

# ERGA Assembly Report

v24.10.15

Tags: Other[INVALID TAG]

TxID	394696
ToLID	<b>fSebMen1_combined</b>
Species	Sebastes mentella
Class	Actinopteri
Order	Perciformes

Genome Traits	Expected	Observed
Haploid size (bp)	796,775,342	1,655,338,225
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:

- . Observed Haploid size (bp) has >20% difference with Expected
- . Not 90% of assembly in chromosomes for collapsed

## Curator notes

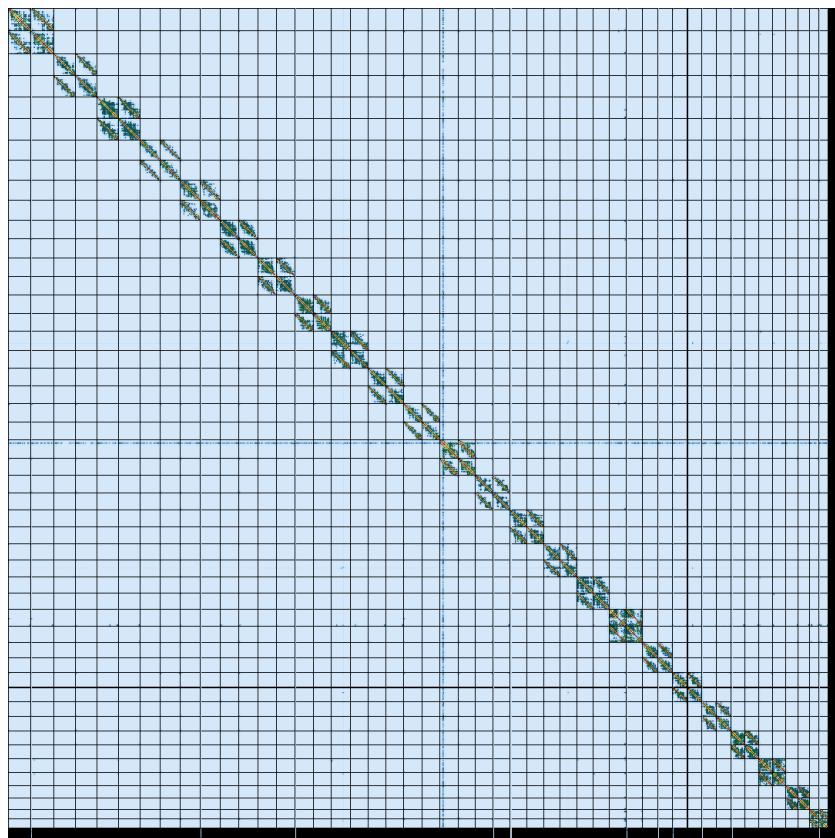
- . Interventions/Gb: 129
- . Contamination notes: ""
- . Other observations: "The assembly of *Sebastes mentella* (fSebMen1) has been generated as part of the Vertebrate Genomes Project (<https://vertebratogenomesproject.org/>). Both haplotypes were analyzed and manually improved using Pretext. "

## Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	1,655,313,025	1,655,338,225
GC %	40.86	40.86
Gaps/Gbp	485.71	561.82
Total gap bp	160,800	186,000
Scaffolds	888	759
Scaffold N50	34,728,012	35,475,877
Scaffold L50	22	22
Scaffold L90	44	42
Contigs	1,692	1,689
Contig N50	6,886,770	6,922,373
Contig L50	69	69
Contig L90	300	301
QV	58.0938	58.0857
Kmer compl.	99.4868	99.4868
BUSCO sing.	2.8%	
BUSCO dupl.	94.8%	
BUSCO frag.	0.7%	
BUSCO miss.	1.8%	

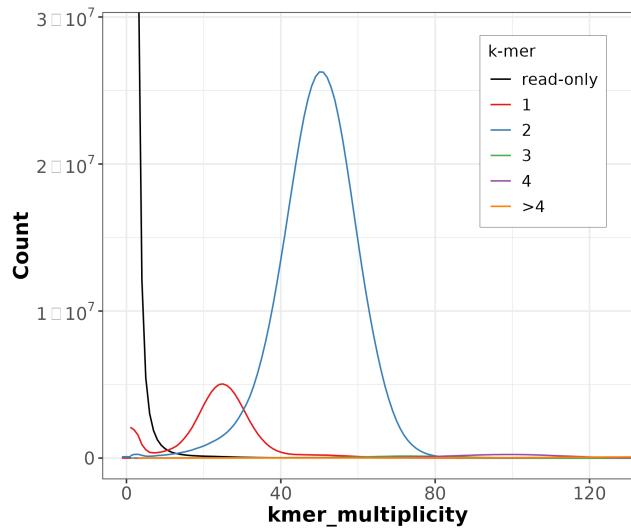
BUSCO: 5.8.2 (euk\_genome\_met, metaeuk) / Lineage: actinopterygii\_odb12 (genomes:75, BUSCOs:7207)

