

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	55448
ToLID	1aCymNodo2
Species	Cymodocea nodosa
Class	Magnoliopsida
Order	Alismatales

Genome Traits	Expected	Observed
Haploid size (bp)	187,263,864	359,377,510
Haploid Number	7 (source: direct)	18
Ploidy	4 (source: direct)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q52

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
- . Observed Ploidy is different from Expected

Curator notes

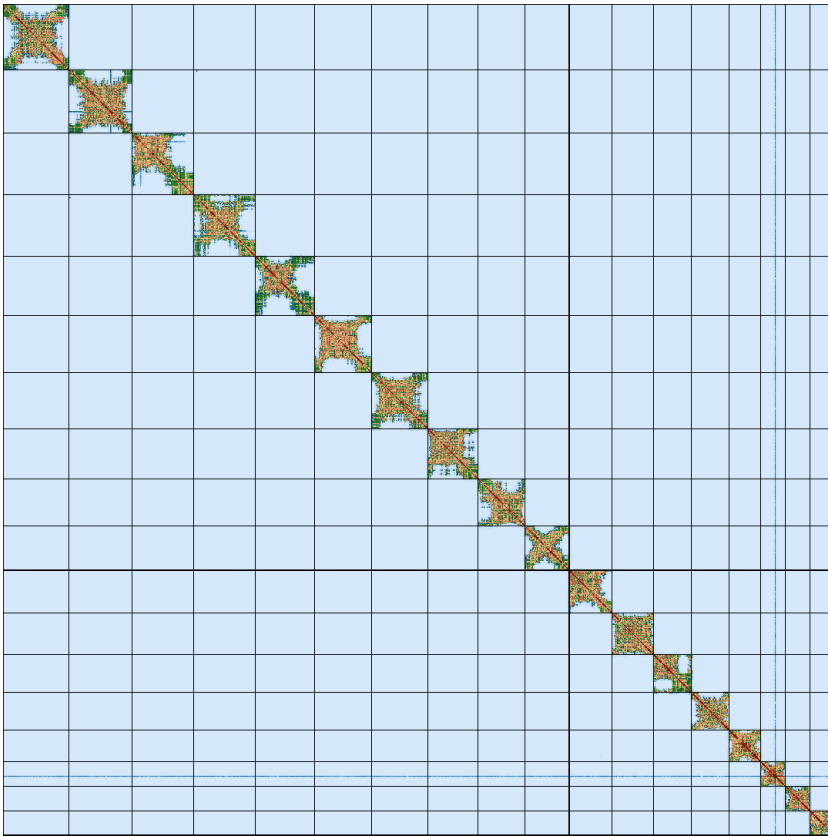
- . Interventions/Gb: 143
- . Contamination notes: "A low number of bacterial contigs were removed post-assembly"
- . Other observations: "23 haplotigs were removed "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	360,243,928	359,377,510
GC %	33.22	33.23
Gaps/Gbp	113.81	133.56
Total gap bp	4,100	7,100
Scaffolds	31	23
Scaffold N50	20,283,553	24,187,662
Scaffold L50	7	7
Scaffold L90	15	15
Contigs	72	71
Contig N50	9,103,177	9,097,069
Contig L50	14	14
Contig L90	34	34
QV	52.5339	52.5849
Kmer compl.	92.1191	92.0497
BUSCO sing.	95.3%	95.7%
BUSCO dupl.	3.9%	3.5%
BUSCO frag.	0.0%	0.0%
BUSCO miss.	0.8%	0.8%

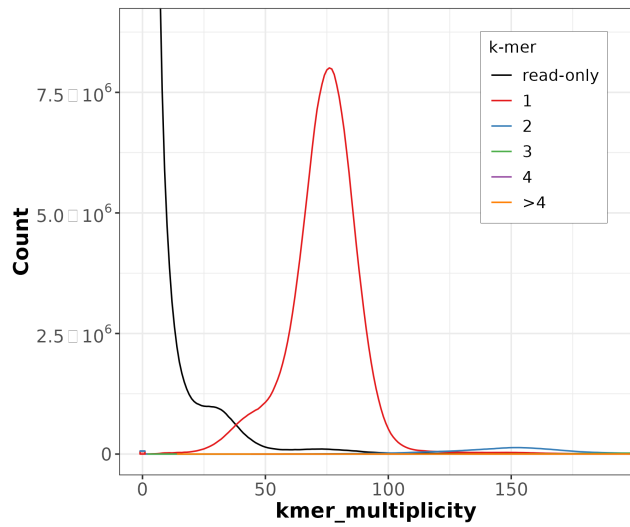
BUSCO: 5.4.3 (euk_genome_met, metaeuk) / Lineage: eukaryota_odb10 (genomes:70, BUSCOs:255)

