



Universidad
Zaragoza

Trabajo Fin de Grado

Validación del análisis del transcriptoma de biofilms de *Anabaena* sp. PCC7120 e identificación de nuevas dianas moduladas por proteínas FUR (*ferric uptake regulator*)

Validation of transcriptomic analysis of *Anabaena* sp. PCC7120 and identification of novel targets modulated by FUR proteins (*ferric uptake regulator*)

Autor/es

Alodia Liesa Delgado

Director/es

Irene Oliván-Muro y María F. Fillat Castejón

Departamento de Bioquímica y Biología
Molecular/Facultad de Ciencias

2023



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

María Francisca Fillat Castejón e Irene Oliván Muro

CERTIFICAN

Que la memoria titulada “Validación del análisis del transcriptoma de biofilms de *Anabaena* sp. PCC7120 e identificación de nuevas dianas moduladas por proteínas FUR (*ferric uptake regulator*)” presentada por Alodia Liesa Delgado ha sido realizada bajo su dirección en el Departamento de Bioquímica y Biología Molecular y Celular de la Facultad de Ciencias.

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

Zaragoza a 13 de junio de 2023

María Francisca Fillat Castejón

Irene Oliván-Muro

ÍNDICE

1. Resumen.....	1
1. Abstract.....	1
2. Introducción	2
2.1. Cianobacterias.....	2
2.2. <i>Anabaena</i> sp. PCC7120	3
2.3. Papel de los reguladores FUR (<i>ferric uptake regulator</i>) en <i>Anabaena</i> sp. PCC 7120.....	3
2.4. Características de los biofilms bacterianos	4
2.5. Biofilms de cianobacterias	4
2.6. Biofilms de <i>Anabaena</i> sp. PCC7120.....	5
2.7. Potenciales aplicaciones de los biofilms de <i>Anabaena</i> sp. PCC7120	5
3. Objetivos del trabajo.....	6
4. Materiales y métodos	6
4.1. Cultivo y valoración de la formación de biofilms.....	6
4.1.1. Preparación e inoculación de cultivos para la obtención de biofilms	6
4.1.2. Cuantificación de biofilms.....	6
4.2. Técnicas de trabajo con DNA	7
4.2.1. Extracción de DNA genómico	7
4.2.2. Reacción en cadena de la polimerasa (PCR).....	7
4.2.3. Electroforesis de DNA en geles de agarosa	7
4.2.4. Purificación y cuantificación de los productos de PCR	8
4.2.5. Ensayos de retardo en gel (EMSA).....	8
4.2.6. <i>Real time</i> RT-PCR	8
4.3. Técnicas bioinformáticas	9
4.3.1. Búsqueda de secuencias	9
4.3.2. Búsqueda de operones en los genes del transcriptoma	9
4.3.3. Búsqueda <i>in silico</i> de potenciales cajas de unión de las distintas proteínas FUR.....	9
4.3.4. Manejo de datos <i>ómicos</i> en hojas de cálculo	10
5. Resultados y discusión.....	10
5.1. Validación de los resultados de transcriptómica comparada de biofilms de <i>Anabaena</i> sp. PCC7120 frente a células planctónicas.	10
5.2. Estudio de la correlación entre la expresión génica y respuesta a diversos estreses.....	11
5.2.1. Estudio de la correlación entre la expresión génica en biofilms y en respuesta a la desecación.....	12
5.2.2. Estudio de la correlación entre la expresión génica en biofilms y en respuesta a la deficiencia de nitrógeno combinado	12
5.3. Estudio de la correlación entre la expresión génica en biofilms y ante la desregulación de parálogos FUR.....	13

5.4. Búsqueda de potenciales secuencias de unión a parálogos FUR en genes cuya expresión varía en el biofilm.....	14
5.5. Selección de genes candidatos a estudio por EMSA	14
5.6. Resultados de los ensayos de cambio en la movilidad electroforética (EMSA).....	17
5.7. Capacidad de formación de biofilms en cepa WT (silvestre) y VCS2770 (cepa de sobreexpresión de FurB) en ausencia y presencia de nitrógeno combinado	21
5.8. Consideraciones finales	23
6. Conclusiones	23
6. Conclusions.....	23
7. Bibliografía	24
Anexos	26
Anexo.1: Cultivo de <i>Anabaena</i> sp. PCC7120 y medios de cultivo	26
Anexo.2: <i>real time</i> RT-PCR	27
Anexo 3: Matrices de peso para la búsqueda de potenciales cajas FUR.....	28
Anexo 4: Preparación de EMSA.....	30
Anexo 5: Negativos para EMSA.....	33

Abreviaturas

cDNA- (DNA complementario al mRNA)

Ct- Umbral de ciclos (*Cycle threshold*)

DEG- Gen diferencialmente expresado (*Differentially expressed gene*)

DNA– Ácido desoxirribonucleico

EMSA- Ensayo de cambio de la movilidad electroforética (*Electrophoretic mobility shift assay*)

EDTA - ácido etilendiaminotetraacético

EPS- Sustancia extrapolimérica

FUR- *Ferric uptake regulator*

Gly– Glicina

H₂O MQ- Agua Mili-Q

Lys– Lisina

min- minutos

mL– mililitros

Nº - Número

pb– Pares de bases

PSA- Persulfato de amonio

PSI– Fotosistema I

PSII- Fotosistema II

p/v– Peso/volumen

RNA– Ácido ribonucleico

TBE– Tris, borato, EDTA

TEMED- tetrametiletilenediamina

s- Segundos

µL - microlitros

WT- Wild-Type (silvestre)

1. RESUMEN

Anabaena sp. PCC7120 es un microorganismo ubicuo que lleva a cabo la fotosíntesis oxigénica. Se ha demostrado su capacidad de hacer frente a diversos estreses como la desecación, debido a la síntesis polisacáridos extracelulares, y ante el déficit de nitrógeno combinado, gracias a la formación de heterocistos. Parte de la regulación de la respuesta a estos estreses parece estar mediada por los parálogos de la familia FUR (*ferric uptake regulator*), siendo la formación de biofilms un mecanismo habitual para hacer frente a diversos estreses. Los biofilms son agrupaciones de microorganismos embebidos en una matriz extracelular que se inducen ante condiciones adversas. Como casi todos los microorganismos, *Anabaena* es capaz de formar biofilms. Sin embargo, las condiciones y mecanismos moleculares que inducen su formación son todavía desconocidos.

Para conocerlos mejor, se partió del transcriptoma comparativo de un biofilm de esta cianobacteria frente a cultivos planctónicos y en un primer paso se realizó una validación del mismo. Tras ello, se estudiaron los solapamientos en la respuesta transcripcional en la formación de biofilm y ante dos estreses (desecación y deficiencia de nitrógeno combinado), así como con los transcriptomas de desregulación de los parálogos FUR, obteniendo solapamiento génico considerable en todos los casos. Dada la posible implicación de proteínas FUR en la formación de biofilms, se estudió la regulación por estos reguladores de genes potencialmente relacionados con este proceso mediante ensayos de cambio de la movilidad electroforética (EMSA), observándose unión directa en 10 nuevas dianas regulatorias.

Por último, se evaluó la capacidad de formar biofilms por *Anabaena* en dos condiciones, falta de nitrógeno combinado y sobreexpresión de FurB. Los resultados muestran que la falta de nitrógeno combinado afecta significativamente la capacidad de la cianobacteria para formar biofilms.

1. ABSTRACT

Anabaena sp. PCC7120 is a ubiquitous microorganism that carries out oxygenic photosynthesis. It has been demonstrated that this organism is capable of coping with various stresses such as desiccation, due to the synthesis of extracellular polysaccharides, and nitrogen deprivation through heterocyst formation. Part of the regulation of the response to these stresses appears to be mediated by paralogs of the FUR (*ferric uptake regulator*) family, being biofilm formation a common mechanism to cope with various stresses. Biofilms are aggregations of microorganisms embedded in an extracellular matrix that are induced under adverse conditions. Like most microorganisms, *Anabaena* has the ability to form biofilms; however, the conditions and molecular mechanisms that induce their formation are still unknown.

To better understand these mechanisms, the transcriptome of the biofilm compared to planktonic cultures was analyzed, and a validation of the biofilm transcriptome was performed as a first step. Then, overlapping responses were studied in biofilm formation and two stresses (desiccation and combined nitrogen deficiency), as well as the transcriptomes of the deregulation of FUR paralogs, showing considerable overlap in all cases. Given a potential regulation between the FUR paralogs and biofilm formation, the regulation of genes potentially related to the process was studied using electrophoretic mobility shift assays (EMSA), revealing direct binding to 10 new regulatory targets.

Finally, the ability of *Anabaena* to form biofilms was evaluated under two conditions: nitrogen deprivation and overexpression of FurB. It was found that nitrogen deprivation significantly affects the capacity to form biofilm.

2. INTRODUCCIÓN

2.1. Cianobacterias

Las cianobacterias son bacterias gram negativas autótrofas, capaces de llevar a cabo la fotosíntesis oxigénica. Estas bacterias son una parte fundamental de muchos ecosistemas, ya que se encargan de la fijación de alrededor del 25% del carbono de la Tierra y algunas cepas son capaces de fijar nitrógeno atmosférico cuando no existe disponibilidad de nitrógeno combinado (1). La fijación de nitrógeno pueden realizarla como bacterias de vida libre o en forma simbiótica, con helechos del género *Azolla* por ejemplo. Por ello, y por su gran capacidad de adaptación a diferentes condiciones ambientales, han colonizado la mayoría de los ambientes acuáticos y terrestres, incluso el desierto (2,3). Además, se considera que las cianobacterias son el ancestro evolutivo de los cloroplastos (4).

Este grupo de microorganismos presenta una gran diversidad morfológica, desde formas unicelulares como las del género *Synechocystis* hasta filamentosas como *Anabaena* spp. o *Spirulina* spp. Sin embargo, sus características metabólicas, fisiológicas y estructurales son similares entre todas ellas.

Las cianobacterias poseen una pared celular que les permite controlar el transporte de sustancias, así como resistir a la presión de turgencia. Esta pared está compuesta por dos membranas: plasmática y externa, y entre ellas se encuentra una capa gruesa de peptidoglicano. Además, en algunos tipos se encuentra una capa mucilaginosa rodeando a la membrana externa, compuesta por fibras de celulosa y proteínas, así como apéndices externos como pueden ser pilis o fimbrias (5).

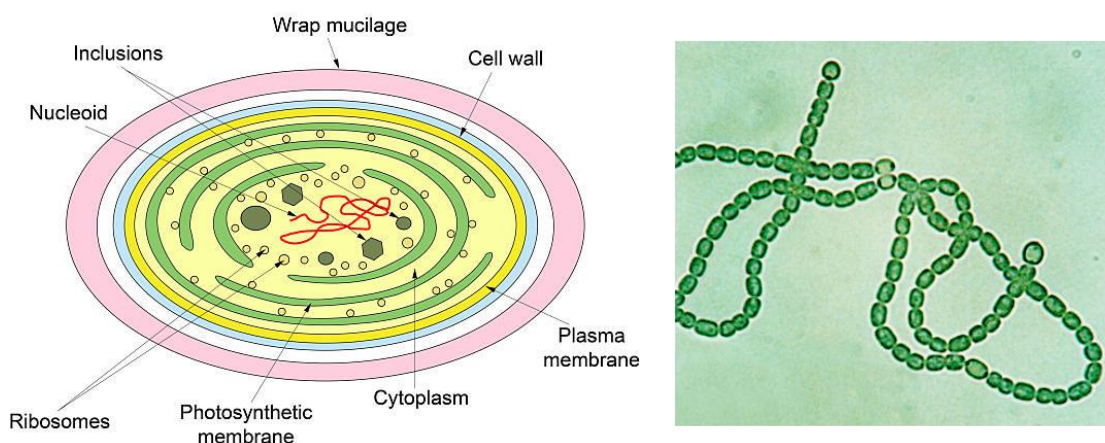


Figura 1: A Esquema de cianobacteria donde se representan los diferentes componentes celulares. Obtenido de Luiz Botter. (<https://es.123rf.com/>) B: Morfología de *Anabaena* sp. PCC7120. Obtenido de (6)

Su aparato fotosintético está compuesto por dos fotosistemas (PSI y PSII), con una estructura parecida a la de plantas superiores, que les permiten absorber energía lumínica y transformarla en energía química. La mayor diferencia en este proceso es la existencia de ficobilisomas en cianobacterias, que son complejos de pigmentos que contienen ficocianina y ficoeritrina, que actúan como antenas del PSII y absorben distintas longitudes de onda que la clorofila a, lo que les permite sobrevivir ante bajas intensidades de luz (5).

Las cianobacterias producen numerosos metabolitos secundarios que pueden usarse como cosméticos, medicamentos y vitaminas (7). Una de sus aplicaciones más prometedoras es el uso como biofertilizantes ya que pueden mejorar las características fisicoquímicas de los suelos y la cantidad de nutrientes minerales que se encuentran en suelos áridos con poca disponibilidad de nutrientes (8).

Sin embargo, estas bacterias también pueden tener efectos negativos sobre el ambiente, ya que pueden producir floraciones o *blooms*, que consisten en el crecimiento desmesurado de cianobacterias, siendo un problema para el ecosistema donde se da. Además, la producción de *blooms* a menudo lleva asociada la producción de cianotoxinas, que pueden causar daño tanto a humanos como mamíferos y otros organismos. Las más frecuentes son las microcistinas provenientes de *Microcystis aeruginosa* (9).

2.2. *Anabaena* sp. PCC7120

Anabaena sp. PCC7120, también denominada *Nostoc* sp. PCC7120, es una cianobacteria filamentosa fijadora de nitrógeno, siendo el modelo de estudio para este tipo de cianobacterias. Además, es una cianobacteria multicelular que, en condiciones de carencia de nitrógeno combinado, es capaz de diferenciar células especializadas denominadas heterocistos, donde se fija el nitrógeno atmosférico vía nitrogenasa, que es muy sensible al O₂. Para proteger a la nitrogenasa del O₂, durante la diferenciación a heterocistos se da un importante cambio en el patrón de expresión génica de las células vegetativas (10): se forman dos envolturas celulares gruesas, una interna formada por glicolípidos y otra externa por polisacáridos, impidiendo la difusión de oxígeno al interior de estas células; se desmantela el fotosistema II para evitar la producción de O₂ durante la fotosíntesis; y no se produce la fijación de CO₂ en las células, por lo que las reservas energéticas provienen de las células vegetativas vecinas y pasan a través de los septos, que son conexiones célula-célula (6).

2.3. Papel de los reguladores FUR (*ferric uptake regulator*) en *Anabaena* sp. PCC 7120

Anabaena sp. PCC7120 tiene altos requerimientos de metales como pueden ser hierro, zinc o manganeso. El hierro es el metal más abundante en las proteínas del PSI. El zinc y manganeso son elementos indispensables para el mantenimiento del PSII, mientras que la enzima nitrogenasa también tiene un alto contenido de metales, concretamente hierro, molibdeno y cobalto (11).

