ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	661986
ToLID	**jhAmpOper1**
Species	Amphisbetia operculata
Class	Hydrozoa
Order	Leptothecata

Genome Traits	Expected	Observed
Haploid size (bp)	264,223,418	276,098,716
Haploid Number	6 (source: ancestor)	15
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q49

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

. Observed Haploid Number is different from Expected

Curator notes

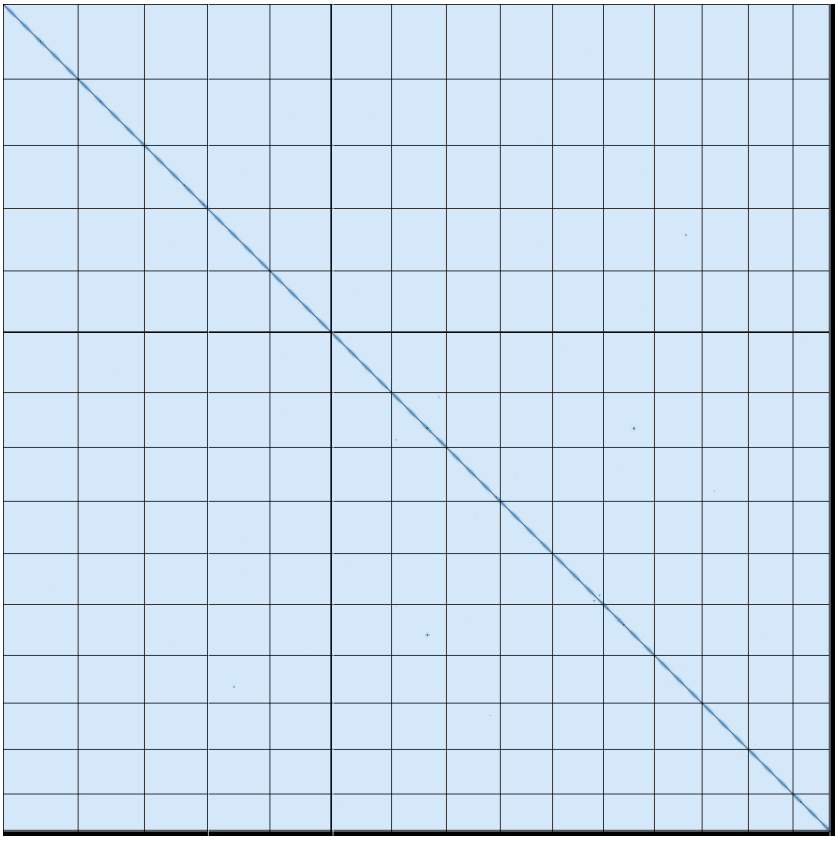
. Interventions/Gb: 54
. Contamination notes: ""
. Other observations: "The assembly of Amphisbetia operculata (jhAmpOper1) is based on 518X ONT data and 175X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial ONT assembly generation with Hifiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. In total, 220 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 12.263 Mb (with the largest being 0.682 Mb). Additionally, 62 regions totaling 4.237 Mb (with the largest being 0.637 Mb) were identified as haplotypic duplications and removed. Mitochondrial genome was assembled using ptGAUL. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 4 haplotypic regions and 1 contaminant sequence were removed, totaling 1.081Mb and 0.02Mb, respectively (with the largest being 0.771Mb and 0.02Mb). Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	277,179,129	276,098,716
GC %	34.52	34.52
Gaps/Gbp	61.33	76.06
Total gap bp	1,700	2,828
Scaffolds	69	64
Scaffold N50	18,172,485	18,172,485
Scaffold L50	6	7
Scaffold L90	13	14
Contigs	86	85
Contig N50	16,034,271	16,034,271
Contig L50	8	8
Contig L90	18	18
QV	49.3456	49.7301
Kmer compl.	91.0382	91.0021
BUSCO sing.	89.7%	90.1%
BUSCO dupl.	1.9%	1.5%
BUSCO frag.	4.0%	4.0%
BUSCO miss.	4.4%	4.4%

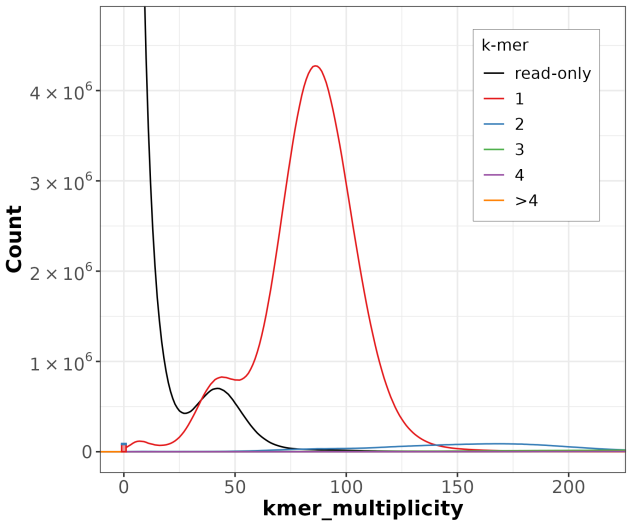
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: cnidaria_odb12 (genomes:17, BUSCOs:3203)

