

ERGA Assembly Report

v24.10.15

Tags: ATLASEa[INVALID TAG]

TxID	30804
ToLID	fLizRam5
Species	Chelon ramada
Class	Actinopteri
Order	Mugiliiformes

Genome Traits	Expected	Observed
Haploid size (bp)	673,345,717	674,117,278
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: 7.7.Q47

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: 18
- . Contamination notes: ""
- . Other observations: "The assembly of Chelon ramada (fLizRam5) is based on 45X PacBio data and 176X Arima Hi-C data generated as part of the ATLASEa programme (<https://www.atlasea.fr>). The assembly process included the following steps: initial PacBio assembly generation with Hifiiasm, 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 0.06 Mb (with the largest being 0.06 Mb). Additionally, 44 regions totaling 3.317 Mb (with the largest being 0.418 Mb) were identified as haplotypic duplications and removed. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 1 haplotypic region was removed, totaling 0.077Mb. The telomeric pattern TTAGGG was identified with TelFinder and used to generate the Pretext Telomeres tracks. Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. "

Quality metrics table

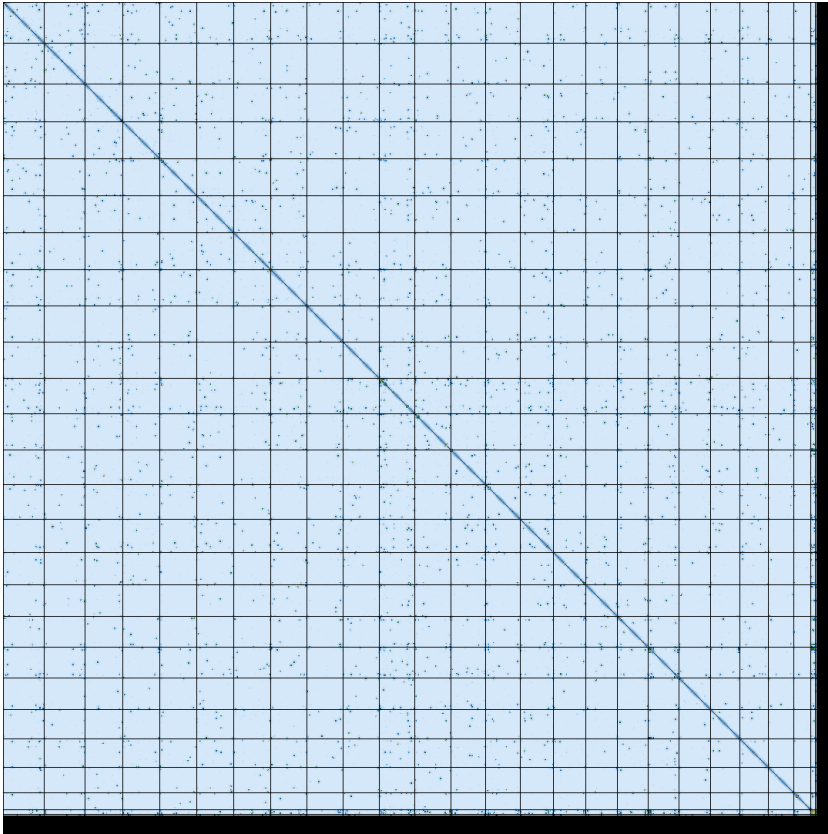
Metrics	Pre-curation collapsed	Curated collapsed
Total bp	674,921,595	674,117,278
GC %	42.33	42.32
Gaps/Gbp	66.67	80.1
Total gap bp	4,500	6,400
Scaffolds	170	153
Scaffold N50	28,868,819	28,915,000
Scaffold L50	12	12
Scaffold L90	22	22
Contigs	215	207
Contig N50	27,180,000	27,180,000
Contig L50	12	12
Contig L90	25	25
QV	47.0017	47.0177
Kmer compl.	86.6436	86.6382
BUSCO sing.	99.1%	98.1%
BUSCO dupl.	0.3%	0.3%
BUSCO frag.	0.1%	0.7%
BUSCO miss.	0.6%	0.9%

Warning! BUSCO versions or lineage datasets are not the same across results:

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

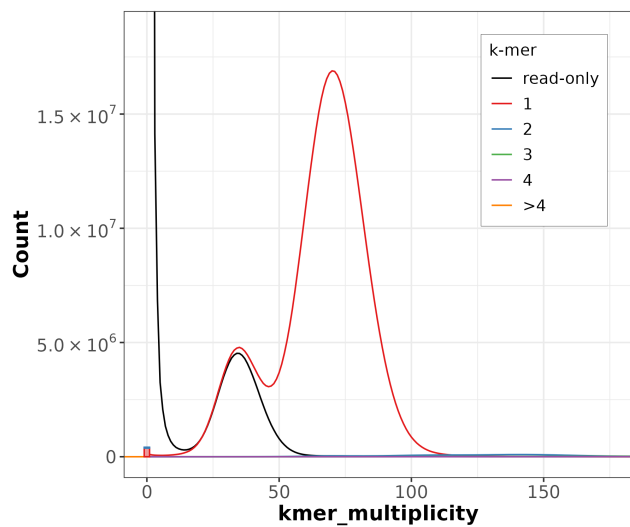
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

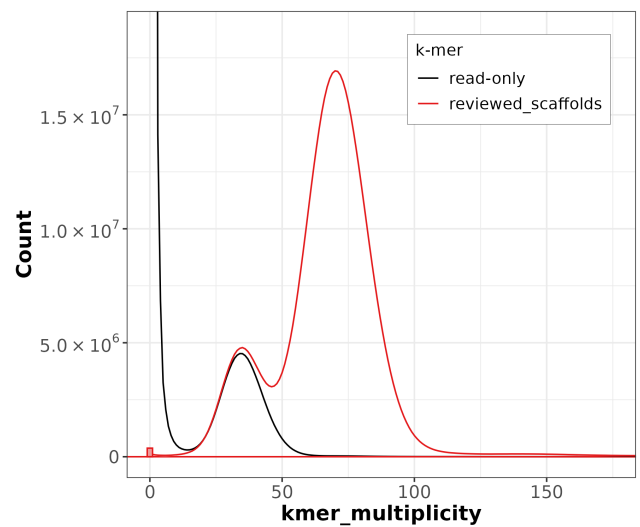


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

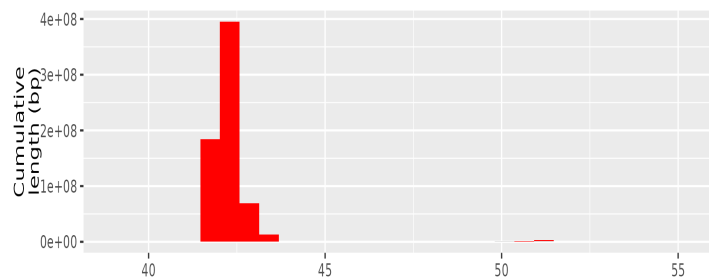


