

Análisis *in silico* de familias multigénicas en *Rhizobium etli* CFN42 y comparación con otras rhizobias

Julie E. Hernández-Salmerón, Gustavo Santoyo

Laboratorio de recombinación y diversidad genómica. Instituto de Investigaciones Químico Biológicas, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Mich. Ciudad Universitaria, Edif. A. C.P. 58030

Resumen

En el presente trabajo se analizaron *in silico* el número de copias y la función de las familias multigénicas de *Rhizobium etli* CFN42. Además, se compararon con aquellas descritas en *Rhizobium leguminosarum*, *Rhizobium sp.* NGR234, *Bradyrhizobium japonicum*, *Sinorhizobium meliloti* y *Mesorhizobium loti*. Los resultados sugieren que *R. etli* presenta la mayor cantidad de genes reiterados y que no hay una correlación directa entre el número de repeticiones y el tamaño del genoma. Así mismo, se determinó que el mayor porcentaje de las repeticiones encontradas en *R. etli* se localizaron en plásmidos, cuyas funciones están fuertemente asociadas a la simbiosis con plantas.

Palabras clave: secuencias reiteradas, rhizobias, conversión génica.

Abstract

In the present study, the number of copies and the function of the multigenic families of *Rhizobium etli* CFN42, were *in silico* analyzed. In addition, they were compared with those described in *Rhizobium leguminosarum*, *Rhizobium sp.* NGR234, *Bradyrhizobium japonicum*, *Sinorhizobium meliloti* and *Mesorhizobium loti*. The results suggest that *R. etli* displays the greater amount of repeated genes and that there is not a direct correlation between the number of repetitions with the size of the genome. Also, we determined that a great percentage of this repetitions found in *R. etli* is located in plasmids, whose functions are strongly associated to the symbiosis with plants.

Key words: reiterated sequences, rhizobia, gene conversion.

Introducción

La recombinación homóloga se define como el proceso por el cual dos secuencias de ADN idénticas o altamente similares intercambian información genética. Este es un mecanismo de intercambio genético que juega importantes papeles durante la reparación del ADN, la segregación de cromosomas, la generación de variación genética, por mencionar algunos (West, 2003; Mariani, 2004). Así, los eventos de recombinación homóloga requieren la presencia de secuencias repetidas.

Se ha sugerido que dicha característica es común en diversos genomas bacterianos (Romero *et al.*, 1999) y que tales reiteraciones pueden estar localizadas en cromosomas o plásmidos, o en ambos, lo que podría generar diversos rearrreglos genómicos. Un resultado de un evento de recombinación entre secuencias repetidas del tipo entrecruzamiento o crossover, generaría amplificaciones o deleciones (Orozco-Mosqueda, *et al.* 2009). En una deleción, se podrían perder genes con funciones importantes que pondrían en riesgo la viabilidad celular, mientras que en las amplificaciones se incrementaría el número de copias de un segmento de ADN. Por otra parte, un segundo resultado sería un evento de conversión génica, la cual se define como la donación unidireccional de información genética entre las dos moléculas que interactúan (Szostak *et al.*, 1983). Cabe destacar que la conversión génica está asociada a la evolución concertada de familias multigénicas, así como la variación antigénica en bacterias patógenas que ayudan a evadir el sistema inmune del huésped (Santoyo y Romero, 2005).

Las reiteraciones multigénicas y los procesos de recombinación homóloga han sido ampliamente estudiados en sistemas eucariotes y procariotes, tales como: *Saccharomyces cerevisiae* (Kuzminov, 2001), *Escherichia coli* (Takahashi y Kobayashi, 1990), *Salmonella typhimurium* (Alani *et al.*, 1994) y *Bacillus subtilis* (Belle *et al.*, 2004). Sin embargo, otros grupos bacterianos como las bacterias fijadoras de nitrógeno (conocidas como rhizobias) han recibido poca atención. Una probable razón es por su importancia agronómica, aunque sin duda representan un sistema interesante dentro del área de la recombinación homóloga por la gran cantidad de secuencias repetidas que presentan sus genomas (Romero *et al.*, 1999; González *et al.*, 2006).

En este trabajo se analizó *in silico* la presencia, número de copias y función de familias multigénicas del genoma de *R. etli*. Además, hacemos un análisis comparativo con las reiteraciones encontradas en otros genomas de rhizobias representativas (*R. leguminosarum*, *Rhizobium sp.* NGR234, *Bradyrhizobium japonicum*, *Sinorhizobium meliloti* y *Mesorhizobium loti*). Es de resaltar que se analizaron únicamente familias multigénicas por su importancia biológica; aunque existen otro tipo de reiteraciones, tales como transposones, secuencias de inserción o aquellas no codificantes.

