

## Trabajo Fin de Grado

Identificación de genes implicados en la síntesis de biofilms en *Anabaena sp.* PCC7120 y análisis de su modulación por la familia de proteínas FUR (*ferric uptake regulator*)

Identification of genes involved in biofilm formation in *Anabaena sp.* PCC7120 and analysis of their regulation by the FUR (ferric uptake regulator) protein family

Autor

Germán Alonso Tolo

Directores

María Francisca Fillat Castejón

Jorge Guío Martínez





**Departamento de  
Bioquímica y Biología  
Molecular y Celular  
Universidad Zaragoza**

María Francisca Fillat Castejón y Jorge Guío Martínez

CERTIFICAN

Que la memoria titulada “Identificación de genes implicados en la síntesis de biofilms en *Anabaena* sp. PCC7120 y análisis de su modulación por la familia de proteínas FUR (*ferric uptake regulator*)” presentada por Germán Alonso Tolo ha sido realizada bajo su dirección, en el Departamento de Bioquímica y Biología Molecular.

A su juicio, reúne todos los requisitos para ser presentada por su autor para la defensa de su Trabajo Fin de Grado

Zaragoza a 25 de junio de 2021

María Francisca Fillat Castejón

Jorge Guío Martínez



## **Agradecimientos**

Antes de abordar la memoria de mi Trabajo de Fin de Grado me gustaría mostrar mi agradecimiento a aquellas personas que lo han hecho posible. A María Fillat, Marisa Peleato y Emma Sevilla, por enseñarme tanto a lo largo de estos años y lograr que la Ciencia me apasione. A Jorge Guío, por su infinita paciencia conmigo, por lo mucho que me ha enseñado y por su constante esfuerzo para que este Trabajo saliese adelante. Y a Ainhoa, por hacer más amenas todas estas horas en el laboratorio y conseguir que los interminables pipeteos de agua se hiciesen entretenidos.

A mis padres, Nieves y Germán, por su esfuerzo y cariño y por su capacidad para soportarme incluso en mis peores momentos de agobio. Y a mis abuelos, Ramón y Avelina, por quererme y cuidarme siempre, por creer incondicionalmente en mí y por haberme hecho el niño y la persona más feliz del mundo. Muchas gracias por todo.



# Índice

<b>1. Resumen/abstract</b>	<b>1</b>
1.1. Resumen . . . . .	1
1.2. Abstract . . . . .	1
<b>2. Introducción</b>	<b>2</b>
2.1. Las cianobacterias . . . . .	2
2.2. Los biofilms en cianobacterias . . . . .	2
2.2.1. Composición y características de los biofilms en cianobacterias . . . . .	2
2.2.2. Mecanismos implicados en la producción de biofilms en cianobacterias . . . . .	3
2.2.3. Interés biológico del estudio de la producción de biofilms en cianobacterias . . . . .	5
2.3. Las proteínas FUR y su relación con la formación de biofilms . . . . .	6
<b>3. Objetivos</b>	<b>7</b>
<b>4. Materiales y métodos</b>	<b>7</b>
4.1. Técnicas de trabajo con DNA . . . . .	7
4.1.1. Reacción en cadena de la polimerasa (PCR) . . . . .	7
4.1.2. Electroforesis de DNA . . . . .	7
4.1.3. Purificación y cuantificación de los productos de PCR . . . . .	7
4.2. Ensayos de retardo en gel (EMSA) . . . . .	8
4.3. Herramientas bioinformáticas . . . . .	8
<b>5. Resultados y discusión</b>	<b>9</b>
5.1. Identificación de genes potencialmente implicados en la formación de biofilms en <i>Anabaena sp.</i> PCC 7120 . . . . .	9
5.2. Análisis bioinformático de la presencia de cajas FUR en regiones promotoras de genes potencialmente implicados en la formación de biofilms en <i>Anabaena sp.</i> PCC7120 . . . . .	13
5.3. Estudio de la modulación de genes potencialmente implicados en la formación de biofilms en <i>Anabaena sp.</i> PCC7120 por las proteínas FUR . . . . .	16
5.4. Discusión global y perspectivas futuras . . . . .	22
<b>6. Conclusiones</b>	<b>23</b>
<b>7. Bibliografía</b>	<b>24</b>





## Abreviaturas

- FUR: *ferric uptake regulator protein family*.
- Fur: *ferric uptake regulator*.
- Zur: *zinc uptake regulator*.
- PerR: *peroxide stress regulator*.
- EMSA: ensayo de retardo en gel.
- EPS: sustancias poliméricas extracelulares.
- Kdo: ácido cetodesoxioctonoico.
- PCR: reacción en cadena de la polimerasa.
- dNTP: desoxinucleótidos trifosfato.
- TBE: Tris, borato, EDTA.
- EDTA: ácido etilendiaminotetraacético.
- TEMED: tetramethylethylenediamine.
- DTT: 1,4-ditiotreitol.
- PSA: persulfato de amonio.
- BLAST: basic local alignment search tool.



## 1. Resumen/abstract

### 1.1. Resumen

Las cianobacterias son microorganismos fotosintéticos surgidos hace 3.500 millones de años y que, como otras bacterias, pueden crecer formando biofilms, agregados microbianos en los que viven rodeadas de una matriz de sustancias poliméricas extracelulares. Estos biofilms cianobacterianos son interesantes por sus aplicaciones en el tratamiento de aguas residuales, la biorremediación o la producción de biocombustibles.

Aunque los mecanismos reguladores de la formación de biofilms en bacterias heterótrofas han sido caracterizados, en el caso de las cianobacterias el conocimiento es todavía limitado. Las cianobacterias presentan una familia de reguladores transcripcionales, las proteínas FUR (*ferric uptake regulator*), implicados en la respuesta a estreses abióticos. Controlan principalmente la homeostasis de metales como el hierro y el zinc, aunque también otros procesos como la respuesta frente a estrés oxidativo o la deficiencia de nitrógeno. Dado que los biofilms se sintetizan en respuesta a condiciones de estrés, se ha estudiado la modulación de su formación en la cianobacteria *Anabaena sp.* PCC7120 por parte de las proteínas FUR.

La realización de una revisión bibliográfica junto con estudios de homología han permitido identificar más de 100 genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120. Además, se ha analizado mediante ensayos de retardo en gel (EMSA) la regulación de estos genes por parte de las proteínas FUR, lo que ha permitido identificar 54 genes directamente regulados por estas proteínas. Se pone así de manifiesto que las proteínas FUR juegan un papel clave en la regulación de la formación de biofilms en *Anabaena sp.* PCC7120.

### 1.2. Abstract

Cyanobacteria are photosynthetic microorganisms which appeared 3.500 million years ago. As many other bacteria, cyanobacteria can grow forming biofilms, microbial aggregates in which cyanobacteria live surrounded by a matrix of extracellular polymeric substances. These cyanobacterial biofilms are interesting because of their applications in residual water treatment, bioremediation or in the production of biofuels.

Although the mechanisms involved in the regulation of biofilm formation in heterotrophic bacteria are precisely described, little is known about the factors controlling biofilm formation in cyanobacteria. Cyanobacteria present a family of transcriptional regulators, FUR (*ferric uptake regulator*) proteins, which participate in the abiotic stress response. FUR proteins mainly control the homeostasis of metals like iron and zinc, but also other events such as the response against oxidative stress or nitrogen starvation. As biofilms are formed under stress conditions, the regulation of their formation by FUR proteins has been studied in the cyanobacteria *Anabaena sp.* PCC7120.

The bibliographic revision performed, altogether with homology studies, allowed identifying more than 100 genes potentially involved in biofilm formation in *Anabaena sp.* PCC7120. Moreover, the regulation of these genes by FUR proteins was studied by electrophoretic mobility shift assays (EMSA), leading to the identification of 54 genes directly regulated by these proteins. Thus, the regulating role of FUR proteins in biofilm formation in *Anabaena sp.* PCC7120 is demonstrated.

## 2. Introducción

### 2.1. Las cianobacterias

Las cianobacterias son organismos procariotas fotosintéticos que habitan la mayoría de los ambientes con luz. Surgieron hace unos 3.500 millones de años y su aparición supuso un cambio en la atmósfera, pues comenzaron a liberar  $O_2$  a la misma (1). Estos organismos se encuentran entre los más abundantes del planeta y su crecimiento está principalmente limitado por la disponibilidad de fósforo y hierro (2). Por el contrario, ante una deficiencia de nitrógeno, muchas cianobacterias son capaces de fijar  $N_2$  atmosférico mediante la enzima nitrogenasa, sensible al  $O_2$ . Dado que en la fotosíntesis se produce  $O_2$ , las cianobacterias han desarrollado dos estrategias para que la nitrogenasa no se vea afectada. Hay cianobacterias que realizan la fotosíntesis durante el día y fijan  $N_2$  por la noche, mientras que algunas cianobacterias filamentosas separan el proceso espacialmente y fijan el  $N_2$  en el heterocisto, una célula especializada con el ambiente anaerobio necesario para ello (3). Esta capacidad para fijar el  $N_2$  permite a las cianobacterias formar asociaciones simbióticas con otros organismos, a los que aportan materia nitrogenada (1).

La inoculación de cianobacterias se ha aplicado a la mejora de suelos áridos y a la fertilización de la tierra, sobre todo mediante su enriquecimiento en N (1). En esta fertilización es habitual el uso del helecho *Azolla* en simbiosis con la cianobacteria *Anabaena azollae* (4). A pesar de sus utilidades, las cianobacterias también pueden resultar perjudiciales. Es el caso de los *blooms*, crecimientos desmedidos de estos organismos. Destacan los de *Microcystis spp.*, peligrosos por la presencia de toxinas dañinas para humanos, mamíferos y otros organismos (5).

### 2.2. Los biofilms en cianobacterias

Los microorganismos, en la naturaleza, pueden vivir formando biofilms, agregados microbianos que habitualmente se forman en la interfase entre un sólido y un líquido (6). En un único biofilm habitan distintas especies bacterianas formando consorcios estables y dentro de ellos se generan gradientes físico-químicos y son habituales los fenómenos de comunicación celular y de transferencia génica horizontal (7).

#### 2.2.1. Composición y características de los biofilms en cianobacterias

Las bacterias de un biofilm suponen menos de un 10 % del peso seco del mismo al estar formados principalmente por una matriz de sustancias poliméricas extracelulares (EPS). Estos EPS son principalmente exopolisacáridos, y pueden permanecer unidos covalentemente a la superficie celular o ser liberados al medio (7). Entre los polisacáridos de los biofilms aparecen, por una parte, homopolisacáridos sintetizados a partir de sacarosa mediante actividades enzimáticas tipo invertasa, y, por otra parte, heteropolisacáridos, constituidos por diversos tipos de monosacáridos y que son sintetizadas por de enzimas del grupo de las glicosiltransferasa (8).

En el caso de las cianobacterias, los EPS están constituidos principalmente por heteropolisacáridos (8) y tienen características distintivas del resto de EPS bacterianos. Por una parte, presentan dos moléculas de ácido urónico, y es habitual también la presencia de grupos sulfato, lo cual confiere a los EPS una naturaleza aniónica que les permite actuar como quelantes de metales (9). Sin embargo, aproximadamente un 12 % del peso seco de los EPS está constituido por polisacáridos de naturaleza hidrófoba como consecuencia de la esterificación de grupos acetilo o la presencia de dominios peptídicos o deoxiazúcares como fucosa o ramnosa (8).

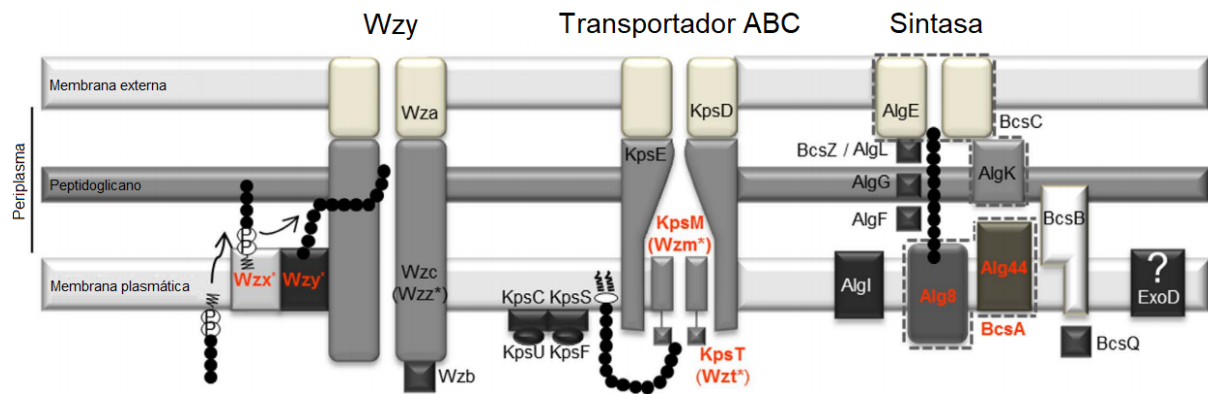
De los heteropolisacáridos que constituyen los EPS cianobacterianos, aproximadamente el 75 % contienen 6 o más tipos de monosacáridos (10). En total, se han identificado 12 tipos de monosacáridos que constituyen estos polímeros, entre los que aparecen las hexosas glucosa, galactosa, manosa y fructosa, las pentosas ribosa, xilosa y arabinosa, las deoxihexosas fucosa, ramnosa y metil-ramnosa, y los ácidos glucurónico y galacturónico (10). En los EPS, además, aparecen moléculas de otra naturaleza, como polipéptidos y lípidos. Los polipéptidos son ricos en glicina, alanina, valina, leucina, isoleucina y fenilalanina, confieren estabilidad mecánica al biofilm y aumentan su capacidad adherente (11). Respecto a los lípidos, permiten que las células resistan a las altas tensiones superficiales del agua que rodea el biofilm, facilitando así el crecimiento bacteriano en superficies (11).

La producción de todas estas moléculas y la consecuente formación de biofilms es uno de los principales motivos que permite a las cianobacterias sobrevivir en ambientes hostiles (8). Por una parte, la producción de EPS permite a las cianobacterias sobrevivir a la desecación, pues sus componentes hidrófilos retienen agua y regulan la pérdida y captación de la misma (12). Además, a los biofilms se les asocian otros papeles protectores. Por una parte, las propiedades hidrófobas de algunos EPS se asocian con la capacidad de las cianobacterias para crecer sobre superficies sin ser arrastradas por el agua (13). Por otra parte, se ha observado un efecto protector frente a la biomineralización (14), y la presencia de pigmentos y aminoácidos tipo microsporina sugieren que el biofilm pueda proteger a las células de las radiaciones ultravioleta (15). Finalmente, los EPS pueden jugar un papel relevante en el control de la homeostasis de los metales. El carácter aniónico de muchos EPS podría favorecer el enriquecimiento del medio en metales escasos y esenciales para las cianobacterias (16), y del mismo modo podría permitir quelar metales tóxicos para estas (17).

La única función de los biofilms cianobacterianos plenamente confirmada empíricamente es la de protección frente a la desecación (12). El resto de funciones, por el contrario, siguen estudiándose. Del mismo modo, la información sobre el proceso de formación de biofilms es escasa, y siguen investigándose activamente los mecanismos implicados en el proceso, lo cual ha dado lugar a la descripción de varios modelos que explican la síntesis y secreción de los EPS.

### 2.2.2. Mecanismos implicados en la producción de biofilms en cianobacterias

Son numerosos los estudios que tratan de describir la genética y la bioquímica del proceso de producción de EPS en bacterias. En cianobacterias, sin embargo, este campo no ha sido explorado exhaustivamente, y, como se ha mencionado, la información disponible es todavía limitada. Los estudios llevados a cabo en bacterias muestran una gran variabilidad en las vías de producción de EPS, pero sin embargo los mecanismos implicados en todos estos procesos parecen estar conservados evolutivamente. Tradicionalmente, se describen cuatro etapas en el proceso de síntesis y exportación de los EPS (8). En primer lugar, los monosacáridos se activan en el citosol y, tras ello, se forman las unidades repetidas por la unión de los monosacáridos a un *carrier* por acción de glicosiltransferasas a nivel de la membrana plasmática. Posteriormente, estas unidades repetidas se polimerizan en el periplasma y el polímero formado se exporta finalmente a la superficie celular (8). Comparando la información disponible sobre la síntesis y secreción de EPS en bacterias y la información genética de diversas cianobacterias, se han propuesto tres sistemas para la producción de EPS en cianobacterias: el sistema Wzy dependiente, el sistema dependiente de transportadores de tipo ABC y el sistema dependiente de sintasa (Figura 1).



**Figura 1.** Funcionamiento de los sistemas Wzy dependiente, dependiente de transportador ABC y dependiente de sintasa. Las proteínas marcadas con un asterisco participan también en la síntesis del antígeno O, y las marcadas con una interrogación tienen una función desconocida. Adaptado de Pereira *et al.* (18).

En el caso del sistema Wzy dependiente, las glicosiltransferasas transfieren los azúcares activados a un *carrier* lipídico, formándose así las unidades repetidas de azúcares que conformarán posteriormente los EPS (8). Estas unidades de azúcares unidas al *carrier* son volteadas por una proteína Wzx, de modo que pasan a localizarse en el periplasma. Allí, la proteína Wzy une estas unidades repetidas de azúcares, formándose el polisacárido que posteriormente será secretado (8). Este proceso requiere de otra proteína auxiliar, Wzc, que parece controlar la longitud del polímero en síntesis. La proteína Wzc presenta actividad ATPasa y autoquinasa, y su estado de fosforilación está además regulado por Wzb, una fosfatasa (19). Finalmente, implicada en este sistema aparece Wza, que forma un canal a través del que se exporta el polisacárido a la superficie celular. Este proceso parece requerir de la unión física de Wza y Wzc (8).

En el caso del sistema dependiente de transportador ABC, los polisacáridos se sintetizan completamente en el citoplasma. En este sistema, el *carrier* al que se unen los polisacáridos es una molécula de ácido 3-deoxi-D-mano-octo-2-ulosónico (ácido cetodesoxioctonoico, Kdo), que es sintetizado por las proteínas KpsC, KpsF, KpsS y KpsU. Una vez ha sido sintetizado, el polisacárido es exportado por el transportador ABC, constituido por las proteínas KpsM y KpsT, que permiten su salida al periplasma. Desde el periplasma, finalmente, el polisacárido es externalizado a través de las proteínas KpsE y KpsD, una polisacárido copolimerasa y una proteína de la membrana externa de exportación de polisacáridos, respectivamente (18).

Finalmente, se encuentra el sistema dependiente de sintasa, descrito para la síntesis de alginato y celulosa. En este caso, la polimerización y exportación de los polisacáridos se dan simultáneamente y los azúcares son activados mediante la unión de una molécula de di-GMP. Estos azúcares activados son reconocidos por Alg44 en el caso de la síntesis de alginato y por BcsA en el caso de la síntesis de celulosa mediante motivos de unión a c-di-GMP, y son transferidos a Alg8 o BcsA, las sintasas de alginato y celulosa, respectivamente, que unen el azúcar a la cadena de polisacárido en crecimiento. En el caso de la síntesis de alginato están implicadas las proteínas AlgF, AlgG y AlgI, que modifican el polímero en el periplasma, y la proteína AlgL que degrada el alginato que se pueda acumular. En la síntesis de celulosa aparece una proteína BcsZ que cumple con la misma función que AlgL, así como otras proteínas esenciales para el proceso pero con función desconocida, BcsB y BcsQ. Finalmente, la exportación del alginato depende de la proteína de andamiaje AlgK y de la porina AlgE, mientras que en el caso de la celulosa el proceso está mediado por BcsC, que combina ambas funciones (18).

### 2.2.3. Interés biológico del estudio de la producción de biofilms en cianobacterias

El estudio de los biofilms y su producción resulta muy interesante por las numerosas aplicaciones de los mismos. Una de las más importantes es el tratamiento de aguas residuales (20). Las cianobacterias asimilan los nutrientes de estas aguas y, junto con el carbono fijado a partir del  $\text{CO}_2$ , producen biomasa. De este modo, compuestos como amonio, nitratos y nitritos, urea y aminoácidos son captados por las cianobacterias y, en consecuencia, el nitrógeno no se disipa en la atmósfera sino que queda retenido en forma de biomasa. La utilización de biofilms en el tratamiento de estas aguas resulta muy ventajosa, pues elimina los problemas de separación de las algas en suspensión y las aguas ya tratadas. Además de eliminar los compuestos nitrogenados, el uso de biofilms cianobacterianos permite atajar otros problemas. Su actividad fotosintética produce un aumento del pH y, en consecuencia, precipitan los fosfatos y disminuye la viabilidad de las bacterias fecales (21).

