad superficial. Las células de *M. xanthus* carecen de flagelos estando equipadas con dos motores distintos, ocho sistemas quimio detectores similares al canónico Che y un número récord de 42 SD-RR (Jenal & Galperin 2011) de manera similar a lo encontrado en este estudio. Por otra parte, algunos patógenos importantes, como *Shigella dysenteriae* y *Mycobacterium leprae*, no codifican ningún SD-RR, a pesar de la longitud de sus secuencias genómicas relativamente grandes (4.5 y 3.27 Mpb, respectivamente). Por lo que, se podría especular que además del tamaño del

genoma, también es importante considerar los estilos de vida de las bacterias y el hábitat en donde se desarrollan.

En vista de que la búsqueda en el genoma de *A. baldaniorum* Sp245 mostró que ocurren dos sistemas canónicos de quimiotaxis homólogos a los estudiados en *A. brasilense* Sp 7 (Bible *et al.*, 2012; Mukherjee *et al.*, 2016, 2019) y de *R. sphaeroides* (Porter *et al.*, 2008), se realizó un árbol filogenético incluyendo las secuencias de los genes descritos. Se localizaron los genes homólogos que codifican para CheY de *Azospirillum* Sp7, CheY3 y CheY4, de igual manera los caracterizados en *R. sphaeroides* CheY1-CheY6, género del grupo que pertenece a hidrobacterias de la familia de las Rhodospirillaceae, un clado de procariontes que se originó en ambientes acuáticos, descrito como un genoma con bajo nivel de genes obtenidos por mecanismos de transferencia horizontal (HGT) (Wisniewski-Dyé *et al.* 2011). Los datos obtenidos indican las relaciones evolutivas entre los SD-RR encontrados Fig 13, que muestran un gran número de ramas indicando la divergencia de la evolución de las secuencias. Las bacterias del género *Azospirillum* se asocian con la raíz de las plantas terrestres, sin embargo, prácticamente todos sus parientes cercanos son acuáticos. La mayoría de los genes de mantenimiento de *Azospirillum* son ortólogos en sus parientes acuáticos cercanos, mientras que la mitad de su genoma ha sido obtenido de organismos terrestres. En contraste a los genomas de sus parientes cercanos, *A. baldaniorum* Sp245 posee siete replicones. Replicones con diferentes características y que no son usuales en proteobacterias indicando un potencial para la plasticidad del genoma (Wisniewski-Dyé *et al.* 2011), como se constató en nuestro estudio en lo que se refiere al gen que codifica a CheY-like.

Los genes ancestrales usualmente codifican funciones “house-keeping” y son conservados a lo largo de distancias evolutivas. Bajo esta premisa se realizó un segundo árbol incluyendo secuencias de CheY’s de varias bacterias. Los datos obtenidos nos mostraron que CheY-like se ubica en una rama diferente de los CheY’s homólogos de *A. brasilense* Sp7, *E. coli* y de *R. sphaeroides* (Fig.14) implicados en quimiotaxis y movilidad (Wisniewski-Dye *et al.* 2011). Se localiza en una rama que agrupa las otras cepas secuenciadas a partir de los aislados obtenidos en Brasil y la India y que contienen la misma región genómica del “cluster” de cinco genes y únicamente una secuencia correspondiente al genoma de *Azospirillum halopraeferens* bacteria aislada de un pasto forrajero de Pakistán (Reinhold *et al.*, 1987). Por lo que, como se ha sugerido se considera un gen adquirido por herencia horizontal, con funciones de adaptación a la rizosfera, un ambiente rico en aminoácidos, carbohidratos, iones inorgánicos y metabolitos secundarios exudados de las raíces de las plantas (Wisniewski-Dyé *et al.* 2011).

