ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	430084
ToLID	**fTriLyr1**
Species	Trigla lyra
Class	Actinopteri
Order	Perciformes

Genome Traits	Expected	Observed
Haploid size (bp)	636,266,564	652,883,168
Haploid Number	24 (source: ancestor)	13
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q46

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

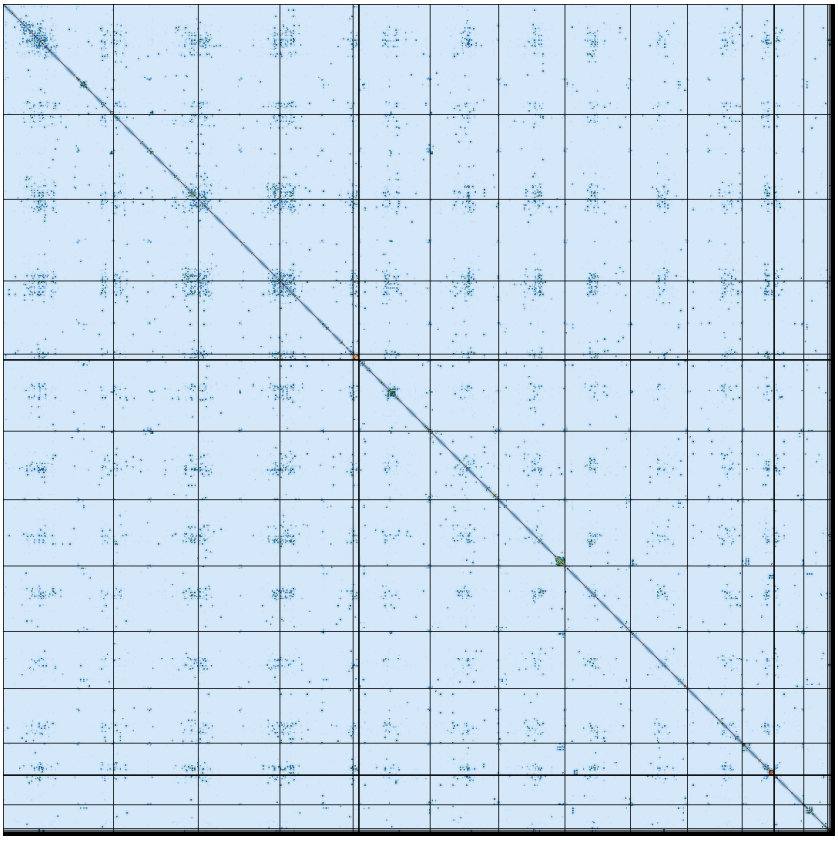
. Interventions/Gb: 21
. Contamination notes: ""
. Other observations: "The assembly of Trigla lyra (fTriLyr1) is based on 54X PacBio data and 448X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial PacBio assembly 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, 128 regions totaling 8.972 Mb (with the largest being 0.93 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, 4 haplotypic regions were removed, totaling 2.099Mb (with the largest being 1.442Mb). 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	655,005,517	652,883,168
GC %	42.37	42.38
Gaps/Gbp	280.91	292.55
Total gap bp	18,400	20,100
Scaffolds	112	102
Scaffold N50	55,319,458	55,319,458
Scaffold L50	5	5
Scaffold L90	11	11
Contigs	296	293
Contig N50	12,264,985	12,264,985
Contig L50	18	18
Contig L90	76	76
QV	46.5254	46.5412
Kmer compl.	88.8045	88.7824
BUSCO sing.	98.8%	99.0%
BUSCO dupl.	0.7%	0.4%
BUSCO frag.	0.1%	0.1%
BUSCO miss.	0.5%	0.5%

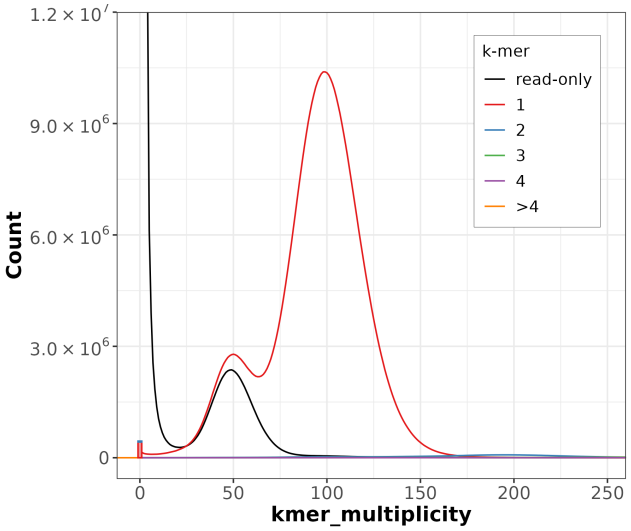
BUSCO: 6.1.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