El hecho de que el nitrógeno de las aguas quede retenido en los biofilms lleva a una segunda aplicación de los mismos, en este caso a nivel agrícola. Los biofilms cianobacterianos pueden utilizarse directamente como fertilizantes o pueden agregarse a los suelos para enriquecerlos en nitrógeno gracias a la capacidad de algunas cianobacterias para fijar el  $\text{N}_2$  atmosférico. Además, se ha visto que los EPS mejoran las propiedades del suelo, aumentando su capacidad de retención de agua y su resistencia a la erosión, lo cual contribuye a evitar la desertificación (21). Las cianobacterias, además, secretan fitohormonas como el ácido indolacético y otros metabolitos secundarios que inducen las respuestas contra patógenos en las plantas, protegiéndolas así de plagas y enfermedades. Esta producción de fitohormonas se ve aumentada en los biofilms (22). Además, los biofilms cianobacterianos son ricos en nutrientes, de modo que constituyen un nicho idóneo para el desarrollo de bacterias heterótrofas, pudiendo promover así el crecimiento de bacterias beneficiosas como las del género *Rhizobium* (22).

Del mismo modo que la biomasa producida en el tratamiento de aguas residuales puede utilizarse como fertilizante, esta puede utilizarse también como biomasa. Aparece así otro de los usos de los biofilms cianobacterianos, la producción de biocombustibles, que permite además revalorizar los productos derivados del tratamiento de aguas. En este contexto, además, es de interés de la producción de  $\text{H}_2$ , una alternativa limpia a los combustibles fósiles, pues su combustión genera únicamente agua. Las cianobacterias son organismos prometedores en la producción de hidrógeno, ya sea mediante una actividad hidrogenasa o mediante la actividad nitrogenasa responsable de la fijación del  $\text{N}_2$ , y que libera también  $\text{H}_2$  (23). La producción de  $\text{H}_2$  podría mejorarse notablemente si se llevase a cabo en biofilms, al poder construirse ensamblados bien definidos que contuviesen las poblaciones de cianobacterias adecuadas y con las modificaciones genéticas necesarias para la optimización del proceso (21).

Resulta también de interés la aplicación de los biofilms cianobacterianos a nivel de biorremediación. Por una parte, el carácter aniónico de los EPS permite utilizarlos como quelantes de metales pesados para retirar estos elementos de los medios (21). Se ha observado también la capacidad de los biofilms cianobacterianos para oxidar pesticidas y herbicidas, así como para acumular compuestos insecticidas. Es interesante, del mismo modo, la aplicación de los biofilms en la eliminación de vertidos de petróleo, no tanto por su acción directa sobre el proceso sino por su efecto beneficioso sobre el crecimiento de las bacterias que encargadas de la degradación de los hidrocarburos, a las que proveen de una fuente de nitrógeno gracias su capacidad para fijar el  $\text{N}_2$  atmosférico (22).

Dadas las múltiples aplicaciones de los biofilms y el escaso conocimiento de sus vías de producción y de la regulación de su formación, resulta interesante abordar el estudio sobre cómo las cianobacterias son capaces de formarlos. La caracterización del proceso permitiría obtener nuevas cepas de cianobacterias modificadas genéticamente para su aplicación en los procesos descritos, de modo que un estudio más profundo de las rutas bioquímicas de síntesis y de su regulación es de gran interés.

### 2.3. Las proteínas FUR y su relación con la formación de biofilms

Las proteínas FUR (*ferric uptake regulator*) son una superfamilia de proteínas ubicuas en organismos procariotas. Son reguladores de la expresión génica y actúan uniéndose a regiones del DNA ricas en A y T denominadas cajas FUR (24). Se caracterizan por la presencia de un dominio rico en histidinas con la secuencia HHHXHX<sub>2</sub>CX<sub>2</sub>C en la región C-terminal, siendo X un residuo cualquiera, habitualmente ácido aspártico (24). Tradicionalmente se ha descrito que las proteínas FUR regulan la expresión génica en forma dimérica, reprimiéndola y utilizando metales divalentes como correpresores. Además, muchas de ellas presentan un sitio de unión a otro metal, Zn<sup>2+</sup>, el cual presenta una función estructural (25). Sin embargo, más recientemente, se ha visto que las proteínas FUR tienen también la capacidad de actuar como activadores transcripcionales y de unirse al DNA en ausencia del metal correpresor (25). Los estudios de cristalografía de rayos X de la proteína Fur de *Campylobacter jejuni* muestran que el sitio de unión al DNA en ausencia del metal correpresor está rotado 180° respecto de orientación en presencia del metal, lo cual plantea la posibilidad de que la proteína se una a distintas regiones del DNA en ausencia y presencia de metales (24).

*Anabaena sp.* PCC7120 presenta tres parálogos FUR, FurA, FurB y FurC. FurA es el principal regulador de la homeostasis del hierro (24), aunque se ha observado que además regula otros procesos como el metabolismo del N<sub>2</sub>, la fotosíntesis y la morfología celular, actuando como un regulador global (26). Respecto a FurB, es una proteína Zur, un regulador del metabolismo del zinc, pero está también implicada en otros procesos como la regulación del estrés oxidativo (27). Finalmente, FurC es una proteína semejante a las proteínas PerR, relacionadas con la respuesta al estrés oxidativo, pero presenta actividades que no se habían asociado antes a las proteínas PerR, al participar en el remodelado de la maquinaria fotosintética tras la fotooxidación (28) y en el metabolismo del N<sub>2</sub> (29).

Dado el papel que juegan las proteínas FUR en la respuesta a estreses físico-químicos, y que los biofilms se sintetizan en muchas ocasiones como protección ante estos estreses, cabe pensar que las proteínas FUR pueden jugar un rol importante en la regulación de su formación. Los estudios con estirpes de *Anabaena sp.* PCC7120 que sobreexpresan las proteínas FUR muestran que la producción de biofilms está relacionada con estas proteínas. Aunque en la estirpe de sobreexpresión de FurB no hay cambios notables en la producción de biofilms, las estirpes de sobreexpresión de FurA y de FurC producen entre 3 y 4 veces más biofilm que la estirpe silvestre. Ante estreses salinos y en deficiencia de nitrógeno, la formación de biofilms aumenta también en la estirpe de sobreexpresión de FurC. Este aumento, ante los estreses nombrados, es algo más sutil cuando se sobreexpresa FurA, mientras que cuando la proteína sobreexpresada es FurB cesa la formación de biofilms por parte de la cianobacteria. A pesar de estas observaciones, los datos sobre el control de la formación de biofilms por parte de las proteínas FUR son limitados, lo cual abre un amplio campo de estudio todavía por explorar (30).



### 3. Objetivos

El propósito de este trabajo es el de analizar el papel de las proteínas FUR en la regulación de la formación de biofilms en *Anabaena sp.* PCC7120. Así, se establecen los siguientes objetivos:

- Identificar genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120.
- Localizar por métodos bioinformáticos secuencias de unión para las proteínas FUR en las regiones promotoras de los genes potencialmente implicados en la formación de biofilms.
- Analizar mediante ensayos de retardo en gel (EMSA) la capacidad de las proteínas FUR para unirse a las regiones promotoras de estos genes.

### 4. Materiales y métodos

#### 4.1. Técnicas de trabajo con DNA

##### 4.1.1. Reacción en cadena de la polimerasa (PCR)

La PCR se utilizó para amplificar, a partir del DNA genómico de *Anabaena sp.* PCC7120, los promotores de los genes que se analizaron mediante ensayos de retardo en gel (EMSA). Cada mezcla de reacción contenía 72  $\mu\text{L}$  de agua miliQ estéril, 10  $\mu\text{L}$  de *buffer* 10x (Tris-HCl 75 mM pH 9,0, KCl 50 mM,  $(\text{NH}_4)_2\text{SO}_4$  20 mM), 3  $\mu\text{L}$  de  $\text{MgCl}_2$  50 mM, 2  $\mu\text{L}$  de mezcla de dNTPs 10 mM, 2,5  $\mu\text{L}$  de cebador directo 20  $\mu\text{M}$ , 2,5  $\mu\text{L}$  de cebador reverso 20  $\mu\text{M}$ , 1  $\mu\text{L}$  de *Taq* DNA polimerasa 5 U/ $\mu\text{L}$  (*Biotools*) y 100 ng de DNA genómico. La amplificación se llevó a cabo en un termociclador *Thermal Cycler 2720* (*Applied Biosystems*). El programa del termociclador y los oligonucleótidos utilizados se indican en las secciones 8.1 y 8.2 de Material Suplementario, respectivamente.

##### 4.1.2. Electroforesis de DNA

Se llevó a cabo una electroforesis de DNA en gel de agarosa al 1 % (p/v) con un 1 % (v/v) de bromuro de etidio para comprobar la correcta amplificación de los promotores y la ausencia de productos inespecíficos de PCR. Para ello disolvieron 0,5 g de agarosa en 50 mL de tampón TBE (Tris-HCl 90 mM pH 8, ácido bórico 90 mM y EDTA 2 mM) calentando la mezcla en un microondas. Una vez disuelta, la mezcla se dejó enfriar y, antes de su solidificación, se añadieron 50  $\mu\text{L}$  de bromuro de etidio 0,5 g/mL. Para realizar la electroforesis se tomaron 4  $\mu\text{L}$  de cada muestra a analizar y se mezclaron con 3  $\mu\text{L}$  de tampón de carga (Tris-HCl 50 mM pH 8, 30 % de glicerol (v/v) y 0,25 % (p/v) de azul de bromofenol), tras lo que se cargaron individualmente en el gel de agarosa. La electroforesis se realizó en tampón TBE durante 25 minutos utilizando una fuente *BioRad PowerPac300*. Para visualizar los resultados se utilizó un equipo *GelDoc200* (*BioRad*) y el programa *MultiAnalist 1.1* (*BioRad Laboratories*, 1997).

##### 4.1.3. Purificación y cuantificación de los productos de PCR

Para eliminar los reactivos de la PCR, que podrían interferir más tarde en el trabajo, los promotores se purificaron utilizando el *kit GFXTM PCR DNA and Gel band Purification* (*GE Healthcare*), siguiendo las indicaciones del fabricante. Finalmente, se determinó la concentración de los productos de PCR utilizando un equipo *NanoVue PlusTM* (*GE Healthcare*).

## 4.2. Ensayos de retardo en gel (EMSA)

El estudio de la unión *in vitro* de FurA, FurB y FurC a los distintos promotores se realizó mediante ensayos de retardo en gel (*electrophoretic mobility shift assay*, EMSA). Estos ensayos se basan en el hecho de que la unión de las proteínas FUR a fragmentos de DNA diana produce una migración del DNA retardada con respecto del DNA libre. Para garantizar que la unión de las proteínas al DNA es específica, los ensayos se realizan en presencia de un DNA competidor, un fragmento interno del gen *pkn22* (*ifpkn22*) con el que no hay unión específica.

Los EMSA se realizaron en geles de poliacrilamida no desnaturalizantes al 6 % para permitir la migración de los complejos DNA-proteína. Los geles se prepararon mezclando 5,79 mL de agua miliQ estéril, 2 mL de acrilamida:bisacrilamida (30:0,8 p/v), 0,93 mL de *Running Buffer* 10x (30,28 g/L Tris-HCl, 142 g/L glicina, pH 8,5), 1,4 mL de glicerol al 50 %, 50  $\mu$ L de PSA al 10 % y 30  $\mu$ L de TEMED. Además, en la mezcla de los geles para ensayar FurA y FurC, se añadió  $MnCl_2$  a una concentración de 100  $\mu$ M. Una vez preparados, los geles se precorrieron durante al menos 1 hora a 60 V y 4 °C en el *buffer* utilizado para preparar los geles. En el caso de los geles para el estudio de FurA y FurC, el *buffer* además contenía  $MnCl_2$  100  $\mu$ M.

Cada mezcla de reacción se preparó mezclando 2  $\mu$ L de *buffer* de unión 10x (Bis-Tris 10 mM pH 7,5, KCl 40 mM, glicerol 5 % (v/v)), 1  $\mu$ L de BSA a 1 mg/mL, 1  $\mu$ L de DTT 20 mM y 1  $\mu$ L de  $MnCl_2$  2mM. Sobre estas mezclas se añadió cada promotor y proteína a estudiar, además del DNA inespecífico, y se completó con agua miliQ estéril hasta un volumen final de 20  $\mu$ L. Para FurA, se añadieron 100 ng de *ifpkn22* y 100 ng de promotor, y la unión DNA-proteína se estudió en presencia de FurA a 100 nM, 150 nM y 200 nM. En el caso de FurB, se añadieron 50 ng de *ifpkn22* y 50 ng de promotor, y la unión DNA-proteína estudió presencia de FurB a 100 nM, 200 nM y 300 nM. En el caso de FurC, se añadieron 50 ng de *ifpkn22* y 50 ng de promotor, y la unión DNA-proteína estudió en presencia FurC a 100 nM, 175 nM y 250 nM. En todos los casos, además, se realizó el estudio en ausencia de proteína, y se introdujo un control negativo y otro positivo en cada gel. Las mezclas se incubaron a temperatura ambiente durante 30 minutos y, tras esto, se mezclaron con 3  $\mu$ L de *buffer* de carga (Tris-HCl 50 mM pH 8, glicerol 30 % (v/v) y azul de bromofenol 0,05 % (p/v)). A continuación, y tras limpiar los pocillos, se cargaron los 20  $\mu$ L de cada *mix* y se llevó a cabo la electroforesis a 90 V y 4 °C durante 90 minutos. Finalizadas las electroforesis, los geles se tiñeron con 1  $\mu$ g/mL de *SyBr<sup>®</sup> Safe Fluorescent Dye* durante 10 minutos, en agitación y oscuridad, para luego ser revelados en un transiluminador *BioRad Gel Doc 2000*.

## 4.3. Herramientas bioinformáticas

La búsqueda de genes de *Anabaena sp.* PCC7120 potencialmente implicados en la formación de biofilms a partir de genes de otras especies se llevó a cabo realizando un BLAST de proteínas. Para ello, se obtuvo de KEGG ([www.genome.jp/kegg/](http://www.genome.jp/kegg/)) la secuencia de aminoácidos de las proteínas identificadas en estas otras especies, y se identificó el homólogo en *Anabaena sp.* PCC7120 utilizando la herramienta BLAST del NCBI ([www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)). En los genes identificados, por homología o directamente de la literatura, se llevó a cabo una búsqueda de cajas FUR. Para ello, desde KEGG, se obtuvo la secuencia promotora de los genes, y realizó una búsqueda de cajas con la herramienta FIMO de MEMEsuite ([meme-suite.org/meme/](http://meme-suite.org/meme/)) utilizando como *cut off* un p-valor de  $10^{-4}$ . Las matrices de peso utilizadas en la búsqueda de cajas FUR se adjuntan en la sección 8.3 de Material Suplementario.

## 5. Resultados y discusión

### 5.1. Identificación de genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC 7120

Con la intención de identificar genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120 se llevó a cabo un revisión bibliográfica que permitió encontrar genes relacionados con la formación de biofilms, tanto en *Anabaena sp.* PCC7120 como en otros microorganismos, en cuyo caso se buscó el gen homólogo en *Anabaena sp.* PCC7120 mediante un BLAST de proteínas. Los resultados de la búsqueda se resumen en la Tabla 1.

Grupo de genes	Genes	Número de genes
Producción de EPS	Glicosiltransferasas	27
	Biosíntesis de polisacáridos	4
	Proteínas modificadoras de monosacáridos	7
	Biosíntesis de lipopolisacáridos	1
	Biosíntesis de poliaminoácidos	1
Sistemas de síntesis de EPS	Sistema Wzy dependiente	18
	Sistema dependiente de transportador de tipo ABC	16
	Sistema dependiente de sintasa	3
Factores de transcripción		1
Sistema de la cianoexosortasa		2

**Tabla 1.** Funciones desempeñadas por los genes de *Anabaena sp.* PCC7120 potencialmente implicados en la formación de biofilms.

En la Tabla 1 se indica, de forma resumida, el resultado de la búsqueda de genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120. Entre los genes aparecen, por una parte, genes codificantes de enzimas implicadas en la producción de EPS, principalmente genes de glicosiltransferasas, proteínas de biosíntesis de polisacáridos y proteínas transformadoras de monosacáridos, así como genes codificantes de proteínas implicadas en la síntesis de lipopolisacáridos y poliaminoácidos como la cianoficina. Por otra parte, se han identificado también genes de los sistemas de síntesis de biofilms previamente descritos; genes codificantes de proteínas del sistema Wzy dependiente (Wza, Wzb, Wzc, Wzx y Wzy), genes codificantes de proteínas del sistema dependiente de transportador de tipo ABC (KpsC, KpsD, KpsE, KpsF, KpsM, KpsS, KpsT y KpsU) y genes codificantes de proteínas del sistema dependiente de sintasa (Alg8 y Alg44). Finalmente, se han encontrado genes codificantes de factores de transcripción y de proteínas del sistema de la cianoexosortasa, proteasas que reconocen señales del extremo C-terminal de las proteínas y que las unen a la membrana plasmática. Son, por tanto, proteínas que podrían estar implicadas en la fijación de los componentes proteicos de los biofilms a la superficie de las cianobacterias.

Todos los genes identificados en la revisión bibliográfica, sus organismos de origen y los genes homólogos en *Anabaena sp.* PCC7120 se recogen en la Tabla 2, junto con la anotación de todos ellos obtenida de KEGG y MicrobesOnline.

