



UNIVERSITAT DE  
BARCELONA

## Nuevos aspectos de la regulación de la virulencia en cepas enteroagregativas de *Escherichia coli*

Alejandro Prieto Durán

**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tdx.cat](http://www.tdx.cat)) i a través del Dipòsit Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tdx.cat](http://www.tdx.cat)) y a través del Repositorio Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tdx.cat](http://www.tdx.cat)) service and by the UB Digital Repository ([diposit.ub.edu](http://diposit.ub.edu)) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITAT DE  
BARCELONA

Universitat de Barcelona, Facultat de Biologia  
Departament de Genetica, Microbiologia y Estadistica  
Seccion de Microbiologia, Virologia y Biotecnologia

**Nuevos aspectos de la regulaci3n de la virulencia en cepas  
enteroagregativas de *Escherichia coli***

Memoria presentada por Alejandro Prieto Dur3n para optar al grado  
de Doctor por la Universidad de Barcelona

Programa de Doctorado EEES (Espacio Europeo de Educaci3n Superior)  
de Biotecnologia

Visto bueno del director y tutor

Visto bueno del codirector

Prof. Antonio Ju3rez Gim3nez

Dr. M3rio H3ttener Queiroz

Memoria presentada por

Alejandro Prieto Dur3n

Barcelona, 2020



## Agradecimientos

Todavía no llego a comprender del todo bien el sentimiento de nostalgia que me invade al escribir estas líneas. Entiendo que el hecho de estar proyectando estas palabras supone cerrar una etapa de mi vida en la que, a pesar de tanto esfuerzo y sacrificio realizados, he sido tremendamente feliz. Quienes me conocen saben que soy una persona sencilla, tímida e introvertida, a la que no se le da nada bien expresar sus sentimientos. Utilizaré este pequeño apartado para dar voz a los mismos y con el fin de que, con el paso de los años, cuando vuelva a leer estas líneas, las palabras que las conforman me transporten a estos momentos de inmensa felicidad y, sobre todo, a las personas que lo hicieron posible. Por todo ello, a pesar de que los agradecimientos sean la parte más corta de la tesis, no es la menos importante. Más bien, todo lo contrario.

Toda esta etapa comenzó en el año 2010. Creo que, junto a mis compañeros y amigos que empezamos el grado de Bioquímica ese mismo año, íbamos viendo pasar las diferentes asignaturas de la carrera sin tener del todo claro que sería de nuestro futuro. En mi caso, esa situación cambió a mediados de 2012, por culpa (en el mejor sentido de la palabra), de la asignatura de Biología Celular y Molecular de los Microorganismos (BCMM).

Me fascinó que aquellos diminutos seres vivos, a pesar de ser de las formas de vida más “sencillas” que podamos encontrar en nuestro planeta, sean capaces de albergar en su interior mecanismos tan complejos que, a pesar de años y años de evolución científica y tecnológica, nunca llegaremos a comprenderlos a fondo. Da la triste casualidad, además, que estas líneas, y una buena parte de esta tesis doctoral, se han escrito mientras estamos viviendo la dura pandemia mundial del COVID-19, una muestra más de la fuerza de la naturaleza en su versión más microscópica.

Desde las prácticas de laboratorio de la asignatura BCMM (donde justo después comencé prácticas voluntarias en el departamento de Microbiología) hasta ahora, he tenido la suerte de cruzarme con muchas personas a las que quiero dedicarles unas palabras de agradecimiento por todo lo que me han aportado a lo largo de este tiempo.

En primer lugar, me gustaría dar las gracias al director de la tesis, Antonio Juárez, por dar la oportunidad a un chico de 20 años (todavía imberbe y con menos de media carrera universitaria finalizada) de poder empezar a trabajar en su laboratorio. Ha sido una experiencia muy gratificante y enriquecedora, en la que he aprendido y he crecido muchísimo tanto a nivel científico como a nivel personal. A pesar de estar en mil frentes, siempre me he sentido muy apoyado en la dirección y gestión de esta tesis, con consejos y opiniones constructivas, siempre guiando y nunca imponiendo criterios. Agradecerle también por todos los esfuerzos realizados para mantener el grupo de investigación en las mejores condiciones posibles con constante búsqueda de financiación externa.

A continuación, me gustaría agradecer al codirector de la tesis, Mário Hüttener, todos estos años vividos juntos. Él sabe más que nadie todo el esfuerzo y horas de dedicación que ha significado esta tesis doctoral. Muchísimas gracias por acompañarme desde el trabajo de final de grado (en el que no tuvimos mucha suerte con las toxinas GhoS y



GhoT) hasta estos momentos. Toda esta etapa ha sido un proceso de mucho trabajo, pero de un continuo aprendizaje. Gracias por apoyarme en todo momento, por estar siempre disponible y dispuesto a ayudar, por tu paciencia y por tu gran sentido del humor, siempre presente y tan necesario.

Gracias a los dos por haber sido unos magníficos directores de este trabajo, y por permitir que todavía mantenga intacta la misma ilusión por la ciencia que el primer día (a pesar de las dificultades que todo ello conlleva en la actualidad en nuestro país).

Me gustaría agradecer también a Sonia Aznar por toda su inestimable ayuda prestada durante todos estos años (protocolos experimentales, pedidos, etc), a David, el más madrugador del departamento, por este último año tan divertido compartiendo poyata, experimentos y bandejas de dulces, y a Mari, por recibirnos siempre con los brazos abiertos cada vez que bajábamos al laboratorio del Parc Científic.

A Manu, gran compañero de fatigas trabajando codo con codo durante todos estos años, que además de ser un gran científico y el chico para todo, es un gran portero de balonmano, ya sea con el 4, el 13 o el 28 a la espalda. Muchísimas gracias por estos años en el laboratorio, si nos ponemos a sumar las horas totales que hemos pasado juntos... ¡mare!, jajajaja. Gracias por intentar poner un poco de orden en el caos de mi poyata, soportar los gritos sordos, canciones desafinadas, sonidos guturales a cualquier momento y a cualquier hora, por ser testigos coetáneos de la famosa inhalación de HCl y ser receptor del regalo del Picasso en forma de gel de agarosa, por ser espía y avisarme de lo que se avecinaba por el pasillo del departamento, por dejarme acompañarte como buen *stalker* cada vez que ibas a revelar un gel... en fin, mil historias divertidas. Ha sido un auténtico placer compartir todo este tiempo contigo. *Ayayayayay...*

Agradecer también a todos los estudiantes que han ido pasando a lo largo de todos estos años por el laboratorio 3. Ha sido un gran placer compartir este tiempo con vosotros. A la Blanca Margarit (per totes les incomptables PCRs que vam posar a punt tot just quan començava, però sobretot pel teu bon rotllo i la teva simpatia), al Joan Espelt, la Eleonora Tassinari, a la Vanessa Kùchler (*Vani Vani*, thanks a lot for the amazing year that we spend together in the lab, I hope everything is fine), al Enric Vera (por esos dos años tan divertidos que pasamos juntos, y por esos croissanitos de chocolate tan ricos que nos traías), a la Maria Bosch (per la teva simpatia i per que malgrat “sigue siendo un jefe serio” vam poder crear el laboratorio gozadera), a la Sara Moreno (no estic ben segur del tot, però em sona que eixa xiqueta era del Grau de Gandia i una gran fallera... moltes gràcies per aquest últim any, espere que et vaja molt bé a la teua futura tesi doctoral) a Carmen, (*grazie mille per tutto e per avermi insegnato a migliorare il mio italiano*) y a Jon, un auténtico crack en todos los sentidos.

A los Hulios, muchas gracias por el grupazo (inclusivos pero intolerantes) que hemos formado durante todo este tiempo. Vosotros más que nadie sabéis cuanto esfuerzo y sacrificio supone nuestro trabajo en el laboratorio día a día, pero gracias a vosotros todo se ha hecho mucho más llevadero. Cada uno de vosotros sois uno de los principales motivos que ha hecho que esta etapa, de la que me siento tan orgulloso, haya valido tanto

la pena. Vosotros también formáis parte de este trabajo. Por todos los innumerables momentos divertidos que hemos compartido juntos (casas rurales, cenas, comidas, cumpleaños, amigos invisibles, *xconfessions*, juegos de los papelitos super originales sin repeticiones de palabras...). Gracias y más gracias infinitas de todo corazón a Sandra, Eva, Sara, Pedro, Gala, Manu, Genoveva, Cris, Álvaro, Óscar y Sergio. ¡Sois geniales!

També m'agradaria donar les gràcies a l'altra part del laboratori 3 (Edu i Gemma), que malgrat vagin perdent poc a poc el seu territori en partits de pàdel, són molt bona gent i els riures estan més que garantits amb ells...*pon la Tusa!* Y también agradecer a todos los otros compañeros que ya marcharon, Fran, Lucía, Montse, Aurora y Albert Jr.

A nivel de departamento, también me gustaría dar las gracias a todos los compañeros que nos hemos ido cruzando cada día por los pasillos y zonas comunes (estudiantes de grado, de máster, doctorandos, *postdocs*, técnicos...) y a todo el personal de secretaría (Manolo, Bea, Rosario, Susana y Mónica) por su amabilidad e inestimable ayuda, siempre con una sonrisa.

A Da, (Dalilili), mai oblidaré quan vam començar junts aquesta aventura al 2012! Ha plogut molt des de que tractaves d'identificar la funció d'aquells gens putatius a *Salmonella*. Després els nostres camins es van separar, però de ben segur que aquella experiència al laboratori va contribuir al teu creixement orgànic infinit! Malgrat la distància física que ha hagut aquests últims anys, cada cop que ens hem retrobat a qualsevol punt del món (Milà, Frankfurt o Sri Lanka) la confiança mai ha decaïgut, i aquí seguim, com al primer dia. *Kala kala, kele kele, kili kili!*

A Óscar Zhu, ¡¿*Doctol ya?! Mi biólogo molecular de plantas, microbiólogo, biotecnólogo molecular y profesor favorito... ¡4 en 1! Muchas gracias por todos estos años de amistad... *joteljoteljotel*. Agradeceros a ti y a Marta por todos los momentos vividos y todos los que vendrán en el futuro.*

A l'Estelita, (*testis* :) moltes gràcies per la teva simpatia i generositat sempre, i per aquests meravellosos viatges que hem fet durant aquests anys (*Pão Bimbao* per Portugal i *Baguette Trip* als Midis Pyrénées).

Al Josep Tarragó, sempre li dic, però vull que quedi plasmat fins l'eternitat... ets molt gran, Josep!!! Moltes gràcies per tots aquests esmorzars "quasi" diaris durant aquests anys de tesi. Ha estat un gran plaer començar cada dia de laboratori amb tu amb un cafè al *Penya*.

A la Cris, pel seu somriure sempre present, senzillesa i simpatia, i tots els moments viscuts (viatges a Cantabria, França, la mítica nit al casino de Montecarlo sopant entre pans cutres entre la *jet-set*, a Grècia, a Cinque Terre...).

A la Nuri, la definició de la paraula "intensidad" hecha persona, en cualquiera de sus múltiples versiones y actualizaciones. Muchas vivencias y momentos que hemos compartido juntos. Muchas gracias por enseñarme (bajo ningún tipo de coacción) la vía de *wnt* en la remodelación del hueso y que la osteoporosis es una enfermedad multifactorial compleja. Por toda nuestra confianza y cariño.

A la Laura Pineda (Pine o Pinedeta), moltíssimes gràcies per tot el teu suport incondicional i pel teu caràcter tan afable i divertit, que genera tantes situacions còmiques que tant em fan riure. Gràcies per la teva bondat, per tots els viatges que hem fet i tots els moments viscuts. Sempre has sigut i seràs una persona molt important per mi (i no gràcies a la teva memòria precisament jajaja).

I a la resta de Bioqninges, Borja, Samantha, Pau, Negre... malgrat que ens veiem poc, sempre és un plaer recordar tots els moments que hem viscut junts.

A Mario y Bruno, muchos años ya a las espaldas... creo que hasta *suenan Michel*... A Javi, por las comidas y charlas en la facultad de Física, y al resto de integrantes VTP remanentes del Xaloc. Edu, muchas gracias por traerme siempre a casa y las divertidas conversaciones posteriores sobre la vida, y como no a Mateo, Lluís (*shshsh*), Arturo y Melbu.

A *iPAD-Dell* (Cristóbal, Bruno y Penas) por los partidos de pádel y *volley* playa utilizados como excusa para lo que de verdad importa, las actualizaciones de carpesano.

A Miriam, moltes gràcies per tota la confiança generada en aquests últims anys malgrat la teua tímidesa (*ejem ejem*). Per la teua simpatia, per estar sempre disposada a ajudar, per ensenyar-me moltes paraules i expressions alcoianes, i pels pantalons del Valencia Basket, un dels millors regals que m'han fet mai.

Agradecerle también, juntamente con Loli y Esther, el gran cariño recibido al permitir que un intruso les invadiera constantemente en Casanova 29 :)

A María, agradecerle todo el afecto recibido durante este tiempo. Pocas veces te cruzas en la vida con personas tan alegres, enérgicas y simpáticas. Gracias por haber conseguido, a pesar de mi gran timidez, haberme hecho sentir como en casa desde el primer momento que pisé Alcoy.

A Lirios, periodista con un futuro inmenso, muchas gracias por tu simpatía, por tu sonrisa y por apoyarnos mutuamente en la guerra contra la *tata* terremoto. Ah, y agradecer y felicitar también a *@Srta.freckles* el hecho de escribir tan bien.

También agradecerle a la abuelita Genoveva, por ser una persona tan buena, tierna, bondadosa y luchadora. Y como no a su *alter ego*, Chispita, la perrita más bonita e inteligente del mundo.

A Jordi, també conegut com l'organitzador (*consigo mismo*) de la mítica i inoblidable festa de la Gamba Roja. Moltes gràcies pel teu increïble sentit de l'humor. Segur que compartirem molts més esmorzars, dinars a la casita, etapes de la Volta Ciclista a Espanya, voltes en patineta elèctric per Mariola, visites i èxits de l'Alcoyano al Viejo Collao i molts més títols del millor equip de la història, el Real Madrid je je je.

Moltes gràcies també als abuelitos Luis y Pepita, per la vostra simpatia, vitalitat, sentit de l'humor, per aquests dinars familiars tant divertits (la increïble sepieta, els abisinios, la fideuà...) i per ensenyar-me els esmorzars on vaig poder conèixer el famós *tortillongui* alcoià.

Gracias de todo corazón a todos por haberme dejado construir un gran puente hasta Alcoy que ayuda a conectar todo, casa y familia, familia y casa :)

A la tita Mari (o *María Josefa de los Dolores*, como le gusta que le llamen), al tío Jose y a los “*prims*” Marta (a pesar de casi nunca coincidir en el autobús porque siempre se le quemaba la tostada, muchas gracias por los chicles mañaneros y todas tus historias con las que me muero de risa) y Álvaro (un auténtico crack aunque haya cambiado las magdalenas por la crema de cacahuete). Muchas gracias por vuestro apoyo, cariño y por estar siempre ahí, junto con la alegría y vitalidad del avi *Lelo* y la Juani. Os quiero mucho.

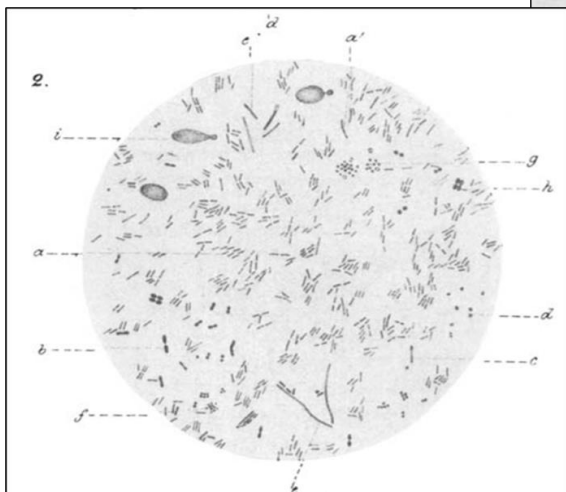
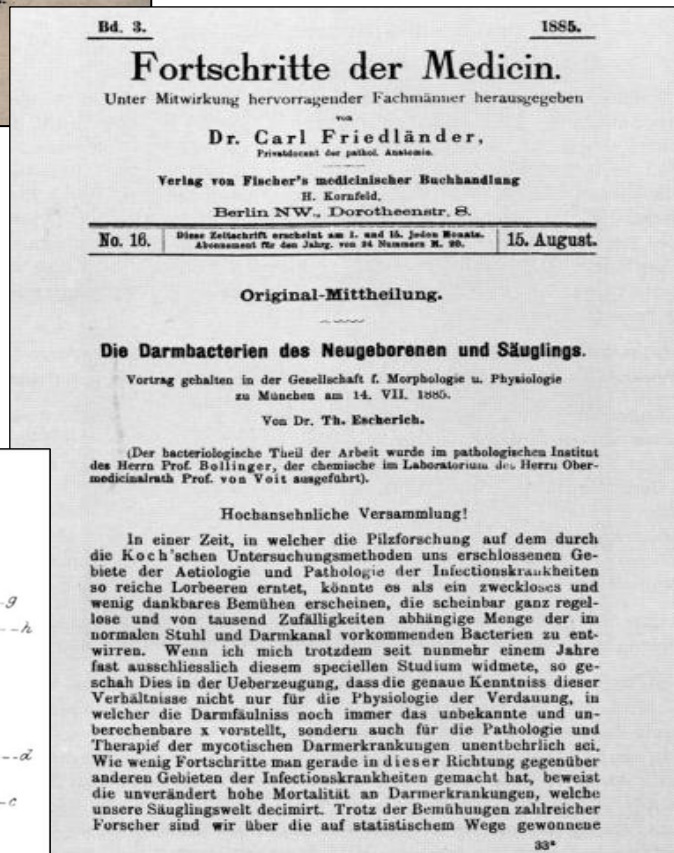
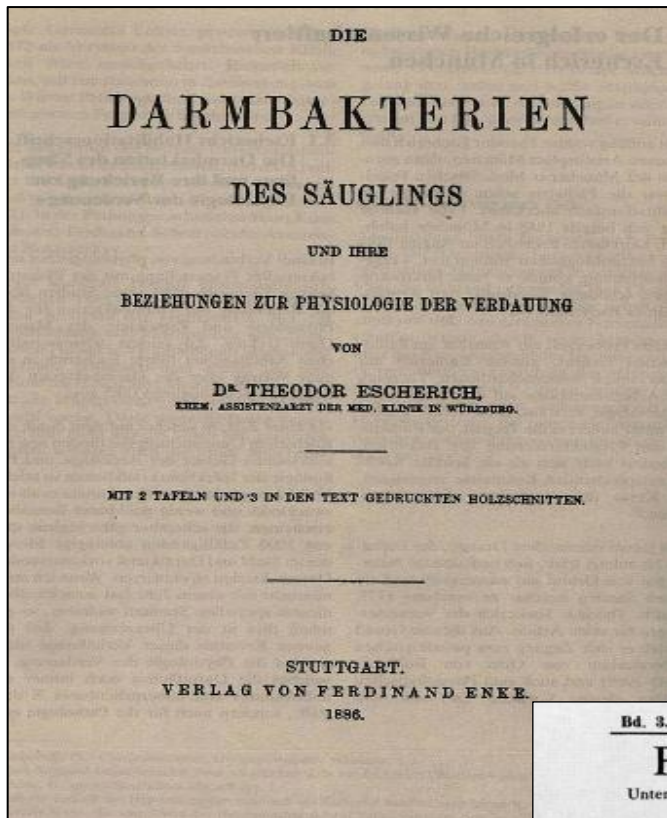
A Iván y Cristina, *tete* y *cuñi*, agradecerles por su bondad, todo vuestro cariño y apoyo continuo, y esas cenas tan ricas que nos preparan :)

Para finalizar, me gustaría agradecer absolutamente todo el apoyo incondicional a mis padres. Por cuidarme tanto y tan bien, por hacerme la vida tan fácil. Significáis absolutamente todo para mí.

Por último, darle las gracias al espectrofotómetro de la cocina de fase 1 del departamento de Microbiología, por actuar como la perfecta *Celestina* que me permitió conocer a Genoveva. Muchas gracias por tu gran vitalidad y tu gran sentido del humor. Por convertir cualquier problema ajeno en algo tuyo y luchar como la que más por solucionarlo. Gracias por enseñarme el significado de la palabra *desfici*. Pero, sobre todo, y a pesar de que la vida no te lo haya puesto nada fácil, por ser una luchadora incansable, tan buena persona y llenar mi vida de tanta felicidad y alegría. Eres increíble en todos los sentidos. Te quiero muchísimo, *job*.

Dentro de unos años, cuando vuelva a leer estas líneas, sus recuerdos me transportarán a estos momentos. Y estoy seguro de que se me escapará una gran sonrisa, y porque no, alguna que otra lagrimilla. Pero todo ha valido mucho la pena. Mucho. Gracias a cada uno de vosotros por hacerme la vida más dulce. Y ya sabéis lo que dicen, a pesar de las dificultades que nos encontremos en el camino... *love for everyone*.

Y ahora, turno de tratar de aportar un granito más de arena a la investigación que Theodor Escherich inició por el 1885...



La presente tesis doctoral ha estado financiada mediante la beca Ayuda del Programa de Formación de Profesorado Universitario FPU15/04398 del Ministerio de Educación, Cultura y Deporte del Gobierno de España.



## Sinopsis

The genomes of Gram-negative bacteria encode paralogues and/or orthologues of global modulators. The nucleoid-associated proteins H-NS and Hha are an example: several enterobacteria such as *Escherichia coli* or *Salmonella* harbor H-NS, Hha and their corresponding paralogues, StpA and YdgT proteins, respectively. Remarkably, the genome of the pathogenic enteroaggregative *E. coli* strain 042 encodes, in addition to the *hha* and *ydgT* genes, two additional *hha* paralogues, *hha2* and *hha3*. We show in this work that there exists a strong correlation between the presence of these paralogues and the virulence phenotype of several *E. coli* strains. *hha2* and *hha3* predominate in some groups of intestinal pathogenic *E. coli* strains (enteroaggregative and Shiga toxin-producing isolates), as well as in the widely distributed extraintestinal ST131 isolates. Because of the relationship between the presence of *hha2/hha3* and some virulence factors, we have been able to provide evidence for Hha2/Hha3 modulating the expression of the antigen 43 pathogenic determinant. We show that tracking global modulators or their paralogues/orthologues can be a new strategy to identify bacterial pathogenic clones and propose PCR amplification of *hha2* and *hha3* genes as a virulence indicator in environmental and clinical *E. coli* isolates.

The enteroaggregative *E. coli* strain 042 also carries, in addition to the *hns* and *stpA* genes, a third gene encoding an *hns* paralogue (termed *hns2*). This novel H-NS paralogue is widespread within pathogenic strains of the *Enterobacteriaceae* bacterial family. We present information about the biological function of this novel H-NS paralogue. Compared with H-NS, H-NS2 expression levels are lower and, in addition, its expression exhibits specific features: it is sensitive to the growth temperature and to the nature of the culture medium. Transcriptomic analysis has shown that the H-NS2 protein targets a subset of the genes targeted by H-NS. Genes targeted by H-NS2 correspond mainly with horizontally transferred (HGT) genes and are also targeted by the Hha protein, a fine-tuner of H-NS activity. Growth temperature significantly influences the effect of the H-NS2 protein on the expression of *E. coli* 042 genes. In addition, we also show that H-NS2 expression can be modified by a reversible C-T transition in the promoter region of the *hns2* gene, thus facilitating that increased H-NS2 levels can compensate for H-NS loss.



We also assessed the role of virulence plasmid pAA2 in the pathogenesis of *E. coli* 042, with special emphasis on the role of the AggR transcriptional regulator. To date, the existing information on the role of this regulator on the virulence of the enteroaggregative strains of *E. coli* had been obtained mainly by analyzing the effect of mutations in the *aggR* gene on the overall expression pattern of the 042 strain. Our interest in analyzing the interactions of the pAA2 plasmid with the regulatory networks of the *E. coli* 042 strain led us to observe an unexpected phenotype, which has allowed us to better understand the role of the AggR regulator. We found that insertion of a genetic sequence in the 3'UTR region of the *aggR* gene or in the IS element located downstream it results in a notable increase of AggR expression, generating a phenotype that is different from that of the wild type strain (decreased growth rate, increased cellular autoaggregation, increased bacterial motility, increased frequency of pAA2 plasmid conjugation and higher plasmid instability). We also show that higher amounts of AggR protein are correlated with overexpression of genes involved in several metabolic pathways (arginine degradation and  $\beta$ -oxidation of fatty acids), whose products are involved in immunomodulation processes and bacterial pathogenesis in the intestinal tract. We hypothesize that these factors can be related to the ability of strain 042 to colonize the gut.

These results demonstrate that in the 3'UTR region of *aggR* gene are located important genetic sequences involved in post transcriptional regulation of the *aggR* gene expression levels.

# ÍNDICE

<b>1. INTRODUCCIÓN</b> .....	1
1.1 <i>Escherichia coli</i> .....	3
1.1.1 Patotipos diarreicos de <i>E. coli</i> .....	6
1.1.1.1 <i>E. coli</i> enteropatógena .....	7
1.1.1.2 <i>E. coli</i> enterohemorrágica .....	8
1.1.1.3 <i>E. coli</i> enteroinvasiva .....	11
1.1.1.4 <i>E. coli</i> enterotoxigénica .....	12
1.1.1.5 <i>E. coli</i> con adherencia difusa .....	13
1.1.1.6 <i>E. coli</i> adherente invasiva .....	14
1.1.2 <i>E. coli</i> enteroagregativa y su cepa prototipo 042 .....	15
1.1.2.1 Factores de virulencia de la cepa <i>E. coli</i> 042: sistemas de secreción y toxinas entéricas .....	18
1.1.3 El plásmido pAA2 y el regulador transcripcional AggR .....	21
1.1.4 Un ejemplo de la heterogeneidad genómica de las cepas EAEC: el brote en 2011 de SUH en Alemania .....	29
1.1.5 Genómica y filogenia de <i>E. coli</i> .....	32
1.2 Proteínas asociadas al nucleoide como moduladores globales de la expresión génica bacteriana .....	36
1.2.1 La proteína H-NS ( <i>Histone-like nucleoid-structuring protein</i> ) .....	38
1.2.1.1 Dominios estructurales de la proteína H-NS .....	40
1.2.1.1.1 Dominio N-terminal .....	40
1.2.1.1.2 Dominio C-terminal .....	42
1.2.1.2 Oligomerización de la proteína H-NS .....	42
1.2.1.3 Unión al ADN por parte de H-NS: efectos topológicos .....	44
1.2.1.4 Regulación de la expresión génica por parte de la proteína H-NS .....	48
1.2.1.5 H-NS como represor de la expresión no deseada de ADN xenogénico ..	52
1.2.1.6 Mecanismos de contra silenciamiento de la expresión génica .....	54
1.2.1.7 Efecto de la proteína H-NS sobre la transcripción espuria .....	57
1.2.1.8 La proteína StpA y otros parálogos de H-NS .....	58
1.2.2 Familia de proteínas tipo Hha .....	61
1.3 Regulación post-transcripcional por secuencias 3' UTR en bacterias .....	68
<b>2. OBJETIVOS</b> .....	3
<b>3. MATERIALES Y MÉTODOS</b> .....	75
3.1 Cepas y plásmidos .....	77
3.2 Medios de cultivo .....	81

## ÍNDICE

3.3 Métodos de esterilización .....	82
3.4 Antibióticos.....	82
3.5 Mantenimiento de los microorganismos.....	83
3.6 Métodos de transferencia génica.....	83
3.6.1 Transformación de células competentes mediante tratamiento en frío con CaCl <sub>2</sub> y posterior choque térmico .....	83
3.6.2 Electroporación .....	84
3.7 Técnicas experimentales con ADN.....	84
3.7.1 Amplificación de ADN mediante la reacción en cadena de la polimerasa (PCR).....	84
3.7.2 Purificación de fragmentos de ADN amplificados por PCR.....	86
3.7.3 Aislamiento de ADN plasmídico .....	86
3.7.4 Cuantificación de ácidos nucleicos .....	87
3.7.5 Electroforesis de ADN en geles de agarosa .....	87
3.7.6 Secuenciación de fragmentos de ADN.....	88
3.7.7 Mutagénesis bacteriana: inactivación de genes cromosómicos utilizando fragmentos de PCR (reemplazamiento alélico).....	88
3.7.8 Construcción de fusiones de genes cromosómicos y plasmídicos con la secuencia codificante para el epítipo 3XFLAG .....	92
3.7.9 Construcción de fusiones de genes cromosómicos con el gen reportero <i>lacZ</i> mediante la recombinasa FLP .....	93
3.7.10 Clonaje de genes en el vector de sobreexpresión pLATE.....	94
3.7.11 Clonaje de genes en el vector pLG338.30 para complementar una mutación determinada en <i>trans</i> .....	96
3.7.12 Construcción de fusiones transcripcionales con el gen reportero <i>lacZ</i> en el plásmido pFZY1.....	97
3.8 Técnicas experimentales con ARN.....	98
3.8.1 Aislamiento de ARN .....	98
3.8.2 <i>Real time</i> qRT-PCR.....	99
3.8.3 5'RACE ( <i>rapid amplification of cDNA ends</i> ).....	100
3.8.4 <i>Northern Blot</i> .....	102
3.8.5 RNA-seq.....	104
3.9 Técnicas experimentales con proteínas.....	105
3.9.1 Obtención de extractos proteicos totales .....	105
3.9.2 Análisis del secretoma bacteriano .....	106
3.9.3 Fraccionamiento celular .....	106
3.9.4 Electroforesis de proteínas en geles de poliacrilamida en presencia de SDS .....	107
3.9.5 Detección de proteínas mediante tinción con Azul de Coomassie.....	108

3.9.6 Detección de proteínas mediante tinción de plata.....	109
3.9.7 Inmunodetección de proteínas específicas por <i>Western Blot</i> .....	109
3.9.8 Sobreexpresión y purificación de proteínas .....	111
3.9.9 Identificación de proteínas mediante LC-MS/MS .....	112
3.9.10 Aislamiento de flagelos.....	112
3.9.11 Ensayos de estabilidad proteica <i>in vivo</i> .....	113
3.10 Ensayos fisiológicos .....	113
3.10.1 Curvas de crecimiento y recuentos de viables .....	113
3.10.2 Ensayos de autoagregación .....	114
3.10.3 Ensayos de motilidad .....	114
3.10.4 Ensayos de resistencia al choque ácido.....	115
3.10.5 Ensayos de resistencia al choque frío.....	115
3.10.6 Ensayos de persistencia bacteriana ante concentraciones letales de antibiótico.....	115
3.10.7 Experimentos de conjugación bacteriana en medio líquido.....	116
3.10.8 Curado del plásmido pAA2.....	116
3.10.9 Ensayos de formación de <i>biofilm</i> .....	117
3.10.10 Ensayos de análisis de actividad $\beta$ -galactosidasa.....	118
<b>4. RESULTADOS</b> .....	119
4.1. Los genes <i>hha2</i> y <i>hha3</i> , nuevos parálogos del gen <i>hha</i> en la cepa <i>E. coli</i> 042 .	121
4.1.1 Identificación de nuevos parálogos del gen <i>hha</i> en el genoma de la cepa <i>E. coli</i> 042 .....	122
4.1.2 Distribución de los parálogos <i>hha2/hha3</i> entre diferentes cepas de <i>E. coli</i> .	123
4.1.2.1 Distribución de los alelos <i>hha2/hha3</i> entre cepas enteroagregativas de <i>E. coli</i> .....	126
4.1.2.2 Distribución de los genes <i>hha2/hha3</i> entre cepas de <i>E. coli</i> productoras de beta-lactamasas de amplio espectro.....	129
4.1.2.3 Distribución de los alelos <i>hha2/hha3</i> entre la colección de cepas de <i>E. coli</i> ECOR .....	130
4.1.2.4 Distribución de los genes <i>hha2/hha3</i> entre cepas ambientales de <i>E. coli</i> .....	131
4.1.3 Análisis genómico de la distribución de los genes <i>hha2</i> y <i>hha3</i> en <i>E. coli</i> mediante herramientas bioinformáticas .....	133
4.1.4 Papel regulador de las proteínas Hha, Hha2 y Hha3 sobre los genes <i>agn43</i> y <i>shf</i> .....	136
4.1.5 Perfil de expresión de los genes <i>hha</i> , <i>hha2</i> y <i>hha3</i> .....	139
4.1.6 Análisis del perfil de interacción proteica de Hha2 con otras proteínas de la cepa <i>E. coli</i> 042 .....	141

4.1.7 Efectos de las mutaciones en los genes <i>hha</i> y <i>hha2</i> sobre el perfil transcriptómico global de la cepa <i>E. coli</i> 042 .....	145
4.2 Descripción de un nuevo parálogo del gen <i>hns</i> en la cepa <i>E. coli</i> 042: el gen <i>hns2</i> .....	152
4.2.1 Identificación de un nuevo parálogo del gen <i>hns</i> en el genoma de <i>E. coli</i> 042 .....	153
4.2.2 Distribución del gen <i>hns2</i> en enterobacterias.....	155
4.2.3 Obtención de mutantes de los genes <i>hns</i> y <i>hns2</i> en la cepa <i>E. coli</i> 042.....	158
4.2.4 Efecto de los genes <i>hns</i> y <i>hns2</i> sobre la cinética de crecimiento de la cepa <i>E. coli</i> 042 .....	160
4.2.5 Análisis de la capacidad de complementación de la mutación <i>hns</i> por parte del gen <i>hns2</i> .....	161
4.2.6 Análisis del perfil de expresión del gen <i>hns2</i> y de su producto génico .....	162
4.2.7 Perfil de interacción de la proteína H-NS2 con otras proteínas de la cepa <i>E. coli</i> 042 .....	173
4.2.8 Efectos de las mutaciones en los genes <i>hns</i> y <i>hns2</i> sobre el transcriptoma global de la cepa <i>E. coli</i> 042 .....	181
4.2.8.1 Validación de los datos RNA-seq mediante qRT-PCR.....	186
4.2.8.2 Efecto los genes <i>hns</i> y <i>hns2</i> sobre el control de la transcripción espuria .....	188
4.2.9 Validación a nivel fenotípico de los datos obtenidos por RNA-seq sobre las cepas WT, $\Delta hns$ , $\Delta hns2$ y $\Delta hns\Delta hns2$ .....	189
4.2.9.1 Efecto de los genes <i>hns</i> y <i>hns2</i> en la resistencia al estrés por disminución del pH .....	189
4.2.9.2 Efecto de los genes <i>hns</i> y <i>hns2</i> en la persistencia a antibióticos .....	190
4.2.9.3 Efecto de los genes <i>hns</i> y <i>hns2</i> en la resistencia al choque frío .....	192
4.2.10 Identificación de mutaciones en el gen <i>hns2</i> que aumenten la capacidad de compensar la pérdida de H-NS.....	193
4.2.10.1 Experimento de evolución dirigida sobre un mutante <i>hns</i> en la cepa <i>E. coli</i> 042.....	197
4.2.11 Efecto de la transición C→T en el promotor del gen <i>hns2</i> sobre la expresión de su producto génico.....	203
4.2.12 Efecto de la transición C→T sobre el promotor del gen <i>hns2</i> : determinación del inicio de transcripción del gen <i>hns2</i> .....	206
4.2.13 Efecto de la dosis génica del gen <i>hns2</i> al complementar la mutación $\Delta hns$ .....	207
4.2.14 Efecto de la temperatura sobre el papel regulador de las proteínas H-NS y H-NS2.....	210
4.3 Nuevos datos del papel del plásmido pAA2 y del regulador AggR en la virulencia de la cepa <i>E. coli</i> 042 .....	218

4.3.1 Curado del plásmido pAA2. Efecto sobre el transcriptoma global de la cepa <i>E. coli</i> 042 .....	219
4.3.2 El gen <i>aggR</i> y su región 3'UTR: regulación de la virulencia en la cepa <i>E. coli</i> 042.....	226
4.3.2.1 Expresión transcripcional del gen <i>aggR</i> .....	226
4.3.2.2 Análisis del perfil de expresión de la proteína AggR.....	229
4.3.2.3 Alteración del fenotipo bacteriano al generar una inserción nucleotídica entre el codón de finalización de traducción y la región 3'UTR del gen <i>aggR</i> . 230	
4.3.2.3.1 Análisis del contenido proteico intracelular de la cepa <i>E. coli</i> 042 <i>aggR</i> + <i>FRT3</i> ' <i>UTR</i> .....	233
4.3.2.3.2 Análisis del secretoma de la cepa <i>E. coli</i> 042 <i>aggR</i> + <i>FRT3</i> ' <i>UTR</i> ... 236	
4.3.2.4 Análisis del transcrito del gen <i>aggR</i> mediante <i>Northern Blot</i> .....	238
4.3.2.5 Análisis genómico de la región 3'UTR del gen <i>aggR</i> .....	241
4.3.2.6 Análisis de la estabilidad del fenotipo generado al alterar la región 3'UTR del gen <i>aggR</i> .....	248
4.3.2.7 Efectos del plásmido pAA2 y del gen <i>aggR</i> en la formación de <i>biofilm</i> por parte de la cepa <i>E. coli</i> 042 .....	252
4.3.2.8 Efecto de la inserción nucleotídica en el extremo 3'UTR del gen <i>aggR</i> sobre el transcriptoma global de la cepa <i>E. coli</i> 042 .....	255
4.3.2.9 Correlación entre los datos transcriptómicos obtenidos vía RNA-seq y el fenotipo observable de la cepa <i>E. coli</i> 042 <i>aggR</i> + <i>FRT3</i> ' <i>UTR</i> .....	264
4.3.2.9.1 Análisis de la motilidad bacteriana.....	264
4.3.2.9.2 Análisis de la síntesis flagelar.....	266
4.3.2.9.3 Análisis de la frecuencia conjugativa del plásmido pAA2 .....	268
<b>5. DISCUSIÓN</b> .....	271
5.1 Los parálogos <i>hha2</i> y <i>hha3</i> en <i>E. coli</i> .....	273
5.2 H-NS2, un tercer parálogo de H-NS en la cepa <i>E. coli</i> 042.....	275
5.3 Importancia del plásmido pAA2 en la virulencia bacteriana y nuevos aspectos de la regulación del gen <i>aggR</i> : papel de la región 3'UTR.....	282
<b>6. CONCLUSIONES</b> .....	289
<b>7. BIBLIOGRAFÍA</b> .....	293
<b>8. ANEXO INTRODUCCIÓN</b> .....	333
A.I.1 <i>E. coli</i> enteropatógena .....	335
A.I.2 <i>E. coli</i> enterohemorrágica.....	338
A.I.3 <i>E. coli</i> enteroinvasiva .....	339
A.I.4 <i>E. coli</i> enterotoxigénica.....	340
A.I.5 <i>E. coli</i> con adherencia difusa.....	343
A.I.6 <i>E. coli</i> adherente invasiva.....	343

<b>9. ANEXO</b> .....	345
9.1 Capítulo 1: Los genes <i>hha2</i> y <i>hha3</i> , nuevos parálogos del gen <i>hha</i> en la cepa <i>E. coli</i> 042 .....	347
9.2 Capítulo 2: Descripción de un nuevo parálogo del gen <i>hns</i> en la cepa <i>E. coli</i> 042: el gen <i>hns2</i> .....	366
9.3 Capítulo 3: Nuevos datos del papel del plásmido pAA2 y del regulador AggR en la virulencia de la cepa <i>E. coli</i> 042 .....	381
<b>10. PUBLICACIONES</b> .....	385

# **1. INTRODUCCIÓN**



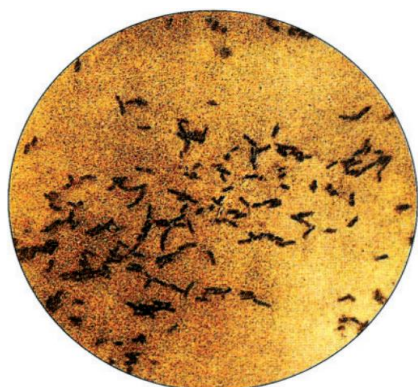


## 1.1 *Escherichia coli*

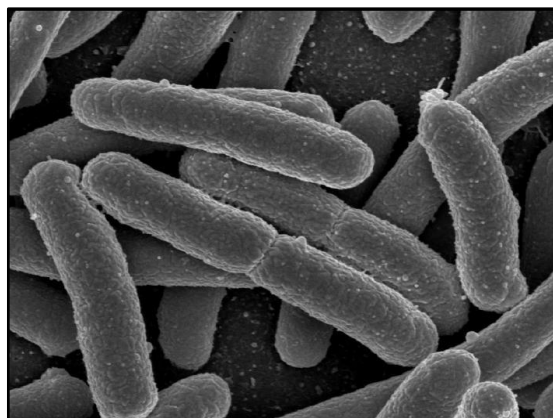
*Escherichia coli* (*E. coli*) es una especie bacteriana perteneciente a la familia *Enterobacteriaceae*, clasificada en la clase *gamma* englobada dentro del filo *Proteobacteria*. *E. coli* es una bacteria gramnegativa que presenta morfología en forma de bacilo y un tamaño de 0,5-1  $\mu\text{m}$  (figura 1.1), con un metabolismo anaerobio facultativo, no genera esporas y presenta motilidad a través de una flagelación peritrica. Es un microorganismo capaz de fermentar la lactosa, catalasa positivo y oxidasa negativo.

Las primeras anotaciones científicas de dicha bacteria datan de 1885, cuando el pediatra y microbiólogo alemán Theodor Escherich logró aislar a este microorganismo “de rápido crecimiento” durante un estudio sobre la relación entre el microbioma de un recién nacido y la fisiología de la digestión. Dicho microorganismo fue bautizado como “*Bacterium coli commune*”, conocido actualmente como *Escherichia coli* (Escherich, 1885; revisado en 1988).

A



B



**Figura 1.1.** (A) Imagen de microscopía óptica donde se representa la preparación y tinción con violeta de genciana en anilina acuosa de *Escherichia coli*. Imagen adaptada de (Escherich, 1885; revisado y adaptado en 1988). (B) Microfotografía realizada mediante microscopía electrónica de rastreo de un cultivo de *E. coli*. Fuente de la imagen: *Rocky Mountain Laboratories, NIAID, NIH* ([www.niaid.nih.gov/about/rocky-mountain-laboratories](http://www.niaid.nih.gov/about/rocky-mountain-laboratories)). En ambas imágenes se puede observar la morfología en forma de bacilo que presenta esta bacteria.

Debido a diferentes factores, tales como (i) la posibilidad de hacer crecer a esta bacteria de una forma rápida en el laboratorio bajo condiciones controladas de crecimiento, (ii) a que la cepa *E. coli* K-12 poseyese el plásmido conocido como factor F (Lederberg y col., 1952) y, (iii) la relativa facilidad de manipular dicha bacteria genéticamente y su elevada versatilidad, *E. coli* se ha convertido en el microorganismo más ampliamente estudiado,

## 1. INTRODUCCIÓN

especialmente desde el crecimiento exponencial de la genética bacteriana a partir de 1950 y de la biología molecular a partir de 1970. Toda esta investigación realizada sobre esta especie bacteriana ha permitido situar a *E. coli* en el epicentro como modelo de estudio tanto en el ámbito de la investigación básica (elucidación de la naturaleza del código genético (Crick y col., 1961), de la replicación (Lehman y col., 1958) y transcripción (Stevens y col., 1960) del ADN, los primeros modelos mecanísticos sobre la regulación de la expresión génica (Jacob y Monod, 1961) etc.) como en el ámbito de la investigación aplicada, utilizándose como herramienta para el desarrollo de múltiples aplicaciones biotecnológicas tales como la síntesis de proteínas recombinantes utilizadas en el ámbito farmacéutico (insulina, eritropoyetina, hormona de crecimiento, taxol, antibióticos, anticuerpos, aminoácidos, vitaminas, etc. (Kamionka, 2011)) y la síntesis de compuestos utilizados en la industria química (fenol, etanol, biocombustibles, etc. (Liu y Khosla, 2010; Hildebrand, 2013; Kim y col., 2014)). Además, debido a los grandes avances relacionados con la secuenciación masiva y el ensamblaje de genomas en la época post-genómica, disponemos de una notable cantidad de genomas secuenciados de este microorganismo. La comparación de los mismos ha permitido corroborar la gran diversidad genómica que presenta esta bacteria, con un núcleo de genoma (común a todas las cepas de *E. coli*) de algo más de 2.000 genes y un pangenoma de más de 16.000 genes (Kaas y col., 2012), aunque dichos valores (estimados) se incrementarán a medida que se secuencien más genomas de esta especie.

*E. coli* es uno de los microorganismos constituyentes del microbioma intestinal de los mamíferos (normalmente reside en las capas mucosas del ciego y del colon). Concretamente, las cepas comensales de *E. coli* representan entre el 0,1-1% de la biomasa de la microbiota humana (Balows y col., 1992; Eckburg y col., 2005). Este relativo bajo porcentaje es debido principalmente al hecho de que el intestino humano es un ambiente principalmente anóxico, en el que se favorece esencialmente el crecimiento de bacterias que presentan un metabolismo anaerobio estricto, tales como *Bacteroides* y *Firmicutes*, géneros que representan el 90% de la población microbiana intestinal (Eckburg y col., 2005; Claesson y col., 2009). *E. coli* también forma parte, aunque en una menor proporción que en el caso del microbioma humano, de la microbiota intestinal de pájaros, reptiles, peces (Hartl y Dykhuizen, 1984; Leimbach y col., 2013).

A pesar de que muchas de las cepas de esta especie bacteriana residen en el tracto gastrointestinal humano como bacterias comensales, algunos linajes de las mismas han

adquirido elementos genéticos adicionales que codifican para diversos factores de virulencia. Por ello, la cepa que los adquiere se convierte en un microorganismo patógeno para el ser humano y otros muchos animales domésticos (Leimbach y col., 2013). Esta elevada plasticidad genética ha provocado que *E. coli* sea una bacteria capaz de colonizar muchos huéspedes distintos y adaptarse a un gran número de condiciones de crecimiento diferentes, hecho que la ha convertido en uno de los mejores modelos para estudiar un amplio número de mecanismos evolutivos, adaptativos y patogénicos (Kawecki y col., 2013).

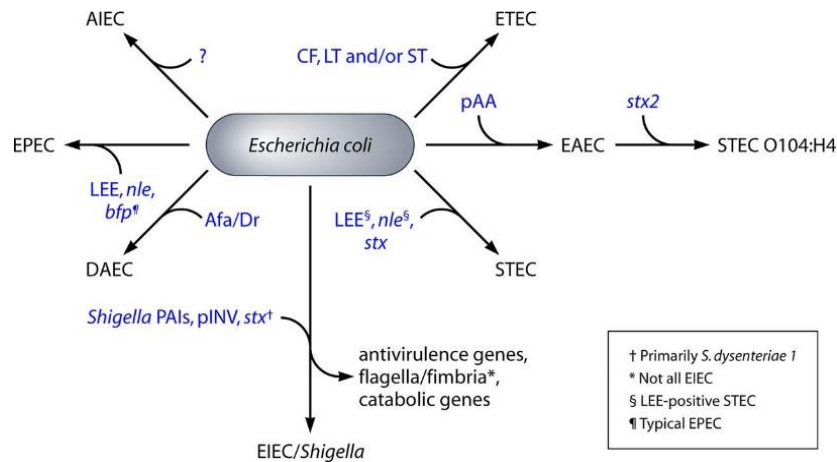
En función del tipo de infección que causan, las cepas patógenas de *E. coli* se agrupan en patotipos. Las infecciones causadas por los distintos patotipos de *E. coli* cursan con manifestaciones clínicas que abarcan tanto patologías extraintestinales (infecciones del tracto urinario, septicemia, meningitis, peritonitis, neumonía, etc.) como patologías intestinales, esencialmente diarreas agudas, aunque también son responsables de cuadros diarreicos persistentes e incluso crónicos.

Dentro del patotipo responsable de causar enfermedades extraintestinales (ExPEC, de *extraintestinal pathogenic E. coli*) se engloban los subtipos UPEC (*uropathogenic E. coli*), NMEC (*neonatal meningitis-associated E. coli*), SEPEC (*sepsis-associated E. coli*) y APEC (*avian pathogenic E. coli*) (Sarowska y col., 2019).

Por otro lado, los patotipos asociados con enfermedades intestinales se engloban dentro del grupo DEC (*diarrheagenic E. coli*). Dentro de dicho colectivo existen 7 patotipos, cada uno de ellos definido por sus propiedades de virulencia: *E. coli* enterotoxigénica (ETEC, *enterotoxigenic E. coli*), enterohemorrágica (EHEC, *enterohemorrhagic E. coli*), enteroinvasiva (EIEC, *enteroinvasive E. coli*), enteropatógena (EPEC, *enteropathogenic E. coli*), con adherencia difusa (DAEC, *diffusely-adherent E. coli*), adherente invasiva (AIEC, *adherent invasive E. coli*) y enteroagregativa (EAEC, *enteroaggregative E. coli*) (Nataro y Kaper, 1998). Las cepas pertenecientes a cada uno de estos patotipos de *E. coli* presentan características distintivas relacionadas con su perfil filogenético, su epidemiología, mecanismos de patogénesis, factores de virulencia y los cuadros clínicos generados y el tratamiento necesario para hacer frente a cada tipo de infección. Los factores de virulencia de las cepas patógenas de *E. coli* se encuentran codificados principalmente en elementos genéticos móviles tales como plásmidos, transposones, bacteriófagos e islas de patogenicidad, que presentan la capacidad de movilizarse entre diferentes cepas, generando de esta manera combinaciones específicas de elementos de

## 1. INTRODUCCIÓN

virulencia que han provocado la aparición (a lo largo de la historia evolutiva) de los diferentes patotipos diarreicos de *E. coli* mencionados anteriormente (figura 1.2).

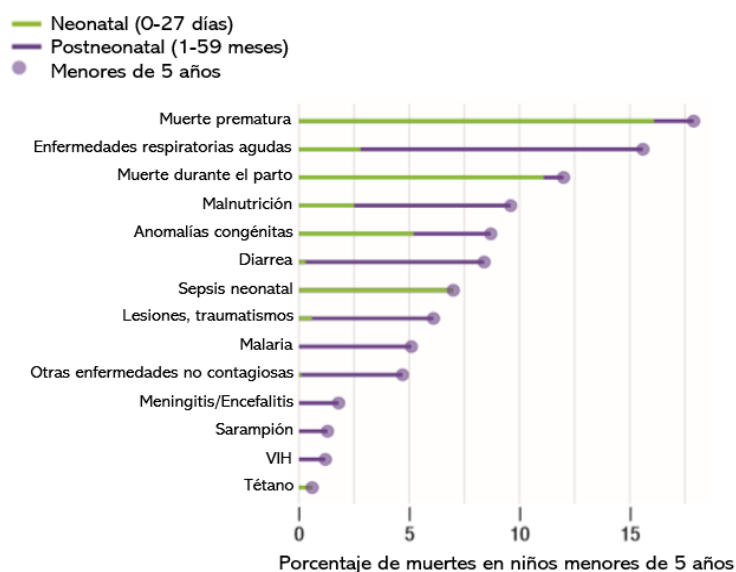


**Figura 1.2.** Esquema general de los procesos de adquisición de genes a través de mecanismos de transferencia horizontal génica que han dado lugar a la aparición de los diferentes patotipos diarreicos de *E. coli* a lo largo de la evolución. Los principales elementos móviles implicados en esta transferencia génica son transposones, profagos y plásmidos. Imagen adaptada de (Croxen y col., 2010).

### 1.1.1 Patotipos diarreicos de *E. coli*

Las enfermedades diarreicas son un problema de salud pública mundial, especialmente en los países en vías de desarrollo, siendo una de las principales causas de mortalidad en niños menores de cinco años de edad. *Escherichia coli* es, junto con rotavirus, uno de los principales agentes etiológicos responsables de estos cuadros diarreicos que provocan una gran tasa de mortalidad en países subdesarrollados (World Health Statistics, 2018) (figura 1.3).

Por lo general, estas diarreas son consecuencia de la exposición a alimentos o agua contaminados con heces humanas procedentes de aguas residuales, fosas sépticas o letrinas. Dentro de cada patotipo diarreico de *E. coli* se agrupan cepas que poseen algunos factores de virulencia comunes y causan enfermedades con una sintomatología y patología muy similares.



**Figura 1.3.** Principales causas de mortalidad en niños menores de 5 años en todo el mundo en 2016. Datos extraídos de (World Health Statistics 2018: Monitoring health for the SDGs).

En este primer apartado de la Introducción se refieren los diferentes patotipos de *E. coli*, con especial énfasis al patotipo enteroagregativo de *E. coli*, al que pertenece la cepa 042, objeto de estudio de esta Tesis Doctoral. A fin de no alargar excesivamente este apartado, se incluyen en el Anexo Introdutorio (como información complementaria) detalles más específicos de los diferentes mecanismos de virulencia de los diferentes patotipos diarreicos de *E. coli*.

### 1.1.1.1 *E. coli* enteropatógena

*E. coli* enteropatógena (EPEC) fue el primer patotipo diarreico de *E. coli* caracterizado (Robins-Browne, 1987). En países en vías de desarrollo, dicho patotipo de *E. coli* es uno de los principales agentes causantes de cuadros diarreicos agudos que, a diferencia de otros patógenos diarreicos como rotavirus, adenovirus, *Campylobacter* y *Salmonella*, pueden convertirse en cuadros persistentes que pueden ir acompañados de fiebre y vómitos y, en casos graves, pueden provocar muerte prematura en bebés (Nguyen y col., 2006).

El principal elemento definitorio de este patotipo de *E. coli* es la presencia de la isla de patogenicidad LEE (*locus of enterocyte effacement*) (McDaniel y Kaper, 1997; Robins-Browne y Hartland, 2002). Esta isla genómica es un elemento esencial para la virulencia de la bacteria. Tiene un tamaño de 40 kilobases aproximadamente y codifica para un sistema macromolecular encargado de secretar un gran número de proteínas efectoras responsables de la unión de la bacteria a las células epiteliales intestinales y provocar

## 1. INTRODUCCIÓN

grandes lesiones en los enterocitos. Las diferentes cepas que forman parte de este patotipo se pueden clasificar como cepas “típicas” o “atípicas” en función de la presencia o ausencia del plásmido pEAF (*adherence factor plasmid*). El único reservorio conocido para las cepas “típicas” de EPEC (tEPEC) es el humano, a diferencia de las cepas “atípicas” (aEPEC) que además de humanos, también pueden infectar a otros animales como perros, conejos, monos y ovejas (Moura y col., 2009; Sekse y col., 2011) (figura 1.4).

Globalmente, el sello distintivo de la infección por parte de EPEC es la formación de lesiones de “adhesión y borrado” (A/E, *attaching and effacing*) en el epitelio intestinal y la secreción de proteínas efectoras que presentan toxicidad sobre los enterocitos. La histopatología A/E se caracteriza por el borrado de las microvellosidades intestinales, la adherencia de las bacterias a la membrana plasmática apical del enterocito y la acumulación de actina polimerizada debajo de las bacterias adheridas a la superficie de las células eucariotas (Knutton y col., 1987) (figura A.I.3 del Anexo Introductorio de esta memoria).

La aparición de cuadros diarreicos en el huésped infectado por parte de este patotipo se produce por la sucesión de los procesos responsables de la unión de la bacteria al enterocito y la posterior secreción de las proteínas efectoras. El daño que la bacteria causa en las microvellosidades del intestino cuando ésta se une al enterocito provoca una disminución de la superficie en la que se encuentran los canales responsables de la absorción de electrolitos. Además, las proteínas efectoras EspF, EspH, Tir y Map tienen la capacidad de inhibir el transportador de sodio-glucosa SGLT1, responsable de la absorción de fluidos en el intestino (Dean y col., 2006; Dean y Kenny, 2009), mientras que las proteínas EspG y EspG2 también están implicadas en alteraciones en el transporte de iones de Cl<sup>-</sup> en la membrana apical (Gill y col., 2007). Todo ello provoca grandes alteraciones de la homeostasis entre la absorción y el flujo de líquidos y electrolitos en el intestino, generando la consecuente enfermedad diarreica.

### **1.1.1.2 *E. coli* enterohemorrágica**

Las cepas enterohemorrágicas de *E. coli* (EHEC) son patógenos de carácter zoonótico de gran importancia en la salud pública. Sus infecciones evolucionan generando diversas fisiopatologías, entre las que destacan el dolor abdominal, la diarrea sanguinolenta y cuadros febriles. Este patotipo engloba a todas las cepas de *E. coli* caracterizadas por la



producción de toxinas muy similares a la toxina tipo 1 (Stx) producida por *Shigella dysenteriae*. Ya que las toxinas Stx activas pueden detectarse utilizando el test de toxicidad en células Vero, a las cepas de *E. coli* productoras de toxinas Shiga (STEC, de *Shiga-toxin producing E. coli*) también se les denomina como *E. coli* verotoxigénicas (VTEC, de *Verocytotoxin-producing Escherichia coli*). Actualmente, ambas anotaciones se utilizan indistintamente.

El patotipo EHEC tiene una gran importancia en el ámbito clínico ya que es el principal agente etiológico causante de colitis hemorrágica (CH) que puede derivar en el síndrome urémico hemolítico (SUH) (Nguyen y Sperandio, 2012), el cual tiene lugar en aproximadamente entre el 5 y 10% de los casos, y cuyas manifestaciones clínicas son anemia hemolítica, trombocitopenia y fallo renal.

En 1983, el serotipo O157:H7 de *E. coli* fue identificado por primera vez a través de su asociación con brotes de colitis hemorrágica (Nataro y col., 1998). Desde entonces, dicha cepa ha sido caracterizada como la cepa prototipo del patotipo EHEC principalmente debido a que, aunque las cepas STEC que causan infecciones en los seres humanos pertenecen a un amplio número de serotipos, la mayoría de los brotes y casos de SUH asociados a este patotipo de *E. coli* han sido atribuidos al serotipo O157:H7. Por todo ello, debido a la gran relevancia clínica de este serotipo, es común hablar de dos categorías dentro del patotipo STEC: STEC O157 y STEC no-O157. El principal reservorio del patotipo enterohemorrágico de *E. coli* es el ganado bovino (figura 1.4), donde, a diferencia de los seres humanos, las infecciones por parte de este patógeno son asintomáticas (Wray y col., 2000). El contagio con dicho patógeno tiene lugar a través de múltiples maneras, principalmente a través del consumo de alimentos contaminados, como frutas y hortalizas frescas (lechuga, rábanos, alfalfa, etc.), zumos de naranja y manzana no pasteurizados, carne parcialmente cocinada, agua contaminada, leche recién ordeñada y productos derivados como mayonesa, yogur y quesos, etc., o a través del contacto directo con animales y humanos infectados (Lee, 2004) (figura 1.4).

Tal y como se ha mencionado anteriormente, la presencia de la toxina Shiga es el principal factor de virulencia característico de este patotipo de *E. coli*. Los genes que codifican las toxinas se encuentran codificados en profagos integrados en el cromosoma bacteriano de manera que, ante una situación de estrés, los fagos portadores de la toxina inducen el ciclo lítico, provocando la lisis de la bacteria, lo cual conlleva la liberación de la toxina Shiga al medio extracelular (Neely y Friedman, 1998).



## 1. INTRODUCCIÓN

Las toxinas Shiga producidas por *E. coli* se pueden clasificar en dos tipos, toxina Shiga 1 (Stx1) y toxina Shiga 2 (Stx2). Existen 3 subtipos diferentes de Stx1 (a, c y d), y 7 subtipos de Stx2 (de la “a” a la “g”) (Scheutz y col., 2012). La toxina Shiga 1 producida por *E. coli* enterohemorrágica difiere de la toxina Shiga de *Shigella dysenteriae* en de uno a siete aminoácidos, y comparte un 60% de homología con la toxina Shiga 2. Las cepas EHEC pueden presentar una única variante de la toxina Shiga o una combinación de ambas. Sin embargo, se ha observado que las cepas que codifican el gen *stx2* tienen una mayor prevalencia a la hora de provocar infecciones que derivan en SUH con más frecuencia y de una manera más severa (Boerlin y col., 1999), siendo los subtipos Stx2a y en menor medida Stx2d y Stx2c, los comúnmente asociados a SUH (Friedrich y col., 2002; Feng y col., 2011).

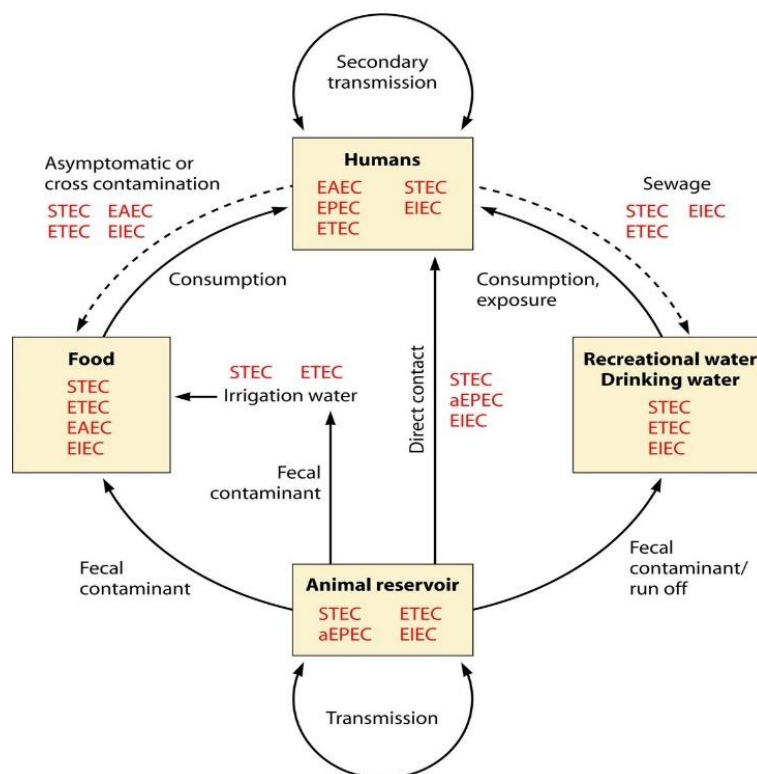
La toxina Shiga está compuesta por dos subunidades principales, designadas A y B (O'Brien y col., 1992; Paton y Paton, 1998). La subunidad A tiene una actividad de ARN N-glicosidasa contra el ARNr 28S (Endo y col., 1988), lo cual provoca la inhibición de la síntesis proteica y produce la apoptosis celular (Sandvig, 2001; Karmali y col., 2010). Por su parte, la subunidad B forma un pentámero que se une al receptor glicolípídico globotriaosilceramida-3 (Gb3) (Lingwood y col., 1987). Esta especificidad determina dónde la toxina Shiga media su fisiopatología. Debido a que los enterocitos no expresan el receptor Gb3, el efecto de la toxina Shiga sobre las células epiteliales intestinales se ejerce a través de otros mecanismos de unión. Se desconoce exactamente el proceso a través del cual la toxina Stx es capaz de pasar del lumen intestinal a través de la barrera epitelial. Una hipótesis es que la inflamación inducida por EHEC al infectar al hospedador pueda proporcionar a la toxina la oportunidad de travesar la barrera epitelial (Schüller, 2011); otra hipótesis es que las bacterias EHEC puedan travesar la barrera epitelial internalizadas en células M, un tipo de enterocitos especializados en la captación de antígenos lumbinales situados en la superficie de los parches de Peyer (Etienne-Mesmin y col., 2011). La secreción de la toxina tipo Shiga también provoca la desregulación de los canales de iones de membrana presentes en la membrana epitelial del intestino, hecho que conduce a la pérdida de iones y de una gran cantidad de agua generando grandes desequilibrios en el flujo de electrolitos a nivel intestinal, dando lugar a los cuadros diarreicos.

Otra toxina importante presente en el patotipo EHEC es la citotoxina subtilasa (SubAB). Dicha toxina se ha encontrado tanto en plásmidos de virulencia (Paton y col., 2004) como

en islas de patogenicidad cromosómicas (Tozzoli y col., 2010; Michelacci y col., 2013). En ensayos *in vitro* realizados con células Vero, se ha podido observar que la proteína SubAB recombinante ejerce efectos más tóxicos que las propias Stx1 o Stx2 (Paton y col., 2004); además, la administración de la proteína pura a ratones es letal y causa hemorragias en el intestino delgado y una enfermedad similar al SUH, que incluye daño renal, hemorragias intestinales, anemia hemolítica y trombocitopenia (Wang y col., 2007; Furukawa y col., 2011).

### **1.1.1.3 *E. coli* enteroinvasiva**

Tanto las cepas enteroinvasivas de *E. coli* (EIEC) como las de *Shigella spp.* son patógenos intracelulares facultativos y potenciales agentes etiológicos de la disentería bacilar (o shigelosis). Los síntomas característicos en personas infectadas con estos microorganismos son diarrea acuosa, con sangre, moco y dolor abdominal. Las cepas EIEC se asocian más a brotes que a casos aislados, en los que la transmisión vía fecal-oral puede ser de persona a persona o por ingestión de alimentos o agua contaminados (figura 1.4). *Shigella*, originalmente conocida como *Bacillus dysenteriae*, fue descrita por primera vez por Kiyoshi Shiga en 1897 durante una epidemia en Japón, donde la bacteria infectó a más de 91.000 personas y causó una tasa de mortalidad de más del 20% (Shiga, 1898; Trofa y col., 1999). El patotipo EIEC fue descubierto unos 50 años después y comparte propiedades bioquímicas, genéticas y patogénicas con el género *Shigella* (Nataro y Kaper., 1998; Lan y col., 2004). EIEC y *Shigella* son patógenos altamente invasivos que utilizan el medio intracelular de las células epiteliales intestinales en el intestino grueso como su nicho replicativo. Estos patógenos son capaces de adaptarse con relativa facilidad a los diferentes cambios de las condiciones ambientales a los que se enfrentan durante el curso de la infección, incluidos el bajo pH gástrico, los cambios de temperatura, la disponibilidad de oxígeno y el estrés oxidativo, así como también la osmolaridad (Marteyn y col., 2012). La infección por parte de este patógeno es un proceso de varias etapas que implica la penetración de la barrera epitelial, la inducción de la muerte de los macrófagos, la invasión de las células epiteliales intestinales, la supresión de la respuesta inmune y el movimiento infectivo intra e intercelular, provocando de esta manera grandes lesiones en la integridad del epitelio intestinal.



**Figura 1.4.** Descripción general de los potenciales reservorios y los modos de transmisión de las cepas patógenas entéricas de *E. coli*. Las mismas se pueden encontrar en varios reservorios animales entre los que se propagan con relativa facilidad. Además, la materia fecal puede contaminar los alimentos (debido a una mala práctica de cocción, contaminación por parte de manipuladores de alimentos sintomáticos o asintomáticos como consecuencia de una higiene de manos inadecuada, etc.), el agua de riego o el agua potable, donde los humanos podemos quedar expuestos después de su consumo o por contacto directo con animales infectados o con otros humanos (transmisión secundaria). Imagen adaptada de (Croxen y col., 2013).

#### 1.1.1.4 *E. coli* enterotoxigénica

El patotipo enterotoxigénico de *E. coli* (ETEC) incluye un grupo muy heterogéneo de cepas que difieren entre ellas en muchas características. Dichas cepas son uno de los principales responsables de los cuadros de diarrea del viajero y diarrea endémica en la mayoría de los países subdesarrollados, donde presentan una amplia distribución (Qadri y col., 2005). En algunos casos la diarrea causada por ETEC puede ser leve, sin presencia de moco o sangre; en otros casos, los cuadros diarreicos pueden ser abundantes causando deshidratación grave, generando también cuadros de cefalea, fiebre, dolor abdominal, náuseas y vómitos (Dalton y col., 1999; Konishi y col., 2011). La infección por parte de ETEC se produce a través de la vía fecal-oral por el consumo de agua o alimentos contaminados (figura 1.4). Además, las cepas del patotipo ETEC provocan grandes pérdidas en la industria porcina ya que son importantes patógenos de lechones recién

nacidos y también en fase post-destete (transición), provocando una gran morbilidad y mortalidad.

El principal factor de virulencia de las cepas ETEC lo constituye la secreción de enterotoxinas, tanto termoestables (ST: *heat-stable enterotoxin*) como termolábiles (LT: *heat-labile enterotoxin*). Estas cepas también expresan diferentes factores de colonización (CF) necesarios para la adherencia de la bacteria al epitelio intestinal. A diferencia de la mayoría de los otros tipos de patógenos, casi todos los factores de virulencia conocidos dentro del patotipo ETEC se codifican principalmente en los plásmidos de virulencia (Gaastra y Svennerholm, 1996). Los detalles de los mecanismos de virulencia de este patotipo se detallan en el Anexo Introductorio de esta memoria.

Pathotype	Host(s)	Site of colonization	Disease(s)	
eEPEC	Children <5 yr, adults at high inocula	Small intestine	Profuse watery diarrhea	
aEPEC	Adults, children	Distal ileum, colon	Watery diarrhea, hemorrhagic colitis, HUS	
STEC			Watery diarrhea, hemorrhagic colitis, HUS	
EIEC/Shigella	Children <5 yr, adults, travelers, immunocompromised persons	Colon	Shigellosis/bacillary dysentery, potential HUS	
EAEC	Adults	Small intestine and/or colon	Traveler's diarrhea, HUS (stx+)	
ETEC	Children	Small intestine	Persistent diarrhea	
	Immunocompromised persons		Persistent diarrhea	
DAEC	Children <5 yr, travelers	Small intestine	Watery diarrhea	
AIEC	Children (increasing in severity from 18 mo to 5 yr), adults	Intestine (uncharacterized location)	Persistent watery diarrhea in children, speculated to contribute to Crohn's disease in adults	
AIEC	Adults, children	Small intestine	Crohn's disease	
Pathotype	Known reservoir(s)/source(s) of contamination	Treatment	Adhesion	Genetic identifiers
eEPEC	Humans	Oral rehydration, antibiotics for persistent cases	Attaching and effacing	eae+, bfp+, stx-
aEPEC	Humans, animals	Hydration, supportive for HUS	Attaching and effacing	eae+, stx-
STEC	Humans, animals, food, water			
EIEC/Shigella	Humans, animals, food, water	Oral rehydration, antibiotics	NA (invasive)	ipaH+, ial+, stx+ (S. dysenteriae)
EAEC	Food, occasionally adult carriers	Antibiotics, oral rehydration	Stacked brick and/or invasive	aatA+, aaiC+, other candidates
ETEC	Food, water, humans, animals	Antibiotics, oral rehydration, potentially probiotics	CF mediated	CFs, LT, ST
		Fluoroquinolones		
DAEC	Unknown	Rehydration	Diffuse adherent and/or invasive	No uniform markers
AIEC	Unknown	Antibiotics, surgical resection	NA (invasive)	Uncharacterized

**Tabla 1.1.** Tabla resumen de la descripción general de los diferentes patotipos entéricos de *E. coli*. Se indica el huésped principal, el lugar de colonización de la bacteria, la enfermedad que causa, los posibles lugares de contaminación y los potenciales reservorios transmisores de la bacteria, el tratamiento recomendado ante las infecciones, los mecanismos a través de los cuales las bacterias se adhieren al intestino y los marcadores genéticos utilizados para identificar cada uno de los patotipos descritos. Tabla adaptada de (Croxen y col., 2013).

### 1.1.1.5 *E. coli* con adherencia difusa

El patotipo de *E. coli* con adherencia difusa (DAEC) engloba a aquellas cepas diarreicas de *E. coli* que se unen a las células epiteliales intestinales sin utilizar los patrones clásicos de adherencia, como la adherencia localizada A/E (propia de los patotipos EPEC o STEC) o en forma de microcolonias (como el patotipo EPEC). Dichas cepas tienen la capacidad de unirse con una adherencia difusa a las células epiteliales HEp-2 en cultivo, donde la

## 1. INTRODUCCIÓN

adherencia bacteriana se produce en toda la superficie de la célula epitelial siguiendo un patrón disperso. El patotipo DAEC está asociado con cuadros diarreicos acuosos que pueden volverse persistentes en niños pequeños, mientras que los adultos suelen ser portadores asintomáticos (Le Bouguéec, 1999). Se desconoce cuál es el reservorio de DAEC y su principal mecanismo de transmisión. Los detalles de los mecanismos de virulencia de este patotipo se detallan en el Anexo Introductorio de esta memoria.

### **1.1.1.6 *E. coli* adherente invasiva**

El patotipo adherente invasivo de *E. coli* (AIEC) ha sido reconocido como uno de los agentes causantes de la enfermedad de Crohn, una enfermedad inflamatoria intestinal que afecta principalmente al intestino delgado. Sin embargo, no existe un agente etiológico responsable único de esta enfermedad. La hipótesis actual es que la enfermedad está causada por una combinación de factores, entre los que se incluyen factores genéticos del huésped, la microbiota intestinal, factores ambientales y microorganismos patógenos entéricos, tales como *Mycobacterium avium subsp. paratuberculosis*, *Campylobacter*, citomegalovirus o la propia *E. coli* adherente invasiva (Mann y Saeed, 2012).

El patotipo AIEC se define por la capacidad que presentan las bacterias de adherirse e invadir células epiteliales y replicar en el interior de las mismas y en los macrófagos. Dicho patotipo no expresa los factores de virulencia comunes encontrados en otras cepas patógenas de *E. coli*, por lo que el componente genético de su fenotipo inflamatorio e invasivo no está descrito completamente (Nash y col., 2010).

La organización del genoma AIEC es similar al de otras cepas patógenas de *E. coli*, con grandes regiones del genoma central interrumpidas por islas genómicas (probablemente adquiridas por transferencia horizontal (Miquel y col., 2010)), donde se encuentran los factores de virulencia que contribuyen a la adherencia bacteriana, la invasión y la supervivencia intracelular (Martinez-Medina y col., 2009; Nash y col., 2010; Krause y col., 2011). Entre estos factores de virulencia destaca el sistema de secreción tipo VI (necesario para la secreción de proteínas al medio extracelular), varios genes que codifican adhesinas, entre los que destaca el pili de tipo I, la proteína de adhesión FimH (proteína muy importante para la adhesión al epitelio intestinal) y largas fibrillas polares que juegan un papel crucial en la invasión, además de genes implicados en la adquisición de hierro (Nash y col., 2010). Los detalles del proceso de infección por parte de este patotipo se detallan en el Anexo Introductorio de esta memoria.

### **1.1.2 *E. coli* enteroagregativa y su cepa prototipo 042**

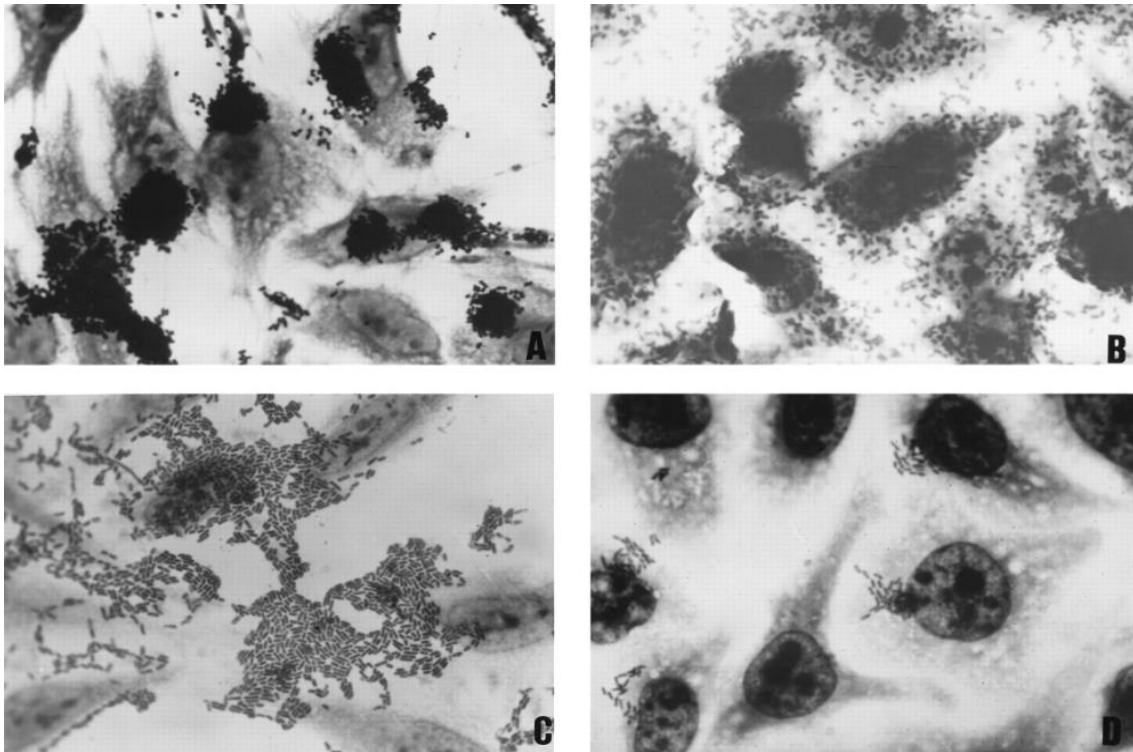
El patotipo enteroagregativo de *E. coli* (EAEC) fue descrito por primera vez en 1987 (Nataro y col., 1987), en un estudio cuyo objetivo era estudiar el diferente patrón de adherencia a células epiteliales Hep-2 por parte de las diferentes cepas de *E. coli* aisladas en un estudio clínico cuyos pacientes eran niños de Chile con diarrea. En 1995 se demostró una asociación significativa del patotipo EAEC con la aparición de diarreas cuando en un estudio clínico con voluntarios pudo ponerse de manifiesto que la cepa *E. coli* 042 fue responsable de causar cuadros diarreicos en la mayoría de los voluntarios sometidos a dicho estudio (Nataro y col., 1995). Diferentes ensayos clínicos posteriores han mostrado la asociación de diferentes cepas EAEC con casos de diarreas infantiles endémicas tanto en países en vías de desarrollo como en países industrializados, con diarreas persistentes en enfermos de sida, con la denominada diarrea de viajero en países subdesarrollados, con brotes diarreicos asociados a agua y alimentos contaminados y con simples casos esporádicos de diarrea (Huang y col., 2006). Las manifestaciones clínicas de las infecciones causadas por cepas EAEC cursan con diarreas acuosas/mucosas ocasionalmente con pequeñas hemorragias, poca fiebre y sin vómitos (Nataro y Kaper, 1998). Además, pacientes con una infección activa por parte de este patógeno muestran señales de inflamación intestinal (elevada presencia de citoquinas proinflamatorias y lactoferrina fecal) (Steiner y col., 1998; Harrington y col., 2005; Strauman y col., 2010).

Las cepas EAEC se pueden diferenciar de otros patotipos intestinales de *E. coli* debido a que presentan un patrón de adherencia a células Hep-2 en forma de “apilado de ladrillos” (Nataro y col., 1987), muy diferente al patrón de adherencia difuso propio del patotipo DAEC, o al patrón de adherencia basado en la formación de microcolonias propio del patotipo EPEC (figura 1.5).

Desde el estudio referido anteriormente (Nataro y col., 1995), se ha adoptado la cepa *E. coli* 042 como la cepa modelo de referencia de este patotipo bacteriano. En 2010 se anotó la secuencia completa del genoma de esta cepa y se caracterizaron sus factores de virulencia (Chaudhuri y col., 2010).

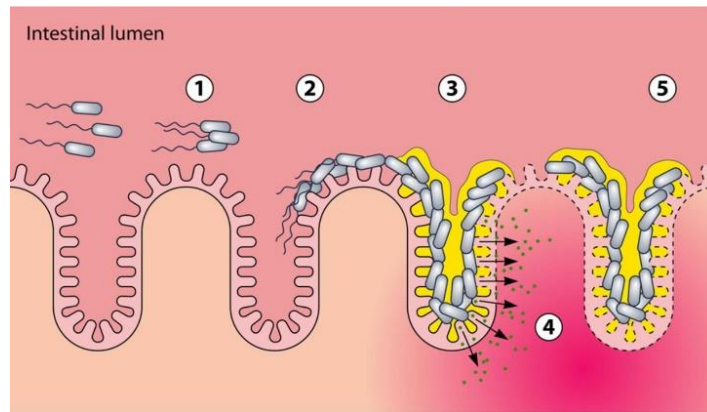


## 1. INTRODUCCIÓN



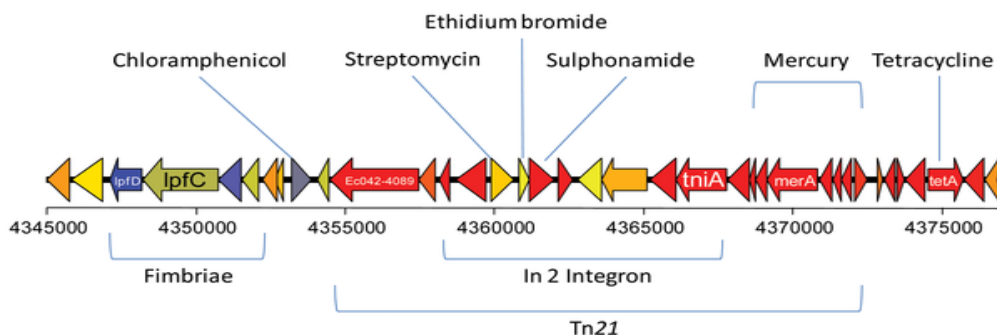
**Figura 1.5.** Patrón de adherencia específico de los patotipos EPEC, DAEC y EAEC a células epiteliales Hep-2. (A) Adherencia localizada de células EPEC. Las bacterias son capaces de formar acúmulos densos de microcolonias en la superficie de las células Hep-2. (B) Adherencia difusa propia del patotipo DAEC. Las células se dispersan sobre la superficie de las células Hep-2. (C) Patrón de adherencia autoagregativo específico de las células EAEC. Las bacterias se agregan entre ellas y se adhieren a las células epiteliales en forma de “apilado de ladrillo”. (D) De manera análoga a lo representado en la imagen (A), en el patrón de adherencia típico del patotipo EPEC las células bacterianas también presentan la capacidad de agruparse en forma de microcolonias alrededor de las células epiteliales de una manera menos compactada. Imagen adaptada de (Scaletsky y col., 1999).

El potencial patogénico de las cepas EAEC es debido a la capacidad que presentan las mismas para adherirse a las células intestinales de la mucosa del íleo y del colon, producir enterotoxinas y citotoxinas, y de inducir procesos inflamatorios, provocando de esta manera una gran toxicidad sobre la mucosa intestinal, generando lesiones definidas por la dilatación de las criptas intestinales, la vesiculación de las microvellosidades y la extrusión de las células epiteliales de la mucosa entérica (Henderson y col., 1999) (figura 1.6).



**Figura 1.6.** Representación esquemática de las etapas implicadas en el mecanismo patogénico de las cepas EAEC. (1) Aglutinación de las bacterias EAEC planctónicas. (2) Adherencia al epitelio intestinal e inicio de la colonización del intestino. (3) Formación de biopelículas bacterianas en la superficie del epitelio intestinal. (4) Secreción de toxinas bacterianas, que producen lesiones en el epitelio y un aumento de la secreción iónica. (5) Formación de biopelículas adicionales sobre el epitelio dañado. Imagen adaptada de (Hebbelstrup Jensen y col., 2014).

Los diferentes aislamientos clínicos de este patotipo enteroagregativo presentan resistencia múltiple a varios antibióticos: tetraciclina, espectinomicina, cloranfenicol, estreptomycin y ampicilina, tal y como se ha descrito en estudios epidemiológicos realizados sobre dichos aislamientos (Yamamoto y col., 1992; Vila y col., 2000). De manera específica, en la cepa *E. coli* 042, la multiresistencia de este patógeno es debida a la inserción del transposón *Tn2411* en su cromosoma, que le aporta genes que codifican por elementos que otorgan resistencia al cloranfenicol (gen *cat*), a la tetraciclina (gen *tetA*), a la estreptomycin y espectinomicina (casetes *aadA1*), a la sulfonamida (gen *suI*), al bromuro de etidio (gen *emrE*) y al mercurio (operón *merRTPCAD*) (Chaudhuri y col., 2010) (figura 1.7).



**Figura 1.7.** Representación esquemática del transposón *Tn2411* de la cepa *E. coli* 042. En la figura se representan los genes que codifican los elementos de resistencia al cloranfenicol, a la tetraciclina y al mercurio, a los que la bacteria es resistente. En esta imagen también se representa el integrón *In 2*, que codifica para genes que confieren resistencia a otros antibióticos (estreptomycin, espectinomicina, sulfonamida) y resistencia a elementos mutagénicos como el bromuro de etidio y el cloruro de mercurio. Imagen adaptada de (Chaudhuri y col., 2010).



## 1. INTRODUCCIÓN

El genoma de la cepa *E. coli* 042 tiene un tamaño de 5,2 Mb y codifica 4886 genes, de los cuales 93 codifican para tARNs, 22 para rARNs y 110 son pseudogenes. Además, un 11% de los genes codifican para proteínas hipotéticas conservadas sin función conocida y un 10% codifican para elementos móviles tales como transposasas, integrasas y elementos genéticos relacionados con bacteriófagos. También se identificaron 78 islas genómicas, las cuales tienen unas secuencias altamente divergentes entre ellas y se encuentran distribuidas de forma variable entre las diferentes cepas de *E. coli* secuenciadas. Estas islas genómicas, designadas como ROD (*regions of difference*) tienen un tamaño total de 1,26 Mb y representan el 24% del genoma de *E. coli* 042. Dichas regiones génicas codifican para un gran número de factores de virulencia, enzimas metabólicas, proteínas sin función atribuida y elementos móviles tales como profagos o el propio transposón conjugativo *Tn2411*, presente en la ROD66 (Chaudhuri y col., 2010).

### **1.1.2.1 Factores de virulencia de la cepa *E. coli* 042: sistemas de secreción y toxinas entéricas**

El genoma de la cepa *E. coli* 042 contiene genes que codifican para sistemas de secreción (TSS) correspondientes a los 6 tipos distintos descritos hasta la fecha (del tipo I al tipo VI). Dichos sistemas, funcionales en su totalidad, incluyen proteínas estructurales, proteínas chaperonas y proteínas efectoras (Chaudhuri y col., 2010).

La cepa 042 codifica para tres copias del T1SS (dos copias cromosómicas y una copia plasmídica). El sistema de secreción tipo I plasmídico permite la expresión de la dispersina (codificada en el gen *aap*, *antiaggregation protein*), una proteína de superficie celular que se une mediante enlaces no covalentes al lipopolisacárido (LPS) de la bacteria, neutralizando de esta manera la fuerte carga negativa del mismo permitiendo que las fimbrias AAF (que presentan a su vez una elevada carga positiva) puedan extenderse a lo largo de la superficie bacteriana y unirse así a los receptores correspondientes para mediar la unión bacteria-célula intestinal (Harrington y col., 2006). Respecto a los otros dos sistemas de secreción tipo I codificados en el cromosoma de *E. coli* 042, el primero presenta homología con el T1SS de EHEC O157:H7 y es responsable de la secreción de la toxina RTX; el segundo es responsable de la secreción del péptido con actividad antibiótica microcina H47 (Rodríguez y col., 1999).

El sistema de secreción tipo II del cromosoma de *E. coli* 042, transcrito a partir de un ARNm policistrónico (Yang y col., 2007), se describió por primera vez en la cepa ETEC

H10407 como el sistema responsable de la secreción de la enterotoxina LT (Tauschek y col., 2002). Aunque dicho sistema de secreción se encuentra altamente conservado entre bacterias patógenas, ni la cepa 042 (ni muchas otras cepas) codifican para la toxina LT. Sin embargo, de manera específica, *E. coli* 042, a través de este sistema, secreta al medio extracelular la proteína YghJ, una proteína con homología al factor de colonización AcfD de *Vibrio cholerae* (Parsot y col., 1991).

El sistema de secreción tipo III se encarga de producir el sistema flagelar e inyectar proteínas efectoras en la célula huésped (Pallen y Gophna, 2007). Dicho sistema flagelar confiere el serotipo H18 a la cepa *E. coli* 042. La cepa *E. coli* 042 también codifica en su cromosoma otros dos T3SS no flagelares: los sistemas ETT2 y *eip*. La isla de patogenicidad ETT2 contiene un gran número de genes ortólogos presentes en las islas de patogenicidad 1, 2 y 3 de *Salmonella*, entre los que destacan la proteína YgeH (codificada en el gen ortólogo al que codifica por la proteína reguladora HilA, regulador global de la SPI-1) e YgeK (codificada en el gen ortólogo al que codifica por la proteína reguladora SsrB, regulador global de la SPI-2) (Ren y col., 2004). La isla de patogenicidad *eip* también contiene otro gen que codifica para una proteína homóloga a HilA (la proteína EilA en *eip*), además de proteínas homólogas a las proteínas efectoras SipB y SipD de la SPI-1 de *Salmonella* (EipB y EipD, respectivamente) y la chaperona SicA (EipA en *eip*) (Ren y col., 2004). Tanto las proteínas YgeH y EilA de la cepa *E. coli* 042 son plenamente funcionales ya que son capaces de complementar un fenotipo *hila*<sup>-</sup> en *Salmonella* Typhimurium (Hüttener y col., 2014).

El sistema de secreción tipo IV codifica por el aparato necesario para que se produzca la conjugación del plásmido pAA2 (sistema de transferencia conjugativo tipo-F). El plásmido pAA2 codifica para una maquinaria conjugativa completa muy similar a la de los plásmidos pUT189 y R100 (Chaudhuri y col., 2010), sistema común a la mayoría de los plásmidos IncFII.

La cepa *E. coli* 042 codifica un gran repertorio de sistemas de secreción del tipo V. Concretamente, su genoma codifica para 14 genes que codifican proteínas englobadas dentro de la familia de autotransportadores AT-1, entre los que destacan la mucinasa Pic y la toxina plasmídica Pet, proteínas del tipo serín proteasas que una vez secretadas al medio extracelular juegan un papel importante en la colonización bacteriana. Además, *E. coli* 042 posee tres copias del autotransportador Antígeno-43 (EC042\_2242, EC042\_4511, EC042\_4803). El antígeno 43 es una proteína de membrana externa

## 1. INTRODUCCIÓN

implicada en el contacto célula-célula cuando las bacterias forman *biofilm* (Danese y col., 2000). Las dos primeras copias (EC042\_2242 y EC042\_4511) son muy similares entre sí (87% de identidad); la tercera copia (EC042\_4803) es más divergente, con una identidad del 72% y 65% respecto a las copias EC042\_2242 y EC042\_4511, respectivamente. Otro sistema de secreción tipo 5 que *E. coli* 042 codifica en su genoma es el sistema TPS (*two-partner secretion system*), formado por los genes *yaeT* e *ytfM* que codifican para una proteína de membrana externa y su sustrato, aunque la función de estas proteínas en la virulencia de la cepa 042 se desconoce (Stegmeier y col., 2007; Knowles y col., 2009). Otro sistema de secreción del tipo 5 que contiene *E. coli* 042 en su genoma es el formado por autotransportadores triméricos AT-2. Concretamente, esta cepa codifica para un gen ortólogo al que codifica por la proteína UpaG en la cepa uropatogénica CFT073 de *E. coli*. De manera análoga a la familia de autotransportadores AT-1, la proteína UpaG está implicada en la interacción célula-célula, en la formación de *biofilm* y en la adherencia a las células epiteliales. Además, como otros miembros de la familia AT-2, dicha proteína es capaz de unirse a las proteínas de la matriz extracelular (Valle y col., 2008).

En referencia al sistema de secreción tipo VI, *E. coli* 042 contiene tres copias del mismo en su genoma. Dos de estos sistemas (EC042\_4524 a 4555 y EC042\_4562 a 4577), se encuentran separados por apenas 4 Kb y su expresión depende del regulador transcripcional plasmídico AggR (Dudley y col., 2006) (ver apartado 1.1.3). El tercer T6SS (EC042\_0210 a 0226) codificado dentro del genoma de *E. coli* 042 se encuentra insertado el locus de ARNt *aspV*. Este T6SS está ampliamente distribuido entre cepas patógenas de *E. coli* de los patotipos UPEC, APEC y EAEC (Brzuszkiewicz y col., 2006; Chen y col., 2006; Johnson y col., 2007; Touchon y col., 2009).

Además, el cromosoma de *E. coli* 042 codifica dos toxinas diferentes: ShET1 y HlyE. HlyE es una toxina de 34 kDa, cuya estructura secundaria se encuentra enriquecida por hélices- $\alpha$  y es capaz de oligomerizar en estructuras más complejas para formar una toxina capaz de crear poros en la superficie de las células infectadas. Esta toxina tiene la capacidad de provocar efectos citolíticos y citopáticos sobre células humanas en cultivo *in vitro* (Mueller y col., 2009). Sin embargo, dicha toxina también se encuentra presente en cepas no patógenas de *E. coli* (von Rhein y col., 2008), por lo que su papel en la patogénesis de la cepa *E. coli* 042 parece tener un rol secundario. Por el contrario, la toxina ShET1 es una proteína formada por 2 subunidades (codificadas en los genes *setA*

y *setB*) que oligomerizan para formar un complejo formado por una subunidad de SetA y un pentámero de SetB, dando lugar a la forma activa de la toxina que se encarga de provocar un incremento en la secreción intestinal vía AMPc y GMPc (Henderson y col., 1999).

### 1.1.3 El plásmido pAA2 y el regulador transcripcional AggR

Una característica genómica común a la mayor parte de cepas EAEC es la presencia de un plásmido de virulencia denominado pAA (*plasmid of Aggregative Adherence*). Concretamente, la cepa *E. coli* 042 posee un plásmido de este tipo, denominado pAA2. Su tamaño supera las 100 Kb de masa molecular y pertenece al grupo de incompatibilidad IncFIIA (Chaudhuri y col., 2010). La mayoría de los genes del plásmido de virulencia pAA2 a los que al menos puede atribuirse una función se clasifican en tres categorías funcionales: una gran variedad de genes de virulencia hipotéticos, genes que forman parte del aparato de conjugación (y transferencia) del plásmido y genes que conforman la maquinaria de replicación del mismo. Concretamente, el plásmido pAA2 contiene 152 pautas abiertas de lectura, de las cuales 32 son pseudogenes, 23 son proteínas hipotéticas conservadas sin función atribuida, 7 proteínas hipotéticas sin homología previa con ninguna proteína descrita en las bases de datos, 18 elementos móviles que codifican para elementos de inserción y transposasas, 17 genes caracterizados como factores de virulencia putativos y 55 genes responsables de la replicación, mantenimiento y la transferencia del plásmido. A diferencia de la gran mayoría de los plásmidos tipo-F que presentan múltiples replicones, el plásmido pAA2 tiene un único replicón identificable regulado por las proteínas RepA y CopB (Chaudhuri y col., 2010).

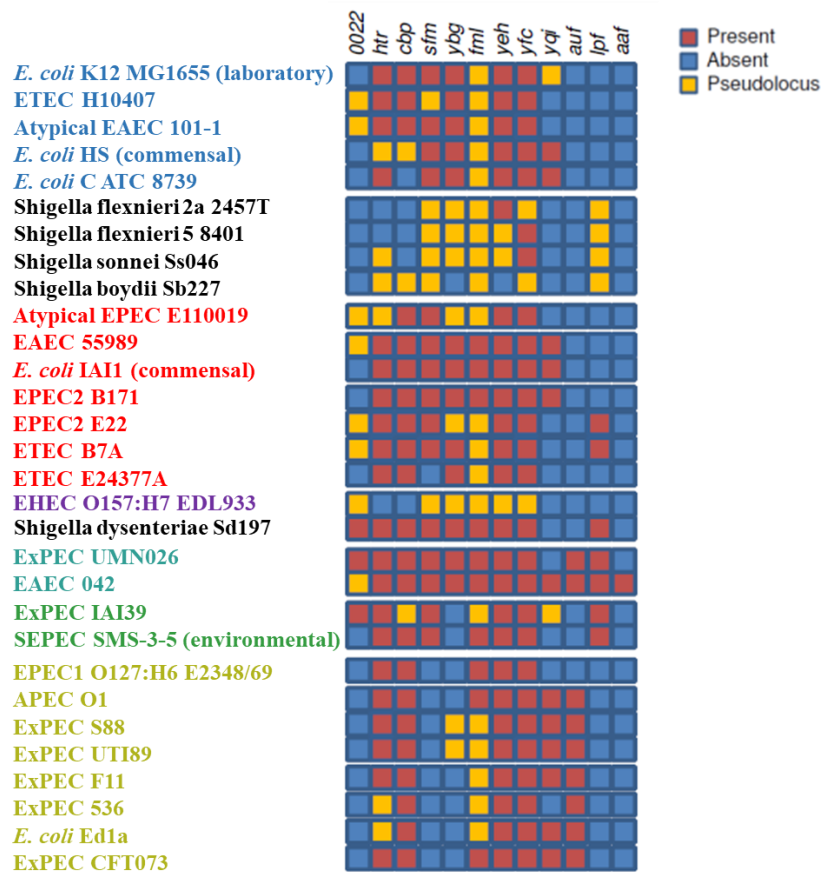
El plásmido de virulencia pAA2 codifica para las toxinas Pet (*plasmid-encoded toxin*) y EAST-1. La proteína Pet es una proteasa secretable perteneciente a la familia de serín proteasas. Una vez secretada, la proteína es internalizada por las células huésped y tiene la capacidad de promover la degradación de las proteínas del citoesqueleto (concretamente de la espectrina) de las células que la internalizan (Villaseca y col., 2000; Navarro-García y col., 2001). La toxina Pet se encuentra regulada de manera sinérgica por parte de las proteínas CRP y Fis (Rossiter y col., 2011). A pesar de su importante papel en la patogénesis de las cepas enteroagregativas de *E. coli*, la presencia de la proteasa Pet se encuentra restringida a un pequeño número de cepas (Czeczulin y col., 1999). La proteína EAST-1, codificada por el gen *astA*, se localiza en el plásmido pAA2

## 1. INTRODUCCIÓN

de manera adyacente al gen *pet* y es capaz de provocar un incremento en la secreción de aniones en la célula infectada (Ménard y col., 2004). Sin embargo, esta proteína también está presente en cepas enteroagregativas que no producen diarreas en estudios con voluntarios (Nataro y col., 1995; Ménard y col., 2004), por lo que la función de esta toxina parece no jugar un papel principal a la hora de provocar diarreas en los pacientes infectados.

El plásmido pAA2 también codifica para una familia de fimbrias, denominadas fimbrias adherentes (AAF). A ellas, hay que sumar otros 11 tipos de fimbrias diferentes codificados en el cromosoma de *E. coli* 042 (figura 1.8). El potencial que presenta dicha cepa de expresar este gran repertorio de fimbrias puede deberse a que los diferentes sistemas alélicos estén implicados en diferentes mecanismos adaptativos asociados a la colonización de diferentes hospedadores o diferentes nichos dentro de un mismo hospedador (Humphries y col., 2005).

Existen 5 variantes alélicas de las fimbrias AAF codificadas en los plásmidos pAA de virulencia de las distintas cepas de *E. coli* enteroagregativas: AAF/I, AAF/II, AAF/III, AAF/IV y AAF/V (Nataro y col., 1992; Czeczulin y col., 1997, Bernier y col., 2002, Boisen y col., 2008, Jonsson y col., 2017). De manera más concreta, en el plásmido pAA2 de la cepa 042 se encuentra presente la clase de fimbrias AAF/II. Los genes requeridos para la síntesis de estas fimbrias adherentes se hallan codificados en dos operones distintos (*aafDA* y *aafCB*), cuya expresión es dependiente del regulador transcripcional AggR.



**Figura 1.8.** Distribución filogenética de los *loci* fimbriales presentes en cepas patógenas y no patógenas de *E. coli* y *Shigella*. Los *loci* que codifican los distintos sistemas fimbriales presentan una distribución diferencial entre la filogenia de *E. coli*. Las cepas pertenecientes a los diferentes grupos filogenéticos están marcadas de colores diferentes: en azul están marcadas las cepas pertenecientes al grupo filogenético A, en rojo al grupo B1, en verde claro al B2, en azul celeste al D1, en verde oscuro al D2 y en morado al grupo E. En negro se remarcan las cepas de *Shigella*. Imagen adaptada de (Chaudhuri y col., 2010).

La proteína AggR se encuentra codificada en el plásmido pAA2 de la cepa *E. coli* 042 y en otros plásmidos pAA de adherencia agregativa. Dicha proteína es un importante factor de transcripción perteneciente a la familia de reguladores transcripcionales AraC/XylS (Nataro y col., 1994; Gallegos y col., 1997). Las proteínas pertenecientes a este grupo de reguladores transcripcionales se caracterizan por presentar un dominio C-terminal de 99 aminoácidos (altamente conservados) y por regular diversas funciones celulares incluyendo el metabolismo, la respuesta al estrés y la síntesis de factores de virulencia. Diferentes estudios epidemiológicos realizados con cepas EAEC sugieren que las cepas que expresan la proteína AggR son más propensas a causar enfermedad diarreica que las cepas que no codifican por este regulador transcripcional. De esta manera, se ha propuesto el término “EAEC típico” para describir las cepas que codifican para la proteína AggR y

## 1. INTRODUCCIÓN

sus genes diana (el regulón AggR), y el término “EAEC atípicas” a aquellas que no codifican el gen *aggR* (Jiang y col., 2002; Huang y col., 2006).

La gran mayoría de reguladores tipo AraC presentan mecanismos de autorregulación (Yu y DiRita, 1999; Munson y Scott, 2000; Ibarra y col., 2003). En concordancia con esto, la proteína AggR también es capaz de autoactivar su propia expresión. Para ello, este factor transcripcional es capaz de unirse en dos sitios diferentes de su propio promotor, situados antes (caja -35) y después (nucleótidos +27 a +58) del inicio de transcripción del gen (Morin y col., 2010). Además, su expresión también se encuentra regulada positivamente por la proteína asociada al nucleoide Fis, y regulada negativamente por H-NS (Morin y col., 2010) y por la proteína Aar (*AggR activated regulator*). Esta pequeña proteína (63 aminoácidos y 7,23 kDa) pertenece a la familia de reguladores ANR (*AraC Negative Regulators*), que incluye a un gran número de reguladores de la expresión génica presentes en un gran número de cepas patógenas gramnegativas entre las que destacan *Vibrio* spp., *Salmonella* spp., *Shigella* spp., *Yersinia* spp., *Citrobacter* spp., cepas de *E. coli* pertenecientes a los patotipos ETEC y EAEC, y miembros de la familia *Pasteurellaceae* (Santiago y col., 2014; Santiago y col., 2016). A pesar de no presentar capacidad de unión al ADN, la proteína Aar es capaz de modular la expresión de AggR uniéndose a esta proteína inhibiendo de esta manera su capacidad de unirse a los promotores de los genes que regula (Santiago y col., 2016).

Los primeros genes descritos como genes regulados por la proteína AggR fueron las fimbrias *aafA* de tipo II (Nataro y col., 1994). En años posteriores se acabó de definir el regulón completo de AggR, formado principalmente por el gen *aap*, cuyo producto génico es la dispersina (Sheikh y col., 2002), y los genes que codifican por el sistema de secreción de esta proteína (sistema de secreción tipo I denominado Aat), necesario para su exportación a la superficie bacteriana (Nishi y col., 2003). La proteína AggR también regula la expresión de un gran número de genes cromosómicos que se hallan presentes en un operón de la isla de patogenicidad localizada en el *locus pheU* (Dudley y col., 2006). Esta isla de patogenicidad mide 117 Kb, y AggR se encarga de regular la expresión de varios genes que codifican el sistema de secreción tipo VI denominado Aai. Todo ello pone de manifiesto que, a pesar de estar codificado en el plásmido pAA2, el regulador transcripcional AggR no solamente se limita a modular la expresión de genes plasmídicos, sino que también es capaz de regular la expresión de genes situados en cromosoma de la bacteria. En 2013 se definió el regulón completo de AggR en la cepa



*E. coli* 042 mediante un análisis transcriptómico realizado por *microarray* (Morin y col., 2013). En este estudio se demostró que AggR se encarga de activar la expresión de al menos 44 genes, de los cuales 16 son proteínas hipotéticas. De manera más específica, dentro de los genes plasmídicos regulados por AggR (21 genes en total), AggR controla la expresión de su propio gen, del gen *aar* (que codifica por la proteína represora de la expresión de AggR), de 5 genes pertenecientes al sistema de secreción de la dispersina (*aatA*, gen que codifica para la proteína de membrana externa con homología a TolC; *aatB*, gen que codifica para una proteína hipotética de membrana; *aatC*, gen que codifica para el transportador ABC de unión a ATP; *aatD*, gen que codifica para una proteína putativa de membrana necesaria para la secreción de la dispersina y la permeasa, codificada en el gen *aatP*), la propia dispersina (*aap*), de 4 genes implicados en la biosíntesis de fimbrias del tipo II (*aafA*, que codifica para la subunidad mayor de la fimbria adherente agregativa de tipo II; *aafB*, que codifica para la adhesina; *aafC*, gen que codifica para la proteína acomodadora de la fimbria adherente agregativa tipo II y *aafD*, gen que codifica para la chaperona del sistema), de los genes EC042\_pAA003 y EC042\_pAA004 (genes implicados en la formación de *biofilm*), del gen *shf* (cuyo producto es una desacetilasa de polisacárido), y de 7 genes adicionales que codifican para proteínas putativas (Morin y col., 2013).

A nivel cromosómico, la proteína AggR se encarga de regular la expresión de 23 genes, localizados en dos regiones separadas (el primer *cluster* alberga los genes del EC042\_3181 al EC042\_3187 y el otro *cluster* de la posición EC042\_4562 a la EC042\_4583, además de una pauta abierta de lectura aislada (EC042\_4006).

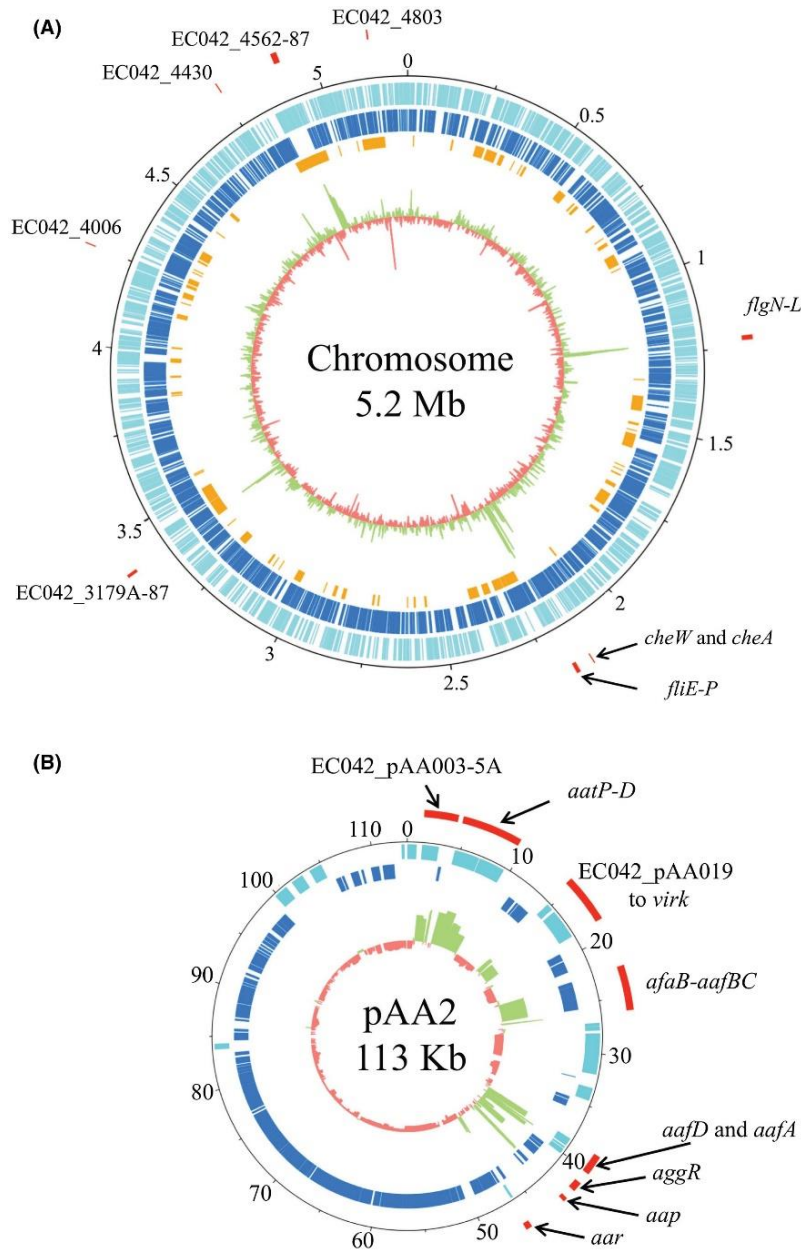
El gen EC042\_3181 es un gen homólogo al que codifica para el activador transcripcional PerC, presente en cepas enteropatógenas de *E. coli*. El mismo se encarga de regular la expresión de la isla de patogenicidad LEE1. El gen EC042\_3182 codifica para una nucleasa tipo ParB, altamente conservada en *E. coli*, *Shigella spp.*, y *Klebsiella pneumoniae*. EC042\_3183 codifica para una hipotética proteína citoplasmática presente en los patotipos EAEC y el patotipo patogénico aviar de *E. coli*. La proteína EC042\_3187 presenta dominios conservados característicos de la familia de helicasas UvrD, presentes en *E. coli*, *Vibrio spp.*, *Salmonella spp* y *Citrobacter rodentium*. De manera análoga, la ORF EC042\_4006 (gen *yicS*) codifica para una proteína hipotética altamente conservada en *E. coli*, *Salmonella spp.*, *Shigella spp.*, y *Citrobacter rodentium*.



## 1. INTRODUCCIÓN

Los otros 20 genes cromosómicos regulados por AggR se encuentran en la isla de patogenicidad PAI-1, insertada de manera adyacente al ARNt *pheU*. Toda esta región incluye el sistema de secreción tipo VI denominado Aai (ORFs de EC042\_4562 a EC042\_4577), entre los que se encuentra la proteína efectora secretada AaiC (codificada en el gen EC042\_4564). La región comprendida entre los genes EC042\_4580 al EC042\_4583 también forma parte de la PAI-1 regulada por AggR, aunque estos genes no forman parte del sistema de secreción tipo VI. Concretamente, estos genes codifican para cuatro proteínas hipotéticas altamente conservadas dentro del genoma de otras cepas patógenas de *E. coli*, incluida la cepa O104:H4 C227-11, cepa codificante para la toxina Shiga y responsable del brote de SUH en Europa en 2011 (apartado 1.1.4).

Sin embargo, como en los análisis de la expresión génica global mediante *microarrays* los diseños de las sondas y la anotación del genoma pueden influir en el resultado de los mismos, además de presentar algunos problemas de detección de transcritos de baja abundancia, en 2019 se repitió el análisis para caracterizar el regulón de AggR mediante la metodología RNA-seq de alto rendimiento (Yasir y col., 2019). Mediante esta nueva aproximación, además de los genes plasmídicos regulados por AggR reportados por *microarray* (Morin y col., 2013), se identificaron nuevos genes del plásmido pAA2 regulados por este activador transcripcional. Concretamente, se identificaron nuevos pseudogenes de función desconocida (EC042\_pAA006, EC042\_pAA019, EC042\_pAA033, EC042\_pAA051, EC042\_pAA053, EC042\_pAA061 y EC042\_pAA105), además de *virK* (gen que codifica para una proteína de virulencia con una alta homología a la proteína VirK de *Shigella flexneri*) como nuevos genes regulados por la proteína AggR (figura 1.9).

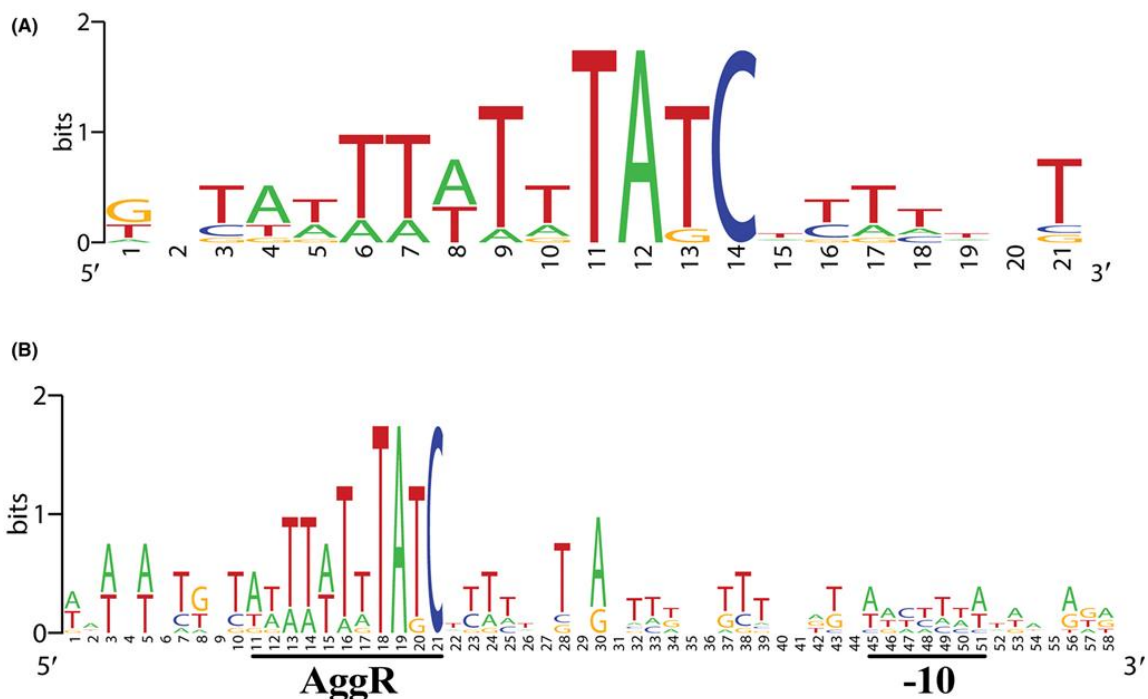


**Figura 1.9.** Expresión génica diferencial entre la cepa *E. coli* 042 salvaje y el mutante *aggR* determinada por RNA-seq, tanto de los genes presentes en el núcleo cromosómico (A) como en el plásmido pAA2 (B). Las líneas rojas más externas identifican algunos de los genes expresados diferencialmente (etiquetados con su nombre o número de gen), seguidos de la localización en la que se encuentran (etiquetada en Kb o en Mb). Los genes anotados están indicados en función de su orientación (cadena *sense* y *antisense*, marcados en azul claro y azul oscuro respectivamente). Las regiones cromosómicas de diferencia (ROD) identificadas por (Chaudhuri y col., 2010) se muestran en naranja. Los genes regulados positivamente por AggR se representan en verde y los genes reprimidos por este factor transcripcional se representan en rojo. Es destacable que los genes regulados por AggR se localizan en una mitad del plásmido, mientras que los genes necesarios para la replicación del plásmido y la transferencia conjugativa del mismo están ubicados en la otra mitad y son independientes del control de AggR. Adaptado de (Yasir y col., 2019).

## 1. INTRODUCCIÓN

A nivel cromosómico, además de los genes identificados mediante el *microarray*, en el mutante *aggR* se encontraron reprimidos genes asociados con la motilidad flagelar, mientras que los genes homólogos del antígeno 43 (Agn43) (EC042\_4803, *flu1* y *flu2*) se encontraron sobreexpresados. Este hecho que sugiere que la proteína AggR podría estar implicada en la regulación de la motilidad y la agregación celular (Yasir y col., 2019).

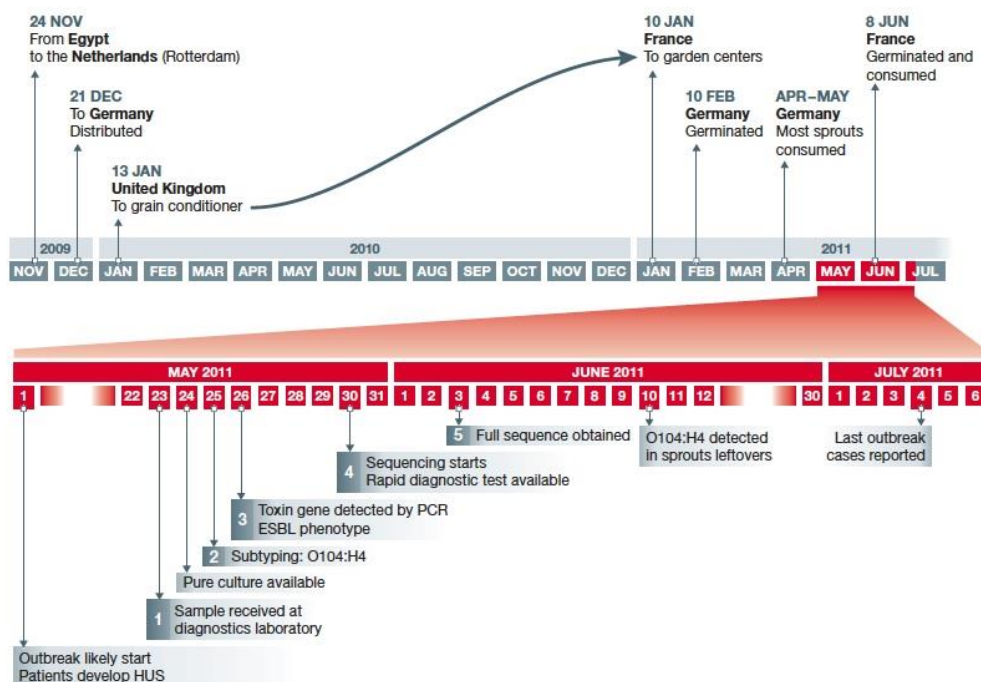
En este mismo trabajo también se determinó la arquitectura de los promotores regulados por AggR y la secuencia consenso de unión de esta proteína al ADN. Para establecer esta secuencia consenso se analizaron con detalle los dos *clusters* génicos responsables de la biosíntesis de las fimbrias AAF/II (*aafDA* y *aafCB*) en la cepa *E. coli* 042. Dicha secuencia consenso también se encontró presente en los promotores de los genes que codifican para las otras 4 clases de fimbrias presentes en los genomas de las cepas EAEC (AAF/I, AAF/III, AAF/IV y AAF/V). Globalmente se pudo determinar que la secuencia de unión de AggR responsable de la expresión de los promotores regulados por esta proteína se encuentra localizada entre 21-22 nucleótidos aguas arriba de la caja -10 y está formada por una caja TATC altamente conservada flanqueada aguas arriba por una secuencia rica en A/T (Yasir y col., 2019) (figura 1.10).



**Figura 1.10.** (A) Secuencia nucleotídica consenso de unión de la proteína AggR al ADN. Dicha secuencia está formada por una caja TATC flanqueada por una secuencia rica en A y T. (B) Organización estructural de los promotores regulados por el factor transcripcional AggR. Se indica el lugar de unión de dicha proteína separado unos 21-22 nucleótidos de la caja -10 del gen regulado. Imagen adaptada de (Yasir y col., 2019).

### 1.1.4 Un ejemplo de la heterogeneidad genómica de las cepas EAEC: el brote en 2011 de SUH en Alemania

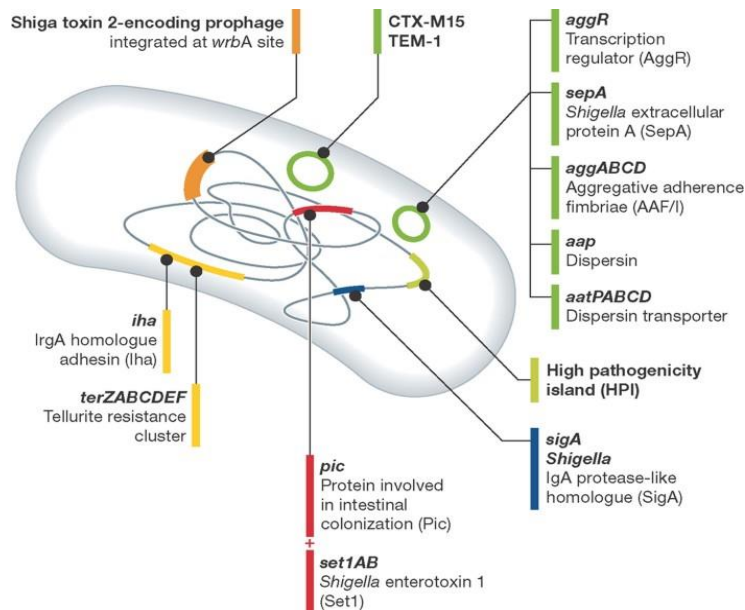
Una característica de las cepas enteroagregativas de *E. coli* es su elevada heterogeneidad genómica. Ello implica que muchos de los factores de virulencia característicos de este patotipo están ausentes en muchas cepas EAEC, lo cual ha dificultado hasta la fecha la comprensión de los mecanismos implicados en la patogénesis de estas bacterias. Un ejemplo de la gran variabilidad genómica de las cepas EAEC hace referencia al brote de intoxicación alimentaria (cuyo origen fueron unas semillas de soja contaminadas) que tuvo lugar en 2011 en Alemania (figura 1.11). El agente etiológico responsable de dicho brote fue una cepa de *E. coli* productora de toxina Shiga. Este brote infeccioso se saldó con 3.842 pacientes hospitalizados, de los cuales 855 desarrollaron síndrome urémico hemolítico (SUH) que provocó una elevada mortalidad (54 personas fallecieron) (Buchholz y col., 2011).



**Figura 1.11.** Evolución temporal de los sucesos que desencadenaron el brote de SUH en Europa en la primavera de 2011. En la línea de arriba (color gris) se representa el viaje de las semillas sospechosas desde los campos donde se cultivaron hasta las instalaciones en Alemania y Francia donde germinaron. En la línea inferior (a escala ampliada el fragmento de calendario en color rojo, que se extiende durante algo más de dos meses) se representan la serie sucesos ocurridos durante el brote epidémico. Imagen adaptada de (Karch y col., 2012).

## 1. INTRODUCCIÓN

La cepa responsable de esta gran intoxicación, serotipada como O104:H4, presentaba el patrón de adherencia a células Hep-2 en forma de apilado de ladrillos (patrón característico del patotipo EAEC), además de dar positivo para los genes más representativos de dicho patotipo, como *aggR*, *aggA*, *set1*, *pic* y *aap* (Bielaszewska y col., 2011). Notoriamente, la cepa causante de este brote también contenía un profago que codificaba el gen de la toxina 2 de Shiga, *stx2*, factor de virulencia característico del patotipo STEC pero totalmente ausente en el cromosoma de las cepas EAEC (figura 1.12).

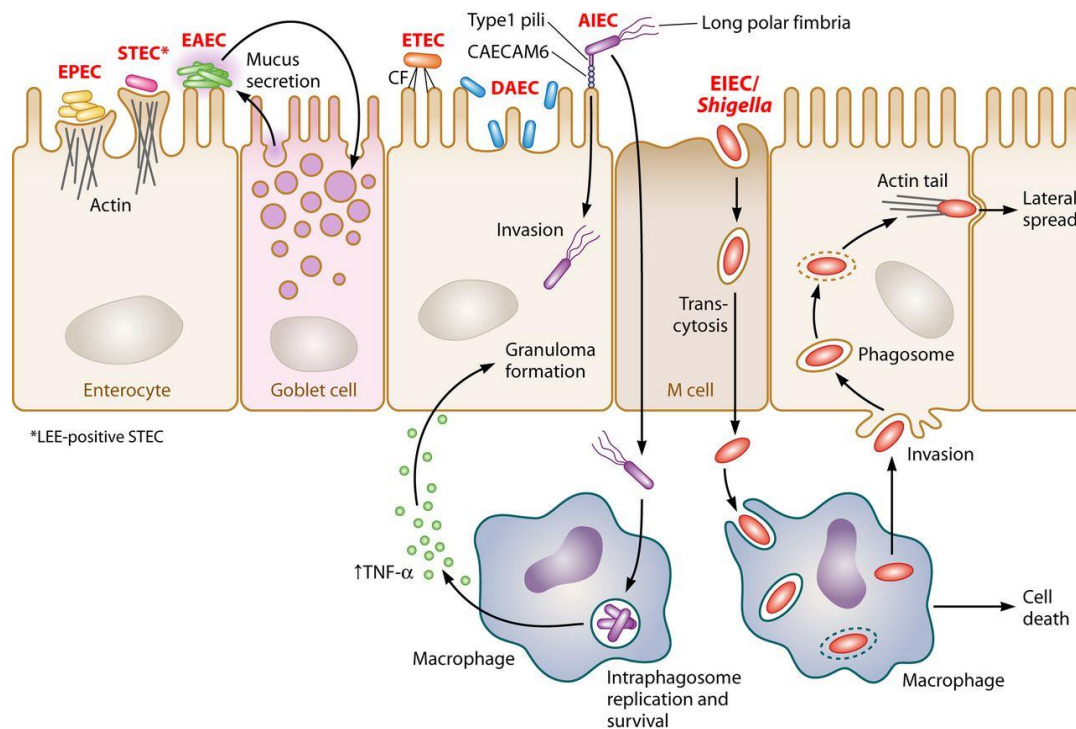


**Figura 1.12.** Representación esquemática de los genes codificados por la cepa *E. coli* O104:H4 (EAEC/STEC), cepa responsable del brote de SUH en Europa durante el año 2011. Se destacan los genes: *stx2* (que codifica para la toxina Stx2); *pic*, *sigA* y *sepA* (que codifican para las proteínas SPATE); Pic, proteína implicada en la colonización intestinal; SigA, un homólogo de Pet, con actividad citotóxica; SepA, (un factor de colonización de *Shigella*), *set1AB* (que codifica ShET1, una holotoxina AB5); *iha* (que codifica para Iha, una adhesina STEC que es un homólogo de IrgA); el regulador transcripcional AggR; el operón *aggABCD*, que codifica para las fimbrias de tipo I; la dispersina (codificada en el gen *aap*) y su sistema transportador (codificado en el operón *aatPABCD*); *lpf1-2* (que codifica para la proteína Lpf de STEC); *terZABCDE* (operón que codifica para un grupo de genes que aportan resistencia al telurito) y *CTX-M15* y *TEM-1* (genes de resistencia a antibióticos). Imagen adaptada de (Karch y col., 2012).

A lo largo de la historia también se han reportado otros brotes (aunque de menor intensidad) provocados por cepas híbridas EAEC/STEC. En 1992 tuvo lugar en Francia un brote con diez casos de SUH. Este brote se asoció a otra cepa EAEC, cuyo serotipo era O111:H2, que también contenía el gen *stx2* (Boudailliez y col., 1997). En 1999, se diagnosticó en Japón otro caso de SUH asociado a una cepa EAEC (cuyo serotipo era O86:HNM) que también codificaba la toxina Stx2 (Iyoda y col., 2000). En 2012, tuvo



lugar otro pequeño brote infeccioso de SUH en Irlanda. En este caso, la cepa responsable también fue una cepa EAEC (serotipada como O111:H21) que también era portadora de la toxina Stx2 (Dallman y col., 2012). Todos estos brotes muestran el elevado potencial de las cepas EAEC para adquirir elementos de virulencia adicionales a través del intercambio genético, poniendo de manifiesto la elevada heterogeneidad de este patotipo bacteriano.



**Figura 1.13.** Representación esquemática de los patrones de adherencia y los mecanismos de virulencia de las cepas diarreas de *E. coli*. *E. coli* enteropatógena (EPEC) (representada en amarillo) y *E. coli* productora de toxina Shiga LEE-positiva (STEC/EHEC) (representada en rosa), son patógenos extracelulares que se unen al epitelio intestinal y provocan lesiones en las microvellosidades, generando las lesiones características A/E. Debido a la presencia del pili BFP, EPEC es capaz de formar microcolonias, dando como resultado un patrón de adherencia localizada (LA). *E. coli* enterotoxigénica (ETEC) (representada en naranja) utiliza factores de colonización (CF) para la unión a las células intestinales del huésped. *E. coli* enteroagregativa (EAEC) (representada en verde) forma *biofilms* en la mucosa intestinal, donde las bacterias se adhieren entre sí para formar un patrón de adherencia agregativo (AA) conocido como "apilado de ladrillo". *E. coli* con adherencia difusa (DAEC) (representada en azul) se dispersa sobre las superficies de las células intestinales, dando lugar a un patrón de adherencia difusa. *E. coli* invasiva adherente (AIEC) (representada en púrpura) coloniza las mucosas intestinales de los pacientes con enfermedad de Crohn y es capaz de invadir las células epiteliales y de replicarse dentro de los macrófagos. AIEC utiliza el pili tipo I para adherirse a las células intestinales, y las fimbrias polares largas, que contribuyen a la invasión. *E. coli* enteroinvasiva (EIEC)/*Shigella* (representada en rojo), son patógenos intracelulares que penetran el epitelio intestinal a través de las células M para obtener acceso a la submucosa. Además, escapan de los macrófagos de la submucosa por inducción de muerte celular de los mismos para invadir basolateralmente las células epiteliales del colon humano. Imagen adaptada de (Croxen y col., 2013).

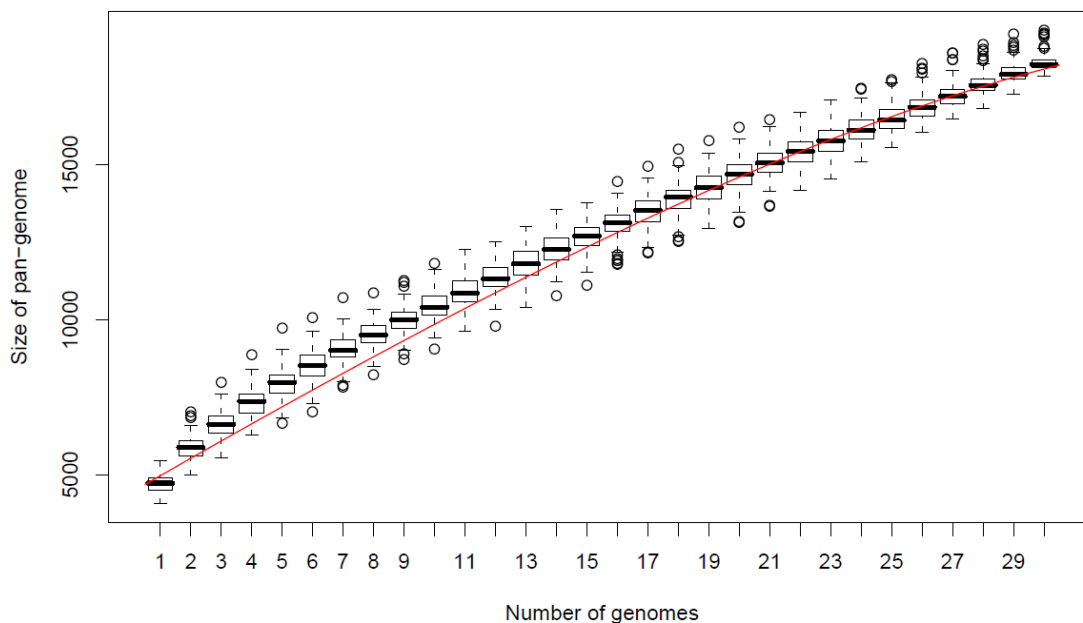
## 1. INTRODUCCIÓN

En la figura 1.13 se representa de forma esquemática y resumida los diferentes procesos implicados en la unión a las células epiteliales intestinales por parte de las cepas pertenecientes a los diferentes patotipos diarreicos de *E. coli* y los procesos que posteriormente desencadenan en la evolución del proceso infeccioso.

### **1.1.5 Genómica y filogenia de *E. coli***

Los tamaños de los genomas de las diferentes cepas de *E. coli* pueden diferir en más de un millón de pares de bases entre las variantes comensales y las patógenas. Mientras que las cepas de laboratorio estándar tienen genomas de ~4,5 millones de pares de bases y 4.000 genes, las cepas patógenas de *E. coli* pueden tener genomas de más de 5,9 millones de pares de bases y 5500 genes (Blattner y col., 1997; Lukjancenko y col., 2010; de Muinck y col., 2013).

Al igual que en otras especies, en las cepas de *E. coli* se consideran un conjunto de genes compartidos y conservados, llamado núcleo del genoma, y un grupo de genes flexibles que se encuentran presentes de forma variable entre las diferentes cepas, denominado pan-genoma. De los 16.400 genes que forman parte del pan-genoma de *E. coli* (conjunto sumatorio de todos los genes diferentes presentes en el total de cepas de *E. coli*) (figura 1.14), tan sólo 2000 genes son genes homólogos y comunes a todas las cepas (Kaas y col., 2012). Por ello, no es de extrañar que, a pesar de formar parte de la misma especie bacteriana, las cepas *E. coli* K-12 (comensal), *E. coli* O157:H7 (enterohemorrágica) y *E. coli* CFT073 (uropatogénica) únicamente compartan un 39,2% de los genes (Welch y col., 2002).

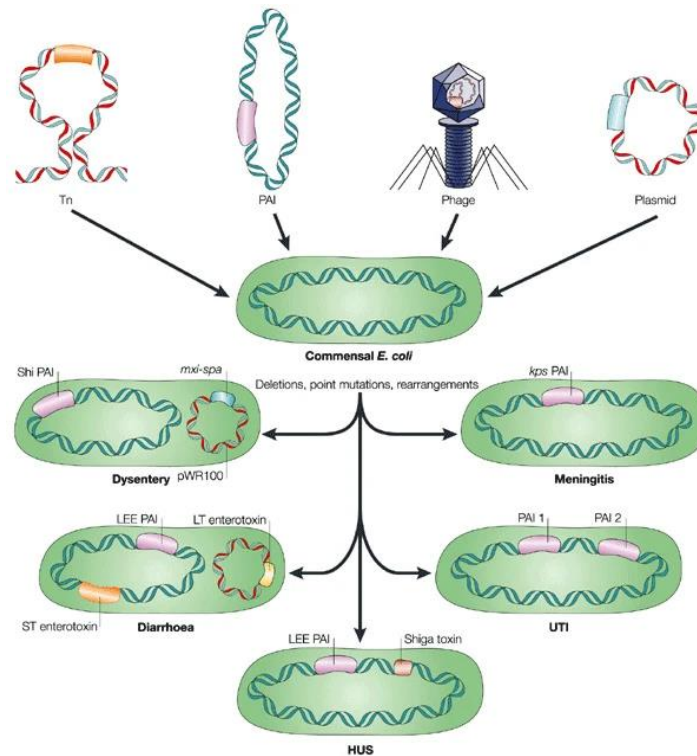


**Figura 1.14.** Representación mediante *Box-plot* del tamaño del pan-genoma de *E. coli* en función del número de genomas secuenciados. A medida que se van secuenciando nuevos genomas de *E. coli*, nuevos genes se van añadiendo al conjunto total de genes diferentes encontrados en esta especie. Imagen adaptada de (Chaudhuri y col., 2010).

La elevada y variada capacidad patogénica de *E. coli* de causar diferentes tipos de infecciones deriva en gran medida del conjunto de genes flexibles adquiridos a través de mecanismos de ganancia y pérdida de material genético en varios puntos críticos en todo el genoma (Touchon y col., 2009). La gran mayoría de los factores de virulencia que se encuentran en las cepas patógenas de *E. coli* se hallan asociados a elementos genéticos móviles. Estos elementos genéticos, como los transposones, las secuencias de inserción, los bacteriófagos y los plásmidos, pueden existir integrados en el cromosoma o como elementos externos que se mantienen mediante autorreplicación dentro del nuevo huésped, aportando nuevos genes que proporcionan nuevos rasgos, características y ventajas evolutivas (figura 1.15). El impacto de estos mecanismos de transferencia de nuevos genes es tan grande que, por ejemplo, en el caso de la cepa *E. coli* 042, se estima que casi una cuarta parte de su contenido genómico está formado por islas genómicas adquiridas por transferencia horizontal (Chaudhuri y col., 2010).

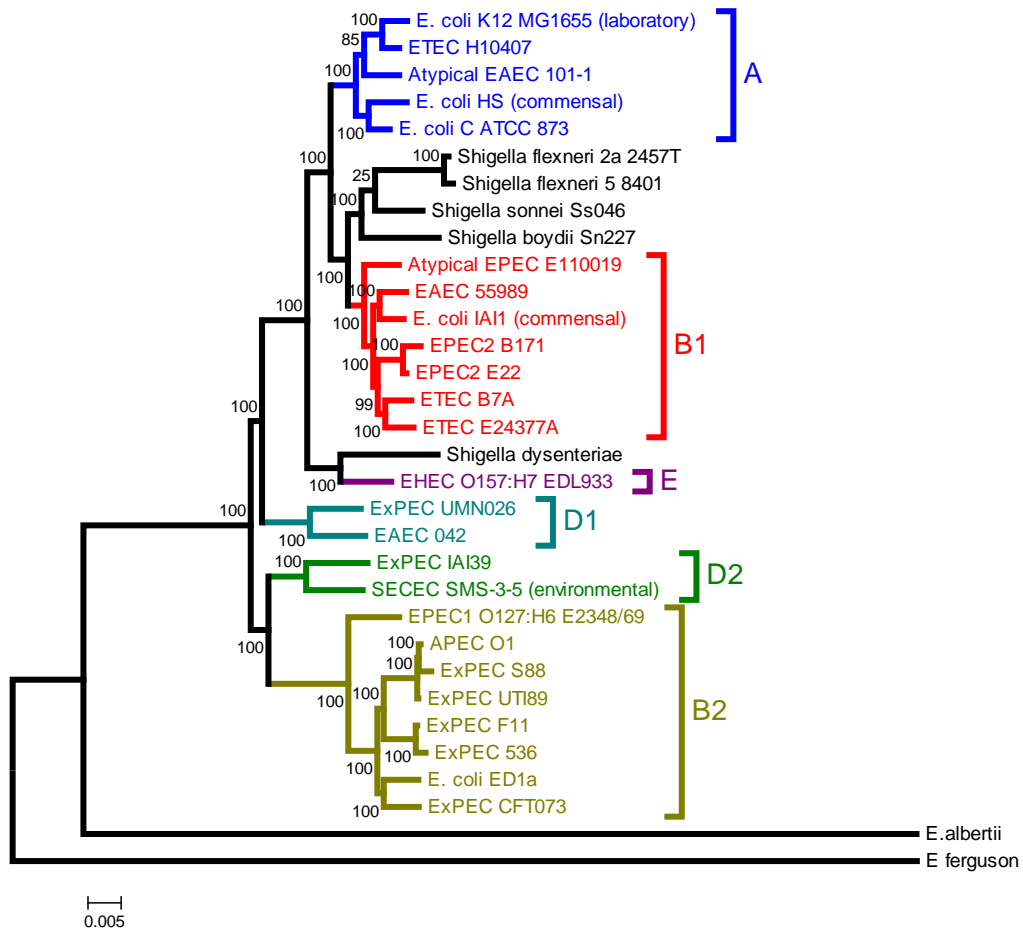


## 1. INTRODUCCIÓN



**Figura 1.15.** Contribución de los elementos genéticos móviles en la evolución de las cepas patógenas de *E. coli*. Los factores de virulencia de *E. coli* pueden estar codificados por varios elementos genéticos móviles, incluidos los transposones (*Tn*) (por ejemplo, la enterotoxina estable al calor (ST) de ETEC), plásmidos (por ejemplo, la enterotoxina termolábil (LT) de ETEC y los factores de invasión de EIEC), bacteriófagos (por ejemplo, la toxina Shiga de EHEC) e islas de patogenicidad (PAI), como el *locus* de adhesión y borrado de enterocitos (LEE) de EPEC/EHEC. Además, a nivel intracelular, grandes deleciones en el genoma, mutaciones puntuales u otros reordenamientos de ADN también pueden contribuir a la virulencia bacteriana. Dichos cambios genéticos pueden dar lugar a la aparición de cepas patógenas de *E. coli* capaces de causar diarrea, disentería (EIEC), síndrome urémico hemolítico (EHEC y STEC-EAEC), infecciones del tracto urinario (UPEC) y meningitis (MNEC). Imagen adaptada de (Kaper y col., 2004).

A nivel poblacional, las diferentes cepas de *E. coli* se pueden clasificar en 5 grandes grupos filogenéticos (A, B1, B2, D y E) (figura 1.16). La gran mayoría de cepas comensales se agrupan principalmente en el filogrupo A; sin embargo, no todas las cepas pertenecientes a un mismo patotipo de *E. coli* se agrupan juntas: las cepas del patotipo ETEC se agrupan en los grupos filogenéticos A y B1 (Sahl y col., 2011), mientras que las cepas del patotipo EAEC lo hacen en los grupos filogenéticos A, B1, B2 y D (Okeke y col., 2010). Este hecho demuestra la gran diversidad y heterogeneidad de los genotipos de las diferentes cepas de *E. coli* patógenas a pesar de pertenecer a un mismo patotipo.



**Figura 1.16.** Relaciones filogenéticas entre los genomas secuenciados de diferentes cepas de *E. coli* y *Shigella*. Las cepas de *E. coli* se pueden agrupar en 5 grupos filogenéticos principales: A (azul), B1 (rojo), B2 (verde claro), D (azul turquesa) y E (morado). *Shigella*/EIEC también forman un filogrupo adicional (negro). Las cepas de un mismo patotipo no siempre se agrupan en el mismo grupo filogenético. Los genomas de *Escherichia albertii* y *Escherichia fergusonii* se incluyen como secuencias de grupos externos. Imagen adaptada de (Chaudhuri y col., 2010).

Toda esta pluralidad presente en los genomas de *E. coli* se hace patente a medida que se van secuenciando más y más genomas de esta especie bacteriana. Por todo ello, dada la gran variabilidad genómica que presentan estas bacterias, se hace difícil clasificar una cepa únicamente a través de su patotipo. Sin embargo, a pesar de la dificultad que ello supone, es informativo considerar los tipos de patología principales como un marco clasificatorio para tener una visión general de las cepas patógenas *E. coli*.

### **1.2 Proteínas asociadas al nucleoide como moduladores globales de la expresión génica bacteriana**

Una de las características más importantes de las células procariotas es la ausencia de un núcleo celular definido. Sin embargo, a pesar de que las bacterias carecen de un núcleo celular envuelto por una membrana, su material genético (ADN cromosómico circular) y los tipos de macromoléculas asociados al mismo se encuentran compactados en un gran complejo (similar a la cromatina característica del núcleo celular eucariota) denominado nucleoide bacteriano. Dicho complejo es esencial para permitir que las grandes dimensiones del ADN cromosómico (en comparación con el tamaño de la célula que lo contiene) pueda confinarse en una estructura que sea compatible con su replicación, segregación, expresión y regulación génica.

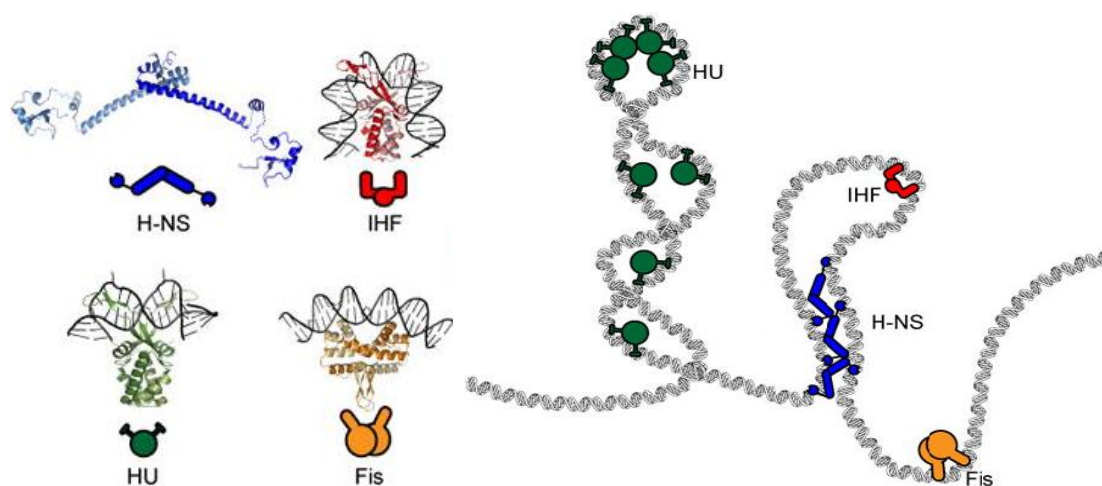
Los principales factores que permiten que el cromosoma bacteriano de *E. coli* (4,6 Mb y 1,6 mm de longitud, en el caso de la cepa de laboratorio K-12) ocupe únicamente una cuarta parte del volumen intracelular (Dame, 2005) son el superenrollamiento del ADN (Trun y Marko, 1998), el comportamiento polimérico del mismo (Pelletier y col., 2012) y las proteínas asociadas al nucleoide (PANs) (Dillon y Dorman, 2010).

Las PANs son a menudo reconocidas como proteínas tipo histona ya que poseen propiedades similares a sus homólogos eucariotas tales como la carga eléctrica, la capacidad de unión al ADN y su elevada abundancia a nivel intracelular. Poseen como características comunes el hecho de que sus unidades monoméricas presentan un peso molecular relativamente bajo, así como la capacidad de las mismas de formar dímeros u oligómeros de elevado grado de complejidad. Además, la mayoría de las PANs tienen la capacidad de alterar la conformación del ADN: son capaces de flexionarlo, envolverlo y conformar puentes de ADN-proteína-ADN (Dillon y Dorman., 2010). No obstante, en comparación con las histonas eucariotas, las PANs interactúan con el ADN cromosómico de forma más lábil.

