ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	202583
ToLID	**fAcaPal1**
Species	Acantholabrus palloni
Class	Actinopteri
Order	Labriformes

Genome Traits	Expected	Observed
Haploid size (bp)	747,024,719	759,627,933
Haploid Number	24 (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.Q48

The following metrics were automatically flagged as below EBP recommended standards or different from expected:

. Kmer completeness value is less than 90 for collapsed

Curator notes

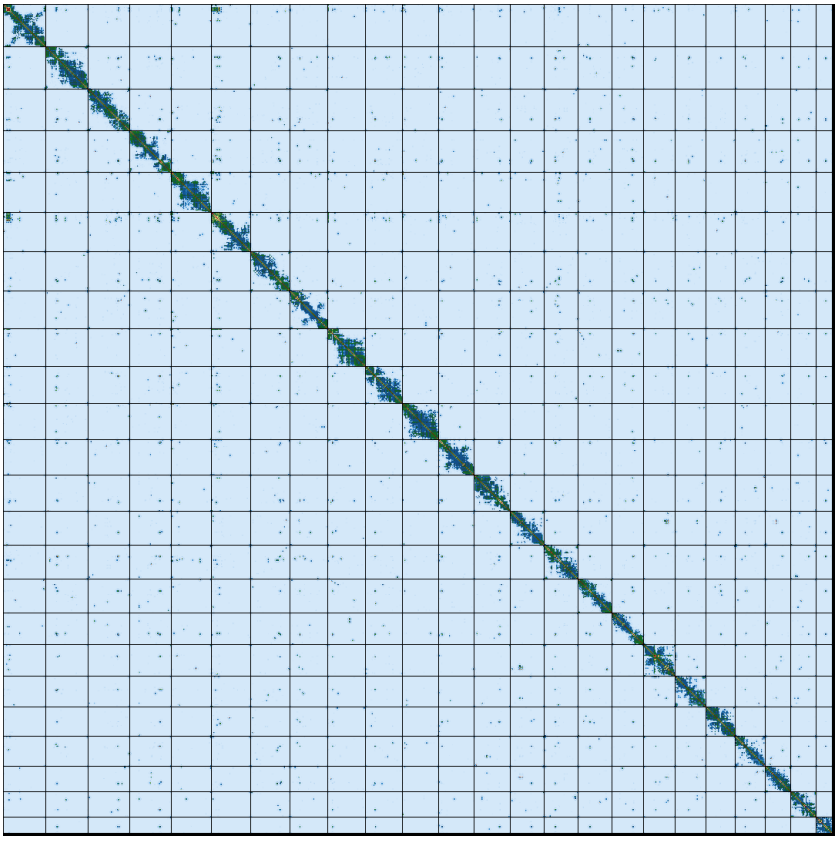
. Interventions/Gb: 39
. Contamination notes: ""
. Other observations: "The assembly of Acantholabrus palloni (fAcaPal1) is based on 36X PacBio data and 109X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial 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, 101 regions totaling 5.373 Mb (with the largest being 0.286 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 0.557Mb. The telomeric pattern TTAGGG was identified with TelFinder and used to generate the Pretext Telomeres tracks. 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	760,165,017	759,627,933
GC %	41.53	41.53
Gaps/Gbp	242.05	259.34
Total gap bp	18,400	21,600
Scaffolds	67	57
Scaffold N50	32,960,166	32,960,166
Scaffold L50	11	11
Scaffold L90	21	21
Contigs	251	254
Contig N50	25,137,971	25,137,971
Contig L50	13	13
Contig L90	38	39
QV	48.4719	48.4702
Kmer compl.	89.4359	89.4276
BUSCO sing.	99.1%	98.2%
BUSCO dupl.	0.5%	0.3%
BUSCO frag.	0.0%	0.8%
BUSCO miss.	0.4%	0.7%

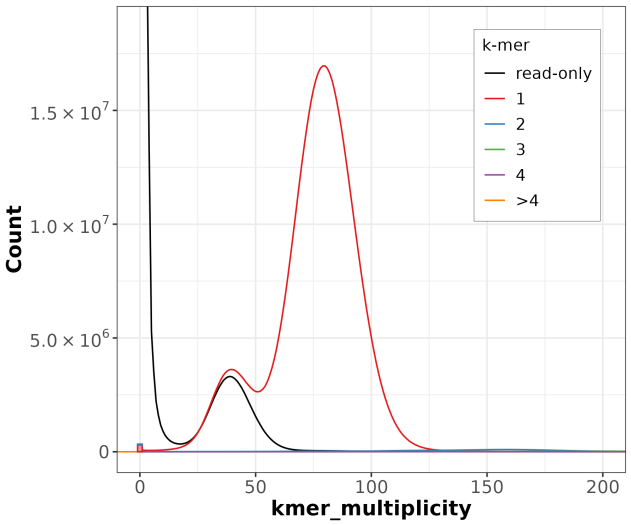
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 6.1.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