Paper	Organismo	Proteína	Gen		Homólogo <i>Anabaena</i> sp. PCC7120		Identidad (%)
			Gen	Proteína	Gen	Anotación	
(31)	<i>Synechocystis</i> PCC6803	Exopolysaccharide export protein; KpsE	<i>slf0923 (epsB)</i>	<i>slf0923 (epsB)</i>	<i>slf5222</i>	Polysaccharide biosynthesis tyrosine autokinase	25,66
		Polysaccharide biosynthesis/export protein; GumB/KpsD	<i>slf1581 (gumB)</i>	<i>slf1581 (gumB)</i>	<i>slf4388</i>	SLBB domain-containing protein	39,18
		Exopolysaccharide synthesis protein; ExoD	<i>slf1875 (exoD)</i>	<i>slf1875 (exoD)</i>	<i>slf1787</i>	Exopolysaccharide biosynthesis protein	56,59
		Lipopolysaccharide transport system permease protein	<i>slr0977</i>	<i>slr0977</i>	<i>slf0917</i>	ABC transporter permease	30,11
		Lipopolysaccharide transport system ATP-binding protein	<i>slr0982 (rbbB)</i>	<i>slr0982 (rbbB)</i>	<i>slf0916</i>	ABC transporter ATP-binding protein	43,41
		ABC-2 type transport system ATP-binding protein	<i>slf0575 (rbbB)</i>	<i>slf0575 (rbbB)</i>	<i>slf4486</i>	ABC transporter ATP-binding protein	65,69
		Lipopolysaccharide transport system permease protein	<i>slf0574 (rbbA)</i>	<i>slf0574 (rbbA)</i>	<i>slf4485</i>	ABC transporter permease	67,65
		Factor sigma J; SigJ		<i>slr0277 (sigJ)</i>	*	*	*
		Similar to polysaccharide biosynthesis/export protein		<i>slf2290</i>	*	*	*
		Dolichyl-1-phosphate-mannose synthase		<i>slf2770</i>	*	*	*
		Mannose-6-phosphate isomerase		<i>slf4342</i>	*	*	*
		Glucosyltransferase		<i>slf4420</i>	*	*	*
		UDP-N-acetyl-D-mannosamine transferase		<i>slf4422</i>	*	*	*
Similar to glucosyltransferase		<i>slf4719</i>	*	*	*		
Putative glycosyl transferase		<i>slr0074</i>	*	*	*		
Probable glycosyl transferase		<i>slf3057</i>	*	*	*		
Probable glycosyl transferase		<i>slf3058</i>	*	*	*		
Similar to polysaccharide export protein		<i>slf3059</i>	*	*	*		
Similar to acetyltransferase		<i>slf3061</i>	*	*	*		
Probable glycosyl transferase		<i>slf3062</i>	*	*	*		
Probable glycosyl transferase		<i>slf3063</i>	*	*	*		
Probable glycosyl transferase		<i>slf3064</i>	*	*	*		
Probable polysaccharide biosynthesis protein		<i>slf3065</i>	*	*	*		
Polysaccharide polymerization protein		<i>slf3066</i>	*	*	*		
Probable glycosyl transferase		<i>slf3067</i>	*	*	*		
Probable glycosyl transferase		<i>slf3068</i>	*	*	*		
Probable glycosyl transferase		<i>slf3069</i>	*	*	*		
Probable glycosyl transferase		<i>slf3070</i>	*	*	*		
Probable glycosyl transferase		<i>slf3071</i>	*	*	*		
Probable polysaccharide biosynthesis protein		<i>slf3072</i>	*	*	*		
Probable glycosyl transferase		<i>slf3073</i>	*	*	*		
Similar to glucosyl-1-phosphate transferase		<i>slf4823</i>	*	*	*		
Glycosyltransferase		<i>slf5223</i>	*	*	*		
(18)	<i>Synechocystis</i> PCC6803	Alg8	<i>slf1566</i>	<i>slf0776</i>	Glycosyltransferase	54,73	
		Dolicho-1-phosphate mannose transferase; Alg8	<i>slf1004</i>	<i>slf0143</i>	Glycosyltransferase	59,52	
		Probable glycosyltransferase; Alg8	<i>slf5056</i>	<i>slf0042</i>	Glycosyltransferase	31,07	
		1,2-diacetylglucosyl-3-beta-glucosyltransferase; Alg8	<i>slf1377</i>	<i>slf4933</i>	Glycosyltransferase	54,3	
		Hyd family of secretion protein; Alg44	<i>slf1181</i>	<i>slf4240</i>	Hyd family efflux transporter periplasmic adaptor subunit	41,75	
		Exopolysaccharide synthesis protein; ExoD	<i>slf1875 (exoD)</i>	<i>slf1787</i>	Exopolysaccharide biosynthesis protein	56,59	
		Exopolysaccharide export protein; KpsE	<i>slf0923 (epsB)</i>	<i>slf5222</i>	Polysaccharide biosynthesis tyrosine autokinase	25,66	
		Polysaccharide biosynthesis/export protein; GumB/KpsD	<i>slf1581 (gumB)</i>	<i>slf4388</i>	SLBB domain-containing protein	39,18	
		KpsC	<i>slf2115</i>	-	-	-	
		KpsF	<i>slf2111</i>	-	-	-	
(33)	<i>Synechocystis</i> PCC6803	KpsS	<i>slf2115</i>	-	-	-	
		KpsU	<i>slf2122</i>	-	-	-	
		Polysialic acid transport protein; KpsM	<i>slf2107 (kpsM)</i>	<i>slf0917</i>	ABC transporter permease	66,31	
		Polysialic acid transport ATP-binding protein; KpsT	<i>slf2108</i>	<i>slf0916</i>	ABC transporter ATP-binding protein	55,32	

Tabla 2. Resultados de la revisión bibliográfica y la búsqueda de genes homólogos en *Anabaena* sp. PCC7120.



(34)	<i>Anabaena</i> sp. PCC7120	Genes Wzb putativos	<i>alr5068</i>	*	*	*	*
			<i>alr3436</i>	*	*	*	*
			<i>alr1067</i>	*	*	*	*
			<i>all5222</i>	*	*	*	*
			<i>all4432</i>	*	*	*	*
		Genes Wzc putativos	<i>alr3059</i>	*	*	*	*
			<i>alr2856</i>	*	*	*	*
			<i>alr2833</i>	*	*	*	*
			<i>all0059</i>	*	*	*	*
			<i>all0493</i>	*	*	*	*
			<i>all4430</i>	*	*	*	*
		Genes Wzx putativos	<i>alr3072</i>	*	*	*	*
			<i>alr3065</i>	*	*	*	*
			<i>alr2857</i>	*	*	*	*
			<i>all2290</i>	*	*	*	*
			<i>all5073</i>	*	*	*	*
			<i>all4428</i>	*	*	*	*
		Genes Wzy putativos	<i>alr3690</i>	*	*	*	*
			<i>alr3060</i>	*	*	*	*
			<i>alr2861</i>	*	*	*	*
			<i>alr4485</i>	*	*	*	*
			<i>all4219</i>	*	*	*	*
			<i>alr2373</i>	*	*	*	*
		Genes KpsM putativos	<i>alr1500</i>	*	*	*	*
			<i>alr1491</i>	*	*	*	*
			<i>alr0076</i>	*	*	*	*
			<i>alr1196</i>	*	*	*	*
			<i>all0917</i>	*	*	*	*
			<i>alr3464</i>	*	*	*	*
		Genes KpsF putativos	<i>alr2432</i>	*	*	*	*
			<i>all5222</i>	*	*	*	*
			<i>all4432</i>	*	*	*	*
		Genes KpsE putativos	<i>alr3059</i>	*	*	*	*
			<i>alr2856</i>	*	*	*	*
	<i>alr2833</i>	*	*	*	*		
	<i>all0059</i>	*	*	*	*		
	<i>all4388</i>	*	*	*	*		
Genes KpsD putativos	<i>alr2294</i>	*	*	*	*		
	<i>all0495</i>	*	*	*	*		
	<i>all5222</i>	*	*	*	*		
	<i>all4432</i>	*	*	*	*		
Genes PCP (polysaccharide copolymerase protein) putativos	<i>alr3059</i>	*	*	*	*		
	<i>alr2856</i>	*	*	*	*		
	<i>alr2833</i>	*	*	*	*		
	<i>all0059</i>	*	*	*	*		
	<i>all0493</i>	*	*	*	*		

Tabla 2 (cont.). Resultados de la revisión bibliográfica y la búsqueda de genes homólogos.

## 5.2. Análisis bioinformático de la presencia de cajas FUR en regiones promotoras de genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120

Una vez identificados los genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120, se analizó si formaban parte de operones utilizando la información disponible en MicrobesOnline (<http://www.microbesonline.org/>). Así, se encontraron nuevos genes potencialmente implicados en la formación de biofilms que formaban operones con los previamente identificados y que por tanto se regularían conjuntamente. Con ello, se identificaron un total de 108 genes que podrían participar en la formación de biofilms en *Anabaena sp.* PCC7120.

Además, se llevó a cabo un cribado para eliminar genes cuyo estudio no resultase interesante. Parte de los genes identificados habían sido detectados por homología a partir de genes de bacterias alejadas evolutivamente de las cianobacterias, de modo que algunos de ellos tenían funciones no relacionadas con la formación de biofilms en *Anabaena sp.* PCC7120. Así, se eliminó del estudio el gen *alr2887*, perteneciente a un operón (*alr2884-alr2940*) codificante de proteínas relacionadas con el control de los ritmos circadianos, y el gen *alr3699*, perteneciente también a un operón relacionado en este caso con la formación del heterocisto (*alr3698-alr3701*). También se eliminaron los genes *alr1105* y *alr5241*, que codifican una arsenato reductasa y una metiltransferasa, respetivamente, y que no resultan por tanto de interés en el estudio realizado. No obstante, sí que se conservó en el estudio el gen *alr0267*, pues aunque se trata de un gen implicado en la formación del heterocisto, se sospecha que está también involucrado en la formación de biofilms (35).

Tras la identificación de unidades transcripcionales implicadas en la formación de biofilms se llevó a cabo la búsqueda de cajas FUR en las regiones promotoras de las mismas. Para ello, se tomó la región intergénica previa a cada unidad transcripcional. No obstante, estudios previos han mostrado que las cajas FUR pueden aparecer dentro de las secuencias codificantes del gen de interés o del gen previo a dicha región intergénica (36). Por ello, junto con la región intergénica, se tomó un fragmento de secuencia codificante de 65 nucleótidos, tanto del gen de interés como del gen previo a la región intergénica de estudio.

En el caso de las cajas detectadas en las regiones intergénicas de genes divergentes (con sentidos de lectura opuestos), hay que tener en cuenta que la región promotora de ambos genes se localiza en la misma región intergénica, de modo que una caja localizada más cerca del gen anterior que del gen de interés podría regular al primero de los genes, en lugar de al segundo. Por ello, en estos casos, las regiones intergénicas se dividieron en tres segmentos de igual longitud y se consideró que el segmento más cercano a cada uno de los genes regulaba a dicho gen, mientras que el segmento central se consideró que podría regular ambos. Por tanto, en el caso de los genes divergentes, se aceptaron como posibles cajas FUR aquellas cajas localizadas en los dos tercios de región intergénica más próximos al gen de estudio, y se rechazaron las que aparecían en el tercio de región intergénica más alejado de la unidad transcripcional de estudio.

El resultado de la búsqueda de cajas FUR, así como el p-valor de cada predicción, se adjuntan en la Tabla 3. Las secuencias intergénicas en las que se detectaron cajas FUR aparecen indicadas en la sección 8.4 de Material Suplementario.

Operón	Gen	Anotación	Caja FurA		Caja FurB		Caja FurC	
			Secuencia caja	p-valor	Secuencia caja	p-valor	Secuencia caja	p-valor
<i>alI0042</i>	<i>alI0042</i>	Glycosyltransferase	ACTGGTTCAGGATTTACTT	3.57e-05				
<i>alI0059</i>	<i>alI0059</i>	Exopolysaccharide transport family protein	ATACAAATATTCAGATAT	4.81e-05	TCAGATAATATTCATTTATTTTAAAC	1.93e-05		
<i>alI0267</i>	<i>alI0267</i>	Heterocyst-specific attachment factor; HesF	TATTAATAAATTCCTGATAT	5.57e-05				
<i>alI0277</i>	<i>alI0277</i>	Sigma J factor; SigJ			GTTATCAGGAATAAATATTATACAG	4.81e-05		
<i>alI0492</i>	<i>alI0492</i>	Glycerol-3-phosphate acyltransferase						
<i>alI0492</i>	<i>alI0492</i>	Polysaccharide biosynthesis protein	GCTTTTGAAAAATATAAAT	9.33e-05				
<i>alI0496</i>	<i>alI0496</i>	Polysaccharide biosynthesis/export family protein						
<i>alI0495</i>	<i>alI0495</i>	Cyanoexosortase B system-associated protein	AGTAAAAATTTGCGTAAAT	1.64e-05				
<i>alI0497</i>	<i>alI0497</i>	Cyanoexosortase B	ATATAATCTGAATTTTTTGA	5.03e-05				
<i>alI0657</i>	<i>alI0657</i>	RibB; UDP-glucuronate decarboxylase						
<i>alI0658</i>	<i>alI0658</i>	UDP-glucose 6-dehydrogenase			ATGGTTTAGAACCAATATCAAAGC	2.32e-05		
<i>alI0776</i>	<i>alI0776</i>	Glycosyltransferase	ATCAATTCAAAAATTCAAAA	5.49e-05				
<i>alI0917</i>	<i>alI0917</i>	Lipopolysaccharide transport system ATP-binding protein						
<i>alI0916</i>	<i>alI0916</i>	Lipopolysaccharide transport system permease protein			AAGGTTGTTAACCGTTATATTACTA	2.03e-05		
<i>alI0919</i>	<i>alI0919</i>	Colanic acid biosynthesis glycosyltransferase						
<i>alI1490</i>	<i>alI1490</i>	ABC transporter ATP-binding component						
<i>alI1491</i>	<i>alI1491</i>	ABC transporter permease			CAACGTCGAATGATTTTAATTAGA	4.6e-06	CAACGTCGAATGATTTTA	8.41e-05
	<i>alI2283</i>	Anti-sigma factor antagonist						
	<i>alI2284</i>	Sigma-B activity negative regulator						
	<i>alI2285</i>	Sugar transferase						
	<i>alI2286</i>	Glucosyltransferase						
	<i>alI2287</i>	Glucosyltransferase						
	<i>alI2288</i>	Glucosyltransferase						
	<i>alI2289</i>	Glucosyltransferase						
	<i>alI2290</i>	Similar to polysaccharide biosynthesis export protein						
	<i>alI2291</i>	Glycosyltransferase						
	<i>alI2292</i>	Probable glycosyltransferase						
	<i>alI2293</i>	Polysaccharide pyruvyl transferase	ACTTATGGTAGTTATATCT	9.32e-06	CAGAATAATTTTATTTATCAACTGA	3.03e-05		
<i>alI2294</i>	<i>alI2294</i>	Polysaccharide export outer membrane protein	AGATAAATACTACCATAAGT	9.32e-06	TCAGTTGATAAATAAATAATTTCTG	3.03e-05		
<i>alI2296</i>	<i>alI2296</i>	GumC family protein						
<i>alI2296</i>	<i>alI2296</i>	Capsular polysaccharide biosynthesis protein						
<i>alI2372</i>	<i>alI2372</i>	ABC transporter ATP-binding protein						
<i>alI2373</i>	<i>alI2373</i>	ABC-2 type transport system permease protein	ATTTAAGGAAAATTTTTATG	2.52e-06				
<i>alI2374</i>	<i>alI2374</i>	Unknown protein						

**Tabla 3.** Resultado de la búsqueda de cajas FUR en los genes identificados en la revisión bibliográfica y la búsqueda bioinformática. Los genes indicados en negrita son los genes identificados en la búsqueda bibliográfica, mientras que el resto son los detectados en el estudio de operones. En la fila correspondiente a cada unidad transcripcional se indican las secuencias de las cajas FUR predichas, así como el p-valor de la predicción.

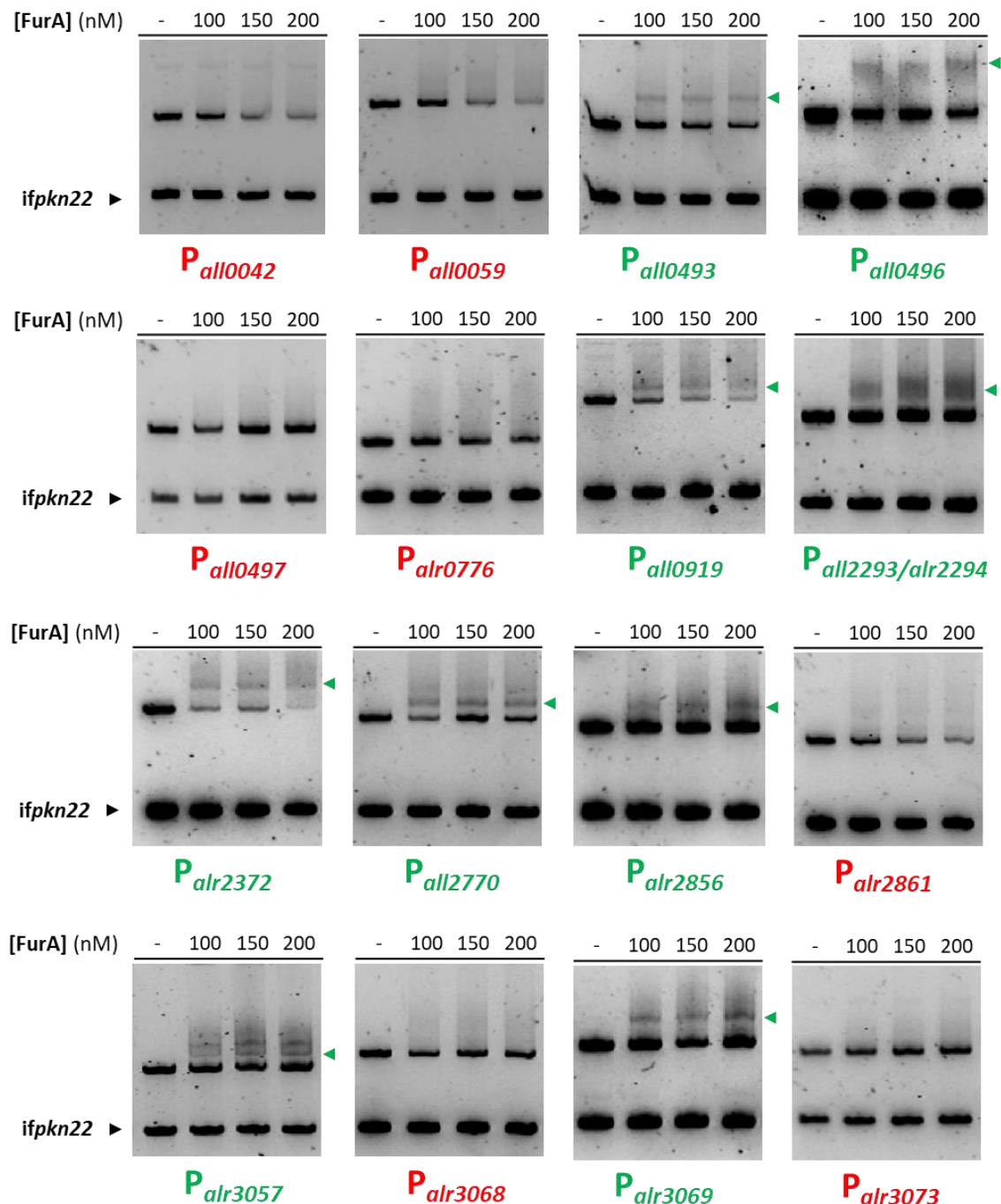


<i>alI2770</i>	<i>alI2770</i>	Dolichol-phosphate-mannose synthase	AGAAAAATATTGAAGATTT	4.79e-05	TCTGACAAACAATTAAATTTCAATAATT	4.1e-05	
<i>alI2832</i>	<i>alI2832</i>	Putative glycosyltransferase					
<i>alI2833</i>	<i>alI2833</i>	Polysaccharide biosynthesis tyrosine autokinase					
<i>alI2856</i>	<i>alI2856</i>	Polysaccharide biosynthesis tyrosine autokinase	AGATAATGAGAAATTTTAA	1.04e-05			
<i>alI2861</i>	<i>alI2861</i>	Unknown protein	AAGAAAAATATTCATGATTT	9.73e-05			
<i>alI3057</i>	<i>alI3057</i>	Probable glycosyl transferase	ACATCATGAGAAATAAAACG	8.64e-05			
<i>alI3060</i>	<i>alI3060</i>	Probable glycosyl transferase					
<i>alI3058</i>	<i>alI3058</i>	Probable glycosyl transferase					
<i>alI3059</i>	<i>alI3059</i>	Similar to polysaccharide export protein					
<i>alI3060</i>	<i>alI3060</i>	O-antigen ligase					
<i>alI3061</i>	<i>alI3061</i>	Similar to acetyltransferase	ATTTAATGAGTAATTTATA	1.23e-07	ATTCAAAAATAAACATTTCTTAAATTC	6.33e-05	
<i>alI3068</i>	<i>alI3068</i>	Probable glycosyl transferase	TTTTAGTATTTGGATAAAAAT	9.89e-05	AGAAAATTAATAATGATATCATCATC	3.21e-05	
<i>alI3069</i>	<i>alI3069</i>	Probable glycosyl transferase					
<i>alI3070</i>	<i>alI3070</i>	Probable glycosyl transferase					
<i>alI3071</i>	<i>alI3071</i>	Probable glycosyl transferase					
<i>alI3073</i>	<i>alI3073</i>	Probable glycosyl transferase	AAATAAATATGCCTAAAAT	8.7e-05			
<i>alI3074</i>	<i>alI3074</i>	UDP-3-O-[β-D-hydroxymyristoyl] glucosamine N-acetyltransferase					
<i>alI3464</i>	<i>alI3464</i>	Glutamine-fructose-6-phosphate-aminotransferase	ACATCATGAGAAATTTATA	5.71e-05	CAAGATCATGAGAAATTTA	3.76e-05	
<i>alI3509</i>	<i>alI3509</i>	GDP-mannose 4,6-dehydratase	ATTTCTCAAGAATCAAAT	7.94e-05	ATTTAGAATAATTTGTCTTAAAAAATA	4.71e-05	
<i>alI3756</i>	<i>alI3756</i>	Unknown protein					
<i>alI3757</i>	<i>alI3757</i>	Glycosyltransferase					
<i>alI4219</i>	<i>alI4219</i>	ABC transporter ATP-binding component	AAGTGACATCTCAAAAAGC	4.92e-05			
<i>alI4239</i>	<i>alI4239</i>	Toxin secretion ABC transporter ATP-binding protein	AGTAAAAAGCTTCTTAAAGC	7.71e-06			
<i>alI4240</i>	<i>alI4240</i>	HlyD family efflux transporter periplasmic adaptor subunit					
<i>alI4388</i>	<i>alI4388</i>	Polysaccharide biosynthesis/export protein					
<i>alI4422</i>	<i>alI4422</i>	UDP-N-acetyl-D-mannosamine transferase					TATAGTCAAGATTTAAAATG
<i>alI4423</i>	<i>alI4423</i>	Glycosyltransferase					
<i>alI4424</i>	<i>alI4424</i>	Probable glycosyltransferase					
<i>alI4425</i>	<i>alI4425</i>	Unknown protein					
<i>alI4426</i>	<i>alI4426</i>	Probable glycosyltransferase	TGAAATATTTACAAAAAC	6.84e-05			
<i>alI4428</i>	<i>alI4428</i>	O-antigen ligase family protein	TATAAAAAGTTTGTATATAT	3.95e-05			
<i>alI4429</i>	<i>alI4429</i>	Sugar transferase					
<i>alI4430</i>	<i>alI4430</i>	Flippase					
<i>alI4431</i>	<i>alI4431</i>	Glycosyltransferase					
<i>alI4432</i>	<i>alI4432</i>	Polysaccharide biosynthesis tyrosine autokinase					
<i>alI4485</i>	<i>alI4485</i>	ABC transporter permease	TCATAAATCTTCTGATAT	1.84e-05			
<i>alI4486</i>	<i>alI4486</i>	ABC transporter ATP-binding protein					
<i>alI4487</i>	<i>alI4487</i>	Glycosyl transferase involved in cell wall biogenesis					
<i>alI4488</i>	<i>alI4488</i>	O-antigen biosynthesis protein					
<i>alI5073</i>	<i>alI5073</i>	Putative inorganic carbon transporter	ATTTTCATGAAAATTTCTAT	1.73e-05			
<i>alI5222</i>	<i>alI5222</i>	Polysaccharide biosynthesis tyrosine autokinase	ATFAAAGGTATATTTTAT	2.93e-05	TATACTTAGAATCATTAATAAACGTTTA	2.59e-05	AATAATCATGAATAAAAATG
<i>alI5223</i>	<i>alI5223</i>	Glycosyltransferase	TCATAAAAACAGATTAATAAT	7.57e-05			CATTTTATTCATGATTTATT
<i>alI7196</i>	<i>alI7196</i>	Two-component response regulator	ACTTATGGAGATTTAATAAT	2.62e-05			

Tabla 3 (cont.). Resultado de la búsqueda de cajas FUR en los genes identificados en la revisión bibliográfica y la búsqueda bioinformática.