Paradójicamente, aunque el hierro es necesario para procesos vitales en la célula, entre los que se encuentra la fotosíntesis, el exceso de este nutriente genera especies reactivas de oxígeno derivadas de la reacción de Mehler, que pueden comprometer la viabilidad (12). Por tanto, es necesario controlar estrictamente la incorporación de metales para minimizar la producción de especies reactivas de oxígeno. Para ello, existen las proteínas FUR (*ferric uptake regulator*), que actúan como reguladores transcripcionales uniéndose a las denominadas como cajas FUR, secuencias ricas en A y T que se encuentran en la región promotora de los genes diana. Estas proteínas se caracterizan por poseer dos dominios: uno rico en histidinas en C-terminal y un dominio de unión al DNA en N-terminal (13). Para llevar a cabo su función, las proteínas FUR forman dímeros en presencia de un metal correpresor y, tras ello, se unen a la caja FUR inhibiendo la expresión de sus genes diana. Sin embargo, más recientemente, se ha conocido que esas proteínas también pueden actuar como activadores. Además, una vez deja de estar presente la molécula correpresora, el dímero se disocia del DNA permitiendo la transcripción (13).

Existen diferentes parálogos de las proteínas FUR, como pueden ser las Zur (reguladores de la homeostasis de zinc), Mur (relacionada con el manganeso) o PerR (relacionadas con la respuesta ante el estrés por peróxido). Concretamente, *Anabaena* sp. PCC 7120 posee tres parálogos FUR: FurA, FurB y FurC, cada uno implicado en diferentes procesos (14). FurA es una proteína esencial que está relacionada con la regulación de la homeostasis del hierro y la protección frente al estrés oxidativo (15). Más recientemente se ha podido conocer su papel en la regulación de la fotosíntesis, el metabolismo del nitrógeno y otros procesos, por lo que se puede hablar de un regulador global (15). FurB es una proteína de tipo Zur, ya que está implicada en la homeostasis de zinc, que también participa en la respuesta frente al estrés oxidativo (16). FurC es una proteína de tipo PerR, que controla la respuesta al estrés por peróxido, también implicada en la fotosíntesis y metabolismo del nitrógeno (17).

2.4. Características de los biofilms bacterianos

Los biofilms son agregados microbianos que se encuentran adheridos o asociados a una superficie biótica o abiótica y embebidos en una matriz. Las células que componen un biofilm y las células planctónicas presentan características distintas, siendo las más significativas las diferencias en el nivel de resistencia a los antibióticos, cambios en la actividad metabólica y la existencia de una matriz extracelular en el biofilm (18). Esta matriz está formada por agua (97%), células microbianas (2-5%), sustancias poliméricas extracelulares tanto homo- como heteropolisacáridos (1-2%), proteínas (<1-2%, incluyendo enzimas), DNA y RNA (<1-2%) y por último iones (unidos o libres) (19).

En un biofilm, varias especies bacterianas pueden formar consorcios estables y experimentan un fenómeno llamado percepción del cuórum, donde las bacterias regulan la expresión de sus genes en función de la densidad celular (20). Para comunicarse, las bacterias producen y liberan señales químicas conocidas como autoinductores; al detectar una mínima cantidad de estas, se produce una alteración de la expresión genética como respuesta (21). En *Anabaena* sp. PCC7120, la percepción del cuórum se activa por la presencia de un autoinductor llamado N-acilhomoserina lactona (22).

Las etapas básicas para la formación de los biofilms son las siguientes: agregación y adhesión, en la que las bacterias se agregan unas con otras y se adhieren a las superficies bióticas o abióticas; crecimiento y acumulación, durante la cual las colonias bacterianas se expanden por crecimiento y reclutan a otras células; y, por último, cuando se dan las condiciones adecuadas, se produce la desagregación (18).

2.5. Biofilms de cianobacterias

Las cianobacterias son grandes productoras de sustancias extracelulares, cuyos componentes mayoritarios son exopolisacáridos (EPS), moléculas directamente implicadas en la formación de biofilms, por constituir parte importante de su matriz (23). Los EPS procedentes de cianobacterias son heteropolisacáridos de alto peso molecular compuestos por diferentes tipos de residuos de monosacáridos, normalmente 6 o más tipos, y son sintetizados por acción combinada de diferentes glicosiltransferasas. Los monosacáridos más comunes en EPS de cianobacterias son glucosa, xilosa, arabinosa, galactosa y fucosa (24), y también contienen otras moléculas como ácido urónico y sulfatos, generando un carácter aniónico. Además, su síntesis se ve afectada por parámetros nutricionales y ambientales, así como específicos de especie (23).

Los EPS pueden estar asociados a otras moléculas como pueden ser péptidos, DNA y lípidos, que también juegan un papel en la formación y desarrollo del biofilm (23). Los lípidos contribuyen a la resistencia a altas tensiones superficiales del agua que se encuentra en los alrededores del biofilm, facilitando el crecimiento en superficies. Los péptidos son importantes gracias a sus propiedades cohesivas, permitiendo que el biofilm sea estable durante un largo periodo de tiempo (25).

Los biofilms de cianobacterias proporcionan una ventaja adaptativa para la supervivencia de estas bacterias ante condiciones adversas, como puede ser la desecación. Los EPS que forman los biofilms contienen ácido urónico, que es hidrofílico, lo que promueve una alta absorción y retención de agua por parte del biofilm, permitiendo que las bacterias puedan desarrollarse correctamente (26). Por otra parte, los biofilms contienen EPS con grupos sulfatos, lo que contribuye a la adsorción de iones en ambientes salinos, favoreciendo la supervivencia. Además, estos grupos sulfatados, gracias a su carácter aniónico, pueden quelar metales pesados, así como contaminantes emergentes (27).

2.6. Biofilms de *Anabaena* sp. PCC7120

En la literatura apenas existe información disponible sobre la formación de biofilms por parte de *Anabaena* sp. PCC7120, así como la formación de sus EPS. Recientemente, en el grupo de investigación de Regulación genética y fisiología de las cianobacterias (*Cyanofur*) se han obtenido los resultados de transcriptómica comparada de biofilm formado por cultivos axénicos de esta cianobacteria frente a cultivos planctónicos. Este es el primer trabajo de esta índole realizado, lo que supone una herramienta útil y muy necesaria para revelar los diferentes mecanismos y rutas metabólicas implicadas en la formación y mantenimiento de los biofilms en cianobacterias filamentosas fijadoras de nitrógeno.

Una primera aproximación del análisis del transcriptoma fue la clasificación en diferentes categorías funcionales como se muestra en la Figura 2.

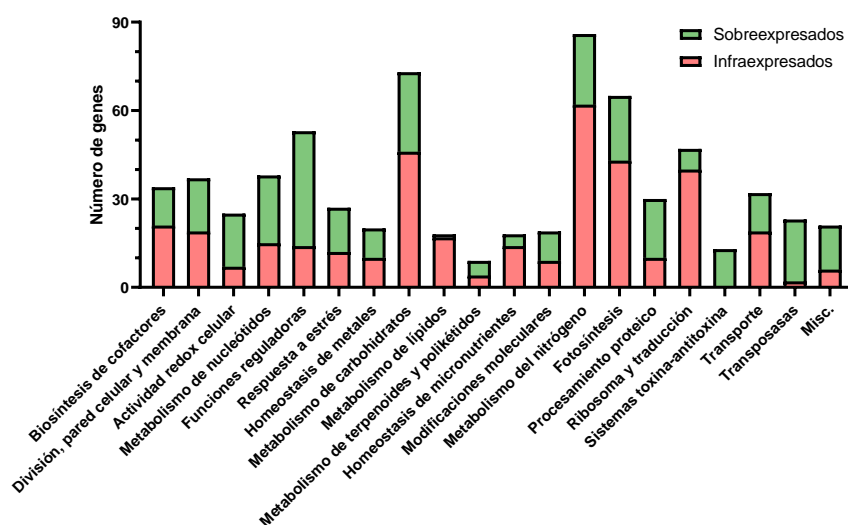


Figura 2: Representación de las categorías funcionales de los genes desregulados ante la formación del biofilm. Obtenido de tesis doctoral de Irene Oliván Muro.

En esta gráfica (*Figura 2*) se muestra la cantidad de genes que varía en el biofilm con respecto al de las cianobacterias en la fase líquida de un mismo cultivo según su categoría funcional, y cómo es ese cambio: sobreexpresión o infraexpresión. Un gran número de genes desregulados tienen relación con el metabolismo de los carbohidratos, metabolismo del nitrógeno, fotosíntesis y ribosoma y traducción, siendo estos procesos vitales para la supervivencia de *Anabaena* sp. PCC7120.

2.7. Potenciales aplicaciones de los biofilms de *Anabaena* sp. PCC7120

Los biofilms de cianobacterias y, más en concreto, los de *Anabaena* sp. PCC7120, tienen numerosas potenciales aplicaciones en distintos campos. El primero de ellos es el tratamiento de aguas residuales: las cianobacterias pertenecientes al biofilm pueden asimilar nutrientes como son el CO₂, fósforo inorgánico y compuestos nitrogenados, amonio, nitratos, nitritos, urea y aminoácidos, para la producción de biomasa (28). Esto resulta beneficioso para la eliminación de fósforo y compuestos nitrogenados, factores clave para la eutrofización de las aguas.

Su uso en biorremediación de metales pesados y contaminantes emergentes también es de interés, ya que la carga negativa de sus EPS favorece la bioadsorción y bioacumulación de estas moléculas. Además, pueden degradar pesticidas y herbicidas, por otros mecanismos moleculares diferentes a los EPS (28).

Por último, en el campo de la agricultura podría usarse como biofertilizante debido a su capacidad de fijación de nitrógeno y la posibilidad de retener sustancias de provecho para la planta con sus EPS(28).

3. OBJETIVOS DEL TRABAJO

El propósito de este trabajo es analizar los datos del transcriptoma comparativo de biofilms de *Anabaena* sp. PCC7120 frente a cultivos planctónicos y, a partir de ellos, encontrar nuevas dianas potencialmente reguladas por las proteínas FUR (*ferric uptake regulator*), así como evaluar el efecto de una serie de estreses sobre la formación de biofilms. Por tanto, los objetivos del trabajo son los siguientes:

- Análisis de los datos de RNAseq del biofilm formado por la cianobacteria filamentosa *Anabaena* sp. PCC7120.
- Validación de los datos mediante *real-time* RT-PCR de una selección de genes.
- Estudio de la posible relación entre el estrés por desecación y falta de nitrógeno combinado con la formación de biofilms.
- Identificación de potenciales cajas FUR y análisis de la capacidad de unión de las proteínas FUR a los promotores de una selección de genes de interés cuya transcripción varía significativamente en el biofilm.

4. MATERIALES Y MÉTODOS

4.1. Cultivo y valoración de la formación de biofilms

4.1.1. Preparación e inoculación de cultivos para la obtención de biofilms

Durante la realización de este trabajo se han preparado cultivos formadores de biofilms en dos condiciones: presencia y ausencia de nitrógeno combinado en el medio de cultivo, que se suministró en forma de NaNO_3 . Se establecieron cultivos tanto de la estirpe silvestre de *Anabaena* sp. PCC7120 (WT) como de un mutante que sobreexpresa el parólogo Zur (FurB), denominado VCS2770. En ambos casos se partió de un cultivo en fase exponencial. Se llevaron a cabo tres lavados con medio fresco en esterilidad BG11 (*Anexo.1: Cultivo de Anabaena sp. PCC7120 y medios de cultivo*), para los cultivos en condiciones normales y BG11₀ (*Anexo.1: Cultivo de Anabaena sp. PCC7120 y medios de cultivo*) para los destinados a condiciones de deficiencia de nitrógeno, ya que se eliminan los restos de NaNO_3 . Después de la centrifugación, se resuspendieron en el volumen necesario para ajustar la densidad de cultivo a 0,15 DO_{750nm}. Además, en el caso de la cepa VCS2770, el medio se suplementó con el antibiótico neomicina (50 µg/ml) como presión selectiva que permitiera mantener el plásmido de sobreexpresión en la estirpe.

Los cultivos se inocularon en placas de 8 pocillos *Ibidi µ-Slide 8-well* de polímero sin tratamiento de superficie de cultivo, con un volumen de 300 µL por pocillo. Tras ello, se dejaron crecer las células durante 10 días bajo 5 µM de fotones de irradiación, 28 °C en condiciones de crecimiento estático. Después del periodo de incubación es común obtener tanto células sésiles (que se adhieren a la placa y forman los biofilms) como células planctónicas (que no se adhieren a la placa y están en solución) como, las cuales se retiraron. Seguidamente, se tiñeron los pocillos que contienen el biofilm con 200 µL de cristal violeta (0,25% p/v en H₂O mili-Q esterilizada) durante 2 minutos. Debido a su carga positiva, este reactivo permite la tinción de elementos con carga negativa, tanto de las células de *Anabaena* sp. PCC 7120 sésiles como de la matriz extracelular que sirve de andamiaje para el biofilm. A continuación, se retiró el cristal violeta y se realizaron tres lavados con H₂O mili-Q para eliminar los restos de colorante que adheridos de forma inespecífica a los pocillos.

4.1.2. Cuantificación de biofilms

La cantidad de biofilm en cada pocillo fue evaluada mediante cuantificación por imagen de la intensidad de los biofilms teñidos con cristal violeta, utilizando el software de procesamiento de imágenes ImageJ v1.53. La biomasa de biofilm se expresa como valores de densidad integrada bruta. Para normalizar estos datos se dividió el valor correspondiente a cada pocillo por el promedio de los pocillos de la cepa WT, que es el control para el experimento.

4.2. Técnicas de trabajo con DNA

4.2.1. Extracción de DNA genómico

La amplificación de promotores requiere de DNA genómico de *Anabaena* sp. PCC7120, procedente de cultivo líquido. Su obtención se llevó a cabo mediante un protocolo adaptado del método descrito por Cai y Wolk (1990). En primer lugar, se centrifugaron 50 ml de cultivo en fase exponencial a 4000 rpm y temperatura ambiente durante 5 min. Tras ello, se descartó el sobrenadante y se resuspende el precipitado celular en el volumen restante para repartirse en dos tubos Eppendorf de 2 mL con tapón de seguridad. A continuación, se centrifugaron los tubos Eppendorf a 12000 rpm y 4 °C durante 5 min y se descartó por completo el sobrenadante. El precipitado de células obtenido se resuspendió en 400 µL de Tris-HCl 10 mM pH 8 con EDTA 0,1 mM y se añadieron 20 µL de SDS al 10%, 450 µL de fenol-cloroformo (1:1) y 3 puntas de espátula de perlas de vidrio y se realizaron cuatro ciclos de 1 min. de agitación en vortex y 1 min. de reposo en hielo, consiguiendo así la rotura de las células. Seguidamente, el lisado obtenido se centrifugó durante 15 min. a 12700 rpm y 4 °C. La fase superior acuosa se transfirió a tubos Eppendorf nuevos para llevar a cabo la extracción con 1 vol. de fenol, 1 vol. de fenol-cloroformo (1:1) y 2 veces con 1 vol. de cloroformo hasta conseguir una fase acuosa completamente incolora. En este paso, añadir un volumen se refiere a añadir un volumen de reactivo igual al de la muestra y su posterior centrifugación. En todos los lavados, se descartó la fase orgánica inferior. Posteriormente, se añadió a la fase acuosa 1 vol. de acetato de sodio 3 M pH 5.2 y 2 vol. de etanol absoluto frío y se incubó durante 1 h a -80 °C para conseguir la precipitación del DNA. Una vez transcurrido este tiempo, se centrifugó durante 15 min a 4 °C y velocidad máxima. El DNA obtenido se lavó con etanol al 70% preenfriado, se secó al aire y se resuspendió en 50 µL de H₂O MQ estéril. Por último, el DNA obtenido se cuantificó con un equipo *NanoVue Plus TM* (GE Healthcare).

