



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**Maestría y Doctorado en Ciencias Bioquímicas**

**“Caracterización de los elementos genéticos que regulan la expresión del sistema CRISPR-Cas en *Salmonella enterica* serovar Typhi”**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

**M. en C. LILIANA MEDINA APARICIO**

TUTOR PRINCIPAL

Dr. ISMAEL HERNÁNDEZ LUCAS  
[Instituto de Biotecnología, UNAM](#)

MIEMBROS DEL COMITÉ TUTOR

Dr. GUILLERMO GOSSET LAGARDA  
[Instituto de Biotecnología, UNAM](#)

Dr. JUAN MIRANDA RÍOS  
[Instituto de Investigaciones Biomédicas, UNAM](#)

Cuernavaca, Mor. Noviembre, 2018.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

El presente proyecto se realizó en el Departamento de Microbiología Molecular del Instituto de Biotecnología (UNAM), bajo la asesoría del Dr. Ismael Hernández Lucas y como parte del Programa de Maestría y Doctorado en Ciencias Bioquímicas de la misma Universidad.

El Comité tutorial estuvo integrado por:

**Dr. Ismael Hernández Lucas**

(Instituto de Biotecnología-UNAM)

**Dr. Guillermo Gosset Lagarda**

(Instituto de Biotecnología-UNAM)

**Dr. Juan Miranda Ríos**

(Instituto de Investigaciones Biomédicas-UNAM)

El Jurado designado para el Examen de Grado fue conformado por:

**Dr. Dimitris Georgellis**

(Instituto de Fisiología Celular-UNAM)

**Dr. David René Romero Camarena**

(Centro de Ciencias Genómicas-UNAM)

**Dr. Lorenzo Patrick Segovia Forcella**

(Instituto de Biotecnología-UNAM)

**Dr. Mario Soberón Chávez**

(Instituto de Biotecnología-UNAM)

**Dr. Mario Enrique Zurita Ortega**

(Instituto de Biotecnología-UNAM)

Este trabajo se llevó a cabo gracias al financiamiento de la Dirección General de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México (DGAPA/UNAM, IN-203215 e IN-203618).

Para la realización de esta investigación, la alumna recibió beca CONACyT (No. CVU/Becario: 215141/220853) y beca otorgada por la Dirección General de Asuntos del Personal Académico (DGAPA/UNAM) en la modalidad de Obtención de Grado/Doctorado como parte del proyecto IN-203215 e IN-203618.

## **AGRADECIMIENTOS**

Al Dr. Ismael Hernández Lucas, porque has sido mi tutor académico y de vida. Por todo el conocimiento que me compartiste durante estos años. Por creer en mí y hacer que ante las caídas me diera cuenta que era capaz de levantarme. Porque gracias a tus exigencias logré terminar este proyecto y ahora tengo las armas para seguir en este camino de la ciencia.

A los miembros de mi comité tutorial Dr. Juan Miranda Ríos y Dr. Guillermo Gosset Lagarda por compartir sus puntos de vista, conocimientos, críticas y sugerencias que ayudaron al desarrollo de este proyecto, y sobre todo por apoyarme en cada tutorial a pesar de que no siempre los resultados eran positivos.

Al jurado de examen: Dr. Mario E. Zurita Ortega, Dr. Mario Soberón Chávez, Dr. Dimitris Georgellis, Dr. Lorenzo P. Segovia Forcella y Dr. David R. Romero Camarena por sus valiosos comentarios y sugerencias que contribuyeron al mejoramiento del manuscrito final que aquí se presenta. También agradezco su disposición y apoyo para la realización oportuna de los trámites correspondientes.

Al Dr. Edmundo Calva Mercado y Dr. José Luis Puente García por estar siempre pendientes de mi desarrollo académico a pesar de que no fui su alumna directa, siempre recibí guía, apoyo y palabras de aliento de su parte cuando lo necesite. Tienen mi admiración, respeto y profundo agradecimiento.

A la Dra. Alejandra Vázquez Ramos, M. C. Marcos Fernández Mora, Biol. Francisco Javier Santana Estrada, Dra. Lucía Perezgasga Ciscomani, Ing. Leticia Olvera Rodríguez y Mario Alberto Caro Bermudez por su colaboración y apoyo técnico brindado para la realización de este proyecto.

A la Dra. Rosa María Gutiérrez Ríos y Dra. Sonia Dávila Ramos por la disposición y compromiso mostrado para la realización de todos los análisis bioinformáticos requeridos en este trabajo de investigación.

A Rebeca Herrera Trujillo, Elvira Villa Herrera, Mario R. Cruz Jarillo, José Luis Gamma Ferrer, María Xochitl González Candelario, Raúl Ríos Muñoz, Héctor Díaz Estrada y Miguel A. Trujillo González por su contribución en la preparación oportuna de todos los medios de cultivo y materiales necesarios para todos los experimentos realizados en este proyecto.

A Lic. Amapola Blanco Zabala, Rosalva González Arenas, Gloria Villa Herrera y Antonio Bolaños Guillen por su apoyo en la realización eficiente y oportuna de los trámites derivados de becas, inscripciones, apoyos económicos, cotizaciones, etc. Su ayuda fue muy valiosa.

A los doctores Ana Lucía Gallego, José Augusto Ramírez, Magdalena Wiesner, José Miguel Villarreal, Carmen Guadarrama, Miguel Ángel De la Cruz, Abraham Medrano, Alejandro Huerta, Verónica Martínez y Víctor Antonio García, quienes en su mayoría eran estudiantes de doctorado cuando ingresé al laboratorio y siempre tuvieron la disposición de guiarme experimental y conceptualmente. Por su paciencia al enseñarme y responder todas mis dudas por más simples que fueran. Gracias a eso pude afrontar las dificultades que se presentaron en cada experimento. Son un gran ejemplo para mí, tanto a nivel personal como profesional.

A todos los que han formado parte del “Lucas team”: Esteban, Noé, Víctor, Abi, Claudio, Fercha, Erika y Cynthia por su apoyo incondicional, ayuda, complicidad y todos los momentos de risas que hicieron más amenas las largas jornadas en el laboratorio. Porque, a pesar de que siguieron su camino en otro lado, sigo contando con su valiosa amistad. A los miembros actuales: Sara Karina, Dinorina, Mirellita, Isaac y Dani, agradezco su confianza, que me dejen formar parte de sus proyectos y que me permitan ayudarlos. Gracias a su entusiasmo y dedicación yo también aprendo mucho de ustedes.

A mis compañeros de laboratorio: Marcos, Alexis, Magda, Jessi, Deya, Mónica, Lucy, Zeus, Fer, Isela, Richard, Inés, Steph, Alvarito, Fabi, Haydeé, Carmen, Panchis, Víctor B., Marcos F., Claudia S., Diego, Gloria y Grecia, por hacer que los días de trabajo fueran más amenos con sus chistes, risas y pláticas, por brindarme su ayuda cuando lo he requerido. También agradezco a todos los que formaron parte del grupo Calva/Puente durante mi estancia en el Instituto, fui afortunada de convivir con ustedes y siempre recibir sus muestras de compañerismo.

A mi familia adquirida en Cuernavaca: Cess, Ale (China), Bere y Edna Lore (Amigui), son unas amigas increíbles que no me han dejado caer en los peores momentos. Gracias por enseñarme a sobrevivir fuera de casa, escuchar todos mis problemas, por sus consejos, por todas las locuras que hemos compartido y los lugares que hemos conocido juntas, Ohana!

A Lore, Crispin (Bro), Clau P., Emi, Ale Vázquez, Nadia y Luz porque además de mis compañeros, se convirtieron en mis amigos y me han dado fortaleza en momentos difíciles.

A mis amigos doctores Arleka, Poncho y Francia (Roomie) porque caminamos juntos en este trayecto llamado doctorado, compartimos el estrés y las frustraciones de experimentos fallidos y becas concluidas. Sin embargo, me alentaron a seguir adelante y me apoyaron en lo que se requiriera, ya fuera un tip

para hacer un experimento, un reactivo, un consejo o simplemente una taza de café.

A Doña Au por adoptarme en el “tupper club”, preocuparse por mi bienestar y procurar que comiera a mis horas. A Adrián, por el servicio de transporte que brinda y su paciencia al esperarme cuando se hacía tarde.

Al resto de mis amigos del Instituto, por brindarme una palabra de aliento o un momento de alegría.

Agradezco especialmente a mi familia por creer en mí. A mis padres por apoyarme incondicionalmente, por no criticar mis sueños ni el tiempo que haya tardado en alcanzarlos. Porque aun cuando la vida nos puso momentos muy difíciles nunca impidieron que yo siguiera mi camino. A mis hermanos, por nunca reclamarme que no haya estado en casa para ayudarlos, sino procurar no darme molestias para que pudiera enfocarme en el trabajo. Gracias por su infinito cariño.

# ÍNDICE GENERAL

	Página
1. LISTA DE ABREVIATURAS .....	1
2. RESUMEN .....	3
3. ABSTRACT .....	4
4. INTRODUCCIÓN .....	5
4.1 <i>Salmonella</i> .....	5
4.2 Fiebre tifoidea .....	5
4.3 Patogénesis de <i>S. Typhi</i> .....	6
4.4 El sistema inmunológico CRISPR-Cas .....	14
4.4.1 Mecanismo de acción del sistema CRISPR-Cas .....	15
4.4.2 Papel del sistema CRISPR-Cas en diversos procesos biológicos bacterianos .....	24
5. ANTECEDENTES .....	26
5.1 El sistema CRISPR-Cas en <i>Salmonella enterica</i> serovar Typhi .....	26
6. HIPÓTESIS .....	29
7. OBJETIVOS .....	29
7.1 Objetivo general .....	29
7.2 Objetivos particulares .....	29
8. MATERIALES Y MÉTODOS .....	30
8.1 Cepas bacterianas, medios de cultivo y condiciones de crecimiento. ....	30
8.2 Manipulación de DNA .....	30
8.3 Generación de mutantes por recombinación homóloga .....	31
8.4 Construcción de fusiones transcripcionales .....	31
8.5 Mutagénesis sitio dirigida .....	32
8.6 Ensayos de actividad específica CAT .....	33
8.7 Determinación de actividad $\beta$ -galactosidasa .....	34
8.8 Purificación de RNA .....	35
8.9 RT-PCR .....	36
8.10 Northern Blot .....	37

8.11	Primer extension .....	38
8.12	Purificación de H-NS y Lrp.....	38
8.13	Ensayos de retardamiento en gel .....	39
8.13.1	EMSAs con H-NS .....	39
8.13.2	EMSAs con Lrp.....	40
8.14	Purificación y electroforesis de proteínas de membrana externa .....	40
9.	RESULTADOS.....	42
9.1	LeuO induce la expresión del locus <i>cse1-cse2-cas7-cas5-cas6e-cas1-cas2</i> - CRISPR en <i>S. Typhi</i> .....	42
9.2	LeuO interactúa con dos zonas de la región reguladora de <i>cse1</i> .....	45
9.3	El medio mínimo N promueve la expresión del promotor <i>cse1</i> .....	47
9.4	H-NS y Lrp están implicados en la regulación del promotor <i>cse1</i> en N-MM... .....	48
9.5	La nucleasa Cas3 es inducida en medio mínimo N y reprimida por H-NS..	56
9.6	Elementos genéticos positivos implicados en la expresión transcripcional de <i>cas3</i> .....	60
9.7	La cadena sentido del locus CRISPR-Cas codifica el operón <i>cse1-cse2-</i> <i>cas7-cas5-cas6e-cas1-cas2</i> -CRISPR y el RNA <i>scse2</i> .....	62
9.8	Los RNAs <i>ascas2-1</i> y <i>ascse2-1</i> están presentes en la hebra antisentido del locus CRISPR-Cas .....	66
9.9	H-NS regula negativamente el RNA antisentido <i>ascse2-1</i> .....	68
9.10	El sistema CRISPR-Cas de <i>S. Typhi</i> es inducido en pH 7.5.....	70
9.11	El locus CRISPR-Cas de <i>S. Typhi</i> está implicado en la síntesis de proteínas de membrana externa.....	72
10.	DISCUSIÓN.....	77
11.	CONCLUSIONES .....	82
12.	PERSPECTIVAS .....	83
13.	BIBLIOGRAFÍA .....	86
	ANEXO I Tablas.....	101
	ANEXO II Artículos publicados.....	110



## 1. LISTA DE ABREVIATURAS

<b>°C</b>	Grados Celsius
<b>BSA</b>	Albúmina de suero bovino
<b>CAMPs</b>	Péptidos catiónicos antimicrobianos
<b>Cas</b>	Proteínas asociadas a CRISPR
<b>CAT</b>	Cloranfenicol acetiltransferasa
<b>CoA</b>	Coenzima A
<b>CRISPR</b>	Clúster de secuencias repetidas palindrómicas cortas interespaciadas regularmente
<b>crRNA</b>	RNA derivado de CRISPR
<b>DNA</b>	Ácido desoxirribonucleico
<b>DTT</b>	Ditiotreitol
<b>EDTA</b>	Ácido etilen-diaminotetraacético
<b>EMSA</b>	Ensayo de cambio en la movilidad electroforética
<b>g</b>	Gramos
<b>H-NS</b>	Proteína tipo histona para la estructuración del nucleoide
<b>IPTG</b>	Isopropil- $\beta$ -D-tiogalactopiranósido
<b>Kb</b>	Kilobase
<b>LPS</b>	Lipopolisacárido
<b>Lrp</b>	Proteína reguladora de respuesta a leucina
<b>LTTR</b>	Regulador transcripcional tipo LysR
<b>mA</b>	Miliamperes
<b>mg</b>	Miligramos
<b>min</b>	Minutos
<b>ml</b>	Militros
<b>mM</b>	Milimolar

<b>M</b>	Molar
<b>ng</b>	Nanogramos
<b>nm</b>	Nanómetros
<b>nt</b>	Nucleótido
<b>OD</b>	Densidad Óptica
<b>OMP</b>	Proteína de membrana externa
<b>ORF</b>	Marco de lectura abierta
<b>PAGE</b>	Electroforesis en gel de poliacrilamida
<b>PAM</b>	Motivo adyacente al protoespaciador
<b>pb</b>	Pares de bases
<b>PBS</b>	Amortiguador salino de fosfatos
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>RNA</b>	Ácido ribonucleico
<b>rpm</b>	Revoluciones por minuto
<b>RT-PCR</b>	Reacción en cadena de la polimerasa con transcriptasa reversa
<b>s</b>	Segundos
<b>SDS</b>	Duodecil-sulfato de sodio
<b>SPI</b>	Isla de Patogenicidad de <i>Salmonella</i>
<b>TAE</b>	Tris-acetato-EDTA
<b>TBE</b>	Tris-borato-EDTA
<b>T3SS</b>	Sistema de secreción tipo III
<b>µg</b>	Microgramos
<b>µl</b>	Microlitros
<b>µM</b>	Micromolar
<b>U</b>	Unidades
<b>wt</b>	Cepa silvestre

## 2. RESUMEN

El sistema CRISPR-Cas presente en bacterias y arqueas proporciona inmunidad adaptativa y heredable contra fagos o plásmidos. Este programa genético también ha sido involucrado en formación de biofilm, desarrollo de cuerpos fructíferos, reparación de DNA y virulencia. A pesar de que se han realizado múltiples investigaciones para caracterizar la funcionalidad de CRISPR-Cas en diversos microorganismos, existe un número limitado de reportes acerca de los factores transcripcionales y señales ambientales implicadas en su control genético. Debido a lo anterior, en este trabajo analizamos la organización y regulación transcripcional del sistema CRISPR-Cas presente en el patógeno de humanos *Salmonella enterica* serovar Typhi. Los resultados obtenidos establecen que la proteína LeuO, miembro de la familia LysR, está involucrada en la regulación positiva de este programa genético, mientras que su expresión es reprimida por los reguladores globales H-NS y Lrp. Además, se identificaron cinco unidades transcripcionales en el locus CRISPR-Cas: *cas3*, el operón *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR*, el RNA pequeño *scse2* y los RNAs antisentido *ascas2-1* y *ascse2-1*. Los cinco transcritos se expresan en medio mínimo N y su actividad es dependiente de pH. En este estudio también se definió que CRISPR-Cas es necesario para la producción de las proteínas de membrana externa OmpC y OmpF, sugiriendo que está implicado en el mantenimiento de la envoltura celular bacteriana. Por lo tanto, nuestros resultados demuestran que el sistema CRISPR-Cas de *S. Typhi* es capaz de participar en procesos biológicos esenciales diferentes a la inmunidad bacteriana.

### 3. ABSTRACT

The CRISPR-Cas system is a mechanism in bacteria and archaea capable of providing adaptive and heritable immunity against phages and plasmids. This system has also been implicated in biofilm formation, fruiting body development, DNA repair and virulence. Significant progress has been made regarding the functionality of this cluster in different microorganisms. However, there are few studies about the genetic elements and environmental signals that promote its expression. Hence, in this work we analyzed CRISPR-Cas transcriptional organization and regulation in the human pathogen *Salmonella enterica* serovar Typhi. The results showed that the nucleoid-associated proteins H-NS and Lrp repress the expression of this system and the LysR-type regulator LeuO participates in its positive control. Five transcriptional units were also identified in the CRISPR-Cas locus: *cas3*, the *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR operon, the small RNA *scse2* and the antisense RNAs *ascas2-1* and *ascse2-1*. These five transcriptional units are expressed in N-minimal medium and up-regulated at pH 7.5. Remarkably, it was also established that the synthesis of the outer membrane proteins OmpC and OmpF is CRISPR-Cas-dependent, demonstrating the involvement of this system in maintenance of cell envelope integrity. Therefore, our data provides evidence that the *S. Typhi* CRISPR-Cas system has a role in biological essential processes beyond bacterial immunity.

## 4. INTRODUCCIÓN

### 4.1 *Salmonella*

*Salmonella* es un género bacteriano perteneciente a la familia *Enterobacteriaceae* de las  $\gamma$ -proteobacterias, se compone por bacilos Gram-negativos, anaerobios facultativos, no esporulados y móviles. Este género se clasifica en dos especies: *Salmonella enterica* (*S. enterica*) y *Salmonella bongori* (*S. bongori*). *S. enterica* se divide en seis subespecies: *enterica* (I), *salamae* (II), *arizonae* (IIIa), *diarizonae* (IIIb), *houtenae* (IV) e *indica* (VI) (Le Minor & Popoff, 1987). De acuerdo al esquema Kauffman-White, basado en los antígenos capsulares (Vi), flagelares (H) y al lipopolisacárido (O), se han identificado más de 2,600 serovariedades pertenecientes a *S. enterica* subsp. *enterica* (Brenner *et al.*, 2000; Issenhuth-Jeanjean *et al.*, 2014).

Los serovares de *S. enterica* son capaces de infectar aves, reptiles y mamíferos, incluyendo al humano. Los cuadros clínicos asociados con la infección en el humano por *Salmonella* son la salmonelosis (gastroenteritis) causada principalmente por *S. enterica* serovar Typhimurium (*S. Typhimurium*), fiebre paratifoidea originada por *S. enterica* serovar Paratyphi (*S. Paratyphi*) y la fiebre tifoidea cuyo agente causal es *S. enterica* serovar Typhi (*S. Typhi*) (Layton & Galyov, 2007; Raffatellu *et al.*, 2008).

### 4.2 Fiebre tifoidea

La fiebre tifoidea o fiebre entérica es una enfermedad sistémica que se caracteriza por la presencia de fiebre de manera cíclica (39-40°C), dolor de cabeza, pérdida del apetito, escalofríos, vómito, hepatomegalia y esplenomegalia (Crump *et al.*, 2015; Huang & DuPont, 2005).

Esta infección es común en regiones con poco desarrollo económico y una infraestructura de salud pública limitada (Dougan & Baker, 2014). Cada año se registran 21 millones de casos de fiebre tifoidea a nivel mundial, de los cuales 128,000 a 161,000 pacientes pierden la vida al no recibir un tratamiento médico

adecuado (<http://www.who.int/immunization/diseases/typhoid/en/>) (Buckle *et al.*, 2012). Los continentes Asiático, Africano y Americano son los que presentan una mayor incidencia (Kothari *et al.*, 2008). En el caso de México, el sistema Nacional de Vigilancia Epidemiológica (SINAVE) reportó para el año 2017, más de 44,500 casos de fiebre tifoidea, siendo los estados de Tamaulipas, Sinaloa, Veracruz, Zacatecas y Chiapas los que presentaron el mayor número de reportes (<http://www.gob.mx/salud/acciones-y-programas/direccion-general-de-epidemiologia-boletin-epidemiologico>)

El tratamiento actual de esta infección bacteriana consiste en la administración de azitromicina, fluoroquinolonas (ciprofloxacino, ofloxacino) o cefalosporinas de tercera (ceftriaxona, cefixime) y cuarta generación (cefepime) (Crump *et al.*, 2015; Vala *et al.*, 2016). La administración oportuna de antibióticos ante la presencia de los síntomas de este padecimiento, junto con mejoras en la salud pública y el suministro de agua potable, podrían conducir al control local de la fiebre entérica. Debido a que esta enfermedad continúa siendo un problema de salud mundial, se ha generado un interés por el estudio de las estrategias moleculares que utiliza *S. Typhi* durante su ciclo de infección, con la finalidad de desarrollar estrategias eficaces que ayuden a la erradicación de la fiebre tifoidea.

### **4.3 Patogénesis de *S. Typhi***

A pesar de que la fiebre tifoidea es un padecimiento prevalente a nivel mundial, los mecanismos moleculares utilizados por *S. Typhi* para establecer la infección han sido poco estudiados. Lo anterior se atribuye a que esta bacteria infecta exclusivamente al humano, en consecuencia, no existe un modelo animal adecuado para estudiar la patogénesis de *S. Typhi*. Por lo tanto, la mayoría de los trabajos experimentales realizados para comprender los mecanismos generales de virulencia, se han llevado a cabo utilizando al patógeno relacionado *S. Typhimurium* (House *et al.*, 2001). Esta bacteria causa una enfermedad sistémica similar a la fiebre tifoidea en ratones genéticamente susceptibles, como los

C57BL/6, C3H/HeJ ó BALB/c (Keestra-Gounder *et al.*, 2015), mientras en el humano solo provoca una gastroenteritis localizada (Zhang *et al.*, 2003).

El ciclo de infección de *S. Typhi* inicia vía oral con la ingestión de alimentos o agua contaminados con materia fecal. La dosis requerida para provocar fiebre tifoidea varía de  $10^3$  a  $10^6$  bacterias y el periodo de incubación, generalmente, es de 1-2 semanas (Hornick *et al.*, 1970).

Dentro del hospedero, *S. Typhi* sobrevive a diversas condiciones de estrés, tales como la acción de la lisozima y los péptidos catiónicos antimicrobianos (CAMPs) presentes en la boca, así como la disminución de la disponibilidad de oxígeno a lo largo del tracto digestivo (Spector & Kenyon, 2012). Para la sobrevivencia bajo condiciones anaeróbicas se sugiere que *S. Typhi* emplea a la enzima nitrato reductasa para utilizar el nitrato como aceptor final de electrones en ausencia de oxígeno (Contreras *et al.*, 1997).

Cuando *Salmonella* llega al estómago, resiste al pH ácido gracias a la respuesta de tolerancia ácida (ATR), la cual ha sido ampliamente estudiada en *S. Typhimurium* (Álvarez-Ordóñez *et al.*, 2012; Audia *et al.*, 2001; Rychlik & Barrow, 2005). En el caso de *S. Typhi* se ha demostrado que la presencia de pH ácido induce la expresión de diversas proteínas de membrana externa y de FliC (flagelina) (Chander *et al.*, 2004; Jindal *et al.*, 2011).

Posteriormente, el patógeno migra hacia el intestino delgado, lugar donde está expuesto a otros tipos de estreses, incluyendo la prevalencia de condiciones anaeróbicas, competencia con la microbiota, incremento en la presión osmótica y presencia de moléculas dañinas tales como CAMPs, bacteriocinas, IgA neutralizantes y la bilis (Rychlik & Barrow, 2005; Spector & Kenyon, 2012). En respuesta a un incremento en la presión osmótica *S. Typhi* muestra la expresión diferencial de 889 genes de 4,474 evaluados. Los genes inducidos o reprimidos incluyen enzimas del metabolismo, proteínas de membrana, factores de virulencia, proteínas de secreción, proteínas de función desconocida y reguladores transcripcionales (Huang *et al.*, 2007). Dentro de este último grupo, se ha validado la participación de los factores sigma RpoS y RpoE, debido a que cepas de *S.*

Typhi carentes de manera individual en los genes correspondientes presentan una menor tasa de crecimiento respecto a la cepa silvestre en condiciones de alta osmolaridad (Du *et al.*, 2011). Adicionalmente, se sugiere que dicha bacteria mantiene la turgencia celular reprimiendo la expresión de la proteína de membrana externa OmpF y manteniendo la expresión de OmpC, ya que esta última forma poros de menor diámetro, por lo que disminuye la entrada de solutos hacia el periplasma (Graeme-Cook *et al.*, 1989; Martínez-Flores *et al.*, 1999). No obstante, cepas de *S. Typhi* deficientes en *ompC* y *ompF* presentan una tasa de crecimiento similar a la cepa silvestre en presencia de altas concentraciones de NaCl (0.3 M) (Villarreal *et al.*, 2014). En el caso de *S. Typhimurium*, la ausencia simultánea de *ompC* y *ompF* ocasiona que la bacteria presente atenuación en el modelo del ratón (Chatfield *et al.*, 1991)

Respecto a la presencia de péptidos catiónicos antimicrobianos, los cuales se insertan en la membrana bacteriana formando poros (Zhang *et al.*, 2001), uno de los mecanismos bacterianos contra la acción de estas moléculas, consiste en la modificación del lípido A con el fin de evitar su anclaje (Guo *et al.*, 1997; Rychlik & Barrow, 2005). Entre los elementos genéticos de *S. Typhi* involucrados en dicho proceso se encuentran los sistemas de dos componentes PhoP/PhoQ, PmrA/PmrB y el gen *pqaB*, los cuales le confieren resistencia a melitina, protamina y polimixina B (Baker *et al.*, 1997; Baker *et al.*, 1999).

Además de los CAMPs, la bacteria debe contender con la presencia de la bilis en el intestino. Este compuesto induce la producción de ROS en *S. Typhi*, por lo que el patógeno resiste al estrés oxidativo promoviendo la producción de enzimas como la superóxido dismutasa (SOD) y catalasa (Walawalkar *et al.*, 2016). Se ha demostrado el papel de la cascada de regulación *ItrR-ompR-ompC* en la resistencia de *S. Typhi* a sales biliares (Villarreal *et al.*, 2014) y una cepa de *S. Typhi* deficiente en *luxS* (enzima implicada en quorum sensing) presenta mayor sensibilidad al estrés inducido por bilis comparada con la cepa silvestre (Walawalkar *et al.*, 2016). La evaluación de una librería de mutantes de *S. Typhi* obtenida mediante la metodología TraDIS (transposon directed insertion-site



sequencing), sugiere la participación de los genes *waaL*, *waaK* y *waaI* (enzimas involucradas en la síntesis del lipopolisacárido), así como de *pagP* (Proteína para acilación del lípido A), *hupA* (subunidad HU $\alpha$  del regulador HU), *mrcA-mrcB* (proteínas asociadas a membrana), *acrAB/tolC* (bomba de flujo para la expulsión de sales biliares), *seqA* (proteína de unión a sitios GATC hemimetilados), *dam* (DNA adenin metiltransferasa) y *phoP/phoQ* (sistema de dos componentes) en la tolerancia de dicha bacteria a bilis (Langridge *et al.*, 2009). De manera interesante, el papel de *acrAB/tolC*, *seqA*, *dam* y *phoP/phoQ* en la resistencia a bilis de *S. Typhimurium* ha sido ampliamente estudiado y las cepas carentes en dichos elementos genéticos se encuentran atenuadas en el modelo del ratón (Heithoff *et al.*, 1999; Lacroix *et al.*, 1996; Miller *et al.*, 1989; Nishino *et al.*, 2006; Prieto *et al.*, 2007; Prieto *et al.*, 2004; van Velkinburgh & Gunn, 1999; Walawalkar *et al.*, 2016).

Una vez que la bacteria sobrevive a las condiciones adversas del intestino delgado, procede a establecer la infección. *Salmonella* se adhiere a las células epiteliales del íleon y células M de las placas de Peyer a través de estructuras denominadas fimbrias (Ramos *et al.*, 2004; Zhang *et al.*, 2000). El genoma de *S. Typhi* posee un repertorio único de operones fimbriales, incluyendo 12 fimbrias tipo chaperona-usher (*stg*, *sth*, *bcf*, *fim*, *saf*, *sef*, *sta*, *stb*, *stc*, *std*, *ste* y *tcf*), un pili tipo IV y un operón ortólogo a *agf* (Townsend *et al.*, 2001), de los cuales se ha demostrado que *stg*, *fim*, *sef* y el pili tipo IV tienen un papel fundamental en la adherencia de *S. Typhi* a células epiteliales (Dufresne *et al.*, 2018; Forest *et al.*, 2007; Zhang *et al.*, 2000).

En esta etapa, la alta osmolaridad, pH alcalino y disminución en las concentraciones de oxígeno del lumen intestinal, generan la activación del sistema de secreción tipo III (T3SS) y de proteínas efectoras codificadas en la isla de patogenicidad 1 de *Salmonella* (SPI-1) tanto en *S. Typhi* como en *S. Typhimurium* (Hansen-Wester & Hensel, 2001; Huang *et al.*, 2007; Leclerc *et al.*, 1998). En este último serovar se ha determinado que el proceso de invasión es mediado principalmente por los genes contenidos en esta isla de patogenicidad, ya que el T3SS transloca los efectores presentes en SPI-1 hacia la célula hospedera. Las

proteínas traslocadas son capaces de inducir rearrreglos en el citoesqueleto de actina dando lugar a la formación de ondulamiento (ruffling), evento que promueve la internalización bacteriana (Cardona-Castro & Sánchez-Jiménez, 2003; Figueroa & Verdugo, 2005; Zhang *et al.*, 2002; Zhou & Galán, 2001). Diversos genes identificados en la invasión de *S. Typhi* a células del epitelio intestinal son homólogos a genes presentes en SPI-1 de *S. Typhimurum* tal es el caso de *hilA* (activador principal de SPI-1), *sipBCDA* (proteínas efectoras), así como *invC*, *invA*, *invE*, *invG* y *prgH* (componentes del aparato de secreción T3SS) (Daigle, 2008; Eichelberg & Galán, 2000; Hermant *et al.*, 1995; Leclerc *et al.*, 1998; Weinstein *et al.*, 1998). Por esta razón, se asume que SPI-1 también es fundamental en el mecanismo de invasión utilizado por *S. Typhi*. Otros elementos genéticos de este serovar implicados en la invasión a células epiteliales corresponden a: Fur (regulador transcripcional), PhoP/PhoQ (Sistema de regulación de dos componentes), el factor sigma  $\sigma^{38}$  (RpoS), el pili Tipo IV, FlhD-FlhC y FliA (activadores del flagelo), así como diferentes fimbrias tipo usher (Stg, Sth, Fim, Saf, Sef, Sta, Stb, Stc, Std, Tcf) (Dufresne *et al.*, 2018; Eichelberg & Galán, 2000; Leclerc *et al.*, 2013; Lee *et al.*, 2007; Zhang *et al.*, 2000).

Una vez que *S. Typhi* se encuentra dentro de las células del epitelio intestinal, produce una exotoxina A<sub>2</sub>B<sub>5</sub> conocida como toxina tifoidea (Spano *et al.*, 2008), la cual es exportada extracelularmente induciendo daño en las células adyacentes. La toxina tifoidea presenta actividad de DNasa I, provocando distensión y detención del ciclo celular (Galán, 2016; Spano, 2016; Spano *et al.*, 2008). Por otra parte, la presencia de la cápsula Vi y una disminución en la expresión del flagelo, evitan que *S. Typhi* sea reconocida por TLR4 y TLR5 (Toll-like receptor) de las células hospederas, esto reduce la producción de IL-8 (Interleucina 8), por lo que el reclutamiento de neutrófilos se inhibe y la respuesta inflamatoria del hospedero disminuye (Raffatellu *et al.*, 2005; Sharma & Qadri, 2004; Wilson *et al.*, 2008).

*S. Typhi* se internaliza en el epitelio y es fagocitada por los macrófagos residentes en la lámina propia, dentro de los cuales se alberga en vacuolas

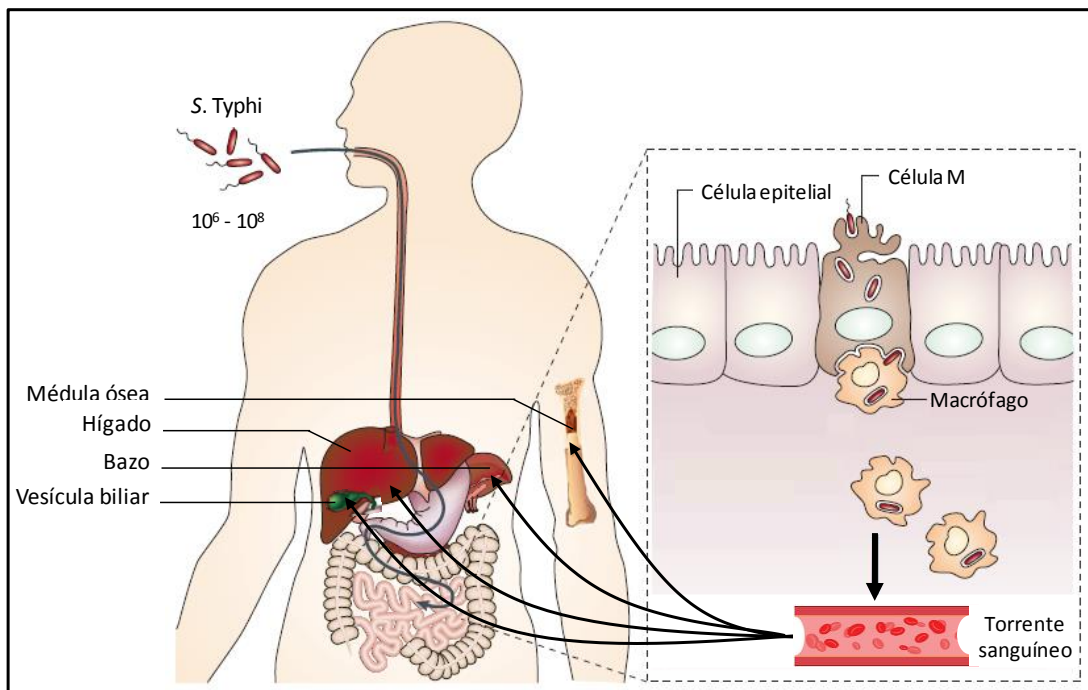
denominadas SCVs (*Salmonella* containing vacuole). Éstas se caracterizan por contener concentraciones limitadas de  $Mg^{2+}$ ,  $Fe^{2+}$ ,  $Ca^{2+}$  y fosfato, así como un pH ácido (Knodler & Steele-Mortimer, 2003), condiciones a las que debe adaptarse el patógeno. Además, tiene que contener con la producción de especies reactivas de oxígeno por parte del macrófago, así como evitar que la SVC se fusione con los lisosomas presentes en este tipo celular (Spano, 2016). Se ha demostrado la participación de diferentes elementos genéticos en la sobrevivencia y replicación de *S. Typhi* durante su estadio en macrófagos, tal es caso de Fur (regulador del consumo y transporte de hierro), PprZ (proteína fosfatasa Ser/Thr 2C), PrkY y PrkX (proteínas cinasas Ser/Thr), MgtBC (sistema de transporte de magnesio), MliC (inhibidor de lisozima tipo C unido a membrana), PagC y PagD (proteínas de membrana externa), Std y Stc (fimbrias tipo chaperona usher), FlhD-FlhC (activadores del flagelo), FliC (flagelina), WaaQGP (proteínas involucradas en la síntesis de LPS), GppA (enzima implicada en la síntesis de ppGpp), RpoE (factor sigma  $\sigma^{38}$ ), AcrA (componente de bomba de flujo), PhoP (regulador de respuesta), PgtE (proteasa de membrana externa) y PqaB (proteína implicada en la resistencia a CAMPs) (Baker *et al.*, 1999; Daigle, 2008; Daigle *et al.*, 2001; Dufresne *et al.*, 2018; Faucher *et al.*, 2006; Faucher *et al.*, 2008; Leclerc *et al.*, 2013; Sabbagh *et al.*, 2012; Zhang *et al.*, 2016).

Experimentos de transcriptómica realizados cuando *S. Typhi* se encuentra dentro de células fagocíticas detectaron la expresión de los siguientes genes pertenecientes a SPI-2: *ssaD*, *ssaJ*, *ssaN*, *ssaTS*, *ssaV* (componentes del aparato de secreción); *ssrA/ssrB* (Sistema de dos componentes activador de SPI-2); *sscA*, *sscB*, *ssaE* (chaperonas); así como *sseED* y *sseB* (proteínas efectoras) (Faucher *et al.*, 2006). Sin embargo, cepas de *S. Typhi*  $\Delta$ *sseB* o  $\Delta$ *ssrB* son capaces de invadir y replicarse dentro de macrófagos humanos de manera similar a la cepa silvestre (Forest *et al.*, 2010). Por lo que el papel de cada uno de los componentes de SPI-2 en la virulencia de este serovar debe ser esclarecido, ya que en *S. Typhimurium* se ha demostrado que el T3SS y los efectores presentes en SPI-2 son fundamentales para que la bacteria sobreviva dentro de células fagocíticas

(Cirillo *et al.*, 1998; Haraga *et al.*, 2008; Hensel *et al.*, 1998; Waterman & Holden, 2003).

Después de un periodo de adaptación dentro del macrófago (3-4 h), la bacteria es capaz de replicarse en esta célula eucarionte, provocando su lisis y diseminándose por el torrente sanguíneo o mediante macrófagos no activados (Cardona-Castro & Sánchez-Jiménez, 2003). En sangre, *S. Typhi* debe enfrentar la opsonización mediada por el complemento. La bacteria utiliza la cápsula Vi para impedir la inserción del componente 3 (C3), de esta manera impide el reconocimiento por los receptores CR3 de macrófagos (Wilson *et al.*, 2011).

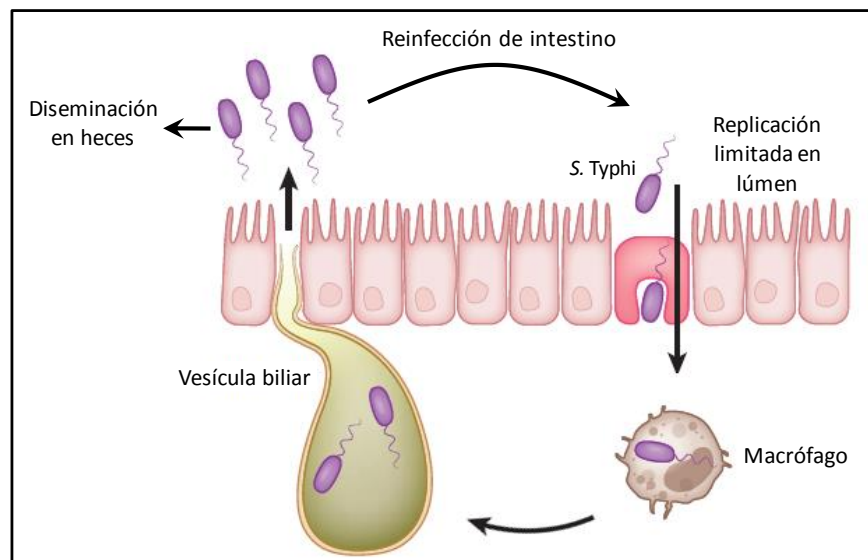
Si el paciente no recibe tratamiento farmacológico o la infección no es controlada por el sistema inmune, *S. Typhi* se disemina a bazo, hígado, médula ósea y vesícula biliar (Figuroa & Verdugo, 2005). Además, es capaz de reinfectar las placas de Peyer, causando necrosis del tejido y perforación intestinal, lo que provoca la muerte del paciente (Crump *et al.*, 2015; Everest *et al.*, 2001). En la figura 4.1 se resume el ciclo de infección de *S. Typhi*.



**Figura 4.1. Ciclo de infección de *Salmonella enterica* serovar *Typhi*.** *Salmonella* es ingerida en alimentos o agua contaminados. Pasa por el estómago e invade las células M del intestino delgado. Posteriormente, es fagocitada por macrófagos en los cuales se replica. La bacteria se

disemina por torrente sanguíneo a médula ósea, hígado, bazo y vesícula biliar, estableciendo una infección sistémica (modificada de Haraga *et al.*, 2008).

Con la administración del antibiótico adecuado la mayoría de los pacientes con fiebre tifoidea se recuperan de la fase aguda de este padecimiento. Sin embargo, del 3-5% de las personas infectadas con *S. Typhi*, desarrollan infección crónica en la vesícula biliar (Gonzalez-Escobedo *et al.*, 2011). En estos casos la bacteria accede a dicho órgano a través de macrófagos, lo coloniza y es capaz de persistir dentro de él hasta por 30 años, convirtiendo al individuo infectado en portador asintomático crónico (Gunn *et al.*, 2014; Sinnott & Teall, 1987). En este tipo de pacientes, *S. Typhi* es secretada con la bilis hacia el intestino delgado, lugar que puede reinfectar o ser secretada a través de las heces (Dougan & Baker, 2014; Gunn *et al.*, 2014) (Figura 4.2). Por lo tanto, los portadores constituyen un reservorio crucial para la diseminación del patógeno al ambiente (Gonzalez-Escobedo *et al.*, 2011). Esto es una estrategia fundamental que utiliza *S. Typhi* para sobrevivir y poder provocar nuevas infecciones. Se ha observado que la presencia de cálculos biliares favorece la persistencia de este patógeno, debido a que es capaz de crear una biopelícula sobre dichas estructuras (Crawford *et al.*, 2010). Sin embargo, los mecanismos moleculares utilizados por esta bacteria para establecerse y permanecer en vesícula biliar no han sido descritos.



**Figura 4.2. Mecanismo de persistencia y diseminación de *S. Typhi* en portadores asintomáticos crónicos.** *Salmonella* se disemina a vesícula biliar a través de macrófagos, en este

órgano puede permanecer hasta por 30 años y junto con la bilis es secretada continuamente al intestino delgado, compartimento que puede reinfectar o diseminarse al medio ambiente mediante heces (modificada de Dougan & Baker, 2014).

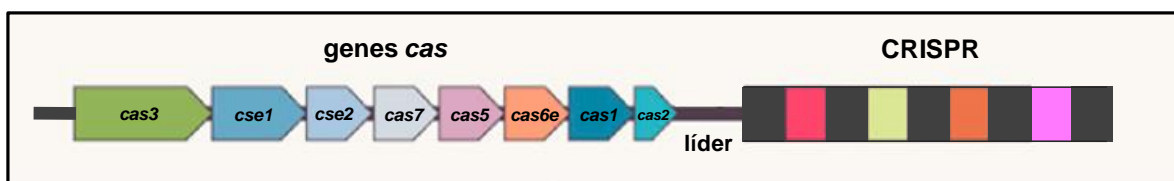
Como se describió anteriormente *Salmonella* utiliza diferentes mecanismos moleculares para sobrevivir a las condiciones ambientales presentes dentro del hospedante. Sin embargo, cuando esta bacteria se encuentra en vida libre también debe contender con diversos tipos de estreses tales como competencia, carencia de nutrientes, presión osmótica, temperatura y la infección por fagos.

#### **4.4 El sistema inmunológico CRISPR-Cas**

Los fagos constituyen una de las mayores amenazas para los organismos procariontes, ya que son capaces de erradicar entre el 4-50% de la población bacteriana que se genera diariamente en el mar (Breitbart & Rohwer, 2005). Debido a lo anterior, bacterias y arqueas han desarrollado diversas estrategias para sobrevivir al ataque viral. Dentro de estos mecanismos de defensa se encuentran: impedir la absorción del fago, interrumpir la inyección de su material genético, los sistemas de restricción/modificación incluyendo una variante recientemente descrita llamada DISARM (Defense Island System Associated with Restriction-Modification), así como la inducción de infecciones abortivas y el sistema CRISPR-Cas (Barrangou *et al.*, 2007; Labrie *et al.*, 2010; Ofir *et al.*, 2018).

El sistema CRISPR-Cas (Clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins) es el único mecanismo de defensa en procariontes capaz de proporcionar inmunidad adaptativa y heredable contra virus o plásmidos (Barrangou *et al.*, 2007; Marraffini & Sontheimer, 2008; Sorek *et al.*, 2013). Este mecanismo de defensa está integrado por un conjunto de proteínas conocidas como Cas, las cuales tienen dominios funcionales de nucleasas, helicasas e integrasas (Haft *et al.*, 2005; Jansen *et al.*, 2002a; Makarova *et al.*, 2006). Adyacente a los genes *cas* se localiza una secuencia líder no codificante, rica en adeninas y timinas que contiene el promotor del locus CRISPR (Hale *et al.*, 2008; Jansen *et al.*, 2002a; Lillestol *et al.*, 2006; Pul *et al.*, 2010), dicho locus está

compuesto por secuencias palindrómicas repetidas (28-40 pb) separadas por fragmentos únicos de tamaño similar (29-72 pb), llamados espaciadores (Figura 4.3) (Jansen *et al.*, 2002b). Algunos de estos fragmentos poseen homología con elementos genéticos externos como virus o plásmidos (Bolotin *et al.*, 2005; Mojica *et al.*, 2005; Pourcel *et al.*, 2005).



**Figura 4.3. Representación esquemática del locus CRISPR-Cas.** Se muestran los genes cas, la secuencia líder y el arreglo CRISPR. Los rectángulos negros representan las secuencias repetidas y los espaciadores se presentan como rectángulos de colores (rojo, verde, naranja, rosa) (modificada de Hille *et al.*, 2018).

Los locus CRISPR-Cas están presentes en el 87% de los genomas de arqueas y 45% de los genomas bacterianos secuenciados (Grissa *et al.*, 2007) y actualmente se dividen en dos clases principales. La Clase I está compuesta por los sistemas Tipo I, Tipo III y Tipo IV, los cuáles se caracterizan por la presencia de Cas3, Cas10 y Csf1 (Cas8-like), respectivamente. Los sistemas Tipo I están organizados en los subtipos I-A, I-B, I-C, I-D, I-E, I-F y I-U, mientras que el Tipo III se clasifica en III-A, III-B, III-C y III-D. No se ha reportado subdivisión para el Tipo IV. La Clase 2 de sistemas CRISPR-Cas se conforma por los tipos II, V y VI, cuyas proteínas características son Cas9, Cas12 (Cpf1) y Cas13, respectivamente. El Tipo II se divide en tres subtipos (II-A, II-B y II-C), mientras el Tipo V contiene seis subtipos (V-A, V-B, V-C, V-D, V-E y V-U), y el Tipo VI se divide en los subtipos VI-A, VI-B1, VI-B2 y VI-C (Koonin *et al.*, 2017; Makarova *et al.*, 2011; Makarova *et al.*, 2015).

#### 4.4.1 Mecanismo de acción del sistema CRISPR-Cas

El mecanismo de resistencia mediada por el sistema CRISPR-Cas ha sido demostrado en varios microorganismos y se divide en tres etapas: a) Adaptación o adquisición de espaciadores, b) Expresión o biogénesis de los crRNAs y c)

Interferencia. Uno de los sistemas mejor caracterizados es el Tipo I-E presente en *E. coli* K12, cuyo mecanismo de acción se describe a continuación:

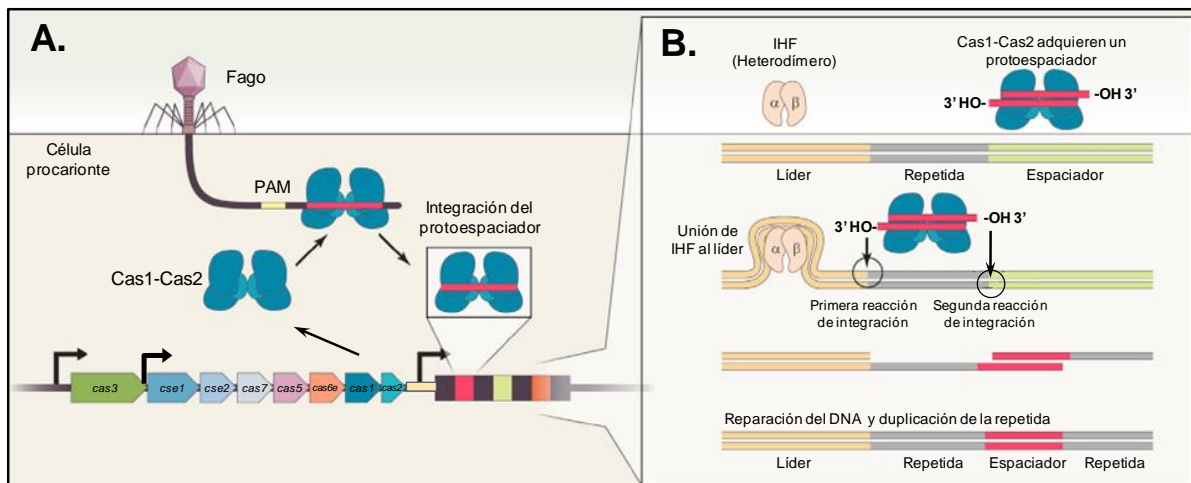
a) Adaptación/adquisición de espaciadores

La etapa de adaptación inicia cuando el sistema CRISPR-Cas detecta la entrada de material genético extraño derivado de fagos o plásmidos. Cas1 y Cas2 son fundamentales para la adquisición e integración de espaciadores (Nuñez *et al.*, 2014; Yosef *et al.*, 2012) y corresponden a una nucleasa/integrasa metal-dependiente y a una endorribonucleasa, respectivamente (Babu *et al.*, 2011; Beloglazova *et al.*, 2008; Makarova *et al.*, 2006; Wiedenheft *et al.*, 2009). Estas proteínas se ensamblan en un complejo formado por dos dímeros de Cas1 que flanquean un dímero de Cas2 (Cas1<sub>2</sub>-Cas2-Cas1<sub>2</sub>) (Nuñez *et al.*, 2014). Este complejo selecciona un fragmento de DNA del invasor, llamado protoespaciador. Se ha sugerido que estos fragmentos derivan de la degradación mediada por RecBCD en las horquillas de replicación truncadas del elemento genético exógeno (Levy *et al.*, 2015).

La selección del protoespaciador está determinada por la presencia de una secuencia conservada de 2-5 nucleótidos denominada PAM (Protospacer Adjacent Motif) (Bhaya *et al.*, 2011; Mojica *et al.*, 2009), con la cual interactúan los residuos Tyr165, Trp170, Tyr188, Tyr217 y Lys211, ubicados en el extremo carboxilo de Cas1 (Wang *et al.*, 2015). El fragmento de DNA foráneo se escinde, probablemente por la actividad nucleasa de Cas1, pero el mecanismo exacto y los elementos involucrados en este proceso no han sido descritos. Sin embargo, la cristalización del complejo Cas1-Cas2 unido a un protoespaciador, revela que el fragmento resultante de 33 nucleótidos está compuesto por 23 pb de DNA de doble cadena (dsDNA), flanqueado por 5-nt de cadena sencilla en cada uno de los extremos 3', los cuales contienen grupos OH (Wang *et al.*, 2015). Estos grupos hidroxilo son utilizados por cada dímero de Cas1 para catalizar la integración del protoespaciador en la secuencia repetida del CRISPR adyacente al líder (Datsenko *et al.*, 2012; Swarts *et al.*, 2012; Yosef *et al.*, 2012). Este evento parece ocurrir a través de un mecanismo de dos pasos: el primero inicia cuando la



proteína IHF se une a la secuencia líder e induce un doblez en el DNA (Nuñez *et al.*, 2016; Yoganand *et al.*, 2017). Esto le permite a uno de los dímeros de Cas1 posicionar un extremo 3'-OH del protoespaciador cerca del fosfato del primer nucleótido de la secuencia repetida, para realizar un ataque nucleofílico que induce un corte y ligación con una de las cadenas de dicha secuencia. En el siguiente paso, el segundo dímero de Cas1 y el 3'-OH restante del protoespaciador realizan otro ataque nucleofílico en la base 28 de la cadena antisentido de la secuencia repetida, lo que da como resultado un espaciador integrado y flanqueado en ambos lados por una cadena sencilla de la repetida del CRISPR (Arslan *et al.*, 2014; Nuñez *et al.*, 2015). Finalmente, la DNA polimerasa I complementa las cadenas simples y una DNA ligasa finaliza el proceso, el cual implica la duplicación de la primer secuencia repetida (Ivančić-Baće *et al.*, 2015; Yosef *et al.*, 2012) (Figura 4.4). Diversos estudios han establecido que la primera base del nuevo espaciador integrado es un remanente del PAM (Datsenko *et al.*, 2012; Goren *et al.*, 2012; Swarts *et al.*, 2012).



**Figura 4.4. Adquisición de espaciadores.** **A.** Al detectar la entrada de material genético proveniente de un fago, el complejo Cas1-Cas2 es capaz de adquirir un fragmento de este DNA (protoespaciador), el cual es seleccionado por la presencia del PAM. **B.** IHF se une a la región líder, induciendo un doblez en el DNA que es reconocido por Cas1-Cas2 para integrar el protoespaciador en la secuencia repetida adyacente. Los extremos 3'-OH del protoespaciador son utilizados por Cas1 para realizar las reacciones de integración. Las cadenas sencillas son complementadas por la DNA polimerasa y una ligasa finaliza el proceso. La integración del nuevo espaciador conlleva la duplicación de la secuencia repetida (Modificada de Hille *et al.*, 2018).

Es claro que Cas1 tiene un papel fundamental en la adquisición de espaciadores. Sin embargo, Cas2 también es necesaria en este proceso, debido a que es la encargada de unir los dos dímeros de Cas1 para producir un complejo funcional Cas1-Cas2. Adicionalmente, esta proteína coloca los dímeros de Cas1 a la distancia adecuada para obtener un protoespaciador de 33 nucleótidos. Cas2 a través de sus residuos de arginina también está implicada en la unión y estabilización del fragmento de DNA exógeno (Wang *et al.*, 2015).

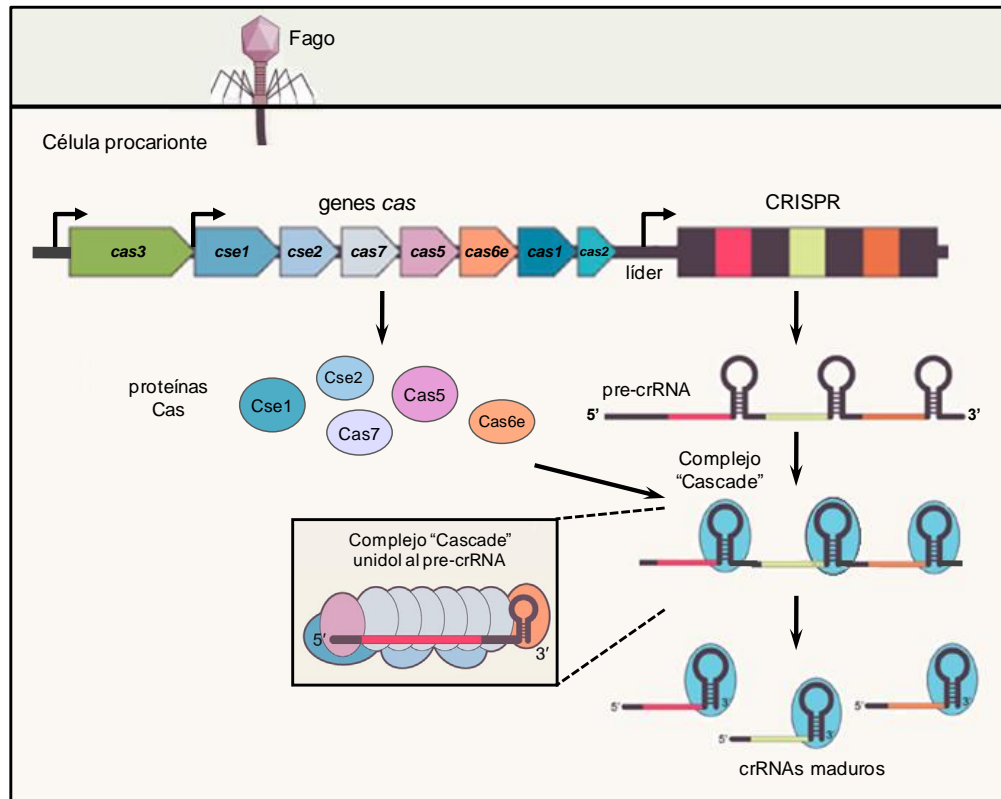
La adquisición de un nuevo espaciador proveniente de un elemento foráneo para el que no existe inmunidad previa, se conoce como adaptación “naïve”. En una segunda reinfección por el mismo invasor, la presencia de este espaciador promueve la adquisición de nuevos fragmentos, este proceso se denomina adaptación “primed” (Datsenko *et al.*, 2012; Swarts *et al.*, 2012).

Debido a que la inserción de espaciadores en el locus CRISPR se lleva a cabo en el extremo proximal al líder, estos fragmentos proveen información cronológica acerca de las infecciones ocurridas a lo largo de la historia evolutiva bacteriana (Al-Attar *et al.*, 2011; Bhaya *et al.*, 2011).

#### b) Expresión o biogénesis de los crRNAs

El resultado de la adaptación es la adquisición de DNA foráneo en las secuencias CRISPR. Posteriormente, cuando las bacterias enfrentan nuevos ataques del mismo invasor, los genes *cas* y las secuencias CRISPR se expresan. Las proteínas Cse1, Cse2, Cas7, Cas5 y Cas6e se ensamblan en un complejo multienzimático conocido como “Cascade” (CRISPR-associated complex for antiviral defense), el cual está conformado por un monómero de Cse1, dos de Cse2, seis de Cas7, uno de Cas5 y uno de Cas6e (Brouns *et al.*, 2008; Jore *et al.*, 2011; Wiedenheft *et al.*, 2011). El locus CRISPR se transcribe como un RNA precursor largo (pre-crRNA) (Al-Attar *et al.*, 2011; Brouns *et al.*, 2008; Pougach *et al.*, 2010), y las repetidas adquieren una estructura secundaria tipo tallo-asa (Kunin *et al.*, 2007), que promueve la unión del complejo “Cascade”. Posteriormente, el pre-crRNA es procesado por la endonucleasa Cas6e en

pequeños RNAs guía (crRNAs) (Brouns *et al.*, 2008). Cada crRNA está compuesto por 7 nucleótidos de la secuencia repetida 5', un espaciador de 33 pb y 21 nucleótidos de la repetida 3' (Jore *et al.*, 2011; Sinkunas *et al.*, 2013) (Figura 4.5).

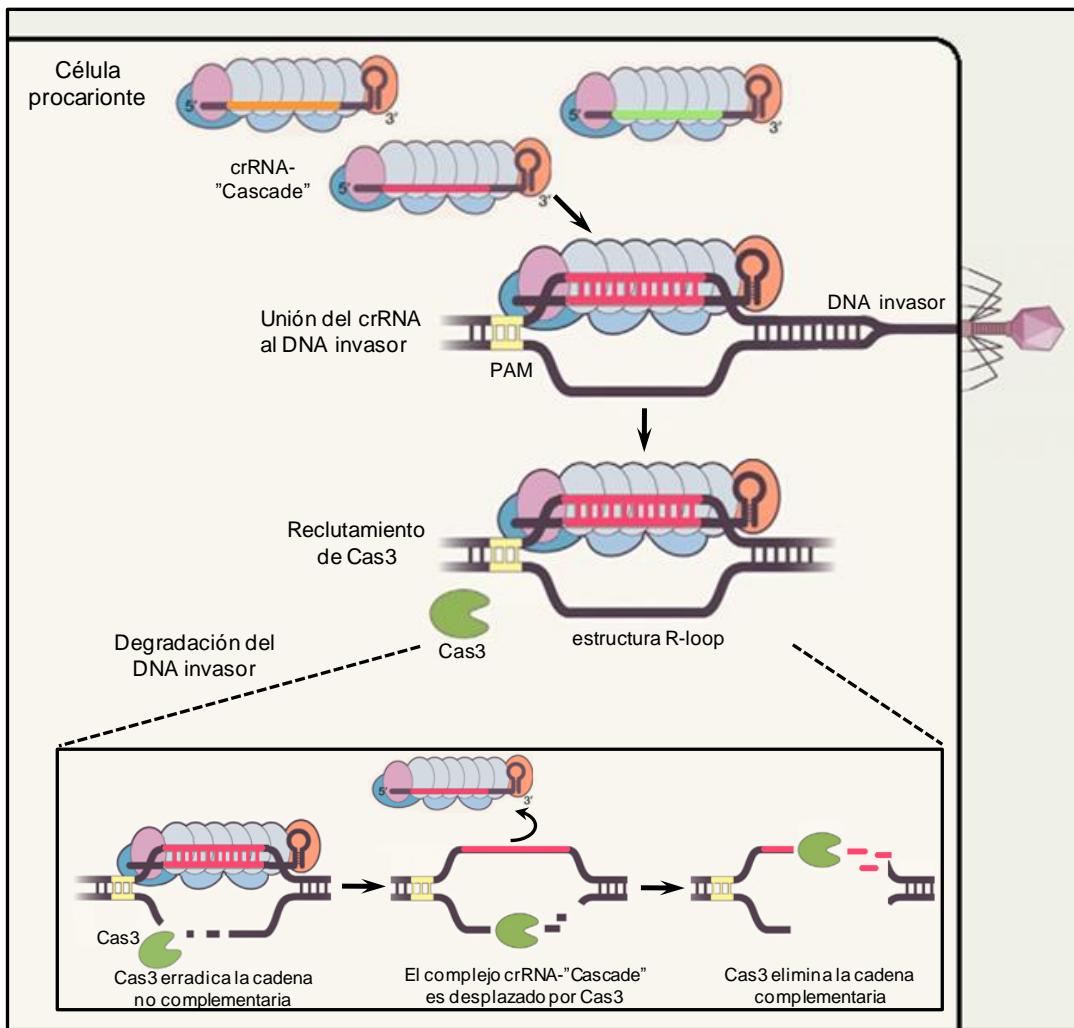


**Figura 4.5. Expresión o biogénesis de los crRNAs.** Los genes *cas* y el arreglo CRISPR se expresan cuando el procarionte enfrenta un nuevo ataque viral. El locus CRISPR se transcribe formando un largo pre-crRNA, en el cual las secuencias repetidas forman una estructura tallo-asa que es reconocida por el complejo "Cascade", conformado por Cse1, Cse2, Cas7, Cas5 y Cas6e. La nucleasa Cas6e procesa el transcrito produciendo los crRNAs maduros, a los cuales permanece unido "Cascade". Este complejo crRNA-"Cascade" es necesario para realizar la interferencia del elemento genético externo (modificada de Hille *et al.*, 2018).

### c) Interferencia

En la etapa de expresión se generan los crRNAs, los cuales son fundamentales en la fase de interferencia, ya que estos pequeños RNAs guían al complejo "Cascade" hacia el ácido nucleico foráneo. El pequeño crRNA con homología al material genético invasor hibrida con su cadena complementaria, desplazando a la hebra opuesta del DNA foráneo, es decir, se forma una estructura "R-loop", cuya presencia promueve el reclutamiento de Cas3 (Jore *et*

*al.*, 2011; Mulepati & Bailey, 2013; Westra *et al.*, 2012). Esta proteína contiene un dominio N-terminal HD de nucleasa de cadena sencilla (ssDNAsa) y un dominio de helicasa en su extremo carboxilo (Sinkunas *et al.*, 2011). Mediante su actividad nucleasa degrada a la cadena no complementaria en dirección 3'→5' (Mulepati & Bailey, 2011; Sinkunas *et al.*, 2011; Westra *et al.*, 2012), y con su dominio de helicasa dependiente de ATP promueve el desplazamiento de crRNA-Cascade del ácido nucleico invasor, generando una nueva cadena sencilla de DNA, la cual queda disponible para la degradación por Cas3 (Howard *et al.*, 2011; Sinkunas *et al.*, 2013; Westra *et al.*, 2012). Por lo tanto, el material genético foráneo es erradicado y la bacteria sobrevive a la infección (Figura 4.6).



**Figura 4.6. Etapa de interferencia.** Los crRNAs guían al complejo “Cascade” al DNA invasor. El crRNA con homología a este material genético foráneo hibrida con la cadena complementaria, por lo que se forma una estructura R-loop y la nucleasa Cas3 es reclutada. Cas3 degrada la cadena no

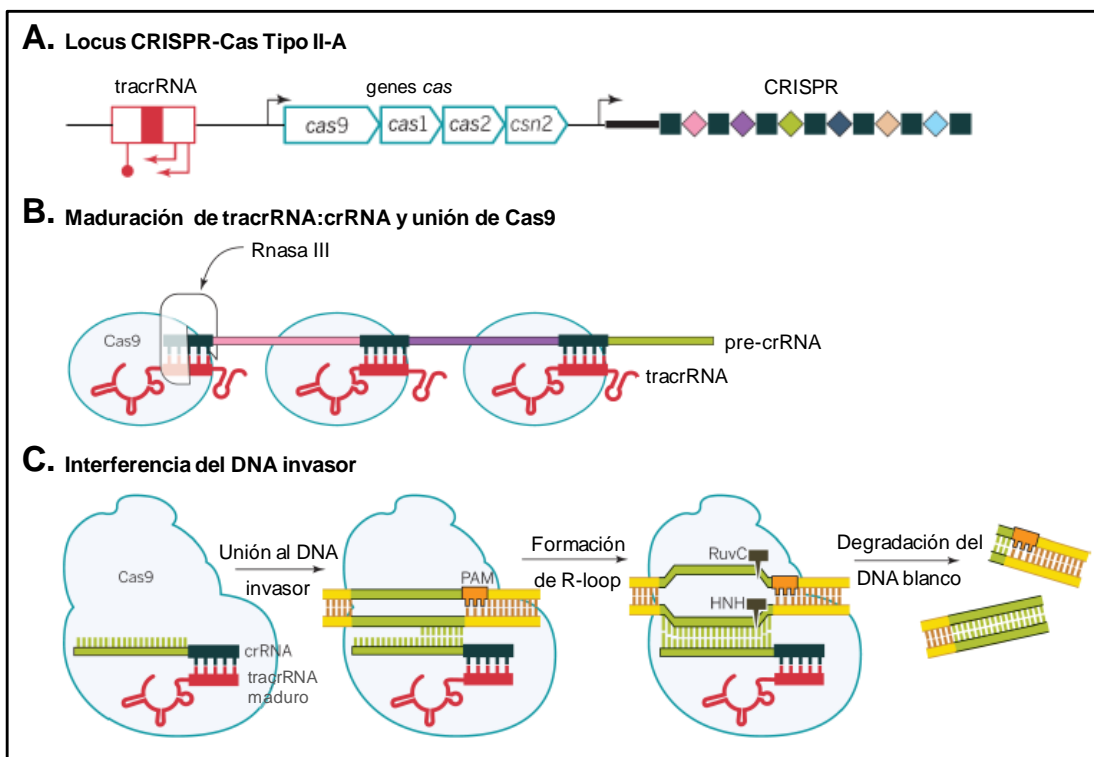
complementaria en el DNA invasor. Posteriormente, desplaza al complejo crRNA-“Cascade” y erradica la segunda cadena del ácido nucleico. De esta manera el procarionte resiste el ataque viral (modificada de Hille *et al.*, 2018).

Para que el sistema CRISPR-Cas pueda reconocer eficientemente al material genético foráneo y degradarlo, además de tener un espaciador con homología a este elemento, es necesario que detecte la presencia de un PAM adyacente a la región que hibrida con el espaciador (Fischer *et al.*, 2012; Semenova *et al.*, 2011; Westra *et al.*, 2012) (Figura 4.6). La ausencia de PAMs en el locus CRISPR-Cas del genoma hospedero evita que los crRNAs reconozcan los espaciadores endógenos, previniendo eventos de autoinmunidad (Marraffini & Sontheimer, 2010; Westra *et al.*, 2013).

El tipo de sistema CRISPR-Cas previamente descrito se encuentra ampliamente distribuido en bacterias y arqueas (Makarova *et al.*, 2011; Makarova *et al.*, 2015). Sin embargo, en la actualidad diversos grupos de trabajo han centrado su atención en los sistemas Tipo II-A de la Clase 2, debido a su potencial como herramienta en ingeniería genética. Con esta estrategia se han generado mutantes en plantas, insectos, ratones y en una gran cantidad de organismos eucariontes y procariontes (Doudna & Charpentier, 2014; Hryhorowicz *et al.*, 2017; Torres-Ruiz & Rodriguez-Perales, 2017).

Para la caracterización de los sistemas Tipo II-A se ha utilizado como modelo a la bacteria *Streptococcus pyogenes*, en la cual se ha sugerido que el sistema CRISPR-Cas funciona en el control de profagos. La estructura genética del clúster CRISPR-Cas de este microorganismo consta de un RNA *trans*-activador (*tracrRNA*), corriente abajo se encuentran los genes *cas9*, *cas1*, *cas2* y *csn2*, finalmente se localiza un locus CRISPR (Figura 4.7A) (Chylinski *et al.*, 2013; Deltcheva *et al.*, 2011). El mecanismo de acción de este sistema se lleva a cabo de la siguiente manera: la adquisición e integración de espaciadores en el arreglo CRISPR es mediada por Cas1 y Cas2 como se mencionó anteriormente. Posteriormente, para el procesamiento del pre-crRNA se requiere la presencia del *tracrRNA*, el cual posee 25 nt complementarios a las secuencias repetidas del pre-crRNA por lo que ambas estructuras hibridan y forman dúplex de RNA que son

estabilizados por Cas9 y escindidos por la enzima RNAsa III (Figura 4.7B) generándose híbridos crRNA-tracrRNAs que guían a la proteína Cas9 al elemento invasor (Deltcheva *et al.*, 2011; Hille & Charpentier, 2016). Esta enzima es capaz de degradar el DNA foráneo debido a que posee dos diferentes dominios con actividad nucleasa, con su dominio HNH degrada la cadena de DNA complementaria al crRNA, mientras que el dominio tipo RuvC se encarga de eliminar la hebra opuesta del ácido nucleico invasor (Figura 4.7C) (Jinek *et al.*, 2012).



**Figura 4.7. Mecanismo de acción del sistema CRISPR-Cas Tipo II-A.** **A.** Organización genética del locus CRISPR-Cas presente en *S. pyogenes*. El tracrRNA se presenta en rojo, los genes *cas* con borde azul y, en el locus CRISPR los cuadros negros representan las secuencias repetidas y los espaciadores se ilustran en rombos de colores. **B.** Generación de los híbridos tracrRNA:crRNA por la acción de la enzima RNAsa III. **C.** El crRNA unido al tracrRNA guían a la nucleasa Cas9 al material genético foráneo. Esta enzima degrada la doble cadena de DNA gracias a sus dominios RuvC y HNH (modificada de Doudna & Charpentier, 2014).

Como se describió, CRISPR-Cas es un sistema inmune eficiente en la degradación de material genético proveniente de virus o plásmidos. Sin embargo, los bacteriófagos han desarrollado diversas estrategias para contrarrestar la actividad de este sistema de defensa procariótico. Los fagos pueden escapar al

reconocimiento por CRISPR-Cas mutando las secuencias PAM o la región de homología con el espaciador correspondiente (Deveau *et al.*, 2008). No obstante, las bacterias contienden con esto mediante la adquisición de múltiples espaciadores, asegurando una defensa exitosa (Datsenko *et al.*, 2012; Fineran *et al.*, 2014).

