ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	1320720
ToLID	**ecLepCelt1.1**
Species	Leptometra celtica
Class	Crinoidea
Order	Comatulida

Genome Traits	Expected	Observed
Haploid size (bp)	376,345,443	364,470,239
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: 7.7.Q61

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

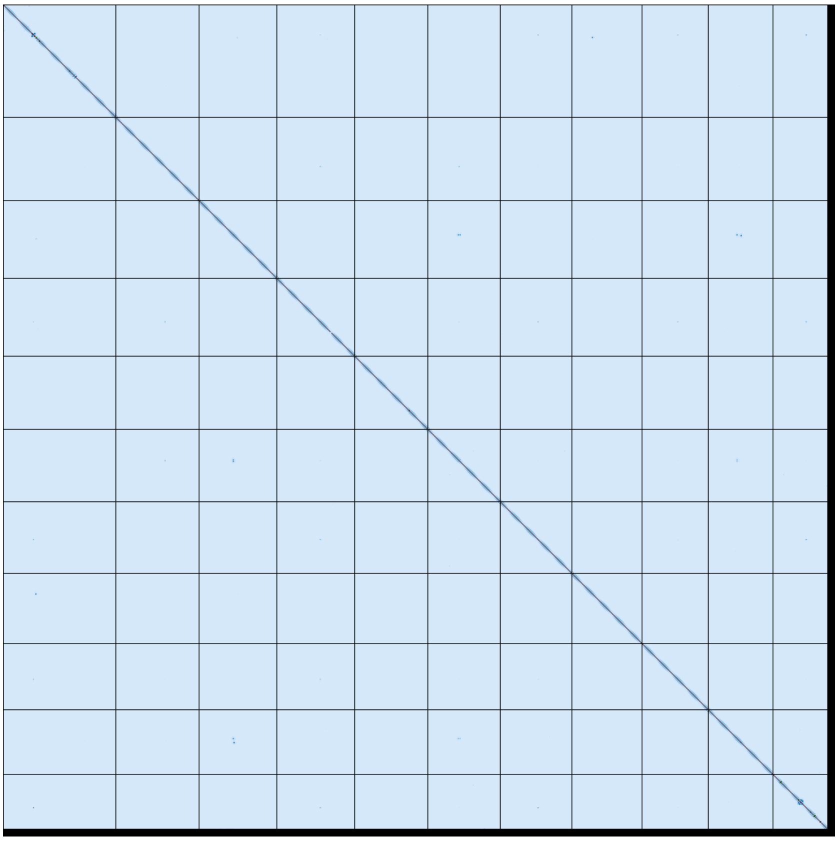
. Interventions/Gb: 183
. Contamination notes: "34 contigs were identified as bacterial sequences and removed, totaling 1.15 Mb, with the largest being 125 Kb. "
. Other observations: "The structure of chromosome 1 is uncertain and might correspond to two chromosomes (break around 21.1Mb). The Hi-C signal was very weak and insufficient to validate the structure. We removed 141 haplotypic regions, totaling 39 Mb, with the largest being 3 Mb. "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	365,931,043	364,470,239
GC %	35.27	35.27
Gaps/Gbp	95.65	181.08
Total gap bp	3,500	10,400
Scaffolds	74	48
Scaffold N50	31,218,100	32,031,400
Scaffold L50	6	5
Scaffold L90	11	10
Contigs	109	114
Contig N50	14,841,870	13,203,000
Contig L50	8	9
Contig L90	26	29
QV	52.1469	61.6572
Kmer compl.	86.7497	86.9424
BUSCO sing.	95.0%	95.2%
BUSCO dupl.	0.3%	0.1%
BUSCO frag.	3.4%	3.4%
BUSCO miss.	1.3%	1.3%

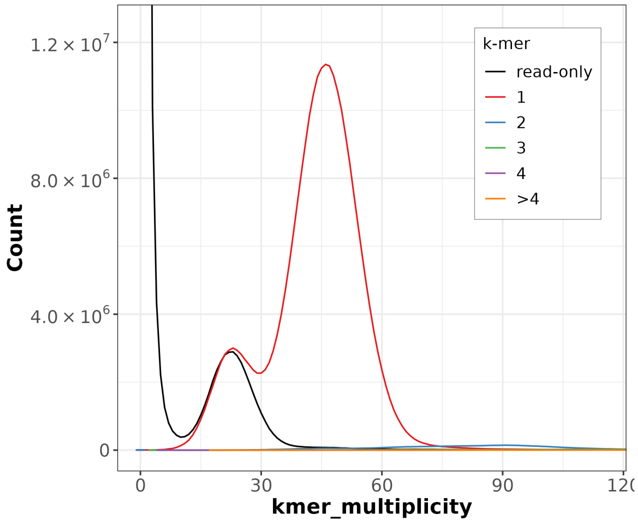
BUSCO: 5.4.3 (euk_genome_met, metaeuk) / Lineage: metazoa_odb10 (genomes:65, BUSCOs:954)

