

19  
2ej. 03081



UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO

UNIDAD ACADÉMICA DE LOS CICLOS PROFESIONAL Y DE  
POSGRADO DEL COLEGIO DE CIENCIAS Y HUMANIDADES  
CENTRO DE INVESTIGACION SOBRE FIJACION  
DE NITROGENO

MECANISMOS DE VARIACION CONCERTADA  
ENTRE SECUENCIAS REITERADAS EN EL GENOMA  
DE *Rhizobium etli* CFN42.

T E S I S

QUE PARA OBTENER EL GRADO DE  
DOCTOR EN INVESTIGACION  
BIOMÉDICA BÁSICA  
P R E S E N T A :  
CESAR RODRIGUEZ SANCHEZ

CUERNAVACA, MOR.

1998.

262231

TESIS CON  
FALLA DE ORIGEN



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.

PAGINACION

DISCONTINUA

## **AGRADECIMIENTOS**

Al Centro de Investigación sobre Fijación de Nitrógeno, UNAM ; por brindarme un espacio, con el material y equipos necesarios para desarrollar esta investigación.

Al Dr. David Romero por proponer y asesorar nuevamente un proyecto interesante e importante en el área de genética molecular.

A los integrantes del jurado, Dra. Carmen Gómez Eichelmann, Dra. Alicia González Manjarrez, Dra. Esperanza Martínez Romero, Dr. Rafael Palacios de la Lama, Dr. Daniel Piñero Dalmau, Dr. Luis Servín González y Dr. David Romero Camarena por las críticas y sugerencias realizadas a la tesis.

Al CONACYT por el apoyo económico a esta investigación con los grants No 4321-N9406 y No 030355 del Programa de Apoyo a la División de Estudios de Posgrado (UNAM).

Al Dr. Rafael Palacios por la recomendación otorgada para recibir el apoyo económico del CONACYT (REF: 940030).

Al CONACYT por el apoyo económico recibido (REF: 940030) a través del programa de Apoyo a los Profesores e Investigadores para la Obtención del Doctorado del Fondo para la Creación de Cátedras Patrimoniales de Excelencia.

A Javier Rivera y a María de la Paz Salas por su apoyo técnico .

A Mónica, César Daniel y Sonia, por crecer unidos en el proyecto más interesante en la vida y formar a la familia Rodríguez Rosenblueth.

A mis padres, hermanos y a todos los integrantes de la familia.

A todos los integrantes del Departamento de Genética Molecular.

Al personal académico, técnico, laboratoristas y administrativo del CIFN, que colabora para que este centro de investigación siga adelante.

## INDICE

Resumen.....	3
Abstract.....	4
Introducción.....	5
Genoma de <i>Rhizobium</i> .....	13
Planteamiento del problema.....	20
Artículo : Multiple recombination events maintain sequence identity among members of the nitrogenase multigene family in <i>Rhizobium etli</i> .	
Resultados adicionales.....	25
Discusión y Perspectivas.....	37
Conclusiones.....	47
Bibliografía.....	48
Fig 1. Modelo de conversión génica.....	8
Fig 2. Modelo de Meselson-Rading.....	10
Fig 3. Modelo de ruptura en doble cadena y reparación.....	12
Fig 4. Estructura y dinámica del plásmido de <i>R. etli</i> CFN42.....	21
Fig 5. Mapa de los genes <i>flaA</i> y <i>flaB</i> de <i>R. melilot</i> 1021.....	26
Fig 6. Construcción de los plásmidos pCRS8 y pCRS9.....	34
Fig. 7. Recombinación mediada por RNA en <i>Saccharomyces cerevisiae</i> .....	44
Tabla I. Genes indispensables o del metabolismo, reiterados en Rhizobiaceae...	18
Tabla II. Genes simbióticos reiterados en Rhizobiaceae.....	19
Tabla III. Cepas de <i>Rhizobium</i> construidas con los plásmidos pCRS8 y pCRS9...	28
Tabla IV. Genes reiterados con alta identidad en Rhizobiaceae.....	32

## RESUMEN

Una característica distintiva del genoma de *Rhizobium* es la presencia de secuencias reiteradas, las cuales frecuentemente constituyen familias multigénicas. Comúnmente se asume que conversión génica aparente entre los elementos reiterados podría tender a variación concertada, manteniendo un alto grado de identidad en secuencia nucleotídica entre sus miembros. La operación de este mecanismo fue demostrada en este trabajo. Se emplearon diferentes construcciones genéticas para establecer el papel de conversión génica aparente como un mecanismo de homogenización entre los miembros de la familia multigénica de la nitrogenasa (*nifH*) del pSim de *R. etli* CFN42. Los resultados muestran que una inserción de 28 pares de bases dentro de la reiteración *nifHa*, puede ser corregida por eventos múltiples de recombinación, incluyendo conversión génica aparente, la cual se presenta a una frecuencia acumulativa de  $8 \times 10^{-5}$ . El proceso de corrección requiere: 1) de la presencia funcional del gene *recA*; 2) de copias adicionales del gene *nifH* para su formación y 3) tiende a un reestablecimiento preciso de la secuencia nucleotídica silvestre. Las frecuencias de conversión génica aparente en *cis* ( $2.8 \times 10^{-5}$ ) o en *trans* ( $4.8 \times 10^{-5}$ ), indican que este evento ocurre frecuentemente a través de interacciones intermoleculares. Además, la alta frecuencia observada de entrecruzamientos múltiples, sugiere que el plásmido simbiótico está comprometido repetidamente en eventos de recombinación, en una situación semejante a la recombinación de fagos o recombinación entre plásmidos pequeños, de alto número de copias.

## ABSTRACT

A distinctive characteristic of the *Rhizobium* genome is the frequent finding of reiterated sequences, which often constitute families. Interestingly, these families usually maintain a high degree of nucleotide sequence identity. It is commonly assumed that apparent gene conversion between reiterated elements might lead to concerted variation among its members. However, the operation of this mechanism has not yet been demonstrated in the Rhizobiaceae. In this work, we employed different genetic construction to address the role of apparent gene conversion as a homogenizing mechanism between members of the plasmid-located nitrogenase multigene family in *Rhizobium etli*. Our results show that a 28-bp insertion into one of the nitrogenase reiterations can be corrected by multiple recombination events, including apparent gene conversion. The correction process was dependent on the presence of both a wild-type *recA* gene and wild-type copies of the nitrogenase reiterations. Frequencies of apparent gene conversion to the wild-type nitrogenase reiterations were the same when the insertion to be corrected was located either in *cis* or in *trans*, indicating that this event frequently occurs through intermolecular interactions. Interestingly, a high frequency of multiple crossovers was observed, suggesting that these large plasmid molecules are engaging repeatedly in recombination events, in a situation akin to phage recombination or recombination among small, high-copy number plasmids.

## **INTRODUCCIÓN**

### **Evolución concertada**

La recombinación génica opera para formar y reformar los genomas de todos los organismos. Los arreglos genéticos dentro y entre las secuencias reiteradas, proveen por tanto una potente fuerza evolutiva que sirve tanto para promover diversidad como para conservar la identidad genética (Lloyd y Low, 1996 ).

En los genomas nucleares, mitocondriales y de cloroplastos se han encontrado una variedad de mecanismos de intercambios no recíprocos entre secuencias repetidas (Dover, 1987). Estos pueden causar una ganancia o pérdida en el número de secuencias repetidas e influir finalmente en la expansión de una variante reiterada a través de una población sexual (Dover,1992). La unidad de ganancia -y pérdida- puede contener un gen completo, generando familias multigénicas, o bien ser mucho más pequeña, generando parches de secuencias repetidas dentro del gen y sus secuencias control (Dover, 1992).

En plantas y animales se han descrito a las familias génicas como fluidas, porque su naturaleza está continuamente cambiando, en posición, abundancia y composición. Algunas familias contienen numerosas copias consistentes de secuencias de función desconocida (familias no-génicas), otras son transcritos de RNA funcional (RNA's ribosomales y de transferencia) o de RNA mensajero que se traducen a proteínas (globinas, histonas inmunoglobulinas, actinas y proteínas MHC) (Dover, 1982). Estos procesos de variación se presentan dentro de los cromosomas y entre los cromosomas, y son potencialmente capaces de acelerar, retardar o mantener constante el rango de diferenciación de DNA entre poblaciones



(Dover,1987).

Una característica importante de estas familias es que, sin importar su número de copias, función y distribución en el genoma (tandem o interdispersas), exhiben una inesperada homogeneidad en secuencia dentro y entre los individuos de las especies (Dover, 1982). Así, en una familia que está distribuida entre varias especies, se observa frecuentemente heterogeneidad sustancial en sus elementos entre las especies, pero una homogeneidad muy alta dentro del individuo (John y Miklos, 1988). Este patrón de variación, se le conoce como evolución concertada (Dover, 1986).

El mecanismo molecular que participa en las familias multigénicas para permitir que evolucionen concertadamente se denomina "Molecular Drive" (Dover, 1982). Este involucra a los procesos por los cuales las mutaciones son capaces de extenderse a través de una familia (homogenización) y a través de mecanismos de selección o deriva génica fijarse en una población. Dentro de ellos, se encontraron una variedad de mecanismos de recombinación que llevan a una transferencia no recíproca de DNA dentro y entre los cromosomas, como son conversión génica, entrecruzamiento desigual, transposición, replicación deslizada ("slippage") e intercambios mediados por RNA (Dover,1986).

Como mecanismo de evolución concertada la recombinación por conversión génica es importante porque : 1) mantiene el número de genes ; 2) puede tener direccionalidad y 3) actúa como un mecanismo de corrección (Li y Graur, 1991). La descripción de conversión génica se presenta en el modelo de la Fig.1. (A) Participan los cromosomas homólogos 1 y 2. Cada DNA duplex consiste de dos cadenas antiparalelas a y b, con los extremos 3' marcados por las flechas. Los genes

difieren en cuatro pares de bases. (B) Una ruptura en el gene 1 es alargada por una endonucleasa. (C) El extremo 3' de la cadena 1b invade al duplex 2, desplazando la cadena 2a dentro de una asa que pudiera extenderse indefinidamente hacia la izquierda de la cadena 1b que esta siendo reparada. (D) La cadena 1b es resintetizada para el extremo 3' por adición de bases nucleotídicas complementarias para la cadena 2b. El hueco en la cadena 1a es reparado por adición de bases complementarias para el asa en la cadena 2a. (E) El extremo 3' de la cadena 1b es ligado para el extremo 5' de la parte intacta de 1b. Formandose ahora dos puntos de entrecruzamiento. (F) Para cada punto de entrecruzamiento las cadenas cruzadas son ahora rotas y religadas como se muestran en la figura. Los sitios 2 y 3 del gene 1 han sido ahora convertidos a la secuencia del gene 2.

Una gran variedad de estudios muestran que a través de la recombinación homóloga puede obtenerse tanto entrecruzamiento (crossing-over) como conversión génica. El intercambio recíproco de un linaje a dos o más genes, se explica fácilmente por rompimiento y reunión de moléculas de DNA parental. La conversión génica se define como un reemplazamiento aparente de un gene por su alelo, y no se explica fácilmente por rompimiento y reunión. En ambos procesos de recombinación está involucrada la resolución del intermediario de Holliday. Una de los dos alternativas para resolver esta estructura causa entrecruzamiento de los genes que bordean la unión; en el otro camino la reparación de los pares de bases no complementarios en los heteroduplex causa conversión génica (Shibata *et al.*, 1995). Los modelos de recombinación caen en dos grandes categorías dependiendo de cómo ocurre el inicio de la recombinación y sobre quién se da la transferencia de información.

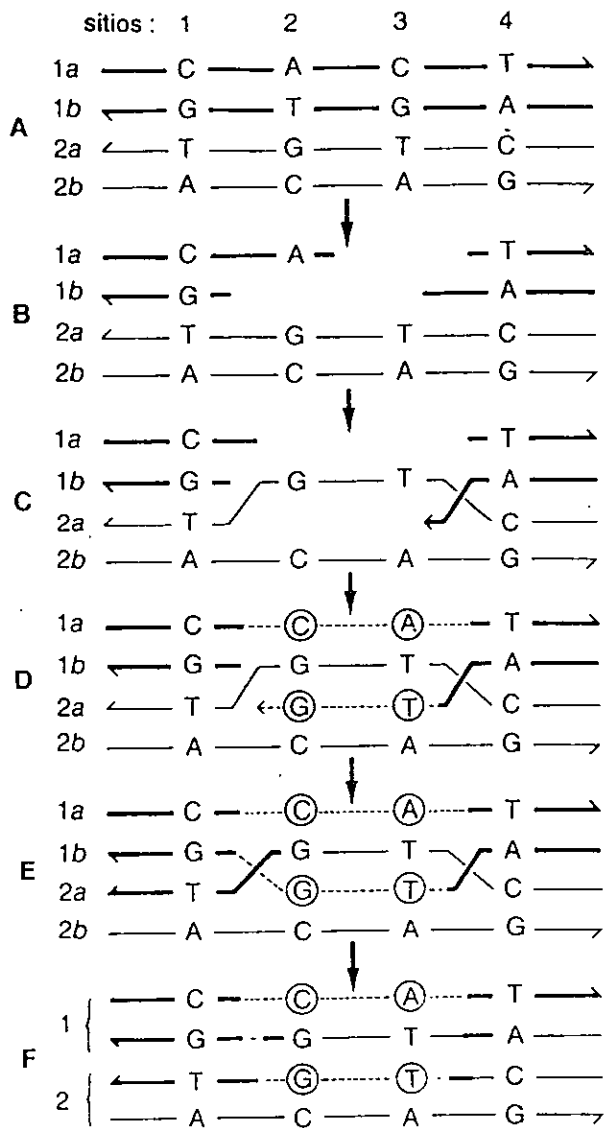


Fig. 1. Modelo de Conversión Génica.

(Tomado de Futuyma, 1986)

En la Fig. 2, se presenta el modelo de Meselson-Radding (Stahl, 1994). La recombinación se inicia cuando una endonucleasa realiza la ruptura de una cadena en el DNA duplex de la secuencia A (a). Esta cadena es desplazada del duplex de DNA por la acción de la DNA polimerasa (b). La cadena invade al DNA duplex de la secuencia B y desplaza a la cadena residente en la cadena de igual polaridad (c). La actividad de una nucleasa degrada a la cadena desplazada (d). La ligación con la secuencia A, forma estructuralmente una unión Holliday, que es inicialmente genéticamente asimétrica, ya que sólo uno de los dos duplex tiene una región de DNA heteroduplex. La unión puede desplazarse; esto lleva a un heteroduplex simétrico, estos DNA heteroduplex pueden estar en ambos duplex (e). La resolución de la unión Holliday ocurre en dirección horizontal o vertical (f). Kourilsky (1986) reporta que al realizarse la corrección del heteroduplex entre DNA's en los cuales el 8% de los pares de bases no son complementarios, detecta en la corrección del DNA heteroduplex que están involucrados bloques de secuencia, más que nucleótidos solos y algunas veces estos bloques de secuencia, generan parches de la secuencia parental.

El segundo modelo (Fig. 3), descrito por Szostak *et al.* (1983), se le ha llamado de ruptura en doble cadena y reparación (DSB). Un DNA duplex de la secuencia A es cortado por una endonucleasa (a). Se presenta degradación exonucleolítica en dirección 5' a 3', dejando expuestos los extremos 3' (b). Los extremos 3' invaden al DNA duplex de la secuencia B que es homólogo a la secuencia A, e inician la síntesis de DNA que reemplaza al DNA perdido, el resultado final es que la nueva información es colocada en una posición semejante en la secuencia A. La ligación

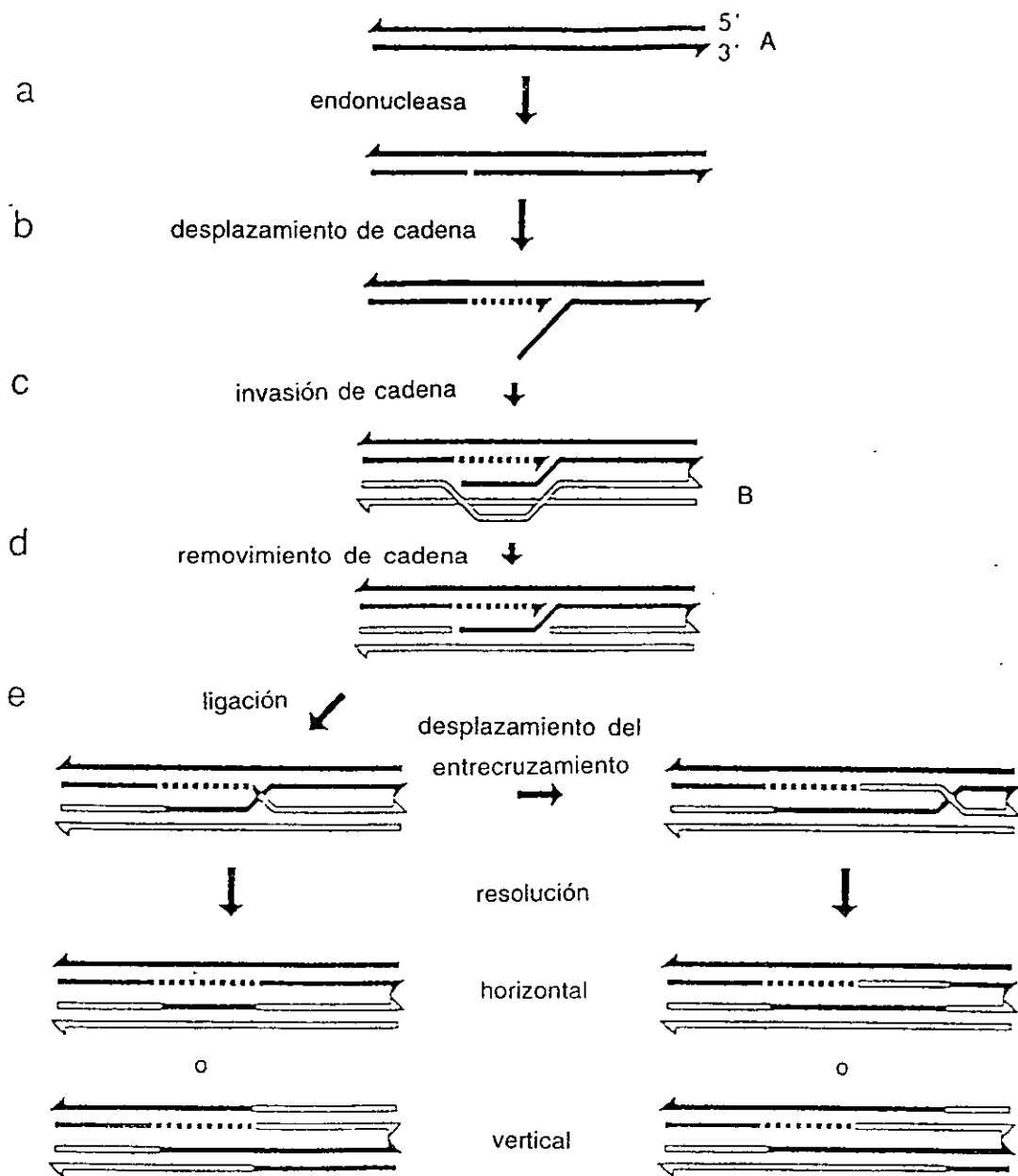


Fig 2. Modelo de Meselson-Radding

(Tomado de Stahl, 1994)

completa crea una estructura de cuatro cadenas en la cual los duplex de DNA son mantenidos juntos por un par de uniones Holliday (c). Las uniones pueden resolverse verticalmente u horizontalmente; si se resuelven en el mismo sentido no hay intercambio, pero los marcadores aledaños al sitio recombinador pueden manifestar conversión. Si se resuelven en sentidos opuestos, resulta el intercambio y hay conversión para los marcadores cercanos al sitio. Finalmente si se resuelven por topoisomerasas no hay intercambios y el producto es igual al original.

Entre las familias multigénicas en hongos filamentosos (*Ascobolus immersus*, *Neurospora crassa*, *Sordaria fumicola* y *S. brevicolis*), levaduras (*Saccharomyces cerevisiae* y *Schizosaccharomyces pombe*) y en eucariotes superiores se han descrito eventos de conversión génica responsables de evolución concertada entre secuencias repetidas arregladas en tandem o dispersas dentro de un mismo cromosoma o en cromosomas diferentes (Hastings, 1988 ; Dover, 1993).

Los procariotes habitualmente poseen un bajo numero de reiteraciones. Blattner *et al.* (1997) determinaron que un 2.5% del genoma de *E. coli* se encuentra en forma de reiteraciones. Aunque la proporción de las secuencias reiteradas sea menor con respecto a la encontrada en eucariotes, la presencia de familias multigénicas en algunos procariotes ha permitido detectar conversión génica entre los genes *rm* de *E. coli* (Harvey *et al.*, 1988), en *Acinetobacter calcoaceticus* entre los genes *catIJF* y *pcaIJF* (Kowalchuk *et al.*, 1995) y en *Salmonella typhimurium* entre los genes *tuf* (Abdulkarim y Hughes, 1996).

