



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**DOCTORADO EN CIENCIAS BIOMÉDICAS**  
**CENTRO DE CIENCIAS GENÓMICAS**

**ANÁLISIS GENÉTICO Y MOLECULAR DE UN NUEVO SISTEMA DE  
REPLICACIÓN: SISTEMA *REPAC* PRESENTE EN EL PRL7 DE *RHIZOBIUM*  
*LEGUMINOSARUM* 3841**

**TESIS**  
**QUE PARA OPTAR EL GRADO DE:**  
**DOCTOR EN CIENCIAS**

**PRESENTA:**  
**GABRIELA PÉREZ SEGURA**

**DIRECTOR DE TESIS**  
**DR. MIGUEL ÁNGEL CEVALLOS GAOS**  
**CENTRO DE CIENCIAS GENÓMICAS**

**COMITÉ TUTOR**

**CUERNAVACA MOR. ENERO 2013**



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.

**Este trabajo fue realizado bajo la dirección del Dr. Miguel Ángel Cevallos Gaos en el Departamento de Ingeniería Genómica del Centro de Ciencias Genómicas de la Universidad Nacional Autónoma de México**

**Durante la realización de este trabajo de Doctorado se contó con una beca del Consejo Nacional de Ciencia y Tecnología (CONACyT) CVU/ Becario:233755/ 212978.**

**El trabajo que aquí presento contó con los apoyos económicos del CONACyT (00000000010009) y de la UNAM DGAPA/PAPIIT (IN205611).**

**Con amor a mis Padres  
y hermano  
que son el motor  
de mi vida**

# AGRADECIMIENTOS

A mi mentor y amigo el Dr. Miguel Ángel Cevallos Gaos (CCG-UNAM)

A mi comité tutorial:

Dr. Lorenzo Segovia Forcella (IBT-UNAM)

Dr. Enrique Merino Pérez (IBT-UNAM)

A los miembros del Jurado:

Dra. Gloria Soberón Chavéz (IIB-UNAM)

Dr. Miguel Ángel C. Cevallos Gaos (CCG-UNAM)

Dra. Elda Guadalupe Espín Ocampo (IBT-UNAM)

Dra. Rosa Laura Camarena Mejía (IIB-UNAM)

Dr. Víctor Manuel González Zúñiga (CCG-UNAM)

Al personal académico del laboratorio de Ingeniería Genómica del CCG-UNAM

Dr. José Guillermo Dávila Ramos

Dr. Víctor Manuel González Zúñiga

Dr. Cesar Rodríguez Sánchez

Dra. Magdalena Wienser Reyes

Q. I. Patricia Bustos Arcos

M. en C. Rosa Isela Santamaría Gutiérrez

M. en B. María de los Ángeles Pérez Oseguera

M. en C. María Soledad Juárez Ramírez



# ÍNDICE GENERAL

ÍNDICE DE FIGURAS.....	I
RESUMEN.....	II
ABSTRACT.....	III
<b>CAPITULO 1: INTRODUCCIÓN.....</b>	<b>1</b>
EL GENOMA BACTERIANO.....	1
EL GENOMA CENTRAL.....	2
La replicación del cromosoma bacteriano.....	4
La segregación del cromosoma.....	6
LOS CROMOSOMAS SECUNDARIOS.....	7
EL GENOMA FLEXIBLE.....	8
Las islas genómicas.....	8
Los profagos.....	10
Los elementos transponibles.....	11
Los plásmidos.....	12
La replicación de los plásmidos.....	15
La segregación de los plásmidos.....	19
La segregación de los plásmido de alto número de copias.....	20
La segregación de los plásmidos de bajo número de copias.....	21
Los sistemas de adicción.....	25
LA TRANSFERENCIA HORIZONTAL DE GENES.....	25
La transformación.....	26
La conjugación.....	28
La transducción.....	29
LOS REPLICONES <i>repABC</i> .....	30
El sistema de partición de los operones <i>repABC</i> .....	35
El módulo de replicación de los operones <i>repABC</i> .....	36
<b>CAPITULO 2: ANTECEDENTES.....</b>	<b>40</b>
<b>CAPITULO 3: JUSTIFICACIÓN Y PLANTEAMIENTO DEL PROBLEMA... </b>	<b>43</b>
<b>CAPITULO 4: .....</b>	<b>44</b>
OBJETIVO GENERAL.....	44
OBJETIVOS PARTICULARES.....	44
<b>CAPITULO 5: DISCUSIÓN Y PERSPECTIVAS.....</b>	<b>45</b>
BIBLIOGRAFÍA.....	50

➤ **ANEXO 1:**

**The *repAC* replication system of the *Rhizobium leguminosarum* pRL7 plasmid is functional: Implications regarding the origin and evolution of *repABC* plasmids. (2012)**

Gabriela Pérez-Segura, Ángeles Pérez-Oseguera, Miguel A. Cevallos.

Plasmid doi: 10.1016/j.plasmid.2012.08.003

➤ **ANEXO2:**

**The replication origin of *repABC* plasmid. (2011)**

Ramón Cervantes-Rivera, Francisco Pedraza-López, Gabriela Pérez-Segura y Miguel A. Cevallos. *BMC. Microbiology* **11**: 158



## ÍNDICE DE FIGURAS

<b>Figura 1:</b> Genoma bacteriano y sus divisiones.....	2
<b>Figura 2:</b> Replicación, condensación, y segregación del cromosoma bacteriano.....	6
<b>Tabla 1:</b> Propiedades conferidas por los plásmidos.....	14
<b>Figura 3:</b> Estructura del origen de replicación.....	16
<b>Figura 4:</b> Representación gráfica de un plásmido replicándose por un mecanismo theta .....	17
<b>Figura 5:</b> Representación gráfica de un plásmido replicándose por un mecanismo círculo rodante .....	19
<b>Figura 6:</b> Diversidad de la organización genética de las regiones de replicación/segregación de los replicones <i>repABC</i> .....	33
<b>Figura 7:</b> Representación gráfica de los cromosomas y los 6 plásmidos de Rlv3841.....	41
<b>Figura 8:</b> Representación esquemática de la localización y tamaño de los genes <i>repAC</i> dentro del plásmido pRL7.....	42
<b>Figura 9:</b> Grafica que nos muestra los niveles de expresión de los genes <i>repA</i> , <i>repB</i> y <i>repC</i> , presentes en cada uno de los plásmidos de la cepa de <i>R. etli</i> CFN42....	49

## RESUMEN

El genoma bacteriano consta de un cromosoma principal y algunas veces puede contener cromosomas secundarios y/o uno o más plásmidos. El cromosoma principal es la parte del genoma más conservada dentro de una especie, contiene la mayoría de los genes esenciales y es el replicón más grande dentro de la bacteria. El cromosoma secundario, contiene al menos un gen esencial, pero su forma de replicarse y segregarse hacia las células hijas es a través de mecanismos encontrados solo en plásmidos. Los plásmidos son moléculas que pueden ser dispensables para la bacteria, no codifican para funciones esenciales y se mantienen en la célula de una forma independiente a la del cromosoma ya que controlan su propia replicación y segregación. Dentro de las alfa proteobacterias existe un sistema de replicación y segregación que ha sido ampliamente estudiado: el sistema *repABC*, éste generalmente se encuentra en los cromosomas secundarios y en los plásmidos grandes de bajo número de copias y contiene todos los elementos que se requieren para el mantenimiento del replicón. Durante este trabajo se analizó la funcionalidad de una nueva configuración de los genes de este operón, presente en el plásmido pRL7 de *R. leguminosarum* 3841, donde sólo se encuentran los genes *repA* y *repC*. Este análisis demostró que este operón puede mantener, aunque de forma inestable, la replicación de un vector suicida dentro de *Rhizobium*. Los genes *repA* y *repC* de este plásmido conforman una sola unidad transcripcional. RepC es esencial para la replicación y el origen de replicación se encuentra dentro de su región codificante. Para el caso de RepA se observó que retiene parcialmente su capacidad de reprimir la expresión del operón, Estas observaciones nos sugieren contemplar al operón *repAC*, como un módulo que pudiera utilizarse en la generación y la evolución de nuevos sistemas de replicación/segregación, en la medida de que el operón adquiera genes por recombinación o transposición que le permitan una segregación estable.

## ABSTRACT

The bacterial genome comprises a main chromosome and sometimes secondary chromosomes and/or one or more plasmids. The main chromosome is the molecule with higher degree of sequence conservation within species. It contains most of the genes essential and is the largest replicon within the organism. The secondary chromosome contains at least one essential gene, but its mechanism of replication and segregation resemble to those found in plasmids. Plasmids are molecules that encode functions that are not essential for the bacteria. These molecules are maintained in the cell in an independent form, because they are capable to control its own replication and segregation

The *repABC* system within alpha-proteobacteria is a replication and segregation system that has been amply studied: it is generally inside the secondary chromosomes and large plasmids of low copy number, We have analyzed the functionality of a new configuration of this operon present in the plasmid pRL7 of *R. leguminosarum* 3841, where only *repA* and *repC* genes are present. This analysis showed that this operon can maintain, although unstable, the replication of a suicide vector in *Rhizobium*. The *repA* and *repC* genes of this plasmid form a single transcriptional unit. RepC is essential for replication and the replication origin is located within the coding region. In the case of RepA it was observed that it partially retains its ability to repress expression of the operon. These observations suggest *repAC* operon as a module that could be used in the generation and development of new replication/segregation systems, as long as it acquires the operon genes that allow stable segregation by recombination or transposition.



# CaPÍTULO 1

## INTRODUCCIÓN

### EL GENOMA BACTERIANO

El genoma bacteriano contiene toda la información biológica hereditaria necesaria para la construcción y el mantenimiento de cada microorganismo. Esta información se almacena en la célula en forma de repicones, los cuales son los elementos que tienen la capacidad de duplicarse y de segregarse a las células hijas. El genoma bacteriano está dividido en genoma central y en genoma flexible.

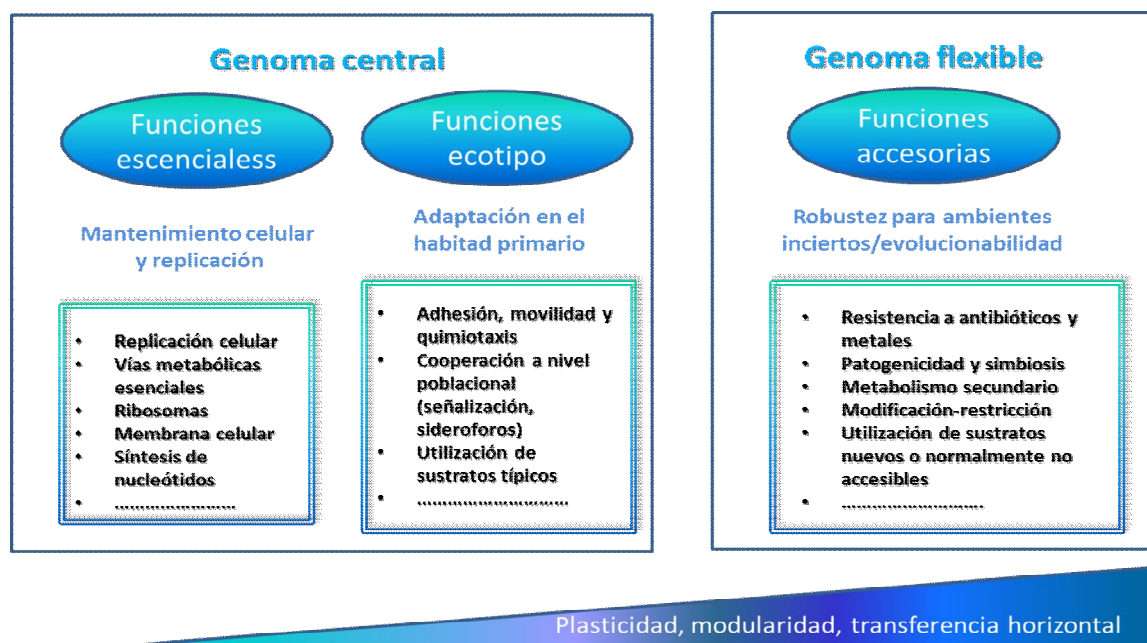
El genoma central se caracteriza por tener una conservación genética entre los miembros de una misma especie, este se hereda principalmente por transferencia vertical [Uchiyama, 2008]. A su vez el genoma central se divide en dos: la primera parte contiene genes involucrados en el mantenimiento celular y la replicación del mismo, también posee todos aquellos genes que participan en las vías metabólicas esenciales, la síntesis de los ribosomas, y en la síntesis, mantenimiento y funcionamiento de la membrana celular etc.

La segunda parte del genoma central codifica las bases del ecotipo, que son las funciones necesarias para la adaptación de la bacteria a su hábitat primario. Entre estas funciones podemos mencionar a la motilidad, la adhesión, la quimiotaxis, entre otras [Holger *et al.*, 2008]. El genoma flexible codifica para las funciones accesorias que proporcionan a la bacteria capacidades para poder habitar nichos diferentes al primario y por ende tiene un enorme impacto evolutivo. Entre las funciones accesorias podemos mencionar aquellas involucradas en la resistencia a los antibióticos y a los metales pesados, en la patogenicidad y en la simbiosis. Esta fracción del genoma es proclive a transferirse horizontalmente.

En la *figura 1* podemos ver un resumen de las funciones codificadas por cada una de las subdivisiones del genoma [Holger *et al*, 2008].

## EL GENOMA CENTRAL

El genoma central de un organismo presenta un contenido de G+C característico [Andrzej *et al*, 2012], reside básicamente en el cromosoma principal y posee un alto grado de sintenia.



**Figura 1: Genoma bacteriano y sus divisiones.** Se muestran algunas características codificadas, el grado de plasticidad, modularidad y transferencia horizontal que se puede observar en cada una de las subdivisiones [Holger *et al*, 2008].

Las modificaciones en el genoma central se deben básicamente a las mutaciones puntuales, aunque también puede sufrir inserciones, deleciones y otro tipo de rearrreglos. Esta parte del genoma bacteriano se perpetúa a través de herencia vertical [Rocha, 2008].

El número de cambios en la secuencia nucleotídica de los genes incrementa, conforme aumente la distancia del gen al origen de replicación del cromosoma, lo que nos indica que las regiones que rodean al origen de replicación son más estables y conservadas, comparadas con las que rodean al término de la replicación **[Anne-Brit Kolsto, 1997]**. La mayoría de los genes localizados en el cromosoma están organizados en operones. En la mayoría de los casos, los genes que codifican para enzimas metabólicamente relacionadas tienden a agruparse. El orden de los genes, en muchos casos, sigue el orden que corresponde a las actividades enzimáticas en las vías metabólicas, y frecuentemente están codificados en el mismo operón **[Rocha, 2008]**.

El tamaño del cromosoma puede variar dependiendo de la especie y del ambiente en que ésta habite. Entre los cromosomas bacterianos existe un amplio rango de tamaño: entre los más pequeños podemos mencionar el de *Candidatus Hodgkinia cicadicola* (0.143 Mpb) y el endosimbionte de insectos *Candidatus Carsonella rudii* (0.159Mpb) cuyos tamaños llegan a ser incluso menores que de algunos megaplásmidos presentes en otros organismos. Los genomas muy pequeños se han asociado a estilos particulares de vida donde los ambientes son muy estables **[Moya et al., 2009]**. Por otro lado los cromosomas más grandes que se han reportado son los de *Stigmatella aurantica*(10.260Mpb) y de *Streptosparangium roseum* (10.341Mpb) que son bacterias con estilos de vida más complejos y que habitan en ambientes mucho más diversos **[Scheneiker et al., 2007]**.

El cromosoma se encuentra en un alto grado de compactación que se logra a través de un complejo núcleo-proteico llamado nucleoide bacteriano. A modo de ejemplo, el cromosoma de *E. coli* se encuentra compactado cerca de unas 1000 veces, y ocupa solo un quinto del volumen celular **[Holmes y Cozarelli, 2000]**. El nucleoide es una entidad muy bien organizada (estructurado dentro de pequeños dominios superenrollados) y altamente dinámica que permite la transcripción, la replicación y la segregación, con una notable regulación espacial y temporal **[Rocha 2008]**.

La compactación del nucleóide está determinada por 3 eventos celulares, que son: la densidad de macromoléculas en el citoplasma bacteriano [Lerman, 1971]; el superenrollamiento negativo de ADN, el cual es esencial para la sobrevivencia de la bacteria y por la condensación debida a proteínas estructurales (condensinas) que se unen al ADN provocando un remodelamiento topológico del ADN, el cual varía en cada fase de crecimiento. Algunas proteínas se unen inespecíficamente al ADN, mientras que otras se unen a dominios específicos. Las mutaciones que afectan los genes que codifican estas proteínas provocan cromosomas poco condensados y que segregan anormalmente, incrementando considerablemente la frecuencia de células anucleadas y de cromosomas guillotizados específicos [Bouffartigues *et al.*, 2007].

### **La replicación del cromosoma bacteriano**

El cromosoma generalmente posee una sola copia por célula, aunque en algunos casos se ha encontrado poliploidia. La replicación del cromosoma bacteriano inicia en un sitio único del ADN llamado origen de replicación (*oriC*), en el caso del cromosoma de *E. coli* el *oriC* está rodeado por los genes: *dnaA*, que codifica para la proteína de inicio de replicación; *gidAB*, que codifican para proteínas de división; el gen *thdF*, que codifica para una proteína de partición y el gen *dnaN*, que codifica una subunidad de la ADN polimerasa [Sibley *et al.*, 2006].

El *oriC* se caracteriza, en la mayoría de los casos estudiados, por ser una región del cromosoma rica en nucleótidos A/T y con sitios de pegado para la proteína iniciadora de la replicación, la cual al unirse de una forma multimérica a estos sitios (en *E. coli* estos sitios se llaman cajas DnaA) desestabiliza la doble cadena del ADN, permitiendo así la entrada de la maquinaria proteica encargada de la replicación del cromosoma o replisoma.



En el cromosoma también existe una región de terminación de la replicación que se caracteriza por la presencia de sitios *ter* en localizaciones específicas y antipodas en el cromosoma al *oriC*.

Tanto el *oriC* como las secuencias de terminación guardan posiciones subcelulares definidas que varían dependiendo de la fase del ciclo celular [**Ghosh et al., 2006; Toro y Shapiro, 2010**].

Cada nueva ronda de replicación debe iniciarse justo antes de que la célula empiece a dividirse, es por eso que debe de estar finamente regulada y sincronizada con la división celular. Todas las especies bacterianas estudiadas hasta el momento poseen mecanismos de control del inicio de la replicación, enfatizando así la importancia de la regulación de este proceso.

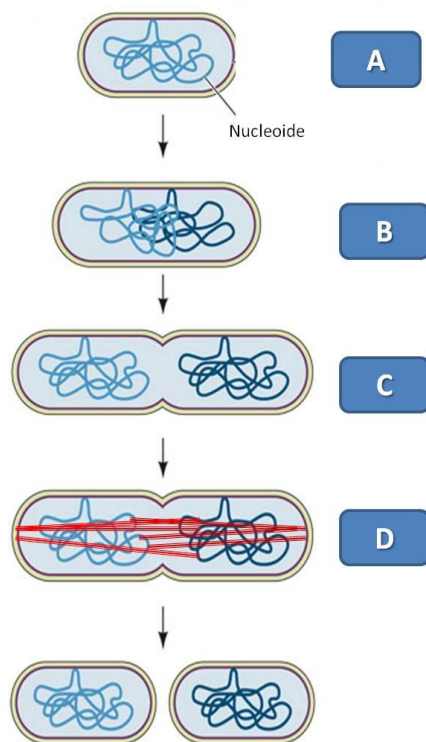
Para el caso del cromosoma de *E.coli*, tanto el *oriC* como la proteína DnaA son los blancos de regulación. La regulación a nivel de *oriC* es a través de un mecanismo de secuestro temporal de los orígenes recién replicados en la membrana citoplasmática mediado por una proteína denominada SeqA, la cuál es una proteína que reconoce específicamente las cadenas hemimetiladas propias de los orígenes de replicación recién duplicados. La regulación de DnaA ocurre a nivel de su actividad y de su disponibilidad.

Existen sitios de unión a DnaA a lo largo de todo el cromosoma además de los localizados en el origen, los cuales titulan la disponibilidad de la proteína para unirse al origen. La unión de DnaA a sus sitios de pegado (cajas DnaA) depende del tipo de nucleótido unido a la proteína, ya que solo se puede unir a estos sitios en presencia de ATP. Una vez que el replisoma se ha ensamblado en el origen, se estimula la actividad ATPasa de DnaA, la hidrólisis de ATP a ADP induce cambios conformacionales en la proteína, que provocan que DnaA pierda su capacidad de unirse al origen; esta capacidad se regenera cuando el ADP es sustituido por ATP [**Mott y Berger 2007**].

## La segregación del cromosoma

Para una transmisión exitosa de la información contenida en el cromosoma, se requiere de una coordinación precisa entre la replicación y la división celular. Algunas de las características principales para lograr este objetivo están resumidas en la *figura 2* y son las siguientes:

- La localización de uno o pocos sitios específicos dentro de la célula donde se lleve a cabo el inicio la replicación
- La fuerza de la replicación y el crecimiento de la bacteria promueven que cada una de las copias del cromosoma se vaya a cada célula hija
- La condensación de las regiones del cromosoma recién replicadas en posiciones definidas dentro de la célula de la célula
- Presencia de las maquinarias moleculares previamente localizadas que aseguran la separación de cada una de las copias del cromosoma a cada célula hija a través del *septum* [Ghosh *et al.*, 2006].



**Figura 2: Replicación, condensación y segregación del cromosoma bacteriano.**

- A)** puntos específicos donde se lleve a cabo la replicación
- B)** fuerza de la replicación y crecimiento de la bacteria provocando separación de las copias
- C)** condensación de las copias en posiciones específicas
- D)** maquinarias moleculares que aseguran la presencia de una copia del cromosoma en cada una de las células hijas

Los cromosomas emplean mecanismos de segregación activa o de partición que se conocen como sistemas Par. Estos sistemas son un mecanismo de transporte para el ADN y, en general, cuentan con tan solo 3 componentes: un sitio de partición que actúa en *cis* (*parS*), que tiene una función análoga a la de un centrómero eucariote y de 2 proteínas que actúan en trans (ParA y ParB). El sitio de partición marca el lugar del ADN al cual se va a unir específicamente una de las proteínas (ParB), la cual funciona como una proteína adaptadora entre el sitio de partición y la segunda proteína del sistema (ParA), la cual es una NTPasa que a través de la unión e hidrólisis de ATP, lleva a cabo la reacción de segregación [Vecchiarelli *et al.*, 2012]. Aunque es un sistema importante en la segregación del cromosoma, se ha demostrado que no es esencial, ya que mutaciones en los genes *parA* y *parB*, no son letales para la bacteria y producen solo una pequeña cantidad de células anucleadas [Bouffartigues *et al.*, 2007]. .

## **LOS CROMOSOMAS SECUNDARIOS**

El análisis de los genomas bacterianos ha permitido localizar a algunos genes esenciales en grandes replicones secundarios a los cuales se les ha dado el nombre de crómidos debido a que tienen un conjunto de propiedades diferentes a los que presentan los cromosomas primarios y pocos genes que son compartidos entre especies distantes. Estos replicones tienen una composición nucleotídica y de uso de codones muy parecida a la del cromosoma, contienen genes esenciales que se encuentran en el cromosoma de otras especies, pero su sistema de mantenimiento y de replicación es diferente. Estos sistemas se asemejan más a los utilizados por otros elementos del genoma bacteriano más pequeños llamados plásmidos y de los cuales hablaremos a detalle mas adelante. El tamaño de los crómidos generalmente es más pequeño que el cromosoma principal, pero más grandes que los plásmidos [Harrison *et al.*, 2010].

Se observa una conservación genética de sus genes y una sintenia compartida solo entre miembros del mismo género bacteriano, pero ninguno de sus genes son universales o únicos.

Los crómidos se encuentran en una porción minoritaria de las especies bacterias, pero aquellas que los tienen los mantienen por grandes periodos evolutivos. Existe una hipótesis que dice que tener crómidos puede ser ventajoso para una bacteria con un genoma grande ya que favorece una replicación más rápida.

## **EL GENOMA FLEXIBLE**

El genoma flexible presenta una alta plasticidad y modularidad de los elementos genéticos que contiene. Los genes que lo componen suelen intervenir en funciones como la resistencia a metales y a los antibióticos, la simbiosis, la patogenicidad, el metabolismo secundario, y la utilización de sustratos nuevos o pocos comunes entre otras muchas cosas. Esta fracción suele estar sujeta a un alto grado de ganancia y pérdida de genes, presenta una composición nucleotídica diferente a la del cromosoma, además de una pérdida de sintenia de los elementos genéticos en las especies cercanamente relacionadas.

El genoma flexible se hereda tanto por transferencia vertical y por herencia horizontal. Este genoma flexible le da una estructura genética con apariencia de mosaico al genoma bacteriano **[Mazur y Koper, 2012]**.

El genoma flexible está compuesto básicamente, pero no únicamente, por islas genómicas, profagos, plásmidos y transposones.

### **Las islas genómicas**

El análisis del genoma bacteriano indica que el cromosoma no es completamente homogéneo en el sentido de que existen secuencias, en regiones particulares, que no tienen el mismo origen filogenético.

Estas secuencias o elementos genéticos poseen un diferente contenido de G+C, se adquirieron por transferencia lateral y se les llama: islas genómicas. Estos elementos genéticos difieren entre cepas bacterianas cercanamente relacionadas, y en general, se les atribuye la capacidad de movilizarse o que la tuvieron en el pasado **[Schmidt & Hensel, 2004]**. Estas secuencias contribuyen a la diversificación y a la adaptación de los microorganismos, tienen un gran impacto en la plasticidad del genoma y le confieren a las bacterias propiedades que las hacen capaces de competir por nuevos nichos ecológicos.

Las islas genómicas son muy variadas en lo que respecta a su organización y su funcionalidad. Entre sus características principales podemos mencionar que:

- Son segmentos de ADN relativamente grandes en donde su tamaño varía de 10 a 200 kb
- Pueden reconocerse por estadísticas nucleotídicas ya que generalmente difieren del resto del cromosoma.
- En la mayoría de los casos se insertan en los genes que codifican para los tRNA y están flanqueados por secuencias directas repetidas, perfectas o casi perfectas de 16 a 20 pb, que son sitios de integración específica de las islas en sus sitios blancos o para su escisión enzimática
- Generalmente albergan genes que codifican para integrasas o factores relacionados con sistemas de conjugación de plásmidos y fagos (de los cuales hablaremos con detalle más adelante) involucrados en la transferencia de estas islas, también podemos encontrar muchos genes crípticos, de los cuales su función no se ha descrito.
- Algunas veces contienen elementos que proveen una ventaja selectiva a la bacteria hospedera **[Juhas *et al.*, 2009]**.

Debido a que las islas genómicas se encuentran integradas en el cromosoma bacteriano, se heredan principalmente por herencia vertical. Se replican y se segregan a las células hijas como parte del el cromosoma principal, aunque también hay una pequeña proporción de transferencia horizontal mediada por fagos.

## Los profagos

Dentro del genoma de la bacteria podemos encontrar regiones que tienen un origen claramente viral. Estas regiones pueden contener genes que pueden conferir resistencia a los antibióticos, genes que codifican factores de virulencia u otros que le permiten a la bacteria portadora una mejor adaptación al medio.

A estas regiones las llamaremos profago, pueden constituir del 10% hasta el 20% del genoma bacteriano, y son los que mayormente contribuyen a las diferencias entre individuos de la misma especie. Estas secuencias llegan a las diferentes bacterias a través de virus los cuales se denominan bacteriófagos [Casjens, 2003]. Los fagos son las formas de vida en la tierra más abundantes ( $\sim 10^{30}$  partículas) y los que más rápido se replican ( $10^{25}$  infecciones cada segundo) además de que presentan una enorme variabilidad. Los genomas de los bacteriófagos pueden ser de ADN de cadena sencilla o doble y también los podemos encontrar de ARN; su tamaño puede variar de poco a muchos cientos de kilobases. Dentro de sus genes esenciales podemos mencionar una replicasa específica, los genes que codifican para los componentes del fago que secuestran la maquinaria replicativa del huésped y los genes que codifican para el empaquetamiento de su genoma en una cápsula de proteínas (cápside).

Los bacteriófagos se clasifican en dos tipos: en virulentos y en temperados. Los primeros replican vigorosamente y lisan a la célula hospedera, mientras que los segundos tienen un estado alternativo inactivo, no lítico llamado lisogénico.

En muchos casos de lisogenia, el genoma del fago se mantiene de manera estable integrado al cromosoma del huésped. En otros casos el genoma del fago se mantiene replicando autónomamente independiente del cromosoma de una forma lineal o circular sin dañar a la célula. La conversión lisogénica es la adquisición de un nuevo fenotipo debido a la presencia de un profago en la célula hospedera. La estructura tipo mosaico de los fagos se debe a la recombinación con otros profagos u otros elementos móviles que se encuentran en la misma bacteria hospedera.

Un estímulo ambiental, como pudieran ser los agentes que dañan al ADN, puede provocar el cambio del estado inactivo a la replicación virulenta, e inducir así la lisis celular. Durante este proceso el ADN de las células hospederas puede empaquetarse accidentalmente en la cápside del fago con lo que podría posteriormente transferirse a una nueva célula hospedera; a este proceso que se le ha llamado transducción.

La habilidad de los bacteriofagos de transducir ADN de una célula hospedera a otra está limitada sólo a fagos grandes de ADN de doble cadena con un genoma de 50-100kb [**Frost, 2005**].

### **Los elementos transponibles**

Otros elementos importantes del genoma flexible son los elementos transponibles, los cuáles son secuencias de ADN con extremos definidos, que se mueven de un sitio a otro, en la misma o diferente molécula de ADN, a través de un evento genético llamado transposición [**Craig, 1996**]. Este mecanismo de transposición puede ser conservativo si la molécula o el sitio donde se encontraba el elemento transponible queda sin ninguna copia de este, por lo tanto el número de copias del transposón dentro de la célula no cambia y se conserva.

Muchos elementos transponibles se transponen de manera no conservativa, dejando una copia del elemento en la molécula de ADN donadora, provocando así que el número de copias del transposón aumente dentro de la misma célula.

Estos elementos también pueden ser mediadores de la translocación de segmentos de ADN que normalmente no son parte de estos elementos transponibles. Muchos elementos transponibles tienen una conformación simple ya que solo codifican para una proteína o complejo de proteínas llamado transposasa, el cual se encarga de llevar a cabo la transposición y algunos otros elementos individuales que pueden o no codificar para otras proteínas.

En los extremos de estos elementos transponibles generalmente se encuentran pequeñas secuencias repetidas invertidas (IRs) que son características para cada uno de estos elementos y en la mayoría de los casos son los sitios de unión para la transposasa. Estos sitios generalmente tienen un tamaño de 15-40 pb, pero pueden alcanzar hasta 150pb y tener muchos sitios de unión a la transposasa.

Si el elemento transponible sólo codifica las funciones necesarias para la transposición; es de un tamaño de 1-2 kb; tiene IRs cortas que delimitan los extremos de esta secuencia y poseen un marco de lectura abierto que codifica para la transposasa, se les llama secuencias de inserción (IS) **[Chandler y Mahillon, 2002]**. Pero si además de estas funciones, el elemento codifica para otras no involucradas en el proceso de transposición, como por ejemplo resistencia a algún antibiótico, el elemento se le conoce como transposón, El tamaño de los transposones es alrededor de 3 - 4 kb. Los elementos IS bacterianos son los elementos transponibles más simples, su capacidad genética está delimitada a 1 o 2 genes, se han identificado aproximadamente 20 familias de IS en eubacterias y arqueas.

