ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	2496499
ToLID	**xgFisRose1**
Species	Fissurella rosea
Class	Gastropoda
Order	Lepetellida

Genome Traits	Expected	Observed
Haploid size (bp)	417,191,445	424,203,741
Haploid Number	9 (source: ancestor)	16
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q50

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

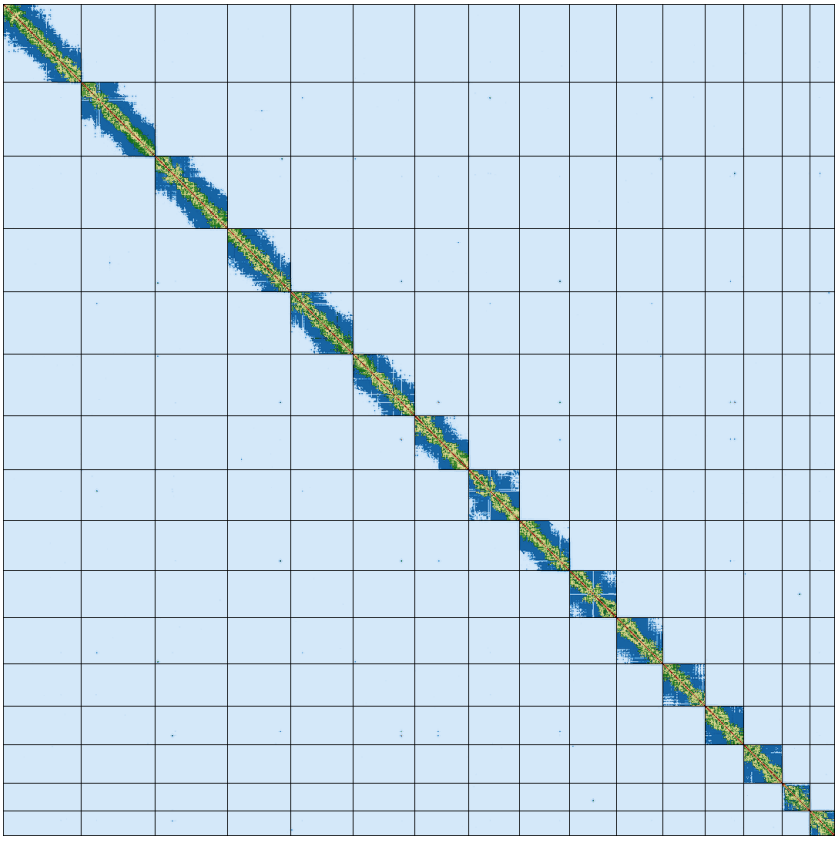
. Interventions/Gb: 12
. Contamination notes: ""
. Other observations: "The assembly of Fissurella nubecula (xgFisRose1) is based on 58X PacBio data and 191X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial 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. In total, 42 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 6.03 Mb (with the largest being 1.925 Mb). Additionally, 20 regions totaling 1.556 Mb (with the largest being 0.401 Mb) were identified as haplotypic duplications and removed. Mitochondrial genome was assembled using OATK. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 1 haplotypic region was removed, totaling 0.16Mb. 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	424,363,042	424,203,741
GC %	37.56	37.56
Gaps/Gbp	65.98	66.01
Total gap bp	2,800	3,000
Scaffolds	20	20
Scaffold N50	27,497,492	27,497,492
Scaffold L50	7	7
Scaffold L90	14	14
Contigs	48	48
Contig N50	17,017,951	17,017,951
Contig L50	10	10
Contig L90	25	25
QV	50.7859	50.789
Kmer compl.	81.7261	81.7226
BUSCO sing.	98.6%	98.6%
BUSCO dupl.	0.4%	0.4%
BUSCO frag.	0.3%	0.3%
BUSCO miss.	0.7%	0.7%

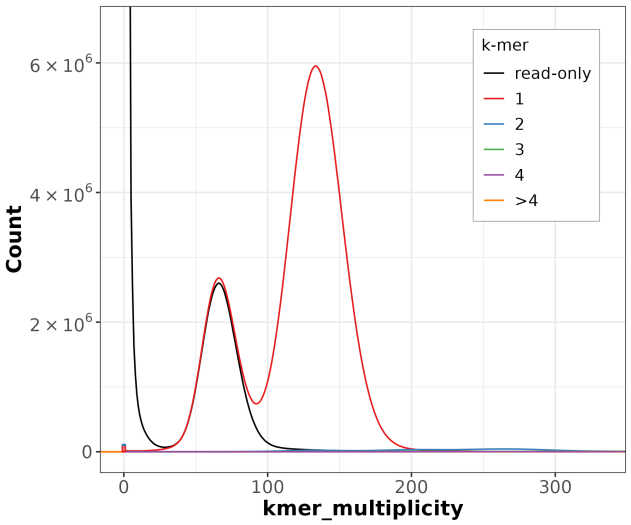
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: mollusca_odb12 (genomes:36, BUSCOs:4421)

