ERGA Assembly Report

v24.10.15

Tags: Other[INVALID TAG]

TxID	28791
ToLID	**fCypLut1_hap2**
Species	Cyprinella lutrensis
Class	Actinopteri
Order	Cypriniformes

Genome Traits	Expected	Observed
Haploid size (bp)	802,987,641	892,865,398
Haploid Number	25 (source: direct)	25
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:

. Kmer completeness value is less than 90 for collapsed

Curator notes

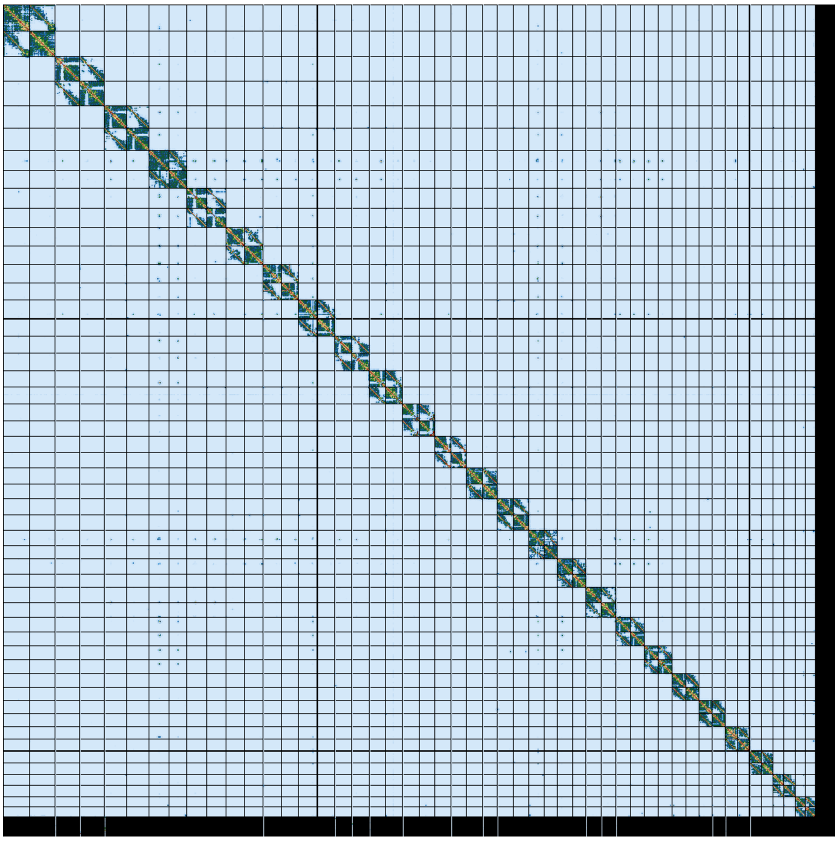
. Interventions/Gb: 154
. Contamination notes: ""
. Other observations: "The assembly of Cyprinella lutrensis (fCypLut1) was generated as part of the Vertebrate Genomes Project (https://vertebrategenomesproject.org/). Both haplotypes were analyzed and manually improved using Pretext. "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	914,424,475	892,865,398
GC %	38.58	38.55
Gaps/Gbp	874.87	919.51
Total gap bp	160,000	164,200
Scaffolds	404	394
Scaffold N50	36,156,036	34,822,651
Scaffold L50	10	11
Scaffold L90	22	23
Contigs	1,204	1,215
Contig N50	1,165,438	1,141,313
Contig L50	208	216
Contig L90	656	663
QV	57.6084	57.5851
Kmer compl.	74.7533	74.0873
BUSCO sing.	94.9%	97.6%
BUSCO dupl.	0.5%	0.5%
BUSCO frag.	1.2%	0.2%
BUSCO miss.	3.5%	1.6%

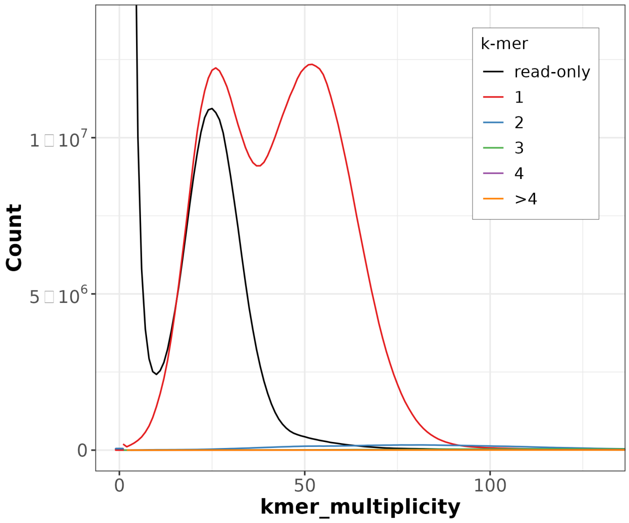
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 5.8.2 (euk_genome_met, metaeuk) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

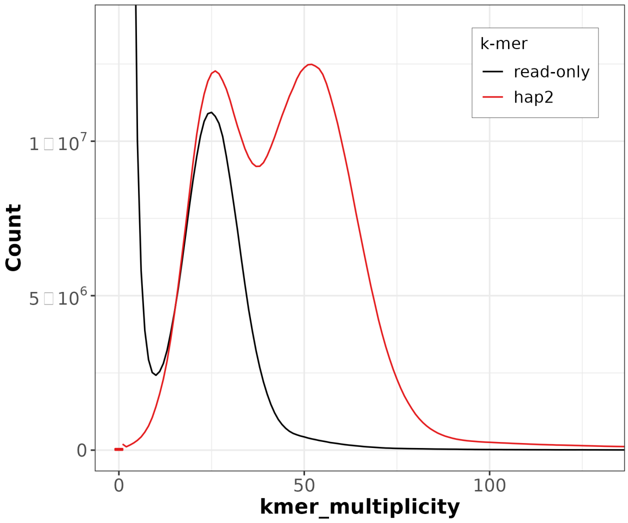


collapsed [LINK]

K-mer spectra of curated assembly

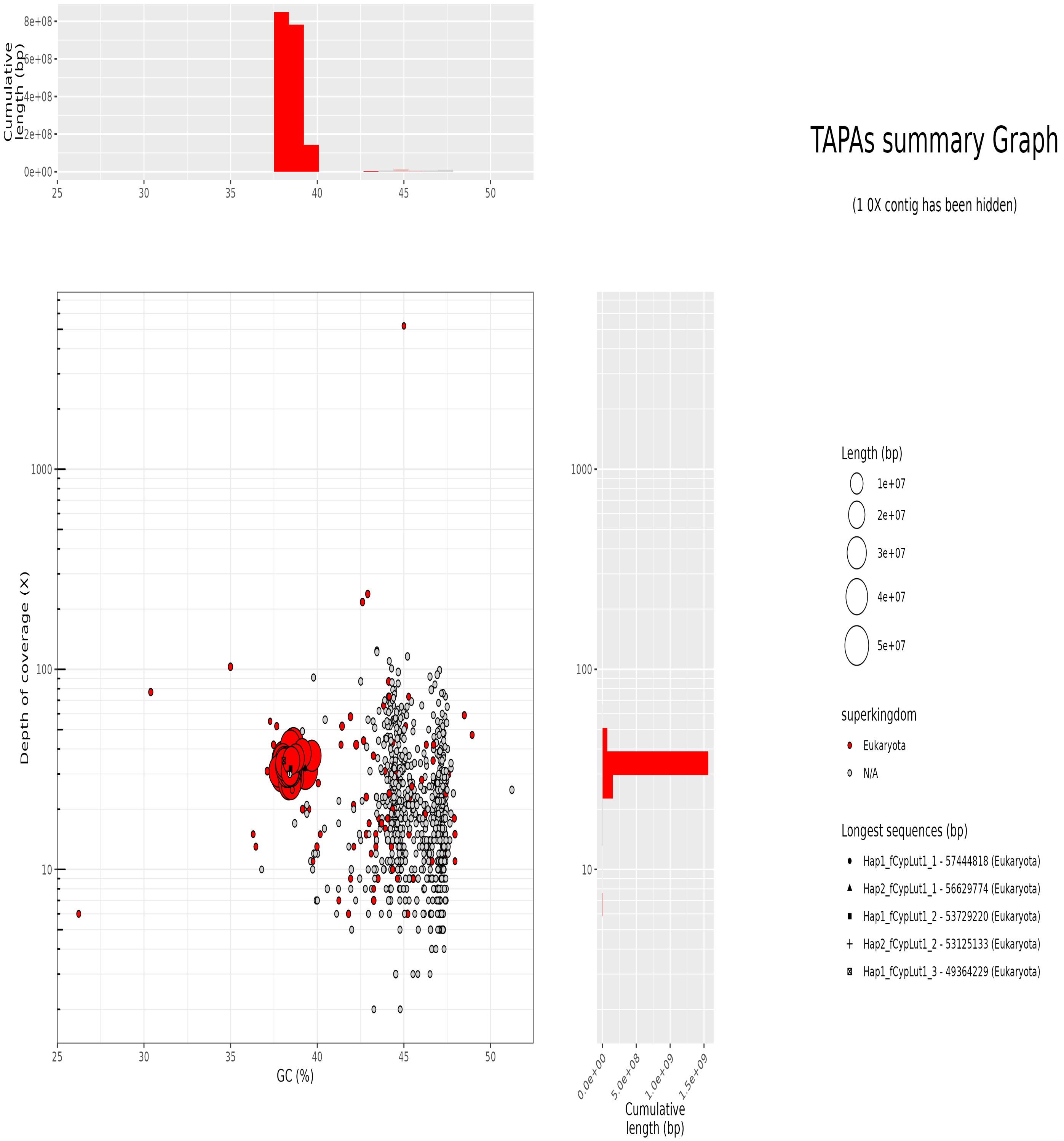


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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: 2025-10-20 09:00:11 CEST