ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	72676
ToLID	**eaCosTenu1**
Species	Coscinasterias tenuispina
Class	Asteroidea
Order	Forcipulatida

Genome Traits	Expected	Observed
Haploid size (bp)	550,548,842	538,324,313
Haploid Number	15 (source: ancestor)	22
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q49

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

Curator notes

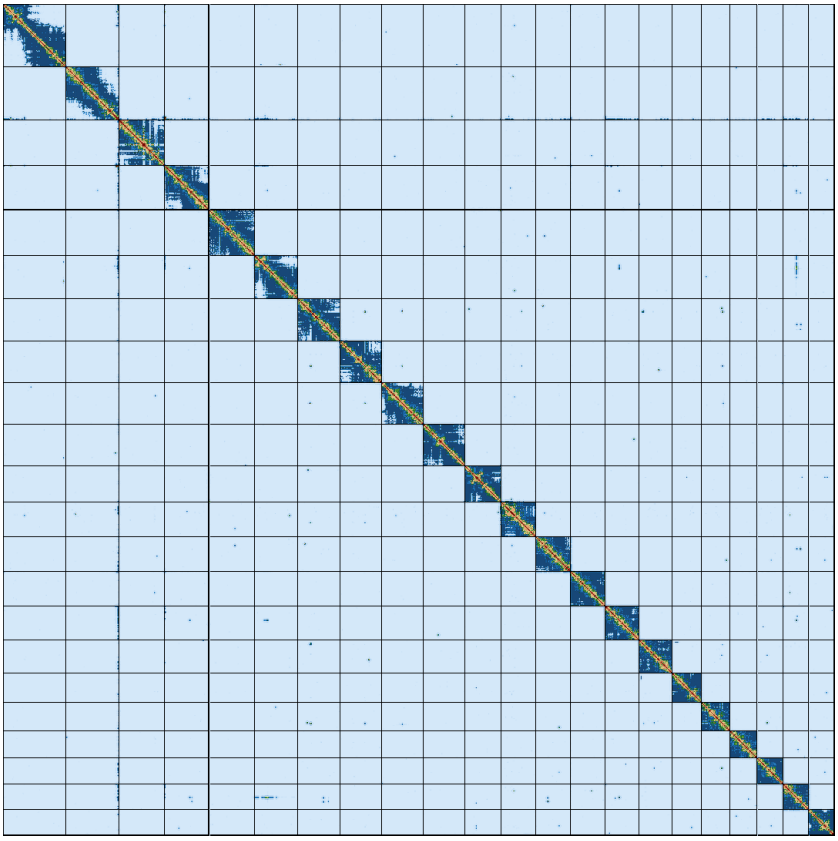
. Interventions/Gb: 86
. Contamination notes: ""
. Other observations: "The assembly of Coscinasterias tenuispina (eaCosTenu1) is based on 34X PacBio data and 205X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial PacBio assembly generation with Hifiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. No contigs were found to be contaminants (bacterial, archaeal, or viral). Additionally, 241 regions totaling 22.339 Mb (with the largest being 1.151 Mb) were identified as haplotypic duplications and removed. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 6 haplotypic regions were removed, totaling 4.947Mb (with the largest being 2.512Mb). 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	544,769,663	538,324,313
GC %	39.81	39.82
Gaps/Gbp	313.89	338.09
Total gap bp	17,100	20,500
Scaffolds	59	45
Scaffold N50	26,697,252	26,945,586
Scaffold L50	10	9
Scaffold L90	20	19
Contigs	230	227
Contig N50	5,077,107	5,030,271
Contig L50	33	33
Contig L90	109	108
QV	49.9259	49.9115
Kmer compl.	89.9418	89.6932
BUSCO sing.	97.6%	97.8%
BUSCO dupl.	0.7%	0.6%
BUSCO frag.	1.2%	1.2%
BUSCO miss.	0.4%	0.4%

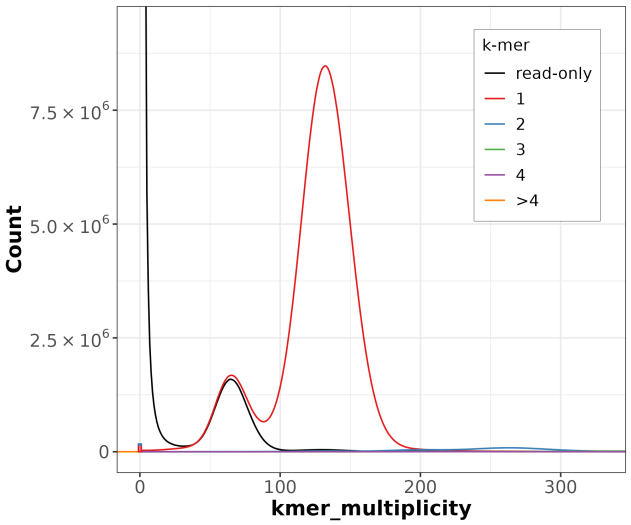
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: metazoa_odb12 (genomes:206, BUSCOs:672)