Métodos

Se realizó la búsqueda de genes repetidos reportados en la literatura, empleando la base de datos PUBMED (<http://www.ncbi.nlm.nih.gov/pubmed>) del National Center for Biotechnology

Autor de correspondencia: Dr. Gustavo Santoyo, Laboratorio Biológicas, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Mich. Ciudad Universitaria, Edif. A. C.P. 58030. Email: gsantoyo@yahoo.com, gustavo_santoyo@yahoo.com

Information (NCBI). Posteriormente, se corroboró la presencia de cada copia de las reiteraciones por búsquedas tipo BLAST del NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), obteniendo el nombre del gen repetido, localización genómica, función y su número de copias. Así mismo, se obtuvo y analizó la secuencia completa del genoma de cada rhizobia: *R. etli* (NC_007761.1), *Rhizobium* sp. NGR234 (NC_012587.1), *R. leguminosarum* (NC_008380.1), *B. japonicum* (NC_004463.1), *S. meliloti* (NC_003047.1) y *M. loti* (NC_002678.2) para corroborar la localización genómica de cada familia multigénica.

Resultados y discusión

Los genomas de bacterias fijadoras de nitrógeno, comúnmente conocidos como rhizobias, presentan diferencias importantes en el tamaño y en arquitectura genómica (Tabla 1). Algunas bacterias contienen varios plásmidos como el caso de *Rhizobium etli*, que cuenta con un cromosoma y seis plásmidos, mientras que el genoma de *Bradyrhizobium japonicum* está compuesto de un solo cromosoma y carece de plásmidos. El análisis realizado reveló que la bacteria *R. etli* posee el mayor número de genes repetidos, con 31, mientras que *R. leguminosarum* y *Rhizobium* sp. NGR234 presentan únicamente 12. El genoma más grande analizado en este trabajo fue el de *B. japonicum*, en el que se encontraron 21 reiteraciones que pertenecen a familias génicas, lo que sugiere que no existe una correlación entre el número de repeticiones y el tamaño del genoma (Figura 1). Sin embargo, es de resaltar que *R. etli* es la rhizobia más estudiada en el área de la recombinación homóloga y dinámica de genomas, por lo que no se descarta la posibilidad de que, conforme se estudien en más detalle otras rhizobias, éstas contengan un mayor número de repeticiones génicas (Quinto *et al.*, 1982; Romero *et al.*, 1991 y 1995).

Las diferentes reiteraciones encontradas se localizaron en diversos replicones, ya sea cromosoma o plásmido, o en ambos.

Tabla 1. Contenido y estructura genómica en rhizobias representativas.

Especie	Plásmidos (pb)	Cromosoma (pb)	Genoma total (pb)
<i>Bradyrhizobium japonicum</i>	0	9,105,828	9,105,828
<i>Rhizobium leguminosarum</i>	6 (2,694,167)	5,057,142	7,751,309
<i>Mesorhizobium loti</i>	2 (560,226)	7,036,071	7,596,297
<i>Rhizobium</i> sp. NGR234	2 (2,966,198)	3,925,702	6,891,900
<i>Sinorhizobium meliloti</i>	2 (3,037,559)	3,654,135	6,691,694
<i>Rhizobium etli</i>	6 (2,148,620)	4,381,608	6,530,228

Esto sugiere que dichas reiteraciones pueden ser blanco para llevar a cabo eventos de recombinación homóloga, y así, generar eventos de conversión génica y/o diversos rearrreglos genómicos (Orozco-Mosqueda *et al.*, 2009). Estos datos están en acuerdo con González *et al.* (2006) donde reportan más de 130 secuencias repetidas presentes en los diferentes replicones del genoma de *R. etli* CFN42. Es importante remarcar que no todas las secuencias repetidas encontradas por González *et al.* (2006) forman parte de familias multigénicas, y que además, podrían ser secuencias pequeñas (aproximadamente 100 pb) que no serían objetivos para llevar a cabo eventos de recombinación homóloga. Se ha sugerido que el mínimo requerido para llevar a cabo eventos de recombinación homóloga a una alta frecuencia es de 300 pb o más en *E. coli* y *R. etli* (Santoyo *et al.*, 2005). Por lo tanto, la presencia de familias multigénicas, que pueden ser genes de más de 1000 pb y con una alta similitud es más probable que participen en eventos de evolución concertada o rearrreglos genómicos (Hansen *et al.*, 2000; Guo *et al.*, 2003; Santoyo y Romero, 2005).

