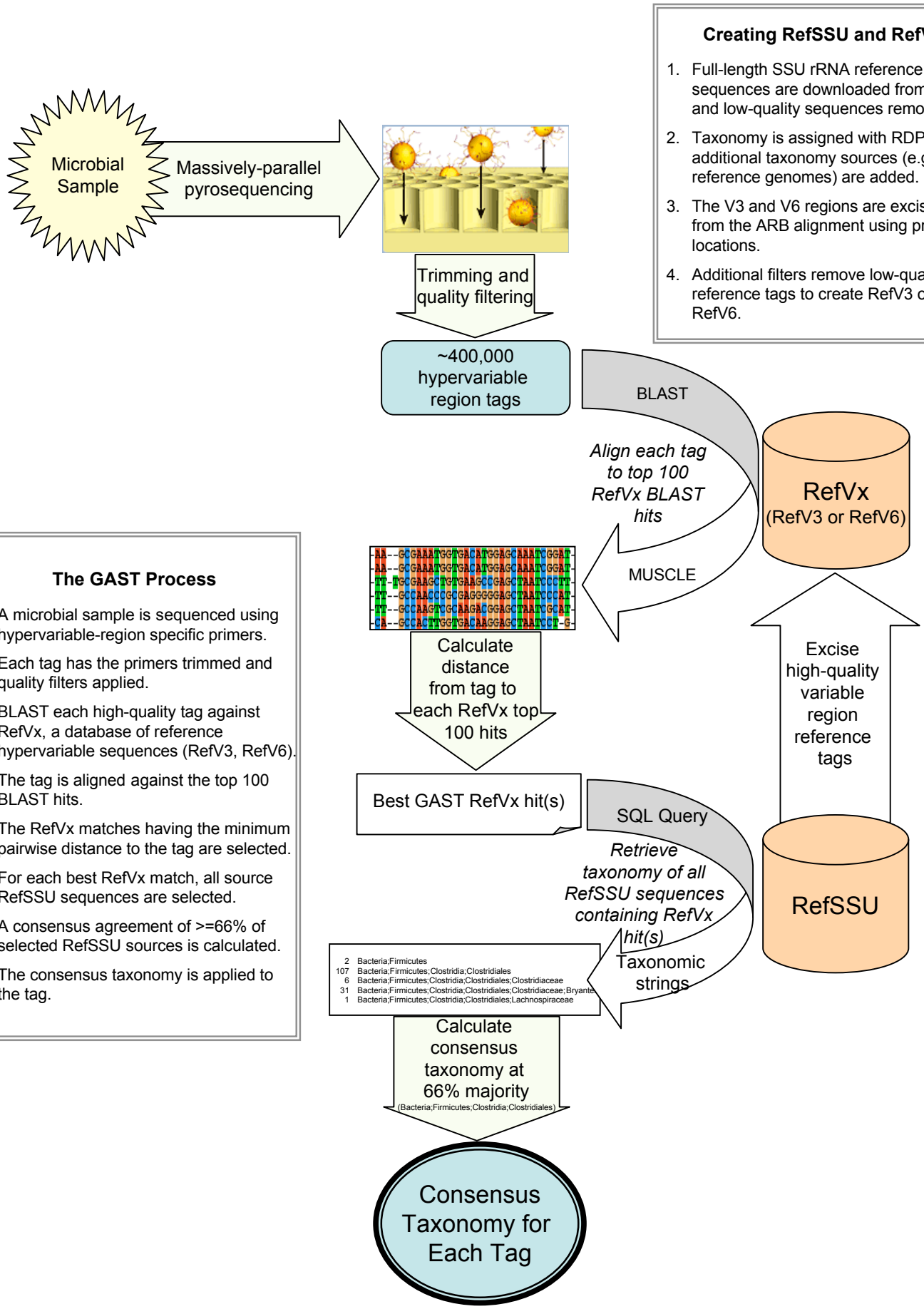
# Global Alignment for Sequence Taxonomy (GAST)




### Creating RefSSU and RefVx

1. Full-length SSU rRNA reference sequences are downloaded from SILVA and low-quality sequences removed.
2. Taxonomy is assigned with RDP, additional taxonomy sources (e.g. reference genomes) are added.
3. The V3 and V6 regions are excised from the ARB alignment using primer locations.
4. Additional filters remove low-quality reference tags to create RefV3 or RefV6.

### The GAST Process

1. A microbial sample is sequenced using hypervariable-region specific primers.
2. Each tag has the primers trimmed and quality filters applied.
3. BLAST each high-quality tag against RefVx, a database of reference hypervariable sequences (RefV3, RefV6).
4. The tag is aligned against the top 100 BLAST hits.
5. The RefVx matches having the minimum pairwise distance to the tag are selected.
6. For each best RefVx match, all source RefSSU sequences are selected.
7. A consensus agreement of >=66% of selected RefSSU sources is calculated.
8. The consensus taxonomy is applied to the tag.