Así mismo, se identificaron tres secuencias que de acuerdo con el análisis realizado en este trabajo se incluyen en SD-RR, secuencias proteicas que en la región C-terminal poseen un motivo rico en aminoácidos de arginina (Figura 16), motivo característico de una nueva familia de proteínas CheYlike (denominadas CLE), identificadas y caracterizadas en *C. crescentus*, las cuales modulan la actividad

flagelar (Nesper *et al*, 2017). Esto en respuesta a la unión del segundo mensajero di-GMPc al C-terminal, en esta conformación de unión al di-GMPc, les permiten a las proteínas Cle interactuar con el motor de flagelo para modular la actividad motora. Lo que les habilita adoptar una función particular e interferir con la quimiotaxis y promover el rápido anclaje a la superficie de las células móviles (Figura 4). El sitio de unión del di-GMPc de las proteínas CLE se localizó a 28 aminoácidos de su dominio receptor conservado. Esta secuencia contiene un repetido en tándem (Y/F)XGPXR/K)R (Nesper *et al*, 2017). El arreglo en tándem de los motivos de arginina conservados junto con la afinidad de unión y estequiometría determinada para CleD, sugiere que todas las proteínas Cle se unen al dímero de di-GMPc y que los residuos conservados de arginina juegan un papel importante en el proceso de la transición del estado móvil al sedentario (Figura 4). Este análisis y la presencia en el genoma de *A. baldaniorum* sugiere que esos homólogos de las proteínas Cle, podrían eventualmente tener una función similar, lo que abre nuevas vías de estudio en nuestro modelo de trabajo.

Como se describió en la introducción las proteínas CheY muestran una versatilidad funcional, que sugiere una adaptación a los nichos ecológicos de las bacterias. Funciones que requieren que CheY sea activado por fosforilación, cómo se constata en el alineamiento mostrado en la Fig. 17, en CheY-Like están presentes los motivos requeridos para esta modificación. De manera interesante en el contexto genómico de la región de señalización se localizan colindantes dos genes que presuntivamente codifican para dos cinasas histidínicas Figura 6, que eventualmente realizarían esta función. Es interesante especular cuál sería la función de CheY-like~ P.

Por otro lado, en la identificación de vías de señalización que controlan la dispersión de la biopelícula en *Vibrio cholerae* identificaron un componente de la motilidad. CheY3, es uno de los cinco RR específicos de quimiotaxis. Notablemente, CheY3 es el único gen *cheY* homólogo requerido para quimiotaxis. La mutante de *cheY3* muestra un máximo de formación de biopelícula similar a la cepa silvestre, así como de biomasa, sin embargo, 21% de la biomasa permanece a las 16 hrs indicando un defecto en la dispersión de la biopelícula. Posteriormente realizan un análisis midiendo las frecuencias de giro y velocidades de nado de la cepa silvestre y  $\Delta cheY3$  observando notables diferencias, la cepa silvestre, gira una vez cada 1 segundo, la mutante  $\Delta cheY3$  gira una vez cada 40 segundos, estos resultados indican que una baja frecuencias de giro en la mutante es responsable del defecto en la dispersión de la biopelícula. Se propone que, si las células no pueden cambiar frecuentemente su dirección de motor, podrían quedar atrapadas por las mallas de la matriz de la biopelícula que compromete su capacidad de escape (Andrew *et al*, 2020).

También se ha reportado otra función alterna a quimiotaxis e involucrada en la formación de biopelícula de la proteína CheY en *Shewanella oneidensis*, el caso del regulador CheY3, el cual participa en el núcleo del sistema de quimiotaxis y que se

une al rotor del flagelo para modificar la dirección de nado. Además se ha descrito que fosforilado también juega un papel esencial durante la biogénesis de la biopelícula asociada a superficies sólidas. Además de manera interesante se describe que la fosforilación de CheY3 es independiente de CheA3 durante este proceso. Se mostró que la CheY3 interactúa con al menos 4 diguanilatos ciclasas y con MxdA, el efector de di-GMPc, probablemente disparando la síntesis de exopolisacáridos por la maquinaria de Mxd. Estos resultados respaldan la propuesta de que la formación de biopelícula involucrarían a reguladores CheY en otras bacterias (Boyeldieu A. *et al*, 2020). Con los datos anteriores se sugiere que múltiples CheY's no solamente intervienen en promover los sesgos rotacionales del motor flagelar y quimiotaxis, sino también para promover comportamientos que dependen sobre la movilidad tales como "swarming", adherencia a superficies, así como a cambios de estilos de vida, como la contribución de CheY4 y CheY6 de *A. brasilense* Sp7 que participan en la formación de biopelícula, adherencia y colonización a las raíces de trigo (Ganusova *et al.*, 2021).

