

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE CIENCIAS MARINAS
INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS



Diferencias genéticas en el genoma del cloroplasto de *Dunaliella salina* SQ y sus implicaciones en la creación de sistemas de expresión de proteínas recombinantes.

T E S I S

QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS
PARA OBTENER EL GRADO DE

**DOCTOR EN CIENCIAS EN ECOLOGÍA MOLECULAR y
BIOTECNOLOGÍA**

PRESENTA

HAYDEE LOPEZ RODRIGUEZ

ENSENADA, BAJA CALIFORNIA, MEXICO. OCTUBRE 2017

Resumen

Las microalgas son organismos de interés biotecnológicos, ya que producen una amplia gama de metabolitos que pueden ser utilizados en diferentes ramas de la industria. En las últimas dos décadas ha surgido interés en este grupo como fabricas para la producción de proteínas recombinantes, principalmente expresión en el cloroplasto. En este trabajo se utilizó a la microalga *Dunaliella salina* para formular la creación de distintos casetes de expresión para la producción de proteínas recombinantes en el cloroplasto. Uno de los requisitos para la creación de sistemas de expresión, es tener un marcador que sea efectivo y nos permita discernir colonias que contienen el transen de las silvestres, en el caso de *D. salina*, se sabe que naturalmente es resistente a diferentes antibióticos, razón por la cual es difícil encontrar un marcador de selección apropiado. En base a los datos obtenidos en este trabajo, se propone utilizar como marcador de selección el herbicida norflurazon, ya que los resultados obtenidos muestran que a una concentración de 5 μ M es capaz de inhibir completamente el crecimiento de esta microalga. La estructura y diseño de los casetes de expresión realizados en este trabajo se basaron en la secuencia reportada del genoma del cloroplasto de la cepa de *D. salina* CCAP 19/18 proveniente de Australia. Se sabe que la expresión de trasngenes en cloroplasto depende del uso de regiones reguladoras específicas para cada especie. En este trabajo se utilizó una cepa aislada de San Quintín, Baja California, México y se asumió que no existía diferencia en el genoma de ambas cepas, sin embargo, después de realizar la secuenciación y anotación del genoma del cloroplasto de esta cepa, se obtuvo que existe una cobertura del 62% con un 96% de identidad, encontrándose la mayor parte de las diferencias en regiones no codificantes. Estas regiones son importantes en la creación de los sistemas de expresión en cloroplasto ya que de esto depende la expresión de los transgenes, por lo que la información generada en el presente estudio podrá ser utilizada para la creación de sistemas de expresión en el cloroplasto de la cepa *D. salina* SQ.

FACULTAD DE CIENCIAS MARINAS
INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS
POSGRADO EN ECOLOGÍA MOLECULAR y BIOTECNOLOGÍA

DIFERENCIAS GENÉTICAS EN EL GENOMA DEL CLOROPLASTO
DE *Dunaliella salina* SQ Y SUS IMPLICACIONES EN LA CREACIÓN
SISTEMAS DE EXPRESIÓN DE PROTEÍNAS RECOMBINANTES.

T E S I S

QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS PARA
OBTENER EL GRADO DE

DOCTOR EN CIENCIAS

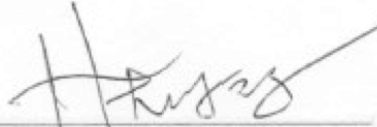
PRESENTA

HAYDEE LOPEZ RODRIGUEZ

Aprobada por:



Dr. José Luis Stephano Hornedo
Director de tesis


Dra. Ivone Giffard Mena
Sinodal
Dra. Sawako Hori-Oshima
Sinodal
Dr. Gerardo Enrique Medina Basulto
Sinodal
Dra. Edna Sánchez Castrejón
Sinodal

Agradecimientos

Al Dr. Stephano por haber confiado en mí para realizar este trabajo, por enseñarme tantas cosas, sin las cuales no estaría donde estoy ahora.

A mi comité, Dra. Ivone Giffard Mena, Dra. Sawako Hori, Dr. Gerardo Medina Basulto y la Dra. Edna Sánchez Castrejón, por sus consejos y enseñanzas.

Al Dr. Luis Manuel Enríquez Paredes, por ser un gran consejero y por su ayuda para poder terminar este grado, creo que a veces él tenía más confianza en mí que la que yo misma me tenía. Muchas gracias.

Al CONACyt por haberme otorgado la beca para poder realizar este trabajo.

A Dante.

A todos mis compañeros de laboratorio por haber sufrido juntos a causa de “La Dunaliella”.

Al Dr. Julio V. Suarez por su apoyo con el trabajo de esta tesis.

Al Dr. Juan Guillermo Vaca Rodríguez y a la Dra. Lus Mercedes López Acuña por haber ayuda a mi permanencia en el posgrado.

A Angélica por su ayuda en todo el papeleo y por siempre atendernos con una sonrisa.

Y finalmente a mi familia por el gran apoyo que han sido durante toda mi vida, pero en especial a mi madre, que me dejó antes de tiempo y no llegó a ver este momento que tanto anhelaba, no es lo mismo si ella.

Contenido

Resumen.....	I
Agradecimientos	III
Lista de tablas	VII
Lista de figuras	VIII
Prólogo	1
 CAPÍTULO I Construcciones de casetes para la expresión de proteínas recombinantes en el cloroplasto de <i>D. salina</i> SQ.....	 9
1. Introducción.....	10
1.1 Ingeniería de cloroplastos en microalgas.....	10
1.2. Promotores y regiones no traducidas (UTR's).....	12
1.3. Marcadores de selección	14
1.4. Optimización de codones	15
1.5. Métodos de transformación	15
1.6. Producción de proteínas terapéuticas en el cloroplasto de microalgas	16
2. Hipótesis	18
3. Objetivo general.....	18
3.1 Objetivos específicos.....	18
4. Metodología	19
4.1 Construcción 1.....	19
4.1.1. Síntesis de promotor y UTR's.....	19
4.1.2 Clonación del marcador de selección <i>nat-1</i> en el plásmido pUC57-utr's	20
4.1.3. Clonación del gen que codifica para Interleucina-2 en el plásmido pUC57-utr's.....	22
4.1.4 Unión de las construcciones IL-2 y Nat-1	24
4.1.5 Clonación de las regiones de recombinación homólogas en el plásmido pIL2-Nat	26
4.2. Construcción 2.....	29
4.2.1 Síntesis del casete de expresión	30
4.2.2 Clonación de la región de recombinación homóloga 5' en el plásmido pEGFP	31
4.2.3 Clonación de la región de recombinación homóloga 3' en el plásmido pGFP5	32
4.3. Transformaciones	34
4.3.1. Transformación del plásmido pDsCh	34
4.3.2. Transformación del plásmido pDsEGFP.....	36
4.3.3. Detección de transformantes de <i>D. salina</i>	37
5. Resultados y discusiones	39
5.1. Síntesis de regiones reguladoras	39
5.2. Clonación del marcador de selección <i>nat-1</i> en el plásmido pUC57-utr's.....	39
5.3. Clonación del gen que codifica para Interleucina-2 en el plásmido pUC57-utr's.....	40
5.4. Unión de las construcciones IL-2 y Nat-1	43

5.5.	<i>Clonación de las regiones de recombinación homólogas en el plásmido pIL2-Nat</i>	46
5.6.	<i>Clonación de la región 5' homóloga en el plásmido pEGFP</i>	51
5.7.	<i>Clonación de la región 3' homóloga en el vector pGFP5</i>	52
5.8.	<i>Transformación del plásmido pDsCh</i>	55
5.9.	<i>Transformación del plásmido pDsEGFP</i>	58
6.	Conclusiones	61
CAPÍTULO II Norflurazon como agente de selección en la microalga <i>D. salina</i> SQ		62
1.	Introducción	63
1.1.	<i>Antibióticos como marcadores de selección en <i>D. salina</i></i>	63
1.2.	<i>Herbicidas como marcadores de selección en <i>D. salina</i></i>	64
2.	Hipótesis	68
3.	Objetivo	68
3.1.	<i>Objetivos particulares</i>	68
4.	Metodología	69
4.1.	<i>Susceptibilidad de <i>D. salina</i> SQ al herbicida norflurazon</i>	69
4.2.	<i>Ligación del gen pds F128V en el plásmido pDsEGFP</i>	70
4.3.	<i>Transformación con el plásmido pGFP/PDS</i>	72
5.	Resultados y discusión	73
5.1.	<i>Resistencia de <i>D. salina</i> SQ al herbicida norflurazon</i>	73
5.2.	<i>Ligación del gen pdsF128V en el plásmido pDsEGFP</i>	75
5.3.	<i>Transformación con el vector pGFP/PDS</i>	79
6.	Conclusiones	84
CAPÍTULO III Anotación y análisis del genoma del cloroplasto de <i>D. salina</i> SQ		85
1.	Introducción	86
1.1.	<i>Estructura de los genomas de cloroplasto</i>	86
1.2.	<i>Aplicaciones de los genomas de cloroplasto</i>	87
1.3.	<i>Genoma del cloroplasto de otras cepas de <i>D. salina</i></i>	88
2.	Hipótesis	93
3.	Objetivo	93
3.1.	<i>Objetivos particulares</i>	93
4.	Metodología	94
4.1.	<i>Recolección del espécimen</i>	94
4.2.	<i>Extracción de DNA de cloroplastos</i>	94
4.3.	<i>Secuenciación</i>	95
4.4.	<i>Anotación y visualización del genoma</i>	95
5.	Resultados y discusión	97

5.1. Organización del genoma del cloroplasto	97
5.2. Estructuras repetitivas.....	104
5.3. Frecuencia relativa de uso de codones.....	106
5.4. Comparación de promotores	108
6. Conclusiones.....	114
Conclusiones finales	115
Perspectivas	117
Referencias	120
Anexos	129
Descripción de plásmidos	129
Primers	130
Secuencia del vector pDsCh.....	131
Secuencia del vector pDsEGFP	133
Secuencia del vector pGFP/PDS.....	135

Lista de tablas

Tabla	Descripción	Página
Tabla I	Frecuencia de uso de codones de <i>D. salina</i> CCAP 19/18	41
Tabla II	Características de los genomas de cloroplasto de las cepas de <i>D. salina</i>	99
Tabla III	Genes presentes en el genoma del cloroplasto de <i>D. salina</i> SQ	101
Tabla IV	Diferencia de intrones entre las cepas de <i>D. salina</i>	102
Tabla V	Descripción de plásmidos	129
Tabla VI	Lista de primers	130

Lista de figuras

Figura	Pie de figura	Página
1	Ventajas de la utilización de las microalgas como sistemas de expresión de proteínas recombinantes.	2
2	Morfología de <i>D. salina</i> .	4
3	Representación esquemática de la integración de los casetes de expresión de proteínas recombinantes en cloroplasto.	12
4	Casete para la expresión de interleucina-2 y el gen de resistencia <i>nat-1</i> en el cloroplasto de <i>D. salina</i> .	19
5	Construcción il-2-NatR. Unión del gen recombinante con el marcador de selección.	24
6	Casete de expresión de la proteína EGFP y el gen <i>nat-1</i> optimizados para cloroplasto.	31
7	Secuencia de promotor, 5' UTR y 3' UTR del gen <i>atpA</i>	39
8	Electroforesis de la digestión para comprobar la clonación del gen <i>nat-1</i> en el plásmido pUC57-utr's.	40
9	Secuencia del gen il-2 sintetizado.	42
10	Electroforesis de la digestión el gen <i>il-2</i> del plásmido pUC57-il2.	42
11	Electroforesis de la PCR para comprobar la clonación del gen <i>il-2</i> en el plásmido pUC57-utr's.	43
12	Electroforesis de la digestión de plásmido pNat-utr con las enzimas <i>Nde</i> I y <i>BstX</i> I.	44
13	Electroforesis de la PCR para agregar los sitios de restricción de las enzimas <i>Nde</i> I y <i>BstX</i> I al gen <i>il-2</i> con promotor y UTR's.	44
14	Electroforesis de la digestión del plásmido de las clonas transformadas para comprobar la presencia del inserto.	45
15	Amplificación de las regiones homólogas 5' y 3'.	46
16	Digestión del plásmido pIL2-Nat para recombinar la región 3' homóloga.	47
17	Digestión con la enzima <i>BstX</i> I para comprobar la recombinación de la región 3' homóloga.	48
18	Digestión con la enzima <i>Nru</i> I para comprobar la recombinación de la región 5' homóloga.	50
19	Casete de expresión del plásmido pDsCh.	50
20	Mapa del plásmido pDsCh.	51
21	Digestión con la enzima <i>EcoR</i> V para comprobar la unión de la región 5' homóloga.	53
22	Digestión con las enzimas <i>Nru</i> I y <i>EcoR</i> I para comprobar la unión de la región 3' homóloga.	54
23	Mapa del plásmido pDsEGFP.	55
24	Amplificación con los primers <i>Atpa2F/Atpa2R</i> y <i>atpaF/atpaR</i> .	56
25	Amplificación con los primers <i>NatF/NatR</i> .	57
26	Nueva amplificación con los primers <i>NatF/NatR</i> .	58

27	Clonas transformadas con el plásmido pDsEGFP.	59
28	Amplificación con los primers InF/InR.	59
29	Ruta de síntesis de carotenos en <i>D. salina</i> .	66
30	Diseño de la placa de cultivo para comprobar la susceptibilidad de <i>D. salina</i> al herbicida norflurazon.	70
31	Efecto de diferentes concentraciones del herbicida norflurazon en el crecimiento de <i>D. salina</i> .	74
32	Amplificación del gen <i>pdsF128V</i> y digestión del vector pDsEGFP con las enzimas <i>Dra</i> III/ <i>Afe</i> I.	76
33	Digestión con la enzima <i>Mfe</i> I.	77
34	Secuencia de aminoácidos de la enzima Fitoeno desaturasa F128V	78
35	Mapa de vector pGFP/PDS.	78
36	Representación esquemática de la integración de transgenes en cloroplasto mediante recombinación homóloga.	88
37	Mapa del genoma de <i>Dunaliella salina</i> CCAP 19/18.	90
38	Lista de los genes de cloroplasto encontrados en los contigs generados en la secuenciación del genoma de la cepa <i>D. salina</i> CONC-001.	92
39	Mapa del genoma de <i>Dunaliella salina</i> SQ.	98
40	Alineamiento de los genomas de las cepas de <i>D. salina</i> CCAP 19/18 y SQ.	100
41	Comparación de la estructura de los genomas de <i>D. salina</i> SQ (arriba) y <i>D. salina</i> CCAP 19/18 (abajo)	103
42	Comparación de la posición de los límites de las regiones LSC, Ira, IRb y SCC entre las cepas de <i>D. salina</i> CCAP 19/18 y SQ.	104
43	Representación gráfica del número y tamaño de las regiones repetitivas en los cloroplastos de las cepas de <i>D. salina</i> .	105
44	Representación gráfica de la frecuencia de uso de codones de ambas cepas de <i>D. salina</i> .	107
45	Alineamiento de la región del promotor-5'UTR del gen <i>rbcL</i> de las cepas de <i>D. salina</i> SQ y CCAP 19/18.	109
46	Alineamiento de la región del promotor-5'UTR del gen <i>psbA</i> de las cepas de <i>D. salina</i> SQ y CCAP 19/18.	110
47	Estructura secundaria del 5' UTR del gen <i>psbA</i> de las cepas SQ (izquierda) y CCAP 19/18 (derecha).	111
48	Alineamiento de la región 3'UTR del gen <i>psbA</i> de las cepas de <i>D. salina</i> SQ y CCAP 19/18.	112
49	Alineamiento de la región 3'UTR del gen <i>atpA</i> de las cepas de <i>D. salina</i> SQ y CCAP 19/18.	113
50	Propuesta de nuevos casetes de expresión para lograr la expresión de proteínas recombinantes en el cloroplasto de <i>D. salina</i> SQ.	118

Prólogo

Las microalgas son un grupo polifilético que incluye organismos eucariontes y procariontes (cianobacterias, también conocidas anteriormente como algas verde-azules) que se caracterizan por ser fotosintéticos fijadores de CO₂. Este grupo es de gran interés para la industria biotecnológica dado que son fuente de una gran variedad de metabolitos, entre los que se incluyen clorofilas, β -carotenos y ácidos grasos, principalmente. Las microalgas también son utilizadas en la industria papelera, en la formulación de cosméticos, así como suplementos alimenticios y terapéuticos (principalmente por sus carbohidratos, β -glucanos) (Brasil et al, 2017). Especialmente en las últimas décadas se han considerado como una alternativa para la producción de biocombustibles (biogás, biodiesel, biohidrógeno y bioetanol) (Wijffels y Barbosa, 2010, Koller et al, 2014); así como una potencial plataforma para la producción de proteínas recombinantes (Specht et al, 2010).

Las microalgas como biorreactores ofrecen ventajas sobre otros sistemas de expresión de proteínas recombinantes eucariotes, ya que tienen la capacidad de duplicar su biomasa en 24 h y pueden ser expresadas en núcleo, cloroplasto o mitocondria (Fig. 1). A diferencia de las bacterias, las microalgas son capaces de realizar modificaciones post-traduccionales, por lo cual pueden ser utilizadas para la producción de proteínas glucosiladas (Gong et al, 2011). Finalmente, algunas microalgas verdes son consideradas como organismos GRAS (Generally

Recognized as Safe, por sus siglas en inglés), lo cual significa que son aptas para consumo humano, y por lo tanto podrían ser utilizadas como una fuente de administración oral de proteínas terapéuticas, disminuyendo o incluso eliminando los pasos de la purificación de las proteínas (Mayfield et al, 2007).

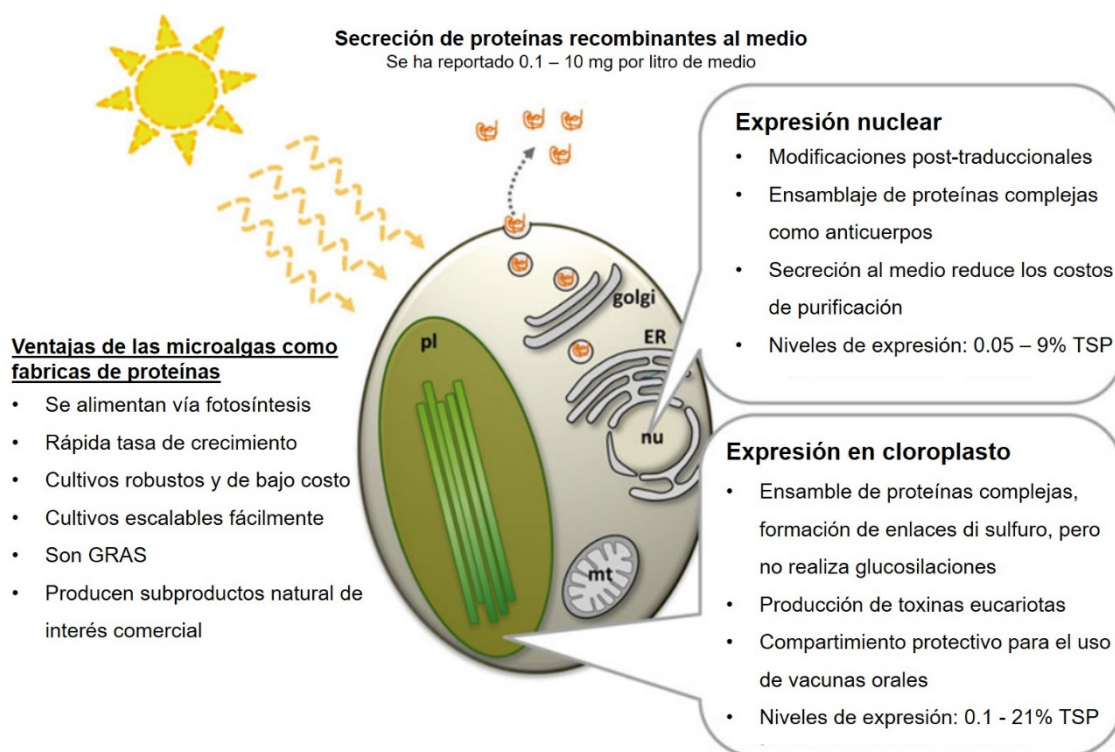


Figura 1. Ventajas de la utilización de las microalgas como sistemas de expresión de proteínas recombinantes. *ER* retículo endoplásmico, *mt* mitocondria, *nu* núcleo, *pl* cloroplasto. Tomado y modificado de Hempel y Maier, 2016.

Otra de las ventajas de la producción de proteínas recombinantes en microalgas es la reducción de costos. Estudios basados en la producción de anticuerpos recombinantes muestran que el costo por gramo de un anticuerpo funcional es \$150, \$ 0.05 y \$ 0.002 US dólares, producido en células de mamífero, planta y microalgas, respectivamente (Mayfield et al, 2003; Potvin y Zhang, 2010).

Los cloroplastos son los orgánulos preferidos para la expresión de proteínas recombinantes, ya que presentan altos niveles de expresión y a diferencia de la transfección nuclear es posible dirigir la inserción de secuencias de DNA por recombinación homóloga. En estos compartimientos se ha logrado obtener una cantidad de proteína recombinante que equivale al 20.9 ± 1.2 % de la concentración total de proteínas solubles (TSP; total soluble proteins) (Surzycki et al, 2009) aunque la mayoría reporta niveles del 0.16 – 5 % de TSP (Gong et al, 2011). *Chlamydomonas reinhardtii* es uno de las microalgas más estudiadas para la producción de proteínas recombinantes, incluso en su cloroplasto se han expresado diferentes proteínas de interés biotecnológico (Hempel y Maier, 2016). Otras microalgas en las que se han expresado proteínas recombinantes en el cloroplasto incluyen *Haematococcus pluvialis* (Gutiérrez et al, 2012) y *Dunaliella tertiolecta* (Georgianna et al, 2013). Sin embargo, para el caso particular de *Dunaliella salina*, a la fecha no se ha reportado un sistema de expresión a nivel del cloroplasto.

La microalga verde *D. salina* es uno de los organismos eucarióticos más halotolerantes que existen en el planeta, puede crecer en salinidades desde 0.05 M hasta 5 M (Ben-Amotz y Avron, 1990) (Fig. 2). Carece de pared celular rígida, cuenta con una membrana celular elástica cubierta de una superficie mucosa y es capaz de acumular hasta un 50 % de su peso seco de glicerol (Ben-Amotz y Avron 1980, 1990). Estas características aunadas a la presencia de bombas de sodio en la membrana celular que eliminan el sodio intracelular (Lee 2008), le permiten resistir los cambios osmóticos que ocurren en ambientes hipersalinos (Hosseini Tafreshi y Shariati, 2009).

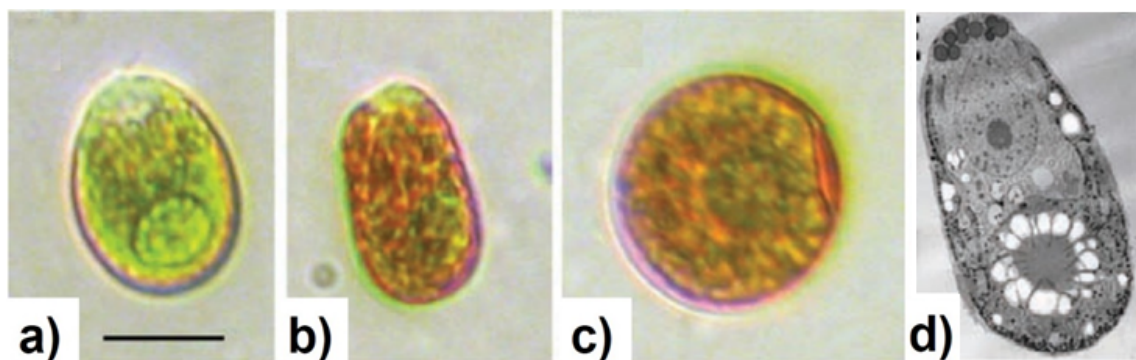


Figura 2. Morfología de *D. salina*. a) Célula vegetativa motil en 1.5 M NaCl; b) célula vegetativa motil en 4.5 M NaCl; c) célula vegetativa no motil en 4.5 M NaCl; d) microfotografía. La barra representa 10 μm . Tomada y modificada de W. Polle et al, 2009.

En la actualidad esta microalga es la principal fuente comercial que existe de β -carotenos, estos pigmentos los almacena en el cloroplasto y pueden llegar a representar hasta un 14 % de su peso seco (Hosseini Tafreshi y Shariati, 2009).

D. salina también ha sido utilizada para la elaboración de biodiesel a partir de microalgas (Yao et al, 2016) y es considerada un buen candidato para la producción de proteínas recombinantes (Geng et al., 2003). El uso de *D. salina* como sistema de expresión de proteínas recombinantes ofrece varias ventajas; por ser un organismo halófilo puede ser cultivado a altas concentraciones de NaCl, lo que reduce la contaminación por otros organismos; es fotosintética y los requerimientos de nutrientes son mínimos, lo que hace que el medio donde se cultiva sea barato. A diferencia de *C. reinhardtii*, que es una microalga de agua dulce, *D. salina* crece en agua de mar, por lo cual no se requiere de agua que potencialmente puede ser usada para consumo humano. La falta de pared celular la hace un protoplasto natural, facilitando su transformación y presenta un solo cloroplasto, el cual puede ser empleado para crear un sistema de expresión a gran escala de proteínas recombinantes (Feng et al, 2014). *D. salina* actualmente se utiliza como suplemento para consumo humano por lo cual en un futuro la creación de sistemas de expresión de proteínas recombinantes en esta microalga, pudiera funcionar como vehículo para la entrega de vacunas orales. Además, una vez establecido este sistema de expresión, puede ser utilizado para la producción de diferentes proteínas recombinantes de interés biotecnológico.

Para poder llevar a cabo la producción de proteínas recombinantes es necesario contar con promotores adecuados. En *D. salina* se han probado una serie de promotores tanto endógenos como exógenos para la expresión de proteínas recombinantes en núcleo. El promotor CAMV 35S de plantas fue

utilizado para la expresión del gen reportero *gus* (Geng et al, 2004; Tan et al, 2005) y del gen *b/le* que confiere resistencia a zeocina (Sun et al, 2005). También se utilizó el promotor ubiquitin- Ω para la expresión de un antígeno de la superficie del virus de hepatitis B (Geng et al, 2003), sin embargo, los niveles de expresión fueron bajos y transitorios. Posteriormente se realizaron trabajos con promotores constitutivos endógenos, sin embargo, a pesar de ser endógenos, se obtuvo baja actividad y se vio afectado el crecimiento de las células transformadas (Li et al, 2010). Para superar estos problemas se trabajó con promotores inducibles. Uno de estos fue el promotor del gen que codifica para la enzima Nitrato reductasa, la cual es clave en el metabolismo de nitrógeno y cuya transcripción es inducida por nitrato y reprimida por amonio. Bajo la regulación de este promotor se expresó la proteína verde fluorescente (EGFP) (Li et al, 2008) y el gen *bar*, el cual confiere resistencia al herbicida Fosfinotricina (Li et al, 2007). En ambos casos se obtuvieron transformantes y se comprobó la expresión de ambos genes. Otro promotor estudiado es el del gen que codifica para la enzima Anhidrasa Carbónica duplicada I (DCA I), el cual es regulado por los niveles de NaCl que existen en el medio. Para evaluar la regulación de este promotor en la expresión del gen *bar* se utilizó un gradiente de salinidad de 0.5 M hasta 2 M, siendo en esta última concentración donde alcanzó su pico máximo de expresión (Li et al, 2010).

La expresión de proteínas recombinantes en *D. salina* solo se ha realizado a nivel nuclear y a la fecha no se conoce de un sistema en el cloroplasto. A

diferencia de la expresión en el núcleo, en cloroplasto depende del uso de promotores endógenos. La elección del promotor y las regiones reguladoras no traducidas (UTR's) son factores cruciales para que se lleve a cabo la expresión de genes recombinantes en cloroplasto (Specht et al, 2010). En el alga *C. reinhardtii* se han realizado estudios sobre el uso de promotores y UTR's heterólogos y a pesar de haber logrado la expresión de la proteína D1, los datos sugieren que la mejor forma de lograr la expresión de transgenes en cloroplasto, es utilizar regiones reguladoras endógenas (Gimpel y Mayfield, 2013). En ese mismo trabajo, Gimpel y Mayfield (2013) recomiendan que, si se va a trabajar con una especie que no cuenta con el genoma del cloroplasto secuenciado, es necesario buscar la forma de obtener las regiones reguladoras y recombinación homóloga del organismo en cuestión. El genoma del cloroplasto de *D. salina* cepa CCAP 19/18 fue secuenciado y reportado por Smith et al (2010) y hasta la fecha es el único reportado para esta especie.

En este trabajo se utilizará la cepa de *D. salina* SQ, la cual fue aislada de lagunas hipersalinas ubicadas en la región de San Quintín, Ensenada, B.C. México. Esta cepa fue aislada con la finalidad de ser utilizada para la producción de lípidos y escalar su producción en sistemas de cultivo de lagunas abiertas en esta zona. A diferencia de las cepas de otras regiones, las cepas nativas están aclimatadas a las condiciones del medio en esta región, facilitando las condiciones de cultivo. Adicionalmente, con esta cepa se están realizando estudios de producción de β -carotenos y otras caracterizaciones moleculares.

Con base en el genoma del cloroplasto de la cepa *D. salina* CCAP 19/19, en el presente trabajo se pretende utilizar promotores y secuencias reguladoras para la creación de un vector que permita la expresión de proteínas recombinantes y marcadores de selección en el cloroplasto de *D. salina* SQ

CAPÍTULO I

Construcciones de casetes para la expresión de proteínas recombinantes en el cloroplasto de *D. salina* SQ

1. Introducción

1.1 Ingeniería de cloroplastos en microalgas

Los cloroplastos de plantas y algas contienen su propio material genético, ya que estos organelos son producto de una endosimbiosis que surgió hace aproximadamente 1.5 mil millones de años cuando una cianobacteria fue fagocitada por una célula eucariota heterotrófica (Bock, 2015). Como producto de esta endosimbiosis surgieron una serie de cambios en los genomas de ambos organismos; tales como, transferencia genética del cloroplasto a núcleo y racionalización del genoma del cloroplasto mediante la eliminación de genes redundantes o no indispensables. Estos cambios dieron como resultado la reducción de tamaño de los genomas de los cloroplastos (Bock, 2015; Adem et al, 2017).

Hace aproximadamente dos décadas se empezó la ingeniería genética de cloroplastos; el genoma del cloroplasto de la planta del tabaco (*Nicotiana tabaccum*) fue modificado para conferir resistencia a un herbicida y al comparar los resultados de los niveles de expresión obtenidos en cloroplasto, con la expresión de este mismo transgen en el genoma nuclear, se observó que los niveles de expresión eran cientos de veces más altos en cloroplasto (Daniell et al, 1998).

La síntesis de productos recombinantes en el cloroplasto ofrece varias ventajas. Presenta altos niveles de expresión, se pueden expresar proteínas utilizando operones, presenta recombinación homóloga, por lo cual los genes

recombinantes pueden ser integrados en sitios específicos del genoma, como regiones intergénicas. La recombinación homóloga también evita que se integren otras regiones de los vectores (Daniell et al, 2004). A diferencia del núcleo, en cloroplasto no ocurre silenciamiento de los genes recombinantes, lo cual es un problema constante en las transformaciones nucleares (Shamriz y Ofoghi, 2016). El cloroplasto también puede ser utilizado para la expresión de proteínas que puedan ser tóxicas si se encuentran en el citosol de la célula (Bogard, 2000).

El cloroplasto de las algas unicelulares es particularmente interesante, ya que estos organismos generalmente contienen solo uno por célula, lo que facilita su ingeniería de genética (Purton, 2007). Además, ofrecen una serie de ventajas sobre la expresión de cloroplasto en plantas, el tiempo de crecimiento es menor y se lleva a cabo en fotobiorreactores cerrados lo cual reduce la contaminación de los cultivos, evitando la fuga de organismos genéticamente modificados al ambiente, lo cual es una preocupación constante cuando se utilizan cultivos de plantas. Algunas plantas producen metabolitos tóxicos secundarios, como la papa (*Solanum tuberosum*) y el tabaco, a diferencia de algunas microalgas, como *C. reinhardtii*, *Chlorella* sp., *H. pluvialis* y *D. salina* que no producen toxinas (Purton et al, 2013).

Para llevar a cabo la expresión de proteínas recombinantes en cloroplasto, es necesario que los casetes de expresión cuenten con los siguientes elementos: el gen de interés regulado por un promotor, 5' y 3' UTR endógenos; un marcador de selección regulado por un promotor, 5' y 3' UTR endógenos y secuencias

flanqueantes que permitan la integración del vector en una región específica del genoma del cloroplasto mediante recombinación homóloga (Fig. 3).

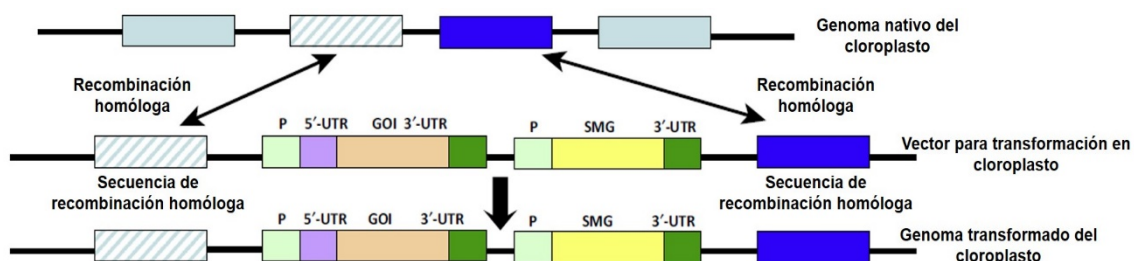


Figura 3. Representación esquemática de la integración de los casetes de expresión de proteínas recombinantes en cloroplasto. P: promotor; GOI: gen de interés; SMG: marcador de selección. Modificado de Jin y Daniell, 2015.

1.2. Promotores y regiones no traducidas (UTR's)

La expresión de proteínas en plantas se lleva a cabo mediante dos RNA polimerasas, NEP (Nucleus-Encoded Polymerase, por sus siglas en inglés) y PEP (Plastid-Encoded Polymerase, por sus siglas en inglés); codificadas en el genoma nuclear y en el genoma del cloroplasto, respectivamente. Sin embargo, la enzima PEP parece ser la única RNA polimerasa involucrada en la transcripción en el cloroplasto de algas verdes; mutantes de *C. reinhardtii* que presentan alteraciones en los genes *rpo* (que codifican para las subunidades de PEP) no son viables, lo que indica que los genes esenciales para el funcionamiento del cloroplasto son transcritos por PEP (Smith y Purton, 2002).

En *C. reinhardtii* se han identificado 3 tipos de promotores: 1) promotores típicos dependientes de σ -70 con elementos -10 y -35, 2) promotores que carecen de la región -35 y 3) una secuencia interna en algunos genes que codifican para RNA de transferencia, pero no han sido caracterizados (Klein, 2008). La eficiencia transcripcional, la estabilidad del transcrito y la eficiencia de traducción, tienen un gran impacto en la producción de proteínas endógenas en algas, y el 5' UTR de los mRNA es clave en cada uno de estos eventos (Mayfield et al, 2007). Los 5' UTR's contienen elementos *cis* que se requieren para la traducción de la región codificante (Barnes et al, 2005). Elementos dentro de los 5' y 3' UTR's tienen efecto sobre la estabilidad de los transcritos, probablemente mediante la interacción de factores proteicos que actúan de manera *trans* (Salvador et al, 2004).

De manera general, los promotores controlan los niveles de transcripción, el 5' UTR regula la estabilidad del mRNA y la iniciación de la traducción, y el 3'UTR está involucrado en la terminación de la transcripción y estabilidad del mRNA (Rasala et al, 2011). La expresión de genes recombinantes en cloroplasto se lleva a cabo bajo promotores y regiones 5' y 3' UTR's endógenas; es muy importante que las secuencias sean especie específicas para que se puedan lograr altos niveles de expresión; ya que, a diferencia de los altos niveles de conservación encontrados en las regiones codificantes, las regiones intergénicas y los UTR's son altamente variables en los genomas de los cloroplastos (Ruhlman et al, 2010).

1.3. Marcadores de selección

Los antibióticos son los marcadores de selección más utilizados, el gen *aph6* (de *Acinetobacter baumannii*) que confiere resistencia a kanamicina y el gen *aadA* (de *Escherichia coli*) que genera resistencia a estreptomicina y espectinomicina han sido utilizados en múltiples especies de algas (Specht et al, 2017). El gen *nat-1* que genera resistencia al antibiótico nourseotricina ha sido utilizado como marcador de selección en la diatomea *Phaeodactylum tricornutum* (Zaslavskaia et al, 2000), en esta misma diatomea se desarrolló un casete de expresión de proteínas recombinantes en cloroplasto que contiene el gen *cat*, el cual confiere resistencia a cloranfenicol (Xie et al, 2014). Otra manera de seleccionar las clonas transformantes, es utilizando mutantes no fotosintéticas, de tal manera que al integrarse en el transgen el marcador de selección permita que se recupere la habilidad de realizar fotosíntesis (Shamriz y Ofoghi, 2016), aunque esto solo es posible en algas mixotrofas. Algunas microalgas no son susceptibles a los antibióticos más comunes, como es el caso de *Platymonas subcordiformis*, la cual es naturalmente resistente a los antibióticos espectomicina, estreptomicina y kanamicina; por esta razón, los herbicidas también han sido utilizados como marcadores de selección para transformación en cloroplasto. El gen *bar* que confiere resistencia al herbicida Basta fue utilizado como marcador de selección en esta microalga (Cui et al, 2014).

