

# ERGA Assembly Report

v24.04.03\_beta

Tags: ATLASea[INVALID TAG]

TxID	55715
ToLID	<b>xbVenVerr1.1</b>
Species	Venus verrucosa
Class	Bivalvia
Order	Venerida

Genome Traits	Expected	Observed
Haploid size (bp)	1,900,759,880	2,087,867,859
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: 7.8.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
- . BUSCO single copy value is less than 90% for collapsed
- . Assembly length loss > 3% for collapsed

### Curator notes

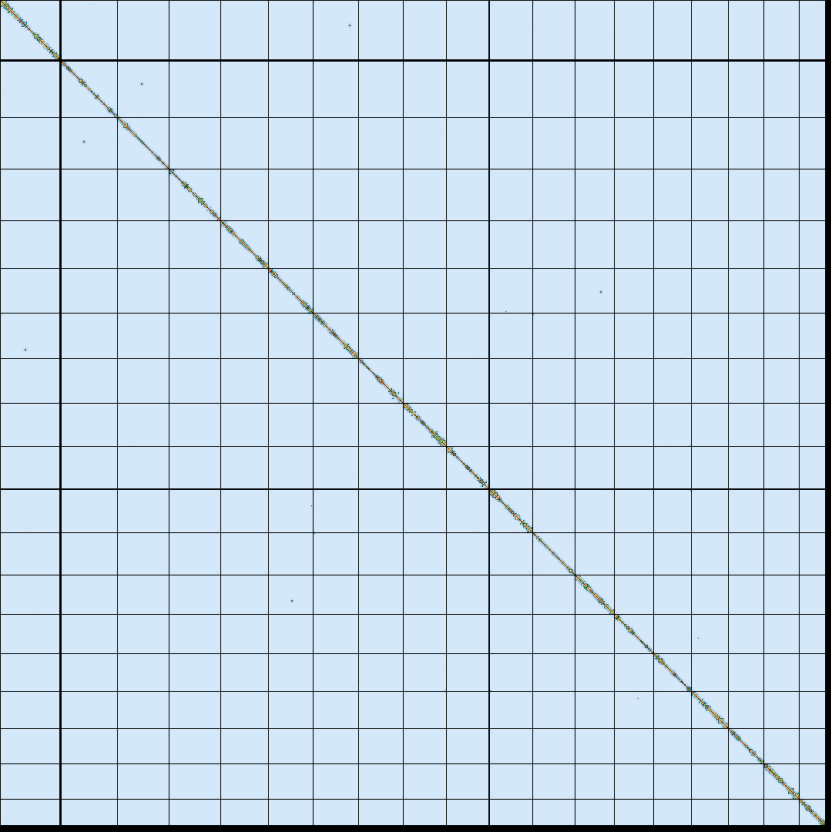
- . Interventions/Gb: None
- . Contamination notes: "One contig was tagged Contaminant during the manual curation"
- . Other observations: "Remove numerous haplotypic duplications during the manual curation"

## Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	2,211,205,762	2,087,867,859
GC %	35.25	35.24
Gaps/Gbp	130.25	128.84
Total gap bp	28,800	31,400
Scaffolds	221	177
Scaffold N50	112,710,031	108,501,084
Scaffold L50	9	9
Scaffold L90	17	17
Contigs	509	446
Contig N50	14,062,000	15,529,060
Contig L50	45	42
Contig L90	143	134
QV	47.11	47.1299
Kmer compl.	66.1492	63.6095
BUSCO sing.	77.2%	79.6%
BUSCO dupl.	4.8%	2.0%
BUSCO frag.	4.4%	4.4%
BUSCO miss.	13.6%	14.0%