En esta lucha continua de bacterias versus virus, estos últimos también codifican pequeñas proteínas conocidas como anti-CRISPR (Acr), las cuales son capaces de interrumpir el proceso de interferencia mediado por el sistema CRISPR-Cas (Bondy-Denomy *et al.*, 2013; Hille *et al.*, 2018). Hasta la fecha se han reportado 22 familias distintas de proteínas Acr, siendo las proteínas AcrF1, AcrF2 y AcrF3, presentes en los fagos tipo Mu que infectan a *Pseudomonas aeruginosa*, las mejor caracterizadas (Zhu *et al.*, 2018). AcrF1 y AcrF2 interactúan directamente con el complejo tipo “Cascade” del sistema CRISPR-Cas Tipo I-F de *P. aeruginosa* bloqueando su unión al DNA invasor y por lo tanto evitando la degradación. Por otra parte, AcrF3 se une a Cas3 impidiendo que esta nucleasa elimine el material genético foráneo (Bondy-Denomy *et al.*, 2015; Bondy-Denomy *et al.*, 2013). Recientemente, se han reportado profagos presentes en el genoma de *Listeria monocytogenes* que codifican las proteínas AcrIIA1, AcrIIA2, AcrIIA3 y AcrIIA4 capaces de inhibir la actividad de la nucleasa Cas9 (Rauch *et al.*, 2017). Más aún, se ha encontrado evidencia de que algunos fagos contienen una proteína tipo H-NS, con la cual podrían reprimir la expresión del sistema CRISPR-Cas bacteriano y realizar la infección (Skenneron *et al.*, 2011).

Es evidente la existencia de una continua competencia virus-bacterias. Sin embargo, gracias a los mecanismos moleculares desarrollados por ambos participantes su prevalencia ha sido posible a lo largo de la evolución. Es fundamental continuar el estudio de esta interacción lo cual sin duda generará conocimiento novedoso sobre esta estrecha relación biológica ancestral.

#### 4.4.2 Papel del sistema CRISPR-Cas en diversos procesos biológicos bacterianos

Es claro que en diversos organismos procariontes CRISPR-Cas constituye un sistema inmune que protege a estas poblaciones de la depredación por fagos y la transferencia horizontal de genes. Sin embargo, este sistema también ha sido relacionado con otras funciones biológicas. Por ejemplo, en *Myxococcus xanthus* el clúster CRISPR-Cas participa en el proceso de esporulación y la producción de exopolisacárido (Viswanathan *et al.*, 2007; Wallace *et al.*, 2014). En *Streptococcus mutans*, cepas deficientes en las proteínas Cas muestran menor tolerancia al estrés oxidativo y a choque térmico (Serbanescu *et al.*, 2015). Adicionalmente, los genes *cas* de dicha bacteria se inducen durante el desarrollo de biofilm en presencia de almidón o sacarosa (Klein *et al.*, 2010; Louwen *et al.*, 2014). En este sentido, el sistema CRISPR-Cas de *Pseudomonas aeruginosa* PA14, está implicado en la inhibición de la formación del biofilm y la movilidad tipo swarming (Zegans *et al.*, 2009). También en *P. aeruginosa*, un espaciador similar a *lasR*, Cas3 y el complejo Csy1-Csy4, están involucrados en la degradación del mRNA correspondiente al regulador de quorum-sensing LasR, lo que promueve una disminución en la respuesta inflamatoria de células humanas de pulmón y el establecimiento de la infección (Li *et al.*, 2016).

Otras bacterias patógenas como *Campylobacter jejuni* y *Neisseria meningitidis* requieren de Cas9 para la invasión y replicación en células epiteliales de intestino y pulmón, respectivamente (Louwen *et al.*, 2013; Sampson *et al.*, 2013). En el caso de *Legionella pneumophila*, agente causal de la enfermedad del legionario en el humano, la actividad de RNAsa y DNAsa de Cas2 es requerida para que la bacteria infecte diferentes especies de amebas tales como *Acanthamoeba castellanii*, *Hartmannella vermiformis* y *Willaertia magna*, nichos necesarios para la replicación del patógeno (Gunderson & Cianciotto, 2013; Gunderson *et al.*, 2015).

En microorganismos comensales como *E. coli* K-12, la nucleasa Cas1 ha sido implicada en la reparación de ácidos nucleicos, ya que es capaz de procesar



intermediarios de Holliday e interactuar con RecB, RecC y RuvB, elementos implicados en reparación de DNA. Más aún, cepas deficientes en *cas1* presentan mayor sensibilidad al daño en el DNA provocado por la exposición a mitomicina C o luz UV (Babu *et al.*, 2011). En *Xenorhabdus nematophila* el RNA denominado NiID (codificado en el locus CRISPR) y la proteína Cas6e son necesarios para que esta bacteria pueda colonizar nematodos del género *Steinernema* (Veesenmeyer *et al.*, 2014).

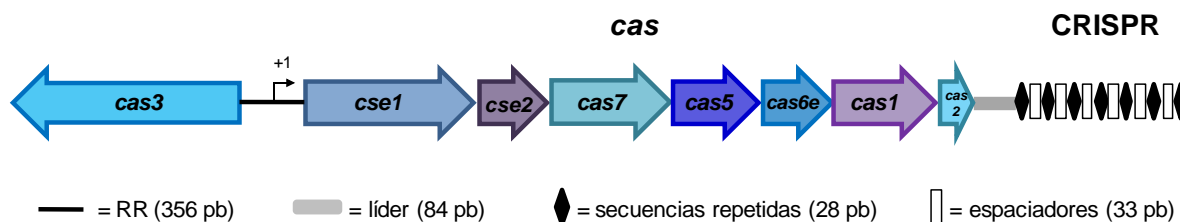
A diferencia de los fenotipos mencionados, en el patógeno de humanos *Francisella novicida* se describió mecanísticamente el papel que desempeña el sistema CRISPR-Cas en virulencia, en este caso la proteína Cas9, el tracrRNA y un RNA llamado scaRNA (small CRISPR-associated RNA) degradan el mRNA de la lipoproteína membranal BLP disminuyendo la síntesis de esta proteína, lo cual provoca que el patógeno presente una alta resistencia a antibióticos y evada el sistema inmune innato del hospedero, produciendo la infección en el modelo del ratón (Sampson *et al.*, 2014; Sampson *et al.*, 2013).

El papel de CRISPR-Cas como mecanismo de defensa ha sido demostrado en diferentes microorganismos, pero en otros, como los mencionados anteriormente, este sistema participa en la fisiología bacteriana. Por lo tanto, es necesario determinar la función de este locus en otros comensales, patógenos y simbioses con el objetivo de definir la diversidad de procesos en los cuales puede estar implicado este programa genético.

## 5. ANTECEDENTES

### 5.1 El sistema CRISPR-Cas en *Salmonella enterica* serovar Typhi

El sistema CRISPR-Cas se encuentra ampliamente distribuido en bacterias y arqueas. En la familia *Enterobacteriaceae*, el análisis de 228 genomas secuenciados reveló que este locus está presente en el 57% de dichas bacterias. *Salmonella enterica* serovar Typhi contiene un locus CRISPR-Cas Tipo I-E constituido por una región reguladora 5' de 356 pb; ocho genes *cas* (*cas3*, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, *cas2*), una región líder (84 pb); siete secuencias repetidas de 28 pb y seis espaciadores de 33 pb (Figura 5.1).



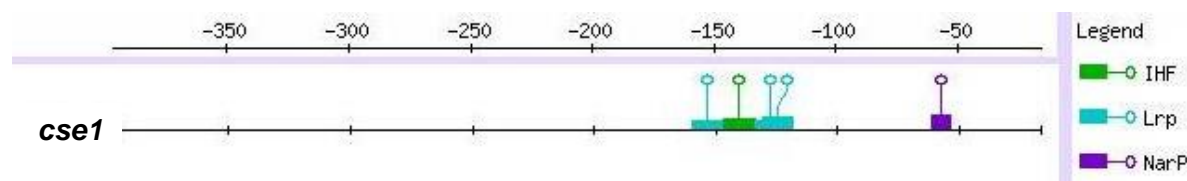
**Figura 5.1. El locus CRISPR-Cas de *S. Typhi*.** Representación esquemática del sistema CRISPR-Cas Tipo I-E presente en *Salmonella enterica* serovar Typhi IMSS-1. La flecha negra representa el inicio de transcripción del gen *cse1*.

Análisis *in silico* de los espaciadores de la cepa *S. Typhi* IMSS-1, utilizando el algoritmo BLASTn, determinaron que no tienen homología con alguna secuencia de fagos o virus disponible en “DDBJ data bank”. Adicionalmente, estudios del locus CRISPR-Cas presente en diferentes serovares de *Salmonella enterica*, demuestran que la adquisición de espaciadores es limitada, sugiriendo que este sistema presenta poca actividad inmunológica en este género bacteriano (Shariat *et al.*, 2015). Más aún, en experimentos de nuestro laboratorio, utilizando una cepa silvestre de *S. Typhi* conteniendo un plásmido por 114 generaciones, no fue posible detectar la presencia de nuevos espaciadores en el arreglo CRISPR. Sin embargo, la alta conservación del locus CRISPR-Cas en *Salmonella*, sugiere que podría estar participando en un proceso biológico diferente a la defensa contra material genético foráneo (Shariat *et al.*, 2015; Touchon & Rocha, 2010). Más aún, ensayos de transcriptómica detectaron la expresión del gen *cas7* cuando *S. Typhi* se encuentra dentro de macrófagos de humanos (Faucher *et al.*, 2005). Además,

perfiles de expresión realizados con *S. Typhi* proveniente de sangre de pacientes con fiebre tifoidea demostraron que *cas3* se induce en dicha condición (Louwen *et al.*, 2014; Sheikh *et al.*, 2011). Por lo tanto, el sistema CRISPR-Cas podría estar participando en el ciclo de infección de este patógeno.

Una manera de identificar la función del clúster CRISPR-Cas en *S. Typhi*, consiste en determinar los factores transcripcionales y señales ambientales implicados en el control positivo y negativo de su expresión genética. No obstante, existe un número limitado de reportes sobre la regulación transcripcional de este sistema inmune en procariontes. Respecto a esto, nuestro grupo de trabajo definió, mediante experimentos de proteómica, que la sobreexpresión del regulador LeuO, miembro de la familia tipo LysR, induce la expresión de Cas7, el tercer componente del locus *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR* y mediante fusiones transcripcionales con la región promotora de *cse1*, se demostró que dicho ORF también es regulado positivamente por LeuO (Hernández-Lucas *et al.*, 2008). Sin embargo, no se ha determinado si este factor transcripcional es capaz de regular al resto de los genes *cas* presentes en *S. Typhi*.

Además de los datos sobre el papel ejercido por LeuO en la inducción de *cse1*, se realizaron análisis bioinformáticos de la región reguladora de este gen. Utilizando el programa *matrix-scan* se analizaron 400 pb corriente arriba del ATG de *cse1* y comparando con las PSSMs (Position-specific scoring matrices) disponibles en RegulonDB, se determinó la presencia de sitios de unión para los reguladores IHF, Lrp y NarP (Figura 5.2).



**Figura 5.2. Posibles proteínas reguladoras involucradas en la regulación transcripcional del locus CRISPR-Cas.** Sitios de unión determinados *in silico* para los factores transcripcionales IHF, Lrp y NarP en la región intergénica 5' de *cse1*. Los números representan las coordenadas respecto al inicio de traducción de *cse1*.

El mecanismo de regulación ejercido por LeuO en el control del promotor *cse1*, su papel en la regulación positiva de los diferentes genes *cas*, la presencia de diversas unidades transcripcionales en el locus CRISPR-Cas, la posible participación de las proteínas IHF, Lrp, NarP u otros reguladores en el control transcripcional de este sistema, las señales que influyen en su expresión, así como su papel biológico, no han sido definidos en *Salmonella enterica* serovar Typhi. Debido a lo anterior, en este estudio se plantea la siguiente hipótesis y objetivos de trabajo.

## 6. HIPÓTESIS

El sistema CRISPR-Cas de *Salmonella enterica* serovar Typhi es modulado por reguladores transcripcionales tales como LeuO, Lrp, IHF y NarP, así como por condiciones intracelulares propias de la célula hospedante. Por lo tanto, podría participar en procesos biológicos diferentes a la inmunidad bacteriana.

## 7. OBJETIVOS

### 7.1 Objetivo general

Caracterizar la regulación transcripcional y función biológica del sistema CRISPR-Cas Tipo I-E presente en *Salmonella enterica* serovar Typhi.

### 7.2 Objetivos particulares

- Elucidar el mecanismo ejercido por LeuO en el control transcripcional del sistema CRISPR-Cas de *S. Typhi*.
- Determinar el papel de IHF, Lrp y NarP en la regulación positiva o negativa de los componentes del clúster CRISPR-Cas de *S. Typhi*.
- Identificar las unidades transcripcionales presentes en el locus CRISPR-Cas de *S. Typhi*.
- Establecer las señales ambientales que influyen en la expresión del locus CRISPR-Cas de *S. Typhi*.
- Determinar el proceso biológico en el cual está involucrado el sistema CRISPR-Cas en *S. Typhi*.

## 8. MATERIALES Y MÉTODOS

### 8.1 Cepas bacterianas, medios de cultivo y condiciones de crecimiento.

Las cepas bacterianas utilizadas en el presente proyecto se enlistan en la Tabla 1 del Anexo I. Las cepas de *S. Typhi* y *E. coli* se cultivaron a 37°C bajo condiciones aeróbicas en medio LB (1% de triptona, 0.5% de extracto de levadura y 1% de NaCl, pH 7.4). *S. Typhi* también fue cultivada en medio rico MA (7 g de caldo nutritivo, 1 g de extracto de levadura, 2 ml de glicerol, 3.75 g de K<sub>2</sub>HPO<sub>4</sub> y 1.3 g de KH<sub>2</sub>PO<sub>4</sub> por litro) (Kawaji *et al.*, 1979) y medio mínimo-N (N-MM) [KCl 5 mM, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 7.5 mM, K<sub>2</sub>SO<sub>4</sub> 0.5 mM, KH<sub>2</sub>PO<sub>4</sub> 1 mM, MgCl<sub>2</sub> 200 µM, glicerol 0.5%, casaminoácidos 1% y Tris-HCl (pH 7.5) 100 mM] (Deiwick *et al.*, 1999). Para los experimentos de pH se utilizó Tris-HCl (pH 7.5) a una concentración final de 200 mM y el valor de pH se ajustó con HCl o NaOH.

Cuando fue necesario, los medios se suplementaron con kanamicina (Km), 30 µg/ml; tetraciclina (Tc), 12 µg/ml; ampicilina (Ap), 200 µg/ml o X-gal (5-bromo-4-cloro-3-indolil-β-D-galactopiranosido), 40 µg/ml.

### 8.2 Manipulación de DNA

El DNA cromosómico de las cepas se aisló mediante el estuche comercial “Promega, Madison”. El diseño de oligonucleótidos para las diferentes reacciones en cadena de la polimerasa (PCRs) se realizó con el programa OLIGO 6 y fueron generados en la Unidad de Síntesis del Instituto de Biotecnología (UNAM), en la Tabla 3 del Anexo I se enlistan los oligonucleótidos utilizados en este trabajo. Las reacciones de PCR se amplificaron en el termociclador GeneAmp PCR System 9700 (Roche Applied Biosystems). Las enzimas de restricción, ligasas, polimerasas y nucleótidos, se obtuvieron de las compañías New England Biolabs, Invitrogen, Stratagene y Thermo Scientific. Los plásmidos y productos de PCR fueron purificados con “High Pure Plasmid Isolation Kit” y “High Pure PCR Purification Kit” (Roche Applied Science, Germany), respectivamente. Las

muestras de DNA fueron secuenciadas con el sistema Perkin Elmer/Applied Biosystems 377-18 del Instituto de Biotecnología (UNAM).

### 8.3 Generación de mutantes por recombinación homóloga

Para generar las mutantes en los genes *lrp*, *hns*, *ssrB*, *phoP*, *ompR*, *crp*, *cpxRA*, *ihf*, *narP*, *arcA*, *pmrAB* y *fnr*, se utilizó el método descrito por Datsenko y Wanner (2000). De manera resumida, la cepa *S. Typhi* IMSS-1 silvestre se transformó con el plásmido termosensible pKD46 (Ap<sup>r</sup>), el cual contiene el sistema de recombinasa Red. La bacteria se creció a 30°C y se seleccionaron aquellas colonias resistentes a ampicilina.

Posteriormente, se realizaron PCRs utilizando como molde el plásmido pKD4 que contiene las secuencias FRT adyacentes al gen de resistencia a kanamicina (Km) y oligos correspondientes al gen que se desea inactivar. Los primers contenían parte de la secuencia del gen de Km y del ORF de interés. Los productos de PCR fueron transformados en la cepa IMSS-1 albergando el plásmido pKD46. Las transformaciones se sembraron en agar LB con Km incubándose a 37°C. Las colonias obtenidas se estriaron en LB-Km y LB-Ap y se seleccionaron aquellas colonias resistentes a Km y sensibles a Ap, indicando la ausencia del plásmido pKD46 y la presencia del interposón de Km.

El DNA genómico de las mutantes fue amplificado y secuenciado con el objetivo de corroborar la eliminación deseada.

### 8.4 Construcción de fusiones transcripcionales

Se amplificó DNA de diferentes longitudes de las regiones reguladoras de *cas3*, *cse1*, *scse2*, *ascas2-1* y del locus CRISPR-Cas utilizando la enzima *Pfu Turbo*<sup>®</sup> DNA polymerase (Stratagene). Los fragmentos de DNA se purificaron y digirieron con las enzimas de restricción *BamH I-Kpn I*, *BamH I-Hind III* o *BamH I-Xho I*, para ser clonadas al plásmido pKK232-9 (Km<sup>r</sup>) o pKK232-8 (Ap<sup>r</sup>), los cuales contienen al gen reportero *cat*. En el caso de la región reguladora de *ascse2-1*,

después de ser amplificada se digirió con la enzima *BamH* I y se ligó al plásmido pRS415 ( $Ap^r$ ), que alberga al reportero *lacZ*. Para fusiones transcripcionales con una longitud menor a 100 pb, se sintetizaron oligonucleótidos complementarios (dúplex) conteniendo extremos cohesivos tanto en la región 5' como en la 3' (Tabla 3). Cada primer se diluyó a una concentración de 50 ng/ $\mu$ l y 500 ng de cada oligonucleótido se colocaron en un tubo eppendorf. La mezcla de primers se incubó a 95°C por 10 min y 30 min a temperatura ambiente para generar la doble cadena de DNA, la cual se clonó al vector correspondiente.

Los plásmidos recombinantes se transformaron por choque térmico en células *E. coli* DH5 $\alpha$ , seleccionándose las clonas por la resistencia al antibiótico correspondiente. Los plásmidos fueron purificados y posteriormente secuenciados en la Unidad de Síntesis y Secuenciación del Instituto de Biotecnología (UNAM).

Cada fusión transcripcional fue transformada en las cepas correspondientes de *S. Typhi* IMSS-1 para realizar cinéticas de expresión y determinar la actividad específica de CAT o  $\beta$ -galactosidasa.

## 8.5 Mutagénesis sitio dirigida

Las sustituciones en los sitios de pegado de LeuO y en la caja -10 de la región promotora de *cse1* se realizaron con el kit de Stratagene (QuickChange®XL Site-Directed Mutagenesis Kit).

Los PCRs fueron sintetizados con oligonucleótidos complementarios conteniendo la sustitución deseada (Tabla 3) y como templado el plásmido pKK9-*cse1*. Las condiciones utilizadas para la amplificación fueron las siguientes:

95°C	3 min	} 18 ciclos
95°C	3 min	
50°C	3 min	
68°C	10 min	
68°C	10 min	



Las reacciones de PCR fueron digeridas con 2  $\mu$ l de enzima *Dpn* I, para eliminar el DNA metilado utilizado como templado, posteriormente los plásmidos sintetizados *in vitro* se transformaron en la cepa DH5 $\alpha$  de *E. coli*, las clonas se seleccionaron por resistencia al antibiótico correspondiente y perfiles de plásmido tipo Eckhardt fueron realizados para seleccionar las clonas con el vector recombinante correcto, los cuales fueron purificados y las mutaciones se confirmaron mediante secuenciación.

Los plásmidos con la mutación correcta fueron transformados en *S. Typhi* conteniendo pFM*Trc*12 y *S. Typhi* con pFM*Trcleu*O-50, para realizar los ensayos de actividad CAT correspondientes.

## 8.6 Ensayos de actividad específica CAT

Las cepas de *S. Typhi* fueron cultivadas en 5 ml de medio LB durante 16 h a 37°C, 0.5 ml de estos preinóculos fueron utilizados para inocular 100 ml de medio nutritivo MA por duplicado, con o sin 50  $\mu$ M de IPTG, según corresponda. En el caso de medio mínimo N, tres mililitros de un cultivo de 16 h a 37°C de las cepas a evaluar, fueron centrifugados, lavados e inoculados en matraces conteniendo 50 ml de dicho medio. Los matraces se incubaron a 37°C, 200 rpm. Se tomaron 1.5 ml de los cultivos de una OD<sub>(595nm)</sub> de 0.4, 0.6 y 0.8, 1 ml de OD<sub>(595nm)</sub> de 1 y 0.6 ml de OD<sub>(595nm)</sub> 1.2-1.4. Las muestras se centrifugaron y lavaron con TDDT (Tris-HCl 50 mM, pH 7.8 y DL-Ditiotreitol 30  $\mu$ M) resuspendiéndolas en 0.6 ml de éste reactivo. Posteriormente, fueron sonicadas 3 min (pulsos de 9.9 s ON - 9.9 s OFF) y centrifugadas 15 min a 12,000 rpm transfiriendo el sobrenadante a tubos eppendorf limpios.

Con el sobrenadante se realizó la determinación de proteínas totales y actividad de CAT. Las proteínas totales se cuantificaron por el método espectrofotométrico utilizando el BCA Protein Assay Kit Reagent (Pierce). 10  $\mu$ l de cada muestra fueron colocados por duplicado en una placa de ELISA de 96 pozos (Costar), agregando 200  $\mu$ l de la mezcla de reacción del kit. Las placas se incubaron 30 min a 37°C y se midió la concentración de proteínas en un lector

Ceres 900-C (Bio-Tek Instruments Inc.). Las proteínas de la muestra reducen el  $\text{Cu}^{2+}$  que interacciona con el BCA (Ácido bicinconínico) formando un compuesto colorimétrico detectable a 562 nm.

La actividad de CAT se determinó colocando por duplicado 5  $\mu\text{l}$  de cada extracto en una placa de ELISA de 96 pozos (Costar) y agregando 200  $\mu\text{l}$  de una mezcla conteniendo 1 mM de DNTB (ácido 5,5'-ditiobis-2-nitrobenzoico), 0.1 mM de acetil-coenzima A (acetil-CoA) y 0.1 mM de cloranfenicol, los tres componentes disueltos en Tris-HCl, pH 7.8 al 0.1 M. La absorbancia se midió a 412 nm cada 5 s durante 5 minutos utilizando el lector Ceres. La enzima CAT transfiere el grupo acetilo de la acetil-CoA al cloranfenicol inactivándolo y el CoA reducido (CoA-SH) reacciona con el DNTB produciendo disulfuro de CoA, ácido tionitrobenzoico y un equivalente molar de 5-tio-2-nitrobenzoato, compuesto que absorbe a una longitud de 412 nm. Una unidad de CAT equivale a 1  $\mu\text{mol}$  de cloranfenicol acetilado por minuto.

Finalmente, la actividad específica CAT se obtuvo a través de la siguiente ecuación:

$$\text{Actividad específica} = \frac{\text{Promedio CAT}}{[\text{Proteínas}]} \quad (\mu\text{mol min}^{-1}\text{mg}^{-1})$$

### 8.7 Determinación de actividad $\beta$ -galactosidasa

Las muestras fueron colectadas como se mencionó anteriormente en los ensayos de CAT. Las pastillas se lavaron con 0.8 ml de buffer Z 1X y se resuspendieron en 0.6 ml de este amortiguador. La concentración de proteínas fue determinada por el método de Lowry. Se colocaron 20  $\mu\text{l}$  de cada muestra por duplicado y del control (Buffer Z 1X) en una placa de ELISA de 96 pozos (Costar), se añadieron 100  $\mu\text{l}$  de la solución de reacción D ( $\text{Na}_2\text{CO}_3$  al 3.4%, NaOH 0.17 N, tartrato de sodio y potasio al 1.7%, sulfato cúprico al 0.85%), mezclando suavemente. La placa se incubó 10 min a temperatura ambiente, después se agregaron 100  $\mu\text{l}$  de reactivo folin-ciocalteu al 16.9%, se mezcló e incubó por 30 min a temperatura ambiente. La concentración de proteínas se midió a una

absorbancia de 562 nm en un lector Ceres 900-C (Bio-Tek instruments Inc.). El protocolo utilizado compara la absorbancia de las muestras con una curva estándar de concentraciones de BSA, previamente calculada. La concentración de proteína se ajusta en un intervalo de 0.15 a 0.3 mg/ml, una vez ajustada la concentración, se determinó la actividad  $\beta$ -galactosidasa.

Para la cuantificación de la actividad  $\beta$ -galactosidasa se colocaron por duplicado 20  $\mu$ l de la muestra en una placa de 96 pozos, se agregaron 100  $\mu$ l de solución A (agua destilada, Buffer Z 1X, lisozima 0.22 mg/ml, Triton al 0.22%,  $\beta$ -mercaptoetanol 16 mM) y la placa se incubó por 10 minutos a 37°C; posteriormente se agregó 100  $\mu$ l de ONPG (o-nitrofenil  $\beta$ -D-galacto pironosido, 0.99 ml de agua), que al entrar en contacto con la enzima  $\beta$ -galactosidasa vira a un color amarillo. La absorbancia de las muestras fue medida con un lector tipo Ceres a 415 nm cada 5 s durante 5 minutos.

La actividad específica  $\beta$ -galactosidasa se obtiene al dividir el valor de la actividad de la enzima entre la concentración de proteína de cada muestra; el valor se expresa en U/mg de proteína.

## 8.8 Purificación de RNA

La purificación de RNA total para los experimentos de RT-PCR y “primer extension” se realizó por el método de fenol-ácido como se describe a continuación:

Las cepas bacterianas IMSS-1 + pFM*Trc*12, IMSS-1 + pFM*Trc*leuO-50 e IMSS-1  $\Delta$ cas-CRISPR fueron inoculadas en 100 ml de medio MA a 37°C, 200 rpm y 10 ml de cultivo a una OD<sub>595nm</sub> de 0.6 fueron colectados. En el caso de *S. Typhi* IMSS-1 silvestre, IMSS-1  $\Delta$ cas-CRISPR e IMSS-1 + pKK *cas3* (-83/-25) se incubaron en 50 ml de medio N-MM a 37°C, 200 rpm y se colectaron 20 ml de cada cultivo a una OD<sub>595nm</sub> de 1.3.

Las muestras se centrifugaron a 7,000 rpm por 5 min a 4°C. La pastilla de células proveniente de 10 ml se resuspendió en 0.5 ml de solución A (SDS al

0.5%, acetato de sodio pH 4.8 20 mM y EDTA pH 8 10 mM, todos fueron disueltos en H<sub>2</sub>O libre de RNAsas), mientras el paquete celular de 20 ml de cultivo se mezcló en 1 ml de solución A y se dividió en dos tubos eppendorf, para seguir con el procedimiento. Las muestras se vortexearon vigorosamente y 0.5 ml de fenol-ácido a 65°C fueron adicionados, se agitó por 15 s en vortex y se incubó a 65°C por 10 min, agitando a intervalos de 50 s. Posteriormente, se centrifugó 12 min a 12,000 rpm, la fase acuosa fue transferida a un tubo eppendorf para adicionarle 1 ml de etanol absoluto y centrifugar 5 min a 13,000 rpm. Al precipitado obtenido se le agregó 1 ml de etanol al 70% y nuevamente se centrifugó 2:30 min a 13,000 rpm para retirar todo el sobrenadante y añadir 0.4 ml de la solución A, resuspendiendo completamente el precipitado. Después, 0.5 ml de fenol-ácido fueron adicionados y se agitó vigorosamente, se centrifugó por 10 min a 13,000 rpm y la fase acuosa se transfirió a un tubo eppendorf para agregar 0.4 ml de cloroformo:alcohol isoamílico (24:1), se vortexeo y centrifugó 10 min a 13,000 rpm. La fase acuosa fue colocada en un tubo eppendorf y se adicionó 1 ml de etanol absoluto. La mezcla fue centrifugada 5 min a 13,000 rpm, el etanol fue decantado y se añadió 1 ml de etanol al 70%. Posteriormente, se centrifugó por 2:30 min a 13,000 rpm y el disolvente fue retirado con la ayuda de una micropipeta. El precipitado se secó a temperatura ambiente y con vacío por 10 min en el equipo SpeedVac SC100 (Savant). Las muestras se resuspendieron en 50 µl de H<sub>2</sub>O libre de RNAsas, juntando los duplicados si fuera el caso y finalmente, se adicionaron 20 U de inhibidor de RNAsa (Ribolock, Thermo Scientific).

La cuantificación de RNA obtenido se realizó a 260 nm y la integridad fue observada en un gel de agarosa al 1.5 %/TAE 1X.

## **8.9 RT-PCR**

Para los experimentos de RT-PCR se utilizó RNA total purificado de las cepas *S. Typhi* IMSS-1, IMSS-1  $\Delta$ cas-CRISPR, IMSS-1 + pFM*Trc*12 ó IMSS-1 + pFM*Trc*leuO-50. 20 µg de RNA fueron digeridos con 1 U de DNAsa I, RNAsa-free (Thermo Scientific) siguiendo las indicaciones recomendadas por el fabricante. Se

realizaron PCRs de *cse2*, *cas1* y *cas2* utilizando el RNA tratado con DNasa I como molde para corroborar la ausencia de DNA.

Para sintetizar el cDNA se utilizó el Kit RevertAid H Minus First Strand cDNA Synthesis (Fermentans), 1 µg de RNA y oligonucleótidos específicos (CRISPR+503 *XhoI*-R para el operón *cas*-CRISPR, *ascas1*-5556 *KpnI*-R en el caso de *ascas2-1* y *ascse1*-34 *BamHI*-R para *ascse2-1*). La reacción se llevó a cabo a 42°C por 60 min. Se utilizaron 2 µl del cDNA y los correspondientes primers específicos para amplificar los distintos fragmentos pertenecientes a las unidades transcripcionales del locus CRISPR-Cas evaluadas.

20 µl de cada producto de PCR fueron colocados en un gel de agarosa al 0.8%, 1.0% ó 1.8% con TAE 0.5X. La electroforesis se realizó a 130 volts por 80 min. Finalmente, los fragmentos de DNA fueron visualizados mediante tinción con bromuro de etidio.

### 8.10 Northern Blot

Para los ensayos de Northern Blot se obtuvo RNA total de las cepas IMSS-1 + pFM*Trc12*, IMSS-1 + pFM*TrcleuO-50* e IMSS-1  $\Delta$ *cas*-CRISPR cultivadas en MA. 8 µg de RNA fueron separados por electroforesis en un gel desnaturizante de agarosa (formaldehído 16%, MOPS 1X). Para determinar los tamaños moleculares se colocó un marcador de RNA (0.24–9.5 Kb RNA ladder; Invitrogen). Posteriormente, los RNAs fueron transferidos a una membrana de nylon (GE Healthcare). La transferencia se realizó por 24 h con buffer alcalino (0.01 N de NaOH y 3 M de NaCl).

Las membranas fueron prehibridadas e hibridadas a 65°C con buffer Rapid Hyb durante 12 h (GE Healthcare). Al término de la hibridación, se lavaron con SSC 2X/SDS 0.1% por 15 min, SSC 1X/SDS 0.1% por 15 min, SSC 0.5X/SDS 0.1% por 15 min y finalmente con SSC 0.1X/SDS 0.1% por 30 min. Finalmente, las membranas fueron incubadas durante 48 h con una película Kodak X-Omat LS.

Las sondas se marcaron con radioactividad utilizando el kit Rediprime y el isótopo [ $\alpha$ - $^{32}\text{P}$ ]dCTP.

### 8.11 Primer extension

Con el objetivo de determinar el inicio de la transcripción de *cas3* se realizaron experimentos de “primer extension”. 40  $\mu\text{g}$  de RNA total provenientes de la cepa IMSS-1+pKK *cas3* (-83/-25) y el oligonucleótido pKK8-R, marcado radioactivamente con [ $\gamma$ - $^{32}\text{P}$ ]ATP, se desnaturalizaron a 65°C durante 5 min y se dejaron enfriar a temperatura ambiente. Posteriormente, la síntesis de cDNA fue realizada con el estuche comercial Maxima H Minus Strand cDNA Synthesis kit (Thermo Scientific), siguiendo las indicaciones recomendadas por el fabricante.

El producto de reacción se precipitó con 0.5 ml de N-butanol y se centrifugó 20 min a 14,000 rpm. Al precipitado obtenido se le colocaron 0.7 ml de etanol absoluto y se centrifugó 5 min a 14,000 rpm. El sobrenadante fue retirado y se adicionaron 0.7 ml de etanol al 70%. La mezcla fue centrifugada a 14,000 rpm por 5 min y se decantó el disolvente. Finalmente, la muestra se secó en el equipo SpeedVac SC100 (Savant) y se resuspendió en 8  $\mu\text{l}$  de una solución que contiene formamida al 95%, EDTA 20 mM, azul de bromofenol al 0.05% y xileno cianol FF al 0.05%. El resultado se analizó por electroforesis en geles de secuencia (poliacrilamida 8%, urea 8M). Se realizó una reacción de secuencia por el método de Sanger utilizando como templado el plásmido pKK *cas3* (-83/-25) y el oligonucleótido pKK8-R marcado radioactivamente con [ $\gamma$ - $^{32}\text{P}$ ]ATP. Esta reacción sirvió para localizar el inicio de transcripción en el genoma.

### 8.12 Purificación de H-NS y Lrp

100 ml de medio LB fueron incubados con las cepas *E. coli* BL21 conteniendo el plásmido pMDH-NS o con *E. coli* BL21 albergando el pT6-Lrp-His<sub>6</sub>. Las bacterias se incubaron a 37°C (200 rpm) hasta alcanzar una OD<sub>595nm</sub> de 0.4. Posteriormente, la expresión se indujo durante 4 h con 0.1% de L-arabinosa.

Los cultivos (100 ml) se centrifugaron a 8,000 rpm durante 5 min (4°C), las células fueron resuspendidas en 5 ml de buffer de unión (8 M urea, 100 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris-HCl pH 8.0) y fueron sonicadas. Los restos celulares fueron eliminados centrifugando 5 minutos a 12,000 rpm. Los lisados obtenidos se colocaron en una columna de Níquel Ni-NTA Agarose (Quiagen). Posteriormente, se realizaron diez lavados con un buffer de urea 8 M pH 6.5. Finalmente, las muestras se eluyeron en urea 8 M pH 4.5.

Utilizando un gel SDS-PAGE se seleccionaron las fracciones que contenían H-NS-Myc-His6 o Lrp-His6. Las fracciones seleccionadas fueron dializadas durante 2 h en un casete Slyde-A-Lyzer10K (Pierce) a 4°C con el buffer de diálisis, el cual contiene 50 mM Tris-HCl pH 8.0, 10 mM MgCl<sub>2</sub>, 20% glicerol, 500 mM NaCl, 0.1% Tritón X-100 y 4 M de urea. El mismo buffer, conteniendo diferentes concentraciones de urea (1 M y 0.2 M), fue utilizado posteriormente para realizar la diálisis durante 2 h en cada concentración de urea. Finalmente, las muestras fueron incubadas en un amortiguador sin urea que contenía 30 mM de Tris-HCl pH 8.0, 10 mM de MgCl<sub>2</sub>, 20% de glicerol, 500 mM de NaCl, 0.1% de Tritón X-100 y 3 mM de EDTA. La concentración de proteína fue determinada por el método de Bradford.

## **8.13 Ensayos de retardamiento en gel**

### **8.13.1 EMSAs con H-NS**

Mediante PCR se amplificaron diferentes fragmentos de la región intergénica 5' de *cse1*, *cas3* y *ascse2-1* para realizar ensayos de retardo con la proteína H-NS Myc-His. 100 ng de DNA se incubaron con concentraciones crecientes de la proteína H-NS en presencia de buffer de unión al 1X (40 mM de HEPES, 8 mM de MgCl<sub>2</sub>, 50 mM de KCl, 1 mM de DTT, 0.05% de NP40 y 0.1 mg ml<sup>-1</sup> de BSA). Las mezclas se incubaron por un periodo de 20 min a temperatura ambiente y los complejos DNA-proteína fueron separados en electroforesis de gel de poliacrilamida al 6%/TBE 0.5X a 70 volts. Los complejos DNA-proteína fueron visualizados mediante tinción con bromuro de etidio.

Como control negativo, se utilizaron 242 pb de la región estructural del gen rRNA 16S de *Salmonella enterica* serovar Typhi.

### 8.13.2 EMSAs con Lrp

Para realizar ensayos de retardo con la proteína Lrp-His<sub>6</sub> se utilizaron diferentes fragmentos de la región promotora de *cse1* amplificados por PCR. Como control negativo se empleó nuevamente la región estructural del rRNA 16S. El DNA (100 ng) fue incubado con concentraciones crecientes de Lrp en presencia del buffer de unión 1X (20 mM de Tris pH 8.0, 0.4 mM de EDTA, 50 mM de NaCl, 1 mM de MgCl<sub>2</sub>, 12.5% (v/v) de glicerol, 100 ng/ml de BSA y 0.1 mM de DTT). Las reacciones se incubaron por 18 min a temperatura ambiente. Posteriormente, se realizó una electroforesis en gel de poliacrilamida 6%/TBE 0.5X a 70 volts. Los complejos DNA-proteína fueron visualizados con bromuro de etidio.

## 8.14 Purificación y electroforesis de proteínas de membrana externa

La purificación de proteínas de membrana externa se realizó de las cepas *S. Typhi* IMSS-1 e IMSS-1  $\Delta$ cas-CRISPR. Como preinóculo se creció cada cepa en 5 ml de LB en tubos de ensaye durante 14 h a 37°C, 200 rpm. 3 ml del preinóculo fueron centrifugados a 7,000 rpm durante 5 min y la pastilla se lavó agregando 0.9 ml de N-MM. Después de centrifugar 5 min a 7,000 rpm, se desechó el sobrenadante, el paquete celular se resuspendió en 0.5 ml de N-MM para inocular matraces de 250 ml conteniendo 50 ml del medio mencionado. Cada cultivo se mantuvo a 37°C/200 rpm hasta una OD<sub>595nm</sub> de 0.6 ó 1.3.

Posteriormente, se colectaron 15 ml de cada cultivo y se centrifugaron 8 min/7,000 rpm/4°C. Las pastillas se resuspendieron en 1 ml de Na<sub>2</sub>HPO<sub>4</sub> 10 mM (pH 7.2) y se centrifugaron 2 min a 12,000 rpm a temperatura ambiente, se eliminó el sobrenadante y el precipitado se resuspendió en 0.6 ml de Na<sub>2</sub>HPO<sub>4</sub> 10 mM (pH 7.2). Los homogenizados se sonicaron durante 9 min en el equipo Sonipred 150 con pulsos de 9.9 s ON y 4.5 s de reposo, y se centrifugaron 2 min a 12000 rpm.



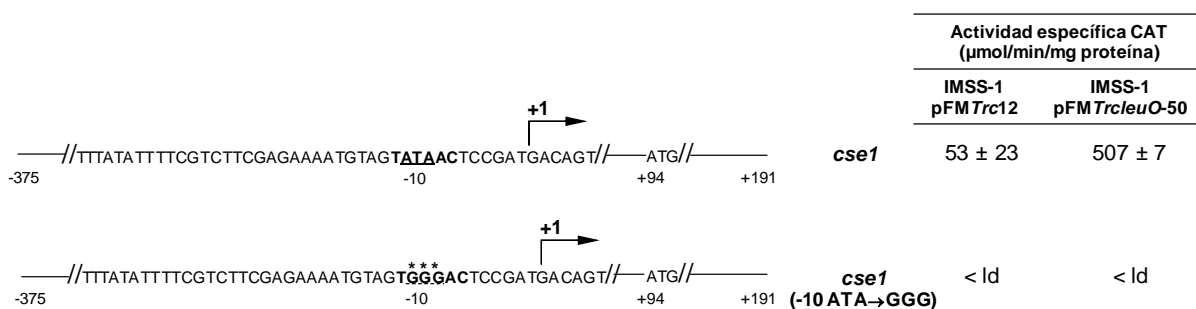
El sobrenadante obtenido fue transferido a un tubo eppendorf y se centrifugó 1 h a 12000 rpm, 4°C. El sobrenadante fue desechado, las proteínas de membrana se solubilizaron en 0.5 ml de Na<sub>2</sub>HPO<sub>4</sub> 10 mM (pH 7.2)/Triton X-100 al 2% y se incubaron a 37°C por 30 min. Transcurrido el tiempo de incubación, se centrifugó 1 h/12000 rpm/4°C, el sobrenadante se eliminó y la pastilla se lavó con 0.5 ml de Na<sub>2</sub>HPO<sub>4</sub> 10 mM (pH 7.2). Posteriormente se centrifugó 1 h/12000 rpm/4°C, se desechó el sobrenadante y finalmente las proteínas se resuspendieron en 50 µl de PBS (Búfer Fosfato Salino) 1X, pH 7.4.

La concentración de proteínas fue calculada con el estuche BCA Protein Assay Reagent (Pierce). 15 µg de cada muestra fueron analizados por electroforesis en geles de poliacrilamida 12% SDS-PAGE. La electroforesis de los geles se realizó a un voltaje de 15 mA durante 16 horas. Las proteínas de membrana externa se visualizaron mediante tinción con azul de Coomassie.

## 9. RESULTADOS

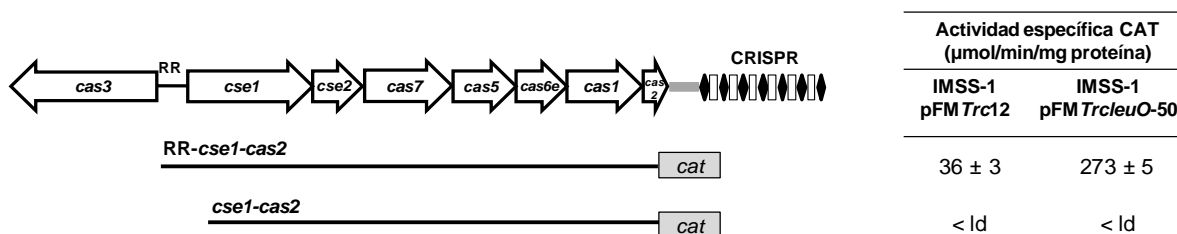
### 9.1 LeuO induce la expresión del locus *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR* en *S. Typhi*

Previamente se estableció mediante experimentos de “primer extension” y 5’ RACE el inicio de transcripción de *cse1*, el cual se localiza 94 pb corriente arriba de su ATG y su expresión es inducida por el regulador LeuO (Hernández-Lucas *et al.*, 2008). Para validar la presencia de un solo inicio de la transcripción dependiente de LeuO, se realizó mutagénesis sitio dirigida de la caja -10 del promotor *cse1* (ATA→GGG). Fragmentos de PCR conteniendo la secuencia de la región reguladora silvestre y la sustitución ATA→GGG fueron clonados independientemente río arriba del gen reportero *cat* y su actividad se evaluó en la cepa de *S. Typhi* IMSS-1 conteniendo el pFM*Trc12* (plásmido control) o pFM*TrcleuO-50* (plásmido que sobreexpresa a la proteína LeuO, ver Tabla 2 en Anexo I) cultivadas en medio rico MA. Los resultados confirman que la presencia de LeuO induce la expresión de *cse1*. Sin embargo, no se observó actividad específica de CAT con la mutación en la caja -10 de este gen, en presencia o ausencia de LeuO, indicando que un solo inicio de transcripción dependiente de este regulador modula la actividad genética de *cse1* en *S. Typhi* IMSS-1 (Figura 9.1).



**Figura 9.1. *cse1* posee un inicio de transcripción dependiente de LeuO.** Mutagénesis sitio dirigida del promotor *cse1*. La secuencia silvestre ATA (subrayada) de la caja -10 fue sustituida por GGG (asteriscos). Las construcciones con la región silvestre (*cse1*) y la sustitución (*cse1* -10 ATA→GGG) fueron evaluadas en las cepas IMSS-1 + pFM*Trc12* e IMSS-1 + pFM*TrcleuO-50* cultivadas en medio MA a OD<sub>595nm</sub>: 1.3. 50  $\mu\text{M}$  de IPTG fueron utilizados para inducir los plásmidos correspondientes, <Id: < límite de detección.

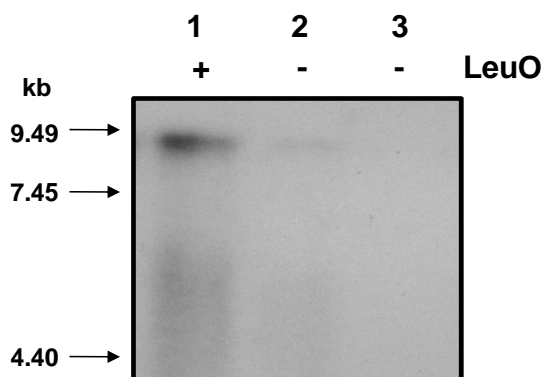
Basado en estos resultados, se definió si el promotor de *cse1* es capaz de controlar la transcripción de los genes que se encuentran corriente abajo: *cse2*, *cas7*, *cas5*, *cas6e*, *cas1* y *cas2*. Para responder lo anterior, se generaron dos fragmentos de DNA: el primero contenía la región reguladora de *cse1* así como la región codificante completa de los genes *cse1* a *cas2* (RR-*cse1-cas2*) y el segundo fragmento carecía del promotor *cse1*, albergando únicamente la región codificante del gen *cse1* hasta *cas2* (*cse1-cas2*). Ambas regiones fueron clonadas en el vector pKK232-9, el cual contiene al gen reportero *cat* y su expresión fue evaluada en *S. Typhi* en presencia o ausencia de LeuO. Los resultados con la fusión RR-*cse1-cas2* mostraron expresión mínima en ausencia de LeuO, mientras en la cepa que sobreexpresa este regulador, la actividad transcripcional incrementó notablemente. El plásmido con la fusión *cse1-cas2*, la cual carece de la región promotora de *cse1*, no mostró actividad de CAT en ausencia o presencia de LeuO (Figura 9.2). Estos datos indican que los genes *cse1-cse2-cas7-cas5-cas6e-cas1-cas2* constituyen un operón dependiente de LeuO.



**Figura 9.2. El clúster *cse1-cse2-cas7-cas5-cas6e-cas1-cas2* forma un operón dependiente de LeuO en *S. Typhi* IMSS-1.** Organización del locus CRISPR-Cas. RR: región reguladora. Bajo el esquema, se muestran los fragmentos de DNA fusionados al reportero *cat*. Las columnas de la derecha indican la actividad específica obtenida para cada construcción en las cepas IMSS-1 + pFMTrc12 e IMSS-1 + pFMTrcleuO-50, cultivadas en medio MA con 50 µM de IPTG a OD<sub>595nm</sub>: 0.6. <ld: < límite de detección.

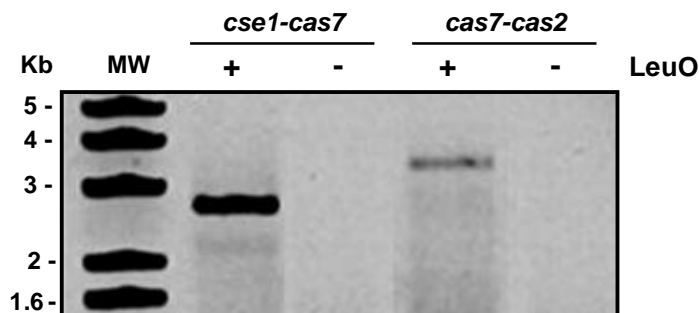
Con el objetivo de establecer si el arreglo CRISPR, el cual se encuentra río abajo de los genes mencionados forma parte del operón regulado por LeuO, se realizaron experimentos de Northern Blot. Se obtuvo RNA total de las cepas IMSS-1 + pFMTrc12 e IMSS-1 + pFMTrcleuO-50 y se transfirió a una membrana de nylon. Utilizando como sonda un fragmento de DNA de las secuencias CRISPR, fue posible detectar una señal de 7-9 Kb con el RNA total de IMSS-1 + pFMTrcleuO-50 y una señal débil del mismo tamaño fue observada en IMSS-1 +

pFM*Trc12* (Figura 9.3). Como control de especificidad de la sonda, se utilizó RNA de la cepa IMSS-1  $\Delta$ *cas*-CRISPR que carece del locus CRISPR-Cas completo, y como se esperaba, no se obtuvo señal alguna. Estos resultados sugieren que en *S. Typhi* IMSS-1 las secuencias CRISPR en conjunto con los genes *cas* constituyen un solo mRNA, cuya actividad es inducida por la proteína reguladora LeuO.



**Figura 9.3. Los genes *cas* junto con el arreglo CRISPR se transcriben en un RNA policistrónico.** Identificación del mRNA *cas*-CRISPR mediante Northern Blot en IMSS-1 conteniendo el pFM*TrcleuO*-50 inducido con 50  $\mu$ M de IPTG (línea 1), IMSS-1 + pFM*Trc12* (línea 2) e IMSS-1  $\Delta$ *cas*-CRISPR (línea 3). La hibridación se realizó con una sonda de la región CRISPR.

Para validar la organización genética del locus CRISPR-Cas observada en los ensayos de Northern Blot, se realizaron RT-PCRs. RNA total de las cepas IMSS-1 + pFM*Trc12* e IMSS-1 + pFM*TrcleuO*-50 inducidos con IPTG, junto con un oligonucleótido específico localizado en las secuencias CRISPR, fueron utilizados para sintetizar el cDNA. Con este templado fue posible amplificar los fragmentos *cse1-cas7* (2.9 Kb) y *cas7-cas2* (3.4 Kb) a partir de las cepa que sobreexpresa LeuO (Figura 9.4). No se obtuvo amplificación de los fragmentos mencionados en ausencia de expresión del regulador, confirmando que el locus *cse1-cse2-cas7-cas5-cas6e-cas1-cas2* junto con las secuencias CRISPR conforman un solo RNA policistrónico dependiente de LeuO, el cual denominamos operón *cas*-CRISPR.



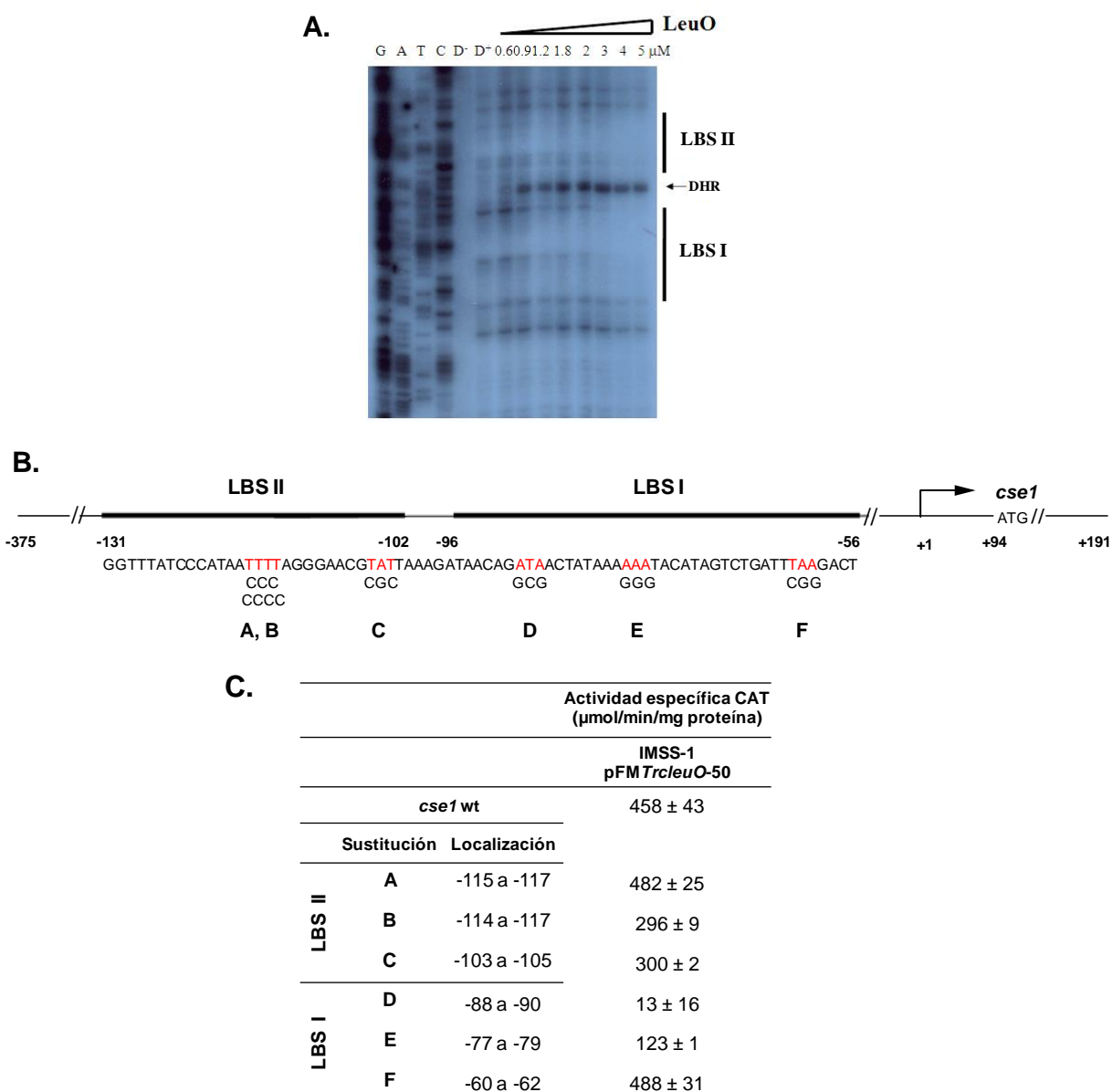
**Figura 9.4.** El locus *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR* constituye una unidad transcripcional dependiente de LeuO. RT-PCR del operón *cas-CRISPR* utilizando RNA total de IMSS-1 + pFM*TrcleuO-50* (+) e IMSS-1 + pFM*Trc12* (-) con 50  $\mu$ M de IPTG. Los fragmentos *cse1-cas7* (2.9 Kb) y *cas7-cas2* (3.4 Kb) fueron amplificados con oligonucleótidos específicos (ver Tabla 3 en Anexo I). MW: marcador de peso molecular.

## 9.2 LeuO interactúa con dos zonas de la región reguladora de *cse1*

Como se mencionó en el apartado anterior, LeuO promueve la expresión del operón *cas-CRISPR*. Por lo tanto, para establecer si LeuO regula de manera directa a este sistema se realizaron ensayos de protección a la enzima DNasa (Footprinting), con la correspondiente proteína purificada y la región reguladora completa de *cse1*. Los resultados demostraron que este factor transcripcional se une a dos zonas de la región intergénica 5' de *cse1* (Figura 9.5A). Los sitios de unión se localizan de la base -56 a -96 y de -102 a -131, respecto al inicio de transcripción de este gen (Figura 9.5B), los cuales fueron denominados LBS I y LBS II, respectivamente por sus siglas en inglés (LeuO binding site).

Debido a que las regiones de interacción identificadas eran de 40 y 30 pb, se determinaron los nucleótidos relevantes para la regulación positiva de *cse1* mediada por LeuO. Considerando que LeuO se une preferentemente a zonas ricas en A+T, seis regiones con alto contenido en dichos nucleótidos fueron sustituidas por G+C, tres de estos sitios se localizaban en LBS I y los tres restantes en LBS II (Figura 9.5B). Los plásmidos conteniendo las sustituciones se evaluaron en *S. Typhi* IMSS-1 sobreexpresando LeuO. Los resultados de la expresión transcripcional mostraron que los cambios realizados en las bases -115 a -117 (A) y -60 a -62 (F) presentan los mismos valores de actividad que la región silvestre. Una disminución de 38% en la expresión de *cse1* fue observada con las

sustituciones realizadas en las bases -114 a -117 (B) y -103 a -105 (C). Mientras un decremento del 76% en la actividad específica fue detectado con los cambios en -77 a -79 (E). Sin embargo, los nucleótidos ubicados en la región -88 a -90 (D) son los de mayor relevancia para la actividad transcripcional mediada por LeuO, ya que la mutación de estas bases disminuye los niveles de expresión de *cse1* en un 97% (Figura 9.5C). La información obtenida indica que tanto LBS I como LBS II son relevantes para la inducción transcripcional de *cse1* dependiente de LeuO.



**Figura 9.5. LeuO regula directamente la expresión de *cse1*.** **A.** Ensayo de protección a la enzima DNasa utilizando la región reguladora de *cse1* y concentraciones crecientes de la proteína purificada LeuO. Las líneas verticales muestran los sitios de unión para este regulador (LBS II, LBS

I), DHR: regiones de hipersensibilidad a la DNAsa por sus siglas en inglés, D<sup>-</sup> y D<sup>+</sup> indican sin y con DNAsa respectivamente, las letras G, A, T, C corresponden a la secuencia nucleotídica de la región reguladora de *cse1*. **B.** Localización de los sitios de interacción de LeuO (LBS II y LBS I). Los números corresponden a las coordenadas respecto al inicio de la transcripción de *cse1*. Se muestran las sustituciones realizadas en LBS II (A, B, C) y LBS I (D, E, F), las secuencias silvestres se presentan en rojo. **C.** Perfil transcripcional de las mutaciones en LBS II y LBS I, evaluadas en la cepa IMSS-1 conteniendo el pFM*Trc*leuO-50 inducido con 50 µM de IPTG. Los ensayos fueron realizados en medio MA a OD<sub>595nm</sub>: 1.3.

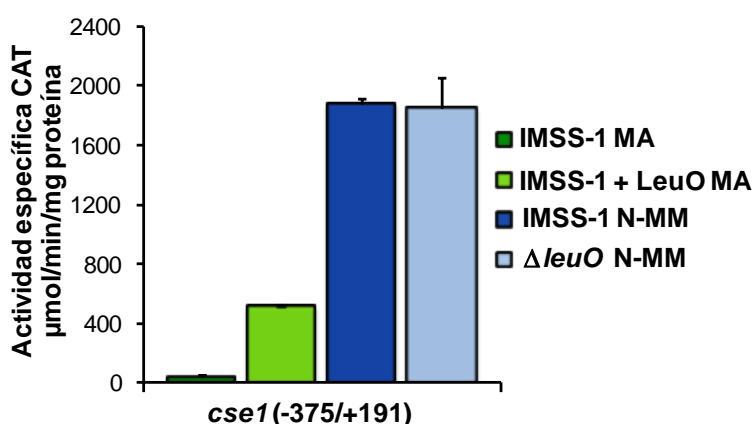
Además del papel en el control de genes involucrados en detoxificación, virulencia y respuesta a estrés (Dillon *et al.*, 2012; Guadarrama *et al.*, 2014; Hernández-Lucas *et al.*, 2008), nuestros resultados demuestran que LeuO está implicado en la regulación directa y positiva del sistema CRISPR-Cas en *Salmonella enterica* serovar Typhi.

### 9.3 El medio mínimo N promueve la expresión del promotor *cse1*

Como se mencionó anteriormente, en medio rico MA fue posible detectar expresión del operón *cas*-CRISPR en *S. Typhi*, mediante la sobreexpresión de LeuO. Esto fue realizado debido a que no se conoce una condición durante la vida libre o patogénesis de esta bacteria, donde se exprese LeuO de manera natural e induzca al locus CRISPR-Cas. Por lo tanto, es necesario identificar condiciones ambientales en las cuales se induzca el sistema CRISPR-Cas. Esta información será de gran utilidad para poder determinar el proceso biológico en el cual pudiera estar participando este clúster genético en *S. Typhi*. En ese sentido, se ha reportado que *cas7*, tercer componente del locus *cse1-cse2-cas7-cas5-cas6-cas1-cas2*-CRISPR, se expresa cuando *S. Typhi* se encuentra dentro de macrófagos de humano (Faucher *et al.*, 2005). Considerando que el medio mínimo N (N-MM) se utiliza para el estudio de la regulación de genes que se expresan en dicha célula eucarionte, tales como los de la isla de patogenicidad 2 de *Salmonella* (SPI-2) (Deiwick *et al.*, 1999), se evaluó la actividad transcripcional de la fusión que contiene la región reguladora de *cse1* [*cse1* (-375/+191)] en esta condición de crecimiento. Los ensayos de actividad transcripcional fueron realizados en la cepa *S. Typhi* silvestre y demostraron que el promotor *cse1*, el cual dirige la transcripción del operón *cas*-CRISPR, se expresa mayoritariamente en medio

mínimo N, comparada con la expresión mínima que presenta en medio rico MA (Figura 9.6).

Para determinar si LeuO es el activador del promotor *cse1* en N-MM, se analizó la actividad de la fusión *cse1* (-375/+191) en la cepa IMSS-1  $\Delta leuO$  y los niveles de expresión observados fueron muy similares a los de IMSS-1 silvestre (Figura 9.6), indicando que en N-MM *cse1* se expresa de manera independiente a LeuO, lo que sugiere la participación de otros elementos genéticos en la regulación positiva del promotor *cse1* en la condición de crecimiento evaluada.



**Figura 9.6. *cse1* se expresa en medio mínimo N de manera independiente a LeuO.** Perfil transcripcional de la fusión *cse1* (-375/+191), conteniendo las bases -375 a +191 respecto al inicio de transcripción de dicho gen, la cual fue evaluada en *S. Typhi* IMSS-1 cultivada en MA (barra verde fuerte), *S. Typhi* IMSS-1 + pFMT*rcleuO*-50 + 50 μM de IPTG en MA (barra verde claro), así como *S. Typhi* IMSS-1 (barra azul fuerte) e IMSS-1  $\Delta leuO$  (barra azul claro) cultivadas en N-MM. Las muestras fueron colectadas a OD<sub>595nm</sub>: 1.3.

#### 9.4 H-NS y Lrp están implicados en la regulación del promotor *cse1* en N-MM

Los datos anteriores sugieren la presencia de factores transcripcionales implicados en la inducción de *cse1* en medio mínimo N. Para identificar reguladores positivos, se evaluó la actividad transcripcional del promotor *cse1* en mutantes individuales de SsrB, PhoP, OmpR, CRP, CpxRA, IHF, NarP, ArcA, PmrAB o FNR, considerando que algunos de ellos como SsrB, PhoP y OmpR están implicados en la regulación de SPI-2, la cual se expresa en N-MM; mientras CRP, CpxRA, IHF, NarP, ArcA, PmrAB y FNR participan en el control de otros genes de virulencia. Los ensayos correspondientes se realizaron en N-MM y en



ninguno de los fondos genéticos analizados se observó un cambio en los valores de actividad, comparados con los de la cepa silvestre de *S. Typhi* (datos no mostrados), indicando que estos reguladores no están implicados en el control positivo de *cse1* en N-MM.

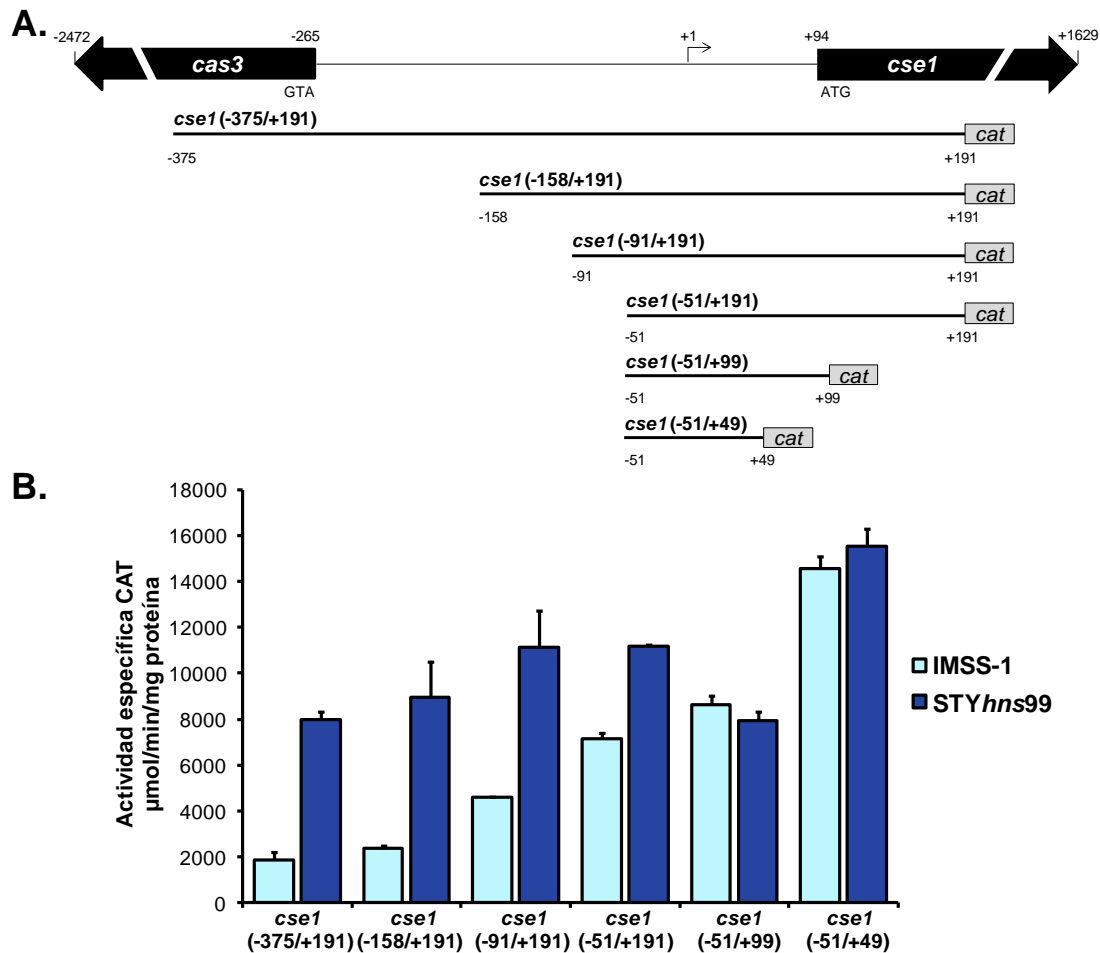
Debido a que no fue posible identificar activadores de la expresión de *cse1*, se llevó a cabo una búsqueda de factores transcripcionales implicados en su regulación negativa. Previos estudios realizados en nuestro laboratorio revelan que los genes regulados por LeuO son comúnmente reprimidos por el regulador global H-NS (De la Cruz *et al.*, 2007; Gallego-Hernandez *et al.*, 2012; Hernández-Lucas *et al.*, 2008). Por dichas razones se evaluó si esta proteína asociada al nucleoide participa en el control negativo de *cse1* en N-MM. Para esto se amplificaron una serie de fragmentos de diferente longitud de la región intergénica 5' de *cse1* y se fusionaron al gen reportero *cat* (Figura 9.7A). Los plásmidos obtenidos se introdujeron en un fondo silvestre de *S. Typhi* IMSS-1 y en la cepa deficiente en *hns* (STY*hns99*). Al evaluar la actividad específica, los valores de expresión de las fusiones *cse1* (-375/+191), *cse1* (-158/+191), *cse1* (-91/+191) y *cse1* (-51/+191) fueron de 7,990, 8,930, 11,150 y 11,160 unidades, respectivamente en el fondo mutante en *hns*, mientras en la cepa IMSS-1 silvestre estas construcciones presentaron 1,860, 2,380, 4,610 y 7,370 unidades respectivamente (Figura 9.7B), revelando que H-NS regula de manera negativa al promotor *cse1* en la condición de crecimiento analizada.

Es relevante hacer notar que los niveles de actividad en la cepa *S. Typhi* IMSS-1 con la construcción más larga *cse1* (-375/+191) fueron menores a los detectados con las fusiones *cse1* (-158/+191), *cse1* (-91/+191) y *cse1* (-51/+191) (Figura 9.7B), demostrando que en estas tres construcciones se eliminaron regiones implicadas en la represión de *cse1*, específicamente las bases -375 a -52, respecto al inicio de transcripción. Más aún, en el fondo STY*hns99* la actividad específica de *cse1* (-375/+191) también presentó un decremento comparado con los valores de expresión obtenidos con las fusiones más cortas *cse1* (-158/+191), *cse1* (-91/+191) y *cse1* (-51/+191) (Figura 9.7B), sugiriendo la participación de otro

---

elemento genético diferente a H-NS en la regulación negativa de *cse1*, el cual modula la expresión de este gen en los nucleótidos -375 a -52.

Con estos experimentos se determinaron zonas de regulación negativa para *cse1* río arriba de su inicio de transcripción. Adicionalmente, regiones de DNA ubicadas corriente abajo del inicio de la transcripción de este gen fueron eliminadas, obteniendo las fusiones *cse1* (-51/+99) y *cse1* (-51/+49). La primer construcción presentó niveles de actividad de 8,640 unidades en la cepa *S. Typhi* IMSS-1 y 7,910 unidades en *STYhns99*. No obstante, el plásmido *cse1* (-51/+49) mostró 14,580 unidades CAT en el fondo silvestre y 15,520 unidades en la cepa deficiente en *hns* (Figura 9.7B), indicando que entre las bases +50 a +99, respecto al inicio de transcripción de *cse1*, también se encuentran elementos de regulación negativa. Todos estos datos sugieren que además de H-NS, existen otros factores silenciando a *cse1* en los nucleótidos -375 a -52 y +50 a +99, es decir, tanto corriente arriba como río abajo de su inicio de transcripción.

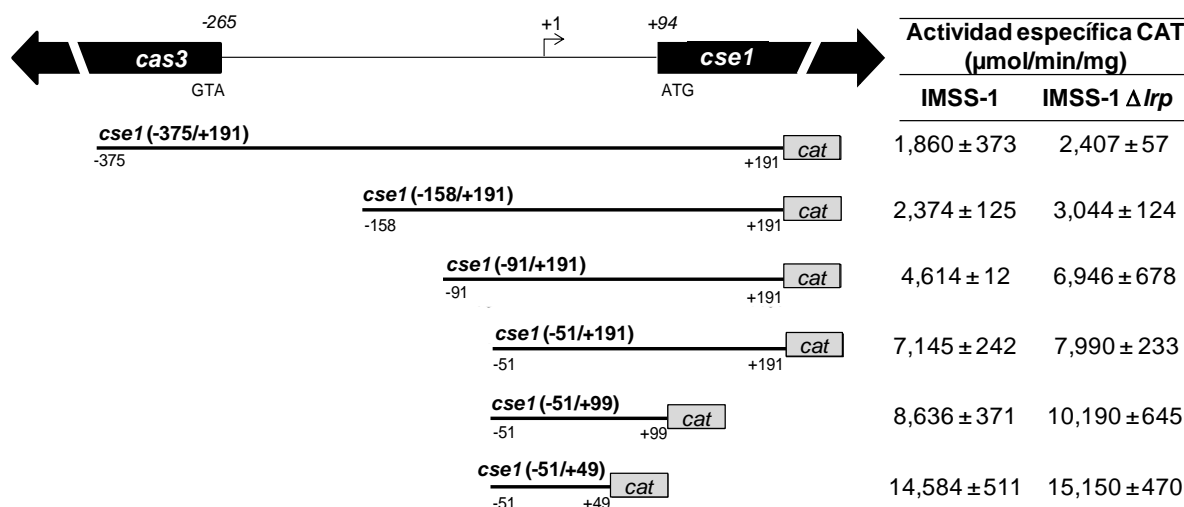


**Figura 9.7. *cse1* es regulado negativamente por H-NS. A.** Representación esquemática de fusiones transcripcionales al reportero *cat* conteniendo fragmentos de diferente longitud de la región reguladora de *cse1*. Los números corresponden a las coordenadas respecto al inicio de transcripción de este gen. **B.** Perfil transcripcional de las construcciones de la región reguladora de *cse1* en los fondos genéticos *S. Typhi* IMSS-1 (azul claro) y STYhns99 (azul rey), cultivadas en N-MM a  $OD_{595nm}$ : 1.3.