En el análisis estructural de la proteína CheY-like, se identifican aminoácidos conservados D59, D14 Y D15 que se han descrito que participan en la fosforilación respecto a CheY de *E. coli* (Fig. 19). Este RR de quimiotaxis se considera un modelo para el mecanismo de la activación mediado por la cinasa. Estudios bioquímicos, genéticos y estructurales sobre CheY indican que la fosforilación induce a un cambio estructural de una conformación inactiva a activa. Cinco aminoácidos conservados son importantes para fosforilación y cambio conformacional subsecuentemente a la fosforilación. El residuo D57 se determinó como el sitio de fosforilación, mientras que D12 y D13 son requeridos para la unión del cofactor magnesio. En el caso de T87 y K109 que en *E. coli* intervienen en eventos post-fosforilación así como las posiciones T87 y Y106 sufren movimientos coordinados que permiten las uniones H (Lee *et al.*, 2001), estos últimos aminoácidos no se encontraron conservados en CheY-like de *A. baldaniorum* Sp245, en su lugar se identificaron los aminoácidos S89, M81 y V108. figura 19.

Estudios realizados en *R. sphaeroides*, en los que se exploró la contribución de las proteínas CheY's que controlan la rotación del motor flagelar. Con tal fin, se realizaron mutaciones puntuales en los genes *cheYs*. Estas fueron en los residuos de aminoácidos equivalentes a los de CheY de *E. coli*. Las mutaciones en el sitio de fosforilación D57A y D57N se construyeron con el fin de observar inhibición de la fosforilación, mientras D13K, X95V y Y106W con el fin de activar la conformación de los CheY's.

Las cepas que contenían los alelos *cheY3* (D10K), *cheY4* (D10K), *cheY6* (D10K) o *cheY6* (D10K) corresponden a fenocopias de sus correspondientes mutantes  $\Delta cheYs$ . Lo que indicó que, las mutaciones similares a D13K en *E. coli* causaron que las proteínas CheY3, CheY4 y CheY6 no funcionen y que por lo tanto se requiere la unión al cofactor  $Mg^{+2}$  de forma estricta. Esto contrasta con *E. coli*, donde

la mutante CheY (D13K) causa un fenotipo muy diferente (rotación CW) a la mutante por  $\Delta cheY$  (rotación en sentido CCW), (Porter *et al.*, 2006).

El rol del residuo en la posición 106W de los CheY3, CheY4 y CheY6 V104W mostró que, estas mutaciones no alteraron la quimiotaxis, la frecuencia de interrupción o la fosforilación *in vitro*. Esto sugiere que a diferencia de CheY de *E. coli*, los CheY's de *R. sphaeroides* son tolerantes las sustituciones en la posición 106 (Porter *et al.*, 2006).

Con estos resultados se demuestra que no todas las proteínas CheYs de los modelos estudiados experimentan cambios conformacionales idénticos tras la fosforilación ya que las mutaciones que activan a CheY de *E. coli* no funcionan de manera similar las CheY's de *R. sphaeroides*.

De tal manera, que fue importante en este estudio lograr la mutación por eliminación del gen  $\Delta cheY-L$  (con la colaboración de la QFB Ma. Luisa Xiqui), Figura 13. Lo que en estudios posteriores a este permitirá dilucidar la participación del gen en la formación de biopelícula y movilidad.

