

ERGA Assembly Report

v24.10.15

Tags: ATLASEa[INVALID TAG]

TxID	994942
ToLID	xbCarOrna1
Species	Caribachlamys ornata
Class	Bivalvia
Order	Pectinida

Genome Traits	Expected	Observed
Haploid size (bp)	1,230,504,543	1,291,897,826
Haploid Number	19 (source: ancestor)	19
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q49

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

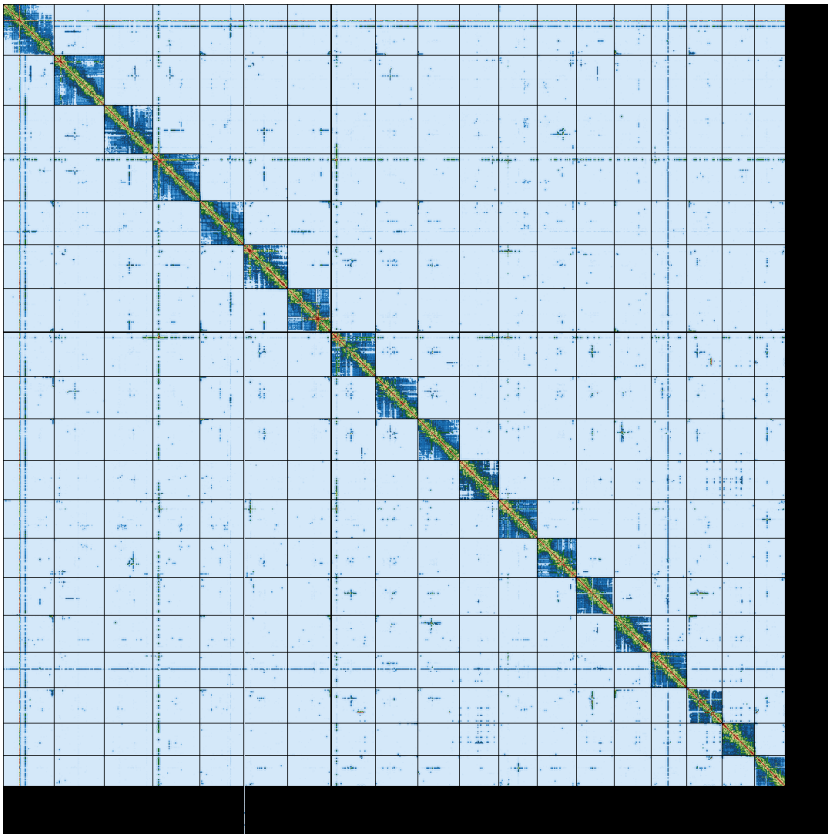
- . Interventions/Gb: 73
- . Contamination notes: ""
- . Other observations: "The assembly of Caribachlamys ornata (xbCarOrna1) is based on 51X PacBio data and 167X Arima Hi-C data generated as part of the ATLASEa programme (<https://www.atlasea.fr>).The assembly process included the following steps: initial PacBio assembly generation with Hifiasm in Hi-C integrated assembly mode, removal of contaminant sequences using Context and Hi-C based scaffolding with YaHS. In total, 9 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 3.6 Mb (with the largest being 0.915 Mb). The mitochondrial genome was assembled using OATK. Curation was performed first on both haplotypes simultaneously using Pretext, and then in a second round individually. 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	1,311,260,172	1,291,897,826
GC %	37	37.31
Gaps/Gbp	188.37	319.68
Total gap bp	24,700	46,700
Scaffolds	214	516
Scaffold N50	61,972,184	65,011,873
Scaffold L50	10	10
Scaffold L90	19	18
Contigs	461	929
Contig N50	6,780,434	4,247,000
Contig L50	60	83
Contig L90	202	326
QV	49.9485	49.3358
Kmer compl.	65.4258	63.428
BUSCO sing.	94.1%	98.9%
BUSCO dupl.	5.5%	0.8%
BUSCO frag.	0.0%	0.1%
BUSCO miss.	0.4%	0.2%