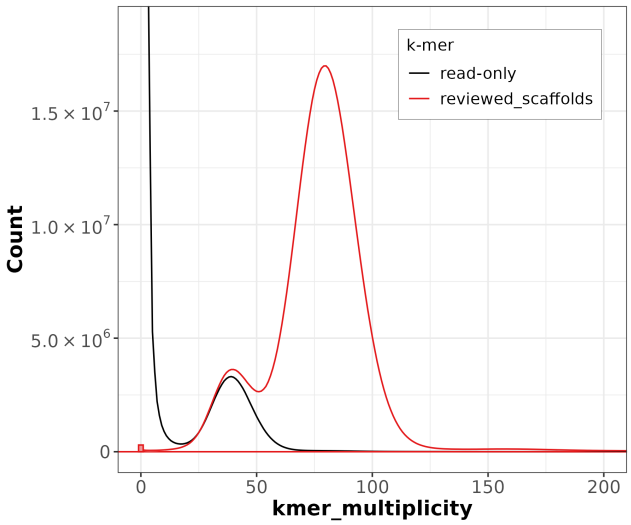


collapsed [LINK]

K-mer spectra of curated assembly

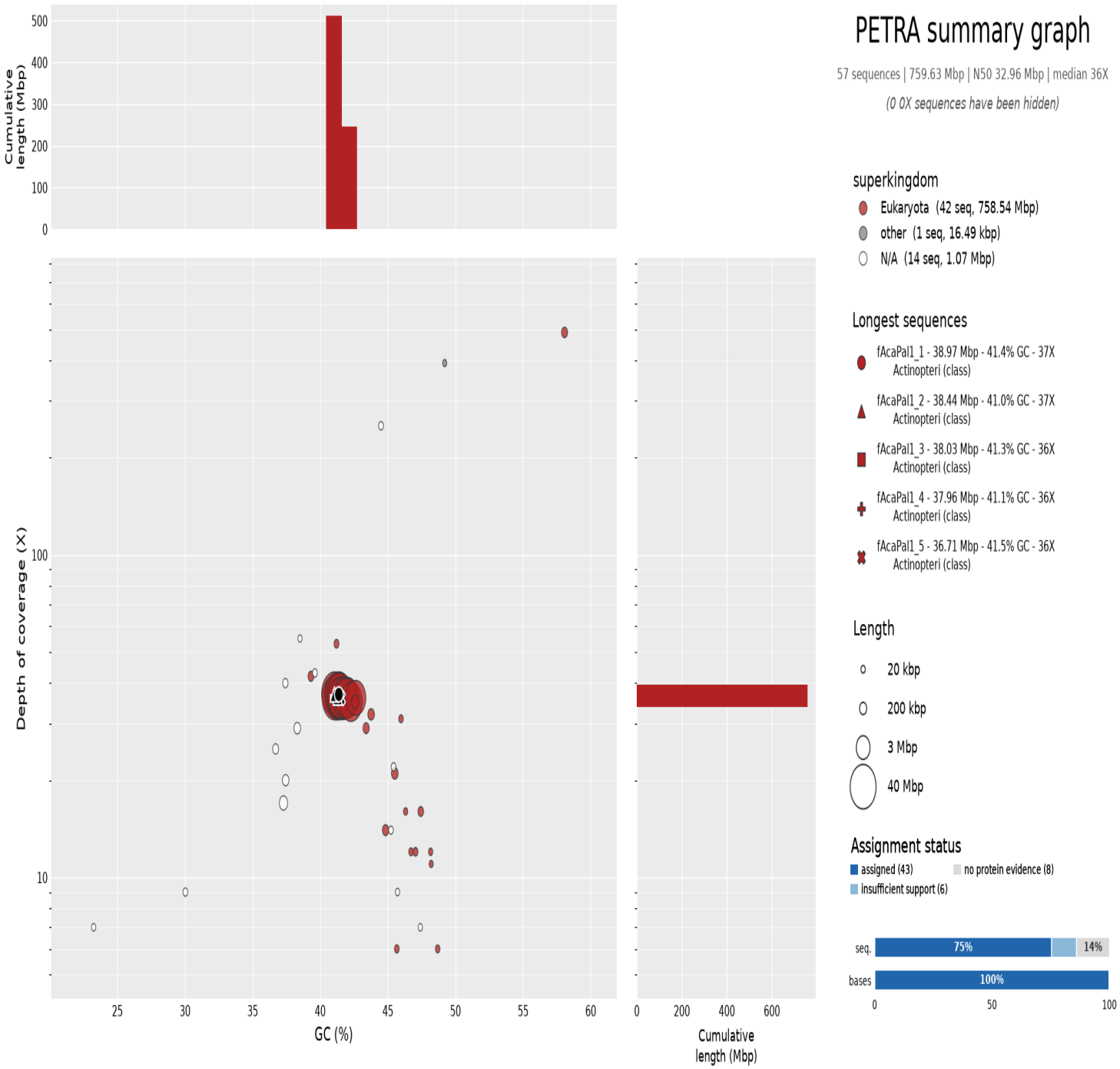


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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	36	109

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:* 0.2.4
  |_ *key param:* NA
- **PretextView**
  |_ *ver:* 1.0.7
  |_ *key param:* NA

Submitter: Sophie Layac
Affiliation: Genoscope

Date and time: 2026-09-16 10:58:33 CEST