

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	1890680
ToLID	gfCryBron
Species	Cryptotrichosporon brontae
Class	Tremellomycetes
Order	Trichosporonales

Genome Traits	Expected	Observed
Haploid size (bp)	12,798,796	18,902,778
Haploid Number	4 (source: ancestor)	16
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	NA

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 5.6.Q61

The following metrics were automatically flagged as below EBP recommended standards or different from expected:

- . Observed Haploid size (bp) has >20% difference with Expected
- . Observed Haploid Number is different from Expected
- . Assembly length loss > 3% for pri
- . More than 1000 gaps/Gbp for pri

Curator notes

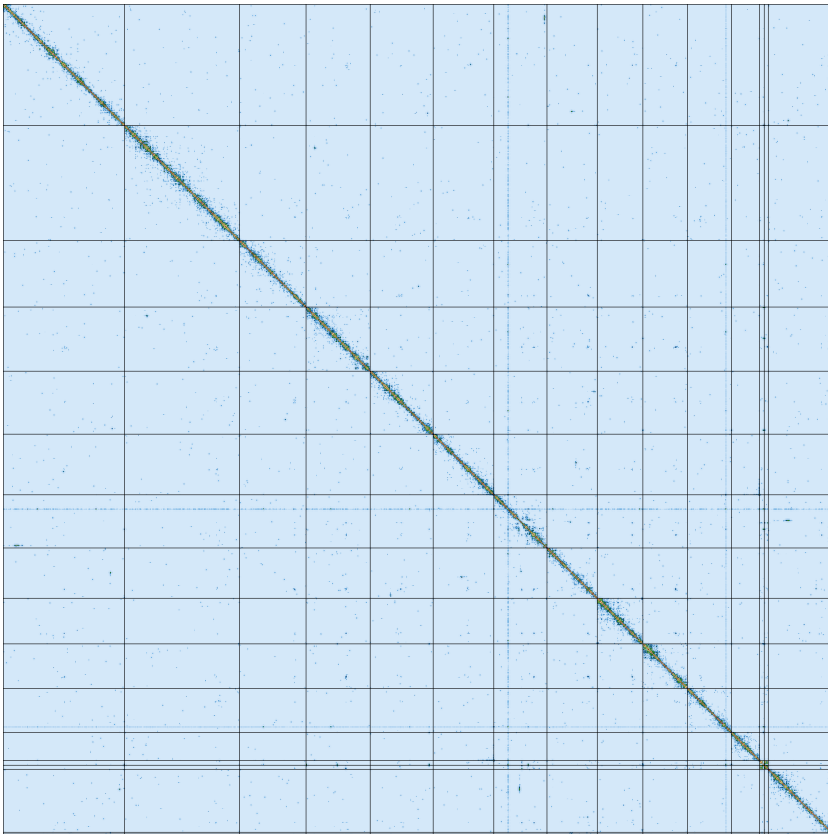
- . Interventions/Gb: None
- . Contamination notes: "No presence of contaminants."
- . Other observations: "Severely downsampled for Hifiasm. I've ran Hifiasm multiple times with no difference in results. Haploid number is close to the expected (12-14). Please note it's a very small genome when looking at the Hi-C contact map."

Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	21,861,153	18,902,778
GC %	65.06	68.36
Gaps/Gbp	0	6,877.3
Total gap bp	0	26,000
Scaffolds	285	24
Scaffold N50	177,658	1,433,913
Scaffold L50	39	5
Scaffold L90	167	12
Contigs	285	154
Contig N50	177,658	193,145
Contig L50	39	32
Contig L90	167	97
QV	61.0153	61.0
Kmer compl.		99.43
BUSCO sing.	96.1%	96.0%
BUSCO dupl.	0.1%	0.1%
BUSCO frag.	0.7%	0.9%
BUSCO miss.	3.2%	3.1%