Los elementos IS se encuentran tanto en los organismos procariotas como en los eucariotas, y son los más comunes de todos los elementos transponibles. Muchos de los elementos transponibles se integran en muchos sitios de los replicones bacterianos aunque se ha observado que existe un sesgo hacia ciertas regiones de estos replicones **[Bennett, 2004]**.

### **Los plásmidos**

Los plásmidos son moléculas de ADN de doble cadena, de tamaño muy variable, pero generalmente más pequeños que el cromosoma. En su mayoría los plásmidos son circulares y superenrollados, aunque también existen lineales como los que se han encontrado en bacterias de los géneros *Borrelia*, *Streptomyces*, *Thiobacillus*, *Nocardia*, *Rhodococcus* y algunas *Escherichias* por mencionar algunos ejemplos.



Los plásmidos tienen un número de copias característico dentro de la célula hospedera; los plásmidos pequeños suelen tener un alto número de copias (muchas veces más de 100 copias por célula), mientras que los más grandes están limitados a una o a pocas copias por célula. Los plásmidos no codifican para funciones esenciales y se mantienen en la célula de una forma independiente a la del cromosoma ya que controlan su propia replicación.

Algunos plásmidos son capaces de moverse horizontalmente dentro de su misma especie y entre especies diferentes, mecanismo que hace que los plásmidos tengan sus propias historias evolutivas, independientes de la del cromosoma bacteriano. Los plásmidos codifican para una gran variedad de funciones que promueven la sobrevivencia y extienden el espectro de ambientes en los cuales puede vivir la bacteria portadora de estos elementos genéticos. Entre los nuevos fenotipos que puede proporcionar un plásmido están enumerados en la *tabla 1*:

Los plásmidos se clasifican en grupos de incompatibilidad. Se dice que dos plásmidos pertenecen al mismo grupo de incompatibilidad cuando no pueden coexistir establemente en la misma línea celular. Este fenómeno se puede dar cuando dos plásmidos comparten los sistemas de replicación, los de segregación o de la regulación de alguno de estos procesos [Ebersbach et al., 2005].

Si en la célula coexisten dos plásmidos compatibles, cada uno de ellos puede mantener su número de copias característico, sin que haya ningún tipo de interferencia entre uno y otro.

En contraste, cuando los plásmidos son incompatibles, la combinación del número de copias de ambos plásmidos es aproximadamente igual a si cada plásmido estuviera presente solo. Puesto que existe una selección al azar para la replicación de cada copia; si uno de ellos excede el número de copias, el menos abundante tiende a diluirse más rápidamente y a perderse en la población. Una de las estrategias utilizadas por los plásmidos para evadir la incompatibilidad es la presencia de más de un sitio de replicación, que puede funcionar alternativamente, para evitar así ser desplazados.

Propiedades de resistencia	Propiedades metabólicas	Factores que modifican el estilo de vida de la célula hospedera
Resistencia a antibióticos aminoglucósidos, $\beta$ -lactámicos, cloranfenicol, tetraciclinas, Sulfamidas , trimetoprima	Producción de antibióticos y bactericinas	Producción de toxinas
Resistencia a metales pesados como: iones de mercurio, níquel, cobalto, bismuto, zinc, plata.	Metabolismo de carbohidratos simples y complejos	Producción de antígenos de colonización
Resistencia a aniones tóxicos como: arsenato, borato, cromato,	Metabolismo de proteínas y proteínas	Síntesis de hemolisinas
Resistencia a agentes intercalantes	Fijación de nitrógeno y utilización de citrato	Producción de suero de resistencia
Resistencia a daño por radiación	Síntesis de Dnasa extracelular	virulencia
Resistencia a bacteriófagos y bactericidas	Producción de sulfito de hidrogeno	Producción de capsula

**Tabla 1: Propiedades conferidas por los plásmidos a la población bacteriana**

Los plásmidos deben garantizar su transmisión vertical adecuada a las células hijas, para poder mantenerse establemente en la población; esto lo hacen a través de varios mecanismos **[Pinto et al., 2012]**:

1. Una replicación controlada, que permite asegurar que cada una de las células hijas contendrá el mismo número de copias del plásmido que la célula parental (Control del número de copias).
2. Sistemas de partición activa, los cuales físicamente distribuyen los plásmidos recién duplicados a las dos nuevas hijas (Segregación).

3. Sistemas de adición de plásmidos, en donde ocurre una muerte post-segregacional de las células hijas que no tienen ni una copia del plásmido. La bacterias que pierden el plásmido mueren debido a la presencia de moléculas con acción tóxica, en donde la presencia del plásmido les proporciona una protección ante la toxina (sistema toxina-antitoxina).
4. Recombinación sitio específica, la cuál convierte los multímeros de plásmidos en monómeros a través de un mecanismo de recombinación sitio específica (Sistemas de resolución de multímeros).
5. Transferencia conjugativa; mecanismo que les permite a algunos plásmidos recolonizar células curadas.

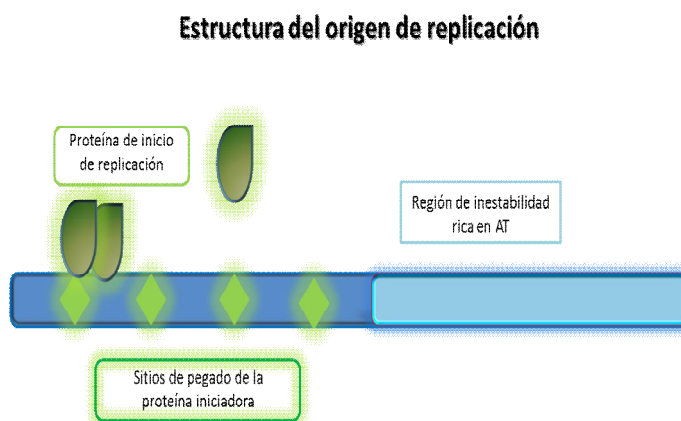
### La replicación de los plásmidos

Una de las principales características de los plásmidos es que controlan su propia replicación [del Solar y Espinoza, 2000]. Una vez que un plásmido entra a una nueva bacteria, debe replicarse hasta alcanzar el número de copias característico de cada plásmido, pero coordinando su replicación con el crecimiento y el ciclo celular de la bacteria hospedera.

Típicamente las funciones involucradas en la replicación y su control, se encuentran codificadas dentro en una región de 1-3 kb conocida como replicón básico, la cuál es la región mas pequeña del plásmido capaz de replicar con el mismo número de copias que el plásmido completo [Kollek *et al.*, 1978]; aquí generalmente podemos encontrar los siguientes elementos (*Figura 3*):

- **Origen de replicación** generalmente llamado (*oriV*). Es característico para cada replicón; es la región mínima en *cis* requerida para soportar una replicación autónoma del plásmido y que contiene los sitios de pegado para la proteína de inicio de replicación, además de una región rica en AT en donde las dos cadenas de ADN se separan para iniciar la síntesis de ADN.

- **El gene que codifica para la proteína iniciadora de la replicación.** La proteína iniciadora usualmente llamada Rep, se une a sitios específicos en el origen de replicación, provocando la apertura de las 2 cadenas y recluta el replisoma, el cual es el complejo enzimático encargado de sintetizar las nuevas copias de cada plásmido.
- **Elementos de control del número de copias.** Estos elementos regulan la tasa de replicación, para mantener dentro de la célula el número de copias característico de cada plásmido [Summers, 1998]. Pueden ser una proteína (plásmido  $\lambda$ dv) o pequeños ARNs antisentido, los cuales proveen un circuito auto-regulatorio, inhibiendo la replicación si el número de copias del plásmido aumenta demasiado (ejem. plásmido ColE1). En otros casos existen ARNs pequeños que pueden regular la traducción de la proteína iniciadora como es el caso del plásmido pTi81 y R1, Existen ejemplos en el que el número de copias del plásmido se regula por un conjunto de pequeñas secuencias de ADN repetidas localizadas en el origen de replicación, llamadas iterones, las cuales titulan la cantidad de proteína iniciadora disponible, además de que promueven que las moléculas recién replicadas permanezcan juntas y bloquean nuevas rondas de replicación, hasta que empiece un nuevo ciclo de división celular, regulando así la replicación. (ejem. Plásmido P1 y F) [del Solar y Espinosa, 2000].

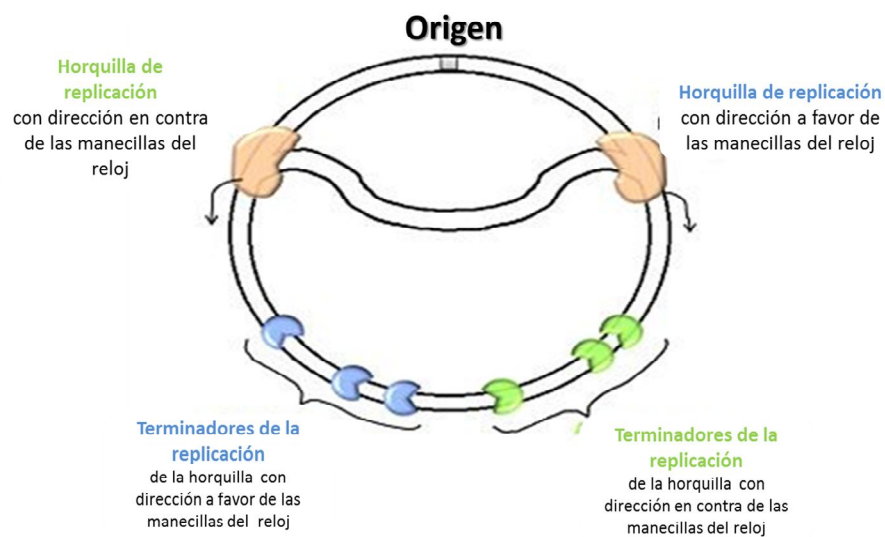


**Figura 3: Estructura del origen de replicación.**

Los rombos verdes representan los sitios de pegado a la proteína de inicio de replicación (gotas verdes), la cual es capaz de polimerizar en el origen, provocando la apertura de la región localizada río abajo, la cual es rica en nucleótidos A/T (rectángulo azul cielo), una vez formada la horquilla, el replisoma será capaz de establecerse, para comenzar la replicación.

Existen tres mecanismos generales para la replicación de los replicones circulares:

1.- **Tipo theta**, es el mecanismo más frecuente en los cromosomas de todas las bacterias y en los plásmidos de las bacterias Gram-negativas. Este mecanismo involucra una apertura las cadenas parentales en el origen de replicación mediado por la proteína de inicio de replicación; la síntesis de un cebador de ARN complementario a esta zona; y la síntesis de ADN a partir de una extensión covalente de este cebador; la cuál es continua en una cadena (cadena líder) y discontinua en la otra (cadena retrasada). La síntesis puede ser uni- o bidireccional (ver *figura 4*). En algunos casos, no se requiere de una proteína como molécula iniciadora un ejemplo de ello es el plásmido ColE1 en donde un RNA funciona directamente como cebador para el inicio de la replicación [Tomizawa, 1984]. El plásmido P1 de *E. coli* icodifica para su propia proteína iniciadora (RepA), pero el sistema no funciona adecuadamente sin la presencia de DnaA, la proteína iniciadora del cromosoma de la célula hospedera [Wickner y Chatteraj, 1987].

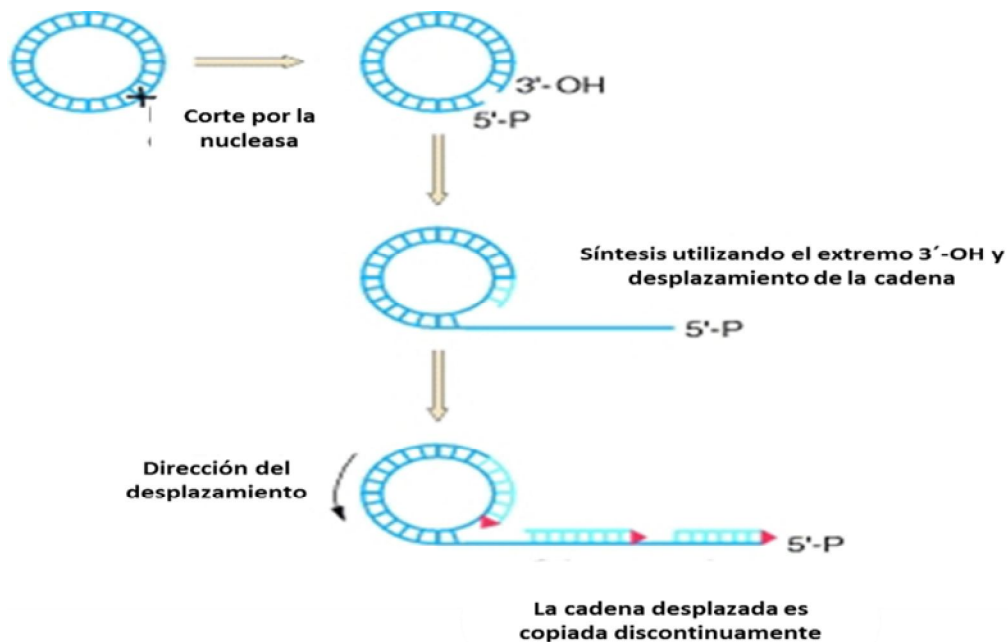


**Figura 4: Representación gráfica de un plásmido replicándose por un mecanismo theta.** El avance de las horquillas hace ver una letra  $\theta$ , de la que recibe el nombre el mecanismo. Se muestran el avance de los replisomas en direcciones opuestas. También se muestra la localización de las regiones terminadoras de la replicación.  
<http://www.proteopedia.org/wiki/index.php/Sandbox20>

2.- **Desplazamiento de cadena**, este mecanismo se asocia frecuentemente a plásmidos de amplio rango de huésped, los cuales requieren para el inicio de la replicación principalmente a tres proteínas codificadas en ellos mismos: una proteína iniciadora de la replicación, una helicasa replicativa y una primasa. La replicación ocurre en 2 orígenes adyacentes localizados asimétricamente, uno en cada cadena. Cuando la proteína de inicio se une a estas regiones, promueve la apertura de la doble cadena con la ayuda de la helicasa replicativa y, en ese momento, la primasa sintetiza el cebador necesario para la síntesis inicial de cada una de las cadenas, de forma continua resultando en el desplazamiento de la cadena complementaria.

La replicación de estos plásmidos es independiente de los factores de la célula hospedera que actúan en las primeras etapas de la replicación del cromosoma. Se ha propuesto que esta independencia es lo que le brinda su carácter de amplio rango de huésped.

3.- **Círculo rodante**, mecanismo mayoritariamente encontrado en plásmidos multicopia de organismos Gram-positivos. Es un proceso asimétrico unidireccional, en donde la síntesis de la cadena líder y de la retardada, no se encuentran acopladas. Una de las características que distingue a este mecanismo es que la cadena líder sintetizada permanece covalentemente unida a la cadena parental. El extremo 3' libre requerido para el inicio de la replicación se origina por un corte específico en el plásmido en una de las cadenas. Dicho corte lo lleva a cabo una proteína denominada Rep. La síntesis de la cadena líder usando este extremo produce un desplazamiento de la cadena parental y se termina la síntesis por varias reacciones específicas de transferencia realizadas por la proteína Rep. Una vez que se terminó la síntesis de la cadena líder, la proteína Rep se inactiva y se generan intermediarios de plásmido de cadena sencilla, para los cuales la RNA polimerasa sintetiza cebadores para la síntesis de la cadena retardada. Estos últimos pasos involucran solo proteínas de la célula hospedera (*ver figura 5*).



**Figura 5: Representación gráfica de un plásmido replicándose por un mecanismo de círculo rodante.** Representación gráfica de un replicón, duplicándose mediante un mecanismo círculo rodante, donde se muestra el corte de cadena sencilla, la síntesis de la cadena líder desplazando la otra y la duplicación de la cadena retardada usando cebadores de ARN. <http://geminiviridae.wordpress.com/replication-strategy/>

## La segregación de los plásmidos

Los plásmidos deben asegurar una transmisión estable a la progenie, y las estrategias a seguir para este propósito dependerán del número de copias del plásmido: si los plásmidos son de bajo número de copias usan una partición activa [Hayes y Barilla 2006]. En contraste, si son de alto número de copias estos se distribuyen pasivamente entre las células hijas [Summers, 1998]. En la mayoría de los plásmidos los genes encargados de heredar el material genético de una forma estable se encuentran agrupados en una pequeña región de la molécula y en algunos casos su regulación está coordinada [Bingle y Thomas 2001].

La probabilidad de que un plásmido no entre a una célula hija es de 0.5 y si  $n$  es el número de plásmidos de la célula en división, la probabilidad de que ninguno no entre a la célula hija es de  $(0.5)^n$ . Debido a que cada división produce 2 células hijas, la probabilidad (Frecuencia de segregación) de que una de ellas no reciba ningún plásmido es :  $P_0=2(0.5)^n=2^{(1-n)}$  . Esta frecuencia debe de ser lo suficientemente baja para que el plásmido parezca estable. Las cinéticas teóricas de pérdida de plásmido muestran que los plásmidos de alto número de copias tienen niveles altos de estabilidad aún si se segregan pasivamente, en contraste de aquellos que lo hacen con bajo número de copias, en los que es esencial que posean un sistema de segregación activo para alcanzar niveles de estabilidad altos [Sumers 1998].

#### ❖ La segregación de los plásmido de alto número de copias

Los plásmidos multicopia se distribuyen pasivamente a las células hijas, si el número de copias de un plásmido de la célula en división es lo suficientemente alto (mayor a 20) permite la difusión libre de las copias al momento de la división celular. La probabilidad de que una de las células hijas se quede sin plásmido es muy baja ( $<10^{-6}$ ), pero si el número de copias es menor la posibilidad aumenta [Summers, 1998].

Los plásmidos de alto número de copias como el plásmido ColE-1 de *E. coli*, pueden formar multímeros y esto es una causa importante de inestabilidad de los plásmidos ya que provoca un aumento en la frecuencia de células libres de plásmidos.

La proporción de oligómeros de un plásmido en un cultivo está determinada por:

- La tasa de recombinación intermolecular
- La fase de crecimiento en la que se encuentran las células que contienen los multímeros.



Existe una ventaja replicativa de los dímeros sobre los monómeros, el dímero replica 2 veces más rápido, debido a que contiene 2 orígenes de replicación lo que lleva a una rápida acumulación clonal de los dímeros en la célula. La presencia de dímeros reducen la estabilidad de los plásmidos, debido a que la célula “cuenta” el número de orígenes de replicación sin tomar en cuenta en que cada dímero posee dos orígenes de replicación lo que trae como consecuencia que el número de copias disminuya, y que aumente la proporción de células sin plásmidos [Summers 1998]. Erradicar estos multimeros es de vital importancia para la adecuada transmisión de los plásmidos multicopia y por ello tienen sitios de recombinación donde las proteínas hospederas actúan convirtiendo los dímeros en monómeros; existen también proteínas sensoras como es el caso de Rcd del plásmido ColE-1, las cuales, provocan la inhibición de la división celular si aún existen multimeros en la célula [Chant y Summers, 2007].

Para que estos plásmidos pueden coexistir establemente dentro de la célula hospedera, además cuentan con sistemas de control de número de copias, los cuales detectan y corrigen fluctuaciones en el número de copias característico de cada plásmido [Summers, 1998].

#### ❖ Segregación de los plásmidos de bajo número de copias

Para que los plásmidos de bajo número de copias sean heredados adecuadamente a las células hijas se requiere de un mecanismo de distribución activo, proceso en el cuál las moléculas recién duplicadas se separan activamente a los extremos de la célula en división a través de una maquinaria celular nucleoprotéica para asegurarse que al menos haya una copia en cada una de las células hijas. El *locus* que codifica para los elementos participantes en este proceso consta de solo 3 componentes:

- a) el sitio o sitios en el plásmido que funcionarán como secuencias centroméricas (*parC*, *parS* o *parH*), haciendo referencia a los centromeros de los cromosomas eucariótico, ya que marca el ADN blanco a ser segregado

- b) la primera de las proteínas codificadas en este loci es una NTPasa (ParA, ParM o TubZ) que utiliza la unión e hidrólisis de ATP o GTP para llevar a cabo la reacción de segregación
- c) el segundo gen de este *locus* codifica para la proteína de unión al sitio centromérico (ParB, ParG o ParR). Usualmente se une a esta región como dímero

Los sistemas de partición de los plásmidos se clasifican en tres grupos, dependiendo del tipo de NTPasa que tengan:

**Tipo I:** Son sistemas que contienen una ATPasa tipo Walker (debido al motivo de unión de ATP presente en la proteína) [Ogura y Hiraga 1983]. Entre las mejor caracterizadas de esta familia están las proteínas ParA de los plásmidos F y P1. A su vez este tipo de sistemas de segregación se subdivide en dos grupos: Los sub-sistemas de **Tipo Ia:** en el que las ATPasas contienen un dominio hélice-vuelta-hélice de unión a ADN en el extremo amino, a través del cual ejerce una autoregulación de la expresión del loci *par*. En la mayoría de los subsistemas analizados de este tipo, se observa que la localización del sitio *parS* es río debajo del operon de partición. En los sub-sistemas del **Tipo Ib** las ATPasas no presentan ningún dominio de unión a ADN, en este subsistema, la proteína de unión al centromero es más pequeña comparada con la presente en el sub-sistema Ia. En la mayoría de los casos estudiados el sitio centromérico colocaliza con la región promotora del operón, al unirse la proteína centromérica al sitio de partición regula la transcripción del operon, ya que reprime al promotor [Gerdes 2000]. Los loci *par* tipo Ia que codifican ATPasas con dominio de unión a ADN en el extremo amino, solo se han encontrado presentes en plásmidos. Como contraparte los loci que contienen las ATPasas pequeñas sin este dominio de unión a ADN se han encontrado tanto en plásmidos como en muchos cromosomas bacterianos [Livny *et al.*, 2007]

Para ambos subsistemas las proteínas de unión a la región centromérica (ParB) contactan su sitio *parS* correspondiente a través de un dominio de unión a ADN localizado en el extremo carboxilo de la proteína [Gerdes *et al.* 2010].

El sistema de segregación tipo ParA es el más ampliamente distribuido en el mundo microbiano [Vecchiarelli *et al.*, 2012]. Este sistema puede distribuir muchas copias de un plásmido [Szardenings *et al.*, 2011]. El mecanismo molecular de cómo actúa el sistema, no se tiene muy claro. Por lo que es un importante tema de investigación y debate.

Existen 2 modelos que tratan de explicar el mecanismo: el primero llamado “tracción por filamentos” (*spanning filament*) postula que ParA unida a ATP polimeriza en filamentos cortos, helicoidales e inestables; utilizando al nucleóide como andamio para su formación, cuando un extremo de los filamentos interactúa con ParB unida al plásmido, el filamento empieza a despolimerizarse arrastrando a los plásmidos a los polos del nucleóide [Ringgaard *et al.*, 2009]. Sin embargo, esta actividad no se ha demostrado directamente.

Las condiciones de polimerización de las diferentes ParAs varía dramáticamente dependiendo del sistema, y hasta la fecha no se han encontrado un conjunto común de parámetros entre los diferentes sistemas, que sean los necesarios para polimerización y por lo tanto no se tiene una clara imagen de cómo funciona este sistema *in vivo* en diferentes organismos [Ptacin *et al.*, 2010].

El segundo modelo llamado “difusión engranaje” (*diffusion ratchet*), postula que ParA unida a ATP se une inespecíficamente al nucleóide en configuración de dímeros o pequeños oligómeros, formando grupos discretos de ParA a lo largo del nucleóide. Cuando ParB unida al plásmido interactúa con las ParAs que se encuentran en estos grupos, promueve la actividad ATPasa de ParA, la cual sufre cambios conformacionales que provocan que se desprenda del nucleóide y libere al complejo ParB-plásmido. Este complejo es capturado por otras ParAs que se encuentran en el mismo grupo o grupos vecinos de ParA que se encuentran distribuidos a lo largo del nucleóide.

El movimiento de los complejos ParB-plásmido a través de los diferentes grupos de ParAs, junto con el movimiento del nucleóide durante la división celular, promueven la movilización del plásmido a los polos de la célula. **[Vecchiarelli et al., 2012]**.

**Tipo II:** Sistemas que presentan ATPasas tipo actina. El sistema mejor estudiado de este tipo es el que se encuentra presente en el plásmido R1, el sistema de segregación de este plásmido consta de tres elementos: *parC*, *parR*, *parM*. El gen *parM* codifica para una ATPasa tipo actina, que se ensambla en filamentos dinámicamente inestables **[Garner et al., 2004]**. El gen *parR* codifica a una proteína que se une de una forma cooperativa a secuencias repetidas localizadas dentro de *parC*, formando un complejo que se une a los filamentos de ParM y los estabiliza. La inestabilidad dinámica de los filamentos de ParM permite la segregación del plásmido mediante un mecanismo de “búsqueda y captura”, en donde los elementos se organizan en una estructura bipolar parecida a un huso de filamentos, en donde su elongación provoca la segregación de los plásmidos **[Garner et al., 2007]**.

**Tipo III:** Sistemas que presentan GTPasas tipo tubulina. Este sistema se identificó recientemente, estos módulos de segregación se encontraron en grandes plásmidos del complejo de *Bacillus cereus* y el ejemplo más estudiado es el plásmido pBtoxis de *Bacillus thuringiensis*. El locus de partición de este plásmido codifica para 2 proteínas: TubR (11.6kDa), la cuál es una proteína de unión a ADN y TubZ (54.4 kDa) la cuál es una GTPasa homóloga tanto a tubulina como a FtsZ **[Larsen et al., 2007]**. TubR se une como dímero a una secuencia *cis* en el ADN conformada por 4 iterones y que se localiza río arriba de los genes *tubR* y *tubZ*. Los monómeros de TubZ se unen al GTP y se ensamblan en protofilamentos. La hidrólisis del GTP provoca la translocación del plásmido debido al crecimiento direccional de los filamentos, provocando el movimiento del filamento dentro de la célula en forma de cinta (*treadmilling*), debido a la elongación de los filamentos por un lado y su contracción por el lado opuesto.

Los filamentos de TubZ secuestran el complejo formado por TubR y el plásmido dando como resultado la traslocación del plásmido a los polos de la célula [Laser *et al.*, 2007; Barilla, 2010].

### Los sistemas de adicción

Para mantener a los plásmidos también existen sistemas de estabilidad post-segregacional denominados sistemas toxina-antitoxina (TA), involucran a una proteína tóxica estable y una antitoxina inestable, la cual puede ser una proteína o un ARN antisentido. De manera que si la célula pierde al plásmido y con ello los genes que codifican para la toxina y la antitoxina, la biosíntesis de ambos se detiene; al ser la antitoxina de vida media más corta, no es capaz de neutralizar la toxina que tiene una vida media más larga, lo que provoca la muerte de la célula que perdió el plásmido. Existe una gran versatilidad en las funciones de los sistemas TA ya que pueden actuar como mediadores de la respuesta a estrés y de la muerte celular programada, en la generación de células tolerantes llamadas “*persisters*” (conocidas como variantes, en estado latente, que son altamente tolerantes a los antibiótico, sin poseer los genes para la resistencia a dichos antibióticos) [Lewis, 2010].

Los sistemas TA también los podemos encontrar en algunos cromosomas donde muestran un origen filogenético más complejo y su papel fisiológico no está bien definido [Hu *et al.*, 2010; Yuan *et al.*, 2011].

## LA TRANSFERENCIA HORIZONTAL DE GENES

Se ha observado que en el mundo bacteriano además de la transferencia vertical de material genético, existe también una transferencia horizontal de este material, la cual puede darse a través de 3 mecanismos principales:

1.-Transformación: en donde las células toman moléculas de ADN que se encuentran en el medio que las rodea y las integran a su genoma

2.-Conjugación: Involucra la transferencia directa de ADN a través de contactos célula-célula

3.-Transducción: en el cuál la transferencia ocurre a través de un bacteriófago

No todas las especies bacterianas presentan todos los mecanismos de transferencia genética. Algunas especies bacterianas son naturalmente transformables y en muchas otras la transformación solo se ha observado después de un pre-tratamiento artificial a las células. La conjugación se ha estudiado mejor en las  $\alpha$ -proteobacterias, pero esto no implica que la conjugación no exista o sea poco frecuente en otros grupos bacterianos.

Análisis genómicos comparativos de bacterias de vida libre, indican que la transferencia horizontal de genes es la principal fuente de diversidad genómica intra-especie [Lerat *et al.*, 2005].

## **LA TRANSFORMACIÓN**

La transformación bacteriana involucra la adquisición de ADN desnudo proveniente del medio extracelular. El número de especies bacterianas en el que ocurre la transformación natural es muy limitado, se conocen alrededor de 40 especies bacterianas naturalmente transformables, distribuidas en todos los grupos taxonómicos.

En muchas de estas especies, la competencia es un estado fisiológico transitorio, conocido como competencia que está finamente controlado por procesos específicos del organismo, entre los que podemos mencionar el *quorum sensing* y la percepción de señales nutricionales [Chen y Dubnau, 2004]. Se ha observado que la competencia generalmente ocurre al final de la fase logarítmica, justo antes de que las células entren a fase estacionaria.

El desarrollo de competencia en esta etapa de crecimiento se asocia a una falta de nutrientes y a una acumulación en el medio de factores de competencia, los cuales a través de un sistema regulatorio de dos componentes estimulan la expresión de los genes requeridos para la competencia. Puesto que los niveles de estos factores dependen de la densidad celular, solo se observa competencia cuando se alcanza una alta densidad poblacional. Por lo tanto, es posible que muchas más especies sean competentes a la transformación, sin embargo las condiciones en las que la desarrollan aún se desconocen **[Claverys y Martin, 2003]**.

Para que ocurra el fenómeno de transformación, luego de que las células estén en estado de competencia, deben, en primer lugar unir a un receptor de la superficie de la célula un fragmento de ADN. un fragmento de ADN se une a receptores de la superficie celular. Posteriormente, el transporte del ADN al interior de la célula obliga a que solo una de las cadenas de la molécula sea importada a su interior; la otra cadena se degrada y los productos de esta degradación se liberan al ambiente extracelular en el caso de bacterias Gram positivas) o al espacio periplásmico en las bacterias Gram-negativas).

Las proteínas involucradas en la importación de la molécula de ADN a la célula, son proteínas altamente conservadas, tanto en organismos Gram-negativos como en Gram-positivos. Los sistemas de competencia los conforman proteínas que presentan homología con proteínas involucradas en el ensamblaje del pili tipo IV (T4P) y a los sistemas de secreción tipo II (T2SSs). Forman una estructura que se expande parcialmente a lo largo de toda la periferia celular y están funcionalmente acoplados a un complejo de translocación de ADN localizado en la membrana citoplásmica **[Chen y Dubnau, 2004]**. En algunas especies el proceso es selectivo para moléculas de ADN provenientes de la misma especie, las cuales contienen secuencias cortas especie-específicas.