### 5.3. Estudio de la modulación de genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120 por las proteínas FUR

Tras localizar las posibles cajas FUR en las regiones promotoras de los genes de estudio, se analizó la unión del correspondiente parálogo FUR a ellas mediante ensayos de retardo en gel (EMSA). Los resultados de estos ensayos se recogen en las Figuras 2, 3 y 4.



**Figura 2.** Resultados de los ensayos de retardo en gel con FurA. Los promotores en los que se produjo unión aparecen marcados en verde y el complejo aparece señalado con una flecha en la parte derecha, mientras que los promotores con los que no se ha detectó unión aparecen marcados en rojo. En la parte superior de cada gel se indican las concentraciones de proteína y en el lateral izquierdo se indica la posición de la banda correspondiente con el DNA inespecífico. Como control positivo se empleó el promotor del gen *all1691* (*furA*) (26) y como control negativo el promotor del gen *alr1911* (*nifJ*) (26).

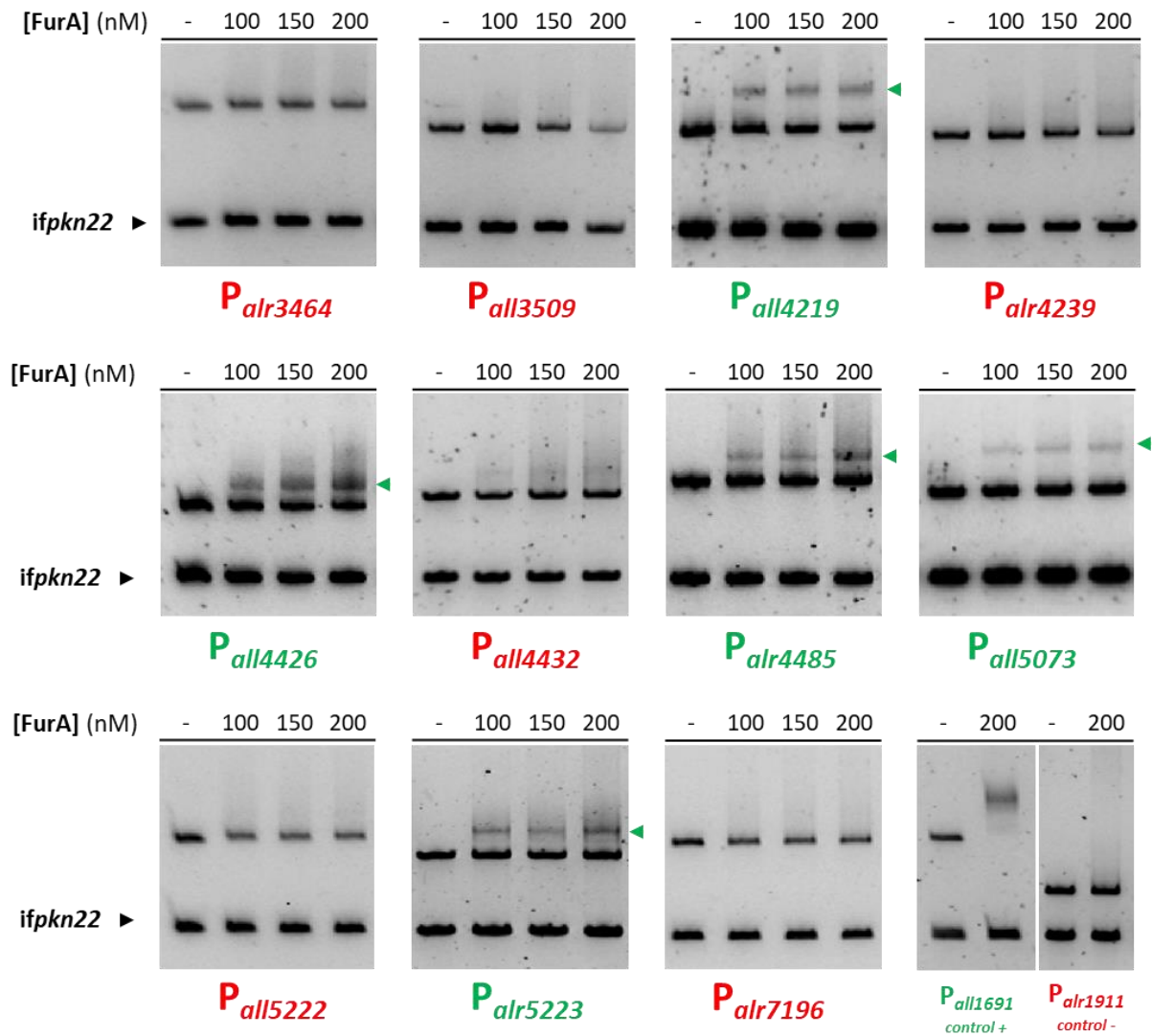


Figura 2 (cont.). Resultados de los ensayos de retardo en gel con FurA.

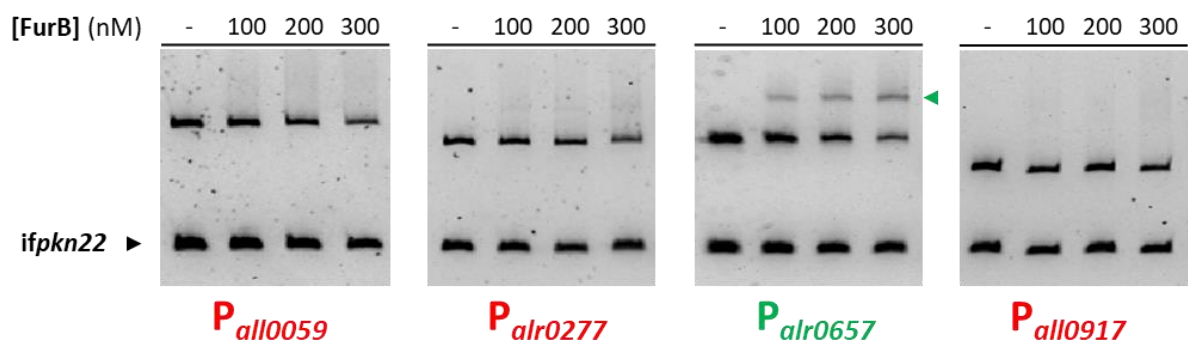


Figura 3. Resultados de los ensayos de retardo en gel con FurB. Los promotores en los que se produjo unión aparecen marcados en verde y el complejo aparece señalado con una flecha en la parte derecha, mientras que los promotores con los que no se ha detectó unión aparecen marcados en rojo. En la parte superior de cada gel se indican las concentraciones de proteína y en el lateral izquierdo se indica la posición de la banda correspondiente con el DNA inespecífico. Como control positivo se empleó el promotor del gen *all4725* (27) y como control negativo el promotor del gen *alr1911* (*nifJ*) (26).

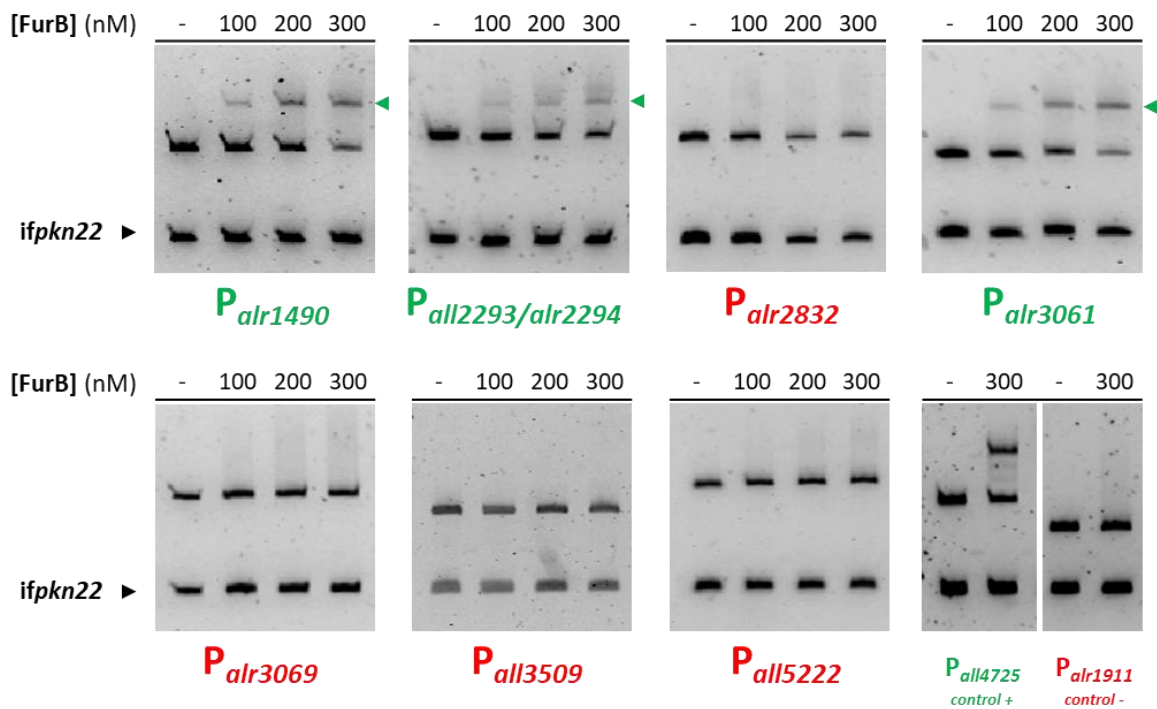


Figura 3 (cont.). Resultados de los ensayos de retardo en gel con FurB.

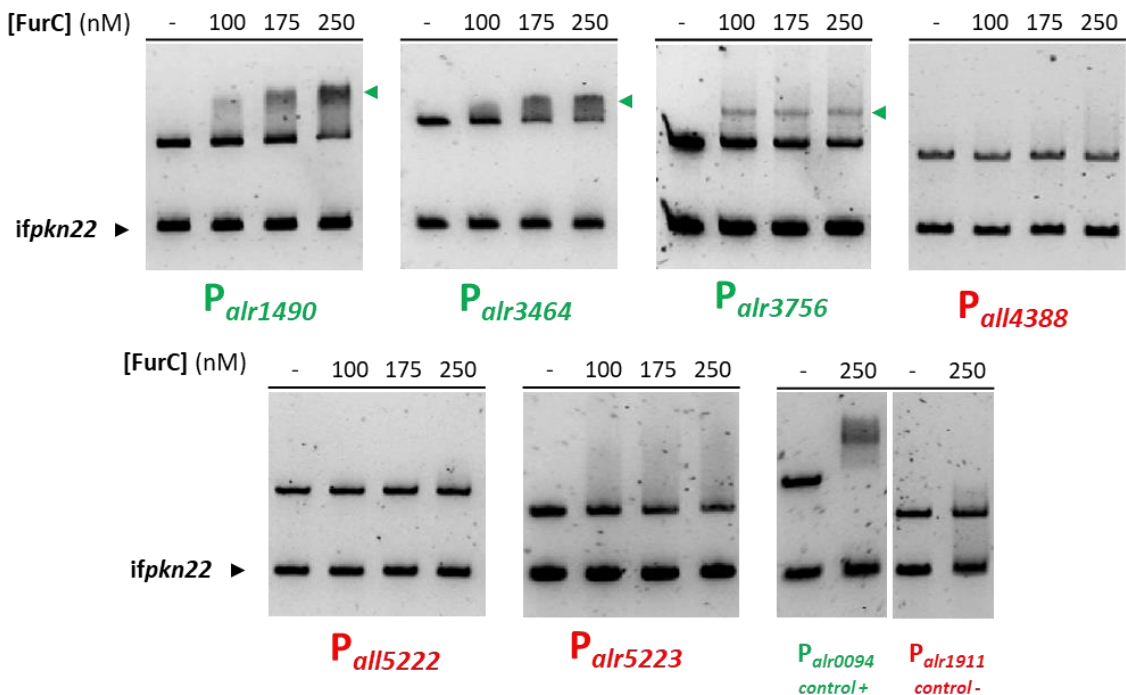


Figura 4. Resultados de los ensayos de retardo en gel con FurC. Los promotores en los que se produjo unión aparecen marcados en verde y el complejo aparece señalado con una flecha en la parte derecha, mientras que los promotores con los que no se ha detectó unión aparecen marcados en rojo. En la parte superior de cada gel se indican las concentraciones de proteína y en el lateral izquierdo se indica la posición de la banda correspondiente con el DNA inespecífico. Como control positivo se empleó el promotor del gen *alr0094* (*hetZ*) (29) y como control negativo el promotor del gen *alr1911* (*nifJ*) (26).

Como se observa en las Figuras 2, 3 y 4, FurA se unió a 14 de los 27 promotores ensayados, FurB a 4 de los 11 promotores ensayados y FurC a 3 de los 6 promotores ensayados.

Dado que varios de estos promotores regulan la expresión de operones, estos resultados han permitido identificar un total de 54 genes implicados en la síntesis de biofilms en *Anabaena sp.* PCC7120 directamente regulados por las proteínas FUR. Para profundizar en la función que desempeñan se integró la información depositada sobre los mismos en las bases de datos y la procedente de los estudios de homología, considerando que la función de la proteína de la cianobacteria y de la de su homóloga son la misma cuando su identidad es superior al 30 % (37). Los genes implicados en la síntesis de biofilms directamente regulados por las proteínas FUR junto con la función desempeñada por los mismos se recoge en la Tabla 4.

Unidad transcripcional	Gen	Proteína codificada	Resultado EMSA		
			FurA	FurB	FurC
all0493-all0492	all0492	Glicerol-3-fosfato aciltransferasa	+		
	all0493	Wzc putativa/polisacárido copolimerasa putativa	+		
all0496-all0495	all0495	Wza putativa/KpsD putativa	+		
	all0496	Proteína asociada al sistema de la cianoxosortasa	+		
alr0657-alr0658	alr0657	RfbB; UDP-glucuronato descarboxilasa		+	
	alr0658	GDP-manosa-6-deshidrogenasa		+	
all0919	all0919	Glicosiltransferasa de síntesis de ácido colánico	+		
alr1490-alr1491	alr1490	KpsT putativa		+	+
	alr1491	KpsM putativa		+	+
all2293-all2283	all2283	Agonista del factor anti sigma	+	+	
	all2284	Regulador negativo del factor sigma B	+	+	
	all2285	Glicosiltransferasa	+	+	
	all2286	Glicosiltransferasa	+	+	
	all2287	Glicosiltransferasa	+	+	
	all2288	Glicosiltransferasa	+	+	
	all2289	Glicosiltransferasa	+	+	
	all2290	Wzx putativa	+	+	
	all2291	Glicosiltransferasa	+	+	
	all2292	Probable glicosiltransferasa	+	+	
alr2294-alr2296	alr2293	Piruvil transferasa	+	+	
	alr2294	Wza putativa/KpsD putativa	+	+	
	alr2295	GumC; proteína de síntesis de lipopolisacáridos	+	+	
alr2372-alr2374	alr2296	Proteína de síntesis de polisacáridos capsulares	+	+	
	alr2372	KpsT putativa	+		
	alr2373	KpsM putativa	+		
all2770	alr2374	Proteína desconocida	+		
	all2770	Dolicol-fosfato-manosa sintasa	+		
	alr2856	Wzc putativa/KpsE putativa/polisacárido copolimerasa putativa	+		
alr3057-alr3060	alr3057	Probable glicosiltransferasa	+		
	alr3058	Probable glicosiltransferasa	+		
	alr3059	Wzc putativa/KpsE putativa/polisacárido copolimerasa putativa	+		
	alr3060	Wzy putativa	+		
alr3069-alr3071	alr3061	Acetiltransferasa		+	
	alr3069	Probable glicosiltransferasa	+		
	alr3070	Probable glicosiltransferasa	+		
alr3464	alr3071	Probable glicosiltransferasa	+		
	alr3464	KpsF putativa			+
	alr3756	Proteína desconocida			+
all4219	alr3757	Glicosiltransferasa			+
	all4219	KpsM putativa	+		
alr4239-alr4240	alr4239	KpsT putativa	+		
	alr4240	Alg44	+		
all4388	all4388	Wza putativa/KpsD putativa			+
all4426-all4422	all4422	UDP-N-acetil-D-manosamina transferasa	+		
	all4423	Glicosiltransferasa	+		
	all4424	Probable glicosiltransferasa	+		
	all4425	Proteína desconocida	+		
	all4426	Probable glicosiltransferasa	+		
alr4485-alr4488	alr4485	KpsM putativa	+		
	alr4486	KpsT putativa	+		
	alr4487	Glicosiltransferasa implicada en la síntesis de la pared celular	+		
	alr4488	Proteína de biosíntesis del antígeno-O	+		
all5073	all5073	Wzy putativa	+		
alr5223	alr5223	Glicosiltransferasa	+		

**Tabla 4.** Recopilación de las funciones desempeñadas por los genes potencialmente implicados en formación de biofilms regulados por las proteínas FUR.