Además de influir en la arquitectura del ADN dado su rol estructural permitiendo compactar el ADN, las PANs también ejercen una función moduladora de la expresión génica, puesto que muestran influencia sobre la replicación, transcripción y traducción de una gran variedad de genes implicados, en muchos casos, en la respuesta a estímulos ambientales (por ejemplo, cambios en la osmolaridad, la temperatura o la disponibilidad

de nutrientes) o genes implicados en virulencia (McLeod y Johnson, 2001). Por todo ello, el nucleoide bacteriano es una estructura dinámica de vital importancia para la célula.

En general, las PANs se han estudiado principalmente en bacterias gramnegativas. Las principales son: H-NS (*Histone-like Nucleoid-Structuring protein*), CRP (*Cyclic AMP receptor protein*), IHF (*Integration Host Factor*), Fis (*Factor Inversion Stimulation*) y HU (*Heat-Unestable Nucleoid Protein*) (Dame, 2005) (figura 1.17). La unión de las PANs al ADN puede producirse de diversas maneras, ya sea reconociendo una secuencia consenso, como en el caso de las proteínas CRP, Fis e IHF, o bien reconociendo dominios estructurales. En este último caso, las PANs pueden unirse a regiones de ADN curvadas ricas en adenina (A) y en timina (T), como en el caso de H-NS y HU (Dillon y Dorman, 2010). En general, debido a la sustancial afinidad inespecífica de unión al ADN, se estima que la mayoría de PANs celulares se encuentran unidas al ADN cromosómico con una cobertura de aproximadamente una PAN cada 100 pares de bases de ADN (Li y col., 2009).



**Figura 1.17.** Representación esquemática de las proteínas asociadas al nucleoide H-NS, IHF, HU y Fis con capacidad unir, envolver o flexionar al ADN, contribuyendo así a la organización del cromosoma bacteriano. H-NS en forma de dímeros (representados en azul) es capaz de unirse y formar puentes en el ADN. HU (representada en verde) es capaz de introducir una flexión de  $\sim 90^\circ$  al ADN y es capaz de envolverlo alrededor de sí mismo, promoviendo así interacciones de ADN de corto alcance. La unión de IHF (representado en rojo) al ADN induce un cambio de sentido en el ADN que cambia la trayectoria del mismo. Fis (representada en naranja) es otra proteína asociada al nucleoide con capacidad de unir y flexionar el ADN. Imagen adaptada de (Badrinarayanan y col., 2015).

### **1.2.1 La proteína H-NS (*Histone-like nucleoid-structuring protein*)**

La proteína H-NS es una proteína con capacidad de unión al ADN ampliamente distribuida entre bacterias, mayoritariamente gramnegativas. Esta proteína se une a regiones del ADN ricas en A/T (Grainger y col., 2006). H-NS ha sido estudiada y caracterizada de forma intensiva en las enterobacterias, principalmente en *E. coli* y en *Salmonella*, aunque también se expresa de forma generalizada en otros géneros y especies cercanos dentro de la subdivisión  $\gamma$  de las proteobacterias. Sin embargo, también existen análogos de la proteína H-NS en especies bacterianas más alejadas evolutivamente a las enterobacterias, como en el caso de la proteína BpH3 de *Bordetella pertussis* (Goyard y Bertin, 1997), MvaT/U dentro del género *Pseudomonas* (Castang y col., 2008), Lsr2 en *Mycobacteria* (Gordon y col., 2010) y Rok, en el género *Bacillus* (Smits y Grossman., 2010). Globalmente, la proteína H-NS, sus parálogos, ortólogos y otras proteínas equivalentes son considerados como una familia de proteínas.

H-NS es una proteína termoestable de pequeño tamaño (137 aminoácidos, con una masa molecular de 15,6 kDa). En *E. coli*, H-NS es una proteína muy abundante intracelularmente, con un nivel aproximado de 20.000 copias de la molécula por célula (Azam y Ishihama, 1999). Dicha proteína fue caracterizada por primera vez en los años 70 como una proteína asociada al nucleoide bacteriano con capacidad de estimular la transcripción *in vitro* del ADN en condiciones limitantes de ARN polimerasa (Jacquet y col., 1971; Cukier-Kahn y col., 1972) (en esta primera publicación H-NS se nombró inicialmente como H1). La proteína H-NS se pudo aislar por primera vez siguiendo un protocolo utilizado para aislar proteínas tipo histona en eucariotas (Varshavsky y col., 1977).

La gran mayoría de bacterias únicamente contienen una única copia cromosómica del gen *hns*. Sin embargo, existen otras especies bacterianas como *Shigella flexneri* y *Pseudomonas putida* que pueden llegar a codificar en su genoma hasta tres y cinco genes tipo *hns*, respectivamente (Tendeng y Bertin, 2003).

En *E. coli*, los niveles de expresión de la proteína H-NS se mantienen relativamente constantes a lo largo de las distintas fases del crecimiento bacteriano (Schröder y Wagner, 2002). Todo ello es debido a que existe una correlación positiva entre la transcripción del gen *hns* y la replicación del cromosoma (Free y Dorman, 1995). Además, a pesar de que

H-NS se expresa de manera similar tanto a 25°C como a 37°C, afecta de manera diferencial a la expresión de los genes que regula en función de la temperatura de crecimiento de la bacteria (Göransson y col., 1990). El gen *hns* se transcribe como un ARNm monocistrónico independientemente de los genes que lo flanquean y se autorregula negativamente a nivel de la transcripción (Falconi y col., 1993). La autorrepresión de *hns* por su propio producto génico se ejerce de manera muy intensa cuando se detiene el movimiento de la horquilla de replicación cromosómica (ya sea por cuestiones del ciclo celular o por fármacos inhibitorios de la replicación celular), lo que sugiere que la transcripción de *hns* es sensible a la progresión del ciclo celular bacteriano (Free y Dorman, 1995). También se ha reportado que las proteínas asociadas al nucleóide Fis y CRP pueden activar la expresión del gen *hns* (Falconi y col., 1996), junto con la proteína CspA en condiciones de choque frío (La Teana y col., 1991), además de por un incremento en la presión hidrostática del medio (Welch y col., 1993). La transcripción de *hns* también es sensible al regulador dependiente de hierro Fur (Troxell y col., 2011). Además, la expresión del gen *hns* también se encuentra negativamente regulada a un nivel post-transcripcional por parte de *dsrA*, un pequeño ARN regulador dependiente de Hfq (Majdalani y col., 2005). Todo ello sugiere que, bajo determinadas condiciones ambientales, los niveles de expresión de la proteína H-NS pueden variar.

La proteína H-NS ejerce un papel esencial en la estabilidad del cromosoma bacteriano. Estudios iniciales sugirieron que H-NS se disponía de una manera más o menos uniforme sobre el cromosoma (Azam y col., 2000; Atlung y Hansen, 2002). Sin embargo, en estudios posteriores en los que se utilizaron técnicas de microscopía de alta resolución (microscopía de localización fotoactivada y microscopía de reconstrucción estocástica) para determinar la localización subcelular de H-NS, se pudo observar como la proteína se dispone agrupada en forma de unos pocos *clusters* a lo largo del cromosoma, a diferencia de otras PANs, como HU, Fis o IHF, que se distribuyen de una manera más homogénea a lo largo del mismo (Wang y col., 2011).

A diferencia de algunos reguladores transcripcionales que reconocen secuencias de nucleótidos específicas para poder unirse al ADN, la unión de la proteína H-NS al ADN se consideró inicialmente como una unión independiente de una secuencia nucleotídica específica. La proteína H-NS presenta preferencias de unión a regiones curvadas de ADN de doble cadena ricas en A/T (Owen-Hughes y col., 1992). En *E. coli*, existe un gran número de genes que presentan estructuras curvadas en sus respectivos promotores

## 1. INTRODUCCIÓN

(Bracco y col., 1989). No obstante, en 2007 se identificaron 2 motivos de unión idénticos de alta afinidad (TCGATATATT) en el gen *proU* de *E. coli* (Bouffartigues y col., 2007). Paralelamente, se identificó una secuencia consenso de unión al genoma de 10 pb por parte de H-NS (5'-tCG(a/t)T(a/t)AATT-3'). Esta secuencia está ampliamente distribuida por todo el genoma de *E. coli*, especialmente en la región promotora de un gran número de genes que presentan mecanismos de regulación por parte de H-NS (Lang y col., 2007). En 2009, mediante análisis de DNAsa *footprinting* y espectroscopía de resonancia magnética nuclear (RMN), se logró identificar otra secuencia de también 10 pb (TCCTTACATT) a la que el extremo C-terminal de H-NS se unía con una alta afinidad (Sette y col., 2009). También se han identificado otros dos sitios de unión de H-NS dentro del promotor de LEE5 (*locus of enterocyte effacement 5*), presente en la isla de patogenicidad de *E. coli* enteropatógena (EPEC) (Bhat y col., 2014). Estos dos sitios de unión se encuentran en las posiciones -110 (TCGATAATTA) y -138 (TGCATATAAA) del promotor de LEE5 y presentan una elevada similitud con la secuencia consenso descrita anteriormente y referenciada en (Lang y col., 2007). Todas estas secuencias presentan un alto porcentaje de A y T, lo que pone de manifiesto la alta afinidad de esta proteína por estas regiones ricas en dichos nucleótidos.

La principal función reguladora de la proteína H-NS y otros miembros de dicha familia de proteínas es reprimir la transcripción (Dorman, 2007). De manera específica, debido a que H-NS se une a regiones del cromosoma ricas en A/T, la represión transcripcional que ejerce H-NS va dirigida principalmente contra el ADN adquirido de manera horizontal (cuyo contenido en A y T es significativamente superior al contenido en A/T del *core* del genoma propio de la bacteria) (Lucchini y col., 2006, Navarre y col., 2006, Oshima y col., 2006). Por todo ello, H-NS es una proteína clave para mantener la homeostasis celular, ya que la expresión descontrolada del ADN adquirido de manera horizontal puede dar lugar grandes alteraciones fisiológicas y a la consecuente pérdida del *fitness* celular (Ali y col., 2014).

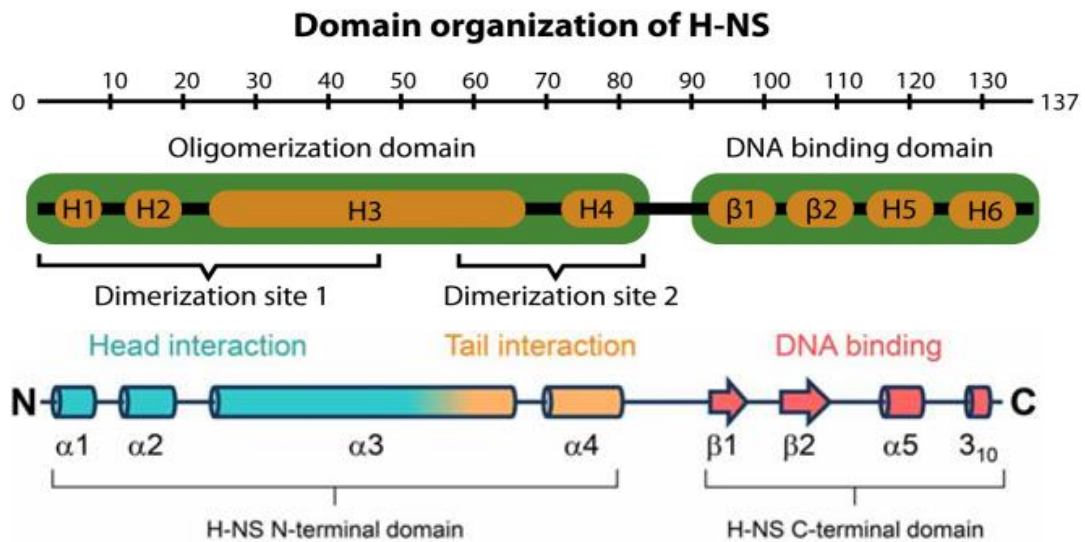
### **1.2.1.1 Dominios estructurales de la proteína H-NS**

#### **1.2.1.1.1 Dominio N-terminal**

La proteína H-NS está formada por dos dominios estructurales unidos por un *linker* flexible. El dominio N-terminal está compuesto por 83 aminoácidos y está formado por 4 hélices alfa, responsables de la oligomerización de la proteína y la formación de



estructuras filamentosas (Arold y col., 2010). Las dos primeras hélices  $\alpha$  son de pequeño tamaño, H1 (residuos 2 al 7) y H2 (residuos del 11 al 18), seguidas de una tercera  $\alpha$ -hélice mayor, H3 (residuos del 22 al 67). Estas tres hélices forman la interfaz de dimerización de H-NS. La cuarta  $\alpha$ -hélice, H4 (residuos del 72 al 83), junto con la parte distal de la H3, dan lugar a un segundo sitio de dimerización (residuos 57 al 83) con una conformación estructural de hélice giro hélice (Leonard y col., 2009; Arold y col., 2010; Ali y col., 2012) (figura 1.18). Además, los 20 primeros aminoácidos del dominio N-terminal también están implicados en la capacidad represora o de silenciamiento génico que ejerce la proteína H-NS. Adicionalmente, esta región participaría, junto con el dominio C-terminal, en el reconocimiento de las regiones curvadas de ADN y por ende en la unión a los ácidos nucleicos (Dorman, 2004).



**Figura 1.18.** Representación esquemática de la estructura secundaria y de la organización de los dominios estructurales de la proteína H-NS. En el extremo N-terminal se encuentra el dominio de oligomerización e interacción con otras proteínas, seguido de la región flexible, que conecta los dos dominios funcionales de H-NS y les permite actuar de manera independiente. En el extremo C-terminal se encuentra el dominio de unión al ADN. H y  $\beta$  representan  $\alpha$ -hélices y láminas- $\beta$ , respectivamente. Se representan también los dos sitios de dimerización dentro del dominio de oligomerización de la proteína (Adaptado de Winardhi y col., 2015; Grainger, 2016).

Aunque se trata de una región bastante conservada entre los distintos miembros de la familia de proteínas H-NS, el grado de conservación del dominio N-terminal es inferior al grado de conservación del dominio C-terminal de unión a ácidos nucleicos (Tendeng y Bertin, 2003). Esta elevada variabilidad del dominio N-terminal de la proteína H-NS está determinada por las diferentes necesidades adaptativas que presentan las bacterias que codifican el gen *hns*, como pueden ser los mecanismos adaptativos a baja temperatura



## 1. INTRODUCCIÓN

(Tendeng y col., 2003), o la diferente especificidad o afinidad de interacción con otras proteínas o el ADN (Bloch y col., 2003).

### 1.2.1.1.2 Dominio C-terminal

El dominio C-terminal de la proteína H-NS (situado entre los aminoácidos 90-137) es el dominio responsable de la interacción con el ADN y presenta una secuencia de aminoácidos altamente conservada en la familia de proteínas tipo H-NS. Dicho dominio presenta una secuencia consenso (TWTGX<sub>1</sub>GX<sub>2</sub>X<sub>3</sub>P), correspondiente a los aminoácidos del 108 y al 116 de la proteína. En *E. coli* y *Salmonella*, los residuos X<sub>1</sub> y X<sub>2</sub> son glutamina y arginina, respectivamente. Experimentos de resonancia magnética nuclear han demostrado que el motivo glutamina-glicina-arginina (QGR) es capaz de insertarse en el surco menor del ADN rico en AT (Gordon y col., 2011). Este dominio de unión al ADN con preferencia en regiones ricas en A y T, denominado *AT-hook motif*, está presente en un gran número de proteínas, como por ejemplo en las proteínas eucariotas HMG (*high mobility group*) (Aravind y Landsman, 1998; Cui y Leng, 2007) y en proteínas de unión al ADN en plantas (Meijer y col., 1996).

A nivel de estructura secundaria, dicho dominio estructural está formado por 2 láminas  $\beta$  ( $\beta$ <sub>1</sub> y  $\beta$ <sub>2</sub>), una hélice  $\alpha$  ( $\alpha$ <sub>5</sub>), y una hélice  $3_{10}$  (Gordon y col., 2011) (esta última es un motivo de estructura secundaria poco común entre las proteínas) (figura 1.18). La región clave responsable de la interacción con el ADN es un bucle presente entre la lámina  $\beta$ <sub>2</sub> y la hélice  $\alpha$  que forma el motivo estructural *AT-hook* que permite la unión de la proteína al surco menor del ADN en regiones ricas en A y T.

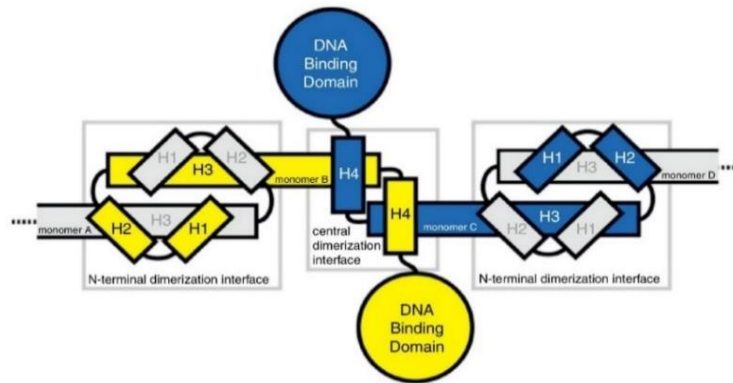
### 1.2.1.2 Oligomerización de la proteína H-NS

Al igual que sucede con otros reguladores globales de la expresión génica, los monómeros de la proteína H-NS oligomerizan en estructuras más complejas para dar lugar a la forma activa de la molécula necesaria para poder realizar sus funciones biológicas. Este proceso de oligomerización se inicia con la formación de la unidad estructural dimérica, y depende de la concentración de proteína efectiva, del tipo de catión presente en el medio y de la temperatura (Smyth y col., 2000). A continuación, se forman los oligómeros, cuya formación recae en la región central de la proteína (residuos 46 al 90) (Bloch y col., 2003). La estructura de H-NS en forma oligomérica se resolvió en 2010 (Arold y col., 2010), donde se determinó la autoasociación del dominio de oligomerización (aminoácidos 1-83) de la proteína H-NS de *S. Typhimurium*. A través de una dimerización simultánea en

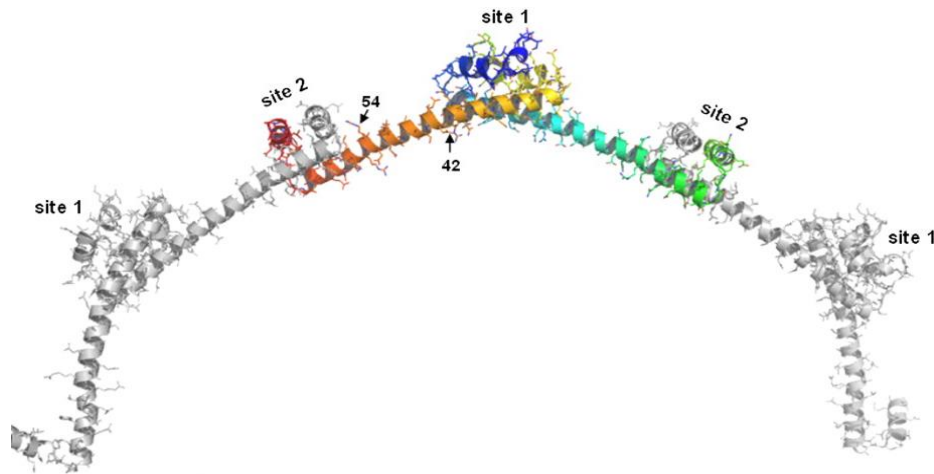
las dos interfaces de los dos sitios de dimerización del dominio N-terminal, la proteína H-NS va oligomerizando formando filamentos en una conformación *head to head/tail to tail*, conformación en la que los monómeros que forman los dímeros de H-NS se orientan de una manera antiparalela, lo que conlleva la creación de andamios superhelicoidales (Arold y col., 2010; Ali y col., 2012) (figura 1.19). Estudios con moléculas individuales de H-NS en solución han demostrado que la proteína es capaz de oligomerizar de manera espontánea en ausencia de ADN. Como se comentaba anteriormente, la oligomerización de la proteína depende de la concentración de la misma. Cuando la concentración de la proteína H-NS es de 10  $\mu\text{M}$ , se ve favorecida la formación de dímeros; por el contrario, cuando la concentración de H-NS es muy elevada (344  $\mu\text{M}$ ), H-NS es capaz de formar oligómeros de hasta 20 unidades de proteína (Spurio y col., 1997; Smyth y col., 2000). Para la formación de estas estructuras de orden superior, además del dominio de oligomerización, también es indispensable la región conectora (*linker*, residuos 64 al 89), ya que proteínas truncadas en las que se había deleciónado la región conectora no eran capaces de formar oligómeros mayores (Smyth y col., 2000), viéndose afectada también la unión cooperativa de H-NS al ADN (Giangrossi y col., 2014).

El dominio de oligomerización de H-NS, además de permitir la interacción entre un monómero de la proteína con otro idéntico, también permite la interacción con otras proteínas que presenten este mismo dominio o dominios parecidos, hecho que afectará a la capacidad moduladora de la expresión génica por parte de H-NS (Dorman y col., 1999). La capacidad reguladora diferencial que presenta la proteína H-NS puede modificarse en función de las proteínas con las que interactúe. Inicialmente, el heterómero mejor caracterizado de H-NS con otra proteína ha sido el que se establece con su proteína paróloga StpA (Williams y col., 1996; Free y col., 1998; Dillon y Dorman, 2010; Dorman, 2014) (ver apartado 1.2.1.8). La proteína StpA es susceptible a la degradación por parte de proteasas de la familia Lon, especialmente en ausencia de interacción con H-NS. En este proceso de formación de heterodímeros H-NS-StpA, H-NS proporciona a StpA un mecanismo de protección contra la degradación por parte de estas proteasas (Johansson y Uhlin, 1999; Johansson y col., 2001). Además, la proteína H-NS también puede interactuar, entre otros, con el regulador transcripcional y chaperona de ARN Hfq (Brescia y col., 2003; Dole y col., 2004), la proteína flagelar FliG (Dorman, 2004) y las proteínas tipo Hha (Nieto y col., 2000; Madrid y col., 2002; Paytubi y col., 2004; Ali y col., 2013).

A



B



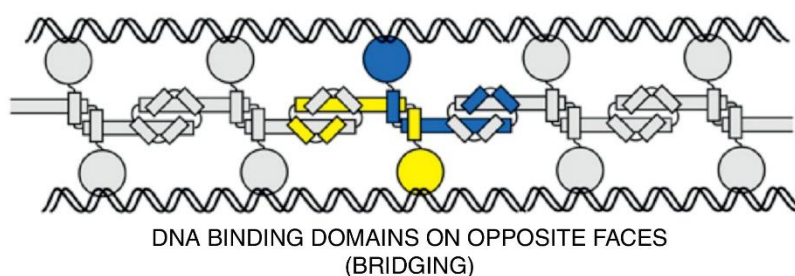
**Figura 1.19.** (A) Representación esquemática de un dímero H-NS/H-NS. Los dos monómeros de H-NS (azul y amarillo) se encuentran unidos a través del mecanismo "head to head /tail to tail". Las hélices H1, H2 y H3 forman la interfaz de dimerización N-terminal. El dominio de dimerización central está formado por un motivo hélice-giro-hélice formado entre el extremo terminal de las hélices H3 y H4. Adaptado de (Ali y col., 2012). (B) Representación estructural del mecanismo de oligomerización de H-NS. En esta imagen se representa un oligómero de tres dímeros relacionados a través del dominio de dimerización N-terminal de cada monómero (*site 1*), conectados a su vez entre sí a través de su interfaz central de oligomerización (*site 2*). El dímero central se representa coloreado, en el que los dos dominios de dimerización de H-NS se encuentran enfrentados entre sí de manera antiparalela. Adaptado de (Arold y col., 2010).

### 1.2.1.3 Unión al ADN por parte de H-NS: efectos topológicos

La proteína H-NS se une al ADN de una manera cooperativa en forma de largos complejos macroproteicos (Badaut y col., 2002). La importancia de estos elevados grados de polimerización radica en el hecho de que mutaciones en el dominio de oligomerización de H-NS inhiben la capacidad de silenciar la expresión génica por parte de esta proteína (aunque retenga la capacidad de unión al ADN) (Ueguchi y col., 1996; Castang y Dove,

2010). Además, también se ha podido comprobar que la presencia de proteínas que alteran la capacidad de oligomerizar por parte de H-NS (como la proteína 5.5 del fago T7 o proteínas H-NS truncadas) también inhiben su capacidad de silenciamiento de la expresión génica (Williamson y Free, 2005; Ali y col., 2011) (una descripción más detallada de estos mecanismos se expone en el apartado 1.2.1.6).

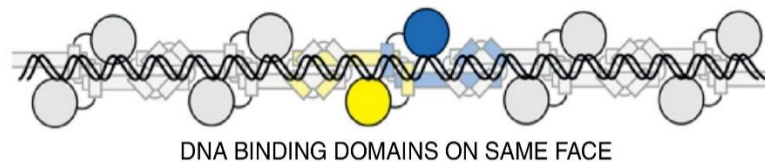
Estudios de microscopía de fuerza atómica (AFM, *atomic force microscopy*) con moléculas individuales de H-NS han demostrado dos posibles modos de unión de H-NS al ADN *in vitro*. El primer modelo describe que la proteína es capaz de formar puentes entre varias cadenas de ADN adyacentes entre sí a nivel inter e intracatenario (Dame y col., 2005). Dicha conformación tiene lugar cuando los dominios de unión al ADN de los monómeros que forman el oligómero de H-NS se orientan de manera antiparalela a lo largo de la hélice de ADN (Arold y col., 2010; Ali y col., 2012) (figura 1.20). Este tipo de unión formando puentes entre el ADN mediante grandes complejos proteicos también se ha identificado en otras proteínas de la familia H-NS, como StpA en *E. coli* y *Salmonella* (Lim y col., 2012), MvaT y MvaU de *Pseudomonas aeruginosa* (Winardhi y col., 2012; Winardhi y col., 2014) y Lsr2 de *Mycobacterium tuberculosis* (Qu y col., 2013).



**Figura 1.20.** Representación esquemática del modelo de unión de H-NS al ADN mediante formación de puentes (*bridging mode*). H-NS puede interactuar en forma de dímero con dos hebras de ADN para formar puentes entre las cadenas nucleotídicas cuando los dominios de unión al ADN de los monómeros de H-NS se disponen de manera antiparalela. Imagen adaptada de (Ali y col., 2012).

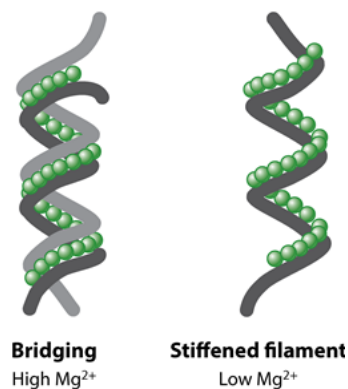
En el segundo modelo descrito, denominado *stiffened filament*, las moléculas de H-NS se disponen de manera extendida sobre el ADN en forma de filamentos dándole una conformación de mayor rigidez (Amit y col., 2003) (figura 1.21).

## 1. INTRODUCCIÓN



**Figura 1.21.** Representación esquemática del modelo de unión de H-NS al ADN mediante el mecanismo *stiffening*. Imagen adaptada de (Ali y col., 2012).

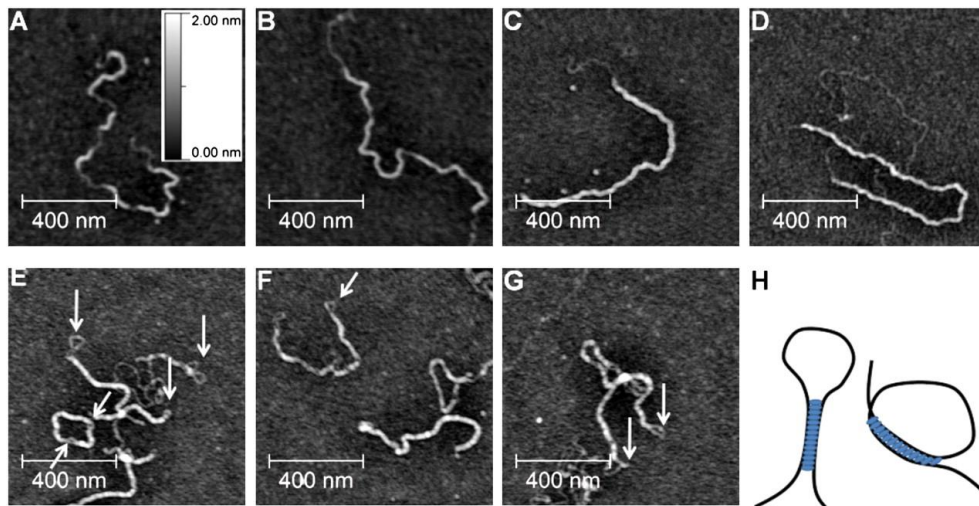
Se ha descrito que la transición de un mecanismo de unión a otro depende de la concentración de iones divalentes (concretamente magnesio) en el medio. A una concentración de 5 mM de  $MgCl_2$  o superior, H-NS es capaz de unirse al ADN en *bridging mode*. Este tipo de unión es muy estable y no es susceptible a cambios en la temperatura o el pH del medio. Sin embargo, en ausencia de cationes divalentes en el medio, o cuando la concentración de los mismos es baja (menos de 1 mM), H-NS se une al ADN y lo recubre en forma de filamentos (*stiffened mode*), dando lugar a una estructura ácido nucleico-proteína rígida (Liu y col., 2010; Singh y col., 2016) (figuras 1.22 y 1.23).



**Figura 1.22.** Representación esquemática de los dos modelos de unión de la proteína H-NS al ADN en función de la concentración de iones magnesio en el medio. Imagen adaptada de (Singh y col., 2016).

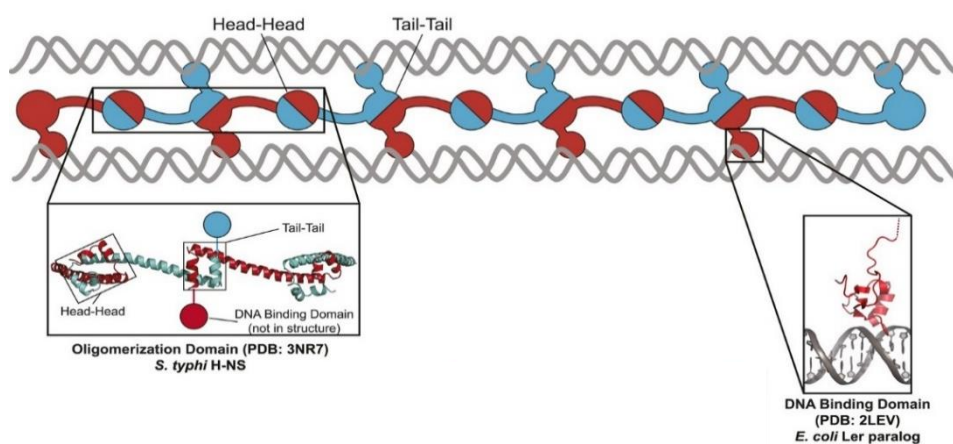
Sin embargo, cabe destacar que a las concentraciones intracelulares de  $Mg^{2+}$  (entre 1-2 mM) coexisten ambos métodos de unión al ADN por parte de H-NS, tanto el modo de unión formando filamentos como el modo de unión formando puentes entre las hélices de ADN (Liu y col., 2010).





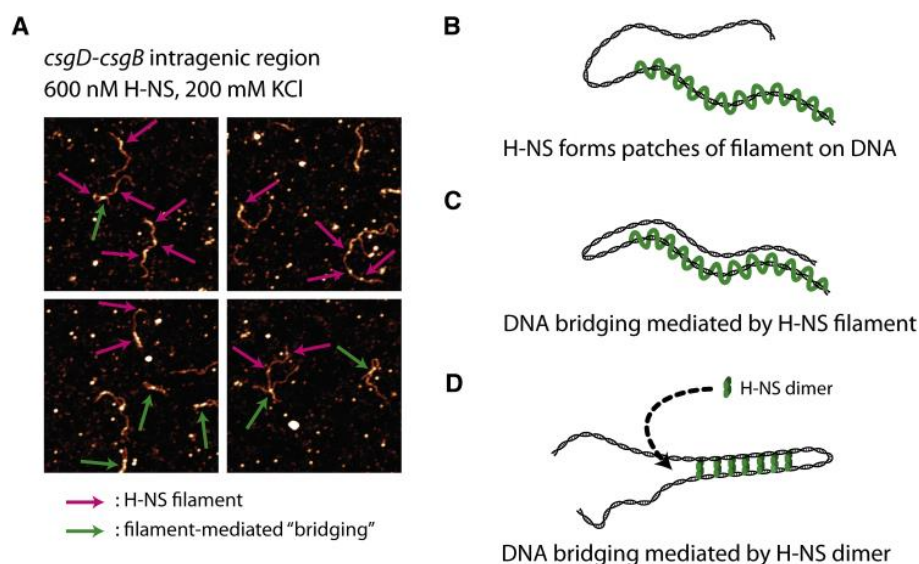
**Figura 1.23.** Imagen de microscopía de fuerza atómica (AFM) de los diferentes complejos de unión formados entre el ADN y H-NS. Representación de las distintas conformaciones que adoptan los complejos ADN/H-NS en presencia o ausencia de iones  $Mg^{2+}$ . (A y B) Interacción durante 40 minutos de la proteína H-NS con el ADN en ausencia de iones  $Mg^{2+}$ . (C y D) Interacción durante 4 horas de la proteína H-NS con el ADN en ausencia de iones  $Mg^{2+}$ . En ambos casos, H-NS se une a lo largo de la cadena de ADN adquiriendo la conformación *stiffening*. (E-F-G) Interacción durante 40 minutos de la proteína H-NS con el ADN bajo una concentración de iones  $Mg^{2+}$  de 10 mM. En presencia de iones  $Mg^{2+}$ , la proteína H-NS se une a la cadena de ADN adquiriendo la conformación *bridging*. Las flechas en blanco indican las estructuras en forma de bucle formadas. (H) Representación esquemática de las diferentes estructuras en forma de bucle observadas en *bridging mode*. Imagen adaptada de (Liu y col., 2010).

Se ha descrito también que la proteína H-NS unida al ADN en forma de filamentos (*stiffened filament*) es capaz de mediar también una unión al ADN en forma de puentes en función de las condiciones del medio en el que tenga lugar la unión proteína-ácido nucleico (Winardhi y col., 2015) (figuras 1.24 y 1.25).



**Figura 1.24.** Representación esquemática del modelo que combina los dos mecanismos descritos para la unión de H-NS al ADN: formación de puentes de ADN (*bridging mode*) a partir de la unión de H-NS en forma de filamentos (*stiffening mode*). En este modelo, H-NS se une en forma de dímero al surco menor del ADN rico en A/T a través del dominio de unión al ADN C-terminal, y puede formar filamentos por los contactos de “cabeza a cabeza” y “cola a cola” a través del dominio de oligomerización N-terminal. Imagen adaptada de (Kotlajich y col., 2015).

## 1. INTRODUCCIÓN

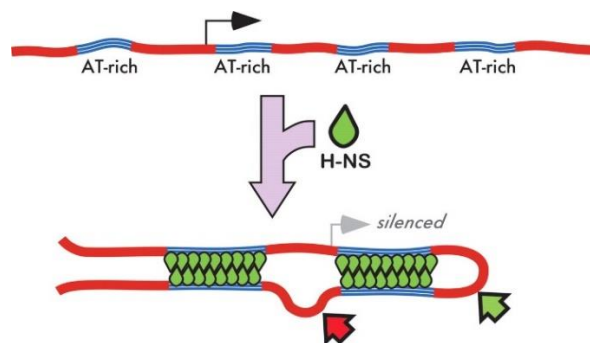


**Figura 1.25.** La unión cooperativa al ADN por parte de H-NS en *stiffening mode* media en la organización de grandes moléculas de ADN en diversas conformaciones dependiendo del nivel de saturación de los enlaces H-NS/ADN y de las diferentes condiciones ambientales. (A) Bajo condiciones de 200 mM de KCl y en ausencia de  $MgCl_2$ , H-NS forma pequeños parches de filamentos (flechas magentas). Estos filamentos pueden asociarse posteriormente con segmentos de ADN desnudos para formar puentes de ADN (flechas verdes). (B-C-D) Representación esquemática del proceso de formación de puentes en el ADN mediado por la proteína H-NS unida al ADN en conformación *stiffening*. La unión de los dímeros de H-NS al ADN, cada uno de los cuales proporciona dos sitios de unión al ADN, permitirán la unión de dos sitios de ADN alejados para formar un puente entre ellos. Imagen adaptada de (Winardhi y col., 2015).

### 1.2.1.4 Regulación de la expresión génica por parte de la proteína H-NS

El inicio del proceso de la transcripción implica la unión de la ARN polimerasa a una secuencia de ADN (promotor) del gen diana. A esto le sigue el desenrollamiento del ADN (formación del complejo abierto) y la creación de la cadena de ARN naciente (iniciación de la transcripción). La unión inicial de la ARN polimerasa está mediada por determinadas secuencias nucleotídicas responsables del contacto directo de la enzima con el ADN. La subunidad  $\sigma^{70}$  de la ARN polimerasa contacta con las secuencias consenso representativas de la caja  $-35$  (5'-TTGACA-3') y la caja  $-10$  (5'-TATAAT-3'). Además, en algunos promotores, el dominio C-terminal de la subunidad  $\alpha$  de la ARN polimerasa ( $\alpha$ CTD) es capaz de unirse a secuencias ricas en A/T situadas aguas arriba del promotor nativo (Lee y col., 2012). En contraposición a estos eventos, al unirse al ADN, H-NS es capaz de alterar la topología del mismo, restringiendo de esta manera las interacciones ARN polimerasa-promotor. Ello se demostró por primera vez en los *loci proU* y *bgl* (genes que están sujetos a represión por parte de H-NS) (Nagarajavel y col., 2007).

Se han propuesto varios mecanismos a través de los cuales H-NS puede reprimir la transcripción. El primer mecanismo reportado de silenciamiento génico mediado por H-NS se basa en la formación de puentes de ADN entre regiones promotoras. Estas estructuras permitirían atrapar e inmovilizar la ARN polimerasa. De esta manera, H-NS podría unirse a secuencias curvadas de ADN ricas en A y T y formar complejos nucleoproteicos entre dos hélices de ADN adyacentes, formando de esta manera un bucle represor que impediría el acceso de la ARN polimerasa (Navarre y col., 2007; Dillon y Dorman, 2010) (figura 1.26).

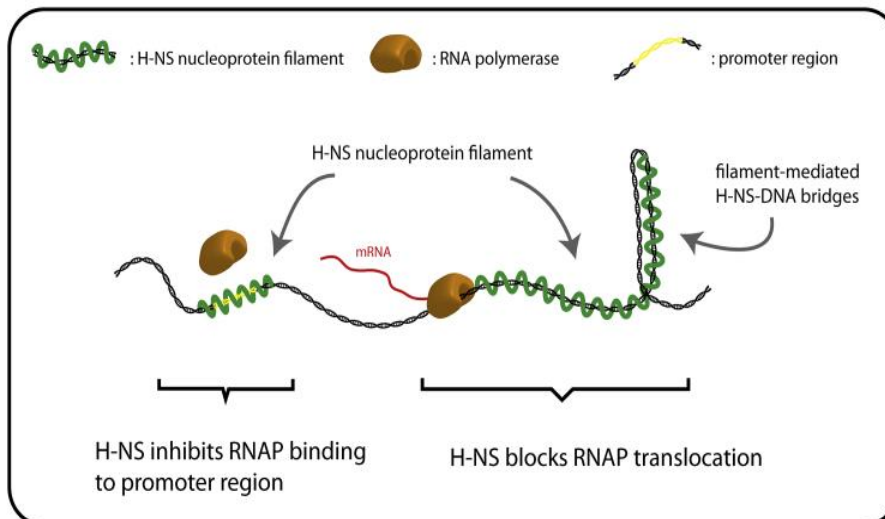


**Figura 1.26.** Silenciamiento génico mediado por H-NS al unirse al ADN en *bridging mode*. H-NS es capaz de polimerizar a lo largo de la cadena nucleotídica y formar puentes entre cadenas de ADN adyacentes, creando bucles que impiden el acceso de la ARN polimerasa al promotor. Imagen adaptada de (Navarre y col., 2007).

Sin embargo, estudios posteriores han demostrado que el mecanismo mediante el cual H-NS es capaz de provocar el silenciamiento génico es mediante la formación de filamentos rígidos sobre el ADN. Ello se pudo comprobar ya que se ha podido observar que mutantes defectivos para la formación de filamentos en el momento de unirse al ADN también son incapaces de silenciar la expresión génica *in vivo* (Lim y col., 2012). De esta manera, H-NS puede formar un filamento continuo sobre una secuencia promotora, creando de esta manera una barrera física que bloquea de manera muy efectiva la accesibilidad a dicho promotor por parte de la ARN polimerasa. Estos filamentos también se pueden formar aguas abajo de la región promotora del gen regulado; de esta manera, el filamento nucleoproteico actuaría como obstáculo físico que impide el proceso de elongación de la ARN polimerasa. Por último, estas estructuras filamentosas también podrían dar lugar a la formación de estructuras represoras en forma de bucle, generándose de esta manera un obstáculo más fuerte para translocación de la ARN polimerasa en comparación con únicamente el filamento en sí (Lim y col., 2012) (figura 1.27).

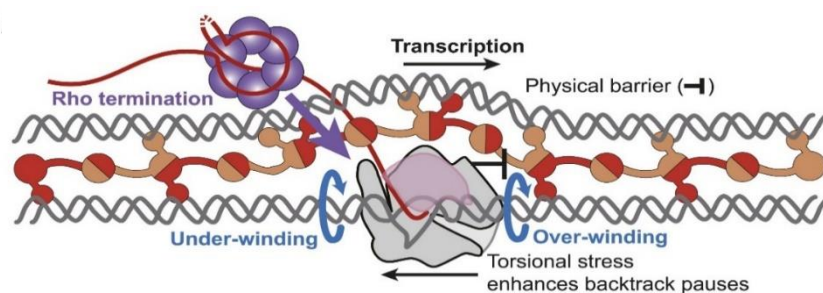


## 1. INTRODUCCIÓN



**Figura 1.27.** Representación esquemática del proceso de silenciamiento génico mediado por nucleofilamentos de H-NS. H-NS es capaz de inhibir la transcripción mediante la inhibición de la unión de la ARN polimerasa a la región promotora o bloqueando la translocación de la enzima a lo largo del ADN. Todo ello es posible mediante la formación de los filamentos nucleoproteicos que H-NS crea en la región promotora o aguas abajo del promotor del gen diana. Además, este filamento puede asociarse con ADN desnudo para formar puentes de ADN, lo que puede restringir aún más la accesibilidad de la ARN polimerasa al ADN. Adaptado de (Winardhi y col., 2015).

Además, estas estructuras filamentosas que forman puentes entre ácidos nucleicos pueden llegar a restringir el superenrollamiento del ADN, creando de esta manera un dominio topológico cerrado en el que se acumula estrés torsional durante la transcripción que promueve la pausa necesaria para que tenga lugar la terminación transcripcional dependiente de Rho (Kotlajich y col., 2015) (figura 1.28).

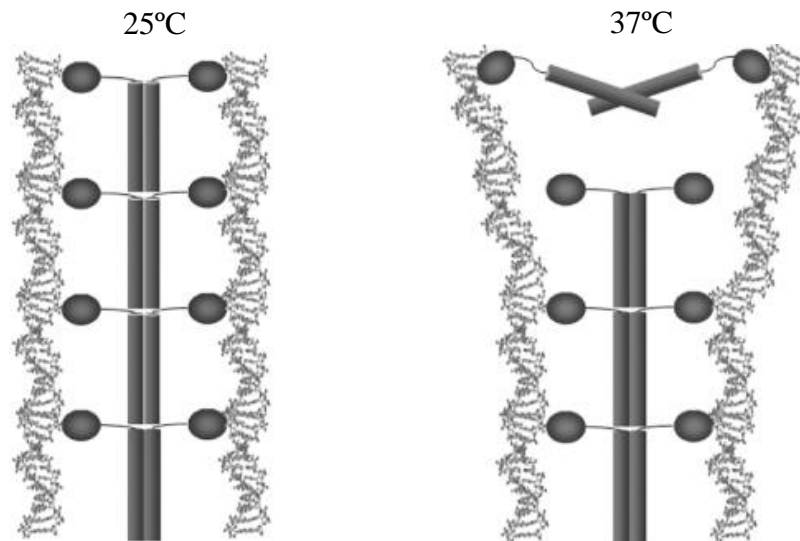


**Figura 1.28.** Las estructuras filamentosas de H-NS/ADN que forman puentes nucleoproteicos entre sí generan una barrera física para el avance del proceso de elongación de la ARN polimerasa. Además, estos puentes crean una tensión conformacional adicional en el macrocomplejo generado que, sumado al impedimento estérico con el que se encuentra la ARN polimerasa, provocan pausas en el proceso de la transcripción que facilitarían la unión de Rho al transcrito naciente y el consecuente proceso de terminación de la transcripción mediado por esta helicasa. Imagen adaptada de (Kotlajich y col., 2015).

Rho es un helicasa homohexamérica de ARN dependiente de ATP que se une a 80 nucleótidos de ARN no estructurado rico en citosina (Peters y col., 2011). Una vez unido al transcrito nascente de ARN, Rho es capaz de translocar de 5'→3' a lo largo del ARN hasta que llega al complejo de elongación, donde tiene la capacidad de terminar la transcripción. La tasa de polimerización del complejo de elongación determina la ventana cinética durante la cual Rho puede actuar sobre la transcripción, de manera que aumentos en la frecuencia y la potencia de los eventos que pausen la transcripción pueden influir en la terminación de la transcripción mediada por Rho (Jin y col., 1992; Kotlajich y col., 2015). La delección de *hns* combinada con mutaciones en *rho* o con la inhibición química de la proteína es letal para la bacteria, lo que sugiere un papel funcional de H-NS en la terminación de la transcripción dependiente de Rho (Tran y col., 2011; Peters y col., 2012).

Al igual que los diferentes mecanismos de unión de H-NS al ADN están influidos por múltiples factores externos, estas estructuras represoras de la expresión génica también se ven afectadas por cambios en los componentes en el medio, de manera que una disminución del pH del medio o de los iones magnesio del mismo favorecerá la formación de filamentos de H-NS y, por tanto, la represión de la expresión génica (Amit y col., 2003; Liu y col., 2010). Igualmente notorio es el efecto de la temperatura, donde a 37°C se observa una gran disminución en la formación de filamentos de H-NS sobre el ADN. La termorregulación dependiente de H-NS se debe al efecto que la temperatura ejerce sobre los diferentes estados de oligomerización de H-NS. Mientras que la estructura secundaria sería básicamente la misma a alta o baja temperatura, la disposición de los dímeros y oligómeros es diferente en función de la temperatura. A elevada temperatura se facilitarían la formación de dímeros, pero no la formación de estructuras de orden superior. De esta manera, se estaría generando una mayor inestabilidad de la unión H-NS-ADN, lo que facilitaría el acceso de la ARN polimerasa, su unión al ADN y la consiguiente expresión génica. En estas condiciones de elevada temperatura, que pueden darse cuando la bacteria infecta células humanas, la formación de estos filamentos se ve muy desfavorecida. Ello tiene mucha importancia a nivel fisiológico, ya que en condiciones de un proceso infeccioso a la bacteria le interesa desreprimir la expresión de genes implicados en la virulencia, muchos de los cuales se encuentran regulados por H-NS. En cambio, a temperaturas ambientales de unos 25°C, se produce la formación de oligómeros de mayor orden, hecho que favorece el silenciamiento génico (Ono y col., 2005) (figura 1.29).

## 1. INTRODUCCIÓN



**Figura 1.29.** Representación esquemática del mecanismo propuesto para el modelo de actuación de H-NS en respuesta a un aumento de la temperatura de 25°C a 37°C. A 25°C, la proteína H-NS es capaz de unirse a las cadenas de ADN de una manera cooperativa basada en una estructura oligomérica de alto orden. En cambio, debido al aumento de la temperatura, a 37°C se produce el cambio conformacional en uno de los dímeros de H-NS. El dímero ya no puede interactuar con el oligómero de proteínas de alto orden y, por lo tanto, ya no se une al ADN de manera cooperativa. La pérdida resultante en la afinidad de unión al ADN afecta la topología del mismo y lo hace más accesible a la ARN polimerasa, lo que da lugar a la transcripción génica con la consecuente pérdida del silenciamiento génico. Adaptado de (Ono y col., 2005).

Además de lo mencionado anteriormente, H-NS también presenta la capacidad de regular la expresión génica a nivel post-transcripcional. Puesto que dicha proteína también muestra una elevada afinidad de unión al ARN, H-NS es capaz de modular positivamente las eficiencias de iniciación de la traducción en moléculas de ARNm que poseen señales de iniciación de traducción deficientes (Park y col., 2010).

### **1.2.1.5 H-NS como represor de la expresión no deseada de ADN xenogénico**

Uno de los papeles fundamentales de la represión de la expresión génica ejercida por parte de la proteína H-NS es controlar la expresión de genes localizados en regiones correspondientes a ADN adquirido de manera horizontal (HGT, *horizontal gene transfer*). Los mecanismos de transferencia génica horizontal son responsables de una parte importante de la gran variabilidad genética que se da entre las diferentes bacterias, incluso entre cepas muy cercanas entre sí. Dichos mecanismos de transferencia génica horizontal se han observado en todos los tipos de bacterias, tanto en bacterias comensales como en bacterias patógenas. Teniendo en cuenta que las bacterias patógenas necesitan adaptarse a condiciones ambientales específicas dentro del hospedador, el papel de la transferencia

génica intra e interespecífica ha jugado un importante papel en la evolución de las bacterias virulentas. Muchos de los genes asociados a la virulencia en la mayoría de los patógenos han sido adquiridos por estos mecanismos. Ejemplos de ello son algunos elementos clave para la colonización de sus hospedadores (Park y Zhang, 2012; Harrison y Brockhurst, 2012), y genes de resistencia a elementos antimicrobianos (Bennet, 2008) o a metales pesados (Pal y col., 2015). Sin embargo, la expresión descontrolada de estos genes adquiridos *de novo* puede provocar la alteración de la homeostasis celular al generar una alteración en los circuitos regulatorios existentes previamente. H-NS actúa como un factor clave que ha permitido la incorporación progresiva de estos genes en las redes regulatorias de la célula sin comprometer el *fitness* celular.

Tal y como se ha mencionado anteriormente, las secuencias xenogénicas adquiridas mediante mecanismos HGT muestran una composición de bases y un uso de codones diferentes al del núcleo del genoma de la célula huésped receptora: en la mayoría de los casos, los genes adquiridos por HGT son relativamente ricos en adenina y timina (Lawrence y Ochman, 1997; Daubin y col., 2003). La capacidad que presenta H-NS para unirse selectivamente y silenciar secuencias ricas en A/T (proceso conocido como silenciamiento xenogénico (Navarre y col., 2007)), convierte a esta proteína en uno de los principales reguladores de la virulencia en *E. coli* y *Salmonella*, así como en otras especies de los géneros *Yersinia* y *Vibrio* (Dillon y Dorman, 2010).

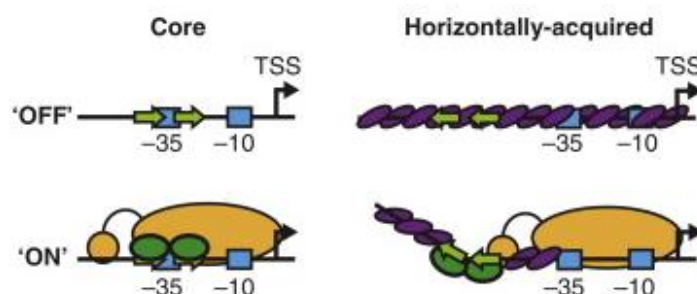
En 2001 se realizó un primer análisis para intentar definir el regulón de H-NS en *E. coli*. En dicho estudio se demostró que la mutación del gen *hns* afectaba aproximadamente al 5% del transcriptoma de la cepa K-12 de *E. coli* (Hommais y col., 2001). En este análisis se observó que varios genes anotados como ADN cuyo origen era la transferencia génica horizontal se encontraban desregulados (principalmente genes implicados en la adhesión celular). El problema radicaba en el hecho de que la cepa utilizada fue una cepa “de laboratorio”, por lo que no incluía en su genoma muchos genes de virulencia, principalmente islas de patogenicidad. En 2006 se realizó un estudio con un *array* dirigido de manera específica contra genes implicados en la virulencia de una cepa patógena de *E. coli*, concretamente la cepa uropatógena 536. En este caso, se pudo observar como todos los *loci* de virulencia descritos hasta la fecha se vieron afectados por la mutación en el gen *hns* (Muller y col., 2006). De manera análoga, en sendos estudios transcriptómicos realizados mediante *microarrays* sobre el mutante *hns* en *S. Typhimurium* (Lucchini y col., 2006; Navarre y col., 2006), se observó que más de 400

## 1. INTRODUCCIÓN

genes del genoma de *Salmonella* se encontraban desregulados ante la ausencia de la proteína H-NS. La gran mayoría de estos genes eran ricos en A/T (el contenido de G/C de los genes silenciados por H-NS era del 48%, mientras que el del genoma global es del 52%), el 90% de los cuales eran genes adquiridos por transferencia horizontal. En estudios de inmunoprecipitación de la cromatina combinados con estos *microarrays* (“*ChIP-on-chip*”), se pudo determinar que la proteína H-NS se unía a más de 400 regiones del genoma de *Salmonella* (incluyendo las 5 islas de patogenicidad y su plásmido de virulencia) donde, además de unirse a regiones promotoras, también se podía unir a regiones codificantes ricas en A/T dentro de los propios genes (Navarre y col., 2006).

### **1.2.1.6 Mecanismos de contra silenciamiento de la expresión génica**

Sin embargo, todo este silenciamiento génico mediado por la proteína H-NS tiene que estar finamente regulado ya que, en algún momento, los genes cuya expresión génica es reprimida por parte de H-NS, necesitarán expresarse. Esta “desrepresión” es especialmente importante cuando la expresión de dichos genes sea necesaria para que las bacterias patógenas puedan poner en marcha su maquinaria de adaptación al hospedador (esencialmente para que las bacterias sobrevivan y prosperen) y para sintetizar los factores de virulencia que iniciarán el proceso infeccioso. Por ello, es evidente que la represión transcripcional ejercida por H-NS no se debe de prolongar de manera indefinida. Debido a la acción de diferentes factores de transcripción como por ejemplo Fis e IHF, o bajo determinadas condiciones ambientales como la temperatura o la osmolaridad, se generan mecanismos que dan lugar a la desestabilización de los complejos nucleoproteícos formados por H-NS y la consecuente desrepresión de los genes silenciados (Navarre y col., 2007) (figura 1.30). Todos estos procesos reciben el nombre de contra silenciamiento.



**Figura 1.30.** Mecanismos de activación de la expresión de genes del núcleo del genoma (*core genome*) y contra silenciamiento de genes adquiridos por transferencia horizontal. Los genes del núcleo del genoma se encuentran en un estado de no expresión (estado “OFF”) hasta que un activador transcripcional (representado mediante óvalos verdes) se une a posiciones conservadas en los promotores de dichos genes (representados con flechas verdes haciendo referencia al lugar de inicio de la transcripción (TSS)) e interactúa directamente con la ARN polimerasa (amarillo) para permitir la transcripción y la consecuente expresión del gen (estado “ON”). Por el contrario, los genes HGT se encuentran silenciados por la proteína H-NS (representada en círculos lilas), la cual se une y polimeriza a lo largo de la secuencia promotora impidiendo de esta manera la unión y actuación de la ARN polimerasa. Sin embargo, cuando la expresión de estos genes es necesaria, un factor de contra silenciamiento de la transcripción (representado en círculos verdes) se puede unir a un sitio alternativo al promotor para desestructurar el complejo H-NS-ADN, permitiendo de esta manera la unión de la ARN polimerasa y la consecuente transcripción y expresión génica. Imagen adaptada de (Will y col., 2015).

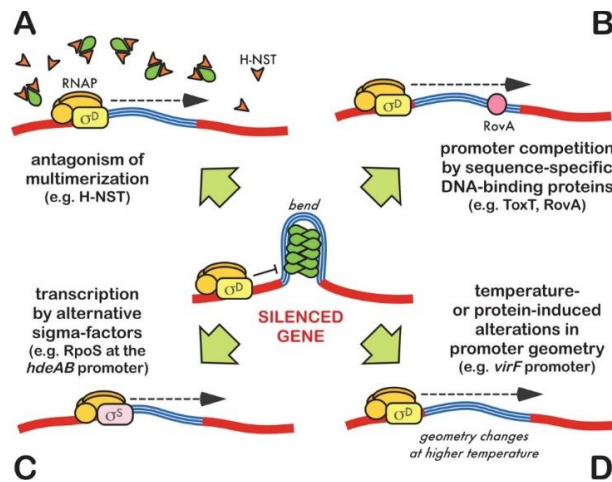
Se han propuesto varios mecanismos de contra silenciamiento de la represión ejercida por parte de la proteína H-NS (Navarre y col., 2007). Dichos mecanismos quedan resumidos en la figura 1.31.

Uno de los principales mecanismos de contra silenciamiento del papel represor de H-NS es el ejercido por algunas proteínas con capacidad de unión al ADN que son capaces de antagonizar las funciones de silenciamiento realizadas por H-NS. El mecanismo propuesto para la actuación de estas proteínas incluye principalmente la competición con H-NS por los sitios de unión al ADN (como en el caso de las proteínas Ler y SsrB) (Walthers y col., 2011; Whinardhi y col., 2014) y la formación de complejos heteroméricos que inhiben la función represora de H-NS. Este último mecanismo es el ejercido principalmente por proteínas con capacidad de interactuar con H-NS interrumpiendo su estado oligomérico, como es el caso de la proteína 5.5 del fago T7 (Liu y Richardson, 1993; Ali y col., 2011) o las proteínas de la familia H-NST, proteínas tipo H-NS (presentes principalmente en pequeñas islas genómicas de los patotipos UPEC, EAEC y ETEC) pero truncadas de manera natural a las que les falta el dominio de unión al ADN y la región flexible (Williamson y Free., 2005). Otro mecanismo de contra silenciamiento es el ejercido por factores de transcripción que se unen al ADN con una



## 1. INTRODUCCIÓN

alta afinidad y compiten contra H-NS por la unión a los ácidos nucleicos, como es el caso de ToxT en *Vibrio cholerae* o SlyA/RovA en *Yersinia* y *E. coli* (Yu y DiRita, 2002; Tran y col., 2005).



**Figura 1.31.** Mecanismos de contra silenciamiento génico descritos para una gran variedad de promotores regulados por H-NS. (A) La interrupción de los complejos represores generados por H-NS por parte de antagonistas de la multimerización tales como H-NST, Ler, SsrB o la proteína 5.5 del fago T7. (B) La competencia por la unión al ADN por parte de proteínas que presentan una capacidad específica de unión al ADN de alta afinidad, tales como ToxT o RovA. (C) La activación de ciertos promotores por factores sigma alterativos, como el promotor *spv* de *Salmonella* o el promotor *hdeAB* de *E. coli*, activados por el factor  $\sigma$  RpoS. (D) Cambios en la arquitectura del promotor debido a la unión de proteínas específicas o cambios ambientales que interrumpen la formación del complejo H-NS, como ocurre en el promotor *virF* en *Shigella*. Imagen adaptada de (Navarre y col., 2007).

Además, bajo ciertas condiciones, las secuencias promotoras de ciertos genes se encuentran reguladas positivamente por el factor sigma alternativo (RpoS) de la ARN polimerasa. Estos factores, asociados con la ARN polimerasa E, tienen la capacidad de poder activar la expresión de promotores de genes adquiridos por HGT aún siendo sometidos al proceso de silenciamiento ejercido por H-NS. Ejemplos de estos *loci* son *spv*, *LEE1* (Laaberki y col., 2006), *asr* (Seputiene y col., 2004), *gadBC* (De Biase y col., 1999), *proU* (Rajkumari y Gowrishankar, 2001), *hdeAB* (Shin y col., 2005), *csgBA* (Olsen y col., 1993), *spvR* (Robbe-Saule y col., 1997), *csiD* (Marschall y col., 1998), y *yciEFG* (Navarre y col., 2006).

Este antagonismo contra el silenciamiento xenogénico provocado por H-NS también tiene lugar bajo ciertas condiciones en las que se producen alteraciones en la topología y arquitectura del promotor. Estas alteraciones, causadas por proteínas de unión al ADN o por cambios ambientales (como por ejemplo, un incremento en la temperatura), provocan

que se interrumpa la formación de los complejos H-NS-ADN. Un claro ejemplo de ello se da en el caso del promotor *virF* de *Shigella* (Falconi y col., 1998).

Cabe destacar que todos estos mecanismos de contra silenciamiento descritos no son excluyentes entre ellos y pueden coexistir a la hora de regular la expresión de un mismo promotor.

### **1.2.1.7 Efecto de la proteína H-NS sobre la transcripción espuria**

Como se comentaba anteriormente, la ARN polimerasa que se expresa de manera constitutiva en bacterias asociada al factor  $\sigma^{70}$ , únicamente necesita reconocer una secuencia parcial que coincide con la secuencia consenso (5'-TATAAT-3) y coincidencias parciales con secuencias accesorias (todas ellas ricas en A y T) para iniciar la transcripción. Por todo ello, y por una cuestión probabilística, en *E. coli*, los genes con un contenido en A y T mayor del 60% contienen secuencias nucleotídicas capaces de dirigir y llevar a cabo la transcripción tanto *in vitro* como *in vivo* (Decker y Hinton, 2013). A lo largo de la historia evolutiva, la evolución natural de los microorganismos (y de manera específica en el caso de *E. coli*) ha permitido que la presencia de estas secuencias “tipo promotor” que puedan dar lugar a una transcripción aberrante sean mínimas en el interior de los genes, donde el porcentaje de contenido A/T es un 5-9% menor que en regiones intergénicas (Blattner y col., 1997).

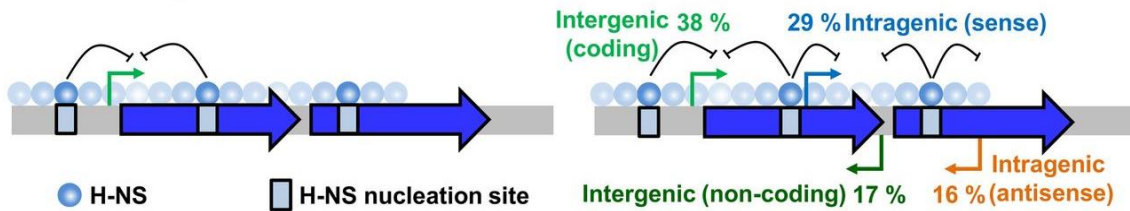
Sin embargo, los mecanismos para evitar la transcripción espuria (transcripción aberrante a partir de promotores intragénicos no específicos) no son eficientes al 100%, por lo que a lo largo de la historia evolutiva se han seleccionado de manera positiva mecanismos implicados en evitar la formación de estos transcritos aberrantes. El principal mecanismo protector ante estos transcritos es el acople entre los procesos de transcripción y traducción. Como la mayoría de los codones de inicio de la traducción y los sitios de unión al ribosoma se encuentran localizados de manera específica y rara vez ocurren por casualidad, la mayoría de los transcritos espurios no se traducen al no contener estas secuencias en el orden necesario para su traducción.

A pesar de ello, existen mecanismos adicionales para suprimir la transcripción génica en promotores intragénicos que darían lugar a transcritos aberrantes y su traducción indeseada. En *E. coli*, uno de estos mecanismos supresores es el complejo nucleoproteico represor que ejerce la proteína H-NS cuando se une al ADN y polimeriza en estructuras proteicas de un orden mayor que son capaces de provocar la represión de la transcripción



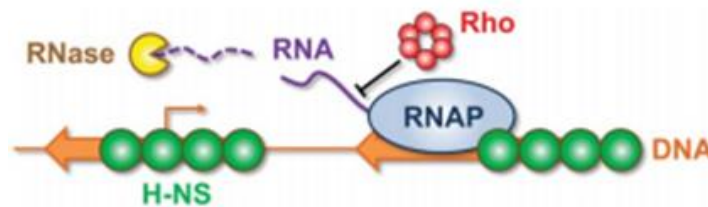
## 1. INTRODUCCIÓN

espuria (Grainger, 2016) (figura 1.32). Por todo ello, se ha descrito que mutaciones en el gen *hns* provocan una transcripción descontrolada de moléculas de ARN de dentro de los genes con un elevado contenido de A y T (Lamberte y col., 2017).



**Figura 1.32.** La proteína H-NS es capaz de silenciar la expresión de un gran número de genes al reprimir la transcripción canónica. Además, otra función principal de esta proteína es silenciar la transcripción espuria de ARN no codificantes. Al suprimir la transcripción intragénica, con ello se evita el gasto innecesario de recursos celulares para dar lugar a transcritos no deseados, la formación de estructuras de ácido nucleico peligrosas y una eventual transferencia con los procesos de transcripción realmente funcionales. Imagen adaptada de (Singh y col., 2014).

Junto con H-NS, el factor de terminación de la transcripción Rho, presente en el 90% de las bacterias, también es capaz de reconocer y finalizar la transcripción de ARN no codificante, proporcionando a la bacteria otro mecanismo de protección ante la transcripción no deseada (D'Heygère y col., 2013) (figura 1.33).

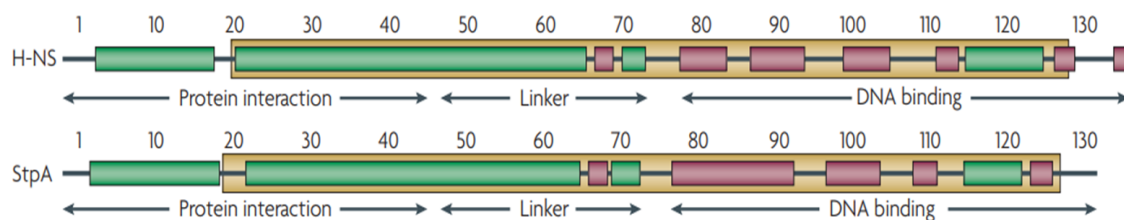


**Figura 1.33.** La proteína H-NS puede suprimir la actividad de promotores espurios dentro de los genes y puede impedir la elongación de la transcripción. Aquellos transcritos que hayan superado la represión de la transcripción inicial suelen estar sujetos a una terminación prematura de la misma por parte de Rho. Finalmente, los transcritos resultantes pueden ser objeto de degradación por parte de las ARNasas celulares. Imagen adaptada de (Wade y Grainger, 2017).

### 1.2.1.8 La proteína StpA y otros parálogos de H-NS

Los genomas de diferentes bacterias entéricas (*Salmonella*, *Escherichia* y *Shigella*), codifican un segundo parálogo de la proteína H-NS, la proteína StpA (*Supressor of td-mutant phenotype A*). El gen *stpA* codifica para una proteína de 134 aminoácidos y 15,3 kDa, que presenta unos niveles de similitud del 67% y una identidad del 58% a nivel de secuencia aminoacídica respecto H-NS, además de presentar unos dominios estructurales muy similares: un dominio N-terminal involucrado en la dimerización e interacción con otras proteínas y un dominio C-terminal implicado en la unión al ADN. Dichos dominios

presentan una estructura secundaria muy parecida a los de la proteína H-NS (Zhang y Belfort, 1992; Williams y col., 1996; Cusick y Belfort, 1998; Dorman, 2004) (figura 1.34).



**Figura 1.34.** Representación esquemática de la comparación de los dominios estructurales y las estructuras secundarias predichas para las proteínas H-NS y StpA de *Salmonella Typhimurium*. De color verde se representan las hélices- $\alpha$  y de color rosa oscuro los dominios en forma de bucle desestructurados. En dorado se indica el dominio tipo histona H-NS extraído de la base de datos Pfam. Dicho dominio proteico es común para H-NS y StpA. Imagen adaptada de (Dillon y Dorman, 2010).

De manera análoga a H-NS, la proteína StpA tiene capacidad de unirse al ADN (Sonden y Uhlin, 1996; Uyar y col., 2009) en regiones ricas en A/T con curvatura intrínseca y formar puentes entre cadenas adyacentes (Dame y col., 2000; Dame y col., 2005). Además, también presenta actividad de chaperona de ARN (Zhang y col., 1995).

En *E. coli*, el patrón de expresión de StpA y H-NS es diferente. Mientras que los niveles de H-NS son más o menos constantes a lo largo de las diferentes etapas de la curva de crecimiento bacteriano, la expresión de StpA depende de la fase de crecimiento. Los niveles de expresión de StpA se mantienen bajos al inicio de la fase exponencial, pero aumentan hasta un 50% más en la fase estacionaria. Además, la expresión de este parálogo de H-NS se ve inducida cuando las bacterias crecen bajo condiciones de estrés o cuando crecen en medio mínimo bajo una limitación en el número de nutrientes del medio (Free y Dorman, 1997). La diferencia entre los niveles intracelulares de ambas proteínas puede deberse a que H-NS y StpA se autoregulan y a la vez muestran una regulación cruzada negativa, en la que H-NS es capaz de reprimir la transcripción del gen *stpA* de una manera más fuerte e intensa que la represión que StpA puede ejercer sobre el gen *hns* (Johansson y Uhlin, 1999).

StpA presenta la capacidad de formar heterodímeros tanto con H-NS (Williams y col., 1996; Free y col., 1998) como con proteínas tipo Hha, como por ejemplo con su parálogo YdgT (Paytubi y col., 2004). Es una proteína relativamente estable en un fondo genético salvaje, pero se degrada rápidamente por acción de la proteasa Lon en ausencia de H-NS,

## 1. INTRODUCCIÓN

un efecto que se exagera aún más en ausencia de YdgT o Hha (Johansson y Uhlin, 1999; Paytubi y col., 2004). La formación de estos heterocomplejos proteicos con H-NS y proteínas tipo Hha proporciona a StpA un mecanismo de protección contra la degradación por parte de la proteasa Lon.

Mediante un análisis de *ChIP-chip* a lo largo del cromosoma de *E. coli* se pudo determinar que los perfiles de unión de H-NS y StpA eran prácticamente superponibles. La eliminación de StpA no dio lugar a ningún cambio significativo en el perfil de unión de H-NS; por el contrario, la supresión de *hns* redujo de manera significativa el número de regiones a las que se podía unir StpA (aproximadamente dos tercios de las regiones de unión de StpA requerían de la presencia de H-NS para que se produjera la unión a los ácidos nucleicos). Una explicación de ello es la necesidad que presenta StpA de formar heterodímeros con H-NS para unirse a regiones específicas del ADN. Otra posibilidad es que los cambios locales en la arquitectura del nucleoide provocados por la ausencia de H-NS causaron una disminución en la afinidad de StpA para unirse a ciertas regiones del cromosoma (Uyar y col., 2009).

Muy pocos fenotipos han podido ser atribuidos a la mutación del gen *stpA* en cepas de *E. coli* que contienen una copia salvaje del gen *hns* (Zhang y col., 1996; Bertin y col., 2001). Ello puede ser debido a los mecanismos de compensación ejercidos por H-NS, además de la baja expresión de StpA en las células de fondo genético salvaje. Sin embargo, se ha podido observar que la expresión de StpA se incrementa aproximadamente 20 veces en un mutante de H-NS. A pesar de ello, debido a su inestabilidad en ausencia de H-NS, la proteína StpA no se acumula en mutantes *hns* más allá del 10% de los niveles observados para H-NS (Sonnenfield y col., 2001). Sin embargo, StpA puede complementar parcialmente la inactivación del gen *hns* y actuar como reservorio molecular de H-NS (Shi y Bennett, 1994; Williams y col., 1996; Bertin y col., 2001).

A pesar de converger en la mayoría de los genes regulados, existen algunos genes regulados de manera diferencial por StpA y H-NS en *E. coli*. Un ejemplo de ello es la regulación del gen *ompF* a través del ARN antisentido *micF* complementario al ARN mensajero de *ompF*. Mientras que H-NS incrementa de manera indirecta los niveles de OmpF mediante la inactivación de la transcripción de *micF*, StpA permite la expresión de *ompF* en *E. coli* mediante la unión y desestabilización de este ARN antisentido (Deighan y col., 2000).

De manera análoga a H-NS, la proteína StpA también presenta la capacidad de formar un filamento rígido (*stiffening mode*) a lo largo del ADN. De este modo, una vez que la proteína StpA forma un filamento proteico sobre los ácidos nucleicos, provoca la generación de conformaciones locales totalmente inaccesibles a cualquier tipo de enzima, generado de esta manera un mecanismo de silenciamiento de la expresión génica al impedir el acceso de la ARN polimerasa a su estructura promotora.

Sin embargo, a diferencia del filamento H-NS-ADN, el filamento StpA-ADN es en gran medida insensible a los cambios fisiológicos en las concentraciones de sal, de iones magnesio o de cambios en el pH y en la temperatura del medio. Todo ello parece indicar que la proteína StpA es más resistente a cambios en los estímulos ambientales que la proteína H-NS. Dado que filamento formado por StpA sobre el ADN es más estable que el formado por H-NS, y esta estabilidad se mantiene en condiciones de estrés osmótico o de elevada temperatura, StpA parece estar implicada en mecanismos de resistencia a cambios en la osmolaridad del medio y a cambios bruscos de temperatura cuando la modulación por parte de la proteína H-NS deja de ser funcional (Lim y col., 2012).

En algunas cepas patógenas extraintestinales de *E. coli*, además de los genes *hns* y *stpA*, se ha logrado identificar un tercer gen parálogo a dichos genes. Este gen, denominado *hfp*, se ha caracterizado inicialmente en la cepa uropatógena de *E. coli* 536. La proteína Hfp es capaz de formar heterocomplejos con la proteína H-NS, y actúa como reservorio molecular de H-NS cuando las bacterias crecen a bajas temperaturas (Müller y col., 2010).

### **1.2.2 Familia de proteínas tipo Hha**

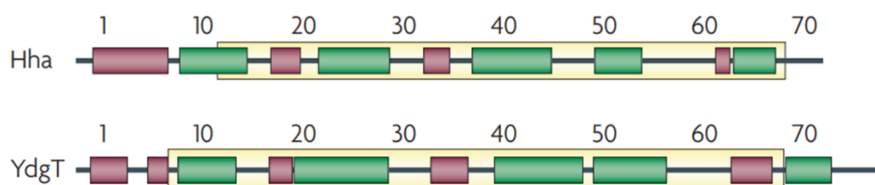
La familia de proteínas tipo Hha está formada por un grupo de proteínas de bajo peso molecular (aproximadamente 8 kDa) que juegan un papel importante en la regulación de la expresión génica en enterobacterias (principalmente se encuentran asociadas a la regulación de elementos de virulencia bacterianos). La proteína prototipo de esta familia, Hha (*High Hemolytic Activity*), fue identificada originalmente como un regulador negativo del gen plasmídico que codifica para la producción de la toxina alfa-hemolisina (*hlyA*), presente en cepas patógenas de *E. coli* (Godessart y col., 1988; Nieto y col., 1991). El gen *hha* presenta unos niveles de transcripción y expresión más elevados cuando las células se encuentran en fase exponencial; en cambio, esta elevada expresión disminuye cuando las células llegan a la fase estacionaria. Además, factores ambientales tales como

## 1. INTRODUCCIÓN

la osmolaridad del medio y la temperatura afectan también a los niveles intracelulares de Hha (Mouriño y col., 1998).

Paralelamente al descubrimiento y caracterización de la proteína Hha, también se descubrió otra proteína miembro de esta familia: la proteína ortóloga a Hha en *Yersinia enterocolitica*, denominada YmoA. Ambas proteínas comparten una identidad aminoacídica del 82%. Esta proteína se caracterizó inicialmente como un regulador negativo del factor transcripcional VirF, y posteriormente como un importante regulador de un gran número de factores de virulencia en *Yersinia*, entre los que destaca la invasina (Cornelis y col., 1991; de la Cruz y col., 1992; Ellison y col., 2003). La gran mayoría de aislamientos de cepas de *E. coli*, *Shigella* y *Salmonella* expresan al menos un parálogo adicional de la proteína Hha: la proteína YdgT (a veces denominada Cnu). A nivel de secuencia aminoacídica, YdgT presenta una similitud del 66% y una identidad del 35% respecto a la proteína Hha. Además, algunas cepas también presentan parálogos adicionales codificados en elementos genéticos móviles tales como plásmidos conjugativos (Forns y col., 2005). A diferencia de H-NS, donde genes parálogos y ortólogos se encuentran en un gran número de proteobacterias gramnegativas más allá del linaje entérico, la presencia de proteínas tipo Hha se encuentra únicamente restringida a las *Enterobacteriaceae* (Madrid y col., 2007).

A nivel estructural, Hha es una proteína enriquecida en hélices- $\alpha$  (figura 1.35), ordenadas entre sí de manera helicoidal y estabilizadas por enlaces hidrofóbicos. La distribución de las cargas en la proteína Hha no es homogénea: mientras que las primeras hélices del extremo N-terminal de la molécula son ricas en residuos básicos que conforman una superficie cargada positivamente, la cara opuesta de la molécula es mayormente acídica (Paytubi y col., 2011).



**Figura 1.35.** Representación esquemática comparativa de los dominios estructurales y las estructuras secundarias predichas para las proteínas Hha e YdgT de *Salmonella Typhimurium*. En color verde se representan las hélices- $\alpha$  y en rosa oscuro los dominios en forma de bucle desestructurados. En dorado se indica el dominio tipo Hha extraído de la base de datos Pfam. Dicho dominio proteico es común tanto para Hha como para su parálogo YdgT. Imagen adaptada de (Dillon y Dorman, 2010).

Los miembros de la familia de proteínas similares a Hha comparten una redundancia funcional significativa (Nieto y col., 2002). Los parálogos de Hha, como YmoA e YdgT, pueden complementar a los mutantes *hha* de *E. coli* en la regulación de varios *loci*; recíprocamente, Hha puede reemplazar funcionalmente a YmoA en mutantes de *Yersinia*, aunque el grado de complementación observado depende del número de copias de la proteína (Mikulskis y Cornelis, 1994; Balsalobre y col., 1996; Nieto y col., 2002; Paytubi y col., 2004; Forns y col., 2005). Debido a su redundancia funcional, las mutaciones individuales en *hha* o *ydgT* en *E. coli* o *Salmonella* son bien toleradas por parte de las bacterias y no dan lugar a grandes cambios en el *fitness* celular. En los *loci* que muestran alguna alteración regulatoria, el efecto es casi siempre más notorio para la mutación *hha* que para la mutación *ydgT*. Concretamente, en el caso de *Salmonella*, mutantes *hha* muestran defectos leves de crecimiento y aumento en la expresión de genes codificados dentro de SPI-1 (responsables de la invasión bacteriana de células epiteliales) al verse afectada la expresión de su regulador principal HilA (Fahlen y col., 2001), así como de genes codificados dentro de SPI-2 requeridos para la supervivencia intracelular (Olekhovich y Kadner, 2007; Silphaduang y col., 2007). También cabe destacar que Hha desempeña un papel importante en la regulación de la conjugación y la transposición. Mutaciones en el gen *hha* provocan un aumento en la frecuencia de conjugación plasmídica, mientras que la sobreexpresión de Hha conduce a una disminución en la frecuencia de conjugación, un efecto que se ha observado en más de una familia de plásmidos (Mikulskis y Cornelis, 1994; Nieto y col., 1998; Forns y col., 2005). La sobreexpresión de Hha también provoca un aumento de la transposición por un mecanismo hasta ahora no comprendido, similar a lo que se ha observado en el caso de H-NS en la transposición de los transposones *IS903*, *Tn10* y *Tn552* en *E. coli* (Mikulskis y Cornelis, 1994; Balsalobre y col., 1996; Shiga y col., 2001; Swingle y col., 2004).

Estas alteraciones reguladoras observadas en los mutantes sencillos se potencian en un doble mutante *hha ydgT* a un grado considerablemente mayor que las contribuciones aditivas de ambos mutantes individuales (Paytubi y col., 2004; Silphaduang y col., 2007). El análisis transcriptómico realizado para analizar el efecto de las mutaciones combinadas de los genes *hha* e *ydgT* en *S. Typhimurium* mostró la expresión alterada de un gran número de genes que se encuentran localizados en regiones de ADN adquiridas a través de mecanismos de transferencia horizontal. Además, muchos de estos genes coinciden con genes cuya expresión se encontraba alterada en un mutante *hns*, entre los que destacan

## 1. INTRODUCCIÓN

genes que forman parte de las islas de patogenicidad SPI-1, 2, 3 y 5, así como un gran número de otros genes presentes en pequeñas islas genómicas (Navarre y col., 2006; Vivero y col., 2008).

Todas estas observaciones indican que las moléculas similares a Hha son altamente redundantes con respecto a su función y que, en fenotipos salvajes de *E. coli* y *Salmonella*, Hha es el efector funcional dominante, mientras que YdgT aparentemente desempeña un papel secundario o de respaldo en ausencia de Hha, similar al descrito para StpA y H-NS en *E. coli*.

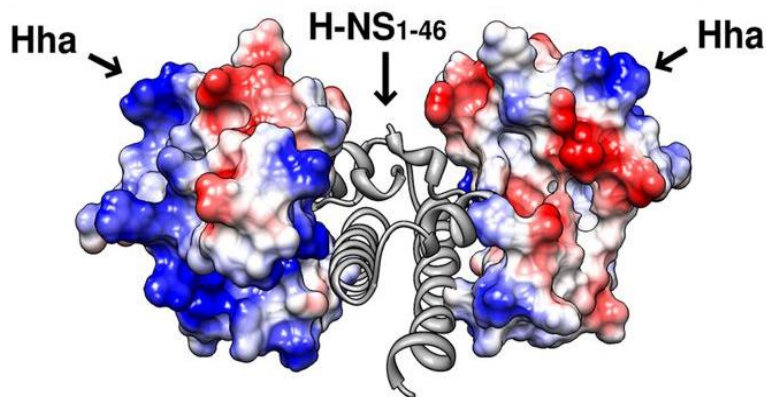
Para estudiar el mecanismo específico de actuación de la proteína Hha se utilizó como modelo la regulación ambiental del operón hemolítico *hlyCABD* de *E. coli*, el cual codifica para la síntesis y secreción de la toxina alfa-hemolisina (Nieto y col., 1991; Carmona y col., 1993). Estos estudios determinaron que Hha no era capaz de unirse a secuencias específicas de la región reguladora del operón *hlyCABD*, pero sí de interaccionar con la proteína H-NS, la cual sí se une de manera específica a secuencias presentes en la región reguladora del operón (Madrid y col., 2002). En un doble mutante *hha hns* se produce una gran expresión de la hemolisina codificada en este operón; ello nos indica que el complejo nucleoproteico formado por H-NS/Hha/ADN presenta la capacidad de modular y reprimir la expresión del operón *hly*, de manera que en ausencia de ambas proteínas se produce una gran desrepresión de la represión ejercida por el complejo macroproteico (Nieto y col., 2000).

Tanto la demostración *in vitro* de la interacción de Hha con H-NS, como la superposición de los genes controlados por Hha/YdgT y por H-NS, sugiere la existencia de interacciones directas entre las proteínas de estas dos familias para regular la expresión génica.

Estudios transcriptómicos globales realizados por *microarray* también han puesto de manifiesto que algunos de los genes silenciados directamente por H-NS, entre los que destaca *proU*, no se ven afectados por mutaciones en *hha* e *ydgT* (Vivero y col., 2008). Ello es de esperar ya que los niveles de Hha/YdgT son mucho más bajos que los niveles de H-NS, y por esto es imposible que estas proteínas puedan actuar en un complejo proteico junto con H-NS en una relación molar de 1:1 (o 1:2) en todos los promotores; más bien sugiere que Hha y YdgT pueden actuar selectivamente con H-NS sólo en ciertos promotores, mientras que en otros la interacción no es posible.



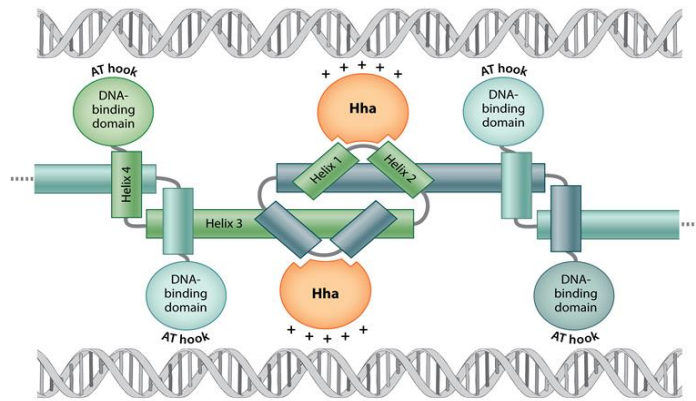
A pesar de que la secuencia aminoacídica completa de los miembros de la familia de proteínas tipo Hha y H-NS no es significativamente similar, las proteínas tipo Hha y el dominio N-terminal de H-NS presentan pequeños bloques con una alta identidad a través de los cuales tiene lugar la interacción proteica (Nieto y col., 2002) (figura 1.36).



**Figura 1.36.** Representación esquemática del complejo de unión de la proteína Hha con el dominio N-terminal de H-NS. Dos monómeros de Hha se unen en dos sitios simétricos a cada lado del dímero de H-NS. Los residuos cargados positivamente, altamente conservados y representados en color azul, quedan expuestos en la superficie cuando se forma el complejo con H-NS, hecho que favorece la interacción del complejo proteico con el ADN. Imagen adaptada de (Ali y col., 2013).

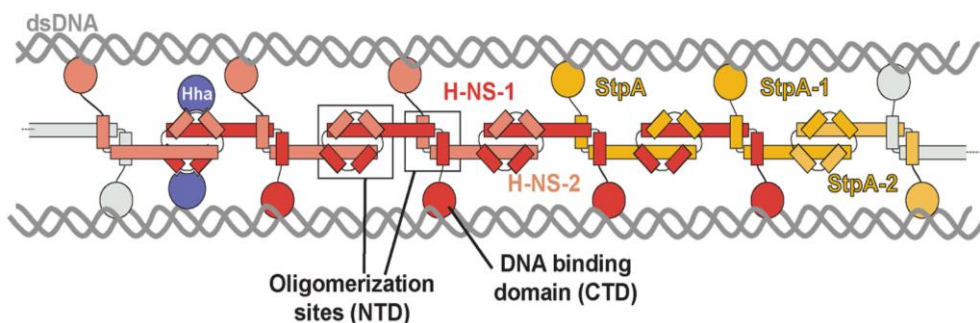
Estudios de resonancia magnética nuclear con el dominio N-terminal de la proteína H-NS han permitido identificar el aminoácido arginina R12 (altamente conservado entre las proteínas H-NS y StpA de enterobacterias) como el residuo esencial para la interacción con Hha (García y col., 2006). En el caso de Hha, el residuo clave en la interacción con H-NS es el ácido aspártico en la posición 48 (D48), además de los residuos R16, E25, R50 y P64 (Nieto y col., 2002). En el caso de YdgT, los residuos claves para la interacción con H-NS son la serina 15 (S15) y la glutamina en la posición 70 (Q70) (Yun y col., 2012). Cualquier cambio en los residuos anteriormente nombrados alteran la capacidad de unión de los complejos proteicos al ADN. Todo ello pone de manifiesto la capacidad que tienen las proteínas de la familia Hha de modular la actividad de unión de H-NS al ADN (figura 1.37).

## 1. INTRODUCCIÓN



**Figura 1.37.** Representación esquemática del complejo de unión de H-NS/Hha al ADN. Los monómeros de Hha (representados en naranja) flanquean ambos lados del dominio N-terminal de H-NS. Esta disposición proyecta las superficies de unión al ADN de Hha y H-NS en lados opuestos del oligómero. Imagen adaptada de (Singh y col., 2016).

La interacción entre las proteínas Hha e YdgT con la proteína H-NS no es únicamente exclusiva a estas dos proteínas, sino que también es extensible a otros miembros de ambas familias. Se han descrito interacciones entre Hha y StpA, YdgT y StpA (Paytubi y col., 2004) y entre YmoA y H-NS (Ellison y Miller, 2006) (figura 1.38).



**Figura 1.38.** Representación esquemática del modelo predicho para la interacción entre H-NS, Hha y StpA en la formación de nucleofilamentos sobre el ADN y la posterior formación de estructuras en forma de puente entre dos cadenas de ácidos nucleicos. Imagen adaptada de (Boudreau y col., 2018).

Todo ello pone de manifiesto la existencia de una relación funcional entre las proteínas de ambas familias. Esta interacción no se limita únicamente a procesos moduladores de la expresión génica, sino que también es responsable de otros mecanismos, como por ejemplo la interacción de Hha o YdgT con la proteína StpA, donde dicha interacción protege a StpA de la degradación proteolítica por parte de Lon (Paytubi y col., 2004), de la misma manera que lo hace su interacción con H-NS (Johansson y col., 2001).

El mecanismo por el cual las moléculas tipo Hha son capaces de regular la expresión génica en los *loci* regulados por H-NS sigue sin estar claro y existe un debate considerable

en cuanto a si las proteínas similares a Hha son capaces de realizar sus acciones uniéndose al ADN directamente o a través de una interacción con H-NS. A pesar de que la proteína Hha presenta motivos de hélice-giro-hélice que son frecuentes en proteínas que tienen la capacidad de unirse al ADN, muy pocos estudios han mostrado la capacidad de esta proteína de unirse al ADN por sí misma de manera independiente a H-NS (Fahlen y col., 2001; Sharma y Zuerner, 2004; Olekhnovich y Kadner, 2007; Sharma y Bearson, 2013; Sharma y Casey, 2014). A excepción de la unión observada de Hha al ADN en el caso del gen *rtsA* de *S. Typhimurium* (Olekhnovich y Kadner, 2007), para obtener las preparaciones de proteína Hha utilizadas en el resto de ensayos de unión *in vitro* no se utilizó un fondo genético  $\Delta hns$ , por lo que no se puede garantizar que Hha no se haya copurificado con H-NS y los resultados de unión al ADN se produzcan de manera indirecta vía unión H-NS. Además, los estudios realizados en *E. coli* que indicaron que Hha puede afectar la expresión génica en ausencia de H-NS, se realizaron en células que expresaban StpA, por lo que no es descartable que Hha pudiera estar ejerciendo su función reguladora uniéndose a esta proteína de manera alternativa.

Lo que parece evidente es que la interacción de las proteínas tipo Hha con las proteínas de la familia H-NS permite el establecimiento de complejos reguladores que presentan propiedades específicas que las proteínas tipo H-NS por sí solas no tendrían (Ali y col., 2013).

Recientemente se ha descrito que las interacciones entre Hha y el dominio N-terminal de la proteína H-NS incrementan la afinidad de unión de H-NS al ADN, además de incrementar la capacidad que presenta esta proteína para formar puentes entre cadenas de ADN, incluso en ausencia de iones  $Mg^{2+}$  (Ali y col., 2013; van der Valk y col., 2017). De esta manera, al incrementar la capacidad del macrocomplejo proteico para formar puentes entre cadenas de ADN, tanto Hha como StpA son capaces de promover eventos de pausa de la ARN polimerasa una vez H-NS se encuentra unida al ADN en forma de nucleofilamentos (Boudreau y col., 2018). Todo ello refuerza el concepto de que la interacción de las proteínas tipo Hha con las proteínas tipo H-NS es esencial para la regulación de la expresión génica en enterobacterias en respuesta a cambios en las condiciones ambientales del medio, donde el principal grupo de genes regulados por los complejos Hha-H-NS son genes adquiridos a través de mecanismos de transferencia génica horizontal (Baños y col., 2009; Ali y col., 2013).

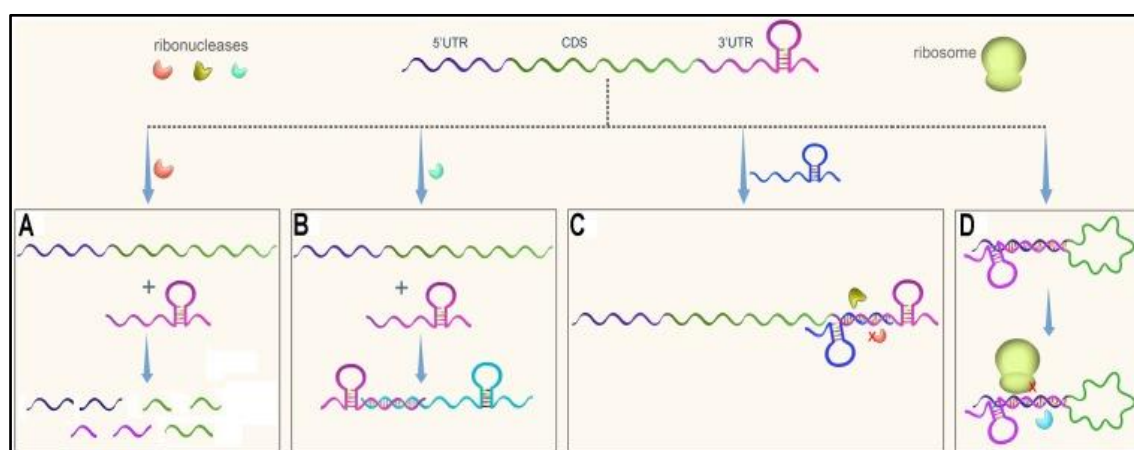
### 1.3 Regulación post-transcripcional por secuencias 3' UTR en bacterias

El papel principal de las moléculas de ARNm es transmitir la información genética. No obstante, además de la secuencia codificante de los correspondientes polipéptidos, dichas moléculas contienen en sus extremos no traducidos, tanto en el extremo 5'UTR (5' *untranslated region*) como en el extremo 3'UTR (3' *untranslated region*), secuencias que pueden jugar un papel muy importante en la regulación post-transcripcional de la expresión génica. En eucariotas, la interacción de proteínas y/o moléculas de ARN tiene lugar en ambas regiones (Pesole y col., 2001). En cambio, en procariotas, los mecanismos clásicos de regulación post-transcripcional han sido asociados principalmente a interacciones en la región 5'UTR (Winkler, 2005; Waters y Storz, 2009; Storz y col., 2011), región en la que se producen mecanismos de atenuación y donde actúan termómetros de ARN y ribointerruptores. Además, en dicha región también se unen proteínas de unión al ADN para modular la accesibilidad de los ribosomas a los RBSs (*ribosome binding site*).

En el mundo procariota se ha considerado tradicionalmente que las regiones 3'UTR contenían principalmente terminadores transcripcionales tanto dependientes como independientes de Rho. No obstante, en los últimos años, debido al desarrollo y a la mayor disponibilidad de una tecnología de secuenciación de alta eficiencia, ha empezado a ponerse de manifiesto que secuencias bacterianas 3'UTR juegan papeles importantes en la regulación de la expresión génica más allá de la terminación transcripcional (revisado en (Ren y col., 2017)).

Entre los diferentes papeles que juegan en bacterias las regiones 3'UTR, uno de los fundamentales es convertirse en reservorios de pequeños ARNs regulatorios (Miyakoshi y col., 2015), los cuales se generan bien por un proceso de procesamiento del ARN o por el eventos de inicio de la transcripción a partir de promotores internos presentes en la pauta abierta de lectura del gen. Algunos de tales ARNs generados a partir de secuencias 3'UTR pueden actuar en *trans* sobre otros genes, o interaccionar con la secuencia 5'UTR de su propio promotor para modular la expresión del gen en cuestión (Ren y col., 2017) (figura 1.39). Un ejemplo de ello tiene lugar en la región 3'UTR del gen *icaR* de *Staphylococcus aureus*, que es capaz de interaccionar con su región 5'UTR para regular su traducción (Ruiz de los Mozos y col., 2013). Las secuencias 3'UTR también presentan

la capacidad de regular la expresión génica al permitir la unión de proteínas que se unen al ARN, las cuales a su vez pueden mediar la unión de otras proteínas efectoras (Mayr, 2017). Teniendo en cuenta que en bacterias gramnegativas la chaperona Hfq se asocia a muchos de los pequeños ARNs regulatorios facilitando su papel regulador al protegerlos de la degradación ribonucleotídica (Kavita y col., 2018), ha podido ponerse de manifiesto que algunas secuencias 3'UTR bacterianas presentan dianas para Hfq (Chao y col., 2012).



**Figura 1.39.** Mecanismos implicados en la regulación post-transcripcional de la expresión génica por parte de la región 3'UTR en bacterias. (A) La región 3'UTR puede ser reconocida por una ribonucleasa que inicia la degradación del ARN mensajero. (B) La región 3'UTR puede comportarse como un reservorio de pequeños ARNs que presentan capacidad de regular la expresión génica. (C) La región 3'UTR puede ser diana de pequeñas moléculas de ARN reguladoras que actúen sobre dicha región protegiéndola de la degradación o, en sentido contrario, estimulando su degradación. (D) La región 3'UTR puede interactuar con la región 5'UTR de su propio gen, estimulando la iniciación de su traducción o incrementando la estabilidad del ARN mensajero del mismo. Imagen adaptada de (Ren y col., 2017).

Otros estudios han puesto de manifiesto que secuencias 3'UTR pueden modular la expresión génica actuando en *cis*. En este caso, la regulación de la degradación del ARNm es un mecanismo importante de control de la expresión génica a nivel post-transcripcional, ya que el procesamiento endonucleotídico inicial de la molécula de ARNm dictará la degradación del transcrito catalizada por exoribonucleasas (Laalami y col., 2013). Teniendo en cuenta que los ribosomas protegen al ARNm de la degradación endonucleolítica (Deneke y col., 2013), las regiones no traducidas pueden ser dianas de las ribonucleasas. Aunque las regiones 5'UTR suelen estar más implicadas en la inestabilidad del transcrito que las regiones 3'UTR (Selinger y col., 2003), también se han reportado diferentes casos de inestabilidad asociada a la región 3'UTR, como en el caso del gen *aceA* de *C. glutamicum* (Maeda y Wachi, 2012), cuya región 3'UTR es diana de la ARNasa E/G, o la región 3'UTR del gen *hild* de *Salmonella enterica*, cuya

## 1. INTRODUCCIÓN

secuencia es diana del degradosoma, donde la actividad nucleolítica de este complejo reduce la expresión de *hilD* (Lopez-Garrido y col., 2014). Asimismo, la región 3'UTR del gen *hmsT* de *Y. pestis* (que codifica para una diguanilato ciclasa que estimula la formación de *biofilm* en este microorganismo), de 283 nucleótidos, modula negativamente la degradación del ARNm de *hmsT*, implicando una PNPasa (Zhu y col., 2016). Estudios posteriores han identificado en *Y. pestis* nuevas regiones 3'UTR de longitud significativa y de elevado porcentaje AU. Las mismas parecen también potenciar la degradación del ARNm de los genes situados aguas arriba (Zhao y col., 2018). Por el contrario, determinadas regiones 3'UTR también pueden ayudar a incrementar la vida media de un determinado ARNm al ser dianas de otros ARNs pequeños con capacidad de proteger al transcrito del complejo de degradación (Ren y col., 2017) (figura 1.39).

## **2. OBJETIVOS**





En el momento del inicio de esta tesis doctoral, nuestro grupo de investigación, analizando el genoma de la cepa enteroagregativa *E. coli* 042 pudo comprobar que, además de los genes *hha* e *ydgT* (presentes en la mayoría de las enterobacterias), el cromosoma de dicha cepa contenía dos genes muy similares al gen *hha*, a los que se les denominó *hha2* y *hha3*. También pudo comprobarse que, además del gen *hns* y su parálogo *stpA*, la cepa *E. coli* 042 también codificaba un nuevo parálogo adicional de *hns*, que se denominó *hns2*. Una parte de los objetivos de esta tesis doctoral se han centrado en tratar de caracterizar el producto génico de estos nuevos parálogos con el fin de entender qué papel biológico presentan estas copias extra de los genes *hha* y *hns* dentro del cromosoma y las vías reguladoras de la bacteria.

También en el momento del inicio de este trabajo, el grupo de investigación estaba estudiando el papel de la alarmona (p)ppGpp en la expresión del gen *aggR* del plásmido pAA2 de la cepa *E. coli* 042. Tal y como se ha comentado en la Introducción de esta memoria, el producto del gen *aggR*, la proteína AggR, es un importante regulador de la virulencia bacteriana. Al realizar fusiones transcripcionales del gen *aggR* con el gen reportero *lacZ*, situando dicho gen marcador en el extremo 3' del gen *aggR*, pudimos observar cómo los clones obtenidos presentaban un fenotipo muy diferente al de la cepa salvaje. Dedicamos una parte de nuestro trabajo a estudiar este fenómeno y los cambios fenotípicos derivados del mismo, lo que la llevado a describir nuevos aspectos de la regulación de la expresión del gen *aggR* por secuencias localizadas aguas abajo de su pauta abierta de lectura.



### **3. MATERIALES Y MÉTODOS**



### 3.1 Cepas y plásmidos

Las diferentes cepas bacterianas utilizadas en este trabajo se detallan en la siguiente tabla:

Cepa	Descripción	Referencia
<i>E. coli</i> 042	Aislamiento clínico: Cm <sup>r</sup> Sm <sup>r</sup> Tc <sup>r</sup>	Prof. I.Henderson
<i>E. coli</i> 042 $\Delta$ lacZ	<i>E. coli</i> 042 mutante sencillo para el gen <i>lacZ</i>	Este trabajo
<i>E. coli</i> 042 $\Delta$ lacZ $\Delta$ cat	<i>E. coli</i> 042 mutante doble para los genes <i>lacZ</i> y <i>cat</i>	Este trabajo
<i>E. coli</i> MG1655	<i>K-12 F<sup>-</sup> <math>\lambda</math><sup>-</sup> ilvG<sup>-</sup> rfb-50 rph-1</i>	Guyer y col., 1981 Blattner y col., 1987
<i>E. coli</i> XL1 Blue	<i>recA1 endA1 gyrA96 thi1 hsdR17 (rk-mk+) supE44 relA1 <math>\lambda</math>- lac- (F' proA+ B+lacIqlacZ<math>\Delta</math>M15 Tn10) Tcr</i>	Bullock y col., 1987
<i>E. coli</i> 5K Rif	<i>F-, hsdR, hsdM, rpsL, thr, thi, leu, lac,</i> resistencia espontánea a Rifampicina	Nieto y col., 1998
<i>E. coli</i> BL21 (DE3)	<i>hsdS, gal, (<math>\lambda</math>cIts857, ind1, Sam7, nin5, lac-UV5-T7 gene)</i>	Studier y Moffatt, 1986
<i>E. coli</i> DH5 $\alpha$	<i>fhuA2 lac(del)U169 phoA glnV44 <math>\Phi</math>80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	Taylor y col., 1993
<i>E. coli</i> Ed1a	Cepa comensal de <i>E. coli</i>	Touchon y col., 2009
Colección LREC (Laboratorio de referencia de <i>E. coli</i> )	60 cepas de <i>E. coli</i> (59 cepas de origen humano y 1 de origen aviar) de los patotipos EAEC, ExPEC, STEC/VTEC, ETEC, tEPEC, EIEC, AIEC aisladas en España, Alemania, Francia, Dinamarca y EE.UU	Prof. Jorge Blanco
Colección EAEC de LREC (Laboratorio de referencia de <i>E. coli</i> )	56 cepas del patotipo EAEC aisladas en España, Alemania y Brasil	Prof. Jorge Blanco
Colección de cepas de <i>E. coli</i> productoras de ESBL	25 cepas de <i>E. coli</i> aisladas de humanos y animales de compañía en Alemania	Prof. Trinad Chakraborty
Colección ECOR	Colección de Referencia de <i>E. coli</i> con 72 cepas salvajes aisladas de diferentes hospedadores y áreas geográficas	Ochman y Selander, 1984b
Cepas ambientales <i>stx2</i> positivas	84 cepas de <i>E. coli</i> de origen ambiental portadoras del gen <i>stx2</i>	Dra. Maite Muniesa
Cepas ambientales <i>stx</i> negativas	88 cepas de origen ambiental no portadoras del gen que codifica para la toxina Stx	Dra. Maite Muniesa
<i>E. coli</i> 042 $\Delta$ hha	<i>E. coli</i> 042 mutante sencillo para el gen <i>hha</i>	Este trabajo
<i>E. coli</i> 042 $\Delta$ hha $\Delta$ hha2	<i>E. coli</i> 042 mutante doble para los genes <i>hha</i> y <i>hha2</i>	Este trabajo
<i>E. coli</i> 042 $\Delta$ hha $\Delta$ hha2 $\Delta$ hha3	<i>E. coli</i> 042 triple mutante de los genes <i>hha</i> , <i>hha2</i> y <i>hha3</i>	Este trabajo
<i>E. coli</i> 042 <i>hha::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína Hha	Este trabajo
<i>E. coli</i> 042 <i>hha2::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína Hha2	Este trabajo
<i>E. coli</i> 042 <i>hha3::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína Hha3	Este trabajo
<i>E. coli</i> 042 <i>ydgT::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína YdgT	Este trabajo

### 3. MATERIALES Y MÉTODOS

<i>E. coli</i> 042 $\Delta hns$	<i>E. coli</i> 042 mutante sencillo para el gen <i>hns</i>	Hüttener y col., 2014
<i>E. coli</i> 042 $\Delta hns2$	<i>E. coli</i> 042 mutante sencillo para el gen <i>hns2</i>	Este trabajo
<i>E. coli</i> 042 $\Delta stpA$	<i>E. coli</i> 042 mutante sencillo para el gen <i>stpA</i>	Este trabajo
<i>E. coli</i> 042 $\Delta hns$ $\Delta hns2$	<i>E. coli</i> 042 mutante doble para los genes <i>hns</i> y <i>hns2</i>	Este trabajo
<i>E. coli</i> 042 $\Delta hns$ $\Delta stpA$	<i>E. coli</i> 042 mutante doble para los genes <i>hns</i> y <i>stpA</i>	Este trabajo
<i>E. coli</i> MG1655 $\Delta hns$	<i>E. coli</i> MG1655 mutante sencillo para el gen <i>hns</i>	Este trabajo
<i>E. coli</i> 042 $\Delta lacZ$ <i>hns::lacZ</i>	<i>E. coli</i> 042 $\Delta lacZ$ con la fusión transcripcional cromosómica del gen reportero <i>lacZ</i> bajo el control del promotor del gen <i>hns</i>	Este trabajo
<i>E. coli</i> 042 $\Delta lacZ$ <i>hns2::lacZ</i>	<i>E. coli</i> 042 $\Delta lacZ$ con la fusión transcripcional cromosómica del gen reportero <i>lacZ</i> bajo el control del promotor del gen <i>hns2</i>	Este trabajo
<i>E. coli</i> 042 <i>hns::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS	Este trabajo
<i>E. coli</i> 042 <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> MG1655 <i>hns::3XFLAG</i>	<i>E. coli</i> MG1655 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS	Este trabajo
<i>E. coli</i> 042 $\Delta cat$ <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 $\Delta cat$ con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> 042 $\Delta cat$ $\Delta hns$ <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 $\Delta cat$ mutante sencillo para el gen <i>hns</i> con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> 042 $\Delta cat$ $\Delta lon$ <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 $\Delta cat$ mutante sencillo para el gen <i>lon</i> con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> 042 $\Delta cat$ $\Delta hns$ $\Delta lon$ <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 $\Delta cat$ mutante doble para los genes <i>hns</i> y <i>lon</i> con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> 042 $\Delta cat$ $\Delta stpA$ <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 $\Delta cat$ mutante sencillo para el gen <i>stpA</i> con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> 042 $\Delta cat$ $\Delta stpA$ $\Delta lon$ <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 $\Delta cat$ mutante doble para los genes <i>stpA</i> y <i>lon</i> con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> 042 $\Delta hns$ <i>phns2*</i>	<i>E. coli</i> 042 mutante sencillo para el gen <i>hns</i> con la transición C→T generada en la caja -10 del promotor del gen <i>hns2</i>	Este trabajo
<i>E. coli</i> 042 $\Delta hns$ <i>phns2*</i> <i>hns2::3XFLAG</i>	<i>E. coli</i> 042 mutante sencillo para el gen <i>hns</i> con la transición C→T generada en la caja -10 del promotor del gen <i>hns2</i> con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína H-NS2	Este trabajo
<i>E. coli</i> 042 pAA2-	<i>E. coli</i> 042 curada del plásmido pAA2	Este trabajo
<i>E. coli</i> 042 <i>aafA::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína AafA	Este trabajo
<i>E. coli</i> 042 $\Delta lacZ$ <i>aggR::lacZ</i>	<i>E. coli</i> 042 $\Delta lacZ$ con la fusión transcripcional cromosómica del gen reportero <i>lacZ</i> bajo el control del promotor del gen <i>aggR</i>	Hüttener y col., 2018



<i>E. coli</i> 042 $\Delta lacZ$ <i>aggR3'UTR::lacZ</i>	<i>E. coli</i> 042 $\Delta lacZ$ con la fusión transcripcional cromosómica del gen reportero <i>lacZ</i> situada después del codón de terminación de la traducción del gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 <i>aggR::3XFLAG</i>	<i>E. coli</i> 042 con el epítipo 3XFLAG insertado en el extremo C-terminal de la proteína AggR	Hüttener y col., 2018
<i>E. coli</i> 042 <i>aggR::Km</i>	<i>E. coli</i> 042 con un casete de resistencia a kanamicina insertado en la región codificante del gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 <i>aggR+Km-r3'UTR</i>	<i>E. coli</i> 042 con un casete de resistencia a kanamicina insertado después del codón de terminación de la traducción del gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 $\Delta aggR$	<i>E. coli</i> 042 mutante sencillo para el gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 <i>aggR+FRT3'UTR</i>	<i>E. coli</i> 042 con la secuencia FRT insertada después del codón de terminación de la traducción del gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 $\Delta 3'UTRaggR$	<i>E. coli</i> 042 mutante para la región 3'UTR del gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 $\Delta IRL ISIA$	<i>E. coli</i> 042 mutante para la región IRL del elemento de inserción <i>ISIA</i> situado aguas abajo del gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 $\Delta 3'UTRaggR \Delta IRL ISIA$	<i>E. coli</i> 042 mutante para la región 3'UTR del gen <i>aggR</i> y la región IRL del elemento de inserción <i>ISIA</i> situado aguas abajo del gen <i>aggR</i>	Este trabajo
<i>E. coli</i> 042 $\Delta ISIA$	<i>E. coli</i> 042 mutante para el elemento de inserción <i>ISIA</i> situado aguas abajo del gen <i>aggR</i>	Este trabajo
<i>S. Typhimurium</i> SV5015	<i>Salmonella enterica</i> serovar Typhimurium SL1344 <i>his+</i>	Prof. Josep Casadesús

**Tabla 3.1.** Listado de las diferentes cepas bacterianas utilizadas durante el desarrollo de este trabajo.

### 3. MATERIALES Y MÉTODOS

Los diferentes plásmidos utilizados en este trabajo se detallan a continuación:

Plásmido	Descripción	Referencia
pKD3	<i>bla</i> (Cb <sup>r</sup> ) FRT <i>cat</i> FRT PS1 PS2 oriR6K Cm <sup>r</sup>	Datsenko y Wanner, 2000
pKD4	<i>bla</i> (Cb <sup>r</sup> ) FRT <i>ahp</i> FRT PS1 PS2 oriR6K Km <sup>r</sup>	Datsenko y Wanner, 2000
pKD46	<i>bla</i> (Cb <sup>r</sup> ) P <sub>BAD</sub> <i>gam bet exo</i> pSC101 oriTS	Datsenko y Wanner, 2000
pCP20	<i>bla</i> (Cb <sup>r</sup> ) <i>cat</i> $\lambda$ <i>cI857 IP<sub>R</sub>flp</i> pSC101 oriTS	Cherepanov y col., 1995
pSUB11	R6KoriV 3XFLAG-Km <sup>r</sup>	Uzzau y col., 2001
pKG136	<i>ahp</i> FRT <i>lacZY</i> + <i>t<sub>his</sub></i> oriR6K	Ellermeier y col., 2002
pLATE 51	P <sub>T7</sub> <i>lacO rep</i> (pMB1) <i>bla</i> (Cb <sup>r</sup> ) 6xHis N-terminal	Thermo Scientific
pLATE 31	P <sub>T7</sub> <i>lacO rep</i> (pMB1) <i>bla</i> (Cb <sup>r</sup> ) 6xHis C-terminal	Thermo Scientific
pLATE51 6HIS-Hha2	Gen <i>hha2</i> clonado con el epítipo 6Histidinas en el extremo N-terminal bajo el promotor de la ARN polimerasa T7 en el plásmido pLATE51	Este trabajo
pLATE31 Hha2-6HIS	Gen <i>hha2</i> clonado con el epítipo 6Histidinas en el extremo C-terminal bajo el promotor de la ARN polimerasa T7 en el plásmido pLATE31	Este trabajo
pLATE51 6HIS-H-NS2	Gen <i>hns2</i> clonado con el epítipo 6Histidinas en el extremo N-terminal bajo el promotor de la ARN polimerasa T7 en el plásmido pLATE51	Este trabajo
pLATE31 H-NS2-6HIS	Gen <i>hns2</i> clonado con el epítipo 6Histidinas en el extremo C-terminal bajo el promotor de la ARN polimerasa T7 en el plásmido pLATE31	Este trabajo
pLG338.30	oriPSC101 Cb <sup>r</sup>	Cunningham y col., 1993
pLG338.30 <i>hns</i>	Gen <i>hns</i> clonado en el vector pLG338.30	Este trabajo
pLG338.30 <i>hns2</i>	Gen <i>hns2</i> clonado en el vector pLG338.30	Este trabajo
pLG338.30 <i>hns2</i> *	Gen <i>hns2</i> clonado en el vector pLG338.30 bajo el control del promotor con la transición C→T generada en la caja -10	Este trabajo
pLG338.30 <i>aggR</i>	Gen <i>aggR</i> clonado en el vector pLG338.30	Este trabajo
pLG338.30 <i>aggR</i> + <i>FRT3</i> ' <i>UTR</i>	Gen <i>aggR</i> con la cicatriz FRT insertada después del codón de terminación del gen clonado en el vector pLG338.30	Este trabajo
pFZY1	F <sup>r</sup> <i>lac</i> replicon <i>lacZYA</i> + Cb <sup>r</sup>	Koop y col., 1987
pFZY1 <i>phns2</i> wt	Región promotora del gen <i>hns2</i> clonada en el vector pFZY1 generando una fusión transcripcional con el gen <i>lacZ</i>	Este trabajo
pFZY1 <i>phns2</i> *	Región promotora del gen <i>hns2</i> con la transición C→T generada en la caja -10 clonada en el vector pFZY1 generando una fusión transcripcional con el gen <i>lacZ</i>	Este trabajo
pCR2.1-TOPO	P <sub>lac</sub> , <i>oriF1</i> , Cb <sup>r</sup> , Km <sup>r</sup>	Invitrogen

**Tabla 3.2.** Listado de los diferentes plásmidos utilizados durante el desarrollo de este trabajo.

### 3.2 Medios de cultivo

Los diferentes medios de cultivo bacteriano (y su composición) utilizados a lo largo de este trabajo fueron los siguientes:

- LB (Luria-Bertani): medio de cultivo rico utilizado de manera rutinaria en el crecimiento en suspensión de *E. coli*. Su composición es 10 gr/l de NaCl, 10 gr/l de peptona trípica de caseína y 5 gr/l de extracto de levadura.
- LB agar: medio de cultivo sólido utilizado para el crecimiento bacteriano en placa. Se obtiene mediante la adición de agar a una concentración de 15 gr/l al medio LB.
- LB agar-X-Gal: medio LB agar suplementado con X-Gal (*5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside*), utilizado a una concentración final de 40 µg/ml con la finalidad de analizar la actividad β-galactosidasa de las bacterias sembradas en placa.
- LB agar motilidad: medio de cultivo semisólido utilizado para realizar ensayos de motilidad bacteriana en placa. Se obtiene al añadir 3 gr/l de agar al medio LB.
- Medio SOB (*Super Optimal Broth*): medio de cultivo líquido utilizado para hacer crecer *E. coli* en los experimentos de mutagénesis por reemplazamiento alélico. Su composición es 20 gr/l de peptona trípica de caseína, 5 gr/l de extracto de levadura, 0,19 gr/l de KCl, 0,58 gr/l de NaCl y 20 mM de sales de Mg<sup>2+</sup> (10 mM de MgCl<sub>2</sub> y 10 mM de MgSO<sub>4</sub>).
- Medio SOC (*Super Optimal Broth with Catabolite repression*): medio de cultivo líquido utilizado para la recuperación de las células después de ser sometidas a un pulso eléctrico tras el proceso de transformación bacteriana mediante electroporación. Se obtiene al suplementar el medio SOB con glucosa 0,45% (p/v).
- 2xYT: medio de cultivo líquido utilizado para realizar los experimentos de conjugación plasmídica. Su composición es 16 gr/l de peptona trípica de caseína, 10 gr/l de extracto de levadura y 5 gr/l de NaCl.
- DMEM: Medio de cultivo comercial (*Gibco*) utilizado para el crecimiento y mantenimiento de células eucariotas. Se utilizó para realizar ensayos de *biofilm* y de crecimiento bacteriano, suplementado con una solución de glucosa al 0,45% (p/v) como fuente de carbono.

### 3. MATERIALES Y MÉTODOS

- **M9:** Medio de cultivo líquido utilizado para el crecimiento de *E. coli* en condiciones limitantes de nutrientes. Su composición es la siguiente: 2 mM de Na<sub>2</sub>HPO<sub>4</sub>, 22 mM de KH<sub>2</sub>PO<sub>4</sub>, 8,5 mM de NaCl, 18,5 mM de NH<sub>4</sub>Cl, 2 mM de MgSO<sub>4</sub>, 0,1 mM de CaCl<sub>2</sub>. Para suplementar este medio mineral (y generar un medio de cultivo mínimo) con una fuente de carbono se utilizó glucosa a una concentración final de 0,4% (p/v).

### 3.3 Métodos de esterilización

A excepción de los compuestos termolábiles, los medios de cultivo anteriormente descritos, junto con el material de plástico y vidrio utilizados a lo largo de este trabajo, fueron esterilizados por calor húmedo a elevada presión en el autoclave a 121°C y 1 atmósfera de presión durante 20 minutos. En el caso de los medios de cultivo que no pudieron ser esterilizados a elevadas temperaturas, se esterilizaron mediante filtración haciendo uso de filtros que presentaban un tamaño de poro de 0,22 µm.

### 3.4 Antibióticos

Los antibióticos utilizados en este trabajo (con la finalidad de suplementar a los medios de cultivo anteriormente descritos) han sido un elemento esencial a la hora de seleccionar los diferentes mutantes utilizados. La preparación de dichos antibióticos (concentraciones *stock* y de uso, medio de preparación y método de esterilización de los mismos) queda resumida en la siguiente tabla:

Antibiótico*	Medio de preparación	Método de esterilización	Concentración <i>stock</i> (mg/ml)	Concentración de uso (µg/ml)
<b>Ampicilina</b>	Agua <i>milli-Q</i>	Filtro de 0,22 µm	100	100/50
<b>Carbenicilina</b>	Agua <i>milli-Q</i>	Filtro de 0,22 µm	100	100/50
<b>Cloranfenicol</b>	Etanol 100%	-	100	25
<b>Kanamicina</b>	Agua <i>milli-Q</i>	Filtro de 0,22 µm	50	50
<b>Rifampicina</b>	Metanol 100 %	-	50	50
<b>Novobiocina</b>	Agua <i>milli-Q</i>	Filtro de 0,22 µm	15	3,3

\*todos los antibióticos utilizados fueron adquiridos de la casa comercial *Sigma-Aldrich*.

**Tabla 3.3.** Listado de los antibióticos utilizados para suplementar los diferentes medios de cultivo utilizados durante este trabajo, donde se detalla el medio de preparación y el método de esterilización de los mismos, además de las concentraciones *stock* y de uso final.

### **3.5 Mantenimiento de los microorganismos**

Todas las cepas bacterianas utilizadas en este trabajo se mantuvieron congeladas a  $-80^{\circ}\text{C}$  en una solución de medio de cultivo rico LB suplementado con glicerol al 20% (v/v). Para reconstituir los correspondientes cultivos, se descongelaron las cepas específicas en placas de LB agar suplementadas con el antibiótico correspondiente. Las placas fueron incubadas a  $30^{\circ}\text{C}$  o  $37^{\circ}\text{C}$  (dependiendo de los requerimientos del microorganismo a utilizar) durante 12-16 horas.

### **3.6 Métodos de transferencia génica**

La incorporación o introducción de material genético de manera exógena a las células bacterianas se realizó mediante dos métodos diferentes: transformación de células competentes con  $\text{CaCl}_2$  a  $4^{\circ}\text{C}$  y posterior choque térmico a  $42^{\circ}\text{C}$  (Cohen y col., 1972), o transformación mediante electroporación (este último es el método que mayor eficiencia de transformación nos permite obtener) (Dower y col., 1988).

#### **3.6.1 Transformación de células competentes mediante tratamiento en frío con $\text{CaCl}_2$ y posterior choque térmico**

A partir de un cultivo crecido durante 16 horas en medio LB a  $30^{\circ}\text{C}/37^{\circ}\text{C}$ , se realizó una dilución 1:100 del mismo en medio de cultivo fresco y se dejó crecer en constante agitación a  $30^{\circ}\text{C}/37^{\circ}\text{C}$  hasta una fase exponencial temprana ( $\text{D.O}_{600\text{nm}}$  de 0,2-0,3). Posteriormente, se incubaron las células en hielo y se centrifugaron a 4.000 rpm durante 10 minutos a  $4^{\circ}\text{C}$ . Seguidamente, se resuspendió el sedimento celular con 20 ml de  $\text{CaCl}_2$  50 mM frío, y se repitió el mismo protocolo, realizándose tres lavados celulares sucesivos con el mismo volumen de  $\text{CaCl}_2$ . Después del último lavado, el sedimento final se resuspendió con 150  $\mu\text{l}$  de  $\text{CaCl}_2$ . La suspensión celular se incubó durante 30 minutos en hielo antes de realizar el protocolo de transformación. A continuación, se mezclaron 60  $\mu\text{l}$  de estas células competentes con 1-10  $\mu\text{l}$  (100 ngr-1  $\mu\text{gr}$ ) de la solución de ADN plasmídico a transformar. Dicha mezcla se incubó durante 30 minutos adicionales en hielo y a continuación se sometió a un choque térmico a  $42^{\circ}\text{C}$  durante 2 minutos. Finalmente, se añadió medio LB fresco a la suspensión celular hasta completar un volumen final de 1 ml y se incubó durante 1 hora a  $30/37^{\circ}\text{C}$  en agitación constante para permitir la expresión de los marcadores fenotípicos transferidos mediante el choque térmico, que nos permitirán seleccionar las células correctamente transformadas al plaquearlas en medio sólido (placas de LB-agar) suplementado con el agente selectivo correspondiente.

#### **3.6.2 Electroporación**

La electroporación es un método de transformación bacteriana que permite la captación de ADN exógeno por parte de las bacterias mediante la creación de poros en la membrana celular (permeabilización) provocados por una descarga eléctrica. Es un método más eficiente que el tratamiento celular con CaCl<sub>2</sub> frío ya que permite obtener hasta 10<sup>9</sup>-10<sup>10</sup> transformantes/μg de ADN transferido (Dower y col., 1988).

A partir de un cultivo bacteriano crecido durante 16 horas a 30°C/37°C, se realizó una dilución 1:100 del mismo en medio de cultivo fresco y se dejó crecer en constante agitación a 30/37°C hasta una densidad óptica (D.O<sub>600nm</sub>) de 0,6-0,8. Posteriormente, se incubaron las células en hielo y se centrifugaron a 4.000 rpm durante 10 minutos a 4°C. Seguidamente, se resuspendió el sedimento celular con 20 ml de glicerol al 10% (v/v) frío, y se realizaron varias centrifugaciones sucesivas a 4.000 rpm a 4°C para tratar de disminuir la fuerza iónica de la suspensión celular (en total, tres lavados sucesivos con glicerol frío). Finalmente, se mezclaron 60 μl de la suspensión celular con 1 μgr de ADN, y se mantuvieron en hielo 30 minutos hasta el momento de la electroporación. La suspensión celular con el ADN se transfirió a una cubeta de electroporación pre esterilizada (*Electroporation Cuvette*, 1mm Gap, MβP) y se sometió a un pulso eléctrico de 1250 V en el electroporador *Electroporator 2510* (Eppendorf®). Una vez sometidas las células al pulso eléctrico, se añadieron 940 μl de medio de cultivo fresco (para completar un volumen final de 1 ml de suspensión celular) y la mezcla se incubó durante 1 hora a 30/37°C para permitir la expresión de los marcadores fenotípicos transferidos. Posteriormente, se plaqueó la suspensión celular en placas de LB agar suplementadas con los antibióticos correspondientes para seleccionar únicamente aquellas células que hubieran adquirido los marcadores específicos transferidos mediante la electroporación.

### **3.7 Técnicas experimentales con ADN**

#### **3.7.1 Amplificación de ADN mediante la reacción en cadena de la polimerasa (PCR)**

Como molde a amplificar mediante la reacción en cadena de la polimerasa (PCR) se utilizó ADN proveniente de un aislamiento plasmídico o bien una suspensión de colonias bacterianas disueltas en 100 μl de agua estéril desionizada. En función del uso posterior al que se quería dar al fragmento amplificado mediante la técnica de PCR, se utilizaron diferentes ADN polimerasas comerciales:

- Cuando dichos fragmentos de ADN amplificado fueron requeridos para su posterior uso (por ejemplo, para realizar el clonaje de un gen específico) se utilizó la polimerasa de alta fidelidad de copia *Phusion Hot Start II High-Fidelity DNA Polymerase* (*Thermo Scientific*). Esta enzima posee actividad polimerasa 5'→3', además una actividad exonucleasa 3'→5', que le otorga la capacidad de alta fidelidad de copia con una tasa de error de  $4.4 \times 10^{-7}$ . El protocolo diseñado para el uso de dicha polimerasa se realizó para un volumen de 50 µl con los siguientes componentes:
  - 1 µl de solución de dNTPs (dATP, dGTP, dCTP, dTTP, *Thermo Scientific*) a una concentración final de 0,2 mM.
  - 10 µl de MgCl<sub>2</sub> a una concentración final de 1,5 mM.
  - 0,5 µl de oligonucleótido *forward* y 0,5 µl de oligonucleótido *reverse* a una concentración final de 0,5 µM.
  - 0,5 µl de la solución comercial de *Phusion Hot Start II High-Fidelity DNA Polymerase* a una concentración de 2 U/µl.
  - 35 µl de agua desionizada.
  - 1-10 pgr de ADN molde a amplificar.
  
- En cambio, cuando las PCRs se diseñaron con la finalidad de realizar el genotipado de una determinada cepa (por ejemplo, confirmar una mutagénesis), para la amplificación de los fragmentos de ADN no fue necesario el uso de una polimerasa de alta fidelidad de copia. En este caso se utilizó la *DreamTaq Green PCR Master Mix* (*Thermo Scientific*). Esta solución comercial ya contiene MgCl<sub>2</sub> y dNTPs a una concentración final de 2 mM y 0,2 mM respectivamente, además de la enzima Taq ADN Polimerasa. El protocolo recomendado por el fabricante para el uso de la *DreamTaq Green PCR Master Mix* se realizó para una reacción de 25 µl con los siguientes componentes:
  - 12,5 µl de solución *Dream Taq Master Mix 2x*.
  - 1 µl de oligonucleótido *forward* y 1 µl de oligonucleótido *reverse* a una concentración final de 0,5 µM.
  - 9,5 µl agua desionizada.
  - 10-100 pgr de ADN molde a amplificar.

El tiempo y la temperatura utilizados en el ciclo de hibridación de los oligonucleótidos al molde a amplificar variaban en función de la longitud y secuencia de los oligonucleótidos utilizados, mientras que el tiempo de extensión por parte de la ADN polimerasa variaba



### 3. MATERIALES Y MÉTODOS

en función del tamaño del fragmento a amplificar y de la velocidad de amplificación de la polimerasa utilizada. En el caso de la *DreamTaq Green PCR Master Mix*, el programa básico utilizado fue de una primera etapa de desnaturalización a 95°C durante 5 minutos, y 30 ciclos repetidos iniciados con una breve etapa (30 segundos) de desnaturalización del ADN a 95°C, hibridación de los oligonucleótidos durante 30 segundos a 55°C y un tiempo de extensión variable entre 30 segundos y 3 minutos (en función del tamaño del gen diana a amplificar, asumiendo una velocidad de amplificación por parte de la polimerasa de 1 minuto/kilobase) a 72°C. El ciclo terminaba con una última etapa de extensión (para acabar de amplificar el ADN no amplificado previamente) de 10 minutos a 72°C. En cambio, en el caso *Phusion Hot Start II High-Fidelity DNA Polymerase*, el programa de amplificación utilizado consistía en una primera etapa de desnaturalización a 98°C durante 5 minutos, y 25 ciclos repetidos iniciados con una etapa de 30 segundos de desnaturalización del ADN a 98°C, una etapa de hibridación de los oligonucleótidos durante 30 segundos a 55°C y una etapa de extensión variable de 15 segundos a 2 minutos (en función del tamaño del gen diana a amplificar, teniendo en cuenta que la velocidad de polimerización de la enzima es de 15-30 segundos/kilobase) a 72°C. El ciclo terminaba con una última etapa de extensión de 10 minutos a 72°C para acabar de amplificar el ADN residual no amplificado durante los ciclos mencionados anteriormente.

#### **3.7.2 Purificación de fragmentos de ADN amplificados por PCR**

Los productos de PCR obtenidos fueron purificados mediante el kit comercial *GeneJET PCR Purification Kit (Thermo Scientific)* siguiendo las instrucciones indicadas por el fabricante. El fundamento del kit utilizado se basa en la adsorción específica de los fragmentos de ADN a purificar en una membrana de sílica con el fin de eliminar los restos de oligonucleótidos, dNTPs, enzimas y sales de los diferentes componentes utilizados para realizar la reacción de amplificación por PCR.

#### **3.7.3 Aislamiento de ADN plasmídico**

El ADN plasmídico fue aislado utilizando el kit comercial *GeneJET Plasmid Miniprep Kit (Thermo Scientific)* siguiendo las instrucciones recomendadas por el fabricante. Dicho aislamiento se basa en el protocolo de lisis alcalina seguido de la neutralización del medio en presencia de una elevada concentración de sal con el fin de provocar la ruptura celular, la desnaturalización y precipitación del ADN cromosómico y las proteínas, además de la liberación del ADN plasmídico. Seguidamente, el material plasmídico se hizo pasar por

una columna con una membrana de sílica en la que el ADN quedó adsorbido, se lavó posteriormente con etanol al 70% para eliminar restos de elementos contaminantes y finalmente fue eluído en tampón de elución (Tris-HCl 10 mM a pH 8,5).

### 3.7.4 Cuantificación de ácidos nucleicos

Para llevar a cabo un análisis cuantitativo y cualitativo del ADN se realizó un análisis espectrofotométrico del mismo. La cuantificación de la concentración y de la pureza del ADN total de cada muestra purificada se realizó mediante la determinación de la absorbancia a  $D.O_{260nm}$  en el espectrofotómetro *NanoDrop® ND-1000 UV-Vis Spectrophotometer ND-1000 (Thermo Scientific)*. Además, también se midió la absorbancia a  $D.O_{280nm}$ , ya que la relación  $D.O_{260nm}/D.O_{280nm}$  informa sobre el grado de pureza del ADN. Si dicha relación es inferior a 2, es indicativo de que hay contaminación proteica de la muestra. Paralelamente, también se cuantificó el ratio  $D.O_{260nm}/D.O_{230nm}$  con el fin de anotar la presencia de residuos fenólicos u orgánicos en la muestra.

### 3.7.5 Electroforesis de ADN en geles de agarosa

Los productos de PCR obtenidos se analizaron mediante electroforesis en geles de agarosa en una cubeta horizontal. La matriz de agarosa se preparó con tampón TBE 0,5x (45 mM de Tris, 45 mM de ácido bórico y 1 mM de EDTA a un pH final de 8,3). La concentración de agarosa utilizada en los geles varió en función de los tamaños del ADN a analizar: para tamaños inferiores a 1 Kb se utilizaron geles con un 2% (p/v) de agarosa, mientras que para tamaños superiores se utilizó una concentración de agarosa de un 0,8% (p/v). Conjuntamente con las muestras a analizar, también se cargaron los marcadores de peso molecular correspondientes, que permiten realizar una estimación del tamaño de las moléculas de ADN presente en las muestras. En función del tamaño de los fragmentos a analizar y del porcentaje de agarosa utilizado para realizar la electroforesis se utilizaron los siguientes marcadores:

- *GeneRuler 1 Kb DNA Ladder (Thermo Scientific)* para analizar fragmentos de 1 a 10 Kb en un gel de agarosa al 0,8% (p/v).
- *GeneRuler 100 pb DNA Ladder (Thermo Scientific)* para analizar fragmentos de hasta 1 Kb en un gel de agarosa al 2% (p/v).

La electroforesis de ADN se llevó a cabo aplicando un voltaje de 100 V durante 30 minutos. Tras la electroforesis, los geles fueron sumergidos durante 20 minutos en una solución de 0,5  $\mu$ g/ml de bromuro de etidio (*Sigma Aldrich*) disuelto en tampón TBE

### 3. MATERIALES Y MÉTODOS

0,5x. Una vez teñidos los geles, se observaron bajo la luz ultravioleta y se tomaron las imágenes correspondientes utilizando el equipo *Molecular Imager® Gel Doc™ XR+* (Bio-Rad).

#### **3.7.6 Secuenciación de fragmentos de ADN**

Cuando fue necesario disponer de la secuencia específica de un fragmento cromosómico amplificado por PCR, o comprobar el correcto clonaje de un inserto en un vector, se enviaron las muestras correspondientes al servicio de secuenciación llevado a cabo por la empresa *Eurofins Genomics* en el Parc Científic de Barcelona. Los componentes necesarios para cada muestra enviada a secuenciar fueron los siguientes:

- 5 µl de producto de PCR previamente purificado a una concentración de 20-80 ngr/µl, o 5 µl de plásmido a una concentración de 80-100 ngr/µl.
- 5 µl del oligonucleótido complementario a la cadena de ADN a secuenciar a una concentración final de 5 µM.

El método de secuenciación utilizado ha sido el método de secuenciación por ciclo (una modificación del método tradicional de secuenciación por Sanger) haciendo uso del aparato de secuenciación *ABI3730XL DNA Analyzer*.

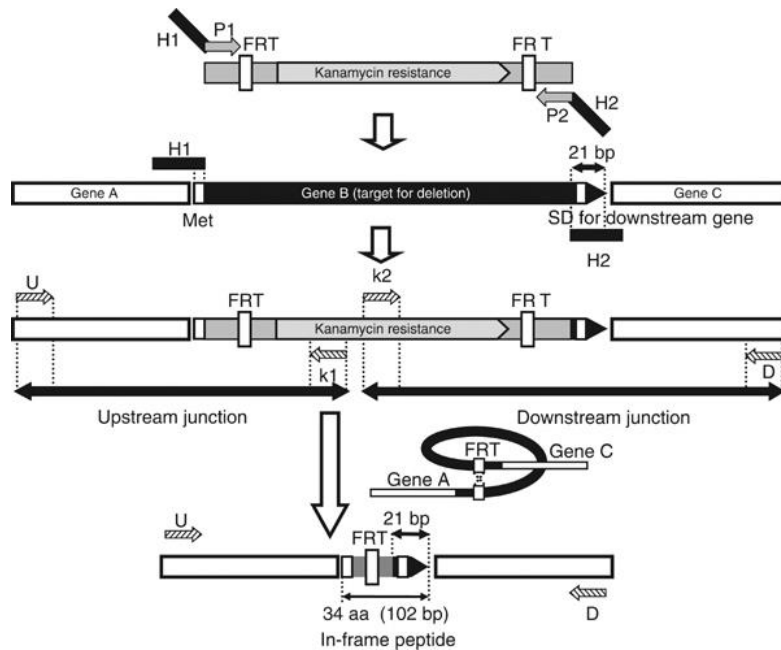
#### **3.7.7 Mutagénesis bacteriana: inactivación de genes cromosómicos utilizando fragmentos de PCR (reemplazamiento alélico)**

A diferencia de otros microorganismos como, por ejemplo, *Saccharomyces cerevisiae*, donde los genes cromosómicos pueden interrumpirse fácilmente mediante la transformación celular con un ADN lineal con homología a la secuencia cromosómica a interrumpir (proceso de recombinación mitótica, altamente eficiente en levaduras), la gran mayoría de las bacterias no se pueden transformar con ADN lineal puesto que éste es fácilmente degradado por parte de las exonucleasas citoplasmáticas.

En el año 2000, Datsenko y Wanner describieron un protocolo mediante el cual es posible reemplazar de manera dirigida una secuencia cromosómica por un gen de resistencia a un antibiótico generado mediante una única reacción de PCR (Datsenko y Wanner, 2000).

Siguiendo este protocolo, para obtener dicho producto de PCR, se utilizaron dos oligonucleótidos de aproximadamente 60 nucleótidos de extensión que presentaban homología tanto a las regiones adyacentes al gen cromosómico a ser inactivado como a un plásmido que porta un casete de resistencia a un antibiótico (plásmido pKD4, resistencia a kanamicina, o plásmido pKD3, resistencia a cloranfenicol). Este producto

de PCR (gen de resistencia a antibiótico cuyos extremos 5' y 3' están formados por secuencias nucleotídicas homólogas al lugar del genoma donde se quiere realizar el reemplazamiento alélico) se utilizó para seleccionar las bacterias transformantes (bacterias que han introducido en su genoma este producto de PCR en lugar de su gen original) ya que éstas adquirirán una resistencia nueva (figura 3.1).



**Figura 3.1.** Representación esquemática de la estrategia utilizada en el protocolo de mutagénesis bacteriana descrito por (Datsenko y Wanner, 2000). Dicho protocolo consiste en generar un producto de PCR mediante el uso de oligonucleótidos que presentan regiones homólogas al fragmento cromosómico a reemplazar (H1 y H2, de 40 nucleótidos) y a un plásmido (pKD3 o pKD4), mediante las colas P1 y P2 (con un tamaño de 20 nucleótidos). La amplificación de dicho plásmido mediante PCR generará un producto que porta un casete de resistencia a antibiótico ( $Cm^r$  en el caso del plásmido pKD3 o  $Km^r$  en el caso del plásmido pKD4). La recombinasa Red presente en el plásmido pKD46 catalizará la reacción de recombinación entre el producto de PCR (previamente purificado) con el molde cromosómico a través de las regiones de homología introducidas en las colas H1 y H2. Opcionalmente, el casete de resistencia exógeno introducido en el cromosoma se puede eliminar con la ayuda de la recombinasa FLP, que catalizará la reacción de recombinación entre las repeticiones directas FRT que flanquean los extremos del casete de resistencia amplificado en el producto de PCR. Imagen adaptada de (Baba y col., 2006).

Para realizar dicha amplificación por PCR, se utilizó el set de oligonucleótidos p1/p2 específicos para cada gen a mutagenizar (cuya secuencia está detallada en las tablas 9.1, 9.11 y 9.19 del Anexo de esta memoria), utilizando como ADN molde a amplificar el plásmido pKD4 (o bien el plásmido pKD3) siguiendo el protocolo de PCR descrito en el apartado 3.7.1.

Después de la amplificación del fragmento de PCR, dicho producto fue tratado con la enzima de restricción DpnI (*Thermo Scientific*), con el fin de digerir de manera específica

### 3. MATERIALES Y MÉTODOS

el ADN molde (ADN de los plásmidos pKD3 (o pKD4)) utilizado en la PCR. Ello es posible ya que la enzima de restricción DpnI tiene la capacidad de reconocer y digerir la secuencia G<sup>m6</sup>ATC únicamente cuando ésta se encuentra metilada, pero no es capaz de digerir el ADN de nueva síntesis amplificado *in vitro* ya que éste no se encuentra metilado.

Para obtener los clones mutantes, previamente se tuvo que transformar en la cepa objeto de la mutagénesis el plásmido pKD46. Dicho plásmido, presente en un bajo número de copias, contiene la recombinasa del fago  $\lambda$  *Red Recombinase* bajo el control de un promotor inducible por arabinosa (P<sub>BAD</sub>). Dicha enzima es la proteína responsable de catalizar la reacción de recombinación entre el producto de PCR obtenido y el lugar del cromosoma bacteriano donde se va a dirigir la mutagénesis. Además de permitir la recombinación del fragmento de PCR con su región cromosómica *target*, la recombinasa Red también inhibe la exonucleasa V celular, permitiendo de esta manera la entrada de ADN linear exógeno al citoplasma celular sin ser degradado.

Para realizar el protocolo experimental, se hizo crecer la cepa a mutagenizar en medio SOB a 30°C suplementado con carbenicilina 100 µg/ml (con el objetivo de seleccionar el crecimiento celular únicamente de las células con el plásmido pKD46 y forzar a las células a mantener el plásmido), L-arabinosa 10 mM (para inducir la expresión de la recombinasa bajo el control del promotor P<sub>BAD</sub>) y 20 mM de iones Mg<sup>2+</sup> (10 mM de MgSO<sub>4</sub> y 10 mM MgCl<sub>2</sub>) (el magnesio actúa como el ion necesario para el evento de recombinación homóloga) hasta una D.O<sub>600nm</sub> de 0,6-1. Seguidamente se obtuvieron células electrocompetentes (siguiendo el protocolo descrito en el apartado 3.6.2), ya que para poder introducir ADN exógeno con alta eficiencia dentro de bacterias es necesario inducir el estado de competencia de las mismas. Se electroporaron 60 µl de células electrocompetentes con 1 µg del ADN obtenido al realizar la PCR utilizando como molde el plásmido pKD3/pKD4. Después de someter las células al pulso eléctrico, se resuspendieron en 1 ml de medio SOC y se incubaron a 37°C en agitación constante a 600 rpm para permitir la recombinación del marcador fenotípico en el cromosoma bacteriano y la recuperación de las células después del choque eléctrico. La selección de mutantes se realizó a 37°C en placas de LB-agar suplementadas con kanamicina 50 µg/ml o cloranfenicol 25 µg/ml (para seleccionar aquellos clones en los que el casete de resistencia a antibiótico amplificado por PCR había recombinado con el gen diana de la mutagénesis). A las colonias resistentes a los marcadores transferidos se les realizó una

PCR con oligonucleótidos que hibridaban aguas arriba y aguas abajo al gen mutagenizado para comprobar que realmente el casete marcador de la mutagénesis se había insertado de manera específica en el lugar deseado (serie de oligonucleótidos p1 up-p2 down, cuya secuencia nucleotídica se encuentra indicada en las tablas 9.1, 9.11 y 9.19 del Anexo de esta memoria).

Sin embargo, ya que algunos genes bacterianos se cotranscriben juntos en un operón o en un *cluster* génico, la disrupción por mutagénesis insercional de una pauta abierta de lectura en el operón puede afectar a la expresión de los genes situados aguas abajo del gen objeto de la mutagénesis. Ello nos ofrecería un fenotipo que puede enmascarar el verdadero fenotipo derivado de la mutación en sí (efectos polares). Por tanto, una vez confirmada la mutación de nuestro gen específico, y aprovechando que el gen de resistencia amplificado e introducido en el cromosoma bacteriano se encontraba flanqueado por secuencias FRT, este gen de resistencia pudo ser fácilmente eliminado gracias a la transformación de la cepa mutante con el plásmido pCP20, que aporta resistencia a ampicilina y cloranfenicol y codifica para la recombinasa FLP (figura 3.1). Las células transformadas con dicho plásmido se sembraron en placas de LB-agar suplementadas con carbenicilina a una concentración de 50 µg/ml y se incubaron a 30°C durante toda la noche. Las colonias transformantes se sembraron de nuevo en placas de LB-agar-carbenicilina 50 µg/ml a 30°C, repitiéndose el proceso durante dos días sucesivos. A dicha temperatura se favorece la expresión de la recombinasa FLP plasmídica, que se encarga de catalizar la reacción de recombinación entre las repeticiones directas FRT, eliminando de esta manera el casete marcador de la mutagénesis y dejando una secuencia FRT residual en el cromosoma (los oligonucleótidos p1/p2 fueron diseñados de manera que la escisión del casete de resistencia creara una delección en pauta de lectura en el fragmento cromosómico disrupcionado). Después de realizar una PCR para confirmar la eliminación del gen de resistencia a antibiótico insertado (mediante oligonucleótidos internos al casete de resistencia insertado en el genoma (KT) y oligonucleótidos que hibridaban en la región aguas arriba y aguas debajo de la región mutagenizada (oligonucleótidos p1 up-p2 down), se realizó una resiembra de los clones en placas de LB a 42°C. Con este paso se produjo el curado del plásmido pCP20 ya que éste posee un origen de replicación termosensible por lo que, a lo largo de diferentes pasajes a elevada temperatura, la presencia del plásmido se fue diluyendo y perdiendo entre las diferentes generaciones bacterianas.