4.2.2. Reacción en cadena de la polimerasa (PCR)

La PCR se utilizó con el objetivo de amplificar promotores de diversos genes de *Anabaena* sp. PCC7120 para poder ser utilizados en la realización de ensayos de retardo en gel (EMSA). La mezcla de reacción utilizada para amplificar cada uno de los promotores contiene los siguientes reactivos: 10 µL de Buffer 10x (Tris-HCl 75 mM pH 9,0, KCl 50 mM, (NH₄)₂SO₄), 2 µL de dNTPs 10 mM (dATP, dGTP, dTTP, dCTP), 72 µL de H₂O MQ en condiciones estériles, 3 µL de MgCl₂ 50 mM, 5 µL de cebador directo 20 µM, 5 µL de cebador reverso 20 µM, 2 µL de DNA genómico de *Anabaena* sp. PCC7120 a 50 ng/µl y 1 µL de polimerasa 5U/µl (*Biotools*). La PCR se llevó a cabo en el termociclador *Thermal Cycler 2720* (*Applied Biosystems*). El programa utilizado, así como los cebadores usados y su temperatura de hibridación, se encuentran en el *Anexo 4: Preparación de EMSA*.

4.2.3. Electroforesis de DNA en geles de agarosa

Para comprobar la correcta amplificación de los promotores y ausencia de productos inespecíficos de PCR, se llevaron a cabo electroforesis en gel de agarosa al 2% y se usó como agente de tinción para el DNA el bromuro de etidio (BrEt). Para la preparación del gel de agarosa, se mezcló 1 g de agarosa con 50 mL de TBE 1x (Tris-HCl 90 mM pH 8, ácido bórico 90 mM y EDTA 2 mM), y se calentó hasta que la agarosa se fundió completamente. Tras enfriar la mezcla, cuando alcanza un valor en torno a unos 60 °C, se añadieron 50 µL de BrEt 1000x (0,5g/ml) y se dejó solidificar hasta la realización de la electroforesis. Para cargar cada uno de los pocillos del gel, se preparó la siguiente mezcla: 5 µL de H₂O mq, 3 µL del DNA amplificado y 2 µL de tampón de carga (Tris-HCl 50 mM pH 8, 30% de glicerol (v/v) y 0,25% (p/v) de azul de bromofenol). Finalmente, se corrió la electroforesis en tampón TBE a un voltaje constante de 90 V durante 30 minutos en una fuente *BioRad PowerPac 3000*. Para la visualización de los resultados se utilizó un equipo *GelDoc2000*.

4.2.4. Purificación y cuantificación de los productos de PCR

Para la purificación de los promotores se utilizó el kit *Illustra GFX PCR DNA and Gel Band Purification Kit* (Cytiva), de acuerdo con las instrucciones del fabricante. Una vez obtenidos los promotores puros, sin ningún tipo de impureza proveniente de la PCR, se cuantificó la cantidad de DNA en el equipo *NanoVue Plus TM* (GE healthcare).

4.2.5. Ensayos de retardo en gel (EMSA)

El ensayo de retardo en gel (EMSA, *Electrophoretic Mobility Shift Assay*) fue utilizado para comprobar la capacidad de las proteínas FurA, FurB y FurC de unirse *in vitro* a las regiones promotoras de los genes seleccionados. Esta técnica se basa en que, si existe formación de complejo entre la proteína y la región promotora, esto provoca una migración retardada respecto al promotor libre. Para comprobar que esta unión es específica, se introdujo como DNA competidor el fragmento interno del gen *pkn22*, para el cual las proteínas FUR no tienen afinidad específica.

Los geles utilizados para el ensayo son geles de poliacrilamida no desnaturalizantes al 6 %. Los reactivos utilizados para la preparación de un gel son: 2,89 mL H₂O mili-Q, 700 µL glicerol 50 %, 1 mL, acrilamida – bisacrilamida (30:0,8 p/v) 30 %, 50 µL PSA al 10% p/v, 30 µL TEMED, 0,93 mL de tampón de electroforesis 10x (30,28 g/L Tris-HCl, 142 g/L glicina pH 8,5). En el caso de FurA y FurC, los geles también contuvieron 5,1 µL de MnCl₂ 4 H₂O (100 mM). Los geles se precorrieron a 60 V durante un mínimo de 30 minutos y a 4 °C en tampón de electroforesis 1X. En el caso de FurA y FurC en el tampón se introdujo 1 mL de MgCl₂ 4 H₂O (100 mM).

La mezcla de reacción fue ligeramente diferente dependiendo del parólogo FUR analizado. En el caso de FurA y FurC, cada pocillo contenía 2 µL de tampón de unión 10X, 1 µL de albúmina sérica bovina 1 mg/ml, 1 µL de DTT (ditiotreitól) 20 mM, 1 µL de MnCl₂ (10 mM), 100 ng de promotor y 100 ng de DNA competidor. En el caso de FurC la composición es similar, salvo que se utilizaron 50 ng de promotor y de DNA competidor. En los ensayos con FurB, el MnCl₂ se sustituyó por 1 µL de ZnSO₄ (10 mM) y se utilizaron 50 ng de promotor y 50 ng de DNA inespecífico competidor. Además, las cantidades de proteína final introducida en cada pocillo variaron según la proteína sometida a estudio: FurA (100, 200 y 300 nM); FurB (100, 250 y 500 nm) y FurC (100, 200 y 300 nM). Estas proteínas fueron proporcionadas por el grupo de investigación (Irene Oliván y Jorge Guío). En todos los casos, se incubó el promotor correspondiente sin proteína (0 nM) como control negativo. Por último, en cada mezcla de reacción se añadieron H₂O Milli-Q hasta alcanzar un volumen final de 20 µl. Esta mezcla se incubó durante 30 minutos para favorecer la formación de los complejos DNA-proteína.

Para cargar la mezcla a en cada uno de los pocillos, se añadieron 3 µL de tampón de carga (Tris-HCl 50 mM, pH 8, glicerol 30% (v/v) y azul de bromofenol 0,05% (p/v)) en la mezcla de reacción y se cargaron 12 µL por pocillo en los ensayos de FurA, o 18 µL en el caso de FurB y FurC. Se dejó correr la electroforesis a 90 V durante 110 min y a 4 °C. Una vez finalizada la electroforesis, los geles se tiñeron en 50mL de running buffer con 3 µL de *Sybr Safe Fluorescent Dye* 1 µg/ml durante 10 minutos en agitación y oscuridad. Para visualizar el resultado, se revela el gel en el transiluminador *Gel Doc 2000* (BioRad).

4.2.6. Real time RT-PCR

Previamente a este trabajo, en el grupo de investigación se había llevado a cabo un análisis de transcriptómica comparada de biofilms de *Anabaena* sp. PCC7120 frente a cultivos planctónicos. Para realizar la validación de estos datos, se partió de cDNA obtenido a partir de las mismas muestras de mRNA empleadas para el estudio de transcriptómica, que se retrotranscribieron previamente en el grupo de investigación. En cada ensayo de *real time* RT-PCR, se realizaron 3 réplicas biológicas (3 cDNAs

de biofilm y 3 cDNAs de cultivo planctónico) y 3 réplicas técnicas para cada gen estudiado. Esta técnica se llevó a cabo en placas de 96 pocillos (*Applied Biosystems™ Placa de reacción de 96 pocillos MicroAmp™ Optical Fast*) con una mezcla final de 30 µL en cada pocillo, donde se añadieron: 12,5 µL de *PowerSYBRGreen Master mix*, 0,2 µL de cebador directo y la misma cantidad de reverso ambos a 25µM, 7,1 µL de agua libre de RNasa y 10 µL de cDNA a una dilución previa de 1/50 en agua libre de RNasa, ya que está calibrado que la prueba queda dentro de los rangos sensibilidad con esta dilución. Durante la manipulación de estas muestras hay que tener especial cuidado con las nucleasas, que son ubicuas, por lo que deben llevar guantes, usar material libre de RNAsas y aplicar NaOH 0,4 M en todas las superficies que se vayan a usar.

Los datos obtenidos de cada gen se normalizaron respecto a la expresión del gen constitutivo *rpnB*, que codifica para la RNasa P de *Anabaena* sp. PCC7120 y cuyos niveles de expresión se mantienen invariables durante el crecimiento del cultivo en distintas condiciones (29). El programa usado, el tratamiento de los datos obtenidos, así como los cebadores utilizados se encuentran en el *Anexo 2*.

4.3. Técnicas bioinformáticas

4.3.1. Búsqueda de secuencias

En el presente trabajo se analizaron como regiones promotoras aquellas localizadas desde 500 pb *upstream* hasta 50 pb *downstream* con respecto al ATG de cada gen de interés, en base a la anotación del genoma disponible en las bases de datos KEGG (<https://www.genome.jp/kegg/>) y Cyanobase (<http://genome.microbedb.jp/CyanoBase>). Este rango se empleó por ser considerado en la literatura como suficientemente amplio para contener todas las posibles secuencias reguladoras de la expresión de los genes de estudio.

4.3.2. Búsqueda de operones en los genes del transcriptoma

Para la realización del trabajo, es necesario conocer los operones que se encuentran en dicho genoma, ya que los genes pueden estar regulados directamente por su secuencia promotora o en forma conjunta con otros genes formando un operón, donde la región promotora de los genes se encuentra delante del primer gen perteneciente al operón. Ante la ausencia de datos experimentales en la mayoría de los casos de formación de operones en *Anabaena* sp. PCC 7120, se recurrió a la base de datos MicrobesOnline (<https://microbesonline.org/>) (30). Esta base de datos está formada a partir de una herramienta bioinformática, contrastada empíricamente, que realiza predicciones en base a distintos factores: distancia entre nucleótidos, genes conservados cerca en otros genomas, correlación de los patrones de expresión, pertenencia a la misma categoría funcional. Junto a esta predicción muestra una probabilidad que va comprendida entre 0 y 1, donde 0 sería muy poco probable que se diera ese operón y 1 muy probable.

Por tanto, al tratarse de operones predichos, hay que comprobar en el transcriptoma que se dé un cambio de expresión significativo de forma similar en todos los genes que forman el operón. Además, si un determinado gen tiene posibilidad de pertenecer a un operón y motivo de unión a una proteína FUR, debe comprobarse si el primer gen del operón posee una caja de unión a esta proteína FUR, como criterio menos restrictivo, ya que, aunque exista caja en el primer gen del operón, este gen puede estar regulado individualmente.

4.3.3. Búsqueda *in silico* de potenciales cajas de unión de las distintas proteínas FUR

Dentro del genoma de *Anabaena* sp. PCC 7120 existen distintas cajas de unión a los parálogos FUR. Para conocer si en la secuencia *upstream* de un determinado gen existe dicha secuencia de unión se ha utilizado la herramienta bioinformática FIMO (<https://meme-suite.org/meme/tools/fimo>)(31). Esta herramienta busca coincidencias entre los motivos de unión de cada proteína FUR (FurA, FurB y FurC),

que se introducen en forma de matrices de peso (*Anexo3*) y las secuencias de interés introducidas, que en este caso se trata de todas aquellas secuencias promotoras que regulan la expresión de uno o más genes diferencialmente expresados en el biofilm. Si existe un motivo de unión a la proteína en la región promotora, es posible que se encuentre regulada por dicha proteína. Para comprobar este hecho, deben realizarse ensayos de retardo en gel (4.2.5. *Ensayos de retardo en gel (EMSA)*). Como criterio para la selección, se buscan coincidencias con un p-valor inferior a 10^{-4} .

4.3.4. Manejo de datos ómicos en hojas de cálculo

Para el presente trabajo, se dispuso de libros de cálculo proporcionados por el grupo de investigación conteniendo los operones de *Anabaena* sp. PCC7120 predichos por *Microbe sOnline* y con anotaciones sobre los distintos genes diferencialmente expresados en el biofilm. También se generaron, según el apartado anterior, hojas con los resultados de las búsquedas de FIMO para cada parólogo FUR. En este apartado, se resumen los procedimientos y fórmulas usadas para manejar los datos ómicos en estas hojas de cálculo, ya que se han identificado más de 1000 genes diferencialmente expresados en biofilms respecto a cultivos planctónicos y su manejo uno a uno de forma convencional resultaría inviable:

- Para simplificar el proceso de búsqueda y recuperación de datos de secuencias, operones y cajas FUR de los genes relacionados con el biofilm se utilizaron hojas de cálculo independientes debido a su gran número de genes. Se utilizó las siguientes funciones: =BUSCAR.X (permite buscar un valor en una columna de una hoja y si se encuentra devolver otro parámetro determinado en otra columna especificada) y =CONTAR.SI (permite contar cuántas veces se encuentra determinado valor en una lista).
- Selección de genes con posible relación con los biofilms: Para empezar, se realizó una selección de genes relacionados con el biofilm en base a estudios previos en la literatura. Para ello, se llevó a cabo una búsqueda selectiva con términos clave que tenían relación con el biofilm dentro de las anotaciones previamente realizadas en el grupo de investigación sobre cada gen: *Desiccation, hydric, water, EPS, polysaccharide, exopolysaccharide, cell wall, envelope, matrix, osmotic, biofilm y drought*. En la hoja de cálculo, la función que se utilizó fue la siguiente: SI (ESNUMERO(HALLAR(“palabra clave”; celda que contiene descripción));1;0). Con esta función condicional, se logra que, si existe la palabra clave en la descripción, se asigne un 1 y si no lo está un 0. Por tanto, a cada gen se le asignó una puntuación correspondiente al número de términos clave con los que se relaciona.

5. RESULTADOS Y DISCUSIÓN

5.1. Validación de los resultados de transcriptómica comparada de biofilms de *Anabaena* sp. PCC7120 frente a células planctónicas.

El punto de partida para realizar este trabajo han sido los resultados de un estudio de transcriptómica comparada (RNAseq) de biofilms de *Anabaena* sp. PCC7120 con respecto a células en fase líquida del mismo cultivo, que se obtuvieron anteriormente en el grupo de investigación. Para ello, se secuenció el mRNA total obtenido de tres réplicas biológicas tanto de los biofilms formados por *Anabaena* sp. PCC7120 como de las células que permanecieron en la fase líquida de cada cultivo. Así pues, tras llevar a cabo esta técnica ómica es necesario validar los datos obtenidos. Dicha validación se realizó mediante *real time* RT-PCR.

Debido a la gran cantidad de genes cuya expresión cambia en el transcriptoma del biofilm, superior a 1000, no se pueden validar todos ellos, sino que se hizo una selección de una serie de genes representativos. Para ello, se escogieron genes representativos pertenecientes a diversas categorías funcionales y que presentaran cambios transcripcionales tanto negativos como positivos.

En base a estos criterios, los genes seleccionados se muestran en la *Tabla 1*. Los resultados obtenidos, recogidos en la *Figura 3* y *Tabla 1*, resultaron favorables: todos los genes muestran cambios transcripcionales superando el límite de ± 2 , por lo que consideran significativos. Además, todos los genes estudiados mediante *real time*-RT-PCR muestran unos niveles de transcripción que se encuentran en correspondencia con los datos de expresión génica obtenidos mediante RNA-seq.