# HiC contact map of curated assembly

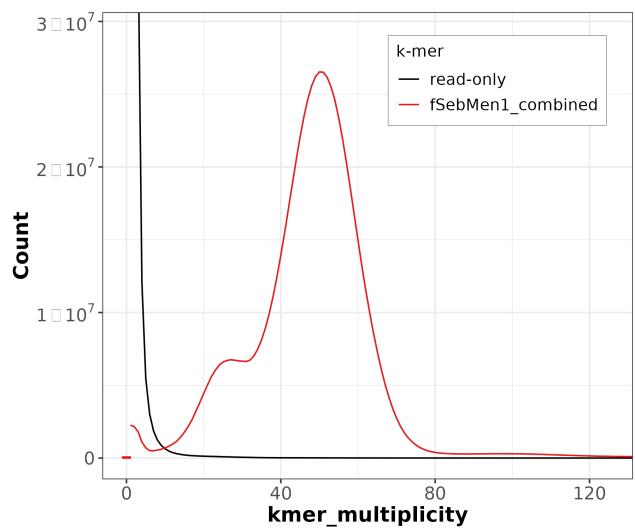


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

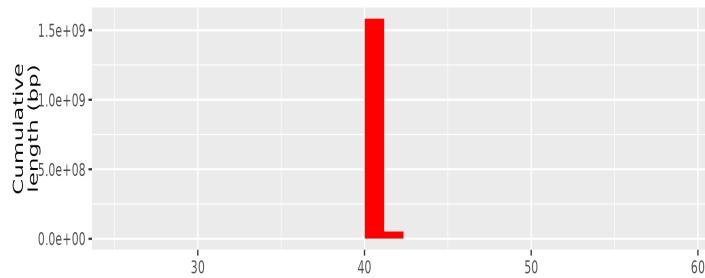


Distribution of k-mer counts per copy numbers found in asm

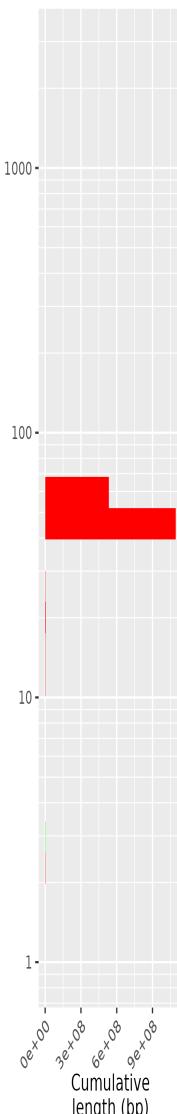
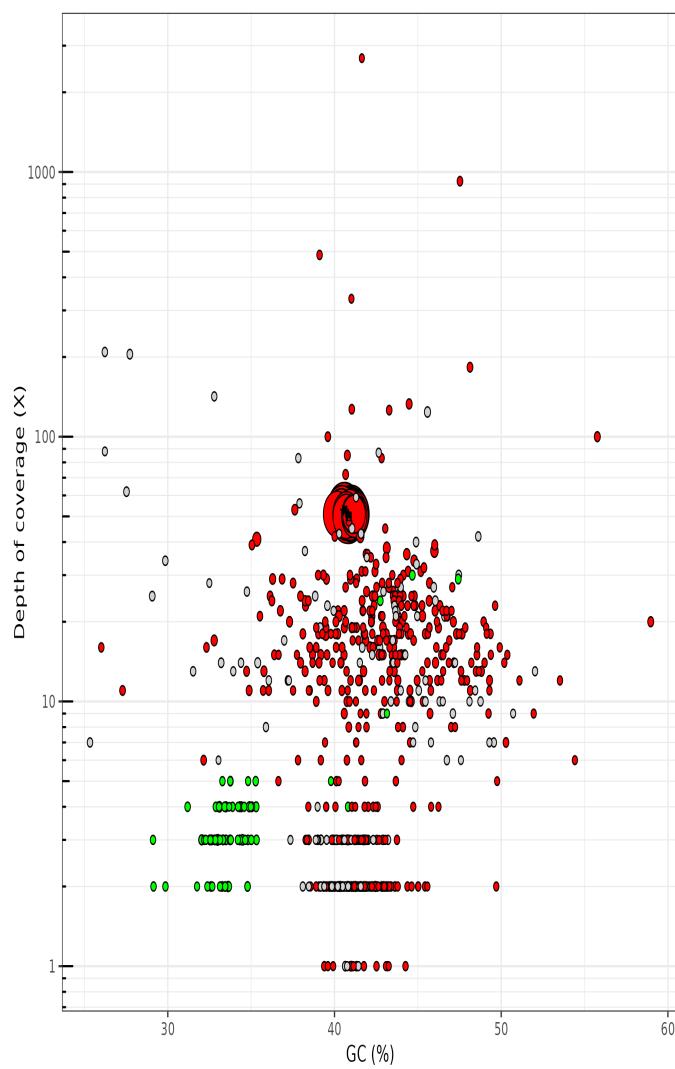


Distribution of k-mer counts coloured by their presence in reads/assemblies

# Post-curation contamination screening



TAPAs summary Graph



## superkingdom

- Bacteria
- Eukaryota
- N/A

## Length (bp)

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

## Longest sequences (bp)

- Hap1\_fSebMen1\_1 - 44889610 (Eukaryota)
- ▲ Hap2\_fSebMen1\_1 - 44862014 (Eukaryota)
- Hap1\_fSebMen1\_2 - 43573437 (Eukaryota)
- + Hap2\_fSebMen1\_2 - 43031409 (Eukaryota)
- ▣ Hap1\_fSebMen1\_3 - 42669276 (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: 2025-10-06 14:03:29 CEST