ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	94921
ToLID	**fBotLun1**
Species	Bothus lunatus
Class	Actinopteri
Order	Pleuronectiformes

Genome Traits	Expected	Observed
Haploid size (bp)	528,054,120	555,807,845
Haploid Number	19 (source: ancestor)	20
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q45

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

. Observed Haploid Number is different from Expected

Curator notes

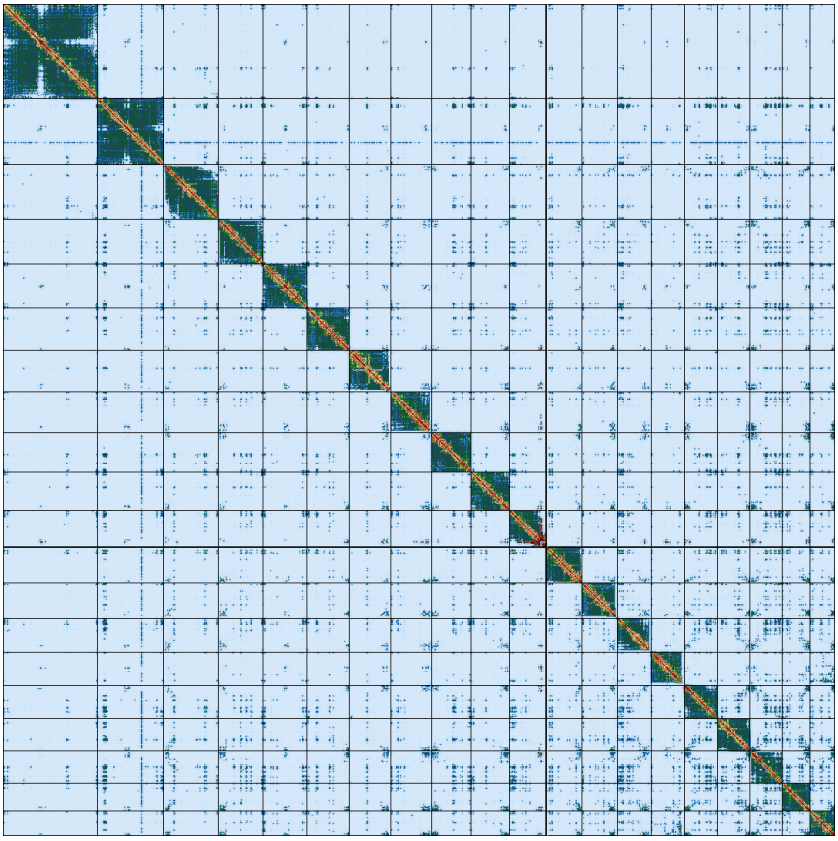
. Interventions/Gb: 0
. Contamination notes: ""
. Other observations: "The assembly of Bothus lunatus (fBotLun1) is based on 60X ONT data and 145X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial ONT assembly generation with Hifiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. In total, 1 contigs was identified as contaminants (bacterial, archaeal, or viral), totaling 0.018 Mb (with the largest being 0.018 Mb). Additionally, 24 regions totaling 1.687 Mb (with the largest being 0.38 Mb) were identified as haplotypic duplications and removed. Finally, the primary assembly was analyzed and manually improved using Pretext. 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	555,802,319	555,807,845
GC %	44.71	44.71
Gaps/Gbp	1.8	17.99
Total gap bp	100	1,900
Scaffolds	34	25
Scaffold N50	27,280,616	27,280,616
Scaffold L50	8	8
Scaffold L90	18	18
Contigs	35	35
Contig N50	27,280,616	27,280,616
Contig L50	8	8
Contig L90	18	18
QV	45.0199	45.0212
Kmer compl.	93.1944	93.195
BUSCO sing.	98.8%	98.8%
BUSCO dupl.	0.2%	0.2%
BUSCO frag.	0.1%	0.1%
BUSCO miss.	0.9%	0.9%

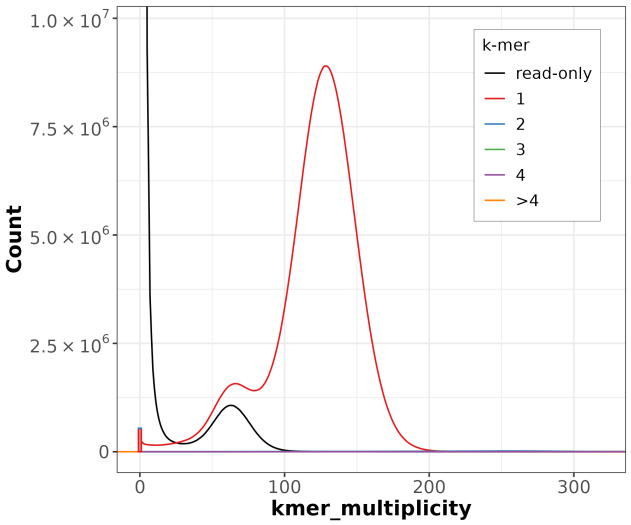
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