## LA CONJUGACIÓN

La conjugación de los plásmidos involucra contactos célula a célula, donde el ácido nucléico permanece dentro de una envoltura celular que lo protege de la degradación por las nucleasas extracelulares y los metales pesados. Los genes que se requieren para la transferencia conjugativa, son finamente regulados incluso cuando el plásmido llega a una nueva célula hospedera, y son transitoriamente activados para poder así colonizar nuevas células vecinas que aún no contengan el plásmido [Bingle y Thomas, 2001].

En las bacterias Gram-negativas los plásmidos conjugativos promueven la síntesis de un *pilus* extracelular el cuál reconoce a la célula receptora y se encarga del contacto célula a célula. Éste puede ser largo y flexible (expresados por el plásmido F) o corto y rígido (expresado por los plásmidos de los grupos de Incompatibilidad N, P y W) y algunos plásmidos codifican tanto para *pilis* largos como para cortos (grupos de incompatibilidad I1, I2, I5, B, K y Z). La transferencia se inicia cuando ocurre un corte de cadena sencilla en el origen de transferencia del plásmido (*OriT*) y la cadena cortada se transfiere a la célula hospedera.

La cadena que no se cortó lleva a cabo una ronda de replicación y así la célula donadora mantiene el plásmido intacto. Durante la transferencia la proteína que cortó el plásmido permanece unida al extremo 5' de la cadena procesada, la recirculación del plásmido transferido depende del *oriT*.

Existen plásmidos pequeños que no cuentan con el conjunto completos de genes necesarios para la conjugación, pero que son capaces de utilizar la maquinaria de los plásmidos conjugativos, para poderse transferir a las células vecinas, a este tipo de plásmidos se les llama movilizables, poseen un *oriT* activo y un conjunto pequeño de genes específicos del plásmido, llamados *mob* cuyos productos se encargan del corte del plásmido y del inicio de su transferencia.

En las bacterias Gram-positivas se ha encontrado una mayor diversidad de sistemas conjugativos: los estudiados se encuentran en los géneros *Bacillus*, *Clostridium*, *Streptomyces*, *Nocardia*, y *Staphylococcus*.



En este último género también se han encontrado pequeños plásmidos mobilizables con funciones similares a los encontrados en las Gram negativas.

En las primeras etapas de la movilización en estos sistemas, existe la formación del *pilus* conjugativo. En algunos de estos géneros las feromonas estimulan la conjugación. Se han reportado casos en donde la transferencia puede ocurrir a otras especies, y otros en los cuales se puede permitir la movilización del cromosoma o de plásmidos no conjugativos. Existen plásmidos promiscuos o también llamados de amplio rango de huésped, que pueden transferirse o movilizarse entre bacterias que pertenecen a especies muy diferentes, y que además pueden mantenerse establemente en ellas **[Thomas y Summers, 2008]**. En algunos casos se ha observado que son capaces de movilizar ADN de bacterias Gram negativas a Gram positivas y en algunos casos también a levaduras, demostrando así que la transferencia horizontal de genes entre géneros y reinos puede ocurrir **[Bingle y Thomas 2001]**.

Para asegurar la integridad física de los plásmidos al ser transferidos a las especies cercanas, existen estrategias de protección a los sistemas de restricción-modificación de la célula hospedera. Algunos de estos plásmidos han modificado los sitios de reconocimiento para evadir las enzimas de restricción de las células hospederas, desarrollado mecanismos que pueden inactivar los sistemas de defensa de la célula que los recibe o inhibir la inducción de la respuesta a estrés (respuesta SOS).

## **LA TRANSDUCCIÓN**

La transducción es un mecanismo de transferencia de material genético de una bacteria donadora a una receptora, a través de la participación de bacteriófagos. La transducción se divide en 2 categorías, la especializada (restringida) y la generalizada.

La principal diferencia entre ellas radica en el tipo de molécula de ADN que se empaqueta en la cápside del fago. En la transducción generalizada se pueden empaquetar fragmentos de ADN de la bacteria donadora de aproximadamente el mismo tamaño del genoma del fago. En comparación, una partícula de transducción especializada surge a partir de la escisión incorrecta de un profago del genoma de la célula hospedera, empaquetando ADN hospedero junto con su propio genoma. El fago se adhiere a la célula receptora e inyecta el material genético que se encuentra dentro de la partícula transductante [Miller, 2001].

El ADN importado puede tener varios destinos: si la bacteria receptora y donadora son lo suficientemente relacionadas filogenéticamente, el ADN transducido puede incorporarse al genoma del huésped a través de algún sistema de recombinación y permanecer ahí establemente. Si el material genético que introduce el fago no cumple con este requerimiento acaba por degradarse.

En el caso de la transducción especializada el DNA de origen no viral puede mantenerse junto con el genoma del fago siempre y cuando este último posea un sistema de replicación independiente y funcional en la célula hospedera. [Fineran *et al.*, 2009].

## LOS REPLICONES *repABC*

Dentro de las *alfaproteobacterias* existe un sistema de replicación y segregación ampliamente distribuido: el sistema *repABC*. Éste generalmente se encuentra en los cromosomas secundarios y en los plásmidos grandes de bajo número de copias, pertenecientes a miembros de los ordenes Rhizobiales, Rhodospirilales y Rhodobacterales [Cevallos *et al.*, 2008; Petersen *et al.*, 2009].

El sistema *repABC* contiene todos los elementos que se requieren para la replicación y un mantenimiento estable de los replicones que lo portan, incluyendo elementos encargados de la regulación de estos procesos y de la incompatibilidad del replicón.

El sistema de replicación/ partición *repABC* consiste en una sola unidad transcripcional **[Cevallos et al., 2008; Petersen et al., 2009]**: el operon *repABC*. Contrario a lo que se encuentra en la mayoría de replicones en donde los genes encargados de la replicación y la segregación, se encuentran en *loci* diferentes, donde se expresan y se regulan independientemente.

Los operones *repABC* contienen 3 genes que codifican para proteínas y que siempre se encuentran en el mismo orden relativo (*repA*, *repB*, *repC*); un pequeño gen regulatorio que codifica para un ARN antisentido (ctRNA, countertranscript RNA), el cual se transcribe en dirección opuesta al operón, y al menos una secuencia centromérica llamada *parS* **[Cervantes-Rivera et al., 2010; Ramirez-Romero et al. 2001; Venkova-Canova et al., 2004; Pappas and Winans, 2003]**.

Se ha encontrado que los genes *repA*, *repB*, *repC* de cada replicon se expresan a partir de uno o varios promotores: el plásmido p42d de *R. etli*, pSymA de *S. meliloti* o pTAVi tienen operones *repABC* que se transcriben a partir de un solo promotor, mientras que el plásmido pTiR10 de *A. tumefaciens* se transcribe utilizando hasta 4 promotores **[Pappas y Winans 2003a]**. Los productos de los genes *repA* y *repB* presentan similitud a proteínas involucradas en la segregación activa de algunos plásmidos y cromosomas, pertenecientes a la superfamilia de proteínas de partición MinD/ParA. Sin embargo, pertenecen a un clado diferente de ParA y ParB **[Cevallos et al., 2008; Pinto et al., 2012]**.

El origen de replicación se localiza dentro el gen *repC*, el cual además, codifica para la proteína iniciadora **[Cervantes-Rivera et al., 2011; Pinto et al., 2011]**. El plásmido pTiR10 codifica un cuarto gen, *repD*, el cual mide 228 nucleótidos y se localiza entre los genes *repA* y *repB*, el plásmido pTiC58, presenta un marco de lectura abierta en la misma posición pero completamente divergente en su secuencia **[Chai y Winans, 2005b]**. El *loci* de partición (*parS*) de estos replicones lo podemos encontrar río abajo del gen *repC* (p42d/pTAV320), río arriba del gen *repA* (pSymA) o dentro de genes (dentro de *repD* del pTir10).

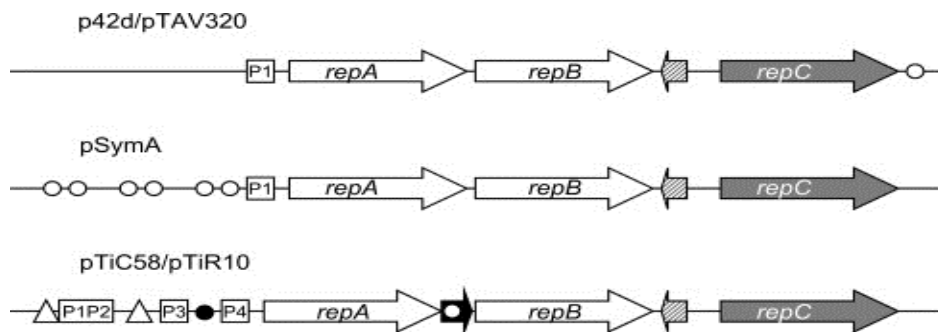
El ctRNA se localiza en la región intergénica entre los genes *repB* y *repC*. Se ha demostrado que juega un papel en el control de la replicación del plásmido, ya que regula postranscripcionalmente a RepC. Todos estos elementos los podemos ver esquematizados en la figura 6. Los plásmidos *repABC* contienen al menos 4 elementos involucrados en la incompatibilidad como son las proteínas RepAB, RepC, el ctRNA y la secuencia *parS*. [Venkova-Canova *et al.*, 2004; Chai y Winnans 2005<sup>a</sup>; MacLellan *et al.*, 2005 ; Cervantes-Rivera *et al.*, 2011].

El análisis filogenético de los 56 genomas completos Rhizobiales que se encontraban anotados en el GeneBank hasta el 2011, mostró que los operones *repABC* se encuentran presentes en 14 cromosomas secundarios y 93 plásmidos, tienen un tamaño que varía de las 4.1 kb hasta las 3.7 Mb.

Los cromosomas primarios tienen al menos una copia del gen *dnaA* que codifica para la proteína de inicio de replicación, y un locus *parAB* que es un sistema de segregación activa y no se encuentra ligado a *dnaA*. Sin embargo, se ha observado que el gen *dnaA* y el loci *parAB* no se encuentran presentes en ningún cromosoma secundario, ni en ninguno de los plásmidos *repABC*. El gen *repC*, se encuentra presente en todos cromosomas secundarios y en dos terceras partes de los plásmidos, pero no se encuentra presente en ninguno de los cromosomas principales [Pinto *et al.*, 2012].

Los árboles filogenéticos de los operones *repABC* muestran cladas intercaladas de los operones encontrados en cromosomas secundarios y en plásmidos, haciendo imposible distinguir los unos de los otros por simple análisis de secuencia. Aparentemente hay un sesgo en el que el sistema *repABC* que hace que sólo se encuentre presente en los cromosomas secundarios y en los plasmidos, y que el gen *dnaA* y el loci *parAB* solo en los cromosomas principales.

Dentro de un mismo replicon se pueden encontrar más de un operón *repABC*, se ha encontrado que al menos 68 plásmidos tienen al menos un locus completo *repABC*, 5 tienen 2 operones completos y 6 plásmidos tienen un locus completo y un *cassette* adicional incompleto [Slater *et al.*, 2009; Pinto *et al.*, (2012)].



**Figura 6. Diversidad de la organización genética de las regiones de replicación / partición de los replicones *repABC*.** Las flechas blancas representan los genes *repA* y *repB*, son las proteínas implicadas en la partición, y en la regulación negativa de la transcripción del operón. Los círculos blancos indican la posición de los sitios *parS*. Los círculos negros indican cajas *vir* involucradas en la regulación. Flechas sombreadas indican la posición del gen de ARN antisentido. Flechas grises indican las posiciones de los genes *repC*, que codifican a las proteínas limitantes para la replicación. La flecha negra representa el gen *repD*. Las cajas cuadradas indican las posiciones de promotor, y triángulos, la posición de las cajas *tra*, involucradas en la regulación del operón [Cevallos et al., 2008].

Esta familia de plásmidos incluye muchos grupos de incompatibilidad, ya que más de un tipo de plásmido *repABC* pueden coexistir en la misma bacteria. Los casos más evidentes con esta peculiaridad son *R. etli* CFN42 y *R. leguminosarum* 3841 cada una contiene 6 grandes plásmidos los cuales todos pertenecen a la familia *repABC*. Dos replicones de *R. etli* CFN42 y uno de *R. leguminosarum* contienen dos operones *repABC* [Young et al., 2006; Gonzalez et al., 2006].

En algunas especies los plásmidos *repABC* constituyen un cantidad significativa del genoma de la bacteria, como es el caso de *R. leguminosarum* 3841, en la cual los plásmidos *repABC* constituyen el 35% de su genoma y el de *R. etli* CFN42 en la que estos plásmidos constituyen aproximadamente el 32% de su genoma [Young et al., 2006; Gonzalez et al., 2006].

Análisis recientes muestran que las proteínas codificadas en este operón poseen distinto origen filogenético y dado que la estructura del operón está completamente conservada, es una evidencia fuerte de que hubo desplazamiento de genes *in situ*. [Slater *et al.*, 2009; Castillo-Ramírez *et al.*, 2009].

Los genomas de las *alfaproteobacterias* pueden contener genes *repAB* no asociados a un ctRNA o con el gen *repC*, de la misma manera plásmidos en los que su replicación depende de un módulo *ctRNA-repC*, sin la participación de genes *repAB* [Izquierdo *et al.*, 2005]. Algunas *alfaproteobacterias* contienen genes *repC* que se encuentran asociados con otros genes que no están involucrados en las funciones de replicación o segregación [Slater *et al.*, 2009; Castillo-ramírez *et al.*, 2009].

Algunos replicones *repABC* responden a una regulación de percepción del *quórum*, como es el caso del plásmido pTiR10 en donde la actividad de sus promotores P1-P3 tiene la influencia positiva de la proteína efectora TraR unida a una octanoil homoserin lactona. Esta regulación ocurre cuando este complejo se une a las cajas *tra* localizadas entre los promotores P1 y P3 ( *figura 6*), lo que trae como consecuencia el incremento del número de copias del plásmido [Pappas y Winans 2003<sup>a</sup>; Pappas y Winanas 2003b].

Otro elemento importante en la regulación del número de copias de este plásmido es la proteína VirG, la cual pertenece a un sistema de dos componentes. Esta proteína modula la actividad del operón tanto positiva como negativamente uniéndose a las cajas *vir* localizadas entre los promotores P3 y P4 del operón [Cho y Winans, 2005]. Se ha observado que proteínas RepA que se encuentran en la misma bacteria no presentan más de un 61% de identidad y el 51% para el caso de las proteínas RepB. Esta observación podría interpretarse argumentando que con una identidad mayor se podría inducir el fenómeno de incompatibilidad de los sistemas. [Castillo-Ramirez *et al.*, 2009; Pinto *et al* 2012].

## **EL SISTEMA DE PARTICIÓN DE LOS OPERONES *repABC***

Los replicones *repABC* debido a su bajo número de copias requieren de un mecanismo activo eficiente de segregación. Las proteínas RepA, RepB y el sitio *parS* (análogo al centrómero) proveen este mecanismo cuyo funcionamiento se ha estudiado en los plásmidos de *Agrobacterium tumefaciens* pTIB6S3, *Paracoccus Versutus* pTAV320, *Rhizobium etli* p42d, *Sinorhizobium meliloti* 1021 pSymA. Se ha demostrado que las remociones o las mutaciones en los genes *repA* y/o *repB* disminuyen considerablemente la estabilidad del plásmido [Tabata *et al.*, 1989; Bartosik *et al.*, 1998; Ramírez-Romero *et al.*, 2001; Maclellan *et al.*, 2006; Gallie *et al.*, 1985].

RepA y RepB son homólogos de miembros de sistemas de segregación bien caracterizados. RepA del plásmido Ti tiene una similitud del 47% con la proteína SopA del plásmido F y 43% con la proteína ParA del bacteriófago P1; RepB es similar a SopB del plásmido F y con ParB del plásmido P1. Existen similitudes en el tamaño de las proteínas que lo componen y en la secuencia del dominio de ATPasa con el de dominio de RepA, estos datos nos sugieren que el sistema de segregación de los replicones *repABC* pertenece a los sistemas de segregación tipo 1a. Las proteínas RepA de las *alfaproteobacterias* llevan a cabo la autoregulación del operón, se unen a la región operadora que se encuentra localizada cerca del inicio de la transcripción del gen *repA*. Esto se ha demostrado en los plásmidos p42d de *R. etli* y el pTiR10 de *A. tumefaciens* R10 [Ramírez-Romero *et al.*, 2001; Pappas y Winans 2003].

Otro elemento importante de los sistemas de segregación es el loci *parS*. Para el caso de los replicones *repABC*, se han identificado los loci *parS* de los plásmidos pTAV320, p42d, pSymA, pTiC58 y pTiR10 el cual consiste en una o más copias de una secuencia palindrómica consenso de 16pb (GTTNNCNGCNGNNAAC) y cumple con los siguientes 3 requisitos [Bartosik *et al.*, 1998; Soberón *et al.*, 2004; Maclellan *et al.*, 2005; Chai y Winnans, 2005b].

1. Es esencial para la estabilidad del plásmido
2. Es el sitio de unión de RepB
3. Es un fuerte factor de incompatibilidad con respecto al plásmido parental del que proviene cuando es puesto a competir en *trans*.

Se han hecho mutaciones puntuales en el sitio *parS* de pSymA y se observó que se elimina la incompatibilidad y que se disminuye la afinidad por RepB reduciendo así la estabilidad del plásmido [Macllenan *et al.*, 2005].

El número de sitios *parS* presentes puede variar ampliamente, algunos pueden tener solo un sitio *parS* como es el caso del plásmido p42d; o como los plásmidos pTAV320 y pTiR10 que contienen dos de estos elementos, mientras que el pSymA puede contener hasta seis elementos *parS*. Sin embargo, tan solo un sitio *parS* basta para la segregación adecuada del plásmido [Bartosik *et al.*, 2001; Chai y Winans 2005b; MacLellan *et al.*,2005; Sóberon *et al.*,2004]. En muchos operones *repABC* no se encontraron secuencias *parS* consenso, lo que sugiere que estos sitios son muy divergentes y se escapan a la detección bioinformática [Cevallos *et al.*, 2008].

## **EL MÓDULO DE REPLICACIÓN DE LOS OPERONES *repABC***

El último gen del operón es *repC*, el cual es esencial para la replicación del plásmido. El marco abierto de lectura de este gene expresado a partir de un promotor exógeno constitutivo es la región mínima necesaria para mantener la replicación de un vector suicida y es capaz de replicar con alto número de copias, en una cepa curada del plásmido parental [Bartosik *et al.*, 1998; Ramírez-Romero *et al.*, 2000; Cervantes-Rivera *et al.*, 2011; Pinto *et al.*, 2011].

La proteína RepC es la única del operón que no tiene un homólogo con una función asociada en las bases de datos. Las proteínas tipo RepC sólo se encuentran en las *alfaproteobacterias*, y al igual que otras secuencias en el



operón, el gen *repC* puede encontrarse en diferentes contextos genómicos. Sin embargo, en la mayoría de los casos se encuentra asociado al gen *ctRNA* y río debajo de los genes *repA* y *repB* en un mismo operón [Cevallos *et al.*, 2002]. RepC es la proteína iniciadora de la replicación y su función es unirse al origen de replicación (*oriV*) del replicón, el cual se localiza dentro de la secuencia codificante de esta proteína.

Dentro de esta secuencia se encuentra una región amplia rica en A+T de aproximadamente 150 nucleótidos, común en los orígenes de replicación, y que facilita la apertura de la doble cadena de ADN para formar así la burbuja de replicación [Cervantes-Rivera *et al.*, 2011]. Se ha demostrado que esta unión es cooperativa, RepC multimeriza en su sitio de unión a ADN y debido a su similitud estructural con los factores transcripcionales de la familia MarR, se ha sugerido que RepC se une al ADN como dímero [Pinto *et al.*, 2011].

Las predicciones de su estructura secundaria y de la conservación de aminoácidos predicen que RepC básicamente presenta 2 dominios, el dominio N-terminal (NTD) de los residuos 1-265 y el C-terminal (CTD) de los residuos 295-439, ambos dominios unidos por 30 aminoácidos hidrofóbicos, no estructurados y poco conservados; la región suficiente para unirse al ADN reside en los residuos 26-158 [Pinto *et al.*, 2011]. El papel del CTD aún no está claro, pero se sabe que los últimos 39 aminoácidos de RepC del p42d juega un papel importante en la incompatibilidad del plásmidos [Cervantes-Rivera *et al.*, 2011].

No se ha podido identificar que RepC actúe en *trans*, se han analizado su papel en los plásmidos PTiR10 [Pinto *et al.*, 2011] y p42d [Cervantes-Rivera *et al.*, 2011], y cuando se sobreexpresa causa un incremento en el número de copias de plásmido que la alberga, pero no aumenta el número de copias de un segundo plásmido dependiente de *repC* presente en la misma célula, ni es capaz de sostener la replicación de un plásmido que solo lleve el origen de replicación, confirmando que su actividad solo es en *cis*.

Esta preferencia en *cis* puede deberse a una baja expresión de la proteína, o a que presente una inestabilidad proteolítica, poca difusión dentro de la célula, secuestro de la proteína por interacciones con ADN no específico, o por la membrana celular [Pinto *et al.*, 2011]. Esta inhabilidad de RepC de actuar en *trans* puede ser deberse a que el péptido nascente de RepC sea atrapado por el origen de replicación *oriV* durante la traducción o también puede ser secuestrado inmediatamente después de que sea liberado por el ribosoma.

La región intergénica entre los genes *repB-repC* llamada *inca* de aproximadamente 150 pb es un fuerte determinante de incompatibilidad, esta región es importante ya que dentro de ella se encuentran 2 elementos importantes para el correcto funcionamiento de la replicación:

1. Aquí se encuentra codificado en la cadena complementaria al operón, un pequeño ARN antisentido de tan solo 59 nt (ctRNA) y el cuál regula los niveles de RepC y es importante para llevar a cabo la adecuada replicación del plásmido [Venkova-Canova *et al.*, 2004]. El ctRNA se estructura en un tallo de 14nt de largo, y en la punta forma una asa de 6 nt, del lado derecho contiene un extremo no estructurado de 13 nt y del lado izquierdo una región no estructurada rica en uracilos de 9 nt de largo, presenta todas las características esenciales para un terminador intrínseco de la transcripción [Cervantes-Rivera *et al.*, 2010].
2. En esta región el ARN mensajero forma una estructura tallo-asa grande que cuando no se encuentra el ctRNA, el sitio de unión a ribosoma y el codón de inicio del gen *repC* permanecen en cadena sencilla, permitiendo la traducción de *repC*, pero si se encuentra presente el ctRNA, este interacciona con el mARN del operon en su región complementaria, provoca que la secuencia líder del gen *repC* forme una estructura tallo asa en el mARN, la cual funciona como un terminador intrínseco de la transcripción, impidiendo así la traducción del *repC*.

Por lo tanto el ctRNA juega un papel importante en la regulación de la síntesis de RepC a nivel transcripcional. La vida media de este ctRNA es corta: de 5 minutos aproximadamente y presenta una síntesis constitutiva, lo que es común para los ARN's antisentido con una función semejante, lo que les permite controlar fácilmente las oscilaciones en la síntesis de la proteína iniciadora y mantener así los niveles adecuados de RepC para que el plásmido se mantenga en la célula sin problemas [**Cervantes-Rivera et al., 2010**].

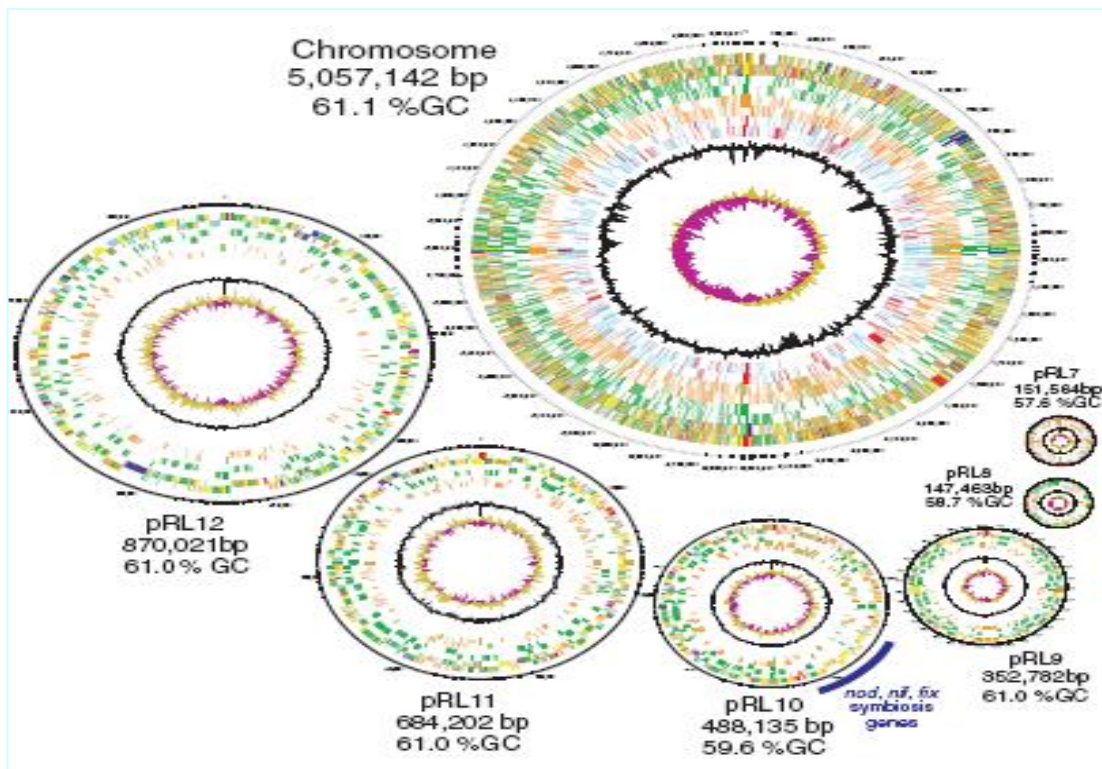
# CAPITULO 2

## ANTECEDENTES

Nuestro modelo de estudio para este trabajo fue *Rhizobium leguminosarum biovar viciae* cepa 3841 (Rlv 3841), una *alfaproteobacteria* simbiote fijadora de nitrógeno de algunas leguminosas como el chícharo y el frijol. Esta cepa presenta una arquitectura genómica multipartita. El tamaño de su genoma es de 7.75 Mpb y se encuentra repartido en un cromosoma y seis plásmidos grandes de bajo número de copias (pRL7-pRL12). El cromosoma mide 5.055 Mpb y tiene un 61.1% de contenido de GC mientras que el plásmido más pequeño (pRL8) mide 147 Kb y posee un contenido de GC de tan solo el 68.7%, más bajo comparado con el del cromosoma. Por su parte, el plásmido más grande (pRL12) presenta un tamaño de 870 Kb y su porcentaje de contenido de GC es muy parecido al del cromosoma, ya que cuenta con un 61%. En la *figura 7* podemos observar el tamaño, el contenido de GC del cromosoma y los 6 plásmidos de esta cepa [Young *et al.* 2006].

El plásmido pRL10 es el plásmido simbiótico puesto que alberga los genes encargados de la simbiosis con la planta (*nod*, *nif* y *fix*). Los plásmidos pRL7 y pRL8 son transferibles por conjugación [Young *et al.*, 2006].

Los 6 plásmidos de Rlv3841 tienen sistemas de replicación basados en el operón *repABC*. Existe una alta divergencia entre cada RepA y RepB presente en cada uno de los plásmidos, las identidades de aminoácidos varían en un intervalo de 41-61% para el caso de la proteína RepA, un 30%-43% para el caso de la proteína RepB y para el caso de RepC se observa de un 55% a un 68%. Estas variaciones podrían explicar la coexistencia en una misma cepa.

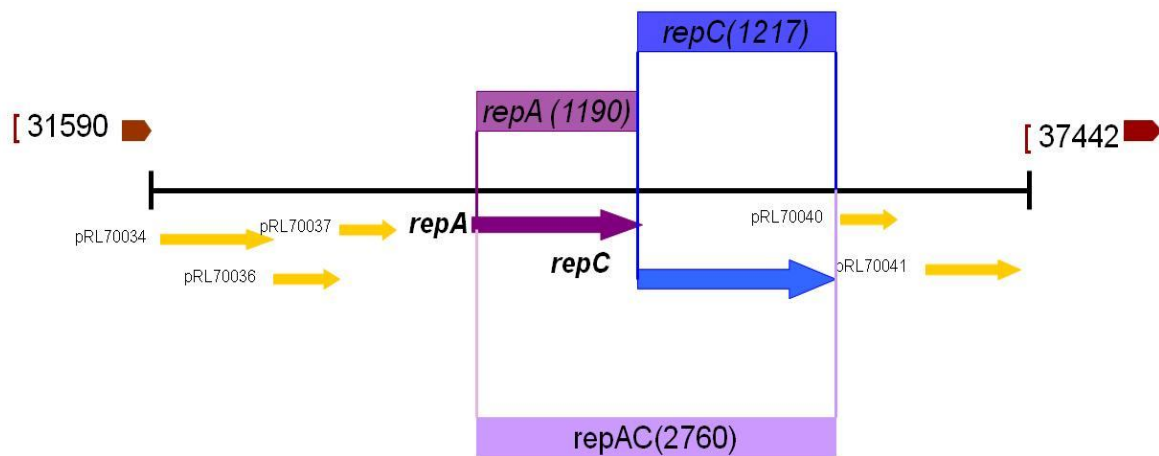


**Figura 7:** Representación gráfica del cromosoma y los 6 plásmidos de Rlv3841. Los plásmidos son representados a la misma escala relativa y el cromosoma a una cuarta parte de esta escala. [Young et al. 2006]

Los replicones que contienen 2 operones *repABC* se encuentran comúnmente dentro de las alfa proteobacterias y presentan un bajo grado de identidad de secuencia entre ellos [Cevallos et al., 2008]. Esta observación sugiere que la presencia de dos operones *repABC* en el mismo replicón no produce inestabilidad de éste en la bacteria hospedera.

Se ha demostrado que algunas combinaciones de cointegración plásmido-plásmido o plásmido-cromosoma son estables y no causan problema a la bacteria [Mavingui et al., 2002; Guo et al., 2003]. Alternativamente se han encontrado plásmidos que presentan nuevas configuraciones de los genes *repA*, *repB*, *repC*. Un ejemplo de ello lo constituye el plásmido pMNOD01 de *Methylobacterium nodulans* ORS2060 que contiene un gen *repC* seguido de un gen que codifica para una pequeña proteína hipotética, el cual es seguido de los

genes *repAB*; todos ellos codificados en la misma cadena de ADN. Otro claro ejemplo de una nueva configuración está presente en el plásmido pRL7 nuestra cepa de estudio *R. leguminosarum 3841*, este plásmido alberga 2 operones *repABC* completos y una versión trunca del operon en donde hay una ausencia del gen *repB* y solo se encuentra el gen *repA* seguido del gen *repC* [Young et al., 2006], como se muestra en la figura 8.



**Figura 8: Representación esquemática de la localización y tamaño de los genes *repAC* dentro del plásmido pRL7.** Mostrando la presencia de otros seis genes en la misma dirección. El gen *repA* se encuentra representado por una flecha morada y mide 1190nt, mientras que el gen *repC* se encuentra representado con una flecha azul y mide 1217 nt, dando un tamaño total del operon de 2760nt.

# CAPITULO 3

## JUSTIFICACION Y PLANEAMIENTO DEL PROBLEMA

La existencia de nuevas configuraciones de los genes *repA*, *reB* y *repC* del operon *repABC*, nos sugieren que existió un intercambio de genes y mecanismos de recombinación genética, que tuvieron un papel crucial en la generación de nuevos sistemas de replicación dentro de las *alfaproteobacterias*.