En procariotes solamente se han reportado altos niveles de reiteración en *Halobacterium* (Sapienza y Doolittle, 1982), *Streptomyces* (Birch *et al.*, 1990),

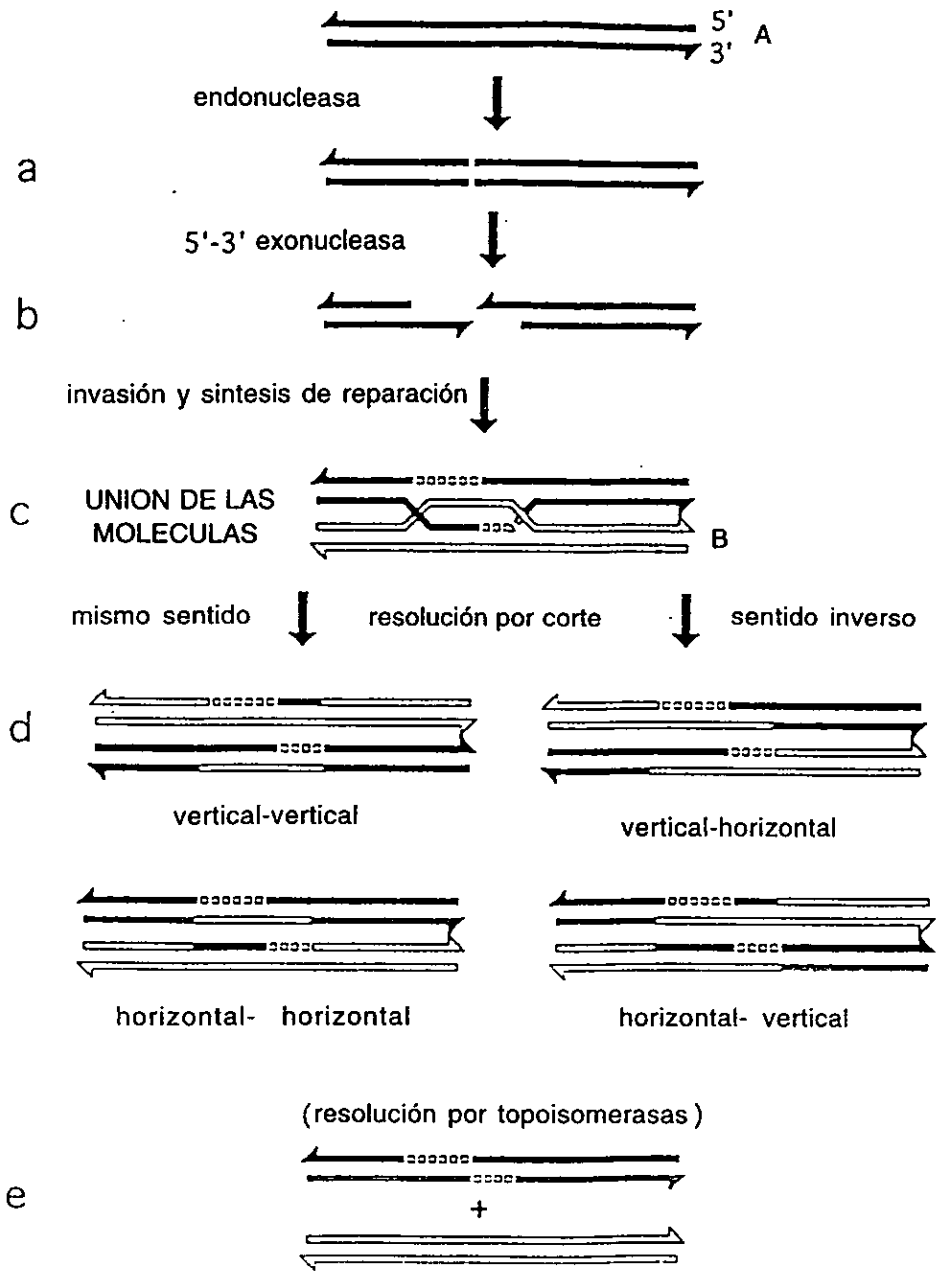


Fig 3. Modelo de ruptura en doble cadena y reparación.

(Tomado de Stahl, 1996)

*Xanthomonas* (Leach *et al.*, 1990), *Agrobacterium* y *Rhizobium* (Flores *et al.*, 1987; Flores *et al.*, 1988). Sin embargo, en estos organismos no se ha detectado entre las secuencias reiteradas la participación de alguno de los mecanismos de homogenización de secuencia. Por las características que posee *Rhizobium* y la importancia que tiene la recombinación mediante conversión génica en la evolución concertada, resulta relevante estudiar este mecanismo de recombinación en este grupo bacteriano.

### **El Genoma de *Rhizobium***

La característica biológica predominante de *Rhizobium* spp. es su capacidad para establecer simbiosis con plantas leguminosas. Debido a esa relación, la mayoría de los estudios se han enfocado al entendimiento de las interacciones con la planta (Fischer, 1994; Schultze *et al.*, 1994). Actualmente, las bacterias del género *Rhizobium* son atractivas como modelo para estudiar la estructura y dinámica del genoma, porque gran parte del genoma está en forma de plásmidos grandes que contienen una alta cantidad de secuencias reiteradas (dentro de los plásmidos y con otras regiones del genoma) y la frecuente presencia de rearrreglos genómicos.

Los plásmidos en *Rhizobium* representan entre el 25% y 50% del genoma (Long, 1989; Martínez *et al.*, 1990; Sobral *et al.*, 1991; García de los Santos *et al.*, 1996). Estos plásmidos tienen un tamaño entre 100 -1600 kb y su número varía entre dos a diez plásmidos. La mayoría de las especies de *Rhizobium* tienen genes involucrados en nodulación y fijación del nitrógeno localizados en un plásmido



*Xanthomonas* (Leach *et al.*, 1990), *Agrobacterium* y *Rhizobium* (Flores *et al.*, 1987; Flores *et al.*, 1988). Sin embargo, en estos organismos no se ha detectado entre las secuencias reiteradas la participación de alguno de los mecanismos de homogenización de secuencia. Por las características que posee *Rhizobium* y la importancia que tiene la recombinación mediante conversión génica en la evolución concertada, resulta relevante estudiar este mecanismo de recombinación en este grupo bacteriano.

### **El Genoma de *Rhizobium***

La característica biológica predominante de *Rhizobium* spp. es su capacidad para establecer simbiosis con plantas leguminosas. Debido a esa relación, la mayoría de los estudios se han enfocado al entendimiento de las interacciones con la planta (Fischer, 1994; Schultze *et al.*, 1994). Actualmente, las bacterias del género *Rhizobium* son atractivas como modelo para estudiar la estructura y dinámica del genoma, porque gran parte del genoma está en forma de plásmidos grandes que contienen una alta cantidad de secuencias reiteradas (dentro de los plásmidos y con otras regiones del genoma) y la frecuente presencia de rearrreglos genómicos.

Los plásmidos en *Rhizobium* representan entre el 25% y 50% del genoma (Long, 1989; Martínez *et al.*, 1990; Sobral *et al.*, 1991; García de los Santos *et al.*, 1996). Estos plásmidos tienen un tamaño entre 100 -1600 kb y su número varía entre dos a diez plásmidos. La mayoría de las especies de *Rhizobium* tienen genes involucrados en nodulación y fijación del nitrógeno localizados en un plásmido

grande, el plásmido simbiótico ó pSim. A los otros plásmidos se les ha denominado crípticos. García de los Santos *et al.* (1996) analizaron la información contenida en algunos de los llamados plásmidos crípticos, encontrándose que pueden portar genes relevantes para simbiosis y vida saprofítica. En algunos plásmidos "crípticos" se han localizado genes que participan en la biosíntesis de polisacáridos de superficie (exopolisacáridos y lipopolisacáridos) y secuencias relacionadas a procesos metabólicos (citrato sintasa, transporte de ácidos dicarboxílicos y biosíntesis de tiamina). Además, se ha descrito que en los plásmidos crípticos pueden encontrarse genes que participan en el catabolismo de algunos compuestos, como aquellos exudados por la planta (trigonelina, estaquidrina, carnitina y calistegina), utilización de compuestos aromáticos (ácido protocatéquico, quinato y catecol), y otros azúcares diversos (rafinosa, acetoacetato, ramnosa, sorbitol, etc.), además de genes de chaperoninas (Rusanganwa y Gupta, 1993) y genes de melanina (Cubo *et al.*, 1988).

La disponibilidad de estas funciones adicionales, contribuye a la diversidad metabólica de estos organismos, y probablemente confiere en ellos una ventaja adaptativa sobre las bacterias del suelo. Por ejemplo, en *R. etli* las mutaciones en los genes de lipopolisacáridos, causan alteraciones en la superficie celular afectando algunos aspectos de la vida libre, como la morfología, motilidad y agregación; además en simbiosis forman nódulos inefectivos (Brom *et al.*, 1992)

En *Rhizobium etli*, Flores *et al.* (1987) han estimado la presencia de 700 elementos reiterados, pertenecientes a 200 familias diferentes. Niveles similares de reiteración de DNA fueron también observados para otros miembros de la familia

*Rhizobiaceae*, como *R. meliloti* y *Agrobacterium tumefaciens*. El análisis realizado excluye a secuencias reiteradas menores de 300 pares de bases (pb). Recientemente, F. J. de Bruijn, (1993) detectó la presencia de secuencias extragénicas de 38 pb llamadas REP (Repetitive extragenic palindromic) y secuencias mayores (124 - 127 pb) referidas como ERIC (Enterobacterial repetitive intergenic consensus). Osteras *et al.*, (1995) reportaron la presencia de dos elementos intergénicos de secuencias palindrómicas de 108 pb y 109 pb, específicas para *Rhizobium*, denominadas RIME (*Rhizobium*-specific intergenic mosaic elements).

En el genoma de diferentes especies de *Rhizobium* las reiteraciones incluyen a secuencias de inserción (IS) y secuencias con alta homología similares en estructura están relacionadas a IS. Este es un componente extremadamente variado, incluyendo diferentes tipos (Dusha *et al.*, 1987, Hahn y Hennecke, 1987, Judd y Sadowsky, 1993, Kaluza *et al.*, 1985, Laberge *et al.*, 1995, Mazurier *et al.*, 1996, Perret *et al.*, 1997, Priefer *et al.*, 1980, Rice *et al.*, 1994, Ruvkun *et al.*, 1982, Soto *et al.*, 1992, Wheatcroft y Laberge, 1991, Wheatcroft y Watson, 1988). La importancia de las secuencias de inserción para definir a la clase de secuencias reiteradas ha sido puesta en relieve con la secuenciación del plásmido simbiótico de *Rhizobium* spp. NGR234. En este replicón, aproximadamente un 18% de la secuencia corresponde a secuencias de inserción (Freiberg *et al.*, 1997).

Estas secuencias pueden bordear a genes funcionales, como en *R. meliloti*, donde un ORF con similitud a IS se encuentra en el extremo 3' de *nodQ1* (Schwedock y Long, 1994), o en *A. caulinodans*, donde dos elementos repetidos diferentes

bordean a *nodD* (Geelen *et al.*, 1995). Se ha encontrado también en *R. leguminosarum* bv. *viciae* una estructura parecida a un transposón, el cual codifica para un sistema de restricción-modificación (Rochepeau *et al.*, 1997).

En ocasiones, las secuencias repetidas incluyen regiones regulatorias largas, como las relacionadas a *nif* y al extremo amino de la nitrogenasa (Better *et al.*, 1983; Watson y Schofield, 1985; Buck *et al.*, 1986). Se ha observado similitud en pequeños sectores entre los genes *nfe* involucrados en competitividad para nodulación (Soto *et al.*, 1993) y los genes *mos* que producen rhizopinas (Murphy *et al.*, 1988; Murphy *et al.*, 1993; Saint *et al.*, 1993).

Earl *et al.* (1987) al estudiar el operón *fixABCX* (relacionado al transporte de electrones), localizaron una duplicación en el extremo 5' del operón, que incluye la región regulatoria y 32 codones del gene *fixA*. Rostas *et al.* (1986) reportaron originalmente la existencia de tres cajas *nod* en *R. meliloti*, estos datos se han extendido con el hallazgo de otras cinco cajas *nod* (Baev, 1991; Baev y Kondorosi, 1992; Plazanet *et al.*, 1995; Barnett *et al.*, 1996; Balleza, 1998) involucradas en la regulación de genes de nodulación.

También se han detectado reiteraciones funcionales que constituyen familias multigénicas de genes completos y operones, involucrados en funciones esenciales y del metabolismo (Tabla I) y genes que participan en el proceso simbiótico (Tabla II). Estas secuencias están interdispersas entre los plásmidos y el cromosoma (sólo los genes *fla* de *R. meliloti* están arreglados en tandem) . La distribución no es al azar, hay secuencias localizadas solo en plásmidos y otras son exclusivas del cromosoma. Muchas de estas secuencias se presentan usualmente en una copia en otros

organismos; de hecho, solamente en *Rhizobium* se ha detectado reiteración de los genes *ftsZ*, *groEL* y *rpoH* (Fischer *et al.*, 1993; Margolin y Long, 1994; Narberhaus *et al.*, 1997; Rusanganwa y Gupta, 1993).

Finalmente hay genes reiterados con funciones relacionadas pero no idénticas. Schwedock y Long (1992) reportaron al par *nodPQ* - *saa*, donde *nodPQ* está involucrado en la sulfatación de un factor de nodulación, mientras *saa* participa en el paso de sulfatación para la biosíntesis de cisteína y metionina. El par *nodM* - *glmS*, donde *nodM* sintetiza la glucosamina necesaria para la biosíntesis del factor *nod*, mientras *glmS* elabora la glucosamina requerida por función celular (Marie *et al.*, 1992).

Presumiblemente por la presencia de secuencias reiteradas, el genoma de algunas cepas de *Rhizobium* está sujeto a frecuentes rearrreglos genómicos que constituyen una causa importante de variabilidad en la población.

En *Rhizobium etli* se han caracterizado rearrreglos genómicos en cuya generación participan reiteraciones tanto cromosomales como plasmídicas (Flores *et al.*, 1988), ocasionando entre los plásmidos eventos de cointegración (Brom *et al.*, 1991) y translocación (Martínez-Salazar, 1991); a su vez se han detectado amplificaciones y deleciones entre secuencias específicas (Romero *et al.*, 1991; Romero *et al.*, 1995).

En algunas especies de *Rhizobium* se han identificado amplicones (regiones específicas de DNA que tienen el potencial para ser amplificadas o deletadas como una unidad) tanto en el cromosoma como en grandes plásmidos (Flores *et al.*, 1993; Romero *et al.*, 1995; Stepkowski *et al.*, 1993; Mavingui *et al.*, 1998).

Tabla I. Genes indispensables o del metabolismo, reiterados en Rhizobiaceae.

GENES	ESPECIE	NUMERO DE REITERACIONES Y LOCALIZACION	AUTORES
<i>flaAyB</i>	<i>Sinorhizobium meliloti</i> , cepas 1021 y 10406	2, ambas en el cromosoma	Pleier y Schmitt, 1989; Bergmann <i>et al.</i> , 1991
<i>ftsZ</i>	<i>Sinorhizobium meliloti</i> 1021	2, ambas en el cromosoma	Margolin <i>et al.</i> , 1991; Margolin y Long, 1994
<i>groEL</i>	<i>Sinorhizobium meliloti</i> 1021	5, hay 2 en cada megaplásmido y 1 en el cromosoma	Rusanganwa y Gupta, 1993; Ogawa y Long, 1993.
<i>groEL</i>	<i>Bradyrhizobium japonicum</i> 110 spc4	5, en el cromosoma	Fischer <i>et al.</i> , 1993
<i>rpoH</i>	<i>Bradyrhizobium japonicum</i> 110 spc4	3, en el cromosoma	Narberhaus <i>et al.</i> , 1997
<i>rpoN</i>	<i>Bradyrhizobium japonicum</i> 110 spc4	2, en el cromosoma	Küllik <i>et al.</i> , 1991
<i>rm</i>	En varias especies de Rhizobia	Variable, dependiendo de la especie (2-3) se asumen en el cromosoma	Kündig <i>et al.</i> , 1995
<i>csA</i>	<i>Rhizobium tropici</i> CFN299	2, una en cromosoma y otra en el pSim	Pardo <i>et al.</i> , 1994; Hernández-Lucase <i>et al.</i> , 1995

Las siglas de los genes corresponden a: *fla* = flagelina; *fts* = inicio del septo de división celular; *gro* = chaperoninas; *rpo* = factor sigma ; *rmn* = ribosomales y *csA* = citrato sintasa

Tabla II. Genes simbióticos reiterados en Rhizobiaceae.

GENES	ESPECIE	NUMERO DE REITERACIONES Y LOCALIZACION	AUTORES
<i>fixK</i>	<i>R.leguminosarum</i> bv. <i>viceae</i> VF39	2, una en el pSim y otra en el plásmido C.	Schlüter <i>et al.</i> , 1997.
<i>fixK</i>	<i>Rhizobium etli</i> CFN42	2, una en el pSim y otra en el plásmido F.	Girard, comunicación personal
<i>fixK</i>	<i>Bradyrhizobium japonicum</i> 110spc4	2, en el cromosoma	Anthamatten <i>et al.</i> , 1992
<i>fixK</i>	<i>Sinorhizobium meliloti</i> 2011	2, en plásmido Sim	Renalier <i>et al.</i> , 1987
<i>nifH</i>	<i>Rhizobium etli</i> CFN42	3, en el plásmido Sim	Quinto <i>et al.</i> , 1985
<i>nifH</i>	<i>Azorhizobium caulinodans</i> ORS571	2, en el cromosoma	Norel y Elmerich, 1987
<i>nifH</i>	<i>Rhizobium</i> sp. NGR234	2, en el plásmido Sim	Badenoch-Jones <i>et al.</i> , 1991
<i>fixN</i>	<i>Sinorhizobium meliloti</i> 2011	2, en el plásmido	David <i>et al.</i> , 1987
<i>fixNOPQ</i>	<i>R.leguminosarum</i> bv. <i>viceae</i> VF39	2, uno en el pSim y otro en plásmido críptico	Schlüter <i>et al.</i> , 1997
<i>fixNOPQ</i>	<i>Rhizobium etli</i> CFN42	2, uno en el pSim y la otra en el pf.	Girard, comunicación personal
<i>nodD</i>	<i>Sinorhizobium meliloti</i> , cepas 1021 y 41	3, en el plásmido Sim	Honma y Ausubel, 1987
<i>nodPQ</i> - <i>saa</i>	<i>Sinorhizobium meliloti</i> 2011	en el plásmido Sim	Schwedock y Long, 1992
<i>nodM</i> - <i>glmS</i>	<i>R.leguminosarum</i> bv. <i>viceae</i> 8401	en el plásmido Sim	Marie <i>et al.</i> , 1992

Las siglas corresponden a: *fix* y *nif* = fijación de nitrógeno; *nod* = nodulación.

En el plásmido simbiótico de *R. etli* se han detectado cuatro amplicones en la región que contiene los genes simbióticos. Su tamaño varía de 90 a 175 Kb (Fig. 4) y los eventos de recombinación que participan son dependientes del gen *recA* (Romero *et al.*, 1995). También ocurren a alta frecuencia eventos que translocan el amplicón *nod-nif* del pSim al plásmido a (pa) y cointegraciones entre el pSim-pa y pSim-plásmido b (Romero *et al.*, 1998). En *R. leguminosarum* bv. trifolii también se han detectado amplicones (Stepkowski *et al.*, 1993). Mavingui *et al.* (1997) han generado amplicones artificiales al azar (RDA) en *R. tropici* y a su vez en la misma especie ha estudiado un amplicón *nod-nif* (Mavingui *et al.*, 1998). Balleza (1998) mediante el análisis de secuencia nucleotídica, identificó secuencias con alta homología (reiteraciones) y teóricamente propuso la presencia de seis amplicones en el regulón *nod-nif* de *R. meliloti* y ocho amplicones en el pSim de *R. spp* NGR234.

### **Planteamiento del Problema**

El hecho de que en *Rhizobium* haya reiteraciones que participan en la dinámica del genoma abre la posibilidad de que algunos de los mecanismos de "molecular drive" estén actuando. Se conoce que la familia multigénica de la nitrogenasa participa activamente en recombinación y además entre sus miembros se presenta una alta homogenidad en secuencia nucleotídica dentro de la especie y entre los individuos de la especie, lo que parece implicar variación concertada para esta familia.



En el plásmido simbiótico de *R. etli* se han detectado cuatro amplicones en la región que contiene los genes simbióticos. Su tamaño varía de 90 a 175 Kb (Fig. 4) y los eventos de recombinación que participan son dependientes del gen *recA* (Romero *et al.*, 1995). También ocurren a alta frecuencia eventos que translocan el amplicón *nod-nif* del pSim al plásmido a (pa) y cointegraciones entre el pSim-pa y pSim-plásmido b (Romero *et al.*, 1998). En *R. leguminosarum* bv. trifolii también se han detectado amplicones (Stepkowski *et al.*, 1993). Mavingui *et al.* (1997) han generado amplicones artificiales al azar (RDA) en *R. tropici* y a su vez en la misma especie ha estudiado un amplicón *nod-nif* (Mavingui *et al.*, 1998). Balleza (1998) mediante el análisis de secuencia nucleotídica, identificó secuencias con alta homología (reiteraciones) y teóricamente propuso la presencia de seis amplicones en el regulón *nod-nif* de *R. meliloti* y ocho amplicones en el pSim de *R. spp* NGR234.

### **Planteamiento del Problema**

El hecho de que en *Rhizobium* haya reiteraciones que participan en la dinámica del genoma abre la posibilidad de que algunos de los mecanismos de "molecular drive" estén actuando. Se conoce que la familia multigénica de la nitrogenasa participa activamente en recombinación y además entre sus miembros se presenta una alta homogenidad en secuencia nucleotídica dentro de la especie y entre los individuos de la especie, lo que parece implicar variación concertada para esta familia.