HiC contact map of curated assembly

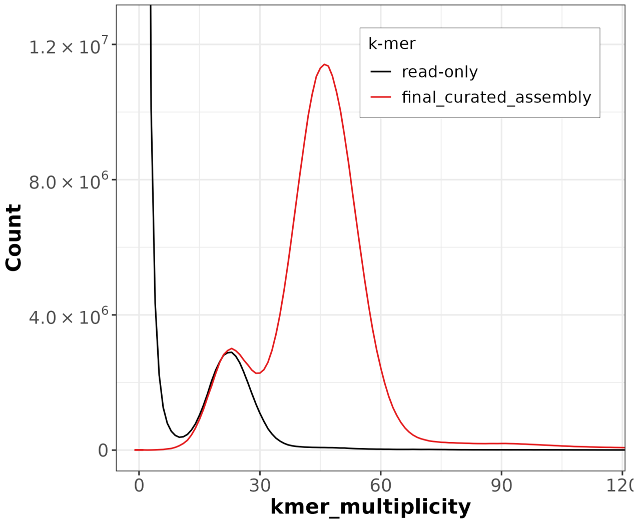


collapsed [LINK]

K-mer spectra of curated assembly

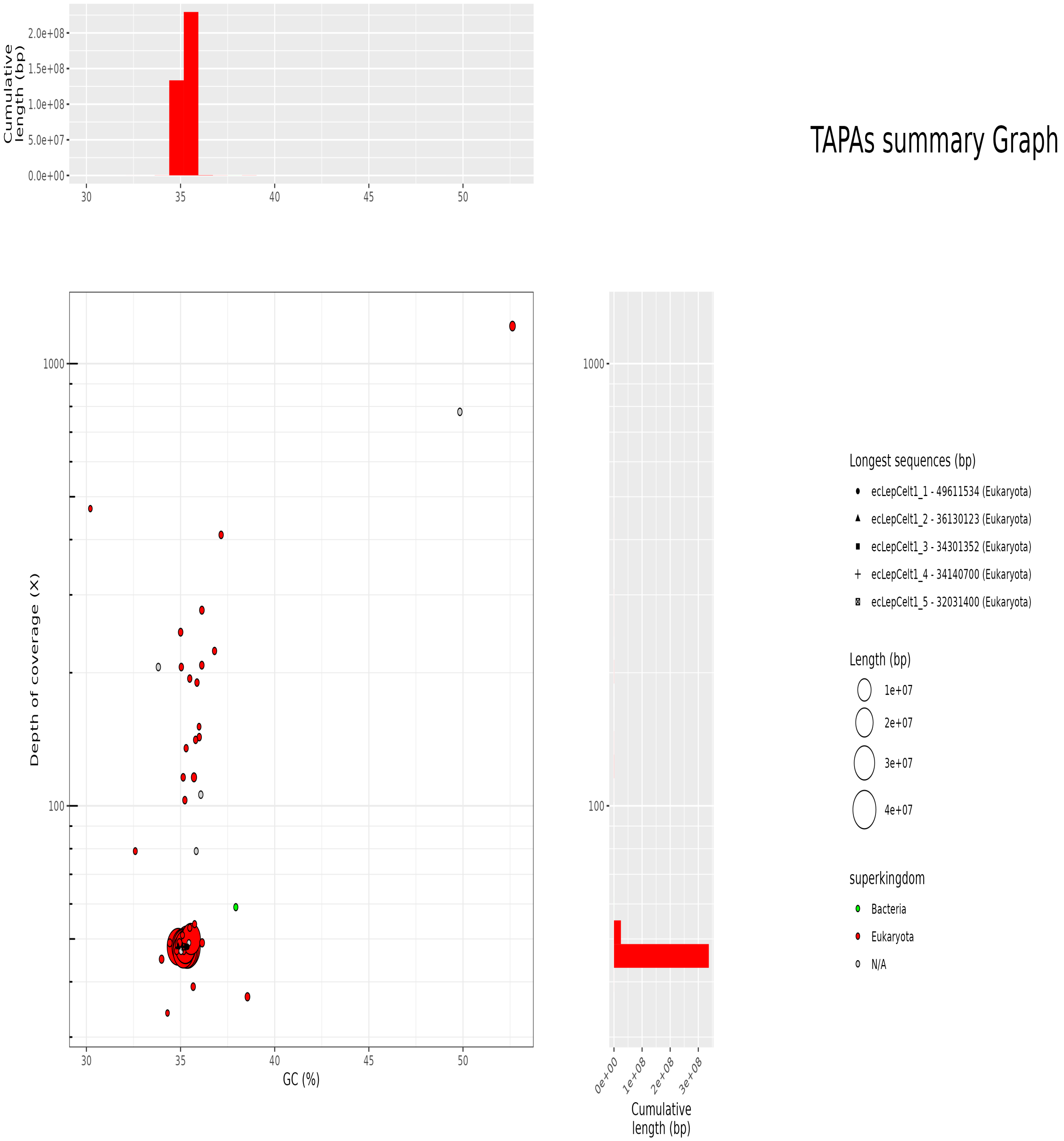


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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
Coverage	47	234

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: 2024-11-26 15:02:13 CET