

# ERGA Assembly Report

v24.10.15

Tags: Other[INVALID TAG]

TxID	394696
ToLID	<b>fSebMen1_hap2</b>
Species	Sebastes mentella
Class	Actinopteri
Order	Perciformes

Genome Traits	Expected	Observed
Haploid size (bp)	796,775,342	818,977,083
Haploid Number	24 (source: ancestor)	24
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

## EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q58

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

### Curator notes

- . Interventions/Gb: 178
- . Contamination notes: ""
- . Other observations: "The assembly of *Sebastes mentella* (fSebMen1) has been generated as part of the Vertebrate Genomes Project (<https://vertebratogenomesproject.org/>). Both haplotypes were analyzed and manually improved using Pretext. "

# Quality metrics table

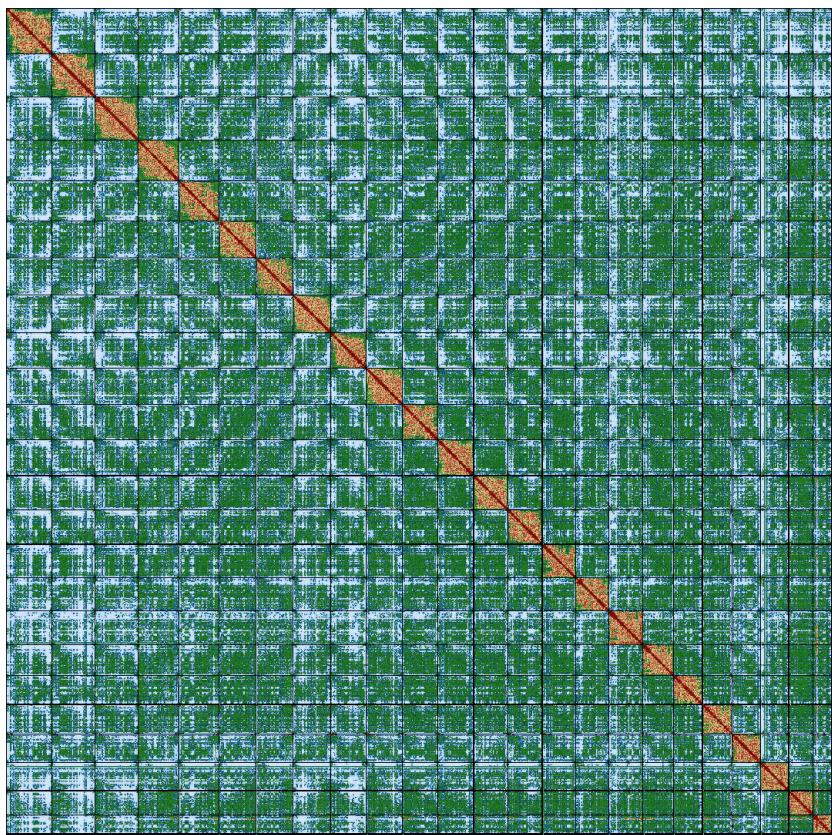
Metrics	Pre-curation collapsed	Curated collapsed
Total bp	832,662,030	818,977,083
GC %	40.88	40.84
Gaps/Gbp	488.79	579.99
Total gap bp	81,400	95,000
Scaffolds	369	246
Scaffold N50	35,252,739	34,938,992
Scaffold L50	11	11
Scaffold L90	22	21
Contigs	776	721
Contig N50	6,966,712	6,694,446
Contig L50	35	37
Contig L90	145	155
QV	58.4465	58.4326
Kmer compl.	93.8798	93.3219
BUSCO sing.	96.2%	97.7%
BUSCO dupl.	0.4%	0.7%
BUSCO frag.	0.8%	0.2%
BUSCO miss.	2.6%	1.3%

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)

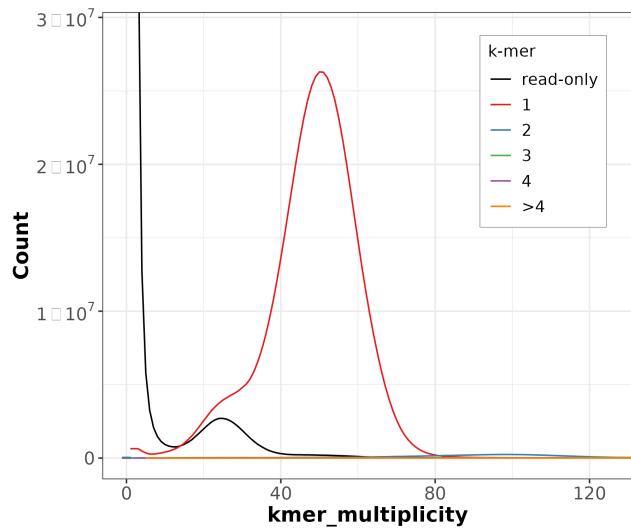
BUSCO: 6.0.0 (euk\_genome\_min, miniprot) / Lineage: actinopterygii\_odb12 (genomes:75, BUSCOs:7207)

