

Universidad de Cantabria
Facultad de Medicina
Departamento de Biología Molecular



**“Ingeniería metabólica de
Synechococcus elongatus PCC 7942”**

Memoria presentada para optar al grado de Doctor por
María del Carmen Santos Merino

El presente trabajo ha sido realizado en el Departamento de Biología Molecular de la Universidad de Cantabria y en el Instituto de Biomedicina y Biotecnología de Cantabria, bajo la dirección del catedrático Fernando de la Cruz Calahorra, gracias a una beca predoctoral FPI (Formación de Personal Investigador) concedida por el Ministerio de Economía, Industria y Competitividad de referencia BES-2012-057387.

Durante este periodo se ha realizado una estancia de tres meses en el laboratorio del Prof. José Luis García (Centro de Investigaciones Biológicas, Madrid, España), gracias a una ayuda de movilidad predoctoral para la realización de estancias breves en centros de I+D financiada por el Ministerio de Economía, Industria y Competitividad.

D. Fernando de la Cruz Calahorra, Catedrático de Genética, perteneciente al departamento de Biología Molecular de la Universidad de Cantabria,

CERTIFICA:

Que Dña. María del Carmen Santos Merino ha realizado bajo su dirección el trabajo que lleva por título "Ingeniería metabólica de *Synechococcus elongatus* PCC 7942".

Considera que dicho trabajo se encuentra terminado y reúne los requisitos necesarios para su presentación como memoria de Doctorado, al objeto de poder optar al grado de Doctor por la Universidad de Cantabria

En Santander, a 12 de mayo de 2017.

Fdo. Fernando de la Cruz Calahorra

A mi familia

Agradecimientos

Después de haber escrito tantas páginas llenas de información científica, esta es para mí, sin duda, la parte más difícil de escribir. No soy una persona que se caracterice por expresar sus sentimientos, pero no podía terminar esta tesis sin agradecer a la gente que corresponde todo lo que ha hecho por mí.

En primer lugar, mi principal agradecimiento es para mi director, Fernando. Sin apenas conocerme me diste la oportunidad de trabajar en tu laboratorio, gracias por confiar en mí durante estos años. Espero no haberte defraudado y haber cumplido las expectativas. He aprendido mucho a lo largo de estos años.

En segundo lugar, agradecer a dos personas que me han ayudado mucho en el mundo “cianobacteril”. Por un lado, a Mapi, gracias por tener abiertas las puertas de tu despacho siempre que he necesitado tu ayuda. Por animarme, incluso ni cuando yo misma creía en mí. He aprendido mucho de ti y es algo que es de agradecer. Y por otro, a David, sin él esta tesis no podría haberse llevado a cabo. Gracias por haberme traspasado todos tus conocimientos sobre cianobacterias. Espero que tu legado continúe.

A mis compañeros de laboratorio, a los que aún siguen y a los que se fueron. A María, porque desde el principio confiaste en mí, aún más que yo misma. Por escucharme, por tus borderías (en algunos casos bien merecidas), por los fines de semana en tu compañía. En fin, gracias por todo, me llevo una amiga de verdad. A MdT, por tu ayuda en toda la parte bioinformática de esta tesis, por estar siempre dispuesta a colaborar, por las conversaciones en el 00.03. A Lillo, por los momentos compartidos tanto dentro como fuera del laboratorio, por estar siempre dispuesto a ayudar. A Atha, por traer esa locura griega al laboratorio. Se te echa de menos Sathanasia. A Ruli, por tu ayuda con el microscopio, por aportar siempre tu granito de sabiduría. A Irene, porque a pesar de estar poquito tiempo con nosotros, me ayudaste más de lo que puedes imaginarte. A las nuevas generaciones, Claudia y Jorge, por poner este toque de alegría al laboratorio. A Carolina, por preguntar siempre como lo llevo, por aguantar mis “chapas” como tú bien sabes. A los técnicos, Ana, Raquel y Carlos, gracias por vuestra ayuda a lo largo de todos estos años. A Alejandro, por tu ayuda dentro y fuera del laboratorio, por tus bromas y chistes que son capaces de alegrarte en los peores días. A los que se fueron pocos después de mi llegada: Val, Jorge, Esther, Inma.

A toda la gente de la planta 2, gracias vuestra ayuda en todos estos años. A Coral, por ayudarme desde el primer momento, por estar en las buenas y en las malas, por compartir tantas comidas y tantos planes fuera del laboratorio. Porque sé que me llevo una amiga. A Laura, por toda tu ayuda en la parte de ácidos grasos, por este reencuentro después de tantos años, por adentrarme en el mundo del swing. Gracias por toda tu ayuda. A Lolu, por llevar siempre la alegría andaluza contigo, por hacer las horas de comida más amenas. A toda la gente con la que he compartido planes durante estos años: Iago, Omar, Yolanda, Lucía, Paula, Patry...

A la gente de cocina, a Rodrigo, Alicia, Sheila y Natalia, por la preparación sin queja de todos los medios de cultivo para las cianos. Creedme que me habéis permitido realizar muchos más experimentos de los que habría hecho si me lo hubiera tenido que preparar yo misma.

A los técnicos del IBBTEC. A Rachel, por la alegría que trasmites siempre, por tus anécdotas, por hacer más amenas las horas de laboratorio. A Mati y Sandra, por nuestras conversaciones de pasillo y ascensor.

En general, gracias a toda la gente que forma parte del departamento de Biología Molecular y del IBBTEC. Ha sido un placer trabajar con vosotros. Gracias por tener siempre las puertas de los laboratorios abiertas.

A Choni, por acogerme en su laboratorio durante mi estancia en Alicante. Gracias por enseñarme tanto en tan poco tiempo. A Javier, porque a pesar del poco tiempo disponible siempre tenías un momento para resolver mis dudas. A Carina, por tu ayuda científica, por tu alegría en el laboratorio que hacían más amenos los días. A Raquel, Paloma, Aida, Jose, gracias a todos.

A José Luis, por aceptar mi estancia en tu laboratorio. A Juan Nogales, por enseñarme todo sobre los modelos metabólicos, por seguirme ayudándome desde la distancia. A Blas, por amenizarnos la mañana con tus listas de música en Spotify. A Lucía, por tu ayuda diaria en el ordenador, por tus conversaciones en la comida. A Irene, por ser tan buen alumna, espero haberte ayudado y enseñado todo lo que sabía. Gracias por hacerme sentir una científica de verdad. A toda la gente del CIB, por la buena acogida durante los tres meses en Madrid. Gracias a vosotros el tiempo se pasó volando.

A mis amigas. A Jenny, porque a pesar del tiempo y la distancia, nuestra amistad sigue igual que el primer día. Por demostrarme que puedo contar contigo tanto en mis victorias como en mis derrotas. A Patry, porque desde aquel primer día de Universidad nuestra relación solo ha ido a mejor. A Raquel, porque nuestra amistad empezó de la manera más tonta, pero ahí

seguimos aunque sea viviendo en países vecinos. A Ana, porque la Universidad nos unió y espero que eso siga así por muchos años. Gracias por las tardes compartidas durante mi estancia en Madrid. Gracias a todas por aguantar mi carácter, a veces, no me soporto ni yo. Os quiero, no cambiéis nunca.

Esta tesis tiene una dedicatoria especial, para una persona que lamentable ya no está aquí, mi madre. Casi al mismo tiempo empezamos una batalla, sin embargo, la tuya fue más dura que la mía y no pudiste salir victoriosa. Tu ausencia supuso un antes y un después en mi vida, tanto que estuve a punto de “mandar a la mierda” todo, incluido esta tesis. Sin embargo, me di cuenta que tú nunca lo habrías consentido, y es por ello que la cogí con más fuerzas como forma de agradecerte todo lo que hiciste por mí a lo largo de tu vida. Siempre tuviste una palabra de aliento, incluso cuando yo ya me había dado por vencida. Cada día que pasa te echo más de menos. Allí donde estés, espero que te sientas orgullosa de lo que he conseguido. Te quiero y nunca te olvido.

A mi padre, que es un luchador, ha sabido salir adelante a pesar de los reveses que le ha dado la vida. A mi hermana Laura, porque aunque somos tan diferentes, siempre nos tenemos la una a la otra. Porque además de mi hermana, eres mi amiga. Te quiero. A mis hermanos, Fran e Isaac, por ayudarme siempre, por creer en mí.

A mis tíos, Mari y Juan Luis, por acogerme durante mi estancia en Madrid, por hacerme sentir como una hija más. A mis primos, Rubén, Noemí, Alma, Alba, por hacerme más amenos los fines de semana en Madrid, por enseñarme esa gran variedad de juegos de mesa, un mundo desconocido para mí. A mis primines, Adrián y Lola, por hacerme volver a mi infancia, por darme tanto cariño.

A mi tía Carmina y a mis primos Begoña y Alberto, porque siempre me habéis animado y apoyado para conseguir todo lo que me proponía.

Maris

Contenido

| | |
|--|----|
| 1. Abreviaturas..... | 21 |
| 2. Introducción..... | 25 |
| 2.1. Cianobacterias..... | 25 |
| 2.1.1. Aspectos generales y taxonomía..... | 25 |
| 2.1.2. <i>Synechococcus elongatus</i> PCC 7942..... | 27 |
| 2.1.2.1. Descripción general..... | 27 |
| 2.1.2.2. Se7942 modelo para la producción de compuestos de interés biotecnológico 29 | |
| 2.1.2.2.1. Bioproducción de alcoholes en Se7942..... | 29 |
| 2.1.2.2.2. Bioproducción de azúcares en Se7942..... | 31 |
| 2.1.2.2.3. Bioproducción de ácidos orgánicos en Se7942..... | 32 |
| 2.1.2.2.4. Bioproducción de hidrógeno en Se7942..... | 33 |
| 2.1.2.2.5. Bioproducción de ácidos grasos en Se7942..... | 34 |
| 2.1.2.2.6. Otros compuestos de interés biotecnológico..... | 34 |
| 2.2. Evolución experimental..... | 36 |
| 2.2.1. Evolución experimental en microorganismos..... | 38 |
| 2.2.1.1. Evolución experimental en <i>Escherichia coli</i> | 39 |
| 2.2.1.2. Evolución experimental en otros microorganismos..... | 44 |
| 2.2.1.3. Evolución experimental en cianobacterias..... | 46 |
| 2.3. Síntesis de ácidos grasos..... | 47 |
| 2.3.1. Metabolismo de ácidos grasos en <i>Escherichia coli</i> | 48 |
| 2.3.1.1. Síntesis de ácidos grasos en <i>Escherichia coli</i> | 48 |
| 2.3.1.2. Degradación de ácidos grasos en <i>Escherichia coli</i> | 50 |
| 2.3.2. Metabolismo de ácidos grasos en cianobacterias..... | 52 |
| 2.3.2.1. Síntesis de ácidos grasos en cianobacterias..... | 52 |
| 2.3.2.2. Degradación de ácidos grasos en cianobacterias..... | 56 |
| 2.3.2.3. Modificaciones para mejorar la producción de ácidos grasos..... | 57 |
| 2.3.2.3.1. Modificaciones para mejorar la producción de ácidos grasos en Se7942. 57 | |
| 2.3.2.3.2. Modificaciones para mejorar la producción de ácidos grasos en Ss6803. 58 | |
| 2.3.2.3.3. Modificaciones para mejorar la producción de ácidos grasos en Ss7002. 59 | |
| 2.4. Modelos metabólicos..... | 60 |
| 2.4.1. Reconstrucción a escala genómica de un modelo metabólico..... | 60 |
| 2.4.2. Interés biotecnológico de los modelos metabólicos..... | 62 |
| 2.4.3. Modelos metabólicos de cianobacterias..... | 63 |

| | |
|--|-----|
| 3. Objetivos | 69 |
| 4. Materiales y Métodos | 73 |
| 4.1. Cepas bacterianas..... | 73 |
| 4.1.1. <i>Escherichia coli</i> | 73 |
| 4.1.2. Cianobacterias..... | 73 |
| 4.2. Plásmidos..... | 77 |
| 4.3. Oligonucleótidos..... | 82 |
| 4.4. Técnicas de biología molecular | 85 |
| 4.4.1. Extracción y purificación de ADN de <i>E. coli</i> | 85 |
| 4.4.2. Cuantificación del ADN..... | 86 |
| 4.4.3. Reacción en cadena de la polimerasa (PCR) | 86 |
| 4.4.4. Electroforesis de ADN en gel de agarosa | 87 |
| 4.4.5. Digestión de ADN mediante enzimas de restricción..... | 87 |
| 4.4.6. Fosforilación y desfosforilación..... | 88 |
| 4.4.7. Reacciones de ligación | 88 |
| 4.4.8. Métodos de clonaje..... | 89 |
| 4.4.8.1. Ensamblaje isotérmico (“Isothermal assembly”) | 89 |
| 4.4.8.2. “Biobricks” | 90 |
| 4.4.8.3. “QuickChange” modificado | 91 |
| 4.5. Aislamiento de ADN genómico total de Se7942 para su secuenciación | 92 |
| 4.6. Aislamiento de ARN de Se7942 para su análisis transcriptómico | 93 |
| 4.7. Secuenciación de ADN y ARN total de Se7942..... | 94 |
| 4.7.1. Secuenciación de ADN total | 94 |
| 4.7.2. Secuenciación de ARN total | 94 |
| 4.8. Técnicas microbiológicas..... | 95 |
| 4.8.1. Medios de cultivo | 95 |
| 4.8.2. Conjugaciones | 97 |
| 4.8.3. Transformación de ADN | 99 |
| 4.8.4. Introducción de mutaciones en el genoma de Se7942..... | 100 |
| 4.8.5. Cálculo del tiempo de generación de Se7942..... | 101 |
| 4.9. Técnicas analíticas usadas en Se7942 | 102 |
| 4.9.1. Cálculo del espectro de absorción de Se7942..... | 102 |
| 4.9.2. Determinación del contenido en pigmentos en Se7942..... | 103 |
| 4.9.3. Determinación del contenido en ácidos grasos de Se7942 | 103 |

| | |
|--|-----|
| 4.9.4. Medida de la actividad de los promotores en Se7942 usando luminiscencia | 104 |
| 4.9.5. Medida de la formación de biofilms en Se7942..... | 104 |
| 4.9.6. Análisis del pilus de Se7942 por microscopía electrónica..... | 105 |
| 4.10. Análisis bioinformáticos para el estudio de las poblaciones evolucionadas..... | 105 |
| 4.10.1. Identificación de mutaciones en poblaciones evolucionadas | 105 |
| 4.10.2. Modelado comparativo de la estructura tridimensional de proteínas | 106 |
| 4.10.3. Herramientas bioinformáticas para el alineamiento de secuencias | 106 |
| 4.10.4. Procesamiento de datos obtenidos en el RNAseq | 106 |
| 4.11. Reconstrucción del modelo metabólico de Se7942..... | 107 |
| 4.11.1. Herramientas bioinformáticas utilizadas para la reconstrucción del modelo..... | 107 |
| 4.11.2. Proceso de reconstrucción metabólica del modelo <i>iMS705</i> | 108 |
| 5. Resultados..... | 119 |
| 5.1. Evolución experimental de Se7942 | 119 |
| 5.1.1. Experimento de evolución en Se7942..... | 119 |
| 5.1.2. Descripción de los genes que presentan mutaciones en las poblaciones evolucionadas..... | 126 |
| 5.1.2.1. Mutación #2: Polipéptido de pilina similar a PilA..... | 126 |
| 5.1.2.2. Mutación #7: Proteína hipotética conservada | 127 |
| 5.1.2.3. Mutación #8: ATP sintasa F1, subunidad alfa..... | 129 |
| 5.1.2.4. Mutación #9: Proteína hipotética conservada | 130 |
| 5.1.2.5. Mutación #12: Promotor de la shikimate kinasa | 132 |
| 5.1.2.6. Mutaciones #13 y #14: Ácido graso de cadena larga-CoA ligasa | 133 |
| 5.1.2.7. Mutación #16: Región intergénica entre dos genes que codifican para un polipéptido de unión al ficobilisoma..... | 135 |
| 5.1.2.8. Mutación #19: Proteína de ensamblaje a fimbrias, similar a PilC..... | 136 |
| 5.1.2.9. Mutación #20: Proteína de motilidad "Twitching" | 138 |
| 5.1.2.10. Mutación #21: Histidina kinasa..... | 139 |
| 5.1.2.11. Mutación #24: Región intergénica entre un gen que codifica para una proteína hipotética conservada y un gen que codifica para la ARNt i(6)A37 tiotransferasa MiaB..... | 141 |
| 5.1.3. Comprobación de la presencia de las mutaciones identificadas en la cepa C11-G ₁₂₄₈ | 141 |
| 5.1.4. Clonación de las mutaciones en la cepa silvestre de Se7942 | 144 |
| 5.1.5. Efecto de las mutaciones en el tiempo de generación | 151 |
| 5.1.6. Efecto de las mutaciones en la expresión génica..... | 156 |

| | |
|--|-----|
| 5.1.7. Efecto de las mutaciones en otras características | 171 |
| 5.2. Síntesis de ácido α -linolénico en Se7942 | 176 |
| 5.2.1. Condiciones óptimas de inducción de los promotores utilizados para clonar las desaturasas productoras de ácido α -linolénico | 176 |
| 5.2.2. Producción de ácido α -linolénico mediante la sobreexpresión de desaturasas de Ss7002 en Se7942 | 177 |
| 5.2.3. Efecto de la modificación de la expresión de los genes de síntesis y degradación de ácidos grasos en Se7942 | 182 |
| 5.2.3.1. Efecto de la sobreexpresión del gen <i>fabD</i> en Se7942 en la producción de ácidos grasos..... | 183 |
| 5.2.3.2. Efecto de la sobreexpresión y delección del gen <i>fabH</i> en Se7942..... | 183 |
| 5.2.3.3. Efecto de la sobreexpresión del gen <i>fabB</i> en el contenido de ácidos grasos de Se7942..... | 188 |
| 5.2.3.4. Efecto de la sobreexpresión del gen <i>fabF</i> en el contenido de ácidos grasos de Se7942..... | 191 |
| 5.2.3.5. Efecto de la delección del gen <i>fadD</i> en el contenido de ácidos grasos de Se7942..... | 193 |
| 5.2.4. Análisis de dobles mutantes de Se7942 productores de ácido α -linolénico | 195 |
| 5.2.5. Análisis de un triple mutante de Se7942 productor de ácido α -linolénico | 199 |
| 5.3. Modelo metabólico de Se7942 | 203 |
| 5.3.1. Construcción y características del modelo metabólico <i>iMS705</i> de Se7942..... | 203 |
| 5.3.2. Comparación de <i>iMS705</i> con otros modelos metabólicos de Se7942..... | 206 |
| 5.3.3. Comprobación <i>in silico</i> del efecto de las modificaciones realizadas en Se7942 para la producción de ácido α -linolénico..... | 212 |
| 5.3.3.1. Modificación <i>in silico</i> de la expresión del gen <i>fabH</i> | 213 |
| 5.3.3.2. Sobreexpresión <i>in silico</i> de <i>fabF</i> | 215 |
| 5.3.3.3. Delección <i>in silico</i> de <i>fadD</i> | 220 |
| 5.3.4. Identificación <i>in silico</i> de nuevos mutantes “knockouts” para mejorar la producción de ácido α -linolénico | 222 |
| 6. Discusión | 228 |
| 6.1. Evolución experimental de Se7942 | 228 |
| 6.1.1. Estudio de las cepas obtenidas mediante evolución experimental..... | 229 |
| 6.1.2. Análisis del tiempo de generación de las cepas evolucionadas y los mutantes construidos..... | 230 |
| 6.1.3. Estudio de la transformabilidad natural en las cepas evolucionadas | 232 |
| 6.1.4. Efecto de las mutaciones sobre la expresión génica..... | 236 |
| 6.2. Síntesis de ácido α -linolénico en Se7942 | 239 |

| | |
|--|-----|
| 6.2.1. Ventajas y desventajas del uso de las cianobacterias para la producción de ácidos grasos..... | 239 |
| 6.2.2. Selección del organismo donador de los genes para la producción de ácido α -linolénico | 239 |
| 6.2.3. Selección del promotor adecuado para la expresión de los genes de producción de ácido α -linolénico | 240 |
| 6.2.4. Estudio del efecto de diferentes condiciones en la producción de ácido α -linolénico..... | 242 |
| 6.2.5. Efecto de la modificación de la expresión de los genes de síntesis y degradación de ácidos grasos en Se7942 | 243 |
| 6.3. Modelo metabólico de Se7942 | 248 |
| 6.3.1. Reconstrucción del modelo metabólico de Se7942..... | 248 |
| 6.3.2. Capacidad predictiva de genes esenciales del modelo metabólico de Se7942 | 249 |
| 6.3.3. Análisis <i>in silico</i> de las modificaciones probadas experimentalmente en Se7942 para mejorar la producción de ácido α -linolénico | 250 |
| 6.3.4. Identificación <i>in silico</i> de nuevos “knockouts” para mejorar la producción de ácido α -linolénico | 251 |
| 7. Conclusiones | 257 |
| 8. Bibliografía | 261 |

Abreviaturas

1. Abreviaturas

| | |
|-------------------|---|
| Abs | Absorbancia |
| ACP | Proteína transportadora de grupos acilo |
| ADN | Ácido desoxirribonucleico |
| ALA | Ácido α -linolénico |
| APC | Aloficocianina |
| ARN | Ácido ribonucleico |
| Ap | Ampicilina |
| ATP | Adenosín trifosfato |
| cAMP | Adenosín monofosfato cíclico |
| Cm | Cloranfenicol |
| Chl | Clorofila |
| Chla | Clorofila a |
| CoA | Coenzima A |
| CRP | Proteína de unión a cAMP |
| Ct | Carotenoides |
| C-terminal | Carboxi-terminal |
| dATP | Desoxiadenosina trifosfato |
| DAP | Ácido diaminopimélico |
| DE | Desviación estándar |
| DO | Densidad óptica |
| dCTP | Desoxicitidina trifosfato |
| dGTP | Desoxiguanosina trifosfato |
| DTT | Ditiotreitol |
| dTTP | Desoxitimidina trifosfato |
| FB | Ficobiliproteínas |
| Gm | Gentamicina |
| Indels | Inserciones y deleciones |
| IPTG | Isopropil- β -D-tiogalactopiranosido |
| kb | Kilobase |
| Km | Kanamicina |
| LA | Ácido linoleico |
| LB | Medio de cultivo Luria-Bertani |
| MCS | Sitio múltiple de clonaje |
| NAD | Nicotinamida adenina dinucleótido |
| Neo | Neomicina |
| ORF | Marco de lectura abierto |
| pb | Pares de bases |
| PC | Ficocianina |
| PCR | Reacción en cadena de la polimerasa |
| PEG | Polietilenglicol |
| p/v | Relación peso/volumen |
| RBS | Sitio de unión al ribosoma |
| RPKM | Reads por kilobase por millón de reads mapeados |
| rpm | Revoluciones por minuto |
| Se7942 | <i>Synechococcus elongatus</i> PCC 7942 |
| Se6301 | <i>Synechococcus elongatus</i> PCC 6301 |
| Ss6803 | <i>Synechocystis</i> sp. PCC 6803 |
| Ss7002 | <i>Synechococcus</i> sp. PCC 7002 |

1. Abreviaturas

| | |
|-------------|--------------------------------------|
| Sm | Estreptomicina |
| SN1 | Sitio neutral 1 |
| SN2 | Sitio neutral 2 |
| SNPs | Polimorfismos de un solo nucleótido |
| T2SS | Sistema de secreción tipo II |
| T4P | Pilus tipo IV |
| Tc | Tetraciclina |
| Tp | Trimetoprima |
| URL | Unidades Relativas de Luminiscencia |
| v/v | Relación volumen/volumenIntroducción |

Introducción

2. Introducción

2.1. Cianobacterias

2.1.1. Aspectos generales y taxonomía

Las cianobacterias son un grupo de bacterias fototróficas, Gram negativas, conocidas comúnmente como algas verde-azuladas debido a los pigmentos fotosintéticos que poseen, clorofila a y ficobiliproteínas (Stanier and Cohen-Bazire, 1977). Pertenecen al Reino de las eubacterias y al *phylum* Cianobacteria (Rippka et al., 1979).

Las cianobacterias constituyen un grupo morfológicamente muy diverso, al igual que sus contenidos en G+C (30 - 62 %) y los tamaños de sus genomas (1,5 - 8 Mb) (Beck et al., 2012). Atendiendo a la envuelta celular, se les considera bacterias Gram negativas debido a la presencia de una membrana externa y una capa de peptidoglicano, aunque su pared celular posea también características de las paredes típicamente Gram positivas (Hoiczky and Hansel, 2000).

La clasificación más aceptada de las cianobacterias las divide en 5 secciones atendiendo a criterios morfológicos y a cómo se realiza la división celular (Rippka et al., 1979). La Sección I incluye estirpes unicelulares que se dividen por fisión binaria o por gemación, perteneciendo a esta división, entre otros, los géneros *Synechococcus* y *Synechocystis*. La Sección II está constituida por estirpes unicelulares que se reproducen por fisión múltiple y/o fisión binaria y a la que pertenecen los géneros *Dermocarpa* y *Pleurocapsa*, entre otros. En la Sección III aparecen estirpes que, como *Pseudanabaena* y *Spirulina*, son filamentosas no formadoras de heterocistos. Las Secciones IV y V la conforman especies filamentosas formadoras de heterocistos, representando la Sección IV especies que presentan división en un plano (como los géneros *Anabaena* y *Nostoc*), mientras que en la Sección V se clasifican especies con división en más de un plano (por ejemplo, el género *Fischerella*). Sin embargo, hoy día la clasificación de las cianobacterias está en revisión, debido al empleo de marcadores quimiotaxonómicos y el análisis comparativo de macromoléculas. De este modo, se ha demostrado que únicamente la Sección II, y las Secciones IV y V conjuntamente, forman grupos filogenéticos verdaderos (Giovannoni et al., 1988).

Se considera a las cianobacterias como principales responsables de la acumulación de oxígeno atmosférico a lo largo de la evolución de la Tierra (Olson, 2006). Se estima que su origen

es muy primitivo (aproximadamente hace unos 2.500 millones de años), desempeñando un papel crucial en la evolución del ambiente geoquímico y de la vida en la Tierra. Por un lado, fueron las promotoras de la transformación de la atmósfera desde un estado reductor inicial hasta su actual estado oxidante (por su producción de O₂), y por otro, fueron las causantes de la aparición de la capa de ozono protectora de la radiación ultravioleta (Buick, 1992; Schopf, 1993). Según la teoría endosimbiótica, las cianobacterias fueron precursores de los cloroplastos en las células vegetales (Gray, 1992). Los estudios filogenéticos realizados han encontrado pruebas a favor de la endosimbiosis y de la posible existencia de una cianobacteria ancestral común a las cianobacterias actuales y a los cloroplastos de plantas (Giovannoni et al., 1988).

En la actualidad, las cianobacterias pueden hallarse en hábitats muy diversos, estando presentes en la práctica totalidad de ecosistemas estudiados: aguas oceánicas y continentales, suelos y ambientes extremos, aguas termales y desiertos (Tandeau de Marsac and Houmard, 1993). Posibles explicaciones de su ubicuidad son los escasos requerimientos nutricionales que poseen, así como su capacidad para vivir de forma simbiótica con plantas y hongos, lo que les permite estar presentes en otros nichos que por sí solos no podrían colonizar (Waterbury, 2006). Cabe destacar la existencia de especies de cianobacterias extremófilas, capaces de vivir en ambientes extremos de temperatura (en aguas termales o en la Antártida), de salinidad (tanto en agua dulce como salada) o de pH (en aguas ácidas o alcalinas) (Waterbury, 2006).

Desde el punto de vista metabólico, las cianobacterias viven en condiciones fotoautotróficas, aunque algunas pueden crecer además en condiciones heterotróficas en presencia de alguna fuente de carbono. Debido a su ubicuidad y a su capacidad de crecimiento fotoautotrófico, las cianobacterias son los principales organismos responsables de la fotosíntesis y la fijación de CO₂ en medios marinos (Herrero and Flores, 2008). También son considerados importantes componentes de los ecosistemas globales, ya que son responsables de entre el 20 - 30 % de la producción total de oxígeno (Pisciotta et al., 2010). Por todo ello, la importancia de las cianobacterias en el ciclo de carbono es fundamental. Por otro lado, cabe destacar su papel en el ciclo del nitrógeno, ya que algunos géneros de cianobacterias son capaces de fijar nitrógeno atmosférico en los heterocistos (Herrero et al., 2001). Además, son los organismos más simples que presentan ritmo circadiano y cuyo mecanismo molecular ha sido ampliamente estudiado (Ditty et al., 2003; Golden, 2003; Iwasaki and Kondo, 2004; Johnson, 2004; Johnson et al., 2008; Williams, 2007). La evolución del ritmo circadiano podría derivar de la necesidad de realizar procesos celulares que son vulnerables a la luz (Johnson, 2010).

En las últimas décadas se han desarrollado diversas herramientas moleculares para la investigación genética y modificación de cianobacterias: ADN exógeno puede ser transferido a cepas unicelulares y filamentosas mediante transformación natural (Grigorieva and Shestakov, 1982; Shestakov and Khyen, 1970; Stevens and Porter, 1980), electroporación (Thiel and Poo, 1989) y conjugación con vectores autónomos o integrativos (Marraccini et al., 1993; Wolk et al., 1984). Hay disponibles diversos marcadores de selección negativa y positiva y construcciones reporteras para medir actividad de promotores o para estudios de localización celular (Berla et al., 2013). Se conocen, además, varios sistemas de restricción-modificación presentes en cianobacterias (Stucken et al., 2013; Zhao et al., 2006). Actualmente, se han secuenciado y anotado al menos 39 genomas de cianobacterias unicelulares y filamentosas (Cyanobase, <http://genome.kazusa.or.jp/cyanobase>), lo cual facilita llevar a cabo modificaciones genéticas y análisis transcriptómicos y proteómicos.

Además de todo lo expuesto, que denota la gran importancia de este *phylum*, las cianobacterias han sido identificadas como nueva fuente de compuestos bioactivos (Abed et al., 2009). Destaca especialmente la producción de compuestos naturales con actividad antibacteriana, antifúngica, antiviral, antiprotozoal y molusquicida (Niedermeyer, 2015). Por otro lado, ciertas cianobacterias son capaces de sintetizar y acumular bioplásticos degradables (Campbell et al., 1982; Sudesh et al., 2002). Estos organismos son capaces de llevar a cabo la degradación de compuestos tóxicos como fenantreno y naftaleno (Kumar et al., 2009), así como la descontaminación acuática de nitritos, de fosfatos y de metales pesados (El-Bestawy, 2008; Fawzy and Issa, 2016; Roy et al., 2015).

En este trabajo se ha utilizado como modelo de estudio la cianobacteria *Synechococcus elongatus* PCC 7942, cuyas características principales se describen en el siguiente apartado.

2.1.2. *Synechococcus elongatus* PCC 7942

2.1.2.1. Descripción general

Synechococcus elongatus PCC 7942 (a partir de ahora, Se7942), anteriormente denominada *Anacystis nidulans* R2, fue descrita por Nägeli (Nägeli, 1849). Se caracteriza por ser fotoautótrofo obligado, de agua dulce, unicelular, se divide por fisión binaria, y pertenece a la sección I de la clasificación morfológica (Rippka et al., 1979). Las células de Se7942 tienen forma

de bacilo y su tamaño oscila entre 1 - 2 μm (Yu et al., 2015). Estas células forman colonias redondas en placas de agar y crecen homogéneamente en medio líquido (Bazin, 1971).

El genoma de Se7942 está formado por un cromosoma circular de 2,7 Mb, un plásmido grande (pANL) de aproximadamente 46,3 Kb y un plásmido pequeño (pANS) de 8,6 Kb (Chen et al., 2008). Esto convierte a Se7942 en uno de los organismos fotosintéticos de vida libre con un tamaño cromosómico más reducido. El contenido en G+C del genoma de Se7942 está alrededor del 55,43 % (Beck et al., 2012). Además, su genoma está totalmente secuenciado desde el año 2004 (número de acceso en GenBank NC_007604).

Se7942 es similar a *Synechococcus elongatus* PCC 6301 (a partir de ahora Se6301). Se6301 tiene un genoma 322 pb mayor al de Se7942 y presenta una inversión cromosómica de 188.6 kb, responsable de que Se6301 no sea transformable naturalmente (Sugita et al., 2007). Además, existen dos regiones del genoma de Se6301 que no están presentes en el genoma de Se7942. La primera de ellas incluye una secuencia de 52 pb (posición nucleotídica de 704.012 a 704.063) y la segunda una secuencia de 243 pb (de 1.783.884 a 1.784.126) en el genoma de Se6301. A pesar de estas diferencias, algunos autores consideran que Se6301 y Se7942 pertenecen a la misma especie (Golden et al., 1989; Wilmotte and Stam, 1984).

Las principales características que hacen que esta cianobacteria haya sido elegida como modelo de estudio, son su relativamente corto tiempo de generación (5 h) (Ihlenfeldt and Gibson, 1975; Yu et al., 2015) y su competencia natural para ser transformada con ADN exógeno. Fue la primera cianobacteria en la que se demostró la existencia de un sistema de transformación natural (Shestakov and Khyen, 1970) y las condiciones para conseguir la máxima eficacia han sido bien definidas (Golden and Sherman, 1984). Además de la transformación natural, existen diversas herramientas de manipulación genética disponibles para Se7942. Un ejemplo de ello son los plásmidos lanzadera capaces de replicar en *E. coli* y en Se7942 (Golden and Sherman, 1983), así como vectores lanzadera suicidas, capaces de integrarse en el genoma de Se7942, por recombinación homóloga (Tsinoremas et al., 1994). A lo largo de los años, Se7942 ha servido como modelo de estudio en muchos campos, destacando la adquisición de carbono inorgánico (Bonfil et al., 1998; Gabay et al., 1998; Maeda et al., 2000; Omata et al., 2001; Tchernov et al., 2001), el transporte y regulación de compuestos nitrogenados (Lee et al., 1999), la respuesta a la carencia de hierro (Michel et al., 2001; Webb et al., 1994), la adaptación a diferentes estreses de nutrientes (Collier and Grossman, 1992; Sauer et al., 1999) y a variaciones ambientales tanto de temperatura como de intensidad de luz (Schelin et al., 2002; van Waasbergen et al., 2002). Por último, es importante mencionar que Se7942 es el modelo

procariótico más utilizado para el estudio del ritmo circadiano (Golden, 2003; Golden et al., 1998).

2.1.2.2. Se7942 modelo para la producción de compuestos de interés biotecnológico

En la actualidad, muchos procesos industriales dependen de fermentaciones de bacterias heterotróficas, que producen productos como vitaminas, enzimas y aminoácidos. La viabilidad económica de estos esquemas de producción está limitada por el coste de las fuentes de carbono y energía utilizadas en los procesos fermentativos (Lau et al., 2015). Las cianobacterias, dotadas de un sistema fotosintético para fijar dióxido de carbono a una forma reducida, constituyen la maquinaria biosintética ideal para la producción sostenible de productos químicos y biocombustibles (Lai and Lan, 2015). A diferencia de las bacterias heterotróficas, las cianobacterias solo necesitan luz solar, dióxido de carbono, agua y una mínima cantidad de nutrientes para crecer, eliminando los costes derivados de las fuentes de carbono y energía y los medios de crecimiento complejos (Ducat et al., 2011b). La luz solar es el recurso más fácilmente disponible y barato de la tierra, siendo las cianobacterias capaces de transformar entre el 3 y el 9 % de la energía solar en biomasa (Dismukes et al., 2008).

Considerando lo anteriormente descrito, junto a los recientes avances en tecnologías de manipulación genética y de ingeniería metabólica, las cianobacterias son uno de los candidatos más atractivos para su uso en aplicaciones biotecnológicas.

A continuación, se describen los principales compuestos que se han conseguido producir en la cianobacteria modelo de estudio de esta tesis Se7942.

2.1.2.2.1. Bioproducción de alcoholes en Se7942

Algunas cianobacterias producen pequeñas cantidades de etanol heterotróficamente durante el metabolismo anaerobio nocturno (Heyer and Krumbein, 1991; Stal and Moezelaar, 1997). Sin embargo, es necesario realizar ingeniería genética para incrementar la productividad y activar la producción bajo condiciones fotosintéticas (Lai and Lan, 2015).

El primer informe de producción mediante la expresión heteróloga de la ruta de síntesis de etanol fue llevado a cabo por (Deng and Coleman, 1999). Estos autores introdujeron los genes

2. Introducción

que codifican para la piruvato descarboxilasa (*pdc*) y la alcohol deshidrogenasa II (*adhII*) de *Zymomonas mobilis* en Se7942. Para ello, clonaron estos genes bajo el control del promotor del operón *rbcLS* propio de Se7942 en un vector derivado del plásmido pANS, capaz de replicar de forma autónoma en Se7942. Con ello consiguieron una producción de 0,02 mg/L de etanol tras 7 días de cultivo. En este caso, el sustrato para la producción del etanol era el piruvato (**Figura 2-1, A**).

Alternativamente, el etanol puede producirse desde el acetyl-CoA vía acetaldehído bajo condiciones aeróbicas. Para ello, (Lan et al., 2013) clonaron el gen *pduP* que codifica para una deshidrogenasa aldehído CoA-acilante de la ruta de degradación del 1,2-propanediol de *Salmonella enterica* y el gen *yqhD* que codifica para una alcohol deshidrogenasa dependiente de NADPH de *E. coli*. Un esquema de la ruta se representa en la **Figura 2-1, B**. La cepa resultante fue capaz de producir 182 mg/L de etanol tras 10 días de cultivo. A pesar de que la producción de etanol mediante esta ruta es más alta, su principal desventaja es la necesidad de dos factores reductores. La cantidad de etanol producido, por tanto, estaría sujeta a la disponibilidad en la célula de dichos factores reductores.

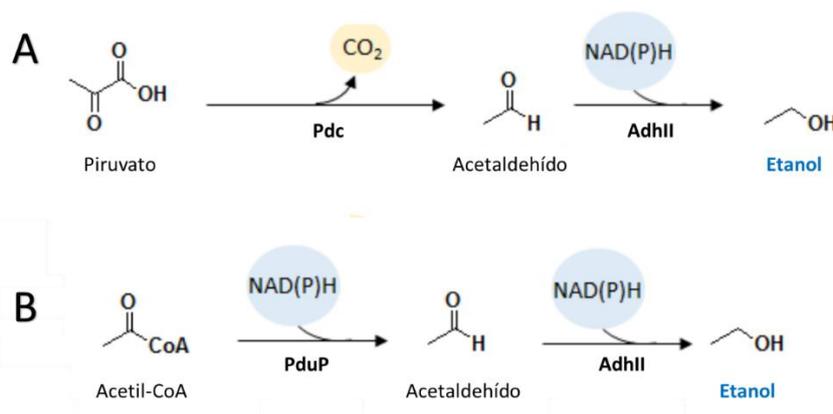


Figura 2-1. Representación esquemática de la producción de etanol. (A) Ruta de síntesis de etanol utilizando como sustrato el piruvato. **(B)** Ruta de síntesis de etanol a partir del acetyl-CoA. NAD(P)H, factor reductor. Modificado de (Lai and Lan, 2015).

Además de etanol, otros alcoholes de cadena corta y de cadena larga han conseguido producirse en nuestra cianobacteria modelo Se7942. Un resumen de estos compuestos producidos, así como de las características de las cepas productoras, se muestra en la **Tabla 2-1**.

Tabla 2-1. Alcoholes producidos en Se7942, así como los promotores y genes usados, la producción obtenida, el metabolito central del que se parte y la referencia bibliográfica. Modificado de (Lai and Lan, 2015).

| Alcohol producido | Promotor/es usado/s | Gen/es expresados | Producción (mg/L) | Días de cultivo | Metabolito central | Referencia |
|-------------------|--|--|-------------------|-----------------|-------------------------|-------------------------|
| Isopropanol | <i>P_{lacO1}</i> | <i>thl-atoAD- adc-adh-pta</i> | 33,1 | 14 | Acetato | (Hirokawa et al., 2016) |
| Isopropanol | <i>P_{trc}</i> | <i>thl-atoAD- adc-adh</i> | 146 | 15 | Acetil-CoA | (Hirokawa et al., 2015) |
| Isopropanol | <i>P_{trc}</i> | <i>thl-atoAD- adc-adh</i> | 26,5 | 9 | Acetil-CoA | (Kusakabe et al., 2013) |
| 1-Butanol | <i>P_{trc}/P_{lacO1}</i> | <i>ter/nphT7, pduP, yqhD, ctr, hbd</i> | 317 | 12 | Acetil-CoA | (Lan et al., 2013) |
| 1-Butanol | <i>P_{trc}/P_{lacO1}</i> | <i>ter/nphT7, pduP, yqhD, ctr, hbd</i> | 30 | 17 | Acetil-CoA | (Lan and Liao, 2012) |
| 1-Butanol | <i>P_{trc}/P_{lacO1}</i> | <i>ter/atoB, adhE2, crt, hbd</i> | 14,5 | 7 | Acetil-CoA | (Lan and Liao, 2011) |
| Isobutanol | <i>P_{trc}</i> | <i>kvid, YqhD</i> | 450 | 6 | Piruvato | (Atsumi et al., 2009) |
| 2-Metilbutanol | <i>P_{trc}</i> | <i>kivD, yqhD, cimA, leuBCD</i> | 178 | 12 | Piruvato/ Acetil-CoA | (Shen and Liao, 2012) |
| 1,2-Propanediol | <i>P_{trc}</i> | <i>sADH, yqhD, mgsA</i> | 150 | 10 | Piruvato | (Li and Liao, 2013) |
| 2,3-Butanediol | <i>P_{lacO1}</i> | <i>alsS, alsD, adh</i> | 2.380 | 20 | Piruvato | (Oliver et al., 2013) |
| Glicerol | <i>P_{trc}</i> | <i>gpp1</i> | 1.170 | 20 | DHAP* | (Wang et al., 2015) |

*DHAP: dihidroxiacetona fosfato

2.1.2.2.2. Bioproducción de azúcares en Se7942

Los azúcares tienen un interés particular, ya que solo pueden ser producidos fotosintéticamente, no utilizando otras estrategias metabólicas (Ducat et al., 2011b). Las proteínas involucradas en la síntesis de sacarosa en cianobacterias han sido descritas recientemente (Kolman et al., 2015). La síntesis de sacarosa involucra la acción secuencial de una sacarosa-fosfato sintasa (SPS) y la sacarosa-fosfato fosfatasa (SPP) (Martinez-Noel et al., 2013). Básicamente, estas enzimas catalizan la conversión de UDP-glucosa y fructosa-6-fosfato en sacarosa. La hidrólisis de sacarosa en fructosa y glucosa es catalizada por una invertasa (INV) (McEwen et al., 2013).

Se7942 produce naturalmente sacarosa, así como fructosa y glucosa, en respuesta a estrés osmótico (Kerby et al., 1990). De esta manera, Se7942 acumula sacarosa en el citoplasma

2. Introducción

hasta 280 mM para compensar la presión osmótica externa (Suzuki et al., 2010). Sin embargo, no es capaz de secretar ninguno de los azúcares que produce. Es por ello que esta cepa ha sido modificada genéticamente para mejorar la producción de glucosa y fructosa, así como para que sea capaz de transportar estos azúcares al exterior de la célula. En la **Figura 2-2** se muestra un esquema de las principales modificaciones realizadas en Se7942 para la producción de azúcares. Niederholtmeyer *et al.* modificaron genéticamente a Se7942 con los genes *invA* (invertasa) y *glf* (transportador de fructosa) de *Zymomonas mobilis*. Tras 5 días de cultivo consiguieron una producción de 45 mg/L de una mezcla de fructosa y glucosa (Niederholtmeyer et al., 2010). Más recientemente, (Ducat et al., 2012) clonaron en Se7942 el transportador de sacarosa, *cscB*, de *E. coli*, y deleccionaron además los genes *invA* y *glgC* en esta cepa. Estos genes codifican para la invertasa y la glucógeno sintasa, respectivamente. De esta manera consiguieron una producción de 3 g/L de sacarosa tras 7 días de cultivo. Recientemente, (Duan et al., 2016) sobreexpresaron conjuntamente en Se7942 el gen *cscB* de *E. coli* y el gen *sps* propio de Se7942, consiguiendo una producción de 1.5 g/L de sacarosa tras 7 días de cultivo.

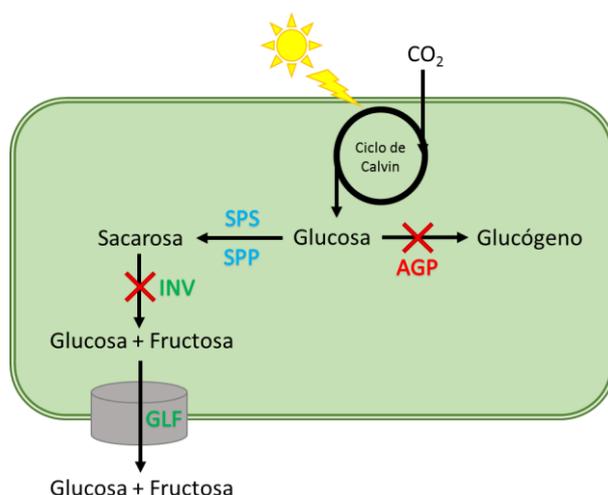


Figura 2-2. Esquema de las modificaciones realizadas en Se7942 para la producción de azúcares. La glucosa es obtenida como producto en el ciclo de Calvin, el cual es realizado gracias a la presencia de CO₂ y luz solar. Esta glucosa es transformada por SPS y SPP en sacarosa. La sacarosa es hidrolizada por INV en glucosa y fructosa, posteriormente transportados por GLF al exterior de la célula. La producción de glucosa puede mejorarse con la delección de AGP que impide el uso de la sacarosa para la producción de glucógeno. En rojo, se representan aquellas dianas que han sido objeto de eliminación; en azul, aquellos genes involucrados en la síntesis de azúcares y cuya expresión no se ha modificado; en verde, aquellas dianas que han sido objeto de sobreexpresión; con una cruz roja, aquellas dianas que han sido deleccionadas. AGP, ADP glucosa pirofosforilasa (gen *glgC*); SPS, sacarosa-fosfato sintasa; SPP, sacarosa-fosfato fosfatasa; INV, invertasa (gen *invA*); GLF, transportador facilitador de glucosa (genes *glf* o *cscB*).

2.1.2.2.3. Bioproducción de ácidos orgánicos en Se7942

El principal ácido orgánico que se ha producido en Se7942 es el D-lactato. Este ácido es una forma isomérica del lactato y es usado como materia prima para la síntesis de bioplásticos

que tienen infinidad de aplicaciones comerciales (Li et al., 2013). También es utilizado para la producción de cosméticos, perfumes, revestimientos, adhesivos y tinta para impresión (Abdel-Rahman et al., 2011; Zhang et al., 2010).

En la síntesis de lactato, el piruvato procedente del ciclo de Calvin es transformado a lactato por la acción de la lactato deshidrogenasa (LDH) y éste es secretado al exterior de la célula por acción de transportadores de lactato (LLD) (**Figura 2-3**). Cabe destacar que Se7942 no posee ninguna de estas enzimas, por tanto, la producción de D-lactato requiere la expresión heteróloga de estas proteínas. (Niederholtmeyer et al., 2010) fueron los primeros en modificar Se7942 para la producción de D-lactato. Para ello, sobreexpresaron los genes *ldhA* y *lldP* de *E. coli* consiguiendo una producción final de 55 mg/L de D-lactato en 4 días. Más recientemente, (Li et al., 2015) consiguieron producir 829 mg/L de D-lactato en 10 días, mediante la expresión de la lactato deshidrogenasa LdhD modificada de *Lactobacillus bulgaricus* y la sobreexpresión del transportador de lactato LldP de *E. coli*.

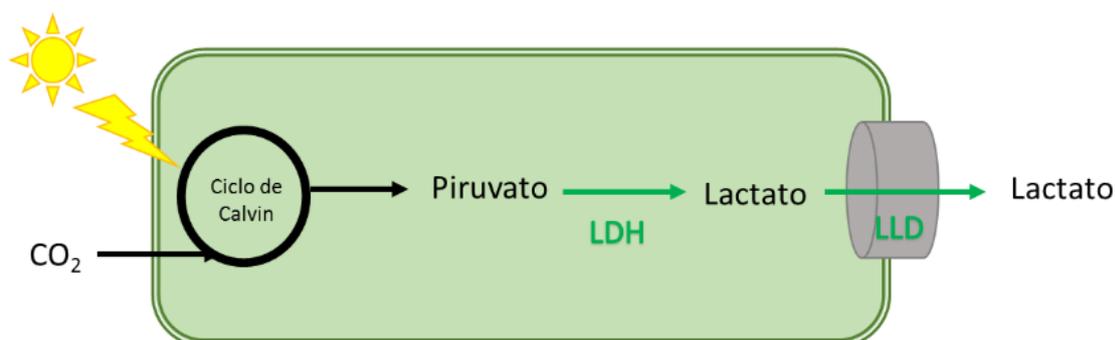


Figura 2-3. Esquema representativo de las modificaciones realizadas en Se7942 para la producción de lactato. El piruvato necesario para la síntesis de lactato es obtenido del ciclo de Calvin. Este piruvato es transformado por LDH en lactato. Este lactato es secretado al medio a través de LLD. En verde se representan aquellas dianas que han sido objeto de sobreexpresión para producir lactato. LDH, lactato deshidrogenasa (genes *ldhA* o *ldhD*); LLD, transportador de lactato (gen *lldP*).

2.1.2.2.4. Bioproducción de hidrógeno en Se7942

El hidrógeno molecular es uno de los potenciales futuros recursos energéticos como alternativa a los combustibles fósiles que existen actualmente. Sus principales ventajas como combustible son numerosas: es ecológico, eficiente, renovable y durante su producción y uso no se genera dióxido de carbono, aunque sí pequeñas cantidades de óxidos de nitrógeno (NO_x) (Dutta et al., 2005). Por estos motivos se ha modificado Se7942 para mejorar la producción fotobiológica de hidrógeno. Para ello, (Ducat et al., 2011a) expresaron la hidrogenasa, HydA de *Clostridium acetobutylicum* en esta cianobacteria no fijadora de nitrógeno. Esta hidrogenasa cataliza la reacción reversible $2 \text{H}^+ + 2 \text{e}^- \rightarrow \text{H}_2$, obteniendo los electrones directamente de la

ferredoxina. Con esta simple modificación se consigue un aumento de más de 500 veces la producción de hidrógeno frente a la cianobacteria silvestre.

2.1.2.2.5. Bioproducción de ácidos grasos en Se7942

La producción de ácidos grasos en cianobacterias, especialmente en Se7942, aparece descrita en el **apartado 2.3.2** de esta tesis.

2.1.2.2.6. Otros compuestos de interés biotecnológico

La principal olefina con interés biotecnológico que se ha sintetizado en Se7942 es el etileno. Las olefinas, comúnmente denominadas alquenos, son compuestos químicos que contienen por lo menos un doble enlace carbono-carbono. Este compuesto es un material bruto crítico en los procesos industriales, ya que a partir de él se producen varios productos petroquímicos. La fabricación de este compuesto lleva asociado un gran coste energético, así como una alta producción de CO₂ (uno de los gases responsables del efecto invernadero) y requiere petróleo, un recurso finito (Ungerer et al., 2012). Para evitar todo esto, se decidió utilizar Se7942 como organismo productor de etileno. (Takahama et al., 2003) consiguieron sintetizar bioetileno fotoautotóficamente mediante la introducción del gen *efe* de *Pseudomonas syringae*. Esta enzima EFE (enzima formadora de etileno) cataliza la formación de etileno desde 2-oxoglutarato, un intermediario en el ciclo de Krebs. El principal problema es que la cepa productora presentaba problemas de crecimiento y tenía un fenotipo verde-amarillo, producido por el estrés metabólico del consumo del 2-oxoglutarato y de la liberación del carbono como etileno. Más recientemente, (Jindou et al., 2014) modificaron genéticamente un complejo enzimático formador de etileno del tomate usando cohesinas-doquerinas de las bacterias *Clostridium thermocellum* y *Acetivibrio cellulolyticus*. Este complejo fue introducido en Se7942 mejorando la producción de etileno y disminuyendo los efectos nocivos causados por dicha producción.

Otro compuesto de interés sintetizado en Se7942 ha sido la dihidroxiacetona (DHA), usada frecuentemente en síntesis orgánica en las industrias cosméticas (da Silva et al., 2009). (Wang et al., 2015) han modificado a Se7942 para que produzca DHA a partir de glicerol usando

la GDH (Glicerol deshidrogenasa) de *Klebsiella pneumoniae*, consiguiendo una producción de 78,6 mg/L en 16 días de cultivo.

2.2. Evolución experimental

La evolución experimental es una estrategia de investigación que ofrece la oportunidad de estudiar procesos evolutivos de forma experimental a tiempo real. En otras palabras, podríamos definir la evolución experimental como el estudio de los cambios evolutivos en poblaciones experimentales como consecuencia de las condiciones impuestas por el investigador (Kawecki et al., 2012).

La esencia del experimento de evolución es conceptualmente simple. Durante decenas, centenas o incluso miles de generaciones, una población es expuesta a una nueva condición medioambiental. Al mismo tiempo, la misma población es mantenida en las condiciones medioambientales ancestrales, de tal manera que actúan como control del experimento. Normalmente, solo se cambia una variable medioambiental para mantener el experimento tan simple como sea posible. Sin embargo, es factible cambiar varios factores ambientales simultáneamente (Garland and Rose, 2009). El cambio en el entorno medioambiental podría causar tanto cambios a nivel genotípico, como a nivel fenotípico, como a ambos niveles (Barrick and Lenski, 2013). Más concretamente, podría decirse que las nuevas condiciones medioambientales experimentales proporcionan nuevas condiciones selectivas y, por consiguiente, promueven la evolución. Nuevas variantes genéticas (producidas mediante recombinación, mutación u otros procesos) pueden ser una ventaja o una desventaja en las condiciones alteradas, produciendo una reproducción diferencial y un incremento de los genotipos más favorecidos (Garland and Rose, 2009).

Si bien es cierto que el científico controla las condiciones ambientales, no por ello impone directamente la selección, ya que esta ocurre de forma natural (Buckling et al., 2009). Entre las principales condiciones definidas por el investigador, las más ampliamente ensayadas han sido cambios en factores particulares, como la temperatura (Bennett and Lenski, 1993), la nutrición (Kolss et al., 2009), otros factores estresantes (Dhar et al., 2011), presencia de parásitos (Zbinden et al., 2008) o competiciones (Santos, 1997; terHorst, 2011).

A lo largo de la historia, un amplio rango de organismos han sido estudiados con este fin, incluyendo nemátodos, moscas de la fruta y ratones, pero los organismos más fáciles y más productivos para trabajar son los microorganismos, fundamentalmente bacterias, levaduras y virus (Buckling et al., 2009).

Cabe destacar que la evolución experimental ha sido usada en varias ramas de la medicina y la tecnología. Durante décadas, ha sido el método elegido para el desarrollo de vacunas atenuadas contra enfermedades causadas por virus y bacterias, como la polio, la tuberculosis, la fiebre amarilla, el sarampión, las paperas y la rubeola. Por consiguiente, la evolución experimental ha contribuido a salvar millones de vidas humanas (Kawecki et al., 2012). Más recientemente, la evolución experimental combinada con la secuenciación genómica y el mapeo genético han sido utilizadas para identificar mutaciones que confieren resistencia a medicamentos contra patógenos, antes incluso de que las mutaciones aparecieran en la naturaleza (Hunt et al., 2010). Por otro lado, la evolución experimental también ha tenido un alto potencial en otras áreas de la biotecnología. Por ejemplo, en el área de los agentes biológicos para control de plagas, ha sido utilizados para desarrollar cepas de microorganismos (Kolodny-Hirsch, 1997) y metazoos (Dion et al., 2011) más virulentos frente a patógenos de plantas. La evolución experimental también ha sido usada para producir biocatalizadores y agentes de biocontrol con otras propiedades deseables, como alta resistencia térmica (de Crecy et al., 2009). Además, esta metodología puede ayudar a entender los procesos evolutivos producidos por especies invasoras (Lee et al., 2007), así como dar lugar a rescates evolutivos de poblaciones locales en extinción (Bell and Gonzalez, 2011). Además, los estudios de evolución experimental podrían arrojar luz en escenarios de respuestas bióticas al cambio global, como pueden ser las respuestas evolutivas de las algas a niveles elevados de CO₂ (Collins et al., 2006). Desde una perspectiva con aplicación en ingeniería, la evolución experimental ha sido usada como complemento o suplemento a los métodos de ingeniería clásica para generar, por ejemplo, organismos que produzcan biocombustibles (Wallace-Salinas and Gorwa-Grauslund, 2013).

Actualmente, se está trabajando en la mejora de los diseños de los experimentos de evolución. Por una parte, han comenzado a utilizarse simulaciones por ordenador que pueden ser usadas como una herramienta útil para el diseño experimental en combinación con la evolución experimental, como la estimación del poder de la selección artificial en los experimentos (Kessner and Novembre, 2015). Los experimentos de evolución requieren una sustancial inversión de tiempo y una labor de mantenimiento de las poblaciones, por ello se están diseñando sistemas automáticos para el cultivo y mantenimiento de estas poblaciones (Horinouchi et al., 2014; Winkler and Kao, 2014).

Como toda metodología experimental, la evolución experimental tiene sus limitaciones. En primer lugar, llevar a cabo a experimentos de evolución bajo un ambiente controlado puede que conduzca a la simplificación de las condiciones encontradas en la naturaleza, que a menudo

son más complejas. En segundo lugar, para observar las respuestas a la selección en un periodo de tiempo determinado, suelen utilizarse condiciones extremas para facilitar los procesos evolutivos, lo que puede dar lugar a patrones fenotípicos y genotípicos no observados en la naturaleza. Adicionalmente, los organismos modelo se han restringido a aquellos que tienen tiempos de generación rápidos en los experimentos que empiezan con poblaciones clonales, porque si no las mutaciones no ocurrirían tan deprisa como para proporcionar el material para la selección. Esto podría ser remediado, en parte, mediante la utilización de métodos que aumentan la variación genética dentro de las poblaciones (por ejemplo, introduciendo mecanismos deficientes de reparación de ADN). A pesar de estas limitaciones, los experimentos de evolución han contribuido significativamente en áreas de estudio de ecología y evolución (Hoang et al., 2016).

2.2.1. Evolución experimental en microorganismos

Los microbios fueron ignorados en gran medida por los biólogos evolutivos durante muchas décadas (Adams and Rosenzweig, 2014). Con el tiempo, sin embargo, su utilidad para los experimentos de evolución fue más evidente. El año 1950 se publica el primer artículo en el que describen la evolución de poblaciones bacterianas (Novick and Szilard, 1950).

Los experimentos de evolución en los que se utilizan microorganismos ofrecen incomparables oportunidades para evaluar la evolución de nuevas cualidades (Beaumont et al., 2009; Elena and Lenski, 2003). Los microorganismos que han sido usados en experimentos de evolución incluyen muchas bacterias y virus, así como algas unicelulares y hongos (Dragosits and Mattanovich, 2013; Kawecki et al., 2012).

Dentro de las características que hacen a estos organismos óptimos para este tipo de experimento destacan sus cortos tiempos de generación y los grandes tamaños de las poblaciones que ocupan pequeños espacios, que favorecen una rápida evolución (Kawecki et al., 2012). También destaca la facilidad con la que los ambientes pueden ser controlados y manipulados. Otra importante particularidad es que la mayoría de los microbios pueden ser almacenados congelados en un estado de no-evolución y posteriormente ser revividos. Esta propiedad permite la comparación directa de organismos de diferentes generaciones, por competición del genotipo ancestral frente al evolucionado para medir el “fitness” relativo, incluso frente a las formas intermedias (Dykhuizen and Hartl, 1983; Lenski et al., 1991). De

hecho, realizando ensayos simultáneos con organismos procedentes de diferentes generaciones, se consigue minimizar los efectos de fluctuaciones incontroladas que pueden confundir la interpretación de la recopilación de datos a diferentes tiempos (Kawecki et al., 2012). También pueden realizarse repeticiones donde la evolución es reiniciada desde generaciones intermedias para probar si un resultado de interés, como el origen de una nueva función, tuvo su origen en cambios tempranos (Blount et al., 2008; Meyer et al., 2012; Woods et al., 2011). Cabe destacar que muchos microorganismos tienen genomas relativamente simples y bien estudiados, permitiendo tanto la manipulación genética como la identificación de dianas genéticas mediante selección (Buckling et al., 2009). Por otro lado, muchos microbios presentan una reproducción asexual, dando lugar a individuos genéticamente iguales, lo que mejora la precisión de las réplicas experimentales (Elena and Lenski, 2003; Gerrish and Lenski, 1998).

Si comparamos los experimentos de evolución usando microorganismos frente a organismos pluricelulares, existen diferencias en varios sentidos. La mayoría de experimentos con microorganismos comienzan con un solo clon y depende de nuevas mutaciones para generar variaciones. La evolución a menudo tiene lugar por barridos selectivos consecutivos, aunque también pueden ser importantes las interacciones dependientes de frecuencia, así como la interferencia clonal (Rozen et al., 2009; Treves et al., 1998). Por el contrario, la evolución experimental utilizando organismos no microscópicos es causada normalmente por variaciones genéticas ya presentes en la población inicial, y los alelos son, con frecuencia, recombinados mediante sexo (Barrick and Lenski, 2013).

Cabe destacar, que muchos microorganismos son de gran importancia para los humanos, no solo como patógenos sino por sus numerosos efectos en ecosistemas esenciales. Por ello, es crucial entender los mecanismos y las dinámicas de la evolución microbiana (Elena and Lenski, 2003).

2.2.1.1. Evolución experimental en *Escherichia coli*

El experimento más conocido desarrollado hasta el momento en bacterias (concretamente en *Escherichia coli*) es el realizado por el científico norteamericano Richard Lenski. Este experimento conocido por sus siglas en inglés "LTEE" ("Long-Term Experimental Evolution"; <http://myxo.css.msu.edu/index.html>), fue iniciado el 15 de Febrero de 1988 y continúa hasta el momento (55.000 generaciones). El diseño experimental es relativamente

2. Introducción

sencillo, ya que consiste en el crecimiento continuado de 12 poblaciones de la bacteria *Escherichia coli* B en matraces idénticos conteniendo medio mínimo Davis suplementado con glucosa 25 mg/L como única fuente de carbono (DM25). Estos matraces son incubados a 37 °C y 120 rpm.

Estas poblaciones iniciales eran idénticamente iguales, excepto en un marcador neutral que distinguía seis líneas de las otras seis: la capacidad/incapacidad de crecer en arabinosa (Ara). Por tanto, 6 líneas eran Ara-, incapaces de crecer en L-arabinosa y su color era rojo cuando crecían en arabinosa tetrazolio (AT; medio compuesto por agar convencional, rico en arabinosa, con un colorante llamado tetrazolio). Estas líneas fueron designadas como Ara-1, Ara-2, Ara-3, Ara-4, Ara-5 y Ara-6. Por otro lado, utilizó 6 líneas Ara+, que se obtuvieron a partir de un mutante aislado por crecimiento de 10^9 células en medio mínimo agar con arabinosa. Estas colonias aparecían de color blanco cuando crecían en medio AT. Estas otras seis líneas fueron designadas como Ara+1, Ara+2, Ara+3, Ara+4, Ara+5 y Ara+6. La diferencia de color entre las colonias permitía el recuento de las mismas cuando se mezclaban los cultivos en placas (Blount et al., 2008).

Cada día, un 1 % de cada población era transferido a un matraz con medio fresco. Para *E. coli* B, este era un nuevo ambiente al cual estaba pobremente adaptado. En cada ciclo de 24 h de crecimiento habían transcurrido 6,6 generaciones. Después de cada 100 generaciones se tomaba una muestra de cada población para congelar junto a la cepa ancestral *E. coli* B. Un esquema del experimento se representa en la **Figura 2-4**.

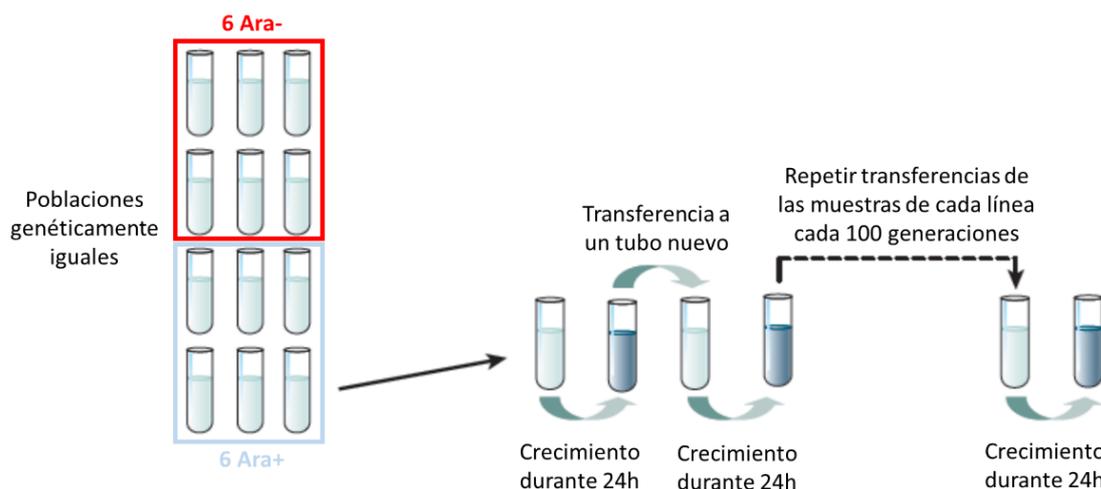


Figura 2-4. Esquema del experimento de evolución "LTEE" realizado por Richard Lenski. Este experimento de evolución parte de 12 poblaciones de *E. coli* idénticas, diferenciándose 6 de ellas respecto a las otras 6 en una mutación Ara. Estos cultivos se crecían durante 24 horas, después de las cuales eran transferidos a medio fresco. El experimento continuaba hasta que los cultivos alcanzaban las 100 generaciones, que además de seguir creciéndose, se congelaban muestras. Modificada de (Buckling et al., 2009).

Dado que estas 12 poblaciones eran inicialmente genéticamente idénticas, salvo en el marcado neutral Ara, el único recurso de selección natural que podía actuar era la mutación *de novo*. La adaptación de cada población al ambiente fue ensayada directamente por competición directa entre la cepa evolucionada frente a la cepa ancestral en ensayos realizados bajo las mismas condiciones de crecimiento del experimento. Para ello, se crecieron individualmente muestras de las poblaciones ancestrales y evolucionadas. Se tomó una pequeña cantidad de células de cada población y se mezclaron en proporción 1:1 en el mismo tubo y se realizó el ensayo de competición durante 24 h. Pasado ese tiempo, se tomaron muestras de la mezcla y se crecieron en medio sólido AT (descrito previamente) que permite distinguir las células de las poblaciones evolucionadas y de la ancestral en base al color que presentaban. La proporción entre ambas células fue utilizada para calcular el “fitness” de las líneas evolucionadas respecto al ancestro.

Para entender qué es el “fitness” y qué importancia tiene en los experimentos de evolución, se describe a continuación cómo se calcula dicho parámetro. El “fitness” de un organismo evolucionado es generalmente expresado en comparación con su ancestro. El “fitness” relativo se mide en un ensayo de competición entre el organismo ancestral y los organismos evolucionados. A menos que se diga lo contrario, los ensayos de competición se realizan en las mismas condiciones en las que se lleva a cabo el ensayo de evolución propiamente dicho (Buckling et al., 2009). En el experimento de evolución de Lenski los ensayos de competición se realizan como se describe a continuación, un procedimiento similar al usado con otros microorganismos. Los dos competidores se crecieron separadamente en el ambiente de competición para asegurarse que ambos estaban aclimatados a las condiciones del ensayo. A continuación, se mezclaron (normalmente en ratios 1:1) y se diluyeron (100 veces en este caso) en el ambiente de competición. Las densidades iniciales a tiempo 0 ($t = 0$) fueron estimadas diluyendo y sembrando las células en el medio AT que distingue las células de la línea evolucionada y ancestral por el color (descrito previamente). Después de un día ($t = 1$) (corresponde al ciclo de transferencia serial en el experimento de evolución), se calculaba el número de células plaqueándolas, como antes, en el medio con el indicador de color. Un esquema del proceso se representa en la **Figura 2-5** (Bennett et al., 1990; Elena and Lenski, 2003).

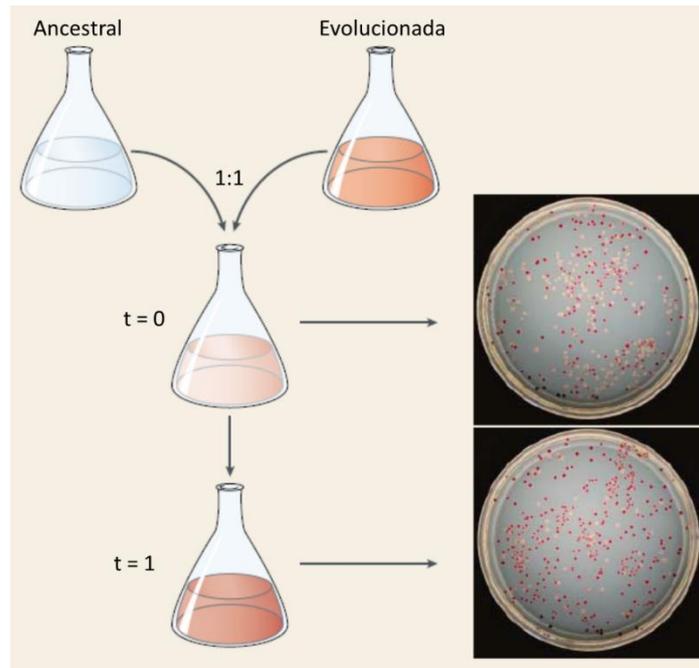


Figura 2-5. Esquema representativo del experimento para medir el “fitness”. Las células de las cepas ancestral y evolucionada se mezclan en un ratio 1:1. Se plaquea a tiempo 0 la mezcla del cultivo. Se deja crecer 24 horas más a 37 °C (t=1) y se plaquea. El número de células obtenidas a tiempo 0 y 1 nos va a permitir calcular el “fitness”, usando la **Ecuación 2-1**. Modificado de (Elena and Lenski, 2003).

La tasa de crecimiento de cada competidor es calculada como el logaritmo en base 2 del ratio entre la densidad final y la inicial (ajustada al factor de dilución usado en el plaqueo) (**Ecuación 2-1**). Por tanto, podemos definir el “fitness” relativo como la proporción de las tasas de crecimiento de la línea evolucionada y ancestral.

Ecuación 2-1. Ecuación para el cálculo del “fitness”. El “fitness” (W) se calcula como la división del logaritmo en base 2 del cociente entre el número de células de la cepa evolucionada ($N_{\text{evolucionada}}$) a tiempo 1 y 0, entre el logaritmo en base 2 del cociente entre el número de células de la cepa ancestral ($N_{\text{ancestral}}$) a tiempo 1 y 0. Si el valor de W es igual a 1, indica que ambas cepas son iguales en cuanto a “fitness” en esas condiciones, es decir, la cepa evolucionada ni ha mejorado ni ha empeorado. Si el valor es menor o mayor que 1, indica que la cepa evolucionada ha empeorado o mejorado su “fitness” frente a la cepa ancestral, respectivamente. Tomada de (Bennett et al., 1990).

$$W = \frac{\log_2 \frac{N_{\text{evolucionada}}(t = 1)}{N_{\text{evolucionada}}(t = 0)}}{\log_2 \frac{N_{\text{ancestral}}(t = 1)}{N_{\text{ancestral}}(t = 0)}}$$

En el año 1994, se publicaron los resultados de experimentos de evolución de las primeras 10,000 generaciones. Estos resultados demostraron que el “fitness” relativo de las 12 poblaciones experimentales había aumentado aproximadamente un 40 %, pero que la mayor parte de ese aumento había tenido lugar en las primeras 2.000 generaciones de evolución. También se observó que el tamaño celular casi se había duplicado durante 10.000 generaciones,

siguiendo una trayectoria similar al “fitness”, con un aumento más rápido durante las 2.000 generaciones. De manera importante, las trayectorias evolutivas del “fitness” y del tamaño celular, y la relación entre ambos, fue sutilmente diferente entre las poblaciones iniciales, sugiriendo que el azar y la contingencia, así como la selección natural, tuvieron papeles importantes en la evolución de estas poblaciones.

A lo largo de los 28 años que lleva realizándose el experimento de evolución en *E. coli*, se han estudiado muchas características en estas cepas evolucionadas (Lang and Desai, 2014). Una de ellas fue el estudio del tiempo y el modo de la evolución genética, es decir, qué mutaciones surgieron en estas cepas y en qué momento aparecieron. Para ello, secuenciaron los genomas de clones de *E. coli* de diferentes generaciones y los compararon con la secuencia de la cepa ancestral con el fin de identificar las mutaciones en los clones evolucionados. El estudio de las mutaciones a lo largo de las diferentes generaciones mostró que existía una discordancia entre el grado de cambios genómicos y la mejora en el “fitness”. En otras palabras, aunque la adaptación frena rápidamente, la evolución genética permanece casi constante durante 20.000 generaciones (Barrick et al., 2009). También se llevó a cabo un estudio sobre los sistemas de control genético del superenrollamiento para ver si habían cambiado durante el experimento de evolución. Se observaron cambios paralelos en la topología del ADN en las diferentes poblaciones, donde el nivel de superenrollamiento aumentaba, normalmente en las primeras 2.000 generaciones. Se descubrieron dos mutaciones en los genes *topA* y *fis* que controlan el superenrollamiento, que resultaron beneficiosas en los ensayos de competición. Estos experimentos indican que el control del superenrollamiento del ADN puede ser una diana clave en las poblaciones de bacterias evolucionadas (Crozet et al., 2005). Durante el experimento de evolución, surgió una población con una característica emergente muy importante: la capacidad de usar citrato como fuente de energía. La aparición de mutantes espontáneos que usen citrato (Cit+) es extraordinariamente rara, y tuvo lugar aproximadamente en la generación 33.000 (Blount et al., 2008). Para estimar el origen de este carácter Cit+ se analizaron y secuenciaron miles de clones de diferentes generaciones. Se descubrió entonces, que el origen de esta nueva función radica en la duplicación en tándem de parte del operón del citrato aguas abajo de un promotor expresado aeróbicamente, permitiendo así la expresión de un transportador de citrato previamente silenciado (Blount et al., 2012). Otro rasgo que surgió en las poblaciones evolucionadas en la generación 45.000 fue un cambio en la susceptibilidad a la infección por fagos. Por un lado, seis líneas evolucionadas tanto Ara+ como Ara-, presentaron susceptibilidad a la infección por el fago T6, mientras que la mayoría de ellas presentaron resistencia a la infección por el fago λ (Meyer et al., 2010). La resistencia al fago λ surge como un efecto

pleiotrópico de una mutación beneficiosa que desregula una vía metabólica no usada. El efecto opuesto puede atribuirse a la pérdida de la resistencia al fago T6. Mientras que la resistencia a este fago está causada por una única mutación presente en el gen *tsx*, la pérdida de esta resistencia estaría causada por la pérdida de dicha mutación. Es por ello, que dicha pérdida podría mejorar probablemente el coste de esta resistencia (Meyer et al., 2010; Meyer et al., 2012).

2.2.1.2. Evolución experimental en otros microorganismos

Tras el éxito del experimento de evolución iniciado por Lenski usando *E. coli* como organismo modelo, otros investigadores decidieron realizar experimentos similares en otros microorganismos. A continuación, se describen otros ejemplos de experimentos de evolución realizados en otros microorganismos, así como los resultados más destacados obtenidos a lo largo de dichos ensayos.

(Rainey and Travisano, 1998) llevaron a cabo un experimento utilizando a *Pseudomonas fluorescens* en dos ambientes diferentes, uno con y otro sin heterogeneidad espacial. El ambiente heterogéneo venía determinado por el cambio en la composición del medio de cultivo cuando no existía agitación. En este nicho, los microorganismos iniciaron una competición adaptativa en búsqueda de oxígeno. Después de varios días creciendo, las bacterias procedentes del medio homogéneo y heterogéneo fueron plaqueadas en medio sólido, con el fin de analizar la morfología de las colonias en cada una de las condiciones de crecimiento. Las colonias creciendo en presencia de heterogeneidad espacial dieron lugar a un rango diverso de morfología (**Figura 2-6, A**), mientras que las colonias que crecieron en ausencia de este no presentaron dicho fenómeno (**Figura 2-6, B**). Estas diferencias en los morfotipos de las diferentes colonias se deben a la diferente posición que ocupaban en medio líquido durante el experimento (Spiers et al., 2002).

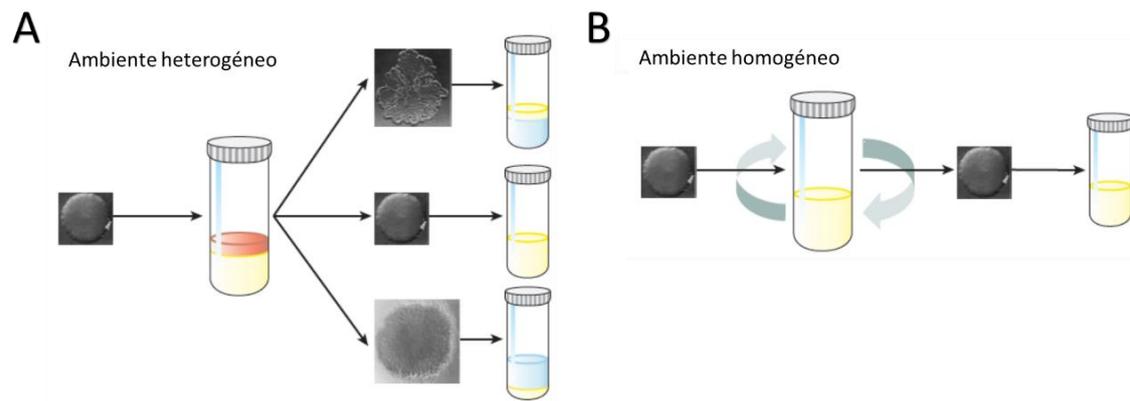


Figura 2-6. Esquema del proceso de evolución adaptativa de *Pseudomonas fluorescens* en un ambiente heterogéneo. (A) Morfología de células creciendo en un ambiente heterogéneo, favorecido por la ausencia de agitación. Las diferencias en esta morfología se deben al lugar que ocupan en el tubo durante el crecimiento, indicado por el anillo de color amarillo en cada tubo. **(B)** Morfología de células creciendo en un ambiente homogéneo, producido por una agitación continua. Todas presentan la misma morfología. Modificado de (Buckling et al., 2009).

(Velicer and Yu, 2003) iniciaron un experimento de evolución con la bacteria social *Myxococcus xanthus*, cuyas células son capaces de cooperar en busca de nutrientes y formar cuerpos fructíferos multicelulares. Este fenómeno conocido como “motilidad-S” requiere la presencia de un pilus extracelular tipo IV. La cepa ancestral era incapaz de formar este pilus, por lo que evolucionó desarrollando una nueva base mecánica para llevar a cabo esta cooperación. Esto trajo como consecuencia la aparición de nuevas morfologías en las colonias, favorecida, en parte, por la mejora en la producción de una matriz de fibras extracelulares que unía las células. En la **Figura 2-7** se muestran las morfologías de varias cepas evolucionadas, así como de la cepa ancestral y la cepa silvestre.

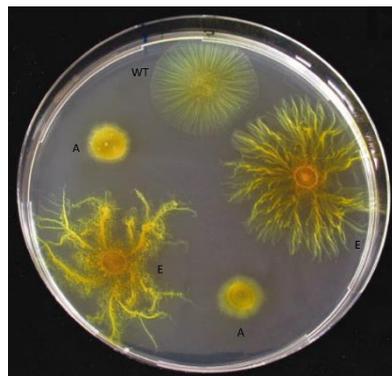


Figura 2-7. Morfología de las colonias de la bacteria *Myxococcus xanthus*. Cepa ancestral que carece de motilidad (A); cepa silvestre (WT) y evolucionadas (E) ambas poseen la capacidad de realizar motilidad. Modificado de (Velicer and Yu, 2003).

2.2.1.3. Evolución experimental en cianobacterias

Pocos estudios existen sobre evolución experimental en cianobacterias, sin embargo, en los últimos cuatro años ha aumentado ampliamente el número de publicaciones en esta área.

Dos experimentos de evolución se han llevado a cabo en *Synechocystis* sp. PCC 6803 (a partir de aquí, Ss6803), una de las cianobacterias modelo más ampliamente utilizada. En el primero de ellos, (Hosoda et al., 2014) llevaron a cabo un experimento de evolución en un medio rico. Ss6803 es incapaz de crecer en un medio bioquímicamente rico, ya que ciertos aminoácidos son tóxicos para las células a una concentración de 1 mM aproximadamente. Estos investigadores crecieron esta cianobacteria en un medio con 0,1 mM de los aminoácidos tóxicos, luego transfirieron los cultivos seriadamente a medios en los que iban aumentando gradualmente la concentración de los aminoácidos hasta llegar a 1 mM. A las 84 generaciones, las células evolucionadas crecían igual en medio con y sin aminoácidos tóxicos. Sin embargo, a 181 generaciones estas cepas crecían más rápido en el medio con aminoácidos tóxicos. Al secuenciar estas cepas observaron mutaciones en genes relacionados con las membranas celulares y los carboxisomas. En el segundo de ellos, (Uchiyama et al., 2015) llevaron a cabo un experimento de evolución en Ss6803 para que fuera capaz de crecer en medio ácido (pH < 5,75). Se crecieron las células a pH 6,0 y se fue bajando gradualmente el pH hasta llegar a 5,5 durante 3 meses. Se secuenciaron las cepas obtenidas, identificándose 11 SNPs (polimorfismos de un solo nucleótido) en las cepas adaptadas a medio ácido.

2.3. Síntesis de ácidos grasos

La síntesis de ácidos grasos es una de las rutas con mayor prevalencia en los organismos. En las células eucariotas y procariotas, los ácidos grasos son los precursores de una gran variedad de importantes bloques de construcción como los fosfolípidos, esfingolípidos y esteroides. Con la modificación del grado de saturación de los fosfolípidos en las membranas celulares, la fluidez de estas estructuras puede ser alterada, lo cual permite una adaptación a los cambios de temperatura (Janssen and Steinbuchel, 2014).

Debido a la importancia del proceso, la ruta de la biosíntesis de ácidos grasos está altamente conservada entre los diferentes reinos. En la primera etapa, se forma malonil-CoA por carboxilación de acetil-CoA consumiéndose ATP. Posteriormente, se intercambia el coenzima A por una proteína transportadora de grupos acilo (ACP) dando lugar a malonil-ACP. Estas reacciones forman parte de la etapa de iniciación de la síntesis de ácidos grasos. A continuación, comienza la etapa cíclica de elongación. En la primera vuelta del ciclo, el malonil-ACP se condensa con acetil-CoA dando lugar a coenzima A y acetoacetil-ACP. Este último es reducido a butiril-ACP, que entrará en el ciclo otra vez por condensación con malonil-CoA. La síntesis de ácidos grasos termina cuando se alcanza una cierta longitud en la cadena, y el acil-ACP es usado para la síntesis de fosfolípidos de membrana (Janssen and Steinbuchel, 2014).

Pese a la similitud existente en la ruta general, la evolución ha dado lugar a diferentes enzimas y diferentes organizaciones de las mismas. En animales y hongos, todos los pasos en la biosíntesis de ácidos grasos son llevados a cabo por sintasas de ácidos grasos tipo I, las cuales forman un complejo proteico multifuncional (Leibundgut et al., 2008). Por otro lado, las sintasas de ácidos grasos de tipo II predominan en procariotas y en los cloroplastos de plantas, en los cuales tiene lugar la síntesis de ácidos grasos *de novo*. Las sintasas de ácidos grasos de tipo II consisten en una colección de enzimas que no están organizadas en un solo gen u operón (White et al., 2005). Tanto en *E. coli* como en las cianobacterias, los genes que codifican para estas enzimas se denominan genes *fab* (del inglés “*f*atty *a*cid *b*iosynthesis”).

Las arqueas carecen de lípidos tradicionales basados en ácidos grasos, por lo que se creía que no poseían sintasas de ácidos grasos (Kates, 1992). Sin embargo, Lombard *et al.* publicaron la existencia de un nuevo tipo de sintasas de ácidos grasos ACP independiente, basándose en alineamientos frente a sintasas tipo II bacterianas (Lombard et al., 2012).

2.3.1. Metabolismo de ácidos grasos en *Escherichia coli*

2.3.1.1. Síntesis de ácidos grasos en *Escherichia coli*

Como se ha mencionado en el apartado anterior, la síntesis de ácidos grasos saturados comienza con la etapa de iniciación. La primera reacción es catalizada por el complejo acetil-CoA carboxilasa (AccABCD), que se encarga de producir malonil-CoA (**Figura 2-8, Reacción 1**). En esta reacción, se añade un grupo HCO_3^- al acetil-CoA dando lugar a malonil-CoA. En *E. coli*, cuatro proteínas forman el complejo AccABCD: biotina carboxilasa, proteína transportadora de carboxi-biotina, carboxilasa-CoA y carboxil transferasa (Choi-Rhee and Cronan, 2003). La actividad de este complejo multiproteico es inhibida por acil-ACP con cadenas de entre 6 y 20 átomos de carbono. De este modo, se evita una acumulación de ácidos grasos que no son usados para la síntesis de membranas (Davis and Cronan, 2001).

La segunda reacción de la etapa de iniciación es catalizada por la malonil-CoA ACP transacilasa (MCAT) (**Figura 2-8, Reacción 2**). Esta enzima, FabD, cataliza la transferencia de la fracción malonil al ACP, dando lugar a malonil-ACP (Joshi and Wakil, 1971). En *E. coli*, la delección del gen *fabD* se ha visto que es letal (Verwoert et al., 1994), y la sobreexpresión de dicho gen tiene como consecuencia la alteración de la composición de ácidos grasos. Más concretamente, la proporción de ácido palmitoleico (C16:1) disminuye, mientras que la proporción de ácido *cis*-vaccénico (C18:1n7) aumenta (Magnuson et al., 1992).

A continuación comenzaría el ciclo de elongación, cuya primera reacción sería catalizada por la enzima β -cetoacil sintasa III (FabH) (**Figura 2-8, Reacción 3**). El malonil-CoA se combina con una molécula de acetil-CoA en la primera reacción de condensación, formando β -cetoacil-ACP. La sobreexpresión de FabH da lugar a un aumento de ácidos grasos de cadena corta, C14 y C16 (Tsay et al., 1992). La delección de este gen conduce a un aumento de los niveles de ácidos grasos C18, pero tiene un efecto negativo sobre el crecimiento y disminuye el tamaño de las células (Yao et al., 2012).

Tras la reacción de condensación, el β -cetoacil-ACP es reducido a β -hidroxiacil-ACP usando NADPH (Toomey and Wakil, 1966). Esta reacción reversible es catalizada por la β -cetoacil-ACP reductasa (FabG) (**Figura 2-8, Reacción 4**). A continuación, el β -hidroxiacil-ACP es deshidratado por las β -hidroxiacil-ACP deshidratasas (FabA/FabZ) dando lugar a enoil-ACP (**Figura 2-8, Reacción 5**) (Heath and Rock, 1996). La siguiente reacción es catalizada por FabI,

una enoil-ACP reductasa que reduce el enoil-ACP a acil-ACP a expensas de NADPH o NADH (**Figura 2-8, Reacción 6**) (Bergler et al., 1996).

En la primera ronda descrita anteriormente, el acetil-CoA representa el inicio del ciclo. En este punto el ciclo comienza otra vez con la reacción de condensación del acil-ACP con otro malonil-CoA. En este caso la reacción es catalizada por cualquiera de las β -oxoacil sintasas I o II (FabB o FabF) (**Figura 2-8, Reacción 7**) (Edwards et al., 1997). La delección del gen *fabF* da lugar a un mutante sensible a temperatura (Garwin et al., 1980), mientras que la sobreexpresión es letal (Subrahmanyam and Cronan, 1998). Por otro lado, la delección del gen *fabB* da lugar a un mutante auxótrofo para ácidos grasos insaturados (Cronan et al., 1969) y la sobreexpresión de este gen incrementa el porcentaje de C18:1n-7 (Clark et al., 1983).

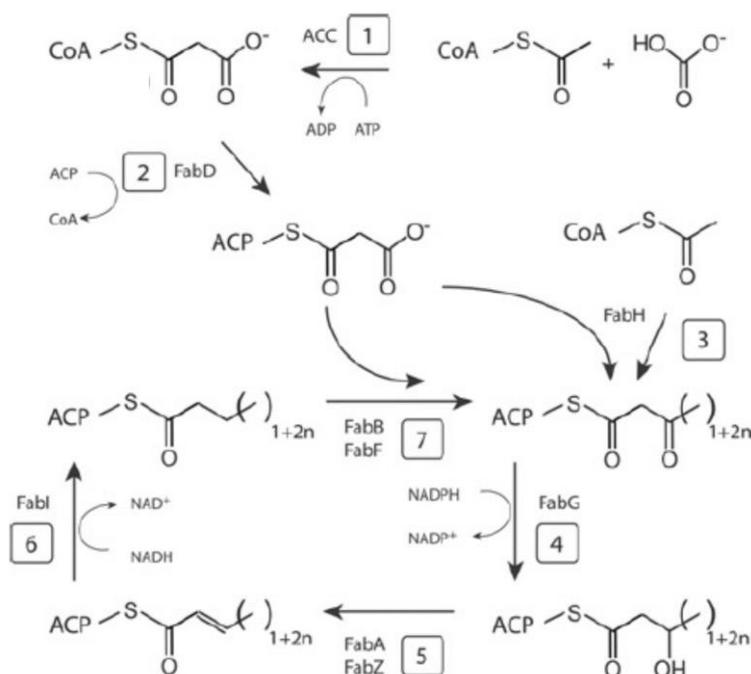


Figura 2-8. Ruta de síntesis de ácidos grasos saturados en *E. coli*. Las reacciones 1 y 2 de esta ruta están englobadas en la etapa de iniciación. El resto de reacciones pertenecen a la fase cíclica de elongación. ACC, acetil-CoA carboxilasa; FabD, malonil-CoA ACP transacilasa; FabH, β -cetoacil-ACP sintasa II; FabB, β -cetoacil-ACP sintasa I; FabF, β -cetoacil-ACP sintasa II; FabG, β -cetoacil-ACP reductasa; FabA/FabZ, C; FabI, enoil acil-ACP reductasa. Tomada de (Chan and Vogel, 2010).

La ruta descrita anteriormente detalla la producción de ácidos grasos saturados. En la **Figura 2-9**, se muestra un esquema de la síntesis de ácidos grasos insaturados en *E. coli*, el cual es ligeramente diferente al de síntesis de ácidos grasos saturados. El primer paso en la producción de ácidos grasos saturados es catalizado por FabA, que cataliza tanto una deshidratación como una reacción de isomerización. La función de isomerización es exclusivamente realizada sobre sustratos de 10 átomos de carbono (β -hidroxidecanoil-ACP).

2. Introducción

FabA convierte esta molécula en *trans*-2-decenoil-ACP y posteriormente isomeriza este intermediario a *cis*-3-decenoil-ACP. A continuación, FabB cataliza la reacción de condensación de este *cis*-3-decenoil-ACP con otro malonil-ACP, introduciéndose los dos grupos carbono cerca del grupo carbonilo. Seguidamente, tienen lugar tres o cuatro ciclos de elongación, tal y como tienen lugar en la ruta de síntesis de ácidos grasos saturados, para dar lugar a *cis*-9-palmitoleil-ACP (C16:1n7) y *cis*-11-vaccenoil-ACP (C18:1n7), respectivamente.

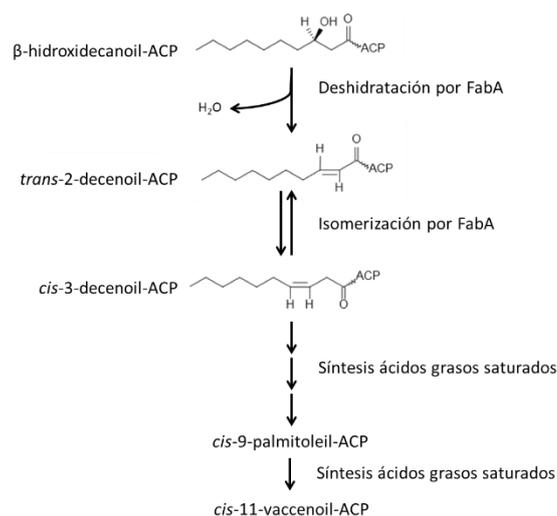


Figura 2-9. Ruta de síntesis de ácidos grasos insaturados en *E. coli*. FabA no solo puede reducir el precursor β -hidroxidecanoil-ACP, sino que también isomeriza el doble enlace desde el *trans*-2-decenoil-ACP a *cis*-3-decenoil-ACP. La posterior elongación de cadena por la acción de otras sintasas de ácidos grasos da lugar a la formación de C16:1 y C18:1.

2.3.1.2. Degradación de ácidos grasos en *Escherichia coli*

Las enzimas que participan en esta ruta de degradación pertenecen a la familia de genes *fad* (“*f*atty *a*cid *d*egradation”). La degradación de ácidos grasos puede comenzar de dos maneras, según la procedencia de estos ácidos grasos. En el caso de que los ácidos grasos provengan del medio externo, (**Figura 2-10, Paso 1**), primero tienen que atravesar la membrana externa por acción de un transportador de la familia FadL. El ácido graso se une al transportador y por un cambio conformacional, el poro se abre y los ácidos grasos difunden al periplasma (Lepore et al., 2011). El transporte del periplasma al citosol lo lleva a cabo FabD y está acoplado a la formación de una molécula de acil-CoA a expensas de ATP (Weimar et al., 2002). Si éstos proceden de la ruta de síntesis de ácidos grasos, se encuentran en forma de acil-ACP en el citoplasma (**Figura 2-10, Paso 2**). El grupo ACP es dissociado de la molécula acil por la acción de una tioesterasa (TesA), obteniéndose así una molécula de ácido graso libre y consumiéndose una molécula de agua. A continuación, actúa la proteína FadD, que, como en el caso anterior,

produce también una molécula de acil-CoA. El acil-CoA formado en ambos casos es consumido en la ruta de β -oxidación, por lo que una delección del gen *fabD* daría lugar a una acumulación de ácidos grasos libres en el citosol (Pech-Canul et al., 2011).