#### **3.7.8 Construcción de fusiones de genes cromosómicos y plasmídicos con la secuencia codificante para el epítipo 3XFLAG**

Siguiendo el protocolo de mutagénesis bacteriana descrito anteriormente (con una ligera modificación), fue posible la construcción de fusiones génicas con la secuencia codificante para el epítipo 3XFLAG en el extremo C-terminal del gen específico objeto de marcaje (Uzzau y col., 2001). Dicho protocolo presenta una gran funcionalidad y versatilidad ya que permite inmunodetectar la expresión de una proteína determinada mediante un anticuerpo monoclonal comercial dirigido contra la etiqueta 3XFLAG sin necesidad de generar anticuerpos específicos contra la proteína objeto de estudio.

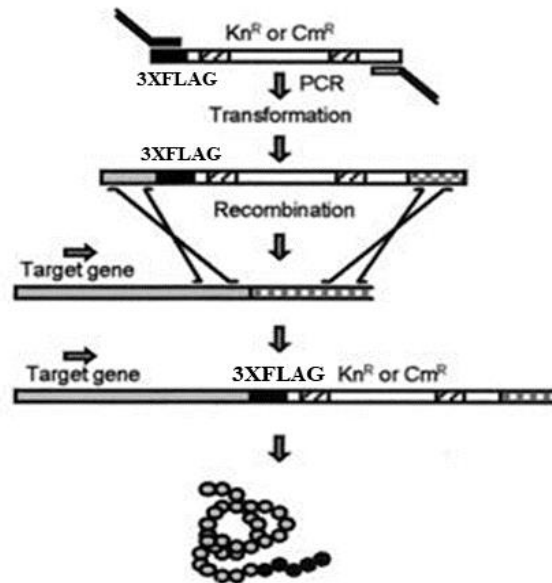
Para realizar la fusión de la etiqueta 3XFLAG en un gen específico se utilizaron las series de oligonucleótidos 3xp1/p2 (cuya secuencia está indicada en las tablas 9.1, 9.11 y 9.19 del apartado Anexo de este trabajo). En el caso del oligonucleótido 3xp1 “*forward*”, el extremo 3’ del oligonucleótido se diseñó de manera que fuera complementario a los 20 primeros nucleótidos de la secuencia codificante del epítipo 3XFLAG presente en el plásmido pSUB11 (“*priming site 1*”), mientras que el extremo 5’ fue diseñado para que fuera complementario a los 40 nucleótidos anteriores al codón *stop* del gen a fusionar con el epítipo 3XFLAG. En el caso del oligonucleótido 3xp2 “*reverse*”, los oligonucleótidos fueron diseñados con la finalidad de que el extremo 3’ fuera complementario al “*priming site 2*” del plásmido pSUB11, mientras que el extremo 5’ del oligonucleótido se diseñó de manera que fuera complementario a los 40 nucleótidos posteriores al codón de finalización de la traducción del gen en el que queríamos insertar el epítipo.

Con estos oligonucleótidos, y utilizando como molde para la PCR el plásmido pSUB11, se obtuvo un fragmento de ADN que contenía un casete de resistencia a kanamicina flanqueado aguas arriba por la región cromosómica situada antes del codón de finalización de la traducción del gen objeto de marcaje seguida por la secuencia codificante para el epítipo 3XFLAG, y aguas abajo por la secuencia nucleotídica complementaria situada tras el codón de parada de la traducción del gen objeto de la mutagénesis (figura 3.2). El protocolo utilizado para realizar dicha PCR se especifica en el apartado 3.7.1.

El reemplazamiento alélico de este fragmento de PCR con el lugar cromosómico objeto de la mutagénesis se realizó tal y como se explica en el apartado 3.7.7. Los mutantes fueron seleccionados en placas de LB-agar suplementadas con kanamicina 50 µg/ml. La



inserción del epítipo 3XFLAG fue confirmada mediante una PCR utilizando oligonucleótidos que hibridaban en una región situada aguas arriba y aguas abajo del gen diana de la mutagénesis (serie de oligonucleótidos p1 up-p2 down).



**Figura 3.2.** Representación esquemática del protocolo de fusión cromosómica de la etiqueta 3XFLAG en el extremo 3' de un gen específico objeto de la mutagénesis (*target gene*). El módulo de ADN que contiene la secuencia codificante del epítipo 3XFLAG (marcado en negro) y un casete de resistencia a antibiótico ( $Km^r$ ) flanqueado por 2 secuencias FRT (marcadas a rayas), es amplificado mediante el uso de oligonucleótidos que presentan secuencias homólogas a la región situada antes del codón *stop* del gen de interés (marcado en gris) y a la región situada aguas abajo de dicho codón (región marcada a cuadros). El producto de PCR generado se integra en el cromosoma de la bacteria de manera análoga a lo indicado en el apartado de mutagénesis bacteriana descrito anteriormente. Imagen adaptada de (Uzzau y col., 2001).

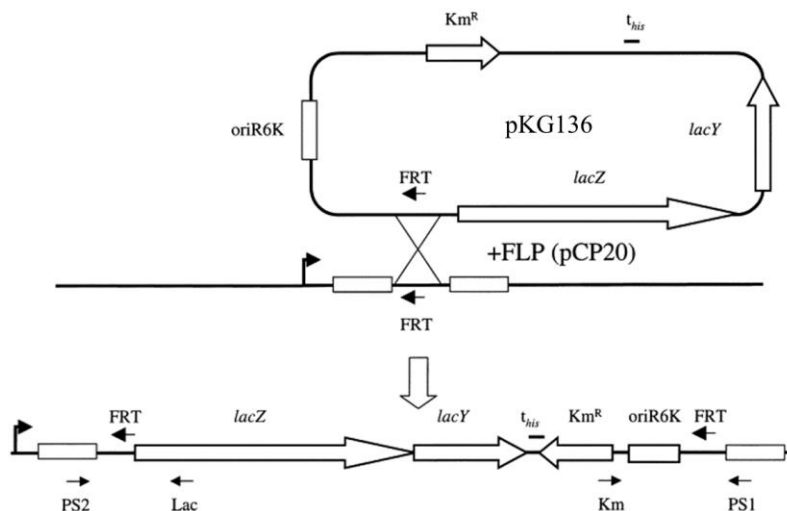
### 3.7.9 Construcción de fusiones de genes cromosómicos con el gen reportero *lacZ* mediante la recombinasa FLP

Utilizando parte del protocolo descrito por (Datsenko y Wanner, 2000), se desarrolló un protocolo que permite generar fusiones cromosómicas con el gen reportero *lacZ* (Ellermeier y col., 2002). Para ello, a los clones que contenían una secuencia FRT residual en el lugar del cromosoma en el que se quería insertar el gen reportero (como consecuencia de la inserción y posterior eliminación del casete de resistencia a antibiótico integrado), se les transformó de nuevo el plásmido pCP20. Las células transformadas con dicho plásmido se crecieron posteriormente en medio SOB a 30°C suplementado con carbenicilina 50  $\mu\text{g/ml}$  y 20 mM de  $\text{Mg}^{2+}$  ( $\text{MgSO}_4$  y  $\text{MgCl}_2$ , a una concentración 10 mM). A dichas bacterias se les electroporó posteriormente el plásmido pKG136, el cual contiene el gen *lacZ* (sin promotor) precedido por una secuencia FRT. Por acción de la

### 3. MATERIALES Y MÉTODOS

recombinasa FLP, codificada en el plásmido pCP20, se produjo la recombinación del gen reportero *lacZ* en el cromosoma a través de las secuencias FRT. Ya que dicho gen *lacZ* porta asociado un gen de resistencia a la kanamicina, y que el plásmido pKG136 es incapaz de replicarse en una célula que no codifique por la proteína  $\lambda$ pir, los clones seleccionados en presencia del antibiótico fueron aquellos que habían integrado el fragmento plasmídico en su cromosoma, generándose de esta manera la fusión del gen *lacZ* bajo el control del promotor del gen mutagenizado previamente (figura 3.3).

Sin embargo, ya que *E. coli* es una bacteria *lac*<sup>+</sup>, para poder realizar estas fusiones cromosómicas previamente se tuvo que mutagenizar el gen *lacZ* de su genoma. Dado que este proceso genera una cicatriz FRT residual extra en el cromosoma, al integrar de nuevo el gen *lacZ* a través del plásmido pKG136, se tuvo que comprobar posteriormente en qué lugar del genoma (secuencia FRT residual) se había insertado el gen reportero.

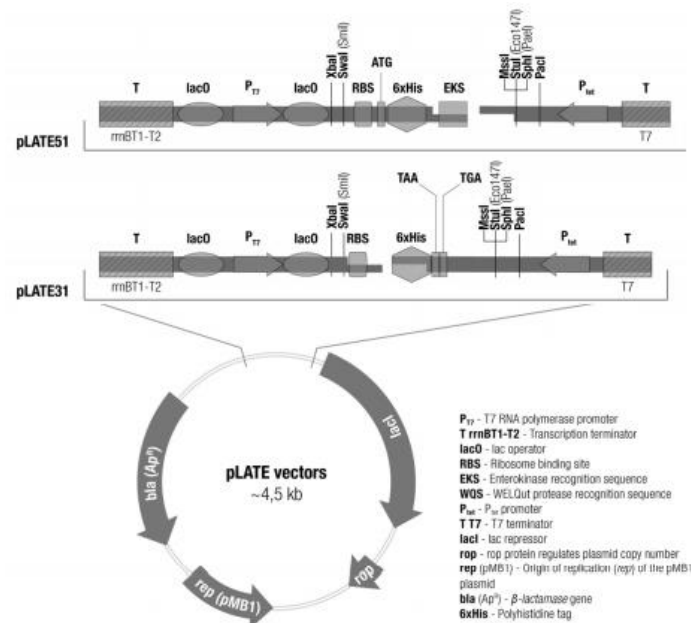


**Figura 3.3.** Representación esquemática del protocolo de fusión del gen reportero *lacZ* en el cromosoma bacteriano. El gen reportero *lacZ* presente en el plásmido pKG136 se integra en la secuencia FRT generada previamente en el cromosoma bacteriano mediante el protocolo de mutagénesis basado en el reemplazamiento alélico. La recombinasa FLP presente en el plásmido pCP20 se encarga de catalizar dicha reacción que permite obtener fusiones cromosómicas con el gen reportero *lacZ*. Imagen adaptada de (Ellermeier y col., 2002).

#### 3.7.10 Clonaje de genes en el vector de sobreexpresión pLATE

Para realizar la sobreexpresión de proteínas a las que se les ha añadido un epítipo de 6 histidinas, se utilizaron distintos plásmidos recombinantes (*pLATE aLICator LIC cloning and expression vector* (Thermo Fisher Scientific)), en los que el gen objeto de estudio se encuentra bajo el control de un promotor dependiente de la ARN polimerasa del fago T7. Con esta finalidad, el gen de interés a clonar se amplificó mediante la polimerasa de

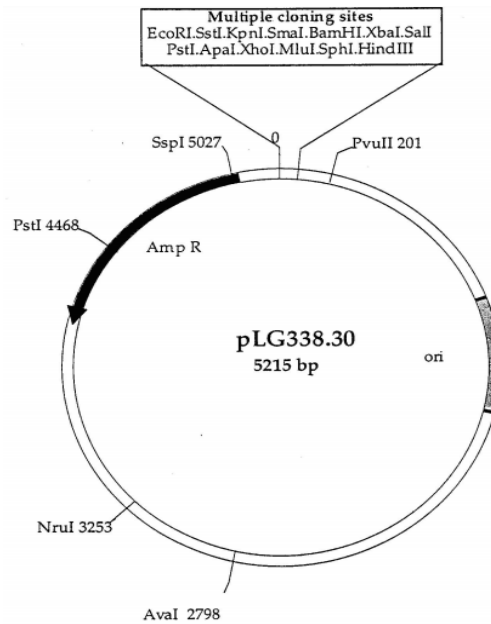
alta fidelidad de copia *Phusion Hot Start II High-Fidelity DNA Polymerase* (Thermo Scientific) siguiendo el protocolo descrito en el apartado 3.7.1. Los fragmentos amplificados por PCR fueron purificados posteriormente mediante el kit *GeneJET PCR Purification Kit* (Thermo Scientific). La ligación de los fragmentos de PCR en los diferentes vectores utilizados (pLATE31 para clonar el gen con la cola de 6 histidinas en el extremo C-terminal y pLATE51 para colocar la cola de 6 histidinas en el extremo N-terminal de la proteína) se realizó siguiendo las instrucciones indicadas por el fabricante (figura 3.4). El fundamento de la ligación se basa en el método de clonaje LIC (*Ligation Independent Cloning*), en el que se utiliza la actividad de exonucleasa 3'– 5' de la ADN polimerasa T4 para generar extremos de 14-21 nucleótidos cohesivos específicos de cadena sencilla tanto en los vectores de clonaje pLATE como en los insertos de ADN específicos (Aslanidis y Jong, 1990). El plásmido y el inserto con los extremos cohesivos se transformaron en células de *E. coli* DH5 $\alpha$ , donde se cerraron los enlaces covalentes entre el plásmido y el inserto para generar la molécula del plásmido cerrado con el gen objeto de interés clonado. A los clones transformantes se les realizó una PCR para comprobar la correcta inserción del gen en el vector. Posteriormente, dichos plásmidos se secuenciaron para corroborar la correcta inserción del gen en el plásmido, tanto a nivel de clonaje en correcta pauta de lectura como la correcta inserción de las colas de 6 histidinas en los extremos amino (N-) y carboxi (-C) terminal de la proteína recombinante.



**Figura 3.4.** Representación esquemática de los vectores pLATE utilizados para clonar y sobreexpresar proteínas de manera inducible con colas de 6 histidinas en los extremos N- y C-terminal de las proteínas recombinantes (plásmidos pLATE 51 y pLATE 31, respectivamente).

#### **3.7.11 Clonaje de genes en el vector pLG338.30 para complementar una mutación determinada en *trans***

Con el fin de complementar una mutación de un gen específico del cromosoma con otro gen aportado en *trans*, se utilizó el plásmido de bajo número de copias pLG338.30 (Cunningham y col., 1993) (figura 3.5). El fragmento específico a clonar (región promotora, región codificante del gen y región terminadora) se amplificó mediante la polimerasa de alta fidelidad de copia *Phusion Hot Start II High-Fidelity DNA Polymerase* (*Thermo Scientific*). A los oligonucleótidos utilizados se les añadió en su extremo 5' las colas correspondientes a las dianas nucleotídicas reconocidas por parte de las enzimas de restricción EcoRI (en el oligonucleótido *forward*) y BamHI (en el oligonucleótido *reverse*). Los fragmentos amplificados por PCR fueron purificados mediante el kit *GeneJET PCR Purification Kit* (*Thermo Scientific*). Seguidamente, tanto el fragmento de PCR obtenido como el plásmido pLG338-30 fueron digeridos con las enzimas de restricción EcoRI/BamHI (*Thermo Scientific*) siguiendo las instrucciones recomendadas por el fabricante. Los productos de digestión volvieron a ser purificados con el mismo kit citado anteriormente. El plásmido digerido y purificado fue tratado con la enzima fosfatasa alcalina *FastAP* (*Fermentas*) con el fin de desfosforilar el extremo 5' del ADN para evitar la re-circularización del mismo. Después de dicha reacción, el fragmento plasmídico volvió a ser purificado. Seguidamente se realizó la reacción de ligación entre el fragmento de PCR y el plásmido, ambos elementos digeridos con enzimas de restricción comunes. Para realizar esta reacción se utilizó la ADN ligasa T4 (*Thermo Scientific*), siguiendo las instrucciones indicadas por el fabricante. El producto de ligación se transformó en células de *E. coli* DH5 $\alpha$ . A los clones transformantes se les realizó una PCR para comprobar la presencia del inserto en el plásmido pLG338.30 haciendo uso de oligonucleótidos que hibridaban aguas arriba y aguas abajo de la región de clonaje múltiple donde se había insertado el fragmento clonado (figura 3.5). A los clones que presentaban el inserto se les realizó un aislamiento plasmídico con el kit *GeneJET Plasmid Miniprep Kit* (*Thermo Scientific*). Dichos plásmidos se secuenciaron posteriormente para corroborar el correcto clonaje del gen objeto de interés en el plásmido.

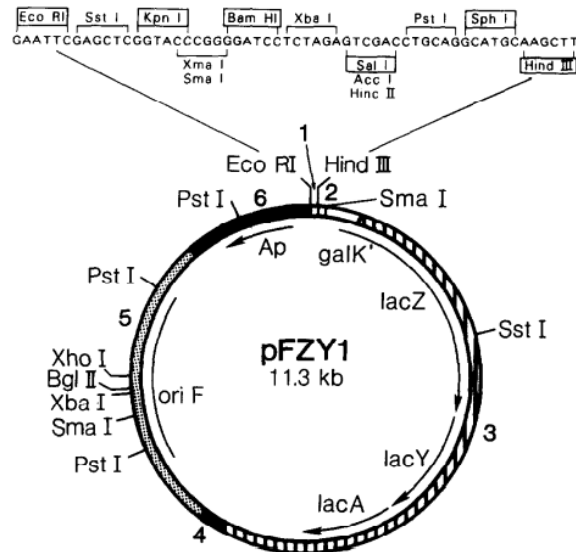


**Figura 3.5.** Mapa del vector de bajo número de copias pLG338.30. Se indica el gen de resistencia a ampicilina (*Amp R*), el origen de replicación del plásmido (*ori*) y el sitio de clonaje múltiple en el que están presentes las dianas nucleotídicas reconocidas por parte de las enzimas de restricción EcoRI y BamHI utilizadas para realizar el clonaje del gen de interés. Imagen adaptada de (Cunningham y col., 1993).

### 3.7.12 Construcción de fusiones transcripcionales con el gen reportero *lacZ* en el plásmido pFZY1

Para generar una fusión transcripcional de un determinado promotor con un gen reportero se utilizó el plásmido pFZY1 (Koop y col., 1987). Este plásmido contiene el gen *lacZ* carente de región promotora precedido por una región de clonaje múltiple (figura 3.6). De manera análoga al apartado anterior, se amplificó la región promotora del gen objeto de estudio mediante oligonucleótidos que contenían en sus extremos 5' las dianas nucleotídicas reconocidas por parte de las enzimas de restricción EcoRI/BamHI. Tanto el producto amplificado como el plásmido fueron digeridos con las mismas enzimas de restricción. A continuación se realizó la reacción de ligación entre ambos fragmentos genómicos, reacción que fue posteriormente transformada a células de *E. coli* DH5 $\alpha$ . Los plásmidos de los clones transformantes se secuenciaron posteriormente para confirmar que el promotor del gen de interés se había clonado en correcta pauta de lectura respecto al gen reportero *lacZ* presente en el plásmido pFZY1.

### 3. MATERIALES Y MÉTODOS



**Figura 3.6.** Mapa del plásmido de bajo número de copias pFZY1 utilizado para obtener fusiones transcripcionales con los promotores de los genes objeto de estudio. Se indica el gen de resistencia a ampicilina (*Ap*), el origen de replicación del plásmido (*ori F*) y el gen *lacZ* sin promotor flanqueado aguas arriba por el sitio de clonaje múltiple en el que están presentes las dianas reconocidas por las enzimas de restricción *EcoRI* y *BamHI*, utilizadas para realizar el clonaje del promotor del gen objeto de estudio. Imagen adaptada de (Koop y col., 1987).

## 3.8 Técnicas experimentales con ARN

### 3.8.1 Aislamiento de ARN

Para realizar el aislamiento del ARN total a partir de los correspondientes cultivos bacterianos, se crecieron las cepas a analizar hasta la fase de crecimiento exponencial ( $D.O_{600nm}$  de 0,4) o el inicio de la fase estacionaria ( $D.O_{600nm}$  de 2,0). Seguidamente, se añadió 1 ml del compuesto *stop solution* (etanol 95% (v/v) y fenol 5% (v/v)) a 5 ml del cultivo bacteriano correspondiente con la finalidad de detener toda actividad metabólica bacteriana. Dicha mezcla se centrifugó a 6.000 rpm durante 10 minutos. Se descartó el sobrenadante y el sedimento celular se congeló a  $-80^{\circ}C$  hasta su posterior uso.

El ARN bacteriano se aisló a partir del *pellet* de los correspondientes cultivos haciendo uso del kit *Tripure Isolation Reagent (Roche)*, siguiendo las instrucciones recomendadas por el fabricante. El fundamento del protocolo utilizado se basa en que este producto es una mezcla de tiocianato de guanidinio (potente agente caotrópico, desnaturizante) y fenol. En su presencia, las células se lisan rápidamente, se solubilizan sus componentes y se inactivan las ribonucleasas. Seguidamente, se añadió cloroformo con la consiguiente generación de dos fases: una fase acuosa (con el ARN) y una fase orgánica (con las

proteínas desnaturalizadas), mientras que el ADN quedaba en la interfase. A continuación, el ARN fue precipitado mediante isopropanol y la posterior adición de etanol al 75% (v/v). Una vez aislada la totalidad del ARN, las potenciales trazas de ADN genómico se digirieron mediante DNAsa I (*Turbo DNA-free, Ambion*). Posteriormente, se purificó el ARN de los restos de enzima y tampón utilizados para degradar las trazas de ADN haciendo uso de una solución de P:C:I (fenol:cloroformo:alcohol isoamílico, a unas proporciones 25:24:1). El ARN purificado se precipitó seguidamente mediante la adición de 3 volúmenes de etanol al 100% y 1/10 parte (del volumen total a precipitar) de acetato sódico 3 M pH 5,2, durante 16 horas a -20°C. Por último, la total integridad del ARN fue comprobada mediante electroforesis en un gel de agarosa al 1,2% (p/v).

### 3.8.2 Real time qRT-PCR

Con el fin de cuantificar los niveles de ARNm de los diferentes genes de interés se utilizó la técnica de qRT-PCR. Una vez aislado el ARN bacteriano total, se midió la concentración y la calidad del mismo haciendo uso del espectrofotómetro *Nano-Drop 1000 (Thermo Fisher Scientific)*. Seguidamente, se retrotranscribió 1 µg del ARN obtenido mediante el kit *High-capacity cDNA Reverse Transcription kit (Applied Biosystems)*, siguiendo las instrucciones indicadas por el fabricante. Este kit contiene dNTPs, la transcriptasa reversa *MultiScribe*, MgCl<sub>2</sub> y oligonucleótidos aleatorios que permiten retrotranscribir todo el ARN de la muestra a ADN complementario (ADNc). El programa utilizado para generar el ADNc a partir de la muestra de ARN fue: 10 minutos a 25°C, 120 minutos a 37°C y 5 minutos finales a 85°C con el fin de inactivar la enzima retrotranscriptasa.

La PCR a tiempo real se realizó utilizando el fluorocromo *SYBR green PCR master mix 2x (Life Technologies)* junto con el ADNc obtenido previamente, haciendo uso del detector de fluorescencia *ABI Prism 7700 sequence detection system (Applied Biosystems)*. Cada reacción se preparó de la siguiente manera:

- 10 µl de 2x *SYBR Green PCR Master Mix (Life Technologies)* (mezcla que contiene el colorante *SYBR green*, la ADN polimerasa *AmpliTaq Gold*, dNTPs y los tampones correspondientes).
- 1 µl de oligonucleótido *forward* a una concentración de 10 µM.
- 1 µl de oligonucleótido *reverse* a una concentración de 10 µM.
- 8 µl del ADNc diluido 1:100.



### 3. MATERIALES Y MÉTODOS

Para realizar la qPCR a tiempo real se utilizó un programa que consistió en una etapa inicial de 10 minutos a 95°C, seguida por 40 ciclos consecutivos iniciados mediante una etapa de desnaturalización a 95°C durante 15 segundos y una etapa de hibridación y extensión a 60°C durante 1 minuto. Las reacciones se llevaron a cabo en placas de 96 pocillos *MicroAmp* cubiertas con *film* óptico adhesivo (*Applied Biosystems*).

Los oligonucleótidos específicos dirigidos a los genes de interés a analizar se diseñaron mediante la herramienta electrónica *Primer Blast* de la base de datos del NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Estos oligonucleótidos se diseñaron de manera que cumplieran las siguientes premisas: que tuvieran una longitud de 20-21 nucleótidos, que presentaran un contenido de GC igual o mayor al 50%, que no formaran estructuras secundarias, que mostraran una temperatura de hibridación de entre 60-62°C y que permitieran amplificar un fragmento específico que tuviera un tamaño de entre 80 y 150 pares de bases.

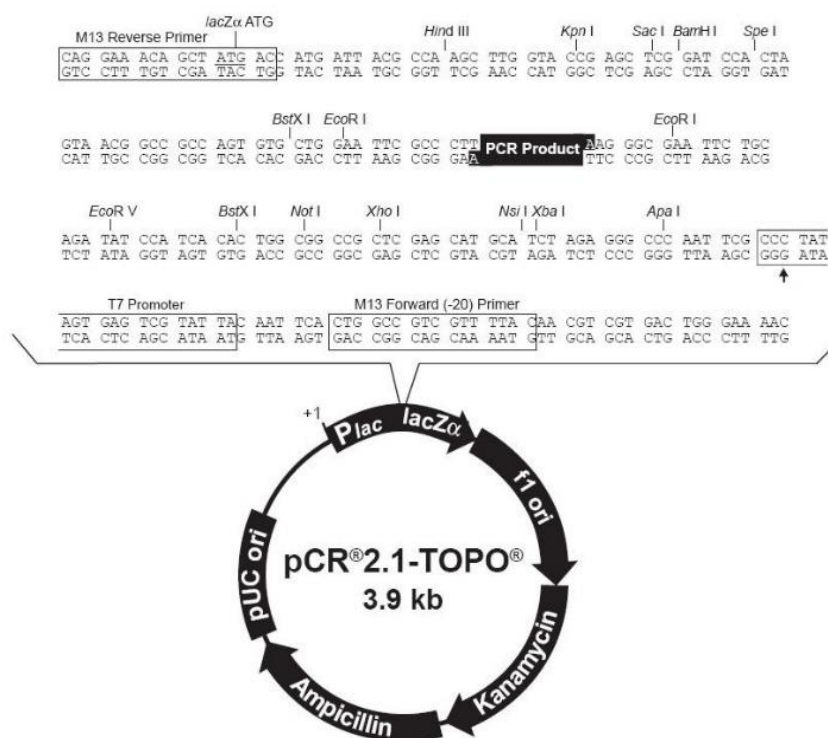
La cuantificación relativa de la expresión de los genes objeto de estudio en los distintos mutantes utilizados respecto a la cepa salvaje se realizó mediante el método Ct comparativo haciendo uso del *software StepOne v2.3* (*Applied Biosystems*). La cantidad relativa del ADNc analizado fue normalizada utilizando el gen *gapA* como gen interno referencia. Dicho gen codifica para la enzima gliceraldehido-3-fosfato deshidrogenasa (GAPDH) y muestra una expresión constante para todas las condiciones analizadas. Cada muestra se analizó por triplicado y se analizó con detalle la *melting curve*. Un solo pico en esta curva es sinónimo de la amplificación específica de un único amplicón; la presencia de varios picos es un elemento indicativo de amplificaciones inespecíficas por parte de los oligonucleótidos utilizados. Los resultados obtenidos se representaron como la expresión relativa del gen de interés en el fondo genético mutante en las condiciones analizadas respecto a la expresión del mismo gen en la cepa salvaje bajo las mismas condiciones de cultivo.

#### **3.8.3 5'RACE (*rapid amplification of cDNA ends*)**

Con el fin de identificar el primer nucleótido (+1) correspondiente al inicio de transcripción de un determinado ARN mensajero, se realizaron experimentos de 5'RACE siguiendo el protocolo descrito por (Wagner y Vogel, 2005) con ligeras modificaciones. Para ello, a partir de 12 µgr de ARN total del cultivo bacteriano correspondiente aislado siguiendo el protocolo descrito en el apartado 3.8.1, se eliminó el extremo 5' trifosfato de

los ácidos nucleicos mediante el uso de 25 unidades de la pirofosfatasa *tobacco acid* (*TAP*, *Epicentre*). Dicha reacción se realizó en un volumen final de 50  $\mu$ l en un tampón formado por 50 mM de acetato sódico pH 6, 10 mM de EDTA, 1%  $\beta$ -mercaptoetanol y 0,1% de Tritón X-100, además de 0,5  $\mu$ l de inhibidor de RNAasas (*Ambion*), y se dejó incubar durante 60 minutos a 37°C. La reacción se detuvo mediante extracción con P:C:I (25:24:1), y los ácidos nucleicos se precipitaron posteriormente mediante la adición de etanol absoluto y acetato sódico 0,3 M pH 5,7. El ARN precipitado se disolvió en agua y seguidamente se le añadió 500 pmoles del adaptador de 5' ARN A3 (5' GAU AUG GCG GAA UUC CUG UAG AAC GAA CAC UAG AAG AAA 3'), desnaturalizado previamente durante 5 minutos a 95°C y posteriormente enfriado en hielo rápidamente. El adaptador A3 fue ligado a los extremos 5' del ARN desfosforilado durante 12 horas a 17°C haciendo uso de 50 unidades de ADN ligasa T4 en un tampón cuya composición era de 50 mM de Tris-HCl pH 7,9, 10 mM de MgCl<sub>2</sub>, 4 mM de DTT, 150  $\mu$ M de ATP y 10% de DMSO. De manera análoga al paso anterior, la reacción se detuvo mediante extracción con P:C:I (25:24:1), y los ácidos nucleicos se precipitaron posteriormente mediante etanol al 100% (v/v) y acetato sódico 3 M a pH 5,2 durante 16 horas a -20°C. Seguidamente, se retrotranscribieron 5  $\mu$ gr del ARN a ADNc haciendo uso del kit *High-capacity cDNA Reverse Transcription kit* (*Applied Biosystems*), siguiendo las instrucciones recomendadas por el fabricante. La retrotranscripción se realizó mediante tres etapas de 10 minutos a 25°C, 120 minutos a 37°C y 5 minutos a 85°C. El producto resultante de la retrotranscripción fue amplificado por PCR mediante un oligonucleótido *forward* (correspondiente a la región complementaria al adaptador A3 unido al extremo 5' del ARNm) y un oligonucleótido *reverse* (que hibrida en la región codificante del gen objeto de estudio) mediante el uso de la polimerasa *DreamTaq Hot Start Green DNA Polymerase* (*Thermo Scientific*), siguiendo las instrucciones recomendadas por el fabricante. Dicho producto de PCR se analizó electroforéticamente en un gel de agarosa al 2% (p/v), donde los amplicones de banda única se purificaron mediante el kit *GeneJET PCR Purification Kit* (*Thermo Scientific*). Posteriormente, el producto de PCR se precipitó y se clonó en el vector pCR2.1 TOPO-vector (*Invitrogen*), que se transformó seguidamente a bacterias competentes comerciales NEB 5 alpha (*New England Biolabs*). Los clones resultantes de la transformación (colonias blancas en placas de LB-agar suplementadas con XGal 40  $\mu$ gr/ml) fueron objeto de una PCR mediante los oligonucleótidos M13 *forward*-M13 *reverse* para confirmar la presencia del inserto clonado (figura 3.7).

### 3. MATERIALES Y MÉTODOS



**Figura 3.7.** Representación esquemática del vector pCR2.1-TOPO (*Invitrogen*) utilizado para clonar el ADNc correspondiente al gen objeto de estudio para determinar su inicio de transcripción mediante el experimento de 5'RACE. Se indica el lugar donde se ha clonado el inserto rompiendo de esta manera la pauta abierta de lectura del gen *lacZα* (dando lugar a colonias blancas en placas de LB-agar XGal 40 μgr/ml), los genes de resistencia a kanamicina y ampicilina y los orígenes de replicación pUC y f1 que contiene el vector. También se indica el lugar de hibridación de los oligonucleótidos M13 *forward*-M13 *reverse* utilizados para comprobar mediante PCR la presencia del inserto y realizar su posterior secuenciación.

De los clones positivos se aisló el plásmido correspondiente mediante *Miniprep* utilizando el kit comercial *GeneJET Plasmid Miniprep Kit* (*Thermo Scientific*). Dichos plásmidos fueron secuenciados, identificándose el inicio de transcripción (+1) del gen objeto de estudio como el primer nucleótido situado después de la secuencia del adaptador de ARN ligado al extremo 5' del ARN mensajero.

#### 3.8.4 Northern Blot

Para poder determinar el nivel de expresión de un determinado ARN mensajero, para analizar posibles patrones de degradación del mismo y para tratar de identificar la presencia de algún ARN antisentido contra el ARN mensajero objeto de estudio, se realizaron ensayos de *Northern Blot*. Para ello, el ARN aislado a partir del protocolo 3.8.1, se mezcló con tampón de carga para ARN 2x (*Bromophenol Blue* 2% (p/v), *Xylene Cyanol* 2% (p/v), 0,5 M EDTA pH 8,0, SDS 20% (p/v) y formamida 0,1% (v/v)). La mezcla se hirvió durante 5 minutos a 95°C para deshacer las estructuras secundarias del

ARN y se incubó rápidamente a 4°C para evitar que se formaran de nuevo estas estructuras. Posteriormente se realizó la electroforesis del ARN en un gel de poliacrilamida en condiciones desnaturalizantes en tampón TBE *RNAse free* comercial 1x (*Sigma Aldrich*), utilizando la cubeta de electroforesis *Protean II xi Cell* (*Bio-Rad*). La composición del gel de poliacrilamida utilizado fue de 30 ml de poliacrilamida al 6% (PAA (19:1), *Bio-Rad*) más urea 8,3 M (*Bio-Rad*), 300 µl de APS 10% (*Bio-Rad*) y 30 µl de TEMED (*Sigma Aldrich*). Paralelamente a las muestras de ARN, también se cargaron en el gel los marcadores de peso molecular pUC19 y ΦX174 (*Thermo Scientific*), previamente marcados radioactivamente con  $\gamma$ -<sup>32</sup>P. La electroforesis se realizó durante 2 horas y 30 minutos a 300 V. Pasado este tiempo, se realizó la electrotransferencia del ARN del gel de poliacrilamida a una membrana de Nylon (*Amersham Hybond XL, GE Healthcare*) haciendo uso del aparato *Trans Blot Cell* (*Bio-Rad*) durante 16 horas en tampón TBE 1x a 4°C. Una vez realizada la electrotransferencia, la membrana (con el ARN transferido) se incubó durante 1 hora a 80°C (*baking*) con el fin de fijar los ácidos nucleicos a la misma. Pasado este tiempo, se lavó la membrana con agua *milli-Q* y se incubó durante 1 hora adicional con la solución de hibridación (*ROTI®Hybri-Quick, Carl Roth*) en tubos de hibridación en agitación constante de 200 rpm a 42°C. Seguidamente, se añadieron 25 µl de la sonda marcada radioactivamente en su extremo 5' con el isótopo radioactivo  $\gamma$ -<sup>32</sup>P, cuya reacción de marcaje (y la de los marcadores de peso molecular utilizados previamente en la electroforesis) se especifica a continuación:

- 2 µl de la sonda (oligonucleótido específico a hibridar con el gen de interés) a una concentración de 10 µM / 2 µl del marcador de peso molecular pUC19/ΦX174 (*Thermo Scientific*)
- 2 µl de tampón *Reaction Buffer A* 10x (*Thermo Scientific*)
- 1 µl de enzima PNK (*Polynucleotide-kinase, Thermo Scientific*)
- 12 µl de H<sub>2</sub>O *RNAase free* (*Thermo Scientific*)
- 3 µl de ATP- $[\gamma$ <sup>32</sup>P] (10 µCi/µl) (*Perkin Elmer*)

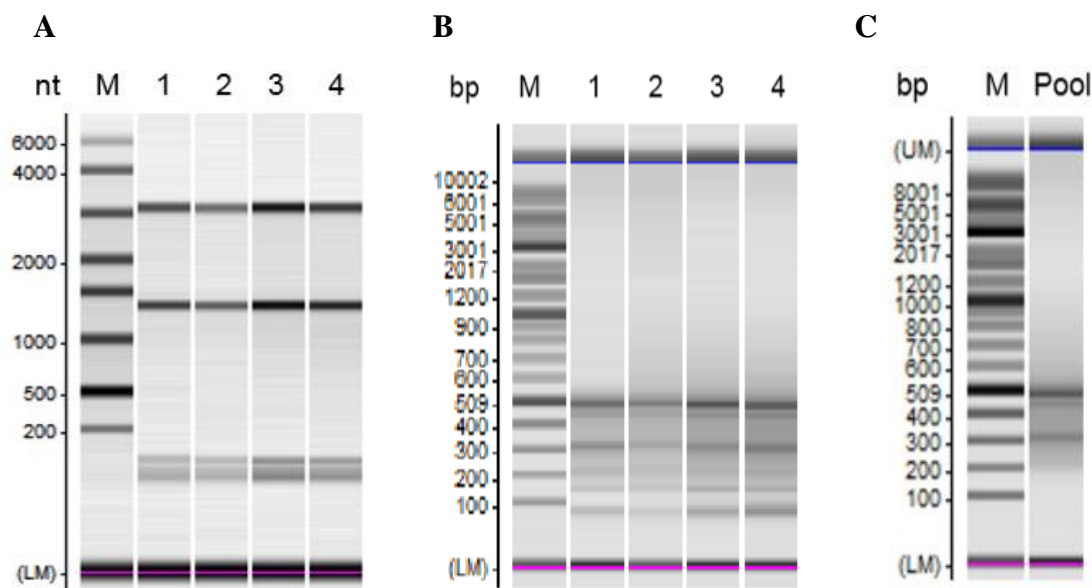
La reacción de marcaje se realizó durante 1 hora a 37°C, más 5 minutos posteriores de inactivación de la enzima PNK a 95°C. Seguidamente, se purificaron las sondas marcadas radioactivamente del resto de ATP- $[\gamma$ <sup>32</sup>P] no unido a los oligonucleótidos mediante las columnas *Sphadex G-25* (*GE Healthcare*), siguiendo las instrucciones indicadas por el fabricante.

### 3. MATERIALES Y MÉTODOS

La membrana junto con la sonda marcada radioactivamente con  $\gamma$ -<sup>32</sup>P se dejó incubando durante 16 horas a 42°C en agitación constante. Pasado este tiempo, se descartó el tampón de hibridación junto con el resto de los oligonucleótidos que no hibridaron con el ARN presente en la membrana, y se realizaron tres lavados (de 20 minutos cada uno) de la membrana con diluciones decrecientes del tampón de lavado SSC 20x (*Saline-Sodium Citrate, Sigma-Aldrich*), cuya composición es 3 M de NaCl, 0,3 M de citrato de sodio (pH 7,0) y 0,1% SDS (los lavados se hicieron con las diluciones 5x, 1x y 0,1x de las sales SSC descritas anteriormente). Finalmente, se expuso la membrana de *Nylon* radioactiva sobre una pantalla (*Imaging Screen K, Bio-Rad*) durante 16 horas dentro de un casete (*Four Square Casette, Spectroline*). Las imágenes de los filtros radiactivos se digitalizaron mediante el sistema *FLA-5100 imaging system (Fujifilm)*.

#### 3.8.5 RNA-seq

Para analizar el efecto de una determinada mutación sobre el perfil global de expresión génica de las bacterias se utilizó la técnica de RNA-seq. El protocolo se realizó en colaboración con la empresa *Vertis Biotechnologie AG (Freising-Weihenstephan, Alemania)*. El ARN de los correspondientes cultivos bacterianos fue aislado haciendo uso del kit *RNA isolation kit (Ambion)*. Después de comprobar la calidad y pureza del ARN mediante electroforesis, se eliminaron las potenciales trazas de moléculas de ARNr de la muestra utilizando el kit *Ribo-Zero rRNA removal kit for bacteria (Illumina)*. La muestra libre de ARNr se sometió a ultrasonidos (2 pulsos de 30 segundos a 4°C). Después de la fragmentación, se ligaron adaptadores a los extremos 3' de las moléculas de ARN. Seguidamente, se sintetizó la primera cadena de ADNc utilizando la transcriptasa reversa M-MLV. Este ADNc fue purificado, y posteriormente se le unió el adaptador 5' *Illumina TruSeq sequencing adapter*. Dicho ADNc fue amplificado por PCR hasta una concentración de 10-20 ng/μl utilizando una ADN polimerasa de alta fidelidad de copia y 15 ciclos de amplificación. El ADNc amplificado por PCR fue fraccionado en fragmentos de 250-500 pb haciendo uso del kit *Agencourt AMPure XP kit (Beckman Coulter Genomics)* (figura 3.8). El *pool* de ADNc fue finalmente secuenciado mediante el sistema *Illumina NextSeq 500* utilizando un tamaño de lectura de 75 pb.



**Figura 3.8.** Análisis por electroforesis de la integridad del ARN aislado de los cultivos bacterianos (A), del ADNc después de ser amplificado por PCR (B), y del *pool* de ADNc utilizado en el experimento de secuenciación masiva mediante *Illumina NextSeq 500* (C). *Vertis Biotechnologie AG*.

### 3.9 Técnicas experimentales con proteínas

#### 3.9.1 Obtención de extractos proteicos totales

Con el fin de poder analizar el contenido total de proteína intracelular bacteriana, se obtuvieron extractos proteicos totales de los cultivos crecidos bajo las condiciones específicas a analizar. Para obtener el extracto proteico total de cada cultivo, se centrifugó 1 ml del mismo a 13.400 rpm, se descartó el sobrenadante y se resuspendió el sedimento celular en tampón de carga de proteínas (glicerol 5% (v/v),  $\beta$ -mercaptoetanol 2,5% (v/v), SDS 1,15% (p/v), Tris-HCl 31 mM a pH 6,6 y azul de bromofenol 0,05% (p/v)) en función del nivel de crecimiento del cultivo. Por ello, para poder normalizar la concentración de proteína obtenida en función de la biomasa del cultivo utilizado, se añadió tampón de carga de proteínas en función de la siguiente relación:

$$\frac{D. 0600nm \text{ cultivo}}{4} = \text{volumen de tampón de carga de proteínas (ml)}$$

Una vez resuspendido el sedimento celular en el tampón de carga de proteínas, se hirvió la muestra a 100°C durante 10 minutos para facilitar la lisis celular y la desnaturalización de las proteínas del extracto. Las proteínas de dichos extractos se analizaron posteriormente mediante electroforesis en geles de poliacrilamida en presencia de SDS como agente desnaturante de las proteínas.



#### **3.9.2 Análisis del secretoma bacteriano**

Para poder estudiar el patrón de las proteínas secretadas por las células bacterianas, se precipitaron dichas proteínas secretadas y liberadas por las bacterias al medio de cultivo y se analizaron posteriormente mediante electroforesis en gel de acrilamida. Con este fin, se crecieron los correspondientes cultivos bacterianos hasta una  $D.O_{600nm}$  de 2,0 y se centrifugaron 10 ml de los mismos a 6.000 rpm durante 10 minutos a temperatura ambiente. El sobrenadante obtenido se purificó mediante un filtro de 0,22  $\mu m$  (*Millex-GP*), con el fin de eliminar los restos celulares no sedimentados durante el proceso de centrifugación. A continuación, las proteínas fueron precipitadas mediante la adición de ácido tricloroacético (TCA) al 45% (p/v), y la posterior incubación durante 1 hora en hielo. Seguidamente, se centrifugaron las muestras durante 30 minutos a 13.400 rpm a 4°C. Se descartó el sobrenadante, se añadió 0,5 ml de acetona fría y se volvió a centrifugar la muestra en las mismas condiciones que las anteriores. Finalmente, se retiró la acetona y se dejó secar el sedimento proteico durante 10 minutos. Por último, se resuspendió el sedimento con tampón de carga de proteínas 1x. Debido a la presencia de restos de ácido tricloroacético utilizado para precipitar las proteínas, al añadir el tampón de carga, el indicador de pH azul de bromofenol viró de color azul a amarillo. Para neutralizar el pH ácido, se añadió Tris-Base saturado (1 M) hasta que el tampón de carga volvió al color azul inicial. Las proteínas de dichos extractos se analizaron posteriormente mediante electroforesis en geles de poliacrilamida en condiciones desnaturalizantes.

#### **3.9.3 Fraccionamiento celular**

Con el fin de obtener el perfil proteico correspondiente a cada una de las diferentes fracciones de la célula (citoplasma, periplasma, membrana interna y membrana externa), se realizó el protocolo de fraccionamiento celular. Por ello, a partir 2 ml del cultivo correspondiente a analizar crecido en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento, se centrifugaron las células a 13.400 rpm y se realizó un lavado con 1 ml de Tris-HCl 20 mM pH 7,4. El *pellet* celular se resuspendió en tampón de lisis (Tris-HCl 50 mM pH 7,4, NaCl 150 mM, EDTA 1 mM y Tritón X-100 1% (v/v)). La mezcla se incubó durante 10 minutos a temperatura ambiente y posteriormente se centrifugó durante 2 minutos a 13.400 rpm. El sedimento celular se resuspendió en 2 ml de H<sub>2</sub>O fría y se incubó durante 10 minutos a 4°C. Seguidamente, se centrifugó la muestra a 12.000 rpm durante 5 minutos a 4°C. El sobrenadante obtenido en dicha centrifugación



correspondía a la fracción de proteínas del periplasma. El sedimento celular se resuspendió en 2 ml de Tris-HCl 20 mM pH 7,4. Dicha suspensión se lisó con ultrasonidos en el sonicador *Labsonic P (B. Braun Biotech International)* mediante dos ciclos seguidos de 30 segundos y 50% de amplitud. A continuación, se centrifugó la muestra a 5.000 rpm durante 10 minutos a 4°C. El sobrenadante obtenido se volvió a centrifugar a 10.000 rpm durante 30 minutos adicionales a 4°C. El sobrenadante obtenido en esta última centrifugación correspondía a la fracción de proteínas presentes en el citoplasma bacteriano. El sedimento, correspondiente a la fracción de las membranas bacterianas, se resuspendió en 2 ml de H<sub>2</sub>O a 4°C, a los que posteriormente se añadió 200 µl del detergente *lauroyl sarcosine* 20% (p/v). Seguidamente, se centrifugó la muestra a 15.000 rpm durante 30 minutos a 4°C. El *pellet*, correspondiente a la fracción de membrana externa, se resuspendió en 40 µl de tampón de carga de proteínas. El sobrenadante contenía la fracción proteica correspondiente a la membrana interna. Las tres fracciones solubles de los sobrenadantes correspondientes a las fracciones proteicas de periplasma, citoplasma y membrana interna se precipitaron de manera análoga al apartado anterior (3.9.2). Las fracciones obtenidas se guardaron a -20°C para ser analizadas por electroforesis en geles de poliacrilamida en condiciones desnaturalizantes.

#### **3.9.4 Electroforesis de proteínas en geles de poliacrilamida en presencia de SDS**

La electroforesis de proteínas en geles de poliacrilamida se ha realizado con 4 objetivos principales: i) analizar cualitativamente las proteínas que forman parte de los diferentes extractos analizados; ii) comprobar que la cantidad total de proteína presente en los extractos celulares era equivalente; iii) realizar la inmunodetección de proteínas transferidas a una membrana mediante anticuerpos específicos; iv) llevar a cabo la identificación de proteínas mediante LC-MS/MS.

Debido a que la gran mayoría de las proteínas analizadas en este trabajo presentaban un peso molecular relativamente bajo, para analizar dichas proteínas se realizaron las electroforesis en geles desnaturalizantes de poliacrilamida en Tris-Tricina-SDS. Este tipo de geles representan una mejor herramienta que los geles bifásicos en Tris-Glicina-SDS para separar con una mayor resolución proteínas cuyo peso molecular se encuentra comprendido entre 5 y 50 kDa (Schägger y von Jagow, 1987). De esta manera, los geles utilizados estaban compuestos por 3 fases: fase de compactación, fase espaciadora y fase de resolución. La primera y segunda fase ocupaban una cuarta parte del gel cada una,

### 3. MATERIALES Y MÉTODOS

mientras que la última fase (la fase de resolución) ocupaba la segunda mitad del gel. La composición de este tipo de geles queda detallada en la tabla 3.4.

Componentes del gel	Fase de resolución 16,5%T 3%C	Fase espaciadora 10%T 3%C	Fase de compactación 4%T 3%C
Acrilamida/Bis 40%T, 3,3%C ( <i>Sigma Aldrich</i> )	1,2 ml	0,36 ml	0,18 ml
Tampón de gel (Tris-HCl 3 M y SDS 0,3% (p/v) pH 8,45)	1 ml	0,5 ml	0,45 ml
Glicerol 50% (v/v)	0,8 ml	-	-
Agua bidestilada	-	0,64 ml	1,23 ml
APS 10% (p/v) ( <i>Bio-Rad</i> )	20 µl	10 µl	15 µl
TEMED ( <i>Sigma Aldrich</i> )	4 µl	2 µl	3 µl

%T: Porcentaje de acrilamida + bisacrilamida total

%C: Porcentaje de bisacrilamida respecto a T

**Tabla 3.4.** Composición de las diferentes fases (resolución, espaciadora y compactación) de los geles de poli(acrilamida) trifásicos utilizados para resolver y analizar electroforéticamente los diferentes extractos proteicos bacterianos aislados durante este trabajo.

Para realizar la electroforesis se utilizaron dos tampones distintos: tampón Ánodo (Tris-HCl 0,2 M pH 8,8) y tampón Cátodo (Tris-HCl 0,25 M, Tricina 0,1 M y SDS 0,4% (p/v), a un pH final de 8,25). Paralelamente a las muestras proteicas a analizar, también se cargó en el gel un marcador de peso molecular que permite estimar el tamaño de las proteínas presentes en las muestras. El marcador utilizado fue *Page Ruler Prestained Protein Ladder* (*Thermo Scientific*), que presenta 10 bandas preteñidas correspondientes a diferentes proteínas con pesos moleculares correspondidos entre 10 y 180 kDa.

La electroforesis se inició a 50 V de potencia hasta que todas las muestras se igualaron en la fase de compactación. A continuación, se subió el voltaje hasta 120 V y se realizó la electroforesis durante 2 horas adicionales, hasta que el marcador del frente de electroforesis presente en el tampón de carga de proteínas llegó al final del gel.

#### 3.9.5 Detección de proteínas mediante tinción con Azul de Coomassie

Para realizar la visualización de las proteínas de manera rutinaria, los geles de poli(acrilamida) resueltos electroforéticamente se tiñeron con Azul de Coomassie. La tinción de proteínas se realizó incubando los geles durante 30 minutos sumergidos en la solución de tinción en agitación suave y a temperatura ambiente. La composición de dicha solución de tinción es Azul brillante de Coomassie R-250 (*Bio-Rad*), ácido acético glacial 10% (v/v) (*Bio-Rad*) y metanol 20% (v/v) (*Bio-Rad*). Seguidamente, se realizaron

diferentes lavados del gel con ácido acético glacial al 10% (v/v) para eliminar el exceso de colorante no unido a las proteínas. Una vez teñidos los geles, se tomaron las imágenes correspondientes utilizando el equipo *Molecular Imager® Gel Doc™ XR+* (Bio-Rad) y se digitalizaron con el *software ImageLab* (Bio-Rad).

### **3.9.6 Detección de proteínas mediante tinción de plata**

En los casos en los que la tinción con Azul de Coomassie no fuera lo suficientemente sensible para visualizar las bandas proteicas, los geles de poliacrilamida fueron teñidos con nitrato de plata, una tinción que presenta una mayor sensibilidad. Esta tinción se llevó a cabo a temperatura ambiente y en agitación suave. En primer lugar, se introdujo el gel en una solución de fijación (metanol 40%) durante 10 minutos. Transcurrido ese tiempo, se realizaron dos lavados con agua bidestilada de 5 minutos cada uno, y a continuación se introdujo el gel en una solución de tiosulfato de sodio al 0,02% (p/v) durante 1 minuto. Seguidamente, se realizaron dos lavados con agua bidestilada de 20 segundos cada uno y posteriormente se incubó el gel durante 10 minutos en una solución de nitrato de plata al 0,1% (p/v). Una vez finalizado este tiempo, se introdujo el gel en la solución de tinción compuesta por 7,5 gr de carbonato de sodio y 5 ml de tiosulfato de sodio al 0,02% (p/v), ambos disueltos en 250 ml de agua bidestilada, a los que se añadió 12,5 µl de formaldehído al 37% (v/v). El gel se mantuvo en esta solución hasta que aparecieron las diferentes bandas proteicas teñidas. La tinción se detuvo añadiendo 1 ml de ácido cítrico al 2% (p/v). Finalmente, se realizó un lavado durante 10 minutos con agua bidestilada y se tomaron las fotografías correspondientes con el equipo *Molecular Imager® Gel Doc™ XR+* (Bio-Rad).

### **3.9.7 Inmunodetección de proteínas específicas por Western Blot**

Con el fin de poder analizar el nivel de expresión de una proteína específica, las proteínas separadas electroforéticamente en gel de poliacrilamida se transfirieron a un soporte sólido (membrana de PVDF, fluoruro de polivinilideno), para posteriormente poder ser inmunodetectadas mediante un anticuerpo específico dirigido contra dicha proteína.

La membrana de PVDF utilizada (*Immun-Blot™ PVDF Membrane for Protein Blotting 0,2 µm*, Bio-Rad) es altamente hidrofóbica, por lo que antes de realizar la transferencia de proteínas del gel de poliacrilamida a la membrana, ésta se sumergió durante 1 minuto en metanol absoluto. De esta manera, mediante el tratamiento con el disolvente orgánico se consigue eliminar la hidrofobicidad de la membrana por lo que se incrementa la capacidad

### 3. MATERIALES Y MÉTODOS

de unión de las proteínas a la misma. A continuación, tanto el gel como la membrana se equilibraron en tampón de transferencia (48 mM de Tris-HCl, 39 mM de glicina, 20% de metanol (v/v), 1,3 mM de SDS) durante 5 minutos. La transferencia de las proteínas del gel de poliacrilamida a la membrana de PVDF se realizó utilizando el sistema de transferencia semi-seco mediante el equipo *Trans-Blot-SD-Semy-Dry Transfer Cell* (Bio-Rad). Para ello, el gel de poliacrilamida se colocó sobre la membrana y el conjunto se colocó entre papeles de filtro (*Whatman 3 MM*), tres por cada lado. La transferencia se realizó a 15 V durante 45 minutos. Una vez transcurrido este tiempo, la membrana se incubó con el tampón de bloqueo (PBS-Tritón X-100 0,2% (v/v) (*Sigma Aldrich*) + leche descremada en polvo 5% (p/v) (*Scharlau*)) durante 1 hora, con el fin de eliminar posibles interacciones inespecíficas del anticuerpo. A continuación, se descartó la solución de bloqueo, y se incubó la membrana con el anticuerpo primario (diluido en PBS-Tritón X-100 0,2% (v/v) + leche descremada en polvo 3% (p/v)) dirigido contra la proteína a detectar en agitación constante durante 16/18 horas a 4°C. Posteriormente, se eliminó el exceso de anticuerpo no unido a la membrana mediante tres lavados sucesivos de 10 minutos con PBS-Tritón X-100 0,2% (v/v) a temperatura ambiente. Seguidamente se incubó la membrana durante 45 minutos a temperatura ambiente con el anticuerpo secundario (dirigido contra el anticuerpo primario) conjugado con la peroxidasa HRP (*horseradish peroxidase, Promega*) (dicho anticuerpo fue previamente diluido 1:2500 en PBS-Tritón X-100 0,2% (v/v)). Por último, se eliminó el exceso de anticuerpo secundario a través de tres lavados con PBS-Tritón X-100 0,2% (v/v) de 30 minutos cada uno a temperatura ambiente. Finalmente, para realizar el revelado quimioluminiscente de la membrana, se incubó la misma durante 5 minutos con la solución de revelado compuesta por luminol y peróxido de hidrógeno (a concentraciones equimolares), siguiendo las indicaciones recomendadas por el fabricante (*Amershan ECL Plus, Western Blotting Detection System, GE Healthcare*). De esta manera, la peroxidasa conjugada al anticuerpo secundario catalizará la reacción de oxidación del luminol dando lugar a la aparición de la quimioluminiscencia. La detección de las bandas correspondientes a la unión del anticuerpo a la proteína objeto de estudio se realizó mediante el analizador de imágenes *Molecular Imager ChemiDoc XRS System* y el software *QuantityOne 4.6.6* (Bio-Rad).

### 3.9.8 Sobreexpresión y purificación de proteínas

Para realizar la sobreexpresión y posterior purificación de proteínas se utilizaron los vectores de la serie pLATE (*Thermo Scientific*). Estos plásmidos, que contienen el gen de interés clonado bajo un promotor dependiente de la ARN polimerasa del fago T7, se transformaron en células capaces de expresar dicha ARN polimerasa bajo el control de un promotor *lac* inducible por IPTG (en este caso se utilizó la cepa *E. coli* BL21 (DE3)). Las colonias obtenidas de la transformación se utilizaron para inocular 1 litro de medio de cultivo LB suplementado con carbenicilina 100 µg/ml (elemento marcador del plásmido pLATE) a una D.O<sub>600nm</sub> inicial de 0,05. Dicho cultivo se incubó posteriormente a 37°C en agitación constante (200 rpm). Cuando los cultivos llegaron a la fase exponencial (D.O<sub>600nm</sub> de 0,4), se añadió IPTG (*isopropyl-β-D-thiogalactoside*) a una concentración final de 1 mM, y se continuó con la incubación a 37°C durante 4 horas adicionales. Después de la incubación, se centrifugaron las células a 10.000 rpm durante 45 minutos a 4°C. Se descartó el sobrenadante y el sedimento celular se resuspendió en 20 ml de tampón A50 (20 mM de HEPES, 100 mM de KCl, 5 mM de MgCl<sub>2</sub> y 50 mM de imidazol) junto con un *cocktail* de inhibidores de proteasas (*cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Roche*). La suspensión celular disuelta en tampón se lisó en el homogenizador celular *GEA Lab Homogenizer PandaPLUS 1000* durante 3 ciclos repetidos a una presión de 600 bar. El lisado celular se centrifugó durante 45 minutos a 13.000 rpm a 4°C. El sobrenadante (libre de células) que contenía todas las proteínas intracelulares expresadas por la bacteria (entre las que se encontraba la proteína sobreexpresada con las 6 histidinas en el extremo N/C-terminal) se mezcló con 1 ml de resina de Ni<sup>2+</sup>-NTA agarosa (*Qiagen*) y se dejaron interaccionar durante 16 horas a 4°C. La mezcla se decantó en una columna de propileno (*Prepacked Gravity Flow Columns for Ion Exchange, Bio-Rad*). A continuación, se realizaron dos lavados de la resina con tampón A50. Para recoger y realizar la elución de las proteínas unidas a la matriz de Ni<sup>2+</sup>-NTA agarosa se utilizó un tampón con una concentración mayor de imidazol (elemento competidor contra las 6 histidinas por los sitios de unión a la matriz de níquel). Por ello, se obtuvieron 5 fracciones correspondientes a la elución de la resina con tampón A200 (20 mM de HEPES, 100 mM de KCl, 5 mM de MgCl<sub>2</sub> y 200 mM de imidazol). Las diferentes fracciones obtenidas se guardaron a -20°C para ser analizadas mediante electroforesis en geles de poliacrilamida y su posterior tinción con Azul de Coomassie.

#### **3.9.9 Identificación de proteínas mediante LC-MS/MS**

Con el fin de realizar la identificación de una proteína específica presente en un extracto proteico se utilizó el protocolo descrito por (Henzel y col., 1993). Para ello, se realizó una electroforesis de la muestra de interés en un gel de poliacrilamida y se tiñó con Azul de Coomassie. La banda de interés, correspondiente a la proteína específica a identificar, se recortó y se transfirió a un tubo de microcentrífuga. La identificación de la proteína de interés se realizó en colaboración con la “Plataforma de Proteómica del Parc Científic de Barcelona”. Para ello, la banda proteica se lavó mediante bicarbonato amónico ( $\text{NH}_4\text{HCO}_3$ ) a una concentración final de 50 mM haciendo uso del filtro Amicon Ultra (*Millipore*) (cuyo tamaño de poro es de 3 kDa), con el fin de eliminar posibles compuestos que pudieran interferir en la reacción de identificación. Seguidamente, la muestra fue tratada con DTT 20 mM durante 1 hora a 60°C y se alcalinizó con 25 mM de yodoacetamina durante 30 minutos a 30°C. Posteriormente, la muestra se digirió con 200 ng de tripsina (*Promega*) durante 16 horas a 37°C. Seguidamente, la mezcla de péptidos resultante fue purificada y secada al vacío. De esta manera, los péptidos digeridos se analizaron mediante un cromatógrafo líquido (*nanoAcquity*) acoplado a un espectrómetro de masas *TQ-Orbitrap Velos* (*Thermo Fisher Scientific*), donde los péptidos separados y eluidos por la cromatografía fueron sometidos a ionización por electrospray. Los espectros de masas de los diferentes péptidos obtenidos se analizaron mediante el *software Thermo Proteome Discoverer (v.1.4.1.14)* utilizando la base de datos *Swiss-Prot*.

#### **3.9.10 Aislamiento de flagelos**

Para realizar el aislamiento de los flagelos de las diferentes cepas objeto de estudio, se utilizaron cultivos de 250 ml (repartidos en dos matraces erlenmeyers con 125 ml de cultivo cada uno) crecidos durante toda la noche a 37°C en estático. Una vez crecidos, los diferentes cultivos se centrifugaron a 5.500 rpm durante 20 minutos a 4°C. A continuación, se descartó el sobrenadante y el sedimento celular se resuspendió en 5 ml de tampón Tris-HCl 0,1 M a pH 7,4 utilizando una varilla de vidrio y una fuerte agitación. Una vez resuspendido, el sedimento celular líquido se hizo pasar a través de una jeringuilla de 10 ml y una aguja de 21 G (*Polnet*). Dicho proceso se repitió en 5 ocasiones. La suspensión resultante se centrifugó a 8.200 rpm durante 30 minutos a 4°C. Seguidamente, se recogió el sobrenadante y se volvió a centrifugar a 12.000 rpm durante

20 minutos a 4°C. Finalmente, se realizó una ultracentrifugación del sobrenadante obtenido a 35.000 rpm durante 60 minutos a 4°C. Se descartó el sobrenadante, y el sedimento, correspondiente al aislamiento flagelar, se resuspendió en 75 µl de tampón Tris-HCl 0,1 M + EDTA 0,2 mM. El correcto aislamiento de los flagelos se analizó mediante electroforesis en gel de poliacrilamida y posterior tinción con Azul de Coomassie y tinción de plata.

### 3.9.11 Ensayos de estabilidad proteica *in vivo*

Los ensayos de estabilidad proteica *in vivo* se realizaron siguiendo el protocolo descrito por (Geuskens y col., 1992; Johansson y Uhlin, 1999). La estabilidad de la proteína analizada se monitorizó después de inhibir la síntesis proteica total mediante la adición de cloranfenicol (a una concentración final de 25 µgr/ml) a los cultivos crecidos hasta el inicio de la fase estacionaria de crecimiento en medio LB. Se tomó una muestra de cada cultivo a tiempo 0, 1, 2 h después de añadir el antibiótico para extraer el contenido proteico total (apartado 3.9.1) y realizar un *Western Blot* para analizar los niveles de expresión de la proteína estudiada.

## 3.10 Ensayos fisiológicos

### 3.10.1 Curvas de crecimiento y recuentos de viables

Cuando se hace crecer una cepa bacteriana en un medio de cultivo definido, ésta lo hace a una velocidad que es característica de esta cepa, del medio de cultivo utilizado y de las condiciones del entorno. El crecimiento de esta población será equilibrado mientras no se produzca una limitación de alguno de los nutrientes o no se acumule algún componente tóxico a una concentración demasiado elevada. Este tipo de crecimiento se puede definir matemáticamente como una reacción de primer orden en la que la velocidad de aumento del número de bacterias en un momento determinado es proporcional al número de bacterias (N) presentes en ese momento:

$$\circ \quad dN/dt = \mu \cdot N,$$

siendo  $\mu$  un parámetro constante y específico a cada bacteria denominado tasa de crecimiento. La fórmula matemática que permite calcular la tasa de crecimiento de un determinado cultivo bacteriano es la siguiente:

$$\circ \quad \mu = (\log N - \log N_0) \cdot 2,3 / \Delta t$$



### 3. MATERIALES Y MÉTODOS

Para poder determinar este parámetro cinético, se realizaron curvas de crecimiento en las que se definió la fase de crecimiento exponencial de los cultivos correspondientes con el fin de poder calcular su tasa de crecimiento. Para realizar las curvas de crecimiento de las cepas estudiadas, se realizó un cultivo previo de la bacteria correspondiente en 10 ml de medio LB a 37°C en agitación constante (200 rpm) durante 16 horas. Al día siguiente, se realizó una dilución 1:100 de dicho cultivo en 25 ml de medio fresco y se determinó la cinética de crecimiento celular a la temperatura correspondiente. Para ello, en cada intervalo de tiempo determinado, se recogió una alícuota de 1 ml de cultivo y se transfirió a una cubeta de plástico de 1,5 ml para espectrofotómetro (*Quimikals Analítica S.L*), para posteriormente cuantificar la absorbancia (o turbidez) del cultivo a una longitud de onda de 600 nm en el espectrofotómetro *Genesys 10S UV-Vis (Thermo Scientific)*.

Para realizar los recuentos de células viables (ufc/ml) de cada una de las cepas estudiadas bajo las diferentes condiciones de crecimiento analizadas, se tomaron alícuotas de 1 ml de cultivo en cada fracción de tiempo correspondiente y se realizó un banco de diluciones con NaCl 0,9% (p/v), sembrando por duplicado 100 µl de las diluciones específicas en placas de LB-agar.

#### **3.10.2 Ensayos de autoagregación**

Para realizar los ensayos de cuantificación de la autoagregación bacteriana, se crecieron los cultivos correspondientes en 20 ml de medio LB durante 16 horas a 25°C o a 37°C en agitación constante (200 rpm). Una vez crecidos, se transfirieron 10 ml de cada cultivo a un tubo de ensayo y se mantuvieron en estático, tomándose alícuotas de 100 µl de la superficie del cultivo cada 20 minutos para medir la densidad óptica de la muestra en el espectrofotómetro. La autoagregación se representó como el descenso de la absorbancia a 600 nm como consecuencia de la sedimentación de las células hacia la base del tubo.

#### **3.10.3 Ensayos de motilidad**

Los experimentos para analizar la motilidad bacteriana en placa se realizaron a partir de cultivos de las cepas correspondientes crecidos en 10 ml de medio rico a 37°C en agitación constante durante 16 horas. Para realizar los ensayos de motilidad se utilizaron placas que contenían medio LB y medio M9, en ambos casos suplementadas con 0,3% de agar, en las cuales se inoculó en el centro de las mismas 5 µl del cultivo bacteriano crecido previamente. Las placas se incubaron a 37°C hasta que se pudo observar el halo de motilidad producido por las bacterias.

En el caso de los ensayos realizados con medio LB 0,3% agar, las medidas del halo de motilidad se realizaron 16 horas después de la inoculación, mientras que en el caso de las placas con medio M9 0,3% agar las medidas para cuantificar el halo de motilidad bacteriano se realizaron a las 24, 48 y 72 horas post-inoculación.

#### **3.10.4 Ensayos de resistencia al choque ácido**

Los ensayos de resistencia a pH ácido se realizaron siguiendo el protocolo descrito por (Chang y Cronan, 1999). Las diferentes cepas sometidas al *shock* por estrés ácido se crecieron en medio LB hasta el inicio de la fase estacionaria de crecimiento (D.O<sub>600nm</sub> 2,0). Seguidamente, se añadió HCl 6 N hasta que el pH del cultivo correspondiente descendió a 3,2. Después de la adición del ácido, los cultivos se incubaron de nuevo a 37°C en agitación constante de 200 rpm durante 0, 40 y 120 minutos. En los correspondientes intervalos de tiempo, se tomó una muestra de cultivo, se realizó un banco de diluciones con NaCl 0,9% (p/v) y se plaquearon las células en placas de LB-agar para realizar el recuento de unidades formadoras de colonias (ufc/ml) y cuantificar de esta manera cómo había afectado el estrés ácido a la viabilidad celular.

#### **3.10.5 Ensayos de resistencia al choque frío**

Para realizar los ensayos de tolerancia al choque por disminución de la temperatura, las diferentes cepas sometidas al estrés se crecieron en medio LB hasta el inicio de la fase estacionaria de crecimiento (D.O<sub>600nm</sub> 2,0). Seguidamente, se disminuyó la temperatura de los cultivos hasta 4°C durante intervalos de tiempo de 30 minutos, 1 hora y 4 horas. Pasado este tiempo, se tomó una muestra de cada cultivo, se realizó un banco de diluciones con NaCl 0,9% (p/v) y se plaquearon las células en placas de LB-agar para realizar el recuento de unidades formadoras de colonias (ufc/ml) y ver cómo había afectado la brusca disminución de la temperatura del medio a la viabilidad celular.

#### **3.10.6 Ensayos de persistencia bacteriana ante concentraciones letales de antibiótico**

Los ensayos de tolerancia a concentraciones letales de un antibiótico (al que las células son sensibles) se realizaron siguiendo el protocolo descrito por (Harms y col., 2017). Las diferentes cepas sometidas al ensayo de persistencia se crecieron en LB hasta el inicio de la fase estacionaria (D.O<sub>600nm</sub> 2,0). Una vez las células llegaron a este punto de crecimiento, se traspasaron 10 ml del cultivo a sendos tubos de propileno de 50 ml y se

### 3. MATERIALES Y MÉTODOS

añadió carbenicilina a una concentración final de 100 µgr/ml. Después de la adición del antibiótico, los cultivos se incubaron a 37°C en agitación constante (200 rpm) durante 1, 2, 3, 4 y 5 horas, tomándose una muestra de cada cultivo correspondiente en los intervalos de tiempo indicados. Se realizaron diluciones seriadas de cada muestra con NaCl 0,9% (p/v) y se plaquearon las células en placas de LB-agar para realizar el recuento ufc/ml y comprobar cómo había afectado la adición del antibiótico a la viabilidad celular.

#### **3.10.7 Experimentos de conjugación bacteriana en medio líquido**

Los experimentos para cuantificar la tasa de conjugación del plásmido pAA2 de la cepa *E. coli* 042 se realizaron a partir de cultivos crecidos durante 16 horas a 37°C en medio 2xYT, tanto de la cepa donadora del plásmido como de la receptora. Primeramente, se atemperaron 900 µl del medio de cultivo fresco 2xYT a la temperatura en que se realizarían los experimentos de conjugación (25°C o 37°C). Seguidamente, se añadieron 50 µl del cultivo de la cepa donadora del plásmido y se incubó a 600 rpm durante 30 minutos (en el caso que la temperatura de conjugación fuera de 37°C) o 60 minutos (en el caso que la temperatura de conjugación fuese de 25°C). Pasado este tiempo de incubación, se añadieron 200 µl del cultivo de la cepa receptora y la mezcla se incubó a la temperatura correspondiente en estático para facilitar el contacto entre la célula donadora y la receptora. Finalmente, la conjugación se detuvo dando un golpe de vórtex a la suspensión e incubando las células en hielo durante 1 minuto. Por último, se sembró la suspensión celular en medio de cultivo sólido selectivo para los marcadores fenotípicos específicos transferidos en la conjugación (placas de LB-agar suplementadas con el antibiótico correspondiente). La frecuencia de conjugación se calculó mediante la relación células transconjugantes/células donadoras, por lo que previamente a cada experimento se calculó el número de células donadoras mediante la siembra sobre medio sólido de un banco de diluciones en NaCl 0,9% (p/v) del cultivo “*overnight*” inicial.

#### **3.10.8 Curado del plásmido pAA2**

Puesto que el plásmido pAA2 no codifica para ningún determinante de resistencia a antibióticos, como paso previo a la realización del protocolo de curado del plásmido pAA2 de la cepa *E. coli* 042, se tuvo que introducir en el plásmido de la bacteria un casete de resistencia a un antibiótico que nos permitiera seleccionar aquellas células que hubieran perdido el plásmido respecto a las que no. El protocolo utilizado para introducir el marcador de resistencia (concretamente un gen de resistencia a la kanamicina) es el

descrito en el apartado 3.7.8 de esta memoria. Una vez obtenido el clon con el módulo de resistencia insertado en su plásmido, se realizó un cultivo “*overnight*” del mismo a 37°C y se realizaron subcultivos seriados en 10 ml en medio LB sin antibióticos durante 10 días seguidos, utilizando como inóculo una dilución 1:1000 del cultivo del día anterior. Una vez llegado al décimo día de subcultivo, se plaquearon diluciones seriadas del mismo en placas de LB-agar suplementado con el antibiótico novobiocina a una concentración de 3,3 µg/ml. Las colonias aisladas que crecieron se picaron en paralelo en placas de LB-agar y placas de LB-agar suplementado con kanamicina 50 µg/ml. A las colonias sensibles a kanamicina, pero viables en LB-agar, se les realizó una PCR para analizar la presencia/ausencia de varios genes localizados en diferentes regiones del plásmido para confirmar la pérdida total del mismo.

### **3.10.9 Ensayos de formación de *biofilm***

Para estudiar la formación de *biofilm* por parte de la cepa *E. coli* 042 se utilizó el protocolo descrito por (Sheikh y col., 2001). La superficie abiótica donde se realizaron estos ensayos fueron placas de poliestireno (*Nunclon Delta Surface, ThermoScientific*). Las diferentes cepas utilizadas en estos ensayos se hicieron crecer previamente en medio rico LB a 37°C hasta la fase estacionaria y se inocularon en las siguientes cantidades de medio en función del tipo de placa utilizada:

- Si la placa utilizada era de 96 pocillos, se utilizaron 200 µl de medio y 2 µl del cultivo bacteriano inicial en fase estacionaria.
- Si la placa utilizada era de 24 pocillos, se utilizaron 800 µl de medio y 8 µl del cultivo bacteriano inicial en fase estacionaria.

Las placas con las bacterias y el medio fresco se incubaron durante toda la noche a 37°C. Pasado este tiempo, se retiró el medio de la placa con cuidado de no alterar la monocapa celular formada sobre la superficie de los distintos pocillos. En el caso de la placa de 24 pocillos, se realizaron 2 lavados con 1 ml de PBS; en el caso de la placa de 96 pocillos, se realizaron 2 lavados con 200 µl de PBS. Posteriormente, se añadió a cada pocillo de la placa de 24 pocillos 100 µl de cristal violeta 0,5% (p/v) para teñir la monocapa bacteriana (en el caso de la placa de 96 se añadió 50 µl de cristal violeta). Dejamos interaccionar el colorante con la monocapa durante 5 minutos a temperatura ambiente, para después retirar el colorante y realizar varios lavados con agua destilada para eliminar las células planctónicas. Sobre la placa de 96 pocillos (con la que se realizó la cuantificación del

### 3. MATERIALES Y MÉTODOS

*biofilm* generado), se añadieron 200  $\mu$ l de etanol al 95% (v/v) y se incubó 2 minutos a temperatura ambiente con la finalidad de solubilizar el *biofilm* formado en los pocillos. A continuación, se recogieron 150  $\mu$ l de cada pocillo y se transfirieron a una placa nueva para realizar medidas espectrofotométricas a una D.O de 570 nm.

#### 3.10.10 Ensayos de análisis de actividad $\beta$ -galactosidasa

Con el fin de estudiar la expresión transcripcional de los diferentes genes objeto de estudio fusionados con el gen reportero *lacZ* se realizó la prueba de la  $\beta$ -galactosidasa (Miller, 1992). Para determinar la actividad  $\beta$ -galactosidasa se crecieron los cultivos correspondientes hasta la fase de crecimiento exponencial (D.O<sub>600nm</sub> de 0,4) o el inicio de la fase estacionaria (D.O<sub>600nm</sub> de 2,0), y se recogieron alícuotas de 1 ml de cada uno de los cultivos. De cada una de las alícuotas, se tomaron 100  $\mu$ l y se añadieron a 900  $\mu$ l de tampón Z (60 mM de Na<sub>2</sub>HPO<sub>4</sub>, 40 mM de NaH<sub>2</sub>PO<sub>4</sub>, 10 de mM KCl, 1 mM de MgSO<sub>4</sub>, 50 mM  $\beta$ -mercaptoetanol). Para provocar el lisado de las células, se añadieron 25  $\mu$ l de SDS 0,1% (p/v) y 50  $\mu$ l de cloroformo, y se agitó la mezcla en el vórtex durante 15 segundos. Tras la ruptura celular, se atemperó la mezcla a 28°C durante 5 minutos. A continuación, se añadió a cada muestra 200  $\mu$ l de ONPG (orto-nitrofenil- $\beta$ -D-galactopiranosido, *Sigma-Aldrich*) preparado a una concentración de 4 mg/ml en tampón Z y se incubó la reacción a 28°C hasta la aparición de un color amarillo (D.O.<sub>420nm</sub> entre 0,3 y 0,9). Para detener la reacción, se añadió a cada muestra 0,5 ml de Na<sub>2</sub>CO<sub>3</sub> a una concentración final de 1 M, y se dejaron reposar las muestras unos minutos en oscuridad. Finalmente, se midieron las absorbancias de las diferentes muestras a dos longitudes de onda (420 y 550 nm) y se cuantificó la actividad  $\beta$ -galactosidasa utilizando la siguiente fórmula:

$$\text{Actividad } \beta - \text{galactosidasa (U.M)} = \frac{1000 \cdot (\text{D.O. } 420\text{nm} - 1.75 \cdot \text{D.O. } 550\text{nm})}{T \cdot V \cdot \text{D.O. } 600\text{nm}}$$

U.M.: Unidades Miller.

D.O<sub>420nm</sub> y D.O<sub>550nm</sub>: absorbancias de la muestra a las longitudes de onda de 420 y 550 nm una vez detenida la reacción al añadir Na<sub>2</sub>CO<sub>3</sub> 1 M.

D.O<sub>600nm</sub>: absorbancia del cultivo a 600 nm en el momento en el que se tomó la muestra.

T: tiempo transcurrido (minutos) desde que se añadió el ONPG hasta que se detuvo la reacción.

V: volumen de muestra (ml) utilizado para determinar la actividad  $\beta$ -galactosidasa.

## **4. RESULTADOS**





# Capítulo 1

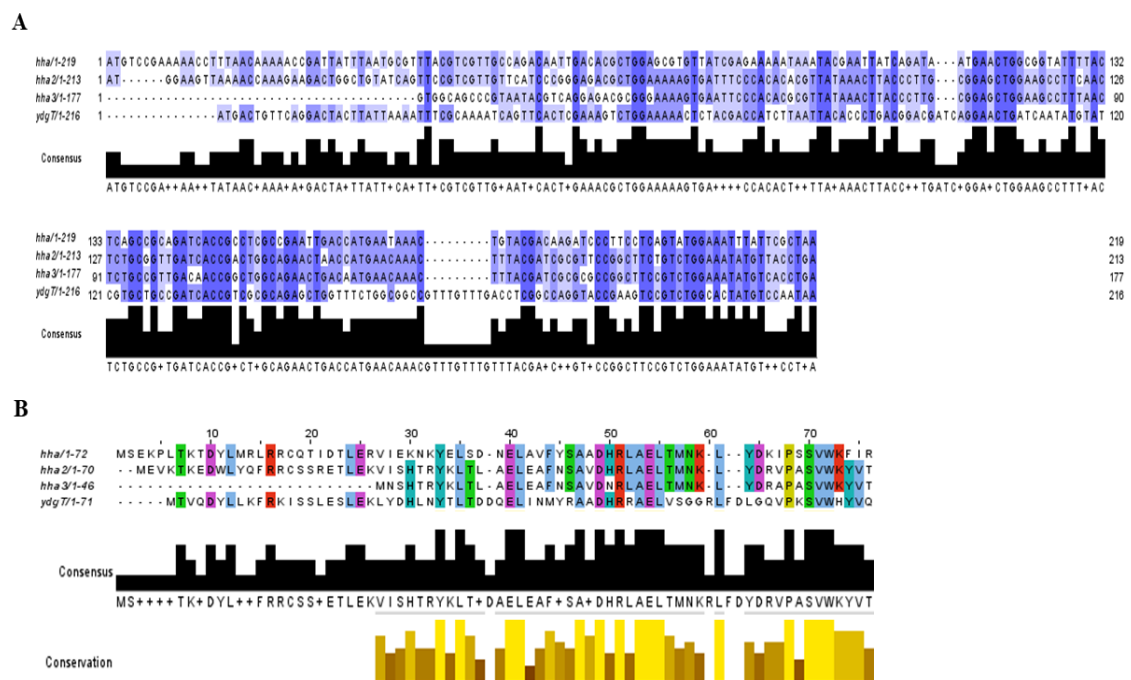


## **4.1. Los genes *hha2* y *hha3*, nuevos parálogos del gen *hha* en la cepa *E. coli* 042**

#### 4. RESULTADOS

##### 4.1.1 Identificación de nuevos parálogos del gen *hha* en el genoma de la cepa *E. coli* 042

Mediante una búsqueda de secuencias homólogas a la secuencia nucleotídica y aminoacídica del gen que codifica para la proteína Hha (NCBI-EC042\_0498 y Uniprot-D3GV89) utilizando la herramienta BLAST (*Basic Local Alignment Search Tool*, programa que permite identificar regiones de homología entre diferentes secuencias biológicas) en las bases de datos de Uniprot (<http://www.uniprot.org/blast/>) y NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), se logró identificar dos nuevos genes parálogos al gen *hha* en el genoma de la cepa *E. coli* 042, los genes EC042\_4516 y EC042\_4796. Los mismos fueron denominados *hha2* y *hha3*, respectivamente. Estos dos nuevos genes se suman al gen *ydgT*, primer gen parálogo de *hha* descrito que está presente en todas las cepas de *E. coli* referenciadas hasta la fecha. En la figura 4.1 se representa el alineamiento de las secuencias de nucleótidos y aminoácidos de los diferentes genes tipo *hha* y las correspondientes proteínas presentes en el cromosoma de *E. coli* 042.



**Figura 4.1.** Alineamiento de las secuencias nucleotídicas (A) y aminoacídicas (B) de los genes *hha*, *hha2*, *hha3* y *ydgT*, codificados en el cromosoma de la cepa *E. coli* 042. Los bloques en negro representan la secuencia consenso global de los 4 genes analizados. A nivel de la secuencia nucleotídica, en color azul intenso se representan las posiciones cuyas secuencia de nucleótidos se mantienen más conservadas. La intensidad de color disminuye a medida que los nucleótidos varían entre los diferentes genes. A nivel de la secuencia proteica, se representan en diferente color los distintos aminoácidos conservados entre las diferentes posiciones aminoacídicas entre las 4 proteínas. Alineamiento múltiple realizado mediante el programa Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

Tal y como se comentaba en el apartado 1.2.2 de la Introducción, las proteínas de la familia Hha son proteínas de bajo peso molecular implicadas en la regulación de la expresión génica en enterobacterias. En la tabla 4.1 se describen las características referentes al tamaño de los genes y las proteínas correspondientes a los alelos tipo *hha* presentes en el cromosoma de *E. coli* 042.

Gen (identificador NCBI locus tag)	Tamaño gen (pares de bases)	Tamaño proteína (aminoácidos)	Peso Molecular (kDa)
<i>hha</i> (EC042_0498)	219	72	8,62
<i>hha2</i> (EC042_4516)	213	70	8,39
<i>hha3</i> (EC042_4796)	177	58	6,46
<i>ydgT</i> (EC042_1794)	216	71	8,41

**Tabla 4.1.** Tabla resumen en la que se indica el tamaño de los genes *hha*, *hha2*, *hha3* e *ydgT*, además del número de aminoácidos que conforman cada proteína correspondiente y su peso molecular.

Respecto a la proteína Hha, la proteína Hha2 presenta un nivel de similitud del 78% y un nivel de identidad del 58%; la proteína Hha3 muestra un nivel de similitud del 82% y un nivel de identidad del 63%; por último, la proteína YdgT, presenta unos niveles de similitud e identidad en relación a la proteína Hha del 67% y del 38%, respectivamente.

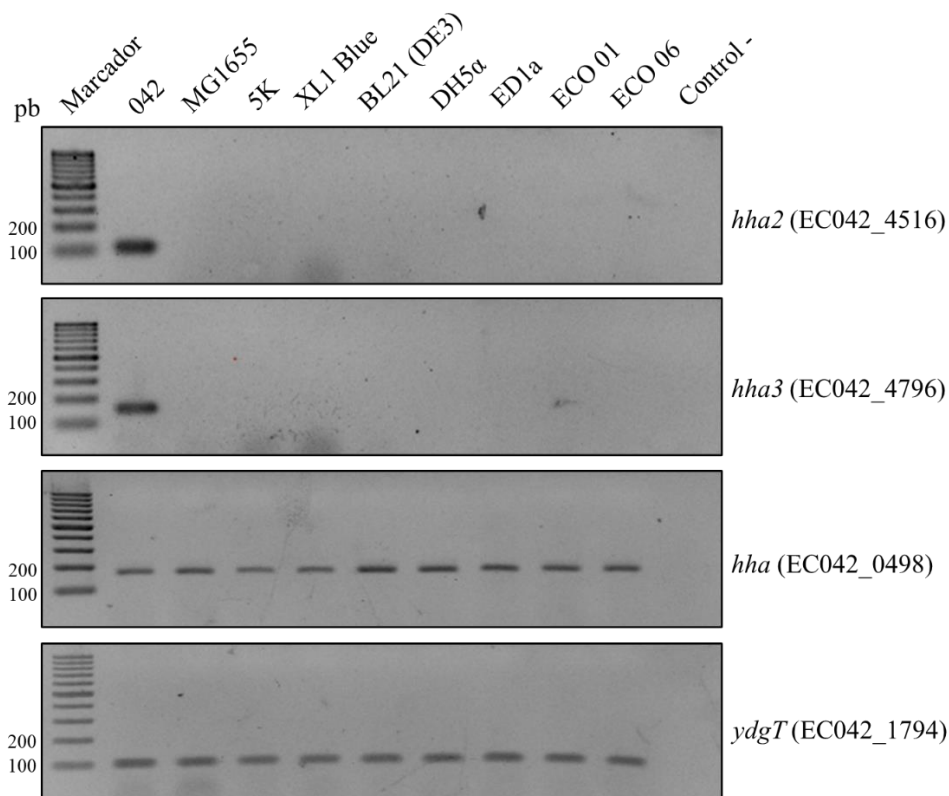
El mayor número de aminoácidos conservados entre las 4 proteínas se concentran en la región C-terminal de las mismas (figura 4.1 B). La proteína Hha3 parece ser una forma truncada de las proteínas Hha/Hha2 ya que carece de dominio N-terminal, de manera que la región codificante de dicha proteína tiene una alta homología únicamente con el dominio C-terminal de ambas proteínas.

#### 4.1.2 Distribución de los parálogos *hha2/hha3* entre diferentes cepas de *E. coli*

Una primera pregunta que se planteó fue si los genes *hha2/hha3* son específicos únicamente de la cepa 042, de cepas enteroagregativas de *E. coli* o, alternativamente, si se encuentran ampliamente distribuidos entre diferentes cepas de los diferentes patotipos de *E. coli*. Para poder analizar la distribución de estos nuevos parálogos entre diferentes cepas de *E. coli*, se diseñaron oligonucleótidos específicos con el fin de detectar cada gen mediante la técnica de reacción en cadena de la polimerasa (PCR). Debido al elevado grado de similitud entre las secuencias nucleotídicas de los diferentes genes parálogos, se utilizó el programa *Primer Blast* (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) para poder diseñar oligonucleótidos que permitieran detectar de manera específica los genes

#### 4. RESULTADOS

*hha2* y *hha3* sin dar una reacción cruzada con el gen *hha*. La secuencia de los oligonucleótidos utilizados y su lugar de hibridación se encuentran detallados en la tabla 9.1 y en la figura 9.2 presentes en el Anexo de esta memoria. Para confirmar la correcta especificidad de los oligonucleótidos diseñados, se realizó una PCR con todos ellos con el fin de detectar los genes *hha*, *hha2*, *hha3* e *ydgT* en las cepas *E. coli* 042, *E. coli* MG1655, *E. coli* 5K, *E. coli* XL1Blue, *E. coli* BL21(DE3), *E. coli* DH5 $\alpha$ , *E. coli* ED1a y dos cepas comensales de la colección ECOR (ECO 01 y ECO 06) (figura 4.2).

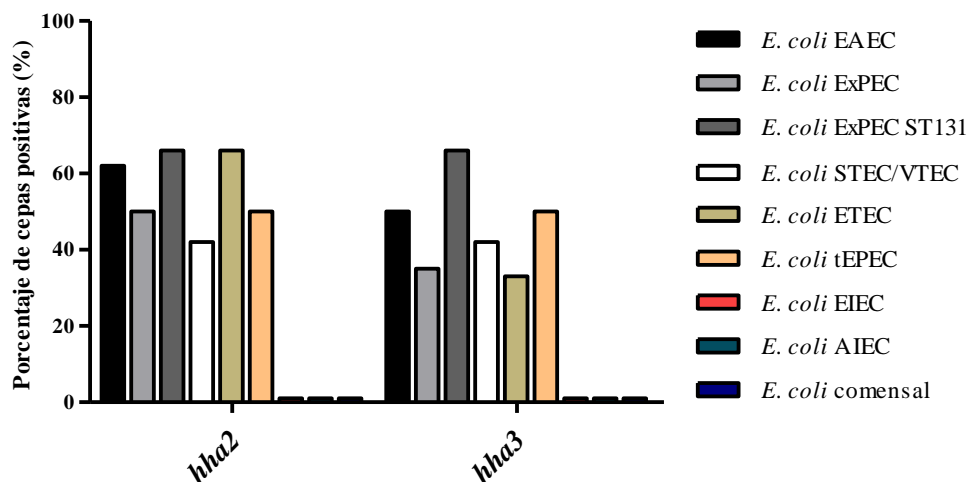


**Figura 4.2.** Amplificación por PCR de los genes *hha2*, *hha3*, *hha* e *ydgT* en diferentes cepas de *E. coli*. Se indica también el tamaño en pb (pares de bases) de las bandas del marcador de peso molecular utilizado (*GeneRuler 100 bp DNA Ladder, Thermo Scientific*) entre las que se encuentran los diferentes productos de PCR obtenidos.

Tal y como se puede observar, los oligonucleótidos diseñados nos han permitido amplificar y detectar de manera específica los alelos *hha2* y *hha3* en la cepa *E. coli* 042, pero no en el resto de cepas de *E. coli* utilizadas en este análisis. Por el contrario, y tal y como era de esperar, se han podido detectar tanto los genes *hha* como *ydgT* en todas las cepas estudiadas. A excepción de la cepa *E. coli* 042, el resto de las cepas que fueron analizadas mediante este análisis por PCR son cepas comensales o cepas de uso rutinario en laboratorio. Por ello, el hecho de que los genes *hha2* y *hha3* no pudieran ser detectados en cepas no virulentas de *E. coli* pero si en la cepa *E. coli* 042, nos llevó a plantear la

hipótesis de que la presencia de estos parálogos pudiera estar restringida únicamente en cepas patógenas de *E. coli*. Por ello, el siguiente paso fue analizar por PCR la distribución de los genes *hha2* y *hha3* entre cepas pertenecientes a diferentes patotipos de *E. coli* haciendo uso de la colección LREC (Laboratorio de Referencia de *E. coli*), una colección de 60 cepas de *E. coli* de origen humano entre las que se incluyen 16 cepas enteroagregativas (EAEC), 21 cepas productoras de toxina Shiga (STEC/VTEC), 3 cepas enterotoxigénicas (ETEC), 2 cepas enteropatógenas típicas (tEPEC), 2 cepas enteroinvasivas (EIEC), 1 cepa adherente invasiva (AIEC), 14 cepas patógenas extraintestinales (ExPEC), de las cuales 6 pertenecen al grupo clonal ST131, y 1 cepa comensal de *E. coli*. Los resultados del análisis por PCR para detectar la presencia/ausencia de los genes *hha2/hha3* en estas cepas se muestran en la figura 4.3.

Patotipo	Número de cepas analizadas	Número de cepas positivas para <i>hha2</i>	Número de cepas positivas para <i>hha3</i>
EAEC	16	10 (62%)	8 (50%)
ExPEC	14	7 (50%)	5 (35%)
ExPEC ST131	6	4 (66%)	4 (66%)
STEC/VTEC	21	9 (42%)	9 (42%)
ETEC	3	2 (66%)	1 (33%)
tEPEC	2	1 (50%)	1 (50%)
EIEC	2	0	0
AIEC	1	0	0
Comensal	1	0	0



**Figura 4.3.** Distribución de los genes *hha2* y *hha3* entre cepas de diferentes patotipos de *E. coli* representados en la colección LREC.

#### 4. RESULTADOS

Los resultados obtenidos indican que un 62% de las cepas EAEC analizadas contenían el gen *hha2* en su genoma, mientras que en el caso del gen *hha3* el porcentaje es inferior, un 50%. A nivel de las cepas STEC/VTEC, la presencia de los alelos *hha2* y *hha3* disminuye (42% para ambos genes) respecto al patotipo EAEC. En referencia a las cepas patógenas extraintestinales de *E. coli*, un 50% de las cepas estudiadas resultaron positivas para el gen *hha2* y un 35% para el gen *hha3*. Respecto al grupo clonal ST131 perteneciente a este patotipo, el 66% de las cepas analizadas contenían los alelos *hha2* y *hha3*. En la tabla 9.3 del apartado Anexo se describe con más detalle serotipo y el grupo filogenético al que pertenece cada cepa analizada de la colección LREC y el resultado del análisis por PCR realizado para analizar la presencia o ausencia de los genes *hha2* y *hha3*. También se analizó la presencia de los genes *hha* e *ydgT* en todas las cepas de la colección LREC, resultando, tal y como era de esperar, todas las bacterias positivas para la presencia de ambos genes.

##### **4.1.2.1 Distribución de los alelos *hha2/hha3* entre cepas enteroagregativas de *E. coli***

Dada la elevada prevalencia de los genes *hha2* y *hha3* en el patotipo enteroagregativo de *E. coli*, se analizó por PCR la presencia de estos parálogos en una colección más extensa de cepas EAEC (concretamente 56 cepas más). Además, también se trató de correlacionar la presencia de estos parálogos con la presencia de diferentes factores de virulencia expresados por estas cepas. Concretamente, además de los alelos *hha2* y *hha3*, en estas 56 cepas EAEC también se realizó un análisis por PCR con el fin de detectar los genes que codifican por factores de virulencia típicos de las cepas enteroagregativas de *E. coli*:

- *aata* (*antiaggregation protein transporter*)
- *aggA* (subunidad de la fimbria AAF/I)
- *aafA* (subunidad de la fimbria AAF/II)
- *agg3A* (subunidad de la fimbria AAF/III)
- *aggR* (regulador transcripcional AggR)
- *astA* (*aggregative heat-stable toxin 1* (toxina EAST1)) \*
- *aap* (*anti-aggregation protein* (Dispersina))
- *set1A* (*Shigella enterotoxin 1 mucinase*)
- *irp2* (*Yersinia bactin biosynthesis*)
- *pet* (*serine protease Pet*)
- *sfh* (*polysaccharide deacetylase*)
- *sat* (*secreted autotransporter toxin*) \*\*
- *pic* (*serine protease Pic*)
- *agn43* (*autotransporter antigen 43*)

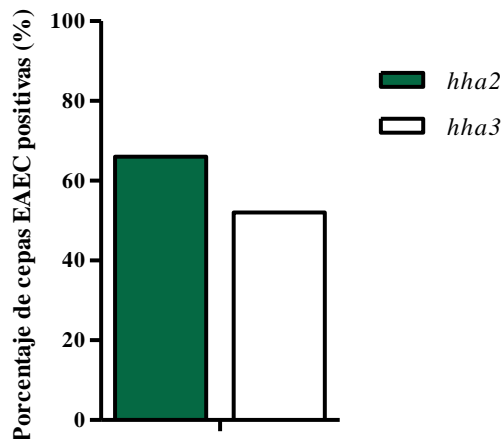
\* *astA*: Además del patotipo EAEC, la presencia del gen *astA* también se asocia a los patotipos EHEC y EPEC.

\*\* *sat*: la presencia del gen *sat* también se asocia mayoritariamente con cepas ExPEC.

En la tabla 9.4 del Anexo se indican todos los resultados de los análisis de PCR realizados en colaboración con el grupo LREC (Laboratorio de Referencia de *E. coli*, dirigido por el Profesor Jorge Blanco), además del serotipo y el grupo filogenético al que pertenecen las 56 cepas EAEC analizadas. Cabe destacar que, tal y como ocurriera con el análisis anterior, todas las cepas analizadas contenían tanto los genes *hha* como *ydgT* en su cromosoma.

A nivel global, se ha podido constatar la elevada prevalencia de los parálogos *hha2* y *hha3* en el patotipo EAEC. Concretamente, el gen *hha2* se ha encontrado presente en 37 (66%) de las 56 cepas analizadas, de las cuales 10 contenían el gen *hha2* únicamente y 27 lo contenían en combinación con *hha3*. Respecto a las cepas positivas para la presencia del alelo *hha3* (29 cepas en total, el 52% de las cepas analizadas), únicamente 2 de ellas tenían el alelo *hha3* en solitario, mientras que las otras 27 cepas restantes también contenían el gen *hha2* en su genoma. De esta manera, los dos alelos se hallaron presentes de manera conjunta en 27 cepas (48% del total). Al realizar el test exacto de Fisher sobre los resultados de los análisis por PCR a nivel de la presencia conjunta de cada uno de los 14 factores de virulencia de EAEC y los genes *hha2/hha3*, se pudo comprobar que existía una fuerte correlación entre la presencia de los genes *aggA*, *astA*, *shf*, *sat* y *agn43* y los alelos *hha2* y *hha3*, entre los que se ha encontrado una fuerte asociación significativa. En la figura 4.4 se representa la prevalencia de estos 5 genes de virulencia en las 56 cepas EAEC analizadas con relación a la presencia de un único alelo tipo *hha* (*hha2+/hha3-* o *hha2-/hha3+*), de ambos alelos (*hha2+/hha3+*) o ningún alelo (*hha2-/hha3-*).

A



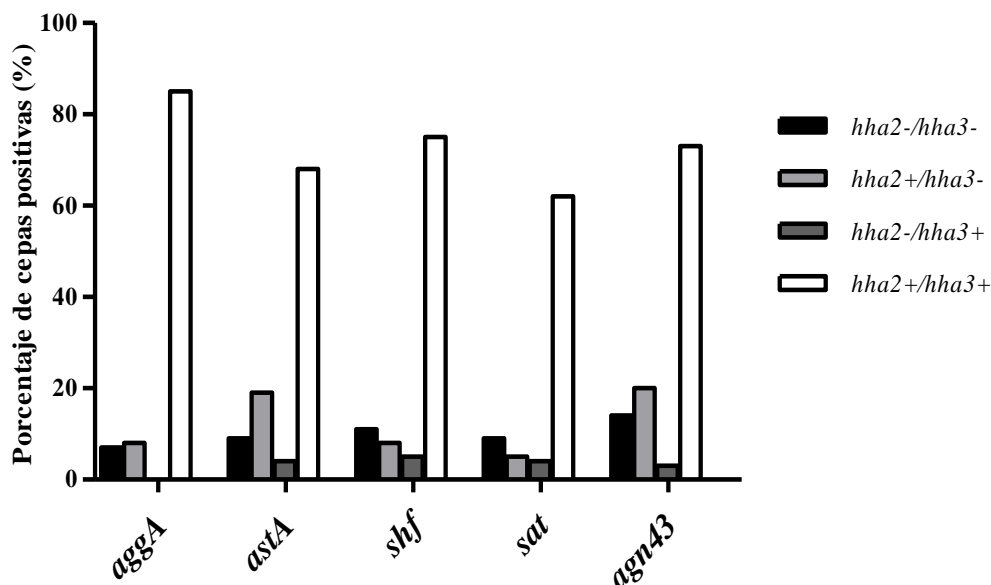


4. RESULTADOS

B

		Genes de virulencia de <i>Escherichia coli</i> enteroagregativa (EAEC)						
		<i>aggA</i>	<i>astA</i>	<i>shf</i>	<i>sat</i>	<i>agn43</i>	<i>hha2</i>	<i>hha3</i>
TOTAL:56 cepas EAEC		13 23%	22 39%	23 41%	23 41%	36 64%	37 66%	29 52%
17 cepas <i>hha2-/hha3-</i>		1 6%	2 12%	3 18%	2 12%	5 29%	-	-
10 cepas <i>hha2+/hha3-</i>		1 10%	4 40%	2 20%	6 60%	7 70%	10 18%	-
2 cepas <i>hha2-/hha3+</i>		0	1 50%	1 50%	1 50%	1 50%	-	2 3,5%
27 cepas <i>hha2+/hha3+</i>		11 41%	15 56%	17 63%	14 52%	23 85%	27 48%	27 48%
Test de exactitud de Fisher	P valor	<i>hha2-/hha3-</i> vs <i>hha2+/hha3+</i>						
		0.01	0.004	0.004	0.007	0.0003		
Test de exactitud de Fisher	P valor	<i>hha2-/hha3-</i> vs <i>hha2+</i> y/o <i>hha3+</i>						
		0.04	0.005	0.02	0.003	0.0005		

C



**Figura 4.4.** (A) Análisis de la presencia de los genes *hha2* y *hha3* entre las 56 cepas EAEC de la colección LREC. (B) Análisis de la distribución de los 5 genes (*aggA*, *astA*, *shf*, *sat* y *agn43*) y los alelos *hha2* y *hha3* entre los genomas de las 56 cepas EAEC de la colección LREC. El número de arriba de cada celda indica el número de cepas que tienen el gen analizado y el de abajo indica el porcentaje de la presencia del gen relativo a todas las cepas analizadas (primera fila, 56 cepas EAEC) o relativo a la presencia o ausencia de los genes *hha2/hha3* (4 filas siguientes). También se indica el *P* valor (últimas dos filas) obtenido al realizar la prueba exacta de Fisher para analizar estadísticamente la presencia conjunta de los 5 genes de virulencia analizados junto con *hha2/hha2*. (C) Representación gráfica de la distribución de los 5 genes de virulencia que muestran asociación significativa con la presencia de *hha2/hha3* en función de si la cepa codifica únicamente el alelo *hha2* (*hha2+/hha3-*), *hha3* (*hha2-/hha3+*), ambos (*hha2+/hha3+*) o ninguno de ellos (*hha2-/hha3-*).

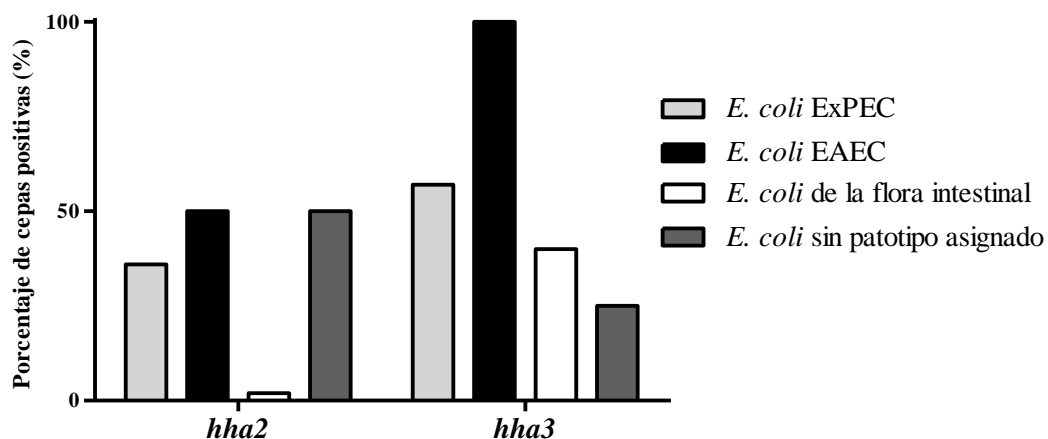
Tal y como se puede observar en la anterior figura, la presencia de cada uno de los 5 genes que codifican por factores de virulencia característicos de las cepas EAEC (*aggA*, *astA*, *shf*, *sat* y *agn43*) ocurre con una mayor incidencia en cepas que contienen de manera conjunta la combinación de los dos alelos parálogos de *hha* (*hha2*+/*hha3*+).

#### 4.1.2.2 Distribución de los genes *hha2/hha3* entre cepas de *E. coli* productoras de beta-lactamasas de amplio espectro

Para completar el análisis, también se estudió la presencia de los genes *hha2* y *hha3* en una colección de 25 aislamientos de *E. coli* (principalmente cepas ExPEC productoras de beta-lactamasas de amplio espectro (*ESBL*)) provenientes de humanos y animales de compañía (todas las cepas fueron aisladas en Alemania en colaboración con el Profesor Trinad Chakraborty). En dicha colección, el alelo *hha2* se encontró presente en el 32% de las cepas analizadas (8 cepas positivas respecto a las 25 analizadas), mientras que el gen *hha3* se detectó en el 52% de las mismas (13 cepas positivas respecto a las 25 analizadas) (figura 4.5).

Categoría patogénica	Número de cepas analizadas	Número de cepas positivas para <i>hha2</i>	Número de cepas positivas para <i>hha3</i>
ExPEC	14	5 (36%)	8 (57%)
EAEC	2	1 (50%)	2 (100%)
Flora intestinal	5	0 (0%)	2 (40%)
N.A*	4	2 (50%)	1 (25%)

\*No se ha podido asignar un patotipo específico



**Figura 4.5.** Distribución de los genes *hha2* y *hha3* entre 25 cepas de *E. coli* de origen humano y animal (principalmente del patotipo ExPEC) productoras de beta-lactamasas de amplio espectro.

Cabe destacar que dichos alelos son predominantes en las cepas ExPEC/EAEC y son menos frecuentes en aquellos aislamientos considerados como flora intestinal. Ya que los genomas de esta colección de cepas están secuenciados, se correlacionaron los datos

#### 4. RESULTADOS

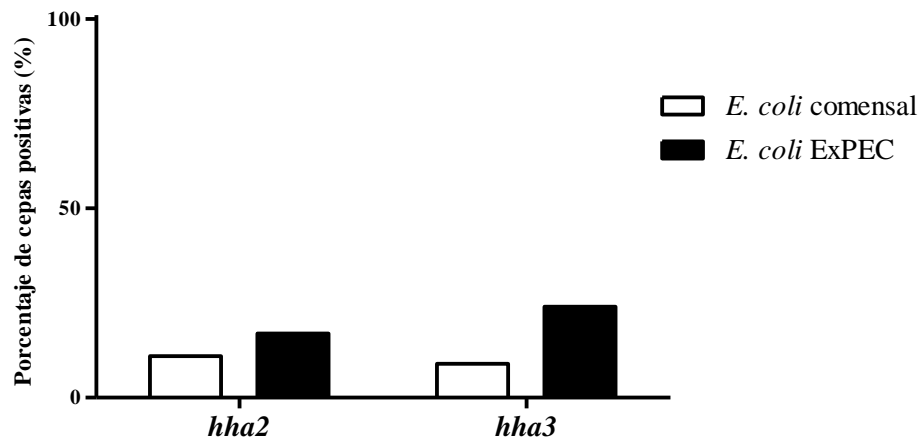
obtenidos mediante PCR con los datos genómicos (tabla 9.5 del Anexo de esta memoria, en la cual se indica también el origen de las cepas utilizadas, la fuente de aislamiento de las mismas, el patotipo y el grupo filogenético al que pertenecen, su secuencia tipo y el tipo de gen *ESBL* que contienen). Globalmente se ha podido constatar que existe una gran concordancia entre el análisis realizado por PCR y la detección *in silico* (mediante BLASTN, con un *cutoff* del 90% a nivel de identidad y 80% a nivel de cobertura) de los genes *hha2/hha3*.

##### **4.1.2.3 Distribución de los alelos *hha2/hha3* entre la colección de cepas de *E. coli* ECOR**

Además de analizar la presencia de los alelos *hha2* y *hha3* en una gran variedad de cepas de *E. coli* patógenas, también se quiso analizar la prevalencia de estos parálogos en cepas comensales. Para ello, se utilizó la colección ECOR, un conjunto de 72 cepas salvajes de *E. coli* aisladas entre 1973 y 1983 a partir de una gran variedad de hospedadores animales localizados en diferentes áreas geográficas (Ochman y Selander, 1984a). Esta colección representa ampliamente la gran variabilidad genotípica presente en *E. coli* (Ochman y Selander, 1984b). Aunque inicialmente se afirmó que ninguna de las cepas ECOR era patógena (Milkman y McKane, 1995), varias publicaciones posteriores han inferido la presencia de varias cepas virulentas entre las cepas ECOR en base a su MLEE (*multilocus enzyme electrophoresis*) (Pupo y col., 1997). Además, entre varias cepas ECOR se identificaron varios determinantes de virulencia propios de las cepas uropatógenas, como *pap*, *hly*, *kps* y *sfa*, aunque se desconocía si dichos genes eran activos o no (Marklund y col., 1992). Posteriormente se estableció que 29 de las 72 cepas de la colección ECOR presentaban el *status* de cepas ExPEC, ya que dichas cepas codifican por dos o más de los siguientes genes de virulencia: *pap*, *sfa/foc*, *afa*, *aer* y *kps* (Clermont y col., 2011).

Tal y como se puede observar en la figura 4.6, la presencia de los alelos *hha2* y *hha3* en el global de las cepas ECOR es más prevalente en las cepas del patotipo ExPEC que en las cepas comensales no virulentas. En la tabla 9.6 del Anexo se representan de manera específica los resultados de los análisis por PCR realizados a cada una de las cepas de la colección ECOR.

Patotipo	Número de cepas analizadas	Número de cepas positivas para <i>hha2</i>	Número de cepas positivas para <i>hha3</i>
Comensal	43	5 (11%)	4 (9%)
ExPEC	29	5 (17%)	7 (24%)



**Figura 4.6.** Análisis por PCR de la distribución de los genes *hha2* y *hha3* entre las 72 cepas de la colección ECOR.

#### 4.1.2.4 Distribución de los genes *hha2/hha3* entre cepas ambientales de *E. coli*

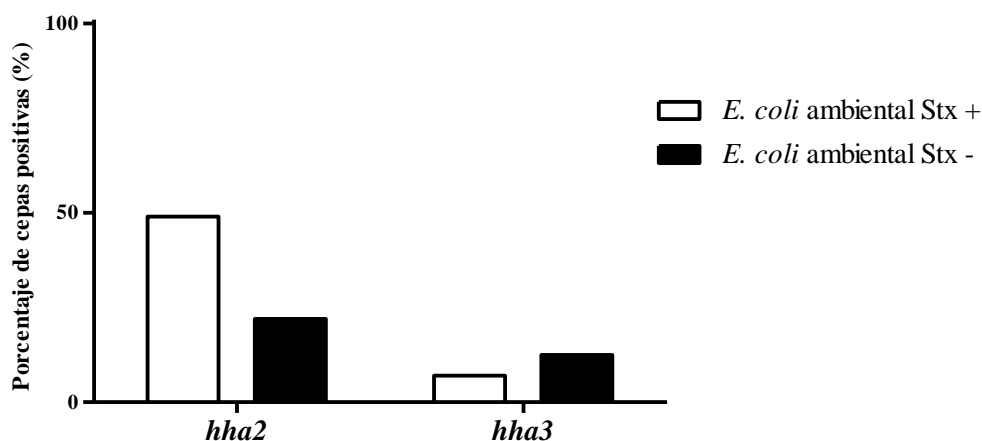
Finalmente, con el objetivo de obtener una visión lo más amplia posible sobre la distribución de ambos parálogos del gen *hha* en *E. coli*, se decidió investigar la presencia de estos genes en el genoma de cepas ambientales de esta especie bacteriana. Para realizar este estudio, se utilizó una colección de 172 cepas de *E. coli* de origen ambiental, de las cuales un total de 84 aislamientos correspondían a cepas que codifican el gen *stx2*. Concretamente, estas cepas se aislaron de muestras de aguas residuales sin tratar de origen urbano (en su mayoría contaminadas por desechos fecales humanos) o de muestras de aguas residuales de tres mataderos diferentes (ganado, cerdo y un matadero mixto de ganado vacuno, cordero y cabra). Las cepas se aislaron de las muestras de agua sólo en base a la presencia de *stx2*. Las muestras se aislaron siguiendo el protocolo descrito por (Blanch y col., 2003) en colaboración con el grupo dirigido por la Dra. Maite Muniesa. En la tabla 9.7 del Anexo se describen las diferentes cepas analizadas, indicando su origen, serotipo, y la presencia o ausencia de la toxina *stx1* (además del tipo de gen *stx2* que presentan todas las cepas aisladas).

El resto de las cepas analizadas en este estudio (88), correspondían a cepas de *E. coli* de origen ambiental no productoras de la toxina Stx. En la figura 4.7 se representan los

#### 4. RESULTADOS

resultados de los análisis por PCR realizados para analizar la presencia de los genes *hha2/hha3* en este conjunto de cepas ambientales.

Producción de toxina Stx	Número de cepas analizadas	Número de cepas positivas para <i>hha2</i>	Número de cepas positivas para <i>hha3</i>
+	84	41 (49%)	6 (7%)
-	88	19 (22%)	11 (12,5%)



**Figura 4.7.** Análisis por PCR de la distribución de los alelos *hha2/hha3* en cepas ambientales de *E. coli* en función de si las cepas son productoras de la toxina Stx2 o no (Stx+/-).

A pesar de su origen ambiental, existe un sesgo significativo en la distribución de los alelos tipo *hha* entre ambos conjuntos de cepas de *E. coli*. Globalmente, mientras que el gen *hha2* se encuentra presente en 41 de las 84 cepas analizadas productoras de la toxina Stx2 (49%), el porcentaje se reduce (22%) cuando se analiza la presencia de este parálogo en cepas no productoras de la toxina Shiga. En el caso del gen *hha3*, la presencia de este parálogo en cepas productoras de Stx es inferior al de *hha2* (7%), mientras que en cepas no productoras de la toxina Shiga el porcentaje de la presencia del alelo *hha3* es ligeramente superior (12,5%).

De manera más específica, el gen *hha2* se ha encontrado presente de manera individual en 35 de las 84 cepas de *E. coli* productoras de Stx (41,6%), mientras que el gen *hha3* por sí solo no estaba presente de manera individual en ninguna de las cepas analizadas. Ambos alelos se hallaron presentes de manera conjunta en 6 de las 84 cepas testadas (7,1%). Por el contrario, al considerar los 88 aislamientos ambientales no productores de la toxina Stx, el alelo *hha2* solo estaba presente de manera individual en diez de las cepas estudiadas (11,3%). Estos resultados muestran unas diferencias significativas para la presencia del alelo *hha2* en las cepas productoras de la toxina Stx2 respecto las que no

producen dicha toxina (test exacto de Fisher,  $P = 0,000019$ ). En este último caso, el gen *hha3* solo se encontraba presente de manera individual en 2 cepas (2,3%) mientras que la combinación de ambos alelos resultó positiva para 9 de las 88 cepas analizadas (10,2%) (tabla 9.7 del Anexo).

### 4.1.3 Análisis genómico de la distribución de los genes *hha2* y *hha3* en *E. coli* mediante herramientas bioinformáticas

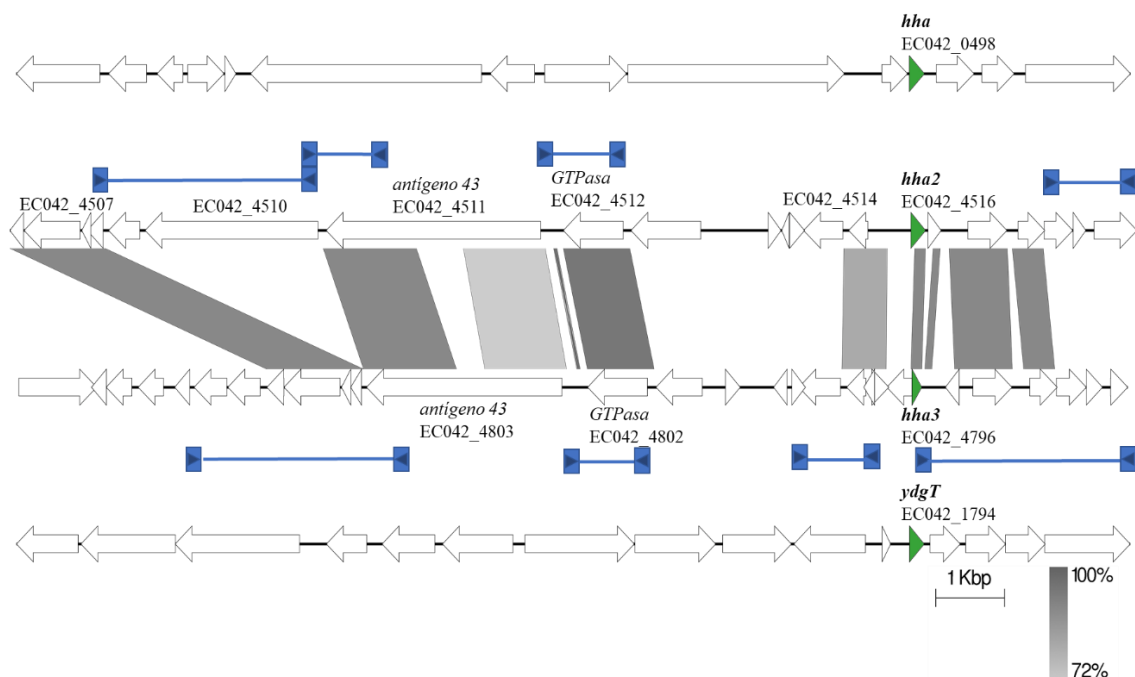
La parte final de este estudio de genómica realizado con el fin de analizar la distribución de los genes *hha2* y *hha3* consistió en realizar una búsqueda mediante BLASTN (90% de identidad y 80% de cobertura) con la secuencia de los genes *hha2* y *hha3* sobre todas las cepas bacterianas secuenciadas hasta la fecha. Ello ha permitido comprobar como la presencia de estos dos parálogos queda limitada principalmente a la especie *Escherichia coli*, siendo totalmente nula su presencia en el género *Salmonella*. Ambos parálogos se identificaron principalmente en cepas patógenas de *E. coli* a excepción de las cepas comensales SE15, O9H5 y ED1a (esta última cepa contiene una delección en el gen *hha2* justo en el lugar de hibridación de los oligonucleótidos utilizados para detectar dicho gen (figura 9.2 del Anexo)).

Comparando entre diferentes cepas, mientras que la posición de los genes *hha* e *ydgT* se mantiene relativamente constante entre todos los genomas (tal y como era de esperar para genes del núcleo central del genoma (*core genome*)) (figura 9.8 del Anexo), la posición de los genes *hha2* y *hha3* varía mucho entre las diferentes cepas secuenciadas, y los mismos siempre se encuentran flanqueados por secuencias de inserción. Analizando de manera específica la localización de los genes *hha*, *ydgT*, *hha2* y *hha3* en el cromosoma de *E. coli* 042, mediante un análisis de comparación lineal de genomas se compararon las regiones aguas arriba y aguas abajo de los bloques genómicos que contienen los 4 genes tipo *hha* (resultados mostrados en la figura 4.8).

En dicha figura se puede observar cómo los entornos génicos de los genes *hha2* y *hha3* son muy similares entre ellos, pero son totalmente diferentes a los entornos génicos de los genes *hha* e *ydgT* (entornos que a su vez también son muy diferentes entre sí). A nivel de organización génica, tanto el parálogo *hha2* como el parálogo *hha3* se encuentran asociados a un bloque de genes comunes (flanqueados por varias secuencias de inserción y transposasas) entre los que destaca el gen *agn43* (*antígeno 43*). Dicho gen es un importante factor de virulencia de *E. coli* y codifica para una proteína de membrana

#### 4. RESULTADOS

externa implicada en el contacto célula con célula muy importante para la formación de *biofilm* (Danese y col., 2000). De manera específica, el gen *hha2* se localiza en una región génica en la que se encuentra la copia del gen *agn43* EC042\_4511, mientras que el gen *hha3* se encuentra asociado a la copia *agn43* EC042\_4803. Curiosamente, el genoma de *E. coli* 042 codifica para otra copia del gen *agn43* (EC042\_2242) (Chaudhuri y col., 2010). Sorprendentemente, la tercera copia de este factor de virulencia también se encuentra asociada a una copia de un pseudogen tipo *hha*, que en las bases de datos no aparece como una pauta abierta de lectura ya que parece ser una forma truncada del gen *hha2*.



**Figura 4.8.** Organización de las diferentes regiones del genoma de *E. coli* 042 en las que se encuentran los genes *hha*, *hha2*, *hha3* e *ydgT*, representados de color verde. En azul se representan todas las secuencias de inserción presentes en dicha región. Se muestran también las regiones de homología con un porcentaje de identidad superior al 72%. Alineamiento de genómica comparativa realizado mediante el programa *EasyFig* (Sullivan y col., 2011) haciendo uso de las secuencias génicas correspondientes al cromosoma de *E. coli* 042 (FN554766.1) extraídas de *GenBank*.

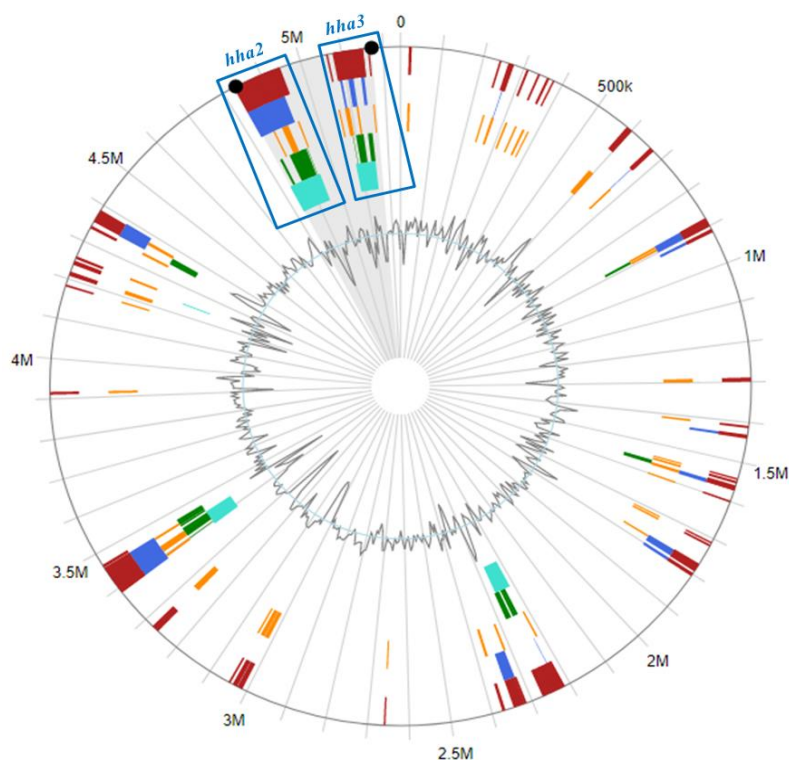
El hecho de que las regiones en las que se localizan los genes *hha2* y *hha3* estén enriquecidas en secuencias de inserción de transposasas y elementos móviles puede ser un elemento indicador de que estos bloques de genes, más que pertenecer al *core* central del genoma, hayan sido adquiridos por la bacteria a lo largo de la historia evolutiva mediante mecanismos de transferencia génica horizontal (HGT). Por ello, mediante el *software* Island Viewer 4 (<https://www.pathogenomics.sfu.ca/islandviewer/>) una interfaz



que a través de 5 métodos de predicción distintos (*Integrated*, *IslandPick*, *IslandPath-DIMOB*, *SIGI-HMM* e *Islander*) permite predecir si una determinada región del genoma bacteriano pertenece a una isla de patogenicidad (Bertelli y col., 2017), se analizó si la localización de los genes *hha2* y *hha3* concurría en zonas definidas como islas genómicas de patogenicidad (PAI).

Las islas de patogenicidad son segmentos de ADN bacteriano que portan uno o más genes de virulencia, los cuales han sido adquiridos en bloque a partir de una fuente externa. Estos segmentos génicos heredados en bloque están presentes en las cepas patógenas y se encuentran ausentes (o distribuidos esporádicamente) en las cepas comensales de una misma especie o especies relacionadas. Dichos fragmentos son unidades compactas que ocupan una gran región del cromosoma bacteriano y están asociados a genes de ARNt y/o secuencias de inserción, por lo que sus extremos suelen estar frecuentemente flanqueados por repeticiones directas, y presentan además genes correspondientes a elementos móviles tales como transposasas, elementos de inserción e integrasas (Gal-Mor y Finlay, 2006). Los principales determinantes de virulencia codificados en las PAI son sistemas de secreción, factores de virulencia implicados en la evasión del sistema inmune del hospedador, factores de adherencia y colonización, además de toxinas y genes implicados en la adquisición de nutrientes (Schmidt y Hensel, 2004).

En la figura 4.9 se representan las diferentes islas de patogenicidad identificadas bioinformáticamente en el cromosoma de la cepa *E. coli* 042 mediante el programa predictor *Island Viewer 4*. Puede comprobarse como los genes *hha2* y *hha3* se localizan en dos islas de patogenicidad diferentes (aunque muy próximas entre ellas en el genoma).



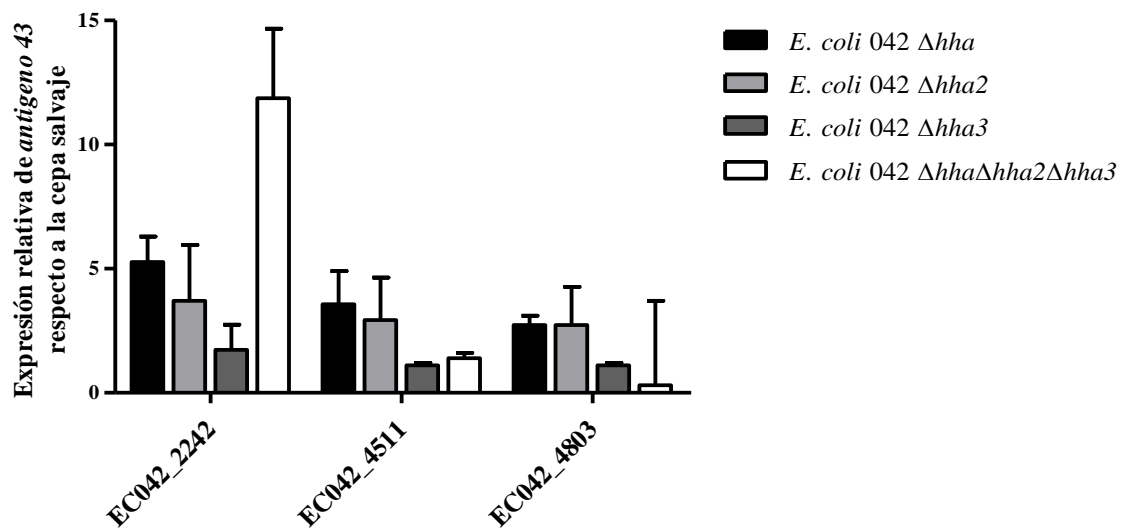
**Figura 4.9.** Representación esquemática de las diferentes islas de patogenicidad distribuidas a lo largo del genoma de la cepa *E. coli* 042. Recuadradas de color azul se muestran las islas de patogenicidad en las que se hallan los genes *hha2* y *hha3*. Los cuadros internos de color rojo, azul, amarillo, verde y azul claro indican las diferentes predicciones realizadas por diferentes programas (*Integrated*, *IslandPick*, *IslandPath-DIMOB*, *SIGI-HMM* e *Islander*) englobados en el software global *Island Viewer 4*.

El gen *hha2* se encuentra asociado a una isla de patogenicidad de gran tamaño (117 Kb) en la que se encuentra un sistema de secreción tipo VI, mientras que el gen *hha3* se localiza en una pequeña isla de patogenicidad (6 Kb) muy próxima a una isla mayor (72 Kb) en la que se encuentran genes implicados en la metabolización del ascorbato, el operón *fecBCD* (implicado en la metabolización del hierro) y reguladores transcripcionales de la familia TetR.

#### 4.1.4 Papel regulador de las proteínas Hha, Hha2 y Hha3 sobre los genes *agn43* y *shf*

El análisis genómico comparativo indicado en el apartado anterior concuerda con la correlación mostrada en la figura 4.4 para la presencia conjunta de los parálogos *hha2/hha3* junto con el gen *agn43*. Por ello, dada la coexistencia de estos genes, se planteó la hipótesis de si los genes tipo *hha* podrían ejercer un papel regulador sobre la expresión de este factor de virulencia. Con este fin, se construyeron mutantes isogénicos sencillos para los genes *hha*, *hha2* y *hha3* y un triple mutante *hha*-nulo (cepa que combina

las tres mutaciones de los parálogos tipo *hha*) y se cuantificó por qRT-PCR el efecto de estas mutaciones sobre la expresión de las tres copias de los genes *agn43* codificadas en el genoma de la cepa *E. coli* 042. Para obtener estos datos, los cultivos correspondientes se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento (D.O<sub>600nm</sub> 2,0), se extrajo el ARN y se cuantificó el nivel de expresión del ARNm de las tres copias del gen *agn43* en los 4 mutantes analizados respecto el nivel de expresión de dichos genes en la cepa salvaje de *E. coli* 042 (figura 4.10).



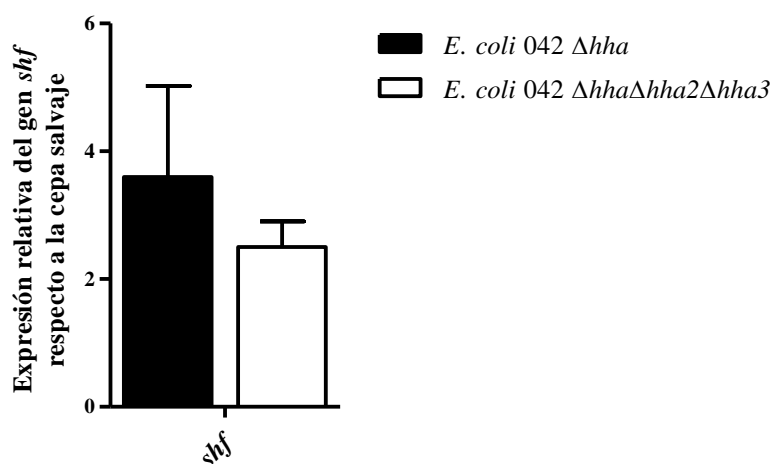
**Figura 4.10.** Expresión relativa de las tres copias del gen *agn43* presentes en el cromosoma de la cepa *E. coli* 042 en los mutantes *hha*, *hha2*, *hha3* y el triple mutante (*hha**hha2hha3*) respecto a la cepa *E. coli* 042 salvaje. La expresión de los genes se analizó por qRT-PCR, utilizando los oligonucleótidos 2242 RT Fw/Rv, 4511 RT Fw/Rv, 4803 RT Fw/Rv. Los resultados fueron normalizados mediante la detección del gen interno *gapA* mediante los oligonucleótidos *gapA*042 RT Fw/Rv. Se representan los valores correspondientes a 3 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Se ha podido detectar un importante efecto de la mutación en los genes *hha* y *hha2* sobre la expresión de las tres copias del gen *antígeno 43* (principalmente sobre las copias EC042\_2242 y EC042\_4511, siendo la mutación sobre el gen *hha* la que ejerce un efecto más pronunciado), mientras que la mutación del gen *hha3* apenas tiene efecto sobre la expresión de este factor de virulencia. De manera notoria, se ha podido observar una gran desregulación en la expresión de la copia de EC042\_2242 cuando se combinan las tres mutaciones en los genes *hha* (un incremento de la transcripción 12 veces superior al del nivel presente en la cepa salvaje). Este efecto sumatorio de las mutaciones combinadas

#### 4. RESULTADOS

de los genes tipo *hha* únicamente se observa en esta copia, curiosamente la única copia del gen *agn43* que no lleva asociada ningún gen tipo *hha* completo.

También se analizó el efecto de los genes tipo *hha* sobre la expresión de otro factor de virulencia del patotipo enteroagregativo de *E. coli* cuya presencia mostraba una asociación significativa con los parálogos *hha2* y *hha3*, el gen *shf*. Dicho gen, descrito inicialmente en *Shigella flexneri*, juega un papel muy importante en la modificación de polisacáridos necesarios para la formación de *biofilm* por parte de la cepa *E. coli* 042 (Fujiyama y col., 2008). A diferencia del gen *antígeno 43*, la cepa *E. coli* 042 únicamente contiene una única copia del gen *shf*. En este caso, se analizó por qRT-PCR el nivel de expresión del gen *shf* en un mutante *hha* sencillo y en el triple mutante *hha*, *hha2* y *hha3* en condiciones de crecimiento en medio rico (LB) a 37°C al inicio de la fase estacionaria de crecimiento ( $D_{600nm}$  2,0) de los correspondientes cultivos bacterianos (figura 4.11).



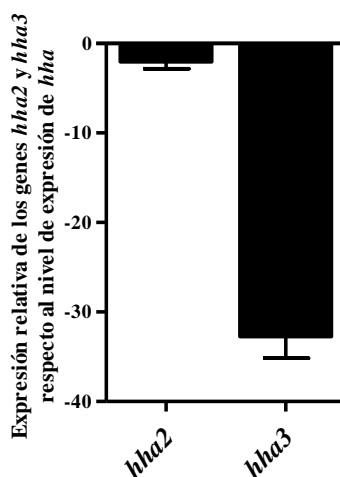
**Figura 4.11.** Expresión transcripcional del gen *shf* en los mutantes *hha* y *hha2hha3*. La expresión del gen *shf* se evaluó por qRT-PCR utilizando los oligonucleótidos *shf* RT Fw/Rv, y se representa de manera relativa al nivel de expresión del gen en la cepa *E. coli* 042 salvaje. Los resultados fueron normalizados mediante la detección del gen interno *gapA* haciendo uso de los oligonucleótidos *gapA042* RT Fw/Rv. Se representan los valores correspondientes a 3 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Analizando los resultados obtenidos, la sobreexpresión observada del gen *shf* en ausencia de la proteína Hha indica el papel represor que esta proteína ejerce sobre la expresión de este factor de virulencia. Estos resultados convergen con los resultados mostrados previamente sobre el efecto de Hha a nivel de la expresión de *antígeno 43*. En cambio, a

diferencia de lo observado con la copia EC042\_2242 de *antígeno 43*, los parálogos *hha2* y *hha3* no parecen tener un papel modulador en la expresión del gen *shf*.

#### 4.1.5 Perfil de expresión de los genes *hha*, *hha2* y *hha3*

Ya que el nivel de expresión de algunos de los moduladores globales de la expresión génica en bacterias es crítico para su función, se decidió analizar el nivel de expresión de los propios parálogos de *hha* en la cepa *E. coli* 042. Para ello, se realizaron análisis por qRT-PCR con la finalidad de cuantificar el nivel de expresión de los parálogos *hha2* y *hha3* (en función del nivel de expresión de *hha*) en medio LB a 37°C al inicio de la fase estacionaria de crecimiento (D.O<sub>600nm</sub> 2,0) (figura 4.12).



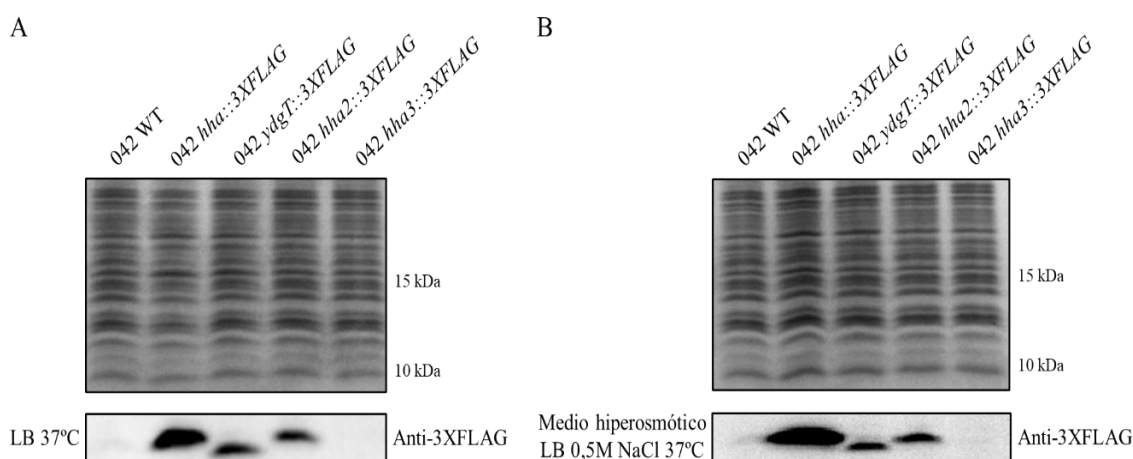
**Figura 4.12.** Expresión transcripcional de los genes *hha2* y *hha3* representada de manera relativa al nivel de expresión del gen *hha* en la cepa *E. coli* 042 salvaje. La expresión de los genes se evaluó por qRT-PCR utilizando la serie de oligonucleótidos *hha*, 4516 y 4796 *forward/reverse*. Los resultados fueron normalizados mediante la detección del gen interno de referencia *gapA* haciendo uso de los oligonucleótidos *gapA042 RT Fw/Rv*. Se representan los valores correspondientes a 3 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Mientras que en medio de crecimiento rico a 37°C la expresión del gen *hha2* es ligeramente inferior a la del gen *hha* (2 veces menor), el nivel de transcripción del gen *hha3* es notablemente menor al de su parálogo (concretamente 32 veces inferior).

Paralelamente, y con el fin de poder obtener un perfil de expresión proteica de las proteínas Hha, Hha2, Hha3 e YdgT, se construyeron las fusiones cromosómicas de los genes mencionados anteriormente con el epítipo 3XFLAG con el fin de inmunodetectar la expresión de dichas proteínas con un anticuerpo específico dirigido contra dicho

#### 4. RESULTADOS

epítipo añadido mediante técnicas de ingeniería genética (siguiendo el protocolo descrito en el apartado 3.7.8 de Materiales y Métodos). Con esta finalidad, se crecieron las cepas *E. coli* 042 *hha::3XFLAG*, *hha2::3XFLAG*, *hha3::3XFLAG* e *ydgT::3XFLAG* en medio rico LB y medio hiperosmótico (medio LB con una concentración de 0,5 M NaCl, condiciones en las que se induce una mayor expresión de la proteína Hha ((Mouriño y col., 1998)) y se obtuvieron extractos proteicos totales de los cultivos crecidos hasta el inicio de la fase estacionaria de crecimiento ( $D.O_{600nm}$  2,0) a 37°C. Para comprobar el nivel de expresión de las proteínas Hha::3XFLAG, Hha2::3XFLAG, Hha3::3XFLAG e YdgT::3XFLAG en las diferentes condiciones analizadas, se realizó una inmunodetección (mediante *Western Blot*) contra dichos extractos utilizando un anticuerpo cuya diana es la secuencia aminoacídica del epítipo 3XFLAG (figura 4.13).



**Figura 4.13.** Análisis por *Western Blot* de la expresión de las proteínas Hha::3XFLAG, Hha2::3XFLAG, Hha3::3XFLAG e YdgT::3XFLAG en medio rico LB (A) y LB 0,5 M NaCl (B) a 37°C en el inicio de la fase estacionaria de crecimiento ( $D.O_{600nm}$  2,0). El panel superior muestra la tinción de Coomassie de cada uno de los extractos proteicos analizados y representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas.

Globalmente, se puede observar como la expresión de la proteína Hha es superior a la de sus otros parálogos Hha2 e YdgT, (que muestran unos niveles de expresión similares), mientras que la expresión proteica de la proteína Hha3 es prácticamente indetectable. Además, tal y como se describió por parte de (Mouriño y col., 1998), la expresión proteica de Hha se encuentra estimulada cuando las bacterias crecen en medio LB con estrés hiperosmótico (0,5 M de NaCl), mientras que la expresión de los otros dos parálogos se mantiene constante sin verse afectada por la osmolaridad del medio. De manera análoga

a los resultados observados en medio LB, la expresión de la proteína Hha3 sigue siendo muy baja en estas condiciones.

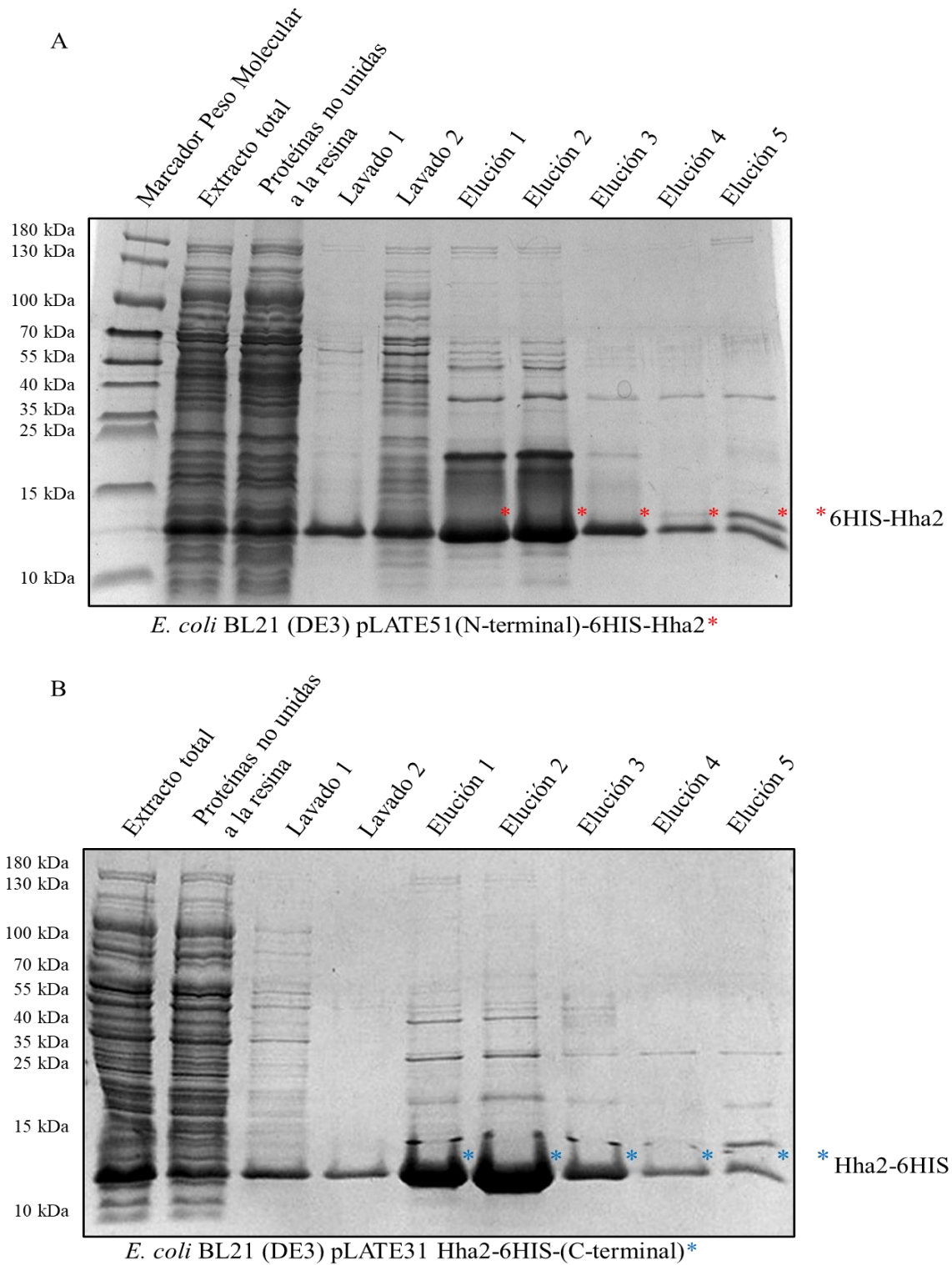
#### **4.1.6 Análisis del perfil de interacción proteica de Hha2 con otras proteínas de la cepa *E. coli* 042**

Ya que la proteína Hha es capaz de interactuar con la proteína H-NS formando heterocomplejos que presentan una gran capacidad para regular eficientemente genes adquiridos por transferencia génica horizontal (Nieto y col., 2001; Baños y col., 2009; Ali y col., 2013), se decidió investigar con qué proteínas son capaces de interactuar los parálogos tipo Hha codificados en el cromosoma de la cepa *E. coli* 042. Sin embargo, debido a que el nivel de expresión del gen *hha2* es muy superior al del gen *hha3* (cuyo nivel de expresión es ínfimo) (figuras 4.12 y 4.13)), se decidió analizar únicamente las posibles parejas de interacción de la proteína Hha2. Además, a diferencia de las proteínas YdgT y Hha3, la proteína Hha2 sí contiene los 4 aminoácidos altamente conservados entre las proteínas Hha (E25, D48, E53, D61) (figura 4.1), de los cuales los residuos E25 y D48 son responsables de la interacción con H-NS (De Alba y col., 2011).