Los resultados anteriores sugieren la participación de factores transcripcionales adicionales a H-NS en el control de *cse1*, para identificarlos se realizaron análisis bioinformáticos. Utilizando el algoritmo BLASTp se seleccionaron los 18 ortólogos de *cse1* con el mayor E-value y un “threshold” superior a  $10^{-5}$ . Posteriormente, las regiones reguladoras de cada uno de los ortólogos (400 pb corriente arriba del respectivo inicio de traducción) fueron analizadas con el programa Oligo-Analysis para buscar motivos conservados. Los resultados demostraron la presencia de un motivo TAATAAA, esta secuencia se comparó utilizando el algoritmo TOM TOM con las bases de datos DPINTERACT y REGTRANSDB v4, y se encontró que tiene alta identidad ( $P = 0.0080$ ) con un sitio

de unión para el regulador global Lrp (Leucline response protein) en *E. coli*. De manera interesante, el motivo TAATAAA se localizó 221 pb corriente arriba y 66 pb río abajo del inicio de transcripción de *cse1* en *S. Typhi* IMSS-1. En *Methanospirillum hungatei* están presentes dos sitios de Lrp, uno a 38 y el segundo a 351 pb río arriba del ATG de *cse1*. En el caso de *Granulibacter bethendens* un motivo de Lrp fue localizado 50 pb corriente arriba del ATG de *cse1*. Para *Desulfovibrio desulfuricans*, *Desulfococcus oleovorans* y *Marinomonas MWYL1* se identificó un sitio Lrp a ~400 pb río arriba respecto al inicio de traducción del correspondiente gen *cas*. La presencia de motivos para Lrp en las regiones reguladoras de diversos ortólogos de *cse1* presentes en  $\alpha$ -,  $\delta$ - y  $\gamma$ -proteobacteria, así como en arqueas, sugiere la participación de la proteína Lrp en la regulación transcripcional del sistema CRISPR-Cas.

Para evaluar si Lrp realmente participa en el control de *cse1*, se construyó una cepa de *S. Typhi* IMSS-1 deficiente en este regulador (*IMSS-1  $\Delta$ lrp*). Posteriormente, las fusiones conteniendo diferentes fragmentos de la región reguladora de *cse1*, fueron transformadas en dicho fondo genético y se realizaron los ensayos de CAT. Los resultados mostraron un incremento en los niveles de expresión de las fusiones *cse1* (-375/+191), *cse1* (-158/+191), *cse1* (-91/+191), *cse1* (-51/+191) y *cse1* (-51/+99) (2,410, 3,040, 6,950, 7,990 y 10,190 unidades, respectivamente) en la cepa *IMSS-1  $\Delta$ lrp* comparados con su actividad en el fondo silvestre (1,860, 2,380, 4,610, 7,150 y 8,640 unidades, respectivamente) (Figura 9.8), por lo tanto Lrp es otro regulador negativo de *cse1*.



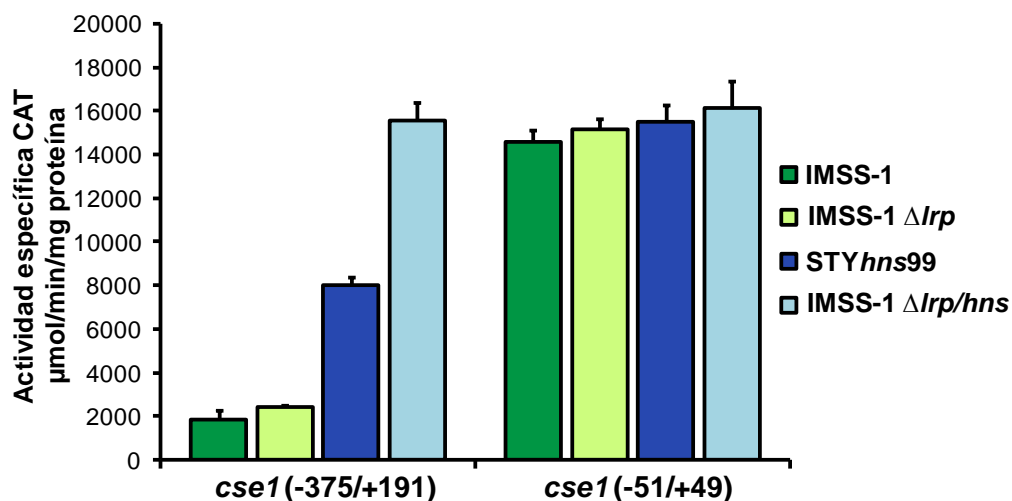
**Figura 9.8.** Lrp reprime la expresión de *cse1*. En el esquema se presentan fusiones transcripcionales al reportero *cat* de diferente longitud de la región reguladora de *cse1*. Los números corresponden a las coordenadas respecto al inicio de transcripción de *cse1*. Las columnas de la derecha muestran el perfil transcripcional de las construcciones en los fondos genéticos *S. Typhi* IMSS-1 y *S. Typhi* IMSS-1 Δ*lrp* cultivadas en N-MM a OD<sub>595nm</sub>: 1.3.

Para confirmar el papel tanto de H-NS como de Lrp en la represión del sistema CRISPR-Cas, se generó una cepa de *S. Typhi* doble mutante *lrp/hns* y se evaluó la fusión que contiene la región intergénica 5' completa de *cse1* [*cse1* (-375/+191)]. Las unidades CAT del plásmido *cse1* (-375/+191) fueron de 15,560 en el fondo genético Δ*lrp/hns*, mientras en las cepas IMSS-1 silvestre, IMSS-1 Δ*lrp* y STY*hns99* se observaron valores de 1,860, 2,410 y 7,990 unidades CAT, respectivamente (Figura 9.9), validando la participación de ambas proteínas en la regulación negativa del promotor *cse1*.

Es relevante mencionar que la actividad de la región reguladora completa de *cse1* en los fondos genéticos IMSS-1 Δ*lrp* y STY*hns99* incrementa 30 y 330% respectivamente, comparada con la expresión en la cepa silvestre, sugiriendo que H-NS tiene un papel principal respecto a Lrp en el silenciamiento del promotor *cse1*. Razón por la cual en ausencia de *lrp*, HNS es capaz de reprimir al 95% la expresión de dicho promotor. Este fenotipo también se puede atribuir a que H-NS regule negativamente a *lrp*. Por lo tanto, solo en la cepa deficiente de *hns* se induciría la actividad necesaria de *lrp* para reprimir la expresión de *cse1*, observándose claramente su papel como regulador negativo en esta condición.

Debido a lo anterior se obtienen 7,990 unidades CAT en la cepa *STYhns99*, mientras que en ausencia de ambos reguladores (*IMSS-1 Δlrp/hns*) la expresión de *cse1* corresponde a 15,560 unidades CAT. Sin embargo, dicha hipótesis requiere ser validada.

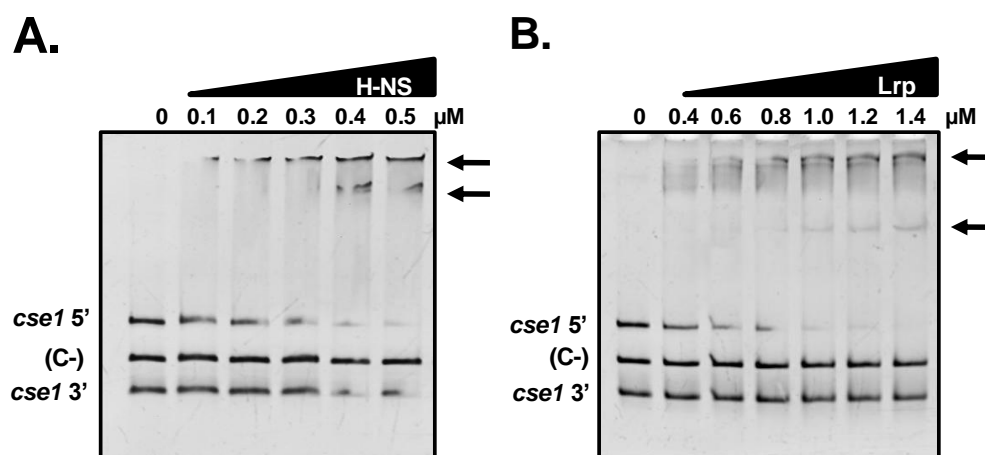
En este trabajo también se determinó que los nucleótidos de -51 a +49 contienen el promotor completo de *cse1* sin ningún elemento de regulación negativa, ya que al evaluar la construcción *cse1* (-51/+49) en los fondos *S. Typhi* silvestre, *Δlrp*, *STYhns99* y *Δlrp/hns* se obtuvieron valores de actividad de 14,580, 15,150, 15,520 y 16,120 unidades, respectivamente (Figura 9.9). Estos datos también confirman que la región reguladora de *cse1* posee dos zonas de regulación negativa, localizadas en los nucleótidos -375 a -52 y de +50 a +191, respecto a su inicio de transcripción.



**Figura 9.9. Lrp y H-NS son represores de *cse1*.** Perfil transcripcional de las fusiones *cse1* (-375/+191) y *cse1* (-51/+49), evaluadas los fondos genéticos *S. Typhi* IMSS-1 (barras verde fuerte), *S. Typhi* IMSS-1  $\Delta$ lrp (barras verde claro), *STYhns99* (barras azul fuerte) y *S. Typhi* IMSS-1  $\Delta$ lrp/hns (barras azul claro) cultivadas en N-MM a OD<sub>595nm</sub>: 1.3.

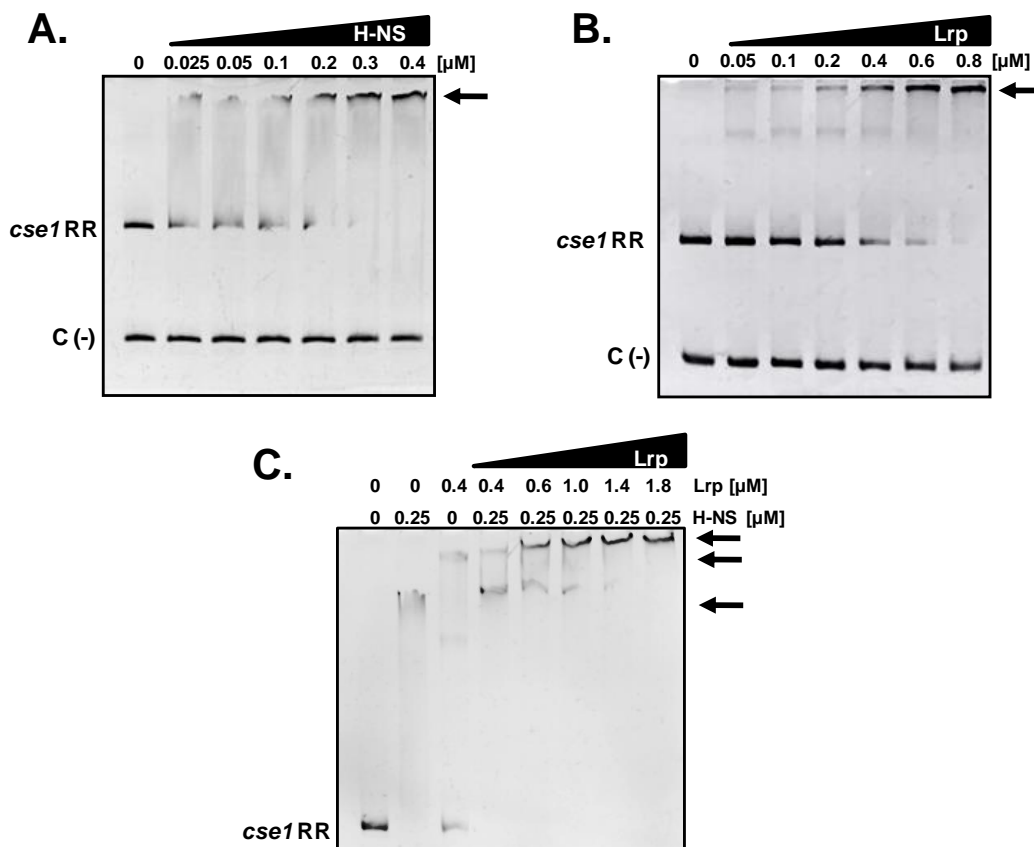
Considerando que se identificaron secuencias corriente arriba y corriente abajo del +1 de transcripción de *cse1*, implicadas en el control negativo de su expresión y, que los factores transcripcionales H-NS y Lrp participan en su represión, se examinó si estos reguladores son capaces de interactuar con las regiones de DNA identificadas. Para resolver este objetivo se realizaron experimentos de interacción DNA-proteína (EMSA) con las correspondientes

proteínas purificadas y diferentes fragmentos de la región reguladora de *cse1*. Los resultados mostraron que tanto H-NS como Lrp interactúan con dos fragmentos: el primero denominado *cse1* 5', se encuentra corriente arriba del inicio de la transcripción de *cse1* (nucleótidos -375 a -52) y el segundo corresponde a *cse1* 3', localizado corriente abajo del +1 de este gen (nucleótidos +50 a +191) (Figura 9.10). Por lo tanto, H-NS y Lrp reprimen directamente la expresión de *cse1*, uniéndose tanto río arriba como río abajo del inicio de la transcripción de este gen.



**Figura 9.10. H-NS y Lrp reprimen directamente la expresión de *cse1*.** EMSAs con las proteínas purificadas H-NS (A), Lrp (B) y los fragmentos *cse1* 5' (nucleótidos -375 a -52) y *cse1* 3' (nucleótidos +50 a +191). Una región de 242 pb del gen 16S rRNA de *S. Typhi* fue utilizada como control negativo (C-) en ambos experimentos. Las flechas indican los complejos DNA-proteína, los cuales se visualizaron mediante el uso de bromuro de etidio.

Con la finalidad de elucidar si H-NS y Lrp son capaces de interactuar simultáneamente con la región reguladora completa de *cse1* (*cse1* RR), se realizaron experimentos de EMSA. En primer lugar verificamos que la interacción individual de H-NS o Lrp con el fragmento *cse1* RR fuera específica y como se muestra en la figura 9.11A y 9.11B, cada uno de los reguladores interactúa con dicha región de DNA y no con el fragmento de 16S rRNA (C-). Posteriormente, utilizando ambas proteínas se observa que el complejo H-NS-*cse1* RR se modifica en presencia de concentraciones crecientes de Lrp, visualizándose una nueva banda, la cual tiene un peso molecular mayor al de los complejos obtenidos con H-NS y Lrp de manera independiente (Figura 9.11C). Estos experimentos demuestran que ambas proteínas se unen de manera simultánea a la región reguladora de *cse1* para reprimir la expresión del sistema CRISPR-Cas en N-MM.

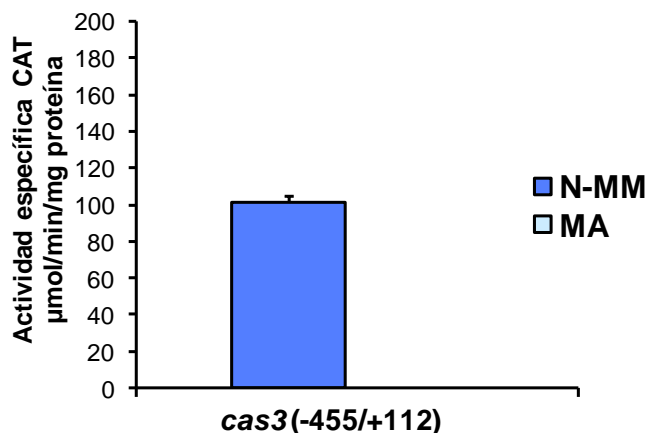


**Figura 9.11. H-NS y Lrp interactúan simultáneamente con *cse1*.** EMSAs con las proteínas purificadas H-NS (A), Lrp (B) y el fragmento de la región reguladora completa de *cse1* (*cse1* RR, nt -455 a +191, respecto al inicio de transcripción). Una región de 242 pb del gen 16S rRNA de *S. Typhi* fue utilizada como control negativo (C-) en ambos experimentos. C. EMSA del fragmento *cse1* RR incubado con 0.25  $\mu$ M de H-NS y concentraciones crecientes de Lrp. Las flechas indican los complejos DNA-proteína, los cuales se visualizaron mediante el uso de bromuro de etidio.

### 9.5 La nucleasa Cas3 es inducida en medio mínimo N y reprimida por H-NS

La organización genética del clúster CRISPR-Cas de *S. Typhi* demuestra que *cse1* se encuentra divergente al gen *cas3*, por lo tanto, comparten la misma región reguladora. Por esta razón, se evaluó si dicho gen *cas* presenta expresión en N-MM al igual que *cse1*. Una fusión de la región reguladora completa de *cas3* fue construida [*cas3* (-455/+112)] y transformada en la cepa silvestre de *S. Typhi*. Los ensayos correspondientes mostraron que *cas3* presenta expresión de 100 unidades CAT en medio mínimo N y su actividad transcripcional fue nula en MA (Figura 9.12).



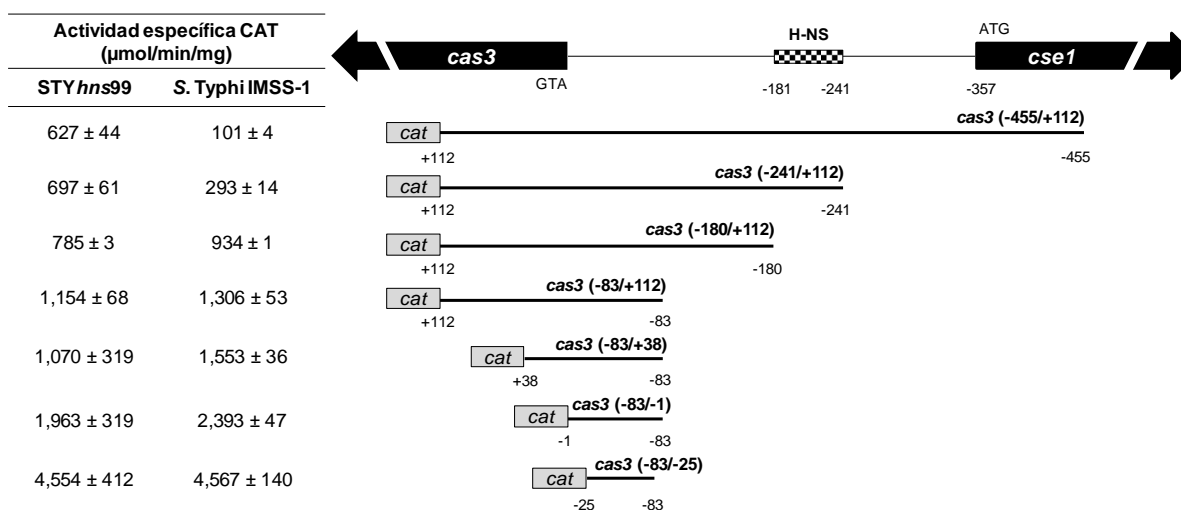


**Figura 9.12. *cas3* se expresa en N-MM.** Perfil transcripcional de la fusión *cas3* (-455/+112) evaluada en la cepa *S. Typhi* IMSS-1, cultivada en N-MM (barra azul fuerte) y MA (barra azul claro) a OD<sub>595nm</sub>: 1.3.

La modesta expresión que presenta el promotor *cas3* sugiere que está reprimido en N-MM. Para identificar zonas implicadas en su regulación negativa, se diseñaron fusiones transcripcionales al gen reportero *cat* de diferentes longitudes de la región intergénica 5' de *cas3* y se analizaron en *S. Typhi* IMSS-1 cultivada en N-MM. Las construcciones *cas3* (-241/+112), *cas3* (-180/+112) y *cas3* (-83/+112) presentaron 290, 930 y 1,306 unidades, respectivamente (Figura 9.13). Estos valores son mayores a los obtenidos con la fusión *cas3* (-455/+112) (101 unidades), indicando que las bases -455 a -84, respecto al ATG de *cas3*, están implicadas en la regulación negativa de dicho promotor. Adicionalmente, las fusiones *cas3* (-83/+38), *cas3* (-83/-1) y *cas3* (-83/-25), en las cuales se eliminaron regiones ubicadas hacia el extremo 3' de la zona reguladora completa de *cas3*, presentaron 1,550, 2,390 y 4,570 unidades CAT, respectivamente (Figura 9.13), mostrando que los nucleótidos ubicados de -24 a +112 respecto al inicio de traducción de *cas3*, también constituye una zona de regulación negativa para este gen. Por lo tanto, al igual que su gen divergente *cse1*, la expresión de *cas3* es silenciada en secuencias ubicadas tanto corriente arriba como corriente abajo de su ATG.

Los experimentos anteriores demostraron que la actividad transcripcional de *cas3* se encuentra reprimida en N-MM y considerando que H-NS y Lrp son represores del promotor *cse1*, se analizó si estas proteínas son las responsables

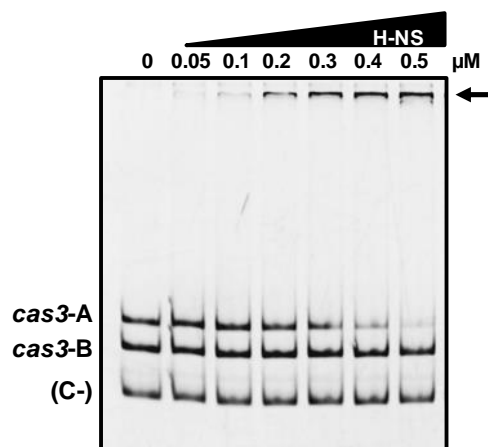
de la regulación negativa de *cas3*. Ensayos de actividad específica CAT fueron realizados utilizando las cepas IMSS-1  $\Delta lrp$  y STY*hns99*, así como las diversas construcciones de la región promotora del gen de interés. Con estos experimentos fue posible definir que solamente H-NS participa en el silenciamiento de *cas3*, ya que los valores de expresión de las fusiones *cas3* (-455/+112) y *cas3* (-241/+112) incrementaron en la cepa deficiente en *hns* (630 y 690 unidades CAT, respectivamente), comparados con los niveles observados en el fondo silvestre (101 y 290 unidades CAT) (Figura 9.13). De manera interesante, al evaluar la construcción *cas3* (-180/+112), la actividad transcripcional fue similar en la *S. Typhi* IMSS-1 y la cepa mutante STY*hns99* (Figura 9.13), indicando que H-NS reprime a *cas3* mediante los nucleótidos -241 a -181, respecto al ATG de dicho gen.



**Figura 9.13. La expresión de *cas3* es reprimida en medio mínimo N por H-NS.** El gen *cas3* se encuentra divergente a *cse1* y comparten una región intergénica 5' de 356 pb. El rectángulo con cuadrícula indica el sitio de represión por H-NS (nt -241 a -181). Bajo el esquema se presentan fusiones transcripcionales al reportero *cat* de diferente longitud de la región reguladora de *cas3*. Los números representan las coordenadas respecto al inicio de traducción. Las columnas de la izquierda muestran la actividad específica CAT de las construcciones en los fondos genéticos *S. Typhi* IMSS-1 y STY*hns99*, cultivadas en N-MM a  $OD_{595nm}$ : 1.3.

Para demostrar si las bases ubicadas de -241 a -181 contienen un sitio de unión para H-NS, se llevaron a cabo experimentos de EMSA con la correspondiente proteína purificada y los fragmentos *cas3*-A (nucleótidos -241 a +112) y *cas3*-B (nucleótidos -180 a +112). Los datos demuestran que H-NS se une al fragmento *cas3*-A y que no interactúa con *cas3*-B (Figura 9.14). Por lo tanto,

este regulador global silencia directamente la expresión de *cas3* a través de las bases -241 a -181.



**Figura 9.14. H-NS interactúa con la región reguladora de *cas3*.** EMSA con la proteína purificada H-NS y los fragmentos *cas3*-A (nucleótidos -241 a +112) y *cas3*-B (nucleótidos -180 a +112). Una región de 242 pb del gen 16S rRNA de *S. Typhi*, fue utilizada como control negativo (C-). La flecha indica el complejo DNA-proteína, el cual fue visualizado mediante el uso de bromuro de etidio.

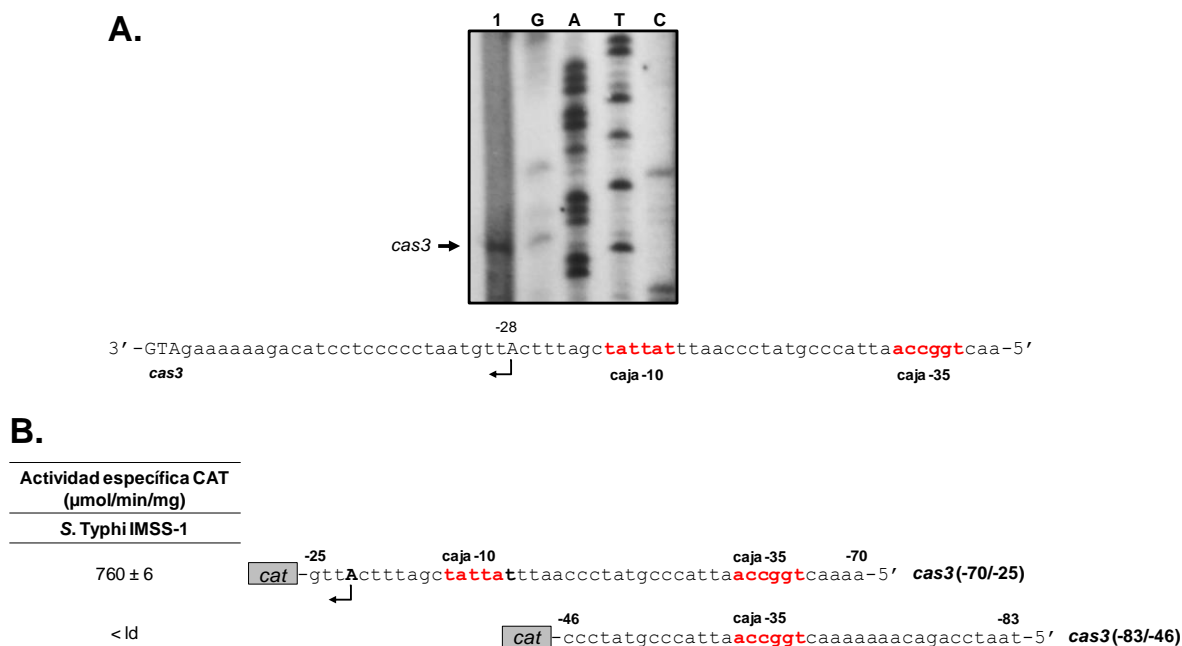
Es relevante mencionar que los datos sobre la regulación negativa de *cas3* sugieren que existe, al menos, otro factor transcripcional distinto de H-NS, implicado en su represión. Esto se debe a que al evaluar la región reguladora completa de *cas3* [*cas3* (-455/+112)], en el fondo mutante STY*hns99*, se observaron valores de 630 unidades CAT. Mientras que la fusión *cas3* (-83/-25), la cual carece de las zonas de control negativo ubicadas corriente arriba y corriente abajo del ATG de *cas3*, presenta 4,550 unidades CAT en el fondo IMSS-1 silvestre (Figura 9.13). Considerando que H-NS interactúa con los nucleótidos -241 a -181 respecto al inicio de traducción de *cas3*, y que dos zonas de regulación negativa fueron identificadas para dicho gen, las cuales se ubican en las bases -455 a -84 y de -24 a +112, es evidente que elementos de regulación adicionales reprimen la expresión de *cas3* a través de esta última región y mediante los nucleótidos -180 a -84. Con el fin de identificar factores transcripcionales implicados en el silenciamiento de *cas3*, las fusiones *cas3* (-455/+112), *cas3* (-241/+112) y *cas3* (-83/+112) fueron evaluadas en las cepas de *S. Typhi* IMSS-1  $\Delta$ *ssrB*,  $\Delta$ *phoP*,  $\Delta$ *ompR*,  $\Delta$ *crp*,  $\Delta$ *cpxRA*,  $\Delta$ *ihf*,  $\Delta$ *narP*,  $\Delta$ *arcA*,  $\Delta$ *pmrAB* y  $\Delta$ *fnr*. No obstante, los valores de actividad transcripcional observados en las diferentes mutantes fueron similares a los obtenidos en la cepa silvestre (datos no mostrados).

En resumen, *cas3* conforma una unidad transcripcional independiente en medio mínimo N, cuya expresión es reprimida en regiones ubicadas tanto corriente arriba como río abajo de su inicio de traducción. H-NS participa en el silenciamiento de este gen, interactuando con los nucleótidos -241 a -181. Sin embargo, existe otro regulador negativo controlando la expresión de *cas3* en N-MM mediante las bases -180 a -84 y -24 a +112, respecto a su ATG.

### **9.6 Elementos genéticos positivos implicados en la expresión transcripcional de *cas3***

La información anterior muestra la caracterización transcripcional de la regulación negativa de *cas3*. En este estudio también se analizó la presencia de elementos involucrados en su regulación positiva. En primer lugar se localizó el inicio de la transcripción mediante experimentos de “primer extension”, en los cuales se utilizó RNA de *S. Typhi* IMSS-1 conteniendo el plásmido *cas3* (-83/-25) y cultivada en N-MM. Esta construcción fue seleccionada debido a que presenta el máximo valor de actividad detectado para *cas3*, ya que se encuentra desprovista de regulación negativa. Los resultados demostraron que el inicio de transcripción de *cas3* se ubica a -28 pb de su ATG, con secuencias consenso -10 y -35 para el factor  $\sigma^{70}$  (Figura 9.15A).

Con el objetivo de validar la región promotora de *cas3*, se construyeron los plásmidos *cas3* (-70/-25) y *cas3* (-83/-46). El primero contiene el inicio de la transcripción, así como las cajas -10 y -35; mientras la segunda construcción carece del +1 de la transcripción así como de la caja -10. Ambos plásmidos fueron evaluados en la cepa silvestre de *S. Typhi* cultivada en N-MM. Solamente se detectó actividad específica con la fusión *cas3* (-70/-25) (Figura 9.15B), confirmando que el promotor de *cas3* se encuentra en los nucleótidos -70 a -25, respecto a su inicio de traducción.

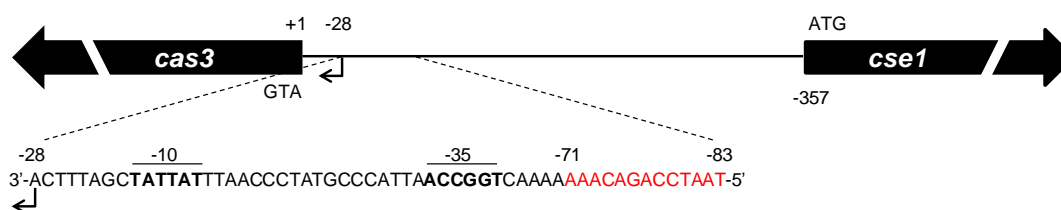


**Fig 9.15. El promotor de *cas3* se localiza en los nucleótidos -70 a -25 respecto a su ATG. A.** “Primer extension” realizado con RNA de la cepa *S. Typhi* IMSS-1, conteniendo el plásmido *cas3* (-83/-25) y cultivada en N-MM a  $OD_{595\text{nm}}$ : 1.3. Línea 1: reacción de síntesis de cDNA realizada con un oligonucleótido del vector pKK232-8 (pKK-R). Las letras G, A, T, C corresponden a la secuencia nucleotídica de la región reguladora de *cas3*. La flecha indica el inicio de la transcripción cuya localización se muestra en la secuencia bajo la imagen, también se ilustran las cajas -10 y -35 (en rojo) para el factor  $\sigma^{70}$ . **B.** Fusiones transcripcionales al reportero *cat* diseñadas para validar el promotor de *cas3*. La flecha indica el inicio de la transcripción de *cas3*, las cajas -10 y -35 se encuentran en rojo. La columna de la izquierda muestra la actividad específica de las construcciones en la cepa *S. Typhi* IMSS-1, cultivada en N-MM a  $OD_{595\text{nm}}$ : 1.3. < Id: < límite de detección.

A pesar de que la construcción *cas3* (-70/-25) contiene todos los elementos necesarios para la expresión del promotor *cas3*, las unidades CAT observadas (760) fueron seis veces menores que la actividad específica presentada por la fusión *cas3* (-83/-25) (4,570 unidades CAT) (Figura 9.13 y 9.15B), indicando que las bases -83 a -71 (5'-TAATCCAGACAAA-3') constituyen una región de regulación positiva para la expresión de *cas3* en N-MM. Los 13 nucleótidos mencionados se encuentran adyacentes a la caja -35 del promotor *cas3*, ubicación característica del elemento UP, el cual corresponde a sitios de unión para la subunidad  $\alpha$  de la RNA polimerasa. Por esta razón, se llevaron a cabo EMSAs con fragmentos de DNA, albergando la región de regulación positiva identificada y la subunidad  $\alpha$  purificada. Sin embargo, no se observó interacción entre estos elementos, sugiriendo que las bases 5'-TAATCCAGACAAA-3' corresponden a

motivos de reconocimiento para un activador. Considerando la participación de LeuO en la regulación positiva del promotor *cse1*, se analizó el papel de este regulador en el control de *cas3*, encontrando que la sobreexpresión de LeuO no induce su actividad transcripcional (datos no mostrados).

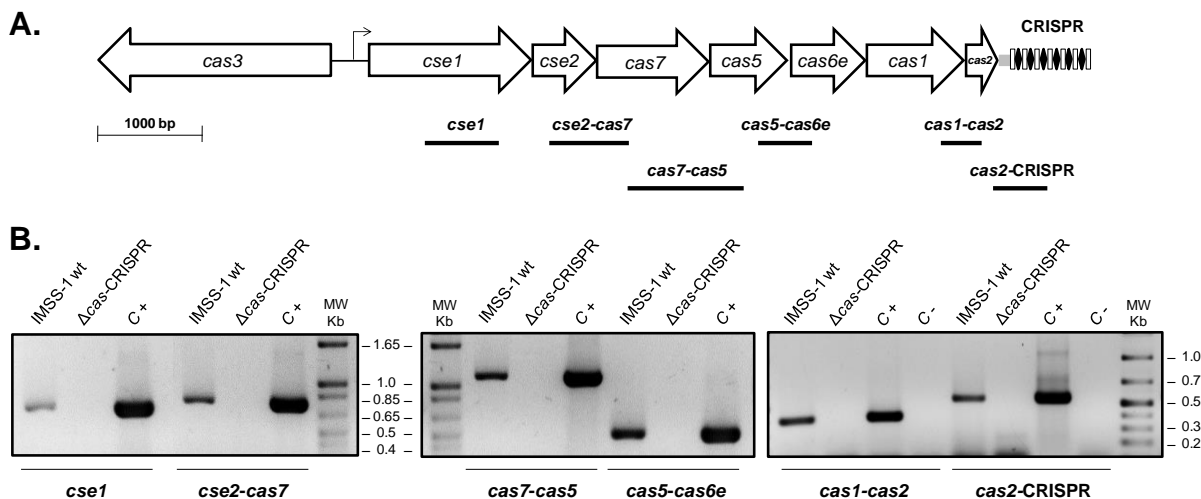
Con los experimentos realizados fue posible definir el promotor del gen que codifica para la nucleasa Cas3 en *S. Typhi*: el inicio de transcripción se localiza en la base -28 respecto al ATG y posee cajas -10 y -35 para el factor  $\sigma^{70}$ . Adicionalmente, los nucleótidos -71 a -83 están involucrados en su regulación positiva (Figura 9.16).



**Figura 9.16. Elementos genéticos necesarios para la expresión de *cas3* en *S. Typhi*.** *cas3* se encuentra divergente a *cse1* y posee un inicio de transcripción (flecha) a 28 pb corriente arriba de su ATG, con secuencias consenso -10 y -35 (en negritas) para el factor  $\sigma^{70}$ . 13 nucleótidos ubicados de -71 a -83 (en rojo), respecto al inicio de traducción, son necesarios para la expresión de *cas3* en N-MM.

## 9.7 La cadena sentido del locus CRISPR-Cas codifica el operón *cse1-cse2-cas7-cas5-cas6-cas1-cas2*-CRISPR y el RNA *scse2*

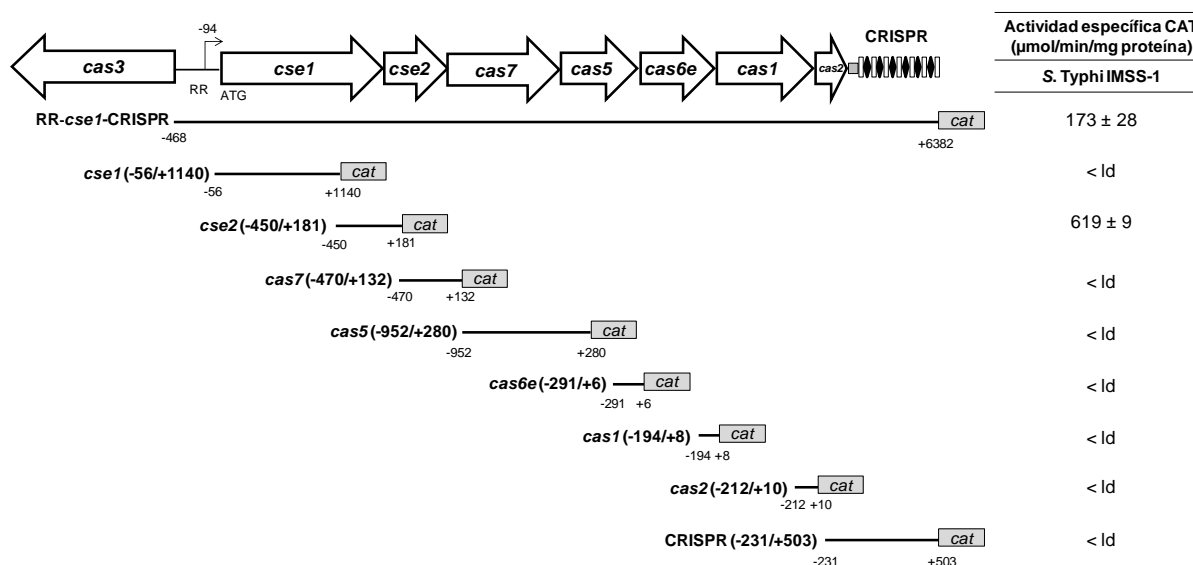
Los datos presentados anteriormente indican que los promotores de los genes divergentes *cas3* y *cse1* se inducen en medio mínimo N. Con el objetivo de evaluar si el promotor de *cse1* controla la transcripción de los genes *cas* y el arreglo CRISPR en esta condición de expresión, se realizaron experimentos de RT-PCR. Utilizando RNA total aislado de la cepa silvestre de *S. Typhi*, se sintetizó un cDNA con un oligonucleótido ubicado en las secuencias CRISPR (extremo 3' del locus CRISPR-Cas). Este cDNA y oligonucleótidos específicos fueron utilizados para amplificar, mediante PCR, los fragmentos *cse1*, *cse2-cas7*, *cas7-cas5*, *cas5-cas6*, *cas1-cas2* y *cas2*-CRISPR (Figura 9.17), demostrando que los genes *cas* junto con el arreglo CRISPR comprenden un operón en medio mínimo N.



**Figura 9.17. El locus *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR* forma un operón en medio mínimo N. A.** Organización del locus CRISPR-Cas presente en *S. Typhi*. Bajo el esquema genético se muestran los fragmentos de DNA amplificados en los experimentos de RT-PCR. **B.** RT-PCR utilizando RNA total de la cepa *S. Typhi* IMSS-1, cultivada en N-MM (OD<sub>595nm</sub>: 1.3). El cDNA fue sintetizado con un oligonucleótido de las secuencias CRISPR. Los fragmentos *cse1* (704 pb), *cse2-cas7* (756 pb), *cas7-cas5* (1104 pb), *cas5-cas6e* (511 pb), *cas1-cas2* (389 pb) y *cas2-CRISPR* (512 pb) fueron amplificados con primers específicos. RNA total de la cepa *S. Typhi*  $\Delta cas$ -CRISPR fue utilizado como control de especificidad. C+: Control positivo, fragmentos amplificados con DNA genómico como templado. C-: Control negativo, reacción de síntesis de cDNA utilizando RNA de *S. Typhi* IMSS-1 sin la enzima transcriptasa reversa. MW: marcador de peso molecular.

Para validar los datos de RT-PCR, se elaboró una fusión que contenía el promotor *cse1*, los genes *cse1-cse2-cas7-cas5-cas6e-cas1-cas2* y las secuencias CRISPR (RR-*cse1*-CRISPR), así como construcciones independientes conteniendo cada una de las regiones intergénicas 5' de los diferentes genes *cas* y del arreglo CRISPR. Todas las fusiones fueron transformadas en *S. Typhi* IMSS-1. Los perfiles de actividad específica realizados en N-MM demostraron que la construcción RR-*cse1*-CRISPR presenta actividad transcripcional (Figura 9.18), indicando que el promotor *cse1* es capaz de transcribir a los genes *cas* que se encuentran corriente abajo (*cse2-cas7-cas5-cas6e-cas1-cas2*) junto con el locus CRISPR. Más aun, la expresión de las fusiones independientes con las regiones intergénicas 5' de *cas7*, *cas5*, *cas6e*, *cas1*, *cas2* y CRISPR fue nula (Figura 9.18), sugiriendo que no existen promotores presentes en estas secuencias de DNA. Sin embargo, con la construcción que contiene 450 pb corriente arriba y 181 pb río abajo del ATG de *cse2* [*cse2* (-450/+181)], se detectó actividad transcripcional

(620 unidades CAT, Figura 9.18), demostrando que existe una región promotora dentro de este fragmento.

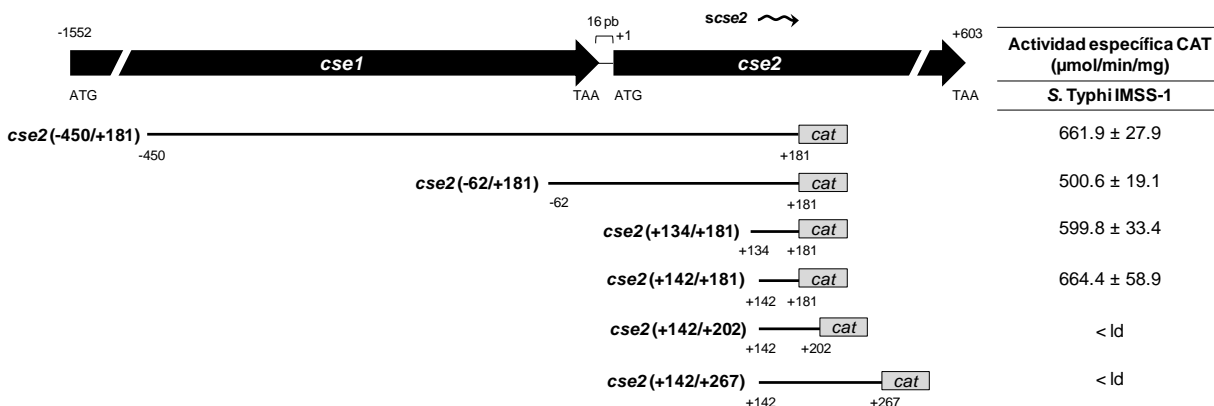


**Figura 9.18. El promotor *cse1* controla la transcripción del locus *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR* en medio mínimo N.** Organización del sistema CRISPR-Cas en *S. Typhi*, RR: región reguladora de *cse1* (356 pb), la flecha indica el inicio de la transcripción. Bajo el esquema se muestran fusiones transcripcionales al reportero *cat*, las cuales fueron nombradas de acuerdo al gen *cas* evaluado. Los números en cada construcción representan las coordenadas respecto al ATG del gen correspondiente. En el caso de RR-*cse1*-CRISPR, las coordenadas se muestran respecto al ATG de *cse1*. Los números en la fusión CRISPR se presentan de acuerdo al primer nucleótido de la secuencia líder de este arreglo. La columna de la derecha corresponde a la actividad específica CAT de las construcciones, evaluadas en la cepa *S. Typhi* IMSS-1 cultivada en N-MM a OD<sub>595nm</sub>: 1.3. <Id: < límite de detección.

Con la finalidad de determinar el tamaño de la unidad transcripcional identificada en *cse2*, se construyeron fusiones al reportero *cat* de diferente longitud, cuya expresión fue evaluada en N-MM. Las construcciones cortas *cse2* (-62/+181), *cse2* (+134/+181) y *cse2* (+142/+181) presentaron niveles de expresión similares a la fusión larga *cse2* (-450/+181) (Figura 9.19), indicando que existe un promotor dentro de la región codificante de *cse2*, específicamente entre las bases +142 a +181 respecto al ATG de este gen. Notablemente, con los plásmidos *cse2* (+142/+202) y *cse2* (+142/+267) no se obtuvo actividad transcripcional (Figura 9.19), sugiriendo la presencia de un terminador río abajo del nucleótido +181. Basado en este resultado, el promotor identificado en la región intragénica de *cse2* no controla la expresión de los elementos genéticos localizados corriente abajo de este gen: *cas7*, *cas5*, *cas6e*, *cas1*, *cas2* y CRISPR. Por lo tanto, el promotor



presente en las bases +142 a +181 dirige la expresión de un posible RNA pequeño, razón por la cual esta unidad transcripcional fue nombrada *scse2* (sense *cse2* RNA).



**Figura 9.19. Identificación de la unidad transcripcional *scse2*.** El gen *cse2* se encuentra río abajo de *cse1*. La flecha ondulada representa al RNA *scse2*. Bajo el diagrama se presentan fusiones transcripcionales al reportero *cat* de diferente longitud. Los números corresponden a las coordenadas respecto al ATG de *cse2*. La columna de la derecha muestra la actividad específica CAT, obtenida al evaluar las construcciones en *S. Typhi* IMSS-1 cultivada en N-MM a OD<sub>595</sub>: 1.3. <ld: < límite de detección.

Mediante análisis bioinformáticos utilizando MEME-MAST (Bailey *et al.*, 2009) se analizó la región +142 a +181 respecto al ATG de *cse2* y se identificaron secuencias consenso -10 (CGCCATATG) y -35 (GTTGAA) para el factor  $\sigma^{32}$  de *E. coli*, con un e-value de  $10^{-6}$ . Se diseñó una fusión transcripcional que carecía de la caja -35 identificada [*cse2* (+154/+181)] y al analizarla en *S. Typhi* IMSS-1, cultivada en N-MM, se observó una expresión ocho veces menor comparada con las unidades CAT de *cse2* (+142/+181) (Figura 9.20), validando que en los nucleótidos +142 a +181 están presentes los elementos necesarios para la expresión del promotor *scse2*.



**Figura 9.20. El promotor *scse2* está presente en los nucleótidos +142 a +181.** Secuencia de nucleótidos contenida en las fusiones *cse2* (+142/+181) y *cse2* (+154/+181). Se muestran las cajas -10 y -35 para el factor  $\sigma^{32}$ . Los números bajo las secuencias corresponden a las coordenadas

respecto al ATG de *cse2*. La columna de la derecha muestra la actividad específica CAT obtenida al evaluar las construcciones en *S. Typhi* IMSS-1 cultivada en N-MM a OD<sub>595nm</sub>: 1.3.

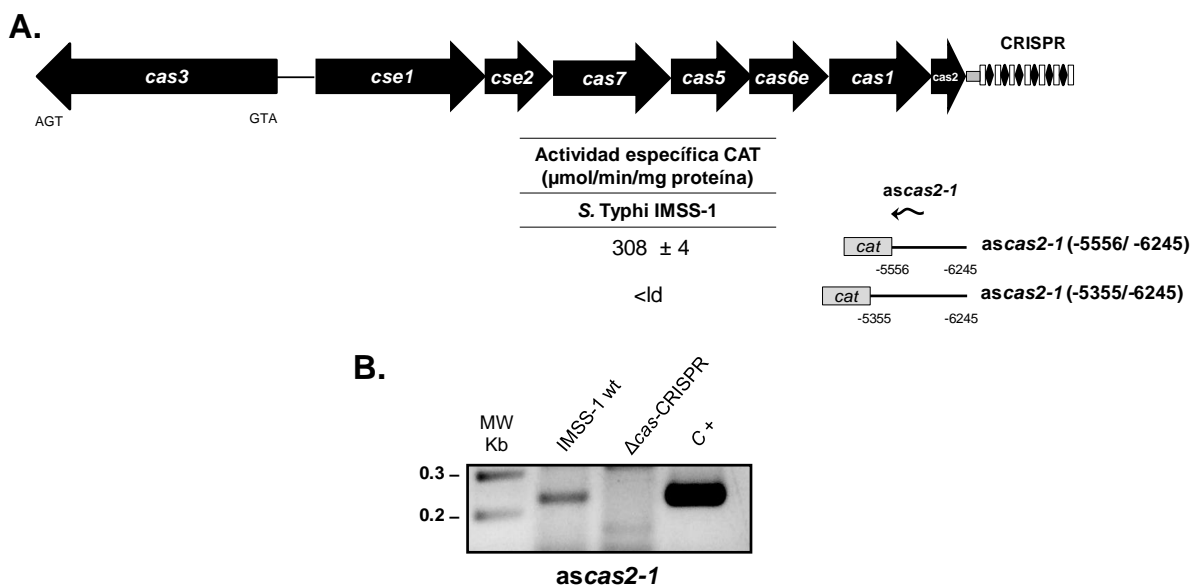
Por lo tanto, en la cadena sentido del locus CRISPR-Cas de *S. Typhi*, un promotor localizado en la región codificante de *cse2* dirige la transcripción del RNA *scse2* y el promotor *cse1* controla la expresión del RNA policistrónico *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR* (operón *cas-CRISPR*).

### 9.8 Los RNAs *ascas2-1* y *ascse2-1* están presentes en la hebra antisentido del locus CRISPR-Cas

Debido a que un RNA fue identificado en la hebra líder del locus CRISPR-Cas de *S. Typhi*, específicamente en la región codificante de *cse2*, se analizó la presencia de otras unidades transcripcionales en la cadena antisentido de dicho locus. Para lo anterior, se evaluaron diferentes fusiones transcripcionales de la cadena reversa del arreglo CRISPR-Cas en N-MM. Los resultados demuestran la presencia de dos promotores independientes. El primero promueve la transcripción de una unidad denominada *ascas2-1*, que fue identificada al analizar una fusión que contenía parte de la cadena antisentido de *cas2* y *cas1* [*ascas2-1* (-5556/-6245)]. Esta construcción presentó 310 unidades CAT en el fondo *S. Typhi* silvestre (Figura 9.21A).

Con el objetivo de definir la longitud de este transcrito, se construyó la fusión *ascas2-1* (-5355/-6245). Sin embargo, su actividad específica fue nula en N-MM (Figura 9.21A), sugiriendo que el transcrito *ascas2-1* se encuentra entre las bases -5556 a -6245, respecto al ATG de *cas3*. Para delimitar con mayor precisión el tamaño de esta unidad transcripcional, se realizaron ensayos de RT-PCR con RNA total de *S. Typhi* cultivada en N-MM. El cDNA fue sintetizado con el oligonucleótido *ascas1-5556 KpnI-R* (Tabla 3), ubicado en la base -5556 respecto al ATG de *cas3*. Diversos PCRs fueron realizados utilizando el oligo mencionado y de manera individual, diferentes primers localizados en regiones corriente arriba. De este modo, se amplificó un fragmento de DNA con oligonucleótidos específicos, localizados en las bases -5556 a -5775 (Figura 9.21B). No obstante, usando primers ubicados de -5556 a -5858 no se obtuvo amplificación, indicando

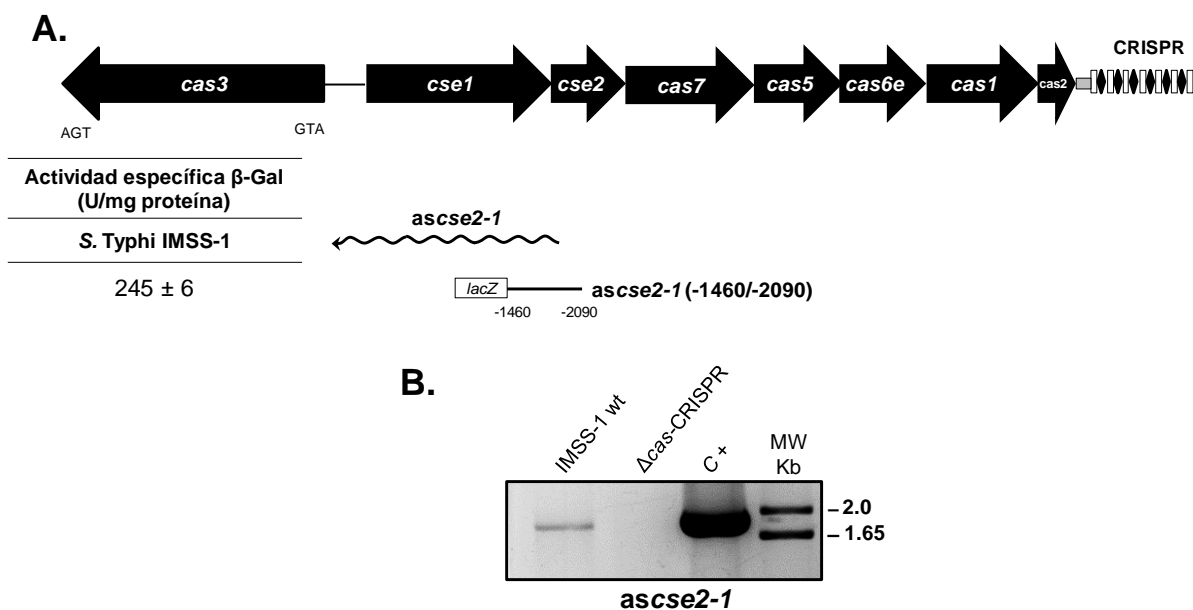
que el RNA *ascas2-1* se encuentra en los nucleótidos -5556 a -5775, respecto al inicio de traducción de *cas3* y tiene un tamaño aproximado de 220 pb.



**Figura 9.21. La unidad transcripcional *ascas2-1* se localiza en la cadena antisentido del locus CRISPR-Cas.** **A.** Representación esquemática del sistema CRISPR-Cas de *S. Typhi*. La flecha ondulada representan al RNA antisentido *ascas2-1*. Los números en cada fusión corresponden a las coordenadas respecto al ATG de *cas3*. La columna de la izquierda muestra el perfil de expresión de las construcciones correspondientes. Los ensayos transcripcionales se realizaron en la cepa *S. Typhi* IMSS-1 cultivada en N-MM a  $OD_{595\text{nm}}:1.3$ . **B.** RT-PCR de *ascas2-1*. RNA total de la cepa IMSS-1 wt en N-MM ( $OD_{595\text{nm}}: 1.3$ ) fue utilizado para la síntesis de cDNA, con el cual se amplificó un fragmento de 220 pb correspondiente al RNA *ascas2-1*. El control positivo (C+) corresponde al fragmento amplificado con DNA genómico como templado. Como control de especificidad se utilizó RNA de *S. Typhi*  $\Delta\text{cas-CRISPR}$ . MW: marcador de peso molecular.

Continuando con la búsqueda de unidades transcripcionales presentes en la cadena antisentido del clúster CRISPR-Cas, se localizó un segundo promotor dentro de la fusión *ascse2-1* (-1460/-2090), la cual contiene los nucleótidos -1460 a -2090, respecto al ATG de *cas3*. Esta construcción presentó niveles de expresión de 250 unidades  $\beta$ -galactosidasa en la cepa *S. Typhi* IMSS-1 (Figura 9.22A). La longitud de este transcrito se definió mediante experimentos de RT-PCR. La síntesis de cDNA fue realizada con RNA total de la cepa IMSS-1 y diferentes oligos localizados corriente abajo del nucleótido -1460. Utilizando el primer *ascse1-34 BamHI-R* (Tabla 3), ubicado en la base -34 pb respecto al inicio de traducción de *cas3*, se logró sintetizar un cDNA. Para la amplificación por PCR, este cDNA, junto con primers ubicados en los nucleótidos -34 y -1958, generaron

un fragmento de 1,925 pb, que abarca parte de la cadena antisentido *cse2* a *cse1*, razón por la cual esta unidad fue nombrada *ascse2-1* (Figura 9.22B).



**Figura 9.22. El RNA *ascse2-1* se transcribe en la cadena antisentido del locus CRISPR-Cas.**  
**A.** Representación esquemática del sistema CRISPR-Cas de *S. Typhi*. La flecha ondulada representa a la unidad transcripcional *ascse2-1*. Los números en la fusión corresponden a las coordenadas respecto al ATG de *cas3*. La columna de la izquierda muestra el perfil transcripcional de la construcción correspondiente, realizado en la cepa *S. Typhi* IMSS-1 cultivada en N-MM a  $OD_{595nm}:1.3$ . **B.** RT-PCR de *ascse2-1*. RNA total de la cepa IMSS-1 wt en N-MM ( $OD_{595nm}: 1.3$ ) fue utilizado para la síntesis del cDNA, con el cual se amplificó un fragmento de 1,925 pb correspondiente al RNA *ascse2-1*. El control positivo (C+) corresponde al fragmento amplificado con DNA genómico como templado. Como control de especificidad se utilizó RNA de *S. Typhi*  $\Delta cas$ -CRISPR. MW: marcador de peso molecular.

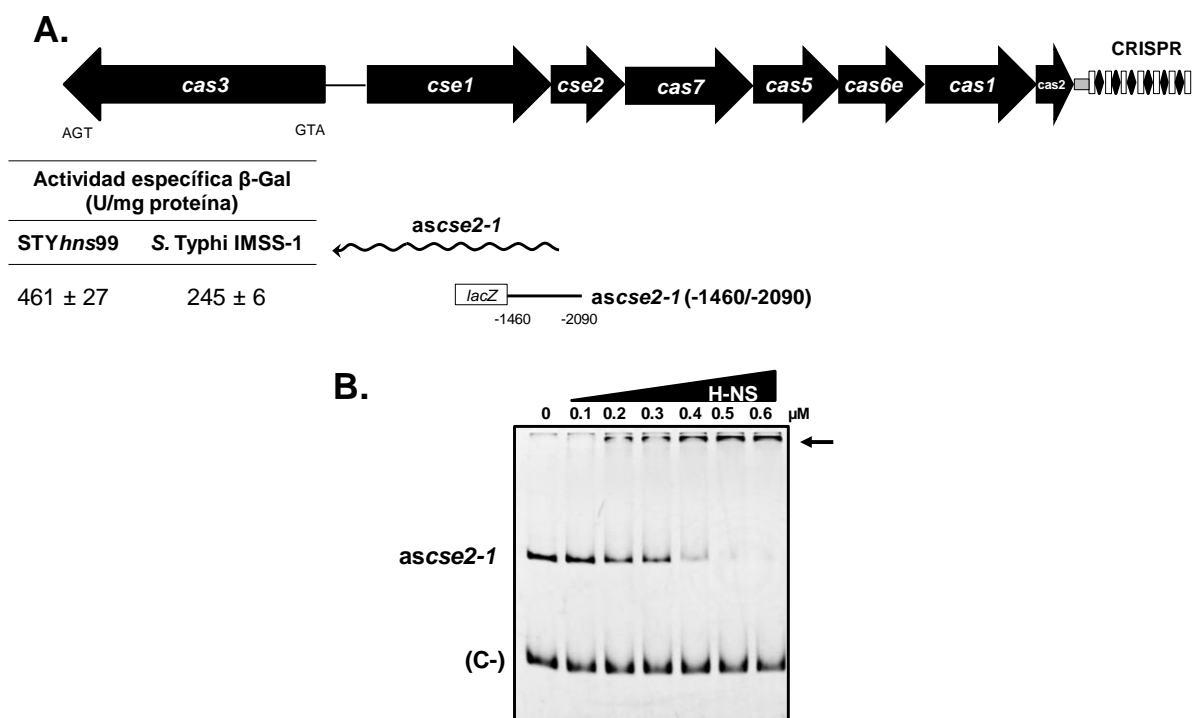
Los datos presentados hasta el momento demuestran la presencia de cinco unidades transcripcionales en el locus CRISPR-Cas de *S. Typhi* IMSS-1: el operón *cas*-CRISPR y *scse2* se localizan en la cadena sentido, mientras que *ascas2-1* y *ascse2-1* corresponden a RNAs presentes en la cadena antisentido. Adicionalmente, el gen *cas3* se transcribe como una unidad independiente y divergente al clúster genético *cas*-CRISPR.

### 9.9 H-NS regula negativamente el RNA antisentido *ascse2-1*

Debido a que H-NS participa en la regulación negativa de las unidades transcripcionales *cas3* y el operón *cas*-CRISPR, se evaluó si dicha proteína también reprime la expresión de los tres RNAs identificados: *scse2*, *ascas2-1* y

*ascse2-1*. Los perfiles transcripcionales realizados en la cepa deficiente en *hns*, conteniendo de manera independiente las construcciones *cse2* (-450/+181), *ascas2-1* (-5556/-6245) y *ascse2-1* (-1460/-2090) mostraron que solamente el antisentido *ascse2-1* se encuentra regulado negativamente por esta proteína, ya que la expresión de la respectiva fusión fue mayor en la ausencia de *hns* que en la cepa silvestre (Figura 9.23A). Mientras que la expresión de los promotores *scse2* y *ascas2-1* fue similar en los fondos genéticos *S. Typhi* IMSS-1 y *STYhns99* (datos no mostrados).

Mediante experimentos EMSA con la proteína H-NS purificada y la región reguladora de *ascse2-1*, se observó que este factor transcripcional es capaz de interactuar con dicho fragmento de DNA (Figura 9.23B), indicando que H-NS reprime de manera directa al antisentido *ascse2-1*.



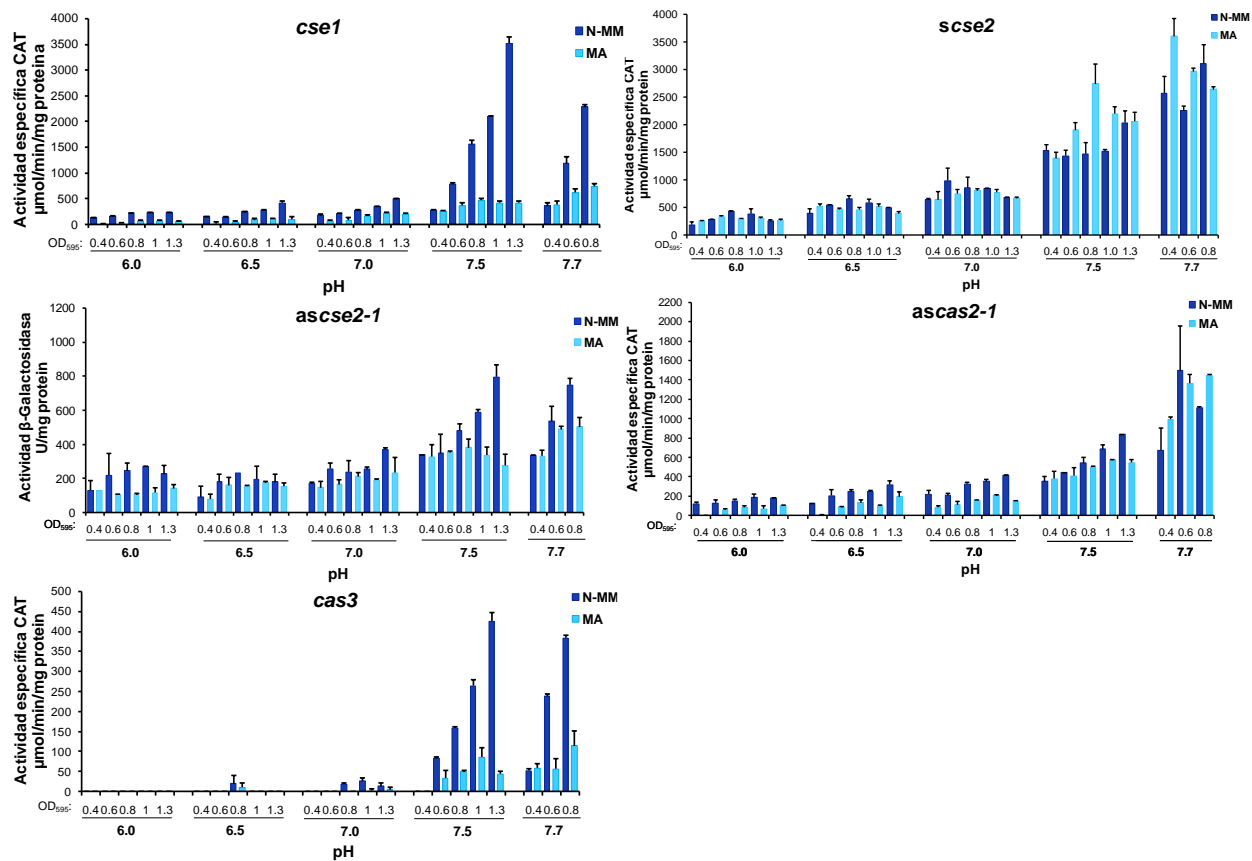
**Figura 9.23. H-NS reprime el promotor *ascse2-1*.** **A.** Representación esquemática del sistema CRISPR-Cas de *S. Typhi*. La flecha ondulada representa a la unidad transcripcional *ascse2-1*. Los números en la fusión corresponden a las coordenadas respecto al ATG de *cas3*. Las columnas de la izquierda muestran el perfil transcripcional de la construcción correspondiente, realizado en la cepa *S. Typhi* IMSS-1 y *STYhns99* cultivadas en N-MM a  $OD_{595nm}:1.3$ . **B.** EMSA realizado con concentraciones crecientes de H-NS y un fragmento de 630 pb que contiene la región reguladora de *ascse2-1* (nucleótidos -1460 a -2090, respecto al ATG de *cas3*). Un fragmento de 242 pb del gen 16S rRNA de *S. Typhi*, fue utilizado como control negativo (C-). La flecha indica el complejo DNA-proteína visualizado mediante el uso de bromuro de etidio.

### 9.10 El sistema CRISPR-Cas de *S. Typhi* es inducido en pH 7.5

Como se describió anteriormente, el sistema CRISPR-Cas de *S. Typhi* contiene cinco unidades transcripcionales que se expresan en medio mínimo N, por lo que es relevante identificar el componente presente en esta condición de crecimiento, responsable de la inducción del locus CRISPR-Cas.

Para cumplir el objetivo planteado, se evaluó la influencia de la fuente de carbono que contiene el N-MM (glicerol), la cual fue reemplazada por glucosa. Al evaluar los promotores *cse1*, *scse2* y *cas3* se observó que su actividad transcripcional fue similar en ambas fuentes de carbono (datos no mostrados). Por lo tanto, el glicerol no es la señal que promueve la expresión de las unidades presentes en el sistema CRISPR-Cas de *S. Typhi*.

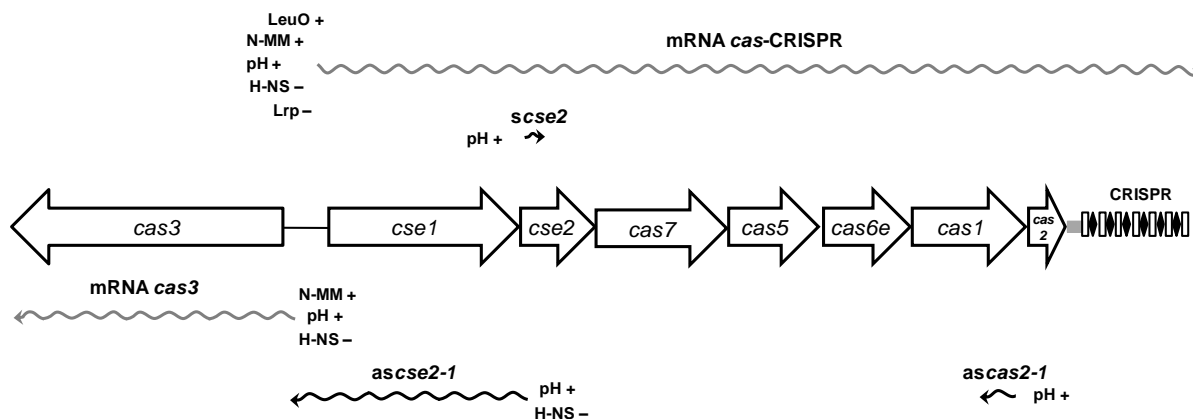
Posteriormente, se analizó el efecto del pH presente en medio mínimo N (7.5) en la actividad del locus CRISPR-Cas. Los niveles de actividad de los promotores *cse1*, *scse2*, *ascse2-1*, *ascas2-1* y *cas3* en pH 6.0-6.5 fueron menores, comparados con la expresión detectada a pH 7.5-7.7 en N-MM (Figura 9.24). Para confirmar que el pH es el responsable de la inducción transcripcional de las cinco unidades del sistema CRISPR-Cas, los promotores correspondientes se evaluaron en medio rico MA a pH 6.0, 6.5, 7.0, 7.5 y 7.7. A pesar de que este medio es diferente al N-MM, la expresión de *cse1*, *scse2*, *ascse2-1*, *ascas2-1* y *cas3* también aumentó en pH 7.5-7.7, respecto a la actividad detectada a pH 6.0-7.0 (Figura 9.24), validando que el pH 7.5 es una señal ambiental que promueve la expresión de las cinco unidades transcripcionales presentes en el sistema CRISPR-Cas.



**Figura 9.24. El pH induce la expresión transcripcional del sistema CRISPR-Cas.** Perfiles transcripcionales de *S. Typhi* IMSS-1 conteniendo de manera independiente las fusiones de los promotores *cse1*, *scse2*, *ascse2-1*, *ascas2-1* y *cas3*. Los ensayos fueron realizados en N-MM (barras azul marino) y MA (barras azul claro) a diferentes pHs (6.0, 6.5, 7.0, 7.5 y 7.7).

Debido a que todas las unidades transcripcionales identificadas en el locus CRISPR-Cas se sobreexpresan en pH 7.5, se evaluó el papel de este sistema en la resistencia de *Salmonella* a esta condición. Se realizaron curvas de crecimiento en N-MM a pH 7.5, comparando a *S. Typhi* silvestre con una cepa deficiente en el clúster CRISPR-Cas. Los resultados mostraron la misma tasa de crecimiento para ambos fondos genéticos, indicando que el sistema CRISPR-Cas no es necesario para la replicación de este patógeno a pH 7.5 (datos no mostrados).

En conclusión, el sistema CRISPR-Cas de *S. Typhi* presenta cinco unidades transcripcionales, los elementos proteicos (LeuO, H-NS, Lrp) y señales (medio mínimo N, pH) implicados en su regulación se muestran en la figura 9.25.