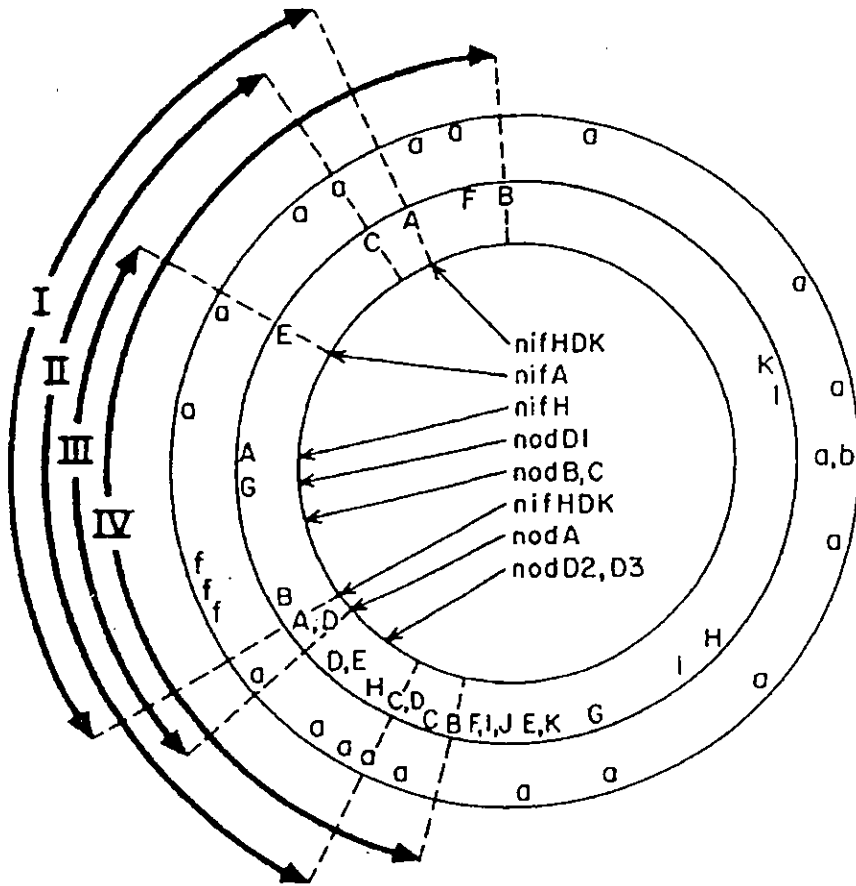


Fig 4. Estructura y dinámica del plásmido simbiótico de *R. etli* CFN42. La estructura general del pSim (390 Kb) está presentada en tres círculos concéntricos; 1 Kb = 0.93°. La posición de algunos genes simbióticos está indicada por flechas en el círculo interno. El círculo siguiente indica la posición de las familias reiteradas de DNA. Cada familia está indicada por una letra (A-K). El círculo externo indica regiones del pSim que tienen similitud con otros replicones de la cepa: pa, pb y pf. Las estructuras marcadas con I - IV indican los cuatro amplicones detectados en el pSim. (Tomado de Romero *et al.*, 1998a).

En particular, los sitios de recombinación del amplicón I del pSim de *R. etli* CFN42 (Fig. 4) son los genes de la familia multigénica de la nitrogenasa. Romero *et al.* (1991) han detectado eventos de amplificación ( $10^{-3}$ ) y deleción ( $10^{-4}$ ) entre los operones *nifHDK*. La familia *nifH* está compuesta de dos reiteraciones de 5Kb localizadas en orientación directa, las cuales comprenden a los operones *nifHDK* (regiones *nifH* a y b) separados por 120 Kb. El tercer elemento de la familia de la nitrogenasa es una reiteración de 1.5Kb que esta localizada a la mitad del amplicón (región *nifHc*); consiste del gen *nifHD\** (*nifD\** es un gene truncado que comprende los primeros 70 aminoácidos de la proteína NifD). La región *nifHc* esta en orientación invertida con respecto a los operones *nifHDK*. Este arreglo de los operones *nifHDK* permite estudiar la dinámica recombinacional de esta familia multigénica plasmídica.

En *R. etli* CFN42, los miembros de esta familia multigénica mantienen 100% de identidad en secuencia para las tres reiteraciones *nifH* (Quinto *et al.*, 1985). También *Azorhizobium caulinodans* posee dos reiteraciones *nifH* idénticas en secuencia (Norel y Elmerich, 1987) y *Rhizobium spp.* NGR234 posee dos operones *nifHDK* idénticos en secuencia (Badenoch-Jones *et al.*, 1989). A pesar de la divergencia entre estas familias, al comparar los genes *nifH* entre *R. etli* y *Rhizobium spp.* NGR234, la similitud es mantenida y presentan una identidad de 87 %, pero los miembros dentro de cada especie son idénticos.

¿Cómo explicar este nivel de conservación de secuencia?

Se podría argumentar que el origen de estas reiteraciones fue reciente y que ha ocurrido una rápida dispersión de los genes *nifH* entre los miembros de la especie. Otra alternativa es que las interacciones recombinacionales (como son

entrecruzamientos sencillos y dobles entrecruzamientos o conversión génica aparente) entre los elementos reiterados pudieran mantener variación concertada entre estos miembros. Sin embargo, la operación de estos mecanismos no ha sido demostrada para organismos de la familia de las Rhizobiaceas.

La característica de *Rhizobium* al poseer un genoma dinámico, abre la posibilidad de estudiar alguno de los mecanismos de recombinación homóloga que participen en la homogenización de una familia multigénica. Desde el punto de vista de la evolución concertada, la conversión génica ofrece un mecanismo interesante de transferencia no recíproca de DNA que permite evaluar su capacidad para homogenizar o corregir secuencias de DNA reiterado interdisperso en el cromosoma o plásmidos. Debido a lo anterior, nos enfocaremos de manera particular a:

**1. Establecer si el proceso de conversión génica opera entre las secuencias reiteradas *nifH* localizadas en el plásmido simbiótico de *R. etli* CFN42.**

Los resultados con respecto al objetivo anterior se encuentran en el artículo "Multiple recombination events maintain sequence identity among members of the nitrogenase multigene family in *Rhizobium etli*", que se presenta a continuación y en la parte correspondiente a resultados adicionales se presenta un análisis teórico para encontrar evidencia de la presencia del mecanismo de conversión génica en otra familia multigénica en *Rhizobium*.

Otro aspecto interesante en recombinación, es el incremento de

recombinación homóloga mediante rupturas en doble cadena. Así, al introducir un sitio de restricción específico en un miembro de la familia *nifH*, la recombinación iniciaría en ese sitio y consecuentemente ocurriría la reparación de esta zona mediante conversión génica. Estas dos características, son un punto interesante de evaluar, por estas razones se va a:

**2.Explorar el efecto de ruptura en doble cadena sobre la dinámica de conversión génica para la familia multigénica *nifH*.**

Los avances con respecto al segundo objetivo se presentan en la parte de resultados adicionales.

Multiple Recombination Events Maintain Sequence Identity Among Members of the Nitrogenase Multigene Family in *Rhizobium etli*

César Rodríguez and David Romero

Departamento de Genética Molecular, Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

Manuscript received December 3, 1997

Accepted for publication March 2, 1998

## ABSTRACT

A distinctive characteristic of the *Rhizobium* genome is the frequent finding of reiterated sequences, which often constitute multigene families. Interestingly, these families usually maintain a high degree of nucleotide sequence identity. It is commonly assumed that apparent gene conversion between reiterated elements might lead to concerted variation among its members. However, the operation of this mechanism has not yet been demonstrated in the Rhizobiaceae. In this work, we employed different genetic constructions to address the role of apparent gene conversion as a homogenizing mechanism between members of the plasmid-located nitrogenase multigene family in *Rhizobium etli*. Our results show that a 28-bp insertion into one of the nitrogenase reiterations can be corrected by multiple recombination events, including apparent gene conversion. The correction process was dependent on the presence of both a wild-type *recA* gene and wild-type copies of the nitrogenase reiterations. Frequencies of apparent gene conversion to the wild-type nitrogenase reiterations were the same when the insertion to be corrected was located either in *cis* or in *trans*, indicating that this event frequently occurs through intermolecular interactions. Interestingly, a high frequency of multiple crossovers was observed, suggesting that these large plasmid molecules are engaging repeatedly in recombination events, in a situation akin to phage recombination or recombination among small, high-copy number plasmids.

GENE duplication, leading to the formation of multigene families in eukaryotes, has been widely recognized as a mechanism for the generation of new functions (OHNO 1970; JOHN and MIKLOS 1988). Interestingly, members of multigene families tend to vary in a concerted way, keeping a high similarity between their members (JOHN and MIKLOS 1988; DOVER 1993). Concerted evolution between members of tandemly-arranged multigene families in eukaryotes has been amply documented. Conservation in a nucleotide sequence is thought to occur predominantly through frequent unequal exchanges between its members (PETES 1980; SZOSTAK and WU 1980; WILLIAMS and STROBECK 1985). A consequence of this mechanism is the frequent expansion and contraction of the tandem array. However, recent determinations show that an alternate mechanism, gene conversion, plays a major role in achieving homogenization in tandem multigene families (GANCLOFF *et al.* 1996). Gene conversion has been defined as the non-reciprocal transfer of sequence information between homologous or homocologous DNA sequences. Frequent events of gene conversion are also responsible for concerted evolution between members of dispersed gene families (JACKSON and FINK 1981; KLEIN and PETES

1981). Unequal exchanges are less adequate to achieve homogenization in these cases, due to the high likelihood of rearrangement of single-copy DNA flanked by dispersed members.

Studies about the occurrence of similar processes in prokaryotes have been hindered by the paucity of reiterated elements in enterobacterial genomes. A recent determination, based on the sequence of the whole *Escherichia coli* genome (BLATTNER *et al.* 1997), shows that reiterated elements constitute about 2.5% of the genome. The most conspicuous families of long reiterated elements in *E. coli* and *Salmonella typhimurium* are the *rm* operons, *tuf* genes, and different types of insertion sequences (BACHELLIER *et al.* 1996; DEONIER 1996; BLATTNER *et al.* 1997), which are commonly arranged as dispersed reiterations. Typically, a high level of nucleotide sequence identity is observed among members of each family. Several reports indicate that homogenization between the reiterated *rm* operons in *E. coli* (HARVEY and HILL 1990), duplicated flagellin genes (OKAZAKI *et al.* 1993) or the *tuf* reiterations in *S. typhimurium* (ABDULKARIM and HUGHES 1996) may be achieved through apparent gene conversion.

Extensive DNA reiteration is found in the genomes of bacteria belonging to the symbiotic nitrogen-fixing genus *Rhizobium* (reviewed by ROMERO *et al.* 1997). These genomes may carry as much as 700 representatives of long, reiterated elements, belonging to 200 different families (FLORES *et al.* 1987). This high level of

Corresponding author: David Romero, Depto. de Genética Molecular, Centro de Investigación sobre Fijación de Nitrógeno-UNAM, Apartado Postal 565-A, 62100 Cuernavaca, Morelos, México. E-mail: dromero@cfm.unam.mx

Genetics 149: 7413-7420 (June, 1998)

AUTHOR: Please answer

reiteration has been observed both in the chromosomal and in the large plasmids that are typical of this genus (FLORES *et al.* 1987; GIRARD *et al.* 1991). In fact, an analysis of the sequence of the whole symbiotic plasmid (pSym, 536.1 kb) of *Rhizobium* sp. NGR234 reveals that reiterated elements constitute about 18% of the plasmid genome (FREIBERG *et al.* 1997).

Besides transposable elements, multigenic families have been observed for several housekeeping genes, such as the *fla* (BERGMAN *et al.* 1991), *flz* (MARGOLIN and LONG 1994), *groEL* (FISCHER *et al.* 1993), *rpoN* (KÜNDIG *et al.* 1993) and citrate synthase genes (PARDO *et al.* 1994; HERNÁNDEZ-LUCAS *et al.* 1995). Several genes involved solely in the symbiotic process are also reiterated, such as *nifHDK* (BADENOCII-JONES *et al.* 1989; NOREL and ELMERICH 1987; QUINTO *et al.* 1985), *fixN* (DAVID *et al.* 1987; SCHLÖTER *et al.* 1997) and the *nodD* genes (SCHULTZE *et al.* 1994). The reiteration mode is usually dispersed; only the *fla* gene reiterations are arranged in tandem (BERGMAN *et al.* 1991).

High levels of nucleotide identity are common between these reiterations. A considerable fraction of the reiterated class (about 70%) in the pSym of *Rhizobium* sp. NGR234 is comprised of identical reiterations (FREINERG *et al.* 1997). Other multigenic families, such as *fla* (BERGMAN *et al.* 1991), *nifHDK* (BADENOCII-JONES *et al.* 1989; NOREL and ELMERICH 1987; QUINTO *et al.* 1985) and the citrate synthase genes (PARDO *et al.* 1994; HERNÁNDEZ-LUCAS *et al.* 1995) also exhibit high identities (over 95%) in nucleotide sequence. A trivial explanation for the occurrence of such high levels of identity would be a relatively recent evolutionary origin. This explanation seems unlikely, because (i) in some cases, such as the *nifHDK* family of *Rhizobium etli*, the sequences are ancient enough to be present in every member of the species (SEGOVIA *et al.* 1993), and (ii) although high levels of nucleotide sequence identity are seen between reiterations within a species, a lower level is seen between species themselves. These data have led to the suggestion that some homogenizing mechanism, conceivably gene conversion, is operating between reiterated sequences in *Rhizobium*.

So far, no studies have been published about the occurrence of gene conversion-like events between reiterations in *Rhizobium*. The only data concerning this phenomenon were obtained during a study of phage crosses in *Rhizobium meliloti*, where apparent gene conversion is claimed to occur at a low frequency (ORNSZ *et al.* 1980), close to the one for spontaneous mutation. If the low frequency observed also applies to recombinational interactions between reiterations, apparent gene conversion would be an inefficient homogenizing mechanism in *Rhizobium*.

In this work, we address the role of apparent gene conversion as an homogenizing mechanism between members of the nitrogenase multigenic family in *R. etli*. Our results indicate that multiple recombination events,

including apparent gene conversion, play an important role in maintaining sequence identity among members of this family.

## MATERIALS AND METHODS

**Bacterial strains, plasmids and media:** The bacterial strains and plasmids employed are listed in Table 1. *E. coli* strains were grown at 37° in LB medium (MILLER 1972), and *R. etli* strains were grown at 30° in PY medium (NOEL *et al.* 1984). Antibiotics were added at the following concentrations: kanamycin (Km), 30 µg ml<sup>-1</sup> (*E. coli*) or at variable concentrations depending on the purpose (*R. etli*, see below); nalidixic acid (Nal), 20 µg ml<sup>-1</sup> (*R. etli*); spectinomycin (Sp), 100 µg ml<sup>-1</sup> (*E. coli* or *R. etli*); and tetracycline (Tc), 5 to 10 µg ml<sup>-1</sup> (*R. etli*) or 10 µg ml<sup>-1</sup> (*E. coli*).

**Plasmid construction:** All DNA manipulations were carried out under standard protocols (SAMBRINK *et al.* 1989) using restriction enzymes, bacterial alkaline phosphatase and T4 polynucleotide ligase from Amersham Corp. (Arlington Heights, IL). To construct a hybrid Km<sup>r</sup>/Sp<sup>r</sup> cassette, a 1074-bp fragment, containing the Sp<sup>r</sup> gene from pHP45- $\Delta$ Sp (Table 1) was obtained through PCR amplification employing oligonucleotides o4 and o5 (Table 2); these oligonucleotides contain BglII sites. This PCR product was digested with BglII and ligated into BamHI-restricted pSUP5011 (a plasmid containing Tn3-mob. SIMON 1984), replacing the 1.6-kb mob fragment. The resulting plasmid, which carries a Tn3 Km<sup>r</sup>/Sp, was called pCRS1. From pCRS1, digestion of the BglII sites already existing in the IS50s led to the release of a 3814-bp BglII fragment, containing a promoterless Km<sup>r</sup> gene and the Sp<sup>r</sup> gene with its own promoter.

To facilitate further subcloning steps of this fragment, the polylinker-containing vector pIC20H (MARSH *et al.* 1984) was digested with SacI and religated, removing a segment carrying XbaI, BglII, XbaI, EcoRV, ClaI and EcoRI sites from the polylinker; this plasmid was called pCRS2. The 3814-bp BglII fragment from pCRS1 was ligated into the unique BamHI site of pCRS2, generating pCRS3. From pCRS3, the promoterless Km<sup>r</sup> gene and the Sp<sup>r</sup> gene with its own promoter were excised as a 3864-bp HindIII fragment, through digestion of two HindIII sites on the polylinker.

To generate a transcriptional fusion between this fragment and the *nifD* gene from *R. etli*, plasmid pEM15 (MURPHY *et al.* 1988) was used. This plasmid is a derivative of pSUP205 (SIMON *et al.* 1983), containing one of the *nifHDK* operons of *R. etli* as an EcoRI fragment (*nifHDK* region 2). Ligation of the 3864 bp HindIII fragment from pCRS3 into one of the HindIII sites present in pEM15 provokes an interruption of the *nifD* coding sequence (codon 139) creating, in the proper orientation, a transcriptional fusion between *nifD* and the promoterless Km<sup>r</sup> gene. The resulting plasmid was called pCRS4.

To generate a polar insertion in *nifH*, a 28-bp double-stranded oligonucleotide (I-SceI, Table 2) with overhanging, compatible BglII ends, was ligated into the unique BglII site of pCRS4, thus interrupting the *nifH* coding sequence (codon 147); this plasmid was named as pCRS5. Finally, to generate plasmid derivatives able to replicate in *R. etli*, plasmids pCRS4 and pCRS5 were digested with EcoRI and the resulting fragments were ligated separately onto the unique EcoRI site of the broad-host range plasmid pRK7813 (JONES and GUTTERSON 1987); the resulting plasmids were called pCRS6 and pCRS7, respectively.

**Construction of *Rhizobium etli* strains carrying the *nifH::I-SceI* and *nifD::Km/Sp* alleles:** Introduction of the *nifH::I-SceI* and *nifD::Km/Sp* alleles into *R. etli* was carried out by an in

TABLE 1  
Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant properties	Source or reference
<i>Rhizobium etli</i>		
CFN42	Wild-type strain	QUINHO <i>et al.</i> (1985)
CE3	Sm <sup>r</sup> derivative from CFN42	NOVI <i>et al.</i> (1984)
CFNX89	CE3 derivative lacking the pSym	BRUM <i>et al.</i> (1992)
CFNX236	CE3 derivative carrying a <i>nifD::Km/Sp</i> allele on the pSym	This study
CFNX237	CE3 derivative carrying the <i>nifH::I-SceI</i> and <i>nifD::Km/Sp</i> allelic combination on the pSym	This study
CFNX238	CE3 derivative carrying plasmid pCRS6	This study
CFNX239	CE3 derivative carrying plasmid pCRS7	This study
CFNX240	CFNX89 derivative carrying plasmid pCRS6	This study
CFNX241	CFNX89 derivative carrying plasmid pCRS7	This study
CFNX242	A <i>recA::ΩCm</i> derivative from strain CFNX237	This study
<i>Escherichia coli</i>		
HB101	Host for recombinant plasmids	BOYER and ROULIANI-DUSSOIX (1969)
S17-1	C600::RP-4-2 (Tc::Mu) (Km::Tn7)	SIMON (1984)
Plasmids		
pEM15	A pSUP205 derivative carrying <i>nifHDK</i> region from <i>R. etli</i> CFN42	MORETT <i>et al.</i> (1988)
pHP45ΩSp	Source of the ΩSp cassette	PRENTKI and KRISCH (1984)
pIC20H	Polylinker-containing vector	MARSH <i>et al.</i> (1984)
pMC71A	Plasmid carrying a constitutive <i>nifA</i> gene from <i>Klebsiella pneumoniae</i> , Cm <sup>r</sup>	BUCHANAN-WOLLASTON <i>et al.</i> (1981)
pMS22	Suicide vector carrying the <i>recA::ΩCm</i> allele from <i>R. etli</i>	J. MARTÍNEZ-SALAZAR, unpublished results
pRK2013	Conjugation helper	FIGURSKI and HELINSKI (1979)
pRK7813	Broad-host range plasmid, Tc <sup>r</sup>	JONES and CUTTERSON (1987)
pSUP205	Suicide vector, Cm <sup>r</sup> , Tc <sup>r</sup>	SIMON <i>et al.</i> (1983)
pSUP5011	A pSUP205 derivative carrying <sup>lc</sup> <del>ΩD-mob</del>	SIMON (1984) ← lc
pCRS1	Intermediate plasmid carrying a hybrid Tn5 Km/Sp	This study
pCRS2	Derivative from pIC20H with a deletion in the polylinker	This study
pCRS3	Derivative from pCRS2 containing a promoterless Km <sup>r</sup> gene and the Sp <sup>r</sup> gene with its own promoter	This study
pCRS4	Derivative from pEM15 carrying the <i>nifD::Km/Sp</i> allele	This study
pCRS5	Derivative from pCRS4 carrying the <i>nifH::I-SceI</i> and <i>nifD::Km/Sp</i> allelic combination	This study
pCRS6	Derivative from pRK7813 carrying the <i>nifD::Km/Sp</i> allele	This study
pCRS7	Derivative from pRK7813 carrying the <i>nifH::I-SceI</i> and <i>nifD::Km/Sp</i> allelic combination	This study

*in vivo* gene replacement procedure (SIMON *et al.* 1983). To that end, plasmid pCRS4 was introduced by transformation into *E. coli* S17-1 and the transformants were mated with *R. etli* CE3 as a recipient. Double recombinants were selected as Nal<sup>r</sup> Sp<sup>r</sup> Tc<sup>r</sup> transconjugants. To verify that the desired gene replacement has occurred, double recombinants were analyzed by Southern blot hybridization against the appropriate *nif* and ΩSp probes. This procedure yielded strain CFNX236, which carries the *nifD::Km/Sp* allele in *nif* region. A. Introduction of the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination was carried out in the same way, but employing pCRS5 as the donor. To ensure coinheritance of both markers, double recombinants were analyzed by PCR amplification using oligonucleotides o2 and o3 (see Table 2 and Figure 1); this primer pair only yielded a PCR product upon integration of the *nifH::I-SceI* allele. A strain containing both the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination in *nif* region was called

CFNX237. A *recA::ΩCm* derivative from strain CFNX237 was generated by a homogenization procedure developed in our laboratory (MARTÍNEZ-SALAZAR *et al.* 1991; ROMERO *et al.* 1995), using pMS22 as the source of the *recA* allele. This procedure gave strain CFNX242.

To generate *R. etli* derivatives carrying these allelic combinations onto small, self-replicating plasmids, *E. coli* S17-1 derivatives carrying pCRS6 or pCRS7 were mated with either *R. etli* CE3 (pSym<sup>+</sup>) or *R. etli* CFNX89 (pSym<sup>-</sup>) as recipients. Transconjugants were selected by their resistance to both nalidixic acid and spectinomycin. These crosses produced strains CFNX238 to CFNX241 (see Table 1).

PCR amplification and nucleotide sequencing: PCR amplifications were carried out using AmpliTaq DNA polymerase in a DNA Thermal Cycler 480 (Perkin-Elmer, Inc., Norwalk, CT). PCR conditions consisted of 30 cycles of 92° for 1 min, 56° for 1 min, and 72° for 1 min, except for amplifications



TABLE 2  
Oligonucleotides used in this study

Name	Sequence	Source: complementary nucleotides	Accession no.
o1	5'-GAGGACCTGCTCAAGCCCGCTAC-3'	<i>nifH</i> : 423-446	M10587
o2	5'-CGCTAGGGATAACACGGTAATATA-3'	<i>I-SceI</i> : 1-24	See below
o3	5'-CATCTTCCTGAGCTCGCCGCTGCTG-3'	<i>nifH</i> : 878-855	M10587
o4	5'-CAAGATCTCCTGATAGTTTGGCTGTGAG-3'	ΩSp: 1784-1756	M60473
o5	5'-TAAGATCTCAGTGGCCGTTTTCATGGCT-3'	ΩSp: 712-740	M60473
<i>I-SceI</i>	5'-CATCCCTAGGGATAACACGGTAATATA-3' 3'-GCCATCCCTATTGCCAATTATCTAG-5'	Not applicable*	Not applicable*

Nucleotides corresponding to the *IgfII* site in o4 and o5 are underlined.

\* Commercial oligonucleotide (Boehringer-Mannheim, Germany).

employing primer pairs o1-o4 and o2-o4, that were done by 30 cycles of 95° for 1 min, 55° for 1 min, and 72° for 2 min. PCR products for nucleotide sequencing were purified using Centri-Sep spin columns (Princeton Separations Inc., Adelphia, NJ). Nucleotide sequencing was performed with an Applied Biosystems Inc. model 373A automated DNA sequencer and a Taq Dye-Deoxy Terminator cycle sequencing kit as specified by the manufacturer (Applied Biosystems Inc., Foster City, CA).

Filter blot hybridization and determination of plasmid profiles: Genomic DNA was digested with *Bam*HI, electrophoresed in 1% agarose gels, blotted onto nitrocellulose (Hybond N+), and hybridized under stringent conditions using Amersham's Rapid-hyb buffer as specified by the manufacturer (Amersham Corp.). Plasmid profiles were obtained by an in-gel lysis method (ECKHARDT 1978), blotted onto nitrocellulose and hybridized similarly. Hybridizations with oligonucleotide probes were done in a sodium chloride-sodium citrate solution using standard procedures (AUSUNEL *et al.* 1987). Most probes were linearized and labelled with <sup>32</sup>P-α-CTP by a random priming procedure (FELMNERG and VOGELSTEIN 1983) using a Rediprime DNA labelling system (Amersham Corp.). Oligonucleotide probes were labelled with <sup>32</sup>P-γ-ATP using T4 polynucleotide kinase.

Molecular characterization of *Km<sup>r</sup>* derivatives: To ascertain the molecular events leading to the formation of each *Km<sup>r</sup>* derivative, single-colony isolates were initially screened for the presence and location of the *nifH::I-SceI* allele. To that end, genomic DNA of each isolate was subjected to PCR employing either the o2-o3 primer pair or the o1-o3 pair (Figure 1). Primer o2 has a sequence that matches the *nifH::I-SceI* allele, while primers o1 and o3 bind to specific points in the *nifH* sequence (Table 2). Thus, reactions with the o2-o3 pair gave an amplified product only if the *nifH::I-SceI* allele was still present, while those with pair o1-o3 served as a positive control for the PCR. To determine the location of the *nifH::I-SceI* allele in the *Km<sup>r</sup>* derivatives that still carried this allele, further PCR amplifications were made with primer pairs o2-o4 and o1-o4 (Figure 1 and Table 2). The first pair of primers gave a positive control for the *nifH::I-SceI* allele, while the second pair provided a positive control for these reactions.

*Km<sup>r</sup>* derivatives were also characterized by determining pSyn size in Eckhardt-type gels. This analysis allows us to distinguish wild-type plasmids (390 kb) from amplified (510 kb) or deleted (270 kb) derivatives, as described previously (ROMERO *et al.* 1991; ROMERO *et al.* 1995). Additionally, genomic DNA of each derivative was digested with *Bam*HI, blotted, and probed against a *nifH*-specific probe. Under these condi-

tions, strain CFNX237 shows three hybridizing bands of 13 kb (*nif* region a harboring the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination), 5.6 kb (*nif* region b) and 4.5 kb (*nif* region c). Derivatives harboring a tandem amplification preserve the same three bands, but show an additional 8.8-kb band, representing the join point; stoichiometry of these bands is also typical, where the band corresponding to *nif* region c is more abundant than the rest. Band pattern is also altered in the derivative carrying an inversion, showing three bands of 10.4, 6.5 and 5.6 kb; the first two bands are join points for this rearrangement, while the last corresponds to *nif* region b. The derivative carrying a deletion shows a single, *nifH*-positive band of 13.6 kb; this band is the join point. Location of the *nifH::I-SceI* and *nifD::Km/Sp* alleles was verified through hybridization with allele-specific probes.

## RESULTS

**Experimental design:** The main objective of this work was to study the relative role of recombination *vs.* apparent gene conversion in the maintenance of sequence identity in a reiterated multigene family. To that end, we chose the nitrogenase multigene family of *R. etli* as a model. All members of this family are located in a single 390-kb plasmid, the symbiotic plasmid (pSyn). This family is composed of two identical direct reiterations of about 5 kb (*nif* regions a and b), which are *nifHDK* operons; these operons are 120 kb apart on the pSyn. The third element of this family (*nif* region c) is located in the middle of this zone, and consists of an identical reiteration, 1.5 kb long, harboring a complete *nifH* gene and a truncated *nifD* gene in an inverted orientation *vis a vis* *nif* regions a and b. Homologous recombination between *nif* regions a and b leads to frequent genomic rearrangements, such as deletions (ROMERO *et al.* 1991; ROMERO *et al.* 1995). This particular arrangement allows us to study the recombinational dynamics of a plasmidic multigene family.

Since expression of these *nif* regions under *ex planta* conditions does not confer any scorable phenotype, we modified *nif* region a by inserting a promoterless *Km<sup>r</sup>* cassette into the *nifD* gene, as described in MATERIALS

change to M

see out copy

24

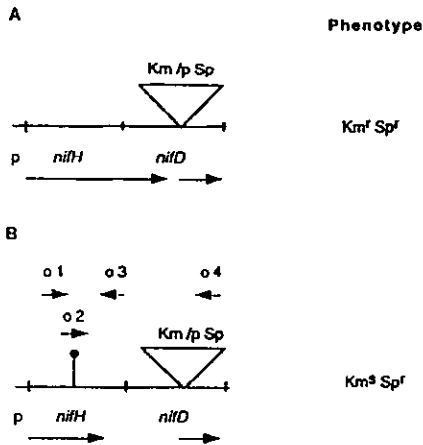


FIGURE 1.—Experimental design. Both parts of the figure represent *nif* region a. Symbols are as follows: the *nifH* promoter (p); the *nifD::Km/Sp* allele (triangle symbol); and the *nifH::I-SceI* allele (stick and ball symbol). Arrows beneath the figures indicate the expected transcripts in each case. (a) In this case, cells become  $Km^r$  due to transcription of the  $Km^r$  gene from the *nifH* promoter. (b) Introduction of the *nifH::I-SceI* allele leads to a  $Km^r$  phenotype, due to the introduction of additional termination codons. Small arrows in this figure indicate the location of specific oligonucleotides (o1–o4) used for characterization.

AND METHODS. This cassette also carries a  $Sp^r$  gene with its own promoter. As shown in Figure 1a, expression of the  $Km^r$  gene in this *nifD::Km/Sp* allele should be under the control of the *nifH* promoter. This construct was then modified by the insertion of a 28-bp oligonucleotide into the *nifH* gene (the *nifH::I-SceI* allele, see MATERIALS AND METHODS). This *nifH::I-SceI* allele leads to alterations in the translational reading of the *nifH* gene, because in-frame reading of this insertion causes misreading of two termination codons (UAG and UAA) present in the oligonucleotide. Additionally, since this insertion provokes a +1 frameshift, two additional stop codons are uncovered (UAA and UGA) at positions matching codons 185–186 and 214–215 of the wild-type *nifH* sequence, respectively. As shown in Figure 1b, the *nifH::I-SceI* allele should block, by polarity, the expression of the *nifD::Km/Sp* allele, thus leading to a  $Km^r$  phenotype. Selection for  $Km^r$  derivatives give us a positive system to identify events that lead to the loss or relocation of the 28-bp insertion, conceivably via recombination with the other members of this multigene family.

Initial tests of the functionality of this system were done in *E. coli*. To that end, we introduced plasmids pCRS4 (carrying the *nifD::Km/Sp* allele; Table 1) and

pCRS5 (carrying the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination; Table 1) into *E. coli* HB101. In this system, expression directed by the *nifH* promoter depends on  $\sigma^{54}$  and the NifA activator protein (VALDERRAMA *et al.* 1996). As expected, *E. coli* strains harboring either pCRS4 or pCRS5 were sensitive to kanamycin ( $50 \mu\text{g ml}^{-1}$ ). Upon introduction of a second, compatible plasmid carrying the constitutively-activated *Klebsiella pneumoniae nifA* gene (pMC71A) into the strain harboring pCRS4, cells became resistant to kanamycin. In contrast, cells carrying both pCRS5 and pMC71A remained sensitive to kanamycin. Thus, these results indicate that expression of the  $Km^r$  gene in the *nifD::Km/Sp* allele is dependent on the *nifH* promoter and that the introduction of the *nifH::I-SceI* allele blocks that expression.

Construction of *R. etli* strains containing on the pSym either the *nifD::Km/Sp* allele (strain CFNX236, Table 1) or the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination (strain CFNX237, Table 1) was done by allelic replacement (see MATERIALS AND METHODS). In this host, maximal expression from the *nifH* promoter is achieved under microaerobic conditions (VALDERRAMA *et al.* 1996). However, basal transcription from this promoter under aerobic conditions was enough to confer to strain CFNX236 a low-level resistance to kanamycin ( $3 \mu\text{g ml}^{-1}$ ).

As expected, strain CFNX237, carrying the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination was sensitive to kanamycin. Thus, loss or relocation of the *nifH::I-SceI* allele can be detected in strain CFNX237 by scoring the frequency of  $Km^r$  derivatives.

Theoretically, recombinational repair of the *nifH::I-SceI* allele to yield a  $Km^r$  derivative can arise either by sister-strand exchanges or through intramolecular exchanges. As shown in Figure 2, a sister-strand crossover between *nif* regions a and b leads to the formation of a large tandem duplication, where the join point carries the *nifD::Km/Sp* allele but lacks the *nifH::I-SceI* allele. Alternatively, removal of the *nifH::I-SceI* allele by sister-strand exchanges can come about from either double recombination or gene conversion. Both processes generate a non-rearranged pSym, lacking the *nifH::I-SceI* allele but maintaining the *nifD::Km/Sp* allele (Figure 2). This kind of product was called apparent gene conversion, since a formal distinction of which process is participating (double recombination or gene conversion) is not possible when the recombining sequences are in direct orientation (SEGALL and ROTH 1994).

Intramolecular exchanges might also be responsible for the generation of selectable  $Km^r$  derivatives. As shown in Figure 3 (left part), an intramolecular crossover generates a true gene conversion recombinant. However, this event cannot be distinguished from the non-rearranged class generated by sister-strand exchanges (see above); therefore, all these are also scored as apparent gene convertants. Intramolecular exchange between *nif* regions a and c generates an inversion (Figure 3, center). This recombinant has a  $Km^r$  phenotype

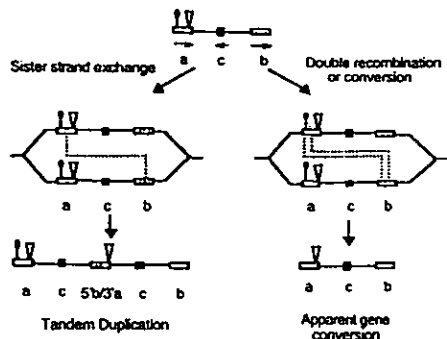


FIGURE 2.—Formation of  $Km^r$  derivatives by sister-strand exchanges. *nif* regions a, b and c are shown as rectangles with white, stippled or black shading, respectively, the *nifH::I-SceI* allele as a stick and ball symbol and the *nifD::Km/Sp* allele as a triangle symbol. (Left part) Recombination between *nif* regions a and b leads to the formation of a large tandem duplication, where the join point (indicated as a rectangle with mixed shading) carries the *nifD::Km/Sp* allele but lacks the *nifH::I-SceI* allele. (Right part) Both double recombination or gene conversion can generate a non-rearranged pSym, lacking the *nifH::I-SceI* allele but maintaining the *nifD::Km/Sp* allele. This kind of product was called apparent gene conversion, since making a formal distinction between which process (double recombination or gene conversion) is participating is not possible when the recombining sequences are in direct orientation. Only the selectable products are shown.

due to the relocation of the *nifD::Km/Sp* allele, which is now fused to the *nifH* region c promoter. Finally, intramolecular double recombination may also produce a  $Km^r$  recombinant, due to a relocation of the *nifH::I-SceI* allele (Figure 3, right).

Correction of a small insertion in the *nif* multigene family is achieved by multiple recombination events: To evaluate the frequency of correction in this multigene family, strain CFNX297, carrying the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination on the pSym, was grown overnight in rich media and plated in media containing a low concentration of kanamycin ( $3 \mu\text{g ml}^{-1}$ ).  $Km^r$  derivatives were found at a high frequency ( $344 \times 10^{-6}$ ). To identify the molecular events responsible for the generation of  $Km^r$  derivatives, 51 colonies were randomly chosen from seven independent selection experiments and purified as single-colony isolates. These derivatives were characterized by a combination of PCR, sizing of the resulting pSym and Southern blot hybridization against specific probes and then assigned to specific classes according to the criteria described in MATERIALS AND METHODS.

As shown in Figure 4, derivatives belonging to class I (tandem duplication) were the most abundant, comprising 74% of the observed products. This class may have resulted from a sister-strand exchange (Figure 2,

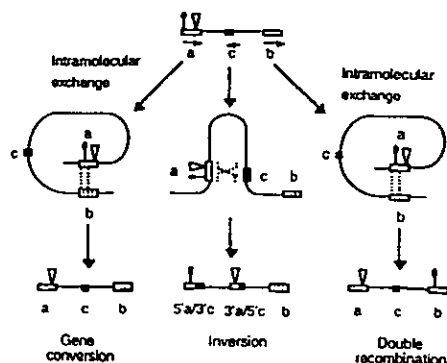


FIGURE 3.—Intramolecular exchanges might also be responsible for the generation of selectable  $Km^r$  derivatives. All symbols are as in Figure 2. (Left part) An intramolecular crossover generates a true gene conversion recombinant. However, this event cannot be distinguished from the non-rearranged class generated by sister-strand exchanges (see Figure 2); therefore, all these are also scored as apparent gene convertants. (Center part) An intramolecular exchange between *nif* regions a and c generates an inversion. This recombinant has a  $Km^r$  phenotype due to the relocation of the *nifD::Km/Sp* allele, which is now fused to the *nifH* region c promoter. (Right part) Intramolecular double recombination may also produce a  $Km^r$  recombinant, due to a relocation of the *nifH::I-SceI* allele. Only the selectable products are shown.

left). Class II derivatives (apparent gene conversions) constitute 14% of the total isolates analyzed. As explained before, these may originate either through interactions involving sister strands (double crossover or gene conversion, Figure 2, right) or from intramolecular exchanges (true gene conversion, Figure 3, left). Inversions (class III, Figure 4) were very scarce in this sample, being represented by a single derivative (ca. 2%). This class is readily explained by assuming a reciprocal intramolecular exchange (Figure 3, center; see DISCUSSION).

The remaining 10% of the derivatives are divided into three additional classes (IV to VI, Figure 4). These were generated most likely through a combination of apparent gene conversion plus additional recombination events. For instance, class IV is a tandem duplication similar but not identical to class I. Unlike class I, class IV derivatives have additionally lost the *nifH::I-SceI* allele, thus being also a case of apparent gene conversion. Class V is similar to the simple class II derivatives (apparent gene conversion), but it harbors an additional *nifD::Km/Sp* allele in *nif* region b. Class VI is an apparent gene conversion that carries a large deletion of the pSym; deletions similar to this have been observed previously (ROMERO *et al.* 1991). No derivatives attributable to intramolecular double recombination (Figure 3,

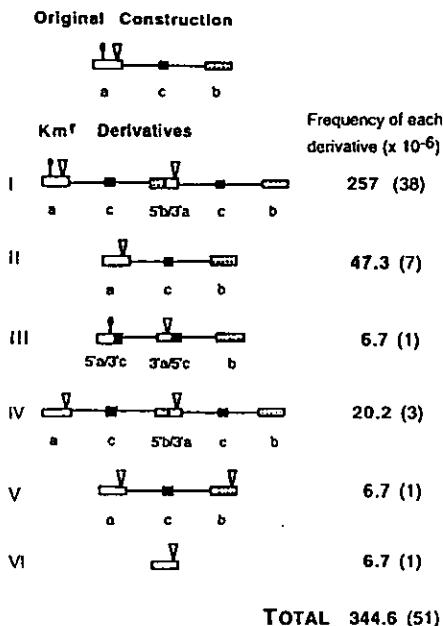


FIGURE 4.—Molecular events leading to the correction of a small insertion in the *nif*/multigene family. Roman numerals indicate specific classes. Numbers in bold denote the frequency of clones in each class, while numbers in parentheses are the number of Km<sup>r</sup> clones represented in the corresponding class. All symbols are as in Figure 2.

right) were found, despite the finding of other complex, double-exchange events. Possible reasons for the absence of this class are presented in the DISCUSSION.

Thus, these results indicate that: (i) a variety of recombination events are participating in the correction of a small insertion in a member of this multigene family; (ii) tandem duplication is a major contributor for the observed correction, followed significantly by apparent gene conversion; and (iii) multiple recombination events were frequently found.

**Role of the *recA* gene in recombinational correction:** The participation of the *recA* gene in the processes leading to correction of the *nifH::I-SceI* allele was evaluated by introducing the *recA::ΩCm* allele into strain CFNX237 as described in MATERIALS AND METHODS. Introduction of the *recA::ΩCm* allele provokes a 50-fold reduction in the frequency of isolation of Km<sup>r</sup> derivatives (from  $344 \times 10^{-6}$  in strain CFNX237 to  $6 \times 10^{-6}$  in strain CFNX242). Characterization of twenty Km<sup>r</sup> isolates obtained from strain CFNX242 did not reveal any loss or relocation of the *nifH::I-SceI* allele. Although not

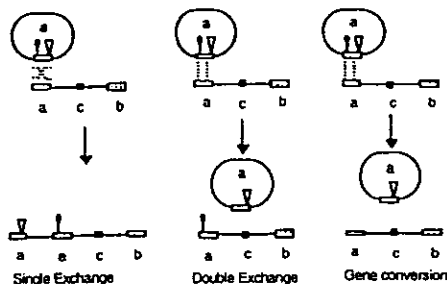


FIGURE 5.—Formation of Km<sup>r</sup> derivatives by intermolecular exchanges. Intermolecular exchanges can form selectable Km<sup>r</sup> derivatives either by a double exchange (center part) or by a true gene conversion event (right part). The product shown at the left part (single exchange) is not selectable, because it confers only a low-level resistance to kanamycin.

characterized further, these derivatives are more likely due to spontaneous mutations that restore transcription of the Km<sup>r</sup> gene. Thus, the *recA* gene participates in the formation of tandem duplications and apparent gene conversions.

Recombinational correction in the *nif* multigene family occurs frequently by intermolecular exchanges: A particular limitation of the system employed so far is that correction of the *nifH::I-SceI* allele might be generated either by intermolecular exchanges (sister-strand events), by intramolecular recombination or through a combination of both processes. However, a clear distinction of the role of intermolecular exchanges can be achieved through the incorporation of the allele whose correction is to be scored into a separate plasmid. As shown in Figure 5, correction in this system occurs only by intermolecular exchanges. This process has three separate outcomes. Single exchanges provoke the incorporation of the small plasmid into the pSym (Figure 5, left). The other two alternatives (double exchange or gene conversion) maintain the small plasmid and the pSym as separate entities, but the small plasmid now lacks the *nifH::I-SceI* allele (Figure 5, center and right). Although the expected products for double exchange or gene conversion events are different, it is not feasible to achieve a distinction between these alternatives, due to the possibility of cosegregation between the small plasmid (generated by a double exchange) and a second, uninvolved copy of the pSym.

To evaluate the role of intermolecular exchanges in the observed recombinational correction, small, broad host-range plasmids containing either the *nifD::Km/Sp* allele (pCRS6, Table 1) or the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination (pCRS7, Table 1) were constructed as described in MATERIALS AND METHODS. Conjugative transfer of each plasmid into a wild-type *R. etli*

strain (CE3) generated strains CFNX238 and CFNX239 (Table 1). Strain CFNX238 (harboring plasmid pCRS6) showed a high-level resistance to kanamycin ( $70 \mu\text{g ml}^{-1}$ ), while strain CFNX239 (carrying pCRS7) remained  $\text{Km}^r$ .  $\text{Km}^r$  derivatives from strain CFNX239 were obtained at a frequency of  $28 \times 10^{-6}$ . A total of 21 separate isolates (obtained from four independent selection experiments) were classified under the criteria described in MATERIALS AND METHODS. All these derivatives showed a structure compatible with apparent gene conversion (Figure 5, right). The absence of the single-exchange class was expected, due to the low resistance to kanamycin observed when the *nifH::I-SceI* and *nifD::Km/Sp* allelic combination was integrated on the pSym. Similar to the previous system, the double-exchange class was absent from this sample (see DISCUSSION).

The frequency of the apparent gene conversion class in this system ( $28 \times 10^{-6}$ ) is similar to the one estimated before ( $48 \times 10^{-6}$ ). These results suggest that this class is frequently generated through intermolecular exchanges.