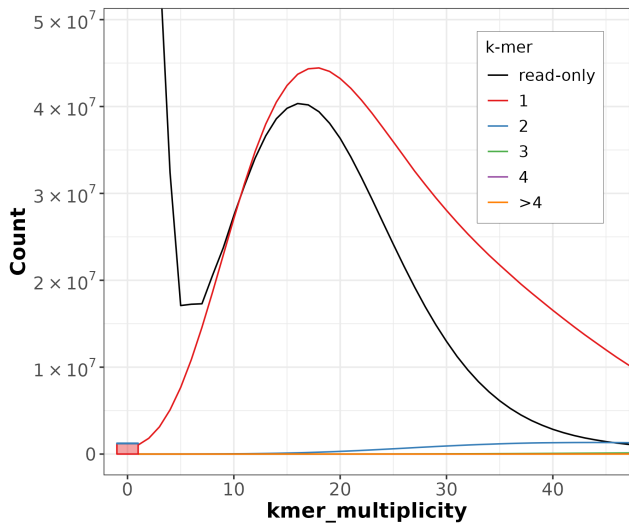
BUSCO 5.4.3 Lineage: mollusca\_odb10 (genomes:7, BUSCOs:5295)

# HiC contact map of curated assembly

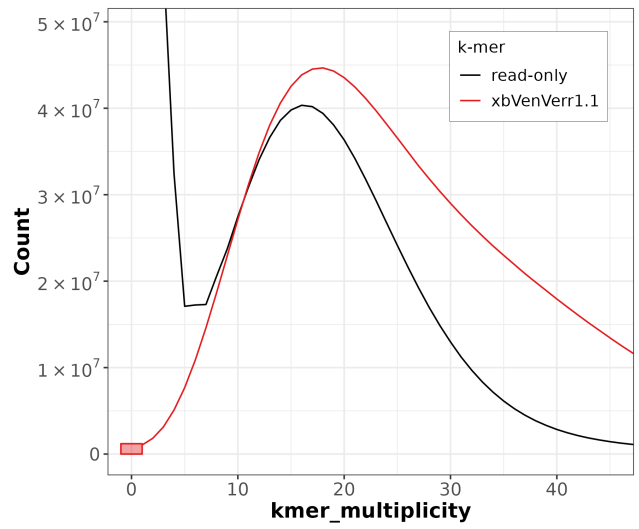


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

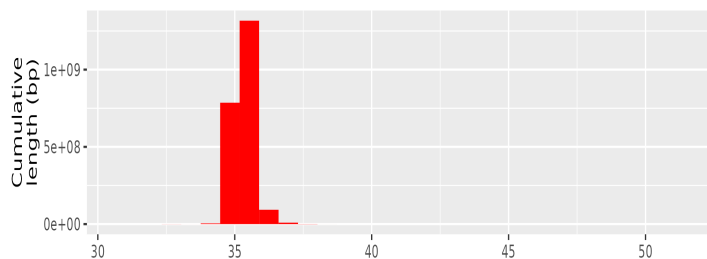


Distribution of k-mer counts per copy numbers found in asm

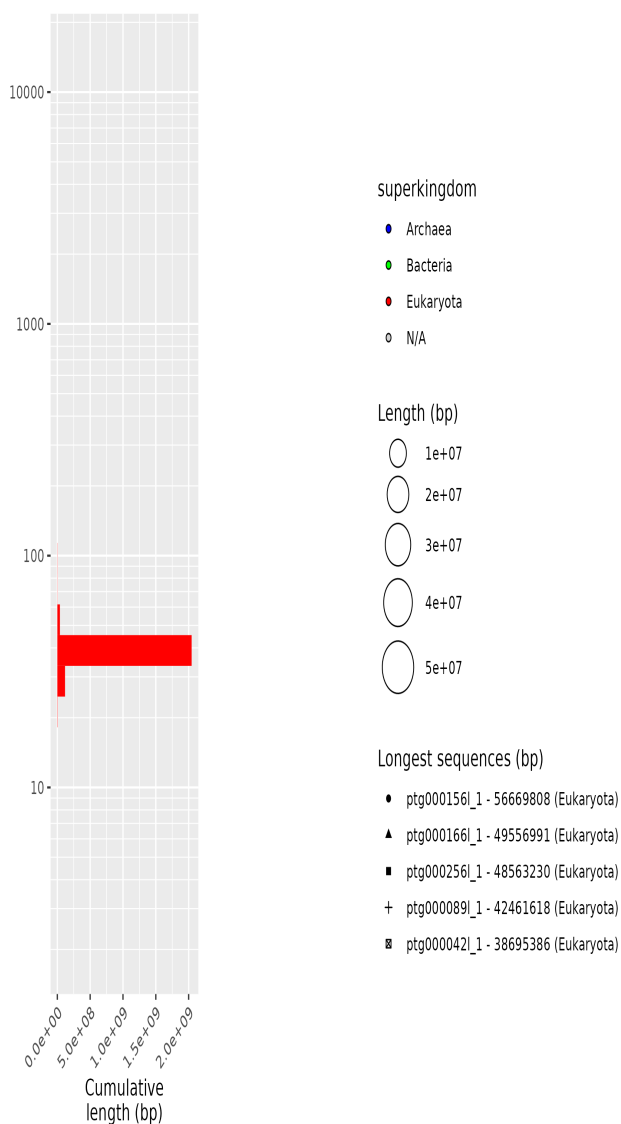