Tabla 1: Genes seleccionados y resultado de la validación

Nombre de gen	Cambio transcripcional en el transcriptoma	Cambio transcripcional (real time-RT PCR)	Desviación estándar	Descripción	Categoría funcional
<i>sjdR</i> (<i>alr0248</i>)	23,66	12,24	$\pm 5,3$	Regulador del disco de unión septal	División celular y control de forma
<i>tonB2</i> (<i>all3585</i>)	16,23	4,28	$\pm 0,973$	Transportador TonB2	Transporte
<i>furC/perR</i> (<i>alr0957</i>)	2,62	1,79	$\pm 0,362$	Regulador transcripcional	Funciones reguladoras
<i>oprB_I</i> (<i>alr0834</i>)	-3,61	-15,51	$\pm 3,340$	Porina	Biosíntesis de pared y mantenimiento
<i>iacT</i> (<i>all4026</i>)	-3,76	-6,32	$\pm 3,871$	Proteína receptora de hierro en membrana externa	Transporte hierro/cobre
<i>ndhO</i> (<i>asr4321</i>)	-5,21	-3,08	$\pm 0,734$	Subunidad O de la NAD(P)H-quinona oxidorreductasa	Fotosíntesis y respiración

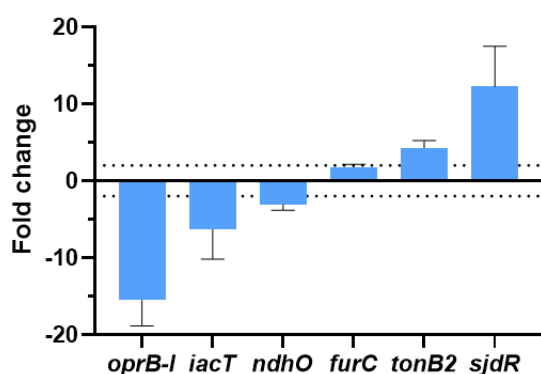


Figura 3: Representación del *fold change* promedio obtenido para cada uno de los genes, así como la desviación típica. El límite a partir del que el cambio se considera significativo se encuentra a ± 2 veces y queda indicado mediante una línea discontinua.

5.2. Estudio de la correlación entre la expresión génica y respuesta a diversos estreses

La formación de biofilms en microorganismos aumenta ante diversos estreses con el fin de favorecer la supervivencia de las células que lo conforman. En la literatura, existe información sobre cómo *Anabaena* sp. PCC7120 responde ante el estrés por desecación y la falta de nitrógeno. Por tanto, uno de los objetivos de este trabajo, es estudiar la posible relación entre la respuesta transcripcional a estos estreses y la formación de biofilms en esta cianobacteria.

5.2.1. Estudio de la correlación entre la expresión génica en biofilms y en respuesta a la desecación.

Estudios previos demostraron que *Anabaena* sp. PCC7120 tiene un cierto nivel de tolerancia a la desecación (32). Además, en condiciones de desecación, se puede inducir la formación de biofilms (18). Para estudiar la relación que existe entre ambas condiciones (estado de biofilm y desecación), se ha estudiado el solapamiento entre las respuestas génicas en ambos casos, representando gráficamente mediante diagramas de Venn los genes que se encuentran diferencialmente expresados (DEGs) tanto positiva como negativamente en ambos casos. Los genes que se encuentran desregulados en desecación se han obtenido a partir de la literatura publicada sobre este tema (32,33).

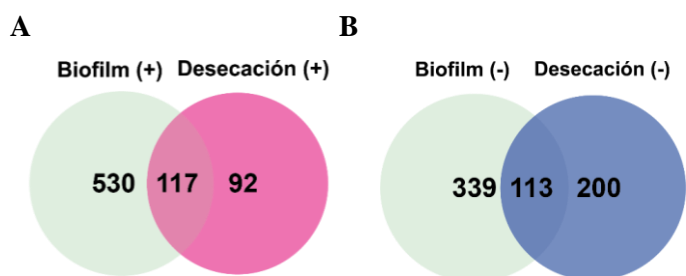


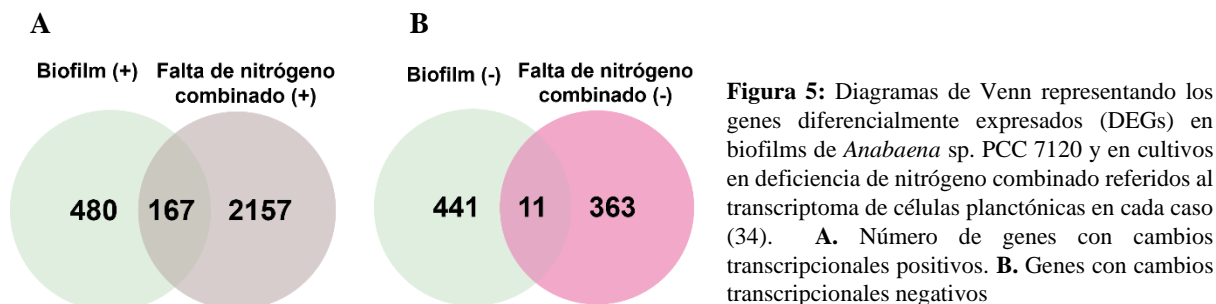
Figura 4: Diagramas de Venn representando los genes diferencialmente expresados (DEGs) en biofilms de *Anabaena* sp. PCC 7120 y en condiciones de desecación referidos al transcriptoma de células planctónicas en cada caso (32,33). **A.** Número de genes con cambios transcripcionales positivos. **B.** Genes con cambios transcripcionales negativos.

De los 313 genes cuya expresión disminuye en condiciones de desecación (*Figura 4B*) hay 133 que también presentan un cambio negativo en el biofilm, lo que indica que un 36% de los genes cuya transcripción se ve afectada en estrés hídrico también varía negativamente en el biofilm. Esto representa un 18% de la suma total de DEGs del biofilm. Por otro lado, en el caso de los 209 genes cuya expresión aumenta en condiciones de desecación (*Figura 4A*), se han identificado 117 genes cuya expresión también aumenta en el biofilm, lo que supone que un 62% de los genes inducidos por desecación, también lo están en el biofilm. En este caso, los 117 genes corresponden al 10% del total de DEGs del biofilm.

Los datos obtenidos muestran un solapamiento considerable de los genes cuya expresión varía en el mismo sentido, tanto aumentando como disminuyendo, en condiciones de desecación y en el biofilm. Así pues, este resultado sugiere que hay una serie de rutas metabólicas activadas o reprimidas en respuesta al estrés por desecación que coinciden con aquellas implicadas en la formación de biofilms y viceversa.

5.2.2. Estudio de la correlación entre la expresión génica en biofilms y en respuesta a la deficiencia de nitrógeno combinado

La respuesta a la deficiencia de nitrógeno incluye cambios metabólicos y morfológicos, con la formación de heterocistos en *Anabaena* sp. PCC7120 (17). Asimismo, en la literatura se encuentran disponibles estudios de RNA-seq sobre *Anabaena* sp. PCC7120 en condiciones de deficiencia de nitrógeno combinado (34). En cuanto al biofilm, el metabolismo del nitrógeno es una de las categorías funcionales donde más genes se encuentran desregulados respecto a las células de *Anabaena* pláctónicas (2.6. *Biofilms de Anabaena* sp. PCC7120). Por tanto, una de las primeras aproximaciones para poder llegar a conocer la relación entre la formación de biofilms y la respuesta de *Anabaena* sp. PCC7120 ante la falta de nitrógeno combinado es el estudio conjunto de genes que cambian su expresión en ambas condiciones (DEGs) tanto positiva como negativamente. Para ello, se ha estudiado el solapamiento génico en las diferentes condiciones, lo que queda representado mediante diagramas de Venn comparando la cantidad de genes que cambian tanto positiva como negativamente.



En el caso de genes desregulados positivamente (*Figura 5A*), 2324 genes están desregulados positivamente ante falta de nitrógeno combinado, de los cuales 167 también presentan cambio transcripcional positivo en el biofilm, lo que indica que un 7,7% de los genes que están desregulados en deficiencia de nitrógeno también lo están en el biofilm. Esos 167 genes representan un 15,1% de los DEGs totales que varían en el biofilm. En el caso de genes desregulados negativamente (*Figura 5B*), 374 genes disminuyen su expresión ante falta de fuente de nitrógeno combinado. Sin embargo solamente hay 11 genes cuya expresión disminuye tanto en el biofilm como en déficit de nitrógeno combinado, lo que supone un 4,4% de los genes que varían su expresión ante falta de nitrógeno combinado. Esto supone un 1% de los DEGs totales del biofilm.

Los datos de obtenidos muestran cómo se produce un mayor solapamiento de genes cuya expresión aumenta en ambas condiciones con respecto a los que disminuyen. Si se atiende al porcentaje de DEGs del biofilm que varían en cada condición, la mayor parte de los genes que están desregulados ante falta de nitrógeno sufrirán sobreexpresión. Por tanto, el patrón de solapamiento difiere entre el estrés por desecación (número similar de DEGs comunes que aumentan y disminuyen) y por falta de nitrógeno combinado, dónde solamente 11 DEGs comunes aumentan frente a los 167 que disminuyen.

5.3. Estudio de la correlación entre la expresión génica en biofilms y ante la desregulación de parálogos FUR

Dado que las proteínas FUR son los principales reguladores de defensa a estreses en microorganismos procariotas, y que estudios previos del grupo de investigación han encontrado una potencial relación entre su expresión y la formación de biofilms, se propuso como uno de los objetivos de este trabajo analizar la relación entre la respuesta transcripcional a la desregulación de los tres parálogos FUR y la formación de biofilms.

Para ello, se emplearon los datos de transcriptómica comparativa frente a la estirpe silvestre de diferentes variantes de *Anabaena* sp. PCC 7120 en los que una de las proteínas FUR se encuentra desregulada (delección condicional “*turn off*” de *furA* (15), delección de *furB* (36) y sobreexpresión de *FurC* (17)), que se compararon con los datos obtenidos de la transcriptómica diferencial de los biofilms frente a cultivos planctónicos.

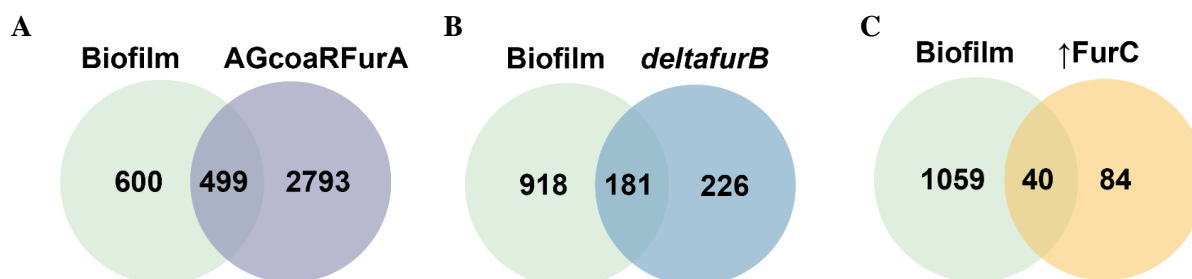


Figura 6: Diagramas de Venn representando los genes diferencialmente expresados (DEGs) en biofilms de *Anabaena* sp. PCC 7120 y en transcriptomas de desregulación de parálogos FUR con respecto a células planctónicas WT en ambos casos. **A:** Transcriptoma de delección condicional “*turn off*” de *furA* (15), **B:** Transcriptoma de delección total de *furB* (36), **C:** Transcriptoma de sobreexpresión de *FurC* (17).

Se observa que la desregulación de FurA afecta a la transcripción de un mayor número de genes, en comparación con las estirpes que tienen desregulados los otros parálogos. También hay un mayor número de genes (499) cuya transcripción cambia tanto en el biofilm como en la estirpe de delección de *furA* (Figura 6A); estos resultados están en consonancia con el papel esencial que juega FurA en la cianobacteria, cuyo regulón es el más extenso de los tres parálogos (15). Esto corresponde a un 45% de DEGs en el biofilm desregulados en ausencia de FurA. En el caso de FurB (Figura 6B) también existe un número considerable de genes (181) cuya expresión génica cambia en el biofilm y en respuesta a la ausencia del regulador. Aunque este número de genes únicamente corresponde al 18,6% de los DEGs identificados en el biofilm, es de destacar que aproximadamente un 45% de los DEGs cuya transcripción está afectada en el mutante de delección de FurB, también lo están en el biofilm. Este hecho está en concordancia con los resultados de la literatura que indican que el nivel celular de FurB tiene consecuencias en la capacidad de formación de biofilms en *Anabaena*. Además, FurB también parece estar relacionada con la regulación del estrés por desecación, y con el proceso de formación de biofilm, ya que se da un solapamiento de la respuesta transcripcional y regulación de genes potencialmente relacionados con la síntesis de EPS, necesarios en la formación de biofilms (36). Por último, en el caso de FurC (Figura 6C) se encontraron 40 genes cuya transcripción varía en ambos transcriptomas, siendo un 3,6% de los DEGs de la formación del biofilm y aproximadamente el 32% de los DEGs identificados en la estirpe que sobreexpresa FurC. Esta proteína está relacionada con la regulación del metabolismo del nitrógeno, una de las categorías funcionales que más varía en la formación del biofilm.

A la vista de los resultados, las redes de expresión génica que condicionan la formación de biofilms parecen guardar relación con FurA o FurB en mayor medida que con FurC, ya sea de forma directa o indirecta, pero en cualquier caso cabría esperar un impacto en la capacidad de formación de biofilms debido a la cantidad de genes cuya expresión se altera en este modo de crecimiento que se ven afectados con la desregulación de estos reguladores.

5.4. Búsqueda de potenciales secuencias de unión a parálogos FUR en genes cuya expresión varía en el biofilm

Todos los cambios transcripcionales observados en el apartado anterior no implican una regulación directa por las proteínas FUR, sino que en algunos casos se podría tratar de regulación indirecta mediante la participación de otros reguladores o RNAs antisentido, así como por mecanismos de compensación. Por lo tanto, es crucial conocer la presencia de potenciales cajas FUR en las regiones promotoras de los DEGs del biofilm, así como la expresión de cuántos DEGs potencialmente regulan la presencia de esas cajas, ya que pueden regular más de un gen en el caso de operones.

Para FurA, se identificaron 361 cajas, que hacen referencia al número total de secuencias promotoras en las que se predice una única caja FUR. Como algunas de estas cajas se ubican al inicio de un operón, existen 440 potenciales DEGs regulados cuya expresión estaría siendo determinada por esas secuencias promotoras. En el caso de FurB, se identificaron 194 cajas únicas, que influyen sobre 226 DEGs. Por último, en el caso de FurC, se identificaron 194 cajas únicas que tienen influencia sobre 232 DEGs.

Por tanto, existe un mayor número de cajas únicas y DEGs potencialmente regulados por parte de FurA, aunque tanto FurB como FurC también podrían tener una implicación en la regulación del biofilm debido a la cantidad de cajas y DEGs potencialmente regulados.