## Conclusiones

- Se identificaron 41 genes *cheY* que codifican presuntivamente para proteínas reguladores de respuesta de dominio único REC en *A. brasilense* Sp7 y 45 en *A. baldaniorum* Sp245.
- La proteína que se investiga en este proyecto tiene conservado los aminoácidos que participan en la fosforilación.
- Se concluye que el gen *cheY* que codifica a la proteína CheY fue obtenida por transferencia horizontal de genes.
- En *A. baldaniorum* Sp245 ocurren tres genes que presentan homología con proteínas CLE efectoras del di-GMPc

## Perspectivas

- Es importante caracterizar a los RR de dominio único en *A. baldaniorum* Sp245, ya que están presentes genes en el genoma que presuntamente codifican para estas proteínas, por lo tanto, resulta interesante analizar el fenotipo de la mutación *A. baldaniorum* Sp245  $\Delta$ *cheY*.
- En el análisis comparativo entre la proteína CheY de *E. coli* y CheY-like *A. baldaniorum* Sp245 se lograron identificar los aminoácidos que se han descrito en CheY *E. coli* que participan en la fosforilación, por lo tanto, purificar la proteína CheY-like y realizar la reacción de fosforilación determinaría el mecanismo de activación de la proteína.
- Los sistemas de señalización son esenciales para las bacterias para mantener su homeostasis y adaptarse a cambios del medio en el que se encuentre, por lo que realizar ensayos de pull-down para determinar la interacción con la proteína FliM del flagelo polar y DgcE contribuiría a comprender los mecanismos de regulación en los que está involucrada la proteína CheY-like.
- Recientes estudios han reportado una nueva clase de proteínas CLE en *C. crescentus*, las cuales también poseen como dominio único el REC, estas proteínas se caracterizan por poseer una región rica en arginina en su extremo carboxilo terminal, por lo tanto, caracterizar los genes encontrados en el genoma de *A. baldaniorum* Sp245 que presuntamente codifican a las proteínas que contienen la región rica Rs como las proteínas CLE sería interesante para comprender y continuar a estudiar su relación con otras bacterias.

## Bibliografía

- Baikalov I, Schroder I, Kaczor-Grzeskowiak M, Grzeskowiak K, Gunsalus RP, Dickerson RE. (1996) Estructura del regulador de respuesta de *Escherichia coli* NarL. *Bioquímica*.
- Bible AN, Khalsa-Moyers GK, Mukherjee T, Green CS, Mishra P, Purcell A, Aksenova A, Hurst GB, Alexandre G. (2015) Metabolic adaptations of *Azospirillum brasilense* to oxygen stress by cell-to-cell clumping and flocculation. *Appl Environ Microbiol* 81:8346–8357. doi:10.1128/AEM.02782-15.
- Belyakov, A., Burygin, G., Arbatsky, N., Shashkov, A. et al. (2012) Identification of an O-linked repetitive glycan chain of the polar flagellum flagellin of *Azospirillum brasilense* Sp7. *Carbohydr. Res.* 361C, 127–132. doi: 10.1016/j.carres.2012.08.019
- Borland, S., Oudart, A., Prigent-Combaret, C., Brochier-Armanet, C., & Wisniewski-Dyé, F. (2015). Genome-wide survey of two-component signal transduction systems in the plant growth-promoting bacterium *Azospirillum*. *BMC Genomics*, 16(1), 833
- Boyer M, Haurat J, Samain S, Segurens B, Gavory F, et al. (2008) Bacteriophage prevalence in the genus *Azospirillum* and analysis of the first genome sequence of an *Azospirillum brasilense* integrative phage. *Appl Environ Microbiol* 74: 861–874
- Boyeldieu, A., Chaouche, A.A., B, M., Honoré, F.A., Méjean, V., and Cécile Jourlin-Castelli, C. (2020). The phosphorylated regulator of chemotaxis is crucial throughout biofilm biogenesis in *Shewanella oneidensis*. *Biofilms and Microbiomes* 6:54. doi.org/10.1038/s41522-020-00165-5
- Bridges, A.A., Fei, C., and Bassler, B.L. (2020). Identification of signaling pathways, matrix-digestion enzymes, and motility components controlling *Vibrio cholerae* biofilm dispersal. *PNAS* 2020. 117: 32639–32647. .pnas.org/cgi/doi/10.1073/pnas.2021166117
- Carlos A. Barba (2014) Los sistemas de dos componentes circuitos moleculares versátiles. *Revista especializada en Ciencias Químico-Biológicas*, 17(1):62-76.
- Caro-Quintero A, Deng J, Auchtung J, Brettar I, Hofle MG, et al. (2011) Unprecedented levels of horizontal gene transfer among spatially co-occurring *Shewanella* bacteria from the Baltic Sea. *ISME J* 5: 131–140
- Croes, C., Moens, S., Bastelaere, E., Vanderleyden, J., and Michiels, K. (1993) Polar flagellum mediates *Azospirillum brasilense* adsorption to wheat roots. *Microbiology* 139, 2261–2269. doi: 10.1099/00221287-139-9-2261
- Del Campo AM, Ballado T, de la Mora J, Poggio S, Camarena L, Dreyfus G. (2007) El control quimiotáctico de los dos sistemas flagelares de *Rhodobacter sphaeroides* está mediado por diferentes conjuntos de proteínas CheY y FlIM. *J. Bacteriol.* ; 189 : 8397–8401
- Phosphorylation-Independent Regulation of the Diguanylate Cyclase WspR Nabanita De1 Michelle Pirruccello. (2008). Petya Violinova Krasteva1 Narae Bae1Rahul Veera Raghavan1 Holger Sondermann1\**PLoS Biol.* 6: e67-e67 DOI: 10.1371/journal.pbio.0060067