1.4. Optimización de codones

El uso de codones varía entre especies, incluso difiere entre los organelos de un mismo organismo. Por lo general, los genes heterólogos tienen una frecuencia de codones diferente al organismo que se pretende utilizar como sistema de expresión, lo que puede generar problemas para la expresión de estos genes. La ausencia de RNA's de transferencia compatibles con el uso de codones del gen recombinante puede resultar en la incorporación de aminoácidos incorrectos o en el truncamiento de la cadena polipeptídica (Shamriz y Ofoghi, 2016). En diferentes trabajos de expresión de proteínas recombinantes en el cloroplasto de *C. reinhardtii* se ha demostrado que la frecuencia del uso de codones afecta los niveles de expresión (Franklin et al, 2002; Mayfield et al, 2003; Mayfield y Schultz, 2004; Jones et al, 2013; Tran et al, 2013). Por lo tanto, para solucionar este problema, es necesario optimizar los codones de los genes recombinantes, de tal manera que los codones raros sean remplazados por codones que utiliza con mayor frecuencia la maquinaria de traducción del cloroplasto.

1.5. Métodos de transformación

La pared celular de las algas representa una barrera física que impide la fácil incorporación de DNA foráneo, por lo que es necesario generar protoplastos; este paso puede ser limitante ya que la composición de la pared de las algas puede variar entre especies (Doron et al, 2016). La microalga *D. salina* carece de pared celular rígida, siendo un protoplasto natural. El método más utilizado para llevar

a cabo la transformación de cloroplastos es la biobalística (Bock, 2015), el cual utiliza macropartículas de oro o tungsteno cubiertas con el DNA que se desea ingresar; estas partículas son disparadas a alta velocidad, penetrando a las células. El alto costo del equipo de laboratorio requerido para llevar a cabo la biobalística ha hecho que se busquen alternativas más económicas. La agitación mediante vórtex de protoplastos en presencia de micropartículas de vidrio y el plásmido a transformar, es un método eficiente para la transformación de cloroplastos (Economou et al, 2014).

Una vez llevado a cabo el protocolo de transformación, la integración de los genes recombinantes ocurre mediante recombinación homóloga en el sitio previamente elegido para su integración; estas regiones que flanquean al gen recombinante deben tener un tamaño mínimo de 1 kb para asegurar una recombinación exitosa (Spetch, et al, 2017).

1.6. Producción de proteínas terapéuticas en el cloroplasto de microalgas

Las proteínas terapéuticas son un sector en crecimiento de la industria farmacéutica. La mayoría de estas proteínas se producen en bacterias, levaduras o células de mamífero, y en algunos casos en plantas y animales transgénicos (Rasala et al, 2010). El sistema de producción de proteínas terapéuticas en cloroplasto de microalgas ofrece varias ventajas, como reducción de costos,

métodos de cultivos sencillo en fotobiorreactores, compartimentalización de proteínas, bioseguridad y fácil escalamiento de cultivos.

En el cloroplasto de *C. reinhardtii* se ha llevado a cabo la síntesis de anticuerpos monoclonales completos tipo IgG1 que reconocen la toxina del ántrax (Tren et al, 2009), anticuerpos de tipos scFv (anticuerpos de cadena sencilla) fusionados con la exotoxina A de *Pseudomonas aeruginosa* con actividad antitumoral (Tran et al, 2012) y siete proteínas humanas terapéuticas, eritropoyetina, los dominios 10 y 14 de fibronectina, interferón β 1, proinsulina, isoforma 21 del factor de crecimiento vascular endotelial (VEGF), proteínas de alta movilidad del grupo de caja 1 (HMGB1) (Rasala et al, 2010).

La proteína recombinante Interleucina-2 es un agente terapéutico utilizado para el tratamiento de melanoma metastático y cáncer metastático renal (Rosenberg, 2014), fue producida por primera vez en *E. coli* en 1983 (Devos et al, 1983) y después de ser aprobado su uso por la FDA (Federal Drugs Administration, por sus siglas en inglés) aún se produce de manera recombinante en este sistema bajo el nombre comercial Proleukin®. En su estado nativo, es una proteína glucosilada, sin embargo, estas glucosilaciones no son necesarias para que esta proteína sea reconocida por su receptor, lo que la hace un buen candidato para ser producida en cloroplasto.

2. Hipótesis

La construcción de casetes de expresión utilizando promotores endógenos permitirá la expresión de proteínas recombinantes y marcadores de selección en el cloroplasto de *Dunaliella salina* SQ.

3. Objetivo general

Crear un sistema de expresión de proteínas recombinantes en el cloroplasto de la microalga *Dunaliella salina* SQ.

3.1 Objetivos específicos

- Construcción de un vector de expresión en el cloroplasto de *D. salina*.
- Transfección de *D. salina* con el casete de expresión creado.
- Expresión del gen de resistencia (*nat-1*).
- Expresión de una proteína recombinante (Interleucina-2; EGFP)

4. Metodología

4.1 Construcción 1

La construcción 1 consiste en formar un casete para la expresión de la proteína Interleucina-2 de humano optimizado para su expresión en cloroplasto y del gen que confiere resistencia al antibiótico nourseotricina (*nat-1*), ambos genes regulados bajo el promotor del gen *atpA* y el 3' UTR del gen *psbA* de *D. salina*. Las regiones de recombinación homóloga que permitirán la integración del casete en el genoma del cloroplasto de *D. salina* pertenecen a una zona intergénica entre los genes *psaA* y *psbA* (Fig. 4).

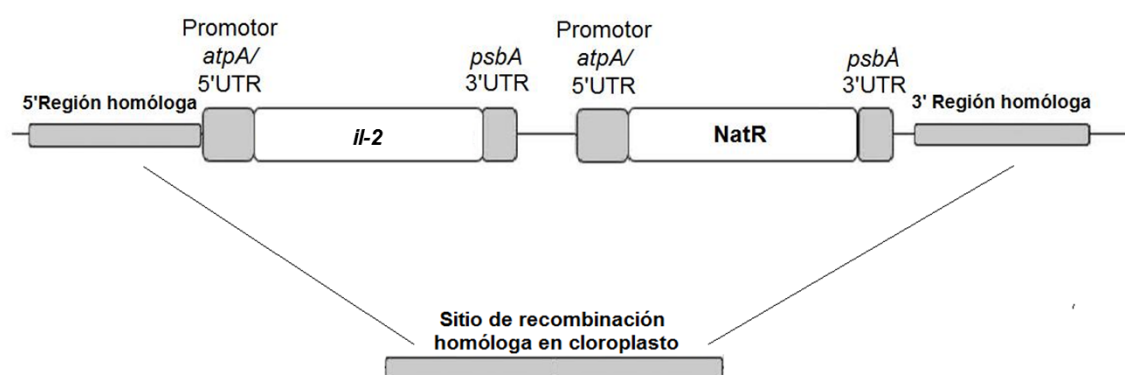


Figura 4. Casete para la expresión de interleucina-2 y el gen de resistencia nat-1 en el cloroplasto de *D. salina*.

4.1.1. Síntesis de promotor y UTR's

Se hizo un análisis de secuencias disponibles del cloroplasto de *D. salina* y se obtuvieron las secuencias del promotor, región no traducida (UTR) del extremo 5' y el UTR de la región 3' del gen *atpA*. Se agregaron dos sitios de reconocimiento de las enzimas *Nco* I y *Pst* I separado las regiones del promotr-

5'UTR del 3' UTR. La construcción fue enviada para su síntesis a la empresa Genscript (Piscataway NJ, USA), la cual envió el fragmento clonado en el plásmido pUC57, a este plásmido se le denominó pUC57-utr's (trabajo previamente realizado en el laboratorio del Dr. Stephano por el MC. Ibrahim Guzmán).

4.1.2 Clonación del marcador de selección *nat-1* en el plásmido pUC57-utr's

El gen que codifica para la enzima nourseotricina acetiltransferasa y confiere resistencia al antibiótico nourseotricina fue elegido como agente de selección. El gen *nat-1* fue aislado del plásmido pYL16 (Werner Bioagents) con los primers Nat1-for y Nat1-rev (Anexo primers). El primer Nat1-for fue diseñado con el sitio de reconocimiento de la enzima de restricción *Nco* I mientras que al primer Nat1-rev se le adicionó el sitio de reconocimiento de la enzima de restricción *Pst* I. La reacción en cadena de la polimerasa (PCR) se realizó con la enzima *Taq* platinum high fidelity (Invitrogen), los componentes fueron 1X de buffer high fidelity, 2 mM MgSO₄, 0.2 mM de dNTP's, 0.5 µM de cada primer (Nat1-for y Nat1-rev), 150 ng del plásmido pYL16, 1 U de *Taq* platinum high fidelity en un volumen final de 50 µl. Las condiciones de reacción fueron 2 min a 94 °C, 30 ciclos de 94 °C 15 s, 60 °C 30s y 68 °C 35 s, seguido de un ciclo 5 min a 68 °C. Posteriormente 25 µl de la reacción fueron visualizados en un gel de agarosa al 1%. El resto de la reacción se purificó utilizando el kit comercial QIAquick Gel Extraction (Qiagen) siguiendo las instrucciones del kit.

El plásmido pUC57-utr's y el fragmento amplificado del gen *nat-1* fueron digeridos con las enzimas *Nco* I y *Pst* I (New England Biolabs), las reacciones se realizaron con 1X de NEBuffer 3, 1 µg de plásmido o 500 ng de producto de PCR, 10 U de *Nco* I, 10 u de *Pst* I en un volumen final de 20 µl. Las reacciones se incubaron 1 h a 37 °C y después se purificaron utilizando el kit comercial QIAquick PCR purification (Qiagen).

La reacción de ligación se llevó a cabo utilizando la enzima T4 DNA ligasa (Fermentas), la cual contenía 1X de DNA ligase buffer, 500 ng del plásmido pUC57-utr's digerido, 150 ng de la PCR del gen *nat-1*, 5 U de T4 DNA ligasa en un volumen final de reacción de 10 µl. La reacción se incubó 2 h a 37 °C. Posteriormente la ligación fue utilizada para transformar células químicamente competentes de *E. coli* DH5α. La reacción de ligación se mezcló con células competentes, se incubó en hielo durante 15 min, se le dio un choque térmico (1 min 42 °C y 1 min en hielo) y se agregaron 250 µl de medio SOC (0.5 % extracto de levadura, 2 % triptona, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ y 20 mM glucosa). La mezcla se incubó 1 h a 37 °C 250 rpm y se sembró en placas con medio LB agar (10 % triptona, 0.5 % extracto de levadura, 1 % NaCl y 1.5 agar %) con carbenicilina (50 µg/ml). Las placas se incubaron durante 12 h a 37 °C.

Se seleccionaron 3 colonias y se resembraron en 10 ml de medio LB líquido (10 % triptona, 0.5 % extracto de levadura y 1 % NaCl) con carbenicilina, se incubaron durante 12 h a 37 °C y 250 rpm. Se purificó el plásmido utilizando el kit comercial Wizard Plus SV Minipreps DNA Purification (Promega) siguiendo las

especificaciones de la casa comercial. Para comprobar la clonación del gen *nat-1*, los plásmidos purificados fueron digeridos con las enzimas *Nco* I y *Pst* I. Las reacciones contenían 1X de NEBuffer 3, 1 µg de plásmido, 10 U de *Nco* I y 10 u de *Pst* I en un volumen final de 20 µl. Las reacciones se incubaron 1 h a 37 °C y fueron visualizadas en un gel de agarosa al 1%. Este plásmido fue denominado pNat-utr.

4.1.3. Clonación del gen que codifica para Interleucina-2 en el plásmido pUC57-utr's

El gen que codifica para la proteína Interleucina-2 humana fue optimizado para su expresión en el cloroplasto de *D. salina*. Para la optimización del gen se seleccionaron las secuencias de 79 genes codificantes de cloroplasto, las cuales fueron compactadas en un solo archivo fasta y posteriormente se sometieron a la página de EMBOSS/cusp (www.bioinformatics.nl/cgi-bin/emboss/cusp). El programa desplegó la lista de codones más utilizados en el cloroplasto de *D. salina*. En base a estos resultados se hizo una sustitución de codones utilizando los más frecuentes. También se agregaron las secuencias de reconocimiento de restricción de las enzimas *Nco* I y *Pst* I, 6 codones de histidina para su posterior purificación mediante una matriz con níquel y el sitio de reconocimiento de la enzima enterocinasa para remover el marcaje de polihistidinas después de la purificación de la proteína. La secuencia fue enviada a sintetizar a la empresa Genscript (Piscataway NJ, USA), la cual envió el fragmento clonado en el plásmido pUC57, a este plásmido se le denominó pUC57-il2.

Los plásmidos pUC57-utr's y pUC57-il2 fueron digeridos con las enzimas *Nco* I y *Pst* I (New England Biolabs), las reacciones contenían 1X de NEBuffer 3, 1 µg de plásmido, 10 U de *Nco* I y 10 u de *Pst* I, en un volumen final de 20 µl. Las reacciones se incubaron 1 h a 37 °C. La digestión del plásmido pUC57-utr's se purificó utilizando el kit comercial QIAquick PCR purification (Qiagen). La digestión del plásmido pUC57-il2 se corrió en un gel de agarosa al 1% y el fragmento correspondiente al gen *il-2* fue extraído del gel utilizando el kit comercial QIAquick Gel Extraction (Qiagen).

La reacción de ligación del gen *il-2* en el plásmido pUC57-utr's se llevó a cabo utilizando la enzima T4 DNA ligasa (Fermentas), la cual contenía 1X de DNA ligase buffer, 500 ng del plásmido pUC57-utr's digerido, 125 ng gen *il-2* y 5 U de T4 DNA ligasa en un volumen final de 10 µl. La reacción se incubó 2 h a 37 °C y posteriormente se utilizó para transformar células químicamente competentes de *E. coli* DH5α. La reacción de ligación se mezcló con las células competentes, se incubó en hielo durante 15 min, se dio choque térmico (1 min 42 °C, 1 min en hielo) y se agregaron 250 µl de medio SOC. Se incubó 1 h a 37 °C 250 rpm y se sembró en placas con medio LB agar con 50 µg/ml de carbenicilina, las cuales se incubaron durante 12 h a 37 °C.

Se seleccionaron 3 colonias y se resembraron en 10 ml de medio LB líquido con carbenicilina, se incubaron durante 12 h a 37 °C y 250 rpm. Se extrajo plásmido utilizando el kit comercial Wizard Plus SV Minipreps DNA Purification (Promega) siguiendo las especificaciones de la casa comercial. Para comprobar los plásmidos tenían clonado el gen *il-2* se realizó una PCR con los primers Pro-

for y UTR-rev utilizando la enzima *Taq* platinum high fidelity (Invitrogen), los componentes de la reacción 1X de buffer high fidelity, 2 mM MgSO₄, 0.2 mM de dNTP's, 0.5 µM de cada primer (Pro-for y UTR-rev), 200 ng de plásmido, 1 U de *Taq* platinum high fidelity y 38.2 µl de agua destilada, obteniendo un volumen final de 50 µl. Las condiciones de reacción fueron 2 min a 94 °C, 30 ciclos de 94 °C 15 s, 60 °C 30s y 68 °C 1:30 min, seguidos de un ciclo 5 min a 68 °C. Posteriormente las reacciones fueron analizadas en un gel de agarosa al 1%. Este plásmido fue nombrado pIL2-utr.

4.1.4 Unión de las construcciones IL-2 y Nat-1

La construcción pIL2-utr, la cual contiene el gen *il-2* con el promotor y los UTR's fue clonada en el vector pNat-utr (Fig. 5). Para realizar la clonación de esta construcción se realizó una PCR de la construcción il2-UTR's para agregar los sitios de restricción de las enzimas *Nde* I y *Bst*X I.

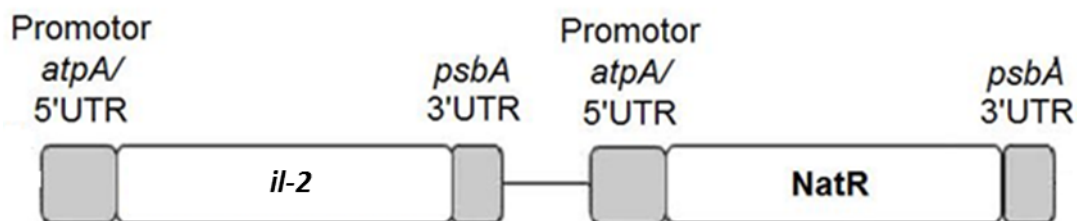


Figura 5. Construcción il-2-NatR. Unión del gen recombinante con el marcador de selección.

Esta PCR se realizó utilizando la enzima TopTaq (Qiagen), los componentes de la reacción 1X de buffer TopTaq PCR buffer, 0.2 mM de dNTP's, 0.5 μ M de cada primer (PrNde-F y UTRBst-R), 135 ng de plásmido y 1.5 U de TopTaq DNA polimerasa, en un volumen final de 50 μ l. Las condiciones de reacción fueron 3 min a 94 °C, 30 ciclos de 94 °C 30 s, 65 °C 30s y 72 °C 1:30 min, seguidos de un ciclo 10 min a 72 °C. Después se corrió un gel de agarosa al 1% y la banda correspondiente al producto de PCR fue purificada utilizando el kit comercial QIAquick Gel Extraction (Qiagen). Posteriormente tanto el producto de PCR purificado como el plásmido pNat-utr fueron digeridos con las enzimas *Nde* I y *Bst*X I (New England Biolabs). Las reacciones contenían 1X de NEBuffer 3, 1 μ g de plásmido pNat-utr o 500 ng del producto de PCR purificado, 10 U de *Nde* I y 10 u de *Bst*X I, en un volumen final de 30 μ l. Las reacciones se incubaron 3 h a 37 °C. El plásmido digerido se corrió en un gel de agarosa al 1 % y la banda correspondiente al plásmido fue purificada del gel utilizando el método de Freeze-Squeeze. Una vez que se escindió la banda del gel con una navaja, esta se fragmentó y se congeló durante 30 min a -80 °C, después se pulverizó y se traspasaron a una columna (Millipore) y se centrifugaron 10 min a 6000 rpm. El DNA fue precipitado del sobrenadante utilizando 2.2 volúmenes de etanol y 1:10 de acetato de sodio (3M pH 5.2). La mezcla se refrigeró a -80 °C por 2 h, se centrifugó (13,000 rpm, 4 °C, 30 min), se decantó el sobrenadante y se agregó 1 ml de etanol al 70 %, se centrifugó nuevamente (13,000 rpm, 4 °C, 15 min) y se retiró el sobrenadante. El pellet de DNA se secó y se resuspendió en 25 μ l de agua destilada.

La digestión del fragmento de PCR fue purificada utilizando el kit comercial QIAquick PCR purification (Qiagen). Después de las respectivas purificaciones se realizó la reacción de ligación utilizando la enzima T4 DNA ligasa (Fermentas). La ligación se llevó a cabo con 100 ng de plásmido y 70 ng de producto de PCR. Las condiciones de la reacción de ligación, así como la transformación de *E. coli* DH5 α fueron como se describe en los aparados anteriores. Para comprobar que las clonas transformadas contenían el inserto, se purificó el plásmido utilizando el kit comercial Wizard Plus SV Minipreps DNA Purification (Promega) siguiendo las instrucciones del manual de usuario. Estos plásmidos fueron digeridos con las enzimas *Nde* I y *Bst*XI (New England Biolabs) como se describe previamente. El producto de la digestión fue analizado en un gel de agarosa al 1 %. El plásmido obtenido fue nombrado pIL2-Nat.

4.1.5 Clonación de las regiones de recombinación homólogas en el plásmido pIL2-Nat

Como región de recombinación homóloga se seleccionó una región intergénica para evitar la interrupción de genes (Fig. 4). Las regiones de recombinación homóloga 5' y 3' se obtuvieron mediante PCR a partir de DNA de *D. salina* SQ. Para clonar estas regiones se utilizó el kit In-Fusion HD (Clontech).

La microalga *D. salina* se cultivó en medio modificado (10 mM KNO₃, 50 mM NaHCO₃, 5 mM MgSO₄·7H₂O, 0.4 mM KH₂PO₄, 2 μ M FeCl₃·6H₂O, 5 μ M EDTA, 7 μ M MnCl₂·4H₂O, 1 μ M CuCl₂·2H₂O, 1 μ M ZnCl₂, 1 μ M CoCl₂·6H₂O, 1 μ M (NH₄)Mo₇O₂₄·4H₂O, 185 μ M H₃Bo₃, 0.2 mM CaCl₂) con 1.5 M NaCl y se le

suministró luz de manera constante hasta que alcanzó una densidad 1×10^8 células/ml. El DNA se purificó con el kit AxyPrep Multisource Genomic DNA miniprep (Axygen) siguiendo las instrucciones para cultivos celulares. Una vez extraído el DNA total de *D. salina* fue utilizado como template para amplificar las regiones homólogas 5' y 3'. Primero se amplificó mediante PCR la región 3'. La reacción contenía 1X de buffer TopTaq PCR buffer, 0.2 mM de dNTP's, 0.5 μ M de cada primer (Infusion-1 y InPvu-R), 500 ng de DNA total y 1.5 U de TopTaq DNA polimerasa, obteniendo un volumen final de 50 μ l. Las condiciones de reacción fueron 3 min a 94 °C, 30 ciclos de 94 °C 30 s, 65 °C 30s y 72 °C 1:30 min, seguidos de un ciclo 10 min a 72 °C. Los primers utilizados en esta PCR contienen una región homóloga a los extremos libres del plásmido donde se va a clonar. El producto de PCR fue purificado utilizando el kit NucleoSpin Gel and PCR Clean-up (Clontech).

El plásmido pIL2-Nat fue digerido con las enzimas *Hind* III y *Pvu* II para crear extremos libres y poder clonar la región 3' homóloga. La doble digestión se llevó a cabo con 1X NEBuffer 4, 1 μ g de plásmido, 20 U *Hind* III (NewEngland Biolabs) y 10 U *Pvu* II (NewEngland Biolabs) en un volumen final de 30 μ l; esta reacción fue incubada 3 h a 37 °C. La reacción se corrió en un gel de agarosa al 1 % y la banda correspondiente al plásmido fue extraída utilizando el método de Freeze-Squeeze.

Los productos de PCR y el plásmido purificado fueron utilizados para realizar la reacción de ligación la cual se llevó a cabo utilizando el kit In-Fusion HD (Clontech); siguiendo las instrucciones incluidas. Posteriormente se

transformaron 2.5 µl de la reacción de ligación en células competentes DH5α y se sembraron en placas de LB agar con 50 µg/ml de carbenicilina.

Se seleccionaron la clonas que crecieron en las placas de agar, se inocularon en medio LB líquido con 50 µg/ml de carbenicilina y se purificó el plásmido utilizando el kit QUIAGEN Plasmid Mini, siguiendo las instrucciones proporcionadas. Para comprobar que se había clonado la región 3' en el plásmido se hizo una digestión con la enzima *BstX* I. La reacción contenía 1X NEBuffer, 1 µg de plásmido y 10 U de *BstX* I, en un volumen final de 20 µl y se incubó 3 h a 37 °C. Finalizada la digestión fue evaluada en un gel de agarosa al 1 %. Este plásmido fue nombrado pIN3'. Se seleccionó una clona positiva para continuar con la ligación de extremo de recombinación homóloga 5'.

El sitio de recombinación homóloga 5' fue amplificado a partir del DNA total extraído de *D. salina* utilizando los primers InfNde-F y Inf2Nde-R, los cuales contienen una región homóloga al vector pIN3'. La reacción contenía 1X de buffer TopTaq PCR buffer, 0.2 mM de dNTP's, 0.5 µM de cada primer (InfNde-F y Inf2Nde-R), 500 ng de DNA total y 1.5 U de TopTaq DNA polimerasa, obteniendo un volumen final de 50 µl. Las condiciones de reacción fueron 3 min a 94 °C, 30 ciclos de 94 °C 30 s, 65 °C 30s y 72 °C 1:30 min, seguidos de un ciclo 7 min a 72 °C. El producto de PCR fue purificado utilizando el kit NucleoSpin Gel and PCR Clean-up (Clontech).

Para clonar la región 5' homologa el plásmido pIN3' fue digerido con la enzima *Nde* I, la cual lo linearizaba después de la región 5'UTR donde se encontraba en gen *il-2*. La reacción de digestión contenía 1 µg de plásmido, 1X

NEBuffer 3 y 10 U *Nde* I en un volumen final de 20 µl. La digestión se incubó 3 h a 37 °C. La digestión fue precipitada con etanol y acetato de sodio.

Los productos de PCR y el plásmido purificado fueron utilizados para realizar la reacción de ligación la cual se llevó a cabo utilizando el kit In-Fusion HD (Clontech); siguiendo las instrucciones incluidas. Posteriormente se transformaron 2.5 µl de la reacción de ligación en células competentes DH5α y se sembraron en placas de LB agar con 50 µg/ml de carbenicilina.

Se seleccionaron la clonas que crecieron en las placas de agar, se inocularon en medio LB líquido con 50 µg/ml de carbenicilina y se purificó el plásmido utilizando el kit QUIAGEN Plasmid Mini siguiendo las instrucciones proporcionadas. Para comprobar si estaba clonada la región homóloga 5', se realizó una digestión con la enzima *Nru* I. La reacción de digestión contenía 1X NEBuffer 3, 1 µg de plásmido, 10 U de *Nru* I (NewEngland Biolabs) y 12 µl de agua. La reacción se incubó 3 h a 37 °C y después se visualizó en un gel de agarosa al 1 %. Una de las clonas que salió positiva fue enviada a secuenciar. En esta secuenciación se corroboró la presencia y la correcta secuencia de cada una de las partes clonadas. El plásmido resultante fue denominado pDsCh.

4.2. Construcción 2

Esta construcción consiste en la expresión de un gen reportero y el gen *nat-1* que confiere resistencia al antibiótico nourseotricina. El gen reportero codifica para la proteína EGFP (Enhanced Green Fluorescent Protein, por sus

siglas en ingles) y está regulado bajo el promotor y 5' UTR del gen *rbcL* de *D. salina*, y el 3' UTR del gen *psbA* (Fig. 6). El gen *nat-1* está regulado está regulado bajo el promotor y 5' UTR del gen *psbA* de *D. salina*, y el 3' UTR del gen *atpA* (Fig. 6). Las regiones de recombinación homólogas fueron las mismas utilizadas en la construcción 1. A diferencia de la construcción 1, en este casete el gen reportero y el marcador de selección se encuentran regulados bajo diferentes promotores y UTR's.

4.2.1 Síntesis del casete de expresión

El gen que codifica para la proteína EGFP fue optimizado para su expresión en el cloroplasto de *D. salina*. En esta ocasión el gen *nat-1* también fue optimizado para su expresión en el cloroplasto de *D. salina*. Como elementos reguladores de la expresión de la proteína EGFP se utilizó el promotor y la región 5'UTR del gen *rbcL* y el 3'UTR del gen *psbA*. Para lograr la expresión del gen *nat-1* se utilizó el promotor y el 5'UTR del gen *psbA* y el 3'UTR del gen *atpA*. Los promotores y UTR's fueron tomados de la secuencia del cloroplasto de *D. salina* CCAP 19/18 (número de acceso de NCBI GQ250046.1). También se le agregaron sitios de restricción de diferentes enzimas entre cada uno de los elementos (Fig.6). Este casete de expresión fue enviado a sintetizar a la empresa Genscript (Piscataway, NJ, EU). La secuencia fue enviada clonada en el vector pUC57, a este vector se le llamó pEGFP.

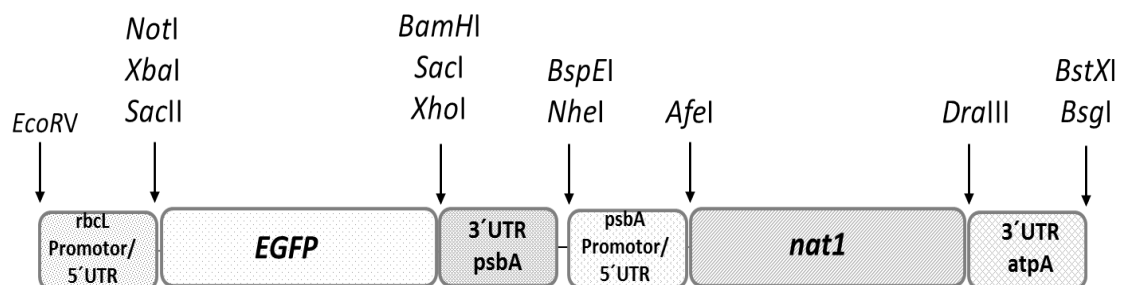


Figura 6. Casete de expresión de la proteína EGFP y el gen *nat-1* optimizados para cloroplasto. Se muestran la ubicación de los sitios de restricción de diferentes enzimas.

4.2.2 Clonación de la región de recombinación homóloga 5' en el plásmido pEGFP

La región 5' homóloga fue amplificada a partir del plásmido pDsCh, bajo las siguientes condiciones, 1X de buffer Phusion HF PCR, 0.2 mM de dNTP's, 0.5 μ M de cada primer (Hom5GFP-F y Hom5GFP-R), 150 ng de plásmido pDsCh y 1 U de Phusion DNA polimerasa, obteniendo un volumen final de 50 μ l. Las condiciones de reacción fueron 30 s a 98 °C, 30 ciclos de 98 °C 10 s, 65 °C 30s y 72 °C 33 s, seguidos de un ciclo 10 min a 72 °C. El producto de PCR fue purificado utilizando el kit NucleoSpin Gel and PCR Clean-up (Clontech).

El plásmido pEGFP fue digerido con la enzima *EcoR* V, la reacción contenía 1X NEBuffer 3, 1 μ g de plásmido, 20 U de *EcoR* V (NewEngland Biolabs) y 14 μ l de agua. La reacción se incubó 3 h a 37 °C y se purificó con el kit NucleoSpin Gel and PCR Clean-up (Clontech).

La reacción de recombinación con el kit Infusion-HD para integrar la región 5' homóloga dentro del plásmido pEGFP se llevó a cabo con 93 ng de PCR purificado y 100 ng de vector digerido, siguiendo las instrucciones incluidas en el kit. Posteriormente esta reacción se transformó en células competentes de *E. coli* DH5 α . Se seleccionaron 10 clonas transformadas, se crecieron en medio LB con carbenicilina (50 μ g/ml) durante 14 h a 37 °C y se les purificó el plásmido con el kit GeneJet Plasmid Miniprep (ThermoFisher) siguiendo las instrucciones proporcionadas. Para comprobar que la región homóloga se había clonado se hizo una digestión con la enzima *EcoR* V. La reacción contenía 1X NEBuffer 3, 1 μ g de plásmido, 20 U de *EcoR* V (NewEngland Biolabs) y 12.5 μ l de agua destilada. La reacción se incubó 3 h a 37 °C y se visualizó en un gel de agarosa al 1%. El vector resultante fue nombrado pGFP5.

4.2.3 Clonación de la región de recombinación homóloga 3' en el plásmido pGFP5

La región 3' de recombinación homóloga fue amplificada utilizando los primers Hom3GFP-F y Hom3GFP-R, estos primers contenían secuencias homólogas al vector pGFP5 para poder insertarla mediante recombinación homóloga utilizando el kit Infusion-HD (Clontech). La amplificación se llevó a cabo utilizando CloneAmp HiFi PCR Premix (Clontech). La reacción contenía 12.5 μ l de CloneAmp HiFi PCR Premix, 0.5 μ M de cada primer (Hom3GFP-F y Hom3GFP-R), 80 ng de plásmido pDsCh y 9 μ l de agua destilada, obteniendo un volumen

final de 25 µl. Las condiciones de reacción fueron 30 s a 98 °C, 30 ciclos de 98 °C 10 s, 65 °C 15s y 72 °C 7 s, seguidos de un ciclo 5 min a 72 °C. El producto de PCR fue purificado utilizando el kit NucleoSpin Gel and PCR Clean-up (Clontech).

El plásmido pEGFP5 fue digerido con la enzima *BstX* I, la reacción contenía 1X NEBuffer 3, 1 µg de plásmido y 10 U de *BstX* I (NewEngland Biolabs), en un volumen final de 20 µl. La reacción se incubó 2 h a 37 °C y se purificó con el kit NucleoSpin Gel and PCR Clean-up (Clontech).

La reacción de recombinación con el kit Infusion-HD se llevó a cabo con 95 ng de PCR purificado y 98 ng de vector digerido, siguiendo las instrucciones incluidas en el kit. Posteriormente con esta reacción se transformaron células competentes de *E. coli* DH5α. Se seleccionaron clonas transformadas y se crecieron en medio LB con carbenicilina (50 µg/ml) durante 14 h a 37 °C para extraerles plásmido con el kit GeneJet Plasmid Miniprep (ThermoFisher). Para comprobar la inserción de la región homóloga 3' se realizó una doble digestión con las enzimas *Nru* I y *EcoR* I; la reacción contenía 1X NEBuffer 3, 10 U de *Nru* I, 10 U de *EcoR* I, 1 µg de plásmido y 17 µl de agua destilada para obtener un volumen final de reacción de 30 µl. La reacción se incubó 2 h a 37 °C y posteriormente se evaluó en un gel de agarosa al 1 %. La clona 2 positiva se envió a secuenciar para comprobar la correcta secuencia de ambas regiones homologas. El plásmido resultante fue denominado pDsEGFP.

4.3. Transformaciones

4.3.1. Transformación del plásmido pDsCh

La transformación del plásmido pDsCh se llevó a cabo utilizando la técnica de biobalística. Las células de *D. salina* se crecieron a una densidad de 1.5×10^6 células/ml en medio modificado con 250 mM de NaCl bajo luz las 24 h. Las células se centrifugaron y se resuspendieron en medio modificado quedando a una densidad de 1×10^8 células/ml. Se sembraron 1×10^8 células en placas de medio modificado con agar.

Las placas de dejaron secar bajo campana mientras se preparaban las micropartículas de tungsteno (BioRad). Se resuspendieron 10 mg de micropartículas en 500 μ l de etanol y se centrifugaron durante 30 s a 14,000 rpm; se retiró el exceso de etanol, se adicionaron nuevamente 250 μ l de etanol y se dio vórtex durante 2 minutos. Las partículas se centrifugaron durante 30 s a 14,000 rpm, se retiró el etanol y se lavaron 2 veces con 250 μ l de agua destilada estéril, después de la última centrifugada se resuspendieron en 150 μ l de agua destilada estéril. Para cubrir las micropartículas con el plásmido pDsCh, se tomaron 50 μ l de partículas resuspendidas en agua destilada estéril y mientras se daba vórtex moderado se adicionaron 5 μ l de plásmido (1 μ g/ μ l), 50 μ l de 2.5 M CaCl_2 y 20 μ l de espermidina (free base) estéril 0.1 M. Una vez agregados estos componentes se dio vórtex moderado durante 3 min y se centrifugó durante 30 s, se retiró el sobrenadante, se adicionaron 250 μ l de etanol y se dio vórtex por 1 min. Posteriormente se centrifugó nuevamente durante 30 s, se

resuspendió en 50 µl de etanol y se almacenó en hielo hasta su uso. Todos los componentes utilizados en la transfección se utilizaron estériles. Se dispararon 10 µl de micropartículas recubiertas del plásmido en cada caja a 1000 psi. Después del disparo se agregaron 5 ml de medio modificado a cada caja, se selló con parafilm y se incubaron toda la noche con iluminación constante. Después se removieron las células de las placas y se sembraron en placas nuevas con medio modificado y 500 ng/ml del antibiótico nourseotricina. Las placas se incubaron con iluminación constante durante 15 días y las colonias que aparecieron se resembraron en placas nuevas con medio modificado y 500 ng/ml de nourseotricina.

Para comprobar la correcta inserción del casete dentro del genoma del cloroplasto se realizó una PCR con los primers NatF y NatR; utilizando como template DNA genómico de clonas transformadas. La extracción de DNA se hizo utilizando el kit AxyPrep Multisource Genomic DNA miniprep (Axygen) siguiendo las instrucciones para cultivos celulares. Se agregaron 2 ml de medio modificado con 250 mM de NaCl a las cajas de Petri para remover las células y este sobrenadante fue utilizado para llevar a cabo la extracción. Como control para corroborar la correcta extracción de DNA de las muestras, se diseñaron 2 juegos de primers que amplificaban los 2 exones del gen *atpA*. Los primers atpaF y atpaR amplifican el parte del primer exón y los primers Atpa2F y Atpa2R amplifican una parte del segundo exón del mismo gen. Las reacciones de PCR contenían 1X de buffer Phusion HF 5X PCR buffer, 0.2 mM de dNTP's, 0.5 µM de cada primer, 500 ng de DNA genómico de *D. salina* (transformada y control),

1 U de Phusion DNA polimerasa y 35.5 μ l de agua destilada, obteniendo un volumen final de 50 μ l. Las condiciones de reacción fueron 30 s a 98 °C, 30 ciclos de 98 °C 10 s, 65 °C 30s y 72 °C 30 s, seguidos de un ciclo 10 min a 72 °C. Las reacciones se evaluaron en un gel de agarosa al 1 %.

