

Genome-wide Gene Order Distances Support a United Gram-Positive Bacteria. C. H. House¹ and S. T. Fitz-Gibbon², ¹Penn State Astrobiology Research Center, 218 Deike Building, University Park, PA. chrishouse@psu.edu
² IGPP Center for Astrobiology, University of California Los Angeles, CA 90095-1567.

During the past decade, considerable effort has been placed on comparing the similarity of organisms with genome-wide methods or, at least, with methods that use more than a single gene [1]. Ever since researchers first identified that gene order information was conserved between humans and mice [2], there has been growing interest in using gene order to estimate the difference between genomes or to solve phylogenetic problems [3-4]. Here, we have attempted to use a monte carlo approach to look for small genome-wide instances of gene order conservation.

The pair-wise distances between 143 prokaryotic genomes were determined using a monte carlo method to estimate the conservation of gene order. The method was based on repeatedly selecting six non-adjacent random orthologs from each of two genomes and determining if the chosen orthologs were in the same order. The raw distances were then corrected for gene order convergence using an adaptation of the Jukes-Cantor model. We have compared the distances found to those of ortholog gene content and small subunit rRNA. The gene order distances are reasonably well correlated with the divergence of rRNA ($R^2 = 0.24$), especially at rRNA Jukes-Cantor distances of less than 0.2 ($R^2 = 0.52$). Overall, gene content is only weakly correlated with rRNA divergence ($R^2 = 0.04$), however, it is especially strongly correlated at rRNA Jukes-Cantor distances of less than 0.1 ($R^2 = 0.67$). Using the gene order distances, the relations of prokaryotes were studied using neighbor joining, agreement subtrees, and a new proposed hierarchical tree building strategy that relies on the fact that shorter distances are known with higher confidence than longer distances.

Consistently, our trees show the Actinobacteria as a sister group to the bulk of the Firmicutes. In fact, the level of gene order support was found to be considerably greater for uniting these two phyla than for uniting any of the proteobacterial classes together. The results are supportive of the idea that the Actinobacteria and Firmicutes are a monophyletic group, which in turn implies a single origin for the gram-positive cell.

References:

- [1] House, C.H. (2009) *Methods in Molecular Biology*, 532: 141-161.
- [2] Nadeau H. and Taylor B. (1984) *Proc Natl Acad Sci U.S.A.* 81, 814-818.
- [3] Sankoff, D., Leduc, G., Antoine, N., Paquin, B., Lang, B.F., Cedergren, R. (1992) *Proc Natl Acad Sci. U.S.A.* 89, 6575-6579.

- [4] Kunisawa T. (2003) *J Theor Biol.* 222, 495-503.