ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	152550
ToLID	**fParZvo1**
Species	Parablennius zvonimiri
Class	Actinopteri
Order	Blenniiformes

Genome Traits	Expected	Observed
Haploid size (bp)	756,171,665	791,974,536
Haploid Number	19 (source: ancestor)	24
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q44

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

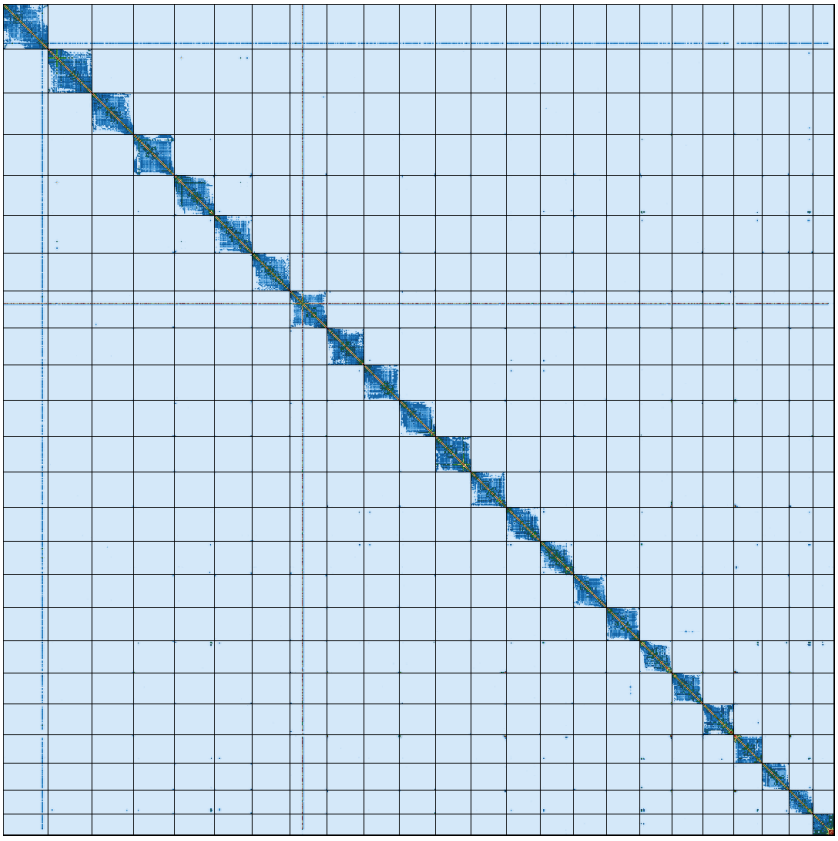
. Interventions/Gb: 18
. Contamination notes: ""
. Other observations: "The assembly of Parablennius zvonimiri (fParZvo1) is based on 38X ONT data and 168X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial ONT 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, 37 regions totaling 25.084 Mb (with the largest being 4.542 Mb) were identified as haplotypic duplications and removed. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 3 haplotypic regions were removed, totaling 1.865Mb (with the largest being 1.318Mb). 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	793,837,058	791,974,536
GC %	41.35	41.35
Gaps/Gbp	26.45	35.35
Total gap bp	2,100	3,700
Scaffolds	37	30
Scaffold N50	34,205,391	34,205,391
Scaffold L50	11	11
Scaffold L90	21	21
Contigs	58	58
Contig N50	30,559,291	30,559,291
Contig L50	12	12
Contig L90	26	26
QV	44.3379	44.35
Kmer compl.	75.3072	75.2349
BUSCO sing.	98.8%	98.9%
BUSCO dupl.	0.4%	0.4%
BUSCO frag.	0.1%	0.1%
BUSCO miss.	0.7%	0.7%

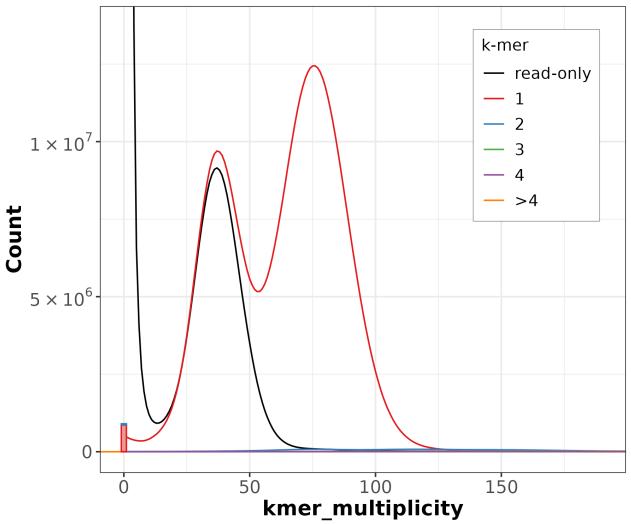
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