La degradación de las moléculas de acil-CoA continúa en un ciclo en el que se realizan los pasos inversos a los de la ruta de síntesis de ácidos grasos, dando lugar a la liberación de una molécula de acetil-CoA en cada ciclo. El ciclo comienza con la oxidación del acil-CoA a enoil-CoA catalizada por FadE, reduciendo el FAD a FADH₂ (Campbell and Cronan, 2002). La enzima FadB cataliza la hidratación de enoil-CoA a β -hidroxiacil-CoA y además oxida este intermediario a β -cetoacil-CoA (Yang et al., 1988). FadA cataliza el paso final del ciclo, en el cual se forman un acetil-CoA y un acil-CoA (reducido en dos átomos de carbono) (Feigenbaum and Schulz, 1975).

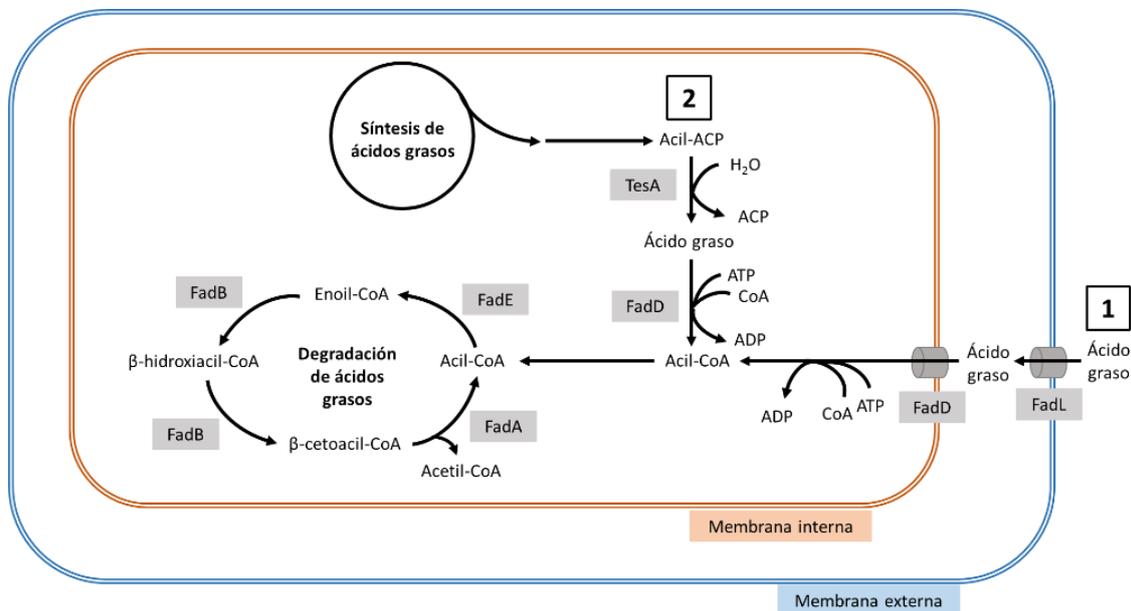


Figura 2-10. Ruta de degradación de ácidos grasos saturados en *E. coli*. TesA, tioesterasa; FadL, transportador de ácidos grasos de cadena larga; FadD, ácido graso de cadena larga-CoA ligasa; FadE, acil-CoA deshidrogenasa; FadB, enoil-CoA hidratasa; FadA, acetil-CoA aciltransferasa.

El catabolismo de los ácidos grasos insaturados no se conoce en profundidad, pero se cree que participan, además de las proteínas de la degradación de ácidos grasos saturados, las enzimas FadH (2,4-dienoil-CoA reductasa) (Tu et al., 2008) y FadM (tioesterasa II) (Nie et al., 2008).

2.3.2. Metabolismo de ácidos grasos en cianobacterias

2.3.2.1. Síntesis de ácidos grasos en cianobacterias

Poco se sabe sobre la ruta de síntesis de ácidos grasos saturados en cianobacterias, así como de las proteínas que participan en ella. Por esta razón, uno de los objetivos de esta tesis ha sido el estudio de dichos componentes.

En cianobacterias, así como en cloroplastos, los ácidos grasos son sintetizados utilizando sintetasas de ácidos grasos de tipo II (Froehlich et al., 1990). Como ocurre en el caso de *E. coli*, la ruta de síntesis de ácidos grasos saturados está dividida en dos etapas: etapa de iniciación y etapa de elongación. Además, habría una etapa adicional para la síntesis de ácidos grasos insaturados a partir de los saturados, la cual sería llevada a cabo por las desaturasas, que serán tratadas posteriormente en este apartado de la tesis.

En la **Tabla 2-2** se muestran las proteínas que participan en la etapa de iniciación de la ruta de biosíntesis de ácidos grasos en las principales cianobacterias. Para todas las cianobacterias, excepto para *Nostoc* sp. PCC 7120, todas las proteínas están codificadas por un sola copia del gen. Además, en esta cianobacteria la subunidad AccC de la acetil-CoA carboxilasa también está codificada por dos genes.

Tabla 2-2. Lista de genes que participan en la etapa de iniciación de la síntesis de ácidos grasos saturados en cianobacterias.

| Cepa | <i>accABCD</i> * | <i>fabD</i> ** | <i>fabH</i> |
|---|--|------------------|------------------|
| <i>Synechococcus elongatus</i> PCC 7942 | Synpcc7942_1595 Synpcc7942_2564 Synpcc7942_1379 Synpcc7942_1956 | Synpcc7942_1456 | Synpcc7942_1455 |
| <i>Synechococcus</i> sp. PCC 7002 | SYNPCC7002_A2443 SYNPCC7002_A0052 SYNPCC7002_A2127 SYNPCC7002_A0058 | SYNPCC7002_A2259 | SYNPCC7002_A0525 |
| <i>Synechocystis</i> sp. PCC 6803 | sll0728 slr0435 sll0053 sll0336 | slr2023 | slr1511 |
| <i>Nostoc</i> sp. PCC 7120 | alr5285 all5057 alr0939/alr2364 all2364 | alr0240 | alr0239 |
| <i>Spirulina platensis</i> NIES-39 | NIES39_M02550 NIES39_O04490 NIES39_O03320 NIES39_J05880 | NIES39_N01200 | NIES39_N01190 |

*El número de acceso de cada gen está en orden según la subunidad a la que corresponda. En el caso de haber dos copias del mismo gen, aparecen separadas por una barra inclinada.

**En el caso de que haya dos copias del mismo gen, aparecen separadas por una barra inclinada.

Stapleton *et al.* aislaron y caracterizaron la proteína malonil-CoA ACP transacilasa (FabD) de *Anabaena variabilis* (Stapleton and Jaworski, 1984). Más recientemente, se ha resuelto la estructura de la proteína FabD de la cianobacteria Ss6803 (Liu et al., 2013), mostrando un 40 % de identidad de secuencia con la proteína de *E. coli*. Por otro lado, (Kuo and Khosla, 2014) reconstituyeron la síntesis de ácidos grasos de Ss7002 en *E. coli*, para estudiar en más detalle esta ruta y la actividad de las enzimas que participan en ella. En este estudio se demostró que, aunque las subunidades individuales de los dos sistemas son cualitativamente intercambiables, son diferentes desde un punto de vista cuantitativo. Mientras que en *E. coli* la síntesis de ácidos grasos está limitada por su deshidratasa FabZ y en menor medida, por su enoil reductasa FabI, en Ss7002 el proceso está limitado por la proteína que inicia la síntesis de ácidos grasos, FabH. (Gu et al., 2016) establecieron que la especificidad de la enzima FabH por sustratos de cadena media-corta da lugar a una composición de ácidos grasos en la célula con estas mismas características.

En la **Tabla 2-3** se resumen las proteínas que actúan durante la etapa de elongación de la ruta de biosíntesis de ácidos grasos saturados en las principales. En ninguna de las cianobacterias que aparecen en esta tabla se encontró el gen que codifica para la proteína FabA, y solo en una de ellas, en Ss7002, aparece el gen *fabB*.

(Kuo and Khosla, 2014) demuestran, que mientras FabF era esencial para la síntesis de ácidos grasos, FabB era prescindible. En *E. coli*, FabB tiene un papel importante controlando el grado de insaturación de los ácidos grasos producidos. En Ss7002, este papel está controlado por las desaturasas, por lo que los autores especularon que FabB podría tener un papel adicional en los organismos de crecimientos rápido, aún sin descubrir. Cabe destacar que de las cinco cianobacterias analizadas en la **Tabla 2-3** solo una de ellas tiene dicha proteína, Ss7002. Por otro lado, (Moche et al., 2001) consiguieron cristalizar la proteína FabF de Ss6803, la cual muestra una homología de secuencia respecto a esta misma de *E. coli* de un 50 %.

Nanjo *et al.* desarrollaron un mutante de Ss6803 en el que se eliminó el gen *fabI* (Nanjo et al., 2010). La composición de ácidos grasos de dicho mutante no varió respecto a la de la cepa silvestre, pero era incapaz de crecer a temperaturas por encima de los 40 °C. Este resultado demuestra que FabI es necesario para la aclimatación de esta cianobacteria a condiciones de alta temperatura.

2. Introducción

Tabla 2-3. Lista de genes que participan en la etapa de elongación de la síntesis de ácidos grasos saturados en cianobacterias.

| Cepas | <i>fabG*</i> | <i>fabZ</i> | <i>fabA</i> | <i>fabI</i> | <i>fabB</i> | <i>fabF*</i> |
|---|-------------------------|------------------|-------------|------------------|------------------|-------------------|
| <i>Synechococcus elongatus</i> PCC 7942 | Synpcc7942_0684 | Synpcc7942_0930 | - | Synpcc7942_0126 | - | Synpcc7942_0537 |
| <i>Synechococcus</i> sp. PCC 7002 | SYNPCC7002_A2367 | SYNPCC7002_A0070 | - | SYNPCC7002_A1676 | SYNPCC7002_A1002 | SYNPCC7002_A1023 |
| <i>Synechocystis</i> sp. PCC 6803 | slr1994/slr0886/slr0330 | sl11605 | - | slr1051 | - | slr1332/sll1069 |
| <i>Nostoc</i> sp. PCC 7120 | all3805/all0968/alr1894 | alr2271 | - | all4391 | - | all4286/alr3343 |
| <i>Spirulina platensis</i> NIES-39 | NIES39_00260 | NIES39_L04560 | - | NIES39_C02750 | - | CHLREDRAFT_122688 |

*En el caso de que haya dos o más copias del mismo gen, aparecen separadas por una barra inclinada.

Tras la producción de ácidos grasos saturados tiene lugar la etapa de insaturación, en la que se introducen dobles enlaces en ácidos grasos C16 y C18. Mientras que muchas bacterias solo poseen una ruta anaeróbica para introducir insaturaciones mediante una deshidratasa (FabA), las cianobacterias tienen desaturasas que disponen de clústeres de hierro-sulfuro y requieren oxígeno y electrones proporcionados por una proteína donadora de electrones no caracterizada (Beld et al., 2015). Estas enzimas pertenecen al grupo de las acil-lípido desaturasas y están asociadas a las membranas tilacoides de estos organismos. La función principal de estas enzimas es convertir un enlace sencillo entre dos carbonos (C-C) en un doble enlace (C=C) en cadenas de ácidos grasos.

Un sistema de clasificación de cianobacterias en base a su composición de ácidos grasos fue propuesto por (Kenyon, 1972; Kenyon et al., 1972) y mejorado por (Murata et al., 1992), determinándose tres grupos. Recientemente, se ha actualizado este sistema de clasificación dividiendo a las cianobacterias en cuatro grupos, subdividiendo el grupo 3 en dos subgrupos (Los and Mironov, 2015). Los organismos del grupo 1 son capaces de introducir un solo doble enlace en la posición $\Delta 9$ de los ácidos grasos, normalmente ácidos grasos C16 o C18. A este grupo pertenece Se7942. En las cianobacterias del grupo 2, el ácido palmítico C16 y el ácido esteárico C18 son desaturados en las posiciones $\Delta 9$ y $\Delta 12$. Pertenecen a este grupo *Prochlorococcus marinus*. Al grupo 3 pertenecen las cianobacterias capaces de introducir tres dobles enlaces, subdividiéndose a su vez en dos grupos. Al grupo 3α , pertenecen las cianobacterias que introducen estos tres dobles enlaces en las posiciones $\Delta 9$, $\Delta 12$ y $\Delta 15$ y al grupo 3γ las que los introducen en las posiciones $\Delta 6$, $\Delta 12$ y $\Delta 15$, produciendo ácido α -linolénico (ALA) y ácido γ -linolénico, respectivamente. Al grupo 3α pertenecen Ss7002 y *Nostoc* sp. PCC 7120. Al grupo 3γ pertenece *Spirulina platensis* NIES-39. Finalmente, en el grupo 4, se encuentran las cianobacterias capaces de introducir cuatro dobles enlaces en las posiciones $\Delta 6$, $\Delta 9$, $\Delta 12$ y $\Delta 15$. A este grupo pertenece Ss6803.

El orden en el que actúan estas desaturasas está estrictamente determinado y se representa en la **Figura 2-11**. El primer doble enlace es introducido por la desaturasa $\Delta 9$ convirtiendo el ácido esteárico en oleico y el ácido palmítico en palmitoleico. Esta enzima está codificada por el gen *desC*. La inserción del segundo doble enlace es catalizado por la desaturasa $\Delta 12$, codificada por el gen *desA*. El ácido oleico pasa a linoleico y el ácido palmitoleico a hexadecadienoico. Sobre el ácido linoleico podrían actuar dos desaturasas, $\Delta 6$ o $\Delta 15$, dando lugar a ácido γ -linolénico o ácido α -linolénico, respectivamente. La $\Delta 6$ desaturasa está codificada por el gen *desD* y la desaturasa $\Delta 15$ por el gen *desB*. Finalmente, se produciría ácido

estearidónico a partir del ácido γ -linolénico o del ácido α -linolénico, por acción de la desaturasa $\Delta 15$ o $\Delta 6$, respectivamente (Tocher et al., 1998) (Los and Murata, 2004).

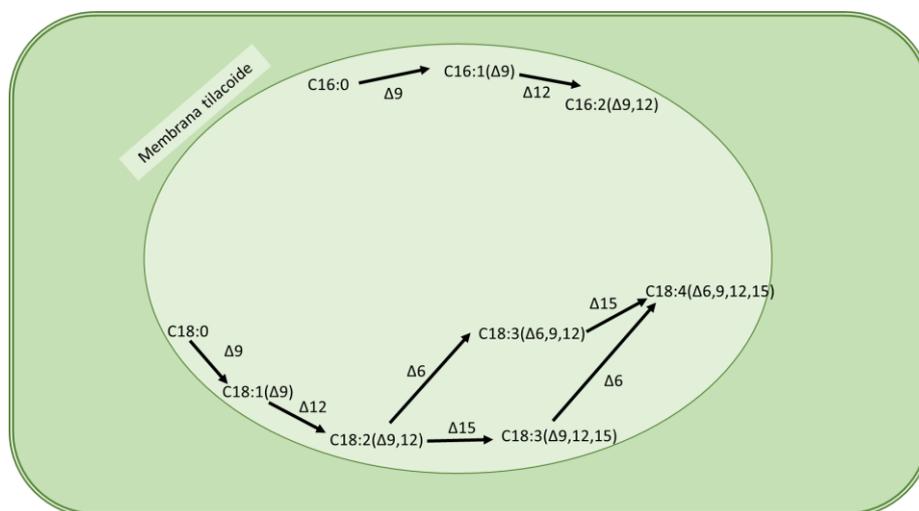


Figura 2-11. Esquema de las principales reacciones de desaturación que tienen lugar en las cianobacterias. $\Delta 9$, desaturasa $\Delta 9$; $\Delta 12$, desaturasa $\Delta 12$; $\Delta 15$, desaturasa $\Delta 15$; $\Delta 6$, desaturasa $\Delta 6$. C16:0, ácido palmítico; C16:1($\Delta 9$), ácido palmitoleico; C16:2($\Delta 9,12$), ácido hexadienoico; C18:0, ácido esteárico; C18:1($\Delta 9$), ácido oleico; C18:2($\Delta 9,12$), ácido linoleico; C18:3($\Delta 9,12,15$), ácido α -linolénico; C18:3($\Delta 6,9,12$), ácido γ -linolénico; C18:4($\Delta 6,9,12,15$), ácido estearidónico.

2.3.2.2. Degradación de ácidos grasos en cianobacterias

Las enzimas que participan en la degradación de ácidos grasos (β -oxidación) en *E. coli* están bien caracterizadas, como se ha expuesto en el apartado 2.3.1.2. En cianobacterias se cree que no existen homólogos de estas enzimas, salvo en el caso de FadD, que está caracterizada en varias especies de cianobacterias. Sin embargo, según la base de datos KEGG parece que existen dos cianobacterias, *Cyanothece* sp. PCC 7425 y *Acaryochloris marina*, que contienen enzimas de esta ruta de β -oxidación de ácidos grasos. Hasta el momento, ninguna de estas enzimas ha sido verificada experimentalmente. Por otro lado, Beld *et al.* realizaron un psi-blast para buscar homólogos de estas enzimas en cianobacterias, como Ss6803, no encontrando homólogos, salvo FadD (Beld et al., 2015). Estos autores sugirieron que Ss6803 podría utilizar otro mecanismo alternativo para degradar ácidos grasos, ya que son capaces de transportar y degradar ácidos grasos que se encuentran presentes en el medio de cultivo. La función de FadD ha sido bien caracterizada en Se7942 y Ss6803 (Kaczmarzyk and Fulda, 2010), así como en Ss7002 (Ruffing, 2014).

2.3.2.3. Modificaciones para mejorar la producción de ácidos grasos

2.3.2.3.1. Modificaciones para mejorar la producción de ácidos grasos en Se7942

La primera modificación genética llevada a cabo en Se7942 para cambiar el perfil de ácidos grasos, fue la introducción del gen *desA* que codifica para la desaturasa $\Delta 12$. Wada *et al.* transformaron Se7942 con el gen *desA* procedente de Ss6803, clonado en un vector de sobreexpresión (Wada et al., 1990). El transformante con esta construcción era capaz de producir ácidos grasos C16:2($\Delta 9,12$) y C18:2($\Delta 9,12$), un 5 y 6 %, respectivamente. Sin embargo, este vector era inestable, perdiéndose después de varias generaciones y con ello la capacidad de producción de este tipo de ácidos grasos. Años más tarde, Gombos *et al.* introdujeron este gen en el cromosoma de Se7942, mejorando así la estabilidad de la construcción. En este caso se produjo un 19 % de C16:2($\Delta 9,12$) y un 2 % de C18:2($\Delta 9,12$) (Gombos et al., 1997). Sakamoto *et al.* introdujeron en el cromosoma de Se7942 además del gen *desA*, el gen *desB* de Ss6803 que codifica para la desaturasa $\Delta 15$, consiguiendo así una producción del 5 % de ácido α -linolénico (Sakamoto et al., 1994).

En los últimos años, los trabajos de síntesis de ácidos grasos en cianobacterias, y especialmente en Se7942, se están centrando en la mejora de la producción de ácidos grasos libres. Ruffing *et al.* modificaron Se7942 delecionando el gen *fadD* y expresando una tioesterasa truncada, *TesA*, de *E. coli* (Ruffing and Jones, 2012). Con estas modificaciones se consiguió aumentar unas 30 veces la secreción de ácidos grasos libres. Siguiendo en esta línea, Ruffing *et al.* modificaron Se7942 delecionando *fadD* y sobreexpresando la tioesterasa (*Fat1*), la ribulosa-1,5-bifosfato carboxilasa/oxigenasa (*RuBisCO*) y las cuatro subunidades de la acetil-CoA carboxilasa de *C. reinhardtii* (Ruffing, 2013a). Estos tres genes sobreexpresados catalizan las reacciones de los tres pasos limitantes en la síntesis de ácidos grasos, consiguiendo con ello una mejora de unas 20 veces en la producción de ácidos grasos libre, alcanzando un valor de 40 mg/L.

Más recientemente, Ruffing *et al.* llevaron a cabo un análisis de “RNAseq” en cepas modificadas para la producción de ácidos grasos libres (Ruffing, 2013b), las cuales aparecen descritas en (Ruffing and Jones, 2012). Para cada cepa, se realizó un análisis de “RNAseq” para identificar aquellos genes que se estaban expresando diferencialmente durante la producción de ácidos grasos libres. Con este análisis se identificaron 15 genes diana que podrían utilizarse para reducir los efectos tóxicos de la producción de ácidos grasos libres. Kato *et al.* descubrieron

un tipo de transporte RND (resistencia-nodulación-división) en Se7942 con capacidad para exportar ácidos grasos libres durante la caracterización de un mutante productor de este tipo de ácidos grasos (Kato et al., 2015). Estos autores construyeron un mutante de Se7942 delecionando *fadD* y sobreexpresando *TesA*, en una cepa en la que se habían delecionado previamente los genes para el transporte del nitrato, consiguiendo así un crecimiento lento y limitado por nitrógeno en cultivos en discontinuo. Durante el experimento, surgió un mutante capaz de crecer en nitrato y además de secretar 45 mg/L de ácidos grasos libres. Al analizar este mutante se identificaron dos mutaciones puntuales responsables de un incremento en la expresión de los genes que componen este nuevo sistema de transporte y, como consecuencia, un mayor transporte de ácidos grasos. Por último, Kato *et al.* además de realizar las dos modificaciones más utilizadas para producir ácidos grasos libres, deleción de *FadD* y sobreexpresión de *TesA*, eliminaron el gen *wzt* del cromosoma de Se7942 (Kato et al., 2016). Esta deleción tuvo como resultado la pérdida de la capa hidrofílica de O-antígeno, aumentando la secreción de ácidos grasos libres, mitigando la fotoinhibición, y aumentando la densidad celular final de los cultivos. Todo ello permite que este mutante alcance una producción de aproximadamente 100 mg/L de ácidos grasos libres, la cantidad más alta hasta el momento obtenida en Se7942.

2.3.2.3.2. Modificaciones para mejorar la producción de ácidos grasos en Ss6803

Pocos estudios de producción de ácidos grasos libres se han llevado a cabo en Ss6803. Hasta el momento el nivel más alto de producción de ácidos grasos libres que se ha conseguido en esta cianobacteria, es de 197 mg/L, frente a la cepa silvestre que tan solo produce 1.8 mg/L. Liu *et al.* construyeron este mutante tras una serie de seis modificaciones genéticas consecutivas, en las que se introdujeron tioesterasas de *E. coli* y plantas y se sobreexpresaron los genes que codifican para la acetil-CoA carboxilasa (Liu et al., 2011). También se delecionaron los genes de síntesis de polihidroxibutarato (PHB), los genes de síntesis de cianoficina, el gen *fadD* y otros 3 genes más (que codifican para una proteína de superficie, una proteína de unión a penicilina y una fosfato acetiltransferasa). Recientemente, Cheah *et al.* analizaron el efecto de los ciclos de luz-oscuridad en la producción de ácidos grasos libres, viéndose esta significativamente reducida en comparación con los experimentos realizados en luz continua (Cheah et al., 2015).

Por otro lado, Chen *et al.* analizaron en Ss6803 la sobreexpresión de las desaturasas $\Delta 6$ y $\Delta 15$ endógenas y exógenas procedentes de dos hongos, *Gibberella fujikuroi* y *Mortierella alpina* (Chen et al., 2014). Con estas modificaciones se consiguió aumentar la producción de ácido α -linolénico y ácido estearidónico y disminuir la acumulación de ácido linoleico y ácido γ -linolénico.

2.3.2.3.3. Modificaciones para mejorar la producción de ácidos grasos en Ss7002

Ss7002 es otra cianobacteria utilizada y modificada para mejorar la producción de ácidos grasos libres. Ruffing *et al.* manipularon genéticamente Ss7002 para incrementar la producción de ácidos grasos libres, sobreexpresando la RuBisCO de Se7942, dos tioesterasas (una de *E. coli* y otra de *C. reinhardtii*) y delecionando el gen *fadD* (Ruffing, 2014). Todas estas modificaciones dieron lugar a un incremento en la producción de ácidos grasos libres de más de tres veces, alcanzándose concentraciones de ácidos grasos libres secretados superiores a los 130 mg/L.

No se han publicado más resultados de experimentos relacionados con ácidos libres secretados en esta cianobacteria. Las investigaciones más recientes se han centrado en la modificación del perfil de ácidos grasos libres producidos con el objetivo de incrementar los ácidos grasos de cadena corta, C12 y C14. El primer trabajo sobre esta línea fue presentado por (Work et al., 2015), los cuales modificaron genéticamente Ss7002 para que produjera ácido láurico (C12:0). Para ello, inhibieron la síntesis de glucógeno delecionando el gen *glgC* e inhibieron también la síntesis de proteínas mediante la privación de nitrógeno al cultivo. Además, la modificaron por delección de *fadD* y sobreexpresión de una tioesterasa derivada de *Umbellularia californica* (Laurel de California). Más recientemente, (Gu et al., 2016) se centraron tanto en la producción de ácido láurico como de ácido mirístico (C14:0). Para ello, deleccionaron en la cepa Ss7002 el gen *fabH* y lo sustituyeron por el de una diatomea, *Chaetoceros* sp. GSL56, que acumula elevados niveles de ácidos grasos C14. De esta manera, consiguieron que la cianobacteria produjera C12:0, no producido anteriormente por esta cepa. Por otro lado, consiguieron aumentar al menos cinco veces la producción de C14:0, dependiendo este nivel del aumento de la fuerza del promotor bajo el que se clonara *fabH* y de las condiciones de cultivo. Cabe destacar la importancia que tienen estos ácidos grasos de cadena corta, puesto que son más deseables para la producción de biodiesel.

2.4. Modelos metabólicos

2.4.1. Reconstrucción a escala genómica de un modelo metabólico

Un campo emergente de la biología de sistemas es la reconstrucción a escala genómica de modelos metabólicos, conocidos como GEMREs (“GEnome-scale Metabolic Reconstruction”). Estas reconstrucciones son bases de datos estructuradas de conocimiento específico que contienen información detallada del organismo en estudio, tales como la estequiometría de las reacciones metabólicas, su reversibilidad, las relaciones entre genes, proteínas y transformaciones bioquímicas, y los datos bioquímicos y fisiológicos disponibles para el organismo analizado (Feist et al., 2009; Thiele and Palsson, 2010).

Estas reconstrucciones pueden ser, además, transformadas a modelos matemáticos a escala genómica (GEMs, “GEnome-scale Models”), que permiten el cálculo de las capacidades metabólicas y el fenotipo de un organismo a partir de su genotipo. Contrariamente a los modelos cinéticos y estocásticos, que requieren gran cantidad de datos cinéticos, los GEMs permiten obtener información metabólica detallada, en términos de distribución de flujo sujetos a una serie de asunciones (Orth et al., 2010). Por lo tanto, la reconstrucción y el análisis basado en restricciones (COBRA, “COnstraint-Based Reconstruction and Analysis”) (Hyduke et al., 2011; Lewis et al., 2012) es escalable a escala genómica y se basa en la aplicación de restricciones impuestas por el genoma (límites internos) y el medio ambiente (límites de los sistemas). Así, describe el conjunto de estados fenotípicos viables en estas condiciones. Un método cada vez más aceptado para analizar estos GEMs es el análisis del balance de flujos (FBA, “Flux Balance Analysis”). FBA es un enfoque matemático para el estudio de los modelos metabólicos en términos de distribución de flujos y se basa en la resolución de un problema de optimización lineal mediante la maximización o minimización de una función objetivo $\mathbf{Z} = \mathbf{c}^T \cdot \mathbf{v}$, sujeta a un conjunto de restricciones.

Matemáticamente, la aplicación de FBA puede representarse como sigue:

$$\max \mathbf{c}^T \cdot \mathbf{v}$$

$$\text{sujeto a } \mathbf{S} \cdot \mathbf{v} = \mathbf{0}$$

$$\mathbf{v}_{\min} \leq \mathbf{v} \leq \mathbf{v}_{\max}$$

Donde \mathbf{Z} es una combinación lineal de los flujos metabólicos, \mathbf{c}^T es un vector con n elementos que contiene ceros para todas las entradas excepto para la reacción (o reacciones) que forma parte de la función objetivo (por ejemplo, tasa de crecimiento), y que indica cómo cada reacción contribuye a la función objetivo. \mathbf{S} es una matriz $m \times n$ que contiene todos los coeficientes estequiométricos en el modelo de m metabolitos y n reacciones. El vector \mathbf{v} tiene n elementos, que representan los valores de flujo individuales para cada reacción que se va a determinar. Las limitaciones $\mathbf{S} \cdot \mathbf{v} = \mathbf{0}$ corresponden a la conversión de la masa en estado estacionario. Los vectores \mathbf{v}_{\min} y \mathbf{v}_{\max} son vectores con n elementos cada uno, que representan el límite más bajo (lb, “lower bounds”) y más altos (ub, “upper bounds”) en los flujos, respectivamente. La salida de este problema de optimización lineal es un vector de distribución de flujo particular \mathbf{v} , que maximiza o minimiza la función objetivo. Una explicación más detallada de todos estos parámetros y de la lógica del sistema se puede encontrar en (Palsson, 2006).

Los datos de transcriptómica, proteómica, metabolómica y fluxómica que se recogen durante los experimentos “ómicos”, se pueden utilizar como andamio en la construcción de los GEMs. En paralelo, los progresos realizados en bioinformática y en la aplicación de análisis computacional, a escala genómica, han contribuido a la mejor comprensión de los sistemas biológicos. La correcta contextualización y el uso efectivo de este enorme volumen de datos suponen retos matemáticos y estadísticos. Es por ello que el desarrollo de métodos para integrar este conjunto de datos es un desafío que los científicos computacionales están abordando actualmente.

En los últimos 15 años, los GEMs han ganado un atractivo en investigación. En el año 2000 fue publicado el primer modelo GEM para *E. coli* MG1655 (Edwards and Palsson, 2000). Este modelo ha continuado actualizándose, publicándose en 2007 el modelo *iAF1260* que contiene 1.260 genes, el 29 % de los genes anotados, 1.039 metabolitos y 1.387 reacciones (Feist et al., 2007). Usando datos de genes esenciales y de otro tipo, este modelo fue actualizado en 2011 a *iJO1366*, conteniendo 1.366 genes, el 32 % de los genes anotados, 1.136 metabolitos y 1.473 reacciones (Orth et al., 2011). Dentro de los microorganismos eucariotas, la levadura *Saccharomyces cerevisiae* es el organismo modelo mejor caracterizado. Desde el primer GEM publicado en 2003 (*iFF708*) (Forster et al., 2003), el número de metabolitos y reacciones ha aumentado de 584 a 1.353 y de 1.145 a 1.566, respectivamente, en el modelo recientemente publicado *iTO977* (Osterlund et al., 2013). Actualmente, están disponibles más de 100 GEMs y el número de organismos con GEMs reconstruidos está creciendo rápidamente debido a los avances en tecnologías de alto rendimiento y en herramientas para la recogida automática de datos y la reconstrucción de “drafts” (Henry et al., 2010; Thiele and Palsson, 2010).

2.4.2. Interés biotecnológico de los modelos metabólicos

Los GEMs han ampliado el espacio metabólico analizable computacionalmente y el alcance de sus aplicaciones (Lewis et al., 2012; McCloskey et al., 2013). Esta expansión proporciona herramientas para abordar estudios de ingeniería metabólica, obtener nueva información biológica, predecir nuevos fenotipos celulares, estudiar procesos evolutivos y modelos de interacciones interespecies (Liao et al., 2012; McCloskey et al., 2013; Monk et al., 2014).

Los GEMs bacterianos se han utilizado en ingeniería metabólica para manipular y optimizar rutas bioquímicas nativas con el objetivo de producir compuestos relevantes a nivel industrial y terapéutico. Dentro de las estrategias utilizadas se incluyen delecciones de genes (Fong et al., 2005), aumento y disminución de la expresión de genes (Fowler et al., 2009), así como mapeo de datos de alto rendimiento para identificar cuellos de botella o rutas competidoras (Lee et al., 2007). Además, hay que destacar que se ha conseguido la integración de rutas no nativas para la producción de compuestos que no son producidos en la cepa hospedadora o en muy poca cantidad (Jung et al., 2011; Yim et al., 2011). Para la predicción de estas nuevas rutas biosintéticas se han usado complejos algoritmos computacionales (Kim et al., 2015). Los métodos más avanzados han permitido la identificación de nuevas estrategias para acoplar el crecimiento bacteriano a la producción del compuesto diana (Kim et al., 2015; Lewis et al., 2012; Zhang and Hua, 2015).

El uso de los GEMs ha permitido también ampliar nuestro conocimiento actual sobre ciertas funciones celulares. Esto es debido a que existen muchos aspectos de las funciones bacterianas que están aún sin caracterizar. Por ejemplo, en *E. coli*, la bacteria más conocida y estudiada, se desconoce la función del 34 % de sus genes (Orth et al., 2011). La función de marcos de lectura abiertos (ORFs) aún desconocida, puede determinarse por comparación directa del crecimiento obtenido de los mutantes delectivos *in silico* frente a datos obtenidos *in vivo*. Las discrepancias entre los resultados experimentales y las predicciones podrían indicar ausencia de información o discrepancias funcionales (McCloskey et al., 2013).

Por otro lado, los GEMs pueden utilizarse para predicciones fenotípicas mediante el análisis basado en restricciones (Palsson, 2006). De esta manera, se obtendría un espacio de soluciones de fenotipos potenciales que permitiría al organismo sobrevivir en un fondo genético y medioambiental dado (Ibarra et al., 2002).

El contenido genético, así como el fenotipo de las bacterias se encuentra constantemente adaptándose a las condiciones medioambientales impuestas. Las adaptaciones ocurren por eliminación de reacciones individuales debidas a mutaciones con pérdida de función, alteraciones en la expresión génica y capacidad enzimática y alteraciones en la cinética enzimática. También pueden tener lugar mediante la adición de nuevas reacciones por transferencia horizontal de genes, duplicación génica y mutaciones con ganancia de función. Se ha demostrado el uso de GEMs en el modelado de la evolución microbiana a través de la eliminación y adición de nuevo contenido en las redes metabólicas (McCloskey et al., 2013).

Por último, cabe destacar el creciente interés de los GEMs para entender mejor como tienen lugar la interacción hospedador-patógeno. Esta metodología se utiliza con el objetivo de desarrollar antimicrobianos mejorados (Lebeis and Kalman, 2009), utilizar microorganismos en bioremediación (Singh et al., 2011), así como comprender y manipular el microbioma del intestino humano para mejorar la salud (Walter et al., 2011).

2.4.3. Modelos metabólicos de cianobacterias

Las reconstrucciones metabólicas de organismos fototróficos están infrarrepresentadas cuando se compara con microorganismos heterotróficos. Un resumen de las reconstrucciones existentes para Ss6803 se muestra en la **Tabla 2-4**. La primera aplicación de FBA en el metabolismo de cianobacterias, en Ss6803 en concreto, fue llevado a cabo por (Shastri and Morgan, 2005), seguida por una extensión del modelo por (Hong and Lee, 2007). Ambas reconstrucciones están enfocadas únicamente en el metabolismo central del carbono (93 y 56 reacciones, respectivamente). Los modelos contienen un ciclo de Krebs incompleto, ya que las cianobacterias carecen de los genes que codifican para el complejo de la 2-oxoglutarato deshidrogenasa, responsable de que esta ruta sea cíclica. En su lugar, se ha incluido una desviación de esta ruta hacia la síntesis del glioxilato. Esta desviación se incluye porque a pesar de que se desconocen los genes que codifican para las enzimas que intervienen en dichas reacciones, existen evidencias experimentales sobre la actividad de las mismas (Yang et al., 2002). El primer modelo a gran escala de Ss6803 fue proporcionado por (Fu, 2009), y contiene 633 genes, alrededor del 20 % de los genes. Sin embargo, la reconstrucción solo presentaba una pequeña curación manual (basada en información genética y bioquímica) y un flujo real restringido al metabolismo central, principalmente por el uso de una función de biomasa con un número limitado de metabolitos. El modelo también incluía una desviación a la ruta de síntesis

2. Introducción

de glioxilato. (Knoop et al., 2010) presentaron una reconstrucción mejorada de este modelo, al incluir ciertas rutas metabólicas ausentes en el modelo de (Fu, 2009). Sin embargo, el número de genes se redujo con respecto al anterior, al contener tan solo un 11 % de los genes de Ss6803. La reconstrucción incluía la primera representación detallada de la ruta fotorrespiratoria en una reconstrucción metabólica. Poco después, tres reconstrucciones adicionales fueron publicadas (Montagud et al., 2010; Montagud et al., 2011; Yoshikawa et al., 2011), cada una de las cuales incorporaba un ciclo de Krebs incompleto y una desviación hacia la ruta de síntesis del glioxilato. El análisis de (Yoshikawa et al., 2011) también compara las soluciones de (Knoop et al., 2010) y (Montagud et al., 2010) bajo condiciones de crecimiento heterotrófico. Una construcción más fue presentada por (Nogales et al., 2012), haciendo uso de una función de biomasa mejorada. El modelo incluye la ruta completa de síntesis de ácidos grasos, incluyendo los ácidos grasos omega-3, y de los pigmentos fotosintéticos (clorofil a, filoquinona, licopeno, β -carotenoides, zeaxantina, equinona, α -tocoferol y β -tocoferol). Este nivel de detalle es lo que permitió definir una ecuación de biomasa más precisa. Destaca también la comparación del potencial metabólico de *Cyanothece* sp. ATCC 51142 y Ss6803 realizada por (Saha et al., 2012), que proporciona información valiosa sobre las funciones biológicas en sus nichos y también abre nuevas vías para la producción económica de biocombustibles. El modelo más reciente es el descrito por (Knoop et al., 2013), que incluye actualizaciones en el ciclo de Krebs, la desviación hacia la ruta de síntesis del glioxilato, así como el papel de la fotorrespiración en el crecimiento celular.

Tabla 2-4. Reconstrucciones metabólicas para la cianobacteria Ss6803.

| Nombre del modelo | Genes | Reacciones | Metabolitos | % de genes anotados | Referencia |
|-------------------|-------|------------|-------------|---------------------|----------------------------|
| - | - | 93 | 46 | - | (Shastri and Morgan, 2005) |
| - | 78 | 56 | 54 | 2.5 % | (Hong and Lee, 2007) |
| - | 633 | 831 | 704 | 20 % | (Fu, 2009) |
| - | 337 | 380 | 291 | 11 % | (Knoop et al., 2010) |
| <i>iSyn669</i> | 669 | 882 | 790 | 21 % | (Montagud et al., 2010) |
| <i>iSyn811</i> | 811 | 956 | 911 | 26 % | (Montagud et al., 2011) |
| - | 376 | 493 | 465 | 12 % | (Yoshikawa et al., 2011) |
| <i>iJN678</i> | 678 | 863 | 795 | 21 % | (Nogales et al., 2012) |
| <i>iSyn731</i> | 731 | 1.156 | 996 | 23 % | (Saha et al., 2012) |
| - | 677 | 759 | 601 | 21 % | (Knoop et al., 2013) |

Se detalla el número de genes, reacciones y metabolitos presentes en el modelo, así como el porcentaje de genes que tiene el modelo respecto al total de genes anotados para esa cepa. También se indica el nombre del modelo, en los casos en los que se le haya dado nombre, así como la referencia donde se describe dicho modelo. Modificado de (Knoop et al., 2013).

Además de los modelos de Ss6803, se ha reconstruido el metabolismo de otras cianobacterias. Se resumen en la **Tabla 2-5**. Todos ellos corresponden a cianobacterias

unicelulares, excepto el modelo de *Nostoc* sp. PCC 7120, una cianobacteria multicelular filamentosa (Malatinszky et al., 2016). Es interesante mencionar el número de reconstrucciones de especies de *Cyanothece*, a pesar de que su uso como modelo para la producción de compuestos de interés biotecnológico no está muy extendido.

Del organismo modelo utilizado en esta tesis, Se7942, se han desarrollado dos reconstrucciones metabólicas. La primera (Triana et al., 2014) incluye una ecuación de biomasa y demuestra su aplicabilidad mediante la simulación de condiciones de crecimiento autotróficas. La segunda (Broddrick et al., 2016) utiliza datos específicos de metabolismo fotoautotrófico, así como datos de esencialidad de genes descritos previamente por (Rubin et al., 2015). Estos datos de esencialidad permitieron grandes mejoras en la precisión del modelo. Además, este modelo describe expresamente la absorción de fotones y el auto-sombreado, consiguiendo así modelar con precisión la luz como un metabolito. El auto-sombreado es consecuencia directa de la posición de las células en el cultivo respecto a la fuente de luz. De esta manera, las células más cercanas a la fuente de luz absorben un exceso de fotones, impidiendo que las células que se encuentran en el interior del cultivo alcancen su máxima tasa de crecimiento.

Tabla 2-5. Reconstrucciones metabólicas para varias cianobacterias modelo.

| Cianobacteria | Nombre del modelo | Genes | Reacciones | Metabolitos | % de genes anotados | Referencia |
|---|-------------------|-------|------------|-------------|---------------------|---------------------------|
| <i>Cyanothece</i> sp. ATCC 51142 | <i>iCce806</i> | 806 | 667 | 587 | 16 % | (Vu et al., 2012) |
| <i>Synechococcus</i> sp. PCC 7002 | <i>iSyp611</i> | 611 | 591 | 544 | 19 % | (Hamilton and Reed, 2012) |
| <i>Spirulina platensis</i> C1 | <i>iAK692</i> | 692 | 875 | 837 | 11 % | (Klanichui et al., 2012) |
| <i>Cyanothece</i> sp. ATCC 51142 | <i>iCyt773</i> | 773 | 946 | 811 | 15 % | (Saha et al., 2012) |
| <i>Cyanothece</i> sp. PCC 7424 | <i>iCyc792</i> | 792 | 1.242 | 1.107 | 14 % | (Mueller et al., 2013) |
| <i>Cyanothece</i> sp. PCC 7425 | <i>iCyn731</i> | 731 | 1.306 | 1.160 | 14 % | (Mueller et al., 2013) |
| <i>Cyanothece</i> sp. PCC 7822 | <i>iCyj826</i> | 826 | 1.258 | 1.110 | 14 % | (Mueller et al., 2013) |
| <i>Cyanothece</i> sp. PCC 8801 | <i>iCyp752</i> | 752 | 1.172 | 994 | 17 % | (Mueller et al., 2013) |
| <i>Cyanothece</i> sp. PCC 8802 | <i>iCyh755</i> | 755 | 1.161 | 973 | 17 % | (Mueller et al., 2013) |
| <i>Synechococcus elongatus</i> PCC 7942 | <i>iSyf715</i> | 715 | 851 | 838 | 26 % | (Triana et al., 2014) |
| <i>Spirulina platensis</i> NIES-39 | - | 620 | 746 | 673 | 24 % | (Yoshikawa et al., 2015) |
| <i>Synechococcus</i> sp. PCC 7002 | - | 728 | 742 | 684 | 23 % | (Hendry et al., 2016) |
| <i>Synechococcus elongatus</i> PCC 7942 | <i>iJB785</i> | 785 | 850 | 768 | 29 % | (Broddrick et al., 2016) |

2. Introducción

| Cianobacteria | Nombre del modelo | Genes | Reacciones | Metabolitos | % de genes anotados | Referencia |
|-----------------------------------|-------------------|-------|------------|-------------|---------------------|----------------------------|
| <i>Nostoc</i> sp. PCC 7120 | - | - | 804 | 777 | - | (Malatinszky et al., 2016) |
| <i>Synechococcus</i> sp. PCC 7002 | iSyp821 | 821 | 723 | 777 | 26 % | (Qian et al., 2016) |

Se detalla el nombre del modelo, número de genes, reacciones y metabolitos, porcentaje de genes del modelo respecto al total de los genes anotados para esa bacteria y referencia. Aparecen por orden cronológico.

Objetivos

3. Objetivos

1. Determinar la función de las mutaciones en cepas de Se7942 obtenidas por evolución experimental y que han resultado en un aumento de su velocidad de crecimiento.
 - a. Análisis de las mutaciones aparecidas durante el experimento de evolución.
 - b. Clonación de los genes que incluyen las mutaciones detectadas y estudio de su relación con la disminución del tiempo de generación.
 - c. Efecto de las mutaciones en la expresión de genes.
 - d. Efecto de las mutaciones en otras características: transformabilidad natural, variaciones en el espectro de absorción, composición de pigmentos, formación de biofilms.
2. Ingeniería metabólica de Se7942 para la producción de ácido alfa-linolénico.
 - a. Análisis de promotores bajo los que se clonarán las desaturasas involucradas en la síntesis de ácido alfa-linolénico.
 - b. Modificación genética de Se7942 con las desaturasas de la cepa Ss7002 para la producción de ácido alfa-linolénico.
 - c. Identificación de genes involucrados en la síntesis de ácidos grasos saturados en Se7942 que mejoran la producción de ácido alfa-linolénico.
 - d. Validación de las dianas identificadas y análisis de su efecto en la producción de ácido alfa-linolénico.
3. Desarrollo de un modelo metabólico de Se7942.
 - a. Construcción de un nuevo modelo metabólico de Se7942 y comparación con los actualmente publicados.
 - b. Identificación de dianas para mejorar la producción de ácido alfa-linolénico en Se7942.

Materiales y Métodos

4. Materiales y Métodos

4.1. Cepas bacterianas

4.1.1. *Escherichia coli*

Las cepas de *Escherichia coli* utilizadas se detallan en la **Tabla 4-1**.

Tabla 4-1. Cepas de *E. coli*.

| Nombre | Genotipo | Referencia |
|--------------|--|------------------------|
| DH5 α | <i>F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ- [Nx^R]</i> | (Grant et al., 1990) |
| β 2150 | <i>ΔdapA::(erm-pir) thrB1004, pro, thi, strA, hsdS, lacZΔM15, (F' lacZΔM15 lacI^q, traD36, proA+, proB+) [Em^R Sm^R]</i> | (Demarre et al., 2005) |

4.1.2. Cianobacterias

Las cianobacterias se detallan en la **Tabla 4-2**.

Tabla 4-2. Cepas de cianobacterias.

| Nombre | Fenotipo | Características | Construcción | Origen |
|---|-----------------|---|--|----------------------------|
| <i>Synechococcus elongatus</i> PCC 7942/2 (NC_007604) | - | Estirpe salvaje de Se7942, la cual contiene pANL (AF441790) pero carece de pANS (NC_004990) | - | Pasteur Culture Collection |
| <i>Synechococcus</i> sp. PCC 7002 (NC_010475.1) | - | Estirpe salvaje de Ss7002 | - | Pasteur Culture Collection |
| GRPS1 | Sm ^R | Se7942 cuyo gen <i>rps12</i> ha sido sustituido por la variante alélica <i>rps12-R43</i> . Esto hace que el fenotipo de este organismo sea Sm ^R . Esta cepa carece del plásmido pANS | - | (Matsuoka et al., 2001) |
| C11 | - | Estirpe evolucionada de Se7942 que contiene las mutaciones que aparecen en la Tabla 5-1 | Evolución experimental | Esta tesis |
| C12 | - | Estirpe evolucionada de Se7942 que contiene las mutaciones que aparecen en la Tabla 5-1 | Evolución experimental | Esta tesis |
| MSM1 | Sm ^R | Derivado de Se7942 que contiene la mutación <i>rps12-R43</i> de la cepa GRPS1 y por ello es resistente a Sm ^R | Transformación natural con el plásmido pMSM2 | Esta tesis |

4. Materiales y métodos

| Nombre | Fenotipo | Características | Construcción | Origen |
|--------|------------------|--|---|------------|
| MSM2 | Neo ^R | MSM1 que contiene aguas abajo del gen Synpcc7942_0894 la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM22 | Esta tesis |
| MSM3 | Sm ^R | MSM2 que contiene la mutación en el gen Synpcc7942_0894 (mutación #12, Tabla 5-1) de la estirpe C11 | Transformación natural con el plásmido pMSM24 | Esta tesis |
| MSM4 | Neo ^R | MSM1 que contiene aguas arriba del gen Synpcc7942_2114 la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM80 | Esta tesis |
| MSM5 | Sm ^R | MSM4 que contiene la mutación en el gen Synpcc7942_2114 (mutación #21, Tabla 5-1) de la estirpe C11 | Transformación natural con el plásmido pMSM82 | Esta tesis |
| MSM6 | Neo ^R | MSM1 que contiene aguas abajo del gen Synpcc7942_0918 la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM4 | Esta tesis |
| MSM7 | Sm ^R | MSM6 que contiene la mutación en el gen Synpcc7942_0918 (mutación #14, Tabla 5-1) de la estirpe C12 | Transformación natural con el plásmido pMSM10 | Esta tesis |
| MSM8 | Neo ^R | MSM5 que contiene aguas abajo del gen Synpcc7942_0894 la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM22 | Esta tesis |
| MSM9 | Sm ^R | MSM8 que contiene la mutación en el gen Synpcc7942_0894 (mutación #12, Tabla 5-1) de la estirpe C11 | Transformación natural con el plásmido pMSM24 | Esta tesis |
| MSM10 | Neo ^R | MSM7 que contiene aguas abajo del gen Synpcc7942_0918 la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM4 | Esta tesis |
| MSM11 | Sm ^R | MSM10 que contiene la mutación en el gen Synpcc7942_0918 (mutación #14, Tabla 5-1) de la estirpe C12 | Transformación natural con el plásmido pMSM10 | Esta tesis |
| MSM12 | Neo ^R | MSM7 que contiene aguas abajo del gen Synpcc7942_0894 la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM22 | Esta tesis |
| MSM13 | Sm ^R | MSM12 que contiene la mutación en el gen Synpcc7942_0894 (mutación #12, Tabla 5-1) de la estirpe C11 | Transformación natural con el plásmido pMSM24 | Esta tesis |
| MSM14 | Neo ^R | MSM13 que contiene aguas arriba del gen Synpcc7942_2242 la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM80 | Esta tesis |

| Nombre | Fenotipo | Características | Construcción | Origen |
|--------|------------------------------------|--|--|------------|
| MSM15 | Sm ^R | MSM14 que contiene la mutación en el gen <i>Synpcc7942_2114</i> (mutación #21, Tabla 5-1) de la estirpe C12 | Transformación natural con el plásmido pMSM82 | Esta tesis |
| MSM16 | Sp/Sm ^R | Derivado de Se7942 que contiene en el sitio neutral 1 (SN1) la construcción <i>PnrsB::desA</i> | Transformación natural con el plásmido pMSM14 | Esta tesis |
| MSM17 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>PnrsB::desA-desB</i> | Transformación natural con el plásmido pMSM16 | Esta tesis |
| MSM18 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>PnrsB::luxAB</i> | Transformación natural con el plásmido pMSM51 | Esta tesis |
| MSM19 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>PnrsB::fabB</i> | Transformación natural con el plásmido pMSM90 | Esta tesis |
| MSM20 | Neo ^R | Derivado de Se7942 que contiene integrado en el plásmido pANL la construcción <i>PnirA::fabB</i> | Conjugación de <i>E. coli</i> a Se7942 con el plásmido pMSM130 | Esta tesis |
| MSM21 | Neo ^R | Derivado de Se7942 que contiene integrado en el plásmido pANL la construcción <i>PnirA::luxAB</i> | Conjugación de <i>E. coli</i> a Se7942 con el plásmido pMSM132 | Esta tesis |
| MSM22 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>PnrsB::riboswitch-luxAB</i> | Transformación natural con el plásmido pMSM142 | Esta tesis |
| MSM23 | Neo ^R , Sm ^R | MSM1 en la que se ha delecionado en algunas copias del cromosoma el gen <i>Synpcc7942_1455</i> por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM182 | Esta tesis |
| MSM24 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>Ptrc::fabD</i> | Transformación natural con el plásmido pMSM196 | Esta tesis |
| MSM25 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>Ptrc::luxAB</i> | Transformación natural con el plásmido pMSM197 | Esta tesis |
| MSM26 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>Ptrc::desA-desB</i> | Transformación natural con el plásmido pMSM201 | Esta tesis |
| MSM27 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>Ptrc::fabB</i> | Transformación natural con el plásmido pMSM202 | Esta tesis |
| MSM28 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>Ptrc::fabH</i> | Transformación natural con el plásmido pMSM228 | Esta tesis |

4. Materiales y métodos

| Nombre | Fenotipo | Características | Construcción | Origen |
|--------|---|---|--|------------|
| MSM29 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>Ptrc::fabF</i> | Transformación natural con el plásmido pMSM234 | Esta tesis |
| MSM30 | Sp/Sm ^R | Derivado de Se7942 que contiene en SN1 la construcción <i>PnrsB::riboswitch-fabH</i> | Transformación natural con el plásmido pMSM236 | Esta tesis |
| MSM31 | Cm ^R | Derivado de Se7942 que contiene en el sitio neutral 2 (SN2) la construcción <i>Ptrc::fabD</i> | Transformación natural con el plásmido pMSM252 | Esta tesis |
| MSM32 | Cm ^R | Derivado de Se7942 que contiene en SN2 la construcción <i>Ptrc::fabF</i> | Transformación natural con el plásmido pMSM253 | Esta tesis |
| MSM33 | Cm ^R | Derivado de Se7942 que contiene en SN2 la construcción <i>PnrsB::fabB</i> | Transformación natural con el plásmido pMSM255 | Esta tesis |
| MSM34 | Neo ^R , Sm ^R , Sp ^R | MSM28 en la que se ha delecionado en algunas copias del cromosoma el gen <i>Synpcc7942_1455</i> por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM182 | Esta tesis |
| MSM35 | Neo ^R , Sm ^R , Sp ^R | MSM30 en la que se ha delecionado en algunas copias del cromosoma el gen <i>Synpcc7942_1455</i> por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM182 | Esta tesis |
| MSM36 | Neo ^R , Sm ^R , Sp ^R | MSM23 que contiene en SN1 la construcción <i>Ptrc::desA-desB</i> . | Transformación natural con el plásmido pMSM197 | Esta tesis |
| MSM38 | Sp/Sm ^R , Cm ^R | MSM26 que contiene en SN2 la construcción <i>Ptrc::fabD</i> | Transformación natural con el plásmido pMSM252 | Esta tesis |
| MSM39 | Sp/Sm ^R , Cm ^R | MSM26 que contiene en SN2 la construcción <i>Ptrc::fabF</i> . | Transformación natural con el plásmido pMSM253 | Esta tesis |
| MSM40 | Sp/Sm ^R , Cm ^R | MSM26 que contiene en SN2 la construcción <i>PnrsB::fabB</i> . | Transformación natural con el plásmido pMSM255 | Esta tesis |
| MSM41 | Sp/Sm ^R , Cm ^R , Neo ^R | MSM39 en la que se ha delecionado en algunas copias del cromosoma el gen <i>Synpcc7942_1455</i> por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM182 | Esta tesis |
| MSM42 | Sm ^R , Neo ^R | MSM1 la que se ha delecionado del cromosoma el gen <i>Synpcc7942_0918</i> por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM266 | Esta tesis |
| MSM43 | Sp/Sm ^R , Neo ^R | MSM26 la que se ha delecionado del cromosoma el gen <i>Synpcc7942_0918</i> por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM266 | Esta tesis |

| Nombre | Fenotipo | Características | Construcción | Origen |
|--------|---|--|--|------------|
| MSM44 | Sp/Sm ^R , Neo ^R | MSM29 la que se ha deleciónado del cromosoma el gen Synpcc7942_0918 por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM266 | Esta tesis |
| MSM45 | Sp/Sm ^R , Cm ^R , Neo ^R | MSM39 la que se ha deleciónado del cromosoma el gen Synpcc7942_0918 por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM266 | Esta tesis |
| MSM46 | Sp/Sm ^R , Neo ^R | MSM29 en la que se ha deleciónado en algunas copias del cromosoma el gen Synpcc7942_1455 por la construcción <i>rps12-nptI</i> | Transformación natural con el plásmido pMSM182 | Esta tesis |

4.2. Plásmidos

Los plásmidos utilizados en esta tesis se especifican en la **Tabla 4-3**, con sus características fenotípicas y una somera descripción de cada uno.

Tabla 4-3. Plásmidos.

| Nombre | Replicón | Fenotipo | Descripción | Construcción | Referencia |
|----------------------------|------------|---|--|--------------|---|
| pRL443 | RP4 | Ap ^R , Tc ^R | Derivado de RP4, Km ^S | - | (Elhai et al., 1997) |
| pSB1K3 | pMB1 | Km ^R | Vector de clonaje de "BioBricks" | - | http://parts.igem.org/Part:pSB1K3 |
| pSB1C3 | pMB1 | Cm ^R | Vector de clonaje de "BioBricks" | - | http://parts.igem.org/Part:pSB1C3 |
| pSyn_1/D-TOPO [®] | pMB1 | Sp/Sm ^R | Vector integrativo en SN1 de Se7942 que contiene el promotor <i>PnrsB</i> de <i>Synechocystis</i> sp. PCC 6803, inducible por níquel | - | GeneArt, Invitrogen |
| pUAGC280 | pMB1 | Ap ^R , Sp/Sm ^R | Vector integrativo en SN1 de Se7942 con el promotor <i>Ptrc</i> , inducible por IPTG, y el represor LacI | - | (Moronta-Barrios et al., 2013) |
| pAM1580 | ColE1 like | Ap ^R , Cm ^R | Vector integrativo en SN2 de Se7942 que contiene los orfs reporteros <i>luxAB</i> sin promotor | - | (Andersson et al., 2000) |
| pDEP6 | pMB1 | Km ^R | pSB1K3::(<i>rep</i> _{pANL} - <i>oriT</i> _{RP4}) | - | (Encinas et al., 2014) |
| pDEP52 | pMB1 | Km ^R , Cm ^R | pSB1C3::(<i>rps12-nptI</i>) | - | Tesis David Encinas Pisa |
| pSU2007::Tn <i>lux</i> | R388 | Tp ^R , Km ^R , Gm ^R | pSU2007:: <i>luxCDABE</i> | - | (Fernandez-Lopez et al., 2005) |

4. Materiales y métodos

| Nombre | Replicón | Fenotipo | Descripción | Construcción | Referencia |
|--------|----------|-----------------------------------|--|---|------------|
| pMSM1 | pMB1 | Km ^R , Cm ^R | pDEP52 con una región de 900 pb aguas abajo del gen <i>fadD</i> (Synpcc7942_0918), clonada a la derecha de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 1 y 2 para amplificar el inserto; oligos 3 y 4 para amplificar el vector | Esta tesis |
| pMSM2 | pMB1 | Cm ^R | pSB1C3 con una región de 1896 pb de la cepa GRPS1 de Se7942 que contiene el gen <i>rps12-R43</i> | Construido por "Biobrick". Inserto amplificado con oligos 5 y 6, clonado con enzimas de restricción <i>EcoRI</i> y <i>PstI</i> | Esta tesis |
| pMSM4 | pMB1 | Km ^R , Cm ^R | pMSM1 con el gen mutado <i>fadD</i> (Synpcc7942_0918) de la cepa C12, clonado a la izquierda de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 7 y 8 para amplificar el gen <i>fadD</i> ; oligos 9 y 10 para amplificar el vector | Esta tesis |
| pMSM10 | pMB1 | Cm ^R | pMSM4 con la eliminación de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 11 y 2 para amplificar el inserto; oligos 3 y 12 para amplificar el vector | Esta tesis |
| pMSM14 | pMB1 | Sp/Sm ^R | pSyn_1/D-TOPO [®] con el gen <i>desA</i> (SYNPCC7002_A2756) de Ss7002 | Gen <i>desA</i> amplificado con los oligos 13 y 14; clonado mediante enzimas de restricción <i>EcoRI</i> y <i>NdeI</i> | Esta tesis |
| pMSM16 | pMB1 | Sp/Sm ^R | pMSM14 con el gen <i>desB</i> (SYNPCC7002_A0159) de Ss7002 | Gen <i>desB</i> amplificado con los oligos 15 y 16; clonado mediante enzimas de restricción <i>NotI</i> y <i>NdeI</i> | Esta tesis |
| pMSM18 | pMB1 | Km ^R , Cm ^R | pDEP52 con una región de 501 pb aguas abajo del gen <i>aroK</i> (Synpcc7942_0894), clonada a la derecha de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 17 y 18 para amplificar el inserto; oligos 3 y 4 para amplificar el vector | Esta tesis |
| pMSM22 | pMB1 | Km ^R , Cm ^R | pMSM18 con el gen mutado <i>aroK</i> (Synpcc7942_0894) de la cepa C11, clonado a la izquierda de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 19 y 20 para amplificar el gen <i>aroK</i> ; oligos 9 y 10 para amplificar el vector | Esta tesis |
| pMSM24 | pMB1 | Cm ^R | pSB1C3 con el gen <i>aroK</i> (Synpcc7942_0894) mutado y una región de 501 pb aguas abajo de este gen | Construido por "Isothermal Assembly". Oligos 19 y 18 para amplificar el inserto; oligos 3 y 10 para amplificar el vector | Esta tesis |

| Nombre | Replicón | Fenotipo | Descripción | Construcción | Referencia |
|---------|----------|-----------------------------------|---|--|------------|
| pMSM51 | pMB1 | Sp/Sm ^R | pSyn_1/D-TOPO [®] con los genes <i>luxAB</i> de <i>Photorhabdus luminescens</i> tomados de pSU2007::Tnlux | Construido por "Isothermal Assembly". Oligos 21 y 22 para amplificar los genes <i>luxAB</i> ; oligos 23 y 24 para amplificar el vector | Esta tesis |
| pMSM76 | pMB1 | Km ^R , Cm ^R | pDEP52 con el gen mutado <i>sasA</i> (Synpcc7942_2242) de la cepa C11, clonado a la derecha de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 25 y 26 para amplificar el gen <i>sasA</i> ; oligos 3 y 4 para amplificar el vector | Esta tesis |
| pMSM80 | pMB1 | Km ^R , Cm ^R | pMSM76 con una región de 996 pb aguas arriba del gen <i>sasA</i> (Synpcc7942_2242), clonada a la izquierda de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 27 y 28 para amplificar el inserto; oligos 9 y 10 para amplificar el vector | Esta tesis |
| pMSM82 | pMB1 | Cm ^R | pMS80 con la eliminación de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 27 y 29 para amplificar el inserto; oligos 30 y 10 para amplificar el vector | Esta tesis |
| pMSM86 | pMB1 | Sp/Sm ^R | pSyn_1/D-TOPO [®] en el que se ha sustituido el promotor <i>PnrsB</i> por el promotor <i>PnirA</i> de Se7942 | Construido por "Isothermal Assembly". Oligos 31 y 32 para amplificar el promotor; oligos 33 y 34 para amplificar el vector | Esta tesis |
| pMSM90 | pMB1 | Sp/Sm ^R | pSyn_1/D-TOPO [®] con el gen <i>fabB</i> (SYNPCC7002_A1002) de Ss7002 | Gen <i>fabB</i> amplificado con los oligos 35 y 36; clonado mediante enzimas de restricción <i>EcoRI</i> y <i>NdeI</i> | Esta tesis |
| pMSM97 | pMB1 | Sp/Sm ^R | pMSM86 con el gen <i>fabB</i> (SYNPCC7002_A1002) de Ss7002 | Gen <i>fabB</i> amplificado con los oligos 35 y 36; clonado mediante enzimas de restricción <i>EcoRI</i> y <i>NdeI</i> | Esta tesis |
| pMSM125 | pMB1 | Sp/Sm ^R | pMSM86 con los genes <i>luxAB</i> de <i>Photorhabdus luminescens</i> tomados de pSU2007::Tnlux | Construido por "Isothermal Assembly". Oligos 37 y 38 para amplificar los genes <i>luxAB</i> ; oligos 23 y 24 para amplificar el vector | Esta tesis |

4. Materiales y métodos

| Nombre | Replicón | Fenotipo | Descripción | Construcción | Referencia |
|---------|----------|--------------------------------------|--|--|------------|
| pMSM130 | pMB1 | Km ^R | pDEP6 con la construcción <i>PnirA::fabB</i> de pMSM97 | Construido por "Biobrick". Fragmento <i>PnirA::fabB</i> amplificado con oligos 39 y 40, clonado con enzimas de restricción <i>XbaI</i> y <i>PstI</i> | Esta tesis |
| pMSM132 | pMB1 | Km ^R | pDEP6 con la construcción <i>PnirA::luxAB</i> de pMSM125 | Construido por "Isothermal Assembly". Oligos 41 y 42 para amplificar el fragmento <i>PnirA::luxAB</i> ; oligos 43 y 44 para amplificar el vector | Esta tesis |
| pMSM134 | pMB1 | Sp/Sm ^R | pSyn_1/D-TOPO [®] en el que se ha sustituido el RBS de este vector por el riboswitch y el RBS descrito en (Nakahira et al., 2013) | Construido por "QuickChange" modificado. Se amplificó el vector con los oligos 45 y 46 | Esta tesis |
| pMSM142 | pMB1 | Sp/Sm ^R | pMSM134 con los genes <i>luxAB</i> de <i>Photorhabdus luminescens</i> tomados de pSU2007::Tn <i>lux</i> | Construido por "Isothermal Assembly". Oligos 47 y 48 para amplificar los genes <i>luxAB</i> ; oligos 49 y 50 para amplificar el vector | Esta tesis |
| pMSM176 | pMB1 | Km ^R , Cm ^R | pDEP52 con el gen <i>plsX</i> (Synpcc7942_1454) de la cepa Se7942, clonado a la izquierda de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 51 y 52 para amplificar el gen <i>plsX</i> ; oligos 9 y 10 para amplificar el vector | Esta tesis |
| pMSM182 | pMB1 | Km ^R , Cm ^R | pMSM176 con el gen <i>fabD</i> (Synpcc7942_1456) de la cepa Se7942, clonado a la derecha de <i>rps12-nptI</i> | Construido por "Biobrick". Gen <i>fabD</i> amplificado con oligos 53 y 54, digerido con las enzimas de restricción <i>XbaI</i> y <i>NotI</i> . Vector pMSM176 digerido con <i>SpeI</i> y <i>NotI</i> | Esta tesis |
| pMSM196 | pMB1 | Ap ^R , Sp/Sm ^R | pUAGC280 con el gen <i>fabD</i> (Synpcc7942_1456) de Se7942 | Gen <i>fabD</i> amplificado con los oligos 55 y 56; clonado mediante enzimas de restricción <i>EcoRI</i> y <i>BamHI</i> | Esta tesis |
| pMSM197 | pMB1 | Ap ^R , Sp/Sm ^R | pUAGC280 en el que se han clonado los genes <i>luxAB</i> de <i>Photorhabdus luminescens</i> tomados de pSU2007::Tn <i>lux</i> | Construido por "Isothermal Assembly". Oligos 57 y 58 para amplificar los genes <i>luxAB</i> ; oligos 59 y 60 para amplificar el vector | Esta tesis |

| Nombre | Replicón | Fenotipo | Descripción | Construcción | Referencia |
|---------|------------|---|---|---|------------|
| pMSM201 | pMB1 | Ap ^R , Sp/Sm ^R | pUAGC280 en el que se han clonado los genes (<i>desA</i> (SYNPCC7002_A2756) y <i>desB</i> (SYNPCC7002_A0159) de Ss7002 | Gen <i>desA</i> amplificado con los oligos 61 y 14; gen <i>desB</i> amplificado con oligos 15 y 62. Clonados simultáneamente mediante enzimas de restricción. Gen <i>desA</i> clonado con <i>EcoRI</i> y <i>NotI</i> ; Gen <i>desB</i> clonado con <i>NotI</i> y <i>BamHI</i> | Esta tesis |
| pMSM202 | pMB1 | Ap ^R , Sp/Sm ^R | pUAGC280 con el gen <i>fabB</i> (SYNPCC7002_A1002) de Ss7002 | Gen <i>fabB</i> amplificado con los oligos 63 y 64; clonado mediante enzimas de restricción <i>EcoRI</i> y <i>BamHI</i> | Esta tesis |
| pMSM226 | ColE1 like | Ap ^R , Cm ^R | pAM1580 en el que se ha eliminado los genes <i>luxAB</i> | Construido por "QuickChange" modificado. Se amplificó el vector con los oligos 65 y 66 | Esta tesis |
| pMSM228 | pMB1 | Ap ^R , Sp/Sm ^R | pUAGC280 con el gen <i>fabH</i> (Synpcc7942_1455) de Se7942 | Gen <i>fabH</i> amplificado con los oligos 67 y 68; clonado mediante enzimas de restricción <i>EcoRI</i> y <i>SalI</i> | Esta tesis |
| pMSM232 | pMB1 | Sp/Sm ^R | pMSM134 en el que se han introducido nuevos sitios de corte para enzimas de restricción | Construido por "QuickChange" modificado. Se amplificó el vector con los oligos 69 y 46 | Esta tesis |
| pMSM234 | pMB1 | Ap ^R , Sp/Sm ^R | pUAG280 con el gen <i>fabF</i> (Synpcc7942_0537) de Se7942 | Gen <i>fabF</i> amplificado con los oligos 70 y 71; clonado mediante enzimas de restricción <i>EcoRI</i> y <i>BamHI</i> | Esta tesis |
| pMSM236 | pMB1 | Sp/Sm ^R | pMSM232 en el que se ha clonado el gen <i>fabH</i> (Synpcc7942_1455) de Se7942 | Gen <i>fabH</i> amplificado con los oligos 72 y 73; clonado mediante enzimas de restricción <i>KpnI</i> y <i>NdeI</i> | Esta tesis |
| pMSM240 | ColE1 like | Ap ^R , Cm ^R | pMSM226 en el que se ha introducido nuevos sitios de corte para enzimas de restricción | Construido por "QuickChange" modificado. Se amplificó el vector con los oligos 74 y 75 | Esta tesis |
| pMSM249 | ColE1 like | Ap ^R , Cm ^R | pMSM240 en el que se ha eliminado el terminador T4 que contenía | Digerido con <i>HindIII</i> y religado consigo mismo | Esta tesis |
| pMSM252 | ColE1 like | Ap ^R , Cm ^R | pMSM249 con la construcción <i>P_{trc}::fabD</i> de pMSM196 | Fragmento <i>P_{trc}::fabD</i> amplificado con oligos 76 y 77, clonado con enzimas de restricción <i>SpeI</i> y <i>KpnI</i> | Esta tesis |

4. Materiales y métodos

| Nombre | Replicón | Fenotipo | Descripción | Construcción | Referencia |
|---------|------------|-----------------------------------|---|---|------------|
| pMSM253 | ColE1 like | Ap ^R , Cm ^R | pMSM249 con la construcción <i>Ptrc::fabF</i> de pMSM234. | Fragmento <i>Ptrc::fabF</i> amplificado con oligos 76 y 78, clonado con enzimas de restricción <i>SpeI</i> y <i>KpnI</i> | Esta tesis |
| pMSM255 | ColE1 like | Ap ^R , Cm ^R | pMSM249 con la construcción <i>PnrsB::fabB</i> de pMSM90 | Fragmento <i>PnrsB::fabB</i> amplificado con oligos 79 y 80, clonado con enzimas de restricción <i>SpeI</i> y <i>KpnI</i> | Esta tesis |
| pMSM266 | pMB1 | Km ^R , Cm ^R | pMSM1 con una región de 933 pb aguas arriba del gen <i>fadD</i> (<i>Synpcc7942_0918</i>), clonada a la izquierda de <i>rps12-nptI</i> | Construido por "Isothermal Assembly". Oligos 9 y 10 para amplificar el inserto; oligos 81 y 82 para amplificar el vector | Esta tesis |

Todos los vectores y cepas construidos o modificados en esta tesis fueron verificados mediante secuenciación de ADN (servicio de secuenciación "Sanger" de la empresa "Stab Vida" llamado "You Tube It").

4.3. Oligonucleótidos

Los oligonucleótidos fueron sintetizados por Sigma-Aldrich o por IDT en función de la logística del instituto. Para los procedimientos de clonaje (Tabla 4-4) y secuenciación (Tabla 4-5) habituales se pidieron desalados.

Tabla 4-4. Oligonucleótidos utilizados para la construcción de los dispositivos genéticos.

| Nº | Nombre | Secuencia nucleotídica (5' → 3')* |
|----|--------------------------|--|
| 1 | H2 F | TAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAAATGACCCTCGGTACTCCTCTGCAGC |
| 2 | H2 R | GGAATTGCAGCTAAAACCGGCAGCATTGCAGGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTT |
| 3 | amplif vector H1 F | GCAGGCTTCTCGCTCACTGAC |
| 4 | amplif vector H1 R | CATTTGATGCTCGATGAGTTTTTCTAA |
| 5 | <i>rpsL</i> bioF | GTTTCTTCAATTGCGGCCGCTTCTAGAGGCGGTAGTGGGCAAAGCCTTG |
| 6 | <i>rpsL</i> bioR | GTTTCTTCTGCAGCGGCCGCTACTAGTACGCTTCCCCATGCTCTAGG |
| 7 | long chain F | CATTAACCTATAAAAATAGGCTTTCGCTAAGGATGATTTTCGGCTAGCGCTCCTCTGAGGGCCACAC |
| 8 | long chain R | GAAGACGCTAAATCCAGCTGAGCGTAGGAGCAACCCCTATTAACTCGCCGATTCAAACATCCCGTCG |
| 9 | amplif vector acil Coa F | TAGGGGTTGCTCCTACGCTC |
| 10 | amplif vector acil Coa R | GAAATCATCTTAGCGAAAGCCTATTTTTATAG |
| 11 | H2 sin colas F | AATGACCCTCGGTACTCCTCTGCAGC |
| 12 | Long chain H2 R | GATCGTCCGCTTAGCTGCAGAGGAGTACCGAGGGTCATTTAACTCGCCGATTCAAACATCCCGTCG |
| 13 | <i>desA</i> F | TTCAATTGCAAGGAGCCCTTACCATGACATCAGTTACTGTGCGC |
| 14 | <i>desA</i> R | TTCCATATGCGGCCGCGTACGTACACCCTTTAGCGTTTTTGATGTTGCG |

| Nº | Nombre | Secuencia nucleotídica (5' → 3')* |
|----|---------------------------|--|
| 15 | desB F | TACGCGGCCGCGAAGGAGCCCTTACCATGCAAAGTACAGTCCGCTCC |
| 16 | desB R | ATGCATATGGTACGTACACCCTTTATTTTTTCGGCTGGTAATACAAG |
| 17 | ZH2 sik F | TAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTCTAAGTTTAGAAA GTCGATCGCCAGAAGAGTTGCC |
| 18 | ZH2 sik R | AAAGGGCAGGGTGGTGACACCTTGCCCGTTTTTTTGC GGCGCAGCCA ATCTCAGTCCGGCAGAAG |
| 19 | shikimate kinasa F | CATTAACCTATAAAAAATAGGCTTTCGCTAAGGATGATTTCAAACCTGGC CTCATCTGCTCCAGGC |
| 20 | shikimate kinas R | GAAGACGCTAAATCCAGCTGAGCGTAGGAGCAACCCCTAGGGCGCA GCTTGGAGGAGACCCC |
| 21 | luxAB IA F | GAGGTGGTGTGAAGCTTGAATTCGAAGGAGCCCTTACCATGAAATT TGAAACTTTTTGCTTAC |
| 22 | luxAB IA R | CTGCCCGCAGGCACTCGATGCATATGGTACGTACACCCTTCTACATGT GGTACTTTTTAATATTATCATC |
| 23 | vector lux F | AAGGGTGTACGTACCATATGCATCG |
| 24 | vector lux R | GGTGAAGGGCTCCTTGAATTC |
| 25 | histidine kinase F | TAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTCTAAAGCCATCCTA CTTACCAACGAATTC |
| 26 | histidine kinase R | AAAGGGCAGGGTGGTGACACCTTGCCCGTTTTTTTGC GGCGCATCGC CAATTAAGAAATTATC |
| 27 | ZH1 histidine kinase F | CATTAACCTATAAAAAATAGGCTTTCGCTAAGGATGATTTCTGTGATCCGT TCTGCCACCCTG |
| 28 | ZH1 histidine kinase R | GAAGACGCTAAATCCAGCTGAGCGTAGGAGCAACCCCTACTAGCGG AACATACTGCTCACTGAGC |
| 29 | ZH1 HK homología con HK R | TACCGAAGCGATCGCGAATTCGTTGGTAAGTAGGATGGCTCTAGCGG AACATACTGCTCACTGAGC |
| 30 | HK sin colas F | AGCCATCCTACTTACCAACGAATTC |
| 31 | PnirA F | GTGAAAGGGCAGATCACCAAGGTAGTCGGCAAATAACCTACAGGTT AAAAACCATGCTAATGGCTC |
| 32 | PnirA R | CGTACACCCTTGGTGAAGGGCTCCTTGAATTCGAAGCTTACAAAGCA GCTTGCATACTCGGC |
| 33 | amplif vector promotor F | AAGCTTGAATTCGAAGGAGCC |
| 34 | amplif vector promotor R | TTATTTGCCGACTACCTTGGTGATCT |
| 35 | fabB F | TTCGAATTCGAAGGAGCCCTTACCATGGTGC GGGTAGTGATAACGG |
| 36 | fabB R | TTCCATATGGTACGTACACCCTTTCAAAAATATTTTTGAAGGGCGATC |
| 37 | luxAB IA nirA F | AGCTGCTTTGTAAGCTTCGAATTCGAAGGAGCCCTTACCATGAAATTT GGAAACTTTTTGCTTAC |
| 38 | luxAB IA nirA R | CTGCCCGCAGGCACTCGATGCATATGGTACGTACTGCAGCCCTTCTA CATGTGGTACTTTTTAATATT |
| 39 | PnirA-fabB F | GTTTCTTGAATTCGCGGCCGCTTCTAGAGACAGGTTAAAAACCATGC TAATGGCTC |
| 40 | PnirA-fabB R | GTTTCTTCTGCGCGGCCGCTACTAGTAAAGGCCAGCTTTTCGACT GAGC |
| 41 | PnirA-luxAB-pDEP6 IA F | CCGAAGCAGGGTTATGCAGCGGAAAAGCGCTGCTACTAGAACAGGTT AAAAACCATGCTAATGGCTC |
| 42 | PnirA-luxAB-pDEP6 IA R | TGACACCTTGCCCGTTTTTTTGC GGACTGCAGCGGCCGCAAGGCCCA GTCTTTCGACTGAGC |
| 43 | pDEP6 abierto F | GCGGCCGCTGCAGTCC |
| 44 | pDEP6 abierto R | TCTAGTAGCAGCGCTTTTCCGC |
| 45 | Riboswitch F | TGGCAGCACCTGCTAAGGAGGCAACAAGGGATCCACCATATGCATC GAGTGCCTGG |
| 46 | Riboswitch R | AGGGCATCAAGACGATGCTGGTATCACCCGAATTCGAAGCTTACACC ACCTC |
| 47 | luxAB RS IA F | GCCCTTGGCAGCACCTGCTAAGGAGGCAACAAGGGATCCATGAAATT TGAAACTTTTTGCTTACATAC |
| 48 | luxAB RS IA R | ACCACCGCCTACTGCCCGCAGGCACTCGATGCATATGGTTTAGGTAT ATTCCATGTGGTACTTCTAATATTATC |
| 49 | amplif vector RS F | ACCATATGCATCGAGTGCCTGG |

4. Materiales y métodos

| Nº | Nombre | Secuencia nucleotídica (5' → 3')* |
|----|----------------------|---|
| 50 | amplif vector RS R | GATCCCTTGTGCTCCTTAGCAG |
| 51 | CDS1319 F | <i>CATTAACCTATAAAAAATAGGCTTTTCGCTAAGGATGATTTTCATGACGCGT</i> GCGCGGATC |
| 52 | CDS1319 R | <i>GAAGACGCTAAATCCAGCTGAGCGTAGGAGCAACCCCTACTAGCTCA</i> CCACTGGCTCAGTACC |
| 53 | CDS1321 bioF | <i>GTTTCTTCGAATTCGCGGCCGCTTCTAGAGATGGCTAAAACGGTGTGG</i> GTGTTTC |
| 54 | CDS1321 bioR | <i>GTTTCTTCCTGCAGCGGCCGCTACTAGTATTAGACCGTGAGGTCTGCT</i> GCTGTTTC |
| 55 | fabD F | <i>CATGGAATTCATGGCTAAAACGGTGTGGGTG</i> |
| 56 | fabD R | <i>TAGAGGATCCTTAGACCGTGAGGTCTGCTGCTG</i> |
| 57 | luxAB Ptrc F | <i>AACAATTTACACAGGAAACAGACCATGGAATTCATGAAATTTGGAAA</i> CTTTTTGCTTAC |
| 58 | luxAB Ptrc R | <i>CCCCCCTCGAGGTCGACTCTAGAGGATCCCTACATGTGGTACTTTTTA</i> ATATTATCATC |
| 59 | pUAGC839 F | GGATCCTTAGAGTCGACCTCG |
| 60 | pUAGC839 R | GAATTCATGGTCTGTTTCCTG |
| 61 | desAB F | <i>CATGGAATTCATGACATCAGTTACTGTGCGC</i> |
| 62 | desAB R | <i>TAGAGGATCCTTATTTTTTCGGCTGGTAATACAAG</i> |
| 63 | fabB pUAGC839 F | <i>CATGGAATTCATGGTGCGGGTAGTGATAACGG</i> |
| 64 | fabB pUAGC839 R | <i>TAGAGGATCCTCAAAAATATTTTTGAAGGGCGATC</i> |
| 65 | pAM1580 sin luxAB F | TGGTGTCCCTGTTGATACGGG |
| 66 | pAM1580 sin luxAB R | GGTTACCCGTGACGGGCTACAC |
| 67 | fabH pUAGC839 F | <i>CATGGAATTCCTGACTCGACCTGGCGTTGG</i> |
| 68 | fabH pUAGC839 R | <i>CGAGGTCGACCTAAACCACCGTGCCCAAC</i> |
| 69 | Riboswitch F new | <i>TGGCAGCACCTGCTAAGGAGGCAACAAGGGTACCACCATATGATCG</i> AGTGCCCTGG |
| 70 | fabF F | <i>CATGGAATTCATGACTGAAACCGGACGCCAG</i> |
| 71 | fabF R | <i>TAGAGGATCCTTAGGGATGGAATTTCCGGAAGG</i> |
| 72 | fabH F | <i>TTCCGGTACCTTGACTCGACCTGGCGTTGG</i> |
| 73 | fabH R | <i>TTCCATATGCTAAACCACCGTGCCCAAC</i> |
| 74 | pAM1580 ER F | <i>TGGGGTACCGGCCGCGGCCGCAACATTTCTTATTTGTTGGTATTACG</i> ATAAG |
| 75 | pAM1580 ER R | <i>CCTACTAGTGGCCATTTAAATACGTTTAACTGATGCTGAAGGGG</i> |
| 76 | Ptrc-fabD NS2 F | <i>GGCCACTAGTTTGACAATTAATCATCCGGCTCG</i> |
| 77 | Ptrc-fabD NS2 R | <i>GGCCGGTACCTTAGACCGTGAGGTCTGCTGCTG</i> |
| 78 | Ptrc-fabF NS2 R | <i>GGCCGGTACCTTAGGGATGGAATTTCCGGAAGG</i> |
| 79 | PnrsB-fabB NS2 F | <i>GGCCACTAGTCGGTCTGATCTTAGCGGGGG</i> |
| 80 | PnrsB-fabB NS2 R | <i>GGCCGGTACCAAGGCCAGTCTTTGACTGAGC</i> |
| 81 | ZH1 delección fadD F | <i>CATTAACCTATAAAAAATAGGCTTTTCGCTAAGGATGATTTTCAGCTTGGGT</i> CGATGCCCCG |
| 82 | ZH1 delección fadD R | <i>GAAGACGCTAAATCCAGCTGAGCGTAGGAGCAACCCCTATCGGTCTA</i> CTCTCCTCAGCCAGC |

*En cursiva, las colas empleadas para el método de clonaje "Isothermal Assembly" con homología al vector donde se van a clonar; en negrita, la región que hibrida con la secuencia a amplificar; subrayado, las dianas para enzimas de restricción.

Tabla 4-5. Oligonucleótidos utilizados para comprobación y secuenciación.

| Nº | Nombre | Secuencia nucleotídica (5' → 3') |
|----|----------------------|----------------------------------|
| 83 | compr mut rpsL F | CCAGCAGCTTGCAGTGTGACG |
| 84 | compr mut rpsL R | GCTGCCAACGTCGTGATAATG |
| 85 | compr mut ac graso F | CCAAGGCACGTTGCTGGCAG |
| 86 | compr mut ac graso R | GCGTTGACTCAAGCCAAAGAAGG |

| Nº | Nombre | Secuencia nucleotídica (5' → 3') |
|-----|------------------------------|-------------------------------------|
| 87 | compr mut sik kinasa F | CGGAGATCGGCTTGGGCATAG |
| 88 | compr integr sik kinasa R | CAGCTCTAGCCCCGAGCAAG |
| 89 | compr mut pilA F | GGCAGGAATACAAAAGAGTGCAC |
| 90 | compr mut pilA R | AAGCCCAATGCGGACGGAG |
| 91 | compr mut 7 hyp prot F | GAGCCAGGAACCGTTTGGTAGAG |
| 92 | compr mut 7 hyp prot R | CAGTGCCAGCGTTCTTCCTC |
| 93 | compr mut histidine kinase F | GCTCTCGTCGCCACGACC |
| 94 | compr mut histidine kinase R | CGCCACAGAGCAGGGTTTAC |
| 95 | compr mut 23 hyp prot F | CGCCGCCGCACTATTGAGC |
| 96 | compr mut 23 hyp prot R | GGCAGCCGAAGGTGGTGATG |
| 97 | bioF | TGCCACTTGACGTCTAAGAA |
| 98 | bioR | ATTACCGCCTTTGAGTGAGC |
| 99 | compr integr sik kinasa F | GGACTCCGCCAGCAGCTTCC |
| 100 | syn F | AGTCGGCAAATAACCTCGG |
| 101 | syn R | GCCAGGCATCAAATAAAACG |
| 102 | compr integr 1F | GCTTGCCTTCTATGGTTCGG |
| 103 | compr integr 1R | CCGCTCCTGCTGCACAGG |
| 104 | compr Ptrc F | GTGGACCGCTTGTGCAACTC |
| 105 | compr Ptrc R | GCTTGGCAGACCGTGGTG |
| 106 | compr pAM1580 F | CGCTTCCCACGCTGAGAGG |
| 107 | compr pAM1580 R | CGCCATAAACTGCCAGGCATC |
| 108 | compr pAM1580 R new | TGCTGGGTAGTTCTCCGCTGC |
| 109 | presencia ZH1 R | CAAAAAGCCCAACTCCGCC |
| 110 | presencia ZH2 F | CCGATACCAGGATCTTGCCATCC |
| 111 | compr CDS1319 F new | GTCGGCTTGGTCAACATCGGC |
| 112 | compr CDS1321 R | CACTACGCCTTCGGTGTGGC |
| 113 | Cm end | CTCGATATCAAATTACGCCCCGCTGCC |
| 114 | Cm pro | AGTTGATCGGGCAGTAAGAGGTTCCAACCTTCACC |
| 115 | syn 2 F | GCAGATCAGTTGGAAGAATTTGTC |

4.4. Técnicas de biología molecular

4.4.1. Extracción y purificación de ADN de *E. coli*

- ADN plasmídico de cultivo bacteriano. Se utilizó “GeneJET plasmid miniprep kit” de Fermentas, siguiendo el protocolo del fabricante.
- Purificación de ADN de gel agarosa. Tras cortar una banda de un gel de electroforesis se utilizó “GeneJET gel extraction kit” de Fermentas, siguiendo el protocolo del fabricante.

- Purificación de ADN de una reacción enzimática o PCR. Se utilizó “GeneJET PCR purification kit” de Fermentas, siguiendo el protocolo del fabricante.

4.4.2. Cuantificación del ADN

El ADN se cuantificó por medidas de absorbancia a 260 nm utilizando el espectrofotómetro “NanoDrop 2000” de Thermo Scientific. Siempre se comprobó la pureza del ADN extraído teniendo en cuenta la relación 260/280 nm ($\sim 1,8$) y 260/230 nm (~ 2), ya que puede condicionar reacciones posteriores.

4.4.3. Reacción en cadena de la polimerasa (PCR)

En la amplificación de moléculas de ADN destinadas al clonaje es de capital importancia que la copia de dichas moléculas sea lo más fiel posible. Por ello, se utilizaron polimerasas comerciales de alta fidelidad (ADN polimerasa “Vent” de BioLabs y “Phusion High-Fidelity DNA Polymerase” de Thermo Scientific), para amplificar fragmentos utilizados en clonación o utilizados en secuenciación tipo “Sanger”.

Las PCRs empleadas para la comprobación de clones se realizaron con una ADN polimerasa obtenida en nuestro laboratorio cuya tasa de error es similar a la “Biotaq™ DNA Polymerase” de Bioline ($1 - 20 \cdot 10^{-5}$ errores/pb). La mezcla de reacción con esta polimerasa fue: 5 μ l de tampón para PCR 10x (500 mM KCl, 100 mM TrisHCl, 1 % Tritón, pH 9), 2 μ l de MgCl₂ 50 mM, 1 μ l de mezcla de nucleótidos 10 mM total, 0,5 μ l de cada cebador 100 μ M, 0,5 μ l de polimerasa, la colonia bacteriana o el ADN a analizar y agua Milli-Q® hasta un volumen final de 50 μ l.

Tres termocicladores fueron utilizados durante este trabajo: “UNO II” de Biometra, “C1000 Thermal Cycler” de BioRad y “TProfessional Thermocycler” de Biometra. En general, el programa utilizado, con modificaciones en función de la polimerasa usada y la molécula a amplificar, fue:

- i) Un paso inicial de desnaturalización de 5 min a 95 °C,
- ii) 30 iteraciones de 3 pasos: 30 s a 95 °C, 30 s a la temperatura de hibridación ADN/cebador

y 1 min por kb de ADN amplificado a 72 °C.

iii) Un último paso de elongación 10 min a 72 °C.

iv) Mantenimiento de la reacción a 4 °C.

Sin embargo, cuando utilizamos la polimerasa “Phusion High-Fidelity DNA Polymerase”, el programa utilizado fue el siguiente:

i) Un paso inicial de desnaturalización de 30 s a 98 °C,

ii) 30 iteraciones de 3 pasos: 5-10 s a 98 °C, 10 - 30 s a la temperatura de hibridación ADN/cebador y 15 - 30 s por kb de ADN amplificado a 72 °C.

iii) Un último paso de elongación 5 - 10 min a 72 °C.

iv) Mantenimiento de la reacción a 4 °C.

4.4.4. Electroforesis de ADN en gel de agarosa

Dependiendo de las longitudes de las moléculas a discernir se emplearon geles de agarosa al 1 o 2 %. La agarosa se disolvió en tampón TBE 0,5x (Tris-HCl 45 mM, ácido bórico 45 mM, EDTA 0,5 mM pH 8,2). Tras hervir y disolver la mezcla y atemperar la solución hasta 60 - 70 °C, se añadieron 5 µl de “SafeView Classic” (ABM) por cada 100 ml de agarosa. Las muestras de ADN se mezclaron en una proporción 5:1 con tampón de carga 6x (Azul de bromofenol 0,25 % (p/v), glicerol 30 % (v/v) en tampón TBE 0,5x). En función de los tamaños moleculares se emplearon distintos marcadores moleculares: “HyperLadder I” (BioLabs) para muestras de hasta 10 kb, y “Generuler 100 pb DNA ladder plus” (Fermentas) para tamaños entre 100 y 3.000 pb. La electroforesis se realizó en cubetas horizontales (BioRad) sumergiendo el gel en tampón TBE 0,5x, y aplicando voltajes entre 100 - 120 V. Los geles se visualizaron en un transiluminador UV “Gel Doc 2000” (BioRad) y se analizaron mediante el programa “Quantity One” (BioRad).

4.4.5. Digestión de ADN mediante enzimas de restricción

Se utilizaron enzimas de restricción de Thermo Scientific. En un volumen de 20 µl, se mezclan 2 µl del tampón de reacción 10x, 1 µl de enzima, 500 ng de ADN y agua Milli-Q® hasta

completar. La incubación se realizó a 37 °C durante 1 h, salvo para enzimas “FastDigest” que necesitan 10 min. La inactivación de la enzima se realizó por calor en condiciones dependientes de la enzima, la mayoría durante 10 min a 65 °C.

4.4.6. Fosforilación y desfosforilación

La proporción de moléculas de vector recircularizadas en una ligación tras un proceso de digestión disminuye si eliminamos los grupos fosfato terminales. Por ello, el ADN de los vectores se defosforiló utilizando fosfatasas alcalinas de Thermo Scientific “Shrimp alkaline phosphatase” y “FastAP thermostable alkaline phosphatase”. Para una reacción de 20 µl, 1 pmol de ADN, 2 µl de tampón, 1 µl de enzima y agua Milli-Q®, incubada 30 min a 37 °C (10 min en el caso de “FastAP”).

La ligación del vector desfosforilado con el inserto tendrá éxito siempre y cuando este tenga sus extremos fosforilados. Por eso, cuando las moléculas procedían directamente de una PCR, se procedió a la fosforilación de los mismos. Para ello se utilizó “T4 polynucleotide kinase” de Fermentas: 1 - 50 pmol de ADN, 2 µl de tampón, 2 µl de ATP 10 mM, 1 µl de enzima y agua Milli-Q hasta un volumen de 20 µl incubado durante 20 min a 37 °C, e inactivando a 75 °C durante 10 min. En insertos procedentes de una digestión este paso no es necesario.

4.4.7. Reacciones de ligación

Las reacciones de ligación se realizaron con “T4 DNA ligase” de Fermentas: 20 - 100 ng de vector, inserto en una relación molar de entre 1:1 y 5:1 frente al vector, 1 µl de ligasa, 2 µl del tampón y agua hasta un volumen total de 20 µl. Si la ligación era entre extremos romos se añadió, además, 2 µl de polietilenglicol 4000 al 50 %. Si bien se indica realizar la incubación 10 min a 22 °C, generalmente se incubó durante la noche. Se inactivó la enzima a 65 °C durante 10 min.

4.4.8. Métodos de clonaje

4.4.8.1. Ensamblaje isotérmico (“Isothermal assembly”)

El ensamblaje isotérmico es una reacción única de múltiples moléculas de ADN con extremos coincidentes (Gibson, 2011).

El método se basa en la siguiente sucesión de reacciones (**Figura 4-1**). La exonucleasa T5 elimina nucleótidos del extremo 5' de dos hebras dobles de ADN con secuencias terminales solapantes. De esta manera, estas dos moléculas presentan extremos terminales complementarios que se unen. Entonces la ADN polimerasa “Phusion High-Fidelity DNA Polymerase” completa los huecos y la ligasa Taq une los cortes. Toda la reacción se realiza a 50 °C; las 3 enzimas funcionan correctamente pero la exonucleasa se va inactivando a esa temperatura lo que evita que degrade más ADN del necesario.

