ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	1491482
ToLID	**fCarAcu1**
Species	Carapus acus
Class	Actinopteri
Order	Ophidiiformes

Genome Traits	Expected	Observed
Haploid size (bp)	450,511,373	466,342,956
Haploid Number	21 (source: ancestor)	18
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q46

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

. Observed Haploid Number is different from Expected

. Kmer completeness value is less than 90 for collapsed

Curator notes

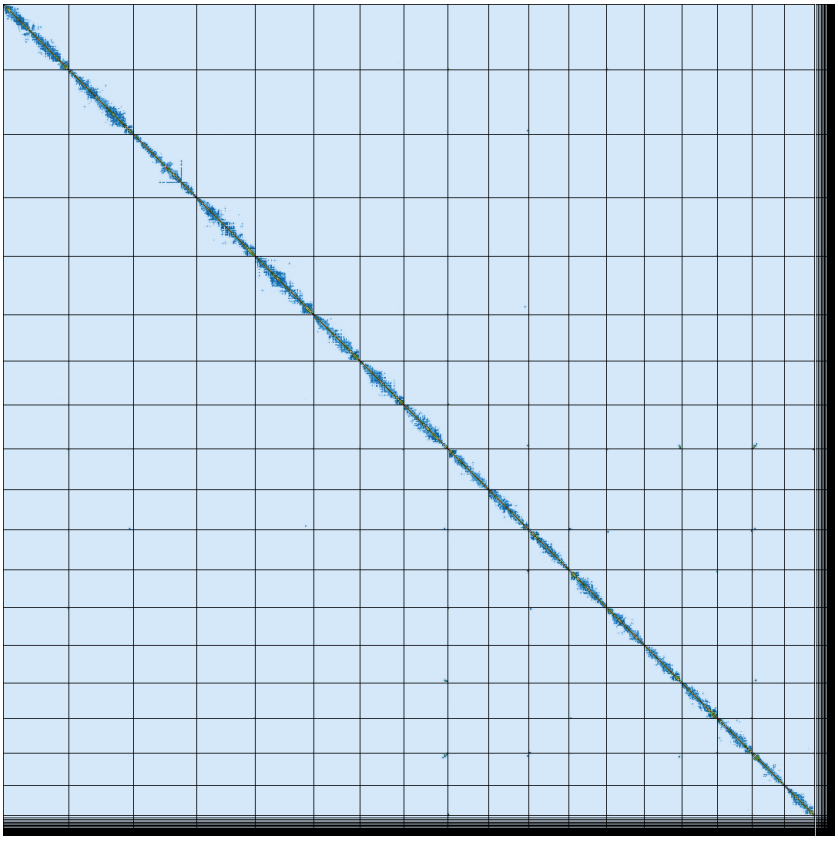
. Interventions/Gb: 0
. Contamination notes: ""
. Other observations: "The assembly of Carapus acus (fCarAcu1) is based on 88X ONT data and 118X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial ONT assembly generation with Hifiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. No contigs were found to be contaminants (bacterial, archaeal, or viral). Additionally, 35 regions totaling 4.273 Mb (with the largest being 0.662 Mb) were identified as haplotypic duplications and removed. Mitochondrial genome was assembled using ptGAUL. 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	466,326,303	466,342,956
GC %	47.59	47.59
Gaps/Gbp	17.16	17.15
Total gap bp	800	800
Scaffolds	56	57
Scaffold N50	24,747,000	24,747,000
Scaffold L50	8	8
Scaffold L90	16	16
Contigs	64	65
Contig N50	22,261,891	22,261,891
Contig L50	9	9
Contig L90	18	18
QV	46.5399	46.5401
Kmer compl.	84.4038	84.4072
BUSCO sing.	97.4%	97.4%
BUSCO dupl.	0.4%	0.4%
BUSCO frag.	0.3%	0.3%
BUSCO miss.	1.9%	1.9%

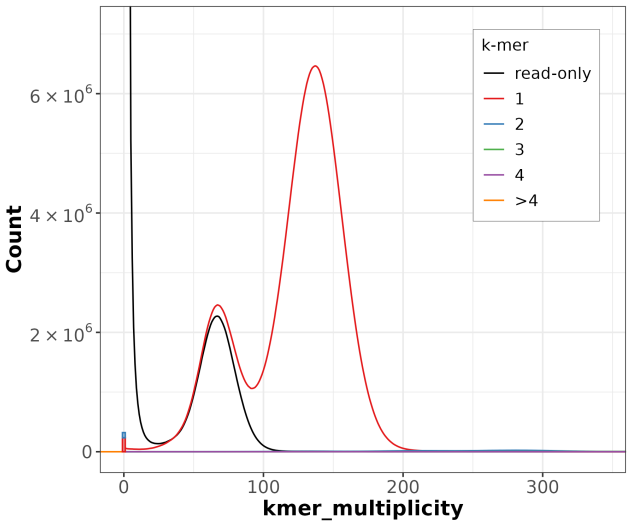
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