Para poder confirmar lo anterior es necesario demostrar la funcionalidad de los nuevos arreglos genéticos como sistemas de replicación.

# CAPITULO 4

## OBJETIVO GENERAL

El objetivo general del trabajo es demostrar que la versión trunca *repAC* presente en el pRL7 del *R. leguminosarum* funciona como un nuevo sistema de replicación.

## OBJETIVOS PARTICULARES

- Determinar si los genes *repAC* al igual que el operon *repABC*, son capaces de mantener la replicación de un vector suicida.
- Determinar si el papel de los genes *repA* y *repC* presentes en la configuración *repAC*, es el mismo que se observa en los genes *repA* y *repC* presentes en el operón *repABC*.
- Determinar la organización de la unidad transcripcional en la que se encuentran los genes *repAC* del pRL7.



# CAPITULO 5

## DISUSIÓN Y PERSPECTIVAS

La replicación y la segregación de material genético son procesos esenciales para la viabilidad bacteriana. En la mayoría de los replicones bacterianos los genes encargados de la replicación y de la segregación se encuentran en *loci* diferentes, cada uno de ellos con sus propios sistemas de regulación. Para el caso de los replicones que cuentan con un sistema de replicación/segregación compacto *repABC*, estas funciones se encuentran co-reguladas, proporcionándole así un estricto control y coordinación de estas funciones. Este rearrreglo genético es poco usual dentro del mundo bacteriano, pareciese estar restringido solo a las alfa-proteobacterias mayoritariamente en el orden de los *Rhizobiales*

Los organismos pertenecientes a este orden, se caracterizan por tener genomas altamente complejos, pueden contener muchos replicones accesorios al cromosoma principal; pueden albergar cromosomas secundarios; megaplásmidos y plásmidos. Estos replicones extras contienen principalmente aquellos genes, que se utilizan en las estrategias de sobrevivencia a diferentes estreses, disminuyendo así la carga genética del cromosomalo que favorece a tiempos de duplicación más rápidos. La mayoría de los replicones extras de este orden bacteriano portan al menos un operón *repABC*, sin embargo, este operón nunca se ha encontrado en cromosomas principales en los cuales predomina la presencia de DnaA como proteína iniciadora de replicación y el sistema ParAB como el de segregación [Pinto et al 2012]. Como se observa, existe un sesgo definido entre los sistemas de replicación y segregación de cromosomas primarios y de los replicones extras presentes en una misma bacteria, ya que se observa una completa ausencia de transferencia de estos genes entre los cromosomas primarios y los replicones secundarios de una misma bacteria.

El porqué de este sesgo no queda claro: ciertamente el tamaño de los replicones no es la limitante para este sesgo debido a que se ha observado que el sistema *repABC* es capaz de replicar moléculas de más de 1Mb. Por lo tanto deben de existir otros factores aun no elucidados responsables de este fenómeno.

Los análisis de los plásmidos presentes en las *Rhizobiales* nos muestran que existen replicones que contienen un operón *repABC* (esta configuración es la más abundante), otros pueden contener hasta 2 operones *repABC*, o 2 operones *repABC* y una versión trunca del operon, o 1 operon *repABC* y una versión trunca del operón, siendo estos últimos los menos abundantes **[Pinto et al., 2012]**. Las versiones truncas que se han encontrado han sido: solo los genes *repAB*, *ctRNA-repC*, o la versión analizada en este trabajo la *repAC*. Es importante resaltar que estas versiones truncas nunca se han encontrado solas en un replicón, siempre se encuentra un operón completo en la misma molécula, lo que nos habla de la importancia de mantener un operón *repABC* íntegro para la adecuada replicación y segregación del replicón.

Como demostramos en este trabajo la versión *repAC* sigue siendo funcional para la replicación del plásmido, pero la segregación no funciona adecuadamente por la ausencia de *repB*. Las proteínas RepAB están filogenéticamente relacionadas a las proteínas ParAB y SopAB pertenecientes a los sistemas de partición de los plásmidos P1 y F respectivamente, los cuales se han estudiado ampliamente. Estos sistemas son altamente eficaces en la segregación de sus replicones, cada uno con su propio sistema de regulación **[Lee y Grossman 2006]**.

Los genes *repA* y *repB* se han encontrado en plásmidos, profagos episomales y cromosomas. En contraparte, el gen *repC* solo se ha encontrado en crómidos y plásmidos. RepC presenta cierto parecido a nivel de secuencia con DnaD de las Gram positivas, proteína del primosoma que se une al origen de replicación y que ayuda a reclutar la helicasa replicativa al origen; también participa en la regulación de la polimerización de DnaA **[Bonilla y Grossman 2012]**.

Se sabe que RepC es esencial para la replicación de los plásmidos que la portan, el origen de replicación del sistema se encuentra dentro de la región codificante de esta proteína y su sobre-expresión provoca un aumento considerable del número de copias del plásmido que la codifica [Pinto *et al.*, 2011; Cervantes-Rivera *et al.*, 2011]. Su mecanismo de acción no ha sido elucidado completamente y es una de las incógnitas más importantes de este sistema.

Una característica poco común y que encontramos en las *Rhizobiales* es que la mayoría de sus plásmidos pertenecen a la familia *repABC* incluso dentro de una misma cepa. Quizá los casos más excepcionales son las cepas *R. etli* CFN42 y *R. leguminosarum* 3841, cada una de ellas con seis grandes plásmidos y todos ellos tienen al menos un locus *repABC*; e incluso algunos contienen hasta 2 operones *repABC* como es el caso de los plásmidos p42a y p42f de *R. etli*, o algunos como el pRL7 de *R. leguminosarum* que contienen hasta 2 operones *repABC* y un operón trunco (*repAC*) [Young *et al.*, 2006; Gonzalez *et al.*, 2006 ]. Esta observación nos muestra que a pesar de tener el mismo sistema de replicación/segregación en varios plásmidos dentro de una misma bacteria, éstos son lo suficientemente divergentes para evitar problemas de incompatibilidad, disminuyendo de este modo la probabilidad de tener interacciones heterológicas entre los sistemas pertenecientes a los diferentes plásmidos que pudieran generar algún tipo de incompatibilidad. Se ha observado que las RepAs presentes en una misma bacteria no pueden tener más de un 61% de identidad entre ellas, mientras que para RepB es el 51%, para el caso de *repC* se ha observado que dentro de una misma bacteria pueden co-existir hasta con una homología mayor al 90% [Pinto *et al.*, 2012].

¿Cómo pueden funcionar en armonía estos sistemas? ¿Existen interacción entre los genes del operón presentes en los diferentes plásmidos o en la misma molécula?, si existen más de un operón en la misma molécula: ¿existirá un orden jerárquico en estos casos? o ¿alguno de los operones lleva la batuta de la

replicación y segregación de la molécula o la molécula se replica y segrega ocupando los 2 o más loci presentes? Estas son unas de las principales interrogantes aún sin resolver sobre el sistema.

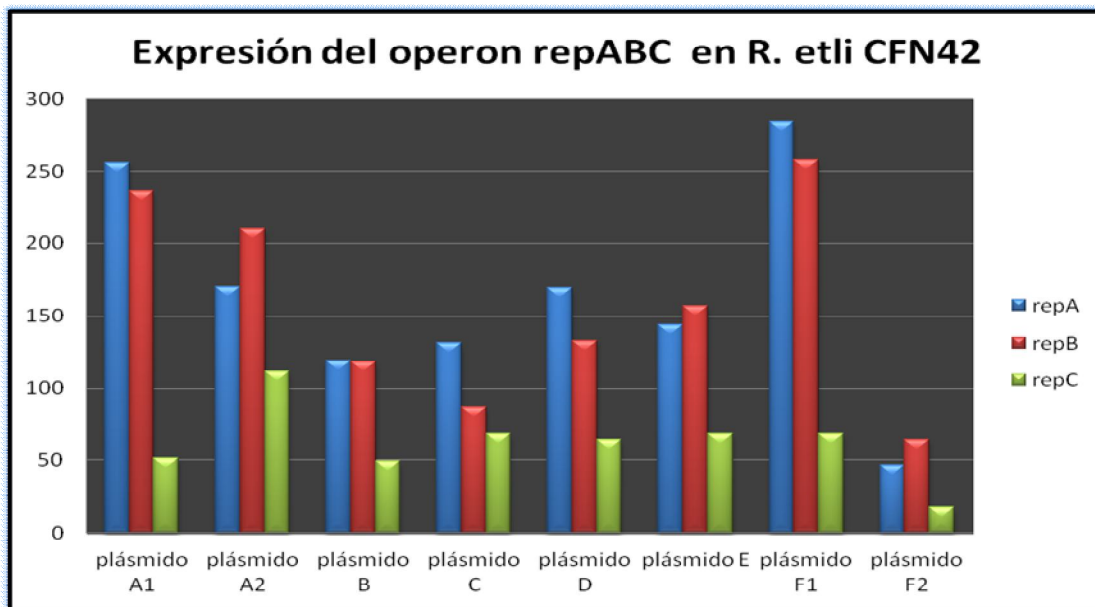
Con la intención de resolver estas incógnitas y aprovechando los estudios del transcriptoma de la cepa de *R. etli* CFN42 que se han hecho dentro del programa de genómica evolutiva del CCG (Tesis Doctoral del Biol. Gamaliel López Leal). Se estudiaron los niveles de transcripción de los genes *repA*, *repB* y *repC*, estos estudios muestran que a nivel transcripcional los genes *repA* y *repB* presentes en cada uno de los plásmidos se expresan a un nivel muy similar, mientras que el gen *repC*, en todos los casos, presenta un nivel de expresión más bajo. Para el caso de los plásmidos que contiene 2 operones *repABC* se muestra que uno de los operones muestra mayores niveles de expresión en todos sus genes comparado con el otro; y la diferencia más clara en lo que respecta a los niveles de expresión es el caso del plásmido p42f (figura 9) **[López-Leal et al. Datos no publicados]**.

Si planteamos la hipótesis de que el operón *repABC* que más se transcribe es el que tiene más oportunidad de dirigir la replicación del plásmido, entonces los análisis de expresión nos podría dar evidencias para identificar cuál de los operones *repABC* de un bireplicón dado es el que es el dominante. Sin embargo, aún se necesitan más estudios para comprender este fenómeno y poder dar una explicación más clara y contundente.

Los análisis de hibridización fluorescente *in situ* de los orígenes de replicación de los replicones *repABC* presentes en *Agrobacterium tumefaciens* (1 cromosoma secundario lineal y 2 plásmidos) y en *Sinorhizobium meliloti* (2 megaplásmidos), mostraron que los orígenes de cada replicón coexisten en los polos ocupando áreas discretas e independientes **[Kahng y Shapiro, 2003]**.

Lo que nos indica que a pesar de tener el mismo sistema de segregación cada uno es capaz de mantener posiciones específicas y diferentes para poder mantenerse establemente.

Aún no es claro es definir cómo llegan a esas posiciones dentro de la célula o si el que guarden posiciones diferentes en el polo se debe a más bien a un fenómeno de exclusión o en el polo residen “anclas” plásmido específicas que hacen que los plásmidos tengan esas localizaciones dentro de la célula bacteriana. Como se ha mencionado, aún hay mucho trabajo por hacer con este sistema, y mucha información que nos pueden aportar para entender mejor los sistemas de replicación de las bacterias.



**Figura 9:** Grafica que nos muestra los niveles de expresión de los genes *repA*, *repB* y *repC* presentes en cada uno de los plásmidos de la cepa de *R. etli* CFN42. Las barras azules representan a los genes *repA*, las rojas a los genes *repB* y las verdes a los genes *repC*. Los valores de expresión están dados en rpnM: el número de lecturas normalizadas con el tamaño de la librería, las condiciones en las que se hizo esta análisis fueron con un crecimiento de la bacteria a 30°C y a 200 r.p.m, se tomaron las muestras cuando las bacterias se encontraban en fase exponencial. [Lopez-leal]

## **BIBLIOGRAFIA**

**Anne-Brit Kolsto**, (1997)

Dynamic bacterial genome organization *Molecular Microbiology* **24**(2):241-248

**Barillà D.**, (2010)

One-way ticket to the cell pole: plasmid transport by the prokaryotic tubulin homolog TubZ. *Proc Natl Acad Sci U S A.* **107**(27):12061-2.

doi: 10.1073/pnas.1007331107.

**Bartosik, D., Baj, J. Włodarczyk, M.**, (1998)

Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. *Microbiology* **144**: 3149-3157.

**Bartosik D., Szymanik M., Wysocka E.**, (2001)

Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J. Bacteriol.* **21**: 6234-6243.

**Bingle LE, Thomas CM.**, (2001)

Regulatory circuits for plasmid survival. *Curr Opin Microbiol.* **2**:194-200. Review.

**Bonilla C.Y. y Grossman A.D.**, (2012)

The primosomal protein DnaD inhibits cooperative DNA binding by the replication initiator DnaA in *Bacillus subtilis*. *Journal of Bacteriology* **194**(18):5110-1117

doi. 10.1128/JB:00958-12.

**Bouffartigues E.,Buckle M., Badaut C., Travers A.,Rimsky S.**, (2007)

H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat. Struct. Mol. Biol.* **14**: 441-448.

**Casjens Sherwood**, (2003)

Prophages and bacterial genomics: what have we learned so far?. *Molecular Microbiology* **49**(2):277-300

doi: 10.1046/j.1365-2958.2003.03580.x

**Castillo-Ramírez S, Vázquez-Castellanos JF, González V, Cevallos MA.**, (2009)

Horizontal gene transfer and diverse functional constraints within a common replication-partitioning system in Alphaproteobacteria: the *repABC* operon. *BMC Genomics.***10**: 536

doi:10.1186/1471-2164-10-536)

**Cevallos MA, Porta H, Izquierdo J, Tun-Garrido C, García-de-los-Santos A, Dávila G, Brom S., (2002)**

Rhizobium etli CFN42 contains at least three plasmids of the repABC family: a structural and evolutionary analysis. *Plasmid*. **48**(2):104-16.

**Cevallos M.A., Cervantes-Rivera R., Gutiérrez-Ríos R.M., (2008)**

The *repABC* plasmid family. *Plasmid*. **60**:19-37.

**Cervantes-Rivera, R., Romero-Lopez, C., Berzal-Herranz, A. & Cevallos, M. A. (2010)**

Analysis of the mechanism of action of the antisense RNA that controls the replication of the repABC plasmid p42d. *J. Bacteriol.* **192**, 3268–3278.

**Cervantes-Rivera, R., Pedraza-Lopez, F., Perez-Segura, G. & Cevallos, M. A. (2011)**

The replication origin of a repABC plasmid. *BMC Microbiol.* **11**, 158

**Chai, Y., Winans S.C., (2005a)**

A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol. Microbiol.* **56**:1574-1585.

**Chai, Y., Winans, S. C., (2005b)**

RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between *repA* and *repB* for plasmid partitioning and autorepression. *Mol. Microbiol.* **58**:1114-1129.

**Chandler M. and Mahillon J. (2002)**

Insertion sequences revisited, in *MobileDNA II* (Craig, N. L., Craigie, R., Gellert, M., and Lambowitz, A. M., eds.), ASM Press, Washington, DC, pp. 305–366.

**Chant E.L., Summers D.K., (2007)**

Indole signalling contributes to the stable maintenance of Escherichia coli multicopy plasmids. *Mol Microbiol.* **63**(1):35-43.

**Chen Inès and David Dabnau, (2004)**

DNA uptake during bacterial transformation. *Nature Reviews* 2: 241-248  
**doi:** 10.1038/nmicro844

**Cho H, Winans S.C., (2005)**

VirA and VirG activate the Ti plasmid repABC operon, elevating plasmid copy number in response to wound-released chemical signals. *Proc Natl Acad Sci U S A.* **102**(41):14843-8.

- Claverys, J.P. & Martin, B.** (2003)  
Bacterial \*competence\* genes signatures of active transformation, or only remnants? *Trends. Microbiol.* 11:161-165.
- Craig, N. L.**, (1996)  
Transposition, in *Escherichia coli and Salmonella: Cellular and Molecular Biology, 2nd ed.* (Neidhardt, F. C., ed.), ASM Press, Washington, D.C, pp. 2339–2362
- del Solar G, Espinosa M.**, (2000)  
Plasmid copy number control: an ever-growing story. *Mol Microbiol.* 3:492-500. Review.
- Ebersbach G, Sherratt DJ, Gerdes K.**, (2005)  
Partition-associated incompatibility caused by random assortment of pure plasmid clusters. *Mol Microbiol.* 56(6):1430-40.
- Fineran P.C., N.K. Petty, and G.P.C. Salmond**, (2009)  
Transduction: Host DNA transfer by bacteriophages. Elsevier Inc 666-679.
- Frost S. Laura, Raphael Leplae, Anne O. Summers and Ariane Toussaint** (2005)  
Mobile genetic elements: The agents of open source evolution. *Nature/reviews* 3:722-732
- Gallie, D. R., Hagiya, M. & Kado, C. I.** (1985)  
Analysis of *Agrobacterium tumefaciens* plasmid pTiC58 replication region with a novel high-copy-number derivative. *J. Bacteriol.* 161, 1034–1041.
- Garner EC, Campbell CS, Mullins RD.**, (2004)  
Dynamic instability in a DNA-segregating prokaryotic actin homolog. *Science.* 306(5698):1021-5.
- Garner EC, Campbell CS, Weibel DB, Mullins RD.**, (2007)  
Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. *Science.* 315(5816):1270-4.
- Gerdes K, Møller-Jensen J, Bugge Jensen R.**, (2000)  
Plasmid and chromosome partitioning: surprises from phylogeny. *Mol Microbiol.* 37(3):455-66.
- Gerdes K, Howard M, Szardenings F.**, (2010).  
Pushing and pulling in prokaryotic DNA segregation. *Cell.* 141(6):927-42.  
**doi:** 10.1016/j.cell.2010.05.033. Review.



- Ghosh SK, Hajra S, Paek A, Jayaram M., (2006)**  
Mechanisms for chromosome and plasmid segregation. *Annu Rev Biochem.* **75**:211-41. Review.
- González V, Santamaría RI, Bustos P, Hernández-González I, Medrano-Soto A, Moreno-Hagelsieb G, Janga SC, Ramírez MA, Jiménez-Jacinto V, Collado-Vides J, Dávila G., (2006)**  
The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci U S A.* **103**(10):3834-9.
- Guo X, Flores M, Mavingui P, Fuentes SI, Hernández G, Dávila G, Palacios R. (2003)**  
Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Res.* 2003 Aug; **13**(8):1810-7.
- Harrison P.W., R. P. Lower, N.K. Kim, y J.P. Young, (2010)**  
Introducing the bacterial “cromid”: not a chromosome, not a plasmid. *Trends in Microbiol.* **18**:141-148.
- Hayes F, Barillà D.(2006)**  
The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat Rev Microbiol.***4**(2):133-43. Review.
- Holger Heuer,Zaid Abdo & Kornelia Smalla, (2008).**  
Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. *FEMS Microbiology Ecol.* **65**:(361-371).  
**doi**:10.1111/j.1574-6941.2008.00539.x
- Holmes V.F. and Cozzarelli N.R, (2000)**  
Closing the ring links between SMC proteins and chromosome partitioning, condensation and supercoiling. *Proc. Natl. Acad. Sci: USA* **97**:1322-1324
- Hu MX, Zhang X, Li EL, Feng YJ., (2010).**  
Recent advancements in toxin and antitoxin systems involved in bacterial programmed cell death. *Int J Microbiol.* 2010:781430.  
**doi**: 10.1155/2010/781430.
- Izquierdo J., Venkova-Canova T., Ramírez-Romero M.A., Téllez-Sosa J., Hernández-Lucas I., Sanjuan J., Cevallos M.A., (2005).**  
An antisense RNA plays a central role in the replication control of a *repC* plasmid. *Plasmid* **54**: 259-277.

**Juhas Mario, Jan Roelof van der Meer, Muriel Gaillard, Rosalinda M. Harding, Derek W. Hood & Derrick W. Crook, (2009).**

Genomic Islands: tools of bacterial horizontal gene transfer an evolution. *FEMS Microbiology Rev.* **33**:376-393.

**Kahng, L. S. & Shapiro, L. (2003)**

Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J. Bacteriol.* **185**, 3384–3391.

**Kollek R, Oertel W, Goebel W., (1978).**

Isolation and characterization of the minimal fragment required for autonomous replication ("basic replicon") of a copy mutant (pKN102) of the antibiotic resistance factor R1. *Mol Gen Genet.* **162**(1):51-7.

**Larsen RA, Cusumano C, Fujioka A, Lim-Fong G, Patterson P, Pogliano J., (2007)**

Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in *Bacillus thuringiensis*. *Genes Dev.* **21**(11):1340-52.

**Lee PS, Grossman AD., (2006)**

The chromosomal partitioning proteins Soj(ParA) and Spo0J(ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in *Bacillus subtilis*. *Mol. Microbiology* **60**: 853-869.

**Lerat E., Daubin V., Ochman H., Moran N.A., (2005)**

Evolutionary Origins of Genomic Repertoires in Bacteria. *PLoS Biol* **3**(5):807-814.

doi:10.1371/journal.pbio.0030130

**Lerman L.S., (1971)**

A transition to a compact form of AND in polymer solutions. *Proc. Natl. Acad. Sci. USA* **68**:1886-1890

**Lewis K., (2010)**

Persister cells. *Annu Rev Microbiol.* **64**:357-72. Review.

doi: 10.1146/annurev.micro.112408.134306.

**Livny J., Yamaichi Y., Waldor M.K., (2007).**

Distribution of centromere-like parS sites in bacteria: insights from comparative genomics. *J Bacteriol* **189**: 8693–8703.

**López-leal Gamaliel** (2013)

Analisis del sigmulon dependiente de RpoH en *R. etli*. Tesis de doctorado del programa de Ciencias Biomédicas CCG-UNAM

**MacLellan, S.R., Smallbone, L.A., Sibley C.D., Finan T. M.**, (2005).

The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. *Mol. Microbiol.* **55**: 611-623.

**Patrick Mavingui; Margarita Flores; Xianwu Guo; Guillermo Dávila; Xavier Perret; William J Broughton; Rafael Palacios** (2002)

Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. *Journal of bacteriology*; **184**(1):171-6.

**Mazur Andrzej & Koper Piotr**, (2012)

Rhizobial plasmids-replication, structure and biological role. *Cent. Eur. J. Biol.* **7**(4):571-586 .

**doi**: 10.2478/s11535-012-0058-8

**Miller R.V.**, (2001)

Environmental bacteriophage-host interactions: factors contribution to natural transduction. *Antonie Van Leeuwenhoek* **79**:141-147

**Mott ML, Berger JM.**, (2007).

DNA replication initiation: mechanisms and regulation in bacteria. *Nat Rev Microbiol.* **5**(5):343-54. Review.

**Moya A., R. Gil., A. Latorre, J. Peretó, M. Garcillán-Barcía and F. de la Cruz**, (2009)

Towards minimal bacterial cells: evolution vs design. *FEMS Microbiol. Rev.* **33**:225-235.

**Ogura T. and Hiraga S.**, (1983)

Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc Natl Acad Sci U S A.* **80**(15): 4784–4788.

**Pappas KM, Winans S.C.** , (2003).

A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol Microbiol.* **48**(4):1059-73.

**Pappas KM, Winans S.C.**, (2003)a.

The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. *Mol Microbiol.* **49**(2):441-55.

**Pappas KM, Winans S.C.**, (2003)b

Plant transformation by coinoculation with a disarmed *Agrobacterium tumefaciens* strain and an *Escherichia coli* strain carrying mobilizable transgenes. *Appl Environ Microbiol.* **69**(11):6731-9.

**Petersen J., Brinkmann H., Pradella S.**, (2009).

Diversity and evolution of *repABC* type plasmids in Rhodobacterales. *Environ Microbiol.* Jul 10.

doi:10.1111/j.1462-2920.2009.01987.x

**Peter M. Bennett** (2004)

Genome plasticity: insertion sequence elements, transposons and integrons, and DNA rearrangement. *Methods Mol Biol.* **266**:71-113.

doi: 10.1385/1-59259-763-7:071

**Pinto, U.M., Flores-Mireles, A.L., Costa, E.D., Winans, S.C.**, (2011).

RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within *repC* and functions only in *cis*. *Mol: Microbiol.* **81**:1593-1606 DOI: 10.1111/j.1365-2958.2011.07789.x.

**Ptacin Jerod L., Steven F. Lee, Ethan C. Garner, Esteban Toro, Michel Eckart, Luis T. Comolli, W.E. Moerner & Lucy Shapiro** (2010)

A spindle-like apparatus guides bacterial chromosome segregation. *Nature Cell Biology.* 12:791-798

doi:10.1038/ncb2083

**Ramírez-Romero, M.A., Téllez-Sosa, J., Barrios, H., Pérez-Oseguera, A., Rosas. V., Cevallos, M.A.**(2001).

RepA negatively autoregulates the transcription of the *repABC* operon of *Rhizobium etli* symbiotic plasmid basic replicon. *Mol. Microbiol.* **42**: 195-204.

**Ringgaard Simon, Jeronen van Zon, Martin Howard and Kenn Gerdes** (2009).

Movement and equipositioning of plasmids by ParA filament disassembly. *PNAS.* **106**(46):19369-19374

**Rocha P.C. Eduardo**, (2008)

The Organization of the bacterial Genome. *Annu.Rev. Genet.* **42**: 211-233.

doi: 10.1146/annurev.genet.42.110807.091653

**Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, Frank B, Gaigalat L, Goesmann A, Groeger C, Gross F, Jelsbak L, Jelsbak L, Kalinowski J, Kegler C, Knauber T, Konietzny S, Kopp M, Krause L, Krug D, Linke B, Mahmud T, Martinez-Arias R, McHardy AC, Merai M, Meyer F, Mormann S, Muñoz-Dorado J, Perez J, Pradella S, Rachid S,**

**Raddatz G, Rosenau F, Rückert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorhölter FJ, Weissman KJ, Welch RD, Wenzel SC, Whitworth DE, Wilhelm S, Wittmann C, Blöcker H, Pühler A, Müller R., (2007).**

Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat Biotechnol.* **25**(11):1281-1289.

**Sibley CD, MacLellan SR, Finan T., (2006).**

The *Sinorhizobium meliloti* chromosomal origin of replication. *Microbiology.* **152**(Pt 2):443-55.

**Slater S.C., Goldman B.S., Goodner B., Setubal J.C., Farrand S.K., Nester E.W., Burr T.J., Banta L., Dickerman A.W., Paulsen I., Otten L., Suen G., Welch R., Almeida N.F., Arnold F., Burton O.T., Du Z., Ewing A., Godsy E., Heisel S., Houmiel K.L., Jhaveri J., Lu J., Miller NM., Norton S., Chen Q., Phoolcharoen W., Ohlin V, Ondrusek D., Pride N., Stricklin S.L., Sun J., Wheeler C., Wilson L., Zhu H., Wood D.W., (2009).**

Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol.* **191**: 2501-2511.

**Schmidt H. & Hensel M., (2004)**

Pathogenicity islands in bacterial pathogenesis. *Clin Microbiol Rev* **17**: 14–56.

**Soberón, N., Venkova-Canova, T., Ramírez-Romero, M.A., Téllez-Sosa, J, Cevallos, M A., (2004).**

Incompatibility and the partitioning site of the *repABC* basic replicon of the symbiotic plasmid from *Rhizobium etli*. *Plasmid* **51**: 203-216.

**Summers D., (1998)**

Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol Microbiol.* **29**(5):1137-45. Review.

**Szardenings F, Guymer D, Gerdes K., (2011)**

ParA ATPases can move and position DNA and subcellular structures. *Curr Opin Microbiol.* **14**(6):712-8. Review.

doi: 10.1016/j.mib.2011.09.008.

**Tabata S, Hooykaas PJ, Oka A., (1989).**

Sequence determination and characterization of the replicator region in the tumor-inducing plasmid pTiB6S3. *J Bacteriol.* **171**(3):1665-72.

**Thomas Christopher M. and David Summers, (2008)**

Bacterial Plasmids. *Encyclopedia of Life Sciences (ELS)*. John Wiley & Sons, Ltd: Chichester.

doi: 10.1002/9780470015902.a0000468.pub2

**Tomizawa Jun-ichi** (1984).

Control of ColE1 plasmid replication: The process binding of RNAI to the primer transcription. *Cell* **38**(3):861-870

**Toro E., Shapiro L.**, (2010).

Bacterial chromosome organization and segregation. *Cold Spring Harb Perspect Biol.* **2**(2):a000349. Review.

**doi:** 10.1101/cshperspect.a000349.

**Uchiyama Ikuc**, (2008)

Multiple genome alignment for identifying the core structure among moderately genomes. *BMC Genomics* 2008 **9**:515

**doi:** 10.1186/1741-2164-9-515

**Vecchiarelli A.G., Mizuuchi K., Funnell B.E.** (2012)

Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. *Mol Microbiol.* **86**(3):513-23.

**doi:** 10.1111/mmi.12017.

**Venkova-Canova, T., Soberón, N.E., Ramírez-Romero, M.A., Cevallos, M.A.**, (2004).

Two discrete elements are required for the replication of repABC plasmid: an antisense RNA and a stem-loop structure. *Mol. Microbiol.* **54**:1431-1444.

**Wickner S.H. and D.K. Chattoraj**, (1987)

Replication of a mini P1 plasmid DNA in vitro requires two initiation proteins, encoded by *repA* gene of phage P1 and *dnaA* gene of *Escherichia coli*. *Proc. Natl. Acad. Sci.* **84**(12):3668-3672

**Yuan J, Yamaichi Y, Waldor MK.**, (2011)

The three vibrio cholerae chromosome II-encoded ParE toxins degrade chromosome I following loss of chromosome II. *J Bacteriol.* **193**(3):611-9.

**doi:** 10.1128/JB.01185-10.

**Young J.P., Crossman L.C., Johnston A.W., Thomson N.R., Ghazoui Z.F., Hull K.H., Wexler M., Curson A.R., Todd J.D., Poole P.S., Mauchline T.H., East A.K., Quail M.A., Churcher C., Arrowsmith C., Cherevach I., Chillingworth T., Clarke K., Cronin A., Davis P., Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Mungall K., Norbertczak H., Rabinowitsch E., Sanders M., Simmonds M., Whitehead S., Parkhill J.**, (2006).

The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol.* **7**:R34.