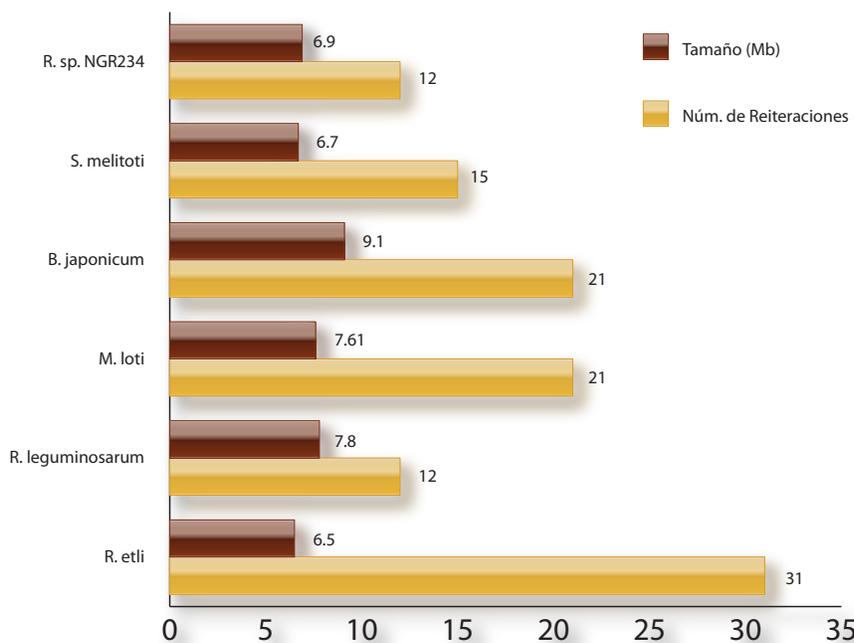


Figura 1. Número de genes repetidos en diferentes especies de rhizobias y comparación con el tamaño de genoma.

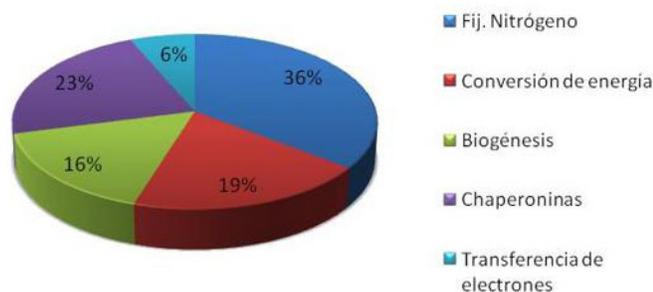


Figura 2. Función de los genes reiterados en *R. etli*.

Las rhizobias pueden vivir de manera saprofitica o formando simbiosis con diversas especies de plantas (Bena *et al.*, 2005). Para ello, requieren de la función de diversos genes que se encuentran localizados principalmente en plásmidos (Brom *et al.*, 1991). Por lo anterior cabe destacar que entre las reiteraciones encontradas están los genes: *fix*, *groEL*, *groES*, *nifH*, *nifD*, *nifK*, *nodD*, *dnaJ*, cuyos productos representan el 36% de las repeticiones y se encuentran relacionados con la nodulación y fijación de nitrógeno (Figura 2); asimismo, es importante destacar que el 66% de genes reiterados en *R. etli* se localizaron en los plásmidos (Figura 3). Hasta el momento se desconoce con precisión la razón del por qué los genomas de rhizobias contienen varias copias de genes que participan en procesos simbióticos. Se ha argumentado que por medio de eventos de conversión génica, que conlleven a la evolución concertada de la familia multigénica se podría evitar la generación de pseudogenes; sin embargo, un pseudogene de la misma manera podría convertir un gen funcional en una copia más sin actividad. Por otra parte, en un trabajo reciente por Guo *et al.*, (2003) han sugerido que la presencia de dos o más genes relacionados con la fijación de nitrógeno (*nifH*) podría tener ventajas desde el punto de vista biotecnológico, ya que por medio de la selección de eventos de amplificación se seleccionarían cepas con una mejor capacidad para fijar nitrógeno, y mejorar de manera indirecta, el crecimiento de la planta.

En conclusión, los genomas de rhizobias pueden ser un excelente sistema para estudiar las reiteraciones génicas y su papel en la dinámica genómica, ya que contienen una gran cantidad de genes o familias multigénicas, principalmente aquellas relacionadas con la simbiosis de plantas. Así mismo, conforme se vayan secuenciando más genomas completos de rhizobias, los análisis comparativos mostrarán un panorama más completo de su estructura genómica.

Agradecimientos

Se agradece a la CIC-UMSNH por financiar los diferentes proyectos del Laboratorio de recombinación y diversidad genómica, así como a los revisores por sus sugerencias.