Para realizar la sobreexpresión y posterior purificación de la proteína Hha2 se utilizaron los plásmidos pLATE31 Hha2-6HIS y pLATE51 6HIS-Hha2, que permiten introducir una cola de 6 histidinas en los extremos C y N-terminal (respectivamente) del gen clonado. Dichos plásmidos se introdujeron en la cepa *E. coli* BL21 (DE3), donde se sobreexpresaron las correspondientes proteínas 6His-Hha2 (N-terminal) y Hha2-6His (C-terminal) bajo el control del promotor de la ARN polimerasa T7 de manera inducible por IPTG. Los extractos proteicos obtenidos se mezclaron con una matriz de Ni<sup>2+</sup>-NTA agarosa donde, tras permitir la unión de las proteínas y lavar las proteínas no unidas, se eluyeron las proteínas retenidas en la matriz mediante la adición de un tampón con concentraciones crecientes de imidazol (200 mM, 500 mM y 1M). El imidazol es un compuesto intermediario en la biosíntesis de la histidina que compite contra las propias histidinas por la unión al níquel<sup>2+</sup>, desplazando de esta manera las proteínas unidas a la resina. Las diferentes fracciones obtenidas en el proceso de purificación se separaron mediante electroforesis en geles desnaturizantes de poliacrilamida en Tris-Tricina-SDS, en los que las proteínas se tiñeron posteriormente con Azul de Coomassie (figura 4.14).

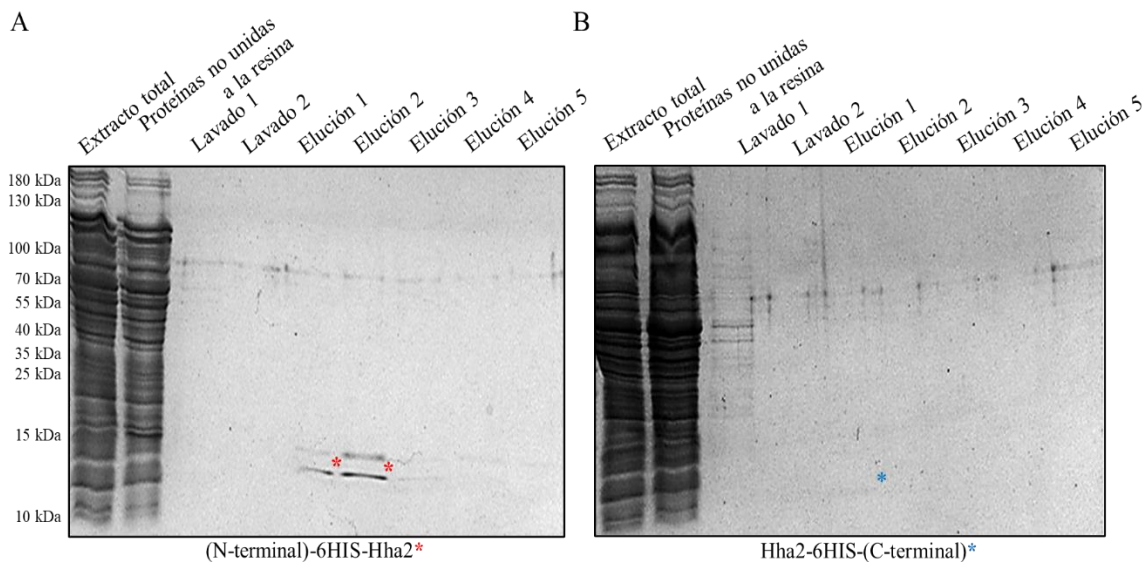


#### 4. RESULTADOS



**Figura 4.14.** Electroforesis en gel de poliacrilamida de las diferentes fracciones proteicas obtenidas en el proceso de purificación de la proteína Hha2 fusionada con un epítipo de 6 histidinas en sus extremos N-terminal (A) y C-terminal (B). El proceso de sobreexpresión proteica se realizó en la cepa *E. coli* BL21 (DE3). Se muestran las fracciones correspondientes al extracto total de proteínas de la célula, las proteínas no unidas a la membrana después de su interacción con la resina de  $\text{Ni}^{2+}$ -agarosa, dos lavados de la resina con imidazol 50 mM, y las eluciones proteicas correspondientes (1-3 se realizaron con imidazol 200 mM, la 4ª elución con 500 mM y la 5ª elución con 1 M de imidazol).

En la fracción correspondiente al extracto proteico total de la cepa *E. coli* BL21 (DE3) conteniendo los plásmidos pLATE51 6HIS-Hha2 y pLATE31 Hha2-6HIS se pueden observar las bandas proteicas relativas a la sobreexpresión de la proteína Hha2 (la proteína muestra un tamaño mayor debido a la adición de la cola de 6 histidinas). En las fracciones correspondientes a la elución de la proteína unida a la resina de níquel, además de la proteína Hha2 purificada (proteína más abundante de la fracción), se pueden observar otras bandas proteicas correspondientes a proteínas de la cepa *E. coli* BL21 (DE3) que han coeluido con la propia proteína Hha2. Sin embargo, ya que el objetivo era poder definir con qué proteínas de la cepa *E. coli* 042 es capaz de interactuar la proteína Hha2, las proteínas eluidas y purificadas con imidazol 200 mM (3 eluciones en total, marcadas con un asterisco en la figura 4.14), se mezclaron con un extracto total de las proteínas de *E. coli* 042, se dejaron interactuar durante 16 horas a 4°C y se purificaron de manera análoga al procedimiento anterior mediante una cromatografía de afinidad con una matriz de Ni<sup>2+</sup>-NTA agarosa. Las diferentes fracciones obtenidas en este nuevo proceso de purificación se separaron mediante electroforesis en geles desnaturalizantes de poliacrilamida en Tris-Tricina-SDS y las proteínas se detectaron mediante tinción con Azul de Coomassie (figura 4.15).



**Figura 4.15.** Electroforesis en gel de poliacrilamida de la purificación de la proteína Hha2 junto con las proteínas de *E. coli* 042. Las proteínas Hha2 (con las 6 histidinas en sus extremos N-terminal (A) o C-terminal (B)) se dejaron interactuar con un extracto proteico total de la cepa *E. coli* 042 y se purificaron mediante cromatografía de afinidad. Se muestran las fracciones correspondientes al extracto total de proteínas de la célula, las proteínas no unidas a la membrana después de su interacción con la resina de Ni<sup>2+</sup>-agarosa, dos lavados de la resina con imidazol 50 mM, y las eluciones proteicas correspondientes (1-3 con imidazol 200 mM, la 4ª elución con 500 mM y la 5ª elución con 1 M de imidazol).

#### 4. RESULTADOS

Con el fin de identificar qué proteínas de *E. coli* 042 habían copurificado con la proteína Hha2 purificada previamente, las fracciones correspondientes a las eluciones 1 y 2 de cada purificación se enviaron al servicio de la “Plataforma de Proteómica del Parc Científic de Barcelona”. Las proteínas presentes en cada una de las eluciones se digirieron con tripsina; los péptidos resultantes se purificaron mediante cromatografía líquida y posterior espectrometría de masas. Mediante los *software* *Thermo Proteome Discoverer* (v.1.4.1.14) y *Sequest HT* se comparó el espectro de masas de cada uno de los péptidos obtenidos para identificar las proteínas correspondientes en la base de datos *Swiss Prot*.

En la tabla 4.2 se indican las proteínas codificadas en el genoma de *E. coli* 042 que han coeluido tanto con la proteína 6HIS-Hha2 como con la proteína Hha2-6HIS durante el proceso de purificación.

Proteínas copurificadas con Hha2	Score*	PSM**	Coverage (%) ***
Proteína de unión al ADN H-NS	2210,63	732	81,75
(N-Terminal)-6HIS-Hha2*	1019,18	465	70,41
Hha2-6HIS-(C-Terminal)*	991,46	402	81,63
Proteína de unión al ADN StpA	330,94	113	63,43
Subunidad $\beta$ de la ARN polimerasa RpoB	284,55	87	60,28
ADN polimerasa I PolA	204,35	60	57,00
Proteína de unión al ADN HU-alpha HupA	155,25	59	46,67
Componente E12-oxoglutarato deshidrogenasa SucA	151,87	52	55,84
Subunidad $\alpha$ de la ARN polimerasa RpoA	148,73	51	82,07
ADN topoisomerasa 1 TopA	120,29	42	50,64
Triptofanasa TnaA	108,43	33	73,46
Poliribonucleotido nucleotidiltransferasa Pnp	107,48	32	50,21
Ribonucleasa G Rng	101,52	30	61,15
Factor de elongación TufA	68,93	24	60,15
Factor de elongación G FusA	65,47	22	40,91
Proteína putativa de unión al ADN H-NS2	65,40	21	55,56
Factor sigma de la ARN polimerasa RpoD	54,71	18	32,79
Proteína de protección del ADN DpsP	41,09	17	71,26
Proteína asociada a la ARN polimerasa RapA	39,39	16	19,63
Factor de terminación/anti-terminación NusG	37,54	15	58,01
Aminopeptidasa citosólica PepA	37,65	15	37,57
Exdesoxiribonucleasa I SbcB	35,27	14	36,00
Proteína putativa YcfP	35,11	14	57,78
Subunidad C alquil hidroperóxido reductasa AhpC	32,46	12	57,75

\*Abundancia relativa de las proteínas presentes en la muestra.

\*\**Peptide Spectrum Match*: Parámetro que muestra el número total de secuencias peptídicas identificadas para cada proteína.

\*\*\* Indicador que muestra el porcentaje de la secuencia de la proteína cubierta por los péptidos identificados por LC/MS (cromatografía líquida/espectrometría de masas).

**Tabla 4.2.** Resultados de la identificación por espectrometría de masas de las proteínas de la cepa *E. coli* 042 que interaccionan con la proteína Hha2. Se representan globalmente los resultados correspondientes a todas las proteínas que han coeluido tanto con la proteína 6HIS-Hha2 como con la proteína Hha2-6HIS.

Tal y como era de esperar, de manera análoga a lo descrito para la proteína Hha debido al elevado grado de similitud estructural entre las dos proteínas, la proteína Hha2 presenta la capacidad de interactuar con un gran número de proteínas, entre las que destacan H-NS y sus parálogos StpA y H-NS2 (proteína no descrita hasta la fecha y objeto de estudio en el siguiente capítulo de la tesis doctoral). Además, la proteína Hha2 también es capaz de interactuar con la subunidad  $\alpha$  de la proteína asociada al nucleóide HU.

#### **4.1.7 Efectos de las mutaciones en los genes *hha* y *hha2* sobre el perfil transcriptómico global de la cepa *E. coli* 042**

Para tratar de profundizar en el papel regulador de la expresión génica ejercido por las proteínas Hha y Hha2 en la cepa *E. coli* 042, se realizó un análisis transcriptómico comparativo por RNA-seq de las cepas *E. coli* 042  $\Delta hha$ , *E. coli* 042  $\Delta hha2$ , *E. coli* 042  $\Delta hha\Delta hha2$  y se compararon con el transcriptoma de la cepa *E. coli* 042 salvaje. Estos ensayos transcriptómicos se realizaron a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento (D.O<sub>600nm</sub> 2,0). El aislamiento del ARN y el procesamiento de las muestras para obtener los datos relativos a la expresión de cada gen se realizaron en colaboración con la empresa *Vertis Biotechnologie AG* (*Freising-Weihenstephan*, Alemania), siguiendo el protocolo descrito en el apartado 3.8.5 de Materiales y Métodos.

Los resultados de este experimento se encuentran publicados en su totalidad en la base de datos GEO (*Gene Expression Omnibus*) *Sequence Read Archive* del NCBI bajo el número de acceso GSE105133, además de en el apartado 9.1 del Anexo de esta memoria, en las tablas 9.9 y 9.10.

En la tabla 4.3 se indican los genes más sobreexpresados (con respecto a la cepa salvaje) en los mutantes  $\Delta hha$  (A),  $\Delta hha2$  (B) y  $\Delta hha\Delta hha2$  (C), respectivamente. Además del nivel de expresión diferencial (*foldchange*) específico para cada gen en cada mutante, se indican también los valores de expresión relativa para dicho gen en el resto de mutantes analizados, a los que hay que sumar también el mutante *hns*. Ello es así porque está asumido de forma general que los genes regulados por la proteína Hha también se encuentran regulados por la proteína H-NS.

#### 4. RESULTADOS

Identificador NCBI ( <i>locus tag</i> )	Fold change			
	$\Delta hha$	$\Delta hha2$	$\Delta hha\Delta hha2$	$\Delta hns$
EC042_RS06355 (hypothetical protein)	26,0	2,0	4,8	7,1
EC042_RS01250 (hypothetical protein)	20,0	2,0	7,8	7,8
EC042_RS20085 (hypothetical protein)	14,2	-	-	-
EC042_RS03935 (hypothetical protein)	13,9	2,3	4,1	3,9
EC042_RS14470 (hypothetical protein)	13,8	4,4	2,0	-
EC042_RS19170 (hypothetical protein)	13,6	-	-	-
EC042_RS21810 (broad-spectrum mercury transporter MerE)	12,2	2,0	-	5,0
EC042_RS09940 (hypothetical protein)	11,7	-	5,0	3,9
EC042_RS24485 (transposase)	11,5	-	-	-
EC042_RS15485 (hypothetical protein)	11,0	-	-	-
EC042_RS27060 (nitrite extrusion protein 2)	10,6	-	-	-
EC042_RS16305 (hypothetical protein)	10,5	-	4,0	3
EC042_RS06730 (hypothetical protein)	10,4	-	-	-
EC042_RS01500 (hypothetical protein)	10,2	-	2,9	3,7
EC042_RS03035 (hypothetical protein)	10,1	-	2,9	3,9
EC042_RS03040 (hypothetical protein)	10,1	-	-	4
EC042_RS25285 (tRNA-Leu)	10,0	-	-	-
EC042_RS08665 (transcriptional regulator)	10,0	-	4,4	4,4
EC042_RS25735 (hypothetical protein)	10,0	-	3,0	5
EC042_RS21265 (hypothetical protein)	9,9	2,0	4,5	4,6
EC042_RS12550 (hypothetical protein)	9,8	-	2,6	5,9
EC042_RS17190 (hypothetical protein)	9,7	-	3,3	2,2
EC042_RS16360 (invasion protein invH)	9,7	-	5,6	10
EC042_RS16320 (hypothetical protein)	9,6	-	3,3	5,9
EC042_RS09160 (division inhibition protein DicB)	9,5	-	4,3	3,5
EC042_RS14290 (hypothetical protein)	9,5	-	-	4,5
EC042_RS26735 (hypothetical protein)	9,2	-	2,4	3,8
EC042_RS03940 (hypothetical protein)	9,1	-	6,4	5,5
EC042_RS11620 (tRNA-Arg)	8,9	-	3,1	12
EC042_RS16950 (tRNA-Phe)	8,8	2,2	4,1	4,1
EC042_RS23700 (hypothetical protein)	8,8	-	2,6	3,2
EC042_RS01830 (hypothetical protein)	8,8	-	3,7	2,3
EC042_RS20080 (hypothetical protein)	8,7	-	4,9	7,7
EC042_RS26640 (hypothetical protein)	8,7	-	-	4,4
EC042_RS17060 (hypothetical protein)	8,7	-	5,2	3,6
EC042_RS09020 (hypothetical protein)	8,6	-	2,9	5,4
EC042_RS03905 (acyl-CoA synthetase)	8,6	-	3,1	4,0
EC042_RS03695 (hypothetical protein)	8,5	-	2,4	2,3



**Tabla 4.3 (A).** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hha$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. También se indican los valores de expresión relativa de dichos genes en los mutantes  $\Delta hha2$ ,  $\Delta hha\Delta hha2$  y  $\Delta hns$ . Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

Considerando en primer lugar los genes sobreexpresados en el mutante sencillo *hha*, es aparente que la gran mayoría de los mismos también se encuentran sobreexpresados en el mutante *hns*. Sin embargo, algunos de tales genes no siguen el patrón clásico descrito hasta la fecha de correulación Hha/H-NS. Algunos de ellos no aparecen sobreexpresados en ningún otro fondo genético mutante y, en los que sí lo están, es aparente que el efecto de la proteína Hha es superior al efecto que se presenta en un mutante H-NS, o incluso en un doble mutante *hha hha2*. Es posible que tales genes (la gran mayoría de los cuales codifican por proteínas hipotéticas sin función atribuida) formen parte de una categoría ya referida en algunas publicaciones, que corresponde a genes regulados por la proteína Hha sin H-NS, mediante un mecanismo que queda por aclarar (Solórzano y col., 2015).

Identificador NCBI ( <i>locus tag</i> )	<i>Fold change</i>			
	$\Delta hha2$	$\Delta hha$	$\Delta hha\Delta hha2$	$\Delta hns$
EC042_RS02730 (copper-transporting P-type ATPase CopA)	4,9	-	-	-
EC042_RS14470 (hypothetical protein)	4,4	13,8	2,0	11,0
EC042_RS20430 (small toxic polypeptide ldrD)	4,2	-	-	-
EC042_RS14785 (small toxic protein ShoB)	4,0	-	-	-
EC042_RS09080 (type I toxin-antitoxin system toxin HokD)	3,9	-	-	-
EC042_RS20445 (small toxic polypeptide ldrD)	3,7	-	-	-
EC042_RS08185 (type I toxin-antitoxin system toxin HokB)	3,6	-	-	-
EC042_RS20435 (small toxic polypeptide ldrD)	3,4	-	-	-
EC042_RS06685 (hypothetical protein)	3,4	-	-	-
EC042_RS20440 (small toxic polypeptide ldrD)	3,3	-	-	-
EC042_RS03760 (YbfA family protein)	3,3	-	3,1	-
EC042_RS14110 (sulfate ABC transporter substrate-binding protein)	3,2	-	-	-
EC042_RS20450 (small toxic polypeptide ldrD)	3,2	-	-	-
EC042_RS26105 (transcriptional regulator)	3,2	3,4	-	-
EC042_RS07830 (putative prophage protein)	2,9	-	-	-
EC042_RS11655 (type I toxin-antitoxin system Hok family toxin)	2,9	-	-	-
EC042_RS22345 (hypothetical protein)	2,7	-	-	-
EC042_RS14790 (hypothetical protein)	2,6	-	-	-

#### 4. RESULTADOS

EC042_RS21410 (hypothetical protein)	2,6	-	-	-
EC042_RS02600 (Hha toxicity modulator TomB)	2,6	-	2,8	3,9
EC042_RS07040 (type I toxin-antitoxin system Hok family toxin)	2,5	-	-	-
EC042_RS14055 (cysteine synthase A CysK)	2,5	-	-	-
EC042_RS11235 (peroxide/acid resistance protein YodD)	2,5	-	-	-
EC042_RS15765 (sulfate adenylyltransferase subunit CysD)	2,5	-	-	-
EC042_RS15815 (putative post-segregation killing protein)	2,5	-	-	-
EC042_RS18425 (tRNA-Met)	2,5	-	3,6	3,3
EC042_RS06630 (stress-induced protein ychH)	2,5	-	-	-
EC042_RS05510 (heat shock protein HspQ)	2,4	-	-	-
EC042_RS21490 (heat shock chaperone IbpA)	2,4	3,5	-	-
EC042_RS12590 (hypothetical protein)	2,4	-	-	-
EC042_RS26755 (hypothetical protein)	2,3	-	-	-
EC042_RS21895 (F0F1 ATP synthase subunit epsilon)	2,3	-	-	-
EC042_RS03935 (hypothetical protein)	2,3	13,9	4,1	3,9
EC042_RS12585 (hypothetical protein)	2,3	-	-	-
EC042_RS11020 (tRNA-Cys)	2,3	-	2,4	3,5
EC042_RS16670 (cell division protein ZapA)	2,3	-	-	-
EC042_RS07855 (tRNA-Gly)	2,2	-	-	-
EC042_RS21415 (type I toxin-antitoxin system toxin TisB)	2,2	-	-	-
EC042_RS13975 (tRNA-Ala)	2,2	-	3,1	4,9
EC042_RS24915 (biofilm peroxide resistance protein BsmA)	2,2	-	-	-
EC042_RS07090 (tRNA-Met)	2,2	2,3	2,6	4,1
EC042_RS16950 (tRNA-Phe)	2,2	8,8	4,1	4,1
EC042_RS24615 (tRNA-Phe)	2,2	6,5	3,2	2,5
EC042_RS05935 (biofilm formation regulator BssS)	2,2	-	2	-
EC042_RS09155 (hypothetical protein)	2,1	-	5,2	6,1
EC042_RS22215 (tRNA-Arg)	2,1	-	-	-
EC042_RS27045 (IS3 family transposase)	2,1	-	-	-
EC042_RS03195 (type I toxin-antitoxin system toxic protein HokeE)	2,1	-	-	4,4
EC042_RS07645 (DUF2534 family protein)	2,1	-	-	-
EC042_RS26425 (transposase)	2,1	-	-	-
EC042_RS08735 (hypothetical protein)	2,1	5,8	2,6	2,8
EC042_RS21780 (GNAT family N-acetyltransferase)	2,1	-	-	-
EC042_RS03860 (hypothetical protein)	2,1	-	-	-
EC042_RS09790 (major outer membrane lipoprotein)	2,1	-	-	-



EC042_RS06355 (hypothetical protein)	2,0	26,0	4,8	7,1
EC042_RS21810 (broad-spectrum mercury transporter MerE)	2,0	12,2	-	5
EC042_RS03600 (tRNA-Gln)	2,0	-	3,7	4,7
EC042_RS23365 (tRNA-Glu)	2,0	-	-	3
EC042_RS19745 (DUF2756 family protein)	2,0	-	-	-
EC042_RS25985 (tRNA-Leu)	2,0	-	2,1	-
EC042_RS04860 (biofilm formation regulator BssR)	2,0	-	-	-
EC042_RS21265 (hypothetical protein)	2,0	9,9	4,5	4,6
EC042_RS11870 (tRNA-Asn)	2,0	4,3	3,0	2,7
EC042_RS01250 (hypothetical protein)	2,0	20,0	7,8	3,9
EC042_RS25975 (tRNA-Leu)	2,0	-	4,1	8,7

**Tabla 4.3 (B).** Listado de los genes sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hha2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. También se indican los valores de expresión relativa de dichos genes en los mutantes  $\Delta hha$ ,  $\Delta hha\Delta hha2$  y  $\Delta hns$ . Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

Considerando en segundo lugar los genes sobreexpresados en el mutante *hha2*, es aparente que en su gran mayoría corresponden a genes no regulados por el sistema Hha-H-NS, lo que sugiere que la proteína Hha2 es capaz de regular la expresión génica de varios genes de una forma independiente a este mecanismo. Dicho sistema regulador queda pendiente de ser estudiado. Es importante mencionar que el número de genes de la cepa 042 sobreexpresados en un mutante sencillo *hha2* es reducido, y los valores de expresión relativa de los genes diferencialmente regulados no son tan elevados como en el caso de la mutación en el gen *hha*. De manera específica, la proteína Hha2 parece regular la expresión de diferentes sistema toxina-antitoxina, entre los que destaca el sistema toxina-antitoxina tipo I LdrD-RdID (cuya sobreexpresión provoca la muerte celular rápida y la condensación nucleoide de la célula huésped (Kawano y col., 2002)) y el sistema toxina-antitoxina tipo I Hok (cuya sobreexpresión provoca colapso del potencial transmembrana y la atenuación de la respiración celular (Gerdes y col., 1986)). Además, de manera análoga a los datos reportados en el caso de la mutación en el gen *hha*, donde se incrementa la expresión de tARNs que codifican por codones raros implicados en la formación de *biofilm* y en la activación de genes fágicos con propiedades tóxicas (García-Contreras y col., 2008), en el mutante *hha2* también se puede observar la sobreexpresión de un gran número de genes que codifican para ARNs de transferencia (tabla 4.3 (B)).

#### 4. RESULTADOS

Identificador NCBI ( <i>locus tag</i> )	Fold change			
	$\Delta hha\Delta hha2$	$\Delta hha$	$\Delta hha2$	$\Delta hns$
EC042_RS25175 (magnesium-translocating P-type ATPase)	23,2	-	-	-
EC042_RS08810 (hypothetical protein)	17,3	-	-	-
EC042_RS08660 (fimbrial protein)	13,0	-	-	10,3
EC042_RS07850 (tRNA-Thr)	12,3	2,7	-	7,0
EC042_RS11295 (hypothetical protein)	10,6	3,1	-	11,4
EC042_RS09275 (hypothetical protein)	10,3	3,4	-	13,3
EC042_RS07145 (hypothetical protein)	10,2	5,2	-	6
EC042_RS08655 (fimbria/pilus periplasmic chaperone)	9,3	5,9	-	7,7
EC042_RS24455 (hypothetical protein)	8,8	7,1	-	4
EC042_RS26995 (hypothetical protein)	8,6	4,8	-	6,1
EC042_RS18795 (hypothetical protein)	8,6	4,1	-	8,8
EC042_RS11730 (hypothetical protein)	8,5	2,8	-	8,5
EC042_RS13835 (tRNA-Arg)	8,5	7,1	-	8,6
EC042_RS08650 (fimbrial biogenesis outer membrane usher protein)	8,5	6,1	-	7,5
EC042_RS15165 (hypothetical protein DNA-binding protein)	8,4	4,9	-	7,2
EC042_RS05780 (tRNA-Ser)	8,4	-	-	17,3
EC042_RS16300 (HilA family transcriptional regulator YgeH)	8,3	6,9	-	7,8
EC042_RS08645 (type 1 fimbrial protein)	8,3	5,4	-	7,8
EC042_RS18175 (hypothetical protein)	8,1	6,1	-	10,3
EC042_RS09105 (hypothetical protein)	8,1	6,4	-	5,5
EC042_RS06765 (tRNA-Tyr)	7,8	-	-	16,3
EC042_RS01250 (hypothetical protein)	7,8	20,1	2,0	3,9
EC042_RS05870 (hypothetical protein)	7,7	5,1	-	10,1
EC042_RS15125 (molybdopterin-guanine dinucleotide biosynthesis protein MobA)	7,7	7,2	-	5,3
EC042_RS07390 (hypothetical protein)	7,7	4,8	-	7,4
EC042_RS21295 (CesD/SycD/LcrH family type III secretion system chaperone)	7,6	4,3	-	15,6
EC042_RS07765 (putative bacteriophage protein RacC)	7,5	3,9	-	9,4
EC042_RS17210 (hypothetical protein)	7,4	3,2	-	6,4
EC042_RS11285 (hypothetical protein)	7,3	5,2	-	9,5
EC042_RS24340 (hypothetical protein)	7,2	5,6	-	4,7
EC042_RS11745 (exodeoxyribonuclease VIII)	7,2	3,0	-	6,6
EC042_RS07745 (hypothetical protein)	7,2	4,4	-	7,5
EC042_RS03185 (hypothetical protein)	7,1	4,4	-	8,1
EC042_RS25650 (putative RadC-like DNA repair protein)	7,1	5,7	-	6,2
EC042_RS07760 (putative phage exodeoxyribonuclease)	7,1	3,6	-	6,6

EC042_RS01795 (reactive chlorine species resistance protein RclC)	7,1	4,3	-	7,0
EC042_RS11335 (hypothetical protein)	7,1	7,3	-	5,2
EC042_RS08640 (type 1 fimbrial protein)	7,1	4,9	-	7,4
EC042_RS21300 (YopB/SseC family type III secretion system)	7,1	6,0	-	7,4
EC042_RS09180 (site-specific integrase)	7,0	3,9	-	8,0

**Tabla 4.3 (C).** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hha\Delta hha2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. También se indican los valores de expresión relativa de dichos genes en los mutantes  $\Delta hha$ ,  $\Delta hha2$  y  $\Delta hns$ . Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

Finalmente, el estudio de los genes con un mayor nivel de sobreexpresión en el doble mutante *hha hha2* permite deducir algunas características de la función reguladora de la proteína Hha2. En este caso, la gran mayoría de dichos genes también se sobreexpresan en el mutante *hns* con valores equivalentes a los obtenidos en el caso del doble mutante  $\Delta hha\Delta hha2$ . Estos son, por tanto, los genes que clásicamente han sido considerados como genes regulados por la proteína Hha de manera conjunta con H-NS. La razón por la que dichos genes no aparezcan como los más sobreexpresados en el estudio realizado con el mutante sencillo *hha* es aparente: en este caso se expresa la proteína Hha2, hecho que ayuda a atenuar la pérdida de la proteína Hha. Ello puede deducirse al comprobar que dichos genes también se sobreexpresan en el mutante  $\Delta hha$  sencillo, pero con un nivel de sobreexpresión significativamente menor.

Por tanto, al igual que sucede con la proteína StpA en el sistema H-NS, la proteína Hha2 presenta, entre sus funciones, la capacidad de actuar como reservorio funcional para atenuar los efectos de la pérdida de la proteína Hha. Adicionalmente, dicha proteína también parece cumplir otras funciones reguladoras independientes de la propia proteína Hha y de H-NS, funciones que quedan pendientes de ser estudiadas.

Teniendo en cuenta lo expuesto anteriormente, consideramos que en la cepa *E. coli* 042, al existir la expresión del gen *hha2* con capacidad de complementar parcialmente la pérdida del gen *hha*, un fenotipo *hha*<sup>-</sup> implicaría inactivar simultáneamente tanto el gen *hha* como el gen *hha2*.

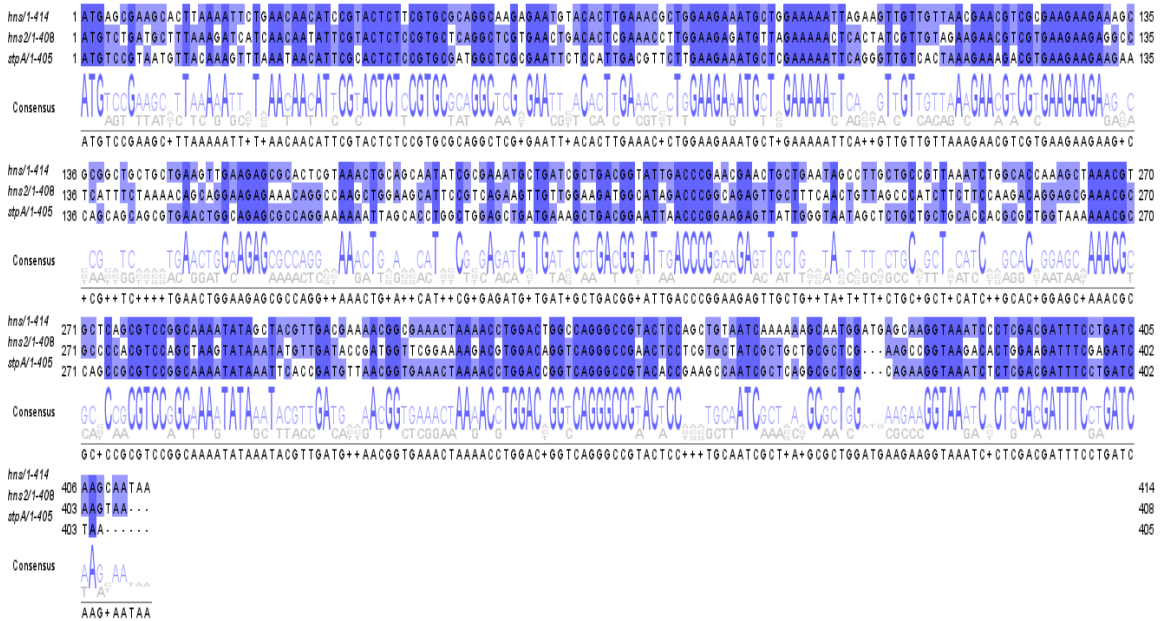
# Capítulo 2



## 4.2 Descripción de un nuevo parálogo del gen *hns* en la cepa *E. coli* 042: el gen *hns2*

### 4.2.1 Identificación de un nuevo parólogo del gen *hns* en el genoma de *E. coli* 042

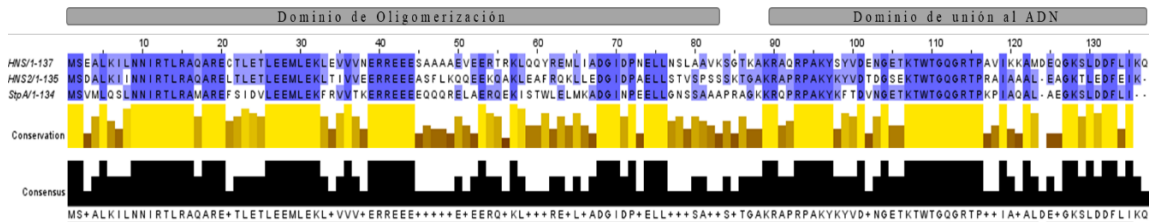
De forma análoga a la estrategia descrita en el capítulo 1 de Resultados de esta memoria en relación a las proteínas tipo Hha, mediante una búsqueda de secuencias homólogas a la secuencia nucleotídica y aminoacídica del gen que codifica para la proteína H-NS (NCBI- EC042\_1292 y Uniprot- D3H2L9) dentro del genoma anotado de la cepa *E. coli* 042 (NC\_017626.1), se logró identificar un nuevo parólogo del gen *hns* no descrito hasta la fecha, el gen EC042\_2834 (nombrado *hns2*). Este gen, codificado en el cromosoma de la cepa *E. coli* 042, se suma al gen *stpA* (Zhang y Belfort, 1992), presente en el genoma de todas las enterobacterias, y de manera específica al gen *hfp*, presente en la cepa uropatógena de *E. coli* 536 (Müller y col., 2010), y al gen *sfh*, presente en el plásmido conjugativo pSF-R27 de la cepa *Shigella flexneri* 2a 2457T (Beloin y col., 2003; Deighan y col., 2003), como genes parálogos identificados del gen *hns*. La figura 4.16 representa el alineamiento múltiple de la secuencia nucleotídica de los genes de los 3 parálogos tipo *hns* presentes en el cromosoma de la cepa *E. coli* 042.



**Figura 4.16.** Alineamiento múltiple de las secuencias nucleotídicas de los genes *hns* (EC042\_1292), *hns2* (EC042\_2834) y *stpA* (EC042\_2865) presentes en el cromosoma de la cepa *E. coli* 042. En color azul intenso se representan las posiciones cuya secuencia de nucleótidos se mantiene más conservada. La intensidad de color disminuye a medida que los nucleótidos varían entre los diferentes genes. Alineamiento múltiple realizado mediante el programa Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

#### 4. RESULTADOS

A nivel de secuencia nucleotídica, el gen *hns2* muestra un porcentaje de identidad del 59% respecto al gen *hns*, mientras que el porcentaje de identidad del gen *stpA* es del 63%. El porcentaje de identidad entre los genes de ambos parálogos es del 57%. En la figura 4.17 se representa el alineamiento múltiple de las secuencias proteicas de los parálogos de H-NS presentes en el cromosoma de *E. coli* 042. A nivel de secuencia proteica, la secuencia aminoacídica de H-NS2 muestra una identidad del 64,4% respecto a la proteína H-NS, superior al 53% que presenta StpA.

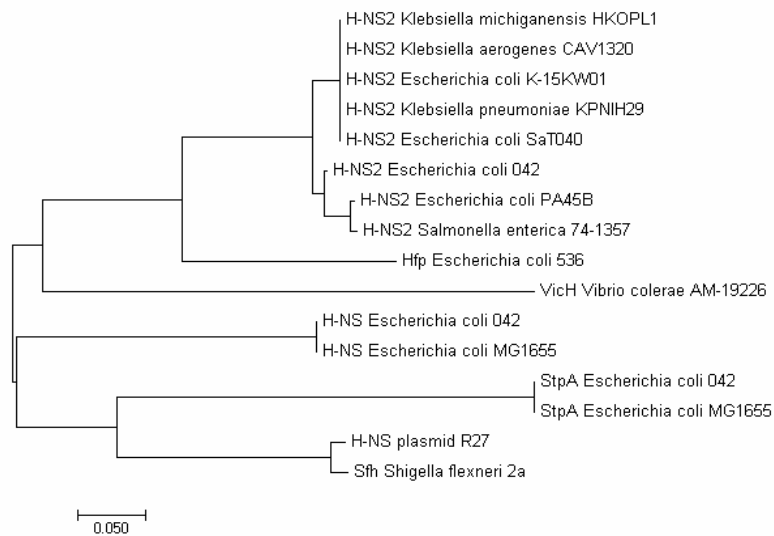


**Figura 4.17.** Alineamiento múltiple de las secuencias aminoacídicas de las proteínas H-NS, H-NS2 y StpA. De color azul intenso se representan las posiciones cuya secuencia de aminoácidos se mantiene más conservada. La intensidad de color disminuye a medida que los aminoácidos varían entre las diferentes proteínas. Los bloques en negro representan la secuencia consenso global para las 3 proteínas analizadas. Alineamiento múltiple realizado mediante el programa Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

En dicha figura 4.17 también se representan los dos dominios estructurales de la proteína H-NS. El dominio N-terminal está compuesto por 83 aminoácidos y es responsable de la oligomerización de la proteína. El dominio de oligomerización de la proteína H-NS2 muestra una mayor homología respecto al de H-NS que el que muestra StpA (identidad del 62.65% *versus* 53.01%), mientras que respecto al dominio C-terminal de la proteína H-NS (aminoácidos del 90 al 137), responsable de la interacción con el ADN, se encuentra más conservado en StpA que en H-NS2 (identidad del 70,5% *versus* 64,44%). En dicho dominio se localiza la secuencia consenso TWGTGQGRTP (aminoácidos del 108 y al 116, ampliamente conservados entre las proteínas H-NS de bacterias gramnegativas). Tanto H-NS2 como StpA contienen dicha secuencia consenso totalmente idéntica a la de la proteína H-NS. En dicha región, el triptófano (W) localizado en la posición 109 está íntimamente relacionado con la interacción con el ADN (Tippner y Wagner, 1995), mientras que la prolina (P) situada en la posición 116 está implicada en la discriminación entre regiones curvadas y no curvadas de ADN (Spurio y col., 1997).

### 4.2.2 Distribución del gen *hns2* en enterobacterias

Mediante herramientas bioinformáticas se analizó la distribución del gen *hns2* entre todos los genomas bacterianos secuenciados. Para ello, se realizó una búsqueda de homología por BLAST con la secuencia nucleotídica del gen *hns2* en la base de datos del NCBI. Curiosamente, se ha podido observar que la presencia de este gen queda restringida únicamente en pocos géneros de la familia *Enterobacteriaceae*, entre los que destacan cepas patógenas de los géneros *Escherichia*, *Klebsiella*, *Serratia* y *Salmonella*. Todas las cepas que contienen el gen *hns2* en su cromosoma también contienen los genes *hns* y *stpA*. En la figura 4.18 se muestra un dendograma que indica la relación evolutiva de la proteína H-NS2 con otros parálogos de la familia H-NS, como la proteína StpA, la proteína Sfh de *Shigella flexneri*, la proteína Hfp de la cepa uropatógena de *E. coli* 536 o la proteína VicH de *Vibrio cholerae*.



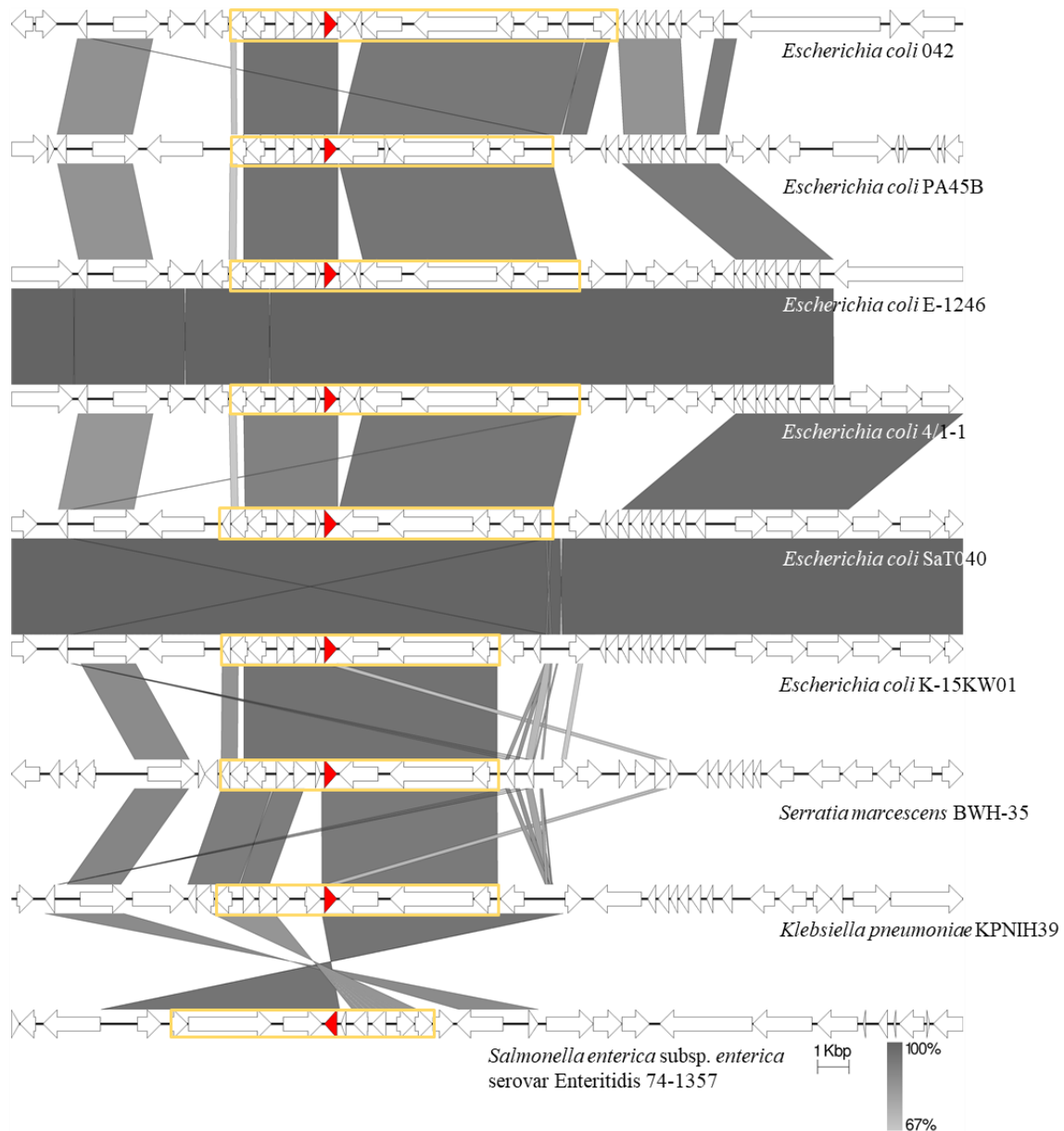
**Figura 4.18.** Dendograma en el que se muestra la relación evolutiva entre las proteínas H-NS, H-NS2 y StpA de diferentes miembros de la familia *Enterobacteriaceae*. El árbol filogenético se construyó mediante el programa MEGA7 (Kumar y col., 2016) haciendo uso del algoritmo *neighbor-joining* utilizando las secuencias aminoacídicas de las diferentes proteínas extraídas de la base de datos del NCBI. Se analizaron las proteínas H-NS2 de *Klebsiella aerogenes* CAV1320 (AKK82112.1), *Klebsiella michiganensis* HKOPL1 (AHW87442.1), *Escherichia coli* K-15KW01 (ANR82355.1), *Escherichia coli* SaT040 (AML15501.1), *Klebsiella pneumoniae* subsp. *pneumoniae* KPNIH29 (AIX70716.1), *Escherichia coli* PA45B (ASW60759.1), *Salmonella* enterica serotipo Enteritidis 74-1357 (ATT16728.1), las proteínas H-NS (CBG35667), H-NS2 (CBG35667) y StpA (CBG35697) de *Escherichia coli* 042, las proteínas H-NS (NP\_417155.1) y StpA (NP\_415753.1) de *Escherichia coli* MG1655, la proteína H-NS del plásmido R27 (NP\_058377.1), la proteína Sfh de *Shigella flexneri* 2a (AAN38840.1), la proteína VicH de *Vibrio cholerae* AM-192226 (ZP\_04962344.1) y la proteína Hfp de *Escherichia coli* 536 (ABG69928). La barra indica 0,050 sustituciones aminoacídicas por posición.



#### 4. RESULTADOS

La homología del gen *hns2* entre todas las cepas que contienen dicho gen en su genoma es de más del 95%, motivo por el cual las cepas que codifican para la proteína correspondiente se agrupan en el mismo *cluster* del dendograma. Además, la proteína H-NS2 presenta una mayor homología con los parálogos Hfp de la cepa UPEC 536 o VicH de *Vibrio cholerae* AM-19226 que la homología que presenta con la propia proteína H-NS o su parálogo StpA. Más alejadas en el dendograma se encuentran la proteína H-NS codificada en el plásmido IncHI1 R27 o la proteína Sfh de *Shigella flexneri*.

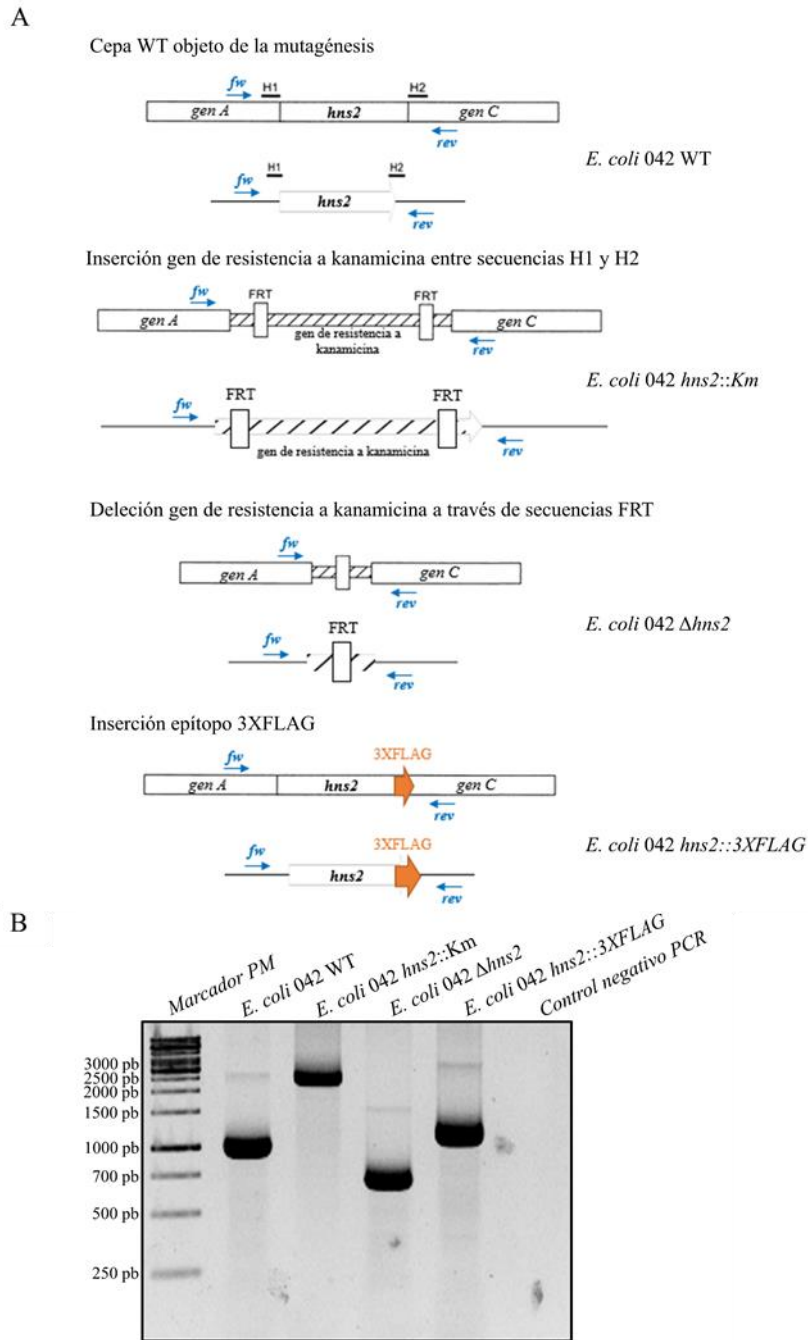
Al realizar una comparación global entre los entornos génicos del gen *hns2* entre las diferentes cepas que contienen dicho gen en su cromosoma, se puede observar cómo, a diferencia de los genes *hha2* y *hha3* descritos anteriormente, el *locus* en el que se encuentra el gen *hns2* no se encuentra flanqueado ni por transposasas, secuencias de inserción o secuencias génicas correspondientes a profagos. Sin embargo, cabe destacar que, en todas las cepas analizadas, la organización de los genes que flanquean al gen *hns2* tanto aguas arriba como aguas abajo es la misma (figura 4.19). Este conjunto de genes, formado por proteínas hipotéticas sin función conocida (a excepción del gen de mayor tamaño del bloque de genes común, que codifica para una proteína integral de membrana de la familia DotA/TraY perteneciente al sistema de secreción tipo IV implicado en la conjugación plasmídica), presenta la misma organización en el genoma de todas las cepas analizadas.



**Figura 4.19.** Organización génica de la región del genoma en la que se localiza el gen *hns2* (representado de color rojo) en las cepas de *E. coli* (042, PA45B, E-1246, 4/1-1, SaT040, K-15K01), *Serratia marcescens* BWH-35, *Klebsiella pneumoniae* KPNIH39 y *Salmonella* Enteritidis 74-1357. Se muestran las regiones de homología entre las diferentes cepas con un porcentaje de identidad superior al 67%. De color amarillo se indica el grupo de genes comunes que flanquean al gen *hns2* presente en todas las cepas analizadas. Alineamiento de genómica comparativa realizado mediante el programa *EasyFig* (Sullivan y col., 2011) haciendo uso de las secuencias génicas extraídas de la base de datos *GenBank* correspondientes al cromosoma de *E. coli* 042 (FN554766.1), *E. coli* PA45B (CP021288.1), *E. coli* E-1246 (CP025573.1), *E. coli* 4/1-1 (CP023844.1), *E. coli* SaT040 (CP014495.1), *E. coli* K-15K01 (CP016358.1), *Serratia marcescens* BWH-35 (CP020507.1), *Klebsiella pneumoniae* KPNIH39 (CP014762.1) y *Salmonella* Enteritidis 74-1357 (CP018642.1).

### 4.2.3 Obtención de mutantes de los genes *hns* y *hns2* en la cepa *E. coli* 042

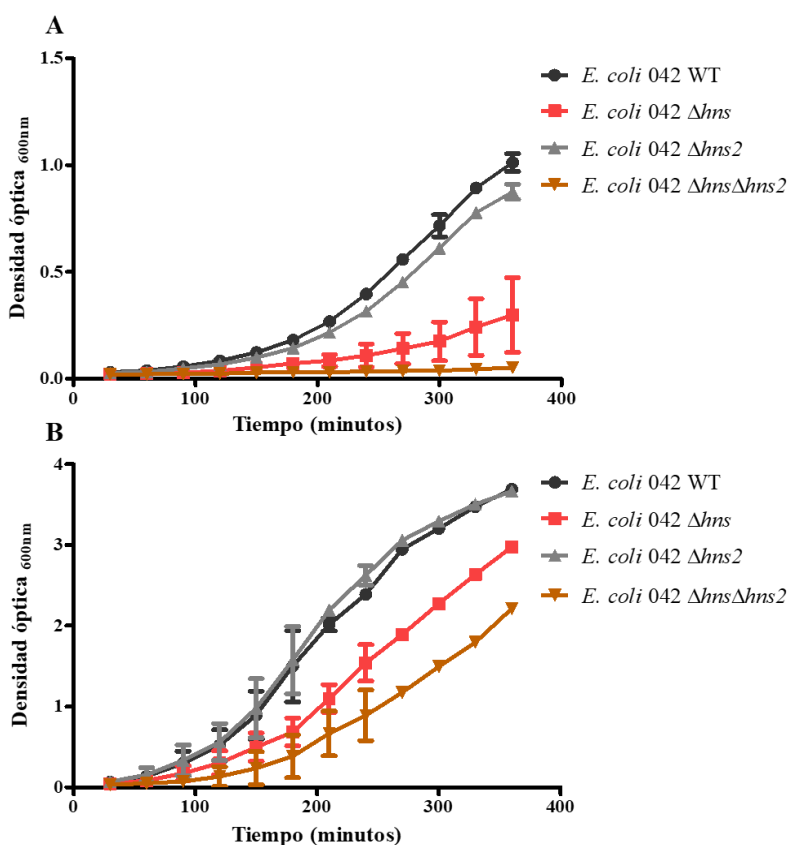
Después de identificar y caracterizar la secuencia de este nuevo parálogo del gen *hns*, el siguiente objetivo fue obtener el mutante del gen *hns2* en la cepa *E. coli* 042. El método utilizado para mutagenizar dicho gen, descrito en el apartado 3.7.7 de Materiales y Métodos, se basa en el protocolo de reemplazamiento alélico descrito por (Datsenko y Wanner, 2000), en el que el gen objeto de la mutagénesis es sustituido por un casete de resistencia a kanamicina. Dicho casete presenta un tamaño de 1,5 Kb, y posteriormente puede ser eliminado del cromosoma mediante la recombinasa FLP codificada en el plásmido pCP20. Paralelamente, para poder obtener un perfil de expresión proteica comparativo entre las proteínas H-NS y H-NS2, se construyeron las fusiones cromosómicas de los genes mencionados anteriormente con el epítipo 3XFLAG (siguiendo el protocolo descrito en el apartado 3.7.8), con el fin de inmunodetectar la expresión de dichas proteínas con un anticuerpo específico dirigido contra el epítipo añadido. La figura 4.20 muestra el análisis de la PCR realizada para confirmar las mutagénesis descritas anteriormente, tanto de la cepa *E. coli* 042 salvaje (WT) utilizada como control, como de las cepas mutantes *E. coli* 042 con el casete de resistencia a kanamicina insertado en la región codificante del gen *hns2* (*E. coli* 042 *hns2::Km*), *E. coli* 042 con la mutación del gen *hns2* después de la eliminación del casete de resistencia a kanamicina insertado previamente (*E. coli* 042  $\Delta$ *hns2*) y *E. coli* 042 con el epítipo 3XFLAG insertado en el extremo 3' del gen *hns2* (*E. coli* 042 *hns2::3XFLAG*). Los oligonucleótidos utilizados para realizar el análisis por PCR de las cepas mutantes (*hns2* p1 up-p2 down) hibridaban aguas arriba y aguas abajo de la región codificante del gen *hns2* (dando lugar a un tamaño de amplicón de 1000 pb en la cepa salvaje), y su secuencia nucleotídica se encuentra detallada en el Anexo de esta memoria en la tabla 9.11.



**Figura 4.20.** (A) Representación esquemática del protocolo utilizado para mutagenizar el gen *hns2* del cromosoma de la cepa *E. coli* 042. Las regiones H1 y H2 indican los lugares de inserción del gen de resistencia a kanamicina en la región codificante del gen *hns2*. En azul se indica el lugar de hibridación de los oligonucleótidos *hns2* p1up-p2down utilizados para realizar el genotipado para confirmar las diferentes mutaciones generadas. (B) Resultado de la PCR realizada con los oligonucleótidos *hns2* p1up-p2down para confirmar las modificaciones genéticas realizadas sobre el gen *hns2*. Se puede observar el incremento del tamaño del amplicón (respecto a la cepa salvaje) al insertar el gen de resistencia a kanamicina (1,5 Kb) en el gen *hns2* (*E. coli* 042 *hns2::Km*), y la posterior disminución del tamaño de dicho amplicón al eliminar dicho casete (*E. coli* 042  $\Delta hns2$ ). También se puede observar la diferencia de tamaño respecto al gen *hns2* de la cepa salvaje al insertar los nucleótidos correspondientes a la secuencia 3XFLAG en el extremo 3' del gen (*E. coli* 042 *hns2::3XFLAG*).

#### 4.2.4 Efecto de los genes *hns* y *hns2* sobre la cinética de crecimiento de la cepa *E. coli* 042

Una vez obtenidos los mutantes isogénicos de los genes *hns* y *hns2*, además del doble mutante combinando ambas mutaciones (*E. coli* 042  $\Delta hns\Delta hns2$ ), se decidió analizar cómo afectaban estas mutaciones a la cinética de crecimiento de las bacterias mutantes y se compararon dichas tasas de crecimiento respecto a la que presenta la cepa salvaje. Las curvas de crecimiento se realizaron en medio líquido rico (LB) tanto a 25°C como a 37°C, a partir de cultivos correspondientes a cada uno de los mutantes analizados crecidos previamente durante 16 horas e inoculados posteriormente (con una dilución 1:100) en medio de cultivo fresco. Para realizar la curva de crecimiento, se tomaron muestras cada 30 minutos de cada cultivo bacteriano y se midió la correspondiente densidad óptica (turbidimetría) en el espectrofotómetro a una longitud de onda de 600 nm (figura 4.21).



**Figura 4.21.** Curva de crecimiento de la cepa *E. coli* 042 salvaje (042 WT) y los correspondientes mutantes *hns* (042  $\Delta hns$ ), *hns2* (042  $\Delta hns2$ ) y *hns hns2* (042  $\Delta hns\Delta hns2$ ) en medio rico LB a 25°C (A) o a 37°C (B). Los cultivos con medio fresco fueron inoculados con una dilución 1:100 a partir de cultivos previos crecidos durante 16 horas en agitación constante. Se representan los valores correspondientes al promedio obtenido al realizar 3 curvas (réplicas biológicas) independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Analizando los resultados de las curvas de crecimiento realizadas, se puede observar cómo a 37°C el efecto de la mutación del gen *hns2* es indetectable cuando se compara respecto a la cepa salvaje. Sin embargo, cuando la mutación en el gen *hns2* se combina con la del gen *hns*, el efecto sobre la tasa de crecimiento del doble mutante es mayor que el del mutante sencillo *hns*. A 25°C, se observan los mismos efectos que a 37°C, pero más notorios: el efecto de la mutación del gen *hns2* provoca una ligera reducción de la tasa de crecimiento respecto a la cepa salvaje, mientras que cuando esta mutación se combina con la mutación en el gen *hns*, el efecto sobre la tasa de crecimiento del doble mutante *hns hns2* es total, generando una situación de crecimiento prácticamente nulo.

El hecho de que a 37°C la mutación sencilla del gen *hns2* apenas tenga un efecto sobre la fisiología de la célula a nivel de tasa de crecimiento nos sugiere que, bajo estas condiciones, la proteína H-NS es capaz de complementar totalmente el efecto de la pérdida de la proteína H-NS2. En cambio, a 25°C H-NS también es capaz de complementar una mutación *hns2*, pero únicamente de manera parcial, ya que el mutante *E. coli* 042  $\Delta hns2$  no es capaz de alcanzar los niveles de crecimiento de la cepa salvaje.

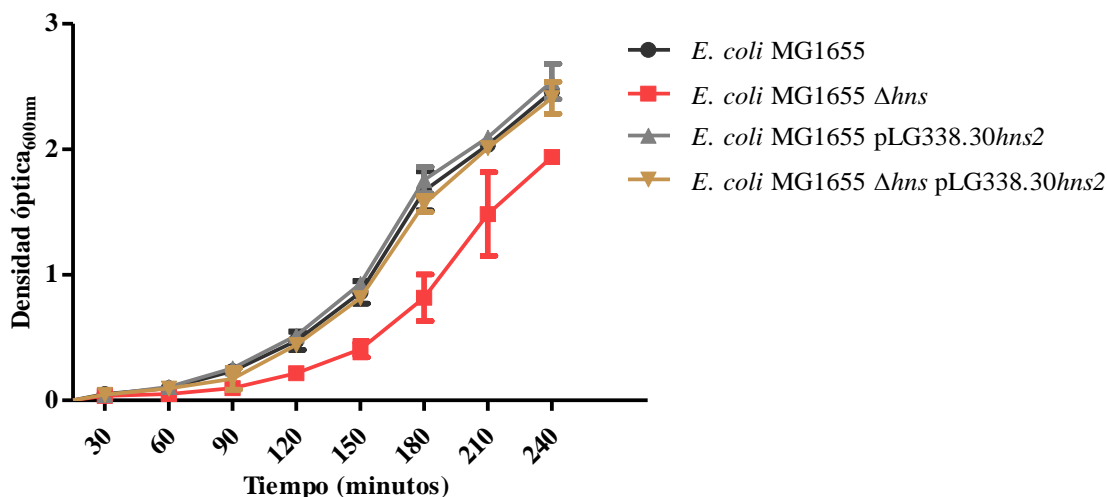
Además, el hecho de que un doble mutante *hns hns2* presente una tasa de crecimiento notablemente inferior a la del mutante sencillo *hns* nos indica que, cuando está presente la proteína H-NS2, ésta es capaz de compensar parcialmente la pérdida de la proteína H-NS, observándose un efecto más pronunciado a 25°C que a 37°C, situación en la que cuando no está presente ninguna de las dos proteínas la afectación sobre la tasa de crecimiento de las bacterias es total.

#### **4.2.5 Análisis de la capacidad de complementación de la mutación *hns* por parte del gen *hns2***

Seguidamente, se trató de analizar si la proteína H-NS2 de la cepa *E. coli* 042 era capaz de complementar la función realizada por la proteína H-NS en una cepa de *E. coli* que no codificara el propio gen *hns2* a nivel cromosómico. Para ello, se utilizó la cepa de laboratorio *E. coli* MG1655, cepa que contiene en su genoma los genes *hns* y *stpA*, pero no *hns2*. Se mutagenizó el gen *hns* de esta cepa (*E. coli* MG1655  $\Delta hns$ ) y se complementó dicho mutante con el gen *hns2* de *E. coli* 042 clonado en un vector de bajo número de copias bajo el control del propio promotor del gen (pLG338.30*hns2*). A continuación, se realizaron las correspondientes curvas de crecimiento en medio LB a 37°C con las cepas *E. coli* MG1655 salvaje, *E. coli* MG1655  $\Delta hns$ , *E. coli* MG1655 pLG338.30*hns2* y

#### 4. RESULTADOS

*E. coli* MG1655  $\Delta hns$  pLG338.30*hns2* (cepa cuya mutación en el gen *hns* fue complementada en *trans* con el gen *hns2* de la cepa *E. coli* 042). En la figura 4.22 se muestran los resultados obtenidos al realizar la curva de crecimiento de las cepas mencionadas anteriormente.



**Figura 4.22.** Curva de crecimiento de las cepas *E. coli* MG1655 salvaje, *E. coli* MG1655  $\Delta hns$ , *E. coli* MG1655 pLG338.30*hns2* y *E. coli* MG1655  $\Delta hns$  pLG338.30*hns2* en medio rico LB a 37°C. Los cultivos con medio fresco fueron inoculados con una dilución 1:100 a partir de cultivos previos crecidos durante 16 horas en medio LB a 37°C en agitación constante. Se representan los valores correspondientes a 3 réplicas biológicas independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

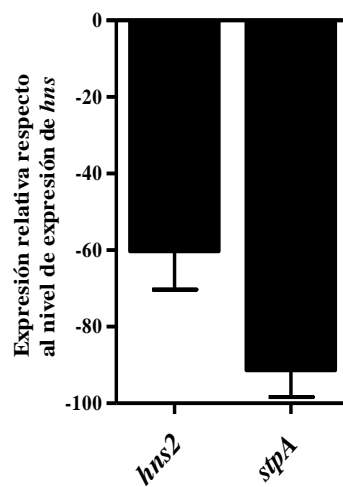
De manera análoga a lo que ocurría con la cepa *E. coli* 042, la mutación en el gen *hns* provoca importantes efectos negativos a nivel de crecimiento en la cepa *E. coli* MG1655. Sin embargo, cuando dicha mutación se complementa con el gen *hns2* de la cepa *E. coli* 042, la cepa mutante es capaz de revertir y alcanzar niveles de crecimiento similares a los de la cepa salvaje. Ello nos indica que la secuencia nucleotídica del gen *hns2* (y por ende la secuencia aminoacídica de la proteína H-NS2) presenta la capacidad de mimetizar y complementar (en parte) las funciones realizadas por la proteína H-NS en la cepa *E. coli* MG1655.

#### 4.2.6 Análisis del perfil de expresión del gen *hns2* y de su producto génico

A continuación, se decidió analizar el perfil de expresión del gen *hns2*. Para ello, se analizó por qRT-PCR el nivel de expresión del ARN mensajero del gen *hns2* (además del otro parálogo, *stpA*) en función del nivel de expresión de *hns* en medio LB a 37°C al inicio de la fase estacionaria de crecimiento (células crecidas hasta un nivel de D.O<sub>600nm</sub> 2,0).



A pesar de presentar funciones regulatorias específicas bajo determinadas condiciones ambientales, si asumimos que en *E. coli* tanto las proteínas StpA (Shi y Bennett, 1994; Williams y col., 1996; Deighan y col., 2000; Bertin y col., 2001) como H-NS2 (figura 4.22) son capaces de realizar funciones análogas a las de la proteína H-NS, no debería tener sentido para la célula destinar recursos energéticos para expresar varias proteínas cuya función es redundante, por lo que hipotetizamos que en condiciones de crecimiento óptimo sin estrés se esperarían identificar bajos niveles de expresión de los dos parálogos estudiados. En la figura 4.23 se representan los resultados de expresión de los genes *hns2* y *stpA* en función del nivel de expresión de *hns* en un fondo genético salvaje.



**Figura 4.23.** Expresión transcripcional (monitorizada por qRT-PCR) de los genes *hns2* y *stpA* representada de manera relativa al nivel de expresión del gen *hns* al inicio de la fase estacionaria de crecimiento en medio LB a 37°C en la cepa *E. coli* 042 salvaje. La expresión de los genes se evaluó utilizando los oligonucleótidos *hns* RT Fw/Rv, *hns2* RT Fw/Rv y *stpA* RT Fw/Rv. Los resultados fueron normalizados mediante la detección del gen interno *gapA* haciendo uso de los oligonucleótidos *gapA042* RT Fw/Rv. Se representan los valores correspondientes a 3 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Mientras que los niveles de expresión del gen *hns* son elevados y se mantienen relativamente constantes a lo largo del crecimiento bacteriano (Hinton y col., 1992; Schröder y Wagner, 2002), se puede observar que los niveles de expresión de los genes *hns2* y *stpA* son muy inferiores a los niveles de expresión de *hns*, siendo los niveles de transcripción del gen *hns2* superiores a los del gen *stpA* en las condiciones analizadas (inicio de la fase de crecimiento estacionaria en medio rico a 37°C).

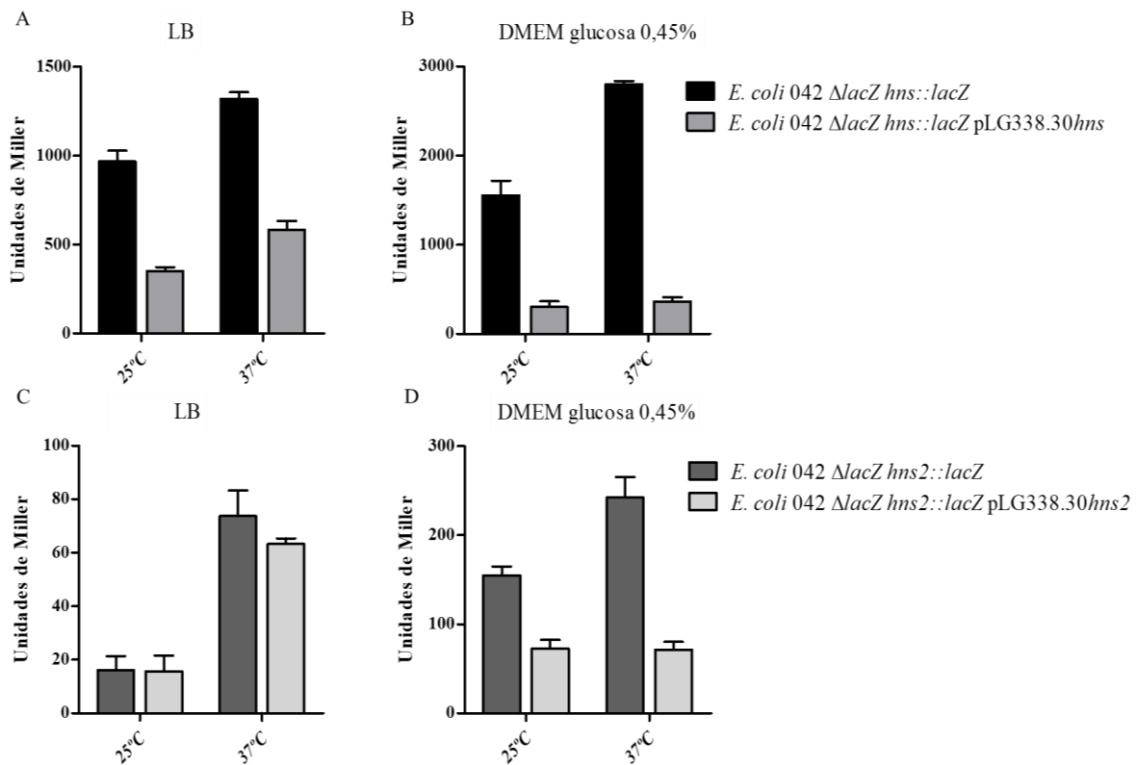
#### 4. RESULTADOS

Para confirmar y complementar los datos relativos a la expresión transcripcional del gen *hns2* comparados con los niveles de expresión de *hns*, se realizaron fusiones transcripcionales cromosómicas de los genes *hns* y *hns2* con el gen reportero *lacZ*. Ya que el gen *hns* se autorregula negativamente (la proteína H-NS reprime la transcripción de su propio gen) (Ueguchi y col., 1993), y al insertar el gen *lacZ* (siguiendo el protocolo descrito por (Ellermeier y col., 2002)) bajo el control del promotor del gen *hns* se tiene que eliminar la secuencia codificante de dicho gen, para poder realizar los análisis transcripcionales mediante ensayos de beta galactosidasa con la presencia de la proteína necesaria para la autorregulación, se complementaron en *trans* las cepas con las fusiones transcripcionales en los genes *hns* y *hns2* con los plásmidos pLG338.30*hns* y pLG338.30*hns2*. Las cepas resultantes se crecieron en medio LB y medio DMEM suplementado con glucosa 0,45% (p/v) a 25°C y 37°C hasta el inicio de la fase estacionaria de crecimiento, momento en el que se cuantificaron los niveles de expresión de los genes *hns* y *hns2* fusionados con el gen reportero *lacZ* mediante ensayos de  $\beta$ -galactosidasa (figura 4.24).

Puede observarse como, en el inicio de la fase estacionaria de crecimiento, la expresión transcripcional del gen *hns* es mucho mayor que la expresión del gen *hns2* (dichos resultados corroboran los resultados obtenidos mediante qRT-PCR representados en la figura 4.23). Cuando se complementa la cepa *E. coli* 042 *hns::lacZ* con la proteína H-NS codificada en el plásmido pLG338.30*hns*, se observa una gran disminución de la expresión transcripcional del gen *hns*, confirmando de esta manera la autorregulación negativa ejercida por esta proteína.

En el caso del gen *hns2*, la autorregulación negativa ejercida por la proteína nativa H-NS2 sobre su propio promotor se observa únicamente en el caso de los cultivos crecidos en medio DMEM, donde además la expresión transcripcional del gen es mucho mayor que en medio LB. En el caso de los cultivos crecidos en medio LB, la expresión transcripcional del gen *hns2* parece independiente de la presencia o ausencia de la propia proteína H-NS2. Ello nos indica que en estas condiciones o no existen los mecanismos autorreguladores a nivel la expresión génica del propio gen, o bien estos mecanismos reguladores están ejercidos por la proteína H-NS. Sin embargo, cabe destacar que el hecho de que los niveles de expresión transcripcional del gen *hns2* (representados en Unidades de Miller) sean tan bajos bajo estas condiciones dificulta observar una disminución de la

expresión transcripcional cuando los mutantes *hns2::lacZ* se complementan con el plásmido pLG338.30*hns2*.

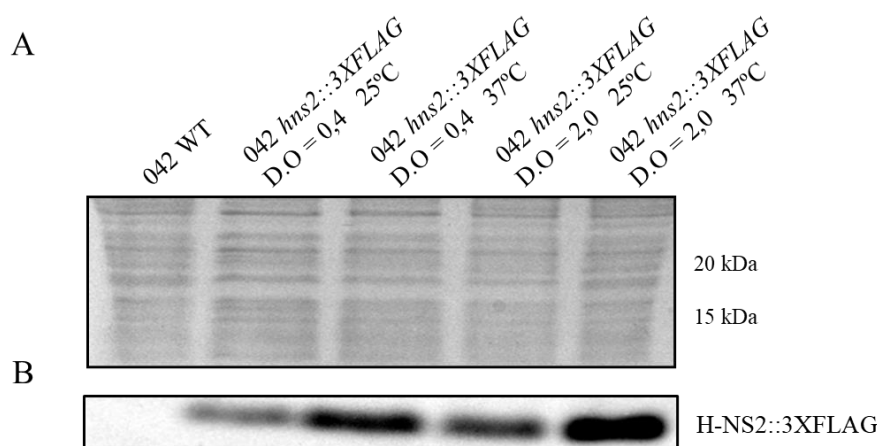


**Figura 4.24.** Expresión transcripcional del gen reportero *lacZ* insertado en el cromosoma de la cepa *E. coli* 042 bajo el control del promotor de los genes *hns* y *hns2* (*E. coli* 042  $\Delta lacZ$  *hns::lacZ* y *E. coli* 042  $\Delta lacZ$  *hns2::lacZ*). Dichas cepas se complementaron también con las proteínas H-NS y H-NS2 codificadas en los plásmidos pLG338.30*hns* y pLG338.30*hns2*, respectivamente. Los niveles de transcripción se cuantificaron mediante ensayos de  $\beta$ -galactosidasa a partir de cultivos crecidos en medio LB (A y C) o medio DMEM suplementado con glucosa al 0,45% (p/v) (B y D) tanto a 25°C como a 37°C hasta el inicio de la fase estacionaria de crecimiento. Se representa el promedio y la desviación estándar correspondientes a tres réplicas biológicas independientes.

El siguiente paso fue tratar de caracterizar los niveles de expresión de la proteína H-NS2. Para ello, se utilizó cepa que contenía el epítipo 3XFLAG insertado en el extremo 3' del gen *hns2*, dando lugar a la proteína H-NS2 fusionada con el epítipo proteico en su extremo C-terminal. Con esta finalidad, se creció la cepa *E. coli* 042 *hns2::3XFLAG* en medio rico LB tanto a 25°C como a 37°C, y se obtuvieron extractos proteicos totales de cultivos en fase exponencial ( $D.O_{600nm} = 0,4$ ) y en el inicio de la fase estacionaria ( $D.O_{600nm} = 2,0$ ). Para comprobar el nivel de expresión de la proteína H-NS2::3XFLAG en las diferentes condiciones de crecimiento analizadas, se realizó una inmunodetección

#### 4. RESULTADOS

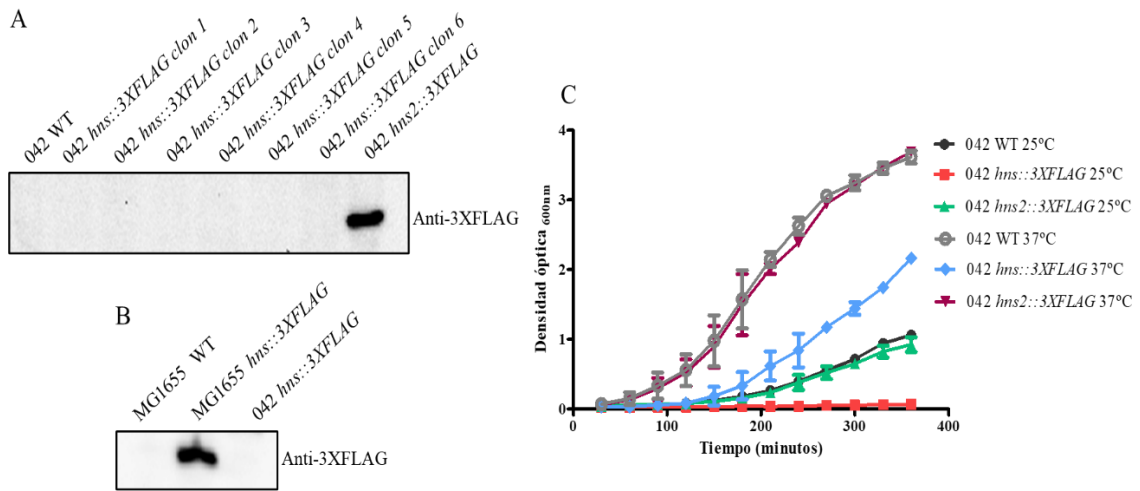
de la proteína H-NS2 presente en dichos extractos utilizando un anticuerpo cuya diana es la secuencia aminoacídica del epítipo 3XFLAG (figura 4.25).



**Figura 4.25.** Análisis por *Western Blot* de la expresión de la proteína H-NS2::3XFLAG en medio rico LB tanto a 25°C como a 37°C en las fases de crecimiento exponencial (D.O<sub>600nm</sub> 0,4) y estacionaria (D.O<sub>600nm</sub> 2,0). (A) Tinción de Coomassie de cada uno de los extractos proteicos analizados mediante el cual se representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas. (B) Resultado de la inmunodetección de la proteína H-NS2::3XFLAG mediante un anticuerpo dirigido contra el epítipo 3XFLAG.

Se puede observar cómo los niveles de expresión de la proteína H-NS2::3XFLAG son superiores a 37°C que a 25°C. Además, la proteína se expresa más cuando las células entran en la fase de crecimiento estacionaria respecto cuando las bacterias se encuentran en fase de crecimiento exponencial.

Para poder comparar el perfil de expresión proteica de H-NS2 con el de H-NS, también se realizó el marcaje con el epítipo 3XFLAG de la proteína H-NS de la cepa *E. coli* 042. Seguidamente se realizó un análisis por *Western Blot* con el fin de inmunodetectar dicha proteína. Para realizar este experimento comparativo, se crecieron 6 clones independientes con el epítipo 3XFLAG insertado en el extremo 3' del gen *hns* (*E. coli* 042 *hns*::3XFLAG, clones 1 al 6) y el clon con el epítipo 3XFLAG insertado en el extremo 3' del gen *hns2* (*E. coli* 042 *hns2*::3XFLAG). Notoriamente, al crecer los cultivos en medio LB para obtener los extractos proteicos para realizar la inmunodetección de las proteínas marcadas con el epítipo 3XFLAG, la cepa *E. coli* 042 *hns*::3XFLAG presentó una tasa de crecimiento extremadamente baja (figura 4.26 (C)).



**Figura 4.26.** (A) Análisis por *Western Blot* de la expresión comparativa de 6 clones independientes de la cepa *E. coli* 042 expresando la proteína H-NS::3XFLAG y de un clon expresando la proteína H-NS2::3XFLAG en medio rico LB a 37°C en el inicio de la fase de crecimiento estacionaria ( $D_{O_{600nm}}$  2,0). (B) Análisis por *Western Blot* de la expresión de la proteína H-NS::3XFLAG en las cepas *E. coli* 042 y *E. coli* MG1655 a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento. (C) Curva de crecimiento de las cepas *E. coli* 042 salvaje (042 WT), *E. coli* 042 *hns*::3XFLAG y *E. coli* 042 *hns2*::3XFLAG en medio LB a 25°C y 37°C. Se representan los valores correspondientes a 3 réplicas biológicas independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

En la figura 4.26 (A) se puede comprobar cómo, además de la baja tasa de crecimiento de la cepa *E. coli* 042 *hns*::3XFLAG respecto a la cepa salvaje (042 WT), sorprendentemente no se observó expresión alguna de la proteína H-NS fusionada con el epítipo 3XFLAG en ninguno de los clones analizados. Para descartar un posible error en el diseño de los oligonucleótidos utilizados para realizar la inserción del epítipo en el gen *hns* (*hns3X* p1-*hns3X* p2, cuya secuencia está detallada en la tabla 9.11 presente en el Anexo de esta memoria), se decidió utilizar esos mismos oligonucleótidos para marcar con la secuencia 3XFLAG el gen *hns* de la cepa *E. coli* MG1655. Siguiendo el mismo protocolo utilizado para marcar en la cepa *E. coli* 042 los genes *hns* y *hns2* (Uzzau y col., 2001), se marcó el gen *hns* en la cepa *E. coli* MG1655, obteniéndose el clon de *E. coli* MG1655 *hns*::3XFLAG. En la propia figura 4.26 (B) se puede observar cómo, a diferencia de la cepa 042, en la cepa MG1655 se pudo detectar una gran expresión de la proteína H-NS.

El hecho de no observar expresión alguna de la proteína H-NS en la cepa *E. coli* 042 puede indicar que, durante el proceso de recombinación homóloga necesario para la inserción de la secuencia codificante del epítipo 3XFLAG antes del codón de *stop* de la

#### 4. RESULTADOS

traducción del gen *hns* de la cepa *E. coli* 042, se esté generando una forma truncada de la proteína que, además de no ser funcional, provoca efectos deletéreos sobre el crecimiento de la bacteria. En la figura 4.27 se indica el alineamiento realizado entre las secuencias nucleotídicas y aminoacídicas del gen *hns* de las cepas *E. coli* 042 y *E. coli* MG1655 con la finalidad de identificar alguna divergencia en la región C-terminal de ambos genes que pudieran explicar por qué en la cepa MG1655 se pudo detectar una gran expresión de la proteína H-NS, mientras que en la cepa 042 no se pudo observar expresión alguna.

#### A

<i>hns_E. coli</i> 042	ATGAGCGAAGCACTTAAAAATTCTGAACAACATCCGTACTCTTCGTGCGCAGGCAAGAGAA	60
<i>hns_E. coli</i> MG1655	ATGAGCGAAGCACTTAAAAATTCTGAACAACATCCGTACTCTTCGTGCGCAGGCAAGAGAA *****	60
<i>hns_E. coli</i> 042	TGTACACTTGAACCGCTGGAAGAAATGCTGGAAAAATTAGAAGTTGTTTAACGAACGCT	120
<i>hns_E. coli</i> MG1655	TGTACACTTGAACCGCTGGAAGAAATGCTGGAAAAATTAGAAGTTGTTTAACGAACGCT *****	120
<i>hns_E. coli</i> 042	CGCGAAGAAGAAAGCGCGGCTGCTGCTGAAGTTGAAGAGCGCACTCGTAAACTGCAGCAA	180
<i>hns_E. coli</i> MG1655	CGCGAAGAAGAAAGCGCGGCTGCTGCTGAAGTTGAAGAGCGCACTCGTAAACTGCAGCAA *****	180
<i>hns_E. coli</i> 042	TATCGCGAAATGCTGATCGCTGACGGTATTGACCCGAACGAACTGCTGAATAGCCTTGCT	240
<i>hns_E. coli</i> MG1655	TATCGCGAAATGCTGATCGCTGACGGTATTGACCCGAACGAACTGCTGAATAGCCTTGCT *****	240
<i>hns_E. coli</i> 042	GCCGTTAAATCTGGCACCAAAGCTAAACGTGCTCAGCGTCCGGCAAAATATAGCTACGTT	300
<i>hns_E. coli</i> MG1655	GCCGTTAAATCTGGCACCAAAGCTAAACGTGCTCAGCGTCCGGCAAAATATAGCTACGTT *****	300
<i>hns_E. coli</i> 042	GACGAAAACGGCGAAACTAAAACCTGGACTGGCCAGGGCCGTA	360
<i>hns_E. coli</i> MG1655	GACGAAAACGGCGAAACTAAAACCTGGACTGGCCAGGGCCGTA *****	360
<i>hns_E. coli</i> 042	AAAGCAATGGATGAGCAAGGTAATCCCTCGACGATTTCTGATCAAGCAATAA	414
<i>hns_E. coli</i> MG1655	AAAGCAATGGATGAGCAAGGTAATCCCTCGACGATTTCTGATCAAGCAATAA *****	414

#### B

H-NS_ <i>E. coli</i> 042	MSEALKILNNIRTLRAQARECTLETLEEMLEKLEVVVNERREEESAAAAEVEERTRKLQQ	60
H-NS_ <i>E. coli</i> MG1655	MSEALKILNNIRTLRAQARECTLETLEEMLEKLEVVVNERREEESAAAAEVEERTRKLQQ *****	60
H-NS_ <i>E. coli</i> 042	YREMLIADGIDPNELLNSLAAVKSGTKAKRAQRPAYSYVDENGETKKTWTGQGRTPAVIK	120
H-NS_ <i>E. coli</i> MG1655	YREMLIADGIDPNELLNSLAAVKSGTKAKRAQRPAYSYVDENGETKKTWTGQGRTPAVIK *****	120
H-NS_ <i>E. coli</i> 042	KAMDEQGKSLDDFLIKQ	137
H-NS_ <i>E. coli</i> MG1655	KAMDEQGKSLDDFLIKQ *****	137

**Figura 4.27.** Alineamiento comparativo de las secuencias nucleotídicas (A) y aminoacídicas (B) del gen *hns* de las cepas *E. coli* 042 y *E. coli* MG1655. Se muestra con un recuadro en negro el único cambio nucleotídico entre las dos secuencias, localizado en la posición 336. Alineamiento realizado mediante el programa Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

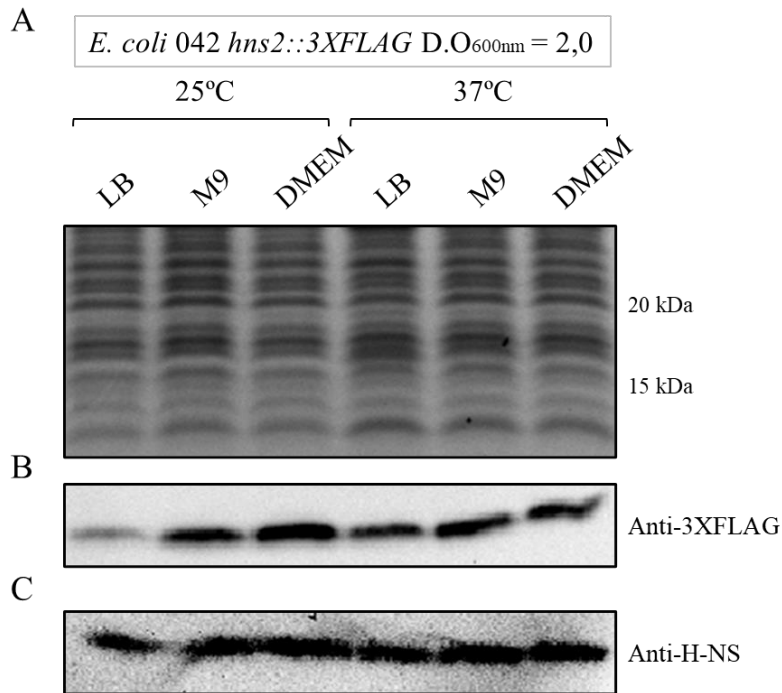
La única diferencia identificada entre la secuencia codificante del gen *hns* de la cepa *E. coli* 042 respecto la cepa MG1655 es un cambio de una guanina (G) por una adenina (A) en la posición 336. Dicho cambio es un cambio sinónimo ya que tanto el codón CAG como el codón CAA codifican para el mismo aminoácido (glutamina, Q), presente en la posición 112 de la secuencia aminoacídica de la proteína H-NS. Por todo ello, dada la total identidad de la secuencia proteica de H-NS entre las cepas de *E. coli* 042 y MG1655, desconocemos el motivo por el cual, al tratar de insertar una secuencia en el extremo 3' del gen *hns* de la cepa *E. coli* 042, se genere una construcción truncada con importantes efectos negativos sobre la fisiología de la bacteria.

Por el contrario, la adición del epítipo 3XFLAG en el extremo C-terminal de la proteína H-NS2 no provoca cambios en la tasa de crecimiento respecto a la cepa salvaje y permite determinar la expresión de la proteína (figuras 4.25 y 4.26 (A)). Además, ya que H-NS2 se expresa más en condiciones de estrés y limitación de nutrientes (situación correspondiente a la fase estacionaria de crecimiento bacteriano), también se decidió analizar la expresión de esta proteína en medio de crecimiento mínimo. Para ello, la cepa con la construcción *hns2::3XFLAG* se creció en medio LB, medio mínimo M9 suplementado con glucosa al 0,40% (p/v) y medio DMEM suplementado con glucosa al 0,45% (p/v), tanto a 25°C como a 37°C. De cada cultivo se obtuvieron los correspondientes extractos proteicos para poder inmunodetectar posteriormente la proteína H-NS2::3XFLAG. A partir de esos mismos extractos proteicos, también se analizó la expresión de la proteína H-NS utilizando anticuerpos específicos dirigidos contra esta proteína (figura 4.28).

A pesar de no poder comparar la expresión de H-NS y H-NS2 utilizando el mismo anticuerpo dirigido contra el epítipo 3XFLAG, sí que se pueden observar diferencias en el patrón de expresión de estas dos proteínas al inmunodetectar H-NS con un anticuerpo específico. Se puede observar cómo, tal y como se ha descrito por parte de (Hinton y col., 1992; Free y col., 1995; Müller y col., 2010), los niveles de expresión de la proteína H-NS se mantienen relativamente constantes independientemente de las condiciones de crecimiento de la bacteria. Por el contrario, los niveles de expresión de la proteína H-NS2 son más variables, expresándose con una mayor intensidad cuando las bacterias crecen en una situación de estrés bajo limitación de nutrientes, como el crecimiento en medios no ricos como son el medio DMEM o el medio mínimo M9.

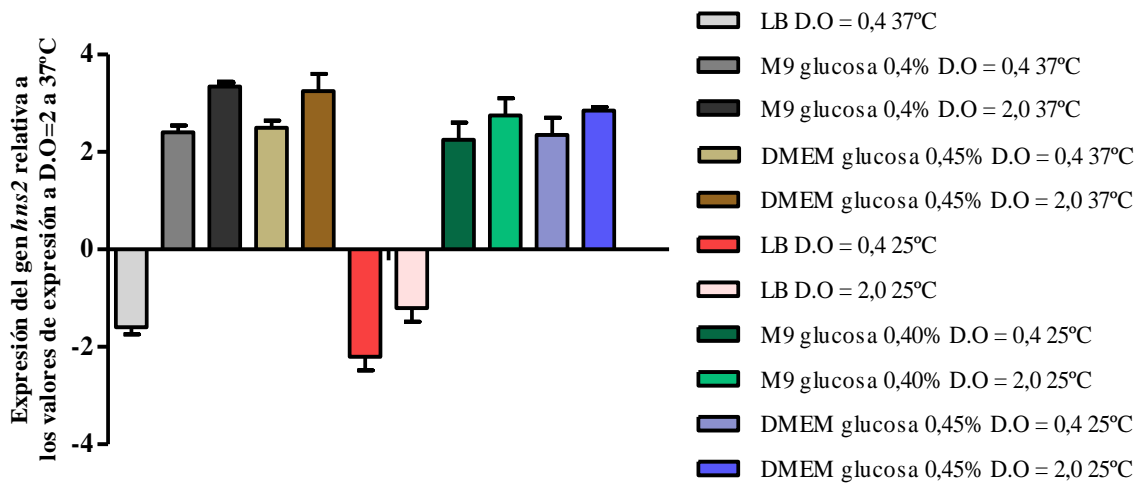


#### 4. RESULTADOS



**Figura 4.28.** Análisis por *Western Blot* de la expresión de las proteínas H-NS2::3XFLAG (B) y H-NS (C) en medio rico LB, medio mínimo M9 (suplementado con glucosa al 0,40% (p/v)) y medio DMEM (suplementado con glucosa al 0,45% (p/v)) tanto a 25°C como a 37°C en el inicio de la fase de crecimiento estacionaria (D.O<sub>600nm</sub> 2,0). El panel de arriba (A) muestra la tinción de Coomassie de cada uno de los extractos proteicos analizados y representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas.

Para acabar de validar los datos de expresión del gen *hns2* en las diferentes condiciones de cultivo analizadas, se realizó un análisis transcripcional por qRT-PCR para correlacionar los niveles de transcripción del gen *hns2* con los niveles de expresión proteica observados en la figura 4.28 (B). En la figura 4.29 se representan los valores correspondientes al nivel de transcripción del gen *hns2* expresados de manera relativa al nivel de expresión de dicho gen bajo condiciones de crecimiento en medio rico (LB) a 37°C cuando las células entran en la fase estacionaria de crecimiento.



**Figura 4.29.** Expresión relativa del gen *hms2* en diferentes condiciones de cultivo (medio rico LB, medio mínimo M9 y medio DMEM, tanto a 25°C como a 37°C, en las fases de crecimiento exponencial y el inicio de la fase estacionaria) respecto a los niveles de expresión del gen *hms2* en medio LB a 37°C en la fase estacionaria de crecimiento. La expresión del gen *hms2* se monitorizó por qRT-PCR, utilizando los oligonucleótidos *hms2* RT Fw/Rv. Los resultados fueron normalizados mediante la detección del gen interno *gapA* mediante los oligonucleótidos *gapA042* RT Fw/Rv. Se representan los valores correspondientes a 3 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

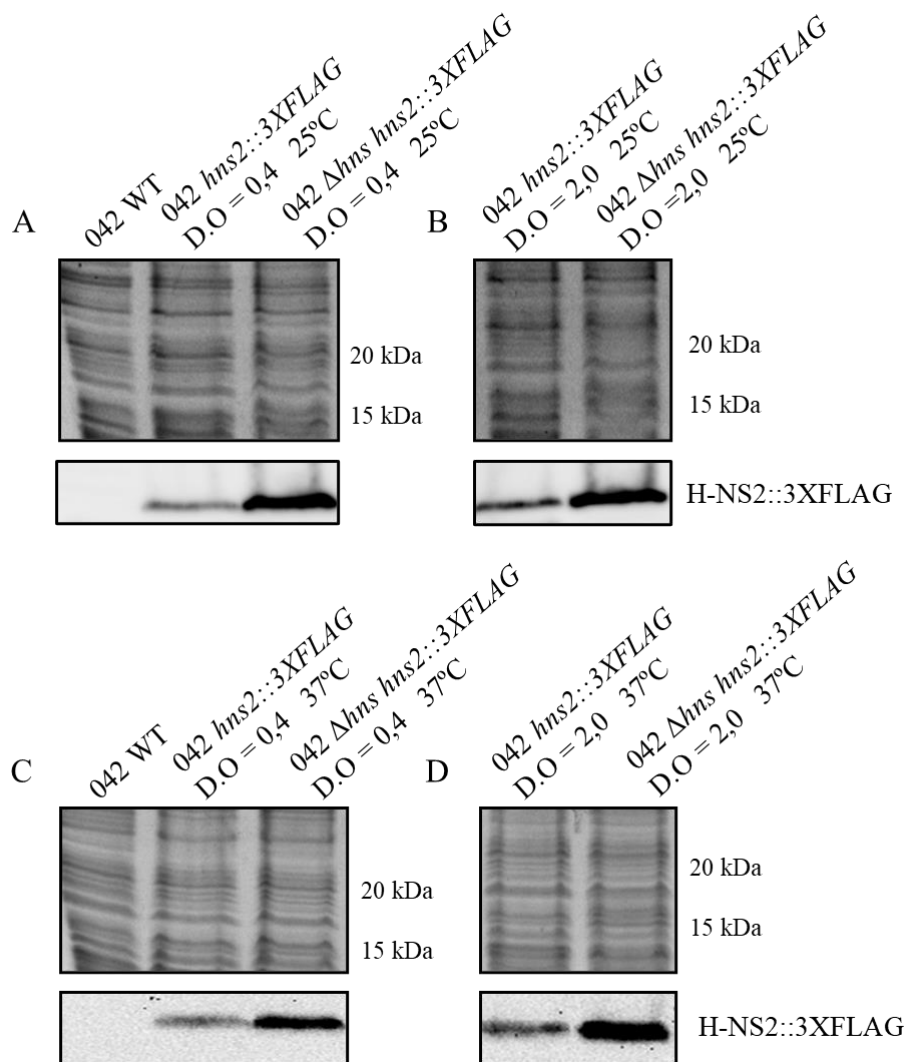
Globalmente, parece existir una buena correlación entre los niveles de transcripción del gen *hms2* monitorizados por qRT-PCR y los niveles de expresión de la proteína detectados mediante *Western Blot* (figura 4.28 (B) y 4.29). A diferencia del gen *hns*, que se expresa de manera constitutiva independientemente de las condiciones del cultivo (temperatura, limitación de nutrientes o fase de crecimiento), la expresión del gen *hms2* es dependiente de la fase de crecimiento bacteriana (expresión mayor en la fase de crecimiento estacionaria respecto a la fase exponencial), de la temperatura (mayor expresión a 37°C que a 25°C) y de la disponibilidad de nutrientes en el medio (ante una situación de limitación y una menor disponibilidad de nutrientes la expresión del gen *hms2* se incrementa).

También se quiso analizar si, tal y como ocurre en el caso de la proteína StpA la cual, ante una situación de ausencia de la proteína H-NS se sobreexpresa y es capaz de compensar parcialmente la inactivación del gen *hns* actuando como un reservorio molecular de esta proteína (Shi y Bennet, 1994; Williams y col., 1996; Deighan y col., 2000), la proteína H-NS2 presenta un comportamiento análogo. Para ello, se mutagenizó el gen *hns* en la cepa *E. coli* 042 que presenta la construcción 3XFLAG en el extremo

#### 4. RESULTADOS

C-terminal de la proteína H-NS2, y se comparó por *Western Blot* la expresión de dicha proteína en un fondo genético salvaje respecto un fondo genético *hns*<sup>-</sup>.

En la figura 4.30 se puede observar cómo, tal y como ocurre en el caso de la proteína StpA, la proteína H-NS2 se sobreexpresa en ausencia de H-NS tanto si las bacterias crecen a 25°C o 37°C, o si se encuentran en la fase de crecimiento exponencial o estacionaria.



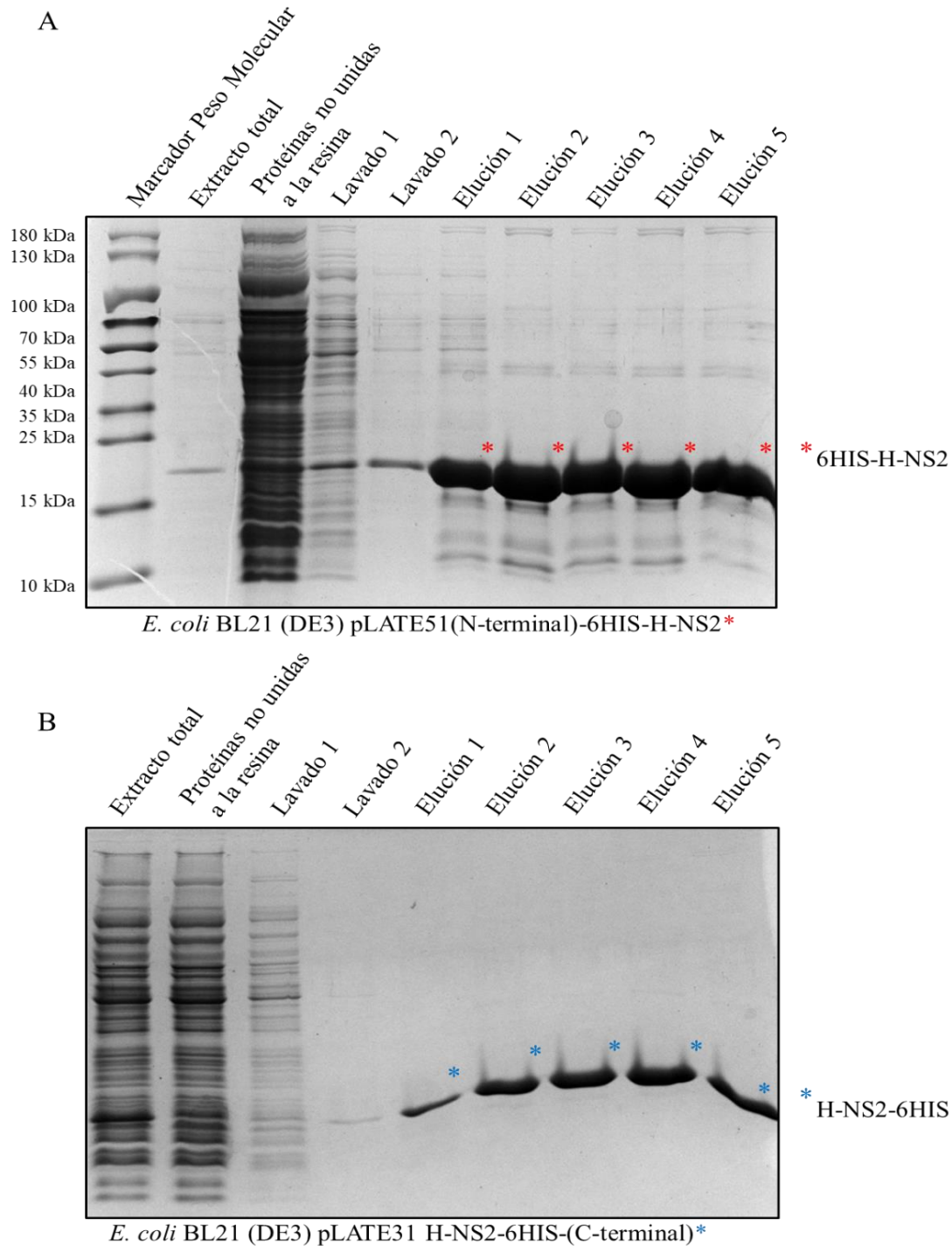
**Figura 4.30.** Análisis por *Western Blot* de la expresión de la proteína H-NS2::3XFLAG en un fondo genético salvaje (042 *hns2*::3XFLAG) y un fondo genético *hns*<sup>-</sup> (042  $\Delta$ *hns* *hns2*::3XFLAG) en medio LB tanto a 25°C (A y B) como a 37°C (C y D), en las fases de crecimiento exponencial (D.O<sub>600nm</sub> 0,4) y estacionaria (D.O<sub>600nm</sub> 2,0). El panel superior muestra la tinción de Coomassie de cada uno de los extractos proteicos analizados y representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas.

### 4.2.7 Perfil de interacción de la proteína H-NS2 con otras proteínas de la cepa *E. coli* 042

Ya que la proteína H-NS es capaz de interactuar con otras proteínas (entre las que destacan Hha, YdgT y StpA) formando de esta manera heterocomplejos proteicos que presentan una gran capacidad para regular eficientemente genes adquiridos por transferencia horizontal (Ali y col., 2013), se decidió investigar con qué otras proteínas es capaz de interactuar la proteína H-NS2. Para ello, se clonó el gen *hns2* en los plásmidos pLATE31 y pLATE51 (*Thermo Scientific*), a través de los cuales se introdujo una cola de 6 histidinas en los extremos C- y N-terminal (respectivamente) del gen clonado. Dichos plásmidos se introdujeron en la cepa *E. coli* BL21 (DE3), donde se sobreexpresaron las correspondientes proteínas 6His-H-NS2 (N-terminal) y H-NS2-6His (C-terminal) bajo el control del promotor de la ARN polimerasa T7 inducible por IPTG. Los extractos proteicos obtenidos se mezclaron con una matriz de Ni<sup>2+</sup>-NTA agarosa y, tras permitir la unión de las proteínas a la resina a través de la interacción por afinidad entre las histidinas y los iones níquel (y posteriormente lavar las proteínas no unidas), se eluyeron las proteínas retenidas en la matriz mediante la adición de un tampón con concentraciones crecientes de imidazol (200 mM, 500 mM y 1 M). Las diferentes fracciones obtenidas en el proceso de purificación se analizaron mediante electroforesis en geles desnaturizantes de poliacrilamida en Tris-Tricina-SDS, en los que las proteínas se tiñeron posteriormente con Azul de Coomassie (figura 4.31).

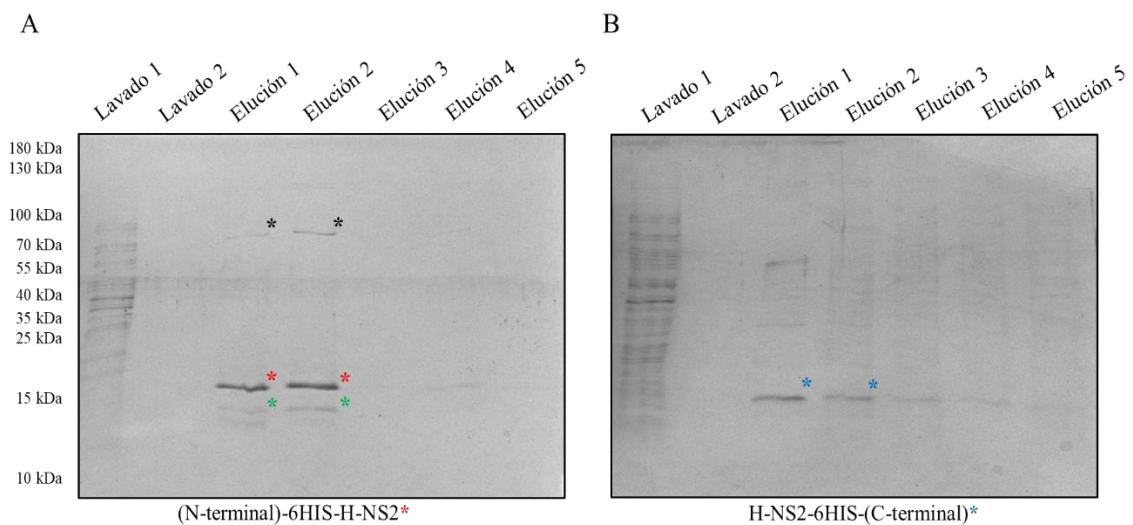
En la fracción correspondiente al extracto proteico total de la cepa *E. coli* BL21 (DE3) conteniendo los plásmidos pLATE51 6HIS-H-NS2 y pLATE31 H-NS2-6HIS se puede observar una banda de mayor intensidad correspondiente a la sobreexpresión de la proteína H-NS2 (la proteína muestra un tamaño mayor al teórico en función de su secuencia aminoacídica debido a la adición de la cola de 6 histidinas). En las fracciones correspondientes a la elución de las proteínas unidas a la resina de níquel, además de la proteína H-NS2 purificada (proteína más abundante de la fracción, siendo la purificación de la proteína con las 6 histidinas en el extremo N-terminal más eficiente ya que ha permitido obtener más proteína purificada que la construcción con las 6 histidinas en el extremo C-terminal), se pueden observar otras bandas proteicas correspondientes a proteínas de la cepa *E. coli* BL21 (DE3) que han coeluido con la propia proteína H-NS2.

#### 4. RESULTADOS



**Figura 4.31.** Electroforesis en gel de poliacrilamida de las diferentes fracciones proteicas obtenidas durante el proceso de purificación de la proteína H-NS2 fusionada con un epítipo de 6 histidinas en sus extremos N-terminal (A) y C-terminal (B). El proceso de sobreexpresión proteica se realizó en la cepa *E. coli* BL21 (DE3). Se muestran las fracciones correspondientes al extracto total de proteínas de la célula, las proteínas no unidas a la resina después de su interacción con la resina de  $\text{Ni}^{2+}$ -agarosa, dos lavados de la resina con imidazol 50 mM, y las eluciones proteicas correspondientes (1-3 se realizaron con imidazol 200 mM, la 4ª elución con 500 mM y la 5ª elución con 1 M de imidazol) en las que se pueden observar las bandas correspondientes a la proteína purificada, identificadas mediante un asterisco.

Sin embargo, ya que el objetivo era poder definir con qué proteínas de la cepa *E. coli* 042 es capaz de interactuar la proteína H-NS2, las proteínas eluidas y purificadas con imidazol 200 mM (3 eluciones en total, marcadas con un asterisco en la figura 4.31), se mezclaron con un extracto proteico total de las proteínas de *E. coli* 042, se dejaron interactuar durante 16 horas a 4°C y se purificaron de manera análoga al procedimiento anterior mediante una cromatografía de afinidad con una matriz de Ni<sup>2+</sup>-NTA agarosa. Las diferentes fracciones obtenidas en este nuevo proceso de purificación se separaron mediante electroforesis en geles desnaturalizantes de poliacrilamida en Tris-Tricina-SDS, en el que las proteínas se tiñeron posteriormente con Azul de Coomassie (figura 4.32).



**Figura 4.32.** Electroforesis en gel de poliacrilamida en el que se muestra el resultado de la purificación de la proteína H-NS2 junto con las proteínas de *E. coli* 042. Las proteínas H-NS2 sobreexpresadas y purificadas con las 6 histidinas en su extremo N-terminal (A) o C-terminal (B) se dejaron interactuar con un extracto proteico total de la cepa *E. coli* 042 y se purificaron mediante cromatografía de afinidad. Se muestran las fracciones correspondientes a dos lavados con imidazol 50 mM de las proteínas no unidas a la resina y las eluciones proteicas correspondientes (1-3 con imidazol 200 mM, la 4ª elución con 500 mM y la 5ª elución con 1 M de imidazol). Con un asterisco en rojo se muestra la proteína H-NS2 con la fusión de las 6 histidinas en su extremo N-terminal; con un asterisco en azul se indica la proteína H-NS2 con las 6 histidinas en su extremo C-terminal.

Con el fin de identificar qué proteínas de *E. coli* 042 habían copurificado con la proteína H-NS2 durante el proceso de sobreexpresión y purificación, las fracciones correspondientes a las eluciones 1 y 2 de ambas purificaciones (figura 4.32 (A) y (B)) se enviaron al servicio de la “Plataforma de Proteómica del Parc Científic de Barcelona”. Las proteínas presentes en cada una de las eluciones se digirieron con tripsina, y los péptidos resultantes se purificaron mediante cromatografía líquida acoplada a un espectrómetro de masas. Mediante los *software Thermo Proteome Discoverer (v.1.4.1.14)*



#### 4. RESULTADOS

y *Sequest HT* se comparó el espectro de masas de cada uno de los péptidos obtenidos para identificar las proteínas correspondientes en la base de datos *Swiss Prot*.

En la tabla 4.4 se indican las proteínas codificadas en el genoma de *E. coli* 042 que han coeluido (y que por tanto interaccionan) con la proteína H-NS2 durante el proceso de purificación (se representan tanto las proteínas que han coeluido con la proteína 6HIS-(N-Terminal)-H-NS2 como las que han coeluido con la proteína H-NS2-6HIS-(C-terminal)).

Proteínas copurificadas con H-NS2	Score*	PSM**	Coverage (%) ***
(N-Terminal)-6HIS-H-NS2*	1810,19	719	71,11
H-NS2-6HIS-(C-Terminal)*	1302,40	535	75,56
Proteasa Lon*	1075,31	396	72,96
Proteína de unión al ADN H-NS*	718,66	298	69,34
Regulador transcripcional YqjI	461,56	221	79,73
ADN topoisomerasa 1	104,95	50	44,62
Proteína de unión al ADN StpA	93,70	45	74,18
Poliribonucleótido nucleotidiltransferasa Pnp	54,29	19	35,72
Helicasa IV	51,51	18	39,47
Proteína asociada a la recombinación RdgC	26,79	9	24,75

\*Abundancia relativa de las proteínas presentes en la muestra.

\*\**Peptide Spectrum Match (PSM)*: Parámetro que muestra el número total de secuencias peptídicas identificadas para cada proteína.

\*\*\* Indicador que muestra el porcentaje de la secuencia de la proteína cubierta por los péptidos identificados por LC/MS (cromatografía líquida/espectrometría de masas).

**Tabla 4.4.** Resultados de la identificación por espectrometría de masas de las proteínas de *E. coli* 042 con las que interacciona la proteína H-NS2. Se representan globalmente los resultados correspondientes a todas las proteínas que han coeluido tanto con la proteína 6HIS-H-NS2 como con la proteína H-NS2-6HIS.

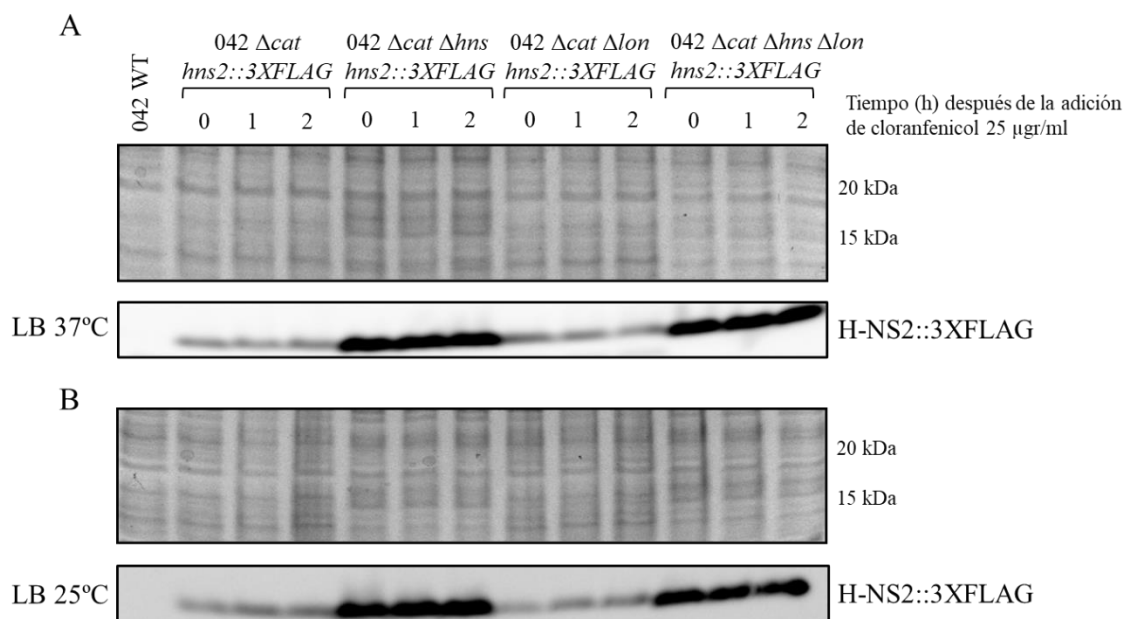
Las proteínas que presentan una mayor abundancia relativa en las eluciones analizadas (además de las dos construcciones de la proteína H-NS2 sobreexpresada y purificada) son H-NS y Lon. Además, también se han podido identificar las proteínas StpA y el regulador transcripcional YqjI como otras proteínas con las que H-NS2 tiene la capacidad de interaccionar. Sin embargo, bajo esta aproximación experimental, no se pudo encontrar interacción alguna entre H-NS2 y la proteína Hha. Ello puede ser debido a que, como la concentración intracelular de la proteína Hha es relativamente baja, para poder identificarla como una proteína de interacción en experimentos de purificación de proteínas por cromatografía de afinidad se requiere la sobreexpresión de la misma (Nieto y col., 2000). Bajo este precepto, en experimentos anteriores realizados por nuestro grupo de investigación con el fin de identificar nuevas parejas de interacción de la proteína Hha, sí se logró identificar también la proteína H-NS2 (resultados no publicados).



Para validar los datos obtenidos analizando las fracciones eluidas, se recortaron del gel representado en la figura 4.32 (A) las dos bandas que presentaban una mayor intensidad (además de la banda correspondiente a la proteína H-NS2 purificada) y se enviaron de nuevo al servicio de la “Plataforma de Proteómica del Parc Científic de Barcelona” donde, de manera análoga a las eluciones líquidas de las proteínas en imidazol 200 mM, se identificaron por espectrometría de masas. Tal y como era de esperar en función de los valores de abundancia relativa observados en la tabla 4.4, dichas proteínas, marcadas con un asterisco en verde y negro en la figura 4.32 (A), corresponden a las proteínas H-NS y Lon, respectivamente.

Este patrón de interacción que presenta la proteína H-NS2 es muy similar al descrito para la proteína StpA. En referencia a la interacción de la proteína StpA con la proteasa Lon, esta proteína es sensible a la degradación por parte de esta familia de proteasas, principalmente en ausencia de H-NS. La formación de los heterodímeros H-NS-StpA aporta a esta última un mecanismo de protección contra la degradación por parte de estas proteasas (Johansson y Uhlin, 1999; Johansson y col., 2001). Por todo ello, ya que H-NS2 también presenta la capacidad de interactuar con la proteasa Lon, se planteó la posibilidad de analizar si esta proteína también es sensible a la degradación proteolítica catalizada por esta proteasa. Con este fin, se realizaron ensayos de estabilidad proteica *in vivo*, siguiendo el protocolo descrito por (Geuskens y col., 1992). La estabilidad de la proteína H-NS2::3XFLAG se monitorizó después de inhibir la síntesis proteica total mediante la adición de cloranfenicol (antibiótico bacteriostático que actúa como un inhibidor de la subunidad 50S del ribosoma) a una concentración final de 25 µg/ml a los cultivos bacterianos crecidos hasta el inicio de la fase estacionaria de crecimiento en LB (ya que la cepa *E. coli* 042 presenta resistencia a cloranfenicol, para poder realizar este experimento previamente se delecionó el gen *cat*, presente en el transposón *Tn2411* y responsable de conferir resistencia a este antibiótico). Se tomó una muestra de cada cultivo a tiempo 0 (antes de añadir el antibiótico) y 1 y 2 horas después de añadir el antibiótico e inhibir la traducción proteica, para posteriormente extraer el contenido proteico total de cada cultivo y realizar una inmunodetección mediante *Western Blot* para analizar los niveles de expresión de la proteína H-NS2 a lo largo del tiempo. La estabilidad de la proteína H-NS2::3XFLAG se analizó en los fondos genéticos salvaje, *hns<sup>-</sup>*, *lon<sup>-</sup>* y *hns<sup>-</sup>lon<sup>-</sup>*. En la figura 4.33 se muestran los resultados de los experimentos de estabilidad de la proteína H-NS2 en medio LB tanto a 37°C como a 25°C.

#### 4. RESULTADOS

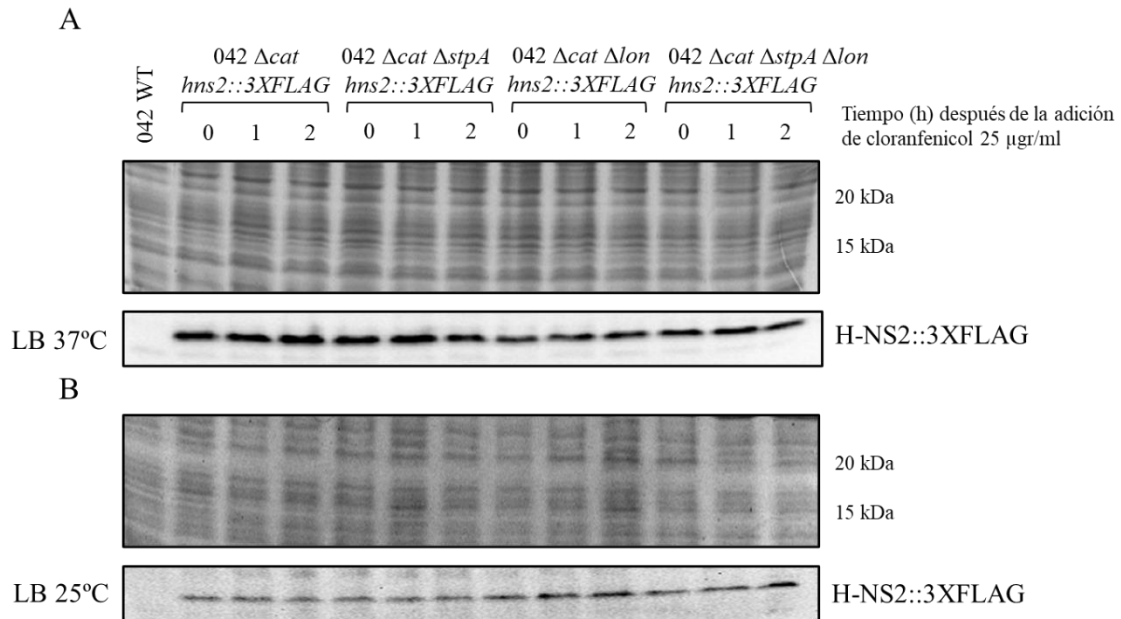


**Figura 4.33.** Análisis por *Western Blot* de la estabilidad proteica de la proteína H-NS2::3XFLAG en un fondo genético salvaje (042 *hns2::3XFLAG*), *hns*<sup>-</sup> (042  $\Delta hns$  *hns2::3XFLAG*), *lon*<sup>-</sup> (042  $\Delta lon$  *hns2::3XFLAG*) y *hns*<sup>-</sup>*lon*<sup>-</sup> (042  $\Delta hns \Delta lon$  *hns2::3XFLAG*). Las condiciones de cultivo analizadas fueron medio de crecimiento rico LB tanto a 37°C (A) como a 25°C (B) de cultivos bacterianos crecidos hasta el inicio de la fase estacionaria de crecimiento. Las muestras se analizaron después de inhibir la síntesis proteica total al añadir cloranfenicol a una concentración final de 25  $\mu$ g/ml. Los paneles superiores muestran la tinción de Coomassie realizada para cada uno de los extractos analizados, representando un control de carga proteica homogénea para cada una de las condiciones analizadas.

Se puede observar cómo, a diferencia de lo descrito en el caso de la proteína StpA, los niveles de expresión de la proteína H-NS2 se mantienen relativamente estables después de inhibir la traducción proteica en un fondo genético salvaje (con la presencia de la proteasa Lon). Dicha estabilidad también se mantiene en un fondo genético sin la proteína H-NS, situación en la que, tal y como se ha descrito anteriormente (figura 4.30), los niveles de expresión de la proteína H-NS2 se incrementan para tratar de compensar la pérdida de la proteína H-NS.

Ya que no se pudo observar degradación alguna de la proteína H-NS2 en un fondo genético *hns*<sup>-</sup>, se quiso analizar si, en la cepa *E. coli* 042, el mecanismo protector sobre la proteína H-NS2 contra degradación por parte de la proteasa Lon lo proporcionaba la presencia de la proteína StpA. Para ello, de manera análoga a los experimentos anteriores, se estudió también la estabilidad proteica de la proteína H-NS2 en un mutante que no expresara la proteína StpA (*E. coli* 042  $\Delta stpA$  *hns2::3XFLAG*). Bajo esta premisa, se

analizó la estabilidad de la proteína H-NS2 en los fondos genéticos salvaje, *stpA*<sup>-</sup>, *lon*<sup>-</sup> y *stpA*<sup>-</sup>*lon*<sup>-</sup> bajo las mismas condiciones experimentales que el experimento anterior.

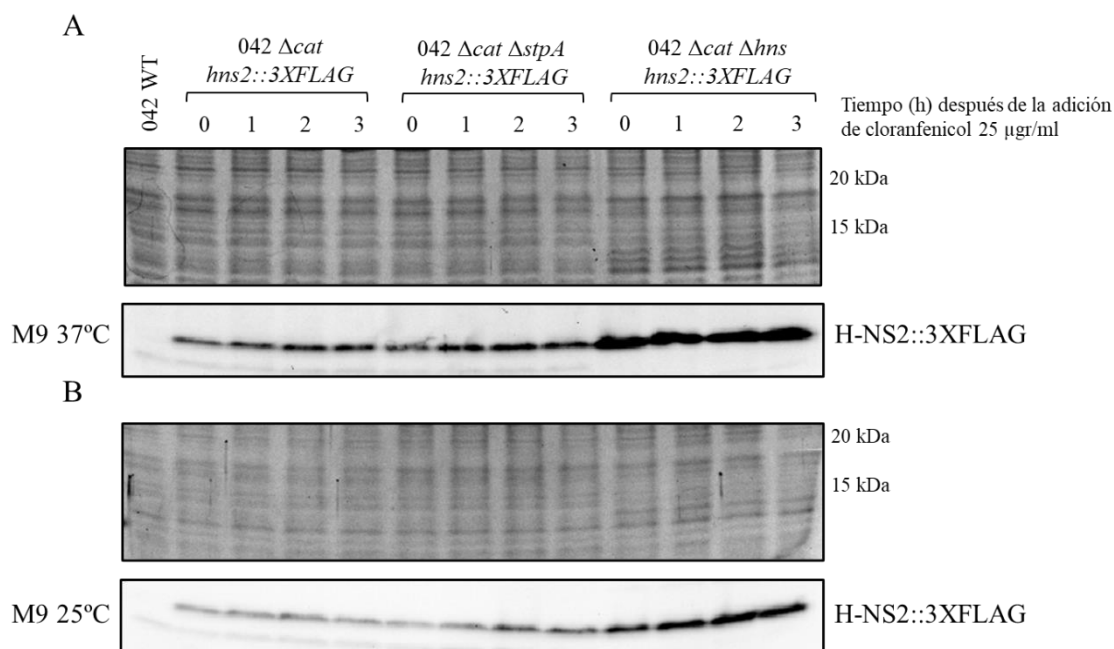


**Figura 4.34.** Análisis por *Western Blot* de la estabilidad proteica de la proteína H-NS2::3XFLAG en un fondo genético salvaje (042 *hns2::3XFLAG*), *stpA*<sup>-</sup> (042  $\Delta stpA$  *hns2::3XFLAG*), *lon*<sup>-</sup> (042  $\Delta lon$  *hns2::3XFLAG*) y *stpA*<sup>-</sup>*lon*<sup>-</sup> (042  $\Delta stpA \Delta lon$  *hns2::3XFLAG*). Las condiciones de cultivo analizadas fueron medio de crecimiento rico LB tanto a 37°C (A) como a 25°C (B) de cultivos bacterianos crecidos hasta el inicio de la fase estacionaria de crecimiento. La cantidad de proteína se analizó después de inhibir la síntesis proteica al añadir cloranfenicol al medio de cultivo a una concentración final de 25  $\mu$ gr/ml. El panel de arriba muestra la tinción de Coomassie de cada uno de los extractos proteicos analizados y representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas.

En este caso (figura 4.34), tampoco se pudo observar degradación alguna de la proteína H-NS2 en ausencia de la proteína StpA. Además, a diferencia de lo que ocurría en el caso del fondo genético *hns*<sup>-</sup> (donde se produce la sobreexpresión de la proteína H-NS2), la ausencia del gen *stpA* no provoca cambio alguno en la expresión de H-NS2.

Ya que no se vio afectación alguna en la estabilidad de la proteína H-NS2 en medio de crecimiento rico independientemente de la presencia o ausencia de las proteínas H-NS o StpA, se optó por analizar la estabilidad de esta proteína en otras condiciones de crecimiento. Concretamente, se realizaron nuevos ensayos de estabilidad proteica *in vivo* a partir de cultivos crecidos en medio mínimo M9 suplementado con glucosa al 0,40% (p/v) (figura 4.35).

#### 4. RESULTADOS



**Figura 4.35.** Análisis por *Western Blot* de la estabilidad proteica de la proteína H-NS2::3XFLAG en un fondo genético salvaje (042 *hns2::3XFLAG*), *stpA*<sup>-</sup> (042  $\Delta stpA$  *hns2::3XFLAG*) y *hns*<sup>-</sup> (042  $\Delta hns$  *hns2::3XFLAG*). Las condiciones de cultivo analizadas fueron medio de crecimiento mínimo M9 (suplementado con glucosa al 0,40% (p/v)) tanto a 37°C (A) como a 25°C (B) de cultivos bacterianos crecidos hasta el inicio de la fase estacionaria de crecimiento. La cantidad de proteína se monitorizó después de inhibir la síntesis proteica al añadir cloranfenicol a una concentración final de 25  $\mu$ g/ml. El panel de arriba muestra la tinción de Coomassie de cada uno de los extractos proteicos analizados y representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas.

Bajo estas condiciones de crecimiento, tampoco se observó afectación alguna a nivel de la tasa de degradación proteolítica de la proteína H-NS2, ya que los niveles de proteína se mantienen estables y constantes incluso 3 horas después de inhibir la traducción proteica. Por todo ello, independientemente de la presencia o ausencia de las proteínas H-NS o StpA, la proteína H-NS2 no es sensible a la degradación proteolítica por parte de la proteasa Lon, por lo que sería necesario realizar más experimentos para tratar de entender cuál es el significado biológico de la interacción física entre H-NS2 y esta proteasa.

### 4.2.8 Efectos de las mutaciones en los genes *hns* y *hns2* sobre el transcriptoma global de la cepa *E. coli* 042

Para tratar de profundizar en el papel regulador de la expresión génica de las proteínas H-NS y H-NS2 en la cepa *E. coli* 042 se analizó el perfil global del transcriptoma derivado de las mutaciones en los genes *hns*, *hns2* y en el doble mutante *hns hns2*, comparándose a su vez con el perfil transcriptómico de la cepa *E. coli* 042 WT. Estos ensayos transcriptómicos se realizaron a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento ( $D.O_{600nm}=2,0$ ), donde el aislamiento del ARN y el procesamiento de las muestras para obtener los datos relativos a la expresión de cada gen se realizaron en colaboración con la empresa *Vertis Biotechnologie AG* (*Freising-Weihenstephan*, Alemania), siguiendo el protocolo descrito en el apartado 3.8.5 de Materiales y Métodos.

Los resultados de este experimento se encuentran publicados en su totalidad en la base de datos GEO (*Gene Expression Omnibus*) *Sequence Read Archive* del NCBI bajo el número de acceso GSE105133, además de en el apartado 9.2 del Anexo de esta memoria, en las tablas 9.12, 9.13 y 9.14.

Teniendo en cuenta que el principal papel regulador de la expresión génica por parte de la proteína H-NS tiene lugar a través de la represión de la misma, en estos análisis transcriptómicos nos centraremos en aquellos genes que en ausencia de la proteína H-NS (o su parálogo H-NS2) su expresión se ve aumentada.

En la tabla 4.5 (A) se representa el listado de genes que muestran un mayor nivel de sobreexpresión en un mutante  $\Delta hns$ . En dicha tabla también se indica el nivel de expresión relativa de dichos genes en los fondos genéticos  $\Delta hha\Delta hha2$ ,  $\Delta hns2$  y  $\Delta hns\Delta hns2$ .

#### 4. RESULTADOS

Identificador NCBI ( <i>locus tag</i> )	<i>Fold change</i>			
	$\Delta hns$	$\Delta hha\Delta hha2$	$\Delta hns2$	$\Delta hns\Delta hns2$
EC042_RS08585 (gadB glutamate decarboxylase beta subunit)	43,0	2,9	3,0	87,8
EC042_RS20290 (gadA glutamate decarboxylase alpha subunit)	37,6	3,1	4,5	59,8
EC042_RS08580 (gadC glutamate/gamma-aminobutyrate antiporter)	34,7	3,4	4,3	68,3
EC042_RS01725 (matB putative fimbrial protein)	33,3	3,3	4,1	108,8
EC042_RS20250 (hdeA putative periplasmic acid stress chaperone)	26,1	-	-	60,3
EC042_RS01730 (matA putative fimbrial transcriptional regulator)	22,7	4,2	4,0	70,0
EC042_RS05780 (tRNA-Ser)	19,3	3,4	4,8	3,0
EC042_RS15335 (stpA DNA-binding protein)	18,6	4,4	3,1	83,7
EC042_RS20255 (hdeD putative acid resistance protein)	17,8	2,5	3,5	36,7
EC042_RS20260 (gadE transcriptional regulator)	16,6	3,2	4,5	32,9
EC042_RS06765 (tRNA-Tyr)	16,3	7,8	5,6	-
EC042_RS13160 (hypothetical protein)	15,7	2,6	-	2,4
EC042_RS21295 (CesD/SycD/LcrH family type III secretion system chaperone)	15,6	7,6	3,9	4,2
EC042_RS01715 (putative fimbrial outer membrane usher protein)	15,1	4,1	4,8	29,0
EC042_RS20245 (acid-activated periplasmic chaperone hdeB)	15,1	-	2,4	37,1
EC042_RS06770 (tRNA-Tyr)	14,8	6,6	3,8	-
EC042_RS04580 (hypothetical protein)	14,5	-	-	2,4
EC042_RS01720 (putative fimbrial protein)	14,2	5,4	7,0	28,0
EC042_RS20090 (hypothetical protein)	14,0	4,0	5,4	25,3
EC042_RS07750 (type I toxin-antitoxin Iar)	13,9	4,2	13,9	5,5
EC042_RS01270 (hypothetical protein)	13,6	3,7	-	6,7
EC042_RS24770 (tRNA-Gly)	13,5	-	-	-
EC042_RS09275 (hypothetical protein)	13,3	10,3	11,8	2,6
EC042_RS01065 (tRNA-Ala)	12,8	3,3	2,0	2,5
EC042_RS18940 (tRNA-Ala)	12,8	3,0	2,1	2,8
EC042_RS26795 (hypothetical protein)	12,2	3,7	6,3	-
EC042_RS16270 (hypothetical protein)	12,2	5,7	4,1	5,0
EC042_RS11015 (tRNA-Leu)	12,1	3,5	3,3	-
EC042_RS11620 (tRNA-Arg)	12,0	3,1	2,7	2,5

**Tabla 4.5 (A).** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. También se indican los valores de expresión relativa de dichos genes en los mutantes  $\Delta hha\Delta hha2$ ,  $\Delta hns2$  y  $\Delta hns\Delta hns2$ . Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.



Los datos de la tabla anterior muestran que la ausencia del gen *hns* provoca la sobreexpresión de una gran número de genes implicados en la respuesta bacteriana a pH ácido (*gadA*, *gadB*, *gadC*, *gadE*, *hdeA*, *hde*, etc), además de genes implicados en la síntesis de fimbrias (*matA*, EC042\_RS01715, EC042\_RS01720, etc) y genes codificantes por ARNs de transferencia, tal y como se ha descrito en el caso del mutante  $\Delta hha$  (García-Contreras y col., 2008) y en el transcriptoma global del mutante  $\Delta hha2$  (tabla 4.3 (B)). Además, muchos de los genes regulados por H-NS también se encuentran diferencialmente expresados en un mutante  $\Delta hns2$  y en el doble mutante  $\Delta hha\Delta hha2$ , hecho que demuestra fenómenos de co-regulación de la expresión génica de un gran número de genes por parte de las proteínas H-NS, H-NS2, Hha y Hha2. Cabe destacar además que, para la gran mayoría de genes, el nivel de sobreexpresión observado en el mutante  $\Delta hns$  se acentúa de manera mucho más notoria en el caso del doble mutante  $\Delta hns\Delta hns2$ . Es importante también mencionar aquí que algunos de los genes regulados por H-NS no están regulados ni por H-NS2 ni por Hha.

En la tabla 4.5 (B) se representan los genes mayormente sobreexpresados bajo el fondo genético  $\Delta hns2$  respecto a la cepa salvaje. En dicha tabla también se indica el nivel de expresión relativa de dichos genes en los fondos genéticos  $\Delta hha\Delta hha2$ ,  $\Delta hns$  y  $\Delta hns\Delta hns2$ .

Identificador NCBI ( <i>locus tag</i> )	Fold change			
	$\Delta hns2$	$\Delta hha\Delta hha2$	$\Delta hns$	$\Delta hns\Delta hns2$
EC042_RS05870 (hypothetical protein)	12,1	7,7	10,1	-
EC042_RS09275 (hypothetical protein)	11,8	10,3	13,3	2,6
EC042_RS07145 (hypothetical protein)	11,3	10,2	6,0	-
EC042_RS26995 (hypothetical protein)	11,3	8,6	6,1	-
EC042_RS18795 (hypothetical protein)	11,2	8,6	8,8	-
EC042_RS11295 (hypothetical protein)	10,3	10,6	11,4	2,0
EC042_RS13835 (tRNA-Arg)	10,1	8,5	8,6	-
EC042_RS24055 (antitoxin hypothetical protein)	9,6	6,9	4,9	-
EC042_RS22745 (DUF3521 domain-containing protein)	9,6	6,2	6,1	-
EC042_RS03185 (hypothetical protein)	9,3	7,1	8,1	-
EC042_RS17205 (hypothetical protein)	9,2	6,9	4,3	-
EC042_RS17210 (hypothetical protein)	9,1	7,4	6,4	-
EC042_RS03005 (envY porin thermoregulatory protein)	9,0	6,5	5,4	-
EC042_RS24455 (hypothetical protein)	8,8	8,8	3,9	-



#### 4. RESULTADOS

EC042_RS17185 (putative ATP/GTP-binding protein)	8,7	6,3	4,8	-
EC042_RS17230 (hypothetical protein)	8,6	6,0	4,4	-
EC042_RS21255 (putative prophage protein)	8,6	5,8	5,9	-
EC042_RS25650 (putative RadC-like DNA repair protein)	8,5	7,1	6,2	-
EC042_RS18375 (hypothetical protein)	8,5	6,5	7,2	-
EC042_RS17220 (putative DNA repair protein)	8,5	6,9	6,1	-
EC042_RS11730 (hypothetical protein)	8,4	8,5	8,5	2,2
EC042_RS24065 (putative plasmid-like protein)	8,4	5,2	5,1	-
EC042_RS17215 (putative antirestriction protein)	8,4	6,9	5,7	-
EC042_RS07390 (hypothetical protein)	8,3	7,7	7,4	-
EC042_RS12065 (hypothetical protein)	8,3	6,2	8,2	2,3
EC042_RS13250 (hypothetical protein)	8,2	5,0	4,5	-
EC042_RS20520 (putative acetyltransferase)	8,1	4,5	4,3	-
EC042_RS23195 (hypothetical protein)	8,0	5,9	5,1	-
EC042_RS17225 (hypothetical protein)	8,0	6,1	4,9	-

**Tabla 4.5 (B).** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. También se indican los valores de expresión relativa de dichos genes en los mutantes  $\Delta hha\Delta hha2$ ,  $\Delta hns$  y  $\Delta hns\Delta hns2$ . Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

Tal y como se puede observar, los genes que presentan un mayor nivel de sobreexpresión en el mutante sencillo  $\Delta hns2$  (mayoritariamente proteínas hipotéticas sin función atribuida adquiridas a través de mecanismos de transferencia génica horizontal), también se encuentran altamente desregulados en el mutante sencillo  $\Delta hns$  y en el doble mutante  $\Delta hha\Delta hha2$ , indicando de nuevo fenómenos de co-modulación de la expresión génica por parte de este grupo de proteínas. Curiosamente, en el caso del doble mutante  $\Delta hns\Delta hns2$ , estos efectos se atenúan hasta no llegarse a detectar expresión diferencial alguna en dicho fondo genético. A diferencia de lo que sucede en el caso de la proteína H-NS, todos los genes regulados por H-NS2 también se encuentran regulados por H-NS y por Hha.

En la tabla 4.5 (C) se representan los genes mayormente sobreexpresados en el fondo genético  $\Delta hns\Delta hns2$  respecto a la cepa salvaje. En dicha tabla también se indica el nivel de expresión relativa de dichos genes en los fondos genéticos  $\Delta hha\Delta hha2$ ,  $\Delta hns$  y  $\Delta hns2$ .

Identificador NCBI ( <i>locus tag</i> )	Fold change			
	$\Delta hns\Delta hns2$	$\Delta hha\Delta hha2$	$\Delta hns$	$\Delta hns2$
EC042_RS01725 (matB putative fimbrial protein)	108,8	3,3	33,3	4,1
EC042_RS08585 (gadB glutamate decarboxylase beta subunit)	87,8	2,9	43,0	3,0
EC042_RS15335 (stpA DNA-binding protein)	83,7	4,4	18,6	3,1
EC042_RS07290 (gsiB hypothetical protein)	77,6	2,9	9,3	3,3
EC042_RS01730 (matA putative fimbrial transcriptional regulator)	70,0	4,2	22,7	4,0
EC042_RS08580 (gadC glutamate/gamma-aminobutyrate antiporter)	68,3	3,4	34,7	4,3
EC042_RS20250 (hdeA putative periplasmic acid stress chaperone)	60,3	-	26,1	-
EC042_RS20290 (gadA glutamate decarboxylase alpha subunit)	59,8	3,1	37,6	4,1
EC042_RS08660 (fimbrial protein)	52,8	13,1	10,3	-
EC042_RS11605 (anti-adapter protein IraM)	41,1	4,9	10,0	-
EC042_RS20245 (acid-activated periplasmic chaperone hdeB)	37,1	-	15,1	2,4
EC042_RS20255 (hdeD putative acid resistance protein)	36,7	2,5	17,8	3,5
EC042_RS20260 (gadE transcriptional regulator)	32,9	3,2	16,6	4,5
EC042_RS07280 (ferritin-like domain-containing protein)	31,3	4,7	8,8	4,8
EC042_RS07285 (ferritin-like domain-containing protein)	30,4	4,1	8,0	3,9
EC042_RS01715 (putative fimbrial outer membrane usher protein)	29,0	4,1	15,1	4,8
EC042_RS01720 (putative fimbrial protein)	28,0	5,4	14,2	7,0
EC042_RS20090 (hypothetical protein membrane protein)	25,3	4,0	14,0	5,4
EC042_RS11225 (transcriptional regulator rcsA)	25,2	2,6	10,6	3,1
EC042_RS17085 (mchD putative microcin H47 activating protein)	25,0	3,4	6,2	2,0
EC042_RS03395 (lipid IVA palmitoyltransferase)	23,3	2,7	8,3	3,9
EC042_RS12270 (wcaF putative colanic acid biosynthesis acetyltransferase)	23,3	2,4	9,8	3,8
EC042_RS12265 (gmd GDP-mannose 4,6-dehydratase)	22,9	3,1	10,5	4,3
EC042_RS24110 (hypothetical protein)	22,6	-	4,2	-
EC042_RS26710 (hypothetical protein)	22,5	-	4,4	-
EC042_RS09025 (cspI cold shock-like protein)	21,0	5,0	10,9	4,4
EC042_RS01710 (fimbria adhesin EcpD)	20,6	2,5	11,3	3,2
EC042_RS17080 (mchE microcin H47 secretion protein colicin V secretion protein CvaA)	20,6	3,6	8,1	3,0
EC042_RS08655 (fimbrial chaperone)	19,6	9,3	7,7	4,8
EC042_RS19635 (aaspG fimbrial adhesin)	19,4	2,6	6,1	3,0

#### 4. RESULTADOS

EC042_RS08860 (cold-shock protein)	18,8	6,4	11,1	5,9
EC042_RS01210 (putative type VI secretion system protein)	18,8	6,0	8,7	6,8
EC042_RS20190 (slp outer membrane lipoprotein)	18,6	-	9,3	-
EC042_RS12275 (wcaE putative colanic acid biosynthesis glycosyl transferase)	17,4	-	7,4	3,0
EC042_RS26340 (hypothetical protein)	16,8	5,9	6,0	2,1
EC042_RS12170 (ugd UDP-glucose 6-dehydrogenase)	16,5	3,1	7,6	3,7
EC042_RS20280 (gadW AraC family transcriptional regulator)	16,4	2,7	7,4	3,0
EC042_RS01705 (putative fimbrial protein)	16,2	3,0	8,3	4,1
EC042_RS17110 (mchS1 putative microcin esterase)	15,4	3,8	4,8	3,1
EC042_RS17075 (mchF microcin H47 secretion/processing ATP-binding protein)	15,4	2,6	5,3	2,3
EC042_RS01240 (putative type VI secretion system protein)	15,3	-	3,8	-

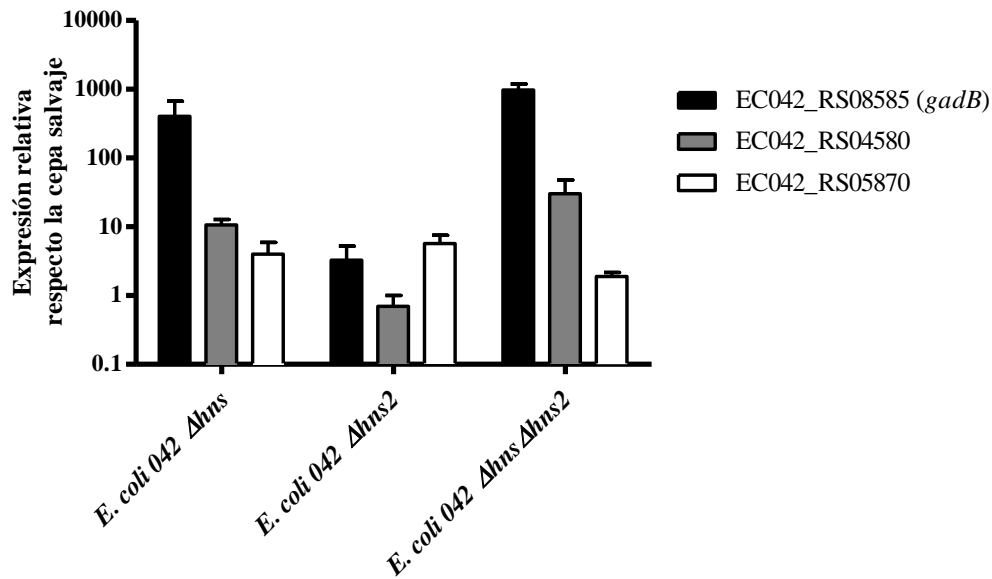
**Tabla 4.5 (C).** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns\Delta hns2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. También se indican los valores de expresión relativa de dichos genes en los mutantes  $\Delta hha\Delta hha2$ ,  $\Delta hns$  y  $\Delta hns2$ . Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

Tal y como se puede observar, los genes cuya expresión se encuentra más desregulada en el doble mutante  $\Delta hns\Delta hns2$  son aquellos que también se encontraban más desregulados en el mutante sencillo  $\Delta hns$ , pero con unos niveles de sobreexpresión mucho más elevados (el efecto de la pérdida de *hns* se exagera aún más en ausencia del gen *hns2*). Cabe destacar, también, el elevado nivel de sobreexpresión del gen *stpA* bajo este fondo genético (83,7 veces más expresado respecto al fondo genético salvaje). En ausencia de los genes *hns* y *hns2*, la proteína StpA se sobreexpresa, muy posiblemente, con el fin de compensar, al menos parcialmente, la pérdida de las proteínas H-NS y H-NS2.

