The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level

Luis M. Rodriguez-R, Santosh Gunturu, William T. Harvey, Ramon Rosselló-Mora, James M. Tiedje, James R. Cole and Konstantinos T. Konstantinidis

Nucleic Acids Research, 2018 1 doi: 10.1093/nar/gky467

www.microbial-genomes.org
“The genome-equivalent of RDP”
Key underlying data is AAI values among Type Genomes*

*Or another genome collection. Type Genomes is based on NCBI’s Type Material database and includes named isolates & some Candidatus.

So, it is very easy to incorporate new Taxonomies

% of genome pairs
in a taxonomic rank

different phyla

Same phylum

Same genus

same species

AAI: 40 50 60 70 80 90 100%

★ Query is novel genus ★ Query is member of existing species

Protologues with key metadata will be highly useful!

Novel taxa can be identified at species, genus & phylum levels

- Novel species <95% AAI
- Novel genus <65% AAI
- Novel phylum <45% AAI
MAGs and SAGs in MiGA

• Not yet available as a single large searchable database (this is coming soon...~20% more MAGs/SAGs than found in NCBI based on manual hunting)

• Many MAGs/SAGs available as separate databases by habitat (e.g., TARA Oceans, Oil spills) or project (e.g., Chattahoochee River, Bio-Gas microbiome).

• Taxonomy of MAGs/SAGs is (currently) based on AAI values against Type Genomes. Taxonomy of Type Genomes is based on NCBI (no corrections applied).

• MAG/SAG quality is check by completeness and contamination using universal protein HMMs, and (optionally) by MyTaxa_Scan (Luo et al., NAR 2014).

• Low quality MAGs flagged by not removed.