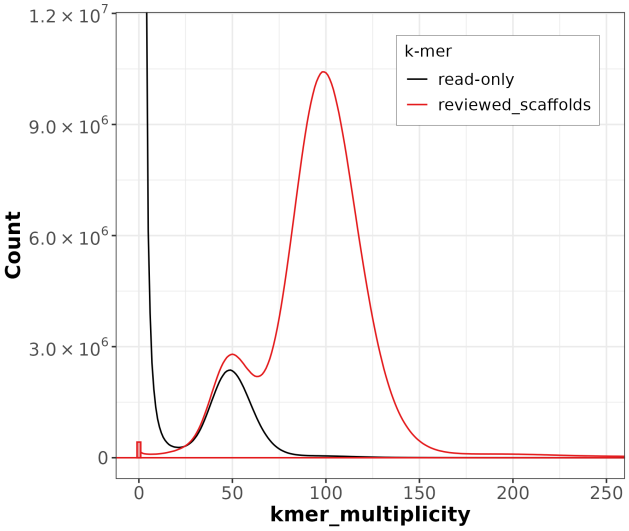


collapsed [LINK]

K-mer spectra of curated assembly

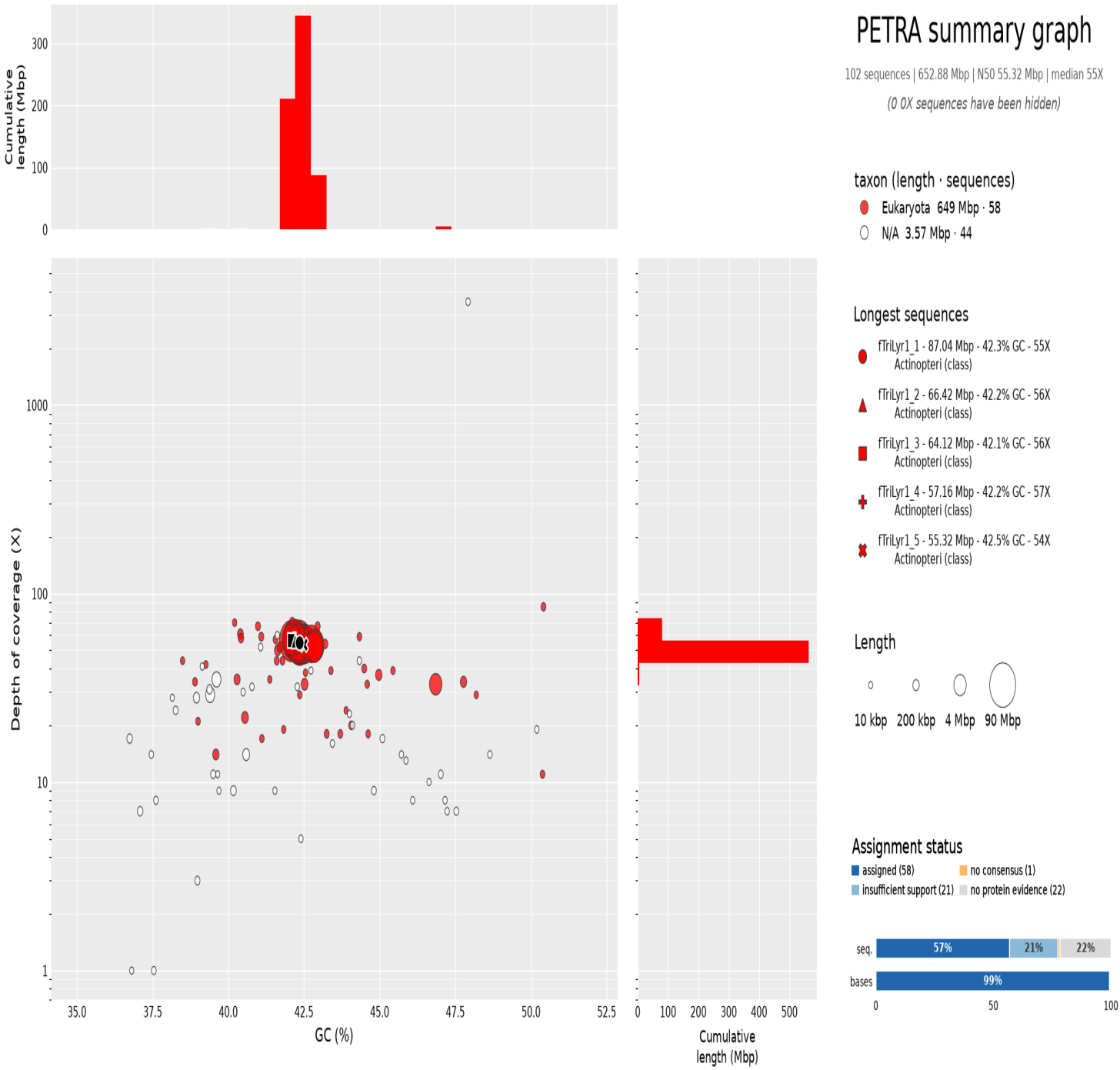


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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	55	450

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: Arnaud Couloux
Affiliation: Genoscope

Date and time: 2026-10-08 06:30:38 CEST