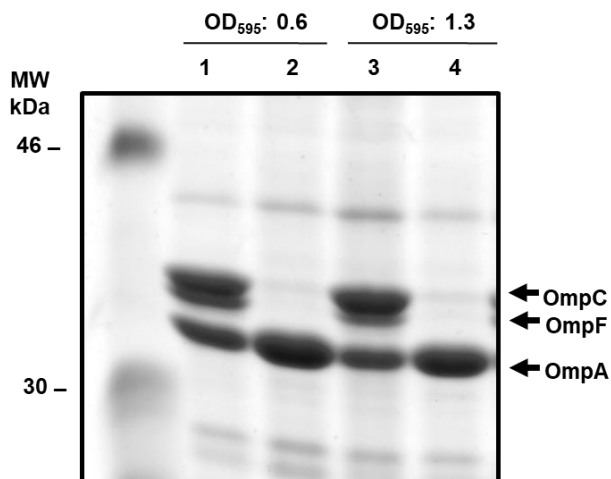
**Figura 9.25. Organización transcripcional y elementos implicados en la regulación del sistema CRISPR-Cas de *S. Typhi*.** Las líneas grises onduladas representan los mRNAs de *cas3* y el operón *cas-CRISPR*. Los RNAs *scse2*, *ascse2-1* y *ascas2-1* se encuentran en líneas negras onduladas. Se muestran los elementos implicados en la regulación positiva (+) y negativa (-) de cada una de las unidades transcripcionales del locus CRISPR-Cas.

### 9.11 El locus CRISPR-Cas de *S. Typhi* está implicado en la síntesis de proteínas de membrana externa

El sistema CRISPR-Cas de *Salmonella enterica* serovar Typhi contiene múltiples unidades transcripcionales presentes en ambas cadenas de DNA, cuya actividad es inducida en pH 7.5. Sin embargo, dichas unidades no se requieren para que esta bacteria sobreviva en el valor de pH mencionado. Por lo tanto, no se conoce el proceso biológico en el cual se encuentra involucrado el sistema CRISPR-Cas en *S. Typhi*. Considerando reportes que demuestran la participación de éste locus en el control de la síntesis de algunas proteínas de membrana externa, como la lipoproteína BLP de *Francisella novicida* (Sampson *et al.*, 2013), se determinó si la deficiencia del locus CRISPR-Cas de *S. Typhi* ocasiona algún cambio en el número o cantidad de proteínas de membrana externa (OMPs). Para lo anterior se purificaron las OMPs de cultivos de *S. Typhi* silvestre y carente en el clúster CRISPR-Cas, cultivadas en N-MM.

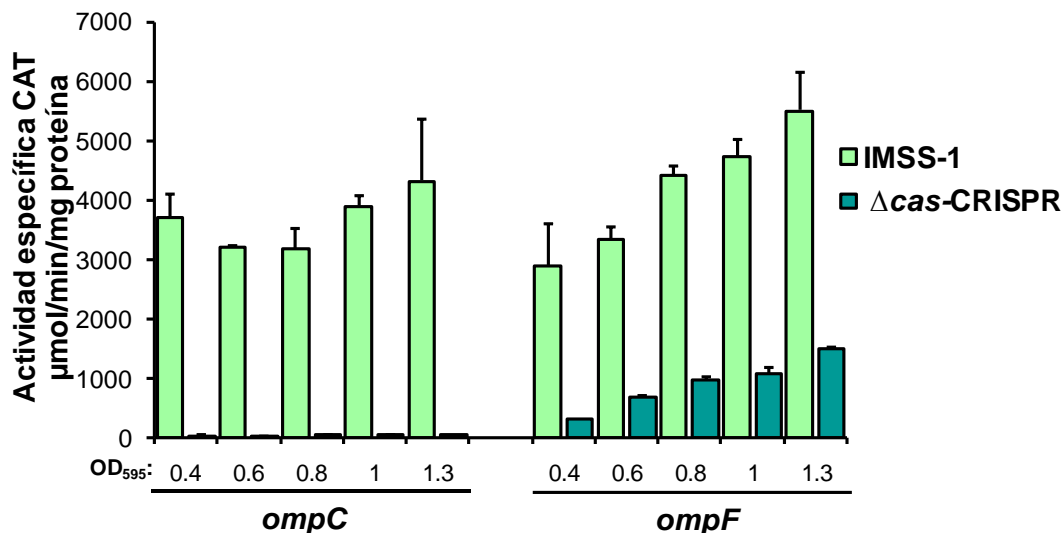
En el perfil de proteínas de membrana externa de *S. Typhi* silvestre, la presencia de las porinas mayoritarias OmpC y OmpF es evidente. Sin embargo, dichas proteínas no se observan en la cepa  $\Delta$ *cas-CRISPR*, indicando que este sistema es necesario para la síntesis de OmpC y OmpF (Figura 9.26).





**Figura 9.26 El sistema CRISPR-Cas es necesario para la producción de OmpC y OmpF en *S. Typhi*.** Perfil de proteínas de membrana externa (OMPs) de las cepas *S. Typhi* IMSS-1 silvestre (líneas 1, 3) y *S. Typhi* IMSS-1  $\Delta cas$ -CRISPR (líneas 2, 4), cultivadas en medio mínimo N a OD<sub>595nm</sub>: 0.6 y 1.3. Las flechas indican las tres OMPs más abundantes: OmpC, OmpF y OmpA. Se incluye el marcador de peso molecular (MW). Las proteínas fueron visualizadas utilizando azul de Coomasie.

Para poder establecer el nivel de regulación en el cual participa el sistema CRISPR-Cas para la síntesis de OmpC y OmpF, se generaron fusiones transcripcionales de las regiones reguladoras de los genes que codifican para estas porinas y su actividad fue analizada en la cepa de *S. Typhi* silvestre y en la deficiente en el locus CRISPR-Cas. Como se observa en la figura 9.27, los valores de expresión obtenidos de *ompC* y *ompF* disminuyeron notablemente en la  $\Delta cas$ -CRISPR, comparados con la transcripción detectada en la cepa silvestre. Estos datos sugieren que el sistema CRISPR-Cas está involucrado en la regulación transcripcional de los genes *ompC* y *ompF*.



**Figura 9.27. El sistema CRISPR-Cas está implicado en la regulación de los promotores *ompC* y *ompF*.** Perfiles transcripcionales de las cepas *S. Typhi* IMSS-1 silvestre (barras verde claro) y *S. Typhi* Δcas-CRISPR (barras verde fuerte), conteniendo de manera independiente fusiones transcripcionales de *ompC* u *ompF*. Los ensayos fueron realizados en N-MM a diferentes valores de O.D.<sub>595nm</sub>.

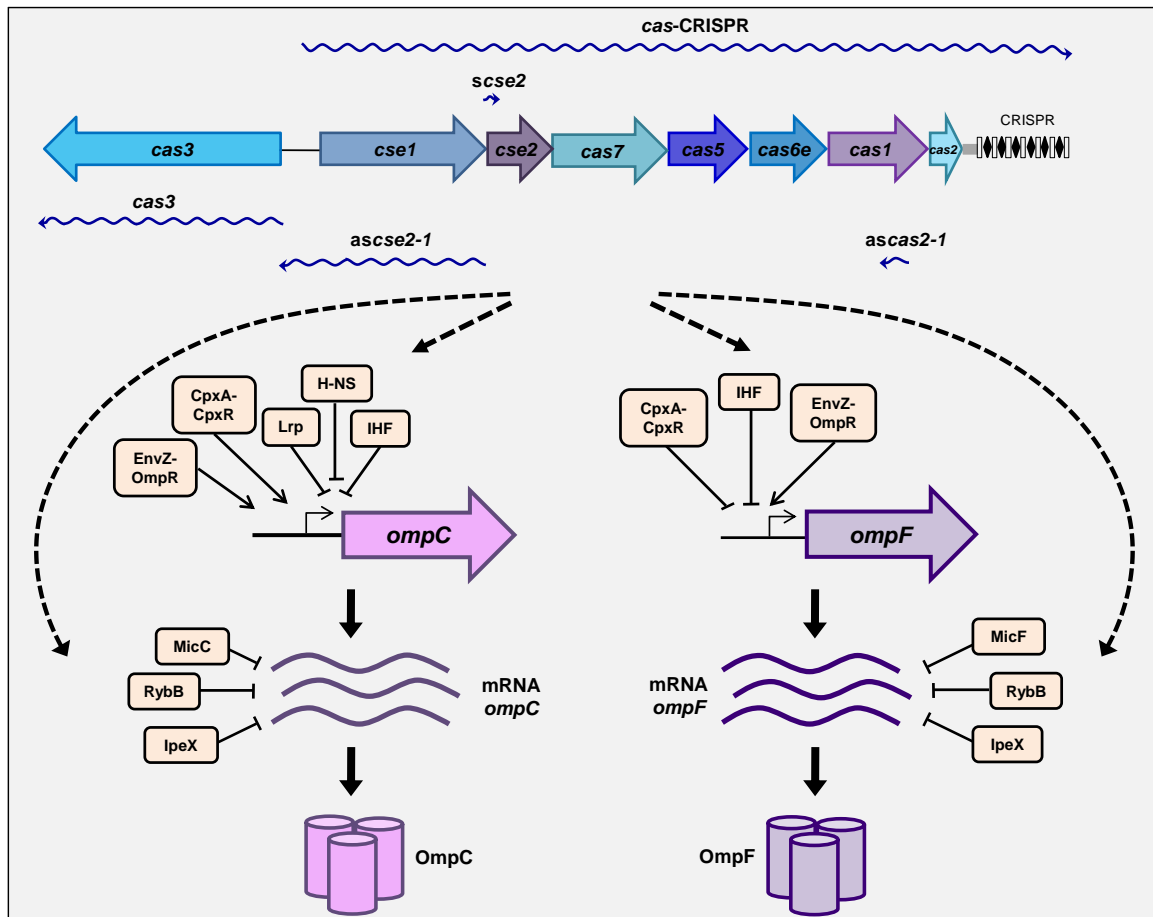
Considerando que diferentes factores transcripcionales como OmpR, CpxRA, IHF, Lrp o H-NS participan como reguladores de la expresión de los promotores *ompC* y *ompF* (Batchelor *et al.*, 2005; De la Cruz & Calva, 2010; Ferrario *et al.*, 1995; Forst *et al.*, 1989; Huang *et al.*, 1990; Puente *et al.*, 1991; Suzuki *et al.*, 1996), además de que los RNAs mensajeros correspondientes a estas proteínas de membrana externa están bajo la modulación post-transcripcional de microRNAs (MicC, MicF, RybB, IpeX) (Castillo-Keller *et al.*, 2006; Chen *et al.*, 2004; Mizuno *et al.*, 1984; Papenfort *et al.*, 2006), es relevante determinar si la regulación ejercida por CRISPR-Cas es de manera directa sobre *ompC* y *ompF*, o bien, a través de sus microRNAs reguladores o proteínas implicadas en su control transcripcional.

Es primordial identificar los componentes del sistema CRISPR-Cas involucrados en el control de *ompC* y *ompF*, es decir, si es el arreglo CRISPR, los RNAs (*scse2*, *ascse2-1* o *ascas2-1*) o las proteínas Cas. Estas últimas podrían participar de manera individual o bien formando un complejo efector. En el caso de que sean las proteínas Cas, éstas podrían interactuar con algún regulador positivo para promover directamente la expresión de los promotores *ompC* y *ompF*. Es

también posible que las proteínas Cas contrarresten el efecto de algún represor proteico de los genes que codifican para las porinas e inducir su transcripción. Otra alternativa sería que las Cas secuestren microRNAs inhibidores de la traducción de los mensajeros *OmpC* y *OmpF*, promoviendo de esta manera la síntesis de las proteínas de membrana externa (Figura 9.28).

Opcionalmente, si las secuencias CRISPR o los RNAs *scse2*, *ascse2-1* o *ascas2-1* estuvieran involucrados en la regulación de las porinas, el mecanismo que se sugiere sería la interacción RNA-RNA de alguno de los componentes del locus CRISPR-Cas con los mensajeros *ompC* y *ompF*. Dicha interacción estabilizaría a los mRNAs, llevándose a cabo su traducción.

Otra posibilidad que se debe considerar, es que los ácidos ribonucleicos pertenecientes al locus CRISPR-Cas (*CRISPR*, *scse2*, *ascse2-1* o *ascas2-1*) interactúen con el mRNA de algún regulador negativo de los genes *ompC* u *ompF*, o bien con los microRNAs que participan en su modulación a nivel post-transcripcional. En ambos casos el híbrido de RNA sería degradado, llevándose a cabo la producción de las proteínas de membrana externa *OmpC* y *OmpF* (Figura 9.28). Es fundamental llevar a cabo estos estudios con el fin de establecer el mecanismo de regulación ejercido por el sistema CRISPR-Cas, en la síntesis de las proteínas de membrana externa *OmpC* y *OmpF* de *S. Typhi*.



**Figura 9.28. Alternativas del mecanismo de acción ejercido por el sistema CRISPR-Cas en la regulación de las porinas OmpC y OmpF en *S. Typhi*.** Uno o varios de los elementos genéticos que conforman el sistema CRISPR-Cas en *S. Typhi* IMSS-1 pueden estar regulando algún factor transcripcional implicado en el control de los promotores *ompC* y *ompF* (líneas rectas punteadas). Otra opción radica en que una de las unidades transcripcionales, presentes en el locus CRISPR-Cas, participe en la modulación de los RNAs pequeños involucrados en el control posttranscripcional de los mensajeros *ompC* y *ompF* (líneas curvas punteadas), promoviendo la síntesis de las proteínas de membrana externa correspondientes.

## 10. DISCUSIÓN

El locus CRISPR-Cas ha sido estudiado principalmente por su función como sistema inmune en bacterias y arqueas. En los últimos años, el interés por este sistema ha aumentado notablemente debido a su utilidad como herramienta de ingeniería genética, con la cual se ha podido editar el genoma de bacterias y organismos eucariontes incluyendo plantas, ratones y DNA humano. Sin embargo, existen pocos estudios acerca de los elementos genéticos y señales que controlan la expresión del sistema CRISPR-Cas, así como de su organización transcripcional. Los datos obtenidos en el presente proyecto de investigación contribuyen en este aspecto.

En este trabajo se realizó la caracterización genética de cinco unidades transcripcionales presentes en el locus CRISPR-Cas de *Salmonella enterica* serovar Typhi (operón *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR*, *scse2*, *ascse2-1*, *ascas2-1* y *cas3*), las cuales se expresan en medio mínimo N y su actividad incrementa en pH 7.5. La presencia de dichos transcritos demuestra que este sistema constituye un operón complejo. De manera interesante, diversos transcritos también han sido identificados en el clúster CRISPR-Cas de una bacteria estrechamente relacionada con *Salmonella*, como lo es *E. coli*. Esta enterobacteria contiene un sistema CRISPR-Cas Tipo I-E y, al igual que *S. Typhi*, se compone por los genes *cas3*, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, *cas2* y un arreglo CRISPR. Se han identificado cuatro unidades transcripcionales independientes en dicho locus. La primera corresponde al RNA mensajero de la nucleasa *cas3*; corriente abajo se encuentran los genes *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*, los cuales conforman un solo RNA policistrónico, cuyo promotor fue nombrado *Pcas*. La tercera unidad corresponde al arreglo CRISPR. Finalmente, el cuarto promotor nombrado anti-*Pcas* fue localizado en la cadena antisentido de la región intergénica de *cas3-cse1* y controla la expresión de un RNA de 150-200 pb (Pul *et al.*, 2010; Westra *et al.*, 2010). En el arquea *Sulfolobus islandicus*, se ha demostrado que los genes *cas* Tipo I-A se transcriben en dos RNA mensajeros: *csa1-cas1* y el operón *acas* (*cas1-cas2-cas4*) (Liu *et al.*, 2015).

Todos estos estudios indican que la presencia de múltiples unidades transcripcionales es una característica de los sistemas CRISPR-Cas de Tipo I. Sin embargo, el número y la composición de los transcritos presentes en este locus, parece ser específico de cada microorganismo.

Respecto a los elementos proteicos implicados en el control transcripcional del sistema CRISPR-Cas de *S. Typhi*, la caracterización del promotor *cse1* (promotor del operón *cas*-CRISPR) demostró que la proteína LeuO participa en su regulación positiva, ya que la sobreexpresión de este regulador induce la actividad de *cse1* a 507 unidades CAT en medio rico MA. No obstante, la expresión de dicho promotor en medio mínimo N es de 1,860 unidades CAT, la cual resultó ser independiente de LeuO, debido a que valores similares de expresión fueron observados tanto en una cepa silvestre como en una  $\Delta leuO$ . Por lo tanto, existe otro regulador no identificado encargado de activar la expresión de *cse1* en esta condición de crecimiento. Opcionalmente, en medio MA puede estar presente un represor, cuya actividad no es contrarrestada por LeuO. Además, se determinó que los factores transcripcionales H-NS y Lrp reprimen de manera directa al operón *cas*-CRISPR. El papel de H-NS en la regulación negativa de otras dos unidades transcripcionales (*cas3* y *ascse2-1*) también fue establecido en *S. Typhi*.

Estudios sobre la regulación del sistema CRISPR-Cas en *E. coli* K12 demostraron la participación de LeuO y BaeR en la regulación positiva del operón *cse1-cse2-cas7-cas5-cas6e-cas1-cas2* (Perez-Rodriguez *et al.*, 2011; Westra *et al.*, 2010). Respecto a la regulación negativa de este transcrito, los factores transcripcionales H-NS y CRP están implicados en su silenciamiento (Pul *et al.*, 2010; Yang *et al.*, 2014). De manera interesante, H-NS también reprime a los promotores *cas3*, *anti-Pcas* y *Pcrispr1*, presentes en el sistema CRISPR-Cas de *E. coli* (Majsec *et al.*, 2016; Pul *et al.*, 2010). Análisis transcripcionales realizados en otros miembros de la familia *Enterobacteriaceae*, establecieron que en *Klebsiella pneumoniae* NTUH-K2044 H-NS participa en la represión de *cas3* (Lin *et al.*, 2016). En el caso del patógeno de plantas *Pectobacterium atrosepticum*, CRP interactúa con la región reguladora de *cas1* activando la transcripción de los genes

*cas1-cas2-3-csy1-csy2-csy3-cas6f* (Patterson *et al.*, 2015). Por otro lado, ensayos de microarreglos realizados en *Yersinia pestis* CO92, indican la participación del regulador global OmpR en el control de *cas1*, ya que la expresión de dicho gen disminuyó en una cepa carente en *ompR*, comparada con los niveles de transcrito obtenidos en la cepa silvestre (Gao *et al.*, 2011; Louwen *et al.*, 2014).

Toda esta información sugiere que hay elementos de regulación comunes para los sistemas CRISPR-Cas presentes en la familia *Enterobacteriaceae*, como es el caso de H-NS, ya que regula más de una unidad transcripcional presente en el locus CRISPR-Cas de *E. coli* K12 y *S. Typhi* IMSS-1. Además, está involucrado en el silenciamiento de *cas3* en *K. pneumoniae* NTUH-K2044. Sin embargo, otros factores transcripcionales son muy específicos según la especie bacteriana. Por ejemplo, Pul y Cols. (2010) reportaron que Lrp no tiene un efecto en la expresión de los genes *cas* de *E. coli*, mientras que esta proteína es un regulador negativo del operón *cas*-CRISPR en *S. Typhi*. A pesar de que CRP participa en la represión del sistema CRISPR-Cas de *E. coli* o en la activación de este locus en *P. atrosepticum* y OmpR es un posible activador de *cas1* en *Y. pestis*, nuestros estudios demuestran que estos reguladores globales no influyen en la expresión del sistema CRISPR-Cas de *S. Typhi*. En el caso de LeuO, su papel como regulador positivo de este clúster genético solo ha sido determinado en *E. coli* y *S. Typhi*. En consecuencia, la regulación transcripcional de este programa genético responde, presumiblemente, a la presencia de múltiples señales ambientales detectadas por diversas proteínas reguladoras, las cuales están involucradas en el ciclo de vida del procarionte hospedero.

El estudio del control de la expresión del sistema CRISPR-Cas de *S. Typhi*, tenía como objetivo proporcionar una idea sobre el papel funcional que desempeña dicho sistema en la fisiología de *Salmonella*. En este aspecto, la actividad de las cinco unidades presentes en el locus CRISPR-Cas de *S. Typhi* incrementa en pH 7.5, este valor de pH es característico de la parte distal del íleon (Evans *et al.*, 1988), compartimento que es invadido por *S. Typhi* para llegar al sistema linfático y diseminarse a hígado, bazo, médula ósea y vesícula biliar,

provocando la muerte del hospedero. Por lo tanto, es relevante determinar si estos elementos genéticos participan en la respuesta bacteriana al ambiente presente en el intestino delgado, o si tienen un papel directo en la invasión de las células M de este órgano humano.

Además de observar la respuesta del sistema CRISPR-Cas a un valor de pH que enfrenta *S. Typhi* durante su ciclo de infección, nuestros resultados también indican que la ausencia de este locus promueve la desaparición de las proteínas de membrana externa mayoritarias OmpC y OmpF. Por lo que es relevante determinar el mecanismo a través del cual CRISPR-Cas está implicado en la producción de las porinas mencionadas. Los resultados obtenidos sugieren que este sistema regula los genes *ompC* y *ompF* a nivel de la transcripción, si lo anterior ocurre de manera directa, constituiría un mecanismo de regulación no descrito hasta el momento para dicho programa genético.

Se ha reportado la participación del locus CRISPR-Cas en el mantenimiento de la envoltura celular o en la respuesta a estrés de membrana en otros microorganismos. Por ejemplo, en el patógeno intracelular *Francisella novicida*, la nucleasa Cas9 junto con los RNAs pequeños tracrRNA y scaRNA son fundamentales en la regulación del transcrito que codifica para la lipoproteína de membrana externa BLP. Lo anterior ocurre cuando el tracrRNA interactúa con el scaRNA, este híbrido forma un complejo ribonucleoproteico con la nucleasa Cas9. Posteriormente, el tracrRNA por complementariedad, hibrida con el RNA mensajero de BLP, llevándose a cabo la degradación de los transcritos de esta lipoproteína, posiblemente mediada por Cas9 o alguna RNasa no identificada (Sampson *et al.*, 2013). La ausencia de BLP promueve que la membrana externa sea menos permeable, confiriendo resistencia a diferentes antibióticos como polimixina B, estreptomycin y kanamicina. Además, las cepas carentes de BLP evitan el reconocimiento por los receptores Toll-like 2, con lo cual la bacteria puede evadir la respuesta del sistema inmune innato del hospedero (Sampson *et al.*, 2014).



En *Streptococcus mutans*, causante de caries dental, cepas deficientes en el locus CRISPR-Cas Tipo II-A son más sensibles al estrés de membrana inducido por detergentes, peróxido de hidrógeno, paraquat y altas temperaturas (Serbanescu *et al.*, 2015). En *E. coli* BW25113, se determinó que el sistema regulador de respuesta a estrés de membrana BaeRS induce la expresión del clúster genético CRISPR-Cas Tipo I-E (Perez-Rodriguez *et al.*, 2011). Es conocido que el estrés de la envoltura celular inducido por fagos al inyectar su material genético, tanto en bacterias como en arqueas, también promueve la activación del sistema CRISPR-Cas (Agari *et al.*, 2010; Fusco *et al.*, 2015; Young *et al.*, 2012). Por lo tanto, todos los datos mencionados sugieren que este clúster genético tiene la habilidad de responder a diversas condiciones que amenazan la integridad de la membrana celular (Ratner *et al.*, 2015).

Debido a que en *S. Typhi* el sistema CRISPR-Cas influye positivamente en la producción de las proteínas de membrana externa OmpC y OmpF, es necesario establecer si este clúster genético responde a señales que inducen la expresión de los genes que codifican para las porinas mencionadas, y si participa indirectamente en alguno de los procesos biológicos en los cuales están involucradas estas proteínas de membrana externa, tales como: virulencia, respuesta a alta osmolaridad, transformación bacteriana por electroporación, resistencia a antibióticos y a la sal biliar deoxicolato de sodio (Chatfield *et al.*, 1991; Kaeriyama *et al.*, 2006; Medeiros *et al.*, 1987; Villarreal *et al.*, 2014). Estudiando estos fenómenos biológicos estableceremos el papel funcional que desempeña el sistema CRISPR-Cas en el género *Salmonella*.

## 11. CONCLUSIONES

La caracterización transcripcional del sistema CRISPR-Cas Tipo I-E presente en la bacteria *Salmonella enterica* serovar Typhi IMSS-1 reveló lo siguiente:

- LeuO es un regulador positivo de la expresión del operón *cas*-CRISPR, este factor transcripcional interactúa con dos zonas de la región intergénica 5' de *cse1* (nucleótidos -56 a -96 y -102 a -131, respecto al inicio de transcripción de *cse1*). Las bases ubicadas en la posición -88 a -90 (ATA) son fundamentales para la actividad transcripcional de este promotor mediada por LeuO.
- H-NS y Lrp participan en la regulación negativa del promotor *cse1*. Ambos factores de transcripción se unen en los nucleótidos -375 a -50 y de +49 a +191, respecto al inicio de transcripción de *cse1*.
- *cas3* es reprimido directamente por H-NS, ya que esta proteína interactúa con los nucleótidos -241 a -181, respecto al ATG de dicho gen.
- El locus CRISPR-Cas está compuesto de cinco unidades transcripcionales: el operón *cas*-CRISPR (*cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR) y *scse2* se encuentran en la hebra sentido, *ascse2-1* y *ascas2-1* están presentes en la cadena antisentido de este locus. Adicionalmente, el gen *cas3* se transcribe como una unidad independiente, la cual es divergente al operón *cas*-CRISPR.
- El regulador global H-NS reprime de manera directa el promotor de *ascse2-1*.
- Las cinco unidades presentes en el locus CRISPR-Cas se expresan en medio mínimo N y la presencia de pH 7.5-7.7 promueve un incremento en su expresión.
- El sistema CRISPR-Cas Tipo I-E de *S. Typhi* IMSS-1 está implicado en la regulación de los genes que codifican para las proteínas de membrana externa OmpC y OmpF.

## 12. PERSPECTIVAS

Los resultados obtenidos en este proyecto de investigación, sobre la caracterización transcripcional de las cinco unidades genéticas presentes en el locus CRISPR-Cas (operón *cas*-CRISPR, *scse2*, *ascse2-1*, *ascas2-1* y *cas3*) del patógeno *S. Typhi*, dan pauta a una primer perspectiva, la cual radica en describir completamente la red de regulación global para este sistema en la bacteria mencionada. Esto se debe a que *LeuO* no funge como el activador principal del sistema CRISPR-Cas, ya que los niveles de actividad observados en medio mínimo N son cuatro veces mayores a los obtenidos en medio rico MA al sobreexpresar *LeuO*. Por lo tanto, deben identificarse el o los factores transcripcionales implicados en la activación del operón *cas*-CRISPR en N-MM. Así mismo, los reguladores involucrados en el control positivo de los transcritos *cas3* y *ascse2-1* también deben ser determinados, al igual que el represor adicional a H-NS, implicado en el silenciamiento de *cas3*. En el caso de los RNAs *scse2* y *ascas2-1*, es relevante identificar los elementos genéticos que participan en su regulación transcripcional tanto positiva como negativa, ya que en el presente estudio no fueron establecidos.

Debido a que la expresión de los cinco transcritos presentes en el locus CRISPR-Cas de *S. Typhi* se induce en pH 7.5, es relevante elucidar la red de regulación implicada en este fenómeno biológico, es decir, identificar los sensores responsables de detectar los cambios en el pH extracelular, así como los reguladores involucrados en la activación de los promotores *cas*-CRISPR, *scse2*, *ascse2-1*, *ascas2-1* y *cas3* en respuesta a la presencia de pH 7.5.

Además de conocer los factores que controlan la transcripción de las unidades presentes en el arreglo CRISPR-Cas tanto en N-MM como en pH 7.5, es necesario evaluar la regulación a nivel postranscripcional. En este aspecto, se debe determinar si los antisentidos *ascse2-1* y *ascas2-1* están implicados en la regulación del RNA mensajero proveniente del operón *cas*-CRISPR, esto considerando que dichas unidades son complementarias al locus *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR. El identificar todos los elementos involucrados en

la regulación del sistema CRISPR-Cas proporcionará una idea del tipo de señales o condiciones a las cuales puede responder este clúster genético, lo cual nos permitirá relacionar la regulación de la expresión con el papel biológico ejercido en *S. Typhi*.

Otra perspectiva consiste en establecer el mecanismo a través del cual el sistema CRISPR-Cas está involucrado en la regulación de las proteínas de membrana externa OmpC y OmpF. Para esto, es de vital importancia complementar la cepa de *S. Typhi* deficiente en el sistema CRISPR-Cas y realizar un perfil de proteínas de membrana externa, para observar que se restablezca la presencia de las porinas OmpC y OmpF. Posteriormente, es necesario mutar, de manera independiente, cada uno de los promotores correspondientes a los diferentes transcritos presentes en el locus CRISPR-Cas, con la finalidad de identificar la unidad o unidades participantes en la regulación de las proteínas de membrana externa de interés. Una vez determinado lo anterior, es relevante demostrar si dicho componente regula de manera directa a *ompC* u *ompF*, o si lo realiza a través de alguna de las proteínas o microRNAs implicados en la regulación transcripcional y postranscripcional de los genes que codifican para estas proteínas de membrana externa.

Otra alternativa para resolver esta última perspectiva consiste en realizar experimentos de RNA-Seq, microarreglos y proteómica de la cepa de *S. Typhi* silvestre y la  $\Delta cas$ -CRISPR, con los cuales podremos observar todos los elementos genéticos relacionados con el sistema CRISPR-Cas y que estén implicados en la regulación y síntesis de las proteínas de membrana externa OmpC y OmpF. De esta manera podremos elucidar el mecanismo a través del cual este programa genético participa en la regulación de las porinas mencionadas. Adicionalmente, este tipo de enfoques genómicos globales nos proporcionará un panorama general de todos los procesos biológicos en los que puede participar CRISPR-Cas tanto en la vida libre como en la patogénesis del género *Salmonella*, lo cual constituye un objetivo primordial en nuestro grupo de trabajo. Considerando que CRISPR-Cas está implicado en la regulación de las

porinas mayoritarias, es posible que su función en *S. Typhi* corresponda al mantenimiento de la membrana celular, por lo que fenómenos como permeabilidad y resistencia a detergentes o especies reactivas de oxígeno, deben ser evaluados en la cepa de *S. Typhi* silvestre y la carente del locus CRISPR-Cas.

Resolver las perspectivas planteadas proporcionará información relevante para establecer los diferentes mecanismos de regulación utilizados por bacterias, para controlar la actividad del sistema CRISPR-Cas de acuerdo a su ciclo de vida. Además, nos ayudará a esclarecer las diferentes funciones que puede desempeñar CRISPR-Cas en la fisiología bacteriana.

### 13. BIBLIOGRAFÍA

- Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S. & Shinkai, A. (2010). Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J Mol Biol*, **395**: 270-81.
- Al-Attar, S., Westra, E. R., van der Oost, J. & Brouns, S. J. (2011). Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol Chem*, **392**: 277-89.
- Álvarez-Ordóñez, A., Prieto, M., Bernardo, A., Hill, C. & López, M. (2012). The Acid Tolerance Response of *Salmonella* spp.: An adaptive strategy to survive in stressful environments prevailing in foods and the host. *Food Res Int*, **45**: 482-92.
- Arslan, Z., Hermanns, V., Wurm, R., Wagner, R. & Pul, U. (2014). Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system. *Nucleic Acids Res*, **42**: 7884-93.
- Audia, J. P., Webb, C. C. & Foster, J. W. (2001). Breaking through the acid barrier: an orchestrated response to proton stress by enteric bacteria. *Int J Med Microbiol*, **291**: 97-106.
- Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G., Binkowski, A., Phanse, S., Joachimiak, A., Koonin, E. V., Savchenko, A., Emili, A., Greenblatt, J., Edwards, A. M. & Yakunin, A. F. (2011). A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. *Mol Microbiol*, **79**: 484-502.
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. & Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*, **37**: W202-8.
- Baker, S. J., Daniels, C. & Morona, R. (1997). PhoP/Q regulated genes in *Salmonella typhi* identification of melittin sensitive mutants. *Microb Pathog*, **22**: 165-79.
- Baker, S. J., Gunn, J. S. & Morona, R. (1999). The *Salmonella typhi* melittin resistance gene *pqaB* affects intracellular growth in PMA-differentiated U937 cells, polymyxin B resistance and lipopolysaccharide. *Microbiology*, **145 ( Pt 2)**: 367-78.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**: 1709-12.
- Batchelor, E., Walthers, D., Kenney, L. J. & Goulian, M. (2005). The *Escherichia coli* CpxA-CpxR envelope stress response system regulates expression of the porins *ompF* and *ompC*. *J Bacteriol*, **187**: 5723-31.
- Beloglazova, N., Brown, G., Zimmerman, M. D., Proudfoot, M., Makarova, K. S., Kudritska, M., Kochinyan, S., Wang, S., Chruszcz, M., Minor, W., Koonin, E. V., Edwards, A. M., Savchenko, A. & Yakunin, A. F. (2008). A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J Biol Chem*, **283**: 20361-71.

- Bhaya, D., Davison, M. & Barrangou, R.** (2011). CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet*, **45**: 273-97.
- Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D.** (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, **151**: 2551-61.
- Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M. F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K. L. & Davidson, A. R.** (2015). Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature*, **526**: 136-9.
- Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R.** (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature*, **493**: 429-32.
- Breitbart, M. & Rohwer, F.** (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol*, **13**: 278-84.
- Brenner, F. W., Villar, R. G., Angulo, F. J., Tauxe, R. & Swaminathan, B.** (2000). *Salmonella* nomenclature. *J Clin Microbiol*, **38**: 2465-7.
- Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V. & van der Oost, J.** (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, **321**: 960-4.
- Buckle, G. C., Walker, C. L. & Black, R. E.** (2012). Typhoid fever and paratyphoid fever: Systematic review to estimate global morbidity and mortality for 2010. *J Glob Health*, **2**: 010401.
- Cardona-Castro, N. & Sánchez-Jiménez, M.** (2003). Mecanismos de interacción de *Salmonella* con la mucosa intestinal. *Infectio*, **7**: 22-9.
- Castillo-Keller, M., Vuong, P. & Misra, R.** (2006). Novel mechanism of *Escherichia coli* porin regulation. *J Bacteriol*, **188**: 576-86.
- Cirillo, D. M., Valdivia, R. H., Monack, D. M. & Falkow, S.** (1998). Macrophage-dependent induction of the *Salmonella* pathogenicity island 2 type III secretion system and its role in intracellular survival. *Mol Microbiol*, **30**: 175-88.
- Contreras, I., Toro, C. S., Troncoso, G. & Mora, G. C.** (1997). *Salmonella typhi* mutants defective in anaerobic respiration are impaired in their ability to replicate within epithelial cells. *Microbiology*, **143** ( Pt 8): 2665-72.
- Crawford, R. W., Rosales-Reyes, R., Ramirez-Aguilar Mde, L., Chapa-Azuela, O., Alpuche-Aranda, C. & Gunn, J. S.** (2010). Gallstones play a significant role in *Salmonella* spp. gallbladder colonization and carriage. *Proc Natl Acad Sci U S A*, **107**: 4353-8.
- Crump, J. A., Sjolund-Karlsson, M., Gordon, M. A. & Parry, C. M.** (2015). Epidemiology, clinical presentation, laboratory diagnosis, antimicrobial resistance, and antimicrobial management of invasive *Salmonella* infections. *Clin Microbiol Rev*, **28**: 901-37.
- Chander, H., Majumdar, S., Sapru, S. & Rishi, P.** (2004). Reactivity of typhoid patients sera with stress induced 55 kDa phenotype in *Salmonella enterica* serovar Typhi. *Mol Cell Biochem*, **267**: 75-82.

- Chatfield, S. N., Dorman, C. J., Hayward, C. & Dougan, G.** (1991). Role of *ompR*-dependent genes in *Salmonella typhimurium* virulence: mutants deficient in both *ompC* and *ompF* are attenuated *in vivo*. *Infect Immun*, **59**: 449-52.
- Chen, S., Zhang, A., Blyn, L. B. & Storz, G.** (2004). MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J Bacteriol*, **186**: 6689-97.
- Chylinski, K., Le Rhun, A. & Charpentier, E.** (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol*, **10**: 726-37.
- Daigle, F.** (2008). Typhi genes expressed during infection or involved in pathogenesis. *J Infect Dev Ctries*, **2**: 431-7.
- Daigle, F., Graham, J. E. & Curtiss, R., 3rd.** (2001). Identification of *Salmonella* Typhi genes expressed within macrophages by selective capture of transcribed sequences (SCOTS). *Mol Microbiol*, **41**: 1211-22.
- Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K. & Semenova, E.** (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun*, **3**: 945.
- Datsenko, K. A. & Wanner, B. L.** (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, **97**: 6640-5.
- De la Cruz, M. A. & Calva, E.** (2010). The complexities of porin genetic regulation. *J Mol Microbiol Biotechnol*, **18**: 24-36.
- De la Cruz, M. A., Fernandez-Mora, M., Guadarrama, C., Flores-Valdez, M. A., Bustamante, V. H., Vazquez, A. & Calva, E.** (2007). LeuO antagonizes H-NS and StpA-dependent repression in *Salmonella enterica ompS1*. *Mol Microbiol*, **66**: 727-43.
- Deiwick, J., Nikolaus, T., Erdogan, S. & Hensel, M.** (1999). Environmental regulation of *Salmonella* pathogenicity island 2 gene expression. *Mol Microbiol*, **31**: 1759-73.
- Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J. & Charpentier, E.** (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, **471**: 602-7.
- Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D. A., Horvath, P. & Moineau, S.** (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol*, **190**: 1390-400.
- Dillon, S. C., Espinosa, E., Hokamp, K., Ussery, D. W., Casadesus, J. & Dorman, C. J.** (2012). LeuO is a global regulator of gene expression in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol*, **85**: 1072-89.
- Doudna, J. A. & Charpentier, E.** (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**: 1258096.
- Dougan, G. & Baker, S.** (2014). *Salmonella enterica* serovar Typhi and the pathogenesis of typhoid fever. *Annu Rev Microbiol*, **68**: 317-36.
- Du, H., Wang, M., Luo, Z., Ni, B., Wang, F., Meng, Y., Xu, S. & Huang, X.** (2011). Coregulation of gene expression by sigma factors RpoE and RpoS in *Salmonella enterica* serovar Typhi during hyperosmotic stress. *Curr Microbiol*, **62**: 1483-9.



- Dufresne, K., Saulnier-Bellemare, J. & Daigle, F. (2018). Functional analysis of the chaperone-usher fimbrial gene clusters of *Salmonella enterica* serovar Typhi. *Front Cell Infect Microbiol*, **8**: 26.
- Eichelberg, K. & Galán, J. E. (2000). The flagellar sigma factor FliA (sigma(28)) regulates the expression of *Salmonella* genes associated with the centisome 63 type III secretion system. *Infect Immun*, **68**: 2735-43.
- Evans, D. F., Pye, G., Bramley, R., Clark, A. G., Dyson, T. J. & Hardcastle, J. D. (1988). Measurement of gastrointestinal pH profiles in normal ambulant human subjects. *Gut*, **29**: 1035-41.
- Everest, P., Wain, J., Roberts, M., Rook, G. & Dougan, G. (2001). The molecular mechanisms of severe typhoid fever. *Trends Microbiol*, **9**: 316-20.
- Faucher, S. P., Curtiss, R., 3rd & Daigle, F. (2005). Selective capture of *Salmonella enterica* serovar Typhi genes expressed in macrophages that are absent from the *Salmonella enterica* serovar Typhimurium genome. *Infect Immun*, **73**: 5217-21.
- Faucher, S. P., Porwollik, S., Dozois, C. M., McClelland, M. & Daigle, F. (2006). Transcriptome of *Salmonella enterica* serovar Typhi within macrophages revealed through the selective capture of transcribed sequences. *Proc Natl Acad Sci U S A*, **103**: 1906-11.
- Faucher, S. P., Viau, C., Gros, P. P., Daigle, F. & Le Moual, H. (2008). The *prpZ* gene cluster encoding eukaryotic-type Ser/Thr protein kinases and phosphatases is repressed by oxidative stress and involved in *Salmonella enterica* serovar Typhi survival in human macrophages. *FEMS Microbiol Lett*, **281**: 160-6.
- Fernandez-Mora, M., Puente, J. L. & Calva, E. (2004). OmpR and LeuO positively regulate the *Salmonella enterica* serovar Typhi *ompS2* porin gene. *J Bacteriol*, **186**: 2909-20.
- Ferrario, M., Ernsting, B. R., Borst, D. W., Wiese, D. E., 2nd, Blumenthal, R. M. & Matthews, R. G. (1995). The leucine-responsive regulatory protein of *Escherichia coli* negatively regulates transcription of *ompC* and *micF* and positively regulates translation of *ompF*. *J Bacteriol*, **177**: 103-13.
- Figuroa, I. & Verdugo, A. (2005). Mecanismos moleculares de patogenicidad de *Salmonella* sp. *Revista Latinoamericana de Microbiología* **47**: 25-42.
- Fineran, P. C., Gerritzen, M. J., Suarez-Diez, M., Kunne, T., Boekhorst, J., van Hijum, S. A., Staals, R. H. & Brouns, S. J. (2014). Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc Natl Acad Sci U S A*, **111**: E1629-38.
- Fischer, S., Maier, L. K., Stoll, B., Brendel, J., Fischer, E., Pfeiffer, F., Dyall-Smith, M. & Marchfelder, A. (2012). An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. *J Biol Chem*, **287**: 33351-63.
- Flores-Valdez, M. A., Puente, J. L. & Calva, E. (2003). Negative osmoregulation of the *Salmonella ompS1* porin gene independently of OmpR in an *hns* background. *J Bacteriol*, **185**: 6497-506.
- Forest, C., Faucher, S. P., Poirier, K., Houle, S., Dozois, C. M. & Daigle, F. (2007). Contribution of the *stg* fimbrial operon of *Salmonella enterica* serovar Typhi during interaction with human cells. *Infect Immun*, **75**: 5264-71.

- Forest, C. G., Ferraro, E., Sabbagh, S. C. & Daigle, F.** (2010). Intracellular survival of *Salmonella enterica* serovar Typhi in human macrophages is independent of *Salmonella* pathogenicity island (SPI)-2. *Microbiology*, **156**: 3689-98.
- Forst, S., Delgado, J. & Inouye, M.** (1989). Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the *ompF* and *ompC* genes in *Escherichia coli*. *Proc Natl Acad Sci U S A*, **86**: 6052-6.
- Fusco, S., Liguori, R., Limauro, D., Bartolucci, S., She, Q. & Contursi, P.** (2015). Transcriptome analysis of *Sulfolobus solfataricus* infected with two related fuselloviruses reveals novel insights into the regulation of CRISPR-Cas system. *Biochimie*, **118**: 322-32.
- Galán, J. E.** (2016). Typhoid toxin provides a window into typhoid fever and the biology of *Salmonella* Typhi. *Proc Natl Acad Sci U S A*, **113**: 6338-44.
- Gallego-Hernandez, A. L., Hernandez-Lucas, I., De la Cruz, M. A., Olvera, L., Morett, E., Medina-Aparicio, L., Ramirez-Trujillo, J. A., Vazquez, A., Fernandez-Mora, M. & Calva, E.** (2012). Transcriptional regulation of the *assT-dsbL-dsbl* gene cluster in *Salmonella enterica* serovar Typhi IMSS-1 depends on LeuO, H-NS, and specific growth conditions. *J Bacteriol*, **194**: 2254-64.
- Gao, H., Zhang, Y., Han, Y., Yang, L., Liu, X., Guo, Z., Tan, Y., Huang, X., Zhou, D. & Yang, R.** (2011). Phenotypic and transcriptional analysis of the osmotic regulator OmpR in *Yersinia pestis*. *BMC Microbiol*, **11**: 39.
- Gonzalez-Escobedo, G., Marshall, J. M. & Gunn, J. S.** (2011). Chronic and acute infection of the gall bladder by *Salmonella* Typhi: understanding the carrier state. *Nat Rev Microbiol*, **9**: 9-14.
- Goren, M. G., Yosef, I., Auster, O. & Qimron, U.** (2012). Experimental definition of a clustered regularly interspaced short palindromic duplicon in *Escherichia coli*. *J Mol Biol*, **423**: 14-6.
- Graeme-Cook, K. A., May, G., Bremer, E. & Higgins, C. F.** (1989). Osmotic regulation of porin expression: a role for DNA supercoiling. *Mol Microbiol*, **3**: 1287-94.
- Grissa, I., Vergnaud, G. & Pourcel, C.** (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res*, **35**: W52-7.
- Guadarrama, C., Villasenor, T. & Calva, E.** (2014). The subtleties and contrasts of the LeuO regulator in *Salmonella* Typhi: Implications in the immune response. *Front Immunol*, **5**: 581.
- Gunderson, F. F. & Cianciotto, N. P.** (2013). The CRISPR-associated gene *cas2* of *Legionella pneumophila* is required for intracellular infection of amoebae. *MBio*, **4**: e00074-13.
- Gunderson, F. F., Mallama, C. A., Fairbairn, S. G. & Cianciotto, N. P.** (2015). Nuclease activity of *Legionella pneumophila* Cas2 promotes intracellular infection of amoebal host cells. *Infect Immun*, **83**: 1008-18.
- Gunn, J. S., Marshall, J. M., Baker, S., Dongol, S., Charles, R. C. & Ryan, E. T.** (2014). *Salmonella* chronic carriage: epidemiology, diagnosis, and gallbladder persistence. *Trends Microbiol*, **22**: 648-55.

- Guo, L., Lim, K. B., Gunn, J. S., Bainbridge, B., Darveau, R. P., Hackett, M. & Miller, S. I.** (1997). Regulation of lipid A modifications by *Salmonella typhimurium* virulence genes *phoP-phoQ*. *Science*, **276**: 250-3.
- Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E.** (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol*, **1**: e60.
- Hale, C., Kleppe, K., Terns, R. M. & Terns, M. P.** (2008). Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA*, **14**: 2572-9.
- Hansen-Wester, I. & Hensel, M.** (2001). *Salmonella* pathogenicity islands encoding type III secretion systems. *Microbes Infect*, **3**: 549-59.
- Haraga, A., Ohlson, M. B. & Miller, S. I.** (2008). *Salmonellae* interplay with host cells. *Nat Rev Microbiol*, **6**: 53-66.
- Heithoff, D. M., Sinsheimer, R. L., Low, D. A. & Mahan, M. J.** (1999). An essential role for DNA adenine methylation in bacterial virulence. *Science*, **284**: 967-70.
- Hensel, M., Shea, J. E., Waterman, S. R., Mundy, R., Nikolaus, T., Banks, G., Vazquez-Torres, A., Gleeson, C., Fang, F. C. & Holden, D. W.** (1998). Genes encoding putative effector proteins of the type III secretion system of *Salmonella* pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. *Mol Microbiol*, **30**: 163-74.
- Hermant, D., Menard, R., Arricau, N., Parsot, C. & Popoff, M. Y.** (1995). Functional conservation of the *Salmonella* and *Shigella* effectors of entry into epithelial cells. *Mol Microbiol*, **17**: 781-9.
- Hernández-Lucas, I., Gallego-Hernández, A. L., Encarnación, S., Fernández-Mora, M., Martínez-Batallar, A. G., Salgado, H., Oropeza, R. & Calva, E.** (2008). The LysR-type transcriptional regulator LeuO controls expression of several genes in *Salmonella enterica* serovar Typhi. *J Bacteriol*, **190**: 1658-70.
- Hille, F. & Charpentier, E.** (2016). CRISPR-Cas: biology, mechanisms and relevance. *Phil Trans R Soc B*, **371**: 20150496.
- Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S. & Charpentier, E.** (2018). The Biology of CRISPR-Cas: Backward and Forward. *Cell*, **172**: 1239-59.
- Hornick, R. B., Greisman, S. E., Woodward, T. E., DuPont, H. L., Dawkins, A. T. & Snyder, M. J.** (1970). Typhoid fever: pathogenesis and immunologic control. *N Engl J Med*, **283**: 686-91.
- House, D., Bishop, A., Parry, C., Dougan, G. & Wain, J.** (2001). Typhoid fever: pathogenesis and disease. *Curr Opin Infect Dis*, **14**: 573-8.
- Howard, J. A., Delmas, S., Ivančić-Baće, I. & Bolt, E. L.** (2011). Helicase dissociation and annealing of RNA-DNA hybrids by *Escherichia coli* Cas3 protein. *Biochem J*, **439**: 85-95.
- Hryhorowicz, M., Lipinski, D., Zeyland, J. & Slomski, R.** (2017). CRISPR/Cas9 immune system as a tool for genome engineering. *Arch Immunol Ther Exp (Warsz)*, **65**: 233-40.
- Huang, D. B. & DuPont, H. L.** (2005). Problem pathogens: extra-intestinal complications of *Salmonella enterica* serotype Typhi infection. *Lancet Infect Dis*, **5**: 341-8.

- Huang, L., Tsui, P. & Freundlich, M. (1990). Integration host factor is a negative effector of *in vivo* and *in vitro* expression of *ompC* in *Escherichia coli*. *J Bacteriol*, **172**: 5293-8.
- Huang, X., Xu, H., Sun, X., Ohkusu, K., Kawamura, Y. & Ezaki, T. (2007). Genome-Wide scan of the gene expression kinetics of *Salmonella enterica* serovar Typhi during hyperosmotic stress. *Int J Mol Sci*, **8**: 116-35.
- Issenhuth-Jeanjean, S., Roggentin, P., Mikoleit, M., Guibourdenche, M., de Pinna, E., Nair, S., Fields, P. I. & Weill, F. X. (2014). Supplement 2008-2010 (no. 48) to the White-Kauffmann-Le Minor scheme. *Res Microbiol*, **165**: 526-30.
- Ivančić-Baće, I., Cass, S. D., Wearne, S. J. & Bolt, E. L. (2015). Different genome stability proteins underpin primed and naive adaptation in *E. coli* CRISPR-Cas immunity. *Nucleic Acids Res*, **43**: 10821-30.
- Jansen, R., Embden, J. D., Gastra, W. & Schouls, L. M. (2002a). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*, **43**: 1565-75.
- Jansen, R., van Embden, J. D., Gastra, W. & Schouls, L. M. (2002b). Identification of a novel family of sequence repeats among prokaryotes. *OMICS*, **6**: 23-33.
- Jindal, G., Tewari, R. & Praveen, R. (2011). Acid induced outer membrane phenotype in *Salmonella* Typhi-a gene product of *fliC*. *Am J Biomed Sci*, **3**: 23-30.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**: 816-21.
- Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., Beijer, M. R., Barendregt, A., Zhou, K., Snijders, A. P., Dickman, M. J., Doudna, J. A., Boekema, E. J., Heck, A. J., van der Oost, J. & Brouns, S. J. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol*, **18**: 529-36.
- Kaeriyama, M., Machida, K., Kitakaze, A., Wang, H., Lao, Q., Fukamachi, T., Saito, H. & Kobayashi, H. (2006). OmpC and OmpF are required for growth under hyperosmotic stress above pH 8 in *Escherichia coli*. *Lett Appl Microbiol*, **42**: 195-201.
- Kawaji, H., Mizuno, T. & Mizushima, S. (1979). Influence of molecular size and osmolarity of sugars and dextrans on the synthesis of outer membrane proteins O-8 and O-9 of *Escherichia coli* K-12. *J Bacteriol*, **140**: 843-7.
- Keestra-Gounder, A. M., Tsolis, R. M. & Baumler, A. J. (2015). Now you see me, now you don't: the interaction of *Salmonella* with innate immune receptors. *Nat Rev Microbiol*, **13**: 206-16.
- Knodler, L. A. & Steele-Mortimer, O. (2003). Taking possession: biogenesis of the *Salmonella*-containing vacuole. *Traffic*, **4**: 587-99.
- Koonin, E. V., Makarova, K. S. & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol*, **37**: 67-78.
- Kunin, V., Sorek, R. & Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol*, **8**: R61.

- Labrie, S. J., Samson, J. E. & Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nat Rev Microbiol*, **8**: 317-27.
- Lacroix, F. J., Cloeckert, A., Grepinet, O., Pinault, C., Popoff, M. Y., Waxin, H. & Pardon, P. (1996). *Salmonella typhimurium* *acrB*-like gene: identification and role in resistance to biliary salts and detergents and in murine infection. *FEMS Microbiol Lett*, **135**: 161-7.
- Langridge, G. C., Phan, M. D., Turner, D. J., Perkins, T. T., Parts, L., Haase, J., Charles, I., Maskell, D. J., Peters, S. E., Dougan, G., Wain, J., Parkhill, J. & Turner, A. K. (2009). Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res*, **19**: 2308-16.
- Layton, A. N. & Galyov, E. E. (2007). *Salmonella*-induced enteritis: molecular pathogenesis and therapeutic implications. *Expert Rev Mol Med*, **9**: 1-17.
- Le Minor, L. & Popoff, M. Y. (1987). Designation of *Salmonella enterica* sp. nov., nom. rev., as the Type and only species of the genus *Salmonella*: Request for an Opinion. *Int J Syst Bacteriol*, **37**: 465-8.
- Leclerc, G. J., Tartera, C. & Metcalf, E. S. (1998). Environmental regulation of *Salmonella typhi* invasion-defective mutants. *Infect Immun*, **66**: 682-91.
- Leclerc, J. M., Dozois, C. M. & Daigle, F. (2013). Role of the *Salmonella enterica* serovar Typhi Fur regulator and small RNAs RfrA and RfrB in iron homeostasis and interaction with host cells. *Microbiology*, **159**: 591-602.
- Lee, H. Y., Cho, S. A., Lee, I. S., Park, J. H., Seok, S. H., Baek, M. W., Kim, D. J., Lee, S. H., Hur, S. J., Ban, S. J., Lee, Y. K., Han, Y. K. & Cho, Y. K. (2007). Evaluation of *phoP* and *rpoS* mutants of *Salmonella enterica* serovar Typhi as attenuated typhoid vaccine candidates: virulence and protective immune responses in intranasally immunized mice. *FEMS Immunol Med Microbiol*, **51**: 310-8.
- Levy, A., Goren, M. G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U. & Sorek, R. (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*, **520**: 505-10.
- Li, R., Fang, L., Tan, S., Yu, M., Li, X., He, S., Wei, Y., Li, G., Jiang, J. & Wu, M. (2016). Type I CRISPR-Cas targets endogenous genes and regulates virulence to evade mammalian host immunity. *Cell Res*, **26**: 1273-87.
- Lillestol, R. K., Redder, P., Garrett, R. A. & Brugger, K. (2006). A putative viral defence mechanism in archaeal cells. *Archaea*, **2**: 59-72.
- Lin, T. L., Pan, Y. J., Hsieh, P. F., Hsu, C. R., Wu, M. C. & Wang, J. T. (2016). Imipenem represses CRISPR-Cas interference of DNA acquisition through H-NS stimulation in *Klebsiella pneumoniae*. *Sci Rep*, **6**: 31644.
- Liu, T., Li, Y., Wang, X., Ye, Q., Li, H., Liang, Y., She, Q. & Peng, N. (2015). Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR de novo spacer acquisition. *Nucleic Acids Res*, **43**: 1044-55.
- Louwen, R., Horst-Kreft, D., de Boer, A. G., van der Graaf, L., de Knecht, G., Hamersma, M., Heikema, A. P., Timms, A. R., Jacobs, B. C., Wagenaar, J. A., Endtz, H. P., van der Oost, J., Wells, J. M., Nieuwenhuis, E. E., van Vliet, A. H., Willemsen, P. T., van Baarlen, P. & van Belkum, A. (2013). A novel link between *Campylobacter jejuni*

- bacteriophage defence, virulence and Guillain-Barre syndrome. *Eur J Clin Microbiol Infect Dis*, **32**: 207-26.
- Louwen, R., Staals, R. H., Endtz, H. P., van Baarlen, P. & van der Oost, J.** (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol Mol Biol Rev*, **78**: 74-88.
- Majsec, K., Bolt, E. L. & Ivančić-Baće, I.** (2016). Cas3 is a limiting factor for CRISPR-Cas immunity in *Escherichia coli* cells lacking H-NS. *BMC Microbiol*, **16**: 28.
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V.** (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct*, **1**: 7.
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J. & Koonin, E. V.** (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*, **9**: 467-77.
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R. & Koonin, E. V.** (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol*, **13**: 722-36.
- Marraffini, L. A. & Sontheimer, E. J.** (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, **322**: 1843-5.
- Marraffini, L. A. & Sontheimer, E. J.** (2010). Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature*, **463**: 568-71.
- Martinez-Flores, I., Cano, R., Bustamante, V. H., Calva, E. & Puente, J. L.** (1999). The *ompB* operon partially determines differential expression of OmpC in *Salmonella typhi* and *Escherichia coli*. *J Bacteriol*, **181**: 556-62.
- Medeiros, A. A., O'Brien, T. F., Rosenberg, E. Y. & Nikaido, H.** (1987). Loss of OmpC porin in a strain of *Salmonella typhimurium* causes increased resistance to cephalosporins during therapy. *J Infect Dis*, **156**: 751-7.
- Miller, S. I., Kukral, A. M. & Mekalanos, J. J.** (1989). A two-component regulatory system (*phoP-phoQ*) controls *Salmonella typhimurium* virulence. *Proc Natl Acad Sci U S A*, **86**: 5054-8.
- Mizuno, T., Chou, M. Y. & Inouye, M.** (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci U S A*, **81**: 1966-70.
- Mojica, F. J., Diez-Villaseñor, C., Garcia-Martinez, J. & Almendros, C.** (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, **155**: 733-40.
- Mojica, F. J., Diez-Villaseñor, C., Garcia-Martinez, J. & Soria, E.** (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol*, **60**: 174-82.

- Mulepati, S. & Bailey, S.** (2011). Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J Biol Chem*, **286**: 31896-903.
- Mulepati, S. & Bailey, S.** (2013). *In vitro* reconstitution of an *Escherichia coli* RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. *J Biol Chem*, **288**: 22184-92.
- Nishino, K., Latifi, T. & Groisman, E. A.** (2006). Virulence and drug resistance roles of multidrug efflux systems of *Salmonella enterica* serovar Typhimurium. *Mol Microbiol*, **59**: 126-41.
- Nuñez, J. K., Bai, L., Harrington, L. B., Hinder, T. L. & Doudna, J. A.** (2016). CRISPR immunological memory requires a host factor for specificity. *Mol Cell*, **62**: 824-33.
- Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W. & Doudna, J. A.** (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat Struct Mol Biol*, **21**: 528-34.
- Nuñez, J. K., Lee, A. S., Engelman, A. & Doudna, J. A.** (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature*, **519**: 193-8.
- Ofir, G., Melamed, S., Sberro, H., Mukamel, Z., Silverman, S., Yaakov, G., Doron, S. & Sorek, R.** (2018). DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat Microbiol*, **3**: 90-8.
- Oropeza, R., Sampieri, C. L., Puente, J. L. & Calva, E.** (1999). Negative and positive regulation of the non-osmoregulated *ompS1* porin gene in *Salmonella typhi*: a novel regulatory mechanism that involves OmpR. *Mol Microbiol*, **32**: 243-52.
- Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J. C. & Vogel, J.** (2006). SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global *omp* mRNA decay. *Mol Microbiol*, **62**: 1674-88.
- Patterson, A. G., Chang, J. T., Taylor, C. & Fineran, P. C.** (2015). Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res*, **43**: 6038-48.
- Perez-Rodriguez, R., Haitjema, C., Huang, Q., Nam, K. H., Bernardis, S., Ke, A. & DeLisa, M. P.** (2011). Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol Microbiol*, **79**: 584-99.
- Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K. A., Djordjevic, M., Wanner, B. L. & Severinov, K.** (2010). Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol*, **77**: 1367-79.
- Pourcel, C., Salvignol, G. & Vergnaud, G.** (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*, **151**: 653-63.
- Prieto, A. I., Jakomin, M., Segura, I., Pucciarelli, M. G., Ramos-Morales, F., Garcia-Del Portillo, F. & Casadesus, J.** (2007). The GATC-binding protein SeqA is required for bile resistance and virulence in *Salmonella enterica* serovar typhimurium. *J Bacteriol*, **189**: 8496-502.

- Prieto, A. I., Ramos-Morales, F. & Casadesus, J. (2004). Bile-induced DNA damage in *Salmonella enterica*. *Genetics*, **168**: 1787-94.
- Puente, J. L., Flores, V., Fernandez, M., Fuchs, Y. & Calva, E. (1987). Isolation of an *ompC*-like outer membrane protein gene from *Salmonella typhi*. *Gene*, **61**: 75-83.
- Puente, J. L., Verdugo-Rodriguez, A. & Calva, E. (1991). Expression of *Salmonella typhi* and *Escherichia coli* *OmpC* is influenced differently by medium osmolarity; dependence on *Escherichia coli* *OmpR*. *Mol Microbiol*, **5**: 1205-10.
- Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N. & Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol*, **75**: 1495-512.
- Raffatellu, M., Chessa, D., Wilson, R. P., Dusold, R., Rubino, S. & Baumler, A. J. (2005). The Vi capsular antigen of *Salmonella enterica* serotype Typhi reduces Toll-like receptor-dependent interleukin-8 expression in the intestinal mucosa. *Infect Immun*, **73**: 3367-74.
- Raffatellu, M., Wilson, R. P., Winter, S. E. & Baumler, A. J. (2008). Clinical pathogenesis of typhoid fever. *J Infect Dev Ctries*, **2**: 260-6.
- Ramos, H. C., Rumbo, M. & Sirard, J. C. (2004). Bacterial flagellins: mediators of pathogenicity and host immune responses in mucosa. *Trends Microbiol*, **12**: 509-17.
- Ratner, H. K., Sampson, T. R. & Weiss, D. S. (2015). I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. *Curr Opin Infect Dis*, **28**: 267-74.
- Rauch, B. J., Silvis, M. R., Hultquist, J. F., Waters, C. S., McGregor, M. J., Krogan, N. J. & Bondy-Denomy, J. (2017). Inhibition of CRISPR-Cas9 with bacteriophage proteins. *Cell*, **168**: 150-8.e10.
- Rychlik, I. & Barrow, P. A. (2005). *Salmonella* stress management and its relevance to behaviour during intestinal colonisation and infection. *FEMS Microbiol Rev*, **29**: 1021-40.
- Sabbagh, S. C., Lepage, C., McClelland, M. & Daigle, F. (2012). Selection of *Salmonella enterica* serovar Typhi genes involved during interaction with human macrophages by screening of a transposon mutant library. *PLoS One*, **7**: e36643.
- Sampson, T. R., Napier, B. A., Schroeder, M. R., Louwen, R., Zhao, J., Chin, C. Y., Ratner, H. K., Llewellyn, A. C., Jones, C. L., Laroui, H., Merlin, D., Zhou, P., Endtz, H. P. & Weiss, D. S. (2014). A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. *Proc Natl Acad Sci U S A*, **111**: 11163-8.
- Sampson, T. R., Saroj, S. D., Llewellyn, A. C., Tzeng, Y. L. & Weiss, D. S. (2013). A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature*, **497**: 254-7.
- Semenova, E., Jore, M. M., Datsenko, K. A., Semenova, A., Westra, E. R., Wanner, B., van der Oost, J., Brouns, S. J. & Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci U S A*, **108**: 10098-103.
- Serbanescu, M. A., Cordova, M., Krastel, K., Flick, R., Beloglazova, N., Latos, A., Yakunin, A. F., Senadheera, D. B. & Cvitkovitch, D. G. (2015). Role of the



- Streptococcus mutans* CRISPR-Cas systems in immunity and cell physiology. *J Bacteriol*, **197**: 749-61.
- Shariat, N., Timme, R. E., Pettengill, J. B., Barrangou, R. & Dudley, E. G.** (2015). Characterization and evolution of *Salmonella* CRISPR-Cas systems. *Microbiology*, **161**: 374-86.
- Sharma, A. & Qadri, A.** (2004). Vi polysaccharide of *Salmonella typhi* targets the prohibitin family of molecules in intestinal epithelial cells and suppresses early inflammatory responses. *Proc Natl Acad Sci U S A*, **101**: 17492-7.
- Sheikh, A., Charles, R. C., Sharmeen, N., Rollins, S. M., Harris, J. B., Bhuiyan, M. S., Arifuzzaman, M., Khanam, F., Bukka, A., Kalsy, A., Porwollik, S., Leung, D. T., Brooks, W. A., LaRocque, R. C., Hohmann, E. L., Cravioto, A., Logvinenko, T., Calderwood, S. B., McClelland, M., Graham, J. E., Qadri, F. & Ryan, E. T.** (2011). *In vivo* expression of *Salmonella enterica* serotype Typhi genes in the blood of patients with typhoid fever in Bangladesh. *PLoS Negl Trop Dis*, **5**: e1419.
- Simons, R. W., Houtman, F. & Kleckner, N.** (1987). Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene*, **53**: 85-96.
- Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. & Siksnys, V.** (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J*, **30**: 1335-42.
- Sinkunas, T., Gasiunas, G., Waghmare, S. P., Dickman, M. J., Barrangou, R., Horvath, P. & Siksnys, V.** (2013). *In vitro* reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J*, **32**: 385-94.
- Sinnott, C. R. & Teall, A. J.** (1987). Persistent gallbladder carriage of *Salmonella typhi*. *Lancet*, **1**: 976.
- Skennerton, C. T., Angly, F. E., Breitbart, M., Bragg, L., He, S., McMahon, K. D., Hugenholtz, P. & Tyson, G. W.** (2011). Phage encoded H-NS: a potential achilles heel in the bacterial defence system. *PLoS One*, **6**: e20095.
- Sorek, R., Lawrence, C. M. & Wiedenheft, B.** (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem*, **82**: 237-66.
- Spano, S.** (2016). Mechanisms of *Salmonella Typhi* host restriction. *Adv Exp Med Biol*, **915**: 283-94.
- Spano, S., Ugalde, J. E. & Galán, J. E.** (2008). Delivery of a *Salmonella Typhi* exotoxin from a host intracellular compartment. *Cell Host Microbe*, **3**: 30-8.
- Spector, M. P. & Kenyon, W. J.** (2012). Resistance and survival strategies of *Salmonella enterica* to environmental stresses. *Food Res Int*, **45**: 455-81.
- Suzuki, T., Ueguchi, C. & Mizuno, T.** (1996). H-NS regulates OmpF expression through *micF* antisense RNA in *Escherichia coli*. *J Bacteriol*, **178**: 3650-3.
- Swarts, D. C., Mosterd, C., van Passel, M. W. & Brouns, S. J.** (2012). CRISPR interference directs strand specific spacer acquisition. *PLoS One*, **7**: e35888.
- Torres-Ruiz, R. & Rodriguez-Perales, S.** (2017). CRISPR-Cas9 technology: applications and human disease modelling. *Brief Funct Genomics*, **16**: 4-12.

- Touchon, M. & Rocha, E. P.** (2010). The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS One*, **5**: e11126.
- Townsend, S. M., Kramer, N. E., Edwards, R., Baker, S., Hamlin, N., Simmonds, M., Stevens, K., Maloy, S., Parkhill, J., Dougan, G. & Baumler, A. J.** (2001). *Salmonella enterica* serovar Typhi possesses a unique repertoire of fimbrial gene sequences. *Infect Immun*, **69**: 2894-901.
- Vala, S., Shah, U., Ahmad, S. A., Scolnik, D. & Glatstein, M.** (2016). Resistance patterns of Typhoid fever in children: a longitudinal community-based study. *Am J Ther*, **23**: e1151-4.
- van Velkinburgh, J. C. & Gunn, J. S.** (1999). PhoP-PhoQ-regulated loci are required for enhanced bile resistance in *Salmonella* spp. *Infect Immun*, **67**: 1614-22.
- Veesenmeyer, J. L., Andersen, A. W., Lu, X., Hussa, E. A., Murfin, K. E., Chaston, J. M., Dillman, A. R., Wassarman, K. M., Sternberg, P. W. & Goodrich-Blair, H.** (2014). NiID CRISPR RNA contributes to *Xenorhabdus nematophila* colonization of symbiotic host nematodes. *Mol Microbiol*, **93**: 1026-42.
- Villarreal, J. M., Becerra-Lobato, N., Rebollar-Flores, J. E., Medina-Aparicio, L., Carbajal-Gomez, E., Zavala-Garcia, M. L., Vazquez, A., Gutierrez-Rios, R. M., Olvera, L., Encarnacion, S., Martinez-Batallar, A. G., Calva, E. & Hernandez-Lucas, I.** (2014). The *Salmonella enterica* serovar Typhi *ltrR-ompR-ompC-ompF* genes are involved in resistance to the bile salt sodium deoxycholate and in bacterial transformation. *Mol Microbiol*, **92**: 1005-24.
- Viswanathan, P., Murphy, K., Julien, B., Garza, A. G. & Kroos, L.** (2007). Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J Bacteriol*, **189**: 3738-50.
- Walawalkar, Y. D., Vaidya, Y. & Nayak, V.** (2016). Response of *Salmonella* Typhi to bile-generated oxidative stress: implication of quorum sensing and persister cell populations. *Pathog Dis*, **74**: ftw090.
- Wallace, R. A., Black, W. P., Yang, X. & Yang, Z.** (2014). A CRISPR with roles in *Myxococcus xanthus* development and exopolysaccharide production. *J Bacteriol*, **196**: 4036-43.
- Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M. & Wang, Y.** (2015). Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. *Cell*, **163**: 840-53.
- Waterman, S. R. & Holden, D. W.** (2003). Functions and effectors of the *Salmonella* pathogenicity island 2 type III secretion system. *Cell Microbiol*, **5**: 501-11.
- Weinstein, D. L., O'Neill, B. L., Hone, D. M. & Metcalf, E. S.** (1998). Differential early interactions between *Salmonella enterica* serovar Typhi and two other pathogenic *Salmonella* serovars with intestinal epithelial cells. *Infect Immun*, **66**: 2310-8.
- Westra, E. R., Pul, U., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M., Van Heereveld, L., Mastop, M., Wagner, E. G., Schnetz, K., Van Der Oost, J., Wagner, R. & Brouns, S. J.** (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol Microbiol*, **77**: 1380-93.

- Westra, E. R., Semenova, E., Datsenko, K. A., Jackson, R. N., Wiedenheft, B., Severinov, K. & Brouns, S. J. (2013). Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet*, **9**: e1003742.
- Westra, E. R., van Erp, P. B., Kunne, T., Wong, S. P., Staals, R. H., Seegers, C. L., Bollen, S., Jore, M. M., Semenova, E., Severinov, K., de Vos, W. M., Dame, R. T., de Vries, R., Brouns, S. J. & van der Oost, J. (2012). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol Cell*, **46**: 595-605.
- Wiedenheft, B., Lander, G. C., Zhou, K., Jore, M. M., Brouns, S. J. J., van der Oost, J., Doudna, J. A. & Nogales, E. (2011). Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature*, **477**: 486-9.
- Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W. & Doudna, J. A. (2009). Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure*, **17**: 904-12.
- Wilson, R. P., Raffatellu, M., Chessa, D., Winter, S. E., Tukel, C. & Baumler, A. J. (2008). The Vi-capsule prevents Toll-like receptor 4 recognition of *Salmonella*. *Cell Microbiol*, **10**: 876-90.
- Wilson, R. P., Winter, S. E., Spees, A. M., Winter, M. G., Nishimori, J. H., Sanchez, J. F., Nuccio, S. P., Crawford, R. W., Tukel, C. & Baumler, A. J. (2011). The Vi capsular polysaccharide prevents complement receptor 3-mediated clearance of *Salmonella enterica* serotype Typhi. *Infect Immun*, **79**: 830-7.
- Yang, C. D., Chen, Y. H., Huang, H. Y., Huang, H. D. & Tseng, C. P. (2014). CRP represses the CRISPR/Cas system in *Escherichia coli*: evidence that endogenous CRISPR spacers impede phage P1 replication. *Mol Microbiol*, **92**: 1072-91.
- Yoganand, K. N., Sivathanu, R., Nimkar, S. & Anand, B. (2017). Asymmetric positioning of Cas1-2 complex and Integration Host Factor induced DNA bending guide the unidirectional homing of protospacer in CRISPR-Cas type I-E system. *Nucleic Acids Res*, **45**: 367-81.
- Yosef, I., Goren, M. G. & Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res*, **40**: 5569-76.
- Young, J. C., Dill, B. D., Pan, C., Hettich, R. L., Banfield, J. F., Shah, M., Fremaux, C., Horvath, P., Barrangou, R. & Verberkmoes, N. C. (2012). Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS One*, **7**: e38077.
- Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H. & O'Toole, G. A. (2009). Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J Bacteriol*, **191**: 210-9.
- Zhang, H., Jia, Y., Xie, X., Wang, M., Zheng, Y., Xu, S., Zhang, W., Wang, Q., Huang, X. & Du, H. (2016). RpoE promotes invasion and intracellular survival by regulating SPI-1 and SPI-2 in *Salmonella enterica* serovar Typhi. *Future Microbiol*, **11**: 1011-24.
- Zhang, L., Rozek, A. & Hancock, R. E. (2001). Interaction of cationic antimicrobial peptides with model membranes. *J Biol Chem*, **276**: 35714-22.