Como se observa en la Tabla 4, el operón *all0493-all0492* está regulado por FurA. El gen *all0492* codifica una glicerol-3-fosfato aciltransferasa, según KEGG, mientras que *all0493*, según los estudios de homología, codifica una polisacárido copolimerasa o una proteína Wzc (34). También está regulado por FurA el operón constituido por *all0495* y *all0496*. El primero de los genes, según los estudios de homología, codifica una proteína Wza o una proteína KpsD (34). En MicrobesOnline, sin embargo, aparece anotado como un gen codificante de una proteína Wza, de modo que es más probable que tenga esta función y forme parte de un sistema de síntesis de EPS Wzy dependiente. Respecto a *all0496*, en el NCBI se describe como una proteína asociada al sistema de la cianoexosortasa, que como se ha descrito anteriormente participa en el anclaje de proteínas a la membrana plasmática. Por su parte, el operón *alr0657-alr0658* está regulado por FurB. El gen *alr0657* se anota en KEGG como un gen codificante de una proteína RfbB, una UDP-glucuronato descaboxilasa, mientras que *alr0658* se describe en la misma base de datos como un gen codificante de una UDP-glucosa 6-deshidrogenasa. Se trataría, por tanto, de proteínas implicadas en la modificación de los residuos glucídicos de los EPS.

El gen *all0919* está regulado por FurA. En el NCBI se describe como un gen codificante de una glicosiltransferasa implicada en la síntesis de ácido colánico, un polímero de la cápsula de algunas bacterias (38). En cuanto al operón *alr1490-alr1491*, la regulación se da tanto por parte de FurB como de FurC. El gen *alr1490*, según KEGG, codifica la proteína de unión a ATP de un sistema de transporte de tipo ABC, mientras que *alr1491*, según los estudios de homología, codifica una proteína KpsM. Dado que las proteínas KpsT de los sistemas de transporte de tipo ABC se describen como proteínas de unión a ATP (39), es posible que el gen *alr1490* codifique para una proteína KpsT y por tanto el operón *alr1490-alr1491* codifique para las proteínas del transportador de tipo ABC de un sistema de síntesis de EPS dependiente de transportador ABC.

Por otra parte, FurA y FurB regulan el operón *all2293-all2283*, entre cuyos genes hay varias glicosiltransferasas, según la información disponible en KEGG y MicrobesOnline (*all2285*, *all2286*, *all2287*, *all2288*, *all2289*, *all2291*, *all2292* y *all2293*). Además, también forma parte de este operón el gen *all2290* que, según los estudios de homología, codifica para la proteína Wzx de un sistema de síntesis de EPS Wzy dependiente (34), y los genes *all2283* y *all2284*, que codifican un agonista de un factor anti-sigma y un regulador negativo de sigma B. Los factores sigma permiten la unión de la RNA polimerasa al DNA para su transcripción, y en el caso de los factores sigma B se ha observado que participan en la regulación de la formación de biofilms en numerosas bacterias (40). Por tanto, este operón podría codificar para proteínas implicadas en síntesis y regulación de la formación de biofilms.

En lo que respecta al operón *alr2294-alr2296*, está regulado tanto por FurA como por FurB. Según los estudios de homología, *alr2294* codifica una proteína Wza o una proteína KpsD(34). Sin embargo, en MicrobesOnline el gen se describe como codificante de una proteína Wza, de modo que es más probable que tenga esta función y el gen codifique una proteína de un sistema de síntesis de EPS Wzy dependiente. Por otra parte, según MicrobesOnline, *alr2295* codifica una proteína GumC en la síntesis de lipopolisacáridos, mientras que *alr2296* se describe en el NCBI como un gen codificante de una proteína implicada en la síntesis de polisacáridos capsulares. Respecto al operón *alr2372-alr2374*, está regulado por FurA. Según MicrobesOnline, *alr2372* es la proteína de unión a ATP de un transportador de tipo ABC que posiblemente, como se ha explicado previamente, sea una proteína KpsT, mientras que *alr2373*, según los estudios de homología, es una proteína KpsM. Para el tercero de los genes del operón no hay ninguna función descrita, pero dado que los dos primeros genes codifican el transportador de tipo ABC de un



sistema de síntesis de EPS dependiente de transportador de tipo ABC este gen podría codificar para un gen con alguna función relacionada con este sistema.

Los genes *all2770* y *alr2856* están regulados FurA. Según KEGG, *all2770* codifica una dolicol-fosfato-manosa sintasa mientras que *alr2856*, según los estudios de homología, puede ser una proteína Wzc, una proteína KpsE o una polisacárido copolimerasa (34). En el NCBI, sin embargo, se describe como una tirosina autoquinasa, de modo que probablemente se trate de una proteína Wzc de un sistema de síntesis de EPS Wzy dependiente. También está regulado por FurA el operón *alr3057-alr3060*. En este caso, *alr3057* y *alr3058* aparecen anotados en MicrobesOnline como genes codificantes de glicosiltransferasas, mientras que *alr3059*, según los estudios de homología, puede ser una proteína Wza, una proteína KpsE o una polisacárido copolimerasa (34). En lo que respecta a *alr3060*, según los estudios de homología, codifica una proteína Wzx perteneciente a un sistema de síntesis de EPS Wzy dependiente (34). Por su parte, el gen *alr3061* está regulado por FurB y, según KEGG, codifica una acetiltransferasa, de modo que en este caso se trataría de una enzima relacionada con la producción de EPS que dotasen a los biofilms de un carácter hidrófobo.

Otro operón regulado por FurA es el *alr3069-alr3071*. Según MicrobesOnline, los tres genes del operón codifican posibles glicosiltransferasas. También se relaciona con la formación de biofilms el gen *alr3464*, regulado por FurC y codificante, según los estudios de homología, de una proteína KpsF (34) implicada en la síntesis del *carrier* al que se unen los polisacáridos en un sistema de síntesis de EPS dependiente de transportador de tipo ABC. FurC puede regular también la expresión del operón *alr3756-alr3757*. La función de *alr3756* es desconocida, pero *alr3757* codifica una glicosiltransferasa, por lo que el operón podría codificar también proteínas implicadas en la formación de biofilms.

El gen *all4219* está regulado por FurA y según los estudios de homología codifica una proteína KpsM (34), una proteína perteneciente a un sistema de síntesis de EPS dependiente de transportador de tipo ABC. También está regulado por FurA el operón *alr4239-alr4240*. En el caso de *alr4239*, según KEGG, codifica la proteína de unión a ATP de un sistema de transporte de tipo ABC que posiblemente, como se ha explicado para los casos anteriores, sea una proteína KpsT, mientras que *alr4240* codifica una proteína Alg44 perteneciente al sistema de síntesis de EPS dependiente de sintasa (18). Otro gen regulado por las proteínas FUR, en concreto por FurC, es *all4388*. Según los estudios de homología, el gen codifica una proteína Wza o una proteína KpsD (34), pero dado que en MicrobesOnline aparece anotado como codificante de una proteína Wza, tendrá más probablemente esta función y codificará un componente de un sistema de síntesis de EPS Wzy dependiente.

FurA también regula la expresión del operón *all4426-all4422*. Salvo *all4425*, cuya función es desconocida, el resto de genes del operón codifican glicosiltransferasas, según la información depositada en KEGG. Por otra parte, el operón *alr4485-alr4488* está regulado también por FurA. Según los estudios de homología, *alr4485* y *alr4486* codifican, respectivamente, una proteína KpsM y una proteína KpsT (34), de modo que constituirían el transportador ABC de un sistema de síntesis de EPS dependiente de transportador de tipo ABC. Respecto a *alr4487* y *alr4488*, según la información obtenida de KEGG y MicrobesOnline, codifican una glicosiltransferasa y una proteína implicada en la síntesis del antígeno-O. Finalmente, FurA regula la expresión de *all5073* y *alr5223*, genes independientes. El gen *all5073*, según los estudios de homología, codifica la proteína Wzy de un sistema de síntesis de EPS Wzy dependiente (34), mientras que *alr5223*, según KEGG, codifica una glicosiltransferasa.

#### 5.4. Discusión global y perspectivas futuras

Dado que las proteínas FUR juegan un papel central en la respuesta a estreses en cianobacterias y que los biofilms se sintetizan en muchas ocasiones en respuesta a estos estreses, podría esperarse que las proteínas FUR actuaran regulando la formación de los mismos. Por ello, se abre la puerta al estudio del papel que juegan las proteínas FUR en la regulación de la formación de biofilms en *Anabaena sp.* PCC7120.

Estudios previos desarrollados con estirpes de sobreexpresión de proteínas FUR han demostrado cómo la sobreexpresión de estos reguladores altera la producción de exopolisacáridos y la formación de biofilms (30). La determinación de la formación de exopolisacáridos en las estirpes de sobreexpresión de proteínas FUR por el método del fenol-sulfúrico mostró que tanto la sobreexpresión de FurA como la de FurC provoca un aumento de la producción de exopolisacáridos, mientras que la sobreexpresión de FurB conduce a una menor producción de los mismos, en comparación con la estirpe silvestre. En el caso de la formación de biofilms, se determinó por el método del cristal violeta, y los resultados fueron semejantes a los de la producción de exopolisacáridos. Las estirpes de sobreexpresión de FurA y de FurC producen más biofilm que la estirpe silvestre, mientras que para la estirpe de sobreexpresión de FurB la producción de biofilm es menor (30).

Partiendo de estas observaciones se inició el estudio del papel de las proteínas FUR en la formación de biofilms en *Anabaena sp.* PCC7120. En primer lugar, se han llevado a cabo una revisión bibliográfica y estudios de homología que han permitido identificar de más de 100 genes potencialmente implicados en este proceso. Con ello, se ha ampliado el conocimiento sobre la formación de biofilms en esta cianobacteria, el cual, a diferencia de lo que ocurre en el caso de las bacterias heterótrofas, es todavía muy limitado.

Por otra parte, la búsqueda de cajas FUR y los ensayos de retardo en gel (EMSA) han permitido identificar 54 genes implicados en la formación de biofilms regulados por esta familia de proteínas, lo que podría explicar los resultados de producción de exopolisacáridos y de formación de biofilms observados en las estirpes de sobreexpresión de estas.

En este trabajo se ha demostrado que FurA regula la expresión de factores de transcripción relacionados con la síntesis de biofilms (el regulador negativo de sigma B y el agonista del factor anti sigma), así como la de numerosas glicosiltransferasas, proteínas de síntesis de polisacáridos, de lipopolisacáridos, polisacárido copolimerasas, proteínas modificadoras de monosacáridos y proteínas del sistema de la cianoexosortasa. Además, regula la expresión de proteínas de los sistemas de síntesis de EPS, como proteínas del sistema Wzy dependiente (Wzx, Wzy, Wza y Wza), proteínas del sistema dependiente de transportador de tipo ABC (KpsE, KpsT y KpsM) y proteínas del sistema dependiente de sintasa (Alg44). Teniendo en cuenta que la estirpe de sobreexpresión de FurA produce más biofilm y más exopolisacáridos que la estirpe silvestre (30), es posible que FurA pueda actuar como activador de la expresión génica de genes de glicosiltransferasas y de proteínas de los sistemas de síntesis de EPS .

Respecto a FurB, regula la expresión de factores de transcripción relacionados con la síntesis de biofilms, al igual que FurA, así como la de glicosiltransferasas y proteínas modificadoras de monosacáridos. FurB regula también la expresión de proteínas pertenecientes a sistemas de síntesis de EPS, como proteínas pertenecientes al sistema Wzy dependiente (Wzx y Wza) o proteínas del sistema dependiente de transportador de tipo ABC (KpsT y KpsM). En este caso, dado que se ha observado que la estirpe de sobreexpresión de FurB produce menos biofilm



y menos exopolisacáridos que la estirpe silvestre, es probable que FurB actúe reprimiendo los genes identificados en este estudio, de modo que su sobreexpresión disminuiría la síntesis de exopolisacáridos y por tanto la formación de biofilms.

Finalmente, FurC también se ve implicada en la regulación de la formación de biofilms. En este caso, regula la expresión de glicosiltransferasas, de proteínas del sistema Wzy dependiente (Wza) y de proteínas del sistema de síntesis de EPS dependiente de transportador de tipo ABC (KpsT, KpsM y KpsF). El papel de FurC en la regulación de la formación de biofilms en *Anabaena sp.* PCC 7120 podría ser el mismo que se ha propuesto para FurA. Es decir, FurC podría actuar como activador de la expresión génica de genes implicados en la formación de biofilms, lo que permitiría explicar por qué la sobreexpresión de FurC da lugar a una mayor síntesis de exopolisacáridos y una mayor formación de biofilms.

Los estudios previos, junto con los resultados obtenidos en este trabajo, ponen de manifiesto que las proteínas FUR desempeñan papeles importantes en la regulación de la formación de los biofilms en *Anabaena sp.* PCC 7120. *A priori*, teniendo en cuenta los resultados de formación de biofilms de las estirpes de sobreexpresión de las proteínas FUR, cabría pensar que FurA y FurC pueden ejercer un papel de activadores transcripcionales en la mayoría de casos, induciendo la expresión de los genes identificados, mientras que FurB actuaría como represor, fundamentalmente. Sin embargo, son necesarios estudios *in vivo* que permitan determinar el efecto de la sobreexpresión de las proteínas FUR sobre la expresión de los genes identificados en el estudio y, por tanto, confirmar su papel como activadores o represores de estos genes.

## 6. Conclusiones

A partir de los resultados obtenidos en el estudio se extraen las siguientes conclusiones:

- *Anabaena sp.* PCC7120 contiene en su genoma 108 genes potencialmente implicados en la formación de biofilms, los cuales han sido identificados mediante una revisión bibliográfica y estudios de homología de proteínas.
- La predicción de cajas de unión de las proteínas FUR ha permitido localizar 54 dianas directas de estos reguladores.
- Las proteínas FUR regulan la formación de biofilms a todos los niveles del proceso, desde la regulación de la expresión de factores de transcripción implicados en el proceso hasta la expresión de las proteínas implicadas en la síntesis y exportación de los EPS.

### Conclusions.

From the results obtained in the study we can conclude:

- *Anabaena sp.* PCC7120 encodes 108 genes potentially involved in biofilm formation, which have been identified developing a bibliographic revision and protein homology studies
- FUR boxes prediction has allowed the identification of 54 direct targets of FUR proteins.
- FUR proteins regulate biofilm formation on all the levels of the process, from the regulation of transcriptional factors expression to the regulation of the expression of the proteins involved in the synthesis and exportation of EPS.

## 7. Bibliografía

1. Whitton BA, Potts M. Introduction to the Cyanobacteria. In: Ecology of Cyanobacteria II. Springer Netherlands; 2012. p. 1-13.
2. Facey JA, Rogers TA, Apte SC, Mitrovic SM. Micronutrients as growth limiting factors in cyanobacterial blooms; a survey of freshwaters in South East Australia. *Aquat Sci.* 2021;83(2).
3. Fay P. Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev.* 1992;56(2):340-73.
4. Yadav RK, Abraham G, Singh YV, Singh PK. Advancements in the utilization of *Azolla-Anabaena* system in relation to sustainable agricultural practices. *P Indian Aca Sci.* 2014;80(2):301.
5. Anderson DM, Glibert PM, Burkholder JM. Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. *Estuaries.* 2002;25(4):704-26.
6. Ríos ADL, Grube M, Sancho LG, Ascaso C. Ultrastructural and genetic characteristics of endolithic cyanobacterial biofilms colonizing Antarctic granite rocks. *FEMS Microbiol Ecol.* 2007;59(2):386-95.
7. Flemming HC, Wingender J. The biofilm matrix. *Nat Rev Microbiol.* 2010;8(9):623-33.
8. Pereira S, Zille A, Micheletti E, Moradas-Ferreira P, Philippis RD, Tamagnini P. Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. *FEMS Microbiol Rev.* 2009;33(5):917-41.
9. Philippis RD, Colica G, Micheletti E. Exopolysaccharide-producing cyanobacteria in heavy metal removal from water: molecular basis and practical applicability of the biosorption process. *Appl Microbiol Biot.* 2011;92(4):697-708.
10. Philippis RD, Sili C, Paperi R, Vincenzini M. Exopolysaccharide-producing cyanobacteria and their possible exploitation: A review. *J Appl Phycol.* 2001;13(4):293-9.
11. Kawaguchi T, Decho AW. Biochemical Characterization of Cyanobacterial Extracellular Polymers (EPS) from Modern Marine Stromatolites (Bahamas). *Preparative Biochem Biotech.* 2000;30(4):321-30.
12. Tamaru Y, Takani Y, Yoshida T, Sakamoto T. Crucial Role of Extracellular Polysaccharides in Desiccation and Freezing Tolerance in the Terrestrial Cyanobacterium *Nostoc commune*. *App Environ Microb.* 2005;71(11):7327-33.
13. Kidron GJ, Yaalon DH, Vonshak A. Two causes for runoff initiation on microbiotic crusts: hydrophobicity and pore clogging. *Soil Sci.* 1999;164(1):18-27.
14. Phoenix VR, Adams DG, Konhauser KO. Cyanobacterial viability during hydrothermal biomineralisation. *Chem Geol.* 2000;169(3-4):329-38.
15. Adhikary SP, Sahu JK. UV protecting pigment of the terrestrial cyanobacterium *Tolypothrix byssoidea*. *J Plant Physiol.* 1998;153(5-6):770-3.
16. Parker DL, Schram BR, Plude JL, Moore RE. Effect of Metal Cations on the Viscosity of a Pectin-Like Capsular Polysaccharide from the Cyanobacterium *Microcystis flos-aquae* C3-40. *App environ microb.* 1996;62(4):1208-13.
17. Philippis RD, Paperi R, Sili C. Heavy metal sorption by released polysaccharides and whole cultures of two exopolysaccharide-producing cyanobacteria. *Biodegradation.* 2006;18(2):181-7.
18. Pereira SB, Mota R, Vieira CP, Vieira J, Tamagnini P. Phylum-wide analysis of genes/proteins related to the last steps of assembly and export of extracellular polymeric substances (EPS) in cyanobacteria. *Scientific Reports.* 2015;5(1).
19. Pereira SB, Santos M, Leite JP, Flores C, Eisfeld C, Büttel Z, et al. The role of the tyrosine kinase *Wzc(sll0923)* and the phosphatase *Wzb (slr0328)* in the production of extracellular polymeric substances (EPS) by *Synechocystis* PCC6803. *Microbiology Open.* 2019;8(6).
20. Rai J, Kumar D, Pandey LK, Yadav A, Gaur JP. Potential of cyanobacterial biofilms in phosphate removal and biomass production. *J Environ Manage.* 2016;177:138-44.

