

ERGA Assembly Report

v24.10.15

Tags: ATLASEa[INVALID TAG]

TxID	150449
ToLID	fAulMac2
Species	Aulostomus maculatus
Class	Actinopteri
Order	Syngnathiiformes

Genome Traits	Expected	Observed
Haploid size (bp)	678,435,034	722,183,757
Haploid Number	22 (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.Q46

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

- . Observed Haploid Number is different from Expected

Curator notes

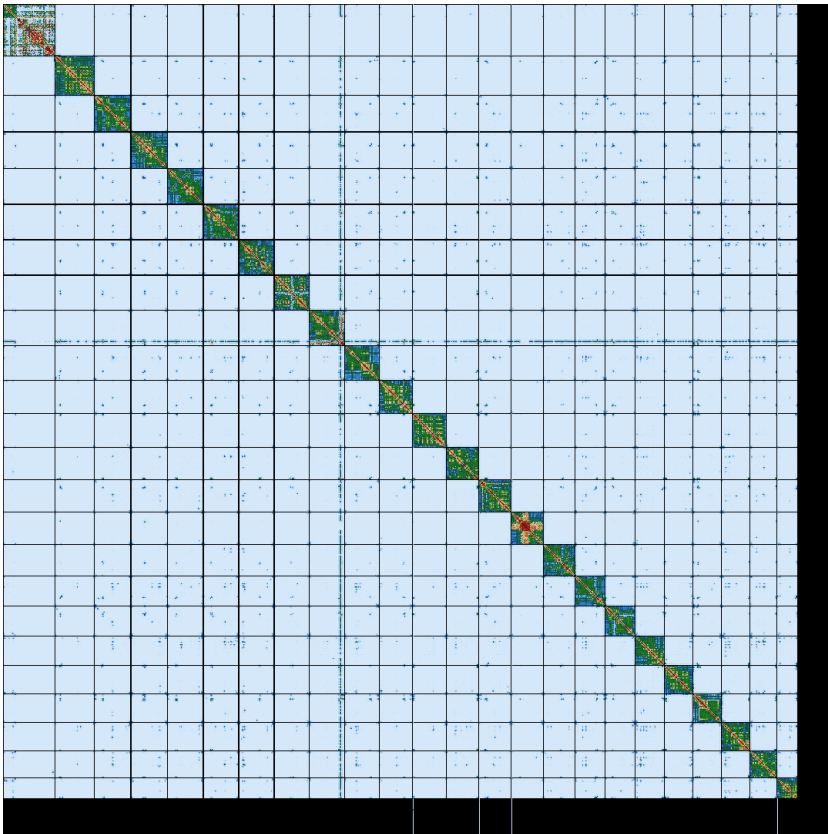
- . Interventions/Gb: 34
- . Contamination notes: ""
- . Other observations: "The assembly of Aulostomus macalus is based on 42X PACBIO data and 144X of Arima Hi-C data generated as part of the ATLASEa programme (<https://www.atlasea.fr>). The assembly process included the following steps: initial ONT assembly generation with Hifiiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. No contaminant sequences were detected. In total, 147 regions totaling 20.416 Mb (with the largest being 4.849 Mb) were identified as haplotypic duplications and removed. The mitochondrial genome was assembled using OATK. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 1 haplotypic region and 6 contaminant sequences were removed, totaling 4.911 Mb and 0.482 Mb, respectively (with the largest being 4.911 Mb and 0.105 Mb). 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	727,567,113	722,183,757
GC %	42.91	42.91
Gaps/Gbp	46.73	62.31
Total gap bp	3,400	5,900
Scaffolds	477	459
Scaffold N50	27,869,090	28,775,080
Scaffold L50	12	12
Scaffold L90	24	23
Contigs	511	504
Contig N50	20,072,396	19,647,333
Contig L50	16	16
Contig L90	39	39
QV	46.6724	46.6844
Kmer compl.	90.2554	90.1743
BUSCO sing.	96.1%	96.7%
BUSCO dupl.	0.8%	0.2%
BUSCO frag.	1.0%	1.0%
BUSCO miss.	2.1%	2.1%