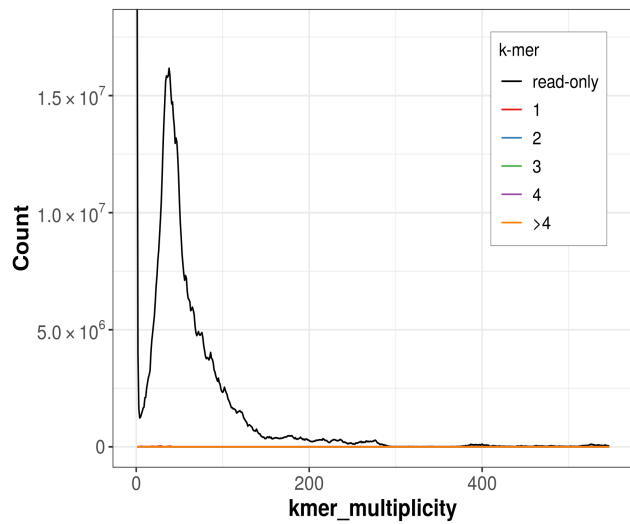
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: basidiomycota_odb10 (genomes:133, BUSCOs:1764)

HiC contact map of curated assembly

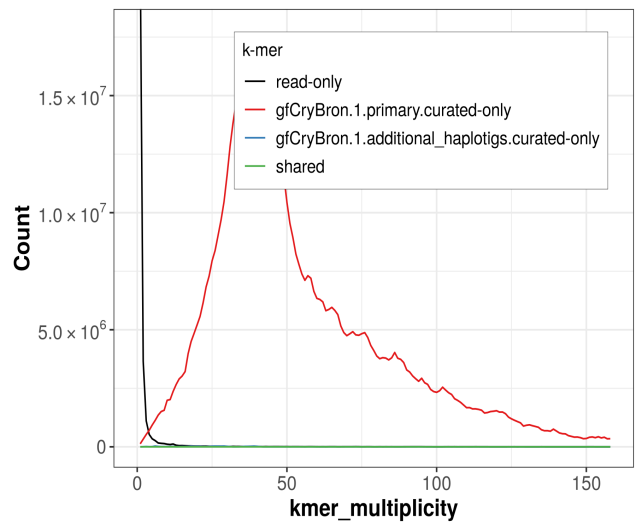


pri [\[LINK\]](#)

