# Genome annotation

## 

We employed the GALBA pipeline for the annotation of gene structures. Specifically, we used protein files from *Saccharomyces cerevisiae*, *Saccharomyces uvarum*, and *Saccharomyces eubayanus* as inputs. These proteins were aligned to the genome using the miniprot {citation} tool, followed by the application of AUGUSTUS {citation} for the structural annotation of the genome. The output GTF files were processed using AGAT for format cleaning and conversion. The quality of the annotation was evaluated using the BUSCO (version: 5.5.0), employing the saccharomycetes\_odb10 database for assessment. For functional annotation, we utilized the web version of eggnog-mapper (http://eggnog-mapper.embl.de/) to upload the S. bayanus protein files. All other parameters were retained as default settings.

Protein BUSCO score:

C:98.0%[S:11.8%,D:86.2%],F:1.4%,M:0.6%,n:2137

2095 Complete BUSCOs (C)

252 Complete and single-copy BUSCOs (S)

1843 Complete and duplicated BUSCOs (D)

29 Fragmented BUSCOs (F)

13 Missing BUSCOs (M)

2137 Total BUSCO groups searched

CDS:

C:98.0%[S:11.6%,D:86.4%],F:1.2%,M:0.8%,n:2137

2094 Complete BUSCOs (C)

247 Complete and single-copy BUSCOs (S)

1847 Complete and duplicated BUSCOs (D)

25 Fragmented BUSCOs (F)

18 Missing BUSCOs (M)

2137 Total BUSCO groups searched