

# ERGA Assembly Report

v24.09.10

Tags: ATLASea[INVALID TAG]

TxID	396355
ToLID	<b>jaEunVerr7.1</b>
Species	Eunicella verrucosa
Class	Anthozoa
Order	Malacalcyonacea

Genome Traits	Expected	Observed
Haploid size (bp)	372,096,591	433,379,001
Haploid Number	6 (source: ancestor)	17
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

## EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q60

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 collapsed
- . Assembly length loss > 3% for collapsed

### Curator notes

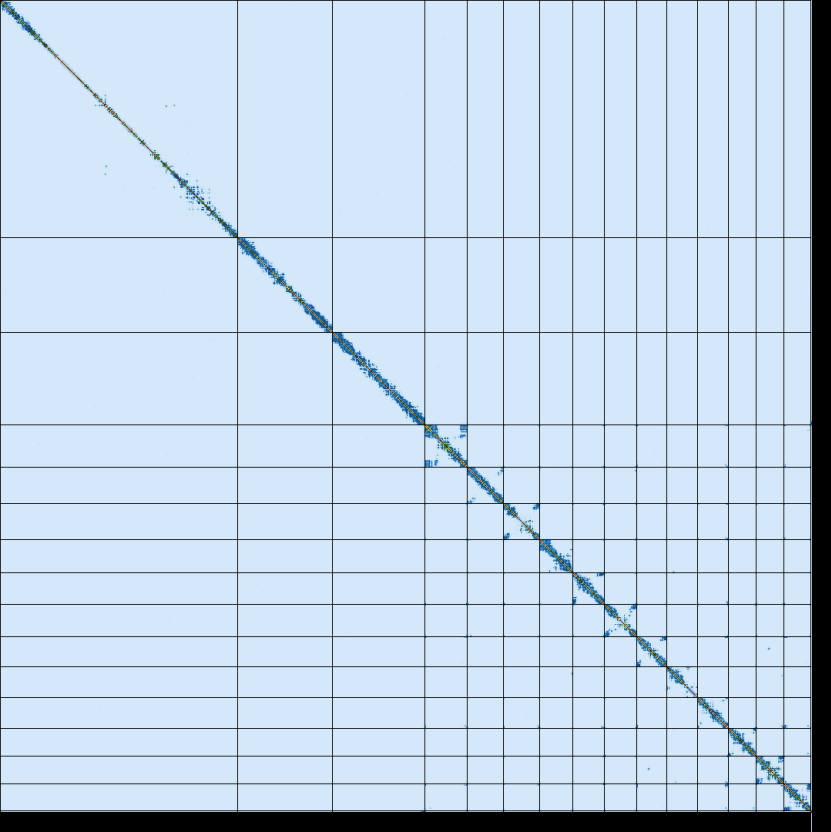
- . Interventions/Gb: 256
- . Contamination notes: ""
- . Other observations: "Numerous haplotypic duplications were removed "

## Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	447,672,612	433,379,001
GC %	37.55	37.56
Gaps/Gbp	0	170.75
Total gap bp	0	11,500
Scaffolds	219	195
Scaffold N50	15,787,020	47,888,948
Scaffold L50	9	3
Scaffold L90	27	13
Contigs	219	269
Contig N50	15,787,020	9,098,906
Contig L50	9	15
Contig L90	27	47
QV	60.462	60.4431
Kmer compl.	77.9179	76.7507
BUSCO sing.	94.5%	95.3%
BUSCO dupl.	1.2%	0.4%
BUSCO frag.	2.7%	2.7%
BUSCO miss.	1.6%	1.6%

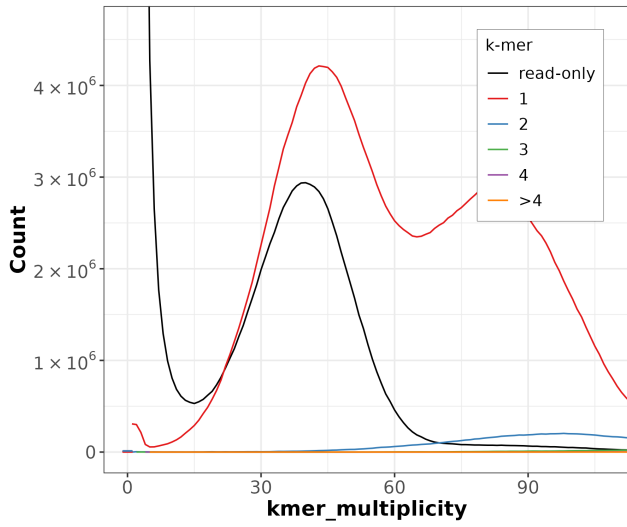
BUSCO 5.4.3 Lineage: eukaryota\_odb10 (genomes:70, BUSCOs:255)