This system allows us to determine if the formation of apparent gene convertants depends on the presence of additional copies of the *nif* region. To that end, plasmid pCRS7 was introduced into strain CFNX89, generating strain CFNX241; since this strain is devoid of the pSym, additional copies of the *nif* region are absent. The frequency of  $\text{Km}^r$  derivatives is reduced threefold (to  $7.1 \times 10^{-6}$ ); these derivatives did not show the correction of the *nifH::I-SceI* allele. Thus, these results, coupled to the *recA* dependency data, indicate that apparent gene conversions are formed through a recombinational process, and not due to a *recA*-dependent, excision repair process.

Correction in the *nif* multigene family leads to a precise restoration in nucleotide sequence: To be useful as a process leading to concerted variation between the *nif* regions, apparent gene conversion must restore the *nifH::I-SceI* allele to an otherwise wild-type sequence. To ascertain if this was the case, three independent class II,  $\text{Km}^r$  derivatives and one class VI derivative (Figure 4) were subjected to PCR amplifications employing oligonucleotides o1 and o4 (Figure 1) as primers. This particular primer pair yielded products of the region encompassing the site where the *nifH::I-SceI* allele was located. These PCR products were sequenced, using oligonucleotide o1 as a primer. In every case, nucleotide sequence was fully restored to wild type. These data were extended by isolating pCRS7 from four independent  $\text{Km}^r$  derivatives obtained as explained in the previous section; these plasmids were then transformed into *E. coli*. Nucleotide sequence from these derivatives also yielded a region identical to the wild-type sequence (data not shown). Thus, apparent gene conversions lead to a precise restoration of the wild-type nucleotide sequence.

## DISCUSSION

In this study we show that multiple recombination events, including tandem duplication, apparent gene conversion and inversion, lead to the correction of a small insertion in the *nif* multigene family. Tandem duplications appeared at a frequency of  $2.8 \times 10^{-7}$ , and were formed through recombination between *nif* region a and *nif* region b. The abundance of this class (75% of the observed events) was expected, since any crossover in the 1-kb region separating the *nifH::I-SceI* and *nifD::Km/Sp* alleles must produce a  $\text{Km}^r$  derivative. Thus, the frequency at which this class is generated reflects the frequency of recombination per kb of available homology in this organism. This estimate agrees well with previous evaluations in an interval of equivalent size (ROMERO *et al.* 1995).

Apparent gene conversion events in the *nif* multigene family were readily detected. The recombinational origin of these events is supported by (i) the strict requirement for a functional *recA* gene, (ii) the dependency on additional copies of the *nif* region for their formation, and (iii) a precise restoration of the wild-type nucleotide sequence. Apparent gene conversions comprised 24% of the observed events, appearing at a cumulative frequency of  $8 \times 10^{-5}$ . This estimate is remarkably similar to the observed frequencies of apparent gene conversion in both *E. coli* (HARVEY and HILL 1990) and *S. typhimurium* (SEGALL and ROTH 1994). Previous estimates obtained in *R. meliloti*, based on data from phage crosses, suggested that the efficiency of apparent gene conversion was lower in this organism than in *E. coli* (OROSZ *et al.* 1980). Our results clearly show that this is not the case, at least for recombinational interactions among members of a multigene family.

It is important to remark that this frequency refers to interactions with a particular outcome, namely, the correction of a specific 28-bp insertion. When all the possible corrections alongside the 5-kb *nif* regions are considered, the fraction of cells in a population with some change caused by apparent gene conversion in the *nif* regions should become substantial. Our results show that transfer of genetic information through apparent gene conversion occurs at a frequency that exceeds the frequency of spontaneous mutation. Therefore, this mechanism must contribute significantly to maintain the genetic homogeneity observed among members of the *nif* multigene family in Rhizobium, resulting in concerted evolution. We suggest that apparent gene conversion should also be relevant to explain other instances of maintenance of sequence identity in Rhizobium, as seen in a variety of multigene families.

Besides demonstrating the occurrence of apparent gene conversion, our work also allows us to draw some useful inferences about the recombination process in Rhizobium. Compared to tandem duplication and ap-

parent gene conversion, inversions are minor contributors to correction of the effects of a small insertion. Only a single inversion, formed by recombination between *nif* region a and *nif* region c was isolated, at a frequency of  $6 \times 10^{-6}$ . Although the occurrence of an inversion allows us to classify this interval as permissive (according to the terminology of SEGALL and ROTH 1994), the scarcity of this class is intriguing. Possible deleterious effects of this rearrangement on strain viability can be ruled out, because the strain harboring an inversion grows as well as the parental strain (data not shown). The infrequent occurrence of inversions cannot be explained by restriction in the size of a homologous region available for recombination. The interval used to generate an inversion is similar in location and size (750 bp) to the one used to generate a tandem duplication (1 kb). Despite this similarity, inversions occur 100-fold less frequently than tandem duplications.

Rarity of the inversion class can be explained by assuming, as has been proposed previously, that recombination in bacteria frequently entails the use of the so-called half-crossing over (MAJIAN and ROTH 1989; KOBAYASHI 1992; SEGALL and ROTH 1994; ROTH *et al.* 1996; YAMAMOTO *et al.* 1996). Following this proposal, a half-crossover between two recombining sequences would generate a recombinant molecule and two broken DNA ends. Tandem duplications and apparent gene conversions can be readily generated by this mechanism; in contrast, formation of an inversion would require two coincidental half-crossovers to produce a viable recombinant, explaining the low frequency of this class. The occurrence of half-crossovers would also explain two additional aspects of this work, namely the preference of intermolecular exchanges for generation of apparent gene convertants and the absence of derivatives attributable to intramolecular double recombination (Figure 3, right). These features might be a consequence of the fact that a single intramolecular half-crossover generates a linear molecule, thus precluding the isolation of a viable recombinant.

Considering the arguments given above, the occurrence of classes attributable to complex, double-exchange events (classes IV to VI, Figure 4) might seem paradoxical. However, we believe that these classes are formed by successive, rather than coincidental, half-crossover events. Common to these classes is that they appear to be composites of two different recombination events, where the product of each event is selectable. For instance, class IV was apparently generated by two selectable events, one a tandem duplication and the other an apparent gene conversion. It is conceivable that recombinants belonging to this class may have formed as successive recombination events during colony development, in a situation akin to phage recombination or recombination among small, high-copy number plasmids (YAMAMOTO *et al.* 1988).

We are indebted to SUSANA BRUM, JAIME MARTÍNEZ-SALAZAR and

RAYANI PALACIOS for critical reviewing of the manuscript, to JOSÉ ENRIQUETE and CÉSAR HERNÁNDEZ for help in preparing the manuscript, to PATRICIA BUSTIN for help with nucleotide sequencing, and to JAVIER RIVERA and MARÍA DE LA PAZ SALAS for skillful technical assistance. Partial financial support for this research was provided by grants No. 4321-NH06 from the Consejo Nacional de Ciencia y Tecnología (México) and No. 030355 from the Programa de Apoyo a las Divisiones de Estudios de Posgrado (Universidad Nacional Autónoma de México). C. R. was a recipient of a scholarship from the Consejo Nacional de Ciencia y Tecnología (México).

## LITERATURE CITED

- ABMULKARIM, F., and D. HUGHES, 1996 Homologous recombination between the *tuf* genes of *Salmonella typhimurium*. *J. Mol. Biol.* 260: 506-522.
- AUSUBEL, F. M., R. BRENT, R. E. KINGSTON, D. D. MORRIS, J. G. SEIDMAN *et al.*, 1987 *Current Protocols in Molecular Biology*, Massachusetts General Hospital, Harvard Medical School, John Wiley & Sons, Inc., New York.
- BACHILLIER, S., E. GILSON, M. HÖNING and C. W. HILL, 1996 Repeated sequences, pp. 2012-2040 in *Escherichia coli* and *Salmonella typhimurium: Cellular and Molecular Biology*, Ed. 2, edited by F. C. NEIDHARDT, R. CURTIS III, J. L. INGRAHAM, E. C. C. LIN, K. B. LUI, JR. *et al.* American Society for Microbiology Press, Washington, DC.
- BADENRICH-JONES, J., T. A. HEILTON, C. MORRISON, K. F. SCOTT and J. SHINE, 1989 Structural and functional analysis of nitrogenase genes from the broad-host-range *Rhizobium* strain ANU240. *Gene* 77: 141-153.
- BERGMAN, K., E. NULTY and L. SU, 1991 Mutations in the two flagellin genes of *Rhizobium meliloti*. *J. Bacteriol.* 173: 3716-3723.
- BIATTNER, F. R., G. PLUNKETT III, C. A. BLOCH, N. T. PERNA, V. BURLAND *et al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. *Science* 277: 1453-1474.
- BOYER, H. W., and D. RULLIANO-DUSSUUX, 1969 A complementation analysis of restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* 41: 459-472.
- BRUM, S., A. GARCÍA DE LOS SANTOS, T. STĘPKOWSKI, M. FLORES, G. DÁVILA *et al.*, 1992 Different plasmids of *Rhizobium leguminosarum* bv. phaseoli are required for optimal symbiotic performance. *J. Bacteriol.* 174: 5183-5189.
- BUCHANAN-WHILLASTON, V., M. C. CANNON, J. L. BEYNON and F. C. CANNON, 1981 Role of the *nifA* gene product in the regulation of *nif* expression in *Albizia julibrissin*. *Nature* 294: 776-778.
- DAVID, M., O. DOMERGUE, P. TOKKONJIC, and D. KAHN, 1987 Transcription patterns of *Rhizobium meliloti* symbiotic plasmid pSym: identification of *nifA*-independent *fix* genes. *J. Bacteriol.* 169: 2239-2244.
- DEMONIK, R. C., 1996 Native insertion sequence elements: locations, distributions and sequence relationships, pp. 2000-2011 in *Escherichia coli* and *Salmonella typhimurium: Cellular and Molecular Biology*, Ed. 2, edited by F. C. NEIDHARDT, R. CURTIS III, J. L. INGRAHAM, E. C. C. LIN, K. B. LUI, JR. *et al.* American Society for Microbiology Press, Washington, DC.
- DIVIK, G. A., 1993 Evolution of genetic redundancy for advanced players. *Curr. Opin. Genet. Dev.* 3: 902-910.
- ECKHARDT, T., 1978 A rapid method for the identification of plasmid deoxyribonucleic acid in bacteria. *Plasmid*. 1: 534-538.
- FENNERG, A. P., and B. VOGELSTEIN, 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6-15.
- FIGURSKI, D. H., and D. R. HELINSKI, 1979 Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proc. Natl. Acad. Sci. USA* 76: 1648-1652.
- FISCHER, H.-M., M. BARST, T. KASPAR, C. ACURIA, F. ARIGNI *et al.*, 1993 One member of a *groES*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. *EMBO J.* 12: 2901-2912.
- FLORES, M., V. GONZÁLEZ, S. BRUM, E. MARTÍNEZ, D. PIRENO *et al.*, 1987 Reiterated DNA sequences in *Rhizobium* and *Agrobacterium* spp. *J. Bacteriol.* 169: 5782-5788.

- FREIBERG, C., R. FELLAY, A. BAIKREH, W. J. BROUGHTON, A. REINHART, *et al.*, 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* 387: 394-401.
- GANGICHY, S., H. ZHU and R. ROTHSCHILD, 1996. Gene conversion plays the major role in controlling the stability of large tandem repeats in yeast. *EMBO J.* 15: 1715-1725.
- GIRARD, M. L., M. FLURIK, S. BRON, D. ROMERO, R. PALACIOS *et al.*, 1991. Structural complexity of the symbiotic plasmid of *Rhizobium leguminosarum* bv. phaseoli. *J. Bacteriol.* 173: 2411-2419.
- HARVEY, S., and C. W. HILL, 1990. Exchange of spacer regions between rRNA operons in *Escherichia coli*. *Genetics* 125: 683-690.
- HERNÁNDEZ-LUCAS, I., M. A. PARDOL, L. SIGOVIA, J. MIRANDA and E. MARTÍNEZ-ROMERO, 1995. *Rhizobium tropici* chromosomal citrate synthase gene. *Appl. Environ. Microbiol.* 61: 3992-3997.
- JACKSON, J. A., and G. R. FINK, 1981. Gene conversion between duplicated genetic elements in yeast. *Nature* 292: 306-311.
- JOHN, B., and G. L. G. MILKES, 1988. *The Eukaryote Genome in Development and Evolution*. Allen & Unwin, London.
- JONES, J. D. G., and N. GUTTERSON, 1987. An efficient mobilizable cosmid vector, pRK7313, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* IV37a. *Gene* 61: 299-306.
- KLEIN, H. L., and T. D. PETER, 1981. Intrachromosomal gene conversion in yeast. *Nature* 289: 144-148.
- KURAYASHI, I., 1992. Mechanisms for gene conversion and homologous recombination: the double-strand break repair model and the successive half crossing-over model. pp. 81-133 in *Advances in Biophysics*, edited by M. KITAMI. Japan Scientific Societies Press and Elsevier, Tokyo and Limerick.
- KÜNDIG, C., H. HENNENCKE and M. GÖTTFRIT, 1993. Correlated physical and genetic map of the *Rhizobium japonicum* 110 genome. *J. Bacteriol.* 175: 613-622.
- MAJIAN, M. J., and J. R. ROTH, 1989. Role of *recBC* function in formation of chromosomal rearrangements: a two-step model for recombination. *Genetics* 121: 433-443.
- MARGOLIN, W., and S. LONG, 1994. *Rhizobium meliloti* contains a novel second homolog of the cell division gene *ftsZ*. *J. Bacteriol.* 176: 2053-2043.
- MARSH, J. L., M. ERPEL and E. J. WYKIS, 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. *Gene* 32: 481-485.
- MARTÍNEZ-SALAZAR, J. M., D. ROMERO, M. L. GIRARD and G. DÁVILA, 1991. Molecular cloning and characterization of the *rocA* gene of *Rhizobium phaseoli* and construction of *rocA* mutants. *J. Bacteriol.* 173: 3035-3040.
- MILLER, J. H., 1972. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- MURPHY, E., S. MORINO and G. ESPIN, 1988. Transcription analysis of the three *nifH* genes of *Rhizobium phaseoli* with gene fusions. *Mol. Gen. Genet.* 213: 499-504.
- NOEL, K. D., F. SÁNCHEZ, L. FERNÁNDEZ, J. LEEJANS and M. A. CEVALLOS, 1984. *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J. Bacteriol.* 158: 148-155.
- NOEL, F., and C. ELANKRICH, 1987. Nucleotide sequence and functional analysis of the two *nifH* copies of *Rhizobium* ORS571. *J. Gen. Microbiol.* 133: 1563-1576.
- OHMI, S., 1970. *Evolution by Gene Duplication*. Springer-Verlag, Berlin.
- OKAZAKI, N., S. MATSUDA, K. SAITO, A. TOMINAGA and M. ENOSHITA, 1993. Conversion of the *Salmonella* phase 1 flagellin gene *fljC* to the phase 2 gene *fljB* on the *Escherichia coli* K12 chromosome. *J. Bacteriol.* 175: 758-766.
- ORKOV, L., A. PAV and G. DALLMANN, 1980. Heterozygosity of phase 16-3 of *Rhizobium meliloti*: moderate level of mismatch repair or gene conversion. *Mol. Gen. Genet.* 179: 163-167.
- PAKID, M. A., J. LACUÑEZ, J. MIRANDA and E. MARTÍNEZ, 1994. Nodulating ability of *Rhizobium tropici* is conditioned by a plasmid-encoded citrate synthase. *Mol. Microbiol.* 11: 315-321.
- PETER, T. D., 1980. Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. *Cell* 19: 765-774.
- PRENTER, P., and H. M. KREMER, 1984. In vitro insertional mutagenesis with a selectable DNA fragment. *Gene* 29: 303-315.
- QUINTO, C., H. DE LA VEGA, M. FLEHER, J. LEMANS, M. A. CEVALLOS *et al.*, 1985. Nitrogenase reductase: a functional multigene family in *Rhizobium phaseoli*. *Proc. Natl. Acad. Sci. USA* 82: 1170-1174.
- ROMERO, D., S. BRON, J. MARTÍNEZ-SALAZAR, M. L. GIRARD, R. PALACIOS *et al.*, 1991. Amplification and deletion of a *nodJif* region in the symbiotic plasmid of *Rhizobium phaseoli*. *J. Bacteriol.* 173: 2435-2441.
- ROMERO, D., J. MARTÍNEZ-SALAZAR, L. GIRARD, S. BRON, G. DÁVILA *et al.*, 1995. Discrete amplifiable regions (amplicons) in the symbiotic plasmid of *Rhizobium etli* CFN42. *J. Bacteriol.* 177: 973-980.
- ROMERO, D., G. DÁVILA and R. PALACIOS, 1997. The dynamic genome of *Rhizobium*, pp. 153-161 in *Bacterial Genomes: Physical Structure and Analysis*, edited by F. J. DE BRUIJN, J. R. LUPSKI and C. WEINSTEIN. Chapman & Hall, New York.
- ROTH, J. R., N. BENSON, T. GALITSKI, K. HAECK, J. G. LAWRENCE *et al.*, 1996. Rearrangements of the bacterial chromosome: formation and applications, pp. 2256-2276 in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Ed. 2, edited by F. C. NEIDHARDT, R. CURTISS III, J. L. INGRAHAM, E. C. C. LIN, K. B. LOW, JR. *et al.* American Society for Microbiology Press, Washington, DC.
- SAMBRINK, J., E. F. FRITZCH and T. MANIATIS, 1989. *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- SCHÜTZER, A., T. PATSCHIKOWSKI, J. QUANT, L. B. SELINGER, S. WEINER *et al.*, 1997. Functional and regulatory analysis of the two copies of the *fixNOQ* operon of *Rhizobium leguminosarum* strain VF39. *Mol. Plant Microbe Interact.* 10: 605-616.
- SCHULTZE, M., E. KUMMEROW, P. RATY, M. BURF and A. KAINIGROST, 1994. Cell and molecular biology of *Rhizobium*-plant interactions. *Int. Rev. Cytol.* 165: 1-75.
- SEFAL, A. M., and J. R. ROTH, 1994. Approaches to half-tetrad analysis in bacteria: recombination between repeated, inverse-order chromosomal sequences. *Genetics* 136: 27-39.
- SIGOVIA, L., J. P. W. YOUNG and E. MARTÍNEZ-ROMERO, 1993. Reclassification of american *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli* sp. nov. *Int. J. Syst. Bacteriol.* 43: 374-377.
- SIMON, R., 1984. High frequency mobilization of gram-negative bacterial replicons by the *in vitro* constructed Tn3-Mob transposon. *Mol. Gen. Genet.* 198: 413-420.
- SIMON, R., U. PRIEDRICH and A. PÜHLER, 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria. *Bio/Technology* 1: 784-791.
- SYKSTAK, J. W., and R. WU, 1980. Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. *Nature* 284: 426-430.
- VALINKKAMA, B., A. DÁVILA, L. GIRARD, E. MORETTI and J. MORA, 1996. Regulatory proteins and *ns*-acting elements involved in the transcriptional control of *Rhizobium etli* reiterated *nifH* genes. *J. Bacteriol.* 178: 3119-3126.
- WILLIAMS, S. M., and C. STROBERG, 1985. Sister chromatid exchange and the evolution of rDNA spacer length. *J. Theor. Biol.* 116: 625-636.
- YAMAHITO, K., H. YOSHIKURA, N. TAKAHASHI and I. KURAYASHI, 1988. Apparent gene conversion in an *Escherichia coli* rec<sup>+</sup> strain is explained by multiple rounds of reciprocal crossing over. *Mol. Gen. Genet.* 212: 393-404.
- YAMAHITO, K., N. TAKAHASHI, Y. FUJITANI, H. YOSHIKURA and I. KURAYASHI, 1996. Occurrence dependence in homologous recombination. *Genetics* 143: 27-36.

## RESULTADOS ADICIONALES

### **Evidencia teórica de la presencia del mecanismo de conversión génica en otra familia multigénica en *Rhizobium*.**

Al detectar la presencia de conversión génica aparente entre los miembros de la familia multigénica de los genes *nifH* en el pSim de *R. etli* CFN42, realizamos un análisis teórico considerando la posibilidad de encontrar evidencia de que este mecanismo ocurra en otras familias multigénicas. Para este fin, se realizó una búsqueda de secuencias reiteradas con alta identidad, que estuvieran presentes en dos cepas de *Rhizobium*. Solamente encontramos un ejemplo representativo, el cual presentamos a continuación.

### **Genes *flaA* y *flaB***

La flagelina aporta el 5% de la proteína de una bacteria y los genes responsables tal vez se expresan a niveles elevados. Los genes estructurales *flaA* y *flaB* codifican dos subunidades estrechamente relacionadas del complejo filamentosos flagelar de *R. meliloti*. (Pleier y Schmitt, 1989). Estos genes en la cepa *R. meliloti* 1021 están constituidos por 1188 (*flaA*) y 1191 (*flaB*) pares de bases y en la cepa *R. meliloti* RU10406 los dos genes miden 1185 pares de bases. Los genes de flagelina se transcriben de manera independiente y están arreglados en tandem con una separación entre ellos de 345 pares de bases (Fig. 5).



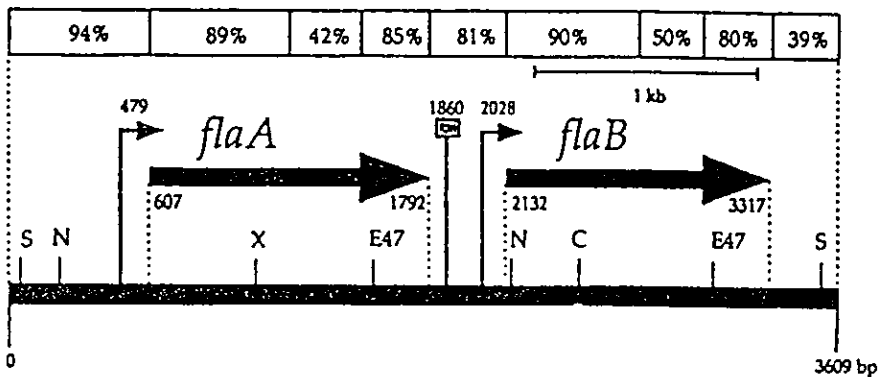


Fig. 5. Mapa de los genes *flaA* y *flaB* de *R. meliloti* 1021. Los genes de flagelina están representados con las flechas grandes. Los números indican la posición en la secuencia nucleotídica. Los sitios de restricción están indicados como: C (*ClaI*), E47 (*Eco47III*), N (*NcoI*), S (*SmaI*) y X (*XhoI*). La comparación porcentual de la identidad de secuencia para la región *flaA-flaB* de la cepa de *R. meliloti* RU10406 está señalada en cada región analizada. El inicio de la transcripción de cada gene está indicado por las flechas pequeñas (479 y 2028). Un sitio de terminación de la transcripción está señalado como un cuadrado (1860) y es una secuencia repetida invertida seguida por residuos de timina (Tomado de Bergman *et al.*, 1991).

