

GENOME TAXONOMY DATABASE



gtdb.ecogenomic.org

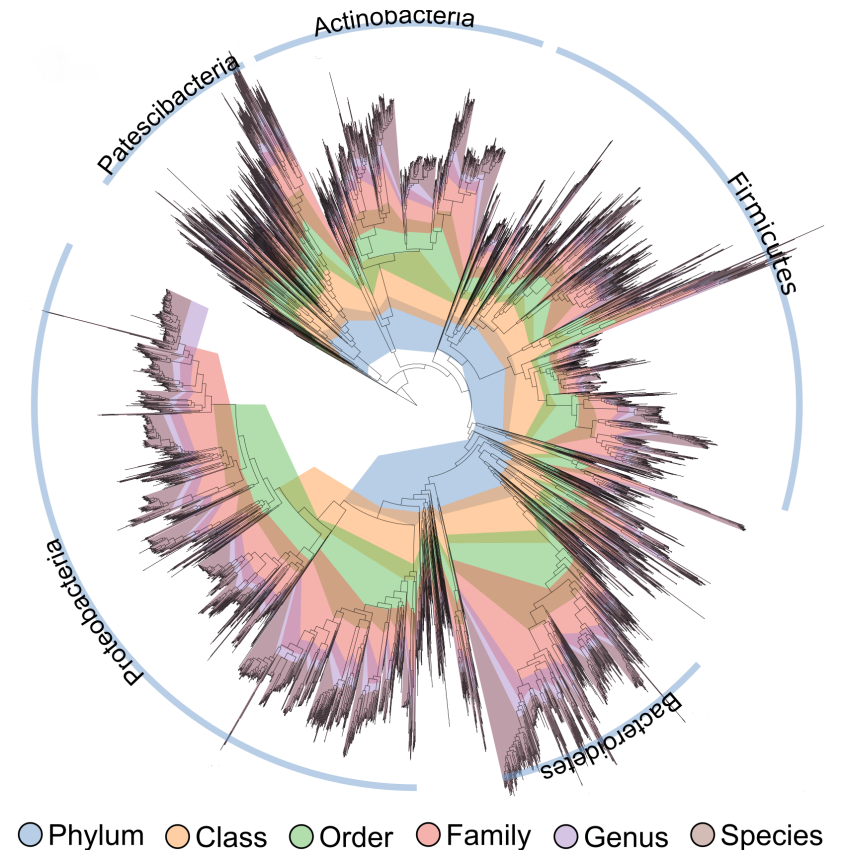
GTDB is an initiative to establish a complete, rank normalized, and phylogenetically consistent genome-based taxonomy for Bacteria and Archaea

GTDB methodology in brief:

- species delineated using average nucleotide identity (ANI)
- higher taxa circumscribe lineages in reference phylogeny
- higher taxa normalized to have similar divergence times
- all genomes have species to domain assignments
- “uncultivated” majority given placeholder names

GTDB data dependencies:

- covers all genomes deposited in the **NCBI Assembly Database**
 - 194,600 genomes as of GTDB R05-RS95
- type material established by consulting **LPSN**
- **SILVA** used to classify 16S rRNA sequences



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

GTDB MEETS SEQCODE



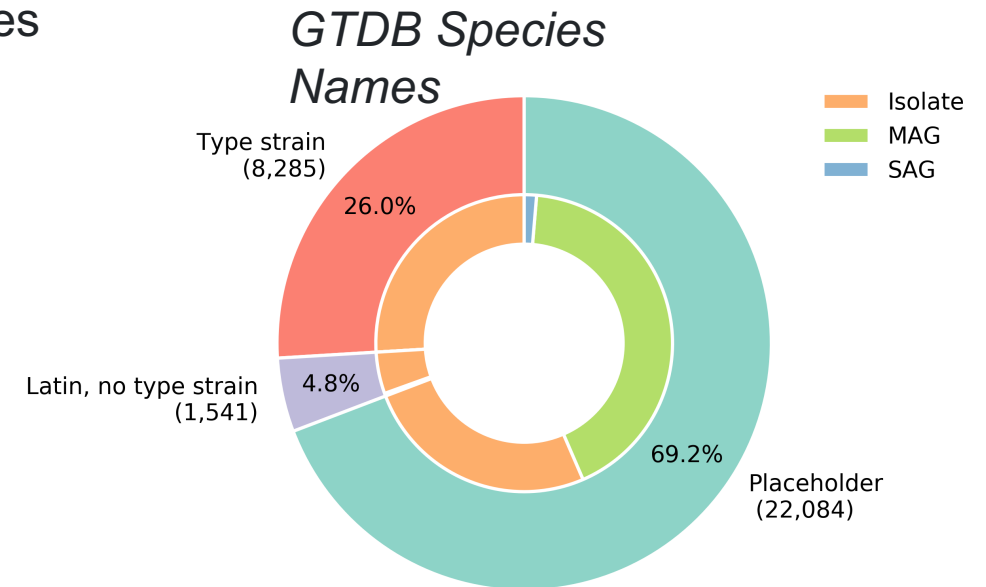
gtdb.ecogenomic.org

Potential advantages:

- type material will be genomes, the entity used by many biologists and bioinformatic resources and tools
- provide stable and less burdensome avenue to naming species
 - ~70% of GTDB species have placeholder names
 - validated species allow for validated higher taxon names
- provide nomenclatural information in format suitable for easy database integration (*i.e.* machine friendly)
 - LPSN must be manually consulted though efforts are underway to make this data more accessible

Potential challenges:

- requires consulting multiple nomenclatural resources to establish priority and type material
 - unless SeqCode resources cover names validated under the ICNP
- SeqCode resources will diverge from ICNP resources on application of nomenclature in addition to taxonomic opinion



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA