ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	706644
ToLID	**ecClaLitt1**
Species	Clarkcomanthus littoralis
Class	Crinoidea
Order	Comatulida

Genome Traits	Expected	Observed
Haploid size (bp)	741,276,809	819,114,027
Haploid Number	4 (source: ancestor)	11
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q41

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
. More than 1000 gaps/Gbp for collapsed

Curator notes

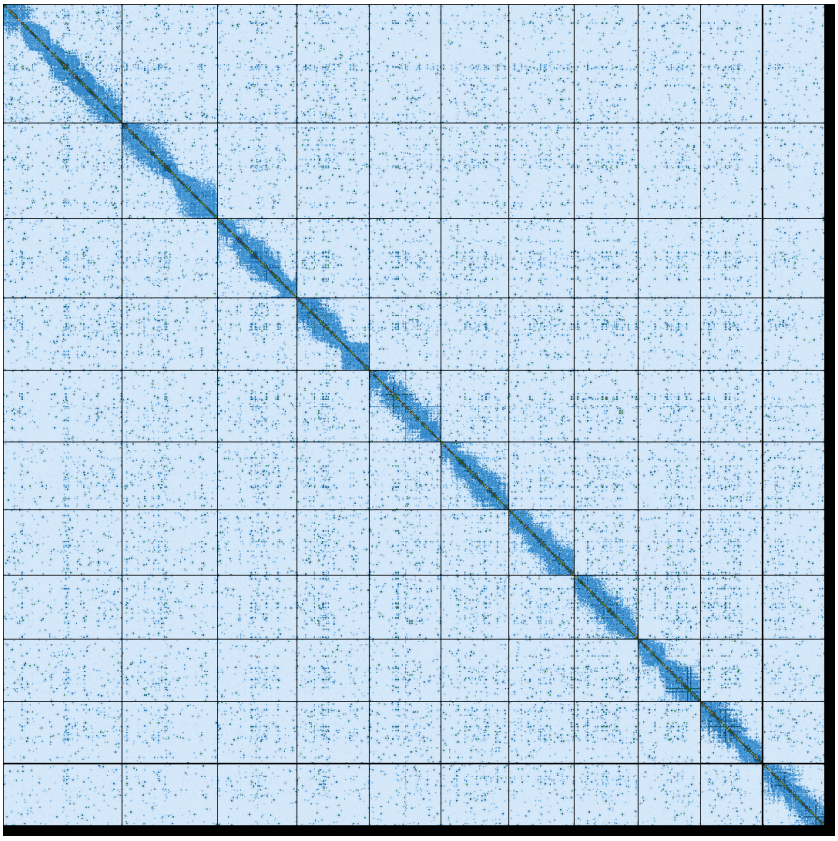
. Interventions/Gb: 128
. Contamination notes: ""
. Other observations: "The assembly of Clarkcomanthus littoralis (ecClaLitt1) is based on 49X PacBio data and 224X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial PacBio+HiC assembly generation with Hifiasm, removal of contaminant sequences using Context, and Hi-C-based scaffolding with YaHS of each haplotype. In total, 3 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 0.069 Mb (with the largest being 0.025 Mb). Mitochondrial genome was assembled using OATK. Finally, Both haplotypes were analyzed and manually improved using Pretext. During manual curation, 20 haplotypic regions were removed, totaling 2.863Mb (with the largest being 0.37Mb). 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	846,971,724	819,114,027
GC %	33.87	33.91
Gaps/Gbp	404.97	1,001.08
Total gap bp	34,300	90,200
Scaffolds	132	250
Scaffold N50	69,060,661	70,203,679
Scaffold L50	5	5
Scaffold L90	10	10
Contigs	475	1,070
Contig N50	4,321,568	1,605,569
Contig L50	59	157
Contig L90	209	506
QV	42.9014	41.7737
Kmer compl.	79.8402	77.7379
BUSCO sing.	95.4%	95.8%
BUSCO dupl.	3.1%	1.8%
BUSCO frag.	0.9%	1.8%
BUSCO miss.	0.6%	0.6%

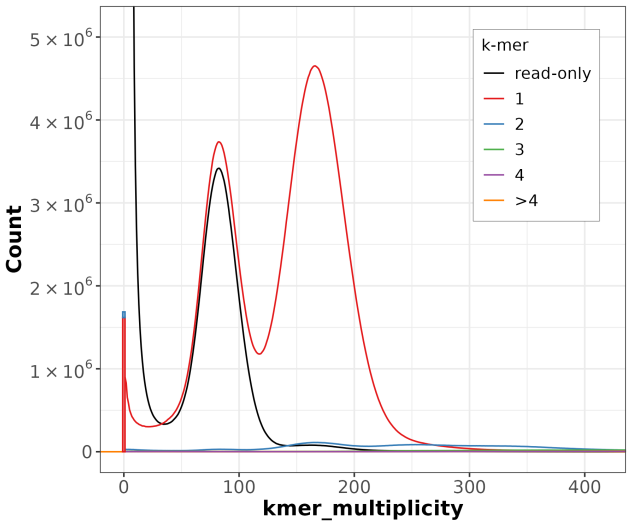
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: metazoa_odb12 (genomes:206, BUSCOs:672)
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: metazoa_odb12 (genomes:206, BUSCOs:672)