En la cepa *R. meliloti* RU10406 los genes *flaA* y *flaB* tienen entre ellos una identidad del 93% y en la cepa *R. meliloti* 1021 los mismos genes presentan una identidad del 85% (Bergman *et al.*, 1991). La identidad se mantiene en las regiones centrales y los cambios que se presentan se localizan en las regiones amino y carboxilo terminal. Al realizar un análisis de secuencia y comparar a los genes de flagelina intercepas 1021 vs RU10406, la identidad entre estos genes baja al 50%. Además, existe en su región interna una divergencia de la secuencia codificadora, pero esta diferencia no se presenta entre los dos genes de flagelina de cada cepa.

Estos datos sugieren que la conversión génica se presenta entre las copias de los genes *flaA* y *flaB*, permitiendo la evolución concertada de los genes de flagelina.

La mayoría de la información con respecto a otras reiteraciones, se tiene por evidencia de hibridización y solamente en algunos casos se reporta la secuencia nucleotídica de los genes repetidos. Hasta el momento la evidencia de la existencia del mecanismo de conversión génica mediante la conservación de identidad de secuencia nucleotídica dentro de las Rhizobiaceae, se restringe a cinco genes completos (reiterados) que presentan una alta identidad entre ellos que va del 94% al 100% (Tabla III).

Estos genes son blancos susceptibles para diseñar en ellos sistemas genéticos, que permitan comprobar experimentalmente conversión génica. De resultar positivos los ensayos con esas secuencias reiteradas este mecanismo de homogenización como un proceso determinante en la dinámica del genoma de *Rhizobium*, se verá apoyado

Tabla III. Genes reiterados con alta identidad en Rhizobiaceae

CEPAS	GENES	IDENTIDAD	BASES	CARACTERISTICAS
<i>Sinorhizobium meliloti</i> 2011	<i>fixK</i>	100%	636	2 copias en el plásmido Sim. Renalier <i>et al.</i> , 1987
<i>Sinorhizobium meliloti</i> 1021	<i>groEL</i>	97%	1638	2 copias en el plásmido Sim. Rusanganwa y Gupta, 1993; Ogawa y Long, 1993.
<i>Rhizobium tropici</i> CFN299	<i>csA</i>	96%	1213	una copia en cromosoma y otra en pSim. Pardo <i>et al.</i> , 1994; Hernández-Lucas <i>et al.</i> , 1995
<i>Rhizobium leguminosarum</i> bv. viceae VF39	<i>fixN</i>	94%	1623	Una copia en el plásmido c y otra en el plásmido d. Schlüter <i>et al.</i> , 1997

**Explorar el efecto de ruptura en doble cadena sobre la dinámica de conversión génica para la familia multigénica *nifH*.**

La recombinación homóloga puede dar inicio mediante ruptura en cadena-sencilla ó ruptura en doble cadena. Si la recombinación se da por ruptura en cadena sencilla, hay asimilación de una cadena y la migración subsecuente de esta cadena. Alternativamente, si ocurre ruptura en doble cadena, es seguida por un mecanismo de reparación que usa una secuencia homóloga sin rupturas como templado.

Szostak *et al.* (1983) reportaron en *S. cerevisiae*, que un plásmido de *E. coli* que porta una secuencia homóloga al DNA de levadura, al realizar *in vivo* ruptura en doble cadena en la secuencia de levadura, el plásmido se integró en el locus cromosomal homólogo y la recombinación se incrementó 3 000 veces. Así, el evaluar el efecto *in vivo* de ruptura en doble cadena en los sistemas intramolecular e intermolecular en el pSim de *R. etli* CFN42, ofrece una manera interesante para profundizar en la dinámica de recombinación y en particular de conversión génica. De presentarse un efecto similar, habrá una mayor proporción de los productos de recombinación obtenidos a baja frecuencia como son las inversiones y creemos que se podrán detectar eventos de conversión génica asociada a recombinación.

El sistema considerado para detectar el efecto de ruptura en doble cadena sobre la dinámica de conversión génica emplea también a la familia multigénica de los *nifH*. La estrategia diseñada consiste en introducir una secuencia artificial específica (megalinker) dentro de una de las reiteraciones *nifH*; esta secuencia será el blanco para una enzima de restricción cuyo sitio de corte sea único en el genoma. El segundo paso consiste en clonar el gene para la enzima de restricción que actúe sobre este sitio, en un vector controlable y compatible con *Rhizobium* e introducirlo en las cepas de *Rhizobium* para evaluar recombinación intramolecular e intermolecular. Si la recombinación homóloga procede preferencialmente bajo el modelo de ruptura en doble cadena y reparación, este sitio, al ser cortado *in vivo*, funcionará como recombinador en el gene *nifH*.

La construcción de la región *nifH* con el megalinker está descrita en el artículo mostrado en esta tesis. El megalinker *1-SceI* contiene las 18 pares de bases

necesarias para ser reconocido como sitio único de restricción de la endonucleasa *1-SceI* del intrón mitocondrial de *Saccharomyces cerevisiae* (Colleaux *et al.*, 1988). Cabe señalar que la presencia natural de este sitio solo se ha detectado, además de en *S. cerevisiae*, en el fago T7 (Thierry *et al.*, 1991). Colleaux *et al.* (1986) determinaron que la presencia al azar de un sitio de 18 pares de bases ocurriría estadísticamente en una cantidad de DNA equivalente a uno en cada 20 genomas humanos. Esto hace sumamente improbable el encontrar otro sitio similar en el genoma de *Rhizobium*.

El ORF *1-SceI* del intrón mitocondrial de levadura (modificado al equivalente del código universal), está clonado en el plásmido pSCM525 bajo el control de un promotor de *E. coli* y produce una proteína idéntica a la transposasa mitocondrial (Colleaux *et al.*, 1986). Para verificar que este ORF es capaz de sintetizar la enzima de restricción correspondiente, se transfirió por conjugación el plásmido pCRS5.1 (la orientación del megalinker está invertida) a la cepa de *E. coli* que tiene el pSCM525. Después de la conjugación, se confirmó la coexistencia de los plásmidos mediante marcadores y patrón de restricción. Al realizar la cinética de inducción con IPTG de la enzima *1-SceI*, se detectó al inicio del experimento (tiempo cero) tanto en los cultivos con-y-sin inductor que el plásmido pCRS5.1 había desaparecido. Este resultado es similar al reportado por Yap y Schiestl (1995) quienes al introducir sitios de corte en doble cadena en secuencias de plásmidos que no tenían homología con secuencias del genoma, el plásmido digerido no se reestablecía por religación y siempre se perdía.

Debido a que el plásmido pSCM525 no se replica en *Rhizobium*, se obtuvo

mediante amplificación por PCR el ORF *1-SceI*. Fue necesario diseñar un sitio de restricción *EcoRI* en el oligonucleótido correspondiente a la región 5' para utilizarlo en la clonación y a su vez mantener la secuencia de Shine-Dalgarno; para el oligonucleótido de la región 3' se usaron los sitios de restricción *SaII* ó *PstI* propios de la secuencia, que están fuera de la región codificadora (Fig. 6a). El producto de PCR fue digerido posteriormente con las enzimas *EcoRI-SaII* ó *EcoRI-PstI* y clonado en el plásmido pSK+ (Fig.6b), para subclonarse posteriormente en dos vectores capaces de replicarse en *Rhizobium*.

Los plásmidos seleccionados contienen el replicón RSF1010, tienen bajo número de copias por célula (12 -13), son de amplio rango de huésped y poseen una buena eficiencia de movilización. Poseen como marcador cloramfenicol, directamente después del promotor pTacLacUV5 (pMMB206) o pTac (pMMB277), tienen los sitios de restricción del polilinker del fago M13mp18 e incorporan un fragmento de *lacZ*. Además tienen el gen *lacIq* que reprime la expresión de estos promotores, asegurando la estabilidad del producto clonado y son inducibles por IPTG ( Morales *et al.*, 1991).

El producto de PCR del ORF *1-SceI*, flanqueado por sitios de restricción *EcoRI-SaII* ó *EcoRI-PstI* fue clonado en los plásmidos pMMB206 y pMMB277. Morales *et al.* (1991) reportaron que en *E. coli* el pTac fue cinco veces mejor en actividad transcripcional que el pTacLacUV5, por lo que decidimos tener las dos alternativas para la inducción de la endonucleasa. Los plásmidos resultantes (fig. 6c), se denominaron pCRS8 (pMMB206+ORF *1-Sce I*) y pCRS9 (pMMB277+ORF *1-SceI*).

Una vez realizadas las construcciones, se introdujeron por conjugación los

plasmidos pCRS8 y pCRS9 de manera independiente a las cepas de *Rhizobium* construidas previamente (CFNX236 a CFNX239, ver artículo). Estas construcciones se verificaron por marcadores y mediante la técnica de Eckhardt se visualizó el perfil de plásmidos. Con el pCRS8 se generaron las cepas CFNX243 a CFNX246 y con el pCRS9 se obtuvieron las cepas CFNX291 a CFNX294 (Tabla IV).

Tabla IV. Cepas de *Rhizobium* construidas con los plásmidos pCRS8 y pCRS9.

CEPAS	CARACTERISTICAS
CFNX 243	Derivada de la CFNX236, con el plásmido pCRS8
CFNX 244	Derivada de la CFNX237, con el plásmido pCRS8
CFNX 245	Derivada de la CE3 , con los plásmidos pCRS6 y pCRS8
CFNX 246	Derivada de la CE3, con los plásmidos pCRS7 y pCRS8
CFNX 291	Derivada de la CFNX236, con el plásmido pCRS9
CFNX 292	Derivada de la CFNX237, con el plásmido pCRS9
CFNX 293	Derivada de la CE3 , con los plásmidos pCRS6 y pCRS9
CFNX 294	Derivada de la CE3 , con los plásmidos pCRS7 y pCRS9

Para realizar las cinéticas de inducción de la endonucleasa *1-SceI*, primero se pusieron cultivos independientes de las cepas CFNX243 a CFNX246 y de CFNX291 a CFNX294. Al día siguiente, en matraces con 50 ml de medio de cultivo PY, se inocularon cultivos a la misma densidad celular y se dejaron crecer durante tres horas. A partir de este momento los cultivos se dividieron . A una serie de cultivos se

les añadió el inductor IPTG (40  $\mu\text{g/ml}$ ) y la otra serie se creció sin inductor. Se tomaron muestras de las dos condiciones experimentales a diferentes tiempos (0, 3, 6, 18 y 24 horas). Para evaluar la frecuencia de las cepas resultantes de las cinéticas de inducción, la última muestra de cada condición fue plateada por dilución. Los controles se platearon en PYSp 100  $\mu\text{g/ml}$  y la condición de selección para los eventos intramoleculares fué PYKm 4  $\mu\text{g/ml}$  y los eventos intermoleculares fueron seleccionados en PYKm 70  $\mu\text{g/ml}$ . Finalmente, se realizó la extracción de DNA total de las muestras (con inducción y sin inducción), para posteriormente hibridizarlas con el detector de *nifH*.

Las frecuencias obtenidas en las cinéticas con inducción y sin inducción para los eventos intramoleculares, indican que los controles seleccionados en  $\text{Km}^4$  se obtuvieron a las frecuencias esperadas de 0.72 para la cepa CFNX243 y de 0.36 para la cepa CFNX291. En contraste, para las cepas que llevan el sitio *1-SceI* (seleccionadas en  $\text{Km}^4$ ) se esperaba una frecuencia menor que la de los controles y se obtuvieron a alta frecuencia 0.97 (CFNX244) y 0.53 (CFNX292). Debido a las frecuencias detectadas en las cepas experimentales, creemos que en las cepas ya había ocurrido reparación del sitio *1-SceI*, antes de iniciar la cinética de inducción.

Para corroborar lo anterior, mediante PCR's se comprobó que la cepas parentales habían perdido el sitio de restricción *1-SceI*. Asimismo, la hibridización con el detector *nifH* de los DNA procedentes de la cinética de inducción, no revelaron las bandas esperadas de 7.8Kb y 5.2Kb producidas por el corte del sitio *1-SceI* en la región *nifHa*. Todas las muestras presentaron el patrón de restricción de los genes



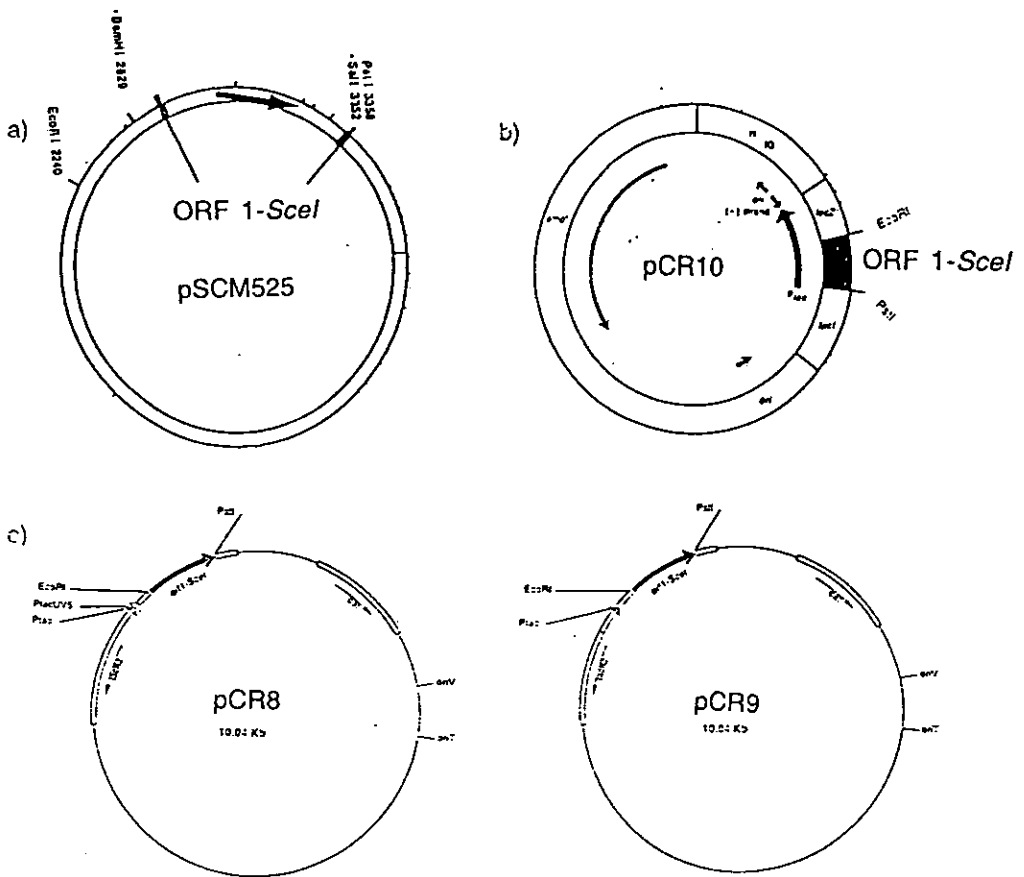


Fig. 6. Plásmidos utilizados para la construcción de los plásmidos pCRS8 y pCRS9. (a) Plásmido pSCM525 (Colleaux et al., 1986), se usó como template para obtener mediante PCR el ORF 1-*Scel*. (b) Plásmido pSK+ tiene clonado el producto de PCR del ORF 1-*Scel*, generando el plásmido pCRS10. (c) Plásmidos pCRS8 (pMMB206) y pCRS9 (pMMB277), llevan clonado el fragmento del ORF 1-*Scel* del pCRS10. Los sitios de restricción son *Pst*I (P), *Eco*RI (E), *Sal*I (S), *Bam*HI (B). Los marcadores de resistencia son cloramfenicol (Cm), ampicilina (Ap). Las regiones promotoras son pTac y pTacLacUV5. Los genes del represor de lactosa (*lacIq*), un fragmento del operón Lac (*lacZ*), el gene del intrón mitocondrial (*1-Scel*).

*nifH* de la cepa parental (*nifHa* con cassette 13kb, *nifHb* 5.6 kb y *nifHc* 4.6kb). Los resultados analizados hasta el momento para los eventos intermoleculares de la cinética de corte, indican que se ha perdido el sitio *1-SceI*, aún antes de realizar las cinéticas.

Con los resultados obtenidos en los eventos intra-e-inter moleculares, podemos establecer que se presentó el fenómeno de inducción cigótica en *Rhizobium*. El control de regulación del promotor de la enzima *1-SceI* no se está dando en los plásmidos (pCRS8 y pCRS9) al momento de la conjugación, y por lo tanto se produjo actividad de la enzima *1-SceI*. Debido a lo anterior, no podemos establecer el origen recombinacional de las bandas que se generaron y el efecto real sobre conversión génica en esos experimentos.

Podemos considerar estos datos positivos en la medida que dan indicio de que el sistema generado funciona en *Rhizobium* y el efecto de ruptura en doble cadena esta dirigido entre las regiones *nifH*. Las alternativas ahora están dirigidas a controlar la expresión de la enzima *1-SceI* para evaluar el efecto de ruptura en doble cadena.

Una opción para el sistema intramolecular, es utilizar la cepa de *R. etli* CFN2001 (carece de pSim y pa) como receptora e introducir por conjugación el plásmido pCRS8 y el pCRS9 de manera independiente, generando las cepas CFNX295 y CFNX296 respectivamente. Posteriormente transferiríamos por conjugación a cada una de las nuevas cepas, el pSim y el pa de las cepas CFNX236 y CFNX237. Finalmente se generarían las cepas CFNX243 y CFNX244 que tienen el pCRS8 y las cepas CFNX291 y CFNX292 que llevan el pCRS9.

Para el sistema intermolecular (plásmidos chicos), se usaría a la cepa de *Rhizobium etli* CFN42 como receptora y por conjugación se introducirán independientemente los plásmidos pCRS8 y pCRS9. Una vez caracterizadas estas construcciones, nuevamente por conjugación introduciríamos de manera independiente los plásmidos pCRS6 y pCRS7 en cada una de las cepas. El resultado final de las construcciones son las cepas CFNX245 y CFNX246 (que llevan el pCRS8) y las cepas CFNX293 y CFNX294 ( que llevan el pCRS9).

Otra estrategia considerada es la clonación del represor *lacIq*, en algunos plásmidos de *R etli*. A partir del plásmido pMMB206 que lleva el gene *lacIq* se diseñarían oligonucleótidos con sitios de restricción y mediante PCR , se obtendría el gene *lacIq*. Posteriormente en un vector suicida que tenga homología con alguno de los plásmidos de la cepa CFN42, se clonaría el producto de PCR del gene *lacIq*. Después por conjugación y mediante recombinación sencilla se integraría este plásmido al fondo genético de *R etli*. Una vez caracterizada la cepa, se introducirían los plásmidos para generar las cepas necesarias para el estudio de la cinética de inducción *1-Scel*.

## DISCUSIÓN Y PERSPECTIVAS

### Recombinación en *Rhizobium etli*

El arreglo particular de los genes *nifH* en *R. etli* permitió detectar amplificaciones y deleciones de manera natural (Brom *et al.*, 1991) y Romero *et al.* (1991) demostraron que la región *nod-nif* de 120kb del pSim flanqueada por los operones *nifH* era la región involucrada en recombinación. En las amplificaciones en tandem, se podrían realizar intercambios desiguales entre otros genes que estén contenidos dentro de la región *nod-nif*, sin alterar la estructura de la zona, permitiendo homogenización de los genes contenidos en la región *nod-nif*. Los productos de deleción, tienen dos alternativas, una es que se pierdan y la otra es que se reintegren en las zonas repetidas o en la región de donde se originaron. Los eventos de integración e intercambios desiguales para otras zonas dentro de la región *nod-nif* no se han evaluado.

La recombinación homóloga en el pSim de *Rhizobium etli* genera un estado dinámico de estructura molecular, produciendo a alta frecuencia rearrreglos genómicos que contribuyen en diferente proporción a la variabilidad estructural del pSim como son los eventos de amplificación (10-3), deleción (10-4) (Romero *et al.*, 1991); translocación (10-4), amplificación y cointegración son reversibles (10-2 -10 -3) (Romero *et al.*, 1998a), conversión génica (10-5) e inversiones (10-6) (véase el artículo). En procariotes, eventos similares de recombinación y a alta frecuencia solamente se han reportado para la familia multigénica de los *rm* en *E. coli* los cuales incluyen conversión génica (Harvey *et al.*, 1988; Harvey y Hill, 1990),

transposición y delección (Hill y Hamish, 1982), duplicación (Hill y Combrato, 1973), intercambio desigual (Lehner y Hill, 1985) e inversión (Hill y Hamish, 1981; Hill y Gray, 1988).

La duplicación de los operones *nifHDK* en *R. etli* CFN42 es la estrategia que le permite a esta cepa incrementar la fijación de nitrógeno (Romero *et al.*, 1988) y la amplificación del amplicón Tipo I promueve la capacidad de fijar nitrógeno en ciertos fondos genéticos (Romero *et al.*, 1997). De entre los productos de recombinación detectados, las inversiones pueden relocalizar genes bajo otros promotores. En *R. etli* CFN42 los promotores de las regiones *nifHa* y *nifHb* son débiles con respecto al promotor de la región *nifHc* que se expresa más fuerte. Al ocurrir un evento de inversión entre cualquiera de las regiones *nifHa* o *nifHb* con la región *nifHc*, el operón quedaría ahora bajo el promotor de la región *nifHc* y esta sería otra alternativa para sobreexpresar los operones *nifHDK* en *R. etli*. Posiblemente este rearrreglo cause un aumento en la fijación de nitrógeno, si no es afectado el arreglo de genes involucrados en simbiosis. Con la finalidad de detectar si se produce este tipo de inversión de manera natural, sería interesante caracterizar cepas provenientes de nódulo y checar el patrón de reiteración *nifH*.