4.3.2. Transformación del plásmido pDsEGFP

La transformación se realizó con micro esferas de vidrio de 425 – 600 μ m de diámetro. El cultivo de *D. salina* se creció hasta alcanzar fase logarítmica y se centrifugaron 100 -150 ml del cultivo a 5000 rpm por 2 min, se retiró el sobrenadante y el paquete celular se lavó tres veces con 25 ml de medio de cultivo centrifugando a 5000 rpm por 2 min. El paquete de células se resuspendió en 800 μ l de medio obteniendo una densidad de 1×10^5 cel/ml y posteriormente se le agregó 300 mg de microesferas, 60 μ g del plásmido y 100 μ l de PEG 20%. La mezcla se agitó utilizando vórtex a 2400 rpm durante 15 s, se reposaron 10 s y después se agitó nuevamente durante 60 s. Se dejó reposar 3 min para que las microesferas se precipitaran. Después de la sedimentación de las microesferas, la mezcla se colocó en 24 h con poca luz y posteriormente se sembró en placas con medio modificado 600 μ g/ml de nourseotricina. Las placas se incubaron con luz intensa constante a 28 °C durante 15 días o hasta que aparecieron colonias. Las colonias obtenidas se resembraron en medio modificado liquido con 250 mM de NaCl y 600 μ g/ml de nourseotricina.

4.3.3. Detección de transformantes de *D. salina*

Para comprobar la integración de los casetes en el genoma del cloroplasto de las células de *D. salina* que sobrevivieron en medio con nourseotricina se realizó una extracción de DNA total. Se tomaron 20 ml de medio de cultivo de las clonas que crecieron en medio con el antibiótico y se centrifugó durante 15 min a 4 °C a 4000rpm, se eliminó el sobrenadante y el paquete de células se resuspendió en 700 µl de buffer CTAB (0.1 M Tris-HCl pH 8, 1.4 M NaCl, 0.01 M EDTA pH 8, 2% CTAB y 2% β-Mercaptoetanol), se sonicó 2 veces en hielo durante pulsos de 45 s, después se incubaron las células a 65 °C durante 90 min; posteriormente se adicionó 1 volumen de fenol/cloroformo/alcohol isoamílico, se dio vórtex 1 min, se centrifugó a 13,000 rpm a 4 °C durante 15 min. La fase acuosa se transfirió a un tubo nuevo y se precipitó el DNA utilizando 2 volúmenes de etanol absoluto frío y se lavó el paquete de DNA con 1 ml de etanol al 70 %. Finalmente, el DNA fue re suspendido en 200 µl de dH₂O (el agua estaba a 65 °C).

Para comprobar la integración del plásmido en *D. salina* se realizó PCR con el DNA obtenido de las clonas transformadas y de la cepa no transformada como control; como control de la amplificación se utilizó la región del gen *atpA*. Las reacciones de PCR contenían 1X de buffer Phusion HF 5X PCR buffer, 0.2 mM de dNTP's, 0.5 µM de cada primer, 500 ng de DNA genómico de *D. salina* (transformada y control), 1 U de Phusion DNA polimerasa y 35.5 µl de agua destilada, obteniendo un volumen final de 50 µl. Las condiciones de reacción fueron 30 s a 98 °C, 30 ciclos de 98 °C 10 s, 65 °C 30s y 72 °C 30 s, seguidos de

un ciclo 10 min a 72 °C. Las reacciones se evaluaron en un gel de agarosa al 1 %.

5. Resultados y discusiones

5.1. Síntesis de regiones reguladoras

La secuencia del promotor, 5' UTR y 3' UTR del gen *atpA* se muestran en la figura 7. Las regiones de los sitios de restricción de las enzimas *Nco* I y *Pst* I se encuentran subrayados y en ese respectivo orden. La longitud de la secuencia es de 977 pares de bases (pb). La secuencia fue clonada en el vector pUC57.

```
GGTACCAATATTAATTTAAATAATATTGTTATTTATTAATAGGTGGTGTCAATTTACGCCACCAAAGCG
CGCACAAGCGCGGAAAAAAAAAAAAAAAAATTTATTAATAATAAATTTTTAATGATCTCTCCGGAATAT
ATAGATAATAACAAATTCAGAAGAAATTTTATAGAAGTTTTAAATATCCTAAAGAATATTTAAATCTTC
TAAAATTTGTTATCTATTTATAACAATTTTATCGGTTTTTTAGAAAAAAAAAAAAAGTATCTCCTAAATT
AAAATGAAAAATTCCTTAGATAAGGGTATGAATGACGACCTTACTTATTTTTTCTTAGGTCGTTAAAAA
TACAGCAAATAGAATAATATATTTAAAAATATAAAGTGTTCGTTTACGACGGAAAGTATAAATTTTAA
ATATATTATTAACAATAGTTTATAAGTTATAAACTATTTGTTTCTATCATATCCCTTTAACAAGAAAA
AAAATTTTTTTCCATGGTATATATATATATATCTGCAGTAATTCTTTATACTTAAGAATTAAGTAGACT
TGTCTACTTAATTCCTTTTAAATAGCATGGATGGAGTAACTCCATCCATGCTATTATCTCTAACAACCC
CAAAAACCAAAGGTTTTTGTGAAGGAAAAAATAGAATATTTTTTCCGCTACGAGCGGGCGATATAAAAA
AATAAGATTTTTTTATCCTAAATTAATTCCTTGTAATAGAATCTCAATTAATGACAAAGATATAAATTAC
AAAAACGGTAGAGTTTCAGGTTTAGATAAATAACCTGGTCCTGCTATATACTAATGTAAGTTAAAATA
TCAGTATATGAACAATATATATATATATTAATAGTCCATGTCCTTCAAAAACGAATAAACGAATAAAC
CAAAGAAAATAAACTTTATTATTTTAAATAAAAATAAAATTTTATTAATTTTTTTTTTTATCAAAAATT
TTTTTGGTACC
```

Figura 7. Secuencia de promotor, 5' UTR y 3' UTR del gen *atpA*

5.2. Clonación del marcador de selección *nat-1* en el plásmido pUC57-utr's

La amplificación de PCR realizada con los primers Nat-for y Nat-rev agregaron los sitios de restricción de las enzimas *Nco* I y *Pst* I, lo que permitió ligar el gen *nat-1* en el plásmido pUC57-utr's. Esto fue comprobado mediante la digestión de los plásmidos obtenidos de las clonas transformantes con las enzimas *Nco* I y *Pst* I, estas digestiones liberaron un fragmento de 570 pb que corresponde al tamaño esperado del gen *nat-1* (Fig. 8).

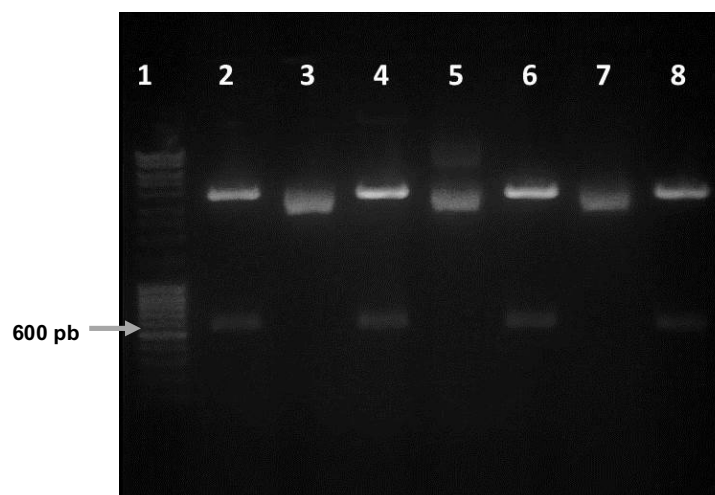


Figura 8. Electroforesis de la digestión para comprobar la clonación del gen *nat-1* en el plásmido pUC57-utr's. Carril 1, marcador de peso molecular; Carriles 2, 4, 6 y 8, digestión con las enzimas *Nco* I y *Pst* I; Carriles 3, 5 y 7, controles de plásmido sin digerir.

5.3. Clonación del gen que codifica para Interleucina-2 en el plásmido pUC57-utr's

Los resultados obtenidos del programa EMBOSS sobre la frecuencia de uso de codones de 79 regiones codificantes de *D. salina* CCAP 19/18 se muestran en la tabla I. Los codones con mayor frecuencia fueron seleccionados para sustituir a sus homólogos en la secuencia del gen *il-2*. La secuencia del gen *il-2* con los codones optimizados para su expresión en el cloroplasto de *D. salina* se muestra en la figura 9.

Tabla I. Frecuencia de uso de codones de *D. salina* CCAP 19/18

Codón	AA*	Fracción	Número de apariciones	Codón	AA*	Fracción	Número de apariciones
GCA	A	0.316	513	CCA	P	0.397	512
GCC	A	0.078	126	CCC	P	0.047	61
GCG	A	0.065	106	CCG	P	0.084	109
GCT	A	0.541	879	CCT	P	0.471	608
TGC	C	0.114	28	CAA	Q	0.917	1185
TGT	C	0.886	217	CAG	Q	0.083	107
GAC	D	0.218	256	AGA	R	0.213	336
GAT	D	0.782	917	AGG	R	0.014	22
GAA	E	0.889	1376	CGA	R	0.133	209
GAG	E	0.111	171	CGC	R	0.072	113
TTC	F	0.276	438	CGG	R	0.008	13
TTT	F	0.724	1147	CGT	R	0.560	881
GGA	G	0.214	385	AGC	S	0.036	91
GGC	G	0.059	106	AGT	S	0.209	534
GGG	G	0.066	118	TCA	S	0.298	762
GGT	G	0.662	1191	TCC	S	0.037	94
CAC	H	0.312	167	TCG	S	0.068	174
CAT	H	0.688	368	TCT	S	0.354	905
ATA	I	0.203	423	ACA	T	0.454	793
ATC	I	0.115	241	ACC	T	0.054	94
ATT	I	0.682	1423	ACG	T	0.067	117
AAA	K	0.921	2337	ACT	T	0.426	744
AAG	K	0.079	201	GTA	V	0.399	640
CTA	L	0.076	227	GTC	V	0.036	58
CTC	L	0.007	21	GTG	V	0.040	64
CTG	L	0.013	38	GTT	V	0.525	841
CTT	L	0.151	449	TGG	W	1.000	352
TTA	L	0.701	2080	TAC	Y	0.197	218
TTG	L	0.052	153	TAT	Y	0.803	889
ATG	M	1.000	536	TAA	*	0.899	71
AAC	N	0.249	517	TAG	*	0.076	6
AAT	N	0.751	1558	TGA	*	0.025	2

La digestión del plásmido pUC57-il2 con las enzimas *Nco* I y *Pst* I liberó un fragmento de 505 pb (Fig. 10). Después de haber realizado la ligación del gen *il-2* dentro del plásmido pUC57-utr's y transformarla en *E. coli* DH5 α , se extrajo plásmido a las clonas que crecieron en las cajas de Petri con medio LB/Carbenicilina. A partir de los plásmidos purificados se realizó una PCR, donde las clonas positivas mostraron la amplificación de un fragmento de 1472 pb, como

se muestra en la figura 11, el cual corresponde en tamaño a las regiones del promotor y los UTR's del gen *atpA* unidos al gen *il-2*.

```
CCATGGCACCACCACCACCACCACGATGACGATGACAAGATGTACCGTATGCAATTATTATCATGTATT
GCTTTATCATTAGCTTTAGTAACTAACTCAGCTCCAAC TTCATCATCAACTAAAAAACTCAATTACAA
TTAGAACACTTATTATTAGACTTACAAATGATTTTAAACGGTATTAACAAC TACAAAAACCCAAAATTA
ACTCGTATGT TAACTTTCAAATTC TACATGCCAAAAAAGCTACTGAATTAAAACACTTACAATGTTTA
GAAGAAGAATTTAAACCATTAGAAGAAGTATTAACTTAGCTCAATCAAAAAACTTCCACTTACGTCCA
CGTGA CTTAATTTCAAACATTAACGTAATTGTATTAGAATTTAAAGGTT CAGAACTACTTTTCATGTGT
GAATACGCTGACGAACTGCTACTATTGTAGAATTCTTAAACCGTTGGATTACTTTCTGTCAATCAATT
ATTTCAACTTTAACTTAACTGCAG
```

El sitio de restricción de las enzimas *Nco* I y *Pst* I se encuentra subrayado

La secuencia de poli histidinas se encuentra marcada en gris

El sitio de reconocimiento de la enzima enterocinasa se encuentra en itálicas

Figura 9. Secuencia del gen *il-2* sintetizado.

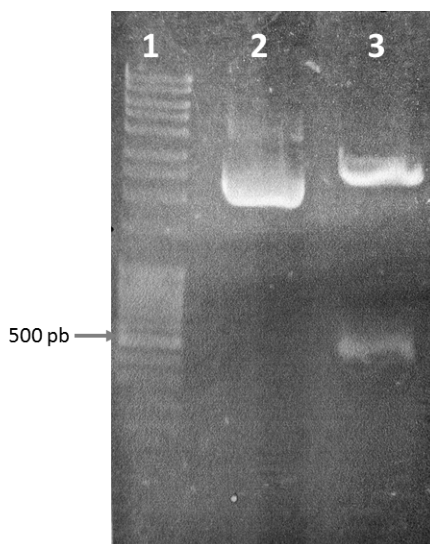


Figura 10. Electroforesis de la digestión el gen *il-2* del plásmido pUC57-*il2*. Carril 1, marcador de peso molecular; Carril 2, plásmido sin digerir; Carril 3, plásmido digerido.

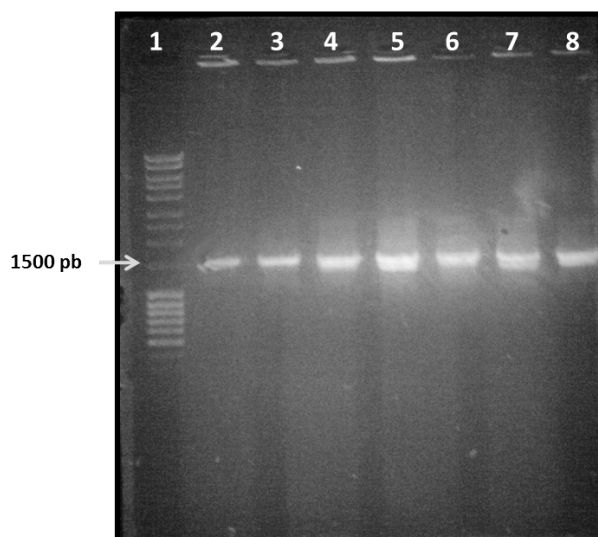


Figura 11. Electroforesis de la PCR para comprobar la clonación del gen *il-2* en el plásmido pUC57-utr's. Carril 1, marcador de peso molecular, Carriles 2-8, PCR de las clonas transformadas.

5.4. Unión de las construcciones IL-2 y Nat-1

Las secuencias del gen *il-2* con su promotor y UTR's fue clonada dentro del vector pNat-utr utilizando las enzimas *Nde* I y *Bst*X I. Para clonar este fragmento dentro del vector pNat-utr, llevo a cabo una digestión con dichas enzimas eliminando un fragmento de 250 pb y la banda correspondiente al plásmido digerido fue recuperada del gel de agarosa (Fig. 12). El gen *il-2* con su promotor y UTR's fueron amplificados mediante PCR con primers que agregaban los sitios de restricción de las enzimas *Nde* I y *Bst*X I. la amplificación de PCR tuvo un tamaño de 1495 pb (Fig.13).

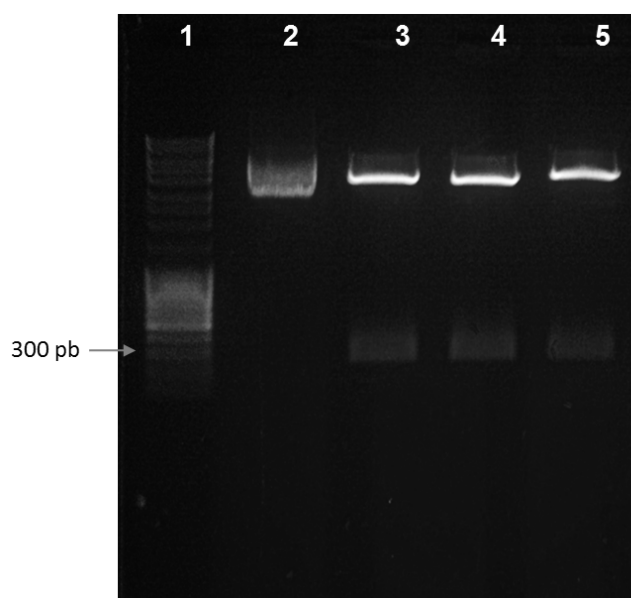


Figura 12. Electroforesis de la digestión de plásmido pNat-utr con las enzimas *Nde* I y *Bst*X I. Carril 1, marcador de peso molecular; Carril 2, plásmido pNat-utr sin digerir; Carriles 3-5, digestiones con las enzimas *Nde*I y *Bst*XI.

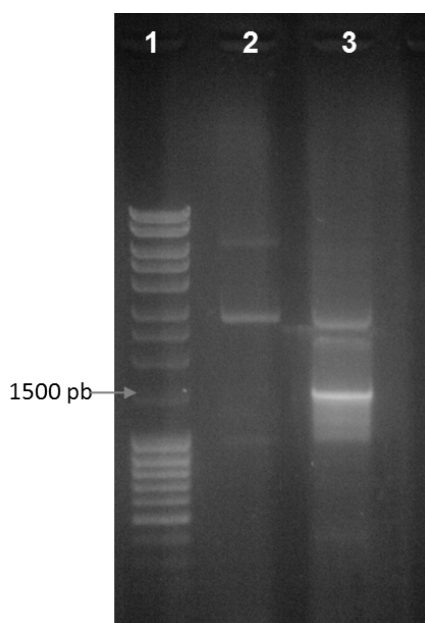


Figura 13. Electroforesis de la PCR para agregar los sitios de restricción de las enzimas *Nde* I y *Bst*X I al gen *il-2* con promotor y UTR's. Carril 1, marcador de peso molecular; Carril 2, control; Carril 3 producto de PCR.

Una vez agregados los sitios de restricción de las enzimas *Nde* I y *Bst*X I mediante PCR, se realizó la digestión enzimática del producto de PCR recuperado del gel de agarosa. Después de las digestiones, del inserto y el plásmido, se realizó la reacción de ligación, la cual fue transformada en *E. coli* y sembrada en medio sólido. Las colonias que aparecieron en medio sólido se crecieron en medio LB líquido para la posterior extracción del plásmido. Para comprobar si esos plásmidos tenían el inserto, se realizó una digestión con las enzimas *Nde* I y *Bst*X I, el resultado de esta digestión se muestra en la figura 12.

Se obtuvo una clona positiva que presentaba el tamaño esperado del inserto (Fig. 14, Carril 2), esta clona fue utilizada para la subsecuente clonación de las regiones de recombinación homólogas.

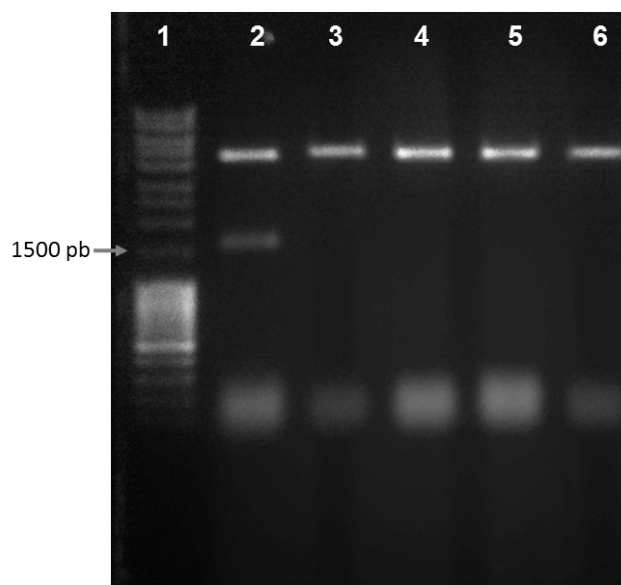


Figura 14. Electroforesis de la digestión del plásmido de las clonas transformadas para comprobar la presencia del inserto. Carril 1, marcador de peso molecular; Carril 2, clona positiva; Carriles 3-6 clonas sin inserto.

5.5. Clonación de las regiones de recombinación homólogas en el plásmido pIL2-Nat

Las regiones de recombinación homóloga fueron obtenidas de DNA total extraído de *D. salina* SQ. La clonación de las regiones de recombinación homóloga dentro del plásmido pIL2-Nat se llevó a cabo utilizando el kit In-Fusion HD, el cual cuenta con una enzima que recombina fragmentos de DNA homólogos. Los primers diseñados para amplificar las regiones de recombinación homóloga que fueron utilizadas para integrar el casete de expresión en el genoma del cloroplasto de *D. salina* contienen una región idéntica a la del plásmido pIL2-Nat. Las amplificaciones con estos primers generaron un producto de 1433 pb para región 5' homóloga mientras que para la región 3' amplificaron un fragmento de 1331 pb (Fig.15).

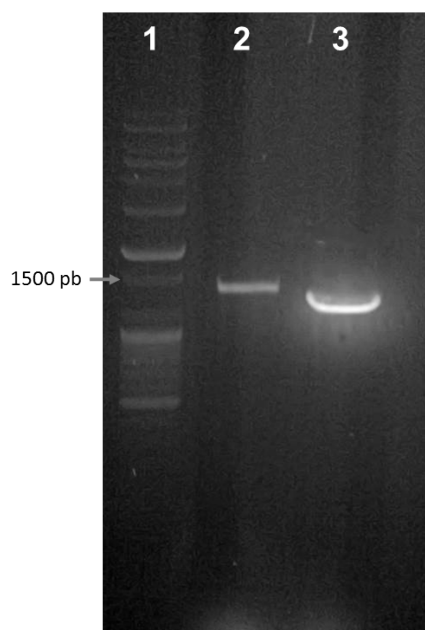


Figura 15. Amplificación de las regiones homólogas 5' y 3'. Carril 1, marcador de peso molecular; Carril 2, región homóloga 5'; Carril 3, región homóloga 3'.

El plásmido pIL2-Nat fue digerido con las enzimas *Hind* III y *Pvu* II para integrar la región homóloga 3', como resultado de esta digestión se eliminó un fragmento de 110 pb. La figura 16 (Carril 3) muestra la doble digestión del plásmido pIL2-Nat. La banda correspondiente al plásmido digerido con las dos enzimas fue purificada para ligar la región 3'homóloga.

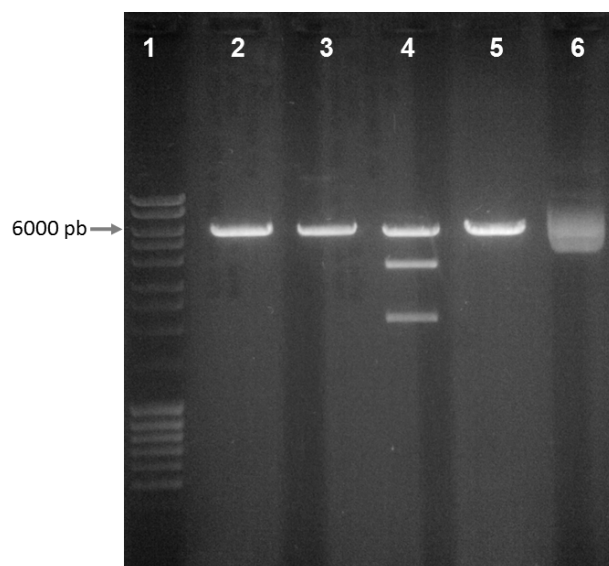


Figura 16. Digestión del plásmido pIL2-Nat para recombinar la región 3' homóloga. Carril 1, marcador de peso molecular; Carril 2, digestión con *Pvu* II; Carril 3, digestión con *Pvu* II y *Hind* III; Carril 4, digestión con *Pvu* I y *Hind* III; Carril 5, digestión con *Hind* III; Carril 6, plásmido sin digerir.

La reacción de recombinación de la región 3' homóloga dentro del plásmido pIL2-Nat fue realizada conforme las especificaciones del kit In-Fusion HD y se obtuvieron clonas transformantes. Se seleccionaron 5 clonas transformadas y se

purificó el plásmido, y se corroboró la presencia de la región 3' homóloga mediante su digestión con la enzima *Bst*XI. Las clonas positivas presentaron dos fragmentos en el gel de agarosa (Fig. 17); el plásmido pIL2-Nat presenta un sitio de reconocimiento para esta enzima y la región 3' homóloga contiene otro sitio. Los tamaños de los fragmentos fueron de 1821 pb y 4762 pb, como se observa en los Carriles 4-8 en la figura 17. El control fue el plásmido pIL2-Nat, como se observa en el Carril 3 de la figura 17; al no existir inserto el plásmido solo se lineariza. La clona 4 fue elegida para continuar con la recombinación de la región 5' homóloga. Este plásmido fue nombrado pIN3'.

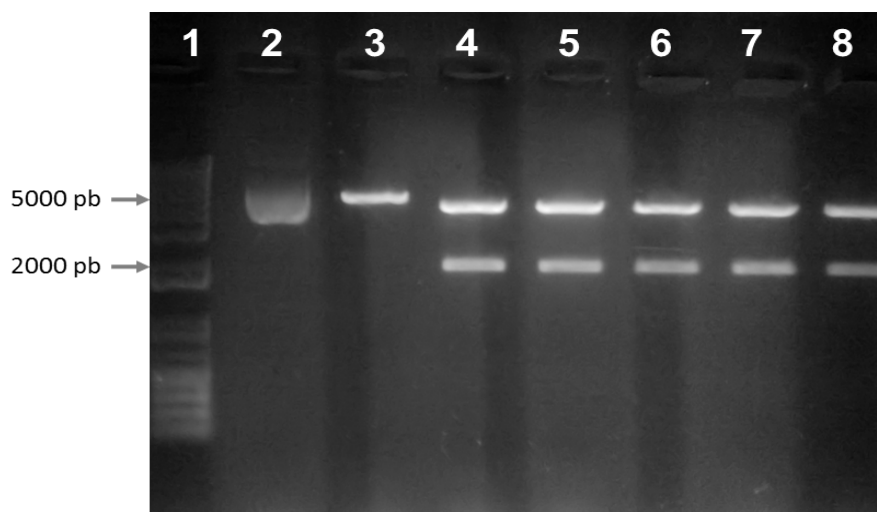


Figura 17. Digestión con la enzima *Bst*XI para comprobar la recombinación de la región 3' homóloga. Carril 1, marcador de peso molecular; Carril 2, plásmido control; Carril 3, digestión del plásmido control; Carriles 4-8, digestión del plásmido de las transformantes positivas.

La región de recombinación homóloga 5' amplificada contiene secuencias homólogas a las de vector pIN3' para recombinarla en el extremo 5', donde se

encuentra el promotor y UTR 5' del gen *il-2*. El plásmido pIN3' fue digerido con la enzima *Nde* I y la reacción fue precipitada sin pasar por el gel de agarosa. Después de amplificar por PCR la región homóloga 5' (Fig. 15, Carril 2) y purificarla, se llevó a cabo la reacción de recombinación con el kit In-Fusion HD. Para comprobar la presencia de la región homóloga 5' homóloga en los plásmidos de las clonas transformadas, se realizó una digestión con la enzima *Nru* I. El plásmido pIN3' tiene solo un sitio de reconocimiento para esta enzima por lo cual solamente se lineariza después de la digestión (Fig. 18, Carril 3) mientras que los plásmidos que contienen la región 5' homóloga presentan dos sitios de restricción, liberando 2 fragmentos uno de 2775 pb y otro de 5247 pb como se puede observar en los Carriles 4 y 5 de la figura 18.

El plásmido de la clona 2 fue enviado a secuenciar. Con esto se comprobó que el casete de expresión que contiene el gen *il-2* y el gen *nat-1* unidos con las regiones del promotor y los UTR's del gen *atpA* de *D. salina*, además de las regiones de recombinación homóloga (Fig. 19). El plásmido resultante fue nombrado pDsCh (Fig. 20).

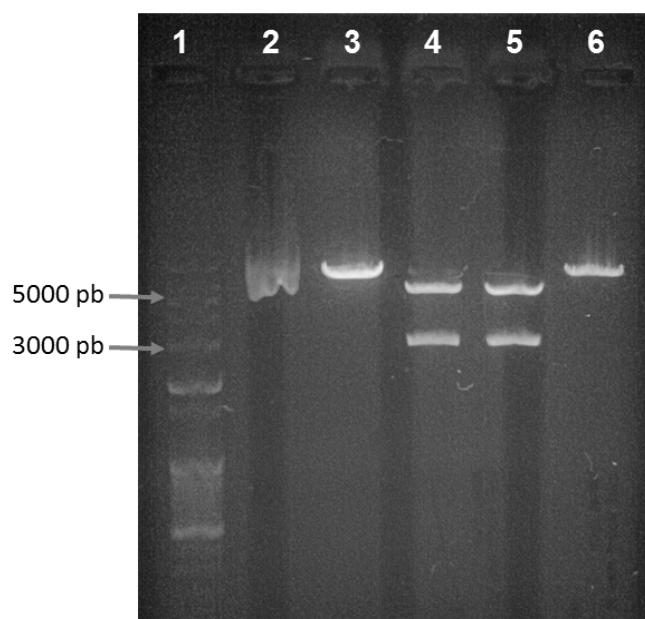


Figura 18. Digestión con la enzima *Nru* I para comprobar la recombinación de la región 5' homóloga. Carril 1, marcador de peso molecular; Carril 2, plásmido control; Carril 3, digestión del plásmido control; Carriles 4-6, digestión del plásmido de las transformantes.

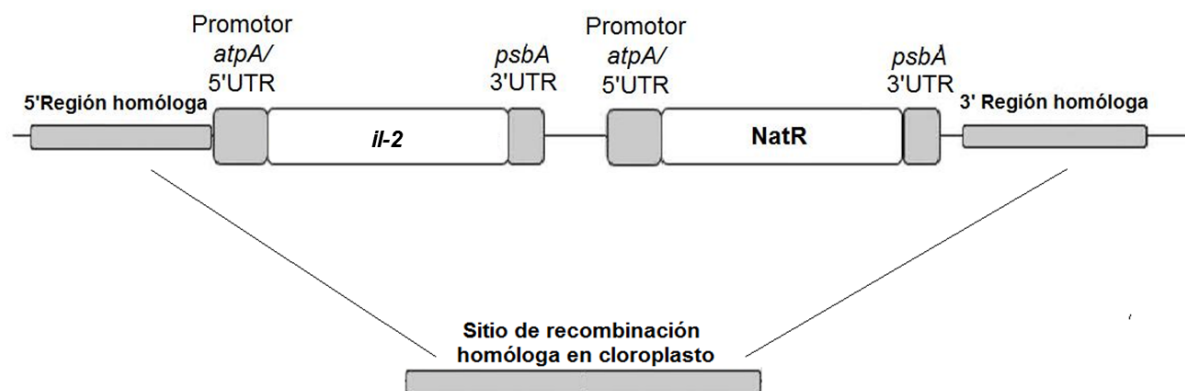


Figura 19. Casete de expresión del plásmido pDsCh.

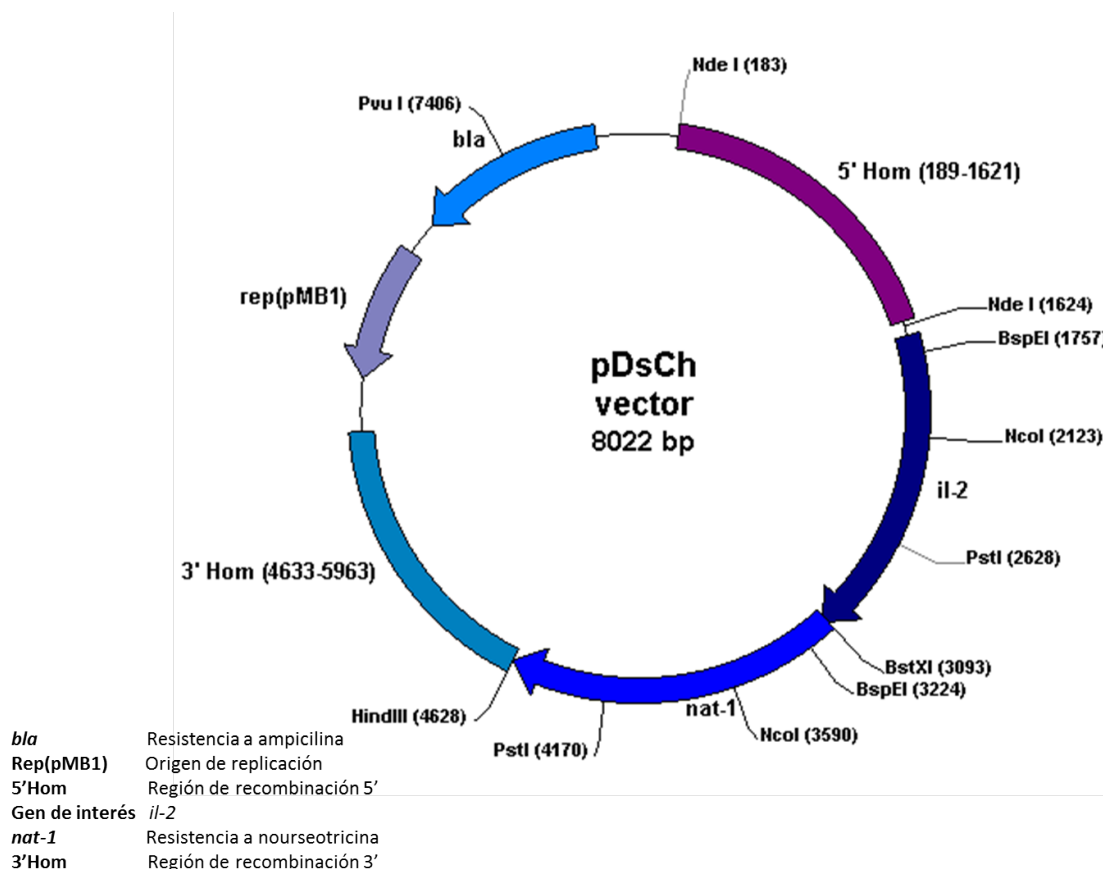


Figura 20. Mapa del plásmido pDsCh.

5.6. Clonación de la región 5' homóloga en el plásmido pEGFP

Después de haber sintetizado el casete que incluía los genes *egfp*, *nat-1* y los nuevos promotores y terminadores, se clonaron las secuencias de recombinación homóloga. La unión se llevó a cabo con el kit de recombinación In-Fusion (Clontech). Para comprobar la correcta integración de la región 5' homóloga se realizó una digestión con la enzima *EcoR* V de los plásmidos obtenidos de las clonas que crecieron en las cajas de LB agar después de la transformación. De las 10 clonas seleccionadas, 7 fueron positivas. Los

plásmidos de las clonas positivas liberaron un fragmento de 1034 pb después de ser digeridas con dicha enzima (Fig. 21). El plásmido obtenido fue nombrado pGFP5.

5.7. Clonación de la región 3' homóloga en el vector pGFP5

Una vez clonada la región 5' homóloga se procedió a clonar la región 3' homóloga. Se eligió una de las clonas positivas y en ese plásmido se unió la región homóloga 3'. La estrategia de integración fue la misma que en la región homóloga 5'. Después de la transformación se tomaron 5 clonas y los plásmidos de estas clonas fueron digeridos con las enzimas *Nru* I y *EcoR* I. Como control negativo se utilizó el plásmido que contenía solo la región homóloga 5'; cuando éste plásmido fue digerido con ambas enzimas se produjeron dos fragmentos, uno de 1027 pb y otro de 6588, mientras que los plásmidos positivos presentaron 3 fragmentos, 1027 pb, 2722pb y 3886 pb (Fig. 22); esto se debe a que la región de recombinación homóloga 3' contiene un sitio de reconocimiento de la enzima *EcoR* I.