5.5. Selección de genes candidatos a estudio por EMSA

Debido a la gran cantidad de datos que se obtienen a partir del transcriptoma, ya que se trata de una ciencia ómica, resulta crucial realizar una selección adecuada de los genes que se estudiarán

posteriormente por EMSA. Siguiendo el procedimiento descrito en 4.3.4. *Manejo de datos ómicos en hojas de cálculo*, se seleccionaron 52 genes de los 1099 DEGs que contiene el transcriptoma. Sin embargo, solo con esta búsqueda no se puede hacer la selección final, ya que el objetivo del trabajo es encontrar la relación que tienen los parálogos FUR con la formación y mantenimiento del biofilm. Por tanto, a partir de estos genes debe buscarse:

- 1) Si en su región promotora contienen cajas FUR para cualquiera de los tres parálogos, así como si se predice que se encuentran formando parte de un operón, y si lo está, si la cabeza del operón contiene caja FUR. Que exista una caja FUR es un indicio que pueda estar regulado por una proteína de este tipo; sin embargo, para comprobar la unión debe realizarse el ensayo de cambio en la movilidad electroforética (EMSA).
- 2) Si su expresión cambia en los transcriptomas de desregulación de los parálogos FurA, FurB y FurC.

No se tuvieron en cuenta genes para los que ya se había realizado previamente EMSA en otros trabajos. La tabla detallando esta información considerada se encuentra en (Anexo 4). La información correspondiente a los genes finalmente seleccionados se resume en la siguiente tabla:

Tabla 2: Unidades transcripcionales consideradas para llevar a cabo los ensayos de EMSA. FC: fold change (cambio transcripcional)

Unidad transcripcional considerada	Miembros unidad transcripcional	Identidad	FC biofilm	Posible regulación por parálogos FUR por presencia de cajas (FC en el transcriptoma correspondiente)		
				FurA	FurB	FurC
<i>alr0894-alr0895</i>	<i>alr0894</i>	Posible policétido ciclase	63,05		+ (+28,75)	+
	<i>alr0895</i>	Alcohol deshidrogenasa	51,49	+	+ (+36,23)	
<i>alr0198-alr0199</i>	<i>alr0198</i>	Proteína desconocida	23,33	+		+
	<i>alr0199</i>	Proteína desconocida	23,90			
<i>alr3415-alr3416</i>	<i>alr3415</i>	Proteína de membrana de tipo insertasa YidC	- 1,64	+		
	<i>alr3416</i>	Proteína de tipo SpoIII-J	- 2,43			
<i>all0802/alr0803-alr0804</i>	<i>all0802</i>	Proteína desconocida	13,58	+ (+2,855)	+ (+6,741)	
	<i>alr0803</i>	Proteína desconocida con homología a histidín-quinasa	14,86			
	<i>alr0804</i>	Posible proteína relacionada con el estrés osmótico	24,59	+	+	
<i>alr1718</i>	<i>alr1718</i>	Posible proteína inducida por estrés con repeticiones KGG / proteína hipotética	10,44		+ (+7,24)	
<i>alr2270-alr2271</i>	<i>alr2270</i>	UDP-3-O-acil N-acetilglucosamina-deacetilasa	- 2,31		+	+
	<i>alr2271</i>	(3R)-hidroximyrastoil-[proteína transportadora de acilo]- deshidratasa	- 2,89			
<i>all0166</i>	<i>all0166</i>	α,α -trehalasa	2,94	+ (+9,458)		
<i>alr0248</i>	<i>alr0248</i>	Regulador de unión del disco septal	23,66			+
<i>asr4653</i>	<i>asr4653</i>	Posible proteína de respuesta a estrés de tipo CsbD/proteína hipotética	11,24	+		

Entre los genes seleccionados destaca el gran aumento de expresión en el biofilm de la unidad transcripcional *alr0894-0895*. El gen *alr0894* codifica una posible policétido ciclasa, ya que presenta dominios homólogos a las proteínas de este tipo según *Pfam*. Estas proteínas están implicadas en la formación de metabolitos secundarios y se predice la presencia de cajas en su secuencia promotora para FurB y FurC. El gen *alr0895* codifica una alcohol deshidrogenasa y contiene en su región promotora cajas para FurA y FurB. Ambos genes se encuentran sobreexpresados ante situaciones de estrés por desecación (32) Según la predicción de operones de Microbes Online, la unidad transcripcional podría abarcar desde *alr0892* hasta *alr0895*; sin embargo, la confianza de la predicción se reduce a 0,63 sobre 1 para *alr0893-alr0894*. Además, el cambio transcripcional es muy diferente entre los 4 genes: *alr0892* (19,07), *alr0893* (38,01), *alr0894* (63,05) y *alr0895* (51,49). Esto ocurre de forma similar en el transcriptoma de delección de *furB*, puesto que no hay cambios significativos para *alr0892-alr0893* (*fold change* fuera del rango de significación (± 2)) pero sí para *alr0894* (+28,75) y *alr0895* (+36,93) (36). Estos datos sugieren que ambos genes constituyen una unidad transcripcional independiente, con sus propias cajas FUR.

La unidad transcripcional *alr0198-alr0199* está formada por los genes *alr0198* y *alr0199*. Ambos genes codifican proteínas desconocidas y según *Pfam* poseen el mismo dominio, DUF6658, cuya función también es desconocida, pero es un dominio ampliamente distribuido en cianobacterias. Los dos genes muestran un gran cambio transcripcional de valor parecido (*alr0198* +23,33 y *alr0199* +23,90), y su expresión aumenta en condiciones de desecación. El objetivo de estudio será el promotor del gen *alr0198*, ya que es el primer gen del operón, para el cual se predice caja para FurA y FurC.

La siguiente unidad transcripcional a estudiar es la formada por los genes *alr3415* y *alr3416*. El gen *alr3415* codifica una insertasa de tipo YidC que promueve la inserción de proteínas en la membrana citoplasmática de la bacteria (37) y que está relacionada con la percepción del cuórum, así como el sistema de la secreción de la bacteria. Además, se predice una caja para FurA en la región promotora de *alr3415*. El gen *alr3416* codifica una proteína de tipo SpoIIJ, que está relacionada con la definición de la forma celular y generación y mantenimiento de la pared celular, siendo una proteína de anclaje. Además, ambos genes tienen cambios transcripcionales parecidos: *alr3415* (-1,64) y *alr3416* (-2,43). En este caso, el gen *alr3415*, se encontraría fuera del límite de significación del transcriptoma del biofilm, que fue fijado en un cambio transcripcional de ± 2 . Sin embargo, comprobando la importancia que tiene la función descrita y que se encuentra formando parte de un posible operón fue incluido en el estudio.

El gen *all0802* codifica una proteína desconocida, que según *Pfam* posee dominios homólogos a la proteína YliH de *Escherichia coli*, tratándose de una proteína reguladora de la formación del biofilm. Además, también posee dominios homólogos a lasinas. Las lasinas son responsables de la biosíntesis de polihidroxialcanoatos (PHAs), que son poliésteres almacenados que sirven como reserva de carbono y energía (38). Cabe destacar que la región promotora de este gen es compartida con el operón formado por los genes *alr0803/alr0804*. Ambos codifican proteínas desconocidas, pero *alr0803* es homóloga a histidín-quinasa transductoras de señales y según *Pfam* *alr0804* posee un dominio relacionado con la protección frente a estrés osmótico. En esta región promotora se predice presencia de caja para FurA y FurB, lo cual concuerda con el hecho de que *all0802* presenta expresión diferencial en los transcriptomas de FurA y FurB.

El gen *alr1718* codifica una proteína sin caracterizar. Según *Pfam*, posee el motivo KGG, un motivo de repetición Lys-Gly-Gly que se encuentra en proteínas expresadas en condiciones de estrés en bacterias. Además, muestra un cambio transcripcional entre la forma de biofilm y células platónicas de 10,44, por lo que su expresión aumenta considerablemente en la forma de biofilm. Se predice la presencia de caja para FurB en la región promotora de este gen, que sufre un cambio transcripcional en el transcriptoma de delección de *furB*.

La unidad transcripcional *alr2270-alr2271* está formada por los genes *alr2270* y *alr2271*. El gen *alr2270* (*lpxC*) codifica la UDP-3-O-[3-hidroxiimiristoil] N-acetilglucosamina deacetilasa, que es una enzima encargada de la biosíntesis de lipopolisacáridos, más en concreto del lípido A. El gen *alr2271* (*fabZ*) codifica la enzima (3R)-hidroximiristoil-acyl-carrier-protein deshidratasa, que está relacionada con la biosíntesis de ácidos grasos. La unidad transcripcional predicha sería de los genes *alr2268-alr2271*. Sin embargo, los genes *alr2268* y *alr2269* no presentan un cambio transcripcional significativo en el biofilm, y la confianza de la predicción de formación de operón entre ellos es limitada, de 0,53 sobre 1. Además, no existe ninguna caja FUR delante de estos genes, por lo que la posible unidad transcripcional estudiada es la formada por los genes *alr2270-alr2271*, ya que su cambio transcripcional en el biofilm frente al cultivo planctónico es muy similar (-2,31 y -2,89) y únicamente posee caja para FurB y FurC el gen *alr2270*.

El gen *all0166* codifica la enzima α,α -trehalasa, que se encarga de la hidrólisis de la trehalosa para generar monómeros de glucosa, y su síntesis aumenta en deshidratación, ya que es un metabolito osmoprotector (33). La unidad transcripcional predicha está formada por los genes *alr0166-all0168*; sin embargo, se predice presencia de caja para FurA delante del gen *all0166*, por lo que será objeto de estudio este gen únicamente. Además, este gen también presenta expresión diferencial en el transcriptoma de delección de *furA*.

El gen *alr0248* (*sjdR*) codifica una proteína reguladora de la unión septal, estructuras necesarias para la comunicación celular entre las células vegetativas, así como entre células vegetativas y heterocistos ya que llevan a cabo el transporte de nutrientes y moléculas reguladoras (39). Al ser proteínas que tienen una estrecha relación con los heterocistos, su expresión también aumenta ante déficit de nitrógeno. De igual modo, su expresión también aumenta en el biofilm, y se predice la presencia de caja para FurA en su región promotora.

El gen *asr4653* codifica una proteína hipotética que según *Pfam* posee el motivo similar a CsbD. CsbD es una proteína de respuesta general a estrés que se ha descrito en *Bacillus subtilis*, cuya expresión está mediada por el factor σ -B. En este caso, se predice la presencia de caja para FurA en su región promotora.

En base a estos resultados, se seleccionaron 10 genes para realizar ensayos de retardo en gel, de los que 8 cambian positivamente en el transcriptoma y 2 de ellos negativamente. La (Figura 7) muestra el *volcano plot* obtenido tras el análisis del transcriptoma del biofilm, donde se señalan las posiciones de los genes seleccionados. Se puede observar que 6 de estos genes se encuentran en el rango superior de significancia estadística.

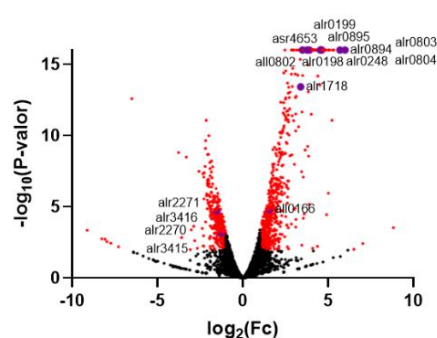


Figura 7: Volcano-Plot realizado a partir de los genes que se encuentran diferencialmente expresados en biofilms de *Anabaena* sp. PCC7120 respecto a cultivos planctónicos. En rojo, se encuentran aquellos que tienen un cambio transcripcional significativo y en negro aquellos cuyo cambio transcripcional no es significativo. En violeta, se indican los genes seleccionados para llevar a cabo el EMSA, y sus unidades transcripcionales. Datos obtenidos directamente a partir del transcriptoma realizado en la tesis doctoral de Irene Oliván Muro.

5.6. Resultados de los ensayos de cambio en la movilidad electroforética (EMSA)

Se realizaron ensayos de cambio en la movilidad electroforética para estudiar la capacidad de unión *in vitro* entre los promotores de los genes de estudio y la/s proteína/s FUR para las cuales presentara una posible caja de unión, con al menos dos réplicas por promotor para asegurar la fiabilidad de los resultados, resumidos en la siguiente tabla:

Tabla 3: Resultado de EMSA para los genes seleccionados.

Unidad transcripcional considerada	Miembros unidad transcripcional	Identidad	FC biofilm	Resultado EMSA		
				FurA	FurB	FurC
<i>alr0894-alr0895</i>	<i>alr0894</i>	Posible policétido ciclasa	63,05		+	-
	<i>alr0895</i>	Alcohol deshidrogenasa	51,49	+	+	
<i>alr0198-alr0199</i>	<i>alr0198</i>	Proteína desconocida	23,33	+		+
	<i>alr0199</i>	Proteína desconocida	23,90			
<i>alr3415-alr3416</i>	<i>alr3415</i>	Proteína de membrana de tipo insertasa YidC	- 1,64	+		
	<i>alr3416</i>	Proteína de tipo SpoIII-J	- 2,43			
<i>all0802/alr0803-0804</i>	<i>all0802</i>	Proteína desconocida	13,58	+	-	
	<i>alr0803</i>	Proteína desconocida con homología a histidín-quinasas	14,87			
	<i>alr0804</i>	Posible proteína relacionada con el estrés osmótico	24,59			
<i>alr1718</i>	<i>alr1718</i>	Posible proteína inducida por estrés con repeticiones KGG / proteína hipotética	10,44		-	
<i>alr2270-alr2271</i>	<i>alr2270</i>	UDP-3-O-acil N-acetilglucosamina deacetilasa	- 2,31		-	+
	<i>alr2271</i>	(3R)-hidroximyristoil-[proteína transportadora de acilo] deshidratasa	- 2,89			
<i>all0166</i>	<i>all0166</i>	α,α -trehalasa	2,94	-		
<i>alr0248</i>	<i>alr0248</i>	Regulador de unión del disco septal	23,66			+
<i>asr4653</i>	<i>asr4653</i>	Posible proteína de respuesta a estrés de tipo CsbD/proteína hipotética	11,24	+		

A continuación, se muestran imágenes representativas de los resultados de los ensayos:

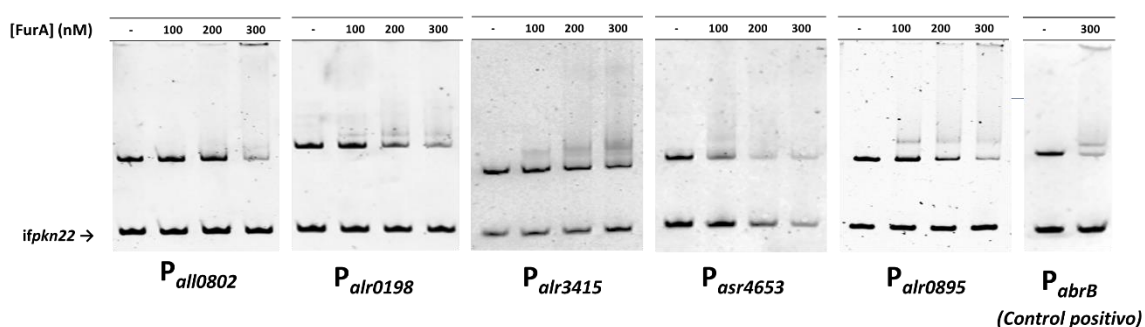


Figura 8: Ensayo de cambio de la movilidad electroforética llevado a cabo para testar la interacción *in vitro* entre FurA y regiones promotoras de una selección de genes. Todos los ensayos fueron realizados con fragmentos de DNA libre o incubados con concentraciones crecientes de FurA (de 0 a 300 nM), separados en gel PAGE al 6%. El fragmento interno del gen *pkn22* (*ifpkn22*) fue usado como DNA competidor no específico. Como control positivo se utilizó la región promotora del gen *abrB*.

