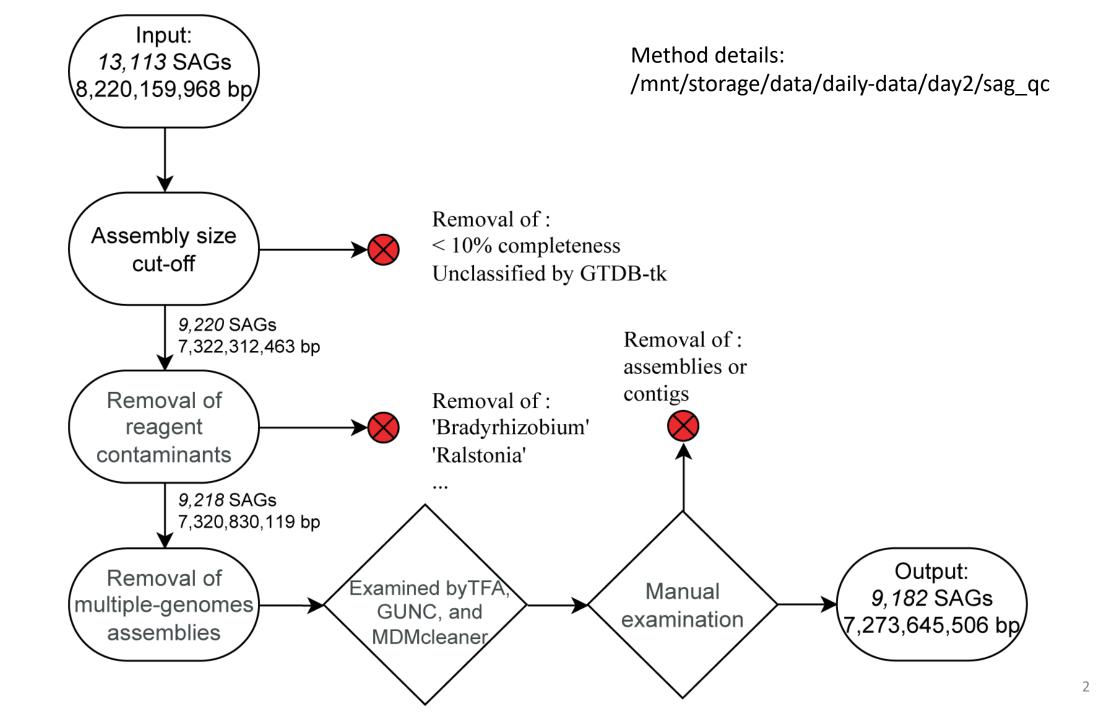
GORG-Dark QC



- **Assembly size cut-off** (CheckM1-estimated completion >=10% + GTDB-tk r202 classified)
- Removal of reagent contaminants (e.g., Bradyrhizobium, BLASTN—nt)
- Assemblies containing sequences from multiple marine prokaryoplankton genomes
 - removal of the entire assembly
 - tetramer PCA, GUNC, MDMcleaner, CheckM + BLASTN—nr
 - removal of individual contigs
 - tetramer PCA outliers + MDMcleaner + BLASTN—nr +
 - keep virus-like: DeepVirFinder2, VIBRANT, Virsorter2 + BLASTN—nr