- 
- Zhang, S., Kingsley, R. A., Santos, R. L., Andrews-Polymenis, H., Raffatellu, M., Figueiredo, J., Nunes, J., Tsolis, R. M., Adams, L. G. & Baumler, A. J. (2003). Molecular pathogenesis of *Salmonella enterica* serotype typhimurium-induced diarrhea. *Infect Immun*, **71**: 1-12.
- Zhang, S., Santos, R. L., Tsolis, R. M., Stender, S., Hardt, W. D., Baumler, A. J. & Adams, L. G. (2002). The *Salmonella enterica* serotype Typhimurium effector proteins SipA, SopA, SopB, SopD, and SopE2 act in concert to induce diarrhea in calves. *Infect Immun*, **70**: 3843-55.
- Zhang, X. L., Tsui, I. S., Yip, C. M., Fung, A. W., Wong, D. K., Dai, X., Yang, Y., Hackett, J. & Morris, C. (2000). *Salmonella enterica* serovar Typhi uses type IVB pili to enter human intestinal epithelial cells. *Infect Immun*, **68**: 3067-73.
- Zhou, D. & Galán, J. (2001). *Salmonella* entry into host cells: the work in concert of type III secreted effector proteins. *Microbes Infect*, **3**: 1293-8.
- Zhu, Y., Zhang, F. & Huang, Z. (2018). Structural insights into the inactivation of CRISPR-Cas systems by diverse anti-CRISPR proteins. *BMC Biol*, **16**: 32.

## ANEXO I

Tabla 1. Cepas bacterianas utilizadas en este estudio

Cepa	Descripción	Referencia
<b><i>E. coli</i></b> BL21(DE3)	F <sup>-</sup> <i>ompT gal (dcm) (lon) hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) met(DE3)</i> .	Novagen
DH5α	Φ80d/ <i>lacZ</i> ΔM15 Δ( <i>lacZYA-argF</i> ) U169 <i>recA1 endA1 hsdR17 (r<sub>k</sub><sup>-</sup>m<sub>k</sub><sup>+</sup>) phoA supE44 λ-thi-1 gyrA96 relA1</i> NaI <sup>r</sup> .	Gibco BRL
<b><i>S. Typhi</i></b> IMSS-1	<i>Salmonella enterica</i> serovar Typhi 9.12, d, serotipo Vi. Aislado clínico mexicano.	(Puente <i>et al.</i> , 1987)
IMSS-1 Δ <i>arcA</i>	<i>S. Typhi</i> IMSS-1 Δ <i>arcA</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>cas</i> -CRISPR	<i>S. Typhi</i> IMSS-1 Δ <i>cas3-cse1-cse2-cas7-cas5-cas6e-cas2-cas1</i> -CRISPR	Este trabajo
IMSS-1 Δ <i>crp</i>	<i>S. Typhi</i> IMSS-1 Δ <i>crp</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>cpxRA</i>	<i>S. Typhi</i> IMSS-1 Δ <i>cpxR-cpxA</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>fnr</i>	<i>S. Typhi</i> IMSS-1 Δ <i>fnr</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>ihf</i>	<i>S. Typhi</i> IMSS-1 Δ <i>himA</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>lrp</i>	<i>S. Typhi</i> IMSS-1 Δ <i>lrp</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>lrp/hns</i>	<i>S. Typhi</i> IMSS-1 Δ <i>lrp, hns99::</i> Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>narP</i>	<i>S. Typhi</i> IMSS-1 Δ <i>narP</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>ompR</i>	<i>S. Typhi</i> IMSS-1 Δ <i>ompR</i>	(Oropeza <i>et al.</i> , 1999)
IMSS-1 Δ <i>phoP</i>	<i>S. Typhi</i> IMSS-1 Δ <i>phoP</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>pmrAB</i>	<i>S. Typhi</i> IMSS-1 Δ <i>pmrA-pmrB</i> , Km <sup>r</sup>	Este trabajo
IMSS-1 Δ <i>ssrB</i>	<i>S. Typhi</i> IMSS-1 Δ <i>ssrB</i> , Km <sup>r</sup>	Este trabajo
IMSS-II	IMSS-1 conteniendo el plásmido pFM <i>Trc12</i>	(Fernandez-Mora <i>et al.</i> , 2004)
IMSS-III	IMSS-1 con el plásmido pFM <i>TrcleuO-50</i>	(De la Cruz <i>et al.</i> , 2007)

IMSS-30	IMSS-1 $\Delta leuO$ , Km <sup>r</sup>	(Hernández-Lucas <i>et al.</i> , 2008)
STY $hns99$	IMSS-I $hns99::Km^r mariner$	(Flores-Valdez <i>et al.</i> , 2003)

**Tabla 2. Plásmidos utilizados en este estudio**

Plásmido	Descripción	Referencia
pFM $Trc12$	Derivado del p $Trc99A$ con un origen de replicación p15A1, Ap <sup>r</sup> .	(Fernandez-Mora <i>et al.</i> , 2004)
pFM $TrcleuO-50$	Derivado del pFM $Trc12$ conteniendo el gen $leuO$ de <i>S. enterica</i> Typhi bajo un promotor inducible por IPTG P $trc$ , Ap <sup>r</sup> .	(Fernandez-Mora <i>et al.</i> , 2004)
pKD4	Derivado de pANTS $\gamma$ conteniendo secuencias FRT adyacentes al gen de resistencia a kanamicina; Km <sup>r</sup> .	(Datsenko & Wanner, 2000)
pKD46	Sistema de recombinasa $\lambda$ Red bajo el promotor inducible por arabinosa, Ap <sup>r</sup> .	(Datsenko & Wanner, 2000)
pKK232-8	Derivado del pBR322 con el gen cloranfenicol acetiltransferasa ( $cat$ ) sin promotor, Ap <sup>r</sup> .	Pharmacia LKB Biotechnology
pKK232-9	Derivado del pKK232-8, con el gen de kanamicina obtenido del pUC4K, clonado Pst I, Km <sup>r</sup> .	(Hernández-Lucas <i>et al.</i> , 2008)
pKK8 $ascas2-1$ (-5556/-6245)	pKK232-8 conteniendo los nucleótidos -5556 a -6245 respecto al ATG de $cas3$ . Esta región forma parte de la cadena antisentido de $cas2-cas1$ e incluye 292 pb de $cas2$ , su región intergenica (2 bp) y 395 pb de $cas1$ , clonado 5' $BamH$ I-3' $Kpn$ I; Ap <sup>r</sup> .	Este trabajo
pKK8 $ascas2-1$ (-5355/-6245)	pKK232-8 conteniendo los nucleótidos -5355 a -6245 respecto al ATG de $cas3$ . Esta región forma parte de la cadena antisentido de $cas2-cas1$ e incluye 292 pb de $cas2$ , su región intergenica (2 bp) y 596 pb de $cas1$ , clonado 5' $BamH$ I-3' $Kpn$ I; Ap <sup>r</sup> .	Este trabajo

pKK8 <i>cas1</i> (-194/+8)	pKK232-8 conteniendo 194 pb corriente arriba y 8 pb corriente abajo del ATG de <i>cas1</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas2</i> (-212/+10)	pKK232-8 conteniendo 212 pb corriente arriba y 10 pb corriente abajo del ATG de <i>cas2</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-455/+112)	pKK232-8 conteniendo 455 pb corriente arriba y 112 pb corriente abajo del ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-241/+112)	pKK232-8 conteniendo 241 pb corriente arriba y 112 pb corriente abajo del ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-180/+112)	pKK232-8 conteniendo 180 pb corriente arriba y 112 pb corriente abajo del ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-83/+112)	pKK232-8 conteniendo 83 pb corriente arriba y 112 pb corriente abajo del ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-83/+38)	pKK232-8 conteniendo 83 pb corriente arriba y 38 pb corriente abajo del ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-83/-1)	pKK232-8 conteniendo los nucleótidos -83 a -1 respecto al ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-83/-25)	pKK232-8 conteniendo los nucleótidos -83 a -25 respecto al ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-83/-46)	pKK232-8 conteniendo los nucleótidos -83 a -46 respecto al ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas3</i> (-70/-25)	pKK232-8 conteniendo los nucleótidos -70 a -25 respecto al ATG de <i>cas3</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas5</i> (-952/+280)	pKK232-8 conteniendo 952 pb corriente arriba y 280 pb corriente abajo del ATG de <i>cas5</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cas6e</i> (-291/+6)	pKK232-8 conteniendo 291 pb corriente arriba y 6 pb corriente abajo del ATG de <i>cas6e</i> , clonado 5' <i>BamH</i> I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo

pKK8 <i>cas7</i> (-470/+132)	pKK232-8 conteniendo 470 pb corriente arriba y 132 pb corriente abajo del ATG de <i>cas7</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 CRISPR (-231/+503)	pKK232-8 conteniendo 231 pb corriente arriba y 503 pb corriente abajo respecto a la primer base de la secuencia líder, clonado 5' <i>Bam</i> H I-3' <i>Xho</i> I, Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse1</i> (-375/+191)	pKK232-8 conteniendo 375 pb corriente arriba y 191 pb corriente abajo del inicio de transcripción de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse1</i> (-158/+191)	pKK232-8 conteniendo 158 pb corriente arriba y 191 pb corriente abajo del inicio de transcripción de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse1</i> (-91/+191)	pKK232-8 conteniendo 91 pb corriente arriba y 191 pb corriente abajo del inicio de transcripción de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse1</i> (-51/+191)	pKK232-8 conteniendo 51 pb corriente arriba y 191 pb corriente abajo del inicio de transcripción de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse1</i> (-51/+99)	pKK232-8 conteniendo 51 pb corriente arriba y 99 pb corriente abajo del inicio de transcripción de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse1</i> (-51/+49)	pKK232-8 conteniendo 51 pb corriente arriba y 49 pb corriente abajo del inicio de transcripción de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse1</i> (-56/+1140)	pKK232-8 conteniendo 56 pb corriente arriba y 1140 pb corriente abajo del ATG de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Hind</i> III; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse2</i> (-450/+181)	pKK232-8 conteniendo 450 pb corriente arriba y 181 pb corriente abajo del ATG de <i>cse2</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse2</i> (-62/+181)	pKK232-8 conteniendo 62 pb corriente arriba y 181 pb corriente abajo del ATG de <i>cse2</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse2</i> (+134/+181)	pKK232-8 conteniendo los nucleótidos +134 a +181 respecto al ATG de <i>cse2</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse2</i> (+142/+181)	pKK232-8 conteniendo los nucleótidos +142 a +181 respecto al ATG de <i>cse2</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo



pKK8 <i>cse2</i> (+142/+202)	pKK232-8 conteniendo los nucleótidos +142 a +202 respecto al ATG de <i>cse2</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse2</i> (+142/+267)	pKK232-8 conteniendo los nucleótidos +142 a +267 respecto al ATG de <i>cse2</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>cse2</i> (+154/+181)	pKK232-8 conteniendo los nucleótidos +154 a +181 respecto al ATG de <i>cse2</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I; Ap <sup>r</sup> .	Este trabajo
pKK8 <i>ompF</i> (-782/+184)	pKK232-8 conteniendo 782 pb corriente arriba y 184 pb corriente abajo del ATG de <i>ompF</i> , Ap <sup>r</sup> .	(Villarreal <i>et al.</i> , 2014)
pKK9 <i>cse1</i>	pKK232-9 conteniendo 375 pb corriente arriba y 191 pb corriente abajo del inicio de transcripción de <i>cse1</i> , clonado 5' <i>Bam</i> H I-3' <i>Kpn</i> I, Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1</i> (-10 ATA→GGG)	Derivado del pKK9- <i>cse1</i> , con una sustitución (ATA→GGG) en los nucleótidos -9 a -11 respecto al inicio de transcripción de <i>cse1</i> , Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1</i> -A	Derivado del pKK9- <i>cse1</i> , con una sustitución (TTT→CCC) en los nucleótidos -115 a -117 respecto al inicio de transcripción de <i>cse1</i> , Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1</i> -B	Derivado del pKK9- <i>cse1</i> , con una sustitución (TTTT→CCCC) en los nucleótidos -114 a -117 respecto al inicio de transcripción de <i>cse1</i> , Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1</i> -C	Derivado del pKK9- <i>cse1</i> , con una sustitución (TAT→CGC) en los nucleótidos -103 a -105 respecto al inicio de transcripción de <i>cse1</i> , Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1</i> -D	Derivado del pKK9- <i>cse1</i> con una sustitución (ATA→GCG) en los nucleótidos -88 a -90 respecto al inicio de transcripción de <i>cse1</i> , Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1</i> -E	Derivado del pKK9- <i>cse1</i> , con una sustitución (AAA→GGG) en los nucleótidos -77 a -79 respecto al inicio de transcripción de <i>cse1</i> , Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1</i> -F	Derivado del pKK9- <i>cse1</i> , con una sustitución (TAA→CGG) en los nucleótidos -60 a -62 respecto al inicio de transcripción de <i>cse1</i> , Km <sup>r</sup> .	Este trabajo
pKK9 <i>ompC</i> (-772/+27)	pKK232-9 conteniendo 772 pb corriente arriba y 27 pb corriente abajo del ATG de <i>ompC</i> , Km <sup>r</sup> .	(Villarreal <i>et al.</i> , 2014)

pKK9 RR- <i>cse1</i> -CRISPR	pKK232-9 conteniendo 347 pb corriente arriba y 6382 pb corriente abajo del inicio de traducción de <i>cse1</i> . Esta región abarca el promotor <i>cse1</i> , los genes <i>cse1-cse2-cas7-cas5-cas6e-cas1-cas2</i> y la secuencia CRISPR, clonado 5' <i>Bam</i> H I-3' <i>Xho</i> I, Km <sup>r</sup> .	Este trabajo
pKK9 RR- <i>cse1-cas2</i>	pKK232-9 conteniendo 347 pb corriente arriba y 5808 pb corriente abajo del inicio de traducción de <i>cse1</i> . Esta región abarca el promotor <i>cse1</i> y los genes <i>cse1-cse2-cas7-cas5-cas6e-cas1-cas2</i> , clonado 5' <i>Bam</i> H I-3' <i>Xho</i> I, Km <sup>r</sup> .	Este trabajo
pKK9 <i>cse1-cas2</i>	pKK232-9 conteniendo 56 pb corriente arriba y 5808 pb corriente abajo del inicio de traducción de <i>cse1</i> . Esta región abarca los genes <i>cse1-cse2-cas7-cas5-cas6e-cas1-cas2</i> , clonado 5' <i>Bam</i> H I-3' <i>Xho</i> I, Km <sup>r</sup> .	Este trabajo
pMDH-NS	pBAD- <i>Myc</i> -His <sub>6</sub> albergando el gen <i>hns</i> de <i>S. Typhi</i> fusionado a <i>Myc</i> -His <sub>6</sub> bajo un promotor inducible por arabinosa, Tc <sup>r</sup> .	(De la Cruz <i>et al.</i> , 2007)
pRS415	Derivado del pNK678 conteniendo los genes <i>lacZ-lacY-lacA</i> sin el promotor del operón <i>lac</i> , Ap <sup>r</sup> .	(Simons <i>et al.</i> , 1987)
pRS <i>ascse2-1</i> (-2090/-1460)	pRS415 conteniendo los nucleótidos -1460 a -2090 respecto al ATG de <i>cas3</i> . Esta región forma parte de la cadena antisentido de <i>cse2-cse1</i> e incluye 181 pb de <i>cse2</i> , su región intergenica (16 bp) y 434 pb de <i>cse1</i> , clonado <i>Bam</i> H I, Ap <sup>r</sup> .	Este trabajo
pT6-LRP-His <sub>6</sub>	pMPM-T6Ω albergando el gen <i>lrp</i> de <i>S. Typhi</i> fusionado a His <sub>6</sub> bajo un promotor inducible por arabinosa, Tc <sup>r</sup> .	Este trabajo

**Tabla 3. Oligonucleótidos utilizados en este estudio**

Oligonucleótido	Secuencia (5'→3')*
16S-F	GCCCAGATGGGATTAGCTAAGT
16S-R	CGGTTATTAACCACAACACC
<i>ascas1-5355 Kpn</i> I-R	<u>GGGGTACCATGAAGATTTGCGGCTGAAGGT</u> CG
<i>ascas1-5556 Kpn</i> I-R	<u>GGGGTACCTCATTAAATCAGTGTATCAGCT</u> CG
<i>ascas1-5775 F</i>	CGGACATCACGATCGGGCTCGG

---



---

<i>ascas1-5858 F</i>	<u>CCCCTGCCGCGAGAACATCTTC</u>
<i>ascas2-6245 BamH I-F</i>	<u>CGGGATCCATTGATTTTCAACAGGTAGAAACG</u>
<i>ascse1-34 BamH I-R</i>	<u>CGGGATCCCGATAATAAATTGGGATACGGG</u>
<i>ascse2-1958 Kpn I-F</i>	<u>GGGGTACCCATCGGGGTGGAGAATTCTG</u>
<i>ascse2-2090 BamH I-F</i>	<u>CGGGATCCATATCATTACGCCATATGGC</u>
<i>cas1-194 BamH I-F</i>	<u>CGGGATCCGGATGCGTGCAAGCGTATCAGCAACATCG</u>
<i>cas1+8 Kpn I-R</i>	<u>GGGGTACCCAGCTCATCAGCGTTTCCTTTTTACC</u>
<i>cas2-212 BamH I-F</i>	<u>CGGGATCCCGACGTAATCCTGCCGAGCCCGATCG</u>
<i>cas2+10 Kpn I-R</i>	<u>GGGGTACCCGATACTCATCGTTAGGTACTCCGATGACC</u>
<i>cas2+177 R</i>	CGCCATAACAACATTTCCCTC
<i>cas2+213 Xho I-R</i>	<u>CCGCTCGAGCGGCTGAAACTCAAACCTGATTCTGTA</u>
<i>cas2+282 Kpn I-F</i>	<u>GGGGTACCTGAAAATCAATAAGTTAGAGATC</u>
<i>cas3-455 BamH I-F</i>	<u>CGGGATCCTCCTGAATGCGGTTATCCAG</u>
<i>cas3-241 BamH I-F</i>	<u>CGGGATCCGAAGACGAAAATATAAATAAAAAG</u>
<i>cas3-180 BamH I-F</i>	<u>CGGGATCCAGTTATCTGTTATCTTTAATACG</u>
<i>cas3-83 BamH I-F</i>	<u>CGGGATCCCTAATCCAGACAAAAAACTGG</u>
<i>cas3 -83/-46 B/K-F</i>	<b>GATCC</b> TAATCCAGACAAAAAACTGGCCAATTACCCGTATCCC <b>GGTAC</b>
<i>cas3 -83/-46 B/K-R</i>	CGGGATACGGGTAATTGGCCAGTTTTTTTGTCTGGATTAG
<i>cas3 -83/-25 B/K-F</i>	<b>GATCC</b> TAATCCAGACAAAAAACTGGCCAATTACCCGTATCCCAATTTATTATCGATTTCAATT <b>GGTAC</b>
<i>cas3 -83/-25 B/K-R</i>	CCAATGAAATCGATAATAAATTGGGATACGGGTAATTGGCCAGTTT TTTTGTCTGGATTAG
<i>cas3 -83/-1 B/K-F</i>	<b>GATCC</b> TAATCCAGACAAAAAACTGGCCAATTACCCGTATCCCAATTTATTATCGATTTCAATTGTAATCCCCCTCCTACAGAAAAAAG <b>GGTAC</b>
<i>cas3 -83/-1 B/K-R</i>	CCTTTTTTCTGTAGGAGGGGATTACAATGAAATCGATAATAAATTGG GATACGGGTAATTGGCCAGTTTTTTTGTCTGGATTAG
<i>cas3 -70/-25 B/K-F</i>	<b>GATCC</b> AAAACTGGCCAATTACCCGTATCCCAATTTATTATCGATTTCAATT <b>GGTAC</b>
<i>cas3 -70/-25 B/K-R</i>	CCAATGAAATCGATAATAAATTGGGATACGGGTAATTGGCCAGTTT <b>G</b>
<i>cas3+38 Kpn I-R</i>	<u>GGGGTACCCAATAGCAGAGGTAGTCCAG</u>
<i>cas3+112 Kpn I-R</i>	<u>GGGGTACCTACCTACCTATAGGTAGATTCC</u>
<i>cas5-952 BamH I-F</i>	<u>CGGGATCCATCGTCGCAAAGTCTGAAGC</u>

---



---

---



---

<i>cas5+280 Kpn I-R</i>	<u>GGGGTACCTGTAATAACGATACTTTTTG</u>
<i>cas6e-291 BamHI-F</i>	<u>CGGGATCCGCATTTTCCGCTCTATCTGGGGAG</u>
<i>cas6e+6 Kpn I-R</i>	<u>GGGGTACCGATACATTATGTTCTCTCCTGCAACGG</u>
<i>cas6e+220 R</i>	GTTTCAATAAGAAACAGGGAG
<i>cas7-470 BamHI-F</i>	<u>CGGGATCCAAGAATGGCGCCGCCATATG</u>
<i>cas7+132 Kpn I-R</i>	<u>GGGGTACCTTGCGACGATACGCGTAAAC</u>
<i>cas7+251 BamHI-F</i>	<u>CGGGATCCGAATTGAAGAAAAATTGCTAAGAG</u>
<i>cas7+286 R</i>	ACTCAGACGAACTCTTAGCAAT
<i>cas7+398 Xho I-R</i>	<u>CCGCTCGAGCGGTTAATCAAATAATTTTCATGGTTAC</u>
<i>cse1-375 BamHI-F</i>	<u>CGGGATCCAACGTCCAGGCAATGGTATG</u>
<i>cse1 (-10 ATA→GGG) F</i>	CGTCTTCGAGAAAATGTAGTGGGACTCCGATGACAGTATTTAAG
<i>cse1 (-10 ATA→GGG) R</i>	CTTAAATACTGTCATCGGAGTCCCCTACTACATTTTCTCGAAGACG
<i>cse1-347 BamHI-F</i>	<u>CGGGATCCCTAGGAGGGGGATTACAATGAAATC</u>
<i>cse1-158 BamHI-F</i>	<u>CGGGATCCACGCTGCGCTATTAAGACTG</u>
<i>cse1-91 BamHI-F</i>	<u>CGGGATCCGATAACTATAAAAAATACATAGTC</u>
<i>cse1-56 BamHI-F</i>	<u>CGGGATCCCTATAGGTAGGTAATATTCCAAATTG</u>
<i>cse1-52 R</i>	ATACAGTCTTAAATCAGACTATGTATTTTTTATAGTTATCTGTTA
<i>cse1-51 BamHI-F</i>	<u>CGGGATCCCTACTTTTTATTTATATTTTCG</u>
<i>cse1+49 Kpn I-R</i>	<u>GGGGTACCTACCTACCTATAGGTAGATTCC</u>
<i>cse1+50 F</i>	ATATTCCAAATTGGCTTAATAAATAGCCCTGCAGGAGTA AAGGT
<i>cse1+99 Kpn I-R</i>	GGGGTACCATCCATACCTTTACTCCTGC
<i>cse1+191 Kpn I-R</i>	<u>GGGGTACCTCCTGAATGCGGTTATCCAG</u>
<i>cse1+794 F</i>	CGAATCACTGCTACTGCAAATGC
<i>cse1+1140 Hind III-R</i>	<u>CCCAAGCTTGAGAAATTGCTCTTCGTGCG</u>
<i>cse1+1498 Sal I-R</i>	<u>ACGCGTCGACGTTTACGATAGCTGCTGGCTAAATCC</u>
<i>cse1 sust-AB F</i>	CGGTTTATCCCATAACCCTAGGGAACGTATTTAAAG
<i>cse1 sust-AB R</i>	CTTTAATACGTTCCCTAGGGTTATGGGATAAACCG
<i>cse1 sust-C F</i>	CCCATAATTTTAGGGAACGCGCTAAAGATAACAG
<i>cse1 sust-C R</i>	CTGTTATCTTTAGCGCGTTCCCTAAAATTATGGG

---



---

---



---

<i>cse1</i> sust-D F	CGATTAAAGATAACAGGCGACTATAAAAAATAC
<i>cse1</i> sust-D R	GTATTTTTTATAGTCGCCTGTTATCTTTAATACG
<i>cse1</i> sust-E F	CAGATAACTATAAAGGGTACATAGTCTGATTTAAG
<i>cse1</i> sust-E R	CTTAAATCAGACTATGTACCCTTTATAGTTATCTG
<i>cse1</i> sust-F F	CATAGTCTGATTCGGGACTGTATCTAC
<i>cse1</i> sust-F R	GTAGATACAGTCCCGAATCAGACTATG
<i>cse2</i> -450 <i>Bam</i> H I-F	<u>CGGGATCCTTCCGCTCATTAGCGTTTCG</u>
<i>cse2</i> -62 <i>Bam</i> H I-F	<u>CGGGATCCTCGTAAACATAAAGCGCGCA</u>
<i>cse2</i> +134/+181 B/K F	<b>GATCC</b> GCGGGCTGAGTTGAAAAGAATGGCGCCGCCATATGGCGTA ATGATAT <b>GGTAC</b>
<i>cse2</i> +134/+181 B/K R	CATATCATTACGCCATATGGCGGCGCCATTCTTTTCAACTCAGCCCG CG
<i>cse2</i> +142 <i>Bam</i> H I-F	<u>CGGGATCCAGTTGAAAAGAATGGCGCCGCCAT</u>
<i>cse2</i> +142/+181 B/K F	<b>GATCC</b> AGTTGAAAAGAATGGCGCCGCCATATGGCGTAATGATAT <b>GG</b> <b>TAC</b>
<i>cse2</i> +142/+181 B/K R	CATATCATTACGCCATATGGCGGCGCCATTCTTTTCAACT <b>G</b>
<i>cse2</i> +142/+202 B/K F	<b>GATCC</b> AGTTGAAAAGAATGGCGCCGCCATATGGCGTAATGATATGT GAAGGACATGACGCTCTCC <b>GGTAC</b>
<i>cse2</i> +142/+202 B/K R	CGGAGAGCGTCATGTCCTTACATATCATTACGCCATATGGCGGCG CCATTCTTTTCAACT <b>G</b>
<i>cse2</i> +154/+181 B/K F	<b>GATCC</b> TGGCGCCGCCATATGGCGTAATGATAT <b>GGTAC</b>
<i>cse2</i> +154/+181 B/K R	CATATCATTACGCCATATGGCGGCGCCAG
<i>cse2</i> +181 <i>Kpn</i> I-R	<u>GGGGTACCATATCATTACGCCATATGGC</u>
<i>cse2</i> +267 <i>Kpn</i> I-R	<u>GGGGTACCGACAAACAGAGCCAGAGCCAT</u>
CRISPR-231 <i>Bam</i> H I-F	<u>CGGGATCCTGGTTACTGGAAATTCCAGC</u>
CRISPR+503 <i>Xho</i> I-R	<u>CCGCTCGAGCGGAACGTCCTGATATGCTGGTGAAC</u>
pKK8-R	GATTTGTCCTACTCAAGC

---



---

\*Los sitios de restricción (*Bam*H I: 5'-GGATCC-3'; *Kpn* I: 5'-GGTACC-3'; *Xho* I: 5'-CTCGAG-3'; *Hind* III: 5'-AAGCTT-3') y nucleótidos adyacentes necesarios para el reconocimiento por la correspondiente endonucleasa se encuentran subrayados. En el caso de oligonucleótidos complementarios, las bases nucleotídicas que restablecen los respectivos sitios de corte se encuentran en negritas.

## **ANEXO II**

# **ARTÍCULOS PUBLICADOS**

# The CRISPR/Cas Immune System Is an Operon Regulated by LeuO, H-NS, and Leucine-Responsive Regulatory Protein in *Salmonella enterica* Serovar Typhi<sup>∇†</sup>

L. Medina-Aparicio, J. E. Rebollar-Flores, A. L. Gallego-Hernández, A. Vázquez, L. Olvera, R. M. Gutiérrez-Ríos, E. Calva, and I. Hernández-Lucas\*

*Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62210, Mexico*

Received 8 December 2010/Accepted 4 March 2011

**Prokaryotes have developed multiple strategies to survive phage attack and invasive DNA. Recently, a novel genetic program denominated the CRISPR/Cas system was demonstrated to have a role in these biological processes providing genetic immunity. This defense mechanism is widespread in the *Archaea* and *Bacteria*, suggesting an ancient origin. In the last few years, progress has been made regarding the functionality of the CRISPR/Cas system; however, many basic aspects of the system remain unknown. For instance, there are few studies about the conditions and regulators involved in its transcriptional control. In this work, we analyzed the transcriptional organization of the CRISPR/Cas system as well as the positive and negative regulators involved in its genetic expression in *Salmonella enterica* serovar Typhi. The results obtained show that in *S. Typhi* the CRISPR/Cas system is a LeuO-dependent operon silenced by the global regulator LRP, in addition to the previously known nucleoid-associated protein H-NS; both LRP and H-NS bind upstream and downstream of the transcriptional start site of *casA*. In this study, relevant nucleotides of the *casA* regulatory region that mediate its LeuO transcriptional activation were identified. Interestingly, specific growth conditions (N-minimal medium) were found for the LeuO-independent expression of the CRISPR/Cas system in *S. Typhi*. Thus, our work provides evidence that there are multiple modulators involved in the genetic expression of this immune system in *S. Typhi* IMSS-1.**

In nature, bacteria and archaea are exposed to multiple viral infections; therefore, several strategies have evolved to inhibit phage predation. The mechanisms to avoid phage infection include the prevention of phage adsorption and DNA injection; the abortive infection and modification-restriction systems are also relevant (20, 37). Recently, a novel genetic program designated the CRISPR/Cas system was shown to provide immunity against virus and exogenous DNAs (7, 9, 42). In general, this system consists of a cluster of proteins (Cas) with functional domains of nucleases, helicases, polymerases, and polynucleotide binding proteins (24, 32, 41); a leader of A/T-rich noncoding sequences (31) is located immediately upstream of a singular cluster of regularly interspaced short palindromic repeats (CRISPR) that are composed of direct repeats separated by spacers that, in some cases, resemble viral or plasmid DNA (8, 49, 51). Accumulating evidence suggests that the CRISPR sequences are transcribed as a large mRNA which is cleaved in the repeat sequences by a Cas endonuclease protein (9, 11, 26), generating small transcripts (crRNAs) that are used by the Cas proteins to target exogenous genetic materials, leading to their degradation (9, 21, 25, 43). Thus, CRISPR/Cas provides immunity against foreign genetic mate-

rial. The incorporation of new spacers into the CRISPR during a phage challenge has been shown to confer specific resistance against the invading virus in future events (7). The presence of a large assortment of short invader sequences in the CRISPR cluster shows the evolutionary history of multiple viral infections in the host, providing evidence of the rapid evolution of this dynamic program (6, 27).

Significant progress has been made regarding the functionality of the CRISPR/Cas system (30, 34, 44, 61). However, this genetic program is not fully characterized; for instance, there are few studies about its transcriptional organization and genetic control (1, 28, 38, 55, 58, 64). In this regard, we previously reported that in *Salmonella enterica* serovar Typhi IMSS-1, the *casA* gene (STY3070) is regulated positively by LeuO; footprinting experiments demonstrated that LeuO binds upstream of the *casA* transcriptional start site, and by use of transcriptional fusions it was shown that *casA* is induced in an H-NS-deficient strain (28). This finding was later observed in *Escherichia coli* K-12, as electrophoretic mobility shift assays (EMSA) and Northern blot experiments showed that LeuO directly and positively regulates the *casA* (*ycgL*) gene, and in an H-NS-deficient strain overexpressing LeuO, the transcriptional *casA* levels were higher than in the wild type overexpressing this regulator (58). Furthermore, Pul et al. (55) determined the  $\sigma^{70}$ -dependent promoters in the 5' region of *casA*, including a *Pcas* promoter and a divergent anti-*Pcas* promoter as well as another one identified in the leader region upstream of the CRISPR I array (*PcrisprI*). Interestingly, the three promoters are repressed by H-NS. Recently, Westra et al. (64) showed that H-NS-mediated *casA* repression can be relieved by the LeuO transcriptional regulator. Together, these data support

\* Corresponding author. Mailing address: Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Cuernavaca, Morelos 62210, Mexico. Phone: (52) (777) 329-1627. Fax: (52) (777) 313-8673. E-mail: ismaelh@ibt.unam.mx.

† Supplemental material for this article may be found at <http://jbb.asm.org/>.

∇ Published ahead of print on 11 March 2011.

the positive and negative roles of LeuO and H-NS in the genetic control of the CRISPR/Cas program in *E. coli* K-12. In the present work, we determined that in *S. Typhi* IMSS-1, the complete CRISPR/Cas system is driven by a single LeuO-dependent promoter, which is directly repressed by H-NS and LRP; both proteins bind upstream and downstream of the *casA* transcriptional start site. In addition, we have identified nucleotides of the regulatory region of *casA* that mediate its LeuO-dependent activation. Specific growth conditions that induce the CRISPR/Cas program in *S. Typhi* IMSS-1 in a LeuO-independent manner are also reported, thus demonstrating that multiple global regulators and specific conditions modulate the transcriptional expression of this genetic program in *Salmonella enterica* serovar Typhi.

## MATERIALS AND METHODS

**Bacterial strains and culture conditions.** The bacterial strains and plasmids used are listed in Table S1 in the supplemental material. *S. Typhi* IMSS-1 was grown in LB (10 g tryptone, 5 g yeast extract, and 10 g NaCl per liter), liquid MA medium (7 g nutrient broth, 1 g yeast extract, 2 g glycerol, 3.75 g  $K_2HPO_4$ , and 1.3 g  $KH_2PO_4$  per liter) (35), or N-minimal medium [5 mM KCl, 7.5 mM  $(NH_4)_2SO_4$ , 0.5 mM  $K_2SO_4$ , 1 mM  $KH_2PO_4$ , 100 mM Tris-HCl (pH 7.5), 200  $\mu$ M  $MgCl_2$ , 0.5% glycerol, and 0.1% Casamino Acids] (15). When required, the following antibiotics were added: kanamycin (Km), 30  $\mu$ g/ml; tetracycline (Tc), 12  $\mu$ g/ml; and ampicillin (Ap), 100  $\mu$ g/ml. *S. Typhi* and *E. coli* strains were grown aerobically at 37°C.

**DNA manipulations.** Plasmid and genomic DNA isolation was performed according to published protocols (57). Primers for PCR amplifications were provided by the oligonucleotide synthesis facility at our institute (see Table S2 in the supplemental material). Restriction enzymes, ligase, nucleotides, and polymerases were obtained from New England BioLabs or Gibco BRL. For sequencing, double-stranded DNA was purified with a High Pure plasmid isolation kit (Boehringer Mannheim, Germany), and sequencing was performed with an automatic Perkin Elmer/Applied Biosystems 377-18 system.

**Northern analysis.** Total RNA was collected from 10 ml of culture by acid-phenol extraction. RNAs were separated in 1% agarose gels, transferred to nylon membranes (GE Healthcare), and probed with PCR probes labeled with a Rediprime kit using  $[\alpha\text{-}^{32}P]dCTP$ . The membranes were prehybridized and hybridized at 65°C, using Rapid Hyb buffer (GE Healthcare), and were washed with 2 $\times$  SSC (1 $\times$  SSC is 0.15 M NaCl plus 0.015 M sodium citrate) plus 0.1% SDS for 15 min, 1 $\times$  SSC plus 0.1% SDS for 15 min, 0.5 $\times$  SSC plus 0.1% SDS for 15 min, and, finally, 0.1 $\times$  SSC plus 0.1% SDS at 65°C for 30 min (57). To determine the molecular weight of the mRNA, a 0.24- to 9.5-kb RNA ladder (Invitrogen) was used.

**Construction of transcriptional reporter fusions.** Oligonucleotides (see Table S2 in the supplemental material) were designed to PCR amplify the complete 5' intergenic region as well as shorter fragments of different lengths of the *casA* regulatory region. PCR fragments were double digested with BamHI-XhoI or BamHI-KpnI and ligated into pKK232-8 or pKK232-9 (see Table S1 in the supplemental material), which contains the promoterless *cat* gene. Fusions were sequenced to verify the correct DNA sequences of the PCR fragments.

**CAT assay.** Chloramphenicol acetyltransferase (CAT) assays were performed as follows (46). *S. Typhi* strains were grown in MA medium supplemented when required with Ap or Km, with or without IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside; 50  $\mu$ M), to optical densities at 595 nm ( $OD_{595}$ ) of 0.4, 0.6, 0.8, 1, and 1.3; at these times, 1.5 ml of bacterial culture was collected by centrifugation and washed with 0.8 ml of TDDT buffer (50 mM Tris-HCl, pH 7.8, 30  $\mu$ M DL-dithiothreitol). Bacterial cells were resuspended in 0.6 ml of TDDT and sonicated on ice for 10-s intervals with 10-s rest periods until the extract was clear. The homogenate was centrifuged, and the supernatant was used for activity measurement. For CAT assays, 5  $\mu$ l of each extract was added in duplicate to a 96-well enzyme-linked immunosorbent assay (ELISA) plate, followed by the addition of 0.2 ml of a reaction mixture containing 1 mM DTNB [5,5'-dithiobis(2-nitrobenzoic acid)], 0.1 mM acetyl-coenzyme A (acetyl-CoA), and 0.1 mM chloramphenicol in 0.1 M Tris-HCl, pH 7.8. The absorbance at 412 nm was measured every 5 s for 5 min, using a Ceres 900 scanning autoreader and microplate workstation. The protein concentration of the cell extracts was determined using the bicinchoninic acid (BCA) protein assay reagent (Pierce). Protein values and the mean rate of product formation by CAT were used to

determine CAT specific activity, in  $\mu$ mol/min/mg protein. To determine the transcriptional activity in N-minimal medium, the bacteria were grown for 16 h, 1 ml was collected, and the CAT protocol described above was used. The results presented in the figures are the means for three independent experiments performed in duplicate.

**Purification of H-NS and LRP proteins.** H-NS and LRP proteins were purified as described previously (16). Briefly, the genes were cloned into the arabinose-inducible vectors pBAD/Myc-His and pMPM-T6, respectively (Invitrogen) (47). *E. coli* BL21(DE3) harboring the cloned gene was grown overnight in LB medium. One milliliter of this preculture was inoculated into 100 ml LB medium and allowed to grow to an  $OD_{600}$  of 0.4 before being induced for 4 h with L-arabinose (Sigma-Aldrich) to a final concentration of 1%. Samples were collected and centrifuged at 8,000 rpm (Thermo Scientific CL31R multispeed centrifuge) for 5 min at 4°C. The cell pellet was resuspended in binding buffer (8 M urea, 100 mM  $NaH_2PO_4$ , 10 mM Tris-HCl, pH 8.0), the cells were sonicated, and the suspension was centrifuged at 12,000 rpm for 5 min at 4°C and kept at 4°C.

A nickel affinity chromatography column (Ni-nitrilotriacetic acid [Ni-NTA] agarose; Qiagen) was equilibrated with binding buffer, the solution containing the protein was loaded onto the column and washed with 10 volumes of urea buffer at pH 8.0 and urea buffer at pH 6.0, and the bound protein was eluted with urea buffer, pH 4.5. Fractions containing H-NS-Myc-His<sub>6</sub> or Lrp-His<sub>6</sub> were resolved by SDS-PAGE. The selected fractions were loaded into a Slide-A-Lyzer 10K cassette (Pierce) and gradually dialyzed at 4°C in a buffer containing 50 mM Tris-HCl, pH 8.0, 10 mM  $MgCl_2$ , 20% glycerol, 500 mM NaCl, 0.1% Triton X-100, and 4 M urea. The same buffer containing different amounts of urea (1 M and 0.2 M) was used to complete the dialysis. The proteins were stored in buffer containing 30 mM Tris-HCl, pH 8.0, 10 mM  $MgCl_2$ , 20% glycerol, 500 mM NaCl, 0.1% Triton X-100, and 3 mM EDTA. Protein concentrations were determined by the Bradford procedure.

**Gel EMSA.** Nonradioactive EMSAs were performed according to a previously described protocol (16). The probes were obtained by PCR, using the primers described in Table S2 in the supplemental material. Each probe (100 ng) was mixed with increasing concentrations of purified protein in the presence of 10 $\times$  H-NS binding buffer (400 mM HEPES, 80 mM  $MgCl_2$ , 500 mM KCl, 10 mM DTT, 0.5% NP-40, and 1 mg/ml bovine serum albumin [BSA]) or 5 $\times$  LRP buffer (100 mM Tris, pH 8.0, 2 mM EDTA, 250 mM NaCl, 5 mM  $MgCl_2$ , 62.5% [vol/vol] glycerol, 500 mg/ml BSA, and 0.5 mM DTT). The mixture was incubated for 20 min at room temperature and then separated by electrophoresis in 6% or 5% native polyacrylamide gels in 0.5 $\times$  Tris-borate-EDTA buffer. The DNA bands were visualized by ethidium bromide staining.

**Site-directed mutagenesis.** To analyze their role in the regulation of the CRISPR/Cas system in *S. Typhi* IMSS-1, LRP-, H-NS- (19), and double H-NS/LRP-deficient strains were obtained. The corresponding genes were deleted by a one-step nonpolar gene inactivation procedure and replaced with selectable antibiotic resistance gene markers (14). Different primer sets were used to verify the presence of the antibiotic resistance gene cassette. Each mutation was further characterized by sequencing to verify the authenticity of the deletion and the presence of the resistance marker.

**Nucleotide sequence accession numbers.** The *casA-cas2* and CRISPR sequences of *S. Typhi* IMSS-1 were deposited in the DDBJ database under accession number HQ655820.

## RESULTS

**A LeuO-dependent promoter is necessary to transcribe the complete CRISPR/Cas operon in *Salmonella enterica* serovar Typhi IMSS-1.** The CRISPR/Cas cluster is found in many prokaryotic genomes, including those of the *Enterobacteriaceae*. In *Salmonella enterica* serovar Typhi IMSS-1, a Mexican clinical strain (52), we detected the presence of a CRISPR/Cas system that resembles the type E (Cse) system found in *S. Typhi* CT18 and *S. Typhi* Ty2 (22, 24), and in this study we analyzed its transcriptional expression. The IMSS-1 cluster contains a 356-bp intergenic 5'-regulatory region and eight *cas* genes (STY3071 and STY3070-STY3069-STY3068-STY3067-STY3066-STY3065-STY3064) that are orthologous to *cas3* and *casA-casB-casC-casD-casE-casI-cas2* in *E. coli* K-12 (9). It also has an 84-bp leader and the CRISPR I sequences, which contain seven 29-bp repeats and six 32-bp spacers (Fig. 1A);



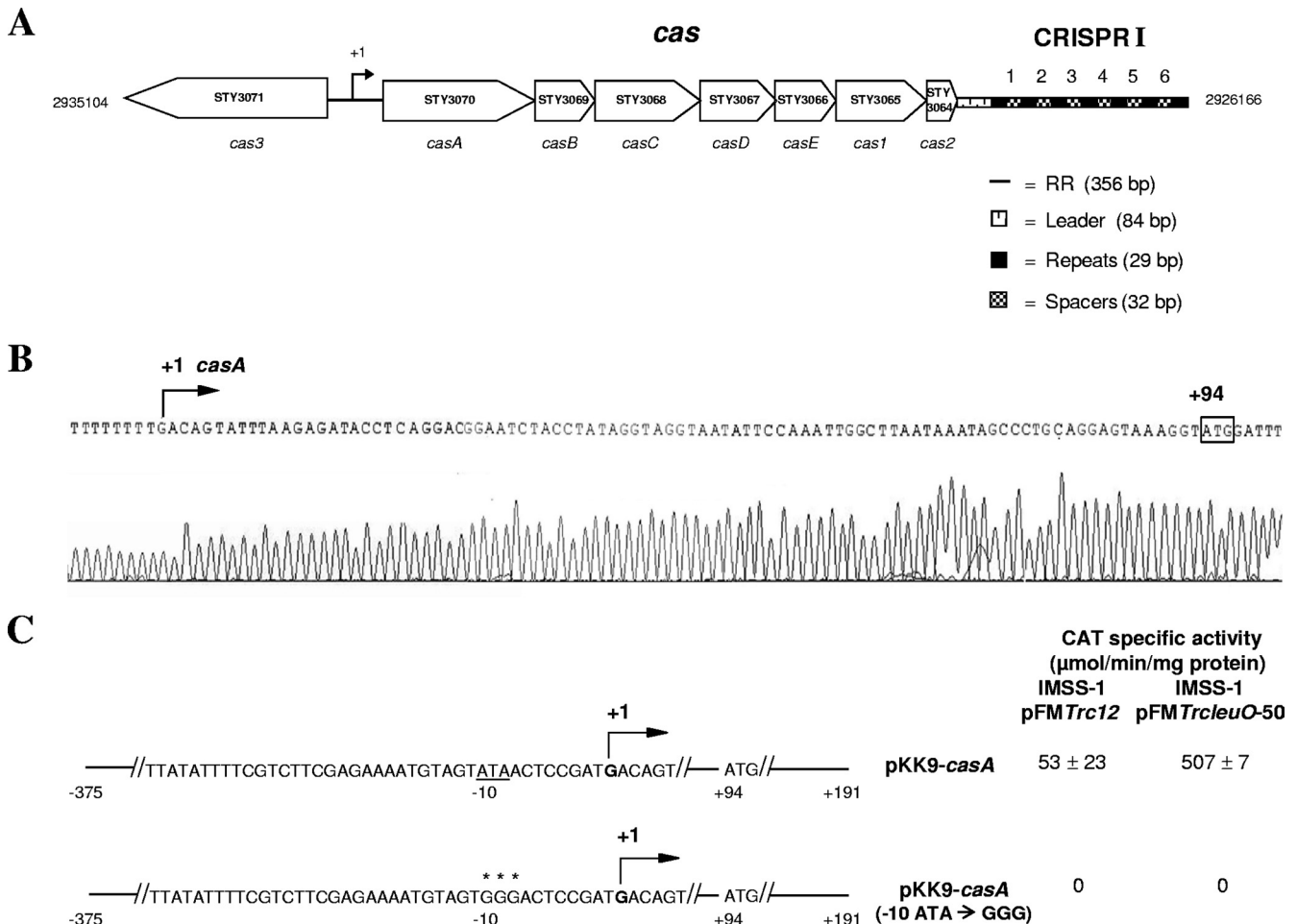


FIG. 1. Genetic organization of the CRISPR/Cas immune system in *S. Typhi* IMSS-1 and identification of the transcriptional start site of *casA* (STY3070). (A) The CRISPR/Cas system is composed of a 5'-regulatory region (RR) of 356 bp, eight genes (*cas3* and *casA-casB-casC-casD-casE-cas1-cas2*), a leader of 84 bp, and the CRISPR I array, containing seven repeats and six spacers. The bent arrow indicates the STY3070 (*casA*) transcriptional start site. The divergent STY3071 (*cas3*) gene is also represented. (B) Mapping of the chromosomal transcriptional start site of the *casA* gene. A 5'-RACE experiment was performed using total RNA from *S. Typhi* harboring the pFMTrcleuO-50 plasmid and induced with 50  $\mu\text{M}$  IPTG. The nucleotide immediately after the polynucleotide tail (T) corresponds to the transcriptional start site (bent arrow), located 94 bp upstream of the ATG translational start codon of *casA*. (C) Site-directed mutagenesis of the -10 box of *casA* was performed by use of a QuikChange Multi site-directed mutagenesis kit (Stratagene), and the wild-type ATA sequence (underlined) was replaced by a GGG sequence (asterisks); the mutation was verified by sequencing. The CAT activities of the wild-type plasmid pKK9-*casA* and the pKK9-*casA* (-10 ATA  $\rightarrow$  GGG) plasmid were determined in the wild-type *S. Typhi* strain overexpressing the pFMTrcleuO-50 plasmid with 50  $\mu\text{M}$  IPTG as well as in *S. Typhi* containing the pFMTrc12 plasmid.

the repeats are identical to those found in *S. Typhi* CT18 and *E. coli* K-12. Analysis of the spacers by use of the Blastn algorithm (3) against a database of phage and virus sequences downloaded from the DDBJ data bank (33) showed no significant similarity with any sequence in the database. In this work, the *S. Typhi cas* genes were named according to their orthologs in *E. coli*.

Previously, we reported the results of primer extension experiments showing that the *S. Typhi casA* (STY3070) regulatory region contains two LeuO-dependent transcription initiation sites, located 94 and 84 bp upstream of the translational ATG start codon (28). To confirm the primer extension results, 5' rapid amplification of cDNA ends (5'-RACE) was performed according to the method of Mendoza-Vargas et al. (48). Briefly, total RNA from the *S. Typhi* IMSS-1 wild-type

strain harboring the pFMTrcleuO-50 plasmid (pFMTrc12 derivative plasmid carrying *leuO* behind an IPTG-inducible *Ptrc* promoter [see Table S1 in the supplemental material]) (18) and induced with 50  $\mu\text{M}$  IPTG was randomly amplified, and in the resulting cDNA, an adapter (homopolynucleotide of A nucleotides) was incorporated by a terminal transferase. This cDNA was used with an oligonucleotide of T nucleotides and specific *casA* primers in a PCR, and the PCR products were purified and sequenced. The results showed that a single chromosomal transcriptional site located 94 bp upstream of the ATG was the only one detected by 5'-RACE (Fig. 1B). Site-directed mutagenesis was performed to modify the -10 sequence (ATA  $\rightarrow$  GGG) of the *casA* initiation site. PCR fragments that contained the entire *casA* regulatory region as well as the -10 substitutions were fused to the *cat* reporter gene,

TABLE 1. The CRISPR/Cas cluster comprises a LeuO-dependent operon in *S. Typhi*<sup>a</sup>

Plasmid	CAT sp act ( $\mu\text{mol}/\text{min}/\text{mg}$ protein)	
	IMSS-1/pFM <i>Trc12</i>	IMSS-1/pFM <i>TrcleuO-50</i>
pKK9-RR- <i>casA-cas2</i>	36 $\pm$ 3	273 $\pm$ 5
pKK9- $\Delta$ RR- <i>casA-cas2</i>	0	0
pKK9-RR- <i>casA</i> -CRISPR	8 $\pm$ 11	137 $\pm$ 0.6

<sup>a</sup> Transcriptional values for the LeuO-dependent *casABCDE12*-CRISPR operon were determined in MA medium. The plasmids pKK9-RR-*casA-cas2*, pKK9- $\Delta$ RR-*casA-cas2*, and pKK9-RR-*casA*-CRISPR were introduced into IMSS-1 harboring plasmid pFM*Trc12* and IMSS-1 containing plasmid pFM*TrcleuO-50*. The CAT specific activity of each strain was determined using samples collected at OD<sub>595</sub> values of 0.4, 0.6, 0.8, and 1 and at 12 h; the values presented are the results obtained at 12 h. IPTG (50  $\mu\text{M}$ ) was used to induce the pFM*TrcleuO-50* plasmid. Data are means  $\pm$  standard deviations for three independent experiments performed in duplicate.

producing plasmids pKK9-*casA* and pKK9-*casA* (−10 ATA → GGG), respectively, which were introduced into the *S. Typhi* IMSS-1 strain containing pFM*Trc12* (a p*Trc99A*[5] derivative containing the p15A1 origin of replication) or pFM*TrcleuO-50* (see Table S1). Experiments with these fusions showed that the overexpression of LeuO induced the transcriptional activity of *casA*, while no CAT activity in the presence of LeuO was observed with the −10 box substitutions, indicating that a unique LeuO-dependent promoter controls the genetic expression of *casA* in *S. Typhi* IMSS-1 (Fig. 1C).

These results prompted us to determine whether the complete CRISPR/Cas system is a LeuO-dependent operon. To resolve this issue, a large DNA fragment harboring the *casA* promoter region and the entire *casA*-to-*cas2* gene cluster (6.15 kb) was first fused to the *cat* reporter gene, creating plasmid pKK9-RR-*casA-cas2*. This fusion was evaluated in an *S. Typhi* IMSS-1 wild-type strain harboring the control vector pFM*Trc12* or the pFM*TrcleuO-50* plasmid. The transcriptional results with pKK9-RR-*casA-cas2* showed minimal activity in the absence of LeuO, while in the presence of LeuO the transcriptional activity increased notably (Table 1). The fusion plasmid pKK9- $\Delta$ RR-*casA-cas2*, which contains the same genetic elements but is devoid of the *casA* promoter region, showed null CAT activity in either the absence or presence of LeuO (Table 1). These data support the notion that *casABCDE12* is a LeuO-dependent operon. To elucidate whether CRISPR is part of the *casABCDE12* LeuO-dependent operon, the fusion plasmid pKK9-RR-*casA*-CRISPR (6.7 kb), containing the *casA* promoter region, the *casA* to *cas2* genes, and the CRISPR sequences, was constructed and independently introduced into IMSS-1/pFM*Trc12* and IMSS-1/pFM*TrcleuO-50*. In the absence of LeuO, only residual activity was observed, which notably contrasted with the transcriptional activity detected with LeuO overexpression (Table 1). A fusion of the complete 5' intergenic region of the CRISPR was constructed, and no activity was observed in the presence or absence of LeuO (data not shown). These results indicate that the complete CRISPR/Cas system in *S. Typhi* comprises a LeuO-dependent operon. The pKK9-*casA*, pKK9-RR-*casA-cas2*, and pKK9-RR-*casA*-CRISPR reporter fusions showed 507 (Fig. 1), 272 (Table 1), and 136 (Table 1) activity units, respectively, with LeuO overexpression. Interestingly, a gradual decrease in activity from *casA* to *cas2* was also found in the

presence of LeuO in *E. coli* K-12 (64). The data presented above support the notion that in *S. Typhi* IMSS-1 a single LeuO-dependent promoter drives CRISPR/Cas expression.

The expression of the LeuO-dependent CRISPR/Cas operon was also analyzed by chromosomal reverse transcription-PCR (RT-PCR). Total RNA isolated from the wild-type strain harboring pFM*TrcleuO-50* and induced with IPTG was used with a specific CRISPR primer (CRISPR-R) located at the end of repeat 7 (see Table S2 in the supplemental material) to obtain a cDNA (6 kb) that included the complete *casABCDE12*-CRISPR operon. Using this template, *casA-casC*- and *casC-cas2*-specific primers were used to amplify the corresponding DNA fragments. The results showed that 2.9-kb and 3.4-kb fragments, containing the *casA-casC* and *casC-cas2* regions, respectively, were amplified (Fig. 2A), further supporting the LeuO-dependent operon structure. Northern experiments were also performed to support the CRISPR/Cas operon organization. Total RNAs from the bacterial strains IMSS-1/pFM*Trc12*, IMSS-1/pFM*TrcleuO-50*, and IMSS-1 $\Delta$ (*casA*-CRISPR) were obtained. The RNAs were transferred, and independent membranes were hybridized with an

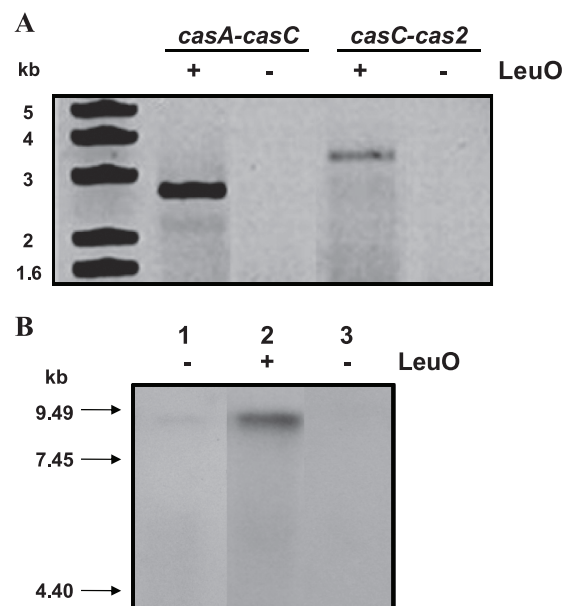


FIG. 2. The CRISPR/Cas cluster comprises a LeuO-dependent operon in *S. Typhi*. (A) The genetic organization of the LeuO-dependent *casABCDE12*-CRISPR operon was determined by chromosomal RT-PCR. *S. Typhi* IMSS-1 containing either the pFM*TrcleuO-50* or pFM*Trc12* plasmid was grown in MA medium with and without 50  $\mu\text{M}$  IPTG, and total RNA was obtained. A 2.9-kb fragment was generated with *casA*- and *casC*-specific primers, and a 3.4-kb fragment was obtained with *casC*- and *cas2*-specific primers. PCRs using RNAs obtained from the strains mentioned above as templates and specific primers for the amplification of the *cas2* and 16S rRNA genes were performed in order to verify the presence of DNA. As expected, no PCR amplification was observed, indicating the absence of DNA in the RNA purification. The left lane shows a DNA molecular size marker. (B) Detection of CRISPR/Cas mRNA by Northern blotting. Chromosomal expression patterns of CRISPR/Cas mRNAs from the bacterial strains IMSS-1/pFM*Trc12* (lane 1), IMSS-1/pFM*TrcleuO-50* (lane 2), and IMSS-1 $\Delta$ (*casA*-CRISPR) (lane 3) are shown. Eight micrograms of total RNA from each sample was loaded and hybridized with an internal CRISPR probe. Sizes of the RNA molecular marker are shown on the left.

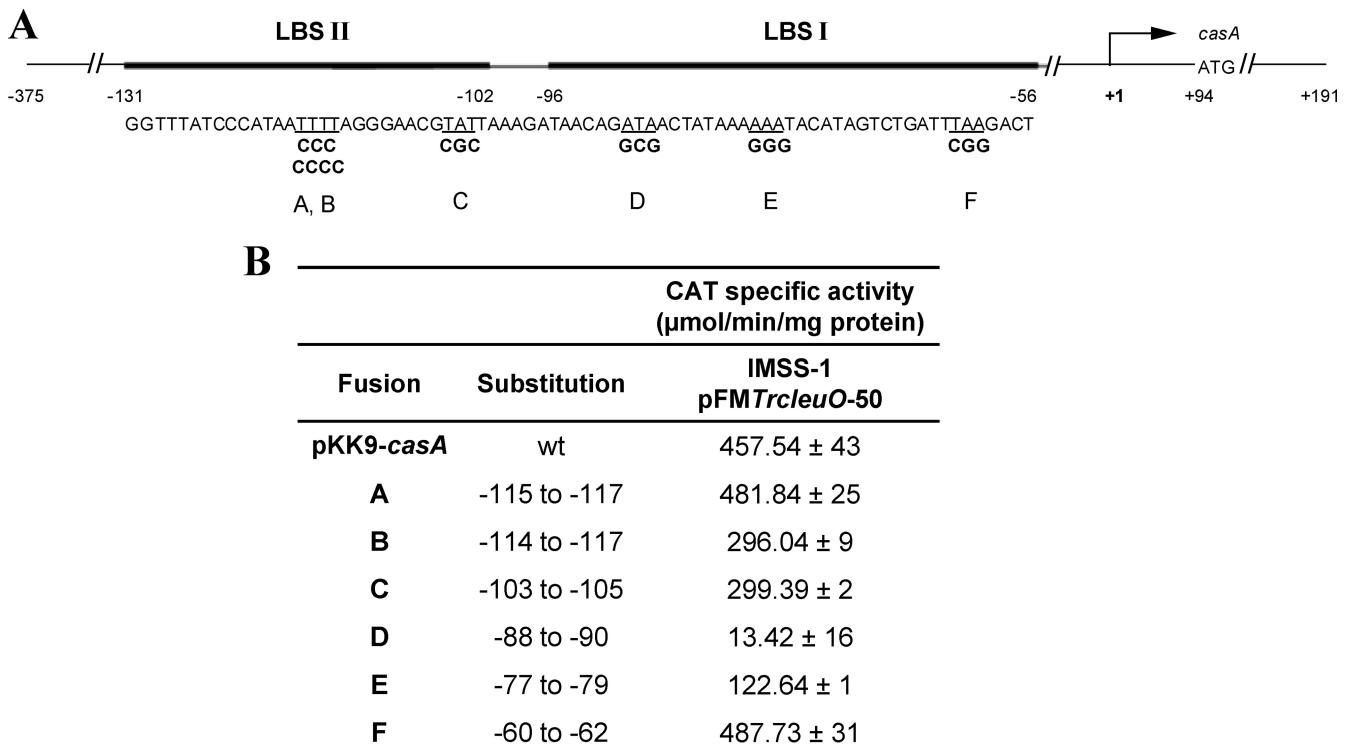


FIG. 3. Mapping of the LeuO recognition sequences in the regulatory region of *casA*. (A) Schematic representation of LBS I and LBS II in the promoter region of *casA*. The entire regulatory region of *casA* was used as a template to obtain mutants in LBS II (A, B, and C) and LBS I (D, E, and F). The substitutions (in bold) were sequenced to verify the correct mutations. The wild-type sequences are underlined. (B) Transcriptional analysis of *casA* promoter fusions with mutations in LBS I and LBS II. *S. Typhi* strain IMSS-1, containing each of the mutations individually as well as the pFMTrcleuO-50 plasmid and induced with 50  $\mu\text{M}$  IPTG, was evaluated for CAT activity. The assays were performed in MA medium at different  $\text{OD}_{595}$  values, and the activities at 12 h are presented. As a positive control, we used the complete wild-type *casA* promoter region. The data are means  $\pm$  standard deviations for three independent experiments performed in duplicate.

internal fragment of the first gene of the operon, the *casA* locus, and with the last part of the operon, the CRISPR sequences. The results showed a null hybridization signal for the *Salmonella*  $\Delta(\text{casA-CRISPR})$  strain, which was devoid of the complete CRISPR/Cas operon; for the IMSS-1/pFMTrc12 strain, a weak signal was detected. In contrast, for the IMSS-1/pFMTrcleuO-50 strain induced with IPTG, a signal of 7 to 9 kb was detected with the *casA* or CRISPR probe. This result demonstrates that the complete CRISPR/Cas system comprises a LeuO-dependent operon. In Fig. 2B, we present only the results obtained with the CRISPR probe, since identical results were obtained with the *casA* probe.

In *E. coli*, the CRISPR/Cas system contains two promoters: the first is a LeuO-dependent promoter (64) located upstream of the *casA* gene, and the second is located in the leader sequence of the CRISPR (55), suggesting that in *E. coli* this system is encoded by two transcriptional units. A comparison between the leader and CRISPR sequences of *E. coli* and *S. Typhi* shows that the leaders differ in length and share a low level of conservation. In fact, the  $-10$  promoter sequence determined for *E. coli* (55) is missing in *S. Typhi* IMSS-1. The low level of conservation in the leader sequences could explain the lack of a CRISPR promoter in *S. Typhi* IMSS-1. It is worth mentioning that in *S. Typhi* and *E. coli* the transcriptional organization of the CRISPR/Cas systems is not conserved, suggesting that the signals or conditions that activate or repress

this genetic program could be different even in these two closely related bacteria.

**Determination of the nucleotides involved in LeuO activation of the CRISPR/Cas genetic system in *S. Typhi* IMSS-1.** As mentioned above, LeuO is a positive activator of the CRISPR/Cas operon in *S. Typhi*; by footprinting experiments, we showed previously that LeuO binds to nucleotides  $-56$  to  $-96$  (LeuO binding site I [LBS I]) and  $-102$  to  $-131$  (LBS II) with respect to the transcriptional start site of *casA* (STY3070) (28). To determine which nucleotides mediate LeuO activation, a collection of substitutions (A, B, C, D, E, and F) were constructed in LBS I and LBS II (Fig. 3A). Several regions in LBS I and LBS II with high A+T content were replaced by G+C nucleotides, considering that LeuO binds preferentially to sites with high A+T content (12). Reporter gene fusions containing three or four substitutions were introduced into the *S. Typhi* IMSS-1 wild-type strain harboring the pFMTrcleuO-50 plasmid. The transcriptional results showed that in the presence of LeuO, the A and F substitutions (nucleotides  $-115$  to  $-117$  and  $-62$  to  $-60$ , respectively) gave similar expression levels to those with the wild-type fusion (Fig. 3B). However, a gradual decrease in activity was detected with the B and C substitutions (nucleotides  $-114$  to  $-117$  and  $-103$  to  $-105$ , respectively) (Fig. 3B). Interestingly, a decrease of 76% was observed with the E substitution (nucleotides  $-77$  to  $-79$ ), and only 3% of wild-type activity remained with the D substitution (nucleo-

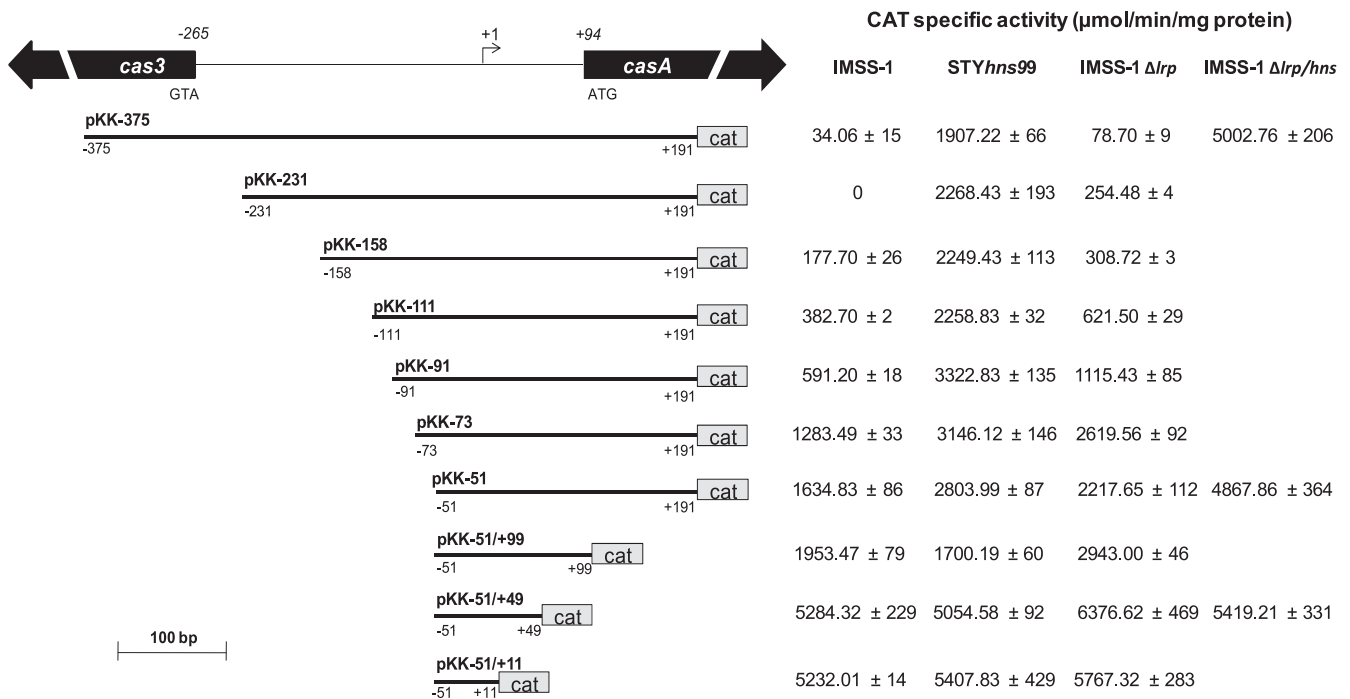


FIG. 4. Promoter fusions harboring different lengths of the *casA* regulatory region. The bent arrow represents the transcriptional start site, located 94 bp upstream of the translational start codon of *casA*; the numbers in italics represent the localization of the *casA* and *cas3* ATG initiation codons relative to the *casA* mRNA start site. Below the diagram of the regulatory region of *casA*, transcriptional fusions are shown. The fusions (in bold) were named according to the transcription initiation site; the numbers below each fusion represent the distances upstream and downstream of the mRNA start site. The right columns represent the expression of the *casA* fusions evaluated in the wild-type *S. Typhi* IMSS-1 strain, the single H-NS (STY*hns99*) and LRP (*Δlrp*) mutants, and the double *Δlrp/hns* mutant strain. The activities were determined at different ODs in rich medium (MA); the data presented correspond to the values obtained at an OD<sub>595</sub> of 1.3. The values are the means ± standard deviations for at least three independent experiments performed in duplicate.

tides -88 to -90) (Fig. 3B). The same constructs were introduced into the *S. Typhi* Δ*leuO* strain harboring pFMT*rleuO*-50, and CAT activity was determined. The results of a single experiment (data not shown) gave the same level of expression for each mutation as that obtained with the wild-type *S. Typhi* strain, indicating that the chromosomal *leuO* gene did not affect the expression of the fusions evaluated. These data indicate that substitutions B, C, D, and E affected the transcriptional *casA* induction mediated by LeuO, suggesting that both LBS I and LBS II are relevant for *casA* LeuO-dependent induction. It will certainly be interesting to determine the impact of base changes on binding of the LeuO protein to elucidate whether the decrease in activity is because the protein is effectively unable to bind to the DNA or binds it in an ineffective manner.

**H-NS partially represses transcriptional expression of the CRISPR/Cas genetic program in *S. Typhi* IMSS-1.** H-NS is a master regulator of the *S. Typhi* genome, and significant evidence has been compiled about the role of this protein in the expression of several genes belonging to the LeuO regulon (13, 16, 28, 59, 60). Previously, we used transcriptional fusions to show basal expression of the *casA* (STY3070) promoter region in the wild-type strain. However, transcriptional activity of this fusion was induced in the presence of LeuO as well as in an H-NS-deficient strain, supporting the notion that H-NS negatively regulates the expression of *casA* (STY3070) in *S. Typhi* IMSS-1 (28). To further characterize the *casA* regulatory re-

gion, a series of promoter fragments harboring progressively shorter upstream sequences of the entire promoter region were generated by PCR, and each of these DNA fragments was fused to the *cat* reporter gene in plasmid pKK232-8 (Fig. 4). The expression levels of these derivatives were examined by CAT activity assay of the wild-type *S. Typhi* IMSS-1 strain as well as the STY*hns99* H-NS-deficient strain (19) (Fig. 4). The CAT activity levels of all of the single fusions analyzed (from pKK-375, which contains the entire *casA* 5' intergenic regulatory region, to pKK-231, pKK-158, pKK-111, pKK-91, pKK-73, and pKK-51) were higher in the H-NS-deficient strain than in the wild-type strain. These results are consistent with the negative role of H-NS in the genetic expression of *casA* (Fig. 4).

Interestingly, the CAT activity values for the H-NS-deficient strain with the pKK-375, pKK-231, pKK-158, and pKK-111 fusions were between 1,900 and 2,200 activity units, but the values with the pKK-91, pKK-73, and pKK-51 fusions were higher (2,800 to 3,300 activity units), indicating that in these three fusions regions of negative regulation were eliminated (Fig. 4). An increased activity in the wild-type strain was also observed with fusions pKK-91, pKK-73, and pKK-51 compared with the longer fusions pKK-375, pKK-231, pKK-158, and pKK-111, supporting the notion that nucleotides in the region from positions -111 to -51 are recognized by negative regulatory elements, in addition to H-NS, repressing *casA* expression (Fig. 4).