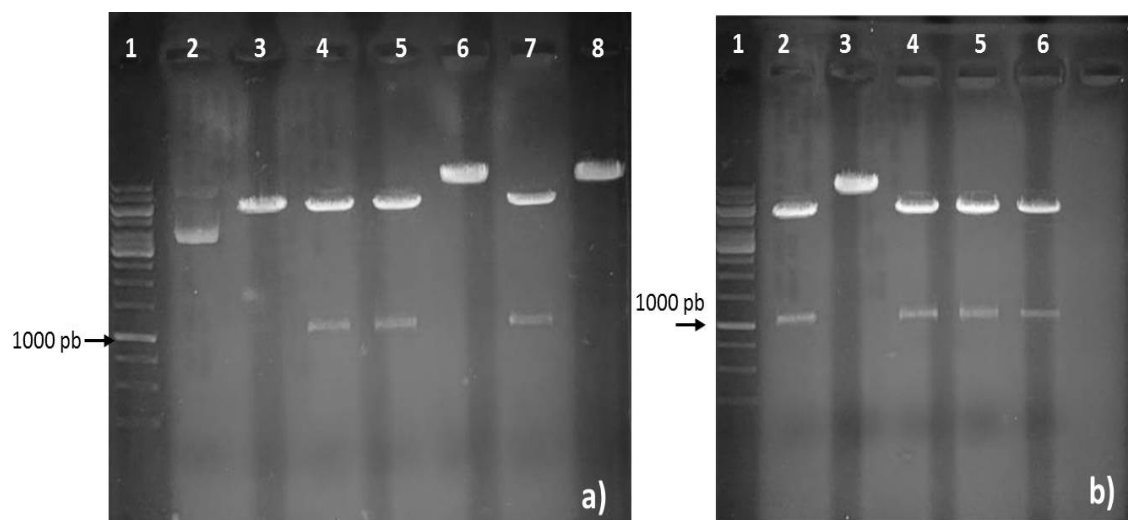


Figura 21. Digestión con la enzima *EcoR* V para comprobar la unión de la región 5' homóloga. a) Carril 1, marcador de peso molecular; Carril 2, plásmido sin digerir; Carril 3, plásmido sin inserto digerido; Carriles 4-8, plásmidos digeridos de las clonas transformadas. b) Carril 1, marcador de peso molecular; Carriles 2-6 plásmidos digeridos de las clonas transformadas.

La clona 2 (Fig. 22, Carril 5) fue enviada a secuenciar para corroborar que ambas regiones de recombinación estaban presentes en el plásmido. El plásmido fue nombrado pDsEGFP (Fig. 23).

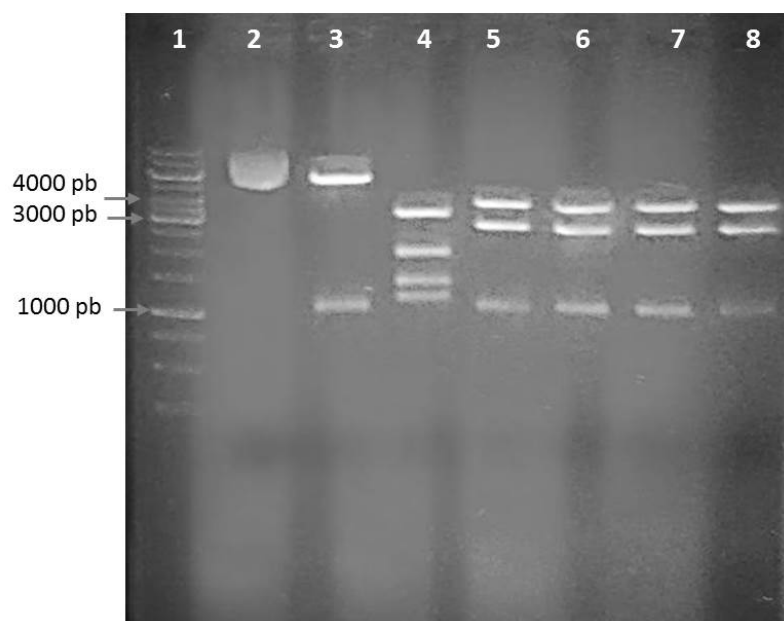


Figura 22. Digestión con las enzimas *Nru* I y *EcoR* I para comprobar la unión de la región 3' homóloga. Carril 1, marcador de peso molecular; Carril 2, plásmido control sin digerir; Carril 3, plásmido control digerido; Carriles 4-8, clonas transformadas.

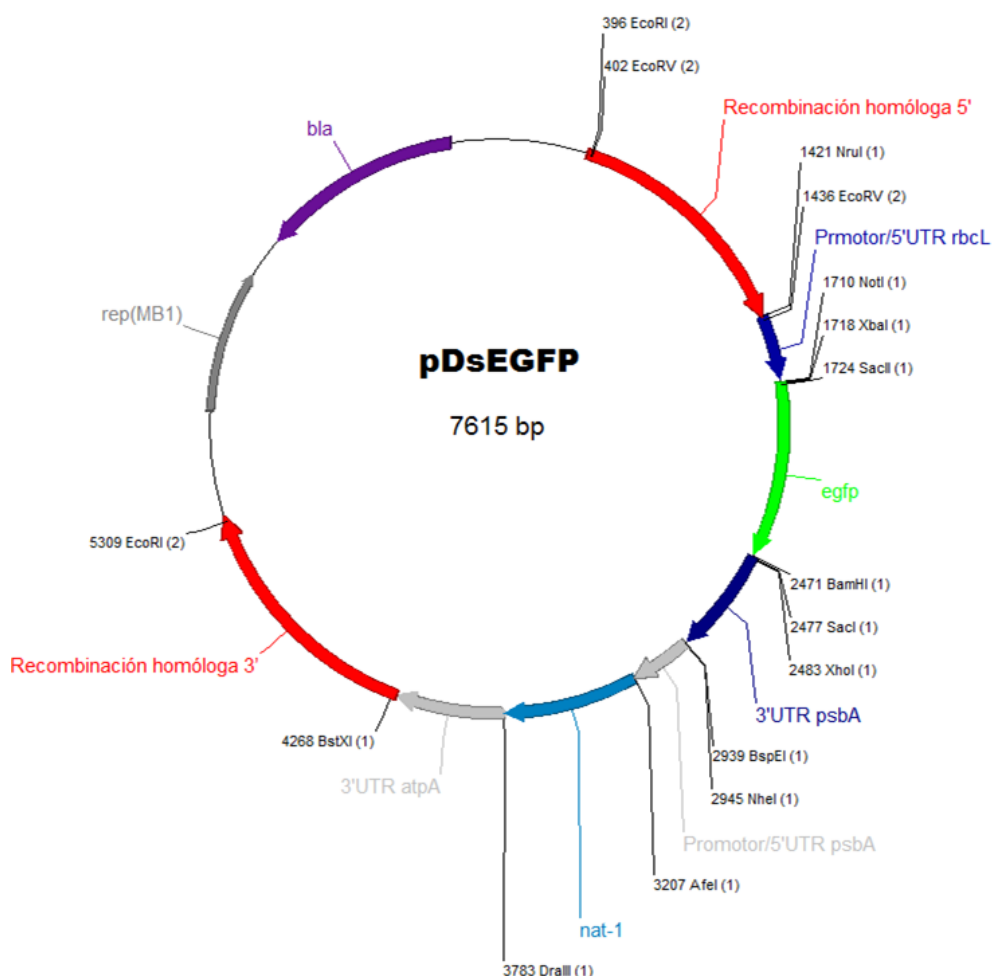


Figura. 23. Mapa del plásmido pDsEGFP. Se muestran cada uno de los elementos que lo componen, así como los sitios de restricción de las enzimas que separan cada elemento.

5.8. Transformación del plásmido pDsCh

Las clonas que crecieron en las placas de medio modificado agar con nourseotricina fueron sometidas a PCR para comprobar la presencia del casete dentro de *D. salina*. La primera amplificación se hizo para corroborar que el DNA estuviera en buen estado, por lo que se utilizaron los primers control

Atpa2F/Atpa2R y atpaF/atpaR (Fig. 24). La amplificación fue positiva dando como resultado una banda de aproximadamente 750 pb con los primers Atpa2F/Atpa2R y 482 pb con los primers atpaF/atpaR. Para comprobar la integración del casete en *D. salina* se amplificó con los primer NatF/NatR, que amplifican el gen *nat* que confiere la resistencia a nourseotricina (Fig. 25). La amplificación fue positiva para ambas clonas (3 y 4) dando una amplificación aproximada 576 pb, pero esta amplificación también fue positiva para la clona sin transformar por lo que se asumió que había contaminación de los reactivos. Se hicieron varias pruebas para detectar la fuente de contaminación y se determinó que fueron las micropipetas que estaban contaminadas con DNA plasmídico. Una vez eliminada la fuente de contaminación se volvió a realizar la misma PCR (Fig. 26) y las clonas potencialmente transformadas arrojaron resultados negativos, por lo que se determinó que la transformación no fue exitosa.

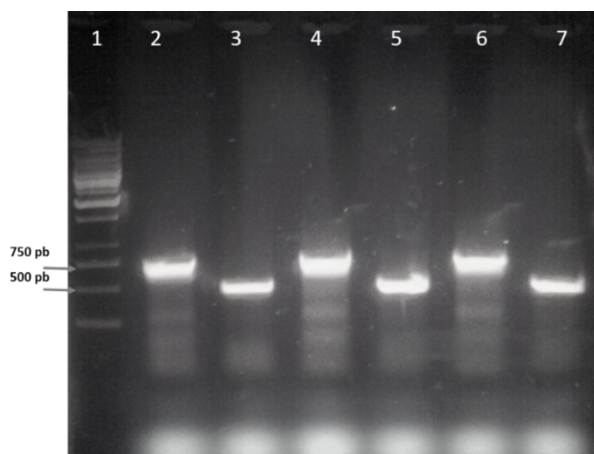


Figura 24. Amplificación con los primers Atpa2F/Atpa2R y atpaF/atpaR. Carril 1, marcador de peso molecular; Carril 2, *D. salina* sin transformar primers

Atpa2F/Atpa2R; Carril 3 *D. salina* sin transformar primers atpaF/atpaR; Carril 4, clona 3 trasformada primers Atpa2F/Atpa2R; Carril 5, clona 3 trasformada primers atpaF/atpaR; Carril 6, clona 4 trasformada primers Atpa2F/Atpa2R; Carril 7, clona 4 trasformada primers atpaF/atpaR.

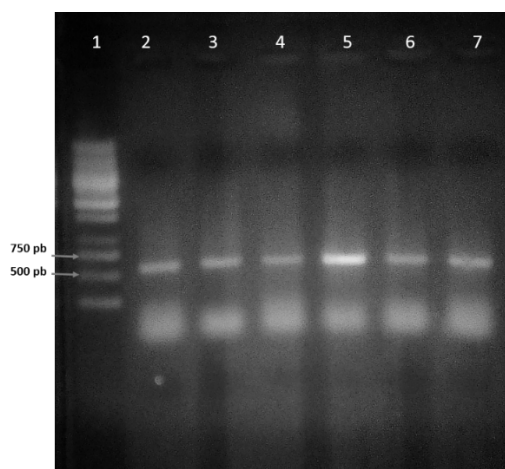


Figura 25. Amplificación con los primers NatF/NatR. Carril 1, marcador de peso molecular; Carriles 2 -3, *D. salina* sin transformar; Carriles 4-5, clona 3 trasformada; Carriles 6-7, clona 4 trasformada.

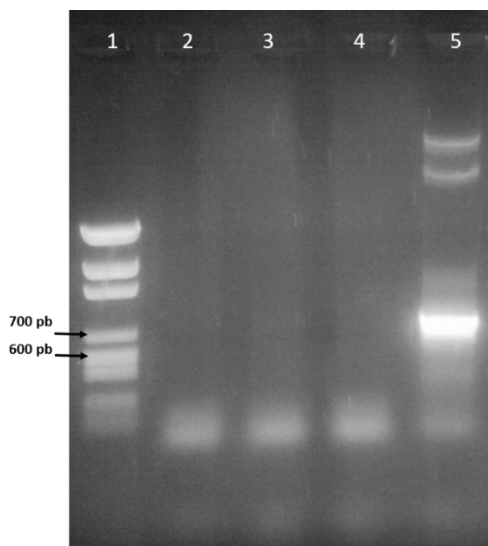


Figura 26. Nueva amplificación con los primers NatF/NatR. Carril 1, marcador de peso molecular; Carriles 2, *D. salina* sin transformar; Carril 3, clona 3 trasformada; Carril 4, clona 4 trasformada y Carril 5, amplificación del plásmido pDsCh.

5.9. Transformación del plásmido pDsEGFP

Las clonas transformadas obtenidas en las placas de medio modificado se crecieron en medio líquido hasta alcanzar la fase logarítmica (Fig. 27), se hizo extracción DNA para detectar si el casete se había integrado en *D. salina* SQ. La PCR control para verificar la integridad del DNA extraído fue exitosa, por lo que se procedió a hacer la amplificación con los primer InF/InR. Como control negativo se utilizó la cepa sin transformar y como control positivo se utilizó el plásmido pDsEGFP, la amplificación esperada era de 1622 pb. Como se puede apreciar en la figura 28, en el Carril del control negativo no aparece ninguna banda como era de esperarse, sin embargo, tampoco aparece en las clonas transformadas, a diferencia del control positivo donde sí se observa la banda de

1622 pb. El procedimiento de transformación se repitió varias veces siempre dando los mismos resultados.

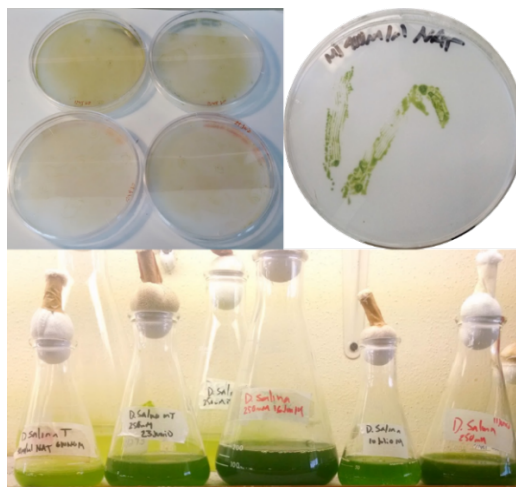


Figura 27. Clonas transformadas con el plásmido pDsEGFP.

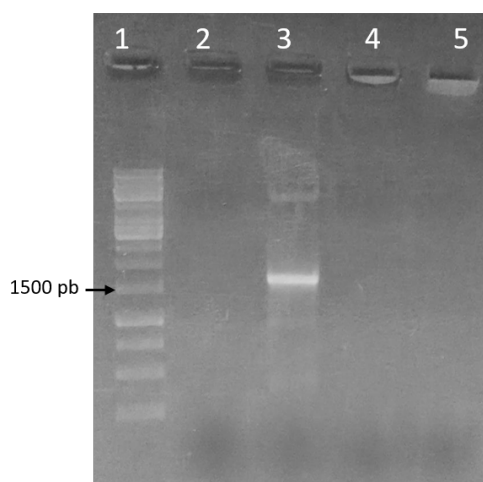


Figura 28. Amplificación con los primers InF/InR. Carril 1, marcador de peso molecular; Carril 2, *D. salina* sin transformar; Carril 3, plásmido control; Carril 4, clona 1 transformada y Carril 5, clona 2 transformada.

En este trabajo se utilizaron dos métodos de transformación y dos plásmidos diferentes, utilizando el mismo marcador de selección. En ninguno de los casos se logró obtener clones de *D. salina* que presentaran el transgen integrado. La alta presencia de clones en el medio con antibiótico fue un factor delimitante en este trabajo, ya que al crecer colonias en medio con antibiótico se asumió que eran cepas que contenían el plásmido recombinante, ya sea integrado en el genoma o bien como episoma. Una vez que se tiene una cepa resistente con el plásmido integrado, el siguiente paso es hacer cultivos subsecuentes para fijar esta característica, lo cual en el caso de microalgas lleva por lo menos 30 días por subcultivo. Previamente se había determinado que el antibiótico nourseotricina podía ser implementado como agente de selección en *D. salina*, sin embargo, se demostró que nuestra cepa es resistente a este antibiótico. La falta de un agente de selección que funcione adecuadamente es una limitante cuando se está trabajando con ingeniería genética, para solucionar este problema se decidió buscar un agente de selección que fuera altamente efectivo, que inhiba completamente el crecimiento de *D. salina*.

6. Conclusiones

Se logró la construcción de dos plásmidos para expresión en cloroplasto de *D. salina*. El plásmido pDsCh contiene el gen que codifica para interleucina 2 de humano, el gen de resistencia al antibiótico nourseotricina, ambos genes regulados por el promotor del gen *atpA* de *D. salina* y las regiones de recombinación homóloga correspondientes a una secuencia intergénica entre los genes *psbA* y *psaA*; y el plásmido pDsEGFP, en el que se encuentra clonado el gen que codifica para la proteína fluorescente EGFP optimizado, regulado por el promotor y 5' UTR del gen *rbcL* y el 3' UTR del gen *psbA*, y el gen de resistencia al antibiótico nourseotricina optimizado, regulado por el promotor y 5' UTR del gen *psbA* y el 3' UTR del gen *atpA*.

No se detectó la integración del casete transfectado en el genoma del cloroplasto de la clonas que crecieron en medio con nourseotricina, por lo cual se concluye que *D. salina* SQ es resistente al antibiótico nourseotricina y este no puede ser utilizado como agente de selección.

CAPÍTULO II

Norflurazon como agente de selección en la microalga *D.*
salina SQ

1. Introducción

1.1. Antibióticos como marcadores de selección en *D. salina*

Los protocolos de transformación requieren el uso de marcadores de selección que permitan discernir las células transformadas de las no transformadas. La disponibilidad de un protocolo fiable de selección de transformantes es un prerequisite para la creación de nuevos sistemas de expresión (Jung y Kim, 2016). Los marcadores de selección más utilizados en microalgas son genes que confieren resistencia a un antibiótico o herbicida, aunque también están aquellos marcadores que restituyen alguna función en células mutantes, por ejemplo, la habilidad de realizar fotosíntesis. La microalga *D. salina* es naturalmente resistente a varios antibióticos que normalmente son utilizados para la selección de transformantes en plantas (Akbari et al, 2014), por esta razón es necesario encontrar un marcador de selección efectivo que permita la selección de transformantes en *D. salina*.

Geng y colaboradores (2003) demostraron que *D. salina* era resistente a los antibióticos estreptomycin, kanamicina, higromicina y G418; mientras que el cloranfenicol a una concentración de 60 µg/ml inhibía su crecimiento. Tan et al (2005) hicieron pruebas de sensibilidad con los mismos antibióticos más la adición de espectinomycin y zeocina, obteniendo como resultados la resistencia a espectinomycin y la inhibición con zeocina (100 µg/ml) y cloranfenicol (400 µg/ml). Estudios preliminares realizados por nuestro grupo de trabajo, encontraron que la cepa *D. salina* SQ (utilizada en este trabajo) es resistente a

cloranfenicol ya que a una concentración de 400 µg/ml la cepa aún continuó creciendo, mientras que el antibiótico tetraciclina no logró inhibir su crecimiento (datos no publicados). El antibiótico nourseotricina es utilizado como agente de selección en la diatomea *Phaeodactylum tricornutum*, inhibiendo su crecimiento a una concentración de 50 µg/ml (Zaslavskaya et al, 2000). Inicialmente en nuestro grupo de trabajo se realizaron pruebas de resistencia con este antibiótico en *D. salina* SQ y se obtuvo que a una concentración de 100 µg/ml inhibía su crecimiento. Posteriormente, el gen *nat-1* fue utilizado como marcador de selección para la creación un casete de expresión de proteínas recombinantes en el cloroplasto de *D. salina*, sin embargo, después de varias rondas de transformación se presentaban clones resistentes, pero no contenían el gen de resistencia integrado en su genoma. Se diseñaron nuevos experimentos para analizar la resistencia de *D. salina* SQ al antibiótico nourseotricina y el resultado obtenido fue que a una concentración 800 µg/ml aun había crecimiento, por lo que se determinó que este antibiótico no es un agente que permite la selección de transformantes.

1.2. Herbicidas como marcadores de selección en *D. salina*

Una alternativa a los antibióticos es el uso de herbicidas como agentes de selección. La resistencia a herbicidas se confiere de dos formas: mediante la creación de cepas mutantes, en las cuales se reemplaza un gen silvestre por uno mutado o bien integrando un gen que confiere resistencia. Fosfinotricina es un herbicida que inhibe a la enzima glutamina sintasa, la cual cataliza la formación

de L-glutamina a partir de L-glutamato y está involucrada en la asimilación de amoníaco (Horlein, 1994). Este compuesto ha sido usado como agente de selección en transformantes nucleares de *D. salina* (Li et al, 2007; Wang et al, 2007; Feng et al, 2009; Li et al, 2010; Jia et al, 2012; Feng et al, 2014b; 2014c; Jung y Kim et al, 2016), el plásmido de transformación contiene el gen *bar* o *pat*, que codifican para la enzima Fosfinotricina acetil transferasa (PPT), encargada de la conversión de la fosfinotricina en N-acetilfosfinotricina, un compuesto no tóxico para las células.

El herbicida norflurazon (4-cloro-5-(metilamino)-2-(α,α,α -trifluoro-m-tolil)-3(2H) piridazinona) interfiere con la síntesis de carotenos inhibiendo a la enzima fitoeno desaturasa (PDS). Esta enzima convierte el fitoeno en zeta-carotenos, los cuales eventualmente darán origen a el resto de los carotenos (Fig. 29). El tratamiento de plantas y algas con norflurazon da como resultado el blanqueamiento de las células (Chamovitz et al, 1991), ya que una de las funciones de los carotenos es proteger a las células fotosintéticas de la fotoinhibición (Ben-Amotz et al, 1989).

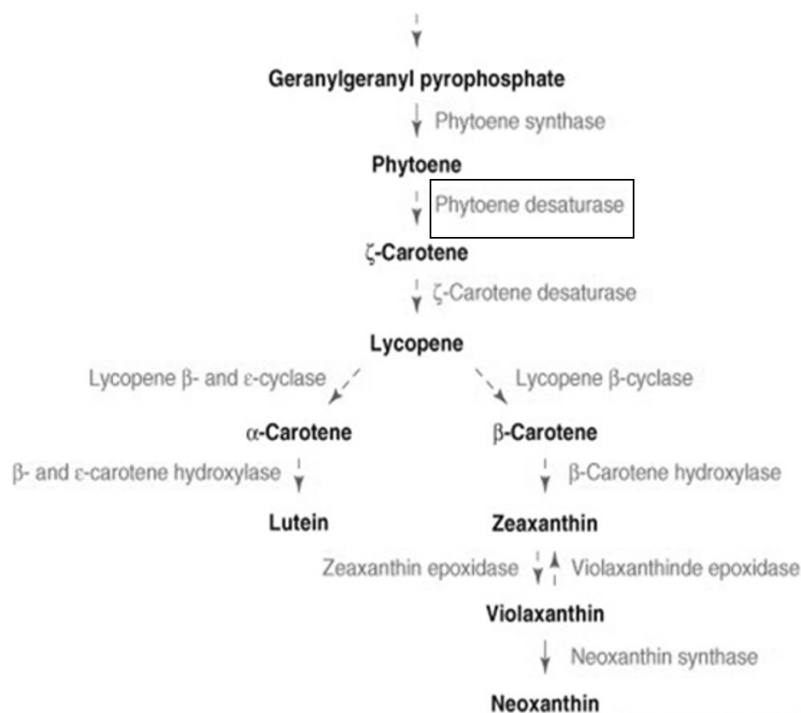


Figura 29. Ruta de síntesis de carotenos en *D. salina*. La enzima fitoeno desaturasa se indica dentro del recuadro. Modificado de Lamers et al, 2008.

El herbicida norflurazon está siendo estudiado como un agente de selección ya que la mayoría de los organismos fotosintéticos son susceptibles a este herbicida. Existen diferentes estudios en los cuales se ha demostrado que mutaciones en el gen que codifica para la enzima PDS confieren resistencia al norflurazon. Tal es el caso de *Chlorella zofingiensis*, donde el aminoácido leucina en la posición 516 fue cambiado por fenilalanina (L516F) (Liu et al, 2014), para *Haematococcus pluvialis* fue en la posición 504 (L504R) (Sharon-Gojman et al, 2015), mientras que en *C. reinhardtii* fue la mutación F131V (Suarez et al, 2014). El Dr. Suarez (Facultad de Ciencias, UABC) también está trabajando con la

misma mutación en la enzima PDS en *D. salina* CCAP 19/30, la cual difiere de posición, siendo F128V (comunicación personal).

El gen que codifica para la enzima PDS F128V de *D. salina* CCAP 19/30 fue donado para comprobar si este gen puede ser utilizado como marcador de selección en la creación de casetes de expresión de proteínas recombinantes en el cloroplasto de *D. salina* SQ.

2. Hipótesis

El herbicida norflurazon puede ser utilizado como agente de selección mediante la resistencia generada por el gen que codifica para la enzima PDSF128V de *D. salina*.

3. Objetivo

Comprobar si el herbicida norflurazon puede ser utilizado como agente de selección en *D. salina* SQ para la producción de proteínas recombinantes

3.1. Objetivos particulares

- Determinar la concentración a la cual el norflurazon inhibe el crecimiento de *D. salina* SQ
- Clonar el gen *pdsF128V* en el vector pDsEGFP
- Transformar el plásmido en *D. salina*

4. Metodología

4.1. Susceptibilidad de *D. salina* SQ al herbicida norflurazon

Se creció un cultivo de *D. salina* SQ en medio modificado hasta que alcanzó fase logarítmica y se utilizó para inocular una placa de 96 pozos con una absorbancia de 0.4-0.5 (630 nm). El diseño de la placa se muestra en la figura 30. El herbicida norflurazon (Pestanal ® Sigma) fue preparado en una solución stock a 10 mM en etanol al 60% el cual se diluyó a una concentración de 10 μ M con agua destilada estéril.

En cada uno de los pozos se inoculó un volumen final de 150 μ l. Los pozos A1, A12, B12, H11 y H12 se dejaron vacíos. Los pozos A2-A11 se inocularon con medio de cultivo; como controles se utilizó el cultivo sin norflurazon (B1-B10), el cultivo con etanol al 60 % para comprobar que no afectaba el crecimiento (B11, C1-C9), la cantidad de etanol fue la misma que se utilizó en las pruebas de sensibilidad. Se evaluaron diferentes concentraciones de norflurazon, 0.1 μ M (C10-D7), 0.2 μ M (D8-E5), 0.5 μ M (E6-F3), 1 μ M (F4-G2), 2 μ M (G3-G12) y 5 μ M (H1-H10). La placa fue incubada en luz intensa constante, el efecto del norflurazon en el crecimiento fue evaluado diariamente durante un periodo de 15 días a 630 nm en lector de placas de ELISA (Stat Fax 2100, Awareness Technology, Inc)



Figura 30. Diseño de la placa de cultivo para comprobar la susceptibilidad de *D. salina* al herbicida norflurazon.

4.2. Ligación del gen *pds* F128V en el plásmido pDsEGFP

El Dr. Julio V. Suarez (Facultad de Ciencias, UABC) donó el plásmido pDunaCp en el cual se encuentra clonado el gen *pdsF128V*, que codifica para la enzima Fitoeno desaturasa con la mutación F128V, la cual confiere resistencia la herbicida norflurazon.

El gen *pdsF128V* fue amplificado a partir del vector pDunaCp con los primers PdsAfe-F y PdsDra-R. Estos primers fueron diseñados con 15 bases homólogas al plásmido pDsEGFP entre las regiones Promotor/5' UTR *psbA* y 3' UTR *atpA* lo anterior con el propósito de clonar el gen *pdsF128V* en lugar del gen *nat-1* con el kit Infusion (Clontech). La reacción de amplificación contenía 1X de buffer Phusion HF PCR buffer, 0.2 mM de dNTP's, 0.5 µM de cada primer (PdsAfe-F y PdsDra-R), 150 ng de plásmido pDunaCp y 1 U de Phusion DNA polimerasa, en un volumen final de 50 µl. Las condiciones de reacción fueron 30 s a 98 °C, 30 ciclos de 98 °C 10 s, 66 °C 30s y 72 °C 48 s, seguidos de un ciclo

10 min a 72 °C. Una fracción de la reacción de PCR se evaluó en un gel de agarosas al 1% y el resto del producto de PCR fue purificado utilizando el kit NucleoSpin Gel and PCR Clean-up (Clontech).

El plásmido pDsEGFP fue digerido con las enzimas *Dra* III y *Afe* I (NewEngland, Biolabs) para remover el gen *nat-1*. La reacción de digestión contenía 1X de NEBuffer 4, 1.5 µg de plásmido, 20 U de *Dra* III, 10 u de *Afe* I y 18.3 µl de agua destilada para obtener un volumen final de 25 µl. La reacción se incubó 3 h a 37 °C. y posteriormente se corrió en un gel de agarosa al 1 %. La banda correspondiente al plásmido digerido fue purificada utilizando el kit GeneJET Gel Extraction (ThermoFisher).

Los productos de PCR y el plásmido purificado fueron utilizados para realizar la reacción de ligación la cual se llevó a cabo utilizando el kit In-Fusion HD (Clontech); siguiendo las instrucciones incluidas. Posteriormente se transformaron 2.5 µl de la reacción de ligación en células competentes DH5α y se sembraron en placas de LB agar con 50 µg/ml de carbenicilina. Se seleccionaron la clonas que crecieron en las placas de agar, se inocularon en medio LB líquido con 50 µg/ml de carbenicilina y se purificó el plásmido utilizando el kit GeneJET Plasmid Miniprep (ThermoFisher), siguiendo las instrucciones proporcionadas.

Para comprobar que se había clonado el gen pdsF128V, se realizó una digestión con la enzima *Mfe* I, ya que existe un sitio de reconocimiento de esta enzima dentro del gen pdsF128V. Las reacciones de digestión contenían 1X Buffer CutSmart, 1 µg de plásmido, 10 U *Mfe* I y se aforó con agua destilada

hasta obtener un volumen de 20 µl. Las reacciones se incubaron 2 h a 37 °C y posteriormente se evaluaron en un gel de agarosa al 1 %. Una de las clonas positivas se envió a secuenciar (Eton Bioscience, San Diego, CA), para comprobar la secuencia del gen *pdsF128V*. Los primer diseñados para la secuenciación fueron PdsF, PdsR, PdsF2 y PdsR2. El plásmido obtenido fue nombrado pGFP/PDS.

4.3. Transformación con el plásmido pGFP/PDS

La transformación se realizó con micro esferas de vidrio de 425 – 600 µm de diámetro. El cultivo de *D. salina* SQ se creció hasta alcanzar fase logarítmica y se centrifugaron 100 -150 ml del cultivo a 5000 rpm por 2 min, se retiró el medio y el paquete celular se lavó tres veces con 25 ml de medio de cultivo centrifugando a 5000 rpm por 2 min. El paquete de células se resuspendió en 800 µl de medio obteniendo una densidad de 1×10^5 cel/ml y posteriormente se le agregó 300 mg de microesferas, 60 µg del plásmido y 100 µl de PEG 20%. La mezcla se agitó en vórtex a 2400 rpm durante 15 s y después se agitó de nuevo durante 60 s. Después de la sedimentación de las microesferas, la mezcla se incubó por 24 h con poca luz y posteriormente se sembró en placas con medio modificado con 5 µM de norflurazon. Las placas se incubaron con luz intensa constante a 28 ° C durante 30 días.

5. Resultados y discusión

5.1. Resistencia de *D. salina* SQ al herbicida norflurazon

Los resultados obtenidos del crecimiento de *D. salina* SQ bajo diferentes concentraciones de norflurazon se muestran en la figura 31. En la gráfica se observa que el control con etanol (línea gris), en el cual se disuelve el herbicida, no afectó el crecimiento de *D. salina*. Mientras que las concentraciones de 0.1, 0.2 y 0.5 μM de norflurazon no afectaron el crecimiento de la microalga, con 1 μM se retrasó el crecimiento, restableciéndose al cuarto hasta que alcanzó una densidad similar a la del control sin herbicida y a la de las concentraciones 0.1, 0.2 y 0.5 μM . En el caso de la concentración de 2 μM se observó que a partir del séptimo día, el crecimiento de la microalga empezó a aumentar, cabe mencionar que esta concentración fue suficiente para inhibir el crecimiento de la microalga *C. reinhardtii* (Suarez et al, 2014). Al utilizar una concentración de 5 μM de norflurazon se observó que la microalga no fue capaz de crecer e incluso la absorbancia disminuyó a través del tiempo. El decremento los valores de absorbancia se relaciona con la inhibición de la ruta de síntesis de carotenos por el herbicida dando como resultado el fotoblanqueamiento de las células. La microalga *D. salina* no es capaz de utilizar fuentes de carbono orgánico como alimento por lo que el fotoblanqueamiento resulta en la muerte de esta microalga (Ben-Amotz et al, 1989).

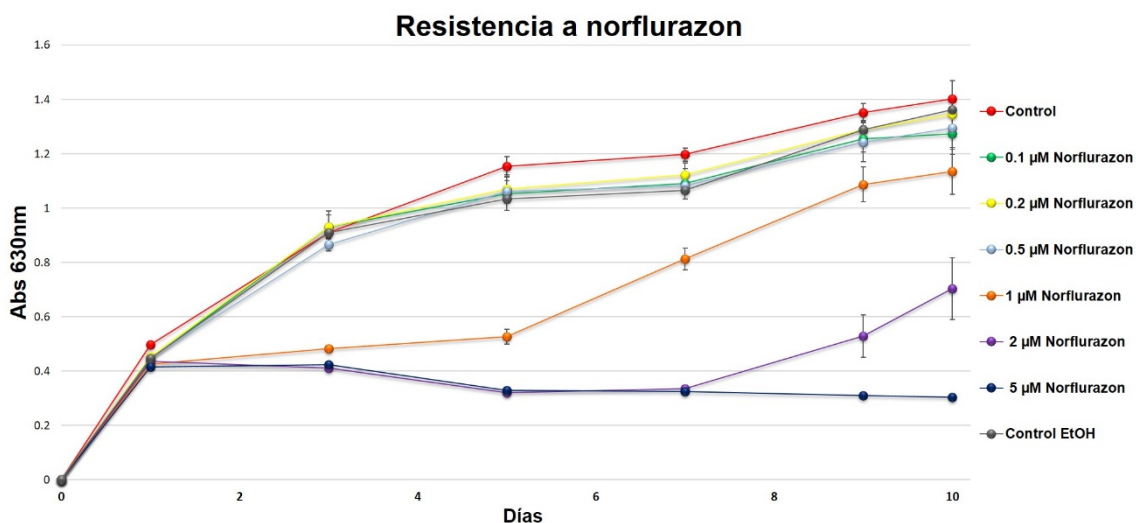


Figura 31. Efecto de diferentes concentraciones del herbicida norflurazon en el crecimiento de *D. salina*. Como controles se utilizó el cultivo sin norflurazon y el cultivo con etanol.

Uno de los principales problemas al generar nuevos sistemas de expresión de proteínas recombinantes es encontrar un marcador de selección adecuado para el organismo que se está utilizando. Se han probado diferentes antibióticos en *D. salina* y se sabe que es naturalmente resistente a estreptomicina, kanamicina, higromicina y G418 (Akari et al, 2014). En algunos trabajos han utilizado cloranfenicol como agente de selección; tal es el caso de Geng y colaboradores (2003) quienes demostraron que una concentración de 60 µg/ml les permite seleccionar a clonas de *D. salina* que fueron transformadas con un plásmido que contenía el gen *cat* (cloranfenicol acetil transferasa). Sin embargo, Tan y colaboradores (2005) encontraron que la concentración de cloranfenicol que inhibe el crecimiento de *D. salina* es 400 µg/ml. En estudios recientes, reportan que la concentración necesaria de cloranfenicol para eliminar a *D. salina*

en un periodo de 6 días es de 1500 µg/ml (Jung y Kim, 2016). Previamente a este trabajo, la cepa con que se trabajó en esta tesis fue sometida a varias pruebas de resistencia a antibióticos, presentó resistencia a los antibióticos ya mencionados, pero también se determinó que era resistente a cloranfenicol y a tetraciclina (datos no publicados). Adicionalmente en este trabajo se estableció que también era resistente al antibiótico nourseotricina.

La discrepancia en resultados sobre la resistencia a cloranfenicol puede deberse a que se utilizaron diferentes cepas, en los trabajos de Sun y Tan, no mencionan la cepa que utilizan, Jung y Kim utilizaron para sus ensayos la cepa KMMCC-1064 obtenida del Centro de Cultivo de Microalgas Marinas en Corea y en nuestro caso utilizamos la cepa SQ aislada en Baja California.

5.2. Ligación del gen *pdsF128V* en el plásmido pDsEGFP

La amplificación del gen *pdsF128V* con un tamaño de alrededor de 1780 pb se presenta en la figura 30. El resto del producto de PCR fue purificado y se obtuvo 750 ng de DNA. La digestión del plásmido pDsEGFP con las enzimas *Dra* III y *Afe* I, liberó un fragmento de 570 pb correspondiente al tamaño del gen *nat-1* (Fig. 32).

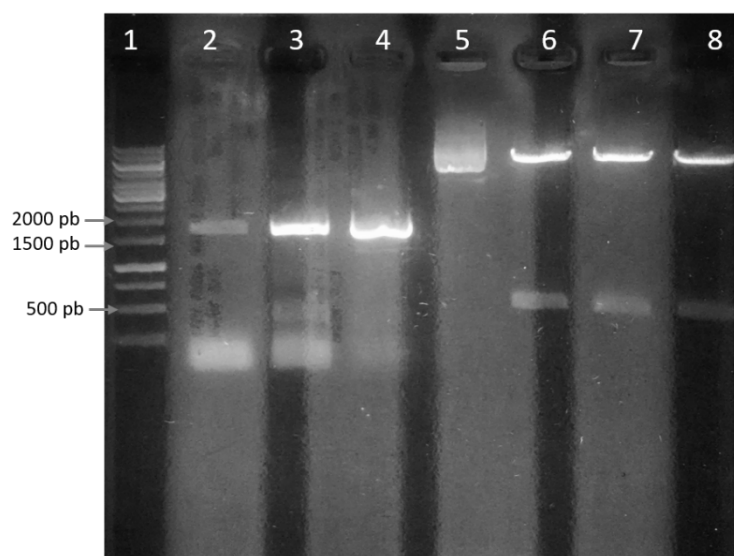


Figura 32. Amplificación del gen *pdsF128V* y digestión del vector pDsEGFP con las enzimas *Dra* III/*Afe* I. Carril 1, marcador de peso molecular; Carriles 2-4, amplificación del gen *pdsF128V*; Carril 5, plásmido pDsEGFP sin digerir; Carriles 6-8, plásmido pDsEGFP digerido.

