Horizontal Gene Transfer Phylogenetics: A Model-Based Approach

Dear Editor,

This manuscript provides the first statistical modeling for the Synteny Index (SI) concept, defined in previous works of ours, measuring similarity between neighborhoods of orthologous genes. The uniqueness and novelty of SI, are the combination of both presence/absence and gene order differences between genomes. Hence, it overcomes the severe limitations associated with more dedicated approaches, focusing on specific mechanisms.

While it was shown that SI provides better resolution of the standard phylogenetic marker, the 16S ribosomal gene, in the deep branches of the tree of life, there was no analytical modeling of the processes affecting SI, a gold standard in modern phylogenetics. Here we tackle this task by focusing on horizontal gene transfer (HGT), modeling it as a Markovian process operating on a genome. Such ש modelling is valuable as it allows observed measures to be transformed into estimates of time periods during evolution, yielding additivity of the measure. To the best of our knowledge, there is no other additivity proof for other gene order/content measures under HGT.

The models we develop are divided into two: an analytical asymptotic model under which we can show additivity of the SI measure, and an empirical model considering real life values with which we subsequently analyze genomic data.

**Developing accurate measures to delineate HGT processes is important for the role they play in various fields of microbiology.** The new modeling developed allows us to explore such properties that are beyond sheer classification.

Prof. Sagi Snir, Ph.D.

The Department of Evolutionary and Environmental Biology

University of Haifa,

Israel