#### 4.2.8.1 Validación de los datos RNA-seq mediante qRT-PCR

Con el fin de validar los datos obtenidos por RNA-seq, se analizó por qRT-PCR el nivel de expresión de varios genes que mostraban niveles de expresión diferencial en los mutantes  $\Delta hns$ ,  $\Delta hns2$  y  $\Delta hns\Delta hns2$ . Concretamente, se seleccionaron los genes EC042\_RS08585 (gen que codifica para la proteína GadB, correspondiente a la subunidad  $\beta$  de la glutamato descarboxilasa), y los genes EC042\_RS04580 y EC042\_RS05870 (genes que

codifican para proteínas hipotéticas). Para obtener estos datos, los cultivos correspondientes se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento ( $D.O_{600nm}=2,0$ ), se extrajo el ARN de los mismos y se cuantificó el nivel de expresión del ARNm de los tres genes objeto de estudio en los tres mutantes analizados respecto al nivel de expresión de dichos genes en la cepa salvaje de *E. coli* 042 (figura 4.36).



**Figura 4.36.** Expresión relativa de los genes EC042\_RS08585, EC042\_RS04580 y EC042\_RS05870 en los mutantes *hns*, *hns2* y *hns hns2* respecto a la cepa *E. coli* 042 salvaje. La expresión de los genes se analizó por qRT-PCR, utilizando los oligonucleótidos RS08585 Fw/Rv, RS04580 Fw/Rv y RS05870 Fw/Rv. Los resultados fueron normalizados mediante la detección del gen interno *gapA* mediante los oligonucleótidos *gapA042* RT Fw/Rv. Se representan los valores correspondientes a 3 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

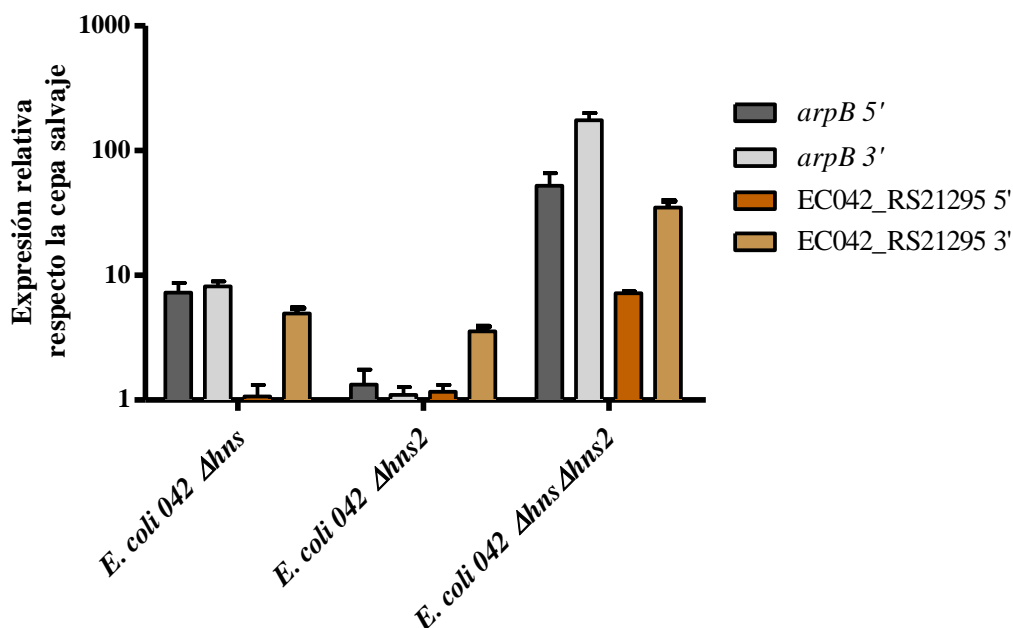
En referencia al gen EC042\_RS08585 (*gadB*), los valores obtenidos por qRT-PCR concuerdan totalmente con los datos obtenidos por RNA-seq. En el mutante *E. coli* 042  $\Delta hns$  se observa una gran incremento en la expresión de dicho gen, mientras que en el caso del mutante  $\Delta hns2$  la sobreexpresión es mucho más tenue. Dicha sobreexpresión es mucho más notoria en el caso del doble mutante  $\Delta hns\Delta hns2$ . En el caso del gen EC042\_RS04580, tanto en el análisis por RNA-seq como por qRT-PCR no se ha podido observar efecto alguno de la mutación en el gen *hns2* sobre la expresión de dicho gen. Por el contrario, sí que se observa una gran desregulación de la expresión en los mutantes  $\Delta hns$  y  $\Delta hns\Delta hns2$ . Sin embargo, existe una discordancia entre los valores de sobreexpresión

#### 4. RESULTADOS

obtenidos mediante el ensayo por qRT-PCR y el RNA-seq: mientras que en el primer caso los niveles más altos de expresión del gen se encuentran en el mutante  $\Delta hns\Delta hns2$ , mediante RNA-seq los valores más altos se observan en el mutante  $\Delta hns$ . En referencia al gen EC042\_RS05870, tanto por RNA-seq como por qRT-PCR se ha podido validar que su expresión se ve altamente incrementada en los mutantes sencillos  $\Delta hns$  y  $\Delta hns2$  (generándose en este último mutante los valores más altos de sobreexpresión), mientras que el efecto se atenúa totalmente en el doble mutante  $\Delta hns\Delta hns2$ .

##### **4.2.8.2 Efecto los genes *hns* y *hns2* sobre el control de la transcripción espuria**

Ya que la proteína H-NS está implicada en reprimir la transcripción espuria a partir de promotores inter e intragénicos (Singh y col., 2014; Grainger, 2016), también se analizó por qRT-PCR los niveles de expresión de un gen (*arpB*) sobre el que se ha demostrado experimentalmente el efecto represor de H-NS sobre la transcripción indeseada a partir de promotores internos presentes en la secuencia codificante de dicho gen (Singh y col., 2014). Para cuantificar los diferentes niveles de transcripción espuria, se analizó mediante qRT-PCR la abundancia relativa de los transcritos de ARNm mediante oligonucleótidos que hibridaban tanto en la región 5' como en la 3' del mismo. Si se produjeran fenómenos de transcripción espuria, se esperaría detectar una cantidad superior de transcrito cuando se utilizan oligonucleótidos que hibridan en el extremo 3', como consecuencia de los eventos de transcripción adicionales generados a partir de promotores internos dentro del gen. Paralelamente, también se analizó este mismo efecto por qRT-PCR en el caso del gen EC042\_RS21295, que codifica para una chaperona del sistema de secreción tipo III CesD/SycD/LcrH, y que mediante RNA-seq se ha podido observar que su expresión se encuentra sobreexpresada tanto en los mutantes sencillos  $\Delta hns$  y  $\Delta hns2$  como en el doble mutante  $\Delta hns\Delta hns2$ . En la figura 4.37 se representan los resultados de la qRT-PCR sobre los dos genes mencionados anteriormente. Tal y como se puede observar, en el caso del gen EC042\_RS21295, para los tres mutantes analizados, existe una mayor abundancia relativa del ARNm cuando la detección se realiza con los oligonucleótidos que hibridan en el extremo 3' del gen. Ello nos indica que, ante la falta de las proteínas H-NS y H-NS2, se produce un exceso de transcrito correspondiente a la región terminal del gen, por lo que dichas proteínas parecen jugar un papel represor de la transcripción no deseada a partir de promotores internos. En el caso del gen *arpB*, también se observa este proceso en el doble mutante  $\Delta hns\Delta hns2$ , pero no así en los mutantes sencillos  $\Delta hns$  y  $\Delta hns2$ .



**Figura 4.37.** Expresión relativa de los genes *arpB* y EC042\_RS21295 en los mutantes *hns*, *hns2* y *hns hns2* respecto a la cepa *E. coli* 042 salvaje. La expresión de los genes se analizó por qRT-PCR, utilizando los oligonucleótidos *arpB* 5' Fw/Rv, *arpB* 3' Fw/Rv, RS21295 5' Fw/Rv y RS21295 3' Fw/Rv. Los resultados fueron normalizados mediante la detección del gen interno de referencia *gapA* utilizando los oligonucleótidos *gapA*042 RT Fw/Rv. Se representan los valores correspondientes a 3 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

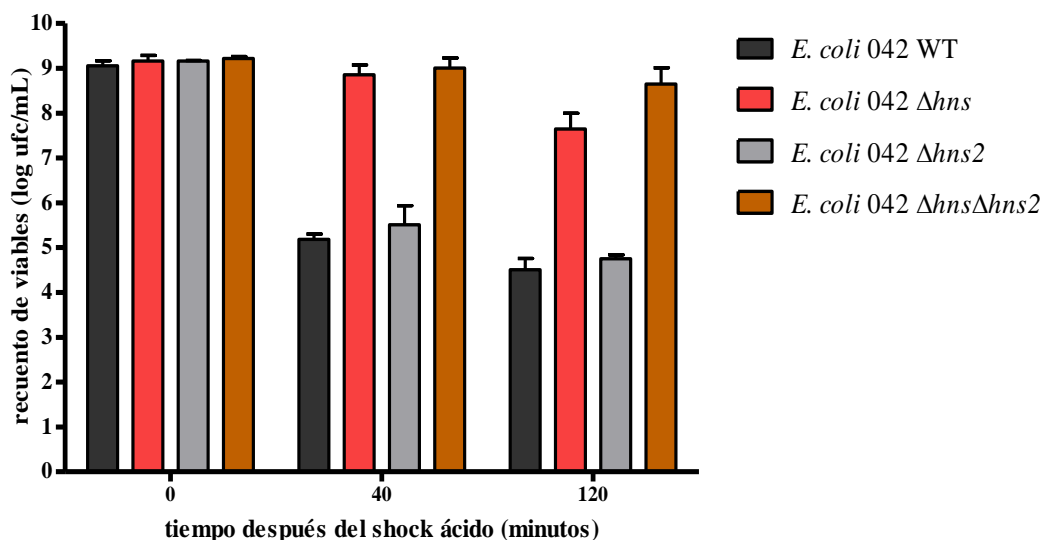
## 4.2.9 Validación a nivel fenotípico de los datos obtenidos por RNA-seq sobre las cepas WT, $\Delta hns$ , $\Delta hns2$ y $\Delta hns \Delta hns2$

### 4.2.9.1 Efecto de los genes *hns* y *hns2* en la resistencia al estrés por disminución del pH

Para poder correlacionar los datos obtenidos sobre el transcriptoma global de los mutantes *hns*, *hns2* y *hns hns2* con un fenotipo observable, se intentó analizar si alguno de los genes u operones desregulados como consecuencia de las mutaciones en los genes tipo *hns* estaban implicados en cambios fisiológicos que se pudieran estudiar experimentalmente. Teniendo en cuenta que los genes pertenecientes a los operones *gad* y *hde* presentaban unos niveles de desregulación muy elevados en el doble mutante deficiente para los genes *hns* y *hns2*, y dichos genes están implicados en la tolerancia bacteriana al estrés por pH ácido (Tucker y col., 2002), se decidió realizar un ensayo de resistencia al estrés ácido sobre las cepas *E. coli* 042 WT, *E. coli* 042  $\Delta hns$ , *E. coli* 042  $\Delta hns2$  y *E. coli* 042  $\Delta hns \Delta hns2$ . Para ello, las diferentes cepas sometidas al *shock* por estrés ácido se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria ( $D.O_{600nm} = 2,0$ ). Una vez

#### 4. RESULTADOS

llegados a este punto, se añadió HCl 6 N a los cultivos hasta que el pH de los mismos descendió a 3,2. Después de la adición del ácido, los cultivos se incubaron durante 0, 40 y 120 minutos, y se estudió la viabilidad celular (resistencia al estrés ácido) mediante siembra en placa y posterior recuento de viables (figura 4.38).



**Figura 4.38.** Representación del recuento de células viables después de someter a los cultivos bacterianos a un estrés ácido por disminución del pH del medio de cultivo. Los cultivos de las cepas de *E. coli* 042 WT, *E. coli* 042  $\Delta hns$ , *E. coli* 042  $\Delta hns2$  y *E. coli* 042  $\Delta hns\Delta hns2$  se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria, momento en el que se disminuyó el pH del medio hasta 3,2 mediante la adición de HCl 6 N. Se realizó el recuento de ufc/ml (unidades formadoras de colonias/ml) de las células resistentes al estrés ácido a tiempo 0, 40 y 120 minutos después de disminuir el pH del medio. Se representa la media y la desviación estándar de tres réplicas biológicas independientes.

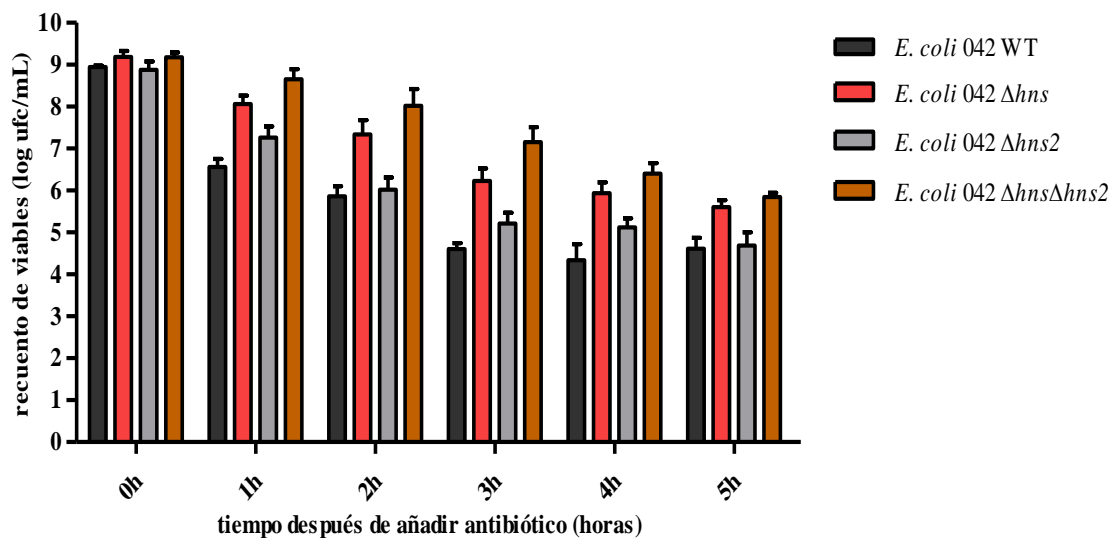
Los resultados obtenidos muestran que el mutante sencillo  $\Delta hns$  y el doble mutante  $\Delta hns\Delta hns2$  presentan una mayor tolerancia a pH ácido que la cepa salvaje y el mutante sencillo *hns2*. Estos resultados concuerdan con la sobreexpresión de los operones *hde* y *gad* observada mediante RNA-seq tanto en el mutante sencillo *hns* como en el doble mutante *hns hns2*, este último presentando unos valores más acentuados de sobreexpresión de los genes implicados en la resistencia al estrés ácido y un fenotipo de mayor resistencia y tolerancia ante pH ácido.

#### 4.2.9.2 Efecto de los genes *hns* y *hns2* en la persistencia a antibióticos

Ya que también en el mutante sencillo  $\Delta hns$  y en el doble mutante  $\Delta hns\Delta hns2$  se pudo observar la sobreexpresión de un gran número de sistemas putativos toxina-antitoxina (tabla 9.14 presente en el Anexo), y dichos sistemas, entre otras funciones, están



implicados en mecanismos de persistencia bacteriana a antibióticos, se decidió realizar un test para analizar el nivel de tolerancia a un antibiótico (al que las bacterias son sensibles) por parte de los diferentes mutantes analizados. La persistencia es un proceso mediante el cual las bacterias evaden la presencia de antibióticos en el medio (sin contener genes que permitan la resistencia a estos antibióticos) al entrar en un estado fisiológicamente inactivo (Harms y col., 2016). Para estudiar este fenómeno, las cepas *E. coli* 042 WT, *E. coli* 042  $\Delta hns$ , *E. coli* 042  $\Delta hns2$  y *E. coli* 042  $\Delta hns\Delta hns2$  se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria ( $D.O_{600nm} = 2,0$ ). Una vez las células llegaron a este punto de crecimiento, se añadió al medio de cultivo carbenicilina (análogo sintético de la ampicilina) a una concentración final de 100  $\mu\text{gr/ml}$ . Después de la adición del antibiótico, los cultivos se incubaron a 37°C en agitación constante durante 1, 2, 3, 4 y 5 horas, y se tomó una muestra de cada cultivo correspondiente en los intervalos de tiempo indicados para realizar siembra en placa y posterior recuento de células viables. Los resultados del ensayo de persistencia a antibióticos se muestran en la figura 4.39.



**Figura 4.39.** Representación del recuento de células viables después de someter a los cultivos bacterianos a un estrés por la presencia de un antibiótico. Los cultivos de las cepas de *E. coli* 042 WT, *E. coli* 042  $\Delta hns$ , *E. coli* 042  $\Delta hns2$  y *E. coli* 042  $\Delta hns\Delta hns2$  se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria, momento en el que se añadió carbenicilina a una concentración final de 100  $\mu\text{gr/ml}$ . Se realizó el recuento de ufc/ml (unidades formadoras de colonias/ml) de las células resistentes al estrés a tiempo 1, 2, 3, 4 y 5 horas después de añadir el antibiótico. Se representa la media y la desviación estándar correspondientes a tres réplicas biológicas independientes.

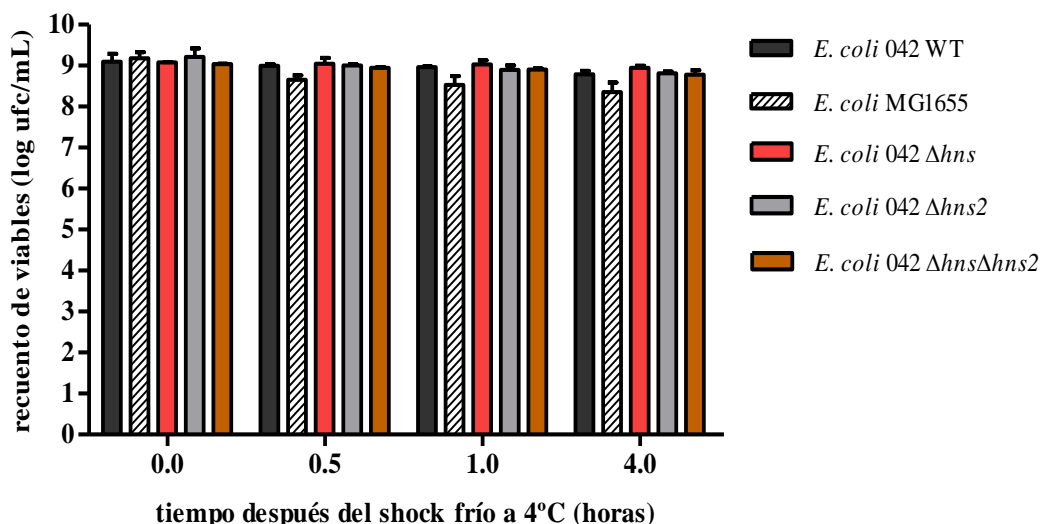
#### 4. RESULTADOS

Tal y como era de esperar, después de añadir el antibiótico al medio se produce una gran disminución en el número de células viables en el cultivo. Este hecho se hace principalmente notorio en el caso de la cepa salvaje y en el mutante sencillo  $\Delta hns2$ . Sin embargo, se puede observar cómo el mutante sencillo  $\Delta hns$  presenta una mayor persistencia ante la presencia de un antibiótico en el medio sobre el cual la bacteria es sensible. Este efecto todavía es más pronunciado en el caso del doble mutante  $\Delta hns\Delta hns2$ , en el que la supervivencia al antibiótico es mayor que en el mutante sencillo  $\Delta hns$ , principalmente durante las 3 primeras horas después de añadir el antibiótico al medio.

Analizando los resultados obtenidos puede apreciarse cómo, de nuevo, los efectos observados a nivel fenotípico en el mutante sencillo  $\Delta hns$  se ven amplificadas en el doble mutante  $\Delta hns\Delta hns2$ , mientras que la mutación sencilla en el gen  $hns2$  no genera efectos fenotípicos significativos.

##### **4.2.9.3 Efecto de los genes *hns* y *hns2* en la resistencia al choque frío**

Otro grupo de genes que presentaban una gran desregulación en el mutante sencillo  $\Delta hns$  y en el doble mutante  $\Delta hns\Delta hns2$  son genes implicados en la respuesta a *shock* frío, concretamente las chaperonas de la familia *csp* (*cold shock protein*). El gen *hns* está implicado en la respuesta adaptativa al *shock* frío a través de un bucle regulatorio ejercido con la proteína CspA (La Teana y col., 1991). Para intentar asociar un fenotipo a esta desregulación, se decidió realizar un ensayo de resistencia a choque por temperatura. Para ello, las diferentes cepas sometidas al *shock* frío (*E. coli* 042 WT, *E. coli* 042  $\Delta hns$ , *E. coli* 042  $\Delta hns2$  y *E. coli* 042  $\Delta hns\Delta hns2$ , además de la cepa de laboratorio *E. coli* MG1655), se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria ( $D.O_{600nm} = 2,0$ ). Una vez llegados a este punto, se disminuyó bruscamente la temperatura de los cultivos hasta 4°C. Los cultivos se incubaron durante 30 minutos, 1 hora y 4 horas y se cuantificó la viabilidad celular (resistencia al choque por temperatura) mediante siembra en placa y posterior recuento de viables (figura 4.40).



**Figura 4.40.** Representación del recuento de células viables después de someter a los cultivos bacterianos a un *shock* por disminución de la temperatura. Los cultivos de las cepas de *E. coli* 042 WT, *E. coli* 042  $\Delta hns$ , *E. coli* 042  $\Delta hns2$ , *E. coli* 042  $\Delta hns\Delta hns2$  y *E. coli* MG1655 se crecieron en medio LB a 37°C hasta el inicio de la fase estacionaria, momento en el que se disminuyó la temperatura de los mismos hasta 4°C. Se realizó el recuento de ufc/ml (unidades formadoras de colonias/ml) de las células resistentes al estrés a tiempos 0, 0,5, 1 y 4 horas después de bajar la temperatura del medio de cultivo hasta 4°C. Se representa la media y la desviación estándar correspondientes a tres réplicas biológicas independientes.

Se puede observar cómo, a diferencia de la cepa *E. coli* MG1655, en la que si se observa una disminución de la viabilidad celular después de someter a los cultivos a la disminución brusca de temperatura de 37°C a 4°C, en el caso de la cepa *E. coli* 042 salvaje (ni de ninguno de los mutantes analizados) no se observa afectación alguna, por lo que no es posible inferir un fenotipo de mayor resistencia al estrés por *shock* frío (en relación a los diferentes mutantes estudiados) bajo estas condiciones experimentales.

#### 4.2.10 Identificación de mutaciones en el gen *hns2* que aumenten la capacidad de compensar la pérdida de H-NS

A fin de comprender mejor la relación funcional entre las proteínas H-NS y StpA, el grupo dirigido por W. Navarre llevó a cabo en 2014 un experimento de evolución dirigida en *Salmonella* Typhimurium sobre un mutante *hns*. De manera análoga a lo que ocurre en *E. coli*, en *S. Typhimurium* un mutante *hns* presenta una tasa de crecimiento notablemente inferior a la de la cepa salvaje. Dicho experimento de evolución dirigida *in vitro* se diseñó con la finalidad de seleccionar clones derivados del mutante *hns* que después de varios días de subcultivos continuos presentaran una tasa de crecimiento superior a la tasa de crecimiento inicial. Una vez seleccionados dichos clones, se trataría de identificar el

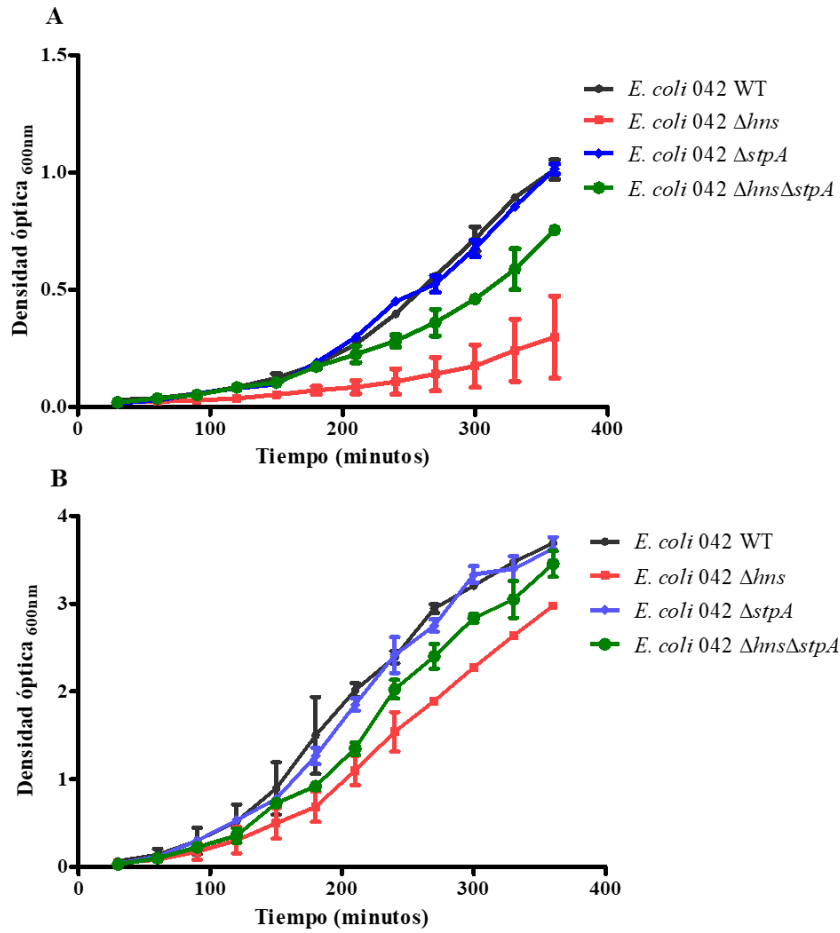
#### 4. RESULTADOS

origen de las mutaciones responsables de la reversión de la tasa de crecimiento (Ali y col., 2014).

En efecto, después de 30 días seguidos de subcultivo (aproximadamente 300 generaciones bacterianas), los autores del trabajo lograron aislar varios clones en los que se produjo una reversión de la tasa de crecimiento de la cepa mutante *hns* hasta niveles comparables a los de la cepa salvaje. Después de secuenciar el genoma entero de tales clones, se pudo observar como en el cromosoma de los mismos se habían producido deleciones de 10-50 Kb en la isla de patogenicidad 1 (SPI-1), mutaciones en el sistema de dos componentes PhoP/PhoQ (encargado de activar genes reprimidos por parte de la proteína H-NS implicados en la virulencia, la resistencia al estrés ácido, la resistencia a péptidos antimicrobianos y en la supervivencia en el interior de los macrófagos) y mutaciones diversas con cambio de sentido (mutaciones puntuales que cambian el aminoácido codificado en el codón determinado) en el gen *stpA* (Ali y col., 2014). El sentido de la selección favorable de dichas mutaciones en el gen *stpA* se interpretó como cambios generados con la finalidad de incrementar la semejanza estructural y funcional de la proteína StpA respecto a la proteína H-NS.

Siguiendo esta premisa, se decidió llevar a cabo un experimento análogo al de estas características en la cepa *E. coli* 042 con la finalidad de entender mejor la función de la proteína H-NS2. Con este experimento se intentaría analizar si la presión selectiva ejercida sobre el gen *hns2* ante la ausencia del gen *hns* (o *stpA*) era capaz de provocar cambios en la secuencia nucleotídica de dicho gen.

Como paso previo a dicho experimento, y con el fin de determinar sobre qué tipo de fondo genético mutante *hns* se realizaría el experimento de evolución dirigida, se decidió analizar la cinética de crecimiento de los mutantes  $\Delta hns$ ,  $\Delta stpA$  y del doble mutante  $\Delta hns\Delta stpA$  en la cepa *E. coli* 042. Las curvas de crecimiento se realizaron en medio líquido rico (LB) tanto a 25°C como a 37°C (figura 4.41).

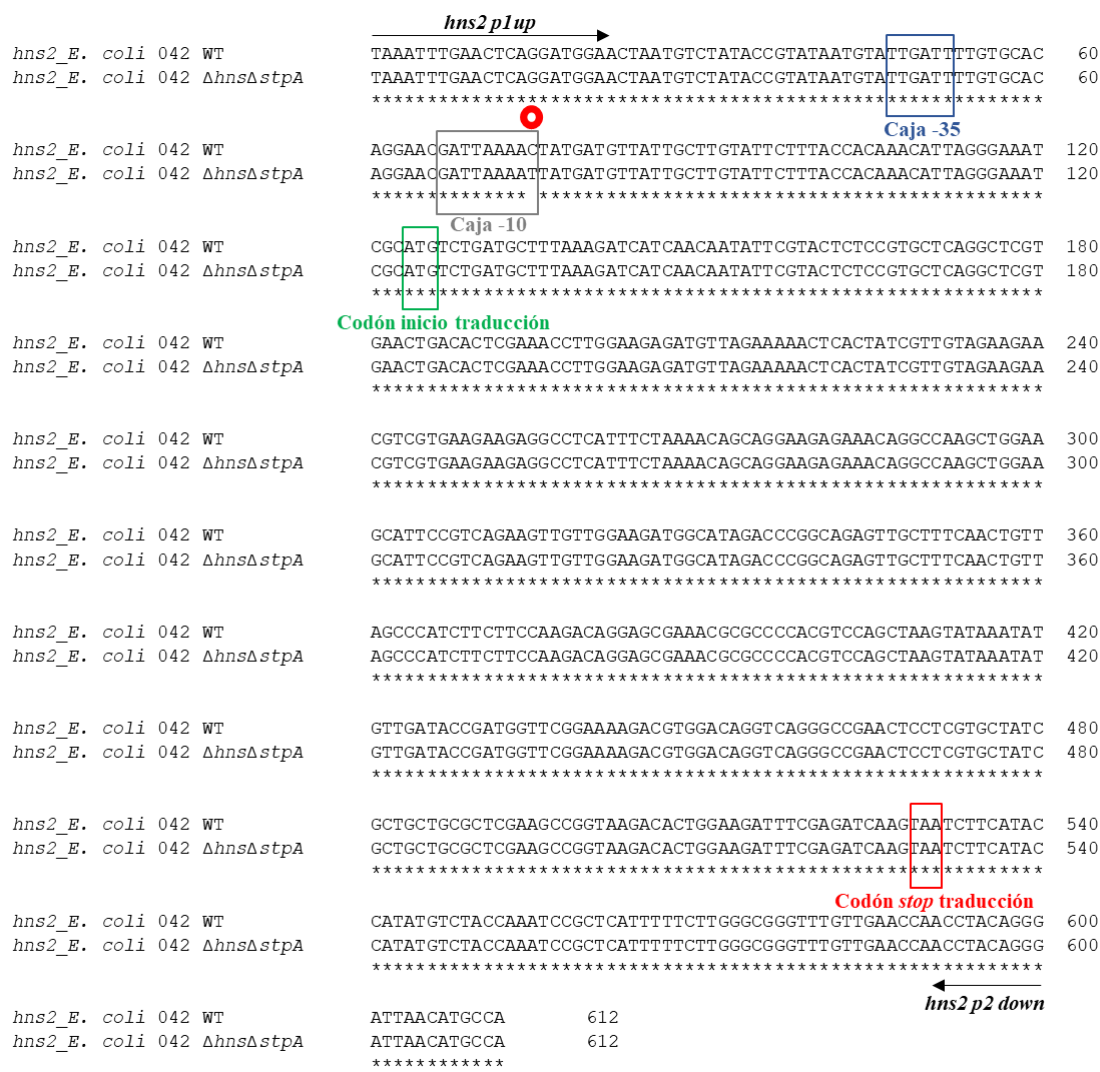


**Figura 4.41.** Curva de crecimiento de la cepa *E. coli* 042 salvaje (042 WT) y los correspondientes mutantes sencillos *hns* (042  $\Delta hns$ ), *stpA* (042  $\Delta stpA$ ) y el doble mutante *hns stpA* (042  $\Delta hns\Delta stpA$ ) en medio rico LB a 25°C (A) o a 37°C (B). Los cultivos con medio fresco fueron inoculados con una dilución 1:100 a partir de cultivos previos crecidos durante 16 horas en agitación constante. Se representan los valores correspondientes al promedio obtenido al realizar 3 réplicas biológicas independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Tanto las curvas realizadas a 25°C como a 37°C siguen un mismo patrón: la cepa que presenta la mutación sencilla del gen *hns* es la que muestra una mayor afectación sobre la tasa de crecimiento. En el lado opuesto, la mutación sencilla del gen *stpA* no provoca afectación alguna sobre la fisiología de las bacterias. Ello parece indicar que la presencia del gen *hns* es indispensable para el correcto *fitness* celular (tal y como se pudo observar en la figura 4.21). Sin embargo, tanto a 25°C como a 37°C, se ha podido observar cómo, de manera sorprendente, la cepa que contiene la doble mutación combinada de los genes *hns* y *stpA* presenta una tasa de crecimiento superior a la del mutante sencillo *hns*, y mucho mayor a la del doble mutante *hns hns2* (representada en la figura 4.21).

#### 4. RESULTADOS

Para intentar comprender este comportamiento, se amplificó por PCR el gen *hns2* tanto en la cepa *E. coli* 042  $\Delta hns\Delta stpA$  como en la cepa salvaje, y se secuenció el fragmento obtenido con el fin de encontrar alguna mutación en la secuencia codificante del gen *hns2* que explicara el porqué de la inesperada elevada tasa de crecimiento que muestra el doble mutante *hns stpA* (figura 4.42).



**Figura 4.42.** Resultado de la secuenciación del gen *hns2* en las cepas *E. coli* 042 salvaje y *E. coli* 042  $\Delta hns\Delta stpA$ . Remarcado en verde se indica el codón de inicio de la traducción del gen *hns2*, en rojo el codón *stop* de la traducción, en gris se indica la caja -10 y en azul la caja -35. La predicción de las secuencias reguladoras se realizó mediante el software BROM (*Sotfberry*), descrito por (Solovyev y Salamov, 2011). Con un círculo en rojo se indica el único cambio nucleotídico observado entre las dos secuencias. Con flechas en negro se indica el lugar de hibridación de los oligonucleótidos utilizados en la secuenciación del gen *hns2*.

Analizando las secuencias obtenidas, puede comprobarse cómo únicamente se había producido un único cambio de un nucleótido en el gen *hns2* en la cepa  $\Delta hns\Delta stpA$  respecto a la cepa salvaje. Dicho cambio, además, se encuentra localizado en una región anterior al codón de inicio de la traducción del gen *hns2*, por lo que no forma parte de la región codificante de dicho gen.

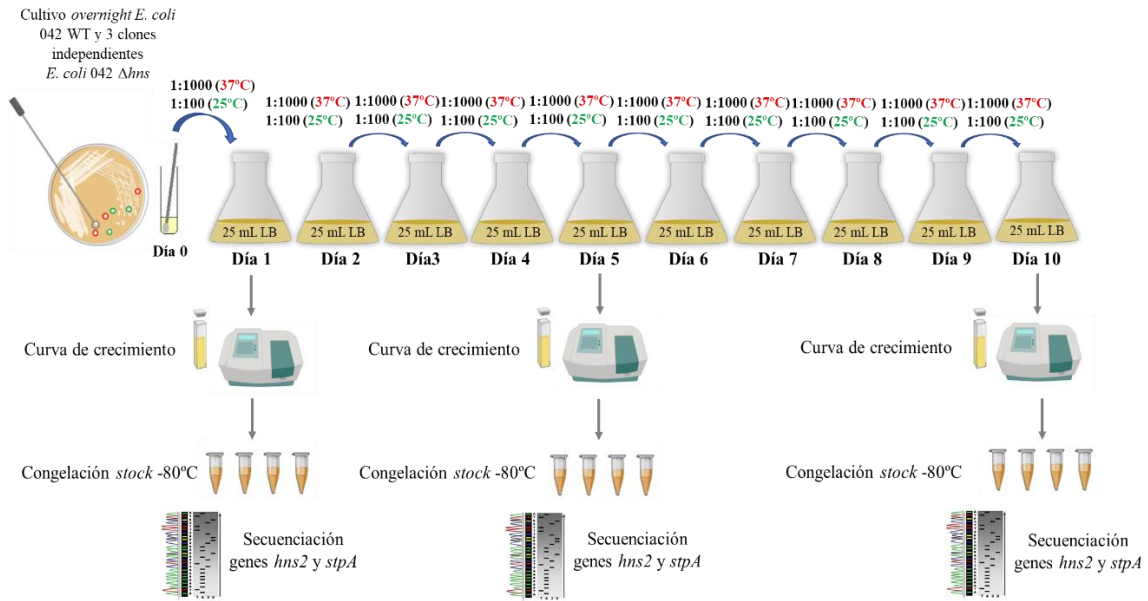
Mediante el algoritmo predictor de secuencias promotoras BPROM (Solovyev y Salamov, 2011), se logró identificar las dos secuencias reguladoras (cajas -35 y -10) en la región promotora del gen *hns2*. De manera notoria, la única mutación detectada en la cepa mutante *E. coli* 042  $\Delta hns\Delta stpA$  respecto a la cepa salvaje (C→T) se localizaba en la última posición de la caja -10 del promotor del gen *hns2*.

#### **4.2.10.1 Experimento de evolución dirigida sobre un mutante *hns* en la cepa *E. coli* 042**

Tal y como se ha comentado anteriormente, se decidió estudiar si mutaciones producidas en el gen *hns2* podrían compensar la pérdida de *fitness* celular asociada a la ausencia del producto génico del gen *hns*. En función de la tasa de crecimiento observada en las figuras 4.21 y 4.41, y debido a la inusual elevada tasa de crecimiento del doble mutante  $\Delta hns\Delta stpA$ , se decidió realizar el experimento de evolución dirigida sobre el mutante sencillo del gen *hns* de la cepa *E. coli* 042. El experimento se realizó durante 10 días de subcultivos continuos en medio LB tanto a 25°C como a 37°C, y se inició a partir de un inóculo crecido durante 16 horas (*overnight*) tanto de la cepa *E. coli* 042 salvaje como de 3 clones independientes mutantes para el gen *hns*. A partir de estos cultivos (correspondientes al día 0), se inocularon con una dilución 1:1000 los cultivos en LB a 37°C y una dilución 1:100 los cultivos en LB a 25°C, iniciándose el día 1 de los subcultivos. De estos cultivos, se realizó la curva de crecimiento y se cuantificó la cinética de crecimiento de los mismos. Al día siguiente, a partir del cultivo correspondiente al día 1 crecido durante toda la noche, se inoculó medio de cultivo fresco LB tanto a 25°C como a 37°C (con diluciones 1:100 y 1:1000 del cultivo del día anterior, respectivamente), iniciándose el día 2 de subcultivos. Este mismo procedimiento se repitió hasta llegar al día 10 de los subcultivos. De manera análoga al día 1, también se monitorizó la cinética de crecimiento de los cultivos de los días 5 y 10. De estos mismos días, se congelaron alícuotas pertenecientes a cada cultivo a -80°C, para posteriormente amplificar por PCR y secuenciar los genes *hns2* y *stpA* de cada cepa. En la figura 4.43 se esquematiza el protocolo utilizado durante el desarrollo de este experimento.

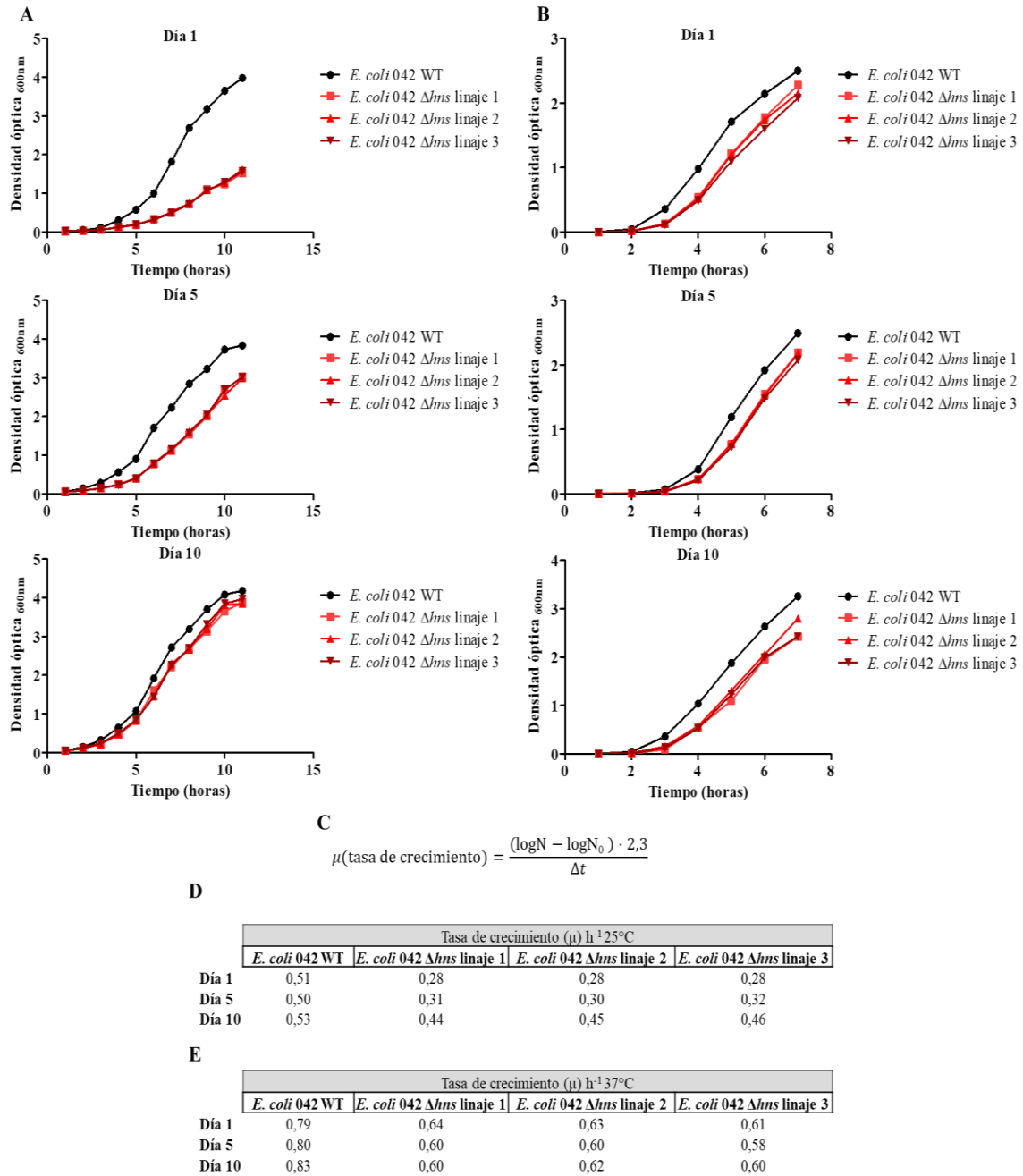


#### 4. RESULTADOS



**Figura 4.43.** Representación esquemática del protocolo de evolución dirigida realizado para seleccionar clones derivados del mutante *hns* de la cepa *E. coli* 042 que presentaran una superior tasa de crecimiento. Para realizar los subcultivos durante 10 días seguidos en medio LB, se seleccionó 1 clon de la cepa *E. coli* 042 WT y 3 clones independientes de la cepa *E. coli* 042  $\Delta hns$ . Los subcultivos en medio fresco se realizaron a partir de un inóculo del cultivo del día anterior (dilución 1:1000 en el caso del cultivo a 37°C y dilución 1:100 en el caso del cultivo a 25°C) durante 10 días seguidos, determinándose la curva de crecimiento de los mismos a los días 1, 5 y 10 de subcultivo. En esos mismos días se congeló una alícuota bacteriana correspondiente a cada cultivo para secuenciar posteriormente los genes *hns2* y *stpA* en dichas muestras.

En la figura 4.44 se representan los resultados de las curvas de crecimiento del experimento de evolución dirigida realizado sobre la cepa *E. coli* 042  $\Delta hns$  después de 10 días de subcultivos continuos, tanto a 25°C como a 37°C.



**Figura 4.44.** Curvas de crecimiento de las cepas *E. coli* 042 WT (negro) y 3 clones independientes de *E. coli* 042  $\Delta hns$  (rojo) en medio LB a 25°C (A) y 37°C (B) durante los días 1, 5 y 10 de subcultivos del experimento de evolución dirigida. Se indica también la fórmula matemática utilizada para calcular la tasa de crecimiento de cada clon específico (C) y los resultados correspondientes a cada cultivo tanto a 25°C (D) como a 37°C (E).

En el caso del experimento realizado a 37°C, se puede observar cómo los 3 clones independientes mutantes *hns* muestran una tasa de crecimiento inferior a la de la cepa salvaje que no presenta variabilidad entre los días 1, 5 y 10 de subcultivos (figura 4.44 (B) y (E)). Sin embargo, cuando el experimento de evolución dirigida tiene lugar a 25°C, en el 5° día de subcultivos ya se puede observar una significativa reversión de la tasa de

#### 4. RESULTADOS

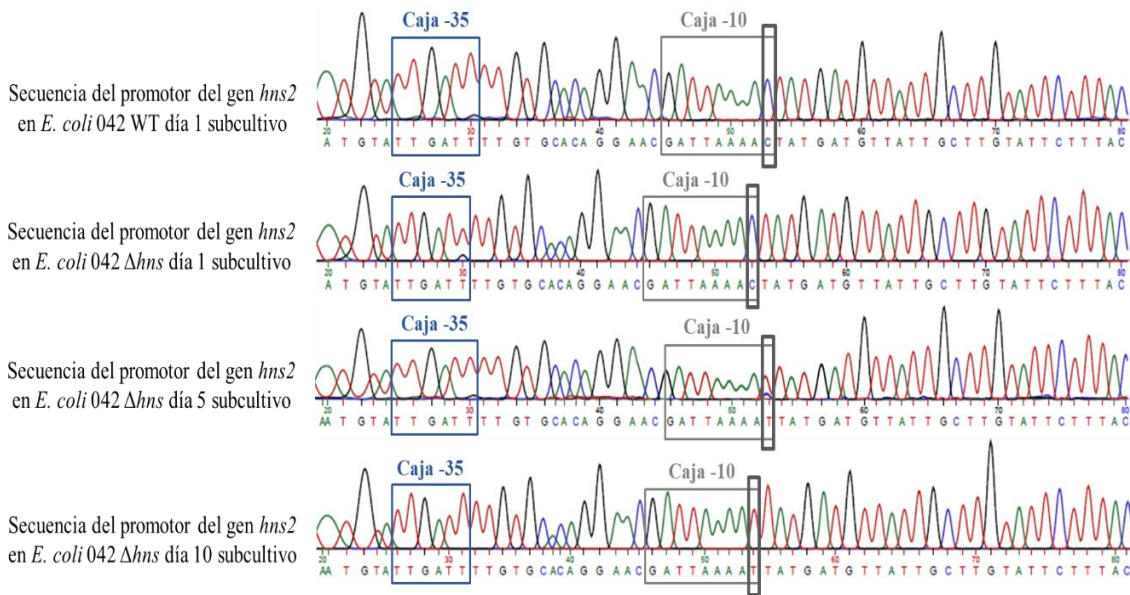
crecimiento de los 3 clones mutantes para el gen *hns* respecto a la tasa de crecimiento observada en el día 1 de subcultivos. A día 10, esta reversión se hace mucho más notable, alcanzando unos niveles de velocidad de crecimiento muy similares a los de la cepa salvaje (figura 4.44 (A) y (D)).

Con la finalidad de identificar posibles mutaciones en los genes *stpA* y *hns2* aparecidas a lo largo de las diferentes generaciones bacterianas que explicaran el porqué de la reversión del fenotipo *hns*<sup>-</sup> hasta alcanzar un nivel de tasa de crecimiento equivalente al de la cepa salvaje por parte de los 3 clones independientes  $\Delta hns$  a 25°C, se decidió secuenciar estos dos genes en los cultivos correspondientes a los días 1, 5 y 10 del experimento de evolución dirigida tanto en el caso de la cepa salvaje como en el de los 3 clones mutantes mencionados anteriormente. A pesar de no observarse diferencias en la tasa de crecimiento de los mutantes *hns* con el paso de los días de subcultivo a 37°C, también se decidió secuenciar los genes *hns2* y *stpA* de los cultivos crecidos bajo estas condiciones experimentales.

Tal y como era de esperar, en los cultivos crecidos a 37°C no se observó ningún cambio nucleotídico en la secuencia de los genes *hns2* y *stpA* en los 3 clones independientes *E. coli* 042  $\Delta hns$  respecto a la secuencia de estos mismos genes en la cepa *E. coli* 042 WT (resultados no mostrados en esta memoria).

En referencia a los cultivos crecidos a 25°C, la secuencia del gen *stpA* tampoco presentaba variación alguna en los clones  $\Delta hns$  revertientes respecto a la secuencia de dicho gen en la cepa salvaje. Sin embargo, en relación con el gen *hns2*, mientras que en la secuencia codificante del gen no se encontró cambio alguno respecto a la cepa salvaje (la secuencia del gen es la misma los días 1, 5 y 10 de subcultivos), en la región promotora de dicho gen se encontró una única (e idéntica) mutación en los 3 clones mutantes *hns* cuya tasa de crecimiento había revertido hasta niveles comparables a los de la cepa salvaje (clones pertenecientes a los días 5 y 10 de subcultivos).

En la figura 4.45 se representan los cromatogramas pertenecientes a la secuenciación de la región promotora del gen *hns2* de la cepa *E. coli* 042 WT (a día 1 de subcultivo) y la cepa *E. coli* 042  $\Delta hns$  (linaje 1) a días 1, 5 y 10 de subcultivos del experimento de evolución dirigida realizado a 25°C.



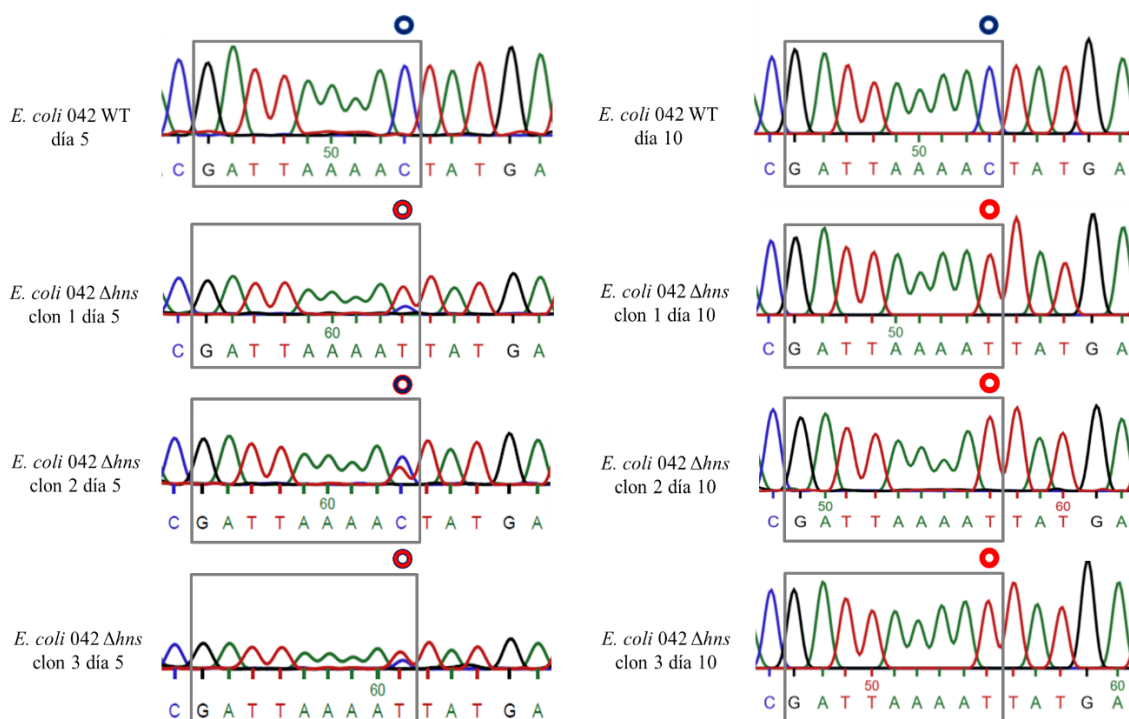
**Figura 4.45.** Cromatograma representativo de la secuenciación del promotor del gen *hns2* en la cepa *E. coli* 042 WT (día 1 de subcultivo) y la cepa *E. coli* 042  $\Delta hns$  linaje 1 (días 1, 5 y 10 de subcultivos) en el caso del experimento de evolución dirigida realizado en medio LB a 25°C. Remarcado en gris se indica la secuencia -10 y en azul la secuencia -35. La predicción de las secuencias reguladoras se realizó mediante el software BPROM (*Sotfberry*), descrito por (Solovyev y Salamov, 2011). Con un rectángulo en negro se indica el único cambio nucleotídico identificado entre las secuencias analizadas.

Curiosamente, la mutación generada en la cepa *E. coli* 042  $\Delta hns$  (a lo largo de los diferentes días de subcultivos a 25°C) en la región promotora del gen *hns2* es la misma mutación que la encontrada en el mutante *E. coli* 042  $\Delta hns\Delta stpA$ , mutante que fenotípicamente presenta una tasa de crecimiento superior a la de un mutante sencillo  $\Delta hns$  (resultados representados en la figura 4.41). Esta mutación, en la que se produce una transición de una citosina (C) por una timina (T), se localiza en la última posición de la caja -10 de la región promotora del gen *hns2*.

En el caso del experimento de evolución dirigida realizado en *Salmonella*, la emergencia temporal de las mutaciones observadas tuvo lugar durante los primeros días de subcultivos. En el día 2 de subcultivo aparecieron las deleciones parciales sobre la isla de patogenicidad SPI-1, mientras que las mutaciones puntuales con cambio de sentido en el gen *stpA* empezaron a aparecer en el día 5 de subcultivos (Ali y col., 2014). En el caso del experimento de evolución dirigida realizado sobre la cepa *E. coli* 042  $\Delta hns$ , durante el 5º día de subcultivos a 25°C ya se pudo observar un aumento significativo en la tasa de crecimiento respecto al día 1, por lo que es plausible que las mutaciones aparecieran en este intervalo tiempo. En la figura 4.45 se representó el cromatograma correspondiente

#### 4. RESULTADOS

únicamente a la secuenciación del gen *hns2* del clon 1 de la cepa *E. coli* 042  $\Delta hns$ , en la que ya aparece la mutación puntual en el promotor del gen. Sin embargo, al observar los cromatogramas correspondientes a los otros dos clones (clones 2 y 3) de *E. coli* 042  $\Delta hns$  a día 5 de subcultivos, se puede apreciar que existe una divergencia en cuanto a la aparición de la transición C $\rightarrow$ T en el promotor del gen *hns2* en función del clon analizado. En la figura 4.46 puede observarse cómo, en el caso del clon 1, existe una población inequívoca que presenta el cambio nucleotídico en el promotor del gen *hns2*, mientras que en el caso de los clones 2 y 3 puede observarse un solapamiento de los picos de fluorescencia correspondientes a cada nucleótido específico, hecho que nos indica que en el cultivo existe una mezcla de poblaciones que contienen el alelo salvaje (C) o el alelo mutado (T). A día 10 de subcultivos, en cambio, las poblaciones de los 3 clones independientes  $\Delta hns$  que presentaban el cambio en el promotor del gen *hns2* han desplazado totalmente a aquellas que contenían el alelo salvaje, tal y como puede apreciarse en el cromatograma, donde existe únicamente un pico de fluorescencia correspondiente al nucleótido (timina) mutado (figura 4.46).



**Figura 4.46.** Cromatograma representativo de la secuencia -10 del promotor del gen *hns2* en la cepa *E. coli* 042 WT y los 3 linajes independientes de la cepa *E. coli* 042  $\Delta hns$  en los días 5 y 10 de subcultivos. Remarcado en gris se indica la secuencia -10, con un círculo azul el nucleótido C correspondiente a la secuencia salvaje del promotor del gen *hns2*, y con un círculo en rojo el nucleótido T correspondiente a la mutación generada durante los continuos subcultivos del mutante *hns* en el experimento de evolución dirigida.

Globalmente, se puede concluir que, en el caso de la cepa *E. coli* 042, a diferencia del caso de *Salmonella* Typhimurium, la presión selectiva generada por la ausencia del gen *hns* selecciona de manera positiva mutaciones en el gen *hns2* que provocan que un mutante *hns* incremente su tasa de crecimiento hasta asemejarla a la de la cepa salvaje. Por tanto, mientras que en el modelo de *Salmonella* pueden seleccionarse positivamente mutaciones con cambio de sentido que dan lugar a variaciones en la secuencia proteica de StpA, en el caso de la cepa *E. coli* 042 no se seleccionan cambios en la secuencia aminoacídica de H-NS2, sino un único tipo de mutación generada en la región promotora del gen, concretamente en la última posición nucleotídica de la secuencia -10.

Sin embargo, respecto a los resultados anteriores, no se puede inferir que la mutación generada en el promotor del gen *hns2* pueda ser el único cambio responsable de la reversión de la tasa de crecimiento del mutante  $\Delta hns$ . En el caso de *Salmonella*, además de las mutaciones en el gen *stpA*, también se produjeron deleciones importantes en varias regiones del genoma (principalmente en genes HGT implicados en la virulencia bacteriana). Por todo ello, se decidió secuenciar masivamente el genoma de uno de los clones mutante  $\Delta hns$  revertiente a los 10 días de subcultivo para compararlo con el genoma de la cepa *E. coli* 042 WT (crecida bajo las mismas condiciones de cultivo). Todo ello tenía como objetivo identificar otros posibles cambios producidos en el genoma bacteriano después de los 10 días de cultivos seriados que, junto con el cambio nucleotídico generado en la caja -10 del promotor del gen *hns2*, pudieran dar respuesta al porqué de la elevada tasa de reversión de las cepas que presentan la mutación en el gen *hns*. La secuenciación masiva de las cepas se realizó en colaboración con el grupo de investigación dirigido por el Profesor Trinad Chakraborty. En los genomas secuenciados no se observaron otros cambios a los que se pudiera atribuir algún efecto en la tasa de crecimiento, más allá de la inserción de fagos crípticos en varios lugares del cromosoma.

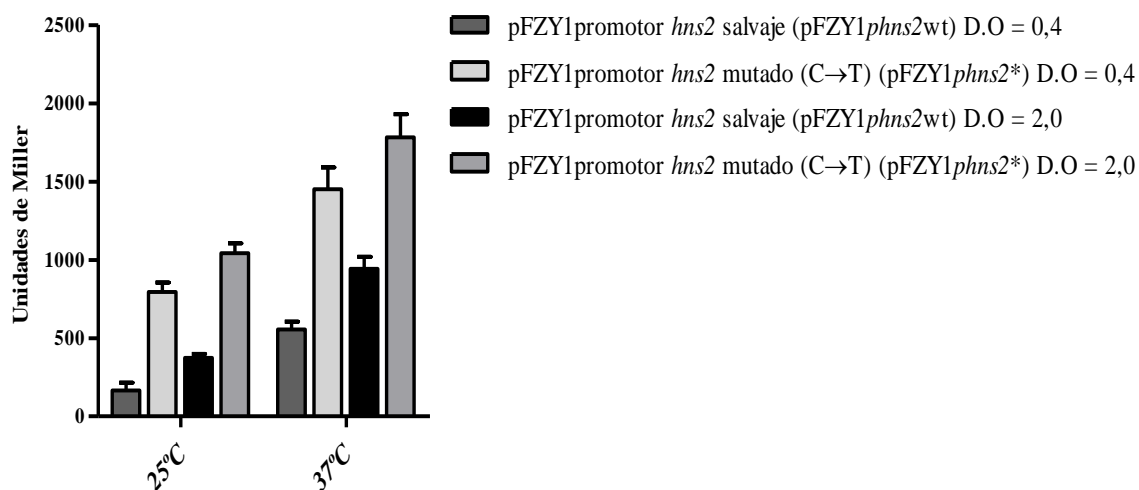
#### **4.2.11 Efecto de la transición C→T en el promotor del gen *hns2* sobre la expresión de su producto génico**

Para comprobar si la mutación puntual generada en el promotor del gen *hns2* a lo largo de las diferentes generaciones bacterianas era capaz de provocar cambios en la expresión transcripcional del gen *hns2*, se realizó una fusión transcripcional con el gen reportero *lacZ* (presente en el plásmido pFZY1) tanto del promotor salvaje del gen *hns2* (*phns2wt*) como del promotor que presenta la transición C→T en la caja -10 (*phns2\**). Una vez generados los vectores correspondientes, se transformaron en la cepa *E. coli* 042  $\Delta lacZ$  y



#### 4. RESULTADOS

se cuantificó la actividad transcripcional de los diferentes promotores mediante ensayos de  $\beta$ -galactosidasa sobre cultivos crecidos en medio LB en la fase de crecimiento exponencial y hasta el inicio de la fase estacionaria de crecimiento, tanto a 25°C como a 37°C (figura 4.47).



**Figura 4.47.** Expresión transcripcional del gen reportero *lacZ* clonado en el plásmido pFZY1 bajo el control del promotor salvaje del gen *hns2* o del promotor con la mutación puntual C→T generada durante el experimento de evolución dirigida. Los niveles de transcripción se cuantificaron mediante ensayos de  $\beta$ -galactosidasa a partir de cultivos de la cepa *E. coli* 042  $\Delta lacZ$  (transformada con los diferentes plásmidos) crecidos en medio LB a 25°C o 37°C en fase exponencial (D.O<sub>600nm</sub>=0,4) o hasta el inicio de la fase estacionaria de crecimiento (D.O<sub>600nm</sub>=2,0). Se representa el promedio y la desviación estándar correspondientes a tres réplicas biológicas independientes.

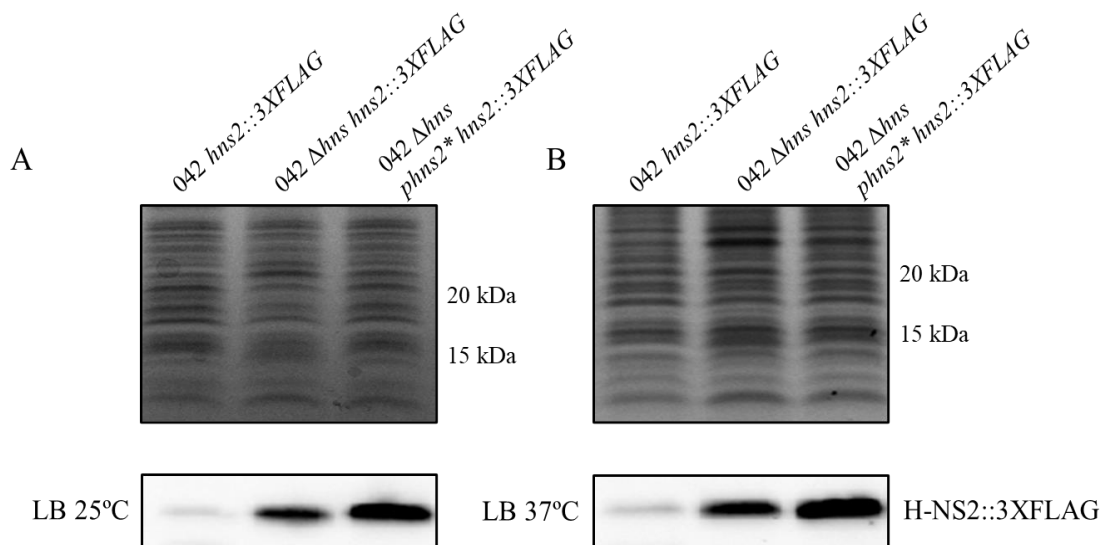
Tal y como se puede observar en la figura 4.47, el nivel de expresión del promotor salvaje del gen *hns2* es superior a 37°C que a 25°C. Estos resultados también se pueden observar en la figura 4.24, donde se analizó mediante una fusión transcripcional a nivel cromosómico los niveles de expresión del gen *hns2*, y en la figura 4.25, donde por *Western Blot* se pudo comprobar que los niveles de expresión de la proteína H-NS2 son superiores cuando las bacterias crecen a 37°C respecto a cuando lo hacen a 25°C. También se puede observar cómo los niveles de expresión del promotor salvaje del gen *hns2* son superiores en el inicio de la fase estacionaria de crecimiento respecto a cuando las bacterias se encuentran en fase de crecimiento exponencial.

En referencia a la expresión comparativa de la actividad transcripcional del promotor salvaje del gen *hns2* respecto el promotor que presenta la mutación puntual (C→T), los niveles de transcripción del promotor mutado son del orden de aproximadamente 2 veces



superiores en comparación con los niveles de expresión del promotor salvaje, independientemente de la temperatura de crecimiento de las bacterias y de la fase de crecimiento en la que se encuentren.

Para acabar de confirmar dichos datos experimentalmente, se compararon los niveles de expresión de la proteína H-NS2 bajo el control del promotor salvaje respecto los niveles de expresión de dicha proteína bajo el control del promotor con la mutación en la caja -10. Para ello, se marcó con el epítipo 3XFLAG la proteína H-NS2 en la cepa *E. coli* 042  $\Delta hns$  (cepa con el promotor salvaje del gen *hns2*) y en la cepa *E. coli* 042  $\Delta hns$  *phns2*\* (cepa que presenta la transición C $\rightarrow$ T generada en la caja -10 del promotor del gen *hns2* después de los 10 días de subcultivos del experimento de evolución dirigida). Dichas cepas se crecieron en medio LB a 25°C y 37°C hasta el inicio de la fase estacionaria de crecimiento, se obtuvieron extractos proteicos totales y se analizó por *Western Blot* los niveles de expresión de la proteína H-NS2::3XFLAG (figura 4.48).



**Figura 4.48.** Análisis por *Western Blot* de los niveles de expresión de la proteína H-NS2::3XFLAG en un fondo genético salvaje (042 *hns2*::3XFLAG), un fondo genético *hns*<sup>-</sup> (042  $\Delta hns$  *hns2*::3XFLAG) y un fondo genético *hns*<sup>-</sup> con la transición C $\rightarrow$ T generada en el promotor del gen *hns2* (042  $\Delta hns$  *phns2*\* *hns2*::3XFLAG). Los cultivos correspondientes se crecieron en medio LB tanto a 25°C (A) como a 37°C (B) hasta el inicio de la fase estacionaria de crecimiento (D.O<sub>600nm</sub> 2,0). El panel de arriba muestra la tinción de Coomassie de cada uno de los extractos proteicos analizados y representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas.

De manera similar a lo mostrado en la figura 4.30, la mutación en el gen *hns* provoca la sobreexpresión de la proteína H-NS2 respecto a un fondo genético salvaje. Además, cuando en un fondo genético *hns*<sup>-</sup> se compara la expresión de la proteína H-NS2 bajo el

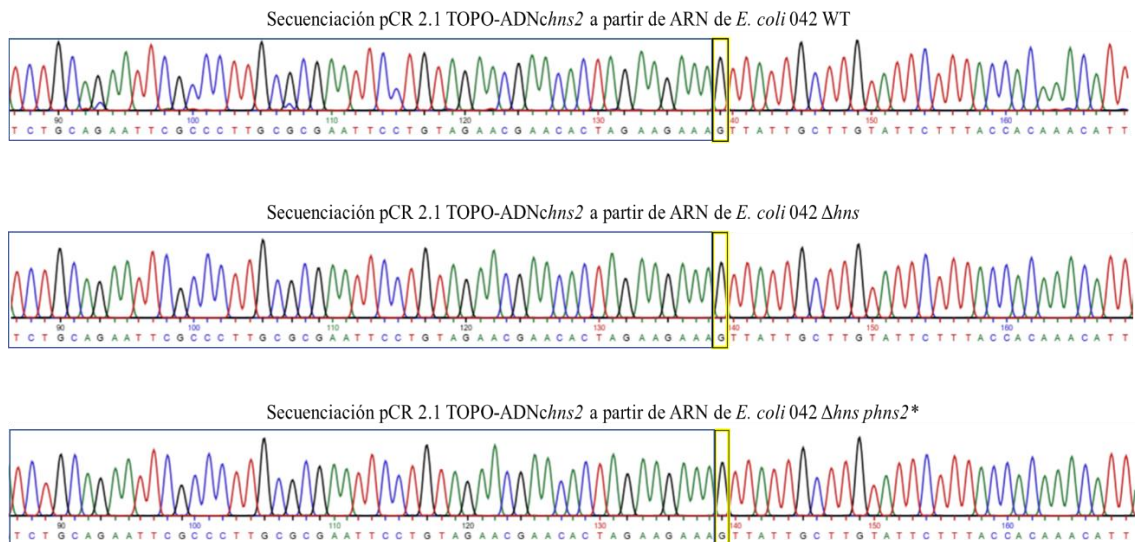
#### 4. RESULTADOS

control del promotor salvaje o bajo el control del promotor que presenta la transición C→T, se puede observar cómo en este último caso se produce un incremento aún mayor en los niveles de expresión de la proteína H-NS2.

Analizando globalmente los resultados obtenidos, todo parece indicar que la mutación en el promotor del gen *hns2* seleccionada de manera positiva por la presión selectiva sobre un fondo genético *hns<sup>s</sup>*, parece ayudar a compensar la pérdida de este gen mediante un incremento en los niveles de transcripción y posterior traducción del gen *hns2* y de su correspondiente producto génico.

#### **4.2.12 Efecto de la transición C→T sobre el promotor del gen *hns2*: determinación del inicio de transcripción del gen *hns2***

Seguidamente, se quiso analizar si la transición C→T generada en el promotor del gen *hns2* era responsable de provocar algún cambio en el inicio de transcripción del gen respecto al promotor salvaje. Para ello, se realizó un experimento de 5' RACE (*rapid amplification of cDNA ends*), donde se amplificó y se secuenció el ADNc correspondiente al gen *hns2* hasta su extremo 5' con el fin de identificar y mapear el primer nucleótido (+1) del ARN mensajero de dicho gen. Dicha amplificación y posterior secuenciación del ADNc del gen *hns2* se realizó en las cepas *E. coli* 042 WT, *E. coli* 042  $\Delta hns$  y *E. coli* 042  $\Delta hns phns2^*$ , a partir de cultivos crecidos en medio LB a 25°C hasta el inicio de la fase estacionaria de crecimiento. Una vez aislado el ARN total de los correspondientes cultivos bacterianos, la base del protocolo experimental utilizado consistió en realizar la ligación de una secuencia adaptadora de ARN (oligonucleótido A3, cuya secuencia es 5'-GAU AUG CGC GAA UUC CUG UAG AAC GAA CAC UAG AAG AAA-3') en el extremo 5' de los ácidos nucleicos. Seguidamente, se retrotranscribió todo el ARN a ADNc, y se amplificó por PCR el ADNc correspondiente al gen *hns2* haciendo uso de un oligonucleótido interno al adaptador unido y un oligonucleótido que hibridaba en el extremo 3' del gen (*hns2-3'RACE Rev*). Dicho ADNc amplificado por PCR se ligó al plásmido pCR 2.1 TOPO-vector. El plásmido generado se secuenció posteriormente, donde el siguiente nucleótido localizado después de la secuencia adaptadora unida al extremo 5' del ARN correspondía al primer nucleótido (+1) del transcrito del gen *hns2*. Concretamente, se secuenciaron 10 clones independientes para cada uno de los tres aislamientos de ARN realizados (cepas WT,  $\Delta hns$  y  $\Delta hns phns2^*$ ). En la figura 4.49 se representan los cromatogramas representativos de la secuenciación del ADNc del gen *hns2* clonado en el plásmido pCR 2.1 TOPO.



**Figura 4.49.** Cromatogramas representativos de la secuenciación del ADNc correspondiente al gen *hns2* amplificado previamente por PCR y clonado en el vector pCR 2.1 TOPO. Remarcado en color azul se muestra la secuencia del adaptador unido al extremo 5' del ARN, posteriormente retrotranscrito a ADNc y amplificado por PCR. Remarcado en amarillo se indica el primer nucleótido identificado después del sitio de unión del adaptador, correspondiente al primer nucleótido (+1) del transcrito del gen *hns2*.

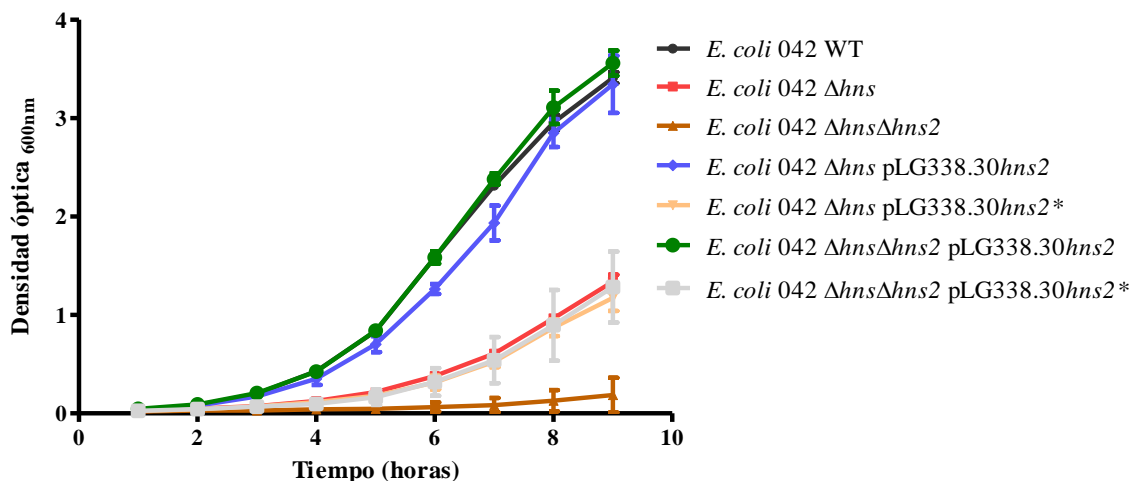
Tal y como se puede observar en la anterior figura, el inicio de transcripción del gen *hns2* es el mismo independientemente de si el promotor del gen contiene la secuencia salvaje o la secuencia con la transición C $\rightarrow$ T, por lo que la transición generada durante el experimento de evolución dirigida no parece provocar cambios en el inicio de transcripción del gen *hns2*. Dicho inicio mapea 7 nucleótidos después de la secuencia de la caja -10, y 41 nucleótidos antes del codón de inicio de traducción del gen *hns2*.

#### 4.2.13 Efecto de la dosis génica del gen *hns2* al complementar la mutación $\Delta hns$

Paralelamente, también se quiso analizar si la capacidad de complementación de una mutación en el gen *hns* en la cepa *E. coli* 042 por parte del gen *hns2* variaba en función de si el gen se encontraba bajo su promotor salvaje o bajo el promotor con la mutación C $\rightarrow$ T. Para ello, además del vector pLG338.30*hns2* (construcción que contiene la secuencia del gen *hns2* clonado bajo el control de su promotor salvaje) utilizado para realizar la complementación de la mutación en el gen *hns* en la cepa *E. coli* MG1655  $\Delta hns$  (figura 4.22), también se construyó el vector pLG338.30*hns2*\*, construcción que contiene la secuencia del gen *hns2* clonado bajo el control del promotor que presenta la transición C $\rightarrow$ T en la caja -10. Dichas construcciones se transformaron a un mutante sencillo  $\Delta hns$

#### 4. RESULTADOS

y a un doble mutante  $\Delta hns\Delta hns2$  de la cepa *E. coli* 042, y se realizaron sendas curvas de crecimiento en medio LB a 25°C para estudiar el nivel de complementación ejercido por estos plásmidos (figura 4.50).



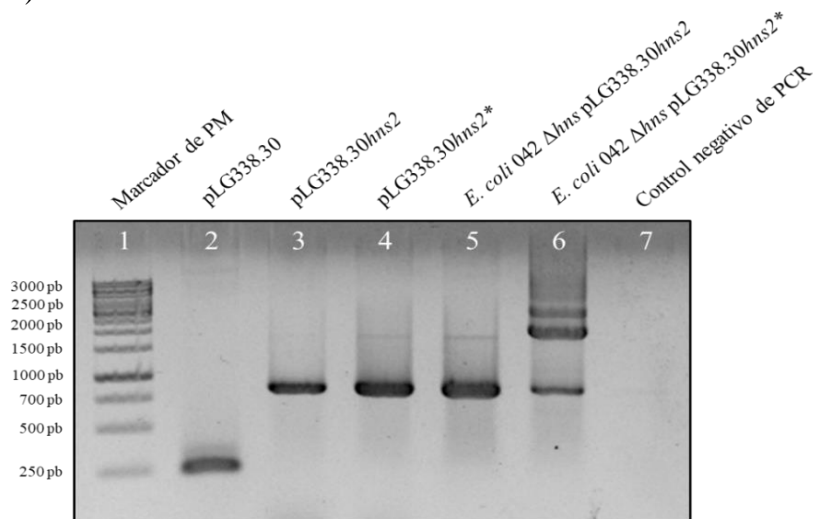
**Figura 4.50.** Curva de crecimiento en medio LB a 25°C de las cepas *E. coli* 042 WT, *E. coli* 042  $\Delta hns$ , *E. coli* 042  $\Delta hns\Delta hns2$ , *E. coli* 042  $\Delta hns$  pLG338.30hns2, *E. coli* 042  $\Delta hns$  pLG338.30hns2\*, *E. coli* 042  $\Delta hns\Delta hns2$  pLG338.30hns2 y *E. coli* 042  $\Delta hns\Delta hns2$  pLG338.30hns2\*. Los cultivos con medio fresco fueron inoculados con una dilución 1:100 a partir de cultivos previos crecidos durante 16 horas en agitación constante. Se representan los valores correspondientes al promedio obtenido al realizar 3 réplicas biológicas independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Tal y como se puede observar, existe una gran diferencia al complementar la mutación sencilla  $\Delta hns$  (o la doble mutación  $\Delta hns\Delta hns2$ ) con el plásmido pLG338.30 con el gen *hns2* clonado con su promotor salvaje o con el promotor que presenta la transición C→T. En el primer caso, la complementación de las mutaciones es prácticamente total, ya que el gran efecto sobre el crecimiento provocado por la falta del gen *hns* se ve totalmente revertido con la complementación en *trans* por parte del gen *hns2*. Sin embargo, cuando la complementación tiene lugar con el gen *hns2* bajo el control del promotor mutado, no existe reversión alguna del crecimiento. Este hecho ocurre principalmente en el caso de la cepa *E. coli* 042  $\Delta hns$  (cepa que también contiene la copia cromosómica del gen *hns2*), por lo que puede inferirse que los elevados niveles de proteína H-NS2 proporcionados por el plásmido cuando el promotor contiene la transición C→T, junto con los niveles de proteína H-NS2 cromosómica pueden llegar a generar cierta toxicidad sobre las bacterias.

Para confirmar que no se hubiera producido otra nueva mutación compensatoria con el fin de aliviar el efecto de un eventual exceso de proteína H-NS2 cuando se encuentra

codificada bajo el control de un promotor con más potencia que el promotor salvaje, se decidió secuenciar tanto el gen *hns2* cromosómico como el gen *hns2* clonado bajo el control del promotor con la transición C→T en el plásmido pLG338.30 en la cepa *E. coli* 042  $\Delta hns$  pLG338.30*hns2*\*, y se comparó con la secuencia nucleotídica del gen *hns2* de la cepa *E. coli* 042  $\Delta hns$  pLG338.30*hns2*.

Al realizar la amplificación por PCR sobre el gen *hns2* cromosómico y su posterior secuenciación pudo comprobarse que, en el caso de las dos cepas analizadas, el gen *hns2* se mantenía con la misma secuencia nucleotídica que la cepa salvaje (resultados no mostrados). Sin embargo, al realizar el análisis por PCR sobre el inserto clonado en el vector pLG338.30*hns2*\* y transformado en la cepa *E. coli* 042  $\Delta hns$ , puede comprobarse como, a diferencia de la construcción que contiene el gen *hns2* clonado bajo su promotor salvaje (*E. coli* 042  $\Delta hns$  pLG338.30*hns2*), la cepa presentaba dos tamaños de amplicón diferentes: además del amplicón original esperado con un tamaño correspondiente al del promotor del gen *hns2* y al propio gen en sí, también se observó un segundo tamaño de amplicón mucho mayor al esperado, el cual no se encontraba presente en los plásmidos originales (pLG338.30*hns2* y pLG338.30*hns2*\*) antes de ser transformados a las bacterias (figura 4.51).



**Figura 4.51.** Resultado de la amplificación por PCR del gen *hns2* clonado en el vector pLG338.30. Se representan los resultados correspondientes a la amplificación del vector pLG338.30 sin inserto clonado (carril 2, utilizado como control), del vector pLG338.30 con el gen *hns2* clonado bajo el control de su promotor salvaje (pLG338.30*hns2*, carril 3) o bajo el control del promotor con la transición C→T (pLG338.30*hns2*\*, carril 4), y de los vectores generados en los carriles 3 y 4 después de ser transformados en la cepa *E. coli* 042  $\Delta hns$  y realizar la curva de crecimiento representada en la figura 4.50 (carriles 5 y 6, respectivamente). En el carril 1 se indica el marcador de peso molecular cargado en el gel y en el carril 7 se muestra el control negativo de PCR.

## 4. RESULTADOS

Al secuenciar dicha construcción, se observó que en la secuencia codificante del gen *hns2* clonado bajo el promotor con la transición C→T se había insertado un elemento de inserción (concretamente de la familia *IS5*), rompiendo totalmente la pauta de lectura codificante para el gen *hns2*, inactivando la totalidad de su producto génico (figura 9.15 representada en el apartado Anexo). El resultado de la secuenciación de los plásmidos utilizados en la curva de crecimiento representada en la figura 4.50 explica por qué al complementar la mutación  $\Delta hns$  con un plásmido que aportaba en *trans* el gen *hns2* bajo el control de un promotor más potente que el salvaje no se conseguían los niveles de complementación deseados. En este sentido, ya que unos elevados niveles de la proteína H-NS2 pueden resultar eventualmente tóxicos para las bacterias, ante este estrés se seleccionan de manera positiva clones en los que el gen clonado se haya inactivado mediante la inserción de un elemento de inserción en la secuencia codificante del mismo.

### **4.2.14 Efecto de la temperatura sobre el papel regulador de las proteínas H-NS y H-NS2**

Como parte final de este apartado, teniendo en cuenta que es a 25°C la temperatura en la que se ha podido poner de manifiesto que un aumento en los niveles de expresión de H-NS2 permiten complementar el fenotipo provocado por la carencia de H-NS, decidimos analizar el efecto de las mutaciones en los genes *hns* y *hns2* sobre el transcriptoma global de la cepa *E. coli* 042 a 25°C. Dichos ensayos transcriptómicos se realizaron a partir de cultivos crecidos en medio LB a 25°C hasta el inicio de la fase estacionaria de crecimiento, donde el aislamiento del ARN y el procesamiento de las muestras para obtener los datos relativos a la expresión de cada gen se realizaron en colaboración con la empresa *Vertis Biotechnologie AG (Freising-Weihenstephan, Alemania)* siguiendo el protocolo descrito en el apartado 3.8.5 de Materiales y Métodos.

Tal y como se puede comprobar, a 25°C, las proteínas H-NS y H-NS2 juegan un papel muy diferente del ya estudiado a 37°C (apartado 4.2.8). Mientras que a 37°C existe una buena correlación entre algunos de los genes que se encuentran sobreexpresados en el fondo genético *hns*<sup>-</sup> y los sobreexpresados en el fondo genético *hns2*<sup>-</sup> (tablas 4.5 (A y B), éste no es el caso a 25°C (tabla 4.8). A esta temperatura, muchos de los genes sobreexpresados en un fondo genético *hns*<sup>-</sup> corresponden a los ya identificados a 37°C, pero con unos niveles de expresión relativa muy superiores (tabla 4.6 (A y B)). Ello claramente indica que, a 25°C, la proteína H-NS juega un papel esencial silenciando la expresión indeseada de un gran número de genes en la cepa *E. coli* 042.



A

Identificador NCBI ( <i>locus tag</i> )	Fold change	
	$\Delta hns$ 37°C	$\Delta hns$ 25°C
EC042_RS08585 (glutamate decarboxylase beta subunit GadB)	43,0	371,1
EC042_RS20290 (glutamate decarboxylase alpha subunit GadA)	37,6	247,8
EC042_RS08580 (glutamate/gamma-aminobutyrate antiporter GadC)	34,7	273,2
EC042_RS01725 (putative fimbrial protein MatB)	33,3	128,5
EC042_RS20250 (putative periplasmic acid stress chaperone HdeA)	26,1	-
EC042_RS01730 (putative fimbrial transcriptional regulator MatA)	22,7	30,6
EC042_RS05780 (tRNA-Ser)	19,3	-
EC042_RS15335 (DNA-binding protein StpA)	18,6	72,1
EC042_RS20255 (putative acid resistance protein HdeD)	17,8	219,3
EC042_RS20260 (transcriptional regulator GadE)	16,6	154,7
EC042_RS06765 (tRNA-Tyr)	16,3	-
EC042_RS13160 (hypothetical protein)	15,7	-
EC042_RS21295 (CesD/SycD/LcrH family type III secretion system chaperone)	15,6	22,9
EC042_RS01715 (putative fimbrial outer membrane usher protein)	15,1	70,0
EC042_RS20245 (acid-activated periplasmic chaperone HdeB)	15,1	187,9
EC042_RS06770 (tRNA-Tyr)	14,8	-
EC042_RS04580 (hypothetical protein)	14,5	13,5
EC042_RS01720 (putative fimbrial protein)	14,2	157,9
EC042_RS20090 (hypothetical protein membrane protein)	14,0	438,8
EC042_RS07750 (type I toxin-antitoxin Iar)	13,9	-
EC042_RS01270 (hypothetical protein)	13,6	88,2
EC042_RS24770 (tRNA-Gly)	13,5	-
EC042_RS09275 (hypothetical protein)	13,3	-
EC042_RS01065 (tRNA-Ala)	12,8	-
EC042_RS18940 (tRNA-Ala)	12,8	-
EC042_RS26795 (hypothetical protein)	12,2	-
EC042_RS16270 (hypothetical protein)	12,2	30,3
EC042_RS11015 (tRNA-Leu)	12,1	-
EC042_RS11620 (tRNA-Arg)	12,0	-

B

Identificador NCBI ( <i>locus tag</i> )	Fold change	
	$\Delta hns$ 25°C	$\Delta hns$ 37°C
EC042_RS07290 (general stress protein)	1137,5	9,3
EC042_RS07280 (ferritin-like domain-containing protein)	736,5	8,8
EC042_RS00760 (fimbrial protein)	718,8	4,3
EC042_RS12270 (colanic acid biosynthesis acetyltransferase WcaF)	581,7	9,8



#### 4. RESULTADOS

EC042_RS12260 (GDP-L-fucose synthase WcaG)	505,7	5,4
EC042_RS24025 (cadaverine/lysine antiporter CadB)	492,4	7,0
EC042_RS20090 (hypothetical protein membrane protein)	438,8	14,0
EC042_RS12275 (colanic acid biosynthesis glycosyltransferase WcaE)	427,6	7,4
EC042_RS08585 (glutamate decarboxylase)	371,1	4,3
EC042_RS17480 (thymidylate kinase)	364,7	4,7
EC042_RS12255 (GDP-mannose mannosyl hydrolase)	352,9	6,1
EC042_RS12280 (putative colanic acid polymerase WcaD)	332,5	6,6
EC042_RS03510 (DUF1266 domain-containing protein)	308,5	4,0
EC042_RS01730 (fimbrial transcriptional regulator MatA)	306,4	22,7
EC042_RS12250 (colanic acid biosynthesis glycosyltransferase WcaI)	294,4	6,8
EC042_RS20240 (MgtC/SapB family protein)	292,1	7,3
EC042_RS01795 (reactive chlorine species resistance protein RclC)	289,9	7,0
EC042_RS12265 (GDP-mannose 4,6-dehydratase)	288,4	10,5
EC042_RS08580 (glutamate/gamma-aminobutyrate antiporter GadC)	273,3	34,7
EC042_RS03225 (LPS O-antigen length regulator)	263,4	6,1
EC042_RS01800 (DUF1471 domain-containing protein)	260,2	6,6
EC042_RS10355 (GGDEF domain-containing protein)	249,9	7,5
EC042_RS20290 (glutamate decarboxylase alpha subunit GadA)	247,8	37,6
EC042_RS07285 (ferritin-like domain-containing protein)	245,6	8,0
EC042_RS22730 (hypothetical protein)	243,5	3,4
EC042_RS13790 (fimbrial protein)	235,5	4,2
EC042_RS08620 (acid stress response transcriptional regulator YdeO)	226,4	7,1
EC042_RS12365 (DUF2314 domain-containing protein)	223,0	6,4
EC042_RS17785 (fimbrial-like protein)	222,5	3,4
EC042_RS20250 (putative acid stress chaperone HdeA)	222,0	26,1
EC042_RS08380 (type VI secretion system PAAR protein)	221,3	5,6
EC042_RS20255 (putative acid resistance protein HdeD)	219,3	17,8
EC042_RS08045 (porin OmpN)	211,6	4,8
EC042_RS22645 (porin OmpL)	205,7	7,1

**Tabla 4.6.** (A) Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT en medio LB a 37°C. Al lado se indican los valores de expresión relativa de dichos genes en el mismo mutante crecido en medio LB a 25°C. (B) Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT en medio LB a 25°C. Al lado se indican los valores de expresión relativa de dichos genes en el mismo mutante crecido en medio LB a 37°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

En relación a la proteína H-NS2, los genes regulados a 37°C por parte de este parólogo (principalmente genes de origen HGT correspondientes a proteínas hipotéticas) no se encuentran regulados a 25°C, por lo que esta proteína también parece jugar un papel modulador de la expresión génica dependiente de la temperatura (tabla 4.7 (A)).

**A**

Identificador NCBI ( <i>locus tag</i> )	Fold change	
	$\Delta hns2$ 37°C	$\Delta hns2$ 25°C
EC042_RS05870 (hypothetical protein)	12,1	-
EC042_RS09275 (hypothetical protein)	11,8	-
EC042_RS07145 (hypothetical protein)	11,3	-
EC042_RS26995 (hypothetical protein)	11,3	-
EC042_RS18795 (hypothetical protein)	11,2	-
EC042_RS11295 (hypothetical protein)	10,3	-
EC042_RS13835 (tRNA-Arg)	10,1	-
EC042_RS24055 (antitoxin hypothetical protein)	9,6	-
EC042_RS22745 (DUF3521 domain-containing protein)	9,6	-
EC042_RS03185 (hypothetical protein)	9,3	-
EC042_RS17205 (hypothetical protein)	9,2	-
EC042_RS17210 (hypothetical protein)	9,1	-
EC042_RS03005 (envY porin thermoregulatory protein)	9,0	-
EC042_RS24455 (hypothetical protein)	8,8	-
EC042_RS17185 (putative ATP/GTP-binding protein)	8,7	-
EC042_RS17230 (hypothetical protein)	8,6	-
EC042_RS21255 (putative prophage protein)	8,6	-
EC042_RS25650 (putative RadC-like DNA repair protein)	8,5	-
EC042_RS18375 (hypothetical protein)	8,5	-
EC042_RS17220 (putative DNA repair protein)	8,5	-
EC042_RS11730 (hypothetical protein)	8,4	-
EC042_RS24065 (putative plasmid-like protein)	8,4	-
EC042_RS17215 (putative antirestriction protein)	8,4	-
EC042_RS07390 (hypothetical protein)	8,3	-
EC042_RS12065 (hypothetical protein)	8,3	-
EC042_RS13250 (hypothetical protein)	8,2	-
EC042_RS20520 (putative acetyltransferase)	8,1	-
EC042_RS23195 (hypothetical protein)	8,0	-
EC042_RS17225 (hypothetical protein)	8,0	-

**B**

Identificador NCBI ( <i>locus tag</i> )	Fold change	
	$\Delta hns2$ 25°C	$\Delta hns2$ 37°C
EC042_RS06740 (nitrate reductase subunit alpha NarZ)	44,9	2,9
EC042_RS06735 (nitrate transporter NarK)	23,6	2,6
EC042_RS06745 (nitrate reductase subunit beta NarH)	21,0	3,2

#### 4. RESULTADOS

EC042_RS06750 (nitrate reductase molybdenum cofactor assembly chaperone NarJ)	13,2	2,4
EC042_RS06755 (respiratory nitrate reductase subunit gamma NarI)	12,8	3,1
EC042_RS03510 (DUF1266 domain-containing protein)	8,2	5,0
EC042_RS10415 (DUF1869 domain-containing protein)	6,7	4,9
EC042_RS16380 (EscR/YscR/HrcR family type III secretion system export apparatus protein)	6,3	2,8
EC042_RS22730 (hypothetical protein)	5,4	5,0
EC042_RS19330 (nitrite reductase large subunit)	4,9	2,9
EC042_RS19990 (nickel ABC transporter)	4,7	3,7
EC042_RS08490 (formate dehydrogenase subunit beta FdxH)	4,6	2,8
EC042_RS03055 (hypothetical protein)	4,5	2,5
EC042_RS08485 (formate dehydrogenase-N subunit alpha FdnG)	4,5	2,1
EC042_RS02000 (galactoside O-acetyltransferase LacA)	4,5	3,2
EC042_RS07770 (hypothetical protein)	4,5	5,9
EC042_RS08480 (sulfate ABC transporter substrate-binding protein)	4,2	-
EC042_RS19335 (nitrite reductase small subunit NirD)	4,0	3,3
EC042_RS28855 (hypothetical protein)	4,0	-
EC042_RS19660 (fimbria/pilus periplasmic chaperone)	3,6	3,9
EC042_RS12365 (yegJ family protein)	3,6	4,1
EC042_RS08380 (type VI secretion system PAAR protein)	3,6	3,0
EC042_RS13110 (ferredoxin-type protein NapF)	3,5	-
EC042_RS08495 (formate dehydrogenase cytochrome b556 subunit)	3,4	2,6
EC042_RS23725 (cytochrome c-552 NrfA)	3,3	-
EC042_RS01960 (2-methylcitrate dehydratase)	3,1	-
EC042_RS28795 (hypothetical protein)	3,1	-

**Tabla 4.7.** (A) Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT en medio LB a 37°C. Al lado se indican los valores de expresión relativa de dichos genes en el mismo mutante crecido en medio LB a 25°C. (B) Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT en medio LB a 25°C. Al lado se indican los valores de expresión relativa de dichos genes en el mismo mutante crecido en medio LB a 37°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

A 25°C, el papel regulador que ejerce la proteína H-NS2 parece ser menor que el observado a 37°C. Bajo estas condiciones, el número de genes cuya expresión se ve desregulada en ausencia del gen *hns2* a 25°C es menor que el número de genes desregulados a 37°C (figura 9.18, Anexo de esta memoria). A 25°C, en el mutante  $\Delta hns2$  principalmente se encuentran altamente desregulados genes implicados en el sistema de la nitrato reductasa (sistema que cataliza la reacción de reducción de nitrito a amonio como fuente asimilable de nitrógeno), además de varias proteínas hipotéticas y enzimas implicadas en diferentes vías metabólicas (tabla 4.7 (B)).

En la tabla 4.8 se representa la relación de genes mayormente desregulados en el mutante  $\Delta hns$  a 25°C y el valor de expresión de dichos genes en el mutante  $\Delta hns2$  crecido bajo las mismas condiciones de cultivo. En dicha tabla se puede observar el notorio efecto que genera la pérdida del gen *hns* cuando las bacterias crecen a baja temperatura, mientras que *hns2* no parece jugar un rol significativo en el silenciamiento de estos genes.

Identificador NCBI ( <i>locus tag</i> )	Fold change	
	$\Delta hns$ 25°C	$\Delta hns2$ 25°C
EC042_RS07290 (general stress protein)	1137,5	-
EC042_RS07280 (ferritin-like domain-containing protein)	736,5	-
EC042_RS00760 (fimbrial protein)	718,8	-
EC042_RS12270 (colanic acid biosynthesis acetyltransferase WcaF)	581,7	-
EC042_RS12260 (GDP-L-fucose synthase WcaG)	505,7	-
EC042_RS24025 (cadaverine/lysine antiporter CadB)	492,4	-
EC042_RS20090 (hypothetical protein membrane protein)	438,8	-
EC042_RS12275 (colanic acid biosynthesis glycosyltransferase WcaE)	427,6	-
EC042_RS08585 (glutamate decarboxylase)	371,1	-
EC042_RS17480 (thymidylate kinase)	364,7	2,0
EC042_RS12255 (GDP-mannose mannosyl hydrolase)	352,9	2,1
EC042_RS12280 (putative colanic acid polymerase WcaD)	332,5	-
EC042_RS03510 (DUF1266 domain-containing protein)	308,5	8,2
EC042_RS01730 (putative fimbrial transcriptional regulator MatA)	306,4	-
EC042_RS12250 (colanic acid biosynthesis glycosyltransferase WcaI)	294,4	-
EC042_RS20240 (MgtC/SapB family protein)	292,1	-
EC042_RS01795 (reactive chlorine species resistance protein RclC)	290,0	-
EC042_RS12265 (GDP-mannose 4,6-dehydratase)	288,4	-
EC042_RS08580 (glutamate/gamma-aminobutyrate antiporter GadC)	273,3	-
EC042_RS03225 (LPS O-antigen length regulator)	263,4	2,2

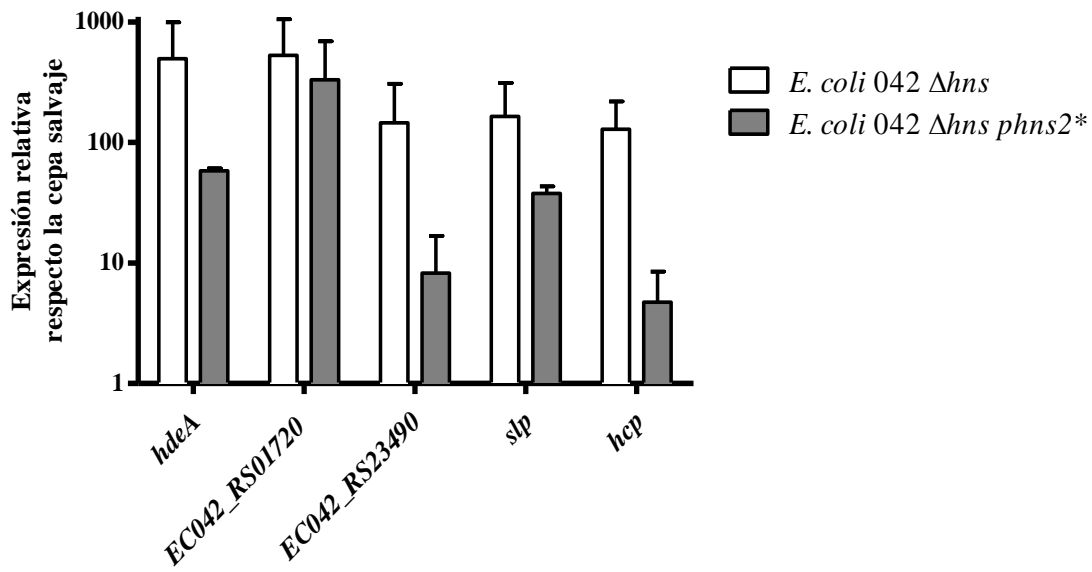
#### 4. RESULTADOS

EC042_RS01800 (DUF1471 domain-containing protein)	260,2	-
EC042_RS10355 (GGDEF domain-containing protein)	250,0	-
EC042_RS20290 (glutamate decarboxylase alpha subunit GadA)	247,8	-
EC042_RS07285 (ferritin-like domain-containing protein)	245,6	-
EC042_RS22730 (hypothetical protein)	243,5	5,4
EC042_RS13790 (fimbrial protein)	235,5	-
EC042_RS08620 (acid stress response transcriptional regulator YdeO)	226,4	-
EC042_RS12365 (DUF2314 domain-containing protein)	223,0	3,6
EC042_RS17785 (fimbrial-like protein)	222,5	-
EC042_RS20250 (putative periplasmic acid stress chaperone HdeA)	222,1	-
EC042_RS08380 (type VI secretion system PAAR protein)	221,3	3,6
EC042_RS20255 (putative acid resistance protein HdeD)	219,4	-
EC042_RS08045 (porin OmpN)	211,6	-
EC042_RS22645 (porin OmpL)	205,7	-

**Tabla 4.8.** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042  $\Delta hns$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT en medio LB a 25°C. Al lado se indican los valores de expresión relativa de dichos genes en el mutante *E. coli* 042  $\Delta hns2$  crecido bajo las mismas condiciones de cultivo.

La falta de efecto de la proteína H-NS2 sobre el silenciamiento de los genes regulados por H-NS a baja temperatura podría ser debido a los bajos niveles de expresión de la proteína H-NS2 a 25°C. Si ello fuese así, debería poder correlacionarse la desregulación de muchos genes en un fondo genético *hns*<sup>-</sup> con la baja tasa de crecimiento, y la atenuación de este fenotipo en un mutante  $\Delta hns$  que presentara la mutación puntual en el promotor del gen *hns2* (*phns2*<sup>\*</sup>) como consecuencia del incremento de la expresión de la proteína H-NS2. Para comprobar dicho supuesto, se analizó por qRT-PCR la expresión de 5 genes en las cepas *E. coli* 042  $\Delta hns$  y *E. coli* 042  $\Delta hns$  *phns2*<sup>\*</sup> (figura 4.52). Concretamente se analizó la expresión de los genes *hdeA* (gen que codifica para la chaperona implicada en la resistencia a estrés ácido HdeA), EC042\_RS01720 (gen que codifica para una proteína tipo fimbria), EC042\_RS23490 (gen que codifica para una proteína hipotética), *slp* (gen que codifica para una lipoproteína de membrana externa) y *hcp* (gen que codifica para una proteína efectora del sistema de secreción tipo VI). Mediante los experimentos de RNA-seq se ha podido comprobar que dichos genes se encuentran altamente sobreexpresados en un mutante  $\Delta hns$  respecto a la cepa salvaje cuando las bacterias crecen a 25°C en medio LB, pero su expresión no se ve alterada en

un mutante  $\Delta hns2$  crecido bajo las mismas condiciones de cultivo (tablas 9.16 y 9.17 Anexo).



**Figura 4.52.** Expresión relativa de los genes *hdeA*, EC042\_RS01720, EC042\_RS23490, *slp* y *hcp* en las cepas *E. coli* 042  $\Delta hns$  y *E. coli* 042  $\Delta hns$  *phns2*\* respecto a la cepa *E. coli* 042 salvaje. La expresión de los genes se analizó por qRT-PCR partir de cultivos crecidos en medio LB a 25°C hasta el inicio de la fase estacionaria de crecimiento, utilizando los oligonucleótidos *hdeA* Fw/Rv, RS01720 Fw/Rv, RS23490 Fw/Rv, *slp* Fw/Rv y *hcp* Fw/Rv. Los resultados fueron normalizados mediante la detección del gen interno *gapA* mediante los oligonucleótidos *gapA*042 RT Fw/Rv. Se representan los valores correspondientes a 2 réplicas biológicas independientes (realizadas con tres réplicas técnicas cada una). Las barras de error representan la desviación estándar de los valores obtenidos en cada una de las dos réplicas realizadas.

Los resultados obtenidos muestran que los elevados niveles de expresión de los 5 genes analizados como consecuencia de la mutación en el gen *hns* se compensan parcialmente cuando dicha mutación se combina con la transición C→T en la caja -10 del promotor del gen *hns2*, es decir, cuando los niveles de expresión de la proteína H-NS2 son superiores a los de un fondo genético salvaje. Ello nos indica que la proteína H-NS2 presenta la capacidad de regular algunos de los genes que regula la proteína H-NS (ayudando a atenuar de esta manera el fenotipo derivado ante la pérdida de la proteína H-NS) cuando dicha proteína se encuentra bajo el control de un promotor que muestra una potencia superior para dirigir la expresión génica respecto a la que muestra el promotor salvaje.

# Capítulo 3



## **4.3 Nuevos datos del papel del plásmido pAA2 y del regulador AggR en la virulencia de la cepa *E. coli* 042**



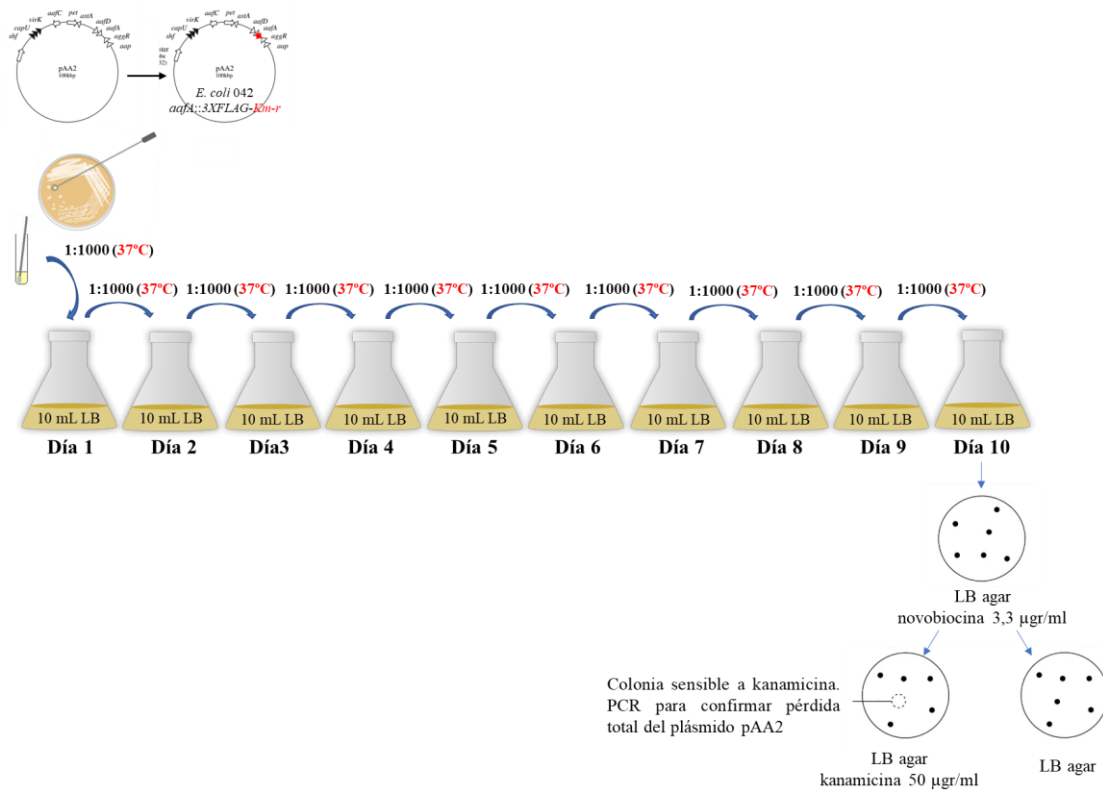
### 4.3.1 Curado del plásmido pAA2. Efecto sobre el transcriptoma global de la cepa *E. coli* 042

Adicionalmente al estudio del papel de los parálogos de las proteínas Hha y H-NS en la biología de la cepa *E. coli* 042, en la presente tesis doctoral decidimos estudiar también el papel del plásmido de virulencia pAA2 bajo una nueva perspectiva. En vez de analizar de manera específica el papel de determinantes de virulencia plasmídicos, intentamos curarlo y ver cómo afectaba la ausencia del plásmido, si era el caso, en la expresión de genes cromosómicos de la cepa *E. coli* 042.

Sin embargo, teniendo en cuenta que el plásmido pAA2 no codifica para ningún determinante de resistencia a antibióticos que nos permitiera trazar qué bacterias contienen el plásmido y cuáles no, como paso previo a la utilización de un protocolo de curado del plásmido pAA2 de la cepa *E. coli* 042, se introdujo en el plásmido un casete de resistencia a un antibiótico que nos permitiera seleccionar aquellas células que hubieran perdido el plásmido respecto a las que no. Para ello, se introdujo el epítipo 3XFLAG en el extremo C-terminal del gen *aafA*, siguiendo el protocolo descrito en el apartado 3.7.8 de Materiales y Métodos. Se decidió utilizar esta metodología ya que la inserción de la etiqueta 3XFLAG conlleva asociada un casete de resistencia a kanamicina y además no implica delecionar ningún gen plasmídico para introducir el casete de resistencia en el mismo (figura 3.2, apartado Materiales y Métodos).

Una vez obtenida la cepa *E. coli* 042 portadora del módulo de resistencia insertado en el plásmido (*E. coli* 042 *aafA::3XFLAG Km-r*), se realizó un cultivo del mismo durante 16 horas a 37°C, a partir del cual se realizaron subcultivos seriados continuos en 10 ml de medio LB sin antibióticos durante 10 días seguidos, utilizando como inóculo una dilución 1:1000 del cultivo del día anterior. Una vez llegado al décimo día de subcultivo, se plaquearon diluciones seriadas del mismo en placas de LB-agar suplementadas con el antibiótico novobiocina a una concentración final del 3,3 µgr/ml. Se decidió utilizar este antibiótico en las placas donde se realizaron las siembras después de los subcultivos ya que la novobiocina es un potente inhibidor de la ADN girasa bacteriana que se ha utilizado tradicionalmente para provocar el curado de plásmidos en diferentes bacterias (McHugh y Swartz, 1977). Las colonias aisladas que crecieron se repicaron en paralelo en placas de LB-agar sin antibiótico y placas de LB-agar suplementadas con kanamicina 50 µgr/ml (figura 4.53).

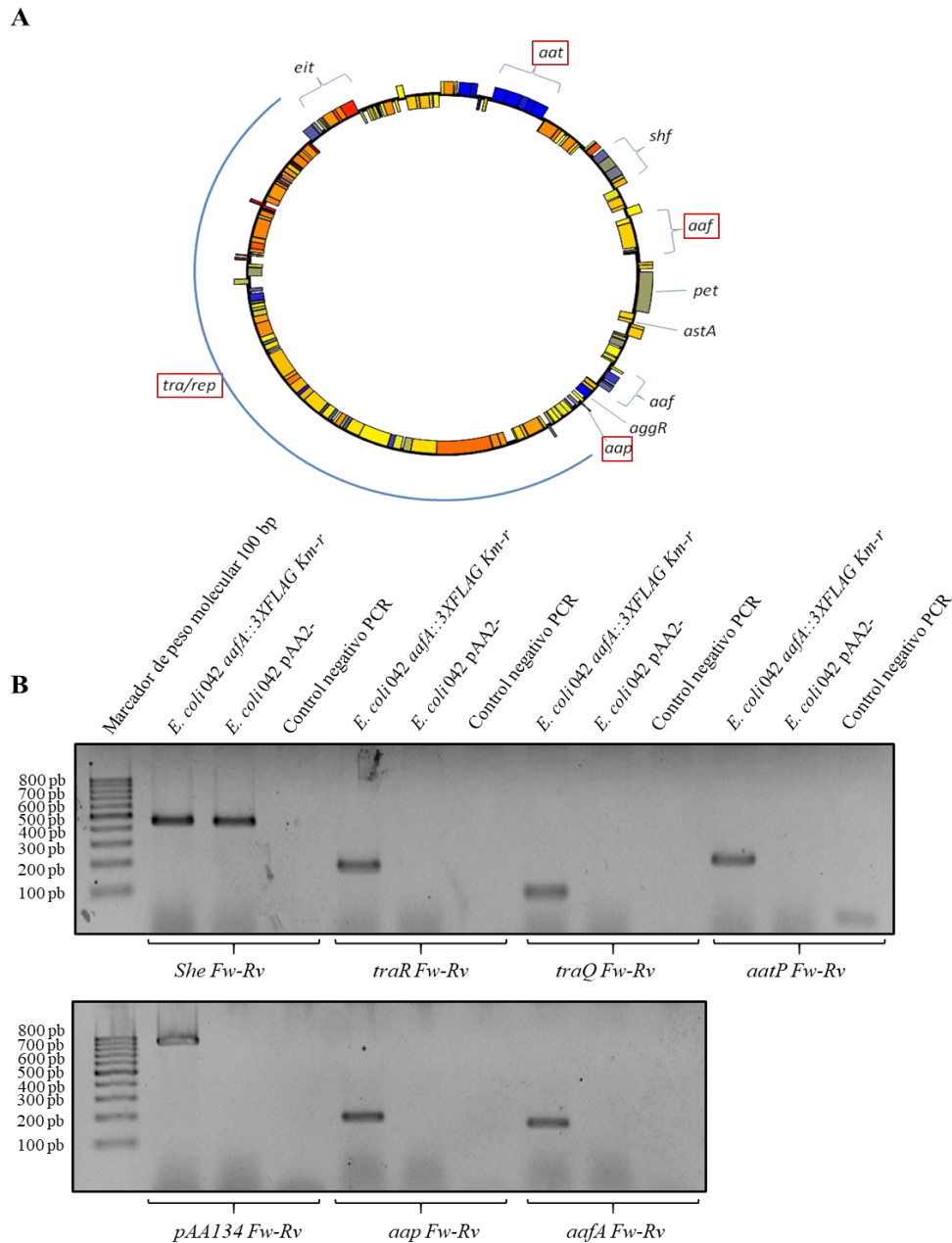
#### 4. RESULTADOS



**Figura 4.53.** Representación esquemática del protocolo utilizado para realizar y detectar el curado del plásmido pAA2 de la cepa *E. coli* 042. Después de introducir un casete de resistencia a kanamicina en el plásmido (concretamente en el extremo 3' después de la secuencia codificante del gen *aafA*) se realizaron subcultivos seriados durante 10 días del clon *aafA::3XFLAG Km-r*, tomando como inóculo una dilución 1:1000 del cultivo del día anterior. Al décimo día de subcultivos, se plaquearon las células en medio LB-agar suplementado con novobiocina a una concentración final de 3,3 µg/ml. Posteriormente, se repicaron las colonias crecidas en dicho medio por duplicado en medio LB-agar y medio LB-agar suplementado con kanamicina 50 µg/ml. Las colonias sensibles a kanamicina fueron utilizadas para analizar mediante PCR la presencia de varios genes plasmídicos (ver texto).

A las colonias sensibles a kanamicina, pero viables en LB, se les realizó un análisis por PCR para confirmar la pérdida del plásmido completo. En dicho análisis por PCR, se comprobó la ausencia total de diferentes genes plasmídicos localizados en diferentes posiciones del mapa genético del plásmido pAA2 con la finalidad de confirmar la pérdida total del mismo. Para ello, se analizó la ausencia de los genes *traQ* y *traR*, implicados en la transferencia conjugativa del plásmido, la dispersina *aap* y la permeasa responsable de su secreción (*aatP*), la metilasa de ADN presente en la *orf* 134 del plásmido, y la subunidad mayor de las fimbrias de tipo II presentes en el plásmido pAA2 (codificadas en el gen *aafA*). Además, también se analizó la presencia del gen cromosómico *she*, gen presente en la isla de patogenicidad homónima de *Shigella flexneri* pero cuya presencia queda restringida al patotipo enteroagregativo de *Escherichia coli* (Czeczulin y col.,

1999). Este último análisis por PCR se realizó con el fin de validar la integridad de la cepa *E. coli* 042 y confirmar la ausencia de posibles contaminaciones con otras bacterias después de los 10 días de subcultivo (figura 4.54).



**Figura 4.54.** (A) Representación esquemática del plásmido pAA2 en la que se indica la localización de los operones codificantes para los principales factores de virulencia presentes en el mismo, además de la región *tra/rep* implicada en la transferencia, replicación y mantenimiento del plásmido. Enmarcados en color rojo se indica la localización de los genes y los operones del plásmido pAA2 analizados por PCR (B) con el fin confirmar la pérdida total del mismo después de realizar subcultivos en medio LB durante 10 días del clon con el casete de resistencia a kanamicina asociado al gen *aafA* (*E. coli* 042 *aafA::3XFLAG Km-r*), dando lugar a la cepa curada del plásmido (*E. coli* 042 pAA2-).

#### 4. RESULTADOS

En la figura 4.54 (A) se puede observar la representación esquemática del plásmido pAA2 de la cepa *E. coli* 042. En una mitad del plásmido se localizan los genes implicados en la maquinaria de replicación y transferencia conjugativa del mismo (*tra/rep*), mientras que en la otra mitad se encuentran los genes de virulencia. Los genes correspondientes a las fimbrias de tipo II (AAF/II) se encuentran codificados en dos operones situados en distintas localizaciones. En el primer operón se localizan los genes *aafA* y *aafR*, además de la chaperona *aafD*; en el segundo operón se encuentra el gen codificante de la proteína AafC (Elias y col., 1999). En la figura 4.54 (B) se representan los resultados del análisis por PCR realizado al único clon que presentaba sensibilidad a kanamicina después de los 10 días de subcultivo. El clon utilizado inicialmente a día 1 (*E. coli* 042 *aafA::3XFLAG Km-r*, resistente a kanamicina) resultó positivo para todos los genes plasmídicos analizados, además de para el gen cromosómico *she*, gen marcador del patotipo EAEC. Después de los 10 días de subcultivo en un medio sin la presión selectiva que forzara a las células a mantener el plásmido, se logró obtener un único clon sensible a kanamicina (*E. coli* 042 pAA2-) cuyo análisis por PCR resultó ser negativo para la presencia de los genes plasmídicos mencionados anteriormente, pero positivo para el gen cromosómico *she*. De esta manera se pudo confirmar el correcto curado del plásmido pAA2 en la cepa *E. coli* 042.

Con el fin de obtener un perfil global del transcriptoma de la cepa *E. coli* 042 en ausencia del plásmido pAA2, se realizó un análisis transcriptómico comparativo por RNA-seq de las cepas *E. coli* 042 WT y *E. coli* 042 pAA2-. Estos ensayos transcriptómicos se realizaron a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento, donde el aislamiento del ARN y el procesamiento de las muestras para obtener los datos relativos a la expresión de cada gen se realizaron en colaboración con la empresa *Vertis Biotechnologie AG (Freising-Weihenstephan, Alemania)*, siguiendo el protocolo descrito en el apartado 3.8.5 de Materiales y Métodos.

En la tabla 4.9 se representa el listado de los genes diferencialmente sobreexpresados en un fondo genético pAA2- respecto a la cepa *E. coli* 042 salvaje. Además, también se indica el nivel de expresión relativa de dichos genes en los mutantes  $\Delta hha$ ,  $\Delta hns$  y  $\Delta hns\Delta hns2$ .

## 4. RESULTADOS

Identificador NCBI ( <i>locus tag</i> )	Fold change			
	pAA2-	$\Delta hha$	$\Delta hns$	$\Delta hns\Delta hns2$
EC042_RS21655 (hypothetical protein)	10,4	6,2	2,1	-
EC042_RS20255 (HdeD family acid-resistance protein)	7,2	3,2	17,8	36,7
EC042_RS20260 (acid resistance transcriptional activator GadE)	7,2	5,0	16,6	32,9
EC042_RS23795 (hypothetical protein)	6,9	4,3	5,3	3,1
EC042_RS20245 (acid-activated periplasmic chaperone HdeB)	6,4	4,0	15,1	37,1
EC042_RS20240 (MgtC/SapB family protein)	5,8	3,6	7,3	8,7
EC042_RS24005 (type V toxin-antitoxin system toxin GhoT)	5,4	5,6	6,6	2,5
EC042_RS20265 (hypothetical protein)	5,4	4,1	4,4	2,2
EC042_RS20250 (acid-activated periplasmic chaperone HdeA)	5,3	-	26,1	60,3
EC042_RS20290 (gadA glutamate decarboxylase alpha subunit)	5,2	5,5	37,6	59,8
EC042_RS08585 (gadB glutamate decarboxylase beta subunit)	5,1	4,3	43,0	87,8
EC042_RS06000 (flagellar basal body rod protein FlgB)	5,0	4,2	-	-
EC042_RS11200 (flagellar motor switch protein FliN)	4,7	-	-	-
EC042_RS06005 (flagellar basal body rod protein FlgC)	4,5	-	-	-
EC042_RS11180 (flagella biosynthesis chaperone FliJ)	4,3	-	-	-
EC042_RS08580 (acid resistance gamma-aminobutyrate antiporter GadC)	4,3	3,9	34,7	68,3
EC042_RS06010 (flagellar hook assembly protein FlgD)	4,1	-	-	-
EC042_RS11165 (flagellar motor switch protein FliG)	3,9	3,0	-	-
EC042_RS09850 (YdiL family protein)	3,9	5,8	5,7	-
EC042_RS11215 (flagellar biosynthesis protein FliQ)	3,7	3,2	2,5	-
EC042_RS20190 (outer membrane lipoprotein Slp)	3,7	3,9	9,3	18,6
EC042_RS11170 (flagellar assembly protein FliH)	3,6	-	-	-
EC042_RS11160 (flagellar basal body M-ring protein FliF)	3,6	2,8	-	-
EC042_RS06015 (flagellar hook protein FlgE)	3,5	3,1	-	-
EC042_RS25790 (anti-adaptor protein IraD)	3,5	4,4	2,8	2,6
EC042_RS15275 (tRNA-Ile)	3,5	2,9	7,5	4,0
EC042_RS16280 (hypothetical protein)	3,5	3,4	4,1	-
EC042_RS03065 (hypothetical protein)	3,5	2,7	6,3	2,7
EC042_RS25785 (hypothetical protein)	3,5	2,9	-	2,2
EC042_RS06025 (flagellar basal-body rod protein FlgG)	3,5	3,6	-	-
EC042_RS20275 (multidrug efflux RND transporter permease subunit)	3,4	4,3	5,0	5,3
EC042_RS06020 (flagellar basal body rod protein FlgF)	3,4	-	-	-
EC042_RS11535 (hypothetical protein)	3,3	5,1	2,7	-
EC042_RS11175 (flagellum-specific ATP synthase FliI)	3,3	3,5	-	-
EC042_RS12450 (PTS galactitol transporter subunit IIB)	3,2	3,2	5,8	3,5
EC042_RS11210 (flagellar type III secretion system pore protein fliP)	3,2	-	-	-

#### 4. RESULTADOS

EC042_RS06030 (flagellar basal body L-ring protein FlgH)	3,1	3,7	-	-
EC042_RS06035 (flagellar basal body P-ring protein FlgI)	3,1	-	-	-
EC042_RS09160 (cell division inhibition protein DicB)	3,1	9,5	3,5	-
EC042_RS05380 (FMNH2-dependent alkanesulfonate monooxygenase)	3,1	3,5	4,8	-
EC042_RS11085 (RNA polymerase sigma factor FliA)	3,1	-	-	-
EC042_RS11220 (flagellar type III secretion system protein fliR)	3,1	4,4	4,1	-
EC042_RS06040 (flagellar assembly peptidoglycan hydrolase FlgJ)	3,0	-	-	-
EC042_RS20235 (heme ABC transporter ATP-binding protein)	3,0	-	-	-
EC042_RS05995 (flagellar basal body P-ring formation protein flgA)	3,0	-	-	-

**Tabla 4.9.** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042 en ausencia del plásmido pAA2, expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. También se indica el valor de expresión relativa de los mismos genes en los mutantes  $\Delta hha$ ,  $\Delta hns$  y  $\Delta hns\Delta hns2$ . Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

Tal y como se puede observar, la pérdida del plásmido pAA2 provoca una gran desregulación en la expresión de un gran número de genes cromosómicos. Los principales grupos de genes que muestran una mayor expresión en ausencia del plásmido pAA2 en relación a cuando dicho plásmido se encuentra presente en la cepa 042 corresponden a genes implicados en la respuesta a estrés ácido (entre los que destacan los genes del operón *hdeAB*, genes que codifican para diferentes isoformas de la glutamato descarboxilasa y el regulador GadE, regulador transcripcional implicado en la respuesta a estrés ácido y activador central de la vía de resistencia a ácido dependiente de glutamato (GDAR, *glutamate-dependent acid resistance*) (Ma y col., 2003)) y genes implicados en la síntesis flagelar Flag-1 (entre los que destacan genes codificados en dos operones independientes (*flg* y *fli*)). Dicho sistema flagelar forma parte del sistema de secreción tipo III y es responsable de conferir el serotipo H18 característico de dicha cepa bacteriana (Chaudhuri y col., 2010)). De manera sorprendente, el perfil transcriptómico de la cepa pAA2- coincide mayoritariamente con el perfil transcriptómico de la cepa  $\Delta hha$  y, en menor medida, al de los mutantes  $\Delta hns$  y  $\Delta hns\Delta hns2$ , a excepción de los genes implicados en la síntesis flagelar, donde no se cumple el patrón de regulación mencionado.

En paralelo con la obtención de estos resultados, se reportó que la inactivación del gen *aar* del plásmido pAA2 afectaba a la expresión de los genes *hns* y *hns2* (Santiago y col., 2017). Este hecho permite interpretar los datos transcriptómicos obtenidos en este trabajo (apartado Discusión).

En referencia a los genes reprimidos en la cepa *E. coli* 042 como consecuencia de la pérdida del plásmido pAA2 (tabla 4.10), la mayor parte de los mismos corresponden a genes cromosómicos regulados por AggR (principalmente genes codificados en la isla cromosómica que conforma el sistema de secreción tipo VI Aai que abarca de los genes EC042\_4562 hasta el EC042\_4583). Estos últimos resultados eran esperables ya que la pérdida del plásmido pAA2 supone la consecuente pérdida de la proteína AggR, por lo que la expresión de todos los genes cromosómicos que dependían de este activador transcripcional se ve alterada negativamente.

<i>E. coli</i> 042 pAA2- vs <i>E. coli</i> 042 WT	
Identificador NCBI ( <i>locus tag</i> )	Fold change
EC042_RS24370/EC042_4564 (type VI secretion system protein AaiC/Hcp2)	-18,3
EC042_RS24365/EC042_4563 (type VI secretion system contractile sheath large subunit)	-17,9
EC042_RS24360/EC042_4562 (type VI secretion system contractile sheath small subunit)	-16,1
EC042_RS24375 (hypothetical protein)	-12,3
EC042_RS24390/EC042_4569 (type VI secretion system tip protein VgrG)	-11,5
EC042_RS24385/EC042_4568 (type VI secretion system baseplate subunit TssG)	-11,4
EC042_RS12055 (type IV toxin-antitoxin system toxin CbtA)	-11,3
EC042_RS24380/EC042_4566 (type VI secretion system baseplate subunit TssF)	-10,9
EC042_RS24395 (hypothetical protein)	-10,5
EC042_RS24450 (hypothetical protein)	-10,5
EC042_RS24435/EC042_4577 (type VI secretion system ATPase TssH)	-10,2
EC042_RS24405/EC042_4572 (type VI secretion system protein TssA)	-10,1
EC042_RS24410/EC042_4573 (type VI secretion protein)	-10,0
EC042_RS24430 (hypothetical protein)	-9,4
EC042_RS24400/EC042_4571 (type VI secretion protein)	-9,2
EC042_RS24415/EC042_4574 (type VI secretion system baseplate subunit TssK)	-8,5
EC042_RS24465 (hypothetical protein)	-8,0
EC042_RS28315 (transposase)	-7,5
EC042_RS24420 (hypothetical protein)	-7,2
EC042_RS17020 (hypothetical protein)	-6,3
EC042_RS24425 (DotU family type IV/VI secretion system protein)	-5,9



## 4. RESULTADOS

EC042_RS17010 (hypothetical protein)	-5,7
EC042_RS24460 (hypothetical protein)	-5,3
EC042_RS27045 (IS3 family transposase)	-4,7
EC042_RS17000 (ParB N-terminal domain-containing protein)	-4,3
EC042_RS24485 (IS66 family transposase)	-3,8
EC042_RS24470 (IS66 family transposase)	-3,8
EC042_RS12580 (hypothetical protein)	-3,8
EC042_RS28895 (IS3 family transposase)	-3,7
EC042_RS17015 (hypothetical protein)	-3,7
EC042_RS21290 (hypothetical protein)	-3,6
EC042_RS24525 (DDE-type integrase/transposase/recombinase)	-3,5
EC042_RS12995 (DUF4222 domain-containing protein)	-3,5
EC042_RS12255 (GDP-mannose mannosyl hydrolase)	-3,5
EC042_RS16990 (hypothetical protein)	-3,3
EC042_RS17030 (AAA family ATPase)	-3,3
EC042_RS17410 (glycolate oxidase subunit GlcE)	-3,2
EC042_RS17005 (phosphoadenosine phosphosulfate reductase)	-3,0
EC042_RS15075 (DNA repair protein RecN)	-2,9
EC042_RS24475 (IS110 family transposase)	-2,9
EC042_RS16515 (putative aminohydrolase SsnA)	-2,9
EC042_RS23675 (c-di-GMP phosphodiesterase PdeC)	-2,8
EC042_RS24965 (hypothetical protein)	-2,8
EC042_RS01275 (integrating conjugative element protein)	-2,7
EC042_RS17415 (glycolate oxidase subunit GlcD)	-2,7
EC042_RS24490 (IS110-like element ISEc45 family transposase)	-2,7
EC042_RS29420 (hypothetical protein)	-2,6
EC042_RS16995 (PerC family transcriptional regulator)	-2,6
EC042_RS07935 (hypothetical protein)	-2,5

**Tabla 4.10.** Listado de los genes cuya expresión resulta mayormente reprimida en el cromosoma de la cepa *E. coli* 042 en ausencia del plásmido pAA2, expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

### 4.3.2 El gen *aggR* y su región 3'UTR: regulación de la virulencia en la cepa *E. coli* 042

#### 4.3.2.1 Expresión transcripcional del gen *aggR*

Además de analizar el efecto del curado del plásmido pAA2 sobre la fisiología de la cepa *E. coli* 042, también se quiso aportar nuevos datos sobre la regulación de la expresión del principal regulador de la expresión de genes de virulencia codificados en este plásmido,

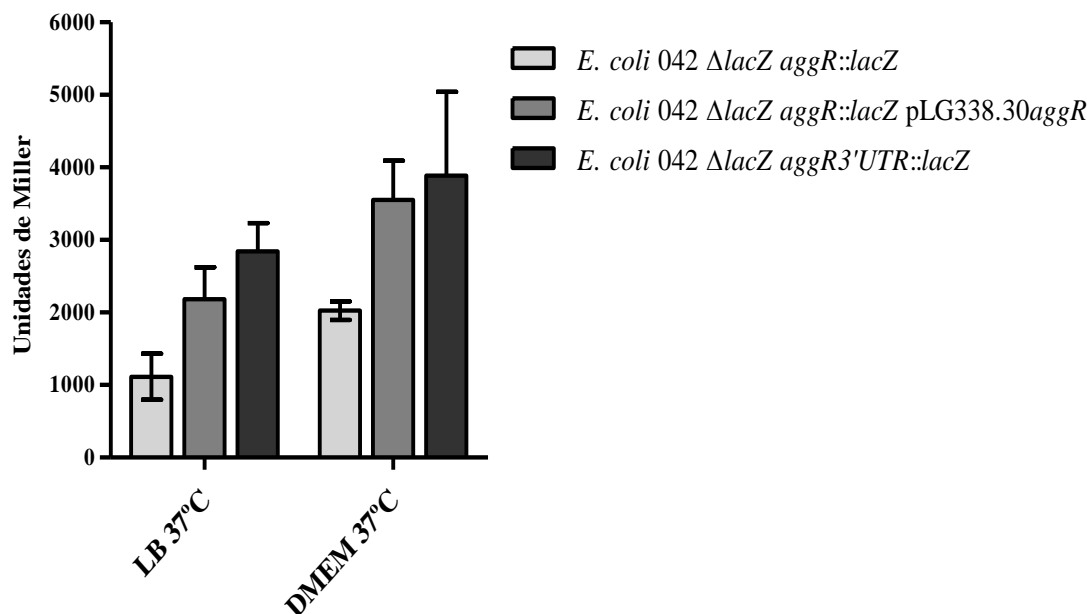
el gen *aggR*. Dicho activador transcripcional, responsable de la expresión de un gran número de factores de virulencia tanto plasmídicos como cromosómicos, presenta una regulación transcripcional dependiente de temperatura y de la fase de crecimiento bacteriano, expresándose a unos niveles superiores a 37°C y en la fase estacionaria de crecimiento respecto a cuando las células crecen a 25°C o se encuentran en la fase de crecimiento exponencial (Hüttener y col., 2018).

Decidimos, en primer lugar, confirmar estos resultados utilizando fusiones transcripcionales cromosómicas del gen reportero *lacZ* con el gen *aggR*. Sin embargo, previamente hay que tener en cuenta que la expresión del gen *aggR* presenta fenómenos de autorregulación por parte de su producto génico (Morin y col., 2010). Por ello, ya que el hecho de generar una fusión transcripcional en el cromosoma de la bacteria con el gen *lacZ* supone eliminar la secuencia codificante del gen diana de la fusión, para realizar estos experimentos se decidió complementar a la cepa que presenta dicha fusión cromosómica con el gen *aggR* aportado en *trans* a través de un plásmido. Con esta finalidad, se decidió introducir el plásmido pLG338.30 con el gen *aggR* clonado bajo el control de su propio promotor en la cepa *E. coli* 042 *aggR::lacZ* (generando de esta manera la cepa *E. coli* 042 *aggR::lacZ* pLG338.30*aggR*).

Sin embargo, a pesar de que el plásmido pLG338.30 sea un vector de bajo número de copias, la dosis génica del gen *aggR* aportada a través del plásmido pLG338.30*aggR* será más elevada que la dosis génica del gen *aggR* presente en condiciones fisiológicas. Por todo ello, y con el fin de analizar los niveles de expresión del gen *aggR* bajo condiciones fisiológicas sin necesidad de incrementar los niveles de proteína AggR de manera exógena, se decidió insertar la fusión transcripcional del gen *lacZ* justo después del codón de *stop* del gen *aggR*, generando la cepa *E. coli* 042 *aggR3'UTR::lacZ*.

De esta manera, se decidió analizar el nivel de expresión transcripcional del gen *aggR* mediante ensayos de  $\beta$ -galactosidasa en las cepas *E. coli* 042 *aggR::lacZ* (cepa que no presenta mecanismos de autoactivación de la expresión de *aggR*), *E. coli* 042 *aggR::lacZ* pLG338.30*aggR* y *E. coli* 042 *aggR3'UTR::lacZ*. Los ensayos se realizaron a 37°C a partir de cultivos crecidos hasta el inicio de la fase de crecimiento estacionaria ( $D.O_{600nm}=2,0$ ), tanto en medio LB como en medio DMEM suplementado con glucosa 0,45% (p/v), condiciones descritas en las que se induce el regulón de AggR (Sheikh y col., 2001; Morin y col., 2013) (figura 4.55).

#### 4. RESULTADOS



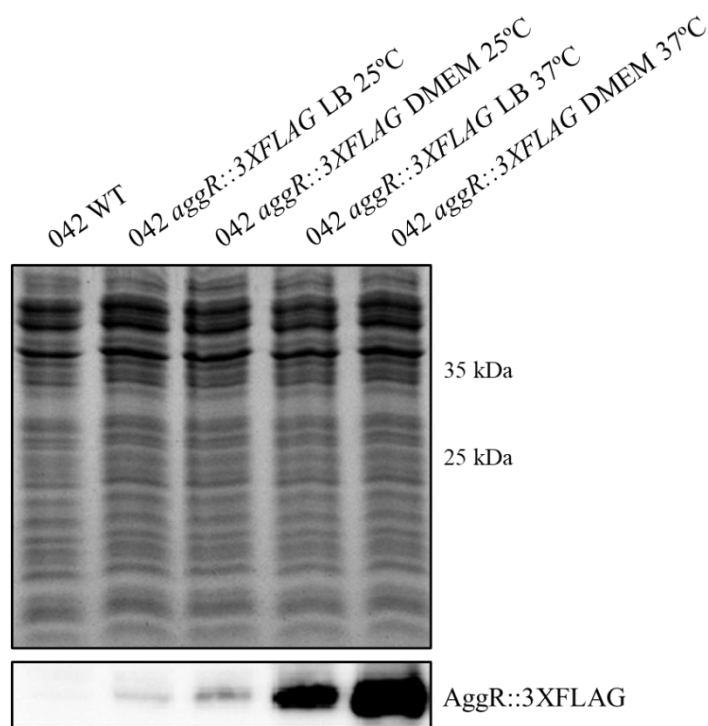
**Figura 4.55.** Expresión transcripcional del gen *aggR* cuantificada mediante ensayos de  $\beta$ -galactosidasa a partir del gen reportero *lacZ* fusionado en el cromosoma de la cepa *E. coli* 042. Dicha expresión transcripcional se cuantificó a partir de cultivos de las cepas *E. coli* 042  $\Delta lacZ$  *aggR::lacZ*, *E. coli* 042  $\Delta lacZ$  *aggR::lacZ* pLG338.30*aggR* y *E. coli* 042  $\Delta lacZ$  *aggR3'UTR::lacZ* crecidos a 37°C en medio LB o medio DMEM suplementado con glucosa 0,45% (p/v) hasta el inicio de la fase estacionaria de crecimiento. Se representa el promedio y la desviación estándar correspondientes a tres réplicas biológicas independientes.

Analizando los resultados representados en la figura 4.55, puede observarse cómo, en efecto, la expresión transcripcional del gen *aggR* es mayor cuando las bacterias crecen en medio DMEM suplementado con glucosa al 0,45% (p/v) respecto a cuando lo hacen en medio rico LB. Dichos resultados concuerdan con resultados publicados previamente (Sheikh y col., 2001; Morin y col., 2013), en los que se indica que el medio DMEM a 37°C son las condiciones inductoras de la expresión de los genes regulados por parte del regulador transcripcional AggR. Además, cuando la cepa que presenta la fusión transcripcional *lacZ* en la secuencia codificante del gen *aggR* se complementa con el plásmido pLG338.30*aggR*, los valores de expresión transcripcional de este regulador se incrementan aproximadamente en 1000 Unidades de Miller. Ello nos indica que la presencia de la proteína AggR es necesaria para la autoactivación de su propio promotor. Dichos datos se confirman al observar los valores de la expresión transcripcional del gen *aggR* cuando la fusión del gen reportero *lacZ* se inserta justo después del codón de *stop* del propio regulador, dejando intacta su secuencia codificante. En este caso, los valores de actividad  $\beta$ -galactosidasa también permiten confirmar la autoactivación del gen *aggR*

por parte de su producto génico; además, dichos valores son ligeramente superiores a los obtenidos en la complementación en *trans* con el plásmido pLG338.30*aggR*.

#### 4.3.2.2 Análisis del perfil de expresión de la proteína AggR

Los estudios de expresión del gen *aggR* se completaron analizando también la expresión de la proteína AggR. Para ello, se marcó con el epítipo 3XFLAG el extremo C-terminal de la proteína AggR, cuantificándose sus niveles de expresión por inmunodetección a partir de cultivos crecidos tanto en medio rico LB como en medio DMEM (suplementado con glucosa al 0,45% (p/v)), incubándose los mismos tanto a 25°C como a 37°C hasta el inicio de la fase estacionaria de crecimiento (figura 4.56).



**Figura 4.56.** Análisis por *Western Blot* de la expresión de la proteína AggR::3XFLAG en medio rico LB y en medio DMEM tanto a 25°C como a 37°C en el inicio de la fase estacionaria de crecimiento ( $D.O_{600nm}$  2,0). El panel superior muestra la tinción de Coomassie de cada uno de los extractos proteicos analizados y representa el control de carga en el que se muestra que la proteína total analizada para cada extracto es similar en todas las condiciones estudiadas.

Puede comprobarse como los niveles de expresión proteica de AggR son totalmente dependientes de la temperatura (niveles de expresión mucho mayores cuando las bacterias crecen a 37°C respecto a 25°C) y del medio de crecimiento de las bacterias (nivel de expresión mayor cuando las bacterias crecen en medio DMEM respecto a cuando lo hacen en medio LB). Estos datos de expresión proteica del regulador transcripcional AggR

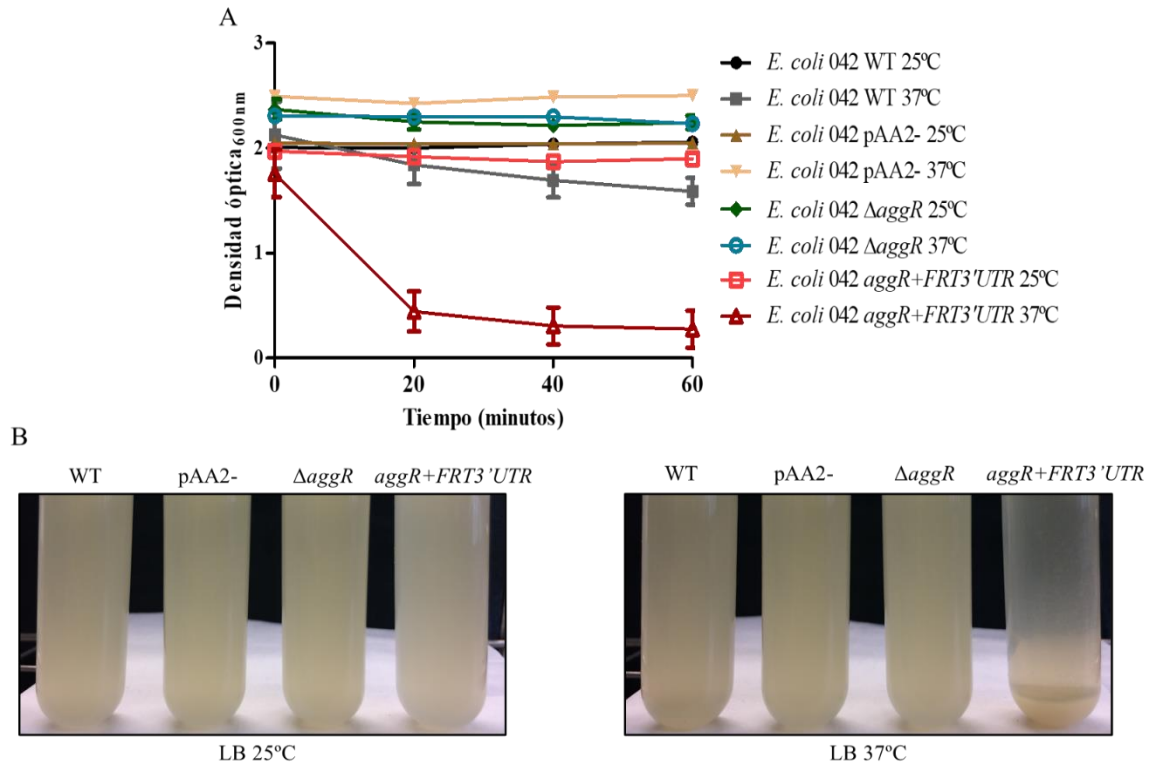
#### 4. RESULTADOS

concuerdan con los perfiles de expresión transcripcional del gen *aggR* previamente descritos por parte de (Sheikh y col., 2001; Morin y col., 2013; Hüttener y col., 2018).

##### **4.3.2.3 Alteración del fenotipo bacteriano al generar una inserción nucleotídica entre el codón de finalización de traducción y la región 3'UTR del gen *aggR***

De manera inesperada, al realizar estrías en placas de LB-agar del clon que contenía el gen reportero *lacZ* insertado justo después del codón de finalización de la traducción del gen *aggR* (*E. coli* 042  $\Delta$ *lacZ aggR3'UTR::lacZ*), las colonias de este mutante presentaban una morfología y un perfil de adherencia a la superficie del agar muy diferente al de la cepa salvaje. Además, cuando a partir de dichas colonias se realizaba un cultivo en medio líquido, rápidamente se podía observar como las células se agrupaban entre ellas formando agregados celulares macroscópicos que sedimentaban en fondo del tubo, presentando un fenotipo altamente autoagregativo. Dicho fenotipo se ha asociado comúnmente con bacterias que presentan mecanismos patogénicos, ya que este proceso está relacionado con incrementos en la tolerancia a agentes antimicrobianos (Corno y col., 2014) y con la capacidad invasiva de la bacteria y su resistencia al sistema inmune del huésped (Galdiero y col., 1988). Todo ello es debido al hecho de que, cuando las células se unen entre ellas y autoagregan, generan un microambiente en el que las bacterias se encuentran protegidas de condiciones ambientales adversas y de las defensas del hospedador, permitiendo su crecimiento y supervivencia a lo largo del tiempo (Trunk y col., 2018). Para analizar con más detalle este fenotipo autoagregativo, se realizaron ensayos para cuantificar la sedimentación celular siguiendo el protocolo descrito en el apartado 3.10.2. Con el fin de descartar que el fenotipo observado fuera debido a posibles efectos polares derivados del gen reportero *lacZ* insertado y del casete de resistencia a kanamicina que dicha construcción lleva asociado, se quiso analizar si los efectos observados se mantenían en la cepa original (cepa que contenía una cicatriz FRT después del codón de terminación de la traducción del gen *aggR*, lugar en el que se inserta el gen *lacZ* siguiendo el protocolo descrito por (Ellermeier y col., 2002)). Para realizar los ensayos de autoagregación, se crecieron los cultivos de las cepas *E. coli* 042 salvaje, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ *aggR* y *E. coli* 042 *aggR*+FRT3'UTR (cepa que contiene la secuencia FRT después del codón *stop* del gen *aggR*) en medio LB durante 16 horas a 25°C y a 37°C en agitación constante. Una vez crecidos, se transfirieron 10 ml de cada cultivo a un tubo de ensayo y se mantuvieron en estático, tomándose alícuotas de la parte

superior del tubo para medir la turbidez de cada muestra en el espectrofotómetro. La autoagregación se representó como el descenso de la absorbancia a 600 nm como consecuencia de la sedimentación de las células hacia la base del tubo (figura 4.57).



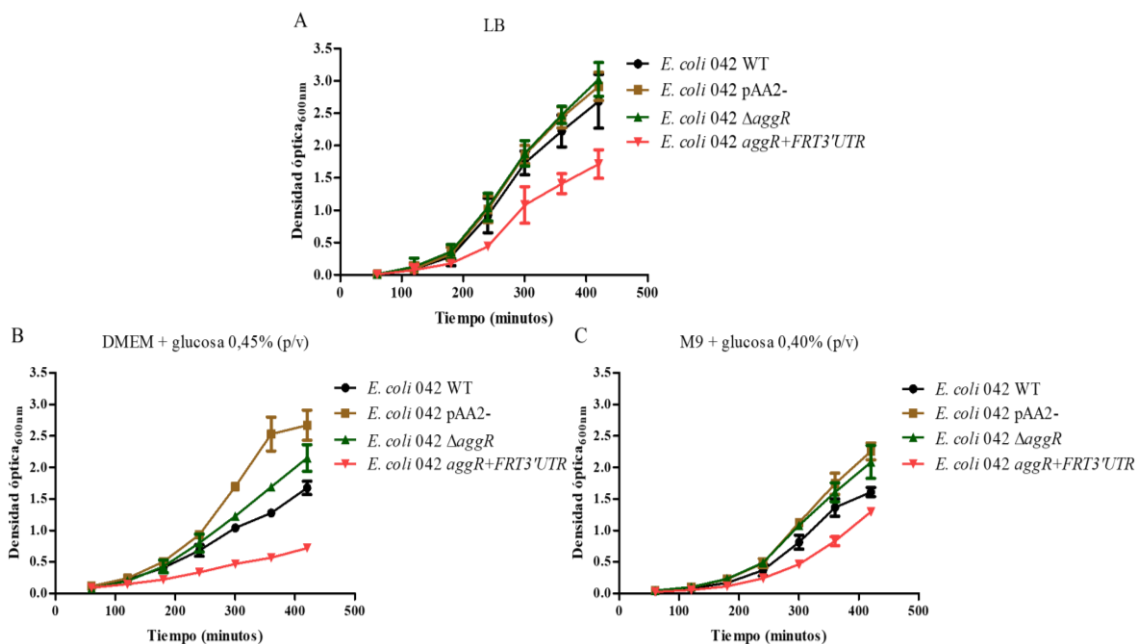
**Figura 4.57.** (A) Representación gráfica de la autoagregación bacteriana cuantificada a través del descenso de la absorbancia observada al dejar reposar en condiciones estáticas los cultivos crecidos en medio LB a 25°C o 37°C de las cepas *E. coli* 042 WT, pAA2-,  $\Delta aggR$  y *aggR*+FRT3'UTR. (B) Imagen representativa de la sedimentación celular (causada por la autoagregación de las bacterias) de los cultivos mencionados anteriormente después de dejarlos reposar en condiciones estáticas durante 60 minutos.

Tal y como se puede observar en la figura 4.57 (A y B), la cepa que presenta la disrupción del extremo 3' del gen *aggR* causada por la adición de una secuencia FRT entre el codón *stop* del gen y su región 3'UTR (*untranslated region*) presenta un fenotipo altamente autoagregativo a 37°C, observado al generarse un notable descenso de la absorbancia de las muestras recogidas en la superficie del tubo como consecuencia de la sedimentación celular hacia la base del mismo. La cepa salvaje también presenta un ligero fenotipo autoagregativo a 37°C, pero a unos niveles muy inferiores a los de la cepa portadora de la construcción *aggR*+FRT3'UTR. Dicho efecto es completamente dependiente de la temperatura de crecimiento de los cultivos, ya que a 25°C el efecto se ve totalmente anulado. Ni la cepa que presenta la delección del gen *aggR* ni la cepa curada del plásmido

#### 4. RESULTADOS

pAA2 presentaban este fenotipo autoagregativo. Ello es debido a que este fenómeno autoaglutinante está mediado por interacciones homotípicas entre las bacterias en las que intervienen múltiples factores, entre los que destacan exopolisacáridos y proteínas de superficie, tales como pili, flagelos y fimbrias. En el caso de la cepa *E. coli* 042, las fimbrias de adherencia agregativa AAF/II se hallan codificadas en el plásmido pAA2 y, además, su expresión se encuentra bajo el control del regulador transcripcional AggR (Nataro y col., 1994), por lo que en ausencia de las fimbrias (en el caso de la cepa *E. coli* 042 pAA2-) y en ausencia de su activador transcripcional (en el caso de la cepa *E. coli* 042  $\Delta$ aggR) no se observa este fenotipo autoagregativo.

Paralelamente, también se quiso analizar a nivel de *fitness* celular cómo afectaban tanto la ausencia del plásmido pAA2, la mutación en el gen *aggR* y la construcción *aggR+FRT3'UTR* a nivel de la tasa de crecimiento bacteriano, comparándose dichos efectos con los valores correspondientes a la cepa salvaje. Para ello, se realizaron curvas de crecimiento a 37°C en medio rico LB, medio DMEM suplementado con glucosa al 0,45 % (p/v) y medio mínimo M9 suplementado con glucosa al 0,40 % (p/v) de las cepas mencionadas anteriormente (figura 4.58).



**Figura 4.58.** Curvas de crecimiento de las cepas *E. coli* 042 salvaje (WT), *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ aggR y *E. coli* 042 *aggR+FRT3'UTR* en medio LB (A), DMEM suplementado con glucosa 0,45% (p/v) (B) y M9 suplementado con glucosa 0,40% (p/v) (C) a 37°C. Se representan los valores correspondientes a 3 réplicas biológicas independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

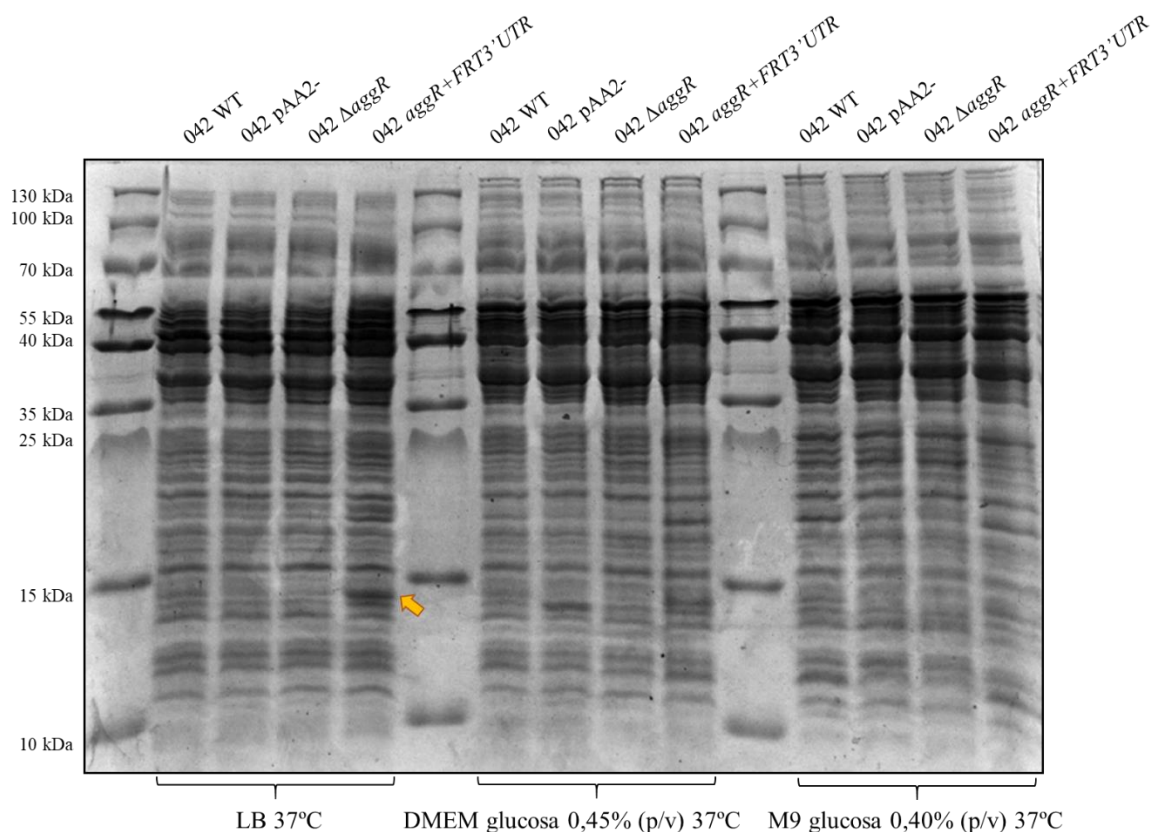


Analizando las cinéticas de crecimiento de los diferentes mutantes estudiados, se puede observar como la pérdida del plásmido pAA2 y la delección del gen *aggR* generan un efecto positivo a nivel de velocidad de crecimiento (en referencia a la cepa salvaje) independientemente de la riqueza de nutrientes del medio de cultivo utilizado (aunque la diferencia de crecimiento respecto a la cepa salvaje se hace más plausible cuando las bacterias crecen en medio DMEM suplementado con glucosa 0,45% (p/v), condiciones inductoras de la expresión de la proteína AggR). En cambio, se observa un efecto totalmente opuesto a nivel de la cinética de crecimiento de la cepa que presenta la inserción de la secuencia FRT en el extremo 3' del gen *aggR*, cuya tasa de crecimiento es inferior a la de la cepa salvaje y mucho menor que la de las cepas pAA2- y  $\Delta aggR$ .

#### **4.3.2.3.1 Análisis del contenido proteico intracelular de la cepa *E. coli* 042 *aggR*+*FRT*3'*UTR***

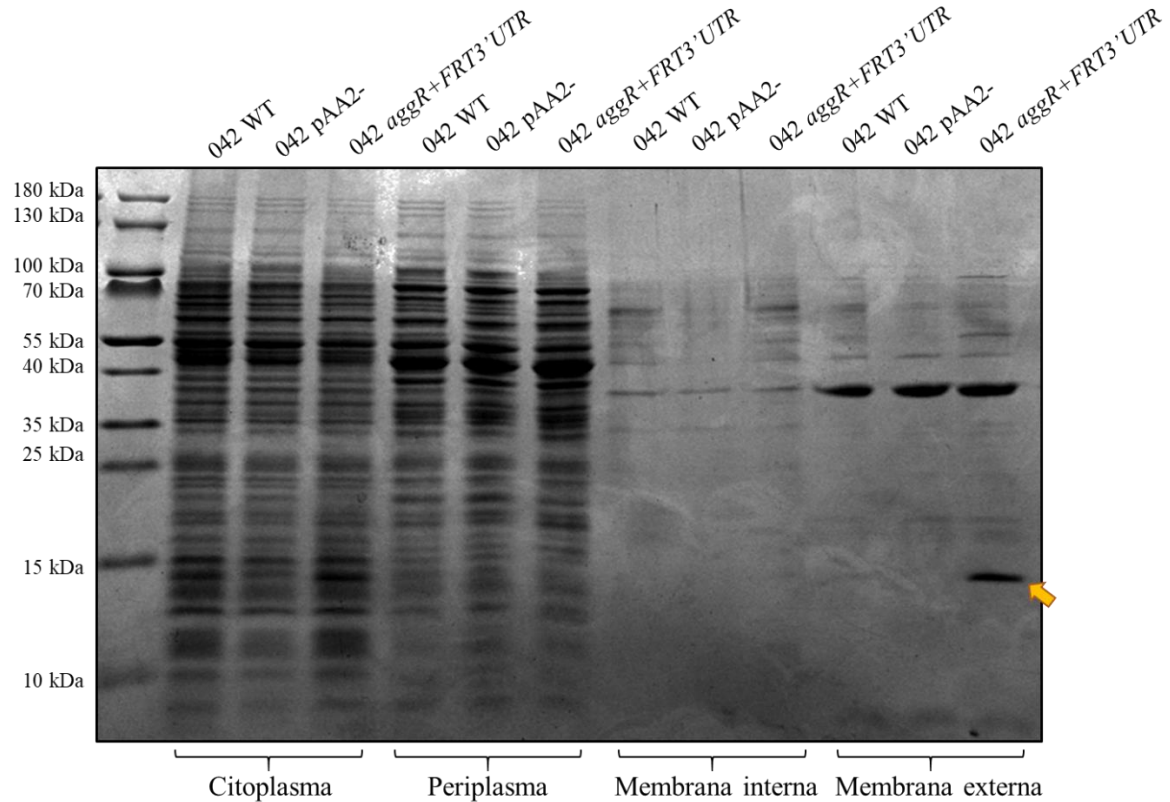
Con el fin de seguir caracterizando la cepa *E. coli* 042 *aggR*+*FRT*3'*UTR*, se decidió analizar su perfil proteico intracelular y se comparó con los perfiles proteicos de la cepa salvaje, la cepa curada del plásmido pAA2 y la cepa mutante  $\Delta aggR$ , a partir de cultivos crecidos a 37°C en medio LB, medio DMEM suplementado con glucosa al 0,45% (p/v) y medio mínimo M9 suplementado con glucosa al 0,40% (p/v). Para analizar el contenido proteico de los diferentes mutantes analizados en cada condición, se crecieron los cultivos correspondientes hasta el inicio de la fase estacionaria de crecimiento ( $D.O_{600nm} = 2,0$ ) a partir de inóculos con una dilución 1:100 de un cultivo previo. Para extraer el contenido proteico de cada cultivo se siguió el protocolo descrito en el apartado 3.9.1 de Materiales y Métodos, con el fin de normalizar la concentración de proteína obtenida en función de la biomasa del cultivo utilizado. Las proteínas de dichos extractos se analizaron posteriormente mediante electroforesis en geles SDS-PAGE y tinción con Azul de Coomassie (figura 4.59).

#### 4. RESULTADOS



**Figura 4.59.** Perfil electroforético del contenido proteico intracelular aislado a partir de cultivos de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ aggR y *E. coli* 042 aggR+FRT3'UTR, crecidas a 37°C en los medios LB, DMEM suplementado con glucosa 0,45% (p/v) y M9 suplementado con glucosa 0,40% (p/v) hasta el inicio de la fase estacionaria de crecimiento. La flecha amarilla indica la proteína sobreexpresada en el fondo genético aggR+FRT3'UTR. Las proteínas se analizaron mediante electroforesis en geles de acrilamida en presencia de SDS y posterior tinción con Azul de Coomassie.

En la figura 4.59 se puede observar que existe una banda correspondiente a una proteína sobreexpresada en la cepa *E. coli* 042 aggR+FRT3'UTR en relación con la cepa salvaje, la cepa curada del plásmido pAA2 y el mutante aggR sencillo. Con el fin de localizar esta proteína en un compartimento celular (citoplasma, periplasma, membrana interna o membrana externa) se decidió realizar un fraccionamiento celular. Para obtener las diferentes fracciones pertenecientes a cada uno de los compartimentos celulares se siguió el protocolo descrito en el apartado 3.9.3 de Materiales y Métodos. Se aisló el contenido proteico de las diferentes fracciones celulares de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2- y *E. coli* 042 aggR+FRT3'UTR a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento (figura 4.60).



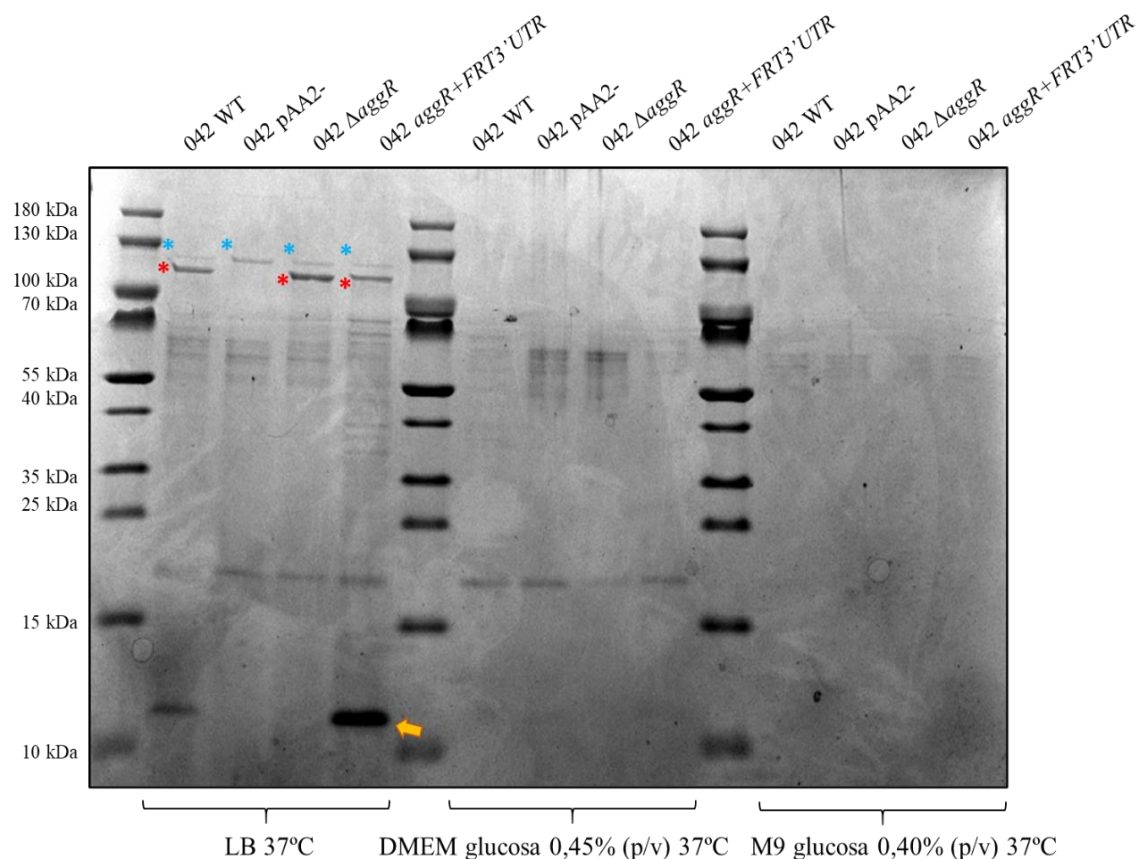
**Figura 4.60.** Perfil electroforético de las proteínas pertenecientes a las fracciones de citoplasma, periplasma, membrana interna y membrana externa de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2- y *E. coli* 042 *aggR+FRT3'UTR*, aisladas a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria. La flecha amarilla indica la proteína de membrana externa sobreexpresada en el fondo genético *aggR+FRT3'UTR*. Las proteínas se analizaron mediante electroforesis en gel SDS-PAGE y posterior tinción con Azul de Coomassie.

En la figura 4.60 se puede observar como la proteína sobreexpresada en la cepa *E. coli* 042 *aggR+FRT3'UTR* se localiza en la fracción celular correspondiente a la membrana externa bacteriana. Con el fin de identificar a qué proteína corresponde la banda proteica sobreexpresada, se recortó la banda de interés del gel de poliacrilamida y se analizó mediante cromatografía de gases acoplada a un espectrómetro de masas. Los resultados mostraron que la proteína objeto de interés sobreexpresada cuando se altera la región comprendida entre la secuencia codificante y la región 3'UTR del gen *aggR* corresponde a la subunidad mayor de las fimbrias de adherencia agregativa de tipo II (AafA). Este resultado está en concordancia con el elevado fenotipo autoagregativo observado en la cepa objeto de estudio, fenotipo que se podría explicar por el elevado incremento en la expresión de las fimbrias agregativas en el fondo genético *aggR+FRT3'UTR* respecto a la cepa salvaje.

#### 4. RESULTADOS

##### 4.3.2.3.2 Análisis del secretoma de la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR*

Paralelamente al análisis del contenido proteico intracelular de los diferentes mutantes estudiados, también se quiso analizar el perfil de proteínas que dichas cepas eran capaces de secretar al medio extracelular. Para ello, se utilizó el protocolo descrito en el apartado 3.9.2 de Materiales y Métodos. Con este fin, se analizó el secretoma de cultivos crecidos hasta el inicio de la fase estacionaria de crecimiento de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ *aggR* y *E. coli* 042 *aggR*+*FRT3*'*UTR* en medio rico LB, medio DMEM suplementado con glucosa al 0,45% (p/v) y medio mínimo M9 suplementado con glucosa al 0,40% (p/v) a 37°C. Las proteínas de dichos extractos se analizaron posteriormente mediante electroforesis en geles SDS-PAGE y tinción con Azul de Coomassie (figura 4.61).



**Figura 4.61.** Perfil electroforético obtenido por SDS-PAGE y posterior tinción con Azul de Coomassie de las proteínas secretadas al medio extracelular por parte de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ *aggR* y *E. coli* 042 *aggR*+*FRT3*'*UTR* crecidas a 37°C en medio LB, medio DMEM suplementado con glucosa 0,45% (p/v) y medio M9 suplementado con glucosa 0,40% (p/v) hasta el inicio de la fase estacionaria de crecimiento. Con unos asteriscos de color rojo y azul se indica la presencia en el secretoma de los cultivos mencionados anteriormente de las proteínas Pet y She, respectivamente. La flecha amarilla indica la proteína sobreexpresada en el fondo genético *aggR*+*FRT3*'*UTR*.

Se ha descrito que la cepa *E. coli* 042 es capaz de secretar al medio extracelular dos proteínas de elevado peso molecular (104 y 116 kDa). La proteína de 104 kDa corresponde al producto codificado en el gen *pet* del plásmido pAA2, y corresponde a un autotransportador con actividad enterotóxica (Eslava y col., 1998). Tal y como se puede observar en la figura 4.61, dicha proteína se encuentra presente en todas las cepas analizadas a excepción de la cepa curada del plásmido pAA2. La otra proteína de elevado peso molecular secretada al medio extracelular por parte de la cepa *E. coli* 042 tiene homología en su dominio N-terminal con la proteína She de *S. flexneri*, una proteína de la familia de proteasas de inmunoglobulinas de tipo A (Rajakumar y col., 1997). Esta proteína se encuentra codificada en el cromosoma de la cepa *E. coli* 042. Su presencia puede observarse en el secretoma de todas las cepas analizadas en la figura 4.61 cuando éstas crecen en medio LB, pero no cuando las bacterias lo hacen en medio mínimo M9 o DMEM.

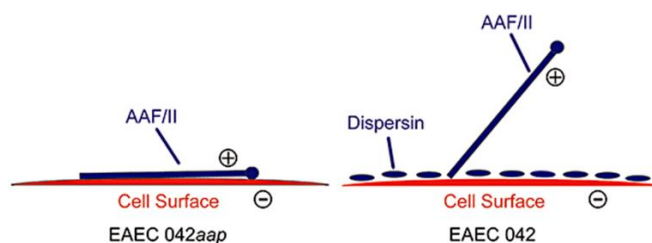
En el caso del secretoma de la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR* crecida en medio LB, se puede observar una proteína con un peso molecular de aproximadamente 10 kDa que aparece sobreexpresada en dicho fondo genético respecto a la cepa *E. coli* 042 WT. Dicha proteína no está presente ni en la cepa *E. coli* 042 pAA2- ni en la cepa mutante para el regulador transcripcional AggR. Con el fin de identificar a qué proteína corresponde la banda proteica sobreexpresada, se recortó dicha banda de interés del gel de poliacrilamida y se identificó mediante cromatografía de gases acoplada a un espectrómetro de masas. Los resultados de la identificación mostraron que la misma corresponde a la proteína dispersina (Aap). Dicha proteína, codificada en el plásmido pAA2, se encuentra bajo el control del regulador transcripcional AggR (Sheikh y col., 2002). Esta proteína es secretada al medio extracelular, pero permanece asociada de manera no covalente al lipopolisacárido (LPS) presente en la superficie celular bacteriana (Velarde y col., 2007).

La sobreexpresión de la subunidad mayor de las fimbrias de adherencia agregativa AafA y de la dispersina Aap en la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR* cuando los cultivos crecen en medio LB a 37°C concuerda con el elevado fenotipo autoagregativo de este mutante representado en la figura 4.57. Debido al elevado punto isoeléctrico de las fimbrias, estas estructuras proteicas presentan una elevada carga positiva en su superficie a pH fisiológico. En ausencia de la dispersina, las elevadas cargas positivas de las fimbrias interaccionarían electrostáticamente con la fuerte carga negativa de la superficie celular, impidiendo su correcto ensamblaje sobre la superficie bacteriana. De esta manera, la



#### 4. RESULTADOS

expresión de la dispersina permite el correcto ensamblaje y extensión de las fimbrias de adherencia agregativa sobre la superficie bacteriana al enmascarar y bloquear la fuerte atracción electrostática entre las fimbrias y la superficie bacteriana (Velarde y col., 2007) (figura 4.62).

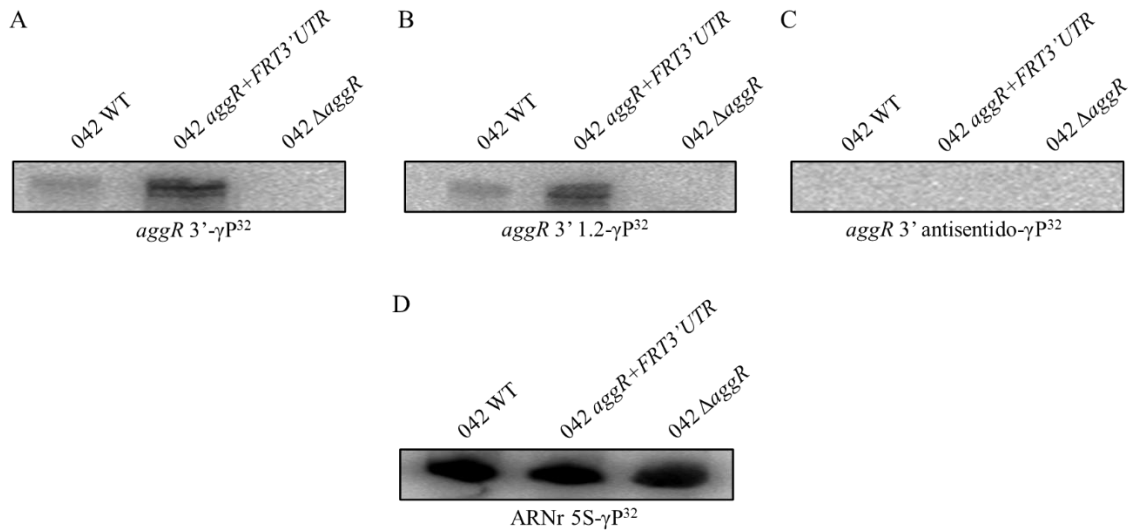


**Figura 4.62.** Representación esquemática del modelo propuesto sobre el papel de la dispersina en la superficie celular de la cepa *E. coli* 042. Dicha proteína es capaz de apantallar la fuerte atracción electrostática entre la carga positiva (+) de las fimbrias AAF/II y la carga negativa (-) de la superficie bacteriana, permitiendo de esta manera el correcto ensamblaje de las fimbrias sobre la membrana externa bacteriana. Imagen adaptada de (Velarde y col., 2007).

#### 4.3.2.4 Análisis del transcrito del gen *aggR* mediante *Northern Blot*

El hecho de que al alterar posibles secuencias reguladoras de la expresión y/o estabilidad del gen *aggR* provoque la sobreexpresión de las proteínas AafA y Aap, cuya expresión génica está mediada por el producto génico de dicho gen, la proteína AggR, puede estar indicando que al alterar estas secuencias génicas se esté provocando un incremento en la expresión de este regulador transcripcional. Para apoyar esta hipótesis, se analizaron mediante *Northern Blot* los niveles de transcripción del ARN mensajero del gen *aggR* en la cepa salvaje y se compararon con los niveles de expresión que se obtienen en la cepa que presenta la alteración en el extremo 3' del gen *aggR*. Para estudiar el transcrito del gen *aggR*, se cultivaron las cepas *E. coli* 042 WT, *E. coli* 042 *aggR*+*FRT3*'*UTR* y *E. coli* 042  $\Delta$ *aggR* (cepa utilizada como control de la especificidad de la sonda utilizada en la hibridación) en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento bacteriano. El contenido de ARN total de cada cultivo se aisló siguiendo el protocolo descrito en el apartado 3.8.1 de Materiales y Métodos, y se resolvió mediante electroforesis en un gel de poliacrilamida al 6% y elevada concentración de urea (8,3 M). Para analizar el nivel de expresión del ARN mensajero del gen *aggR* en los diferentes fondos genéticos se utilizaron dos sondas diferentes de 21 nucleótidos (que hibridaban en diferentes regiones del extremo 3' de la secuencia codificante del gen *aggR*) marcadas radioactivamente en su extremo 5' con ATP- $[\gamma P^{32}]$  (10 $\mu$ Ci/ $\mu$ l). Adicionalmente, también

se utilizó una sonda complementaria a la cadena antisentido del gen *aggR*, con el fin de localizar un posible ARN de interferencia cuya expresión se viera alterada al modificar el extremo 3' del gen *aggR* (figura 4.63).



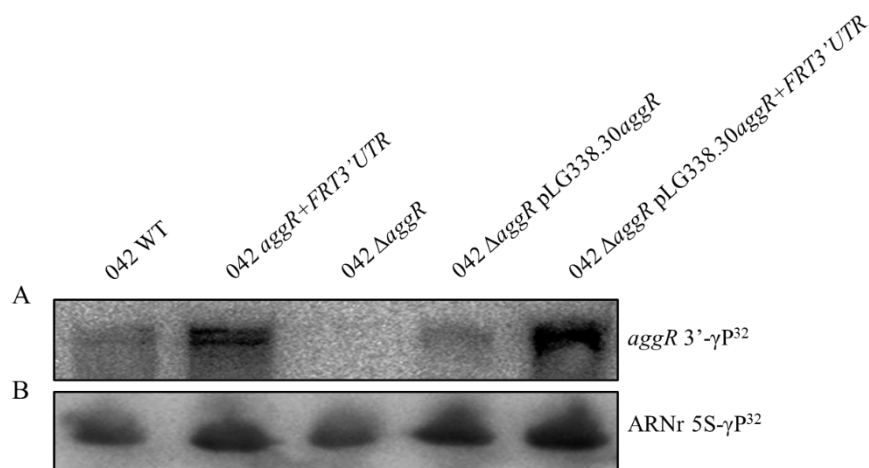
**Figura 4.63.** Análisis por *Northern Blot* de la expresión del ARN mensajero del gen *aggR* en las cepas *E. coli* 042 WT, *E. coli* 042 *aggR*+FRT3'UTR y *E. coli* 042  $\Delta$ *aggR* a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento. Para realizar el experimento, se analizaron 25  $\mu$ gr del ARN correspondiente a cada cepa mediante electroforesis en gel de poliacrilamida al 6% con elevadas concentraciones de urea (8,3 M). Los ácidos nucleicos fueron transferidos a una membrana de *nylon*, donde se realizó la incubación con las sondas correspondientes (*aggR* 3' (A), *aggR* 3' 1.2 (B) y *aggR* 3' antisentido (C)) marcadas radioactivamente con P<sup>32</sup> en su extremo 5'. Como control de la homogeneidad de carga se analizó la expresión del ARNr 5S (D) en los diferentes mutantes estudiados (debido a la elevada expresión de este ARNr, únicamente se utilizaron 10  $\mu$ gr de ARN para incubar con la sonda específica).

Analizando los resultados representados en la figura 4.63, puede observarse como el hecho de alterar el extremo 3' del gen *aggR* provoca un incremento en la expresión del ARN mensajero correspondiente a dicho gen respecto a los niveles de expresión del gen *aggR* en un fondo genético salvaje. En dicho experimento no se encontró expresión alguna de un ARN antisentido que hibridara en la región complementaria sobre la región en la que hibridaron las sondas que sí detectaron expresión de la secuencia codificante del gen *aggR*. La correcta especificidad de las sondas se pudo confirmar ya que en el mutante  $\Delta$ *aggR* no se observó hibridación con las sondas utilizadas. En la figura 9.20 del apartado Anexo se muestra la imagen completa del gel representado en la figura 4.63 después de realizar el revelado de la emisión por parte del isótopo radioactivo unido a las diferentes sondas utilizadas. De manera sorprendente, se puede observar como el tamaño del transcrito detectado correspondiente al gen *aggR* (más de 1.300 pb) es mucho mayor del tamaño esperable debido al tamaño del gen (900 pb).



#### 4. RESULTADOS

En un experimento posterior, se clonó en el plásmido de bajo número de copias pLG338.30 tanto el gen *aggR* con su secuencia 3' salvaje como el gen *aggR* con la secuencia que presenta la inserción en su extremo 3'. Ambos plásmidos se transformaron a un mutante *E. coli* 042  $\Delta$ *aggR*, generando las cepas *E. coli* 042  $\Delta$ *aggR* pLG338.30*aggR* y *E. coli* 042  $\Delta$ *aggR* pLG338.30*aggR*+*FRT3*'*UTR*. Utilizando dichas cepas, se analizó por *Northern Blot* la expresión del ARN mensajero del gen *aggR* clonado en el vector pLG338.30. De esta manera, al clonar las regiones salvaje y alteradas del extremo 3' del gen *aggR*, y analizarlas sobre un fondo genético sin el gen *aggR* cromosómico, se quiso analizar si los efectos fenotípicos observados en el mutante *E. coli* 042 *aggR*+*FRT3*'*UTR* eran generados únicamente como consecuencia de alterar el extremo 3' del gen *aggR*, o bien si al provocar esta alteración se estuvieran produciendo otros efectos secundarios sobre los genes situados aguas abajo en el lugar del cromosoma en el que tuvo lugar la modificación genética. Analizando los resultados representados en la figura 4.64 puede concluirse que, cuando la construcción que presenta la secuencia FRT insertada en el extremo 3' del gen *aggR* se clona en el plásmido pLG338.30, se genera un incremento en la expresión del gen *aggR* que no se observa cuando se analiza la construcción salvaje (además, dicho incremento es superior a la expresión observada en la construcción cromosómica, pero ello podría ser debido al mayor número de copias plasmídicas respecto a la única copia cromosómica del gen *aggR*).

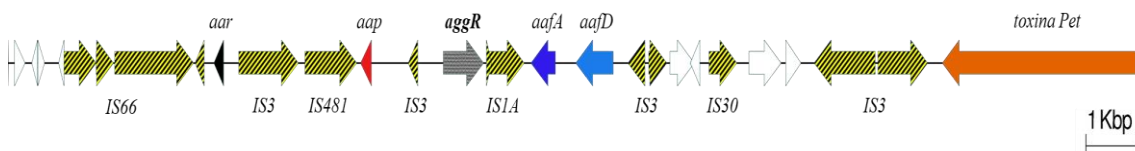


**Figura 4.64.** Análisis por *Northern Blot* de la expresión del ARN mensajero del gen *aggR* en las cepas *E. coli* 042 WT, *E. coli* 042 *aggR*+*FRT3*'*UTR*, *E. coli* 042  $\Delta$ *aggR*, *E. coli* 042  $\Delta$ *aggR* pLG338.30*aggR* y *E. coli* 042  $\Delta$ *aggR* pLG338.30*aggR*+*FRT3*'*UTR* aislado a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento. El nivel de expresión del gen *aggR* se analizó mediante la sonda *aggR* 3' (A) marcada radioactivamente con P<sup>32</sup> en su extremo 5'. Como control de homogeneidad de carga se analizó la expresión del ARNr 5S (B) en los diferentes mutantes estudiados.

Resumiendo los resultados presentados hasta ahora, puede afirmarse que el hecho de alterar la secuencia del extremo 3' del gen *aggR* provoca un incremento en los niveles de ARN mensajero de dicho gen y, en consecuencia, de la cascada regulatoria ejercida por este activador transcripcional, tal y como puede observarse en los incrementos de la expresión de los genes que forman parte del regulón de AggR (genes *aap* y *aafA*) observados en el mutante *E. coli* 042 *aggR*+*FRT3*'*UTR*.

#### 4.3.2.5 Análisis genómico de la región 3'UTR del gen *aggR*

Ya que la alteración de la región 3'UTR del gen *aggR* genera un incremento en la expresión transcripcional de este gen, lo que a su vez provoca grandes cambios en la fisiología de la cepa *E. coli* 042, nos propusimos analizar con más detalle la estructura de dicha región génica. En una primera aproximación, se investigó el entorno génico de la localización del gen *aggR* dentro del plásmido pAA2 de la cepa *E. coli* 042. En dicho análisis se ha podido comprobar que el gen *aggR* se encuentra situado entre un gran número de secuencias de inserción y de transposasas (figura 4.65). Concretamente, *aggR* se encuentra flanqueado aguas arriba por un elemento de inserción de la familia *IS3*, y aguas abajo por un elemento de inserción de la familia *IS1A*.

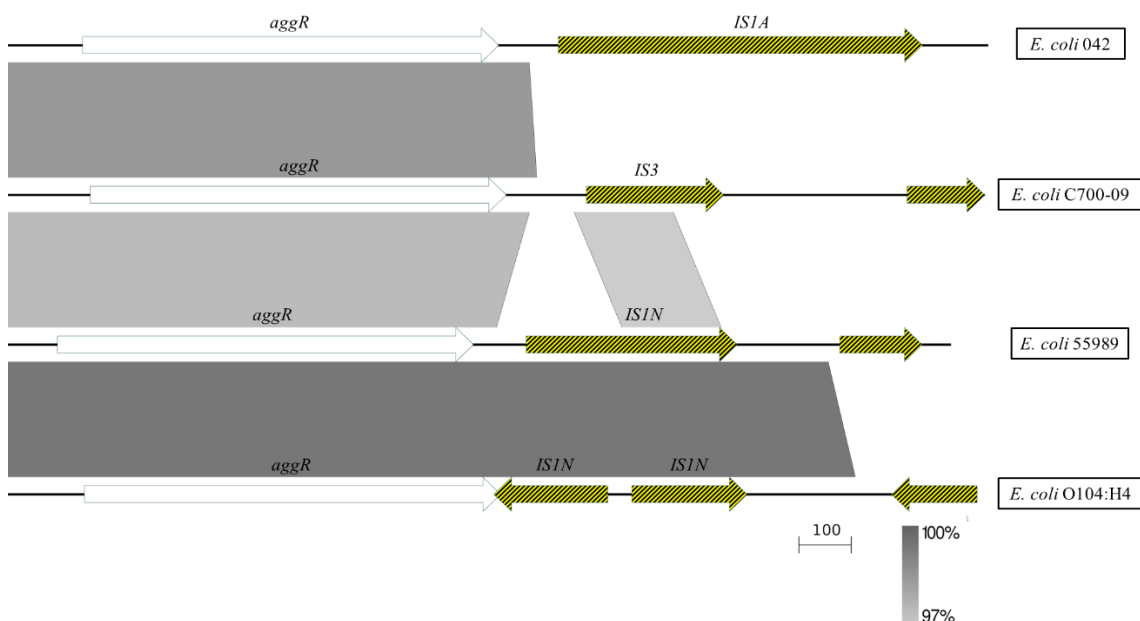


**Figura 4.65.** Representación esquemática de la región génica comprendida entre las posiciones nucleotídicas 28073 a 50530 del plásmido pAA2, correspondiente al entorno génico del gen *aggR* (EC042\_pAA052). En gris se indica el gen *aggR*, y en amarillo y negro se indican las diferentes transposasas y elementos de inserción presentes en la secuencia analizada. También se indican los genes *aar*, *aap*, *aafA*, *aafD* y la toxina Pet, todos ellos situados muy próximos al gen *aggR*. Anotación génica extraída de *GenBank* FN554767.1. Representación realizada mediante el programa *EasyFig* (Sullivan y col., 2011).

Seguidamente se quiso comprobar si, tal y como ocurre en el caso del gen *aggR* de la cepa 042, el hecho de estar rodeado por secuencias de inserción y transposasas también se cumplía en el caso del gen *aggR* presente en otros plásmidos pAA de otras cepas enteroagregativas de *E. coli*. También se quiso delimitar, dentro de la región intergénica entre el gen *aggR* y el elemento de inserción *IS1A* (región que abarca desde el codón de finalización de la traducción del gen *aggR* y el codón de inicio de la traducción de la transposasa *IS1A*), qué región génica forma parte de la región no codificante del gen *aggR*

#### 4. RESULTADOS

(3'UTR) y qué región ya forma parte del elemento de inserción en sí. Para ello, se analizó bioinformáticamente tanto la región codificante del gen *aggR* como la región génica presente aguas abajo de dicho gen en otras cepas de *E. coli*, con el fin de localizar una región presente en todas las cepas portadoras del gen *aggR* que definiera una región 3'UTR común en todas las cepas. Concretamente, se analizó y se comparó la región situada aguas abajo del gen *aggR* en los plásmidos pAA de las cepas enteroagregativas de *E. coli* 042, *E. coli* C700-09, *E. coli* 55989 y en la cepa responsable del brote se SUH en Alemania en 2011, la cepa *E. coli* O104:H4 (figura 4.66). Se ha descrito recientemente que cepa *E. coli* C700-09 contiene una nueva variante del plásmido pAA no descrita hasta la fecha que codifica dos clases de fimbrias agregativas (AAF/III y AAF/V) (Jonsson y col., 2017), mientras que la cepa 55989 ha sido reconocida como la cepa prototipo EAEC que expresa las fimbrias de tipo III (AAF/III) (Bernier y col., 2002). Por el contrario, la cepa O104:H4 contiene una variante del plásmido pAA que codifica para las fimbrias de tipo I (AAF/I) (Rasko y col., 2011).



**Figura 4.66.** Alineamiento realizado mediante herramientas de genómica comparativa del entorno génico existente aguas abajo del gen *aggR* presente en los plásmidos pAA de las cepas *E. coli* 042, *E. coli* C700-09, *E. coli* 55989 y *E. coli* O104:H4. De color blanco se representa el gen *aggR* correspondiente a cada una de las cepas analizadas, y de amarillo y negro se representan las diferentes secuencias de inserción localizadas aguas abajo del gen *aggR*. Alineamiento realizado mediante el programa *EasyFig* (Sullivan y col., 2011) haciendo uso de las secuencias génicas extraídas de la base de datos *GenBank* correspondientes a los plásmidos pAA de *E. coli* 042 (FN554767.1), *E. coli* C700-09 (LT719075), *E. coli* 55989 (CU928159) y *E. coli* O104:H4 (CP011332).

Tal y como se puede observar en la figura anterior, tanto la región promotora como la región codificante de los diferentes genes *aggR* de las diferentes cepas enteroagregativas de *E. coli* analizadas presentan una homología superior al 97%. Dicha homología se mantiene también en una serie de nucleótidos situados aguas abajo de la región codificante de cada gen *aggR*. Concretamente, se ha localizado una región común a todas las cepas de 44 nucleótidos de longitud, la cual hemos asociado como región 3'UTR del gen *aggR*. En el caso de la cepa *E. coli* C700-09, dicha región es unos 7 nucleótidos más larga que en las cepas 55989 y O104:H4, por lo que dicha cepa presenta una homología de 51 nucleótidos respecto la región 3'UTR del gen *aggR* de la cepa *E. coli* 042.

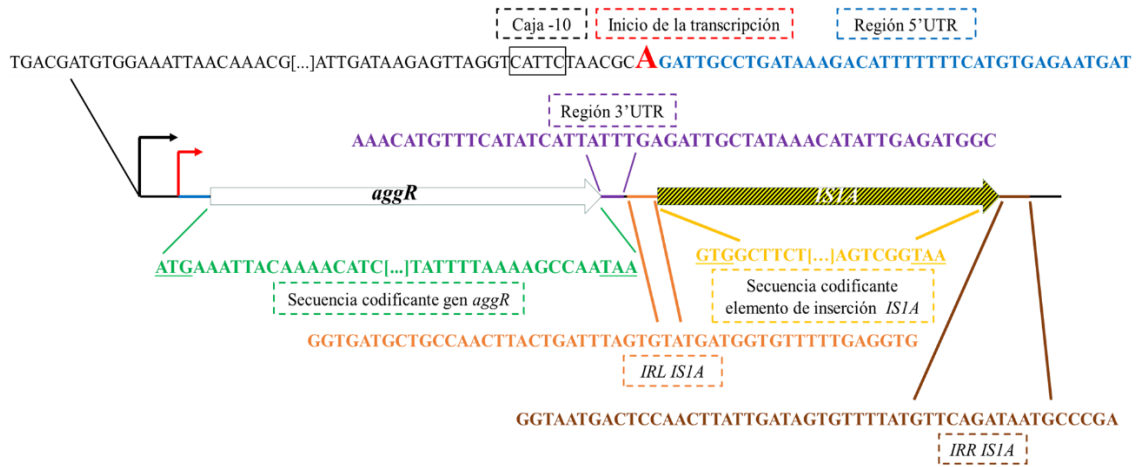
Además, en todas las cepas analizadas, se ha podido observar que el gen *aggR* se encuentra siempre flanqueado por secuencias de inserción (figura 4.66). Los elementos de inserción *IS* (*insertion sequence element*) son fragmentos genéticos de ADN móviles que, a diferencia de los transposones (que también contienen genes accesorios, como por ejemplo genes que portan elementos de resistencia a antibióticos), únicamente codifican para las enzimas responsables de su transposición (Berg y Howe, 1989; Craig y col., 2002). Concretamente, los *IS* suelen tener un tamaño de entre 0,7 Kb y 2,5 Kb y contienen como máximo dos pautas abiertas de lectura que ocupan la totalidad de la longitud del elemento de inserción, que se encuentra flanqueado en ambos extremos por secuencias repetidas invertidas (IR). Dichas pautas abiertas de lectura corresponden a la enzima catalítica (responsable de la ruptura del ADN para permitir el movimiento y transferencia del elemento de inserción) y a proteínas reguladoras de dicha catálisis.

Analizando cada una de las secuencias de inserción situadas aguas abajo del gen *aggR* mediante el programa *IS Finder* (una herramienta bioinformática que permite identificar y caracterizar elementos de inserción) (Siguier y col., 2006)), se ha podido comprobar cómo la secuencia del elemento de inserción situado aguas abajo del gen *aggR* en la cepa *E. coli* 042 no presenta homología con los elementos de inserción situados aguas abajo del gen *aggR* en las otras cepas estudiadas, los cuales si muestran una alta homología entre sí (figura 4.66).

En el caso específico del *IS* localizado aguas abajo del gen *aggR*, dicho elemento móvil se clasifica como *ISIA* y se encuentra situado 7 nucleótidos después del último nucleótido de la región 3'UTR del gen *aggR* (figura 4.67). Dicho elemento de inserción tiene un tamaño de 768 pb, y se encuentra flanqueado por repeticiones terminales invertidas (IRL

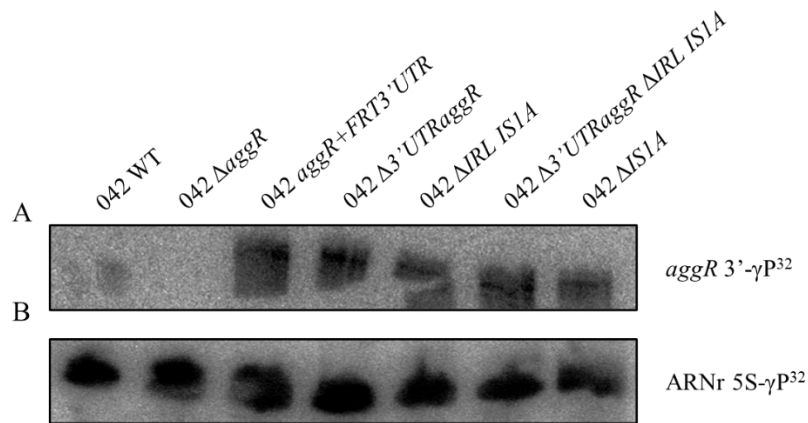
#### 4. RESULTADOS

e IRR) de 50 nucleótidos cada una, entre las que se encuentran dos pautas abiertas de lectura (la primera con un tamaño de 276 pb y la segunda con un tamaño de 378 pb).



**Figura 4.67.** Representación esquemática de los elementos genéticos que conforman el entorno del gen *aggR*. De color blanco se indica el gen *aggR*, y en amarillo y negro se indica el elemento de inserción *ISIA* situado aguas abajo del mismo. Se indica también la región promotora del gen (en negro), las regiones 5'UTR y 3'UTR (azul y violeta, respectivamente), el inicio de transcripción del gen *aggR* determinado experimentalmente (Morin y col., 2010), además de las repeticiones invertidas (izquierda y derecha) que flanquean al elemento de inserción *ISIA* (naranja y marrón, respectivamente). Anotación génica extraída de *GenBank* FN554767.1. Representación esquemática realizada mediante el programa *EasyFig* (Sullivan y col., 2011).

Con la finalidad de determinar qué región situada aguas abajo del gen *aggR* está implicada en la regulación de la expresión del propio gen *aggR*, realizamos cuatro delecciones: la primera, correspondiente a la secuencia 3'UTR del gen *aggR* (marcada en violeta en la figura 4.67). La segunda, correspondiente a la secuencia invertida y repetida izquierda (IRL) del elemento *ISIA* (marcada en naranja en la figura 4.67). La tercera delección combinaba tanto la eliminación de la región 3'UTR del gen *aggR* como la supresión del elemento IRL de *ISIA*. Finalmente, la última delección generada correspondió a toda la secuencia perteneciente al elemento *ISIA*, incluyendo sus repeticiones terminales invertidas (delección marcada en amarillo en la figura 4.67). Una vez generadas dichas delecciones, se analizó el nivel de expresión del ARNm del gen *aggR* en dichos fondos genéticos mediante *Northern Blot* (figura 4.68).



**Figura 4.68.** Análisis por *Northern Blot* de la expresión del ARN mensajero del gen *aggR* en las cepas *E. coli* 042 WT, *E. coli* 042  $\Delta$ aggR, *E. coli* 042 aggR+FRT3'UTR, *E. coli* 042  $\Delta$ 3'UTRaggR, *E. coli* 042  $\Delta$ IRL IS1A, *E. coli* 042  $\Delta$ 3'UTRaggR  $\Delta$ IRL IS1A y *E. coli* 042  $\Delta$ IS1A. El ARN total de las cepas fue aislado a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento bacteriano. El nivel de expresión del gen *aggR* se analizó a partir de 25  $\mu$ gr de ARN total (resuelto electroforéticamente) mediante la sonda *aggR* 3' marcada radioactivamente con P<sup>32</sup> en su extremo 5' (A). Como control de la homogeneidad de carga se analizó la expresión del ARNr 5S (B) en los diferentes mutantes estudiados (debido a la elevada expresión de este ARNr, únicamente se utilizaron 10  $\mu$ gr de ARN para incubar con la sonda específica).

Tal y como se puede observar en todos los casos, la delección de cualquier región situada aguas abajo del gen *aggR* provoca un incremento en los niveles del ARNm del propio gen, por lo que en dicha región génica existen elementos implicados en la regulación de la expresión de este regulador transcripcional. Todo ello pone de manifiesto que la inserción del elemento *IS1A*, seleccionada de manera positiva a lo largo de la historia evolutiva de esta bacteria, está implicada en mecanismos relacionados con la expresión y/o estabilidad del transcrito correspondiente al gen *aggR*.

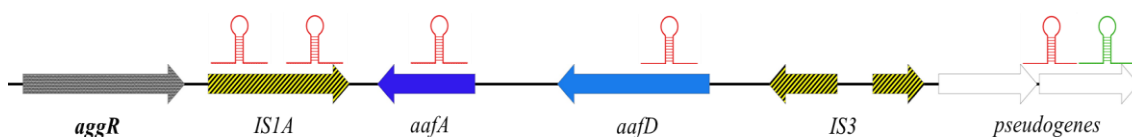
Además, el resultado indicado anteriormente en la figura 9.20 del Anexo sugiere la existencia de un transcrito largo que iría más allá del gen *aggR* donde, alteraciones en el mismo, como por ejemplo la adición de una secuencia FRT, son capaces de provocar un incremento en la estabilidad de ese ARNm. La existencia de dicho transcrito tan largo puede indicar bien la ausencia de un terminador transcripcional próximo al final del gen *aggR*, o bien que, a pesar de que existiera un terminador próximo al gen, el hecho de añadir una secuencia génica adicional a dicha región suponga una alteración o la inactivación del correspondiente terminador transcripcional.

Para tratar de confirmar un supuesto u otro, realizamos un análisis bioinformático de toda la región situada aguas abajo del gen *aggR* en busca de la presencia de terminadores Rho-



#### 4. RESULTADOS

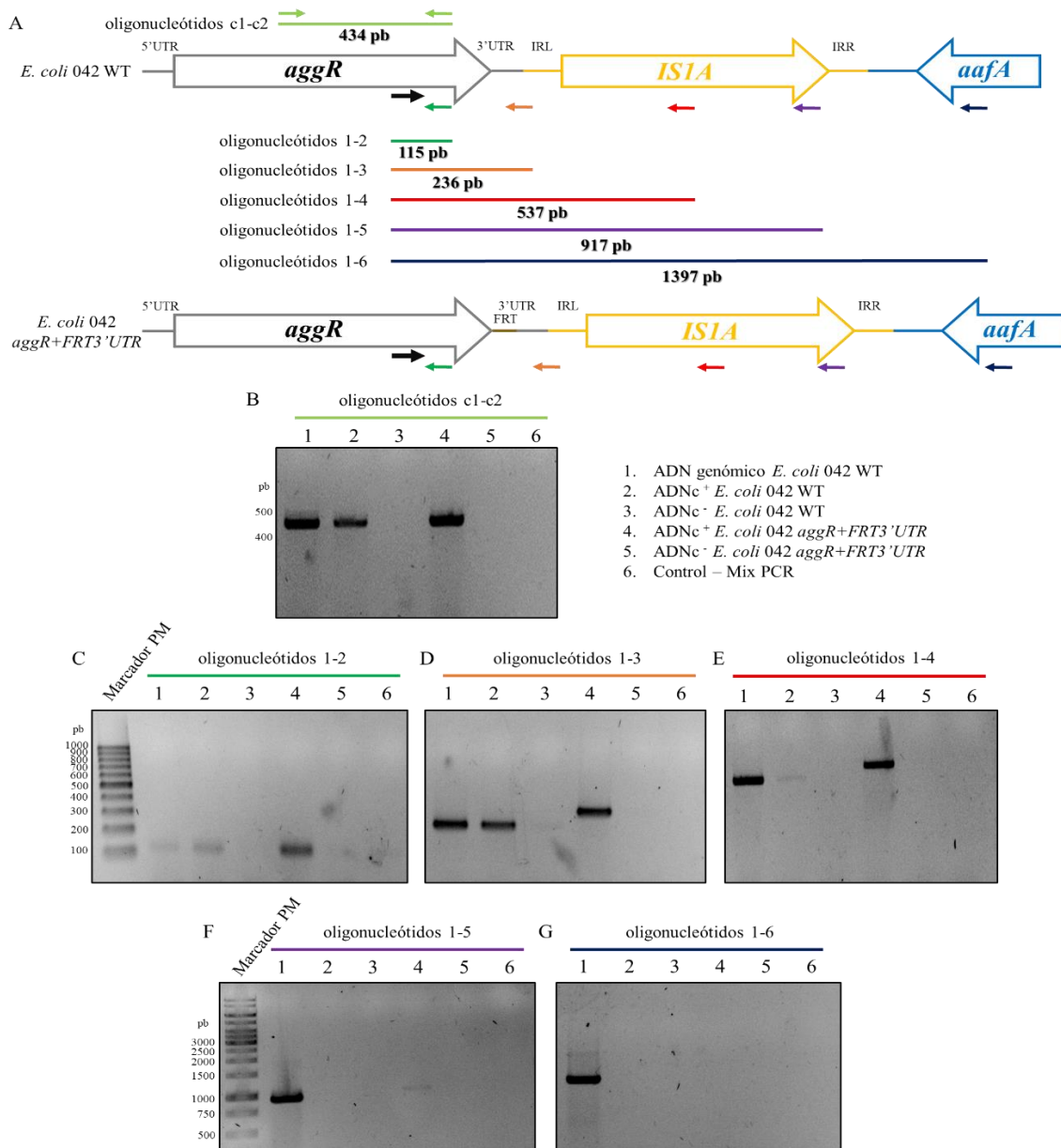
independientes. Analizando la secuencia desde el extremo 3'UTR del gen *aggR* mediante el *software* predictor *ARNold* (Naville y col., 2011), el primer terminador Rho-independiente identificado en la cadena codificante del gen *aggR* se encontró aguas abajo del gen *aafD* y del elemento de inserción *IS3* que le acompaña, unas 3,7 Kb después del codón de terminación de la traducción del gen *aggR* (figura 4.69). Cuando se analizó la presencia de terminadores Rho-dependientes mediante el *software* predictor *RhoTermPredict* (Di Salvo y col., 2019), se lograron identificar dos terminadores próximos al final del gen *aggR* (a 150 y a 733 nucleótidos de distancia del codón de terminación de la traducción del gen, respectivamente), situados dentro del elemento de inserción *IS1A* (figura 4.69).



**Figura 4.69.** Identificación mediante herramientas bioinformáticas de la presencia de terminadores transcripcionales en la región situada aguas abajo de la cadena codificante del gen *aggR*. La presencia de terminadores Rho-independientes (marcado en verde en la figura) se analizó mediante el *software* *ARNold* (Naville y col., 2011), mientras que la presencia de terminadores Rho-dependientes (marcados en rojo en la figura) se analizó mediante el *software* *RhoTermPredict* (Di Salvo y col., 2019). Representación esquemática realizada mediante el programa *EasyFig* (Sullivan y col., 2011) utilizando la anotación genética del plásmido pAA2 extraída de *GenBank* (FN554767.1).

Por todo ello, tanto los resultados obtenidos mediante *Northern Blot* como los obtenidos mediante análisis bioinformáticos indican que el transcrito de ARNm perteneciente al gen *aggR* va más allá de la región 3'UTR del mismo, llegando a extenderse hasta el elemento de inserción *IS1A*. Para tratar de confirmar dichos supuestos, realizamos ensayos de *walking-PCR* con la finalidad de comprobar en qué lugar acababa el transcrito del gen *aggR*. Para realizar dichos experimentos, se crecieron las cepas *E. coli* 042 WT y *E. coli* 042 *aggR*+*FRT*3'UTR hasta el inicio de la fase de crecimiento estacionaria en medio LB a 37°C, se aisló el ARN total de los cultivos y se retrotranscribió en su totalidad hasta ADNc. A partir de este ADNc, se diseñaron varias PCRs fijando un oligonucleótido *forward* en la región terminal de la secuencia codificante del gen *aggR*, mientras que el oligonucleótido *reverse* se fijó en diferentes posiciones situadas aguas abajo del gen *aggR*. Dicha metodología se realizó con la finalidad de identificar alguna combinación de oligonucleótidos a partir de la cual no se detectara amplificación por PCR, sinónimo de que el transcrito de *aggR* no prosigue más allá del punto fijado (figura 4.70).





**Figura 4.70.** Experimento de *walking*-PCR diseñado con la finalidad de analizar el tamaño del transcrito del gen *aggR*. (A) Representación esquemática del gen *aggR* y su entorno génico (elemento de inserción *ISIA* y gen *aafA*), donde se indica el lugar de hibridación de los oligonucleótidos utilizados en los análisis por PCR y el tamaño del amplicón esperado para cada pareja de *primers*. De color negro y con un tamaño mayor se indica el oligonucleótido fijado en el extremo 3' del gen *aggR*. (B) Control de integridad del ADNc obtenido a partir de la retrotranscripción del ARN total de las cepas *E. coli* 042 WT y *E. coli* 042 *aggR*+*FRT3'UTR*. Se ha analizado la calidad del mismo mediante PCR utilizando el ADNc como molde con oligonucleótidos que hibridan en la secuencia codificante del gen *aggR*. (C, D, E, F y G) Resultado del análisis por PCR de la detección del ADNc correspondiente al transcrito del gen *aggR* mediante oligonucleótidos que hibridan al final de la secuencia codificante del gen *aggR* (C), al principio de la región repetida e invertida izquierda del elemento de inserción *ISIA* (D), a mitad del elemento de inserción *ISIA* (E), al final del elemento de inserción *ISIA* (F) y dentro del gen *aafA* (G). Se puede apreciar el incremento del tamaño del amplicón en el caso de la cepa *E. coli* 042 *aggR*+*FRT3'UTR* respecto a la cepa salvaje debido a la inserción de los nucleótidos correspondientes a la cicatriz *FRT* después del codón *stop* del gen *aggR*.

#### 4. RESULTADOS

Se puede observar como en todos los casos analizados en este experimento (tanto en el control de calidad del ADNc como con las diferentes combinaciones de oligonucleótidos utilizados en el *walking*-PCR), los niveles de ADNc detectados referentes al transcrito del gen *aggR* siempre son superiores en el caso del mutante *aggR+FRT3'UTR* respecto a la cepa salvaje, en cuyo caso los niveles de transcrito detectados a través del ADNc son muy tenues (resultados análogos a los mostrados en los *Northern Blots* representados en las figuras 4.63, 4.64 y 4.68).

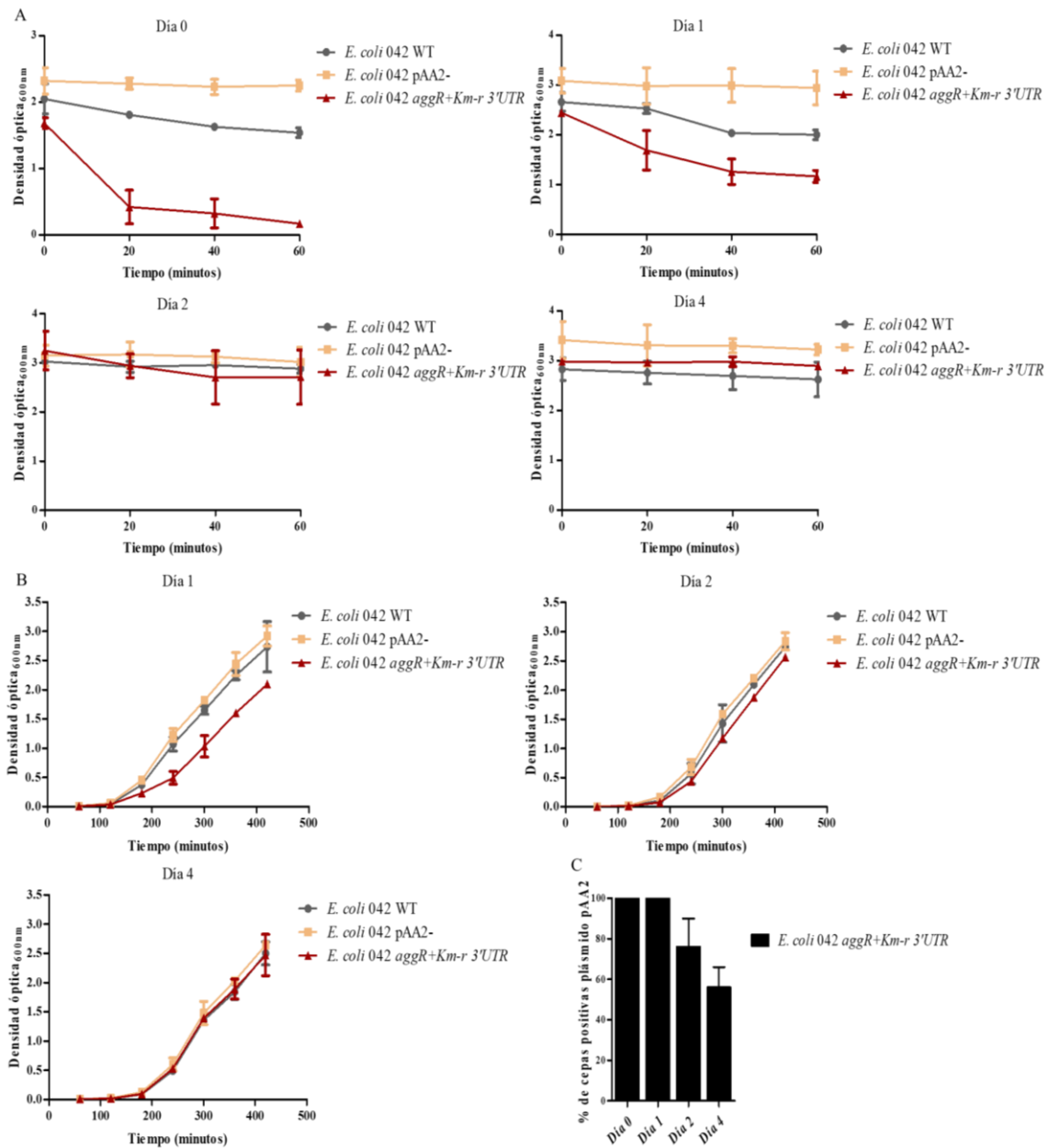
En referencia al tamaño del transcrito del gen *aggR*, puede apreciarse como el mismo se puede detectar con una relativa abundancia (en el caso de la cepa *E. coli* 042 *aggR+FRT3'UTR*) con oligonucleótidos que hibridan dentro de la región codificante del elemento de inserción *ISIA* (figura 4.70 (E)). Cuando se desplazan los oligonucleótidos aguas abajo de dicha región (hasta una posición más terminal del elemento de inserción, (figura 4.70 (F)), también se ha podido detectar el transcrito perteneciente al gen *aggR*, aunque a una cantidad neta mucho menor. Por el contrario, cuando se analiza la presencia del transcrito más allá del elemento de inserción (con oligonucleótidos situados en el gen *aafA*) no se ha podido observar banda alguna del ADNc correspondiente a dicho transcrito (figura 4.70 (G)). A través de esta aproximación experimental se ha podido comprobar como la transcripción del gen *aggR* abarca hasta mucho más allá de su secuencia codificante, logrando extenderse hasta la región terminal del elemento de inserción *ISIA* situado aguas abajo del mismo.

##### **4.3.2.6 Análisis de la estabilidad del fenotipo generado al alterar la región 3'UTR del gen *aggR***

Cabe destacar, sin embargo, que el fenotipo observado en la cepa que presenta la inserción de la secuencia FRT justo después del codón de terminación del gen *aggR* y antes de su secuencia 3'UTR es un fenotipo altamente inestable, cuyas manifestaciones se diluían totalmente con el paso de los días al realizar subcultivos seriados de las bacterias mutantes. Una posible explicación de la inestabilidad de este fenotipo puede ser debida a que el hecho de alterar las secuencias reguladoras del extremo 3' del gen *aggR* (incrementándose de esta manera los niveles de expresión del gen) se estén generando a nivel intracelular unos condicionantes tóxicos y deletéreos (como, por ejemplo, la expresión innecesaria de muchos genes no requeridos en las condiciones de cultivo en el laboratorio), lo cual se traduce en una pérdida de la eficiencia celular. Para evitar esta situación, podrían seleccionarse de manera positiva clones en los que se perdiera el

plásmido pAA2, el propio gen *aggR* en sí o únicamente la construcción generada aguas abajo del gen *aggR*, diluyéndose de esta manera el efecto de los diferentes fenotipos observados. Con el fin de comprobar estos supuestos, se realizaron ensayos de estabilidad plasmídica. Sin embargo, tal y como se explicaba en el apartado 4.3.1, teniendo en cuenta que el plásmido pAA2 no codifica para ningún determinante de resistencia a antibióticos que nos permitiera trazar qué bacterias contienen el plásmido y cuáles no, para poder realizar este experimento en el fondo genético mencionado anteriormente, se trabajó con la cepa original a la que se insertó el casete de resistencia a kanamicina justo después del codón *stop* del gen *aggR* (*E. coli* 042 *aggR+Km-r3'UTR*). Para poder analizar la estabilidad del fenotipo derivado de dicha modificación genética, se realizaron subcultivos seriados durante 4 días en medio LB a 37°C de los clones *E. coli* 042 WT, *E. coli* 042 pAA2- y *E. coli* 042 *aggR+Km-r3'UTR*, tomándose como inóculo una dilución 1:1000 del cultivo correspondiente del día anterior. Durante los días 1, 2 y 4 de subcultivos se analizó la cinética de crecimiento y el fenotipo autoagregativo de las diferentes cepas y, en el caso de la cepa *E. coli* 042 *aggR+Km-r3'UTR*, también se analizó la estabilidad del plásmido pAA2 a lo largo de los diferentes días de subcultivo (figura 4.71). Para ello, se utilizó la metodología descrita en el apartado 3.10.8 de Materiales y Métodos, y se calculó la estabilidad plasmídica en el cultivo en función del número de colonias positivas para la presencia del plásmido pAA2 respecto al número total de colonias analizadas.

#### 4. RESULTADOS

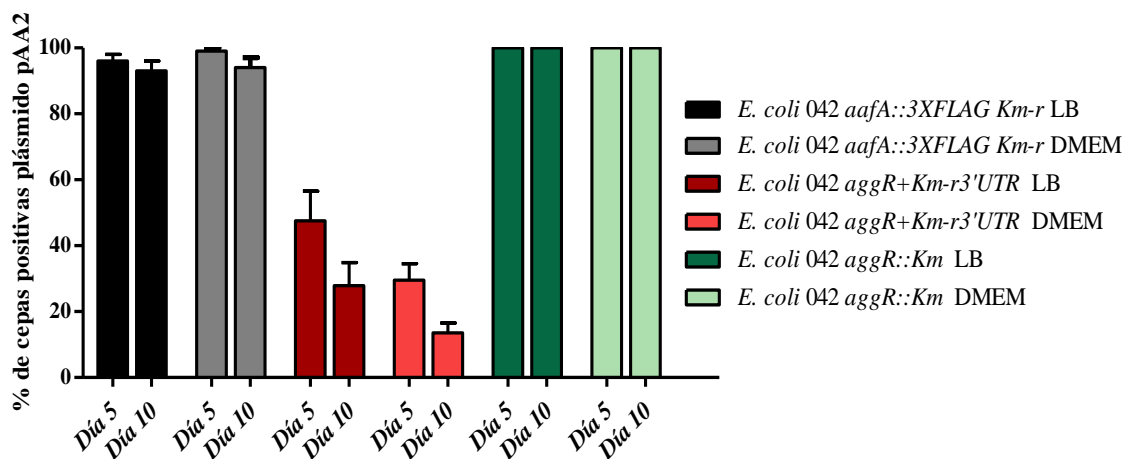


**Figura 4.71.** Análisis de la estabilidad del fenotipo derivado de la alteración de la región 3'UTR del gen *aggR*. (A) Perfil de autoagregación de los diferentes mutantes (*E. coli* 042 WT, *E. coli* 042 pAA2- y *E. coli* 042 *aggR*+*Km-r* 3'UTR) durante los días 0, 1, 2 y 4 de subcultivos seriados en medio LB, realizados a partir de una dilución 1:1000 del inóculo del día anterior. (B) Curvas de crecimiento de los diferentes mutantes estudiados a los días 1, 2 y 4 de subcultivos seriados. (C) Porcentaje de clones positivos para la presencia del plásmido pAA2 tras los diferentes días de subcultivo de la cepa *E. coli* 042 *aggR*+*Km-r* 3'UTR.

Tal y como se puede observar en la figura 4.71, el elevado fenotipo autoagregativo de la cepa *E. coli* 042 *aggR*+*Km-r* 3'UTR va disminuyendo a medida que transcurren los días de subcultivos seriados. De manera análoga, la menor tasa de crecimiento del mutante respecto a la cepa salvaje y la cepa pAA2- va revertiendo con el paso de los días hasta equipararse las cinéticas de crecimiento con las de las dos cepas mencionadas

anteriormente. En referencia a la estabilidad plasmídica, existe una correlación entre la disminución de la intensidad de los fenotipos del mutante y el cada vez menor porcentaje de bacterias que retuvieran el plásmido bajo estas condiciones, confirmándose de esta manera la inestabilidad generada al alterar el extremo 3' del gen *aggR*.

Para acabar de confirmar la selección positiva de aquellos clones que pierden el plásmido pAA2 cuando se incrementan los niveles de expresión del gen *aggR*, se quiso analizar la estabilidad del plásmido pAA2 en los dos supuestos opuestos: condiciones en las que se sobreexpresa el gen *aggR* (*E. coli* 042 *aggR+Km-r3'UTR*), y condiciones en las que no hay expresión del gen (cepa mutante *E. coli* 042 *aggR::Km*). Además, en este experimento también se incluyó la cepa *E. coli* 042 *aafA::3XFLAG Km-r*, cepa con un casete de resistencia a kanamicina insertado en el plásmido pAA2 (para seguir su trazabilidad) pero que contiene el gen *aggR* salvaje. Para analizar la estabilidad plasmídica en los fondos genéticos mencionados anteriormente, se realizaron subcultivos seriados de las cepas correspondientes durante 10 días seguidos en medio LB o medio DMEM suplementado con glucosa al 0,45% (p/v) a 37°C, y se analizó la presencia/ausencia del plásmido pAA2 entre las diferentes poblaciones a días 5 y 10 de subcultivo (figura 4.72).



**Figura 4.72.** Representación del porcentaje de bacterias positivas para la presencia del plásmido pAA2 después de 5 y 10 días de subcultivos seriados en medio LB y DMEM (suplementado con glucosa 0,45 % (p/v)) a 37°C de las cepas *E. coli* 042 *aafA::3XFLAG Km-r*, *E. coli* 042 *aggR+Km-r3'UTR* y *E. coli* 042 *aggR::Km*. Se representa el porcentaje global de la estabilidad plasmídica correspondiente a 3 réplicas biológicas independientes.

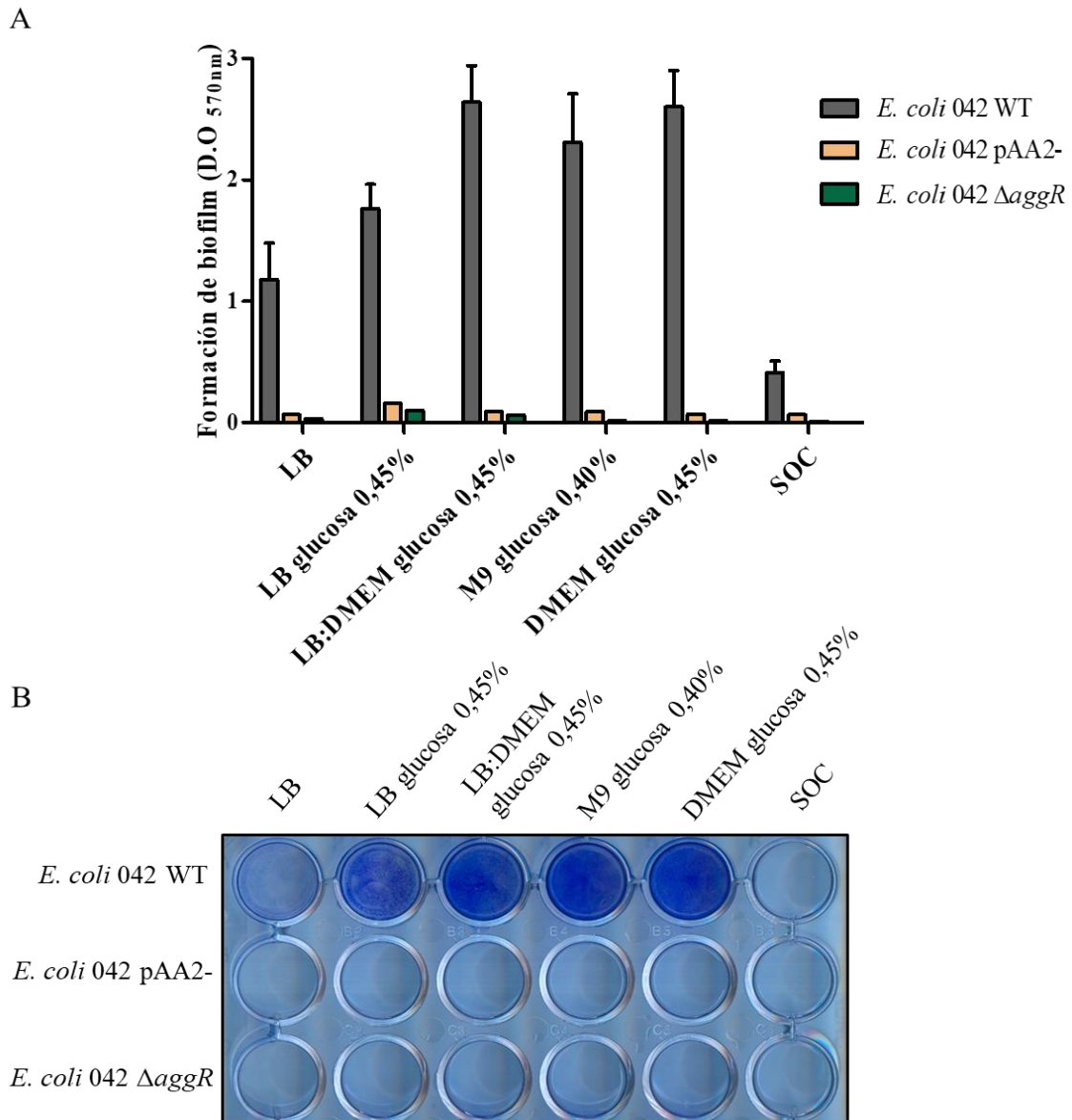
En referencia al fondo genético que muestra una expresión salvaje del gen *aggR* y la cepa mutante para dicho gen, la estabilidad del plásmido pAA2 es prácticamente total a lo largo

#### 4. RESULTADOS

de los diferentes días de subcultivos. Sin embargo, en cuanto a la cepa que presenta la sobreexpresión del gen *aggR*, la estabilidad del plásmido pAA2 disminuye de manera notoria a lo largo de los diferentes días de subcultivo (dicho descenso en la estabilidad plasmídica es mayor cuando las bacterias crecen en medio DMEM respecto a cuando lo hacen en medio rico LB). Por tanto, la inserción de secuencias de ADN en la región 3'UTR del gen *aggR* puede resultar una estrategia útil para obtener variantes de la cepa *E. coli* 042 curadas del plásmido pAA2.

##### **4.3.2.7 Efectos del plásmido pAA2 y del gen *aggR* en la formación de *biofilm* por parte de la cepa *E. coli* 042**

Un aspecto esencial en la virulencia de las cepas enteroagregativas de *E. coli* es la capacidad que presentan dichas cepas para adherirse a las células epiteliales intestinales y la posterior formación de *biofilms* bacterianos. En este sentido, se quiso analizar la importancia del plásmido pAA2 (a nivel global) y del regulador AggR (a nivel particular) en referencia a este proceso patogénico. En una primera prueba, se quiso analizar en diferentes medios de cultivo la formación de *biofilm* por parte de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2- y *E. coli* 042  $\Delta$ *aggR*. Los ensayos para estudiar la formación de *biofilm* se realizaron siguiendo el protocolo descrito en el apartado 3.10.9 de Materiales y Métodos sobre placas de poliestireno a las que se añadieron diferentes medios de cultivo (LB, LB suplementado con glucosa 0,45% (p/v), LB + DMEM suplementado con glucosa 0,45% (p/v) a una proporción 1:1, M9 suplementado con glucosa 0,40% (p/v), DMEM suplementado con glucosa 0,45% (p/v) y SOC) y se inocularon con las bacterias correspondientes. En referencia a la cepa salvaje, puede observarse en la figura 4.73 cómo las condiciones óptimas para la formación de *biofilm* por parte de la cepa *E. coli* 042 son medio de crecimiento mínimo (medio DMEM o medio M9 suplementados con glucosa, o medio DMEM y medio LB mezclados en una proporción 1:1), tal y como se describió por parte de (Sheikh y col., 2001). También se puede observar cómo, en ausencia del gen *aggR* (cepas *E. coli* 042 pAA2- y *E. coli* 042  $\Delta$ *aggR*), la capacidad de formación de *biofilm* por parte de la cepa *E. coli* 042 sobre una superficie abiótica es totalmente nula.



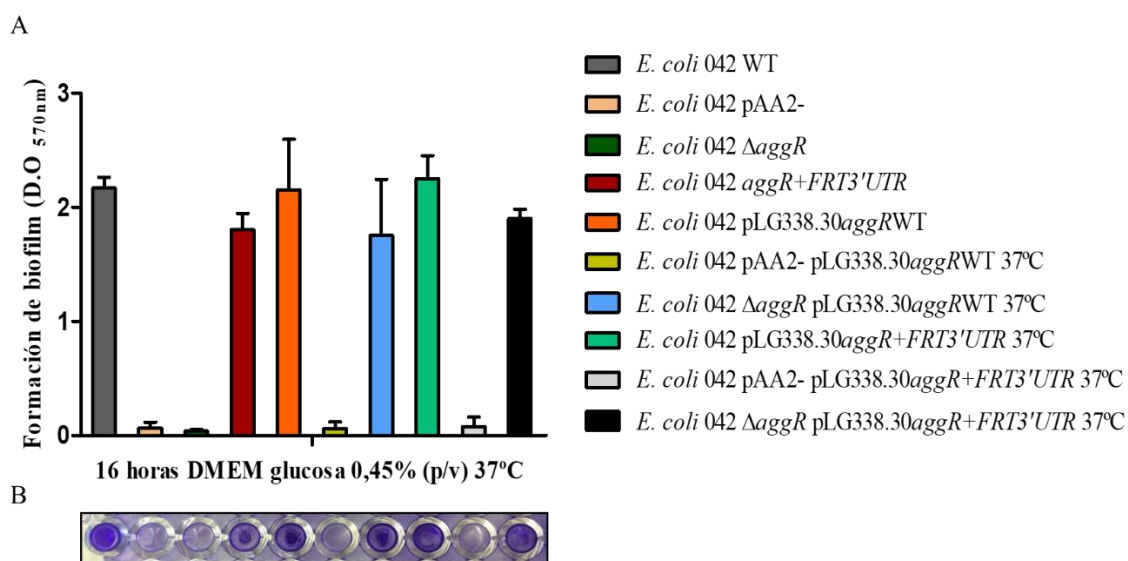
**Figura 4.73.** (A) Análisis de la formación de *biofilm* por parte de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2- y *E. coli* 042  $\Delta$ aggR en los medios LB, LB suplementado con glucosa 0,45% (p/v), LB + DMEM suplementado con glucosa 0,45% (p/v), M9 suplementado con glucosa 0,40% (p/v), DMEM suplementado con glucosa 0,45% (p/v) y SOC. Los ensayos de formación de *biofilm* se realizaron en placas de poliestireno de 96 pocillos. Se representan los valores correspondientes a 3 réplicas biológicas independientes con 6 réplicas técnicas cada una. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas. (B) Imagen representativa del resultado de la tinción con cristal violeta 0,1% (p/v) correspondiente a la formación de *biofilm* por parte de las diferentes cepas estudiadas. Los experimentos cualitativos se realizaron en placas de poliestireno de 24 pocillos.

También se quiso analizar la capacidad de formación de *biofilm* por parte de las bacterias que presentaban la alteración en el extremo 3' del gen *aggR*. Para ello, se analizó la capacidad de formación de *biofilm* en cepas que presentaban tanto la alteración a nivel cromosómico como aquellas que presentaban la construcción clonada en el plásmido



#### 4. RESULTADOS

pLG338.30. Paralelamente, también se quiso analizar si las bacterias curadas del plásmido pAA2 eran capaces de realizar *biofilm* al complementarlas en *trans* con un plásmido que les aportara el gen *aggR* salvaje (pLG338.30*aggR*WT) que dichas bacterias habían perdido junto con el plásmido pAA2. De esta manera, se quiso determinar si la incapacidad de formación de *biofilm* por parte de la cepa curada del plásmido pAA2 era debido a la falta del gen *aggR* o debido a que únicamente las fimbrias AAF/II codificadas a nivel plasmídico eran responsables de la formación de *biofilm* por parte de *E. coli* 042, y no las fimbrias codificadas en el cromosoma de esta cepa (figura 4.74).



**Figura 4.74.** Análisis de la formación de *biofilm* en medio DMEM suplementado con glucosa 0,45% (p/v) por parte de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ *aggR* y *E. coli* 042 *aggR*+*FRT3'*UTR. Dichas cepas fueron complementadas con el plásmido pLG338.30*aggR*WT y con el plásmido pLG338.30*aggR*+*FRT3'*UTR. Los ensayos de formación de *biofilm* se realizaron en placas de poliestireno de 96 pocillos (representados en la imagen (B)), cuantificándose los mismos mediante métodos espectrofotométricos (A) después de solubilizar las bacterias adheridas a la placa. Se representan los valores correspondientes a 3 réplicas biológicas independientes con 6 réplicas técnicas cada una. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Analizando los resultados obtenidos, puede observarse cómo la alteración del extremo 3' del gen *aggR* provoca una ligera disminución en la capacidad de formación de *biofilm* por parte de la cepa *E. coli* 042. Ello puede ser debido al elevado grado de autoagregación y a la formación de acúmulos bacterianos (y la dificultad de solubilizarlos) por parte de esta cepa. La complementación de la mutación cromosómica del gen *aggR* con el gen *aggR* clonado en el plásmido pLG338.30 (tanto con su extremo 3' salvaje como con su extremo 3' alterado por la inserción de la secuencia FRT) funciona correctamente, ya que

la cepa mutante complementada en *trans* es capaz de realizar *biofilm* bacteriano a unos niveles similares a los de la cepa salvaje. Por el contrario, a pesar de complementar la cepa pAA2- en *trans* con los mismos plásmidos funcionales portadores del gen *aggR* con los que se complementó a la cepa  $\Delta$ *aggR*, esta cepa no es capaz de formar *biofilm*. Todo ello indica que son las fimbrias codificadas a nivel plasmídico (y no las fimbrias codificadas en el cromosoma de la cepa) las responsables de la formación de *biofilm* en la cepa *E. coli* 042.

#### 4.3.2.8 Efecto de la inserción nucleotídica en el extremo 3'UTR del gen *aggR* sobre el transcriptoma global de la cepa *E. coli* 042

Teniendo en cuenta que la modificación del extremo 3' del gen *aggR* provoca un aumento en su expresión, lo que implica la sobreexpresión de genes tales como la dispersina *aap* y las fimbrias *aafA*, decidimos analizar cómo afectaba la sobreexpresión del gen *aggR* a nivel global sobre el transcriptoma de la cepa *E. coli* 042. Para ello, se realizó un análisis transcriptómico comparativo por RNA-seq de la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR* respecto a la cepa *E. coli* 042 WT. Estos ensayos transcriptómicos se realizaron a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento, donde el aislamiento del ARN y el procesamiento de las muestras para obtener los datos relativos a la expresión de cada gen se realizaron en colaboración con la empresa *Vertis Biotechnologie AG* (*Freising-Weihenstephan*, Alemania), siguiendo el protocolo descrito en el apartado 3.8.5 de Materiales y Métodos. En la tabla 4.11 se representan los genes mayormente sobreexpresados bajo el fondo genético *aggR*+*FRT3*'*UTR* respecto al fondo genético salvaje.

<i>E. coli</i> 042 <i>aggR</i> + <i>FRT3</i> ' <i>UTR</i> vs <i>E. coli</i> 042 WT	
Identificador NCBI ( <i>locus tag</i> )	Fold change
EC042_RS23795 (hypothetical protein)	11,3
EC042_RS16305 (hypothetical protein)	9,4
EC042_RS17065 (CPBP family intramembrane metalloprotease)	9,4
EC042_RS21655 (hypothetical protein)	7,5
EC042_RS16275 (hypothetical protein)	6,6
EC042_pAA105 (pseudogene)	6,3
EC042_RS16360 (invasion protein)	5,7
EC042_RS16280 (hypothetical protein)	5,7
EC042_RS09160 (cell division inhibition protein DicB)	5,0
EC042_RS12795 (hypothetical protein)	5,0
EC042_RS05725 (trifunctional transcriptional regulator/proline dehydrogenase/L-glutamate gamma-semialdehyde dehydrogenase)	4,9

#### 4. RESULTADOS

EC042_RS09020 (hypothetical protein)	4,7
EC042_RS29440 (hypothetical protein)	4,7
EC042_RS12450 (PTS galactitol transporter subunit IIB)	4,5
EC042_RS07860 (phage lysis protein EssD)	4,4
EC042_RS10140 (N-succinylarginine dihydrolase AstB)	4,3
EC042_RS10145 (succinylglutamate-semialdehyde dehydrogenase AstD)	4,3
EC042_RS10135 (succinylglutamate desuccinylase AstE)	4,2
EC042_RS29405 (hypothetical protein)	4,1
EC042_RS22490 (fatty acid oxidation complex subunit alpha FadB)	4,1
EC042_RS02130 (hydrogen peroxide resistance inhibitor IprA)	4,1
EC042_RS22485 (acetyl-CoA C-acyltransferase FadA)	3,9
EC042_RS05770 (phosphate starvation-inducible protein PhoH)	3,9
EC042_RS03040 (hypothetical protein)	3,8
EC042_RS15365 (glutaredoxin-like protein NrdH)	3,8
EC042_RS09860 (MFS transporter)	3,8
EC042_RS09040 (YdfR family protein)	3,8
EC042_RS03755 (K(+)-transporting ATPase subunit F)	3,8
EC042_RS05765 (hypothetical protein)	3,7
EC042_RS01360 (lateral flagellar basal body component protein lfiE)	3,7
EC042_RS19645 (type 1 fimbrial protein)	3,6
EC042_RS26280 (dispersin export ABC transporter ATP-binding protein AatC)	3,6
EC042_RS03060 (hypothetical protein)	3,6
EC042_RS10155 (succinylornithine/acetylornithine transaminase AstC)	3,6
EC042_RS10150 (arginine N-succinyltransferase AstA)	3,5
EC042_RS20725 (hypothetical protein)	3,5
EC042_RS16260 (helix-turn-helix transcriptional regulator)	3,5
EC042_RS01925 (hypothetical protein)	3,5
EC042_RS17395 (malate synthase G glcB)	3,5
EC042_RS09850 (YdiL family protein)	3,3
EC042_RS29570 (hypothetical protein)	3,3
EC042_RS23860 (phosphonate C-P lyase system protein PhnG)	3,3
EC042_RS01355 (lateral flagellar RpoN-interacting regulatory protein)	3,3
EC042_RS21385 (hypothetical protein)	3,2
EC042_RS26265 (dispersin export ABC transporter permease subunit AatP)	3,2
EC042_RS26275 (dispersin export-associated protein AatB)	3,2
EC042_RS15285 (carbon starvation induced protein CsiD)	3,2
EC042_RS23640 (hypothetical protein)	3,2
EC042_RS12535 (fimbria/pilus periplasmic chaperone)	3,2
EC042_RS11835 (inverse autotransporter)	3,2
EC042_RS15295 (NADP-dependent succinate-semialdehyde dehydrogenase GabD)	3,2
EC042_RS04950 (putrescine/spermidine ABC transporter substrate-binding protein)	3,2
EC042_RS11725 (DUF1391 domain-containing protein)	3,1

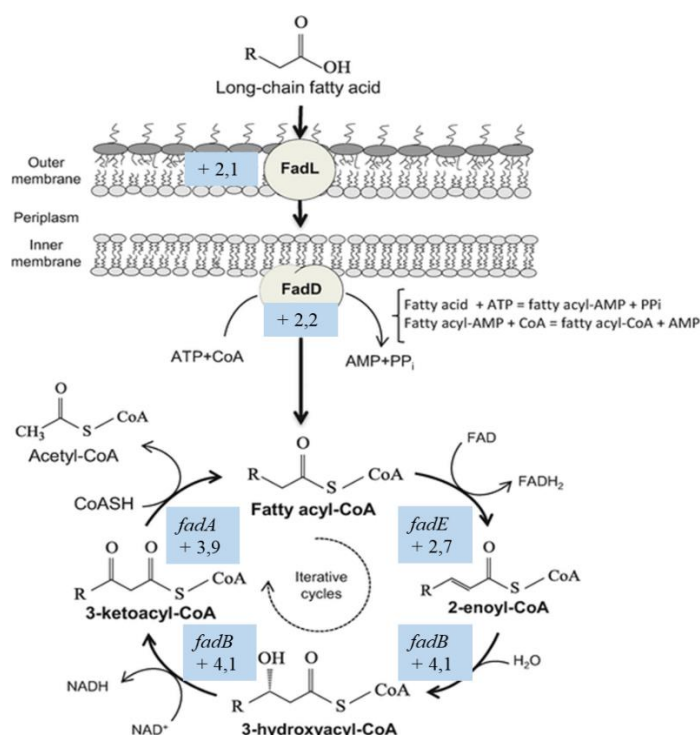
EC042_RS03045 (hypothetical protein)	3,1
EC042_RS20080 (hypothetical protein)	3,1
EC042_RS17070 (hypothetical protein)	3,1
EC042_RS07545 (FAD-binding oxidoreductase)	3,1
EC042_RS02040 (2-keto-4-pentenoate hydratase)	3,1
EC042_RS08465 (leucine-rich repeat domain-containing protein)	3,1
EC042_RS08120 (phosphatidate cytidylyltransferase)	3,1
EC042_RS25730 (N-acetylneuraminic acid outer membrane channel NanC)	3,1
EC042_RS21315 (HilA family transcriptional regulator EilA)	3,1
EC042_RS26285 (dispersin export-associated protein AatD)	3,0
EC042_RS13300 (anaerobic glycerol-3-phosphate dehydrogenase subunit A)	3,0
EC042_RS23955 (alpha-galactosidase)	3,0
EC042_RS17400 (protein GlcG)	3,0
EC042_RS26510/EC042_pAA060 (AggR-activated transcriptional regulator Aar)	2,9
EC042_RS01520 (flagellar export chaperone FliS)	2,9
EC042_RS26270 (dispersin export ABC transporter outer membrane protein aatA)	2,8
EC042_RS05935 (biofilm formation regulator BssS)	2,8
EC042_RS01290 (acyl-CoA dehydrogenase FadE)	2,7
EC042_RS23395 (isocitrate lyase aceA)	2,4
EC042_RS10460 (long-chain-fatty-acid--CoA ligase FadD)	2,2
EC042_RS13815 (long-chain fatty acid transporter FadL)	2,1

**Tabla 4.11.** Listado de los genes mayormente sobreexpresados en el cromosoma de la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR*, expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. De color amarillo se encuentran remarcados los genes implicados en la metabolización de la arginina a través de la vía *ast*; en azul se indican los genes responsables de la metabolización de ácidos grasos de cadena larga a través de la vía *fad*; en verde se indican los genes responsables de catalizar la vía del glioxilato. De color gris se indican los genes flagelares sobreexpresados bajo este fondo genético. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

Tal y como se puede observar en la tabla 4.11, existe una gran variedad funcional de genes sobreexpresados cuando se altera el extremo 3' del gen *aggR*. Entre estos genes se encuentran una gran variedad de genes fágicos, un gran número de proteínas hipotéticas y diferentes genes que codifican para una gran variedad de enzimas metabólicas, entre las que destacan las enzimas pertenecientes a la vía *ast* (implicada en el metabolismo de la arginina), la vía *fad* (implicada en el metabolismo de ácidos grasos de cadena larga) y los genes *glcB* (gen que codifica para la isoforma G de la malato sintasa) y *aceA* (gen que codifica para la isocitrato liasa), ambas enzimas implicadas en la vía del glioxilato.

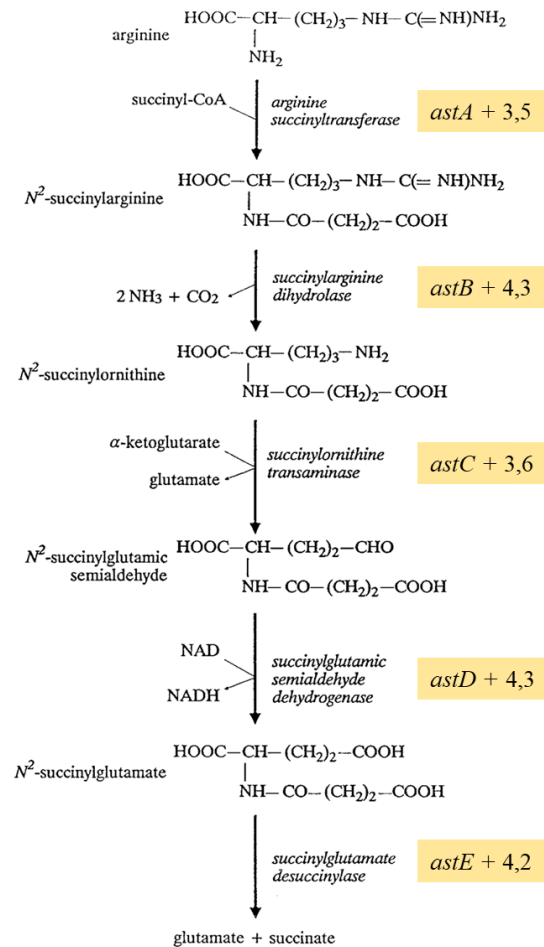
#### 4. RESULTADOS

Los genes de la vía *fad* que muestran un mayor nivel de sobreexpresión son los genes *fadA* y *fadB* (sobreexpresados 4 veces más en un fondo genético *aggR+FRT3'UTR* respecto al nivel de expresión de dichos genes en un fondo genético salvaje). Dichos genes forman parte de un mismo operón (*fadAB*) y presentan actividad multienzimática implicada en el transporte, acilación y en la  $\beta$ -oxidación de los ácidos grasos (figura 4.75). Los otros genes de la vía *fad* sobreexpresados bajo este fondo genético son *fadE*, *fadD* y *fadL* (dichos genes presentan un nivel de sobreexpresión respecto a la cepa salvaje de 2,7, 2,2 y 2,1 veces, respectivamente).



**Figura 4.75.** Representación esquemática de la vía *fad* implicada en el metabolismo de los ácidos grasos de cadena larga. De color azul claro se indican los diferentes genes *fad* responsables de la catálisis de las diferentes reacciones metabólicas. También se indica el valor de sobreexpresión de dichos genes en la cepa *E. coli* 042 *aggR+FRT3'UTR* respecto la cepa *E. coli* 042 WT, valores obtenidos mediante RNA-seq. Imagen adaptada de (Jiménez-Díaz y col., 2017).

La vía AST (*arginine succinyltransferase*), codificada en el operón *astABCDE*, es la vía metabólica responsable de la degradación de la arginina hasta glutamato, succinato, CO<sub>2</sub> y amonio. Dicha vía metabólica permite el crecimiento de *E. coli* utilizando la arginina como única fuente de nitrógeno. Los genes que codifican para las diferentes enzimas de esta vía metabólica (*astA*, *astB*, *astC*, *astD* y *astE*) presentan unos niveles de expresión de entre 3 y 4 veces superiores en un fondo genético *aggR+FRT3'UTR* respecto al nivel de expresión de dichos genes en la cepa salvaje (figura 4.76).

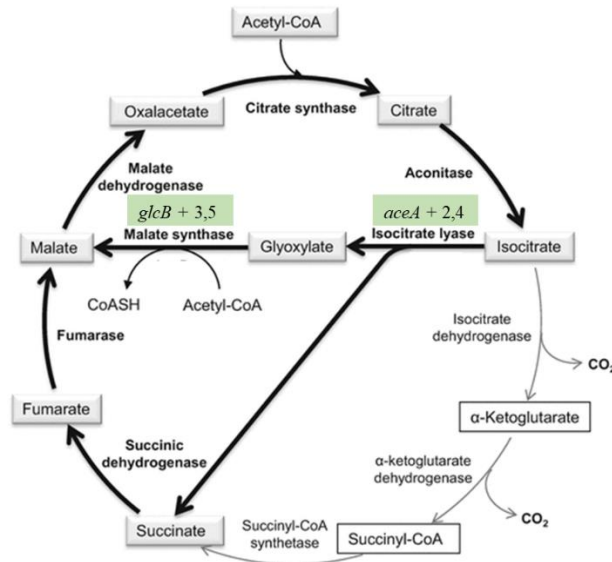


**Figura 4.76.** Representación esquemática de las enzimas y los metabolitos implicados en la vía AST (*arginine succinyltransferase*) responsable del catabolismo de la arginina. Se indican las diferentes reacciones implicadas en la metabolización de dicho aminoácido hasta glutamato y succinato. De color amarillo claro se indican los diferentes genes *ast* responsables de la catálisis de las diferentes reacciones metabólicas. También se indica el valor de sobreexpresión de dichos genes en la cepa *E. coli* 042 *aggR+FRT3'UTR* respecto la cepa *E. coli* 042 WT, valores obtenidos mediante RNA-seq. Imagen adaptada de (Schneider y col., 1998).

En referencia a los genes *glcB* y *aceA* (cuyo producto génico son las enzimas responsables de catalizar la vía del glioxilato), estos genes presentan unos niveles de sobreexpresión del orden de 3,5 y 2,4 veces superiores en la cepa *E. coli* 042 *aggR+FRT3'UTR* respecto a la cepa salvaje (figura 4.77). Dicho ciclo, presente en la gran mayoría de microorganismos y células vegetales (pero no en células animales), es una variante del ciclo del ácido cítrico que permite realizar la síntesis de glucosa a partir de ácidos grasos.

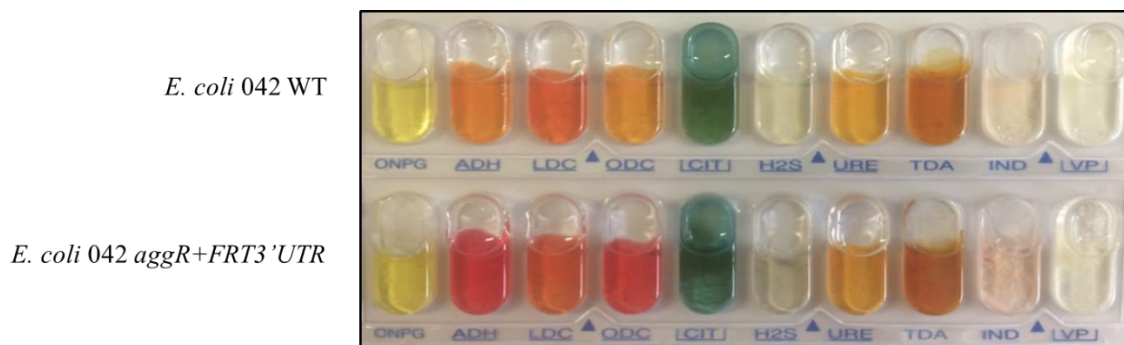


#### 4. RESULTADOS



**Figura 4.77.** Representación esquemática del ciclo del glioxilato implicado en la síntesis de azúcares a partir de ácidos grasos. De color verde claro se indican los genes *glcB* y *aceA* responsables de la catálisis de las diferentes reacciones metabólicas de dicho ciclo. También se indica el valor de sobreexpresión de dichos genes en la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR* respecto a la cepa *E. coli* 042 WT, valores obtenidos mediante RNA-seq. Imagen adaptada de (Jiménez-Díaz y col., 2017).

Con el fin de confirmar las alteraciones metabólicas generadas en la cepa *E. coli* 042 cuando se sobreexpresa del gen *aggR*, se inoculó una tira API 20E™ con las cepas *E. coli* 042 WT y *E. coli* 042 *aggR*+*FRT3*'*UTR* (figura 4.78). La tira API, a través de una serie de pruebas bioquímicas estandarizadas, nos indica varios aspectos del metabolismo de una determinada bacteria.



**Figura 4.78.** Resultado de la inoculación de la tira API 20E™ con las cepas *E. coli* 042 WT y *E. coli* 042 *aggR*+*FRT3*'*UTR*. Las pruebas realizadas determinan las siguientes actividades enzimáticas: ONPG (beta-galactosidasa), ADH (arginina deshidrolasa), LDC (lisina descarboxilasa), ODC (ornitina descarboxilasa), CIT (utilización del citrato), H<sub>2</sub>S (producción de H<sub>2</sub>S), URE (ureasa), TDA (triptófano desaminasa), IND (producción de indol), VP (producción de acetoína (Voges-Proskauer)). La inoculación de la tira API se realizó siguiendo las instrucciones indicadas por el fabricante (*bioMérieux*, Francia).



Puede observarse cómo, de la batería de pruebas bioquímicas realizadas en la tira API, existen reacciones en las que se observan diferencias entre la cepa salvaje y la cepa que presenta la interrupción del extremo 3' del gen *aggR*. Específicamente, las pruebas ADH (medio que contiene L-arginina y mide la actividad de hidrólisis de este aminoácido) y la prueba ODC (medio que contiene L-ornitina y mide la actividad ornitina descarboxilasa), confirman la sobreexpresión de las vías enzimáticas responsables del metabolismo de la arginina.

Además, de manera análoga al caso de la cepa pAA2-, en la cepa *E. coli* 042 *aggR+FRT3'UTR* también se ha encontrado una importante desregulación en la expresión de genes implicados en la síntesis flagelar (concretamente se observó una gran desregulación en la expresión de los genes *lfiE* y *lafK*). Sin embargo, dichos genes no forman parte de los operones *fli* y *flg*, responsables de la síntesis del sistema flagelar Flag-1. En 2005, se describió que la cepa *E. coli* 042 también codifica en su genoma un *cluster* génico adicional responsable de la síntesis de un potencial segundo sistema flagelar: el sistema Flag-2, responsable de la expresión de un sistema de flagelación lateral (Ren y col., 2005). Los genes flagelares sobreexpresados en el fondo genético *aggR+FRT3'UTR* pertenecen a este segundo sistema flagelar. Existen algunas especies de la familia gammaproteobacteria (entre las que destacan *Aeromonas hydrophila* y *Vibrio parahaemolyticus*) que utilizan de manera dual los dos sistemas flagelares: mientras que el sistema flagelar polar es importante para la motilidad de las bacterias en medio líquido, el sistema flagelar lateral es esencial en la motilidad sobre superficies sólidas (Kirov y col., 2002; MacCarter, 2004). En *E. coli*, el sistema flagelar lateral Flag-2 se encuentra codificado de manera restringida únicamente en unas pocas cepas. Dicho sistema presenta, además, una alta homología con el sistema flagelar lateral presente en especies de los géneros *Citrobacter*, *Yersinia*, *Aeromonas* y *Vibrio* (Ren y col., 2005).

En referencia a los genes que forman parte del regulón del activador transcripcional AggR identificados mediante la tecnología de *microarrays* (Morin y col., 2013) y RNA-seq (Yasir y col., 2019) (genes reprimidos en un mutante *aggR*), todos ellos se encuentran sobreexpresados en la cepa que presenta la alteración en el extremo 3' del gen *aggR* (tabla 4.12). Todos estos datos corroboran todavía más el hecho de que en el extremo 3' del gen *aggR* existen secuencias importantes a nivel de la regulación de la expresión del gen, y que la interrupción de las mismas supone un incremento en la expresión de dicho gen y por ende de los productos génicos activados por parte de este regulador transcripcional.

#### 4. RESULTADOS

Regulación <i>aggR</i> (Morin y col., 2013; Yasir y col., 2019)	
<i>E. coli</i> 042 <i>aggR</i> + <i>FRT3'UTR</i> vs <i>E. coli</i> 042 WT	
Identificador NCBI ( <i>locus tag</i> )	Fold change
EC042_pAA105 (pseudo gene)	6,3
EC042_RS26280/EC042_pAA010 (dispersin export ABC transporter ATP-binding protein AatC)	3,6
EC042_RS26265/EC042_pAA007 (dispersin export ABC transporter permease subunit AatP)	3,2
EC042_RS26275/EC042_pAA009 (dispersin export-associated protein AatB)	3,2
EC042_RS11835/EC042_2219 (inverse autotransporter)	3,2
EC042_RS26510/EC042_pAA060 (AggR-activated transcriptional regulator Aar)	2,9
EC042_RS26270/EC042_pAA008 (dispersin export ABC transporter outer membrane protein aatA)	2,8
EC042_RS05935/EC042_1127 (biofilm formation regulator BssS)	2,8
EC042_RS23675/EC042_4430 (c-di-GMP phosphodiesterase PdeC)	2,8
EC042_RS27190/EC042_pAA022 (putative hexosyltransferase CapU)	2,8
EC042_RS26460/EC042_pAA046 (aggregative adherence fimbria II chaperone AafD)	2,4
EC042_RS26345/EC042_pAA023 (virulence factor VirK)	2,3
EC042_RS26465/EC042_pAA048 (aggregative adherence fimbria II major subunit aafA)	2,3
EC042_RS26240/EC042_pAA003 (class I isoprenoid biosynthesis enzyme)	2,3
EC042_RS26245/EC042_pAA004 (isopentenyl-diphosphate Delta-isomerase)	2,3
EC042_RS26325 (hypothetical protein pseudo)	2,1
EC042_RS30530/EC042_pAA055 (dispersin aap)	2,0
EC042_RS26515 (IS630 family transposase)	2,0
EC042_pAA047 (hypothetical protein)	2,0
EC042_RS26335/EC042_pAA021 (polysaccharide deacetylase family protein)	2,0
EC042_RS26330/EC042_4581 (hypothetical protein)	1,8
EC042_RS17030/EC042_3187 (AAA family ATPase)	1,8
EC042_4581 (hypothetical protein)	1,8
EC042_RS24360/EC042_4562 (type VI secretion system contractile sheath small subunit tssB)	1,8
EC042_RS24365/EC042_4563 (type VI secretion system contractile sheath large subunit tssC)	1,8
EC042_RS24370/EC042_4564 (type VI secretion system protein AaiC/Hcp2)	1,7
EC042_RS24375/EC042_4565 (putative type VI secretion protein)	1,7
EC042_RS16990 (hypothetical protein)	1,6
EC042_RS26470/EC042_pAA050 (IS1-like element IS1A family transposase)	1,6
EC042_RS30500/EC042_pAA006 (pseudogene - transposase)	1,6
EC042_RS30525/EC042_pAA053 (IS3 family transposase)	1,6
<b>EC042_RS26480/EC042_pAA052 (aggregative adherence transcriptional regulator AggR)</b>	<b>1,6</b>
EC042_RS30410/EC042_4581A (transposase)	1,5
EC042_RS17010/EC042_3184 (hypothetical protein)	1,5
EC042_RS24435/EC042_4577 (type VI secretion system ATPase TssH)	1,5

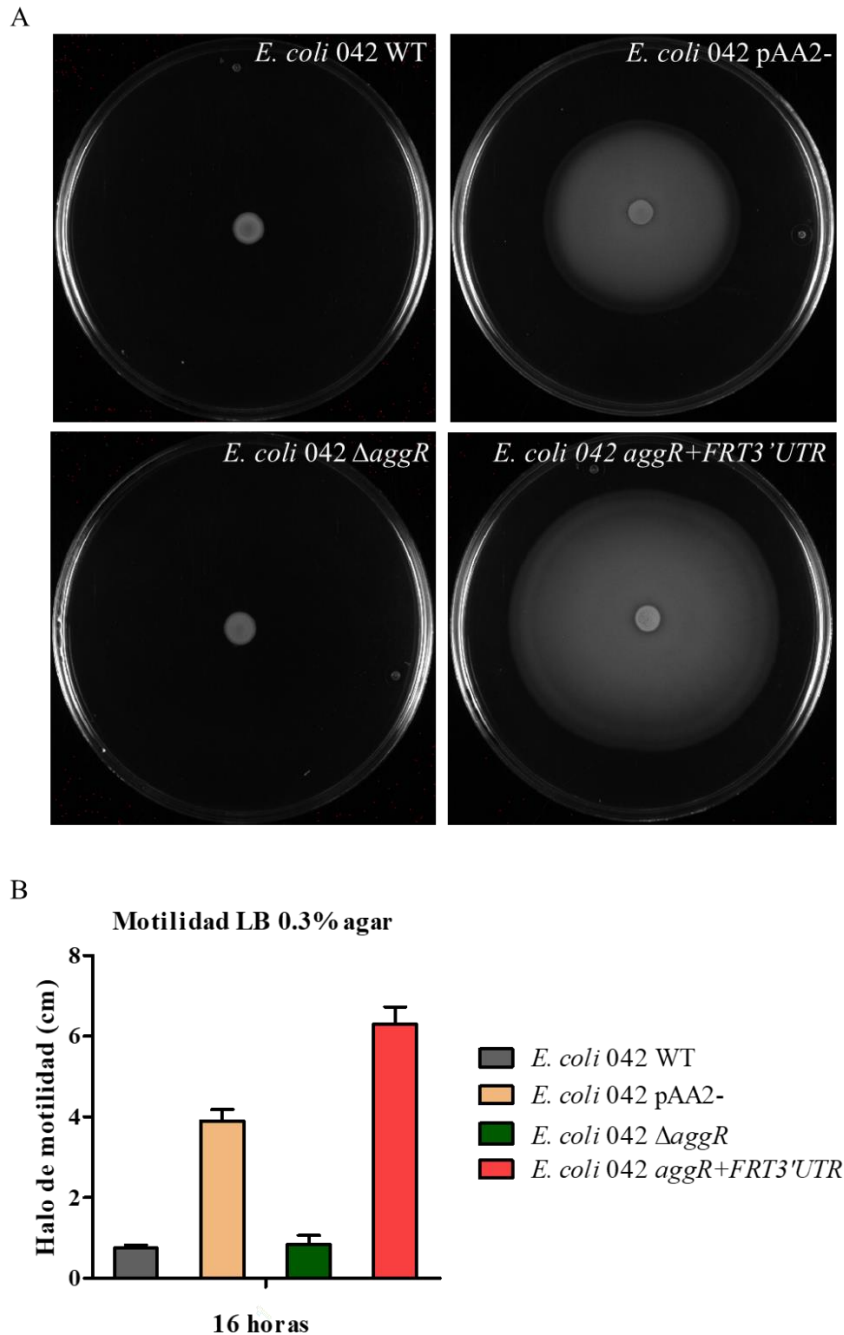
EC042_RS16995/EC042_3181 (PerC family transcriptional regulator)	1,5
EC042_RS24385/EC042_4568 (type VI secretion system baseplate subunit TssG)	1,5
EC042_RS24390/EC042_4569 (type VI secretion system tip protein VgrG TssI)	1,5
EC042_RS17015 (pseudogene)	1,5
EC042_RS12055/EC042_2249 (type IV toxin-antitoxin system toxin CbtA)	1,5
EC042_RS24430/EC042_4576 (putative type VI secretion protein)	1,5
EC042_4581B (hypothetical protein)	1,4
EC042_pAA033 (transposase)	1,4
EC042_RS24490/EC042_4587 (IS110-like element ISEc45 family transposase)	1,4
EC042_RS24460/EC042_4582 (hypothetical protein)	1,4
EC042_RS24380/EC042_4566 (type VI secretion system baseplate subunit TssF)	1,4
EC042_RS17005/EC042_3183 (phosphoadenosine phosphosulfate reductase)	1,4
EC042_RS26390/EC042_pAA031 (aggregative adherence fimbria II usher protein aafC)	1,4
EC042_RS28315/EC042_3188 (transposase)	1,3
EC042_pAA005A (conserved hypothetical protein)	1,3
EC042_RS24475/EC042_4584 (IS110 family transposase)	1,3
EC042_RS24415/EC042_4574 (type VI secretion system baseplate subunit TssK)	1,3
EC042_RS24400/EC042_4571 (type VI secretion protein)	1,3
EC042_RS24395/EC042_4570 (putative type VI secretion protein)	1,3
EC042_RS24405/EC042_4572 (type VI secretion system protein TssA)	1,3
EC042_RS21290/EC042_4006 (putative exported protein)	1,3
EC042_4583 (hypothetical protein)	1,2
EC042_4574A (putative type VI secretion system protein)	1,2
EC042_4578 (pseudogene - transposase)	1,2
EC042_RS24410/EC042_4573 (type VI secretion protein)	1,2
EC042_4580 (hypothetical protein)	1,2
EC042_RS30400 (transposase)	1,2
EC042_RS24425/EC042_4575 (DotU family type IV/VI secretion system protein)	1,2
EC042_RS27045/EC042_4585 (IS3 family transposase)	1,2
EC042_4585A (pseudogene-transposase)	1,1
EC042_RS26385/EC042_pAA030 (aggregative adherence fimbria II minor subunit aafB)	1,1
EC042_RS16985/EC042_3179A (hypothetical protein)	1,1

**Tabla 4.12.** Listado de los genes que forman parte del regulón de AggR (Morin y col., 2013; Yasir y col., 2019) cuya expresión resulta incrementada al alterar el extremo 3' del gen *aggR* respecto al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

### 4.3.2.9 Correlación entre los datos transcriptómicos obtenidos vía RNA-seq y el fenotipo observable de la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR*

#### 4.3.2.9.1 Análisis de la motilidad bacteriana

En el análisis transcriptómico realizado sobre la cepa pAA2- y la cepa que presenta la disrupción del extremo 3' del gen *aggR* se observó una importante desregulación de los genes implicados en la motilidad flagelar y elementos implicados en la síntesis y el ensamblaje de fimbrias (tablas 4.9 y 4.11). El hecho de que la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR* (que presenta unos elevados niveles de expresión del gen *aggR*) muestre un elevado nivel de expresión de genes implicados en la síntesis de flagelo, concuerda con los datos reportados en el caso del mutante *aggR*, donde se ha descrito que dicha mutación provoca una represión de la expresión del sistema flagelar (Yasir y col., 2019). En el caso de la cepa pAA2-, a pesar de no contener la proteína AggR, existe una discordancia con los valores reportados por parte de (Yasir y col., 2019). Ello parece estar indicando que, además del gen *aggR*, existen otros elementos en el plásmido pAA2 que también están implicados en la regulación de la motilidad flagelar de la cepa *E. coli* 042. Con el fin de tratar de atribuir un fenotipo a la desregulación observada en el análisis transcriptómico, se realizaron ensayos para analizar la capacidad de movimiento de las bacterias cuantificando la motilidad flagelar de las distintas cepas de *E. coli* estudiadas. Para realizar dichos experimentos, se utilizó el protocolo descrito en el apartado 3.10.3 de Materiales y Métodos. Para ello, se crecieron las cepas *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ *aggR* y *E. coli* 042 *aggR*+*FRT3*'*UTR* en medio LB durante 16 horas a 37°C, para posteriormente inocular cada una de las cepas en placas de motilidad (LB + agar 0,3% (p/v)). Las placas se incubaron durante 16 horas a 37°C con el fin de observar el grado de motilidad de las diferentes bacterias inoculadas (figura 4.79).



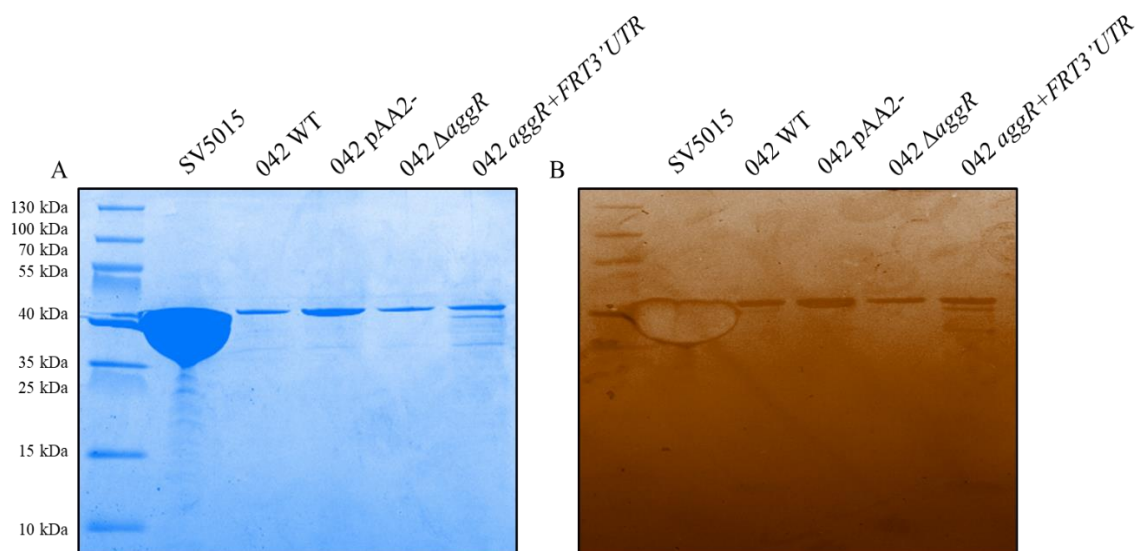
**Figura 4.79.** (A) Fotografía de los halos de motilidad de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ aggR y *E. coli* 042 aggR+FRT3'UTR inoculadas en medio LB agar 0,3% (p/v) a 37°C durante 16 horas. (B) Cuantificación del diámetro del halo de motilidad de las diferentes cepas estudiadas en medio LB agar 0,3% (p/v). Se representan los valores correspondientes a 3 réplicas biológicas independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

Los resultados obtenidos a nivel de fenotipo móvil concuerdan con la desregulación de los genes del sistema flagelar identificados en los fondos genéticos pAA2- y aggR+FRT3'UTR. Dichas cepas presentan una alta motilidad, a diferencia de las cepas WT y  $\Delta$ aggR, en las que no se pudo detectar halo de motilidad alguno.

## 4. RESULTADOS

### 4.3.2.9.2 Análisis de la síntesis flagelar

Para acabar de corroborar los datos obtenidos en el análisis transcriptómico y confirmar que el elevado fenotipo móvil de las bacterias *E. coli* 042 pAA2- y *E. coli* 042 *aggR*+*FRT3*'*UTR* era debido a un incremento en la expresión del sistema flagelar en estos fondos genéticos, se realizó un aislamiento de los flagelos producidos por parte de las diferentes cepas estudiadas. El protocolo utilizado para realizar el aislamiento flagelar, a partir de cultivos crecidos en medio LB en estático a 37°C durante 16 horas, se encuentra detallado en el apartado 3.9.10 de Materiales y Métodos. Como control positivo de que el protocolo de aislamiento funcionara correctamente se utilizó la cepa altamente flagelada de *Salmonella* Typhimurium SV5015. Todos los aislamientos se analizaron mediante electroforesis en geles SDS-PAGE (figura 4.80).



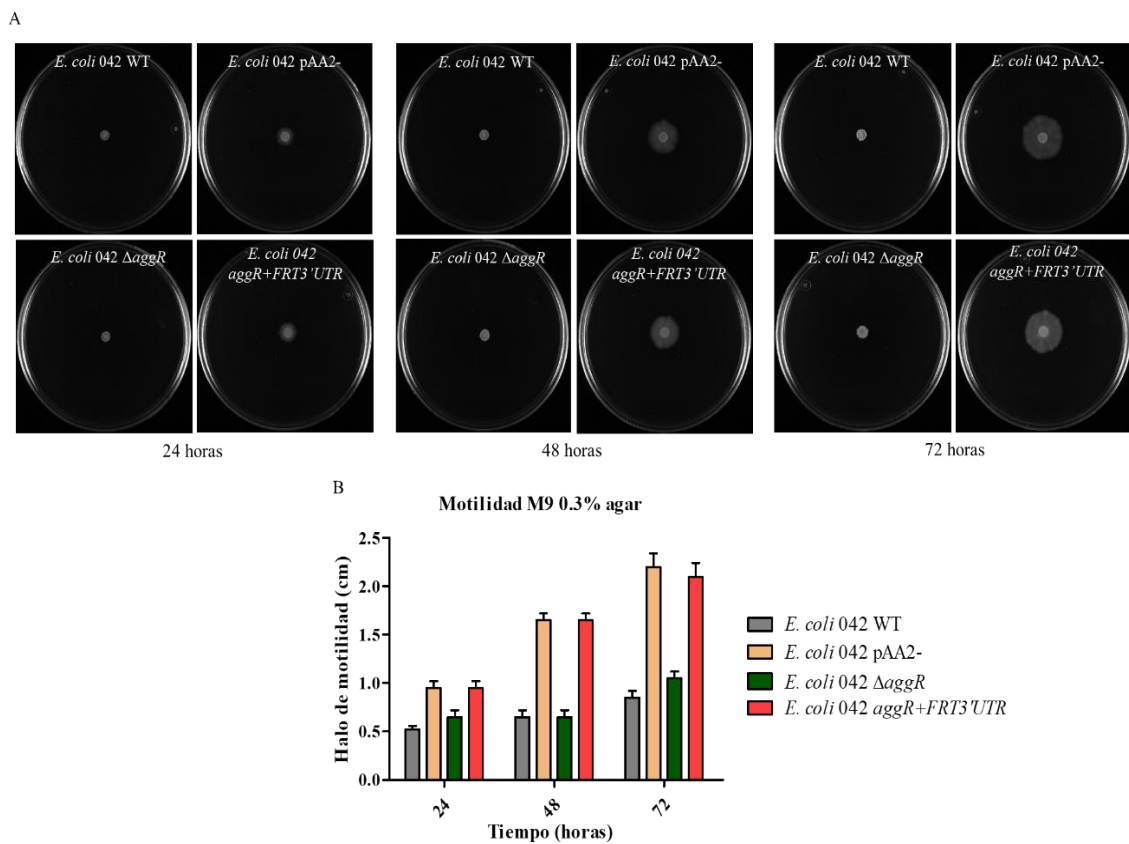
**Figura 4.80.** Tinción con Azul de Coomassie (A) y tinción de plata (B) del aislamiento flagelar de las cepas *Salmonella* Typhimurium SV5015, *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ *aggR* y *E. coli* 042 *aggR*+*FRT3*'*UTR*, crecidas en medio LB a 37°C durante 16 horas en estático. Las proteínas se analizaron mediante electroforesis en geles de acrilamida en presencia de SDS.

Tal y como se puede observar, la cepa de *Salmonella enterica* serovar Typhimurium SV5015 produce una gran cantidad de flagelina, mientras que en referencia a la cepa *E. coli* 042 su nivel de producción de proteína flagelar es muchísimo menor. En cuanto a la cepa pAA2-, sí presenta unos niveles de producción de flagelo superiores a los de la cepa salvaje y a los de la cepa mutante para el gen *aggR*, por lo que el elevado fenotipo móvil de esta cepa podría correlacionarse con una mayor síntesis flagelar. En el caso de la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR*, también se puede observar un ligero incremento en la producción flagelar, además de unas bandas adicionales, que podrían corresponderse



bien a isoformas flagelares (del sistema flagelar polar o del sistema flagelar lateral sobreexpresado en este fondo genético) o a productos de degradación flagelar.

Para acabar de confirmar el elevado fenotipo móvil de las cepas *E. coli* 042 pAA2- y *E. coli* 042 *aggR*+*FRT3*'*UTR*, se realizaron ensayos análogos a los representados en las placas de motilidad en LB (figura 4.79), pero esta vez en placas de medio M9 suplementado con glucosa 0,40% (p/v). En este caso, se incubaron los diferentes mutantes estudiados durante 24, 48 y 72 horas en placas con medio M9 suplementadas con 0,3% de agar (p/v) (figura 4.81).



**Figura 4.81.** (A) Fotografía de los halos de motilidad de las cepas *E. coli* 042 WT, *E. coli* 042 pAA2-, *E. coli* 042  $\Delta$ aggR y *E. coli* 042 *aggR*+*FRT3*'*UTR*, inoculadas en medio M9 agar 0,3% (p/v) a 37°C durante 24, 48 y 72 horas. (B) Cuantificación del diámetro del halo de motilidad de las diferentes cepas estudiadas en medio M9 agar 0,3% (p/v). Se representan los valores correspondientes a 3 réplicas biológicas independientes. Las barras de error representan la desviación estándar de los 3 valores obtenidos en cada una de las réplicas realizadas.

En referencia a la comparación de la motilidad de las bacterias en medio mínimo M9 respecto al medio LB, éstas presentan un fenotipo menos móvil cuando crecen en medio mínimo respecto a cuando lo hacen en medio rico. Sin embargo, y de manera análoga a

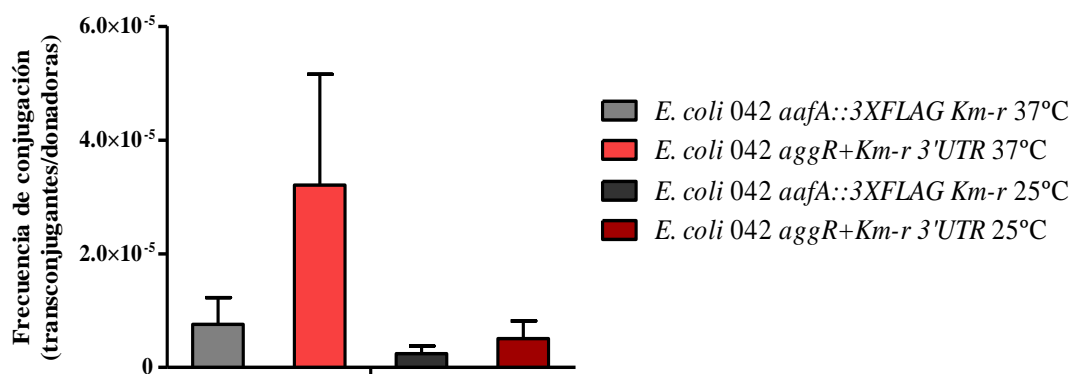


#### 4. RESULTADOS

lo observado en el caso de LB, tanto las cepas pAA2- como *aggR*+*FRT3*'*UTR* presentan un fenotipo más móvil que la cepa salvaje y la cepa mutante *aggR*.

##### 4.3.2.9.3 Análisis de la frecuencia conjugativa del plásmido pAA2

El análisis transcriptómico de la cepa *E. coli* 042 *aggR*+*FRT3*'*UTR* también mostró que bajo este fondo genético tenía lugar la desregulación de la expresión de algunos de los genes *tra* implicados en la transferencia plasmídica mediante conjugación bacteriana (tabla 9.21 Anexo). Con el fin de asociar un fenotipo a esta expresión diferencial de los genes *tra*, se analizó la tasa de conjugación del plásmido que porta la inserción del casete de resistencia a kanamicina después del codón *stop* del gen *aggR* (*E. coli* 042 *aggR*+*Km-r3*'*UTR*) y se comparó con la tasa de conjugación del plásmido de la cepa *E. coli* 042 *aafA*::*3XFLAG* *Km-r*. Como cepa receptora de la conjugación se utilizó la cepa *E. coli* 5K Rif (cepa resistente a rifampicina), por lo que las bacterias transconjugantes (5K Rif<sup>R</sup> pAA2 *aafA*::*3XFLAG* *Km-r* o 5K Rif<sup>R</sup> pAA2 *aggR*+*Km-r3*'*UTR*, células receptoras que habían adquirido el plásmido pAA2 transmitido por conjugación a partir de células donadoras) se seleccionaron en un medio selectivo suplementado con rifampicina y kanamicina. La conjugación se realizó siguiendo el protocolo descrito en el apartado 3.10.7 de Materiales y Métodos en medio 2xYT, tanto a 37°C como a 25°C (figura 4.82).



**Figura 4.82.** Representación gráfica de la frecuencia de conjugación del plásmido pAA2 a 25°C y 37°C. La interacción entre las células donadoras del plásmido pAA2 (*E. coli* 042 *aafA*::*3XFLAG* *Km-r* y *E. coli* 042 *aggR*+*Km-r3*'*UTR*) y las células receptoras (*E. coli* 5K Rif<sup>R</sup>) se realizó en medio líquido 2xYT durante 30 minutos (a 37°C) o 120 minutos (a 25°C). La frecuencia de conjugación se calculó al dividir el número de células transconjugantes obtenidas entre el número inicial de células donadoras del plásmido pAA2. Las barras de error representan las desviaciones estándar obtenidas al calcular el promedio de la frecuencia conjugativa del plásmido pAA2 correspondiente a tres réplicas biológicas independientes.

Analizando los resultados obtenidos, puede comprobarse cómo la cepa que contiene la inserción del casete de resistencia a kanamicina alterando el extremo 3' del gen *aggR* presenta una mayor tasa de transferencia conjugativa del plásmido pAA2 respecto a la cepa salvaje, tanto a 37°C como a 25°C. Este fenotipo altamente conjugativo se puede correlacionar con la sobreexpresión de los genes *tra* identificados mediante RNA-seq en el fondo genético en el que se sobreexpresa el gen *aggR* (tabla 9.21 Anexo).



## **5. DISCUSIÓN**



### 5.1 Los parálogos *hha2* y *hha3* en *E. coli*

Si bien las características de un elevado número de reguladores globales de la expresión génica han sido estudiadas de forma exhaustiva en los últimos años, éste no es el caso de sus correspondientes parálogos y ortólogos. La presencia de ortólogos de genes que codifican para reguladores globales en plásmidos no es infrecuente (Shintani y col., 2015), pero el papel biológico de los productos de tales genes es poco conocido en la mayoría de los casos. En esta memoria se ha podido poner de manifiesto que el estudio de parálogos y/u ortólogos de reguladores globales de la expresión génica puede resultar una herramienta de gran importancia para entender mejor la biología de las células bacterianas y algunas de sus redes reguladoras.

Mediante una búsqueda bioinformática en el genoma de la cepa *E. coli* 042, se ha logrado identificar en el cromosoma de dicha bacteria la presencia de dos nuevos genes parálogos al gen *hha*. Estos dos genes adicionales se denominaron *hha2* y *hha3*. A diferencia de los genes *hha* e *ydgT*, que se encuentran ampliamente distribuidos entre la familia *Enterobacteriaceae* (Madrid y col., 2007), la presencia de los genes *hha2* y *hha3* se encuentra de manera restringida únicamente en la especie *Escherichia coli*.

Los genes *hha2/hha3* se localizan en posiciones variables en el cromosoma de las diferentes cepas de *E. coli* que los contienen, y se encuentran flanqueados por elementos de inserción, transposasas e integrasas, lo que sugiere que dichos genes han sido adquiridos a través de mecanismos de transferencia génica horizontal (HGT). Además, en el caso específico de la cepa *E. coli* 042, dichos parálogos se localizan en regiones cromosómicas que forman parte de islas de patogenicidad. Todo ello no es de extrañar ya que se estima que hasta una cuarta parte del genoma de la cepa *E. coli* 042 proviene de material genético adquirido mediante mecanismos de transferencia horizontal génica en forma de islas genómicas, conformando las denominadas ROD (*regions of difference*) (Chaudhuri y col., 2010).

La asociación de los parálogos *hha2* y *hha3* a patotipos específicos de *E. coli* y, dentro de ellos, a clones altamente virulentos, proporciona una nueva herramienta para el seguimiento de la presencia de este tipo de variantes patógenas de *E. coli*. Tradicionalmente, la identificación de estos clones virulentos se ha realizado mediante una combinación de protocolos clásicos de tipificación y subtipificación (es decir, detección de genes y/o productos génicos asociados con virulencia, serotipados, pruebas

## 5. DISCUSIÓN

antimicrobianas y otros rasgos fenotípicos como actividades bioquímicas o enzimáticas), junto con aproximaciones de genotipificación desarrolladas más recientemente (electroforesis de campo pulsante, tipificación *multilocus* de secuencia o secuenciación masiva de genomas completos (Sabat y col., 2013)).

Además, debido a la relevancia de *E. coli* como un importante indicador de contaminación fecal, se han desarrollado varios métodos específicos para la detección de este microorganismo en muestras de origen ambiental y alimentario. A diferencia de las metodologías de detección clásicas y tediosas que consumen mucho tiempo, como la técnica de fermentación de múltiples tubos, los enfoques como la PCR son rápidos, altamente específicos y sensibles (Paton y Paton, 2003; Buckalew y col., 2006). Para detectar la presencia de *E. coli* en muestras ambientales por PCR, comúnmente se ha analizado la presencia de diferentes genes, como el ADNr 16S, *EF-Tu* (Molaei y col., 2015), *uidA* o *yaiO* (Molina y col., 2015), mientras que para realizar la detección de patotipos específicos de *E. coli* se han diseñado combinaciones específicas de oligonucleótidos con el fin de detectar la presencia de determinados factores de virulencia (Wu y col., 2007; Omar y Barnard, 2014; Ndlovu y col., 2015).

Sin embargo, teniendo en cuenta la elevada plasticidad genética de *E. coli*, y debido al impacto de los mecanismos de transferencia génica horizontal que dan forma al genoma de esta especie bacteriana (Touchon y col., 2009), no es sorprendente que los aislamientos que pertenecen a un patotipo específico de *E. coli* puedan mostrar combinaciones inusuales de factores de virulencia. Por lo tanto, más allá de los factores de virulencia clásicos, los marcadores de virulencia ampliamente distribuidos entre las diferentes variantes patógenas de *E. coli* pueden ayudar a identificar aislamientos virulentos de esta especie.

Como se comentaba anteriormente, un ejemplo de estos marcadores de virulencia pueden ser los parálogos *hha2* y *hha3*, cuya presencia se ha detectado especialmente entre los patotipos EAEC, STEC/VTEC, ExPEC y cepas de origen ambiental productoras de la toxina Stx2. En referencia al patotipo ExPEC, es destacable que los genes *hha2* y *hha3* se asocian a los aislamientos tipo ST131. Teniendo en cuenta que estos genes no se presentan usualmente en cepas comensales o en otros aislamientos no patógenos de *E. coli*, la detección de los genes *hha2/hha3* se puede correlacionar de una forma clara con cepas virulentas de este microorganismo.



La pobre expresión del gen *hha3* en la cepa *E. coli* 042 sugiere que quizás no sea funcional en esta cepa (dicho gen parece ser una forma truncada del dominio C-terminal de los genes *hha* y *hha2*). Con respecto al gen *hha2*, a pesar de que su expresión sea inferior a la del gen *hha*, la misma es detectable a unos niveles comparables a los de la proteína YdgT. Además, la proteína Hha2 presenta un perfil de interacción proteica muy similar al que muestra la proteína Hha, ya que es capaz de interactuar con la proteína H-NS y sus parálogos H-NS2 y StpA.

Se ha podido demostrar una asociación significativa entre la presencia conjunta de los alelos *hha2* y *hha3* con los genes *aggA*, *astA*, *shf*, *sat* y *antígeno 43*, todos ellos factores de virulencia característicos del patotipo enteroagregativo de *E. coli*. En función de ello, se ha podido asociar los genes *hha* y *hha2* a la regulación conjunta de la expresión de los diferentes genes que codifican para el *antígeno 43* en la cepa *E. coli* 042.

Además de contribuir a la regulación de genes específicos, los análisis transcriptómicos realizados en este trabajo han puesto de manifiesto que, de manera análoga a lo que sucede en el caso de otros parálogos de moduladores globales, como por ejemplo la relación que se establece entre las proteínas StpA y Hfp en referencia a H-NS, (Zhang y col., 1996; Muller y col., 2010), la proteína Hha2 actúa como un reservorio de la proteína Hha. El transcriptoma del mutante sencillo  $\Delta hha2$  muestra que dicha proteína es capaz de regular la expresión de un número limitado de genes de manera independiente al sistema comodulado por Hha/H-NS. Dichos genes corresponden a sistemas toxina-antitoxina y a proteínas putativas sin función atribuida, por lo que son necesarios futuros estudios para comprender la función específica de dichos genes y sobre cómo la proteína Hha2 es capaz de ejercer un mecanismo regulatorio de manera independiente a H-NS. Únicamente en un doble mutante  $\Delta hha\Delta hha2$  se pone de manifiesto una desregulación significativa de genes que, además de ser silenciados por Hha/Hha2, también son silenciados por H-NS.

## **5.2 H-NS2, un tercer parálogo de H-NS en la cepa *E. coli* 042**

Por lo que hace referencia al gen *hns2*, los trabajos desarrollados en esta memoria se han centrado en intentar comprender el papel de su producto génico en relación al regulador global H-NS, presente de forma genérica en un amplio grupo de microorganismos gramnegativos. ¿Por qué la cepa *E. coli* 042 ha incorporado en su genoma este parálogo del gen *hns* que no está presente en cepas comensales o incluso en los genomas de otras variantes patógenas de *E. coli*? Los estudios de (i) regulación de su expresión, (ii) análisis

## 5. DISCUSIÓN

de su papel regulador global en la cepa 042, y (iii) evolución dirigida *in vitro*, permiten responder a esta pregunta.

En primer lugar, de forma similar a lo descrito en esta memoria en relación con el caso de la proteína Hha2 y de los anteriormente referidos parálogos de H-NS, StpA y Hfp (Zhang y col., 1996; Muller y col., 2010), la proteína H-NS2 proporciona un reservorio para la proteína H-NS. Las primeras evidencias ya se obtuvieron al comparar la velocidad de crecimiento de los mutantes sencillos *hns* y *hns2* junto con la del doble mutante *hns hns2*. La combinación de las mutaciones conjuntas en los dos alelos tiene un impacto sobre la tasa de crecimiento que es superior a la del mutante *hns* sencillo. Esta suposición inicial se ha corroborado posteriormente con otros resultados que se refieren a continuación.

Un aspecto importante y diferencial entre las proteínas H-NS y H-NS2 es el conjunto de genes regulados por cada una de ellas, donde H-NS se encarga de regular un conjunto de genes significativamente superior al número de genes regulados por H-NS2. De forma interesante, los genes que muestran un mayor nivel de sobreexpresión en el mutante *hns2* coinciden con los genes más sobreexpresados en el caso del doble mutante *hha hha2*. Ello nos indica que la proteína H-NS2 es capaz de modular, junto con Hha (y Hha2), un variado grupo de genes en la cepa *E. coli* 042, la mayoría de función desconocida y de probable origen HGT. Por el contrario, muchos de los genes que se encuentran significativamente sobreexpresados en el mutante sencillo *hns*, únicamente se encuentran ligeramente sobreexpresados en el mutante *hns2*. A diferencia del origen HGT de los genes regulados por la proteína H-NS2, la gran mayoría de genes regulados por la proteína H-NS pertenecen a genes del núcleo del genoma. Ejemplos de ello son los genes que forman parte de los operones *gad*, *mat* y *hde*, genes del *core* cromosómico implicados en la respuesta a estrés ácido. Sin embargo, también es destacable el hecho de que muchos de estos genes presenten una máxima desrepresión en un doble mutante *hns hns2*. Ello nos indica la importancia del alelo *hns2* como un reservorio del gen *hns*, ya que la ausencia conjunta de ambos genes genera efectos drásticos sobre la expresión de un gran número de genes en la cepa *E. coli* 042. Ante esta situación se ha podido observar una gran sobreexpresión del gen que codifica para la proteína StpA, con la probable finalidad de tratar de compensar, al menos parcialmente, la pérdida de las proteínas H-NS y H-NS2.

Por tanto, mientras que la proteína H-NS es capaz de regular principalmente genes del núcleo del genoma (además de también presentar la capacidad de regular eficientemente genes HGT) (Navarre y col., 2006)), H-NS2 se encarga de regular específicamente genes de origen HGT.

El hecho de que H-NS2 fundamentalmente regule la expresión de genes HGT tiene relación con lo observado en el caso de la proteína H-NS codificada por el plásmido IncHI1 R27 (H-NS<sub>R27</sub>) donde, en *Salmonella*, dicha proteína se encarga de silenciar específicamente genes HGT, al igual que la proteína Hha (Baños y col., 2009). Para mantener una correcta homeostasis celular, queda patente que los reguladores globales como H-NS necesitan desarrollar mecanismos para discriminar entre diferentes grupos de genes (principalmente discernir entre genes de origen HGT y genes que forman parte del núcleo del genoma). Proteínas tales como Hha, H-NS2 y H-NS<sub>R27</sub> ayudarían a diferenciar ambos grupos de genes, si bien desconocemos el mecanismo por el cual dichas proteínas son capaces de reconocer las regiones reguladoras de unos genes y de otros no. Ensayos de retardo en gel con diferentes fragmentos de ADN correspondientes a la región reguladora de ambos grupos de genes no permitieron poner de manifiesto que H-NS2 se una preferencialmente a ninguno de ellos, tanto en presencia como en ausencia de la proteína Hha (resultados no publicados del grupo de investigación). Adicionalmente, diferentes análisis realizados *in silico* sobre el grupo de genes regulados por H-NS2 tampoco permitieron observar diferencias ni en la curvatura, el porcentaje de AT o la presencia/ausencia de secuencias consenso de unión de H-NS en la región promotora de dichos genes (resultados no publicados del grupo de investigación).

Un aspecto importante de los estudios de transcriptómica realizados en este trabajo corresponde al impacto de la temperatura sobre el papel de las proteínas H-NS y H-NS2 en la regulación de la expresión génica en la cepa *E. coli* 042. Teniendo en cuenta que esta cepa es virulenta y necesita, por tanto, silenciar un grupo significativo de genes a temperaturas fuera del hospedador, es aparente que H-NS cumple este papel. En función de la temperatura, existe un silenciamiento mayor de la expresión de muchos genes a 25°C que a 37°C. Si bien los genes sobreexpresados en un mutante *hns* son básicamente los mismos a ambas temperaturas, el nivel de sobreexpresión de dichos genes es muy superior a 25°C que a 37°C. El hecho de que a esta última temperatura H-NS siga silenciando al menos en parte la expresión de estos genes se debe a que, para que se alivie completamente el silenciamiento de estos genes, se requieren otros estímulos diferentes

## 5. DISCUSIÓN

de la temperatura (Navarre y col., 2007). Por lo que respecta a la proteína H-NS2, excepto para un pequeño número de genes, su papel regulador a baja temperatura parece ser irrelevante. H-NS2 parece actuar como un factor que ayuda a mantener silenciados genes HGT a elevada temperatura a la espera que otros factores diferentes (factores de contra silenciamiento) determinen la máxima expresión de los mismos.