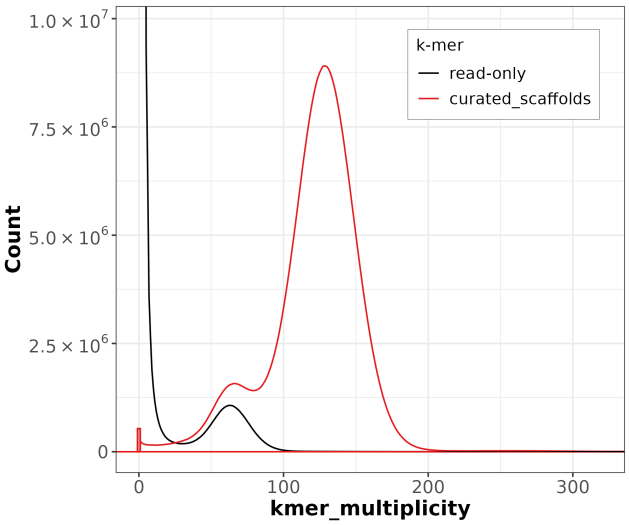


collapsed [LINK]

K-mer spectra of curated assembly

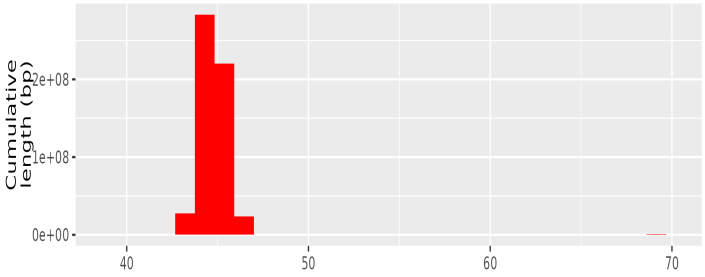


Distribution of k-mer counts per copy numbers found in asm

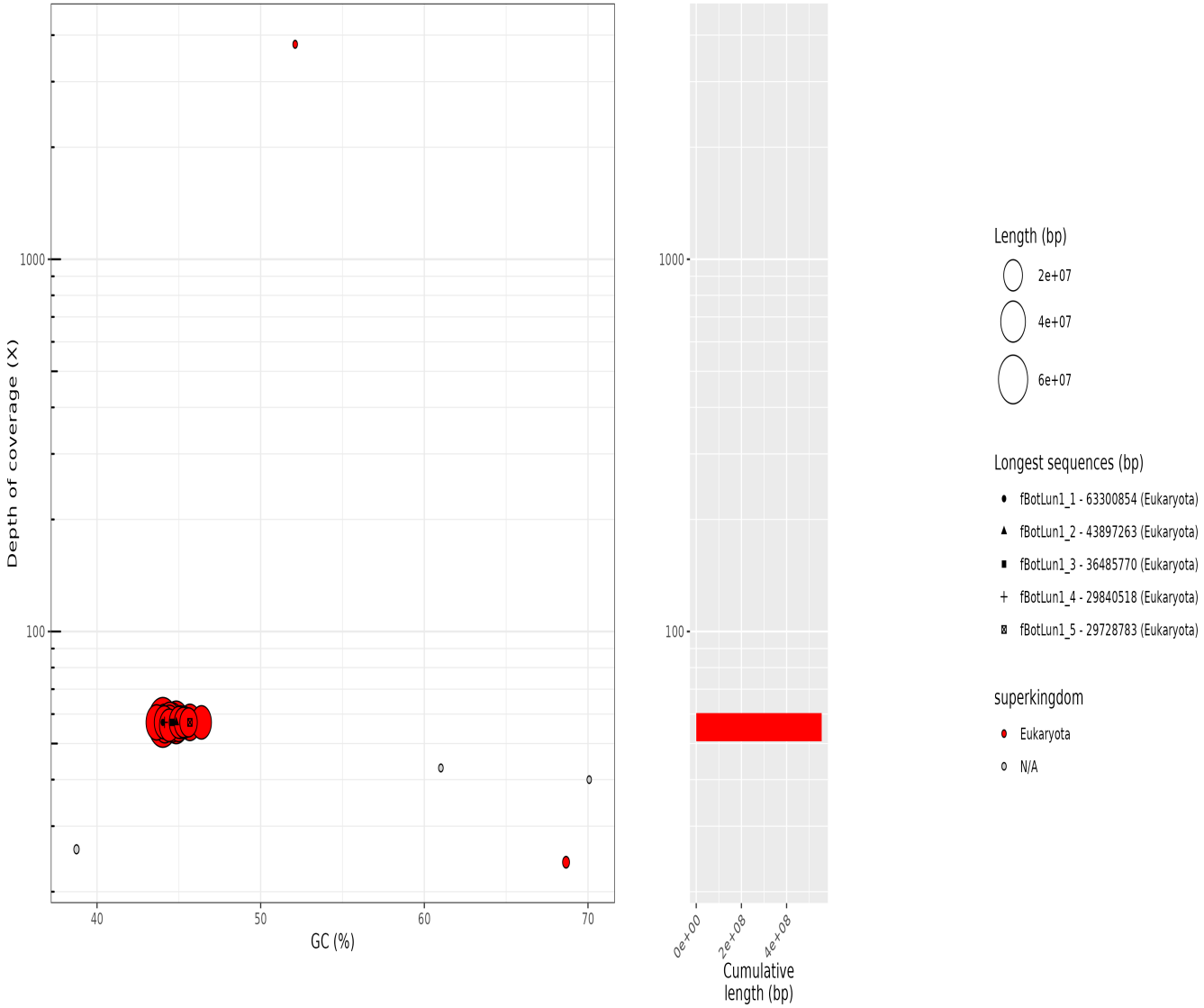


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



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	60	145

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: 2026-03-17 09:11:53 CET