

ERGA Assembly Report

v24.10.15

Tags: ERGA-Pilot

| | |
|---------|----------------------|
| TxID | 30301 |
| ToLID | kaBotSchl1 |
| Species | Botryllus schlosseri |
| Class | Ascidacea |
| Order | Stolidobranchia |

| Genome Traits | Expected | Observed |
|-------------------|-----------------------|-------------|
| Haploid size (bp) | 536,801,721 | 555,430,406 |
| Haploid Number | 16 (source: ancestor) | 17 |
| Ploidy | 2 (source: ancestor) | 2 |
| Sample Sex | unknown | unknown |

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q42

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

- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for pri
- . BUSCO single copy value is less than 90% for pri
- . Assembly length loss > 3% for pri

Curator notes

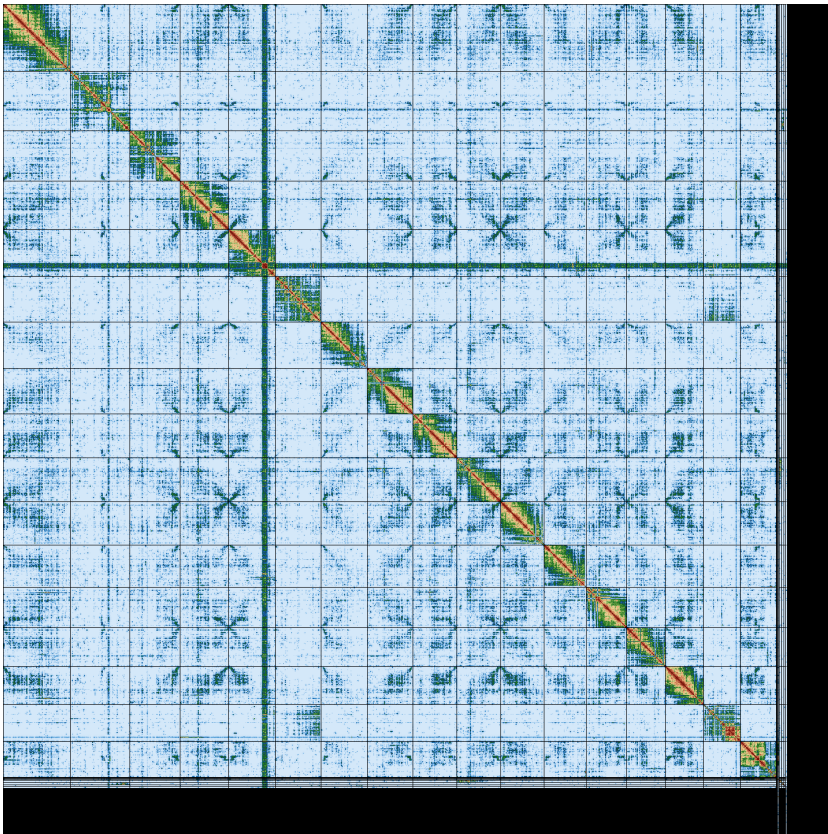
- . Interventions/Gb: 152
- . Contamination notes: "No presence of contaminants."
- . Other observations: "None"

Quality metrics table

| Metrics | Pre-curation pri | Curated pri |
|--------------|---------------------|----------------|
| Total bp | 669,506,172 | 555,430,406 |
| GC % | 40.46 | 40.46 |
| Gaps/Gbp | 648.24 | 898.4 |
| Total gap bp | 86,800 | 59,000 |
| Scaffolds | 560 | 432 |
| Scaffold N50 | 29,857,706 | 29,281,996 |
| Scaffold L50 | 10 | 9 |
| Scaffold L90 | 49 | 17 |
| Contigs | 994 | 931 |
| Contig N50 | 1,847,095 | 1,837,085 |
| Contig L50 | 112 | 95 |
| Contig L90 | 373 | 327 |
| QV | 42.6514 | 42.0554 |
| Kmer compl. | 75.4032 | 65.5143 |
| BUSCO sing. | 84.5% | 89.1% |
| BUSCO dupl. | 7.8% | 1.5% |
| BUSCO frag. | 1.5% | 1.5% |
| BUSCO miss. | 6.3% | 8.0% |

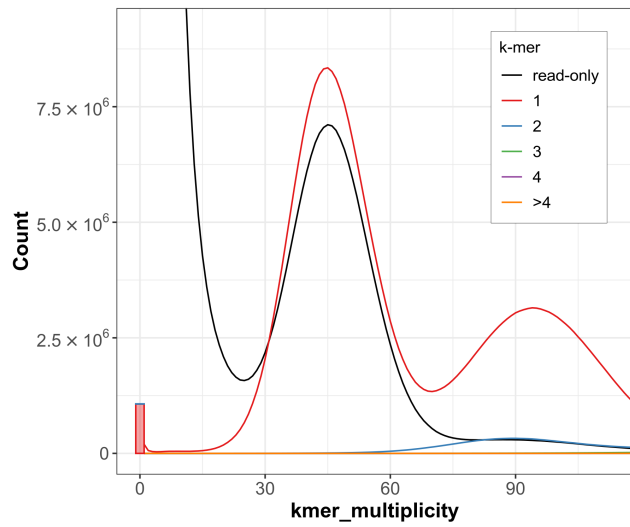
BUSCO: 5.8.2 (euk_genome_aug, augustus) / Lineage: metazoa_odb10 (genomes:65, BUSCOs:954)