El patrón de expresión de la proteína H-NS2 es completamente diferente al de la proteína H-NS. Mientras que esta última proteína presenta un elevado nivel de expresión que es independiente de muchos factores fisiológicos, éste no es el caso de H-NS2. El patrón de expresión de esta proteína es dependiente de la fase de crecimiento, incrementándose su expresión cuando las bacterias entran en la fase de crecimiento estacionaria (baja tasa de crecimiento) respecto a cuando las bacterias se dividen de forma activa (fase de crecimiento exponencial). Ello puede estar relacionado con las condiciones que tienen lugar dentro del hospedador, resaltando de esta manera el papel regulador de H-NS2 prevalentemente a 37°C en condiciones de colonización. En relación con esto, la temperatura también es un factor que influye en el nivel de expresión de esta proteína, reduciéndose significativamente la misma cuando las bacterias crecen a 25°C respecto a cuando lo hacen a 37°C. Sin embargo, independientemente de la temperatura de crecimiento, el nivel de expresión de la proteína H-NS2 siempre es notablemente inferior al de la proteína H-NS. El limitado papel regulador de H-NS2 a baja temperatura puede ser debido a su bajo nivel de expresión en estas condiciones.

El experimento de evolución dirigida diseñado con la finalidad de identificar mutaciones en algunos de los dominios de H-NS2 que provocaran que la proteína se asemejara a H-NS con el fin de mimetizar su función molecular, ha permitido resaltar la importancia de los niveles de expresión de H-NS2 en la fisiología de la cepa *E. coli* 042. Utilizando como modelo el crecimiento de un mutante *hns* de la cepa 042 en medio LB a 25°C, todos los clones obtenidos que incrementaron de manera significativa la tasa de crecimiento con respecto al mutante *hns* inicial presentaron siempre la misma mutación: una transición C→T en el promotor del gen *hns2* que incrementa el nivel de expresión de su producto génico. Además, el doble mutante *hns stpA* que, de forma sorprendente presenta una tasa de crecimiento similar a la de la cepa salvaje (a pesar de contener la mutación en el gen *hns*), también muestra la misma mutación en el promotor del gen *hns2*.

Por tanto, mientras que en *Salmonella* Typhimurium la presión selectiva causada por la mutación en el gen *hns* recae sobre el gen *stpA* (provocando mutaciones con cambio de

sentido que dan lugar a variaciones en la secuencia aminoacídica de la proteína StpA (Ali y col., 2014)), en el caso de la cepa *E. coli* 042, en las condiciones de cultivo analizadas en este trabajo, dicha presión selectiva no provoca cambios en la secuencia aminoacídica de H-NS2, sino que provoca un único cambio en la región promotora del gen, concretamente en la secuencia -10. Dicha secuencia nucleotídica, denominada caja *Pribnow* o caja TATA (por su alto contenido en A y T, hecho que permite separar fácilmente la doble hebra de ADN), es la secuencia reconocida por la subunidad  $\sigma 70$  de la ARN polimerasa, y se localiza 10 nucleótidos antes del inicio de transcripción del gen (Pribnow, 1975; Schaller y col., 1975). En la literatura científica se han reportado varios casos de mutaciones puntuales en la secuencia -10 que provocaban cambios en la expresión del gen situado bajo el control de dicha secuencia promotora. Por ejemplo, se ha descrito que una mutación en el primer nucleótido de la caja *Pribnow* que suponga la aparición de una timina (T) en los genes *ompC* y *ompF*, provoca un incremento en la actividad promotora y un incremento de la expresión de las porinas Omp (Ozawa y col., 1987). Otro ejemplo de ello tiene lugar en el caso de *Klebsiella oxytoca*, donde se ha descrito que mutaciones puntuales en la caja *Pribnow* del gen *blaOXY* provocaban la sobreexpresión de beta lactamasas de amplio espectro respecto a la secuencia salvaje (Fournier y col., 1995).

Estos resultados ayudan a reafirmar la hipótesis de que el gen *hns2* de la cepa *E. coli* 042 representa un reservorio de la proteína H-NS de manera que, si el nivel de expresión de la misma se redujera de forma indeseada, la transición C→T permitiría incrementar la expresión de *hns2*, generando elevados niveles de proteína H-NS2 que puedan compensar una eventual carencia de la proteína H-NS. Sin embargo, este supuesto no tiene que relacionarse únicamente con una poco probable mutación en el gen *hns*. Tal y como se ha descrito por parte de (Dorman, 2014), la incorporación al azar (por parte de las bacterias) de ADN mediante mecanismos de transferencia génica horizontal puede provocar que la concentración intracelular de H-NS se diluya al tener que interaccionar la proteína residente con las secuencias ricas en A/T presentes en el nuevo ADN exógeno incorporado. Por ello, algunos plásmidos incorporan una copia extra de H-NS, para ayudar a mantener elevadas las concentraciones intracelulares de la proteína y evitar que su adquisición reduzca el *fitness* celular (Dorman, 2014). Sin embargo, la incorporación de ADN de origen HGT que no incluya genes *hns* puede llegar a redundar en una reducción del *fitness*. En contraposición a ello, la disponibilidad del gen *hns2* puede

## 5. DISCUSIÓN

garantizar que, frente a estos procesos de incorporación de ADN foráneo, la dilución de la proteína H-NS pueda compensarse con un incremento en la expresión de H-NS2. Por tanto, una hipótesis plausible es que la presencia de genes como *hns2* puede ayudar a facilitar y soportar la incorporación de ADN HGT por parte de las bacterias.

En cualquier caso, la fina regulación de la expresión de la proteína H-NS2 parece ser un factor crítico para la fisiología de la cepa 042. Ante nuestra sorpresa, cuando se intentó complementar un mutante *hns* aportando en *trans* el gen *hns2* clonado en el vector pLG388.30 bajo el control de su promotor salvaje (*phns2*) o bajo el control del promotor que presenta la transición C→T (*phns2\**), se produjo complementación de la tasa de crecimiento y del efecto desregulador de la mutación *hns* únicamente en la cepa *E. coli* 042  $\Delta hns$  pLG388.30*phns2*, pero no en la cepa *E. coli* 042  $\Delta hns$  pLG388.30*phns2\**. Es comprensible la complementación con el primer sistema ya que, a pesar de que el vector pLG388.30 sea un vector de bajo número de copias, el incremento de expresión de H-NS2 dirigido desde el gen *hns2* plasmídico puede ayudar a incrementar los niveles de la proteína de forma suficiente como para ser equivalentes al de una única copia cromosómica con la mutación *phns2\**. Resultó incomprensible el efecto que se obtuvo en la cepa *E. coli* 042  $\Delta hns$  pLG388.30*phns2\** hasta que, al secuenciar el plásmido introducido en la misma, pudo comprobarse que una copia del elemento *IS5* se había insertado en la secuencia codificante del gen *hns2* clonado, inactivándolo en su totalidad. Estos resultados muestran que, si bien incrementos moderados en la expresión del gen *hns2* como los causados por la mutación *phns2\** pueden ser beneficiosos para la fisiología celular en ciertas circunstancias, una expresión más elevada tiene efectos deletéreos, generando estrés, el cual se traduce en movimientos de elementos *IS* y la consecuente selección de clones que portan la inactivación insercional del gen *hns2* bajo el control de *phns2\**.

En este trabajo también se ha podido determinar la interacción *in vitro* de la proteína H-NS2 con la proteasa Lon. Esta interacción ya fue en su momento descrita en el caso de la proteína StpA (Johansson y Uhlin, 1999), donde la presencia de la proteína H-NS proporcionaba un mecanismo protector ante la degradación proteolítica de esta proteína por parte de la proteasa Lon. Ya que la fenilalanina presente en la posición 21 (Phe21) de la secuencia codificante de la proteína StpA es requerida para la sensibilidad de StpA al procesamiento proteolítico por parte de la proteasa Lon (mientras que en la proteína H-NS el aminoácido que se encuentra en dicha posición es una cisteína), el hecho de que

en la proteína H-NS2, el aminoácido localizado en la posición 21 también sea un residuo altamente hidrofóbico como es la leucina (Leu21), apoyó la hipótesis de la sensibilidad de esta proteína ante la acción proteolítica de la proteasa Lon. No obstante, al analizar la estabilidad de H-NS2 tanto en el fondo genético salvaje como en mutantes *hns* y *stpA* (además de *lon*), no se pudo demostrar inestabilidad alguna de la proteína ante la degradación proteolítica. Por todo ello, con la información disponible, no podemos concluir que la interacción de H-NS2 con Lon sea un artefacto, o que no se hayan encontrado las condiciones en las cuales Lon procesa proteolíticamente a H-NS2.

Para finalizar este apartado, cabe destacar que todas las cepas que expresan H-NS2 también expresan H-NS y StpA, lo cual indica que la proteína H-NS2 presenta características reguladoras diferentes a las de StpA. A nivel de secuencia aminoacídica, la proteína H-NS2 muestra un nivel de similitud global a H-NS superior al que muestra la proteína StpA. Sin embargo, cuando se comparan los dominios de dimerización/oligomerización y de unión al ADN de H-NS2 y StpA respecto a los dominios de H-NS, el primero muestra un nivel de similitud superior al que presenta dicho dominio en StpA, pero el segundo no. Este hecho nos sugiere que la proteína H-NS2 presenta propiedades de dimerización/oligomerización más parecidas a H-NS (por ejemplo, al comodular genes con la proteína Hha), y difiere en sus propiedades de unión al ADN (lo que no le permite reconocer algunos de los genes diana de H-NS). Todas las cepas identificadas hasta el momento que contienen el gen *hns2* son virulentas y presentan múltiples determinantes de resistencia a antibióticos (Johnson y col., 2012; Conlan y col., 2014). Las enterobacterias que incorporan el sistema H-NS2 a su genoma se proveen de un elemento importante que permite complementar la función de H-NS bajo diferentes condiciones fisiológicas. Por esta razón, el estudio en profundidad de este regulador puede ayudar a contribuir a desarrollar estrategias específicas para combatir las infecciones causadas por parte de estas cepas.

En función de la información disponible, podemos concluir que los parálogos de H-NS presentan algunas propiedades comunes: (i) a pesar de converger en la regulación de muchos genes comunes, el conjunto de genes diana puede ser diferente al de H-NS; (ii) sus niveles de expresión son inferiores a los de H-NS, presentando además un perfil de expresión diferente; (iii) pueden potenciar la capacidad reguladora de H-NS bien al formar complejos heteroméricos con esta proteína, o reconociendo solamente un subconjunto de los genes ayudando a disminuir la cantidad neta de moléculas de H-NS



## 5. DISCUSIÓN

necesaria para esta finalidad y, (iv) constituyen un reservorio funcional de la proteína H-NS ante situaciones en las que la concentración celular de la misma pueda verse disminuida.

A modo de conclusión de este apartado, es importante destacar que en el análisis de la virulencia en microorganismos patógenos se ha considerado usualmente la presencia o ausencia de los correspondientes determinantes de virulencia. Los resultados presentados en este trabajo en referencia a los diferentes parálogos de Hha y H-NS estudiados, ponen de manifiesto que también es importante considerar la presencia de los alelos específicos de reguladores globales de la expresión génica. La disponibilidad progresiva de secuencias genómicas completas de muchas cepas patógenas puede facilitar el correspondiente análisis de la presencia de parálogos de reguladores globales, y proporcionará una visión más completa de la complejidad de los regulones de virulencia que presentan algunos patógenos.

### **5.3 Importancia del plásmido pAA2 en la virulencia bacteriana y nuevos aspectos de la regulación del gen *aggR*: papel de la región 3'UTR**

Hasta la fecha, la información existente acerca del papel del regulador AggR sobre la virulencia de las cepas enteroagregativas de *E. coli* se había obtenido fundamentalmente analizando el efecto de mutaciones en el gen *aggR* sobre el patrón global de expresión génica de la cepa 042 (Dudley y col., 2006; Morin y col., 2013; Yasir y col., 2019). Nuestro interés en analizar las interacciones del plásmido pAA2 con las redes reguladoras de la cepa *E. coli* 042 nos llevó a observar un fenotipo inesperado, hecho que nos ha permitido comprender mejor el papel del regulador AggR.

En referencia al papel del plásmido pAA2 en la virulencia de la cepa *E. coli* 042, el análisis del transcriptoma global de un clon libre del plásmido nos proporcionó un resultado inesperado: la mayoría de los genes que se encontraban sobreexpresados en la cepa pAA2- coincidían con aquellos que se encontraban sobreexpresados en un mutante  $\Delta hha$  y, en menor medida, en el caso de los mutantes  $\Delta hns$  y  $\Delta hns\Delta hns2$ . Estos resultados se pueden interpretar a partir de un artículo publicado recientemente (Santiago y col., 2017), en el que se estudia el papel de la proteína Aar, codificada por el plásmido pAA2. Dicha proteína pertenece a la familia ANR (reguladores negativos de proteínas tipo AraC) y presenta la capacidad de unirse a AggR para reprimir la capacidad de ésta de activar la transcripción (Santiago y col., 2014; Santiago y col., 2016). Un análisis transcriptómico

realizado sobre el mutante *aar* de la cepa *E. coli* 042 puso de manifiesto que, ante la ausencia de esta proteína, se producía una disminución en la expresión transcripcional de los genes *hns*, *hns2* y *stpA* (Santiago y col., 2017). Teniendo en cuenta que muchos genes de la cepa *E. coli* 042 se encuentran coregulados por H-NS/H-NS2 y Hha/Hha2, parece razonable la explicación de que el curado del plásmido pAA2, que conlleva un genotipo *aar* y la consecuente disminución de la expresión de estos reguladores globales, tenga como consecuencia la desregulación de genes silenciados por Hha/Hha2 y H-NS/H-NS2.

El otro gran grupo de genes cromosómicos cuya expresión también se ha visto afectada ante la ausencia del plásmido pAA2 corresponde a la sobreexpresión de genes implicados en la motilidad flagelar, además de la represión de los genes del cromosoma activados por la proteína AggR (esencialmente el sistema de secreción tipo VI presente en la isla de patogenicidad Aai), debido a la ausencia de este regulador transcripcional como consecuencia de la pérdida del plásmido.

Debido al inesperado fenotipo obtenido al insertar un fragmento de ADN (bien sea un gen de resistencia a un antibiótico o bien la cicatriz nucleotídica residual FRT resultante de la eliminación de dicho gen) en el extremo 3' del gen *aggR*, decidimos continuar esta parte del trabajo estudiando el mecanismo por el cual la expresión del gen *aggR* se veía influenciada por la interrupción de secuencias localizadas en la posición 3'UTR de dicho gen.

Entre los diferentes fenotipos que se han observado al colocar una secuencia nucleotídica en la posición 3'UTR del gen *aggR*, uno de ellos correspondía a la selección de clones que perdían el plásmido pAA2 debido a una inusual inestabilidad plasmídica. Nuestro interés inicial en obtener un clon derivado de la cepa *E. coli* 042 libre de plásmido implicó realizar un gran número de subcultivos para obtener un único clon curado. Paradójicamente, los estudios posteriores realizados sobre el gen *aggR* nos permitieron curar de una manera mucho más fácil el plásmido pAA2 de la cepa *E. coli* 042. Por tanto, la manipulación de las secuencias nucleotídicas situadas aguas abajo del gen *aggR* han permitido desarrollar una estrategia relativamente sencilla para provocar la pérdida del plásmido pAA2. Probablemente, esta estrategia también podría resultar válida para curar otros plásmidos del grupo de pAA que codifiquen para un gen tipo *aggR*.

En este trabajo, se ha podido demostrar que el elevado fenotipo autoagregativo de los clones que presentaban la inserción de la cicatriz FRT entre el codón de finalización de

## 5. DISCUSIÓN

la traducción del gen *aggR* y su secuencia 3'UTR (fondo genético *aggR+FRT3'UTR*), era debido a un notable incremento en la expresión del gen *aggR* y, por tanto, de la expresión de los genes cuya transcripción es dependiente de su producto génico. Adicionalmente al incremento de expresión del gen *aggR* (cuantificado mediante *Northern Blot*), también pudo ponerse de manifiesto la sobreexpresión tanto de la subunidad mayor de las fimbrias de adherencia agregativa de tipo II (AafA) como de la dispersina (Aap) en la membrana externa y secretoma (respectivamente) de la cepa *E. coli* 042 *aggR+FRT3'UTR*.

El análisis transcriptómico comparativo entre la cepa *E. coli* 042 *aggR+FRT3'UTR* y la cepa *E. coli* 042 WT nos ha permitido complementar la información sobre el efecto que la sobreexpresión del gen *aggR* genera en la fisiología de la cepa 042. Además de la esperada inducción de factores de virulencia dependientes de AggR, en dicho fondo genético también se ha podido observar una importante inducción de genes implicados en la síntesis flagelar y en la transferencia conjugativa del plásmido pAA2. En concordancia con estos datos transcriptómicos, la cepa *E. coli* 042 *aggR+FRT3'UTR* presenta una mayor motilidad y una superior tasa de conjugación plasmídica respecto a la cepa salvaje. Ello significa que, si *in vivo* se produjeran condiciones ambientales que conduzcan a la sobreexpresión de AggR, dichas células presentarán una mayor capacidad de transferencia de su plásmido de virulencia a otras células receptoras, lo que puede dar lugar a la aparición clones virulentos a partir de cepas de *E. coli* residentes en el tracto gastrointestinal.

Merece una mención específica el hecho de que, entre los genes sobreexpresados en la cepa *E. coli* 042 *aggR+FRT3'UTR*, se encuentren genes que forman parte de dos vías metabólicas específicas: la vía *ast* de degradación de la arginina, y las vías *fad* y *aceAB*, que permiten degradar ácidos grasos a través de la  $\beta$ -oxidación y a través del ciclo del glioxilato, respectivamente. Hoy en día está establecido que, tanto el metabolismo de la arginina (Gogoi y col., 2016) como el de los ácidos grasos de cadena corta (Sun y O'Riordan, 2013), juegan un papel esencial tanto en la patogénesis microbiana como en la homeostasis del microbioma intestinal (Postler y Ghosh, 2017). El operón *fad* es inducido en respuesta a la presencia de ácidos grasos de cadena larga contactando la membrana celular, pero la inducción del operón independiente de esos ácidos grasos permite el catabolismo de ácidos grasos de cadena corta tales como el acético, láctico, propiónico o butírico, compuestos caracterizados por jugar un papel inmunomodulador

en el intestino (Clark, 1981). Una hipótesis interesante, que trasciende a esta tesis doctoral, es que estas vías metabólicas, inducidas por la sobreexpresión de AggR, puedan jugar un papel importante en la patogénesis de las cepas enteroagregativas de *E. coli*, alterando de esta manera la homeostasis intestinal al reducir la concentración de arginina o la de ácidos grasos de cadena corta. Un artículo muy reciente sugiere que la proteína AggR es capaz de inducir factores de virulencia (no caracterizados hasta la fecha) para producir enfermedad (Boisen y col., 2019). Planteamos como una posible hipótesis que las vías metabólicas identificadas en este trabajo puedan estar implicadas en estos mecanismos.

Un aspecto muy importante por resaltar es el mecanismo por el cual la disrupción de la secuencia nucleotídica correspondiente al extremo 3'UTR del gen *aggR* provoca un aumento de su expresión. En primer lugar, el análisis bioinformático realizado en este trabajo ha demostrado que, asociada a los genes *aggR* de diferentes plásmidos pAA de diferentes cepas enteroagregativas de *E. coli*, se encuentra conservada una secuencia común de aproximadamente 44 pb localizada más allá del último codón de la secuencia codificante de los diferentes genes *aggR* analizados. En el caso del gen *aggR* del plásmido pAA2 de la cepa *E. coli* 042, aguas abajo de su secuencia 3'UTR, se localiza el elemento de inserción *ISIA*, cuya organización estructural se basa en una primera repetición terminal invertida (IRL), seguida por la secuencia codificante de la transposasa presente en este elemento de inserción, y una segunda repetición terminal invertida (IRR). Es importante destacar que en todas las cepas analizadas bioinformáticamente en este trabajo también existe un elemento de inserción situado aguas abajo de su correspondiente gen *aggR*, lo que sugiere que inserciones al azar en esa región se han seleccionado porque presentan un efecto sobre la fisiología de las cepas enteroagregativas de *E. coli* (portadoras de plásmidos tipo pAA) que es seleccionable positivamente.

Los estudios de hibridación de ARN mediante *Northern Blot* mencionados anteriormente, además de confirmar de manera inequívoca el incremento significativo en la concentración intracelular del transcrito de *aggR* al colocar una secuencia FRT más allá del codón de terminación de la traducción del gen, también han demostrado que se produce un aumento en la cantidad neta de dicho transcrito tanto si la secuencia FRT se coloca en la secuencia IRL del elemento *ISIA* como si se coloca más allá de la misma, delecionando el propio elemento de inserción. La explicación más razonable para este hecho es que, por un mecanismo todavía por determinar, la inserción de secuencias FRT

## 5. DISCUSIÓN

en la región terminal del gen *aggR* pudieran estar generando una alteración en la topología de la región 3' del ARNm de dicho gen, aumentando de esta manera su estabilidad.

Dichos análisis por *Northern Blot* también han puesto de manifiesto un resultado inesperado en relación con el elevado tamaño del transcrito detectado perteneciente a la molécula de ARNm del gen *aggR*, un tamaño mucho mayor al que correspondería por la longitud de la secuencia codificante del gen en sí. El gran tamaño de dicho transcrito podría estar relacionado con el hecho de que no exista un terminador transcripcional Rho-independiente aguas abajo de *aggR* hasta 3,7 Kb posteriores al codón *stop* del gen. Los resultados de “*walking-PCR*” apoyan los resultados obtenidos mediante *Northern Blot*, y muestran cómo el transcrito que se inicia en *aggR* continúa aguas abajo del mismo, llegando a penetrar en el elemento de inserción *ISA1* situado aguas abajo del gen. Es razonable pensar que tales transcritos puedan ser objeto de degradación por parte de ARNasas celulares, reduciéndose de esta manera la concentración de los mismos y, por tanto, la expresión de AggR. La modificación de dichas secuencias por la adición de una secuencia FRT alteraría la estructura de tal transcrito, retrasando al menos parcialmente la degradación del ARNm correspondiente, permitiendo una expresión superior de AggR.

Sin embargo, cabe destacar que, a pesar de la ausencia de terminadores transcripcionales Rho-independientes cerca del final del gen *aggR*, sí que se ha logrado identificar la presencia de posibles terminadores Rho-dependientes aguas abajo de dicho gen, facilitados los dos primeros por la inserción del elemento *ISIA*. Una posible hipótesis para interpretar estos resultados sería que, en ausencia de niveles elevados de Rho (que se induce en condiciones de estrés), la transcripción del gen *aggR* seguiría aguas abajo del mismo, penetrando en el elemento *ISIA*. En cambio, en condiciones de estrés, en presencia de Rho, los transcritos del gen *aggR* finalizarían poco después de la secuencia codificante del gen. Si dichos transcritos cortos fuesen más estables, ello provocaría un aumento en la expresión de AggR, induciéndose de esta forma la expresión de los determinantes de virulencia de la cepa 042 regulados por esta proteína (además de las vías metabólicas mencionadas anteriormente). De esta forma, tendría sentido la presencia de dichos elementos de inserción en la región 3'UTR del gen *aggR* de las cepas enteroagregativas de *E. coli*, facilitando en cierta manera la inducción de la expresión de AggR en respuesta a estrés.

Finalmente, es importante resaltar que, en este caso, la secuencia 3'UTR implicada en la regulación de la expresión de *aggR* es insólitamente larga. En otros ejemplos estudiados

en bacterias, la longitud de las regiones 3'UTR varía entre los 63 nucleótidos del gen *aceA* en *C. glutamicum* (Maeda y Wachi., 2012), los 283 nucleótidos del gen *hmsT* en *Y. pestis* (Zhu y col., 2016), y los 310 nucleótidos del gen *hilD* en *S. enterica* (López-Garrido y col., 2014). Es aparente que la inserción del elemento *ISIA* en la región intergénica entre los genes *aggR* y *aafA* extiende dicha región de forma significativa y que, alteraciones en la misma, provocan grandes cambios en la estabilidad del ARNm de *aggR*.

Una cuestión que queda pendiente y trasciende a esta tesis doctoral es el papel de las ARNasas celulares implicadas en la degradación del transcrito correspondiente al gen *aggR*. En el caso del ARN mensajero correspondiente al gen *hilD*, se ha descrito que su región 3'UTR es diana del degradasoma bacteriano, donde tanto las ribonucleasas ARNasa E como la PNPasa son responsables de la degradación de dicho transcrito, cuya terminación transcripcional es Rho-independiente (López-Garrido y col., 2014). En el caso del gen *hmsT* en *Y. pestis*, cuya terminación transcripcional es Rho-dependiente, únicamente la ribonucleasa PNPasa está implicada en la degradación, dependiente de la región 3'UTR, de dicho transcrito (Zhu y col., 2016). En la actualidad, seguimos trabajando en esta problemática para tratar de comprender cómo se regulan a nivel post-transcripcional los niveles de expresión de *aggR*, ya que la expresión de este importante regulador transcripcional es un factor crítico en la fisiología y virulencia de la bacteria *E. coli* 042. Ello es importante porque los mecanismos de regulación post-transcripcional son mecanismos de respuesta rápida muy importantes, que permiten a las bacterias adaptarse ante cualquier cambio ambiental y que actúan en la fase final del proceso modulador de la expresión génica, que al final va a acabar definiendo cuál es la cantidad neta de producto génico activo sometido al proceso de traducción para acabar convirtiéndose en la forma final de la proteína efectiva.





## **6. CONCLUSIONES**



Las conclusiones que se pueden extraer a lo largo de este trabajo son las siguientes:

- 1) La presencia de los parálogos *hha2* y *hha3* es altamente prevalente en variantes patógenas de *E. coli* y mucho menos prevalente en cepas comensales, por lo que pueden ser considerados como buenos marcadores de cepas virulentas de *E. coli*.
- 2) Los genes *hha2* y *hha3* se encuentran asociados a islas de patogenicidad adquiridas a través de mecanismos de transferencia génica horizontal, y su presencia se correlaciona de manera positiva con la presencia de los marcadores moleculares propios del patotipo enteroagregativo de *E. coli*, tales como *aggA*, *astA*, *shf*, *sat* y *agn43*.
- 3) Una de las principales funciones de la proteína Hha2 es la de actuar como un reservorio de la proteína Hha.
- 4) La presencia del gen *hns2*, nuevo parálogo de *hns* identificado en la cepa *E. coli* 042, se encuentra restringida a cepas patógenas de la familia *Enterobacteriaceae*, dentro de los géneros *Escherichia*, *Klebsiella*, *Serratia* y *Salmonella*.
- 5) A diferencia de H-NS, la temperatura determina el grupo de genes regulados por H-NS2. Mientras que H-NS reprime la expresión de un amplio número de genes tanto a 25°C como a 37°C, la función reguladora principal de H-NS2 se ejerce esencialmente a 37°C, donde se encarga de reprimir junto con H-NS una parte de los genes regulados por esta última proteína (principalmente genes de origen HGT).
- 6) De manera análoga a lo que sucede con la proteína Hha2, la proteína H-NS2 actúa como un reservorio molecular de la proteína H-NS. En ausencia del gen *hns*, se produce un incremento de la expresión de H-NS2 mediante un mecanismo que implica una transición C→T en la secuencia -10 del promotor del gen *hns2*, el cual permite compensar parcialmente la pérdida funcional de la proteína H-NS.
- 7) Aspectos comunes de los parálogos de reguladores globales son: (i) menor nivel de expresión que la proteína de referencia; (ii) capacidad de corregular con dicha proteína la expresión de algunos genes; (iii) posibilidad de regular la expresión de otros genes de manera independiente, y (iv) capacidad funcional de actuar como un reservorio molecular del regulador de referencia.
- 8) La pérdida del plásmido pAA2 de la cepa *E. coli* 042 provoca grandes efectos en la fisiología de la bacteria: mayor velocidad de crecimiento, mayor motilidad, incremento en la capacidad de síntesis flagelar e imposibilidad de realizar *biofilm*.

## 6. CONCLUSIONES

- 9) La expresión del gen *aggR* del plásmido pAA2 se encuentra regulada a nivel post-transcripcional. Alteraciones en la región 3'UTR del mismo aumentan su expresión.
- 10) El transcrito del gen *aggR* se extiende más allá de dicho gen, probablemente debido a la inexistencia de un terminador transcripcional cerca de su región codificante.
- 11) Los genes *aggR* de diferentes plásmidos pAA presentan una región 3'UTR común, seguida siempre de la presencia de un elemento de inserción *IS*, lo que sugiere una implicación de tales elementos en la expresión de dichos genes.
- 12) A nivel fenotípico, el aumento de la expresión de *aggR* genera grandes cambios en la cepa *E. coli* 042, tales como la disminución de la tasa de crecimiento, incremento en la autoagregación celular, incremento de la motilidad bacteriana y un incremento en la frecuencia de conjugación y en la inestabilidad del plásmido pAA2.
- 13) Incrementos en los niveles de *aggR* provocan, además de un aumento de la expresión de los genes ya descritos como parte del regulón AggR, incrementos en la expresión de genes que codifican para enzimas de vías metabólicas específicas, las cuales pueden estar relacionadas con la actividad patogénica de las cepas enteroagregativas de *E. coli* en el intestino.

## **7. BIBLIOGRAFÍA**



- Afset, J.E., Bruant, G., Brousseau, R., Harel, J., Anderssen, E., Bevanger, L., & Bergh, K. (2006).** Identification of virulence genes linked with diarrhea due to atypical enteropathogenic *Escherichia coli* by DNA microarray analysis and PCR. *Journal of Clinical Microbiology*, 44(10), 3703–3711.
- Ali, S.S., Beckett, E., Bae, S.J., & Navarre, W.W. (2011).** The 5.5 protein of phage T7 inhibits H-NS through interactions with the central oligomerization domain. *Journal of bacteriology*, 193(18), 4881–4892.
- Ali, S.S., Xia, B., Liu, J., & Navarre, W.W. (2012).** Silencing of foreign DNA in bacteria. *Current Opinion in Microbiology*, 15(2), 175–181.
- Ali, S.S., Whitney, J.C., Stevenson, J., Robinson, H., Howell, P. L., & Navarre, W.W. (2013).** Structural insights into the regulation of foreign genes in *Salmonella* by the Hha/H-NS Complex. *Journal of Biological Chemistry*, 288(19), 13356–13369.
- Ali, S.S., Soo, J., Rao, C., Leung, A.S., Ngai, D.H.M., Ensminger, A.W., & Navarre, W.W. (2014).** Silencing by H-NS potentiated the evolution of *Salmonella*. *PLoS Pathogens*, 10(11), e1004500.
- Amit, R., Oppenheim, A.B., & Stavans, J. (2003).** Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophysical Journal*, 84(4), 2467–2473.
- Aravind, L. & Landsman, D. (1998).** AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Research*, 26(19), 4413–4421.
- Arbibe, L., Kim, D.W., Batsche, E., Pedron, T., Mateescu, B., Muchardt, C., Parsot, C., & Sansonetti, P.J. (2007).** An injected bacterial effector targets chromatin access for transcription factor NF- $\kappa$ B to alter transcription of host genes involved in immune responses. *Nature Immunology*, 8(1), 47–56.
- Arikawa, K., & Nishikawa, Y. (2010).** Interleukin-8 induction due to diffusely adherent *Escherichia coli* possessing Afa/Dr genes depends on flagella and epithelial Toll-like receptor 5. *Microbiology and Immunology*, 54(9), 491–501.
- Arold, S.T., Leonard, P.G., Parkinson, G.N., & Ladbury, J.E. (2010).** H-NS forms a superhelical protein scaffold for DNA condensation. *Proceedings of the National Academy of Sciences*, 107(36), 15728–15732.
- Aslanidis, C., & de Jong, P.J. (1990).** Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Research*, 18(20), 6069–6074.
- Atlung, T., & Hansen, F.G. (2002).** Effect of different concentrations of H-NS protein on chromosome replication and the cell cycle in *Escherichia coli*. *Journal of Bacteriology*, 184(7), 1843–1850.
- Azam, T. A., & Ishihama, A. (1999).** Twelve species of the Nucleoid-Associated Protein from *Escherichia coli*. *Journal of Biological Chemistry*, 274(46), 33105–33113.
- Azam, T.A., Hiraga, S., & Ishihama, A. (2000).** Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes to Cells*, 5(8), 613–626.



## 7. BIBLIOGRAFÍA

- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B. L., & Mori, H. (2006).** Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, 2(1).
- Badaut, C., Williams, R., Arluison, V., Bouffartigues, E., Robert, B., Buc, H., & Rimsky, S. (2002).** The degree of oligomerization of the H-NS nucleoid structuring protein is related to specific binding to DNA. *Journal of Biological Chemistry*, 277(44), 41657–41666.
- Badea, L., Doughty, S., Nicholls, L., Sloan, J., Robins-Browne, R.M., & Hartland, E.L. (2003).** Contribution of Efa1/LifA to the adherence of enteropathogenic *Escherichia coli* to epithelial cells. *Microbial Pathogenesis*, 34(5), 205–215.
- Badrinarayanan, A., Le, T.B., & Laub, M.T. (2015).** Bacterial chromosome organization and segregation. *Annual review of cell and developmental biology*, 31, 171–199.
- Balows, A. (1992).** *The prokaryotes: a handbook on the biology of bacteria: ecophysiology, isolation, identification, applications*. Springer-Verlag.
- Balsalobre, C., Juarez, A., Madrid, C., Mourino, M., Prenafeta, A., & Munoa, F. J. (1996).** Complementation of the hha mutation in *Escherichia coli* by the ymoA gene from *Yersinia enterocolitica*: dependence on the gene dosage. *Microbiology*, 142(7), 1841–1846.
- Baños, R.C., Vivero, A., Aznar, S., García, J., Pons, M., Madrid, C., & Juárez, A. (2009).** Differential regulation of horizontally acquired and core genome genes by the bacterial modulator H-NS. *PLoS Genetics*, 5(6), e1000513.
- Barnich, N., Boudeau, J., Claret, L., & Darfeuille-Michaud, A. (2003).** Regulatory and functional co-operation of flagella and type 1 pili in adhesive and invasive abilities of AIEC strain LF82 isolated from a patient with Crohn's disease. *Molecular Microbiology*, 48(3), 781–794.
- Barnich, N., Carvalho, F.A., Glasser, A.L., Darcha, C., Jantscheff, P., Allez, M., Peeters, H., Bommelaer, G., Desreumaux, P., Colombel, J.F., & Darfeuille-Michaud, A. (2007).** CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. *Journal of Clinical Investigation*, 117(6), 1566–1574.
- Behrens, M., Sheikh, J., & Nataro, J.P. (2002).** Regulation of the overlapping pic/set locus in *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infection and Immunity*, 70(6), 2915–2925.
- Beloin, C., Deighan, P., Doyle, M., & Dorman, C. J. (2003).** *Shigella flexneri* 2a strain 2457T expresses three members of the H-NS-like protein family: characterization of the Sfh protein. *Molecular Genetics and Genomics*, 270(1), 66–77.
- Bennett, P.M. (2008).** Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *British Journal of Pharmacology*, 153(S1), S347–S357.
- Berg, D.E., & Howe, M.M. (1989).** *Mobile DNA*. American Society for Microbiology.

- Bergounioux, J., Elisee, R., Prunier, A.-L., Donnadiou, F., Sperandio, B., Sansonetti, P., & Arbibe, L. (2012).** Calpain activation by the *Shigella flexneri* effector VirA regulates key steps in the formation and life of the bacterium's epithelial niche. *Cell Host & Microbe*, *11*(3), 240–252.
- Bernier, C., Gounon, P., & Le Bouguenec, C. (2002).** Identification of an Aggregative Adhesion Fimbria (AAF) Type III-encoding operon in enteroaggregative *Escherichia coli* as a sensitive probe for detecting the AAF-encoding operon family. *Infection and Immunity*, *70*(8), 4302–4311.
- Bertelli, C., Laird, M.R., Williams, K.P., Lau, B. Y., Hoad, G., Winsor, G.L., & Brinkman, F.S. (2017).** IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research*, *45*(W1), W30–W35.
- Bertin, P., Hommais, F., Krin, E., Soutourina, O., Tendeng, C., Derzelle, S., & Danchin, A. (2001).** H-NS and H-NS-like proteins in Gram-negative bacteria and their multiple role in the regulation of bacterial metabolism. *Biochimie*, *83*(2), 235–241.
- Betis, F., Brest, P., Hofman, V., Guignot, J., Kansau, I., Rossi, B., Servin, A., & Hofman, P. (2003).** Afa/Dr diffusely adhering *Escherichia coli* infection in T84 cell monolayers induces increased neutrophil transepithelial migration, which in turn promotes cytokine-dependent upregulation of decay-accelerating factor (CD55), the receptor for Afa/Dr Adhesins. *Infection and Immunity*, *71*(4), 1774–1783.
- Bhat, A.P., Shin, M., & Choy, H.E. (2014).** Identification of high-specificity H-NS binding site in LEE5 promoter of enteropathogenic *Escherichia coli* (EPEC). *Journal of Microbiology*, *52*(7), 626–629.
- Blanch, A.R., García-Aljaro, C., Muniesa, M., & Jofre, J. (2003).** Detection, enumeration and isolation of strains carrying the stx2 gene from urban sewage. *Water Science and Technology*, *47*(3), 109–116.
- Blattner, F.R., Plunkett, G III., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y. (1997).** The complete genome sequence of *Escherichia coli* K-12. *Science*. 1997; *277*:1453–1474.
- Bloch, V., Yang, Y., Margeat, E., Chavanieu, A., Augé, M. T., Robert, B., Arold, S., Rimsky, S., & Kochoyan, M. (2003).** The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nature Structural & Molecular Biology*, *10*(3), 212–218.
- Boerlin, P., McEwen, S.A., Boerlin-Petzold, F., Wilson, J.B., Johnson, R.P., & Gyles, C.L. (1999).** Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. *Journal of clinical microbiology*, *37*(3), 497–503.
- Boisen, N., Struve, C., Scheutz, F., Krogfelt, K. A., & Nataro, J. P. (2008).** New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. *Infection and Immunity*, *76*(7), 3281–3292.

## 7. BIBLIOGRAFÍA

- Boisen, N., Melton-Celsa, A.R., Hansen, A.M., Zangari, T., Smith, M.A., Russo, L.M., Scheutz, F., O'Brien, A.D., & Nataro, J.P. (2019).** The role of the AggR regulon in the virulence of the Shiga toxin-producing enteroaggregative *Escherichia coli* epidemic O104:H4 strain in mice. *Frontiers in Microbiology*, *10*, 1824.
- Boudailliez, B., Berquin, P., Mariani-Kurkdjian, P., Ilef, D. D., Cuvelier, B., Capek, I., Tribout, B., Bingen, E., & Piussan, C. (1997).** Possible person-to-person transmission of *Escherichia coli* O111 - associated hemolytic uremic syndrome. *Pediatric Nephrology*, *11*(1), 36–39.
- Boudreau, B.A., Hron, D.R., Qin, L., van der Valk, R.A., Kotlajich, M.V, Dame, R.T., & Landick, R. (2018).** StpA and Hha stimulate pausing by RNA polymerase by promoting DNA–DNA bridging of H-NS filaments. *Nucleic Acids Research*, *46*(11), 5525–5546.
- Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., & Rimsky, S. (2007).** H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nature Structural & Molecular Biology*, *14*(5), 441–448.
- Bracco, L., Kotlarz, D., Kolb, A., Diekmann, S., & Buc H. (1989).** Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. *The EMBO Journal*, *8*(13), 4289–4296.
- Brescia, C. C. Mikulecky, P.J., Feig, A.L., & Sledjeski, D.D. (2003).** Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA*, *9*(1), 33–43.
- Bringer, M.A., Barnich, N., Glasser, A.L., Bardot, O., & Darfeuille-Michaud, A. (2005).** HtrA stress protein is involved in intramacrophagic replication of adherent and invasive *Escherichia coli* strain LF82 isolated from a patient with Crohn's disease. *Infection and Immunity*, *73*(2), 712–721.
- Brzuszkiewicz, E., Bruggemann, H., Liesegang, H., Emmerth, M., Olschlager, T., Nagy, G., Albermann, K., Wagner, C., Buchrieser, C., Emody, L., Gottschalk, G., Hacker, J., & Dobrindt, U. (2006).** How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proceedings of the National Academy of Sciences*, *103*(34), 12879–12884.
- Buchholz, U., Bernard, H., Werber, D., Böhmer, M. M., Remschmidt, C., Wilking, H., Deleré, Y., an der Heiden, M., Adlhoch, C., Dreesman, J., Ehlers, J., Ethelberg, S., Faber, M., Frank, C., Fricke, G., Greiner, M., Höhle, M., Ivarsson, S., Jark, U., & Kühne, M. (2011).** German outbreak of *Escherichia coli* O104:H4 associated with sprouts. *New England Journal of Medicine*, *365*(19), 1763–1770.
- Buckalew, D.W., Hartman, L.J., Grimsley, G.A., Martin, A.E., & Register, K.M. (2006).** A long-term study comparing membrane filtration with Colilert® defined substrates in detecting fecal coliforms and *Escherichia coli* in natural waters. *Journal of Environmental Management*, *80*(3), 191–197.

- Bullock, W.O., Fernandez, J.M., & Short, J.M. (1987).** XL1-Blue—a high-efficiency plasmid transforming *recA Escherichia coli* strain with  $\beta$ -galactosidase selection. *Biotechniques*, 5(3), 376-379.
- Carmona, M., Balsalobre, C., Muñoa, F., Mouriño, M., Jubete, Y., Cruz, F., & Juárez, A. (1993).** *Escherichia coli* hha mutants, DNA supercoiling and expression of the haemolysin genes from the recombinant plasmid pANN202-312. *Molecular Microbiology*, 9(5), 1011–1018.
- Castang, S., McManus, H.R., Turner, K.H., & Dove, S.L. (2008).** H-NS family members function coordinately in an opportunistic pathogen. *Proceedings of the National Academy of Sciences*, 105(48), 18947–18952.
- Castang, S., & Dove, S.L. (2010).** High-order oligomerization is required for the function of the H-NS family member MvaT in *Pseudomonas aeruginosa*. *Molecular Microbiology*, 78(4), 916–931.
- Chang, Y.Y., & Cronan, J.E. (1999).** Membrane cyclopropane fatty acid content is a major factor in acid resistance of *Escherichia coli*. *Molecular Microbiology*, 33(2), 249–259.
- Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., & Vogel, J. (2012).** An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *The EMBO Journal*, 31(20), 4005–4019.
- Chassaing, B., Rolhion, N., Vallée, A. de, Salim, S.Y., Prorok-Hamon, M., Neut, C., Campbell, B.J., Söderholm, J.D., Hugot, J.P., Colombel, J.F., & Darfeuille-Michaud, A. (2011).** Crohn disease–associated adherent-invasive *E. coli* bacteria target mouse and human Peyer's patches via long polar fimbriae. *Journal of Clinical Investigation*, 121(3), 966–975.
- Chaudhuri, R.R., Sebahia, M., Hobman, J.L., Webber, M.A., Leyton, D.L., Goldberg, M.D., Cunningham, A.F., Scott-Tucker, A., Ferguson, P.R., Thomas, C.M., Frankel, G., Tang, C.M., Dudley, E.G., Roberts, I.S., Rasko, D.A., Pallen, M.J., Parkhill, J., Nataro, J.P., Thomson, N. R., & Henderson, I.R. (2010).** Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* Strain 042. *PLoS ONE*, 5(1), e8801.
- Chen, S.L., Hung, C.S., Xu, J., Reigstad, C.S., Magrini, V., Sabo, A., Blasiar, D., Bieri, T., Meyer, R.R., Ozersky, P., Armstrong, J.R., Fulton, R.S., Latreille, J.P., Spieth, J., Hooton, T.M., Mardis, E.R., Hultgren, S.J., & Gordon, J.I. (2006).** Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: A comparative genomics approach. *Proceedings of the National Academy of Sciences*, 103(15), 5977–5982.
- Cherepanov, P.P., & Wackernagel, W. (1995).** Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, 158(1), 9–14.
- Claesson, M.J., O'Sullivan, O., Wang, Q., Nikkilä, J., Marchesi, J.R., Smidt, H., de Vos, W. M., Ross, R. P., & O'Toole, P. W. (2009).** Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE*, 4(8), e6669.

## 7. BIBLIOGRAFÍA

- Clark, D. (1981).** Regulation of fatty acid degradation in *Escherichia coli*: analysis by operon fusion. *Journal of Bacteriology*, 148(2), 521-526.
- Clavijo, A.P., Bai, J., & Gómez-Duarte, O.G. (2010).** The longus type IV pilus of enterotoxigenic *Escherichia coli* (ETEC) mediates bacterial self-aggregation and protection from antimicrobial agents. *Microbial Pathogenesis*, 48(6), 230-238.
- Clermont, O., Gordon, D.M., Brisse, S., Walk, S.T., & Denamur, E. (2011).** Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. *Environmental Microbiology*, 13(9), 2468-2477.
- Cohen, S.N., Chang, A.C.Y., & Hsu, L. (1972).** Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-Factor DNA. *Proceedings of the National Academy of Sciences*, 69(8), 2110-2114.
- Conlan, S., Thomas, P.J., Deming, C., Park, M., Lau, A.F., Dekker, J.P., Snitkin, E.S., Clark, T.A., Luong, K., Song, Y., Tsai, Y.C., Boitano, M., Dayal, J., Brooks, S.Y., Schmidt, B., Young, A.C., Thomas, J.W., Bouffard, G.G., Blakesley, R.W., Segre, J.A. (2014).** Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing *Enterobacteriaceae*. *Science Translational Medicine*, 6(254), 254ra126-254ra126.
- Cornells, G.R., Sluiter, C., Delor, I., Geib, D., Kaniga, K., Rouvroit, C.L., Sory, M.P., Vanooteghem, J.C., & Michiels, T. (1991).** ymoA, a *Yersinia enterocolitica* chromosomal gene modulating the expression of virulence functions. *Molecular Microbiology*, 5(5), 1023-1034.
- Corno, G., Coci, M., Giardina, M., Plechuk, S., Campanile, F., & Stefani, S. (2014).** Antibiotics promote aggregation within aquatic bacterial communities. *Frontiers in Microbiology*, 5.
- Craig, N.L., Craigie, R., Gellert, M. & Lambowitz, A.M. (2002).** *Mobile DNA II*. ASM Press.
- Crick, F.H.C., Barnett, L., Brenner, S., & Watts-Tobin, R.J. (1961).** General nature of the genetic code for proteins. *Nature*, 192(4809), 1227-1232.
- Croxen, M.A. (2010).** Molecular mechanisms of *Escherichia coli* pathogenicity. *Nature Reviews. Microbiology*, 8(1).
- Cui, T., & Leng, F. (2007).** Specific recognition of AT-Rich DNA sequences by the mammalian High Mobility Group Protein AT-hook 2: a SELEX study. *Biochemistry*, 46(45), 13059-13066.
- Cukier-Kahn, R., Jacquet, M., & Gros, F. (1972).** Two heat-resistant, low molecular weight proteins from *Escherichia coli* that stimulate DNA-directed RNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 69(12).
- Cunningham, T.P., Montelaro, R.C., & Rushlow, K.E. (1993).** Lentivirus envelope sequences and proviral genomes are stabilized in *Escherichia coli* when cloned in low-copy-number plasmid vectors. *Gene*, 124(1), 93-98.



- Cusick, M.E., & Belfort, M. (2002).** Domain structure and RNA annealing activity of the *Escherichia coli* regulatory protein StpA. *Molecular Microbiology*, 28(4), 847–857.
- Czczulin, J.R., Balepur, S., Hicks, S., Phillips, A., Hall, R., Kothary, M.H., Navarro-Garcia, F., & Nataro, J.P. (1997).** Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infection and immunity*, 65(10), 4135–4145.
- Czczulin J.R., Whittam T.S., Henderson I.R., Navarro-Garcia F., Nataro J.P. (1999).** Phylogenetic analysis of enteroaggregative and diffusely adherent *Escherichia coli*. *Infection and Immunity* 67, 2692–2699.
- D'Heygère, F., Rabhi, M., & Boudvillain, M. (2013).** Phyletic distribution and conservation of the bacterial transcription termination factor Rho. *Microbiology*, 159 (Pt\_7), 1423–1436.
- Dallman, T., Smith, G.P., O'Brien, B., Chattaway, M.A., Finlay, D., Grant, K.A., & Jenkins, C. (2012).** Characterization of a verocytotoxin-producing enteroaggregative *Escherichia coli* serogroup O111:H21 strain associated with a household outbreak in Northern Ireland. *Journal of clinical microbiology*, 50(12), 4116–4119.
- Dalton, C.B., Mintz, E.D., Wells, J.G., Bopp, C.A., & Tauxe, R.V. (1999).** Outbreaks of enterotoxigenic *Escherichia coli* infection in American adults: a clinical and epidemiologic profile. *Epidemiology and Infection*, 123(1), 9–16.
- Dame, R.T., Wyman, C., & Goosen, N. (2000).** H-NS mediated compaction of DNA visualized by atomic force microscopy. *Nucleic Acids Research*, 28(18), 3504-3510.
- Dame, R.T. (2005).** The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Molecular microbiology* 56: 858–870.
- Dame, R.T., Luijsterburg, M.S., Krin, E., Bertin, P.N., Wagner, R., & Wuite, G.J.L. (2005).** DNA Bridging: a property shared among H-NS-like proteins. *Journal of Bacteriology*, 187(5), 1845–1848.
- Danese, P.N., Pratt, L.A., Dove, S.L., & Kolter, R. (2000).** The outer membrane protein, Antigen 43, mediates cell-to-cell interactions within *Escherichia coli* biofilms. *Molecular Microbiology*, 37(2), 424–432.
- Datsenko, K.A., & Wanner, B.L. (2000).** One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12), 6640–6645.
- Daubin, V., Lerat, E. & Perrière, G. (2003).** The source of laterally transferred genes in bacterial genomes. *Genome Biology* 4, R57.
- de Alba, C.F., Solórzano, C., Paytubi, S., Madrid, C., Juárez, A., García, J., & Pons, M. (2011).** Essential residues in the H-NS binding site of Hha, a co-regulator of horizontally acquired genes in enterobacteria. *FEBS Letters*, 585(12), 1765–1770.
- De Biase, D., Tramonti, A., Bossa, F., & Visca, P. (1999).** The response to stationary-phase stress conditions in *Escherichia coli*: role and regulation of the glutamic acid decarboxylase system. *Molecular Microbiology*, 32(6), 1198–1211.

## 7. BIBLIOGRAFÍA

- De la Cruz, F., Carmona, M., & Juárez, A. (1992).** The Hha protein from *Escherichia coli* is highly homologous to the YmoA protein from *Yersinia enterocolitica*. *Molecular Microbiology*, 6(22), 3451–3452.
- De Muinck, E.J., Lagesen, K., Afset, J.E., Didelot, X., Rønningen, K. S., Rudi, K., Stenseth, N.C., & Trosvik, P. (2013).** Comparisons of infant *Escherichia coli* isolates link genomic profiles with adaptation to the ecological niche. *BMC Genomics*, 14(1), 81.
- Dean, P., Maresca, M., Schller, S., Phillips, A.D., & Kenny, B. (2006).** Potent diarrheagenic mechanism mediated by the cooperative action of three enteropathogenic *Escherichia coli*-injected effector proteins. *Proceedings of the National Academy of Sciences*, 103(6), 1876–1881.
- Dean, P., & Kenny, B. (2009).** The effector repertoire of enteropathogenic *E. coli*: ganging up on the host cell. *Current Opinion in Microbiology*, 12(1), 101–109.
- Dean, P., Scott, J.A., Knox, A.A., Quitard, S., Watkins, N.J., & Kenny, B. (2010).** The enteropathogenic *E. coli* effector EspF targets and disrupts the nucleolus by a process regulated by mitochondrial dysfunction. *PLoS pathogens*, 6(6), e1000961.
- Decker, K.B., & Hinton, D.M. (2013).** Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annual Review of Microbiology*, 67(1), 113–139.
- Deighan, P., Free, A., & Dorman, C.J. (2000).** A role for the *Escherichia coli* H-NS-like protein StpA in OmpF porin expression through modulation of micF RNA stability. *Molecular Microbiology*, 38(1), 126–139.
- Deneke, C., Lipowsky, R., & Valleriani, A. (2013).** Effect of ribosome shielding on mRNA stability. *Physical Biology*, 10(4), 046008.
- Deng, W., Puente, J. L., Gruenheid, S., Li, Y., Vallance, B.A., Vazquez, A., Barba, J., Ibarra, J.A., O'Donnell, P., Metalnikov, P., Ashman, K., Lee, S., Goode, D., Pawson, T., & Finlay, B.B. (2004).** Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proceedings of the National Academy of Sciences*, 101(10), 3597–3602.
- Di Salvo, M., Puccio, S., Peano, C., Lacour, S., & Alifano, P. (2019).** RhoTermPredict: an algorithm for predicting Rho-dependent transcription terminators based on *Escherichia coli*, *Bacillus subtilis* and *Salmonella enterica* databases. *BMC Bioinformatics*, 20(1), 117.
- Dillon, S.C., & Dorman, C.J. (2010).** Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature reviews Microbiology* 8: 185–95.
- Dole, S., Klingen, Y., Nagarajavel, V., & Schnetz, K. (2004).** The protease Lon and the RNA-binding protein Hfq reduce silencing of the *Escherichia coli* bgl operon by H-NS. *Journal of Bacteriology*, 186(9), 2708–2716.
- Dorman, C.J., Hinton, J.C., & Free, A. (1999).** Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends in Microbiology* 7: 124–128.



- Dorman, C.J. (2004).** H-NS: a universal regulator for a dynamic genome. *Nature reviews Microbiology* 2: 391–400.
- Dorman, C.J. (2007).** H-NS, the genome sentinel. *Nature Reviews Microbiology*, 5(2), 157–161.
- Dorman, C.J. (2014).** H-NS-like nucleoid-associated proteins, mobile genetic elements and horizontal gene transfer in bacteria. *Plasmid* 75C: 1–11.
- Dower, W.J., Miller, J.F., & Ragsdale, C.W. (1988).** High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Research*, 16(13), 6127–6145.
- Drago-Serrano, M.E, Gavilanes Parra, S., & Angel Manjarrez-Hernández, H. (2006).** EspC, an autotransporter protein secreted by enteropathogenic *Escherichia coli* (EPEC), displays protease activity on human hemoglobin. *FEMS Microbiology Letters*, 265(1), 35–40.
- Dudley, E.G., Thomson, N.R., Parkhill, J., Morin, N.P., & Nataro, J.P. (2006).** Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative *Escherichia coli*. *Molecular Microbiology*, 61(5), 1267–1282.
- Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K.E., & Relman, D.A. (2005).** Diversity of the human intestinal microbial flora. *Science* (New York, N.Y.), 308(5728), 1635–1638.
- Egile, C., Loisel, T.P., Laurent, V., Li, R., Pantaloni, D., Sansonetti, P.J., & Carlier, M.F. (1999).** Activation of the Cdc42 effector N-Wasp by the *Shigella flexneri* Icsa protein promotes actin nucleation by Arp2/3 complex and bacterial actin-based motility. *The Journal of Cell Biology*, 146(6), 1319–1332.
- Elias, W.P., Jr, Czczulin, J.R., Henderson, I.R., Trabulsi, L.R., & Nataro, J.P. (1999).** Organization of biogenesis genes for aggregative adherence fimbria II defines a virulence gene cluster in enteroaggregative *Escherichia coli*. *Journal of bacteriology*, 181(6), 1779–1785.
- Ellermeier, C.D., Janakiraman, A., & Slauch, J.M. (2002).** Construction of targeted single copy lac fusions using lambda Red and FLP-mediated site-specific recombination in bacteria. *Gene*, 290(1–2), 153–61.
- Ellison, D.W., Young, B., Nelson, K., & Miller, V.L. (2003).** YmoA negatively regulates expression of invasin from *Yersinia enterocolitica*. *Journal of Bacteriology*, 185(24), 7153–7159.
- Ellison, D.W., & Miller, V.L. (2006).** Regulation of virulence by members of the MarR/SlyA family. *Current Opinion in Microbiology*, 9(2), 153–159.
- Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, T., & Igarashi, K. (1988).** Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. *European Journal of Biochemistry*, 171(1–2), 45–50.

## 7. BIBLIOGRAFÍA

- Escherich, T. & Bettelheim K.S. (1988).** The intestinal bacteria of the neonate and breast-fed infant. *Reviews of Infectious Diseases*, 10(6).
- Eslava, C., Navarro-García, F., Czeczulin, J.R., Henderson, I.R., Cravioto, A., & Nataro, J.P. (1998).** Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infection and immunity*, 66(7), 3155–3163.
- Etienne-Mesmin, L., Chassaing, B., Sauvanet, P., Denizot, J., Blanquet-Diot, S., Darfeuille-Michaud, A., Pradel, N., & Livrelli, V. (2011).** Interactions with M cells and macrophages as key steps in the pathogenesis of enterohemorrhagic *Escherichia coli* infections. *PLoS one*, 6(8), e23594.
- Faherty, C., Harper, J.M., Shea-Donohue, T., Barry, E.M., Kaper, J.B., Fasano, A., & Nataro, J.P. (2012).** Chromosomal and plasmid-encoded factors of *Shigella flexneri* induce secretogenic activity *ex vivo*. *PLoS ONE*, 7(11), e49980.
- Fahlen, T.F., Wilson, R.L., Boddicker, J.D., & Jones, B.D. (2001).** Hha is a negative modulator of transcription of hilA, the *Salmonella enterica* serovar Typhimurium invasion gene transcriptional activator. *Journal of Bacteriology*, 183(22), 6620–6629.
- Falconi, M., Higgins, N.P., Spurio, R., Pon, C.L., & Gualerzi, C.O. (1993).** Expression of the gene encoding the major bacterial nucleoid protein H-NS is subject to transcriptional auto-repression. *Molecular Microbiology*, 10(2), 273–282.
- Falconi, M., Brandi, A., La Teana, A., Gualerzi, C.O., & Pon, C.L. (1996).** Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of *hns* expression. *Molecular Microbiology*, 19(5), 965–975.
- Falconi, M., Colonna, B., Prosseda, G., Micheli, G., & Gualerzi, C.O. (1998).** Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of virF promoter to transcriptional repressor H-NS. *The EMBO Journal*, 17(23), 7033–7043.
- Farfan, M.J., Cantero, L., Vidal, R., Botkin, D.J., & Torres, A.G. (2011).** Long Polar Fimbriae of enterohemorrhagic *Escherichia coli* O157:H7 bind to extracellular matrix proteins. *Infection and Immunity*, 79(9), 3744–3750.
- Fasano, A., Noriega, F.R., Maneval, D.R., Chanasongcram, S., Russell, R., Guandalini, S., & Levine, M.M. (1995).** *Shigella* enterotoxin 1: an enterotoxin of *Shigella flexneri* 2a active in rabbit small intestine in vivo and in vitro. *Journal of Clinical Investigation*, 95(6), 2853–2861.
- Fasano, A., Noriega, F.R., Liao, F.M., Wang, W., & Levine, M.M. (1997).** Effect of *Shigella* enterotoxin 1 (ShET1) on rabbit intestine in vitro and in vivo. *Gut*, 40(4), 505–511.
- Feng, P.C.H., Jinneman, K., Scheutz, F., & Monday, S.R. (2011).** Specificity of PCR and serological assays in the detection of *Escherichia coli* Shiga toxin subtypes. *Applied and Environmental Microbiology*, 77(18), 6699–6702.

- Fernandez-Prada, C.M., Hoover, D.L., Tall, B.D., Hartman, A.B., Kopelowitz, J., & Venkatesan, M.M. (2000).** *Shigella flexneri* IpaH7.8 facilitates escape of virulent bacteria from the endocytic vacuoles of mouse and human macrophages. *Infection and Immunity*, 68(6), 3608–3619.
- Forns, N., Banos, R.C., Balsalobre, C., Juarez, A., & Madrid, C. (2005).** Temperature-dependent conjugative transfer of R27: role of chromosome- and plasmid-encoded Hha and H-NS proteins. *Journal of Bacteriology*, 187(12), 3950–3959.
- Fournier, B., Lu, C.Y., Lagrange, P.H., Krishnamoorthy, R., & Philippon, A. (1995).** Point mutation in the Pribnow box, the molecular basis of beta-lactamase overproduction in *Klebsiella oxytoca*. *Antimicrobial Agents and Chemotherapy*, 39(6), 1365–1368.
- Free, A., & Dorman, C.J. (1995).** Coupling of *Escherichia coli* hns mRNA levels to DNA synthesis by autoregulation: implications for growth phase control. *Molecular Microbiology*, 18(1), 101–113.
- Free, A., & Dorman, C.J. (1997).** The *Escherichia coli* stpA gene is transiently expressed during growth in rich medium and is induced in minimal medium and by stress conditions. *Journal of Bacteriology*, 179(3), 909–918.
- Free, A., Williams, R.M., & Dorman, C.J. (1998).** The StpA protein functions as a molecular adapter to mediate repression of the bgl operon by truncated H-NS in *Escherichia coli*. *Journal of bacteriology* 180: 994–7
- Friedrich, A.W., Bielaszewska, M., Zhang, W., Pulz, M., Kuczius, T., Ammon, A., & Karch, H. (2002).** *Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *The Journal of Infectious Diseases*, 185(1), 74–84.
- Fujiyama, R., Nishi, J., Imuta, N., Tokuda, K., Manago, K., & Kawano, Y. (2008).** The shf gene of a *Shigella flexneri* homologue on the virulent plasmid pAA2 of enteroaggregative *Escherichia coli* 042 is required for firm biofilm formation. *Current Microbiology*, 56(5), 474–480.
- Furukawa, T., Yahiro, K., Tsuji, A. B., Terasaki, Y., Morinaga, N., Miyazaki, M., Fukuda, Y., Saga, T., Moss, J., & Noda, M. (2011).** Fatal hemorrhage induced by subtilase cytotoxin from Shiga-toxigenic *Escherichia coli*. *Microbial Pathogenesis*, 50(3–4), 159–167.
- Gaastra, W. & Svennerholm, A.M. (1996).** Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends in Microbioly.* 4:444–452.
- Galdiero, F., Carratelli, C.R., Nuzzo, I., Bentivoglio, C., & Galdiero, M. (1988).** Phagocytosis of bacterial aggregates by granulocytes. *European Journal of Epidemiology*, 4(4), 456–460.
- Gallegos, M.T., Schleif, R., Bairoch, A., Hofmann, K., & Ramos, J.L. (1997).** Arac/XylS family of transcriptional regulators. *Microbiology and molecular biology reviews: MMBR*, 61(4), 393–410.
- Gal-Mor, O., & Finlay, B. B. (2006).** Pathogenicity islands: a molecular toolbox for bacterial virulence. *Cellular Microbiology*, 8(11), 1707–1719.

## 7. BIBLIOGRAFÍA

**García, J., Madrid, C., Juárez, A., & Pons, M. (2006).** New roles for key residues in helices H1 and H2 of the *Escherichia coli* H-NS N-terminal Domain: H-NS dimer stabilization and Hha binding. *Journal of Molecular Biology*, 359(3), 679–689.

**García-Contreras, R., Zhang, X.S., Kim, Y., & Wood, T.K. (2008).** Protein translation and cell death: the role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. *PLoS one*, 3(6), e2394.

**Garmendia, J., Frankel, G., & Crepin, V.F. (2005).** Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation. *Infection and immunity*, 73(5), 2573–2585.

**Gaytán, M.O., Martínez-Santos, V.I., Soto, E., & González-Pedrajo, B. (2016).** Type Three Secretion System in attaching and effacing pathogens. *Frontiers in Cellular and Infection Microbiology*, 6.

**Gerdes, K., Bech, F.W., Jørgensen, S. T., Løbner-Olesen, A., Rasmussen, P. B., Atlung, T., Boe, L., Karlstrom, O., Molin, S., & von Meyenburg, K. (1986).** Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the *E. coli* relB operon. *The EMBO journal*, 5(8), 2023–2029.

**Geuskens, V., Mhammedi-Alaoui, A., Desmet, L. & Toussaint, A. (1992).** Virulence in bacteriophage Mu: a case of trans-dominant proteolysis by the *Escherichia coli* Clp serine protease. *The EMBO Journal*, 11(13), 5121–5127.

**Ghosal, A., Bhowmick, R., Banerjee, R., Ganguly, S., Yamasaki, S., Ramamurthy, T., Hamabata, T., & Chatterjee, N.S. (2009).** Characterization and studies of the cellular interaction of Native Colonization Factor CS6 Purified from a clinical isolate of enterotoxigenic *Escherichia coli*. *Infection and Immunity*, 77(5), 2125–2135.

**Giangrossi, M., Wintraecken, K., Spurio, R., & de Vries, R. (2014).** Probing the relation between protein–protein interactions and DNA binding for a linker mutant of the bacterial nucleoid protein H-NS. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(2), 339–345.

**Gill, R.K., Borthakur, A., Hodges, K., Turner, J.R., Clayburgh, D.R., Saksena, S., Zaheer, A., Ramaswamy, K., Hecht, G., & Dudeja, P.K. (2007).** Mechanism underlying inhibition of intestinal apical Cl<sup>-</sup>/OH<sup>-</sup> exchange following infection with enteropathogenic *E. coli*. *Journal of Clinical Investigation*, 117(2), 428–437.

**Girón, J. A., Torres, A. G., Freer, E., & Kaper, J. B. (2002).** The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. *Molecular Microbiology*, 44(2), 361–379.

**Glasser, A.L., Boudeau, J., Barnich, N., Perruchot, M.H., Colombel, J.F., & Darfeuille-Michaud, A. (2001).** Adherent invasive *Escherichia coli* strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death. *Infection and Immunity*, 69(9), 5529–5537.

- Godessart, N., Muñoa, F.J., Regue, M., & Juárez, A. (1988).** Chromosomal mutations that increase the production of a plasmid-encoded haemolysin in *Escherichia coli*. *Journal of general Microbiology* 134: 2779–2787.
- Gogoi, M., Datey, A., Wilson, K.T., & Chakravorty, D. (2016).** Dual role of arginine metabolism in establishing pathogenesis. *Current Opinion in Microbiology*, 29, 43–48.
- Goldberg, M. B., & Theriot, J. A. (1995).** *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proceedings of the National Academy of Sciences*, 92(14), 6572–6576.
- Goluszko, P., Selvarangan, R., Popov, V., Pham, T., Wen, J.W., & Singhal, J. (1999).** Decay-accelerating factor and cytoskeleton redistribution pattern in HeLa cells infected with recombinant *Escherichia coli* strains expressing Dr family of adhesins. *Infection and Immunity*, 67(8), 3989–3997.
- Gómez-Duarte, O.G., Ruiz-Tagle A., Gómez D.C., Viboud, G.I., Jarvis, K.G., Kaper, J.B., & Girón, J.A. (1999).** Identification of IngA, the structural gene of longus type IV pilus of enterotoxigenic *Escherichia coli*. *Microbiology-SGM*, 145, (Pt):1809–1816.
- Göransson, M., Sondén, B., Nilsson, P., Dagberg, B., Foreman, K., Emanuelsson, K., & Uhlin, B. E. (1990).** Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature*, 344(6267), 682–685.
- Gordon, G.R., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., Navarre, W.W., Xia, B., & Liu, J. (2010).** Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America*, 107(11).
- Gordon, B.R., Li, Y., Cote, A., Weirauch, M.T., Ding, P., Hughes, T.R., Navarre, W.W., Xia, B., & Liu, J. (2011).** Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proceedings of the National Academy of Sciences*, 108(26), 10690–10695.
- Goyard, S., & Bertin, P. (1997).** Characterization of BpH3, an H-NS-like protein in *Bordetella pertussis*. *Molecular Microbiology*, 24(4), 815–823.
- Grainger, D.C., Hurd, D., Goldberg, M. D., & Busby, S.J.W. (2006).** Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Research*, 34(16), 4642–4652.
- Grainger, D.C. (2016).** Structure and function of bacterial H-NS protein. *Biochemical Society Transactions*, 44(6), 1561–1569.
- Grys, T.E., Siegel, M.B., Lathem, W.W., & Welch, R.A. (2005).** The StcE protease contributes to intimate adherence of enterohemorrhagic *Escherichia coli* O157:H7 to host cells. *Infection and Immunity*, 73(3), 1295–1303.
- Guignot, J., Segura, A., & Tran Van Nhieu, G. (2015).** The serine protease EspC from enteropathogenic *Escherichia coli* regulates pore formation and cytotoxicity mediated by the Type III secretion system. *PLoS pathogens*, 11(7), e1005013.



## 7. BIBLIOGRAFÍA

- Guttman, J.A., Li, Y., Wickham, M.E., Deng, W., Vogl, A.W., & Finlay, B.B. (2006).** Attaching and effacing pathogen-induced tight junction disruption in vivo. *Cellular Microbiology*, 8(4), 634–645.
- Guttman, J.A., Samji, F.N., Li, Y., Deng, W., Lin, A., & Finlay, B.B. (2007).** Aquaporins contribute to diarrhoea caused by attaching and effacing bacterial pathogens. *Cellular Microbiology*, 9(1), 131–141.
- Guyer, M.S., Reed, R.R., Steitz, J.A., & Low, K.B. (1981).** Identification of a sex-factor-affinity site in *E. coli* as  $\gamma\delta$ . *Cold Spring Harbor Symposium on Quantitative Biology*, 45, 135-140
- Handa, Y., Suzuki, M., Ohya, K., Iwai, H., Ishijima, N., Koleske, A. J., Fukui, Y., & Sasakawa, C. (2007).** *Shigella* IpgB1 promotes bacterial entry through the ELMO–Dock180 machinery. *Nature Cell Biology*, 9(1), 121–128.
- Hardwidge, P.R., Deng, W., Vallance, B.A., Rodriguez-Escudero, I., Cid, V.J., Molina, M., & Finlay, B.B. (2005).** Modulation of host cytoskeleton function by the enteropathogenic *Escherichia coli* and *Citrobacter rodentium* effector protein EspG. *Infection and immunity*, 73(5), 2586–2594.
- Harms, A., Maisonneuve, E., & Gerdes, K. (2016).** Mechanisms of bacterial persistence during stress and antibiotic exposure. *Science*, 354(6318), aaf4268.
- Harms, A., Fino, C., Sørensen, M. A., Semsey, S., & Gerdes, K. (2017).** Prophages and growth dynamics confound experimental results with antibiotic-tolerant persister cells. *MBio*, 8(6).
- Harper, J., Kreisberg, R.B., Clements, J.D., Strauman, M.C., Nataro, J.P., & Marohn, M. (2011).** Induction of increased permeability of polarized enterocyte monolayers by enterotoxigenic *Escherichia coli* heat-labile enterotoxin. *The American Journal of Tropical Medicine and Hygiene*, 84(3), 451–455.
- Harrington, A., Darboe, N., Kenjale, R., Picking, W.L., Middaugh, C.R., Birket, S., & Picking, W.D. (2006).** Characterization of the interaction of single tryptophan containing mutants of IpaC from *Shigella flexneri* with phospholipid membranes. *Biochemistry*, 45(2), 626–636.
- Harrington, S.M., Dudley, E.G., & Nataro, J.P. (2006).** Pathogenesis of enteroaggregative *Escherichia coli* infection. *FEMS Microbiology Letters*, 254(1).
- Harrison, E., & Brockhurst, M.A. (2012).** Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in Microbiology*, 20(6), 262–267.
- Hartl, D.L., & Dykhuizen, D.E. (1984).** The population genetics of *Escherichia coli*. *Annual Review of Genetics* 18:1, 31-68.
- Hebbelstrup Jensen, B., Olsen, K.E.P., Struve, C., Kroghfelt, K.A., & Petersen, A M. (2014).** Epidemiology and clinical manifestations of enteroaggregative *Escherichia coli*. *Clinical Microbiology Reviews*, 27(3), 614–630.

- Henderson, I.R., Hick, S., Navarro-Garcia, F., Elias, W.P., Philips, A.D., & Nataro, J.P. (1999).** Involvement of the enteroaggregative *Escherichia coli* plasmid-encoded toxin in causing human intestinal damage. *Infection and Immunity*, 67(10).
- Henzel, W.J., Billeci, T.M., Stults J.T., Wong, S.C., Grimley, C., & Watanabe, C. (1993).** Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proceedings of the National Academy of Sciences of the United States of America*, 90(11), 5011–5015.
- High, N., Mounier, J., Prévost, M.C., Sansonetti, P.J. (1992).** IpaB of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole. *EMBO Journal* 11, 1991–1999.
- Hilbi, H., Moss, J.E., Hersh, D., Chen, Y., Arondel, J., Banerjee, S., Flavell, R.A., Yuan, J., Sansonetti, P.J., & Zychlinsky, A. (1998).** *Shigella*-induced apoptosis is dependent on caspase-1 which binds to IpaB. *Journal of Biological Chemistry*, 273(49), 32895–32900.
- Hildebrand, A., Schlacta, T., Warmack, R., Kasuga, T., & Fan, Z. (2013).** Engineering *Escherichia coli* for improved ethanol production from gluconate. *Journal of Biotechnology*, 168(1), 101–106.
- Hinton, J.C.D., Santos, D.S., Seirafi, A., Hulton, C.S.J., Pavitt, G.D., & Higgins, C.F. (1992).** Expression and mutational analysis of the nucleoid-associated protein H-NS of *Salmonella typhimurium*. *Molecular Microbiology*, 6(16), 2327–2337.
- Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J.P., Danchin, A., & Bertin, P. (2001).** Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Molecular Microbiology*, 40(1), 20–36.
- Huang, D.B. (2006).** A review of an emerging enteric pathogen: enteroaggregative *Escherichia coli*. *Journal of Medical Microbiology*, 55(10), 1303–1311.
- Huang, D.B., Nataro, J.P., DuPont, H.L., Kamat, P.P., Mhatre, A.D., Okhuysen, P.C., & Chiang, T. (2006).** Enteroaggregative *Escherichia coli* is a cause of acute diarrheal illness: a meta-analysis. *Clinical Infectious Diseases*, 43(5), 556–563.
- Hudault, S., Spiller, O.B., Morgan, B.P., & Servin, A.L. (2004).** Human diffusely adhering *Escherichia coli* expressing Afa/Dr adhesins that use human CD55 (decay-accelerating factor) as a receptor does not bind the rodent and pig analogues of CD55. *Infection and immunity*, 72(8), 4859–4863.
- Humphries, A., DeRidder, S., & Baumler, A.J. (2005).** *Salmonella* enterica Serotype Typhimurium fimbrial proteins serve as antigens during infection of mice. *Infection and Immunity*, 73(9), 5329–5338.
- Hüttener, M., Dietrich, M., Paytubi, S., & Juárez, A. (2014).** HilA-like regulators in *Escherichia coli* pathotypes: the YgeH protein from the enteroaggregative strain 042. *BMC Microbiology*, 14(1), 268.



## 7. BIBLIOGRAFÍA

- Hüttener, M., Prieto, A., Espelt, J., Bernabeu, M., & Juárez, A. (2018).** Stringent response and AggR-dependent virulence regulation in the enteroaggregative *Escherichia coli* strain 042. *Frontiers in microbiology*, 9, 717.
- Ibarra, J.A., Villalba, M.I., & Puente, J.L. (2003).** Identification of the DNA binding sites of PerA, the transcriptional activator of the bfp and per operons in enteropathogenic *Escherichia coli*. *Journal of Bacteriology*, 185(9), 2835–2847.
- Iguchi, A., Thomson, N.R., Ogura, Y., Saunders, D., Ooka, T., Henderson, I.R., Harris, D., Asadulghani, M., Kurokawa, K., Dean, P., Kenny, B., Quail, M A., Thurston, S., Dougan, G., Hayashi, T., Parkhill, J., & Frankel, G. (2009).** Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *Journal of Bacteriology*, 191(1), 347–354.
- Iwai, H., Kim, M., Yoshikawa, Y., Ashida, H., Ogawa, M., Fujita, Y., Muller, D., Kirikae, T., Jackson, P.K., Kotani, S., & Sasakawa, C. (2007).** A bacterial effector targets Mad2L2, an APC inhibitor, to modulate host cell cycling. *Cell*, 130(4), 611–623.
- Iyoda, S., Tamura, K., Itoh, K., Izumiya, H., Ueno, N., Nagata, K., Togo, M., Terajima, J., & Watanabe, H. (2000).** Inducible stx2 phages are lysogenized in the enteroaggregative and other phenotypic *Escherichia coli* O86:HNM isolated from patients. *FEMS Microbiology Letters*, 191(1), 7–10.
- Jacob, F., & Monod, J. (1961).** Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3, 318-356.
- Jacquet, M., Cukier-Kahn, R., Pla, J., & Gros, F. (1971).** A thermostable protein factor acting on in vitro DNA transcription. *Biochemical and Biophysical Research Communications*, 45(6), 1597–1607.
- Jansson, L., Tobias, J., Lebens, M., Svennerholm, A.M., & Teneberg, S. (2006).** The major subunit, CfaB, of Colonization Factor Antigen I from enterotoxigenic *Escherichia coli* is a glycosphingolipid binding protein. *Infection and Immunity*, 74(6), 3488–3497.
- Jiang, Z.D., Greenberg, D., Nataro, J.P., Steffen, R., & DuPont, H.L. (2002).** Rate of occurrence and pathogenic effect of enteroaggregative *Escherichia coli* virulence factors in international travelers. *Journal of clinical microbiology*, 40(11), 4185–4190.
- Jimenez-Diaz, L., Caballero, A., & Segura, A. (2017).** Pathways for the degradation of fatty acids in bacteria. In: Rojo F. (eds) *Aerobic utilization of hydrocarbons, Oils and Lipids. Handbook of Hydrocarbon and Lipid Microbiology. Springer International Publishing*, 291-313.
- Jin, D.J., Burgess, R.R., Richardson, J.P., & Gross, C.A. (1992).** Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proceedings of the National Academy of Sciences of the United States of America*, 89(4).
- Johansson, J., & Uhlin, B.E. (1999).** Differential protease-mediated turnover of H-NS and StpA revealed by a mutation altering protein stability and stationary-phase survival of *Escherichia coli*. *Proceedings of the National Academy of Sciences* 96,10776–10781.

- Johansson, J., Eriksson, S., Sondén, B., Wai, N., Uhlin, B.E., Johansson, R., & Sonde, B. (2001).** Heteromeric interactions among nucleoid-associated bacterial proteins: localization of StpA-stabilizing regions in H-NS of *Escherichia coli*. *Journal of bacteriology* 183, 2343–2347.
- Johnson, T.J., Kariyawasam, S., Wannemuehler, Y., Mangiamale, P., Johnson, S.J., Doetkott, C., Skyberg, J.A., Lynne, A.M., Johnson, J.R., & Nolan, L.K. (2007).** The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. *Journal of bacteriology*, 189(8), 3228–3236.
- Johnson, J.R., Porter, S.B., Zhanel, G., Kuskowski, M.A., & Denamur, E. (2012).** Virulence of *Escherichia coli* clinical isolates in a murine sepsis model in relation to sequence type ST131 status, fluoroquinolone resistance, and virulence genotype. *Infection and Immunity*, 80(4), 1554–1562.
- Jønsson, R., Struve, C., Boll, E.J., Boisen, N., Joensen, K.G., Sørensen, C.A., Jensen, B.H., Scheutz, F., Jenssen, H., & Krogfelt, K.A. (2017).** A novel pAA virulence plasmid encoding toxins and two distinct variants of the fimbriae of enteroaggregative *Escherichia coli*. *Frontiers in Microbiology*, 8, 263.
- Kaas, R.S., Friis, C., Ussery, D.W., & Aarestrup, F.M. (2012).** Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes. *BMC Genomics*, 13(1), 577
- Kamionka, M. (2011).** Engineering of therapeutic proteins production in *Escherichia coli*. *Current Pharmaceutical Biotechnology*, 12(2), 268–274.
- Kaper, J.B., Nataro, J.P., & Mobley, H.L. (2004).** Pathogenic *Escherichia coli*. *Nature Reviews Microbiology* 2: 123-40.
- Karch, H., Denamur, E., Dobrindt, U., Finlay, B.B., Hengge, R., Johannes, L., Ron, E.Z., Tønjum, T., Sansonetti, P.J., & Vicente, M. (2012).** The enemy within us: lessons from the 2011 European *Escherichia coli* O104:H4 outbreak. *EMBO Mol. Med.* 4, 841–848.
- Karmali, M.A., Mascarenhas, M., Shen, S., Ziebell, K., Johnson, S., Reid-Smith, R., Isaac-Renton, J., Clark, C., Rahn, K., & Kaper, J.B. (2003).** Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. *Journal of Clinical Microbiology*, 41(11), 4930–4940.
- Karmali, M.A., Gannon, V., & Sargeant, J.M. (2010).** Verocytotoxin-producing *Escherichia coli* (VTEC). *Veterinary Microbiology*, 140(3–4), 360–370.
- Kavita, K., de Mets, F., & Gottesman, S. (2018).** New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Current Opinion in Microbiology*, 42, 53–61.
- Kawano, M., Oshima, T., Kasai, H., & Mori, H. (2002).** Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in *Escherichia coli*. *Molecular Microbiology*, 45(2), 333–349.

## 7. BIBLIOGRAFÍA

- Kawecki T.J., Lenski R.E., Ebert D., Hollis B., Olivieri I., & Whitlock M.C. (2013).** Experimental evolution. *Trends in Ecology & Evolution* 27:547–560.
- Kenny, B., DeVinney, R., Stein, M., Reinscheid, D.J., Frey, E.A., & Finlay, B.B. (1997).** Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell*, 91(4), 511–520.
- Kim, D.W., Lenzen, G., Page, A.L., Legrain, P., Sansonetti, P.J., & Parsot, C. (2005).** The *Shigella flexneri* effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes. *Proceedings of the National Academy of Sciences*, 102(39), 14046–14051.
- Kim, B., Park, H., Na, D., & Lee, S. Y. (2014).** Metabolic engineering of *Escherichia coli* for the production of phenol from glucose. *Biotechnology Journal*, 9(5), 621–629.
- Kirov, S.M., Tassell, B.C., Semmler, A.B.T., O'Donovan, L.A., Rabaan, A.A., & Shaw, J.G. (2002).** Lateral flagella and swarming motility in *Aeromonas* species. *Journal of Bacteriology*, 184(2), 547–555.
- Knowles, T.J., Scott-Tucker, A., Overduin, M., & Henderson, I.R. (2009).** Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nature Reviews Microbiology*, 7(3), 206–214.
- Knutton, S. (1987).** Adhesion of enteropathogenic *Escherichia coli* to human intestinal enterocytes and cultured human intestinal mucosa. *Infection and Immunity*, 55(1).
- Konishi, N., Obata, H., Monma, C., Nakama, A., Kai, A., & Tsuji, T. (2011).** Bacteriological and epidemiological characteristics of enterotoxigenic *Escherichia coli* isolated in Tokyo, Japan, between 1966 and 2009. *Journal of Clinical Microbiology*, 49(9), 3348–3351.
- Koop, A.H., Hartley, M.E., & Bourgeois, S. (1987).** A low-copy-number vector utilizing beta-galactosidase for the analysis of gene control elements. *Gene*, 52(2–3), 245–256.
- Kotlajich, M.V, Hron, D.R., Boudreau, B.A., Sun, Z., Lyubchenko, Y.L., & Landick, R. (2015).** Bridged filaments of histone-like nucleoid structuring protein pause RNA polymerase and aid termination in bacteria. *ELife*, 4.
- Krause, D.O., Little, A.C., Dowd, S.E., & Bernstein, C.N. (2011).** Complete genome sequence of adherent invasive *Escherichia coli* UM146 isolated from Ileal Crohn's disease biopsy tissue. *Journal of Bacteriology*, 193(2), 583–583.
- Kumar, S., Stecher, G., & Tamura, K. (2016).** MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874.
- La Teana, A., Brandi, A., Falconi, M., Spurio, R., Pon, C.L., & Gualerzi, C.O. (1991).** Identification of a cold shock transcriptional enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS. *Proceedings of the National Academy of Sciences*, 88(23), 10907–10911.

- Laaberki, M.H., Janabi, N., Oswald, E., & Repoila, F. (2006).** Concert of regulators to switch on LEE expression in enterohemorrhagic *Escherichia coli* O157:H7: Interplay between Ler, GrlA, HNS and RpoS. *International Journal of Medical Microbiology*, 296(4–5), 197–210.
- Laalami, S., Zig, L., & Putzer, H. (2014).** Initiation of mRNA decay in bacteria. *Cellular and Molecular Life Sciences*, 71(10), 1799–1828.
- Lamberte, L.E., Baniulyte, G., Singh, S.S., Stringer, A.M., Bonocora, R.P., Stracy, M., Kapanidis, A.N., Wade, J.T., & Grainger, D.C. (2017).** Horizontally acquired AT-rich genes in *Escherichia coli* cause toxicity by sequestering RNA polymerase. *Nature Microbiology*, 2(3), 16249.
- Lan, R., Alles, M.C., Donohoe, K., Martinez, M.B., & Reeves, P.R. (2004).** Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infection and immunity*, 72(9), 5080–5088.
- Lang, B., Blot, N., Bouffartigues, E., Buckle, M., Geertz, M., Gualerzi, C.O., Mavathur, R., Muskhelishvili, G., Pon, C.L., Rimsky, S., Stella, S., Babu, M.M., & Travers, A. (2007).** High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Research*, 35(18), 6330–6337.
- Lawrence, J.G., & Ochman, H. (1997).** Amelioration of bacterial genomes: rates of change and exchange. *Journal of Molecular Evolution*, 44(4), 383–397.
- Le Bouguéneq C. (1999).** Diarrhea-associated diffusely adherent *Escherichia coli*. *Clinical microbiology reviews*, 12(1), 180–181.
- Lederberg, J., Cavalli, L.L., & Lederberg, E.M. (1952).** Sex compatibility in *Escherichia coli*. *Genetics*, 37(6), 720–730.
- Lee, S. (2004).** Microbial safety of pickled fruit and vegetables and hurdle technology. *Internet Journal of Food Safety*, 4, pp: 21–32.
- Lee, D.J., Minchin, S.D., & Busby, S.J.W. (2012).** Activating transcription in bacteria. *Annual Review of Microbiology*, 66(1), 125–152.
- Lehman I.R. (1958).** Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *The Journal of Biological Chemistry*, 233(1).
- Leimbach, A., Hacker, J., & Dobrindt, U. (2013).** *E. coli* as an all-rounder: the thin line between commensalism and pathogenicity. *Current Topics in Microbiology and Immunology*, 358. (pp. 3–32).
- Leonard, P.G., Ono, S., Gor, J., Perkins, S.J., & Ladbury, J.E. (2009).** Investigation of the self-association and hetero-association interactions of H-NS and StpA from enterobacteria. *Molecular Microbiology*, 73(2), 165–179.
- Li, G.W., Berg, O.G., & Elf, J. (2009).** Effects of macromolecular crowding and DNA looping on gene regulation kinetics. *Nature Physics*, 5(4), 294–297.

## 7. BIBLIOGRAFÍA

**Lim, C.J., Lee, S.Y., Kenney, L.J., & Yan, J. (2012).** Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing. *Scientific Reports*, 2(1), 509.

**Lingwood, C.A., Law, H., Richardson, S., Petric, M., Brunton, J.L., DeGrandis, S. & Karmali, M. (1987).** Glycolipid binding of purified and recombinant *Escherichia coli* produced verotoxin in vitro. *Journal of Biological Chemistry* 262, 8834–8839.

**Liu, Q., & Richardson, C.C. (1993).** Gene 5.5 protein of bacteriophage T7 inhibits the nucleoid protein H-NS of *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 90(5), 1761–1765.

**Liu, T., & Khosla, C. (2010).** A balancing act for taxol precursor pathways in *E. coli*. *Science*, 330(6000), 44–45.

**Liu, Y., Chen, H., Kenney, L.J., & Yan, J. (2010).** A divalent switch drives H-NS/DNA-binding conformation between stiffening and bridging modes. *Genes & Development*, 24(4), 339–344.

**López-Garrido, J., Puerta-Fernández, E., & Casadesús, J. (2014).** A eukaryotic-like 3' untranslated region in *Salmonella enterica* hliD mRNA. *Nucleic Acids Research*, 42(9), 5894–5906.

**Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., & Hinton, J.C.D. (2006).** H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathogens*, 2(8), e81.

**Ludwig, A., Bauer, S., Benz, R., Bergmann, B., & Goebel, W. (1999).** Analysis of the SlyA-controlled expression, subcellular localization and pore-forming activity of a 34 kDa haemolysin (ClyA) from *Escherichia coli* K-12. *Molecular Microbiology*, 31(2), 557–567.

**Lukjancenko, O., Wassenaar, T.M., & Ussery, D.W. (2010).** Comparison of 61 Sequenced *Escherichia coli* Genomes. *Microbial Ecology*, 60(4), 708–720.

**Ma, Z., Gong, S., Richard, H., Tucker, D.L., Conway, T., & Foster, J.W. (2003).** GadE (YhiE) activates glutamate decarboxylase-dependent acid resistance in *Escherichia coli* K-12. *Molecular Microbiology*, 49(5), 1309–1320.

**Ma, C., Wickham, M.E., Guttman, J.A., Deng, W., Walker, J., Madsen, K.L., Jacobson, K., Vogl, W.A., Finlay, B.B., & Vallance, B.A. (2006).** *Citrobacter rodentium* infection causes both mitochondrial dysfunction and intestinal epithelial barrier disruption in vivo: role of mitochondrial associated protein (Map). *Cellular Microbiology*, 8(10), 1669–1686.

**Madrid, C., Nieto, J.M., Paytubi, S., Falconi, M., Gualerzi, C.O., & Juarez, A. (2002).** Temperature- and H-NS-dependent regulation of a plasmid-encoded virulence operon expressing *Escherichia coli* hemolysin. *Journal of Bacteriology*, 184(18), 5058–5066.

**Madrid, C., Garcia, J., Pons, M., & Juarez, A. (2007).** Molecular evolution of the H-NS protein: interaction with Hha-like proteins is restricted to *Enterobacteriaceae*. *Journal of Bacteriology*, 189(1), 265–268.



- Maeda, T., & Wachi, M. (2012).** 3' Untranslated Region-dependent degradation of the *aceA* mRNA, encoding the glyoxylate cycle enzyme isocitrate lyase, by RNase E/G in *Corynebacterium glutamicum*. *Applied and Environmental Microbiology*, 78(24), 8753–8761.
- Majdalani, N., Vanderpool, C.K., & Gottesman, S. (2005).** Bacterial small RNA regulators. *Critical Reviews in Biochemistry and Molecular Biology*, 40(2), 93–113.
- Mann, E.A. & Saeed, S.A. (2012).** Gastrointestinal infection as a trigger for inflammatory bowel disease. *Current Opinion in Gastroenterology*, 28(1).
- Marklund, B., Tennent, J. M., Garcia, E., Hamers, A., Baga, M., Lindberg, F., Gaastra, W., & Normark, S. (1992).** Horizontal gene transfer of the *Escherichia coli* pap and prs pili operons as a mechanism for the development of tissue-specific adhesive properties. *Molecular Microbiology*, 6(16), 2225–2242.
- Marschall, C., Labrousse, V., Kreimer, M., Weichart, D., Kolb, A., & Hengge-Aronis, R. (1998).** Molecular analysis of the regulation of *csiD*, a carbon starvation-inducible gene in *Escherichia coli* that is exclusively dependent on  $\sigma^S$  and requires activation by cAMP-CRP. *Journal of Molecular Biology*, 276(2), 339–353.
- Marteyn, B., Gazi, A., & Sansonetti, P. (2012).** *Shigella*: a model of virulence regulation in vivo. *Gut Microbes*, 3(2), 104–120.
- Martinez-Medina, M., Mora, A., Blanco, M., López C., Alonso, M.P., Bonacorsi, S., Nicolas-Chanoine, M.H., Darfeuille-Michaud, A., Garcia-Gil, & Blanco, J. (2009).** Similarity and divergence among adherent-invasive *Escherichia coli* and extraintestinal pathogenic *E. coli* strains. *Journal of Clinical Microbiology JCM*, 47:3968–3979.
- Mayr, C. (2017).** Regulation by 3'-Untranslated Regions. *Annual Review of Genetics*, 51(1), 171–194.
- McCarter, L.L. (2004).** Dual flagellar systems enable motility under different circumstances. *Journal of Molecular Microbiology and Biotechnology*, 7(1–2), 18–29.
- McDaniel, T.K., Jarvis, K.G., Sonnenberg, M.S., & Kaper, J.B. (1995).** A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proceedings of the National Academy of Sciences*, 92(5), 1664–1668.
- McDaniel, T.K., & Kaper, J.B. (1997).** A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Molecular Microbiology*, 23(2), 399–407.
- McHugh, G.L., & Swartz, M.N. (1977).** Elimination of plasmids from several bacterial species by Novobiocin. *Antimicrobial Agents and Chemotherapy*, 12(3), 423–426.
- McKane, M., & Milkman, R. (1995).** Transduction, restriction and recombination patterns in *Escherichia coli*. *Genetics*, 139(1), 35–43.
- McLeod, S. M., & Johnson, R. C. (2001).** Control of transcription by nucleoid proteins. *Current Opinion in Microbiology*, 4(2), 152–159.

## 7. BIBLIOGRAFÍA

- Meijer, A.H., van Dijk, E.L. & Hoge, J.H.C. (1996).** Novel members of a family of AT hook-containing DNA-binding proteins from rice are identified through their *in vitro* interaction with consensus target sites of plant and animal homeodomain proteins. *Plant Molecular Biology* 31, 607–618.
- Mellies, J.L., Barron, A.M. & Carmona, A.M. (2007).** Enteropathogenic and enterohemorrhagic *Escherichia coli* virulence gene regulation. *Infection and Immunity IAI*, 75(9), 4199-4210.
- Ménard, L.P., Lussier, J.G., Lépine, F., de Sousa, C.P., & Dubreuil, J.D. (2004).** Expression, purification, and biochemical characterization of enteroaggregative *Escherichia coli* heat-stable enterotoxin 1. *Protein Expression and Purification*, 33(2), 223–231.
- Michelacci, V., Tozzoli, R., Caprioli, A., Martínez, R., Scheutz, F., Grande, L., Sánchez, S., & Morabito, S. (2013).** A new pathogenicity island carrying an allelic variant of the Subtilase cytotoxin is common among Shiga toxin producing *Escherichia coli* of human and ovine origin. *Clinical Microbiology and Infection*, 19(3), E149–E156.
- Mikulskis, A.V., & Cornetis, G.R. (1994).** A new class of proteins regulating gene expression in enterobacteria. *Molecular Microbiology*, 11(1), 77–86.
- Miller, J. H. (1992).** A Short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria / by Jeffrey H. Miller. *Cold Spring Harbor Laboratory Press*.
- Miquel, S., Peyretailade, E., Claret, L., de Vallée, A., Dossat, C., Vacherie, B., Zineb, E.H., Segurens, B., Barbe, V., Sauvanet, P., Neut, C., Colombel, J.F., Medigue, C., Mojica, F.J. M., Peyret, P., Bonnet, R., & Darfeuille-Michaud, A. (2010).** Complete genome sequence of Crohn's disease-associated adherent-invasive *E. coli* Strain LF82. *PLoS ONE*, 5(9), e12714.
- Miyakoshi, M., Chao, Y., & Vogel, J. (2015).** Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Current Opinion in Microbiology*, 24, 132–139.
- Molaei, N., Abtahi, H., Ghannadzadeh, M. J., Karimi, M., & Ghaznavi-Rad, E. (2015).** Application of Reverse Transcriptase-PCR (RT-PCR) for rapid detection of viable *Escherichia coli* in drinking water samples. *Journal of environmental health science & engineering*, 13, 24.
- Molina, F., López-Acedo, E., Tabla, R., Roa, I., Gómez, A., & Rebollo, J. E. (2015).** Improved detection of *Escherichia coli* and coliform bacteria by multiplex PCR. *BMC Biotechnology*, 15(1), 48.
- Morin, N., Tirling, C., Ivison, S. M., Kaur, A.P., Nataro, J.P., & Steiner, T.S. (2010).** Autoactivation of the AggR regulator of enteroaggregative *Escherichia coli* *in vitro* and *in vivo*. *FEMS Immunology & Medical Microbiology*, 58(3), 344–355.
- Morin, N., Santiago, A.E., Ernst, R.K., Guillot, S.J., & Nataro, J.P. (2013).** Characterization of the AggR regulon in enteroaggregative *Escherichia coli*. *Infection and immunity*, 81(1), 122–132.



- Mounier, J., Popoff, M.R., Enninga, J., Frame, M.C., Sansonetti, P.J., & Van Nhieu, G.T. (2009).** The IpaC carboxyterminal effector domain mediates Src-dependent actin polymerization during *Shigella* invasion of epithelial cells. *PLoS Pathogens*, 5(1), e1000271.
- Moura, R.A., Sircili, M.P., Leomil, L., Matte, M.H., Trabulsi, L.R., Elias, W.P., Irino, K., & Pestana de Castro, A.F. (2009).** Clonal relationship among atypical enteropathogenic *Escherichia coli* strains isolated from different animal species and humans. *Applied and Environmental Microbiology*, 75(23), 7399–7408.
- Mouriño, M., Balsalobre, C., Madrid, C., Nieto, J.M., Prenafeta, A., Muñoa, F.J., & Juárez, A. (1998).** Osmolarity modulates the expression of the Hha protein from *Escherichia coli*. *FEMS Microbiology Letters*, 160(2), 225–229.
- Mueller, M., Grauschopf, U., Maier, T., Glockshuber, R., & Ban, N. (2009).** The structure of a cytolytic  $\alpha$ -helical toxin pore reveals its assembly mechanism. *Nature*, 459(7247), 726–730.
- Müller, C.M., Dobrindt, U., Nagy, G., Emody, L., Uhlin, B.E., & Hacker, J. (2006).** Role of histone-like Proteins H-NS and StpA in expression of virulence determinants of uropathogenic *Escherichia coli*. *Journal of Bacteriology*, 188(15), 5428–5438.
- Müller, C.M., Schneider, G., Dobrindt, U., Emödy, L., Hacker, J., & Uhlin, B.E. (2010).** Differential effects and interactions of endogenous and horizontally acquired H-NS-like proteins in pathogenic *Escherichia coli*. *Molecular Microbiology*, 75(2), 280–293.
- Munson, G.P., & Scott, J.R. (2002).** Rns, a virulence regulator within the AraC family, requires binding sites upstream and downstream of its own promoter to function as an activator. *Molecular Microbiology*, 36(6), 1391–1402.
- Nagarajavel, V., Madhusudan, S., Dole, S., Rahmouni, A.R., & Schnetz, K. (2007).** Repression by binding of H-NS within the transcription unit. *Journal of Biological Chemistry*, 282(32), 23622–23630.
- Nash, J.H., Villegas, A., Kropinski, A.M., Aguilar-Valenzuela, R., Konczy, P., Mascarenhas, M., Ziebell, K., Torres, A.G., Karmali, M.A., & Coombes, B.K. (2010).** Genome sequence of adherent-invasive *Escherichia coli* and comparative genomic analysis with other *E. coli* pathotypes. *BMC Genomics*, 11(1), 667.
- Nataro, J.P., Kaper, J.B., Robins-Browne, R., Prado, V., Vial, P., & Levine, M.M. (1987).** Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *The Pediatric Infectious Disease Journal*, 6(9), 829–831.
- Nataro, J.P., Deng, Y., Maneval, D.R., German, A.L., Martin, W.C., & Levine, M.M. (1992).** Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. *Infection and Immunity*, 60(6), 2297–304.
- Nataro, J.P., Yikang, D., Yinggang, D., & Walker, K. (1994).** AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *Journal of Bacteriology*, 176(15), 4691–4699.

## 7. BIBLIOGRAFÍA

- Nataro, J.P., Deng, Y., Cookson, S., Cravioto, A., Savarino, S.J., Guers, L.D., Levine, M.M., Tacket, C.O. (1995).** Heterogeneity of enteroaggregative *Escherichia coli* virulence demonstrated in volunteers. *The Journal of Infectious Diseases*, 171(2), 465-468.
- Nataro, J.P., & Kaper, J.B. (1998).** Diarrheagenic *Escherichia coli*. *Clinical Microbiology Reviews*, 11, 142-201.
- Navarre, W.W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S.J., & Fang, F.C. (2006).** Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science*, 313(5784).
- Navarre, W.W., McClelland, M., Libby, S.J., & Fang, F.C. (2007).** Silencing of xenogenic DNA by H-NS--facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes & Development*, 21(12), 1456-1471.
- Navarro-Garcia, F., Canizalez-Roman, A., Luna, J., Sears, C., & Nataro, J.P. (2001).** Plasmid-encoded toxin of enteroaggregative *Escherichia coli* is internalized by epithelial cells. *Infection and Immunity*, 69(2), 1053-1060.
- Navarro-Garcia, F., Serapio-Palacios, A., Vidal, J.E., Salazar, M.I., & Tapia-Pastrana, G. (2014).** EspC promotes epithelial cell detachment by enteropathogenic *Escherichia coli* via sequential cleavages of a cytoskeletal protein and then focal adhesion proteins. *Infection and Immunity*, 82(6), 2255-2265.
- Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A., & Gautheret, D. (2011).** ARNold: A web tool for the prediction of Rho-independent transcription terminators. *RNA Biology*, 8(1), 11-13.
- Ndlovu, T., Le Roux, M., Khan, W., & Khan, S. (2015).** Co-detection of virulent *Escherichia coli* genes in surface water sources. *PLOS ONE*, 10(2), e0116808.
- Neely, M.N., & Friedman, D.I. (1998).** Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Molecular Microbiology*, 28(6), 1255-1267.
- Nguyen, R.N., Taylor, L.S., Tauschek, M., & Robins-Browne, R.M. (2006).** Atypical enteropathogenic *Escherichia coli* infection and prolonged diarrhea in children. *Emerging Infectious Diseases*, 12(4), 597-603.
- Nguyen, Y., & Sperandio, V. (2012).** Enterohemorrhagic *E. coli* (EHEC) pathogenesis. *Frontiers in Cellular and Infection Microbiology*, 2.
- Nieto, J.M., Carmona, M., Bolland, S., Jubete, Y., Cruz, F., & Juárez, A. (1991).** The hha gene modulates haemolysin expression in *Escherichia coli*. *Molecular Microbiology*, 5(5), 1285-1293.
- Nieto, J.M., Prenafeta, A., Miquelay, E., Torrades, S., & Juarez, A. (1998).** Sequence, identification and effect on conjugation of the rmoA gene of plasmid R100-1. *FEMS Microbiology Letters*. 169:59-66.

- Nieto, J.M., Madrid, C., Prenafeta, A., Miquelay, E., Balsalobre, C., Carrascal, M., & Juárez, A. (2000). Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and H-NS. *Molecular and General Genetics MGG*, 263(2), 349–358.
- Nieto, J.M., Madrid, C., Miquelay, E., Parra, J. L., Rodriguez, S., & Juarez, A. (2002). Evidence for direct protein-protein interaction between members of the Enterobacterial Hha/YmoA and H-NS families of proteins. *Journal of Bacteriology*, 184(3), 629–635.
- Nishi, J., Sheikh, J., Mizuguchi, K., Luisi, B., Burland, V., Boutin, A., Rose, D.J., Blattner, F.R., & Nataro, J.P. (2003). The export of coat protein from enteroaggregative *Escherichia coli* by a specific ATP-binding cassette transporter system. *Journal of Biological Chemistry*, 278(46), 45680–45689.
- Nougayrede, J.P., & Donnenberg, M.S. (2004). Enteropathogenic *Escherichia coli* EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway. *Cellular Microbiology*, 6(11), 1097–1111.
- O'Brien, A.D., Tesh, V.L., Donohue-Rolfe, A., Jackson, M.P., Olsnes, S., Sandvig, K., Lindberg, A.A., & Keusch, G.T. (1992). Shiga toxin: biochemistry, genetics, mode of action, and role in pathogenesis. *Current Topics in Microbiology and Immunology*; 180:65-94.
- Ochman, H. & Selander, R.K. (1984a). Evidence for clonal population structure in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 81(1), 198–201.
- Ochman, H. & Selander, R.K. (1984b). Standard reference strains of *Escherichia coli* from natural populations. *Journal of Bacteriology*, 157(2), 690–693.
- Okeke, I.N., Wallace-Gadsden, F., Simons, H.R., Matthews, N., Labar, A.S., Hwang, J., & Wain, J. (2010). Multi-Locus sequence typing of enteroaggregative *Escherichia coli* isolates from nigerian children uncovers multiple lineages. *PLoS ONE*, 5(11), e14093.
- Olekhnovich, I.N., & Kadner, R.J. (2007). Role of nucleoid-associated proteins Hha and H-NS in expression of *Salmonella enterica* activators HilD, HilC, and RtsA required for cell invasion. *Journal of Bacteriology*, 189(19), 6882–6890.
- Olsén, A., Arnqvist, A., Hammar, M., Sukupolvi, S., & Normark, S. (1993). The RpoS Sigma factor relieves H-NS-mediated transcriptional repression of *csgA*, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Molecular Microbiology*, 7(4), 523–536.
- Omar, K.B., & Barnard, T.G. (2014). Detection of diarrhoeagenic *Escherichia coli* in clinical and environmental water sources in South Africa using single-step 11-gene m-PCR. *World Journal of Microbiology and Biotechnology*, 30(10), 2663–2671.
- Ono, S., Goldberg, M.D., Olsson, T., Esposito, D., Hinton, J.C.D., & Ladbury, J.E. (2005). H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. *Biochemical Journal*, 391(2), 203–213.

## 7. BIBLIOGRAFÍA

**Oshima, T., Ishikawa, S., Kurokawa, K., Aiba, H., & Ogasawara, N. (2006).** *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Research*, 13(4), 141–153.

**Owen-Hughes, T.A., Pavitt, G.D., Santos, D.S., Sidebotham, J.M., Hulton, C.S.J., Hinton, J.C.D., & Higgins, C.F. (1992).** The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell*, 71(2), 255–265.

**Ozawa, Y., Mizuno, T., & Mizushima, S. (1987).** Roles of the Pribnow box in positive regulation of the *ompC* and *ompF* genes in *Escherichia coli*. *Journal of Bacteriology*, 169(3), 1331–1334.

**Pal, C., Bengtsson-Palme, J., Kristiansson, E., & Larsson, D.G. (2015).** Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC genomics*, 16, 964.

**Pallen, M.J., & Gophna, U. (2007).** Bacterial flagella and Type III Secretion: case studies in the evolution of complexity. *Gene and Protein Evolution* (pp. 30–47).

**Park, H.S., Ostberg, Y., Johansson, J., Wagner, E.G.H., & Uhlin, B.E. (2010).** Novel role for a bacterial nucleoid protein in translation of mRNAs with suboptimal ribosome-binding sites. *Genes & Development*, 24(13), 1345–1350.

**Park, H., Valencia-Gallardo, C., Sharff, A., Van Nhieu, G.T., & Izard, T. (2011).** Novel vinculin binding site of the IpaA invasin of *Shigella*. *Journal of Biological Chemistry*, 286(26), 23214–23221.

**Park, C., & Zhang, J. (2012).** High expression hampers horizontal gene transfer. *Genome Biology and Evolution*, 4(4), 523–532.

**Parsot, C., Taxman, E., & Mekalanos, J.J. (1991).** ToxR regulates the production of lipoproteins and the expression of serum resistance in *Vibrio cholerae*. *Proceedings of the National Academy of Sciences*, 88(5), 1641–1645.

**Paton, A.W. & Paton, J.C. (1998).** Detection and characterization of Shiga toxicogenic *Escherichia coli* by using multiplex PCR assays for *stx1*, *stx2*, *eaeA*, enterohemorrhagic *E. coli* *hlyA*, *rfbO111*, and *rfbO157*. *Journal of Clinical Microbiology*, 36(2), 598–602.

**Paton, A.W. & Paton, J.C. (2003).** Detection and characterization of STEC in stool samples using PCR. *Methods in Molecular Medicine*, 73, 45–54.

**Paton, A.W., Srimanote, P., Talbot, U.M., Wang, H., & Paton, J. C. (2004).** A new family of potent AB5 cytotoxins produced by Shiga toxicogenic *Escherichia coli*. *Journal of Experimental Medicine*, 200(1), 35–46.

**Paytubi, S., Madrid, C., Forn, N., Nieto, J.M., Balsalobre, C., Uhlin, B.E., & Juárez, A. (2004).** YdgT, the Hha paralogue in *Escherichia coli*, forms heteromeric complexes with H-NS and StpA. *Molecular Microbiology*, 54(1), 251–263.

**Paytubi, S., García, J., & Juárez, A. (2011).** Bacterial Hha-like proteins facilitate incorporation of horizontally transferred DNA. *Open Life Sciences*, 6(6).

- Peiffer, I., Guignot, J., Barbat, A., Carnoy, C., Moseley, S.L., Nowicki, B.J., Servin, A.L., & Bernet-Camard, M.F. (2000).** Structural and functional lesions in brush border of human polarized intestinal caco-2/TC7 cells infected by members of the Afa/Dr diffusely adhering family of *Escherichia coli*. *Infection and Immunity*, 68(10), 5979–5990.
- Pelletier, J., Halvorsen, K., Ha, B.Y., Paparcone, R., Sandler, S.J., Woldringh, C.L., Wong, W.P., & Jun, S. (2012).** Physical manipulation of the *Escherichia coli* chromosome reveals its soft nature. *Proceedings of the National Academy of Sciences*, 109(40), E2649–E2656.
- Perdomo, O.J., Cavaillon, J.M., Huerre, M., Ohayon, H., Gounon, P., & Sansonetti, P.J. (1994).** Acute inflammation causes epithelial invasion and mucosal destruction in experimental shigellosis. *The Journal of Experimental Medicine*, 180(4), 1307–1319.
- Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F., & Liuni, S. (2001).** Structural and functional features of eukaryotic mRNA untranslated regions. *Gene*, 276(1–2), 73–81.
- Peters, J.M., Mooney, R.A., Grass, J.A., Jessen, E.D., Tran, F., & Landick, R. (2012).** Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes & Development*, 26(23), 2621–2633.
- Peters, J.M., Vangeloff, A.D., & Landick, R. (2011).** Bacterial transcription terminators: The RNA 3'-End Chronicles. *Journal of Molecular Biology*, 412(5), 793–813.
- Pilonieta, M.C., Boderó, M.D., & Munson, G.P. (2007).** CfaD-dependent expression of a novel extracytoplasmic protein from enterotoxigenic *Escherichia coli*. *Journal of Bacteriology*, 189(14), 5060–5067.
- Postler, T.S., & Ghosh, S. (2017).** Understanding the Holobiont: How microbial metabolites affect human health and shape the immune system. *Cell Metabolism*, 26(1), 110–130.
- Pribnow, D. (1975).** Nucleotide Sequence of an RNA Polymerase Binding Site at an Early T7 Promoter. *Proceedings of the National Academy of Sciences of the United States of America*, 72(3), 784–788.
- Pupo, G.M., Karaolis, D.K., Lan, R., & Reeves, P.R. (1997).** Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. *Infection and immunity*, 65(7), 2685–2692.
- Qadri, F., Svennerholm, A.M., Faruque, A.S., & Sack, R.B. (2005).** Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clinical Microbiology Reviews*, 18(3), 465–483.
- Qu, Y., Lim, C.J., Whang, Y.R., Liu, J., & Yan, J. (2013).** Mechanism of DNA organization by *Mycobacterium tuberculosis* protein Lsr2. *Nucleic Acids Research*, 41(10), 5263–5272.



## 7. BIBLIOGRAFÍA

- Rajakumar, K., Sasakawa, C., & Adler, B. (1997).** Use of a novel approach, termed island probing, identifies the *Shigella flexneri* she pathogenicity island which encodes a homolog of the immunoglobulin A protease-like family of proteins. *Infection and Immunity*, 65(11), 4606-4614.
- Rajkumari, K., & Gowrishankar, J. (2001).** In vivo expression from the RpoS-dependent P1 promoter of the osmotically regulated proU Operon in *Escherichia coli* and *Salmonella enterica* serovar *Typhimurium*: activation by rho and hns mutations and by cold stress. *Journal of Bacteriology*, 183(22), 6543–6550.
- Rasko, D.A., Webster, D.R., Sahl, J.W., Bashir, A., Boisen, N., Scheutz, F., Paxinos, E.E., Sebra, R., Chin, C.S., Iliopoulos, D., Klammer, A., Peluso, P., Lee, L., Kislyuk, A.O., Bullard, J., Kasarskis, A., Wang, S., Eid, J., Rank, D., Waldor, M.K. (2011).** Origins of the *E. coli* strain causing an outbreak of Hemolytic–Uremic Syndrome in Germany. *New England Journal of Medicine*, 365(8), 709–717.
- Ren, C.P., Chaudhuri, R.R., Fivian, A., Bailey, C.M., Antonio, M., Barnes, W.M., & Pallen, M.J. (2004).** The ETT2 gene cluster, encoding a Second Type III Secretion System from *Escherichia coli*, is present in the majority of strains but has undergone widespread mutational attrition. *Journal of Bacteriology*, 186(11), 3547–3560.
- Ren, C.P., Beatson, S.A., Parkhill, J., & Pallen, M.J. (2005).** The Flag-2 locus, an ancestral gene cluster, is potentially associated with a novel flagellar system from *Escherichia coli*. *Journal of Bacteriology*, 187(4), 1430–1440. <https://doi.org/10.1128/JB.187.4.1430-1440.2005>
- Ren, G.X., Guo, X.P., & Sun, Y.C. (2017).** Regulatory 3' Untranslated Regions of bacterial mRNAs. *Frontiers in Microbiology*, 8.
- Robbe-Saule, V., Schaeffer, F., Kowarz, L., & Norel, F. (1997).** Relationships between H-NS,  $\sigma^S$ , SpvR and growth phase in the control of spvR, the regulatory gene of the *Salmonella* plasmid virulence operon. *Molecular and General Genetics MGG*, 256(4), 333–347.
- Robins-Browne, R.M. (1987).** Traditional enteropathogenic *Escherichia coli* of infantile diarrhea. *Clinical Infectious Diseases*, 9(1), 28–53.
- Robins-Browne, R.M. & Hartland, E.L. (2002).** *Escherichia coli* as a cause of diarrhea. *Journal of Gastroenterology and Hepatology*, 17: 467-475.
- Rodríguez, E., Gaggero, C., & Laviña, M. (1999).** The structural gene for microcin H47 encodes a peptide precursor with antibiotic activity. *Antimicrobial Agents and Chemotherapy*, 43(9), 2176-2182.
- Rossiter, A.E., Browning, D.F., Leyton, D.L., Johnson, M.D., Godfrey, R.E., Wardius, C.A., Desvaux, M., Cunningham, A.F., Ruiz-Perez, F., Nataro, J.P., Busby, S.J.W., & Henderson, I.R. (2011).** Transcription of the plasmid-encoded toxin gene from enteroaggregative *Escherichia coli* is regulated by a novel co-activation mechanism involving CRP and Fis. *Molecular Microbiology*, 81(1), 179–191.

- Roy, K., Hilliard, G.M., Hamilton, D.J., Luo, J., Ostmann, M.M., & Fleckenstein, J.M. (2009).** Enterotoxigenic *Escherichia coli* EtpA mediates adhesion between flagella and host cells. *Nature*, 457(7229), 594–598.
- Roy, K., Kansal, R., Bartels, S.R., Hamilton, D.J., Shaaban, S., & Fleckenstein, J.M. (2011).** Adhesin degradation accelerates delivery of heat-labile toxin by enterotoxigenic *Escherichia coli*. *Journal of Biological Chemistry*, 286(34), 29771–29779.
- Ruiz de los Mozos, I., Vergara-Irigaray, M., Segura, V., Villanueva, M., Bitarte, N., Saramago, M., Domingues, S., Arraiano, C.M., Fechter, P., Romby, P., Valle, J., Solano, C., Lasa, I., & Toledo-Arana, A. (2013).** Base pairing interaction between 5'- and 3'-UTRs controls icaR mRNA Translation in *Staphylococcus aureus*. *PLoS Genetics*, 9(12), e1004001.
- Sabat, A. J., Budimir, A., Nashev, D., Sá-Leão, R., van Dijk, J. M., Laurent, F., Grundmann, H., Friedrich, A. W., & on behalf of the ESCMID Study Group. (2013).** Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Eurosurveillance*, 18(4), 20380.
- Sabui, S., Ghosal, A., Dutta, S., Ghosh, A., Ramamurthy, T., Nataro, J.P., Hamabata, T., & Chatterjee, N.S. (2010).** Allelic variation in colonization factor CS6 of enterotoxigenic *Escherichia coli* isolated from patients with acute diarrhoea and controls. *Journal of Medical Microbiology*, 59(7), 770–779.
- Sahl, J.W., Steinsland, H., Redman, J.C., Angiuoli, S.V., Nataro, J.P., Sommerfelt, H., & Rasko, D.A. (2011).** A comparative genomic analysis of diverse clonal types of enterotoxigenic *Escherichia coli* reveals pathovar-specific conservation. *Infection and Immunity*, 79(2), 950–960.
- Saldana, Z., Erdem, A.L., Schuller, S., Okeke, I.N., Lucas, M., Sivananthan, A., Phillips, A. D., Kaper, J.B., Puente, J.L., & Giron, J.A. (2009).** The *Escherichia coli* Common Pilus and the Bundle-Forming Pilus act in concert during the formation of localized adherence by enteropathogenic *E. coli*. *Journal of Bacteriology*, 191(11), 3451–3461.
- Salinger, N., Kokona, B., Fairman, R., & Okeke, I.N. (2009).** The plasmid-encoded regulator activates factors conferring lysozyme resistance on enteropathogenic *Escherichia coli* strains. *Applied and environmental microbiology*, 75(1), 275–280.
- Samadder, P., Xicohtencatl-Cortes, J., Saldaña, Z., Jordan, D., Tarr, P.I., Kaper, J.B., & Girón, J.A. (2009).** The *Escherichia coli* ycbQRST operon encodes fimbriae with laminin-binding and epithelial cell adherence properties in Shiga-toxigenic *E. coli* O157:H7. *Environmental Microbiology*, 11(7), 1815–1826.
- Sandvig, K. (2001).** Shiga toxins. *Toxicon* 39, 1629–1635.
- Santiago, A.E., Ruiz-Perez, F., Jo, N.Y., Vijayakumar, V., Gong, M.Q., & Nataro, J.P. (2014).** A large family of antivirulence regulators modulates the effects of transcriptional activators in Gram-negative pathogenic bacteria. *PLoS Pathogens*, 10(5), e1004153.
- Santiago, A.E., Yan, M.B., Tran, M., Wright, N., Luzader, D.H., Kendall, M.M., Ruiz-Perez, F., & Nataro, J.P. (2016).** A large family of anti-activators accompanying XylS/AraC family regulatory proteins. *Molecular Microbiology*, 101(2), 314–332.



## 7. BIBLIOGRAFÍA

- Santiago, A.E., Yan, M.B., Hazen, T.H., Sauder, B., Meza-Segura, M., Rasko, D. A., Kendall, M.M., Ruiz-Perez, F., & Nataro, J.P. (2017).** The AraC Negative Regulator family modulates the activity of histone-like proteins in pathogenic bacteria. *PLoS Pathogens*, 13(8), e1006545.
- Sarowska, J., Futoma-Koloch, B., Jama-Kmiecik, A., Frej-Madrzak, M., Ksiaczczyk, M., Bugla-Ploskonska, G., & Choroszy-Krol, I. (2019).** Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: recent reports. *Gut Pathogens*, 11(1), 10.
- Scaletsky, I.C.A., Pedroso, M.Z., Oliva, C.A.G., Carvalho, R.L.B., Morais, M.B., & Fagundes-Neto, U. (1999).** A localized adherence-like pattern as a second pattern of adherence of classic enteropathogenic *Escherichia coli* to HEp-2 cells that is associated with infantile diarrhea. *Infection and Immunity*, 67(7).
- Schägger, H., & von Jagow, G. (1987).** Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, 166(2), 368–379.
- Schaller, H., Gray, C., & Herrmann, K. (1975).** Nucleotide sequence of an RNA polymerase binding site from the DNA of bacteriophage fd. *Proceedings of the National Academy of Sciences of the United States of America*, 72(2), 737–741.
- Scheutz, F., Teel, L.D., Beutin, L., Pierard, D., Buvens, G., Karch, H., Mellmann, A., Caprioli, A., Tozzoli, R., Morabito, S., Strockbine, N.A., Melton-Celsa, A.R., Sanchez, M., Persson, S., & O'Brien, A.D. (2012).** Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *Journal of Clinical Microbiology*, 50(9), 2951–2963.
- Schmidt, H., & Hensel, M. (2004).** Pathogenicity Islands in Bacterial Pathogenesis. *Clinical Microbiology Reviews*, 17(1), 14–56.
- Schneider, B.L., Kiupakis, A.K., & Reitzer, L.J. (1998).** Arginine catabolism and the arginine succinyltransferase pathway in *Escherichia coli*. *Journal of Bacteriology*, 180(16).
- Schröder, O., & Wagner, R. (2002).** The bacterial regulatory protein H-NS - A versatile modulator of nucleic acid structures. *Biological Chemistry*, 383(6), 945–960.
- Schüller, S. (2011).** Shiga toxin interaction with human intestinal epithelium. *Toxins*, 3(6), 626–639.
- Sekse, C., Sunde, M., Lindstedt, B.A., Hopp, P., Bruheim, T., Cudjoe, K.S., Kvitle, B., & Urdahl, A.M. (2011).** Potentially human-pathogenic *Escherichia coli* O26 in norwegian sheep flocks. *Applied and Environmental Microbiology*, 77(14), 4949–4958.
- Selinger, D.W., Saxena, R.M., Cheung, K.J., Church, G.M., & Rosenow, C. (2003).** Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome research*, 13(2), 216–223.

- Seputiene, V., Suziedelis, K., Normark, S., Melefors, O., & Suziedeliene, E. (2004).** Transcriptional analysis of the acid-inducible *asr* gene in enterobacteria. *Research in Microbiology*, 155(7), 535–542.
- Servin A.L. (2005).** Pathogenesis of Afa/Dr diffusely adhering *Escherichia coli*. *Clinical microbiology reviews*, 18(2), 264–292.
- Sette, M., Spurio, R., Trotta, E., Brandizi, C., Brandi, A., Pon, C.L., Barbato, G., Boelens, R., & Gualerzi, C.O. (2009).** Sequence-specific recognition of DNA by the C-terminal domain of nucleoid-associated protein H-NS. *Journal of Biological Chemistry*, 284(44), 30453–30462.
- Sharma, V.K., & Zuerner, R.L. (2004).** Role of *hha* and *ler* in transcriptional regulation of the *esp* operon of enterohemorrhagic *Escherichia coli* O157:H7. *Journal of Bacteriology*, 186(21), 7290–7301.
- Sharma, V.K., & Bearson, B.L. (2013).** Hha controls *Escherichia coli* O157:H7 biofilm formation by differential regulation of global transcriptional regulators FlhDC and CsgD. *Applied and Environmental Microbiology*, 79(7), 2384–2396.
- Sharma, V.K., & Casey, T.A. (2014).** Determining the relative contribution and hierarchy of *hha* and *qseBC* in the regulation of flagellar motility of *Escherichia coli* O157:H7. *PLoS ONE*, 9(1), e85866.
- Sheikh, J., Czeczulin, J.R., Harrington, S., Hicks, S., Henderson, I.R., Le Bouguéneq, C., Gounon, P., Phillips, A., & Nataro, J.P. (2002).** A novel dispersin protein in enteroaggregative *Escherichia coli*. *Journal of Clinical Investigation*, 110(9), 1329–1337.
- Sheikh, J., Hicks, S., Dall’Agnol, M., Phillips, A. D., & Nataro, J. P. (2008).** Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Molecular Microbiology*, 41(5), 983–997.
- Shi, X., & Bennett, G.N. (1994).** Plasmids bearing *hfq* and the *hns*-like gene *stpA* complement *hns* mutants in modulating arginine decarboxylase gene expression in *Escherichia coli*. *Journal of Bacteriology*, 176(21), 6769–6775.
- Shiga, K. (1898).** Ueber den Erreger der Dysenterie in Japan, Vorläufige Mitteilung. *Zentralbl Bakteriol Microbiol Hyg*, vol. 23 (pg. 599-600).
- Shiga, Y., Sekine, Y., Kano, Y., & Ohtsubo, E. (2001).** Involvement of H-NS in transpositional recombination mediated by IS1. *Journal of Bacteriology*, 183(8), 2476–2484.
- Shin, M., Song, M., Rhee, J.H., Hong, Y., Kim, Y.J., Seok, Y.J., Ha, K.S., Jung, S.H., & Choy, H.E. (2005).** DNA looping-mediated repression by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping. *Genes & development*, 19(19), 2388–2398.
- Shintani, M., Sanchez, Z. K., & Kimbara, K. (2015).** Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Frontiers in microbiology*, 6, 242.

## 7. BIBLIOGRAFÍA

- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., & Chandler, M. (2006).** ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Research*, *34*(90001), D32–D36.
- Silphaduang, U., Mascarenhas, M., Karmali, M., & Coombes, B.K. (2007).** Repression of intracellular virulence factors in *Salmonella* by the Hha and YdgT nucleoid-associated proteins. *Journal of Bacteriology*, *189*(9), 3669–3673
- Sinclair, J.F., & O'Brien, A.D. (2002).** Cell surface-localized nucleolin is a eukaryotic receptor for the adhesin intimin- $\gamma$  of enterohemorrhagic *Escherichia coli* O157:H7. *Journal of Biological Chemistry*, *277*(4), 2876–2885.
- Singh, S.S., Singh, N., Bonocora, R.P., Fitzgerald, D.M., Wade, J.T., & Grainger, D.C. (2014).** Widespread suppression of intragenic transcription initiation by H-NS. *Genes & development*, *28*(3), 214–219.
- Singh, K., Milstein, J. N., & Navarre, W.W. (2016).** Xenogeneic Silencing and Its Impact on Bacterial Genomes. *Annual Review of Microbiology*, *70*(1), 199–213.
- Skoudy, A., Mounier, J., Aruffo, A., Ohayon, H., Gounon, P., Sansonetti, P., & Tran Van Nhieu, G. (2000).** CD44 binds to the *Shigella* IpaB protein and participates in bacterial invasion of epithelial cells. *Cellular Microbiology*, *2*(1), 19–33.
- Smits, W.K., & Grossman, A.D. (2010).** The transcriptional regulator Rok binds A+T-Rich DNA and is involved in repression of a mobile genetic element in *Bacillus subtilis*. *PLoS Genetics*, *6*(11), e1001207.
- Smyth, C.P., Lundback, T., Renzoni, D., Siligardi, G., Beavil, R., Layton, M., Sidebotham, J. M., Hinton, J. C. D., Driscoll, P.C., Higgins, C.F., & Ladbury, J.E. (2000).** Oligomerization of the chromatin-structuring protein H-NS. *Molecular Microbiology*, *36*(4), 962–972.
- Soderblom, T., Oxhamre, C., Wai, S.N., Uhlen, P., Aperia, A., Uhlin, B.E., & Richter-Dahlfors, A. (2005).** Effects of the *Escherichia coli* toxin cytolysin A on mucosal immunostimulation via epithelial Ca<sup>2+</sup> signalling and Toll-like receptor 4. *Cellular Microbiology*, *7*(6), 779–788.
- Sohel, I., Puente, J.L., Ramer, S.W., Bieber, D., Wu, C. Y., & Schoolnik, G.K. (1996).** Enteropathogenic *Escherichia coli*: identification of a gene cluster coding for bundle-forming pilus morphogenesis. *Journal of Bacteriology*, *178*(9), 2613–2628.
- Solórzano, C., Srikumar, S., Canals, R., Juárez, A., Paytubi, S., & Madrid, C. (2015).** Hha has a defined regulatory role that is not dependent upon H-NS or StpA. *Frontiers in microbiology*, *6*, 773.
- Solovyev, V., & Salamov, A. (2011).** Automatic annotation of microbial genomes and metagenomic sequences. In *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies* (Ed. R.W. Li), *Nova Science Publishers*, 61-78.

- Sonden, B., & Uhlin, B.E. (1996).** Coordinated and differential expression of histone-like proteins in *Escherichia coli*: regulation and function of the H-NS analog StpA. *EMBO Journal*, 15(18), 4970–4980.
- Sonnenfield, J.M., Burns, C.M., Higgins, C.F., & Hinton, J.C.D. (2001).** The nucleoid-associated protein StpA binds curved DNA, has a greater DNA-binding affinity than H-NS and is present in significant levels in hns mutants. *Biochimie*, 83(2), 243–249.
- Spangler, B.D. (1992).** Structure and function of cholera toxin and the related *Escherichia coli* heat-labile enterotoxin. *Microbiological Reviews*, 56(4), 622–647.
- Spurio, R., Falcon, M., Brandi, A., Pon, C.L., Gualerzi, C.O. (1997).** The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *The EMBO Journal*, 16(7), 1795–1805.
- Stegmeier, J. F., Glück, A., Sukumaran, S., Mäntele, W., & Andersen, C. (2007).** Characterization of YtfM, a second member of the Omp85 family in *Escherichia coli*. *Biological Chemistry*, 388(1).
- Steiner, T.S., Lima, A.A.M., Nataro, J.P. & Guerrant, R.L. (1998).** Enterotoxigenic *Escherichia coli* produce intestinal inflammation and growth impairment and cause interleukin-8 release from intestinal epithelial cells. *Journal of Infectious Diseases*, 177(1).
- Stevens, A. (1960).** Incorporation of the adenine ribonucleotide into RNA by cell fractions from *E. coli* B. *Biochemical and Biophysical Research Communications*, 3(1), 92–96.
- Storz, G., Vogel, J., & Wassarman, K.M. (2011).** Regulation by small RNAs in bacteria: expanding frontiers. *Molecular Cell*, 43(6), 880–891.
- Strauman, M.C., Harper, J.M., Harrington, S.M., Boll, E.J., & Nataro, J.P. (2010).** Enterotoxigenic *Escherichia coli* disrupts epithelial cell tight junctions. *Infection and Immunity*, 78(11), 4958–4964.
- Studier, F.W., & Moffatt, B.A. (1986).** Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, 189(1), 113–130.
- Sullivan, M.J., Petty, N.K., & Beatson, S.A. (2011).** Easyfig: a genome comparison visualizer. *Bioinformatics*, 27(7), 1009–1010.
- Sun, Y., & O’Riordan, M.X.D. (2013).** Regulation of bacterial pathogenesis by intestinal short-chain fatty acids. *Advances in Applied Microbiology*, 85, 93–118.
- Swingle, B., O’Carroll, M., Haniford, D., & Derbyshire, K.M. (2004).** The effect of host-encoded nucleoid proteins on transposition: H-NS influences targeting of both IS903 and Tn10. *Molecular Microbiology*, 52(4), 1055–1067.
- Tauschek, M., Gorrell, R.J., Strugnell, R.A., & Robins-Browne, R.M. (2002).** Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 99(10), 7066–7071.

## 7. BIBLIOGRAFÍA

- Tauschek, M., Strugnell, R.A., & Robins-Browne, R.M. (2002).** Characterization and evidence of mobilization of the LEE pathogenicity island of rabbit-specific strains of enteropathogenic *Escherichia coli*. *Molecular Microbiology*, 44(6), 1533–1550.
- Taylor, R.G, Walker, D.C. & McInnes, R.R. (1993).** *E. coli* host strains significantly affect the quality of small-scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research*, 21(7), 1677–1678.
- Tendeng, C., & Bertin, P.N. (2003).** H-NS in Gram-negative bacteria: a family of multifaceted proteins. *Trends in microbiology* 11: 511–8.
- Tendeng, C., Soutourina, O.A., Danchin, A., & Bertin, P.N. (2003).** MvaT proteins in *Pseudomonas* spp.: a novel class of H-NS-like proteins. *Microbiology* 149: 3047–50.
- Tippner, D., & Wagner, R. (1995).** Fluorescence analysis of the *Escherichia coli* transcription regulator H-NS reveals two distinguishable complexes dependent on binding to specific or nonspecific DNA sites. *Journal of Biological Chemistry*, 270(38), 22243–22247.
- Touchon, M., Hoede, C., Tenailon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M. El, Frapy, E., Denamur, E. (2009).** Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genetics*, 5(1), e1000344.
- Tozzoli, R., Caprioli, A., Cappannella, S., Michelacci, V., Marziano, M.L., & Morabito, S. (2010).** Production of the Subtilase AB<sub>5</sub> cytotoxin by Shiga toxin-negative *Escherichia coli*. *Journal of Clinical Microbiology*, 48(1), 178–183.
- Tran, H.J., Heroven, A.K., Winkler, L., Spreter, T., Beatrix, B., & Dersch, P. (2005).** Analysis of RovA, a transcriptional regulator of *Yersinia pseudotuberculosis* virulence that acts through antirepression and direct transcriptional activation. *Journal of Biological Chemistry*, 280(51), 42423–42432.
- Tran, L., van Baarsel, J.A., Washburn, R.S., Gottesman, M.E., & Miller, J.H. (2011).** Single-gene deletion mutants of *Escherichia coli* with altered sensitivity to Bicyclomycin, an Inhibitor of Transcription Termination Factor Rho. *Journal of Bacteriology*, 193(9), 2229–2235.
- Trofa, A.F., Ueno-Olsen, H., Oiwa, R., & Yoshikawa, M. (1999).** Dr. Kiyoshi Shiga: discoverer of the Dysentery Bacillus. *Clinical Infectious Diseases*, 29(5), 1303–1306.
- Troxell, B., & Hassan, H.M. (2013).** Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Frontiers in Cellular and Infection Microbiology*, 3, 59.
- Trun, N.J., & Marko, J.F. (1998).** Architecture of a bacterial chromosome. *American Society for Microbiology Reviews.*, 64, 276–283.
- Trunk, T., Khalil, H.S., & Leo, J.C. (2018).** Bacterial autoaggregation. *AIMS Microbiology*, 4(1), 140–164.
- Tucker, D.L., Tucker, N., & Conway, T. (2002).** Gene expression profiling of the pH response in *Escherichia coli*. *Journal of Bacteriology*, 184(23), 6551–6558.



- Ueguchi, C., Suzuki, T., Yoshida, T., Tanaka, K., & Mizuno, T. (1996).** Systematic mutational analysis revealing the functional domain organization of *Escherichia coli* nucleoid protein H-NS. *Journal of Molecular Biology*, 263(2), 149–162.
- Uyar, E., Kurokawa, K., Yoshimura, M., Ishikawa, S., Ogasawara, N., & Oshima, T. (2009).** Differential binding profiles of StpA in Wild-Type and hns mutant cells: a comparative analysis of cooperative partners by chromatin Immunoprecipitation-Microarray analysis. *Journal of Bacteriology*, 191(7), 2388–2391.
- Uzzau, S., Figueroa-Bossi, N., Rubino, S., & Bossi, L. (2001).** Epitope tagging of chromosomal genes in *Salmonella*. *Proceedings of the National Academy of Sciences*, 98(26), 15264–15269.
- Valle, J., Mabbett, A.N., Ulett, G.C., Toledo-Arana, A., Wecker, K., Totsika, M., Schembri, M.A., Ghigo, J.M., & Beloin, C. (2008).** UpaG, a new member of the trimeric Autotransporter family of adhesins in uropathogenic *Escherichia coli*. *Journal of Bacteriology*, 190(12), 4147–4161.
- van der Valk, R.A., Vreede, J., Qin, L., Moolenaar, G.F., Hofmann, A., Goosen, N., & Dame, R.T. (2017).** Mechanism of environmentally driven conformational changes that modulate H-NS DNA-bridging activity. *ELife*, 6.
- Varshavsky, A.J., Nedospasov, S.A., Bakayev, V.V., Bakayeva, T.G., & Georgiev, G.P. (1977).** Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Research*, 4(8), 2725–2746.
- Velarde, J.J., Varney, K.M., Inman, K.G., Farfan, M., Dudley, E., Fletcher, J., Weber, D.J., & Nataro, J.P. (2007).** Solution structure of the novel dispersin protein of enteroaggregative *Escherichia coli*. *Molecular Microbiology*, 66(5), 1123–1135.
- Vidal, J.E., & Navarro-García, F. (2008).** EspC translocation into epithelial cells by enteropathogenic *Escherichia coli* requires a concerted participation of type V and III secretion systems. *Cellular Microbiology*, 10(10), 1975–1986.
- Vila, J., Vargas, M., Henderson, I.R., Gascón, J., & Nataro, J.P. (2000).** Enteroaggregative *Escherichia coli* virulence factors in traveler's diarrhea strains. *Journal of Infectious Diseases*, 182(6), 1780–1783.
- Villaseca, J.M., Navarro-García, F., Mendoza-Hernandez, G., Nataro, J.P., Cravioto, A., & Eslava, C. (2000).** Pet Toxin from enteroaggregative *Escherichia coli* produces cellular damage associated with fodrin disruption. *Infection and Immunity*, 68(10), 5920–5927.
- Vivero, A., Baños, R.C., Mariscotti, J.F., Oliveros, J.C., García-del Portillo, F., Juárez, A., & Madrid, C. (2008).** Modulation of horizontally acquired genes by the Hha-YdgT proteins in *Salmonella enterica* Serovar Typhimurium. *Journal of Bacteriology*, 190(3), 1152–1156.
- von Rhein, C., Bauer, S., Simon, V., & Ludwig, A. (2008).** Occurrence and characteristics of the cytolysin A gene in *Shigella* strains and other members of the family *Enterobacteriaceae*. *FEMS Microbiology Letters*, 287(2), 143–148.

## 7. BIBLIOGRAFÍA

- Wade, J.T., & Grainger, D.C. (2017).** Spurious transcription and its impact on cell function. *Transcription*, 9(3), 182–189.
- Wagner, E.G., & Vogel, J. (2005).** Approaches to identify novel non-messenger RNAs in bacteria and to investigate their biological functions: functional analysis of identified non-mRNAs. In Handbook of RNA Biochemistry. Hartmann, R.K., Bindereif, A., Schön, A., and Westhof, E. (eds). Weinheim, Berlin: Wiley-VCH, pp. 614–642.
- Walthers, D., Li, Y., Liu, Y., Anand, G., Yan, J., & Kenney, L.J. (2011).** *Salmonella enterica* response regulator SsrB relieves H-NS silencing by displacing H-NS bound in polymerization mode and directly activates transcription. *Journal of Biological Chemistry*, 286(3), 1895–1902.
- Wang, H., Paton, J.C., & Paton, A.W. (2007).** Pathologic changes in mice induced by subtilase cytotoxin, a potent new *Escherichia coli* AB<sub>5</sub> toxin that targets the endoplasmic reticulum. *The Journal of Infectious Diseases*, 196(7), 1093–1101.
- Wang, W., Li, G.W., Chen, C., Xie, X.S., & Zhuang, X. (2011).** Chromosome organization by a Nucleoid-Associated Protein in live bacteria. *Science*, 333(6048), 1445–1449.
- Wang, F., Jiang, Z., Li, Y., He, X., Zhao, J., Yang, X., Zhu, L., Yin, Z., Li, X., Wang, X., Liu, W., Shang, W., Yang, Z., Wang, S., Zhen, Q., Zhang, Z., Yu, Y., Zhong, H., Ye, Q., Yuan, J. (2013).** *Shigella flexneri* T3SS effector IpaH4.5 modulates the host inflammatory response via interaction with NF- $\kappa$ B p65 protein. *Cellular Microbiology*, 15(3), 474–485.
- Waters, L.S., & Storz, G. (2009).** Regulatory RNAs in Bacteria. *Cell*, 136(4), 615–628.
- Welch, R.A., Burland, V., Plunkett, G., Redford, P., Roesch, P., Rasko, D., Buckles, E.L., Liou, S.R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C., Perna, N.T., Mobley, H.L.T., Donnenberg, M.S., & Blattner, F.R. (2002).** Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 99(26), 17020–17024.
- Welch, T.J., Farewell, A., Neidhardt, F.C., Bartlett, D.H. (1993).** Stress-response of *Escherichia coli* to elevated hydrostatic pressure. *Journal of Bacteriology* 175(22):7170–7177
- Will, W.R., Navarre, W.W., & Fang, F.C. (2015).** Integrated circuits: how transcriptional silencing and counter-silencing facilitate bacterial evolution. *Current Opinion in Microbiology*, 23, 8–13.
- Williams, R.M., Rimsky, S., & Buc, H. (1996).** Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. *Journal of Bacteriology*, 178(15), 4335–4343.
- Williamson, H.S., & Free, A. (2005).** A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS antagonist. *Molecular Microbiology*, 55(3), 808–827.



- Winardhi, R.S., Fu, W., Castang, S., Li, Y., Dove, S.L., & Yan, J. (2012).** Higher order oligomerization is required for H-NS family member MvaT to form gene-silencing nucleoprotein filament. *Nucleic Acids Research*, *40*(18), 8942–8952.
- Winardhi, R.S., Castang, S., Dove, S.L., & Yan, J. (2014).** Single-Molecule study on histone-like nucleoid-structuring Protein (H-NS) paralogue in *Pseudomonas aeruginosa*: MvaU bears DNA organization mode similarities to MvaT. *PLoS ONE*, *9*(11), e112246.
- Winardhi, R.S., Yan, J., & Kenney, L.J. (2015).** H-NS regulates gene expression and compacts the nucleoid: insights from single-molecule experiments. *Biophysical Journal*, *109*(7), 1321–1329.
- Winkler, W.C. (2005).** Riboswitches and the role of noncoding RNAs in bacterial metabolic control. *Current Opinion in Chemical Biology*, *9*(6), 594–602.
- World Health Organization. (2018).** WHO Press, Geneva, Switzerland. *World Health Statistics* (2018): Monitoring health for the SDGs.
- Wray, C., McLaren, I.M., Randall, L.P., & Pearson, G.R. (2000).** Natural and experimental infection of normal cattle with *Escherichia coli* 0157. *Veterinary Record*, *147*(3), 65–68.
- Wu, X.Y., Chapman, T., Trott, D.J., Bettelheim, K., Do, T.N., Driesen, S., Walker, M. J., & Chin, J. (2007).** Comparative analysis of virulence genes, genetic diversity, and phylogeny of commensal and enterotoxigenic *Escherichia coli* isolates from weaned pigs. *Applied and Environmental Microbiology*, *73*(1), 83–91.
- Xicohtencatl-Cortes, J., Monteiro-Neto, V., Ledesma, M.A., Jordan, D.M., Francetic, O., Kaper, J.B., Puente, J.L., & Girón, J.A. (2007).** Intestinal adherence associated with type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. *Journal of Clinical Investigation*, *117*(11), 3519–3529.
- Xicohtencatl-Cortes, J., Monteiro-Neto, V., Saldana, Z., Ledesma, M.A., Puente, J.L., & Giron, J.A. (2009).** The type 4 pili of enterohemorrhagic *Escherichia coli* O157:H7 are multipurpose structures with pathogenic attributes. *Journal of Bacteriology*, *191*(1), 411–421.
- Yamamoto, T., Echeverria, P., & Yokota, T. (1992).** Drug resistance and adherence to Human intestines of enteroaggregative *Escherichia coli*. *Journal of Infectious Diseases*, *165*(4), 744–749.
- Yang, J., Baldi, D.L., Tauschek, M., Strugnell, R.A., & Robins-Browne, R.M. (2007).** Transcriptional regulation of the yghJ-pppA-yghG-gspCDEFGHIJKLM cluster, encoding the Type II Secretion pathway in enterotoxigenic *Escherichia coli*. *Journal of Bacteriology*, *189*(1), 142–150.
- Yasir, M., Icke, C., Abdelwahab, R., Haycocks, J.R., Godfrey, R.E., Sazinas, P., Pallen, M. J., Henderson, I.R., Busby, S., & Browning, D.F. (2019).** Organization and architecture of AggR-dependent promoters from enteroaggregative *Escherichia coli*. *Molecular microbiology*, *111*(2), 534–551.