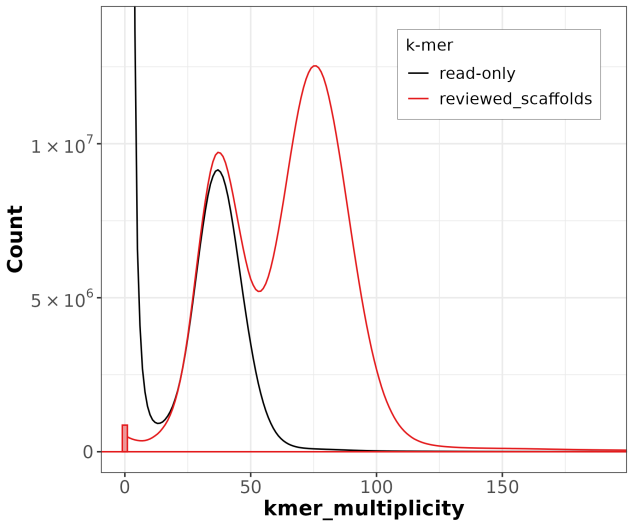


collapsed [LINK]

K-mer spectra of curated assembly

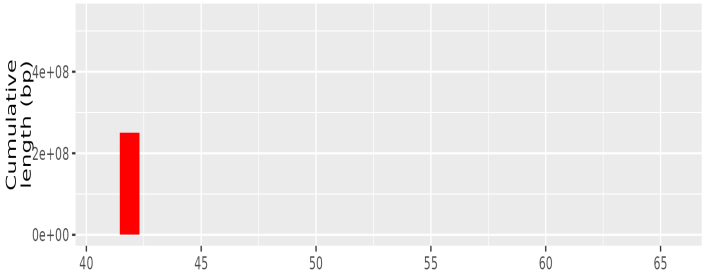


Distribution of k-mer counts per copy numbers found in asm

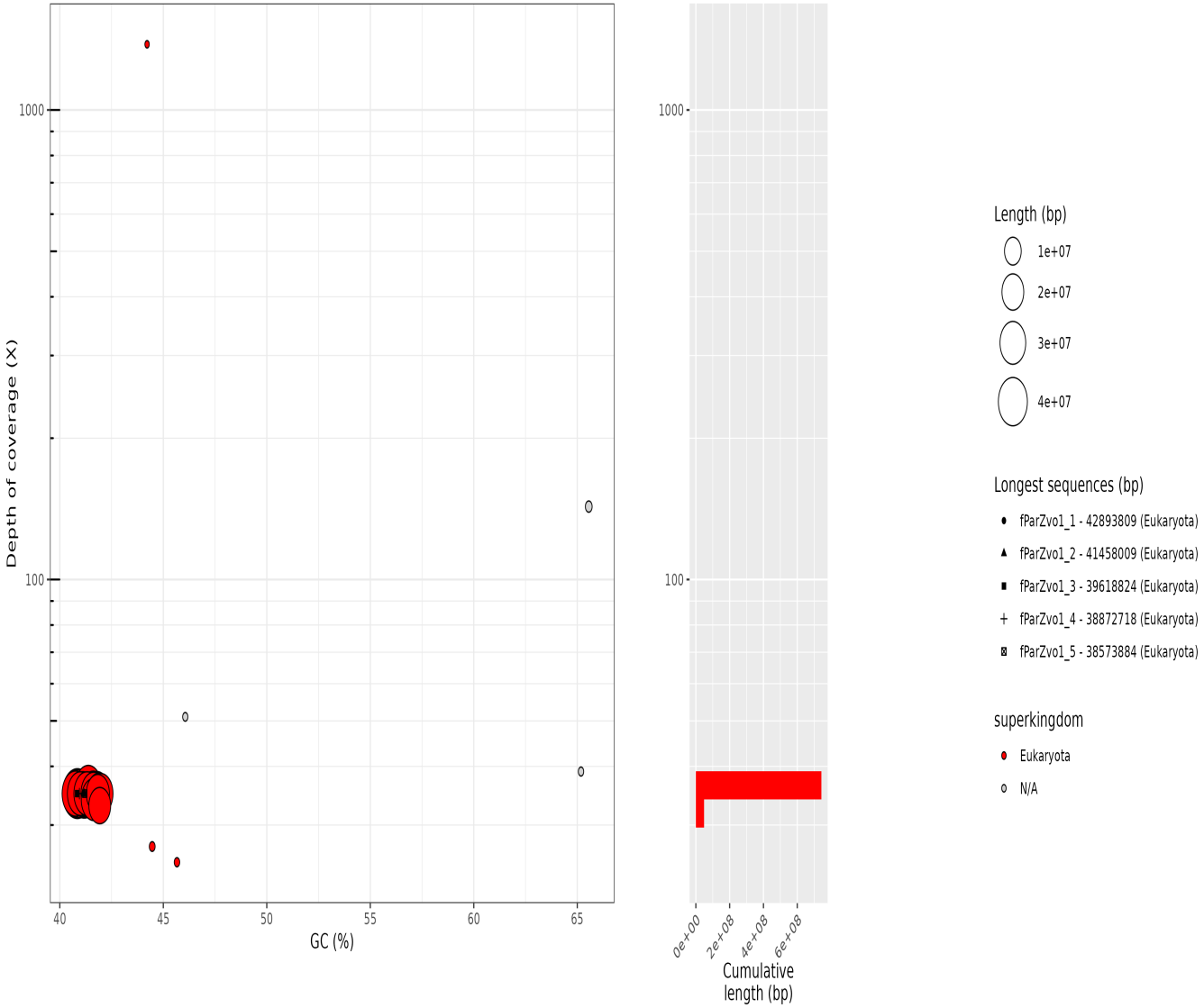


Distribution of k-mer counts coloured by their presence in reads/assemblies

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	38	168

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: 2026-03-26 09:08:23 CET