ANALYSIS OF FUNCTION AND EVOLUTION OF ORTHOLOGS IN RHIZOBIALES GENOMES
Humberto Peralta, Gabriela Guerrero, Alejandro Aguilar, and Jaime Mora

Program of Functional Genomics of Prokaryotes, Center for Genomic Sciences, National University of Mexico
Av. Universidad, Chamilpa, Cuernavaca, Morelos, Mexico. CP 62210  E-mail: peralta@cgg.unam.mx

Background

The Rhizobiales is a versatile alpha proteobacterial order. They show genomic diversity, several lifestyles and ability to engage in symbiotic or pathogenic relationships. The common orthologs in the Rhizobiales genomes represent an ideal gene set to detect their evolutive tendencies and to assess the importance of chromosomal gene order conservation (or synteny). Also, we consider that some differences in the proteins respond to adaptation to specific niches or conditions; we propose a new parameter, the species signature, to detect these changes. We explored the relations of function and orthology, synteny, codon bias, evolution, and the physicochemical properties of the amino acids encoded in the common orthologs of five chromosomes of Rhizobiales, namely: Sinorhizobium meliloti, Rhizobium etli and Mesorhizobium loti (plant symbionts), Agrobacterium tumefaciens (a tumor-inducing plant pathogen), and Brucella melitensis (an intracellular mammal pathogen).

Results. Relationships of orthology and synteny

Fig. 1 shows the chromosome comparison of S. meliloti and A. tumefaciens. The red dots denote syntenic orthologs, and the blue ones nonsyntenic orthologs. At right, in red lines, microsyntenic regions.

Fig. 2 shows the strategy to obtain common orthologs; red syntenic genes; light and dark blue, genes without or partial synteny. Functional differences of the common orthologs are depicted in Fig. 3 left, and at the right those for nonsyntenic orthologs and nonorthologs. The syntenic products showed increased sequence conservation (Fig. 4A), and open organization (4B) and more reduced synonymous and nonsynonymous substitution rates (4D). Two main phylogenetic relationships were related to synteny (4C).

The species signature and the evolutive rates

The correlations between the species signature and nonsynonymous (Fig. 6A) and synonymous substitution rates (Fig 6B) were very high, and better than identity or similarity (not shown), demonstrating the usefulness of our parameter to extract evolutive information at protein level.

Physicochemical properties of amino acids

Figs. 7A and 7B show the best pair relationships of amino acids encoded in the common orthologs, as assessed with the Atkinley matrices (PNAS 102:893): the usefulness of these matrices is revealed in Fig. 7C, with the products showing a different tendency (in red) in many transmembranal segments; a clustering analysis revealed the overall matrices relationships, and products with similar profile (Fig. 8, P S A V C denote Polarity, Secondary structure, Amino acid composition, Volume, and Charge matrices, respectively). Clear functional differences were found with the sign obtained with the matrices for the products (Fig. 9), with Polarity as example: informational functions with positive sign, and metabolism with negative one.

Relationships of nonsynonymous substitution

In order to detect functional differences regarding evolutive rate, the common orthologs were divided in low or high nonsynonymous substitution (dN) groups; low dN genes showed abundance in metabolism and informational functions; in contrast, high dN genes were abundant in cellular processes and hypothetical functions (Fig. 10).

Conclusions

The chromosomal arrangement is related to evolutive trends and also has a functional significance. The species signature is an useful parameter to extract evolutive information from amino acid sequences. The nonsynonymous substitution is related to the codon bias (MGC3, CAI) in the Rhizobiales; and strikingly, to physicochemical properties of the amino acid residues.

We thank help by Julio Martinez, and support from DGAPA/UNAM IN215307 grant.