Distribution of k-mer counts per copy numbers found in asm

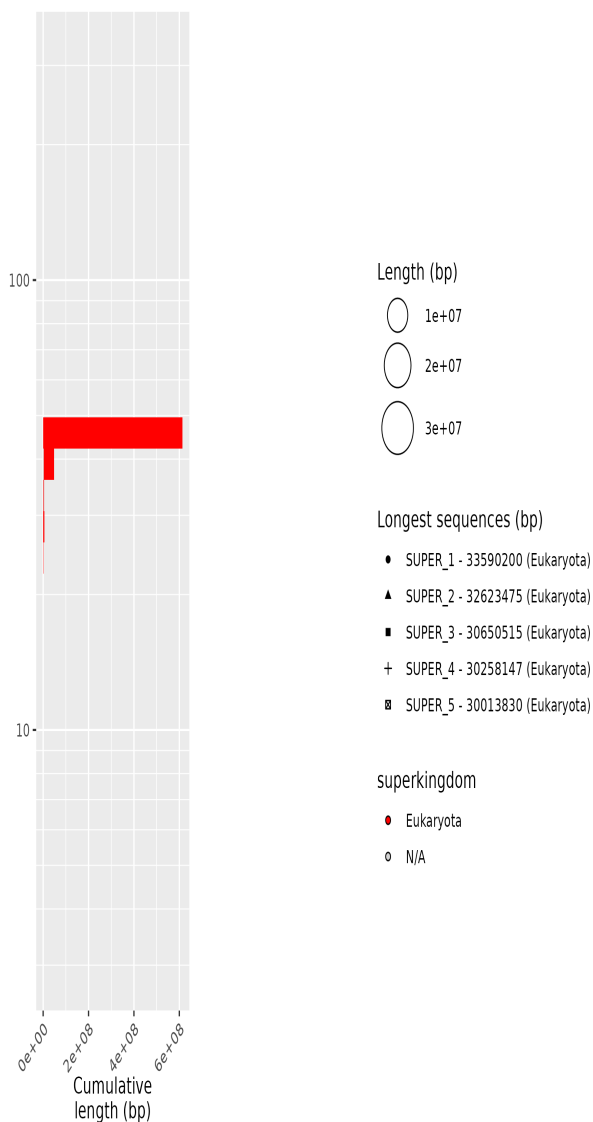
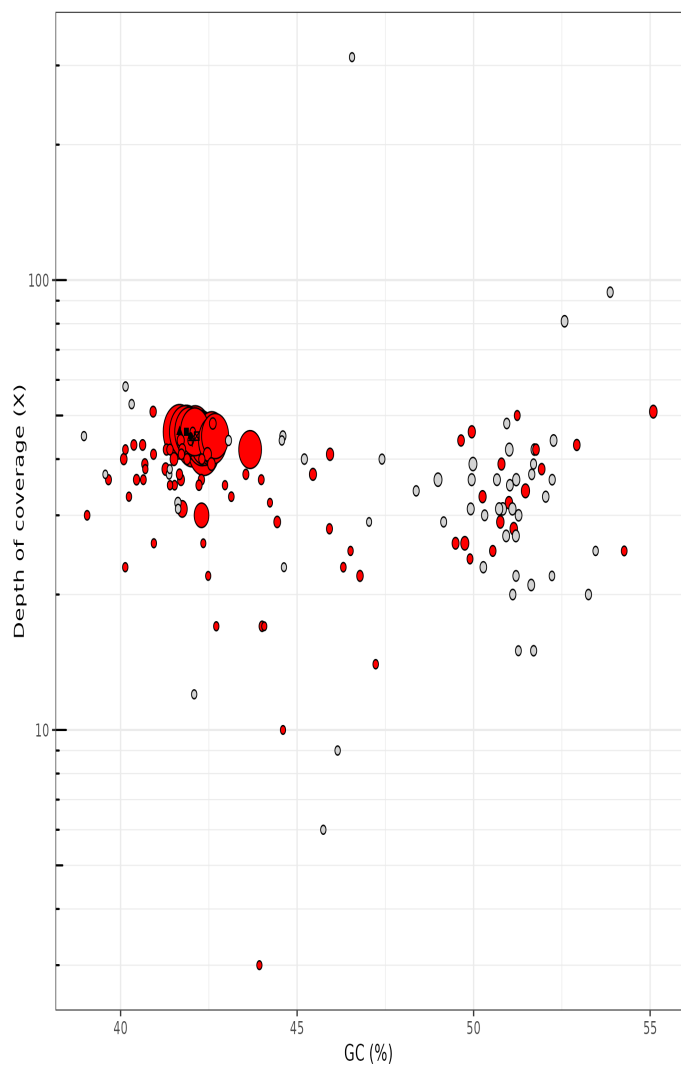


Distribution of k-mer counts coloured by their presence in reads/assemblies

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	45	177

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: Sophie Layac

Affiliation: Genoscope

Date and time: 2026-08-01 08:34:53 CEST