En el caso de la región promotora del gen *asr4653*, se observa retardo del inespecífico a mayores concentraciones de FurA; sin embargo, como aparece más tarde que el retardo específico, el resultado se considera igualmente positivo, aunque sugiere una menor afinidad.

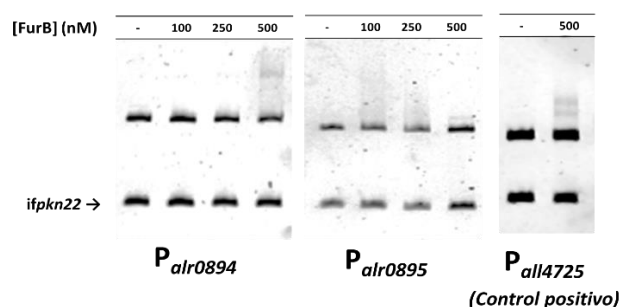


Figura 9: Ensayo de cambio de la movilidad electroforética llevado a cabo para evaluar la interacción *in vitro* entre FurB y regiones promotoras de una selección de genes. Todos ensayos fueron realizados con fragmentos de DNA libre o incubados con concentraciones crecientes de FurB (de 0 a 500 nM), separados en gel PAGE al 6%. El fragmento interno del gen *pkn22* (*ifpkn22*) fue usado como DNA competidor no específico. Como control positivo se utilizó la región promotora del gen *all4725*.

La proteína FurB en estos ensayos se encontraba poco activa, incluso frente a su control positivo, por lo que para intentar solventar este problema la concentración de FurB (nM) introducida fue mayor que en el resto de los ensayos.

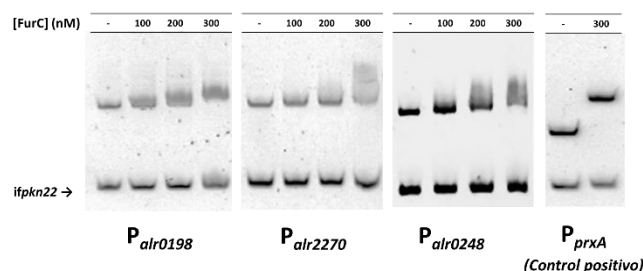


Figura 10: Interacción *in vitro* entre FurC y regiones promotoras de una selección de genes. Los ensayos se realizaron con fragmentos de DNA libre o incubados con la concentración indicada de FurC (0 a 300 nM) y separados en gel PAGE al 6%. El fragmento interno del gen *pkn22* (*ifpkn22*) fue usado como DNA competidor no específico. Como control positivo se utilizó la región promotora del gen *prxA*.

En el caso de *alr0198*, de forma similar al caso del promotor de *asr4653* con FurA, puede apreciarse un cierto arrastre del inespecífico a concentraciones mayores. Sin embargo, se aprecia formación del complejo con el DNA específico a concentraciones menores, indicando que hay interacción con FurC.

Por tanto, los genes cuya secuencia promotora interaccionó con FurA fueron *alr0198*, *alr3415*, *all0802*, *asr4653* y *alr0895*. En el caso de FurB, se observó interacción con los promotores de *alr0894* y *alr0895* y para FurC con los promotores de *alr0198*, *alr2270* y *alr0248*.

La unidad transcripcional *all0802*, que ha resultado positiva en los ensayos de EMSA frente a FurA codifica una proteína desconocida que posee homología con la proteína YliH. Esta proteína se encuentra en *E.coli*, microorganismo también gram-negativo, y se trata de un regulador global de muchos genes implicados en la represión del catabolismo y respuesta a estrés, que también regula la toma y excreción de moléculas tipo autoinductor- 2 (AI-2) (40). Además, *All0802* también posee homología con dominios correspondientes a fasinas, específicamente a PhaF y PhaI, que se encargan de regular la expresión de los genes encargados de la síntesis de PHA (polihidroxicanoatos)(41), que podrían ser usados como fuente de reserva ante condiciones de estrés. Por tanto, *All0802* podría tener un papel en *Anabaena* como posible regulador que activaría otros genes también implicados en la formación del biofilm. Cabría mencionar también las consecuencias regulatorias sobre *alr0803* y *alr0804*. En el caso de *Alr0803* es particularmente de interés por ser una posible histidín-quinasa transductora de señales, que podría participar en alguna cascada regulatoria que controlara la formación de biofilms en respuesta a algún estímulo.

La unidad transcripcional *alr3415-alr3416*, también positiva en los ensayos de EMSA frente a FurA, codifica una insertasa de tipo YidC, que promueve la inserción de proteínas en la membrana citoplasmática de la bacteria (37) y que está relacionada con la percepción del cuórum, así como el sistema de secreción de la bacteria, podría guardar relación con la secreción de componentes de matriz u otras proteínas que condicionen (positiva o negativamente) la formación de biofilms. Esta proteína corresponde a la familia universalmente conservada YidC/Oxa2/Alb3. El gen *alr3416* codifica una

proteína de tipo spoIII-J, que está relacionada con la definición de la forma celular y generación y mantenimiento de la pared celular. Además, la expresión de ambos genes disminuye en los biofilms.

La unidad transcripcional formada por *alr4653*, que también ha resultado positiva en los ensayos EMSA para FurA, codifica una proteína con función desconocida que posee un dominio CsbD. CsbD es una proteína de respuesta general a estrés que se ha descrito en *Bacillus subtilis*, cuya expresión está mediada por el factor σ -B. Además, probablemente, es extracelular, ya que se encuentran ancladas a la superficie celular mediante interacción mediada por cationes, unión característica de bacterias gram-negativas. La evidencia de su localización en la membrana extracelular se ha obtenido a través de estudios anteriores que lograron purificar esta proteína a partir de un aislamiento selectivo de las proteínas de membrana (42). Podría así Alr4653 estar relacionada con la percepción o respuesta a estrés, potencialmente afectando al proceso de formación de biofilms.

FurA y FurC se unen al promotor de la unidad transcripcional *alr2270(lpxC)-alr2271(fabZ)*, que está compuesta por los genes que codifican a dos enzimas: una deacetilasa y una deshidratasa, respectivamente, implicadas en la síntesis del lipopolisacárido A, con función en la pared celular. Existe una disminución de su expresión en el biofilm. Este hecho no es aislado, ya que tanto la unidad transcripcional *alr3415-alr3416* como la unidad transcripcional *alr2270-2271* codifican proteínas con funciones relacionadas con la pared celular y en el transcriptoma la mayoría de genes relacionados con la pared celular siguen la misma tendencia; es decir, disminuir su expresión en el biofilm. Esto podría tener relación con la necesidad de alterar la rigidez de membrana durante la formación de biofilms.

En el caso de la unidad transcripcional *alr0198-alr0199*, se han obtenido ensayos positivos de EMSA frente a FurA y FurC. Estos genes codifican varias proteínas desconocidas con dominios también desconocidos pero conservados en cianobacterias. Estas proteínas probablemente también están asociadas a la membrana, ya que han podido ser aisladas de la fracción de proteínas de membrana extracelular (42). Para ser traslocadas a la superficie externa, estas proteínas poseen un péptido señal específico. Además, la expresión de ambas se incrementa durante la desecación, lo que sugiere que su función está relacionada con la capacidad de la cianobacteria para hacer frente a este estrés y, por tanto la formación de biofilms, debido al solapamiento génico en desecación (*Figura 4*). Sin embargo, debido a la falta de información sobre los dominios de estas proteínas, serían necesarios más estudios para conocer con certeza su función exacta. El gen *alr0248* ha resultado positivo en los ensayos EMSA frente a FurC y codifica una proteína reguladora de unión del disco septal (*sjdR*). La unión septal es una estructura ubicada en las cianobacterias multicelulares que facilita la comunicación entre células. En esta zona existe una capa de peptidoglicano atravesada por ciertas proteínas que permiten la unión física de células adyacentes a través de la membrana de las diferentes células, y facilitando el transporte de metabolitos entre ellas (43). SjdR está estrechamente relacionada con la membrana plasmática y cumple una función importante en el proceso de intercambio celular. Se ha observado que su expresión aumenta durante la formación del biofilm, lo que sugiere que desempeña un papel crucial en el intercambio entre células, función relevante en los biofilms donde se altera la comunicación celular (20). Además, el perfil de expresión de las proteínas de membrana que se relacionan con la membrana plasmática difiere de aquellas proteínas que están relacionadas con la pared celular, como por ejemplo las proteínas Alr2270-Alr2271 y Alr0198-Alr0199. Mientras que las proteínas relacionadas con el mantenimiento de la membrana aumentan, aquellas relacionadas con la formación de la pared celular disminuyen. Esto sugiere que la formación del biofilm implica cambios en la forma, así como en la rigidez celular.

Respecto a la unidad transcripcional formada por *alr0894-alr0895*, se han obtenido ensayos positivos con FurB tanto para la región de promotora de *alr0894* como para la de *alr0895*. Además, también el resultado fue positivo para la interacción de FurA con la secuencia promotora de *alr0895*. En base a la

predicción de *Microbes Online*, la unión de *alr0894* podría afectar a la expresión de un operón constituido por éste y *alr0895*, mientras que de forma independiente podría haber regulación a nivel de *alr0895*. El primer gen, *alr0894*, codifica una policétido ciclasa. En la naturaleza hay enzimas de este tipo implicadas en la síntesis de lactonas (44), siendo la molécula de percepción del cuórum en *Anabaena* sp. PCC7120 una N-acilhomoserina lactona (22). A esto se le suma el alto cambio transcripcional que muestra entre la forma de biofilm y células planctónicas. Por tanto, esto podría sugerir que esta enzima esté implicada en la síntesis de la molécula responsable de la percepción del cuórum, con posibles consecuencias relevantes para el comportamiento de las células en el biofilm. El gen *alr0895* codifica una alcohol deshidrogenasa, cuya expresión aumenta en el biofilm de forma considerable con respecto a las células planctónicas. En otras bacterias gram negativas, como pueden ser *Acinetobacter baumannii* (45) o *Pseudomonas aeruginosa* (46), también se da un aumento considerable de la expresión de genes que codifican alcohol deshidrogenasa. Por lo que previsiblemente, la alcohol deshidrogenasa tendría una función conservada en la formación de biofilms, actuando como molécula clave capaz de modular el sistema de percepción del cuórum (45).

Cabe mencionar que en el marco de este trabajo se han realizado ensayos sobre un número reducido de genes debido a las limitaciones de tiempo, por lo que para poder elucidar el papel de cada proteína FUR en el control de la formación de biofilms sería necesario llevar a cabo más experimentos, así como otros ensayos ómicos, ya que los resultados obtenidos en este trabajo resultan prometedores y abren la puerta a un campo en el que profundizar. Además, existen genes ya positivos en EMSA para algún parálogo FUR según trabajos anteriores que podrían guardar relación con la formación de biofilms, por lo que este hecho apoyaría todavía más el papel de las proteínas FUR en este proceso.

5.7. Capacidad de formación de biofilms en cepa WT (silvestre) y VCS2770 (cepa de sobreexpresión de FURB) en ausencia y presencia de nitrógeno combinado

Los biofilms están relacionados con la respuesta ante diversos estreses, siendo un estrés de interés la deficiencia de nitrógeno. Se ha observado que existe un cierto solapamiento entre los DEGs en ausencia de nitrógeno combinado y ante la formación de biofilm (5.2.2. *Estudio de la correlación entre la expresión génica en biofilms y en respuesta a la deficiencia de nitrógeno combinado*). Además, existe una alta cantidad de genes pertenecientes al metabolismo del nitrógeno que se encuentran desregulados ante la formación del biofilm (2.6. *Biofilms de Anabaena sp. PCC7120*). Por lo tanto, existen indicios de una posible relación entre el déficit de nitrógeno combinado y la formación de biofilm.

Se realizaron ensayos de producción de biofilms con las estirpes silvestre (*Anabaena* sp. PCC7120) y de sobreexpresión de FurB (VCS2270). Según la literatura, al nivel celular de FurB está relacionado con la alteración de la capacidad de formación de biofilms (36). Sin embargo, la delección de *furB* compromete la viabilidad de *Anabaena* en deficiencia de nitrógeno, ya que forma menos heterocistos y su patrón de desarrollo parece incorrecto (36), por lo que sólo se empleó en este caso la estirpe de sobreexpresión. El resultado obtenido tras el cultivo de biofilms durante 10 días fue el siguiente:

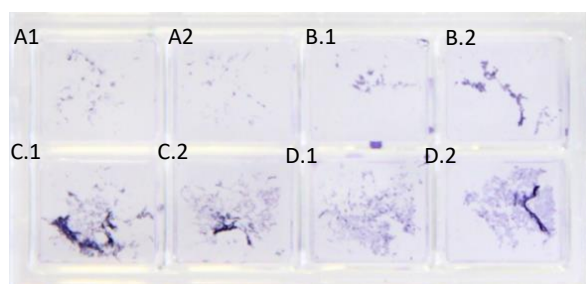


Figura 11: Formación de biofilms por *Anabaena* sp. PCC7120 (A.1, A.2, C.1 y C.2) y VCS2270 (B.1, B.2, D.1 y D.2) en presencia (A.1, A.2, B.1 y B.2) o ausencia (C.1, C.2, D.1 y D.2) de nitrógeno combinado.

Tabla 4: Resultados de la cuantificación de los biofilms formados por las estirpes silvestre y de sobreexpresión de FurB (VCS2770) en condiciones normales (CN) o en deficiencia de nitrógeno (-N). Se muestra la ratio de DIB (densidad integrada bruta), proporcional a la biomasa de biofilm, de cada pocillo con respecto al control (cepa)

Cepa	Condición	Ratio de DIB con respecto al control
WT	CN	1,00±0,83
WT	-N	1,47±0,24
VCS2770	CN	8,87±1,47
VCS2770	-N	9,00±0,29

Tabla 5: Resultados de test paramétricos t-test desapareados realizados mediante el programa GraphPad Prism 8.0.1

Parejas de datos	p-valor (nivel de significación)
WT vs WT (-N)	0,017
Sobreexpresión FurB vs sobreexpresión FurB (-N)	0,0013
WT vs Sobreexpresión FurB	0,158
WT (-N) vs sobreexpresión FurB (-N)	0,92

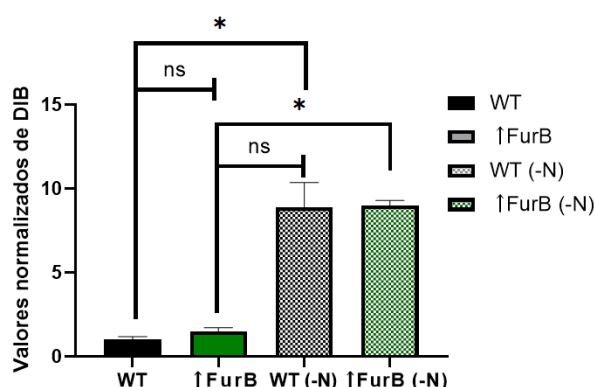


Figura 12: Representación de la DIB (Densidad integrada bruta) normalizada respecto a la WT (estirpe silvestre) en condiciones normales. En la gráfica se muestra los valores para todas las condiciones empleadas WT (estirpe silvestre), ↑ FurB (sobreexpresión de FurB) y (-N), falta de nitrógeno combinado. Además, se muestra la significancia entre datos: Ns (no significativa) y * (significativa)

Las cepas formaron significativamente más biofilm en ausencia de una fuente de nitrógeno combinado (p -valor $< 0,05$), con $8,9 \pm 1,5$ veces más formación de biofilm para la cepa de *Anabaena* sp. 7120 silvestre y $9,0 \pm 0,3$ veces más para la cepa de sobreexpresión de FurB (VCS2270). Por tanto, este resultado muestra que la ausencia de nitrógeno favorece la formación de biofilms.