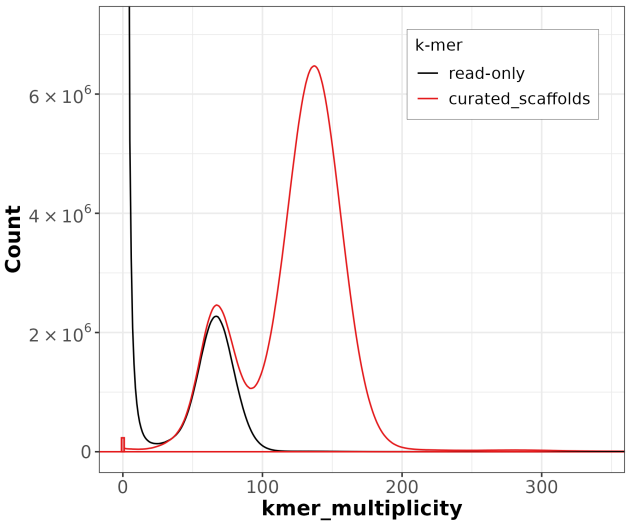


collapsed [LINK]

K-mer spectra of curated assembly

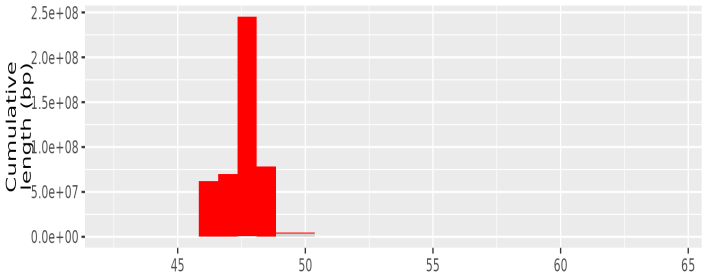


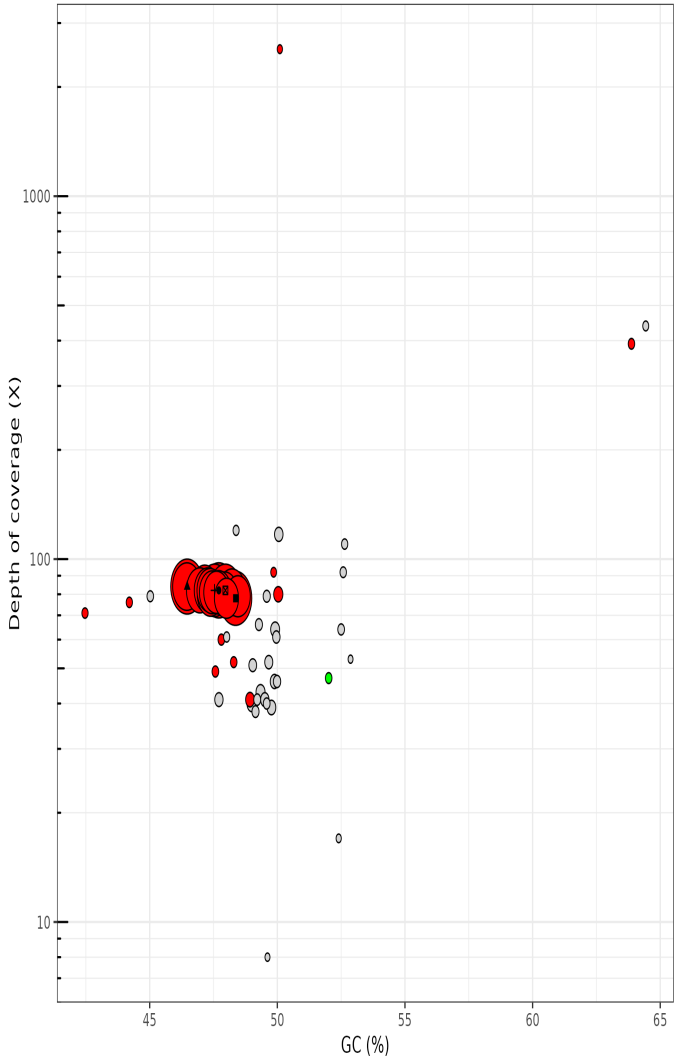
Distribution of k-mer counts per copy numbers found in asm

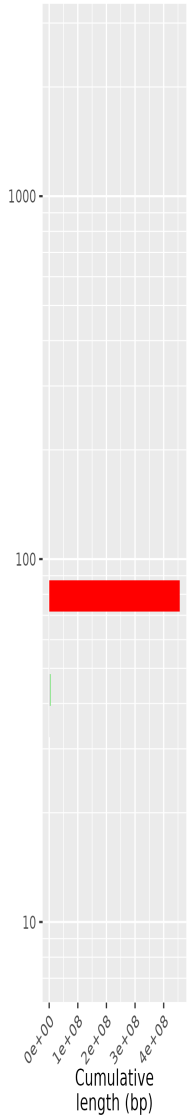


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening

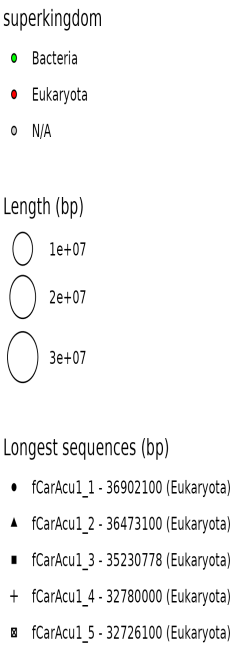






TAPAs summary Graph

(1 0X contig has been hidden)



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	88	118

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: 2026-06-09 11:51:40 CEST