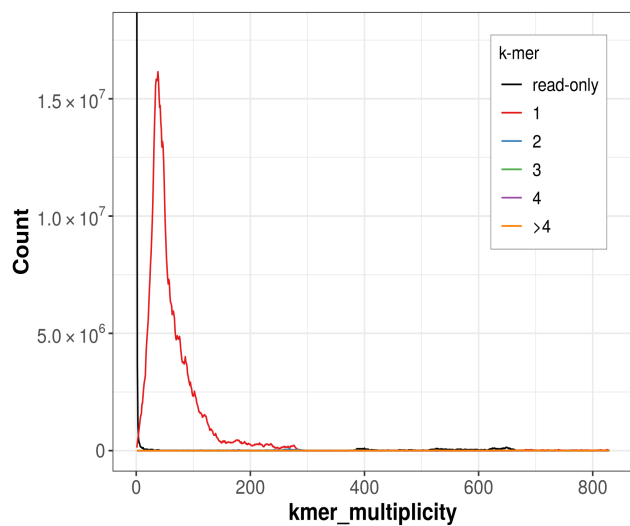
K-mer spectra of curated assembly



Distribution of k-mer counts per copy numbers found in asm

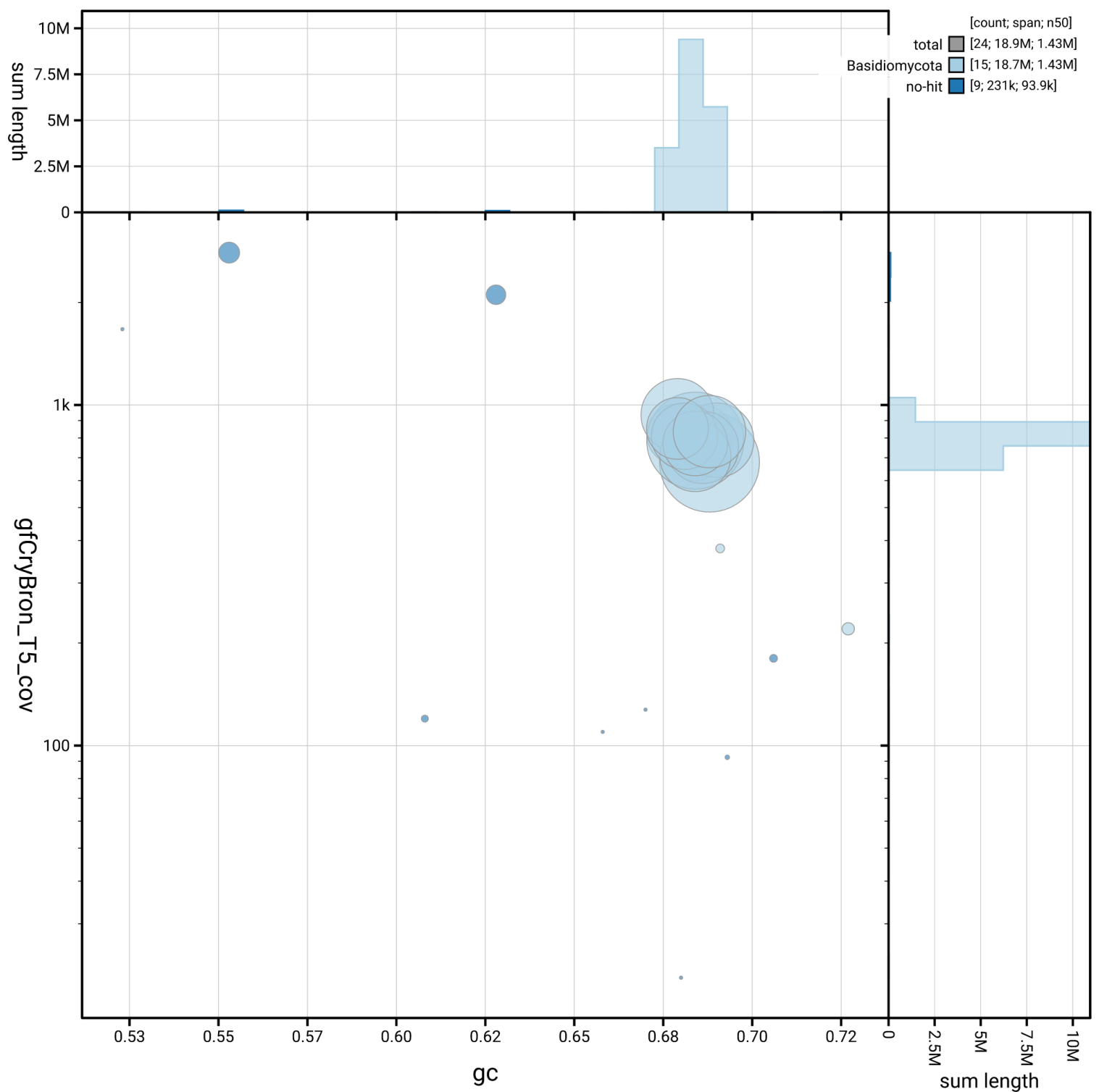


Distribution of k-mer counts coloured by their presence in reads/assemblies



Distribution of k-mer counts per copy numbers found in asm

Post-curation contamination screening



pri. Bubble plot circles are scaled by sequence length, positioned by coverage and GC proportion, and coloured by taxonomy. Histograms show total assembly length distribution on each axis.

Data profile

Data	HiFi	Bionano	OmniC
Coverage	36x	NA	180x

Assembly pipeline

- **Hifiasm**
 - |_ *ver*: 0.20.0-r639
 - |_ *key param*: NA
- **purge_dups**
 - |_ *ver*: 1.2.5
 - |_ *key param*: NA
- **YaHS**
 - |_ *ver*: 1.2.2
 - |_ *key param*: NA

Curation pipeline

- **sanger-tol/curationpretext**
 - |_ *ver*: v1.5.0-g5fd8e6e
 - |_ *key param*: NA
- **Nextflow**
 - |_ *ver*: 25.04.2
 - |_ *key param*: NA
- **pretextmap**
 - |_ *ver*: 0.1.9
 - |_ *key param*: NA
- **HiGlass**
 - |_ *ver*: 0.10.4
 - |_ *key param*: NA

Submitter: André E. R. Soares

Affiliation: SciLifeLab

Date and time: 2025-10-29 16:42:38 CET