

ERGA Assembly Report

v24.10.15

Tags: Other[INVALID TAG]

TxID	27695
ToLID	fProLin1_hap1
Species	Prochilodus lineatus
Class	Actinopteri
Order	Characiformes

Genome Traits	Expected	Observed
Haploid size (bp)	1,179,458,285	1,448,379,579
Haploid Number	27 (source: direct)	27
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q57

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

- . Observed Haploid size (bp) has >20% difference with Expected
- . Kmer completeness value is less than 90 for collapsed

Curator notes

- . Interventions/Gb: 23
- . Contamination notes: ""
- . Other observations: "The assembly of Prochilodus lineatus (fProLin1) was generated as part of the Vertebrate Genomes Project (<https://vertebrategenomesproject.org/>). Both haplotypes were analyzed and manually improved using Pretext. "

Quality metrics table

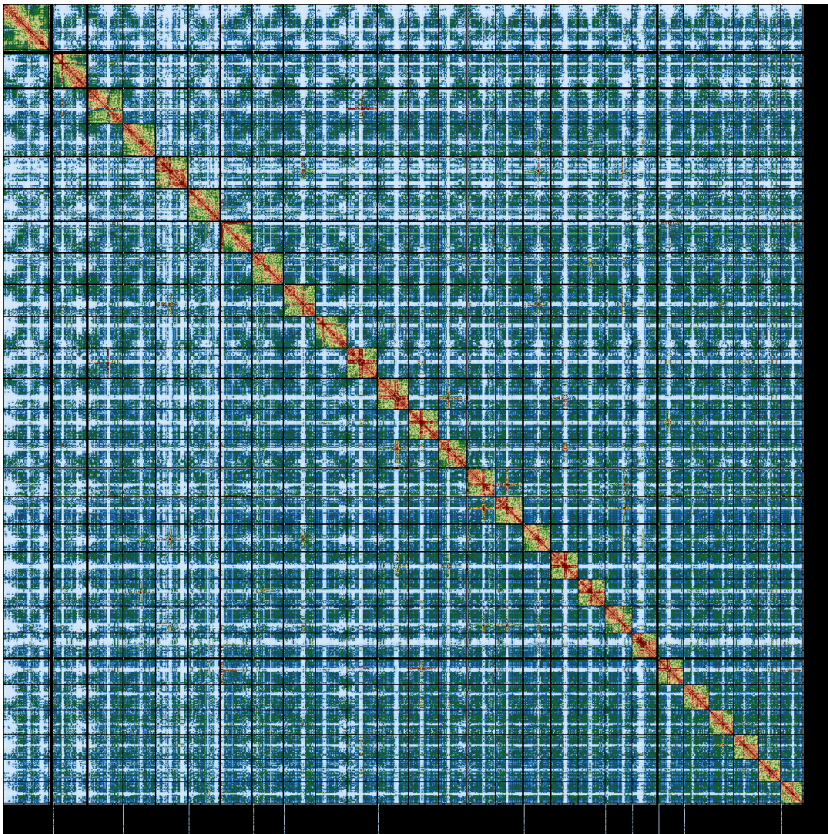
Metrics	Pre-curation collapsed	Curated collapsed
Total bp	1,433,312,037	1,448,379,579
GC %	42.04	42.02
Gaps/Gbp	673.96	695.95
Total gap bp	193,200	201,600
Scaffolds	853	811
Scaffold N50	47,300,995	51,531,475
Scaffold L50	14	13
Scaffold L90	30	26
Contigs	1,819	1,819
Contig N50	6,189,296	5,995,764
Contig L50	59	60
Contig L90	481	490
QV	57.1482	57.3558
Kmer compl.	75.2466	75.2307
BUSCO sing.	94.0%	96.7%
BUSCO dupl.	1.7%	1.9%
BUSCO frag.	1.3%	0.2%
BUSCO miss.	3.1%	1.2%

Warning! BUSCO versions or lineage datasets are not the same across results:

BUSCO: 5.8.2 (euk_genome_met, metaeuk) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

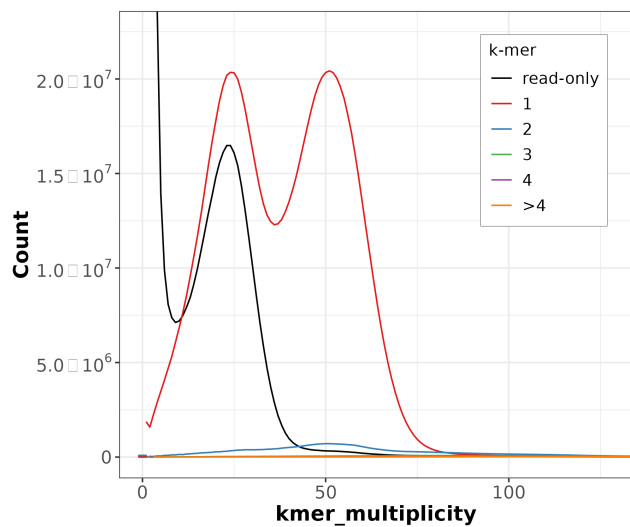
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

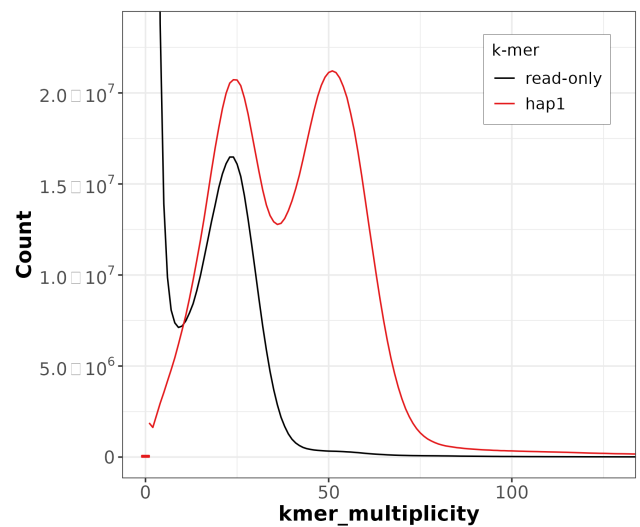


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

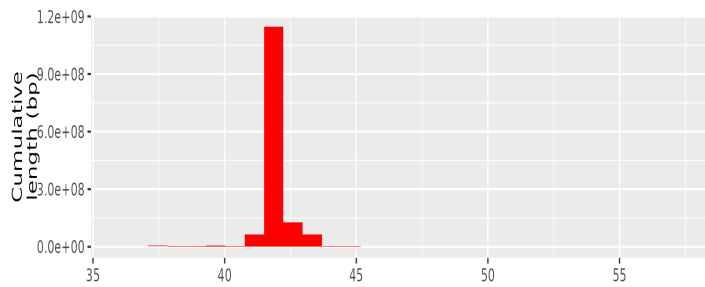


Distribution of k-mer counts per copy numbers found in asm



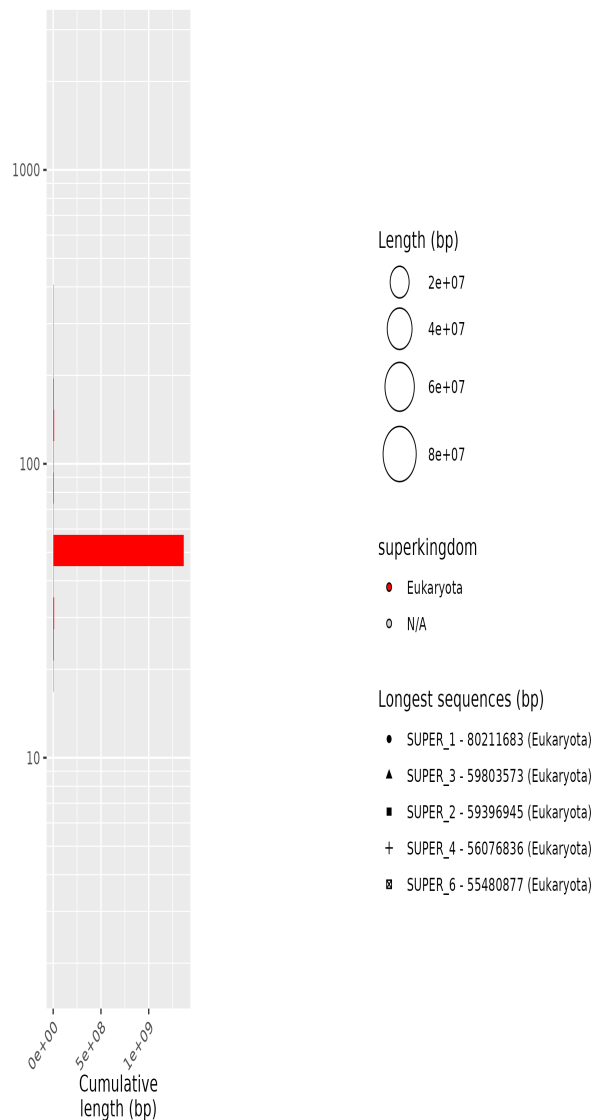
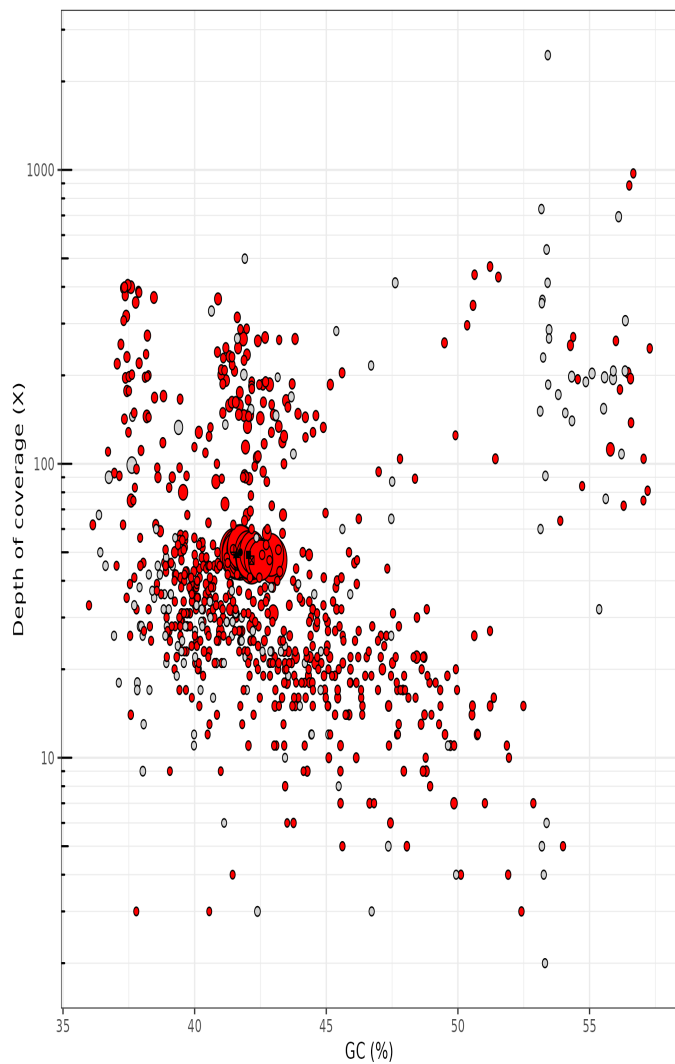
Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph

(1 0X contig has been hidden)



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	NA	NA

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: Benjamin Istace

Affiliation: Genoscope

Date and time: 2025-12-04 05:10:25 CET