# HiC contact map of curated assembly

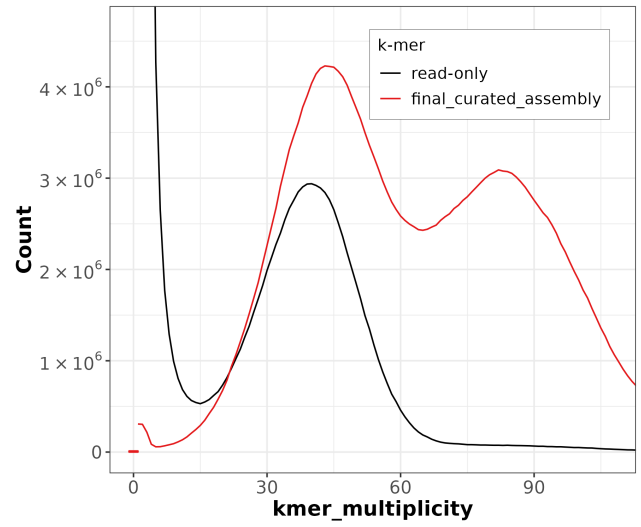


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

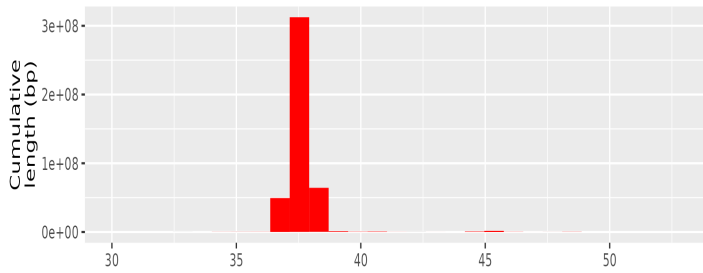


Distribution of k-mer counts per copy numbers found in asm

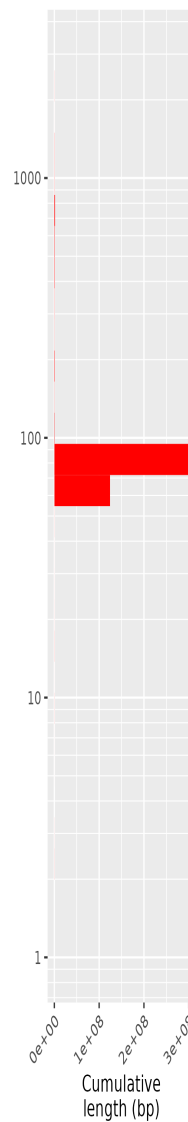
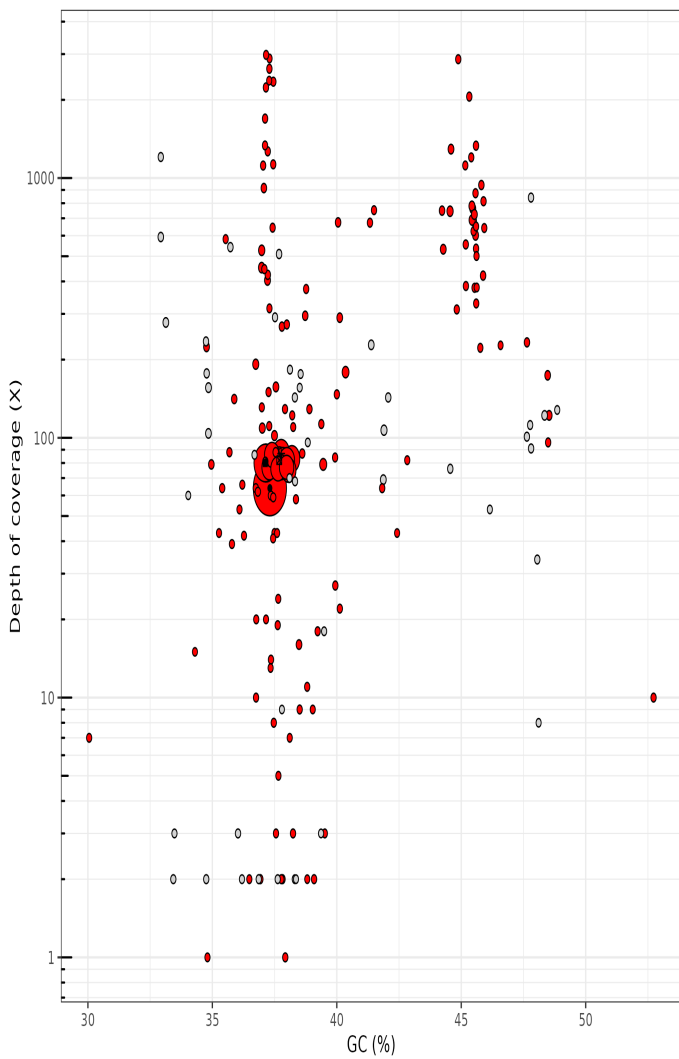


Distribution of k-mer counts coloured by their presence in reads/assemblies

# Post-curation contamination screening



## TAPAs summary Graph



### Longest sequences (bp)

- jaEunVerr\_1 - 123790362 (Eukaryota)
- ▲ jaEunVerr\_2 - 49523602 (Eukaryota)
- jaEunVerr\_3 - 47888948 (Eukaryota)
- + jaEunVerr\_4 - 22201749 (Eukaryota)
- ▣ jaEunVerr\_5 - 19014071 (Eukaryota)

### superkingdom

- Eukaryota
- N/A

### Length (bp)

- 3.0e+07
- 6.0e+07
- 9.0e+07
- 1.2e+08

**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	77	117

# 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: Caroline Belser

Affiliation: Genoscope

Date and time: 2024-10-18 10:39:18 CEST