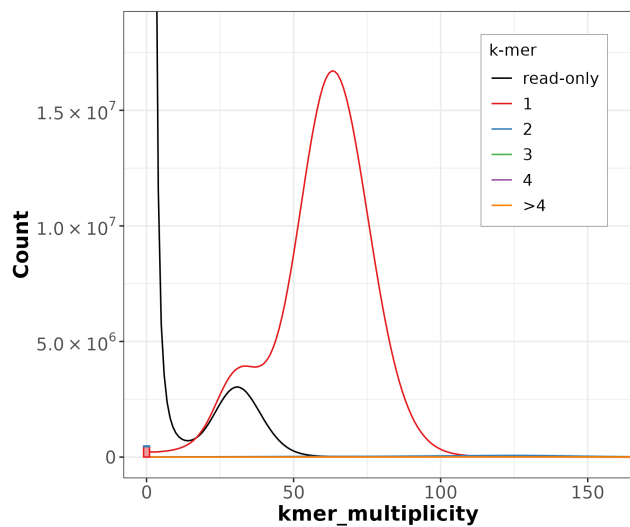
BUSCO: 5.8.2 (euk_genome_met, metaeuk) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

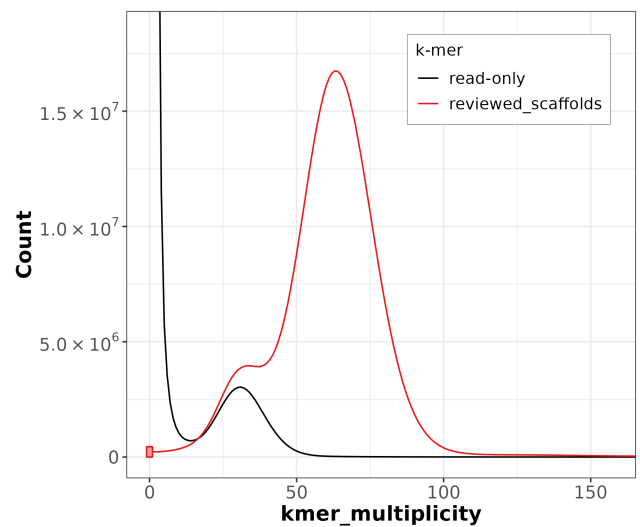


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

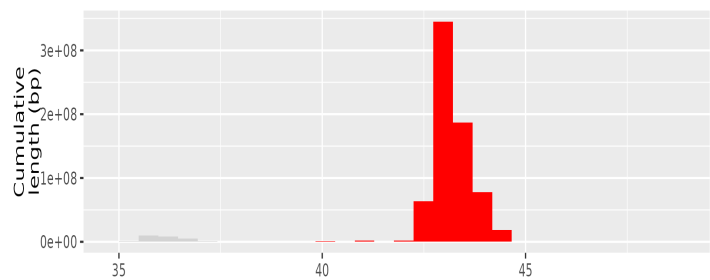


Distribution of k-mer counts per copy numbers found in asm

