

# ERGA Assembly Report

v24.10.15

Tags: Other[INVALID TAG]

TxID	8097
ToLID	<b>fMyoScol_h1</b>
Species	<i>Myoxocephalus scorpius</i>
Class	Actinopteri
Order	Perciformes

Genome Traits	Expected	Observed
Haploid size (bp)	611,232,883	726,603,694
Haploid Number	19 (source: direct)	19
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

## EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q53

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

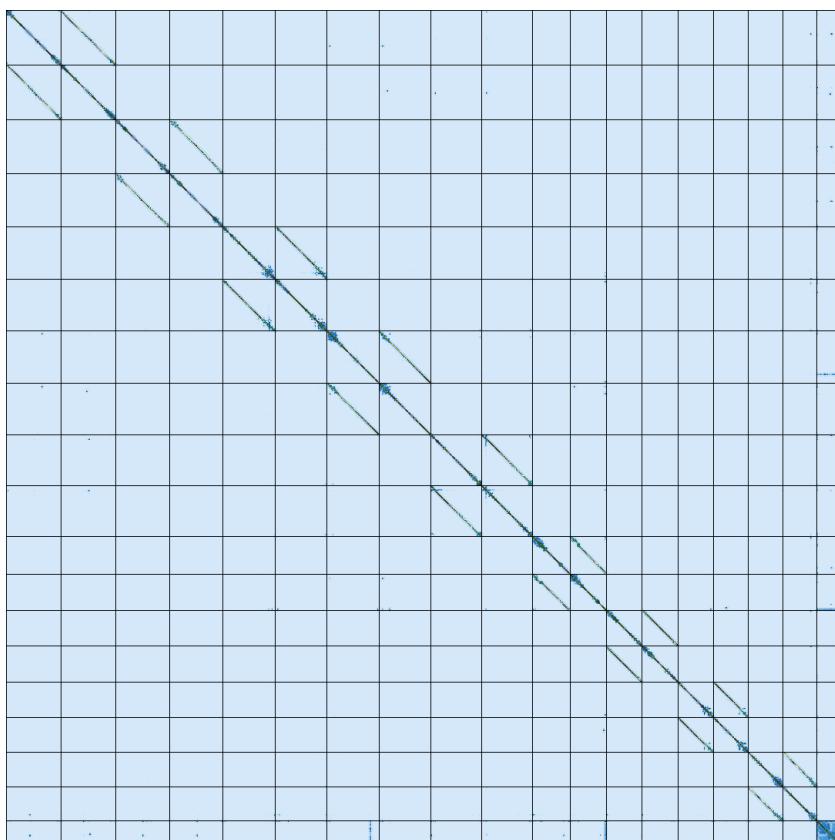
- . Interventions/Gb: 120
- . Contamination notes: ""
- . Other observations: "The assembly of *Myoxocephalus scorpius* (fMyoScol.1) was generated as part of the Vertebrate Genomes Project (<https://vertebratogenomesproject.org/>). Both haplotypes were analyzed and manually improved using Pretext. Chromosome 18 contains a large region (0-8.4 Mb) where the organization of contigs is uncertain. "

## Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	715,405,851	726,603,694
GC %	42.51	42.51
Gaps/Gbp	680.73	750.07
Total gap bp	97,400	109,000
Scaffolds	97	60
Scaffold N50	36,375,788	37,154,042
Scaffold L50	8	8
Scaffold L90	17	17
Contigs	584	605
Contig N50	3,080,030	2,948,135
Contig L50	63	73
Contig L90	259	282
QV	53.9733	53.9826
Kmer compl.	89.2493	89.3927
BUSCO sing.	96.1%	96.5%
BUSCO dupl.	0.4%	0.4%
BUSCO frag.	0.8%	0.7%
BUSCO miss.	2.7%	2.4%

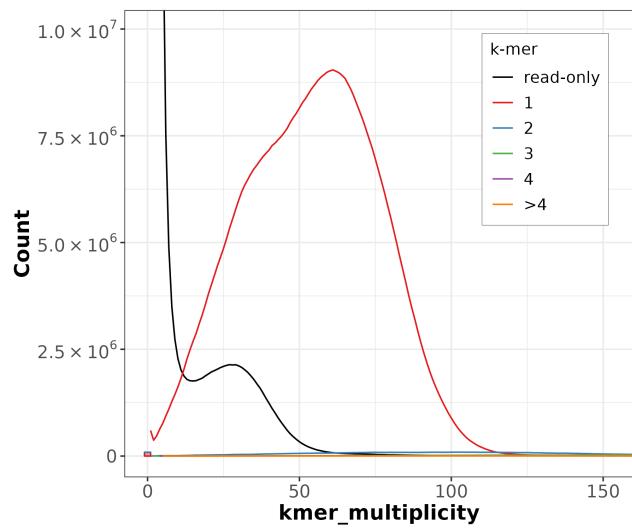
BUSCO: 5.8.2 (euk\_genome\_met, metaeuk) / Lineage: actinopterygii\_odb12 (genomes:75, BUSCOs:7207)