HiC contact map of curated assembly

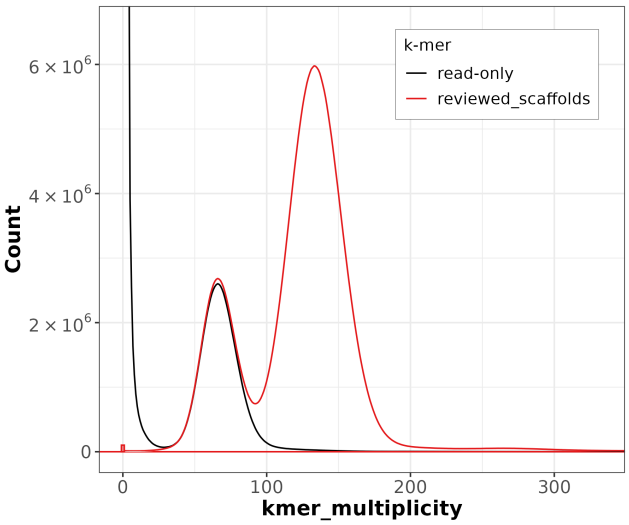


collapsed [LINK]

K-mer spectra of curated assembly

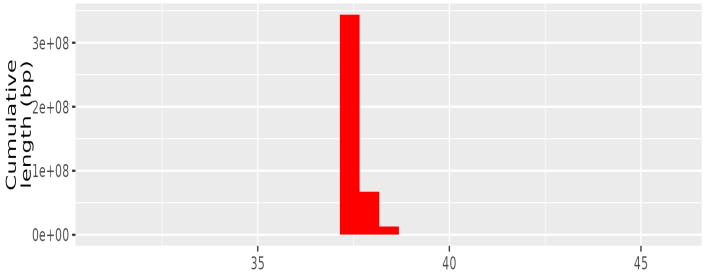


Distribution of k-mer counts per copy numbers found in asm

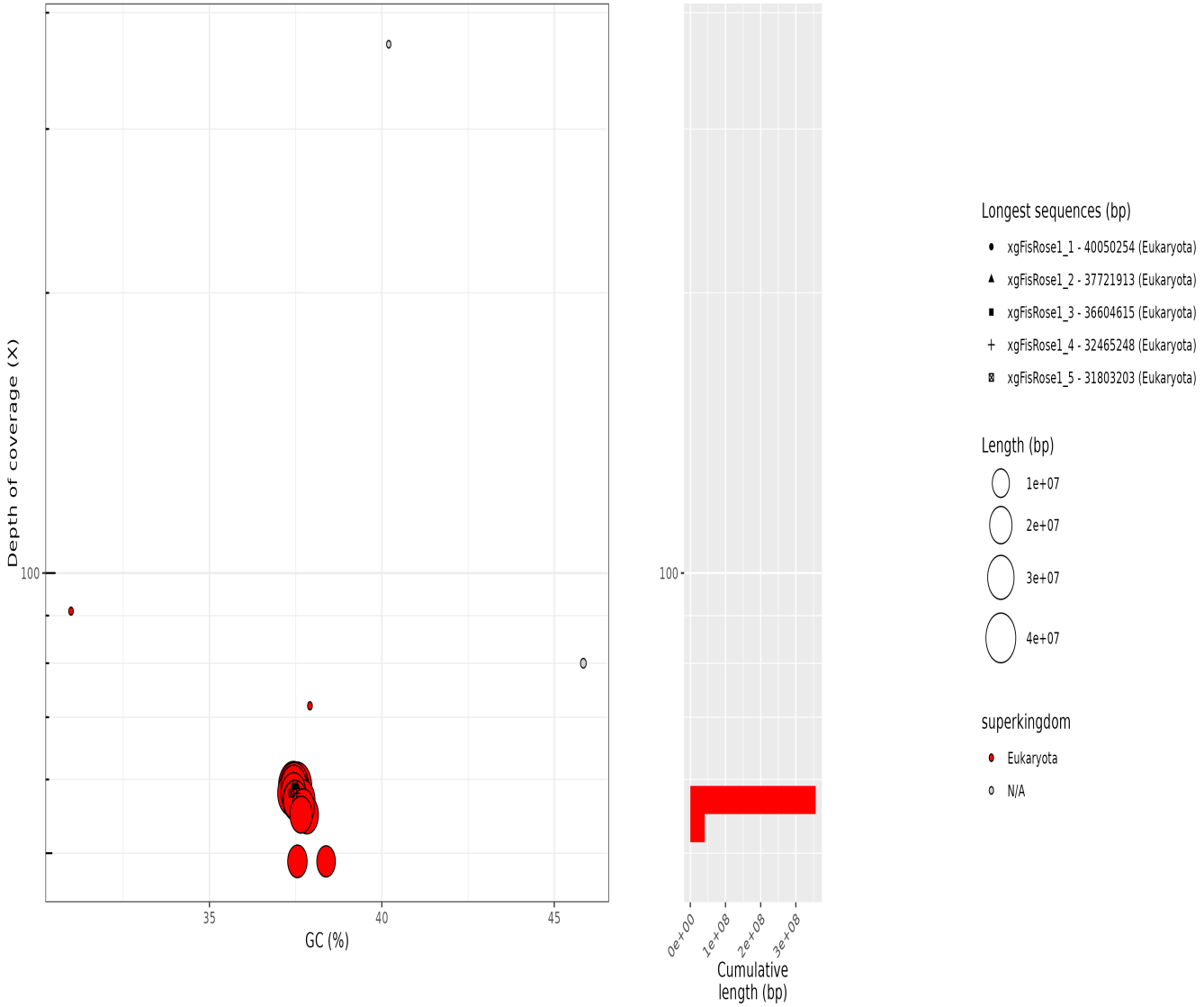


Distribution of k-mer counts coloured by their presence in reads/assemblies

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	58	190

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: 2026-06-17 21:52:53 CEST