21. Roeselers G, van Loosdrecht MCM, Muyzer G. Phototrophic biofilms and their potential applications. *J Appl Phycol.* 2007;20(3):227-35.
22. Bharti A, Velmourougane K, Prasanna R. Phototrophic biofilms: diversity, ecology and applications. *J Appl Phycol.* 2017;29(6):2729-44.
23. Schütz K, Happe T, Troshina O, Lindblad P, Leitao E, Oliveira P, et al. Cyanobacterial H<sub>2</sub> production, a comparative analysis. *Planta.* 2004;218(3):350-9.
24. Fillat MF. The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. *Arch Biochem and Biophys.* 2014;546:41-52.
25. Sevilla E, Bes MT, Peleato ML, Fillat MF. Fur-like proteins: Beyond the ferric uptake regulator (Fur) paralog. *Arch Biochem Biophys.* 2021;701:108770.
26. González A, Bes MT, Barja F, Peleato ML, Fillat MF. Overexpression of FurA in *Anabaena sp.* PCC7120 Reveals New Targets for This Regulator Involved in Photosynthesis, Iron Uptake and Cellular Morphology. *Plant Cell Physiol.* 2010;51(11):1900-14.
27. Sein-Echaluce VC, González A, Napolitano M, Luque I, Barja F, Peleato ML, et al. Zur (FurB) is a key factor in the control of the oxidative stress response in *Anabaena sp.* PCC7120. *Environ Microbiol.* 2014;17(6):2006-17.
28. Sevilla E, Sarasa-Buisan C, González A, Cases R, Kufryk G, Peleato ML, et al. Regulation by FurC in *Anabaena* Links the Oxidative Stress Response to Photosynthetic Metabolism. *Plant Cell Physiol.* 2019;60(8):1778-89.
29. Sarasa-Buisan C, Guio J, Broset E, Peleato ML, Fillat MF, Sevilla E. FurC (PerR) from *Anabaena sp.* PCC7120: a versatile transcriptional regulator engaged in the regulatory network of heterocyst development and nitrogen fixation. *Environ Microbiol.* 2021.
30. Sandoval A. Papel de los reguladores transcripcionales FUR (*Ferric Uptake Regulator*) en la síntesis de exopolisacáridos y formación de biofilms en *Anabaena sp.* PCC7120 [Tesis doctoral]. Universidad de Zaragoza; 2021.
31. Kehr JC, Dittmann E. Biosynthesis and Function of Extracellular Glycans in Cyanobacteria. *Life.* 2015;5(1):164-80.
32. Yoshimura H, Okamoto S, Tsumuraya Y, Ohmori M. Group 3 sigma factor gene, *sigJ*, a key regulator of desiccation tolerance, regulates the synthesis of extracellular polysaccharide in cyanobacterium *Anabaena sp.* strain PCC7120. *DNA Res.* 2007;14(1):13-24.
33. Ciebiada M, Kubiak K, Daroch M. Modifying the Cyanobacterial Metabolism as a Key to Efficient Biopolymer Production in Photosynthetic Microorganisms. *Int J Mol Sci.* 2020;21(19):7204.
34. Pereira SB, Mota R, Santos CL, Philippis RD, Tamagnini P. Assembly and Export of Extracellular Polymeric Substances (EPS) in Cyanobacteria. In: *Adv Bot Res.* Elsevier; 2013. p. 235-79.
35. Oliveira P, Pinto F, Pacheco CC, Mota R, Tamagnini P. HesF, an exoprotein required for filament adhesion and aggregation in *Anabaena sp.* PCC7120. *Environ Microbiol.* 2014;17(5):1631-48.
36. López-Gomollón S, Hernández JA, Angarica VE, Peleato ML, Fillat MF. Cross-talk Between Iron and Nitrogen Regulatory Networks in *Anabaena (Nostoc) sp.* PCC 7120: Identification of Overlapping Genes in FurA and NtcA Regulons. *J Mol Biol.* 2007;374(1):267-81.
37. Rost B. Protein structures sustain evolutionary drift. *Fold Des.* 1997;2:S19-24.
38. Hanna A, Berg M, Stout V, Razatos A. Role of Capsular Colanic Acid in Adhesion of Uropathogenic *Escherichia coli*. *Appl Environ Microb.* 2003;69(8):4474-81.
39. Pavelka MS, Hayes SF, Silver RP. Characterization of KpsT, the ATP-binding component of the ABC-transporter involved with the export of capsular polysialic acid in *Escherichia coli* K1. *J Biol Chem.* 1994;269(31):20149-58.
40. Jäger S, Jonas B, Pfanzelt D, Horstkotte MA, Rohde H, Mack D, et al. Regulation of Biofilm Formation by sigma B is a Common Mechanism in *Staphylococcus epidermidis* and is not Mediated by Transcriptional Regulation of *sarA*. *Int J Artif Organs.* 2009;32(9):584-91.

---

MATERIAL SUPLEMENTARIO

---

## 8. Material suplementario

### 8.1. Programa del termociclador para la PCR

1. Desnaturalización inicial del DNA: 5 min a 95 °C
2. Desnaturalización del DNA: 30 s a 95 °C
3. Hibridación del DNA: 30 s a una temperatura 2 °C inferior a la T<sub>m</sub> de cada par de cebadores
4. Etapa de elongación: 45 s a 72 °C
5. Elongación final: 10 min a 72 °C
6. Conservación: a 4 °C

Los pasos segundo a cuarto se repitieron cíclicamente un total de 30 veces. La temperatura de hibridación (T<sub>m</sub>) dependió de la pareja de oligonucleótidos empleados, siendo siempre 2 °C inferior a la temperatura de hibridación de los oligonucleótidos. En la sección 8.2 de Material Suplementario se detalla tanto la T<sub>m</sub> como la secuencia de cada uno de los oligonucleótidos utilizados.

### 8.2. Oligonucleótidos utilizados para la amplificación de las regiones promotoras de los genes potencialmente implicados en la formación de biofilms en *Anabaena sp.* PCC7120

Oligonucleótido	Secuencia 5' → 3'	T <sub>m</sub> (°C)	Secuencia amplificada
Pall0042 up	TTTACCGATATAGTTTAACTGG	60	Promotor del gen <i>all0042</i> de <i>Anabaena sp.</i> PCC7120
Pall0042 dw	AACCGCAATGCGATACTCAG		
Pall0059 up	TCGCCAGGAAATAATCAGGG	60	Promotor del gen <i>all0059</i> de <i>Anabaena sp.</i> PCC7120
Pall0059 dw	ACGCTACCAATTCATCAGTAG		
Palr0267 up	TTTGAGGATGGTATTCCTAGC	60	Promotor del gen <i>alr0267</i> de <i>Anabaena sp.</i> PCC7120
Palr0267 dw	ATTTGAGACTGCTATGTTAAGC		
Palr0277 up	TTAGCCTAACGGGATGCAAG	60	Promotor del gen <i>alr0277</i> de <i>Anabaena sp.</i> PCC7120
Palr0277 dw	TACCATCAGTTTGCATAGAGG		
Pall0493 up	GATACTGATGGCAGAAAATAAAG	62	Promotor del operón <i>all0493-all0492</i> de <i>Anabaena sp.</i> PCC7120
Pall0493 dw	GTTTAATTCCTGACCTTGTTGC		
Pall0496 up	AGAAAGCCAGCTACTTTCCC	60	Promotor del operón <i>all0496-all0495</i> en <i>Anabaena sp.</i> PCC7120
Pall0496 dw	AAGTAGTTAAGTGCTTTTGTTTG		
Pall0497 up	TACCCAGGAAAGTAAAGTAGC	60	Promotor del gen <i>all0497</i> de <i>Anabaena sp.</i> PCC7120
Pall0497 dw	TTCGTGTTCTGTGCTGATGG		
Palr0657 up	TCCTAACACTACGACAGCAG	60	Promotor del operón <i>alr0657-alr0658</i> de <i>Anabaena sp.</i> PCC7120
Palr0657 dw	CGCCAGTCACCAAAATCTC		
Palr0776 up	ACCAGCCGGATTTGTAGTG	60	Promotor del gen <i>alr0776</i> de <i>Anabaena sp.</i> PCC7120
Palr0776 dw	CAACCAAATGACCAAAGATAAC		
Pall0917 up	GCGAGTAGTAGGTAAACTGC	60	Promotor del operón <i>all0917-all0916</i> de <i>Anabaena sp.</i> PCC7120
Pall0917 dw	TGCCTAAGCATACTCTCAGG		

Pall0919 up	TTCCAGAAGTTCTTATCTATGC	60	Promotor del gen <i>all0919</i> de <i>Anabaena sp.</i> PCC7120
Pall0919 dw	GATTGACAACGACGTTTCTG		
Palr1490 up	TGACGATTGCGGCATAAAGC	60	Promotor del operón <i>alr1490-alr1491</i> de <i>Anabaena sp.</i> PCC7120
Palr1490 dw	TGATGTAATTGTCAACATAGGG		
Pall2293 up	ACAATTAAGGTTTGGACATGG	60	Promotor del operón <i>all2293-all2283</i> y <i>alr2294-</i> <i>alr2296</i> de <i>Anabaena sp.</i> PCC7120
Pall2293 dw	GTCGAATAAGTATTTACTAAACC		
Palr2372 up	TGAGGATAGATATCTCATTTG	60	Promotor del operón <i>alr2372-alr2374</i> de <i>Anabaena sp.</i> PCC7120
Palr2372 dw	GACGGCTTCAACAGTACCG		
Pall2770 up	ATTCATCTGCTTACCATCACC	60	Promotor del gen <i>all2770</i> de <i>Anabaena sp.</i> PCC7120
Pall2770 dw	TGCGGGTACTGGCAACAATG		
Palr2832 up	TACCCATCTCAAAAACGCC	60	Promotor del operón <i>alr2832-alr2833</i> de <i>Anabaena sp.</i> PCC7120
Palr2832 dw	TCATGGACTAGAGCAACTTTC		
Palr2856 up	TGATAGGTACTGATCTAGGTG	60	Promotor del gen <i>alr2856</i> de <i>Anabaena sp.</i> PCC7120
Palr2856 dw	CTGTTGGAAGTTAATCAGGTC		
Palr2861 up	CAGCATCAAACAGCTTATCTTG	60	Promotor del gen <i>alr2861</i> de <i>Anabaena sp.</i> PCC7120
Palr2861 dw	TGCCAAACAACTATAAATGCC		
Palr3057 up	AGCAGATTTTTCAAGCCACTAG	60	Promotor del operón <i>alr3057-alr3060</i> de <i>Anabaena sp.</i> PCC7120
Palr3057 dw	ACGCATTACTCACACCCTAG		
Palr3061 up	CACAGAATAGTCTTTGCTTGG	60	Promotor del gen <i>alr3061</i> de <i>Anabaena sp.</i> PCC7120
Palr3061 dw	TCTCCATGTGCGAGTACTAC		
Palr3068 up	CCGCTAAAAGATTAAAAATGGC	60	Promotor del gen <i>alr3068</i> de <i>Anabaena sp.</i> PCC7120
Palr3068 dw	TGCGGTCACTGGAACCATC		
Palr3069 up	TGCTCAGGATTTAATTATCTGG	60	Promotor del operón <i>alr3069-alr3071</i> de <i>Anabaena sp.</i> PCC7120
Palr3069 dw	TAATGGTGATAGATGATTTTTGC		
Palr3073 up	TCTCATGTATTGTTTCATACCTG	60	Promotor del operón <i>alr3073-alr3074</i> de <i>Anabaena sp.</i> PCC7120
Palr3073 dw	ATTACAATGATTTCTAAATCTGAG		
Palr3464 up	TAGCTTTGTAGGGTGGGTAG	60	Promotor del gen <i>alr3464</i> de <i>Anabaena sp.</i> PCC7120
Palr3464 dw	GTTGCCCAACGTGTATGACC		
Pall3509 up	TTACCCAATCCCCATTACCC	60	Promotor del gen <i>all3509</i> de <i>Anabaena sp.</i> PCC7120
Pall3509 dw	GCAATAAATCCTGCTGCTCC		
Palr3756 up	CGATCGCACCTGAGCAATAG	62	Promotor del operón <i>alr3756-alr3757</i> de <i>Anabaena sp.</i> PCC7120
Palr3756 dw	GGAAACAATGCCATCCATTTCC		
Pall4219 up	GGGTTGAAGTAAAGCAAGCTG	60	Promotor del gen <i>all4219</i> de <i>Anabaena sp.</i> PCC7120
Pall4219 dw	CAGGATTTGCAGTTGTGAGG		
Palr4239 up	AATAGTGGAATTTTACTTTGGG	60	Promotor del operón <i>alr4239-alr4240</i> de <i>Anabaena sp.</i> PCC7120
Palr4239 dw	ACAAGAAATCTTGAAAAGCGC		
Pall4388 up	GATCAATCAGAAGTATAGCCG	60	Promotor del gen <i>all4388</i> de <i>Anabaena sp.</i> PCC7120
Pall4388 dw	GCTGGCTGTGATAGAAATTTTC		
Pall4426 up	GACATTGGCGTTATTTGGCTAC	62	Promotor del operón <i>all4426-all4422</i> de <i>Anabaena sp.</i> PCC7120
Pall4426 dw	TCTAGCAATTCCTTGAAATAGTC		
Pall4432 up	GAGTGCTTTGGAGTTAAGTG	60	Promotor del operón <i>all4432-all4426</i> de <i>Anabaena sp.</i> PCC7120
Pall4432 dw	CCTTGTAAGAGTCCATTGTC		
Palr4485 up	GAAACTTCAGTTAAGCAACCAC	62	Promotor del operón <i>alr4485-alr4488</i> de <i>Anabaena sp.</i> PCC7120
Palr4485 dw	CTGCACTGCGCCAGTTTG		
Pall5073 up	CACGGGAATCAGGACATCG	60	Promotor del gen <i>all5073</i> de <i>Anabaena sp.</i> PCC7120
Pall5073 dw	TAGCTAGAACTGTTTTAGAGG		
Pall5222 up	GGGAACAGCAACAGGTTGTG	60	Promotor del gen <i>all5222</i> de <i>Anabaena sp.</i> PCC7120
Pall5222 dw	ACCAATGATGTTAACC GAATAC		

<i>Pall5223</i> up	TTGAAAACCTGGGAATTAAGAGG	60	Región promotora común al gen <i>all5222</i> y el gen <i>alr5223</i> de <i>Anabaena</i> sp. PCC7120
<i>Pall5223</i> dw	TACAATGTACCGATTTTGTTG		
<i>Palr7196</i> up	ACTCTTGGCTGGCAATTGAG	60	Promotor del gen <i>alr7196</i> de <i>Anabaena</i> sp. PCC7120
<i>Palr7196</i> dw	GCATCTAATCTTAGTCCGGC		
<i>Pall1691</i> up	CATATGACTGTCTACACAAATACTTCG	72	Promotor del gen <i>all1691</i> ( <i>furA</i> ) de <i>Anabaena</i> sp. PCC7120
<i>Pall1691</i> dw	GGATCCCTAAAGTGGCATGAGCG		
<i>Palr1911</i> up	GCCTACTCTGCGAGTTCTCCG	68	Promotor del gen <i>alr1911</i> ( <i>nifJ</i> ) de <i>Anabaena</i> sp. PCC7120
<i>Palr1911</i> dw	GGCCTGTGAGAGTTGCTGCAC		
<i>Pall4725</i> up	CTCCGGTGGCACAGGTATTGGC	72	Promotor del gen <i>all4725</i> de <i>Anabaena</i> sp. PCC7120
<i>Pall4725</i> dw	CAGTGTTGCAGTCCGACGCAACCG		
<i>Palr0094</i> up	GCGTTTAGTTTATCCGCAAA	62	Promotor del gen <i>all0094</i> ( <i>hetZ</i> ) de <i>Anabaena</i> sp. PCC7120
<i>Palr0094</i> dw	CTCAAGCATTGTTGTAGCCG		

**Tabla 5.** Listado de oligonucleótidos utilizados en la amplificación de las regiones promotoras de los genes potencialmente implicados en la formación de biofilms.

**8.3. Matrices de peso para la búsqueda de cajas FUR**

0.642857 0.142857 0.000000 0.214286  
0.428571 0.000000 0.357143 0.214286  
0.500000 0.000000 0.142857 0.357143  
0.571429 0.000000 0.000000 0.428571  
0.785714 0.000000 0.214286 0.000000  
0.642857 0.000000 0.000000 0.357143  
1.000000 0.000000 0.000000 0.000000  
0.214286 0.000000 0.142857 0.642857  
0.142857 0.214286 0.000000 0.642857  
0.000000 0.642857 0.000000 0.357143  
0.071429 0.000000 0.000000 0.928571  
0.000000 0.642857 0.357143 0.000000  
0.928571 0.071429 0.000000 0.000000  
0.714286 0.000000 0.000000 0.285714  
0.214286 0.071429 0.071429 0.642857  
0.785714 0.000000 0.000000 0.214286  
0.785714 0.000000 0.000000 0.214286  
0.642857 0.000000 0.357143 0.000000  
0.000000 0.285714 0.000000 0.714286

Matriz de peso utilizada para la búsqueda de cajas de unión de FurA.



0.333333 0.266667 0.266667 0.133333  
0.666667 0.066667 0.066667 0.200000  
0.266667 0.000000 0.266667 0.466667  
0.333333 0.200000 0.133333 0.333333  
0.600000 0.200000 0.133333 0.066667  
0.000000 0.000000 0.000000 1.000000  
0.066667 0.000000 0.933333 0.000000  
0.933333 0.000000 0.000000 0.066667  
0.066667 0.000000 0.066667 0.866667  
0.933333 0.000000 0.000000 0.066667  
0.933333 0.000000 0.000000 0.066667  
0.000000 0.200000 0.000000 0.800000  
0.133333 0.333333 0.400000 0.133333  
0.866667 0.000000 0.066667 0.066667  
0.066667 0.000000 0.000000 0.933333  
0.200000 0.000000 0.000000 0.800000  
0.733333 0.133333 0.066667 0.066667  
0.000000 0.066667 0.000000 0.933333  
0.133333 0.800000 0.000000 0.066667  
0.800000 0.066667 0.066667 0.066667  
0.533333 0.000000 0.133333 0.333333  
0.400000 0.066667 0.066667 0.466667  
0.666667 0.133333 0.000000 0.200000  
0.400000 0.133333 0.066667 0.400000  
0.400000 0.200000 0.133333 0.266667

Matriz de peso utilizada para la búsqueda de cajas de unión de FurB.

0.111111 0.500000 0.000000 0.388889  
0.666667 0.000000 0.222222 0.111111  
0.777778 0.000000 0.000000 0.222222  
0.666667 0.166667 0.055556 0.111111  
0.444444 0.222222 0.333333 0.000000  
0.000000 0.000000 0.111111 0.888889  
0.000000 0.944444 0.055556 0.000000  
1.000000 0.000000 0.000000 0.000000  
0.333333 0.000000 0.000000 0.666667  
0.500000 0.111111 0.166667 0.222222  
0.888889 0.055556 0.000000 0.055556  
0.166667 0.611111 0.000000 0.222222  
0.333333 0.166667 0.444444 0.055556  
1.000000 0.000000 0.000000 0.000000  
0.166667 0.388889 0.222222 0.222222  
0.277778 0.000000 0.000000 0.722222  
0.222222 0.111111 0.055556 0.611111  
0.000000 0.166667 0.055556 0.777778  
0.277778 0.000000 0.722222 0.000000

Matriz de peso utilizada para la búsqueda de cajas de unión de FurC.