## The *repAC* replication system of the *Rhizobium leguminosarum* pRL7 plasmid is functional: Implications regarding the origin and evolution of *repABC* plasmids

Gabriela Pérez-Segura, Ángeles Pérez-Oseguera, Miguel A. Cevallos\*

Programa de Genómica evolutiva, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apartado Postal 565-A, Cuernavaca, Morelos, Mexico

### ARTICLE INFO

#### Article history:

Received 26 March 2012

Accepted 26 August 2012

Available online 4 September 2012

Communicated by Dr. Dhruba K. Chattoraj

#### Keywords:

Alpha-proteobacteria

Plasmid segregation

Plasmid evolution

Origin of replication

### ABSTRACT

The *repABC* replication/partitioning systems are commonly found in alpha-proteobacteria plasmids and in secondary chromosomes. All of the elements required for their replication and stable maintenance are encoded within a single transcription unit: the *repABC* operon. The *repC* gene encodes an initiator protein, while *RepA*, *RepB* and centromere-like sequence (*parS*) direct plasmid segregation. Strains containing two or more *repABC* plasmids are a common feature in some alpha proteobacteria groups, indicating that the *repABC* plasmid family embraces several incompatibility groups. Genes encoded within *repABC* operons are highly dynamic: each one possess its own distinctive phylogeny and homologous recombination events are common within these operons. Additionally, alpha-proteobacterial genomes contain *repAB* genes not associated with the *ctrRNA* or with *repC* as well as plasmids whose replication depends on a *ctrRNA-repC* module without the participation of *repAB* genes. Some alphaproteobacteria have *repC* genes clustered with other genes that are not involved in replication/partitioning functions. These atypical associations of genes could have an important role in the origin and diversification of new plasmids. Here we evaluated the functionality and possible evolutionary consequences of one of these atypical gene associations: the *repAC* genes present in the *Rhizobium leguminosarum* plasmid pRL7. The *repAC* genes are organized in an operon and they are capable of sustaining replication but in an unstable manner. *RepC* was essential for replication, and the origin of replication resides within its coding region. In contrast, *RepA* plays a minor role in the negative regulation of its own transcription.

© 2012 Elsevier Inc. All rights reserved.

### 1. Introduction

Genome sequencing projects have shown that *repABC* replication/partitioning systems are common in alpha-proteobacterial plasmids. They are also frequently detected on secondary chromosomes and on large low copy-number plasmids in members of the Rhizobiales, Rhodospirillales, and Rhodobacterales bacterial orders (Cevallos et al., 2008; Petersen et al., 2009). These replication/partitioning systems are named after *repABC* because all of the ele-

ments required for their replication and stable maintenance in proliferating cells are encoded within a single transcription unit, the *repABC* operon. All *repABC* operons include three protein-encoding genes in the same relative order (*repA*, *repB*, and *repC*), a small antisense RNA regulatory gene, mentioned hereafter as counter-transcribed RNA (*ctrRNA*), and at least one centromere-analog (*parS*) (Cervantes-Rivera et al., 2010; Ramírez-Romero et al., 2001; Pappas and Winans, 2003).

*RepA* is a DNA binding protein and a Walker-type ATPase of the MinD/ParA superfamily with the putative capacity to polymerize into filaments. *RepB* is a DNA binding protein that interacts with the *parS* site and forms

\* Corresponding author. Fax: +52 7773 17 55 81.

E-mail address: [mac@ccg.unam.mx](mailto:mac@ccg.unam.mx) (M.A. Cevallos).



oligomers with its cognate RepA (Cevallos et al., 2008). Both proteins have similarities with other plasmid- and chromosome-partitioning proteins. RepA and RepB play dual roles in plasmid maintenance: they bind *parS* for in plasmid partitioning and an operator sequence in the negative transcriptional regulation of their own operon (Bartosik et al., 2001; Chai and Winans, 2005b; Pappas and Winans, 2003; Ramírez-Romero et al., 2000; Soberón et al., 2004). RepC, a protein that does not show similarities with members of other protein families, is crucial for plasmid replication. The origin of replication of *repABC* plasmids resides within the *repC* coding sequence (Cervantes-Rivera et al., 2011; Pinto et al., 2011). The region located between *repB* and *repC* genes encodes a small antisense RNA (ctrRNA) that has a crucial role in modulating RepC levels in the cell (Venkova-Canova et al., 2004; MacLellan et al., 2005; Chai and Winans, 2005a; Cervantes-Rivera et al., 2010).

Despite their apparent structural homogeneity, *repABC* operons have highly diverse DNA sequences, and some of the elements are present in only a subset of the family members. Generally, these operons vary with respect to the presence of peptide-encoding minigenes, the number and class of the regulatory elements involved in operon transcription, and the number and position of *parS* sequences (Cevallos et al., 2008; Pappas, 2008).

Plasmid stability depends on a delicate balance between proteins and DNA sites involved in plasmid replication and segregation. The introduction of a new plasmid containing the same or similar replication or segregation genes may disrupt this balance and provoke plasmid incompatibility. Strains within alpha-proteobacteria frequently contain several *repABC* replicons, indicating that *repABC* plasmids belong to several incompatibility groups (Cevallos et al., 2008; Petersen et al., 2009).

Plasmids containing two *repABC* operons are also commonly found, but with a low degree of sequence identity between them (Cevallos et al., 2008). This observation suggests that the presence of two *repABC* operons on the same plasmid does not induce instability. Furthermore, cointegrates between two different *repABC* plasmids created under laboratory conditions can also be stable (Mavingui et al., 2002; Guo et al., 2003).

Recent analyses indicate that proteins encoded by *repABC* operons have a distinct phylogeny and that homologous recombination events are common within these operons (Slater et al., 2009; Castillo-Ramírez et al., 2009). Additionally, alpha-proteobacterial genomes contain *repAB* genes not associated with the ctrRNA or with *repC* as well as plasmids whose replication depends on a ctrRNA-*repC* module without the participation of *repAB* genes (Izquierdo et al., 2005). Some alpha-proteobacteria have *repC* genes clustered with other genes that are not involved in replication/partitioning functions. Alternatively, new configurations of *repA*, *repB* and *repC* genes have been identified, such as those found on the plasmid pMNOD01 from *Methylobacterium nodulans* ORS 2060, containing a *repC* gene followed by a small hypothetical protein-encoding gene that is followed by the *repAB* genes; all of these genes are encoded in the same DNA strand. Another example involves the plasmid pRL7 from *Rhizobium leguminosarum*

bv. *viciae* 3841, which contains two *repABC* operons and a “truncated” version containing only the *repA* and *repC* genes (Young et al., 2006). These observations suggest that gene shuffling and gene recombination have a crucial role in the generation of new replication systems in alpha-proteobacteria.

It is necessary to demonstrate the functionality of these “atypical” gene clusters to support this hypothesis. In the work presented here, we show that the “truncated” *repAC* version of the *repABC* operon, present in the plasmid pRL7 from *R. leguminosarum* bv. *viciae* 3841, is able to sustain the replication of a suicide vector it is cloned into. We also determined the minimal region that is sufficient to demonstrate this ability, and show that the origin of replication of this replication system resides within the *repC* gene, as has been demonstrated for other *repABC* systems (Cervantes-Rivera et al., 2011; Pinto et al., 2011).

## 2. Material and methods

### 2.1. Bacterial strains and growth conditions

*Rhizobium etli* CFNX107 (*R. etli* CE3 *recA::ΩSpc<sup>R</sup>*, *Sm<sup>R</sup>*, derivative, cured of plasmids p42d and p42a) (Martínez-Salazar et al., 1991) and *R. leguminosarum* bv. *viciae* 3841 (Johnston and Beringer, 1975) cell cultures were grown in PY medium supplemented with 10 mM CaCl<sub>2</sub> at 30 °C (Noel et al., 1984), S17-1 (Simon et al., 1983), and *Escherichia coli* DH5α (Hanahan, 1983) cell cultures were grown in Luria-Bertani medium at 37 °C. Nalidixic acid (20 µg/ml), spectinomycin (100 µg/ml), gentamicin (30 µg/ml), and chloramphenicol (30 µg/ml) were added to the culture media when appropriate.

### 2.2. DNA isolation and manipulation

Genomic DNA from *R. leguminosarum* bv. *viciae* 3841 was isolated using a DNA isolation kit (Amersham), and plasmid DNA using the High Pure Plasmid isolation Kit (Roche). DNA was digested and ligated under the conditions specified by the enzyme manufacturers (New England Biolabs, Fermentas).

### 2.3. Construction of suicide vectors pDOP and pDOABTi

The backbone of plasmid pDOP was constructed by deleting a superfluous 889 bp DNA fragment from plasmid pBC SK+ (Stratagene) using PCR. To obtain this derivative a PCR product was made using primers DELBC-U and DELBC-L (Table 1), Platinum *Taq* DNA polymerase High Fidelity from Invitrogen, and plasmid pBC SK+ as template. These primers add *MluI* and *BglIII* restriction sites on one side of the amplification product and *MluI* and *NsiI* sites on the other side. The PCR product was digested with *MluI* and circularized with T4 DNA ligase (Fermentas) to obtain plasmid pΔBC. To make plasmid pΔBC mobilizable, a 784 bp region, reported to confer the maximal efficiency of conjugation of RK2 plasmid, was obtained by PCR utilizing primers ORIT-U and ORIT-L (Table 1), plasmid RK2 as template, and the *Taq* DNA polymerase mentioned above (Guiney

**Table 1**  
Primers used in this work.

Primer name	Sequence
DELBC-U	TAACGCGTATGCATAAAGAACGTGGACTCC
DELBC-L	CGACGCGTAGATCTCTTATTACGGCGTAGC
ORIT-U	GGATCCCGGCTTGGCTTGCTCG
ORIT-L	GGATCCTTTCCGCTGCATAACCC
PARDE-U	ATGCATCCCTTGGTCAAATGGG
PARDE-L	ACGCGTAGTACGCCATCAGGACG
repABTI58-U	ACCGGTCTACGTGCAGATTTCAC
repABTI58-L	ATGCATAA CCCTGACTATTGGTC
repCRLe-U	TCTAGAAAGGAACAGCTATGCAGACTGAACACA
repCRLe-L	AGGCTTCTACGCCGACCTGGTTC
repACRLe-U	TCTAGAAACGCGATGGAATTG
repACRLe-L	AGGCTTATCGCCGTGATATTTAT
RiL6.5-up	AATCATTTAGACCTCTGGTCTCTCAAGT
RiL6.5-low	AGACTACGGCCGCTACGCCGACCTGGTTCG
repACXba-U	TCTAGAAACGCGATGGAATTG
repACSaL-L	GTCGACCCGCTCCAGGATGCAGGC
repCMIu-U	TCTAGACTTTACGGGTGTCAC
repACHindIII-U	AGGCTTAACGCGATGGAATTGCGT
repACKpn-L	GGTACCATCGCCGTGATATTTAT
repAKpn-L	GGTACTCTAGTCTGCATTTTCGTC

and Yakobson, 1983). These primers introduce *Bam*HI sites on each side of the PCR product. This amplification product was digested with *Bam*HI (Fermentas) and inserted into the *Bgl*III site of plasmid p $\Delta$ BC, to generate plasmid pDO. Finally, the 790 bp minimal stabilization region (*parDE*) of plasmid RK2 was obtained using PCR with primers PARDE-U and PARDE-L (Table 1) and plasmid RK2 as template, and incorporated into pDO plasmid using *Mlu*I and *Nsi*I sites present in both plasmid pDO and the PCR product to obtain pDOP (Roberts and Helinski, 1992).

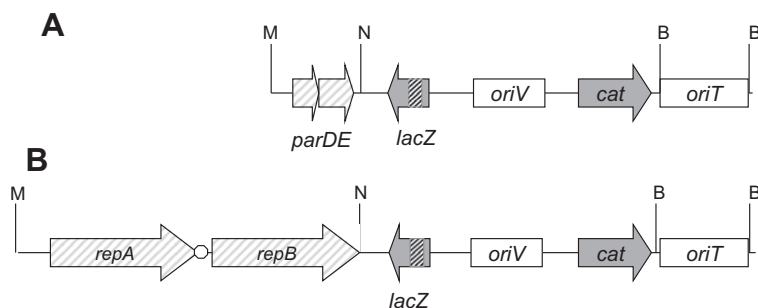
To construct pDOABTi, the segregation system of *Agrobacterium tumefaciens* plasmid pTiC58 was PCR amplified with primers repABTI58-U and repABTI58-L (Table 1), and genomic DNA of *A. tumefaciens* C58 as template (Chai and Winans, 2005b). The PCR product, containing on one side a *Mlu*I site and on the other a *Nsi*I site, was then introduced identical restriction sites of pDO to produce pDOABTi. Both vectors were incapable of autonomous replication in *Rhizobium* (data not shown). Linear maps of these vectors are presented in Fig. 1. pDOP and pDOABTi repli-

cate in *E. coli* using the origin of replication of their parental plasmid (pBC SK+) suggesting that the plasmid copy-number of these new vectors in *E. coli* is around ~300–500 copies per cell. Plasmids pDOP and pDOABTi could be introduced into *Rhizobium* by conjugation.

#### 2.4. Construction of pDOP derivatives

Two PCR products both containing the *repAC* genes from plasmid pRL7 but differing in the length of the upstream and downstream regions were obtained using Platinum High Fidelity *Taq* Polymerase (Invitrogen). Total DNA from *R. leguminosarum* bv. *viciae* 3841 was used as a template, and primers were designed according to the genome sequence described recently (Young et al., 2006) (Table 1). The first PCR product included a region of approximately 6.5 kb located between nucleotides 30,310–36,950 of the plasmid pRL7 and was obtained using the RiL6.5-up and RiL6.5-low primers. The second amplification product was created utilizing the repACRLe-U and repACRLe-L primers and it encompassed a 2,760-bp region that began 206 bp upstream of the *repA* initiation codon and ended 159 bp downstream of the *repC* stop codon. A third PCR amplification product containing *repC* with a Shine-Dalgarno sequence that added 6 bp upstream of the *repC* initiation codon was obtained in a similar manner using the repCRLe-U and repCRLe-L primers. The amplification products were cloned using the TOPO-TA cloning kit (Invitrogen), and the products were subcloned into pDOP. Inserts were cloned after digestion with *Xba*I and *Hind*III, except for the construct containing the 6.5-kb product the digestions were with *Xba*I and *Kpn*I. Inserts were introduced into the same relative orientation with respect to the Plac promoter present in pDOP. The construct containing the 6.5-kb fragment was named pAC6.5; the construct containing the 2.76-kb fragment was named pAC200. The construct carrying only *repC* was named pACrepC.

To evaluate the possible influence of the Plac promoter on the replication properties of the constructs described here, derivatives containing a  $\Omega$ gentamicin resistance interposon ( $\Omega$ -Gm<sup>R</sup>, obtained from plasmid pBSL142; Alexeyev et al., 1995) at the *Xba*I site, located between



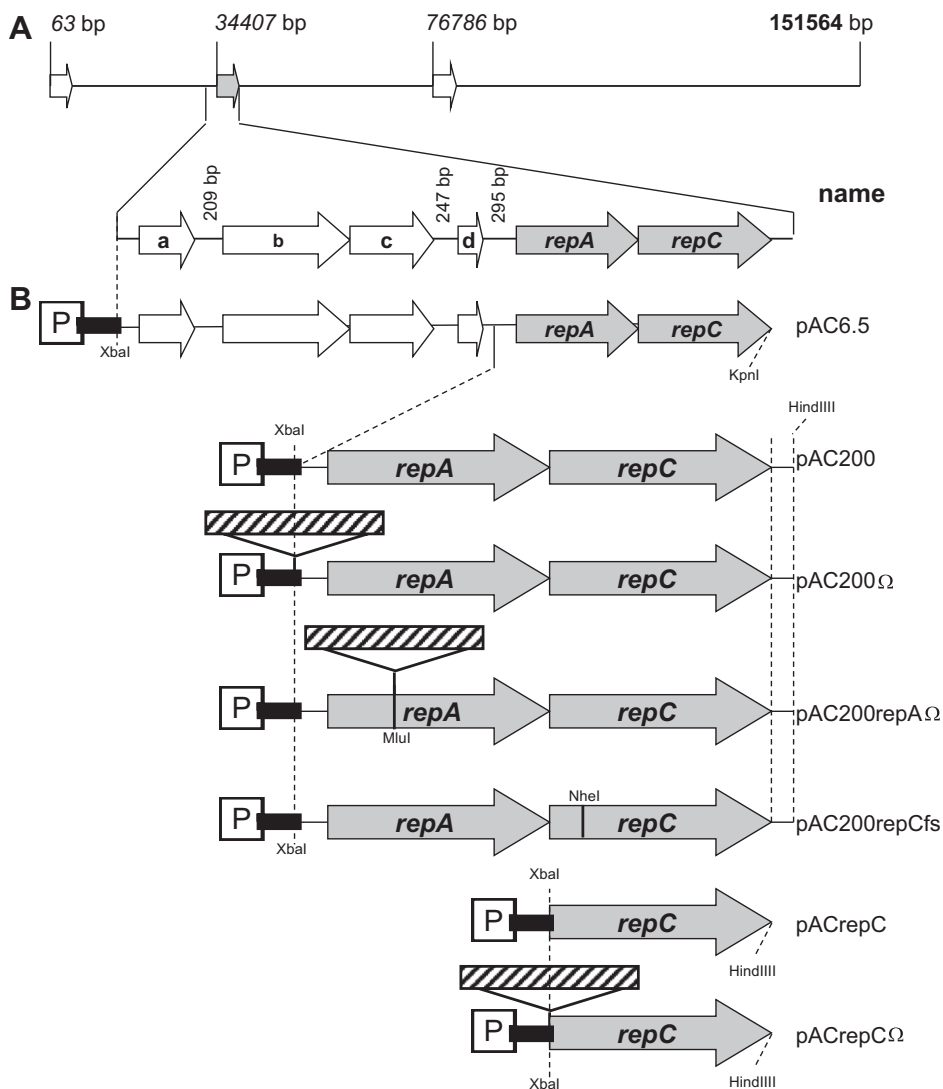
**Fig. 1.** Linear maps of: (A) pDOP (4272 bp) and (B) pDOABTi (6256 bp). Both plasmids contain the transfer origin of RK2 plasmid (*oriT*); a gene conferring chloramphenicol resistance (*cat*); a ColE1 type origin of replication; a gene encoding the *lacZ*  $\alpha$  peptide. M represents the position of the *Mlu*I site, and N, the position of a *Nsi*I restriction site. B marks the positions of *Bam*HI/*Bgl*III chimeric sites. pDO contains the RK2 stabilization region (*parDE*) and pDOABTi possess the segregation system of the *A. tumefaciens* pTiC58 plasmid (*repA* and *repB*). The white circle between *repA* and *repB* indicates the position of a gene encoding a small peptide (*repD*), which contains a centromeric-like region, *parS*. Vectors contain a Multi Cloning Site region within the gene encoding the *lacZ*  $\alpha$  peptide (dashed rectangle) with the following restriction sites: SacI, SacII, BstXI, EagI, NotI, XbaI, SpeI, BamHI, SmaI, PstI, EcoR1, EcoRV, HindIII, Clal, Sall, XhoI, ApaI and KpnI. Underlined restriction sites are not unique in pDOABTi.

the Plac promoter and the insert, were obtained (Fig. 2, Table 2). The pAC200  $\Omega$  plasmid was obtained by introducing a  $\Omega$ -spectinomycin resistance interposon ( $\Omega$ -Sp<sup>R</sup>, obtained from the pBSL175 plasmid; Alexeyev et al., 1995) into the *Mlu*I site present in the *repA* gene of plasmid pAC200 (Table 2). To evaluate the role of RepC, a frame-shift mutation was introduced at the *Nhe*I restriction site located at positions 171–175 bp within the *repC* gene in the pAC200 construct. To this end, the pAC200 construct was digested with *Nhe*I, and the protruding 5' ends were filled with *Taq* polymerase High Fidelity (Invitrogen). The

reaction mixture contained dNTPs and the buffer, provided by the manufacturer, and the reaction ran for 1 h at 72 °C. The obtained products were ligated and used to transform *E. coli*. The resulting construct was named pAC200repCfs. The constructs used in this work are listed in Table 2.

### 2.5. Construction of *repAC::gusA* fusions

Two *repAC::gusA* fusions were constructed: the first *gusA* fusion was obtained after amplifying by PCR a 2760 bp region that embrace the *repA* gene, including



**Fig. 2.** (A) Linear representation of the *R. leguminosarum* 3841 pRL7 plasmids indicating the relative positions of the *repABC* operons (white arrows) and the *repAC* genes (gray arrows). Numbers in italics indicate the position of the initiation codons of the *repA* genes as annotated in GenBank (NC\_008382). The number in bold at the right indicates the total length of the pRL7 plasmid. Below, a magnification of the 6.5-kb region containing the *repAC* genes in which the white arrows show the relative positions of the genes located upstream of the *repAC* genes and transcribed in the same direction as the *repAC* genes (gray arrows). (B) Graphic representation of the gene content of the constructs utilized in this work. Gray arrows represent the *repAC* genes, and white arrows indicate the genes located upstream of *repA*: (a) pRL70033, hypothetical protein; (b) pRL70034, similar to IstA IS21 transposase family; (c) pRL70036, similar to IstB IS21 transposase family; (d) pRL70037, hypothetical protein. Numbers between white arrows show sizes of intergenic regions. Boxed P shows the relative position of the vector's Plac promoter. Dashed boxes indicate the positions of  $\Omega$ -cassettes and the resistance they confer:  $\Omega$ -Gm<sup>R</sup>, gentamicin resistant; and  $\Omega$ -Sp<sup>R</sup>, spectinomycin resistant. The black line within the *repC* gene of the pAC200repCfs construct indicates the relative position of the *Nhe*I restriction site used to introduce a frame-shift mutation.

**Table 2**

Vectors and constructs used in this work.

Plasmid	Relevant characteristics	Source or reference
pCR2.1-TOPO pBSL142	Cloning vector, Amp <sup>R</sup> . Ω-gentamycin donor Amp <sup>R</sup> , Gm <sup>R</sup> .	Invitrogene Alexeyev et al. (1995).
pBSL175	Ω-spectinomycin donor Amp <sup>R</sup> , Spc <sup>R</sup> .	Alexeyev et al. (1995).
pBBR1MCS-5	Gm <sup>R</sup> cloning vector replicable in <i>Rhizobium</i> .	Kovach et al. (1995).
pBBMCS53	pBBR1MCS-5 derivative carrying a promoterless β-glucuronidase gene, Gm <sup>R</sup> .	Corvera et al. (1999).
pRK415	Tc <sup>R</sup> cloning vector replicable in <i>Rhizobium</i> and compatible with pBBMCS53.	Keen et al. (1988)
pDOP	Suicide vector, <i>oriT</i> , <i>parDE</i> , Cm <sup>R</sup> .	This work
pDOABTi	Suicide vector, <i>oriT</i> , <i>repADB</i> , Cm <sup>R</sup> .	This work
pAC6.5	pDOP derivative that contains a DNA fragment spanning from position 30310–36801. Embraces genes with loci tags: RL70033, RL70034, RL70036, RL70037, RL70038 ( <i>repA</i> ), RL70039 ( <i>repC</i> ).	This work
pAC200	pDOP derivative containing the <i>repAC</i> genes plus 203 bp upstream <i>repA</i> initiation codon and 159 bp downstream the <i>repC</i> stop codon.	This work
pDOABTi200	pDOABTi derivative containing same insert of construct pAC200.	This work
pAC200Ω	Same as pAC200 but containing an Ω-gentamycin interposon in the <i>XbaI</i> site.	This work
pAC200repAΩ	Same as pAC200 but containing an Ω-spectinomycin interposon in the <i>MluI</i> site located within <i>repA</i> .	This work
pACrepC	pDOP derivative containing the <i>repC</i> gene and a Shine-Dalgarno sequence 6 pb upstream the initiation codon of <i>repC</i> .	This work
pACrepCΩ	Same as pACrepC but containing an Ω-gentamycin interposon in the <i>XbaI</i> site.	This work
pGUSACSal	pBBMCS53 derivative carrying a <i>repAC::gusA</i> fusion.	This work
pGUSACSalΩSp <sup>R</sup>	Similar to pGUSACSal but containing an Ω-spectinomycin interposon within the <i>repA</i> coding region.	This work
pGUSACMlu	pBBMCS53 derivative carrying the last 458 pb of the <i>repA</i> gene (starting from the <i>MluI</i> site located within <i>repA</i> ) and the first 1080 pb of <i>repC</i> fused to a <i>gusA</i> promoterless gene.	This work
pRK415repAC	pRK415 derivative, containing <i>repA</i> , 206 bp of its upstream region, <i>repC</i> and 159 bp of its downstream region.	This work
pRK415repA	pRK415 derivative, containing the complete <i>repA</i> gene and 206 bp of its upstream region.	This work

206 bp upstream its initiation codon, and the first 1080 bp of the *repC* gene. The PCR product, made with primers repACXba-U and repACSal-L and *R. leguminosarum* 3841 total DNA as template, was cloned into plasmid pCR2.1-TOPO using TOPO-TA cloning kit (Invitrogen) to generate plasmid pTOPOACgus. Taking into account that the primers introduce an *XbaI* site at one end of the amplification product and a *Sall* site at the other, the pTOPOACgus insert was excised with *XbaI* and *Sall* and introduced into plasmid pBBMCS53 (Corvera et al., 1999), in the appropriate orientation, upstream its *gusA* promoter-less gene, to produce plasmid pGUSACSal. To assemble the second *gusA* fusion a PCR product containing the last 458 bp of the *repA* gene (starting from the *MluI* site located within *repA*) and the first 1080 bp of *repC* was obtained with primers repCmlu-U and repACSal-L, and genomic DNA of strain 3841 as template. The amplification product was cloned into plasmid pJET1.2 (Fermentas) following the recommendations included in the CloneJET kit produced by Fermentas to obtain plasmid pJETCmlu. The insert of this plasmid was removed with *XbaI* and *Sall* and inserted into plasmid pBBMCS53 upstream its *gusA* promoter-less gene, to produce plasmid pGUSACMlu. Also, a pGUSACSal derivative named pGUSACSalΩSpc<sup>R</sup> was obtained introducing a Ω-Spc<sup>R</sup> cassette into the *MluI* restriction site located within the *repA* coding region. Plasmid pBSL175 was the donor of the Ω-Spc<sup>R</sup> cassette. Constructs were introduced into *R. etli* CFNX107 by mating and their β-glucuronidase activities were measured as described below.

## 2.6. Construction of pRK415 derivatives

Plasmid pRK415 is a broad-host-range vector with a low plasmid copy-number (1–4 copies per cell) compatible with plasmid pBBMCS53. This vector was used to construct two derivatives: the first one, pRK415repAC, contains an insert embracing a 2,760-bp DNA region that includes 206 bp of upstream of the *repA* initiation codon and ended 159 bp downstream of the *repC* stop codon. The insert was obtained by PCR, using primers repACHindIII-U and repACKpn-L and plasmid pAC6.5 as template. Insert of pRK415repAC is almost identical to the insert present in construct pAC200, only differing in the restriction sites used in the cloning process. The second construct, pRK415repA, possesses a 1397 bp insert, containing the complete *repA* gene and 206 bp of its upstream region. This DNA region was obtained by PCR with primers repACHindIII-U and repAKpn-L and plasmid pAC6.5 as template. In both cases plasmid pJET1.2 was used as intermediary vector during the PCR cloning processes (CloneJET kit, Fermentas). The two derivatives were introduced into *R. etli* CFNX107/pGUSACSal by conjugation.

## 2.7. β-glucuronidase activity measurements

Overnight cultures of the strains carrying *gusA* fusions were grown in the presence of gentamicin. Cultures were diluted in fresh PY medium to OD<sub>540</sub> 0.2 and were grown to a final OD<sub>540</sub> 0.4. One milliliter of this culture was cen-

trifuged and resuspended in a salt wash solution (0.4 mM MgSO<sub>4</sub> / 1.2 mM K<sub>2</sub>PO<sub>4</sub>) supplemented with chloramphenicol (100 µg/ml). Quantitative β-glucuronidase assays were performed with p-nitrophenyl glucuronide as substrate as described previously (Wilson et al., 1992). Data were normalized to OD of the culture. The results presented here are the mean of three independent experiments.

### 2.8. Bacterial mating

pDOABTi200 and pDOP derivatives were introduced into *R. etli* CFNX107 using *E. coli* S17–1 as the donor strain (Simon et al., 1983). Strains were grown in antibiotic-free liquid medium until the stationary phase was reached. The two strains were then mixed at a donor-to-recipient ratio of 1:1 (vol) on antibiotic-free PY plates (donor: 9 × 10<sup>7</sup> cells/receptor: 2 × 10<sup>8</sup>), and the plates were incubated at 30 °C overnight. Cells were resuspended in PY medium, and serial dilutions were plated on PY plus nalidixic acid and chloramphenicol.

### 2.9. Plasmid profiles

The plasmid profiles of five transconjugants of each cross were visualized on agarose gels according to the protocol reported by Hynes and McGregor (1990).

### 2.10. Plasmid stability

To estimate the stability of the plasmid constructs, stationary phase cultures, grown in the presence of chloramphenicol, were washed and then diluted 100-fold in fresh medium without selection for the plasmid to an initial OD of 0.2 at 620 nm. Cultures were incubated with agitation (250 rpm), for 24 h at 30 °C. Samples taken at 0, 6, 8, 10, and 24 h were serially diluted and inoculated onto plates in the absence of selective drugs. One hundred colonies from each time point were chosen and poked onto plates with and without the selective antibiotic. Loss rates (percent per generation) were calculated by the following equation:  $[1 - (F/I)^{1/N}] \times 100\%$ , where *I* is the initial percentage of plasmid-containing cells and *F* is the percentage of plasmid-containing cells after *N* generations.

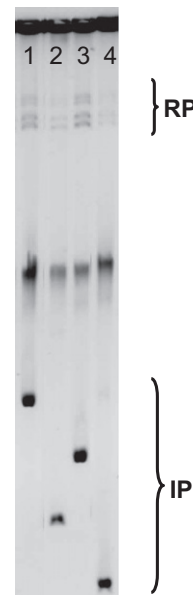
## 3. Results and discussion

*R. leguminosarum* bv. *viciae* 3841 is able to fix nitrogen in symbiosis with a restricted number of legumes of the tribe Viciae such as lentils, peas, vetches and *Lathyrus* spp. The genome of *R. leguminosarum* bv. *viciae* 3841 comprises a chromosome of 5.05 Mb and six large plasmids ranging between 147 and 870 kb in size, which all contain *repABC* operons. Notably, plasmid pRL7 contains two *repABC* operons that have different sequences and a third truncated operon; these three operons are encoded by the same DNA strand in the first half of the plasmid (Young et al., 2006).

The truncated operon contains the *repA* and *repC* genes without traces of *repB* or the cTRNA, and is positioned between the complete *repABC* operons (Fig. 2A). For this rea-

son, from here on we will refer to this truncated version as *repAC*. The *repAC* genes are the last two genes of a six-gene cluster encoded by the same DNA strand, suggesting that the *repAC* genes are part of a larger transcription unit.