HiC contact map of curated assembly

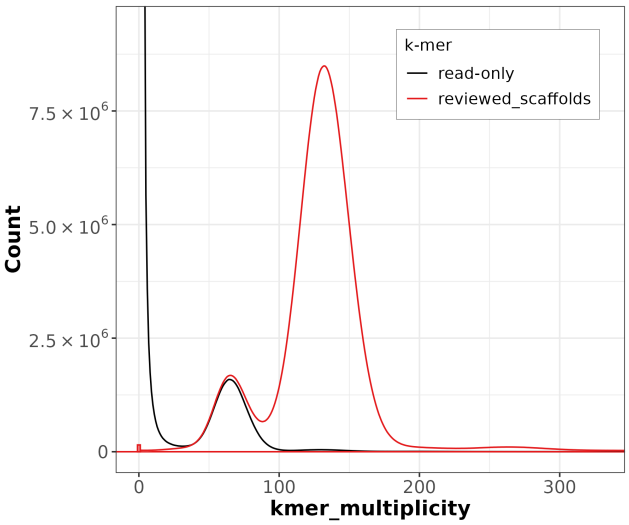


collapsed [LINK]

K-mer spectra of curated assembly

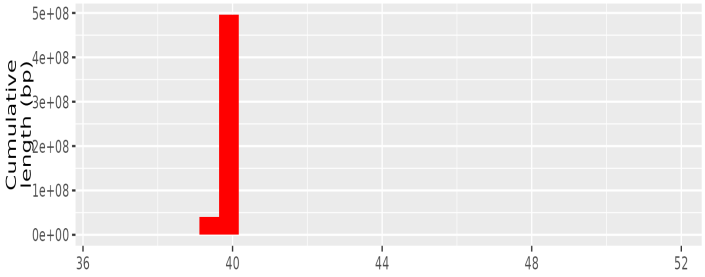


Distribution of k-mer counts per copy numbers found in asm

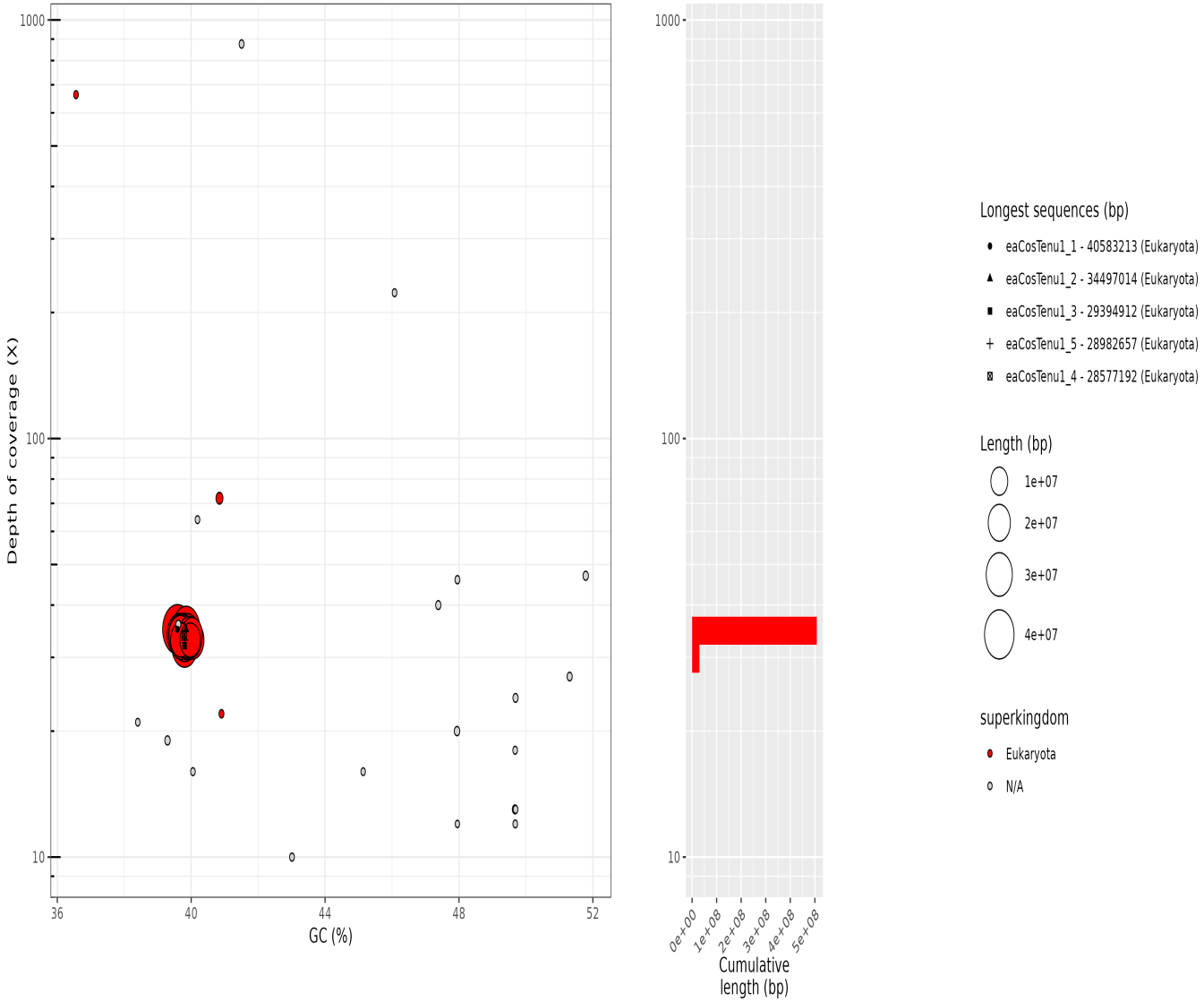


Distribution of k-mer counts coloured by their presence in reads/assemblies

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	34	207

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: Mohamed Meftah Saoues
Affiliation: Genoscope

Date and time: 2026-05-12 20:15:28 CEST