#### 8.4. Secuencias intergénicas en las que se detectaron cajas FUR

*a110042*

ACAATGCCAGTGACTGTTTCTTGTACCCGACCGTCTGGATAAATGATGAATTCTAATGTTTCCATACTCT  
TAGCCAACCACAATAAGTCCGCTTGAATTTTACTGCCTTAGCAGCAAGCTATATGTATAGCTCCAGGAAC  
ATTGTTACTTAGATACAAACTTGCCATTTTTTAACTAATGTTCCATCTCTCAAATAAAATATCTCAGAAA  
ATTTACAAAAATTATGAATTTTATAAGTCAACACATCTGTTATTAGTTAGTTTCTCTTAACCTAATCCAT  
AAGTCAAGGTAATCTCGTGCTGATGGTTTTGGCGGTAAAAGCTTTGAAAAATCGGCACTAGGGGTAGTAT  
AGCTAGCTGTGTGACTGTAAAAATAGCAAATTTATTGAAAATTCTTATCATTATCATCGGCTTTGACTT  
ATTCTCGGATGGGCAATACTGCACTGTGTTAGGAGCAGTTAAGGATTACAAGAGCAAAAAATGCTGTAATC  
TCAACATTAAGTATTTATCTATACAAAAAACATTACAAATGATGAGGTAGTATCCATGAGACAAGAGAT  
TTTTGGCAGATTTCACTGTTAGTAGTATTGGTGTCAACGTTCTGCTACTTAGTATCACTAACGCATTACA  
CTGCTAATGGGAGAATATATTTCCTTTTAGATTTTAGGACTTAACCTCTAATGCCTATTAGATCCTAACGT  
CCTTAATTAATTTCTTTTTTAAAGCAAAAAATCTCTTAACCTTTTTTATCAAGCCACATTGAGGGACTAGGG  
TTAAATTAACCTTTAAAGAAGCTATTTTTTAACAAAAACATGGTAACTGAGATTGCCAGTTAGTTCTAATT  
CATCATTTCTTTAATTTTAAAGTAACTGTGTAAATCTCATTTGTTTAACTTTACCGATATAGTTTAAAC  
TGGTATTCTAAAATATTAGTTTCTCTTTAATTGTTAATTTATTTGCAGAGGTAAGCTATGTTTTGAAAA  
AGATTGAACTAGAAGTATATTATAGAAATTAAGAAAGAGAACCAGAACTTTTTTCGGGTGTTCTCTGTC  
TTCTCCTCAAGGGTTTATCATTATGAACCCATTAATCATCAATGGTATACTACAAACTGGTTGAAGATTT  
ACTTTAAAAATAGCTAAGAGATTAACAGGTAAAGTTTTCTTTGCCTTGCCAAAATTCCTGCATTCAGAGT  
TTTTAGTATATCAATACTTTATGAAAAATCAGGTATCTTTCCATAAGTTGACTGACAATTTGCAGTCCAT  
GTATTTAACTGAGTATCGCATTGCGGTTTTAATTCCTTGCCGCAATGAAGCCTT

*a110059*

AAAAAGCAAAGAATATTGAAGGTTACGTAGCCTTTTGGAAAGGTGTAAGGTTGAGACTAAATAAGCATC  
ATCAGCCTAACAACTAACTTTGTGGAGACGCGAGCAATCGCGTCTTTTTTTTTATCATAATAAATTTTTT  
TCATCTACCATCAGTTTCTCTTGTGCGCCAGGAAATAATCAGGGAATTTCAAGTAAGCTTTAAATTTTATA  
TTCAAATTTAGTTGTAATCAGATAATATTCATTATTTTTTAACTAGAAATTAAGTTAACAGTGAATC  
AACCTATACCTTGGGAAAATTTAATTATTTGTTTCATGTATATTTTTTTCGAGAATATCAGATTTACTATG  
ATGTTTAGTTAAATAAATAATGCAATATTTAATTGTAATTTATTTAGCGATCGCCTACACATACACTATC  
CCTAAAAGCTATACAAATATTCAAGATATCGCCTTACATACTAATATGTACAAGTAATACAAGTCACTTA  
ATTACAATCAAGTGAATTAATAATTTTTGTTAATATCCTGCCATATAAAAAATAAAGTGTCTACT  
GATGAATTGGTAGCGTTATTTATAACAACCTCTTCTTTAAAGAAGAACCTGAATATTACTTCTAACATA  
TTAATAAGGATATCGTGAAAAAATATCAGAGAAGTTGACACTACAAACACTCTTGTTTTAAGCTCAATGT  
GCTTGTTATTCATTAACCTATGTAATCAATAACTTCAGTTTTTATACAAGTAATAGTTGTAATTAATCAA  
GTATATATGTGTCCGTAATATAACTTTTTAATGGAAAGATTATGTTAAGAGAGTATTTAATTAATAATAT  
TTATTTTTTCTAGAACTGACCGATATCCGAGGGATTGTATAGTTGGCTAACATCAGTTTAAAGCAAGGAC  
AATTTGTTACCAATTCCTTTGCCAAATAAATGGAAAT

*alr0267*

ATCCCTGTACCACTACCAATCTTCGCCATTAATTGCTCACGACGCTGCCGATATTCTGCTTGCATAACCG  
TTAATCTTGAAATTTTTTGTATAATTTTACACTTTTGGCAACTATTATGTCTGAAAGTTACTATCTTGCA  
GTAGTTGCTAATTTAAGTATTTAACCAATATTTATTTCAACTCAAGTGTTAGTACAAGAAATCAGACAT  
CTTTTGAGGATGGTATTCCTAGCATAAATTAGTTTCGACAGTTGAAATATTTATTTGCTATGACACAAA  
TAATTTTTCTAAGTAATTTTAAATCCGTCTAATTTTTGTGTTTGGATTTGCATTATTTAACTATATATT  
TTGTGATTTATGCTGCTTTGTGCAGTTTTATTAATAAATGCTGATATAACTATTAGTTTTTATCCTTTGC  
ACAAATTAAGAAAAGGATTTTTTAGCCAATTTTTTCACGAATTAATTTTATGTCTGTTGCATTGAAGAC  
GGTTTGCTTAACATAGCAGTCTCAAATTACATCTGGTAAAACCTAACTTCTAATCACTATATCCCATTCC  
ATCAAGTGAATTGGCATCATTATCAAAGAGATACGAAGCAATATGTATATCTAGATATCTTGACATGAAA  
AATCATGATTTGAGGATTTAGGTATTGTTCAAAACAACATTGATTTTTTCGGAAATATGTTTAAATTA  
CCGTAAAAATTTATATGTAAAAATTTGATATTTCTCAAAAAATAGGAGTGTGGATACAGTAATTTCTAACT  
TAAAATAATGTTATTAGCTAACAGAAGCCAACATGAATTATCAGCTAAAACTTTTAAGACAACAAGTTTA

AGAAACCACCATGCCAGAAACGCAAAGCTTTTCCAAGAGAATAGATAAATTCGCTTAATACATCTTTTGCA  
AGATC

*alr0277*

TTCAAATAGGTTTTTTGTCTAATAAACCGACAATCATTAGTAGCAGTATTACTATAGCTGCATGAATCAA  
TTCAAAATACCCTACGGGAAGCGCAAGCGCTACAAAATTAAGACAGCCACCCACAAGGGGCGTTAGCCT  
AACGGGATGCAAGTCCGGTTTTTTAACCTAACTTTTTCTCGTAAACTTTCCGTTTAAACAACCCAGTTT  
TTTTTAACAACTGGGGTTTTAGTTTTATGTATGTATTTCCAGCAGAAATAACTGAGTTTGTATCAGGA  
ATAAATATTATACAGAGAATATTTTTTCATTGAATAAAATAAAAATACTTATAACAGAGATAACTTATAAAT  
CATGCGTACAAACACTGAACGGCAGAGTCAGCCCCCTGGTAAGTTCAGTGAACGACGATGGTTTATTCA  
GGCGTACTAAGGAAAATAATGACAGCTTTGAAGAACATATTACGCAACTCCAATTTGGTGAAAAAAGGA  
GTTACCACAAGGGAATTTTTATATGGCAGCAAGTGAGTCCTCTATGCAAACCTGATGGTATGGAGTTATTG  
CGCTCATACCACCATAA

*al10493*

ATAAATGATGAATAATAGCTACTAAATGTGACAATATCTTGTGGATTGACCGCATAAAATTTTCATTAATT  
TTATATGTCTCTTTTATGACCCACCAATTGTTAAGAGCTATATTATGCTTTTGAAAAATATAAATGGA  
TTGGATTAGC

*al10496*

CAGCACCTCTGGAGAGTGAAATAGAGGGAGAAAGCCAGCTACTTTCCCTGAGATTACTGATTAAACCAT  
CATTGCATGAGTATTATTGCGTAGAAAAAGTGGACTGAGGAAGCATCTTAGTAAAAATATTTCCCTCAGT  
CCTCACCATGTTTTAAATTGAGCGAATTTGAATTTATAAGTAGCTTCTGGGTATCTAGCAGAGAAAATTT  
ACCAATACTTCACATACTTGGACAGAGTGAATAAAATTTACTGTACAGATTCTAGGGTATAAACGTAGATA  
GATTGACAATTTTACTGAACCTATAGTGTCAATTTTTCCAGTAAAATTTGCGTAAATATAGCAATTTTT  
ACTAATGATCCCTTTATCTAAATTTCTCAAACAAAAGCACTTAACTACTTTAGTAGTGTGTTGTTGTTATT

*al10497*

CCAAAGCAGTAGGAGATTTTCTGCAAAACCAGCATCCACTACCCAGGAAAGTAAAGTAGCATAAGTTTC  
AGTGCTGCATTTCTTTAAATACCCACAAGGTTATATCGTAATAGGTAATTGGTGTTTTTATAATTACCTA  
TTACCAATGGTACCCTTAATTTGAACTCCCTCAAAGAGAGTTGATCAACTTATATATTCTGATTTTTTG  
AACACTTAAATTTGAATTGCTTACGCTTATGGCACTTCAGCAACAGATTAGAACCCGAAACACGGATCA  
ATTACTAGGCATAGGCATTTTAGG

*alr0657*

AACCGCAACGCCTGCAAACAGGTAATATCACCTCTGTAGCTCCTAACACTACGACAGCAGTATAAACACC  
CCTACACCCTTACACCCCATACCCTTAATTCCTGACAACGTACTCTAACAAGACTTTATTTTGTGTAAG  
ATTTGATGAATTATTGATCAAAAGACACAAATACTTCATGATGTTTGTATTTTGCTAAGTGAATTAACA  
AAAATCTAAAGAAAATGTAACACCCCAAGCAAAAATCTGGAATGAGTATTGTTCCCTGCTGTATTGTCAT  
ATTTGGGCTAGATGGTTTAGAACCAATATCAAAAGCTACTAACAATATCTATTAAGTAACTGGGAAAATA  
TGAGAATTTGGTGACTGGCGGTGCAGGATTTATGGTTCTCATCTTATTGATCGGCTGATACC

*alr0776*

TATTTTTGACTAGCTGCGATCGCCTCTGCAATAGAATTGACTTGTACCCTGTCTTGTGTTTCGCATCAAAAT  
ACTACAAAAATTTTTACCACACTCAGCAAATATTGTATTGTGTTGAGGCTAATCAATTCAAAATTCAA  
AAATAATTAGTCCCATAGTATTGAAGCGGTAACAAACAGATGGGTGCGTTCTTTTTAGGGTTAATATTG  
TTATCTTTGGTCATTTGGTTGGGATTACTATTTTT

*al10917*

ACAAGGTTGTTAACCGTTATATTACTACCCGTTACGCTCAAAAACACTTGCAAAATTTGTCATAATATTA  
 TCCTTTCTGTGATTGATAAACGATAGGATACAACATCAAACCTTTTATCAACTCCCATACCCAAAACAAGT  
 ATTTTCAGAACAAAGGGTCAAAAACGATCTCTGAATTACTACTTAAATCATTACCTTACAAATCTTGAGG  
 TTCCACAATCATGCTGCCTGAGATAGTTTATCAGCCTGAGAGTATGCTTAGGCATCCGATACAATTTGTT  
 AACT

*a110919*

GCATGGCGTAAATTTCTTAAACATCCCCTGGTAGTTGGGAGAGAATTTCTTCGGTTAGAGACATATAAA  
 TATTATCAAGTTTATCTTTCTTCCAGAAGTTCTTATCTATGCAATATAAACCAATACAGTGCAGTTAAGA  
 ATACATTATGAAATTAATACATTTTGGTTAAAAAAGATTATAGATATCTTACCAATAAAATCAGCAAGTT  
 TAGCTAACACTGAATATCTACGTAATACGCTAGTAATTAATGTTTCATCAAAAAACTACATTATATGATT  
 TTTCTACTGATATTTAAAAATCTATCTCCAGAGAGATTCTCTGTTTTCTATTGAACAATCATAAATCAT  
 GATATATACTCATTTCGTGTGCAATAAGAGCAATTAAATATTTTAAATGGGTGCAACTTTTCAGGTTTAC  
 TTAGGTAAATTATTGCAAGCTTATGGTGATGAAGAGTTGAGATAAAAAACATGAAGATTGCTGTTTTAGTT  
 GACAAATTTCCAGTTATTTTCAGAAACGTTTCGTTGTCAATCAAAT

*alr1490*

ATGGCTGCAATTTCCAAGCAGACTATGATAAAAAATACCAATAATTTCCAATCCATTTAATCCACATGATAC  
 TTACCTATATTTGTTAATTCAGAAGTTGATAAATTTGTGATGATTGAAATAATTACTACAATAGATGTATA  
 GATCAGCATTGTGTACTGGGATAACCAATTTGGGAAGATAAATTTGCAAAATTTGTCTCTTGTAGAGTTTGAA  
 TATTTACGAGCCATCAAAGTTAATGGGCATTGAAGTTTAAAAGTTAACAGGGTCAGCCCTCTAGAAATA  
 CAAATCCATAACCAATCCAGAGCCATACATCAATTTTATTGACGATTGCGGCATAAAGCATATAAAAAAT  
 TACAAAATGAAGAATACCCATATCAACGTGTGAATGATTTAATTAGAGTCAATTTGGTTTTATTTTCC  
 ATATTGTCTCAGTATGCTTTCTCCACCTTTGGTAGTAAAATGCGACAAAAGATATATAAAGGATCTGCGGG  
 TTAATAGTGTCTAACCAGCCGGTTTTCGACTCCGCTCAACCCTCATCTGTGCGAGAGTTGAGCGGAGTCG  
 AACTCGGTTTACTGGTACTTTATTTTCATGCAAGTCCCTAAGCTATTACGAGTGTCCCCAGTTTAAAC  
 AAAGAAGTCTGGTATTTACGTTGCTTCTTCATATCCCCCTATGTTGACAATTACATCAGCTTTAACCTAA  
 CTTTTCACTAAAGCATAAAACATCAAATACTTCTTGTATCTGAATAGGTAAATAGAATCAATTTTCATCTA  
 TTTTTTTGGAAGTATTTATCAGAATACGTATTATTCATACATATTTTAAATCATTAAATAAACATGATAT  
 TTAATTAGAAGATTTATACTTATTTTTGTAAAAATAAGAGAAAATAAATGAGTATACAAAGAAGTATTTT  
 GATTTGTTATTTGTCTTCATTCCTTGAGCAAGGTTTTAGCCTATGAAAATTTCCCATGAGTTAATTCCTG  
 CTGCTTCTTTAAACTTAAATGAGGTGACAATCGAGGT

*a112293*

GCTAAACACGGTGTAGATAGACTCAAGCTGACAAGAATTGTGGTGACAATTAAGGTTTGGACATGGTAA  
 GACCCAGTATTTTCTATTTCTATGTAATAAATAAGTTATTTTATGACAGTAGATTACGGTAATAAGCATA  
 AGTAAATATTTTGGTACCAATATATACTTTATTTCTTAAAGATTTTGAATCAGAATAATTTTATTTAT  
 CAACTGAATAGATACTTATGGTAGTTATATCTCAATCTTTTTTCTAAGTAATAACATTGAGAAAATTTTA  
 CATAAAGATTACTTTTATGTAATTACACTACGTATTTTATATTGCAAAATTAATCAATAACTAGACTTT  
 TTTATACTGATATAGCCAATAACATGAGAAGAGAGTAATAGAGATTCATCACTCTGTAAATCAAACATGG  
 GACTCACCGTTGGTTTAGTAAATACTTATTCGACTTTAAATATTGGTGATGCGGCAATTTA

*alr2294*

TAAATTGCCGCATACCAATATTTAAAGTCGAATAAGTATTTACTAAACCAACGGTGAGTCCCATGTTTG  
 ATTAACAGAGTGATGAATCTCTATTACTCTCTTCTCATGTTATTGGCTATATCAGTATAAAAAAGTCTAG  
 TTATTGATTAATTTTGAATATAAAATACGTAGTGAATTACATAAAAAGTAATCTTGTATGTAAAATTTT  
 CTCAATGTTATTACTTAGAAAAAAGATTGAGATATAACTACCATAAGTATCTATTTCAGTTGATAAATAAA  
 ATTATTCTGATTCTAAAATCTTTAAGAGAATAAAGTATATATTGGTACCAAAATATTTACTTATGCTTAT

TACCGTAATCTACTGTCATAAAAATAACTTATTTATTACATAGAAATAGAAAATACTGGGTCTTACCATGT  
CCAAACCTTTAATTGTCACCACAATTCTTGTGACGCTTGAGTCTATCTACACCGTGTTTAGC

*alr2372*

AATTGTTCCCTCTTCAATAGTAATTGAATGAAACTTGTACTCAATTTTGTACTTATTAGATTCCATAAATA  
GCAGACATCATATAGTTTGACATAAGTTGTCTCCAAAATTATGAGTACTTAGAGATGGAAATCCCGGAGA  
CAAGCCTGTCAATAATAAAAAGGTAAAAGCCTGTAAATATTAGCAACAAAAATTGTGACCGAATTTTATCC  
TTACTCCCCATTTAACACGAGATTTAACAAGGCTACATAATCTCTCTCTGAGAGCCTGAGATATAAATT  
TTATGAGGATAGAGTATCTCATTGTAGTTTATTTGGAAATCAATAATTTAAGGAAATTTTATGTAAA  
ATCTGATTGATCTGTTTTAAATGCAGAGTAGATTAGGTAGGAATGTGTAGTTTACCAAGTCATGCTGCAA  
AAATGGGTAATATTCCACTAGAAACAGACCAGCAGATGCGTAGTTAATCAATCACTCAAAATAGAAGAATT  
AGCTGAGGCGTTGTTGAATTTCTCCATCTCAAGGGATTGGCGAACTACTTGGCGAAGATCCCCAATCT  
CCAGCCACTCATAAAAATAGCCGATATGTAGCTATGCGATCGCCTGACATAACACGACTCAACTACAAT  
ATATATAGATATTGTTGAGAAATTGTATAGTTAGGGTCTAGCTGTCATGGCTC

*a112770*

ATTGCTCTGTTAGGCGATCGCTGTACCCCTTGTGTGATAGCCTTGCTTCTAAAATTCTCCGACATCGTTT  
TTCCTTCTGTGCCTTCATCGCTTGTGTTGCGATTGTATTACCTGTATTTTCTCATGCAGATGATGTACTA  
ATTTGCATCCATCTTGAAAGAAAATTCATCTGCTTACCATCACCTATAGTCTAGATTGATGATGATTTGT  
CATCTAGTCAAATTCACATATATGTATATTTGTCAATAGATAAATCGTAAATTGTATAATAATTCAAATA  
ACAACATTACATTCGTGATTCAACAGCATTATAGATAAATCTAGTGCCTAGTAACGTCTTTGAGTCGGGA  
AAATTTAGATACAGAGATAAAAATACTTGTATCCGATAGCTGATTTACCCGAGAAAAATATGAAGATTTT  
ATGAAGAATATCTACTGACTCTATGTATACTTAAGCAGATACAGTTAATGTCATAAAAAGAATCTTCTACC  
ACCTGAATTTCTAATACATGAGTATCAATAAAACTCAGTCATTGTTGCCAGTACCCGCAGGTAATTTACAA  
GTTTCTGAGTT

*alr2832*

ACCAACTACGAGCCGAGCAAGACAAATATTTTGTCCCGAGAAAAAGTTCTGGTAACTGTTGGTTAAGGGAA  
CACCAAAAAATGAATTAGCCACAATTGATTGTTTGGTAGGGTGCCTCAGTATGAAGATTTCTGAGTATAG  
TTAGGTTCTATCGCACTGACGCACCCTACTGGATAGTCTATTACCAATTACCAATTACCCATCTCAAAAA  
CGCCCCCTTCCATCAAAAAGGGGGGCTGAAGCTATCTGACAACAATTAATTTTATAATTTAAATAGCAAAA  
TAATTTCAAAGATAGAATCAATTTTGTAGTTTATTTGTAATACTATTTTACTAGCGCGCATCACTTCT  
GATTTCTACAAAATTCGATAAATGTATCAAAAAATAGTTAAAAAATATCATCCAAAATACCAAAAATAATGA  
CTGATATCTGCTGAAATATTTAAGGGAAATACTTACAACTTAGACTGATATGAAAGTTGCTCTAGTCC  
ATGATTATTTGACCCAGCGCGGTGGGGCAGAACGGGTATTTGAATT

*alr2856*

TAATAAAAGCGTAAAACCTTAGCAAAAGTGGGATGTATGATGCAGTAGCTAACGTCTAGCAACATATATC  
AAGATTTTACGTTGGACTGTGTGACAACACACACGGCTTTTATTTGATCTGCTAGATAAAATAGATATTAC  
GCCAGACAAAAGAACAGTGAAAATACGTAGGCTATTTCTGGGGAAAAAGCTTGATAATCTTTTATTTAAATC  
TTTATTGATTATTTAAATCAGGTTGCATTCAATCAGTAGCAGTAATTAAGGAAAATAAATCAGCGAACAA  
TGATTACAGTTTACAGCTTTAATCAATGGAAAAACCTAGCATAATGCTTGAAAAAAGCCCGGACACTCG  
AAGCTTAATTTTGTAGAGTGAACAATGCCTACAATACCCCTAGTTTGTAGGTAAGTACTGATCTAGGTGGATTT  
AAAAAATTTCTAGGAATTTCTATAGGACTCATACTTGATTTTTGAAATATACGTAGGGGAGCCACTGCGTT  
GGTGAAGCAGTACGGTCTTGGGGTTTCCCAAGTGGAGTAAGTCTGAAAGGGTTTCCCGCTTGGAGCA  
AGTGGCGTTGGGCAATGCCACCCTACAGATAATGAGAATTTTTTAGAAATCAAATCGGATTCCTATATA  
ACTGAATAAATGTTGAATATTTCTTTATTTGGCTGCATTACTGTTCTACATCAATATAAATGAAATAAC  
TAATTTTTTAGTAGTGTCAATCTTTAGGGTATTATCTGAACTATGCCTTCCCAACAAGACGATCAAGACCT  
GATTAACCTCCAACAGTATTGGCTGATTGTAAAAAG