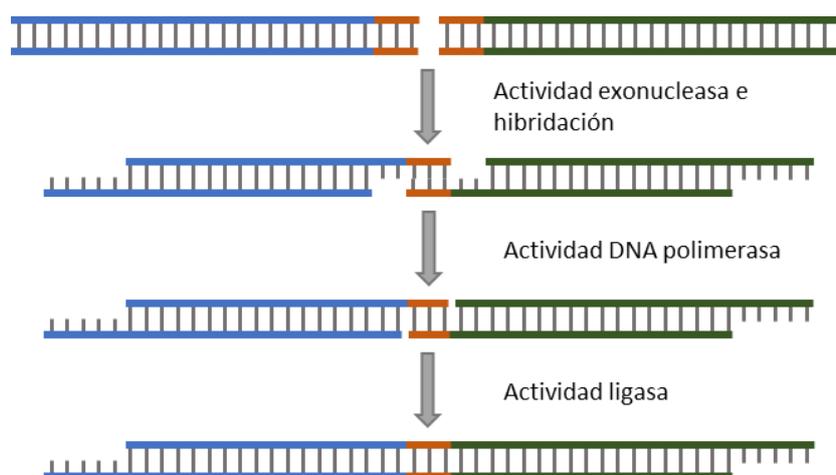


Figura 4-1. Proceso de ensamblaje isotérmico de dos fragmentos de ADN. La exonucleasa elimina nucleótidos en la dirección 5' -> 3', dejando porciones de cadena sencilla, lo que hace que hibriden los extremos homólogos entre ambos fragmentos (en naranja). La ADN polimerasa añade nucleótidos en los extremos 3' libres y la ligasa sella las mellas.

El procedimiento experimental partió de PCRs para conseguir los fragmentos con los extremos cohesivos. Los amplificados se corrieron en gel y se purificaron las bandas adecuadas.

Se prepararon 6 ml de tampón mezclando 3 ml de 1M Tris-HCl pH 7,5, 150 µl de 2 M MgCl₂, 60 µl 100 mM dGTP, 60 µl 100 mM dCTP, 60 µl 100 mM dTTP, 60 µl 100 mM dATP, 300 µl 1M DTT, 1,5 g PEG 8000, 300 µl 100 mM NAD y agua Milli-Q® hasta el volumen total. Se dividió en 18 alícuotas de 320 µl, siendo congeladas a -20 °C para su posterior uso. Utilizando una de estas

alícuotas se preparó la mezcla de reacción añadiendo 1,2 µl de exonucleasa T5 (Epicentre), 20 µl de polimerasa “Phusion High-Fidelity DNA Polymerase” (Thermo Scientific), 160 µl de “Taq DNA ligase” (BioLabs) y 700 µl de agua Milli-Q®. Se prepararon alícuotas de 15 µl en tubos de PCR y se guardaron a -20 °C (siendo recomendable usarlas antes de un mes).

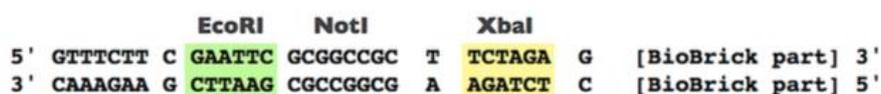
El siguiente paso fue mezclar en un volumen de 5 µl los distintos fragmentos a la misma concentración molar. Si se estaba construyendo un vector se utilizó como referencia el uso de aproximadamente 100 ng del fragmento portador del origen de replicación y/o la resistencia antibiótica. Estos 5 µl se añadieron a una de las alícuotas de mezcla de reacción y se incubó la mezcla durante 1 h a 50 °C.

Por último, el vector resultante se introdujo en la cepa deseada mediante electroporación, microdializándolo o purificándolo previamente.

4.4.8.2. “Biobricks”

El sistema “Biobrick” (Shetty et al., 2008) usa un vector que contiene puntos de corte para cinco enzimas de restricción, divididos en prefijos (*EcoRI*, *NotI* y *XbaI*) y sufijos (*SpeI*, *NotI*, *PstI*) (Figura 4-2).

PREFIJO ESTÁNDAR



SUFIJO ESTÁNDAR

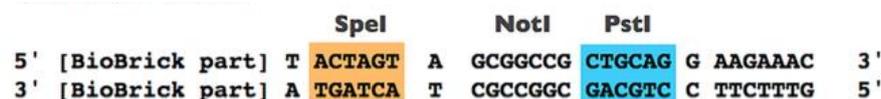


Figura 4-2. Prefijo y sufijo estándar presentes en los Biobrick. En la parte superior de la imagen se muestran las dianas de las enzimas de restricción que conforman el prefijo. En verde aparece la diana para *EcoRI*, en blanco la diana para *NotI* y en amarillo la diana para *XbaI*. En la parte inferior se observan las dianas de las enzimas de restricción presentes en el sufijo. En naranja se muestra la diana para *SpeI*, en blanco la diana para *NotI* y en azul la diana para *PstI*.

El segmento de ADN a clonar se amplifica mediante PCR usando un oligonucleótido directo que contiene los sitios de restricción del prefijo y un oligonucleótido reverso que contiene los del sufijo. Este fragmento se clona en un vector “Biobrick” cortando, tanto el inserto como el vector, con una enzima de restricción del sufijo y otra del prefijo. También puede

clonarse en un vector “Biobrick” que contenga ya un fragmento. Si se quiere clonar a la izquierda del fragmento, es necesario cortar el vector con *EcoRI* y *XbaI* e inserto con *EcoRI* y *SpeI*, y si se quiere clonar a la derecha, habrá que cortar el vector con *SpeI* y *PstI* e inserto con *XbaI* y *SpeI*. El corte que deja la enzima *XbaI*, es compatible con el que deja *SpeI*, pero su unión no da lugar a un nuevo punto de corte para una enzima de restricción. Esto permitiría clonar un fragmento, pero por otro lado impediría su futura eliminación de la construcción. Un esquema de este proceso se representa en la **Figura 4-3**.

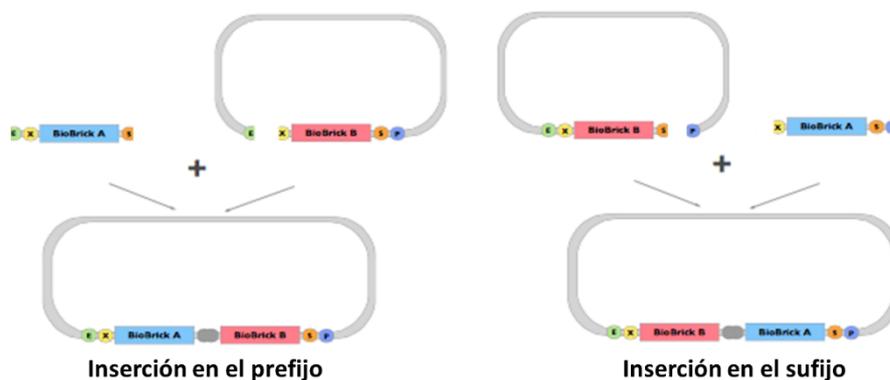


Figura 4-3. Construcción de plásmidos mediante el método Biobrick. En círculos se representan las enzimas de restricción presentes en el vector, donde E representa a *EcoRI*, X representa *XbaI*, S representa *SpeI* y P representa *PstI*. El cuadrado redondeado gris representa la cicatriz que se forma entre la unión del sitio de corte de *XbaI* con *SpeI*. Los rectángulos rosas y azules representan las partes que componen estas construcciones.

4.4.8.3. “QuickChange” modificado

Cuando era necesario realizar modificaciones en un plásmido recombinante (cambio de promotor, adición de dianas para enzimas de restricción o adición de un riboswitch), se adaptó la metodología “QuickChange” de Stratagene (Bok and Keller, 2012). Este método permite cambiar varios de nucleótidos de la secuencia de un plásmido, para lo que sólo es necesaria una PCR de la que se obtiene directamente una molécula circular de DNA. Por tanto, es imprescindible el uso de una ADN polimerasa que no provoque desplazamiento de cadena. Por ello, se modificó dicho método, generando una molécula lineal de DNA con extremos romos que posteriormente fue ligada y convertida en circular (**Figura 4-4**).

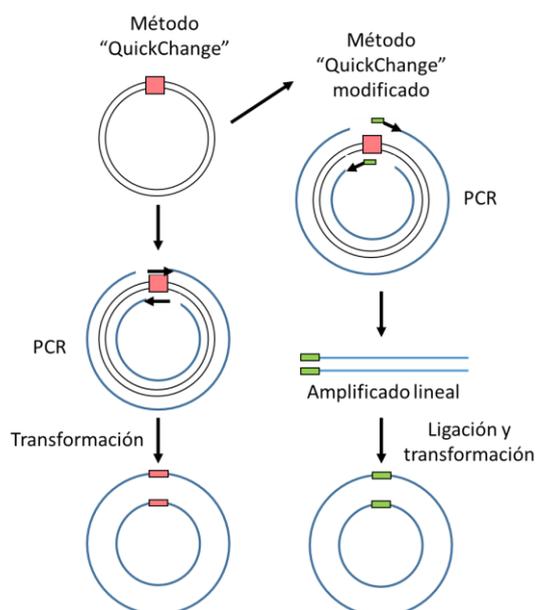


Figura 4-4. Mutagénesis dirigida mediante el método QuickChange. En el método tradicional “QuickChange” (izquierda) la mutación deseada (rectángulo rojo) se incorpora en dos cebadores complementarios. Estos se usan en una reacción de amplificación que genera una molécula circular con dos mellas, una en cada cadena y en posiciones diferentes. Las mellas son reparadas por la maquinaria celular. En la modificación del método (derecha de la figura), cada cebador contiene parte de la región a insertar (rectángulos verdes) y se genera un producto lineal que se liga *in vitro* y se transforma.

El quid del método radica en el diseño de los cebadores para realizar una PCR con el cambio que se quiere realizar en su secuencia. Primeramente, se fosforilaron los oligonucleótidos siguiendo el protocolo descrito en el apartado 4.4.6 y luego se realizó una PCR con estos cebadores fosforilados. Se limpiaron los fragmentos de PCR usando los kits descritos en el apartado 4.4.1 y se trataron con *DpnI* para eliminar el ADN molde. La molécula lineal resultante se sometió a una reacción de ligación siguiendo el protocolo descrito en el apartado 4.4.7. En dicha reacción de ligación, fue necesario usar bajas concentraciones de ADN con el fin de maximizar la ligación intramolecular. Finalmente, la molécula resultante se introdujo mediante electroporación en la cepa deseada siguiendo el procedimiento descrito en el apartado 4.8.3.

4.5. Aislamiento de ADN genómico total de Se7942 para su secuenciación

La obtención de ADN total de Se7942 se realizó siguiendo el protocolo descrito por (Tamagnini et al., 1997). Se resuspendieron 50 ml de un cultivo saturado de Se7942 en 500 μ l de Tris-HCl 50 mM, pH 8,0, 10 mM de EDTA. A esta mezcla se añadieron 0,5 g de bolas de vidrio

de diámetro 0,6 µm, 25 µl de SDS y 500 µl de fenol-cloroformo (1:1, v/v). Las células se lisaron usando un homogeneizador ("Fast prep-24", M.P Biomedical), mediante tres ciclos de 30 s a 5.800 rpm, intercalando incubaciones de 2 min en hielo. Posteriormente, la muestra resultante se centrifugó a 14.000 rpm durante 15 min. Se extrajo la fase superior acuosa y se lavó dos veces con el mismo volumen de cloroformo. A continuación, se añadió 1/10 del volumen de acetato sódico 3 M (pH 5,2) y 2,5 volúmenes de etanol al 100 % para precipitar el ADN. La mezcla se guardó a -20 °C durante al menos 5 h. Finalmente, se centrifugó a 14.000 rpm durante 30 min a 4 °C, se lavó el precipitado con etanol al 70 % y se dejó secar. Cuando el ADN se utilizó como molde para PCR, se resuspendió en 150 µl de agua Milli-Q®. Cuando el ADN se utilizó para secuenciación, se resuspendió en 100 µl de buffer TE (10 mM Tris, 1 mM EDTA, pH 8,0) y se trató con 1 µl de ARNasa A (10 mg/ml, Thermo Scientific) durante 1 h a 37 °C. Después se realizó una limpieza del ADN con el kit de purificación "DNeasy Blood & Tissue" (Qiagen) y el ADN se resuspendió, finalmente, en 150 µl de TE.

4.6. Aislamiento de ARN de Se7942 para su análisis transcriptómico

La obtención de ARN de Se7942 se hizo siguiendo el protocolo descrito para *Brucella abortus* (Sangari et al., 2010). Se recogieron 50 ml de un cultivo de Se7942 a una DO₇₂₀ de 0,8 y se trataron con 2 vol de la solución "RNAprotect Cell Reagent" (Qiagen). Se mezclaron usando un vórtex y se dejaron 5 min a temperatura ambiente. A continuación, se centrifugaron a 4,500 rpm durante 15 min a 4 °C y se eliminó el sobrenadante por absorción con una pipeta (tras este paso, se puede congelar la muestra). Se resuspendió el pellet en 300 µl de TE y 45 µl de lisozima (20 mg/ml, Roche). Se mezclaron por agitación con vórtex a temperatura ambiente en ciclos de 10 s cada 2 min durante 10 - 15 min, periodo en el que tenía lugar la incubación con la lisozima añadida. Después, se añadieron 15 µl de Zwittergente (Calbiochem) al 10 % en agua libre de ADNasa/ARNasa (Ambion) y 6 µl de proteinasa K (20 mg/ml, Roche). Se incubó la mezcla durante 1 h a 37 °C. Tras este tiempo, se realizó una limpieza del ARN con el kit de purificación "RNeasy Mini Kit" (Qiagen) que se eluyó finalmente en 30 µl de agua libre de ADNasa/ARNasa. Posteriormente, se realizó un paso adicional para eliminar el posible ADN que pudiese quedar en la extracción. Para ello, se utilizó el kit "TURBO DNA-free" (Ambion). Una vez tratada la muestra fue congelada a -80 °C hasta su uso.

4.7. Secuenciación de ADN y ARN total de Se7942

4.7.1. Secuenciación de ADN total

La concentración de ADN total en las muestras a secuenciar se midió utilizando “Nanodrop 2000” (Thermo Scientific) y “Qubit 2.0 Fluorometer” (Invitrogen). Para su secuenciación, se fragmentó 1 µg de ADN con el ultrasonicador “Bioruptor Nex gen” (Diagenode), usando un programa de 20 ciclos de 30 s a 4 °C. Se comprobó la calidad de la muestra utilizando un “Bioanalyzer 2100” (Agilent Technologies) con el kit “DNA 1000” (Agilent Technologies). A continuación, se prepararon y validaron las librerías de ADN utilizando el kit “TruSeq DNA Sample Preparation” y “Bioanalyzer” junto con el chip “High Sensitivity DNA” (Agilent Technologies). Las librerías se prepararon con un tamaño de inserto de 400 pb y se colocaron adaptadores para realizar lecturas de secuencia pareadas (“paired-end”) de 101 bases de longitud (reads). La cuantificación de las librerías se realizó mediante Real Time PCR utilizando “StepOnePlus Real-Time PCR System” (Applied Biosystems). Se utilizó el kit “TruSeq PE Cluster v5-CS-GA” (Illumina) para preparar el “Flow Cell”. Finalmente, las muestras fueron secuenciadas por la empresa “BGI” utilizando “Hiseq 2500 Sequencing System” de “Illumina”. La calidad de los reads fue analizada con “PRINSEQ” (Schmieder and Edwards, 2011). A continuación, estos reads fueron utilizados en “Bowtie2” (Langmead and Salzberg, 2012) para indexar y mapear contra el genoma de referencia (CP000100.1). Seguidamente, se utilizó “SAMtools” (Li et al., 2009) para generar un archivo “BAM”, el cual fue visualizado utilizando el software “Artemis” (Rutherford et al., 2000).

4.7.2. Secuenciación de ARN total

Se analizó la calidad del ARN total obtenido utilizando el equipo “Experion Automated Electrophoresis System” (Bio-Rad), junto con el kit de “Experion RNA Analysis” (Bio-Rad). A continuación, se sometió al ARN a un proceso de enriquecimiento de ARNm, mediante el kit “MICROBE express” (Invitrogen). El éxito de este proceso se comprobó utilizando nuevamente el equipo “Experion Automated Electrophoresis System” junto con el kit de “Experion RNA Analysis”. A continuación, se prepararon las librerías pertinentes, con un tamaño de read de 37 pb, utilizando el kit “ScriptSeq mRNA-Seq Library Preparation Kit Illumina-Compatible”

(Epicentre). La calidad de estas librerías se analizó mediante “Agilent 2100 Bioanalyzer”. Seguidamente, se cuantificó las librerías mediante Real Time PCR utilizando “StepOnePlus Real-Time PCR System” (Applied Biosystems). A continuación, se utilizó el kit “TruSeq PE Cluster v5-CS-GA” (Illumina) para preparar el “Flow Cell”. Finalmente, las muestras fueron secuenciadas por la empresa “BGI” utilizando “Hiseq 2500 Sequencing System” de “Illumina”. Los datos obtenidos de la secuenciación fueron cargados en Bowtie 2 (Langmead and Salzberg, 2012) para indexar y mapear contra el genoma de referencia (CP000101). Seguidamente, se utilizó SAMtools (Li et al., 2009) para generar un archivo “BAM”, el cual fue visualizado utilizando el software “Artemis” (Rutherford et al., 2000).

4.8. Técnicas microbiológicas

4.8.1. Medios de cultivo

Para el cultivo de *E. coli* en medio líquido se utilizó el medio Luria-Bertani (LB) (triptona 10 g/L, extracto de levadura 5 g/L, NaCl 5 g/L, Pronadisa) (Sambrook, 1989). Este medio fue suplementado con 1,5 % agar (p/v) para los cultivos sólidos. En el caso de crecimiento en medios selectivos, se añadieron antibióticos a las concentraciones indicadas: ácido nalidíxico (Nx, Sigma-Aldrich) 20 µg/ml, estreptomina (Sm, Apollo) 300 µg/ml, ampicilina (Ap, Apollo) 100 µg/ml, kanamicina (Km, Sigma-Aldrich) 50 µg/ml y cloranfenicol (Cm, Sigma-Aldrich) 25 µg/ml. En el caso de la cepa β 2150, se añadió ácido diaminopimélico (DAP) 30 mM al LB. Los cultivos líquidos se crecieron a 37 °C en agitación (120-150 rpm) en matraces de cristal, y el crecimiento en sólido se realizó en placas de Petri de 90x14 mm en incubadores a 37 °C.

Para la conservación de las cepas, se centrifugaron los cultivos en fase estacionaria y se resuspendieron en una solución de glicerol-peptona (50 % de glicerol (v/v) y 0,75 % de peptona (p/v)). Las cepas se guardaron a -20 °C ó a -80 °C.

El cultivo de la cianobacteria Se7942 se realiza normalmente en medio BG11 (Rippka et al., 1979), a 30 °C, con iluminación de luz blanca fría constante (60 µmoles de fotones·m⁻²·s⁻¹) y agitación orbital (100 rpm) en caso de cultivos líquidos. En otras ocasiones, los cultivos se crecieron en el incubador “Multi-cultivator MC 1000-OD” (Photon Systems Instruments) a 30 °C con burbujeo constante de CO₂ al 3 %. La fuente de iluminación en este caso son LEDs del tipo “Cool White” (blanca fría) y la intensidad utilizada varió en función de las necesidades de cada

experimento. El medio de cultivo BG11 se prepara a partir de un concentrado de sales 100x, una solución 1 M de K_2HPO_4 y $NaNO_3$. Para 1 L de concentrado de sales 100x se necesita 7,5 g de $MgSO_4 \cdot 7H_2O$, 3,6 g de $CaCl_2 \cdot 2H_2O$, 0,6 g de ácido cítrico, 9,3 mg de Na_2EDTA , 0,6 g de citrato férrico-amónico, 2 g de Na_2CO_3 , 286 mg de H_3BO_3 , 181 mg de $MnCl_2 \cdot 3H_2O$, 22,2 mg de $ZnSO_4 \cdot 7H_2O$, 39 mg de $Na_2MoO_4 \cdot 2H_2O$, 7,9 mg de $CuSO_4 \cdot 5H_2O$ y 4,99 mg de $CoCl_2 \cdot 6H_2O$. La mezcla se autoclavó 20 min a 121 °C y se conservó en la nevera protegida de la luz, ya que es fotosensible.

Para la obtención de 1 L de medio BG11 líquido se emplearon 10 ml de concentrado de sales, 200 µl de una solución 1 M de K_2HPO_4 , 1,5 g de $NaNO_3$ y 990 ml de agua Elix®. La mezcla se autoclavó 20 min a 121 °C previo uso.

Para la obtención de 1 L de medio BG11_o líquido se emplearon 10 ml de concentrado de sales, 200 µl de una solución 1 M de K_2HPO_4 y 990 ml de agua Elix®. La mezcla se autoclavó 20 min a 121 °C previo uso.

Para la obtención de 1 L de medio BG11 sólido se emplearon 2 botellas. En la primera se introdujeron los siguientes elementos: 10 ml de concentrado de sales, 200 µl de una solución 1 M de K_2HPO_4 , 1,5 g de $NaNO_3$ y 490 ml de agua Elix®. En la segunda botella se añadieron 10 g de Bacto™-agar (Becton Dickinson) y 500 ml de agua Elix®. Ambas botellas se autoclavaron 20 min a 121 °C, para finalmente mezclar ambos preparados a la hora de realizar placas de medio BG11 sólido, cuando la temperatura del agar fundido era inferior a 60 °C.

Para la preparación de medio BG11 selectivo, se añadieron al medio los antibióticos necesarios en las concentraciones que se indican a continuación. En medio líquido, las concentraciones de antibióticos fueron: neomicina (Neo, Sigma) 5 µg/ml si el gen de resistencia está localizado en un plásmido y 1 µg/ml si está en el cromosoma, cloranfenicol (Cm) 5 µg/ml, espectinomicina (Sp, Apollo) 10 µg/ml y estreptomicina (Sm) 10 µg/ml. En medio sólido, las concentraciones fueron: neomicina (Neo) 5 µg/ml, cloranfenicol (Cm) 10 µg/ml, espectinomicina (Sp) 10 µg/ml y estreptomicina (Sm) 50 µg/ml. En medio sólido con filtro de nitrocelulosa (superficie en la que se encuentran las células durante el proceso de conjugación o transformación natural) las concentraciones fueron: neomicina (Neo) 25 µg/ml, cloranfenicol (Cm) 10 µg/ml, espectinomicina (Sp) 20 µg/ml y estreptomicina (Sm) 50 µg/ml.

El cultivo de Ss7002 se realiza de forma rutinaria en medio A+, a una temperatura estándar de 30 °C, con iluminación de luz blanca fluorescente constante (60 µmoles de fotones·m⁻²·s⁻¹) y con agitación orbital (100 rpm) en caso de cultivos líquidos.

Para la preparación del medio A+ es necesaria una solución de metales P1 1000x cuya composición es la siguiente: 34,26 g de H_3BO_3 , 4,32 g de $MnCl_2 \cdot 4H_2O$, 0,315 g de $ZnCl_2$, 0,03 g de MoO_3 (85 %), 0,003 g de $CuSO_4 \cdot 5H_2O$, 0,01215 g de $CoCl_2 \cdot 6H_2O$ y agua Elix[®] hasta 1 L. La mezcla se autoclavó 20 min a 121 °C y se conservó en la nevera protegida de la luz, ya que es fotosensible.

Para la preparación del medio de cultivo A+ líquido se necesitan 18 g de NaCl, 0,6 g de KCl, 1 g de $NaNO_3$, 5 g de $MgSO_4 \cdot 7H_2O$, 1 ml de KH_2PO_4 al 5 % (p/v), 7,2 ml de $CaCl_2$ al 3,7 % (p/v), 20 ml de $Na_2EDTA \cdot 4H_2O$ al 0,3 % (p/v), 1 ml de $FeCl_3 \cdot 6H_2O$ al 0,389 % (p/v en HCl 0,1 N), 10 ml de Trizma Base pH 8,2 al 1 % (p/v), 1 ml de la solución de metales P1 y agua Elix[®] hasta 1 L. La mezcla se autoclavó previo uso mediante protocolo estándar. El medio fue suplementado con una solución de Vitamina B12 a una concentración de 4 mg/L. Esta vitamina es termolábil, por lo que se esterilizó mediante filtración (filtros con un tamaño de poro de 0,22 μm) y se añadió al medio inmediatamente previo a su uso, pues es fotosensible.

Para la obtención de 1 L de medio A+ sólido se emplearon 2 botellas. En la primera se introdujeron los siguientes elementos: 18 g de NaCl, 0,6 g de KCl, 1 g de $NaNO_3$, 5 g de $MgSO_4 \cdot 7H_2O$, 1 ml de KH_2PO_4 al 5 % (p/v), 7,2 ml de $CaCl_2$ al 3,7 % (p/v), 20 ml de $Na_2EDTA \cdot 4H_2O$ al 0,3 % (p/v), 1 ml de $FeCl_3 \cdot 6H_2O$ al 0,389 % (p/v en HCl 0,1 N), 10 ml de Trizma Base pH 8,2 al 1 % (p/v), 1 ml de la solución de metales P1 y agua Elix[®] hasta 500 ml. En la segunda botella se añadieron 15 g de Bacto[™]-agar (Becton Dickinson) y 500 ml de agua Elix[®]. Ambas botellas se autoclavaron 20 min a 121 °C, para finalmente mezclar ambos preparados a la hora de realizar placas de medio BG11 sólido, cuando la temperatura del agar fundido era inferior a 60 °C. En este caso, como ocurre en la preparación de medio A+ líquido, fue necesario añadir una solución de Vitamina B12 a una concentración de 4 mg/L previo a su utilización en cultivos.

4.8.2. Conjugaciones

Para las conjugaciones entre cepas de *E. coli*, se crecieron los cultivos a 37 °C en agitación (150 rpm) durante toda la noche en 10 ml de LB suplementado con el antibiótico correspondiente. Se lavaron los cultivos de donadores y receptores para eliminar los restos de antibióticos y se midió la DO_{600} . Donadores y receptores se mezclaron en un volumen total de 4 ml en relación 1:1. La mezcla se centrifugó durante 5 min a 4.000 rpm. El sobrenadante se desechó y el precipitado celular fue resuspendido en 400 μl de LB. Se preparó previamente una placa de 24 pocillos, añadiendo 1 ml de LB agar (suplementado con DAP si era necesario) a cada

pocillo y dejando solidificar y secar. Un volumen de 15 μl de la mezcla conjugativa se depositó sobre un pocillo con LB agar de dicha placa preparada previamente atemperada. La placa de conjugación se incubó 1 h a 37 °C. Pasado ese tiempo, se interrumpió la conjugación, resuspendiendo las células en 1 ml de PBS 1x estéril. Se hicieron diluciones seriadas y se sembraron 100 μl de las mismas en placas con los antibióticos correspondientes, seleccionando células donadoras, receptoras y transconjugantes. La frecuencia de conjugación se expresó como el número de transconjugantes por célula donadora.

Para las conjugaciones de *E. coli* a Se7942, se emplearon cultivos de cianobacterias crecidos hasta una DO_{720} máxima de 0,5 como receptores y cultivos saturados de *E. coli* como donadores. Antes de comenzar el protocolo es necesario cuantificar la concentración de clorofila (Chl) del cultivo/s de Se7942. Para ello, se centrifugó 1 ml de un cultivo a 13.000 rpm durante 5 min. Se eliminó el sobrenadante y se resuspendió el pellet en 1 ml de metanol 100 %. Se agitó la mezcla en un vórtex durante 1 o 2 min y se centrifugó de nuevo durante 5 min a 13.000 rpm. Se recogió el sobrenadante y se añadió a una cubeta de espectrofotometría de 1 cm de paso, para posteriormente medir la DO_{665} en un espectrofotómetro “UV-1603” (Shimadzu). El blanco de absorbancia (Abs) se realizó con metanol solo. La concentración de clorofila por ml se calculó utilizando la siguiente ecuación: $\mu\text{g Chl/ml} = \text{Abs}_{665} \cdot \epsilon(\text{Chl}) \cdot 10$, donde $\epsilon(\text{Chl})$ se corresponde con el coeficiente de extinción molar de la Chl (13,43). Una vez determinada la cantidad de clorofila por mililitro, se calculó el volumen necesario de cultivo para cada conjugación, siendo 15 μg la cantidad de clorofila necesaria.

Para iniciar el protocolo de conjugación propiamente dicho, se lavaron los cultivos de cianobacterias 2 veces con BG11 y se resuspendieron en 100 μl de BG11. El cultivo de *E. coli* se lavó también dos veces con LB estéril (en caso de usar la cepa β 2150, suplementado con DAP30). Posteriormente, se hicieron diluciones seriadas de estos cultivos en volúmenes de 1 ml. Se realizaron las mezclas conjugativas con 15 μg Chl (resuspendidos en 100 μl de BG11) y 100 μl de la dilución correspondiente del cultivo de *E. coli*. Se preparó también una mezcla control que solo contenía 15 μg Chl (resuspendidos en 100 μl de BG11) y 100 μl BG11. Se sembraron las mezclas conjugativas sobre un filtro de nitrocelulosa de 0,45 μm de tamaño de poro (Millipore), colocados sobre placas de medio BG11 sólido suplementado con 5 % de LB y, en caso de usar la cepa β 2150, también se añadió DAP30. Las placas con las mezclas conjugativas se incubaron a 30 °C, en condiciones de oscuridad durante la primera hora y a continuación con luz blanca a 60 $\mu\text{mol foton}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. A las 24 h, los filtros conjugativos se traspasaron a placas de BG11 y se incubaron en las mismas condiciones de temperatura y luz descritas anteriormente. Tras 24 h, estos mismos filtros se cambiaron nuevamente, en esta ocasión a BG11 selectivo, suplementado

con los antibióticos correspondientes. Tras una incubación de 7 a 14 días en las condiciones previamente descritas, aparecen colonias transconjugantes.

4.8.3. Transformación de ADN

La transformación de las cepas de *E. coli* con ADN se realizó mediante electroporación, dadas la alta eficiencia del método (Dower et al., 1988) y las facilidades logísticas del laboratorio.

Para la preparación de las bacterias electrocompetentes se crecieron cultivos en LB hasta una DO_{600} de 0,4-0,6, y tras exponerlas 30 min a hielo, se recogieron por centrifugación a 4.000 rpm durante 15 min a 4 °C en una centrifuga "5810 R" de Eppendorf. A continuación, se realizaron 2 lavados en agua Milli-Q® autoclavada y a 6 °C. Seguidamente, se realizó un último lavado, esta vez con glicerol 10 % (v/v) también a 6 °C en las mismas condiciones de centrifugación. Después del último lavado, las células se resuspendieron en un volumen de glicerol 10 % de tal manera que se concentrasen 100 veces respecto al volumen original del cultivo. A continuación, se prepararon alícuotas de 50 µl que fueron congeladas inmediatamente en un baño de hielo seco. Estas alícuotas se conservaron a -80 °C hasta su uso.

Por su parte, el ADN que se va a electroporar tiene que estar libre de sales. El ADN procedente de una purificación se empleó directamente, pero el procedente de reacciones enzimáticas se microdializó antes de su uso. La muestra de ADN se dejó reposar durante 30 min sobre un filtro de nitrocelulosa "Millipore GS" de 0,05 µm de tamaño de poro, que a su vez flotaba en una placa de Petri con agua Milli-Q®.

La transformación propiamente dicha se realizó mezclando una alícuota de células competentes y el ADN deseado. Cuando se trató de ADN plasmídico purificado se usaron 1-2 µl, mientras que en el caso de una ligación se usó todo el volumen de la reacción. Esta mezcla se añadió a una cubeta de electroporación de 0,2 cm "Gene Pulser" (BioRad), previamente enfriada en hielo. Se sometió a un pulso eléctrico (2,5 kV/cm, capacitancia 25 µF y 200 Ω) usando un electroporador "MicroPulser" de BioRad. Rápidamente esta mezcla fue resuspendida en 1 ml de medio LB estéril precalentado a 37 °C, y se pasó el resultante a un tubo Eppendorf de 1,5 ml y se incubó a 37 °C en agitación durante 1 - 2 h, en función de los genes de resistencia a expresar. Finalmente, se plaquearon 200 µl del cultivo en una placa de Petri con LB agar suplementado con los antibióticos adecuados para seleccionar los transformantes.

La transformación natural de Se7942 se realizó según el protocolo descrito anteriormente por (Golden and Sherman, 1984). Se empleó un cultivo de Se7942 crecido hasta una DO_{720} de 0,5 como receptor de la transformación y se determinó la cantidad de clorofila como se indica en el apartado **4.8.2**. Los cultivos se lavaron dos veces con medio BG11 para finalmente resuspender 10 μg Chl en 100 μl de BG11 y añadirles 500 ng del ADN transformante. En este caso, al igual que en la conjugación, fue necesario preparar una muestra control que consistió en 10 μg Chl del cultivo (resuspendidos en 100 μl de BG11) sin ADN. La mezcla transformante se dejó a 30 °C, en agitación orbital a 200 rpm y en oscuridad. Tras un periodo mínimo de 5 h y máximo de 24 h, se sembraron las mezclas transformantes en un filtro de nitrocelulosa 0,45 μm (Millipore) sobre una placa de medio BG11. Esta placa se mantuvo en condiciones controladas de temperatura (30 °C) y de luminosidad (luz blanca a 60 μmol fotones $\cdot\text{m}^{-2}\cdot\text{s}^{-1}$). Al cabo de 24 h, el filtro se cambió a una placa de medio BG11 suplementada con los antibióticos correspondientes. Esta placa se incubó en las mismas condiciones que la anterior. Las colonias transformantes se visualizaron en 7 - 14 días.

4.8.4. Introducción de mutaciones en el genoma de Se7942

El reemplazamiento génico mediado por el alelo *rps12-R43*, se basa en el uso de una variante de la proteína S12 de la subunidad 30S del ribosoma que confiere resistencia a Sm (pero es recesiva frente al alelo silvestre) (Funatsu and Wittmann, 1972; Lederberg, 1951). En cianobacterias, este método está descrito para Se7942. La cepa GRPS1 es un derivado de Se7942-R2, en la cual el gen *rps12-wt* ha sido sustituido por *rps12-R43* (codifica una variante K43R de la proteína), confiriendo un fenotipo Sm^{R} (Matsuoka et al., 2001). En nuestro caso introducimos el alelo *rps12-R43* en el genoma de Se7942 mediante transformación natural, seleccionando los transformantes por ser Sm^{R} (**Figura 4-5, A**). Para introducir la mutación deseada, se construye un primer vector con un gen de resistencia a Km y el marcador de selección negativa *rps12*, flanqueados por el gen mutado a sustituir y el gen aguas arriba o abajo a este gen (en la misma disposición en la que se encuentran en el cromosoma de Se792) (**Figura 4-5, B**). Se utiliza el gen *rps12-wt* de Ss6803 para evitar las recombinaciones que se darían en el caso de usar el de Se7942 (Takahama et al., 2004), clonado bajo el promotor del gen *psbA1* de Se7942. Con este plásmido se llevó a cabo una primera transformación natural, seleccionando a los transformantes con Km, lo que hace que el intermediario obtenido tenga un fenotipo Km^{R} . En este primer proceso, además de integrarse en el genoma de Se7942 el gen Km, se inserta también el gen *rps12*, por lo que el fenotipo del transformante merodiploide es Sm^{S} , ya que la

copia *wt* del gen *rps12* es dominante frente a las mutadas (Lederberg, 1951). Seguidamente, los mutantes obtenidos se utilizaron en una segunda transformación natural con un segundo plásmido que contenía los dos genes, tal y como se encontraban en el cromosoma de Se7942 (Figura 4-5, C). La única salvedad es que uno de los genes presentaba una mutación con respecto a la cepa silvestre. Esta última transformación natural se seleccionó con Sm, ya que al desaparecer del cromosoma el gen *wt rps12* introducido, el gen *rps12-R43* vuelve a manifestar el fenotipo Sm^R.

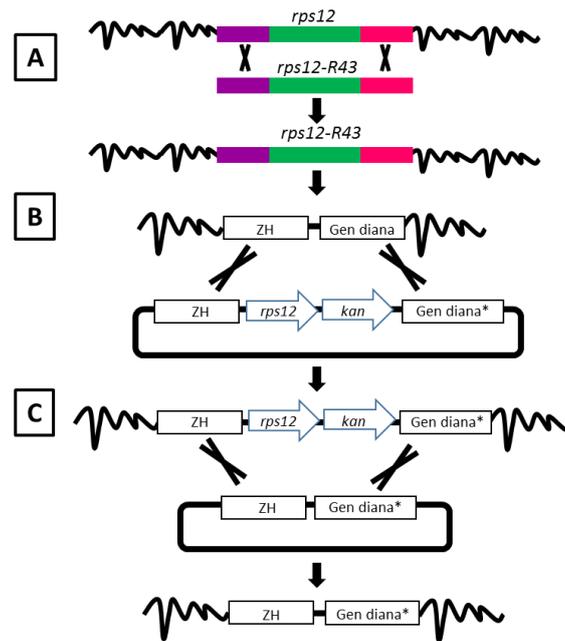


Figura 4-5. Esquema del sistema de introducción de mutaciones o deleciones utilizado. Los cromosomas y los plásmidos se representan como ondas y rectángulos redondeados, respectivamente. Los genes y las zonas homólogas (ZH) se representan como rectángulos. Los sitios de recombinación entre secuencias homólogas se muestran con cruces. Nota: en el caso de querer deleccionar un gen, el gen diana representado en el esquema sería una zona de homología aguas abajo al gen a deleccionar.

Como Se7942 contiene entre 3-4 copias de su cromosoma (Griese et al., 2011), hay que segregar los mutantes obtenidos. Para ello, se subcultivó la bacteria Se7942 varias veces en medio sólido suplementado con el antibiótico correspondiente. Posteriormente, se realizó análisis de las colonias obtenidas mediante PCR, para esclarecer si se habían obtenido mutantes puros o merodiploides.

4.8.5. Cálculo del tiempo de generación de Se7942

Las colonias que contenían la construcción deseada fueron se crecieron en 60 ml de medio BG11 hasta una DO_{720} de $\sim 0,6$ a la temperatura e intensidad de luz correspondiente para cada experimento. Los cultivos se inocularon a una DO_{720} entre 0,04-0,08 y se dejaron crecer hasta una DO_{720} entre 0,8-1. La DO_{720} fue monitorizada en el “Multi-cultivador MC 1000-OD” (Photon Systems Instruments). Para el cálculo del tiempo de generación se realizaron al menos 3 grupos de medidas para cada cepa. Los valores obtenidos fueron representados en gráficas semilogarítmicas. La tasa de crecimiento (K') fue calculada en la fase exponencial temprana de cada curva de crecimiento, que se corresponde con la pendiente de la recta obtenida en dicha representación. De esta manera, el tiempo de generación fue calculado como $\ln 2/K'$. Se realizó un análisis estadístico de los tiempos de generación utilizando un análisis de varianza de un factor (ANOVA), aplicando el test estadístico Tukey-Kramer para comparaciones múltiples usando el software “Prism” (GraphPad Software). Los resultados se presentan como la media \pm SD ($n \geq 5$), considerando un nivel de $p < 0,001$ como significativo.

4.9. Técnicas analíticas usadas en Se7942

4.9.1. Cálculo del espectro de absorción de Se7942

La medida del espectro de absorción de célula completa se realizó siguiendo el protocolo descrito por (Tillich et al., 2014) con pequeñas modificaciones. Las medidas de absorbancia se llevaron a cabo en medio BG11 líquido entre las longitudes de onda 380 nm y 800 nm a temperatura ambiente en cubetas de plástico con un espectrofotómetro “T70+ UV/VIS Spectrometer” (PG Instruments) con una línea base de BG11 y usando pasos de 1 nm. Se tomó 1 ml del cultivo, independientemente de la DO del cultivo. Para normalizar la representación cada valor de absorbancia se dividió por el valor de la absorbancia a 750 nm. Para ello se utilizó la siguiente ecuación: $Abs_N = Abs_x/Abs_{750}$, donde Abs_N , Abs_x y Abs_{750} son las absorbancias normalizada, a una longitud de onda determinada y a 750 nm, respectivamente.

A partir de los datos obtenidos en este espectro, se calculó también el contenido relativo de ficobiliproteínas (FB) y carotenoides (Ct) respecto al de clorofila a (Chla). El contenido relativo de FB se calculó como el cociente entre la Abs_{630} y la Abs_{680} , mientras que el de Ct como el cociente entre la Abs_{495} y la Abs_{680} .

4.9.2. Determinación del contenido en pigmentos en Se7942

La extracción de pigmentos totales se llevó a cabo siguiendo el método del metanol al 100 % (v/v) (Arnon et al., 1974). Se tomó 1 ml del cultivo a una DO_{750} de 0,65, se centrifugó durante 5 min a 10.000 rpm y a 4 °C. Se eliminó el sobrenadante, se resuspendió el pellet en 1 ml de metanol 100 % y se dejó a -20 °C durante 1 h. Tras ese tiempo, se centrifugó a 10.000 rpm durante 10 min a 4 °C y se midió la absorbancia del sobrenadante a 652 nm y 665 nm. Se usó como blanco para tomar las medidas metanol al 100 %. La concentración de clorofila a (Chla) se calculó usando la siguiente ecuación: $Chla (\mu\text{g/ml}) = 16,29 \cdot Abs_{665} - 8,54 \cdot Abs_{652}$, donde Abs_{665} y Abs_{652} son las absorbancias a 665 y 652 nm, respectivamente.

Para medir el contenido en carotenoides se utilizó el sobrenadante obtenido y se midió a 470 nm. Para calcular la concentración de dicho pigmento se utilizó la siguiente ecuación: $Ct (\mu\text{g/ml}) = [(1.000 \cdot Abs_{470}) - (2,86 \cdot Chla)]/245$, donde Abs_{470} es la absorbancias a 470 nm.

La extracción de ficobiliproteínas se llevó a cabo mediante sonicación en hielo. Se tomaron 5 ml del cultivo a una DO_{750} de 0,65, se centrifugó durante 5 min a 4.000 rpm. Se eliminó el sobrenadante, se resuspendió el pellet en 5 ml de PBS 0,01 M pH 7,4 y se sonicó durante 5 min. Tras ese tiempo, se centrifugó a 4.000 rpm durante 10 min y se midió la absorbancia del sobrenadante a 615 nm y 652 nm. Estas longitudes de onda corresponden a la absorción máxima de la ficocianina (PC) y aloficocianina (APC), respectivamente. La concentración de cada pigmento se calculó utilizando las siguientes ecuaciones, donde Abs_{615} y Abs_{652} , corresponden a las absorbancia a 615 y 652 nm, respectivamente.

$$PC (\text{mg/ml}) = ((Abs_{615} - 0,474 \cdot Abs_{652}))/5,34$$

$$APC (\text{mg/ml}) = ((Abs_{652} - 0,208 \cdot Abs_{615}))/5,09$$

$$\text{Total ficobiliproteínas (mg/ml)} = PC + APC$$

4.9.3. Determinación del contenido en ácidos grasos de Se7942

Para el estudio de la naturaleza y composición de los ácidos grasos presenten en las cianobacterias se convirtieron dichos ácidos grasos en ésteres metílicos o FAMES, siguiendo el protocolo descrito por (McInroy and Kloepper, 1995). Se añadieron 1,25 ml del reactivo de saponificación (120 g NaOH, 400 ml metanol, 400 ml agua Elix®) a cada pellet, se agitaron durante 30 s y se calentaron hasta 100 °C en un baño durante 5 min. Se retiraron las muestras, se agitaron durante 30 s y se devolvieron de nuevo las muestras al baño a 100 °C durante 25

min. Pasado ese tiempo, se enfriaron las muestras en un baño de agua fría y se añadieron 2,5 ml del reactivo de metilación (384 ml HCl, 550 ml metanol, 266 ml agua Elix®). Se agitaron las muestras durante 30 s y se trasladaron a un baño a 80 °C durante 10 min. Tras ello, se enfriaron nuevamente las muestras en un baño de agua fría, se añadieron 1,25 ml del reactivo de extracción (200 ml éter etílico, 200 ml hexano) y se agitaron durante 10 min en un agitador rotatorio a 35 - 40 rpm. Se centrifugaron los tubos 3 min a 3.000 rpm, para separar la muestra en dos fases, descartándose la inferior. Se recogió la fase superior en un tubo eppendorf de 1,5 ml, al que previamente se había añadido 0,5 ml del reactivo de lavado (7,2 g NaOH, 600 ml agua Elix®). Se centrifugaron las muestras durante 5 min a 7.000 rpm. Se transfirió el sobrenadante a un vial de cristal de cromatografía de gases y se encapsuló. Se conservaron las muestras a -20 °C hasta que fueron analizadas por cromatografía de gases en la empresa colaboradora Biomar (León).

4.9.4. Medida de la actividad de los promotores en Se7942 usando luminiscencia

La actividad de los promotores se midió utilizando como gen reportero la luciferasa de *Photorhabdus luminescens* en un cultivo en crecimiento siguiendo el método detallado en (Eisenhut et al., 2012). Para determinar la bioluminiscencia producida por un determinado promotor, se tomaron muestras cada h y se ajustó la DO_{720} de 1 ml de cultivo con medio BG11 a 0,5. Se transfirieron 100 μ l por triplicado de cada una de las muestras a una placa microtiter opaca de 96 pocillos (Thermo Fisher Scientific). Antes de comenzar a medir, se preparó una solución stock 100 mM de decanal en metanol, que se diluyó justo antes de añadirla a la placa con BG11 para dar lugar a una solución de uso 2 mM de decanal. Para empezar a medir, se añadieron 100 μ l de una solución de uso 2 mM de decanal y la placa se colocó inmediatamente en el lector de placas “Wallac Victor 3 1420 Multilabel Counter” (Perkin Elmer). La bioluminiscencia se midió cada 10 s durante 15 min a 30 °C. La máxima emisión de luz (alrededor del minuto 10 después de empezar a medir) se usó como valor de luminiscencia y se relativizó a la DO_{720} . Los resultados se representan en unidades relativas de luminiscencia (URL).

4.9.5. Medida de la formación de biofilms en Se7942

Para medir la formación de biofilms de cultivos de Se7942 se utilizó el ensayo de adherencia cristal violeta descrito para Ss6803 (Fisher et al., 2013). Para ello, se inocularon los

cultivos a una DO_{720} de 0,1 en 2 ml de medio BG11 en tubos de cristal. En cada réplica experimental, se crecieron los cultivos por triplicado de forma estática durante 96 h a 30 °C a 60 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. La adherencia de los cultivos se midió de la siguiente manera: se decantó de cada uno de los tubos el medio de cultivo que contenía las células no adheridas. A continuación, se lavó cada tubo cuidadosamente 2 veces con 2 ml de BG11. Se tiñeron los tubos con cristal violeta al 1 % (v/v) disuelto en agua Milli-Q® durante 15 min. Pasado ese tiempo, se lavaron los tubos suavemente 3 veces con agua Milli-Q®. Se resuspendieron las células adheridas en 1 ml de DMSO y se agitó vigorosamente utilizando el vórtex. Tras ello, se midió la DO_{630} de la mezcla de resuspensión en DMSO en un espectrofotómetro “UV-1603” (Shimadzu), utilizando como blanco DMSO. Cada cepa fue analizada por triplicado y los resultados que se muestran son las medias de tres experimentos independientes.

4.9.6. Análisis del pilus de Se7942 por microscopía electrónica

Los cultivos de Se7942 seleccionados para su visualización por microscopía electrónica fueron crecidos hasta llegar a una DO_{720} de 2. Una porción de estos cultivos fue depositada sobre rejillas de Fombarde recién recubiertas a temperatura ambiente durante 30 min y fueron teñidos negativamente con una solución de 1 % (p/v) de ácido fosfotúngstico. Las micrografías fueron tomadas con una magnificación nominal de x20,000 y x25,000 con una cámara GATAN modelo ORIUS SC 1000 CCD, utilizando el microscopio electrónico JEOL JEM-1011 operando a 80 kV en el Servicio de Microscopía Electrónica de la Fundación Instituto de Investigación Marqués de Valdecilla (IDIVAL).

4.10. Análisis bioinformáticos para el estudio de las poblaciones evolucionadas

4.10.1. Identificación de mutaciones en poblaciones evolucionadas

Los “reads” sin ensamblar obtenidos de la secuenciación de las cepas evolucionadas (resultados obtenidos en el apartado **4.7.1**), así como de las cepas control fueron alineados frente a la referencia del genoma de Se7942 anotada en GenBank (número de acceso es NC_007604). Para ello, se utilizó el software gratuito “breseq” (Deatherage and Barrick, 2014) para detectar las mutaciones presentes en las poblaciones experimentales.

4.10.2. Modelado comparativo de la estructura tridimensional de proteínas

El modelado comparativo de proteínas es una técnica predictiva que permite obtener una aproximación de la estructura atómica de una cadena polipeptídica en base a estructuras conocidas de proteínas relacionadas (casi siempre homólogas). En primer lugar, utilizamos el servidor interactivo “HHpred” para buscar estructuras de proteínas ya resueltas que tengan homología con la proteína a modelar (Soding et al., 2005). A continuación, utilizamos el programa “Modeller” (Eswar et al., 2006) o “PHYRE2” (Kelley et al., 2015) para el modelado estructural comparativo de nuestra proteína de interés. Se utilizó como molde la estructura tridimensional resuelta con mayor porcentaje de identidad obtenido con “HHpred”, obtenidas siempre de “Protein Data Bank” (PDB) (Berman et al., 2000).

4.10.3. Herramientas bioinformáticas para el alineamiento de secuencias

Para la visualización de los alineamientos de los “reads” obtenidos tanto de la secuenciación del ADN como del ARN total de Se7942 y cepas derivadas de esta cianobacteria, se usó el software gratuito “Artemis” (Carver et al., 2012).

Para el alineamiento múltiple de secuencias de proteínas se utilizó el programa “MUSCLE”, que produce resultados más precisos y de manera más rápida que “T-Coffee” y “MAFFT” (Edgar, 2004).

4.10.4. Procesamiento de datos obtenidos en el RNAseq

Se utilizó el programa SAMtools, para cuantificar la expresión génica de los archivos obtenidos en el apartado **4.7.2**, representándola como RPKM (lecturas por kilobase de transcrito por millón de lecturas mapeadas). Dado que solo se disponía de una réplica de cada una de las muestras, se dividió cada una de ellas en tres partes, para poder simular que teníamos tres réplicas biológicas y así poder calcular la expresión diferencial. El análisis estadístico para calcular la expresión diferencial frente a la cepa control fue realizado con el paquete DESeq

(Anders and Huber, 2010) y el software R (www.r-project.org/), que usa un modelo estadístico basado en una distribución binomial negativa.

4.11. Reconstrucción del modelo metabólico de Se7942

4.11.1. Herramientas bioinformáticas utilizadas para la reconstrucción del modelo

Para la reconstrucción del modelo metabólico *in silico* de Se7942 se utilizaron las herramientas bioinformáticas que se describen a continuación, así como el proceso de configuración inicial de éstas.

A. Equipo.

- a. Un ordenador personal capaz de ejecutar Matlab.
- b. Matlab, versión 7.0 o superior a éste (The MathWorks Inc., Natick, MA). Matlab es un programa que permite realizar cálculos numéricos y visualizarlos.
- c. COBRA Toolbox versión 2.0 o superior a ésta (Hyduke et al., 2011). COBRA Toolbox es un conjunto de secuencias de comandos específicos para llevar a cabo la reconstrucción y análisis de modelos basados en restricciones y que corren en el entorno Matlab.
- d. La biblioteca de programas libSBML 4.0.1 o superior a ésta.
- e. SBML Toolbox versión 3.1.1. o superior para MATLAB que permita leer y escribir modelos en formato SBML.
- f. Un solucionador de programación lineal. COBRA Toolbox soporta varios solucionadores de acceso abierto y comercial. Para este trabajo se utilizó GLPK (<https://www.gnu.org/software/glpk/>).
- g. Un programa que pueda utilizarse para reunir los datos. Este trabajo se llevó a cabo en Microsoft Excel 2013.

B. Configuración del equipo.

- a. Instalar Matlab.
- b. Instalar libSBML, el SBML Toolbox, y seleccionar el solucionador GLPK.
- c. Instalar COBRA Toolbox 2.0.
- d. Iniciar Matlab y navegar al directorio de COBRA Toolbox.
- e. Guardar la ruta.
- f. Inicializar COBRA Toolbox escribiendo `initCobraToolbox`.

4.11.2. Proceso de reconstrucción metabólica del modelo *iMS705*

La reconstrucción del modelo metabólico de Se7942 se llevó a cabo en cuatro etapas consecutivas, comunes en cualquier reconstrucción: i) construcción del “draft”; ii) curación manual del “draft”; iii) conversión a modelo matemático; y iv) evaluación y análisis; tal y como muestra la **Figura 4-6**.

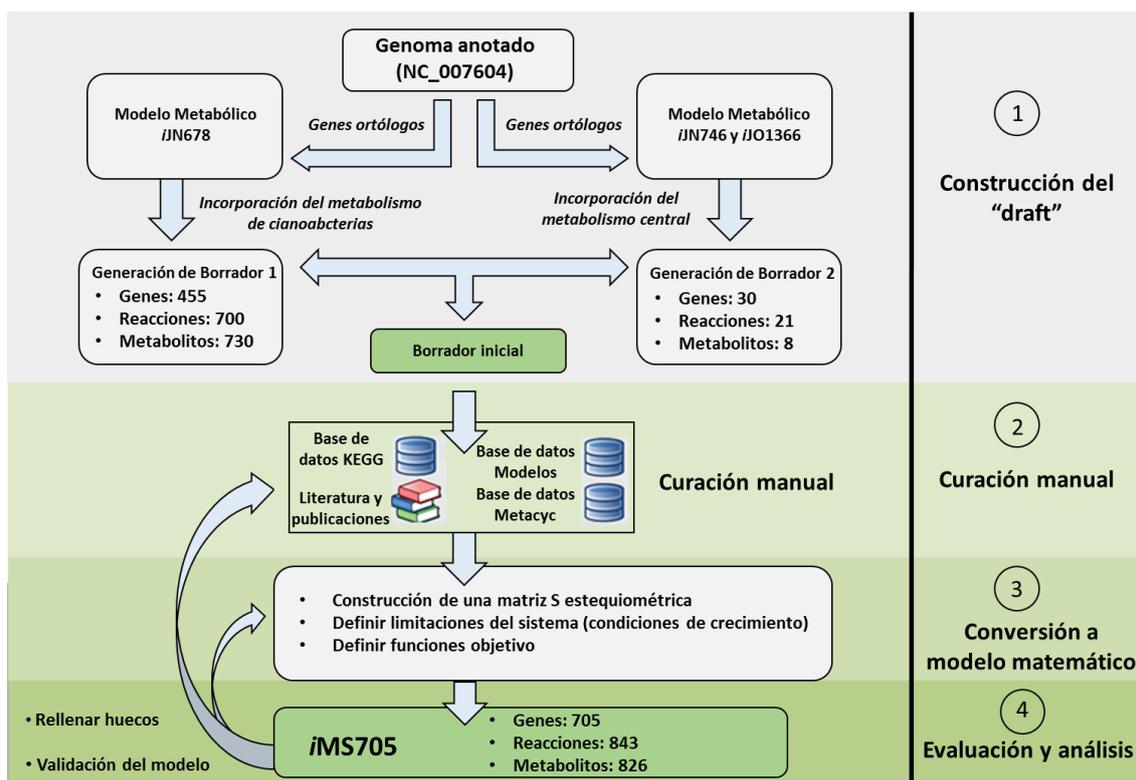


Figura 4-6. Proceso de reconstrucción de un modelo metabólico a escala genómica. Representación detallada de las etapas llevadas a cabo en la reconstrucción del modelo *iMS705* de Se7942.

Etapa 1: Construcción del “draft” inicial

Como elemento indispensable para llevar a cabo la reconstrucción metabólica de un organismo dado, es necesario disponer de una anotación genómica de alta calidad. En nuestro caso, disponíamos del genoma secuenciado y anotado en GenBank de Se7942, cuyo número de acceso es NC_007604.

Para la construcción del primer “draft” del modelo, se utilizó un modelo metabólico de alta calidad existente de una cianobacteria muy cercana filogenéticamente a Se7942, Ss6803 (*iJN678*, (Nogales et al., 2012)). En primer lugar, se procedió a identificar los genes ortólogos

entre Se792 y Ss6803 mediante BLASTp recíproco. Para ello, se comparó la secuencia de aminoácidos de las proteínas codificadas en el genoma de Se7942 con las de Ss6803. Se consideró como genes ortólogos aquellos que obtuvieron doble coincidencia en el análisis y que compartieran un mínimo de un 40 % de identidad a lo largo de toda la secuencia, un 80 % de cobertura (% de aminoácido que el BLASTp puede alinear), y un 10^{-20} del valor esperado (representa el número de aciertos que esperaríamos hallar por casualidad dada la calidad de la alineación y el tamaño de la base de datos).

La creación de este “draft” inicial se obtuvo a partir del programa GEMSiRV-MrBac (<http://sb.nhri.org.tw/GEMSiRV/en/GEMSiRV>). Este programa genera un “draft” del modelo de la cepa Se7942 mediante la asignación de las funciones (reacciones) presentes en el modelo de Ss6803 a los genes ortólogos previamente identificados en Se7942. Para ello, se introdujo la lista previamente creada de estos genes ortólogos al servidor de MrBac. El “draft” resultante fue exportado a Excel como herramienta de análisis de datos que permitió manejarlos más fácilmente. Con esta estrategia se identificaron funciones conservadas entre cianobacterias, que se asocian a 445 genes, 700 reacciones y 730 metabolitos, como puede verse en la **Figura 4-6**.

Adicionalmente se llevó a cabo otro segundo “draft” utilizando otros dos modelos metabólicos de alta calidad de *Pseudomonas putida* KT2440 (*i*JN746, (Nogales et al., 2008)) y de *Escherichia coli* MG1655 (*i*JO1366, (Orth et al., 2011)). Como en el caso anterior, se identificaron los genes ortólogos entre Se7942 y estas dos cepas mediante BLASTp recíproco, utilizando las restricciones descritas anteriormente. De este segundo “draft” se excluyeron los genes que ya se habían detectado en la comparativa con Ss6803. Se generó este segundo “draft” con el programa GEMSiRV-MrBac y se exportó a Excel. Mediante esta estrategia se identificaron mayoritariamente funciones involucradas en el metabolismo central, las cuales se encuentran altamente conservadas entre bacterias. Un total de 30 genes, 21 reacciones y 8 metabolitos se obtuvieron con dicha comparación (**Figura 4-6**).

Ambos “draft” generados se unieron en un único archivo Excel, con el cual se procedió a la segunda etapa, la curación manual.

Etapa 2: Curación manual de la versión inicial del modelo

Aunque la versión inicial del modelo aporta una información muy aproximada de las capacidades metabólicas de Se7942 basado en su contenido génico, no se pueden establecer con absoluta certeza ciertas propiedades específicas en este microorganismo, tales como la

localización de las reacciones, o la especificidad de sustratos y cofactores, ya que estas propiedades son generalmente específicas de especie. Debido a esto, con el fin de reflejar en la reconstrucción las verdaderas capacidades metabólicas de Se7942, es necesaria una revisión manual de las reacciones incluidas en el “draft” inicial. Esta revisión se llevó a cabo mediante la utilización de recursos bioinformáticos como BLAST, bases de datos como BRENDA, KEEG o Metacyc entre otros, así como a partir de la literatura científica existente sobre Se7942 y especies filogenéticamente relacionadas. La revisión manual se realizó sistemáticamente siguiendo los pasos que se describen a continuación:

1) Revisión manual de la versión inicial con la ayuda de la literatura existente

La revisión manual de cada reacción se llevó a cabo mediante la comprobación de los sustratos y cofactores correctos de cada enzima, puesto que pueden diferir entre organismos. Posteriormente las reacciones fueron balanceadas tanto en masa como en carga. Además se verificó la correcta estequiometría y direccionalidad de la reacción, así como la localización subcelular de cada metabolito.

2) Verificación de la asociación gen-proteína-reacción (GPR)

El “draft” automático asigna a cada reacción el gen o genes que codifican la enzima responsable. Es por tanto crítico verificar manualmente si los genes asociados a una reacción son los correctos, puesto que errores en la asociación de la GPR cambiarán los resultados de los estudios *in silico*. La comprobación de la asociación de la GPR a cada reacción en Se7942 se llevó a cabo con la ayuda de la anotación del genoma. Para ello, se realizaron alineamientos BLASTp en la base de datos del NCBI utilizando la secuencia de aminoácidos de las proteínas codificadas por cada gen. Se asumieron determinados criterios, los mismos que en la etapa 1, que junto con la revisión de la literatura científica existente, corroboraron si dicha enzima estaba presente en el metabolismo de Se7942. Además, la revisión de la literatura permitió verificar si la proteína funcional era un complejo enzimático, si la proteína o complejo participaba en más de una reacción, y si más de una proteína llevaba a cabo la misma función.

3) Aportar referencias sobre la información actual de cada reacción

Los datos reunidos de algunos organismos relacionados con el organismo de estudio, así como los datos específicos existentes sobre el mismo, se incluyen en la reconstrucción en forma

de referencias y notas. Así, se revisó la información actual de cada reacción de la reconstrucción de Se7942 en recursos bioinformáticos como BRENDA y/o revisando la literatura existente. Estos datos permitieron revisar las evidencias y material de apoyo para la inclusión de reacciones y genes presentes en Se7942.

4) Asignación del valor de confianza a cada reacción

El valor de confianza proporciona de manera rápida la información y evidencias actualmente disponibles del microorganismo a modelar. Por ello, a cada reacción de la reconstrucción se le asignó un valor de confianza con un rango de 0 a 4, donde 0 es el valor de evidencias más bajo y 4 el más alto (**Tabla 4-6**).

Tabla 4-6. Sistema empleado actualmente para la asignación del valor de confianza de cada reacción. Tomado de (Thiele and Palsson, 2010).

| Tipo de evidencia | Valor de confianza | Ejemplos |
|--------------------|--------------------|---|
| Datos bioquímicos | 4 | Evidencias directas para funciones de productos de genes y reacciones bioquímicas: ensayos bioquímicos, purificación de proteínas. |
| Datos genéticos | 3 | Evidencias directas o indirectas para la función de un gen: caracterización de un gen <i>knockout</i> , caracterización y sobreexpresión de un gen <i>knockin</i> . |
| Datos fisiológicos | 2 | Evidencias indirectas para reacciones bioquímicas basadas en datos fisiológicos: productos de secreción o componentes del medio que sirven como evidencias para el transporte y reacciones metabólicas. |
| Datos del modelo | 1 | No existen evidencias disponibles, pero la reacción se requiere en el modelo. La función incluida es una hipótesis y necesitaría verificarse mediante experimentación. |

5) Añadir capacidades metabólicas específicas de cada especie

Para incluir las capacidades metabólicas específicas de Se7942 carentes en Ss6803, se obtuvo una lista de genes no comunes entre ambas especies. Manualmente, se revisó esta lista de genes utilizando recursos bioinformáticos como BLAST, bases de datos como BRENDA, KEGG o Metacyc entre otros, así como a partir de la literatura científica existente sobre Se7942. En el caso de la cepa Se7942, el análisis de estos genes permitió reconstruir gran parte de las rutas de síntesis de pigmentos no fotosintéticos y de compuestos de la envoltura celular e identificar transportadores de membrana.

6) Construcción de la función objetivo: la reacción de biomasa (BOF, “Biomass Objective Function”)

BOF es una reacción única de cada especie que contiene todos los precursores esenciales que constituyen el contenido de la biomasa de un organismo. Esta reacción describe la velocidad a la que la biomasa celular se sintetiza en las proporciones correctas, por lo que simula el crecimiento celular (Feist and Palsson, 2010). Puesto que se asume que las bacterias optimizan su tasa de crecimiento en estado estacionario, la predicción de flujos *in silico*, se computa maximizando la reacción BOF, resolviendo así el conjunto de ecuaciones lineales generado al usar el análisis FBA. Dado que actualmente existe suficiente información sobre la composición macromolecular de Se7942, se construyó una BOF específica que se detalla en la **Tabla 4-7**. Solo en el caso de la composición del “pool” soluble y de iones se utilizó la información descrita para el modelo de Ss6803 (Nogales et al., 2012).

Tabla 4-7. Composición macromolecular de la biomasa de Se7942 presente en el modelo metabólico iMS705.

| Categoría | Compuesto | mmol/g peso seco | Compuesto | mmol/g peso seco |
|------------------|------------------|------------------|-----------|------------------|
| Aminoácidos | ala-L | 0,430429 | leu-L | 0,502770 |
| | arg-L | 0,276912 | lys-L | 0,107005 |
| | asn-L | 0,112680 | met-L | 0,062237 |
| | asp-L | 0,209755 | phe-L | 0,142064 |
| | cys-L | 0,044171 | pro-L | 0,223670 |
| | gln-L | 0,249622 | ser-L | 0,246328 |
| | glu-L | 0,232409 | thr-L | 0,212999 |
| | gly | 0,293231 | trp-L | 0,071686 |
| | his-L | 0,070297 | tyr-L | 0,100144 |
| | ile-L | 0,222897 | val-L | 0,272311 |
| ADN | datp | 0,010891 | dgtp | 0,013525 |
| | dctp | 0,013520 | dttp | 0,010858 |
| ARN | ctp | 0,036897 | utp | 0,030575 |
| | gtp | 0,048778 | atp | 0,038565 |
| Glicógeno | Glicógeno | 1,233494 | | |
| Peptidoglicano | Peptidoglicano | 0,025228 | | |
| Lipopolisacárido | Lipopolisacárido | 0,025228 | | |
| Lípidos* | mgdg140 | 0,001625 | dgdg181 | 0,003665 |
| | mgdg141 | 0,001509 | sqdg140 | 0,000052 |
| | mgdg160 | 0,055958 | sqdg160 | 0,002711 |
| | mgdg161 | 0,043884 | sqdg161 | 0,001497 |

| Categoría | Compuesto | mmol/g peso seco | Compuesto | mmol/g peso seco |
|-------------------|--------------------------|------------------|--------------------------------|------------------|
| | mgdg180 | 0,003831 | sqdg180 | 0,000071 |
| | mgdg181 | 0,008939 | sqdg181 | 0,000353 |
| | dgdg140 | 0,000819 | pg140 | 0,000152 |
| | dgdg141 | 0,000156 | pg160 | 0,013059 |
| | dgdg160 | 0,019690 | pg161 | 0,008513 |
| | dgdg161 | 0,013218 | pg180 | 0,000404 |
| | dgdg180 | 0,001287 | pg181 | 0,003006 |
| "Pool" soluble | Putrescina | 0,033270 | 5,10-Metilen tetrahidrofolato | 0,000223 |
| | Espermidina | 0,006744 | 5-Metiltetrahidrofolato | 0,000223 |
| | Acetil-CoA | 0,000279 | Tiamina difosfato | 0,000223 |
| | CoA | 0,000168 | Hemo O | 0,000223 |
| | Malonil-CoA | 0,000031 | Protohemo | 0,000223 |
| | NAD | 0,001787 | Glutación reducido | 0,000223 |
| | NADH | 0,000045 | Adenosilcobalamina | 0,000223 |
| | NADP | 0,000112 | Undecaprenil difosfato | 0,000055 |
| | NADPH | 0,000335 | 10-Formil tetrahidrofolato | 0,000223 |
| | FAD | 0,000223 | Corismato | 0,000223 |
| | 5,6,7,8-Tetrahidrofolato | 0,000223 | 8-Amino-7-oxononanoato | 0,000223 |
| Riboflavina | 0,000223 | | | |
| Iones inorgánicos | K ⁺ | 0,170400 | Mn ²⁺ | 0,003029 |
| | NH ₄ | 0,011360 | MoO ₄ ²⁺ | 0,003029 |
| | Mg ²⁺ | 0,007573 | CO ²⁺ | 0,003029 |
| | Ca ²⁺ | 0,004544 | Zn ²⁺ | 0,003029 |
| | Fe ²⁺ | 0,006816 | SO ₄ | 0,003787 |
| | Fe ³⁺ | 0,006816 | Pi | 0,003787 |
| | Cu ²⁺ | 0,003029 | Na ²⁺ | 0,003787 |
| Pigmentos | Clorofila a | 0,014146 | β-Criptoxantina | 0,000039 |
| | Filoquinona | 0,001768 | Caloxantina | 0,000271 |
| | β-Caroteno | 0,002013 | Nostoxantina | 0,000077 |
| | Zeaxantina | 0,001471 | | |
| Gasto de energía | ATP | 53,368511 | H ₂ O | 48,150193 |
| | ADP | 53,35 | H ⁺ | 53,35 |
| | Pi | 53,3462 | PPi | 0,097082 |

*Lípidos: mgdg, monogalactosil-diacilglicerol; dgdg, digalactosil-diacilglicerol; sqdg, sulfoquinovosil-diacilglicerol; pg, fosfatidilglicerol-didodecanoil.

7) Completar la reconstrucción con reacciones no asociadas a genes

En el “draft” inicial que se generó automáticamente, se asocian genes a todas las reacciones por similitud de secuencia. Sin embargo, a veces no se conocen los genes que codifican reacciones bioquímicas bien conocidas en el microorganismo de estudio. Estas reacciones se denominan reacciones huérfanas. Por otro lado, es bien conocida la existencia de reacciones espontáneas en sistemas biológicos. Es por esto que estas reacciones huérfanas y espontáneas están ausentes en los “draft” y es necesaria su inclusión en un paso posterior. Así, un total de 63 reacciones huérfanas y 10 espontáneas fueron añadidas a la reconstrucción.

8) Completar la reconstrucción con transportadores y reacciones de intercambio

Es necesario introducir reacciones que representen los transportadores de los metabolitos que se están intercambiando entre los diferentes compartimentos celulares o entre estos y el exterior. Este paso es importante para determinar los nutrientes potencialmente utilizables por el microorganismo de estudio. Es importante contar con evidencias experimentales para verificar que existen dichos transportadores. También se deben incluir las reacciones de transporte de metabolitos que se intercambian por difusión como el O₂ o el CO₂. Un total de 115 reacciones de transporte fueron incluidas en el modelo.

Etapa 3: Conversión de la reconstrucción a modelo matemático

Esta etapa se lleva a cabo en dos pasos.

1) Conversión a modelo matemático

La conversión a formato matemático se realizó de forma automática utilizando Matlab. De esta forma se construyó la matriz estequiométrica constituida por 843 reacciones y 826 metabolitos.

2) Selección de la función objetivo

Para resolver los problemas de la programación lineal usando la aproximación FBA es necesario optimizar una función objetivo como se ha indicado en la introducción. A menudo se utiliza como objetivo la BOF con el fin de estimar los flujos metabólicos que maximizan la tasa

de crecimiento del organismo de estudio. Así, nosotros seleccionamos la reacción de la biomasa (BOF) como objetivo.

Etapas 4: Validación del modelo y análisis

Para determinar la fiabilidad del modelo, comprobar posibles errores durante la reconstrucción, o la identificación de huecos entre rutas metabólicas, es necesario una comprobación y validación del mismo. Todos los análisis llevados a cabo en este trabajo se realizaron con el paquete de software COBRA Toolbox (Hyduke et al., 2011), en el entorno de computación de Matlab (The MathWorks Inc., Natick, MA).

1) Denominación del modelo final

El modelo *iMS705* se nombró de acuerdo con la convención para modelos metabólicos en donde *i* es sinónimo de *in silico*, MS hace referencia a las iniciales del constructor del modelo (María Santos), y 705 indica el número de genes incluidos en la reconstrucción.

Resultados

5. Resultados

5.1. Evolución experimental de Se7942

5.1.1. Experimento de evolución en Se7942

Como se ha mencionado en la sección **2.2** del capítulo introducción, la evolución experimental se define como el estudio de los cambios evolutivos en poblaciones experimentales como consecuencia de las condiciones impuestas por el investigador (Kawecki et al., 2012). En nuestro caso, la población experimental utilizada fue la cianobacteria Se7942 y las condiciones impuestas fueron el cultivo continuado bajo alta intensidad de luz ($1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$), alta temperatura ($41 \text{ }^\circ\text{C}$) y una alta concentración de CO_2 (5 %). Se partió de dos poblaciones de Se7942 genéticamente idénticas, a las que denominamos C11 y C12, que no estaban adaptadas a las condiciones en las que se iba a llevar a cabo el experimento de evolución. Estas poblaciones se tomaron de la Colección de Cultivos del Pasteur (PCC), en la cual se mantenían en medio BG11 a $25 \text{ }^\circ\text{C}$, a baja intensidad de luz ($65 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$), y sin burbujeo de CO_2 . Se cultivaron seriadamente en las nuevas condiciones, y se realizaron pases cada 6 días (aproximadamente cada 24 generaciones) diluyendo el cultivo 1:130. Se tomaron muestras en las generaciones 0, 824 y 1248 en el caso de la línea C11 y en las generaciones 0, 728 y 1168 en el caso de la línea C12, tal y como se muestra en la **Figura 5-1**, para su posterior análisis.

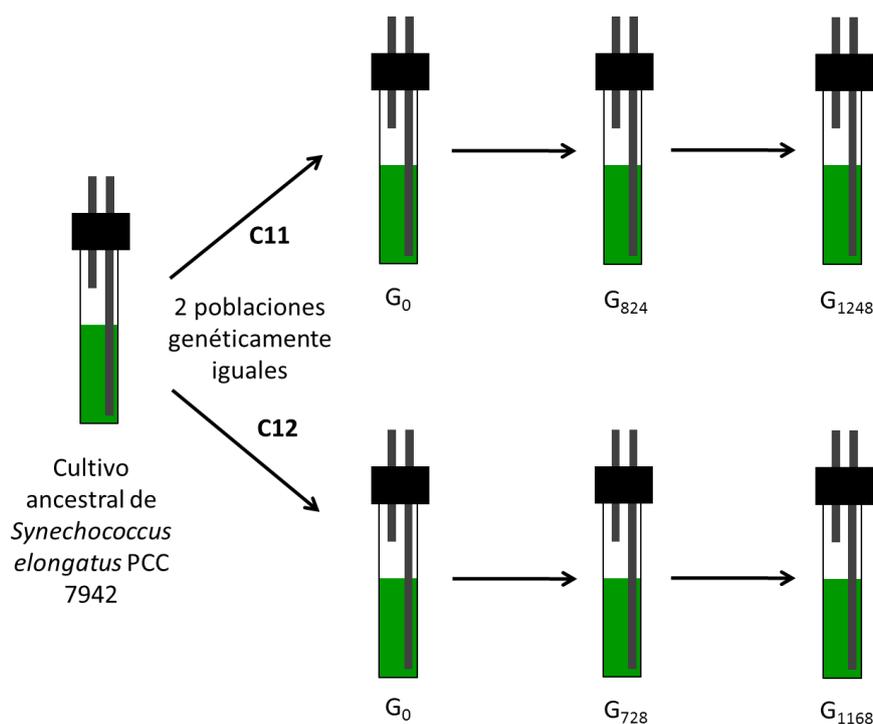


Figura 5-1. Esquema del experimento de evolución de Se7942. 2 poblaciones de Se7942 genéticamente idénticas se diluían 1:130 cada 6 días en medio fresco. Se tomaron muestras para su posterior análisis en las generaciones indicadas (G_n).

Una vez realizado el experimento de evolución, se procedió a analizar el tiempo de generación que presentaban las líneas C11 y C12 en las generaciones 1248 y 1168, respectivamente. Estos experimentos se realizaron en las mismas condiciones en las que se había llevado a cabo el experimento de evolución (41 °C, 1.313 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ y 5 % CO_2) utilizando como control la cepa ancestral de Se7942. Una de las curvas de crecimiento realizada con estas cepas se muestra en la **Figura 5-2**. Calculamos el tiempo de generación de cada una de las cepas (ver apartado **4.8.5**). Para la cepa ancestral (G_0), el tiempo de generación fue de 8,6 h, para C11- G_{1248} , 3,13 h y para C11- G_{1168} , 2,95 h. Por tanto, podemos afirmar que las cepas C11 y C12 han mejorado su tiempo de generación significativamente como consecuencia del experimento de evolución.

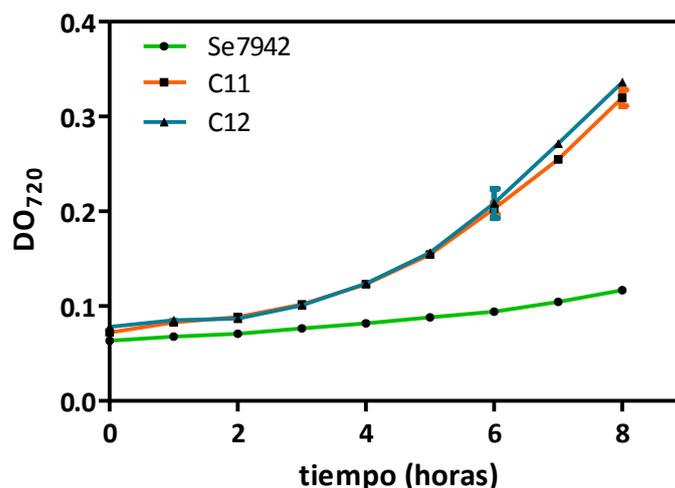


Figura 5-2. Curva de crecimiento de la cepa Se7942 y de las cepas evolucionadas C11 y C12. En verde, se representa la cepa ancestral Se7942; en naranja, la cepa evolucionada C11 generación 1248 y en azul, la cepa evolucionada C12 generación 1168. Cada punto corresponde a la media de tres valores de medida de la $DO_{720} \pm DE$.

Una vez observamos que las cepas C11 y C12 habían mejorado su tiempo de generación, se secuenciaron tanto las poblaciones G_{824} y G_{1248} de la línea C11, como las G_{728} y G_{1168} de la línea C12 para identificar SNPs y/o indels responsables de dicho fenotipo. En la **Figura 5-3**, se muestra un esquema de las mutaciones que aparecen en las cepas secuenciadas. Como se observa en C11- G_{824} y C12- G_{728} , las mutaciones que aparecieron en las poblaciones no se habían fijado, ya que su frecuencia era menor al 100 %. Sin embargo, ya en C11- G_{1248} y C12- G_{1168} , alguna de estas mutaciones están fijadas y otras se han seleccionado significativamente. Como controles en la secuenciación utilizamos dos cepas, la primera corresponde a una muestra de la cepa silvestre de Se7942 que se congeló al inicio del experimento y a la que en la **Figura 5-3** denominamos congelada. El segundo control es una muestra de cultivo de la cepa silvestre, misma cepa con la que se inició el experimento, que continuó creciendo en condiciones de la PCC y se secuenció junto a las cepas evolucionadas. Cabe destacar que la cepa con la que se inició el experimento, a la que en la **Figura 5-3** denominamos congelada, ya presentaba ciertas mutaciones respecto a la secuencia del genoma de Se7942 anotada en GenBank (NC_007604). Por otro lado, la cepa silvestre que se cultivó de forma continuada a lo largo del experimento solo presenta una mutación seleccionada significativamente respecto a las que tenía la cepa congelada. En la **Tabla 5-1**, que discutiremos a continuación, se muestran todas estas mutaciones en detalle.

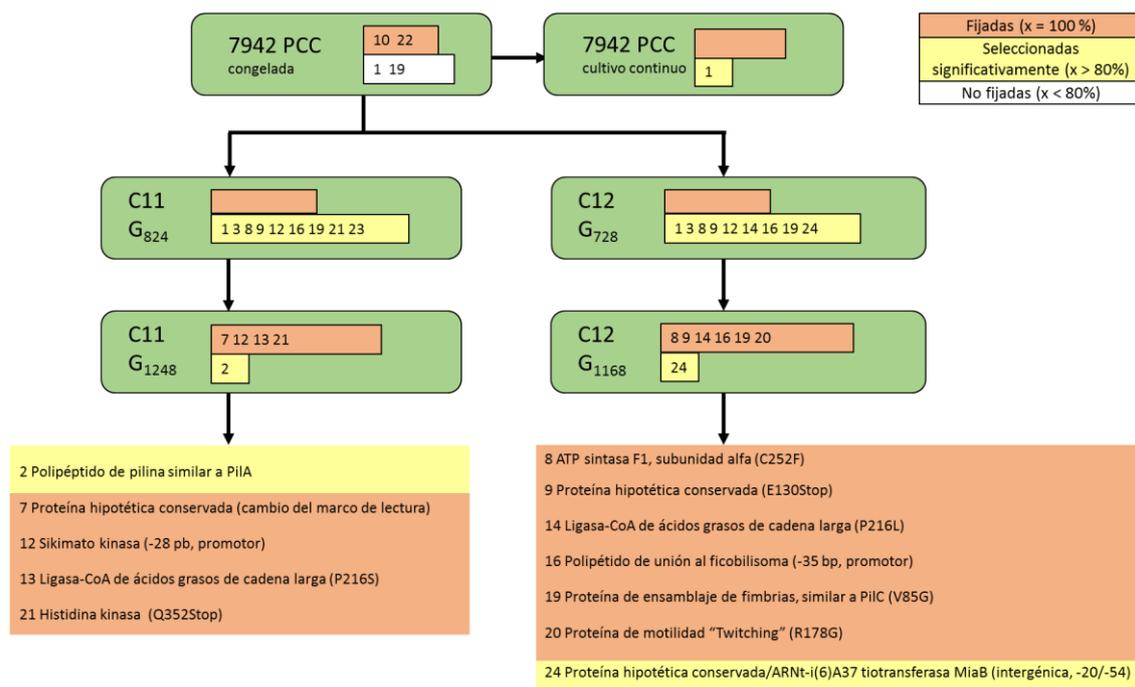


Figura 5-3. Esquema de las mutaciones presentes en los genomas de las poblaciones utilizadas en el experimento de evolución. En total son 6 poblaciones: congelada, cultivada continuamente, C11-G₈₂₄, C11-G₁₂₄₈, C12-G₇₂₈ y C12-G₁₁₆₈. En naranja, se muestran las mutaciones fijadas (proporción = 100 %). En amarillo, se muestran las mutaciones seleccionadas significativamente (proporción > 80 %). En blanco, se muestran las mutaciones no fijadas (proporción < 80 %). Solo se representan las mutaciones no presentes en las cepas de partida, congelada y cultivo continuo en laboratorio, y que han surgido en el experimento de evolución.

Para obtener las mutaciones, SNPs e indels, que se representan en la **Figura 5-3** y en la **Tabla 5-1**, se mapearon las secuencias de las dos cepas control y de las cepas evolucionadas frente a la secuencia referencia del GenBank (NC_007604) con el software "breseq" (ver apartado 4.10.1). De las nueve mutaciones de la cepa C11-G₈₂₄, solo dos (#12 y #21), se mantuvieron en la generación 1248. Por otro lado, en la línea C12, cinco de las nuevas mutaciones presentes en la generación 728 se mantuvieron en la generación 1168. Estas mutaciones corresponden a #8, #9, #14, #19 y #24. Cabe destacar que la mutación #1 que estaba presente en los dos controles, congelada y cultivo continuo, se mantenga en generaciones tempranas de las cepas evolucionadas y desaparezca en generaciones tardías. En su lugar aparecen otras mutaciones en este mismo gen, mutaciones #2 y #3. Existen otras mutaciones, sombreadas en verde en la **Tabla 5-1**, que se encuentran en todas las cepas secuenciadas con una frecuencia de aparición del 100 % y que obviamente no se han representado en la **Figura 5-3**. Estas mutaciones pueden deberse a errores de secuenciación presentes en la secuencia de referencia anotada en GenBank o a mutaciones presentes en la cepa de Se7942 que tenemos en nuestro laboratorio. En la **Tabla 5-1** sombreadas en naranja, se muestran dos mutaciones únicamente presentes en la cepa control que se mantuvo congelada, #10 y #22. Estas mutaciones pueden ser errores de secuenciación, ya que como se observa en esta misma tabla,

no están presentes en la cepa que se mantuvo en cultivo continuo en el laboratorio. En el apartado siguiente describiremos los genes en los que se encuentran las principales mutaciones en las líneas C11-G₁₂₄₈ y C12-G₁₁₆₈.

5. Resultados

Tabla 5-1. Mutaciones de las poblaciones obtenidas en el experimento de evolución. Sombreadas en azul se muestran aquellas mutaciones presentes en las cepas control (congelada y cultivo continuo) que se conservan en algunas cepas evolucionadas. Sombreadas en naranja se muestran dos mutaciones únicamente presentes en la cepa control que se mantuvo congelada. Sombreadas en verde se muestran aquellas mutaciones que estaban en las cepas con las que se inició el experimento y se han conservado en todas las cepas evolucionadas a lo largo de las diferentes generaciones. Sombreadas en blanco se muestran las mutaciones que han surgido a lo largo del experimento de evolución. En la columna donde se indica el tipo mutación, si la inserción era de gran tamaño se indica con un “+” seguida de número de nucleótidos de la inserción, pero no la secuencia de la misma. En la columna de anotación, si la inserción o deleción afecta a una proteína codificante se indica la posición de la mutación, así como la longitud de la secuencia codificante separadas por /; si es un SNP, se muestra en rojo y subrayado el cambio en el aminoácido que ha tenido lugar; si afecta a una región intergénica, se indica la posición de la mutación respecto a ambos genes separada por /. En la columna del gen, la flecha indica la dirección del gen en el cromosoma de Se7942.

| Nº | Posición | Mutación | Congelada | Cultivo continuo | C11-G ₈₂₄ | C11-G ₁₂₄₈ | C12-G ₂₈ | C12-G ₁₁₆₈ | Anotación | Gen | Descripción |
|----|----------|----------|-----------|------------------|----------------------|-----------------------|---------------------|-----------------------|------------------------------|---|---|
| 1 | 49,248 | Δ3 bp | 11,50% | 97,80% | 52,80% | - | 61,30% | - | Codificante (112-114/435 nt) | Synpcc7942_0048 ← | Polipéptido de pilina similar a PilA |
| 2 | 49,250 | +GATAAC | - | - | - | 88,52% | - | - | Codificante (+IV 39) | Synpcc7942_0048 ← | Polipéptido de pilina similar a PilA |
| 3 | 49,333 | T → C | - | - | 47,4% | - | 58,5% | - | L10P (CIG → CCG) | Synpcc7942_0048 ← | Polipéptido de pilina similar a PilA |
| 4 | 92,906 | G → A | 13,40% | 100% | 100% | 100% | 100% | - | P145L (CC → CIC) | Synpcc7942_0095 ← | Regulador transcripcional de dos componentes, familia de hélice alada |
| 5 | 92,978 | T → C | 100% | 100% | 100% | 100% | 100% | - | Q121R (CAG → CCG) | Synpcc7942_0095 ← | Regulador transcripcional de dos componentes, familia de hélice alada |
| 6 | 159,223 | A → T | 44,30% | 100% | 100% | 100% | 100% | - | L113* (TIG → TAG) | Synpcc7942_0157 ← | Proteína hipotética conservada |
| 7 | 326,326 | +T | - | - | - | 100% | - | - | Codificante (214/1245 nt) | Synpcc7942_0329 ← | Proteína hipotética conservada |
| 8 | 331,632 | G → T | - | - | 42,80% | - | 39,40% | 100% | C252F (TGC → TIC) | Synpcc7942_0336 → → cc7942_0329 ← Synpcc7942_0329 ← | ATP sintasa F1, subunidad alfa |
| 9 | 441,449 | C → A | - | - | 42,30% | - | 41,00% | 100% | E130* (GAG → TAG) | Synpcc7942_0452 ← | Proteína hipotética conservada |
| 10 | 641,282 | A → C | 100% | - | - | - | - | - | P29P (CCA → CCG) | Synpcc7942_0647 → | Proteína hipotética conservada |
| 11 | 893,872 | C → G | 87,70% | 100% | 100% | 100% | 100% | - | G85R (GGT → CGT) | Synpcc7942_0887 ← | SSU proteína ribosomal S12P |
| 12 | 905,445 | C → T | - | - | 9,40% | 100% | 1,90% | - | Intergénica (28/+40) | Synpcc7942_0894 ← / ← Synpcc7942_0895 | Shikimato kinasa/Proteína hipotética conservada |
| 13 | 924,724 | C → T | - | - | - | 100% | - | - | P216S (CCA → TCA) | Synpcc7942_0918 → | Ácido graso de cadena larga-CoA ligasa |

| Nº | Posición | Mutación | Congelada | Cultivo continuo | C11-G ₂₄ | C11-G ₂₄₈ | C12-G ₇₂₈ | C12-G ₁₁₆₈ | Anotación | Gen | Descripción |
|----|-----------|----------|-----------|------------------|---------------------|----------------------|----------------------|---|--|---|-------------|
| 14 | 924,725 | C → T | - | - | - | 20,10 % | 100 % | P216L (C C A → C T A) | Synpcc7942_0918 → | Ácido graso de cadena larga-CoA ligasa | |
| 15 | 924,962 | T → C | 100 % | 100 % | 100 % | 100 % | 100 % | L295P (C I G → C C G) | Synpcc7942_0918 → | Ácido graso de cadena larga-CoA ligasa | |
| 16 | 1,062,322 | C → A | - | - | 40,80 % | 40,10 % | 100 % | Intergénica (+13/ 35) | Synpcc7942_1049 → / → Synpcc7942_1050 | Polipéptido de unión al fibobilisoma/Polipéptido de unión al fibobilisoma | |
| 17 | 1,787,336 | A → C | - | 100 % | 100 % | - | 100 % | V107G (G I G → G G G) | Synpcc7942_1716 ← | Diguanilato ciclasa/Fosfodiesterasa con sensor(es) PAS/PAC | |
| 18 | 2,026,943 | A → G | 64,10 % | 100 % | 100 % | 100 % | 100 % | V312A (G I G → G C G) | Synpcc7942_1954 ← | CTP sintasa | |
| 19 | 2,146,088 | A → C | 2,40 % | - | 38,90 % | 36,40 % | 100 % | V85G (G I T → G G T) | Synpcc7942_2069 ← | Proteína de ensamblaje de fimbrias, similar a PilC | |
| 20 | 2,146,923 | G → C | - | - | - | - | 100 % | R178G (G G T → G G T) | Synpcc7942_2070 ← | Proteína de motilidad "Twitching" | |
| 21 | 2,195,880 | C → T | - | - | 12,30 % | - | - | Q352* (C A G → I A G) | Synpcc7942_2114 → | Histidina kinasa | |
| 22 | 2,296,973 | C → G | 100 % | - | - | - | - | R3P (C G C → C C C) | Synpcc7942_2228 ← | SSU proteína ribosomal S19P | |
| 23 | 2,441,072 | +A | 7,60 % | 100 % | 54,30 % | 100 % | - | Codificante (302/1035 nt) | Synpcc7942_2373 ← | Proteína hipotética conservada | |
| 24 | 2,441,393 | +18 pb | - | - | 40,10 % | 29,10 % | 80,70 % | Intergénica (20/ 54) | Synpcc7942_2373 ← / → Synpcc7942_2374 | Proteína hipotética conservada/ ARNt i(6)A37 tiotransferasa MiaB | |

5.1.2. Descripción de los genes que presentan mutaciones en las poblaciones evolucionadas

A continuación describiremos las mutaciones presentes en las poblaciones C11-G₁₂₄₈ y C12-G₁₁₆₈. Analizamos solo las mutaciones presentes en estas poblaciones, ya que sabemos que presentan una reducción en el tiempo de generación (**Figura 5-2**) y por ello, sería interesante determinar su función sobre este fenotipo.

5.1.2.1. Mutación #2: Polipéptido de pilina similar a PilA

La mutación #2 de la población C11 es una inserción de seis nucleótidos en la posición 49.250 pb del genoma de Se7942. Esta mutación afecta al gen Synpcc7942_0048, que codifica para una pilina similar a PilA. El cambio que tiene lugar es una inserción de dos aminoácidos (isoleucina y valina) tras el aminoácido 39 de la proteína sin cambiar su marco de lectura. Su abundancia en la población fue de un 88 %. Se7942, al igual que otras cianobacterias, posee otra copia extra de este gen en su cromosoma, Synpcc7942_0049.

PilA es una subunidad del pilus tipo IV (fimbrias), frecuentes en la superficie de bacterias gram-negativas. Las fimbrias están involucradas en virulencia, absorción del ADN, motilidad “twitching” y adhesión. Aunque el ensamblaje del pilus es un proceso complejo, los pili son principalmente homopolímeros de proteínas PilA, las cuales son añadidas o eliminadas vía ATPasas para extender o retraer el pilus (Schuergers and Wilde, 2015).

Como puede observarse en la **Figura 5-4**, la mutación #2 de C11 afecta a una hélice alfa, dando lugar a un cambio en el plegamiento de la proteína, pero no afectando a su hidrofobicidad. Este cambio podría afectar a la unión entre cadenas en el proceso de ensamblaje del pilus. (Sokol, 2015) mutó una de las dos copias del gen *pilA* en Se7942 y observó que cuando están presentes las copias silvestres y mutadas de dicho gen se impedía el ensamblaje del pilus. Este autor sugiere además que existe una interacción tóxica entre ambos tipos de proteínas que da lugar a una incapacidad para ensamblar el pilus.

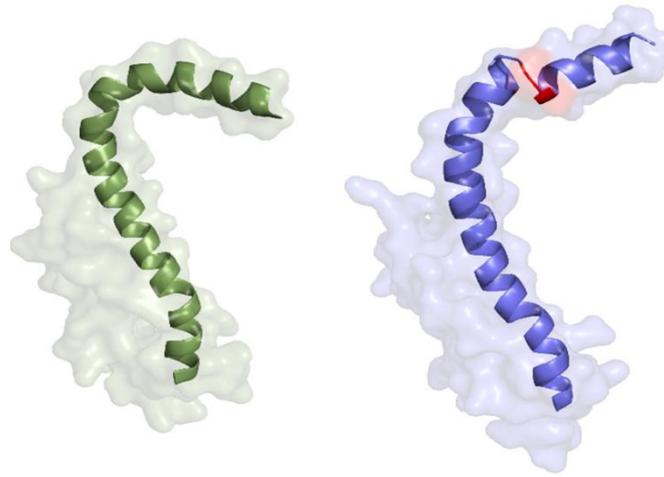


Figura 5-4. Predicción de la estructura de la proteína PiIA de la cepa silvestre (izquierda) y de la cepa C11 (derecha). Obtenido mediante modelado comparativo utilizando como modelo la estructura 2M7G de la proteína PiIA de *Geobacter sulfurreducens* de PDB. En rojo, la inserción de dos aminoácidos. Realizado con “Modeller”.