La participación de rondas múltiples de recombinación contribuyen de manera significativa en la homogenización de secuencia de las reiteraciones *nifH* mediante conversión génica aparente (véase el artículo). En *E. coli* este fenómeno sólo ha sido detectado en recombinación de plásmidos pequeños con el gene de *kan/neo*. (Yamamoto *et al.*, 1988) y en el gene de *tet* (Doherty *et al.*, 1983).

El haber detectado conversión génica aparente, nos llevó a considerar otras

posibilidades para este mecanismo que no han sido exploradas en este organismo y por lo tanto, los estudios de conversión génica se podrían extender a distintos aspectos de este proceso como son: tamaño de la región que puede ser convertida en la misma zona, conversión génica asociada a recombinación, direccionalidad, polaridad, conversión mediada por transcritos de RNA y regiones funcionales cortas que puedan ser convertidas.

**¿ Cual es el tamaño de la región que puede ser susceptible de conversión génica, considerando inserciones o deleciones ?**

En eucariotes se ha determinado que la región de conversión puede ser desde 20 pares de bases (pb) hasta 2 kilobases (kb) (Dover, 1987). Para procariotes se ha establecido en plásmidos pequeños (pBR) regiones de conversión desde 10 pb hasta 283 pb (Doherty *et al.*, 1983; Fishel *et al.*, 1986; Fishel y Kolodner, 1989; Rubnitz y Subramani, 1996; Yamamoto *et al.*, 1988). En el cromosoma, las zonas convertidas entre reiteraciones naturales son de 90 pb a 270 pb para *rm* en *E. coli* y *Salmonella* (Harvey y Hill, 1990; Mattatall y Sanderson, 1996; Mattatall *et al.*, 1996), mayores a 400 pb entre los espacios de *rrn* en *Neisseria* (Zhang *et al.*, 1992). La recombinasa específica del cambio de fase de los genes de flagelina convierte un mínimo de 565 pb hasta 1013 pb en *Salmonella* (Okazaki *et al.*, 1993). Kowalchuk *et al.* (1995) reportaron en *Acinetobacter calcoaceticus* la conversión de una región de 1598 pb entre los genes *cat IJF* y *pca IJF*.

Debido a la extensión de homología entre los genes *nifH* y posiblemente del operón *nifHDK*, podríamos especular que en este organismo la conversión podría

abarcara una región de una kb o mayor. La estrategia inicial para detectar el tamaño que puede ser convertido, es eliminar zonas de DNA entre el sitio del megalinker *1-SceI* y el cassette. Las deleciones se harían en la zona de 700 bp comprendida entre los marcadores utilizados (megalinker y cassette). Se realizarían pequeñas deleciones, las cuales se incrementarían mediante intervalos hasta cubrir la zona. Para evaluar el efecto de las inserciones en conversión génica, se usaría la misma región. Después del megalinker *1-SceI* queda un sitio de *Bgl II* el cual puede usarse para ir añadiendo oligonucleótidos ó regiones de DNA de diferente tamaño e incrementar la heterología en la secuencia.

### **¿ Ocurre conversión génica asociada a recombinación?**

En el sistema diseñado, la conversión génica asociada a recombinación puede detectarse en el sistema intramolecular, entre las secuencias reiteradas que están en orientación inversa en las regiones *nifHa* y *nifHc* (véase el artículo, Fig. 3). Al producirse recombinación entre la región *nifHa* y *nifHc* pueden ser recuperados dos productos posibles. La inversión genera dos bandas recombinantes, una banda sería la recombinante 5'c y la región 3'a con el cassette y el producto recíproco mantendría el linker en la posición original región 5'a y la región 3'c; pero si en el evento de inversión ocurre conversión génica la presencia del linker no sería detectada. Debido a que las inversiones ocurren a una frecuencia de  $10^{-6}$ , la limitante de este evento es incrementar el número de células en las cuales se pueda detectar la inversión para posteriormente identificar la presencia o ausencia del megalinker. De manera similar se puede evaluar la participación de conversión

génica asociada a recombinación entre la región *nifHc* y *nifHb* (véase el párrafo abajo).

**¿Existe direccionalidad en este sistema para conversión génica?**

Seguramente distintas regiones del genoma recombinan a diferentes frecuencias y algunos eventos de recombinación tienen sitios preferenciales de inicio o de término. El sistema diseñado para evaluar conversión génica fue insertado en la región *nifHa* (véase el artículo); el arreglo particular de los genes de la nitrogenasa en el pSim , permite ahora insertar el sistema en la región *nifHb*. Una manera similar de evaluar la contribución de la región *nifHa* y *nifHc* sobre conversión génica , es caracterizar los productos posibles de recombinación y la proporción en que se presenten con respecto a la inserción en *nifHa*, para determinar si hay direccionalidad.

En este caso los productos de delección serían mayoritarios, ya que la recombinación entre la región 5'a y 3'b con cassette, son seleccionados , el producto recíproco que sería una amplificación, la cual no es seleccionable a menos que primero ocurra conversión génica y posteriormente se amplifique (rondas múltiples de recombinación). La otra opción es la inversión entre la región *nifHc* y *nifHb* , en donde se producirían dos bandas recombinantes 5'c y 3'b con el cassette y la banda recíproca sería 5'b con el megalinker y 3'c, a la vez podría ocurrir conversión génica asociada a recombinación. Además, de los productos mencionados se presentaría conversión génica y rondas múltiples de recombinación.

Una estrategia para evaluar la contribución de la región *nifHc* a cada una de



las regiones *nifHa* y *nifHb*, consiste en introducir el sistema diseñado en las dos regiones. En este sistema todos los productos posibles de recombinación intramolecular son recuperados, de tal manera que sólo por inversiones, conversión génica y conversión génica asociada a recombinación se obtendrían clonas kanamicina resistentes y se podría evaluar la contribución específica de cada uno de ellos a la dinámica recombinacional de la familia de los *nifH*.

### **¿Existe polaridad en este sistema para conversión génica ?**

Los estudios de conversión génica en hongos han establecido que es común que se establezca co-conversión en alelos adyacentes, lo cual implica que el proceso involucra la presencia de un gradiente de conversión a través de la secuencia más que en un punto en el DNA.

La construcción del sistema diseñado para evaluar conversión génica, no tiene marcadores a los lados del megalinker y del cassette que permitan evaluar la existencia de polaridad en esta zona: Para determinar la presencia de polaridad necesitamos insertar marcadores, los cuales podrían ser los diseñados para las inserciones y además se tendrían que diseñar oligonucleotidos antes del megalinker (que no modifiquen la funcionalidad de los codones de paro ya establecidos), así con la presencia o ausencia de estos marcadores podríamos determinar si hay polaridad.

## ¿ La conversión génica puede ocurrir mediante RNA ?

La recombinación entre un transcrito reverso (cDNA) y su homólogo cromosomal ha sido propuesta en eucariotes como un mecanismo que está involucrado en la excisión de intrones del DNA y también participa en conversión génica entre secuencias reiteradas dispersas . Sin embargo, hay que considerar que el evento de conversión génica no esté asociado con el intercambio de marcadores externos, para que el donador de la información genética pueda ser un cDNA.

Derr y Strathern (1993) en *Saccharomyces cerevisiae* , diseñaron un sistema de recombinación que consiste en un plásmido donador del DNA complementario, que es homólogo al gene cromosomal utilizado como blanco (Fig. 7) y demostraron que la recombinación entre un transcrito reverso y su homólogo cromosomal participan en conversión génica. Este tipo de recombinación requiere de la expresión de la transcriptasa reversa para la formación de pseudogenes. Así, cualquier estructura que afecte al transcrito, su abundancia y su capacidad para servir como sustrato para transcriptasa reversa necesariamente influye en el rango de recombinación mediada por RNA.

En una colección de 63 cepas de Rhizobia se identificaron retrones, los cuales codifican a la transcriptasa reversa, estos elementos genéticos fueron localizados en las siguientes especies; *Bradyrhizobium* spp (5), *R. trifolii* (2) *R. loti* (2) y *R. sp* (1) (Rice *et al.*, 1993). Seis retrones producidos por las cepas de Rhizobia son únicos, presentan heterogenidad en secuencia nucleotídica y muestran poca similitud con los retrones de *E. coli* y myxobacterias (Rice *et al.*, 1993).

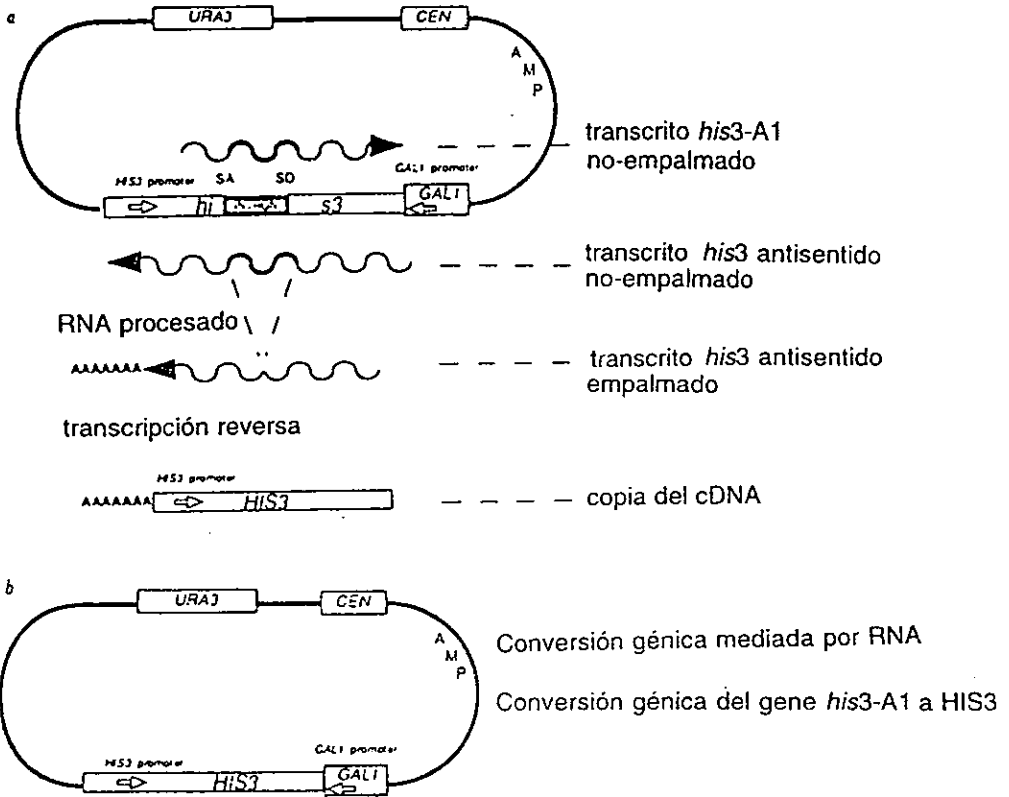


Fig. 7. Recombinación mediada por RNA en *Saccharomyces cerevisiae*. a) El plásmido GAL1-*his3*-A1 actúa como donador de HIS3 cDNA. El gene de levadura *his3* está interrumpido por un intrón artificial insertado en el sitio único MscI en la orientación no empalmable al promotor de HIS3, pero en orientación empalmable con el promotor GAL1. Los transcritos son líneas onduladas y el intrón es una línea gruesa. SD = donador de empalme; SA = aceptor de empalme. La formación de HIS3 cDNA requiere la transcripción reversa del transcrito antisentido *his3* empalmado. Este cDNA está involucrado en un evento subsecuente de recombinación. b) La conversión génica mediada por RNA de la secuencia *his3*-A1 del plásmido, resulta en su conversión a HIS3 (Tomado de Derr y Strathern, 1993).

El hallazgo de los retrones en *Rhizobium*, ofrece la posibilidad de detectar y diseñar un mecanismo para evaluar la contribución de la transcriptasa reversa en conversión génica. La propuesta, puede ser factible si detectamos primeramente la presencia de msDNA (multicopy single-stranded DNA) en *R. etli* y diseñamos en los genes *nifH* (no-retroelemento) un sistema de recombinación mediada por RNA, lo cual ofrece una vía distinta para conversión génica entre los miembros de la familia multigénica de la nitrogenasa.

Esta expectativa se ve apoyada por los estudios de Derr y Strathern (1993), con respecto a que conversión génica ocurre en un rango apreciable, entre un transcrito reverso de un no-retroelemento y su alelo cromosomal en *S. cerevisiae*. Además, los estudios de la actividad transcripcional de los genes simbióticos en condiciones microaeróbicas, indican que los transcritos de los genes *nifH* están presentes (Girard *et al.*, 1996), de manera que los RNA de interés podrían ser los templados para la actividad de transcriptasa reversa y producir los cDNA.

**¿Se podrían extender los estudios de conversión génica en regiones funcionales cortas en extensión ?**

Un caso interesante en *Rhizobium* es la presencia de secuencias de aproximadamente 400 bp presentes en diferentes genes simbióticos donde las regiones regulatorias están conservadas y los primeros 24 aminoácidos de la región amino terminal de *nifH* son similares (Better *et al.*, 1983; Watson y Schofield, 1985). Las secuencias serían un mecanismo de control de genes que se expresan en

bacteroide y que están coordinadamente regulados con los genes de fijación.

Para los genes *nfe* su expresión resulta en un incremento en la capacidad competitiva y eficiencia en nodulación, la cual puede iniciarse antes del disparo de los genes de fijación (Soto *et al.*, 1993). Los genes *mos* se expresan sólo cuando los genes simbióticos están funcionando, y están involucrados en la síntesis de rhizopinas (Murphy *et al.*, 1988; Murphy *et al.*, 1993; Saint *et al.*, 1993). Estas funciones auxiliares que promueven la interacción simbiótica, aseguran la sobrevivencia de la bacteria al proveer un sustrato selectivo de crecimiento, en el inicio de los procesos simbióticos.

Este punto es interesante en la medida que a través de duplicaciones se generarían secuencias de regulación para el control de zonas importantes en la relación simbiótica. Por la función que desempeñan algunos de estos genes, resulta de interés evaluar si el mecanismo de conversión génica se presenta en estas regiones reguladoras de genes importantes en el establecimiento de la relación simbiótica y así preservar la organización estructural de genes presentes en el nódulo.

Los estudios de conversión génica en *Rhizobium* se pueden extender a diferentes aspectos de la vida libre y de la relación simbiótica. Por el momento, he abordado los puntos principales en los cuales el avance de los trabajos permite realizar estudios relacionados con este mecanismo. Finalmente, el resolver parte o algunos de los temas sugeridos anteriormente, permitiría comprender los procesos de la mecánica recombinacional de los genes *nifH* y en el contexto evolutivo conllevaría al entendimiento de la dinámica de las familias multigénicas.

## CONCLUSIONES

1. Los eventos de conversión génica aparente constituyeron el 14% de los productos de recombinación que participaron en la corrección de una inserción de 28 pares de bases, localizada en un miembro de la familia multigénica de la nitrogensa.
2. Una variedad de eventos múltiples de recombinación (que incluyen conversión génica aparente), se presentaron en una proporción del 10%.
3. Las duplicaciones en tandem fueron los eventos de recombinación que contribuyeron en mayor proporción (74%) para la corrección observada.
4. Las inversiones se presentaron a baja frecuencia y en una proporción menor (2%).
5. La participación de recombinación en el origen de conversión génica aparente en la familia multigénica *nifH* es apoyado por (a) El requerimiento estricto del gene funcional *recA*; (b) la dependencia de copias adicionales silvestres de la región *nifH* para su formación y (c) el restablecimiento preciso de la secuencia nucleotídica silvestre.
6. La proporción en que se presentaron los eventos de recombinación pueden ser explicados por el modelo de recombinación de entrecruzamiento a la mitad.

## BIBLIOGRAFÍA

Abdulkarim, F. y D. Hughes. 1996. Homologous recombination between the *tuf* genes of *Salmonella typhimurium*. J. Mol. Biol., 260 : 506 - 522.

Anthamatten, D., Scherb, B. y H. Hennecke. 1992. Characterization of a *fix* LJ-regulated *Bradyrhizobium japonicum* gene sharing similarity with the *Escherichia coli* *fnr* and *Rhizobium meliloti* *fixK*. J. Bacteriol., 174 : 2111 - 2120.

Badenoch-Jones, J., Hulton, T.A., Morrison, C.M., Scott, K.F. and J. Shine. 1989. Structural and functional analysis of nitrogenase genes from the broad-host range *Rhizobium* strain ANU240. Gene., 77 : 110 -153.

Baev, N., Endre, G., Petrovics, G., Banfalvi, Z. y A. Kondorosi. 1991. Six nodulation genes of *nod* box locus 4 in *Rhizobium meliloti* are involved in nodulation signal production: *nod M* codes for D-glucosamine synthetase. Mol. Gen. Genet., 228 : 113 - 124.

Baev, N. y A. Kondorosi. 1992. Nucleotide sequence of the *Rhizobium meliloti* *nodL* gene located in locus n5 of the *nod* regulon. Plant. Mol. Biol., 18 : 843-846.

Balleza Mejía, Daniel. 1998. Análisis de secuencias de ADN reiterado y predicción de rearrreglos genómicos en *Rhizobium*. Tesis de Maestría en Investigación Biomédica Básica. UACP y P del CCH. UNAM, CIFN. Cuemavaca, Mor. México. 65 pp.

Barnett, M.J., Rushing, B.G., Fisher, R.F. y S.R. Long. 1995. Transcriptional start sites for *syrM* and *nodD* 3 flank an insertion sequence relic in *Rhizobium meliloti*. J. Bacteriol., 178 : 1782 - 1787.

Blattner, F.R., Plunkett 3rd, G., Bolch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. y Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science.*, 277 : 1453 - 1474.

Better, M., Lewis, B., Corbin, D., Ditta, G. and D. R. Helinski. 1983. Structural relationships among *Rhizobium meliloti* symbiotic promoters. *Cell.*, 35 : 479 - 485 .

Bergman, K., Nulty, E. y L. Su. 1991. Mutations in two flagellin genes of *Rhizobium meliloti* . *J. Bacteriol.*, 173 : 3716 - 3723 .

Birch, A., Häusler, A. y R. Hütter. 1990. Genome rearrangement and genetic instability in *Streptomyces.spp.* *J. Bacteriol.*, 172 : 4138 - 4142.

Brom, S., García de los Santos, A., Girard, M.L., Dávila, G., Palacios, R. y D. Romero. 1991. High-frequency rearrangements in *Rhizobium leguminosarum* bv. *phaseoli* plasmids. *J. Bacteriol.*, 173 : 1344 - 1346.

Brom, S., García de los Santos, A., Stepkowsky, T., Flores, M., Dávila, G., Romero, D. y R. Palacios. 1992. Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. *J. Bacteriol.*, 174 : 5183 - 5189.

Buck, M., Miller, S., Drummond, M. y R. Dixon. 1986. Upstream activator sequences are in the promoters of nitrogen fixation genes. *Nature.*, 320 : 374 - 378.

Colleaux, L., d'Auriol, L., Betemier, M., Cottarel, G., Jacquier, A., Galibert, F. y B. Dujon. 1986. Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *E.coli* as a specific double strand endonuclease. *Cell.*, 44 : 521 - 523.



Colleaux, L., d'Auriol, L., Galibert, F. y B. Dujon .1988. Recognition and cleavage site of the intron-encoded omega transposase. P. N. A. S. USA., 85 : 6022-6026.

David, M., Domergue, O., Pognonec, P. y D. Kahn. 1987. Transcription patterns of *Rhizobium meliloti* symbiotic plasmid pSym: Identification of *nifA*-Independent *fix* genes. J. Bacteriol., 169 : 2239 - 2244.

de Bruijn, F.J. 1992. Use of repetitive (Repetitive Extragenic Palindromic and Enterobacterial Repetitive Intergenic Consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. Appl. Environ. Microbiol., 58 : 2180 - 2187.

Derr, L.K. y J.N. Strathern. 1993. A role for reverse transcripts in gene conversion. Nature, 341: 170 -173.

Doherty, M.J., Morrison, P.T. y R. Kolodner. 1983. Genetic recombination of bacterial plasmid DNA. Physical and genetic analysis of the products of plasmid recombination in *Escherichia coli*. J. Mol. Biol., 167 : 539 - 560.

Dover, G.1982. Molecular drive: a cohesive mode of species evolution. Nature., 299 : 111 - 116.

Dover, G. A.1986. Molecular drive in multigene families : how biological novelties arise, spread and are assimilated. Trends Genet., 2 : 159 -165.

Dover, G.A. 1987. DNA turnover and the molecular clock. J. Mol. Evol., 26 : 47 - 58.

Dover, G. A.1992. Observing development through evolutionary eyes: A practical approach. BioEssays., 14 : 281- 287.

Dover, G.A. 1993. Evolution of genetic redundancy for advanced players. *Curr. Opin. Genet. Dev.*, 3 : 902 - 910.

Dusha, I., Kovalenko, S., Banfalvi, Z. y A. Kondorosi. 1987. *Rhizobium meliloti* insertion element ISRm2 and its use for identification of the *fixX* gene. *J. Bacteriol.*, 169 : 1403 - 1409.

Earl, C.D., Ronson, C.W. y F.M. Ausubel. 1987. Genetic and structural analysis of the *Rhizobium meliloti fixA, fixB, fixC* and *fixX* genes. *J. Bacteriol.*, 169 : 1127 - 1136.

Fischer, H.M., Babst, M., Kaspar, T., Acuña, G., Arigoni, F. y H. Hennecke. 1993. One member of a *groES-EL*-like chaperonin multigene family in *Bradhyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. *EMBO J.*, 12 : 2901 - 2912.

Fischer, H-H. 1994. Genetic regulation of nitrogen fixation in Rhizobia. *Microbiol. Rev.*, 58 : 352 - 386.

Fishel, R., Siegel, E.C. y R. Kolodner. 1986. Gene conversion in *Escherichia coli*: resolution of heteroallelic mismatched nucleotides by co-repair. *J. Mol. Biol.*, 188 : 147 - 157.