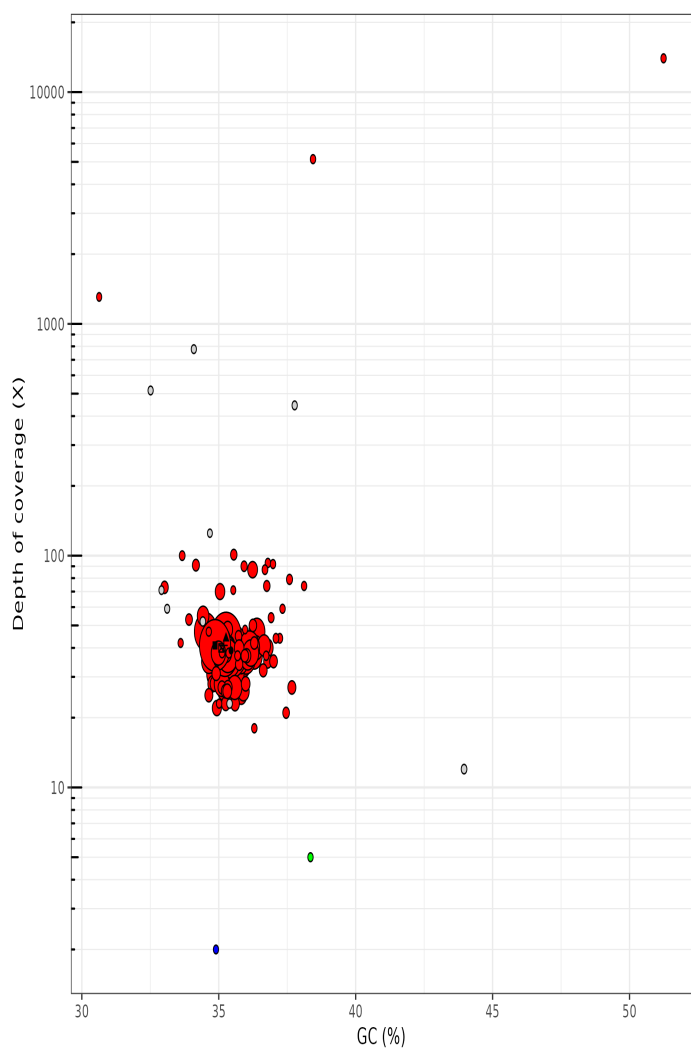


Distribution of k-mer counts coloured by their presence in reads/assemblies

# 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	PACBIO Hifi
Coverage	42X

# 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

- **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

Submitter: Jean-Marc Aury

Affiliation: Genoscope

Date and time: 2024-07-04 14:24:35 CEST