5.1.2.2. Mutación #7: Proteína hipotética conservada

La mutación #7 de C11 afecta a una proteína hipotética conservada. Consiste en una inserción de un nucleótido en la posición 326.326 pb del genoma de Se7942, y está presente en el 100 % de la población analizada. Esta inserción da lugar a un cambio del marco de lectura produciendo un codón de STOP temprano, truncando la proteína y quizás, alterando la expresión del operón entero debido al cambio en la estabilidad del ARNm (**Figura 5-5**).

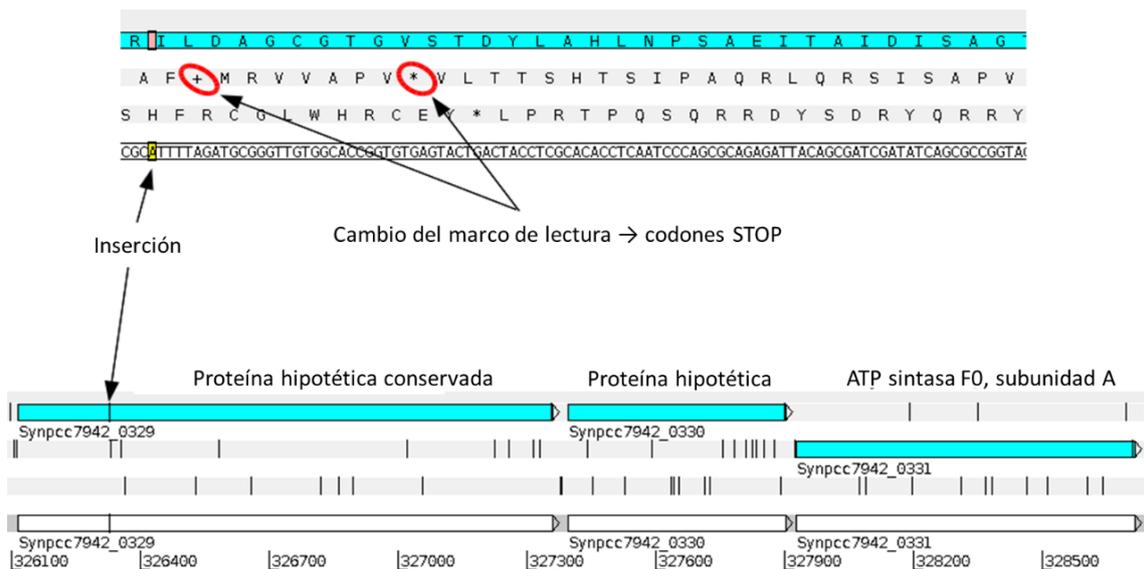


Figura 5-5. Mutación #7 de C11. En la imagen superior se muestra la secuencia de aminoácidos de la proteína y la secuencia de nucleótidos del gen, así como el lugar que ocupa la inserción. Se muestran tres marcos de lectura, correspondiendo el azul al de la proteína silvestre, y los otros dos al de la proteína mutada (el de arriba, si la proteína se transcribe por la cadena directa; el de abajo, si se transcribe por la cadena reversa). En la imagen inferior se muestra la posición que ocupa dicha inserción en el genoma de Se7942. En azul, se muestra el genoma de la cepa silvestre; en blanco, el de la cepa C11. Imagen obtenida de “Artemis”.

5. Resultados

No existe información anotada sobre esta proteína en las bases de datos. En la **Tabla 5-2**, se muestran los resultados obtenidos al realizar una búsqueda usando BLASTp. Como se puede observar, esta proteína está altamente conservada en cianobacterias, ya que todas las proteínas que tienen homología con la de Se7942 pertenecen a dichos organismos. Podría tratarse de una metiltransferasa tipo 12, ya que presenta una alta homología con esta proteína presente en cianobacterias, como por ejemplo, *Oscillatoria* sp. PCC 6506 o *Fischerella* sp. JSC-11, entre otros.

Tabla 5-2. Resultados de BLASTp realizado con la proteína codificada por Synpcc7942_0329.

| Número de acceso | Nombre de la proteína | Organismo | Identidad | E-valor |
|------------------|--|---|-----------|----------|
| Q31RF8_SYNE7 | Proteína hipotética no caracterizada | <i>Synechococcus elongatus</i> (cepa PCC 7942) (<i>Anacystis nidulans</i> R2) | 100 % | 0,0 |
| Y1184_SYNP6 | Proteína no caracterizada syc1184_c | <i>Synechococcus</i> sp. (cepa ATCC 27144 / PCC 6301 / SAUG 1402/1) (<i>Anacystis nidulans</i>) | 100 % | 0,0 |
| Q3M9V3_ANAVT | Proteína hipotética no caracterizada | <i>Anabaena variabilis</i> (cepa ATCC 29413 / PCC 7937) | 62 % | 1,0E-142 |
| Q8Z0S1_NOSS1 | AlI0012 proteína | <i>Nostoc</i> sp. (cepa PCC 7120 / UTEX 2576) | 62 % | 1,0E-141 |
| B4WI79_9SYNE | Dominio de la familia metiltransferasa | <i>Synechococcus</i> sp. PCC 7335 | 61 % | 1,0E-140 |
| D8G4U4_9CYAN | Metiltransferasa tipo 12 | <i>Oscillatoria</i> sp. PCC 6506 | 60 % | 1,0E-139 |
| B4VSN4_9CYAN | Dominio de la familia metiltransferasa | <i>Microcoleus chthonoplastes</i> PCC 7420 | 61 % | 1,0E-139 |
| G6FU31_9CYAN | Metiltransferasa tipo 12 | <i>Fischerella</i> sp. JSC-11 | 61 % | 1,0E-138 |
| F4XLK0_9CYAN | Dominio de la proteína metiltransferasa | <i>Lyngbya majuscula</i> 3L | 60 % | 1,0E-138 |
| A0ZF02_NODSP | Proteína hipotética no caracterizada | <i>Nodularia spumigena</i> CCY9414 | 61 % | 1,0E-137 |
| B8HPB0_CYAP4 | Metiltransferasa tipo 11 | <i>Cyanothece</i> sp. (cepa PCC 7425 / ATCC 29141) | 60 % | 1,0E-135 |
| B5W7S2_SPIMA | Metiltransferasa tipo 11 | <i>Arthrospira maxima</i> CS-328 | 58 % | 1,0E-135 |
| D4ZWE2_SPIPL | Proteína hipotética no caracterizada | <i>Arthrospira platensis</i> NIES-39 | 58 % | 1,0E-134 |
| B2J051_NOSP7 | Metiltransferasa tipo 11 (EC 2.1.1.64) | <i>Nostoc punctiforme</i> (cepa ATCC 29133 / PCC 73102) | 59 % | 1,0E-133 |
| B7KKS6_CYAP7 | Metiltransferasa tipo 12 | <i>Cyanothece</i> sp. (cepa PCC 7424) (<i>Synechococcus</i> sp. (cepa ATCC 29155)) | 59 % | 1,0E-132 |
| Q55818_SYNY3 | SII0487 proteína | <i>Synechocystis</i> sp. (cepa PCC 6803 / Kazusa) | 59 % | 1,0E-131 |
| H0PH27_9SYNC | Proteína hipotética no caracterizada sII0487 | <i>Synechocystis</i> sp. PCC 6803 subcepa PCC-P | 59 % | 1,0E-131 |
| H0PCG0_9SYNC | Proteína hipotética no caracterizada sII0487 | <i>Synechocystis</i> sp. PCC 6803 subcepa PCC-N | 59 % | 1,0E-131 |
| H0P045_9SYNC | Proteína hipotética no caracterizada sII0487 | <i>Synechocystis</i> sp. PCC 6803 subcepa GT-I | 59 % | 1,0E-131 |
| F7UTS4_9SYNC | Proteína hipotética no caracterizada sII0487 | <i>Synechocystis</i> sp. PCC 6803 | 59 % | 1,0E-131 |
| D7E2H7_NOSA0 | Metiltransferasa tipo 11 | <i>Nostoc azollae</i> (cepa 0708) (<i>Anabaena azollae</i> (cepa 0708)) | 57 % | 1,0E-131 |
| B7JXM0_CYAP8 | Metiltransferasa tipo 12 | <i>Cyanothece</i> sp. (cepa PCC 8801) (<i>Synechococcus</i> sp. (cepa PCC 8801 / RF-1)) | 58 % | 1,0E-130 |
| F5UCP6_9CYAN | Metiltransferasa tipo 12 | <i>Microcoleus vaginatus</i> FGP-2 | 57 % | 1,0E-130 |

(Shen et al., 2008) han demostrado que la inactivación de la proteína homóloga de Ss6803 y en Ss7002, no afecta al crecimiento a baja intensidad de luz. Por otro lado, a alta intensidad la cepa silvestre presentaba un tiempo de generación menor que estos mutantes. Por tanto, podríamos descartar esta mutación como responsable del fenotipo de crecimiento rápido en la cepa C11.

5.1.2.3. Mutación #8: ATP sintasa F1, subunidad alfa

La mutación #8 de C12 afecta a la subunidad alfa de la ATP sintasa F1. Esta mutación también se encontró en C11-G₈₂₄ y C12-G₇₂₈, en un porcentaje inferior al 50 % (Tabla 5-1). Sin embargo, en C11 se ha perdido posteriormente, lo contrario de lo que ha ocurrido en C12. La mutación que aparece en este gen es un cambio de un nucleótido, G→T, en la posición 331,632 pb del genoma de Se7942, presente en el 100 % de la población analizada. Esta mutación afecta al aminoácido 252 de esta proteína, dando lugar a un cambio de una cisteína por una fenilalanina.

La función de la ATP sintasa F1 es la producción de ATP a partir de ADP en presencia de un gradiente de protones a través de la membrana. El centro catalítico de esta ATP sintasa F1 está formado por tres subunidades α y tres β . La ATP sintasa se localiza en las membranas tilacoidales como una proteína periférica de membrana.

Como se observa en la Figura 5-6, este cambio se produce en un residuo que no está conservado en las F1 ATPasas de cianobacterias.

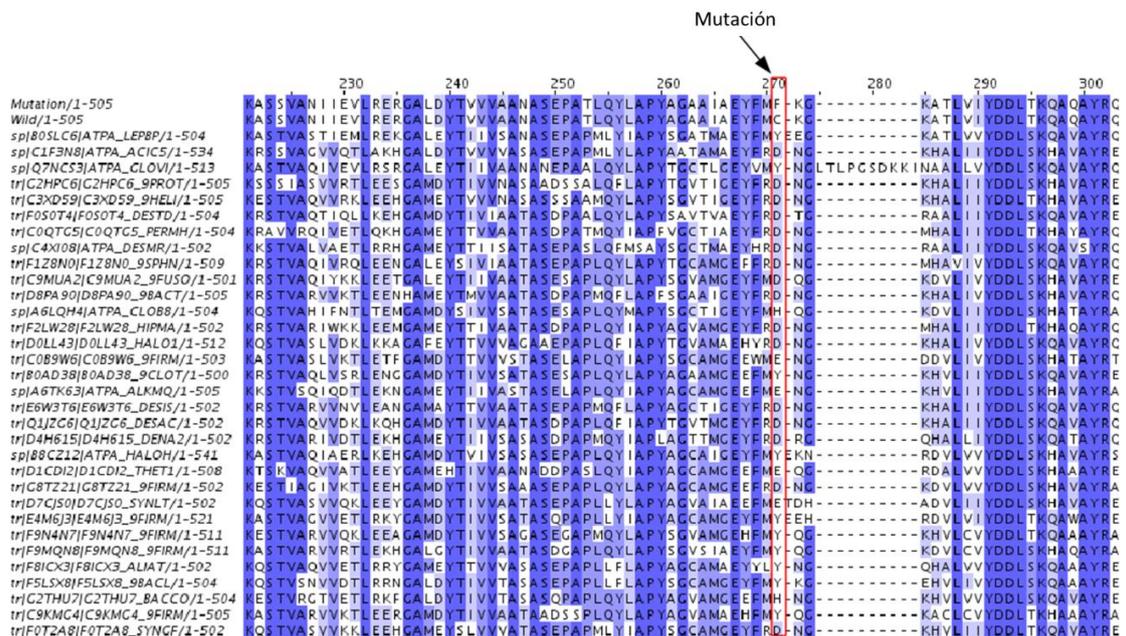


Figura 5-6. Alineamiento de un fragmento de las subunidades alfa de las F1 ATPasas de cianobacterias. En azul oscuro, se representan los aminoácidos conservados entre todas las secuencias; en azul claro, se representan los aminoácidos no conservados en la mayoría de las secuencias; en rojo, la posición del residuo mutado. Realizado con "MUSCLE".

Por otro lado, si atendemos a la estructura tridimensional de esta proteína representada en la **Figura 5-7**, no se aprecian cambios en ninguno de los motivos relevantes, tales como el encargado de la oligomerización o de unión e hidrólisis de ATP (motivo “Walker”).

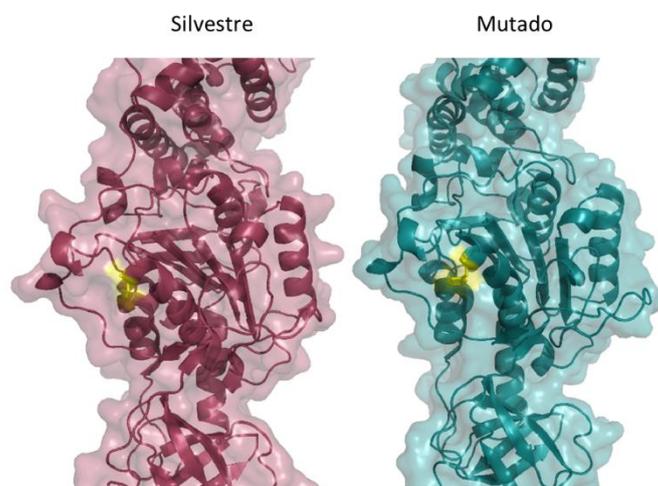


Figura 5-7. Predicción de la estructura tridimensional de la subunidad alfa de la ATPasa F1 de Se7942. Obtenido mediante modelado comparativo utilizando como modelo la estructura 5CDF de la proteína F-ATPasa de *Paracoccus denitrificans* de PDB. A la izquierda se muestra el modelo para la cepa silvestre. A la derecha se muestra el modelo para la cepa mutada. En amarillo, el aminoácido mutado. Realizado con “Modeller”.

5.1.2.4. Mutación #9: Proteína hipotética conservada

La siguiente mutación que analizaremos se encuentra en la cepa C12 y afecta a una proteína hipotética conservada (sin información anotada en GenBank). Esta mutación también se encontraba presente en C11-G₈₂₄ y C12-G₇₂₈, en un porcentaje menor al 50 % (**Tabla 5-1**). Como ocurrió con la mutación #8, en C11 la mutación #9 se ha perdido durante el experimento de evolución y no aparece en la generación 1248.

La mutación que aparece en este gen es un cambio de un nucleótido, una citosina por una adenina, en la posición 441.449 pb del genoma de Se7942 y que está presente en el 100 % de la población analizada. Esta mutación afecta al aminoácido 130 de esta proteína, dando lugar a un codón de STOP temprano (**Figura 5-9**).

Al realizar una búsqueda usando BLASTp, todas las proteínas homólogas son proteínas hipotéticas de otras cianobacterias. Esta familia de proteínas presenta un dominio DUF3155 cuya función se desconoce. La alta conservación entre especies de cianobacterias podría sugerir conservación de función. Sin embargo, (Rubin et al., 2015) determinaron que este gen no era esencial en Se7942 creciendo a 37 °C y en condiciones de baja intensidad de luz (70 μmol

fotones·m⁻²·s⁻¹). Como se observa en la **Figura 5-8**, el gen Synpcc7942_0452 no forma parte de ninguno operón, puesto que los genes que componen su entorno genético son muy diversos.

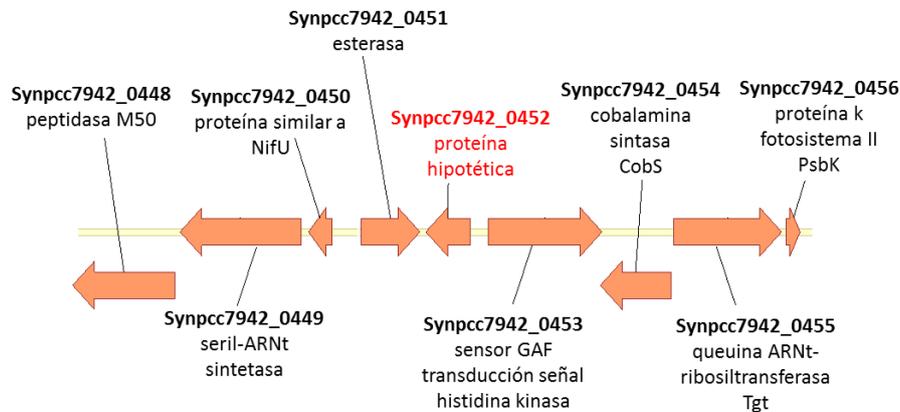


Figura 5-8. Genes que componen el entorno genético de la mutación número 9 que afecta a una proteína hipotética. En rojo, se muestra el gen en el que se encuentra la mutación descrita. Imagen tomada del programa “Vector NTI” (Lu and Moriyama, 2004).

Como se observa en la **Figura 5-9**, el mutante C12 presenta una deleción de 30 aminoácidos en el extremo C-terminal de esta proteína. Por otro lado, en dicho alineamiento, se observa un motivo altamente conservado entre cianobacterias. La deleción se encuentra en una zona no conservada entre ellas, por lo que podría cambiar la actividad pero no la función. En este caso, no fue posible construir un modelo estructural de esta proteína, ya que no se dispone de la estructura de una proteína similar para llevar a cabo una predicción estructural.

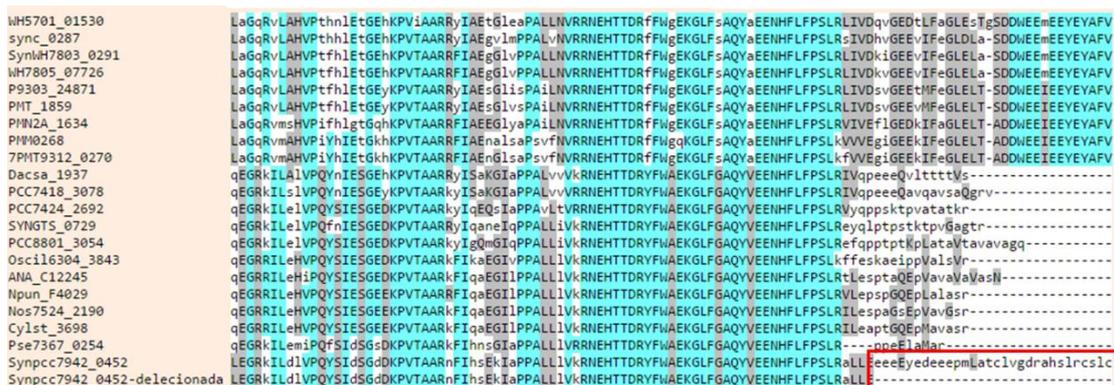


Figura 5-9. Alineamiento de un fragmento de la proteína silvestre (Synpcc7942_0452) y la mutada en C12 frente a proteínas homólogas de cianobacterias. La secuencia de la proteína silvestre es la penúltima y la de la proteína mutada es la última. Sombreado en azul, se representan los aminoácidos conservados entre todas las secuencias; en gris, los aminoácidos conservados en la mayoría de las secuencias; en rojo, la zona deletada en C12. El set de proteínas homólogas de cianobacterias fue tomado de la base de datos “UNIPROT”. El alineamiento fue realizado con “MUSCLE”.

5.1.2.5. Mutación #12: Promotor de la shikimato kinasa

La mutación #12 de C11 se localiza en una región intergénica. Esta mutación apareció en C11-G₈₂₄ y en C12-G₇₂₈ con baja abundancia, 9,40 % y 1,90 %, respectivamente (**Tabla 5-1**). La mutación #12 desapareció de la población C12 durante el experimento de evolución después de la generación 728.

La mutación #12 afecta al promotor de la shikimato kinasa (*aroK*, Synpcc7942_0094). (Vijayan et al., 2011) ya describen esta región como un promotor, proponiendo como punto de inicio de la transcripción la posición 905.450 pb, y el inicio del gen en la posición 905.417 pb (**Figura 5-10**). La mutación se encuentra en la posición 905.445 pb, a -28 pb del codón de inicio, por tanto estaría incluida dentro del promotor. La población C11 presenta esta mutación con una abundancia del 100 %, y da lugar a un cambio de un nucleótido, una citosina por una timina. Este cambio podría estar afectando a la unión de la ARN polimerasa o del ribosoma, y por ello a la expresión del gen.

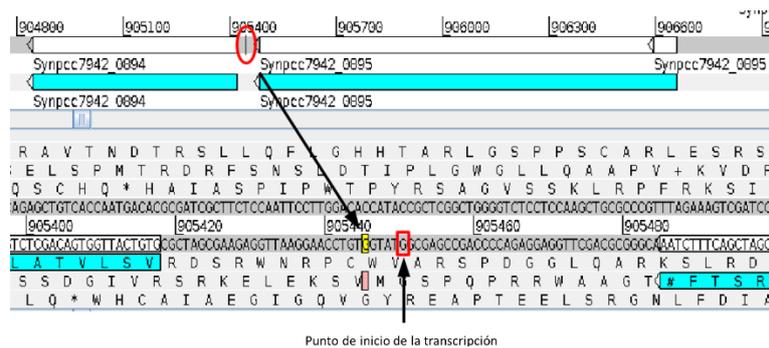


Figura 5-10. Región intergénica con la mutación #12 de C11. En rojo, se muestra la posición del SNP en el extremo 5' del ARNm de la shikimato kinasa. Imagen obtenida de "Artemis".

La shikimato kinasa cataliza la quinta reacción en la ruta del corismato, convirtiendo el shikimato en shikimato 3-fosfato, tal y como se observa en la **Figura 5-11**. La ruta del shikimato tiene como producto final los 3 aminoácidos aromáticos (fenilalanina, triptófano y tirosina).

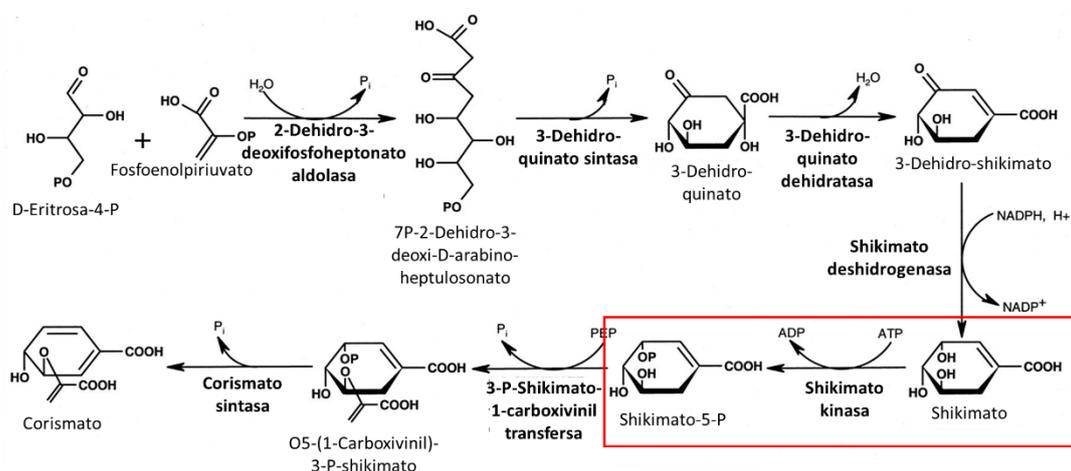


Figura 5-11. Ruta de síntesis del corismato. En el recuadro en rojo aparece la reacción catalizada por la shikimato kinasa. Modificado de (Daugherty et al., 2001)-.

5.1.2.6. Mutaciones #13 y #14: Ácido graso de cadena larga-CoA ligasa

Las poblaciones C11 y C12 presentan un SNP en el gen que codifica para la ácido graso de cadena larga-CoA ligasa. En C11-G₁₂₄₈, este SNP aparece en la posición 924.724 pb con una frecuencia del 100 %, aunque no estaba presente en C11-G₈₂₄. El cambio que tiene lugar (C→G) produce un cambio de aminoácido en la posición 216 de una prolina por una serina. El SNP aparece en la posición 924.725 pb con una frecuencia del 100 % en C12-G₁₁₆₈ y del 20,10 % en C12-G₇₂₈. Al igual que en el caso anterior, el cambio que tiene lugar es C→G, que provoca un cambio también en el aminoácido 216 de una prolina por una leucina (**Tabla 5-1**).

Este gen codifica para FadD, una enzima esencial en el metabolismo de los ácidos grasos y cuya función se describe en la introducción (apartados **2.3.1.2** y **2.3.2.2**). Su función es convertir los ácidos grasos libres en ésteres grasos acil-CoA (intermediarios en la síntesis de lípidos complejos), siendo su actividad mayor cuando los sustratos son ácidos grasos de más de 10 átomos de carbono. Esta reacción es importante para la biogénesis de la membrana y para el almacenamiento de energía.

Se realizó un alineamiento múltiple para observar el grado de conservación del aminoácido mutado entre cianobacterias. Como se observa en la **Figura 5-12**, este aminoácido pertenece a un motivo muy conservado aunque el aminoácido de dicha posición es variable. Por otro lado, no se encontró en ninguna de las secuencias una serina o una leucina en esta posición.

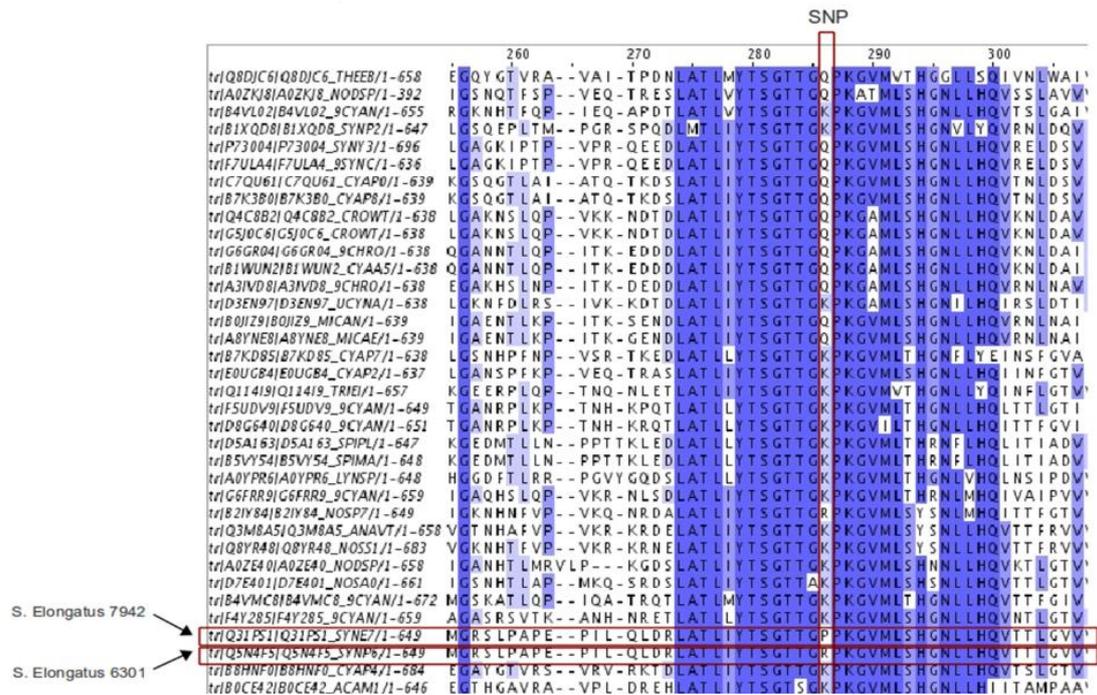


Figura 5-12. Alineamiento de la secuencia de aminoácidos de varias proteínas FadD procedentes de cianobacterias. En azul oscuro, se representan los aminoácidos conservados entre todas las secuencias; en azul claro, se representarían los aminoácidos conservados en la mayoría de las secuencias; en rojo, la posición del residuo mutado. Realizado con “MUSCLE”.

En la **Figura 5-13**, se ha modelado la estructura tridimensional de esta proteína, observándose la posición del motivo altamente conservado. También se ha modelado la posición de la mutación en las cepas evolucionadas, así como la distribución de su carga e hidrofobicidad. Como se observa en la figura, la presencia de las mutaciones #13 y #14 no afectaría ni a la distribución de carga, ni a la hidrofobicidad, por lo que la proteína mutada podría funcionar correctamente. Además, la mutación se encuentra en un “loop” y no en una hélice alfa u hoja beta, lo que podría explicar el hecho de que la mutación no tenga ningún efecto sobre la estructura tridimensional de la proteína.

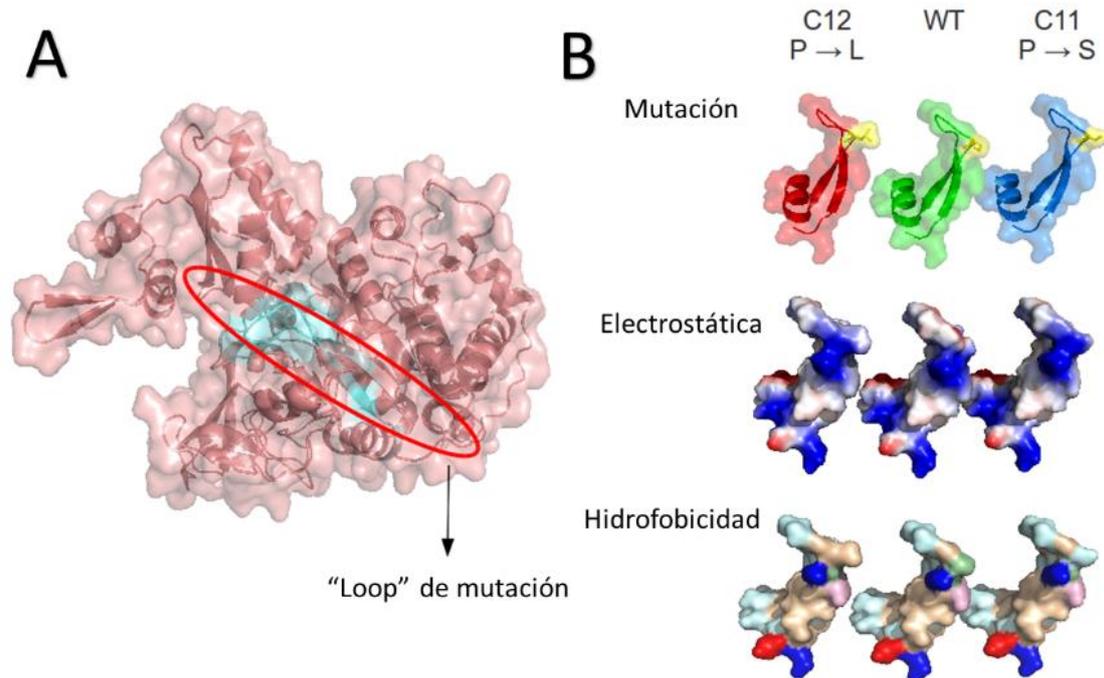


Figura 5-13. Modelo de la estructura tridimensional de la ácido graso de cadena larga-CoA ligasa. (A) Modelo de la estructura de la proteína silvestre FadD obtenido mediante modelado comparativo utilizando como modelo la estructura 1PG4 de la ácido graso de cadena larga-CoA ligasa de *Salmonella enterica* de PDB. En azul, el motivo altamente conservado entre cianobacterias representado en la **Figura 5-12**. (B) Figura detallada del motivo conservado en la cepa silvestre y en las cepas evolucionadas C11 y C12. En la imagen superior, se muestra en amarillo la posición de la mutación dentro del motivo. En la imagen del medio, se muestra la distribución de la carga dentro del motivo (en rojo, residuos con carga negativa; en azul, residuos con carga positiva; en blanco, residuos sin carga). En la imagen inferior, se muestra la distribución de la hidrofobicidad dentro del motivo (en azul claro, residuos menos hidrofílicos; en azul oscuro, residuos más hidrofílicos; en rosa, residuos menos hidrofóbicos; en rojo, residuos más hidrofóbicos). Realizado con "Modeller".

5.1.2.7. Mutación #16: Región intergénica entre dos genes que codifican para un polipéptido de unión al ficobilisoma

Tal y como ocurría con la mutación #12, la mutación #16 se encuentra en una región intergénica de la población C12. Esta mutación también aparecía en C11-G₈₂₄ y en C12-G₇₂₈ con una abundancia en torno al 40 %, perdiéndose en C11-G₁₂₄₈ (**Tabla 5-1**). En la mutación #16 se produce un transversión de C→A en la posición 1.062.322 pb. Como puede verse en la **Figura 5-14**, la mutación se encuentra entre dos genes (Synpcc7942_1049 y Synpcc7942_1050) que codifican para dos polipéptidos de unión al ficobilisoma. La secuencia nucleotídica presenta una identidad del 58 % entre ambos genes, mientras que las proteínas codificadas presentan una identidad del 49 %. La mutación se encuentra a 35 nucleótidos del codón de inicio del gen Synpcc7942_1050. Comparando esta situación con la shikimato kinasa, podríamos pensar que esta mutación afectaría al promotor del gen Synpcc7942_1050. Sin embargo, analizando los datos del transcriptoma de Se7942 (Vijayan et al., 2011), determinamos que esta región

5. Resultados

intergénica a un ARNm policistrónico de los genes Synpcc7942_1048, Synpcc7942_1049 y Synpcc7942_1050, no a un promotor. Para determinar si la mutación afecta a la expresión del ARNm, sería necesario realizar RNAseq. Estos resultados se explican en el apartado 5.1.6.

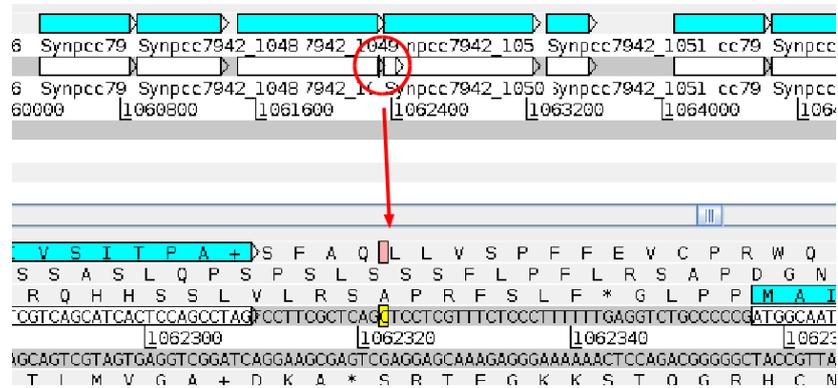


Figura 5-14. Representación de la región intergénica con la mutación #16. En rojo, se muestra la posición del SNP. Imagen obtenida de "Artemis".

La función de estos polipéptidos es el acoplamiento de la ficocianina a la aloficocianina en los núcleos de los ficobilisomas. También determinan el estado de agregación y la localización de las unidades de ficobiliproteínas en forma de disco dentro de los ficobilisomas y modulan sus propiedades espectroscópicas para mediar en una transferencia de energía óptima y directa (Figura 5-15).

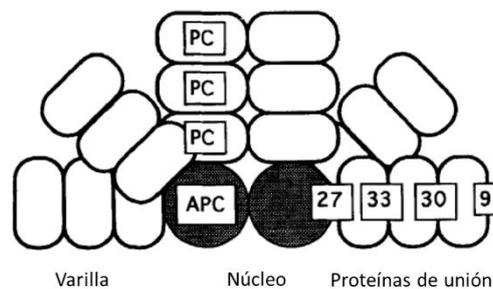


Figura 5-15. Representación esquemática de un ficobilisoma de Se7942. (Guan et al., 2007) determinaron que la subunidad de 27 kDa estaba codificada por el gen Synpcc7942_2030, la de 33 kDa por el gen Synpcc7942_1049, la de 30 kDa por el gen Synpcc7942_1048 y la de 9 kDa por el gen Synpcc7942_1050. Los genes Synpcc7942_1049 y Synpcc7942_1050 pertenecen al mismo operón junto al gen Synpcc7942_1048. El gen Synpcc7942_2030, en cambio, se encuentra aguas arriba de esta región. PC, ficocianina; APC, aloficocianina. Los números que aparecen a la derecha en la imagen corresponden al peso molecular (en kDa) de los polipéptidos de unión al ficobilisoma. Modificado de (Bhalerao et al., 1995).

5.1.2.8. Mutación #19: Proteína de ensamblaje a fimbrias, similar a PilC

La mutación #19 se encuentra en el 100 % de la población de C12-G₁₁₆₈ y afecta a una proteína de ensamblaje a fimbrias similar a PilC. Esta mutación también aparecía en C11-G₈₂₄ y

C12-G₇₂₈, en un porcentaje menor al 40 %, no encontrándose en C11-G₁₂₄₈ (Tabla 5-1). La mutación #19 es un cambio de un nucleótido de A→C en la posición 2.146.923 pb. Afecta al aminoácido 85 de esta proteína de 410 aminoácidos, dando lugar a un cambio de una Val→Gly.

La Figura 5-16 muestra un alineamiento de la secuencia de la proteína silvestre y de la proteína mutada de Se7942, frente a la secuencia de homólogos en otras cianobacterias. El aminoácido mutado en la cepa C12 está altamente conservado entre cianobacterias, por lo que la mutación podría afectar a la actividad de la proteína.



Figura 5-16. Alinamiento de un fragmento de la proteína silvestre y la mutada frente a proteínas homólogas de cianobacterias. Los nombres de la proteína silvestre y de la proteína mutada se indican con un recuadro verde. Con un recuadro rojo se indica la posición que ocupa la mutación en la secuencia. En azul, se representan los aminoácidos conservados entre todas las secuencias; en gris, los aminoácidos conservados en la mayoría de secuencias. El set de proteínas homólogas fue tomado de la base de datos "UNIPROT". El alineamiento fue realizado con "MUSCLE".

Esta proteína de ensamblaje a fimbrias, similar a PilC, está involucrada en la translocación de pilinas tipo IV (PilA) desde el citoplasma al espacio periplásmico. Se encuentra en la membrana interna de las células y podría actuar como una proteína de membrana politópica, atravesando varias veces la membrana. Es necesaria para en la competencia para la transformación de ADN en cianobacterias como Se7942 (Schatz et al., 2013) y Ss6803 (Bhaya et al., 2000). Estos mismos autores determinaron que la disrupción de la proteína PilC, mediante su truncamiento con un gen de resistencia a espectinomicina, suprime la competencia natural. Dado que la mutación afecta a una proteína involucrada en la competencia para la transformación, podría afectar a dicha función.

En este caso, no pudimos construir un modelo estructural con suficiente valor de confianza para determinar si la mutación afectaba a la estructura de la proteína, así como a otras características de la misma.

5.1.2.9. Mutación #20: Proteína de motilidad “Twitching”

La mutación #20 se encuentra en la proteína de motilidad “Twitching” (PilT) en C12-G₁₁₆₈ con una abundancia del 100 % (Tabla 5-1). No apareció en C12-G₇₂₈, ni en ninguna de las poblaciones C11. La mutación #20 es un cambio de un nucleótido (G→C) en la posición 2.146.923 pb del genoma de Se7942. Afecta al aminoácido 178 de esta proteína, de 362 aminoácidos, dando lugar a un cambio de una Arg→Gly.

La Figura 5-17 muestra el alineamiento entre la proteína silvestre y la proteína mutada. El aminoácido mutado está muy conservado en las cianobacterias analizadas, por lo que la mutación podría afectar a la actividad de la proteína.

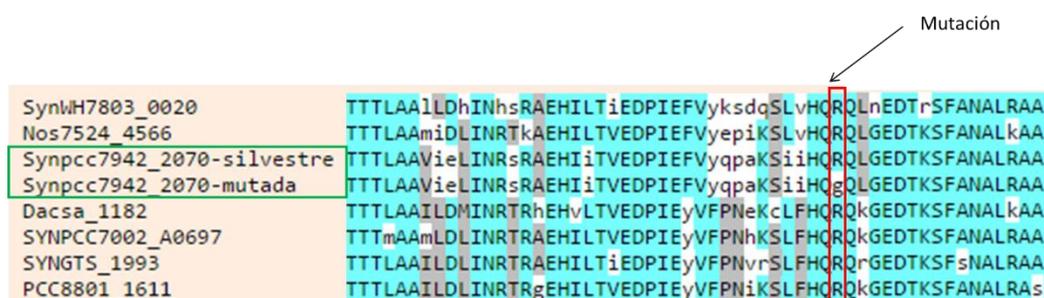


Figura 5-17. Alineamiento de la proteína PilT silvestre y la mutada frente a proteínas homólogas en otras cianobacterias. Las secuencias de la proteína silvestre y de la mutada se indican con un recuadro verde. Con un recuadro rojo se indica la posición que ocupa la mutación en la secuencia. En azul, se representan los aminoácidos conservados entre todas las secuencias; en gris, se representan los aminoácidos conservados en la mayoría de las secuencias. El grupo de proteínas homólogas fue tomado de la base de datos “UNIPROT”. El alineamiento fue realizado con “MUSCLE”.

PilT es una ATPasa de función relacionada con la retracción del pilus. Se localiza en el citoplasma y es requerida tanto para la motilidad tipo “Twitching” como para la competencia en transformación.

Como se observa en la Figura 5-18, A, la mutación #20 se encuentra en una hoja beta. La proteína PilT de Se7942 presenta cuatro regiones conservadas entre homólogos bacterianos, motivo “Walker A” (aminoácidos 132 a 140), caja Asp (aminoácidos 157 a 165), motivo “Walker B” (aminoácidos 250 a 256) y caja His (aminoácidos 272 a 280) (Okamoto and Ohmori, 2002). Estos cuatro motivos están asociados con la unión e hidrólisis de ATP (Jakovljevic et al., 2008). Aunque la mutación se encuentra fuera de estas cuatro regiones pudiera ser esencial, ya que está altamente conservado en cianobacterias.

La distribución de cargas en la proteína silvestre y la mutada es diferente (Figura 5-18, B). El residuo mutado cambia la distribución de carga en zonas adyacentes al mismo. De esta

manera, disminuyen las zonas de carga positiva cercanas al residuo mutado, y aumentan las zonas de carga negativa cercanas al mismo. En la **Figura 5-18, C**, se observa la posición y el tamaño del residuo mutado, siendo este residuo (Gly) menos voluminoso que el de la proteína silvestre (Arg).

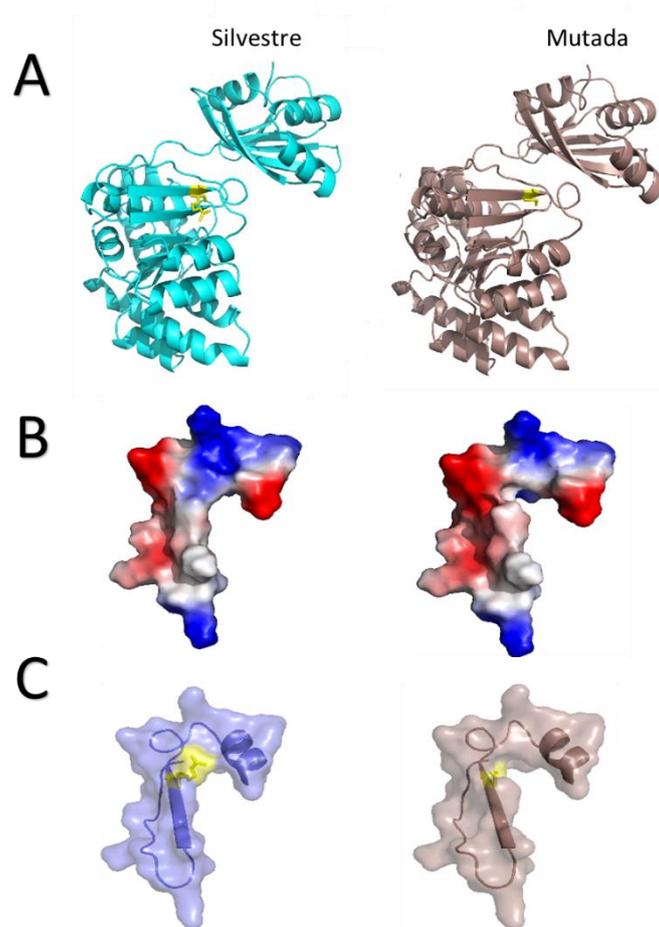


Figura 5-18. Modelado de la proteína de motilidad "Twitching", PilT. (A) estructura tridimensional de la proteína silvestre y mutada utilizando como modelo la estructura 3JVV de la proteína PilT de *Pseudomonas aeruginosa* de PDB. En amarillo, el residuo mutado. **(B)** Distribución de cargas en las proteínas silvestre y mutada. En rojo, residuos con carga negativa; en azul, residuos con carga positiva; en blanco, residuos sin carga. **(C)** Posición y tamaño del residuo mutado en las proteínas silvestre y mutada. En amarillo, el residuo mutado. Realizado con "Modeller".

5.1.2.10. Mutación #21: Histidina kinasa

La mutación #21 aparece al 100 % en la población C11-G₁₂₄₈ y afecta a la histidina kinasa SasA. También aparecía en C11-G₈₂₄ con una abundancia del 12,30 %, pero no en ninguna población C12 (**Tabla 5-1**). La mutación #21 es un cambio de C→T en la posición 2.195.880 pb del genoma de Se7942. Afecta al aminoácido 352 de esta proteína de 387 aminoácidos, dando lugar a un codón de STOP temprano.

Como se observa en la **Figura 5-19**, la delección afecta a la región C-terminal de la proteína SasA, eliminándose 34 aminoácidos. Esta región está altamente conservada en cianobacterias, por lo que su eliminación podría dar lugar a pérdida de actividad de la proteína.



Figura 5-19. Alineamiento del extremo C-terminal de la proteína silvestre SasA y la mutada frente a proteínas homólogas de cianobacterias. Las secuencias de las proteínas silvestre y mutada se indican con un recuadro verde. Con un recuadro rojo se indica la posición de la delección. En azul, se representan los aminoácidos conservados entre todas las secuencias; en gris, aquellos aminoácidos conservados en la mayoría de las secuencias. El grupo de proteínas homólogas fue tomado de la base de datos “UNIPROT”. El alineamiento fue realizado con “MUSCLE”.

El gen *sasA* codifica para una histidina quinasa que, junto con un regulador de respuesta codificado por *rpaA*, forma un sistema regulatorio de dos componentes que ayuda a mantener el control circadiano de la expresión génica (Takai et al., 2006). Respecto a la asociación con el oscilador Kai, la información temporal fluye desde la proteína del ritmo circadiano KaiC a SasA. KaiC mejora el ratio de autofosforilación de SasA en su residuo conservado de histidina 161 (aunque SasA no influye en la actividad autokinasa de KaiC) (Smith and Williams, 2006). SasA transfiere su grupo fosforilo a RpaA, que se activa. La diana de RpaA no se conoce (Mackey and Golden, 2007).

SasA presenta un dominio N-terminal de interacción con KaiC (aminoácidos 1-97) y un segundo dominio C-terminal histidina quinasa (aminoácidos 158-382). Como la delección se encuentra en la región C-terminal de la proteína (a partir del aminoácido 352), podría afectar a la actividad del dominio histidina quinasa. El residuo de histidina-161, que autofosforila a SasA no está afectado por la delección. Tampoco afectaría a la interacción con KaiC.

La **Figura 5-20** muestra un modelo tridimensional de la proteína SasA de la cepa silvestre y la delección #21. La delección afecta tanto a una hoja beta, como a una hélice alfa y un lazo, que desaparecen en la proteína mutada.

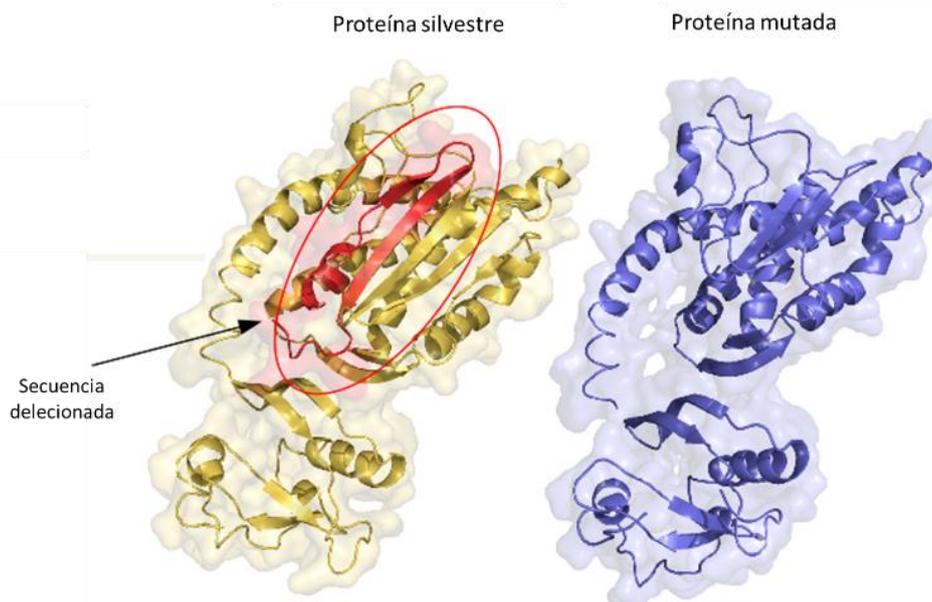


Figura 5-20. Modelo estructural de la proteína SasA silvestre y la deleción #21. Obtenido mediante modelado comparativo utilizando como modelo la estructura 4U70 de la histidina kinasa de *Lactobacillus plantarum* de PDB. A la izquierda, en amarillo, se representa a la proteína silvestre; a la derecha, en azul, se representa a la proteína con la deleción; en rojo, la zona delecionada. Realizado con "PHYRE2".

5.1.2.11. Mutación #24: Región intergénica entre un gen que codifica para una proteína hipotética conservada y un gen que codifica para la ARNt i(6)A37 tiotransferasa MiaB

La mutación #24 se encuentra en una región intergénica de C12-G₁₁₆₈, cuya abundancia en la población C12-G₁₁₆₈ es del 80,70 %. También estaba presente en C11-G₈₂₄ y en C12-G₇₂₈ con una abundancia menor al 41 %, pero no en C11-G₁₂₄₈ (**Tabla 5-1**). La mutación #24 es una inserción de 18 nucleótidos en la posición 2.441.393 pb del genoma de Se7942. Esta región intergénica pertenece a un operón de ARNm, no a un promotor, según los datos del transcriptoma de Se7942 (Vijayan et al., 2011). Para determinar si la mutación afecta a la expresión del ARNm, sería necesario realizar RNAseq. Estos resultados se explican en el apartado 5.1.6.

5.1.3. Comprobación de la presencia de las mutaciones identificadas en la cepa C11-G₁₂₄₈

Dado que C11-G₁₂₄₈ presentaba tan solo cinco mutaciones frente a las nueve mutaciones de C12-G₁₁₆₈, seleccionamos C11-G₁₂₄₈ para identificar las mutaciones responsables del

crecimiento rápido. Con tan solo cinco mutaciones, C11-G₁₂₄₈ ya presenta una tasa de crecimiento máximo respecto a la cepa silvestre, por lo que sería más fácil identificar cuál(es) de ellas eran responsables de dicho fenotipo.

Resecuenciamos mediante el método “Sanger” las 5 mutaciones de C11-G₁₂₄₈, para confirmar los cambios observados en la secuenciación “Illumina”. Se realizó una PCR de cada una de las regiones de C11-G₁₂₄₈ que tenía la mutación, y una PCR extra de una región que contiene una mutación presente en las cepas control del experimento de evolución, la mutación #23. Para amplificar la región de la mutación #2 se usaron los cebadores 89 y 90, para la #7 los cebadores 91 y 92, para la #12 los cebadores 87 y 88, para la #13 los cebadores 85 y 86, para la #21 los cebadores 93 y 94 y para la #23 los cebadores 95 y 96. Los tamaños de banda que se obtienen son de 538 pb, 975 pb, 950 pb, 667 pb, 1094 pb y 1010 pb, respectivamente (**Figura 5-21**). En el caso de la región de la mutación #21, se observan dos bandas en el gel de electroforesis, correspondiendo a esta región únicamente la banda superior.

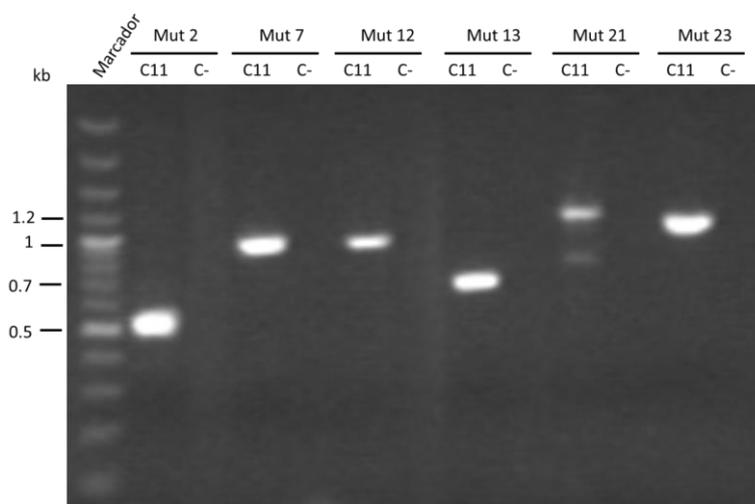


Figura 5-21. Amplificación mediante de PCR de las 5 regiones del genoma de C11-G₁₂₄₈ que contienen las mutaciones responsables de su fenotipo. Para cada mutación se realizó una PCR de un cultivo líquido de C11 y se usó un control negativo sin ADN. Se usó como marcador “Generuler 100 pb DNA ladder plus” (Fermentas).

Los fragmentos obtenidos fueron secuenciados mediante el método de “Sanger”, con los mismos oligonucleótidos utilizados en la amplificación (**Figura 5-22**). De las mutaciones presentes en C11-G₁₂₄₈ todas ellas estaban presentes, salvo la mutación #13 que afecta a la ácido graso de cadena larga-CoA ligasa (**Figura 5-22, D**). Esta mutación se encontraba en la población C11-G₁₂₄₈ con una frecuencia del 100 %, por lo que se podría tratarse de un error durante la secuenciación mediante “Illumina”. Para descartar que esto hubiera ocurrido, decidimos analizar esta región en varias colonias aisladas de la población C11 mediante secuenciación

“Sanger”, para determinar si el problema estaba en la secuenciación o verdaderamente la mutación no estaba presente en la población C11.



Figura 5-22. Alineamiento de las regiones que contienen las mutaciones presentes en C11-G₁₂₄₈. Alineamiento realizado frente al genoma de la cepa silvestre Se7942 anotado en GenBank (NC_007604). En rojo, se indica la posición que ocupa o debería ocupar la mutación. Alineamientos realizados con “AlignX” una aplicación del programa “Vector NTI”.

El siguiente paso fue la amplificación y secuenciación de la región que contendría la mutación #13 de ocho colonias de C11-G₁₂₄₈ con los cebadores 85 y 86 (**Figura 5-23**). Las cuatro primeras colonias analizadas se obtuvieron de un cultivo congelado, que se refrescó y plaqueó justo antes de realizar la PCR. Las otras cuatro colonias procedían de una placa de C11 que se había cultivado seriadamente en el laboratorio. Con este experimento quisimos descartar que esta mutación se hubiera perdido durante el cultivo seriado en el laboratorio, del que se había obtenido el ADN molde para obtener el amplicón cuya secuencia se muestra en **Figura 5-22**, mutación #13. Sin embargo, ninguna colonia contenía la mutación (**Figura 5-23**). Por tanto, no ha sido el cultivo seriado lo que originó la pérdida de dicha mutación, simplemente esta mutación no estaba desde el inicio del experimento y es un error de la secuenciación “Illumina”.

5. Resultados

| | 620 | 630 | 640 | 650 | 660 | 670 | 680 | 690 | 700 |
|-----------|-----|---|-----|-----|-----|-----|-----|-----|-----|
| LCFA | 595 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 1 | 608 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 2 | 462 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 3 | 472 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 4 | 527 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 5 | 557 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 6 | 278 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 7 | 262 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |
| colonia 8 | 278 | TTGCAGCTCGATCGCTTAGCCACTTTGATCTATACTCGGGCACCACAGG | | | | | | | |

Figura 5-23. Alineamiento de la región que contiene la mutación #13. Análisis de dicha región en ocho colonias de C11-G₁₂₄₈ alineadas frente al genoma de la cepa silvestre Se7942 anotado en GenBank (NC_007604). En rojo, se indica la posición que ocupa o debería ocupar la mutación en esa región. Alineamientos realizados con “AlignX” una aplicación del programa “Vector NTI”.

En resumen, de las cinco mutaciones detectadas mediante la secuenciación “Illumina”, solo cuatro de ellas están presentes en C11-G₁₂₄₈. Estas mutaciones son las que analizaremos para ver si son responsables del fenotipo de crecimiento rápido.

5.1.4. Clonación de las mutaciones en la cepa silvestre de Se7942

El método seleccionado para introducir las mutaciones en la cepa silvestre de Se7942 (Matsuoka et al., 2001), se explica en el apartado 4.8.4. Se basa en la cepa GRPS1, que contiene una mutación en el gen *rps12* que le confiere resistencia a estreptomycin. Aunque en nuestro laboratorio disponíamos de esta cepa, no la utilizamos para introducir las mutaciones, ya que su genoma podría contener otras mutaciones que afectaran o enmascararan el efecto de las mutaciones a analizar.

El primer paso en el proceso de identificación del efecto de las mutaciones, fue la modificación de la cepa Se7942 para conferirle resistencia a estreptomycin. Para ello, amplificamos mediante PCR la región que contenía la mutación en el gen *rps12* de la cepa GRPS1 con los cebadores 5 y 6. Cortamos con enzimas de restricción dicho fragmento y lo clonamos en el vector pSB1C3, dando lugar al vector pMSM2 (Tabla 4-3). Se secuenció dicha construcción y el fragmento de PCR utilizando los cebadores 97 y 98 para comprobar la presencia de la mutación en el gen *rps12*.

A continuación, introdujimos por separado en Se7942, mediante transformación natural, el vector pMSM2 y el fragmento de PCR que contiene el gen *rps12*-R43 (Figura 5-24). Los resultados de (Matsuoka et al., 2001) mostraban que, aunque las cianobacterias no degradaban el ADN lineal, la eficacia de reemplazamiento de genes usando ADNs lineales era menos efectiva que la de ADNs circulares. Esto explicaría el hecho de que obtuvimos menos colonias cuando se

utilizó el fragmento de PCR en la transformación natural que cuando se utilizó el vector circular, a pesar de usar en ambos casos la misma cantidad de ADN en el proceso de transformación.

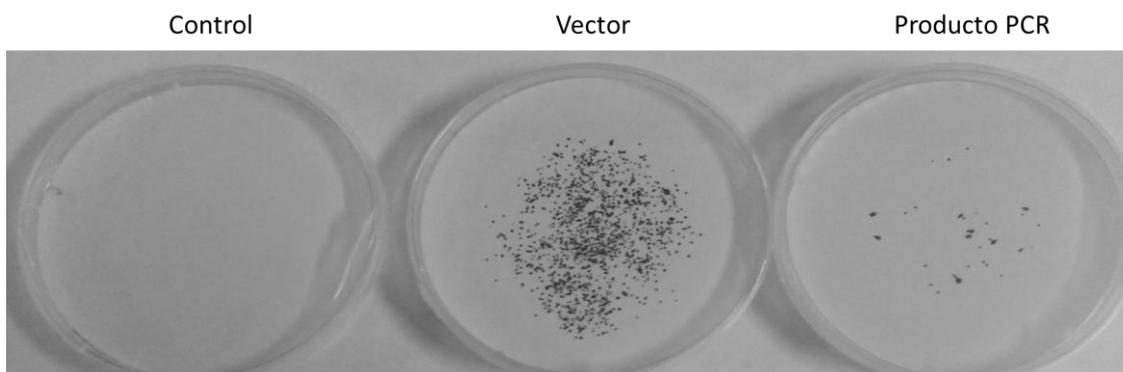


Figura 5-24. Resultados de la transformación natural para introducir la mutación en el gen *rps12* de Se7942. A la izquierda, se muestra la transformación control en la que no se ha añadido ADN. En el centro, la transformación utilizando el vector pMSM2. A la derecha, la transformación utilizando el producto de PCR que contiene dicha mutación.

Se subcultivaron varias colonias de cada uno de los transformantes dos veces en medio sólido suplementado con Sm50. Se amplificó y secuenció con los cebadores 5 y 6 la región que contenía el gen *rps12-R43* de dos transformantes obtenidos con el fragmento de PCR y otros dos de la transformación con el vector (**Figura 5-25**). De las cuatro colonias analizadas solo dos contenían la mutación, una procedente de la transformación con el fragmento de PCR y otra de la transformación con el vector. Esta mutación es un cambio de A→G en la posición 128 pb del gen, que da lugar a un cambio de aminoácido de Lys→Arg en el residuo 43 (apartado 4.8.4). La ausencia de esta mutación en alguna de las colonias, se debe a que la presencia de Sm50 en el medio favorece la aparición de otras mutaciones en este mismo gen y que confieren resistencia a estreptomycin. Concretamente, estas colonias presentaban un cambio en el nucleótido 167 de G→C, que daba lugar a un cambio en el residuo 56 de esta proteína Rps12 de Arg→Pro.

| | 88 | 90 | 100 | 110 | 120 | 130 | 140 | 150 | 160 |
|-------------------|----|-----------------|-----------------|------------|-----------------|-------|-------------|-----------|-----|
| <i>rps12</i> | 88 | CGGCGTGGCGTTTGC | ACTCGCGTCTACACC | ACGACCCCGA | AAAGCCCAACTCGGC | CTTGC | CGGAAAGTGGC | ACGGGTTCC | |
| colonia 9 PCR | 88 | CGGCGTGGCGTTTGC | ACTCGCGTCTACACC | ACGACCCCGA | AAAGCCCAACTCGGC | CTTGC | CGGAAAGTGGC | ACGGGTTCC | |
| colonia 10 PCR | 88 | CGGCGTGGCGTTTGC | ACTCGCGTCTACACC | ACGACCCCGA | AAAGCCCAACTCGGC | CTTGC | CGGAAAGTGGC | ACGGGTTCC | |
| colonia 5 vector | 88 | CGGCGTGGCGTTTGC | ACTCGCGTCTACACC | ACGACCCCGA | AAAGCCCAACTCGGC | CTTGC | CGGAAAGTGGC | ACGGGTTCC | |
| colonia 11 vector | 88 | CGGCGTGGCGTTTGC | ACTCGCGTCTACACC | ACGACCCCGA | AAAGCCCAACTCGGC | CTTGC | CGGAAAGTGGC | ACGGGTTCC | |

Figura 5-25. Alineamiento de la secuencia que contiene la mutación en el gen *rps12*. Dos colonias obtenidas de la transformación utilizando el fragmento de PCR (colonia 9 y 10) y dos utilizando el vector (colonias 5 y 11), alineadas frente al genoma de la cepa silvestre Se7942 anotado en GenBank (NC_007604). En rojo, se indica la posición que ocupa o debería ocupar la mutación en esa región. Alineamientos realizados con "AlignX" una aplicación del programa "Vector NTI".

Para comprobar que en el caso de la transformación natural usando el vector, la inserción de la región con el gen *rps12* mutado había tenido lugar mediante una recombinación doble, realizamos una PCR usando los cebadores 113 y 114 que amplifican el gen de resistencia a

cloranfenicol del vector. En la **Figura 5-26, A**, se observa una banda de amplificación muy marcada en el control positivo, que es el vector pMSM2, usado en la transformación natural para introducir la mutación. Sin embargo, en las dos colonias analizadas se observa también una banda tenue a la altura del control positivo. Como este resultado era un poco confuso, decidimos analizar también la resistencia a cloranfenicol de estos mutantes, tal y como se observa en la **Figura 5-26, B**. Se cultivaron los transformantes y la cepa silvestre en medio BG11 líquido en concentraciones crecientes de dicho antibiótico. Tanto los mutantes como la cepa silvestre crecen a las mismas concentraciones de cloranfenicol y a partir de 3 $\mu\text{g/ml}$ ninguno de ellos crece. En los derivados de Se7942 resistentes a dicho antibiótico, utilizamos una concentración de cloranfenicol 5 $\mu\text{g/ml}$, tal y como hemos explicado en el apartado **4.8.1** de materiales y métodos. Esto confirma que la mutación se ha introducido por recombinación doble y que el vector usado en la transformación natural no se integró en el cromosoma.

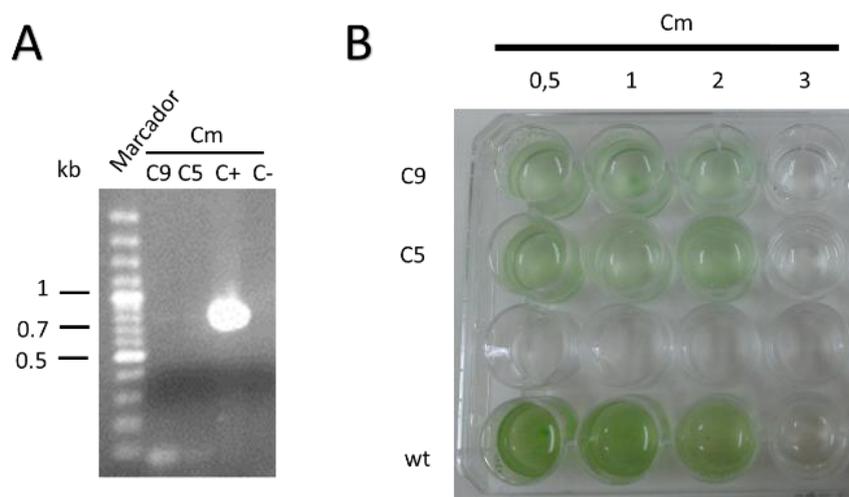


Figura 5-26. Análisis de los transformantes con la mutación en el gen *rps12* que confiere resistencia a estreptomicina. (A) Gel de electroforesis de la PCR para analizar la presencia del gen de resistencia a cloranfenicol en los transformantes con la mutación en el gen *rps12*. C9, colonia 9 obtenida por transformación natural con el fragmento de PCR que contiene la mutación en dicho gen; C5, colonia 5 obtenida por transformación natural con el vector pMSM2; C+, vector pMSM2 usado como control positivo de la PCR; C-, control negativo. Se usó como marcador “Generuler 100 pb DNA ladder plus” (Fermentas). **(B)** Análisis de la resistencia/sensibilidad a diferentes concentraciones de cloranfenicol de los transformantes con la mutación en el gen *rps12*. C9 y C5, representan lo mismo que en el panel A; wt, cepa silvestre de Se7942. Las concentraciones de cloranfenicol se muestran en la parte superior de la imagen encima de cada pocillo y las unidades son $\mu\text{g/ml}$.

Seleccionamos al mutante obtenido mediante transformación natural con el fragmento de PCR que contiene la mutación en el gen *rps12*, al que denominamos MSM1 (ver **Tabla 4-2**), para clonar en él las mutaciones de C11. Decidimos clonar las mutaciones #12 y #21. Descartamos la #2 porque el gen mutado codifica para la proteína PilA, que como ya se ha observado en otras cianobacterias afecta a la transformabilidad. También descartamos la #7 porque, como se ha mencionado en el apartado **5.1.2.2**, su inactivación disminuye la tasa de

crecimiento a alta intensidad lumínica (Shen et al., 2008). Por otro lado, decidimos analizar el efecto de la mutación #14, ya que, aunque no está presente en C11 y sí en C12, está relacionada con el metabolismo de ácidos grasos, otro de los temas sobre los que versa esta tesis.

La primera mutación que introdujimos fue la #12, que corresponde a la mutación en el promotor de la shikimato kinasa. Seguimos el método descrito por (Matsuoka et al., 2001) y por ello, realizamos dos construcciones, pMSM22 y pMSM24, con las que realizamos dos transformaciones naturales consecutivas.

En la primera de las transformaciones usamos el vector pMSM22, que contiene el gen y el promotor de la shikimato kinasa de C11 y el gen *agua* abajo a este, y entre ellos la construcción *rps12-kan* (ver **Tabla 4-3**). Se introdujo dicha construcción mediante transformación natural en la cepa MSM1, sustituyendo de esta manera el gen y el promotor de la shikimato kinasa silvestre por el mutado y por un gen de resistencia a kanamicina y el gen *rps12*, dando lugar a la cepa MSM2. Las colonias obtenidas se subcultivaron dos veces en medio sólido suplementado con Neo5 para lograr la segregación de los mutantes. Se analizaron mediante PCR para comprobar la correcta inserción de la construcción en el cromosoma con los cebadores 99 y 88 (**Figura 5-27**). De las 16 colonias analizadas, 12 fueron positivas en la PCR dando una banda de aproximadamente 2,2 kb, mientras que en el resto de colonias no se obtuvo ninguna banda. En el control de la PCR, la cepa silvestre (WT) se observa una banda esperable de menor tamaño, 588 pb.

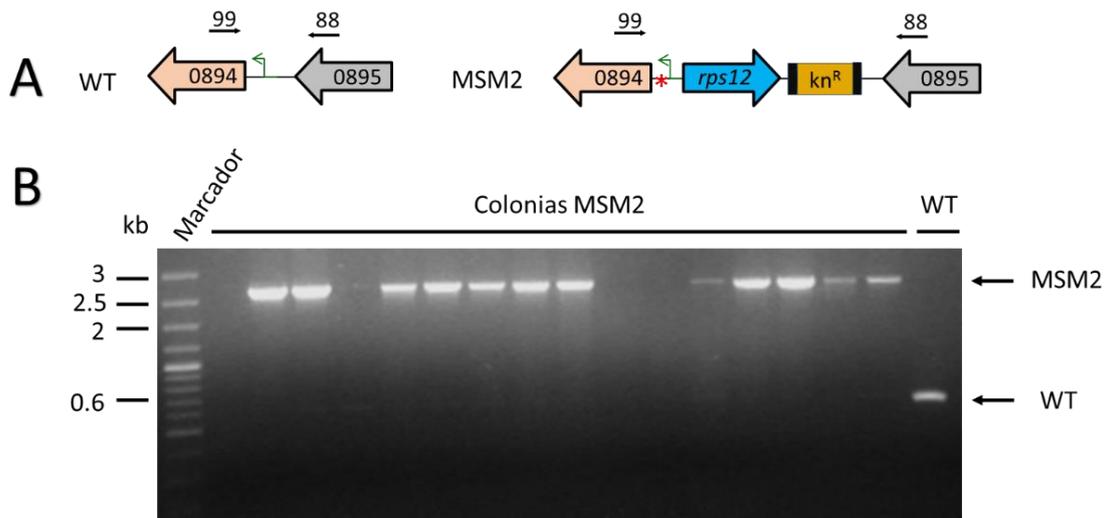


Figura 5-27. Análisis de las colonias del mutante MSM2. (A) Mapa de la región del gen de la shikimato kinasa (0894) del cromosoma de la cepa silvestre (WT) así como del mutante MSM2. Las flechas de color naranja y gris representan a los genes Synpcc7942_0894 y Synpcc7942_0895, respectivamente. La flecha azul representa al gen *rps12* de Ss6803. El rectángulo marrón representa el gen de resistencia a kanamicina (*kan*^R). La flecha verde representa al promotor del orf de la shikimato kinasa. Las flechas y los números por encima del mapa representan los oligonucleótidos usados en el análisis por PCR. La mutación en el promotor de la shikimato kinasa se representa como un asterisco de color rojo en MSM2. (B) Gel de electroforesis de los productos de la PCR utilizada para analizar la región del promotor y el gen de la shikimato kinasa de 16 colonias de MSM2 y de la cepa silvestre (WT). La flecha superior de la derecha indica la posición de la banda correspondiente a la mutación en el promotor de la shikimato kinasa de C11 y la construcción *rps12-kan*; la flecha inferior de la derecha indica la posición de la banda correspondiente a cepa silvestre. Se usó como marcador “Generuler 100 pb DNA ladder plus” (Fermentas).

Se realizó un segundo análisis en el que se examinó el cambio de la resistencia a antibióticos de las colonias obtenidas. Se tomaron colonias transformantes individuales de MSM2 y se replicaron en placas de BG11+Neo5, BG11+Sm50 y BG11, como se puede ver en la **Figura 5-28**. La cepa MSM1 sobre la que hemos introducido esta mutación era resistente a Sm. Sin embargo, al introducir la mutación en el promotor de la shikimato kinasa también hemos introducido la construcción *rps12-kan* que le confiere resistencia a Neo, pero haciendo que MSM2 pierda su resistencia a Sm. Como se ha explicado en el apartado **4.8.4**, cuando en una misma cepa se encuentra una copia silvestre del gen *rps12* y una copia mutada, la copia silvestre es dominante sobre la mutada, de ahí, que la cepa MSM2 sea sensible a Sm.

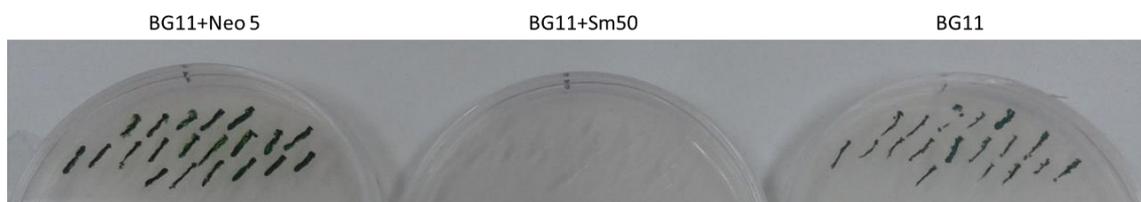


Figura 5-28. Réplicas de los transformantes MSM2 en medio BG11 con diferentes antibióticos (Neo5 y Sm50) y sin ellos.

Seleccionamos una colonia positiva para las dos técnicas de comprobación utilizadas y realizamos sobre ella la segunda transformación natural. En la cepa resultante, MSM3, los marcadores de selección (*rps12* y *kan*) son eliminados, quedando la región tal y como se encontraba en la cepa silvestre, con la excepción de la mutación en el promotor de la shikimato kinasa.

En esta segunda transformación, utilizamos el vector pMSM24 y seleccionamos los transformantes en Sm50. Las colonias obtenidas se subcultivaron dos veces en medio sólido suplementado con Sm50 para lograr la correcta segregación de los mutantes. Se analizó la inserción de la construcción en el cromosoma mediante PCR con los cebadores 99 y 88 (**Figura 5-29**). De las 16 colonias, todas fueron positivas en la PCR, dando el amplicón esperado, de aproximadamente 588 pb, al igual que el control positivo de la PCR, la cepa silvestre (WT).

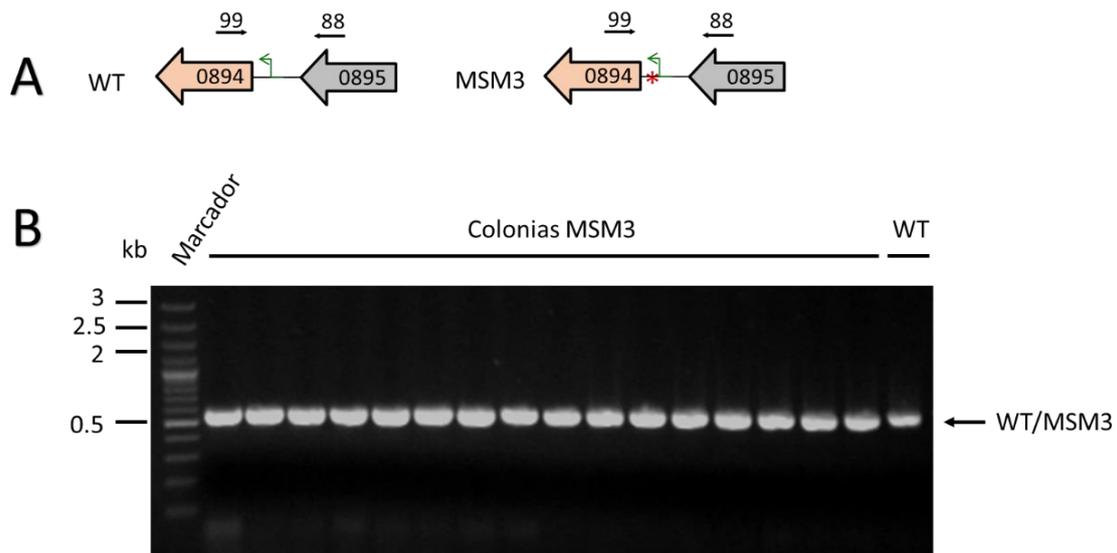


Figura 5-29. Análisis de las colonias del mutante MSM3. (A) Mapa de la región del promotor y el gen de la shikimato kinasa (0894) del cromosoma de la cepa silvestre (WT), así como del mutante MSM3. La única diferencia entre ambos es la mutación en el promotor de la shikimato kinasa representada por un asterisco de color rojo en MSM3. Las flechas de color naranja y gris representan a los genes *Synpcc7942_0894* y *Synpcc7942_0895*, respectivamente. La flecha verde representa al promotor del orf de la shikimato kinasa. Las flechas y los números por encima del mapa representan los oligonucleótidos usados en el análisis por PCR. (B) Gel de electroforesis de la PCR para analizar la región del promotor y el gen de la shikimato kinasa de 16 colonias de MSM3 y de la cepa silvestre (WT). La flecha de la derecha indica la posición de la banda correspondiente a la mutación en el promotor de la shikimato kinasa de C11, así como a la cepa silvestre. Se usó como marcador "Generuler 100 pb DNA ladder plus" (Fermentas).

Como en el caso anterior, se hizo un segundo análisis en donde se examinó el cambio de la resistencia a antibióticos de las colonias obtenidas. Se tomaron colonias transformantes individuales de MSM3 y se replicaron en placas de BG11+Sm50, BG11+Neo5 y BG11, como se puede ver en la **Figura 5-30**. La cepa MSM2 sobre la que hemos realizado esta segunda transformación era resistente a Neo, sin embargo al eliminar la construcción *rps12-kan*, la cepa pierde su resistencia a Neo y recupera su resistencia a Sm (apartado **4.8.4**). Por tanto, serán

5. Resultados

positivas aquellas colonias que sean resistentes a Sm, sensibles a Neo y crezcan en BG11. De las colonias analizadas solo una de ellas presentaba resistencia a Neo. Seleccionamos una de las colonias Neo^s para el siguiente paso.

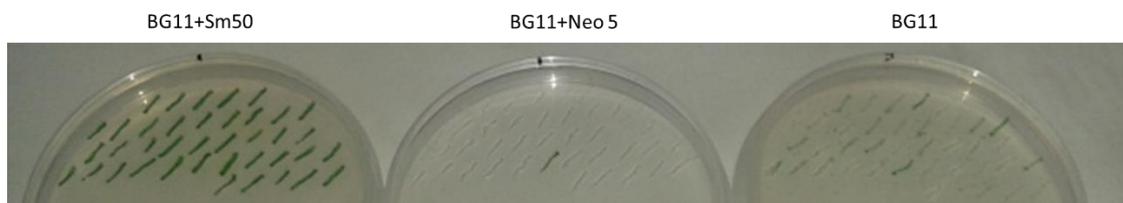


Figura 5-30. Réplicas de los transformantes MSM3 en medio BG11 con diferentes antibióticos (Sm50 y Neo5) y sin ellos.

Se amplificó y secuenció la región de la mutación #12 en las colonias MSM3 con los cebadores 87 y 88. Como se observa en la **Figura 5-31**, las cuatro colonias analizadas presentaban la mutación.

| | 569 | 570 | 580 | 590 | 600 | 610 | 620 | 630 |
|--------------------------|-----|----------------------------------|-----|---------------------------------------|-----|-----|-----|-----|
| romotor+shikimato kinasa | 569 | GACACGCGATCGCTTCTCCAATTCCTTGGACA | C | CATACCGCTCGGCTGGGGTCTCCTCCAAGCTGCGCCC | | | | |
| Colonia 2 | 444 | GACACGCGATCGCTTCTCCAATTCCTTGGACA | C | CATACCGCTCGGCTGGGGTCTCCTCCAAGCTGCGCCC | | | | |
| Colonia 9 | 432 | GACACGCGATCGCTTCTCCAATTCCTTGGACA | C | CATACCGCTCGGCTGGGGTCTCCTCCAAGCTGCGCCC | | | | |
| Colonia 11 | 431 | GACACGCGATCGCTTCTCCAATTCCTTGGACA | C | CATACCGCTCGGCTGGGGTCTCCTCCAAGCTGCGCCC | | | | |
| Colonia 16 | 441 | GACACGCGATCGCTTCTCCAATTCCTTGGACA | C | CATACCGCTCGGCTGGGGTCTCCTCCAAGCTGCGCCC | | | | |

Figura 5-31. Alineamiento de la región que contiene la mutación #12 en colonias de MSM3. Análisis de la secuencia de esta región de cuatro colonias de la cepa MSM3 alineadas frente a la cepa silvestre Se7942 cuya referencia en GenBank es NC_007604. En rojo, se indica la posición que ocupa o debería ocupar la mutación en esa región. Alineamientos realizados con "AlignX" una aplicación del programa "Vector NTI".

Para la construcción del resto de mutantes, realizamos los mismos pasos que hemos utilizado para la construcción del mutante MSM2 y MSM3. En la **Figura 5-32**, se muestra un esquema de las cepas construidas así como de los vectores utilizados para construir cada una de ellas. Se partió de la cepa MSM1 y se obtuvieron tres mutantes simples, cada uno de los cuales contenían cada una de las tres mutaciones mencionadas anteriormente: #12, #14 y #21. Estos mutantes son MSM3, MSM7 y MSM5, respectivamente. Sobre el mutante MSM5, se construyeron dos dobles mutantes, MSM9 y MSM11. El primero de ellos contenía las mutaciones #12 y #21, mientras que el segundo presentaba las mutaciones #14 y #21. A partir del mutante MSM7, se construyó un doble mutante al que denominamos MSM13 y que contenía las mutaciones #12 y #14. Por último, llevamos a cabo la construcción de un triple mutante, MSM15, que contenía las tres mutaciones (#12, #14 y #21).

de crecimiento utilizando lámparas fluorescentes (3,13 h). Por tanto, seleccionamos la intensidad de luz de $280 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ para el crecimiento de los mutantes construidos y el cálculo del tiempo de generación.

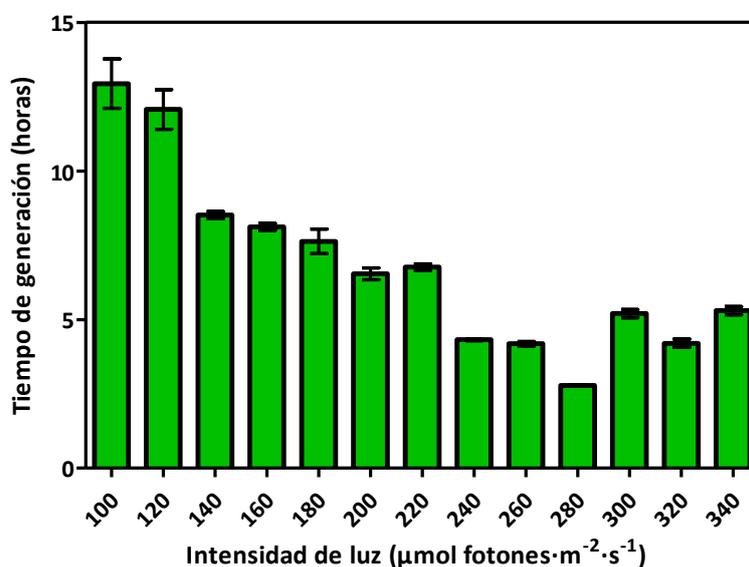


Figura 5-33. Tiempo de generación de C11-G₁₂₄₈ a diferentes intensidades de luz. Los datos corresponden a la media de tres réplicas biológicas independientes y se representan como la media \pm DE.

Realizamos el análisis del tiempo de generación de los mutantes de la **Figura 5-32**, comenzando con el mutante MSM1, que contiene la mutación en el gen *rps12* que confiere resistencia a estreptomicina. Este análisis se realizó para determinar si esta mutación tenía algún efecto sobre el crecimiento de Se7942 y que pudiera enmascarar el producido por las mutaciones de las poblaciones evolucionadas.

No se observaron diferencias significativas en el tiempo de generación entre la cepa silvestre Se7942 y el mutante MSM1 (**Figura 5-34, A**). Por tanto, podemos concluir que la mutación en el gen *rps12* no afecta al tiempo de generación.

A continuación, analizamos el tiempo de generación del triple mutante MSM15 y C11 frente a la cepa silvestre de Se794 (**Figura 5-34, B**). Los tiempos de generación para C11 y MSM15 fueron similares (4,29 h y 4,25 h, respectivamente), pero más bajos que la cepa silvestre. Por tanto, las tres mutaciones insertadas en la cepa MSM15 son suficientes para conseguir el fenotipo de crecimiento rápido de la cepa C11.

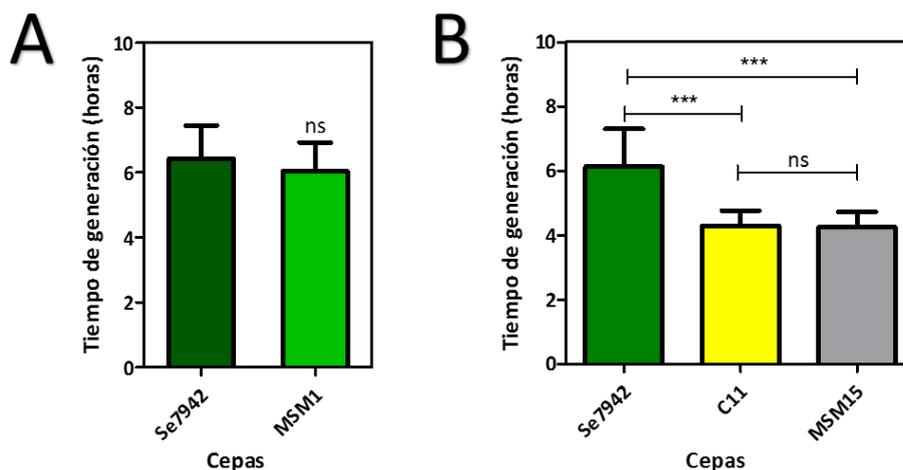


Figura 5-34. Tiempo de generación de los mutantes MSM1 y MSM15. (A) Comparativa del tiempo de generación de la cepa silvestre (Se7942) frente al mutante MSM1. La barra verde oscura representa a la cepa silvestre y la barra verde clara a la cepa MSM1. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. ns (no significativo) $p > 0,05$, test t de Student no pareado. **(B)** Comparativa del tiempo de generación de la cepa silvestre (Se7942) frente a C11 y el triple mutante MSM15, construido con las mutaciones de las cepas evolucionadas. La barra verde oscura representa a la cepa silvestre, la barra amarilla representa a la cepa C11 y la barra gris representa a la cepa MSM15. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. ns (no significativo) $p > 0,05$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa Se7942.

Para identificar que mutaciones presentes en MSM15 eran responsables del crecimiento rápido, realizamos curvas de crecimiento con los mutantes que llevaban cada una de las mutaciones por separado, o sea, MSM3, MSM5 y MSM7 (**Figura 5-35, A**). Tanto MSM3 como MSM5 presentaban diferencias significativas frente a la cepa silvestre, siendo menor esta diferencia en la cepa MSM5. Por otro lado, la cepa MSM7 no presentaba diferencias significativas frente a la cepa silvestre, por lo que la mutación clonada en esta cepa no tiene un efecto significativo sobre la tasa de crecimiento. Realizamos una segunda comparativa de los mutantes individuales (MSM3, MSM5 y MSM7) y del mutante triple (MSM15) frente a la cepa C11 (**Figura 5-35, B**). Los mutantes individuales presentaron diferencias significativas frente a la cepa C11. Solo el mutante triple MSM15 no presentó diferencias significativas, hecho que ya se había discutido en el párrafo anterior.

Si analizamos en conjunto estas dos gráficas, podemos concluir que el efecto de las mutaciones clonadas en los mutantes MSM3 y MSM5, presentes en el mutante triple MSM15, son las responsables del fenotipo de crecimiento rápido observado en la C11-G₁₂₄₈.

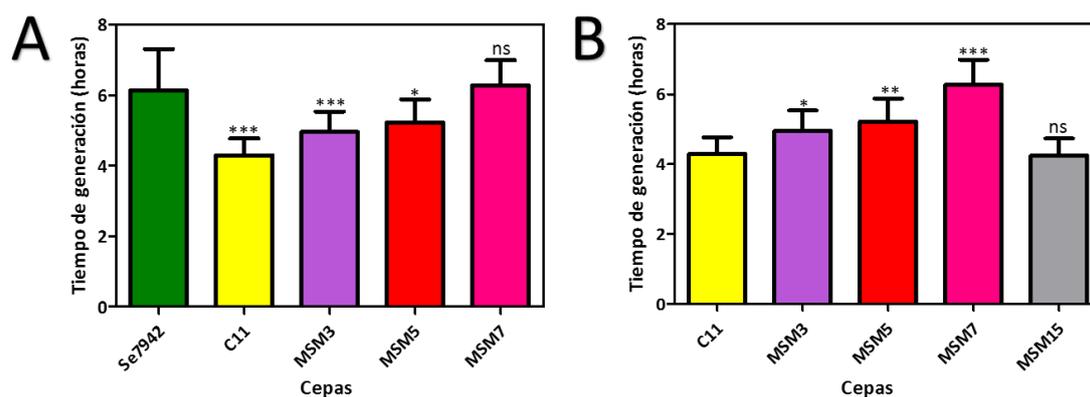


Figura 5-35. Tiempo de generación de los mutantes construidos con las mutaciones de las cepas evolucionadas utilizando luces LED a $280 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. (A) Comparativa del tiempo de generación de la cepa silvestre (Se7942) frente a C11 y los mutantes sencillos MSM3, MSM5 y MSM7. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. ns (no significativo) $p > 0,05$, * $p < 0,05$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa silvestre de Se7942. (B) Comparativa del tiempo de generación de la cepa C11 frente a los mutantes sencillos MSM3, MSM5 y MSM7 y al mutante triple MSM15. Los datos corresponden al menos tres réplicas biológicas independientes y se representan como la media \pm DE. ns (no significativo) $p > 0,05$, * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa C11.

Estos mismos análisis fueron repetidos en el Instituto Pasteur de París en las mismas condiciones en las que se realizó el experimento de evolución, es decir, usando luces fluorescentes con una intensidad de luz de $1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. Se analizó el tiempo de generación de dos mutantes simples MSM3 y MSM7, un mutante doble MSM9 y un mutante triple MSM15 (**Figura 5-36**). Si comparamos el tiempo de generación de estos mutantes frente a la cepa silvestre, los mutantes MSM3, MSM9 y MSM15, al igual que C11, presentaron diferencias significativas respecto a la cepa silvestre (**Figura 5-36, A**). Por otro lado, al comparar el tiempo de generación de todas estas cepas mutantes frente a C11, se observaron diferencias significativas únicamente en la cepa MSM7 (**Figura 5-36, B**). En estas condiciones, con la única mutación presente en el mutante MSM3 se consigue el fenotipo de crecimiento rápido. De ahí que las cepas MSM9 y MSM15, que contienen ambas esta mutación, presenten también una tasa de crecimiento sin diferencias significativas frente a la de la cepa C11.

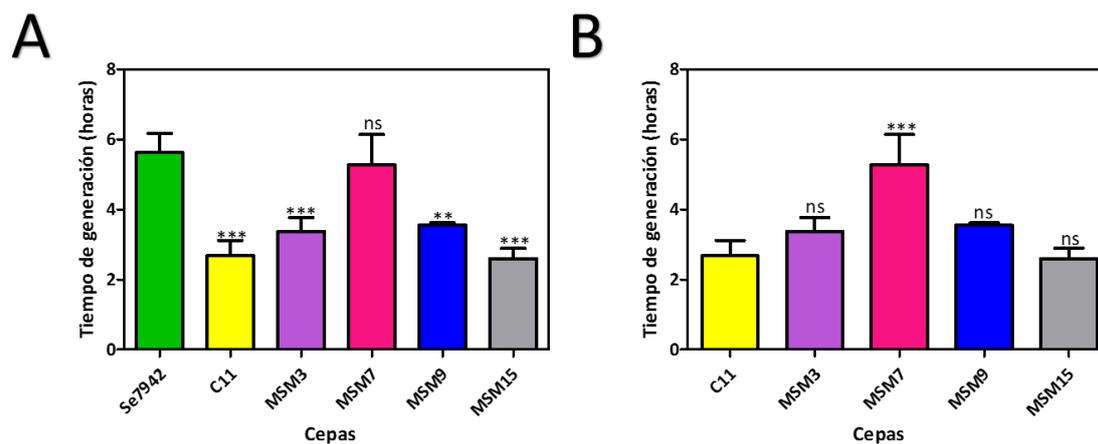


Figura 5-36. Tiempo de generación de los mutantes construidos con las mutaciones de las cepas evolucionadas utilizando lámparas fluorescentes a $1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. (A) Comparativa del tiempo de generación de la cepa silvestre (Se7942) frente a C11 y los mutantes sencillos MSM3 y MSM7, al mutante doble MSM9 y al mutante triple MSM15. Los datos corresponden a la media de al menos tres réplicas biológicas independientes y se representan como la media + DE. ns (no significativo) $p > 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa silvestre de Se7942. (B) Comparativa del tiempo de generación de la cepa C11 frente a los mutantes sencillos MSM3 y MSM7, al mutante doble MSM9 y al mutante triple MSM15. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. ns (no significativo) $p > 0,05$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa C11.

Adicionalmente, se repitieron estas mismas curvas de crecimiento en el “Multi-cultivator MC 1000-OD” que, como se ha comentado anteriormente lleva LEDs instalados, a $1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. También se realizaron en el Instituto Pasteur de París porque nuestro equipo no podía superar intensidades de luz superiores a $1.000 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. Como el caso anterior, primero calculamos el tiempo de generación de todas las cepas frente a la cepa silvestre (Figura 5-37, A), siendo los resultados muy similares a los obtenidos usando luces fluorescentes (Figura 5-36, A). La única diferencia radicaba en que los grados de significación fueron menores en el caso de las cepas MSM3 y MSM15. En la Figura 5-37, B se compara el tiempo de generación de estas mismas cepas frente a C11, presentando solamente diferencias significativas la cepa MSM7. Este mismo resultado se había observado cuando se utilizaban luces fluorescentes (Figura 5-36, B). Sin embargo, si comparamos los resultados obtenidos en la Figura 5-35 y en la Figura 5-37, usando en ambos casos LEDs como fuente de alimentación, observamos que la velocidad de crecimiento es mayor en la primera imagen. Esto se debe a la diferencia en la intensidad de luz utilizada en el experimento. Mientras que para calcular el tiempo de generación en la Figura 5-35, se utilizaron $280 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, para la Figura 5-37 se utilizaron $1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. Por tanto, la diferencia que estamos encontrando se debe a la intensidad lumínica y no al tipo de luz utilizado.