Galperin MY.( 2006) Clasificación estructural de reguladores de la respuesta bacteriana: diversidad de dominios de salida y combinaciones de dominios. *J. Bacteriol.* 188 : 4169–4182.

Ganusova, E. E., Vo, L. T., Abraham, P. E., O'Neal Yoder, L., Hettich, R. L., and Alexandre, G. (2021). The *Azospirillum brasilense* core chemotaxis proteins CheA1 and CheA4 link chemotaxis signaling with nitrogen metabolism. *mSystems* 6, e01354–e01420. doi: 10.1128/mSystems.01354-2

Gao R, Mack TR, Stock AM. ( 2007) Bacterial response regulators: versatile regulatory strategies from common domains. *Trends Biochem. Sci.*; 32 : 225-234

Gordon, V. D., and Wang, L. (2019). Bacterial mechanosensing: the force will be with you, always. *J. Cell Sci.* 132:jcs227694. doi: 10.1242/jcs.227694

Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, et al. (2007) Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science* 316: 1307–1312

Gumerov VM, Ortega DR, Adebali O, Ulrich LE y Zhulin IB (2020) *Nucleic Acid Research* 48: D459 – D464.

Greer-Phillips, S. E., Stephens, B. B., and Alexandre, G. (2004) An energy taxis transducer promotes root colonization by *Azospirillum brasilense*. *J. Bacteriol.* 186, 6595–6604. doi: 10.1128/JB.186.19.6595-6604.200

Guttenplan, S. B., and Kearns, D. B. (2013) Regulation of flagellar motility during biofilm formation. *FEMS Microbiol. Rev.* 37, 849–871. doi: 10.1111/1574-6976.12018

Hazelbauer GL, Falke JJ, Parkinson JS.( 2008). Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* 33:9 –19. <https://doi.org/10.1016/j.tibs.2007.09.014>.

Jung, K., Fried, L., Behr, S., & Heermann, R. (2012). Histidine kinases and response regulators in networks. *Current Opinion In Microbiology*, 15, 118–124.

Lee, S. Y., Cho, H.S., Pelton, J.G., Yan,D., Henderson, R.K., King, D.S., Huang,L., Kustu, S., Berry E.A. and Wemmer, D.E. (2001). Crystal structure of an activated response regulator bound to its target. *Nature Structural Biol.* 8: 52-56

Moens, S., Schloter, M., and Vanderleyden, J. (1996). Expression of the structural gene, *laf1*, encoding the flagellin of the lateral flagella in *Azospirillum brasilense* Sp7. *J. Bacteriol.* 178, 5017–5019. doi: 10.1128/jb.178.16. 5017-5019.1996

Mukherjee T, Kumar D, Burriss N, Xie Z, Alexandre G. (2016). *Azospirillum brasilense* chemotaxis depends on two signaling pathways regulating distinct motility parameters. *J Bacteriol* 198:1764 –1772. doi:10.1128/JB.00020-16.