HiC contact map of curated assembly

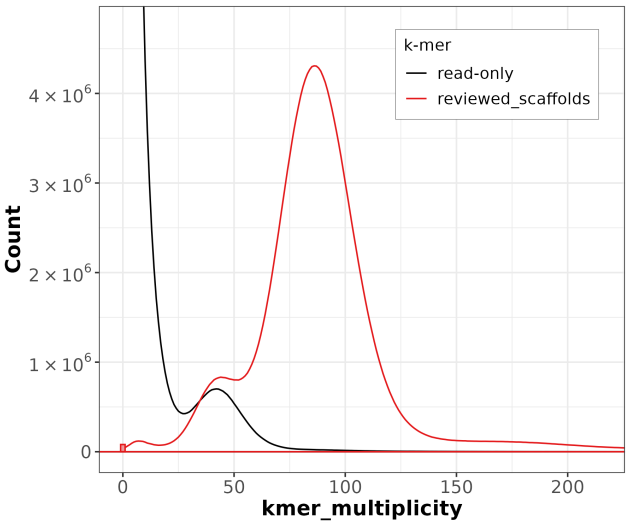


collapsed [LINK]

K-mer spectra of curated assembly

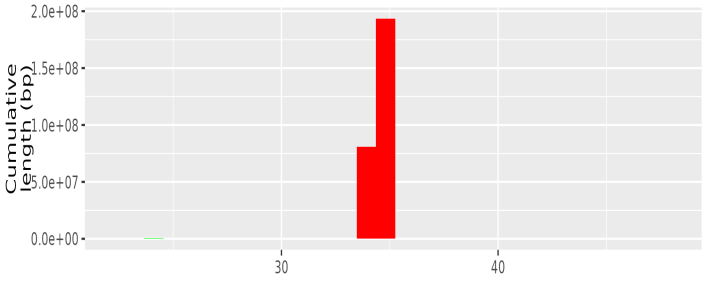


Distribution of k-mer counts per copy numbers found in asm

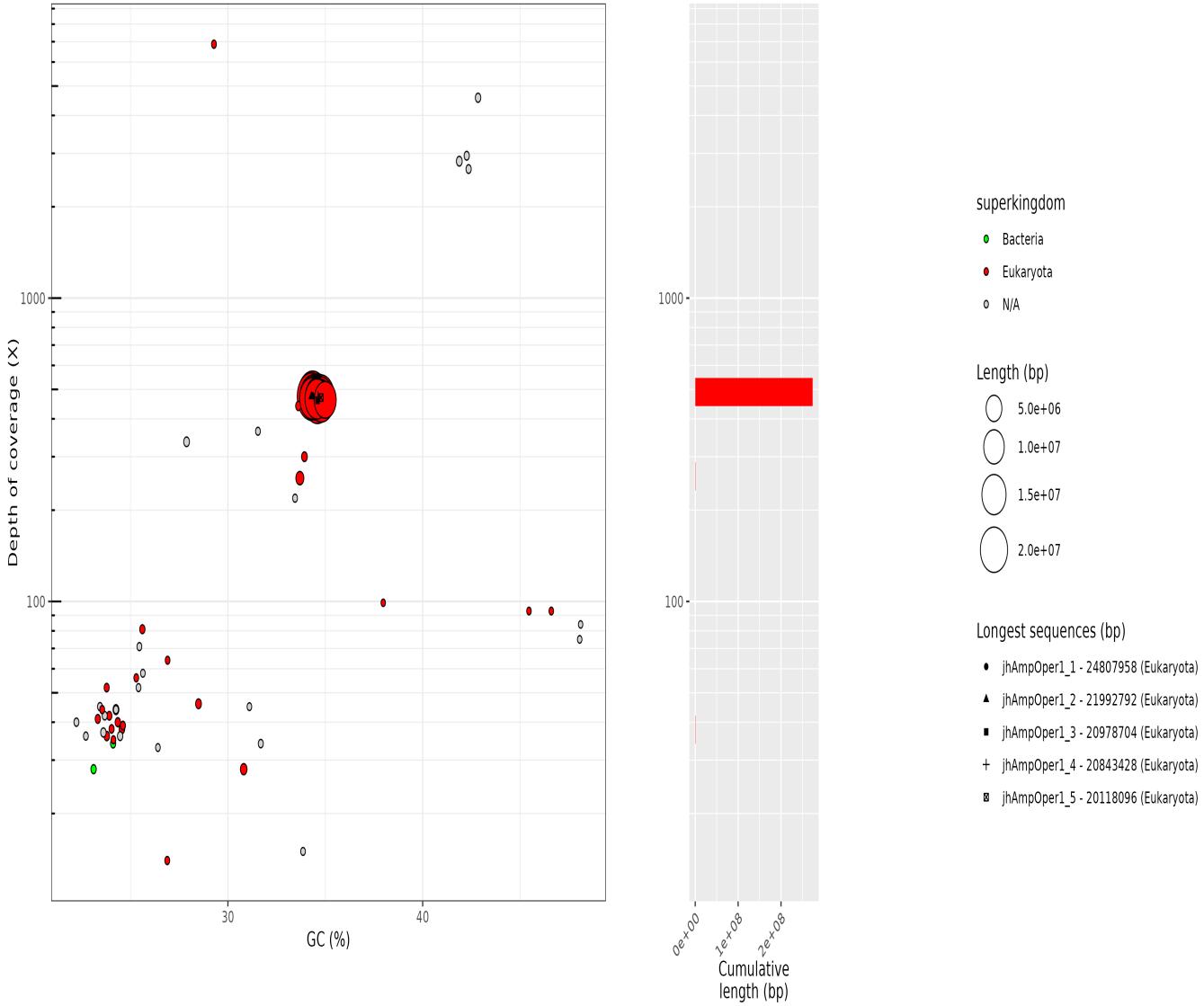


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



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	Long reads	Arima
Coverage	520	176

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: Emilie Teodori
Affiliation: Genoscope

Date and time: 2026-04-27 16:00:36 CEST