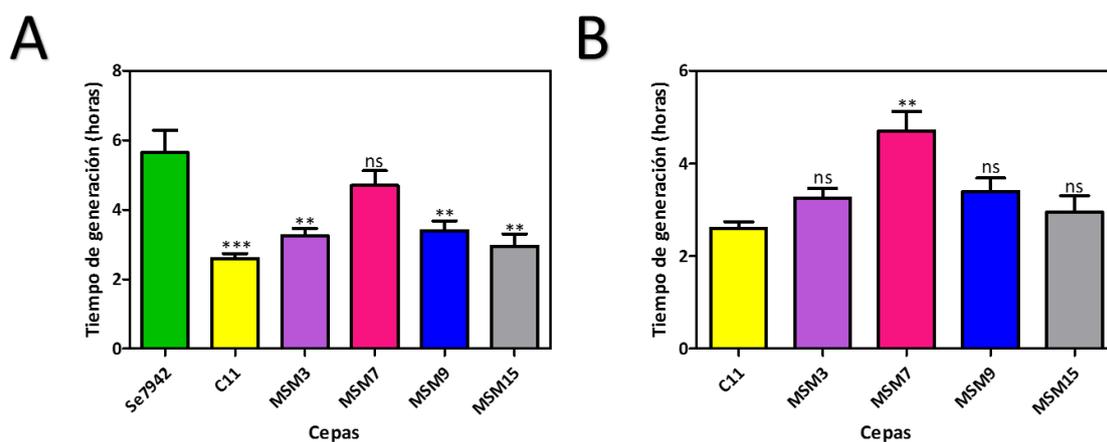


Figura 5-37. Tiempo de generación de los mutantes construidos con las mutaciones de las cepas evolucionadas utilizando luces LED a $1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. (A) Comparativa del tiempo de generación de la cepa silvestre (Se7942) frente a C11 y los mutantes sencillos MSM3 y MSM7, al mutante doble MSM9 y al mutante triple MSM15. Los datos corresponden a la media de al menos tres réplicas biológicas independientes y se representan como la media + DE. ns (no significativo) $p > 0,05$, ** $p < 0,01$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa silvestre de Se7942. (B) Comparativa del tiempo de generación de la cepa C11 frente a los mutantes sencillos MSM3 y MSM7, al mutante doble MSM9 y al mutante triple MSM15. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. ns (no significativo) $p > 0,05$, ** $p < 0,01$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa C11.

5.1.6. Efecto de las mutaciones en la expresión génica

Con el objetivo de identificar genes con expresión diferencial significativa, se llevó a cabo un análisis transcriptómico de C11-G₁₂₄₈ y de C12-G₁₁₆₈ bajo las mismas condiciones en las que se había llevado a cabo el experimento de evolución. Como control de estos experimentos se utilizó la cepa silvestre de Se7942.

En primer lugar, se llevó a cabo un análisis global de los genes que presentaban una expresión diferencial, tanto en C11-G₁₂₄₈ como en C12-G₁₁₆₈. Al no tener réplicas biológicas de cada una de las muestras, se dividió la muestra en tres partes para calcular la expresión diferencial (apartado 4.10.4).

Se consideraron genes con expresión diferencial aquellos cuyo $\log_2\text{FoldChange}$ fuera superior a 1,5 (inducción) o inferior a -1,5 (represión) y un p-valor ajustado inferior a 0,05 (95 % de confianza estadística). De esta manera, se obtuvo que en C11-G₁₂₄₈ aumentaban su expresión el 0,14 % de sus genes, mientras que el 0,92 % la disminuían. En C12-G₁₁₆₈ ninguno de sus genes aumentó su expresión y un 0,22 % disminuyó su expresión. En ambos casos el porcentaje de genes afectados es bajo, por lo que las mutaciones que han surgido durante el proceso de evolución no afectan a la expresión de forma generalizada. En la **Tabla 5-3** y en la **Tabla 5-4**, se muestran los genes afectados en cada caso.

Tabla 5-3. Genes diferencialmente expresados en C11-G₁₂₄₈. Se comparó dicha expresión con la de la cepa silvestre en las mismas condiciones en las que se llevó a cabo el experimento de evolución (41 °C, 1.313 μmol de fotones- $\text{m}^{-2}\cdot\text{s}^{-1}$ con luces fluorescentes). En esta tabla se representan aquellos genes cuyo p-valor ajustado $\leq 0,05$ y el $\log_2\text{FC} \leq -1,5$ o $\geq 1,5$.

| Gen | Nombre del gen | Categoría | p-valor ajustado | $\log_2\text{FC}^*$ |
|-----------------|---|---|------------------|---------------------|
| Synpcc7942_0034 | N-acetilornitina aminotransferasa | Función desconocida | 0,04777705 | -3,290687954 |
| Synpcc7942_0245 | Gliceraldehído-3-fosfato deshidrogenasa (EC 1.2.1.12) | Transporte y metabolismo de carbohidratos | 0,01448052 | -4,5169876 |
| Synpcc7942_0304 | Proteína hipotética | Transporte y metabolismo de coenzimas | 0,0075568 | -5,053165341 |
| Synpcc7942_0337 | ATP sintasa cadena gamma | Producción y conversión de energía | 0,03041854 | 3,039049714 |
| Synpcc7942_0369 | Posible monoamina oxidasa contenedora de flavina | Transporte y metabolismo de aminoácidos | 7,91E-05 | -5,099640307 |
| Synpcc7942_0556 | Regulador transcripcional de dos componentes, familia hélices alado | Mecanismos de transducción de señales | 0,02304049 | -4,656049036 |
| Synpcc7942_0557 | Proteína hipotética | Función desconocida | 0,00059481 | -6,615484157 |
| Synpcc7942_0834 | Proteína hipotética | Función desconocida | 2,23E-05 | -8,716424227 |
| Synpcc7942_1164 | Proteína hipotética | Función desconocida | 0,01590783 | -3,39323936 |
| Synpcc7942_1238 | Permeasa transporte de nitrato nrtB | Transporte y metabolismo de iones inorgánicos | 0,02716149 | 4,07648493 |
| Synpcc7942_1572 | Proteína similar a subunidad deshidrogenasa | Transporte y metabolismo de aminoácidos | 0,02304049 | -4,40532631 |
| Synpcc7942_1574 | Posible glucosidasa | Función desconocida | 0,01590783 | -3,528056382 |
| Synpcc7942_1610 | Transhidrogenasa piridina nucleótido subunidad beta (EC 1.6.1.2) | Función desconocida | 0,03041854 | -3,28293019 |
| Synpcc7942_1611 | Transhidrogenasa nicotinamida nucleótido similar a subunidad alfa | Producción y conversión de energía | 0,03041854 | -4,737553053 |
| Synpcc7942_1612 | Piridina nucleótido transhidrogenasa subunidad alfa | Producción y conversión de energía | 0,00141174 | -5,597359212 |
| Synpcc7942_1661 | Proteína hipotética | Función desconocida | 0,0377896 | -2,972906155 |
| Synpcc7942_1849 | ARN polimerasa sigma factor SigC | Transcripción | 0,0141174 | -6,711244882 |
| Synpcc7942_2182 | Proteína hipotética | Función desconocida | 0,03541122 | -2,988612715 |
| Synpcc7942_2185 | Proteína hipotética | Función desconocida | 0,02304049 | -3,399964955 |

5. Resultados

| Gen | Nombre del gen | Categoría | p-valor ajustado | log ₂ FC* |
|------------------|--|-------------------------------------|------------------|----------------------|
| Synpcc7942_2219 | Proteína ribosomal 30S S8 | Función desconocida | 0,03597078 | 3,302708757 |
| Synpcc7942_2223 | Proteína ribosomal 30S S17 | Traducción | 0,05004095 | 4,031367977 |
| Synpcc7942_2254 | Proteína hipotética | Función desconocida | 0,03041854 | -3,999723651 |
| Synpcc7942_2267 | Proteína hipotética | Función desconocida | 0,00290756 | -4,338742971 |
| Synpcc7942_2386 | Proteína expresión/ formación hidrogenasa HypD | Función desconocida | 0,00450211 | -3,772032901 |
| Synpcc7942_2391 | Similar a oxidoreductasa | Transporte y metabolismo de lípidos | 0,03041854 | -3,53832369 |
| Synpcc7942_2555 | Hidrogenasa NAD-reductora HoxS subunidad beta | Producción y conversión de energía | 0,02460038 | -3,552347207 |
| Synpcc7942_2556 | Hidrogenasa NAD-reductora HoxS subunidad delta | Producción y conversión de energía | 0,0001048 | -5,258559645 |
| Synpcc7942_2557 | Hidrogenasa NAD-reductora HoxS subunidad gamma | Producción y conversión de energía | 0,00290756 | -5,024567694 |
| Synpcc7942_B2632 | Proteína hipotética | Función desconocida | 0,03041854 | -4,488803315 |

*log₂FC, log₂FoldChange. En granate, se representan aquellos genes cuyo log₂FC ≥ 1,5; en amarillo, se representan aquellos genes cuyo log₂FC ≤ -1,5.

Tabla 5-4. Genes diferencialmente expresados en C12-G₁₁₆₈. Se comparó dicha expresión con la de la cepa silvestre en las mismas condiciones en las que se llevó a cabo el experimento de evolución (41 °C, 1.313 μmol de fotones·m⁻²·s⁻¹ con luces fluorescentes). En esta tabla se representan aquellos genes cuyo p-valor ajustado ≤ 0,05 y el log₂FC ≤ -1,5 o ≥ 1,5.

| Gen | Nombre del gen | Categoría* | p-valor ajustado | log ₂ FC* |
|------------------|---------------------|---------------------|------------------|----------------------|
| Synpcc7942_0891 | Proteína hipotética | Función desconocida | 0,00866235 | -4,022487422 |
| Synpcc7942_1707 | Proteína hipotética | Función desconocida | 0,00762227 | -3,957711208 |
| Synpcc7942_2277 | Proteína hipotética | Función desconocida | 0,05714964 | -3,353885543 |
| Synpcc7942_R0027 | Cys-ARNt | ARN | 0,01508504 | -4,37698258 |
| Synpcc7942_R0046 | Val-ARNt | ARN | 0,00051562 | -5,117837151 |

*log₂FC, log₂FoldChange. En amarillo, se representan aquellos genes cuyo log₂FC ≤ -1,5.

Si analizamos en más detalle la actividad que realizan estos genes dentro de la célula, tal y como se muestra en la **Figura 5-38** y **Figura 5-39**, entenderemos mejor a qué función o funciones celulares está afectando dicho cambio en la expresión.

En primer lugar, analizaremos los genes sobreexpresados en C11-G₁₂₄₈ (**Figura 5-38, A**). Hay que destacar el alto porcentaje de genes que codifican para proteínas con función desconocida que se estaban sobreexpresando, un 25 %. Si bien es cierto, que este fenómeno

está asociado con el hecho de que el genoma de Se7942 está compuesto por un porcentaje muy elevado de genes codificantes de proteínas hipotéticas, alrededor del 51 % (**Figura 5-38, C**). Por otro lado, un porcentaje similar de genes que codifican para proteínas de producción y conversión de energía estaban sobreexpresados. Junto a estas dos categorías de genes, existían otras dos más. Un 25 % de genes que codifican para proteínas involucradas en traducción y otro 25 % de genes que codifican para proteínas implicadas en el transporte y metabolismo de iones orgánicos.

Si atendemos ahora a los genes cuya expresión se vio disminuida en C11-G₁₂₄₈, tal y como se muestra en la **Figura 5-38, B**, observamos que como ocurría en el caso anterior un alto porcentaje no tiene función asignada. La suma total de estos genes asciende a un 54 %, mucho más alta que en el caso anterior. Por otro lado, a pesar de que un alto porcentaje de genes codificantes para proteínas de producción y conversión de energía estaban sobreexpresados, un porcentaje del 21 % de estos genes había disminuido su expresión. También hay que destacar que la expresión de un alto porcentaje de genes que codifican implicados en el metabolismo disminuyó, 9 % en el metabolismo de aminoácidos, 4 % en el de carbohidratos y 3 % en el de coenzimas. Además, un porcentaje muy bajo de genes involucrados en el mantenimiento y procesamiento de información también disminuyó su expresión, siendo tan solo un 4 %, y cuya función estaba relacionada con la transcripción. Finalmente, el porcentaje restante de genes, un 4%, estaba implicado en procesos celulares y de señalización, concretamente en mecanismos de transducción de señales.

5. Resultados

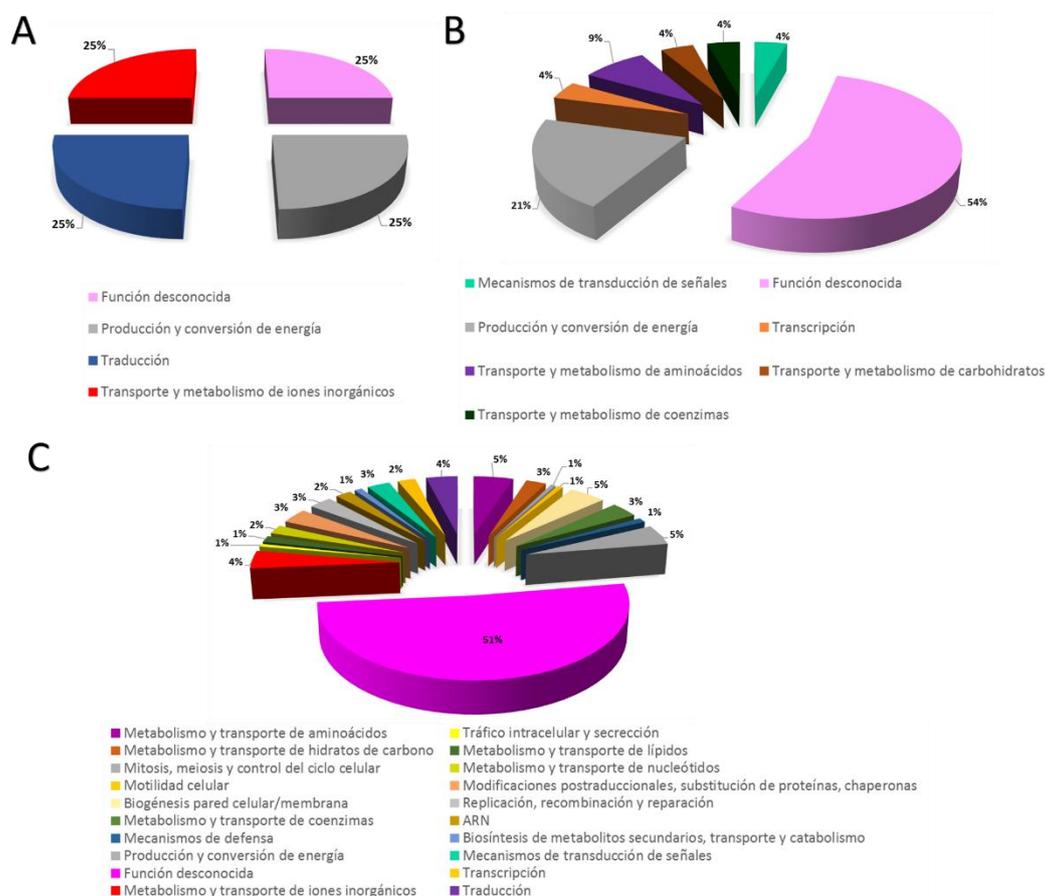


Figura 5-38. Distribución de los genes diferencialmente expresados en C11-G₁₂₄₈ en categorías COG. (A) Genes en los que ha aumentado la expresión. **(B)** Genes en los que ha disminuido la expresión. Los COG a los que pertenecen se muestran en las leyendas bajo los paneles. Más información sobre estos genes en la **Tabla 5-3**. **(C)** Distribución de categorías COG en Se7942.

Si analizamos ahora los genes cuya expresión se vio disminuida en C12-G₁₁₆₈, tal y como se muestra en la **Figura 5-39, A**, ya que en esta cepa no hubo genes que se estuvieran sobreexpresando. Como ocurría en los casos anteriores, un alto porcentaje no tiene función asignada, como consecuencia directa del alto porcentaje de genes del genoma de Se7942 que codifican para proteínas hipotéticas (**Figura 5-39, B**). La suma total de estos genes asciende a un 60 %, el porcentaje más alto de todos los analizados. También hay que destacar que la expresión de un alto porcentaje de genes implicados en el mantenimiento y procesamiento de información también disminuyó, concretamente un 40 % de ellos implicados en el ARN de transferencia.

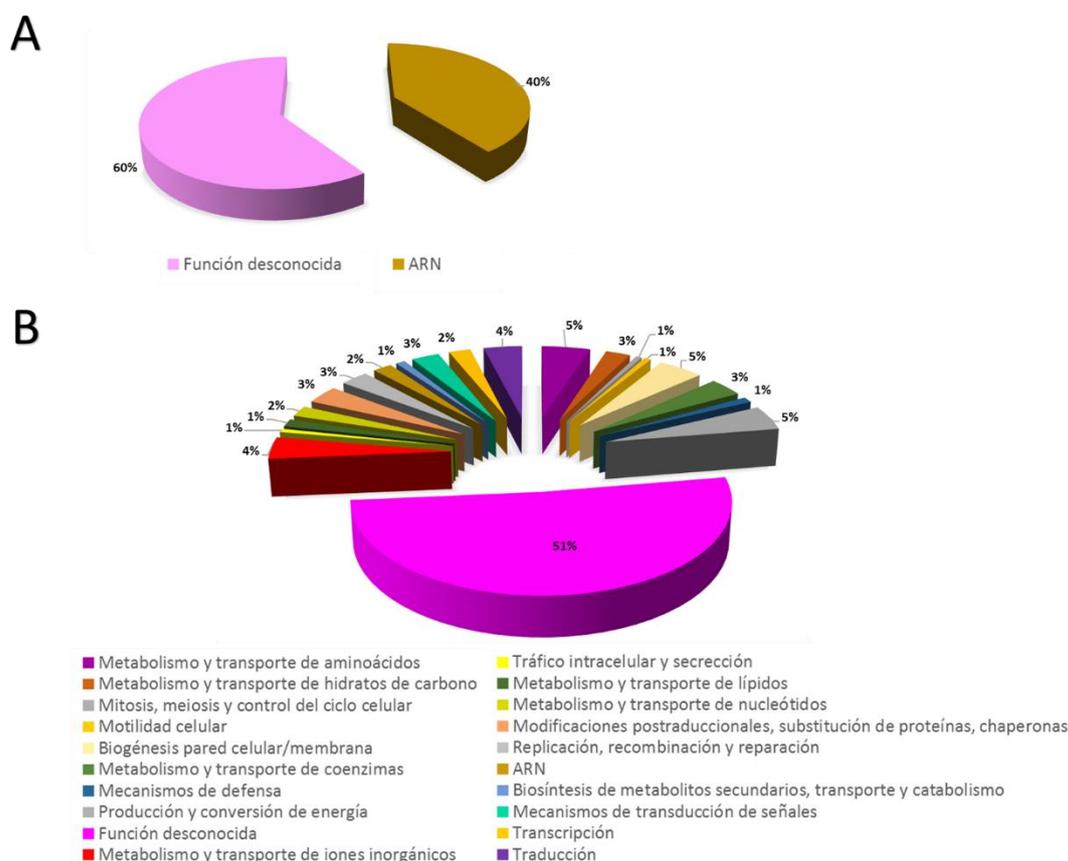


Figura 5-39. Distribución de los genes diferencialmente expresados en C12-G₁₁₆₈ en categorías COG. (A) Genes en los que ha disminuido la expresión. Más información sobre estos genes en la **Tabla 5-4**. **(B)** Distribución de categorías COG en Se7942.

En segundo lugar, analizamos la expresión de los genes que conforman el entorno genético de cada una de las cinco mutaciones presentes en C11-G₁₂₄₈.

Empezaremos analizando el entorno genético de la mutación #2, que afecta al gen codificante de la proteína PilA, tal y como se muestra en la **Figura 5-40**. Lo componen genes codificantes tanto de proteínas hipotéticas como de proteínas con algún dominio funcional conocido (**Figura 5-40, A**). Cabe destacar que existe otra copia del gen *pilA* en esta región (Synpcc7942_0049), hecho que ya se había destacado en el apartado **5.1.2.1**. Por otro lado, si atendemos a los niveles de expresión obtenidos, expresados como niveles de RPKM (del inglés, “reads per kilobase per million”), podemos afirmar que la mutación no afecta a la expresión de ninguna de las copias de los genes *pilA* (Synpcc7942_0048 y Synpcc7942_0049) (**Figura 5-40, B**). Sin embargo, se observa que la mutación disminuye la expresión del gen Synpcc7942_0047. Sin embargo, en la **Tabla 5-3**, no aparece este gen, ya que no cumple los requisitos que el \log_2 FoldChange sea inferior a -1,5 y el p-valor ajustado inferior a 0,05, siendo estos valores -0,579 y 1, respectivamente. Por tanto, concluimos que la mutación #2, que afecta al gen *pilA*, no modifica la expresión del propio gen ni de los genes de su entorno genético.

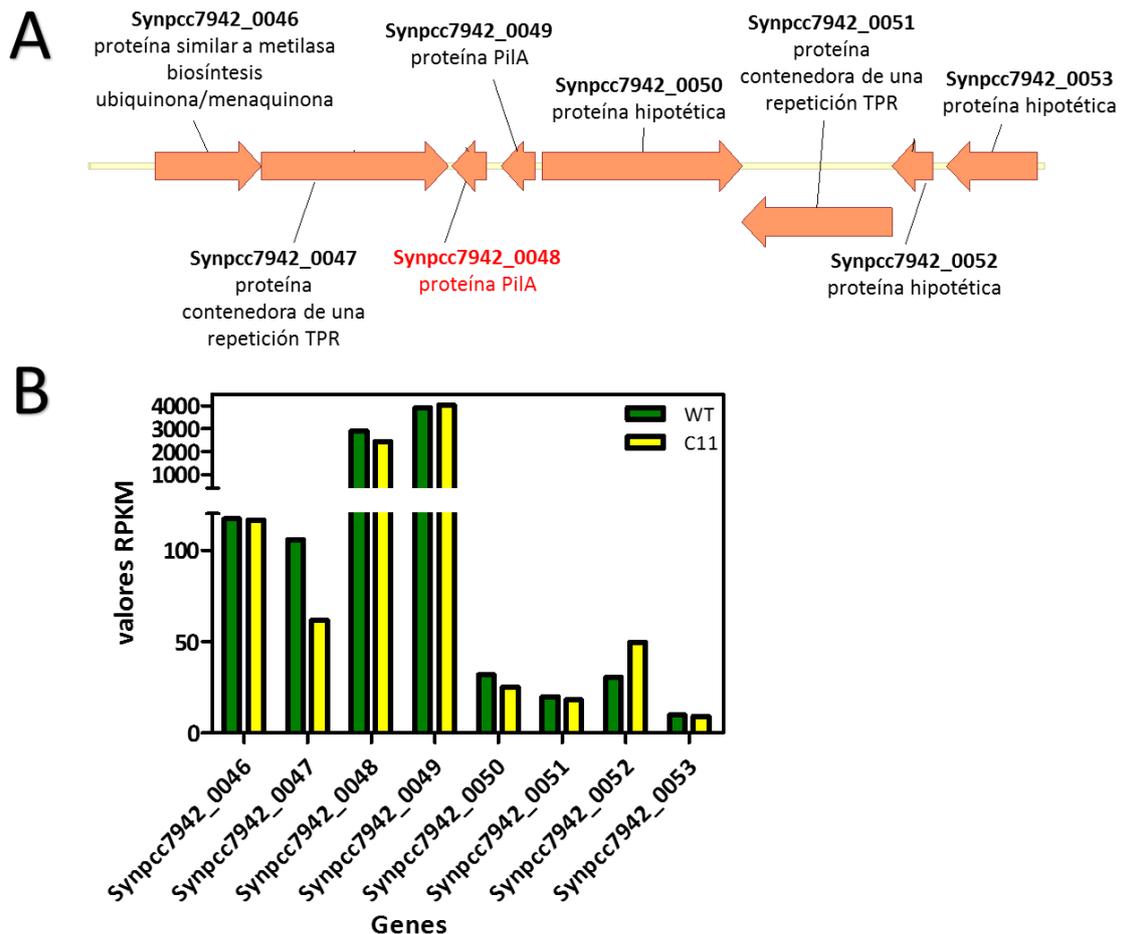


Figura 5-40. Expresión de los genes del entorno genético de la mutación #2. (A) Mapa genético alrededor de *pilA* (En rojo, se especifica el gen en el que se encuentra la mutación descrita). (B) Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético del gen *pilA* en la cepa silvestre (barras verdes) frente a la cepa C11 (barras amarillas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

A continuación, analizaremos el entorno genético de la mutación #7, que afecta a una proteína hipotética conservada, cuya localización en el genoma de Se7942 se representa en la **Figura 5-41**. El entorno genético de dicha mutación está compuesta por genes que codifican para una proteína de división celular, una apoproteína, varias proteínas relacionadas con la síntesis de alofococianina, una proteína de unión a ficobilisomas y por cinco subunidades de la ATP sintasa FoF₁ (**Figura 5-41, A**). Los niveles de RPKM se muestran en la **Figura 5-41, B**. Observamos que los genes Synpcc7942_0325, Synpcc7942_0327 y Synpcc7942_0330 han aumentado su expresión en C11 en comparación con Se7942, aunque en ninguno de ellos el cambio es significativo, ya que el valor del p-valor ajustado es superior a 0,05.

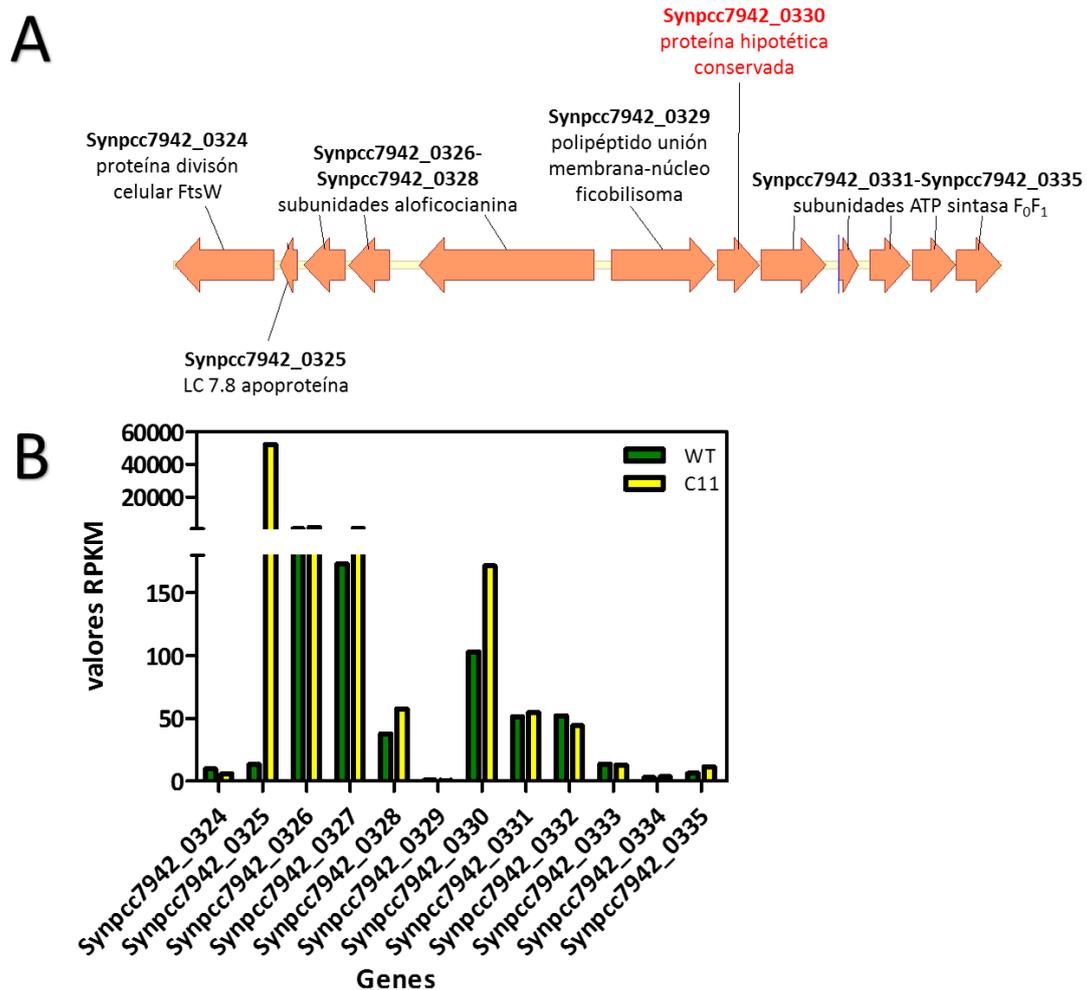


Figura 5-41. Expresión de los genes del entorno genético de la mutación #7. (A) Mapa genético del entorno de Synpcc7942_0330 (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético de Synpcc7942_0330 en la cepa silvestre (barras verdes) frente a la cepa C11 (barras amarillas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

En tercer lugar, examinaremos el entorno genético de la mutación #12 que afecta al promotor del gen de la shikimato kinasa (Synpcc7942_0894, *aroK*). El entorno genético de este gen está compuesto por genes que codifican para proteínas con funciones diversas (**Figura 5-42, A**). Destacan genes que codifican para proteínas implicadas en la división celular (*minD* y *minE*), una sintasa F_0F_1 , una proteína de fotosistema, una proteína hipotética y una proteína similar a ferredoxina. Si atendemos a los niveles de RPKM, los genes Synpcc7942_0893 y Synpcc7942_0894 aumentan su expresión, mientras que los genes Synpcc7942_0892 y Synpcc7942_0898 la disminuyen (**Figura 5-42, B**). En ninguno de ellos el cambio en la expresión génica fue significativo, ya que los valores del \log_2 FoldChange y p-valor ajustado fue superior a 0,05 (1, 1, 0,253 y 0,732, para Synpcc7942_0892, Synpcc7942_0893, Synpcc7942_0894 y Synpcc7942_0898, respectivamente). Solo encontramos diferencias significativas en la expresión de Synpcc7942_0894 cuando comparamos los niveles de expresión usando el p-valor,

5. Resultados

cuyo valor es de 0,0061. Sin embargo, esta medida es menos restrictiva que el p-valor ajustado, por lo que para hacer una comparación más acertada sería necesario repetir el RNAseq con tres réplicas e incluso, complementar dicho análisis con RT-PCR.

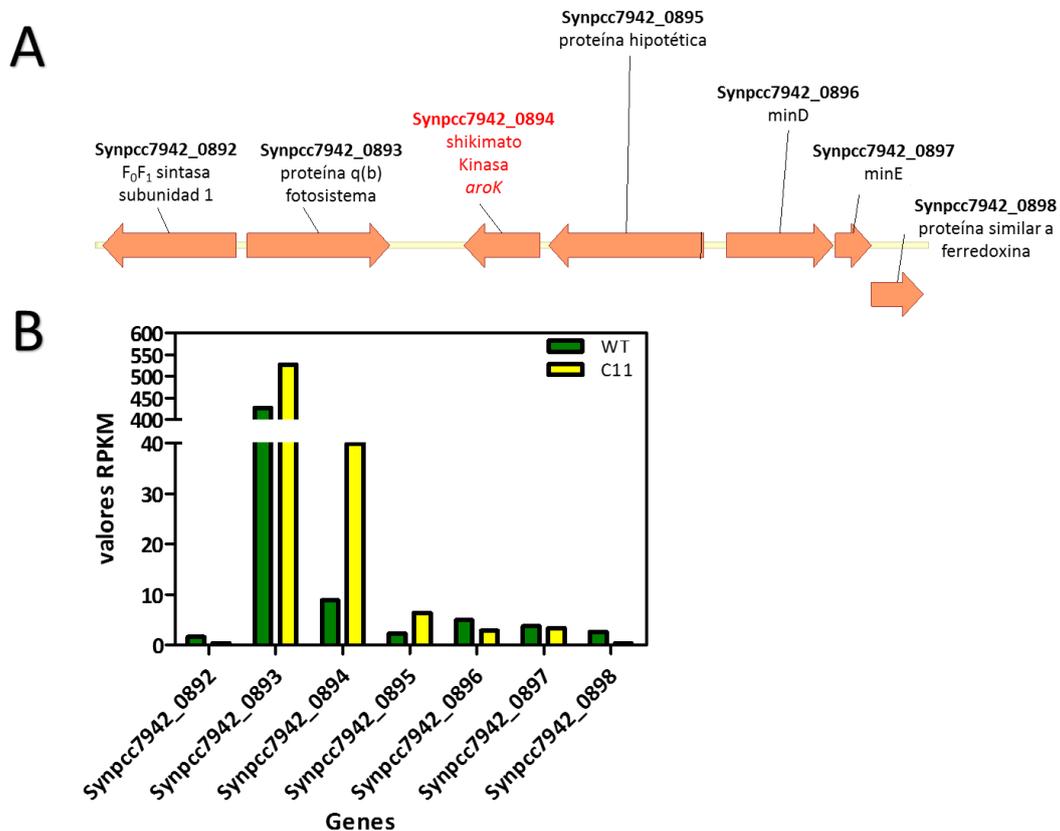


Figura 5-42. Expresión de los genes del entorno genético de la mutación #12. (A) Mapa genético del entorno de *aroK* (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético del gen *aroK* en la cepa silvestre (barras verdes) frente a la cepa C11 (barras amarillas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

Por último, el entorno genético de la mutación #21, que afecta al gen que codifica para la histidina kinasa *SasA*, contiene proteínas hipotéticas conservadas, así como genes que codifican para una metionil ARNt-formiltransferasa, un represor de xilosa, una ribosa-fosfato piroquinasa y un ARNt-amidotransferasa (**Figura 5-43, A**). La expresión de dichos genes, excepto *Synpcc7942_2111*, mostraron menor RPKM en la cepa C11 (**Figura 5-43, B**). En ninguno de ellos la diferencia fue significativa, siendo el p-valor ajustado igual a 1 para todos los genes, salvo para el gen *Synpcc7942_2111* y *Synpcc7942_2115* que dicho valor fue 0,236 y 0,379, respectivamente.

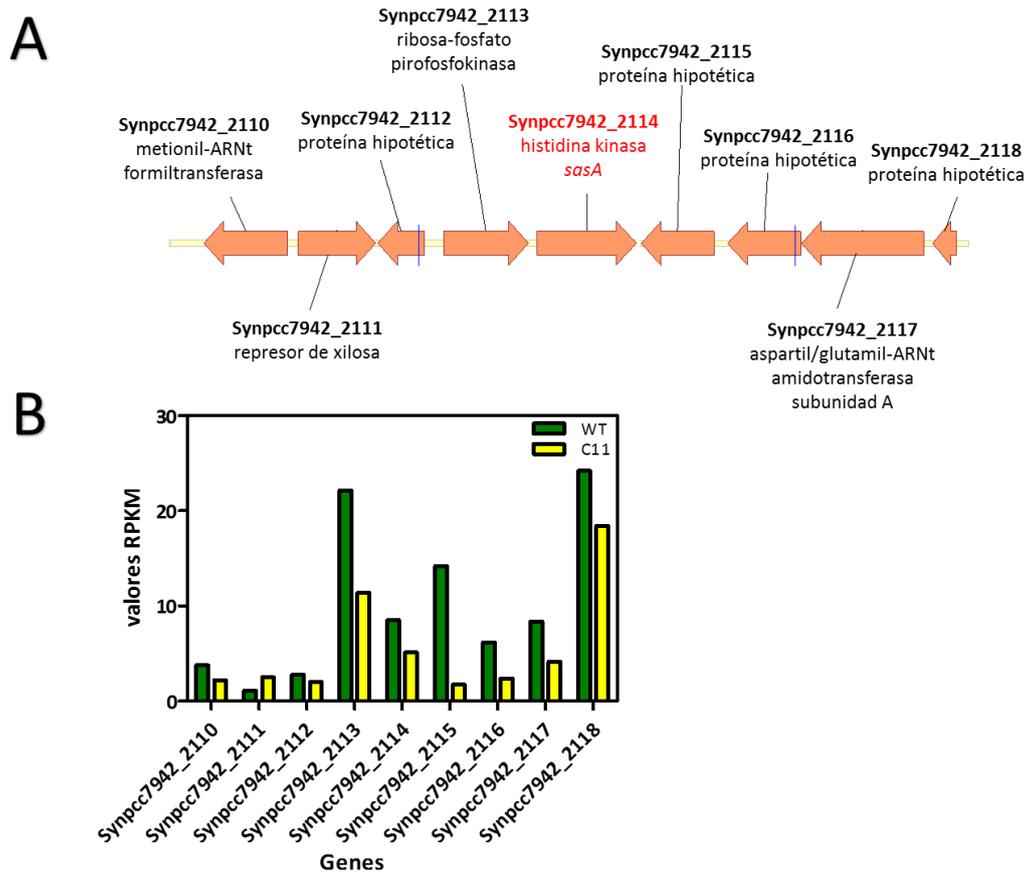


Figura 5-43. Expresión de los genes del entorno genético de la mutación #21. (A) Mapa genético del entorno de *sasA* (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético del gen *sasA* en la cepa silvestre (barras verdes) frente a la cepa C11 (barras amarillas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

A continuación, llevamos a cabo el análisis de la expresión de los genes que conforman el entorno genético de cada una de las siete mutaciones presentes en C12-G₁₁₆₈.

En primer lugar, analizaremos el entorno genético de la mutación #8, que afecta a la subunidad alfa de la ATP sintasa F₀F₁ (**Figura 5-44**). El entorno genético de dicha mutación lo componen genes codificantes de otras subunidades de esta ATP sintasa F₀F₁, dos proteínas hipotéticas y una ferredoxina (**Figura 5-44, A**). En cuatro de estos genes está disminuida la expresión (Synpcc7942_0332, Synpcc7942_0338, Synpcc7942_0339 y Synpcc7942_0340) en C12 y en otros cuatro está aumentada (Synpcc7942_0334, Synpcc7942_0335, Synpcc7942_0336 y Synpcc7942_0337). Ninguno de estos cambios fue significativo, siendo en todos ellos el p-valor ajustado mayor a 0,05.

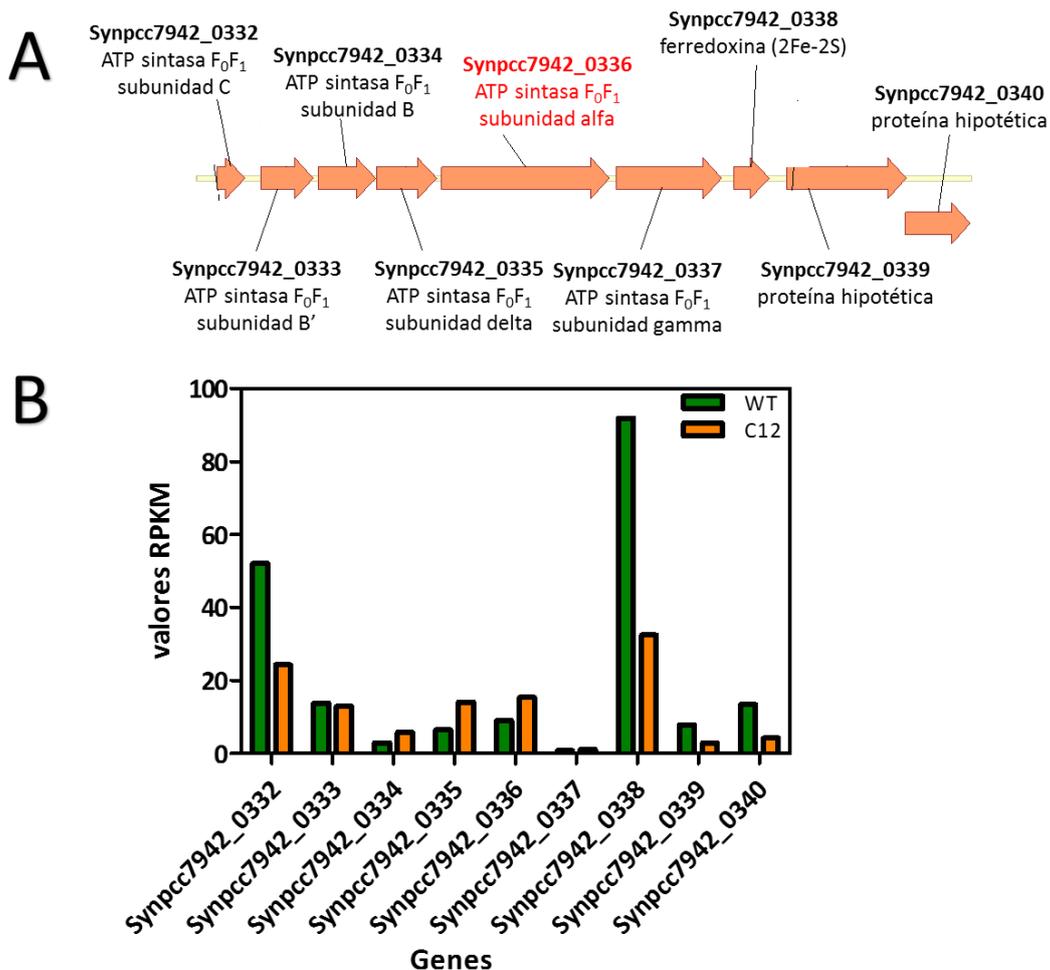


Figura 5-44. Expresión de los genes del entorno genético de la mutación #8. (A) Mapa genético del entorno de Synpcc7942_0336 (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético del gen Synpcc7942_0336 en la cepa silvestre (barras verdes) frente a la cepa C12 (barras naranjas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

A continuación, estudiaremos el entorno genético de la mutación #9, que afecta al gen Synpcc7942_0452, que codifica para una proteína hipotética (**Figura 5-45, A**). En todos ellos disminuye la expresión en comparación con la cepa silvestre (**Figura 5-45, B**). Esta disminución fue más marcada en el caso de los genes Synpcc7942_0448, Synpcc7942_0450, Synpcc7942_0452 y Synpcc7942_0456. Sin embargo, ninguno de estos cambios fue significativo, ya que el p-valor ajustado fue mayor a 0,05 para todos ellos.

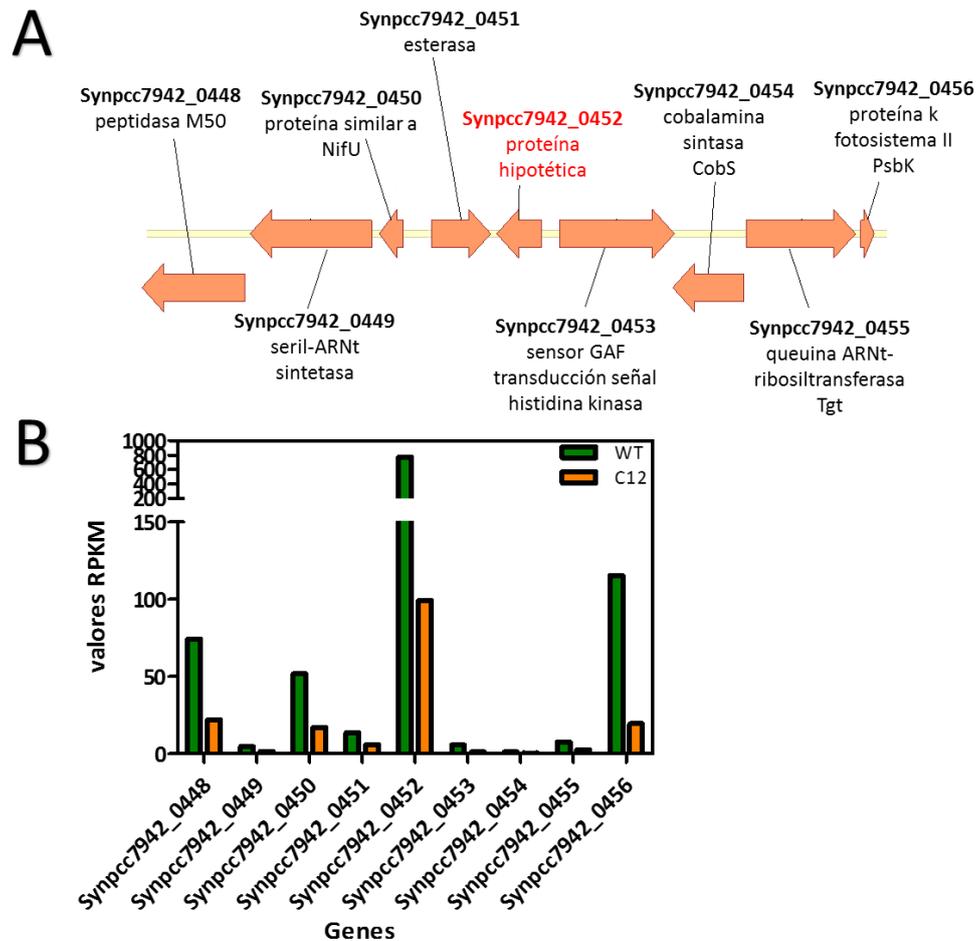


Figura 5-45. Expresión de los genes del entorno genético de la mutación #9. (A) Mapa genético del entorno de Synpcc7942_0452 (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético del gen Synpcc7942_0452 en la cepa silvestre (barras verdes) frente a la cepa C12 (barras naranjas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

En tercer lugar analizaremos el entorno genético de la mutación #13, que afecta a la ácido graso de cadena larga-CoA ligasa (FadD) (**Figura 5-46**). El entorno genético de dicha mutación en su mayoría lo componen genes que codifican para proteínas hipotéticas, aunque también existe un gen de la 3-hidroquinato deshidratasa, un gen del centro de reacción del fotosistema I, una glutamato ARNt ligasa y una proteína similar a ARNt hidroxilasa (**Figura 5-46, A**). La expresión de todos estos genes ha disminuido (**Figura 5-46, B**). Sin embargo, el p-valor ajustado de todos ellos fue 1, salvo en el caso de los genes Synpcc7942_0914 y Synpcc7942_0921, que fue de 0,744 y 0,491, respectivamente.

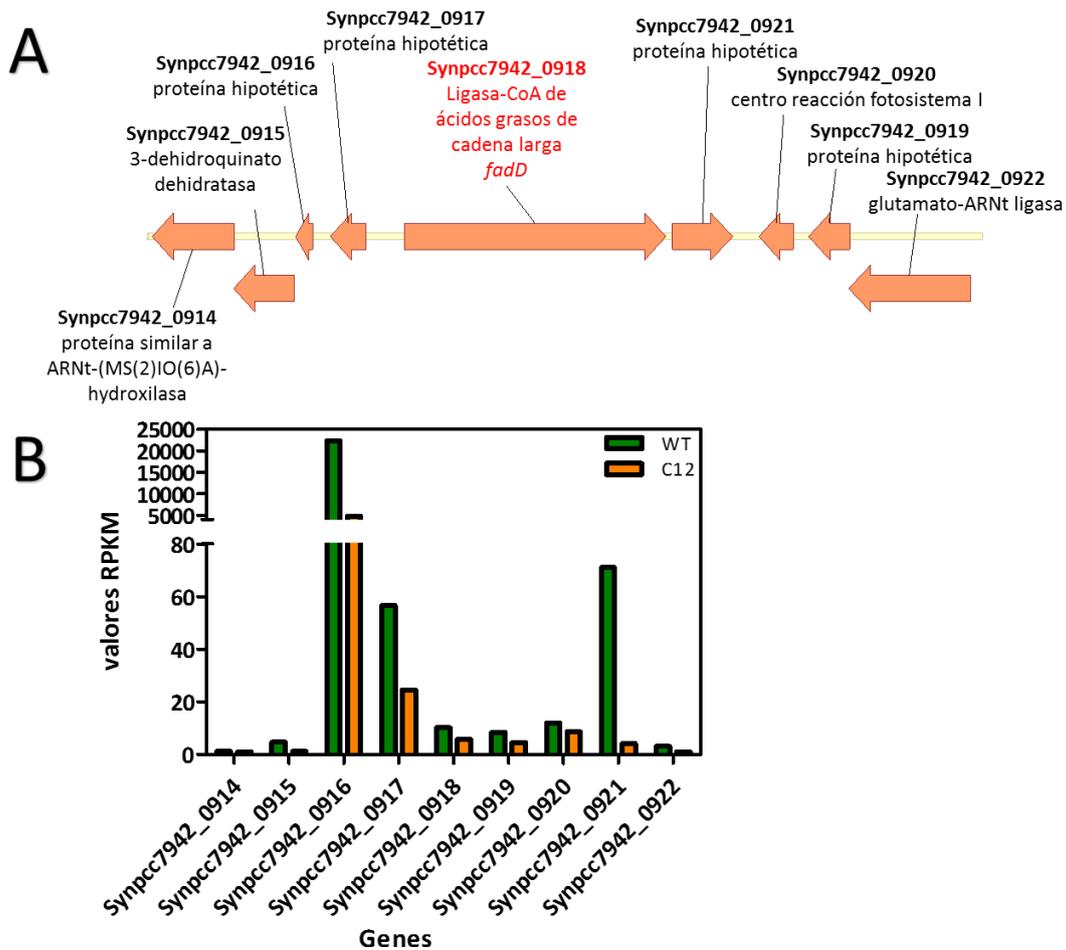


Figura 5-46. Expresión de los genes del entorno genético de la mutación #14. (A) Mapa genético del entorno de *fadD* (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético del gen *fadD* en la cepa silvestre (barras verdes) frente a la cepa C12 (barras naranjas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

La mutación #16 afecta a la región intergénica entre dos genes que codifican para sendos polipéptidos de unión al ficobilisoma (Synpcc7942_1049 y Synpcc7942_1050) (**Figura 5-47**). Su entorno genético en su mayoría lo componen genes que codifican para diferentes subunidades de ficocianina (**Figura 5-47, A**). La expresión de los genes Synpcc7942_1048, Synpcc7942_1049 y Synpcc7942_1053 ha aumentado y la de los genes Synpcc7942_1046, Synpcc7942_1050 y Synpcc7942_1051 ha disminuido (**Figura 5-47, B**). Sin embargo, el p-valor ajustado de todos ellos fue superior a 0,05, siendo 1 para todos ellos, excepto para los genes Synpcc7942_1048 y Synpcc7943_1053, que fue de 0,662.

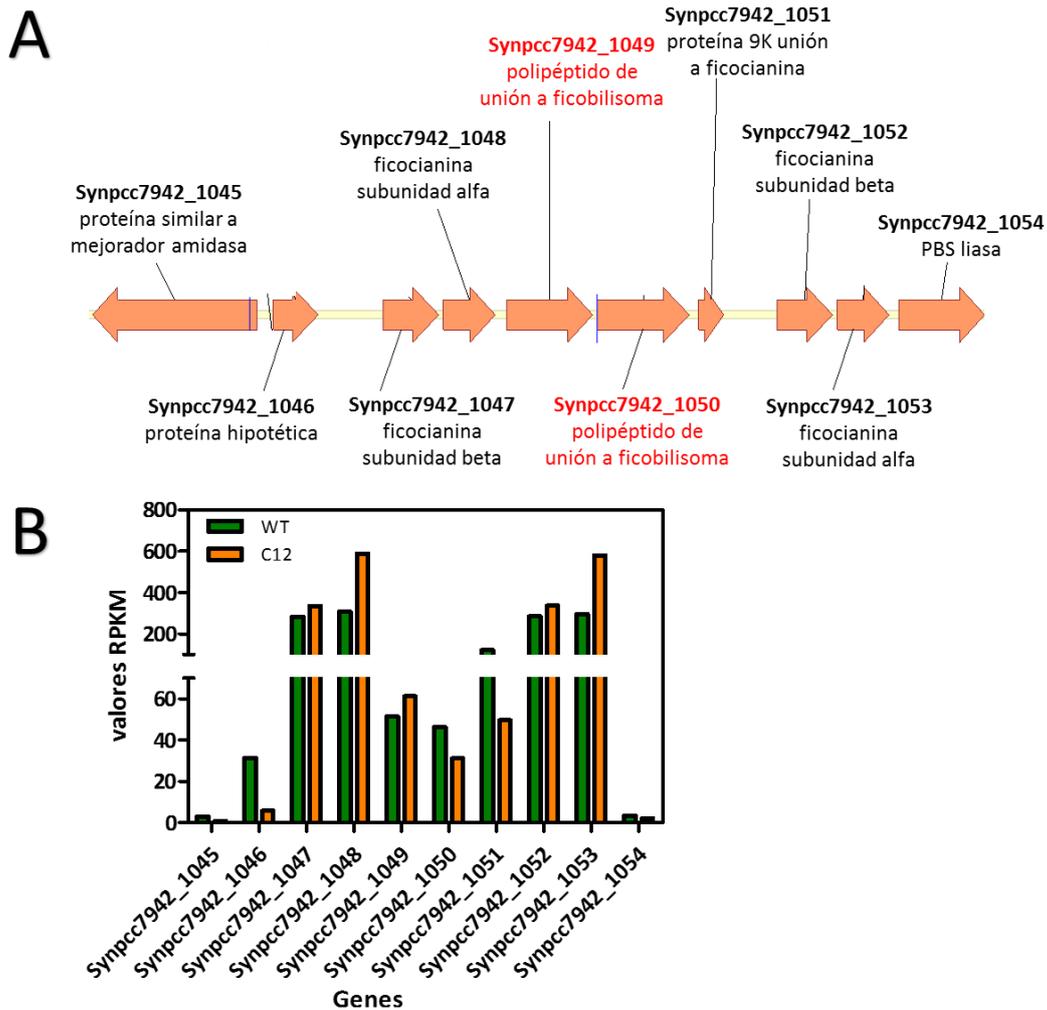


Figura 5-47. Expresión de los genes del entorno genético de la mutación #16. (A) Mapa genético del entorno de Synpcc7942_1049 y Synpcc7942_1050 (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético de esta mutación intergénica en la cepa silvestre (barras verdes) frente a la cepa C12 (barras naranjas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

En quinto lugar, examinaremos el entorno genético de las mutaciones #19 y #20, que afectan a dos genes consecutivos en el genoma de Se7942. El primero de ellos codifica para la proteína PilC, mientras que el segundo codifica para PilT. Su entorno genético está compuesto por genes que codifican para proteínas con funciones diversas (**Figura 5-48, A**), destacando chaperonas moleculares (*dnaJ* y *dnaK*), una ATPasa, una proteína de choque térmico, un ARNt-dihidrouridina sintasa A y varias proteínas hipotéticas. La expresión de todos estos genes ha disminuido (**Figura 5-48, B**), no siendo este cambio significativo, ya que el p-valor ajustado es 1. Las mutaciones #19 y #20 no tienen ningún efecto en la expresión génica de los genes a los que afectan dichas mutaciones, ni a los genes de su entorno genético.

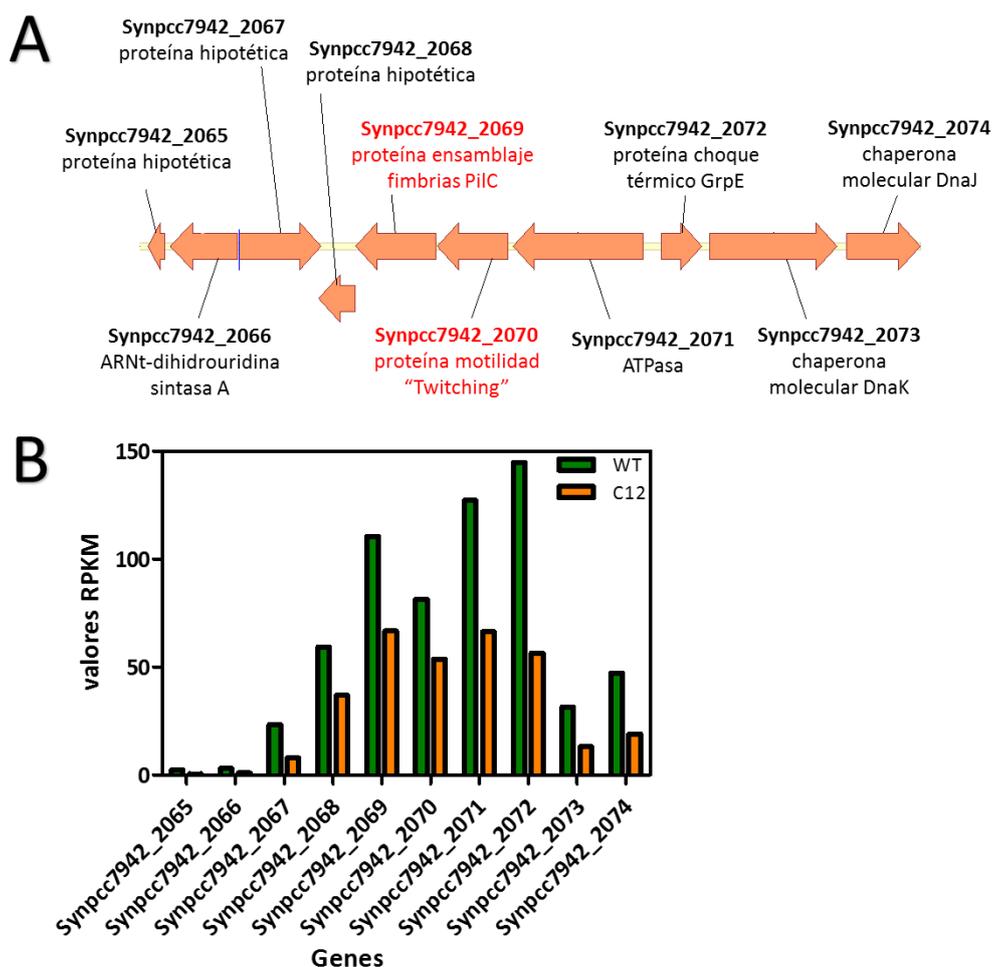


Figura 5-48. Expresión de los genes del entorno genético de las mutaciones #19 y #20. (A) Mapa genético del entorno de Synpcc7942_2069 y Synpcc7942_2070 (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético de estos dos genes en la cepa silvestre (barras verdes) frente a la cepa C12 (barras naranjas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

Por último, examinaremos el entorno genético de la mutación #24, que afecta a la región intergénica entre los genes codificantes de una proteína hipotética y la ARNt i(6)A37 tiotransferasa MiaB. Su entorno genético lo componen genes codificantes para proteínas hipotéticas conservadas y dos proteínas implicadas en la división celular, FtsZ y FtsQ (**Figura 5-49, A**). La expresión de todos ellos disminuye en comparación con los niveles obtenidos en la cepa silvestre (**Figura 5-49, B**), siendo el p-valor ajustado de todos ellos superior a 0,05 (1 para todos, excepto para el gen Synpcc7942_2373 que fue de 0,853).

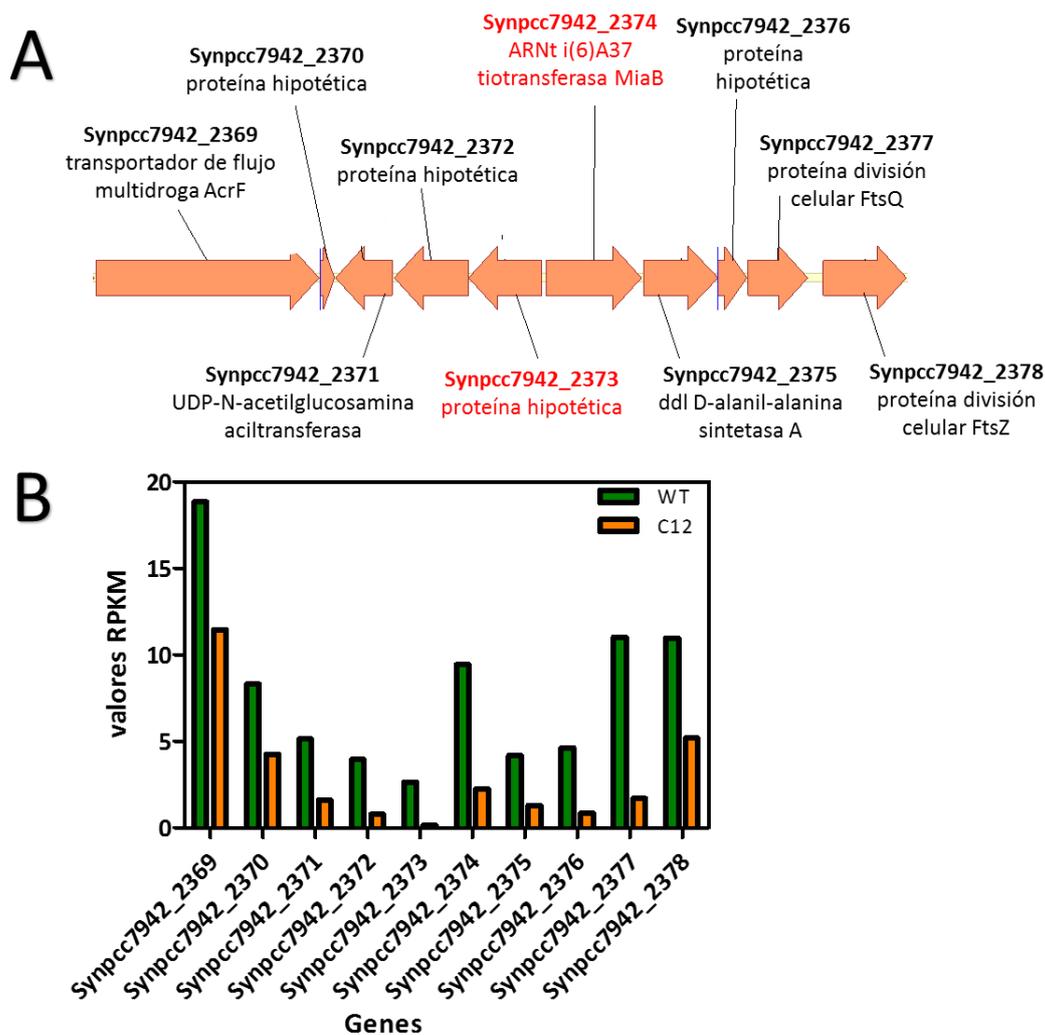


Figura 5-49. Expresión de los genes del entorno genético de la mutación #24. (A) Mapa genético del entorno de Synpcc7942_2373 y Synpcc7942_2374 (en rojo). **(B)** Comparación de los niveles de expresión (medidos como RPKM) de los genes del entorno genético de esta mutación intergénica en la cepa silvestre (barras verdes) frente a la cepa C12 (barras naranjas). RPKM, lecturas por kilobase por millón de lecturas mapeadas.

5.1.7. Efecto de las mutaciones en otras características

Se analizó el efecto de estas mutaciones sobre otras características de Se7942: transformabilidad natural, espectro de absorción de célula completa, composición de pigmentos y formación de biofilms.

La capacidad de Se7942 para absorber ADN exógeno es una característica que se ha utilizado para construir mutantes productores de compuestos de interés industrial. Es por ello que decidimos analizar la capacidad de C11-G₁₂₄₈ y C12-G₁₁₆₈ para ser transformadas naturalmente. Se transformaron estas dos poblaciones y la cepa silvestre utilizando el plásmido pMSM2 y los transformantes se seleccionaron en Sm50. Realizamos además un ensayo control

en el que no añadimos ADN. Los resultados de una de las réplicas en cada condición se muestra en la **Figura 5-50**. Solo se obtuvieron colonias en la transformación natural de la cepa silvestre, no encontrándose ninguna en la transformación de C11-G₁₂₄₈ y C12-G₁₁₆₈. Durante el experimento de evolución, las poblaciones C11 y C12 han perdido la capacidad de transformarse naturalmente. De las mutaciones presentes en estas cepas, algunas están afectando a componentes del pilus, maquinaria involucrada en la toma del ADN. En el caso de C11-G₁₂₄₈, la mutación está afectando a la proteína PilA, mientras que en C12-G₁₁₆₈ hay dos proteínas mutadas, PilC y PilT.

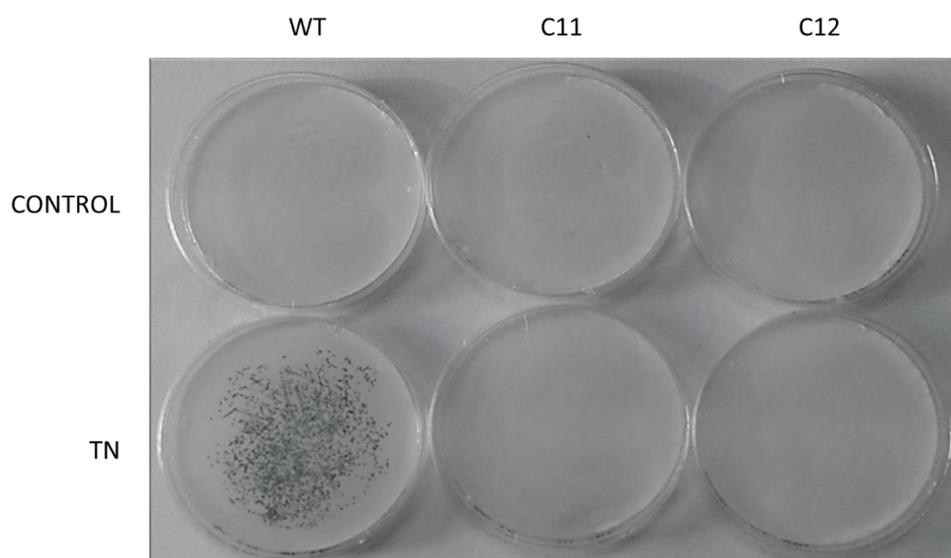


Figura 5-50. Transformación natural de la cepa WT y las poblaciones C11-G₁₂₄₈ y C12-G₁₁₆₈ con el plásmido pMSM2. A la izquierda, se muestra el control y la transformación natural de la cepa silvestre. En el centro, el control y la transformación natural de C11-G₁₂₄₈. A la derecha, el control y la transformación natural de C12-G₁₁₆₈. Control, hace referencia al experimento de transformación natural de cada cepa sin ADN (imágenes superiores). TN, hace referencia a la transformación natural de cada cepa con el plásmido pMSM2 (imágenes inferiores).

Dado que la transformabilidad natural está relacionada con la funcionalidad del pilus tipo IV, analizamos su presencia en la cepa silvestre y las poblaciones C11-G₁₂₄₈ y C12-G₁₁₆₈. Realizamos una tinción negativa de cultivos de estas tres cepas con una solución del 1 % (p/v) de ácido fosfotúngstico (apartado 4.9.6). En la **Figura 5-51** se muestran las micrografías tomadas con el microscopio electrónico de transmisión, observándose que las células de ambas poblaciones evolucionadas están hiperpiliadas. En el caso de C11-G₁₂₄₈, la mutación responsable de este fenotipo sería la inserción de dos codones en el gen *pilA*. En Ss6803, se ha observado que la inactivación de PilA1, proteína homóloga a la inactivada en C11-G₁₂₄₈, provoca una disminución en el número de pili por célula (Bhaya et al., 2000). En el caso de C12-G₁₁₆₈, la cepa presenta mutaciones en los genes *pilC* y *pilT*. En Ss6803, la inactivación de pilT da lugar a células hiperpiliadas (Okamoto and Ohmori, 2002), sin embargo, la inactivación de *pilC* no tiene ningún

efecto sobre el número de pili (Bhaya et al., 2000). Estas imágenes indican que tanto C11-G₁₂₄₈ y C12-G₁₁₆₈ están hiperpiliadas. Sin embargo, se necesita un muestreo más amplio para determinar qué mutaciones son responsables del fenotipo de células hiperpiliadas. Experimentos preliminares sugieren que la mutación en PilT presente en C12-G₁₁₆₈ sería la responsable de dicho fenotipo. En el caso de C11-G₁₂₄₈, no está tan claro que la mutación presente en PilA sea la causante del aumento del número de pili en la célula.

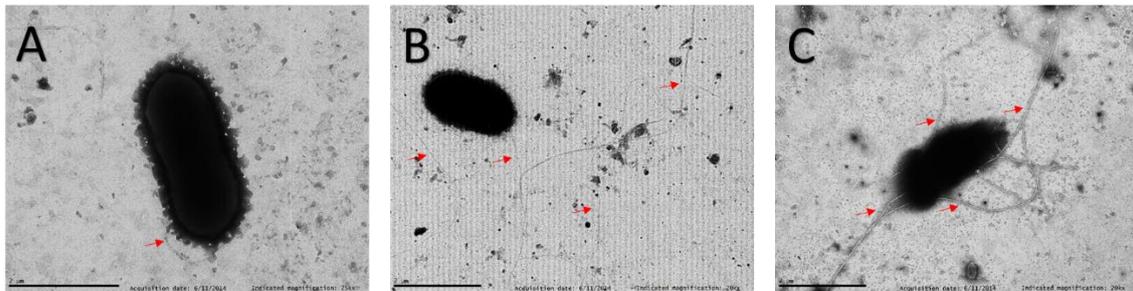


Figura 5-51. Imágenes del microscopio electrónico de transmisión del pilus de varias cepas. Las células fueron teñidas negativamente con una solución del 1 % (p/v) de ácido fosfotúngstico. **(A)** Cepa silvestre (WT). **(B)** Cepa C11-G₁₂₄₈. **(C)** Cepa C12-G₁₁₆₈. Las flechas rojas indican la posición que ocupa el pilus en cada caso. La barra de escala aparece a la izquierda de la imagen, abajo, siendo en todos los casos de 2 μ m.

Otra de las características que estudiamos fue el espectro de absorción de célula completa. Estos experimentos se llevaron a cabo con Se7942 silvestre, con C11-G₁₂₄₈, así como con otros tres mutantes construidos con las mutaciones de C11-G₁₂₄₈ y C12-G₁₁₆₈ (MSM15, MSM3 y MSM5). En la **Figura 5-52** se muestra el espectro de absorción de estos mutantes, correspondiendo cada uno de los picos a los diferentes pigmentos presentes en Se7942. Los picos que aparecen a 440 y 680 nm corresponden a la clorofila a, el pico a 495 nm a los carotenoides y el de 630 nm a las ficobiliproteínas.

Observamos que los espectros de absorción de la cepa evolucionada y los mutantes, con excepción de MSM3, son diferentes a los de la cepa silvestre. En el caso de C11-G₁₂₄₈, los picos de las ficobiliproteínas y el de la clorofila a 680 nm están aumentados. El espectro de absorción de MSM5 es el que más difiere del espectro de las demás cepas.

Dado que la medida de la clorofila a (Chla) en cianobacterias es una forma indirecta de calcular la biomasa, se utilizó el ratio PC/Chla y Ct/Chla para analizar variaciones en dichos pigmentos en los mutantes. Este valor nos da una estimación relativa del contenido de PC y Ct presentes en la célula. El ratio entre la absorbancia de las ficobiliproteínas a 630 nm frente a la absorbancia de la clorofila a 680 nm, fue similar en las cepas silvestre, C11-G₁₂₄₈ y MSM3, en torno a 1,08 - 1,09. En el caso de las cepas MSM15 y MSM5, este valor fue ligeramente superior, 1,14 y 1,117, respectivamente. Por otro lado, en cuanto al ratio entre la absorbancia de los

5. Resultados

carotenoides a 495 nm frente a la absorbancia de la clorofila a 680 nm, fue similar en las cepas silvestre, C11-G₁₂₄₈ y MSM3, en torno a 1,12 – 1,19, como ocurría en el caso anterior. El valor más alto fue obtenido en la cepa MSM5 (1,44) seguido de la cepa MSM15 (1,275).

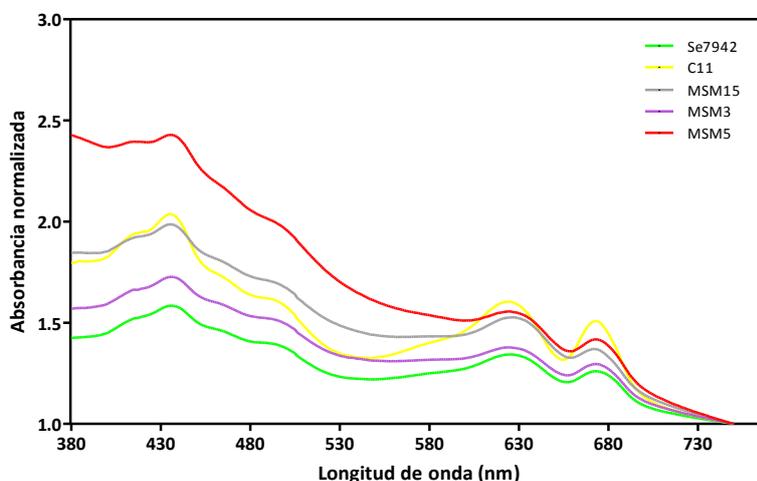


Figura 5-52. Espectro de absorción de célula completa de Se7942, cepa silvestre y derivados. Los datos representan la media de tres réplicas biológicas. El espectro de absorción de las diferentes cepas está normalizado por la DO₇₅₀. El espectro de cada cepa se muestra con un color diferente, indicado en la leyenda de la figura.

Se analizó la concentración de pigmentos de C11-G₁₂₄₈, en los tres mutantes (MSM15, MSM3 y MSM5) y Se7942 silvestre (**Figura 5-53**). En los mutantes MSM15 y MSM5, no se observó un cambio en la concentración de pigmentos analizados, a pesar de que el ratio entre las ficobiliproteínas y la clorofila, calculado a partir del espectro de absorción, fuera diferente al de la cepa silvestre. Lo mismo ocurre con la concentración de carotenoides en estos dos mutantes, que no presenta variaciones, a pesar de que el ratio de carotenoides/clorofila, calculado a partir del espectro de absorción, era mayor.

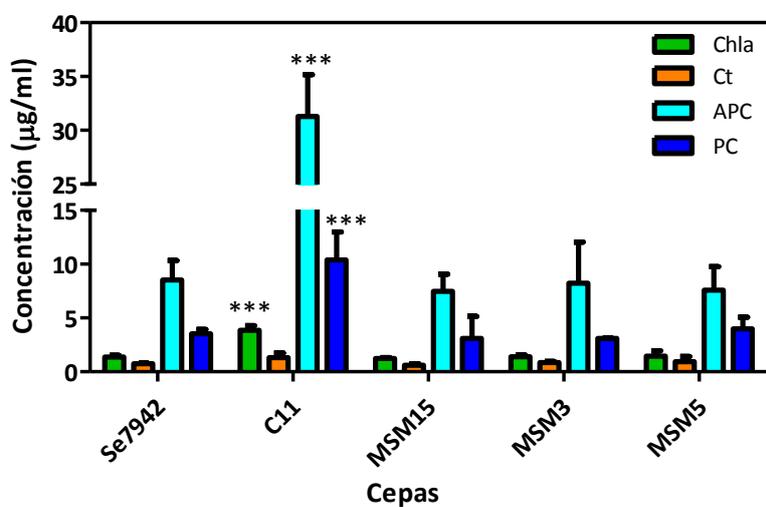


Figura 5-53. Contenido de pigmentos en Se7942, cepa silvestre y derivados. Chla, clorofila a; Ct, carotenoides totales; APC, aloficocianina; PC, ficoeritrina. Los datos corresponden a la media de al menos tres réplicas biológicas y se

representan como la media + DE. *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple, usando como control la cepa Se7942.

En la **Figura 5-52**, habíamos destacado que los picos de las ficobiliproteínas y el de la clorofila a 680 nm estaban aumentados para la cepa C11. Este hecho guarda relación con lo que se observa en la **Figura 5-53**, donde la clorofila y las ficobiliproteínas (APC y PC) presentaban niveles superiores a los de la cepa silvestre, siendo estos aumentos significativos.

En Se7942 se ha demostrado que la inactivación de PilC y/o PilT da lugar a la formación de biofilms (Schatz et al., 2013), involucrando al pilus en la formación de estos biofilms. Dado que C11-G₁₂₄₈ y C12-G₁₁₆₈, presentan mutaciones en algunos de los componentes del pilus, analizamos la formación de biofilms en estas cepas. Añadimos a dicho análisis los mutantes MSM15, MSM3 y MSM5. Como observamos en la **Figura 5-54**, no existen diferencias significativas en la formación de biofilm de las cepas analizadas frente a la cepa silvestre. En el caso de C11-G₁₂₄₈ y C12-G₁₁₆₈, los valores de absorbancia a 630 nm son mayores que los de la cepa silvestre, lo que indicaría una mayor cantidad de biofilms. Sin embargo, las barras error que representan a la desviación estándar son muy grandes. Para extraer conclusiones sería necesario cuantificar estos biofilms utilizando un protocolo menos variable o utilizando técnicas más sensibles y específicas, como puede ser la microscopía electrónica de barrido.

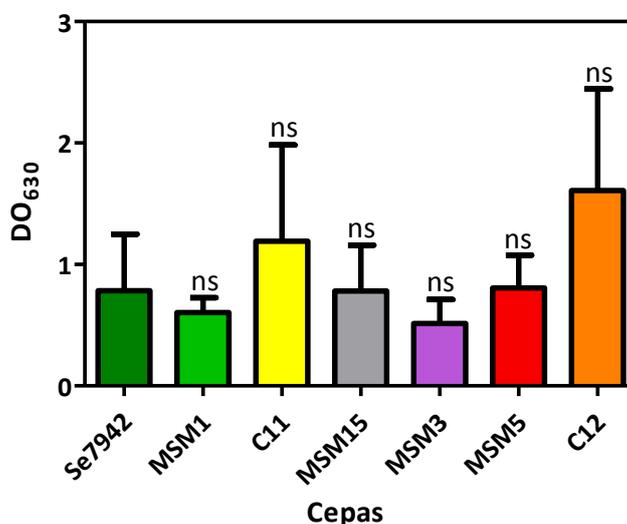


Figura 5-54. Formación de biofilms en Se7942, cepa silvestre y derivados. La medida de la adherencia se realizó por espectrometría de las células fijadas y teñidas con cristal violeta. Los datos corresponden a la media de tres réplicas biológicas y se representan como la media + DE. ns (no significativo) $p > 0,05$, * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa Se7942.

5.2. Síntesis de ácido α -linolénico en Se7942

5.2.1. Condiciones óptimas de inducción de los promotores utilizados para clonar las desaturasas productoras de ácido α -linolénico

Para establecer las condiciones óptimas de producción de ácido α -linolénico (ALA), analizamos la expresión de dos promotores inducibles (*PnrsB* y *Ptrc*) en Se7942. *PnrsB* es un promotor que regula la expresión del operón *nrsBACD* en Ss6803, es inducido por iones Ni^{2+} y presenta una baja transcripción basal en presencia del represor. El sistema de dos componentes NrsRS controla su expresión (Lopez-Maury et al., 2002). El promotor *Ptrc* es un híbrido entre los promotores *trp* y *lacUV5* de *E. coli* y es inducible por IPTG (Amann et al., 1988). Los genes reporteros de la luciferasa (*luxAB*) de *Photorhabdus luminescens* fueron clonados aguas abajo de cada uno de estos promotores. Cada construcción fue introducida en el sitio SN1 del genoma de Se7942 mediante recombinación homóloga (apartado 4.8.3) resultando en las cepas reporteras, MSM18 (*PnrsB::luxAB*) y MSM25 (*Ptrc::luxAB* junto con el gen represor *lacI^q*) (Tabla 4-2).

Para analizar la actividad del promotor *PnrsB*, al cultivo de MSM18 se le añadió una solución 5 μM de NiSO_4 y se midió la luminiscencia durante 72 h después de la inducción (Figura 5-55, A). Se usaron dos controles en este experimento, la misma cepa (MSM18) sin añadir la solución de NiSO_4 y la cepa silvestre de Se7942 que no emite luminiscencia. Como se observa en la Figura 5-55, A, el máximo nivel de expresión de la luciferasa se obtuvo tras 24 h de inducción [3.048 ± 154 Unidades relativas de luminiscencia (URL)]. Este valor fue unas 2,62 más bajo en la cepa MSM18 sin inductor (1.162 ± 148 URL). El hecho de que se produzca luminiscencia en ausencia de níquel, puede deberse a dos motivos. El primero es que el BG11 utilizado lleva Co^{2+} en su composición, que también es un inductor de este promotor. El segundo motivo es que Se7942 no posee en su genoma el operón *nrsRS* que controla la expresión de dicho promotor. Además, este valor continuó creciendo a lo largo de todo el experimento. Este hecho puede explicarse porque el promotor *PnrsB* es inducible también por Co^{2+} , ion metálico presente en el medio BG11 (Rippka et al., 1979). Por otro lado, la cepa silvestre de Se7942 no emite luminiscencia, como cabría esperar.

En el caso del promotor *Ptrc*, se analizaron concentraciones del inductor (IPTG) desde 0 μM a 1 mM. Su actividad se evaluó durante 50 h tras la inducción. En este caso, también usamos dos controles, la cepa MSM25 con la construcción *Ptrc::luxAB* sin añadir IPTG y la cepa silvestre

de Se7942. El máximo nivel de luminiscencia ocurrió tras 24 h de inducción para concentraciones superiores a 50 μM de IPTG (**Figura 5-55, B**). No se observaron diferencias en los niveles de inducción en concentraciones de IPTG de 100 μM , 500 μM y 1 mM (22.367 ± 919 URL, 23.916 ± 391 URL y 25.047 ± 888 URL, respectivamente).

En ambos casos, la máxima actividad luciferasa tuvo lugar tras 24 h de inducción. En el caso del promotor *P_{trc}*, decidimos usar 1 mM de IPTG, ya que es la concentración más utilizada en bibliografía en Se7942.

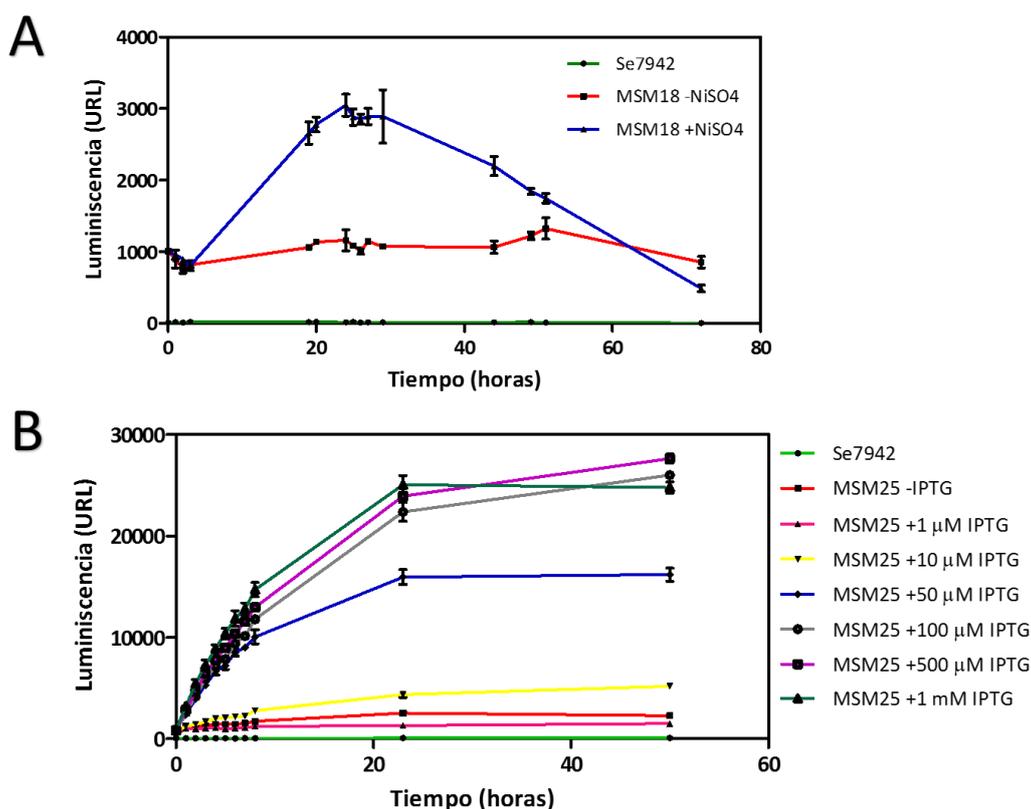


Figura 5-55. Actividad luciferasa en las cepas reporteras de Se7942. (A) Cepa MSM18 añadiendo y sin añadir la solución 5 μM de NiSO_4 . **(B)** Cepa MSM25 inducida con concentraciones crecientes de IPTG. En ambos casos, los datos mostrados representan la media de tres réplicas independientes y se representan como la media \pm DE. Las cepas empleadas se representan en las leyendas de cada panel.

5.2.2. Producción de ácido α -linolénico mediante la sobreexpresión de desaturasas de Ss7002 en Se7942

Para producir ácido α -linolénico en Se7942 es necesario introducir dos desaturasas, DesA y DesB, ya que esta cianobacteria carece de ellas (Wada et al., 1990). DesA introduce un doble enlace en la posición $\Delta 12$ y DesB otro en la posición $\Delta 15$ (ver Introducción, apartado 2.3.2.1). El

primer paso fue seleccionar el organismo del cual se iban a obtener dichos genes. Utilizamos Ss7002, ya que el ARNm de DesA se acumula más rápido en esta cepa (Sakamoto et al., 1997) que en otras cianobacterias, como Ss6803 (Los et al., 1993). Por otro lado, existe información de cómo funcionan estos genes en la producción de ALA en Ss7002 (Ludwig and Bryant, 2011; Ludwig and Bryant, 2012; Sakamoto and Bryant, 1997; Sakamoto and Bryant, 2002; Sakamoto et al., 1997; Sakamoto et al., 1998), además de que no se han utilizado anteriormente en otros organismos para la producción de dicho ácido graso.

Seguidamente, se clonaron estas desaturasas. Para ello, se construyeron tres mutantes de Se7942, con diferentes construcciones insertadas en SN1: MSM16 (*PnrsB::desA*), MSM17 (*PnrsB::desA-desB*) y MSM26 (*Ptrc::desA-desB*) (**Tabla 4-2**). MSM16 contiene la desaturasa DesA, capaz de transformar ácido oleico en ácido linoleico. MSM17 y MSM26 contienen las desaturasas DesA y DesB, capaces de producir ALA, diferenciándose entre sí en el promotor de dichos genes.

El siguiente paso fue determinar las condiciones óptimas para su expresión. El primer análisis de transcripción de estas desaturasas fue llevado a cabo por (Sakamoto and Bryant, 1997) mediante hibridación "Northern Blot". En el caso de *desA*, los transcritos fueron tres veces más abundantes a 22 °C que a 38 °C. Sin embargo en el caso de *desB* solo se detectaron a 22 °C. Más recientemente, (Ludwig and Bryant, 2012) llegaron a las mismas conclusiones utilizando datos de secuenciación ADNc. Se infiere que era necesario trabajar a bajas temperaturas para conseguir niveles altos de expresión de ambas desaturasas, especialmente para DesB. Las primeras condiciones que se probaron para analizar la producción de ácidos grasos fueron bajas temperaturas (22 °C) y baja intensidad de luz (60 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$). Seleccionamos esta intensidad de luz porque es la utilizada normalmente para el crecimiento de Se7942.

Se cultivaron Se7942, Ss7002 y las cepas mutantes (MSM16, MSM17 y MSM26) en las condiciones descritas en el párrafo anterior y se analizó su composición de ácidos grasos. Ss7002 y Se7942 se utilizaron como controles. MSM16, que contiene la construcción *PnrsB::desA* en el sitio SN1, produjo C18:2 (LA, ácido linolénico) (**Tabla 5-5**), mientras que Se7942 no lo produjo. MSM16 produce menos LA que Ss7002 (2.58 % frente a 26.16 %). Este resultado puede deberse a que, o el promotor no es muy fuerte, o el gen *desA* no funciona eficientemente en Se7942, entre otros factores.

Por otro lado, MSM17 y MSM26 produjeron C18:3n-3 (ALA), 2.20 % y 7.49 %, respectivamente. Dado que MSM26 produjo una cantidad de ALA similar a la que produce Ss7002, se descartó la idea de que DesA no fuera funcional en Se7942. En MSM17 y MSM26, se

observó que los niveles de los precursores, C18:1n-9 y C18:2, disminuyeron (**Tabla 5-5**). Además, en el caso de MSM26 se detectó también una reducción de C16:1 (ácido palmitoleico), así como un aumento de C16:0 (ácido palmítico). Cabe destacar que este incremento fue muy marcado en relación a los demás mutantes, pudiéndose ser una de las causas la necesidad de una mayor cantidad de intermediarios de esta ruta con el objetivo de incrementar el producto final, el ALA.

Tabla 5-5. Composición de ácidos grasos de la cepa silvestre de Se7942 y de las cepas modificadas genéticamente para expresar las desaturasas DesA y DesB de Ss7002.

| Cepa | C14:0 | C16:0 | C16:1 | C18:0 | C18:1n-9 | C18:2 | C18:3n-3 |
|---------------|--------|---------|---------|--------|----------|---------|----------|
| Ss7002 | 0,28 ± | 34,45 ± | 20,40 ± | 0,25 ± | 11,50 ± | 26,16 ± | 6,26 ± |
| | 0,26 | 0,44 | 1,12 | 0,21 | 8,29 | 8,30 | 1,80 |
| Se7942 | 1,01 ± | 42,92 ± | 49,37 ± | 0,61 ± | 2,98 ± | 0,00 ± | 0,00 ± |
| | 0,09 | 1,64 | 2,40 | 0,55 | 0,74 | 0,00 | 0,00 |
| MSM16 | 0,98 ± | 46,77 ± | 45,90 ± | 2,81 ± | 7,33 ± | 2,58 ± | 0,00 ± |
| | 0,27 | 4,36 | 3,49 | 2,85 | 2,89 | 0,58 | 0,00 |
| MSM17 | 1,03 ± | 43,47 ± | 46,90 ± | 0,81 ± | 0,67 ± | 1,36 ± | 2,20 ± |
| | 0,14 | 8,16 | 2,55 | 0,74 | 0,21 | 0,98 | 0,59 |
| MSM26 | 1,56 ± | 70,44 ± | 16,22 ± | 2,73 ± | 0,00 ± | 0,00 ± | 7,49 ± |
| | 0,40 | 4,14 | 2,95 | 0,47 | 0,00 | 0,00 | 1,43 |

*Estos datos representan la media de tres o más réplicas de experimentos independientes ± la desviación estándar (DE).

En la **Figura 5-56**, se compara la producción de ALA en nuestras cepas modificadas para la producción de dicho ácido graso y en los controles, Se7942 y Ss7002. El análisis estadístico muestra que no existen diferencias significativas en la producción de ALA entre Ss7002 y el mutante MSM26. Sin embargo, la producción de ALA en MSM17 fue estadísticamente diferente a la de Ss7002 con una $p = 0,0089$. Un bajo nivel de expresión del promotor *PnrsB* en la cepa MSM17 podría ser la razón por la cual la expresión de los genes *desA* y *desB* no produce el mismo incremento en la síntesis de ALA, tal y como sí ocurre en la cepa MSM26. El remplazamiento del promotor *PnrsB* por un promotor más fuerte, *Ptrc*, da lugar a un incremento de 3,4 veces en la producción de ALA.

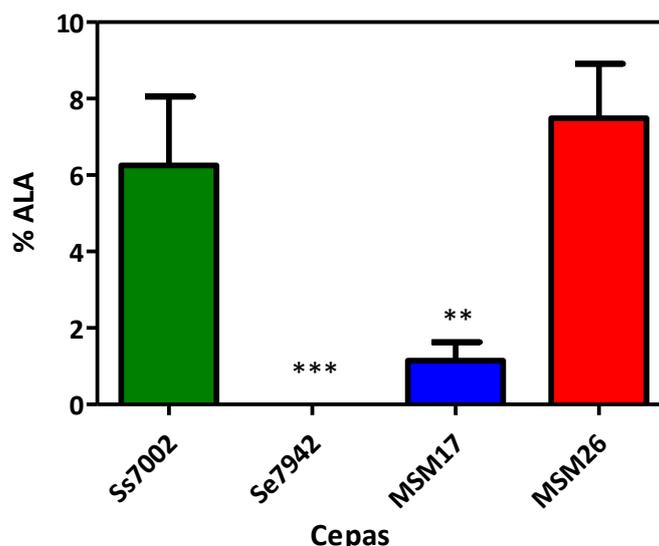


Figura 5-56. Producción de ALA en condiciones de baja intensidad de luz. La barra verde representa a Ss7002, la barra azul a MSM17 y la barra roja a MSM26. Los datos mostrados representan la media de un mínimo de tres biológicas independientes y se representan como la media + DE. **p < 0,01, ***p < 0,001, test t de Student no pareado.

Recientemente, (Yang et al., 2015) han desarrollado una base de datos llamada “Cyanomics” (<http://lag.ihb.ac.cn/cyanomics>) en la que se integran estudios “ómicos” (transcriptómicos y proteómicos) desarrollados en Ss7002. Utilizando estos datos, recogimos en la **Tabla 5-6** aquellas condiciones que favorecen la expresión de las desaturasas en este organismo. De esta manera, podríamos detectar alguna condición que aumentara la producción de ALA.

Tal y como se muestra en la **Tabla 5-6**, en el caso de la desaturasa DesA solo se encontró una condición que podría favorecer la producción de ALA, una alta intensidad de luz. Sin embargo, en el caso de la desaturasa DesB se encontraron seis condiciones que podrían mejorar dicha producción. Entre ellas se encuentra el crecimiento a 22 °C, condición probada en este trabajo y que mejora la producción de ALA (**Tabla 5-5, Figura 5-56**). Las cinco condiciones restantes fueron la limitación de hierro, la limitación de azufre, una alta intensidad de luz, una temperatura de 30 °C y una baja concentración de sal. La única condición común entre ambas desaturasas fue una alta intensidad de luz. Esta condición fue probada experimentalmente en Ss7002 por (Sakamoto et al., 1997), siendo la producción de ALA 3,8 veces más alta bajo estas condiciones lumínicas. Por todo ello, utilizamos dicha condición en los siguientes experimentos junto a la temperatura de 22 °C. La principal diferencia entre los experimentos realizados por (Sakamoto et al., 1997) es que utilizaron una intensidad de 250 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ mientras que en el caso de los datos extraídos de “Cyanomics” utilizaron 900 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. En nuestro caso utilizamos una intensidad de luz de 250 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, ya que en Se7942

intensidades de luz de 500 – 1.000 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ causan fotodaños en las células (Clarke et al., 1995; Kulkarni and Golden, 1995). Además, en Ss7002 con una intensidad de luz de 250 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ya se ve un efecto positivo en la producción de ALA (Sakamoto et al., 1997). Mientras que a 900 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, no existen datos experimentales de la cantidad de ALA producida por Ss7002.