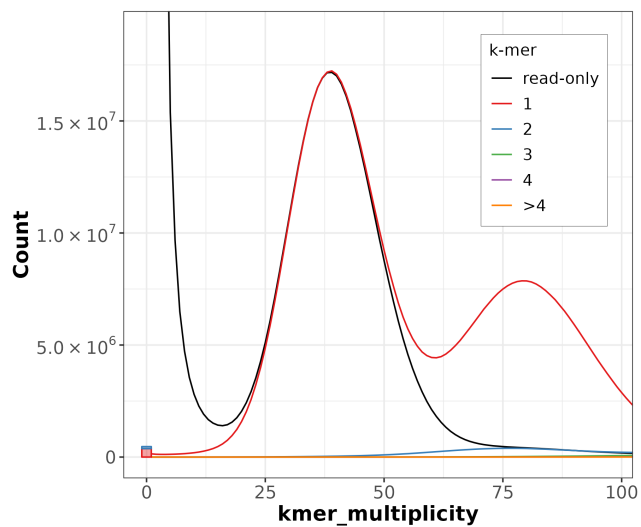
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: mollusca_odb12 (genomes:36, BUSCOs:4421)

HiC contact map of curated assembly

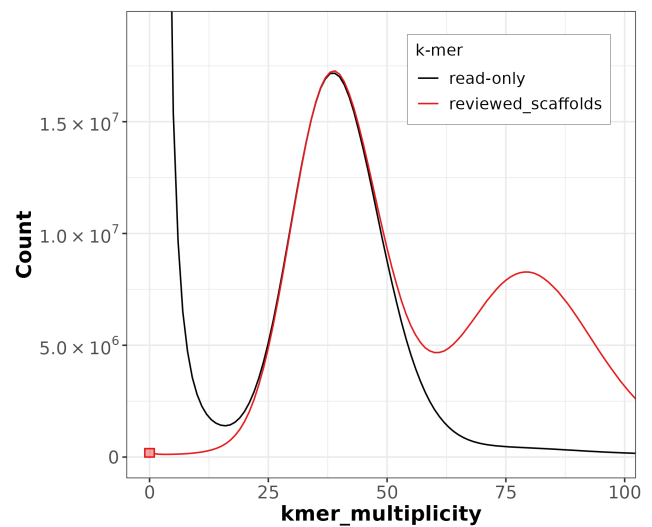


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

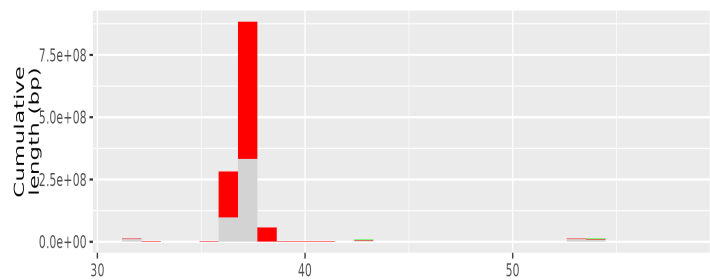


Distribution of k-mer counts per copy numbers found in asm