# HiC contact map of curated assembly

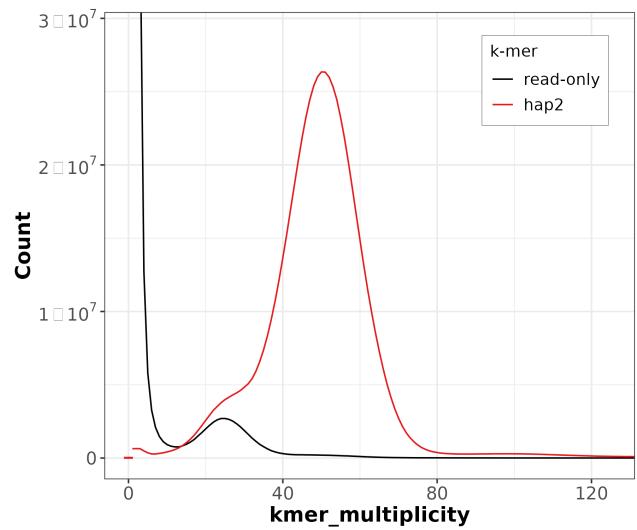


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

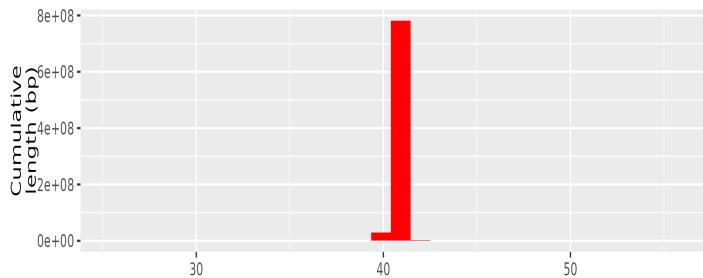


Distribution of k-mer counts per copy numbers found in asm

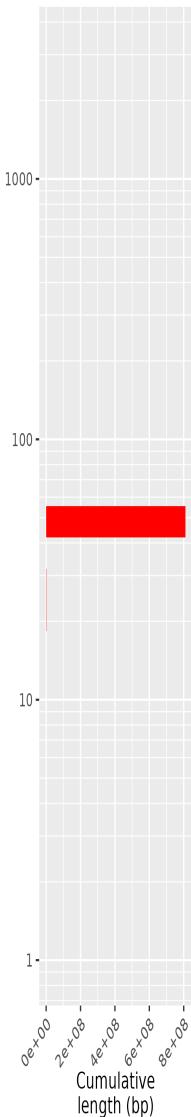
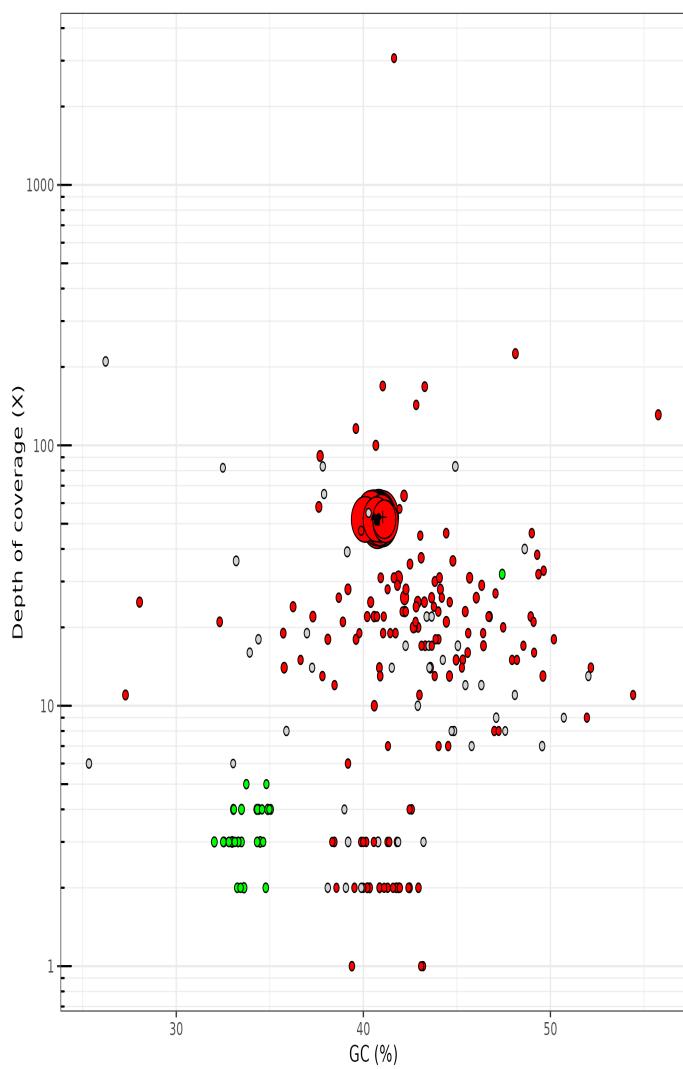


Distribution of k-mer counts coloured by their presence in reads/assemblies

# Post-curation contamination screening



TAPAs summary Graph



superkingdom

- Bacteria
- Eukaryota
- N/A

Length (bp)

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

Longest sequences (bp)

- SUPER\_1 - 44862014 (Eukaryota)
- SUPER\_2 - 43031409 (Eukaryota)
- SUPER\_3 - 42092157 (Eukaryota)
- SUPER\_4 - 40381062 (Eukaryota)
- SUPER\_5 - 39593787 (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	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: 2025-11-24 09:34:22 CET