Tabla 5-6. Perfil transcriptómico de las desaturasas DesA y DesB bajo diferentes condiciones de cultivo.

| Proteína | Condiciones* | Fold-change** | p-valor** |
|----------|---------------------------|---------------|----------------------|
| DesA | Alta intensidad de luz | 2,52 | 3,83e ⁻²⁴ |
| DesB | Limitación de hierro | 2,26 | 1,56e ⁻⁰⁵ |
| | Limitación de azufre | 2,68 | 1,05e ⁻⁰⁷ |
| | Alta intensidad de luz | 6,03 | 3,26e ⁻³⁸ |
| | 30 °C | 2,32 | 0,00010897 |
| | Baja concentración de sal | 2,57 | 0,00017609 |
| | 22 °C | 5,41 | 1,37e ⁻²⁷ |

*Alta intensidad de luz, 900 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$; Limitación de hierro, adición al medio de 720 μM mesilato de deferoxamina (quelante de Fe); Limitación de azufre, sustitución del $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$ por $\text{MgCl}_2\cdot 6\text{H}_2\text{O}$; Baja concentración de sal, medio con 3 mM de NaCl y 0,08 mM de KCl; Temperatura, las condiciones estándar de crecimiento de Ss7002 son 38 °C.

**Los genes con “fold-change” ≥ 2 y “p-valor” $< 0,05$ están considerados significativamente sobreexpresados.

En la **Figura 5-57**, se muestra el análisis de la producción de ALA de Ss7002 y nuestra cepa mutante MSM26, que contiene *P_{trc}::desA-desB*. Estos experimentos se realizaron a baja temperatura (22 °C), y utilizando dos condiciones de luz: alta intensidad (AI), que corresponde a 250 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, y baja intensidad (BI), que corresponde a 60 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. La producción de ALA aumentó del 6,26 % en BI al 14,44 % en AI en la cepa Ss7002 (**Figura 5-57, B**). En cambio, en la cepa MSM26 no se observaron cambios significativos en la producción de ALA al variar las condiciones de iluminación (**Figura 5-57, B**): 7,49 % en BI frente a 8,70 % en AI.

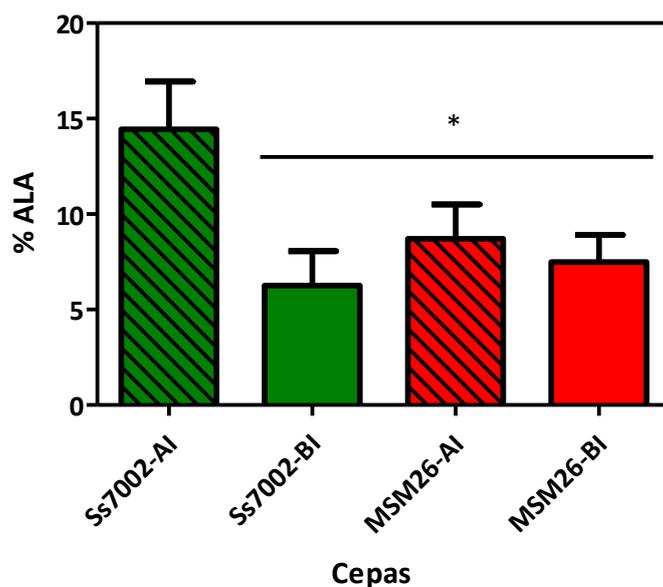


Figura 5-57. Influencia de la intensidad lumínica en la producción de ALA. Comparación de la producción de ALA en diferentes condiciones de luz de la cepa Ss7002 y del mutante MSM26. Las barras verdes representan a la cepa Ss7002 y las barras rojas a la cepa MSM26. AI, alta intensidad de luz, $250 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (barras rayadas); BI, baja intensidad de luz, $60 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (barras lisas). Los datos mostrados representan la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, test t de Student no pareado.

Estos resultados indican que los genes clonados son funcionales en Se7942 y las proteínas que codifican, producen niveles de ALA similares a los de la cepa original, Ss7002. Curiosamente, en la cepa MSM26, la cepa modificada genéticamente con una mayor producción de ALA, el ácido graso C16:1 es consumido para la producción de ALA, al igual que los intermediarios C18 (Tabla 5-5). Por ello, nuestro siguiente objetivo fue aumentar la conversión de los ácidos grasos C16 a C18 mediante la modificación de la expresión de las enzimas involucradas en la ruta de síntesis de ácidos grasos. De esta manera al aumentar los sustratos utilizados en la producción de ALA, podríamos aumentar también la producción de este ácido graso.

5.2.3. Efecto de la modificación de la expresión de los genes de síntesis y degradación de ácidos grasos en Se7942

Para investigar el efecto de las modificaciones en los genes de la ruta de síntesis y degradación de ácidos grasos de Se7942 en la producción de intermediarios de la síntesis de ALA, utilizamos la información descrita para esta ruta en *E. coli* (ver Introducción, apartado 2.3.1.1). De esta manera, analizamos cinco dianas potenciales, los genes *fabD*, *fabH*, *fabB*, *fabF* y *fadD*. En todos los casos, las condiciones de cultivo fueron $30 \text{ }^\circ\text{C}$ y $60 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, y las de inducción (cuando fuera necesario) $22 \text{ }^\circ\text{C}$ y $60 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. La idea era sobreexpresar estos genes junto con las desaturasas, por lo que hubo que usar una baja

temperatura de inducción, a la cual hemos demostrado en el apartado anterior una alta funcionalidad de estas enzimas.

5.2.3.1. Efecto de la sobreexpresión del gen *fabD* en Se7942 en la producción de ácidos grasos

FabD es una malonil-CoA ACP transacilasa (MCAT) que cataliza la transferencia de la fracción malonil al ACP, dando lugar a malonil-ACP (Joshi and Wakil, 1971) (ver Introducción, apartado 2.3.1.1). La sobreexpresión de dicho gen en *E. coli* altera la composición de ácidos grasos, aumentando el ácido *cis*-vaccénico (C18:1n-7) y disminuyendo el ácido palmitoleico (C16:1) (Magnuson et al., 1992).

En este trabajo, hemos modificado Se7942 para que sobreexpresara la proteína FabD, para ver si tiene el mismo efecto observado en *E. coli*. Para ello, construimos la cepa MSM24 que lleva en el SN1 la construcción *P_{trc}::fabD* y analizamos su composición de ácidos grasos. MSM24 no muestra modificaciones en la composición de ácidos grasos (Figura 5-58). Los pequeños cambios observados en varios ácidos grasos, como por ejemplo C14:0 y C18:1n-9, no presentan diferencias significativas con Se7942. Concretamente, no se confirma el efecto observado en *E. coli*, ya que no se produce un aumento de C18:1 y disminución de C16:1. Además, no se favorece la producción de ninguno de los intermediarios necesarios en la producción de ALA. Por ello, descartamos la utilización de esta enzima para mejorar la producción de ALA.

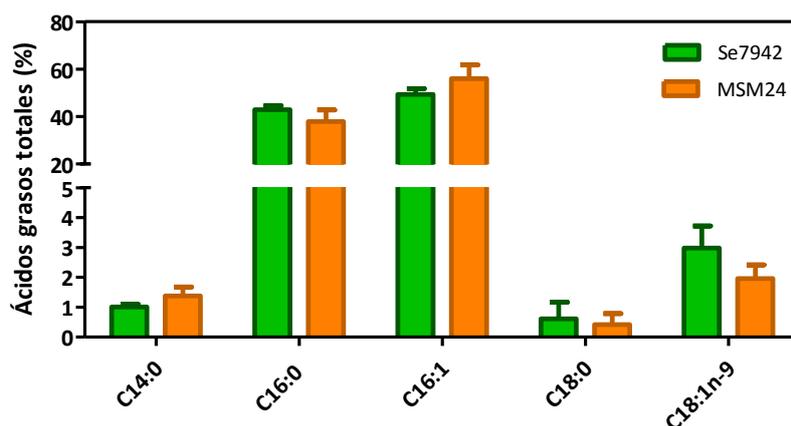


Figura 5-58. Efecto de la sobreexpresión del gen *fabD* en el contenido de ácidos grasos de Se7942. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, test t de Student no pareado.

5.2.3.2. Efecto de la sobreexpresión y delección del gen *fabH* en Se7942

La siguiente diana de la ruta a modificar fue la enzima β -cetoacil sintasa III (FabH), que cataliza la unión del malonil-CoA con acetil-CoA, formando β -cetoacil-ACP (ver Introducción, apartado 2.3.1.1). La sobreexpresión de FabH en *E. coli* aumenta los ácidos grasos C14 y C16 (Tsay et al., 1992), mientras que la delección del gen produce un aumento de los niveles de ácidos grasos C18 (Yao et al., 2012). En Ss7002, (Kuo and Khosla, 2014) sugieren que la expresión endógena de FabH o la sustitución por el ortólogo de *E. coli* debería aumentar el flujo de ácidos grasos.

Construimos dos mutantes de Se7942: MSM28, que sobreexpresaba el gen *fabH* de Se7942 bajo el promotor *P_{trc}* y MSM23 con una delección de *fabH* (Tabla 4-2). Para su construcción, utilizamos el sistema diseñado por (Matsuoka et al., 2001), mediante el cual sustituimos el gen *fabH* de la cepa MSM1 por un gen de resistencia a kanamicina y una copia silvestre del gen *rps12* de Ss6803 (Figura 5-59, A). Después de tres pases de estos transformantes en placas con el antibiótico correspondiente (Neo), analizamos los mutantes mediante PCR usando los oligonucleótidos 111 y 112, observándose dos bandas en el gel de electroforesis. La banda de mayor tamaño (~ 2,6 kb) pertenece a aquellas copias del cromosoma en las que se ha eliminado el gen *fabH*, mientras que la banda de menor tamaño (~ 2,1 kb) corresponde a las copias silvestres del cromosoma (Figura 5-59, B). Este gen ha sido descrito recientemente como un gen esencial por (Rubin et al., 2015). Por esta razón, cuando tratamos de eliminarlo obtenemos un mutante con una delección parcial del gen *fabH*.

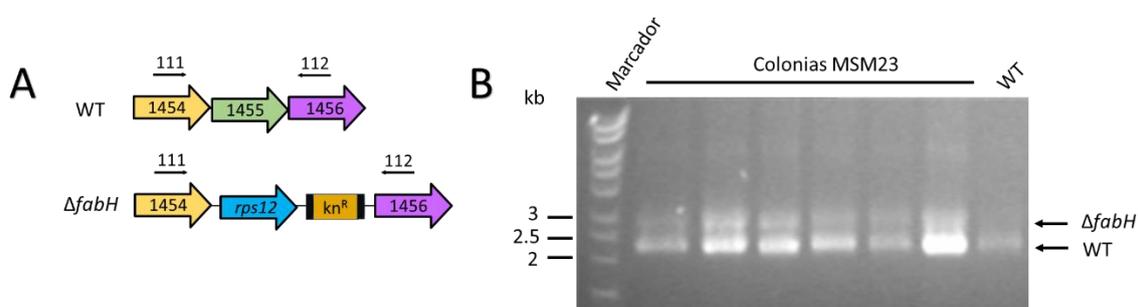


Figura 5-59. Construcción del mutante MSM23 deficiente en el gen *fabH*. (A) Mapa de la región *fabH* (1455) del cromosoma de este mutante así como de la cepa silvestre (WT). Varias copias del cromosoma de este mutante tienen la delección del gen *fabH* y las otras tienen copias silvestres del cromosoma. Las flechas y los números por encima del mapa representan los oligonucleótidos usados en el análisis por PCR. (B) Gel de electroforesis de la PCR para analizar la región *fabH* del genoma de 6 colonias de MSM23 y de la cepa silvestre (WT). La flecha superior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la delección del gen *fabH*; la flecha inferior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la región silvestre del gen *fabH*. Como marcador se utilizó "HyperLadder I" (BioLabs).

Con el objeto de conseguir la delección del gen *fabH*, construimos una serie de mutantes de Se7942 en cuyo cromosoma habíamos introducido una copia extra del gen *fabH* previamente. Construimos dos mutantes, MSM30 y MSM28, que contienen en el sitio SN1 una copia extra de

este gen bajo un promotor inducible, *P_{trc}* y *P_{nrsB}* con un “riboswitch” dependiente de teofilina (Nakahira et al., 2013), respectivamente (**Tabla 4-2**). Sobre estos mutantes, realizamos posteriormente la delección del gen *fabH*.

Antes de realizar dicha delección y su correspondiente análisis, estudiamos el promotor *P_{nrsB}* con el “riboswitch” dependiente de teofilina de la cepa MSM22 (**Tabla 4-2**). Esta cepa lleva en SN1 los genes *luxAB* de *Photorhabdus luminescens* bajo el promotor *P_{nrsB}* y este “riboswitch”. En ausencia de teofilina, el “riboswitch” forma una estructura en forma de lazo con el RBS en el ARNm. El “riboswitch” cambia su conformación cuando se une la teofilina al aptámero, dejando libre el RBS y permitiendo la iniciación de la traducción del gen de interés (**Figura 5-60**).



Figura 5-60. Esquema del modelo funcional del “riboswitch” dependiente de teofilina. En verde, se representa la secuencia del aptámero donde se une la teofilina; en rojo, se representa la secuencia del sitio de unión al ribosoma (RBS); en amarillo, se representa el gen de interés y el codón de inicio de la traducción del mismo, AUG; en gris, se representa el ribosoma; el hexágono verde, representa una molécula de teofilina. Modificado de (Nakahira et al., 2013).

Para analizar la actividad del promotor *P_{nrsB}* con el “riboswitch” de teofilina, se añadió al cultivo de la cepa MSM22 una solución 5 μ M de NiSO_4 y otra 2 mM de teofilina y se midió la luminiscencia durante 96 h tras la inducción (**Figura 5-61**). Se usaron como controles, la misma cepa sin añadir teofilina ni NiSO_4 , o añadiendo solo NiSO_4 , o añadiendo solo teofilina y la cepa silvestre de Se7942 que no emite luminiscencia (**Figura 5-61**). El máximo nivel de luminiscencia se obtuvo tras 66 h de inducción (448 ± 17 URL). Este valor fue aproximadamente 6 veces más alto que en ausencia de inductores (74 ± 16 URL). Además, cuando añadimos un solo inductor (NiSO_4 o teofilina) el nivel de inducción fue similar al de la cepa silvestre de Se7942 y a la cepa MSM22 sin inductores. Cuando comparamos el proceso de inducción de la cepa MSM22 (**Figura 5-61**) frente a la cepa MSM18 (**Figura 5-55, A**), cepa que lleva en el sitio SN1 la construcción *P_{nrsB}::luxAB*, observamos varias diferencias. Por un lado, el máximo nivel de inducción de MSM22 fue más bajo que el de la cepa MSM18 y tuvo lugar más tarde, 66 h frente a 24 h. Por otro lado, la luminiscencia estuvo más controlada en MSM22, porque el nivel de inducción de la cepa MSM22 en ausencia de uno de los inductores fue similar al de la cepa silvestre. Este sistema

nos permitirá obtener la expresión de nuestro gen de interés de forma controlable y en un nivel muy bajo.

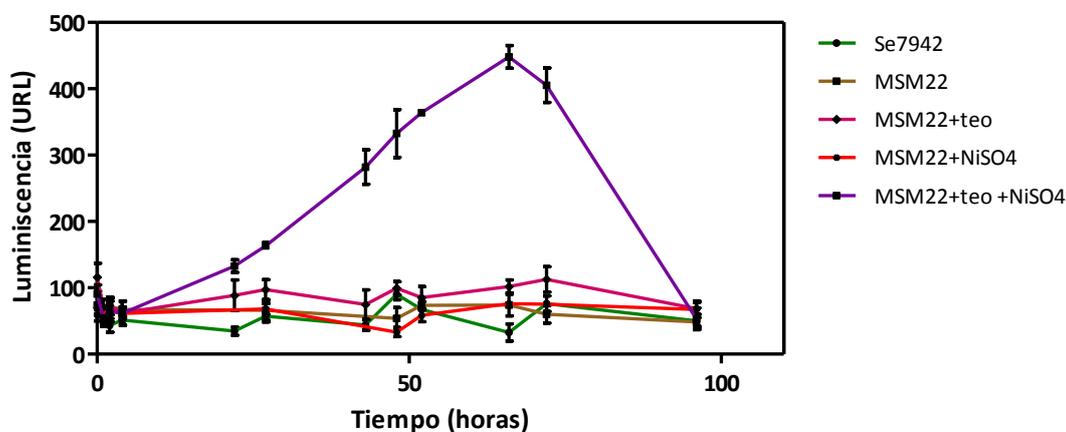


Figura 5-61. Actividad luciferasa de la cepa reportera MSM22. En verde, cepa silvestre de Se7942; en marrón, cepa MSM22 en ausencia de inductores; en rosa, cepa MSM22 en presencia de 2 mM de teofilina; en rojo, cepa MSM22 en presencia de 5 μ M de NiSO₄; en morado, cepa MSM22 en presencia de 5 μ M de NiSO₄ y 2 mM de teofilina. Los datos de luminiscencia a diferentes tiempos son la media de tres réplicas independientes y se representan como la media \pm DE.

Una vez determinado el nivel de inducción del promotor *PnrsB* con el “riboswitch” dependiente de teofilina, llevamos a cabo la delección del gen *fabH* en las cepas MSM30 y MSM28. Como se ha comentado anteriormente, estas cepas contienen una copia extra de este gen en SN1 bajo un promotor inducible, *Ptrc* y *PnrsB* con un “riboswitch” dependiente de teofilina, respectivamente. Se usó el plásmido pMSM182 para la delección del gen *fabH* mediante transformación natural, obteniéndose las cepas MSM34 y MSM35, respectivamente (**Tabla 4-2**). Después de tres pases de los transformantes en Neo, se analizaron mediante PCR usando los oligonucleótidos 111 y 112 (**Figura 5-59, A**). Se utilizaron como control de esta PCR las colonias de la cepa MSM23 obtenidas y analizadas en la PCR de la **Figura 5-59, B**. El mutante MSM23 presenta una delección parcial del gen *fabH*. Los resultados de dicha PCR se muestran en la **Figura 5-62**. Como ocurría en la **Figura 5-59, B**, la banda de mayor tamaño (~2.6 kb) pertenece a aquellas copias del cromosoma en las que se ha eliminado el gen *fabH*, mientras que la banda de menor tamaño (~2.1 kb) corresponde a las copias silvestres del cromosoma. En la **Figura 5-62** puede observarse que, tanto en las colonias de la cepa MSM35 como en las de la cepa control, MSM23, se obtienen copias del cromosoma de ambos tipos. Sin embargo, en el caso de las colonias de la cepa MSM34, solo se observa una banda que corresponde a la delección del gen *fabH*. La eliminación de este gen solo pudo llevarse a cabo cuando la cepa contenía una copia extra del gen bajo un promotor fuerte sin inducir, como es el caso de la cepa MSM34. Sin embargo, cuando usamos un promotor débil, en el caso de la cepa MSM35, la eliminación del

gen *fabH* fue parcial. Esto es lo mismo que ocurrió en la cepa silvestre cuando tratamos de eliminar dicho gen, obteniendo así la cepa MSM23.

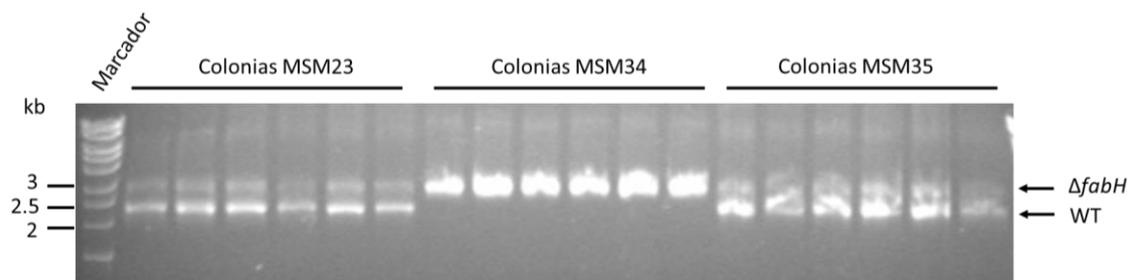


Figura 5-62. Análisis mediante de PCR de la región del gen *fabH*. Se muestran los amplicones obtenidos por PCR con los cebadores 111 y 112, usando ADN molde proveniente de colonias de las cepas MSM23, MSM34 y MSM35. La flecha superior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la deleción del gen *fabH*; la flecha inferior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la región silvestre del gen *fabH*. Se usó como marcador "HyperLadder I" (BioLabs).

Una vez construidas todas estas cepas, analizamos la composición de ácidos grasos de MSM23, que tiene una deleción parcial del gen *fabH*, y MSM28, que sobreexpresa este gen. El mutante MSM28 (**Figura 5-63**, barras amarillas), produce C12:0, ácido graso ausente en la cepa control Se7942. Se observa también un aumento significativo en la producción de otros ácidos grasos de cadena corta (C14:0 y C16:1) y una disminución de ácidos grasos de cadena larga (C18:1). Por el contrario, en el mutante MSM23 (**Figura 5-63**, barras rosas), la producción de ácidos grasos de cadena larga aumenta (C18:0 y C18:1n-9) y disminuye la de C16:1. Podemos decir, por tanto, que el efecto de la deleción parcial del gen *fabH* podría ser beneficioso en la producción de ALA, ya que en este mutante aumentan los intermediarios C18 y la célula podría tener más sustrato para producir nuestro ácido graso de interés.

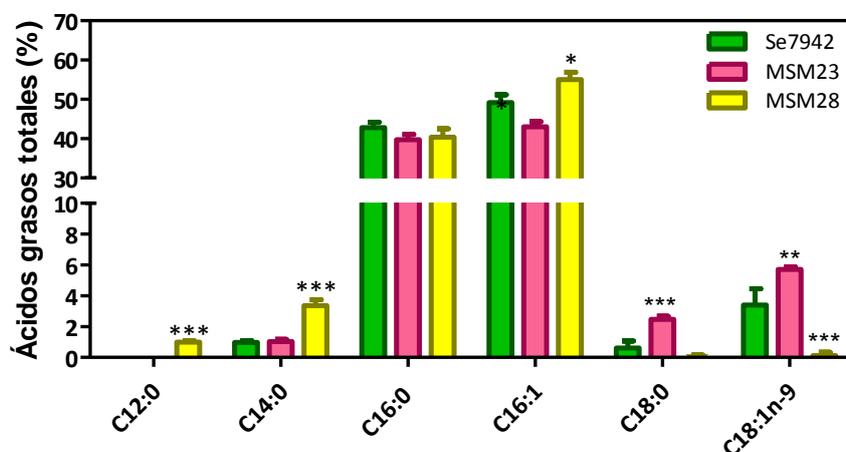


Figura 5-63. Efecto de la deleción parcial del gen *fabH* en el contenido de ácidos grasos de Se7942. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos representan la media de un mínimo

de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, test t de Student no pareado.

5.2.3.3. Efecto de la sobreexpresión del gen *fabB* en el contenido de ácidos grasos de Se7942

FabB es una β -oxoacil sintasa I que cataliza la condensación del acil-ACP con otro malonil-CoA, en la ruta cíclica de síntesis de ácidos grasos (Edwards et al., 1997) (ver Introducción, apartado 2.3.1.1). La delección del gen *fabB* en *E. coli* da lugar a un mutante auxótrofo para ácidos grasos insaturados (Cronan et al., 1969) mientras que la sobreexpresión de este gen aumenta la cantidad de C18:1 (Clark et al., 1983).

Cuando comparamos los genes que codifican para las enzimas que participan en la ruta de síntesis de ácidos grasos en Ss7002 y Se7942, observamos que *fabB* está presente en Ss7002 pero no en Se7942. En la **Tabla 2-3**, ya se mostró que de seis cianobacterias analizadas solo una de ellas, Ss7002, tenía dicho gen. (Kuo and Khosla, 2014) determinaron que FabB no era esencial en Ss7002 y especularon que pudiera tener un papel adicional en esta cianobacteria de crecimiento rápido. Por esta razón y porque su sobreexpresión en *E. coli* aumenta el intermediario C18:1, decidimos identificar la función de este gen en Se7942.

Con objeto de sobreexpresar el gen *fabB* en Se7942, clonamos *fabB* de Ss7002 bajo los promotores *PnrsB* y *Ptrc*. Generamos así dos cepas mutantes, MSM19 y MSM27, respectivamente (**Tabla 4-2**), que tenían la construcción integrada en SN1 del cromosoma de Se7942. En la **Figura 5-64** se muestra la composición de ácidos grasos de estas dos cepas y de la cepa silvestre de Se7942, utilizada como control. En ambos casos, la sobreexpresión del gen *fabB* aumentó significativamente la cantidad de C14:0. Este hecho indicaría que FabB presenta una mayor afinidad por sustratos de cadena corta, tal y como ocurre en *E. coli* y que ha sido descrito por (Edwards et al., 1997). Por otro lado, en la cepa MSM27 (**Figura 5-64**, barras moradas) la cantidad de C16:0 disminuye, mientras que la de C16:1 aumenta. Este efecto no se observa en la cepa MSM19, por lo que podría estar relacionado con la cantidad de FabB que se obtiene al sobreexpresar, ya que MSM19 lleva un promotor más débil que MSM27. No obstante, el efecto de este gen en la producción de ácidos grasos no es muy marcado en ninguna de las dos cepas, por esta razón decidimos utilizar otro sistema para sobreexpresar este gen.

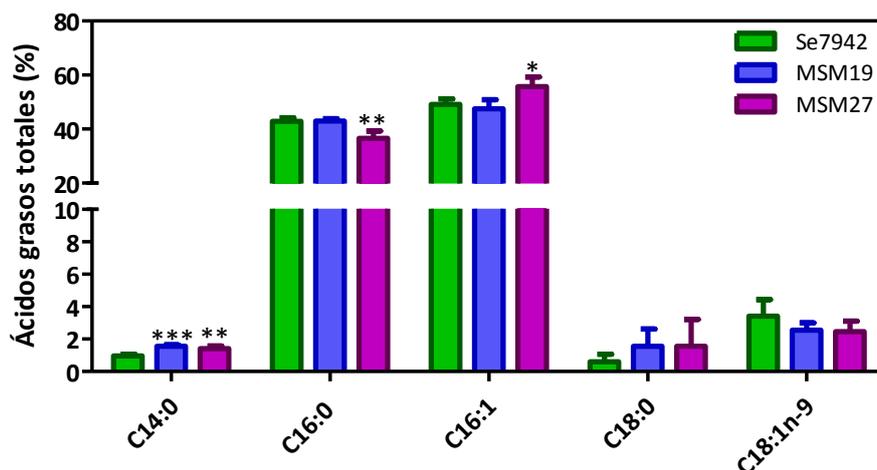


Figura 5-64. Efecto de la sobreexpresión del gen *fabB* en el contenido de ácidos grasos de Se7942. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa Se7942.

Para ello, construimos otro mutante, MSM20, que lleva el gen *fabB* clonado bajo el promotor *PnirA* de Se7942 descrito por (Aichi et al., 2004) (Tabla 4-2). Como en casos anteriores, antes de analizar la composición de ácidos grasos de la cepa MSM20, analizamos la actividad del promotor *PnirA*. Para ello, construimos la cepa MSM21 que lleva clonados los genes *luxAB* de *Photorhabdus luminescens* bajo el promotor *PnirA*. En este caso, la construcción no se introdujo por transformación natural en SN1, sino que se utilizó un plásmido, pMSM132, que permitía introducir la construcción por conjugación. pMSM132 fue construido a partir del plásmido pDEP6, que lleva el origen de transferencia de RP4, para permitir la conjugación de *E. coli* a Se7942. También contiene el origen de replicación del plásmido pANL, que permite recombinar con pANL del receptor y formar cointegrados (Encinas et al., 2014). En presencia de amonio el promotor *PnirA* se encuentra reprimido y en presencia de nitrato se encuentra activo (Aichi et al., 2004). Para analizar la actividad del promotor *PnirA*, se cultivó MSM21 en medio BG11₀ con 3,75 mM (NH₄)₂SO₄ para mantener el promotor reprimido. Cuando el cultivo alcanzó una DO₇₂₀ de ~ 0,5, se lavaron los cultivos en medio BG11₀ y se inocularon las células en diferentes medios. Se utilizó medio BG11, medio BG11₀ con 3,75 mM (NH₄)₂SO₄ y medio BG11₀ con 60 mM KNO₃, siendo los dos primeros controles del experimento. También se usó como control la cepa Se7942 crecida en BG11 que no emite luminiscencia. Se midió la luminiscencia emitida por las células durante 48 h tras la inducción (Figura 5-65).

Como se observa en la Figura 5-65, el máximo nivel de luminiscencia se obtuvo tras 24 h de inducción (1.357.608 ± 17.132 URL). Por otro lado, este promotor presenta un alto nivel de expresión basal cuando está reprimido con amonio, comparable al que presenta la cepa MSM21

en medio BG11. El medio BG11 contiene 17.65 mM de NaNO_3 , de ahí que la expresión del promotor *PnirA* en este medio sea alta. Podemos afirmar que la concentración de amonio utilizada no está reprimiendo por completo este promotor. Por otro lado, el nivel de inducción en BG11₀ con la solución de KNO_3 es muy alta, por encima de los niveles que habíamos obtenido con los promotores *PnrsB* y *Ptrc* (Figura 5-65).

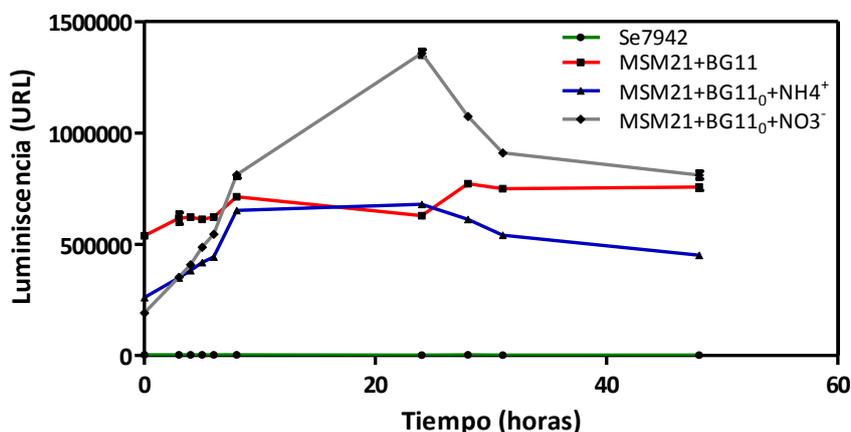


Figura 5-65. Actividad luciferasa en la cepa reportera MSM21. En verde, cepa silvestre de Se7942 crecida en BG11; en rojo, cepa MSM21 en medio BG11; en azul, cepa MSM21 en presencia de NH_4^+ para reprimir el promotor *PnirA*; en gris, cepa MSM21 en presencia de NO_3^- para inducir el promotor *PnirA*. Se muestran los datos de luminiscencia, media a diferentes tiempos. Cada valor es la media de tres réplicas independientes \pm DE.

Una vez determinado el nivel de inducción del promotor *PnirA*, medimos la producción de ácidos grasos en la cepa MSM20 (Figura 5-66). Como en el caso de MSM21, esta cepa también fue construida mediante conjugación usando el plásmido pMSM130. Solo se observó un cambio significativo en la producción de C16:0, disminuyendo ésta en el mutante MSM20, al igual que ocurría en el mutante MSM27 (Figura 5-64). Este mutante lleva la construcción *Ptrc::fabB* clonada en el sitio SN1. En ambos casos el gen *fabB* se está sobreexpresando bajo un promotor inducible fuerte. Además, si comparamos estos resultados con los obtenidos en la Figura 5-64, en la cepa MSM20 no se produce un aumento en C14:0 como había ocurrido en las cepas MSM19 y MSM27. En la cepa MSM19 el gen *fabB* estaba clonado en SN1 bajo el promotor *PnrsB*.

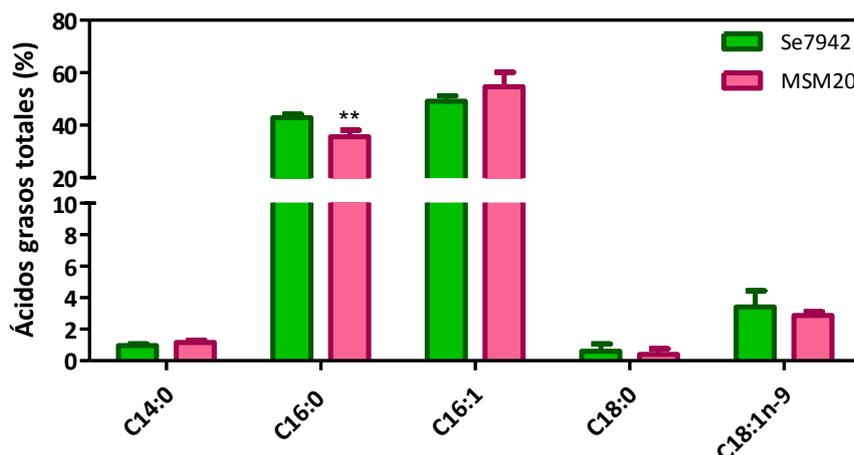


Figura 5-66. Efecto de la sobreexpresión del gen *fabB* en el contenido de ácidos grasos de Se7942. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, test t de Student no pareado.

Con los datos de todas estas cepas en las que sobreexpresamos *fabB* bajo diferentes promotores, podemos concluir que su sobreexpresión no aumenta la producción de intermediarios C18:1 como ocurría en *E. coli*.

5.2.3.4. Efecto de la sobreexpresión del gen *fabF* en el contenido de ácidos grasos de Se7942

FabF es una β -oxoacil sintasa II que cataliza la condensación del acil-ACP con otro malonil-CoA, en la ruta cíclica de síntesis de ácidos grasos (Edwards et al., 1997) (ver Introducción, apartado 2.3.1.1). Se ha demostrado que la sobreexpresión de este gen en *E. coli* es letal debido a que los altos niveles de FabF bloquean el acceso de FabB al producto generado por FabD (Subrahmanyam and Cronan, 1998).

A pesar del efecto letal en *E. coli*, decidimos sobreexpresar el gen *fabF* en Se7942. Como se ha comentado en los apartados 2.3.2.1 y 5.2.3.3, en Se7942 esta reacción solo está catalizada por FabF, ya que esta cianobacteria carece de FabB. Por ello, el efecto de bloqueo visto en *E. coli* al sobreexpresar *fabF*, no debería tener lugar en Se7942. Con dicho fin, generamos el mutante MSM29, que sobreexpresa *fabF* bajo el promotor *P_{trc}* integrado en el sitio SN1 del cromosoma de Se7942 (Tabla 4-2). En la Figura 5-67, se muestra el análisis de la composición de este mutante y de la cepa silvestre Se7942. La sobreexpresión de *fabF* en presencia de 1 mM de IPTG, produce una disminución de ácidos grasos de cadena corta (C14:0, C16:0 y C16:1) y un incremento en los ácidos grasos de cadena larga (C18:0, y C18:1n-9). Estos resultados indican

que el incremento en los niveles de FabF mejora la transformación de los ácidos grasos C14 y C16 en ácidos grasos C18. Como en el caso de la delección parcial de FabH, la sobreexpresión de FabF mejora los intermediarios de la producción de ALA, y combinando esta sobreexpresión con la de las desaturasas podríamos, quizás, aumentar la producción de ALA.

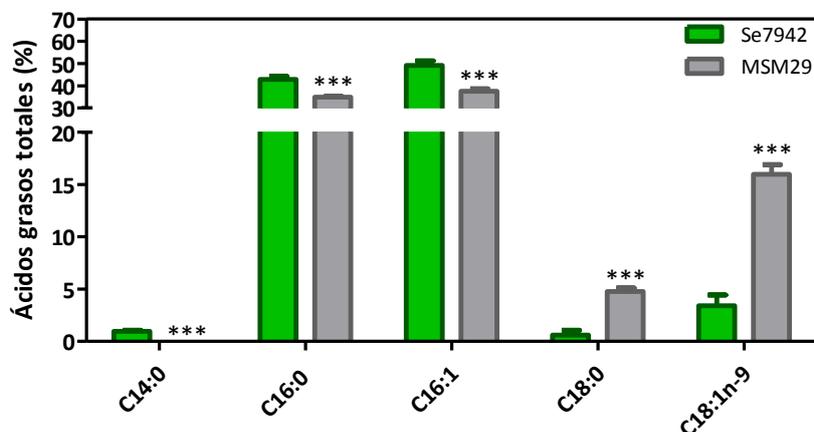


Figura 5-67. Efecto de la sobreexpresión del gen *fabF* en el contenido de ácidos grasos de Se7942. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, test t de Student no pareado.

En *Lactococcus lactis* (Morgan-Kiss and Cronan, 2008), se ha visto que la proteína FabF puede suplir la función de FabH en la síntesis de ácidos grasos. Con el fin de determinar si en Se7942 tenía lugar el mismo fenómeno, delecionamos *fabH* sobre el mutante MSM29 que sobreexpresa el gen *fabF*. Al mutante obtenido con dicha modificación lo denominamos MSM46 (Tabla 4-2). De esta manera podríamos conseguir una delección total del gen *fabH*, en el caso de que FabF sustituyera la función de FabH. Sin embargo, tal y como se observa en la Figura 5-68, la sobreexpresión de *fabF* en Se7942 no permitió la delección completa del gen *fabH*. Esto indicaría que la proteína FabF no puede sustituir la actividad llevada a cabo por FabH en Se7942.

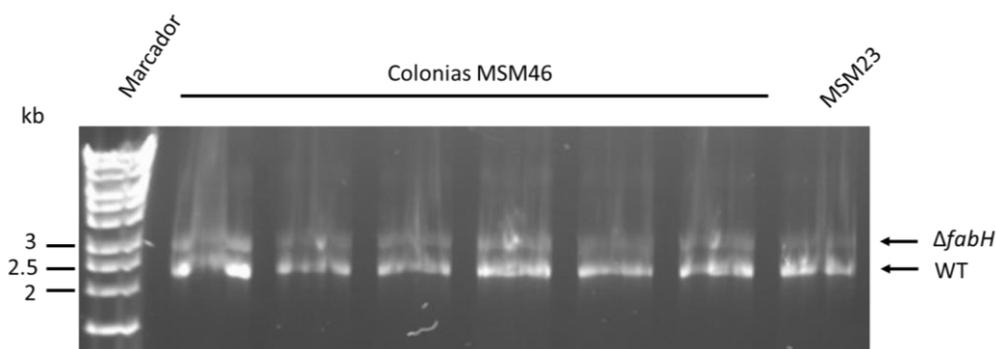


Figura 5-68. Análisis mediante de PCR de la región del gen *fabH*. Se muestran los amplicones obtenidos por PCR usando los cebadores 111 y 1112, usando ADN molde proveniente de colonias de las cepas MSM16 y MSM23. La flecha superior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la delección del gen *fabH*; la flecha inferior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la región silvestre del gen *fabH*. Se usó como marcador "HyperLadder I" (BioLabs).

5.2.3.5. Efecto de la delección del gen *fadD* en el contenido de ácidos grasos de Se7942

La ruta de degradación de ácidos grasos de cadena corta y larga es la β -oxidación. Las enzimas que participan en esta ruta en *E. coli* están bien caracterizadas (ver Introducción, apartado 2.3.1.2), pero en cianobacterias no existen datos bibliográficos ni anotaciones genómicas de dichos genes (ver Introducción, apartado 2.3.2.2), salvo en el caso de *fabD*. Esta enzima cataliza el primer paso de la ruta de degradación de ácidos grasos. En *E. coli*, la delección del gen *fabD*, y con ello el bloqueo de la β -oxidación de ácidos grasos, provoca la acumulación de ácidos grasos libres en el citosol (Pech-Canul et al., 2011). En Se7942, dicha eliminación da lugar a un cambio en el perfil de ácidos grasos de membrana, particularmente un incremento en su grado de saturación (Ruffing and Jones, 2012).

Con esta información, delecionamos el gen *fabD* en Se7942, obteniendo la cepa MSM42 (Tabla 4-2). Para su construcción utilizamos el sistema diseñado por (Matsuoka et al., 2001), mediante el cual sustituimos el gen *fabH* por un gen de resistencia a kanamicina y una copia silvestre del gen *rps12* de Ss6803 (Figura 5-69, A). Después de tres pases de estos transformantes en placas con el antibiótico correspondiente (Neo), analizamos los mutantes mediante PCR usando los oligonucleótidos 109 y 110, observándose una banda en el gel de electroforesis. La banda que aparece en las siete colonias analizadas presentaba un tamaño de $\sim 2,3$ kb, el esperado para la eliminación en el cromosoma del gen *fadD*. Por otro lado, en ninguna de las colonias apareció una banda de mayor tamaño ($\sim 2,8$ kb) correspondiente a las copias silvestres del cromosoma (Figura 5-69, B). En este caso obtuvimos un mutante puro, consiguiendo eliminar todas las copias del gen *fadD* de Se7942.

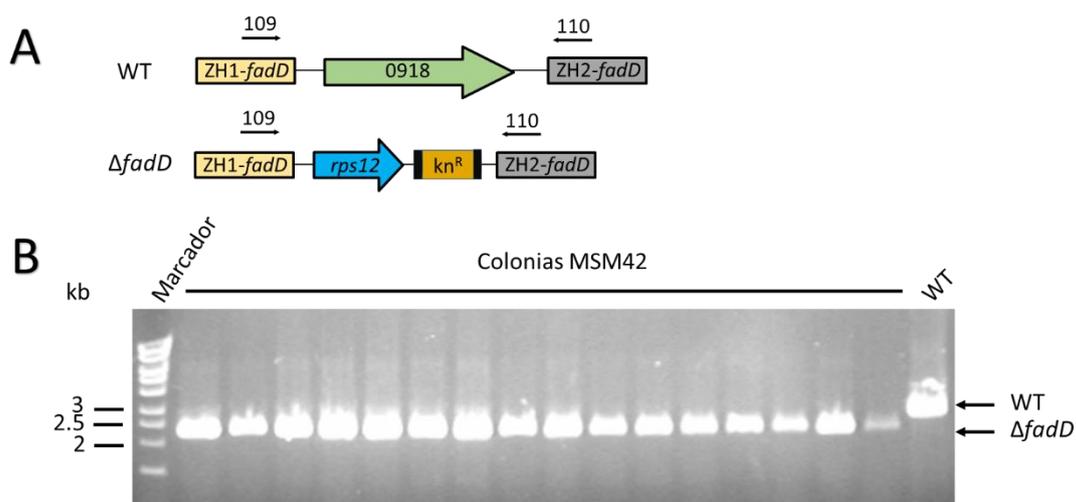


Figura 5-69. Construcción del mutante MSM42 deficiente en el gen *fadD*. (A) Mapa de la región *fadD* (0918) del cromosoma de este mutante así como de la cepa silvestre (WT). Las flechas y los números por encima del mapa representan los oligonucleótidos usados en el análisis por PCR. (B) Gel de electroforesis de la PCR para analizar la región *fadD* del genoma de 16 colonias de MSM42. La flecha superior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la región silvestre del gen *fadD*; la flecha inferior de la derecha indica la posición de la banda correspondiente a las copias del cromosoma con la delección del gen *fadD*. Se usó como marcador "HyperLadder I" (BioLabs).

Una vez analizado el genotipo del mutante MSM42, el siguiente paso fue el análisis de su composición de ácidos grasos. Como en los casos anteriores, se comparó con la cepa silvestre de Se7942 y los datos se muestran en la **Figura 5-70**. En la cepa MSM42 se produjo un aumento significativo de todos los ácidos grasos C18 (C18:0 y C18:1n-9), mientras que la cantidad de C16:1 disminuyó. Esta modificación podría ser beneficiosa para la producción de ALA, ya que los ácidos grasos C18 son utilizados como sustrato en la producción de este ácido poliinsaturado.

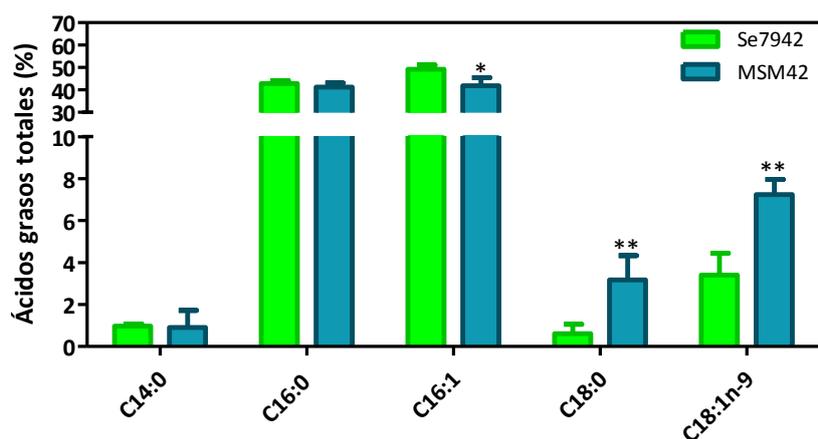


Figura 5-70. Efecto de la delección de *fadD* en la composición de ácidos grasos de Se7942. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, test t de Student no pareado.

5.2.4. Análisis de dobles mutantes de Se7942 productores de ácido α -linolénico

Con los resultados obtenidos en el apartado 5.2.3, decidimos construir tres dobles mutantes productores de ALA. Para ello, usamos la cepa MSM26 (*P_{trc}::desA-desB* en el sitio SN1, ver **Tabla 4-2**) en la que introdujimos las modificaciones capaces de aumentar los ácidos grasos C18. En total, elegimos tres cambios: la delección parcial del gen *fabH* (ver apartado 5.2.3.2), la sobreexpresión del gen *fabF* (ver apartado 5.2.3.4) y la eliminación del gen *fadD* (ver apartado 5.2.3.5). De esta manera, obtuvimos tres mutantes, MSM36, MSM39 y MSM43, respectivamente (**Tabla 4-2**). En el mutante MSM39, la construcción para sobreexpresar el gen *fabF* se introdujo en el sitio SN2. Previamente, comparamos el análisis de ácidos grasos del mutante MSM32 (*P_{trc}::fabF* en el sitio SN2) frente al mutante MSM29 (*P_{trc}::fabF* en el sitio SN1) (ver **Figura 5-67**). No se encontraron diferencias significativas en la composición de ácidos grasos entre estas dos cepas (datos no mostrados). Por tanto, podemos concluir que la localización en el cromosoma de Se7942 (SN1 o SN2) de la construcción *P_{trc}::fabF* no afecta a su sobreexpresión.

Una vez construidos los mutantes, analizamos su composición total de ácidos grasos, y especialmente la producción de ALA. Se crecieron todas las cepas a 30 °C y a una intensidad de luz de 60 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. La inducción se llevó a cabo a 22 °C mediante la adición de 1 mM de IPTG, recogiendo los cultivos 24 h más tarde. Se extrajeron los ácidos grasos y se analizó su composición, resultados que se muestran en la **Tabla 5-7**. En la **Figura 5-71**, se muestra un esquema de las cepas analizadas en este apartado, para facilitar la interpretación de los resultados.

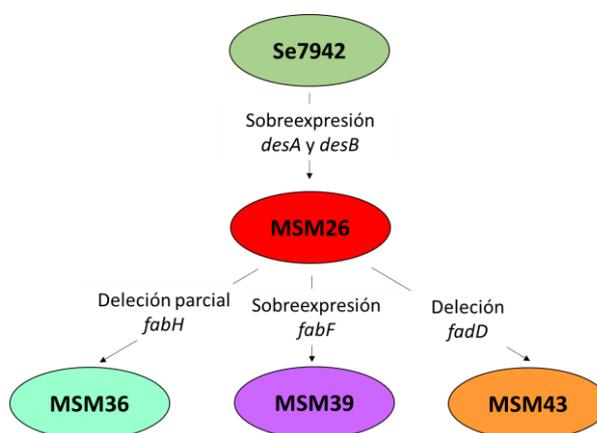


Figura 5-71. Esquema de los dobles mutantes construidos. El código de colores utilizado para representar a cada una de las cepas es el mismo que el de la **Figura 5-72**.

5. Resultados

Los tres dobles mutantes fueron capaces de producir ALA (C18:3n-3): 6,68 % en MSM36, 17,98 % en MSM39 y 11,80 % en MSM43. En las tres cepas, el nivel de C16:1 disminuyó muy significativamente. Por otro lado, no se detectaron C18:1n-9 y C18:2. Este perfil de ácidos grasos ya se observó previamente, cuando solo se sobreexpresaban las desaturasas en la cepa MSM26 (ver **Tabla 5-5**). Sorprende el hecho de que, en todos los mutantes, los niveles de C16:0 estén por encima de los que presenta la cepa silvestre de Se7942. Por otro lado, en el mutante MSM43, aumenta el nivel de C14:0, mientras que en MSM36 disminuye y en MSM39 no está presente este ácido graso. Por último, si atendemos a la concentración de C18:0 presente en los mutantes, observamos que el mayor productor de ALA, MSM39, carece de este ácido graso. Sin embargo, en los otros dos mutantes el nivel está por encima de los niveles de la cepa silvestre.

Tabla 5-7. Composición de ácidos grasos de la cepa silvestre de Se7942 y de las cepas modificadas genéticamente para mejorar la producción de ALA (MSM36, MSM39 y MSM43).

| Cepa | C14:0 | C16:0 | C16:1 | C18:0 | C18:1n-9 | C18:2 | C18:3n-3 |
|---------------|--------|---------|---------|--------|----------|--------|----------|
| Se7942 | 1,01 ± | 42,92 ± | 49,37 ± | 0,61 ± | 2,98 ± | 0,00 ± | 0,00 ± |
| | 0,09 | 1,64 | 2,40 | 0,55 | 0,74 | 0,00 | 0,00 |
| MSM36 | 0,32 ± | 72,70 ± | 16,28 ± | 1,70 ± | 0,00 ± | 0,00 ± | 6,68 ± |
| | 0,63 | 3,21 | 1,74 | 1,97 | 0,00 | 0,00 | 0,39 |
| MSM39 | 0,00 ± | 65,53 ± | 13,56 ± | 0,00 ± | 0,00 ± | 0,00 ± | 17,98 ± |
| | 0,00 | 2,85 | 2,45 | 0,00 | 0,00 | 0,00 | 2,51 |
| MSM43 | 1,25 ± | 64,87 ± | 13,91 ± | 6,34 ± | 0,00 ± | 0,00 ± | 11,80 ± |
| | 0,53 | 3,92 | 2,68 | 0,34 | 0,00 | 0,00 | 0,89 |

Estos datos representan la media de los valores de un mínimo de tres réplicas de experimentos independientes ± la desviación estándar (DE).

Para determinar si existía una mejora en la producción de ALA en estos tres mutantes respecto a la cepa MSM26 utilizada como chasis para la construcción de los mismos, comparamos su producción en la **Figura 5-72**. Como se observa, tanto MSM39 como MSM43 mejoran significativamente la producción de ALA, respecto al control MSM26. En el caso de MSM43, la producción aumentó un 1,5 %, mientras que en MSM39 el aumento fue de un 2,4 %. En el mutante MSM36 no mejora dicha producción, no observándose diferencias significativas con la producción respecto a la cepa control de este experimento. Por tanto, estos resultados nos indicarían que tanto la sobreexpresión del gen *fabF* (mutante MSM39) como la delección del gen *fadD* (mutante MSM43) mejoran la síntesis de ALA en Se7942, siendo la mejora más acusada en el mutante MSM39. Sin embargo, la delección parcial de *fabH* (mutante MSM36), no está afectando a dicha producción.

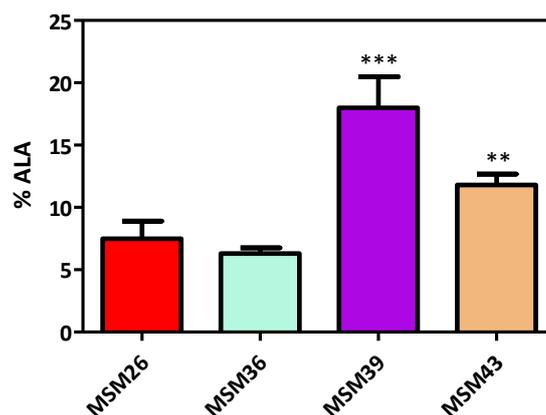


Figura 5-72. Producción de ALA en el mutante simple MSM26 y en los dobles mutantes, MSM36, MSM39 y MSM43. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa MSM26.

Se llevó a cabo el análisis de otros dos dobles mutantes, MSM38 y MSMS40, que contienen junto a la sobreexpresión de las desaturasas de *Ss7002*, la sobreexpresión del gen *fabB* bajo el promotor *PnrsB* y la sobreexpresión del gen *fabD* bajo el promotor *Ptrc* en el sitio SN2, respectivamente (**Tabla 4-2**). A pesar de que la sobreexpresión de los genes en ausencia de las desaturasas no tuviera ningún efecto significativo en la composición de ácidos grasos, quisimos ver si junto a ellas podría verse algún tipo de cambio. Antes de la construcción de estos mutantes, analizamos la composición de dos mutantes que presentaban dichas construcciones en el sitio SN2 en ausencia de las desaturasas, MSM31 y MSM33 (**Tabla 4-2**). Comparamos su composición de ácidos grasos frente a los mutantes que llevaban dichas construcciones en el sitio SN1, MSM19 (**Figura 5-64**) y MSM24 (**Figura 5-58**) frente a MSM31, no observándose cambios significativos (datos no mostrados). Por tanto, como ocurría en el caso de la sobreexpresión de *fabF* la posición que ocupan los genes en el cromosoma de *Se7942* (SN1 o SN2) no afecta a la sobreexpresión.

Una vez comprobado que la posición de las construcciones en un sitio neutral u otro del cromosoma no afectaba, decidimos analizar la composición de los dobles mutantes, MSM38 y MSM40, y cuyos resultados se muestran en la **Tabla 5-8**. En la **Figura 5-73**, se muestra un esquema de las cepas analizadas en este apartado, para facilitar la interpretación de los resultados.

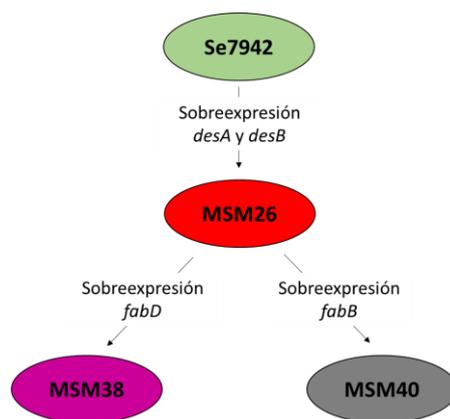


Figura 5-73. Esquema de los dobles mutantes construidos. El código de colores utilizado para representar a cada una de las cepas es el mismo que el de la **Figura 5-74**.

Como puede observarse en la **Tabla 5-8**, tanto en el mutante MSM38 como en MSM40, aumenta la cantidad de C18:0 en comparación con la cepa silvestre, algo que ya habíamos observado en los mutantes MSM36 (deleción parcial de *fabH*) y MSM43 (deleción de *fadD*) (ver **Tabla 5-7**). Además, la proporción de ácido graso C14:0 también aumentó en MSM38 y MSM40, hecho que ya se había dado en MSM43 (ver **Tabla 5-7**). En ambas cepas, la cantidad de C16:0 aumenta, mientras que la de C16:1 disminuye. Este fenómeno ha tenido lugar en todos los mutantes en los que hemos sobreexpresado las desaturasas bajo el promotor *P_{trc}*. Por último, MSM38 y MSM40 producían ALA, a niveles por debajo de MSM23 (7,49 %).

Tabla 5-8. Composición de ácidos grasos de la cepa silvestre de Se7942 y de las cepas modificadas genéticamente para mejorar la producción de ALA (MSM38 y MSM40).

| Cepa | C14:0 | C16:0 | C16:1 | C18:0 | C18:1n-9 | C18:2 | C18:3n-3 |
|--------|----------------|-----------------|-----------------|----------------|----------------|----------------|----------------|
| Se7942 | 1,01 ± 0,09 | 42,92 ± 1,64 | 49,37 ± 2,40 | 0,61 ± 0,55 | 2,98 ± 0,74 | 0,00 ± 0,00 | 0,00 ± 0,00 |
| MSM38 | 1,23 ± 0,93 | 70,09 ± 2,32 | 15,17 ± 2,45 | 3,41 ± 1,09 | 0,00 ± 0,00 | 0,00 ± 0,00 | 7,03 ± 1,72 |
| MSM40 | 2,05 ± 0,37 | 59,04 ± 4,66 | 29,21 ± 1,66 | 3,19 ± 1,59 | 0,00 ± 0,00 | 0,00 ± 0,00 | 2,15 ± 1,11 |

Estos datos representan la media de los valores de un mínimo de tres réplicas de experimentos independientes ± la desviación estándar (DE).

En la **Figura 5-74**, se compara la producción de ALA de los mutantes dobles, MSM38 y MSM40, frente al mutante simple, MSM26, a partir del cual fueron construidos. Como se observa, la producción de ALA no presentó diferencias significativas entre MSM38 y MSM26. Por tanto, la modificación añadida a MSM38 (*P_{trc}::fadD* en el sitio SN2) no tuvo ningún efecto sobre la producción de ALA. Por otro lado, en MSM40 la producción de ALA disminuyó

considerablemente, presentando diferencias significativas frente a MSM26. La sobreexpresión de *fabB* en el mutante MSM40 tiene un efecto negativo notable sobre la producción de ALA, reduciéndose un 63 %. Este efecto podría estar relacionado con lo que especulaban (Kuo and Khosla, 2014) acerca de que FabB podría tener un papel adicional en los organismos de crecimientos rápido, no asociado a la síntesis de ácidos grasos.

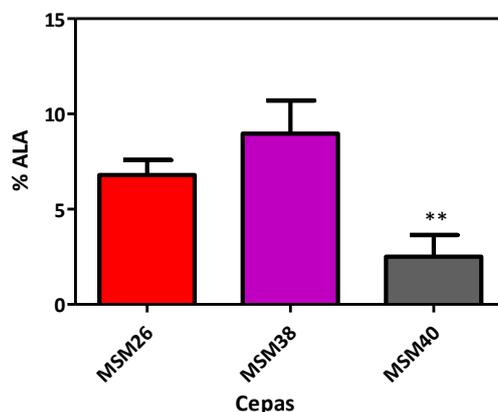


Figura 5-74. Producción de ALA en el mutante simple MSM26 y en los dobles mutantes, MSM38 y MSM40. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa MSM26.

5.2.5. Análisis de un triple mutante de Se7942 productor de ácido α -linolénico

Una vez demostrado que la sobreexpresión del gen *fabF* y la delección del gen *fadD* mejoran la producción de ALA, decidimos construir un triple mutante, MSM45 (Tabla 4-2). Este mutante sobreproduciría FabF y las desaturasas DesA y DesB, mientras que no produciría FadD.

Como en el caso anterior, la cepa se cultivó a 30 °C y a una intensidad de luz de 60 $\mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. La inducción se llevó a cabo a 22 °C mediante la adición de 1 mM de IPTG, recogiendo los cultivos 24 h más tarde. Se extrajeron los ácidos grasos y se analizó su composición. En la Tabla 5-9, se muestra la composición de ácidos grasos de MSM45, así como de la cepa silvestre de Se7942. MSM45 produjo un 22,28 % de ALA, nivel más elevado que MSM39, el mutante doble que producía mayor cantidad de ALA (17,98 %) (Tabla 5-7). Cuando comparamos la composición de ácidos grasos de MSM45 (triple mutante) con MSM26 y los mutantes MSM39 y MSM43 (Figura 5-73), observamos variaciones en la composición. Se producen una disminución de C16:1, así como un aumento en C16:0, aunque menos marcado en este caso. Desaparecen también los ácidos grasos C18:1n-9 y C18:2. En el caso de C14:0, este ácido graso desaparece, tal y como ocurría en la cepa MSM39. Curiosamente, el nivel de C18:0

5. Resultados

aumenta con respecto a la cepa silvestre, al igual que ocurría en la cepa MSM43. Podríamos atribuir esta subida al efecto de la delección del gen *fadD*.

Tabla 5-9. Composición de ácidos grasos de la cepa silvestre de Se7942 y del triple mutante MSM45.

| Cepa | C14:0 | C16:0 | C16:1 | C18:0 | C18:1n-9 | C18:2 | C18:3n-3 |
|--------|--------|---------|---------|---------|----------|--------|----------|
| Se7942 | 1,01 ± | 42,92 ± | 49,37 ± | 0,61 ± | 2,98 ± | 0,00 ± | 0,00 ± |
| | 0,09 | 1,64 | 2,40 | 0,55 | 0,74 | 0,00 | 0,00 |
| MSM45 | 0,00 ± | 54,46 ± | 10,30 ± | 11,59 ± | 0,00 ± | 0,00 ± | 22,28 ± |
| | 0,00 | 1,01 | 0,79 | 2,19 | 0,00 | 0,00 | 1,41 |

Estos datos representan la media de los valores de un mínimo de tres réplicas de experimentos independientes ± la desviación estándar (DE).

Finalmente, en la **Figura 5-75** comparamos la producción de ALA entre el triple mutante (MSM45), los dobles mutantes (MSM39 y MSM43) y la cepa Ss7002. Solo en el caso de la cepa Ss7002 las condiciones de inducción han cambiado, siendo de AI, ya que en ellas se produce una mayor cantidad de ALA (apartado 5.2.2). Como podemos observar en esta gráfica, la cepa MSM45 es la mejor productora de ALA, superando los niveles de producción de los dobles mutantes y de la cepa Ss7002.

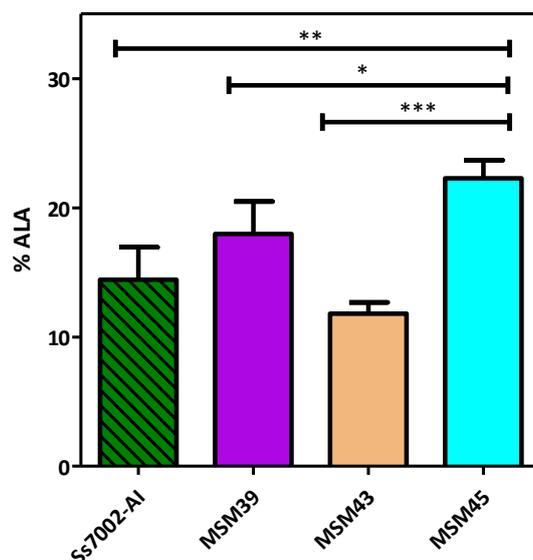


Figura 5-75. Producción de ALA en diferentes mutantes de Se7942. Se compara la producción de ALA en Ss7002 en condiciones de AI, en los mutantes dobles (MSM39 y MSM43) y en el mutante triple (MSM45). El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa Ss7002 en condiciones de AI.

Podemos concluir que la combinación de la sobreexpresión del gen *fabF* y la delección del gen *fadD* con la sobreexpresión de las desaturasas DesA y DesB de Ss7002, permiten la producción de ALA en Se7942, incluso a niveles superiores a los de Ss7002, en condiciones que

favorecen la expresión de dichos genes. El efecto conjunto de estas dos modificaciones, sobreexpresión de *fabF* y la delección de *fadD*, ya se había visto en el mutante MSM44. En la **Figura 5-76**, se compara la composición de ácidos grasos entre dicho mutante, frente a los mutantes simples MSM29 (sobreexpresión de *fabF*) y MSM42 (delección de *fadD*) (**Tabla 4-2**). Como podemos observar en dicha imagen, la composición de los ácidos grasos C16:1, C18:0 y C18:1n-9 presentaba diferencias significativas al comparar la composición de los mutantes simples frente a MSM44. En el caso de C16:1, la concentración es menor en MSM44 que en las otras dos cepas, mientras que la concentración de C18:0 y C18:1n-9 fue mayor. Estos cambios pueden deberse a la suma de efectos de ambas mutaciones.

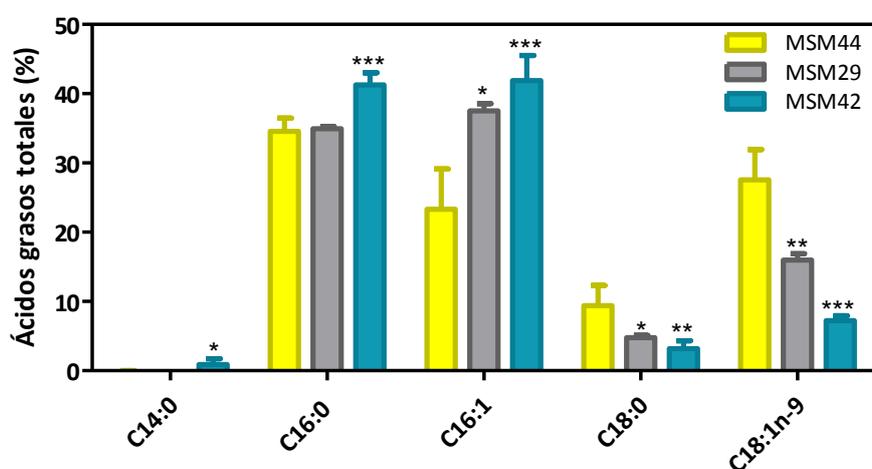


Figura 5-76. Efecto de la sobreexpresión de *fabF* y la delección de *fadD* en el contenido de ácidos grasos de Se7942. El código de colores de las cepas analizadas se muestra en la leyenda de la figura. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnett de comparación múltiple usando como control la cepa MSM44.

Por último, se construyó otro triple mutante, MSM41, que sobreexpresa *fabF*, junto a la delección parcial del gen *fabH* y la sobreexpresión de ambas desaturasas (**Tabla 4-2**). Aunque el efecto de la delección no había producido un aumento en la producción de ALA en el mutante doble MSM36 (**Figura 5-72**), decidimos construir este triple mutante para ver si el efecto conjunto de la sobreexpresión de *fabF* y la delección parcial de *fabH* podían mejorar la producción de ALA. En la **Figura 5-77**, se compara la producción de ALA de la cepa MSM41 con los mutantes dobles MSM39 (sobreexpresión de desaturasas y *fabF*) y MSM36 (sobreexpresión de desaturasas y delección parcial de *fabH*) (**Tabla 4-2**). La producción en MSM41 fue de $10,44 \pm 0,15$ % ALA, mientras que en los otros dos mutantes fue de $17,98 \pm 2,51$ y $6,68 \pm 0,39$, respectivamente. El mutante triple no superó la producción del mutante doble que más producía, MSM39, sin embargo, sí lo hizo respecto al mutante, MSM36. Por tanto, puede

afirmarse, que la combinación de la sobreexpresión de *fabF* con la delección de *fabH* no mejoró la producción de ALA respecto a MSM45 ($22,28 \pm 1,41$).

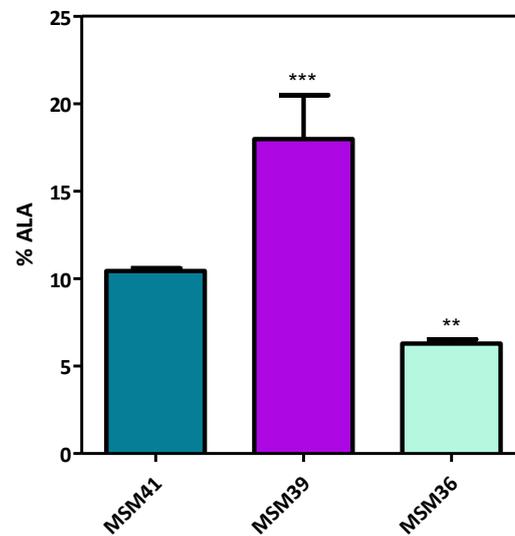


Figura 5-77. Producción de ALA en mutantes de Se7942. La barra azul oscuro representa a la cepa MSM41, la barra morada a la cepa MSM39 y la barra azul claro a la cepa MSM36. Los datos corresponden a la media de un mínimo de tres réplicas biológicas independientes y se representan como la media + DE. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, ANOVA seguida de un test de Dunnet de comparación múltiple usando como control la cepa MSM41.

5.3. Modelo metabólico de Se7942

5.3.1. Construcción y características del modelo metabólico *iMS705* de Se7942

La reconstrucción metabólica de Se7942, denominada *iMS705*, se realizó en base a la anotación de alta calidad de su genoma del GenBank cuyo número de acceso es NC_007604. El genoma de Se7942 consta de 2.695.903 pb y contiene un total de 2.664 genes (ORF). A estos hay que añadir los 50 genes del plásmido pANL cuyo tamaño es de 46.366 pb. Dado que muchas proteínas de Se7942 están anotadas como proteínas hipotéticas, durante el proceso de construcción del modelo se realizó una revisión exhaustiva de la anotación, teniendo en cuenta tanto la literatura científica como bases de datos metabólicas y bioquímicas (ej. KEGG, BRENDA, BIGG, entre otras).

El primer paso de la construcción consistió en la identificación mediante BLASTp recíproco de genes ortólogos entre Se7942 y Ss6803. Se realizó una segunda identificación de genes ortólogos entre Se7942 y *E. coli* y entre Se7942 y *P. putida*. De esta manera, se consiguieron identificar genes de Se7942 con ortólogos en Ss6803 (434), *E. coli* (9) y *P. putida* (21). Esta información fue utilizada para construir una primera versión del modelo de Se7942 mediante asignación automática de reacciones bioquímicas a genes de Se7942, en función de las asignadas a sus ortólogos en Ss6803, *E. coli* o *P. putida*, tal como constan en sus respectivos modelos metabólicos *iJN678*, *iJO1366* y *iJN746*. La primera versión del modelo de Se7942 con los genes ortólogos de Ss6803 contenía un total de 700 reacciones, 730 metabolitos y 455 genes. A esta versión, se le añadieron 21 reacciones, 8 metabolitos y 30 genes contenidos en los modelos de *E. coli* MG1655 (*iJO1366*) y de *P. putida* KT2440 (*iJN1411*).

En la **Figura 5-78**, se muestra un análisis comparativo de las reacciones y metabolitos compartidos entre el modelo de Ss6803 (*iJN746*) y los modelos de *E. coli* MG1655 (*iJO1366*) y de *P. putida* KT2440 (*iJN1411*). De las 863 reacciones presentes en Ss6803, 509 son comunes a *E. coli* y 445 a *P. putida* (**Figura 5-78, A**). Por otro lado, de los 797 metabolitos presentes en Ss6803, 596 son comunes a *E. coli* y 448 a *P. putida* (**Figura 5-78, B**). Esto se explica porque la mayoría de las reacciones y metabolitos que hemos añadido al modelo provienen de Ss6803, organismo con el que se generó el primer “draft”.



Figura 5-78. Diagramas de Venn representando el número de reacciones y metabolitos comunes entre el modelo de Ss6803 (*iJN678*) y los modelos de *E. coli* MG1655 (*iJO1366*) y de *P. putida* KT2440 (*iJN746*). (A) Diagramas de Venn de reacciones comunes. (B) Reacciones de Venn de metabolitos comunes. En verde, se representa el modelo *iJN678*; en rosa, se representa el modelo *iJO1366*; en azul, se representa el modelo *iJN746*. Diagramas realizados con InteractiVenn (Heberle et al., 2015).

De esta versión inicial del modelo se revisaron manualmente un total de 395 reacciones, de las cuales 41 fueron descartadas tras observar una insuficiente identidad de secuencia entre los genes ortólogos identificados automáticamente (ver criterio en apartado 4.11.2 de Material y Métodos) y tras llevar a cabo una revisión en la literatura para estas reacciones (p. ej., se eliminaron reacciones de síntesis de tocoferol y cianoficina inicialmente incluidas en el “draft”). El modelo final (*iMS705*) contiene 669 reacciones de las 721 del modelo inicial.

A continuación, se obtuvo una lista de los genes no comunes entre Se7942 y Ss6803 que se revisó de forma manual. De esta manera, se añadieron al modelo aquellas reacciones específicas de Se7942, un total de 54 reacciones, 68 metabolitos y 74 genes. Esta información fue añadida usando la información presente en la base de datos BIGG Models (King et al., 2016), que recopila la anotación de todos los modelos publicados hasta el momento. Por otro lado, analizando la lista de genes comunes entre estas Se7942 y Ss6803 se identificaron aquellas reacciones que no se habían anotado previamente en el modelo de Ss6803. Este análisis aportó al modelo 57 reacciones, 20 metabolitos y 183 genes, información descrita en el modelo *iJN678*.

A este número hay que añadir una reacción más, la reacción de biomasa cuyos metabolitos y coeficientes estequiométricos se describen en la **Tabla 4-7** de Materiales y Métodos.

En resumen, el modelo final de Se7942 (*iMS705*) contiene 705 genes, cuyos productos génicos participan en 843 reacciones involucrando 826 metabolitos. De estas 843 reacciones, 731 fueron introducidas a partir del draft inicial, de las cuales 3 fueron reacciones espontáneas incluidas basadas en evidencias experimentales. El modelo se completó con un total de 33 reacciones huérfanas incluidas tras una revisión bibliográfica que apoyaba la presencia de éstas en Se7942, pese a que actualmente se desconozca el gen(es) responsables. Así, por ejemplo, se introdujo la reacción llevada a cabo por la enzima gamma caroteno biciclasa, involucrada en el metabolismo de carotenoides (Schafer et al., 2006). Finalmente, se añadieron 49 reacciones de intercambio de metabolitos. Todas las reacciones incluidas en el modelo se asignaron a alguno de los cuatro compartimentos celulares en los que se divide la red metabólica de Se7942, es decir, en espacio extracelular, periplasma, citoplasma y membranas tilacoidales.

Las reacciones se categorizaron en 13 subsistemas, tal y como se muestra en la **Figura 5-79**. Cinco subsistemas incluían un mayor número de reacciones: transportadores, metabolismo de carbohidratos, metabolismo de aminoácidos, metabolismo de lípidos y metabolismo de vitaminas. Cabe destacar que en *iMS705* un alto porcentaje de las reacciones, 13,76 %, están involucradas en el metabolismo de lípidos, lo que refleja el gran potencial metabólico de Se7942 para la síntesis de dichos compuestos. Por otro lado, el metabolismo de aminoácidos también tiene un papel importante, ya que el 13,76 % de las reacciones pertenecen a esta categoría. Algunas de las reacciones involucradas en la síntesis de leucina, arginina y lisina, ya habían sido modeladas anteriormente en *Geobacter metallireducens* y se tomaron del modelo metabólico *iAF987* de este organismo (Feist et al., 2014). Aunque solo representan un 2,02 % en el modelo de Se7942, se incluyeron las reacciones para las síntesis de grupos prostéticos y cofactores no descritas en el modelo de Ss6803 (*iJN678*), pero sí en el de *E. coli* (*iJO1366*).

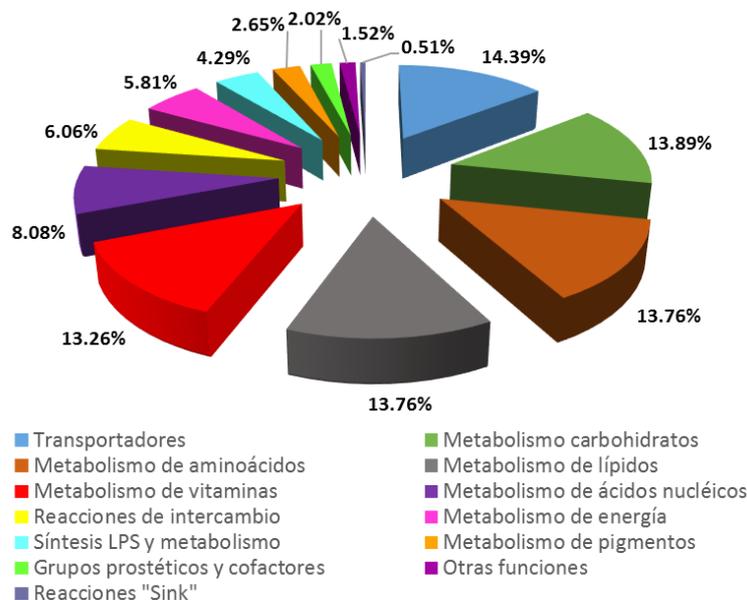


Figura 5-79. Clasificación funcional de las reacciones metabólicas en el modelo de Se7942. Las 843 reacciones del modelo fueron categorizadas en 13 subsistemas, que se indican en la leyenda de la figura.

5.3.2. Comparación de *iMS705* con otros modelos metabólicos de Se7942

Al mismo tiempo que se construía el modelo metabólico *iMS705* de Se7942, fue publicado otro modelo metabólico de Se7942, *iJB785* (Broddrick et al., 2016). Por ello decidimos comparar los modelos existentes hasta el momento, para identificar genes únicos presentes en cada modelo, ver qué función realiza la proteína para la que codifican, y añadirlos en el modelo mejorado de la cianobacteria Se7942.

En primer lugar, analizamos las reacciones y metabolitos de los modelos de Se7942, *iSyf715* (Triana et al., 2014) e *iJB785* (Broddrick et al., 2016), así como del desarrollado en la presente tesis, *iMS705*. Sin embargo, la anotación de las reacciones y metabolitos no era la misma en todos los modelos. Esto se debe a que la nomenclatura utilizada en el modelo *iSyf715* no es la estándar y que utilizan todos los modelos anotados en la base de datos BIGG Models. Como consecuencia, solo pudimos analizar los genes comunes y no comunes presentes en todos ellos, tal y como se muestra en la **Figura 5-80**. Es importante destacar que el número de genes del modelo *iSyf715* no es 715 como (Triana et al., 2014) describen en su artículo, sino 714. Si analizamos el número de genes comunes entre los tres modelos, la cifra asciende a 565 genes, siendo un 80 %, un 79 % y un 72 % de los genes que componen los modelos *iMS705*, *iSyf715* e *iJB785*, respectivamente. El porcentaje de genes comunes es alto en todos los casos, lo que indica que los modelos son muy parecidos. Entre el modelo *iMS705* e *iJB785* existen 53 genes

comunes, no presentes en *iSyf715*, siendo 60 genes comunes entre *iSyf715* e *iJB785* y 34 genes comunes entre *iMS705*, *iSyf715* e *iJB785*. Hay que destacar también que existen genes que están presentes en un único modelo, siendo 53 en el caso de *iMS705*, 55 en *iSyf715* y 107 en *iJB785*.

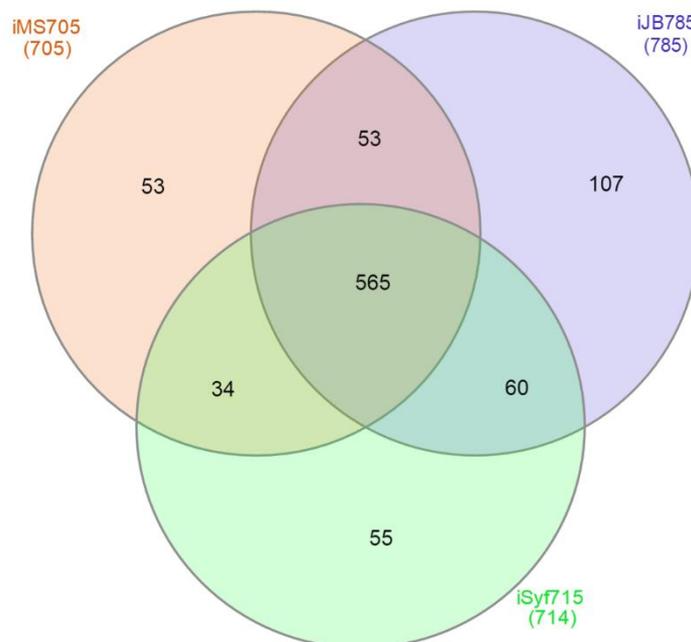


Figura 5-80. Diagrama de Venn representando el número de genes comunes entre los modelos de Se7942 existentes hasta el momento, *iMS705*, *iJB785* e *iSyf715*. Diagramas realizados con InteractiVenn.

Dado que *iMS705* e *iJB785*, son los modelos cuya nomenclatura de reacciones y metabolitos sigue la descrita en la base de datos BIGG Models, hicimos un análisis comparativo de los genes presentes en estos dos modelos. De esta manera, observamos que existen 618 genes comunes entre ambos modelos, mientras que el modelo *iMS705* tiene 87 genes que no están presentes en *iJB785* y éste posee 167 genes no presentes en *iMS705*. Dado que en el modelo *iJB785*, (Broddrick et al., 2016) han modelado los fotones de luz como si fuera un metabolito, decidimos construir sobre él un cuarto modelo metabólico incluyendo parte de los 87 genes no comunes entre *iMS705* e *iJB785*.

El nuevo modelo (*iMS835*) incluye 50 de las 87 reacciones no comunes entre *iMS705* e *iJB785* y contiene, 835 genes, 806 metabolitos y 895 reacciones. Además, corregimos varios errores presentes en el modelo *iJB785*, como la síntesis de glucógeno, la degradación de ácidos grasos y el intercambio de leucina, que estaban mal anotados. Por otro lado, en colaboración con el Dr. Juan Nogales, del CIB, creamos una ecuación de biomasa “core” eliminando aquellos metabolitos que no fueran producto de una reacción esencial. De esta manera, se obtiene una biomasa más simplificada, y por ende, más cercana a la realidad.

A continuación, comparamos la capacidad de estos modelos para predecir genes esenciales. Se utilizaron las restricciones mínimas estandarizadas en los cuatro modelos, previamente usadas por (Broddrick et al., 2016), permitiendo una comparación directa de las tres redes metabólicas. Estas condiciones estaban marcadas por tres valores de “uptake”, siendo la tasa de “uptake” de CO₂ de -1,99 mmol · gDW⁻¹ · h⁻¹, de HCO₃ de -1,99 mmol · gDW⁻¹ · h⁻¹ y de fotones de -14,3 mmol · gDW⁻¹ · h⁻¹. Los resultados obtenidos *in silico* se compararon con los datos obtenidos *in vivo* descritos por (Rubin et al., 2015).

Los resultados de dicha comparación se muestran en la **Tabla 5-10**. El modelo *iJB785* es el que mejor predice los genes esenciales (88 % de aciertos). Le sigue nuestro modelo *iMS835* (84 % de aciertos), seguido a su vez por *iMS705* (64 %) e *iSyf715* (33 %). Si atendemos a la capacidad predictiva de genes no esenciales, en primer lugar tenemos el modelo *iSyf715* (97 % de aciertos), seguido de nuestro modelo *iMS705* (95 %), *iMS835* (90 %) e *iJB785* (85 %). Si analizamos el total de genes correctamente predichos por cada uno de los modelos, en primer lugar tenemos a *iMS705* (79 % de aciertos), seguido de los modelos *iJB785* y *iMS835* (78 % y 75 %, respectivamente), y terminando con *iSyf715* (55 %). Los modelos *iJB785*, *iMS835* e *iSyf715*, son capaces de predecir genes beneficiosos, siendo 5, 5 y 1 los genes beneficiosos predichos, respectivamente. En la categoría de genes beneficiosos se engloban aquellos genes que provocan defectos en el crecimiento cuando son mutados, es decir, aquellos que disminuyen la tasa de crecimiento cuando son mutados. Nuestro modelo *iMS705*, no predijo ningún gen beneficioso.

Por otro lado, comparamos el número total de genes presentes en los modelos *iMS705* e *iMS835*, según la concordancia o no de la predicción de esencialidad *in vivo* frente a *in silico* (**Figura 5-81**). Así, se consideraron concordantes tanto aquellos genes considerados esenciales *in vivo* e *in silico*, como los no esenciales en ambas aproximaciones. Se representaron en verde en la **Figura 5-81**. Sin embargo, para los genes cuya predicción de esencialidad *in silico* no coincidía con los resultados *in vivo* (Rubin et al., 2015), se consideró una no concordancia (representada en rojo en **Figura 5-81**). El modelo *iMS835* es el modelo que más genes predijo correctamente (673, frente a los 587 de *iJB785*, los 474 de *iMS705* y los 376 de *iSyf715*). En cuanto al número de genes predichos incorrectamente, en primer lugar está el modelo *iSyf715* con 306, seguido de *iMS705* con 196 y de *iJB785* con 165, y terminando con *iMS835* con 125. Por último, hay que destacar que el número de genes no analizados, pues no se contaba con información *in vivo* sobre su esencialidad (en azul en la **Figura 5-81**), fue similar en todos los modelos: 37 para *iMS835*, 35 para *iMS705*, 33 para *iJB785* y 32 para *iSyf715*.

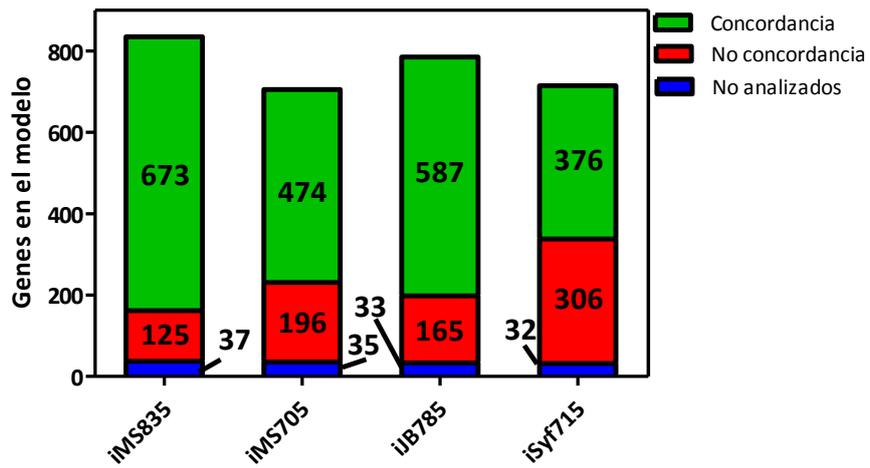


Figura 5-81. Comparación de esencialidad de genes *in vivo* frente a *in silico* de los modelos de Se7942, *iMS835*, *iMS705*, *iJB785* e *iSyf715*. La concordancia/no concordancia de genes está basada en la comparación con los datos *in vivo* de los resultados de RB-TnSeq obtenidos por (Broddrick et al., 2016).

Tabla 5-10. Comparación de genes esenciales entre los cuatro modelos metabólicos de Se7942. Modificado de (Broddrick et al., 2016).

| Categoría de genes | /MS835 | | /MS705 | | /JB785 | | /SY715 | | | | |
|-------------------------|------------------------|------------|------------------------|------------|------------------------|------------|------------------------|------------|-----|-----|------------|
| | Incluidos en el modelo | Predicción | | | |
| Esenciales | 718 | 423 | 340 (84) | 361 | 323 | 457 | 399 | 457 | 360 | 134 | 118 (33) |
| Beneficiosos | 157 | 5 | 0 (0) | 53 | 0 (0) | 5 | 72 | 5 | 56 | 1 | 0 (0) |
| No esenciales | 1.748 | 412 | 287 (90) | 256 | 244 (95) | 323 | 281 | 323 | 266 | 579 | 258 (97) |
| No datos <i>in vivo</i> | 100 | 0 | n.d. | 35 | n.d. | 0 | 33 | 0 | 32 | 0 | n.d. |
| Genes totales | 2.723 | 835 | 627 (79**) | 705 | 474 (71**) | 785 | 785 | 785 | 714 | 714 | 376 (55**) |

*Igual al número de genes correctamente predichos de ser esenciales *in silico*.

**Genes totales correctos/(genes totales incluidos en el modelo – genes en el modelo *in vivo*) x 100 %.

n.d., no determinado.

Si analizamos más en detalle los genes que presentan discordancias entre la esencialidad *in vivo* y la predicción *in silico*, podremos clasificarlos en siete categorías (**Figura 5-82**). Analizamos únicamente dichos genes en los modelos *iMS835* e *iJB785*, con el fin de determinar qué categorías habían mejorado en el modelo *iMS835* tras incluirle parte de los genes no comunes entre *iJB785* e *iMS705*. En la categoría denominada “suposición estado constante”, se incluyen los genes requeridos para que el modelo mantenga un balance de masas, aunque causen una discrepancia. En otras palabras, el organismo posee las enzimas para llevar a cabo una determinada reacción, aunque se sabe que no son funcionales, pero en el modelo son necesarias para su funcionamiento. Al comparar los genes incluidos en esta categoría en los modelos *iMS835* e *iJB785* encontramos un número similar en ambos modelos (3 y 9, respectivamente). Esto indicaría que los nuevos genes incluidos en *iMS835*, procedentes de *iMS705* y ausentes en *iJB785*, están solucionando alguno de los problemas causados por estas enzimas. En segundo lugar tenemos los genes que se encuentran “fuera de alcance”, perteneciendo a esta categoría los genes cuyo impacto en la esencialidad no puede determinarse en el modelo a escala genómica. Dentro de esta categoría suelen incluirse enzimas reguladoras, que no tienen efecto *in silico*, pero sí *in vivo*. En este caso se encuentran 12 genes en el modelo *iJB785* por 4 en *iMS835*. La tercera categoría, denominada genes “no esenciales para los componentes de la biomasa” incluye genes que codifican para componentes de la biomasa que son esenciales para el crecimiento *in silico*, pero son prescindibles *in vivo*. En esta categoría se encuentran 6 genes en el modelo *iMS835* por 16 en *iJB785*. El mayor número de genes en el segundo caso se debe en gran medida al uso de una biomasa “core”, en hemos eliminado aquellos metabolitos de la reacción de biomasa que no eran producto de una reacción esencial. En cuarto lugar tenemos genes con una “isoenzima anotada”, es decir, aquellos genes para los que existe una isoenzima funcional anotada en el genoma que provoca redundancia y causa discrepancia. En el modelo *iMS835* 20 genes pertenecen por 19 en *iJB785*. La siguiente categoría incluye genes que codifican para subunidades no esenciales en una reacción esencial o beneficiosa y que causa discrepancia. El número de genes incluidos en esta categoría son 25 en *iMS835* y 23 en *iJB785*. A continuación, analizamos los genes que suponen una “red bypass”, definiéndose estos como genes cuyo delección puede ser suplida por una ruta alternativa que conecta la red metabólica. Al igual que en el caso anterior, el número de genes incluidos en esta categoría fue similar en ambos modelos (34 en *iMS835* y 36 en *iJB785*). En último lugar tenemos la categoría “hueco conocido” que incluye genes que dan lugar a un hueco en la ruta metabólica, ya que su función o requerimiento metabólico *in vivo* es desconocido. En esta categoría el número de genes en el modelo *iMS835* fue de 33, mientras que en *iJB785* fue de 50.

Analizando en conjunto estos datos, concluimos que las discrepancias en la predicción de la esencialidad de genes en ambos modelos fueron similares, independientemente del número total de genes afectados. Solo en el caso de cuatro categorías, la de suposición de estado constante, la de fuera de alcance, la de biomasa esencial y la de hueco conocido, estos valores fueron inferiores en el modelo *iMS835*. Esto sería consecuencia de las modificaciones realizadas en el modelo *iJB785* y que dieron lugar al modelo *iMS835*, como se explica al inicio de este apartado.

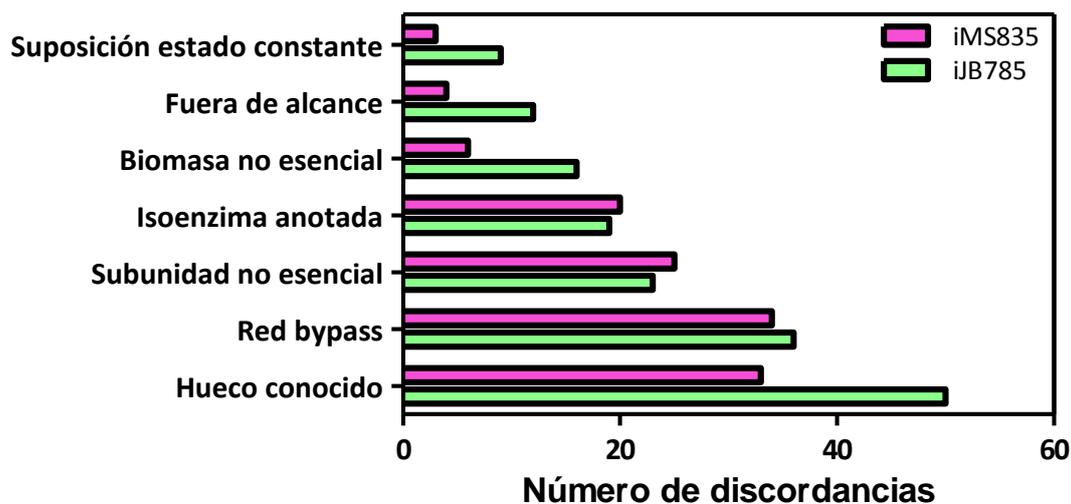


Figura 5-82. Comparación de los genes de los modelos *iMS705* e *iJB785* con discordancia entre la esencialidad *in vivo* y la predicción *in silico*. Modificado de (Broddrick et al., 2016).

5.3.3. Comprobación *in silico* del efecto de las modificaciones realizadas en *Se7942* para la producción de ácido α -linolénico

Para comprobar si el modelo era capaz de predecir modificaciones que mejoraran la producción de ALA, decidimos evaluar *in silico* las modificaciones realizadas *in vivo*. Analizamos, por tanto, el efecto *in silico* de la modificación de la expresión de los genes *fabH*, *fabF* y *fadD* en el modelo metabólico *iMS835*. Previamente, añadimos al modelo las reacciones necesarias para la síntesis de ALA, que fueron tomadas del modelo *iJN678* de *Ss6803* (Nogales et al., 2012). A este modelo lo denominaremos *iMS835_ALA*.

Primeramente, representamos la producción de ALA frente a la tasa de crecimiento (Figura 5-83). La relación entre la producción de ALA y la tasa de crecimiento es un equilibrio de tipo sencillo, ya que a mayor producción de ALA se produce menor crecimiento y viceversa. Además, esta producción está acoplada al crecimiento, ya que cuando no se produce ALA el organismo no es capaz de crecer.

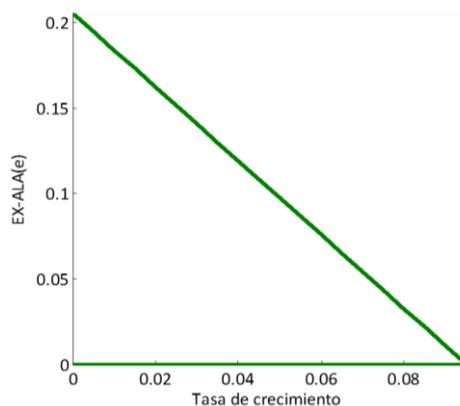


Figura 5-83. Efecto de la producción de ALA sobre la tasa de crecimiento. La producción de ALA se representa en el eje Y como EX_ALA(e). La tasa de crecimiento (k') se representa en el eje X. La línea verde representa la relación entre la producción de ALA y el crecimiento de Se7942. Las unidades para la tasa de crecimiento son h^{-1} y para el EX_ALA(e) son $mmol \cdot gDW^{-1} \cdot h^{-1}$.