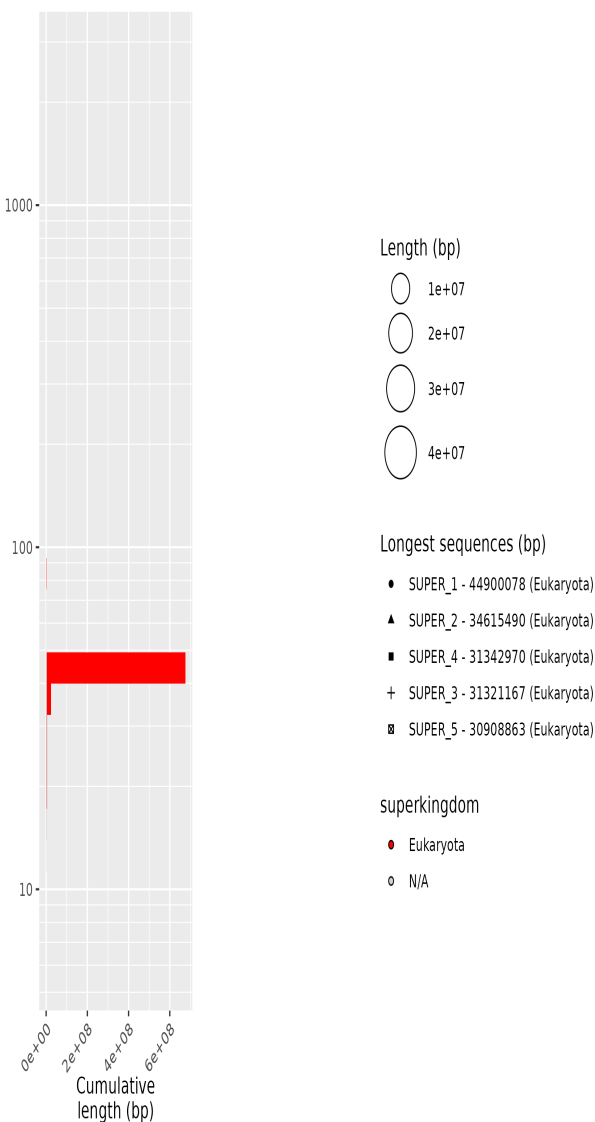
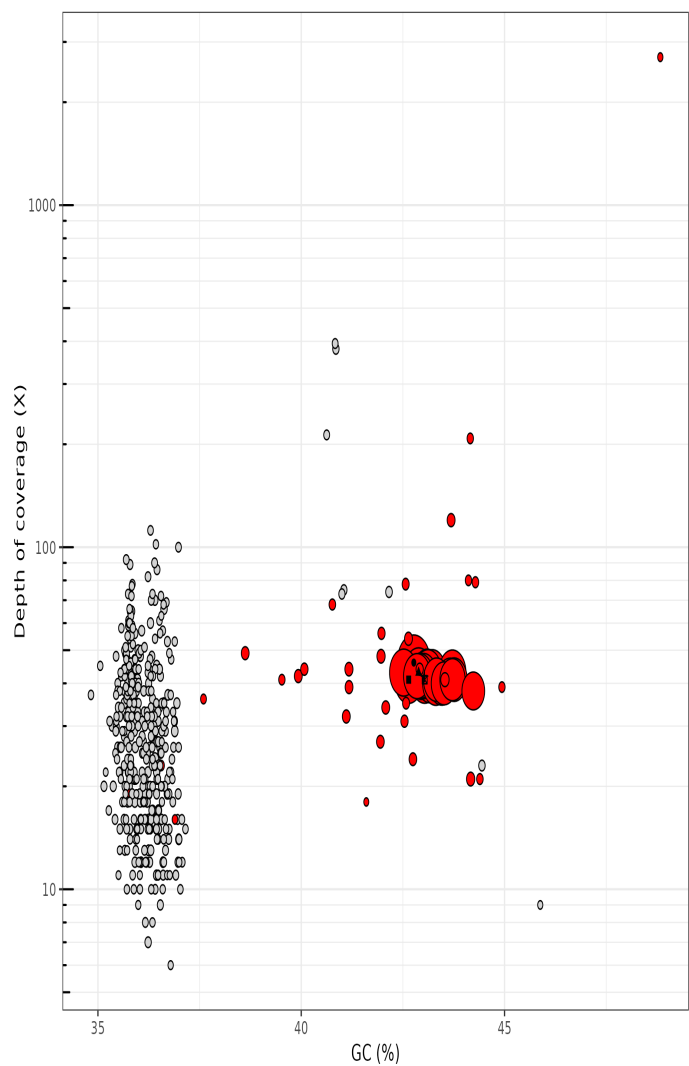


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



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	Arima
Coverage	42	144

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: Adama Ndar

Affiliation: Genoscope

Date and time: 2025-09-01 00:48:15 CEST