Referencias

- Alani E, Reenan RA, Kolodner RD. 1994. Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. *Genetics*, 137: 19-39.
- Belle A, Landthaler M, Shub DA. 2002. Intronless homing: site-specific endonuclease Seg F of bacteriophage T4 mediates localized

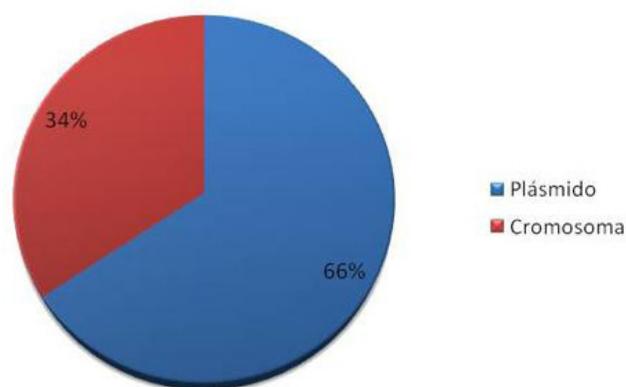


Figura 3. Localización genómica de los genes repetidos en *R. etli*.

marker exclusion analogous to homing endonucleases of group I introns. *Genes & Development*, 16: 351-362.

- Bena G, Lyet A, Huguet T, Olivier I. 2005. Medicago-Sinorhizobium, Symbiotic specificity evolution and the geographic expansion of Medicago. *Journal of Evolutionary Biology*, 18: 1547-1558.
- Brom S, García de los Santos A, Girard ML, Davila G, Palacios R, Romero D. 1991. High-frequency rearrangements in *Rhizobium leguminosarum* bv. *phaseoli* plasmids. *Journal of Bacteriology*, 173: 1344-1346.
- Gonzalez V, Santamaria RI, Bustos P, Hernández-González I, Medrano-Soto A, Moreno-Hagelsieb G, Janga SC, Ramírez MA, Jiménez-Jacinto V, Collado-Vides J, Dávila G. 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proceedings of the National Academy of Science USA*, 103: 3834-3839.
- Guo X, Flores M, Mavingui P, Fuentes SI, Hernández G, Dávila G, Palacios R. 2003. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Research*, 13: 1810-1817.
- Hansen TF, Carter AJ, Chiu CH. 2000. Gene conversion may aid adaptive peak shifts. *Journal of Theoretical Biology*, 207: 495-511.
- Kuzminov A. 2001. DNA replication meets genetic exchange: chromosomal damage and its repair by homologous recombination. *Proceedings of the National Academy of Science USA*, 98: 8461-8468.
- Marians KJ. 2004. Mechanisms of replication fork restart in *Escherichia coli*. *Philos Trans R Soc Lond B Biol Sci*, 29: 71-77.
- Orozco-Mosqueda MC, Altamirano-Hernández J, Farias-Rodríguez R, Valencia-Cantero E, Santoyo G. 2009. Homologous recombination and dynamics of rhizobial genomes. *Research in Microbiology*, 160: 733-741.
- Quinto C, de la Vega H, Flores M, Fernandez L, Ballado T, Soberon G, Palacios R. 1982. Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli*. *Nature*, 299: 724-726.
- Romero D, Martínez-Salazar J, Ortiz E, Rodríguez C, Valencia-Morales E. 1999. Repeated sequences in bacterial chromosomes and plasmids: a glimpse from sequenced genomes. *Research Microbiology*, 150: 735-743.
- Romero D, Martínez-Salazar J, Girard L, Brom S, Dávila G, Palacios R, Flores M, Rodríguez C. 1995. Discrete amplifiable regions (amplicons) in the symbiotic plasmid of *Rhizobium etli* CFN42. *Journal of Bacteriology*, 177: 973-980.

- Romero D, Brom S, Martínez-Salazar J, Girard ML, Palacios R, Dávila G.** 1991. Amplification and deletion of a *nod-nif* region in the symbiotic plasmid of *Rhizobium phaseoli*. *Journal of Bacteriology*, 173: 2435–2441.
- Santoyo G, Martínez-Salazar J, Rodríguez C, Romero D.** 2005. Gene conversion tracts associated with crossovers in *Rhizobium etli*. *Journal of Bacteriology*, 187: 4116-4126.
- Santoyo G, Romero D.** 2005. Gene conversion and concerted evolution in bacterial genomes. *FEMS Microbiology Reviews*, 29: 169-183.
- Szostak J, Orr-Weaver T, Rothstein R, Stahl F.** 1983. The doublestrand-break repair model for recombination. *Cell*, 33: 25-35.
- Takahashi N, Kobayashi I.** 1990. Evidence for the doublestrand break repair model of bacteriophage lambda recombination. *Proceedings of the National Academy of Science USA*, 87: 2790-2794.
- West SC.** 2003. Molecular views of recombination proteins and their control. *Natural Review Molecular Cell Biology*, 4: 435-445.