## 7. BIBLIOGRAFÍA

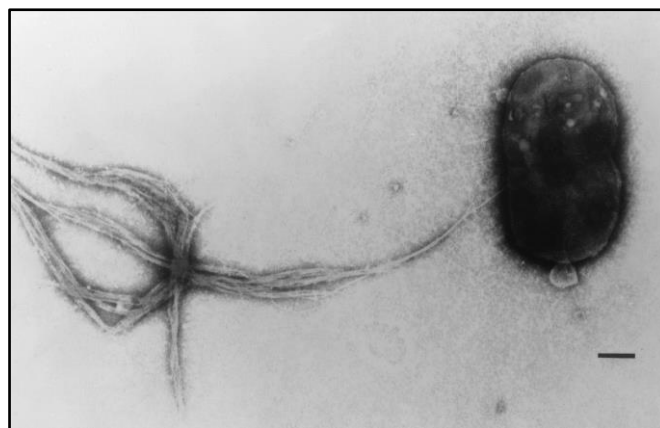
- Yoshida, S., Handa, Y., Suzuki, T., Ogawa, M., Suzuki, M., Tamai, A., Abe, A., Katayama, E., & Sasakawa, C. (2006).** Microtubule-severing activity of *Shigella* is pivotal for intercellular spreading. *Science*, *314*(5801), 985–989.
- Yu, R.R. & DiRita, V.J. (1999).** Analysis of an auto-regulatory loop controlling ToxT, cholera toxin and toxin-coregulated pilus production in *Vibrio cholerae*. *Journal of Bacteriology* *181*: 2584–2592.
- Yu, R.R., & DiRita, V.J. (2002).** Regulation of gene expression in *Vibrio cholerae* by ToxT involves both antirepression and RNA polymerase stimulation. *Molecular Microbiology*, *43*(1), 119–134.
- Yun, S.H., Ji, S.C., Jeon, H.J., Wang, X., Kim, S.W., Bak, G., Lee, Y., & Lim, H.M. (2012).** The CnuK9E H-NS complex antagonizes DNA binding of DicA and leads to temperature-dependent filamentous growth in *E. coli*. *PLoS ONE*, *7*(9), e45236.
- Zhang, A., & Belfort, M. (1992).** Nucleotide sequence of a newly-identified *Escherichia coli* gene, *stpA*, encoding an H-NS-like protein. *Nucleic Acids Research*, *20*(24), 6735–6735.
- Zhang, A., Derbyshire, V., Salvo, J.L., & Belfort, M. (1995).** *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly in vitro. *RNA (New York, N.Y.)*, *1*(8).
- Zhang, A., Rimsky, S., Reaban, M. E., Buc, H., & Belfort, M. (1996).** *Escherichia coli* protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *The EMBO journal*, *15*(6), 1340–1349.
- Zhu, H., Mao, X.-J., Guo, X.P., & Sun, Y.C. (2016).** The *hmsT* 3' untranslated region mediates c-di-GMP metabolism and biofilm formation in *Yersinia pestis*. *Molecular Microbiology*, *99*(6), 1167–1178.
- Zurawski, D.V., Mummy, K.L., Faherty, C.S., McCormick, B.A., & Maurelli, A.T. (2009).** *Shigella flexneri* type III secretion system effectors OspB and OspF target the nucleus to downregulate the host inflammatory response via interactions with retinoblastoma protein. *Molecular Microbiology*, *71*(2), 350–368.

## **8. ANEXO INTRODUCCIÓN**



### A.I.1 *E. coli* enteropatógena

*Escherichia coli* enteropatógena es un patógeno no invasivo que no produce ningún tipo de enterotoxinas. El mecanismo fisiopatológico de esta bacteria se fundamenta en un primer proceso de unión de la bacteria al enterocito a través del pili BFP (*bundle-forming pilus*) (figura A.I.1). Esta estructura proteica es responsable de la unión de las bacterias entre sí al epitelio a través de un patrón de adherencia localizada en forma de microcolonias (Giron y col., 2002). La expresión de este pili adherente está regulada por el operón *perABC* (*plasmid-encoded regulator*) (Mellies y col., 2007). Dicho operón se encuentra codificado, al igual que el pili BFP, en el plásmido pEAF (Sohel y col., 1996), presente únicamente en las cepas “típicas” del patotipo EPEC. Otros elementos implicados en la primera etapa de colonización por parte de las bacterias tEPEC son la proteína LifA (*lymphocyte inhibitory factor*), una proteína de superficie bacteriana de elevado peso molecular implicada en la adherencia *in vitro* a células epiteliales (Badea y col., 2003), y la proteína ECP (*E. coli common pilus*), que actúa como un factor de adherencia accesorio a las células intestinales y en las interacciones bacteria-bacteria (Saldaña y col., 2009). En cuanto a las cepas enteropatógenas atípicas (aEPEC), su mecanismo patogénico también radica en un primer evento de unión al enterocito, pero con un patrón agregativo distinto por el hecho de no disponer del pili BFP. También ha podido comprobarse que muchas de las cepas clínicas aEPEC que también codifican por la proteína LifA son importantes agentes causantes de diarreas en niños menores de 5 años (Afset y col., 2006).



**Figura A.I.1.** Microfotografía realizada mediante microscopía electrónica del pili BFP (*bundle-forming pilus*) de la cepa *E. coli* E2348/69. Dicha estructura está implicada en mediar la adhesión de las bacterias EPEC a los enterocitos intestinales. La barra indica un tamaño de 0,35  $\mu\text{m}$ . Imagen adaptada de (Nataro y Kaper, 1998).

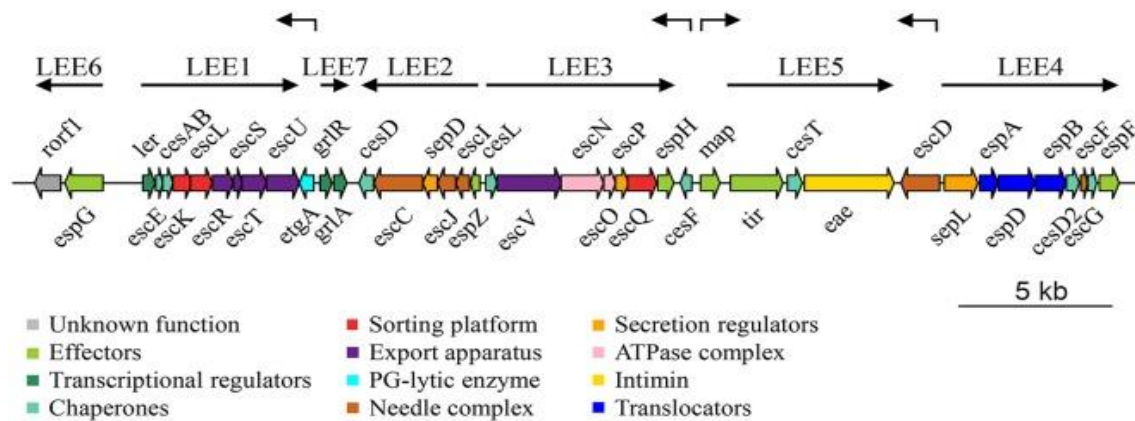
## 8. ANEXO INTRODUCCIÓN

Una vez producida la interacción bacteria-enterocito, la célula bacteriana es capaz de translocar proteínas efectoras al interior de la célula huésped. Todo ello viene facilitado por el sistema de secreción tipo 3 (T3SS) codificado en el cromosoma de la bacteria, concretamente en la isla de patogenicidad LEE (figura A.I.2). En dicho elemento genético se encuentran codificadas las proteínas efectoras que la bacteria secreta al interior del enterocito y sus correspondientes chaperonas, además de la mayoría de los genes necesarios para la regulación de la expresión de todo el entramado macromolecular responsable de la patogénesis de la bacteria (Deng y col., 2004; Iguchi y col., 2009). Dichos elementos reguladores son las proteínas Ler (*LEE-encoded regulator*), GrlA (*global regulator of LEE activator*) y GrlR (*global regulator of LEE repressor*), y el operon *per* (Garmendia y col., 2005).

Una de las principales proteínas efectoras implicadas en el proceso infectivo de las bacterias enteropatógenicas es la proteína Tir (*translocated intimin receptor*). La proteína Tir es un receptor que la bacteria transloca a la célula huésped y, una vez allí, actúa como receptor molecular para la proteína intimina bacteriana (Kenny y col., 1997). La intimina es una proteína transmembrana de 94 kDa codificada por el gen *eae* que, una vez unida a su receptor Tir, provoca una reorganización en los filamentos de actina de la célula epitelial intestinal, dando lugar a la polimerización de la misma en largas cadenas que alteran la morfología del citoesqueleto celular, dañando de esta manera las microvellosidades intestinales provocando su pérdida de función.

Paralelamente al proceso de unión del receptor translocado Tir y la intimina bacteriana, se produce la unión del sistema de secreción tipo III al enterocito y la consecuente translocación de proteínas efectoras al interior de la célula infectada. Una de las principales proteínas efectoras secretada por parte de las cepas EPEC es Map (*mitochondria-associated protein*). Dicha proteína estimula la formación de largos filopodios en la membrana de los enterocitos provocando grandes alteraciones en la barrera epitelial intestinal del huésped infectado, además de provocar disfunciones en las mitocondrias del enterocito mediante la generación de cambios de potencial entre la membrana de la mitocondria y el citoplasma de la célula infectada (Ma y col., 2006). Otra proteína efectora secretada por el T3SS que también se dirige a la mitocondria de la célula huésped es la proteína EspF (Nougayrède y Donnenberg, 2004). Esta proteína multifuncional está implicada en la alteración de las uniones entre las membranas apicales y basolaterales de las células de la barrera epitelial intestinal provocando la

permeabilización de la misma (Guttman y col., 2006), y en la reorganización y pérdida de la nucleolina del nucléolo de la célula provocando el consecuente bloqueo del procesamiento del ARNr celular (Dean y col., 2010). Junto con la proteína EspG, EspF también está implicada en alterar la localización de las acuaporinas celulares, provocando cambios en el flujo de electrolitos del huésped dando lugar a la enfermedad diarreica (Guttman y col., 2007). Además, la proteína EspG también es capaz de interactuar con la tubulina celular provocando la alteración del citoesqueleto de microtúbulos de la célula epitelial (Hardwidge y col., 2005).



**Figura A.I.2.** Representación esquemática de la estructura genética de la isla de patogenicidad LEE de la cepa *Escherichia coli* E2348/69 O127:H6 (EPEC). Los genes se representan como flechas coloreadas en función su categoría funcional propuesta. La organización de los distintos operones de la isla (LEE1 a LEE7) se indica mediante flechas negras sobre los genes LEE, al igual que las unidades transcripcionales individuales (*etgA*, *cesF*, *map* y *escD*). La representación esquemática está realizada a escala (barra de escala 5 Kb). Imagen adaptada de (Gaytán y col., 2016).

Otra proteína efectora muy importante en el mecanismo de patogénesis de EPEC es EspC. A diferencia de las proteínas efectoras anteriores, esta proteína no está codificada en la isla de patogenicidad LEE. Además, es secretada a la célula huésped mediante un mecanismo sinérgico mediado por el sistema de secreción tipo III y el sistema de secreción tipo V (Vidal y Navarro-García, 2008). La proteína EspC es una serina proteasa que actúa como una enterotoxina y está involucrada en la citotoxicidad celular durante la adhesión bacteriana y la generación de poros en la membrana celular eucariota (Navarro-García y col., 2014; Guignot y col., 2015), además de estar implicada en la supervivencia de la bacteria (Drago-Serrano y col., 2006) y conferir resistencia a la lisozima (Salinger y col., 2009).

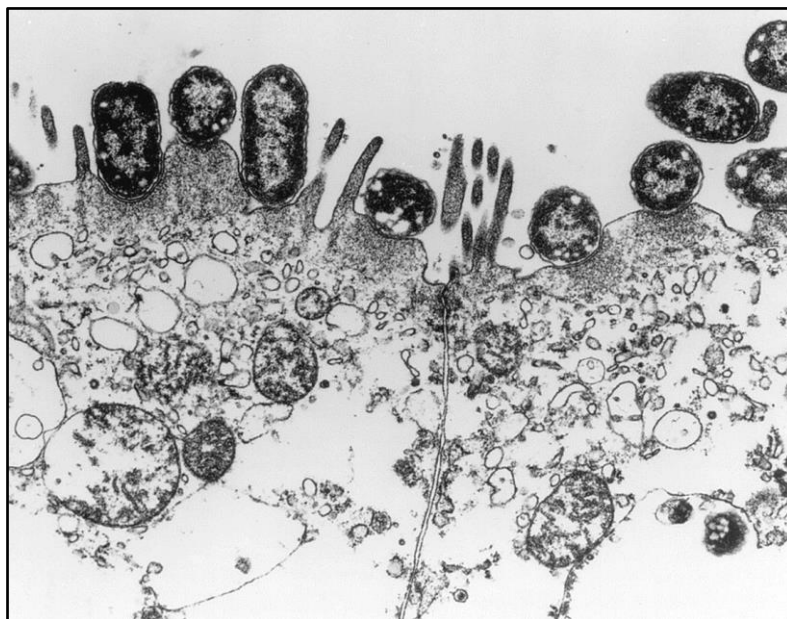


### **A.I.2 *E. coli* enterohemorrágica**

En función de su incidencia y asociación con casos de síndrome urémico hemolítico (SUH) y brotes esporádicos, las cepas pertenecientes al patotipo enterohemorrágico de *E. coli* se clasifican través de su seropatotipo, existiendo hasta la fecha 5 diferentes (de la A a la E). El seropatotipo A está formado por cepas altamente virulentas de los serotipos O157:H7 y O157:H- (H- no móviles) causantes de numerosos brotes y con una alta asociación a SUH. El seropatotipo B incluye cepas de serotipos que han causado brotes ocasionales y son relativamente comunes en casos de colitis hemorrágica (CH) y SUH. El seropatotipo C está formado por cepas de serotipos no implicados en brotes pero que se han aislado de pacientes con CH y SUH. El seropatotipo D incluye cepas de serotipos que nunca se han asociado con SUH, pero que sí se han aislado de pacientes con diarrea y CH. Finalmente, el seropatotipo E incluye muchos serotipos aislados de animales, alimentos y medio ambiente no implicados en casos clínicos en el ser humano (Karmali y col., 2003).

EHEC es capaz de colonizar el intestino grueso, principalmente la región del colon. La interacción de esta bacteria con los enterocitos está mediada por la intimina (la unión de la intimina con su receptor Tir y la nucleolina contribuye en la fijación inicial de la bacteria a la célula huésped) (Sinclair y col., 2002) y por el T3SS. De manera análoga al patotipo enteropatógeno descrito anteriormente, los factores que participan en la unión bacteria-huésped y son responsables de la formación de la lesión intestinal de adhesión y borrado (A/E) típica de estos patotipos virulentos (figura A.I.3), también se encuentran codificados en la isla de patogenicidad LEE (McDaniel y col., 1995). Entre ellos destacan las fimbrias ELF (*E. coli YcbQ laminin-binding fimbria*) y Lpf (*long polar fimbria*), que facilitan la unión de la bacteria a la matriz extracelular a través de la laminina (Samadder y col., 2009; Farfan y col., 2011). Otra proteína importante de unión es la proteína HCP (*hemorrhagic coli pilus*), una proteína multifuncional que forma largas fibras agrupadas que también pueden unirse a las proteínas de la matriz extracelular en las células epiteliales, además de estar implicada en procesos como la motilidad, la formación de *biofilm* y la invasión celular (Xicohtencatl-Cortes y col., 2007; Xicohtencatl-Cortes y col., 2009). Otra proteína importante en la patogénesis de EHEC es la metaloproteasa dependiente de zinc StcE, que presenta actividad mucinasa que le permite a la bacteria reducir la viscosidad de la saliva en la cavidad oral y también ayuda a prevenir los agregados bacterianos de mucina. Además, esta proteína también facilita la unión de la

bacteria a las células epiteliales intestinales al reducir la barrera mucosa en las células huésped (Grys y col., 2005).



**Figura A.I.3.** Microfotografía realizada mediante microscopía electrónica donde se representan las lesiones A/E (*attaching and effacing*) características de los patotipos EPEC y EHEC. Puede observarse la gran alteración de las microvellosidades intestinales como consecuencia de la adherencia de las bacterias a la membrana plasmática apical de los enterocitos. Imagen adaptada de (Nataro y Kaper, 1998).

### **A.I.3 *E. coli* enteroinvasiva**

Durante el proceso infeccioso por parte de las cepas enteroinvasivas de *E. coli*, las bacterias penetran primero en la barrera epitelial a través de las células M por transcitosis. En la submucosa subyacente, la bacteria es fagocitada por macrófagos adyacentes, de los que es capaz de evadir su propia degradación escapando de los fagosomas e induciendo la muerte celular de los macrófagos por piroptosis dependiente de caspasa-1 (Hilbi y col., 1998). La adhesión bacteriana a la célula huésped está mediada por la proteína IpaB y por el complejo IpaBCD, que se unen al receptor de hialuronano CD44 y a la integrina  $\alpha_5\beta_1$ , respectivamente (Skoudy y col., 2000). Las proteínas IpaC, IpgB1, IpgD, IpaA y VirA están implicadas en la reorganización del citoesqueleto celular y en la invaginación de la membrana que promueve la translocación de la bacteria al interior celular dentro de un fagosoma (Handa y col., 2007; Mounier y col., 2009; Park y col., 2011). Posteriormente, la bacteria es capaz de escapar del fagosoma utilizando los efectores IpaB, IpaC, IpaD e IpaH<sub>7,8</sub> (High y col., 1992; Fernandez-Prada y col., 2000; Harrington y col., 2006). Una vez en el interior de la célula huésped infectada, las proteínas OspG, OspF, OspB, OspZ,

OspI, IpaH4.5 e IpaH<sub>9.8</sub> se encargan de inhibir y contrarrestar las respuestas inflamatorias del huésped, como la vía activada por la proteína quinasa activada por mitógeno (MAPK) y las vías NF-κB, así como la producción de citocinas (Kim y col., 2005; Arbibe y col., 2007; Zurawski y col., 2009; Wang y col., 2013). Las proteínas VirG y VirA son las proteínas claves que permiten la motilidad intra e intercelular. VirG provoca la reorganización del citoesqueleto de actina de la célula infectada sobre la superficie bacteriana en forma de cola que proporciona la fuerza propulsora requerida para la motilidad del patógeno (Goldberg y Theriot, 1995; Egile y col., 1999). VirA es otro efector esencial que, con su capacidad para estimular la desestabilización de los microtúbulos, permite que la bacteria se extienda de manera eficiente a través de la densa red citoesquelética intracelular (Yoshida y col., 2006).

Durante todo este proceso infeccioso, la conservación de la integridad epitelial por parte del patógeno es un elemento clave. En un estadio temprano de la infección, a la bacteria le interesa mantener su nicho replicativo intacto para posteriormente promover su colonización. Por ello, en el inicio de la infección no interesa provocar lesiones tempranas en el epitelio. En este sentido, las proteínas IpaB, IpgD, VirA y OspE promueven la supervivencia e integridad de la célula huésped e inhiben las señales apoptóticas (Iwai y col., 2007; Bergounioux y col., 2012).

La respuesta inflamatoria masiva asociada con macrófagos en estado apoptótico, junto con la invasión de las células epiteliales intestinales y la infiltración de leucocitos polimorfonucleares, perforan aún más la barrera epitelial y, en última instancia, conducen a las lesiones características de la patología de la *shigelosis* (Perdomo y col., 1994). Las enterotoxinas *Shigella* 1 y 2 (ShET1 y ShET2) son factores de virulencia adicionales responsables de las manifestaciones clínicas de las infecciones por parte de EIEC/*Shigella*, como la diarrea acuosa (Fasano y col., 1995; Fasano y col., 1997). Las enterotoxinas Pic y SepA, y las citotoxinas Sat y SigA, son factores de virulencia adicionales que contribuyen a la patogénesis de este patotipo bacteriano (Behrens y col., 2002; Faherty y col., 2012).

### **A.I.4 *E. coli* enterotoxigénica**

En las cepas ETEC, los factores de colonización implicados en la interacción con las células hospedadoras se denominan antígenos de Superficie Coli (CS, *Coli surface antigens*). Estas estructuras, fimbriales, afimbriales, helicoidales o fibrilares, se encargan

de mediar la adhesión de las bacterias a receptores de las células epiteliales del intestino delgado (receptores de fibronectina, glicoesfingolípidos y glicoproteínas), lo que permite su colonización por parte de la bacteria (Ghosal y col., 2009; Jansson y col., 2009; Sabui y col., 2010). Además, también existen otras proteínas implicadas en la adherencia de la bacteria a las células epiteliales. Un ejemplo de ello son las proteínas EtpA, una exoproteína de adhesión situada en el extremo del flagelo bacteriano (Roy y col., 2009), o el *longus type IV pilus*, un pili que actúa de manera semejante al BFP de EPEC, además de estar implicado en la autoagregación bacteriana (Gomez-Duarte y col., 1999; Clavijo y col., 2010).

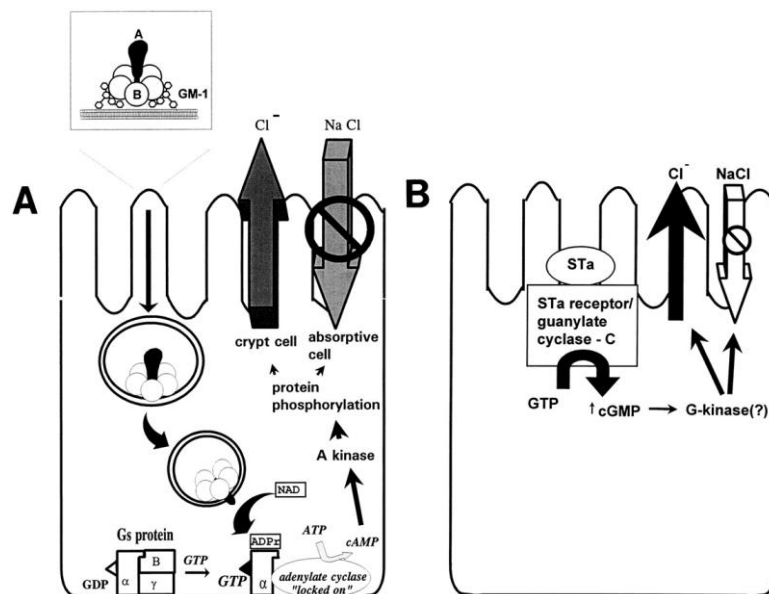
Como se comenta en la Introducción (apartado 1.1.1.4), la presencia de las enterotoxinas LT y ST es el elemento definitorio de este patotipo bacteriano. Se han encontrado dos variantes de ST-I (STp y STh) en bacterias responsables de causar enfermedades en humanos, mientras que la variante ST-II se encuentra en cepas que colonizan animales (Nataro y Kaper, 1998). Estas cepas pueden contener una única copia de la toxina ST o LT (siendo más virulentas las cepas *ST-only* (Qadri y col., 2005), o una combinación de ambas toxinas (LT/ST)).

LT es una toxina de elevado peso molecular que tiene un elevado nivel de similitud (80% de identidad) con la toxina colérica de *V. cholerae*. Cada molécula de toxina se divide en dos subunidades (A y B), de manera que la forma activa de la toxina es una subunidad A y un pentámero de subunidades B. Las cepas ETEC secretan la enterotoxina LT dentro de las vesículas de membrana externa OMV (*outer membrane vesicles*) mediante un sistema de secreción tipo II (Taushek y col., 2002). Las subunidades B pentaméricas se unen a gangliósidos GM1 en balsas lipídicas para liberar la subunidad catalítica A en el interior de la célula infectada. Una vez allí, a través del transporte retrógrado vesicular pasa al citoplasma, donde cataliza la ADP-ribosilación de la subunidad alfa de la proteína G, inhibiendo de esta manera su actividad GTPasa, lo que a su vez provoca un incremento en los niveles de AMPc (Spangler, 1992). Todo ello estimula la secreción de iones cloruro y demás electrolitos a través del canal regulador transmembranal CFTR (*cystic fibrosis transmembrane conductance regulator*). Además, la toxina LT también provoca daños en la barrera intestinal dificultando la absorción e incrementando aún más la pérdida de fluidos por el lumen intestinal (Kreisberg y col., 2011).

La enterotoxina ST es una toxina de tamaño molecular inferior que la proteína LT (después del procesamiento y la secreción a través de las envueltas celulares bacterianas

se convierte en un polipéptido de únicamente 18 o 19 aminoácidos de longitud). Estos pequeños péptidos mimetizan la hormona guanilina, por lo que se unen al receptor de guanililciclase C (GC-C) en el epitelio intestinal. La unión de ST provoca una hiperactivación del receptor que da lugar a un incremento en los niveles de GMPc que a su vez estimularán el receptor CFTR, lo que provoca un incremento de la secreción de iones cloruro en el lumen intestinal, además de inhibir indirectamente el intercambiador de sodio-hidrógeno, alterando de esta manera el correcto flujo de electrolitos y agua en el intestino.

Por tanto, la secreción de iones cloruro acompañada de la secreción de fluidos al lumen intestinal y la malabsorción intestinal provocadas por las toxinas LT y ST, son los eventos que acaban desencadenando los cuadros diarreicos generados por parte del patotipo ETEC (figura A.I.4).



**Figura A.I.4.** Mecanismos de actuación de las toxinas LT y ST secretadas por las bacterias del patotipo ETEC. (A) La toxina LT, al ser internalizada mediante endocitosis, provoca un incremento en los niveles de AMPc en el citoplasma de la célula infectada, lo que dará lugar a un incremento en la secreción de iones  $\text{Cl}^-$  y a la disminución de absorción la iones  $\text{Na}^+$  y  $\text{Cl}^-$ . (B) De manera análoga, la toxina ST provoca un incremento en los niveles citoplasmáticos de GMPc, lo cual provocará un incremento en la secreción de iones  $\text{Cl}^-$  y a la disminución de la absorción de  $\text{NaCl}$ . Imagen adaptada de (Nataro y Kaper, 1998).

Otros factores de virulencia presentes en el patotipo ETEC son la citolisina A, una citotoxina cuyas funciones son la formación de poros, la inducción de apoptosis en los macrófagos y actividad hemolítica (Ludwig y col., 1999; Söderblom y col., 2005), la serin proteasa EatA, implicada en la correcta secreción de la toxina LT (Roy y col., 2011), y la

proteína CexE, una pequeña proteína secretada que facilita el ensamblaje de las fimbrias en la superficie bacteriana (Pilonieta y col., 2007).

### **A.I.5 *E. coli* con adherencia difusa**

Las cepas de *E. coli* con adherencia difusa (DAEC) expresan una serie de proteínas de unión pertenecientes a las familias de adhesinas afimbriales (Afa) y adhesinas fimbriales (Dr). Estas adhesinas se encuentran en la superficie de la membrana externa de la célula bacteriana, confiriéndole a la bacteria su principal mecanismo de patogenicidad (Hudault y col., 2004). Concretamente, las adhesinas Dr y F1845 se unen al receptor DAF (*brush border-associated decay-accelerating factor*), una glicoproteína que presenta un elevado nivel de expresión en la superficie apical de las células epiteliales polarizadas del intestino, además de en las células epiteliales del tracto genito-urinario y las células endoteliales. Después de la unión de la bacteria a la célula huésped, se induce el reordenamiento del citoesqueleto, destruyendo o reorganizando parcialmente las microvellosidades intestinales mediante la desarticulación de la actina F y la villina (Goluszko y col., 1999; Peiffer y col., 2000; Servin, 2005). Después de la unión de la bacteria al receptor DAF, se estimula la secreción de la citoquina proinflamatoria IL-8 a través de la estimulación flagelar del receptor 5 *Toll-like* (TLR5), lo que resulta en la activación de la proteína quinasa activada por mitógeno, las quinasas reguladas por señal extracelular 1 y 2 (ERK1/2), P38 y Jun-C quinasa (Bétis y col., 2003; Arikawa y Nishikawa, 2010). Después de la secreción de IL-8, los leucocitos polimorfonucleares migran a través de la barrera epitelial, promoviendo la producción de TNF- $\alpha$  y IL-1 $\beta$ , lo que provoca a su vez un incremento en la expresión de DAF y la consecuente cascada proinflamatoria que incrementará las lesiones en el tejido intestinal (Bétis y col., 2003).

### **A.I.6 *E. coli* adherente invasiva**

La etapa inicial de la patogénesis de las cepas de *E. coli* adherentes invasivas en el íleon requiere la adhesión de las bacterias a las células huésped. Para ello, las bacterias utilizan el pili tipo 1, para adherirse así al receptor CAECAM6a expresado por las células epiteliales del hospedador (Barnich y col., 2003; Barnich y col., 2007). Después de la adhesión a las células epiteliales, la bacteria AIEC produce las vesículas OMV mediante las cuales es capaz de penetrar proteínas efectoras en el interior de la célula hospedadora. Mediante la expresión de largas fibrillas polares que interactúan con las células M, AIEC es capaz de acceder a la lámina propia (Chassaing y col., 2011). Una vez allí, la bacteria

## 8. ANEXO INTRODUCCIÓN

tiene la capacidad de infectar y replicarse dentro de los fagolisosomas de los macrófagos sin inducir su muerte celular (Glasser y col., 2001; Bringer y col., 2005). La replicación continua de la bacteria en el interior de los macrófagos provoca la secreción de elevados niveles de TNF- $\alpha$ , provocando la inflamación intestinal y la formación de los granulomas presentes en los pacientes con la enfermedad de Crohn (Glasser y col., 2001).



## **9. ANEXO**



## 9.1 Capítulo 1: Los genes *hha2* y *hha3*, nuevos parálogos del gen *hha* en la cepa *E. coli* 042

Tabla 9.1. Listado de los oligonucleótidos utilizados durante el desarrollo del capítulo 1 de Resultados esta memoria.

Figura 9.2. Secuencia nucleotídica y lugar de hibridación de los oligonucleótidos utilizados para detectar los genes *hha*, *hha2*, *hha3* e *ydgT*.

Tabla 9.3. Tabla resumen de la colección de cepas de *E. coli* LREC utilizadas para analizar la presencia o ausencia de los genes *hha2* y *hha3*.

Tabla 9.4. Listado de las cepas enteroagregativas de *E. coli* de la colección LREC en las que se analizó la presencia de los 14 genes marcadores del patotipo EAEC, además de los genes *hha2* y *hha3*.

Tabla 9.5. Listado de las cepas de *E. coli* productoras de  $\beta$ -lactamasas de amplio espectro (ESBL) en las que se analizó tanto por PCR como bioinformáticamente la presencia de los alelos *hha2* y *hha3*.

Tabla 9.6. Listado de las cepas de *E. coli* de la colección ECOR donde se analizó por PCR la presencia de los genes *hha2* y *hha3*.

Tabla 9.7. Listado de cepas de *E. coli* de origen ambiental, productoras de la toxina Stx2, en las que se analizó la presencia de los genes *hha2* y *hha3*.

Figura 9.8. Representación esquemática de la localización de los genes *hha*, *hha2*, *hha3* e *ydgT* en cepas de *E. coli* pertenecientes a diferentes patotipos, cuyo genoma se encuentra totalmente secuenciado y ensamblado en la base de datos del NCBI.

Tabla 9.9. Listado de los genes mayormente sobreexpresados en el fondo genético  $\Delta hha$  respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 37°C.

Tabla 9.10. Listado de los genes mayormente sobreexpresados en el fondo genético  $\Delta hha\Delta hha2$  respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 37°C.

## 9. ANEXO

<b>Oligonucleótido</b>	<b>Secuencia (5'-3')</b>
4516 <i>forward</i>	5' CTTACCCTTGCGGAGCTGGA 3'
4516 <i>reverso</i>	5' TATTTCCAGACAGAAGCCGGAA 3'
4796 <i>forward</i>	5' GAGACGCGGGAAAAGTGAAT 3'
4796 <i>reverso</i>	5' CAGGTGACATATTTCCAGACGG 3'
<i>hha forward</i>	5' GTTTACGTCGTTGCAGACAA 3'
<i>hha reverso</i>	5' TTCCATACTGAGGAAGGGATCT 3'
<i>ydgT forward</i>	5' ATGTATCGTGCTGCCGATCA 3'
<i>ydgT reverso</i>	5' ATTGGACATAGTGCCAGACG 3'
<i>aggR-Fw</i>	5' CTAATTGTACAATCGATGTA 3'
<i>aggR-Rv</i>	5' ATGAAGTAATTCTTGAAT 3'
<i>aggA-Fw</i>	5' TTAGTCTTCTATCTAGGG 3'
<i>aggA-Rv</i>	5' AAATTAATTCCGGCATGG 3'
<i>aafA-Fw</i>	5' ATGTATTTTTAGAGGTTGAC 3'
<i>aafA-Rv</i>	5' TATTATATTGTCACAAGCTC 3'
<i>aap-Fw</i>	5' CTTTTCTGGCATCTTGGGT 3'
<i>aap-Rv</i>	5' GTAACAACCCCTTTGGAAGT 3'
<i>set1A-Fw</i>	5' TCACGCTACCATCAAAGA 3'
<i>set1A-Rv</i>	5' TATCCCCCTTTGGTGGTA 3'
<i>astA-Fw</i>	5' CCATCAACACAGTATATCCGA 3'
<i>astA-Rv</i>	5' GGTCGCGAGTGACGGCTTTGT 3'
<i>irp2-Fw</i>	5' AAGGATTCGCTGTTACCGGAC 3'
<i>irp2-Rv</i>	5' TCGTCGGGCAGCGTTTCTTCT 3'
<i>pet-Fw</i>	5' GACCATGACCTATACCGACAGC 3'
<i>pet-Rv</i>	5' CCGATTTCTCAAACCTCAAGACC 3'
<i>shf-Fw</i>	5' ACTTTCTCCCGAGACATTC 3'
<i>shf-Rv</i>	5' CTTTAGCGGGAGCATTCAT 3'
<i>agg-3A-Fw</i>	5' GTATCATTGCGAGTCTGGTATTCAG 3'
<i>agg-3A-Rv</i>	5' GGGCTGTTATAGAGTAACTTCCAG 3'
<i>aat-Fw</i>	5' CTGGCGAAAGACTGTATCAT 3'
<i>aat-Rv</i>	5' AATGTATAGAAATCCGCTGTT 3'
<i>agn43-Fw</i>	5' ACGCACAACCATCAATAAAA 3'
<i>agn43-Rv</i>	5' CCGCCTCCGATACTGAATGC 3'
<i>sat-Fw</i>	5' ACTGGCGGACTCATGCTGT 3'
<i>sat-Rv</i>	5' AACCTGTAAGAAGACTGAGC 3'
<i>pic-Fw</i>	5' TTCAGCGGAAAGACGAA 3'
<i>pic-Rv</i>	5' TCTGCGCATTCATACCA 3'
<i>hha p1</i>	5' ATGTCCGAAAAACCTTTAACGAAAACCGATTATTTAAT GCGTGTGTAGGCTGGAGCTGTCTTC 3'
<i>hha p2</i>	5' TTAGCGAATAAATTTCCATACTGAGGAAGGGATCTTGT CGTACATATGAATATCCTCCTTAGT 3'
<i>hha 3x p1</i>	5' GTACGACAAGATCCCTTCCTCAGTATGGAAATTTATTC GCGACTACAAAGACCATGACGG 3'
<i>hha 3x p2</i>	5' CAGCCGGTTATAGCTCCGAAAGCGAATTATTTATTAGC GACATATGAATATCCTCCTTAG 3'
<i>hha p1 up</i>	5' GTTAGAATTATTACAACCATGGG 3'

4516 p1	5'ATCACGTTAATTGCAGCATAAAGTGATGAGAGGCTAAT GGAAGTGTAGCTGGAGCTGCTTC 3'
4516 p2	5'AACGCCCGGCACAATACCGCTGTGATTAACGAACGTTT AGGTCATATGAATATCCTCCTTAGT 3'
4516 3x p1	5'TTACGATCGCGTTCCGGCTTCTGTCTGGAAATATGTTA CCGACTACAAAGACCATGACGG 3'
4516 3x p2	5'GCCCGGCACAATACCGCTGTGATTACCGAACGTTT TACATATGAATATCCTCCTTAG 3'
4516 p1 up	5' GTCTGAGCGGTAAACGTATC 3'
4796 p1	5'CGGAACACTAAGGCTGTATTGCGGGACACCTGGTAAG TCGCAGTGTAGGCTGGAGCTGCTTC 3'
4796 p2	5'GATAACACCGGAGTCATCGTAGCGCTGCTTCAATATTC AGGTCATATGAATATCCTCCTTAGT 3'
4796 3x p1	5'TTACGATCGCGCGCCGGCTTCCGTCTGGAAATATGTCA CCGACTACAAAGACCATGACGG 3'
4796 3x p2	5'AACACCGGAGTCATCGTAGCGCTGCTTCAATATTCAGG TGCATATGAATATCCTCCTTAG 3'
4796 p1 up	5' TCGTTCATGTATCAGAGTGC 3'
ydgT 3x p1	5'CCTCGGCCAGGTACCGAAGTCCGTCTGGCACTATGTCC AAGACTACAAAGACCATGACGG 3'
ydgT 3x p2	5'TAAGGCTTTTATAACCGATCGCTACTTTCTTTATTGGCA TATGAATATCCTCCTTAG 3'
ydgT p1 up	5' GCTCTACGCTTGTAGAGCGAC 3'
UP378	5' GCGTTTTGACCATCTTCGT 3'
2242 RT Fw	5' CCACGGATGTCATTGCTCAG 3'
2242 RT Rv	5' TTTCCTTTTCCAGGGTTGCCG 3'
4511 RT Fw	5' GGGGCCTTTAACTACACCCT 3'
4511 RT Rv	5' TTTACACCGGTCTGATGGCT 3'
4803 RT Fw	5' TGCCAGCGGTAAGTTCAGTA 3'
4803 RT Rv	5' CCTTTACTGAGCTGTGTTCTG 3'
shf RT Fw	5' CGCTGGTCTGAAGTGAACGA 3'
shf RT Rv	5' CGCTCAATAAGTCCTGCCGA 3'
gapA042 RT Fw	5' TTTCCGTGCTGCTCAGAAAC 3'
gapA042 RT Rv	5' GTC AACACCAACTTCGTCCC 3'
LP378	5'ACAGGAGCAGTTTCAGACAG 3'
KT	5'CGGCCACAGTCGATGAATCC 3'
hha2 plate51 NT fw	5'GGTGATGATGATGACAAGATGGAAGTTAAAACCAAAG AAGACTGGCTGTAT 3'
hha2 plate51 NT rv	5'GGAGATGGGAAGTCATTAGGTAACATATTTCCAGACA GAAGCCGGAAC 3'
hha2 plate31 CT fw	5'AGAAGGAGATATAACTATGGAAGTTAAAACCAAAGAA GACTGGCTGTATCAG 3'
hha2 plate31 CT rv	5'GTGGTGGTGTGATGGTGTGATGGCCGGTAACATATTTCCAGA CAGAAGCCGGAAC 3'

**Tabla 9.1.** Listado de los oligonucleótidos (anotación y secuencia nucleotídica de 5' a 3') utilizados durante el desarrollo del capítulo 1 de Resultados de esta memoria.

## 9. ANEXO

**Secuencia nucleotídica gen *hha* (EC042\_0498):**

5'-ATGTCGAAAAACCTTTAACAAAAACCGATTATTTAATGCGTTTACGTTCGTTGCCAGACAATTGACACGCTGGAGCGTGTTAT  
 CGAGAAAAATAAATACGAATTATCAGATAATGAACCTGGCGGTATTTACTCAGCCGAGATCACCGCCTCGCCGAATTGACCAT  
 GAATAAACTGTACGACAAGATCCCTTCCTCAGTATGGAAATTATTTCGCTAA -3'

$\xrightarrow{\text{hha forward}}$   
 $\xleftarrow{\text{hha reverso}}$

Tamaño amplicón: 166 pb

**Secuencia nucleotídica gen *ydgT* (EC042\_1794):**

5'-ATGACTGTTCCAGACTACTTATTAATAATTCGCAAAATCAGTTCACCTCGAAAGTCTGGAAAACTCTACGACCATCTTAATTAC  
 ACCCTGACGGACGATCAGGAACTGATCAATATGTATCGTCTGCCGATCACCGTCGCGCAGAGCTGGTTTCTGGCGGCCGTT  
 TGTTTGACCTCGGCCAGGTACCGAAGTCGCTGGCACTATGTCCAATAA -3'

$\xrightarrow{\text{ydgT forward}}$   
 $\xleftarrow{\text{ydgT reverso}}$

Tamaño amplicón: 100 pb

**Secuencia nucleotídica gen *hha2* (EC042\_4516):**

5'-ATGGAAGTTAAACCAAGAAGACTGGCTGTATCAGTTCGCTGTTGTTTCATCCGGGAGACGCTGGAAAAAGTATTCC  
 ACACACGTTATAAECTTACCCTTGCGGAGCTGGAAGCCTTCAACTCTGCGGTTGATCACCGACTGGCAGAACTAACCATGAACA  
 AACTTTACGATCGCGTTCGCGCTTCTGTCTGGAAATATGTTACCTGA -3'

$\xrightarrow{\text{4516 forward}}$   
 $\xleftarrow{\text{4516 reverso}}$

Tamaño amplicón: 107 pb

**Secuencia nucleotídica gen *hha3* (EC042\_4796):**

5'-GTGGCAGCCCGTAATACGTCAGGAGACGCGGAAAAAGTGAATTCCACACGCGTTATAAACTTACCCTTGCGGAGCTGGAA  
 GCCTTAACTCTGCCGTTGACAACCGGCTGGCAGAACTGACAATGAACAAAATTACGATCGCGCGCGGCTTCCGCTCTGGAA  
 ATATGTCACCTGA -3'

$\xrightarrow{\text{4796 forward}}$   
 $\xleftarrow{\text{4796 reverso}}$

Tamaño amplicón: 154 pb

**Figura 9.2.** Secuencia nucleotídica y lugar de hibridación de los oligonucleótidos utilizados para detectar por PCR los genes *hha*, *ydgT*, *hha2* y *hha3*. También se indica el tamaño del producto obtenido al amplificar por PCR cada uno de los genes con los oligonucleótidos específicos.

Patotipo	Código LREC	Serotipo	Sequence type	Grupo filogenético	hha2	hha3
EAEC	1	O3:H2			+	+
EAEC	2	O3:H2			+	+
EAEC	3	O11:H-			-	-
EAEC	4	O17:H-			+	+
EAEC	5	O44:H18			+	+
EAEC	6	ONT:HNT			+	+
EAEC	7	O3:H2			+	+
EAEC	8	O15:H-			-	-
EAEC	9	O17:H18			+	-
EAEC	10	O128:HNT			-	-
EAEC	11	O104:H4			+	-
EAEC	12	O86:H2			-	-
EAEC	13	O17:H18			+	+
EAEC	14	O11:H18			+	+
EAEC	15	ONT:H27			-	-
EAEC	101	O99:H4			-	-
ExPEC	17	O17:H18	ST597	D	+	-
ExPEC	18	O7:H18	ST62	F	-	-
ExPEC	19	O6:H1	ST73	B2	+	+
ExPEC	20	O6:H31	ST127	B2	-	-
ExPEC	21	O4:H5	ST12	B2	+	-
ExPEC	22	O45A:H7	ST95	B2	-	-
ExPEC	23	O18:H7	ST95	B2	-	-
ExPEC	24	O1:H7	ST95	B2	-	-
ExPEC	26	O25b:H4	ST131	B2	-	+
ExPEC	27	O25b:H4	ST131	B2	+	+
ExPEC	28	O25b:H4	ST131	B2	+	+
ExPEC	29	O25b:H4	ST131	B2	-	-
ExPEC	30	O25b:H4	ST131	B2	+	+
ExPEC	31	O16:H5	ST131	B2	+	-
STEC/VTEC	32	O174:H8			+	+
STEC/VTEC	33	O55:H7			+	-
STEC/VTEC	35	O177:[H25]			-	-
STEC/VTEC	36	O111:[H8]			-	-
STEC/VTEC	37	O113:H4			+	+
STEC/VTEC	38	O103:H2			-	-
STEC/VTEC	39	O26:H11			+	+
STEC/VTEC	40	O41:H26			-	-
STEC/VTEC	41	O157:H7			-	-
STEC/VTEC	42	O157:H7			-	-
STEC/VTEC	43	O174:H8			+	+
STEC/VTEC	44	O8:H19			-	-
STEC/VTEC	45	O48:H21			-	-



9. ANEXO

STEC/VTEC	46	O118:H12			+	+
STEC/VTEC	47	O174:H21			+	-
STEC/VTEC	48	O73:H18			-	+
STEC/VTEC	49	O139:H1,H38			+	-
STEC/VTEC	50	O128:HNM			-	-
STEC/VTEC	51	O2: H25			-	-
STEC/VTEC	55	O113			+	+
STEC/VTEC	57	O121:H19			-	-
ETEC	52	O8:H9			+	-
ETEC	58	O78:H11			-	+
ETEC	60	O115:H40			+	-
tEPEC	53	O26:H11			+	+
tEPEC	34	O128ac:[H2]			-	-
EIEC	56	O96:H19			-	-
EIEC	59	O124:H30			-	-
AIEC	25	O83:H1	ST135	B2	-	-
Comensal	16	O16:H48	ST98	A1	-	-

Patotipo	
EAEC	<i>enteroaggregative E. coli</i>
ExPEC	<i>extraintestinal pathogenic E. coli</i>
STEC/VTEC	<i>Shiga toxin (verotoxin) producing E. coli</i>
ETEC	<i>enterotoxigenic E. coli</i>
tEPEC	<i>typical enteropathogenic E. coli</i>
EIEC	<i>enteroinvasive E. coli</i>
AIEC	<i>adherent-invasive E. coli</i>

**Tabla 9.3.** Tabla resumen de la colección de cepas LREC (Laboratorio de Referencia de *E. coli*) divididas por patotipos. Se indica el serotipo de cada cepa, su secuencia tipo, el grupo filogenético al que pertenece cada cepa y el resultado del análisis por PCR realizado para analizar la presencia o ausencia de los genes *hha2* y *hha3*.

Código LREC	Serotipo	Grupo Filogenético	Genes de virulencia de <i>Escherichia coli</i> enteroagregativa (EAEC)																	
			Cepa	O:H		<i>aatA</i>	<i>aggA</i>	<i>aafA</i>	<i>agg3A</i>	<i>aggR</i>	<i>astA</i>	<i>aap</i>	<i>set1A</i>	<i>irp2</i>	<i>pet</i>	<i>shf</i>	<i>sat</i>	<i>pic</i>	<i>agn43</i>	<i>hha2</i>
66	O85:HNT	A		+	-	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-
71	O85:H10	A		+	-	-	-	+	-	+	-	+	-	-	-	-	+	-	-	-
87	O3:H2	A		+	+	-	-	+	+	+	+	+	+	-	+	+	+	-	-	-
95	O15:H-	A		+	-	-	-	+	-	+	+	+	+	-	-	+	-	-	-	-
82	O15:H2	A		+	-	-	-	+	-	+	+	+	+	-	-	+	-	-	-	-
79	O86:H2	A		+	-	-	+	+	-	+	+	+	+	-	-	+	-	-	-	-
80	O86:H2	A		+	-	-	+	+	-	+	+	+	+	-	-	+	-	-	-	-
89	O86:H2	A		+	-	-	+	+	-	+	+	+	+	-	-	+	-	-	-	-
90	O86:H2	A		+	-	-	+	+	-	+	+	+	+	-	-	+	-	-	-	-
12	O86:H2	A		+	-	-	+	+	-	+	+	+	+	-	-	+	-	-	-	-
101	O99:H4	A		+	-	-	-	+	-	+	+	+	+	-	-	+	-	-	-	-
81	O114:H2	A		+	-	-	+	+	-	+	+	+	+	-	-	+	-	-	-	-
3	O11:H-	B1		+	-	+	-	+	+	+	+	+	+	+	-	+	+	-	-	-
88	O86:H27	B1		+	-	-	-	+	-	+	+	+	+	+	+	+	+	-	-	-
10	O128:HNT	B1		+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
100	ONT:HNT	B2		+	-	-	-	+	-	+	-	+	-	-	-	-	+	-	-	-
67	O86:H11	D		+	-	-	-	+	-	+	-	+	-	-	-	-	-	-	-	-
83	O3:H2	A		+	+	-	-	+	+	+	+	+	+	-	+	+	+	+	+	-
8	O15:H-	A		+	-	-	-	-	-	-	+	+	+	-	-	+	-	+	-	-
74	O125:H2	A		+	-	-	-	+	-	+	+	+	+	-	-	+	+	+	+	-
91	ONT:H33	A		+	-	-	-	+	-	+	+	+	+	-	+	+	+	+	+	-
92	ONT:H33	A		+	-	-	-	+	-	+	+	+	+	-	+	+	+	+	+	-
96	ONT:H33	A		+	-	-	-	+	-	+	+	+	+	-	+	+	+	+	+	-

9. ANEXO

98	ONT:H33	A	+	-	-	-	+	-	+	+	+	+	-	+	+	+	+	-
93	O104:H4	B1	+	-	-	+	+	+	+	+	+	+	+	-	+	-	+	-
94	O104:H4	B1	+	-	-	+	+	+	+	+	+	+	+	-	+	+	+	-
76	O111:H-	B1	+			-	+	+	+	+	+	+	-	+	+	-	+	-
99	ONT:HNT	A	+	-	-	-	-	+	+	-	+	+	-	-	-	-	-	+
84	O86:H27	B1	+	-	-	-	+	-	+	+	+	+	+	+	+	+	-	+
13	O17:H18	A	+	-	-	-	+	-	+	+	-	+	+	-	+	+	+	+
1	O3:H2	A	+	+	-	-	+	+	+	-	+	-	+	-	-	+	+	+
2	O3:H2	A	+	+	-	-	+	+	+	-	+	-	-	-	-	+	+	+
77	O3:H2	A	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+
78	O3:H2	A	+	+	-	-	+	+	+	+	+	+	-	-	+	+	+	+
7	O3:H2	A	+	+	-	-	+	+	+	+	+	+	-	+	+	+	+	+
61	O15:H2	A	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+
69	O78:H2	A	+	+	-	-	+	-	+	+	+	+	+	-	+	-	+	+
70	O78:H10	A	+	+	-	-	+	+	+	+	+	+	-	+	+	-	+	+
75	O92:H33	A	+	-	-	-	+	-	+	+	+	+	-	+	+	+	+	+
62	O168:H4	A	+	+	-	-	+	+	+	-	-	+	-	+	-	-	+	+
85	O168:H4	A	+	+	-	-	+	+	+	-	-	+	-	+	-	-	+	+
68	O61:H4	B1	+	-	-	-	+	-	+	+	+	+	+	+	+	+	+	+
11	O104:H4	B1	+	-	-	+	+	+	+	+	+	+	+	-	+	+	+	+
72	ONT:H28	B1	+	-	+	-	+	+	+	+	+	+	+	-	+	+	+	+
73	ONT:H28	B1	+	-	+	-	+	+	+	+	+	+	+	-	+	+	+	+
6	ONT:HNT	B1	+	-	+	-	+	+	+	+	-	+	+	-	+	+	+	+
63	O17:H2	B2	+	+	-	-	+	-	+	+	+	+	+	+	+	+	+	+
14	O11:H18	D	+	+	-	-	+	-	+	+	+	+	+	+	+	+	+	+
64	O15:H18	D	+	-	-	+	+	-	+	+	+	+	-	+	+	+	+	+
65	O15:H-	D	+	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+

		Genes de virulencia de <i>Escherichia coli</i> enteroagregativa (EAEC)																
		<i>aatA</i>	<i>aggA</i>	<i>aafA</i>	<i>agg3A</i>	<i>aggR</i>	<i>astA</i>	<i>aap</i>	<i>set1A</i>	<i>irp2</i>	<i>pet</i>	<i>shf</i>	<i>sat</i>	<i>pic</i>	<i>agn43</i>	<i>hha2</i>	<i>hha3</i>	
TOTAL: 56 cepas EAEC		56	13	6	12	50	22	52	44	49	48	23	23	44	36	37	29	
		-	23%	11%	21%	89%	39%	93%	79%	88%	86%	41%	41%	79%	64%	66%	56%	
17 cepas <i>hha2-/hha3-</i>		17	<b>1</b>	1	6	15	<b>2</b>	16	12	16	12	<b>3</b>	<b>2</b>	12	<b>5</b>	-	-	
		-	<b>6%</b>	6%	35%	88%	<b>12%</b>	94%	71%	94%	71%	<b>18%</b>	<b>12%</b>	71%	<b>29%</b>	-	-	
10 cepas <i>hha2+/hha3-</i>		10	<b>1</b>	0	2	9	<b>4</b>	9	10	10	10	<b>2</b>	<b>6</b>	10	<b>7</b>	10	-	
		-	<b>10%</b>	0	20%	90%	<b>40%</b>	90%	100%	100%	100%	<b>20%</b>	<b>60%</b>	100%	<b>70%</b>	18%	-	
2 cepas <i>hha2-/hha3+</i>		2	<b>0</b>	0	0	1	<b>1</b>	2	1	2	2	<b>1</b>	<b>1</b>	1	<b>1</b>	-	2	
		-	<b>0</b>	0	0	50%	<b>50%</b>	100%	50%	100%	100%	<b>50%</b>	<b>50%</b>	50%	<b>50%</b>	-	3,5%	
27 cepas <i>hha2+/hha3+</i>		27	<b>11</b>	5	4	25	<b>15</b>	25	21	21	24	<b>17</b>	<b>14</b>	21	<b>23</b>	27	27	
		-	<b>41%</b>	19%	15%	93%	<b>56%</b>	93%	78%	78%	89%	<b>63%</b>	<b>52%</b>	78%	<b>85%</b>	48%	48%	
Test de exactitud de Fisher	P valor	<i>hha2-/hha3-</i> vs <i>hha2+/hha3+</i>					<b>0.004</b>					<b>0.004</b>	<b>0.007</b>			<b>0.0003</b>		
Test de exactitud de Fisher	P valor	<i>hha2-/hha3-</i> vs <i>hha2+</i> y/o <i>hha3+</i>					<b>0.005</b>					<b>0.02</b>	<b>0.003</b>			<b>0.0005</b>		

**Tabla 9.4.** Resultados de los análisis por PCR de la distribución, en 56 cepas EAEC de la colección LREC, de los genes *hha2/hha3* (marcados en verde) y los 14 genes que codifican para factores de virulencia típicos de EAEC (marcados en amarillo). También se indica el serotipo y el grupo filogenético al que pertenece cada una de las cepas analizadas. Los resultados del test exacto de Fisher se muestran cuando  $P < 0.05$ . Las diferencias significativas se muestran en negrita.

## 9. ANEXO

Aislamiento	Origen	Material	gen ESBL	Patotipo	Grupo filogenético	Sequence type	<i>hha2</i> PCR	<i>hha2</i> in silico	<i>hha3</i> PCR	<i>hha3</i> in silico
H10	humano	orina	<i>CTX-M-15</i>	NA*	B1	ST-443	+	+	+	+
H38	humano	swab rectal	<i>CTX-M-15</i>	ExPEC	A	ST-44	-	-	+	+
H75	humano	orina	<i>CTX-M-15</i>	ExPEC	B2	ST-131	+	-	+	+#
H89	humano	herida	<i>CTX-M-15</i>	NA*	D	ST-349	-	-	+	+
H99	humano	herida	<i>CTX-M-1</i>	ExPEC	A	ST-410	-	-	-	-
H127	humano	orina	<i>CTX-M-15</i>	EAEC	B2	ST-131	+	+	+	+#
H130	humano	swab cervical	<i>CTX-M-15</i>	ExPEC	A	ST-88	+	+	+	+#
H134	humano	orina	<i>CTX-M-15</i>	ExPEC	B1	ST-58	-	-	+	+
H136	humano	orina	<i>CTX-M-1</i>	ExPEC	D	ST-2509	-	-	-	-
H141	humano	orina	<i>CTX-M-15</i>	ExPEC	D	ST-2141	-	-	+	+
H154	humano	swab rectal	<i>CTX-M-15</i>	Flora intestinal	B1	ST-224	-	-	-	-
H162	humano	orina	<i>CTX-M-1</i>	NA*	A	ST-10	+	+	-	-
V1	perro	orina	<i>CTX-M-1</i>	EAEC	D	ST-354	-	-	+	+
V8	caballo	heces	<i>CTX-M-1</i>	Flora intestinal	D	ST-38	-	+	-	+
V41	perro	cavidad abdominal	<i>CTX-M-1</i>	ExPEC	A	ST-10	+	+	+	-
V63	perro	útero	<i>CTX-M-1</i>	ExPEC	B1	ST-162	-	+	-	-
V64	perro	útero	<i>CTX-M-1</i>	ExPEC	B1	ST-162	-	-	-	-
V89	caballo	heces	<i>CTX-M-1</i>	Flora intestinal	A	ST-10	-	-	-	-
V98	caballo	heces	<i>CTX-M-15</i>	Flora intestinal	A	ST-361	-	-	+	+
V105	perro	orina	<i>CTX-M-1</i>	ExPEC	A	ST-3476	-	-	+	+
V116	perro	herida	<i>CTX-M-1</i>	NA*	A	ST-10	-	-	-	-
V177	perro	hígado	<i>CTX-M-15</i>	ExPEC	A	ST-410	+	+	+	+
V182	caballo	heces	<i>CTX-M-15</i>	Flora intestinal	A	ST-10	-	-	+	+
V205	gato	orina	<i>CTX-M-1</i>	ExPEC	B1	ST-224	-	-	-	-
V292	perro	catéter urinario	<i>CTX-M-15</i>	ExPEC	D	ST-405	+	+	-	-

\*NA: no se ha podido asociar ningún patotipo a la cepa de *E. coli* #presencia de homólogos con una identidad y un *coverage* inferiores al 90% y 80%

**Tabla 9.5.** Listado de las cepas de *E. coli* productoras de  $\beta$ -lactamasas de amplio espectro (ESBL) analizadas durante este trabajo. Se indica el origen de los diferentes aislamientos, el tipo de gen ESBL que codifican, el patotipo y el grupo filogenético al que pertenecen las diferentes cepas y su secuencia tipo. También se indica el resultado de la PCR realizada para identificar la presencia/ausencia de los alelos *hha2/hha3* y el resultado de la comparativa al realizar el análisis *in silico* mediante unos parámetros de identidad y *coverage* superiores al 90 y 80%, respectivamente.

Cepa	Patotipo	<i>hha2</i>	<i>hha3</i>	Cepa	Patotipo	<i>hha2</i>	<i>hha3</i>	Cepa	Patotipo	<i>hha2</i>	<i>hha3</i>
EC01	Comensal	-	-	EC26	Comensal	-	-	EC51	ExPEC	+	+
EC02	ExPEC	-	-	EC27	Comensal	-	-	EC52	ExPEC	-	-
EC03	Comensal	-	+	EC28	Comensal	-	-	EC53	ExPEC	-	-
EC04	Comensal	-	-	EC29	Comensal	-	-	EC54	ExPEC	-	-
EC05	Comensal	+	-	EC30	Comensal	-	-	EC55	ExPEC	-	+
EC06	Comensal	-	-	EC31	Comensal	-	-	EC56	ExPEC	-	+
EC07	Comensal	-	-	EC32	Comensal	-	-	EC57	ExPEC	+	-
EC08	ExPEC	+	-	EC33	Comensal	-	-	EC58	Comensal	-	-
EC09	Comensal	-	+	EC34	Comensal	-	-	EC59	Comensal	+	-
EC10	Comensal	-	+	EC35	ExPEC	-	-	EC60	ExPEC	+	+
EC11	ExPEC	+	-	EC36	ExPEC	-	-	EC61	Comensal	-	-
EC12	Comensal	-	-	EC37	ExPEC	-	+	EC62	ExPEC	-	-
EC13	Comensal	-	-	EC38	ExPEC	-	-	EC63	ExPEC	-	-
EC14	Comensal	-	-	EC39	ExPEC	-	-	EC64	ExPEC	-	-
EC15	Comensal	-	-	EC40	ExPEC	-	-	EC65	Comensal	+	-
EC16	Comensal	-	+	EC41	ExPEC	-	-	EC66	ExPEC	-	-
EC17	Comensal	-	-	EC42	Comensal	-	-	EC67	Comensal	-	-
EC18	Comensal	-	-	EC43	Comensal	+	-	EC68	Comensal	-	-
EC19	Comensal	-	-	EC44	Comensal	+	-	EC69	Comensal	-	-
EC20	ExPEC	-	-	EC45	Comensal	-	-	EC70	Comensal	-	-
EC21	Comensal	-	-	EC46	ExPEC	-	-	EC71	Comensal	-	-
EC22	Comensal	-	-	EC47	Comensal	-	-	EC72	ExPEC	-	+
EC23	Comensal	-	-	EC48	ExPEC	-	-				
EC24	ExPEC	-	-	EC49	ExPEC	-	-				
EC25	Comensal	-	-	EC50	ExPEC	+	+				

**Tabla 9.6.** Listado de las 72 cepas de *E. coli* pertenecientes a la colección ECOR analizadas en este trabajo. Se indica el patotipo correspondiente a cada una de las cepas estudiadas y el resultado del análisis por PCR realizado para identificar la presencia/ausencia de los alelos *hha2* y *hha3*.

9. ANEXO

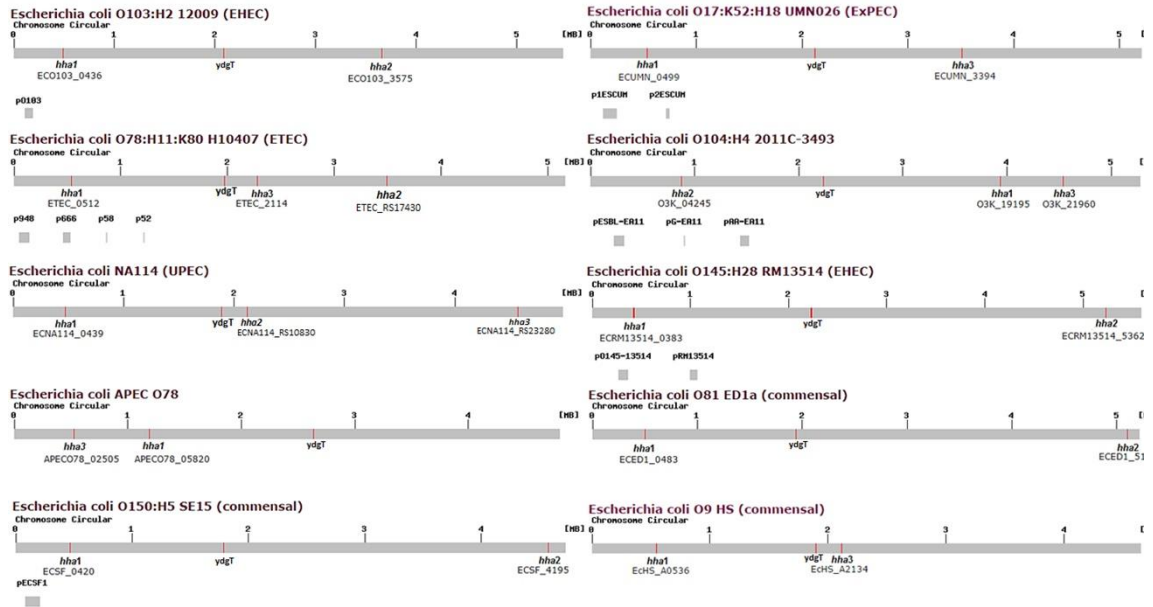
Aislamiento	Origen	Serotipo	<i>stx</i> <sub>1</sub>	<i>stx</i> <sub>2</sub>	<i>eae</i>	<i>hha</i> <sub>2</sub>	<i>hha</i> <sub>3</sub>	Aislamiento	Origen	Serotipo	<i>stx</i> <sub>1</sub>	<i>stx</i> <sub>2</sub>	<i>eae</i>	<i>hha</i> <sub>2</sub>	<i>hha</i> <sub>3</sub>
41 (115)	Ganado	O1:H20	-	<i>stx</i> <sub>2a</sub>	-	-	-	104 <i>stx</i> <sub>2-r</sub>	Ganado	O113:H21	-	<i>stx</i> <sub>2a</sub>	-	-	-
57 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	-	-	105 <i>stx</i> <sub>2</sub> ME	Ganado	O157:H7	-	<i>stx</i> <sub>2c</sub>	+	-	-
58 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	-	106 <i>stx</i> <sub>2</sub> ME	Ganado	O171:H2	-	<i>stx</i> <sub>2a</sub> , <i>stx</i> <sub>2c</sub>	-	+	-
58 <i>stx</i> <sub>2</sub> TAR 43	Ganado	O101:H9	-	<i>stx</i> <sub>2a</sub>	-	-	-	107 <i>stx</i> <sub>2-r</sub>	Ganado	O2:H25	-	<i>stx</i> <sub>2g</sub>	-	-	-
59 <i>stx</i> <sub>2</sub> ME-k2	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	-	108 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	-
60 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2a</sub>	-	+	+	109 <i>stx</i> <sub>2-r</sub>	Ganado	O113:H21	-	<i>stx</i> <sub>2a</sub>	-	+	-
62 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	+	110 <i>stx</i> <sub>2</sub> ME	Ganado	O157:H-	+	<i>stx</i> <sub>2c</sub>	+	-	-
63 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	-	111 <i>stx</i> <sub>2</sub> ME	Ganado	O157:H-	-	<i>stx</i> <sub>2c</sub>	+	+	-
64 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	-	-	113 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	-	-
65 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2a</sub>	-	-	-	115 <i>stx</i> <sub>2-r</sub>	Ganado	O91:H21	-	<i>stx</i> <sub>2a</sub> , <i>stx</i> <sub>2c</sub>	-	-	-
66 <i>stx</i> <sub>2</sub> -ME-8	Ganado	O181:H49	-	<i>stx</i> <sub>2</sub>	-	-	-	116 <i>stx</i> <sub>2-r</sub>	Ganado	O76:H2	+	<i>stx</i> <sub>2c</sub>	-	+	-
67 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	-	124	Ganado	ONT:H2	-	<i>stx</i> <sub>2c</sub> , <i>stx</i> <sub>2d</sub>	-	+	-
68 <i>stx</i> <sub>2-r</sub>	Ganado	O89:H19	-	<i>stx</i> <sub>2c</sub>	-	-	-	125-14	Ganado	O1:H20	-	<i>stx</i> <sub>2a</sub>	-	-	-
69 <i>stx</i> <sub>2-r</sub>	Ganado	O181:H49	-	<i>stx</i> <sub>2a</sub>	-	+	-	126-2	Ganado	O22:H8	+	<i>stx</i> <sub>2c</sub>	-	+	-
70 <i>stx</i> <sub>2</sub> ME-16	Ganado	O162:H7	-	<i>stx</i> <sub>2</sub> , <i>stx</i> <sub>2c</sub>	-	-	-	O3	Ganado	O1:H20	-	<i>stx</i> <sub>2a</sub>	-	-	-
71 <i>stx</i> <sub>2-r</sub>	Ganado	O28:H28	-	<i>stx</i> <sub>2c</sub>	+	-	-	4	Humano	O90:H-	-	<i>stx</i> <sub>2d</sub>	-	+	-
Ex E72	Ganado	O181:H20	-	<i>Stx</i> <sub>2a</sub>	-	+	-	7	Humano	O8:H9	-	<i>stx</i> <sub>2e</sub>	-	+	-
72 <i>stx</i> <sub>2-r</sub>	Ganado	O181:H20	-	<i>stx</i> <sub>2c</sub>	-	+	-	8 <i>stx</i> <sub>2-r</sub>	Humano	O146:H-	-	<i>stx</i> <sub>2a</sub>	-	-	-
73 <i>stx</i> <sub>2-r</sub>	Ganado	O181:H20	-	<i>stx</i> <sub>2c</sub>	-	+	+	51 <i>stx</i> <sub>2-r</sub>	Humano	O146:H21	+	<i>stx</i> <sub>2d</sub>	-	+	+
74 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	+	52 <i>stx</i> <sub>2-r</sub>	Humano	O171:H2	-	<i>stx</i> <sub>2c</sub> , <i>stx</i> <sub>2d</sub>	-	-	-
82 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	-	55 <i>stx</i> <sub>2-r</sub>	Humano	O127:H-	-	<i>stx</i> <sub>2a</sub> , <i>stx</i> <sub>2d</sub>	-	+	-
83 <i>stx</i> <sub>2-r</sub>	Ganado	O2:H25	-	<i>stx</i> <sub>2a</sub>	-	+	-	77 <i>stx</i> <sub>2-r</sub>	Humano	O90:H-	-	<i>stx</i> <sub>2d</sub>	-	+	-
84 <i>stx</i> <sub>2-r</sub>	Ganado	O156:H8	-	<i>stx</i> <sub>2c</sub>	-	-	-	78 <i>stx</i> <sub>2-r</sub>	Humano	O166:H21	+	<i>stx</i> <sub>2c</sub>	-	+	+
84b <i>stx</i> <sub>2-r</sub>	Ganado	O171:H-	-	<i>stx</i> <sub>2a</sub>	-	-	-	79 <i>stx</i> <sub>2-r</sub>	Humano	O90:H-	+	<i>stx</i> <sub>2d</sub>	-	+	-
85 <i>stx</i> <sub>2-r</sub>	Ganado	O171:H2	-	<i>stx</i> <sub>2c</sub>	-	+	-	80 <i>stx</i> <sub>2-r</sub>	Humano	ONT:H9	-	<i>stx</i> <sub>2a</sub> , <i>stx</i> <sub>2e</sub>	-	-	-
86 <i>stx</i> <sub>2-r</sub>	Ganado	O2:H25	-	<i>stx</i> <sub>2g</sub>	-	-	-	81 <i>stx</i> <sub>2-r</sub>	Humano	ONT:H2	-	<i>stx</i> <sub>2a</sub>	-	+	-
87 <i>stx</i> <sub>2-r</sub>	Ganado	O2:H25	-	<i>tx</i> <sub>2a</sub>	-	+	-	118-2	Humano	ONT:H51	-	<i>stx</i> <sub>2a</sub>	-	-	-
88 <i>stx</i> <sub>2-r</sub>	Ganado	O136:H1,H12	-	<i>stx</i> <sub>2g</sub>	-	-	-	121	Humano	O8:H9	-	<i>stx</i> <sub>2e</sub>	-	-	-
<i>Ec</i> 88	Ganado	O157:H7	-	<i>stx</i> <sub>2a</sub>	+	-	-	122-12 <i>stx</i> <sub>2</sub>	Humano	O8:H9	-	<i>stx</i> <sub>2a</sub>	-	+	-



89 stx2-r	Ganado	O2:H25	-	<i>stx2g</i>	-	-	-	42 stx2-r	Mixto	ONT:H48	-	<i>stx2a</i>	-	-	-
90 stx2-r	Ganado	O2:H25	-	<i>stx2g</i>	-	+	-	49 stx2-r	Mixto	O26:H-	-	<i>stx2a, stx2e</i>	-	-	-
91 stx2-r	Ganado	O2:H25	-	<i>stx2g</i>	-	-	-	133stx2-r	Mixto	O90:H-	+	<i>stx2d</i>	-	+	-
92 stx2-r	Ganado	O2:H25	-	<i>stx2g</i>	-	-	-	48 stx2-r	Mixto	ONT:H21	-	<i>stx2e</i>	-	+	-
93 stx2-r	Ganado	O2:H25	-	<i>stx2g</i>	-	+	-	134 stx2-r	Mixto	O2:H25	-	<i>stx2g</i>	-	-	-
95 stx2-r	Ganado	O2:H25	-	<i>stx2g</i>	-	-	-	135stx2Fr212	Mixto	O9:H-	-	<i>stx2a</i>	-	-	-
96 stx2 ME	Ganado	ONT:H-	-	<i>stx2c</i>	-	-	-	140stx2Fr326	Mixto	ONT:H-	-	<i>stx2e</i>	-	+	-
98 stx2-r	Ganado	O136:H1	-	<i>stx2g</i>	-	+	-	141stx2Fr329	Mixto	ONT:H-	-	<i>stx2e</i>	-	-	-
99 stx2-r	Ganado	ONT:HNT	-	<i>stx2c</i>	-	-	-	75 stx2-r	Cerdo	O2:H21	-	<i>stx2e</i>	-	+	-
100 stx2-r	Ganado	O2:H25	-	<i>Stx2a</i>	-	+	-	76 stx2-r	Cerdo	O54:H21	-	<i>stx2e</i>	-	-	-
101 stx2-r	Ganado	O157:H7	-	<i>stx2c</i>	+	-	-	148stx2Fr1.2	Cerdo	O100:H-	-	<i>stx2e</i>	-	-	-
102 stx2 ME	Ganado	O8:H31	-	<i>stx2g</i>	-	+	-	149a stx2	Cerdo	O8:H-	-	<i>stx2e</i>	-	+	-
103 stx2-r	Ganado	O171:H2	-	<i>stx2a</i>	-	-	-	150 stx2-r	Cerdo	O26:H-	-	<i>stx2e</i>	-	-	-

**Tabla 9.7.** Listado de las cepas ambientales de *E. coli* analizadas en este trabajo, aisladas a partir de aguas residuales sin tratar de origen urbano (en su mayoría contaminadas por desechos fecales humanos) o de muestras de aguas residuales provenientes de tres mataderos diferentes (ganado, cerdo y un matadero mixto de ganado vacuno, cordero y cabra). Se indica el origen de cada uno de los aislamientos, el serotipo de cada una de las cepas analizadas, la producción o no de la toxina Stx1, el tipo de gen que codifica para la toxina Stx2, la presencia/ausencia del gen *eae* (gen que codifica para la intimina bacteriana) y el resultado del análisis de la PCR realizada para identificar la presencia/ausencia de los alelos *hha2* y *hha3*.

9. ANEXO



**Figura 9.8.** Representación esquemática de la localización de los genes *hha*, *hha2*, *hha3* e *ydgT* en cepas de *E. coli* pertenecientes a diferentes patotipos cuyos genomas se encuentran totalmente secuenciados y ensamblados en la base de datos del NCBI.

<i>E. coli</i> 042 $\Delta$ <i>hha</i> vs <i>E. coli</i> 042 WT		
<b>Identificador</b>	<b>Descripción</b>	<b>Fold change</b>
EC042_RS06355	hypothetical protein	26,0
EC042_RS01250	hypothetical protein	20,1
EC042_RS20085	hypothetical protein	14,2
EC042_RS03935	hypothetical protein	13,9
EC042_RS14470	hypothetical protein	13,8
EC042_RS19170	hypothetical protein	13,6
EC042_RS21810	broad-spectrum mercury transporter MerE	12,2
EC042_RS09940	hypothetical protein	11,7
EC042_RS24485	IS66 family transposase	11,5
EC042_RS15485	hypothetical protein	11,0
EC042_RS27060	nitrite extrusion protein 2	10,6
EC042_RS16305	hypothetical protein	10,5
EC042_RS06730	hypothetical protein	10,4
EC042_RS01500	hypothetical protein	10,2
EC042_RS03035	hypothetical protein	10,1
EC042_RS03040	hypothetical protein	10,1
EC042_RS25285	tRNA-Leu	10,0
EC042_RS08665	hypothetical protein	10,0
EC042_RS25735	hypothetical protein	10,0
EC042_RS21265	hypothetical protein	9,9
EC042_RS12550	hypothetical protein	9,8
EC042_RS17190	hypothetical protein	9,7
EC042_RS16360	invasion protein invH	9,7
EC042_RS16320	hypothetical protein	9,6
EC042_1722	cell division inhibition protein DicB	9,5
EC042_RS14290	hypothetical protein	9,5
EC042_RS26735	hypothetical protein	9,2
EC042_RS03940	hypothetical protein	9,1
EC042_RS11620	tRNA-Arg	8,9
EC042_RS16950	tRNA-Phe	8,8
EC042_RS23700	hypothetical protein	8,8
EC042_RS01830	hypothetical protein	8,8
EC042_RS20080	hypothetical protein	8,7
EC042_RS26640	hypothetical protein	8,7
EC042_RS17060	hypothetical protein	8,7
EC042_RS09020	hypothetical protein	8,6
EC042_0751	acyl-CoA synthetase	8,6
EC042_RS03695	hypothetical protein	8,5
EC042_RS25655	hypothetical protein	8,5
EC042_RS07960	hypothetical protein	8,5
EC042_RS25660	hypothetical protein	8,5
EC042_RS08415	hypothetical protein	8,5

## 9. ANEXO

EC042_RS17175	hypothetical protein	8,5
EC042_3897	membrane protein	8,4
EC042_3060	helix-turn-helix transcriptional regulator	8,3
EC042_3335	fimbrial-like protein	8,3
EC042_0024	fimbrial protein	8,3
EC042_3062	EscT/YscT/HrcT family type III secretion system	8,2
EC042_RS20145	hypothetical protein	8,2
EC042_RS14520	hypothetical protein	8,2
EC042_4011	HilA family transcriptional regulator EilA	8,1
EC042_RS10655	hypothetical protein	8,1
EC042_RS05120	hypothetical protein	8,1
EC042_RS16555	hypothetical protein	8,0
EC042_RS25640	antirestriction protein	8,0
EC042_4074	type III effector protein	7,9
EC042_RS26395	transposase	7,9
EC042_RS08325	hypothetical protein	7,9
EC042_1599	leucine-rich repeat domain-containing protein	7,9
EC042_0642	LysR family transcriptional regulator	7,6
EC042_RS15625	hypothetical protein	7,6
EC042_RS11150	integrase	7,5
EC042_RS09055	cold shock-like protein CspF	7,5
EC042_RS25625	hypothetical protein	7,5
EC042_RS05810	curli assembly protein CsgF	7,4
EC042_RS26165	hypothetical protein	7,4
EC042_RS09025	cold-shock protein cspI	7,4
EC042_RS01495	flagellin	7,4
EC042_RS24125	Rha family transcriptional regulator	7,3
EC042_RS08605	Putative ABC transporter, ATP-binding/permease protein	7,3
EC042_RS11335	hypothetical protein	7,3
EC042_RS06925	hypothetical protein	7,3
EC042_RS10985	hypothetical protein	7,2
EC042_RS05825	hypothetical protein	7,2
EC042_RS13775	hypothetical protein	7,2
EC042_RS02970	FimA homolog	7,2
EC042_RS16345	Type III secretion system protein	7,2
EC042_RS20195	helix-turn-helix transcriptional regulator	7,2
EC042_RS15125	molybdopterin-guanine dinucleotide biosynthesis protein MobA	7,2
EC042_RS25405	hypothetical protein	7,2
EC042_RS06915	cell division inhibitor	7,2
EC042_RS00740	Putative fimbrial protein	7,1
EC042_RS25200	hypothetical protein	7,1
EC042_RS24455	hypothetical protein	7,1

EC042_RS13835	tRNA-Arg	7,1
EC042_RS08465	hypothetical protein	7,1
EC042_RS03945	hypothetical protein	7,0
EC042_RS16300	Putative invasion genes regulator	6,9
EC042_RS25780	hypothetical protein	6,9
EC042_RS08635	fimbrial protein FimH	6,9
EC042_RS26745	conjugal transfer protein TraL	6,9
EC042_RS25005	hypothetical protein	6,9
EC042_RS02975	fimbrial chaperone SfmC	6,8
EC042_RS02230	hypothetical protein	6,8
EC042_RS15160	hypothetical protein	6,8
EC042_RS11925	hypothetical protein	6,8
EC042_RS19640	pilin chaperone	6,8
EC042_RS25480	hypothetical protein	6,8
EC042_RS25520	hypothetical protein	6,8
EC042_RS14330	hypothetical protein	6,7
EC042_RS12875	hypothetical protein	6,7
EC042_RS02995	Fimbriae biosynthesis transcriptional regulator fimZ	6,7
EC042_RS25665	antitoxin	6,7
EC042_RS01800	hypothetical protein	6,7
EC042_RS25555	hypothetical protein	6,7
EC042_RS08150	hypothetical protein	6,7
EC042_RS16295	hypothetical protein	6,7
EC042_RS26295	hypothetical protein	6,6
EC042_RS22360	hypothetical protein	6,6
EC042_RS11325	sulfoxide reductase heme-binding subunit YedZ	6,6
EC042_RS17230	antitoxin	6,6
EC042_RS25685	hypothetical protein	6,6
EC042_RS12050	antitoxin	6,6
EC042_RS23430	hypothetical protein	6,6
EC042_RS08795	hypothetical protein	6,6
EC042_RS06615	hypothetical protein	6,6
EC042_RS08410	aminopeptidase	6,6
EC042_RS20185	hypothetical protein	6,5
EC042_RS01065	tRNA-Ala	6,5
EC042_RS24615	tRNA-Phe	6,5
EC042_RS22525	tRNA-Ala	6,5
EC042_RS16565	hypothetical protein	6,5

**Tabla 9.9.** Listado de los genes mayormente sobreexpresados en la cepa *E. coli* 042  $\Delta hha$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT a partir de cultivos crecidos en medio LB a 37°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

## 9. ANEXO

<i>E. coli</i> 042 $\Delta$ <i>hha</i> $\Delta$ <i>hha2</i> vs <i>E. coli</i> 042 WT		
Identificador	Descripción	Fold change
EC042_RS25175	magnesium-translocating P-type ATPase	23,2
EC042_RS08810	EC042_RS08810 hypothetical protein	17,3
EC042_RS08660	fimbrial protein	13,1
EC042_RS07850	tRNA-Thr	12,3
EC042_RS11295	hypothetical protein	10,6
EC042_RS09275	hypothetical protein	10,3
EC042_RS07145	hypothetical protein	10,2
EC042_RS08655	fimbria/pilus periplasmic chaperone	9,3
EC042_RS24455	hypothetical protein	8,8
EC042_RS26995	hypothetical protein	8,6
EC042_RS18795	hypothetical protein	8,6
EC042_RS11730	hypothetical protein	8,5
EC042_RS13835	tRNA-Arg	8,5
EC042_RS08650	fimbrial biogenesis outer membrane usher protein	8,5
EC042_RS15165	hypothetical protein DNA-binding protein	8,4
EC042_RS05780	tRNA-Ser	8,4
EC042_RS16300	HilA family transcriptional regulator YgeH	8,3
EC042_RS08645	type 1 fimbrial protein	8,3
EC042_RS18175	hypothetical protein	8,2
EC042_RS09105	hypothetical protein	8,1
EC042_RS06765	tRNA-Tyr	7,8
EC042_RS01250	hypothetical protein	7,8
EC042_RS05870	EC042_RS05870 hypothetical protein	7,7
EC042_RS15125	molybdopterin-guanine dinucleotide biosynthesis protein MobA	7,7
EC042_RS07390	hypothetical protein	7,7
EC042_RS21295	CesD/SycD/LcrH family type III secretion system chaperone	7,6
EC042_RS07765	putative bacteriophage protein racC	7,5
EC042_RS17210	hypothetical protein	7,4
EC042_RS11285	hypothetical protein	7,3
EC042_RS24340	hypothetical protein	7,2
EC042_RS11745	exodeoxyribonuclease VIII	7,2
EC042_RS07745	hypothetical protein	7,2
EC042_RS03185	hypothetical protein	7,1
EC042_RS25650	putative RadC-like DNA repair protein	7,1
EC042_RS07760	exonuclease VIII (putative phage exodeoxyribonuclease)	7,1
EC042_RS01795	reactive chlorine species resistance protein RclC	7,1
EC042_RS11335	hypothetical protein	7,1
EC042_RS08640	type 1 fimbrial protein	7,1
EC042_RS21300	YopB/SseC family type III secretion system	7,1

EC042_RS09180	site-specific integrase	7,0
EC042_RS06905	putative phage exodeoxyribonuclease	6,9
EC042_RS17205	hypothetical protein	6,9
EC042_RS17220	RadC family protein putative DNA repair protein	6,9
EC042_RS17215	putative antirestriction protein	6,9
EC042_RS24355	IS630 family transposase	6,9
EC042_RS24055	type IV toxin-antitoxin system toxin CbtA	6,9
EC042_RS07770	putative bacteriophage protein	6,8
EC042_RS21315	HilA family transcriptional regulator EilA	6,8
EC042_RS15160	hypothetical protein	6,8
EC042_RS06900	excisionase	6,7
EC042_RS08605	ABC transporter ATP-binding protein/permease	6,7
EC042_RS10010	arpB putative type III effector protein	6,7
EC042_RS24165	hypothetical protein	6,7
EC042_RS12565	putative nucleic acid binding protein (putative regulator)	6,7
EC042_RS09170	putative exodeoxyribonuclease	6,7
EC042_RS05815	csgE curli production assembly/transport component	6,7
EC042_RS00145	putative type III effector protein	6,6
EC042_RS21310	putative cell invasion type III effector protein	6,6
EC042_RS09165	Qin prophage; uncharacterized protein	6,6
EC042_RS06770	tRNA-Tyr	6,6
EC042_RS11740	hypothetical protein	6,6
EC042_RS00115	IS1 transposase B	6,6
EC042_RS16100	tRNA-Met	6,6
EC042_RS07135	hypothetical protein	6,5
EC042_RS06400	hypothetical protein	6,5
EC042_RS05810	csgF curli production assembly/transport component	6,5
EC042_RS16295	type III secretion-associated chaperone	6,5
EC042_RS18375	hypothetical protein	6,5
EC042_RS03005	envY porin thermoregulatory protein	6,5
EC042_RS05765	hypothetical protein	6,5
EC042_RS09430	rstA two-component response regulator	6,5

**Tabla 9.10.** Listado de los genes mayormente sobreexpresados en la cepa *E. coli* 042  $\Delta hha\Delta hha2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT a partir de cultivos crecidos en medio LB a 37°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.



## **9.2 Capítulo 2: Descripción de un nuevo parálogo del gen *hns* en la cepa *E. coli* 042: el gen *hns2***

Tabla 9.11. Listado de los oligonucleótidos utilizados durante el desarrollo del capítulo 2 de Resultados de esta memoria.

Tabla 9.12. Listado de los genes mayormente sobreexpresados en el fondo genético  $\Delta hns$  respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 37°C.

Tabla 9.13. Listado de los genes mayormente sobreexpresados en el fondo genético  $\Delta hns2$  respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 37°C.

Tabla 9.14. Listado de los genes mayormente sobreexpresados en el fondo genético  $\Delta hns\Delta hns2$  respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 37°C.

Figura 9.15. Resultado de la secuenciación del gen *hns2* clonado en el plásmido pLG338.30, bajo el control de su promotor salvaje o bajo el control del promotor que presenta la transición C→T, después de ser transformados en la cepa *E. coli* 042  $\Delta hns$ .

Tabla 9.16. Listado de los genes mayormente sobreexpresados en el fondo genético  $\Delta hns$  respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 25°C.

Tabla 9.17. Listado de los genes mayormente sobreexpresados en el fondo genético  $\Delta hns2$  respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 25°C.

Figura 9.18. Categoría funcional y número total de genes sobreexpresados en el transcriptoma global de la cepa *E. coli* 042 en ausencia de los genes *hns* o *hns2* en función de la temperatura de crecimiento de las bacterias (25°C o 37°C).

<b>Oligonucleótido</b>	<b>Secuencia (5'-3')</b>
hns p1	5'ATGAGCGAAGCACTTAAAATTCTGAACAACATC CGTACTCTTGTGTAGGCTGGAGCTGCTTC 3'
hns p2	5'TTATTGCTTGATCAGGAAATCGTCGAGGGATTT ACCTTGCTGCATATGAATATCCTCCTTAGT 3'
hns 3x p1	5'TGAGCAAGGTAAATCCCTCGACGATTCCTGAT CAAGCAAGACTACAAAGACCATGACGG 3'
hns 3x p2	5'CGGGATTTTTAAGCAAGTGCAATCTACAAAAGA TTATTGCCATATGAATATCCTCCTTAG 3'
hns p1 up	5' CCTCACGTGCTGCGAAATCA 3'
hns p2 down	5' GCAGGCCTTCGTTGAATACG 3'
hns2 p1	5'GCTTGTATTCTTTACCACAAACATTAGGGAAAT CGCATGTCTGTGTAGGCTGGAGCTGCTTC 3'
hns2 p2	5'AAAATGAGCGGATTTGGTAGACATATGGTATGA AGATTACTTCATATGAATATCCTCCTTAGT 3'
hns2 3x p1	5'GCTCGAAGCCGGTAAGACACTGGAAGATTTCGA GATCAAGGACTACAAAGACCATGACGG 3'
hns2 3x p2	5'ATGAGCGGATTTGGTAGACATATGGTATGAAGA TACTTGCATATGAATATCCTCCTTAG 3'
hns2 p1 up	5' TAAATTTGAACTCAGGATGGA 3'
hns2 p2 down	5' TGGCATGTTAATCCCTGTAGG 3'
stpA p1	5'TTTTTGTTTTGGCGTTAAAAGGTTTTCTTTATTAT GTCCGTAGTGTAGGCTGGAGCTGCTTC 3'
stpA p2	5'ACGCGCCATAGCAGCGACATCCGGCCTCAGTAA TTAGATCAGCATATGAATATCCTCCTTAGT 3'
stpA p1 up	5' GGAATTAGCGAGCAGAGAGCG 3'
stpA p2 down	5' ACGCACCAGTAATAACGCTGA 3'
lon p1	5'TACCTGGCGGAAATTAACCTAAGAGAGAGCTCT ATGAATCCTGTGTAGGCTGGAGCTGCTTC 3'
lon p2	5'CCTGTTTTTATTAGTGCATTTTGC GCGAGGTCAC TATTTTGCCATATGAATATCCTCCTTAGT 3'
lon p1 up	5' CTATTCTCGGCGTTGAATGTG 3'
lon p2 down	5' TACAACACCCCTTAATAAGGG 3'
hns2_pLATE51 NT fw	5'GGTGATGATGATGACAAGATGTCTGATGCTTTA AAGATCATCAACAATATTCGT 3'
hns2_pLATE51 NT rev	5'GGAGATGGGAAGTCATTACTTGATCTCGAAATC TTCCAGTGTCTTACCGGC 3'
hns2_pLATE31 CT fw	5'AGAAGGAGATATAACTATGTCTGATGCTTTAAA GATCATCAACAATATTCGTACT 3'
hns2_pLATE31 CT rev	5'GTGGTGGTGATGGTGATGGCCCTTGATCTCGAA ATCTTCCAGTGTCTTACCGGC 3'
pLG338 EB Fw	5' CCGCGCACATTTCCCCGAAAA 3'
pLG338 EB Rv	5' GCCAGTGCCAAGCTAGGCCTA 3'
hns2 pLG338 ECORI fw 5	5' CGGAATTCTAAATTTGAACTCAGGATGGA 3'
hns2 pLG338 BAMHI rev 3	5' CGGGATCCTGGCATGTTAATCCCTGTAGG 3'
hns2 prom BAMHI rev 3'	5' CGGGATCCGCGATTTCCCTAATGTTTGTGG 3'
hns pLG338 ECORI fw 5'	5' CGGAATTCTCAGCCACAGGCCCTCAATGA 3'
hns pLG338 BAMHI rev 3'	5' CGGGATCCACATCCGTATCGGTGTTATCC 3'
hns RT Fw	5' TGC GCAGGCAAGAGAATGTA 3'
hns RT Rv	5' AGCAGTTCGTTCCGGGTCAAT 3'
hns2 RT Fw	5' TCGGAAAAGACGTGGACAGG 3'

## 9. ANEXO

hns2 RT Rv	5' AGTGTCTTACCGGCTTCGAG 3'
stpA RT Fw	5' GCACCTGGCTGGAGCTGATGA 3'
stpA RT Rv	5' GTCCAGGTTTTAGTTTCACCG 3'
RS08585 Fw	5' GGTATAGCCTTCACGACCGA 3'
RS08585 Rv	5' CGCGTGTGAAATCGATCAGT 3'
RS04580 Fw	5' TTGAGTCCCTGGTTCCTGAG 3'
RS04580 Rv	5' CACATCATGCCCCGTTGTTATT 3'
RS05870 Fw	5' CTTTAGGAGAAGCACGGAAAGT 3'
RS05870 Rv	5' CAGCCATTCGGACAACAGAA 3'
RS21295 5' Fw	5' GCCAGTAAGTTCGAACGT 3'
RS21295 5' Rv	5' AGAGTTCAGCTTCGTCAAGTC 3'
RS21295 3' Fw	5' ATTACATGGGCCAGTGCCAG 3'
RS21295 3' Rv	5' TCCTCTTCGGTATTGGCTGTC 3'
arpB 5' Fw	5' CTCCTCATAGAATGCACATCG 3'
arpB 5' Rv	5' TGCATGGATGGGTGAAGTGGC 3'
arpB 3' Fw	5' CCTCAGCAAATAAACATCAT 3'
arpB 3' Rv	5' CGTCTTTAGTTCATCAGCACT 3'
hdeA RT Fw	5' GGACGAATCCTTCCAGCCAA 3'
hdeA RT Rv	5' CGATAGCTGGGGTTACGGTT 3'
RS01720 RT Fw	5' ACTACCGCTTTCAGTACGCC 3'
RS01720 RT Rv	5' TCCTTTGTTATCCGCCGTGT 3'
RS23490 RT Fw	5' GCAGCAACGTCGGTTTCG 3'
RS23490 RT Rv	5' GGTCGTGATTGTTGTGGTCC 3'
slp RT Fw	5' TTTGCTTGCCGCATGTAGTT 3'
slp RT Rv	5' ACGTTGATAACCTTCCCGCC 3'
Hcp RT Fw	5' CCGATCCACAGTCTGGTCAA 3'
Hcp RT Rv	5' CCACTTTCGGCAGCATTTC 3'
adaptador RACE-B6	5' GCGCGAATTCCTGTAGA 3'
hns2-3'RACE Rev	5' TCCACGTCTTTTCCGAACCA 3'
M13 <i>forward</i>	5' GTTTTCCAGTCACGAC 3'
M13 <i>reverse</i>	5' CAGGAAACAGCTATGAC 3'
lacZ Rev	5' GATGACCTGCAAGGCGATTA 3'
gapA042 RT Fw	5' TTTCCGTGCTGCTCAGAAAC 3'
gapA042 RT Rv	5' GTC AACACCAACTTCGTCCC 3'
lacZ 042 p1	5'TGAGCGGATAACAATTTACACAGGATACAGCT ATGACTATGGTGTAGGCTGGAGCTGCTTC 3'
lacZ 042 p2	5'AAATACGGGCAGACATGGCCTGCCCGTTATTA TTATTTTTGCATATGAATATCCTCCTTAGT 3'
lacZ 042 p1 up	5' CGTTGGCAGATTCTTTAATGC 3'
lacZ 042 p2 down	5' CCGGTAATAATCCACAGCAGG 3'
cat 042 p1	5'GAGTTATCGAGATTTTCAGGAGCTAAGGAAGCT AAAATGGAGGTGTAGGCTGGAGCTGCTTC 3'
cat 042 p2	5'AGGCGTTTAAGGGCACCAATAACTGCCTTAAAA AAATTACGCCATATGAATATCCTCCTTAGT 3'
cat 042 p1 up	5' GAAAATGAGACGTTGATCGGC 3'
cat 042 p2 down	5' CATAAGCGGCTATTTAACGAC 3'

**Tabla 9.11.** Listado de oligonucleótidos utilizados en el desarrollo del capítulo 2 de esta memoria.

<i>E. coli</i> 042 $\Delta$ <i>hns</i> vs <i>E. coli</i> 042 WT 37°C		
Identificador	Descripción	Fold change
EC042_RS08585	gadB glutamate decarboxylase beta subunit	43,0
EC042_RS20290	gadA glutamate decarboxylase alpha subunit	37,6
EC042_RS08580	gadC glutamate/gamma-aminobutyrate antiporter (extreme acid sensitivity protein)	34,7
EC042_RS01725	matB putative fimbrial protein	33,3
EC042_RS20250	hdeA putative periplasmic acid stress chaperone	26,1
EC042_RS01730	matA putative fimbrial transcriptional regulator	22,7
EC042_RS05780	tRNA-Ser	19,3
EC042_RS15335	stpA DNA-binding protein	18,6
EC042_RS20255	hdeD putative acid resistance protein	17,8
EC042_RS20260	gadE transcriptional regulator	16,6
EC042_RS06765	tRNA-Tyr	16,3
EC042_RS13160	hypothetical protein	15,7
EC042_RS21295	CesD/SycD/LcrH family type III secretion system chaperone	15,6
EC042_RS01715	putative fimbrial outer membrane usher protein	15,1
EC042_RS20245	acid-activated periplasmic chaperone hdeB	15,1
EC042_RS06770	tRNA-Tyr	14,8
EC042_RS04580	hypothetical protein	14,5
EC042_RS01720	putative fimbrial protein	14,2
EC042_RS20090	hypothetical protein membrane protein	14,0
EC042_RS07750	type I toxin-antitoxin Iar	13,9
EC042_RS01270	hypothetical protein	13,6
EC042_RS24770	tRNA-Gly	13,5
EC042_RS09275	hypothetical protein	13,3
EC042_RS01065	tRNA-Ala	12,8
EC042_RS18940	tRNA-Ala	12,8
EC042_RS26795	hypothetical protein	12,2
EC042_RS16270	hypothetical protein	12,2
EC042_RS11015	tRNA-Leu	12,1
EC042_RS11620	tRNA-Arg	12,0
EC042_RS14470	hypothetical protein	11,7
EC042_RS11295	hypothetical protein	11,4
EC042_RS01710	fimbria adhesin EcpD	11,3
EC042_RS04590	hypothetical protein	11,2
EC042_RS25370	hypothetical protein	11,2
EC042_RS08860	cold-shock protein	11,1
EC042_RS23640	hypothetical protein	11,1
EC042_RS09025	cspI cold shock-like protein	10,9
EC042_RS08150	hypothetical protein	10,9
EC042_RS09175	putative phage excisionase	10,9
EC042_RS05535	tRNA-Ser	10,9
EC042_RS08355	transcriptional regulatori protein RcsA	10,6
EC042_RS12265	gmd GDP-mannose 4,6-dehydratase	10,5
EC042_RS06760	hypothetical protein	10,5

## 9. ANEXO

EC042_RS18175	hypothetical protein	10,5
EC042_RS27815	fimbrial protein	10,3
EC042_RS22525	tRNA-Ala	10,2
EC042_RS04230	host cell division inhibitory peptide Kil	10,1
EC042_RS05870	hypothetical protein	10,1
EC042_RS05925	DUF2770 domain-containing protein	10,0
EC042_RS16360	invasion protein	10,0
EC042_RS06350	anti-adaptor protein IraM	10,0
EC042_RS12070	hypothetical protein	10,0
EC042_RS12270	wcaF putative colanic acid biosynthesis	9,8
EC042_RS25990	hypothetical protein	9,8
EC042_RS07630	hypothetical protein	9,7
EC042_RS16100	tRNA-Met	9,6
EC042_RS17605	hypothetical protein	9,6
EC042_RS11740	hypothetical protein	9,5
EC042_RS11285	hypothetical protein	9,5
EC042_RS07765	racC putative bacteriophage protein	9,4
EC042_RS07290	GsiB hypothetical protein	9,3
EC042_RS20190	slp outer membrane lipoprotein	9,3
EC042_RS21855	hypothetical protein	9,2
EC042_RS23405	ShET2/EspL2 family type III secretion system effector protein	9,2
EC042_RS06900	excisionase	9,2
EC042_RS20070	hypothetical protein	8,9
EC042_RS15135	hypothetical protein	8,9
EC042_RS09015	addiction module toxin GnsA/GnsB	8,9
EC042_RS20270	multidrug transporter subunit MdtE	8,9
EC042_RS07280	YciE hypothetical protein	8,8
EC042_RS18795	hypothetical protein	8,8

**Tabla 9.12.** Listado de los genes mayormente sobreexpresados en la cepa *E. coli* 042  $\Delta hns$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT a partir de cultivos crecidos en medio LB a 37°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

<i>E. coli</i> 042 $\Delta$ <i>hns2</i> vs <i>E. coli</i> 042 WT 37°C		
Identificador	Descripción	Fold change
EC042_RS05870	hypothetical protein	12,1
EC042_RS09275	hypothetical protein	11,8
EC042_RS07145	hypothetical protein	11,3
EC042_RS26995	hypothetical protein	11,3
EC042_RS18795	hypothetical protein	11,2
EC042_RS11295	hypothetical protein	10,3
EC042_RS13835	tRNA-Arg	10,1
EC042_RS24055	antitoxin hypothetical protein	9,6
EC042_RS22745	DUF3521 domain-containing protein	9,6
EC042_RS03185	hypothetical protein	9,3
EC042_RS17205	hypothetical protein	9,2
EC042_RS17210	hypothetical protein	9,1
EC042_RS03005	envY porin thermoregulatory protein	9,0
EC042_RS24455	hypothetical protein	8,8
EC042_RS17185	putative ATP/GTP-binding protein	8,7
EC042_RS17230	hypothetical protein	8,6
EC042_RS21255	putative prophage protein	8,6
EC042_RS25650	putative RadC-like DNA repair protein	8,5
EC042_RS18375	hypothetical protein	8,5
EC042_RS17220	putative DNA repair protein	8,5
EC042_RS11730	hypothetical protein	8,4
EC042_RS24065	putative plasmid-like protein	8,4
EC042_RS17215	putative antirestriction protein	8,4
EC042_RS07390	hypothetical protein	8,3
EC042_RS12065	hypothetical protein	8,3
EC042_RS13250	hypothetical protein	8,2
EC042_RS20520	putative acetyltransferase	8,1
EC042_RS23195	hypothetical protein	8,0
EC042_RS17225	hypothetical protein	8,0
EC042_RS24045	hypothetical protein	8,0
EC042_RS08890	hypothetical protein	7,9
EC042_RS01250	hypothetical protein	7,9
EC042_RS24005	toxin of GhoTS toxin-antitoxin pair	7,9
EC042_RS11990	Rha family transcriptional regulator hypothetical protein	7,8
EC042_RS09940	hypothetical protein	7,8
EC042_RS03370	citD citrate lyase acyl carrier protein	7,8
EC042_RS25660	hypothetical protein	7,8
EC042_RS24050	toxin hypothetical protein	7,8
EC042_RS07765	racC putative bacteriophage protein	7,7
EC042_RS12045	hypothetical protein	7,7
EC042_RS21245	prophage derepression protein	7,6
EC042_RS12025	hypothetical protein	7,6
EC042_RS21260	putative prophage DNA primase	7,6

## 9. ANEXO

EC042_RS09165	Qin prophage	7,6
EC042_RS05810	csgF curli production assembly/transport component	7,6
EC042_RS12565	putative nucleic acid binding protein (putative regulator)	7,5
EC042_RS25670	hypothetical protein	7,5
EC042_RS11745	exodeoxyribonuclease	7,5
EC042_RS26380	hypothetical protein	7,5
EC042_RS06575	ABC transporter permease	7,5
EC042_RS11960	hypothetical protein	7,4
EC042_RS03375	citC citrate (pro-3S)-lyase] ligase	7,4
EC042_RS06570	ABC transporter ATP-binding protein	7,3
EC042_RS18810	hypothetical protein	7,3
EC042_RS08885	hypothetical protein	7,3
EC042_RS24165	hypothetical protein	7,3
EC042_RS17195	transcriptional regulator putative DNA-binding protein	7,3
EC042_RS17825	flotillin hypothetical protein	7,3
EC042_RS20195	transcriptional regulator dctR	7,3
EC042_RS23515	hypothetical protein	7,2
EC042_RS18175	hypothetical protein	7,2
EC042_RS21265	hypothetical protein	7,2
EC042_RS03940	hypothetical protein	7,2
EC042_RS25665	antitoxin protein	7,2
EC042_RS24340	hypothetical protein	7,2
EC042_RS12985	hypothetical protein	7,2
EC042_RS07745	hypothetical protein	7,1
EC042_RS25655	hypothetical protein	7,1
EC042_RS06910	phage protein	7,1
EC042_RS01795	reactive chlorine species resistance protein RclC	7,1
EC042_RS24125	AlpA family transcriptional regulator	7,1

**Tabla 9.13.** Listado de los genes mayormente sobreexpresados en la cepa *E. coli* 042  $\Delta hns2$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT a partir de cultivos crecidos en medio LB a 37°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.



<i>E. coli</i> 042 $\Delta$ hns $\Delta$ hns2 vs <i>E. coli</i> 042 WT 37°C		
Identificador	Descripción	Fold change
EC042_RS01725	matB putative fimbrial protein	108,8
EC042_RS08585	gadB glutamate decarboxylase beta subunit	87,8
EC042_RS15335	stpA DNA-binding protein	83,7
EC042_RS07290	GsiB hypothetical protein	77,6
EC042_RS01730	matA putative fimbrial transcriptional regulator	70,0
EC042_RS08580	gadC glutamate/gamma-aminobutyrate antiporter (extreme acid sensitivity protein)	68,3
EC042_RS20250	hdeA putative periplasmic acid stress chaperone	60,3
EC042_RS20290	gadA glutamate decarboxylase alpha subunit	59,8
EC042_RS08660	fimbrial protein	52,8
EC042_RS11605	anti-adapter protein IraM interacts RssB preventing the degradation of the stress sigma factor RpoS	41,1
EC042_RS20245	acid-activated periplasmic chaperone hdeB	37,1
EC042_RS20255	hdeD putative acid resistance protein	36,7
EC042_RS20260	gadE transcriptional regulator	32,9
EC042_RS07280	YciE hypothetical protein	31,3
EC042_RS07285	YciF hypothetical protein	30,4
EC042_RS01715	putative fimbrial outer membrane usher protein	29,0
EC042_RS01720	putative fimbrial protein	28,0
EC042_RS20090	hypothetical protein membrane protein	25,3
EC042_RS11225	transcriptional regulator rcsA	25,2
EC042_RS17085	mchD putative microcin H47 activating protein	25,0
EC042_RS03395	lipid IVA palmitoyltransferase	23,3
EC042_RS12270	wcaF putative colanic acid biosynthesis acetyltransferase	23,3
EC042_RS12265	gmd GDP-mannose 4,6-dehydratase	22,9
EC042_RS24110	hypothetical protein	22,6
EC042_RS26710	hypothetical protein	22,5
EC042_RS09025	cspI cold shock-like protein	21,0
EC042_RS01710	fimbria adhesin EcpD	20,6
EC042_RS17080	mchE microcin H47 secretion protein colicin V secretion protein CvaA	20,6
EC042_RS08655	fimbrial chaperone	19,6
EC042_RS19635	aaspG fimbrial adhesin	19,4
EC042_RS08860	cold-shock protein	18,8
EC042_RS01210	putative type VI secretion system protein	18,8
EC042_RS20190	slp outer membrane lipoprotein EC042_3792	18,6
EC042_RS12275	wcaE putative colanic acid biosynthesis glycosyl transferase	17,4
EC042_RS26340	hypothetical protein	16,8
EC042_RS12170	ugd UDP-glucose 6-dehydrogenase	16,5
EC042_RS20280	gadW AraC family transcriptional regulator	16,4
EC042_RS01705	putative fimbrial protein	16,2
EC042_RS17110	mchS1 putative microcin esterase	15,4

9. ANEXO

EC042_RS17075	mchF microcin H47 secretion/processing ATP-binding protein	15,4
EC042_RS01240	putative type VI secretion system protein	15,3
EC042_RS08650	outer membrane usher protein	14,3
EC042_RS08820	hydrogen peroxide resistance protein	14,0
EC042_RS01800	hypothetical protein	14,0
EC042_RS25450	msbB2 lipid A biosynthesis (KDO)2-(lauroyl)-lipid IVA acyltransferase	13,7
EC042_RS08630	putative molybdopterin-binding oxidoreductase	13,7
EC042_RS27025	type III secretion protein	13,5
EC042_RS10355	putative signal transduction protein membrane protein	13,0
EC042_RS01220	hypothetical protein	13,0
EC042_RS12280	putative colanic acid polymerase wcaD	12,7
EC042_RS25460	glycosyl transferase family 1	12,7
EC042_RS26725	hypothetical protein	12,3
EC042_RS12425	hypothetical protein	12,3
EC042_RS01860	autotransporter outer membrane autotransporter barrel domain-containing protein	12,2
EC042_RS26705	hypothetical protein	12,1
EC042_RS12250	colanic acid biosynthesis fucosyltransferase WcaI	12,0
EC042_RS01205	type VI secretion system contractile sheath large subunit	11,9
EC042_RS21355	adenine deaminase	11,8
EC042_RS10130	ATP-independent periplasmic protein-refolding chaperone	11,7
EC042_RS26335	hypothetical protein	11,6
EC042_RS01865	EAL domain-containing protein	11,4
EC042_RS08635	FimH fimbrial adhesin	11,1
EC042_RS19640	molecular chaperone	11,0
EC042_RS09520	transcription modulator YdgT	11,0
EC042_RS12540	fimbrial protein	10,9
EC042_RS19650	type 1 fimbrial protein	10,8
EC042_RS00440	transcriptional regulator leuO	10,8
EC042_RS05820	csgD csgAB operon transcriptional regulator involved in biofilm formation	10,5
EC042_RS08645	fimbrial protein	10,4
EC042_RS08815	diguanylate cyclase DgcZ	10,4
EC042_RS10230	molecular chaperone HscC	10,3

**Tabla 9.14.** Listado de los genes mayormente sobreexpresados en la cepa *E. coli* 042  $\Delta$ *hns* $\Delta$ *hns2*, expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT a partir de cultivos crecidos en medio LB a 37°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

```

E. coli 042 Δhns pLG338.30 hns2 TAAATTTGAAGCTCAGGATGGAAGCTAATGTCTATACCGTATAATGTATTGATTTTGTGCAC
E. coli 042 Δhns pLG338.30 hns2* TAAATTTGAAGCTCAGGATGGAAGCTAATGTCTATACCGTATAATGTATTGATTTTGTGCAC
*****

E. coli 042 Δhns pLG338.30 hns2 AGGAACGATTAATAATCATGATGTTATTGCTTGTATTCTTTACCACAACATTAGGGAAT
E. coli 042 Δhns pLG338.30 hns2* AGGAACGATTAATAATCATGATGTTATTGCTTGTATTCTTTACCACAACATTAGGGAAT
*****

E. coli 042 Δhns pLG338.30 hns2 CGCATGTCTGATGCTTTAAAGATCATCAACAATATTCTGACTCTCCGTGCTCAGGCTCGT
E. coli 042 Δhns pLG338.30 hns2* CGCATGTCTGATGCTTTAAAGATCATCAACAATATTCTGACTCTCCGTGCTCAGGCTCGT
*****

E. coli 042 Δhns pLG338.30 hns2 GAACTGACACTCGAAACCTTGGAAAGAGATGTTAG-----
E. coli 042 Δhns pLG338.30 hns2* GAACTGACACTCGAAACCTTGGAAAGAGATGTTAGGGAAGTGCAGCAAGTCCCTGATAT
*****

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* GAGATCATGTTTGTCTATCTGGAGCCATAGAACAGGGTTCATCATGAGTCATCAACTTACC

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* TTCCCGCAGCTGAATTCAGCAGTAAGCCCGCTCAGACCAGAAAAGAGATTTTCTTTGTCC

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* CGCATGGACAGATTTCTGCCATGGCAAAACATGGTGAAGTCATCGAGCCGTTTTACC

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* AAGGCTGGTAATGGCCGGACCTTATCCCTGGAAACCATGCTACGATTCACTGCATG

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* CAGCATTGGTACAACCTGAGCGATGGCGCATGGAAGATGCTCTGACGAATCGCCTCC

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* ATGCGTCTGTTGCCCGGTATCCCTGGATAGCGCCTTCCGGACCGCACCACCATATG

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* GTAATCTGCTGCATGGAGAGGAGCAATTTGTCTCAGCGATGCCGGTACCAGGGGGCC

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* CACAGCGCAGGAGCTGGCCGAGGTGGATGTGAGTGGCTGATCGCCAGCGCCCGGCA

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* AAGTAAGAACCTTGAACAGCATCCACGCAAGAACAAAACGGCCATCAACATCGAATACA

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* TGAAGCCAGCATCCGGGCCAGGGTGGAGCACCCTTTCGCATCATCAAGCAGACTTCC

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* GCTTCGTGAAAGCCAGATACAAGGGTTGCTGAAAAACGATAAACCACTGGCGATGTTAT

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* TCACGCTGGCAACCTGTTTCGGCGGACCAATGATACGTCAGTGGGAGAGATCTCACT

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* AAAAAGTGGGATAACGCCTTAAATGGCGAAGAACGGTCTAAATAGGCTGATTCAGGC

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* ATTTACGGGAGAAAAATCGGCTCAAACATGAAGAAATGAAATGACTGAGTCAGCCGAGA

E. coli 042 Δhns pLG338.30 hns2 -----
E. coli 042 Δhns pLG338.30 hns2* AGAATTTCCCGCTTATTCGCACCTTCTTAGAAAACTCACTATCGTTGTAGAAGAACG
*****

E. coli 042 Δhns pLG338.30 hns2 TCGTGAAGAAGAGGCTCATTCTAAAACAGCAGGAAGAAACAGGCCAAGCTGGAAGC
E. coli 042 Δhns pLG338.30 hns2* TCGTGAAGAAGAGGCTCATTCTAAAACAGCAGGAAGAAACAGGCCAAGCTGGAAGC
*****

E. coli 042 Δhns pLG338.30 hns2 ATTCCGTGAGAAGTTGTTGGAAGATGGCATAGACCCGGCAGAGTTGCTTCAACTGTTAG
E. coli 042 Δhns pLG338.30 hns2* ATTCCGTGAGAAGTTGTTGGAAGATGGCATAGACCCGGCAGAGTTGCTTCAACTGTTAG
*****

E. coli 042 Δhns pLG338.30 hns2 CCCATCTTCTTCCAAGACAGGAGCGAAACGCGCCCGTCCAGCTAAGTATAAATATGT
E. coli 042 Δhns pLG338.30 hns2* CCCATCTTCTTCCAAGACAGGAGCGAAACGCGCCCGTCCAGCTAAGTATAAATATGT
*****

E. coli 042 Δhns pLG338.30 hns2 TGATACCGATGGTTCGGAAAAGACGTGGACAGTCAAGGCGCAACTCCTCGTGTATCGC
E. coli 042 Δhns pLG338.30 hns2* TGATACCGATGGTTCGGAAAAGACGTGGACAGTCAAGGCGCAACTCCTCGTGTATCGC
*****

E. coli 042 Δhns pLG338.30 hns2 TGCTGCGCTCGAAGCCGTAAGACACTGGAAGATTTGAGATCAAGTAATCTTCATACCA
E. coli 042 Δhns pLG338.30 hns2* TGCTGCGCTCGAAGCCGTAAGACACTGGAAGATTTGAGATCAAGTAATCTTCATACCA
*****

E. coli 042 Δhns pLG338.30 hns2 TATGCTACCAAATCCGCTCATTCTTCTGGCGGGTTGTTGAACCAACCTACAGGGAT
E. coli 042 Δhns pLG338.30 hns2* TATGCTACCAAATCCGCTCATTCTTCTGGCGGGTTGTTGAACCAACCTACAGGGAT
*****

E. coli 042 Δhns pLG338.30 hns2 TAACATGCCA
E. coli 042 Δhns pLG338.30 hns2* TAACATGCCA
*****

```

**Figura 9.15.** Secuenciación del gen *hns2* clonado en el plásmido pLG338.30 bajo el control del promotor salvaje o el promotor que presenta la transición C→T en la caja -10, después de ser transformados los respectivos plásmidos en la cepa *E. coli* 042 Δ*hns*. Se indica de color rojo la inserción del elemento transponible *IS5* en el plásmido que presenta la mutación en el promotor del gen *hns2*, rompiendo totalmente la pauta de lectura del gen, inactivándolo de esta manera.

## 9. ANEXO

<i>E. coli</i> 042 $\Delta$ hns vs <i>E. coli</i> 042 WT 25°C		
Identificador	Descripción	Fold change
EC042_RS07290	general stress protein	1.137,5
EC042_RS07280	ferritin-like domain-containing protein	736,5
EC042_RS00760	fimbrial protein	718,8
EC042_RS12270	colanic acid biosynthesis acetyltransfer WcaF	581,7
EC042_RS12260	colanic acid biosynthesis wcaG	505,7
EC042_RS24025	cadaverine/lysine antiporter	492,4
EC042_RS20090	DUF2776 domain-containing protein	438,8
EC042_RS12275	colanic acid biosynthesis glycosyltransferas2 WcaE	427,6
EC042_RS27760	hypothetical protein	389,1
EC042_RS08585	glutamate decarboxylase	371,1
EC042_RS17480	thymidylate kinase	364,7
EC042_RS12255	GDP-mannose mannosyl hydrolase	352,9
EC042_RS12280	putative colanic acid polymerase WcaD	332,5
EC042_RS01730	matA putative fimbrial transcriptional regulator	306,4
EC042_RS12250	colanic acid biosynthesis glycosyltransferase WcaI	294,4
EC042_RS20240	MgtC/SapB family protein	292,1
EC042_RS01795	reactive chlorine species resistance protein RclC	289,9
EC042_RS12265	GDP-mannose 4,6-dehydratase	288,4
EC042_RS08580	acid resistance gamma-aminobutyrate antiporter GadC	273,3
EC042_RS03225	LPS O-antigen length regulator	263,4
EC042_RS01800	DUF1471 domain-containing protein	260,2
EC042_RS10355	GGDEF domain-containing protein	249,9
EC042_RS20290	glutamate decarboxylase	247,8
EC042_RS07285	ferritin-like domain-containing protein	245,6
EC042_RS22730	hypothetical protein	243,5
EC042_RS13790	fimbrial protein	235,5
EC042_RS17785	fimbrial-like protein	222,5
EC042_RS20250	acid-activated periplasmic chaperone HdeA	222,1
EC042_RS20255	HdeD family acid-resistance protein	219,3
EC042_RS08045	porin OmpN	211,6
EC042_RS27675	hypothetical protein	210,4
EC042_RS22645	porin OmpL	205,7
EC042_RS01135	hypothetical protein	197,6
EC042_RS03430	hypothetical protein	197,4
EC042_RS12225	colanic acid biosynthesis pyruvyl transferase WcaK	188,9
EC042_RS20245	acid-activated periplasmic chaperone HdeB	187,9
EC042_RS27360	hypothetical protein	185,0
EC042_RS12230	M-antigen undecaprenyl disphosphate flippase	182,1
EC042_RS08635	fimbrial protein	174,0
EC042_RS08655	fimbria/pilus periplasmic chaperone	164,2
EC042_RS02970	type 1 fimbrial protein subunit FimA	164,1
EC042_RS01720	hypothetical protein	157,9
EC042_RS00740	fimbrial protein	157,7
EC042_RS17490	inorganic phosphate transporter PitB	156,1

EC042_RS20260	acid resistance transcriptional activator GadE	154,7
EC042_RS19635	fimbrial protein	153,4
EC042_RS08650	fimbrial biogenesis outer membrane usher protein	151,9
EC042_RS12245	mannose-1-phosphate guanyltransferase cpsB	149,5
EC042_RS20580	YiaB family inner membrane protein	148,7
EC042_RS25245	DUF898 domain-containing protein	146,7
EC042_RS12240	phosphomannomutase CpsG	146,3
EC042_RS12235	undecaprenyl-phosphate glucose phosphotransferase	144,6
EC042_RS09025	cold shock protein CspI	143,1
EC042_RS23490	hypothetical protein	140,9
EC042_RS01205	type VI secretion system contractile sheath large subunit	140,8
EC042_RS22360	hypothetical protein	140,3
EC042_RS18870	multidrug efflux transporter transcriptional repressor AcrS	139,6
EC042_RS12305	low molecular weight protein-tyrosine-phosphatase Wzb	136,7
EC042_RS00745	fimbrial protein	129,1
EC042_RS01725	common pilus major fimbrillin subunit EcpA	128,5

**Tabla 9.16.** Listado de los genes mayormente sobreexpresados en la cepa *E. coli* 042  $\Delta hns$ , expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT a partir de cultivos crecidos en medio LB a 25°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

## 9. ANEXO

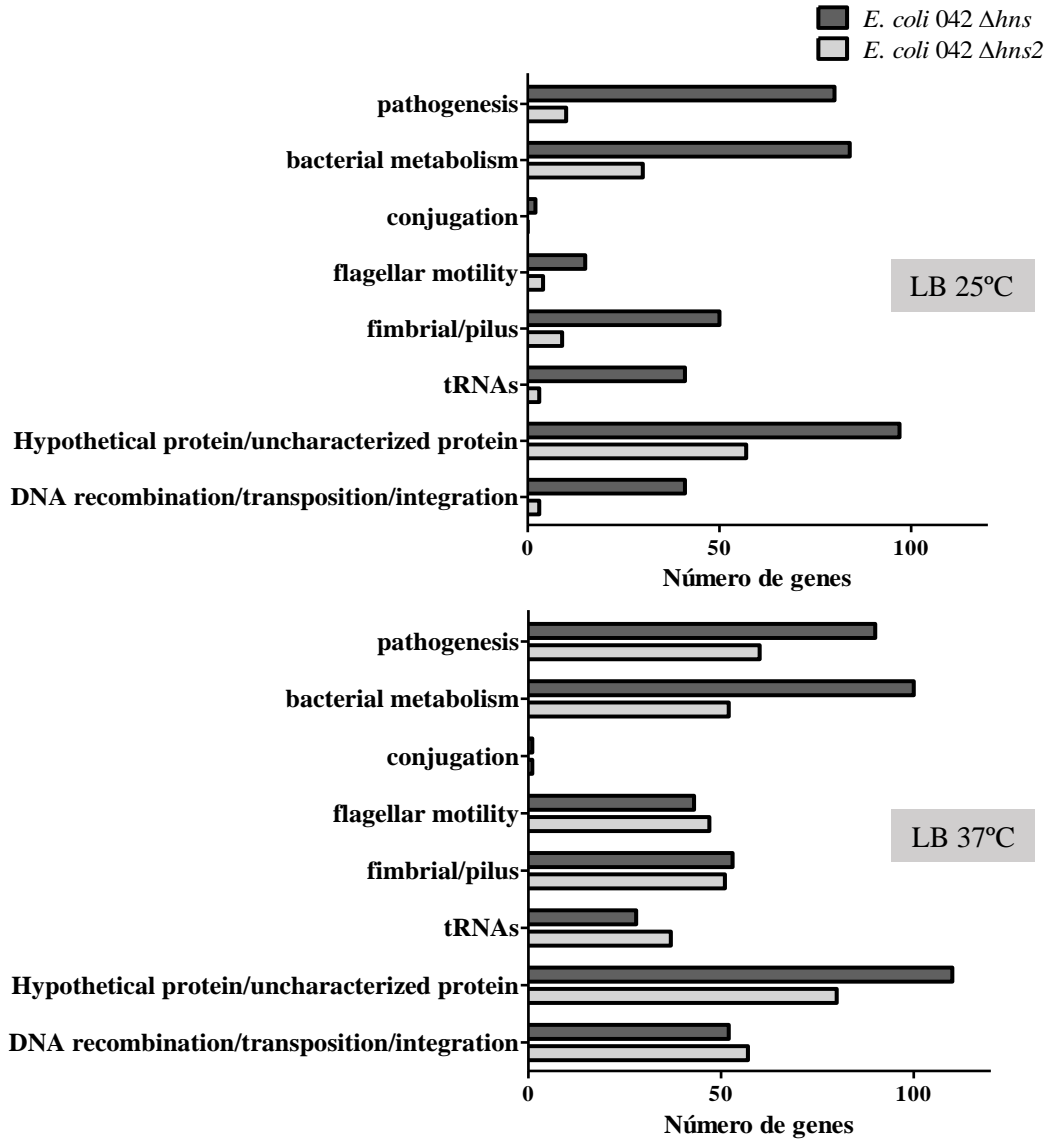
<i>E. coli</i> 042 $\Delta$ <i>hns2</i> vs <i>E. coli</i> 042 WT 25°C		
Identificador	Descripción	Fold change
EC042_RS06740	nitrate reductase subunit alpha NarZ	44,9
EC042_RS06735	nitrate transporter NarK	23,6
EC042_RS06745	nitrate reductase subunit beta NarH	21,0
EC042_RS06750	nitrate reductase molybdenum cofactor assembly chaperone NarJ	13,2
EC042_RS06755	respiratory nitrate reductase subunit gamma NarI	12,8
EC042_RS03510	DUF1266 domain-containing protein	8,2
EC042_RS10415	DUF1869 domain-containing protein	6,7
EC042_RS16380	EscR/YscR/HrcR family type III secretion system export apparatus protein	6,3
EC042_RS22730	hypothetical protein	5,4
EC042_RS19330	nitrite reductase large subunit	4,9
EC042_RS19990	nickel ABC transporter, nickel/metallophore periplasmic binding protein	4,7
EC042_RS08490	formate dehydrogenase subunit beta fdxH	4,6
EC042_RS03055	hypothetical protein	4,5
EC042_RS08485	formate dehydrogenase-N subunit alpha fdnG	4,5
EC042_RS02000	galactoside O-acetyltransferase lacA	4,5
EC042_RS07770	hypothetical protein	4,5
EC042_RS08480	sulfate ABC transporter substrate-binding protein	4,2
EC042_RS19335	nitrite reductase small subunit NirD	4,0
EC042_RS28855	hypothetical protein	4,0
EC042_RS19660	fimbria/pilus periplasmic chaperone	3,6
EC042_RS12365	YegJ family protein	3,6
EC042_RS08380	type VI secretion system PAAR protein	3,6
EC042_RS13110	ferredoxin-type protein NapF	3,5
EC042_RS08495	formate dehydrogenase cytochrome b556 subunit	3,4
EC042_RS23725	cytochrome c-552 nrfA	3,3
EC042_RS01960	2-methylcitrate dehydratase	3,1
EC042_RS28795	hypothetical protein	3,1
EC042_RS13105	nitrate reductase	3,0
EC042_RS19995	nickel ABC transporter permease	2,9
EC042_RS10150	arginine N-succinyltransferase	2,9
EC042_RS08145	hypothetical protein	2,7
EC042_RS20005	nickel ABC transporter ATP-binding protein nikD	2,6
EC042_RS25245	DUF898 domain-containing protein	2,6
EC042_RS06720	transcriptional regulator two-component system response regulator NarL	2,5
EC042_RS13100	periplasmic nitrate reductase subunit alpha	2,5
EC042_RS01835	hypothetical protein	2,5
EC042_RS18105	DUF805 domain-containing protein	2,4
EC042_RS20000	nickel ABC transporter permease subunit NikC	2,4
EC042_RS01950	methylisocitrate lyase prpB	2,4
EC042_RS10140	N-succinylarginine dihydrolase	2,4
EC042_RS06725	nitrate/nitrite two-component system sensor NarX	2,4



EC042_RS10155	succinylornithine/acetylornithine transaminase	2,4
EC042_RS23395	isocitrate lyase	2,3
EC042_RS03045	hypothetical protein	2,3
EC042_RS17485	thymidylate kinase	2,3
EC042_RS13840	OmpH family outer membrane protein	2,3
EC042_RS16435	hypothetical protein	2,2
EC042_RS03225	LPS O-antigen length regulator	2,2
EC042_RS01135	hypothetical protein	2,2
EC042_RS00085	molecular chaperone DnaK	2,2
EC042_RS10135	succinylglutamate desuccinylase	2,1
EC042_RS10145	succinylglutamate-semialdehyde dehydrogenase astD	2,1
EC042_RS19915	acyl carrier protein	2,1
EC042_RS02090	glycosyltransferase family 2 protein	2,1
EC042_RS23715	hypothetical protein DUF485 domain-containing protein	2,1
EC042_RS20580	EC042_2288 GDP-mannose mannosyl hydrolase	2,1
EC042_RS20580	YiaB family inner membrane protein	2,0
EC042_RS17480	thymidylate kinase	2,0
EC042_RS24655	molecular chaperone GroES	2,0
EC042_RS02665	chaperone protein HtpG	2,0
EC042_RS29620	protein YrbN	2,0
EC042_RS20010	nickel import ATP-binding protein Nike	2,0

**Tabla 9.17.** Listado de los genes mayormente sobreexpresados en la cepa *E. coli* 042  $\Delta$ *hns2*, expresados de manera relativa (*fold change*) al nivel de expresión de dichos genes en la cepa *E. coli* 042 WT a partir de cultivos crecidos en medio LB a 25°C. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.





**Figura 9.18.** Número total de genes sobreexpresados en el transcriptoma global de las cepas *E. coli* 042  $\Delta hns$  y *E. coli* 042  $\Delta hns2$  en función de la temperatura del medio de crecimiento de las bacterias (25°C o 37°C). También se indica la categoría funcional de los diferentes genes sobreexpresados como consecuencia de estas mutaciones.

### 9.3 Capítulo 3: Nuevos datos del papel del plásmido pAA2 y del regulador AggR en la virulencia de la cepa *E. coli* 042

Tabla 9.19. Listado de los oligonucleótidos utilizados durante el desarrollo del capítulo 3 de Resultados de esta memoria.

Figura 9.20. Imagen completa del *Northern Blot* representado en la figura 4.63 del apartado 4.3.2.4 del capítulo 3 de Resultados, donde se puede apreciar el elevado peso molecular del transcrito detectado correspondiente al gen *aggR*.

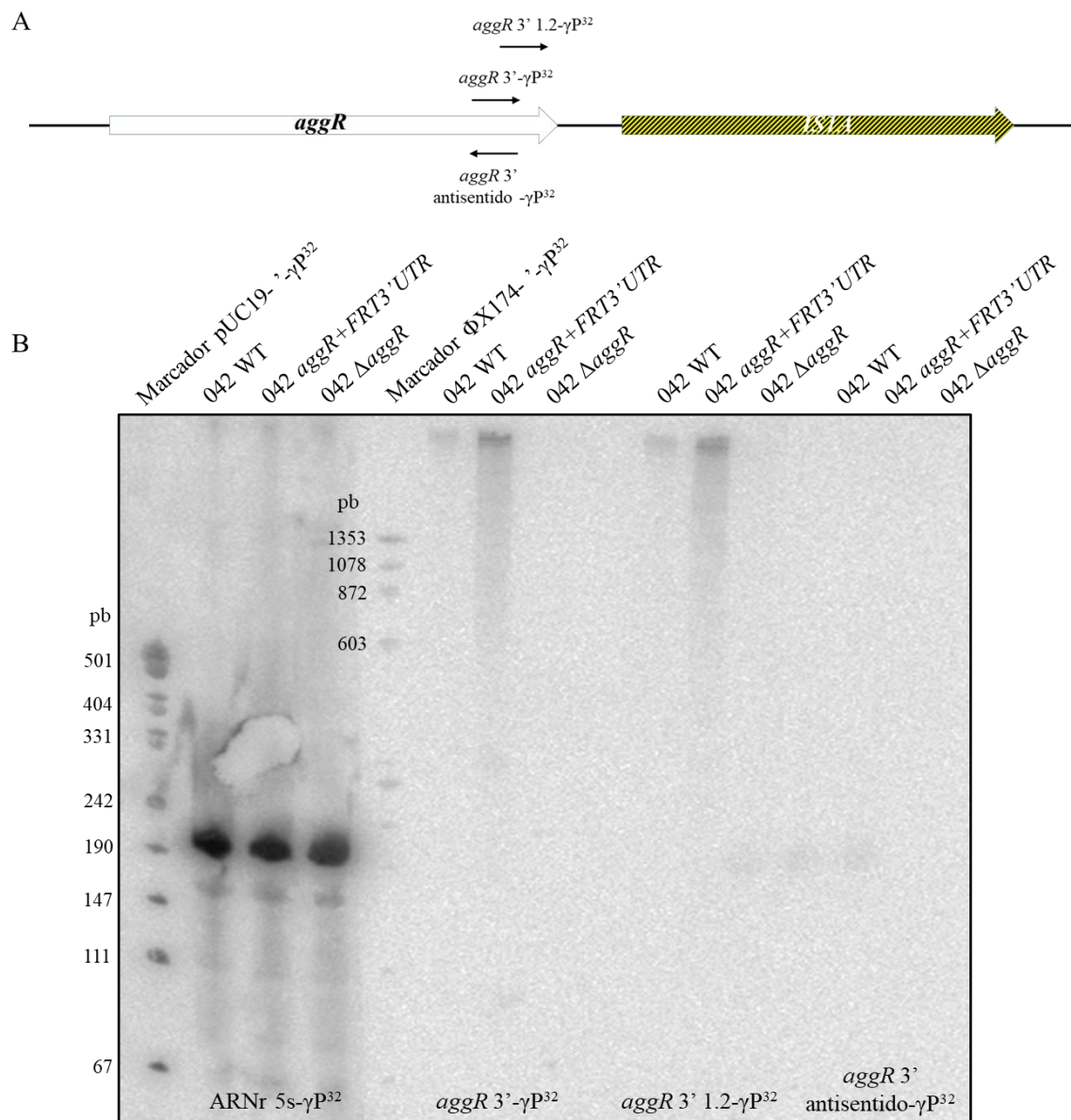
Tabla 9.21. Listado de genes implicados en la transferencia plasmídica sobreexpresados en el fondo genético *aggR+FRT3'UTR* respecto a la cepa *E. coli* 042 WT, cuantificados mediante RNA-seq a partir de cultivos crecidos en medio LB a 37°C.

## 9. ANEXO

Oligonucleótido	Secuencia (5'-3')
aafA3x p1	5'AATATATGAAGTAGTAGTATTAATGCTGAGCTT GTGACAAATGACTACAAGACCATGACGG 3'
aafA 3x p2	5'CCAAAAGTTATTATATTGTCACAAGCTCAGCAT TTAATTTTCATATGAATATCCTCCTTAG 3'
aafA p1 up	5' GCCTCTCCTAGGTTTTCTTAC 3'
aafA p2 down	5' CACTATCAATAAGTTGGAGTC 3'
She Fw	5' GGGTATTGTCCGTTCCGAT 3'
She Rv	5' ACAACGATACCGTCTCCCG 3'
traR Fw	5' GTCAGAACATGTCCGGGAGGA 3'
traR Rv	5' AATATTTTCCGCCGGGCTTC 3'
traQ Fw	5' AAGGATGGCCGCAATCAACT 3'
traQ Rv	5' CCCTGGGTGTCTGGTTTTCAT 3'
aatP Fw	5' AACGGAATTGACAAGCTGGG 3'
aatP Rv	5' GCTTGCTTTCAATCCAAGGC 3'
pAA 134 Fw	5' TCGTCATCACCCGACTGAAA 3'
pAA 134 Rv	5' AACCAGTTATCATTTCGCGGC 3'
aap Fw	5' GCGGTTGGAACGCAGATAAT 3'
aap Rv	5' TGAAGTTGCTTGTTGTGCCA 3'
aafA Fw	5' GGAGCGCAAATATCGACCTG 3'
aafA Rv	5' GCCAGAGTGAATCCTGCTGA 3'
aggR p1	5'ACATTTTTTTTCATGTGAGAATGATATGAAATTA AAACAAAACGTGTAGGCTGGAGCTGCTTC 3'
aggR p2	5'TTATTGGCTTTTAAAATAAGTCAAGAATTGTTTT GGTGTTATCATATGAATATCCTCCTTAGT 3'
aggR p1 up	5' GCTGCAATTAAGATACAACCCC 3'
aggR 3UTR p1	5'ATAACACCAAAAACAATTCTTGACTTATTTTAAA AGCCAATAAGTGTAGGCTGGAGCTGCTTC 3'
aggR 3UTR p2	5'ATATGTTTATAGCAATCTCAAATAATGATATGA AACATGTTTCATATGAATATCCTCCTTAGT 3'
aggR 3UTR p1 up	5' CGCAGATTGCCTGATAAAGAC 3'
aggR 3UTR p2 down	5' TTGCCGTTACGCACCACTCCG 3'
aggR IRL p1	5'TCATATCATTATTTGAGATTGCTATAAACATATT GAGATGGCGTGTAGGCTGGAGCTGCTTC 3'
aggR IRL p2	5'AGTAGCTGAACAGGAGGGACAGCTGATAGAAA CAGAAGCCACCATATGAATATCCTCCTTAGT 3'
aggR IS1a p1	5'ATTATTTGAGATTGCTATAAACATATTGAGATG GCTGAAGTTGTGTAGGCTGGAGCTGCTTC 3'
aggR IS1a p2	5'GTTTAGCGACTCGATGGAATTCGTTGTATAGAT CACTTTGCGCATATGAATATCCTCCTTAGT 3'
aggR IS1a p2 down	5' CTCAATTACACCGGCTATGCC 3'
aggR 3'	5' GAATTGTTTTGGTGTATGCCA 3'
aggR 3' 1.2	5' GGCTTTTAAAATAAGTCAAGA 3'
aggR antisense	5' CCTTATGCAATCAAGAATGAG 3'
5S	5' CTACGGCGTTTCACTTCTGAGTTC 3'
gapA042 RT Fw	5' TTTCCGTGCTGCTCAGAAAC 3'
gapA042 RT Rv	5' GTCAACACCAACTTCGTCCC 3'
aggR walking-PCR c1	5' GTCCGAATTGGTCAAAAGGAA 3'
aggR walking-PCR c2	5' AAGCCTAATGAAATATGATGT 3'
aggR walking-PCR 1	5' CCTTATGCAATCAAGAATGAG 3'
aggR walking-PCR 2	5' AAGCCTAATGAAATATGATGT 3'

aggR walking-PCR 3	5' CTAAATCAGTAAGTTGGCAGC 3'
aggR walking-PCR 4	5' GTGACGTAAAATCGTGTTGAG 3'
aggR walking-PCR 5	5' CGACAGCGACTTCCGTCCCAG 3'
aggR walking-PCR 6	5' CTGCGCTGATGCTGGTATGCG 3'

**Tabla 9.19.** Listado de los oligonucleótidos utilizados durante el desarrollo del capítulo 3 de Resultados esta memoria.



**Figura 9.20.** Resultado del revelado del *Northern Blot* representado en la figura 4.63 presente en el apartado 4.3.2.4 de Resultados. (A) Lugar de hibridación en la secuencia del gen *aggR* de las sondas marcadas radioactivamente con  $\gamma\text{P}^{32}$  con la finalidad de detectar el transcrito perteneciente a dicho gen. (B) Imagen completa del gel después de realizar y digitalizar el revelado de la membrana. Se puede observar el elevado tamaño del transcrito perteneciente al gen *aggR* detectado con las dos parejas de oligonucleótidos (*aggR* 3' y *aggR* 3' 1.2) que hibridan al final de la secuencia codificante del mismo.

<i>E. coli</i> 042 <i>aggR</i> + <i>FRT3</i> ' <i>UTR</i> vs <i>E. coli</i> 042 WT 37°C		
Identificador	Descripción	Fold change
EC042_pAA077	type-F conjugative transfer system pilin assembly thiol-disulfide isomerase TrbB	2,0
EC042_pAA078	type-F conjugative transfer system pilin chaperone TraQ	1,9
EC042_pAA081	conjugal transfer protein TrbE	1,8
EC042_pAA076	P-type conjugative transfer protein TrbJ	1,7
EC042_pAA086	type-F conjugative transfer system protein TraW	1,7
EC042_pAA085	conjugal transfer protein TraU	1,6
EC042_pAA075	conjugal transfer protein TrbF	1,6
EC042_pAA092	conjugal transfer protein TraP	1,5
EC042_pAA096	type IV conjugative transfer system protein traL	1,5
EC042_pAA080	type-F conjugative transfer system pilin traF	1,4
EC042_pAA087	type-F conjugative transfer system protein TrbI	1,4
EC042_pAA088	type IV secretion system protein TraC	1,4
EC042_pAA097	type IV conjugative transfer system pilin TraA	1,3
EC042_pAA083	conjugative transfer system pilin trbC	1,3
EC042_pAA073	conjugal transfer protein TraG	1,3
EC042_pAA074	conjugal transfer protein TraH	1,3
EC042_pAA090	type IV conjugative transfer system protein traV	1,3
EC042_pAA082	type-F conjugative transfer system mating-pair traN	1,2
EC042_pAA098	conjugal transfer relaxosome protein TraY	1,2
EC042_pAA094	type-F conjugative transfer system secretin TraK	1,1
EC042_pAA069	type IV conjugative transfer system coupling protein TraD	1,1
EC042_pAA079	conjugal transfer protein TrbA	1,1

**Tabla 9.21.** Listado de genes implicados en la transferencia plasmídica sobreexpresados en el fondo genético *aggR*+*FRT3*'*UTR* respecto a la cepa salvaje de *E. coli* 042, a partir de cultivos crecidos en medio LB a 37°C hasta el inicio de la fase estacionaria de crecimiento. Los datos transcriptómicos fueron obtenidos mediante la tecnología de RNA-seq en colaboración con la empresa biotecnológica *Vertis Biotechnologie AG*.

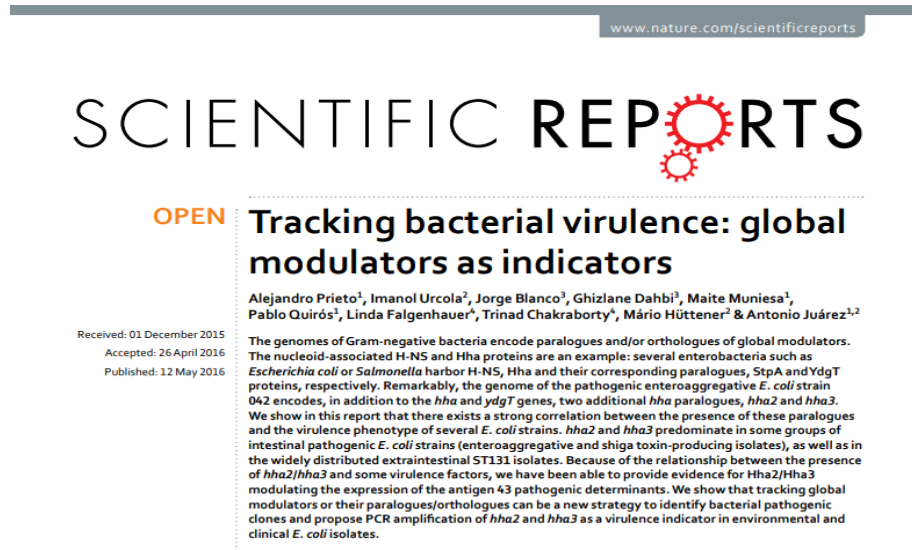
## **10. PUBLICACIONES**





Durante el desarrollo de la presente tesis doctoral se han publicado los siguientes artículos correspondientes a parte de los resultados mostrados en los diferentes capítulos de esta memoria:

**Capítulo 1 (apartado 4.1): Prieto, A., Urcola, I., Blanco, J., Dahbi, G., Muniesa, M., Quirós, P., Falgenhauer, L., Chakraborty, T., Hüttener, M., & Juárez, A. (2016).** Tracking bacterial virulence: global modulators as indicators. *Scientific Reports*, 6(1), 25973.



**Capítulo 2 (apartado 4.2): Prieto, A., Bernabeu, M., Aznar, S., Ruiz-Cruz, S., Bravo, A., Queiroz, M.H., & Juárez, A. (2018).** Evolution of bacterial global modulators: Role of a novel H-NS paralogue in the enteroaggregative *Escherichia coli* strain 042. *mSystems*, 3(3).



RESEARCH ARTICLE  
Host-Microbe Biology



## Evolution of Bacterial Global Modulators: Role of a Novel H-NS Paralogue in the Enteroaggregative *Escherichia coli* Strain 042

A. Prieto,<sup>a</sup> M. Bernabeu,<sup>a</sup> S. Aznar,<sup>a</sup> S. Ruiz-Cruz,<sup>b</sup> A. Bravo,<sup>b</sup> M. H. Queiroz,<sup>a</sup> A. Juárez<sup>a,c</sup>

<sup>a</sup>Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain

<sup>b</sup>Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

<sup>c</sup>Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Barcelona, Spain

**ABSTRACT** Bacterial genomes sometimes contain genes that code for homologues of global regulators, the function of which is unclear. In members of the family *Enterobacteriaceae*, cells express the global regulator H-NS and its paralogue StpA. In *Escherichia coli*, out of providing a molecular backup for H-NS, the role of StpA is poorly characterized. The enteroaggregative *E. coli* strain 042 carries, in addition to the *hns* and *stpA* genes, a third gene encoding an *hns* paralogue (*hns2*). We present in this paper information about its biological function. Transcriptomic analysis has shown that the H-NS2 protein targets a subset of the genes targeted by H-NS. Genes targeted by H-NS2 correspond mainly with horizontally transferred (HGT) genes and are also targeted by the Hha protein, a fine-tuner of H-NS activity. Compared with H-NS, H-NS2 expression levels are lower. In addition, H-NS2 expression exhibits specific features: it is sensitive to the growth temperature and to the nature of the culture medium. This novel H-NS paralogue is widespread within the *Enterobacteriaceae*.

## 10. PUBLICACIONES

Otros artículos derivados de la presente tesis doctoral cuya publicación se encuentra pendiente son los siguientes:

Capítulo 2 (apartado 4.2): Prieto, A., Bernabeu, M., Falgenhauer, L., Chakraborty, T., Hüttener, M., and Juárez, A. Regulated expression of the third H-NS paralogue H-NS2 influences enteroaggregative *Escherichia coli* strain 042 fitness. Manuscrito enviado a la revista *mSystems* (2020).

Capítulo 3 (apartado 4.3): Novel features of AggR-mediated virulence regulation in enteroaggregative *E. coli* strains. Manuscrito en preparación.

A pesar de que su temática se aleje del contenido de este trabajo, durante el periodo correspondiente a esta tesis doctoral, también se han logrado publicar otros artículos derivados de la colaboración con la actividad realizada por nuestro grupo de investigación:

**Hüttener, M., Prieto, A., Aznar, S., Bernabeu, M., Glaría, E., Valledor, A.F., Paytubi, S., Merino, S., Tomás, J., & Juárez, A. (2019).** Expression of a novel class of bacterial Ig-like proteins is required for IncHI plasmid conjugation. *PLOS Genetics*, 15(9), e1008399.

**Bernabeu, M., Sánchez-Herrero, J.F., Huedo, P., Prieto, A., Hüttener, M., Rozas, J., & Juárez, A. (2019).** Gene duplications in the *E. coli* genome: common themes among pathotypes. *BMC Genomics*, 20(1), 313.

**Hüttener, M., Prieto, A., Espelt, J., Bernabeu, M., & Juárez, A. (2018).** Stringent response and AggR-dependent virulence regulation in the enteroaggregative *Escherichia coli* strain 042. *Frontiers in Microbiology*, 9.

**Hüttener, M., Prieto, A., Aznar, S., Dietrich, M., Paytubi, S., & Juárez, A. (2018).** Tetracycline alters gene expression in *Salmonella* strains that harbor the Tn 10 transposon. *Environmental Microbiology Reports*, 10(2), 202–209.

**Tassinari, E., Aznar, S., Urcola, I., Prieto, A., Hüttener, M., & Juárez, A. (2016).** The incC sequence is required for R27 plasmid stability. *Frontiers in Microbiology*, 7.