HiC contact map of curated assembly

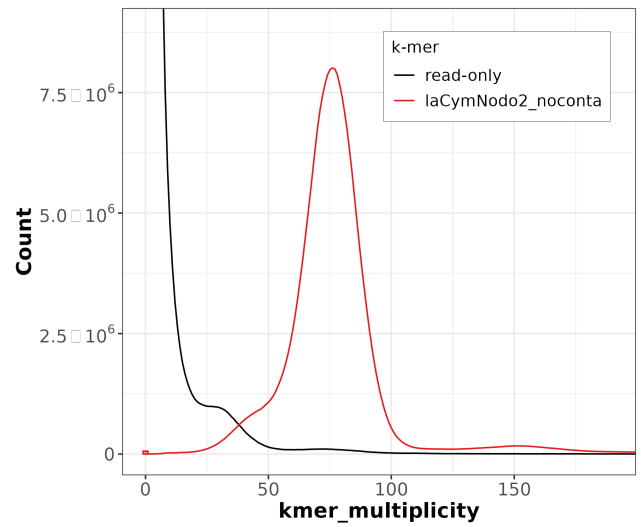


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

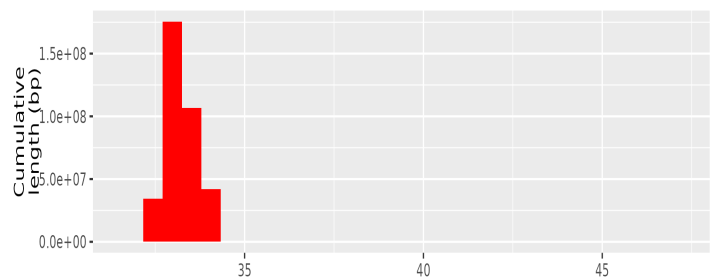


Distribution of k-mer counts per copy numbers found in asm

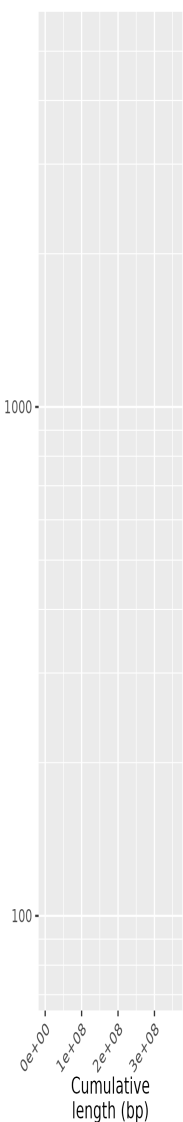
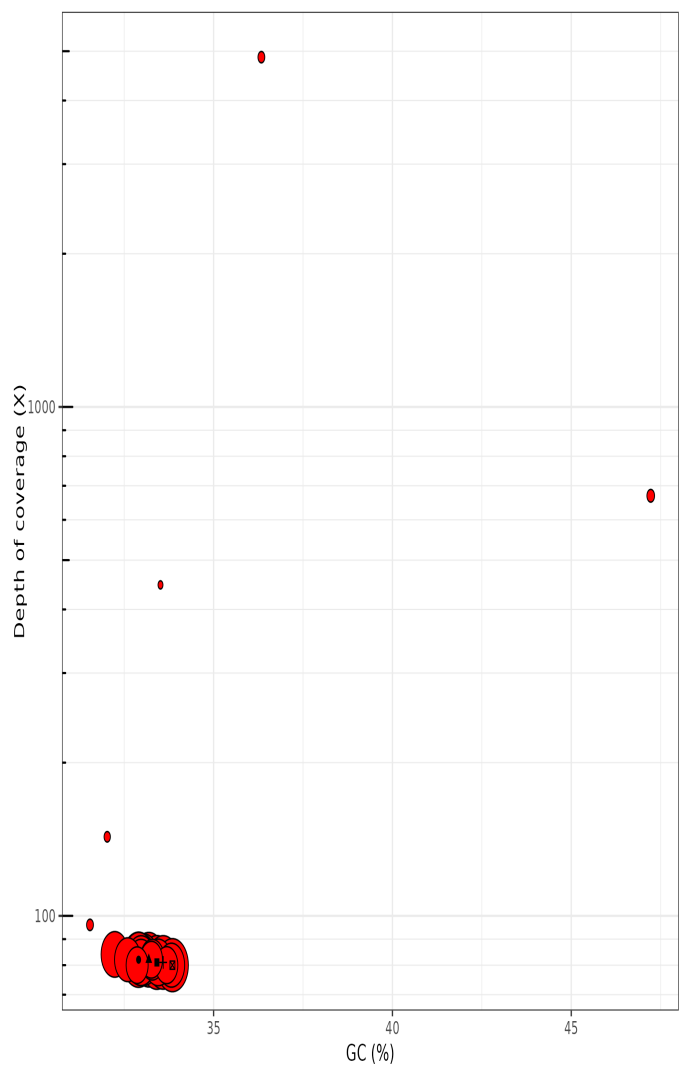


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



- Longest sequences (bp)
- SUPER_1 - 28285218 (Eukaryota)
 - ▲ SUPER_2 - 27559890 (Eukaryota)
 - SUPER_3 - 26636253 (Eukaryota)
 - + SUPER_4 - 26344093 (Eukaryota)
 - ▣ SUPER_5 - 25780390 (Eukaryota)

- Length (bp)
- 1e+07
 - 2e+07

- superkingdom
- Eukaryota

collapsed. 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	PACBIO Hifi	Arima
Coverage	40	100

Assembly pipeline

- **Hifiasm**
 - |_ *ver*: 0.19.5-r593
 - |_ *key param*: NA
- **purge_dups**
 - |_ *ver*: 1.2.5
 - |_ *key param*: NA
- **YaHS**
 - |_ *ver*: 1.2
 - |_ *key param*: NA

Curation pipeline

- **PretextMap**
 - |_ *ver*: 0.1.9
 - |_ *key param*: NA
- **PretextView**
 - |_ *ver*: 0.2.5
 - |_ *key param*: NA

Submitter: Benjamin Istace

Affiliation: Genoscope

Date and time: 2025-01-23 02:47:46 CET