Additional fusions deleting downstream sequences relative

to the transcriptional start site (pKK-51/+11 and pKK-51/+49) showed between 5,000 and 5,200 activity units for the wild type as well as the H-NS mutant strain. However, with the pKK-51/+99 fusion, 1,700 to 1,900 activity units were detected for the wild type and the H-NS mutant, indicating that between nucleotides +49 and +99 with respect to the transcriptional start site, negative regulatory elements repress *casA* expression (Fig. 4). In summary, systematic deletions upstream and downstream of the transcriptional start site of *casA* resulted in increased CAT activity when they were evaluated in an H-NS-deficient strain, indicating that this nucleoid-associated protein partially represses *casA* expression and that other genetic elements in addition to H-NS are involved in the silencing of *casA* in *S. Typhi* IMSS-1.

**LRP is a negative regulator of the CRISPR/Cas operon in *S. Typhi* IMSS-1.** To further characterize the genetic expression of *casA* in *S. Typhi* IMSS-1 and to identify other putative genetic elements involved in the control of *casA*, a bioinformatic analysis was performed. A set of *casA*-orthologous proteins was constructed using the Blastp algorithm (3, 4); the highest E value and an upper threshold of  $10^{-5}$  were considered to select the orthologous proteins. With this method, we detected a set of 21 *casA*-orthologous proteins, which are present in archaea and in alpha-, delta-, and gammaproteobacteria. The 400-bp sequences upstream of the putative translational start sites of the 21 *casA*-orthologous polypeptides were aligned to eliminate redundant sequences. The resulting 18 upstream sequences were analyzed using the Oligo-Analysis program (62). A search on both strands for motifs with a maximum length of 6 nucleotides was performed. After assembling the patterns discovered by the algorithm, we found diverse motifs. A conserved LRP motif (TAATAAA) was present in 10 of the sequences analyzed, and this motif was compared using the TOM TOM algorithm (23) with those reported in the DPINTERACT (56) and REGTRANSDB v4 (36) databases. The results showed that the TAATAAA pattern was preserved in close conservation with an LRP binding site of *E. coli* in both databases ( $P = 0.0080$ ). Interestingly, LRP motifs were located 221 bp upstream and 66 nucleotides downstream of the *casA* transcriptional start site of *S. Typhi* IMSS-1. We also identified that the LRP pattern in *S. Typhi* Ty2 has the same distribution as that in *S. Typhi* IMSS-1. These findings suggest that the LRP protein is perhaps a *casA* negative regulator, especially considering the site found downstream of the transcriptional start site. In *Methanospirillum hungatei*, two LRP sites are present, 38 and 351 bp upstream of the *casA* ATG translational start codon. In the case of *Granulibacter bethendens*, an LRP motif was located 50 bp upstream of the *casA* translational start site. In *E. coli* K-12 and *E. coli* W3110, two LRP motifs were identified whose positions varied significantly, but all were located upstream of the transcriptional start site. In *Desulfovibrio desulfuricans*, *Desulfococcus oleovorans*, and *Marinomonas* MWYL1, an LRP motif was detected ~400 bp upstream of the ATG translational start site. The presence of LRP motifs in the regulatory region of *casA* orthologs belonging to alpha-, delta-, and gammaproteobacteria and to archaea suggested that the LRP protein could be involved in the transcriptional regulation of the CRISPR/Cas system.

To evaluate the role of LRP in the genetic expression of the

*casA* gene, we constructed an *S. Typhi* IMSS-1 strain deficient in *lrp* by using the method of Datsenko and Wanner (14). Transcriptional fusions of different lengths (Fig. 4) were introduced into the *lrp*-deficient strain, and CAT activity assays were performed. The results showed an increased CAT activity in the LRP-deficient strain relative to that in the wild-type strain for most of the fusions evaluated. The effect was most evident with the pKK-91, pKK-73, and pKK-51 fusions (Fig. 4), suggesting a negative role for LRP in the expression of *casA*.

The activities for the pKK-51/+11 and pKK-51/+49 fusions were 5,232 and 5,284 activity units, respectively, with the wild type, whereas with the LRP-deficient strain, the values obtained were 5,767 and 6,376 activity units, respectively. For the pKK-51/+99 fusion, around 1,900 and 2,900 activity units were observed with the wild type and the LRP mutant, respectively (see Fig. 4 for standard deviation values), again supporting the conclusion that negative regulatory elements between nucleotides +49 and +99 contribute to *casA* repression (Fig. 4). Thus, both LRP and H-NS repress the genetic expression of CRISPR/Cas. To evaluate the roles of both proteins, a double *hns/lrp* mutant was constructed, and transcriptional experiments were performed with the wild type, the single *hns* and *lrp* mutants, and the *hns/lrp* double mutant.

The longer transcriptional fusion pKK-375, which contains the complete 5' intergenic regulatory region of *casA*, and the pKK-51 and pKK-51/+49 fusions were introduced into the strains described above. The pKK-375 fusion in the wild-type strain showed a value of 34 activity units; in the single H-NS mutant, an activity of 1,907 activity units was observed; in the *lrp* single mutant, a value of 78 activity units was detected; and the double *hns/lrp* mutant showed 5,002 activity units, hence demonstrating the role of both LRP and H-NS in the repression of *casA* (see Fig. 4 for standard deviation values). Interestingly, the pKK-51 fusion had values of around 1,600, 2,800, 2,200, and 4,800 activity units in the wild type, the *hns* and *lrp* single mutants, and the double *hns/lrp* mutant, respectively. The activity values obtained with pKK-51/+49 were 5,200 for the wild type, 5,000 for the *hns* mutant, 6,300 for the *lrp*-deficient strain, and 5,400 for the double *hns/lrp* mutant strain (Fig. 4). This indicates that the sequence from nucleotides -51 to +49 contains the complete promoter region, without any negative regulatory elements, since with the wild-type strain we obtained the same activity as that observed with the double H-NS/LRP mutant strain. These data also confirm that the *casA* regulatory region contains two main zones of negative regulation, from nucleotides -51 to -375 and from nucleotides +49 to +99, and that H-NS and LRP are the genetic elements that repress *casA* expression. With the data mentioned above, we also showed that in the presence of H-NS, LRP has a minor effect; nevertheless, in the absence of H-NS, LRP mainly represses *casA* expression (Fig. 4), suggesting that LRP could compensate for the absence of H-NS.

It is interesting that previous studies by Pul et al. (55) indicated that in *E. coli* K-12 the LRP protein did not have a significant effect on the expression of the *casA* gene. Thus, in *E. coli* and *S. Typhi*, the negative control of *casA* expression is exerted differently, indicating that the environmental conditions or proteins that control CRISPR/Cas genetic expression in these two related bacteria could be different.

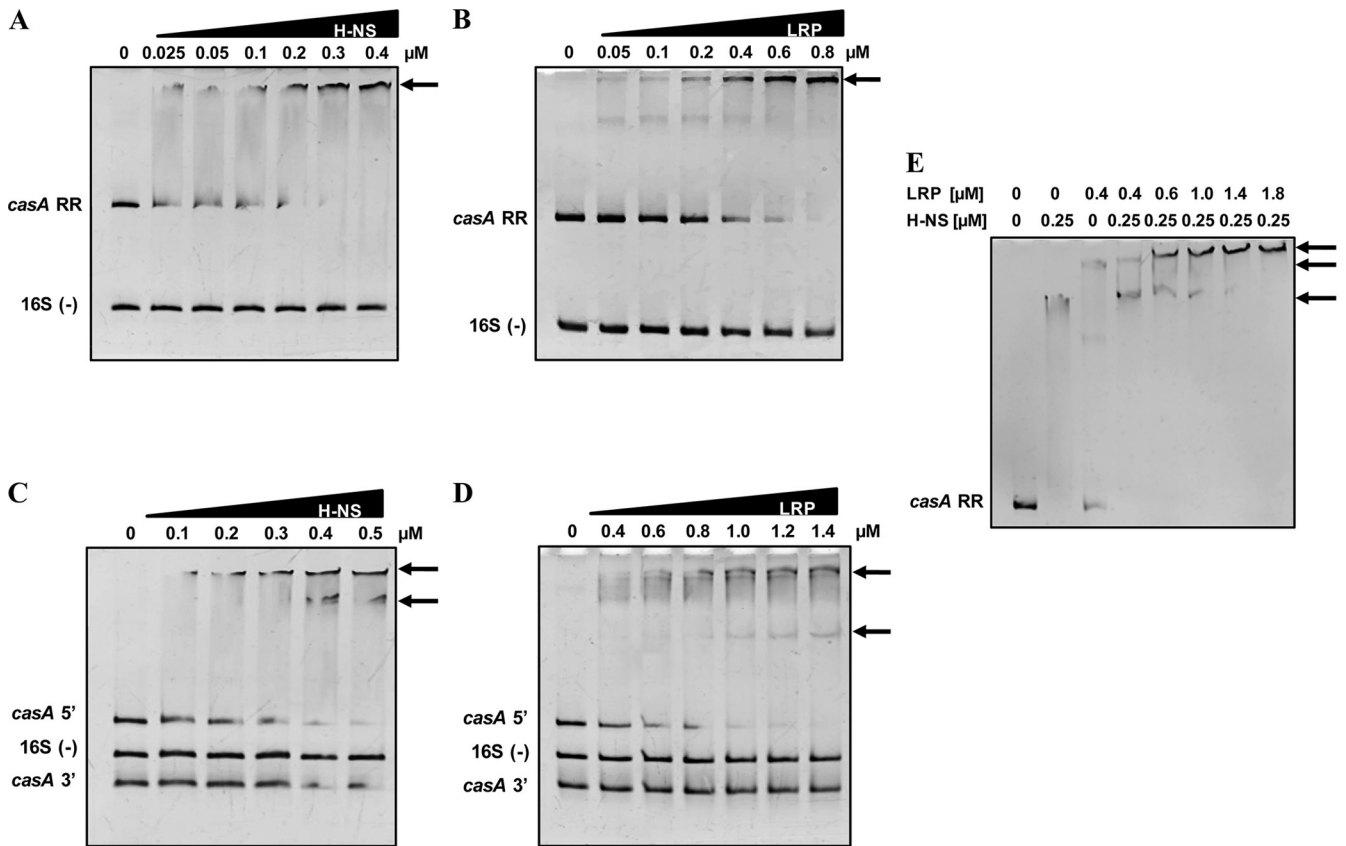


FIG. 5. H-NS and LRP interact with the *casA* regulatory region. EMSAs were performed with the LRP and H-NS purified proteins and with various DNA fragments of the *casA* regulatory region to determine the H-NS and LRP recognition sites. The complete 5' intergenic regulatory region of *casA* (RR) was incubated independently with H-NS (A) or with LRP (B). The *casA* 5' (nucleotides -375 to -52) and 3' (nucleotides +50 to +191) regions were incubated with H-NS (C) or LRP (D), and the *casA* RR was incubated with both proteins (E). Increasing concentrations of purified H-NS-Myc-His<sub>6</sub> or LRP-His<sub>6</sub> protein were incubated with the PCR-generated DNA fragments. The EMSA experiments were performed in 5% or 6% polyacrylamide gels, and gels were stained with ethidium bromide. 16S rRNA gene structural fragments of 242 bp were used as a negative control for the EMSAs. Arrows indicate DNA-protein complexes.

**H-NS and LRP directly repress the expression of the *casA* promoter.** The data presented above showed that sequences upstream and downstream of the transcription initiation site of *casA* are involved in its silencing and that the genetic regulators involved in this repression are LRP and H-NS. To determine whether these proteins bind upstream and downstream of the initiation site, EMSAs were performed with the corresponding purified proteins and with various fragments of the *casA* promoter region. EMSA using the complete regulatory region of *casA* and the H-NS protein showed that H-NS interacted directly with this fragment, since a single complex was observed (Fig. 5A). In the case of EMSA with LRP and the entire *casA* regulatory region, a predominant complex was observed, supporting the direct interaction of LRP with the *casA* promoter region. A second transient band was also visible; however, it disappeared with increased LRP amounts, and only the predominant complex remained in the retardation gel (Fig. 5B). Further EMSAs showed that H-NS and LRP bound to the 5' *casA* region (nucleotides -51 to -375); interestingly, these proteins also interacted with the 3' *casA* region (nucleotides +49 to +191), since we detected two complexes, the highest corresponding to the 5' region and the lowest to the 3'

fragment's interactions with the respective proteins (Fig. 5C and D). That is, EMSAs with H-NS or LRP in the presence of only the 5' *casA* fragment or the 3' *casA* fragment rendered the complexes mentioned above (not shown). It is noteworthy that no strong binding was observed with LRP and the 3' *casA* fragment in the presence of the 5' *casA* fragment (Fig. 5D), suggesting that the latter has a higher affinity for LRP. The 3' *casA* fragment did bind to LRP at 0.6 μM in the absence of the 5' *casA* fragment (data not shown). As a negative control, in these assays we used a 242-bp coding region of the 16S rRNA gene of *S. Typhi* IMSS-1. These results support the hypothesis that H-NS and LRP bind independently to the 5' and 3' *casA* regulatory regions. However, in order to elucidate if the two regulators are able to interact simultaneously with the *casA* regulatory sequence, EMSAs with both proteins were performed, showing that the H-NS-*casA* complex in the presence of increasing amounts of LRP was modified, with a new band visualized. This novel complex was of a higher mass than those obtained with the H-NS and LRP proteins independently, indicating that both regulators bound simultaneously to the same fragment of DNA (Fig. 5E). It is relevant to mention that although the two proteins interacted with

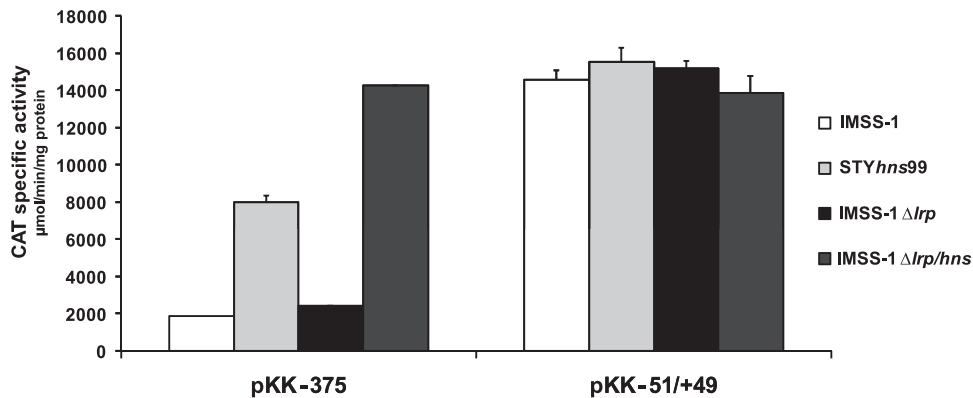


FIG. 6. Transcriptional profile of the *casA* promoter region in N-minimal medium. The expression profiles of the pKK-375 and pKK-51/+49 fusions were evaluated in the wild-type *S. Typhi* IMSS-1, STYhns99, IMSS-1  $\Delta$ lrp, and IMSS-1  $\Delta$ lrp/hns strains. The activities were determined at 16 h in N-minimal medium. The values are the means and standard deviations for at least three independent experiments performed in duplicate.

the same fragment, they did not appear to displace each other, thus demonstrating the simultaneous involvement of H-NS and LRP in the repression of the CRISPR/Cas system in *S. Typhi*.

**The CRISPR/Cas genetic system is induced in *S. Typhi* IMSS-1 grown in N-minimal medium.** The CRISPR/Cas immune system is widely distributed in prokaryotes; however, few studies point out the natural conditions that induce its expression. Therefore, it is of interest to investigate the signals or conditions that activate this genetic system. For *S. Typhi*, the selective capture of genes expressed in macrophages showed that the third component of the CRISPR/Cas system, STY3068 (*casC* ortholog), is induced under these conditions (17), and according to Deiwick et al. (15), a minimal medium deficient in magnesium and phosphate simulates the environmental niche encountered by *Salmonella* in the eukaryotic cell, since several genes of *Salmonella* pathogenicity island 2 that are relevant to the survival of this pathogen inside macrophages are induced in this medium (10).

To elucidate whether the *S. Typhi* CRISPR/Cas program is induced under conditions that simulate the environment inside the macrophage, the expression of the *casA* regulatory region was evaluated in N-minimal medium. The results showed that in N-minimal medium, the pKK-375 fusion gave 1,800 activity units in the wild-type strain, while its activity was only 34 activity units in rich MA medium (Fig. 4 and 6). Interestingly, in N-minimal medium, we observed that the expression values for the pKK-375 fusion in the *hns* and *lrp* single null and double *hns/lrp* null strains were 8,000, 2,400, and 14,000 activity units, respectively, supporting the negative role of H-NS and LRP in the transcriptional expression of *casA* in N-minimal medium (Fig. 6). In the wild type and the *hns*, *lrp*, and double *hns/lrp* mutants, the activity values for the pKK-51/+49 fusion (which contains the promoter region without *cis*-acting negative elements) were around 14,000 to 15,000 (Fig. 6), supporting the conclusion that H-NS and LRP are the main repressors of *casA* expression.

To determine the promoter that activates the expression of *casA* in N-minimal medium, 5'-RACE experiments were performed. RNA from the wild-type strain grown in N-minimal medium was obtained, and a specific primer (*casA* 5'-RACE)

located 83 bp downstream of the ATG initiation codon of *casA* was used for the determination of the transcription initiation site (see Table S2 in the supplemental material). The experiment showed that for cells grown in N-minimal medium, a transcriptional start site was located 94 bp upstream of the ATG translational codon. Interestingly, this initiation site coincides with the LeuO-dependent transcriptional site observed in rich medium (MA medium). These data indicate that transcriptional expression in MA medium in the presence of overexpressed LeuO, as well as that in N-minimal medium in the absence of overexpressed LeuO, is driven by the same promoter. Additional experiments validated this assertion, since with a fusion that harbors a mutation in the Pribnow box [pKK9-*casA* (-10 ATA  $\rightarrow$  GGG)] of the transcriptional start site of *casA*, null CAT activity was detected in the presence of LeuO in either MA medium or N-minimal medium (data not shown). Furthermore, a fusion of 62 bp (nucleotides -51 to +11), which contained 51 and 11 bp upstream and downstream, respectively, of the transcription initiation site of *casA*, showed activity values of 5,232 (Fig. 4) and 16,470 (data not shown) in rich and minimal medium, respectively, supporting the idea that a common promoter drives *casA* expression in both media. To determine whether LeuO is the *casA* activator in N-minimal medium, the pKK-375 fusion was introduced into the IMSS-1  $\Delta$ leuO strain: we found that LeuO does not induce *casA* expression in N-minimal medium, since the activity values in the *leuO*-deficient strain were the same as those in the wild-type strain (data not shown). However, LeuO could increase transcriptional activity in N-minimal medium, as overexpression of the *leuO* gene resulted in a 2-fold induction of *casA* transcription compared to an absence of overexpression in the wild-type *Salmonella* strain (data not shown). These data support the role of LeuO in N-minimal medium and MA rich medium as an inducer of *casA* expression. However, *casA* can also be expressed in a LeuO-independent mode in N-minimal medium, which reveals that additional regulators are needed to fully activate *casA* under these conditions. Future research is needed to identify the full repertoire of genetic elements that control *casA* activation in both N-minimal medium and MA medium.

## DISCUSSION

Three distinct stages are involved in CRISPR/Cas functionality: integration of the spacers, expression of CRISPR/Cas, and interference with the foreign DNA. In this study, we analyzed the transcriptional expression of the CRISPR/Cas system in *S. Typhi* IMSS-1. Previously, we reported that two putative promoters control *casA* expression in *S. Typhi* IMSS-1 and that the initiation sites are located 94 and 84 bp upstream of the ATG translational start codon (28). In this work, we performed 5'-RACE experiments that showed a single chromosomal transcription initiation site (Fig. 1B). By site-directed mutagenesis of the  $-10$  box, we demonstrated that a unique  $+1$  mRNA is located 94 bp upstream of the ATG translational start codon of *casA* (Fig. 1C). The promoter region reported in this work is induced in the presence of the global regulator LeuO. Site-directed mutagenesis of the LeuO binding sites supports the positive role of this protein in the expression of *casA*, and we showed that both LBS I and LBS II are determinants of LeuO induction. Interestingly, the ATA bases located at positions  $-88$  to  $-90$  (LBS I) are key for LeuO activation (Fig. 3B). Previously, we demonstrated that a TATgTcATAT region located at positions  $-125$  to  $-134$  with respect to the transcriptional start site was necessary for the LeuO induction of the *ompS1* gene (16). These data provide evidence that this regulator does not necessarily recognize a consensus sequence (28). Additional efforts are needed to elucidate the rules that LeuO follows for DNA recognition and transcriptional activation.

Using transcriptional fusions, RT-PCR, and Northern experiments, we demonstrated that the entire CRISPR/Cas system is induced as a large polycistronic mRNA in rich medium in the presence of LeuO, supporting an operon organization (Table 1; Fig. 2) that forms part of the LeuO regulon in *S. Typhi*. Interestingly, Northern blot experiments with *E. coli* showed that the *casA-casB-casC-casD-casE-casI-cas2* genes are part of one transcriptional unit (58). It has also been reported by Pul et al. (55) that the CRISPR sequences contain their own promoter in *E. coli*. Thus, in contrast to the CRISPR/Cas operon organization in *S. Typhi*, *E. coli* has two transcriptional units, one driving the genetic expression of *casABCDEI2* and the other inducing CRISPR transcription. This result provides evidence that the genetic control of the CRISPR/Cas program is highly regulated in the *Enterobacteriaceae* and that even in two closely related bacteria, such as *S. Typhi* and *E. coli*, the functional transcriptional organization is different, raising the possibility that the CRISPR/Cas system can be fine-tuned in response to multiple environmental inputs.

Various studies have shown that genes upregulated by LeuO are downregulated by H-NS (28, 29, 39, 40, 58, 64). A recent report showed that *casA* and CRISPR expression is repressed independently by H-NS in *E. coli* (55). In contrast, in the case of *S. Typhi* IMSS-1, a fusion containing the leader and the CRISPR locus showed no effect in an H-NS-deficient strain (data not shown), while H-NS repression was observed in the transcriptional expression of *casA*. These data again point out the differences in CRISPR/Cas transcriptional control in *E. coli* versus *S. Typhi*. Another notable difference is the finding that the global regulatory protein LRP is a negative repressor of *casA* in *S. Typhi* (Fig. 4), while Pul et al. (55) reported that

LRP did not have a significant effect on the expression of the *cas* genes in *E. coli*.

The involvement of H-NS, LRP, and a LysR-type regulator in the control of other *Salmonella* genes has been documented. For instance, the *spv* virulence genes are positively regulated by the LysR-type regulator SpvR and repressed by H-NS and LRP (45, 50). A negative role of H-NS and LRP is also reported for the 16S rRNA genes, where it is postulated that these proteins bind upstream and downstream of the transcription initiation site and work together in a synergic fashion to repress the expression of the 16S rRNA genes (53, 54). Similar to these studies, we found that H-NS and LRP bind upstream and downstream of the transcription initiation site to repress *casA* (Fig. 5C and D). Interestingly, both proteins appear to simultaneously bind the *casA* regulatory region (Fig. 5E), and it is possible that a nucleosome structure formed between H-NS, LRP, and the target promoter is needed to exert repression of CRISPR/Cas, as suggested for the 16S rRNA repression mediated by H-NS and LRP (53, 54).

Future research is needed to elucidate the fine-tuning mechanism of *casA* repression mediated by H-NS and LRP, and it would also be interesting to determine how LeuO in the presence of these proteins induces *casA* activity. The transcriptional results in the absence of H-NS and LRP showed *casA* activity values of around 5,000 in MA medium (Fig. 4), while the expression value obtained in the presence of LeuO was only 500 (Fig. 1C), indicating that LeuO moderately induces *casA* transcriptional expression and that some additional positive regulatory elements or signals are needed to obtain its complete induction. We also found that the *casA* promoter region is highly induced in N-minimal medium, in a LeuO-independent manner (Fig. 6). Nevertheless, LeuO can also enhance *casA* activity in N-minimal medium if it is overexpressed from a synthetic promoter.

The participation of global regulatory proteins such as LeuO, H-NS, and LRP, which have been implicated in virulence metabolism and other biological processes, suggests that the CRISPR/Cas system responds to a wide variety of environmental cues. It was reported that the CRISPR/Cas program is involved in sporulation in *Myxococcus xanthus* (63) and in biofilm formation in *Pseudomonas aeruginosa* (65) and also that housekeeping genes can be targets of CRISPR interference (2). The fact that the CRISPR/Cas system is induced under stress conditions such as growth in N-minimal medium suggests that it could have a role in pathogenesis. Further studies of mutants inactivated in the *cas* genes are needed to evaluate their role in infection and virulence in *Salmonella*.

## ACKNOWLEDGMENTS

This research was supported by grants to I.H.-L. from DGAPA/UNAM (IN214808) and CONACYT (89337 and 127298) and by grants from CONACYT, México (82383), and DGAPA-UNAM (216310) to E.C.

We thank S. J. Jaime-Rodríguez, M. L. Zavala-García, N. Becerra-Lobato, F. J. Santana, M. Fernández-Mora, P. Gaytan, E. Bustos, S. Becerra, J. Yañez, and J. A. Ramirez-Trujillo for technical help and A. Gutierrez-Preciado, A. Medrano-López, J. M. Villarreal, M. A. De la Cruz, R. Oropeza, J. Caballero, M. Dunn, and J. Miranda for critical readings of the manuscript and stimulating discussions.



## REFERENCES

- Agari, Y., et al. 2010. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J. Mol. Biol.* **395**:270–281.
- Aklujkar, M., and D. R. Lovley. 2010. Interference with histidyl-tRNA synthetase by a CRISPR spacer sequence as a factor in the evolution of *Pelobacter carbinolicus*. *BMC Evol. Biol.* **10**:230.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Altschul, S. F., et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Amann, E., O. Birgit, and A. Karl-Josef. 1988. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**:301–315.
- Andersson, A. F., and J. F. Banfield. 2008. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**:1047–1050.
- Barrangou, R., et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**:1709–1712.
- Bolotin, A., B. Quinquis, A. Sorokin, and S. D. Ehrlich. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**:2551–2561.
- Brouns, S. J., et al. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**:960–964.
- Bustamante, V. H., et al. 2008. HilD-mediated transcriptional cross-talk between SPI-1 and SPI-2. *Proc. Natl. Acad. Sci. U. S. A.* **105**:14591–14596.
- Carte, J., R. Wang, H. Li, R. M. Terns, and M. P. Terns. 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22**:3489–3496.
- Chen, C. C., et al. 2003. LeuO-mediated transcriptional derepression. *J. Biol. Chem.* **278**:38094–38103.
- Chen, C. C., M. Y. Chou, C. H. Huang, A. Majumder, and H. Y. Wu. 2005. A cis-spreading nucleoprotein filament is responsible for the gene silencing activity found in the promoter relay mechanism. *J. Biol. Chem.* **280**:5101–5112.
- Datsenko, K., and B. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**:6640–6645.
- Deiwick, J., T. Nikolaus, S. Erdogan, and M. Hensel. 1999. Environmental regulation of *Salmonella* pathogenicity island 2 gene expression. *Mol. Microbiol.* **31**:1759–1773.
- De la Cruz, M. A., et al. 2007. LeuO antagonizes H-NS and StpA-dependent repression in *Salmonella enterica* *ompS1*. *Mol. Microbiol.* **66**:727–743.
- Faucher, S. P., R. Curtiss III, and F. Daigle. 2005. Selective capture of *Salmonella enterica* serovar Typhi genes expressed in macrophages that are absent from the *Salmonella enterica* serovar Typhimurium genome. *Infect. Immun.* **73**:5217–5221.
- Fernández-Mora, M., J. L. Puente, and E. Calva. 2004. OmpR and LeuO positively regulate the *Salmonella enterica* serovar Typhi *ompS2* porin gene. *J. Bacteriol.* **186**:2909–2920.
- Flores-Valdez, M. A., J. L. Puente, and E. Calva. 2003. Negative osmoregulation of the *Salmonella ompS1* porin gene independently of OmpR in an *hns* background. *J. Bacteriol.* **185**:6497–6506.
- Forde, A., and G. F. Fitzgerald. 1999. Bacteriophage defence systems in lactic acid bacteria. *Antonie Van Leeuwenhoek* **76**:89–113.
- Garneau, J. E., et al. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**:67–71.
- Grissa, I., G. Vergnaud, and C. Pourcel. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **35**:W52–W57.
- Gupta, S., J. A. Stamatoyannopoulos, T. L. Bailey, and W. S. Noble. 2007. Quantifying similarity between motifs. *Genome Biol.* **8**:R24.
- Haft, D. H., J. Selengut, E. F. Mongodin, and K. E. Nelson. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1**:e60.
- Hale, C. R., et al. 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**:945–956.
- Haurwitz, R. E., M. Jinek, B. Wiedenheft, K. Zhou, and J. A. Doudna. 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**:1355–1358.
- Heidelberg, J. F., W. C. Nelson, T. Schoenfeld, and D. Bhaya. 2009. Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS One* **4**:e4169.
- Hernández-Lucas, I., et al. 2008. The LysR-type transcriptional regulator LeuO controls the expression of several genes in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* **190**:1658–1670.
- Hommais, F., et al. 2001. Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol. Microbiol.* **40**:20–36.
- Horvath, P., and R. Barrangou. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**:167–170.
- Jansen, R., J. D. van Embden, W. Gastra, and L. M. Schouls. 2002. Identification of a novel family of sequence repeats among prokaryotes. *OMICS* **6**:23–33.
- Jansen, R., J. D. van Embden, W. Gastra, and L. M. Schouls. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43**:1565–1575.
- Kaminuma, E., et al. 9 November 2010. DDBJ progress report. *Nucleic Acids Res.* doi:10.1093/nar/gkq1041.
- Karginov, F. V., and G. J. Hannon. 2010. The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol. Cell* **37**:7–19.
- Kawaji, H., T. Mizuno, and S. Mizushima. 1979. Influence of molecular size and osmolarity of sugars and dextrans on the synthesis of outer membrane proteins O-8 and O-9 of *Escherichia coli* K-12. *J. Bacteriol.* **140**:843–847.
- Kazakov, A. E., et al. 2007. RegTransBase—a database of regulatory sequences and interactions in a wide range of prokaryotic genomes. *Nucleic Acids Res.* **35**:D407–D412.
- Labrie, S. J., J. E. Samson, and S. Moineau. 2010. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**:317–327.
- Lillestøl, R. K., et al. 2009. CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. *Mol. Microbiol.* **72**:259–272.
- Lucchini, S., et al. 2006. H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.* **2**:e81.
- Madhusudan, S., A. Paukner, Y. Klingen, and K. Schnetz. 2005. Independent regulation of H-NS-mediated silencing of the *bgl* operon at two levels: upstream by BglJ and LeuO and downstream by DnaKJ. *Microbiology* **151**:3349–3359.
- Makarova, K. S., N. V. Grishin, S. A. Shabalina, Y. I. Wolf, and E. V. Koonin. 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1**:7.
- Marraffini, L. A., and E. J. Sontheimer. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**:1843–1845.
- Marraffini, L. A., and E. J. Sontheimer. 2010. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**:568–571.
- Marraffini, L. A., and E. J. Sontheimer. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* **11**:181–190.
- Marshall, D. G., B. J. Sheehan, and C. J. Dorman. 1999. A role for the leucine-responsive regulatory protein and integration host factor in the regulation of the *Salmonella* plasmid virulence (*spv*) locus in *Salmonella typhimurium*. *Mol. Microbiol.* **34**:134–145.
- Martínez-Laguna, Y., E. Calva, and J. L. Puente. 1999. Autoactivation and environmental regulation of *bfpT* expression, the gene coding for the transcriptional activator of *bfpA* in enteropathogenic *Escherichia coli*. *Mol. Microbiol.* **33**:153–166.
- Mayer, M. P. 1995. A new set of useful cloning and expression vectors derived from pBlueScript. *Gene* **163**:41–46.
- Mendoza-Vargas, A., et al. 2009. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One* **4**:e7526.
- Mojica, F. J., C. Díez-Villaseñor, J. García-Martínez, and E. Soria. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**:174–182.
- O'Byrne, C. P., and C. D. Dorman. 1994. Transcription of the *Salmonella typhimurium* *spv* virulence locus is regulated negatively by the nucleoid-associated protein H-NS. *FEMS Microbiol. Lett.* **12**:99–106.
- Pourcel, C., G. Salvignol, and G. Vergnaud. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**:653–663.
- Puente, J. L., V. Flores, M. Fernández, Y. Fuchs, and E. Calva. 1987. Isolation of an *ompC*-like outer membrane protein gene from *Salmonella typhi*. *Gene* **61**:75–83.
- Pul, U., et al. 2005. LRP and H-NS cooperative partners for transcription regulation at *Escherichia coli* rRNA promoters. *Mol. Microbiol.* **258**:864–876.
- Pul, U., R. Wurm, and R. Wagner. 2007. The role of LRP and H-NS in transcription regulation: involvement of synergism, allosteric and macromolecular crowding. *J. Mol. Biol.* **366**:900–915.
- Pul, U., et al. 2010. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* **7**:1495–1512.
- Robison, K., A. M. McGuire, and G. M. Church. 1998. A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete *Escherichia coli* K-12 genome. *J. Mol. Biol.* **284**:241–254.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Shimada, T., K. Yamamoto, and A. Ishihama. 2009. Involvement of the

- leucine response transcription factor LeuO in regulation of the genes for sulfa drug efflux. *J. Bacteriol.* **191**:4562–4571.
59. **Stratmann, T., S. Madhusudan, and K. Schnetz.** 2008. Regulation of the *yjyQ-bglJ* operon, encoding LuxR-type transcription factors, and the divergent *yjyP* gene by H-NS and LeuO. *J. Bacteriol.* **190**:926–935.
60. **Ueguchi, C., T. Ohta, C. Seto, T. Suzuki, and T. Mizuno.** 1998. The *leuO* gene product has a latent ability to relieve *bgl* silencing in *Escherichia coli*. *J. Bacteriol.* **180**:190–193.
61. **van der Oost, J., M. M. Jore, E. R. Westra, M. Lundgren, and S. J. Brouns.** 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* **34**:401–407.
62. **van Helden, J., B. André, and J. Collado-Vides.** 1998. Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. *J. Mol. Biol.* **281**:827–842.
63. **Viswanathan, P., K. Murphy, B. Julien, A. G. Garza, and L. Kroos.** 2007. Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J. Bacteriol.* **189**:3738–3750.
64. **Westra, E. R., et al.** 2010. H-NS mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* **77**:1380–1393.
65. **Zegans, M. E., et al.** 2009. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* **191**:210–219.

# CRISPR-Cas system presents multiple transcriptional units including antisense RNAs that are expressed in minimal medium and upregulated by pH in *Salmonella enterica* serovar Typhi

Liliana Medina-Aparicio, Javier E. Rebollar-Flores, América A. Beltrán-Luviano, Alejandra Vázquez, Rosa M. Gutiérrez-Ríos, Leticia Olvera, Edmundo Calva and Ismael Hernández-Lucas\*

## Abstract

The CRISPR-Cas system is involved in bacterial immunity, virulence, gene regulation, biofilm formation and sporulation. In *Salmonella enterica* serovar Typhi, this system consists of five transcriptional units including antisense RNAs. It was determined that these genetic elements are expressed in minimal medium and are up-regulated by pH. In addition, a transcriptional characterization of *cas3* and *asc2-1* is included herein.

## INTRODUCTION

Prokaryotes have developed strategies to survive phage invasion, including the CRISPR-Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins) system that provides specific heritable immunity against these invaders [1, 2]. The expression of CRISPR-*cas* generates small RNAs (crRNAs) that, by homology, recognize genetic material of phage and facilitate cleavage of the invader by Cas proteins [3, 4]. These systems are classified according to the presence of signature Cas proteins [5]. The hallmark of the CRISPR-Cas Type I system is the presence of Cas3, which contains an N-terminal ssDNA nuclease and a C-terminal DExH helicase domains [6, 7]. This protein is involved in cleavage of exogenous target nucleic acids [7, 8]. The Type II system requires Cas9 and a *trans*-activating CRISPR RNA (tracrRNA) for DNA recognition and degradation [9]. Based on its few requirements, this system has been used for genomic engineering in prokaryotes and eukaryotes [10]. The Type III system uses a complex composed of the RAMP proteins and Cas10 nuclease to silence the invader [11, 12].

The CRISPR-Cas systems are widely distributed in bacteria [13] and are also involved in virulence [14, 15], fruiting body development [16], swarming motility, biofilm

formation [17] and response to stress conditions such as the perturbation of the bacterial envelope [18–20] or exposure to  $\gamma$ -radiation [21].

*Salmonella enterica* serovar Typhi (*S. Typhi*) IMSS-1, a clinical strain that causes typhoid fever, contains a Type I-E CRISPR-Cas cluster composed of *cas3*, *cse1-cse2-cas7-cas5-cas6-cas1-cas2*, an 84 bp leader sequence, seven 29 bp repeats and six 32 bp spacers [22]. The conserved genetic organization of the *cas* genes in some *Salmonella* serovars is consistent with its having a biological function in these bacteria [23]. Interestingly, the expression of *cas7* (STY3068) has been detected when *S. Typhi* infects macrophages [24]. Additionally, we previously reported the transcriptional expression of *cse1* from this bacterium in N-minimal medium (N-MM) [22], which promotes the expression of the *Salmonella* pathogenicity island 2 genes [25]. Furthermore, the *S. Typhi* CRISPR-Cas locus is regulated by LeuO, H-NS and LRP [22, 26], which are involved in pathogenesis [27–32]. A role of H-NS and LeuO in the transcriptional regulation of the Type I-E CRISPR-Cas system from *Escherichia coli* K12 has also been reported [33, 34].

Since some evidence for the activity of the CRISPR-Cas system in virulence has been reported, we evaluated the expression of the *S. Typhi* IMSS-1 CRISPR-Cas locus in N-MM,

Received 8 August 2016; Accepted 17 December 2016

**Author affiliation:** Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Cuernavaca, Morelos 62210, Mexico.

**\*Correspondence:** Ismael Hernández-Lucas, ismaelh@ibt.unam.mx

**Keywords:** *S. Typhi*; CRISPR-Cas; regulation; pH; antisense RNA.

**Abbreviations:** Cas, CRISPR-associated; CAT, chloramphenicol acetyltransferase; CRISPR, clustered regularly interspaced short palindromic repeat; CRP, cAMP receptor protein; DTNB, 5,5'-dithiobis(2-nitrobenzoic acid); EMSA, electrophoretic mobility shift assays; MA, medium A; N-MM, N-minimal medium; RT-PCR, reverse transcriptase PCR; *S. Typhi*, *Salmonella enterica* serovar Typhi.

Two supplementary tables and two supplementary figures are available with the online Supplementary Material.

finding five transcriptional units. Two of them, the *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR (*cas*-CRISPR operon) and *scse2* (sense *cse2* RNA), are located in the sense strand, whereas *ascse2-1* (antisense RNA of *cse2* to *cse1*) and *ascas2-1* (antisense RNA of *cas2* to *cas1*) are present on the antisense strand. Additionally, the *cas3* gene is transcribed as an independent unit divergent to the *cas*-CRISPR operon. Interestingly, these five genetic elements showed higher transcriptional expression at pH 7.5, a pH value characteristic of the human terminal ileum [35].

In this work, we also report that H-NS represses *ascse2-1* and *cas3* expression and that a 13 nt sequence (5'-TAATC-CAGACAAA-3') is involved in *cas3* positive control.

## METHODS

### Bacterial strains, plasmids and culture conditions

The bacterial strains and plasmids used in this work are listed in Table S1 (available in the online Supplementary Material). *S. Typhi* [36] and *E. coli* strains were grown aerobically at 37°C in LB (10 g tryptone, 5 g yeast extract and 10 g NaCl per litre), MA (7 g nutrient broth, 1 g yeast extract, 2 ml glycerol, 3.75 g K<sub>2</sub>HPO<sub>4</sub> and 1.3 g KH<sub>2</sub>PO<sub>4</sub> per litre) [37] or N-MM (pH 7.5) {0.37 g KCl, 0.99 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.087 g K<sub>2</sub>SO<sub>4</sub>, 0.14 g KH<sub>2</sub>PO<sub>4</sub>, 0.019 g MgCl<sub>2</sub>, 1 g casamino acids, 5 ml glycerol and 100 mM Tris/HCl to get pH 7.5 per litre [25]} medium. The pH 7.5 of regular N-MM containing 100 mM Tris/HCl decreases to 7.0 at OD<sub>595</sub> 1.3; therefore, for pH experiments, 200 mM Tris/HCl was used to avoid pH changes during bacterial growth. When required, the following antibiotics were added: kanamycin, 30 µg ml<sup>-1</sup>; tetracycline, 12 µg ml<sup>-1</sup>; and ampicillin, 200 µg ml<sup>-1</sup>.

### DNA and RNA manipulations

Plasmid and genomic DNA isolations were carried out according to published protocols [38]. Primers for PCR amplifications were provided by the Oligonucleotide Synthesis Facility at our institute (Table S2). Restriction enzymes, ligase, nucleotides and polymerases were acquired from New England Biolabs, Invitrogen or Thermo Scientific. For sequencing, dsDNA was purified with the High Pure Plasmid Isolation Kit (Roche) and sequenced with an automatic Perkin-Elmer/Applied Biosystems 377-18 system.

Total RNA for reverse transcriptase PCR (RT-PCR) and primer extension experiments was purified from 20 ml of *S. Typhi* N-MM cultures (OD<sub>595</sub> 1.3) using the acid/phenol method. The concentration of RNA was determined by measuring the absorbance at 260 nm and the integrity of RNA was determined in a 1.5 % agarose gel [39].

### Construction of *S. Typhi* mutants

Mutants were obtained by the one-step nonpolar mutagenesis procedure described by Datsenko and Wanner [40]. The target gene was replaced with selectable antibiotic resistance gene markers. The resistance cassette was removed using the pCP20 plasmid. Each mutation was further

characterized by sequencing to verify the authenticity of the deletion.

### Construction of transcriptional reporter constructs

For transcriptional *cat* constructs, oligonucleotides (see Table S2) were designed to amplify DNA fragments of different lengths. PCR products were double-digested with *Bam*HI-*Kpn*I, *Bam*HI-*Hind*III or *Bam*HI-*Xho*I and ligated into pKK232-8 or pKK232-9 (Table S1), which contains the promoterless *cat* gene. For transcriptional *lacZ* reporter constructs, PCR fragments were amplified, digested with *Bam*HI and cloned into pRS415 plasmid [41] (Table S1). All constructs were sequenced to verify the correct DNA sequence of the PCR fragments.

### Chloramphenicol acetyltransferase assays

To determine the genetic expression of the *cat* reporter gene mediated by the *S. Typhi* promoters, chloramphenicol acetyltransferase (CAT) assays were performed according to Martínez-Laguna *et al.* [42]. Briefly, *S. Typhi* strains harbouring the reporters were grown in N-MM to different optical densities. Cells were harvested, centrifuged, washed with 0.8 ml of TDTT buffer [50 mM Tris/HCl and 30 µM DL-DTT (pH 7.8)], resuspended in 0.5 ml of TDTT and sonicated on ice for 9.9 s intervals with 9.9 s rest periods until the extract was clear. The homogenate was centrifuged, and the supernatant was used for activity measurement. For CAT assays, 5 µl of each extract was added in duplicate to a 96-well ELISA plate, followed by the addition of 0.2 ml of a reaction mixture containing 1 mM DTNB [5,5'-dithiobis(2-nitrobenzoic acid)], 0.1 mM acetyl-CoA and 0.1 mM chloramphenicol in 0.1 M Tris/HCl (pH 7.8). The absorbance at 412 nm was measured every 5 s for 5 min using a Ceres 900 scanning autoreader and microplate workstation. The protein concentration of the cell extracts was obtained using the bicinchoninic acid protein assay reagent (Pierce). Protein values and the mean rate of product formation by CAT were used to determine CAT-specific activity as micromoles per minute per milligram of protein.

### β-Galactosidase assays

To determine the genetic expression of the reporter *lacZ* gene mediated by *S. Typhi* promoters, β-galactosidase activity was measured by the Miller method adapted for a microtitre plate format as described by Oropeza *et al.* [43]. Protein concentration was determined by the Lowry method. The results of transcriptional reporters presented are the mean of three independent experiments performed in duplicate.

### Primer extension analysis

To identify transcriptional start sites, primer extension experiments were performed as follows. Forty microgram of total RNA was denatured at 65°C for 5 min and then annealed with the corresponding [γ-<sup>32</sup>P] ATP-labelled primers. Primers were extended with reverse transcriptase at 50°C for 30 min as indicated in the Maxima H Minus Strand cDNA Synthesis Kit (Thermo Scientific) protocol.

The extended products were precipitated and analysed by electrophoresis in 8 % polyacrylamide/8 M urea gels alongside sequencing ladders [43].

### RT-PCR

Experiments were performed according to Villarreal *et al.* [44]. Briefly, 20 µg of total RNA was treated with DNase I (Thermo Scientific), and 1 µg of RNA (DNA-free as assessed by PCR) was used for cDNA synthesis with specific primers as established in the RevertAid H Minus First Strand cDNA Synthesis kit protocol (Thermo Scientific). Finally, PCRs were performed using 2 µl of cDNA as a template, and the reaction products were analysed in 1.0 % or 1.8 % agarose gels. For all the RT-PCR experiments, a cDNA synthesis negative control reaction without reverse transcriptase was carried out.

### Electrophoretic mobility shift assays

The H-NS protein was purified as described previously in De la Cruz *et al.* [45] and Medina-Aparicio *et al.* [22] and non-radioactive electrophoretic mobility shift assays (EMSA) were performed. The probes were obtained by PCR using primers described in Table S2. Each probe (100 ng) was mixed with increasing concentrations of purified protein in the presence of H-NS-binding buffer (40 mM HEPES, 8 mM MgCl<sub>2</sub>, 50 mM KCl, 1 mM DTT, 0.05 % NP40 and 0.1 mg ml<sup>-1</sup> BSA). The mixture was incubated for 20 min at room temperature and then separated by native 6 % PAGE in 0.5× Tris/borate/EDTA buffer. The DNA bands were visualized by ethidium bromide staining.

## RESULTS

### Negative and positive genetic elements involved in *S. Typhi cas3* transcriptional expression

Several studies regarding Cas3 structure, from other bacteria, and its enzymatic mechanism are available [46–48]. However, studies on its transcriptional regulation in *Salmonella* have not been performed. Therefore, a *cas3* transcriptional characterization in *S. Typhi* is reported here.

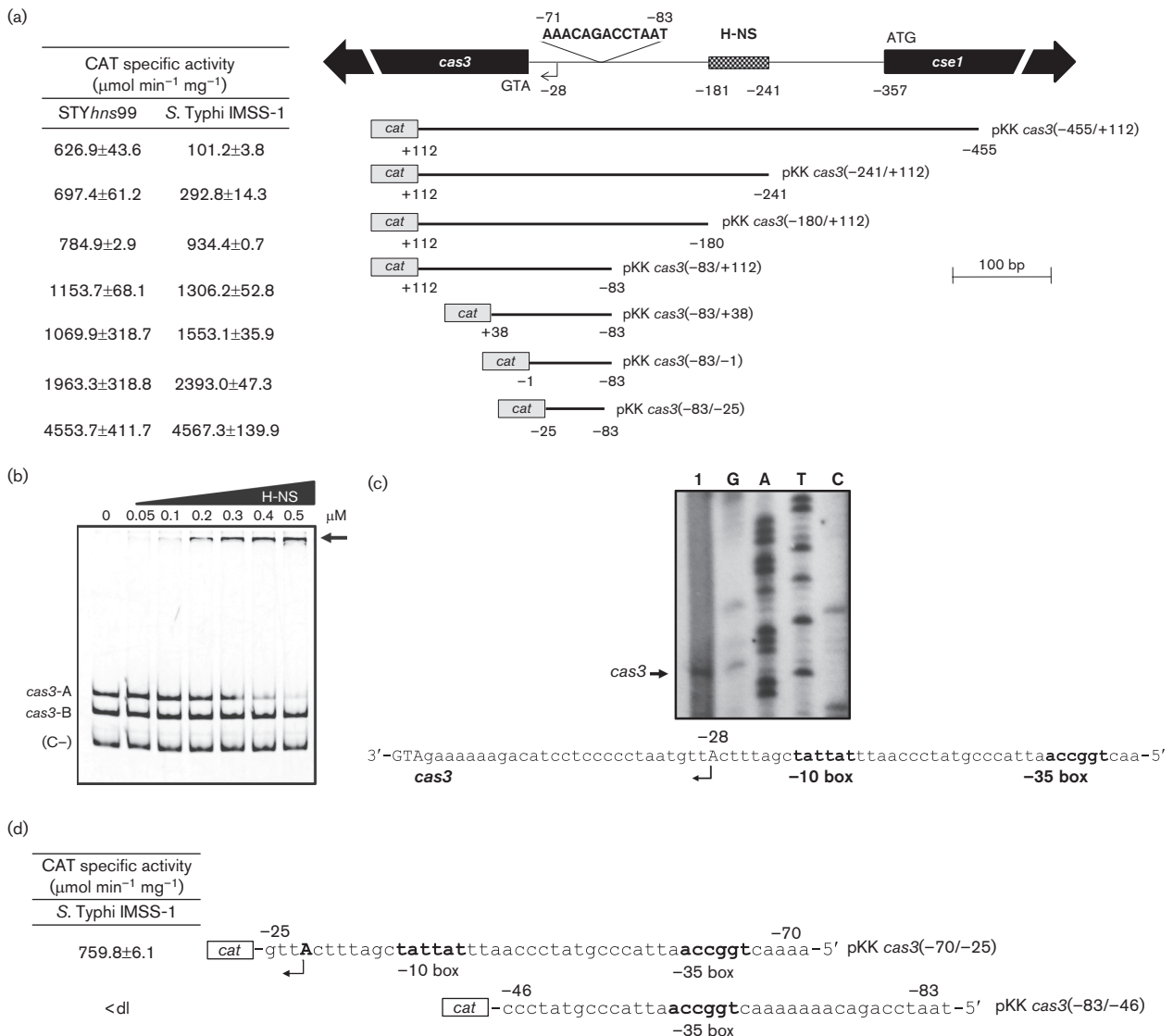
Previously, we have determined that the *cse1* promoter is repressed by the global regulators H-NS and LRP whereas LeuO participates in its positive regulation; also, it can be expressed independently of LeuO in N-MM [22]. In *S. Typhi*, *cas3* is divergent to *cse1*, sharing a 5' intergenic regulatory region of 357 bp (Fig. 1a). Therefore, we evaluated *cas3* expression in N-MM. The pKK *cas3* (–455/+112) transcriptional reporter construct, containing 455 bp upstream and 112 bp downstream of the *cas3* translational start site, was evaluated in the *S. Typhi* IMSS-1 wild-type. *cas3* expression was of 101 CAT units (Fig. 1a). This low transcriptional activity suggests that *cas3* promoter is repressed in N-MM. Thus, to identify *cas3* regions involved in its repression, transcriptional reporters encompassing different lengths of the *cas3* regulatory region were analysed in *S. Typhi* grown in N-MM to OD<sub>595</sub> 1.3. Fig. 1(a) shows that constructs pKK *cas3* (–241/+112), pKK *cas3* (–180/+112) and pKK *cas3* (–83/+112) had activities of 293, 934 and 1306 CAT units, respectively. These values

were higher than those obtained with the longest reporter pKK *cas3* (–455/+112) (101 CAT units), indicating that nucleotides –455 to –82 are involved in *cas3* negative regulation. Additional constructs downstream of this *cas3* repressing region, pKK *cas3* (–83/+38), pKK *cas3* (–83/–1) and pKK *cas3* (–83/–25), produced 1553, 2393 and 4567 CAT units, respectively (Fig. 1a), showing that the DNA region between –24 and +112 is also implicated in *cas3* repression. Thus, two regions (–455 to –82 and –24 to +112 relative to the *cas3* translational start codon) contain DNA information involved in *cas3* silencing, indicating that *cas3* is negatively regulated downstream and upstream of its protein start codon.

In order to determine the transcriptional factors responsible for *cas3* negative regulation, its transcriptional activity was assessed in individual mutants lacking the H-NS [49], LRP, CRP, CpxR, SsrB, IHF, PhoP, PmrA, FNR or OmpR regulators, particularly since some of them (H-NS or CRP) have been previously implicated in the regulation of other CRISPR-Cas systems [33, 50, 51] or in virulence of *Salmonella* (CpxR, SsrB, IHF, PhoP, PmrA, FNR or OmpR) [52–58]. We found that only H-NS represses *cas3* expression (Fig. S1), since an increase of CAT activity using pKK *cas3* (–455/+112) (627 units) was observed in the *hns*-deficient strain (STY*hns99*) with respect to the wild-type (101 CAT units, Fig. 1a). To further characterize H-NS negative regulation, we evaluated the CAT activity of the *cas3* transcriptional reporter constructs in the STY*hns99* strain. The experiments showed higher activity with pKK *cas3* (–241/+112) (697 CAT units) compared to the expression in the wild-type (293 CAT units). However, with plasmid pKK *cas3* (–180/+112), similar CAT levels (780–930 units) were observed in the wild-type and the STY*hns99* strain (Fig. 1a), indicating that H-NS represses *cas3* expression by acting between nucleotides –455 and –181 relative to its ATG codon. These data are in agreement with the previous report that demonstrated the interaction of H-NS with this region in order to also repress the *cas3*-divergent *cse1* gene [22]. The results obtained also showed that H-NS-independent regulation involves bases –180 to +112, since in the wild-type background, the CAT values of this construct were nine times greater than the pKK *cas3* (–455/+112) reporter (Fig. 1a).

The data on *cas3* repression suggest that nucleotides –241 to –181, upstream of the *cas3* translational start site, are more relevant for negative regulation than those at –455 to –240, since CAT activity levels were higher, in the wild-type strain, when the sequence from –241 to –181 [pKK *cas3* (–180/+112)], was removed (Fig. 1a). Thus, it was determined whether these nucleotides corresponded to an H-NS-binding site for *cas3* repression. EMSA using purified H-NS protein and two *cas3* DNA fragments, *cas3*-A (nt –241 to +112) and *cas3*-B (nt –180 to +112), were performed. As expected, H-NS had more affinity to *cas3*-A region than for *cas3*-B, validating that this nucleoid protein directly represses *cas3* expression through binding at bases –241 to –181 (Fig. 1b).

Interestingly, the pKK *cas3* (–83/–25) plasmid, which lacks negative regulatory regions upstream and downstream of



**Fig. 1.** Genetic characterization of *cas3*. (a) The *cas3* gene is divergent to *cse1* and they share a 5' intergenic region of 357 bp (thin line). The bent arrow represents the *cas3* transcriptional start sites located 28 bp upstream of the *cas3* translational start codon. Nucleotides -83 to -71 (5'-TAATCCAGACAAA-3') are involved in *cas3* positive regulation. The hatched rectangle indicates the H-NS repression site located between nucleotides -241 and -181. Below the diagram, CAT transcriptional constructs of different length are shown. They were named (boldface) according to the nucleotides contained in each construct; the numbers below describe the coordinates upstream and downstream relative to *cas3* ATG start codon. The left-hand columns represent the expression of the *cas3* transcriptional reporters evaluated in the wild-type *S. Typhi* IMSS-1 and in the *hns* isogenic mutant (STY*hns99*). CAT-specific activities were measured in N-MM (100 mM Tris/HCl) at OD<sub>595</sub> 1.3. The values are the means±standard deviations of three independent experiments performed in duplicate. (b) Binding of H-NS to *cas3*. EMSA with increasing concentrations of purified H-NS-Myc-His6 and the *cas3*-A (nucleotides -241 to +112), as well as *cas3*-B (nucleotides -180 to +112) fragment. A 16S rRNA fragment of 242 bp was used as a negative control (C-). Arrow indicates the DNA-protein complex. (c) Mapping of *cas3* transcriptional start site. Primer extension experiments were performed with RNA from *S. Typhi* IMSS-1 harbouring pKK *cas3* (-83/-25) grown in N-MM (100 mM Tris/HCl) to OD<sub>595</sub> 1.3. The oligonucleotide of the vector pKK232-8 (pKK-R) was used to synthesize the cDNA which was resolved in an 8% polyacrylamide gel alongside sequencing ladders. The location of the transcriptional start site at 28 bp upstream of the *cas3* ATG start codon is shown in the sequence at the bottom of the figure (bent arrow); the -10 and -35 sequences (boldface) are also indicated. (d) Transcriptional reporter constructs used to validate the *cas3* promoter. The pKK *cas3* (-70/-25) contains the transcriptional start site (bent arrow) and the -10 and -35 sequences (boldface), whereas pKK *cas3* (-83/-46) only comprises the -35 box. Both plasmids were evaluated in *S. Typhi* IMSS-1 wild-type strain grown in N-MM (100 mM Tris/HCl) to OD<sub>595</sub> 1.3 and the results of the CAT assays are shown in the left columns. The values are the means±standard deviations for three independent experiments performed in duplicate. <dl (<detection limit) represents values between 0 and 10 CAT units.

the *cas3* ATG codon (–455 to –82 and –24 to +112), gave activity values of 4567 CAT units in the wild-type strain, whereas in the STY*hns99*, the expression levels with pKK *cas3* (–455/+112) were only of 627 CAT units (Fig. 1a). This suggests that there is another negative regulator for *cas3* besides H-NS.

As a part of the transcriptional characterization of *cas3*, an evaluation of the *cis* genetic elements involved in its positive expression was also performed. First, its transcriptional start site was determined by primer extension, using RNA from *S. Typhi* IMSS-1 harbouring plasmid pKK *cas3* (–83/–25), and grown in N-MM to OD<sub>595</sub> 1.3. The results showed a transcriptional start site localized 28 bp upstream of the *cas3* translational start codon, with consensus –10 and –35 sequences corresponding to the  $\sigma^{70}$  factor (Fig. 1c). To validate the *cas3* promoter region, we constructed pKK *cas3* (–70/–25) containing the transcriptional start site and the –10 and –35 sequences and pKK *cas3* (–83/–46) lacking the transcriptional start site and the –10 box. Both plasmids were evaluated in *S. Typhi* grown in N-MM. CAT activity was obtained only with pKK *cas3* (–70/–25) (760 units, Fig. 1d), validating the presence of the *cas3* promoter between nucleotides –70 and –25 from its translational start codon. Although this construct contains the elements necessary for *cas3* activity, the CAT values observed (760 CAT units) were six times lower than the expression levels of plasmid pKK *cas3* (–83/–25) (4567 CAT units, see Fig. 1a), indicating that nucleotides –83 to –71 (5'-TAATCCAGACAAA-3') are relevant for its transcriptional expression in N-MM (Fig. 1a). These 13 nucleotides are adjacent to the –35 sequence, a characteristic location for the UP element, which corresponds to the binding site for RNA polymerase  $\alpha$ -subunit and contains a high percentage of A/T [59]. EMSA performed with the region involved in *cas3* positive regulation and purified  $\alpha$ -subunit showed no interaction between these two elements (data not shown), suggesting that the 5'-TAATCCAGACAAA-3' nucleotides could be a binding site for a transcriptional activator. Since LeuO positively regulates the *cas3*-divergent *cse1* gene, we evaluated the role of this LysR regulator in *cas3* expression. The results showed that LeuO overexpression did not promote *cas3* transcriptional activity (data not shown).

In summary, we have identified the presence of negative regulatory elements upstream and downstream of the *cas3* translational start site. H-NS participates in its silencing by interacting with nucleotides –241 to –181 from the *cas3* translational initiation site. Additionally, 13 nucleotides located from –83 to –71 are involved in *cas3* positive regulation.

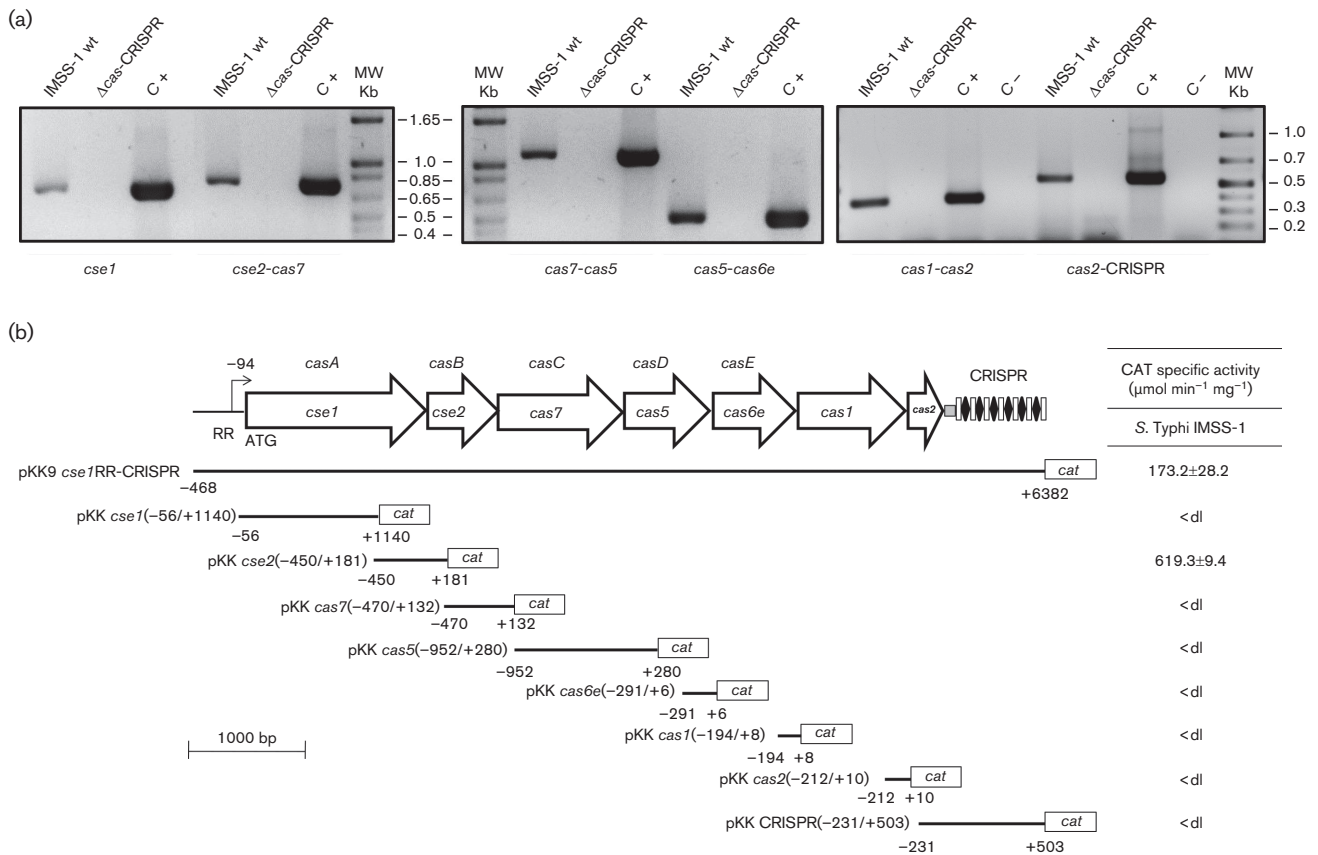
### The sense strand of the *S. Typhi* CRISPR-Cas locus encodes two transcriptional units: the *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR operon and the *scse2* RNA

Previously, we showed that there was minimal expression of the CRISPR-Cas system in rich medium and that LeuO overexpression induces transcriptional activity of the

*cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR operon (*cas*-CRISPR) in this condition [22]. Here, we evaluated by RT-PCR whether the *cse1* promoter drives the genetic expression of the complete *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR locus as a polycistronic RNA in N-MM. Total RNA, isolated from the wild-type *S. Typhi* strain grown in this medium to OD<sub>595</sub> 1.3, was used for synthesizing cDNA with a specific CRISPR oligonucleotide (CRISPR +503 *Xho*I-R) located at the 3' end of the locus (see Table S2). Using this cDNA and specific primers, the fragments *cse1*, *cse2-cas7*, *cas7-cas5*, *cas5-cas6e*, *cas1-cas2* and *cas2*-CRISPR were amplified by PCR (Fig. 2a). The results demonstrated that these genes together with the CRISPR sequences comprise an operon in N-MM.

To further support these data, a transcriptional reporter containing the *cse1* promoter, the *cse1-cse2-cas7-cas5-cas6e-cas1-cas2* genes and the CRISPR sequences, together with independent constructs of each 5' intergenic *cas* regions were obtained and transformed into *S. Typhi* IMSS-1. The results in N-MM at OD<sub>595</sub> 1.3 showed, as expected, that the construct encompassing the *cse1* promoter and the *cse1-cse2-cas7-cas5-cas6e-cas1-cas2*-CRISPR region (pKK9 *cse1*RR-CRISPR) gave CAT activity of 173, indicating that this promoter is able to drive the expression of the downstream *cas* genes together with the CRISPR sequences (Fig. 2b). Moreover, no expression with the independent *cse1* coding region and with the independent *cas7*, *cas5*, *cas6e*, *cas1*, *cas2* and CRISPR 5' intergenic regions was obtained (Fig. 2b). Interestingly, with the *cse2* reporter pKK *cse2* (–450/+181), containing 450 bp upstream and 181 bp downstream of the *cse2* ATG codon, CAT activity values of 619 were observed, suggesting the presence of a promoter within this region (Fig. 2b). Additionally, nucleotides –450 to +181 were also fused to the *lacZ* reporter gene and transcriptional activity of 1549.3±128.5  $\beta$ -galactosidase units was detected (data not shown), supporting the notion that this region contains a promoter.

To determine the size of the *cse2* transcriptional unit, several transcriptional reporters of different lengths were obtained and transformed into *S. Typhi* IMSS-1. The CAT activity results in N-MM demonstrated that the shorter constructs pKK *cse2* (–62/+181), pKK *cse2* (+134/+181) and pKK *cse2* (+142/+181) gave expression levels similar to those of the longest plasmid pKK *cse2* (–450/+181) (Fig. 3a), showing that a promoter is present in the *cse2* intragenic region, between nucleotides +142 and +181 relative to the *cse2* ATG start codon. Surprisingly, the pKK *cse2* (+142/+202) and pKK *cse2* (+142/+267) were not expressed (Fig. 3a), indicating the presence of a terminator sequence downstream of nucleotide +181, an observation that also supports the difference observed in Fig. 2 (b) between the pKK9 *cse1*RR-CRISPR and pKK *cse2* (–450/+181) reporters. Collectively, these data indicate that a promoter is present between nucleotides +142 and +181 relative to the *cse2* translational start site. This short region contains the elements necessary to induce *scse2* expression. Bioinformatic analyses using MEME-MAST [60] showed the presence of

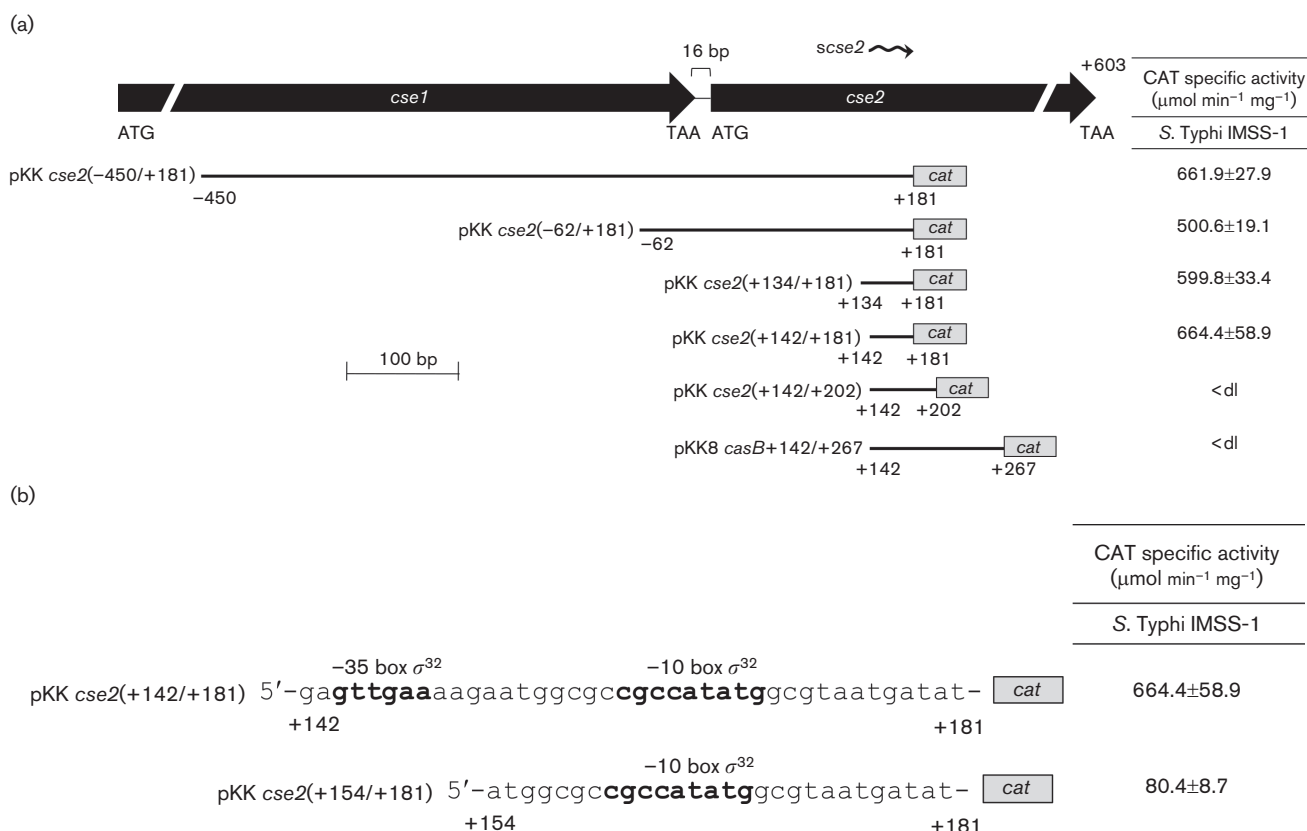


**Fig. 2.** The *cse1-cse2-cas7-cas5-cas6e-cas1-cas2-CRISPR* locus is expressed as an operon in N-MM. (a) Chromosomal RT-PCRs were performed with total RNA from *S. Typhi* IMSS-1 grown in N-MM (100 mM Tris/HCl) to OD<sub>595</sub> 1.3. PCR fragments of the *cse1* intragenic region (704 bp), *cse2-cas7* (756 bp, containing the last 454 bp of *cse2*, a *cse2-cas7* intergenic region of 16 bp and the first 286 bp of *cas7*), *cas7-cas5* (1104 bp, encompassing the last 815 bp of *cas7*, *cas7-cas5* intergenic region of 9 bp and the first 280 bp of *cas5*), *cas5-cas6e* (511 bp, harbouring the last 291 bp of *cas5* and the first 220 bp of *cas6e*), *cas1-cas2* (389 bp that contains the last 210 bp of *cas1*, *cas1-cas2* intergenic region of 2 bp and the first 177 bp of *cas2*) and *cas2-CRISPR* (516 bp, including the last 13 bp of *cas2*, the 84 bp leader sequence and 419 bp of the CRISPR array) were generated using specific primers and, as template, a cDNA synthesized with oligonucleotide CRISPR +503 *Xho*I-R located at the end of the CRISPR sequence. cDNA from *S. Typhi* lacking the complete CRISPR-Cas locus ( $\Delta$ *cas*-CRISPR) was used as negative specific amplification control for RT experiments. The *cse1*, *cse2-cas7*, *cas7-cas5*, *cas5-cas6e*, *cas1-cas2* and *cas2-CRISPR* DNA fragments were synthesized from genomic *S. Typhi* DNA (C+) as positive control of the PCR. Additionally, a cDNA synthesis reaction using RNA from *S. Typhi* IMSS-1 without reverse transcriptase was carried out as negative RT-PCR control (C-). DNA molecular weight markers (MW) are included. (b) Genetic organization of the *S. Typhi cas*-CRISPR operon. The *cse1* 5' intergenic region of 357 bp (regulatory region, RR) contains the promoter of *cas*-CRISPR unit; the bent arrow indicates the *cse1* transcriptional start site located 94 bp upstream of its ATG codon. This polycistronic RNA is composed of *cse1-cse2-cas7-cas5-cas6e-cas1-cas2* genes (the first five genes are also called *casA*, *casB*, *casC*, *casD* and *casE* [3]), a leader sequence of 84 bp (grey rectangle) and the CRISPR I array containing seven repeats (white rectangles) and six spacers (black diamonds). Below CRISPR-Cas diagram, CAT transcriptional constructs of the CRISPR-Cas locus are shown. These were named (boldface) according to the corresponding *cas* gene evaluated; the numbers below each line represent the upstream and downstream coordinates relative to the ATG codon of the respective ORF. For pKK9 *cse1*RR-CRISPR, the coordinates upstream and downstream are from the *cse1* ATG codon; in the case of the CRISPR construct, the name and coordinates upstream and downstream are according to the first nucleotide of the CRISPR leader sequence. The right columns represent the expression of the CRISPR-Cas reporters evaluated in the wild-type *S. Typhi* IMSS-1 strain. CAT-specific activities were measured in N-MM (100 mM Tris/HCl) at OD<sub>595</sub> 1.3. The values are the means±standard deviations for three independent experiments performed in duplicate; <dl (<detection limit) represents values between 0 and 10 CAT units.

