



Analisis y re-anotacion del genoma de *Sodalis glossinidius*, endosimbionte de la mosca tse-tse

Eugeni Belda, Andrés Moya, y Francisco J. Silva

Institut Cavanilles de Biodiversitat i Biologia Evolutiva and Departament de Genètica, Universitat de València

Sodalis glossinidius es uno de los tres endosimbiontes microbianos de la mosca tse-tse junto con el endosimbionte primario *Wigglesworthia glossinida* y el endosimbionte facultativo *Wolbachia pipientis*. *Sodalis glossinidius* representa un estado intermedio entre las bacterias de vida libre y bacterias mutualistas asociadas de manera estricta a un huésped como *Buchnera aphidicola*; el tamaño de su genoma (4.2 megabases) secuenciado recientemente (1) es cercano al de otras bacterias de vida libre. Sin embargo presenta una capacidad codificadora reducida (51 % del genoma) debido a que presenta una gran cantidad de pseudogenes únicamente comparables con el patógeno de humanos *Mycobacterium leprae* (2,3). Con el objetivo de determinar los cambios que tuvieron lugar en el ancestro de *Sodalis glossinidius* durante su adaptación a la vida intracelular hemos llevado a cabo la re-anotacion y re-analisis de su genoma.

Aplicando un criterio de identificación de pseudogenes basado en el análisis por blastx de las regiones intergenicas del genoma de *Sodalis glossinidius* y la predicción de pautas abiertas de lectura con el programa Genewise hemos incrementado el número de pseudogenes de los 972 descritos inicialmente en la anotación original a mas de 1500 pseudogenes. Además, el estatus de varios genes anotados inicialmente como codificantes ha sido cambiado a pseudogenes debido a que los análisis demuestran que se tratan de fragmentos de genes ancestrales de mayor tamaño. Hemos llevado a cabo también la anotación de Secuencias de Insercion, identificando 5 grupos mayoritarios que representan un 2.5 % del genoma. La mayoría de estos elementos son copias parciales de la secuencia de inserción original o contienen transposasas inactivas.

El análisis composicional del genoma de *Sodalis glossinidius* revela la presencia de varias Islas Genomicas, algunas de ellas con características propias de este tipo de elementos como la presencia de tRNA's adyacentes, repeticiones directas flanqueando la región, así como desviaciones composicionales significativas respecto el promedio del genoma y genes de movilidad e integración (transposasas, integrasas, etc.). Las funciones de los productos de genes y pseudogenes han sido re-analizados siguiendo un protocolo similar al utilizado por Gundogdu et al. (4), y el impacto en el metabolismo de la transición entre el estado de vida libre y el estado mutualista ha sido analizado comparando el conjunto de genes funcionales con el conjunto de genes y pseudogenes.

1. Toh H, Weiss BL, Perkin SAH, Yamashita A, Oshima K, Hattori M, Aksoy S (2006) Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res* 16:149-156.
2. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D et al. (2001) Massive gene decay in the leprosy bacillus. *Nature* 409:1007-1011.
3. Gomez-Valero L, Rocha EP, Latorre A, Silva FJ (2007) Reconstructing the ancestor of *Mycobacterium leprae*: the dynamics of gene loss and genome reduction. *Genome Res* 17:1178-1185.
4. Gundogdu O, Bentley SD, Holden MT, Parkhill J, Dorrell N, Wren BW (2007) Re-annotation and reanalysis of the *Campylobacter jejuni* NCTC11168 genome sequence. *BMC Genomics* 8:162.