*alr2861*

CAGTAAAAAACTTTTCTTATACTTTGTTTCAACATATTTATGCACAAAATAAATGATTATGTATAAATGGG  
 CAGATATATATTTCTTTGCTTTAATTAAGTTCTATAATAACTAGATGAGAACCAAATTCGTATGATAAAT  
 ACTAAAGAAAATATTCATGATTTTTTTTATTAAAAGCAATATTATTGATATAAAAAAT

*alr3057*

CTGAGGGTTTACCAGCTCATGCAGGTAGTGTAGCGGTGCGTTTTGCAAGGGCTGAAAGATATGTAGGGGTG  
 TTGGGGTATAGGGGTTTAGGGGTATAGGGGTTCAAATATATTTAAAGAATGAAATGTACTAATCCCCAAT  
 CCCCAGTCCCCAATCCCCAATCCCCACACTGGCAACCCCAAACAACCTACTATATCTATTTATCTATTAA  
 TCGATGCTGAATACCAAGTTACTCGATTTAGAGGCAGCGATCGCATTGTCTCCCTTGCTTTCCCGGAATT  
 GAATTTAACCGCCGAGCAGATTTTTCAAGCCACTAGTCTCACATCATGAGAATAAAACGACTACTGATTG  
 AAGGTTCTGAAGATATTGGGTCGCAAAATAAATCTTAGGCGATCGCTGGGGCAGTCTTAATAGAAAAG  
 CAAAGTACAAGCAACTACAAAAGACAATAAGCCTGTTTCCGCACAAAATACAAAATTAGATGAAGCTATG  
 ATGAAGTTGATTTAAATTAATGAAATTAATTTCTATGTTGCGTAAAGTGCTTTACAAGTGTATATACTCTC  
 TACAATGTAAAGTTTTTGCAAAAGAGGGGAATGATACAGTGACTTTTTAAACCTAGAGATACTACCATATAC  
 CTGTGTGTTCTTATGTAAAGTAGTTATAAAATTTACAACATTTTTGCAGTATCTTCCCTATTGACTTAAT  
 TGCAAACTTGCGGGGTGATGATATCGCCTAGGGTGTGAGTAATGCGTAAACCAGTTCTAACCATCTTTTA  
 CCAGTTCAATCCCTGGCAAACCTTCTATTGGAGGAAT

*alr3061*

TCTTTTATTGGTTTACTTTGCTGGCGGTAGCCGGTTGGGATTGGAACGCCAAGCCCACAACGTAATACAA  
 TTCAAAATTCAAAATAAACATTTCTAAATTCATATACACAGATTAATAATAAGGAAGCATCATGGCTG  
 ATAATTTCTCAAACCTAAGAAAAAACTTTATACTATCCAAGTATTGCGGGGGTTAGCAG

*alr3068*

TTAAAAGTTGGATATTATTACAGTTTCACATGGTTTATATTCCTGAAAAATCTCACAATATTTAATGAGT  
 ATTTATTAAGGGAACATCAAAAAATAAATTTATCCGAAAATACTTATGCCTACTATATCTGTAATTTATTC  
 TGCCATAATGCTGAACGTACTATTTTAGAGACAATATC

*alr3069*

AAAAATGGATATTAACCTAAAATATATATCCTAAGAAAGAAAAATAATCCACTGATGATTAATAATCAAG  
 AAAAATATCTTCCTTTTGGAGGTATAAGGCTTTTTAGTATTTGGATAAAATGGCTTTGGCCGTGATGTTT  
 TAAAACCCATGCCGTGATTCAAGACGCGGACAGAAATTAATAATGATATCATCATCCATCCTCCTTTAGT  
 AATTTATGCCGACGGCAATATAGGCGACAGTAAGTCATCGCACAAACATTCACGCCAACTGCACATCCATG  
 AACGCAGCTAAAAAATGACACAACAAGTAATTAACCTCTACAAGCAAAAATCATCTATCACCATTAGAAAA  
 CCAAAGCG

*alr3073*

TGTCTGTTTTATGGAGTTATAGGTATATTTTTTACAATCGCACTCGCACTTTGAGCAAGGTTTAGAGAAA  
 TGCTTTAAAAGTAATTTTGGAGGTTATGTAGCATATTAGGTTTGTGGTGAGCTACTTTTTGAACAAGC  
 TCTTATGTGGTTTTTGTATTAGTTAACTTATCTTGCGGAGATTTAAAATAAATATGCCTAAAATATCTGTA  
 ATTATTCCCGCCTATAATGCTGAACGCACAATTTTAGAAACGATTAA

*alr3464*

ACTTCTTGAGCATTCCGGTTTACCTGGGTGCAGAAACCACTCGCAGCATGGGTCTAGCTTACTAAGGAAT  
 TCACTACCGATTCTTTTTTGGCTCACAATTAAGTGTCTCGATAGCAAGCCTCTTTAGCTTTGTAGGGTGG  
 GTAGTAATGGCGTTCCGCTCGCCGTAAGCCATCGCCCATCTACAACATCATGAGAAATTTATAAAAAATA

CCTAGTGGTTAAAGGAGCAGGTTGCTCCTTTTTTCTTTCGTAGTCAATAGTCCATAGTAATAACTGTTG  
 ACCACTGACCACTCACAACGAACAACCTGACAAATGACTATACTGGATTTAATTATCCTGAAAAAAGGGTG  
 GTGTGAGCAATGTGTGGCATCGTTGGGTATATCGGAACTCAAGCAGCAACCGAGATTTTGTGGCTGGGC  
 TAGAAAACTAGAGTATCGCGGTTACGATTCCGCAGGTATAGCTACGGTTTGGGAAGGAGATGTAAATTG  
 TGTCCGAGCCAAGGGCAAACCTCCATAACCTGCGTCTAAACTAGAACAACTGGCTACTCCATCTCAAATT  
 GGTATTGGTCATACACGTTGGGCAACTCATGGCAAACCAGAGGAGTATAACGCCCATCCCCATTTAGATA  
 CAGCTATGCGAATCGCTGTTGTGCAGAATGGGATTATTGAAAATTACCGCGAGTTACGCGACGAACTGAA

*a113509*

GATTGCAATAGTTTCCACGCTAGAGCGTGTTACGTCACGTCACCGGTTACCGACAACCTAAAACCTTCACTGATT  
 CCCTCTGCGTCTCTGGCGAAGCGAGAGTACTTTATGACTCTCGCGGATCAATTCTTCACCAATTCTGAC  
 CTTGATGCAAGTCCTATGATGGGGAGTGCTGATTTGTAGGGTGATGTAGTGCGTGTTAAGTAAAGAATCA  
 TTCATTCTTTACCCAATCCCCATTACCCCTTATCCCCAGAGGGGGCCCCACCTTCCCCAGTCCCATAATA  
 AATTTGCCAAATTTAGAATAATTGTTCTAAAAATAACTAGGAAATAATGACGGAACATCATTTGTCACTAT  
 TAACCTGTGGATATGTAAGATTCGTACTTTGATTTTGAACATGGGTGGGTGGGCAATACCCACCAAAA  
 TCATGATACGGTGGGCATTGCCACCCTACAGATACTTAGATTTGTTCAAGAATCAAATCGGATTCCTAT  
 AGAAACTAAGAAAATAAGAATAAAGTAGGTAATAATGGCGAAAATTATAGTTACTGGAGCAGCAGGATTTA  
 TTGCTTCTCACCTGGTAGAAACCCTACT

*alr3756*

TCACTGTTTGGGCAGGAGTTAAGGTTCAAAGCGGAGGTGTTGGAGCGATCGCACCTGAGCAATAGTTTCA  
 AAGTCACGATCACTGTGAACTAAAAGTAAATAAAGCATTCTACCCAGAATATATCAGTATATTAACCTGT  
 ATTTTGTACCTAAGTACAAGAGATTCAGTGTATTCCTCGCTGATATAGCAAAGATTTCCGCCAATATCT  
 AAATTCATGAGTCCCCGACTTCATGTTATTCGCTAATGCGATTCAATTCATTTCTACCGCAGTGCTATCA  
 ATGTACTGATTTTTAACTCTCCGTATATGCTATACAAGCAACGCCAACAAAAGCTGTGGAAATGGATGGC  
 ATTGTTCTAGTTGGGACATT

*a114219*

AAATTTATAACTGCGATCGGTATTGGTGGTGAAGAAGCCCAAAATATTCCTGATTGGGACTAGGGTTG  
 AACTAAGCAAGCTGGATTTAAAAGCACAAAGTGGGGTGGGTAAAGTTGAATCGGGTGTATCACCACCC  
 CTTGCAAAAATAATGAGACTAACATTAACACGCAGTTGCGTTTAATAATGTTACTTGACAAAACATATAC  
 AGTTGATCATGAGAAGCAGTGGTAATATCACTGCTTCTTTGTATTTCAGCGCCCACTACTGATATATTTTG  
 GGCATAACGTAGAGAAATTTATTTAAAACCTTTGAGATGCCAAAAAATGTAAAGTAATATTTACAACGAAT  
 TTGATACCTGACTAAAAAAAAGTGACATCTCAAAAAGCCTTAATCTGGGGCAAAAAGTTTAAAACCTCAC  
 AACTGCAAAATCCTGTT

*alr4239*

CTGACAATTTCTTCACCAAACCCCAAGAAGAGCGCACCCGCAAATTTTTATCACAGATTTCTCTAGTGCAA  
 ACAATTACCTCGATGCAAAAATTCATATCTGGAAATCTTACTCTGGTGTATGTGCGTGATAAACTGCCG  
 TAACTTGGAGATAACAGAGTATGATATCTACTCTCCTAATTGCCAAAAGGGTTAAATATAAACTCTTT  
 CCAGTGGGAGATGCCACATTTAAAATGTTATGTCAGTATTTAAAATAGTATCTCAGGAACGGCAAATT  
 TACTAAATCTTTAAAATTAGTTGCTTAGTTTCAACGATGCACAACCAGTTTAAATAAGGAAAAATTTTTA  
 ACCCTCAGCCGACAGGGAGACTAAGCGAAAATCGCTATCTGGTTTTTTTCGGGTGTAAGTACTGTTCAATT  
 ACATCCACAAAACGGAATGCCCTCAGCAATTTACCATAATTTTCGGGTATTATTTGCATATAAGTTGTAGT  
 TGATACAGTAATATTTACATTGATAACAATAAGGTTTATGTATATAGTAAATAAATTATACATAGTGCTT  
 ATTTAAGAAAAAGCACAAAATCAATCACATCTTACTGACGAAAGCAATAACAATTTTTACCGAATCTTCA  
 AAGAAAAACCAGTGACATAGAAATACTATAAAAAGGGAAAATTAAGTATTAGATGCTTTCAAACATAAGCA  
 ATCTAGACCTTGCTTTAGTTATCCAAACCAGCCAGGGGACTTCTGACTAATTTGTTATGCGGATTTTAAT  
 TTACTCATATAACTACTATCCAGAACCAATTGGTATAGCTCCCTAATGAC



*a114388*

TTATCAACCAGATTCAAAAACAAGACACATCTTTGCATGAAGTAATTGATCAATCAGAAGTATAGCCGTA  
 GCTAGACAGACTAGGACATCAAGCTAAGATAAAACCTGGACAACACAATGATTTCAAGTCTGTCAACAGC  
 TATATCTCAGCAAAAATACTAATGTGAAAATTATATTCACAGATGTTGTTTAGTAAAATCAGGCAACTTT  
 TACTGTAACCTTTACTAAAAATTCAACTTGGGAAAAATATTTGACCGGATTAAGATAGTTTTTCTAGGA  
 AAATGAGAATTGTCGGGAACTTCAATACTTAAGTATAGTCAAGATTAATAATGTTTAGACAATACTTTAC  
 CGATGGTAGTCTCAGAGTTTAAAAAATATGCTTAACACAGGTTGCTGAAATTTCTATCACAGCCAGCTG  
 TGGGTGCAGCTTTGCTAACTGC

*a114426*

TAAGACCAAAGTCTCAACGTCCCTATCTAGAAATGCGTTCATCAGAATACTTAGATATCAAATAGTTACT  
 GAAATATTTACAAAACCTAATGAAAAAATATTGCAATGAATGTTTTACTAATTAATAAATCTGATATTG  
 ATGGTGGTGCAGCTCGTGCAGCCTATAGATT

*a114432*

CAACTAAAATAACGCACCACAGTTTAGGTTTGACAACCGATTTCTGCCGGATAAATAAGGGATAGAAAA  
 AGAGTCGATTTGGGAACAAGTCATGTTTCAATACGTCTGCAATCGGTTAGTCGATTAGTGAAGAGAAGAC  
 AAAAAAGCGATCGCTATAACTTGATTTGATACGTTGGCTAATTAACAGGCTTGTTACGTAAATCCTAT  
 GAGATAGCCAAAGTATAAAAACTTTTCTGACCTACAAGTGTGGGAAAAATAATTCTCTCGTTAAATA  
 ATTTTGCTTTTTCTATATATAAACTTTTAAACAATACTTTGCTGTCTCAGGCTGAAGCCCTGGGAATGGA  
 TACAAACAGGGAAGACAATGCTTCATGAGATTACTAGTAGAACACAAAATCTATGAATTTTCCGTAAAC  
 TTGTTTCATTAAGTGTTTTTACATAAACATGGTATGATATGACTGATTAAGCATATATATCAGGGAATAT  
 CTTTGGTGAATAACAACCTTACTTATCAATCAACAACTCACTAGTATTCCTTAGTTATATTTTAGGAAAA  
 AATTGCCGATCATCGGTAATTTAGCACCAATTAGGTGTTCACTTACAACCTAAAAAATCTTTGAGTGCTTT  
 GGAGTTTAAAGTGTGATAGTCTTTTAGCTCAGTTTTTTTAAAGTCACTTTGTCTTTAGGTCTTTAAACAA  
 AATAAAGAGCATGAGTTGTATGTATTGAAGGTGCTTTCATCAATCAAAGGATTCAAAATCGTGTACTGTA  
 TTCTGTCTGTGATGGGTTAAACCCAAATAGTACACTTTATAAAAAAGTTCGTTAGTTGATATCAACTGA  
 TGAGTAAATCCACCAAAAATTTGTTTCTGATAGTCAACAGAACACATGAATATTGTTAAGTAGAATACA  
 GTGTTAATGGTGCATAAGCACTGGTATAATCTGCTTTCGACAGTCTTTTTTATAAAAGTTTGTATATATG  
 ACAATGGACTCTTTACAAGGTTTTGAAGAAGTAGATATTCACAAAATACGTAGAAGTTCTTCA

*a1r4485*

AAGAACATCAACGGGCTGAACAAGAACGTCAACGGGCTGATAGACTGGCAGAGTATCTGCGTTGACGCGA  
 AGCGGCTTGTCTTTAGACATCGCAAGGGATTGACCCAAACGACCTACCCACGTAAAAAGTAGGGCTGATA  
 GATTGATAGAATAGTTGCGTAGAGATCACTTTTCCAGAGAAACAGGCTAGGCATAGAATTTGCCAATGCC  
 AACTGTTTGTCTTTAGACATTGCCAAAATATATCTAAGAACTTCAGTTAAGCAACCACATATATAGAT  
 GATCCACATATATAGATGATATGGATGTACGAGATGTATATATTTCTCGTATAATTTTGTGGTATTTA  
 ACTCTTAATCAAATTTAAGATTAGTAGTTGCTGACAGAAACTTTAGGAACTGACTAACAAATAAAGT  
 GTCGTCAGGATCTAAATATATCTTGGTTGTTTAGTCTGCGAGCAAGCTGATAAATCTTGTGATATTTCAA  
 AAAAGACTGAGAAATATGAACTTTCACTACTAACCCAATGCAAAAAAATGCACTCTGATAGTATAGAATG  
 CGCGAATAACTCGCTTGTGGTCTGTGGTGGTGTTTGAAGATGGCTCTTGTAATTAATAAATAATTCAA  
 ACTGGCTGCAAGTGCAGCGCTACTGGGAAGTCTGCG

*a115073*

GACATCGTTATTTAGTGAACCTACCATTACAATTTGGCAAAAAATCCGGAAGTTAGCGATATTTGATGAAA  
 ATTTCTATTTCTACTGCATTTTGATGACCGAATGCTTAACGTCAATCAGCATGGATAGGAATTAATAAT  
 TTGCAATTTGAGTTATTACCTTTTTATAATGAGCCGATTATGATTAAGTCCAAGTTCACTGGGGCTTG  
 GTAATCAAGCGCAGCTTATTTTATGAATTTAGTCTGGCAACGATTTACTTTATCTTCTTTACCTCTAAAA  
 CAGTTTCTAGCTACAAG

*al15222/alr5223*

TCAGCTTCTATAGGAGTAGAAATGGATTGTATAAAATTTGTTATTAATCCATTCGTTTAGTCTCATCGGCT  
 GTGTCTATTTGGGAAGTTATAAACTTTACTGATGACACTACCGAACTGCTCAAGCCATTAATTGCAAGC  
 ATAAGAATTTAATCTGGTTTTATGACATGGAAACATCTTTGGTTTTATACTTTTCCAAGTCTATAATGTGT  
 TGATTATATACAATCTAAACAATCCAAACTTCACACATATATCCCATCAAACAGTCCATGCTACATAACC  
 AATAATCATGAATAAAATGCTTTTTCTGCACAGGGGGAACACAAAATAAAAATTACAAAAGCATTACTAAT  
 CATTCATATTCGCTACATGAAGAAGGTATTTCACTTAGTAGTTAACTATTTAGTATTGTTAAATAACTAT  
 GTCGCTAACTTCTCAGTCAATAAGGTATATAACAGGCACTATGTTTCTAGTAGAATTGACGAAAATATAG  
 GATCATACAATGTACCGATTTTGGTTGTTATGAACCTGTTTGTATCTGCTCCAACTTTAATTTTTACTGT  
 ATTGGGGGATCTTCAGATCAATTATTTAATCTCCTTCGCGTTTTCTTGCTTAATCTAAGTGTAAATCACGA  
 GGTATACGATACCATGTTTTAAAGATGCCTGGCATCCTTCATAAAAAGTTCATCTTACTAGGATGTTTTT  
 TAGGTGTTTCTTCATAAAAGGTATATTTTATTAAGTTGTATTAACCTTACATAGTAATCAAGTGTAGTTA  
 AATCTTCTTAAGCGTAATCGCCTTAATTGTTTGGTAAACAACAAAATAATCACCTCTTAATCCCAGTTTT  
 CAAGGTGCAACTACCTTAAATAGACCAATGATGTTAACCGAATACTTATCTCTCAATGACTGATTGATT  
 TACAATTGAAGTGTATCCAAATTAACATTCTACACGGATGATATATACTTAGAATCATTAAACTTTATA  
 ATACGGAGTTGATATCTCAGTTAGGATATTTCCAGTGAAGTAGTTTATTTAGATTTCAGGATATCAACGT  
 AAATAATTAGTGCAACTTAACACAGAAAATTTGTGCAGTCCGTGATTTAATGATTGGCCAAAAGTAAATA  
 TTCATATATATAACTTTACACTCGTGTAGGAAAGTAGAAAACATTTATACTACATCTTAGCCATTTTATT  
 TCCAGATAGTTAGAACACAATGAATAAGCAATCTTACCAAACTTTAAACTCTGAACAAAATGGCAACTCT  
 AAATCTTTCTTGCC

*alr7196*

CTCCTTCCTCGTTAGCCCAGGTGAGGGTGAAAAATGGATATTTAACTAGCAAGAAGTTATATTAACCTCT  
 TGGCTGGCAATTGAGTAAAACTCTTGTTTAAAGATGTTGCTGATGTTAGACAGAATCTCTGTCTAACATC  
 CCCAAATTGTTTGCCTGTAAGGGTTATGCTGCATAAAGGTATTGATAAACTCACCAGTGAAATCGCCCTA  
 CAGCATGGCGGAGCATTGTCGCTAACTCTATCAGTTCAATGAATTTTCATAGCGGTAGATGATTTATATA  
 GCTAATTGCTGTATTATCTGCCAGTGTGAAAATAACAAGATTAATCTCCTAACTTATGGAGATTTAA  
 TATCTATATTTAATCTAATAATCACAATCCTTTTACCACACAGTATGTTAACCAGGCGAAGAGCCGGAC  
 TAAGATTAGATGCCATCGACTTGCGAAAATATAGCCACCCT