To determine whether the *R. leguminosarum* 3841 *repAC* genes are capable of allowing replication, two different PCR products containing the *repAC* genes were obtained and cloned into a vector unable to replicate in *Rhizobium* (pDOP). Both amplification products were cloned in the same direction and were under the influence of the Plac promoter present in the vector. The first construct, pAC6.5, had a 6.64-kb insert and included the cluster of six genes encoded in the same direction described above. The second construct, pAC200, contained complete *repAC* genes plus 206 bp upstream of the *repA* initiation codon and 159 bp downstream of the *repC* stop codon (Fig. 2B). These constructs were introduced by conjugation into *R. etli* CFNX107, a species closely related to *R. leguminosarum* that contains plasmids with *repABC* operons with enough variability to avoid incompatibility problems. Transconjugants were obtained for both constructs at approximately the same frequency (2 × 10<sup>-3</sup>). Moreover, the plasmid profile analysis of at least four transconjugants of each cross, obtained without further purification passages, indicated that the constructs were capable of replicating in *R. etli* CFNX107 (Fig. 3, lanes 1 and 2). These results show that genes present in the insert in the pAC200 construct were sufficient for replication. However, to discard the possible transcriptional influence of the Plac promoter of the vector on the genes present in the insert, a Ω-Gm<sup>R</sup> interposon was introduced between the insert and the Plac promoter of the pAC200 construct to generate the plasmid pAC200 Ω



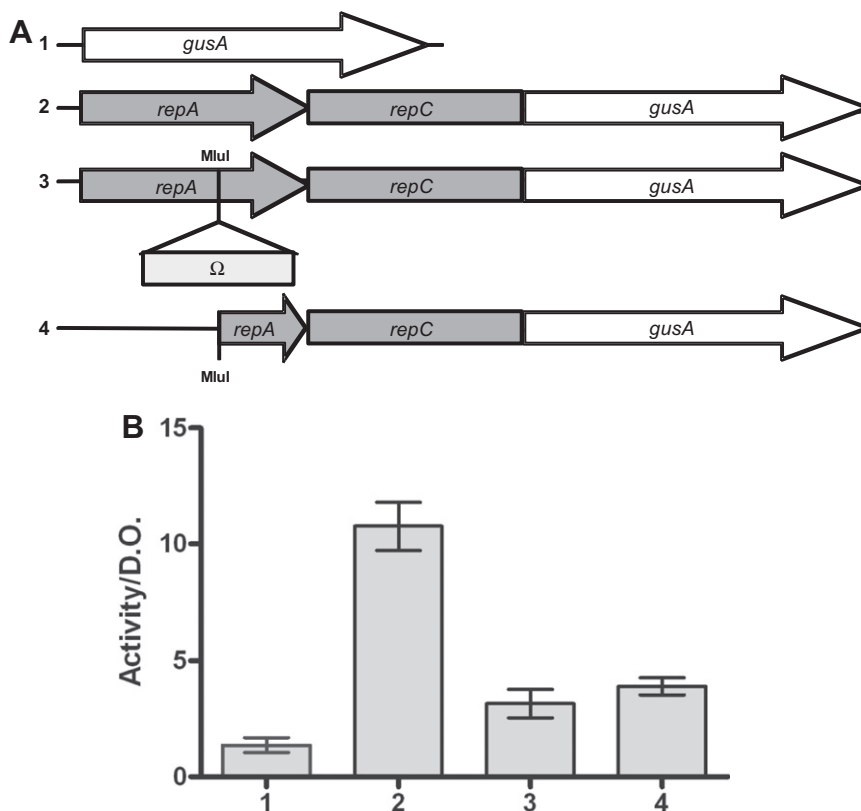
**Fig. 3.** Plasmid profiles (negative image) of *R. etli* CFNX107 transconjugants carrying the following plasmids: pAC6.5 [10.9 Kb] (lane 1), pAC200 [7.0 Kb] (lane 2), pAC200 Ω [8.6 Kb] (lane 3), and pACrepC [5.5 Kb] (lane 4). IP shows the positions of the incoming plasmids, and RP indicates the location of the resident plasmids present in the *R. etli* CFNX107 strain.

(Figs. 2B and 3, lane 3). This plasmid was conjugated into *R. etli* CFNX107, and the frequency of the transconjugants in addition to the plasmid profiles obtained for a sample of the transconjugants indicate that the transconjugants behave in the same manner as the parental construct. This result indicates that the plasmid pAC200 insert contains all of the elements required to replicate as an independent entity in *Rhizobium*, including a cognate promoter.

RepC is essential for the replication of *repABC* plasmids (Bartosik et al., 1998; Ramírez-Romero et al., 2000; Pappas and Winans, 2003). To determine if RepC role is similar in *repAC*, a frame-shift mutation in the *repC* gene of the plasmid pAC200 was obtained as described in the Material and Methods section. The derivative construct, pAC200repCfs, was conjugated to *R. etli* CFNX107; however, transconjugants were not obtained, indicating that RepC plays a crucial role in the replicability of the plasmid pAC200. Moreover, pACrepC, a construct that contains only *repC* with a Shine–Dalgarno sequence that we introduced at the appropriate distance and under the control of the Plac promoter located in the pDOP vector, was capable of replicating by itself in the *R. etli* CFNX107 genetic background (Figs. 2B and 3, lane 4). In contrast, a derivative of this construct containing a  $\Omega$ -Gm<sup>R</sup> interposon between *repC* and the Plac vector was unable to replicate. These observations demonstrate that RepC is the minimal element required for

independent replication and that the origin of replication (*oriV*) resides within *repC*, as has been shown for other *repABC* plasmids (Cervantes-Rivera et al., 2011; Pinto et al., 2011).

To determine if the *repAC* genes are organized in a single transcription unit, a PCR product containing the *repA* gene, its promoter and a fragment of *repC* (1080 bp) was cloned upstream of the promoterless  $\beta$ -glucuronidase gene (*gusA*) of plasmid pBBMCS53, to produce plasmid pGUSACSal. If *repA* and *repC* are organized in an operon, an insertion of a  $\Omega$  interposon within *repA* will cause a polar effect on *repC* expression. With this idea in mind, plasmid pGUSACSal was modified by inserting a  $\Omega$ -Spc<sup>R</sup> interposon into the *Mlu*I site located 727 bp within the *repA* coding-region, to obtain pGUSACSal $\Omega$ Spc<sup>R</sup> (Fig. 4). Both constructs and the empty vector (pBBMCS53) were introduced into *R. etli* CFNX107 and the  $\beta$ -glucuronidase activity of the transconjugants was determined. As show in Fig. 4B, the insertion of a  $\Omega$ -Spc<sup>R</sup> cassette into *repA* reduces the activity of the reporter gene by 2.8-fold. However, a marginal  $\beta$ -glucuronidase activity remains, suggesting the presence of a low-efficiency promoter upstream of *repC* and within a *repA* region located between the *Mlu*I restriction site and the *repA* stop codon. To explore this possibility, and to discard an unwanted influence of the interposon, a new *gusA* fusion was made by cloning a region that begins



**Fig. 4.** (A) Schematic representation of the *gusA* transcriptional fusions: 1, *gusA* (promoter-less) of pBBMCS53 (empty vector); 2, pGUSACSal; 3, pGUSACSal $\Omega$ Spc<sup>R</sup>; 4, pGUSACMlu. (B)  $\beta$ -glucuronidase activities of *repAC*::*gusA* fusions of strains: 1, CFNX107/ pBBMCS53, 2, CFNX107/pGUSACSal; 3, CFNX107/ pGUSACSal $\Omega$ Spc<sup>R</sup>; 4, CFNX107/ pGUSACMlu.

in the *repA* *Mlu*I restriction site and ends in the position 1080 of the *repC* gene in front of the promoterless *gusA* gene of pBBMCS53. This new construct, named pGUSACM-lu, was introduced into CFNX107. The activity of the reporter gene in the transconjugants was measured and showed the same level of activity determined in strain CFNX107/pGUSACSal (Fig. 4). These results support the presence of a low-efficiency promoter upstream of *repC*.

Nevertheless, to evaluate the functional relevance of the promoter located within *repA* and considering that *repC* is essential for replication, an  $\Omega$ -Spc<sup>R</sup> interposon was placed in the *Mlu*I site of *repA* in the pAC200 plasmid. The resulting construct, pAC200repA  $\Omega$  and the unmodified plasmid were introduced to *R. etli* CFNX107 by conjugation. Transconjugants were obtained only with plasmid pAC200, indicating that the insertion of a  $\Omega$ -cassette within *repA* disrupts *repC* function. Together, these observations indicate that *repA* and *repC* belong to the same transcriptional unit and that the putative weak promoter located within the *repA* coding region seems to be irrelevant for plasmid replication.

RepA plays a central role in the negative autoregulation of *repABC* operons, and its activity is greatly enhanced by RepB and *parS* sequence. To assess the participation of RepA in the negative regulation of the *repAC* operon, the  $\beta$ -glucuronidase activity of the *repA::gusA* fusion of CFNX107/pGUSACSal was evaluated when RepA or RepA and RepC were provided *in trans*. To do that, pRK415repA that express RepA from its cognate promoter, and pRK415repAC that contains the whole *repAC* operon, were introduced individually by conjugation into CFNX107/pGUSACSal strain. The  $\beta$ -glucuronidase activities of a transconjugant of each cross were determined (Fig. 5). Results of such experiments indicate that RepA partially retains the capacity to repress its own expression.

Segregation of *repABC* plasmids depends on the correct functioning of RepA and RepB proteins, and the *parS* site (Bartosik et al., 2001; Soberón et al., 2004). As mentioned above, *repAC* lacks *repB* and the *ctrNA* gene, suggesting that *repAC* is unstable. To determine the stability of plasmid pAC200, the construct containing the minimal region capable of independent replication in *R. etli*, a plasmid retention test was performed. Fig. 6 show that plasmid

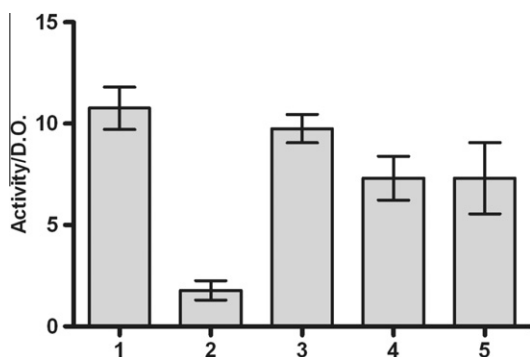


Fig. 5.  $\beta$ -glucuronidase activities of CFNX107 derivatives, carrying: 1, pGUSACSal; 2, pBBMCS53/pRK415 (empty vectors); 3, pGUSACSal/pRK415; 4, pGUSACSal/pRK415A and 5, pGUSACSal/pRK415AC.

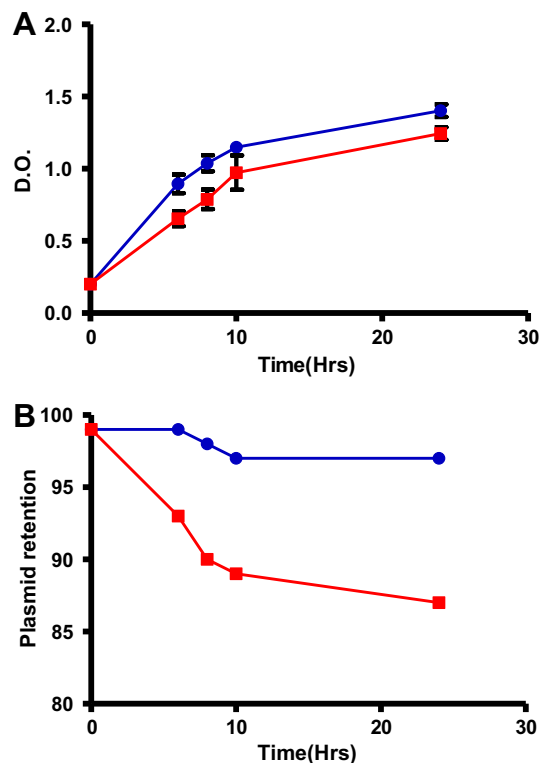


Fig. 6. (A) Growth kinetics of *R. etli* CFNX107/pAC200 (squares), and *R. etli* CFNX107/pDOABTi200 (circles), in PY medium without antibiotics. (B) Plasmid loss rate of *R. etli* CFNX107/pAC200 (squares), and *R. etli* CFNX107/pDOABTi200 (circles).

pAC200 is unstable: a CFNX107 strain carrying pAC200 lost the plasmid at a rate of 0.9% per generation. However, after 10 h cells retaining the plasmid reach a plateau, suggesting a possible integrational rescue. Taking into consideration that plasmid stability also depends on replication efficiency, the insert from plasmid pAC200 was introduced into pDOABTi, a suicide vector containing the segregation system of the *A. tumefaciens* pTiC58 plasmid, to generate construct pDOABTi200. A *R. etli* CFNX107/pDOABTi200 strain had a plasmid loss rate of 0.1% per generation, suggesting that the plasmid instability observed in plasmid pAC200 is mainly related to the absence of an active segregation system and not to its replication abilities.

These observations suggest that *repAC* genes have been retained, in evolutionary terms, in the *R. leguminosarum* 3841 population, relying on the replication and stability properties provided by the *repABC* operons present in plasmid pRL7. However, the ability of *repAC* genes to allow replication may be interpreted as an intermediary stage in the generation and evolution of a new plasmid replication/partitioning system; if the *repAC* genes become closely associated with genes from a functional partitioning system by recombination or transposition, a new functional system may be formed.

In summary, the *repAC* genes present in the *R. leguminosarum* plasmid pRL7 were capable of sustaining replication but in an unstable manner. RepC was essential for replication, and the origin of replication resides within its

coding region. In contrast, RepA probably, only plays a marginal role regulating negatively its transcription. The *repAC* genes may be sources of new plasmids if they acquire a new partitioning system.

## Acknowledgments

This work was supported by the Consejo Nacional de Ciencia y Tecnología [CONACYT, Grant number 100099] and the Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica [PAPIIT, Grant number IN205611-3]. G. P-S was supported during the Ph.D. program (Programa de Doctorado en Ciencias Biomédicas-Universidad Nacional Autónoma de México) by a scholarship from Consejo Nacional de Ciencia y Tecnología. We thank Michael Dunn for his critical review of the manuscript.

## References

- Alexeyev, M.F., Shokolenko, I.N., Croughan, T.P., 1995. Improved antibiotic-resistance gene cassettes and omega elements for *Escherichia coli* vector construction and *in vitro* deletion/insertion mutagenesis. *Gene* 160, 63–67.
- Bartosik, D., Baj, J., Włodarczyk, M., 1998. Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. *Microbiology* 144, 3149–3157.
- Bartosik, D., Szymanik, M., Wysocka, E., 2001. Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J. Bacteriol.* 21, 6234–6243.
- Castillo-Ramírez, S., Vázquez-Castellanos, J.F., González, V., Cevallos, M.A., 2009. Horizontal gene transfer and diverse functional constraints within a common replication-partitioning system in alphaproteobacteria: the *repABC* operon. *BMC Genomics* 10, 536. <http://dx.doi.org/10.1186/1471-2164-10-536>.
- Cervantes-Rivera, R., Pedraza-López, F., Pérez-Segura, G., Cevallos, M.A., 2011. The replication origin of a *repABC* plasmid. *BMC Microbiol.* 11, 158. <http://dx.doi.org/10.1186/1471-2180-11-158>.
- Cervantes-Rivera, R., Romero-López, C., Berzal-Herranz, A., Cevallos, M.A., 2010. Analysis of the mechanism of action of the antisense RNA that controls the replication of the *repABC* plasmid p42d. *J. Bacteriol.* 192, 3268–3278.
- Cevallos, M.A., Cervantes-Rivera, R., Gutiérrez-Ríos, R.M., 2008. The *repABC* plasmid family. *Plasmid* 60, 19–37.
- Chai, Y., Winans, S.C., 2005a. A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol. Microbiol.* 56, 1574–1585.
- Chai, Y., Winans, S.C., 2005b. RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between *repA* and *repB* for plasmid partitioning and autorepression. *Mol. Microbiol.* 58, 1114–1129.
- Corvera, A., Promé, D., Promé, J.C., Martínez-Romero, E., Romero, D., 1999. The *nolL* gene from *Rhizobium etli* determines nodulation efficiency by mediating the acetylation of the fucosyl residue in the nodulation factor. *Mol. Plant Microbe Interact.* 12, 236–246.
- Guiney, D.G., Yakobson, E., 1983. Location and nucleotide sequence of the transfer origin of the broad host range plasmid RK2. *Proc. Natl. Acad. Sci. USA* 80, 3595–3598.
- Guo, X., Flores, M., Mavingui, P., Fuentes, S.I., Hernández, G., Dávila, G., Palacios, R., 2003. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Res.* 13, 1810–1817.
- Hanahan, D., 1983. Studies of transformation of *E. coli* with plasmids. *J. Mol. Biol.* 166, 557–560.
- Hynes, M.F., McGregor, N.F., 1990. Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. *Mol. Microbiol.* 4, 567–574.
- Izquierdo, J., Venkova-Canova, T., Ramírez-Romero, M.A., Téllez-Sosa, J., Hernández-Lucas, I., Sanjuan, J., Cevallos, M.A., 2005. An antisense RNA plays a central role in the replication control of a *repC* plasmid. *Plasmid* 54, 259–277.
- Johnston, A.W.B., Beringer, J.E., 1975. Identification of rhizobium strains in pea root nodules using genetic markers. *J. Gen. Microbiol.* 87, 343–350.
- Keen, N.T., Tamaki, S., Kobayashi, D., Trollinger, D., 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene* 70, 191–197.
- Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., Peterson, K.M., 1995. 2nd Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176.
- MacLellan, S.R., Smallbone, L.A., Sibley, C.D., Finan, T.M., 2005. The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. *Mol. Microbiol.* 55, 611–623.
- Martínez-Salazar, J., Romero, D., Girard, M.L., Dávila, G., 1991. Molecular cloning and characterization of the *recA* gene of *Rhizobium phaseoli* and construction of *recA* mutants. *J. Bacteriol.* 173, 3035–3040.
- Mavingui, P., Flores, M., Guo, X., Dávila, G., Perret, X., Broughton, W.J., Palacios, R., 2002. Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. *J. Bacteriol.* 184, 171–176.
- Noel, K.D., Sánchez, A., Fernández, L., Leemans, J., Cevallos, M.A., 1984. *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J. Bacteriol.* 158, 148–155.
- Pappas, K.M., 2008. Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. *Plasmid* 60, 89–107.
- Pappas, K.M., Winans, S.C., 2003. The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. *Mol. Microbiol.* 49, 441–455.
- Petersen, J., Brinkmann, H., Pradella, S., 2009. Diversity and evolution of *repABC* type plasmids in rhodobacterales. *Environ. Microbiol.* <http://dx.doi.org/10.1111/j.1462-2920.2009.01987.x>.
- Pinto, U.M., Flores-Mireles, A.L., Costa, E.D., Winans, S.C., 2011. RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within *repC* and functions only in *cis*. *Mol. Microbiol.* 81, 1593–1606. <http://dx.doi.org/10.1111/j.1365-2958.2011.07789.x>.
- Ramírez-Romero, M.A., Téllez-Sosa, J., Barrios, H., Pérez-Oseguera, A., Rosas, V., Cevallos, M.A., 2001. RepA negatively autoregulates the transcription of the *repABC* operon of the *Rhizobium etli* symbiotic plasmid basic replicon. *Mol. Microbiol.* 42, 195–204.
- Ramírez-Romero, M.A., Soberón, N., Pérez-Oseguera, A., Téllez-Sosa, J., Cevallos, M.A., 2000. Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. *J. Bacteriol.* 182, 3117–3124.
- Roberts, R.C., Helinski, D.R., 1992. Definition of a minimal plasmid stabilization system from the broad-host-range plasmid RK2. *J. Bacteriol.* 174, 8119–8132.
- Simon, R., Priefer, U., Pühler, A., 1983. A broad host-range mobilization system for *in vivo* genetic engineering: transposon mutagenesis Gram-negative bacteria. *Bio/Technology* 1, 784–791.
- Slater, S.C., Goldman, B.S., Goodner, B., Setubal, J.C., Farrand, S.K., Nester, E.W., Burr, T.J., Banta, L., Dickerman, A.W., Paulsen, I., Otten, L., Suen, G., Welch, R., Almeida, N.F., Arnold, F., Burton, O.T., Du, Z., Ewing, A., Gotsy, E., Heisel, S., Houmiel, K.L., Jhaveri, J., Lu, J., Miller, N.M., Norton, S., Chen, Q., Phoolcharoen, W., Ohlin, V., Ondrusek, D., Pride, N., Stricklin, S.L., Sun, J., Wheeler, C., Wilson, L., Zhu, H., Wood, D.W., 2009. Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *J. Bacteriol.* 191, 2501–2511.
- Soberón, N., Venkova-Canova, T., Ramírez-Romero, M.A., Téllez-Sosa, J., Cevallos, M.A., 2004. Incompatibility and the partitioning site of the *repABC* basic replicon of the symbiotic plasmid from *Rhizobium etli*. *Plasmid* 51, 203–216.
- Venkova-Canova, T., Soberón, N.E., Ramírez-Romero, M.A., Cevallos, M.A., 2004. Two discrete elements are required for the replication of a *repABC* plasmid: an antisense RNA and a stem-loop structure. *Mol. Microbiol.* 54, 1431–1444.
- Wilson, K.J., Huges, S.G., Jefferson, R.A., 1992. The *Escherichia coli* *gus* operon, induction and expression of the *gus* operon in *E. coli* and the occurrence and use of GUS in other bacteria. In: Gallagher, S.R. (Ed.), *GUS Protocols, using the GUS Gene as a Reporter of Gene Expression*. Academic Press, San Diego, CA, pp. 7–23.
- Young, J.P., Crossman, L.C., Johnston, A.W., Thomson, N.R., Ghazoui, Z.F., Hull, K.H., Wexler, M., Curson, A.R., Todd, J.D., Poole, P.S., Mauchline, T.H., East, A.K., Quail, M.A., Churcher, C., Arrowsmith, C., Cherevach, I., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabinowitz, E., Sanders, M., Simmonds, M., Whitehead, S., Parkhill, J., 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol.* 7, R34.



RESEARCH ARTICLE

Open Access

# The replication origin of a repABC plasmid

Ramón Cervantes-Rivera, Francisco Pedraza-López, Gabriela Pérez-Segura and Miguel A Cevallos\*

## Abstract

**Background:** *repABC* operons are present on large, low copy-number plasmids and on some secondary chromosomes in at least 19  $\alpha$ -proteobacterial genera, and are responsible for the replication and segregation properties of these replicons. These operons consist, with some variations, of three genes: *repA*, *repB*, and *repC*. RepA and RepB are involved in plasmid partitioning and in the negative regulation of their own transcription, and RepC is the limiting factor for replication. An antisense RNA encoded between the *repB-repC* genes modulates *repC* expression.

**Results:** To identify the minimal region of the *Rhizobium etli* p42d plasmid that is capable of autonomous replication, we amplified different regions of the *repABC* operon using PCR and cloned the regions into a suicide vector. The resulting vectors were then introduced into *R. etli* strains that did or did not contain p42d. The minimal replicon consisted of a *repC* open reading frame under the control of a constitutive promoter with a Shine-Dalgarno sequence that we designed. A sequence analysis of *repC* revealed the presence of a large A+T-rich region but no iterons or DnaA boxes. Silent mutations that modified the A+T content of this region eliminated the replication capability of the plasmid. The minimal replicon could not be introduced into *R. etli* strain containing p42d, but similar constructs that carried *repC* from *Sinorhizobium meliloti* pSymA or the linear chromosome of *Agrobacterium tumefaciens* replicated in the presence or absence of p42d, indicating that RepC is an incompatibility factor. A hybrid gene construct expressing a RepC protein with the first 362 amino acid residues from p42d RepC and the last 39 amino acid residues of RepC from SymA was able to replicate in the presence of p42d.

**Conclusions:** RepC is the only element encoded in the *repABC* operon of the *R. etli* p42d plasmid that is necessary and sufficient for plasmid replication and is probably the initiator protein. The *oriV* of this plasmid resides within the *repC* gene and is located close to or inside of a large A+T region. RepC can act as an incompatibility factor, and the last 39 amino acid residues of the carboxy-terminal region of this protein are involved in promoting this phenotype.

## Background

Proteins that are involved in the initiation of DNA replication are essential to cells. These proteins recognize the origin of replication, destabilize double-stranded DNA, and recruit the replisome, which is the machinery directly involved in DNA replication [1].

Both the activity and concentration of the initiator proteins are highly regulated because the genetic material needs to be replicated only once per generation. A failure in this process could accelerate the production of new DNA molecules with a concomitant increase in the number of new origins of replication, which could be

used in new rounds of replication and leading to cell death (i.e., “runaway replication”) [2].

Initiator proteins control the replication rate using several mechanisms that limit either their own synthesis or their availability. The initiator proteins can directly auto-regulate the transcription of their own genes or trigger the production of negative regulators, antisense-RNAs or proteins, which are co-transcribed with the initiator genes. The activity of the initiator proteins can be controlled by covalent modifications or by titrating out their availability using DNA sites that resemble origins of replication. In addition, the DNA initiation rate can be controlled by blocking or hiding the origins of replication [3,4].

The initiation of replication of the *Escherichia coli* chromosome and of some of its plasmids has been

\* Correspondence: mac@ccg.unam.mx  
Programa de Genómica Evolutiva, Centro de Ciencias Genómicas,  
Universidad Nacional Autónoma de México, Apartado Postal 565-A,  
Cuernavaca, Morelos, México

studied extensively. However, our knowledge of other bacterial replication systems is limited. Research on new replicons that are not found in *E. coli* or its close relatives would yield new insights into the regulation of initiation of replication in bacteria. The present work concerns *repABC* replicons, which are present on large, low copy-number plasmids and on some secondary chromosomes in at least 19  $\alpha$ -proteobacterial genera. Some bacterial strains contain more than one *repABC* replicon, indicating that this plasmid family encompasses several incompatibility groups [5-7].

The basic replicon of *repABC* plasmids is compact because all of the elements required for replication and segregation are encoded in a single operon, the *repABC* operon [8,9]. However, this operon is controlled by a complex regulatory mechanism. The first two genes of the *repABC* operon encode for proteins belonging to a type Ia segregation system [10]. RepA and RepB have been implicated in the negative transcriptional regulation of the *repABC* operon [9,11].

RepC is a limiting replication factor and thus has been suggested to be the initiator protein [8,12,13]. The members of the *repABC* family contain a centromeric-like sequence (*parS*) in three possible locations: downstream of and close to the stop codon of *repC* [14,15], between *repA* and *repB*, or upstream of *repA* [16,17]. A conserved sequence between the *repB* and *repC* genes is present in all known *repABC* replicons and contains an antisense RNA (ctRNA) gene, the product of which negatively modulates the expression of RepC [18-20]. Regulatory role of the ctRNA depends on its pairing with the *repABC* mRNA. In the absence of the ctRNA, the mRNA section corresponding to the *repB-repC* intergenic region folds into a large stem-loop structure so that the predicted *repC* Shine-Dalgarno (SD) sequence and the *repC* initiation codon remain single-stranded, allowing *repC* translation. In contrast, when the ctRNA hybridizes with the *repABC* mRNA, the *repC* leader sequence forms an intrinsic terminator, blocking *repC* transcription [21].

Many aspects of the biology of these plasmids remain unknown, especially the details of the replication or segregation of these genetic elements. In this paper, we demonstrate the following: A) RepC is the only element encoded in the *repABC* operon of the *Rhizobium etli* p42d plasmid (formally pRetCFN42d) that is necessary and sufficient for plasmid replication. B) RepC is an incompatibility factor. C) The RepC carboxy-terminal region is involved in the incompatibility phenotype. D) The origin of replication of the *repABC* plasmid resides in a large A+T-rich region located at the central section of the *repC* gene.

## Methods

### Plasmids, bacterial strains and growth conditions

The bacterial strains and the plasmids used in this work are described in Table 1. *E. coli* strains were grown at 37°C in Luria-Bertani medium. *Rhizobium* strains were grown at 30°C in PY medium supplemented with 1 mM CaCl<sub>2</sub> [22]. Nalidixic acid (20 µg/ml) and chloramphenicol (30 µg/ml) were added when required. Growth kinetics were made in 500 ml flasks containing, 50 ml of PY medium without antibiotics. Incubation was performed at 30°C and 250 rpm.

### Bacterial mating

pDOP derivatives were introduced into *Rhizobium* by conjugation using *E. coli* S17-1 as a donor strain [23]. The strains were grown in the proper antibiotic-free liquid medium to stationary phase, mixed in a donor-recipient ratio of 1:2 on antibiotic-free PY plates, and incubated at 30°C overnight. The cells were resuspended in PY medium, and serial dilutions were plated on the appropriate selective PY medium.

### Constructions

The primers used in this work are presented in Table 2. The plasmid pDOP-H3 replicon was obtained by cloning a 5.6 Kb *Hind*III fragment containing the complete *repABC* operon from pH3 into the *Hind*III site of pDOP [24]. The inserts of the plasmids used in this work, unless otherwise indicated, were PCR amplified from *repABC* operon cloned in pH3 [24]. Inserts were introduced into pDOP in the sense orientation with respect to the promoter, using restriction sites that were included in the primer sequences.

### Constructs used to determine the minimal replicon

Insert of plasmid pDOP- $\alpha$ C was generated by amplifying the *inca-repC* region with the primers ALFAU2 and Mal-C2Kpn. The *repC* (p42d) gene present in pDOP-C was amplified by PCR with the primers RBS-C and Mal-C2. The *repC* gene of pSymA, present in construct pDOP-CsA, was obtained by PCR with the primers C-SymA and K-Syma-L and the genomic DNA of *S. meliloti* 2011 as the template. The *repC* gene of the linear chromosome of *Agrobacterium tumefaciens* C58 was obtained by PCR with the primers repCAT-BamU and repCATHinL and genomic DNA as the template.