-35 (GTTGAA) and -10 (CGCCATATG) consensus sequences for *E. coli*  $\sigma^{32}$  [61, 62], between nucleotides +142 and +181 (Fig. 3b), with an e-value of  $10^{-6}$ . A transcriptional reporter lacking the -35  $\sigma^{32}$  box was constructed, pKK *cse2*

(+154/+181), and evaluated in *S. Typhi* IMSS-1 grown in N-MM. Its CAT activity was eight times lower (80 CAT units) than the pKK *cse2* (+142/+181) (664 CAT units, Fig. 3b), validating the presence of *scse2* promoter in this region. Notably,





**Fig. 3.** Identification and delimitation of the *scse2* transcriptional unit. (a) The *cse2* gene is composed of 603 bp and localized downstream of *cse1* with a 5' intergenic region of 16 bp. The wavy arrow represents the *scse2* RNA encompassing *cse2* intragenic nucleotides +142 to +181. Below the diagram, CAT transcriptional constructs of different length are shown. They were named (boldface) according to the nucleotides contained in each *cse2* construct; the numbers describing each reporter represent the coordinates upstream and downstream relative to the *cse2* ATG start codon. The right columns represent the expression of the *cse2* plasmids evaluated in the wild-type *S. Typhi* IMSS-1 strain. CAT-specific activities were measured in N-MM (100 mM Tris/HCl) at OD<sub>595</sub>1.3. (b) *scse2* promoter is present in nucleotides +142 to +181. The pKK *cse2* (+142/+181) contains the -10 and -35  $\sigma^{32}$  sequences (boldface), whereas the pKK *cse2* (+154/+181) lacks the -35 box. Both plasmids were evaluated in *S. Typhi* IMSS-1 wild-type strain grown in N-MM (100 mM Tris/HCl) to OD<sub>595</sub> 1.3, and the results of the CAT assays are shown in the right columns. For all the CAT-specific activities, the values are the means±standard deviations of three independent experiments performed in duplicate; <dl (< detection limit) represents values between 0 and 10 CAT units.

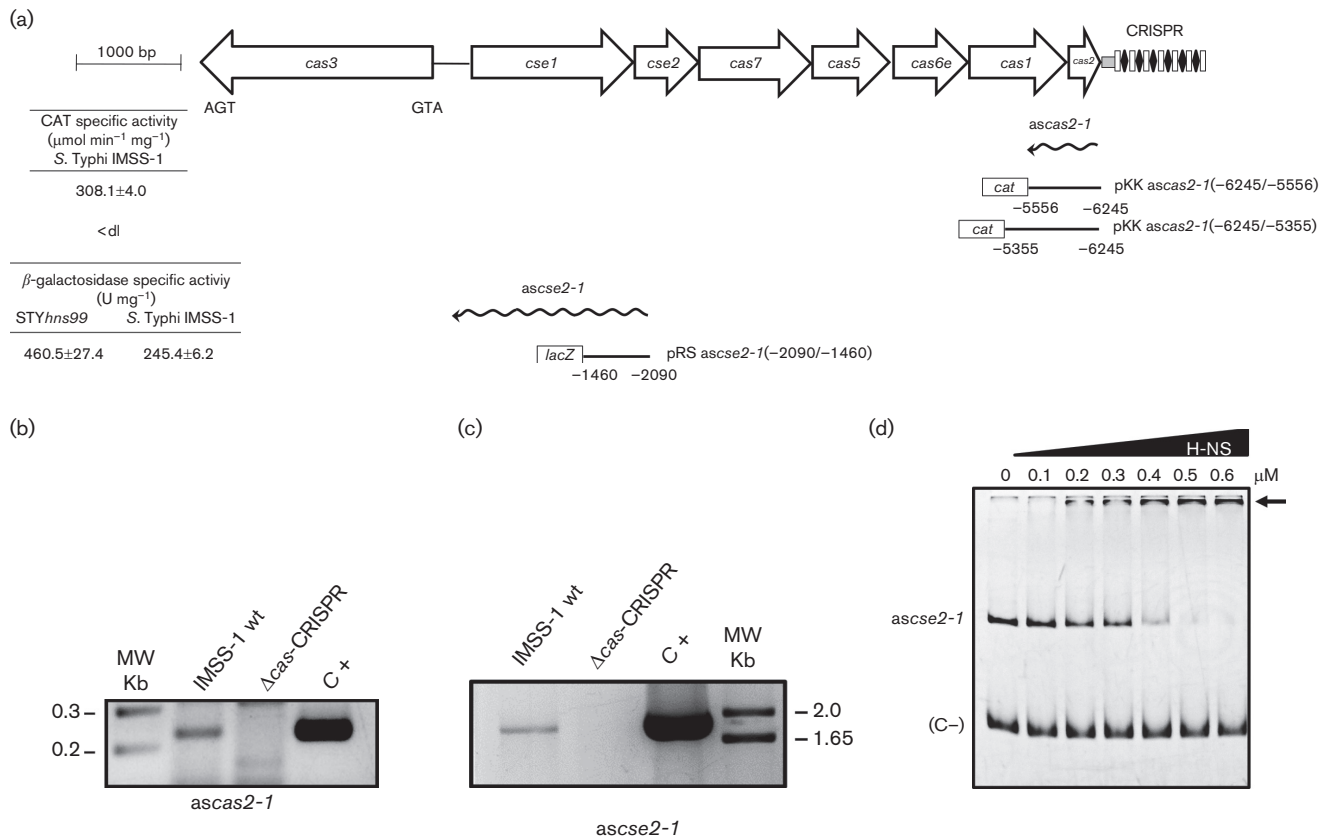
such promoter would be coding for a very small 20 nt RNA, here named *scse2*, whose role in expression warrants future research.

Thus, in the sense strand of the *S. Typhi* CRISPR-Cas locus, the *cse1* promoter drives the genetic expression of the *cse1-cse2-cas7-cas5-cas6-cas1-cas2*-CRISPR operon, whereas another promoter controls the activity of the *scse2* RNA.

### The antisense strand of the *S. Typhi* CRISPR-Cas locus encodes two transcriptional units: *ascse2-1* and *ascas2-1*

Since an RNA in the intragenic region of *cse2* was found and *cas3* is divergent to the *cse1-cse2-cas7-cas5-cas6-cas2-cas1* genes, the question was whether there were other transcriptional units in the antisense strand of the CRISPR-Cas locus, and hence, two independent promoters were found. The first one encompasses part of the *cas2* and *cas1*

antisense strand (*ascas2-1*) since a transcriptional reporter containing this region (nucleotides -6245 to -5556 relative to the *cas3* ATG start codon), namely pKK *ascas2-1* (-6245/-5556), had activity values of 308 CAT units in *S. Typhi* IMSS-1 grown in N-MM (OD<sub>595</sub> 1.3) (Fig. 4a). To determine the length of this transcript, the construct pKK *ascas2-1* (-6245/-5556) was evaluated in the same growth conditions, but no CAT activity was detected (Fig. 4a). This suggests that the *ascas2-1* transcriptional unit is located between -6245 and -5556 (689 bp). To validate this result, a chromosomal RT-PCR was performed. The cDNA was synthesized using *S. Typhi* total RNA and a primer located at the -5556 nucleotide (*ascas1* -5556 *KpnI*-R, see Table S2). For PCR amplification, the primers *ascas1* -5556 *KpnI*-R and *ascas1* -5775 F (Table S2), together with the cDNA as template, rendered a product of 220 bp, indicating the presence of the *ascas2-1* RNA (Fig. 4b).



**Fig. 4.** The *ascas2-1* and *ascse2-1* transcriptional units are present in the antisense strand of the *S. Typhi* CRISPR-Cas loci. (a) Schematic representation of the *S. Typhi* CRISPR-Cas system: wavy arrows represent the *ascas2-1* and *ascse2-1* transcriptional units. Below the diagram, the respective CAT transcriptional constructs for the *ascas2-1* and *lacZ* transcriptional reporters of *ascse2-1* are shown. The plasmids were named (boldface) according to the antisense transcriptional unit evaluated; the numbers below correspond to the nucleotides contained in each construct relative to the ATG start codon of *cas3*. The left-hand columns represent the expression of the *ascse2-1* plasmid evaluated in the wild-type *S. Typhi* IMSS-1 as well as in the *hns* isogenic mutant (STY*hns99*). In the case of *ascas2-1*, only the expression in the wild-type is shown. CAT-specific and  $\beta$ -galactosidase-specific activities were measured in N-MM (100 mM Tris/HCl) at OD<sub>595</sub> 1.3. The values are the means $\pm$ standard deviations of three independent experiments performed in duplicate. <dl (< detection limit) represents values between 0 and 10 CAT units. (b) RT-PCR of *ascas2-1* transcriptional unit. Total RNA was purified from *S. Typhi* IMSS-1 grown in N-MM (100 mM Tris/HCl) to OD<sub>595</sub> 1.3 and the *ascas1* –5556 *Kpn*I-R oligonucleotide was used to generate cDNA. Using this template, a DNA fragment of 220 bp was PCR-amplified with specific primers located at –5556 and –5775 nucleotides upstream of *cas3* ATG codon. (c) RT-PCR of *ascse2-1* RNA. Chromosomal RT-PCR was performed with total RNA from *S. Typhi* IMSS-1 grown in N-MM (100 mM Tris/HCl) to OD<sub>595</sub> 1.3. The cDNA was synthesized with the *ascse1* –34 *Bam*HI-R primer. A PCR fragment of 1925 bp, encompassing nucleotides –34 to –1958 relative to *cas3* translational start site, was amplified using the cDNA as template. For both RT-PCR experiments, the corresponding cDNAs from *S. Typhi* lacking the complete CRISPR-Cas locus ( $\Delta$ *cas*-CRISPR) was used as negative specific amplification control in PCR experiments. As positive control, the *ascas2-1* and *ascse2-1* fragments of 220 bp and 1925 bp, respectively, were synthesized from genomic *S. Typhi* DNA (C+). Additionally, the corresponding cDNA synthesis reactions without reverse transcriptase were carried out as negative RT-PCR control and no DNA amplification was observed (data not shown). DNA molecular weight markers (MW) are included. (d) H-NS interacts with *ascse2-1*. EMSA were performed with increasing concentrations of purified H-NS-Myc-His6 protein and a PCR-generated DNA fragment of *ascse2-1* (630 bp, nucleotides –2090 to –1460 relative to *cas3* ATG codon). A 16S rRNA fragment of 242 bp was used as a negative control (C–). Arrow indicates the DNA-protein complex.

The second promoter named *ascse2-1* was identified between the *cse2* and *cse1* antisense strand, at –2090 to –1460 since a plasmid harbouring this region, pRS *ascse2-1* (–2090/–1460), rendered expression levels of 245  $\beta$ -galactosidase units when it was evaluated in *S. Typhi* grown in N-MM (Fig. 4a). To define the *ascse2-1* transcriptional unit, RT-PCR experiments were performed. Total RNA purified from *S. Typhi* (N-MM, OD<sub>595</sub> 1.3) and the *ascse1* –34 *Bam*HI-R oligonucleotide was

used to generate the cDNA. For PCR amplification, this cDNA together with primers *ascse1* –34 *Bam*HI-R and *ascse2* –1958 *Kpn*I-F rendered a fragment of 1925 bp corresponding to the antisense *ascse2-1* (Fig. 4c).

To further characterize *ascas2-1* and *ascse2-1*, it was determined whether H-NS represses its expression since this nucleoid protein negatively regulates other *S. Typhi* CRISPR-Cas

transcriptional units such as *cas3* and the *cas*-CRISPR operon. Thus, the pKK *ascas21* (−6245/−5556) and pRS *ascse2-1* (−2090/−1460) were evaluated in STYhms99 strain grown in N-MM. The transcriptional profiles showed that H-NS is not implicated in the control of *ascas2-1* since the CAT units obtained in STYhms99 were similar to those observed in the IMSS-1 wild-type strain (Fig. S1). However, the activity levels of the pRS *ascse2-1* (−2090/−1460) were higher in the H-NS-deficient strain (460  $\beta$ -galactosidase units) than in the wild-type (245  $\beta$ -galactosidase units) (Fig. 4a), showing that H-NS silences the expression of the antisense *ascse2-1* RNA. EMSA using purified H-NS protein and an *ascse2-1* fragment (nucleotides −2090 to −1460, 630 bp) demonstrated that this global regulator directly represses *ascse2-1* since it interacted with the DNA region evaluated (Fig. 4d).

Thus, two transcriptional units in the antisense strand of the CRISPR-Cas loci were identified: *ascse2-1* and *ascas2-1*. Moreover, *ascse2-1* is negatively regulated by H-NS.

### pH regulates the genetic expression of the five transcriptional units present in the *S. Typhi* CRISPR-Cas system

The expression of the CRISPR-Cas system has been previously shown to depend on environmental conditions [18, 19]. In addition, as described herein, there are transcriptional units on the sense strand (*cas*-CRISPR operon, *scse2*) and on the antisense strand (*ascse2-1*, *ascas2-1*) of the *S. Typhi* CRISPR-Cas locus that, together with *cas3*, are expressed in N-MM. Hence, in order to identify the specific signal in N-MM responsible for CRISPR-Cas induction, the influence of the carbon source (glycerol) present in this medium was tested by replacing it with glucose. Upon evaluation of the *cse1* (*cas*-CRISPR promoter), *scse2* and *cas3* promoters, their expression was found to be similar in both carbon sources (data not shown) indicating that glycerol is not the signal that promotes CRISPR-Cas expression. Several reports show the induction of *Salmonella* pathogenicity genes in acid pH [63–65], so the influence of pH on CRISPR-Cas transcription was tested. The expression of *cse1*, *scse2*, *ascse2-1*, *ascas2-1* and *cas3* promoters was relatively low at acid pH 6.0–6.5. However, their activity was higher in N-MM pH 7.5 and 7.7 (Fig. 5a). To confirm that, in fact, pH determines this increase, the five transcriptional units were evaluated in rich medium A (MA) at pH 6.0, 6.5, 7.0, 7.5 and 7.7. Even though this medium is significantly different from N-MM, the expression of *cse1*, *scse2*, *ascse2-1*, *ascas2-1* and *cas3* also increased at pH 7.5–7.7 (Fig. S2). In the case of *cse1*, *ascse2-1* and *cas3*, their transcriptional levels were higher in N-MM than those obtained in MA (Figs 5a and S2) illustrating a significant effect of N-MM and pH. On the other hand, the expression of the *scse2* and *ascas2-1* promoters in N-MM and MA was similar, although the effect of pH was maintained (Figs 5a and S2).

Additionally, primer extension experiments to determine the transcriptional start site of the *ascse2-1* unit were performed, by using total RNA from *S. Typhi* IMSS-1 harbouring the pRS *ascse2-1* plasmid grown in N-MM at pH

6.5 and 7.5. For generating the cDNA, we used the *ascse1* −1766 *Bam*HI-R oligonucleotide, located at nucleotide −1766 relative to *cas3* translational initiation site. The results showed an *ascse2-1* transcriptional start site at 2062 bp upstream of the *cas3* ATG codon, whose transcription increased at pH 7.5 compared to pH 6.5 (Fig. 5b).

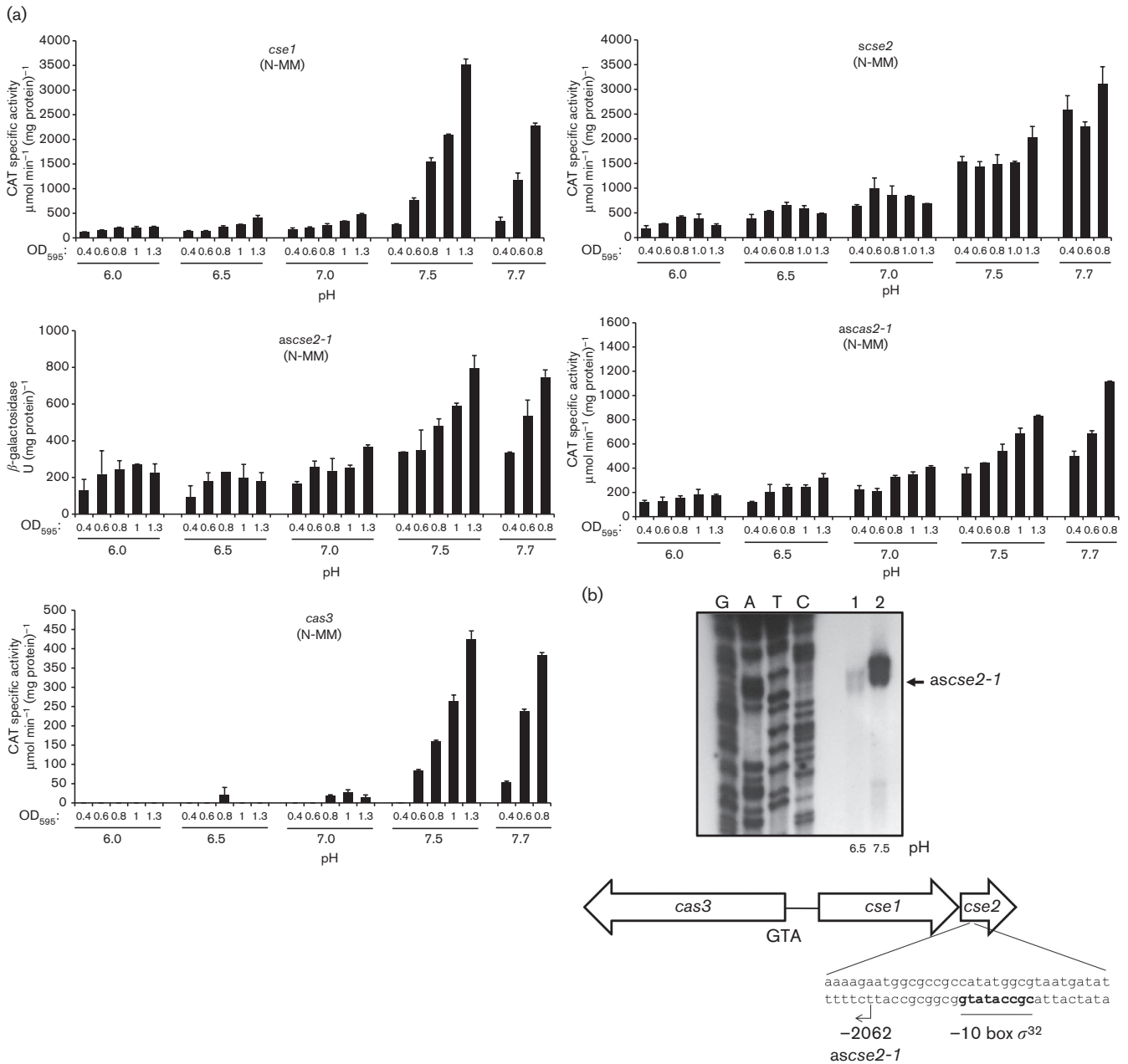
Since all the transcriptional units identified were up-regulated at pH 7.5, we evaluated their role under this condition. Growth rate experiments comparing a strain devoid of the entire CRISPR-Cas locus with the wild-type strain were performed in N-MM at pH 7.5 and 7.7. The results showed the same growth rate in exponential and stationary phase in both strains, thus indicating that the CRISPR-Cas system is not necessary for *S. Typhi* replication in N-MM at pH 7.5–7.7 (data not shown).

## DISCUSSION

The CRISPR-Cas system has mainly been studied for its function in bacterial immunity, and recently, its role in virulence [14], biofilm formation [17] and gene regulation [15] has been documented. However, there are relatively few studies on its transcriptional regulation and the signals that promote its expression. We previously reported that H-NS and LRP silence the *cse1* promoter whereas the LysR-type regulator LeuO induces its expression in rich medium [22, 26]. In this study, we performed a genetic characterization of five transcriptional units present in the *S. Typhi* CRISPR-Cas locus: the *cas*-CRISPR operon, *scse2*, *ascse2-1*, *ascas2-1* and *cas3*. These genetic elements were expressed in N-MM and their activity increased from pH 7.5 to 7.7.

Our data support the view that the CRISPR-Cas system in *S. Typhi* is coded by a complex operon, since it harbours specific independent transcriptional units not previously identified. Furthermore, regulation studies of the CRISPR-Cas Type I-E system present in *E. coli* K12 support our data, since this locus presents diverse transcriptional units: one containing the *cas3* gene; a second unit comprises the *cse1-cse2-cas7-cas5-cas6-cas1-cas2* genes; the CRISPR loci are transcribed from a promoter present in the leader sequence; and an antisense promoter (anti-*Pcas*) was also identified in the 414 bp intergenic region between *cas3* and *cse1* [33, 66, 67]. In *Sulfolobus islandicus*, it has also been demonstrated that the Type I-A CRISPR-Cas system is transcribed as two units: the *csa1-cas1* and the *acas* (*cas1-cas2-cas4*) operons [68]. Thus, all these works support the notion that multiple transcriptional units on sense and antisense strands may be characteristic of the CRISPR-Cas Type I systems of bacteria and archaea.

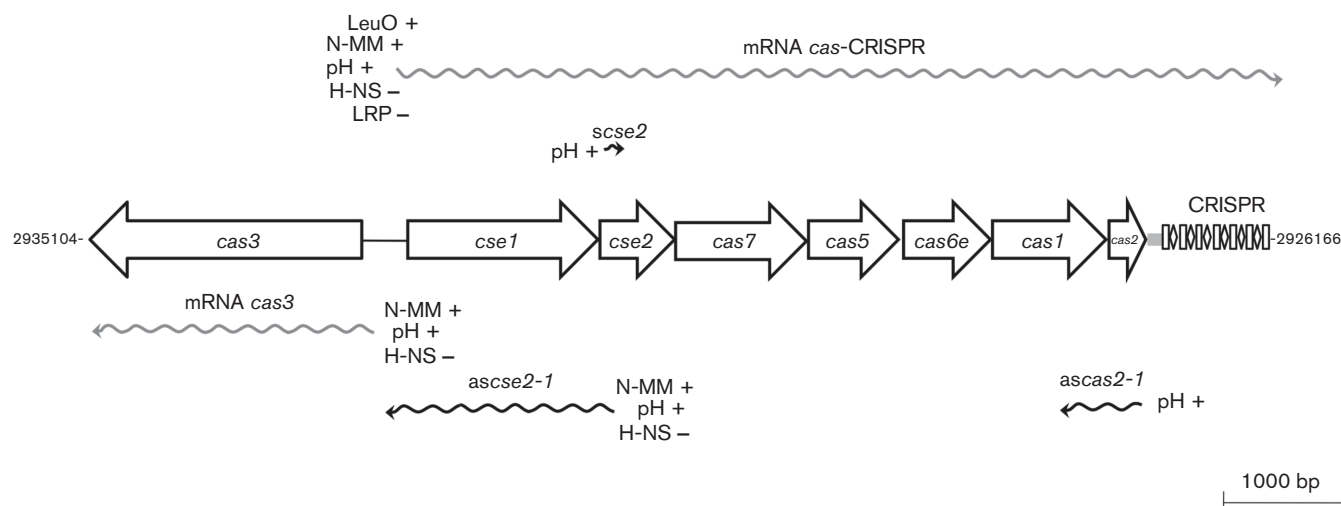
The transcriptional characterization of the *S. Typhi* *cas3* regulatory region is also presented here. This nuclease-helicase protein has an essential role in bacterial immunity mediated by functional CRISPR-Cas Type I systems, since its HD nuclease domain degrades the invader genetic element [7, 8] and recently it has been implicated in the acquisition of new spacer sequences [69, 70]. We have thus demonstrated that, in



**Fig. 5.** *cse1*, *scse2*, *ascse2-1*, *ascas2-1* and *cas3* expression is modulated by pH. (a) Transcriptional profiles of *S. Typhi* IMSS-I harbouring pKK *cse1* (–468/+98), pKK *cse2* (–450/+181), pRS *ascse2-1* (–2090/–1460), pKK *ascas2-1* (–6245/–5556) or pKK *cas3* (–455/+112), grown in N-MM at different pH values (6.0, 6.5, 7.0, 7.5 or 7.7). Tris/HCl (200 mM) was used to avoid a decrease in pH. Transcriptional activities were measured at different OD<sub>595</sub> values. The values are the means±standard deviations of two independent experiments performed in duplicate. (b) Identification of the *ascse2-1* transcriptional start site. Primer extension experiments were performed with total RNA from *S. Typhi* containing the pRS *ascse2-1* (–2090/–1460) plasmid grown at OD<sub>595</sub> 1.3 in N-MM (200 mM Tris/HCl) at pH 6.5 (line 1) or pH 7.5 (line 2). The cDNAs were synthesized with the *ascse1* –1766 *Bam*HI-R oligonucleotide (see Table S2) and then resolved in 8% polyacrylamide gels alongside sequencing ladders. The location of the *ascse2-1* transcriptional start site (bent arrow), at the –2062 nucleotide relative to *cas3* ATG start codon, is shown in the sequence at the bottom of the figure. Using MEME-MAST [60], a –10 consensus sequence for *E. coli* σ<sup>32</sup> CGCCATATG (boldface) was found [61] with an e-value of 10<sup>–6</sup>.

*S. Typhi*, *cas3* is negatively regulated by the global regulatory protein H-NS, which also silences *cse1* [26] and the antisense *ascse2-1*. Interestingly, in the CRISPR-Cas system from *E. coli* K12, H-NS binds to *Pcas*, anti-*Pcas* and *Pcrispr1* promoters,

silencing their expression [33]. Thus, in *E. coli* and *S. Typhi*, H-NS regulates more than one CRISPR-Cas transcriptional unit, suggesting that, in these bacteria, the CRISPR-Cas is silenced in multiple steps.



**Fig. 6.** Transcriptional organization and regulation of the *S. Typhi* CRISPR-Cas system. Wavy grey arrows represent mRNAs of *cas-CRISPR* and *cas3*, whereas the RNA *scse2*, the antisense *ascse2-1* and *ascas2-1* are shown as wavy black arrows. We present all the elements implicated in the positive (+) and negative (–) regulation of each transcriptional unit from the *S. Typhi* CRISPR-Cas locus that have been identified in this work, as well as by Hernández-Lucas *et al.* [26] and Medina-Aparicio *et al.* [22] including the LysR-type regulator LeuO, LRP (leucine-response protein), H-NS (heat-stable nucleoid-structuring protein), N-MM (N-minimal medium) and pH.

Although the divergently transcribed *cse1* and *cas3* genes share a 5' intergenic region, we found that LRP and LeuO only control *cse1* expression. These data showed that *cas3* is transcriptionally regulated by other factors or signals that maintain its expression at low levels in N-MM.

An evaluation to determine transcriptional factors involved in *cas3* regulation, as well as those in *cas-CRISPR* operon, *scse2* and *ascas2-1*, led to the finding that the CpxR, SsrB, IHF, PhoP, PmrA, FNR, CRP, LRP and OmpR regulators are not involved in the transcriptional control of these transcriptional units (Fig. S1). The role of H-NS in the regulation of *scse2* and *ascas2-1* was also evaluated but no effect was observed (Fig. S1). The proteins evaluated include transcriptional factors implicated in virulence [52–58], or proteins involved in regulation of other CRISPR-Cas systems such as cAMP receptor protein (CRP), whose role in positive regulation of *Thermus thermophilus* and *Pectobacterium atrosepticum* CRISPR-Cas systems [50, 71] or in the repression of *E. coli* CRISPR-Cas locus [51] has been established. Therefore, the lack of CRP participation in the control of *S. Typhi* CRISPR-Cas transcriptional units illustrates the differences in the transcriptional regulation of this system [22, 71].

The expression of the transcriptional units defined in this work increases at pH 7.5, a pH value characteristic of the distal ileum [35], a compartment that is invaded and colonized by *S. Typhi*. Thus, it is relevant to determine whether these genetic elements have a role in the response to the environment present in the gastrointestinal tract or directly participate in the invasion of this compartment. Although it

is possible that the CRISPR-Cas systems would offer a mechanism for adaptation to environment changes, experiments of phage resistant or plasmid acquisition in conditions where the *S. Typhi* CRISPR-Cas locus is expressed are necessary to determine whether it participates in bacterial immunity.

Finally, the *S. Typhi* CRISPR-Cas system could be involved in other aspects of the pathogenesis and is a matter of future studies. For instance, the Type II CRISPR-Cas system of the human pathogen *Francisella novicida* downregulates the expression of bacterial lipoprotein, promoting the evasion of the host immune system [15]. In *Campylobacter jejuni* and *Neisseria meningitidis*, the Cas9 protein, a component of the CRISPR-Cas system, is necessary for intracellular survival of these bacteria in human epithelial cells from intestine and lung, respectively [14, 15].

In summary, we have characterized five transcriptional units present in the *S. Typhi* CRISPR-Cas system. In Fig. 6, all the known elements implicated in the transcriptional control of *S. Typhi* CRISPR-Cas locus are shown, providing an overview of its regulation.

#### Funding information

This work was supported by grants from Dirección General de Asuntos de Personal Académico, DGAPA/UNAM (IN203312 and IN203215 to I. H.-L.; IN201513 to E. C.) and Consejo Nacional de Ciencia y Tecnología, CONACYT, México (89337 and 127298 to I. H.-L.; 179946 to E. C.)

#### Acknowledgements

We would like to thank N. Becerra-Lobato, M. L. Zavala-García, V. E. Osio-Becerro, C. Hernández, M. Fernández-Mora, F. J. Santana, P. Gaytan, E. Bustos, S. Becerra and J. Yañez (IBT-UNAM) for technical help and J. M. Villarreal, A. Medrano-López, G. Gosset (IBT-UNAM); J.

Miranda (IIB-UNAM), and M. Dunn (CCG-UNAM) for stimulating discussions and critical reading.

#### Conflicts of interest

The authors declare that there are no conflicts of interest.

#### References

- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007;315:1709–1712.
- Sorek R, Lawrence CM, Wiedenheft B. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem* 2013;82:237–266.
- Brouns SJ, Matthijs MJ, Lundgren M, Westra ER, Slijkhuys RJ et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 2008;321:960–964.
- Hale CR, Zhao P, Olson S, Duff MO, Graveley BR et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 2009; 139:945–956.
- Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 2011;9:467–477.
- Beloglazova N, Petit P, Flick R, Brown G, Savchenko A et al. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J* 2011;30:4616–4627.
- Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P et al. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* 2013;30: 1335–1342.
- Westra ER, van Erp PB, Künne T, Wong SP, Staals RH et al. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by cascade and Cas3. *Mol Cell* 2012;46:595–605.
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011;471:602–607.
- Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 2014;346: 1258096.
- Elmore JR, Sheppard NF, Ramia N, Deighan T, Li H et al. Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. *Genes Dev* 2016;30:447–459.
- Samai P, Pyenson N, Jiang W, Goldberg GW, Hatoum-Aslan A et al. Co-transcriptional DNA and RNA cleavage during Type III CRISPR-Cas immunity. *Cell* 2015;161:1164–1174.
- Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 2007;35:W52–W57.
- Louwen R, Horst-Kreft D, De Boer AG, Van der Graaf L, De Knecht G et al. A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain-Barré syndrome. *Eur J Clin Microbiol Infect Dis* 2013;32:207–226.
- Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL, Weiss DSA. CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 2013;497:254–257.
- Viswanathan P, Murphy K, Julien B, Garza AG, Kroos L. Regulation of dev, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J Bacteriol* 2007;189:3738–3750.
- Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH et al. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J Bacteriol* 2009;191:210–219.
- Perez-Rodriguez R, Haitjema C, Huang Q, Nam KH, Bernardis S et al. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol Microbiol* 2011;79:584–599.
- Ratner HK, Sampson TR, Weiss DS. I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. *Curr Opin Infect Dis* 2015;28:267–274.
- Serbanescu MA, Cordova M, Krastel K, Flick R, Beloglazova N et al. Role of the *Streptococcus mutans* CRISPR-Cas systems in immunity and cell physiology. *J Bacteriol* 2015;197:749–761.
- Williams E, Lowe TM, Savas J, Diruggiero J. Microarray analysis of the hyperthermophilic archaeon *Pyrococcus furiosus* exposed to gamma irradiation. *Extremophiles* 2007;11:19–29.
- Medina-Aparicio L, Rebollar-Flores JE, Gallego-Hernández AL, Vázquez A, Olvera L et al. The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and Leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J Bacteriol* 2011; 193:2396–2407.
- Shariat N, Timme RE, Pettengill JB, Barrangou R, Dudley EG. Characterization and evolution of *Salmonella* CRISPR-Cas systems. *Microbiology* 2015;161:374–386.
- Faucher SP, Curtiss R, Daigle F. Selective capture of *Salmonella enterica* serovar Typhi genes expressed in macrophages that are absent from the *Salmonella enterica* serovar Typhimurium genome. *Infect Immun* 2005;73:5217–5221.
- Deiwick J, Nikolaus T, Erdogan S, Hensel M. Environmental regulation of *Salmonella* pathogenicity island 2 gene expression. *Mol Microbiol* 1999;31:1759–1773.
- Hernández-Lucas I, Gallego-Hernández AL, Encarnación S, Fernández-Mora M, Martínez-Batallar AG et al. The LysR-type transcriptional regulator LeuO controls expression of several genes in *Salmonella enterica* serovar Typhi. *J Bacteriol* 2008;190: 1658–1670.
- Baek CH, Wang S, Roland KL, Curtiss R 3rd. Leucine-responsive regulatory protein (Lrp) acts as a virulence repressor in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 2009;191:1278–1292.
- Espinosa E, Casadesús J. Regulation of *Salmonella enterica* pathogenicity island 1 (SPI-1) by the LysR-type regulator LeuO. *Mol Microbiol* 2014;91:1057–1069.
- Guadarrama C, Villaseñor T, Calva E. The subtleties and contrasts of the LeuO regulator in *Salmonella* Typhi: Implications in the immune response. *Front Immunol* 2014;5:581.
- Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M et al. H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog* 2006;2:e81.
- Marshall DG, Sheehan BJ, Dorman CJ. A role for the leucine-responsive regulatory protein and integration host factor in the regulation of the *Salmonella* plasmid virulence (spv) locus in *Salmonella typhimurium*. *Mol Microbiol* 1999;34:134–145.
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H et al. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 2006;313:236–238.
- Pul U, Wurm R, Arslan Z, Geissen R, Hofmann N et al. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol* 2010;7:1495–1512.
- Westra ER, Pul U, Heidrich N, Jore MM, Lundgren M et al. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol Microbiol* 2010;77:1380–1393.
- Evans DF, Pye G, Bramley R, Clark AG, Dyson TJ et al. Measurement of gastrointestinal pH profiles in normal ambulant human subjects. *Gut* 1988;29:1035–1041.
- Puente JL, Flores V, Fernández M, Fuchs Y, Calva E. Isolation of an *ompC*-like outer membrane protein gene from *Salmonella typhi*. *Gene* 1987;61:75–83.
- Kawaji H, Mizuno T, Mizushima S. Influence of molecular size and osmolarity of sugars and dextrans on the synthesis of outer membrane proteins O-8 and O-9 of *Escherichia coli* K-12. *J Bacteriol* 1979;140:843–847.
- Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989.

39. Gallego-Hernández AL, Hernández-Lucas I, de La Cruz MA, Olvera L, Morett E et al. Transcriptional regulation of the *assT-dsbl-dsbl* gene cluster in *Salmonella enterica* serovar Typhi IMSS-1 depends on LeuO, H-NS, and specific growth conditions. *J Bacteriol* 2012;194:2254–2264.
40. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 2000;97:6640–6645.
41. Simons RW, Houman F, Kleckner N. Improved single and multi-copy lac-based cloning vectors for protein and operon fusions. *Gene* 1987;53:85–96.
42. Martínez-Laguna Y, Calva E, Puente JL. Autoactivation and environmental regulation of *bfpT* expression, the gene coding for the transcriptional activator of *bfpA* in enteropathogenic *Escherichia coli*. *Mol Microbiol* 1999;33:153–166.
43. Oropeza R, Sampieri CL, Puente JL, Calva E. Negative and positive regulation of the non-osmoregulated *ompS1* porin gene in *Salmonella* Typhi: a novel regulatory mechanism that involves OmpR. *Mol Microbiol* 1999;32:243–252.
44. Villarreal JM, Becerra-Lobato N, Rebollar-Flores JE, Medina-Aparicio L, Carbajal-Gómez E et al. The *Salmonella enterica* serovar Typhi *ltrR-ompR-ompC-ompF* genes are involved in resistance to the bile salt sodium deoxycholate and in bacterial transformation. *Mol Microbiol* 2014;92:1005–1024.
45. De La Cruz MA, Fernández-Mora M, Guadarrama C, Flores-Valdez MA, Bustamante VH et al. LeuO antagonizes H-NS and StpA-dependent repression in *Salmonella enterica ompS1*. *Mol Microbiol* 2007;66:727–743.
46. Gong B, Shin M, Sun J, Jung CH, Bolt EL et al. Molecular insights into DNA interference by CRISPR-associated nuclease-helicase Cas3. *Proc Natl Acad Sci USA* 2014;111:16359–16364.
47. Huo Y, Nam KH, Ding F, Lee H, Wu L et al. Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nat Struct Mol Biol* 2014;21:771–777.
48. Mulepati S, Bailey S. Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J Biol Chem* 2011;286:31896–31903.
49. Flores-Valdez MA, Puente JL, Calva E. Negative osmoregulation of the *Salmonella* *ompS1* porin gene independently of OmpR in an *hns* background. *J Bacteriol* 2003;185:6497–6506.
50. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S et al. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J Mol Biol* 2010;395:270–281.
51. Yang CD, Chen YH, Huang HY, Huang HD, Tseng CP. CRP represses the CRISPR/Cas system in *Escherichia coli*: evidence that endogenous CRISPR spacers impede phage P1 replication. *Mol Microbiol* 2014;92:1072–1091.
52. Bijlsma JJ, Groisman EA. The PhoP/PhoQ system controls the intramacrophage type three secretion system of *Salmonella enterica*. *Mol Microbiol* 2005;57:85–96.
53. Fink RC, Evans MR, Porwollik S, Vazquez-Torres A, Jones-Carson J et al. FNR is a global regulator of virulence and anaerobic metabolism in *Salmonella enterica* serovar Typhimurium (ATCC 14028s). *J Bacteriol* 2007;189:2262–2273.
54. Gunn JS, Ryan SS, van Velkinburgh JC, Ernst RK, Miller SI. Genetic and functional analysis of a PmrA-PmrB-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of *Salmonella enterica* serovar Typhimurium. *Infect Immun* 2000;68:6139–6146.
55. Humphreys S, Rowley G, Stevenson A, Anjum MF, Woodward MJ et al. Role of the two-component regulator CpxAR in the virulence of *Salmonella enterica* serotype Typhimurium. *Infect Immun* 2004;72:4654–4661.
56. Lee AK, Detweiler CS, Falkow S. OmpR regulates the two-component system SsrA-SsrB in *Salmonella* pathogenicity island 2. *J Bacteriol* 2000;182:771–781.
57. Mangan MW, Lucchini S, Danino V, Cróinín TO, Hinton JC et al. The integration host factor (IHF) integrates stationary-phase and virulence gene expression in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* 2006;59:1831–1847.
58. Walthers D, Carroll RK, Navarre WW, Libby SJ, Fang FC et al. The response regulator SsrB activates expression of diverse *Salmonella* pathogenicity island 2 promoters and counters silencing by the nucleoid-associated protein H-NS. *Mol Microbiol* 2007;65:477–493.
59. Ross W, Gosink KK, Salomon J, Igarashi K, Zou C et al. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* 1993;262:1407–1413.
60. Bailey TL, Boden M, Buske FA, Frith M, Grant CE et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 2009;37:W202–W208.
61. Cowing DW, Bardwell JC, Craig EA, Woolford C, Hendrix RW et al. Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc Natl Acad Sci USA* 1985;82:2679–2683.
62. Wade JT, Castro Roa D, Grainger DC, Hurd D, Busby SJ et al. Extensive functional overlap between sigma factors in *Escherichia coli*. *Nat Struct Mol Biol* 2006;13:806–814.
63. Beuzón CR, Banks G, Deiwick J, Hensel M, Holden DW. pH-dependent secretion of SseB, a product of the SPI-2 type III secretion system of *Salmonella* Typhimurium. *Mol Microbiol* 1999;33:806–816.
64. Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK et al. An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. *Cell Host Microbe* 2013;14:683–695.
65. Löber S, Jäckel D, Kaiser N, Hensel M. Regulation of *Salmonella* pathogenicity island 2 genes by independent environmental signals. *Int J Med Microbiol* 2006;296:435–447.
66. Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M et al. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol* 2010;77:1367–1379.
67. Shimada T, Yamamoto K, Ishihama A. Involvement of leucine-response transcription factor LeuO in regulation of the genes for sulfa-drug efflux. *J Bacteriol* 2009;14:4562–4571.
68. Liu T, Li Y, Wang X, Ye Q, Li H et al. Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR *de novo* spacer acquisition. *Nucleic Acids Res* 2015;43:1044–1055.
69. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K et al. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun* 2012;3:945.
70. Richter C, Gristwood T, Clulow JS, Fineran PC. *In vivo* protein interactions and complex formation in the *Pectobacterium atropense* subtype I-F CRISPR/Cas system. *PLoS One* 2012;7:e49549.
71. Patterson AG, Chang JT, Taylor C, Fineran PC. Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res* 2015;43:6038–6048.

Edited by: R. Lan and W. Achouak

## MINIREVIEW

# The CRISPR-Cas system in *Enterobacteriaceae*

Liliana Medina-Aparicio<sup>1</sup>, Sonia Dávila<sup>2</sup>, Javier E. Rebollar-Flores<sup>1</sup>,  
Edmundo Calva<sup>1</sup> and Ismael Hernández-Lucas<sup>1,\*</sup>

<sup>1</sup>Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Cuernavaca, Morelos 62210, México and <sup>2</sup>Centro de Investigación en Dinámica Celular, Universidad Autónoma del Estado de Morelos, Av. Universidad No. 1001, Cuernavaca, Morelos 62209, México

\*Corresponding author: Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Cuernavaca, Morelos 62210, México. Tel: +52 777 329-1627; Fax: +52 777 313-8673; E-mail: [ismaelh@ibt.unam.mx](mailto:ismaelh@ibt.unam.mx)

**One sentence summary:** The CRISPR-Cas system in *Enterobacteriaceae*: biological functions, transcriptional regulation, distribution, utility as a molecular marker and as a tool for specific genome editing.

**Editor:** Patrik Bavoil

## ABSTRACT

In nature, microorganisms are constantly exposed to multiple viral infections and thus have developed many strategies to survive phage attack and invasion by foreign DNA. One of such strategies is the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated proteins (Cas) bacterial immunological system. This defense mechanism is widespread in prokaryotes including several families such as *Enterobacteriaceae*. Much knowledge about the CRISPR-Cas system has been generated, including its biological functions, transcriptional regulation, distribution, utility as a molecular marker and as a tool for specific genome editing. This review focuses on these aspects and describes the state of the art of the CRISPR-Cas system in the *Enterobacteriaceae* bacterial family.

**Keywords:** *Enterobacteriaceae*; CRISPR-Cas function; CRISPR-typing; CRISPR-Cas regulation; CRISPR-Cas diversity

## INTRODUCTION

The *Enterobacteriaceae* include human pathogens that are the etiological cause of plague, typhoid fever and shigellosis, illnesses that are still a worldwide health problem (Niyogi 2005; Buckle, Walker and Black 2012; WHO 2016). Other members of this family include plant pathogens that affect crops such as potato, apple and pear, causing important economic losses (Oh and Beer 2005; Czajkowski et al. 2011; De Boer, Li and Ward 2012). Beneficial bacteria of the human intestinal microbiota also include species of this taxonomical family, which are the first microorganisms to colonize the gut of human newborns (Arrieta et al. 2014). It is clear that these ubiquitous bacteria inhabit multiple environments and are exposed to a wide array of stress conditions during their life cycle.

For instance, phage attack is one of the most relevant stress conditions for bacteria, since phages kill between 4% and 50% of the bacteria generated daily in the sea (Breitbart and Rohwer 2005). Mechanisms that have evolved to avoid phage infection include the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated proteins (Cas) system that provides specific heritable immunity against these invaders (Barrangou et al. 2007; Sorek, Lawrence and Wiedenheft 2013). It consists of a set of Cas proteins with functional domains of nucleases, helicases and polynucleotide-binding proteins (Jansen et al. 2002a; Haft et al. 2005; Makarova et al. 2006). Contiguous to the *cas* genes there is an A/T-rich non-coding leader sequence, and downstream of this region lies the CRISPR array, composed of ~30-bp direct repeats separated by similarly sized unique sequences called spacers (Jansen et al. 2002b). The presence of

Received: 4 October 2017; Accepted: 8 January 2018

© FEMS 2018. All rights reserved. For permissions, please e-mail: [journals.permissions@oup.com](mailto:journals.permissions@oup.com)



CRISPR was reported for the first time in 1987, it was observed at the 3' end of the *iap* gene in *Escherichia coli* (Ishino et al. 1987). The official CRISPR name was assigned in 2002, when it was observed that these sequences were associated with the *cas* genes (Jansen et al. 2002a). An analysis of the spacers present in the CRISPR arrays from 67 prokaryotic genomes, including *Enterobacteriaceae* members, demonstrated that these sequences derive from foreign genetic elements such as phages and plasmids (Mojica et al. 2005). The study of spacers in other enterobacteria like *Yersinia pestis* reinforced their extrachromosomal origin and also showed that the addition of new spacers occurs in a polarized fashion, adjacent to the leader sequence (Pourcel, Salvignol and Vergnaud 2005). In 2007, Barrangou et al. demonstrated the function of CRISPR-Cas as a bacterial specific immune system in *Streptococcus thermophilus*, showing that the resistance was obtained by the incorporation of a small fragment of the foreign genetic material into the CRISPR sequences. One year later it was proposed in *E. coli* K12 that, after the expression of CRISPR as pre-crRNA, cleavage by a Cas endonuclease contained in a multiprotein complex known as Cascade (CRISPR-associated complex for antiviral defense) generates small mature RNAs (crRNA), which guide this protein complex to the exogenous genetic element (Brouns et al. 2008). Then, the Cas3 nuclease is recruited for targeting the invader, leading to its degradation (Westra et al. 2012; Mulepati and Bailey 2013). Thus, *E. coli* has been used as a model for elucidating the molecular mechanism of CRISPR-Cas function.

CRISPR-Cas systems are present in 87% and 45% of archaeal and bacterial genomes, respectively (Grissa, Vergnaud and Pourcel 2007) and are classified into two main classes. Class 1 members all use a multisubunit crRNA-effector complex for CRISPR-Cas immunity (Makarova et al. 2015). The members of this Class are Type I, Type III and Type IV systems, which are characterized by the presence of Cas3, Cas10 and Csf1, respectively. The Type I systems are organized in the subtypes I-A, I-B, I-C, I-D, I-E, I-F and I-U, whereas Type III comprises the subtypes III-A, III-B, III-C and III-D. No subtypes have been described for Type IV. In Class 2 systems, a single protein and a tracrRNA perform the biological functions of the multi-subunit crRNA-effector complex for immunity. This class is composed by Type II, Type V and Type VI CRISPR-Cas systems, whose signature proteins are Cas9, Cpf1 (Cas12) and Cas13, respectively. Type II is divided into three subtypes (II-A, II-B and II-C), Type V contains six subtypes (V-A, V-B, V-C, V-D, V-E and V-U), and Type VI is comprised by the VI-A, VI-B1, VI-B2 and VI-C subtypes. The subtypes are defined by specific *cas* genes signature (Makarova et al. 2011; Makarova et al. 2015; Koonin, Makarova and Zhang 2017).

Fundamental knowledge about the function, transcription, distribution and utility of the Class 1, Type I-E and I-F CRISPR-Cas has been generated for members of *Enterobacteriaceae*. This article presents the state of the art of the CRISPR-Cas systems in this bacterial family.

### The CRISPR-Cas immunological system in *Enterobacteriaceae*

The performance of Type I-E and Type I-F as defense mechanisms against mobile genetic elements, such as phages and plasmids, has been best characterized in *E. coli* and *Pectobacterium atrosepticum*, respectively. In particular, the functional model for the *E. coli* Type I-E CRISPR-Cas system is reviewed below, having been divided into three stages: (i) adaptation or

spacer acquisition, (ii) crRNA biogenesis or expression and (iii) interference.

#### Adaptation/spacer acquisition

The adaptation stage initiates when the bacterial CRISPR-Cas system detects invasive genetic material derived from phages or plasmids. Cas1 and Cas2 are fundamental for spacer acquisition and integration (Yosef, Goren and Qimron 2012; Nuñez et al. 2014). They correspond to a metal-dependent nuclease/integrase and an endoribonuclease (Makarova et al. 2006; Beloglazova et al. 2008; Wiedenheft et al. 2009; Babu et al. 2011). These proteins assemble in a complex formed by two Cas1 dimers that sandwich one central Cas2 dimer (Cas<sub>1</sub><sub>2</sub>-Cas<sub>2</sub>-Cas<sub>1</sub><sub>2</sub>) (Nuñez et al. 2014). This complex selects a DNA fragment from the invader (protospacer), presumably derived from the RecBCD-mediated degradation at stalled replication forks of the exogenous genetic element (Levy et al. 2015). The selection of the protospacer is determined by the presence of a protospacer adjacent motif (PAM) (Mojica et al. 2009), composed by the 5'-AAG-3' nucleotides (Swarts et al. 2012; Yosef, Goren and Qimron 2012; Yosef et al. 2013) that interact with the Tyr165, Trp170, Tyr188, Tyr217 and Lys211 residues, located at the C-terminal of Cas1 (Wang et al. 2015). The foreign DNA fragment is then excised, probably by Cas1 nuclease activity, but the precise mechanism and the enzymes involved remain unknown.

However, the study of crystallized Cas1-Cas2 complex bound to a protospacer reveals that the resulting 33-nt fragment is composed of 23 bp dsDNA, flanked by 5-nt 3' single-stranded overhangs with 3'-OH groups (Wang et al. 2015). These hydroxyl groups are used by each Cas1 dimer to catalyze the integration of the protospacer in the leader-proximal repeat of the CRISPR array (Datsenko et al. 2012; Swarts et al. 2012; Yosef, Goren and Qimron 2012). This appears to happen through a two-step mechanism that initiates when IHF binds to the leader sequence inducing DNA bending (Nuñez et al. 2016; Yoganand et al. 2017). This allows the positioning of the first Cas1 dimer and one 3'-OH end of the protospacer in proximity to the phosphate of the first nucleotide in the repeat sequence sense strand, to make a nucleophilic attack that induces cleavage and ligation with the target CRISPR repeat. In the next step, the second Cas1 dimer and the remaining 3'-OH of the protospacer perform another nucleophilic attack in the 28th base of the antisense strand, resulting in a fully integrated spacer flanked at both sites by a single strand of the CRISPR repeat (Arslan et al. 2014; Nuñez et al. 2015). Finally, DNA polymerase I complements the single strands and a DNA ligase completes the process, which involves the duplication of the first repeat sequence (Yosef, Goren and Qimron 2012; Ivančić-Baće et al. 2015). This process is known as naïve adaptation and implies the acquisition of a novel spacer from a foreign element to which no previous immunity exists. The presence of a spacer into the CRISPR array promotes the acquisition of various spacers in a second reinfection from the previously encountered invader: this biological process is termed primed adaptation (Datsenko et al. 2012; Swarts et al. 2012).

It is clear that Cas1 has a fundamental role in naïve and primed adaptation. However, Cas2 is also necessary for these processes, since the Cas2 dimer bridges the two Cas1 dimers to produce a functional Cas1-Cas2 complex; additionally, Cas2 positions the Cas1 dimers at the correct distance to obtain a 33-nt protospacer. Furthermore, Arg residues of Cas2 are implicated in the binding and stabilization of the dsDNA of the exogenous protospacer (Wang et al. 2015), although its nuclease activity is

dispensable for spacer acquisition (Nuñez et al. 2014; Nuñez et al. 2015).

#### Expression/crRNA biogenesis

The result of the adaptation is the acquisition of foreign DNA fragments into the CRISPR sequence. Then, when the bacteria deal with new attacks from the same invader, they express the *cas3*, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, *cas2* genes and the CRISPR array. The gene products are assembled into the multienzymatic Cascade complex conformed by one monomer of Cse1, two of Cse2, six of Cas7, and one of Cas5 and Cas6e (Brouns et al. 2008; Jore et al. 2011; Wiedenheft et al. 2011a). The CRISPR array composed of the repeats and spacer sequences is transcribed as a long precursor crRNA (pre-crRNA) (Brouns et al. 2008; Pougach et al. 2010; Al-Attar et al. 2011) where the repeats acquire a hairpin structure (Kunin, Sorek and Hugenholtz 2007), which functions as a signal for Cascade binding. This pre-crRNA is processed by the Cascade endonuclease Cas6e yielding small guide crRNAs (Brouns et al. 2008). Each crRNA is composed by 7 nucleotides from the 5' repeat sequence, a 33-bp spacer and 21 nucleotides from the 3' repeat (Jore et al. 2011; Sashital, Jinek and Doudna 2011; Sinkunas et al. 2013).

#### Interference

In the interference, the ribonucleoprotein complex Cascade-crRNA, through Cse1, recognizes the PAM sequence in the foreign genetic element (Jore et al. 2011; Sashital, Wiedenheft and Doudna 2012; Hayes et al. 2016). Then, the crRNA base pairs with its complementary DNA strand in the invader (protospacer), promoting the formation of an R-loop structure and Cas3 recruitment (Jore et al. 2011; Westra et al. 2012; Mulepati and Bailey 2013). Cas3 HD single-stranded DNA nuclease domain nicks and degrades the non-complementary strand of the R-loop in a 3' → 5' direction (Mulepati and Bailey 2011; Sinkunas et al. 2011; Westra et al. 2012), and by its ATP-dependent helicase activity promotes the displacing of the crRNA-Cascade from the invader DNA, generating a new single strand DNA available for Cas3 degradation (Howard et al. 2011; Westra et al. 2012; Sinkunas et al. 2013). Thus, the target DNA is eliminated and the bacterium can resist the invasive genetic element.

Foreign genetic elements can escape CRISPR-Cas recognition by mutating PAM sequences or the region with homology to the corresponding spacer (Deveau et al. 2008), but bacteria contend with this through multiple spacer acquisition (primed adaptation), ensuring a successful CRISPR-Cas-mediated defense (Datsenko et al. 2012; Fineran et al. 2014).

#### The Type I-F CRISPR-Cas immunological system

This genetic system has been widely studied in *P. atrosepticum*, where it is composed by CRISPR1, *cas1*, *cas2-3*, *cys1*, *cys2*, *cys3*, *cas6f* and CRISPR3. The Cas2-3 protein contains a Cas2 domain in the N-terminus and a C-terminal Cas3 helicase and HD nuclease domains, while the Cys1, Cys2, Cys3 and Cas6f proteins are part of a multiprotein complex (Cys complex) (Wiedenheft et al. 2011b; Richter et al. 2012). In the adaptation stage Cas1 interacts with the Cas2-3 protein forming a complex (Cas<sub>14</sub>–Cas<sub>2-3</sub>) involved in spacer capture and integration into the CRISPR locus (Fagerlund et al. 2017), thus participating in naïve and primed adaptation. Priming in *P. atrosepticum* is about 500 times more efficient than naïve adaptation (Staals et al. 2016), and since the new acquired spacers, by priming, target both strands of the invader, a highly efficient interference is assured (Richter et al. 2014).

In crRNA biogenesis, *cas1*–*cas2-3*–*cys1*–*cys2*–*cys3*–*cas6f* are expressed as an operon, whereas the CRISPR arrays contain independent promoters and are transcribed to form the pre-crRNA. The Cas6f nuclease alone is responsible for pre-crRNA processing to generate mature crRNAs (Haurwitz et al. 2010; Przybilski et al. 2011; Richter and Fineran 2013). The complex Cas6f-crRNA promotes interaction with Cys3, Cys2 and Cys1 yielding the Cys complex (Richter et al. 2012; Richter and Fineran 2013), which is guided to the target DNA in a crRNA-dependent manner.

Finally, for interference, the Cys complex recruits the Cas2-3 protein (Richter et al. 2012) and the Cas3 nuclease domain degrades the invader genetic element.

The performance of the Type I-E and Type I-F CRISPR-Cas system has been described here in order to illustrate its complexity in fairly well-studied organisms. However, it is fundamental to continue this line of research in other bacteria, in order to have a better understanding of its diversity in prokaryotes.

#### Spacer composition of CRISPRs in Enterobacteriaceae

According to the CRISPR-Cas mechanism as an immunological system, the spacers contained in the CRISPR arrays provide memory of the foreign genetic elements that have invaded to the bacteria. In this regard, several analyses of spacer composition have been performed in members of the *Enterobacteriaceae* family.

The first studies about spacer origin and composition in *Enterobacteriaceae* were performed in *Y. pestis*. The evaluation of 307 *Y. pestis* strains, including genome sequences of reference strains and isolates from China, the former Soviet Union and Mongolia, revealed 131 unique spacers: 59% (77 spacers) had homologous sequences to a prophage located in the *Y. pestis* CO92 genome, whereas 17% (22 spacers) were homologous to other DNA regions, including the IS285 and 23S rRNA genes, of the *Y. pestis* chromosome (Pourcel, Salvignol and Vergnaud 2005; Cui et al. 2008). A subsequent analysis of 100 *Y. pestis* strains from Mongolia identified seven new spacers that matched *Y. pestis* genes that encoded conserved hypothetical proteins, putative phage proteins, *cas3*, as well as genes encoded in the pCD1 plasmid (Riehm et al. 2012). A *Y. pestis* study using strains from a plague foci in Brazil showed 21 new region specific spacers: 11 of them were similar to virus, plasmid and chromosomal sequences (Barros et al. 2014). Regarding other *Yersinia* species, 1902 different spacers were identified in 335 *Yersinia pseudotuberculosis* strains. Only ~10% of the spacers had identity with sequences in the database using BLASTN: these included 38 and 14 spacers that showed significant similarity to prophages of *Yersinia enterocolitica* 8081 and *Y. pestis*, respectively. In this study it was also reported that 32 spacers matched with the cryptic plasmid pYptb32953 of *Y. pseudotuberculosis* IP32953, and one spacer was similar to an *E. coli* plasmid (Koskela et al. 2015). In the case of the *Yersinia frederiksenii* Y225 and the *Yersinia kristensenii* Y231 strains, both contained five spacers with significant identity to *Y. enterocolitica* plasmid pYE854. Finally, in the non-virulent *Yersinia similis* 228 strain, four spacers similar to *Yersinia* virulence plasmids were observed (Tan et al. 2016).

Relevant studies about spacer composition in other *Enterobacteriaceae* genus have also been performed. For instance, in *E. coli* strains from the ECOR collection, 52 different spacers were found. Two of them had similarity with enterobacterial conjugative plasmids, and other two were similar to enterobacterial phages (Mojica et al. 2005). A second study of 100 *E. coli* strains (28 available genomes and 72 strains from the ECOR collection)

identified a total of 745 unique spacers: 7% were homologous to phages (including prophages from the *E. coli* E24377A and *E. coli* SE11 genomes), 3.6% were similar to plasmids and 1.8% shared homology with non-mobile elements (Diez-Villaseñor et al. 2010). These data were extended with the analysis of 290 strains (including 263 natural isolates) that determined 1278 different spacers: 3.9% showed identity with phages, 2.1% with plasmids and 0.87% with chromosomal regions (Touchon et al. 2011). In another study, the evaluation of 185 *E. coli* isolates from Australian animals (kangaroo, Tasmanian devil, reptile, cattle, dog, horse, duck, bird, fish and rodent), humans and environmental waters, determined that the most prevalent spacer in these strains targets the intestinal P7 phage and the *E. coli* pO111 plasmid (Sheludchenko et al. 2015). In respect to pathogenic *E. coli*, 252 Shiga toxin-producing *E. coli* (STEC) strains (including the serogroups O26, O45, O103, O111, O121, O145 and O157) were analyzed and 432 unique spacers were identified, 35 of them matched known phages, plasmids or putative phage regions (Yin et al. 2013). In another work about STEC, 190 strains show 528 distinct spacers: 10 of them had identities with sequences from plasmids of *Salmonella enterica* or *E. coli*, and 1 spacer shared identity to bacteriophage P7. These 11 spacers were present in several STEC strains (Toro et al. 2014). In an analysis of 164 *E. coli* human isolates (82 Uropathogenic *E. coli* (UPEC) and 82 fecal *E. coli* strains), 227 distinct spacers were detected: 5.3% of them shared homology with phages and 13.7% were homologous to plasmids, such as the pB8, R100 and pSB102 that are known to harbor antibiotic resistance genes (Dang et al. 2013).

Regarding the *Salmonella* genus, from 400 clinical isolates, including *S. enterica* serovar Enteritidis, Typhimurium, Newport and Heidelberg, 179 distinct spacers were identified: 10% of the spacers were found in phages or prophages, 2% were similar to plasmids and 15% matched fragments of different bacterial chromosomes, including those of *Salmonella*, *E. coli*, *Klebsiella* sp. and *Erwinia* sp. (Shariat et al. 2015). In 92 *Salmonella* isolates from farms of Wisconsin and Minnesota, 395 unique spacers were found: only seven and two independent spacers shared significant identity with *Salmonella* bacteriophages and enterobacteria-related phages, respectively (Wehnes et al. 2014).

In the case of the *Shigella* genus, 61 isolates (48 *Shigella flexneri*, 11 *Shigella sonnei* and 2 *Shigella dysenteriae*) were evaluated and 32 distinct spacer sequences were founded. From nine spacers matching chromosomal DNA from different microorganisms, one of them was similar to a plasmid, a second one showed similarity to the *Shigella* gene encoding N-acetylmuramoyl-L-alanine amidase and a third spacer matched the *cas2-3* gene of *Yersinia* (Guo et al. 2015). An extensive work was performed using 237 *Shigella* strains encompassing the four species of this genus (*S. flexneri*, *S. sonnei*, *S. dysenteriae* and *Shigella boydii*). The results showed 31 distinct spacers, all of them matched plasmids and phages from a wide range of microorganisms. Even more, 24 of these spacers also had homology with *Shigella* chromosomal sequences (Yang et al. 2015). In the *Shigella* genus, the presence of CRISPR spacers that matched with the *cas* genes, *cas3*, *cas1*, *csy3*, has also been reported (Wang et al. 2016).

A study about spacer composition in *Enterobacteriaceae* was performed with 51 complete genomes from *Escherichia*, *Shigella* and *Salmonella*. The results reveal 594 unique spacers: 4.38% of them showed similarity to bacterial chromosomes, 3.5% with plasmids and 1.18% with viral sequences. Several of these spacers matched bacteriophages and plasmids associated with *Escherichia* and *Salmonella* (Touchon and Rocha 2010).

The evaluation of 24 genomes from different *Klebsiella* species (*Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Klebsiella varicola* and *Klebsiella michiganensis*) containing CRISPR arrays shows that

from 351 unique spacer sequences, 8.0% had similarity with microbial plasmids (these includes 12 spacer sequences that matched to *Klebsiella* plasmids), 1.4% were similar to phages, 22.5% matched prophages and 3.4% show significant identity with sequences of bacterial chromosomes, including one spacer that target a gene encoding for a membrane protein of *Klebsiella* and four spacers similar to *cas* genes (Shen et al. 2017). An exclusive work with 52 *K. pneumoniae* genomes showed that only six of them have CRISPR arrays. The spacer composition reveals that from 116 spacer sequences, 13% showed identity with genes from phages, 8% had homology with plasmids, 5% were similar to genes of *Klebsiella* spp. genomes and 7% were homologous to genes of other bacteria (Ostria-Hernández et al. 2015).

The genome of *Proteus mirabilis* NO-051/03 contains 20 spacers: three of them are homologous to prophage sequences and one spacer matches with a DNA fragment from conjugative IncL/M-type plasmids (D'Andrea et al. 2016).

All the studies described above provide information about spacer origin and composition in different human pathogens belonging to the *Enterobacteriaceae* family. Regarding enterobacterial plant pathogens, in *Erwinia amylovora* the evaluation of 37 isolates identified 454 distinct CRISPR spacers; some of them were homologous to DNA fragments of plasmids (21.8%), viral genes (1.3%) and prophages (4.0%) (Rezzonico, Smits and Duffy 2011). A second study of 85 *E. amylovora* strains showed the presence of 588 unique spacers. 16% shared homology to plasmids, 5% matched with phages (most of them targeted to the *Erwinia tasmaniensis* phage  $\Phi$ Et88), and 2% were similar to bacterial chromosome sequences (McGhee and Sundin 2012). In other *Erwinia*, such as *Erwinia piriflorinigrans* CFBP5888<sup>T</sup>, analysis of 23 spacers reveals that some of them have homology with the *E. tasmaniensis* phage  $\Phi$ Et88 and with prophages of *Erwinia pyrifoliae* DSM12163 and *E. amylovora* ATCC BAA-2158 (Smits et al. 2013).

Regarding *Pectobacterium*, the CRISPR arrays from *P. atrosepticum* SCRI1043, *Pectobacterium carotovorum* subsp. *carotovorum* PCC21 and *Pectobacterium wasabiae* WPP163 were analyzed using CRISPRTarget. 41, 62 and 64 spacers, respectively, were identified. Remarkably, these three genomes contain spacers (two of *P. atrosepticum*, four of *P. carotovorum* and eight of *P. wasabiae*) that matched with prophages associated with *Pectobacterium* genomes. Additionally, *P. atrosepticum* SCRI1043 harbors a self-complementary spacer sharing nucleotide identity with the *eco0560* gene located in the pathogenicity island HAI2 (Biswas et al. 2013).

In general, a discrete amount of spacers from different *Enterobacteriaceae* members appear to derive from extrachromosomal genetic elements such as virus or plasmids. Remarkably, some strains harbor spacers with identity to sequences of their own genome. However, a great amount of *Enterobacteriaceae* spacers have not homology with any sequence in the databanks. Hence, future studies are needed to determine in which *Enterobacteriaceae* genus, the spacers together with the direct repeats and Cas proteins function as an immunological system or they are implicated in other processes such as gene regulation, bacterial physiology or virulence.

### Transcriptional organization and regulation of the *Enterobacteriaceae* CRISPR-Cas systems Type I-E and I-F

The process of CRISPR-Cas immunity has been divided into three stages: adaptation, expression and interference. Since CRISPR-Cas expression is pivotal for adaptation and interference, studies about the genetic elements and signals that control CRISPR-Cas transcription, as well as its transcriptional organization,

have been mainly performed in *Enterobacteriaceae* such as *E. coli* K-12, *S. enterica* serovar Typhi IMSS-1 and *P. atrosepticum* SCRI1043.

The Type I-E CRISPR-Cas system of the commensal *E. coli* K-12 is composed of *cas3*, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1*, *cas2* and the CRISPR1 array (Haft et al. 2005; Touchon and Rocha 2010). This genetic locus contains four independent promoters. One of these is present in the 5' intergenic region of *cas3* controlling the expression of this nuclease. The second promoter, anti-*Pcas*, is found in the antisense strand of the *cas3*–*cse1* intergenic region and drives the expression of a small RNA (150–200 bp) without an assigned function. Another promoter is located upstream of the *cse1* gene (*Pcas*) and controls the transcription of the entire *cse1*–*cse2*–*cas7*–*cas5*–*cas6e*–*cas1*–*cas2* operon. Finally, a regulatory region is located in the leader sequence of the CRISPR array (*Pcrispr1*) and drives pre-crRNA transcription (Pul et al. 2010; Westra et al. 2010).

Positive and negative genetic elements involved in the control of these promoters have been identified. In the case of *Pcas*, the global regulators H-NS and CRP participate in its silencing (Pul et al. 2010; Yang et al. 2014), whereas BaeR and LeuO are involved in its positive regulation (Westra et al. 2010; Perez-Rodriguez et al. 2011). For *cas3*, anti-*Pcas* and *Pcrispr1* promoters, negative regulation mediated by H-NS has been demonstrated (Pul et al. 2010; Majsec, Bolt and Ivančić-Baće 2016). Independent *hns* and *crp* mutants as well as overexpression of LeuO promote activity of the CRISPR-Cas system that induces resistance to phage invasion (Westra et al. 2010; Yang et al. 2014).

In the human pathogen *S. Typhi* IMSS-1, studies about the transcriptional organization and control of its CRISPR-Cas system demonstrated that this locus contains five transcriptional units. Two of them, the *cse1*–*cse2*–*cas7*–*cas5*–*cas6e*–*cas1*–*cas2*-CRISPR (*cas*-CRISPR operon) and *scse2* (*sense cse2* RNA) are transcribed from the sense strand, whereas *ascse2-1* (*antisense RNA of cse2 to cse1*) and *ascas2-1* (*antisense RNA of cas2 to cas1*) are present on the antisense strand. Additionally, the *S. Typhi cas3* gene is transcribed as an independent unit divergent to the *cas*-CRISPR operon (Medina-Aparicio et al. 2017). The transcriptional activity of the *cse1*–*cse2*–*cas7*–*cas5*–*cas6e*–*cas1*–*cas2*-CRISPR polycistronic mRNA is negatively regulated by H-NS and positively induced by LeuO as in *E. coli* K12 (Hernández-Lucas et al. 2008; Medina-Aparicio et al. 2011). However, the Leucine response protein (Lrp) has also been implicated in its repression, although, in contrast, no effect of this regulator was observed in the regulation of the *E. coli* K12 CRISPR-Cas locus (Pul et al. 2010; Medina-Aparicio et al. 2011). In the case of *cas3* and *ascse2-1* transcriptional units, the role of H-NS in silencing their expression has also been demonstrated.