La cepa de sobreexpresión de FurB no formó significativamente (p -valor $> 0,05$) más biofilm, independientemente de la condición. Por ello, no se puede afirmar en base a estos resultados que haya una posible relación entre sobreexpresión de FurB y mayor formación de biofilms, datos lo cual no se corresponde a los datos publicados (36), donde si formó significativamente más biofilm la cepa de sobreexpresión. La diferencia de resultados puede deberse a las condiciones de cultivo o tinción, que pueden no haber sido las mismas, ya que el resultado obtenido se encuentra cercano al límite de significación (p -valor 0,158). Por ejemplo, el cultivo inicial se lavó tres veces incluso para el crecimiento de biofilms en condiciones normales para igualar las condiciones de puesta con los biofilms en deficiencia de nitrógeno combinado.

Los resultados obtenidos en este experimento revelan que la falta de nitrógeno combinado afecta significativamente a la formación de biofilms. Las características de los biofilms de *Anabaena* sp. PCC7120, junto con la capacidad de formar biofilms en mayor medida en condiciones de deficiencia de nitrógeno, indican una adaptación eficiente a entornos con limitaciones nutricionales. Por tanto, los biofilms de *Anabaena* sp. PCC7120 podrían ser una potencial herramienta biotecnológica para el campo de la agricultura, especialmente como biofertilizante, ya que estos microorganismos podrían promover el crecimiento de plantas al proporcionarles los nutrientes necesarios y mejorar su capacidad de absorción de nutrientes del suelo (47). Así pues, los resultados indican que los biofilms de *Anabaena*

sp. PCC7120 podrían resultar de interés como alternativa a los fertilizantes químicos usados actualmente, ya que contribuirían a la preservación del medio ambiente y a una mayor sostenibilidad agrícola debido a su posible bajo coste y reducción de contaminantes emitidos a suelos y acuíferos.

5.8. Consideraciones finales

La validación y análisis de los datos de transcriptómica comparativa de *Anabaena* sp. PCC7120 en comparación con cultivos planctónicos han revelado la existencia de interrelación a nivel de expresión génica entre esta forma de crecimiento y la adaptación a estreses, como la desecación o deficiencia de nitrógeno, así como la desregulación de proteínas implicadas en la respuesta a estreses como son los parálogos FUR. La mayor resiliencia de los biofilms ante estreses, así como los resultados que demuestran una mayor formación en deficiencia de nitrógeno, los convierten en un objeto de estudio atractivo para su posible aplicación en biotecnología, como la biofertilización en suelos pobres y degradados.

La identificación de nuevas dianas de regulación por parte de las proteínas FurA, FurB y FurC, las cuáles corresponden a genes que se expresan de manera diferencial en los biofilms y potencialmente están implicados en su formación, constituye un avance en la comprensión de las rutas moleculares que controlan este proceso, así como el papel de las proteínas en el mismo. Esto sienta la base para la realización de futuros estudios con el fin de profundizar en si la modulación de la actividad de estos parálogos FUR y sus consecuencias sobre la formación de biofilms pueden mejorar las características de interés en los biofilms o limitar el impacto negativo que podrían tener sobre el entorno.

6. CONCLUSIONES

Los resultados obtenidos en el presente trabajo han permitido extraer las siguientes conclusiones:

- Existe en *Anabaena* sp. PCC7120 un solapamiento en la expresión génica diferencial ante el estrés por desecación y la formación de biofilms, así como entre la falta de nitrógeno combinado y la formación de biofilms
- Los parálogos FUR de *Anabaena* sp. PCC7120 tienen implicación en la regulación directa de genes potencialmente relacionados con el desarrollo y mantenimiento de biofilms en esta cianobacteria
- La falta de nitrógeno es un factor condicionante sobre la formación de biofilms en *Anabaena* sp PCC7120

6. CONCLUSIONS

The work carried out leads to the following conclusions:

- There is an overlap in the differential gene expression in *Anabaena* sp. PCC7120 between desiccation stress and biofilm formation, as well as between combined nitrogen deficiency and biofilms formation.
- FUR paralogs in *Anabaena* sp. PCC7120 are directly involved in the regulation of genes potentially related to the development and maintenance of the biofilm in this cyanobacterium.
- Nitrogen deficiency is a determining factor in biofilm formation in *Anabaena* sp. PCC7120.

7. BIBLIOGRAFÍA

1. Veaudor T, Blanc-Garin V, Chenebault C, Diaz-Santos E, Sassi JF, Cassier-Chauvat C, et al. Recent Advances in the Photoautotrophic Metabolism of Cyanobacteria: Biotechnological Implications. *Life (Basel)*. 2020 May 1;10(5).
2. Chen MY, Teng WK, Zhao L, Hu CX, Zhou YK, Han BP, et al. Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation. *ISME J*. 2021 Jan 17;15(1):211–27.
3. Gutu A, Kehoe DM. Emerging Perspectives on the Mechanisms, Regulation, and Distribution of Light Color Acclimation in Cyanobacteria. *Mol Plant*. 2012 Jan;5(1):1–13.
4. Martin W, Kowallik K. Annotated english translation of mereschkowsky's 1905 paper 'Über natur und ursprung der chromatophoren impflanzenreiche.' *Eur J Phycol* 1999 Aug 1;34(3):287–95.
5. Allaf MM, Peerhossaini H. Cyanobacteria: Model Microorganisms and Beyond. *Microorganisms* 2022 Apr;10(4).
6. Flores E, Herrero A. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nature Reviews Microbiology* 2009 8:1 . 2009 Dec 7;8(1):39–50.
7. Abed RMM, Dobretsov S, Sudesh K. Applications of cyanobacteria in biotechnology. *J Appl Microbiol*. 2009;106(1):1–12.
8. Singh JS, Kumar A, Rai AN, Singh DP. Cyanobacteria: A Precious Bio-resource in Agriculture, Ecosystem, and Environmental Sustainability. *Front Microbiol* . 2016 Apr 21 ;7(APR).
9. Huertas MJ, Mallén-Ponce MJ. Dark side of cyanobacteria: searching for strategies to control blooms. *Microb Biotechnol*. 2022 May 1 ;15(5):1321–3.
10. Videau P, Cozy LM. *Anabaena* sp. strain PCC 7120: Laboratory Maintenance, Cultivation, and Heterocyst Induction. *Curr Protoc Microbiol*. 2019 Feb 1;52(1):e71.
11. Ernst A, Black T, Cai Y, Panoff JM, Tiwari DN, Wolk CP. Synthesis of nitrogenase in mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst development or metabolism. *J Bacteriol* 1992 ;174(19):6025–32.
12. Mehler AH. Studies on reactions of illuminated chloroplasts. II. Stimulation and inhibition of the reaction with molecular oxygen. *Arch Biochem Biophys* . 1951 ;34(2):339–51.
13. Sevilla E, Bes MT, Peleato ML, Fillat MF. Fur-like proteins: Beyond the ferric uptake regulator (Fur) paralog. *Arch Biochem Biophys*. 2021 Apr 15 ;701.
14. Hernández JA, López-Gomollón S, Bes MT, Fillat MF, Peleato ML. Three fur homologues from *Anabaena* sp. PCC7120: exploring reciprocal protein-promoter recognition. *FEMS Microbiol Lett* . 2004 Jul 15 ;236(2):275–82.
15. González A, Bes MT, Luisa Peleato M, Fillat MF. Expanding the Role of FurA as Essential Global Regulator in Cyanobacteria. *PLoS One* . 2016 Mar 1 11(3).
16. 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. PCC 7120. *Environ Microbiol* . 2015 Jun 1 ;17(6):2006–17.
17. 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* . 2022 Feb 1 24(2):566–82. Available from: <https://pubmed.ncbi.nlm.nih.gov/33938105/>
18. Sauer K, Stoodley P, Goeres DM, Hall-Stoodley L, Burmølle M, Stewart PS, et al. The biofilm life cycle: expanding the conceptual model of biofilm formation. *Nat Rev Microbiol*. 2022 Oct 1;20(10):608–20.
19. Sutherland IW. The biofilm matrix – an immobilized but dynamic microbial environment. *Trends Microbiol*. 2001 May 1;9(5):222–7.
20. Solano C, Echeverez M, Lasa I. Biofilm dispersion and quorum sensing. *Curr Opin Microbiol*. 2014;18(1):96–104.
21. Miller MB, Bassler BL. Quorum sensing in bacteria. *Annu Rev Microbiol*. 2001;55:165–99.
22. Romero M, Muro-Pastor AM, Otero A. Quorum sensing N-acylhomoserine lactone signals affect nitrogen fixation in the cyanobacterium *Anabaena* sp. PCC7120. *FEMS Microbiol Lett* . 2011 Feb 1;315(2):101–8.
23. Rossi F, De Philippis R. Role of Cyanobacterial Exopolysaccharides in Phototrophic Biofilms and in Complex Microbial Mats. *Life* . 2015 Apr 1 ;5(2):1218.
24. Pereira S, Zille A, Micheletti E, Moradas-Ferreira P, De Philippis R, Tamagnini P. Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly.
25. Kawaguchi T, Decho AW. Biochemical Characterization of Cyanobacterial Extracellular Polymers (EPS) from Modern Marine Stromatolites (Bahamas). <http://dx.doi.org/101080/10826060008544971>. 2008;30(4):321–30.

26. Tamaru Y, Takani Y, Yoshida T, Sakamoto T. Crucial role of extracellular polysaccharides in desiccation and freezing tolerance in the terrestrial cyanobacterium *Nostoc commune*. *Appl Environ Microbiol*. 2005 Nov;71(11):7327–33.
27. Maeda K, Okuda Y, Enomoto G, Watanabe S, Ikeuchi M. Biosynthesis of a sulfated exopolysaccharide, synechan, and bloom formation in the model cyanobacterium *synechocystis* sp. Strain pcc 6803. *Elife*. 2021 Jun 1;10.
28. Roeselers G, Loosdrecht MCMV, Muyzer G. Phototrophic biofilms and their potential applications. *J Appl Phycol*. 2008 Jun ;20(3):227–35.
29. Vioque A. Analysis of the gene encoding the RNA subunit of ribonuclease P from cyanobacteria. *Nucleic Acids Res*. 1992 Dec 12 ;20(23):6331.
30. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics*. 2011 Apr 1 ;27(7):1017–8.
31. Price. MN HKH, AEJ, and AAP. A Novel Method for Accurate Operon Predictions in All Sequenced Prokaryotes. *Nucleic Acids Res*. 2005;880–92.
32. Katoh H, Asthana RK, Ohmori M. Gene Expression in the Cyanobacterium *Anabaena* sp. PCC7120 under Desiccation. *Ecology*. 2004 ; 47(2):164–74. Available from: <https://www.jstor.org/stable/25153042>
33. Higo A, Katoh H, Ohmori K, Ikeuchi M, Ohmori M. The role of a gene cluster for trehalose metabolism in dehydration tolerance of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Microbiology (N Y)*. 2006 Apr ;152(4):979–87.
34. Flaherty BL, Van Nieuwerburgh F, Head SR, Golden JW. Directional RNA deep sequencing sheds new light on the transcriptional response of *Anabaena* sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC Genomics* [Internet]. 2011 Jun 28;12.
35. González A, Bes MT, Peleato ML, Fillat MF. Unravelling the regulatory function of FurA in *Anabaena* sp. PCC 7120 through 2-D DIGE proteomic analysis. *J Proteomics* 2011 May 1 74(5):660–71.
36. Irene Olivan-Muro, Cristina Sarasa-Buisan, Jorge Guio, Jesus Arenas-Bustos, Emma Sevilla, Maria F. Fillat. Unbalancing Zur (FurB)-mediated homeostasis in *Anabaena*.sp.PCC7120: Consequences on metal trafficking, heterocyst development and biofilm formation. Aceptado, pendiente de publicar.
37. Wang P, Dalbey RE. Inserting membrane proteins: The YidC/Oxa1/Alb3 machinery in bacteria, mitochondria, and chloroplasts. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2011 Mar 1;1808(3):866–75.
38. Pfam: Phasin .
39. Schätzle H, Arévalo S, Flores E, Schleiff E. A TonB-Like Protein, SjdR, Is Involved in the Structural Definition of the Intercellular Septa in the Heterocyst-Forming Cyanobacterium *Anabaena*. *mBio*. 2021 Jun 1;12(3).
40. Domka J, Lee J, Wood TK. YliH (BssR) and YceP (BssS) Regulate *Escherichia coli* K-12 Biofilm Formation by Influencing Cell Signaling. *Appl Environ Microbiol* 2006 Apr 72(4):2449.
41. Poly granule associated (IPR008769) - InterPro entry – InterPro.
42. Yoshimura H, Ikeuchi M, Ohomori M. Cell Surface-Associated Proteins in the Filamentous Cyanobacterium *Anabaena* sp. strain PCC 7120. *Microbes Environ* 2012 Dec, 27(4):538.
43. Kieninger AK, Maldener I. Cell-cell communication through septal junctions in filamentous cyanobacteria. *Curr Opin Microbiol* 2021 Jun 1 61:35–41.
44. Gao DW, Jamieson CS, Wang G, Yan Y, Zhou J, Houk KN, et al. A Polyketide Cyclase That Forms Medium-Ring Lactones. *J Am Chem Soc*. 2021 Jan 13;143(1):80–4. Available from: <https://pubmed.ncbi.nlm.nih.gov/33351624/>
45. Zhang K, Yang X, Yang J, Qiao X, Li F, Liu X, et al. Alcohol dehydrogenase modulates quorum sensing in biofilm formations of *Acinetobacter baumannii*. *Microb Pathog*. 2020 Nov;148:104451.
46. Finelli A, Gallant C V., Jarvi K, Burrows LL. Use of In-Biofilm Expression Technology To Identify Genes Involved in *Pseudomonas aeruginosa* Biofilm Development. *J Bacteriol*. 2003 May;185(9):2700.
47. Kollmen J, Strieth D. The Beneficial Effects of Cyanobacterial Co-Culture on Plant Growth. *Life (Basel)* 2022 Feb 1 12(2).