HiC contact map of curated assembly

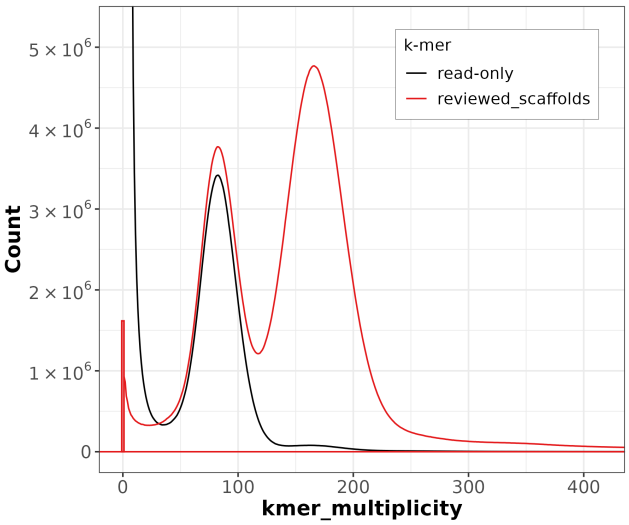


collapsed [LINK]

K-mer spectra of curated assembly

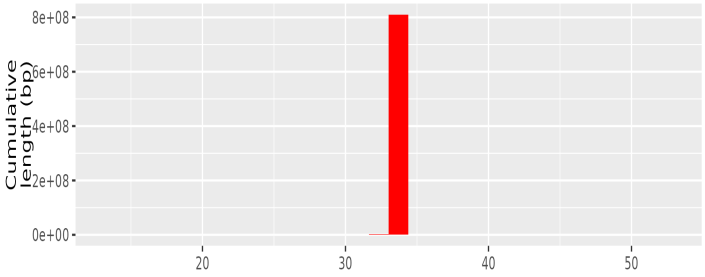


Distribution of k-mer counts per copy numbers found in asm

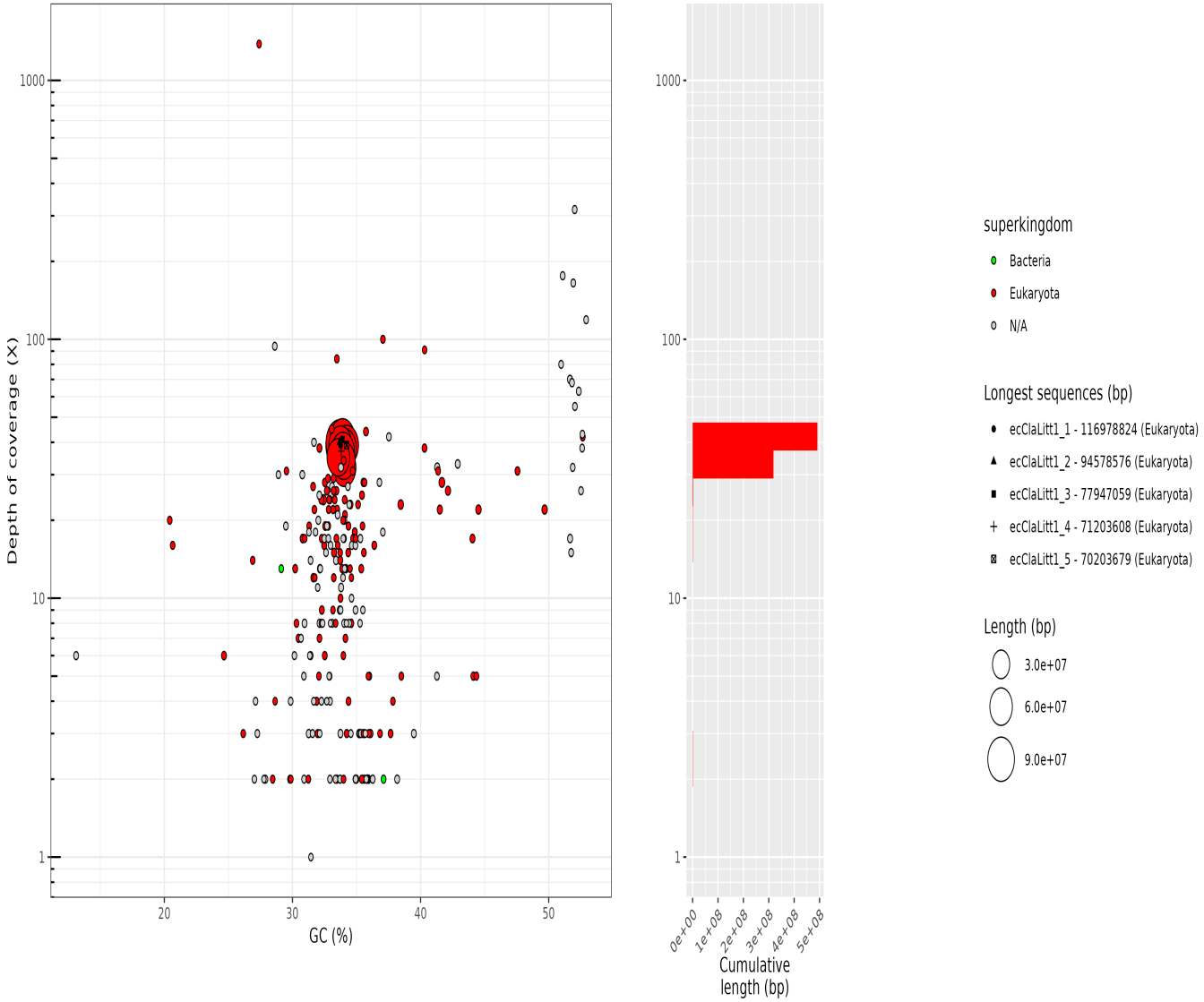


Distribution of k-mer counts coloured by their presence in reads/assemblies

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	45	203

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: Jean-Marc Aury
Affiliation: Genoscope

Date and time: 2026-08-08 15:51:57 CEST