Después de las reacciones de recombinación y la transformación, se purificó el plásmido de 4 clonas que crecieron en las placas con antibiótico. Para comprobar que se encontraba clonado el gen *pdsF128V* se hizo una digestión con la enzima *Mfe* I, ya que el plásmido pDsEGFP no contiene ningún sitio de tiene sitio de restricción interno para esta enzima. En la figura 33 se puede observar el plásmido linearizado debido a la presencia del sitio de restricción de la enzima *Mfe* I, lo que demostró la inserción de una secuencia que tiene un sitio de reconocimiento para esta enzima, probablemente la del gen *pdsF128V*.

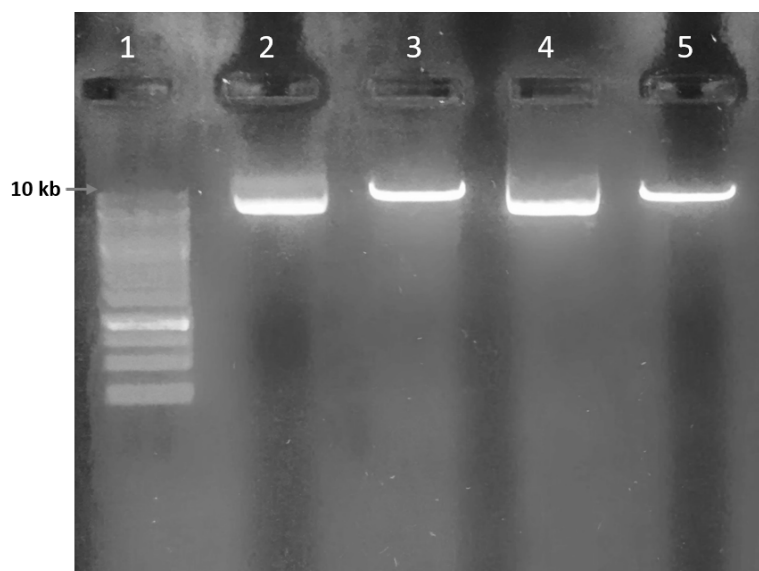


Figura 33. Digestión con la enzima *Mfe* I. Carril 1, marcador de peso molecular; Carril 2, plásmido clona 1 sin digerir; Carril 3, plásmido clona 1 digerido; Carril 4, plásmido clona 2 sin digerir; Carril 5, plásmido clona 2 digerido.

Para corroborar que la secuencia clonada era del gen *pdsF128V*, se envió a secuenciar el plásmido de la clona 1. La secuencia enviada corresponde a la al gen *pdsF128V*, en la figura 34 se muestra la secuencia de los 582 aminoácidos de la proteína PDSF128V y la mutación F128V se encuentra subrayada, la secuencia nativa de la enzima PDS nativa de *D. salina* se encuentra registrada en la base de datos NCBI con el número de acceso ABB51091. El plásmido resultante fue nombrado pGFP/PDS y tiene un tamaño de 8797 pb (Fig. 35).

MQVMQGRAHTQAAYS RQPAQQRTQRRVGR TTRLQVYARDFPAPQFDGTETYQEAVL
 STKLQNA PRPVKPQ RVVIAGAGLAGLSAAKYLSDAGHIPVVLEARDVLGGKVAAWKD
 EDGDWYETGLHIFVGAYPNMQRLFKELNISDRLQWKSHSMIFAMQDKPGEFSRFEFP
 DIPAPWNGVIVILRNNEMLSWPEKIQFAIGLLPAIIFGQKYVEEQDELTVTQWMQKQ
 GVPSRVNHEVFIAMAKALNFINPDELSMTVVLTALNRFLQERHGSKMAFLDGAPPER
 LCQPMVDHFTSKGGELRMNARIKQIVLNEDNSVKHFELLNGEIVEGDVYMSAMPVDI
 MKKLMPQPWKGMPFFQKLNGLEGVPVINIHIWFDRKLSTVDHLLFSRSDLLSVYADM
 STTCKEYTD DKANMLELVFAPAAEWIGRPDEEIVDATMKELEKLPNEIKADQSLAK
 IRKSKVIKTPLSVCKSTAGREKYRPSQKTPISNFYLAGDYTKQKYLASMEGAVFSGK
 LACEQVVEDASTRTGSQSAASSQPALAAASA AVLLAMGAVLAGNASAQVLTEAIWGS
 EVLTEAVTFPWF*

Figura 34. Secuencia de aminoácidos de la enzima Fitoeno desaturasa F128V

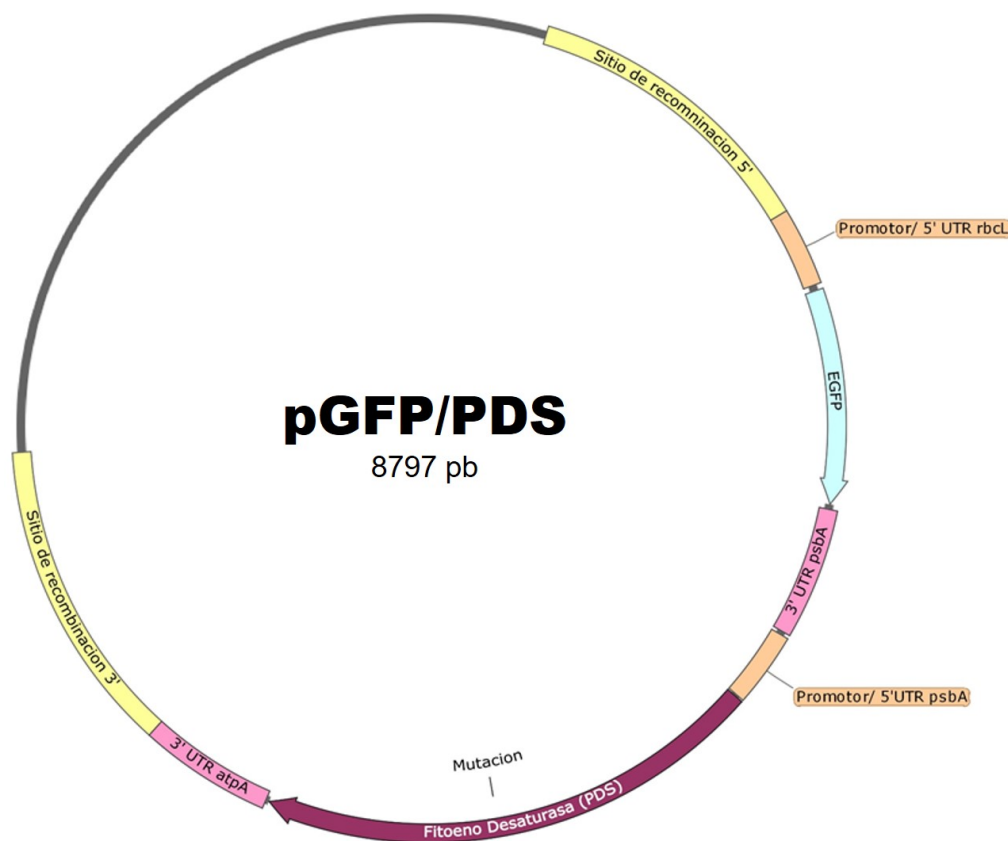


Figura 35. Mapa de vector pGFP/PDS. Amarillo, sitios de recombinación homóloga; naranja, promotor y 5' UTR; rosa, 3' UTR; gen *egfp*, verde y morado, fitoeno desaturasa con la mutación F128V.

5.3. Transformación con el vector pGFP/PDS

Las placas con las células utilizadas para la transformación permanecieron en la incubadora por 30 días con luz alta constante sin presentar crecimiento de ninguna colonia. Algunas cajas permanecieron por 45 días bajo estas condiciones, pero tampoco presentaron ninguna colonia. El proceso de transformación se llevó a cabo en varias ocasiones, pero los resultados fueron siempre los mismos. Esto se puede deber a que el proceso de transformación no es el adecuado para esta microalga, sin embargo, este método está reportado por Feng (et al, 2009). También se hicieron modificaciones del protocolo, se aumentó la velocidad del vórtex y se probaron diferentes concentraciones de PEG, pero el resultado fue el mismo.

En los capítulos I y II se utilizaron dos métodos de transformación y tres plásmidos diferentes sin embargo en ninguno caso se logró obtener clones de *D. salina* SQ que presentaran el transgen integrado. Al no haber presencia de clones transformantes, surgen varias preguntas sobre a qué se debe este resultado, analizando los métodos de transformación, sabemos que el método de biobalística es ampliamente utilizado para la ingeniería de cloroplastos, en *C. reinhardtii* es el método preferido de transformación de cloroplasto mientras que en *D. salina* ha sido utilizado exitosamente para generar transformantes nucleares (Tan et al, 2005). El método de transformación mediante esferas de vidrio también ha sido previamente utilizado en *D. salina* para generar

transformantes nucleares (Feng et al, 2009) y es una alternativa económica para la transformación de cloroplastos de microalgas, como se ha comprobado en *C. reinhardtii* (Economou et al, 2014). Ambos métodos han generado transformantes nucleares de *D. salina* y en *C. reinhardtii* producen transformantes de cloroplasto; una manera de determinar si en nuestra cepa está funcionando el método de transformación es buscar la presencia del plásmido recombinante después de 24 h de transformación en vez de hacer estos análisis después de varios pases en medio de cultivo con antibiótico. Adicionalmente a estos dos métodos, la electroporación también ha sido utilizada en *D. salina* (Geng et al, 2003; Sun et al, 2005) para transformaciones en núcleo, aunque los rendimientos han sido menores que los utilizados en otras áreas por lo que los aseguran que se necesita estandarizar este protocolo. La electroporación no es utilizada comúnmente para la transformación de cloroplasto de microalgas, pero existen reportes en donde ha sido un protocolo exitoso (Xie et al, 2014), por lo que no se descarta que este método pueda ser utilizado en *D. salina* SQ.

Si se asume que el método de transformación permite la entrada del plásmido recombinante, el siguiente paso sería determinar la razón por la cual no se está llevando a cabo la recombinación homóloga. En la diatomea *P. tricornutum* se creó un sistema de transformación eficiente, el éxito se atribuye a las regiones de recombinación seleccionadas, el transgen está dirigido para recombinarse en la región intergénica transcripcionalmente activa entre los genes *trnI-trnA*, ubicados en las regiones repetitivas invertidas, por lo cual se

integra en dos regiones del genoma (Xie et al, 2014). La combinación entre una zona de recombinación intergénica transcripcionalmente activa y la doble integración podría ser una solución para promover la recombinación homóloga en *D. salina* SQ, ya que las regiones utilizadas en este trabajo carecen de regiones codificantes y están alejadas de los genes que las flanquean. El uso de regiones transcripcionalmente activas ofrece ciertas ventajas, porque permiten la creación de transgenes que carezcan de las regiones 5' o 3' UTR o incluso de promotores, por lo que se podría crear un sistema que funcione como un policistrón, de tal manera que además de servir de región de recombinación también podamos utilizar el promotor/UTR de este gen para que se lleve a cabo la transcripción del gen recombinante.

Las condiciones de crecimiento correctas después de la transformación son un factor importante a la hora de obtener clonas en organismos donde se están estableciendo nuevos sistemas de expresión es necesario estandarizar estas condiciones. En *D. tertiolecta* se dificultó el crecimiento de las células transformadas, no hubo crecimiento con los parámetros previos a la transformación sino hasta que fueron modificadas (Georgianna et al, 2013). En el presente trabajo, las condiciones de cultivo después de la transformación fueron diferentes a las previas, se utilizó luz excesiva para generar presión y que solo las células resistentes a norflurazon pudieran sobrevivir, en el trabajo de Georgianna y colaboradores utilizaron cámaras de CO₂ con humidificador para poder obtener clonas transformadas, modificaciones como estas y medios

suplementados con vitaminas pudieran favorecer el crecimiento de las clonas de *D. salina* SQ.

Otro factor importante fue el agente de selección, previamente se había determinado que el antibiótico nourseotricina podía ser implementado como agente de selección en *D. salina*, sin embargo, se demostró que nuestra cepa es resistente a este antibiótico. Esto se pudo observar debido a la gran presencia de clonas que crecían en medio con antibiótico, que en realidad eran falsos positivos. La falta de un agente de selección que funcione adecuadamente es una limitante cuando se está trabajando con ingeniería genética, para solucionar este problema se decidió buscar un agente de selección que fuera altamente efectivo, que inhiba completamente el crecimiento de *D. salina* SQ. En las pruebas realizadas en este capítulo, se demostró que el norflurazon puede ser utilizado como agente de selección, ya que a una concentración de 5 μ M, inhibe completamente el crecimiento de esta microalga.

A pesar de que *D. salina* es un buen candidato para la expresión de proteínas recombinantes, aun se carece de información suficiente para desarrollar este sistema. Como se ha mencionado anteriormente, el conocer el genoma de la especie que se desea modificar es fundamental para realizar los casetes de expresión con promotores y secuencias reguladoras endógenas, el plastoma de *D. salina* CCAP 19/18 se encuentra secuenciado y en base a esto se diseñaron los diferentes plásmidos. En trabajos previos donde se intentó

crear un casete para expresar la hemaglutinina del virus de la influenza A(H1N1) en la cepa SQ (Mendoza-Damián, 2012), como región de recombinación se intentó utilizar una secuencia que incluía el gen *psbA* modificado para resistir al herbicida basta, sin embargo, cuando se hicieron las amplificaciones para aislar esta secuencia de la cepa SQ, los tamaños de las amplificaciones fueron distintos al esperado, una vez secuenciados se evidenció que la estructura de este gen es diferente en la cepa SQ, ya que no se presentaba un intrón como estaba reportado en el genoma de CCAP 19/18. El genoma nuclear de *D. salina* aún no está secuenciado a pesar que el departamento de energía de Estados Unidos está trabajando en el desde el año 2006 (genome.jgi.doe.gov/DunsalCCAP1918/ DunsalCCAP1918.info.html) y solo se cuenta con el genoma del cloroplasto de la cepa CCAP 19/18, por lo cual no se puede determinar si la variación genética que presenta el gen *psbA* se extiende en todo el genoma y si estas diferencias podrían afectar la selección de promotores y UTR's.

6. Conclusiones

Se determinó que una concentración 5 μ M es suficiente para evitar el crecimiento de *D. salina* SQ, por lo cual, el herbicida norflurazon puede ser utilizado como agente de selección en *D. salina* SQ.

El gen que codifica para la enzima fitoeno desaturasa F128V de *D. salina* CCAP 19/30 fue clonado en el plásmido pDsEGFP reemplazando al gen *nat-1*. El nuevo plásmido pGFP/PDS fue utilizado para transformar *D. salina* SQ, sin poder detectar transformantes debido a ausencia de clones.

CAPÍTULO III

Anotación y análisis del genoma del cloroplasto de *D. salina* SQ

1. Introducción

1.1. Estructura de los genomas de cloroplasto

Los cloroplastos son orgánulos capaces de llevar a cabo la fotosíntesis en plantas y algas, proceso mediante el cual los organismos obtienen energía y carbono para llevar a cabo la síntesis de almidón, aminoácidos, pigmentos y ácidos grasos. Los cloroplastos se originaron por endosimbiosis alrededor de 1.5 mil millones de años, evolucionando de bacterias fotosintéticas que vivieron en los ancestros del linaje de algas y plantas (Adem et al, 2017). Debido a su origen, los cloroplastos poseen su propio genoma. El genoma de los cloroplastos modernos contiene alrededor de 55 – 130 genes, los cuales codifican principalmente para la maquinaria de expresión y el aparato fotosintético. El resto de las proteínas que necesita el cloroplasto se encuentran codificadas en el genoma nuclear y son importadas post-traduccionalmente (Jensen y Leister 2014).

La mayoría de los genomas de cloroplastos tienen una organización conservada, presentan una estructura circular dividida en cuatro regiones; una región grande única (Large Single Copy, LSC) y una región pequeña única (Small Single Copy, SSC), las cuales se dividen por dos regiones repetidas (Sablok et al, 2016). Los tamaños de estas regiones varían dependiendo de la especie.

Para entender el funcionamiento del cloroplasto, su contenido genético y los mecanismos involucrados en la expresión génica, se empezaron a secuenciar los genomas completos de estos orgánulos de diferentes organismos. El primer

genoma de cloroplasto secuenciado fue el de la planta de tabaco (*Nicotiana tabacum*) (Shinozaki et al, 1986) y en la actualidad se encuentran alrededor 1520 plastomas reportados en GenBank (<https://www.ncbi.nlm.nih.gov/genome/browse/?report=5>); de los cuales 92 corresponden a algas verdes. El avance en las técnicas de secuenciación (Next-Generation Sequencing; NGS) ha permitido un gran avance en el campo de la genética de cloroplastos.

1.2. Aplicaciones de los genomas de cloroplasto

La información obtenida de la secuenciación de los genomas de cloroplasto es utilizada en diferentes áreas de estudio. Un área importante ha sido la evolución, en la cual esta información se utiliza para realizar estudios filogenéticos, identificación molecular y diversidad genética en plantas y algas (Lemieux et al, 2016; Fučíková et al, 2016; Ng et al, 2017; Zhang et al, 2016; Choi y Park 2015). Otro campo importante es la ingeniería genética, el conocimiento de los genomas completos de los cloroplastos facilita el avance de esta área de estudio. (Gan et al, 2016).

La producción de proteínas recombinantes en los cloroplastos de microalgas depende directamente de la disponibilidad de la secuencia del genoma, ya que la integración de casetes de expresión se da mediante recombinación homóloga (Daniell et al, 2016) (Fig. 36). La integración se diseña para que ocurra en regiones intergénicas evitando la disrupción de genes. Otro punto importante

para llevar a cabo la expresión de genes foráneos en el cloroplasto es necesario utilizar regiones reguladoras y promotores nativos de cada microalga. Por estas razones es esencial conocer el genoma del cloroplasto, ya que esto nos provee la información de las regiones reguladoras, promotores y los sitios de recombinación.

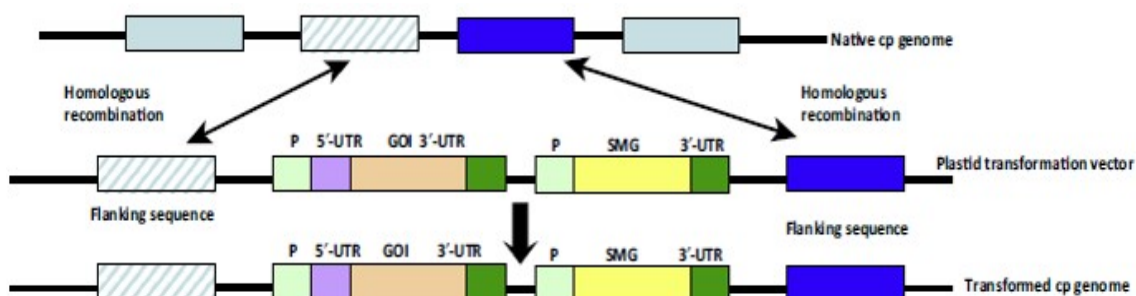


Figura 36. Representación esquemática de la integración de transgenes en cloroplasto mediante recombinación homóloga (Tomado de Jin y Daniell 2015).

1.3. Genoma del cloroplasto de otras cepas de *D. salina*

El primer genoma secuenciado del cloroplasto de la microalga *D. salina* fue reportado por Smith et al, en el 2010. La cepa secuenciada en este trabajo fue CCAP 19/18 aislada de la laguna hipersalina Hutt que se encuentra en el oeste de Australia. El genoma presentó un tamaño de 269 kb con una estructura cuadripartita, compuesto de dos regiones grandes invertidas y las regiones LSC (Large Single Copy) y SSC (Small Single Copy). La secuencia del genoma de

esta cepa se encuentra depositada en GenBank con el número de acceso GQ250046 (Fig. 37). Este genoma está compuesto por 102 genes de los cuales 5 se encuentran repetidos, ignorando las repeticiones en total son 66 genes que codifican para proteínas, 3 genes para RNA ribosomal y 28 genes para tRNA's. Este trabajo demostró que el tamaño tan grande del genoma del cloroplasto se debía a la presencia de grandes regiones de DNA no codificante, más que a un gran número de genes. Los autores de este trabajo remarcan la importancia de conocer la secuencia del genoma de *D. salina* ya que esto permite modificarla genéticamente para la producción de proteínas recombinantes.

Posteriormente, en el 2015 se publicó la secuencia del genoma de la mitocondria de *D. salina* CONC-001 (Del Vasto et al, 2015), esta cepa fue aislada de la laguna hipersalina La Rinconada ubicada en el desierto de Atacama al norte de Chile.

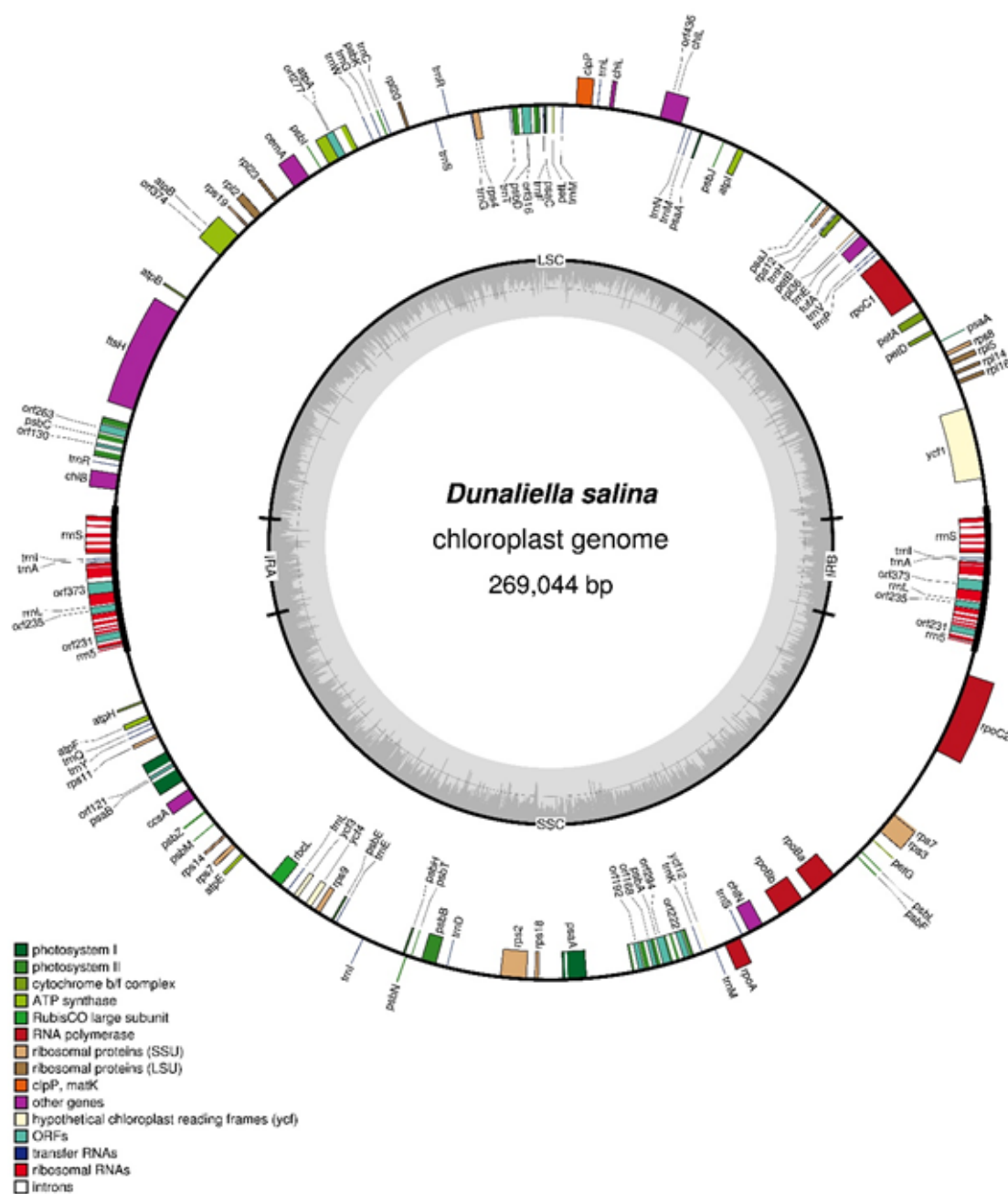


Figura 37. Mapa del genoma de *Dunaliella salina* CCAP 19/18. El mapa fue creado con el software OrganellarGenomeDRAW.

En este trabajo utilizaron los datos obtenidos de esta secuenciación para tratar de ensamblar el genoma del cloroplasto de esta cepa. Los contigs obtenidos tienen un alto contenido de AT, con genes que presentaron intrones y un alto contenido de regiones repetitivas ubicadas en regiones no codificantes. Debido a estas regiones repetitivas les fue difícil ensamblar el genoma completo, por lo que solo pudieron obtener contigs de 0.5 – 8.9 kb, pero con estos datos estiman que el tamaño del genoma del cloroplasto de *D. salina* CONC-001 es de 370 kb aproximadamente, pudiendo incluso tener un mayor tamaño. Ellos consideran que sus datos corresponden con los del genoma previamente secuenciado, siendo uno de los más grandes reportado y que contiene una gran cantidad de regiones repetitivas en regiones no codificantes. Como no lograron ensamblar el genoma completo, solo muestran una lista de los genes que lograron identificar (Fig. 38).

La secuencia del genoma de *D. salina* CCAP 19/18 y la secuencia parcial del genoma de *D. salina* CONC-001, demuestran que los plastomas de esta especie de microalga contienen un gran número de regiones repetitivas, lo que dificulta su análisis. Las diferencias en las regiones intergénicas impiden que el genoma se pueda ensamblar utilizando como referencia el previamente reportado. Estas diferencias pueden afectar si tratamos de utilizar regiones no codificantes, como promotores y sitios de recombinación intergénicos de una cepa de *D. salina* para crear casetes de expresión de proteínas recombinantes en otra cepa de esta microalga. Por esta razón se decidió secuenciar el genoma de *D. salina* SQ, para generar información que permita crear una plataforma de

expresión de proteínas recombinantes para esta cepa en particular. En esta sección se presenta estructura del genoma de *D. salina* SQ y se compara con el genoma previamente reportado de la cepa *D. salina* CCAP 19/18 y las posibles implicaciones en el diseño de casetes para la expresión de proteínas recombinantes.

Part B. Plastid DNA-located genes		Part B. Plastid DNA-located genes	
PtDNA locus	# codons	PtDNA locus	# codons
<i>atpA</i>	503	<i>psbF</i>	43
<i>atpB</i>	479	<i>psbH</i>	82
<i>atpE</i>	134	<i>psbI</i>	36
<i>atpF</i>	173	<i>psbJ</i>	41
<i>atpH</i>	81	<i>psbK</i>	45
<i>atpI</i>	238	<i>psbL</i>	37
<i>ccsA</i>	396	<i>psbM</i>	36
<i>cemA</i>	593	<i>psbN</i>	43
<i>chlL</i>	290	<i>psbT</i>	30
<i>petA</i>	287	<i>psbZ</i>	61
<i>petB</i>	214	<i>rpl2</i>	274
<i>petD</i>	159	<i>rpl5</i>	180
<i>petG</i>	36	<i>rpl14</i>	118
<i>petL</i>	31	<i>rpl16</i>	136
<i>psaA</i>	711	<i>rpl20</i>	111
<i>psaB</i>	734	<i>rpl23</i>	115
<i>psaC</i>	80	<i>rpl36</i>	36
<i>psaJ</i>	40	<i>rps4</i>	251
<i>psbA</i>	352	<i>rps8</i>	143
<i>psbB</i>	507	<i>rps9</i>	195
<i>psbC</i>	461	<i>rps11</i>	129
<i>psbD</i>	351	<i>rps12</i>	128
<i>psbE</i>	81	<i>rps18</i>	139
		<i>rps19</i>	91
		<i>tufA</i>	417
		<i>ycf3</i>	170
		<i>ycf4</i>	219
		<i>ycf12</i>	32

Figura 38. Lista de los genes de cloroplasto encontrados en los contigs generados en la secuenciación del genoma de la cepa *D. salina* CONC-001.

2. Hipótesis

La secuencia del genoma cloroplasto de *D. salina* SQ aislada de las lagunas hipersalinas de San Quintín, Baja California, México es diferente a la secuencia del genoma de *D. salina* CCAP 19/18 aislada de la laguna hipersalina Hutt, Australia.

3. Objetivo

Anotación y análisis del genoma de *D. salina* SQ

3.1. Objetivos particulares

- Anotación de genoma del cloroplasto de *D. salina* SQ.
- Comparación del genoma del cloroplasto de la cepa *D. salina* SQ (México) con el genoma reportado de la cepa *D. salina* CCAP 19/18 (Australia).
- Comparación de las regiones promotor, 5'UTR y 3' de los genes *rbcL* y *psbA* y la región 3' UTR de los genes *atpA* y *psbA*

4. Metodología

4.1. Recolección del espécimen

La cepa de *D. salina* fue obtenida de las lagunas hipersalinas en la costa de San Quintín, Ensenada, Baja California, México; específicamente en las coordenadas 30° 32' 13.91" N, 116° 2' 1.22" O. La muestra fue transportada al laboratorio Meredith Gould, Facultad de Ciencias, UABC y posteriormente se aisló una sola célula utilizando un microscopio modificado para microinyección.

4.2. Extracción de DNA de cloroplastos

La microalga *D. salina* se creció en medio modificado suplementado con 250 mM de NaCl en iluminación constante a una temperatura ambiente. Se inoculó un cultivo de 500 ml, el cual fue incubado en oscuridad total 48 h antes de alcanzar la fase logarítmica. Posteriormente se centrifugó para recuperar las células a 3000 rpm durante 30 min. El medio fue removido y las células se homogenizaron durante 5 minutos en 30 ml de buffer de aislamiento frío (1.25 M NaCl, 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, 0.1% BSA (w/v), 0.1% β -mercaptoethanol (v/v)). El homogeneizado se filtró utilizando una doble capa de la membrana Miracloth (Merck) y el filtrado fue colectado en tubos de 50 ml, posteriormente, se centrifugó 10 minutos a 3000 rpm y se resuspendió el pellet de cloroplastos en 10 ml de buffer de aislamiento frío. Se centrifugó nuevamente durante 10 minutos a 3000 rpm y el paquete de cloroplastos se resuspendió en 10 ml de H₂O destilada estéril (Shi et al, 2012). El DNA de los cloroplastos se

purificó con el kit AxyPrep Multisource Genomic DNA miniprep (Axygen) siguiendo las instrucciones para cultivos celulares

4.3. Secuenciación

El DNA de cloroplasto fue enviado a secuenciar UCSD (San Diego, CA, EU). La secuenciación se llevó a cabo utilizando la tecnología de Illumina MiSeq. Los datos de secuenciación fueron descargados del sitio Basespace de Illumina (<https://basespace.illumina.com/apps/>) y fueron ensamblados utilizando A5 miseq pipeline (Coil et al, 2014) y Ray v2.3.2 (Boisvert et al, 2010).

4.4. Anotación y visualización del genoma

La anotación del genoma se llevó a cabo utilizando RNAweseal (Lang et al, 2007), MFannot (Beck y Lang, 2010) y tRNAscan-SE 1.21 (Schattner et al, 2005) y los resultados validados realizando búsquedas de BLAST. Las regiones repetidas invertidas (IR), las regiones repetidas palíndromas y las regiones repetidas forward fueron identificadas utilizando el programa REPuter (Kurtz et al, 2001) (<https://bibiserv2.cebitec.uni-bielefeld.de/reputer>) siguiendo los siguientes criterios: tamaño ≥ 30 pb, identidad = 90% (distancia de Hamming = 3). El software MAUVE v2.4.0 (Darling et al, 2004) se utilizó para comparar los bloques de genes entre el genoma de la cepa reportada y nuestra cepa. La frecuencia relativa de uso de codones se calculó y se comparó con la que presenta la cepa CCAP 19/18 utilizando el software ChloroMitoCU (Sablok et al, 2017). El mapa del genoma del cloroplasto se generó utilizando el software en

línea OGDRAW (<http://ogdraw.mpimp-golm.mpg.de/cgi-bin/ogdraw.pl>) (Lohse et al, 2007).

Adicionalmente se realizó una comparación entre los promotores utilizados en la última construcción (revisar Figura 6), los cuales fueron sintetizados a partir de la secuencia del cloroplasto de *D. salina* CCAP 19/18 depositada en GenBank. El alineamiento se realizó utilizando el software MEGA7 (Kumar et al, 2016) con el algoritmo ClustalW. La estructura secundaria del 5' UTR del gen *psbA* se realizó utilizando el software Mfold (<http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form>) (Zuker et al, 2003).

5. Resultados y discusión

5.1. Organización del genoma del cloroplasto

La cepa aislada fue cultivada exitosamente en el laboratorio en medio modificado y se logró la separación de los cloroplastos del resto de la célula para tener en mayor proporción DNA del cloroplasto. El material genético fue enviado para su secuenciación, dichos resultados fueron ensamblados obteniendo dos contigs que posteriormente fueron unidos (Tesis de Doctorado de M.C. Dante Alberto Magdaleno Moncayo).

El genoma del cloroplasto de *D. salina* SQ presentó un tamaño de 243,635 pb, siendo 25,409 pb más pequeño que el genoma reportado de la cepa CCAP 19/18 (número de acceso de GenBank GQ250046). El genoma anotado fue sometido a la base de datos de GenBank y se obtuvo el número de acceso KX530454. El genoma presenta una estructura cuádruple, como la mayoría de los genomas de cloroplasto, cuenta con la región LSC (Large Single Copy, por sus siglas en inglés) cuyo tamaño es de 107,815 pb, la región SSC (Small Single Copy, por sus siglas en inglés) de 101,527 pb y las regiones invertidas (IR) de 17,145 pb. Las regiones LSC y SSC tienen un menor tamaño que la cepa previamente reportada mientras que las IR tienen un tamaño mayor (17,145 pb vs 14,409 pb) (Fig. 39). El porcentaje de GC es del 29.73% lo cual está dentro del rango de la mayoría de los genomas de otras microalgas (CCAP 19/19 GC= 32.1 %) (Tabla II).

Figura 39. Mapa del genoma de *Dunaliella salina* SQ. El mapa fue creado con el software OrganellarGenomeDRAW.

Tabla II. Características de los genomas de cloroplasto de las cepas de *D. salina*

Cepa	<i>D. salina</i> CCAP 19/18	<i>D. salina</i> SQ
No. Acceso GenBank	GQ250046.1	KX530454.1
Longitud (pb)	269044	243635
IR (pb)	14409	17145
LSC (pb)	127339	107815
SSC (pb)	112887	101527
% GC	32.08	29.73
Genes codificantes	66	66
Genes rRNA	3	3
Genes tRNA	28	29

El genoma está compuesto por 101 genes, de los cuales 10 se encuentran duplicados (5 se encuentran en las regiones invertidas) y uno por triplicado (Tabla III). Dentro de esos genes, 66 codifican para proteínas, 29 para RNA de transferencia (tRNA) y 3 codifican para RNA ribosomal (rRNA). El gen *trnM* se encuentra por triplicado mientras que en la cepa reportada se encuentra duplicado. Los tres genes que codifican para rRNA se encuentran localizados en las regiones invertidas además de los genes que codifican para tRNA, *trnA* y *trnI*. De los ocho genes que contienen intrones (Tabla IV), *psaA* es transpliced; mientras su primer exón se encuentra codificado en la cadena líder, los cuatro restantes se encuentran codificados en la cadena complementaria, además los primeros dos exones están posicionados en la región LSC y los exones 3 – 5 en la región SSC. A diferencia de la cepa CCAP 19/18, en SQ los genes *atpB* y *chlL*

no presentan intrones y en el caso de *rrnS* y *rrnL* un número menor de intrones (Tabla III).

Al comparar las secuencias de los genomas de ambas cepas de *D. salina* podemos observar que solamente comparten un 62% de cobertura y una identidad del 96% (Fig. 40). Las diferencias entre las secuencias de ambos genomas se dan principalmente en las regiones no codificantes.

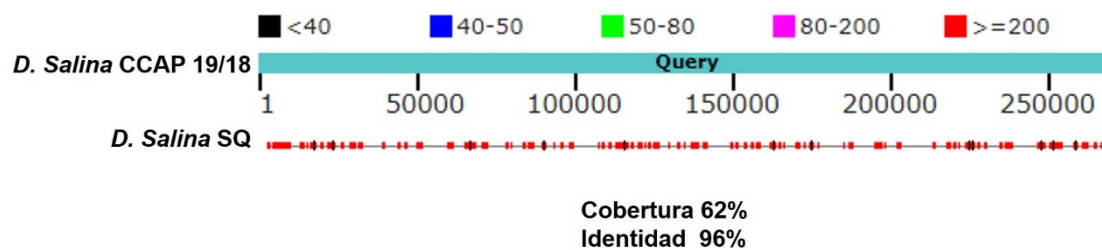


Figura. 40 Alineamiento de los genomas de las cepas de *D. salina* CCAP 19/18 y SQ. En rojo se muestran las regiones que alinean entre los dos genomas.

