

# ERGA Assembly Report

v24.09.10

Tags: ATLASea[INVALID TAG]

TxID	691888
ToLID	<b>eoAstMedi1.1</b>
Species	Astrospartus mediterraneus
Class	Ophiuroidea
Order	Euryalida

Genome Traits	Expected	Observed
Haploid size (bp)	2,187,344,310	2,293,599,108
Haploid Number	18 (source: ancestor)	23
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

## EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q54

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

- . Observed Haploid Number is different from Expected

### Curator notes

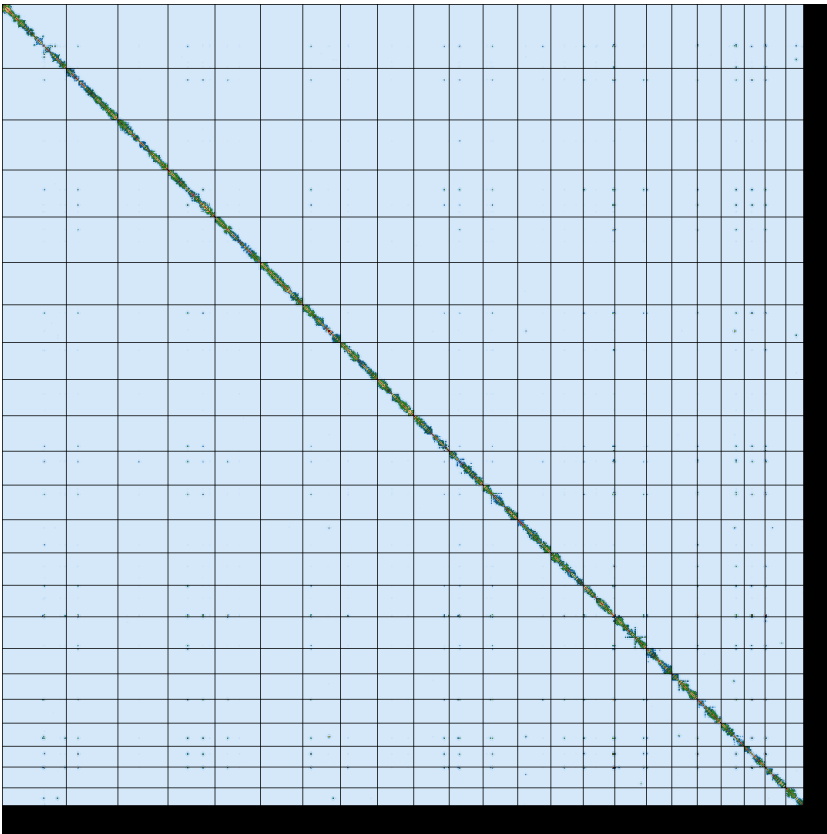
- . Interventions/Gb: 12
- . Contamination notes: "23 bacterial sequences of 2 Mb and one virus sequence of 19 Kb detected. These were removed before scaffolding. "
- . Other observations: "The post-curation contamination screening plot corresponds to the assembly before bacterial sequences are removed. "

## Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	2,312,399,603	2,293,599,108
GC %	37.77	37.78
Gaps/Gbp	435.05	440.36
Total gap bp	100,600	102,500
Scaffolds	1,038	883
Scaffold N50	97,329,811	97,329,811
Scaffold L50	10	10
Scaffold L90	21	21
Contigs	2,044	1,893
Contig N50	3,522,000	3,531,845
Contig L50	192	190
Contig L90	636	629
QV	44.1581	54.7172
Kmer compl.	96.656	96.1969
BUSCO sing.	95.2%	95.2%
BUSCO dupl.	0.9%	0.9%
BUSCO frag.	2.2%	2.2%
BUSCO miss.	1.7%	1.7%

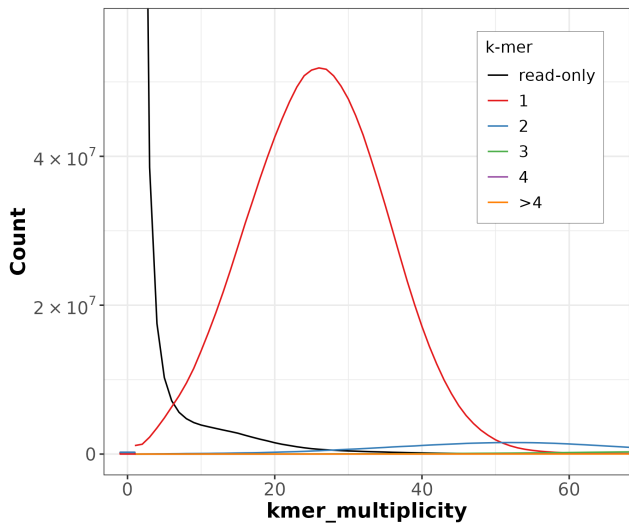
BUSCO 5.4.3 Lineage: metazoa\_odb10 (genomes:65, BUSCOs:954)

