

# ERGA Assembly Report

v24.04.03\_beta

Tags: ATLASea[INVALID TAG]

TxID	357811
ToLID	<b>fApoImb</b>
Species	Apogon imberbis
Class	Actinopteri
Order	Kurtiformes

Genome Traits	Expected	Observed
Haploid size (bp)	794,137,431	819,133,141
Haploid Number	24 (source: ancestor)	18
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:

- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for collapsed

### Curator notes

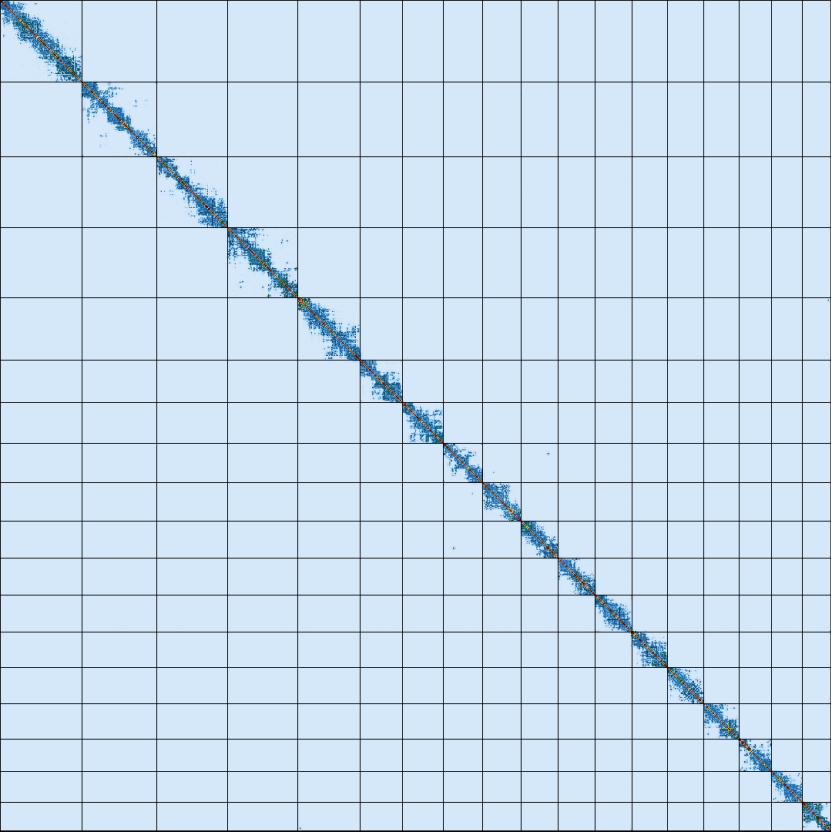
- . Interventions/Gb: 4
- . Contamination notes: "One bacterial sequence of 74 Kb detected. It was removed before scaffolding."
- . Other observations: "Manual curation was minimal because we already had the chromosomes after scaffolding with yahs. The post-curation contamination screening plot corresponds to the assembly before bacterial sequences are removed."

## Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	819,152,605	819,133,141
GC %	39.9	39.9
Gaps/Gbp	3.66	6.1
Total gap bp	300	800
Scaffolds	38	33
Scaffold N50	40,310,727	40,310,727
Scaffold L50	7	7
Scaffold L90	16	16
Contigs	41	38
Contig N50	40,310,727	40,310,727
Contig L50	7	7
Contig L90	16	16
QV	48.2956	48.3336
Kmer compl.	82.7357	82.7356
BUSCO sing.	98.3%	98.3%
BUSCO dupl.	0.6%	0.6%
BUSCO frag.	0.2%	0.2%
BUSCO miss.	0.9%	0.9%