The transcriptional activity of the five genetic elements present in the *S. Typhi* CRISPR-Cas loci is induced by basic pH, a condition encountered by the bacterium during human infection (Medina-Aparicio et al. 2017). Recent data support the role of the  $\sigma^{54}$  factor in the transcriptional control of the *S. Typhimurium cas1* gene (Samuels et al. 2013), suggesting that this genetic element could respond to diverse stress conditions such as nitrogen starvation or the presence of heavy metals.

The *K. pneumoniae* NTUH-K2044 strain contains a set of Type I-E *cas* genes composed by *cas3*, *cse1*, *cse2*, *cas6e*, *cas7*, *cas5*, *cas1* and *cas2* that are flanked by two CRISPR arrays. In this *Enterobacteriaceae*, it has been demonstrated that H-NS indirectly mediates plasmid acquisition, since it is implicated in the negative regulation of *cas3*. Therefore, in the *hns* mutant, the CRISPR-Cas system is induced and the acquisition and stability of engineered plasmids diminishes. Interestingly, the presence of the antibiotic imipenem induces *hns* transcriptional activity, inhibit-

ing the nuclease Cas3 and, thus increasing plasmid acquisition (Lin et al. 2016).

The studies performed in *K. pneumoniae*, *Salmonella* and *E. coli* are fundamental in the characterization, transcriptional organization and regulation of the Type I-E CRISPR-Cas systems. Additionally, studies on the transcription of the subtype I-F CRISPR-Cas have been carried out in *P. atrosepticum* SCRI1043. In this phytopathogen, the CRISPR-Cas system encompasses CRISPR1, which is divergent to the *cas1*–*cas2-3*–*csy1*–*csy2*–*csy3*–*cas6f* operon; CRISPR3 is located downstream of the *cas* genes, and CRISPR2 is separated from this genetic region by a putative toxin-antitoxin system (ECA3686-7). Each CRISPR array contains its own promoter and an additional *cas1* promoter controls the expression of a polycistronic mRNA that includes the *cas* genes and CRISPR3 (Przybilski et al. 2011). Therefore, in *P. atrosepticum* as well as in *S. Typhi*, the CRISPR-*cas* locus forms an operon. Also in *P. atrosepticum*, the CRP-cAMP complex interacts with the *cas1* regulatory region for transcriptional activation (Patterson et al. 2015). Plasmid interference and spacer acquisition was not detectable in the *crp* or *cyaA* (adenylate cyclase that produces cAMP) *P. atrosepticum* mutants, supporting the functional regulatory performance of these genes in the CRISPR-Cas immunity of this bacterium (Patterson et al. 2015). In contrast, CRP is a negative regulator of the *E. coli* K12 CRISPR-Cas expression (Yang et al. 2014).

CRISPR-Cas regulation has also been described in *Serratia* sp. ATCC39006 that contains a Type I-E, I-F and III-A CRISPR-Cas systems (Fig. 2). Patterson et al. (2016) demonstrated that the transcriptional expression of these three immunological systems is controlled by a quorum sensing regulatory network that includes the SmaI and SmaR proteins. SmaI controls the production of AHL (*N*-acyl homoserine lactone), a chemical signal that interacts with the repressor protein SmaR, preventing its negative effect on the genetic expression of the three CRISPR-Cas systems, thus allowing spacer acquisition and interference of foreign genetic elements (Patterson et al. 2016).

Hence, multiple transcriptional factors and signals are involved in the regulation of CRISPR-Cas in *Enterobacteriaceae*. H-NS represses more than one CRISPR-Cas transcriptional unit in *E. coli* K12 and *S. Typhi* IMSS-1 (Pul et al. 2010; Medina-Aparicio et al. 2011). Moreover, this histone-like protein also represses *cas3* activity in *K. pneumoniae* NTUH-K2044 (Lin et al. 2016), and it is suggested that phages utilize their own H-NS-like protein to repress CRISPR-Cas for bacterial predation (Skenneron et al. 2011). Therefore, it is possible that H-NS is a general transcriptional repressor of CRISPR-Cas in *Enterobacteriaceae*. However, specific proteins such as LeuO, Lrp and CRP are also involved in CRISPR-Cas regulation in different members of this proteobacterial family (Westra et al. 2010; Medina-Aparicio et al. 2011; Yang et al. 2014; Patterson et al. 2015). Interestingly, a recent study demonstrated that the absence of HtpG in *E. coli* disrupts CRISPR-Cas immunity, since this chaperone stabilizes the Cas3 nuclease (Yosef et al. 2011), suggesting that this system is also controlled at the post-translational level. The involvement of global regulatory proteins in CRISPR-Cas regulation suggests that this genetic system can be turned on or off in response to distinct environmental conditions and signals.

### The CRISPR-Cas system in enterobacterial-host interaction

It is clear that in bacteria and archaea the CRISPR-Cas genetic elements are an immunological system that avoids phage predation and lateral gene transfer to protect bacterial populations.

However, this system has also been related to other biological processes.

The influence of CRISPR-Cas in bacteria-host interaction has been documented. For instance, CRISPR-Cas has been implicated in the mutualistic association of the enterobacteria *Xenorhabdus* and nematodes of the genus *Steinernema*. This biological association allows the infection and killing of a wide range of insects (Stock and Blair 2008). *Xenorhabdus nematophila* has six CRISPR arrays (A–G) and a Type I-E set of *cas* genes. The CRISPR-F contains a spacer (32 bp) and two repeats (29 bp each) that encode for the *NilD* RNA whose processing is Cas6e dependent. Independent disruption of *cas6e* or *nilD* inhibits the symbiotic interaction of *X. nematophila* with multiple nematode species, demonstrating that Cas6e is required for *NilD* processing, and this mature RNA is fundamental for the bacteria-nematode relationship (Heungens, Cowles and Goodrich-Blair 2002; Veessenmeyer et al. 2014). Therefore, *X. nematophila* CRISPR-Cas system is involved in mutualistic associations.

In the human pathogen *S. enterica* serovar Typhi, the expression of *cas7* (STY3068) has been detected in macrophages infected with this bacterium (Faucher, Curtiss and Daigle 2005). Furthermore, the *S. Typhi* CRISPR-Cas locus contains five promoters that are expressed at pH 7.5 (Medina-Aparicio et al. 2017), a pH value characteristic of the distal ileum (Evans et al. 1988), an intestinal compartment that is invaded and colonized by this pathogen. Moreover, CRISPR-Cas is regulated by *LeuO*, *H-NS* and *Lrp*, which are involved in pathogenesis (Navarre et al. 2006; Baek et al. 2009; Guadarrama, Villaseñor and Calva 2014). These data suggest that CRISPR-Cas could be implicated in the interaction of *S. Typhi* with human cells such as macrophages or intestinal epithelial cells.

Furthermore, in the plant pathogen *P. atrosepticum*, analysis of microarray data showed that its *cas* genes are expressed during potato infection (Przybilski et al. 2011).

Regarding *E. coli*, uropathogenic strains have less spacers and repeats than commensal fecal strains isolated from the same person (Dang et al. 2013). A negative correlation has also been established between the CRISPR content and the pathogenicity of *E. coli* and related strains; moreover, the higher the number of virulence factors within a strain, the lower the amount of repeat content (Garcia-Gutierrez et al. 2015). Fewer spacers were also found in STEC strains associated with outbreaks (Toro et al. 2014). In contrast to the scenario described above, more spacers were found in avian pathogenic *E. coli* strains, than in avian fecal commensal *E. coli* (Fu et al. 2017). It is clear that future studies on the correlation between virulence encoded genes and CRISPR content are needed to determine the role of the CRISPR-Cas system in *E. coli* pathogenesis.

Although the role of CRISPR-Cas system in pathogenic *E. coli* remains unknown, it is suggested that the Type I-E CRISPR-Cas loci of the commensal *E. coli* K12 W3110 can function in DNA repair, since Babu et al. (2011) demonstrated that an *E. coli* K12 strain deficient in *cas1* has increased sensibility to DNA damage (mitomycin C and UV light) compared with the wild-type strain. Moreover, the Cas1 nuclease is able to cleave Holliday junctions, DNA intermediates produced during DNA repair. Additionally, Cas1 interacts with the *RecB*, *RecC* and *RuvB* proteins that are involved in this process. This interesting result prompts to determine whether the role of CRISPR-Cas system in DNA repair is a widespread function in *Enterobacteriaceae* and other microorganisms.

The role of the CRISPR-Cas loci present in different commensal, pathogens and symbionts from the *Enterobacteriaceae* family should be addressed in order to define how this system

influences bacterial physiology, pathogenesis or other host interactions.

### CRISPR-Cas distribution in *Enterobacteriaceae*

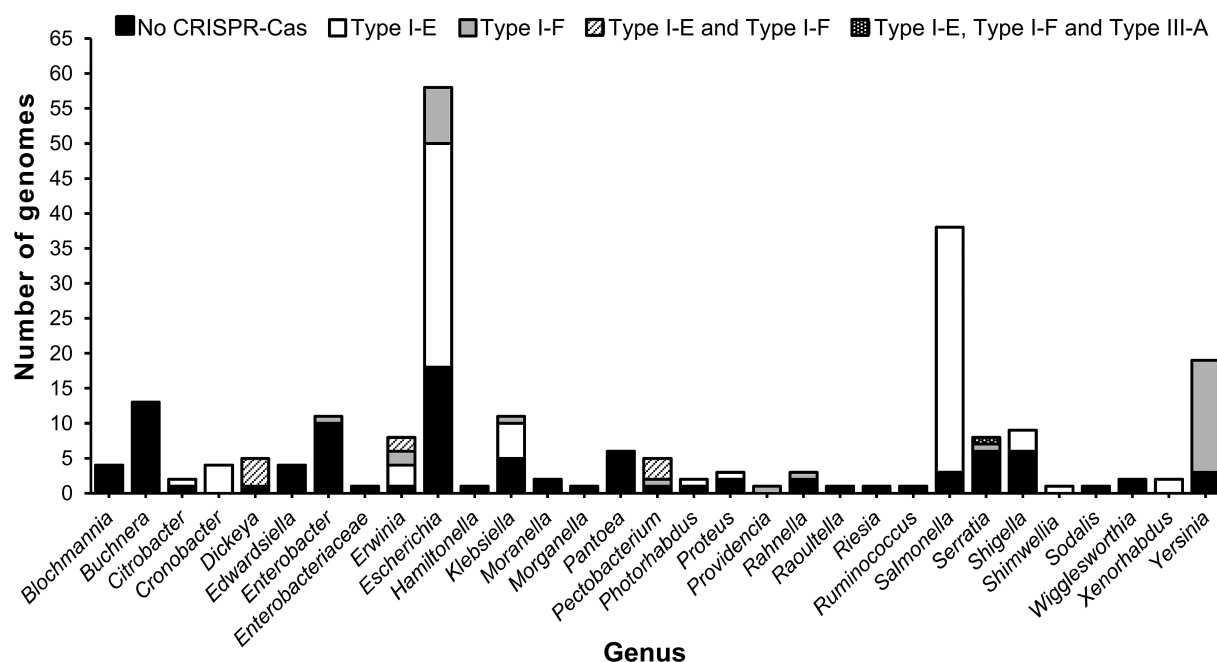
The *Enterobacteriaceae* family is represented by 56 genera (Euzéby 1997). To determine the presence, distribution and frequency of the CRISPR-Cas systems in this family, a bioinformatic search was performed. The analysis of 228 complete enterobacterial genomes available in the NCBI database, corresponding to 31 genera, shows that the Type I-E system alone is present in 38.6% of these genomes. The occurrence of only the Type I-F system was detected in 14%; both subtypes I-E and I-F are contained in 4%; and only one genome (0.44%) had three CRISPR-Cas systems (Type I-E, Type I-F and Type III-A). Remarkably, 43% of the genomes evaluated lacked the *cas* genes (Fig. 1).

These data are represented in a phylogenetic tree that shows the enterobacterial species and their corresponding CRISPR-Cas Type systems (Fig. 2). In the case of *E. coli*, which can be found as a pathogen or commensal in the intestine of a variety of animals including humans, four CRISPR arrays have been detected, CRISPR1, CRISPR2, CRISPR3 and CRISPR4, as well as the *cas* genes, *cas3*, *cse1*, *cse2*, *cas7*, *cas5*, *cas6e*, *cas1* and *cas2*, that belong to the Type I-E (or *E. coli* subtype) and are associated with the CRISPR1 loci. The second set of *E. coli* *cas* genes are *cas1-cas2-3-csy1-csy2-csy3-cas6f* (*csy4*) that together with the CRISPR3 array constitute a CRISPR-Cas Type I-F system. CRISPR2 and CRISPR4 arrays are not associated with *cas* genes (Haft et al. 2005; Diez-Villaseñor et al. 2010; Touchon and Rocha 2010). Either Type I-E or I-F set of *cas* genes were found in 39 of 57 *E. coli* genomes analyzed (Fig. 2). None of the strains evaluated harbor both types of systems, in agreement with previous observations (Diez-Villaseñor et al. 2010). Interestingly, 18 of the 57 *E. coli* completely sequenced genomes analyzed do not have any *cas* genes. The other species of the *Escherichia* genus, *Escherichia fergusonii* ATCC 35469 harbors only one complete Type I-E CRISPR-Cas system (Fig. 2).

CRISPR-Cas systems were also determined in human pathogens of the *Yersinia* genus. Four and twelve genomes of *Y. pseudotuberculosis* and *Y. pestis*, respectively, possess the *cas1-cas2-3-csy1-csy2-csy3-cas6f* genes that together with the CRISPR locus named YPa constitute a Type I-F CRISPR-Cas system (Fig. 2). Indeed, this was first identified in several *Y. pestis* strains, and so was previously known as the Ypest subtype (Haft et al. 2005). A characteristic of this system is that *cas2-3* encodes for a Cas3 protein fused to Cas2 (Makarova et al. 2011). *Yersinia pestis* and *Y. pseudotuberculosis* strains usually harbor two additional CRISPR arrays (YPb and YPc) that are not associated with *cas* genes (Pourcel, Salvignol and Vergnaud 2005; Cui et al. 2008). In the case of *Y. enterocolitica*, three sequenced strains were evaluated and the presence of a CRISPR-Cas system was not observed (Fig. 2), in agreement with previous studies (Tan et al. 2016).

The *Dickeya* genus includes plant pathogens that infect a wide variety of ornamental plants and economically important vegetables such as potato, causing maceration-associated diseases (Czajkowski et al. 2011). *Dickeya dadantii* (Ech703, Ech586, 3937) and *Dickeya zeae* Ech1591 contain both the I-E and I-F CRISPR-Cas subtypes (Fig. 2) (McGhee and Sundin 2012). Interestingly, *Dickeya solani* 3337 does not present any CRISPR-Cas system (Pédrón et al. 2014).

The *Pectobacterium* genus comprises economically important phytopathogens that produce soft-rot and blackleg disease in potatoes (De Boer, Li and Ward 2012). Five complete



**Figure 1.** Distribution of CRISPR-Cas loci in *Enterobacteriaceae*. The presence of this genetic system was determined in 228 fully sequenced *Enterobacteriaceae* genomes. The frequency of CRISPR-Cas Type I-E, Type I-F, both Type I-E and I-F, or three systems (I-E, I-F and III-A) in distinct genera of this family is shown. The height of each column indicates the number of genomes evaluated for each genus.

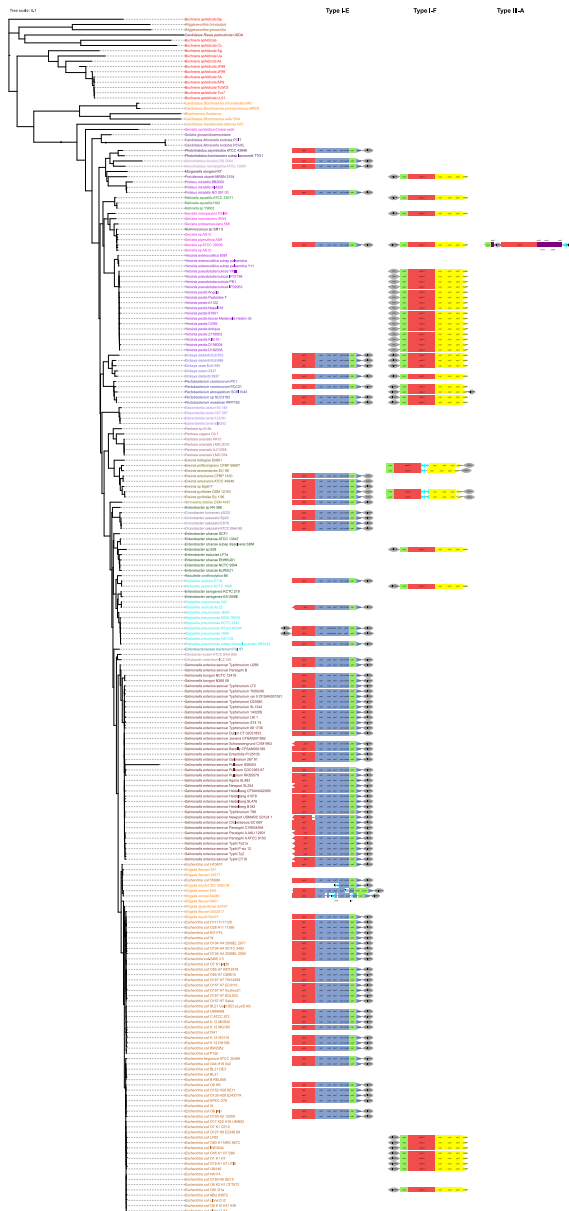
*Pectobacterium* genomes were analyzed, showing that *P. carotovorum* PCC21, *Pectobacterium* sp. SCC3193 and *P. wasabiae* WPP163 harbor a Type I-F and a Type I-E CRISPR-Cas immunological systems (Fig. 2). Additional CRISPR arrays have been reported for *P. carotovorum* PCC21 and *P. wasabiae* WPP163 (Biswas et al. 2013). Another species, *P. atrosepticum* SCRI1043 (formerly *Erwinia carotovora* subsp. *atroseptica*) contains a set of Type I-F *cas* genes and three CRISPR arrays (CRISPR1, CRISPR2 and CRISPR3). CRISPR1 is located upstream *cas1* and CRISPR3 is downstream *cas6f* (Fig. 2), whereas CRISPR2 is not associated with any *cas* gene (Przybilski et al. 2011). The fifth genome corresponds to *P. carotovorum* PC1, which does not have a CRISPR-Cas immunological system.

The *Erwinia* genus is composed mostly of plant pathogens such as *E. amylovora* (CFBP1430, ATCC 49946) and *Erwinia* sp. Ejp617 that show the presence of a Type I-E CRISPR-Cas locus (Fig. 2). *Erwinia tasmaniensis* Et1 99 contains a Type I-F set of genes with an insertion of two ORFs between *cas3* and *csy1* and the CRR4 CRISPR array is located downstream of *cas6f*. In opposite, the location of the CRISPR loci in other Type I-F systems is upstream of *cas1* (Fig. 2). A similar genetic arrangement has been observed in the pear pathogens *E. piriflorinigrans* CFBP 5888T (Smits et al. 2013), *E. pyrifoliae* DSM 12163 and *E. pyrifoliae* Ep1 96. However, in these cases three hypothetical genes were inserted between *cas3* and *csy1* (Fig. 2). The *E. pyrifoliae* strains also carry a Type I-E CRISPR-Cas system. None of the *Erwinia* described above contain additional *cas*-associated CRISPR arrays (Rezzonico, Smits and Duffy 2011; McGhee and Sundin 2012; Smits et al. 2013). Finally, in *Erwinia billingiae* Eb661 the CRISPR-Cas locus was absent.

In *Klebsiella*, which includes nosocomial opportunistic pathogens, species such as *K. oxytoca* E718 have a Type I-E set of *cas* genes associated with CRISPR1, but *K. oxytoca* KCTC 1686 contains a Type I-F *cas* genes downstream to CRISPR4 (Fig. 2) (Shen et al. 2017), demonstrating that it is possible to find either

of these CRISPR-Cas subtypes in the same species. In the case of *K. variicola* At 22, the presence of a Type I-E CRISPR-Cas system was observed, with *cas3* divergent from *cse1* (Fig. 2). An additional orphan CRISPR locus occurs in the last two *Klebsiella* strains mentioned above (Shen et al. 2017). The analysis of eight sequenced genomes of *K. pneumoniae* showed that *K. pneumoniae* subsp. *rhinoscleromatis* SB3432 harbors a canonical Type I-E CRISPR-Cas system (Fig. 2), whereas *K. pneumoniae* NTUH K2044 and *K. pneumoniae* 1084 contain a Type I-E system consisting of CRISPR2-*cas3*-*cse1*-*cse2*-*cas6e*-*cas7*-*cas5*-*cas1*-*cas2*-CRISPR3. It differs from the *E. coli* Type I-E because in that microorganism *cas6e* is located downstream of *cas5*, while in the strains of *K. pneumoniae*, *cas6e* is present between *cse2* and *cas7* (Fig. 2) (Ostria-Hernández et al. 2015; Shen et al. 2017). In *K. pneumoniae* 342, JM45, MGH 78578, KCTC 2242 and HS1128 the presence of CRISPR-Cas was not observed. Therefore, this genetic system is not widely distributed in *K. pneumoniae*, in accordance with previous studies showing that only 6 of 52 *K. pneumoniae* genomes analyzed had a CRISPR-Cas locus (Ostria-Hernández et al. 2015). Remarkably, a Type III CRISPR-Cas system in *K. pneumoniae* strains recovered from dogs, cats, sheep and hedgehogs was present in the multiresistance plasmid pENVA (Schluter et al. 2014).

The *Salmonella* genus is composed of human and animal pathogens as well as non-pathogenic strains such as *Salmonella bongori*. Two CRISPR arrays (CRISPR1 and CRISPR2) have been identified in *Salmonella* and only CRISPR1 is associated with a Type I-E set of *cas* genes (Touchon and Rocha 2010). In 35 of the 38 *Salmonella* genomes analyzed, the Type I-E CRISPR-Cas system is present (Fig. 2). Remarkably, in the *Salmonella enterica* serovars Schwarzengrund, Newport, Paratyphi A and Typhi, *cas3* is divergent to the *cse1*-*cse2*-*cas7*-*cas5*-*cas6e*-*cas1*-*cas2* genes (Fabre et al. 2012; Shariat et al. 2015). Finally, *S. enterica* serovars Pullorum S06004, Javiana and Paratyphi B do not have any *cas* genes (Fig. 2).



**Figure 2.** Phylogenetic tree showing the presence of the CRISPR-Cas systems in *Enterobacteriaceae*. The enterobacterial genus and its corresponding species are represented in different colors and the CRISPR-Cas systems I-E, I-F and III-A are schematized. For the Type I-E CRISPR-Cas system, *cas3* is presented in red, *cse1-cse2-cas7-cas5-cas6e* and *cas2* are blue, and *cas1* is shown in green. In the Type I-F CRISPR-Cas locus, the *cas1* gene is represented in green, *cas2-3* in red and *csy1-csy2-csy3-cas6f* are shown in yellow. The Type III-A are depicted as follows, *cas1* is shown in green, *cas2*, *csm2*, *csm3*, *csm4*, *csm5* and *cas6* are in purple, and *cas10* is in red. The CRISPR arrays are represented in gray and insertions (ORFs or IS elements) are in light blue. The backbone of the phylogenetic tree was obtained with the distance data from the 16S ribosomal gene and the CRISPR-Cas arrays were obtained with the GenCont II program using as seed the *cas1* gene. For the analysis, 228 completely sequenced genomes obtained from the NCBI database were used.

The *Shigella* genus is composed of the human pathogens *S. boydii*, *S. sonnei*, *S. flexneri* and *S. dysenteriae*, all of them causing shigellosis or bacillary dysentery (Hale 1991). The evaluation of nine complete *Shigella* genomes shows that *S. boydii* Sb227 contains an intact *cas3-cse1-cse2-cas7-cas5-cas6e-cas1-cas2* locus. However, *S. boydii* CDC 308394 has deletions of some *cas*

genes. Duplication of *cas* genes and IS insertions in the *cas* array is observed in *S. sonnei* Ss046 and *S. sonnei* 53G. Although all the strains described above contain different *cas* arrangements, the CRISPR1 loci is always located downstream of the *cas2* gene (Fig. 2) (Wang et al. 2016). In *S. flexneri* 301, *S. flexneri* 2457T, *S. flexneri* 8401 and *S. flexneri* 2002017 only the presence of *cas3* and a CRISPR array has been observed, whereas three CRISPR loci are present in *S. dysenteriae* Sd197 (Wang et al. 2016). Additionally, *cas1* and *cas2* point mutations have been detected in this bacterial genus (Yang et al. 2015). Remarkably, all these kinds of genetic rearrangements in the CRISPR-Cas system are a specific feature of this *Enterobacteriaceae* genus.

In the case of the genus *Proteus*, *P. mirabilis* NO-051/03, an isolate from skin and soft tissue infections, contains a Type I-E CRISPR-Cas locus and an orphan CRISPR array (Fig. 2) (D'Andrea et al. 2016). However, *P. mirabilis* BB2000 and HI4320 did not show any CRISPR-Cas locus.

The analysis of seven sequenced genomes from the genus *Serratia* shows that only *S. marcescens* FG194 harbors a Type I-F CRISPR-Cas system (Fig. 2). In *Serratia* spp. ATCC39006, the presence of four CRISPR layouts is reported. CRISPR1 is located downstream from Type I-E *cas* genes, CRISPR2 is grouped with Type I-F *cas* genes, whereas CRISPR3 and CRISPR4 are associated with a Type III-A immunological system (Fig. 2) (Patterson et al. 2016). Remarkably, until now this *Serratia* strain is unique, as it is the only reported enteric bacterium containing three different types of CRISPR-Cas systems.

In general, for various *Enterobacteriaceae* genera such as *Xenorhabdus*, *Shimwellia* and *Cronobacter*, the analysis of two, one and four genomes, respectively, showed that all the species evaluated contain a Type I-E CRISPR-Cas system. In the case of *Photorhabdus* and *Citrobacter*, the presence of a Type I-E CRISPR-Cas was observed in one of the two genomes analyzed for each genus (Fig. 2).

For *Providencia*, the genome analysis of *Providencia stuartii* MRSN 2154 demonstrated that it contains a Type I-F CRISPR-Cas system. For the genus *Rahnella* and *Enterobacter*, three and eleven sequenced genomes, respectively, were evaluated and showed that only *Rahnella aquatilis* ATCC 33071 and *Enterobacter* sp. 638 carry a Type I-F CRISPR-Cas locus (Fig. 2).

In the remaining *Enterobacteriaceae* genera including *Buchnera*, *Wigglesworthia*, *Riesia*, *Blochmania*, *Hamiltonella*, *Sodalis*, *Moranella*, *Morganella*, *Ruminococcus*, *Edwardsiella*, *Pantoea*, *Raoultella* and *Enterobacteriaceae*, the CRISPR-Cas systems were absent (this study).

The results described above show that the predominant CRISPR-Cas system in *Enterobacteriaceae* is the Type I-E, and the presence of both Type I-E and Type I-F is not common. The presence of three distinct CRISPR-Cas systems occurs in only one *Serratia* strain. The distribution analysis also shows that the presence of the CRISPR-Cas array is not regular in some *Enterobacteriaceae* species, for instance *Rahnella aquatilis* ATCC 33071 contains a CRISPR-Cas locus while *Rahnella aquatilis* HX2 does not. The same phenomenon is also observed in other bacteria like *Serratia marcescens*, *Pectobacterium carotovorum*, *Enterobacter* sp., *Klebsiella pneumoniae* or *S. enterica* serovar Pullorum.

### The CRISPR sequences: an alternative phylogenetic tool in *Enterobacteriaceae*

CRISPR sequences are widely distributed in bacterial genomes and composed by direct repeats and spacers. Since spacers show chronological acquisition and a high degree of polymorphism,

they have been successfully used for molecular typing in *Mycobacterium tuberculosis* (Kamerbeek et al. 1997), *S. thermophilus* (Horvath et al. 2008), *Campylobacter jejuni* (Schouls et al. 2003) or *Corynebacterium diphtheriae* (Mokrousov et al. 2007). Studies using CRISPR to determine phylogenetic relationships of closely related taxa, source tracking and geographical origin of different members of the *Enterobacteriaceae* family have also been performed.

*Yersinia pestis*, the causative agent of human plague (Perry and Fetherston 1997), is the first *Enterobacteriaceae* where the use of CRISPR for genotyping was proposed (Pourcel, Salvignol and Vergnaud 2005; Vergnaud et al. 2007). The analysis of CRISPR loci YPa, YPb and YPc (also known as YP1, YP2 or YP3, respectively) of isolates from insects, animals and patients from China and the former Soviet Union, established that strains with similar spacer arrays were associated with specific geographical regions and were useful for proposing a transmission route (Cui et al. 2008). Similar results have been observed with *Y. pestis* strains from Mongolia, Brazil, and clinical samples from patients of Madagascar (Riehm et al. 2012; Barros et al. 2014; Riehm et al. 2015). The CRISPR loci have also been utilized to distinguish *Y. pestis* from the foodborne pathogen *Y. pseudotuberculosis*, because they share only a few spacers (Pourcel, Salvignol and Vergnaud 2005; Vergnaud et al. 2007; Koskela et al. 2015). Thus CRISPR could be useful for source tracking during an outbreak of human plague, since this zoonotic disease continues to be a global health problem, with most cases reported in Africa, Asia and the Americas (WHO 2016).

The *Salmonella* genus includes two species, *S. enterica* and *S. bongori* where *S. enterica* includes more than 2600 serovars (Issenhuth-Jeanjean et al. 2014), many of which cause human and animal diseases. Multiple *Salmonella* outbreaks occur every year around the world and many antigenic and molecular typing methods are available to determine the serovar and the strains that provoke a specific outbreak. One of these methodologies corresponds to CRISPR-Multi-Virulence Locus Sequence Typing (MVLST), which analyzes CRISPR loci and virulence genes (*sseL* and *fimH*) (Liu et al. 2011a). This technique is able to discriminate among *Salmonella* serovars since isolates from the same serovar have similar CRISPR loci or share at least most of their spacer sequences (Fricke et al. 2011; Fabre et al. 2012). In addition, for some *Salmonella* serotypes, CRISPR-MVLST can differentiate individual outbreak strains from non-outbreak-related strains/clones, and allow for the efficient tracking of the pathogen source within the food system (Liu et al. 2011a; Shariat et al. 2013b).

The typing based only on CRISPR arrays allows the characterization of the predominant *Salmonella* populations in dairy farms, in order to prevent dissemination to human hosts (Wehnes et al. 2014), and provides a useful phylogenetic tool for detecting multiple *Salmonella* serovars in asymptomatic animal hosts like the Pennsylvania raccoon (Very et al. 2016). Additionally, the ease of obtaining different length variants of CRISPR1 and CRISPR2 using PCR has allowed the rapid differentiation of *S. Typhimurium* and *Salmonella* Newport outbreak strains, and this simple screening approach is suitable for low-capacity public health laboratories or hospitals (Fabre et al. 2012; Shariat et al. 2013b).

Interestingly, it is suggested that the amplification of only the newly incorporated spacers in CRISPR1 and CRISPR2 (CLSPT, CRISPR locus spacer pair typing) is good enough to discriminate *Salmonella* serovars (Li et al. 2014). In the case of the etiological agents of typhoid fever and paratyphoid fever, *S. Typhi* and *S. Paratyphi A*, respectively, it is known that CRISPR2

amplification allows the distinction of these two serovars (Fabre et al. 2014).

Regarding the highly clonal Enteritidis serovar, one of the major causes of food-borne salmonellosis in the United States, several reports show that CRISPR-MVLST together with PFGE are highly accurate typing methods to discriminate chicken, egg, environmental and human clinical isolates (Liu et al. 2011b; Shariat et al. 2013a). This methodology has also made possible the identification of the predominant Enteritidis sequence types present in specific geographical regions (Deng et al. 2015) or to determine that the origin of human infections in Tahiti was derived from farm animals (Le Hello et al. 2015).

In other serovars such as Typhimurium, a bacterium able to induce gastroenteritis in humans, it has been demonstrated that the CRISPR-MVLST is a very efficient approach for subtyping Brazilian isolates, since it complements and validates the results obtained by PFGE (Almeida et al. 2017). In addition, a relation between CRISPR sequence types and antibiotic resistance was observed in this *Salmonella* serovar present in animals and human isolates (DiMarzio et al. 2013). Furthermore, based on CRISPR polymorphism, the CRISPOL method was developed for Typhimurium, allowing the creation of an open web-accessible database, providing a serotype/spacer dictionary internationally available for strain tracking (Fabre et al. 2012).

The use of CRISPR for typing clinical isolates has been used for other *Salmonella* serovars such as Heidelberg, Virchow and Pullorum, showing that this methodology is very useful for epidemiological *Salmonella* studies (Shariat et al. 2013c; Bachmann et al. 2014; Xie et al. 2017). Several studies had demonstrated the helpfulness of CRISPR arrays for subtyping different *S. enterica* serovars; however, the use of CRISPR-Cas loci for *Salmonella* microevolutionary studies needs to be carefully interpreted since no congruence has been detected between evolutionary patterns obtained from CRISPR or *cas* genes analysis and phylogeny based on whole genome sequence (Timme et al. 2013; Pettengill et al. 2014).

For *E. amylovora*, the causal agent of fire blight, an important disease that causes considerable economic losses of apple and pear crops worldwide (Oh and Beer 2005), low genetic diversity by pulsed-field gel electrophoresis (PFGE), ribotyping or rep-PCR fingerprinting analysis has been observed (McManus and Jones 1995; Jock et al. 2002; Rico et al. 2008). However, using CRISPR arrays named CRR-1, CRR-2 and CRR-3, it was possible to separate *E. amylovora* strains according to their host plants: apple, pear, quince, raspberry and Indian hawthorn. The strains from apple and pear were even classified into three groups depending on their geographical origin: group I included *E. amylovora* strains from Europe, New Zealand and the Middle East, and group II and III contained isolates from the United States (Rezzonico, Smits and Duffy 2011). McGhee and Sundin (2012) differentiated *E. amylovora* strains isolated in the eastern and western United States, and proposed that an *E. amylovora* strain from the eastern USA could be the source of fire blight disease in New Zealand and Europe. Thus, CRISPRs can also be effectively utilized for strain tracking.

Other members of the *Enterobacteriaceae* family are the human pathogens belonging to the genus *Shigella*, which still constitutes a public health problem mainly affecting children within the ages of 1–4 years (Niyogi 2005). Only one study suggested that CRISPR could be a new tool for discrimination of *Shigella* subtypes (Yang et al. 2015). In these organisms, only CRISPR1 is adjacent to the Type I- E *cas* genes, seven additional orphan CRISPR arrays were described and the presence from two to seven CRISPR arrays are detected in *Shigella* strains (Guo et al.



2015; Wang et al. 2016). An evaluation of 237 strains of the four *Shigella* species (*S. sonnei*, *S. flexneri*, *S. dysenteriae* and *S. boydii*) showed that point mutations in the *cas1* and *cas2* genes are characteristic of some *S. boydii* subtypes. Moreover, it was observed that the different *Shigella* subtypes contain particular point mutations in their repeat sequences. Thus, the analysis of the repeat sequences, spacer content and *cas* genes may represent a tool for the detection and typing of *Shigella* during an outbreak (Yang et al. 2015). However, studies using a representative amount of isolates from several geographic regions are needed to validate the advantage in the use of CRISPR loci for *Shigella* distinction.

With regard to *E. coli* pathogenic strains, studies using its CRISPR arrays show that the STEC strains of serotype O113:H21 have a high degree of diversity (Feng et al. 2014). CRISPR polymorphism has been efficiently used for a specific detection of Enterohemorrhagic *Escherichia coli* (EAEC) O104:H4, the bacteria that caused the 2011 outbreaks in Germany and France (Delannoy et al. 2012a). In STEC, including those of O157:H7 and the 'big six' serotypes (O26, O45, O103, O111, O121 and O145), the CRISPR alleles within each serogroup were generally similar in spacer content and order, demonstrating that CRISPR is an effective way of screening STEC strains (Yin et al. 2013). Furthermore, studies with Enterohemorrhagic *E. coli* (EHEC) clearly support the application and utility of the CRISPR loci in identification and molecular serotyping of EHEC O157:H7, O26:H11, O145:H28, O103:H2, O111:H8, O121:H19 and O45:H2 (Delannoy, Beutin and Fach 2012b).

In AIEC (Adherent-invasive *E. coli*), CRISPR-associated proteins Cas5 and Cas6 and a hypothetical protein (LF82034; YP002555197) were uniquely conserved in the B2 AIEC group and define a subset of pathogenic strains belonging to this phylogenetic group (Deshpande et al. 2015). Therefore, in some *E. coli* strains, CRISPR can be used for the benefit of public health (Delannoy, Beutin and Fach 2012b; Delannoy et al. 2015; Ison et al. 2015; Delannoy, Beutin and Fach 2016). However, it was also reported that in a large panel of strains including human and animal samples, there was no obvious link between the presence, size, or content of CRISPR and a specific lifestyle or origin, suggesting caution in using CRISPR as an epidemiological marker in *E. coli* (Touchon et al. 2011). In this sense, studies on the genetic diversity of isolates from Australian animals, humans and environmental waters also show that CRISPR array analysis alone was not effective as a source tracking tool for these *E. coli* (Sheludchenko et al. 2015). Furthermore, comparison of CRISPR1 alleles from *E. coli* strains of fecal and aquatic environments shows that CRISPR alleles had similar spacer repertoires, indicating that these sequences are not sufficient for distinguishing these two *E. coli* populations (Tymensen 2016).

The use of CRISPR for typing should also be applied with caution in *Klebsiella*, since the CRISPR-Cas locus is not present in all the *K. pneumoniae* sequenced genomes (Ostria-Hernández et al. 2015), and CRISPR arrays can vary in the same *Klebsiella* sequence type (Shen et al. 2017).

In summary, the genotyping or subtyping based on CRISPR arrays has been documented and validated in certain *Enterobacteriaceae* genera. Indeed, the results obtained by CRISPR analysis correlate with MLVA, SNP, PFGE or whole-genome single-nucleotide-polymorphism typing methods (Riehm et al. 2012; Shariat et al. 2013a; Shariat et al. 2013b; Shariat et al. 2013c; Deng et al. 2015; Riehm et al. 2015; Very et al. 2016; Almeida et al. 2017; Xie et al. 2017). Thus, the characterization of CRISPR sequences is a pivotal alternative phylogenetic tool in enterobacteria.

## Genome engineering in *Enterobacteriaceae* using its endogenous CRISPR-Cas systems

The CRISPR-Cas system has been used for genome engineering in bacteria, worms, zebrafish, plants, cell lines, mice and rats. To modify the DNA an exogenous Cas9 protein from the type II system and a guide RNA are used (Doudna and Charpentier 2014; Zhang, Wen and Guo 2014; Schmid-Burgk 2017). In *Enterobacteriaceae*, specifically in the plant pathogen *P. atrosepticum*, it is possible to use its endogenous Type I-F CRISPR-Cas system to delete specific chromosomal regions. This was demonstrated by Vercoe et al. (2013) who engineered a plasmid that contains three spacer-repeat units for targeting the chromosomal gene encoding for the N-acylhomoserine lactone synthase (*expI*), an enzyme involved in quorum sensing. Using this plasmid and the natural expression of the *P. atrosepticum* Type I-F CRISPR-Cas system, they showed that chromosomal targeting is highly efficient. With this system, the authors were able not only to modify one gene, but were also able to remove a 100-kb pathogenic island (Vercoe et al. 2013), providing evidence that the Type I-F CRISPR-Cas system is a useful tool for specific genome editing. This methodology was also applied in *Serratia* sp. to create point mutations in the flagella master regulator FlhC (Hampton et al. 2016). Hence, endogenous CRISPR-Cas systems of recalcitrant microorganisms can be used as an alternative genetic tool to modify their genomes. Therefore, it is possible to use a wide range of CRISPR-Cas systems for genome engineering in prokaryotic cells. It is fundamental to continue with CRISPR research, since until now the Type II system is the only editing system that can be applied to many organisms.

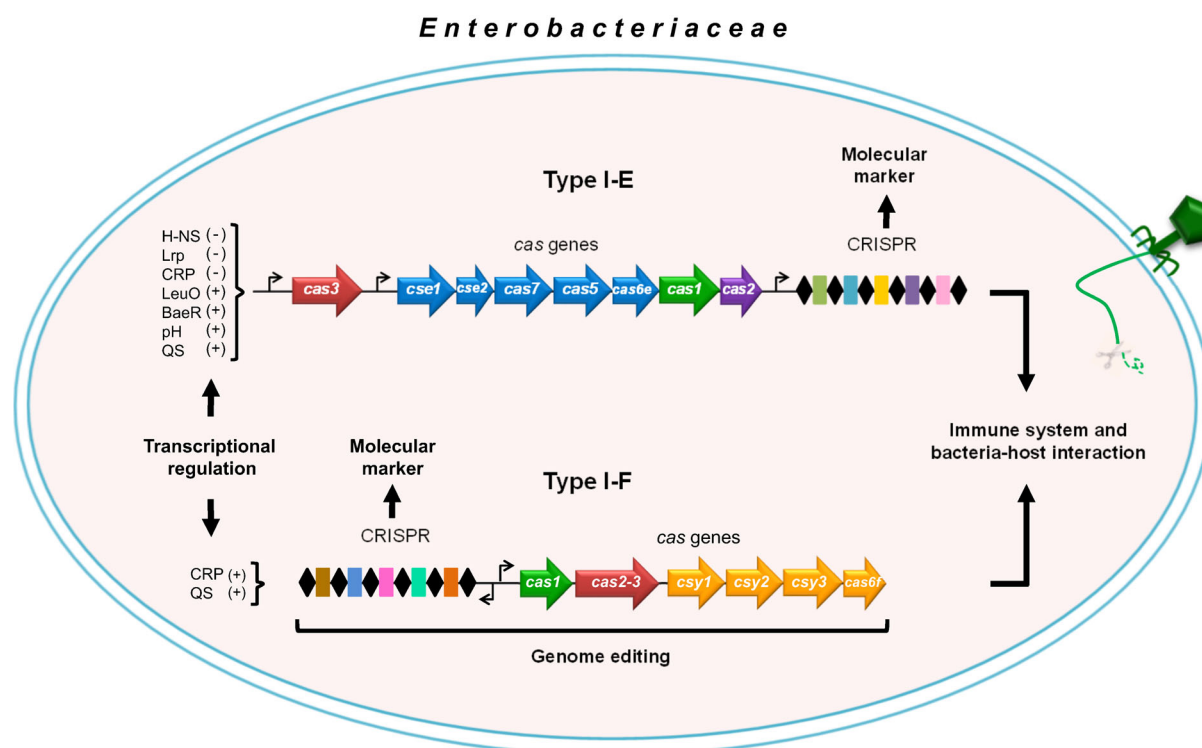
Finally, in Fig. 3, we summarize all the topics developed in this CRISPR-Cas review.

## CONCLUSIONS

In this review, the state of the art of the CRISPR-Cas system in *Enterobacteriaceae* has been described. However, in spite of the valuable amount of information generated so far, fundamental questions still remain, such as why only 43% of the members of this family contain this genetic system. Determining the factors that influence its loss or permanence is a matter for exciting future research. Another remarkable aspect that we need to address is whether other biological functions in the pathogenic or free-living state of *Enterobacteriaceae* are performed by the CRISPR-Cas systems, as well as the biological factors that influence their diversity. Better understanding of the modes of genetic regulation in different isolates should also allow for an improved comprehension in this respect. It is clear that we need to recognize that there is still a way to go before we can understand the plethora of questions and answers that the study of CRISPR-Cas in the *Enterobacteriaceae* family offers.

## ACKNOWLEDGEMENTS

We would like to thank M. Dunn (CCG-UNAM) and to I. Olivares-Casique, S. Mayo-Pérez, M. Navarrete-Flores and D. Angeles-Salazar from IBT-UNAM for stimulating discussions and critical reading. The authors also thank to A. Vázquez, F. J. Santana, L. Perezgasga and M. Fernández-Mora from IBT-UNAM for their useful help along this work.



**Figure 3.** CRISPR-Cas system in *Enterobacteriaceae*. Type I-E and Type I-F are the prevalent CRISPR-Cas systems in the *Enterobacteriaceae* family. They function as an immune system against phages and plasmids, and have also a role in bacteria-host interaction. Regulatory proteins such as Lrp, H-NS, CRP, LeuO and BaeR, as well as external signals like pH or Quorum sensing (QS), are involved in the transcriptional control of these two types CRISPR-Cas loci. Furthermore, the CRISPR sequences are a molecular marker used for typing, and the Type I-F system is useful for editing *Enterobacteriaceae* genomes.

## FUNDING

This work was supported by grants from Dirección General de Asuntos de Personal Académico, DGAPA/UNAM (IN203215 to I. H.-L.) and Consejo Nacional de Ciencia y Tecnología, CONACyT, México (127298 to I. H.-L). L. M.-A. was supported by a predoctoral fellowship from CONACyT, México (No. 2208 53).

**Conflicts of interest.** None declared.

## REFERENCES

- Al-Attar S, Westra ER, van der Oost J et al. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol Chem* 2011;**392**:277–89.
- Almeida F, Medeiros MIC, DdP Rodrigues et al. Molecular characterization of *Salmonella* Typhimurium isolated in Brazil by CRISPR-MVLST. *J Microbiol Methods* 2017;**133**: 55–61.
- Arrieta MC, Stiemsma LT, Amenyogbe N et al. The intestinal microbiome in early life: health and disease. *Front Immunol* 2014;**5**:427.
- Arslan Z, Hermanns V, Wurm R et al. Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system. *Nucleic Acids Res* 2014;**42**:7884–93.
- Babu M, Beloglazova N, Flick R et al. A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. *Mol Microbiol* 2011;**79**:484–502.
- Bachmann NL, Petty NK, Ben Zakour NL et al. Genome analysis and CRISPR typing of *Salmonella enterica* serovar Virchow. *BMC Genomics* 2014;**15**:389.
- Baek CH, Wang S, Roland KL et al. Leucine-responsive regulatory protein (Lrp) acts as a virulence repressor in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 2009;**191**:1278–92.
- Barrangou R, Fremaux C, Deveau H et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007;**315**:1709–12.
- Barros MP, Franca CT, Lins RH et al. Dynamics of CRISPR loci in microevolutionary process of *Yersinia pestis* strains. *PLoS One* 2014;**9**:e108353.
- Beloglazova N, Brown G, Zimmerman MD et al. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J Biol Chem* 2008;**283**:20361–71.
- Biswas A, Gagnon JN, Brouns SJ et al. CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. *RNA Biol* 2013;**10**:817–27.
- Breitbart M, Rohwer F. Here a virus, there a virus, everywhere the same virus? *Trends Microbiol* 2005;**13**:278–84.
- Brouns SJ, Jore MM, Lundgren M et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 2008;**321**:960–4.
- Buckle GC, Walker CL, Black RE. Typhoid fever and paratyphoid fever: systematic review to estimate global morbidity and mortality for 2010. *J Glob Health* 2012;**2**:010401.
- Cui Y, Li Y, Gorge O et al. Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced short palindromic repeats. *PLoS One* 2008;**3**:e2652.
- Czajkowski R, Pérombelon MCM, van Veen JA et al. Control of blackleg and tuber soft rot of potato caused by *Pectobacterium* and *Dickeya* species: a review. *Plant Pathol* 2011;**60**:999–1013.
- D’Andrea MM, Giani T, Henrici De Angelis L et al. Draft genome sequence of *Proteus mirabilis* NO-051/03, representative of a

- multidrug-resistant clone spreading in Europe and expressing the CMY-16 AmpC-Type beta-Lactamase. *Genome Announc* 2016;**4**:e01702-15.
- Dang TN, Zhang L, Zollner S et al. Uropathogenic *Escherichia coli* are less likely than paired fecal *E. coli* to have CRISPR loci. *Infect Genet Evol* 2013;**19**:212-8.
- Datsenko KA, Pougach K, Tikhonov A et al. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun* 2012;**3**:945.
- De Boer SH, Li X, Ward LJ. *Pectobacterium* spp. associated with bacterial stem rot syndrome of potato in Canada. *Phytopathology* 2012;**102**:937-47.
- Delannoy S, Beutin L, Burgos Y et al. Specific detection of enteroaggregative hemorrhagic *Escherichia coli* O104:H4 strains by use of the CRISPR locus as a target for a diagnostic real-time PCR. *J Clin Microbiol* 2012a;**50**:3485-92.
- Delannoy S, Beutin L, Fach P. Use of clustered regularly interspaced short palindromic repeat sequence polymorphisms for specific detection of enterohemorrhagic *Escherichia coli* strains of serotypes O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28, and O157:H7 by real-time PCR. *J Clin Microbiol* 2012b;**50**:4035-40.
- Delannoy S, Beutin L, Fach P. Improved traceability of Shiga-toxin-producing *Escherichia coli* using CRISPRs for detection and typing. *Environ Sci Pollut Res Int* 2016;**23**:8163-74.
- Delannoy S, Mariani-Kurkdjian P, Bonacorsi S et al. Characteristics of emerging human-pathogenic *Escherichia coli* O26:H11 strains isolated in France between 2010 and 2013 and carrying the *stx2d* gene only. *J Clin Microbiol* 2015;**53**:486-92.
- Deng X, Shariat N, Driebe EM et al. Comparative analysis of subtyping methods against a whole-genome-sequencing standard for *Salmonella enterica* serotype Enteritidis. *J Clin Microbiol* 2015;**53**:212-8.
- Deshpande NP, Wilkins MR, Mitchell HM et al. Novel genetic markers define a subgroup of pathogenic *Escherichia coli* strains belonging to the B2 phylogenetic group. *FEMS Microbiol Lett* 2015;**362**, DOI: 10.1093/femsle/fnv193.
- Deveau H, Barrangou R, Garneau JE et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* 2008;**190**:1390-400.
- Diez-Villaseñor C, Almendros C, Garcia-Martinez J et al. Diversity of CRISPR loci in *Escherichia coli*. *Microbiology* 2010;**156**:1351-61.
- DiMarzio M, Shariat N, Kariyawasam S et al. Antibiotic resistance in *Salmonella enterica* serovar Typhimurium associates with CRISPR sequence type. *Antimicrob Agents Chemother* 2013;**57**:4282-9.
- Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 2014;**346**:1258096.
- Euzeby JP. List of bacterial names with Standing in nomenclature: a folder available on the Internet. *Int J Syst Bacteriol* 1997;**47**:590-2.
- Evans DF, Pye G, Bramley R et al. Measurement of gastrointestinal pH profiles in normal ambulant human subjects. *Gut* 1988;**29**:1035-41.
- Fabre L, Le Hello S, Roux C et al. CRISPR is an optimal target for the design of specific PCR assays for *Salmonella enterica* serotypes Typhi and Paratyphi A. *PLoS Negl Trop Dis* 2014;**8**:e2671.
- Fabre L, Zhang J, Guigon G et al. CRISPR typing and subtyping for improved laboratory surveillance of *Salmonella* infections. *PLoS One* 2012;**7**:e36995.
- Fagerlund RD, Wilkinson ME, Klykov O et al. Spacer capture and integration by a Type I-F Cas1-Cas2-3 CRISPR adaptation complex. *Proc Natl Acad Sci USA* 2017;**114**:E5122-8.
- Faucher SP, Curtiss R, 3rd, Daigle F. Selective capture of *Salmonella enterica* serovar Typhi genes expressed in macrophages that are absent from the *Salmonella enterica* serovar Typhimurium genome. *Infect Immun* 2005;**73**:5217-21.
- Feng PC, Delannoy S, Lacher DW et al. Genetic diversity and virulence potential of Shiga toxin-producing *Escherichia coli* O113:H21 strains isolated from clinical, environmental, and food sources. *Appl Environ Microbiol* 2014;**80**:4757-63.
- Fineran PC, Gerritzen MJ, Suarez-Diez M et al. Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc Natl Acad Sci U S A* 2014;**111**:E1629-38.
- Fricke WF, Mammel MK, McDermott PF et al. Comparative genomics of 28 *Salmonella enterica* isolates: evidence for CRISPR-mediated adaptive sublineage evolution. *J Bacteriol* 2011;**193**:3556-68.
- Fu Q, Su Z, Cheng Y et al. Clustered, regularly interspaced short palindromic repeat (CRISPR) diversity and virulence factor distribution in avian *Escherichia coli*. *Res Microbiol* 2017;**168**:147-56.
- Garcia-Gutierrez E, Almendros C, Mojica FJ et al. CRISPR content correlates with the pathogenic potential of *Escherichia coli*. *PLoS One* 2015;**10**:e0131935.
- Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 2007;**35**:W52-7.
- Guadarrama C, Villaseñor T, Calva E. The subtleties and son-trasts of the LeuO regulator in *Salmonella* Typhi: Implications in the immune response. *Front Immunol* 2014;**5**:581.
- Guo X, Wang Y, Duan G et al. Detection and analysis of CRISPRs of *Shigella*. *Curr Microbiol* 2015;**70**:85-90.
- Haft DH, Selengut J, Mongodin EF et al. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol* 2005;**1**:e60.
- Hale TL. Genetic basis of virulence in *Shigella* species. *Microbiol Rev* 1991;**55**:206-24.
- Hampton HG, McNeil MB, Paterson TJ et al. CRISPR-Cas gene-editing reveals RsmA and RsmC act through FlhDC to repress the SdhE flavinylation factor and control motility and prodigiosin production in *Serratia*. *Microbiology* 2016;**162**:1047-58.
- Haurwitz RE, Jinek M, Wiedenheft B et al. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 2010;**329**:1355-8.
- Hayes RP, Xiao Y, Ding F et al. Structural basis for promiscuous PAM recognition in Type I-E Cascade from *E. coli*. *Nature* 2016;**530**:499-503.
- Hernández-Lucas I, Gallego-Hernández AL, Encarnación S et al. The LysR-type transcriptional regulator LeuO controls expression of several genes in *Salmonella enterica* serovar Typhi. *J Bacteriol* 2008;**190**:1658-70.
- Heungens K, Cowles CE, Goodrich-Blair H. Identification of *Xenorhabdus nematophila* genes required for mutualistic colonization of *Steinernema carpocapsae* nematodes. *Mol Microbiol* 2002;**45**:1337-53.
- Horvath P, Romero DA, Coute-Monvoisin AC et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol* 2008;**190**:1401-12.
- Howard JA, Delmas S, Ivančić-Baće I et al. Helicase dissociation and annealing of RNA-DNA hybrids by *Escherichia coli* Cas3 protein. *Biochem J* 2011;**439**:85-95.

- Ishino Y, Shinagawa H, Makino K et al. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 1987;169:5429–33.
- Ison SA, Delannoy S, Bugarel M et al. Genetic diversity and pathogenic potential of attaching and effacing *Escherichia coli* O26:H11 strains recovered from bovine feces in the United States. *Appl Environ Microbiol* 2015;81:3671–8.
- Issenhuth-Jeanjean S, Roggentin P, Mikoleit M et al. Supplement 2008–2010 (no. 48) to the White-Kauffmann-Le Minor scheme. *Res Microbiol* 2014;165:526–30.
- Ivančić-Baće I, Cass SD, Wearne SJ et al. Different genome stability proteins underpin primed and naive adaptation in *E. coli* CRISPR-Cas immunity. *Nucleic Acids Res* 2015;43:10821–30.
- Jansen R, Embden JD, Gaastra W et al. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 2002a;43:1565–75.
- Jansen R, van Embden JD, Gaastra W et al. Identification of a novel family of sequence repeats among prokaryotes. *OMICS* 2002b;6:23–33.
- Jock S, Donat V, Lopez MM et al. Following spread of fire blight in Western, Central and Southern Europe by molecular differentiation of *Erwinia amylovora* strains with PFGE analysis. *Environ Microbiol* 2002;4:106–14.
- Jore MM, Lundgren M, van Duijn E et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* 2011;18:529–36.
- Kamerbeek J, Schouls L, Kolk A et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997;35:907–14.
- Koonin EV, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* 2017;37:67–78.
- Koskela KA, Mattinen L, Kalin-Manttari L et al. Generation of a CRISPR database for *Yersinia pseudotuberculosis* complex and role of CRISPR-based immunity in conjugation. *Environ Microbiol* 2015;17:4306–21.
- Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol* 2007;8:R61.
- Le Hello S, Maillard F, Mallet HP et al. *Salmonella enterica* serotype Enteritidis in French Polynesia, South Pacific, 2008–2013. *Emerg Infect Dis* 2015;21:1045–8.
- Levy A, Goren MG, Yosef I et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* 2015;520:505–10.
- Li H, Li P, Xie J et al. New clustered regularly interspaced short palindromic repeat locus spacer pair typing method based on the newly incorporated spacer for *Salmonella enterica*. *J Clin Microbiol* 2014;52:2955–62.
- Lin TL, Pan YJ, Hsieh PF et al. Imipenem represses CRISPR-Cas interference of DNA acquisition through H-NS stimulation in *Klebsiella pneumoniae*. *Sci Rep* 2016;6:31644.
- Liu F, Barrangou R, Gerner-Smidt P et al. Novel virulence gene and clustered regularly interspaced short palindromic repeat (CRISPR) multilocus sequence typing scheme for subtyping of the major serovars of *Salmonella enterica* subsp. *enterica*. *Appl Environ Microbiol* 2011a;77:1946–56.
- Liu F, Kariyawasam S, Jayarao BM et al. Subtyping *Salmonella enterica* serovar Enteritidis isolates from different sources by using sequence typing based on virulence genes and clustered regularly interspaced short palindromic repeats (CRISPRs). *Appl Environ Microbiol* 2011b;77:4520–6.
- Majsec K, Bolt EL, Ivančić-Baće I. Cas3 is a limiting factor for CRISPR-Cas immunity in *Escherichia coli* cells lacking H-NS. *BMC Microbiol* 2016;16:28.
- Makarova KS, Grishin NV, Shabalina SA et al. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 2006;1:7.
- Makarova KS, Haft DH, Barrangou R et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 2011;9:467–77.
- Makarova KS, Wolf YI, Alkhnbashi OS et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* 2015;13:722–36.
- McGhee GC, Sundin GW. *Erwinia amylovora* CRISPR elements provide new tools for evaluating strain diversity and for microbial source tracking. *PLoS One* 2012;7:e41706.
- McManus PS, Jones AL. Genetic fingerprinting of *Erwinia amylovora* strains isolated from tree-fruit crops and *Rubus* spp. *Phytopathology (USA)* 1995;85:1547–53.
- Medina-Aparicio L, Rebollar-Flores JE, Beltrán-Luviano AA et al. CRISPR-Cas system presents multiple transcriptional units including antisense RNAs that are expressed in minimal medium and upregulated by pH in *Salmonella enterica* serovar Typhi. *Microbiology* 2017;163:253–65.
- Medina-Aparicio L, Rebollar-Flores JE, Gallego-Hernández AL et al. The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and Leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J Bacteriol* 2011;193:2396–407.
- Mojica FJ, Diez-Villaseñor C, Garcia-Martinez J et al. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 2009;155:733–40.
- Mojica FJ, Diez-Villaseñor C, Garcia-Martinez J et al. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* 2005;60:174–82.
- Mokrousov I, Limeschenko E, Vyazovaya A et al. *Corynebacterium diphtheriae* spoligotyping based on combined use of two CRISPR loci. *Biotechnol J* 2007;2:901–6.
- Mulepati S, Bailey S. Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J Biol Chem* 2011;286:31896–903.
- Mulepati S, Bailey S. *In vitro* reconstitution of an *Escherichia coli* RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. *J Biol Chem* 2013;288:22184–92.
- Navarre WW, Porwollik S, Wang Y et al. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 2006;313:236–8.
- Niyogi SK. Shigellosis. *J Microbiol* 2005;43:133–43.
- Nuñez JK, Bai L, Harrington LB et al. CRISPR immunological memory requires a host factor for specificity. *Mol Cell* 2016;62:824–33.
- Nuñez JK, Kranzusch PJ, Noeske J et al. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat Struct Mol Biol* 2014;21:528–34.
- Nuñez JK, Lee AS, Engelman A et al. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* 2015;519:193–8.
- Oh CS, Beer SV. Molecular genetics of *Erwinia amylovora* involved in the development of fire blight. *FEMS Microbiol Lett* 2005;253:185–92.

- Ostria-Hernández ML, Sánchez-Vallejo CJ, Ibarra JA et al. Survey of Clustered regularly interspaced short palindromic repeats and their associated Cas proteins (CRISPR/Cas) systems in multiple sequenced strains of *Klebsiella pneumoniae*. *BMC Res Notes* 2015;**8**:332.
- Patterson AG, Chang JT, Taylor C et al. Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res* 2015;**43**:6038–48.
- Patterson AG, Jackson SA, Taylor C et al. Quorum sensing controls adaptive immunity through the regulation of multiple CRISPR-Cas systems. *Mol Cell* 2016;**64**:1102–8.
- Pédrón J, Mondy S, des Essarts YR et al. Genomic and metabolic comparison with *Dickeya dadantii* 3937 reveals the emerging *Dickeya solani* potato pathogen to display distinctive metabolic activities and T5SS/T6SS-related toxin repertoire. *BMC Genomics* 2014;**15**:283.
- Perez-Rodríguez R, Haitjema C, Huang Q et al. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol Microbiol* 2011;**79**:584–99.
- Perry RD, Fetherston JD. *Yersinia pestis*—etiologic agent of plague. *Clin Microbiol Rev* 1997;**10**:35–66.
- Pettengill JB, Timme RE, Barrangou R et al. The evolutionary history and diagnostic utility of the CRISPR-Cas system within *Salmonella enterica* ssp. *enterica*. *PeerJ* 2014;**2**:e340.
- Pougach K, Semenova E, Bogdanova E et al. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol* 2010;**77**:1367–79.
- Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 2005;**151**:653–63.
- Przybilski R, Richter C, Gristwood T et al. Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA Biol* 2011;**8**:517–28.
- Pul U, Wurm R, Arslan Z et al. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol* 2010;**75**:1495–512.
- Rezzonico F, Smits TH, Duffy B. Diversity, evolution, and functionality of clustered regularly interspaced short palindromic repeat (CRISPR) regions in the fire blight pathogen *Erwinia amylovora*. *Appl Environ Microbiol* 2011;**77**:3819–29.
- Rico A, Fuhrer ME, Ortiz-Barredo A et al. Polymerase chain reaction fingerprinting of *Erwinia amylovora* has a limited phylogenetic value but allows the design of highly specific molecular markers. *Phytopathology* 2008;**98**:260–9.
- Richter C, Dy RL, McKenzie RE et al. Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res* 2014;**42**:8516–26.
- Richter C, Fineran PC. The subtype I-F CRISPR-Cas system influences pathogenicity island retention in *Pectobacterium atrosepticum* via crRNA generation and Csy complex formation. *Biochem Soc Trans* 2013;**41**:1468–74.
- Richter C, Gristwood T, Clulow JS et al. *In vivo* protein interactions and complex formation in the *Pectobacterium atrosepticum* subtype I-F CRISPR/Cas system. *PLoS One* 2012;**7**:e49549.
- Riehm JM, Projahn M, Vogler AJ et al. Diverse genotypes of *Yersinia pestis* caused plague in Madagascar in 2007. *PLoS Negl Trop Dis* 2015;**9**:e0003844.
- Riehm JM, Vergnaud G, Kiefer D et al. *Yersinia pestis* lineages in Mongolia. *PLoS One* 2012;**7**:e30624.
- Samuels DJ, Frye JG, Porwollik S et al. Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the  $\sigma^{54}$  (RpoN) regulon of *Salmonella* Typhimurium LT2. *BMC Genomics* 2013;**14**:602.
- Sashital DG, Jinek M, Doudna JA. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat Struct Mol Biol* 2011;**18**:680–7.
- Sashital DG, Wiedenheft B, Doudna JA. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol Cell* 2012;**46**:606–15.
- Schluter A, Nordmann P, Bonnin RA et al. IncH-type plasmid harboring bla CTX-M-15, bla DHA-1, and qnrB4 genes recovered from animal isolates. *Antimicrob Agents Chemother* 2014;**58**:3768–73.
- Schmid-Burgk JL. Disruptive non-disruptive applications of CRISPR/Cas9. *Curr Opin Biotechnol* 2017;**48**:203–9.
- Schouls LM, Reulen S, Duim B et al. Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. *J Clin Microbiol* 2003;**41**:15–26.
- Shariat N, DiMarzio MJ, Yin S et al. The combination of CRISPR-MVLST and PFGE provides increased discriminatory power for differentiating human clinical isolates of *Salmonella enterica* subsp. *enterica* serovar Enteritidis. *Food Microbiol* 2013a;**34**:164–73.
- Shariat N, Kirchner MK, Sandt CH et al. Subtyping of *Salmonella enterica* serovar Newport outbreak isolates by CRISPR-MVLST and determination of the relationship between CRISPR-MVLST and PFGE results. *J Clin Microbiol* 2013b;**51**:2328–36.
- Shariat N, Sandt CH, DiMarzio MJ et al. CRISPR-MVLST subtyping of *Salmonella enterica* subsp. *enterica* serovars Typhimurium and Heidelberg and application in identifying outbreak isolates. *BMC Microbiol* 2013c;**13**:254.
- Shariat N, Timme RE, Pettengill JB et al. Characterization and evolution of *Salmonella* CRISPR-Cas systems. *Microbiology* 2015;**161**:374–86.
- Sheludchenko MS, Huygens F, Stratton H et al. CRISPR Diversity in *E. coli* isolates from Australian animals, humans and environmental waters. *PLoS One* 2015;**10**:e0124090.
- Shen J, Lv L, Wang X et al. Comparative analysis of CRISPR-Cas systems in *Klebsiella* genomes. *J Basic Microbiol* 2017;**57**:325–36.
- Sinkunas T, Gasiunas G, Fremaux C et al. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* 2011;**30**:1335–42.
- Sinkunas T, Gasiunas G, Waghmare SP et al. *In vitro* reconstitution of cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J* 2013;**32**:385–94.
- Skenneron CT, Angly FE, Breitbart M et al. Phage encoded H-NS: a potential achilles heel in the bacterial defence system. *PLoS One* 2011;**6**:e20095.
- Smits TH, Rezzonico F, Lopez MM et al. Phylogenetic position and virulence apparatus of the pear flower necrosis pathogen *Erwinia piriflorinigrans* CFBP 5888T as assessed by comparative genomics. *Syst Appl Microbiol* 2013;**36**:449–56.
- Sorek R, Lawrence CM, Wiedenheft B. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem* 2013;**82**:237–66.
- Staals RH, Jackson SA, Biswas A et al. Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR-Cas system. *Nat Commun* 2016;**7**:12853.
- Stock SP, Blair HG. Entomopathogenic nematodes and their bacterial symbionts: the inside out of a mutualistic association. *Symbiosis* 2008;**46**:65–75.

- Swarts DC, Mosterd C, van Passel MW et al. CRISPR interference directs strand specific spacer acquisition. *PLoS One* 2012;**7**:e35888.
- Tan SY, Tan IK, Tan MF et al. Evolutionary study of *Yersinia* genomes deciphers emergence of human pathogenic species. *Sci Rep* 2016;**6**:36116.
- Timme RE, Pettengill JB, Allard MW et al. Phylogenetic diversity of the enteric pathogen *Salmonella enterica* subsp. *enterica* inferred from genome-wide reference-free SNP characters. *Genome Biol Evol* 2013;**5**:2109–23.
- Toro M, Cao G, Ju W et al. Association of Clustered regularly interspaced short palindromic repeat (CRISPR) elements with specific serotypes and virulence potential of Shiga toxin-producing *Escherichia coli*. *Appl Environ Microbiol* 2014;**80**:1411–20.
- Touchon M, Charpentier S, Clermont O et al. CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *J Bacteriol* 2011;**193**:2460–7.
- Touchon M, Rocha EP. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS One* 2010;**5**:e11126.
- Tymensen LD. CRISPR1 analysis of naturalized surface water and fecal *Escherichia coli* suggests common origin. *Microbiol Open* 2016;**5**:527–33.
- Veesenmeyer JL, Andersen AW, Lu X et al. NiLD CRISPR RNA contributes to *Xenorhabdus nematophila* colonization of symbiotic host nematodes. *Mol Microbiol* 2014;**93**:1026–42.
- Vercoe RB, Chang JT, Dy RL et al. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet* 2013;**9**:e1003454.
- Vergnaud G, Li Y, Gorge O et al. Analysis of the three *Yersinia pestis* CRISPR loci provides new tools for phylogenetic studies and possibly for the investigation of ancient DNA. *Adv Exp Med Biol* 2007;**603**:327–38.
- Very KJ, Kirchner MK, Shariat N et al. Prevalence and spatial distribution of *Salmonella* infections in the Pennsylvania raccoon (*Procyon lotor*). *Zoonoses Public Health* 2016;**63**:223–33.
- Wang J, Li J, Zhao H et al. Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. *Cell* 2015;**163**:840–53.
- Wang P, Zhang B, Duan G et al. Bioinformatics analyses of *Shigella* CRISPR structure and spacer classification. *World J Microbiol Biotechnol* 2016;**32**:38.
- Wehnes CA, Rehberger TG, Barrangou R et al. Determination of *Salmonella* clustered regularly interspaced short palindromic repeats (CRISPR) diversity on dairy farms in Wisconsin and Minnesota. *J Dairy Sci* 2014;**97**:6370–7.
- Westra ER, Pul U, Heidrich N et al. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol Microbiol* 2010;**77**:1380–93.
- Westra ER, van Erp PB, Kunne T et al. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol Cell* 2012;**46**:595–605.
- WHO. Plague around the world, 2010–2015. *Wkly Epidemiol Rec* 2016;**91**:89–93.
- Wiedenheft B, Lander GC, Zhou K et al. Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* 2011a;**477**:486–9.
- Wiedenheft B, van Duijn E, Bultema JB et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc Natl Acad Sci USA* 2011b;**108**:10092–7.
- Wiedenheft B, Zhou K, Jinek M et al. Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* 2009;**17**:904–12.
- Xie X, Hu Y, Xu Y et al. Genetic analysis of *Salmonella enterica* serovar Gallinarum biovar Pullorum based on characterization and evolution of CRISPR sequence. *Vet Microbiol* 2017;**203**:81–7.
- Yang C, Li P, Su W et al. Polymorphism of CRISPR shows separated natural groupings of *Shigella* subtypes and evidence of horizontal transfer of CRISPR. *RNA Biol* 2015;**12**:1109–20.
- Yang CD, Chen YH, Huang HY et al. CRP represses the CRISPR/Cas system in *Escherichia coli*: evidence that endogenous CRISPR spacers impede phage P1 replication. *Mol Microbiol* 2014;**92**:1072–91.
- Yin S, Jensen MA, Bai J et al. The evolutionary divergence of Shiga toxin-producing *Escherichia coli* is reflected in clustered regularly interspaced short palindromic repeat (CRISPR) spacer composition. *Appl Environ Microbiol* 2013;**79**:5710–20.
- Yoganand KN, Sivathanu R, Nimkar S et al. Asymmetric positioning of Cas1-2 complex and integration host factor induced DNA bending guide the unidirectional homing of protospacer in CRISPR-Cas type I-E system. *Nucleic Acids Res* 2017;**45**:367–81.
- Yosef I, Goren MG, Kiro R et al. High-temperature protein G is essential for activity of the *Escherichia coli* clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. *Proc Natl Acad Sci USA* 2011;**108**:20136–41.
- Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res* 2012;**40**:5569–76.
- Yosef I, Shitrit D, Goren MG et al. DNA motifs determining the efficiency of adaptation into the *Escherichia coli* CRISPR array. *Proc Natl Acad Sci USA* 2013;**110**:14396–401.
- Zhang F, Wen Y, Guo X. CRISPR/Cas9 for genome editing: progress, implications and challenges. *Hum Mol Genet* 2014;**23**:R40–6.