### Constructs carrying *repC* deletions or *repC* fragments

The insert of the plasmid pDOP-C/D1UM with a deletion in its 5'-end was obtained with the oligonucleotides repC-D1U and Mal-C2. The *repC* gene present in the plasmid pDOP-C/RD1L was amplified with the primers RBS-C and repC-D1L. Six plasmids carrying fragments of the *repC* gene were constructed: pDOP-C/F1 insert

**Table 1 Bacterial strains and plasmid used in this work**

Strain	Relevant characteristics	Reference
<i>Rhizobium etli</i> CE3	Streptomycin resistant derivative of CFN42 strain	[20]
<i>R. etli</i> CFNX101	<i>recA</i> :: $\Omega$ -Spectinomycin derivative of CE3	[46]
<i>R. etli</i> CFNX107	<i>recA</i> :: $\Omega$ -Spectinomycin derivative of CE3, lacking plasmid p42a and p42d.	[46]
<i>E. coli</i> S17-1	Plasmid donor in conjugations	[23]
Plasmid	Relevant characteristics	Reference
pDOP	A chloramphenicol resistant suicide vector derived from pBC SK(+), and containing <i>oriT</i>	This work
pDOP-E'	pDOP derivative with the intergenic region <i>repB-repC</i> , the complete <i>repC</i> gene under Plac promoter, and 500 pb downstream <i>repC</i> stop codon.	[22]
pDOP-H3	pDOP derivative carrying a 5.6 Kb <i>HindIII</i> with <i>repABC</i> operon of <i>R. etli</i> plasmid p42d.	This work
pDOP- $\alpha$ C	pDOP derivative with the intergenic region <i>repB-repC</i> and the complete <i>repC</i> gene under Plac promoter.	This work
pDOP-C	pDOP carrying <i>repC</i> gen of plasmid p42d, with a SD sequence (AGGA) and under Plac promoter.	This work
pDOP-C/D1UM	Similar to pDOP-C but with a <i>repC</i> gene carrying a deletion from codon 2 to codon 29	This work
pDOP-C/RD1L	Similar to pDOP-C but with a <i>repC</i> gene carrying a deletion from codon 372 to codon 401	This work
pDOP-F1	pDOP containing a <i>repC</i> fragment from codon 2 to codon 110, with a SD consensus sequence under Plac promoter.	This work
pDOP-C/F1-F2	pDOP containing a <i>repC</i> fragment from codon 2 to codon 209, with a SD consensus sequence under Plac promoter.	This work
pDOP-C/F1-F3	pDOP containing a <i>repC</i> fragment from codon 2 to codon 309, with a SD consensus sequence under Plac promoter.	This work
pDOP-C/F4	pDOP containing a <i>repC</i> fragment from codon 310 to codon 403, with a SD consensus sequence under Plac promoter.	This work
pDOP-C/F4-F3	pDOP containing a <i>repC</i> fragment from codon 210 to codon 403, with a SD consensus sequence under Plac promoter.	This work
pDOP-C/F4-F2	pDOP containing a <i>repC</i> fragment from codon 111 to codon 403, with a SD consensus sequence under Plac promoter.	This work
pDOP-C s/SD	Similar to pDOP-C but without the SD sequence	This work
pDOP-TtMC	Similar to pDOP-C but with a mutant <i>repC</i> gene carrying silent mutations to increase its CG content	This work
pDOP-CBbgIII	Similar to pDOP-C but with <i>repC</i> gene, carrying a frameshift mutation at the <i>BglIII</i> restriction site	This work
pDOP-CSphI	Similar to pDOP-C but with <i>repC</i> gene, carrying a frameshift mutation at the <i>SphI</i> restriction site	This work
pDOP-CA <sub>t</sub> LC	pDOP derivative carrying <i>repC</i> gen of the <i>Agrobacterium tumefaciens</i> C58 linear chromosome, with a SD sequence (AGGA) and under Plac promoter.	This work
pDOP-CsA	pDOP derivative carrying <i>repC</i> gen of the <i>Sinorhizobium meliloti</i> 1021 pSymA, with a SD sequence (AGGA) and under Plac promoter.	This work
pDOP/C420-1209	pDOP with a hybrid <i>repC</i> gene, encoding the first 140 amino acid residues of the pSymA RepC protein and the rest of p42d.	This work
pDOP/C1-420	pDOP with a hybrid <i>repC</i> gene, encoding the first 140 amino acid residues of the p42d RepC protein and the rest of pSymA.	This work
pDOP/C421-840	pDOP with a hybrid <i>repC</i> gene encoding, the first 140 amino acid residues of the pSymA RepC protein, the next 140 amino acid residues from the p42d RepC protein and the rest from the pSymA RepC protein.	This work
pDOP/Cs421-840	pDOP with a hybrid <i>repC</i> gene, encoding the first 140 amino acid residues of the p42d RepC protein, the next 140 amino acid residues from the pSymA RepC protein and the rest from the p42d RepC protein.	This work
pDOP/C841-1209	pDOP derivative with a hybrid <i>repC</i> gene, encoding the first 280 amino acid residues of pSymA RepC and the rest of p42d RepC protein.	This work
pDOP/C1-990	pDOP derivative with a hybrid <i>repC</i> gene, encoding the first 330 amino acid residues of p42d RepC protein and the rest of pSymA RepC protein.	This work
pDOP/C1-1086	pDOP derivative with a hybrid <i>repC</i> gene, encoding the first 362 amino acid residues of p42d RepC protein and the rest of pSymA RepC protein.	This work

was obtained with primers *repC*-F1U and *repC*-F1L. The insert of plasmid pDOP-C/F1-F2 was obtained with primers *repC*-F1U and *repC*-F2L. Inserts of plasmids pDOP-C/F1-F3, pDOP-C/F4, pDOP-C/F4-F3, and pDOP-C/F4-F2 were obtained with the following primer pairs: *repC*-F1U and *repC*-F3L, *repC*-F4U and *repC*-

F4L, *repC*-F3U and *repC*-F4L, *repC*-F2U and *repC*-F4L, respectively.

#### **Construction of a *repC* mutant lacking its Shine-Dalgarno sequence**

The insert of the plasmid pDOP-Cs/SD was acquired by PCR with the primers *repC*d-sSDU and Mal-C2.

**Table 2 Oligonucleotides used in these work**

Name	Sequence
ALFAU2	5'-AGGGTACCCCGCAAAGAAAAGA
Mal-C2Kpn	5'-TCGGTACCTTACCCAGCCCTCAAACC
RBS-C	5'-GGATCCAAGGAAACAGCTATGCAGTCGGGGAATG
repC-F1U	5'-GCGGCCGCGGATCCAATGCAGTCGGGGAATGTA
repC-F1L	5'-ACTAGTCCCGGGAACCCCGACTCCACCAGA
repC-F2U	5'-GCGGCCGCGGATCCAATGGATCGTCCGTAAGGATAG
repC-F2L	5'-ACTAGTCCCGGGGCGCGGAATTCTGCTCGC
repC-F3U	5'-GCGGCCGCGGATCCAATGGTTCCGACCCCTGAAGGG
repC-F3L	5'-ACTAGTCCCGGGGCGCGGTGCATAGTCGCT
repC-F4U	5'-GCGGCCGCGGATCCAATGGGCGTGGGAAGTCGGCGA
repC-F4L	5'-ACTAGTCCCGGGTTACCCAGCCCTCAAACC
repC-D1U	5'-GGATCCAAGGAAACAGCTATGACGCTTGCGCTCGTGC
repC-D1L	5'-GCCAAGCTTTTATATCATCGGGCCAAGC
repC-D2U	5'-GGATCCAAGGAAACAGCTATGGGAAAAGCTGCCGATA
repC-D2L	5'-GCCAAGCTTTTATAGATCCCGCAGATAG
repC-D3U	5'-GGATCCAAGGAAACAGCTATGGAAGTGTCTAAGA
repC-D3L	5'-GCCAAGCTTTTATTCCAGGATGCACGCA
repCd-sSDU	5'-GAAAAGAGCTCCCTCAACGT
Cd-StopC-U	5'-GGATCCTAACAGTCGGGGAATG
Ttrack1-U	5'AGCCCGAGTCCGTGAACGAGTCCGAGCCGCGTCCGAGAAG GAGCAGCACATACAGAATTCAAACCC
Ttrack1-L	5'-CTCGAGGTTCTCCAAGCGGTTACGACACCTCCTCGCGGAG CATCTCCATCTCGTTGAGCAGCTAGTGACCCCTTCAAG
Ttrack2-U	5'-ATCGAGCAGCACATCCAGAATCCAAGCCCGAGTCCGTG
Ttrack2-L	5'-GGAGTTCTGGATGTGCTGCTCGATGTGGGCGCGTTGGT GCTGATCTTCTCGCGGTTCTCGAGTTCTCCAAGCGGTT
repCATBamU	5'-GGATCCAAGGAAACAGCTATGGACAGCACATGTGTAACG
repCATHindL	5'-AAGCTTCTAACCCGCCATGCCACCTC
K-Syma-L	5'-GGTACCTCACGACACCCCGCCG
C-SymA	5'-GGATCCAAGGAAACAGCTATGGAGATTGGAAGTGTGACG
Mal-C2	5'-TCAAGCTTTTACCCAGCCCTCAAACC
AL-2U	5'-ATCGGCACAGCGTTCGGCTTTTCGTCGCCCTC CTGGCGCATCGGAA
1L-BU	5'-GAGAATGCTTTTGCTTCGACCTGCTCGACCCAC TGCTGGCCCGCTCC
2L-CU	5'-GAGCAGGGGGCAAAGGCGAGCTTGAACCGGCC AACAGGCAAAAAGG
BL-3U	5'-GAGCGGTTAGACGCGCAAGCCATTAGCCTTCAGC CGAAGAATGAATCG
AL-2Uc	5'-CGAAAAGCCGAACGCTGTGCCGAT
1L-BUc	5'-CAGGTCGAAGCCAAAAGCATTCTC
2L-CUc	5'-CAAGCTCGGCTTTGCCCTGCTC
BL-3Uc	5'-AATGGCTTGCCGTCTAACCGCTC
Cd-1086L	5'-GTTGATGAAGTAGCCCTTCCAG
SaU-CdL	5'-AACGCCAAGCATAGATCGTACCGT CCCCAGCATAGATCGAACCCAC
SaL-CdU	5'-CTGTTGAGCTTCTATCCGGAGAAC GAGTTGCGTCAGGATGCACAATTG
Cs-1087U	5'-CTGGAAAGGGCTAACTTCATCAA CTCTGCTGGCGGCTATCTGCGCGAT

**Constructs carrying repC frame-shift mutations**

Plasmid pDOP-CBglII, was constructed digesting pDOP-C with *Bgl*II and filling in 5'-overhangs with T4 DNA polymerase (Fermentas). The blunted plasmid was ligated again with T4 ligase (Fermentas). Plasmid pDOP-CSphI was constructed in a similar way but digesting pDOP-C with *Sph*I.

**Construction of a repC gene carrying synonymous mutations in the A+T rich region**

The pDOP-TtMC insert was obtained by an overlap extension PCR as described by Horton et al. (1989) [25]. The first PCR was performed using the primers Ttrack1-U and Mal-C2Kpn, and pH3 DNA as initial template. This product was purified and used as

template for a second PCR with the oligonucleotides Mal-C2Kpn and Ttrack2-U; the amplification product was named T2-U. A third PCR amplification product obtained with the primers RBS-C and Ttrack1-L, and pH3 DNA as the template, was purified and used as a template in a new PCR reaction with the primers RBS-C and Ttrack2-L. The amplification product was named T2-L.

Finally, PCR products T2-U and T2-L were then mixed and used as the template for the last PCR. In this reaction, the primers Mal-C2Kpn and RBS-C were used, and the final PCR product was cloned into pDOP.

#### **Construction of *repC* hybrid genes**

Overlap extension PCR was also employed to obtain *repC* hybrid genes. *RepC* gene amplification products from pSymA were obtained using pDOP-CsA as the template, and the *repC* p42d products were obtained using pH3 as the template. Most of the hybrid genes described here required the overlap of two PCR products. The insert of plasmid pDOP/C420-1209 was obtained using the primers C-SymA and AL-2Uc for the first PCR product and AL-2U and Mal-C2 for the second product. The final PCR product was obtained with the external primers C-SymA and Mal-C2. The insert of plasmid pDOP/C1-420 was constructed with primers RBS-C and 1L-B2c and the primers 1L-B2 and K-SymAL for the first and second PCR products, respectively. These products were combined using the primers RBS-C and K-SymAL. The pDOP/C841-1209 insert was constructed with the primers C-SymA and BL-3Uc for the first PCR product and BL-3U and Mal-C2 for the second. These products were joined in a third PCR with the primers C-SymA and Mal-C2. The hybrid gene in pDOP/C1-990 was acquired with the primers RBS-C and Sal-CdL for the first PCR product and Sal-CdU and Mal-C2 for the second. These PCR products were integrated in a third PCR with the primers RBS-C and Mal-C2. Similarly, the hybrid gene of pDOP/C1-990 was obtained with the primers RBS-C and Cd-1086 for the first amplification product. To obtain the second PCR product, the primers Cs-1087U and Mal-C2 were used, and both PCR products were fused with the primers RBS-C and Mal-C2. The inserts of two of the constructs, pDOP/C421-840 and pDOP/Cs421-840, required the fusion of three PCR products. The hybrid gene located in pDOP/C421-840 required the primers C-SymA and AL-2Uc for the first PCR product, the primers AL-2U and AL-2Uc for the second PCR product, and the primers 2L-CU and K-SymA for the third PCR product. The three PCR products were fused in the final PCR with the primers C-SymA and K-SymA. The hybrid gene present in pDOP/Cs421-840 was obtained using the primers RBS-C and 1L-B2c for the first PCR product, the primers 1L-B2 and B2-3Uc for the second PCR product, and the primers BL-2U and Mal-C2 for

the third PCR product. These PCR products were linked using the primers RBS-C and Mal-C2 in the final PCR. DNA sequences of the inserts of all constructs were obtained to corroborate their correctness.

#### **Plasmid profiles**

The plasmid profiles of four transconjugants from each cross were visualized on agarose gels according to the protocol described by Hynes and McGregor [26].

#### **DNA isolation and manipulation**

Plasmid DNA was isolated using the High Pure Plasmid Isolation Kit (Roche) according to the manufacturer's instructions. Restriction and ligation reactions were performed under the conditions specified by the enzyme manufacturer (Fermentas). PCR was performed using Platinum High Fidelity *Taq* Platinum Polymerase or ThermoLace™ DNA Polymerase (Invitrogen). PCR products were cloned using the TOPO TA Cloning Kit (Invitrogen).

#### **Plasmid incompatibility**

The incompatibility properties of the constructs were determined as described in Ramírez-Romero *et al.* [7].

#### **Plasmid replication in *R. etli***

To determine the replication capabilities of the pDOP derivatives in *R. etli*, the plasmids were introduced into CFNX107 by conjugation. The plasmid profiles of at least four transconjugants from each cross were analyzed. A recombinant plasmid was considered capable of replicating in *R. etli* if the plasmid profiles of the transconjugants showed a new band of the expected size.

#### **Plasmid copy-number determination**

The plasmid copy numbers of the CFNX107 transconjugants containing pDOP derivatives were evaluated as follows: the total DNA of each transconjugant was isolated, digested with *Hind*III endonuclease, resolved in a 1% agarose gel and transferred to Hybond-N+ membranes (Amersham). The blot was then simultaneously hybridized with an  $\Omega$ - spectinomycin cassette located within the *recA* gene (chromosome-encoded) and with a fragment of pDOP; both probes were of the same size and GC content. The hybridization signals were quantified with a PhosphorImager SI (Molecular Dynamics). The plasmid copy-number was calculated from the ratio of the integrated hybridization signal of the recombinant plasmid and the integrated hybridization signal of the chromosome.

#### **Bioinformatics**

Alignments were performed with Clustal-W [27] at the WWW service of the European Bioinformatics Institute <http://www2.ebi.ac.uk/clustalw>. Protein secondary

structure predictions were made with PSIPRED [28] at the WWW service of the Bioinformatics Group, UCL Department Of Computer Science <http://bioinf.cs.ucl.ac.uk/psipred/>. The DNA duplex helical stability profile was calculated using WEB-THERMODYN: sequence analysis software for profiling DNA helical stability [http://www.gsa.buffalo.edu/dna/dk/WEBTHERMODYN/\[29\]](http://www.gsa.buffalo.edu/dna/dk/WEBTHERMODYN/[29]).

## Results

### The *oriV* of p42d is located within the *repC* coding sequence

The basic replicon of *Rhizobium etli* p42d, defined as the smallest DNA region that contains all of the elements required to replicate with the same stability and plasmid copy-number as the parental plasmid, consists of the complete *repABC* operon plus 500 bp downstream of the *repC* stop codon (*inc-beta* region, containing *parS*) and 86 bp upstream of the *repA* initiation codon [8] (Figure 1). pDOP-E', a construct which carries the complete *repC* gene, the intergenic sequence between *repB* and *repC* (*inc-alpha*), and the 500 bp downstream of the *repC* stop codon under a constitutive promoter (Plac promoter), can replicate but does so with a slightly higher plasmid copy-number than the parental plasmid. However, derivatives that lack parts of the gene encoding the antisense RNA were unable to replicate [20].

To identify the minimal region of p42d that is capable of independent replication (putting aside the properties of the parental plasmid), we further explored the region between the *repB* stop codon and the 500 bp downstream of the *repC* stop codon. Three PCR products that possessed parts of this region were amplified and cloned into pDOP, a mobilizable suicide vector, under the control of the Plac promoter, which behaves as a constitutive promoter in *Rhizobium*. The first construct (pDOP- $\alpha$ C) contained the *repB-repC* intergenic region (*inc-alpha*) and the complete *repC* gene. The second construct, pDOP-SDnC, contained the *repC* open reading frame (ORF), including its putative *repC* Shine-Dalgarno (SD) sequence (AGGUG). The third construct contained the *repC* ORF but with a SD sequence that was more similar to the *Rhizobium etli* SD consensus (AGGAA) positioned 6 bp prior to the *repC* initiation codon (pDOP-C). As a control, we introduced a *HindIII* fragment of 5.6 Kb that carried the entire *repABC* of p42d into pDOP conferring it the ability to replicate in *Rhizobium* (Figure 1) [24].

These constructs were introduced by mating into a *recA Rhizobium etli* CFN42 derivative lacking the p42d and p42a plasmids (CFNX107) (Figure 1). Only constructs pDOP-H3, pDOP- $\alpha$ C and pDOP-C were introduced with similar conjugation frequencies, from  $1.6 \times 10^3$  to  $6 \times 10^4$ . However, CFNX107/pDOP-C transconjugants

formed colonies after a longer time period (6-7 days), which was slower than the CFNX107/pDOP- $\alpha$ C and CFNX107/pDOP-H3 transconjugants and the receptor strain CFNX107 (3-4 days). Plasmid profile analyses of the transconjugants showed that the introduced plasmids replicated independently (Figure 2). The analyses also showed that pDOP-C replicated with a higher plasmid copy-number than pDOP-H3 carrying the complete p42d *repABC* operon. This observation was corroborated by measuring the plasmid copy-number of the transconjugants: 6 copies of pDOP-C were present per chromosome instead of 1-2 copies of the control plasmid pDOP-H3 (Figure 3).

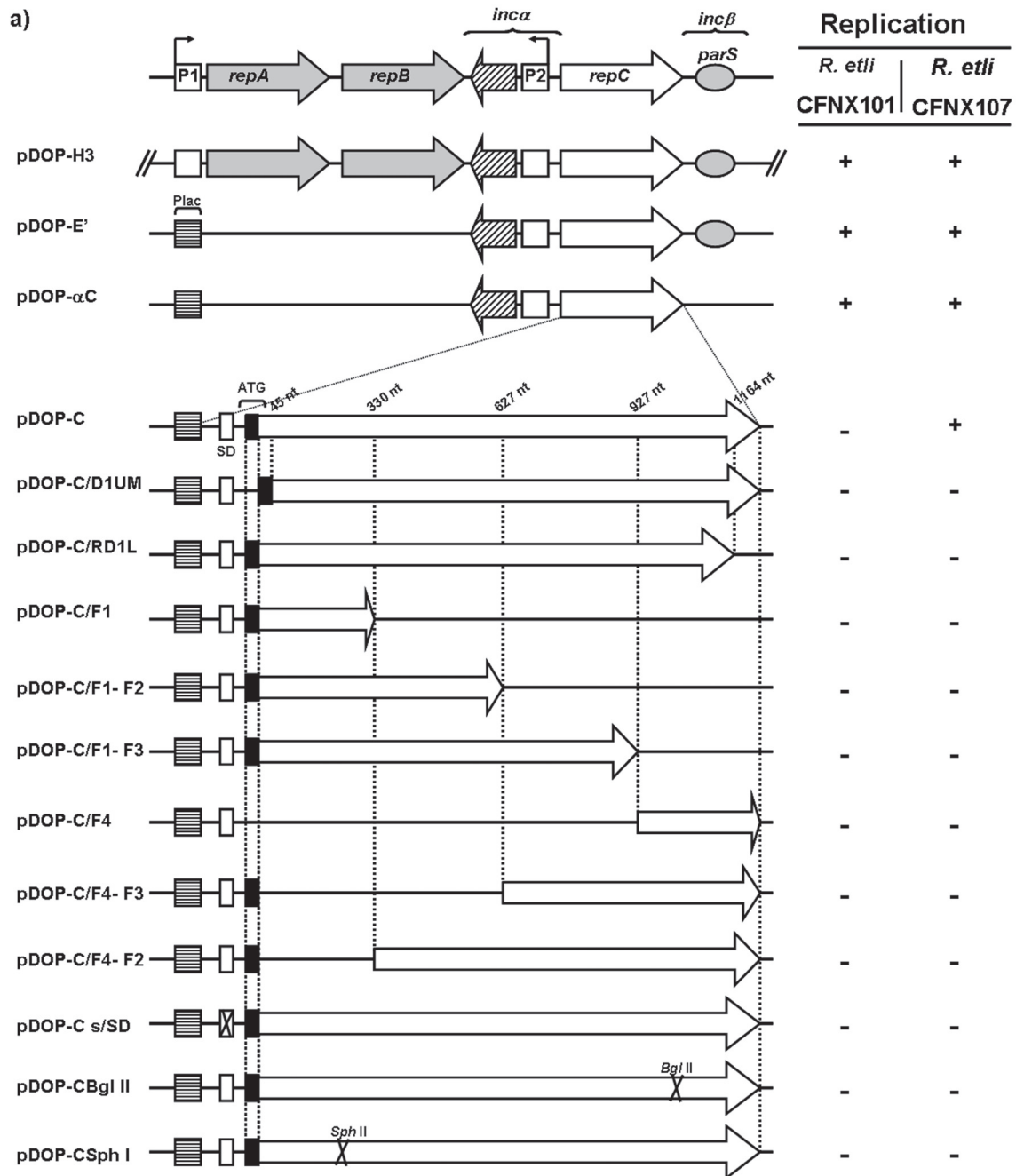
These results indicate that the minimal replicon of p42d consists of a *repC* gene under a constitutive promoter (Plac) and the SD sequence that we introduced and that the origin of replication resides within the *repC*-coding region. However, the growth rate of CFNX107 strain was negatively influenced by the introduction of pDOP-C (see Figure 4).

To prove that RepC is essential for replication, two *repC* deletions and two frame-shift mutants were constructed and cloned into pDOP under the control of the Plac promoter. Plasmid pDOP-C/D1UM contained a *repC* gene with a deletion of 14 codons (from codon 2 to 14), and plasmid pDOP-C/RD1L contained a *repC* gene with a deletion at its 3' end of 14 codons (from codon 388 to 401). The construct pDOP-CBglII possessed a *repC* gene with a frame-shift mutation at nucleotide 948, while plasmid pDOP-CSphI carried a frame-shift mutation at nucleotide 277. All of these constructs contained the same SD sequence as construct pDOP-C and were in the same relative orientation with respect to PLac in the vector. All plasmids were mated into the *R. etli* CFNX107 strain, but no transconjugants were obtained, indicating that the complete RepC product is crucial for replication.

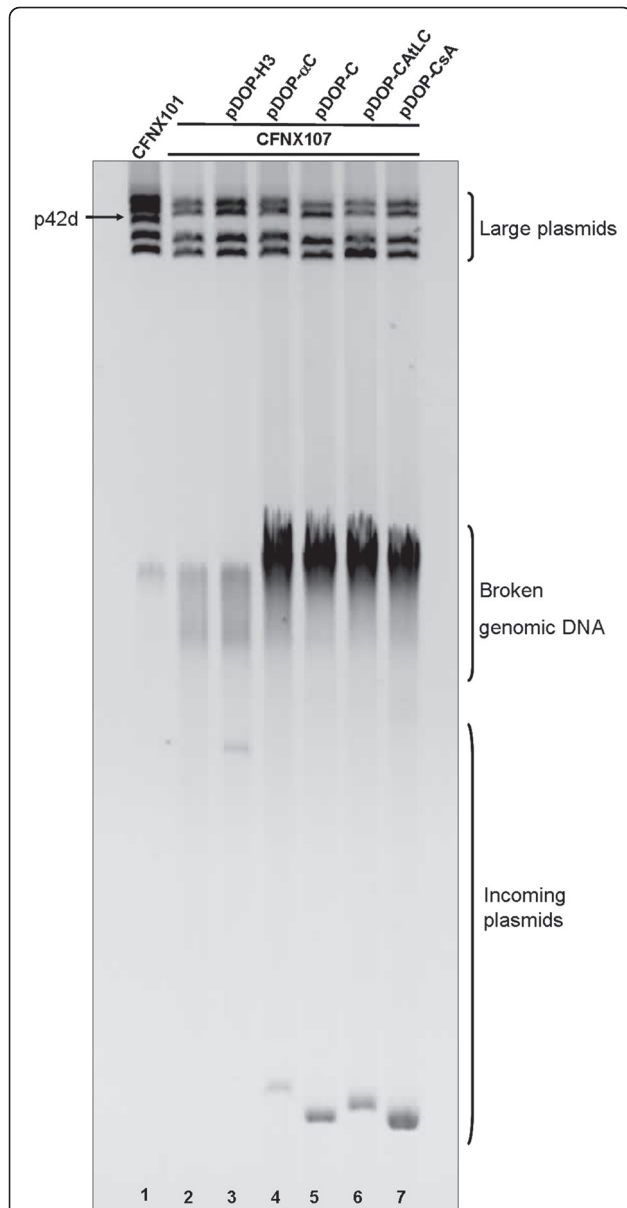
To demonstrate that these observations were not specific to the p42d *repC* sequence, the *repC* genes of *S. meliloti* 1021 pSymA and the *A. tumefaciens* C58 linear chromosome were amplified by PCR and introduced into pDOP under Plac control and downstream of a SD sequence. The recombinant plasmids were conjugated into *R. etli* strain CFNX107, and the plasmid profiles of the transconjugants were analyzed. Both recombinant plasmids were capable of replication in *Rhizobium*, as was pDOP-C (Figure 2). These results clearly suggest that the presence of an origin of replication (*oriV*) within *repC* is a general property of *repABC* operons.

### Analysis of the *repC* sequence: the role of the high A+T content region

To circumscribe the origin of replication (*oriV*) of the *repABC* plasmids, we performed an *in silico* analysis to

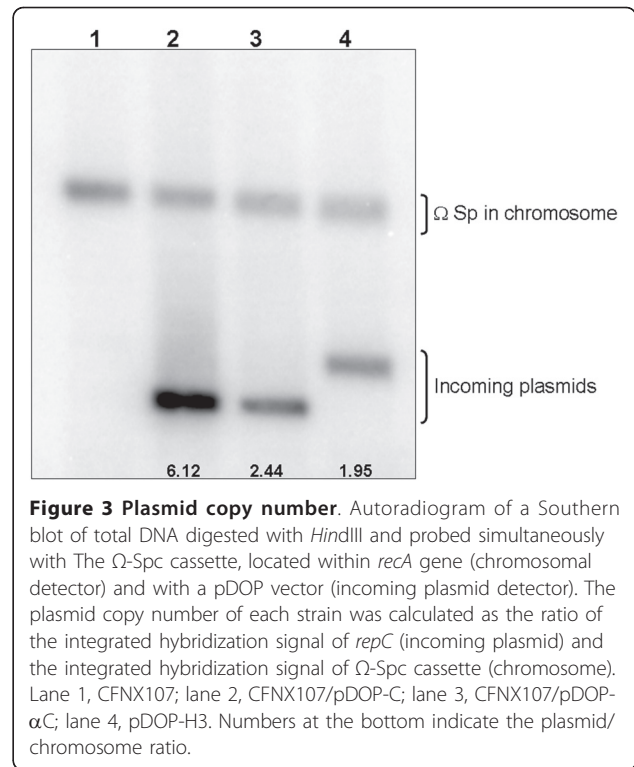


**Figure 1 Linear representation of the constructs used in this work. a)** At the top of the figure the p42d *repABC* operon is shown. Grey arrows represent genes encoding the partitioning proteins and *parS* and the grey ellipse represents the centromeric-like region *parS*. A white arrow shows the relative position of the gene encoding RepC, a protein essential for replication. Dashed arrow represents a gene encoding a small antisense RNA that modulates *repC* expression. Boxed P1 and P2, indicate the position and transcription directions of the promoters found within the *repABC* operon. Brackets indicate regions involved in plasmid incompatibility. Below, graphic representation of the genetic elements present in each one of the constructs used in this work, using the same symbols than above. Square filled with horizontal lines shows the relative position of *pLac*, a constitutive promoter in *Rhizobium*. **b)** A magnification of the *repC* gene and *repC* gene fragments present in the constructs, including the genetic elements introduced by us: white vertical rectangle represent a Shine-Dalgarno (SD) sequence, while the black vertical rectangle shows the initiation codon. Crossed rectangle indicates that the SD sequence was eliminated in that particular construction. Crosses within the white arrows, marked with *Sph*I or *Bgl*II, indicate that inserts of those constructs possess a frame-shift mutation in that specific point. Construct names are listed in the left column and their replication capabilities in strains CFNX101 and CFNX107 are listed in the columns in the right: (+) indicates that the construct is capable of autonomous replication and (-) that the construct does not have this property.



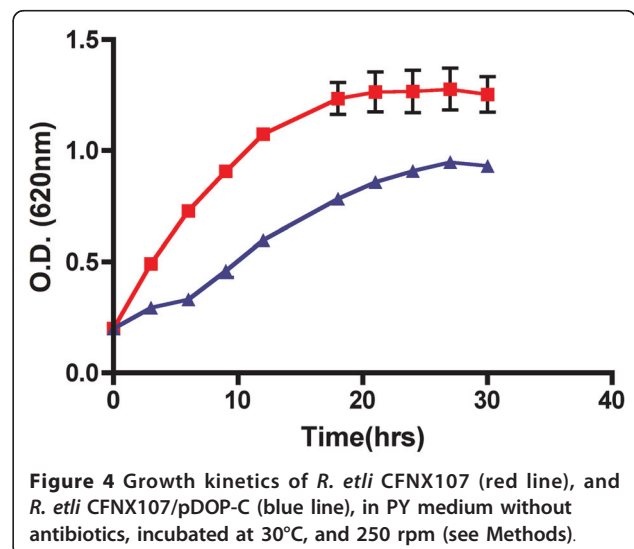
**Figure 2 Plasmid profiles of *Rhizobium etli* CFNX101, and *Rhizobium etli* CFNX107 transconjugants, carrying the following plasmids: pDOP-H3, pDOP- $\alpha$ C, pDOP-C, pDOP-CAtLC, pDOP-CsA.** Brackets at right show the positions of the resident large plasmids, broken DNA, and of the incoming plasmids. Arrow at left shows the location of plasmid p42d, in *R. etli* CFNX101. Negative image of Ethidium bromide stained gel.

search for three sequence features that are characteristic of the *oriV* in low copy-number plasmids: a set of tandem direct repeat sequences (iterons), a region of high A+T content, and DnaA boxes. We only detected a region of high A+T content between positions 450 and 850 of the *repC* coding region. However, we did not find any trace of even highly degenerated direct repeat sequences or of DnaA boxes.