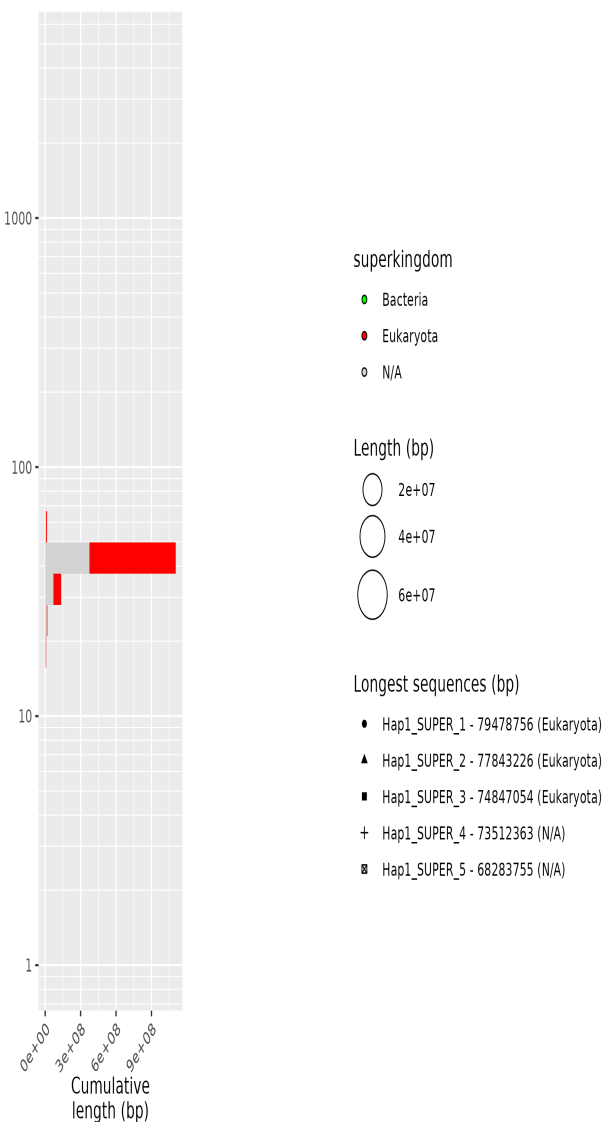
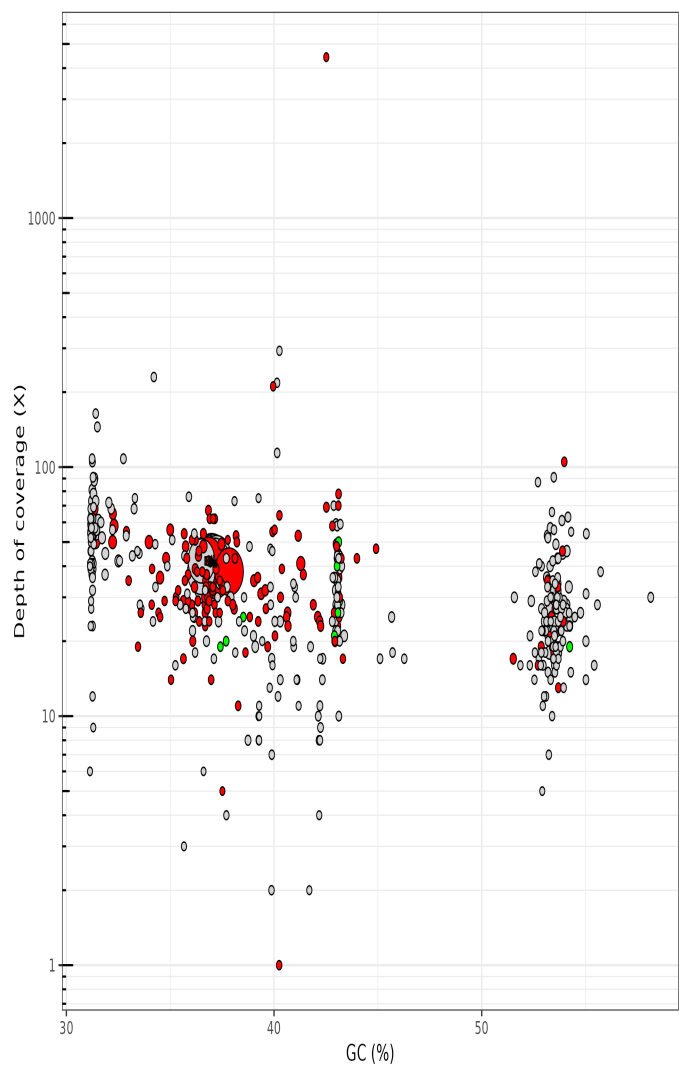
Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph

(10 0X contigs have 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	51	167

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: Caroline Menguy

Affiliation: Genoscope

Date and time: 2026-01-31 00:54:13 CET