5.3.3.1. Modificación *in silico* de la expresión del gen *fabH*

En primer lugar analizamos la modificación *in silico* de la expresión del gen *fabH*, que cataliza una reacción en el modelo (KAS15). En la **Figura 5-84** se representa la relación en el modelo *iMS835_ALA* entre el flujo de esta reacción con la tasa de crecimiento (Panel A) o con la producción de ALA (Panel B). El estado ideal de la célula es la región representada en verde, en la que presenta una alta tasa de crecimiento, aunque la producción de ALA no sea la máxima que podría alcanzar. Esto es debido a que la producción de ALA supone una carga para la célula, y por ello tiene que encontrar un equilibrio entre crecimiento y producción de ALA. Esta zona verde es la interfase entre la zona naranja y la zona azul. La zona naranja representa el estado natural de la célula, ya que comprende todas las soluciones en las que se produce ALA sin afectar a la tasa de crecimiento. Es un espacio de soluciones y no una solución concreta, porque dentro de una población cada una de las células presenta una tasa de crecimiento y una producción de ALA diferente. Dentro de la zona naranja, se representa con una franja amarilla, las soluciones que corresponderían a una delección hipofuncional del gen *fabH*, cuyos resultados *in vivo* se explican en detalle en el apartado **5.2.4**. La zona azul representa el estado de la célula para una producción óptima de ALA, aunque sea a expensas del crecimiento de Se7942. En esta región se incluyen todas las soluciones en las que la célula presenta una menor tasa de crecimiento y una mayor producción de ALA que la cepa silvestre. Dentro de esta zona encontramos una franja roja, región en la cual se encontraría Se7942 cuando se sobreexpresa *fabH* junto a las desaturasas en Se7942. Sin embargo, esta condición no fue probada en esta tesis, porque la sobreexpresión de *fabH in vivo* no aumentaba la producción de intermediarios C18 necesarios

para obtener ALA. Hay que destacar que los datos *in silico* referentes a una delección hipofuncional del gen *fabH* coinciden con los datos *in vivo*. En ambos casos, no se obtiene una mayor producción de ALA cuando se realiza una delección parcialmente funcional de *fabH*.

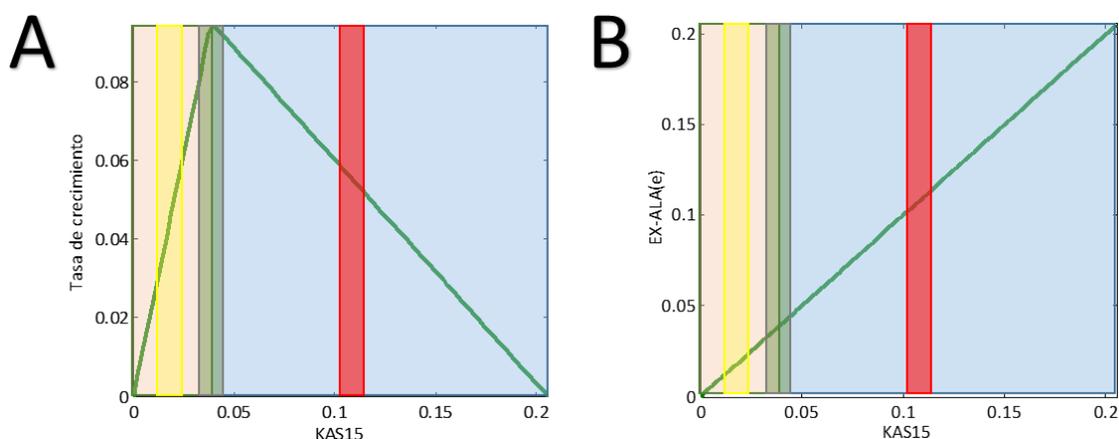


Figura 5-84. Efecto de la tasa de crecimiento y la producción de ALA en el flujo de la reacción KAS15 en *iMS835_ALA*. (A) Relación entre el flujo de KAS15 expresado en $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$ frente a la tasa de crecimiento en h^{-1} . La línea verde representa la relación entre la tasa de crecimiento de Se7942 y el flujo de la reacción KAS15. (B) Relación entre el flujo de KAS15 expresado en $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$ frente a la producción de ALA en $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$. La línea verde representa la relación entre la producción de ALA, representada como EX_ALA(e), y el flujo de la reacción KAS15. En naranja, se representa el estado natural de la célula. En azul, se representa el estado deseado para una producción óptima de ALA. En amarillo, se representa el estado de la célula cuando presenta una delección hipofuncional del gen *fabH* y que se encuentra incluido en el estado naranja. En rojo, se representa el estado de la célula cuando se sobreexpresa *fabH* y que se encuentra incluido en el estado azul. En verde, se representa el estado ideal para la célula, un crecimiento alto para sobrevivir en la población, aunque no una máxima producción de ALA.

Además, simulamos el efecto de la sobreexpresión de *fabH* (FabH-Up) y la delección parcialmente funcional de este mismo gen (FabH-Down) utilizando el modelo *iMS835_ALA*. Para ello, fijamos el flujo de la reacción KAS15 al doble o a la mitad del flujo óptimo para la producción de ALA en cada mutante, respectivamente. Los nuevos modelos construidos se utilizaron para estimar la distribución de flujos en la ruta de síntesis de ácidos grasos utilizando el método de muestreo de Monte Carlo implementado en el paquete COBRA Toolbox 2.0. Esta aproximación permite estimar el flujo más probable para cada una de las reacciones incluidas en esta ruta. Por tanto, es una herramienta útil cuando se comparan dos estados funcionales distintos, al identificar las diferencias más significativas en término de distribución de flujo. Los resultados (Figura 5-85) muestran que la probabilidad de densidad fue más alta para el mutante FabH-Down cuando el flujo a través de la reacción de biomasa (BOF_core) fue cero. Eso estaría indicando que la sobreexpresión de *fabH* tiene un efecto negativo sobre la biomasa (BOF_core) más marcado en este mutante. En el caso del mutante FabH-Up también existe un efecto negativo, puesto que se reduce el flujo de la reacción de biomasa respecto al de la cepa silvestre. El flujo de las desaturasas DesA (DES::12) y DesB(DES::15), está directamente relacionado con el

flujo obtenido para la producción de ácido oleico (EX_OA(e)) y de ALA (EX_ALA(e)), siendo sustrato y producto de estas reacciones, respectivamente. Por otro lado, cabe destacar el aumento de flujo de la desaturasa DesC en la reacción DESAT18a en el mutante FabH-Up, mientras que en el mutante FabH-Down disminuye. Este último resultado, guarda relación con los datos encontrados *in vivo*, en los que vemos una menor producción de ALA en el mutante que presenta una delección hipofuncional del gen *fabH* (apartado 5.2.3.2). Esta disminución podría deberse, a que el flujo que pasa por esta reacción ha disminuido con respecto a la cepa que lleva solo las desaturasas (WT).

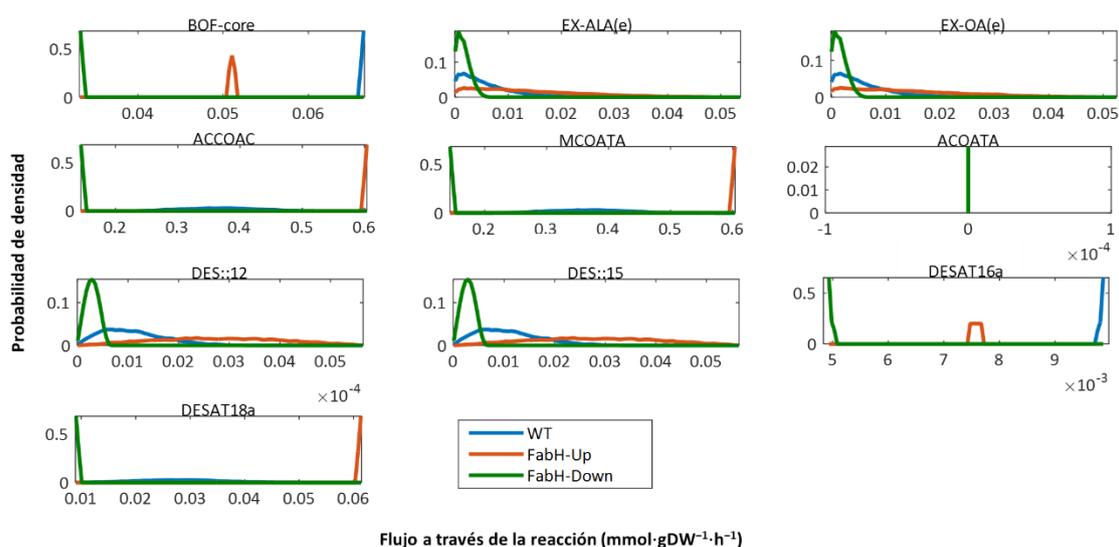


Figura 5-85. Histogramas de muestreo de algunas de las reacciones que componen la ruta de síntesis de ALA en Se7942. En azul, se representa a la cepa silvestre (WT); en naranja, se representa a la cepa en la que se ha sobreexpresado *fabH* (FabH-Up); en verde, se representa a la cepa en la que se ha delecionado parcialmente el gen *fabH* (FabH-Down). Se representa la probabilidad (eje Y) de encontrar un determinado flujo (mmol-gDW⁻¹-h⁻¹) para cada reacción dentro del rango representado (eje X) al maximizar el objetivo (tasa de crecimiento). Reacciones que se representan: “BOF_core”, biomasa; “EX_ALA(e)”, producción de ALA; “EX_OA(e)”, producción de ácido oleico; “ACCOAC”, reacción catalizada por la acetil carboxilasa; “MCOATA”, reacción catalizada por FabD; “ACOATA”, reacción catalizada por FabH; “DES::12”, reacción catalizada por DesA; “DES::15”, reacción catalizada por DesB; “DESAT16a”, reacción catalizada por DesC sobre ácidos grasos C16; “DESAT18a”, reacción catalizada por DesC sobre ácidos grasos C18.

5.3.3.2. Sobreexpresión *in silico* de *fabF*

En segundo lugar, analizamos la modificación *in silico* de la expresión del gen *fabF*, que cataliza diez reacciones en el modelo *iMS835_ALA* denominadas, EAR40y, EAR60y, EAR80y, EAR100y, EAR120y, EAR140y, EAR160y, EAR180x, OGMEACPS y OPMEACPS. Las dos últimas reacciones llevan a cabo reacciones intermediarias en la síntesis de biotina, mientras que las ocho restantes intervienen en la iniciación del ciclo de elongación ácidos grasos de cadena conocida. De esta manera, EAR40y elongaría ácidos grasos de 4 átomos de carbono; EAR60y,

elongaría ácidos grasos de 6 átomos de carbono; EAR80y, de 8; EAR100y, de 10 y así sucesivamente.

En la **Figura 5-86** se representa el flujo de las diez reacciones del modelo *iMS835_ALA* frente a la tasa de crecimiento. Para todas ellas, salvo para OGMEACPS y OPMEACPS, existe un flujo máximo de cada reacción, en torno a 0,05, que se produce el máximo crecimiento. Solo en el caso de la reacción EAR180x, el flujo es menor, siendo aproximadamente 0,01. Esto estaría relacionado con la especificidad de sustrato de la enzima, siendo esta menor para ácidos grasos de 18 átomos de carbono, que para los de 16. Esto estaría asociado con la cantidad de ácidos grasos presentes en la célula *in vivo*, siendo menor la cantidad de los primeros y mayor la de los últimos (ver **Figura 5-64**). Por otro lado, el flujo de las reacciones OGMEACPS y OPMEACPS presenta una relación lineal con la tasa de crecimiento: a mayor flujo, mayor crecimiento. Esto tiene sentido, ya que son reacciones que forman parte de la síntesis de biotina, un cofactor esencial para muchas enzimas metabólicas.

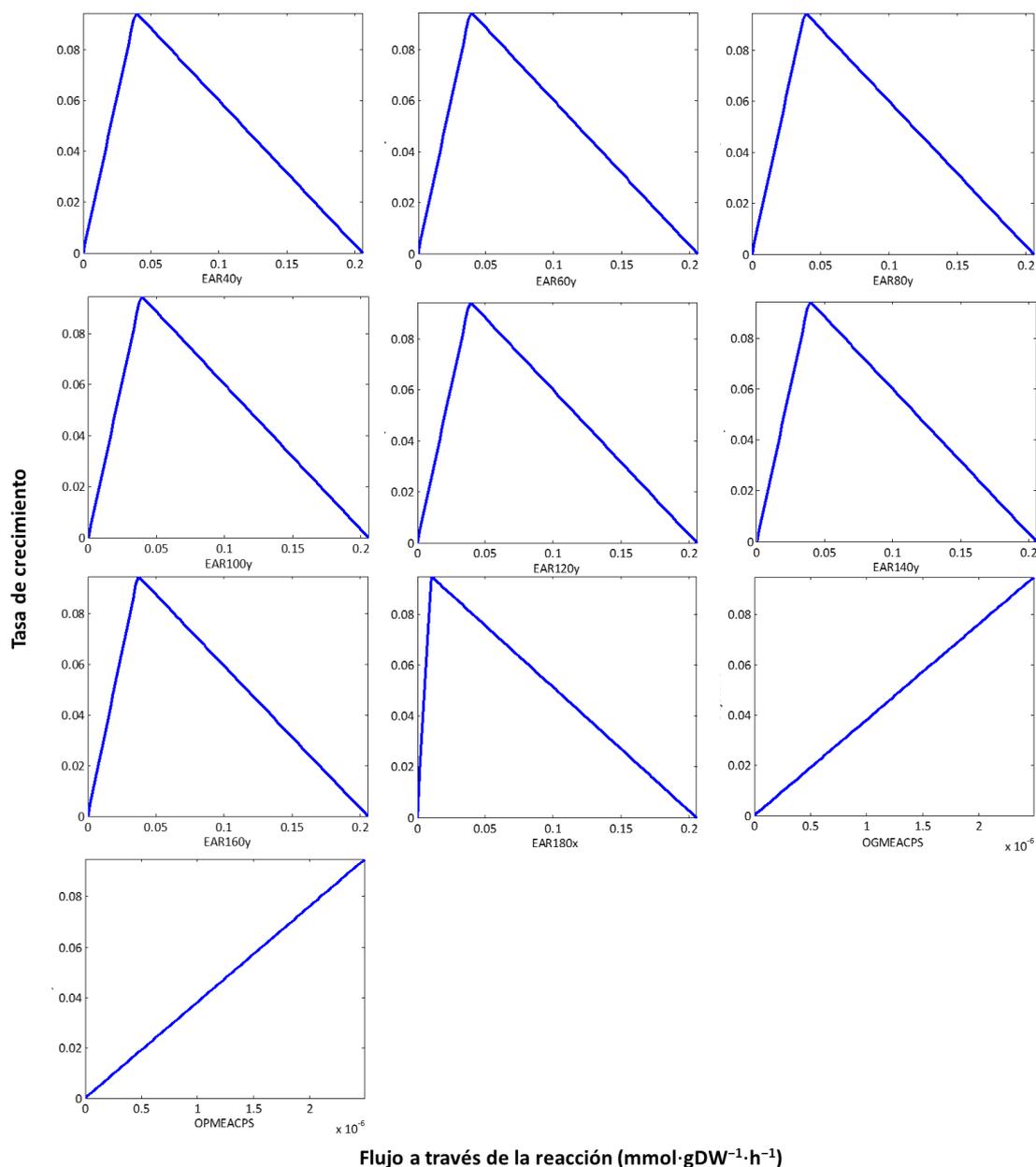


Figura 5-86. Relación entre el flujo de cada una de las reacciones catalizadas por FabF expresado en $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$ frente a la tasa de crecimiento en h^{-1} . Los diferentes paneles representan cada una de las reacciones catalizadas por FabF y cuyo nombre se indica bajo el eje X. Para realizar estas predicciones se utilizó el modelo *iMS835_ALA*.

En la **Figura 5-87** se representa el flujo de cada una de las reacciones catalizadas por FabF en el modelo *iMS835_ALA* frente a la producción de ALA. Para todas ellas, existe una relación lineal entre el flujo de cada reacción y la producción de ALA. Sin embargo, en todos los casos esta relación lineal es positiva, salvo para OGMEACPS y OPMEACPS, que es negativa. En el caso de las relaciones lineales positivas, a mayor flujo de la reacción se produce una mayor producción de ALA. Mientras que en las dos reacciones con una relación lineal negativa, a mayor flujo de la reacción se produce una menor producción de ALA. Esto se debe a que las reacciones

6. Discusión

OGMEACPS y OPMEACPS están involucradas en las síntesis de biotina y no de ácidos grasos. Todas ellas usan el mismo sustrato (acetil-CoA), por lo que al estar produciéndose ALA no quedaría suficiente cantidad de sustrato para las reacciones OGMEACPS y OPMEACPS. Por ello disminuye su flujo al aumentar la producción de ALA. En el mutante que sobreexpresa *fabF*, MSM39, el equilibrio se encuentra en la sobreexpresión puesto aumenta la producción de ALA sin afectar al crecimiento de Se7942.

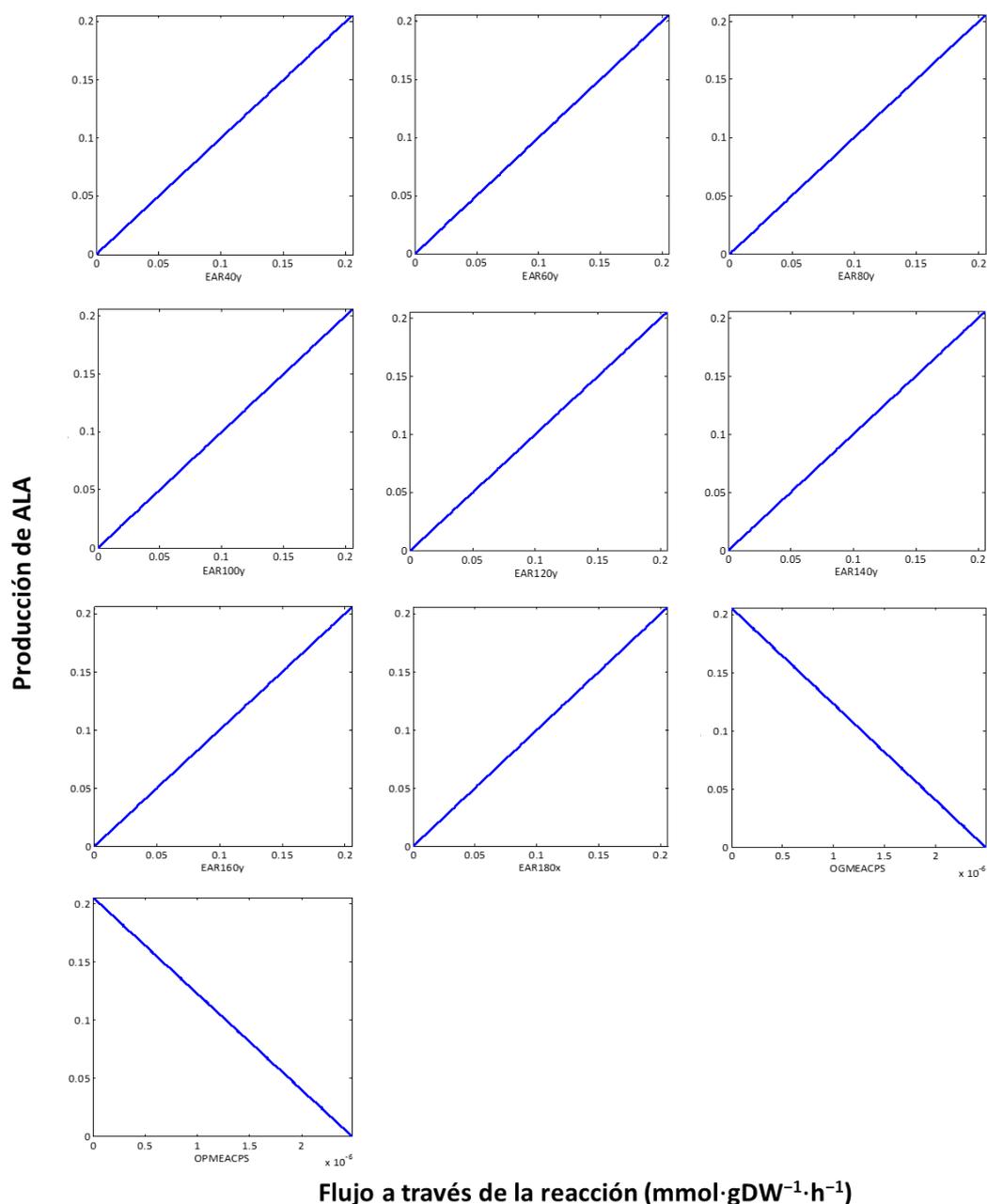


Figura 5-87. Relación entre el flujo de cada una de las reacciones catalizadas por FabF expresado en mmol·gDW⁻¹·h⁻¹ frente a la producción de ALA en mmol·gDW⁻¹·h⁻¹. Los diferentes paneles representan cada una de las reacciones catalizadas por FabF y cuyo nombre se indica bajo el eje X. Para realizar estas predicciones se utilizó el modelo *iMS835_ALA*.

Además de este análisis de flujos en la cepa silvestre, simulamos la construcción de dos mutantes *in silico* que sobreexpresaban *fabF* a dos niveles diferentes. En el primero, FabF-2x, las reacciones catalizadas por FabF presentaban un flujo doble al de la cepa silvestre, WT. En el segundo, FabF-4x, el flujo era cuatro veces superior al de la cepa silvestre. Como en el caso anterior, utilizamos el método de muestreo de Monte Carlo para estimar la distribución de flujos en la ruta de síntesis de ácidos grasos. Como se observa en la **Figura 5-88** (gráfica superior, título BOF_core), la sobreexpresión de *fabF* disminuye la biomasa (BOF_core) en ambas cepas. Por otro lado, la producción de ALA fue similar en la cepa WT y en la cepa FabF-2x, aunque en la cepa FabF-4x, dicha producción disminuyó (gráfica superior, título EX_ALA(e)). Esto es algo que ya habíamos visto en la **Figura 5-87**. Además, los flujos de la acetil carboxilasa (gráfica superior, título ACCOAC), así como de FabD (gráfica central, título MCOATA), aumentaron en ambos mutantes. Esto indicaría que la sobreexpresión de *fabF* simula la ruta de síntesis de ácidos grasos. Cabe destacar, por otro lado, que los flujos de las desaturasas DesA (gráfica central, título DES::12) y DesB (gráfica central, título DES::15), aumentaron en el mutante FabF-2x (pico naranja), mientras que en el mutante FabF-4x (pico verde) se mantuvieron al mismo nivel que la cepa WT (pico azul). Por otro lado, el flujo de la desaturasa DesC para ácidos grasos C16 (gráfica inferior, título DESAT16a) disminuyó en ambos mutantes, mientras que la desaturasa DesC (gráfica inferior, título DESAT18a) para ácidos grasos C18 aumentó. Estos resultados tienen concordancia con lo que ocurre *in vivo* al sobreexpresar *fabF*, y que puede observarse en la **Figura 5-67**.

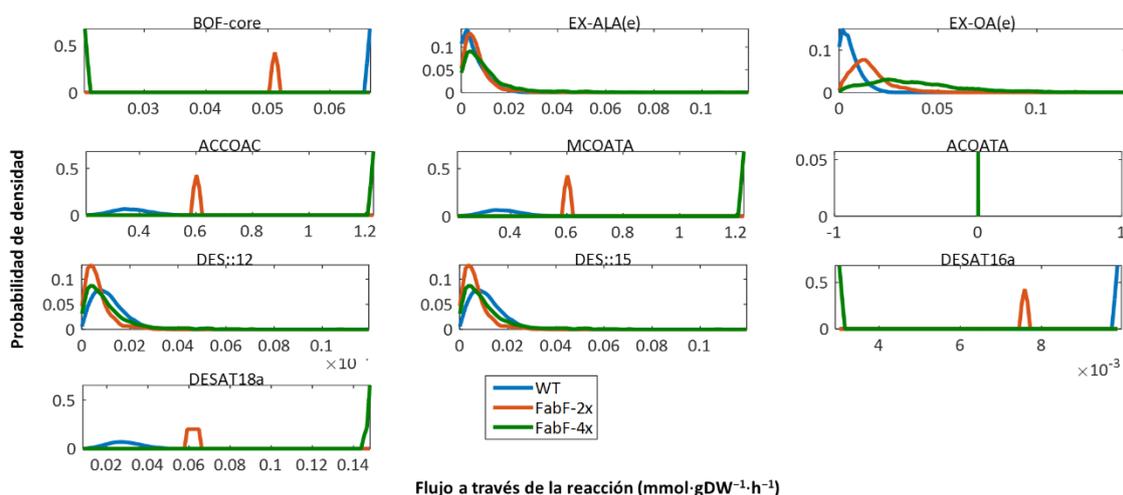


Figura 5-88. Histogramas de muestreo de algunas de las reacciones que componen la ruta de síntesis de ALA en Se7942. En azul, se representa a la cepa silvestre (WT); en naranja, se representa a la cepa en la que se ha sobreexpresado FabF con un flujo que duplica al de la cepa silvestre (FabF-2x); en verde, se representa a la cepa en la que se ha sobreexpresado FabF con un flujo que cuadruplica al de la cepa silvestre (FabF-4x). Se representa la probabilidad (eje Y) de encontrar un determinado flujo ($\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$) para cada reacción dentro del rango representado (eje X) al maximizar el objetivo (tasa de crecimiento). Reacciones que se representan: “BOF_core”, biomasa; “EX_ALA(e)”, producción de ALA; “EX_OA(e)”, producción de ácido oleico; “ACCOAC”, reacción catalizada por la acetil carboxilasa; “MCOATA”, reacción catalizada por FabD; “ACOATA”, reacción catalizada por FabH; “DES::12”, reacción catalizada por DesA; “DES::15”, reacción catalizada por DesB; “DESAT16a”, reacción catalizada por DesC sobre ácidos grasos C16; “DESAT18a”, reacción catalizada por DesC sobre ácidos grasos C18.

5.3.3.3. Deleción *in silico* de *fadD*

En último lugar simulamos el efecto de la deleción *in silico* del gen *fadD*, que cataliza diez reacciones en el modelo, denominadas FACOAL80i, FACOAL100i, FACOAL120i, FACOAL140i, FACOAL160i, FACOAL161i, FACOAL180i, FACOAL181i, FACOAL1821i y FACOAL1831i. Estas catalizan la degradación, respectivamente, de ácidos grasos C8:0, C10:0, C12:0, C14:0, C16:0, C16:1, C18:0, C18:1, C18:2 y C18:3.

En primer lugar, analizamos el flujo que presentan estas reacciones *in silico* en la cepa de Se7942 que sobreexpresa las desaturasas (MSM26), con el fin de establecer los límites que pueden alcanzar las mismas. Representamos solo una de las reacciones, puesto que las gráficas para el resto son iguales. En la **Figura 5-89** se representa el flujo de una de estas reacciones, FACOAL1831i, en el modelo *iMS835_ALA*. Como se observa en el panel A, el flujo máximo de esta reacción solo puede alcanzar valores cercanos a $1,000 \text{ mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$ ya que, a flujos mayores Se7942 no es capaz de crecer. Por otro lado, el flujo de esta reacción no afecta a la producción de ALA, que se mantiene constante en cualquiera de los flujos que se representan en la **Figura 5-89, B**.

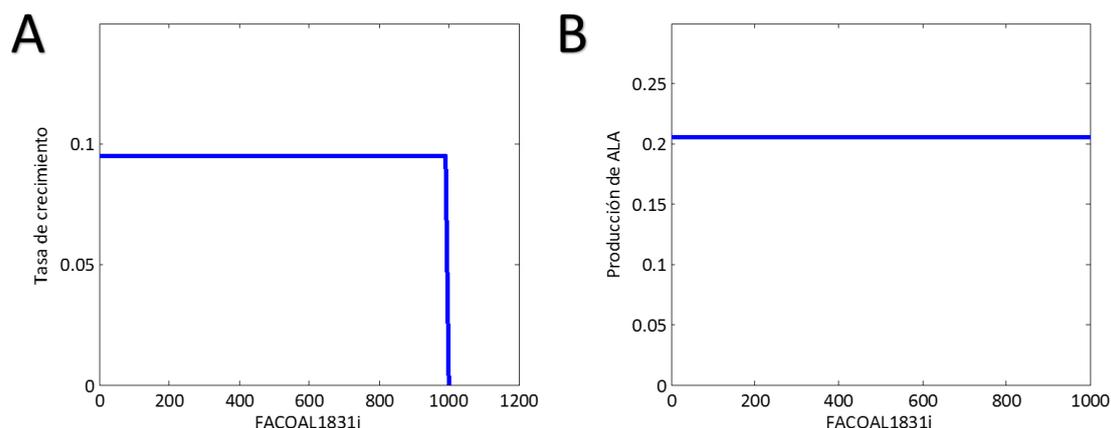


Figura 5-89. Efecto de la tasa de crecimiento y la producción de ALA en el flujo de la reacción FACOAL1831i en *iMS835_ALA*. (A) Relación entre el flujo de FACOAL1831i expresado en $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$ frente a la tasa de crecimiento en h^{-1} . La línea azul representa la relación entre la tasa de crecimiento de Se7942 y el flujo de la reacción FACOAL1831i. (B) Relación entre el flujo de FACOAL1831i expresado en $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$ frente a la producción de ALA en $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$. La línea azul representa la relación entre la producción de ALA.

Junto al análisis de flujos en la cepa silvestre, simulamos a partir del modelo *iMS835_ALA* construcción de un mutante en el que el flujo de todas las reacciones catalizadas por FadD fuera cero. Como en los casos anteriores, utilizamos el método de muestreo de Monte Carlo para estimar la distribución de flujos en la ruta de síntesis de ácidos grasos. Como se observa en la **Figura 5-90**, la delección del gen *fadD in silico* no tuvo ningún efecto sobre la biomasa (gráfica superior, título BOF_core), ya que el flujo de la reacción en la cepa WT y en la cepa DfadD son solapantes. Lo mismo ocurre con los flujos de las reacciones que componen la ruta de síntesis de ácidos grasos.

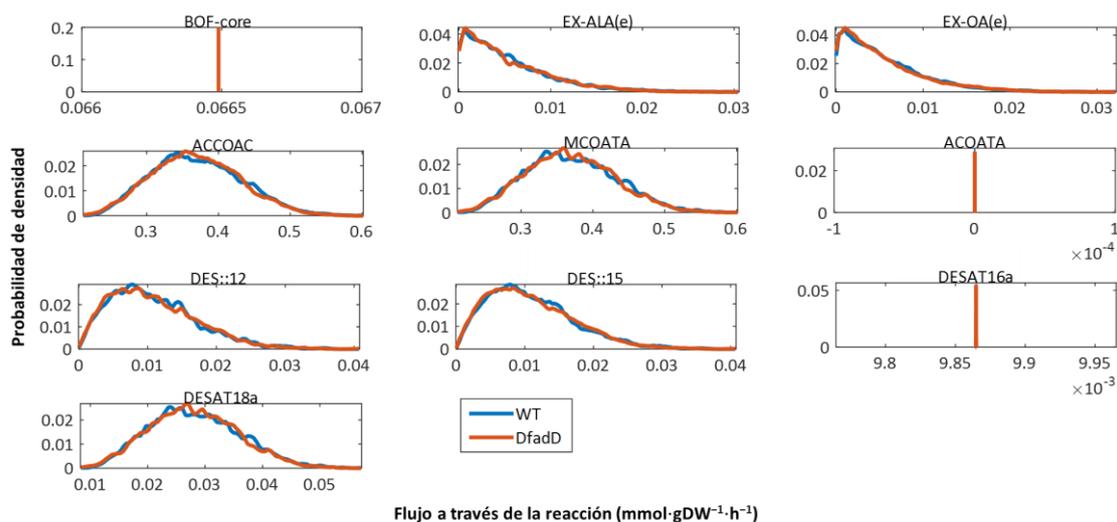


Figura 5-90. Histogramas de muestreo de algunas de las reacciones que componen la ruta de síntesis de ALA en Se7942. En azul, se representa a la cepa silvestre (WT); en naranja, se representa a la cepa en la que se ha deleciónado FadD (DfadD). Se representa la probabilidad (eje Y) de encontrar un determinado flujo ($\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$) para cada reacción dentro del rango representado (eje X) al maximizar el objetivo (tasa de crecimiento). Reacciones que se representan: “BOF_core”, biomasa; “EX_ALA(e)”, producción de ALA; “EX_OA(e)”, producción de ácido oleico; “ACCOAC”, reacción catalizada por la acetil carboxilasa; “MCOATA”, reacción catalizada por FabD; “ACOATA”, reacción catalizada por FabH; “DES::12”, reacción catalizada por DesA; “DES::15”, reacción catalizada por DesB; “DESAT16a”, reacción catalizada por DesC sobre ácidos grasos C16; “DESAT18a”, reacción catalizada por DesC sobre ácidos grasos C18.

5.3.4. Identificación *in silico* de nuevos mutantes “knockouts” para mejorar la producción de ácido α -linolénico

Por último, simulamos la construcción de diferentes mutantes “knockouts” de Se7942 *in silico* para mejorar la producción de ALA utilizando la “toolbox” de Cobra, llamada GDLS (“Genetic Design through Local Search”). En primer lugar, redujimos el tamaño del modelo, eliminando reacciones a través de las que no pasa flujo. A continuación, obtuvimos la lista de reacciones esenciales y no esenciales del modelo reducido. Con esto se pretendía obtener una lista de potenciales reacciones a deleciónar. Esta lista consta de los genes que se han descartado en el modelo reducido, junto a las reacciones no esenciales presentes en el mismo. Esta lista es necesaria para la “toolbox” explicada al inicio de este párrafo.

Una vez construida dicha lista, el siguiente paso fue utilizar dicha “toolbox” y obtener las soluciones. Las diferentes soluciones obtenidas, así como la tasa de crecimiento y la producción de ALA de estas cepas se muestra en la **Tabla 5-11**. Las reacciones G6PDH2r, GND y PGL forman parte de la ruta de las pentosas fosfato. La reacción PDH pertenece al metabolismo del piruvato y la reacción GART, al de purinas. Las reacciones PSERT, PGCD, GHMT2r, ALAD_L y VALTA pertenecen al metabolismo de aminoácidos, de tal manera que la primera participa en el

metabolismo de glicina, serina y treonina, las dos siguientes en el de arginina y prolina, la siguiente en el de alanina, aspartato y glutamato y la última en el de valina, leucina e isoleucina. Todas las reacciones, salvo la reacción GART, usan como sustrato el acetil-CoA. Este metabolito es el usado para iniciar la ruta de síntesis de ácidos grasos en Se7942, por lo que estas rutas están en competición directa con la ruta de síntesis de ácidos grasos. Por tanto, si las bloqueamos total o parcialmente con la delección de uno o varios genes implicados en dichas rutas, conseguiremos aumentar el acetil-CoA, y por ende, la producción de ALA. Esto es lo que ha ocurrido en todos los mutantes “knockouts” obtenidos *in silico*, aumentando en todos ellos la producción de ALA, aunque siempre a expensas de la tasa de crecimiento (Tabla 5-11, columnas 3 y 4). La tasa de crecimiento de la cepa silvestre *in silico* es de 0,094985 h⁻¹, mientras que en todos estos mutantes “knockouts” fue menor, tal y como se observa en la Tabla 5-11.

Tabla 5-11. Mutantes “knockout” obtenidos mediante simulación *in silico* para mejorar la producción de ALA.

| Nº “knockouts” | Nombre del mutante | Tasa de crecimiento | Producción de ALA | Reacciones | Nombre de la enzima |
|----------------|--------------------|---------------------|-------------------|------------|----------------------------------|
| 2 | M2-A | 0,028353 | 0,144198 | G6PDH2r | Glucosa-6-fosfato deshidrogenasa |
| | | | | PSERT | Fosfoferina transaminasa |
| | M2-B | 0,028353 | 0,144198 | GND | Fosfogluconato deshidrogenasa |
| | | | | PGCD | Fosfoglicerato deshidrogenasa |
| | M2-C | 0,028353 | 0,144198 | PGL | 6-fosfoglucolactonasa |
| | | | | PSERT | Fosfoferina transaminasa |
| 3 | M3-A | 0,045814 | 0,106410 | GART | GAR transformilasa-T |
| | | | | GHMT2r | Glicina hidroximetiltransferasa |
| | | | | GND | Fosfogluconato deshidrogenasa |
| 4 | M4-A | 0,020238 | 0,161759 | G6PDH2r | Glucosa-6-fosfato deshidrogenasa |
| | | | | GART | GAR transformilasa-T |
| | | | | PDH | Piruvato deshidrogenasa |
| | | | | PSERT | Fosfoferina transaminasa |
| | M4-B | 0,012951 | 0,17729 | ALAD_L | L-alanina deshidrogenasa |
| | | | | PDH | Piruvato deshidrogenasa |
| | | | | PGL | 6-fosfoglucolactonasa |
| | | | | PSERT | Fosfoferina transaminasa |
| | M4-C | 0,015382 | 0,172268 | ALAD_L | L-alanina deshidrogenasa |
| | | | | GND | Fosfogluconato deshidrogenasa |
| | | | | PGCD | Fosfoglicerato deshidrogenasa |
| | | | | VALTA | Valina transaminasa |
| 5 | M5-A | 0,015335 | 0,172370 | ALAD_L | L-alanina deshidrogenasa |
| | | | | GART | GAR transformilasa-T |

| Nº "knockouts" | Nombre del mutante | Tasa de crecimiento | Producción de ALA | Reacciones | Nombre de la enzima |
|----------------|--------------------------|---------------------|-------------------|------------|-------------------------------|
| 6 | M6-A | 0,010635 | 0,182540 | GND | Fosfogluconato deshidrogenasa |
| | | | | PGCD | Fosfoglicerato deshidrogenasa |
| | | | | VALTA | Valina transaminasa |
| | | | | ALAD_L | L-alanina deshidrogenasa |
| | | | | GART | GAR transformilasa-T |
| | | | | PDH | Piruvato deshidrogenasa |
| | | | | PGL | 6-fosfoglucolactonasa |
| PSERT | Fosfoserina transaminasa | | | | |
| | | | | VALTA | Valina transaminasa |

Por otro lado, si representamos la tasa de crecimiento frente a la producción de ALA en los mutantes de la **Tabla 5-11**, tal y como se muestra en la **Figura 5-91**, observamos que en ninguno de los mutantes obtenidos utilizando GDLS la producción de ALA está acoplada al crecimiento. Para que esto fuera así, el espacio de soluciones de cada mutante, el área representada en rojo en cada gráfica, no debería contener ninguna tasa de crecimiento en la que no hubiera producción de ALA. En otras palabras, nunca existiría un mutante cuya producción de ALA fuera $0 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{h}^{-1}$. Desde un punto de vista industrial, la obtención de un mutante que tenga acoplada la producción de un determinado compuesto al crecimiento es muy deseable. Este acoplamiento significa que solo se necesita un paso para producir el compuesto deseado. De otra manera, se necesitaría un primer paso para producir biomasa (etapa de crecimiento) y luego una etapa para la producción del compuesto en cuestión. La producción de un compuesto en una etapa es más barata y rápida que en dos, es por ello que es necesario hacer un esfuerzo adicional para conseguir dicho acoplamiento.

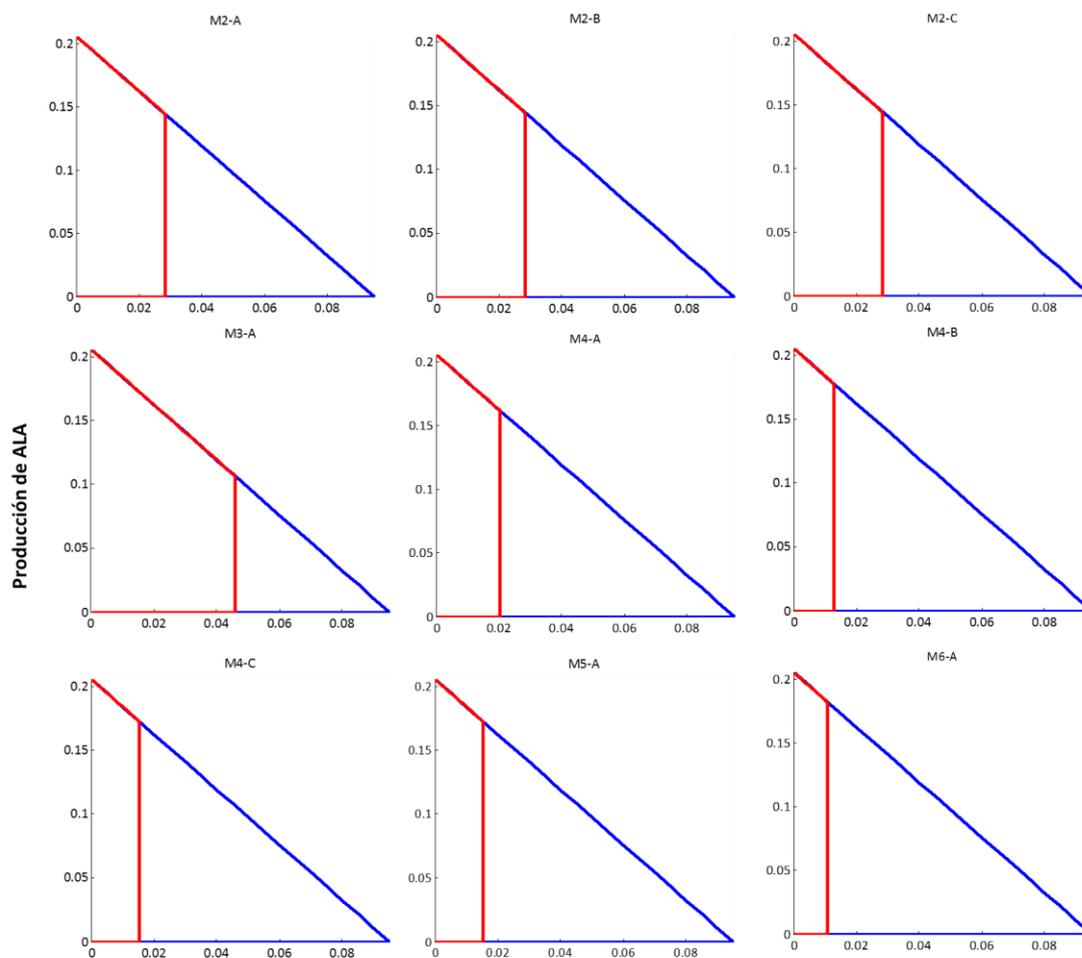


Figura 5-91. Representación de la tasa de crecimiento frente a la producción de ALA de los diferentes mutantes obtenidos *in silico*. Los diferentes paneles representan cada uno de los mutantes obtenidos *in silico* y cuyo nombre se indica sobre cada una de las gráficas. En azul, se representa la cepa silvestre; en rojo, se representa el mutante obtenido *in silico*. Las unidades de la tasa de crecimiento son h^{-1} y las de la producción de ALA son $mmol \cdot gDW^{-1} \cdot h^{-1}$.

Discusión

6. Discusión

En esta tesis doctoral se han llevado a cabo diferentes estrategias y experimentos para manipular la cianobacteria Se7942 por ingeniería genética y metabólica, con el fin de obtener un chasis más atractivo para su uso a nivel industrial. Por un lado, se ha mejorado su tiempo de generación mediante evolución experimental, con el objetivo de acortar los procesos de producción a nivel industrial. Por otro, se ha producido ácido alfa-linolénico, un ácido graso omega-3 de gran importancia nutricional. En tercer lugar, se ha desarrollado un modelo metabólico de Se7942, que nos ha permitido identificar nuevas dianas para incrementar la producción de ácidos grasos omega-3.

6.1. Evolución experimental de Se7942

En el año 1955, Se7942 (denominada en aquella época *Anacystis nidulans*) fue descrita como una cianobacteria de crecimiento rápido por (Kratz and Myers, 1955). Estos autores determinaron que una alta intensidad de luz y una alta temperatura, en torno a 41 °C, eran condiciones ideales para su rápido crecimiento (tiempo de generación de 2 h). Sin embargo, tras su mantenimiento en condiciones de laboratorio, esta cepa ha perdido su capacidad para crecer rápidamente, incluso a temperaturas por encima de los 38 °C. (Yu et al., 2015) demostraron que el tiempo de generación actual a 41 °C para Se7942 es de 4,9 h.

El objetivo de nuestro trabajo fue conseguir que Se7942 recuperase su capacidad de crecimiento rápido. Una cianobacteria que crezca rápido y pueda ser fácilmente manipulada mediante ingeniería genética sería una candidata ideal para una gran cantidad de aplicaciones biotecnológicas, así como para su uso en procesos industriales. Para ello, llevamos a cabo un protocolo de evolución experimental. Esta metodología es ideal cuando se quieren modificar fenotipos complejos, ya que no es necesario hacer suposiciones sobre cuáles o cuántas mutaciones serán beneficiosas. Además, ofrece otras ventajas como que no son necesarias herramientas moleculares complejas, solo es necesaria la capacidad para crecer continuamente los cultivos. Hay que añadir que, al ser la tasa de mutación natural el único recurso de variación genética, las cepas resultantes son variantes genéticas de la cepa ancestral y no se consideran OMGs (Organismos Modificados Genéticamente). Esto ofrece una ventaja

para su uso posterior en producción industrial, así como el hecho de que el experimento de evolución se hace en las condiciones de cultivo que van a usarse posteriormente en dicho proceso.

6.1.1. Estudio de las cepas obtenidas mediante evolución experimental

En este experimento de evolución obtuvimos dos poblaciones, a las que denominamos C11-G₁₂₄₈ y C12-G₁₁₆₈, y cuyo tiempo de generación había mejorado sustancialmente, siendo de 3.13 h y 2.95 h, respectivamente. Por tanto, la evolución es un método que nos ha permitido mejorar el tiempo de generación y cumplir uno de los objetivos planteados.

Por otro lado, cabe destacar que fueron pocas las mutaciones que aparecieron durante el proceso de evolución. Como se puede ver en la **Figura 5-3**, en el caso de C11-G₁₂₄₈ aparecieron únicamente cinco mutaciones, mientras que en C12-G₁₁₆₈ fueron siete. (Yu et al., 2015) describieron una nueva cianobacteria de crecimiento rápido a la que denominaron *Synechococcus elongatus* UTEX 2973. Esta cepa había sido depositada en la colección de cultivos de algas de la Universidad de Texas (UTEX) como *Synechococcus leopoliensis* UTEX 625. Sin embargo, en los últimos años, perdió su capacidad de crecimiento a 38 °C. (Yu et al., 2015) seleccionaron una colonia que era capaz de crecer a 38 °C en un cultivo mixto de *Synechococcus leopoliensis* UTEX 625, y la depositaron en la colección de cultivos de algas UTEX como *Synechococcus elongatus* UTEX 2973. Esta cepa presenta pequeñas diferencias en el genoma respecto a Se7942. Al analizar el número de mutaciones entre los dos genomas, determinaron que un total de 55 SNPs e indels estaban presentes en el genoma de esta cianobacteria. Esta cifra es mucho más alta que la que nosotros hemos obtenido durante el experimento de evolución experimental, lo que hace más difícil la identificación de la/s posible/s mutación/es responsables de dicho fenotipo.

Cabe destacar que, de las cinco mutaciones primeramente detectadas en C11-G₁₂₄₈ mediante la secuenciación “Illumina”, solo cuatro de ellas se confirmaron mediante secuenciación “Sanger”, tal y como se muestra en la **Figura 5-22**. La mutación que afecta al gen de la ácido graso de cadena larga-CoA ligasa no fue detectada utilizando secuenciación “Sanger”, y por tanto se debe un error de la secuencia de “Illumina”.

También hay que destacar el relativamente escaso número de generaciones que necesitamos para mejorar el tiempo de generación de Se7942. En el caso de C11 fueron necesarias 1248, mientras que en el caso de C12, fueron 1168. En el experimento de evolución realizado por Lenski, descrito en el apartado **2.2.1.1**, se necesitaron tan solo 200 generaciones para que mejorara el crecimiento de *E. coli* (Barrick et al., 2009).

6.1.2. Análisis del tiempo de generación de las cepas evolucionadas y los mutantes construidos

La parte más importante de los experimentos relacionados con las cepas evolucionadas fue la identificación de las mutaciones responsables del fenotipo de crecimiento rápido. Una vez realizada la secuenciación completa de las poblaciones C11-G₁₂₄₈ y C12-G₁₁₆₈, sus secuencias se compararon con la cepa silvestre y se identificaron las mutaciones presentes en cada una (**Figura 5-3** y **Tabla 5-1**). Nos resultó sorprendente que con solo 4 mutaciones en C11-G₁₂₄₈ y 7 en C12-G₁₁₆₈, las cepas evolucionadas hubiesen prácticamente duplicado la velocidad de crecimiento de Se7942.

Como C11-G₁₂₄₈ alcanzaba un mayor crecimiento con solo cuatro cambios con respecto a Se7942 silvestre, decidimos construir mutantes individuales para aclarar el rol de cada uno en la tasa de crecimiento. Dos de las cuatro mutaciones quedaron descartadas (un cambio en un gen codificante de una proteína hipotética y otro en el gen *pilA*, que codifica para las subunidades del pilus) y concentramos nuestros esfuerzos en las otras dos que, a priori, parecían más relevantes, *sasA* y *aroK*. Añadimos también a este estudio la mutación en el gen *fadD* de C12-G₁₁₆₈, por tratarse de un gen implicado en la degradación de ácidos grasos. Para completar el estudio, decidimos además, construir mutantes dobles (*sasA-aroK*, *sasA-fadD*, *aroK-fadD*) y el mutante con las mutaciones en los tres genes de interés (para un resumen de los mutantes construidos, ver **Figura 5-32**).

Para construir dichos mutantes, elegimos el método descrito por (Matsuoka et al., 2001), cuya metodología se describe en el apartado **4.8.4**. Este método conlleva el uso de una cepa de Se7942 con una mutación en el gen *rps12*, que confiere resistencia a estreptomicina. Para descartar que esta mutación tuviera un efecto negativo en el crecimiento de Se7942, comparamos el tiempo de generación de la cepa silvestre con el de la cepa MSM1, que contenía la mutación en el gen *rps12*. No se obtuvieron diferencias significativas en el tiempo de

generación entre ambas cepas, tal y como se muestra en la **Figura 5-34, A**. Por tanto, la mutación en el gen *rps12* no afecta al crecimiento de Se7942 y podíamos utilizar este sistema para construir los mutantes individuales sin dar lugar a un enmascaramiento del efecto de las mutaciones clonadas de las cepas evolucionadas.

Una vez construidos los mutantes sobre la cepa MSM1, los derivados resultantes debían ser evaluados en su velocidad de crecimiento. En primer lugar, hubo que determinar qué intensidad de luz sería la apropiada para medir dicho parámetro. Esto se debe a que durante el experimento de evolución se utilizaron como fuente de luz lámparas fluorescentes, mientras que las curvas de crecimiento para calcular el tiempo de generación se realizaron en el “Multi-cultivator MC 1000-OD” que llevan instaladas luces LED. Tal y como se puede observar en la **Figura 6-1**, los espectros de emisión de ambas fuentes de luz son similares. Ambos tipos de lámparas emiten un amplio espectro de luz con una alta liberación de fotones en los rangos espectrales azul y verde, que disminuye a partir de 650 nm. Dada la similitud entre ambos espectros de emisión, los LEDs pueden reemplazar adecuadamente a las luces fluorescentes en el mismo rango de temperatura de color sin causar cambios significativos en la tasa de crecimiento de las cianobacterias, así como en sus propiedades bioquímicas (Schulze et al., 2014).

Cabe destacar que en los últimos años el precio de las luces LEDs ha disminuido significativamente, aunque todavía son de media cuatro veces más caros que las luces fluorescentes. Sin embargo, el coste inicial puede ser compensado por su larga vida útil y a menudo por la mejor eficiencia energética de las luces LEDs respecto a las luces fluorescentes (Ooms et al., 2016).

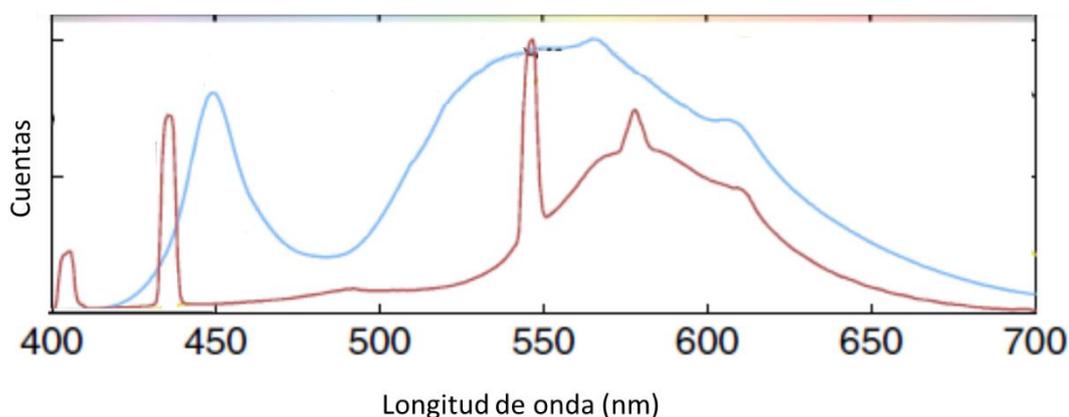


Figura 6-1. Comparativa del espectro de luz emitido por luces fluorescentes y por luces LED. En rojo, se muestra el espectro de emisión de las luces fluorescentes blancas frías utilizadas en el experimento de evolución. En azul, espectro de emisión de las luces LED frías con filtro LP650 del “Multi-cultivator MC 1000-OD”. En cada gráfica, en el eje X se representan las longitudes de onda mientras que en el eje Y se representan los valores relativos del poder óptico irradiado a cada longitud de onda. Modificado de (Ooms et al., 2016).

Analizamos el crecimiento de los diferentes mutantes, comparando el tiempo de generación obtenido a una intensidad de luz de $1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ bajo luces fluorescentes (**Figura 5-35**) y LEDs (**Figura 5-37**). No observamos diferencias en los tiempos de generación obtenidos para la misma cepa con las diferentes fuentes de luz, un resultado similar al descrito por (Schulze et al., 2014).

Por otro lado, analizamos el efecto de la intensidad de luz sobre el tiempo de generación. Para ello, comparamos el tiempo de generación obtenido utilizando luces LED a $280 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (**Figura 5-35**) y a $1.313 \mu\text{mol fotones}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (**Figura 5-37**). A baja intensidad de luz, **Figura 5-35**, solo el mutante MSM15 no presentaba diferencias en el tiempo de generación respecto a C11-G₁₂₄₈. Es decir, son necesarias tanto la mutación en el gen *aroK* (mutación #12) como la presente en *sasA* (mutación #21) para que el tiempo de generación no varíe respecto a C11-G₁₂₄₈. Sin embargo, a alta intensidad de luz, tanto usando luces LED como lámparas fluorescentes, basta una sola mutación, aquella en *aroK* (mutante MSM3), para alcanzar la tasa de crecimiento de C11-G₁₂₄₈ (**Figura 5-37** y **Figura 5-36**).

En resumen, el tiempo de generación no se vio afectado por el tipo de luz utilizada, ya sean luces LED o fluorescentes. Por otro lado, la intensidad de luz sí afectó al tiempo de generación. A baja intensidad de luz son necesarias dos mutaciones, la mutación #12 y la #21, que se encuentra en el gen que codifica para la shikimato kinasa y para la histidina kinasa, respectivamente, mientras que a alta intensidad de luz, solo es necesaria la mutación #12.

6.1.3. Estudio de la transformabilidad natural en las cepas evolucionadas

Las cepas evolucionadas presentaron, sin embargo, la pérdida de una de las características que permite la fácil manipulación genética de Se7942, la transformabilidad natural. Aunque Se7942 puede también recibir ADN mediante conjugación, el protocolo es más trabajoso y, además, es necesaria la comprobación posterior de que las colonias transconjugantes obtenidas son axénicas. Además, la selección de los transconjugantes con el antibiótico es más lenta, relacionada con el tiempo que tardan las bacterias donadoras del plásmido en morir. Datos previos de nuestro laboratorio (Tesis David Encinas Pisa) nos han permitido la comparación directa de la eficiencia de conjugación y transformación de Se7942. En las condiciones experimentales en las que se realizan estas técnicas (ver epígrafes **4.8.2** y **4.8.3**) la frecuencia de conjugación y transformación fue de $1\cdot 10^{-3}$ y $1,2\cdot 10^{-6}$, respectivamente. Además, en esa misma

tesis se demostró que la eficiencia de estos sistemas de introducción de ADN en Se7942 estaba afectada en gran medida por los sistemas de restricción-modificación presentes en este organismo. Sin embargo, una de las ventajas que presenta la introducción de material genético en Se7942 mediante la conjugación es que no tiene limitación en el tamaño del ADN transferido. En muchos casos, esta metodología es elegida cuando se quieren introducir fragmentos de gran tamaño o incluso cuando se quiere obtener merodiploides (Ruffing, 2011).

En bacterias Gram negativas, la competencia natural está relacionada con el pilus tipo IV (T4P) que forma el sistema de secreción tipo II (T2SS). En cianobacterias, los homólogos a las proteínas que forman dicho pilus, están involucrados en la transformación natural. Además de formar la estructura y el motor del pilus, estas proteínas participan en la unión de ADN exógeno en la superficie celular, en el transporte al citoplasma, en el procesamiento del ADN en una hebra simple, y en la protección frente a endonucleasas de restricción (Chen and Dubnau, 2004). Tanto C11-G₁₂₄₈ como C12-G₁₁₆₈ presentan mutaciones en uno o más genes que componen dicha estructura. Para entender mejor cómo pueden estar afectando dichas mutaciones a la transformación natural, vamos a describir cada uno de los componentes de dicha estructura.

El pilus tipo IV consiste en un homopolímero de pequeñas proteínas PilA organizadas helicoidalmente. La extensión o retracción del pilus consiste en la adición o sustracción de unidades de PilA mediante ATPasas de ensamblaje (PilB) y ATPasas de desensamblaje (PilT). Las pilinas PilA (en Se7942 existen dos tipos, PilA1 y PilA2) se sintetizan como prepilinas que son procesadas proteolíticamente por acción de PilD, situada en la membrana interna (Bhaya et al., 2000). PilC es una proteína localizada en la membrana interna y es necesaria para el ensamblaje del pilus, ya que proporciona la base sobre la que se va a construir dicha estructura (Imam et al., 2011). En *Streptococcus pneumoniae*, existen evidencias de interacción directa entre pilus y ADN mediante microscopía electrónica (Laurenceau et al., 2013). Además, estos autores proponen que el pilus es el receptor primario durante la transformación natural. Por tanto, la función del pilus sería captar el ADN extracelular y transportarlo al espacio periplásmico a través del canal formado por la proteína PilQ (Burkhardt et al., 2011). Una vez en el espacio periplásmico, el ADN pasaría de cadena doble a cadena sencilla por acción de ComA (Yoshihara et al., 2001) o ComF (Nakasugi et al., 2006). El ADN captado pasaría al citosol gracias al hipotético canal similar a ComEC propuesto por (Yoshihara et al., 2001). Un esquema de cómo tendría lugar este proceso se muestra en la **Figura 6-2**.

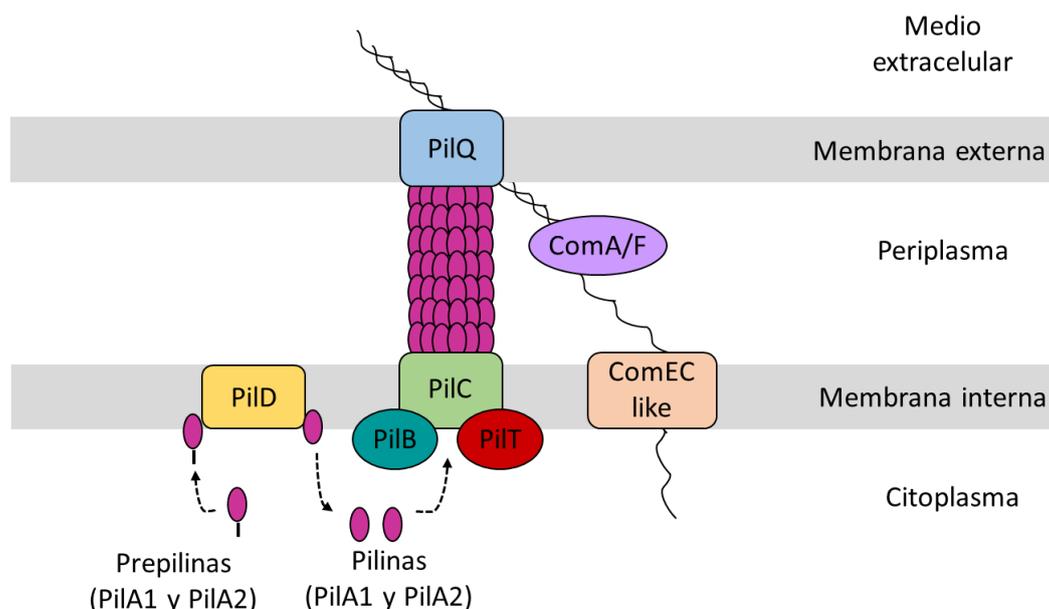


Figura 6-2. Esquema de la maquinaria del pilus T4P de *Se7942*. Se representa el proceso de ensamblaje del pilus, así como las proteínas involucradas en la captación y translocación de ADN al citoplasma. Modelo basado en (Chen and Dubnau, 2004; Nakasugi et al., 2006; Okamoto and Ohmori, 2002; Schuergers and Wilde, 2015; Yoshihara et al., 2001).

En *Ss6803*, la interrupción de los genes *pilA*, *pilT* y otros que codifican componentes de la estructura del pilus, afectan a la formación del pilus, así como a la competencia natural (Bhaya et al., 2000; Yoshihara et al., 2001). La interrupción de cualquiera de los genes implicados en la competencia, *comA*, *comF*, *comE* o *comC*, también da lugar a la pérdida de esta capacidad (Nakasugi et al., 2006).

En el caso de *C11-G₁₂₄₈*, la proteína implicada en la transformación que está mutada es PilA, que presenta una inserción de 6 nucleótidos sin cambios en el marco de lectura (ver **Tabla 5-1** y **Figura 5-4**). Como se ha mencionado en el apartado **5.1.2.1**, la inserción afecta a una hélice alfa, cambiando el plegamiento de la proteína, pero no su hidrofobicidad. Aunque podría pensarse que este cambio afectaría a la unión entre cadenas en el proceso de ensamblaje del pilus, si analizamos las imágenes tomadas con el microscopio electrónico de transmisión (**Figura 5-51**) observamos que la población *C11-G₁₂₄₈* sí presenta pilus en su superficie. Sin embargo, (Sokol, 2015) mutó una de las dos copias del gen *pilA*, obteniendo así copias silvestres y mutadas de PilA. La mutación afectaba al residuo 170 de la proteína codificada por el gen *Synpcc7942_0049*, que da lugar a un cambio de una alanina por un aspártico. Como consecuencia de la mutación, observaron que la presencia de estos dos tipos de proteína PilA impedía el ensamblaje del pilus. Estos autores especulan que la cercanía de esta mutación a un bucle D, el cual es esencial para el ensamblaje del pilus y su función, es responsable de dicho efecto. En nuestro caso ocurre lo contrario, a pesar de tener dos copias diferentes de la proteína

PilA el pilus es capaz de formarse, tal y como se ve en la **Figura 5-51**. Sin embargo, este pilus no es funcional ya que C11-G₁₂₄₈ no puede ser transformada naturalmente, hecho que puede observarse en la **Figura 5-50**.

En el caso de C12-G₁₁₆₈, la explicación de la pérdida de competencia natural es más compleja, por dos motivos. Por un lado, en la población C12-G₁₁₆₈ existen dos mutaciones en sendas proteínas que participan en la formación del pilus, PilC y PilT. Pero, por otro lado las mutaciones son SNPs (ver **Tabla 5-1**, **Figura 5-16** y **Figura 5-17**), siendo más difícil determinar el efecto que pueda tener dicha mutación sobre la actividad de la proteína. En ambos casos, ese SNP da lugar a un cambio de aminoácido, una valina por una glicina en la posición 85 de la proteína PilC y una arginina por una glicina en la posición 178 en el caso de PilT. En el caso de PilC, no pudimos realizar un modelado de la proteína con suficiente confianza para determinar si estaba afectando a su funcionalidad. Sin embargo, en PilT se conoce la existencia de cuatro dominios que están involucrados en la función de la proteína, no encontrándose la mutación en ninguno de ellos.

El residuo mutado en la proteína PilC está altamente conservado entre cianobacterias (**Figura 5-16**). Por tanto, podría ser el responsable de la pérdida de la competencia. En Ss6803, la disrupción de la proteína PilC mediante inserción del transposón Tn5, da lugar a pérdida de competencia (Yoshihara et al., 2001). El caso de la proteína PilT es similar a la anterior. La mutación también afecta a un residuo muy conservado en cianobacterias (**Figura 5-17**), pero sí que tiene un efecto claro en la distribución de la carga. Por ello, podría ser la responsable de la pérdida de la competencia natural. Para averiguar cuál de las mutaciones es la responsable de esta pérdida, se clonarían por separado y se analizaría la eficiencia de transformación natural. Quizás, el efecto no se deba a una de las mutaciones, sino al efecto conjunto de las dos. Esperamos clarificar esto con futuros experimentos.

La cepa reconstruida en el laboratorio con las mutaciones de C11-G₁₂₄₈ puede ser transformada naturalmente. Esto se debe a que las mutaciones que hemos introducido no están relacionadas con la transformabilidad natural. Por ello, la cepa reconstruida es una buena candidata para ser utilizada en ingeniería genética.

En Se7942 la formación de biofilm está causada por inactivación de las proteínas homólogas del sistema T2SS /pilus T4P y está asociada con la deficiencia de proteínas de secreción (Schatz et al., 2013). Estos autores determinaron que la inactivación de los genes Synpcc7942_2071 y Synpcc7942_2069, que codifican para PilB y PilC, respectivamente, da lugar a un fenotipo de formación de biofilm. Sin embargo, la inactivación del gen Synpcc7942_2070,

que codifica para PiIT, no forma biofilm, debido probablemente a la presencia en el genoma de Se7942 de dos genes adicionales que codifican para homólogos de PiIT (Synpcc7942_0847 y Synpcc7942_2349).

6.1.4. Efecto de las mutaciones sobre la expresión génica

En cuanto a los datos de transcriptómica obtenidos de las cepas evolucionadas, cabe destacar que: i) no se encontró ningún efecto a nivel global sobre la expresión de una categoría concreta de genes; ii) el número de genes cuya expresión génica difería significativamente entre las cepas evolucionadas (C11-G₁₂₄₈ y C12-G₁₁₆₈) y la silvestre era pequeño (29 para C11-G₁₂₄₈ y 5 para C12-G₁₁₆₈); iii) la expresión de dichos genes era generalmente inferior en las cepas evolucionadas.

Las mutaciones en los genes *aroK* (Synpcc7942_0894) y *sasA* (Synpcc7942_2114) tienen un efecto directo sobre el crecimiento de Se7942, como hemos observado (**Figura 5-35**, **Figura 5-36** y **Figura 5-37**). AroK mejora la tasa de crecimiento a alta intensidad de luz, mientras que a baja intensidad de luz se requieren ambas mutaciones para lograr un mayor crecimiento. La mutación en *aroK* se localiza en su región promotora. Análisis transcriptómicos preliminares indican que la expresión de este gen podría estar incrementada en C11-G₁₂₄₈. Sin embargo, necesitaríamos contar con más réplicas biológicas en dicho análisis para tener una confianza estadística que permita evaluar la variación en la expresión génica. Por otra parte, ya contamos con el mutante individual con este SNP (MSM3) y planeamos evaluar por RNAseq y/o RT-PCR el efecto de dicha mutación en la expresión de *aroK*. Este gen *aroK* codifica para una enzima de la ruta de síntesis del corismato, que es usado en cianobacterias para la síntesis de plastoquinona (Nowicka and Kruk, 2016; Pfaff et al., 2014). Esta molécula tiene un papel importante en la fotosíntesis, ya que transfiere los electrones del centro de reacción del fotosistema II al complejo citocromo *b₆f*. (Pfaff et al., 2014) observaron que, en Ss6803, al deletar varios genes que codifican para proteínas de la ruta de transformación del corismato en plastoquinona, se reducía el contenido de ésta. Esto tenía como consecuencia una reducida tasa de fotosíntesis, que daba lugar a una disminución en el crecimiento. En nuestro caso, podría estar ocurriendo lo contrario. Es decir, C11-G₁₂₄₈ produciría más plastoquinona, lo que aumentaría la tasa de fotosíntesis y, como consecuencia directa, disminuiría el tiempo de generación. En el futuro sería importante comparar cuantitativamente los niveles de plastoquinona presentes en las cepas mutantes con

respecto a la cepa silvestre, así como medir la tasa de fotosíntesis que posee de cada una de ellas.

El gen *sasA* codifica para una histidina kinasa, que junto a un regulador de respuesta (RpaA), forma un sistema regulatorio de dos componentes que ayuda a mantener el control circadiano de la expresión génica (Takai et al., 2006). En Se7942 se ha visto que la disrupción de *sasA* o *rpaA* da lugar a apagones en el ritmo circadiano o a arritmia, dependiendo de las condiciones de cultivo (Iwasaki et al., 2000; Takai et al., 2006). En nuestro caso, la delección en SasA se encuentra en la región C-terminal (**Figura 5-19**), por lo que podría afectar a la actividad del dominio histidina kinasa. En el modelado de *sasA* que se muestra en la **Figura 5-20**, se aprecia que en la proteína mutada desaparecen una hoja beta, una hélice alfa y un lazo, lo que probablemente conduzca a una pérdida de función de la proteína. Se ha descrito que un mutante *sasA*⁻ de Se7942 crece a la misma velocidad que la cepa silvestre en condiciones de luz continua, mientras que el crecimiento es menor en condiciones de luz-oscuridad (Takai et al., 2006). Dicho mutante no mantenía los ritmos circadianos en luz continua de alta intensidad, pero sí a baja luminosidad, en que reducía el período del ritmo (Iwasaki et al., 2000). Por otro lado, la delección de *sasA* no da lugar a un fenotipo de elongamiento celular, como ocurre con otros elementos del ritmo circadiano (Dong et al., 2010). Sin embargo, no existen estudios que correlacionen la delección de *sasA* con un aumento en la tasa de crecimiento de Se7942. Desafortunadamente, no disponemos de datos experimentales sobre el ritmo circadiano de la población C11-G₁₂₄₈, así como los derivados de Se7942 que contienen dicha mutación. Sería importante, en el futuro, medir el ritmo circadiano de dichas cepas, para dilucidar la relación entre un crecimiento más rápido y el ritmo circadiano.

Tanto en C11-G₁₂₄₈ como en C12-G₁₁₆₈ existen mutaciones que afectan a la expresión de proteínas hipotéticas cuyo efecto no tenemos en cuenta. También existen proteínas cuya expresión afecta directa o indirectamente a la síntesis y composición de pigmentos de la célula. Por ello, decidimos estudiar la composición de pigmentos de C11-G₁₂₄₈, así como de algunos derivados de Se7942 que contenían algunas de las mutaciones de las cepas evolucionadas.

Para ello, analizamos tanto el espectro de célula completa como la composición en µg/ml de cada uno de los pigmentos (**Figura 5-52** y **Figura 5-53**, respectivamente). Aunque los ratios de carotenoides/clorofila y ficobiliproteínas/clorofila calculados a partir del espectro de absorción fueron diferentes en las cepas MSM5 y MSM15, no se observó ningún cambio en la composición cuantitativa de cada uno de los pigmentos. Por otro lado, aunque en C11-G₁₂₄₈ ambos ratios fueron similares a los de la cepa silvestre, los picos de las ficobiliproteínas y el de

la clorofila a 680 nm estaban aumentados en el espectro de absorción en comparación con la cepa silvestre. Cuando cuantificamos cada uno de estos pigmentos, observamos también dicho aumento, por lo que ambos datos estarían correlacionados. Por tanto, alguna de las mutaciones presentes en C11-G₁₂₄₈ y que no hemos clonado en los mutantes que hemos construido está afectando a la composición de los pigmentos, en concreto a la clorofila y a las ficobiliproteínas (APC y PC).

6.2. Síntesis de ácido α -linolénico en Se7942

6.2.1. Ventajas y desventajas del uso de las cianobacterias para la producción de ácidos grasos

La producción microbiana de ácidos grasos ha recibido mucha atención en los últimos años como materia prima para la producción de energía renovable (Ruffing, 2014). Los principales esfuerzos se han centrado en la ingeniería genética de *E. coli*. Sin embargo, la producción de ácidos grasos en *E. coli* requiere de una fuente de carbono (glucosa, por ejemplo), que tiene que ser proporcionada en el medio de cultivo. En la actualidad, estas fuentes de carbono son prohibitivamente caras para ser utilizadas como materia prima para un producto de bajo valor como el combustible. La principal ventaja de usar microorganismos fotosintéticos sobre otras bacterias es su capacidad para usar la fotosíntesis para capturar energía de la luz solar y convertir el CO₂ en productos de interés. Además, las cianobacterias son fáciles de manipular genéticamente, y algunas de ellas han sido ya modificadas para la producción de ácidos grasos (ver apartado 2.3.2.3).

Una de las principales desventajas de utilizar cianobacterias frente a *E. coli*, ha sido su elevado tiempo de generación. En esta tesis demostramos cómo se puede reducir significativamente el tiempo de generación utilizando la evolución experimental. Además, recientemente (Yu et al., 2015) descubierto una cepa de *Synechococcus* de crecimiento rápido. Por tanto, las barreras en el uso de las cianobacterias, y en concreto de Se7942, como chasis para la producción de compuestos de interés industrial están disminuyendo.

6.2.2. Selección del organismo donador de los genes para la producción de ácido α -linolénico

Existen varias cianobacterias capaces de producir ALA de forma natural, entre las que destacan Ss6803 y Ss7002. Como hemos descrito en el apartado 2.3.2.1, la síntesis de ácidos grasos insaturados en cianobacterias es llevada a cabo por desaturasas. Se7942 solo posee la enzima DesC que introduce un doble enlace en la posición $\Delta 9$. Para producir ALA necesita las enzimas DesA y DesB a fin de introducir las dos restantes insaturaciones.

Los genes de Ss6803 ya habían sido utilizados anteriormente en Se7942 para la producción de ALA, obteniéndose una producción máxima del 5 % (Sakamoto et al., 1994). Para mejorar esta eficiencia, nosotros usamos los de Ss7002, que ofrecen ciertas ventajas. La primera es que existe gran cantidad de información sobre dichos genes, en cuanto a condiciones que favorecen una alta expresión (Ludwig and Bryant, 2012; Sakamoto et al., 1997), como se muestra en la **Tabla 5-6**. Por otro lado, los genes de Ss7002 son más eficientes en la conversión de los ácidos grasos C18 en ALA que los de Ss6803, siendo la producción máxima en Ss7002 de un 20 % (Sakamoto et al., 1997) frente al 2,53 % que produce Ss6803 (Chen et al., 2014). Además, el uso de codones en Se7942 y Ss7002 es muy similar, con un promedio de diferencia del 0,418 en el porcentaje de uso de cada codón (Ruffing, 2014), por lo que no es necesario optimizar el uso de codones de los genes a clonar.

Cabe destacar que el uso a nivel industrial de Ss7002 para producir compuestos de interés está limitado por el coste de su medio de crecimiento y no por su tiempo de generación, que es relativamente corto. Esta cianobacteria necesita vitamina B12 para su crecimiento, ya que presenta una auxotrofia natural para la cobalamina (Wilhelm and Trick, 1995). Para un proceso industrial a gran escala, sería excesivamente caro añadir dicha vitamina al medio de cultivo. Recientemente, (Perez et al., 2016) han identificado el origen de esta auxotrofia, que radica en el uso de una metionina sintasa dependiente de cobalamina (MetH) en el último paso de la síntesis de L-metionina. La expresión heteróloga del gen *metE*, que codifica para una metionina sintasa independiente de cobalamina, de *Synechococcus* sp. PCC 73109 ha conseguido mitigar esta auxotrofia y ha posibilitado la construcción de un Ss7002 más apropiado para futuras aplicaciones industriales a gran escala.

6.2.3. Selección del promotor adecuado para la expresión de los genes de producción de ácido α -linolénico

Una vez determinado de qué organismo íbamos a tomar los genes para producir ALA en Se7942, el siguiente paso fue decidir bajo qué promotor íbamos a clonarlos. Varios estudios han determinado que la desaturasa DesB únicamente se transcribe a 22 °C (Ludwig and Bryant, 2012; Sakamoto and Bryant, 1997). Dado que Se7942 no crece a temperaturas por debajo de 25 °C (**Figura 6-3**) y que el gen de la desaturasa DesB necesita bajas temperaturas para transcribirse, fue necesario acoplarle un promotor inducible. Así podríamos crecer los cultivos de Se7942 a 30 °C y únicamente bajaríamos la temperatura a la hora de inducir el promotor. De esta manera,

aunque el cultivo no podría continuar creciendo sería capaz de expresar la proteína de interés. Por otro lado, al depender la transcripción de uno de los genes de la temperatura, tendríamos un sistema controlable, ya que la producción de ALA solo se llevaría a cabo durante el tiempo que tuviéramos los cultivos a baja temperatura.

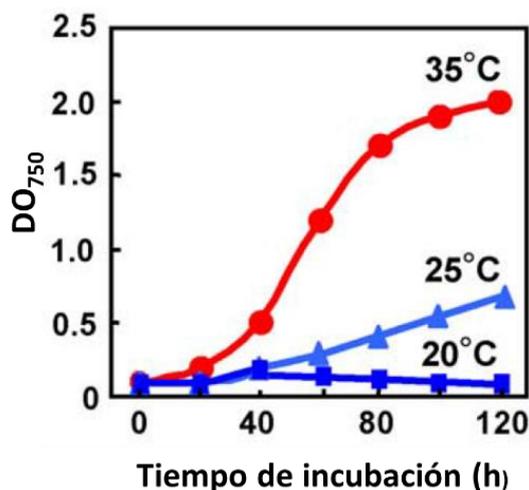


Figura 6-3. Curvas de crecimiento de Se7942 a diferentes temperaturas. Modificado de (Los and Murata, 2004).

Por ello, probamos dos promotores inducibles disponibles en el laboratorio, *PnrsB* y *Ptrc*. *PnrsB* regula la expresión del operón *nrsBACD* en Ss6803, es inducido por iones de Ni^{2+} y exhibe un bajo "leakiness" (Lopez-Maury et al., 2002). *Ptrc* es un híbrido entre los promotores *trp* y *lacUV5* derivados de *E. coli* y es inducible por IPTG (Amann et al., 1988). Como se observa en la **Figura 5-55**, el promotor *Ptrc* es casi 10 veces más fuerte que *PnrsB*. De ahí que se produzca 3,5 veces más ALA cuando se expresan las desaturasas con el promotor *Ptrc* (**Figura 5-56**) en comparación con *PnrsB* (ver **Tabla 5-5**).

La cantidad de ALA producida por Se7942 utilizando las desaturasas de Ss7002 fue más alta que la producida usando las de Ss6803 (Sakamoto et al., 1994). Las enzimas de ambas cianobacterias fueron funcionales en Se7942, produciendo un 7,5 % de ALA con las de Ss7002, frente al 5 % de ALA producido con las de Ss6803 (Sakamoto et al., 1994). Esta diferencia podría deberse a las que las desaturasas de Ss7002 son más eficientes en la producción de ALA que las de Ss6803. También podría deberse a las diferencias en el nivel de expresión de las desaturasas, ya que (Sakamoto et al., 1994) usaron el promotor nativo, mientras que en nuestro caso hemos usado un promotor inducible muy fuerte (*Ptrc*).

6.2.4. Estudio del efecto de diferentes condiciones en la producción de ácido α -linolénico

Además de la baja temperatura, probamos otra de las condiciones que favorecía la expresión de estas desaturasas y por ende, la producción de ALA. (Sakamoto et al., 1997) demuestran que una alta intensidad de luz da lugar a un aumento en la producción de ALA. Un hecho similar fue observado en un análisis de microarray realizado en Ss6803 por (Hihara et al., 2001). Comparamos el efecto conjunto de una alta intensidad de luz y una baja temperatura, tanto en la cepa Ss7002 como en la cepa MSM26 (*Ptrc::desA-desB*), tal y como se observa en la **Figura 5-57**. En el caso de la cepa Ss7002, una alta intensidad de luz favoreció la producción de ALA, sin embargo, en la cepa MSM26 esta condición no tuvo efecto.

El hecho de que la temperatura esté afectando la producción de ALA en la cepa que sobreexpresa las desaturasas de Ss7002, mientras que la luz no lo hace, puede deberse al mecanismo de regulación sobre la expresión génica que se esté dando en cada una de las condiciones. (Sakamoto and Bryant, 1997) sugieren que el transcrito del gen *desB* es más estable a baja temperatura, asociando este efecto con la estructura del ARNm. Los cambios dependientes de temperatura en la estructura secundaria del ARNm han sido descritos como elementos de control del RBS e inicio de la traducción, y estos elementos han sido posteriormente denominados como termómetros de ARN (Kortmann and Narberhaus, 2012). Basándonos en este tipo de regulación, en Ss7002 a bajas temperaturas podría conseguirse una estabilización de los transcritos de *desB*, que daría lugar a una acumulación de ARNm. A altas temperaturas, la estructura secundaria de este ARNm podría cambiar, lo que provocaría una degradación rápida, y por ello, se observarían bajos niveles de estos transcritos (Sakamoto and Bryant, 1997; Sakamoto et al., 1997). Como consecuencia, la cantidad de proteína que habría en la célula sería menor, y la cantidad de ALA producido disminuiría considerablemente.

Por otro lado, el efecto de la alta intensidad de luz no se repite en ambos organismos, Se7942 y Ss7002. Esto indicaría que la regulación no se está dando a nivel del ARNm, como en el caso de la baja temperatura. El efecto podría deberse a un regulador que se activaría en condiciones de alta intensidad de luz, se uniría al promotor y daría lugar a la transcripción del gen. Sin embargo, este posible regulador no estaría presente en Se7942, de ahí que el efecto visto en Ss7002 no se replique en Se7942. A este tipo de regulación se le conoce como regulación positiva, cuyo mecanismo de acción se muestra en la **Figura 6-4**. En cianobacterias, los sistemas de regulación de la expresión de genes inducibles por estrés están mediados por sistemas

Hik/Rre (Histidina kinasa/Regulador de respuesta) (Los et al., 2010). Sin embargo, poco se conoce de la regulación de la expresión de los genes involucrados en la síntesis de ácidos grasos, por ello es difícil establecer el mecanismo por el cual una alta intensidad de luz (una condición de estrés) estaría afectando a la expresión de las desaturasas.

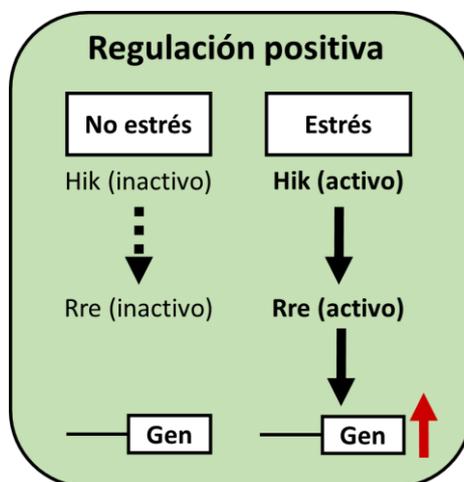


Figura 6-4. Representación esquemática de un sistema de regulación positiva de la expresión de genes inducibles por estrés. La histidina kinasa (Hik) se encuentra inactiva en ausencia de estrés y, como resultado, el correspondiente Regulador de respuesta (Rre) está inactivo. En células sujetas a estrés, Hik es activada por fosforilación y la señal es transferida al Rre relacionado, el cual aumenta la expresión de los genes silenciados bajo condiciones no estresantes. Las flechas negras continuas representan las señales que activan componentes aguas abajo y la flecha negra punteada indica su ausencia. La flecha roja corresponde al incremento de la expresión génica. Adaptado de (Los et al., 2010).

6.2.5. Efecto de la modificación de la expresión de los genes de síntesis y degradación de ácidos grasos en Se7942

Existen varios factores que juegan un papel determinante en los niveles de ALA en la cepa de Se7942 mejorada genéticamente. El primero de ellos, es la cantidad de desaturasas producidas para la síntesis de ALA, cuyo efecto ya hemos discutido en el apartado 6.2.3. Por otro lado, la cantidad y disponibilidad de los sustratos utilizados por estas desaturasas también determina la cantidad de ALA producida.

En el caso concreto de la cepa MSM26 (*P_{trc}::desA-desB*), cuyo perfil de ácidos grasos se muestra en la **Tabla 5-5**, observamos que la producción de ALA consume mayoritariamente los ácidos grasos C18, e incluso parte de los ácidos grasos C16:1. Sin embargo, esta cepa no es capaz de utilizar los ácidos grasos C16:0. Es por ello que decidimos identificar aquellas enzimas que se encargan de transformar los ácidos grasos C16:0 en C16:1 y C18. Para ello, es necesario definir cómo tiene lugar la síntesis de ácidos grasos en cianobacterias, ya que hasta el momento se desconoce.

Aunque las cianobacterias tienen un perfil prometedor para la producción de ácidos grasos insaturados, el principal problema es el limitado conocimiento existente acerca de las enzimas involucradas en la ruta de síntesis de ácidos grasos. Recientemente, (Kuo and Khosla, 2014) determinaron que las enzimas de esta ruta biosintética en Ss7002 tienen la misma función que sus homólogos en *E. coli*. Por esta razón, usamos los datos de *E. coli*, y probamos la función de cinco genes mediante su sobreexpresión o delección. Con solo tres de ellos (delección parcial de *fabH*, sobreexpresión de *fabF* y delección de *fadD*), obtuvimos el mismo perfil de ácidos grasos al observado en *E. coli*. Sin embargo, nuestros datos contradicen lo descrito por (Kuo and Khosla, 2014). Ellos introdujeron los genes de Ss7002 en *E. coli*, y vieron que su función era la misma, hecho que puede deberse a que los genes de Ss7002 estaban regulados de la misma manera que son regulados en *E. coli*. En Se7942 se desconoce totalmente la regulación de los genes implicados en la síntesis de ácidos grasos. Es por ello que no podemos inferir que se está modificando la regulación al sobreexpresar y/o deleccionar estos genes de síntesis de ácidos grasos.

Otro hecho que pone de manifiesto la diferente regulación de los genes implicados en la síntesis de ácidos grasos en *E. coli* y Se7942, es la incapacidad para deleccionar el gen *fabH* en este último. En el caso de *E. coli*, en primera instancia se pensó que este gen era letal (Lai and Cronan, 2003), ya que estos autores no fueron capaces de obtener mutantes deficientes en dicho gen. Sin embargo, años más tarde, (Yao et al., 2012) construyeron una cepa de *E. coli* que carecía del gen *fabH*. Este mutante presenta un reducido tamaño celular, así como una tasa de crecimiento menor a la de la cepa silvestre. (Rubin et al., 2015) describen a este gen como un gen esencial para Se7942, es por ello que cuando tratamos de eliminarlo obtenemos un mutante con una delección parcial (**Figura 5-59**). Solo fuimos capaces de eliminar la copia del gen *fabH* del cromosoma de Se7942 cuando introducimos una copia extra en otro lugar del cromosoma bajo un promotor fuerte (**Figura 5-62**). Esto puede deberse a que este gen presente unos niveles altos de expresión, que solamente se pueden suplir cuando expresamos la proteína bajo un promotor fuerte. Por otro lado, tal y como describen (Kuo and Khosla, 2014), FabH sería el factor limitante en la síntesis de ácidos grasos en Ss7002, hecho que podría estar repitiéndose en Se7942. Recientemente, (Gu et al., 2016) no consiguieron deleccionar este mismo gen en Ss7002, demostrando que también es esencial para la viabilidad de esta cianobacteria.

En *Lactococcus lactis* (Morgan-Kiss and Cronan, 2008), se ha visto que la proteína FabF puede suplir la función de FabH en la síntesis de ácidos grasos. Sin embargo, en Se7942 hay que destacar la imposibilidad de deleccionar *fabH* cuando sobreexpresamos *fabF*, indicando que la función que cumplen estas enzimas en las células no es la misma. El fallo de FabF en el

reemplazamiento de FabH es debido a la distinta afinidad de sustrato que presentan estas enzimas, siendo FabF activa con acetyl-ACP y no con acetyl-CoA, sustrato habitual de FabH (Jackowski et al., 2002; Tsay et al., 1992).

FabB es una β -oxoacil sintasa I que cataliza la condensación del acil-ACP con otro malonil-CoA, en la ruta cíclica de síntesis de ácidos grasos (Edwards et al., 1997) (ver Introducción, apartado 2.3.1.1). La delección del gen *fabB* en *E. coli* da lugar a un mutante auxótrofo para ácidos grasos insaturados (Cronan et al., 1969) y la sobreexpresión de este gen aumenta la cantidad de C18:1 (Clark et al., 1983). Sin embargo, en Se7942 la sobreexpresión de FabB procedente de Ss7002 no tuvo ningún efecto en la composición de ácidos grasos. Esto guarda relación con lo descrito por (Kuo and Khosla, 2014). Estos autores determinaron que FabB no tiene una función en la síntesis de ácidos grasos y especulan que esta proteína podría tener un papel adicional en los organismos de crecimientos rápido, aún sin descubrir.

Un hecho que tiene especial relevancia, obtenido con la sobreexpresión de *fabH*, fue la síntesis de ácido láurico (C12:0), ácido graso que no está presente de forma natural en Se7942. Este mutante (MSM28) también fue capaz de aumentar la producción de ácido mirístico, C14:0. Ambos son ácidos grasos de cadena media saturados, incluyéndose dentro de esta categoría ácidos grasos que contienen entre 6 y 14 átomos de carbono. Entre sus principales aplicaciones destaca, su uso en la fabricación de detergentes y lubricantes debido a su baja viscosidad y a su capacidad para formar espumas. Tras la esterificación con alcoholes, estos ácidos grasos de cadena media pueden ser usados como emulsificadores para comida y cosméticos (Froissard et al., 2015). Cabe destacar también sus valiosas propiedades como precursores de la síntesis de biodiesel, ya que los metil ésteres de ácidos grasos de cadena media mejoran la calidad del combustible (Gu et al., 2016). Actualmente, las principales fuentes de este tipo de ácidos grasos son el aceite de coco y el de palma. Junto a estos, las algas y las diatomeas han sido descritas como buenos sistemas celulares para sintetizar ácidos grasos de 10 a 14 átomos de carbono (Froissard et al., 2015; Gu et al., 2016). Por otro lado, (Gu et al., 2016) han demostrado que las cianobacterias no son buenos productores de estos ácidos grasos de cadena media, como consecuencia de la diferente actividad de FabH sobre ácidos grasos de diferente longitud de cadena. En Se7942, esta enzima tiene afinidad por ácidos grasos de cadena media, es por ello que el sobreexpresar *fabH*, conseguimos aumentar la producción de este ácido graso. Aunque en esta tesis no nos hemos centrado en la producción de biodiesel, podría ser una aplicación futura de este descubrimiento.

Por otro lado, aunque fueron tres las modificaciones de genes implicados en la síntesis de ácidos grasos las que tuvieron un efecto positivo sobre la producción de intermediarios C18, solo dos de ellos mejoraron los niveles de ALA producidos. Tanto la sobreexpresión de *fabF* como la delección de *fadD* junto a la sobreexpresión de las desaturasas permitieron alcanzar altos niveles de ALA. Sin embargo, la delección de parcial del gen *fabH* tuvo el efecto contrario (**Figura 5-36**). Estos resultados amplían nuestros conocimientos actuales sobre el papel de estos genes en la ruta de síntesis de ácidos grasos, así como de los mecanismos moleculares que son capaces de mejorar la síntesis de ácidos grasos poliinsaturados usando cianobacterias.

En Ss7002 el nivel más alto de ALA descrito en bibliografía fue de un 19 % en la cepa silvestre, en condiciones de alta intensidad de luz y baja temperatura (Sakamoto et al., 1997). Por otro lado, (Chen et al., 2014) obtuvieron una cepa de Ss6803 capaz de producir en torno a un 23 % de ALA mediante la sobreexpresión de las desaturasas endógenas $\Delta 12$ y $\Delta 15$. La importancia del trabajo realizado en esta tesis reside en la obtención de los niveles más altos de ALA descritos hasta el momento. El nivel logrado de Se7942 genéticamente modificado fue de un 22,28 %. En futuros trabajos de investigación, sería interesante la combinación de todas estas aproximaciones en Ss6803 o Ss7002, que podría permitir incrementar aún más los niveles de ALA producidos hasta el momento en cianobacterias.

Este es el primer trabajo realizado en cianobacterias que muestra que la producción de ALA puede incrementarse usando modificaciones en las enzimas involucradas en la ruta de síntesis de ácidos grasos. Los genes que codifican para las enzimas que participan en esta ruta están bien caracterizados. Sin embargo, no existe apenas información sobre la actividad y estructura de estas enzimas. El esquema de la posible ruta de síntesis de ácidos grasos que tiene lugar en cianobacterias se muestra en la **Figura 6-5**. Este estudio pone los cimientos para el aumento del contenido de ácidos grasos, así como para la producción de grandes cantidades de ácidos grasos poliinsaturados en cianobacterias. Además, esta cianobacteria modificada podría incluso usarse para la producción de suplementos nutricionales de consumo humano o animal y biodiesel.