HiC contact map of curated assembly

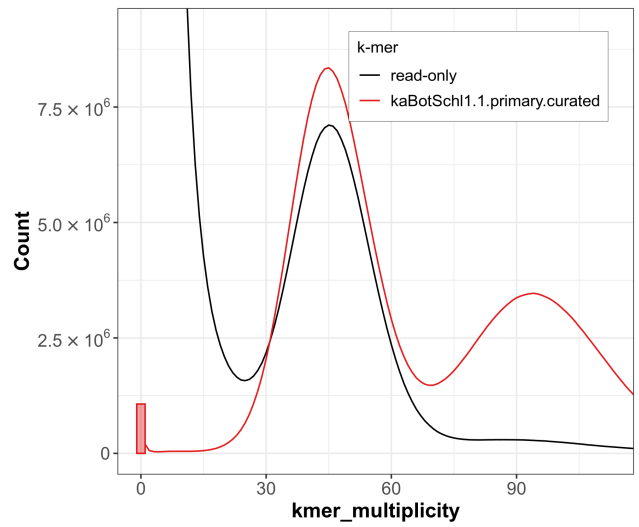


pri [\[LINK\]](#)

K-mer spectra of curated assembly



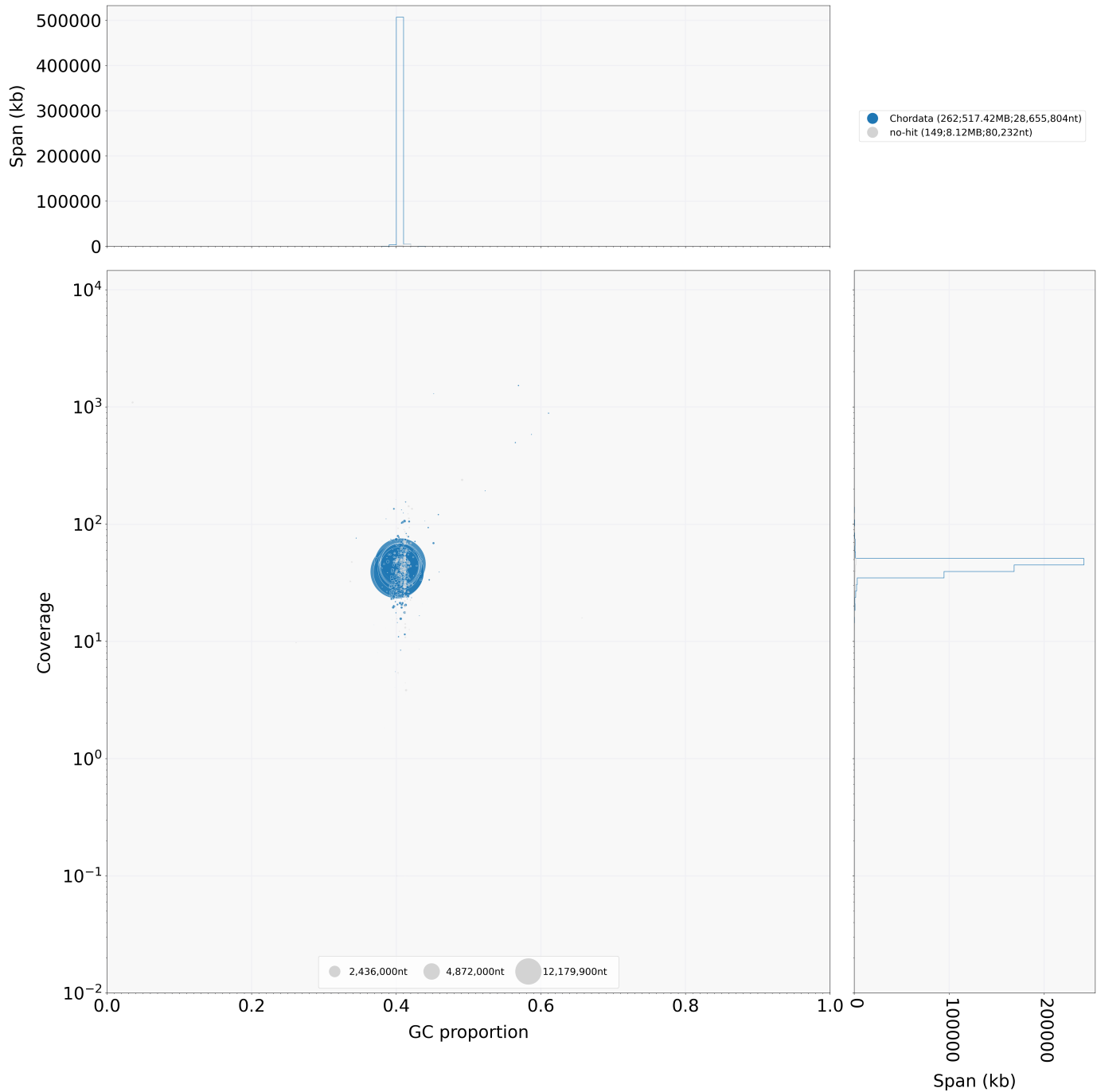
Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening

filename.blobDB.json.bestsum.phylum.p8.span.100.blobplot.bam0



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 | ONT | Bionano | OmniC |
|----------|------|---------|-------|
| Coverage | 135x | NA | 141x |

Assembly pipeline

- **Necat**
 - |_ *ver*: 0.0.1
 - |_ *key param*: NA
- **purge_dups**
 - |_ *ver*: 1.2.5
 - |_ *key param*: NA
- **YaHS**
 - |_ *ver*: 1.1
 - |_ *key param*: NA

Curation pipeline

- **sanger-tol/curationpretext**
 - |_ *ver*: 1.4.1
 - |_ *key param*: NA
- **PretextView**
 - |_ *ver*: 1.0.3
 - |_ *key param*: NA

Submitter: Carmela Gissi

Affiliation: UNIBA

Date and time: 2025-09-23 10:35:47 CEST