Mukherjee *et al.*, (2019), Multiple CheY Homologs Control Swimming Reversals and Transient Pauses in *Azospirillum brasilense*, *Biophysical Journal* <https://doi.org/10.1016/j.bpj.2019.03.006>

Nesper, *et al.* (2017). Cyclic di-GMP differentially tunes a bacterial flagellar motor through a novel class of CheY-like regulators. *eLife*, 6, e28842.

- O'Neal, L., Gullett, J. M., Aksenova, A., Hubler, A., Briegel, A., Ortega, D., *et al.* (2019). Distinct chemotaxis protein paralogs assemble into chemoreceptor signaling arrays to coordinate signaling output. *mBio* 10, e01757–e01719. doi: 10.1128/mBio.01757-19
- O'Neal, L., Vo, L., and Alexandre, G. (2020). Specific root exudate compounds sensed by dedicated chemoreceptors shape *Azospirillum brasilense* chemotaxis in the rhizosphere. *Appl. Environ. Microbiol.* 86, e01026–e01020. doi: 10.1128/AEM.01026-20
- Porter, S. L., Wadhams, G.H. Armitage. J. P. (2006). The CheYs of *Rhodobacter sphaeroides*. *J. Biol. Chem.* 281:32694–32704.
- Porter SL, Wadhams GH, Armitage JP. (2008). *Rhodobacter sphaeroides*: complexity in chemotactic signalling. *Trends Microbiol.* :251-60. doi: 10.1016/j.tim.2008.02.006. Epub 2008 Apr 25. PMID: 18440816.
- Kern D, Volkman BF, Luginbuhl P, Nohaile MJ, Kustu S, Wemmer DE. Estructura de un interruptor fosforilado transitoriamente en la transducción de señales bacterianas. *Naturaleza.* 1999; 402 : 894–898.
- Sant'Anna, F. H., Andrade, D. S., Trentini, D. B., Weber, S. S., & Schrank, I. S. (2011). Tools for genetic manipulation of the plant growth-promoting bacterium *Azospirillum amazonense*. *BMC Microbiology*, 11(1), 1.
- Steenhoudt O, Vanderleyden J (2000) *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiol Rev* 24: 487–506.
- Stock AM, Robinson VL, Goudreau PN. (2000) Transducción de señales de dos componentes. *Annu. Rev. Biochem.* ; 69 : 183–215
- Viruega-Góngora, V. I., Acatitla-Jácome, I. S., Reyes-Carmona, S. R., Baca, B. E., and Ramírez-Mata, A. (2020). Spatio-temporal formation of biofilms and extracellular matrix analysis in *Azospirillum brasilense*. *FEMS Microbiol. Lett.* 367:fnaa037. doi: 10.1093/femsle/fnaa037
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296-W303 (2018).
- Wen, Q., Liu, X., Wang, H., & Lin, J. (2014). A versatile and efficient markerless gene disruption system for *Acidithiobacillus thiooxidans*: application for characterizing a copper tolerance related multicopper oxidase gene. *Environmental microbiology*, 16(11), 3499-3514.
- Wisniewski-Dyé F, Borziak K, Khalsa-Moyers G, Alexandre G, Sukharnikov LO, Wuichet K, *et al.* (2011) *Azospirillum* genomes reveal transition of bacteria from aquatic to terrestrial environments. *PLoS Genet*;7:e102430. <http://dx.doi.org/10.1371/journal.pgen.1002430>
- Zhulin, I. B., and Armitage, J. P. (1993). Motility, chemokinesis, and methylation-independent chemotaxis in *Azospirillum brasilense*. *J. Bacteriol.* 175, 952–958. doi: 10.1128/JB.175.4.952-958.1993