Tabla III. Genes presentes en el genoma del cloroplasto de *D. salina* SQ

Tipo de genes	Genes
Fotosistema I	<i>psaA</i> [*] , <i>psaB</i> [*] , <i>psaC</i> , <i>psaJ</i>
Fotosistema II	<i>psbA</i> [*] , <i>psbB</i> , <i>psbC</i> [*] , <i>psbD</i> [*] , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
Biosíntesis de clorofila	<i>chlB</i> , <i>chlL</i> , <i>chlN</i>
Complejo Citocromo b/f	<i>petA</i> , <i>petB</i> , <i>petD</i> , <i>petG</i> , <i>petL</i>
ATP sintasa	<i>atpA</i> [*] , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> , <i>atpH</i> , <i>atpI</i>
RubisCO	<i>rbcL</i>
RNA polimerasa	<i>rpoA</i> , <i>rpoBa</i> , <i>rpoBb</i> , <i>rpoC1</i> , <i>rpoC2</i>
Proteínas ribosomales (LSU)	<i>rpl14</i> , <i>rpl16</i> , <i>rpl2</i> , <i>rpl20</i> , <i>rpl23</i> , <i>rpl36</i> , <i>rpl5</i>
Proteínas ribosomales (SSU)	<i>rps11</i> , <i>rps12</i> , <i>rps14</i> , <i>rps18</i> , <i>rps19</i> , <i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> , <i>rps8</i> , <i>rps9</i>
RNA ribosomal	<i>rrn5</i> (2x), <i>rrnL</i> (2x), <i>rrnS</i> (2x)
RNA transferencia	<i>trnA</i> (2x), <i>trnC</i> , <i>trnD</i> , <i>trnE</i> (2x), <i>trnF</i> , <i>trnG</i> (2x), <i>trnH</i> , <i>trnI</i> (2x), <i>trnK</i> , <i>trnL</i> (2x), <i>trnM</i> (3x), <i>trnN</i> , <i>trnP</i> , <i>trnQ</i> , <i>trnR</i> (2x), <i>trnS</i> (2x), <i>trnT</i> , <i>trnV</i> , <i>trnW</i> , <i>trnY</i>
Proteínas de función desconocida	<i>ycf1</i> , <i>ycf12</i> , <i>ycf3</i> , <i>ycf4</i>
Otros genes	<i>ccsA</i> , <i>cemA</i> , <i>clpP</i> , <i>ftsH</i> , <i>tufA</i> ,

* Genes que contienen intrones

^ Genes que contienen intrones transpliced

2x Genes duplicados

3x genes triplicados

Tabla IV. Diferencia de intrones entre las cepas de *D. salina*

Cepa	Numero de intrones por gen										Total de intrones
	<i>atpA</i>	<i>atpB</i>	<i>chlL</i>	<i>psaA</i>	<i>psaB</i>	<i>psbA</i>	<i>psbC</i>	<i>psbD</i>	<i>rrnS</i>	<i>rrnL</i>	
<i>D. salina</i> SQ	1	0	0	2	1	5	2	1	2	4	17
<i>D. salina</i> CCAP 19/18	1	1	1	3	1	5	2	1	4	7	26

El arreglo de los genes dentro del genoma del cloroplasto de ambas cepas de *D. salina* es similar, los bloques de genes mantienen el mismo orden solo cambia el orden de algunos marcos abiertos de lectura (ORF's) (Fig. 41). En la figura anterior también se muestra la arquitectura obtenida con el software Mauve, las regiones similares se encuentran del mismo color y están conectadas por líneas; las regiones que se encuentran en diferente posición entre ambos genomas son las que corresponden a los ORF's.

Los límites entre las regiones LSC – IR – SSC – IR son conservados entre ambas cepas. Los genes que se encuentran más próximos a los límites entre estas regiones no cambian, lo único que varía son las distancias a las que se encuentran de ellos (Fig. 42). En el límite LSC – IR se encuentran los genes *chlB* y *rrnS*, en el límite IR – SSC están los genes *rrn5* y *atpH*, en el límite SSC – IR están los genes *rpoC2* y *rrn5*, mientras que en la región IR – LSC se encuentran los genes *rrnS* y *ycf1*.

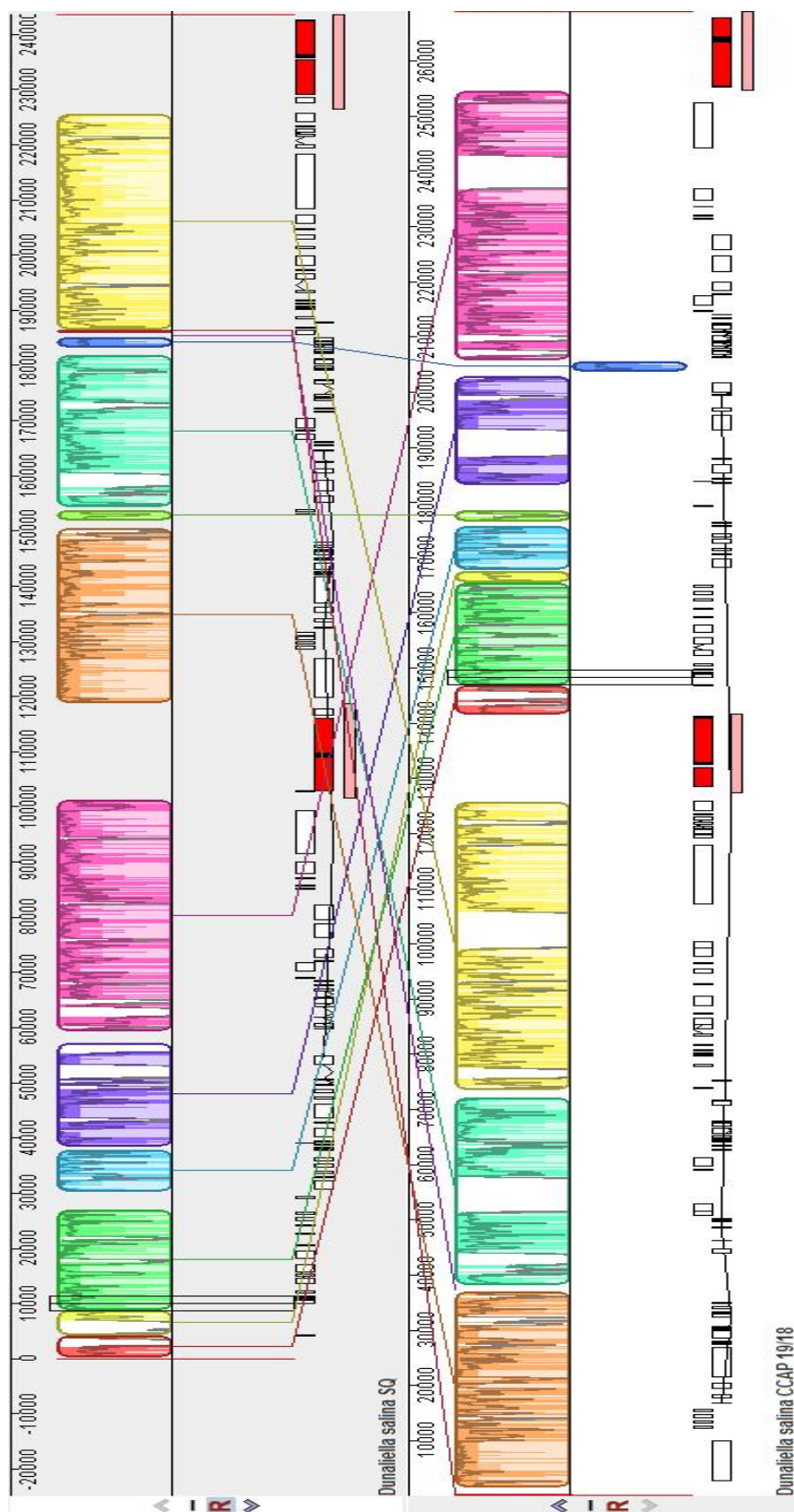


Figura 41. Comparación de la estructura de los genomas de *D. salina* SQ (arriba) y *D. salina* CCAP 19/18 (abajo)

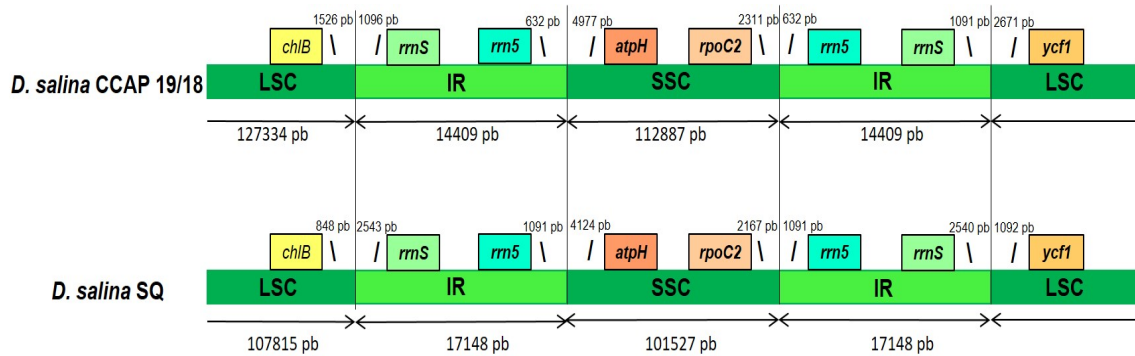


Figura 42. Comparación de la posición de los límites de las regiones LSC, Ira, IRb y SCC entre las cepas de *D. salina* CCAP 19/18 y SQ.

5.2. Estructuras repetitivas

El software REPuter localizó 2,108 pares de repeticiones de una longitud de ≥ 30 pb con una identidad mayor al 90% en el genoma del cloroplasto de *D. salina* SQ, de las cuales 1,185 fueron forward y 923 palindrómicas. Se encontraron 1832 repeticiones en la región SSC, 1211 en la región LSC y 1173 en las regiones IR (Fig. 43). Comparando con el genoma del plástido de la cepa CCAP 19/18, el número de regiones repetitivas dentro del criterio de búsqueda fueron menores, en este genoma se detectaron 5000 repeticiones, sin embargo, la presencia de 2018 repeticiones bajo estos parámetros sigue siendo alto, resultado consistente con la mayoría de los genomas de cloroplasto de microalgas.

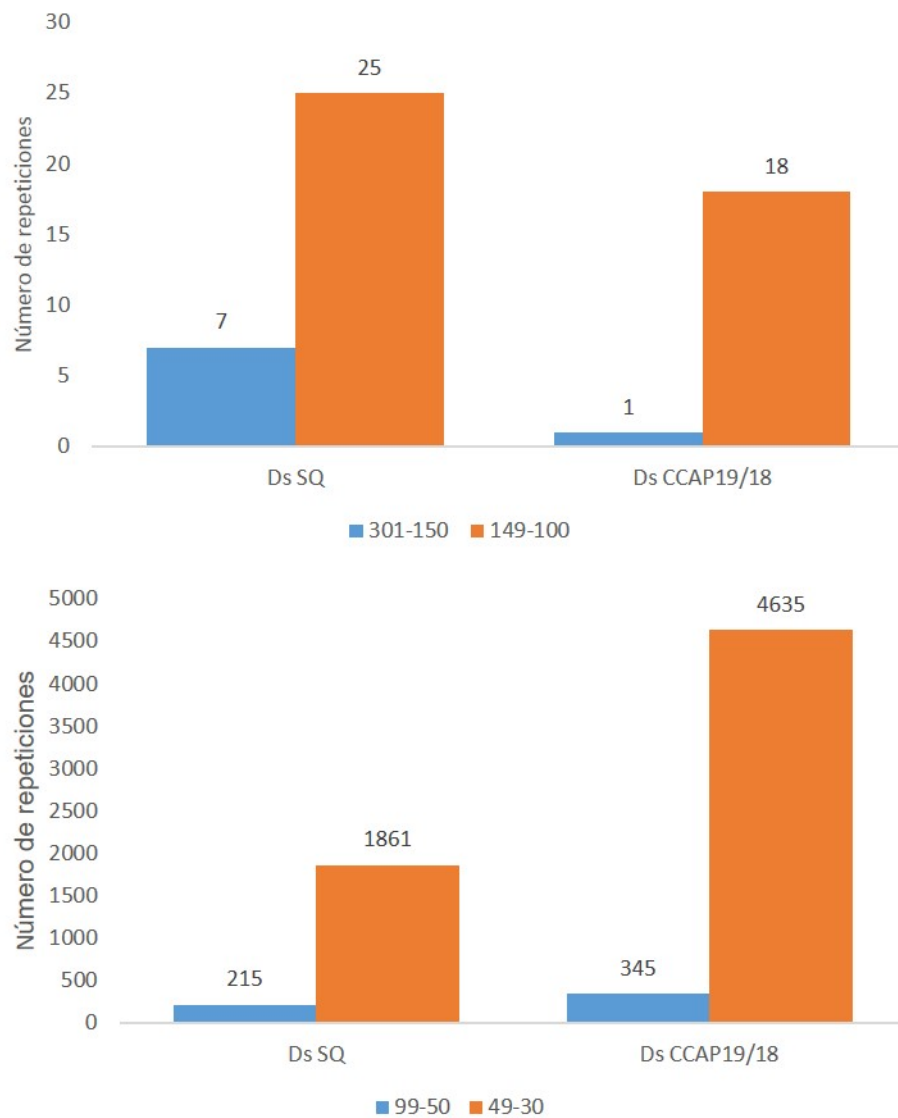


Figura 43. Representación gráfica del número y tamaño de las regiones repetitivas en los cloroplastos de las cepas de *D. salina*.

5.3. Frecuencia relativa de uso de codones

Los resultados obtenidos del análisis de frecuencia relativa de uso de codones demostraron que en todos los casos el codón más utilizado para los diferentes aminoácidos es el mismo en ambas cepas (Fig. 44a). Los porcentajes de GC en general y en la ubicación dentro del codón fueron similares en las dos cepas, los codones más utilizados son los que contienen AT, especialmente en la tercera posición del codón (Fig. 44b). Estos resultados coinciden con el alto contenido de AT presentan ambos genomas.

En el diseño de casetes para la expresión de proteínas recombinantes es importante considerar el alto uso de los codones con adeninas y timinas que tiene el cloroplasto, por lo cual se debe realizar la optimización de codones en los genes recombinantes. Este alto contenido de adeninas y timinas suele ser un problema en la optimización de genes, ya que generalmente los fragmentos grandes con alto número de AT son difíciles de sintetizar porque las cadenas de DNA son menos estables. También se complica la amplificación por PCR, ya que existen polimerasas que tiende a cambiar las adeninas por guaninas, y a pesar de que este cambio resulte en una mutación silenciosa, la baja frecuencia del uso de codones con GC en la tercera posición puede ser un problema en la expresión de los genes recombinantes por el déficit que puede existir de estos tRNA (Mayfield et al, 2003; Heitzer et al, 2007; Mayfield et al, 2007; Surzycki et al, 2009; Specht et al, 2010; Wu et al, 2011; Qin et al, 2012)

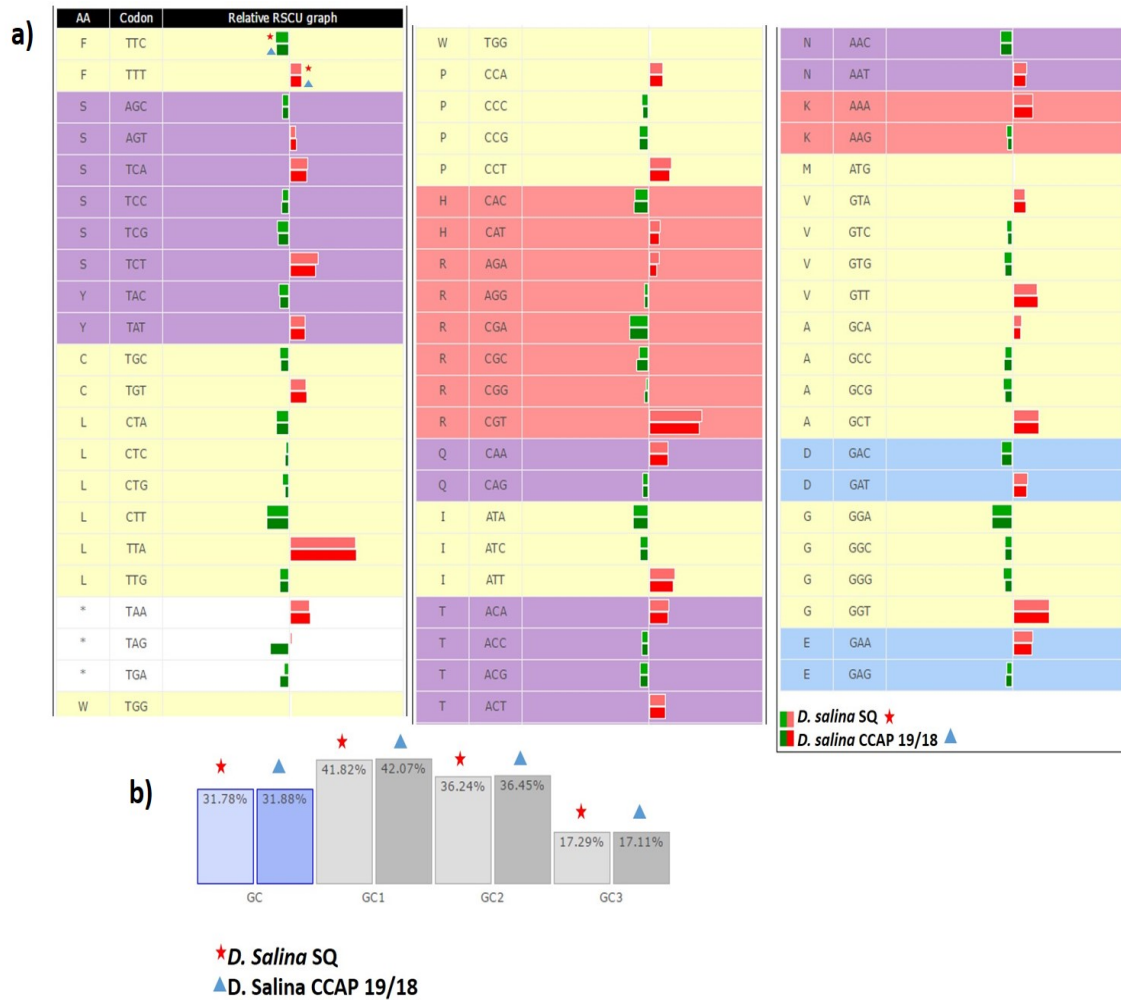


Figura 44. Representación gráfica de la frecuencia de uso de codones de ambas cepas de *D. salina*. a) Comparación del uso de la frecuencia de uso de codones entre las cepas SQ y CCAP 19/18. b) Porcentaje de GC total y porcentaje de GC en las tres ubicaciones del codón.

5.4. Comparación de promotores

Una vez realizado el alineamiento se compararon las secuencias de las regiones de los diferentes promotores y terminadores. En la figura 45 se muestra la imagen de alineamiento que corresponde a la región de promotor-5'UTR del gen *rbcL*, se observan las diferencias entre la secuencia de la cepa CCAP 19/18 y la cepa SQ; en color morado se resalta la posible secuencia del promotor para este gen la cual no se la presenta la cepa SQ, esto no significa que no la tenga, es probablemente se encuentra en una región más distal. Estas diferencias podrían ser un problema si se pretende utilizar promotores de otras cepas para la expresión de proteínas recombinantes.

Al analizar la secuencia correspondiente al promotor-5'UTR del gen *psbA*, se observa que también existen diferencias entre ambas cepas, en este caso se detectaron dos regiones que pueden funcionar como promotores, pero solo una de ellas se encuentra en ambas cepas (Fig. 46). Al comparar las secuencias reguladoras de estas cepas con otras analizadas por Gimpel y Mayfield (2013) se puede observar que la segunda región (la más cercana al ATG) podría ser la secuencia del promotor. Siendo así, esta secuencia podría ser utilizada en ambas cepas para llevar a cabo la expresión de proteínas recombinantes. Adicionalmente, en ambos 5' UTR's del *psbA* se observó el bucle característico que presentan diferentes especies de algas, localizado antes del posible sitio de unión al ribosoma (Fig. 47).

```

RbcL-Plásmido    1  -----TTAGCGGAGCCCTCCCGTGAGGGGGGCTAGGACTTTGTCCGT
RbcL-CCAP       439 TATAATTAGAATTTTAAATTAGCGGAGCCCTCCCGTGAGGGGGGCTAGGACTTTGTCCGT
RbcL-SQ         443 -----

RbcL-Plásmido    43  CGCCTAAGGGCTAGGATCAAGATCCGTCGCGTGAGGGAAACAAAAAGTTGTTTTGTCCA
RbcL-CCAP       499 CGCCTAAGGGCTAGGATCAAGATCCGTCGCGTGAGGGAAACAAAAAGTTGTTTTGTCCA
RbcL-SQ         443 -----ATAAAAGAAAATTAAACAATAA

RbcL-Plásmido    103 GAAGTATTTTTTTCTTATAAAAAAGAAAAAGAA-----AAAAAAATTATA-----
RbcL-CCAP       559 GAAGTATTTTTTTCTTATAAAAAAGAAAAAGAA-----AAAAAAATTATA-----
RbcL-SQ         466 AGCATAATAAGTTATTATAAATACAGGAACGAAATTATATAGAATTATAATTATAAATT

RbcL-Plásmido    150 --AAATTATAAAAGAATTTTATATTCTTTAGTTATCAAAATTATAATTGGTAAAAGATT
RbcL-CCAP       606 --AAATTATAAAAGAATTTTATATTCTTTAGTTATCAAAATTATAATTGGTAAAAGATT
RbcL-SQ         526 GGAATTAGAAAAAATTATATCTTCTTTAATTACCAAAATTATAATTGGTAAAAGATT

RbcL-Plásmido    208 ATTATATTATCGGATAGATTATTTTAGGATCGACAAAAATGTTTCATTTTATAGATAATAA
RbcL-CCAP       664 ATTATATTATCGGATAGATTATTTTAGGATCGACAAAAATGTTTCATTTTATAGATAATAA
RbcL-SQ         586 ATTATATTATCGGATAGATTATTTTAGGATCGACAAAAATGTTTCATTTTATAGATAATAA

RbcL-Plásmido    268 A
RbcL-CCAP       724 A
RbcL-SQ         646 A

```

Figura 45. Alineamiento de la región del promotor-5'UTR del gen *rbcL* de las cepas de *D. salina* SQ y CCAP 19/18. También se muestra la región del plásmido. En color morado se muestra la región del posible promotor para este gen y en negro están marcados los nucleótidos iguales.

psbA-Plásmido	1	CACAAAGTCCTAGCCCTTAGGGGGCGGGGGCGTTTGTCCGTCGCGTGAGAGAAATAACAA
psbA-SQ	1	TGCCAACAAATGATCTTCATAGAATTATATACACAGACTATTGCCGCTCGAAAAGAACA
psbA-CCAP1918	1	CACAAAGTCCTAGCCCTTAGGGGGCGGGGGCGTTTGTCCGTCGCGTGAGAGAAATAACAA
psbA-Plásmido	61	AGTTGTTTTGTCCGTACACAATAAAAAATTATATAAAAAGTTGTTGTTTATAAAAAGGAT
psbA-SQ	61	TTTTATCTT-TTTAATACACAATAAAAAATTATATAAAAAGTTGTTGTTTATAAAAAGGAT
psbA-CCAP1918	61	AGTTGTTTTGTCCGTACACAATAAAAAATTATATAAAAAGTTGTTGTTTATAAAAAGGAT
psbA-Plásmido	121	GCTCTTGCTTTTGATAAATTATCTGGTATTATAATAATACCAGGTTTAAATGAAAAAGCT
psbA-SQ	120	GCTCTTGCTTTTGATAAATTATCTGGTATTATAATAATACCAGGTTTAAATGAAAAAGCT
psbA-CCAP1918	121	GCTCTTGCTTTTGATAAATTATCTGGTATTATAATAATACCAGGTTTAAATGAAAAAGCT
psbA-Plásmido	181	TGAATAATA----AAAAAATAAAAAAGTTATTTTAA--CTTTTTTCGGAGAAAAATCAAA
psbA-SQ	180	TGAATAATATAAAAAAGATAAAAAAGTTATTTTAACTTTTTTCGGAGAAAAATCAAA
psbA-CCAP1918	181	TGAATAATA----AAAAAATAAAAAAGTTATTTTAA--CTTTTTTCGGAGAAAAATCAAA
psbA-Plásmido	235	AATTAATAAAAAATTATAAATT
psbA-SQ	240	AATTAATAAAAAATTACAAATT
psbA-CCAP1918	235	AATTAATAAAAAATTATAAATT

Figura 46. Alineamiento de la región del promotor-5'UTR del gen *psbA* de las cepas de *D. salina* SQ y CCAP 19/18. También se muestra la región del plásmido. En color morado se muestra la región del posible promotor para este gen y en negro están marcados los nucleótidos iguales.

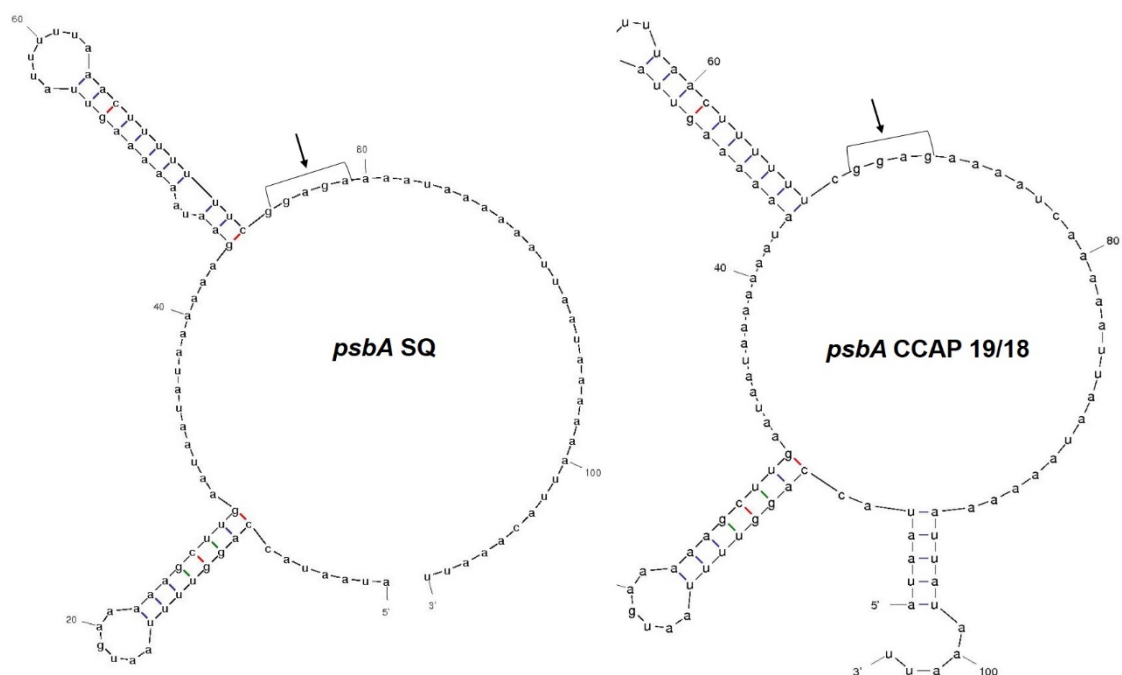


Figura 47. Estructura secundaria del 5' UTR del gen *psbA* de las cepas SQ (izquierda) y CCAP 19/18 (derecha). La flecha indica el sitio de union a ribosma.

También se analizaron las regiones 3'UTR utilizadas en los casetes de expresión diseñados previamente. Tanto las regiones 3' UTR del gen *psbA* como del *atpA* mostraron diferencias entre ambas cepas, ambas regiones tienen un tamaño menor en la cepa SQ (Fig. 48 y Fig. 49). La región 3' UTR está involucrada en el procesamiento y estabilidad del mRNA (Rasala y Mayfield, 2011) y en experimentos donde utilizan distintos 3' UTR's para la expresión de proteínas recombinantes en el cloroplasto de *C. reinhardtii* comprobaron que la acumulación de los mRNA es independiente del 3'UTR y que depende principalmente de la región promotor-5'UTR; por lo que tal vez las diferencias en

estas regiones no sean tan importantes como las que existen en la región promotor-5'UTR a la hora de la expresión de proteínas recombinantes en cloroplasto.

UTR-psbA-Plasmi	1	TTCTTTATACTTAAGAATTAAGTAGACTTGCTACTTAATTCCTTTTAATAGCATGGATG
UTR-psbA-8Q	1	TTCTTTATACTTAAGAATTAAGTAGACAACTCTACTTAATTCCTTTTAATAGCATGGATG
UTR-psbA-CCAP19	1	TTCTTTATACTTAAGAATTAAGTAGACTTGCTACTTAATTCCTTTTAATAGCATGGATG
UTR-psbA-Plasmi	61	GAGTAAACTCCATCCATGCTATTATCTCTAACAACCCCAAAAACCAAGGTTTTGTGAA
UTR-psbA-8Q	60	GAGTAAACTCCATCCATGCTATTATTTATAACAAC-----GGAAAAATAAGATTTTTTTT
UTR-psbA-CCAP19	61	GAGTAAACTCCATCCATGCTATTATCTCTAACAACCCCAAAAACCAAGGTTTTGTGAA
UTR-psbA-Plasmi	121	GGAAAAATAGAATATTTTTTCCGCTACGAGCGGGCGATATAAAAAATAAGATTTTTTTT
UTR-psbA-8Q	96	-----GGAAAAATAAGATTTTTTTT
UTR-psbA-CCAP19	121	GGAAAAATAGAATATTTTTTCCGCTACGAGCGGGCGATATAAAAAATAAGATTTTTTTT
UTR-psbA-Plasmi	181	ATCCTAAATTAATTCCTTGTAAATAGAATCTCAATTAATGACAAAGATATAAATTACAAAA
UTR-psbA-8Q	114	ATCCTAAATTAATTCCTTGTAAATAGAACTCAATTAATTTACACAAATGTAAATTACAAAA
UTR-psbA-CCAP19	181	ATCCTAAATTAATTCCTTGTAAATAGAATCTCAATTAATGACAAAGATATAAATTACAAAA
UTR-psbA-Plasmi	240	ACGGTAGAGTTTCAGGTTTAGATAAAATAACCTGGTCCTGCTATATAACTAATGTAAGTTA
UTR-psbA-8Q	174	AGGATTCAGCTTCAGGTTGAGATGAATAATATGGTTCGCTATCTAAACAAATGTGACTTA
UTR-psbA-CCAP19	240	ACGGTAGAGTTTCAGGTTTAGATAAAATAACCTGGTCCTGCTATATAACTAATGTAAGTTA
UTR-psbA-Plasmi	300	AAATATCAGTATATGAACAATATATATTATATATTAATAGTCCATGTCTTCAAAAACGA
UTR-psbA-8Q	234	AAATATATATATGATTAACAATCTAT-----GTTAATAGTACATGTCTTGAATAACGA
UTR-psbA-CCAP19	300	AAATATCAGTATATGAACAATATATATTATATATTAATAGTCCATGTCTTCAAAAACGA
UTR-psbA-Plasmi	360	ATAAACGAATAAACCAAGAAAAATAAACTTTATTATTTTAAATAAAATAAAATTTTATT
UTR-psbA-8Q	287	ATAAACGA-----AGAAAAAT-----TTCTATATAAAA---GACTTATT
UTR-psbA-CCAP19	360	ATAAACGAATAAACCAAGAAAAATAAACTTTATTATTTTAAATAAAATAAAATTTTATT
UTR-psbA-Plasmi	420	AAATTTTTTTTTTATCAAAAATTTTTTTTTT
UTR-psbA-8Q	325	AAATCCCTTTT-----AGAAATTTTTTCTTT
UTR-psbA-CCAP19	420	AAATTTTTTTTTTATCAAAAATTTTTTTTTT

Figura 48. Alineamiento de la región 3'UTR del gen *psbA* de las cepas de *D. salina* SQ y CCAP 19/18. También se muestra la región del plásmido. En negro están marcados los nucleótidos iguales.

```

UTR-atpA_Plasmid 1 TAAAAAAAAAAAAAAAAATTTAATTTTATAAATTTTATTAATTTAGTTACAAAATCTTCCATT
UTR-atpA_SQ 1 TTAACTTAACCTAAATGAAATTTTATAAATTTGATTAATTT-GTTATAAAATTTTCAATT
UTRatpA_CCAP19/ 1 TAAAAAAAAAAAAAAAAATTTAATTTTATAAATTTTATTAATTTAGTTACAAAATCTTCCATT

UTR-atpA_Plasmid 61 TTCATCGGCCCGCCAAAAGATAGTAAAAAA-----TATACATATCTTTTAGGCGGAG
UTR-atpA_SQ 60 TTCATAGGCCTCCGTCAAAAGATATAAAAAAAAAAATATATATATCTTTTAGGCGGAG
UTRatpA_CCAP19/ 61 TTCATCGGCCCGCCAAAAGATAGTAAAAAA-----TATACATATCTTTTAGGCGGAG

UTR-atpA_Plasmid 115 CCTATCTCCGGTCTTTTATAAAAAGAAAAACCGACATACAAAATACAAAATTGTATGGAG
UTR-atpA_SQ 120 CCTATCTCCGGTCTTTTATAAAAAGAAAAACCGACATACAAAAT-----TGTATGTA
UTRatpA_CCAP19/ 115 CCTATCTCCGGTCTTTTATAAAAAGAAAAACCGACATACAAAATACAAAATTGTATGGAG

UTR-atpA_Plasmid 175 GCGGACAAAACAAATTTGTTGTTTGTCTTAAGGAAAAAATTCAACATATTTAATATA
UTR-atpA_SQ 173 -----TTGTCATAAGCAAAAAAATTCAACCTTTTAAATCTAA
UTRatpA_CCAP19/ 175 GCGGACAAAACAAATTTGTTGTTTGTCTTAAGGAAAAAATTCAACATATTTAATATA

UTR-atpA_Plasmid 235 TATGTACTGTTTACAAATATTTGATATAACTTGATTATCGAAATTTTCAAACAGTACAT
UTR-atpA_SQ 210 TATGTACTGTTTGACAAATTTGATATAACTTTATTATCAAAATTTTCAAACAGTACAT
UTRatpA_CCAP19/ 235 TATGTACTGTTTACAAATATTTGATATAACTTGATTATCGAAATTTTCAAACAGTACAT

UTR-atpA_Plasmid 295 A-ATTATAAATT-ATGTAAATCGATATTTATAAAAAATATTTTATAGTTTATTGAAAC-T
UTR-atpA_SQ 270 ATATTATAATTTTATTAAGATTTGATATTTATAAAAAATCTTTTATAGTTTATTGAAAT
UTRatpA_CCAP19/ 295 A-ATTATAAATT-ATGTAAATCGATATTTATAAAAAATATTTTATAGTTTATTGAAAC-T

UTR-atpA_Plasmid 352 ATTATATTCTAAATATCGAAGA
UTR-atpA_SQ 330 TTTTCTTCTAAATAAATTA
UTRatpA_CCAP19/ 352 ATTATATTCTAAATATCGAAGA

```

Figura 49. Alineamiento de la región 3'UTR del gen *atpA* de las cepas de *D. salina* SQ y CCAP 19/18. También se muestra la región del plásmido. En negro están marcados los nucleótidos iguales.

6. Conclusiones

Se logró la anotación del genoma del cloroplasto de *D. salina* SQ, contiene 101 genes, 4 de fotosistema I, 15 de fotosistema II, 5 involucrados en la síntesis de clorofila, 6 para ATP sintasa, uno de RuBisCO, 5 de RNA polimerasa, 18 de proteínas ribosomales, 6 de RNA ribosomal, 29 tRNA's, 4 de proteínas de función desconocida y 5 de genes diversos.

El genoma del cloroplasto de la cepa SQ (México) es diferente en tamaño (243635 pb), regiones no codificantes y reguladoras al reportado de la cepa CCAP 19/18 (Australia) (269044). La cepa SQ presenta un número menor de genes con intrones, 8 genes presentan intrones mientras que en la cepa CCAP 19/18 10 genes presentan intrones. La cantidad de intrones en la cepa SQ es 17 y en la cepa CCAP 19/18 es 26. En SQ el gen *trnM* está triplicado mientras que en la cepa CCAP 19/18 se encuentra por duplicado. Existen pequeñas variaciones en la arquitectura del genoma, pero la sintenia es la misma y el número de estructuras repetitivas es menor en la cepa SQ.

Las regiones 5'UTR de los genes *rbcL* y *psbA* son diferentes en ambas cepas y solo en el gen *psbA* se pudo detectar la presencia de una secuencia potencial para el promotor. También las regiones 3' UTR de los genes *atpA* y *psbA* son diferentes en ambas cepas. La variabilidad encontrada entre los genomas de cloroplasto de las cepas de *D. salina* dificulta el uso de regiones reguladoras entre cepas para la creación de sistemas de expresión de proteínas recombinantes en cloroplasto.