ANEXOS

ANEXO.1: Cultivo de *Anabaena* sp. PCC7120 y medios de cultivo

Tabla S. 1: Cepas de *Anabaena* sp. PCC7120 y lugar de procedencia

Estirpe	Lugar de procedencia
<i>Anabaena</i> sp. PCC7120 (estirpe salvaje)	Colección de cultivos del Instituto Pasteur
VCS2770 (Estirpe de sobreexpresión de FurB)	Sein-Echaluze VC, Gonzalez A, Napolitano M, Luque I, Barja F, Peleato ML, Fillat 659 MF. Zur (FurB) is a key factor in the control of the oxidative stress response in 660 <i>Anabaena</i> sp. PCC 7120. Environmental microbiology 2014; Epub

Tabla S. 2: Medios de cultivo para *Anabaena* sp.7120

Medio BG11	Medio BG ₀ 11
<ul style="list-style-type: none"> - 1,35g/L de NaNO₃ - 0,738g/L de NaHCO₃ - 0,04g/L de K₂HPO₄. - 0,036 g/L de CaCl₂ · 2 H₂O - 0,075 g/L de MgSO₄ · 7 H₂O - 0,02 g/L de Na₂CO₃ Suplementado con 1 mL/L de solución basal de solución de micronutrientes	<ul style="list-style-type: none"> - 0,738g/L de NaHCO₃ - 0,04g/L de K₂HPO₄. - 0,036 g/L de CaCl₂ · 2 H₂O - 0,075 g/L de MgSO₄ · 7 H₂O - 0,02 g/L de Na₂CO₃ Suplementado con 1 mL/L de solución basal de solución de micronutrientes

Tabla S. 3: Composición micronutrientes suplementados a medio BG11

Micronutrientes
<ul style="list-style-type: none"> - 6 g/L de citrato férrico amónico - 6 g/L de ácido cítrico - 1 g/L de EDTA, - 2,86 g/L de H₃BO₃ - 1,81 g/L de MnCl₂ · 4 H₂O - 0,22 g/L de ZnSO₄ · 7 H₂O - 0,39 g/L de Na₂MoO₄ · 2 H₂O - 0,079 g/L de CuSO₄ · 5 H₂O - 0,0494 g/L de CoCl₂

ANEXO.2: REAL TIME RT-PCR

Tabla S. 4: Oligonucleótidos usados para la amplificación y temperatura de hibridación

Oligonucleótido	Secuencia génica	Tipo de amplificación
sjdR_fw	ACAAATTCTCAGCCACAGGTTAAT	Amplificación de un fragmento interno del gen <i>sjdR</i>
sjdR_rv	TGAGGAGTAATAGCAGGTGCTGTCT	
tonB2_fw	CGATCGCTGCAAGGTAAA	Amplificación de un fragmento interno del gen <i>ionB2</i>
tonB2_rv	GGTGACTGGTGCTGCTAAACC	
oprB_I_fw	CCAATGGGACATACCGTGGTA	Amplificación de un fragmento interno del gen <i>oprB_I</i>
oprB_I_rv	AGCCTGCGGCAAATTCATAT	
ndhO_fw	CGAGATGGTTCGCGCTATC	Amplificación de un fragmento interno del gen <i>ndhO</i>
ndhO_rv	TCGCTGGCTTTTCGCTTCTAC	

Programa usado para real time RT-PCR y tratamiento matemático de los datos para determinar el cambio en la expresión génica.

El programa utilizado para determinar el cambio en la expresión génica fue *QuantStudio 5 Real-Time PCR software*.

- 1- Una vez obtenidos los datos, se revisan los valores de Ct obtenidos, ya que si hay un Ct mayor de 30, no puede darse por válido al estar fuera del límite de sensibilidad. Tras ello se realiza la media de esos datos.
- 2- Cálculo del ΔCt a partir de cada valor de nuestro gen a estudiar respecto al gen *housekeeping*. Para ello, se realiza la siguiente operación, con la media se consigue evitar emparejamientos arbitrarios:

$$\text{Ct del GEN A ESTUDIAR} - \text{MEDIA DE VALORES DE Ct de GEN HOUSEKEEPING}$$

- 3- Cálculo de $\Delta\Delta Ct$, mediante la siguiente operación:

$$\text{Ct DE MUESTRA A ESTUDIAR} - \text{MEDIA DE VALORES DE Ct de MUESTRA CONTROL}$$

- 4- Cálculo del *fold change*, como $2^{-\Delta\Delta Ct}$. En caso de que el valor obtenido sea inferior a uno, se calcula el inverso y se cambia de signo (-1/fold change), para así poder resaltar el carácter negativo del valor (infraexpresión con respecto al control)

Tabla S. 5: Programa utilizado para la Real Time RT-PCR

			40 ciclos		Curva de Melting		
Temperatura (°C)	50	95	95	60	95	60	95
Tiempo	2 min.	10 min.	15 s	1 min.	15 s	1 min.	15 s

ANEXO 3: Matrices de peso para la búsqueda de potenciales cajas FUR

FurA

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				

FurB

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

FurC

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				

Tabla S. 6:Resultado de las cajas FUR predichas para las cabezas de unidad transcripcional o para alguno de los genes contenidos, si existe. Además, se muestra la secuencia consenso y el p-valor de la predicción. N° hace referencia a número de cajas

	FurA			FurB			FurC		
	N°	Secuencia consenso	p-valor	N°	Secuencia consenso	p-valor	N°	Secuencia consenso	p-valor
<i>alr0894</i>	0	no FurA box at p-value <E-04	no box	1	TGATATTGGTAATCT	0,0000138	1	CAAAGTGAAAACCGCTTTA	7.41e-05
<i>alr0895</i>	1	AGAAATAAACTCATTTTCT	0,0000638	1	TGAGAACGTTTTTCG	0,0000448	0	no FurC box at p-value < E-03	no box
<i>alr0198</i>	3	ATATGAATACTCATAAAAA	0,0000378	0	no FurB box at p-value <E-04	no box	1	AATACTCATAAAAAAACCA	4.78e-05
<i>alr3415</i>	3	TAGTAATTTCTGAATAAAT	0,0000113	0	no FurB box at p-value <E-04	no box	0	no FurC box at p-value < E-03	no box
<i>all0802</i>	2	ATATGAAAATAGCTTAAAT	0,0000457	3	TAAAAATAATTATTT	0,0000826	0	no FurC box at p-value < E-03	no box
<i>alr1718</i>	0	no FurA box at p-value <E-04	no box	1	TGAAAATCTTTATAT	0,0000234	0	no FurC box at p-value < E-03	no box
<i>alr2270</i>	0	no FurA box at p-value <E-04	no box	2	TGTTATTCATAATTC	0,0000811	1	CATAGTCAATACGAATTGG	1.15e-06
<i>all0166</i>	1	AAAAATAGTTTGATTCTAC	0,0000552	0	no FurB box at p-value <E-04	no box	0	no FurC box at p-value < E-03	no box
<i>alr0248</i>	0	no FurA box at p-value <E-04	no box	0	no FurB box at p-value <E-04	no box	1	TTATCTGATAAAAAAATCG	9.2e-05
<i>asr4653</i>	1	ATAGAAAGCCTCCACAAGT	0,0000957	0	no FurB box at p-value <E-04	no box	0	no FurC box at p-value < E-03	no box

ANEXO 4: PREPARACIÓN DE EMSA

Tabla S. 7: Candidatos finales a selección para la realización de EMSA. No se incluye casilla para Transcriptoma de FurC, ya que no hay positivos

Genes	Caja FurA	Caja FurB	Caja FurC	Posible Operón	EMSA realiza-do	Trasncrip-toma FurA	Transcrip-toma FurB	Fold change bf
<i>all1939</i>				<i>all1939-all1940</i>				- 2,188
<i>alr0442</i>	X	X				X		- 2,493
<i>alr2271</i>				<i>alr2268-alr2271</i>		X		- 2,894
<i>alr3059</i>	X			<i>alr3057-alr3060</i>		X		- 2,150
<i>alr3060</i>	X			<i>alr3057-alr3061</i>				- 2,909
<i>alr3061</i>						X		- 2,829
<i>alr3062</i>								- 3,277
<i>alr3063</i>		X		<i>alr3063-alr3065</i>		X		- 3,015
<i>alr3064</i>	X		X	<i>alr3063-alr3065</i>				- 3,350
<i>alr3066</i>						X		- 3,031
<i>alr3067</i>								- 3,322
<i>alr3068</i>	X					X		- 3,370
<i>alr3074</i>	X		X	<i>alr3073-alr3074</i>		X		- 2,478
<i>alr3608</i>	X							- 2,286
<i>alr4629</i>	X					X		- 3,011
<i>all0475</i>	X		X	<i>all0475-all0476</i>		X		- 2,557
<i>all0802</i>	X	X						13,548
<i>all0875</i>		X			X	X		13,991
<i>all1471</i>		X		<i>all1471-1479</i>		X		- 156,381
<i>all2342</i>	X			<i>all2340-all2345</i>		X		2,560
<i>all2343</i>	X			<i>all2340-all2345</i>		X		4,411
<i>all3401</i>	X					X		- 2,079
<i>all3489</i>	X					X		- 6,373
<i>all3984</i>		X						4,714
<i>all4446</i>	X			<i>all4443-all4446</i>		X		2,435
<i>all4461</i>				<i>all4461-all4462</i>		X		- 2,678
<i>all4626</i>	X		X				X	7,332
<i>all4719</i>			X			X		- 2,707
<i>all4735</i>			X			X		- 2,016
<i>all4768</i>								4,840
<i>all4782</i>						X	X	17,280
<i>all7367</i>				<i>all7367/7370</i>				3,410
<i>alr0198</i>	X		X	<i>alr0198/alr0199</i>		X	X	23,326
<i>alr0600</i>						X		2,467
<i>alr0834</i>	X	X			X	X	X	- 3,609
<i>alr0896</i>				<i>alr0896-alr0898</i>			X	28,911
<i>alr0897</i>		X			X	X		18,165
<i>alr0898</i>						X	X	40,108
<i>alr1718</i>		X				X	X	10,438
<i>alr2270</i>		X	X	<i>alr2268-alr2271</i>		X		- 2,321
<i>alr3199</i>						X		24,784

Tabla S. 7: Continuación

Genes	Caja FurA	Caja FurB	Caja FurC	Posible Operón	EMSA realizado	Transcrip-toma FurA	Transcrip-toma FurB	Fold change bf
<i>alr3603</i>						X		- 2,340
<i>alr4490</i>		X		<i>alr4489-alr4494</i>		X		- 3,324
<i>alr4504</i>	X			<i>alr4504-4505</i>				2,540
<i>alr4505</i>	X			<i>alr4504-4505</i>				2,706
<i>qsl2047</i>						X		- 2,246
<i>asr0105</i>		X		<i>asr0105-asr0106</i>				4,636
<i>asr1134</i>							X	23,760
<i>asr3848</i>	X			<i>asr3843-asr3848</i>				- 2,556
<i>asr3935</i>								- 2,102
<i>asr4653</i>	X							11,243
<i>asr8513</i>	X					X		3,466

Tabla S.8: Oligonucleótidos usados para la amplificación de promotores y temperatura de hibridación

Oligonucleótido	Secuencia génica	Temperatura de hibridación (T _m)
Palr3415_fw	CTAGTTGGGCTAACTGTTGTAG	49 °C
Palr3415_rv	CCTACTTCGCAGTATATTGAGG	
Palr1718_fw	GAGCCATAACCGCAGATATC	49 °C
Palr1718_rv	CACGTTGCTTGTCTTCATCC	
Palr2270_fw	CAGAACTGCTGAAGTGCTG	49 °C
Palr2270_rv	GTACTGACACCACTATGCAG	
Pasr4653_fw	CTACAGTATCTATGACGGCTTC	49 °C
Pasr4653_rv	CATGAGATGGTTTACAACACTG	
Palr0198_fw	GCGGATAGCTGTCACTTCTAAC	49 °C
Palr0198_rv	CTACAGAGAGATAGGGCTTAAC	
Pall0166_fw	GTTGACCCTGACAACCGC	49 °C
Pall0166_rv	CCTAGTGGTGTGTTGTCCTG	
Palr0895_fw	CAAAGCACAATTCAGCAAGC	49 °C
Palr0895_rv	CACTTCAGTACCGCGTAGAC	

Programa utilizado para realización de PCR

- 1- Desnaturalización inicial del DNA: 5 min a 95 °C
- 2- Desnaturalización del DNA: 30 s a 95 °C
- 3- Hibridación de cebadores: 30 s a la temperatura de hibridación
- 4- Elongación: 45 s a 72 °C
- 5- Elongación final: 10 min a 72 °C
- 6- Conservación a 4 °C

Del paso 2 al paso 4, se repite durante 30 ciclos

ANEXO 5: NEGATIVOS PARA EMSA

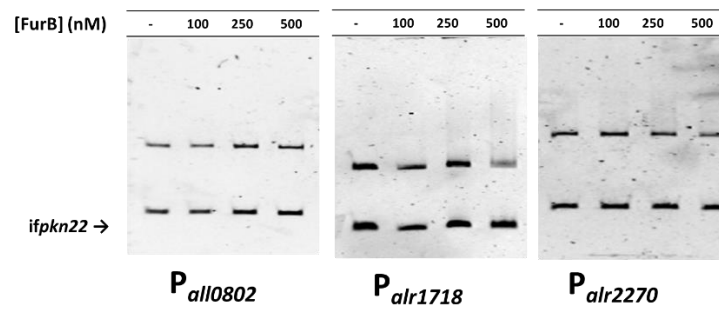


Figura S. 1: Ensayo de cambio de la movilidad electroforética llevado a cabo para evaluar la interacción *in vitro* entre FurB y regiones promotoras de una selección de genes. Todos ensayos fueron realizados con fragmentos de DNA libre o incubados con concentraciones crecientes de FurB (de 0 a 500 nM), separados en gel PAGE al 6%. El fragmento interno del gen *pkn22* fue usado como DNA competidor no específico.

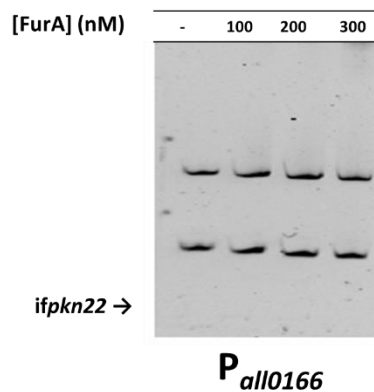


Figura S. 2: Ensayo de cambio de la movilidad electroforética llevado a cabo para testar la interacción *in vitro* entre FurA y regiones promotoras de una selección de genes. Todos los ensayos fueron realizados con fragmentos de DNA libre o incubados con concentraciones crecientes de FurA (de 0 a 300 nM), separados en gel PAGE al 6%. El fragmento interno del gen *pkn22*(*ifpkn22*) fue usado como DNA competidor no específico.

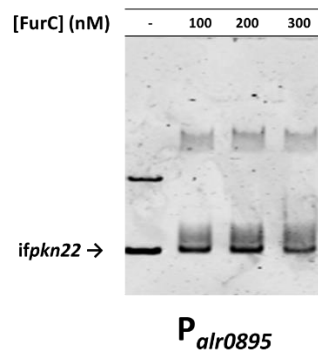


Figura S. 3: Ensayo de cambio de la movilidad electroforética llevado a cabo para testar la interacción *in vitro* entre FurC y regiones promotoras de una selección de genes. Todos los ensayos fueron realizados con fragmentos de DNA libre o incubados con concentraciones crecientes de FurC (de 0 a 300 nM), separados en gel PAGE al 6%. El fragmento interno del gen *pkn22*(*ifpkn22*) fue usado como DNA competidor no específico.