# HiC contact map of curated assembly

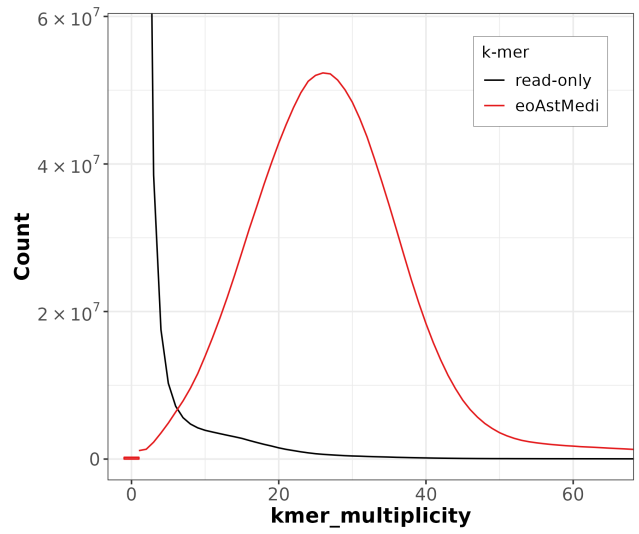


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

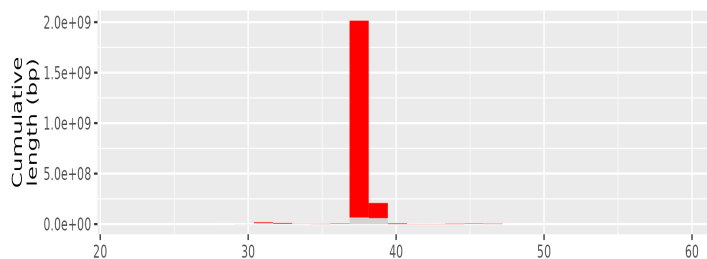


Distribution of k-mer counts per copy numbers found in asm

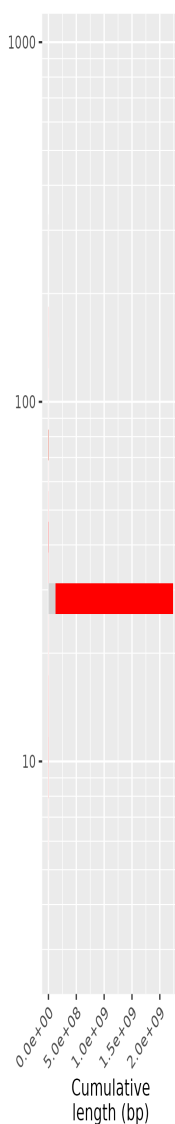
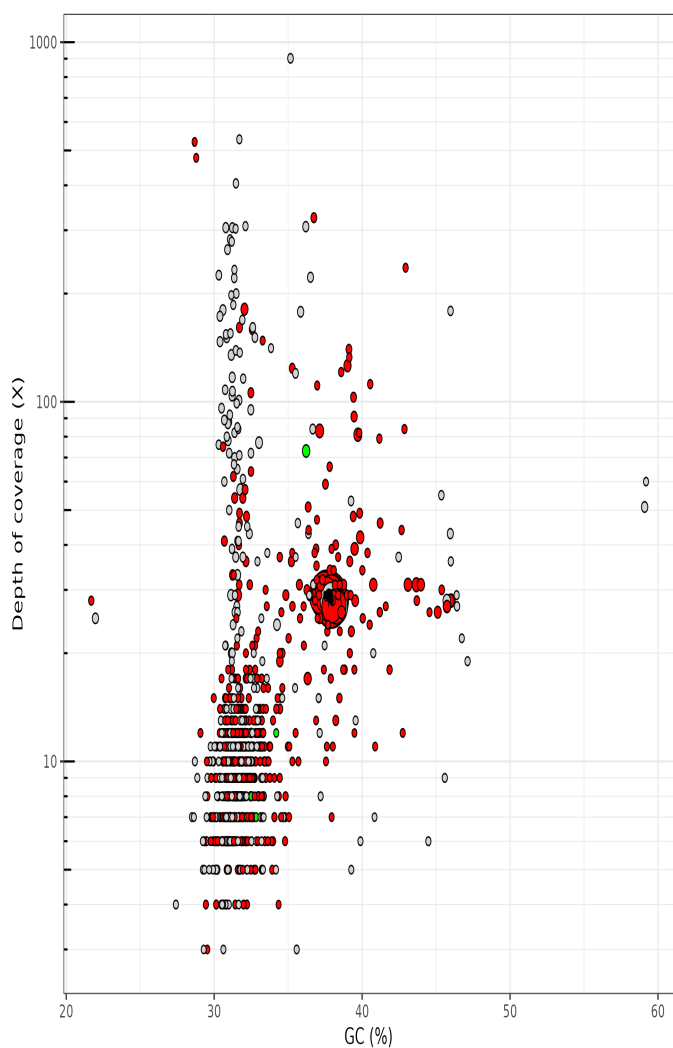


Distribution of k-mer counts coloured by their presence in reads/assemblies

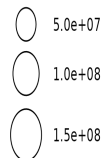
# Post-curation contamination screening



## TAPAs summary Graph



Length (bp)



Longest sequences (bp)

- SUPER\_1 - 176332233 (Eukaryota)
- ▲ SUPER\_2 - 142400392 (Eukaryota)
- SUPER\_3 - 138018593 (Eukaryota)
- + SUPER\_4 - 131069952 (Eukaryota)
- ⊠ SUPER\_5 - 123925664 (Eukaryota)

superkingdom

- Bacteria
- Eukaryota
- N/A

**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 (4-enz)
Coverage	30	20

# 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: Adama Ndar  
Affiliation: Genoscope

Date and time: 2024-11-19 14:47:07 CET