Conclusiones finales

El objetivo general de este trabajo consistía en crear un sistema de expresión de proteínas recombinantes en el cloroplasto de *D. salina* SQ y como se puede apreciar en los capítulos I y II, se crearon tres distintos plásmidos para lograr la expresión de las proteínas interleucina-2, EGFP, nourseotricina acetil transferasa y fitoeno desaturasa F128V. Sin embargo, no se pudo detectar su integración en el plastoma de *D. salina* SQ. El primer problema que se presentó fue la falta de un agente de selección efectivo contra *D. salina*, originalmente mediante estudios previos no publicados se había determinado que el antibiótico nourseotricina inhibía su crecimiento, sin embargo, después de una serie de transformaciones y la alta presencia de clonas, se realizó un nuevo estudio en donde se determinó que el uso de este antibiótico no era adecuado como agente de selección ya que a una concentración de 1 mg/ml aún existía crecimiento de esta microalga. El problema se solucionó al comprobar que el herbicida norflurazon una concentración de 5 μ M inhibe su crecimiento (capítulo II).

Los casetes de expresión se crearon utilizando como referencia el genoma el cloroplasto de *D. salina* depositado en Genbank, el cual pertenece a la cepa CCAP 19/18. En este estudio se utilizó la cepa *D. salina* SQ aislada en la región de San Quintín, Baja California. Al formular el objetivo del trabajo se asumió que el genoma era el mismo, sin embargo, después no poder detectar transformantes, se decidió analizar el genoma de la cepa SQ. En los resultados obtenidos de este análisis se puede apreciar que ambos genomas son diferentes y que estas diferencias se encuentran principalmente en regiones no

codificantes. Se sabe que la expresión de proteínas recombinantes en cloroplasto depende del uso de regiones no traducidas y promotores endógenos, y el utilizar secuencias no nativas suele ser un problema para la expresión de estas proteínas.

Cualquiera de las razones anteriores pudo ser lo que evitó que se llevara a cabo la expresión de las proteínas, adicionalmente también pudo haber sido un problema de integración debido las regiones de recombinación elegidas o la falta de réplicas mediante el método de biobalística. Este trabajo contribuye al conocimiento de la creación de casetes para la expresión de proteínas recombinantes en cloroplasto de *D. salina* SQ y al estudio de las diferencias genéticas, como el número de intrones, regiones repetitivas y disminución de homología en regiones no codificantes, que existen entre los genomas de los orgánulos entre las distintas cepas de *D. salina*.

Perspectivas

A partir de los resultados obtenidos en esta tesis doctoral, se propone la creación de dos nuevos casetes para lograr la expresión de proteínas recombinantes en el cloroplasto de *D. salina*. Como se observó en los resultados obtenidos de los capítulos I y II, la problemática de la expresión de proteínas recombinantes en el cloroplasto de *D. salina* depende de varios factores, en primera instancia se propone modificar el sitio de recombinación homóloga, en lugar de utilizar una región no codificante se propone utilizar secuencias que contengan una región codificante. Como se muestra en la figura 50a, en esta alternativa se pretende utilizar parte del intrón I y el exón II del gen *atpA* como brazo izquierdo de recombinación y la región 3' UTR del mismo gen como brazo derecho de recombinación. El gen a expresar sería el *pdsF128V* que confiere resistencia al herbicida norflurazon, que como se demostró en el capítulo II, funciona como agente de selección. Este gen no contará con un promotor, porque utilizará el nativo del gen *atpA*, la idea es que se genere una especie de policistrón que permita la transcripción de gen *pdsF128V* junto con este gen.

La segunda opción que se propone (Fig. 50b) es utilizar como región de inserción la región no codificante entre los genes *trnE* y *psbH*, pero como regiones de recombinación homóloga regiones codificantes. Como región de recombinación homóloga izquierda se propone la región que contiene a los genes *trnE* y *psbE*, y como región de recombinación derecha la región que contiene los genes *psbH*, *psbN* y *psbT*.

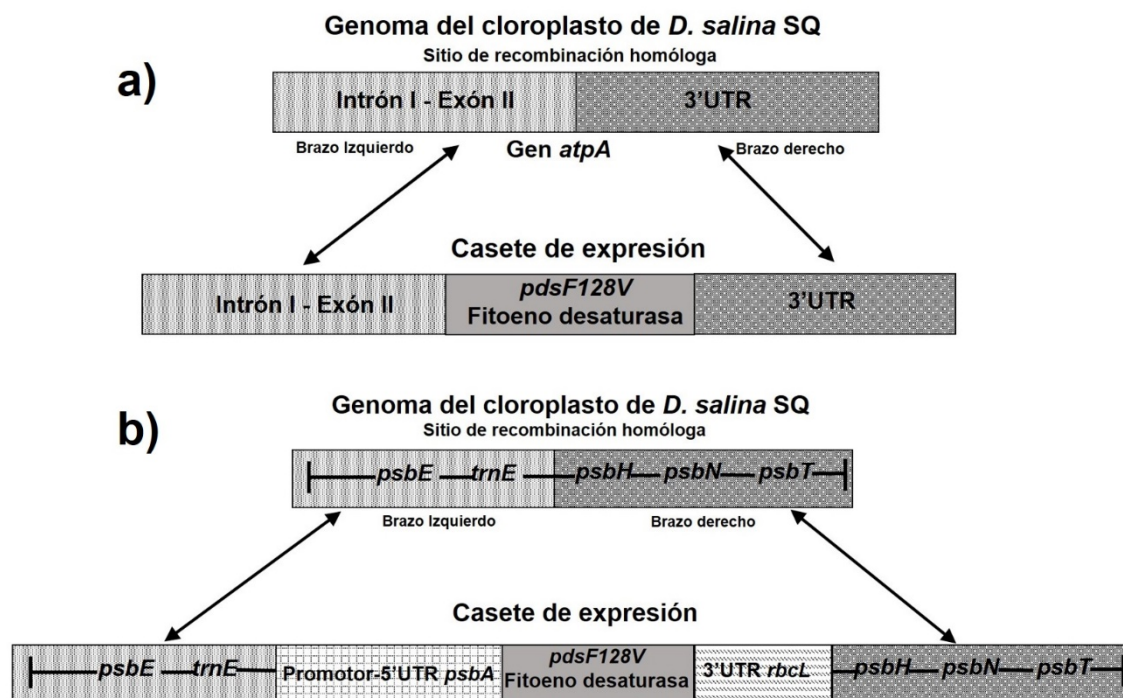


Figura 50. Propuesta de nuevos casetes de expresión para lograr la expresión de proteínas recombinantes en el cloroplasto de *D. salina* SQ. a) Casete de expresión en forma de policistrón. b) Casete de expresión utilizando el promotor y 5' UTR del gen *psbA*, 3'UTR del gen *rbcL* y regiones de recombinación homóloga codificantes para los genes *psbE*, *trnE*, *psbH*, *psbN* y *psbT*.

De igual manera se propone utilizar el *pdsF128V*, pero en esta construcción estará regulado por el promotor y 5'UTR del gen *psbA* de *D. salina* SQ y la región 3'UTR del gen *rbcL* también de *D. salina* SQ. Todas las regiones serán basadas en la secuencia del cloroplasto de *D. salina* SQ, la cual se obtuvo como en esta tesis (capítulo III). Al utilizar solamente el gen que confiere resistencia a norflurazon se disminuyen el número de variables, y se simplifica la

forma de detectar la presencia de la integración del transgen dentro del genoma del cloroplasto, ya que las colonias que crezcan en medio suplementado con norflurazon serán las que contengan el transgen. Estas dos construcciones permiten probar dos estrategias diferentes para lograr la expresión de proteínas recombinantes.

Referencias

- Adem, M., Beyene, D. y Feyissa, T. 2017. Recent achievements obtained by chloroplast transformation. *Plant Methods*, 13(1), p.30.
- Akbari, F., Eskandani, M. y Khosroushahi, A.Y. 2014. The potential of transgenic green microalgae; a robust photobioreactor to produce recombinant therapeutic proteins. *World Journal of Microbiology and Biotechnology*, 30(11), pp.2783-2796.
- Barnes, D., Franklin, S., Schultz, J., Henry, R., Brown, E., Coragliotti, A. y Mayfield, S.P. 2005. Contribution of 5'-and 3'-untranslated regions of plastid mRNAs to the expression of *Chlamydomonas reinhardtii* chloroplast genes. *Molecular Genetics and Genomics*, 274(6), pp.625-636.
- Bartels, P.G. y Watson, C.W. 1978. Inhibition of carotenoid synthesis by fluridone and norflurazon. *Weed Science*, 26(2), pp.198-203.
- Beck, N. y Lang, B. 2010. MFannot, organelle genome annotation webserver. <http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>.
- Ben-Amotz, A. and Avron, M., 1980. Glycerol, beta-carotene and dry algal meal production by commercial cultivation of *Dunaliella*. *Algae biomass: production and use/ [sponsored by the National Council for Research and Development, Israel and the Gesellschaft fur Strahlen-und Umweltforschung (GSF), Munich, Germany]; editors, Gedaliah Shelef, Carl J. Soeder.*
- Ben-Amotz, A., Shaish, A. y Avron, M. 1989. Mode of action of the massively accumulated β -carotene of *Dunaliella bardawil* in protecting the alga against damage by excess irradiation. *Plant Physiology*, 91(3), pp.1040-1044.
- Ben-Amotz, A. and Avron, M., 1990. The biotechnology of cultivating the halotolerant alga *Dunaliella*. *Trends in Biotechnology*, 8, pp.121-126.
- Bock, R. 2015. Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. *Annual Review of Plant Biology*, 66, pp.211-241.
- Bogorad, L. 2000. Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products. *Trends in biotechnology*, 18(6), pp.257-263.
- Boisvert, S., Laviolette, F. y Corbeil, J. 2010. Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. *Journal of computational biology*, 17(11), pp.1519-1533.
- Brasil, B.D.S.A.F., de Siqueira, F.G., Salum, T.F.C., Zanette, C.M. y Spier, M.R. 2017. Microalgae and cyanobacteria as enzyme biofactories. *Algal Research*, 25, pp.76-89.

- Chamovitz, D., Pecker, I. y Hirschberg, J. 1991. The molecular basis of resistance to the herbicide norflurazon. *Plant molecular biology*, 16(6), pp.967-974.
- Choi, K.S. y Park, S. 2015. The complete chloroplast genome sequence of *Aster spathulifolius* (Asteraceae); genomic features and relationship with Asteraceae. *Gene*, 572(2), pp.214-221.
- Coil, D., Jospin, G. y Darling, A.E. 2014. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics*, pbtu661.
- Cui, Y., Qin, S. y Jiang, P. 2014. Chloroplast transformation of *Platymonas* (Tetraselmis) subcordiformis with the bar gene as selectable marker. *PloS one*, 9(6), p. e98607.
- Daniell, H., Datta, R., Varma, S., Gray, S. y Lee, S.B. 1998. Containment of herbicide resistance through genetic engineering of the chloroplast genome. *Nature biotechnology*, 16(4), p.345.
- Daniell, H., Cohill, P.R., Kumar, S., Dufourmantel, N. y Dubald, M. 2004. *Chloroplast genetic engineering* (pp. 443-490). Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Daniell, H., Lin, C.S., Yu, M. y Chang, W.J. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome biology*, 17(1), p.134.
- Darling, A.C., Mau, B., Blattner, F.R. y Perna, N.T. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome research*, 14(7), pp.1394-1403.
- Del Vasto, M., Figueroa-Martinez, F., Featherston, J., Gonzalez, M.A., Reyes-Prieto, A., Durand, P.M. y Smith, D.R. 2015. Massive and widespread organelle genomic expansion in the green algal genus *Dunaliella*. *Genome biology and evolution*, 7(3), pp.656-663.
- Devos, R., Plaetinck, G., Cheroutre, H., Simons, G., Degrave, W., Tavernier, J., Remaut, E. y Fiers, W. 1983. Molecular cloning of human interleukin 2 cDNA and its expression in *E. coli*. *Nucleic Acids Research*, 11(13), pp.4307-4323.
- Doron, L., Segal, N.A. y Shapira, M. 2016. Transgene expression in microalgae—From tools to applications. *Frontiers in plant science*, 7.
- Economou, C., Wannathong, T., Szaub, J. y Purton, S. 2014. A simple, low-cost method for chloroplast transformation of the green alga *Chlamydomonas reinhardtii*. *Chloroplast biotechnology: Methods and protocols*, pp.401-411.
- Feng, S., Xue, L., Liu, H. y Lu, P. 2009. Improvement of efficiency of genetic transformation for *Dunaliella salina* by glass beads method. *Molecular biology reports*, 36(6), p.1433.

- Feng, S., Li, X., Xu, Z. y Qi, J. 2014. *Dunaliella salina* as a novel host for the production of recombinant proteins. *Applied microbiology and biotechnology*, 98(10), pp.4293-4300.
- Feng, S., Feng, W., Zhao, L., Gu, H., Li, Q., Shi, K., Guo, S. y Zhang, N. 2014a. Preparation of transgenic *Dunaliella salina* for immunization against white spot syndrome virus in crayfish. *Archives of virology*, 159(3), pp.519-525.
- Feng, S., Li, S., Li, Q., Shi, K. and Xue, L. 2014b. Preparation of recombinant human canstatin using transgenic *Dunaliella salina*. *Acta Biochim Biophys Sin*, 46(5), pp.428-430.
- Franklin, S., Ngo, B., Efuet, E. y Mayfield, S.P. 2002. Development of a GFP reporter gene for *Chlamydomonas reinhardtii* chloroplast. *The Plant Journal*, 30(6), pp.733-744.
- Fučíková, K., Lewis, P.O. y Lewis, L.A. 2016. Chloroplast phylogenomic data from the green algal order Sphaeropleales (Chlorophyceae, Chlorophyta) reveal complex patterns of sequence evolution. *Molecular phylogenetics and evolution*, 98, pp.176-183.
- Gan, S.Y., Lim, P.E. y Phang, S.M. 2016. Genetic and Metabolic Engineering of Microalgae. In *Algae Biotechnology* (pp. 317-344). Springer International Publishing.
- Geng, D., Wang, Y., Wang, P., Li, W. y Sun, Y. 2003. Stable expression of hepatitis B surface antigen gene in *Dunaliella salina* (Chlorophyta). *Journal of Applied Phycology*, 15(6), pp.451-456.
- Geng, D.G., Han, Y., Wang, Y.Q., Wang, P., Zhang, L.M., Li, W.B. y Sun, Y.R. 2004. Construction of a system for the stable expression of foreign genes in *Dunaliella salina*. *Acta Bot Sin*, 46(3), pp.342-345.
- Georgianna, D.R., Hannon, M.J., Marcuschi, M., Wu, S., Botsch, K., Lewis, A.J., Hyun, J., Mendez, M. y Mayfield, S.P. 2013. Production of recombinant enzymes in the marine alga *Dunaliella tertiolecta*. *Algal Research*, 2(1), pp.2-9.
- Gimpel, J.A. y Mayfield, S.P. 2013. Analysis of heterologous regulatory and coding regions in algal chloroplasts. *Applied microbiology and biotechnology*, 97(10), pp.4499-4510.
- Gong, Y., Hu, H., Gao, Y., Xu, X. y Gao, H. 2011. Microalgae as platforms for production of recombinant proteins and valuable compounds: progress and prospects. *Journal of industrial microbiology & biotechnology*, 38(12), pp.1879-1890.

- Gutierrez, C.L., Gimpel, J., Escobar, C., Marshall, S.H. y Henríquez, V. 2012. Chloroplast genetic tool for the green microalgae *Haematococcus pluvialis* (Chlorophyceae, Volvocales) 1. *Journal of phycology*, 48(4), pp.976-983.
- Heitzer, M., Eckert, A., Fuhrmann, M. y Griesbeck, C. 2007. Influence of codon bias on the expression of foreign genes in microalgae. *Transgenic microalgae as green cell factories*, pp.46-53.
- Hempel, F. y Maier, U.G. 2016. Microalgae as Solar-Powered Protein Factories. In *Advanced Technologies for Protein Complex Production and Characterization* (pp. 241-262). Springer International Publishing.
- Hoerlein, G. 1994. Glufosinate (phosphinothricin), a natural amino acid with unexpected herbicidal properties. In *Reviews of environmental contamination and toxicology* (pp. 73-145). Springer New York.
- Hosseini Tafreshi, A. y Shariati, M. 2009. *Dunaliella* biotechnology: methods and applications. *Journal of Applied Microbiology*, 107(1), pp.14-35.
- Jensen, P.E. y Leister, D. 2014. Chloroplast evolution, structure and functions. *F1000Prime Rep*, 6(40), pp.10-12703.
- Jia, Y., Li, S., Allen, G., Feng, S. y Xue, L. 2012. A novel glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter for expressing transgenes in the halotolerant alga *Dunaliella salina*. *Current microbiology*, 64(5), pp.506-513.
- Jin, S. y Daniell, H. 2015. The engineered chloroplast genome just got smarter. *Trends in plant science*, 20(10), pp.622-640
- Jones, C.S., Luong, T., Hannon, M., Tran, M., Gregory, J.A., Shen, Z., Briggs, S.P. y Mayfield, S.P. 2013. Heterologous expression of the C-terminal antigenic domain of the malaria vaccine candidate Pfs48/45 in the green algae *Chlamydomonas reinhardtii*. *Applied microbiology and biotechnology*, 97(5), pp.1987-1995.
- Jung, H.S. y Kim, D.S. 2016. Utility of the *pat* gene as a selectable marker gene in production of transgenic *Dunaliella salina*. *Fisheries and Aquatic Sciences*, 19(1), p.31.
- Klein, Uwe. 2008. Chloroplast Transcription, In David Stern (ed.), *The Chlamydomonas Sourcebook, Vol. 2: Organellar and Metabolic Processes*. Elsevier. ISBN 0-12-370875-3. Chapter 25. s 893 - 913
- Koller, M., Muhr, A. y Braunegg, G. 2014. Microalgae as versatile cellular factories for valued products. *Algal research*, 6, pp.52-63.
- Kurtz, S., Choudhuri, J.V., Ohlebusch, E., Schleiermacher, C., Stoye, J. y Giegerich, R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic acids research*, 29(22), pp.4633-4642.

- Lang, B.F., Laforest, M.J. y Burger, G. 2007. Mitochondrial introns: a critical view. *Trends in Genetics*, 23(3), pp.119-125.
- Lemieux, C., Otis, C. y Turmel, M. 2016. Comparative chloroplast genome analyses of streptophyte green algae uncover major structural alterations in the Klebsormidiophyceae, Coleochaetophyceae and Zygnematophyceae. *Frontiers in plant science*, 7.
- Li, J., Xue, L., Yan, H., Wang, L., Liu, L., Lu, Y. y Xie, H. 2007. The nitrate reductase gene-switch: a system for regulated expression in transformed cells of *Dunaliella salina*. *Gene*, 403(1), pp.132-142.
- Li, J., Xue, L., Yan, H., Liu, H. y Liang, J. 2008. Inducible EGFP expression under the control of the nitrate reductase gene promoter in transgenic *Dunaliella salina*. *Journal of Applied Phycology*, 20(2), pp.137-145.
- Li, J., Lu, Y., Xue, L. y Xie, H. 2010. A structurally novel salt-regulated promoter of duplicated carbonic anhydrase gene 1 from *Dunaliella salina*. *Molecular biology reports*, 37(2), p.1143.
- Liu, J., Sun, Z., Gerken, H., Huang, J., Jiang, Y. y Chen, F. 2014. Genetic engineering of the green alga *Chlorella zofingiensis*: a modified norflurazon-resistant phytoene desaturase gene as a dominant selectable marker. *Applied microbiology and biotechnology*, 98(11), pp.5069-5079.
- Lohse, M., Drechsel, O. y Bock, R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Current genetics*, 52(5-6), pp.267-274.
- Mayfield, S.P., Franklin, S.E. y Lerner, R.A. 2003. Expression and assembly of a fully active antibody in algae. *Proceedings of the National Academy of Sciences*, 100(2), pp.438-442.
- Mayfield, S.P. y Schultz, J. 2004. Development of a luciferase reporter gene, luxCt, for *Chlamydomonas reinhardtii* chloroplast. *The Plant Journal*, 37(3), pp.449-458.
- Mayfield, S.P., Manuell, A.L., Chen, S., Wu, J., Tran, M., Siefker, D., Muto, M. y Marin-Navarro, J. 2007. *Chlamydomonas reinhardtii* chloroplasts as protein factories. *Current opinion in biotechnology*, 18(2), pp.126-133.
- Mendoza-Damian, C. (2012). *Construcción de un vector para expresar la proteína Hemaglutinina del virus de la influenza A (H1N1) 2009 en el cloroplasto de Dunaliella salina*. Maestría. Universidad Autónoma de Baja California.
- Ng, P.K., Lin, S.M., Lim, P.E., Liu, L.C., Chen, C.M. y Pai, T.W. 2017. Complete chloroplast genome of *Gracilaria firma* (Gracilariaceae, Rhodophyta), with discussion on the use of chloroplast phylogenomics in the subclass Rhodymeniophycidae. *BMC genomics*, 18(1), p.40.

- Potvin, G. y Zhang, Z. 2010. Strategies for high-level recombinant protein expression in transgenic microalgae: a review. *Biotechnology Advances*, 28(6), pp.910-918.
- Purton, S. 2007. Tools and techniques for chloroplast transformation of *Chlamydomonas*. In *Transgenic microalgae as green cell factories* (pp. 34-45). Springer New York.
- Purton, S., Szaub, J.B., Wannathong, T., Young, R. y Economou, C.K. 2013. Genetic engineering of algal chloroplasts: progress and prospects. *Russian journal of plant physiology*, 60(4), pp.491-499.
- Qin, S., Lin, H. y Jiang, P. 2012. Advances in genetic engineering of marine algae. *Biotechnology advances*, 30(6), pp.1602-1613.
- Rasala, B.A., Muto, M., Lee, P.A., Jager, M., Cardoso, R.M., Behnke, C.A., Kirk, P., Hokanson, C.A., Crea, R., Mendez, M. y Mayfield, S.P. 2010. Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of *Chlamydomonas reinhardtii*. *Plant biotechnology journal*, 8(6), pp.719-733.
- Rasala, B.A. y Mayfield, S.P. 2011. The microalga *Chlamydomonas reinhardtii* as a platform for the production of human protein therapeutics. *Bioengineered bugs*, 2(1), pp.50-54.
- Rasala, B.A., Muto, M., Sullivan, J. and Mayfield, S.P. 2011. Improved heterologous protein expression in the chloroplast of *Chlamydomonas reinhardtii* through promoter and 5' untranslated region optimization. *Plant biotechnology journal*, 9(6), pp.674-683.
- Rosenberg, S.A. 2014. IL-2: the first effective immunotherapy for human cancer. *The Journal of Immunology*, 192(12), pp.5451-5458.
- Ruhlman, T., Verma, D., Samson, N. and Daniell, H., 2010. The role of heterologous chloroplast sequence elements in transgene integration and expression. *Plant physiology*, 152(4), pp.2088-2104.
- Sablok, G., Mudunuri, S.B., Edwards, D. y Ralph, P.J. 2016. Chloroplast genomics: Expanding resources for an evolutionary conserved miniature molecule with enigmatic applications. *Current Plant Biology*.
- Sablok, G., Chen, T.W., Lee, C.C., Yang, C., Gan, R.C., Wegrzyn, J.L., Porta, N.L., Nayak, K.C., Huang, P.J., Varotto, C. y Tang, P. 2017. ChloroMitoCU: Codon patterns across organelle genomes for functional genomics and evolutionary applications. *DNA Research*.
- Salvador, M.L., Suay, L., Anthonisen, I.L. y Klein, U. 2004. Changes in the 5'-untranslated region of the *rbcl* gene accelerate transcript degradation more

- than 50-fold in the chloroplast of *Chlamydomonas reinhardtii*. *Current genetics*, 45(3), pp.176-182.
- Schattner, P., Brooks, A.N. y Lowe, T.M. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic acids research*, 33(suppl 2), pp. W686-W689.
- Shamriz, S. y Ofoghi, H., 2016. Outlook in the application of *Chlamydomonas reinhardtii* chloroplast as a platform for recombinant protein production. *Biotechnology and Genetic Engineering Reviews*, 32(1-2), pp.92-106.
- Sharon-Gojman, R., Maimon, E., Leu, S., Zarka, A. y Boussiba, S. 2015. Advanced methods for genetic engineering of *Haematococcus pluvialis* (Chlorophyceae, Volvocales). *Algal Research*, 10, pp.8-15.
- Shi, C., Hu, N., Huang, H., Gao, J., Zhao, Y.J. y Gao, L.Z. 2012. An improved chloroplast DNA extraction procedure for whole plastid genome sequencing. *Plos one*, 7(2), p. e31468.
- Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K. y Ohto, C. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *The EMBO journal*, 5(9), p.2043.
- Smith, A.C. y Purton, S. 2002. The transcriptional apparatus of algal plastids. *European Journal of Phycology*, 37(3), pp.301-311.
- Smith, D.R., Lee, R.W., Cushman, J.C., Magnuson, J.K., Tran, D. y Polle, J.E. 2010. The *Dunaliella salina* organelle genomes: large sequences, inflated with intronic and intergenic DNA. *BMC plant biology*, 10(1), p.83.
- Specht, E., Miyake-Stoner, S. y Mayfield, S. 2010. Micro-algae come of age as a platform for recombinant protein production. *Biotechnology letters*, 32(10), pp.1373-1383.
- Specht, E.A., Karunanithi, P.S., Gimpel, J.A., Ansari, W.S. and Mayfield, S.P., 2017. Host Organisms: Algae. *Industrial Biotechnology: Microorganisms*, pp.605-641.
- Suarez, J.V., Banks, S., Thomas, P.G. y Day, A. 2014. A new F131V mutation in *Chlamydomonas* phytoene desaturase locates a cluster of norflurazon resistance mutations near the FAD-binding site in 3D protein models. *PloS one*, 9(6), p. e99894.
- Sun, Y., Yang, Z., Gao, X., Li, Q., Zhang, Q. y Xu, Z. 2005. Expression of foreign genes in *Dunaliella* by electroporation. *Molecular biotechnology*, 30(3), pp.185-192.

- Surzycki, R., Greenham, K., Kitayama, K., Dibal, F., Wagner, R., Rochaix, J.D., Ajam, T. y Surzycki, S. 2009. Factors effecting expression of vaccines in microalgae. *Biologicals*, 37(3), pp.133-138.
- Tan, C., Qin, S., Zhang, Q., Jiang, P. y Zhao, F. 2005. Establishment of a micro-particle bombardment transformation system for *Dunaliella salina*. *JOURNAL OF MICROBIOLOGY-SEOUL*, 43(4), p.361.
- Tran, M., Zhou, B., Pettersson, P.L., Gonzalez, M.J. y Mayfield, S.P. 2009. Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnology and bioengineering*, 104(4), pp.663-673.
- Tran, M., Van, C., Barrera, D.J., Pettersson, P.L., Peinado, C.D., Bui, J. y Mayfield, S.P. 2013. Production of unique immunotoxin cancer therapeutics in algal chloroplasts. *Proceedings of the National Academy of Sciences*, 110(1), pp. E15-E22.
- W. Polle, J.º.E., Tran, D. y Ben-Amotz, A. 2009. History, distribution, and habitats of algae of the genus *Dunaliella* Teodoresco (Chlorophyceae). In *The Alga Dunaliella* (pp. 1-13). Science Publishers.
- Wang, T., Xue, L., Hou, W., Yang, B., Chai, Y., Ji, X. y Wang, Y. 2007. Increased expression of transgene in stably transformed cells of *Dunaliella salina* by matrix attachment regions. *Applied microbiology and biotechnology*, 76(3), pp.651-657.
- Wijffels, R.H. y Barbosa, M.J. 2010. An outlook on microalgal biofuels. *Science*, 329(5993), pp.796-799.
- Wu, S., Xu, L., Huang, R. y Wang, Q. 2011. Improved biohydrogen production with an expression of codon-optimized hemH and lba genes in the chloroplast of *Chlamydomonas reinhardtii*. *Bioresource technology*, 102(3), pp.2610-2616.
- Xie, W.H., Zhu, C.C., Zhang, N.S., Li, D.W., Yang, W.D., Liu, J.S., Sathishkumar, R. y Li, H.Y. 2014. Construction of novel chloroplast expression vector and development of an efficient transformation system for the diatom *Phaeodactylum tricornutum*. *Marine biotechnology*, 16(5), pp.538-546.
- Yao, S., Lu, J., Sárossy, Z., Baggesen, C., Brandt, A. y An, Y. 2016. Neutral lipid production in *Dunaliella salina*. *Journal of Applied Phycology*, 28(4), pp.2167-2175.
- Zaslavskaya, L.A., Lippmeier, J.C., Kroth, P.G., Grossman, A.R. y Apt, K.E. 2000. Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *Journal of Phycology*, 36(2), pp.379-386.

- Zhang, D., Li, K., Gao, J., Liu, Y. y Gao, L.Z. 2016. The Complete Plastid Genome Sequence of the Wild Rice *Zizania latifolia* and Comparative Chloroplast Genomics of the Rice Tribe Oryzeae, Poaceae. *Frontiers in Ecology and Evolution*, 4, p.88.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*, 31(13), pp.3406-3415.

Anexos

Descripción de plásmidos

Tabla V. Descripción de plásmidos

Plásmido	Descripción
pUC57-utr's	Contiene el promotor y 5'UTR del gen <i>atpA</i> y el 3' UTR del gen <i>psbA</i> de <i>D. salina</i>
pNat-utr	Gen <i>nat-1</i> clonado en el vector pUC57-utr's
pUC57-il2	Contiene el gen <i>il-2</i> de humano optimizado para su expresión en cloroplasto de <i>D. salina</i>
pIL2-utr	Gen <i>il-2</i> optimizado clonado en el vector pUC57-utr's
pIL2-Nat	Contiene el gen <i>il-2</i> unido a el promotor y los UTR's y el gen <i>nat-1</i> que también contiene el promotor y los UTR's
pIN3'	Plásmido pIL2-Nat con la región de recombinación homóloga del extremo 3' clonada
pDsCh	Plásmido para transformar en el cloroplasto de <i>D. salina</i> . Contiene el gen <i>il-2</i> con promotor y UTR's, el gen <i>nat-1</i> con promotor y UTR's y las regiones de recombinación homologas.
pEGFP	Contiene el gen <i>egfp</i> optimizado para su expresión en cloroplasto de <i>D. salina</i> , regulado bajo el promotor y 5' UTR del gen <i>rbcl</i> y el 3' UTR del gen <i>psbA</i> ; el gen <i>nat-1</i> optimizado para su expresión en cloroplasto de <i>D. salina</i> y regulado bajo el promotor y 5' UTR del gen <i>psbA</i> y el 3' UTR del gen <i>atpA</i>
pGFP5	Plásmido pEGFP con la región homóloga 5' clonada
pDsEGFP	Plásmido para transformar en el cloroplasto de <i>D. salina</i> . Contiene el gen <i>egfp</i> optimizado para su expresión en cloroplasto de <i>D. salina</i> , regulado bajo el promotor y 5' UTR del gen <i>rbcl</i> y el 3' UTR del gen <i>psbA</i> ; el gen <i>nat-1</i> optimizado para su expresión en cloroplasto de <i>D. salina</i> y regulado bajo el promotor y 5' UTR del gen <i>psbA</i> y el 3' UTR del gen <i>atpA</i> ; y las regiones de recombinación homóloga 5' y 3'.
pDunaCp	Plásmido que contiene el gen <i>pdsF128V</i> , que confiere resistencia al herbicida norflurazon
pGFP/PDS	Plásmido para transformar en el cloroplasto de <i>D. salina</i> . Contiene el gen <i>egfp</i> optimizado para su expresión en cloroplasto de <i>D. salina</i> , regulado bajo el promotor y 5' UTR del gen <i>rbcl</i> y el 3' UTR del gen <i>psbA</i> ; el gen <i>pdsF128v</i> de <i>D. salina</i> regulado bajo el promotor y 5' UTR del gen <i>psbA</i> y el 3' UTR del gen <i>atpA</i> ; y las regiones de recombinación homóloga 5' y 3'.