# HiC contact map of curated assembly

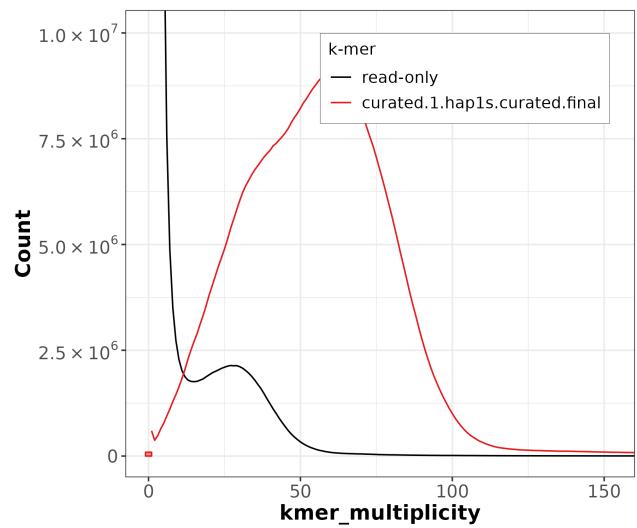


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

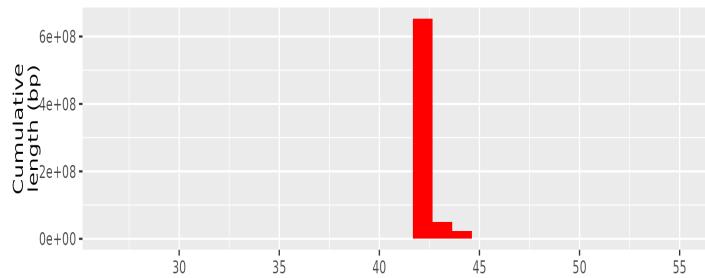


Distribution of k-mer counts per copy numbers found in asm

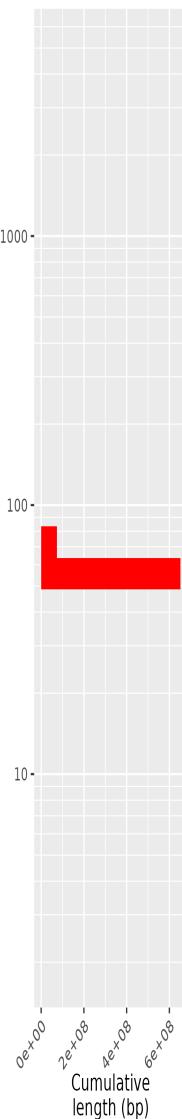
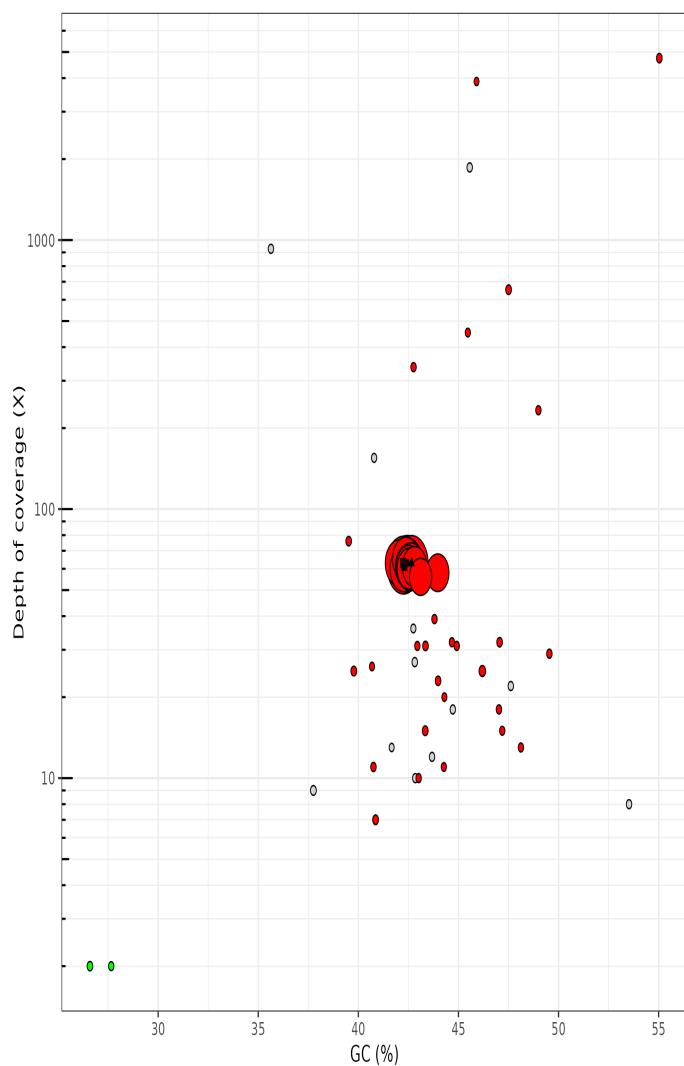


Distribution of k-mer counts coloured by their presence in reads/assemblies

# Post-curation contamination screening



TAPAs summary Graph



superkingdom

- Bacteria
- Eukaryota
- N/A

Length (bp)

- 1e+07
- 2e+07
- 3e+07
- 4e+07
- 5e+07

Longest sequences (bp)

- fMyoSco1\_1 - 57335687 (Eukaryota)
- ▲ fMyoSco1\_2 - 56670893 (Eukaryota)
- fMyoSco1\_3 - 55017562 (Eukaryota)
- + fMyoSco1\_4 - 54655007 (Eukaryota)
- ☒ fMyoSco1\_5 - 53230495 (Eukaryota)

**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	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: Jean-Marc Aury  
Affiliation: Genoscope

Date and time: 2025-08-04 16:32:38 CEST