Ribotaxa: Combined Approaches for Taxonomic Resolution Down to the Species Level from Metagenomics Data Revealing Novelties

Authors : Oshma Chakoory, Sophie Comtet-Marre, Pierre Peyret

Abstract : Metagenomic classifiers are widely used for the taxonomic profiling of metagenomic data and estimation of taxa relative abundance. Small subunit rRNA genes are nowadays a gold standard for the phylogenetic resolution of complex microbial communities, although the power of this marker comes down to its use as full-length. We benchmarked the performance and accuracy of rRNA-specialized versus general-purpose read mappers, reference-targeted assemblers and taxonomic classifiers. We then built a pipeline called RiboTaxa to generate a highly sensitive and specific metataxonomic approach. Using metagenomics data, RiboTaxa gave the best results compared to other tools (Kraken2, Centrifuge (1), METAXA2 (2), PhyloFlash (3)) with precise taxonomic identification and relative abundance description, giving no false positive detection. Using real datasets from various environments (ocean, soil, human gut) and from different approaches (metagenomics and gene capture by hybridization), RiboTaxa revealed microbial novelties not seen by current bioinformatics analysis opening new biological perspectives in human and environmental health. In a study focused on corals' health involving 20 metagenomic samples (4), an affiliation of prokaryotes was limited to the family level with Endozoicomonadaceae characterising healthy octocoral tissue. RiboTaxa highlighted 2 species of uncultured Endozoicomonas which were dominant in the healthy tissue. Both species belonged to a genus not yet described, opening new research perspectives on corals' health. Applied to metagenomics data from a study on human gut and extreme longevity (5), RiboTaxa detected the presence of an uncultured archaeon in semi-supercentenarians (aged 105 to 109 years) highlighting an archaeal genus, not yet described, and 3 uncultured species belonging to the Enorma genus that could be species of interest participating in the longevity process. RiboTaxa is user-friendly, rapid, allowing microbiota structure description from any environment and the results can be easily interpreted. This software is freely available at https://github.com/oschakoory/RiboTaxa under the GNU Affero General Public License 3.0.

Keywords : metagenomics profiling, microbial diversity, SSU rRNA genes, full-length phylogenetic marker

Conference Title : ICBAHM 2022 : International Conference on Bioinformatic Approaches for Human Microbiome

Conference Location : Istanbul, Türkiye

Conference Dates : October 20-21, 2022