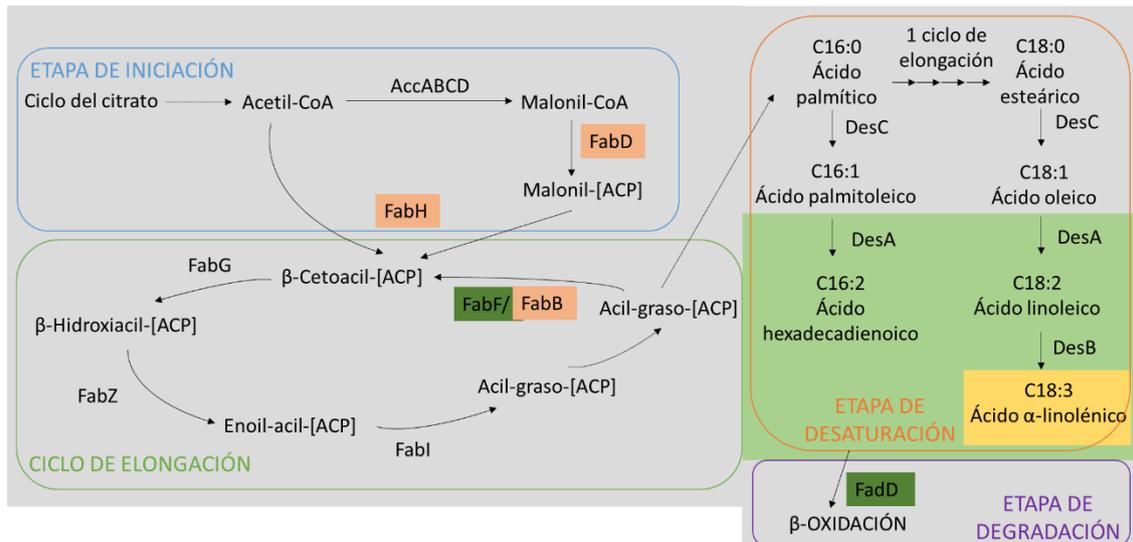


Figura 6-5. Esquema de la posible ruta de síntesis de ácidos grasos en cianobacterias. La ruta de síntesis de ácidos grasos en cianobacterias consta de tres etapas: iniciación, elongación y desaturación. Existe también una etapa de degradación de la que apenas se tiene información. En gris, se muestran las reacciones que forman parte de la ruta nativa presente en Se7942. En verde claro, se muestran las enzimas que fueron necesarias clonar para la producción de ALA. En amarillo, se muestra al ALA. En verde oscuro, se muestran las enzimas de la ruta en las que la expresión del gen que codifica para la misma fue modificada y tuvieron un efecto positivo en la síntesis de ALA. En naranja, se muestran las enzimas de la ruta en las que la expresión del gen que codifica para la misma fue modificada y no tuvieron un efecto positivo en la síntesis de ALA.

6.3. Modelo metabólico de Se7942

6.3.1. Reconstrucción del modelo metabólico de Se7942

En este estudio, hemos reconstruido el modelo metabólico a escala genómica de Se7942, basándonos en el modelo *iJN678* de Ss6803 desarrollado por (Nogales et al., 2012). Se utilizó el modelo de esta cianobacteria, ya que ambas cianobacterias presentan una alta homología en cuanto a composición génica se refiere. Al comparar la secuencia de ambos genomas utilizando la herramienta OrthoANI (Lee et al., 2015), obtenemos un porcentaje de identidad entre ambas secuencias, del 67 %. Este alto porcentaje de identidad nos indica que ambas cianobacterias comparten muchos genes y por ello, muchas actividades metabólicas comunes. El hecho de tener un organismo tan parecido a Se7942, cuyo modelo metabólico ya está desarrollado, fue una ventaja a la hora de construir el modelo de nuestra cianobacteria de estudio, ya que se redujo el tiempo empleado en la anotación de reacciones y sus correspondientes metabolitos. Solo hubo que anotar aquellas reacciones específicas de nuestra cianobacteria, y no las reacciones del núcleo del metabolismo de las cianobacterias.

Uno de los principales problemas a la hora de realizar el modelo metabólico de Se7942 fue el alto porcentaje de genes con función desconocida presentes en su genoma. El porcentaje de este tipo asciende al 51 %, lo que implica que más de la mitad de las proteínas codificadas por el genoma de Se7942 no pueden incluirse en el modelo por desconocerse su función. Sin embargo, con las diferentes bases existentes para analizar homología con proteínas del mismo tipo o de otras cianobacterias, e incluso con bibliografía publicada sobre alguna de estas proteínas, ese porcentaje puede reducirse. En este trabajo, se han utilizado especialmente dos bases de datos. La primera de ellas, “CyanoBase Similarity Search” nos permitió comparar nuestra proteína de función desconocida usando una búsqueda proteica en otras cianobacterias (Fujisawa et al., 2017). La segunda de ellas fue “The Transporter Classification Database (TCDB)” que nos permitió hacer una búsqueda más exhaustiva cuando la proteína hipotética se trataba de un transportador, ya que compara la secuencia del posible transportador con toda una base de datos de transportadores (Saier et al., 2016).

Por otro lado, es necesario mencionar que el modelo que se describe y presenta en esta tesis, *iMS705*, se inició cuando solo estaba publicado el modelo de Se7942, *iSyf715* (Triana et al., 2014). Durante la construcción de *iMS705* fue publicado el modelo *iJB785* de Se7942

(Broddrick et al., 2016). En la **Figura 5-80**, se muestra un análisis comparativo del número de genes presentes en estos tres modelos. Todos ellos tienen en común 565 genes, un número muy alto, teniendo en cuenta el número total de genes que presenta cada modelo (*iMS705*, 705 genes; *iMS785*, 785 genes; *iSyf715*, 714 genes). Por otro lado, hay que destacar que existen genes presentes específicamente en cada modelo. Es por ello, que decidimos obtener una lista de los genes no comunes entre *iMS705* e *iJB785*. Mediante la adición de parte de los genes de esta lista al modelo *iJB785*, así como la modificación de reacciones mal anotadas en el mismo, diseñamos un modelo más preciso, al que denominamos *iMS835*. Este modelo fue el utilizado para analizar la producción de ALA *in silico*.

6.3.2. Capacidad predictiva de genes esenciales del modelo metabólico de Se7942

La validación de los dos modelos desarrollados en esta tesis fue realizada mediante su capacidad para predecir el impacto de la delección de los genes sobre el crecimiento de Se7942. Los resultados obtenidos *in silico* fueron comparados con resultados *in vivo* obtenidos previamente por (Rubin et al., 2015). La predicción del fenotipo de crecimiento de las delecciones génicas es un test canónico de la precisión de los modelos metabólicos y un punto de referencia útil para validar la calidad de los mismos (Becker and Palsson, 2008). La precisión de nuestros modelos fue muy alta, un 79 % para *iMS835* y un 71 % para *iMS705*, siendo la del modelo *iMS835* incluso un poco mayor a la del modelo publicado recientemente *iJB785*, la cual asciende a un 78 % (ver **Tabla 5-10**). Sin embargo, en nuestros modelos este valor aumenta debido al gran número de genes no esenciales que son correctamente predichos. A pesar de que estas tasas de éxito se han calculado de forma correcta, la predicción correcta de genes no esenciales no tiene el mismo peso que la correcta predicción de genes esenciales desde un punto de vista biológico (Becker and Palsson, 2008). En los casos de falsos positivos, genes no esenciales predichos como esenciales por el modelo, un experimento de delección del gen en el laboratorio puede verificar que la predicción es errónea. Sin embargo, en los casos de falsos negativos, genes esenciales predichos como no esenciales por el modelo, es necesario una amplia colección de experimentos para localizar dichos errores.

Se cree que las predicciones erróneas normalmente ocurren por varias razones. Los falsos negativos pueden estar causados por una incompleta definición de la función de biomasa, una

incorrecta definición del medio de crecimiento usado en los experimentos, y/o una acumulación de intermediarios tóxicos. Por otro lado, los falsos positivos pueden estar causados por una rigurosa definición de la función de biomasa, una incorrecta definición en el medio de crecimiento, y/o la presencia de isoenzimas desconocidas en una reacción dada. La función de biomasa es fundamental en la simulación de delección de genes, ya que un gen es predicho como esencial si su delección resulta en una completa deficiencia de flujos hacia una reacción específica. El medio de crecimiento usado para los experimentos es también muy importante porque la esencialidad de los genes es dependiente de cuantos sustratos estén disponibles para usarse. La acumulación de intermediarios tóxicos es difícil de simular con exactitud, ya que en muchos casos se desconoce si la célula es capaz de producir un metabolismo y muchos menos si es capaz de degradarlo. La presencia de isoenzimas desconocidas sugiere que el organismo no está estudiado tan bien como podría estarlo (Becker and Palsson, 2008).

El modelo *iJB785* es capaz de predecir mejor los genes esenciales que nuestro modelo *iMS835*, un 88 % frente a un 84 %. Sin embargo, a nivel de número de discordancias (falsos positivos y falsos negativos), nuestro modelo *iMS835* disminuyó este valor en gran medida respecto a *iJB785*, tal y como se puede ver en la **Figura 5-81**. Es por este motivo, que decidimos usar el modelo *iMS835* para identificar posibles dianas para mejorar la producción de ALA en *Se7942*. La alta consistencia entre las simulaciones y los resultados experimentales indican la utilidad del modelo *iMS835* para simulaciones metabólicas. Otro de las razones para utilizar este modelo es que deriva del modelo *iJB785*, que tiene como novedad el modelado de los fotones de luz como si de un metabolito se tratase (Broddrick et al., 2016). De esta manera, puede incluirse en el modelo la cantidad, calidad y la tonalidad de la luz utilizadas para el crecimiento de *Se7942*, y así mimetizar con más exactitud las condiciones en las que crece esta cianobacteria en el laboratorio.

6.3.3. Análisis *in silico* de las modificaciones probadas experimentalmente en *Se7942* para mejorar la producción de ácido α -linolénico

De las modificaciones probadas *in vivo* e *in silico*, la única que no permitió un aumento en la producción de ALA, respecto a la cepa que solo contiene las dos desaturasas, fue la delección parcial del gen *fabH*. El modelo nos ha permitido encontrar una explicación científica de porque esta modificación no permite conseguir dicho objetivo. Al disminuir el flujo de la reacción catalizada por FabH, KAS15, estamos disminuyendo la probabilidad de una mayor producción

de ALA, tal y como se observa en la **Figura 5-84**. Es por ello que la estrategia a abordar en futuro con este gen sería su sobreexpresión, con el fin aumentar la producción de ALA. Al sobreexpresar este gen estamos aumentando el flujo de la reacción KAS15, y como consecuencia de ello, estamos aumentando el flujo hacia la producción de ALA. En este caso, nos estaríamos moviendo en la zona representado en rojo en la **Figura 5-84**. Sin embargo, la sobreexpresión de *fabH in silico* tiene un efecto negativo sobre el crecimiento, por lo que sería muy importante controlar dichos niveles de sobreexpresión para no afectar negativamente al crecimiento de Se7942.

En cuanto a la sobreexpresión de *fabF*, el efecto visto *in silico* es igual al efecto observado *in vivo*, de tal manera a que mayor flujo de las reacciones catalizadas por FabF tiene lugar una mayor producción de ALA (ver **Figura 5-87**). Además, como en el caso de la sobreexpresión de FabH, la sobreexpresión de FabF también tiene un efecto negativo sobre el crecimiento, tal y como se observa en la **Figura 5-86**. Sin embargo, este hecho no fue observado en los mutantes *in vivo*, que no presentan alteraciones en el crecimiento respecto a la cepa silvestre.

Por último, aunque la delección de FadD *in vivo* tuvo un efecto en el incremento de la producción de ALA, en el modelo no ocurrió tal cosa. Como se observa en la **Figura 5-89**, incluso cuando el flujo de las reacciones catalizadas por FadD es 0, la producción de ALA es máxima. Esto puede deberse a que la ruta de degradación de ácidos grasos en el modelo está ausente, ya que no existen datos *in vivo* que nos hayan permitido incluirla. Es por ello que estaríamos obviando información que estaría afectando a nuestras simulaciones.

6.3.4. Identificación *in silico* de nuevos “knockouts” para mejorar la producción de ácido α -linolénico

Utilizando el modelo iMS835 modificado para la producción de ALA, conseguimos identificar *in silico* posibles dianas para mejorar la producción de este ácido graso. Estas podrían probarse en un futuro *in vivo* para corroborar que verdaderamente, tienen un efecto positivo en la producción de ALA. Las reacciones catalizadas por todos estos genes, salvo la reacción PDH y GART, forman parte del metabolismo de aminoácidos. La reacción PDH pertenece al metabolismo del piruvato y la reacción GART, al de purinas. Todas ellas salvo la reacción GART usan acetil-CoA como sustrato para sintetizar productos finales o intermediarios de la ruta en la que participan. Este mismo metabolito es usado por Se7942 para iniciar la ruta de síntesis de ácidos grasos. Si bloqueamos totalmente o parcialmente estas rutas con la delección de uno o

varios genes cuyos productos proteicos catalizan participan en dichas rutas, evitaremos la competición por el acetyl-CoA. De esta manera, desviaríamos hacia la síntesis de ácidos grasos el sustrato no utilizado por las rutas bloqueadas, con el objetivo de usarlo en la producción de ALA. En el caso la reacción GART, desconocemos cual puede ser el mecanismo por el cual la delección del gen que codifica para la enzima que cataliza esta reacción está favoreciendo la síntesis de ALA.

Por un lado, hay que destacar que todas las delecciones que aumentaron la producción de ALA tuvieron un efecto negativo en el crecimiento de Se7942 (ver **Tabla 5-11** y **Figura 5-91**). Por este motivo, sería interesante ver si en vez de deleccionarse totalmente el gen que codifica para las enzimas que catalizan esta reacción, podría disminuirse el flujo de la misma. De esta manera, podría aumentarse la producción de ALA sin afectar al crecimiento.

Por otro lado, hay que destacar que la relación de todos los mutantes *in silico* con la producción de ALA fue un equilibrio de tipo sencillo (**Figura 6-6, A**). Como se observa en dicha gráfica a mayor tasa crecimiento del organismo, menor producción del compuesto de interés y a la inversa. Este tipo de relación se da cuando la producción de un determinado compuesto no está asociada al crecimiento del organismo que lo produce. Sin embargo, para la producción a escala industrial se precisa un tipo de producción acoplada al crecimiento (**Figura 6-6, B**). En este tipo de sistema existe siempre producción, ya que al depender del crecimiento, el organismo tiene que producir dicho compuesto para crecer. En este caso no existe una relación lineal entre la tasa de crecimiento y la producción, puesto que no todos los organismos que componen una población presentan el mismo nivel de producción. De esta manera, disponemos de un espacio de soluciones que incluye todos los niveles de producción, así como de crecimiento, de los individuos que componen dicha población. La producción acoplada al crecimiento resuelve problemas de estabilidad génica, y permite seleccionar y optimizar organismos para crecer, mucho más fácil que seleccionar organismos para producción (Knoop and Steuer, 2015). En este tipo de sistema de producción, para que la célula logre un crecimiento más rápido, necesita producir el compuesto de interés. Esto permite que la población evolucione de forma natural hacia mayores tasas de producción y como consecuencia, hacia un crecimiento más rápido. De este modo, las cepas o mutantes que producen la molécula diana se pueden seleccionar fácilmente seleccionando las cepas con el crecimiento más rápido. Las mutaciones que tuvieran como resultado una menor producción, darían lugar a una disminución de la tasa y, por ello, acabarían perdiéndose en las poblaciones (Feist et al., 2010). Recientemente, (Erdrich et al., 2014) han podido demostrar que la síntesis de productos puede acoplarse al crecimiento en cianobacterias. Estos autores consiguieron tal efecto, reduciendo la proporción de ATP y NADPH

producido por la cadena de transporte de electrones fotosintéticos, mediante diferentes mutantes “knockouts”. Con ello bloquearon los flujos alternativos para la obtención de estas moléculas con poder reductor, para que el organismo utilizara la síntesis del compuesto de interés para obtener este poder reductor. Hay que destacar, comparado con organismos heterotróficos, que este acoplamiento requiere fundamentalmente el uso de diferentes estrategias que implican un alto número de modificaciones metabólicas. Además, nuestro caso es aún más complejo, ya que DesB, una de las enzimas involucradas en la síntesis de ALA, depende de una baja temperatura para su actividad. Es por ello, que sería interesante en un futuro modificar el gen que codifica para esta proteína, para que pueda ser estable a la temperatura habitual de crecimiento de Se7942.

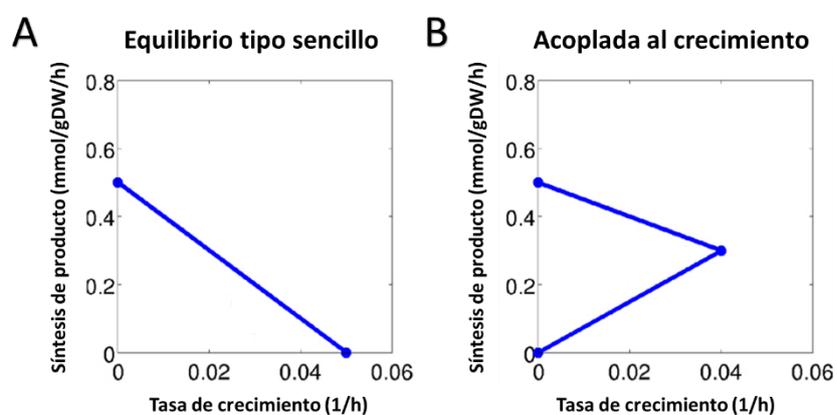


Figura 6-6. Tipos de relación entre la síntesis de un determinado producto y el crecimiento de un organismo. (A) Equilibrio de tipo sencillo. **(B)** Síntesis de un producto acoplada al crecimiento. La línea azul delimita el espacio de soluciones que se puede dar para cada uno de los tipos de producción. Los puntos azules representan los vértices de ese espacio de soluciones. Modificado de (Knoop and Steuer, 2015).

Conclusiones

7. Conclusiones

1. Durante el experimento de evolución realizado en la cianobacteria Se7942 se obtuvieron dos poblaciones evolucionadas, C11-G₁₂₄₈ y C12-G₁₁₆₈, que presentaban menores tiempo de generación que la cepa silvestre (3,13 h y 2,95 h, respectivamente, respecto a 8,6 h) y contenían cuatro y siete mutaciones respecto a la cepa de partida.
2. A alta intensidad solo era necesaria la mutación en el gen *aroK*, mientras a que a baja intensidad fueron necesarias dos mutaciones (*aroK* y *sasA*).
3. El tipo de luz utilizada para el crecimiento de Se7942, luces fluorescentes o luces LEDs, no afectó al tiempo de generación. Sin embargo, la intensidad de luz tuvo un efecto significativo.
4. Las poblaciones evolucionadas C11-G₁₂₄₈ y C12-G₁₁₆₈ han perdido su competencia natural debido mutaciones que afectan a componentes del pilus tipo IV.
5. El promotor que controla la sobreexpresión de *desA* y *desB* en Se7942 tiene un efecto directo sobre la cantidad de ALA producido. Así, el promotor *P_{trc}*, más fuerte, produce más ALA que *P_{nrsB}*.
6. *fabH* es un gen esencial de Se7942 y solo puede delecionarse cuando se introduce una copia extra en el cromosoma bajo un promotor fuerte.
7. La sobreexpresión de *fabF* no permite la delección completa del gen *fabH*, por lo que la función de ambas enzimas no es intercambiable.
8. La sobreexpresión de *fabH* da lugar a la producción de ácido láurico, ácido graso que no está presente de forma natural en Se7942. Por otro lado, provoca un aumento en la producción de ácido mirístico.
9. La clonación del gen *fabB* de Ss7002 en Se7942 no tuvo efecto en la composición de ácidos grasos. Esto corrobora la hipótesis de que este gen no cataliza ninguna reacción de la síntesis de ácidos grasos en Ss7002, pero tiene alguna otra acción enzimática relacionada directamente con el crecimiento rápido.

10. La sobreexpresión de *fabF*, la delección parcial de *fabH* o la delección total de *fadD* aumentan la producción de ácidos grasos C18. Sin embargo, solo la sobreexpresión de *fabF* y la delección total de *fabD*, junto a la sobreexpresión de las desaturasas *desA* y *desB* mejoraron la producción de ALA en comparación con el mutante que solo sobreexpresa las desaturasas. El triple mutante MSM45, alcanzó los niveles de producción más altos de ALA que se han descrito en cianobacterias

11. Se han desarrollado dos metabólicos de Se7942. El primero, *iMS705*, contiene 705 genes, 843 reacciones y 826 metabólicos. El segundo, *iMS835*, contiene 835 genes, 806 metabolitos y 895 reacciones. La capacidad predictiva de genes esenciales en *iMS835* es la más alta de todos los modelos que existen para esta cianobacteria. Además, nos ha permitido entender por qué la delección parcial del gen *fabH* no mejoró la producción de ALA. *iMS835* pudo reproducir *in silico* el efecto obtenido *in vivo* con la sobreexpresión de *fabF*, no así con la delección de *fadD*.

Bibliografía

8. Bibliografía

- Abdel-Rahman, M.A., Y. Tashiro, and K. Sonomoto. 2011. Lactic acid production from lignocellulose-derived sugars using lactic acid bacteria: overview and limits. *J Biotechnol.* 156:286-301.
- Abed, R.M., S. Dobretsov, and K. Sudesh. 2009. Applications of cyanobacteria in biotechnology. *Journal of applied microbiology.* 106:1-12.
- Adams, J., and F. Rosenzweig. 2014. Experimental microbial evolution: history and conceptual underpinnings. *Genomics.* 104:393-398.
- Aichi, M., S. Maeda, K. Ichikawa, and T. Omata. 2004. Nitrite-responsive activation of the nitrate assimilation operon in Cyanobacteria plays an essential role in up-regulation of nitrate assimilation activities under nitrate-limited growth conditions. *Journal of bacteriology.* 186:3224-3229.
- Amann, E., B. Ochs, and K.J. Abel. 1988. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene.* 69:301-315.
- Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome biology.* 11:R106.
- Andersson, C.R., N.F. Tsinoremas, J. Shelton, N.V. Lebedeva, J. Yarrow, H. Min, and S.S. Golden. 2000. Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods in enzymology.* 305:527-542.
- Arnon, D.I., B.D. McSwain, H.Y. Tsujimoto, and K. Wada. 1974. Photochemical activity and components of membrane preparations from blue-green algae. I. Coexistence of two photosystems in relation to chlorophyll a and removal of phycocyanin. *Biochimica et biophysica acta.* 357:231-245.
- Atsumi, S., W. Higashide, and J.C. Liao. 2009. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nature biotechnology.* 27:1177-1180.
- Barrick, J.E., and R.E. Lenski. 2013. Genome dynamics during experimental evolution. *Nature reviews. Genetics.* 14:827-839.
- Barrick, J.E., D.S. Yu, S.H. Yoon, H. Jeong, T.K. Oh, D. Schneider, R.E. Lenski, and J.F. Kim. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature.* 461:1243-1247.
- Bazin, M.J. 1971. 5-fluorouracil resistance in *Anacystis nidulans* Drouet. *British Phycological Journal.* 6:25-28.
- Beaumont, H.J., J. Gallie, C. Kost, G.C. Ferguson, and P.B. Rainey. 2009. Experimental evolution of bet hedging. *Nature.* 462:90-93.
- Beck, C., H. Knoop, I.M. Axmann, and R. Steuer. 2012. The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms. *BMC genomics.* 13:56.
- Becker, S.A., and B.O. Palsson. 2008. Three factors underlying incorrect in silico predictions of essential metabolic genes. *BMC systems biology.* 2:14.
- Beld, J., D.J. Lee, and M.D. Burkart. 2015. Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Molecular bioSystems.* 11:38-59.
- Bell, G., and A. Gonzalez. 2011. Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science.* 332:1327-1330.
- Bennett, A.F., K.M. Dao, and R.E. Lenski. 1990. Rapid evolution in response to high-temperature selection. *Nature.* 346:79-81.

- Bennett, A.F., and R.E. Lenski. 1993. Evolutionary Adaptation to Temperature II. Thermal Niches of Experimental Lines of *Escherichia coli*. *Evolution; international journal of organic evolution*. 47:1-12.
- Bergler, H., S. Fuchsichler, G. Högenauer, and F. Turnowsky. 1996. The Enoyl-[Acyl-Carrier-Protein] Reductase (FabI) of *Escherichia coli*, which Catalyzes a Key Regulatory Step in Fatty Acid Biosynthesis, Accepts NADH and NADPH as Cofactors and is Inhibited by Palmitoyl-CoA. *European Journal of Biochemistry*. 242:689-694.
- Berla, B.M., R. Saha, C.M. Immethun, C.D. Maranas, T.S. Moon, and H.B. Pakrasi. 2013. Synthetic biology of cyanobacteria: unique challenges and opportunities. *Front Microbiol*. 4:246.
- Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. 2000. The Protein Data Bank. *Nucleic acids research*. 28:235-242.
- Bhalerao, R.P., J.L. Collier, P. Gustafsson, and A.R. Grossman. 1995. The structure of phycobilisomes in mutants of *Synechococcus* sp. strain PCC 7942 devoid of specific linker polypeptides. *Photochemistry and Photobiology*. 61:298-302.
- Bhaya, D., N.R. Bianco, D. Bryant, and A. Grossman. 2000. Type IV pilus biogenesis and motility in the cyanobacterium *Synechocystis* sp. PCC6803. *Molecular microbiology*. 37:941-951.
- Blount, Z.D., J.E. Barrick, C.J. Davidson, and R.E. Lenski. 2012. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature*. 489:513-518.
- Blount, Z.D., C.Z. Borland, and R.E. Lenski. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. 105:7899-7906.
- Bok, J.W., and N.P. Keller. 2012. Fast and easy method for construction of plasmid vectors using modified quick-change mutagenesis. *Methods in molecular biology*. 944:163-174.
- Bonfil, D.J., M. Ronen-Tarazi, D. Sultemeyer, J. Lieman-Hurwitz, D. Schatz, and A. Kaplan. 1998. A putative HCO₃⁻ transporter in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *FEBS letters*. 430:236-240.
- Brodrick, J.T., B.E. Rubin, D.G. Welkie, N. Du, N. Mih, S. Diamond, J.J. Lee, S.S. Golden, and B.O. Palsson. 2016. Unique attributes of cyanobacterial metabolism revealed by improved genome-scale metabolic modeling and essential gene analysis. *Proceedings of the National Academy of Sciences of the United States of America*. 113:E8344-E8353.
- Buckling, A., R. Craig Maclean, M.A. Brockhurst, and N. Colegrave. 2009. The Beagle in a bottle. *Nature*. 457:824-829.
- Buick, R. 1992. The antiquity of oxygenic photosynthesis: evidence from stromatolites in sulphate-deficient Archaean lakes. *Science*. 255:74-77.
- Burkhardt, J., J. Vonck, and B. Averhoff. 2011. Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium *Thermus thermophilus* HB27. *The Journal of biological chemistry*. 286:9977-9984.
- Campbell, J., 3rd, S.E. Stevens, Jr., and D.L. Balkwill. 1982. Accumulation of poly-beta-hydroxybutyrate in *Spirulina platensis*. *Journal of bacteriology*. 149:361-363.
- Campbell, J.W., and J.E. Cronan, Jr. 2002. The enigmatic *Escherichia coli* *fadE* gene is *yafH*. *Journal of bacteriology*. 184:3759-3764.
- Carver, T., S.R. Harris, M. Berriman, J. Parkhill, and J.A. McQuillan. 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*. 28:464-469.
- Clark, D.P., D. DeMendoza, M.L. Polacco, and J.E. Cronan. 1983. .beta.-Hydroxydecanoyl thioester dehydrase does not catalyze a rate-limiting step in *Escherichia coli* unsaturated fatty acid synthesis. *Biochemistry*. 22:5897-5902.
- Clarke, A.K., D. Campbell, P. Gustafsson, and G. Öquist. 1995. Dynamic responses of photosystem II and phycobilisomes to changing light in the cyanobacterium *Synechococcus* sp. PCC 7942. *Planta*. 197:553-562.

- Collier, J.L., and A.R. Grossman. 1992. Chlorosis induced by nutrient deprivation in *Synechococcus* sp. strain PCC 7942: not all bleaching is the same. *Journal of bacteriology*. 174:4718-4726.
- Collins, S., D. Sultemeyer, and G. Bell. 2006. Changes in C uptake in populations of *Chlamydomonas reinhardtii* selected at high CO₂. *Plant, cell & environment*. 29:1812-1819.
- Cronan, J.E., Jr., C.H. Birge, and P.R. Vagelos. 1969. Evidence for two genes specifically involved in unsaturated fatty acid biosynthesis in *Escherichia coli*. *Journal of bacteriology*. 100:601-604.
- Crozat, E., N. Philippe, R.E. Lenski, J. Geiselmann, and D. Schneider. 2005. Long-term experimental evolution in *Escherichia coli*. XII. DNA topology as a key target of selection. *Genetics*. 169:523-532.
- Chan, D.I., and H.J. Vogel. 2010. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem J*. 430:1-19.
- Cheah, Y.E., A.J. Zimont, S.K. Lunka, S.C. Albers, S.J. Park, K.F. Reardon, and C.A.M. Peebles. 2015. Diel light:dark cycles significantly reduce FFA accumulation in FFA producing mutants of *Synechocystis* sp. PCC 6803 compared to continuous light. *Algal Research*. 12:487-496.
- Chen, G., S. Qu, Q. Wang, F. Bian, Z. Peng, Y. Zhang, H. Ge, J. Yu, N. Xuan, Y. Bi, and Q. He. 2014. Transgenic expression of delta-6 and delta-15 fatty acid desaturases enhances omega-3 polyunsaturated fatty acid accumulation in *Synechocystis* sp. PCC6803. *Biotechnology for biofuels*. 7:32.
- Chen, I., and D. Dubnau. 2004. DNA uptake during bacterial transformation. *Nature reviews. Microbiology*. 2:241-249.
- Chen, Y., C.K. Holtman, R.D. Magnuson, P.A. Youderian, and S.S. Golden. 2008. The complete sequence and functional analysis of pANL, the large plasmid of the unicellular freshwater cyanobacterium *Synechococcus elongatus* PCC 7942. *Plasmid*. 59:176-192.
- Choi-Rhee, E., and J.E. Cronan. 2003. The biotin carboxylase-biotin carboxyl carrier protein complex of *Escherichia coli* acetyl-CoA carboxylase. *The Journal of biological chemistry*. 278:30806-30812.
- da Silva, G.P., M. Mack, and J. Contiero. 2009. Glycerol: a promising and abundant carbon source for industrial microbiology. *Biotechnology advances*. 27:30-39.
- Daugherty, M., V. Vonstein, R. Overbeek, and A. Osterman. 2001. Archaeal shikimate kinase, a new member of the GHMP-kinase family. *Journal of bacteriology*. 183:292-300.
- Davis, M.S., and J.E. Cronan, Jr. 2001. Inhibition of *Escherichia coli* acetyl coenzyme A carboxylase by acyl-acyl carrier protein. *Journal of bacteriology*. 183:1499-1503.
- de Crecy, E., S. Jaronski, B. Lyons, T.J. Lyons, and N.O. Keyhani. 2009. Directed evolution of a filamentous fungus for thermotolerance. *BMC biotechnology*. 9:74.
- Deatherage, D.E., and J.E. Barrick. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods in molecular biology*. 1151:165-188.
- Demarre, G., A.M. Guerout, C. Matsumoto-Mashimo, D.A. Rowe-Magnus, P. Marliere, and D. Mazel. 2005. A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalph) conjugative machineries and their cognate *Escherichia coli* host strains. *Research in microbiology*. 156:245-255.
- Deng, M.D., and J.R. Coleman. 1999. Ethanol synthesis by genetic engineering in cyanobacteria. *Applied and environmental microbiology*. 65:523-528.
- Dhar, R., R. Sagesser, C. Weikert, J. Yuan, and A. Wagner. 2011. Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. *Journal of evolutionary biology*. 24:1135-1153.
- Dion, E., F. Zele, J.C. Simon, and Y. Outreman. 2011. Rapid evolution of parasitoids when faced with the symbiont-mediated resistance of their hosts. *Journal of evolutionary biology*. 24:741-750.

- Dismukes, G.C., D. Carrieri, N. Bennette, G.M. Ananyev, and M.C. Posewitz. 2008. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Current opinion in biotechnology*. 19:235-240.
- Ditty, J.L., S.B. Williams, and S.S. Golden. 2003. A cyanobacterial circadian timing mechanism. *Annual review of genetics*. 37:513-543.
- Dong, G., Q. Yang, Q. Wang, Y.I. Kim, T.L. Wood, K.W. Osteryoung, A. van Oudenaarden, and S.S. Golden. 2010. Elevated ATPase activity of KaiC applies a circadian checkpoint on cell division in *Synechococcus elongatus*. *Cell*. 140:529-539.
- Dower, W.J., J.F. Miller, and C.W. Ragsdale. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic acids research*. 16:6127-6145.
- Dragosits, M., and D. Mattanovich. 2013. Adaptive laboratory evolution -- principles and applications for biotechnology. *Microbial cell factories*. 12:64.
- Duan, Y., Q. Luo, F. Liang, and X. Lu. 2016. Sucrose secreted by the engineered cyanobacterium and its fermentability. *Journal of Ocean University of China*. 15:890-896.
- Ducat, D.C., J.A. Avelar-Rivas, J.C. Way, and P.A. Silver. 2012. Rerouting carbon flux to enhance photosynthetic productivity. *Applied and environmental microbiology*. 78.
- Ducat, D.C., G. Sachdeva, and P.A. Silver. 2011a. Rewiring hydrogenase-dependent redox circuits in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 108:3941-3946.
- Ducat, D.C., J.C. Way, and P.A. Silver. 2011b. Engineering cyanobacteria to generate high-value products. *Trends in biotechnology*. 29:95-103.
- Dutta, D., D. De, S. Chaudhuri, and S.K. Bhattacharya. 2005. Hydrogen production by Cyanobacteria. *Microbial cell factories*. 4:36.
- Dykhuizen, D.E., and D.L. Hartl. 1983. Selection in chemostats. *Microbiological reviews*. 47:150-168.
- Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*. 32:1792-1797.
- Edwards, J.S., and B.O. Palsson. 2000. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences of the United States of America*. 97:5528-5533.
- Edwards, P., J.S. Nelsen, J.G. Metz, and K. Dehesh. 1997. Cloning of the *fabF* gene in an expression vector and in vitro characterization of recombinant *fabF* and *fabB* encoded enzymes from *Escherichia coli*. *FEBS letters*. 402:62-66.
- Eisenhut, M., J. Georg, S. Klahn, I. Sakurai, H. Mustila, P. Zhang, W.R. Hess, and E.M. Aro. 2012. The antisense RNA *As1_flv4* in the Cyanobacterium *Synechocystis* sp. PCC 6803 prevents premature expression of the *flv4-2* operon upon shift in inorganic carbon supply. *The Journal of biological chemistry*. 287:33153-33162.
- El-Bestawy, E. 2008. Treatment of mixed domestic-industrial wastewater using cyanobacteria. *Journal of industrial microbiology & biotechnology*. 35:1503-1516.
- Elena, S.F., and R.E. Lenski. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature reviews. Genetics*. 4:457-469.
- Elhai, J., A. Vepritskiy, A.M. Muro-Pastor, E. Flores, and C.P. Wolk. 1997. Reduction of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain PCC 7120. *Journal of bacteriology*. 179:1998-2005.
- Encinas, D., M.P. Garcillan-Barcia, M. Santos-Merino, L. Delaye, A. Moya, and F. de la Cruz. 2014. Plasmid conjugation from proteobacteria as evidence for the origin of xenologous genes in cyanobacteria. *Journal of bacteriology*. 196:1551-1559.
- Erdrich, P., H. Knoop, R. Steuer, and S. Klamt. 2014. Cyanobacterial biofuels: new insights and strain design strategies revealed by computational modeling. *Microbial cell factories*. 13:128.

- Eswar, N., B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, and A. Sali. 2006. Comparative protein structure modeling using Modeller. *Current protocols in bioinformatics*. Chapter 5:Unit 5 6.
- Fawzy, M.A., and A.A. Issa. 2016. Bioremoval of heavy metals and nutrients from sewage plant by *Anabaena oryzae* and *Cyanosarcina fontana*. *International journal of phytoremediation*. 18:321-328.
- Feigenbaum, J., and H. Schulz. 1975. Thiolases of *Escherichia coli*: purification and chain length specificities. *Journal of bacteriology*. 122:407-411.
- Feist, A.M., C.S. Henry, J.L. Reed, M. Krummenacker, A.R. Joyce, P.D. Karp, L.J. Broadbelt, V. Hatzimanikatis, and B.O. Palsson. 2007. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular systems biology*. 3:121.
- Feist, A.M., M.J. Herrgard, I. Thiele, J.L. Reed, and B.O. Palsson. 2009. Reconstruction of biochemical networks in microorganisms. *Nature reviews. Microbiology*. 7:129-143.
- Feist, A.M., H. Nagarajan, A.E. Rotaru, P.L. Tremblay, T. Zhang, K.P. Nevin, D.R. Lovley, and K. Zengler. 2014. Constraint-based modeling of carbon fixation and the energetics of electron transfer in *Geobacter metallireducens*. *PLoS computational biology*. 10:e1003575.
- Feist, A.M., and B.O. Palsson. 2010. The biomass objective function. *Current opinion in microbiology*. 13:344-349.
- Feist, A.M., D.C. Zielinski, J.D. Orth, J. Schellenberger, M.J. Herrgard, and B.O. Palsson. 2010. Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metabolic engineering*. 12:173-186.
- Fernandez-Lopez, R., C. Machon, C.M. Longshaw, S. Martin, S. Molin, E.L. Zechner, M. Espinosa, E. Lanka, and F. de la Cruz. 2005. Unsaturated fatty acids are inhibitors of bacterial conjugation. *Microbiology*. 151:3517-3526.
- Fisher, M.L., R. Allen, Y. Luo, and R. Curtiss, 3rd. 2013. Export of extracellular polysaccharides modulates adherence of the *Cyanobacterium synechocystis*. *PloS one*. 8:e74514.
- Fong, S.S., A.P. Burgard, C.D. Herring, E.M. Knight, F.R. Blattner, C.D. Maranas, and B.O. Palsson. 2005. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol Bioeng*. 91:643-648.
- Forster, J., I. Famili, P. Fu, B.O. Palsson, and J. Nielsen. 2003. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome research*. 13:244-253.
- Fowler, Z.L., W.W. Gikandi, and M.A. Koffas. 2009. Increased malonyl coenzyme A biosynthesis by tuning the *Escherichia coli* metabolic network and its application to flavanone production. *Applied and environmental microbiology*. 75:5831-5839.
- Froehlich, J.E., R. Poorman, E. Reardon, S.R. Barnum, and J.G. Jaworski. 1990. Purification and characterization of acyl carrier protein from two cyanobacteria species. *Eur J Biochem*. 193:817-825.
- Froissard, M., M. Canonge, M. Pouteaux, B. Cintrat, S. Mohand-Oumoussa, S.E. Guillouet, T. Chardot, N. Jacques, and S. Casaregola. 2015. Lipids containing medium-chain fatty acids are specific to post-whole genome duplication *Saccharomycotina* yeasts. *BMC evolutionary biology*. 15:97.
- Fu, P. 2009. Genome-scale modeling of *Synechocystis* sp. PCC 6803 and prediction of pathway insertion. *Journal of Chemical Technology & Biotechnology*. 84:473-483.
- Fujisawa, T., R. Narikawa, S.I. Maeda, S. Watanabe, Y. Kanesaki, K. Kobayashi, J. Nomata, M. Hanaoka, M. Watanabe, S. Ehira, E. Suzuki, K. Awai, and Y. Nakamura. 2017. CyanoBase: a large-scale update on its 20th anniversary. *Nucleic acids research*. 45:D551-D554.
- Funatsu, G., and H.G. Wittmann. 1972. Ribosomal proteins. 33. Location of amino-acid replacements in protein S12 isolated from *Escherichia coli* mutants resistant to streptomycin. *Journal of molecular biology*. 68:547-550.

- Gabay, C., J. Lieman-Hurwitz, M. Hassidim, M. Ronen-Tarazi, and A. Kaplan. 1998. Modification of topA in *Synechococcus* sp. PCC 7942 resulted in mutants capable of growing under low but not high concentration of CO₂. *FEMS microbiology letters*. 159:343-347.
- Garland, T., and M.R. Rose. 2009. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*. University of California Press.
- Garwin, J.L., A.L. Klages, and J.E. Cronan, Jr. 1980. Beta-ketoacyl-acyl carrier protein synthase II of *Escherichia coli*. Evidence for function in the thermal regulation of fatty acid synthesis. *The Journal of biological chemistry*. 255:3263-3265.
- Gerrish, P.J., and R.E. Lenski. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica*. 102-103:127-144.
- Gibson, D.G. 2011. Enzymatic assembly of overlapping DNA fragments. *Methods in enzymology*. 498:349-361.
- Giovannoni, S.J., S. Turner, G.J. Olsen, S. Barns, D.J. Lane, and N.R. Pace. 1988. Evolutionary relationships among cyanobacteria and green chloroplasts. *Journal of bacteriology*. 170:3584-3592.
- Golden, S.S. 2003. Timekeeping in bacteria: the cyanobacterial circadian clock. *Current opinion in microbiology*. 6:535-540.
- Golden, S.S., C.H. Johnson, and T. Kondo. 1998. The cyanobacterial circadian system: a clock apart. *Current opinion in microbiology*. 1:669-673.
- Golden, S.S., M.S. Nalty, and D.S. Cho. 1989. Genetic relationship of two highly studied *Synechococcus* strains designated *Anacystis nidulans*. *Journal of bacteriology*. 171:24-29.
- Golden, S.S., and L.A. Sherman. 1983. A hybrid plasmid is a stable cloning vector for the cyanobacterium *Anacystis nidulans* R2. *Journal of bacteriology*. 155:966-972.
- Golden, S.S., and L.A. Sherman. 1984. Optimal conditions for genetic transformation of the cyanobacterium *Anacystis nidulans* R2. *Journal of bacteriology*. 158:36-42.
- Gombos, Z., E. Kanervo, N. Tsvetkova, T. Sakamoto, E.M. Aro, and N. Murata. 1997. Genetic Enhancement of the Ability to Tolerate Photoinhibition by Introduction of Unsaturated Bonds into Membrane Glycerolipids. *Plant physiology*. 115:551-559.
- Grant, S.G., J. Jessee, F.R. Bloom, and D. Hanahan. 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proceedings of the National Academy of Sciences of the United States of America*. 87:4645-4649.
- Gray, M.W. 1992. The endosymbiont hypothesis revisited. *International review of cytology*. 141:233-357.
- Griese, M., C. Lange, and J. Soppa. 2011. Ploidy in cyanobacteria. *FEMS microbiology letters*. 323:124-131.
- Grigorieva, G., and S. Shestakov. 1982. Transformation in the cyanobacterium *Synechocystis* sp. 6803. *FEMS microbiology letters*. 13:367-370.
- Gu, H., R.E. Jinkerson, F.K. Davies, L.A. Sisson, P.E. Schneider, and M.C. Posewitz. 2016. Modulation of Medium-Chain Fatty Acid Synthesis in *Synechococcus* sp. PCC 7002 by Replacing FabH with a *Chaetoceros* Ketoacyl-ACP Synthase. *Frontiers in plant science*. 7:690.
- Guan, X., S. Qin, F. Zhao, X. Zhang, and X. Tang. 2007. Phycobilisomes linker family in cyanobacterial genomes: divergence and evolution. *International journal of biological sciences*. 3:434-445.
- Hamilton, J.J., and J.L. Reed. 2012. Identification of functional differences in metabolic networks using comparative genomics and constraint-based models. *PLoS one*. 7:e34670.
- Heath, R.J., and C.O. Rock. 1996. Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in *Escherichia coli* fatty acid biosynthesis. *The Journal of biological chemistry*. 271:27795-27801.

- Heberle, H., G.V. Meirelles, F.R. da Silva, G.P. Telles, and R. Minghim. 2015. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC bioinformatics*. 16:169.
- Hendry, J.I., C.B. Prasanna, A. Joshi, S. Dasgupta, and P.P. Wangikar. 2016. Metabolic model of *Synechococcus* sp. PCC 7002: Prediction of flux distribution and network modification for enhanced biofuel production. *Bioresource technology*. 213:190-197.
- Henry, C.S., M. DeJongh, A.A. Best, P.M. Frybarger, B. Linsay, and R.L. Stevens. 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature biotechnology*. 28:977-982.
- Herrero, A., and F.G. Flores. 2008. The Cyanobacteria: Molecular Biology, Genomics, and Evolution. Caister Academic Press.
- Herrero, A., A.M. Muro-Pastor, and E. Flores. 2001. Nitrogen control in cyanobacteria. *Journal of bacteriology*. 183:411-425.
- Heyer, H., and W.E. Krumbein. 1991. Excretion of fermentation products in dark and anaerobically incubated cyanobacteria. *Archives of microbiology*. 155:284-287.
- Hihara, Y., A. Kamei, M. Kanehisa, A. Kaplan, and M. Ikeuchi. 2001. DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *The Plant cell*. 13:793-806.
- Hirokawa, Y., Y. Dempo, E. Fukusaki, and T. Hanai. 2016. Metabolic engineering for isopropanol production by an engineered cyanobacterium, *Synechococcus elongatus* PCC 7942, under photosynthetic conditions. *Journal of bioscience and bioengineering*.
- Hirokawa, Y., I. Suzuki, and T. Hanai. 2015. Optimization of isopropanol production by engineered cyanobacteria with a synthetic metabolic pathway. *Journal of bioscience and bioengineering*. 119:585-590.
- Hoang, K.L., L.T. Morran, and N.M. Gerardo. 2016. Experimental Evolution as an Underutilized Tool for Studying Beneficial Animal–Microbe Interactions. *Frontiers in Microbiology*. 7.
- Hoiczky, E., and A. Hansel. 2000. Cyanobacterial cell walls: news from an unusual prokaryotic envelope. *Journal of bacteriology*. 182:1191-1199.
- Hong, S.-J., and C.-G. Lee. 2007. Evaluation of central metabolism based on a genomic database of *Synechocystis* PCC6803. *Biotechnology and Bioprocess Engineering*. 12:165-173.
- Horinouchi, T., T. Minamoto, S. Suzuki, H. Shimizu, and C. Furusawa. 2014. Development of an automated culture system for laboratory evolution. *Journal of laboratory automation*. 19:478-482.
- Hosoda, K., M. Habuchi, S. Suzuki, M. Miyazaki, G. Takikawa, T. Sakurai, A. Kashiwagi, M. Sueyoshi, Y. Matsumoto, A. Kiuchi, K. Mori, and T. Yomo. 2014. Adaptation of a cyanobacterium to a biochemically rich environment in experimental evolution as an initial step toward a chloroplast-like state. *PloS one*. 9:e98337.
- Hunt, P., A. Martinelli, K. Modrzynska, S. Borges, A. Creasey, L. Rodrigues, D. Beraldi, L. Loewe, R. Fawcett, S. Kumar, M. Thomson, U. Trivedi, T.D. Otto, A. Pain, M. Blaxter, and P. Cravo. 2010. Experimental evolution, genetic analysis and genome re-sequencing reveal the mutation conferring artemisinin resistance in an isogenic lineage of malaria parasites. *BMC genomics*. 11:499.
- Hyduke, D., J. Schellenberger, R. Que, R. Fleming, I. Thiele, J. Orth, A. Feist, D. Zielinski, A. Bordbar, N. Lewis, S. Rahmanian, J. Kang, and B. Palsson. 2011. COBRA Toolbox 2.0.
- Ibarra, R.U., J.S. Edwards, and B.O. Palsson. 2002. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*. 420:186-189.
- Ihlenfeldt, M.J., and J. Gibson. 1975. CO₂ fixation and its regulation in *Anacystis nidulans* (*Synechococcus*). *Archives of microbiology*. 102:13-21.
- Imam, S., Z. Chen, D.S. Roos, and M. Pohlschroder. 2011. Identification of surprisingly diverse type IV pili, across a broad range of gram-positive bacteria. *PloS one*. 6:e28919.
- Iwasaki, H., and T. Kondo. 2004. Circadian timing mechanism in the prokaryotic clock system of cyanobacteria. *Journal of biological rhythms*. 19:436-444.

- Iwasaki, H., S.B. Williams, Y. Kitayama, M. Ishiura, S.S. Golden, and T. Kondo. 2000. A kaiC-interacting sensory histidine kinase, SasA, necessary to sustain robust circadian oscillation in cyanobacteria. *Cell*. 101:223-233.
- Jackowski, S., Y.M. Zhang, A.C. Price, S.W. White, and C.O. Rock. 2002. A missense mutation in the fabB (beta-ketoacyl-acyl carrier protein synthase I) gene confers tiolactomycin resistance to *Escherichia coli*. *Antimicrobial agents and chemotherapy*. 46:1246-1252.
- Jakovljevic, V., S. Leonardy, M. Hoppert, and L. Sogaard-Andersen. 2008. PilB and PilT are ATPases acting antagonistically in type IV pilus function in *Myxococcus xanthus*. *Journal of bacteriology*. 190:2411-2421.
- Janssen, H.J., and A. Steinbuchel. 2014. Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid based biofuels. *Biotechnology for biofuels*. 7:7.
- Jindou, S., Y. Ito, N. Mito, K. Uematsu, A. Hosoda, and H. Tamura. 2014. Engineered platform for bioethylene production by a cyanobacterium expressing a chimeric complex of plant enzymes. *ACS synthetic biology*. 3:487-496.
- Johnson, C.H. 2004. Precise circadian clocks in prokaryotic cyanobacteria. *Current issues in molecular biology*. 6:103-110.
- Johnson, C.H. 2010. Circadian clocks and cell division: what's the pacemaker? *Cell cycle*. 9:3864-3873.
- Johnson, C.H., T. Mori, and Y. Xu. 2008. A cyanobacterial circadian clockwork. *Current biology : CB*. 18:R816-R825.
- Joshi, V.C., and S.J. Wakil. 1971. Studies on the mechanism of fatty acid synthesis. XXVI. Purification and properties of malonyl-coenzyme A--acyl carrier protein transacylase of *Escherichia coli*. *Archives of biochemistry and biophysics*. 143:493-505.
- Jung, J.Y., H.S. Yun, J. Lee, and M.K. Oh. 2011. Production of 1,2-propanediol from glycerol in *Saccharomyces cerevisiae*. *J Microbiol Biotechnol*. 21.
- Kaczmarzyk, D., and M. Fulda. 2010. Fatty acid activation in cyanobacteria mediated by acyl-acyl carrier protein synthetase enables fatty acid recycling. *Plant physiology*. 152:1598-1610.
- Kates, M. 1992. Archaeobacterial lipids: structure, biosynthesis and function. *Biochemical Society symposium*. 58:51-72.
- Kato, A., N. Takatani, K. Use, K. Uesaka, K. Ikeda, Y. Chang, K. Kojima, M. Aichi, K. Ihara, K. Nakahigashi, S. Maeda, and T. Omata. 2015. Identification of a Cyanobacterial RND-Type Efflux System Involved in Export of Free Fatty Acids. *Plant & cell physiology*. 56:2467-2477.
- Kato, A., K. Use, N. Takatani, K. Ikeda, M. Matsuura, K. Kojima, M. Aichi, S. Maeda, and T. Omata. 2016. Modulation of the balance of fatty acid production and secretion is crucial for enhancement of growth and productivity of the engineered mutant of the cyanobacterium *Synechococcus elongatus*. *Biotechnology for biofuels*. 9:91.
- Kawecki, T.J., R.E. Lenski, D. Ebert, B. Hollis, I. Olivieri, and M.C. Whitlock. 2012. Experimental evolution. *Trends in ecology & evolution*. 27:547-560.
- Kelley, L.A., S. Mezulis, C.M. Yates, M.N. Wass, and M.J. Sternberg. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*. 10:845-858.
- Kenyon, C.N. 1972. Fatty acid composition of unicellular strains of blue-green algae. *Journal of bacteriology*. 109:827-834.
- Kenyon, C.N., R. Rippka, and R.Y. Stanier. 1972. Fatty acid composition and physiological properties of some filamentous blue-green algae. *Archiv fur Mikrobiologie*. 83:216-236.
- Kerby, N.W., R.H. Reed, and P. Rowell. 1990. Incorporation of ¹⁴C in the cyanobacterium *Synechococcus* PCC 6301 following salt stress. *Planta*. 181:393-398.
- Kessner, D., and J. Novembre. 2015. Power analysis of artificial selection experiments using efficient whole genome simulation of quantitative traits. *Genetics*. 199:991-1005.

- Kim, B., W.J. Kim, D.I. Kim, and S.Y. Lee. 2015. Applications of genome-scale metabolic network model in metabolic engineering. *Journal of industrial microbiology & biotechnology*. 42:339-348.
- King, Z.A., J. Lu, A. Drager, P. Miller, S. Federowicz, J.A. Lerman, A. Ebrahim, B.O. Palsson, and N.E. Lewis. 2016. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic acids research*. 44:D515-522.
- Klanchui, A., C. Khannapho, A. Phodee, S. Cheevadhanarak, and A. Meechai. 2012. iAK692: a genome-scale metabolic model of *Spirulina platensis* C1. *BMC systems biology*. 6:71.
- Knoop, H., M. Grundel, Y. Zilliges, R. Lehmann, S. Hoffmann, W. Lockau, and R. Steuer. 2013. Flux balance analysis of cyanobacterial metabolism: the metabolic network of *Synechocystis* sp. PCC 6803. *PLoS computational biology*. 9:e1003081.
- Knoop, H., and R. Steuer. 2015. A computational analysis of stoichiometric constraints and trade-offs in cyanobacterial biofuel production. *Frontiers in bioengineering and biotechnology*. 3:47.
- Knoop, H., Y. Zilliges, W. Lockau, and R. Steuer. 2010. The metabolic network of *Synechocystis* sp. PCC 6803: systemic properties of autotrophic growth. *Plant physiology*. 154:410-422.
- Kolman, M.A., C.N. Nishi, M. Perez-Cenci, and G.L. Salerno. 2015. Sucrose in cyanobacteria: from a salt-response molecule to play a key role in nitrogen fixation. *Life*. 5:102-126.
- Kolodny-Hirsch, D.M.V., N.A.M. . 1997 Selection of a morphological variant of *Autographa californica* nuclear polyhedrosis virus with increased virulence following serial passage in *Plutella xylostella*. *J. Invertebr. Pathol.* 69:205-211.
- Kolss, M., R.K. Vijendravarma, G. Schwaller, and T.J. Kawecki. 2009. Life-history consequences of adaptation to larval nutritional stress in *Drosophila*. *Evolution; international journal of organic evolution*. 63:2389-2401.
- Kortmann, J., and F. Narberhaus. 2012. Bacterial RNA thermometers: molecular zippers and switches. *Nature reviews. Microbiology*. 10:255-265.
- Kratz, W.A., and J. Myers. 1955. Nutrition and Growth of Several Blue-Green Algae. *American Journal of Botany*. 42:282-287.
- Kulkarni, R.D., and S.S. Golden. 1995. Form II of D1 is important during transition from standard to high light intensity in *Synechococcus* sp. strain PCC 7942. *Photosynthesis research*. 46:435-443.
- Kumar, M.S., G. Muralitharan, and N. Thajuddin. 2009. Screening of a hypersaline cyanobacterium, *Phormidium tenue*, for the degradation of aromatic hydrocarbons: naphthalene and anthracene. *Biotechnology Letters*. 31:1863.
- Kuo, J., and C. Khosla. 2014. The initiation ketosynthase (FabH) is the sole rate-limiting enzyme of the fatty acid synthase of *Synechococcus* sp. PCC 7002. *Metabolic engineering*. 22:53-59.
- Kusakabe, T., T. Tatsuke, K. Tsuruno, Y. Hirokawa, S. Atsumi, J.C. Liao, and T. Hanai. 2013. Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metabolic engineering*. 20:101-108.
- Lai, C.Y., and J.E. Cronan. 2003. Beta-ketoacyl-acyl carrier protein synthase III (FabH) is essential for bacterial fatty acid synthesis. *The Journal of biological chemistry*. 278:51494-51503.
- Lai, M.C., and E.I. Lan. 2015. Advances in Metabolic Engineering of Cyanobacteria for Photosynthetic Biochemical Production. *Metabolites*. 5:636-658.
- Lan, E.I., and J.C. Liao. 2011. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metabolic engineering*. 13:353-363.
- Lan, E.I., and J.C. Liao. 2012. ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 109:6018-6023.
- Lan, E.I., S.Y. Ro, and J.C. Liao. 2013. Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria. *Energy & Environmental Science*. 6:2672-2681.

- Lang, G.I., and M.M. Desai. 2014. The spectrum of adaptive mutations in experimental evolution. *Genomics*. 104:412-416.
- Langmead, B., and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods*. 9:357-359.
- Lau, N.S., M. Matsui, and A.A. Abdullah. 2015. Cyanobacteria: Photoautotrophic Microbial Factories for the Sustainable Synthesis of Industrial Products. *BioMed research international*. 2015:754934.
- Laurenceau, R., G. Pehau-Arnaudet, S. Baconnais, J. Gault, C. Malosse, A. Dujancourt, N. Campo, J. Chamot-Rooke, E. Le Cam, J.P. Claverys, and R. Fronzes. 2013. A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS pathogens*. 9:e1003473.
- Lebeis, S.L., and D. Kalman. 2009. Aligning antimicrobial drug discovery with complex and redundant host-pathogen interactions. *Cell host & microbe*. 5:114-122.
- Lederberg, J. 1951. Streptomycin resistance; a genetically recessive mutation. *Journal of bacteriology*. 61:549-550.
- Lee, C.E., J.L. Remfert, and Y.M. Chang. 2007. Response to selection and evolvability of invasive populations. *Genetica*. 129:179-192.
- Lee, H.M., M.F. Vazquez-Bermudez, and N.T. de Marsac. 1999. The global nitrogen regulator NtcA regulates transcription of the signal transducer PII (GlnB) and influences its phosphorylation level in response to nitrogen and carbon supplies in the Cyanobacterium *synechococcus* sp. strain PCC 7942. *Journal of bacteriology*. 181:2697-2702.
- Lee, I., Y.O. Kim, S.C. Park, and J. Chun. 2015. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *International journal of systematic and evolutionary microbiology*.
- Leibundgut, M., T. Maier, S. Jenni, and N. Ban. 2008. The multienzyme architecture of eukaryotic fatty acid synthases. *Current opinion in structural biology*. 18:714-725.
- Lenski, R.E., M.R. Rose, S.C. Simpson, and S.C. Tadler. 1991. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *The American Naturalist*. 138:1315-1341.
- Lepore, B.W., M. Indic, H. Pham, E.M. Hearn, D.R. Patel, and B. van den Berg. 2011. Ligand-gated diffusion across the bacterial outer membrane. *Proceedings of the National Academy of Sciences of the United States of America*. 108:10121-10126.
- Lewis, N.E., H. Nagarajan, and B.O. Palsson. 2012. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature reviews. Microbiology*. 10:291-305.
- Li, C., F. Tao, J. Ni, Y. Wang, F. Yao, and P. Xu. 2015. Enhancing the light-driven production of D-lactate by engineering cyanobacterium using a combinational strategy. *Scientific reports*. 5:9777.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and S. Genome Project Data Processing. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25:2078-2079.
- Li, H., and J.C. Liao. 2013. Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO₂ to 1,2-propanediol. *Microbial cell factories*. 12:4.
- Li, Y., L. Wang, J. Ju, B. Yu, and Y. Ma. 2013. Efficient production of polymer-grade D-lactate by *Sporolactobacillus laevolacticus* DSM442 with agricultural waste cottonseed as the sole nitrogen source. *Bioresource technology*. 142:186-191.
- Liao, Y.C., M.H. Tsai, F.C. Chen, and C.A. Hsiung. 2012. GEMSiRV: a software platform for GENome-scale metabolic model simulation, reconstruction and visualization. *Bioinformatics*. 28:1752-1758.

- Liu, X., J. Sheng, and R. Curtiss, 3rd. 2011. Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 108:6899-6904.
- Liu, Y., Y. Zhang, X. Cao, and S. Xue. 2013. Cloning, purification, crystallization and preliminary X-ray crystallographic analysis of MCAT from *Synechocystis* sp. PCC 6803. *Acta crystallographica. Section F, Structural biology and crystallization communications*. 69:1256-1259.
- Lombard, J., P. Lopez-Garcia, and D. Moreira. 2012. An ACP-independent fatty acid synthesis pathway in archaea: implications for the origin of phospholipids. *Mol Biol Evol*. 29:3261-3265.
- Lopez-Maury, L., M. Garcia-Dominguez, F.J. Florencio, and J.C. Reyes. 2002. A two-component signal transduction system involved in nickel sensing in the cyanobacterium *Synechocystis* sp. PCC 6803. *Molecular microbiology*. 43:247-256.
- Los, D., I. Horvath, L. Vigh, and N. Murata. 1993. The temperature-dependent expression of the desaturase gene *desA* in *Synechocystis* PCC6803. *FEBS letters*. 318:57-60.
- Los, D.A., and K.S. Mironov. 2015. Modes of Fatty Acid desaturation in cyanobacteria: an update. *Life*. 5:554-567.
- Los, D.A., and N. Murata. 2004. Membrane fluidity and its roles in the perception of environmental signals. *Biochimica et biophysica acta*. 1666:142-157.
- Los, D.A., A. Zorina, M. Sinetova, S. Kryazhov, K. Mironov, and V.V. Zinchenko. 2010. Stress sensors and signal transducers in cyanobacteria. *Sensors*. 10:2386-2415.
- Lu, G., and E.N. Moriyama. 2004. Vector NTI, a balanced all-in-one sequence analysis suite. *Briefings in bioinformatics*. 5:378-388.
- Ludwig, M., and D.A. Bryant. 2011. Transcription Profiling of the Model Cyanobacterium *Synechococcus* sp. Strain PCC 7002 by Next-Gen (SOLiD) Sequencing of cDNA. *Front Microbiol*. 2:41.
- Ludwig, M., and D.A. Bryant. 2012. *Synechococcus* sp. Strain PCC 7002 Transcriptome: Acclimation to Temperature, Salinity, Oxidative Stress, and Mixotrophic Growth Conditions. *Front Microbiol*. 3:354.
- Mackey, S.R., and S.S. Golden. 2007. Winding up the cyanobacterial circadian clock. *Trends in microbiology*. 15:381-388.
- Maeda, S., G.D. Price, M.R. Badger, C. Enomoto, and T. Omata. 2000. Bicarbonate binding activity of the CmpA protein of the cyanobacterium *Synechococcus* sp. strain PCC 7942 involved in active transport of bicarbonate. *The Journal of biological chemistry*. 275:20551-20555.
- Magnuson, K., W. Oh, T.J. Larson, and J.E. Cronan, Jr. 1992. Cloning and nucleotide sequence of the *fabD* gene encoding malonyl coenzyme A-acyl carrier protein transacylase of *Escherichia coli*. *FEBS letters*. 299:262-266.
- Malatinszky, D., R. Steuer, and P.R. Jones. 2016. A comprehensively curated genome-scale two-cell model for the cyanobacterium *Anabaena* sp. PCC 7120. *Plant physiology*.
- Marraccini, P., S. Bulteau, C. Cassier-Chauvat, P. Mermet-Bouvier, and F. Chauvat. 1993. A conjugative plasmid vector for promoter analysis in several cyanobacteria of the genera *Synechococcus* and *Synechocystis*. *Plant Molecular Biology*. 23:905-909.
- Martinez-Noel, G.M., A.C. Cumino, L. Kolman Mde, and G.L. Salerno. 2013. First evidence of sucrose biosynthesis by single cyanobacterial bimodular proteins. *FEBS letters*. 587:1669-1674.
- Matsuoka, M., K. Takahama, and T. Ogawa. 2001. Gene replacement in cyanobacteria mediated by a dominant streptomycin-sensitive *rps12* gene that allows selection of mutants free from drug resistance markers. *Microbiology*. 147:2077-2087.
- McCloskey, D., B.O. Palsson, and A.M. Feist. 2013. Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Molecular systems biology*. 9:661.

- McEwen, J.T., I.M. Machado, M.R. Connor, and S. Atsumi. 2013. Engineering *Synechococcus elongatus* PCC 7942 for continuous growth under diurnal conditions. *Applied and environmental microbiology*. 79:1668-1675.
- McInroy, J.A., and J.W. Kloepper. 1995. Survey of indigenous bacterial endophytes from cotton and sweet corn. *Plant and Soil*. 173:337-342.
- Meyer, J.R., A.A. Agrawal, R.T. Quick, D.T. Dobias, D. Schneider, and R.E. Lenski. 2010. Parallel changes in host resistance to viral infection during 45,000 generations of relaxed selection. *Evolution; international journal of organic evolution*. 64:3024-3034.
- Meyer, J.R., D.T. Dobias, J.S. Weitz, J.E. Barrick, R.T. Quick, and R.E. Lenski. 2012. Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science*. 335:428-432.
- Michel, K.P., E.K. Pistorius, and S.S. Golden. 2001. Unusual regulatory elements for iron deficiency induction of the *idiA* gene of *Synechococcus elongatus* PCC 7942. *Journal of bacteriology*. 183:5015-5024.
- Moche, M., K. Dehesh, P. Edwards, and Y. Lindqvist. 2001. The crystal structure of β -ketoacyl-acyl carrier protein synthase II from *Synechocystis* sp. at 1.54 Å resolution and its relationship to other condensing enzymes1. *Journal of molecular biology*. 305:491-503.
- Monk, J., J. Nogales, and B.O. Palsson. 2014. Optimizing genome-scale network reconstructions. *Nature biotechnology*. 32:447-452.
- Montagud, A., E. Navarro, P. Fernandez de Cordoba, J.F. Urchueguia, and K.R. Patil. 2010. Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. *BMC systems biology*. 4:156.
- Montagud, A., A. Zelezniak, E. Navarro, P.F. de Cordoba, J.F. Urchueguia, and K.R. Patil. 2011. Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium *Synechocystis* sp. PCC6803. *Biotechnology journal*. 6:330-342.
- Morgan-Kiss, R.M., and J.E. Cronan. 2008. The *Lactococcus lactis* FabF fatty acid synthetic enzyme can functionally replace both the FabB and FabF proteins of *Escherichia coli* and the FabH protein of *Lactococcus lactis*. *Archives of microbiology*. 190:427-437.
- Moronta-Barrios, F., J. Espinosa, and A. Contreras. 2013. Negative control of cell size in the cyanobacterium *Synechococcus elongatus* PCC 7942 by the essential response regulator RpaB. *FEBS letters*. 587:504-509.
- Mueller, T.J., B.M. Berla, H.B. Pakrasi, and C.D. Maranas. 2013. Rapid construction of metabolic models for a family of Cyanobacteria using a multiple source annotation workflow. *BMC systems biology*. 7:142.
- Murata, N., H. Wada, and Z. Gombos. 1992. Modes of Fatty-Acid Desaturation in Cyanobacteria. *Plant and Cell Physiology*. 33:933-941.
- Nägeli, C. 1849. Gattungen einzelliger Algen: physiologisch und systematisch Bearbeitet. Friedrich Schulthess.
- Nakahira, Y., A. Ogawa, H. Asano, T. Oyama, and Y. Tozawa. 2013. Theophylline-Dependent Riboswitch as a Novel Genetic Tool for Strict Regulation of Protein Expression in Cyanobacterium *Synechococcus elongatus* PCC 7942. *Plant and Cell Physiology*. 54:1724-1735.
- Nakasugi, K., C.J. Svenson, and B.A. Neilan. 2006. The competence gene, *comF*, from *Synechocystis* sp. strain PCC 6803 is involved in natural transformation, phototactic motility and piliation. *Microbiology*. 152:3623-3631.
- Nanjo, Y., N. Mizusawa, H. Wada, A.R. Slabas, H. Hayashi, and Y. Nishiyama. 2010. Synthesis of fatty acids de novo is required for photosynthetic acclimation of *Synechocystis* sp. PCC 6803 to high temperature. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 1797:1483-1490.
- Nie, L., Y. Ren, and H. Schulz. 2008. Identification and characterization of *Escherichia coli* thioesterase III that functions in fatty acid beta-oxidation. *Biochemistry*. 47:7744-7751.

- Niederholtmeyer, H., B.T. Wolfstader, D.F. Savage, P.A. Silver, and J.C. Way. 2010. Engineering cyanobacteria to synthesize and export hydrophilic products. *Applied and environmental microbiology*. 76:3462-3466.
- Niedermeyer, T.H. 2015. Anti-infective Natural Products from Cyanobacteria. *Planta medica*. 81:1309-1325.
- Nogales, J., S. Gudmundsson, E.M. Knight, B.O. Palsson, and I. Thiele. 2012. Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings of the National Academy of Sciences of the United States of America*. 109:2678-2683.
- Nogales, J., B.O. Palsson, and I. Thiele. 2008. A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: iJN746 as a cell factory. *BMC systems biology*. 2:79.
- Novick, A., and L. Szilard. 1950. Experiments with the Chemostat on spontaneous mutations of bacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 36:708-719.
- Nowicka, B., and J. Kruk. 2016. Cyanobacteria use both p-hydroxybenzoate and homogentisate as a precursor of plastoquinone head group. *Acta Physiologiae Plantarum*. 38:49.
- Okamoto, S., and M. Ohmori. 2002. The cyanobacterial PilT protein responsible for cell motility and transformation hydrolyzes ATP. *Plant & cell physiology*. 43:1127-1136.
- Oliver, J.W., I.M. Machado, H. Yoneda, and S. Atsumi. 2013. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proceedings of the National Academy of Sciences of the United States of America*. 110:1249-1254.
- Olson, J.M. 2006. Photosynthesis in the Archean era. *Photosynthesis research*. 88:109-117.
- Omata, T., S. Gohta, Y. Takahashi, Y. Harano, and S. Maeda. 2001. Involvement of a CbbR homolog in low CO₂-induced activation of the bicarbonate transporter operon in cyanobacteria. *Journal of bacteriology*. 183:1891-1898.
- Ooms, M.D., C.T. Dinh, E.H. Sargent, and D. Sinton. 2016. Photon management for augmented photosynthesis. *Nature communications*. 7:12699.
- Orth, J.D., T.M. Conrad, J. Na, J.A. Lerman, H. Nam, A.M. Feist, and B.O. Palsson. 2011. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism--2011. *Molecular systems biology*. 7:535.
- Orth, J.D., I. Thiele, and B.O. Palsson. 2010. What is flux balance analysis? *Nature biotechnology*. 28:245-248.
- Osterlund, T., I. Nookaew, S. Bordel, and J. Nielsen. 2013. Mapping condition-dependent regulation of metabolism in yeast through genome-scale modeling. *BMC systems biology*. 7:36.
- Palsson, B.Ø. 2006. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press.
- Pech-Canul, A., J. Nogales, A. Miranda-Molina, L. Alvarez, O. Geiger, M.J. Soto, and I.M. Lopez-Lara. 2011. FadD is required for utilization of endogenous fatty acids released from membrane lipids. *Journal of bacteriology*. 193:6295-6304.
- Perez, A.A., Z. Liu, D.A. Rodionov, Z. Li, and D.A. Bryant. 2016. Complementation of Cobalamin Auxotrophy in *Synechococcus* sp. Strain PCC 7002 and Validation of a Putative Cobalamin Riboswitch In Vivo. *Journal of bacteriology*. 198:2743-2752.
- Pfaff, C., N. Glindemann, J. Gruber, M. Frentzen, and R. Sadre. 2014. Chorismate pyruvate-lyase and 4-hydroxy-3-solanesybenzoate decarboxylase are required for plastoquinone biosynthesis in the cyanobacterium *Synechocystis* sp. PCC6803. *The Journal of biological chemistry*. 289:2675-2686.
- Pisciotta, J.M., Y. Zou, and I.V. Baskakov. 2010. Light-dependent electrogenic activity of cyanobacteria. *PLoS one*. 5:e10821.
- Qian, X., M.K. Kim, G. Kenchappa Kumaraswamy, A. Agarwal, D.S. Lun, and G. Charles Dismukes. 2016. Flux balance analysis of photoautotrophic metabolism: Uncovering new biological

- details of subsystems involved in cyanobacterial photosynthesis. *Biochimica et biophysica acta*.
- Rainey, P.B., and M. Travisano. 1998. Adaptive radiation in a heterogeneous environment. *Nature*. 394:69-72.
- Rippka, R., J. Deruelles, J.B. Waterbury, M. Herdman, and R.Y. Stanier. 1979. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology*. 111:1-61.
- Roy, A.S., J. Hazarika, N.A. Manikandan, K. Pakshirajan, and M.B. Syiem. 2015. Heavy metal removal from multicomponent system by the cyanobacterium *Nostoc muscorum*: kinetics and interaction study. *Applied biochemistry and biotechnology*. 175:3863-3874.
- Rozen, D.E., N. Philippe, J. Arjan de Visser, R.E. Lenski, and D. Schneider. 2009. Death and cannibalism in a seasonal environment facilitate bacterial coexistence. *Ecology letters*. 12:34-44.
- Rubin, B.E., K.M. Wetmore, M.N. Price, S. Diamond, R.K. Shultzaberger, L.C. Lowe, G. Curtin, A.P. Arkin, A. Deutschbauer, and S.S. Golden. 2015. The essential gene set of a photosynthetic organism. *Proceedings of the National Academy of Sciences of the United States of America*. 112:E6634-6643.
- Ruffing, A.M. 2011. Engineered cyanobacteria: teaching an old bug new tricks. *Bioengineered bugs*. 2:136-149.
- Ruffing, A.M. 2013a. Borrowing genes from *Chlamydomonas reinhardtii* for free fatty acid production in engineered cyanobacteria. *Journal of Applied Phycology*. 25:1495-1507.
- Ruffing, A.M. 2013b. RNA-Seq analysis and targeted mutagenesis for improved free fatty acid production in an engineered cyanobacterium. *Biotechnology for biofuels*. 6:113.
- Ruffing, A.M. 2014. Improved Free Fatty Acid Production in Cyanobacteria with *Synechococcus* sp. PCC 7002 as Host. *Frontiers in bioengineering and biotechnology*. 2:17.
- Ruffing, A.M., and H.D. Jones. 2012. Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnol Bioeng*. 109:2190-2199.
- Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.A. Rajandream, and B. Barrell. 2000. Artemis: sequence visualization and annotation. *Bioinformatics*. 16:944-945.
- Saha, R., A.T. Versepunt, B.M. Berla, T.J. Mueller, H.B. Pakrasi, and C.D. Maranas. 2012. Reconstruction and comparison of the metabolic potential of cyanobacteria *Cyanothece* sp. ATCC 51142 and *Synechocystis* sp. PCC 6803. *PLoS one*. 7:e48285.
- Saier, M.H., Jr., V.S. Reddy, B.V. Tsu, M.S. Ahmed, C. Li, and G. Moreno-Hagelsieb. 2016. The Transporter Classification Database (TCDB): recent advances. *Nucleic acids research*. 44:D372-379.
- Sakamoto, T., and D.A. Bryant. 1997. Temperature-regulated mRNA accumulation and stabilization for fatty acid desaturase genes in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *Molecular microbiology*. 23:1281-1292.
- Sakamoto, T., and D.A. Bryant. 2002. Synergistic effect of high-light and low temperature on cell growth of the Delta12 fatty acid desaturase mutant in *Synechococcus* sp. PCC 7002. *Photosynthesis research*. 72:231-242.
- Sakamoto, T., S. Higashi, H. Wada, N. Murata, and D.A. Bryant. 1997. Low-temperature-induced desaturation of fatty acids and expression of desaturase genes in the cyanobacterium *Synechococcus* sp. PCC 7002. *FEMS microbiology letters*. 152:313-320.
- Sakamoto, T., D.A. Los, S. Higashi, H. Wada, I. Nishida, M. Ohmori, and N. Murata. 1994. Cloning of omega 3 desaturase from cyanobacteria and its use in altering the degree of membrane-lipid unsaturation. *Plant Mol Biol*. 26:249-263.
- Sakamoto, T., G. Shen, S. Higashi, N. Murata, and D.A. Bryant. 1998. Alteration of low-temperature susceptibility of the cyanobacterium *Synechococcus* sp. PCC 7002 by genetic manipulation of membrane lipid unsaturation. *Archives of microbiology*. 169:20-28.

- Sambrook, J. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- Sangari, F.J., A.M. Cayon, A. Seoane, and J.M. Garcia-Lobo. 2010. Brucella abortus ure2 region contains an acid-activated urea transporter and a nickel transport system. *BMC microbiology*. 10:107.
- Santos, M. 1997. On the contribution of deleterious alleles to fitness variance in natural populations of Drosophila. *Genetical research*. 70:105-115.
- Sauer, J., M. Gori, and K. Forchhammer. 1999. Nitrogen starvation in synechococcus PCC 7942: involvement of glutamine synthetase and NtcA in phycobiliprotein degradation and survival. *Archives of microbiology*. 172:247-255.
- Schafer, L., M. Sandmann, S. Woitsch, and G. Sandmann. 2006. Coordinate up-regulation of carotenoid biosynthesis as a response to light stress in Synechococcus PCC7942. *Plant, cell & environment*. 29:1349-1356.
- Schatz, D., E. Nagar, E. Sendersky, R. Parnasa, S. Zilberman, S. Carmeli, Y. Mastai, E. Shimoni, E. Klein, O. Yeager, Z. Reich, and R. Schwarz. 2013. Self-suppression of biofilm formation in the cyanobacterium Synechococcus elongatus. *Environmental microbiology*. 15:1786-1794.
- Schelin, J., F. Lindmark, and A.K. Clarke. 2002. The clpP multigene family for the ATP-dependent Clp protease in the cyanobacterium Synechococcus. *Microbiology*. 148:2255-2265.
- Schmieder, R., and R. Edwards. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 27:863-864.
- Schopf, J.W. 1993. Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life. *Science*. 260:640-646.
- Schuerger, N., and A. Wilde. 2015. Appendages of the cyanobacterial cell. *Life*. 5:700-715.
- Schulze, P.S., L.A. Barreira, H.G. Pereira, J.A. Perales, and J.C. Varela. 2014. Light emitting diodes (LEDs) applied to microalgal production. *Trends in biotechnology*. 32:422-430.
- Shastri, A.A., and J.A. Morgan. 2005. Flux balance analysis of photoautotrophic metabolism. *Biotechnol Prog*. 21:1617-1626.
- Shen, C.R., and J.C. Liao. 2012. Photosynthetic production of 2-methyl-1-butanol from CO₂ in cyanobacterium Synechococcus elongatus PCC7942 and characterization of the native acetohydroxyacid synthase. *Energy & Environmental Science*. 5:9574-9583.
- Shen, G., H.S. Leonard, W.M. Schluchter, and D.A. Bryant. 2008. CpcM posttranslationally methylates asparagine-71/72 of phycobiliprotein beta subunits in Synechococcus sp. strain PCC 7002 and Synechocystis sp. strain PCC 6803. *Journal of bacteriology*. 190:4808-4817.
- Shestakov, S.V., and N.T. Khyen. 1970. Evidence for genetic transformation in blue-green alga Anacystis nidulans. *Molecular & general genetics : MGG*. 107:372-375.
- Shetty, R.P., D. Endy, and T.F. Knight, Jr. 2008. Engineering BioBrick vectors from BioBrick parts. *Journal of biological engineering*. 2:5.
- Singh, J.S., P.C. Abhilash, H.B. Singh, R.P. Singh, and D.P. Singh. 2011. Genetically engineered bacteria: an emerging tool for environmental remediation and future research perspectives. *Gene*. 480:1-9.
- Smith, R.M., and S.B. Williams. 2006. Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium Synechococcus elongatus. *Proceedings of the National Academy of Sciences of the United States of America*. 103:8564-8569.
- Soding, J., A. Biegert, and A.N. Lupas. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research*. 33:W244-248.
- Sokol, K.A. 2015. Characterization and cellular roles of a bacterial O-GlcNAc transferase in Synechococcus elongatus PCC7942 (Doctoral dissertation). Retrieved from the University of Minnesota Digital Conservancy.

- Spiers, A.J., S.G. Kahn, J. Bohannon, M. Travisano, and P.B. Rainey. 2002. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics*. 161:33-46.
- Stal, L.J., and R. Moezelaar. 1997. Fermentation in cyanobacteria. *FEMS Microbiology Reviews*. 21:179-211.
- Stanier, R.Y., and G. Cohen-Bazire. 1977. Phototrophic prokaryotes: the cyanobacteria. *Annual review of microbiology*. 31:225-274.
- Stapleton, S.R., and J.G. Jaworski. 1984. Characterization and purification of malonyl-coenzyme A:[acyl-carrier-protein] transacylases from spinach and *Anabaena variabilis*. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*. 794:240-248.
- Stevens, S.E., and R.D. Porter. 1980. Transformation in *Agmenellum quadruplicatum*. *Proceedings of the National Academy of Sciences of the United States of America*. 77:6052-6056.
- Stucken, K., R. Koch, and T. Dagan. 2013. Cyanobacterial defense mechanisms against foreign DNA transfer and their impact on genetic engineering. *Biological Research*. 46:373-382.
- Subrahmanyam, S., and J.E. Cronan, Jr. 1998. Overproduction of a functional fatty acid biosynthetic enzyme blocks fatty acid synthesis in *Escherichia coli*. *Journal of bacteriology*. 180:4596-4602.
- Sudesh, K., K. Taguchi, and Y. Doi. 2002. Effect of increased PHA synthase activity on polyhydroxyalkanoates biosynthesis in *Synechocystis* sp. PCC6803. *International journal of biological macromolecules*. 30:97-104.
- Sugita, C., K. Ogata, M. Shikata, H. Jikuya, J. Takano, M. Furumichi, M. Kanehisa, T. Omata, M. Sugiura, and M. Sugita. 2007. Complete nucleotide sequence of the freshwater unicellular cyanobacterium *Synechococcus elongatus* PCC 6301 chromosome: gene content and organization. *Photosynthesis research*. 93:55-67.
- Suzuki, E., H. Ohkawa, K. Moriya, T. Matsubara, Y. Nagaike, I. Iwasaki, S. Fujiwara, M. Tsuzuki, and Y. Nakamura. 2010. Carbohydrate metabolism in mutants of the cyanobacterium *Synechococcus elongatus* PCC 7942 defective in glycogen synthesis. *Applied and environmental microbiology*. 76:3153-3159.
- Takahama, K., M. Matsuoka, K. Nagahama, and T. Ogawa. 2003. Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the *psbAI* locus. *Journal of bioscience and bioengineering*. 95:302-305.
- Takahama, K., M. Matsuoka, K. Nagahama, and T. Ogawa. 2004. High-frequency gene replacement in cyanobacteria using a heterologous *rps12* gene. *Plant & cell physiology*. 45:333-339.
- Takai, N., M. Nakajima, T. Oyama, R. Kito, C. Sugita, M. Sugita, T. Kondo, and H. Iwasaki. 2006. A KaiC-associating SasA-RpaA two-component regulatory system as a major circadian timing mediator in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 103:12109-12114.
- Tamagnini, P., O. Troshina, F. Oxelfelt, R. Salema, and P. Lindblad. 1997. Hydrogenases in *Nostoc* sp. Strain PCC 73102, a Strain Lacking a Bidirectional Enzyme. *Applied and environmental microbiology*. 63:1801-1807.
- Tandeau de Marsac, N., and J. Houmard. 1993. Adaptation of cyanobacteria to environmental stimuli: new steps towards molecular mechanisms. *FEMS microbiology letters*. 104:119-189.
- Tchernov, D., Y. Helman, N. Keren, B. Luz, I. Ohad, L. Reinhold, T. Ogawa, and A. Kaplan. 2001. Passive entry of CO₂ and its energy-dependent intracellular conversion to HCO₃⁻ in cyanobacteria are driven by a photosystem I-generated $\Delta\mu\text{H}^+$. *The Journal of biological chemistry*. 276:23450-23455.
- terHorst, C.P. 2011. Experimental evolution of protozoan traits in response to interspecific competition. *Journal of evolutionary biology*. 24:36-46.

- Thiel, T., and H. Poo. 1989. Transformation of a filamentous cyanobacterium by electroporation. *Journal of bacteriology*. 171:5743-5746.
- Thiele, I., and B.O. Palsson. 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols*. 5:93-121.
- Tillich, U.M., N. Wolter, K. Schulze, D. Kramer, O. Brodel, and M. Frohme. 2014. High-throughput cultivation and screening platform for unicellular phototrophs. *BMC microbiology*. 14:239.
- Tocher, D.R., M.J. Leaver, and P.A. Hodgson. 1998. Recent advances in the biochemistry and molecular biology of fatty acyl desaturases. *Progress in lipid research*. 37:73-117.
- Toomey, R.E., and S.J. Wakil. 1966. Studies on the mechanism of fatty acid synthesis. XV. Preparation and general properties of beta-ketoacyl acyl carrier protein reductase from *Escherichia coli*. *Biochimica et biophysica acta*. 116:189-197.
- Treves, D.S., S. Manning, and J. Adams. 1998. Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Molecular Biology and Evolution*. 15:789-797.
- Triana, J., A. Montagud, M. Siurana, D. Fuente, A. Urchueguia, D. Gamermann, J. Torres, J. Tena, P.F. de Cordoba, and J.F. Urchueguia. 2014. Generation and Evaluation of a Genome-Scale Metabolic Network Model of *Synechococcus elongatus* PCC7942. *Metabolites*. 4:680-698.
- Tsay, J.T., W. Oh, T.J. Larson, S. Jackowski, and C.O. Rock. 1992. Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *The Journal of biological chemistry*. 267:6807-6814.
- Tsinoremas, N.F., A.K. Kutach, C.A. Strayer, and S.S. Golden. 1994. Efficient gene transfer in *Synechococcus* sp. strains PCC 7942 and PCC 6301 by interspecies conjugation and chromosomal recombination. *Journal of bacteriology*. 176:6764-6768.
- Tu, X., P.A. Hubbard, J.J. Kim, and H. Schulz. 2008. Two distinct proton donors at the active site of *Escherichia coli* 2,4-dienoyl-CoA reductase are responsible for the formation of different products. *Biochemistry*. 47:1167-1175.
- Uchiyama, J., Y. Kanesaki, N. Iwata, R. Asakura, K. Funamizu, R. Tasaki, M. Agatsuma, H. Tahara, A. Matsushashi, H. Yoshikawa, S. Ogawa, and H. Ohta. 2015. Genomic analysis of parallel-evolved cyanobacterium *Synechocystis* sp. PCC 6803 under acid stress. *Photosynthesis research*. 125:243-254.
- Ungerer, J., L. Tao, M. Davis, M. Ghirardi, P.-C. Maness, and J. Yu. 2012. Sustained photosynthetic conversion of CO₂ to ethylene in recombinant cyanobacterium *Synechocystis* 6803. *Energy & Environmental Science*. 5:8998-9006.
- van Waasbergen, L.G., N. Dolganov, and A.R. Grossman. 2002. *nbIS*, a gene involved in controlling photosynthesis-related gene expression during high light and nutrient stress in *Synechococcus elongatus* PCC 7942. *Journal of bacteriology*. 184:2481-2490.
- Velicer, G.J., and Y.T. Yu. 2003. Evolution of novel cooperative swarming in the bacterium *Myxococcus xanthus*. *Nature*. 425:75-78.
- Verwoert, II, E.F. Verhagen, K.H. van der Linden, E.C. Verbree, H.J. Nijkamp, and A.R. Stuitje. 1994. Molecular characterization of an *Escherichia coli* mutant with a temperature-sensitive malonyl coenzyme A-acyl carrier protein transacylase. *FEBS letters*. 348:311-316.
- Vijayan, V., I.H. Jain, and E.K. O'Shea. 2011. A high resolution map of a cyanobacterial transcriptome. *Genome biology*. 12:R47.
- Vu, T.T., S.M. Stolyar, G.E. Pinchuk, E.A. Hill, L.A. Kucek, R.N. Brown, M.S. Lipton, A. Osterman, J.K. Fredrickson, A.E. Konopka, A.S. Beliaev, and J.L. Reed. 2012. Genome-scale modeling of light-driven reductant partitioning and carbon fluxes in diazotrophic unicellular cyanobacterium *Cyanothece* sp. ATCC 51142. *PLoS computational biology*. 8:e1002460.

- Wada, H., Z. Compos, and N. Murata. 1990. Enhancement of chilling tolerance of a cyanobacterium by genetic manipulation of fatty acid desaturation. *Nature*. 347:200-203.
- Walter, J., R.A. Britton, and S. Roos. 2011. Host-microbial symbiosis in the vertebrate gastrointestinal tract and the *Lactobacillus reuteri* paradigm. *Proceedings of the National Academy of Sciences of the United States of America*. 108 Suppl 1:4645-4652.
- Wallace-Salinas, V., and M.F. Gorwa-Grauslund. 2013. Adaptive evolution of an industrial strain of *Saccharomyces cerevisiae* for combined tolerance to inhibitors and temperature. *Biotechnology for biofuels*. 6:151.
- Wang, Y., F. Tao, J. Ni, C. Li, and P. Xu. 2015. Production of C3 platform chemicals from CO₂ by genetically engineered cyanobacteria. *Green Chemistry*. 17:3100-3110.
- Waterbury, J.B. 2006. The Cyanobacteria—Isolation, Purification and Identification. In *The Prokaryotes: Volume 4: Bacteria: Firmicutes, Cyanobacteria*. M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt, editors. Springer US, New York, NY. 1053-1073.
- Webb, R., T. Troyan, D. Sherman, and L.A. Sherman. 1994. MapA, an iron-regulated, cytoplasmic membrane protein in the cyanobacterium *Synechococcus* sp. strain PCC7942. *Journal of bacteriology*. 176:4906-4913.
- Weimar, J.D., C.C. DiRusso, R. Delio, and P.N. Black. 2002. Functional role of fatty acyl-coenzyme A synthetase in the transmembrane movement and activation of exogenous long-chain fatty acids. Amino acid residues within the ATP/AMP signature motif of *Escherichia coli* FadD are required for enzyme activity and fatty acid transport. *The Journal of biological chemistry*. 277:29369-29376.
- White, S.W., J. Zheng, Y.M. Zhang, and Rock. 2005. The structural biology of type II fatty acid biosynthesis. *Annual review of biochemistry*. 74:791-831.
- Wilhelm, S.W., and C.G. Trick. 1995. Effects of vitamin B12 concentration on chemostat cultured *Synechococcus* sp. strain PCC 7002. *Canadian Journal of Microbiology*. 41:145-151.
- Wilmotte, A.M.R., and W.T. Stam. 1984. Genetic Relationships among Cyanobacterial Strains Originally Designated as '*Anacystis nidulans*' and Some Other *Synechococcus* Strains. *Microbiology*. 130:2737-2740.
- Williams, S.B. 2007. A circadian timing mechanism in the cyanobacteria. *Advances in microbial physiology*. 52:229-296.
- Winkler, J.D., and K.C. Kao. 2014. Recent advances in the evolutionary engineering of industrial biocatalysts. *Genomics*. 104:406-411.
- Wolk, C.P., A. Vonshak, P. Kehoe, and J. Elhai. 1984. Construction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 81:1561-1565.
- Woods, R.J., J.E. Barrick, T.F. Cooper, U. Shrestha, M.R. Kauth, and R.E. Lenski. 2011. Second-order selection for evolvability in a large *Escherichia coli* population. *Science*. 331:1433-1436.
- Work, V.H., M.R. Melnicki, E.A. Hill, F.K. Davies, L.A. Kucek, A.S. Beliaev, and M.C. Posewitz. 2015. Lauric Acid Production in a Glycogen-Less Strain of *Synechococcus* sp. PCC 7002. *Frontiers in bioengineering and biotechnology*. 3:48.
- Yang, C., Q. Hua, and K. Shimizu. 2002. Metabolic Flux Analysis in *Synechocystis* Using Isotope Distribution from ¹³C-Labeled Glucose. *Metabolic engineering*. 4:202-216.
- Yang, S.Y., J.M. Li, X.Y. He, S.D. Cosloy, and H. Schulz. 1988. Evidence that the *fadB* gene of the *fadAB* operon of *Escherichia coli* encodes 3-hydroxyacyl-coenzyme A (CoA) epimerase, delta 3-cis-delta 2-trans-enoyl-CoA isomerase, and enoyl-CoA hydratase in addition to 3-hydroxyacyl-CoA dehydrogenase. *Journal of bacteriology*. 170:2543-2548.

- Yang, Y., J. Feng, T. Li, F. Ge, and J. Zhao. 2015. CyanOmics: an integrated database of omics for the model cyanobacterium *Synechococcus* sp. PCC 7002. *Database : the journal of biological databases and curation*. 2015.
- Yao, Z., R.M. Davis, R. Kishony, D. Kahne, and N. Ruiz. 2012. Regulation of cell size in response to nutrient availability by fatty acid biosynthesis in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. 109:E2561-2568.
- Yim, H., R. Haselbeck, W. Niu, C. Pujol-Baxley, A. Burgard, J. Boldt, J. Khandurina, J.D. Trawick, R.E. Osterhout, and R. Stephen. 2011. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol*. 7.
- Yoshihara, S., X. Geng, S. Okamoto, K. Yura, T. Murata, M. Go, M. Ohmori, and M. Ikeuchi. 2001. Mutational analysis of genes involved in pilus structure, motility and transformation competency in the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803. *Plant & cell physiology*. 42:63-73.
- Yoshikawa, K., S. Aikawa, Y. Kojima, Y. Toya, C. Furusawa, A. Kondo, and H. Shimizu. 2015. Construction of a Genome-Scale Metabolic Model of *Arthrospira platensis* NIES-39 and Metabolic Design for Cyanobacterial Bioproduction. *PLoS one*. 10:e0144430.
- Yoshikawa, K., Y. Kojima, T. Nakajima, C. Furusawa, T. Hirasawa, and H. Shimizu. 2011. Reconstruction and verification of a genome-scale metabolic model for *Synechocystis* sp. PCC6803. *Appl Microbiol Biotechnol*. 92:347-358.
- Yu, J., M. Liberton, P.F. Cliften, R.D. Head, J.M. Jacobs, R.D. Smith, D.W. Koppelaar, J.J. Brand, and H.B. Pakrasi. 2015. *Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO₂. *Scientific reports*. 5:8132.
- Zbinden, M., C.R. Haag, and D. Ebert. 2008. Experimental evolution of field populations of *Daphnia magna* in response to parasite treatment. *Journal of evolutionary biology*. 21:1068-1078.
- Zhang, C., and Q. Hua. 2015. Applications of Genome-Scale Metabolic Models in Biotechnology and Systems Medicine. *Frontiers in physiology*. 6:413.
- Zhang, J., M. Liu, H. Jin, L. Deng, J. Xing, and A. Dong. 2010. In vitro enhancement of lactate esters on the percutaneous penetration of drugs with different lipophilicity. *AAPS PharmSciTech*. 11:894-903.
- Zhao, F., X. Zhang, C. Liang, J. Wu, Q. Bao, and S. Qin. 2006. Genome-wide analysis of restriction-modification system in unicellular and filamentous cyanobacteria. *Physiological genomics*. 24:181-190.