Fishel, R. y R. Kolodner. 1989. Gene conversion in *Escherichia coli* : the RecF pathway for resolution of heteroduplex DNA. *J. Bacteriol.*, 171 : 3046 - 3052.

Flores, M., González, V., Brom, S., Martínez, E., Piñero, D., Romero, D., Dávila, G. y R. Palacios. 1987. Reiterated DNA sequences in *Rhizobium* and *Agrobacterium* spp. *J. Bacteriol.*, 169 : 5782 - 5788.

Flores, M., González, V., Pardo, M. A., Leija, A., Martínez, E., Romero, D. Piñero, D., Dávila, G. y R. Palacios. 1988. Genomic instability in *Rhizobium phaseoli*. J. Bacteriol., 170 : 1191 - 1196.

Flores, M., Brom, S., Stepkowski, T., Girard, M.L., Dávila, G., et al .1993. Gene amplification in *Rhizobium*: identification and in vivo cloning of discrete amplifiable DNA regions (amplicons) from *Rhizobium leguminosarum* biovar *phaseoli*. P. N. A. S. USA., 90 : 4932 - 4936.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A. y X. Perret. 1977. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature, 387 : 394 - 401.

Futuyma, D.J. 1991. Evolutionary Biology. 2Ed. Sinauer Associates, Inc. Sunderland, Massachusetts, USA. p : 443 - 481.

García de los Santos, A., Brom, S. y D. Romero. 1996. *Rhizobium* plasmids in bacteria-legume interactions. World J. Microbiol. & Biotechnol., 12 : 119 - 125.

Geelen, D., Goethals, K., Van Montagu, M. y M. Holsters. 1995. The *nodD* locus from *Azorhizobium caulinodans* is flanked by two repetitive elements. Gene., 164 : 107 - 111.

Girard, L., Valderrama, B., Palacios, R., Romero, D. y G. Dávila. 1996. Transcriptional activity of the symbiotic plasmid of *Rhizobium etli* is affected by different environmental conditions. Microbiology., 142 : 2847 -2856.

Hahn, M. y H. Hennecke. 1987. Conservation of a symbiotic DNA region in soybean root nodule bacteria. Appl. Environ. Microbiol., 53 : 2253 - 2255.

Hastings, P.J. 1988. Conversion events in fungi. *In*: Genetic Recombination. (Eds.) Kuchertapati, R. y G.R. Smith. ASM Press. Washington DC p: 23 - 72

Harvey, S. y C. W. Hill. 1990. Exchange of spacer regions between rRNA operons in *Escherichia coli*. *Genetics.*, 125 : 683 - 690.

Harvey, S., Hill, C. W., Squires, C. y C.L. Squires. 1988. Loss of the spacer loop sequence from the *rmB* in the *Escherichia coli* K-12 subline that bears the *relA1* mutation. *J. Bacteriol.*, 170 : 1235 - 1238.

Hernández-Lucas, I., Pardo, M.A., Segovia, L., Miranda, J. y E. Martínez-Romero. 1995. *Rhizobium tropici* chromosomal citrate synthase gene. *Appl. Environ. Microbiol.*, 61 : 3992 - 3997.

Hill, C.W. y G. Combrato. 1973. Genetic duplication induced at very high frequency by ultraviolet irradiation in *Escherichia coli*. *Mol .Gen. Genet.*, 127 : 197 - 214.

Hill, C.W. y B.W. Hamish. 1981. Inversions between ribosomal RNA genes of *Escherichia coli*. *P. N. A. S. USA.*, 78 : 197 - 214.

Hill, C.W. y B.W. Hamish. 1982. Transposition of a chromosomal segment bounded by redundant rRNA genes into other rRNA genes in *Escherichia coli*. *J. Bacteriol.*, 149 : 449 - 457.

Hill, C.W. y J.A. Gray. 1988. Effects of chormosomal inversion on cell fitness in *Escherichia coli*. K-12. *Genetics.*, 119 : 771 - 778.

Homma, M.A. y F.M. Ausubel. *Rhizobium meliloti* has three functional copies of the *nodD* symbiotic regulatory gene. *P. N. A. S. USA.*, 84 : 8558 -

John, B. y G.L.G. Miklos. 1988. The Eukaryote Genome in Development and Evolution. Allen & Unwin, London. p : 1 - 25 y 234 - 244.

Judd, A. y M.J. Sadowsky. 1993. The *Bradyrhizobium japonicum* serocluster 123 hyperreiterated DNA region, HSR1, has DNA and aminoacid sequence homology to IS1380, an insertion sequence from *Acetobacter pasteurianus*. Appl. Environ. Microbiol., 59 : 1656 - 1661.

Kaluza, K., Hahn, M. and H. Hennecke. 1985. Repeated sequences similar to insertion elements clustered around the *nif* region of the *Rhizobium japonicum* genome. J. Bacteriol., 162 : 535 - 542.

Kourilsky. P. 1986. Molecular mechanisms for gene conversion in higher cells. Trends Genet., 2 : 60 - 63.

Kowalchuk, G.A., Gregg-Jolly, L.A. y L. N. Omston. 1995. Nucleotide sequences transferred by gene conversion in the bacterium *Acinetobacter calcoaceticus*. Gene., 153 : 111 - 115.

Kullik, I., Fritsche, S., Knobel, H., Sanjuan, J., Hennecke, H. y H-M. Fischer. 1991. *Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the sigma 54 gene (*rpo M*). J. Bacteriol., 173 : 1125 - 1138.

Kündig, C., Beck, C., Hennecke, H. y M Göttfert. 1995. A single rRNA gene region in *Bradyrhizobium japonicum*. J. Bacteriol., 177 : 5151 - 5154.

Laberge, S., Middleton, A.T. y R. Wheatcroft. 1995. Characterization, nucleotide sequence, and conserved genomic locations of insertion sequences *ISRm 5* in *Rhizobium meliloti*. J. Bacteriol., 177 : 3133 - 3142.

Leach, J.E., White, F.F., Rmoas, M.I. y H. Leung. 1990. A repetitive DNA sequence differentiates *Xanthomonas campestris pv.oryzae* from other pathovars of *X. campestris*. Mol. Plant-Microbe. Int., 3 : 238-246.

Lehner, A.F. y C.W. Hill. 1985. Merodiploidy in *Escherichia coli-Salmonella typhimurium* crosses: the role of unequal recombination between ribosomal RNA genes. Genetics, 110 : 365 - 380.

Li, W-H. y D. Graur. 1991. Fundamentals of Molecular Evolution. Sinauer Associates, Inc. Sunderland, Massachusetts, USA. p : 162 -170.

Lloyd, R.G. y K.B. Low. 199 . Homologous Recombination. In *Escherichia coli* and *Salmonella typhimurium* : Cellular and Molecular Biology. 2Ed. Edited by Neidhart, F.C., Curtiss III, R., Ingraham, L., Lin, E.C.C., Low Jr, K.B. *et al* . American Society for Microbiology Press, Washington, D.C. USA. p : 2236 - 2255.

Long, S. R. 1989. *Rhizobium* genetics. Ann. Rev. Genet., 23: 483 - 506.

Marie, C., Bamy, M.A. y A. Downie. 1992. *Rhizobium leguminosarum* has two glucosamine synthetases, *GlmS* and *NodM*, required for nodulation and development of nitrogen-fixating nodules. Mol. Microbiol., 6 : 843 - 851.

Margolin, W., Corbo, J.C. y S. R. Long. 1991. Cloning and characterization of a *Rhizobium meliloti* homolog of the *Escherichia coli* cell division gene *ftsZ*. J. Bacteriol., 173 : 5822 - 5830.

Margolin, W. y S.R. Long. 1994. *Rhizobium meliloti* contains a novel second homolog of the cell division gene *ftsZ*. J.Bacteriol., 176 : 2033 - 2044.

Martínez-Salazar, J.M. 1991. Clonación y caracterización del gen *recA* de *Rhizobium phaseoli*. Tesis de Doctorado en Investigación Biomédica Básica. UACPyP del CCH. UNAM. CIFN . Cuemavaca, Mor. México. 65pp.

Martínez,E., Romero,D. y R. Palacios.1990. The *Rhizobium* genome. Crit. Rev. Plant. Sci., 9 : 59 -93.

Mattatall, N.R y K. E. Sanderson. 1996. *Salmonella typhimurium* LT2 possesses three distinct 23S rRNA intervening sequences. J. Bacteriol., 178 : 2272 - 2278.

Mattatall, N.R., Daines, D.A., Liu, S-L. y K. E. Sanderson. 1996. *Salmonella typhi* contains identical intervening sequences in all seven *rm* genes. J. Bacteriol., 178 : 5323 - 5326.

Mavingui, P., Flores, M., Romero, D., Martínez-Romero, E. y R. Palacios. 1997. Generation of *Rhizobium* strains with improved symbiotic properties by random DNA amplification (RDA). Nature Biotechnol., 15 : 564 - 569.

Mavingui, P., Laeremans, T., Flores, M., Romero, D., Martínez-Romero, E. y R. Palacios. 1998. Genes essential for Nod factor production and nodulation are located on a symbiotic amplicon (AMP*Rtr*/CFN299pc60) in *Rhizobium tropici*. J. Bacteriol., 180 : (En prensa).

Mazurier, S-I., Rigottier-Gois, L. y N. Amarger. 1996. Characterization, distribution, and localization of *ISRI* 12, an insertion sequence element from *Rhizobium leguminosarum* bv. *viciae*. Appl. Environ. Microbiol., 62 : 685 - 693.

Morales, V.M., Bäckman, A. y M. Bagdasarian. 1991. A series of wide-host-range low-copy-number vectors that allow direct screening for recombinants. *Gene*, 97: 39 - 47.

Murphy, P.J., Heycke, N., Trenz, S.P., Ratet, P., de Bruijn, F.J. y J.Schell. 1988. Synthesis of an opine-like compound, a rhizopine, in alfalfa nodules is symbiotically regulated. *P. N. A. S. USA.*, 85 : 9133 - 9137.

Murphy, P.J., Trenz, S.P., Grzemski, W., de Bruijn, F. J. y J. Schell. 1993. The *Rhizobium meliloti* rhizopine *mos* locus is a mosaic structure facilitating its symbiotic regulation. *J.Bacteriol.*, 175 : 5193 - 5204.

Narberhaus, F., Krummenacher, P., Fischer, H-M. y H. Hennecke. 1997. Three disparately regulated genes for sigma<sup>32</sup>-like transcription factors in *Bradyrhizobium japonicum*. *Mol. Microbiol.*, 24 : 93 - 104.

Norel, F. y C. Elmerich. 1987. Nucleotide sequence and functional analysis of the two copies of *Rhizobium* ORS571. *J. Gen. Microbiol.*, 133 : 1563-1576.

Ogawa, J., Brierley, H.L. y S.R. Long. 1991. Analysis of *Rhizobium meliloti* nodulation mutant WL131 : novel insertion sequence *ISRm 3* in *nod G* and altered *nod H* protein product. *J. Bacteriol.*, 173 : 3060 - 3065.

Okazaki, N., Matsuo, S., Saito, K., Tominaga, A. y M. Enomoto. 1993. Conversion of the *Salmonella phase 1* flagellin gene *fliC* to the phase 2 gene *fljB* on the *Escherichia coli* K-12 chromosome. *J. Bacteriol.*, 175 : 758 -766.



Osteras, M., Stanley, J. y T.M. Finan. 1995. Identification of *Rhizobium* -specific intergenic mosaic elements within an essential two-component regulatory system of *Rhizobium* species. *J. Bacteriol.*, 177 : 5485 - 5494.

Pardo, M.A., Lagúnez, J., Miranda, J. y E. Martínez. 1994. Nodulating ability of *Rhizobium tropici* is conditioned by a plasmid-encoded citrate synthase. *Mol. Microbiol.*, 11 : 315 - 321.

Perret, X., Viprey, V., Freiberg, C. y W.J. Broughton. 1997. Structure and evolution of NGRRS-1, a complex, repeated element in the genome of *Rhizobium* sp. strain NGR234. *J. Bacteriol.*, 179 : 7488 - 7496.

Plazannet, C., Réfrégier, G., Demont, N., Truchet, G. y C. Rosenberg. 1995. The *Rhizobium meliloti* region located downstream of the nod box n6 is involved in the specific nodulation of *Medicago lupulina*. *FEMS Microbiol. Letters.*, 133 : 285 - 291.

Priefer, V., Burkardt, H., Klipp, W. y A. Puhler. 1980. ISR1: an insertion element isolated from the soil bacterium *Rhizobium lupini*. *Cold. Spring Harbor Symp. Quant. Biol.*, 45 : 87.

Pleier, E. y R. Schmitt. 1989. Identification and sequence analysis of two related flagellin genes in *Rhizobium meliloti*. *J. Bacteriol.*, 171 : 1467 - 1475.

Plessis, A., Perrin, A., Haber, J.E. y B. Dujon. 1992. Site-specific recombination determined by *1-Scel*, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics.*, 130 : 451- 460.

Quinto, C., De la Vega, H., Flores, M., Leemans, J., Cevallos, M.A., Pardo, M.A., Azpiroz, R., Girard, M.I., Calva, E. y R. Palacios. 1985. Nitrogenase reductase: A functional multigene family in *Rhizobium phaseoli*. *P. N. A. S. USA.*, 82 : 1170-1174.

- Renalier, M-H., Batut, J., Ghai, J., Terzaghi, B., Gherardi, M., David, M., Garnerone, A-M., Vasse, J., Truchet, G., Huguet, T. y P. Bositard. 1987. A new symbiotic cluster on the pSym megaplasmid of *Rhizobium meliloti* 2011 carries a functional *fix* gene repeat and a *nod* locus. J. Bacteriol., 169 : 2231 - 2238.
- Rice, D.J., Somasegaran, P., MacGlashan, K. y B. Bohlool. 1994. Isolation of insertion sequence ISRLdTAL1145-1 from a *Rhizobium* sp (*Leucaena diversifolia*) and distribution of homologous sequence identifying cross-inoculation group relationships. Appl. Environ. Microbiol., 60 : 4394 - 4403.
- Rice, S.A., Bieber, J., Chun, J-Y., Stacey, G. y B.C. Lampson. 1993. Diversity of retron elements in a population of rhizobia and other gram-negative bacteria. J. Bacteriol., 175 : 4250 -4254.
- Rochepeau, P., Selinger, L.B. y M.F. Hynes. 1997. Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum* VF39SM. Mol. Gen. Genet., 256 : 387 - 396.
- Romero, D., Singleton, P.W., Segovia, L., Morett, E., Bohlool, B.B., Palacios, R. y G. Dávila. 1988. Effect of naturally occurring *nif* reiterations on the symbiotic effectiveness in *Rhizobium phaseoli*. Appl. Environ. Microbiol., 54 : 848 - 850.
- Romero, D., Brom, S., Martínez-Salazar, J., Girard, M.L., Palacios, R. y G. Dávila. 1991. Amplification and deletion of a *nod-nif* region in the symbiotic plasmid of *Rhizobium phaseoli*. J. Bacteriol., 173 : 2435-2441.
- Romero, D., Martínez-Salazar, J., Girard, L., Brom, S., Dávila, G., Palacios, R., Flores, M. y C. Rodríguez. 1995. Discrete amplifiable regions (amplicons) in the symbiotic plasmid of *Rhizobium etli* CFN42. J. Bacteriol., 177 : 973 -980.

Romero, D. y R. Palacios.1997. Gene Amplification and genomic plasticity in prokaryotes. *Annu. Rev. Genet.*, 31 : 91 - 111.

Romero, D., Dávila, G. y R. Palacios. 1998a. The Dynamic Genome of *Rhizobium*. In : *Bacterial Genomes, physical structure and analysis*. (Eds) de Bruijn, F.J. *et al.*, 1998a. Chapman & Hall, New York, NY, USA. p : 153 - 161.

Romero, D., Martínez-Salazar, J.,Palacios, R., Rodríguez, C. y E. Valencia-Morales. 1998b. The Dynamic genome of *Rhizobium* : Addressing some basic questions.In : *Highlights in Nitrogen Fixation*. (Eds) Martínez, E. y G. Hernández. Plenum,February 1st - 6th, Cancún, México. p: 1 - 5.

Rostas, K., Kondorosi, E., Horvath, B., Simoncsits, A. y A. Kondorosi. 1986. Conservation of extended promoter regions of nodulation genes in *Rhizobium*. *P. N. A. S. USA.*, 83 : 1757 -1761.

Rubnitz, J. y S. Subramani. 1996. Extrachromosomal and chromosomal gene conversion in mammalian cells. *Mol. Cell. Biol.*, 6 : 1608 - 1614.

Rusanganwa, E. y R.S. Gupta. 1993. Cloning and characterization of multiple *groEL* chaperonin-encoding genes in *Rhizobium meliloti*. *Gene.*, 126 : 67 - 75.

Ruvkun, G.B., Long, S.R., Meade, H.M., van der Bos, R.C. y F.M. Ausubel. 1982. ISRm1 : a *Rhizobium meliloti* insertion sequence that transposes preferentially into nitrogen fixation genes. *J. Mol. Appl. Genet.*, 1 : 405 - 418.

Saint, C.P., Wexler, M., Murphy, P.J. Tempé, J., Tate, M.T. y P.J. Murphy. 1993. Characterization of genes for synthesis and catabolism of a new rhizopine induced in nodules by *Rhizobium meliloti* Rm 220-3 : Extension of the rhizopine concept. *J.Bacteriol.*, 175 : 5205 -5215.

Sapienza, C. y W.F. Doolittle. 1982. Unusual physical organization of the *Halobacterium* genome. *Nature.*, 295 : 384- 389.

Schlüter, A., Patschkowski, T., Quandt, J., Selinger, B., Weidner, S., Krämer, M., Zhou, L., Hynes, M.F. y U. B. Priefer. 1997. Functional and regulatory analysis of the two copies of the *fix* NOPQ operon of *Rhizobium leguminosarum* strain VF39. *Mol. Plant. Microbe-Interact.*, 10 : 605 - 616.

Schultze, M., Kondorosi, E., Ratet, P., Buiré, M. y A. Kondorosi. 1994. Cell and molecular biology of *Rhizobium* -plant interactions. *Int. Rev. Cytol.*, 156 : 1 - 75.

Schwedock, J. y S. Long. 1992. *Rhizobium meliloti* genes involved in sulfate activation: the two copies of *nodPQ* and a new locus, *saa*. *Genetics.*, 132 : 899 - 909.

Schwedock, J. y S.R. Long. 1994. An open reading frame downstream of *Rhizobium meliloti nodQ1* shows nucleotide sequence similarity to an *Agrobacterium tumefaciens* insertion sequence. *Mol. Plant. Microbe-Interact.*, 7 : 151 - 153.

Shibata, T., Nakagawa, K-I. y N. Morishima. 1995. Multi-site-specific endonucleases and the initiation of homologous genetic recombination in yeast. *Adv. Biophys.*, 31 : 77 - 91.

Sobral, B.W.S., Honeycutt, R.J. y A.G. Atherly. 1991. The genomes of the family Rhizobiaceae: Size, stability and rarely cutting restriction endonucleases. *J. Bacteriology.*, 173 : 704 - 709.

Soto, M.J., Zorzano, A., Olivares, J. y N. Toro. 1992. Sequence of ISRm4 from *Rhizobium meliloti* strain GR4. *Gene.*, 120 : 125 -126.

Soto, M.J., Zorzano, A., Mercado-Blanco, J., Lepek, V., Olivares, J. y N. Toro. 1993. Nucleotide sequence and characterization of *Rhizobium meliloti* nodulation competitiveness genes *nfe*. J. Mol. Biol., 229 : 570 - 576.

Stahl, F.W. 1994. The Holliday junction on its thirtieth anniversary. Genetics, 138 : 241 - 246.

Stahl, F. 1996. Meiotic recombination in yeast : Coronation of the double-strand-break repair model. Cell., 87 : 965 - 968.

Stepkowsky, T., Flores, M., Brom, S., Romero, D. y R. Palacios. 1993. Amplification and deletion events in *Rhizobium leguminosarum* bv trifolii ANU843. In :New Horizons in Nitrogen Fixation. (eds) Palacios, R., Mora, J. y W.E. Newton. Kluwer Academic Publishers, Netherlands. p : 654.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. and F.W. Stahl. 1983. The double-strand-break repair model for recombination. Cell., 33 : 25 - 35.

Thierry, A., Perrin, A., Boyer, J., Fairhead, C., Dujon, B., Frey, B. y G. Schmitz. 1991. Cleavage of yeast and bacteriophage T7 genomes at single site using the rare cutter endonuclease *1-SceI*. Nucleic. Acids. Research., 19 : 189 - 190.

Wheatcroft, R. y R.J. Watson. 1988. Distribution of insertion sequence ISRm1 in *Rhizobium meliloti* and other gram-negative bacteria. J. Gen. Microbiol., 134 : 113 - 121.

Wheatcroft, R. y S. Laberge. 1991. Identification and nucleotide sequence of *Rhizobium meliloti* insertion sequence ISRm3 : similarity between the putative transposase encoded by ISRm3 and those encoded by *Staphylococcus aureus* IS256 and *Thiobacillus ferrooxidans* IST2. J. Bacteriol., 173 : 2530 - 2538.

Watson, J.M. and P. R. Schofield. 1985. Species-specific, symbiotic plasmid-located repeated DNA sequences in *Rhizobium trifolii*. Mol. Gen. Genet., 199 : 279 - 289.

Yamamoto, K., Yoshikura, H., Takahashi, N. y I. Kobayashi. 1988. Apparent gene conversion in an *Escherichia coli* rec+ strain is explained by multiple rounds of reciprocal crossing-over. Mol. Gen. Genet., 212 : 393 - 404.

Yap, W.Y. y R.H. Schiestl. 1995. Nature of abortive transformation in *Saccharomyces cerevisiae*. Curr. Genet., 28 : 517 - 520.

Zhang, Q.Y., DeRyckere, D., Lauer, P. y M. Koomey. 1992. Gene conversion in *Neisseria gonorrhoeae* : Evidence for its role in pilus antigenic variation. P. N. A. S. USA., 89 : 5366 - 5370.