Primers

Tabla VI. Lista de primers

Primer	Secuencia 5' - 3'
Nat1-for	CATATACCATGGATGACCACTCTTGACGACAC
Nat1-rev	GATCATCTGCAGGCCGCTCAGGCCG
Pro-for	GCATATAGGCCTGGTACCAATATTAATTTAA
UTR-rev	GATCATGGATCCGGTACCAAAAAAATTTTG
PrNde-F	GTCTCATATGGGTACCAATATTAATTTAAATAAT
UTRBst-R	GTCGCGCCATTGGATTGGGGTACCAAAAAAATTTTG
Infusion-1	CGGCCAGCTGAATTCGATATCGGATAAAAGACCAGGAAC
InPvu-R	CGGGAGGGCTCCGCTAGCATATCCCTCACGGGTCGCG
InfNde-F	CTAAATATCGAAGACCATAGTAACCAATGGTGTCAAGC
Inf2Nde-R	CTTATATTCTGCACCCAGGTGCATATGCCAATCGGGAATTCAC
Hom5GFP-F	GCCAGTGAATTCGATATCGGATAAAAGACCAGGAACCG
Hom5GFP-R	GGGCTCCGCTAAGATATCCCTCACGGGTCGCGATTAACG
Hom3GFP-F	AATATCGAAGACCATGTAACCAATGGTGTCAAGC
Hom3GFP-R	ATATTCTGCACCCAGGTGCCAATCGGGAATTCAC
NatF	ATGGATGACCACTCTTGACGACAC
NatR	GGGGCAGGGCATGCTCATGTAG
atpaF	GGCAATGCGCACACCAGAAG
atpaR	GACCACGACCTACTGGAATC
Atpa2F	CGTGAGCTTATTATTGGTGACCG
Atpa2R	CACGAAGACGGTAACCACGAGC
InF	GGATAAAAGACCAGGAACCG
InR.	CATAACCTTCAGGCATAGCAG
PdsAfe-	GTTAAGTTAAATCACAGCGCTTTAGAACCAAGGAAATGTTACAGC
PdsDra-R	GTTAAGTTAAATCACTGCGTGTTAGAACCAAGGAAATGTTACAGC
PdsF	ATGCAAGTTATGCAAGGTCGTGCTC
PdsR	CATAAACGTTTCAGGTGGAGCACCGTC
PdsF2	CCCTGATGAATTATCAATGACTGTAG
PdsR2	TTAGAACCAAGGAAATGTTACAGCTTC

Secuencia del vector pDsCh

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTG
 TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTGCGGGGCTGGCT
 TAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGGGATAAAAAGACCAGGAACCGACTTA
 GCGCGCCGCCCTCACGCGACGGATCTTGATCCTAGCCCCCTAGGCGACGGACAAACGCCCGCGCCCCCGTG
 AGGGGCCCTTAGGGGGGCGCTTAATTTTTTTAAAAATGTGGAATCAGGGCAACCTGATAGGACTGATCCGT
 GTAGGAGGGGTACGAACCTTCTATCCTAGTCTAAGTGGCTGGCGAATCTAATTGGCCTCGGCTCAGAAAGGT
 CCACTAGAGAAAAGCGAAAGGGAAGTGGAAATTTTATACGGACCCGCCTATAACCTGCCATGAATGGGTGTAG
 TTTACCAAAGGACAAGTTTGGTCTCCCTTCTTGGCCCTATGGATGTCATAAACACTCTACTAGGCGGGTC
 CGAAAGTGGTCAGAACGGAAAAGGCCCTTTCATTTAATTTTCAATTAATTATGTCGTTAGTTTTAGCTAACA
 ACCCTAACATCGAACTTCCGCAAGCATAGAACCAGGTTGACTGGAAAGCTGTGTTTTCGGTCGTTCAATCG
 TTAAGGGAACAAGCCTTTATCTGTTTCAAGTCTGGGGCGGGGGCGCCAGTGAAGTTGCGTGAGGTTCAAA
 AGCAAATTTCTGACCCGTGAGGGCACCTGCTACTTTATTACTGGCTATTTCGTAGAGTTACGTCTAAGTCGG
 TCCGCGAAACGGCGGAAGTAGATGAGCTTATTGTTGTAAACGATTTCAGACAAATGGGCCATCTTTGATGGC
 TTCGTGGGTATGGATCTAGCTCATTGGCAACCTCCTCCTGTTTCGAGGCGGGTTCGCTATATTCTAAAGGC
 TGACAAGCGACGTTGGAGACCGCTTGGTATTCCACAGTTAAGGATCGTGTCTATCCAGGCGGTTGTGCCGG
 AGGCACCTGAACCTGAGTGGGACGTAAAGTTTGAAGGAAATTCCTATGGTTATAGACCAAAGCGTAACAGA
 GCTGATGCTCTTCAACGCACCTTCAATCATCTAGCGGATTCTCCTACGTCACAACCAAACGTTAATCGCGA
 CCCGTGAGGGGGGCGGTAGTAGCGGATATCTCTGGTGGCTTTGATAACATTTCTCATAAGAGTGTTATCGA
 CGCTGTTTCAGGAATTTCTGCTCATTCTTAGTTGAACGCTGGTTAAGAGCTGGAATTTCTGAACCCGTGA
 GGGCGGCGTCTTCGTGGAACCTTCGGCGGGGCGAGGCTGTTCCGCGAGGGCGGGGTTCATGTCTCCTCTTTTGT
 GTAATATCGTCTTACATGGTCTAGGTGACATTATTTATAATTTATATTTAATTTATAATTTATAATTTATA
 TTTATGTTTCTATTTATATTTTATATTTATGTTTCTGGAAGATCCACGGAAGCAAAAAGCCGATAATGGT
 ACGCTATGCGGATGATTTTATCCTTCTTTGTAAGTACTACGAGTCTGCCAAGGACGAACCATATGGGTACC
 AATATTAATTTAAATAATATTGTTATTTATTAATAGGTGGTGTCAATTTACGCCACCAAAGCGCGCACAAG
 CGCGGAAAAAAAAAAAAAAAAAATTTATTAATAATAAATTTTAAATGATCTCTCCGGAATATATAGATAATA
 ACAAAATTCAGAAGAAATTTTATAGAAGTTTAAATATCCTAAAGAATATTTAAATCTCTTAAATTTGTTA
 TCTATTTATAACAATTTTATCGGTTTTTTAGAAAAAAAAAAAGTATCTCCTAAATTAATAAGAAAAATTC
 TTAGATAAGGGTATGAATGACGACCTTACTTATTTTTTCTTAGGTGCTTAAAAATACAGCAAATAGAATA
 ATATATTTAAATATAAAGTGTTTTCGTTTACGACGGAAGTATAAATTTTAAATATATTATTAACAATA
 GTTTATAAGTTATAAACTATTTGTTTCTATCATATCCCTTTAACAAGAAAAAAAAATTTTTTCCATGGCAC
 CACCACCACCACGATGACGATGACAAGATGTACCGTATGCAATTATTATCATGTATTGCTTTATCATT
 AGCTTTAGTAACCTAAGCTCCTCACTTCACTCACTAAAAAACTCAATTACAATTAGAACACTTAT
 TATTAGACTTACAATGATTTTAAACGGTATTAACAATAAAAAACCCAAAATTAACCTCGTATGTTAACT
 TTCAAATTTACATGCCAAAAAAGCTACTGAATTAACAACTTACAATGTTTAGAAGAAGAATTAACACC
 ATTAGAAGAAGTATTAACCTTAGCTCAATCAAAAACTTCCACTTACGTCCACGTGACTTAATTTCAAACA
 TTAACGTAATTGTATTAGAATTAAGGTTTCAGAACTACTTTTCATGTGTGAATACGCTGACGAACTGCT
 ACTATTGTAGAATTTTAAACCGTTGGATTACTTTCTGTCAATCAATTATTTCACTTTAACTTAACGCA
 GTAATTTCTTTATACTTAAGAATTAAGTAGACTTGTCTACTTAATTTCTTTTTAATAGCATGGATGGAGTAAA
 CTCCATCCATGCTATTATCTCTAACAACCCCAAAAACCAAAGGTTTTTGTGAAGGAAAAATAGAATATTT
 TTTCCGCTACGAGCGGGCGATATAAAAAAATAAGATTTTTTTATCCTAAATTAATTTCTGTAATAGAATCT
 CAATTAATGACAAAGATATAAATTACAAAAACGGTAGAGTTTCAGGTTTAGATAAATAACCTGGTCCTGCT
 ATATACTAATGTAAGTTAAATATCAGTATATGAACAATATATATTATATTAATAGTCCATGTCCTCC
 AAAACGAATAAACGAATAAACCAAAGAAAATAAACTTTATTATTTTAAATAAAATAAAATTTTATTAAT
 TTTTTTTTTTATCAAAAATTTTTTGGTACCCCAATCCAATGGTACCAATATTAATTTAAATAATATTGTT
 ATTTATTAATAGGTGGTGTCAATTTACGCCACCAAAGCGCGCACAAAGCGGAAAAAAGAAAAAATTT
 TATTAATAATAAATTTTTAATGATCTCTCCGGAATATATAGATAATAACAAATTCAGAAGAAATTTTTAGA
 AGTTTTAAATATCCTAAAGAATATTTAAATCTTCTAAAATTTGTTATCTATTTATAACAATTTTATCGGT
 TTTTATAGAAAAAAGTATCTCCTAAATTAATTAAGAAAAATTTCTAGATAAGGGTATGAATGACGAC
 CTTACTTATTTTTTCTTAGGTGCTTAAAAATACAGCAAATAGAATAATATATTTAAATATAAAGTGTT

TCGTTTACGACGGAAAGTATAAATTTTAAATATATTATTAAACAATAGTTTATAAGTTATAAACTATTTGT
 TTCTATCATATCCCTTTAACAAGAAAAAAATTTTTTTCATGGATGACCACTCTTGACGACACGGCTTAC
 CGGTACCGCACCAGTGTCCCGGGGACGCCGAGGCCATCGAGGCACTGGATGGGTCTTACCACCGACAC
 CGTCTTCCGCGTCACCGCCACCGGGGACGGCTTCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCGTACCA
 AGGTGTTCCCCGACGACGAATCGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTCCCGGACG
 TTCGTGCGGTACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTTCGTCTCGTACTCCGGCTGGAACCGCCG
 GCTGACCGTCGAGGACATCGAGGTGCCCCCGGAGCACCGGGGGCACGGGGTTCGGGCGCGCGTGTGATGGGGC
 TCGCGACGGAGTTGCCCCGCGAGCGGGGCGCCGGGCACCTCTGGCTGGAGGTACCAACGTCAACGCACCG
 GCGATCCACGCGTACCGGCGGATGGGGTTCACCCTCTGCGGCCTGGACACCGCCCTGTACGACGGCACCGC
 CTCGGACGGCGAGCAGGCGCTCTACATGAGCATGCCCTGCCCCCTGACTGCAGTAATTCTTTATACTTAAGA
 ATTAAGTAGACTTGTCTACTTAATTCCTTTTAAATAGCATGGATGGAGTAACTCCATCCATGCTATTATCT
 CTAACAACCCCCAAAAACCAAGGTTTTTGTGAAGGAAAAATAGAATATTTTTTCCGCTACGAGCGGGCGA
 TATAAAAAAATAAGATTTTTTTATCCTAAATTAATTCCTGTAATAGAATCTCAATTAATGACAAAGATATA
 AATTACAAAAACGGTAGAGTTTCAGGTTTAGATAAAATAACCTGGTCCTGCTATATACTAATGTAAGTTAA
 AATATCAGTATATGAACAATATATATTATATATTAATAGTCCATGTCCTTCAAAAACGAATAAACGAATAA
 ACCAAAGAAAATAAAACTTTTATTATTTTAAATAAAATAAAATTTTTATTATTTTTTTTTTTATCAAAAATT
 TTTTTGGTACCAAGCTTCTACGAGTCTGCCAAGGACGAACCCGCCTTACGCTTTGCCGAAGAATTTTTTAC
 TTCTCCTCGTGAGGGGGCTTAAATTTAAGGAAACAAATGAGCCTGTTGTTGTACACATCACACAGGGTTT
 TGACTTCCTGAATTGCACATATAGACGTGTAATAATTACGGTTATAATAAAGCTTTCACCATTTCGTCCGC
 CGATCTTCGAGGGTGCAGAACTTGTGTAACGAACCCGCCTTGGTACGATGAATCCGTAACCAATGGTGTCA
 AGCACACCATTGTACTGGTAAGTGTTCCTAAGGATAAGCTGAAGTCTGTTTCGTGCAGAATATAAGAAAGTC
 TTCCGCTTCTACCGTGGTTTTCAAGATAAGTGGGCTTGTTAATGCGTTAACGACCCCGCCAAAGTTAAGGG
 TTTTCGCTCAATCGTGTCAAACGATTGATTGTACCAGAGTCTTCTGTCATCTGGATAACTACATTTTCAGAT
 TGACGGTACGCTTACTTCGTAGAGAGCATCAGTCCAAGTCTTGGGCTTGGATAATAGCTCGCTATTTTCAGA
 AAGGTCAAGCTTCCTAGAATAGACTCTCCATGGGTTCTAGGGGTGGGGGCGTCGAACCTCGGATTTGACTCT
 ATTACCACACAGATACTTTTCTAAGGTAGACTATATTCTGGTACGTAGCGATATGTTTAAAGATGATCCGT
 CGCCTGTTGCTAAGGAGTATTGGGCTCGACGCCAACAACTTTTTTTTTTGACCGTAAAGCGGTTCGACCTGC
 TCGCTGGTTTAGACAAAGACTTATCAGCATCACAGGAGCACGTGTGCCGGTTTGTAAATTCATCTTTATAC
 TCCGGTGAAAGGTTGCAGAGACATCACGTTATAGAAAGATCTTGGGTGGAAGGGATACGTTCCGAAACCT
 ACTTCTTGTACATTCTGTTTGTATCGTCGAGTGCATTATGGCGGTGATGCGGATCTTTGGCGAGCTTACT
 TATCGTCTTATAAAGTAACCTATCCGAAGAAGTCAACACGTGGAAAGGTAAGTCCAGTCCAACAGGTTAAC
 CAAAACGATTTCCGTCAGGCGGGTTCGTTTCGGTTTTCTGGAACCCACCTGAATTTAAATTTGGGTGGTTTTAC
 AGGCGGTTCGTGTTTCGTAACATAAACGCGCTGTTCTAATTTATTAGTGCTTCGATAAGTACCGTAAACGT
 ATTTTATACATTATTTATGCCTATAAGAAAACCTACCTTTTTCGTCTCCTTTCTCTTTCGTCTCCTTTATTTT
 CTTCTCTTGTCACTGGAAGATCAGGGCCCAACGTGCCTTGGCCTTCCATTTGTGAATTCCCGATTGGCACC
 AGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGTTTGCATATTGGGCGCTCTTCCGCTTCCTCG
 CTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACG
 GTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
 GTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGC
 TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT
 GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGC
 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCGCTCCAGCTGGGCTGTGTGCAC
 GAACCCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGAGTCCAACCCGGTAAGACA
 CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAG
 AGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAG
 CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
 TTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGG
 GGTCTGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC
 ACCTAGATCCTTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
 CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCATCTGTCTATTTTCGTTTCATCCATAGTTGCCT
 GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCG
 CGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAG
 TGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTTCGC
 CAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTTCGTTTGGTATG

GCTTCATT CAGCTCCGGT TCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGT
 TAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAG
 CACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCC
 ACATAGCAGAACTTTAAAAGTGCTCATCATTGGAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTAC
 CGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACCTGATCTTCAGCATCTTTTACTTTACC
 AGCGTTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATG
 TTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGAT
 ACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCT
 GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCTGTC

Secuencia del vector pDsEGFP

TCGCGCGTTTTCCGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTG
 TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGCTGGCT
 TAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCG
 TAAGGAGAAAATACCGCATCAGGCGCCATTTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG
 CGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCC
 AGGGTTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTCGATATCGGATAAAAGACCAGGAACC
 GACTTAGCGCGCCGCCCTCACGCGACGGATCTTGATCCTAGCCCTAGGCGACGGACAAACGCCCCCGCCC
 CCCGTGAGGGGGCCCTTAGGGGGGCGCTTAATTTTTTTTAAAAATGTGGAATCAGGGCAACCTGATAGGACTG
 ATCCGTGTAGGAGGGGTACGAACCTCTATCCTAGTCTAAGTGGCTGGCGAATCTAATTGGCCTCGGCTCAG
 AAAGGTCCACTAGAGAAAGCGAAAGGGAAGTGAATTTTATACGGACCCGCCTATAACCTGCCATGAATGG
 GTGTAGTTTTACCAAAGGACAAGTTTGGTCTCCCTTCTTGCCCTATGGATGTCATAAACACTCTACTAGG
 CGGGTCCGAAAGTGGTCAGAACGGAAAAGGCCTTTCAATTAATTTTCATTTAATTATGTCGTTAGTTTTAG
 CTAACAACCCTAACATCGAACTTCCGCAAGCATAGAACCAGGTTGACTGGAAAGCTGTGTTTTCGGTTCGTT
 CAATCGTTAAGGGAACAAGCCTTTATCTGTTTCAAGTCTGGGGCGGGGGCGCCAGTGAAGTTGCGTGAGG
 TTCAAAGCAAATTTCTTGACCCGTGAGGGCACCTGCTACTTTTATTACTGGCTATTTCGTAGAGTTACGTCTA
 AGTCGGTCCGCGAAACGGCGGAAGTAGATGAGCTTATTGTTGTAAACGATTTCAGACAAATGGGCCATCTTT
 GATGGCTTCGTGGGTATGGATCTAGCTCATTGGCAACCTCCTCCTGTTTCGACGGCGGGTTCGCTATATTCT
 AAAGGCTGACAAGCGACGTTGGAGACCGCTTGGTATTTCCACAGTTAAGGATCGTGTATCCAGGCGGTTG
 TGCCGGAGGCACCTGAACCTGAGTGGGACGTAAAGTTTGAAGGAAATTCCTATGGTTATAGACCAAAGCGT
 AACAGAGCTGATGCTCTTCAACGCACCTTCAATCATCTAGCGGATTCTCCTACGTCACAACCAAACGTTAA
 TCGCGACCCGTGAGGGATATCTTAGCGGAGCCCTCCCGTGAGGGGGGCTAGGACTTTGTCCGTCGCCTAAG
 GGCTAGGATCAAGATCCGTCGCGTGAGGGAACAAAAAAGTTGTTTTGTCCAGAAGTATTTTTTTCTTATA
 AAAAAAGAAAAAGAAAAAAAATTATAAAATTATAAAAGAATTTTATATTCTTTAGTTATCAAAATTATA
 ATTTGGTAAAAGATTATTATATTATCGGATAGATTATTTTAGGATCGACAAAAATGTTTCATTTTTAGATA
 ATAAAGCGGCCGCTCTAGACCGCGGATGCATCATCATCATCATATGGTTTCTAAAGGTGAAGAATTAT
 TTACAGGTGTTGTTTCTATTTTAGTTGAATTAGATGGTGATGTTAATGGTCATAAATTTTCTGTTTCTGGT
 GAAGGTGAAGGTGATGCTACATATGGTAAATTAACATTAATAATTTATTTGTACAACAGGTAAATTACCTGT
 TCCTTGCCCTACATTAGTTACAACATTAACATATGGTGTTCATGTTTTTCTCGTTATCCTGATCATATGA
 AACAACATGATTTTTTTTAAATCTGCTATGCCTGAAGGTTATGTTCAAGAACGTACAATTTTTTTTAAAGAT
 GATGGTAATTATAAAACACGTGCTGAAGTTAAATTTGAAGGTGATACATTAGTTAATCGTATTGAATTA
 AGGTATTGATTTTTAAAGAAGATGGTAATATTTTAGGTCATAAATTAGAATATAATTATAATTCTCATAATG
 TTTATATTATGGCTGATAAACAAAAAATGGTATTAAGTTAATTTTTAAATTCGTCATAATATTGAAGAT
 GGTTCTGTTCAATTAGCTGATCATTATCAACAAAATACACCTATTGGTGATGGTCCTGTTTTATTACCTGA
 TAATCATTATTTATCTACACAATCTGCTTTATCTAAAGATCCTAATGAAAAACGTGATCATATGGTTTTAT
 TAGAATTTGTTACAGCTGCTGGTATTACTTTAGGTATGGATGAATTATATAAATAAGGATCCGAGCTCCTC
 GAGTTCTTTTATACTTAAGAATTAAGTAGACTTGTCTACTTAATCTTTTTTAATAGCATGGATGGAGTAAAC
 TCCATCCATGCTATTATCTCTAACAACCCCAAAAACCAAAGTTTTTGTGAAGGAAAAAATAGAATATTTT

TTCCGCTACGAGCGGGCGATATAAAAAATAAGATTTTTTTATCCTAAATTAATTCTTGTAATAGAATCTC
 AATTAATGACAAAGATATAAATTACAAAAACGGTAGAGTTTCAGGTTTAGATAAATAACCTGGTCCTGCTA
 TATACTAATGTAAGTTAAATATCAGTATATGAACAATATATATTATATATTAATAGTCCATGTCCTTCA
 AAAACGAATAAACGAATAAACCAAAGAAAATAAACTTTATTATTTTAAATAAAATAAAATTTTATTAATT
 TTTTTTTTTATCAAAAATTTTTTTTTTTCCGGAGCTAGCCACAAAGTCCTAGCCCTTAGGGGGCGGGGGCG
 TTTGTCCGTCGCGTGAGAGAAATAACAAAGTTGTTTTGTCCGGTACACAATAAAAAATTATATAAAAGTTG
 TTGTTTATAAAAAGGATGCTCTTGCTTTTGATAAATTATCTGGTATTATAATAATACCAGGTTTTAATGAA
 AAAGCTTGAATAATAAAAAAATAAAAAAGTTATTTTTAACTTTTTTCGGAGAAAATCAAAAATTAATAAAA
 AATTATAAATTAGCGCTATGACTACTTTAGATGATACTGCTTATCGTTATCGTACTTCTGTTCTCGGTGAT
 GCTGAAGCTATTGAAGCTTTAGATGGTTCTTTTACTACTGATACTGTTTTTCGTGTTACTGCTACTGGTGA
 TGGTTTTACTTTACGTGAAGTTCCTGTTGATCCTCCTTTAACTAAAGTTTTTCCTGATGATGAATCTGATG
 ATGAATCTGATGCTGGTGAAGATGGTGAATCCTGATTCTCGTACTTTTGTTGCTTATGGTGAATCTGATG
 TTAGCTGGTTTTTGTGTTGTTTCTTATTCTGGTTGAATCGTCGTTTAACTGTTGAAGATTTGAAGTTGC
 TCCTGAACATCGTGGTCATGGTGTGGTCTGCTTTAATGGGTTTAGCTACTGAATTTGCTCGTGAACGTG
 GTGCTGGTCATTTATGGTTAGAAGTTACTAATGTTAATGCTCCTGCTATTTCATGCTTATCGTCGTATGGGT
 TTTACTTTTATGTGGTTTAGATACTGCTTTATATGATGGTACTGCTTCTGATGGTGAACAAGCTTTATATAT
 GTCTATGCCTTGTCTTAAACACGCAGTGATTTAACTTAACTTAATAAAAAATTTTTTTTACCGCCCCCCCC
 CTCACGGGGAGGGAGGCTAGGACAAGGTCCGTCGTAAAAAATAAATTTAATTTTATAAATTTTATTA
 ATTTAGTTACAAAATCTTCCATTTTCATCGGCCCGCCAAAAGATAGTAAAAAATATACATATCTTTTAG
 GCGGAGCCTATCTCCGGTCTTTTATAAAAAGAAAAACCGACATACAAAATACAAAATTGTATGGAGGCGGA
 CAAAACAAATTTGTTGTTTTGTCCTAAGGAAAAAATTCACATATTTAATATAATATGTACTGTTTTACA
 ATATTTTGATATAACTTGATTATCGAAATTTTCAAACAGTACATAATTATAAATTATGTAAATCGATATTT
 ATAAAAAATATTTTATAGTTTATTGAACTATTATATTCTAAATATCGAAGACCATGTAACCAATGGTGTC
 AAGCACACCATTGTACTGGTAACTGTTTCTAAGGATAAGCTGAAGTCTGTTTCGTGCAGAATATAAGAAAGT
 CTTCCGCTTCTACCGTGGTTTTCAAGATAAGTGGGCTTGTTAATGCGTTAACGACCCCGCCCAAAGTTAAGG
 GTTTCGCTCAATCGTGTCAAACGATTGATTGTACCAGAGTCTTCTGTCTATCTGGATAACTACATTTTCAGA
 TTGACGGTACGCTTACTTCGTAGAGAGCATCAGTCCAAGTCTTGGGCTTGGATAATAGCTCGCTATTTTCAG
 AAAGGTCAAGCTTCCTAGAATAGACTCTCCATGGGTTCTAGGGGTGGGGGCGTCGAACTCGGATTTGACTC
 TATTACCACACAGATACTTTTCTAAGGTAGACTATATTCTGGTACGTAGCGATATGTTTAAAGATGATCCG
 TCGCTGTTGCTAAGGAGTATTGGGCTCGACGCCAACAACTTTTTTTTTTGACCGTAAAGCGGTGCGACTG
 CTCGCTGGTTTAGACAAAGACTTATCAGCATACAGGAGCAGTGTGCCGGTTTGTAATTCATCTTTATA
 CTCGGGTGAAAGGTTGCAGAGACATCAGTTATAGAAAGATCTTTGGGTGGAAGGGATACGTTTCGGAACC
 TACTTCTTGTAATTCTGTTTGTCTCGTGCAGTGCATTATGGCGGTGATGCGGATCTTTGGCGAGCTTAC
 TTATCGTCTTATAAAGTAACCTCATCCGAAGAAGTCAACACGTGGAAAGGTAAGTCCAGTCCAACAGGTTAA
 CCAAACGATTTCCGTCAGGCGGGTTCTGTTTCGGTTTTCTGGACCCACCTGAATTTAAATTTGGGTGGTTTA
 CAGGCGGTGCTGTTTCGTAATAAATAACGCGCTGTTCTAATTTATTAGTGCTTCGATAAGTACCGTAAACG
 TATTTTATACATTATTTATGCCTATAAGAAAACCTACCTTTTCTGCTCTCCTTTCTCTCTCTCTCTTTATTT
 TCTTCTCTTGTCACTGGAAGATCAGGGCCCAACGTGCCTTGGCCTTCCATTTGTGAATTTCCCGATTGGCAC
 CTGGGTGCGAATATAAGAAAGTCTTCGTGACTGCAGAGGCTGCATGCAAGCTTGGCGTAATCATGGTC
 ATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGT
 GTAAAGCCTGGGGTGCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTTCCAG
 TCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTATTGG
 GCGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTTCGGCTGCGGCGAGCGGTATCAGCTC
 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC
 CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGCGGTTTTTCCATAGGCTCCGCCCCCTGACGA
 GCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
 CCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTC
 CCTTCGGGAAGCTGTGTCACGAACCCCGTTACGCCGAGGTATCTCAGTTCCGTTAGGTGAGTCTGCTC
 CAAGCTGGGCTGTGTCACGAACCCCGTTACGCCGAGGTATCTCAGTTCCGTTAGGTGAGTCTGCTC
 AGTCCAACCCGTTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGCGCAGG
 TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGG
 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCA
 CCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGAT
 CCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATTTTGGTCATGAG

ATTATCAAAAAGGATCTTCACCTAGATCCTTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATAT
 ATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT
 CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCC
 CAGTGCTGCAATGATACCGCGAGACCCACGCTCACC GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG
 GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAA
 GCTAGAGTAAGTAGTTCCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
 ACGCTCGTCGTTTTGGTATGGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCA
 TGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGCGAGTGTTA
 TCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC
 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAA
 TACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
 AAATCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTC
 AGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAA
 TAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTTATCAGGGT
 TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
 TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTA
 TCACGAGGCCCTTTTCGTC

Secuencia del vector pGFP/PDS

TCGCGCGTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTG
 TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCT
 TAAGTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCG
 TAAGGAGAAAATACCGCATCAGGCGCCATTTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG
 CGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAACGCC
 AGGGTTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTCGATATCGGATAAAAAGACCAGGAACC
 GACTTAGCGCGCCGCCCTCACGCGACGGATCTTGATCCTAGCCCCCTAGGCGACGGACAAACGCCCCCGCCC
 CCCGTGAGGGGGCCCTTAGGGGGGCGCTTAATTTTTTTTAAAAATGTGGAATCAGGGCAACCTGATAGGACTG
 ATCCGTGTAGGAGGGGTACGAACCTTCTATCCTAGTCTAAGTGGCTGGCGAATCTAATTGGCCTCGGCTCAG
 AAAGGTCCACTAGAGAAAGCGAAAGGGAAGTGAATTTTATACGGACCCGCCTATAACCTGCCATGAATGG
 GTGTAGTTTTACCAAAGGACAAGTTTGGTCTCCCTTCTTGCCCTATGGATGTCATAAACACTCTACTAGG
 CGGGTCCGAAAGTGGTCAGAACGGAAGGCCTTTCAATTAATTTTCATTTAATTATGTCGTTAGTTTTAG
 CTAACAACCCTAACATCGAACTTCCGCAAGCATAGAACCAGGTTGACTGGAAAGCTGTGTTTTCGCGTCTGTT
 CAATCGTTAAGGGAACAAGCCTTTATCTGTTTGAAGTCTGGGGCGGGGGCGCCAGTGAAGTTGCGTGAGG
 TTCAAAAGCAAATTCTTGACCCGTGAGGGCACCTGCTACTTTTATTACTGGCTATTTCGTAGAGTTACGTCTA
 AGTCGGTCCGCGAAACGGCGGAAGTAGATGAGCTTATTGTTGTAAACGATTTCAGACAAATGGGCCATCTTT
 GATGGCTTCGTGGGTATGGATCTAGCTCATTGGCAACCTCCTCCTGTTTCGACGGCGGGTTCGCTATATTCT
 AAAGGTGACAAGCGACGTTGGAGACCGCTTGGTATTTCCACAGTTAAGGATCGTGTATCCAGGCGGTTG
 TGCCGGAGGCACCTGAACCTGAGTGGGACGTAAAGTTTGAAGGAAATTCCTATGTTTATAGACCAAAGCGT
 AACAGAGCTGATGCTCTTCAACGCACCTTCAATCATCTAGCGGATTCTCCTACGTACAACCAAACGTTAA
 TCGCGACCCGTGAGGGATATCTTAGCGGAGCCCTCCCGTGAGGGGGGCTAGGACTTTGTCCGTGCGCTAAG
 GGCTAGGATCAAGATCCGTGCGGTGAGGGAACAAAAAGTTGTTTTGTCCAGAAGTATTTTTTTCTTATA
 AAAAAAGAAAAAGAAAAAAAATTTATAAAATTATAAAGAATTTTATATTCTTTAGTTATCAAAATTATA
 ATTTGGTAAAGATTATTATATTATCGGATAGATTATTTTAGGATCGACAAAAATGTTTCATTTTTTAGATA
 ATAAAGCGGCCGCTCTAGACCGCGGATGCATCATCATCATCATATGGTTTCTAAAGGTGAAGAATTAT
 TTACAGGTGTTGTTTCTATTTTAGTTGAATTAGATGGTGATGTTAATGGTCATAAATTTTCTGTTTCTGGT
 GAAGGTGAAGGTGATGCTACATATGGTAAATTAACATTAATAATTTATTTGTACAACAGGTAAATTACCTGT
 TCCTTGGCCTACATTAGTTACAACATTAACATATGGTGTTCAATGTTTTTCTCGTTATCCTGATCATATGA
 AACAACATGATTTTTTTTTAAATCTGCTATGCCTGAAGGTTATGTTCAAGAACGTACAATTTTTTTTTAAAGAT
 GATGGTAATTATAAAACACGTGCTGAAGTTAAATTTGAAGGTGATACATTAGTTAATCGTATTGAATTAAG
 AGGTATTGATTTTTAAAGAAGATGGTAATATTTTAGGTGATAAATTAGAATATAATTATAATTCTCATAATG
 TTTATATTATGGCTGATAAACAAAAAATGGTATTAAAGTTAATTTTAAATTCGTCATAATATTGAAGAT

GGTTCTGTTCAATTAGCTGATCATTATCAACAAAATACACCTATTGGTGATGGTCCTGTTTTATTACCTGA
TAATCATTATTTTATCTACACAATCTGCTTTATCTAAAGATCCTAATGAAAAACGTGATCATATGGTTTTAT
TAGAATTTGTTACAGCTGCTGGTATTACTTTAGGTATGGATGAATTATATAAATAAGGATCCGAGCTCCTC
GAGTTCTTTTATACTTAAGAATTAAGTAGACTTGTCTACTTAATTCTTTTTAATAGCATGGATGGAGTAAAC
TCCATCCATGCTATTATCTCTAACAACCCCAAAAACCAAAGGTTTTTGTGAAGGAAAAAATAGAATATTTT
TTCCGCTACGAGCGGGCGATATAAAAAAATAAGATTTTTTTTATCCTAAATTAATTCTTGTAATAGAATCTC
AATTAATGACAAAGATATAAATTACAAAAACGGTAGAGTTTTAGGTTTTAGATAAATAACCTGGTCCTGCTA
TATAACTAATGTAAGTTAAATATCAGTATATGAACAATATATATTATATATTAATAGTCCATGTCTTTCA
AAAACGAATAAACGAATAAACCAAGAAAAATAAACTTTATTATTTTAAATAAAAATAAAATTTTATTAATT
TTTTTTTTTATCAAAAATTTTTTTTTTTCCGGAGCTAGCCACAAAGTCCTAGCCCTTAGGGGGCGGGGGCG
TTTGTCCGTCGCGTGAGAGAAATAACAAAGTTGTTTTGTCCGGTACACAATAAAAAATTATATAAAAGTTG
TTGTTTATAAAAAGGATGCTCTTGCTTTTGATAAATTATCTGGTATTATAATAATACCAGGTTTTAATGAA
AAAGTTGAATAATAAAAAAATAAAAAAGTTATTTTTAACTTTTTTCGGAGAAAAATCAAAAATTAATAAAA
AATTATAAATTAGCGCTGGTATGCAAGTTATGCAAGGTCGTGCTCATACACAAGCTGCATATTCACGCCAA
CCAGCTCAACAACGACTCAACGCCGTGTAGGTCGTACAACCTCGTTTACAAGTATATGCTCGTGATTTTCCA
GCTCCTCAATTCGATGGTACAGAACTTACCAAGAAGCTGTTGCATTATCAACAAAATTACAAAATGCTCC
ACGTCCTGTTAAACCACAACGTGTTGTAATTGCTGGTGCAGGATTAGCTGGTTTTATCTGCTGCCAAATATT
TATCAGATGCTGGTCATATTCCAGTTGTATTAGAAGCTCGTGATGTTCTTGGAGGTAAAGTAGCAGCTTGG
AAAGATGAAGACGGTGATTGGTATGAAACAGGTTTACATATCTTCGTTGGTGCTTATCCTAACATGCAACG
TTTATTTTAAAGAATTAATATCTCAGATCGTCTGCAGTGGAAGTCTCACTCAATGATTTTTGCTATGCAAG
ACAAACCAGGTGAATTTAGTCGTTTTGAATTTCTGATATTCCAGCTCCTTGGAATGGTGTTATTGTAATT
TTACGTAACAATGAAATGTTATCTTGGCCAGAAAAAATTCATTTGCAATTGGTTTTACTTCCTGCTATCAT
TTTCGGTCAAAAATATGTTGAAGAACAAGATGAATTAAGTGAACACAATGGATGCAAAAACAAGGTGTTTC
CATCTCGTGTAATCATGAAGTTTTTATTGCTATGGCTAAAGCTTTAAATTTTATTAACCCTGATGAATTA
TCAATGACTGTAGTTCTTACAGCTTTAAATCGTTTTTTACAAGAACGTCATGGTTCTAAAATGGCTTTTTTT
AGACGGTGCTCCACCTGAACGTTTATGTCAACCAATGGTAGATCACTTCACTTCAAAAAGGTGGAGAATTAC
GTATGAATGCACGTATTAAGCAAATAGTTTTAAATGAAGACAATTCAGTAAAACATTTTGAATTACTTAAT
GGTGAAATTGTTGAAGGTGATGTATACATGTCTGCTATGCCTGTTGATATTATGAAGAAATTAATGCCACA
ACCTTGGAAGGCATGCCATTCTTTCAAAAATTAATGGTTTAGAAGGTGTACCTGTTATTAATATACATA
TTTGGTTTGACCGTAAATTATCAACAGTAGACTCACTTACTTTTTCTCGTTAGATTTACTTTCTGTTTAT
GCTGATATGAGTACTACATGTAAAGAAATACACTGACGATAAAGCTAATATGTTAGAATTAGTATTTGCTCC
AGCTGCCGAATTGGATTGGTCGTCCTGATGAAGAGATTGTAGATGCTACAATGAAAGAACCTGAAAAATTAT
TTCTTAACGAAATTAAGCTGATCAATCATTAGCTAAAATTCGTAAGTCTAAGGTTATAAAAAACACCTTTA
TCAGTATGTAAATCTACCGCGGGTCGTGAAAAATATCGTCCATCACAAAAAACACCAATTAGTAATTTCTA
TCTTGCTGGTGATTACACTAAACAAAAATATTTAGCTTCTATGGAAGGTGCTGTATTTTCAGGTAAATTAG
CTTGTAACAAGTTGTAGAAGACGCTTCAACACGTACTGGTTCTCAAAGTGCAGCTTCTTCACAACCTGCT
TTAGCTGCCGCAAGTGCCGCTGTTTTACTTGCTATGGGTGCTGTATTAGCTGGTAACGCATCAGCACAAAGT
TTTAACAGAAGCTATTTGGGGTCTGAAGTATTAAGTGAAGCTGTAACATTTCTTGTTCTAACACGCAG
TGATTTAACTTAACTTAATAAAAAATTTTTTTTACCGCCCCCCCCCTCACGGGGAGGGAGGCTAGGACAAG
GTCCGTCGTAAAAAATAAATAAATTTTATTAATTTTATTAATTTAGTTACAAAATCTTCCATTTTC
ATCGGCCCCGCCAAAAGATAGTAAAAAATATACATATCTTTTAGGCGGAGCCTATCTCCGGTCTTTTATA
AAAAGAAAAACCGACATACAAAATACAAAATTGTATGGAGGCGGACAAAACAAATTTGTTGTTTTGTCCTA
AGGAAAAAATTCACATATTTAATATAATATGTACTGTTTTACAATATTTTGATATAAATTGATTATCGA
AATTTTCAAACAGTACATAATTATAAATTATGTAAATCGATATTTATAAAAAATATTTTATAGTTTATTGA
AACTATTATATTCTAAATATCGAAGACCATGTAACCAATGGTGTCAAGCACACCATTTGACTGGTAACTGT
TTCTAAGGATAAGCTGAAGTCTGTTTCGTGCAGAAATATAAGAAAGTCTTCCGCTTCTACCGTGGTTTCAAGA
TAAGTGGGCTTGTTAATGCGTTAACGACCCCGCCCAAAGTTAAGGGTTTCGCTCAATCGTGTCAAACGATT
GATTGTACCAGAGTCTTCTGTCATCTGGATAACTACATTTTCAGATTGACGGTACGCTTACTTCGTAGAGA
GCATCAGTCCAAGTCTTGGGCTTGGATAAAGTCTCGCTATTTTCAGAAAGGTCAAGCTTCTAGAAATAGACT
CTCCATGGGTTCTAGGGGTGGGGGCGTCGAAGCTCGGATTTGACTCTATTACCACACAGATACTTTTCTAAG
GTAGACTATATTCTGGTACGTAGCGATATGTTTAAAGATGATCCGTCGCCTGTTGCTAAGGAGTATTGGGC
TCGACGCCAACAACTTTTTTTTTTGACCGTAAAGCGGTGACCTGCTCGCTGGTTTAGACAAAGACTTATC
AGCATCACAGGAGCACGTGTGCCCGGTTTGTAAATTCATCTTTTATACTCCGGTGAAAGGTTGCAGAGACATC
ACGTTATAGAAAGATCTTTGGGTGGAAGGGATACGTTTCGGAAACCTACTTCTTGTACATTCTGTTTGTTCAT

CGTCGAGTGCATTATGGCGGTGATGCGGATCTTTGGCGAGCTTACTTATCGTCTTATAAAGTAACTCATCC
 GAAGAAGTCAACACGTGGAAAGGTAAGTCCAGTCCAACAGGTTAACCAAAACGATTTCCGTCAGGCGGGTT
 CGTTCCGTTTTCTGGACCCACCTGAATTTAAATTTGGGTGGTTTACAGGCGGTCTGTTCGTAACATAAATA
 ACGCGCTGTTCTAATTTATTAGTGCTTCGATAAGTACCGTAAACGTATTTTATACATTATTTATGCCTATA
 AGAAAACCTACCTTTTCTGTCTCCTTTCTCTTCGTCTCCTTTATTTTCTTCTCTTGTCACTGGAAGATCAGG
 GCCCAACGTGCCTTGGCCTTCCATTTGTGAATTCCCGATTGGCACCTGGGTGCAGAATATAAGAAAGTCTT
 CGTCGACTGCAGAGGCCTGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTTCTGTGTGAAATTGTT
 ATCCGCTCACAAATCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTG
 AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTTCCAGTCGGGAAACCTGTCTGTGCCAGCTGCA
 TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCCGCTTCTCTCGCTCACTG
 ACTCGCTGCGCTCGGTCTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC
 ACAGAACTCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA
 GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTTACAGGAGCATCACAAAAATCGACGCTCAAGTC
 AGAGTGGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTTTCTCA
 TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCC
 CCGTTCAGCCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
 TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
 GAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA
 CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGAC
 GCTCAGTGAACGAAAACCTACGTTAAGGGATTTTGGTTCATGAGATTATCAAAAAGGATCTTCACCTAGAT
 CCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC
 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCC
 GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTG
 CAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT
 AGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATT
 CAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCT
 TCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
 AATTCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG
 AGAATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAATACCGCGCCACATAGCA
 GAACTTTTAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTG
 AGATCCAGTTTCATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTACCAGCGTTTC
 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
 TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTT
 GAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTA
 AGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTTCGTC