ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	185886
ToLID	**fPteVol1**
Species	Pterois volitans
Class	Actinopteri
Order	Perciformes

Genome Traits	Expected	Observed
Haploid size (bp)	877,590,350	908,186,800
Haploid Number	21 (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.Q50

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

. Observed Haploid Number is different from Expected

Curator notes

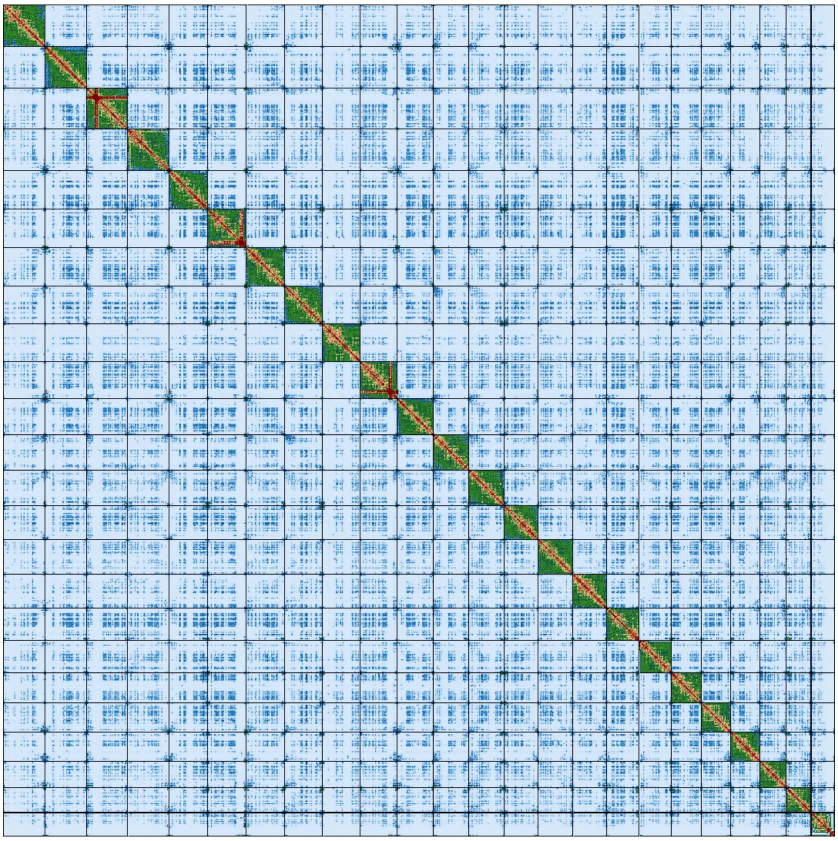
. Interventions/Gb: 17
. Contamination notes: ""
. Other observations: "The assembly of Pterois volitans (fPteVol1) is based on 59X PacBio data and 189X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial PacBio 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, 34 regions totaling 2.103 Mb (with the largest being 0.215 Mb) were identified as haplotypic duplications and removed. During manual curation, 3 haplotypic regions were removed, totaling 4.554Mb (with the largest being 1.967Mb). 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	912,721,347	908,186,800
GC %	40.8	40.8
Gaps/Gbp	74.5	78.18
Total gap bp	6,800	7,800
Scaffolds	36	36
Scaffold N50	38,453,888	39,614,986
Scaffold L50	11	11
Scaffold L90	22	21
Contigs	104	107
Contig N50	35,254,685	35,254,685
Contig L50	12	12
Contig L90	24	24
QV	50.0528	50.0479
Kmer compl.	91.2604	91.1857
BUSCO sing.	99.1%	99.5%
BUSCO dupl.	0.5%	0.2%
BUSCO frag.	0.0%	0.0%
BUSCO miss.	0.3%	0.3%

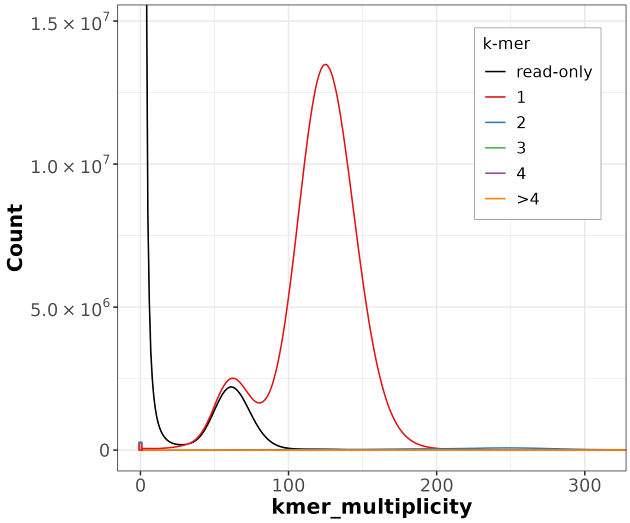
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

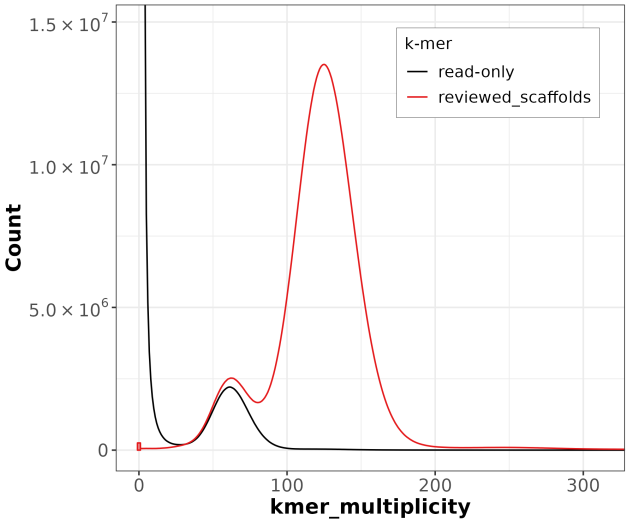


collapsed [LINK]

K-mer spectra of curated assembly

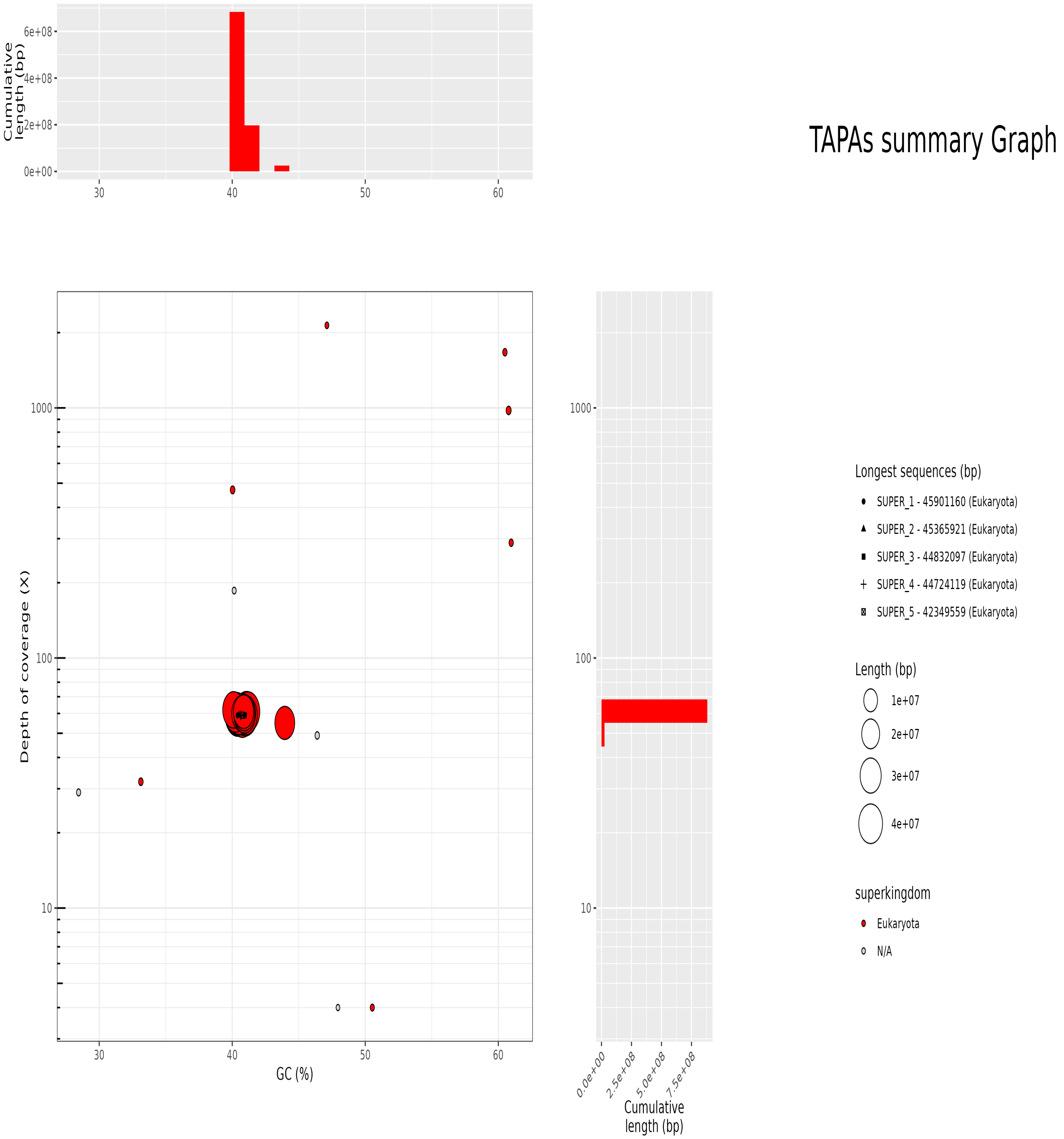


Distribution of k-mer counts per copy numbers found in asm



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	59	189

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: Sophie Layac
Affiliation: Genoscope

Date and time: 2025-11-28 02:04:03 CET