

Poster C-38

Comparison Between Complete Genomes of Vibrio Species



Authors:

Patrícia P. Côgo (*Institute of Computing, University of Campinas*)

João Meidanis (*Institute of Computing, University of Campinas*)

Fabiano Thompson (*Department of Genetics, Institute of Biology, Federal University of Rio de Janeiro*)

Short Abstract: This work proposes a methodology based on the “Genomic Rearrangement Theory” to compare complete genomes of vibrio species. In other words, our focus are large-scale mutations involving entire genes or blocks of consecutive genes, instead of events that affect individual bases.

Long Abstract:

Genome sequencing has made possible the development of new comparative genomic methods, and hence, a revolution in microbial taxonomy. A new, modern microbial taxonomy is being established based on multiple loci or entire genomes [3].

This work proposes a methodology based on the “Genomic Rearrangement Theory” [1] to compare complete genomes of vibrio species. In other words, our focus are large-scale mutations involving entire genes or blocks of consecutive genes, instead of events that affect individual bases.

The methodology is strongly based on the same steps traditionally followed when measuring evolution in Bioinformatics: define the model, identify homolog structures, and employ the model and these structures to find the most parsimonious sequence of events that can explain the differences between the genomes.

Our model is based on the occurrence of four fundamental operations: Fusions, Fissions, Reversals, and Gene Loss (modeled as a special case of fission). This model is similar to the one described by Yancopoulos et al. [4], the major differences being our consideration of gene loss and the way we treat translocations and block interchanges, which are not fundamental events, but rather a result of two successive operations of fission and fusion.

Homolog structures will be identified through profiles, i.e, weighted matrices that describe families of proteins, as in the PROSITE project [2], but with many more profiles, so that we can cover all vibrio proteins.

When two organisms are compared, families with multiple members in one or both genomes will be spliced, guided by a phylogenetic analysis, in order to resolve duplicated genes. After that, the distance between these species, i.e, the minimum number of fundamental operations that

transform a genome into the other, is calculated. Phylogenetic trees are then built from such distances.

Initial experiments have built trees equivalent to the ones obtained from traditional gene markers.

We expect to gain from this study a deeper understanding of the evolution of vibrios and help improve the taxonomy of these organisms.

References

[1] J. Meidanis and Z. Dias. An alternative algebraic formalism for genome rearrangements. *Comparative Genomics Empirical and Analytical Approaches to Gene Order Dynamics Comparative Maps Multigenes Families*, 2000.

[2] C.J.A. Sigrist, L. Cerutti, N. Hulo, A. Gattiker, L. Falquet, M. Pagni, A. Bairoch, and P. Bucher.

PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinformatics*, 3:265–274, 2002.

[3] F.L. Thompson, T. Iida, and J. Swings. Biodiversity of vibrios. *Microbiology and Molecular Biology Reviews*, pages 403–431, 2004.

[4] S. Yancopoulos, O. Attie, and R. Friedberg. Efficient sorting genomic permutations by translocation, inversion and block interchange. *Bioinformatics*, 21(16):3340–3346, 2005.