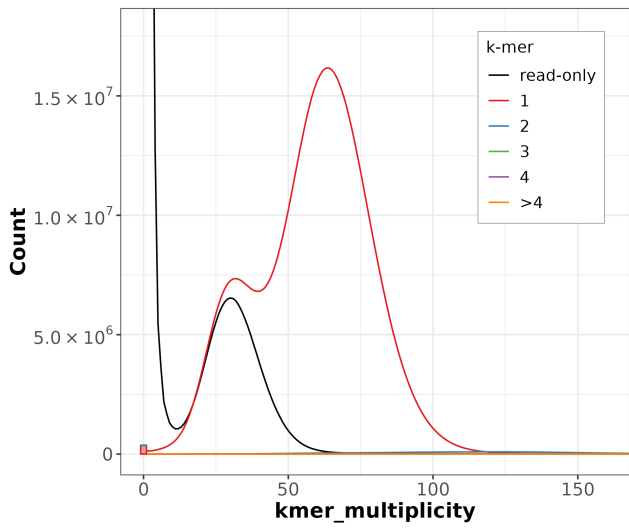
BUSCO 5.4.3 Lineage: actinopterygii\_odb10 (genomes:26, BUSCOs:3640)

# HiC contact map of curated assembly

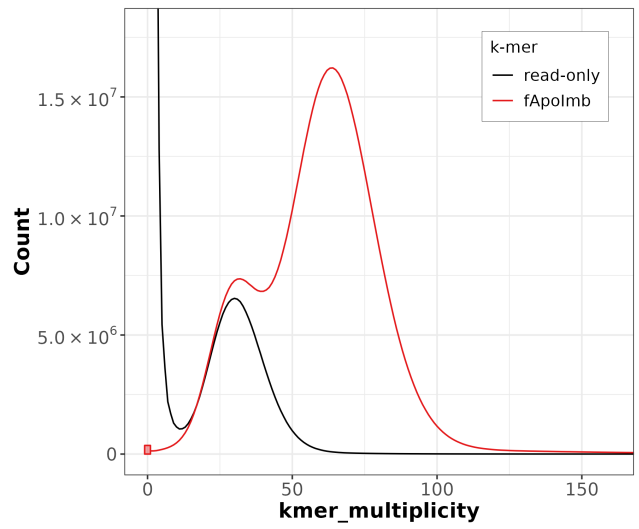


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

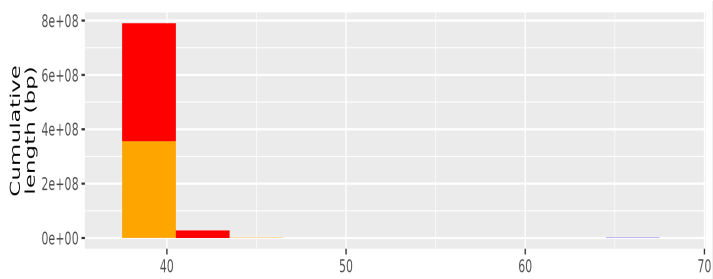


Distribution of k-mer counts per copy numbers found in asm



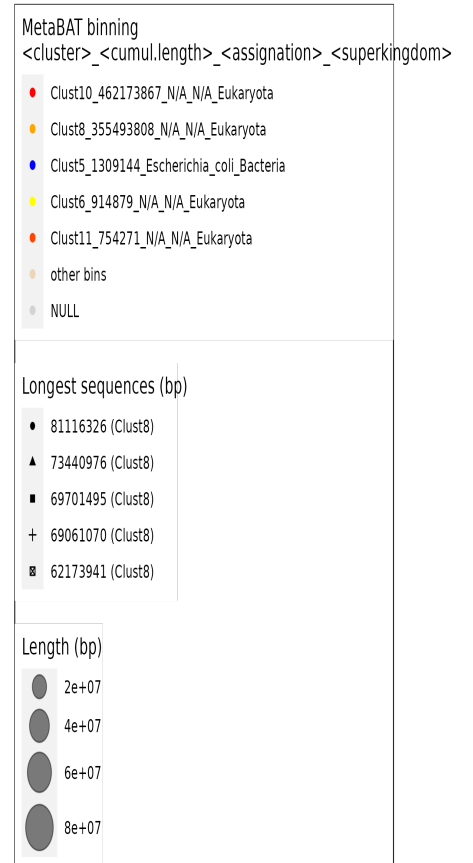
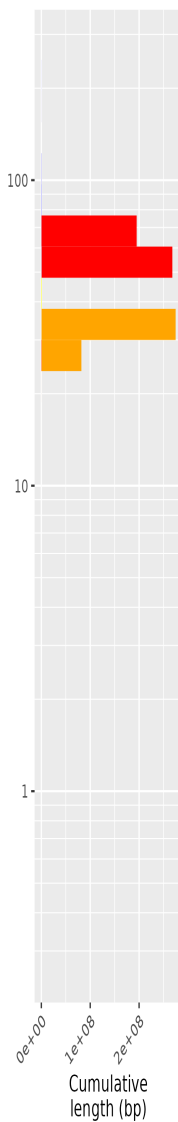
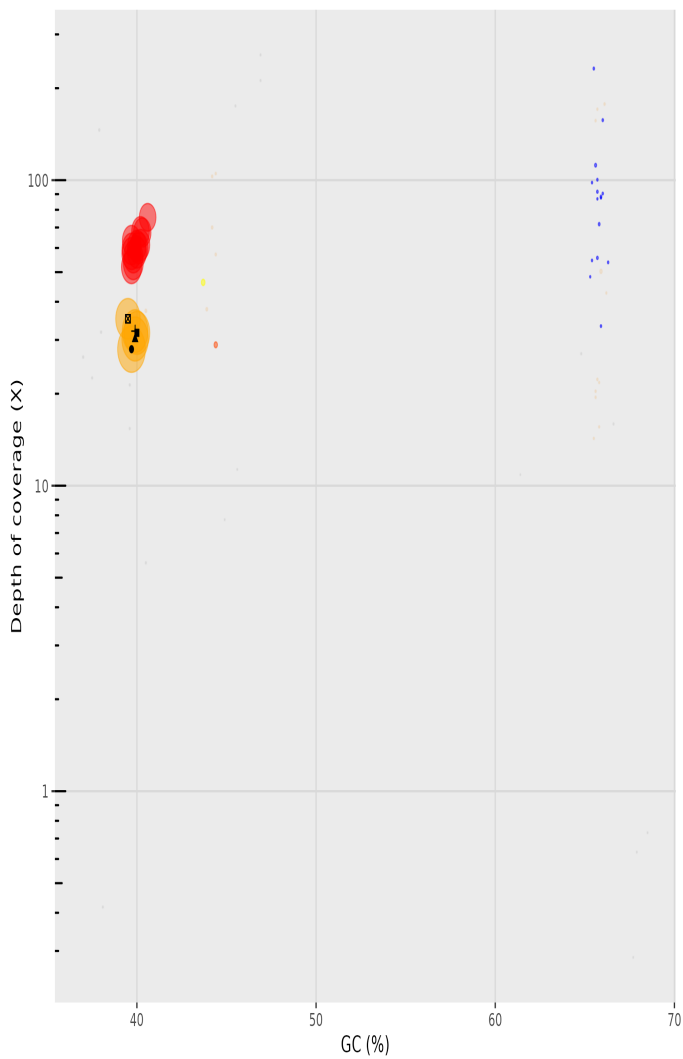
Distribution of k-mer counts coloured by their presence in reads/assemblies

# Post-curation contamination screening



## Decap Summary Graph

(12 0X contigs have 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	PACBIO Hifi	ONT
Coverage	93X	128X

# 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

- **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

Submitter: Adama Ndar  
Affiliation: Genoscope

Date and time: 2024-07-30 15:42:16 CEST