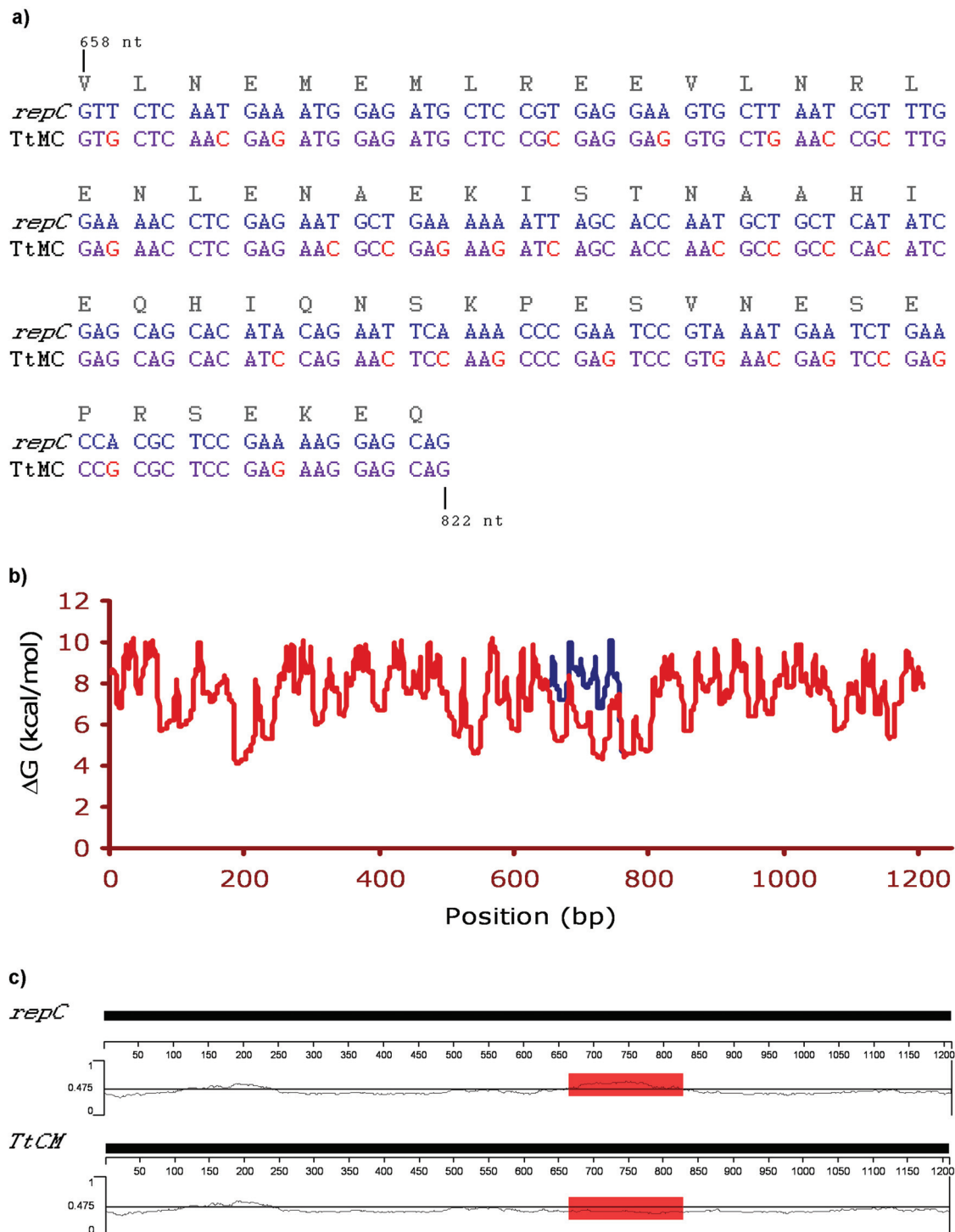
**Figure 3 Plasmid copy number.** Autoradiogram of a Southern blot of total DNA digested with *Hind*III and probed simultaneously with The  $\Omega$ -Spc cassette, located within *recA* gene (chromosomal detector) and with a pDOP vector (incoming plasmid detector). The plasmid copy number of each strain was calculated as the ratio of the integrated hybridization signal of *repC* (incoming plasmid) and the integrated hybridization signal of  $\Omega$ -Spc cassette (chromosome). Lane 1, CFNX107; lane 2, CFNX107/pDOP-C; lane 3, CFNX107/pDOP- $\alpha$ C; lane 4, pDOP-H3. Numbers at the bottom indicate the plasmid/chromosome ratio.

To determine if the high A+T content region has a role in plasmid replication, we constructed a *repC* derivative in which a group of silent mutations were introduced with the aim of altering the A+T content and increase the DNA duplex stability of this region, without disrupting the *repC* product (Figure 5). This *repC* mutant was cloned into pDOP under the Plac promoter and a SD sequence, generating the plasmid pDOP-TtMC. This plasmid could not replicate in *Rhizobium*



**Figure 4 Growth kinetics of *R. etli* CFNX107 (red line), and *R. etli* CFNX107/pDOP-C (blue line), in PY medium without antibiotics, incubated at 30°C, and 250 rpm (see Methods).**





**Figure 5** a) Gene alignment of *repC* and its mutant derivative pDOP-TtMC from position 658 to 822, indicating nucleotide changes introduced into pDOP-TtMC (red letters) to increase the C+G content of this region. Note that the included mutations did not change the RepC protein sequence. **b)** DNA duplex stability expressed as  $\Delta G$  along *repC* gene (red line) and its mutant derivative TtMC (blue line). **c)** Graphic showing A+T content along *repC* gene and its mutant derivative TtMC. A+T average in both genes is the same: 0.475. The A+T rich region of *repC* is boxed. Note that the equivalent region in TtMC, also boxed, the A+T content is above the average.

strains with or without p42d, indicating that the A+T rich region plays a major role in replication.

#### RepC exerts its action in *cis*

The identification of an *oriV* sequence is generally based on its ability to facilitate replication when present on a plasmid that otherwise could replicate only if the appropriate replication factors (e.g., an initiator protein) were provided in *trans*. To more precisely locate the *oriV* within *repC*, we cloned a collection of internal segments of *repC* into the suicide vector pDOP (Figure 1). This collection was conjugated into an *R. etli* strain containing the parental plasmid (CFNX101) as the source of all the *trans* elements required for replication, but we were unable to obtain transconjugants.

To determine if the activation of *oriV* requires transcription (i.e., the *repC* mRNA also acts as a replication primer), we constructed a pDOP derivative that contained a *repC* gene but lacked a SD sequence (pDOP-Cs/SD) (Figure 1). This plasmid was also incapable of replicating in *R. etli* CFNX101. Similarly, the two plasmids with *repC* frame-shift mutations, pDOP-CBglII and pDOP-CSphI, were also conjugated into *R. etli* CFNX101 without success. Overall, these results indicate that RepC exerts its action in *cis*.

#### RepC as an incompatibility factor

Plasmid incompatibility, or the inability of two replicons to coexist in the same cell line, results from the sharing of elements involved in plasmid replication, partitioning or control [30]. The *repC* open reading frame of p42d, when cloned in a vector capable of replicating in *R. etli*, CFNX101, can coexist with p42d [8]. However, all of our attempts to introduce the construct pDOP-C into *R. etli* CFNX101 failed. In contrast, CFNX101 transconjugants carrying a similar construct (pDOP-CsA) that contained the *repC* gene pSymA of *S. meliloti* 2011 were easily obtained. The frequencies with which CFNX101/pDOP-CsymA and CFNX107/pDOP-CsymA transconjugants were obtained were similar (average  $5 \times 10^{-3}$ ). Moreover, the plasmid profiles of the transconjugants showed that pDOP-CsA replicated in these strains as an independent entity. These observations indicate that pDOP-C and its parental plasmid p42d are incompatible, while that of pDOP-CsymA and p42d are compatible.

The RepC protein of *S. meliloti* 2011 pSymA shares 54% identity with the p42d RepC protein, and both proteins have very similar secondary structures (Figure 6). To map the RepC regions of p42d that are involved in plasmid incompatibility, a collection of hybrid genes containing fragments of the *repC* genes from *S. meliloti* pSymA and *R. etli* p42d were constructed. A schematic representation of the hybrid genes and their properties is shown in Figure 7. The hybrid genes were designed so

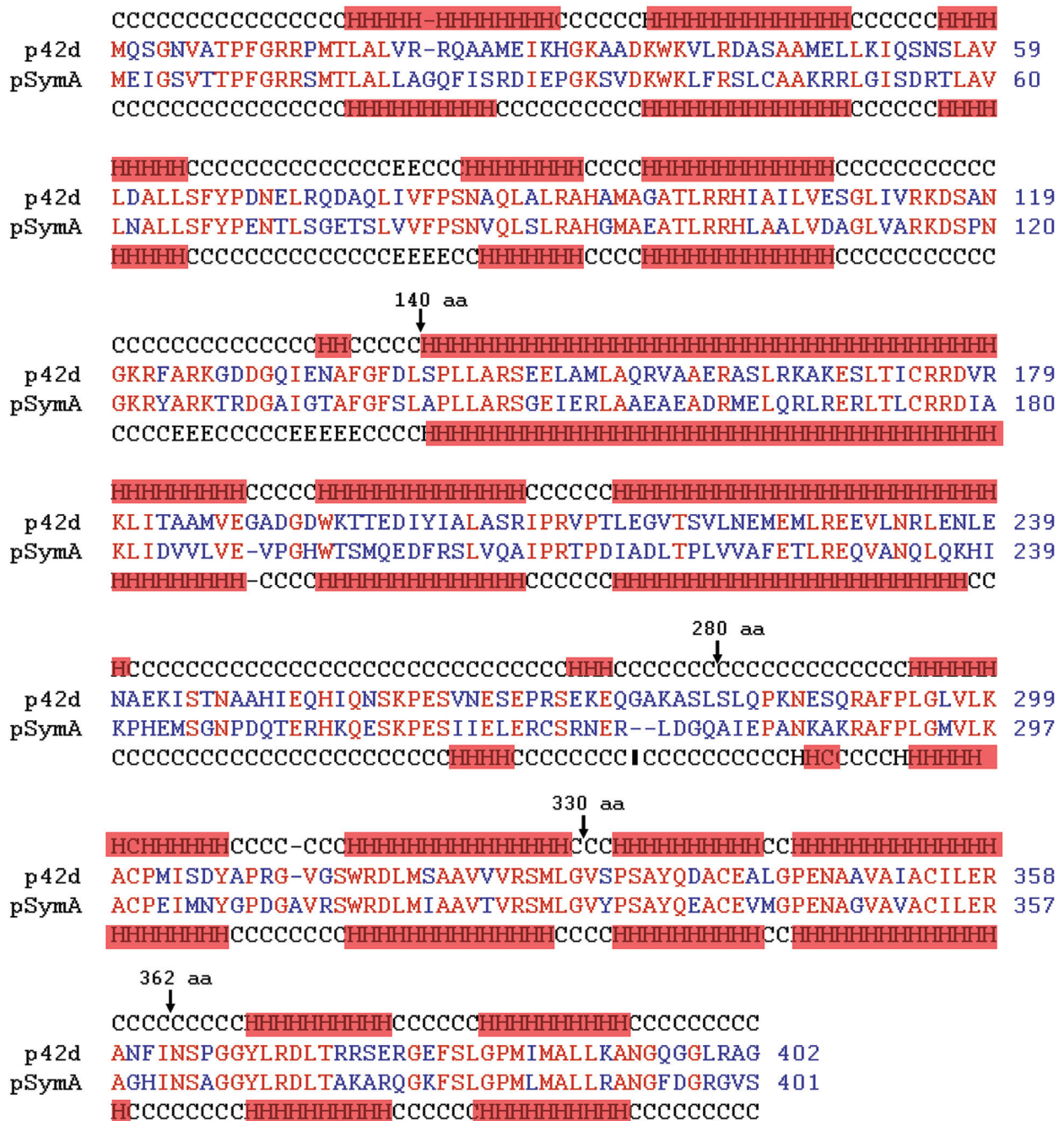
that none of the predicted alpha-helix and beta regions of the *repC* products were disturbed. The hybrid genes were cloned into pDOP under the Plac promoter and transferred by conjugation into *R. etli* CFNX107 to determine their ability to replicate autonomously and into *R. etli* CFNX101 to test if they were able to replicate without the interference of p42d. Two constructs were capable of replicating in both genetic backgrounds: pDOP-C1-990 and pDOP-C1-1086. The rest of the constructs failed to replicate in both strains. The plasmid pDOP-C1-1086 expresses a hybrid protein containing the first 362 amino acid residues (aa) of the p42d RepC protein and the last 39 aa carboxy-terminal region of the pSymA RepC protein. With respect to plasmid incompatibility, this recombinant plasmid behaved the same as plasmid pDOP-CsymA, i.e., it replicated similarly in the strains CFNX101 and CFNX107. This result indicates that the RepC region involved in plasmid incompatibility resides in the last 39 amino acid residues of the protein.

#### Discussion

Plasmids in which the *oriV* is located in the gene encoding an initiation protein are uncommon but not exceptional. The *Enterococcus faecalis* pheromone-responding plasmid pAD1 [31] (Francia, et al., 2004), the *Staphylococcus xylosum* plasmid pSX267 [32], the plasmids pAM $\beta$ 1 and pLS32 from *Bacillus subtilis* [33-35], and the *Staphylococcus aureus* multiresistance plasmids pSK1 and pSK41 [36,37] fall into this category. However, the origins of replication in all of these plasmids have recognizable iterons, and an insert that contains some or all of the iterons from these plasmids is usually capable of driving plasmid replication if the initiator protein is provided in *trans*. The minimal replicon of the p42d plasmid is the *repC* ORF sequence driven by a constitutive promoter (Plac) with an SD sequence that we designed. Frame shift and deletion mutants of the *repC* gene disrupted the capacity for replication of the minimal replicon, indicating that RepC is essential for replication and is likely the initiator protein. To confirm this function, it will be necessary to demonstrate that this protein binds the *oriV*, melts the double-stranded DNA, and recruits the initiation host factor.

A DNA sequence analysis of the *repC* gene clearly showed the absence of iterons or other large, perfect or imperfect, repetitive sequences (>8 bp), which are the typical DNA-binding sites of plasmid initiator proteins [1].

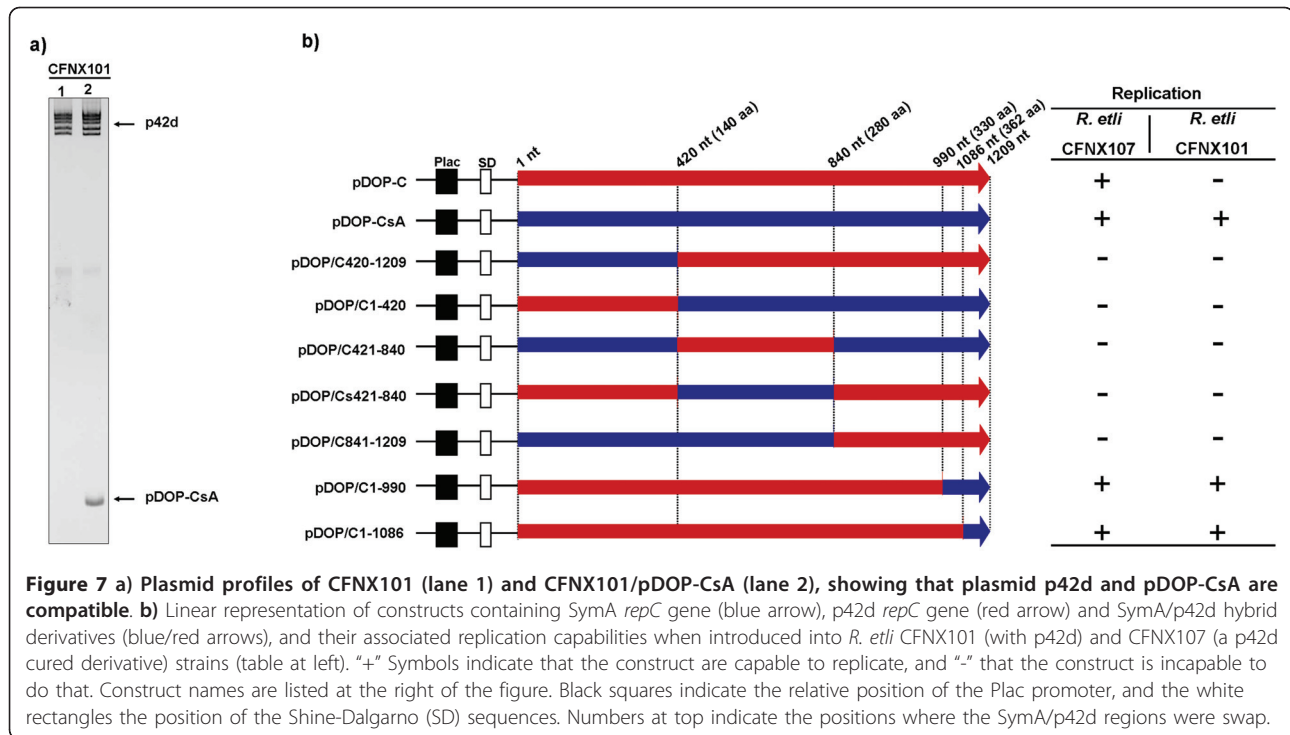
The replication of several bacterial plasmids, such as P1, F, R6K, RK2, Rts1, pMU720, and pSC101, requires a crucial and concerted participation of DnaA and the plasmid-encoded initiator protein. These plasmids contain at least one DnaA box in their *oriV* sequences [38-43]. For other plasmids, DnaA participates only as an accessory, but these plasmids also contain DnaA boxes in their origins of replication (e.g., pR1) [44].



**Figure 6 Protein alignment of p42d RepC from *R. etli* CFN42 and pSymA RepC from *S. meliloti* 1021 and where identical amino acid residues are marked in red. The secondary structures of these proteins are also shown. Coiled regions are marked with C; helical regions are marked with boxed H letters; and with letter E, the stranded regions. Arrows with an associated numbers indicates the positions where the genes were swapped, in the hybrid genes (see table 1).**

However, we failed to identify such DnaA boxes within the *repC*-coding region, suggesting that DnaA does not have a role in p42d replication. A common property of theta-replicating plasmids is an A+T rich region close to the origin of replication,

which is necessary for strand melting and the assembly of host initiation factors [1]. The *repC* ORF sequence of p42d contains a large A+T rich region that is crucial for plasmid replication. A construct carrying silent mutations that partially eliminated the A+T rich region was



unable to promote replication in *R. etli* strains with or without the symbiotic plasmid, indicating that this region is an essential part of the *oriV*. However, a sequence analysis of other *repC* genes located in *repABC* operons revealed that an A+T rich region was present in all of the analyzed plasmids but its relative location was not conserved (data not shown).

The p42d minimal replicon (pDOP-C) has two intriguing properties. First, the construct resulted in enhancing the plasmid copy-number to around six, in contrast parental plasmid, which was maintained at 1-2 copies per chromosome. Second, the strain carrying this construct has a longer duplication time and a lower yield when the cells reach stationary phase than the strain without this construct.

While describing the observed increase in the plasmid copy-number, we must bear in mind that the *repC* gene in pDOP-C was expressed by a constitutive promoter.

In addition, the negative transcriptional regulation of the *repC* gene expression mediated by RepA and RepB was eliminated, and the antisense RNA (ctRNA), which also plays a negative role in the expression of *repC*, was removed. In the absence of these layers of negative regulation, it is expected that the plasmid replication would accelerate resulting in the production of new DNA molecules with a concomitant increase in the number of new origins of replication, which in turn, could be used to promote new rounds of replication, leading to cell death. However, in the present study, with the use of

the minimal replicon (pDOP-C) we did not observe cell death, and the plasmid copy-number increased only moderately. This observation suggests the existence of a posttranslational mechanism that limits RepC activity, thus preventing over-initiation.

Growth kinetics of CFNX101 and CFNX107 were identical (data not shown), however, when pDOP-C was introduced into CFNX107 growth of the bacterium was inhibited. The growth rate and yield diminution observed in strain CFNX107/pDOP-C relative to CFNX107 is not likely caused by the metabolic burden imposed by pDOP-C replication. The size of the parental plasmid (p42d) is approximately 374 Kb, while the size of pDOP-C is approximately 5.57 Kb; even if we take into consideration the 6-fold increase in plasmid copy-number, the amount of DNA required for replication in CFNX107/pDOP-C is several fold lower than the amount of DNA required for replication in CFNX101. Based on these observations it can be hypothesized that RepC, being an initiator protein, must perform three tasks: recognize the origin of replication, unwind the DNA at the origin, and recruit the replisome. An excess of RepC could lead to the formation of more of replication "bubbles". However, if one or more elements of the replisome are suboptimal in the growing cell, then, some replication forks will be stalled resulting in inhibition of cell division and growth.

We demonstrated that pDOP-C was capable of autonomous replication in an *R. etli* strain lacking the

parental plasmid (p42d). However, we could not introduce this construct into an *R. etli* strain harboring the parental plasmid. In contrast, a similar construct that contained the *repC* gene of *S. meliloti* pSymA replicated autonomously with the same behavior in both strains. This result indicates that RepC is an incompatibility factor that prevents the coexistence of p42d and pDOP-C and that the incompatibility phenomenon is replicon-specific. Additionally, a construct (pDOP-C1-1086) expressing a chimeric protein consisting of the amino-terminal region of p42 RepC and 39 aa residues of the carboxy-terminal region of the pSymA RepC protein was capable of replicating as an independent entity with the same efficiency in *R. etli* strains, with or without p42d. This result indicates that the last 39 aa residues of the RepC carboxy-terminal region are directly involved in the incompatibility phenotype. A close inspection of this region in the RepC proteins of pSymA and p42d shows that they share 62.5% of identity, indicating that 15 amino acid residues or less are critical in promoting the incompatibility phenotype. Interestingly, however, in spite of the variations in 15 aa residues, RepC proteins of p42d and pSymA have a similar secondary structure: both possess two alpha helices of ten amino acid residues each, separated by a coiled region of six amino acid residues, in the same relative positions.

Our current hypothesis linking incompatibility and the RepC posttranslational regulation is as follows: RepC, like many other plasmid-encoded initiator proteins, exists in two forms, an active monomer and an inactive dimer, and protein thermodynamics favors dimer formation [1]. The RepC carboxy-terminal region is involved in dimer formation, and the dimerization process is replicon-specific. The introduction of pDOP-C into a strain containing p42d displaces the RepC monomer-dimer equilibrium that favors the inactive form, preventing the establishment of the incoming plasmid. A similar introduction of a construct with the RepC of a compatible plasmid will not affect the monomer-dimer equilibrium and will allow the establishment of the new plasmid.

Another unusual observation was the inability to complement the *repC* ORF *in trans* for replication. One possibility is that the *repC* transcript acts as an RNA primer for replication or assists in DNA melting at the *oriV*. However, the construct pDOP-Cs/SD, which lacks a SD sequence, could not replicate in CFNX101, suggesting that translation is required for the newly synthesized RepC protein to be located at the *oriV*. To the best of our knowledge, the only initiator protein that functions only in *cis* is RepA from prophage N15 [45]. At this

stage we cannot determine which of these possibilities is more likely, and further experiments are needed to resolve these questions.

## Conclusions

RepC is the only element encoded in the *repABC* operon of the *Rhizobium etli* p42d plasmid that is necessary and sufficient for plasmid replication and is likely the initiator protein. The *oriV* of this plasmid resides within the *repC* gene and is located close to or inside of a large A+T region. This architecture is shared by other *repABC* plasmids. Our results also indicate that RepC can act as an incompatibility factor and that the last 39 aa of the carboxy-terminal region of this protein are involved in this phenotype.

## Acknowledgements and Funding

This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT, México) (Grant number: 00000000100099); and by the Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT-UNAM, México) (Grant number IN205611-3) to M.A. C. R. C-R, F. P-L and G P-S were supported during the Ph.D. program (Programa de Doctorado en Ciencias Biomédicas-Universidad Nacional Autónoma de México) with scholarships from Consejo Nacional de Ciencia y Tecnología and Dirección General de Estudios de Posgrado (México). We are greatly indebted to Ángeles Pérez-Oseguera for her technical support, and to Dr. Pallavolu Maheswara Reddy for his critical review of the manuscript.

## Authors' contributions

R C-R conducted the bulk of the experiments and made the constructions; F P-L and G P-S made growth kinetics, plasmid profiles and incompatibility experiments. MAC designed and coordinated the study, and wrote the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

Received: 15 April 2011 Accepted: 30 June 2011

Published: 30 June 2011

## References

1. del Solar G, Giraldo R, Ruiz-Echevarría MJ, Espinosa M, Díaz-Orejas R: Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* 1998, **62**:434-464.
2. Nordström K, Molin S, Light J: Control of replication of bacterial plasmids: genetics, molecular biology, and physiology of the plasmid R1 system. *Plasmid* 1984, **12**:71-90.
3. Paulsson J, Chatteraj DK: Origin inactivation in bacterial DNA replication control. *Mol Microbiol* 2006, **61**:9-15.
4. Zakrzewska-Czerwinska J, Jakimowicz D, Zawilak-Pawlik A, Messer W: Regulation of the initiation of chromosomal replication in bacteria. *FEMS Microbiol Rev* 2007, **31**:378-387.
5. Cevallos MA, Cervantes-Rivera R, Gutiérrez-Ríos RM: The *repABC* plasmid family. *Plasmid* 2008, **60**:19-37.
6. Castillo-Ramírez S, Vázquez-Castellanos JF, González V, Cevallos MA: Horizontal gene transfer and diverse functional constraints within a common replication-partitioning system in Alphaproteobacteria: the *repABC* operon. *BMC Genomics* 2009, **10**:536.
7. Pappas KM: Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. *Plasmid* 2008, **60**:89-107.
8. Ramírez-Romero MA, Soberón N, Pérez-Oseguera A, Téllez-Sosa J, Cevallos MA: Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. *J Bacteriol* 2000, **182**:3117-3124.

9. Pappas KM, Winans SC: The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid repABC operon. *Mol Microbiol* 2003, **49**:441-455.
10. Gerdes K, Moller-Jensen J, Jensen RB: Plasmid and chromosome partition: surprises from phylogeny. *Mol Microbiol* 2000, **37**:455-466.
11. Ramírez-Romero MA, Téllez-Sosa J, Barrios H, Pérez-Oseguera A, Rosas V, Cevallos MA: RepA negatively autoregulates the transcription of the repABC operon of the *Rhizobium etli* symbiotic plasmid basic replicon. *Mol Microbiol* 2001, **42**:195-204.
12. Tabata S, Hooykaas PJ, Oka A: Sequence determination and characterization of the replicator region in the tumor-inducing plasmid pTiB6S3. *J Bacteriol* 1989, **171**:1665-1672.
13. Bartosik D, Baj J, Włodarczyk M: Molecular and functional analysis of pTAV320, a repABC-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. *Microbiology* 1998, **144**:3149-3157.
14. Bartosik D, Szymank M, Wysocka E: Identification of the partitioning site within the repABC-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J Bacteriol* 2001, **183**:6234-6243.
15. Soberón N, Venkova-Canova T, Ramírez-Romero MA, Téllez-Sosa J, Cevallos MA: Incompatibility and the partitioning site of the repABC basic replicon of the symbiotic plasmid from *Rhizobium etli*. *Plasmid* 2004, **51**:203-216.
16. Chai Y, Winans SC: RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between repA and repB for plasmid partitioning and autorepression. *Mol Microbiol* 2005, **58**:1114-1129.
17. MacLellan SR, Zaheer R, Sartor AL, MacLean AM, Finan TM: Identification of a megaplasmid centromere reveals genetic structural diversity within the repABC family of basic replicons. *Mol Microbiol* 2006, **59**:1559-1575.
18. Chai Y, Winans SC: A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol Microbiol* 2005, **56**:1574-1585.
19. MacLellan SR, Smallbone LA, Sibley CD, Finan TM: The expression of a novel antisense gene mediates incompatibility within the large repABC family of alpha-proteobacterial plasmids. *Mol Microbiol* 2005, **55**:611-623.
20. Venkova-Canova T, Soberón NE, Ramírez-Romero MA, Cevallos MA: Two discrete elements are required for the replication of a repABC plasmid: an antisense RNA and a stem-loop structure. *Mol Microbiol* 2004, **54**:1431-1444.
21. Cervantes-Rivera R, Romero-López C, Berzal-Herranz A, Cevallos MA: Analysis of the mechanism of action of the antisense RNA that controls the replication of the repABC plasmid p42d. *J Bacteriol* 2010, **192**:3268-3278.
22. Noel KD, Sanchez A, Fernandez L, Leemans J, Cevallos MA: *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J Bacteriol* 1984, **158**:148-155.
23. Simon R, Priefer U, Pühler A: A broad host-range mobilization system for in vivo genetic engineering transposon mutagenesis in Gram negative bacteria. *Bio/Technology* 1983, **1**:784-791.
24. Ramírez-Romero MA, Bustos P, Girard L, Rodríguez O, Cevallos MA, Dávila G: Sequence, localization and characteristics of the replicator region of the symbiotic plasmid of *Rhizobium etli*. *Microbiology* 1997, **143**:2825-2831.
25. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR: Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 1989, **77**:61-68.
26. Hynes MF, McGregor NF: Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. *Mol Microbiol* 1990, **4**:567-574.
27. Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994, **22**:4673-4680.
28. Jones DT: Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 1999, **292**:195-202.
29. Huang Y, Kowalski D: WEB-THERMODYN: sequence analysis software for profiling DNA helical stability. *Nucl Acids Res* 2003, **31**:3819-3821.
30. Novick RP: Plasmid incompatibility. *Microbiol Rev* 1987, **51**:381-395.
31. Francia MV, Fujimoto S, Tille P, Weaver KE, Clewell DB: Replication of *Enterococcus faecalis* pheromone-respondering plasmid pAD1: location of the minimal replicon and oriV site and RepA involvement in initiation of replication. *J Bacteriol* 2004, **186**:5003-5016.
32. Gering M, Götz F, Brückner R: Sequence and analysis of the replication region of the *Staphylococcus xylosus* plasmid pSX267. *Gene* 1996, **182**:117-122.
33. Bruand C, Ehrlich SD: Transcription-driven DNA replication of plasmid pAMBeta1 in *Bacillus subtilis*. *Mol Microbiol* 1998, **30**:135-145.
34. Tanaka T, Ogura M: A novel *Bacillus natto* plasmid pLS32 capable of replication in *Bacillus subtilis*. *FEBS Lett* 1998, **422**:243-246.
35. Tanaka T, Ishida H, Maehara T: Characterization of the replication region of plasmid pLS32 from the Natto strain of *Bacillus subtilis*. *J Bacteriol* 2005, **187**:4315-4326.
36. Kwong SM, Skurray RA, Firth N: *Staphylococcus aureus* multi-resistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense RNA regulation. *Mol Microbiol* 2004, **51**:497-509.
37. Kwong SM, Skurray RA, Firth N: Replication control of staphylococcal multi-resistance plasmid pSK41: an antisense RNA mediates dual-level regulation of Rep expression. *J Bacteriol* 2006, **188**:4404-4412.
38. Betteridge T, Yang J, Pittard AJ, Praszker J: Role of RepA and DnaA proteins in the opening of the origin of DNA replication of an IncB plasmid. *J Bacteriol* 2004, **186**:3785-3793.
39. Gaylo PJ, Turjman N, Bastia D: DnaA protein is required for replication of the minimal replicon of the broad-host-range plasmid RK2 in *Escherichia coli*. *J Bacteriol* 1987, **169**:4703-4709.
40. Hansen EB, Yarmolinsky MB: Host participation in plasmid maintenance: dependence upon dnaA of replicons derived from P1 and F. *Proc Natl Acad Sci USA* 1986, **83**:4423-4427.
41. Hasunuma K, Sekiguchi M: Replication of plasmid pSC101 in *Escherichia coli* K12: requirement for dnaA function. *Mol Gen Genet* 1977, **154**:225-230.
42. Itoh Y, Terawaki Y: Replication properties of mini-Rts1 derivatives deleted for DnaA boxes in the replication origin. *Plasmid* 1989, **21**:242-246.
43. Kline BC, Kogoma T, Tam JE, Shields MS: Requirement of the *Escherichia coli* dnaA gene product for plasmid F maintenance. *J Bacteriol* 1986, **168**:440-443.
44. Ortega-Jiménez S, Giraldo-Suárez R, Fernández-Tresguerres ME, Berzal-Herranz A, Díaz-Orejas R: DnaA dependent replication of plasmid R1 occurs in the presence of point mutations that disrupt the dnaA box of oriR. *Nucleic Acids Res* 1992, **20**:2547-2551.
45. Mardanov AV, Ravin NV: Functional characterization of the repA replication gene of linear plasmid prophage N15. *Res Microbiol* 2006, **157**:176-183.
46. Martínez-Salazar J, Romero D, Girard ML, Dávila G: Molecular cloning and characterization of the recA gene of *Rhizobium phaseoli* and construction of recA mutants. *J Bacteriol* 1991, **173**:3035-3040.

doi:10.1186/1471-2180-11-158

Cite this article as: Cervantes-Rivera et al.: The replication origin of a repABC plasmid. *BMC Microbiology* 2011 **11**:158.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

