

ERGA Assembly Report

v24.10.15

Tags: ATLASEa[INVALID TAG]

TxID	3088264
ToLID	xbAcaDesh1
Species	<i>Acanthocardia deshayesii</i>
Class	Bivalvia
Order	Cardiida

Genome Traits	Expected	Observed
Haploid size (bp)	1,408,141,178	1,392,406,787
Haploid Number	26 (source: ancestor)	19
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q50

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: 45
. Contamination notes: ""
. Other observations: "The assembly of *Acanthocardia deshayesii* (xbAcaDesh1) is based on 59X PacBio data and 160X 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, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. In total, 8 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 0.286 Mb (with the largest being 0.079 Mb). Additionally, 352 regions totaling 71.457 Mb (with the largest being 4.087 Mb) were identified as haplotypic duplications and removed. Mitochondrial genome was assembled using OATK. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 16 haplotypic regions were removed, totaling 26.195Mb (with the largest being 4.416Mb). Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. "

Quality metrics table

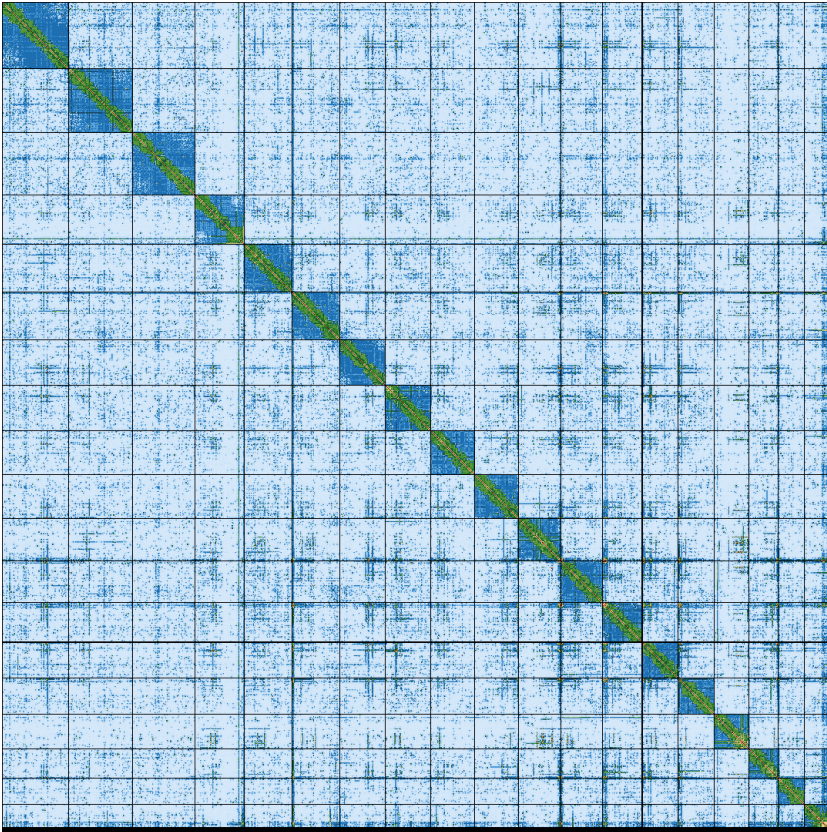
Metrics	Pre-curation collapsed	Curated collapsed
Total bp	1,418,612,480	1,392,406,787
GC %	37.44	37.44
Gaps/Gbp	83.88	96.24
Total gap bp	11,900	16,700
Scaffolds	111	93
Scaffold N50	74,318,164	75,911,673
Scaffold L50	8	8
Scaffold L90	17	17
Contigs	230	227
Contig N50	18,649,377	16,871,271
Contig L50	24	24
Contig L90	92	90
QV	50.0746	50.0641
Kmer compl.	66.8568	66.3489
BUSCO sing.	97.9%	96.9%
BUSCO dupl.	1.7%	0.8%
BUSCO frag.	0.1%	1.8%
BUSCO miss.	0.3%	0.5%

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

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

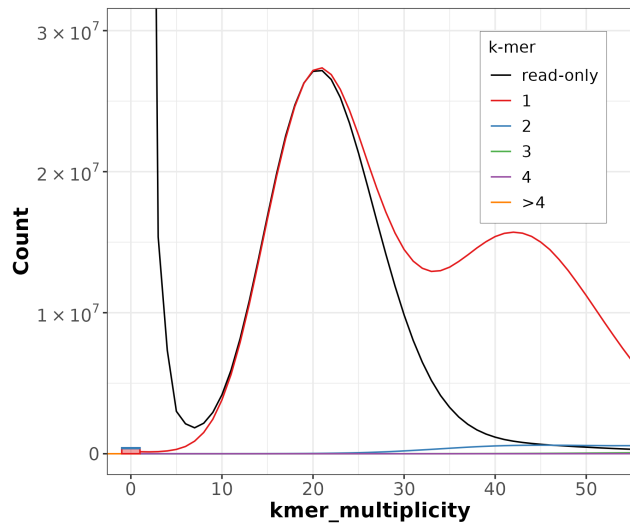
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: mollusca_odb12 (genomes:36, BUSCOs:4421)

HiC contact map of curated assembly

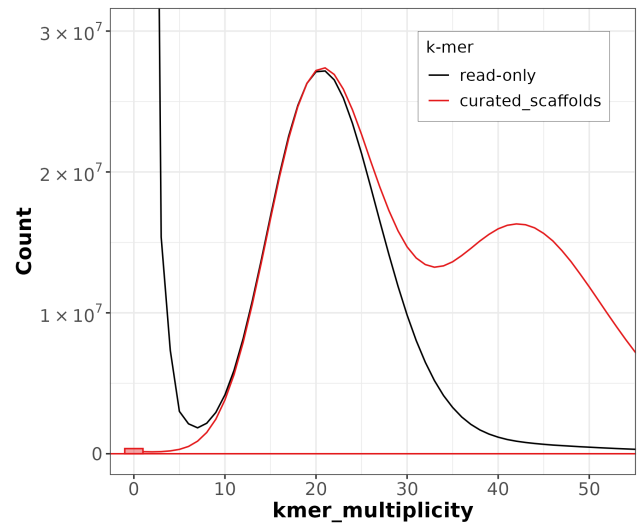


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

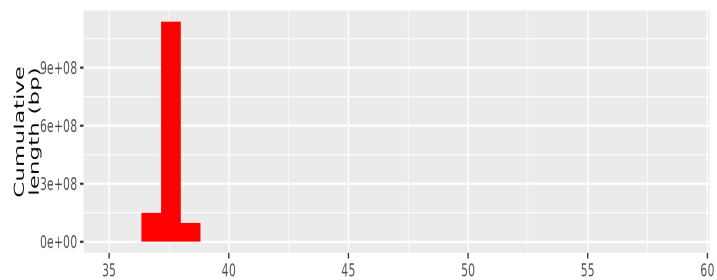


Distribution of k-mer counts per copy numbers found in asm



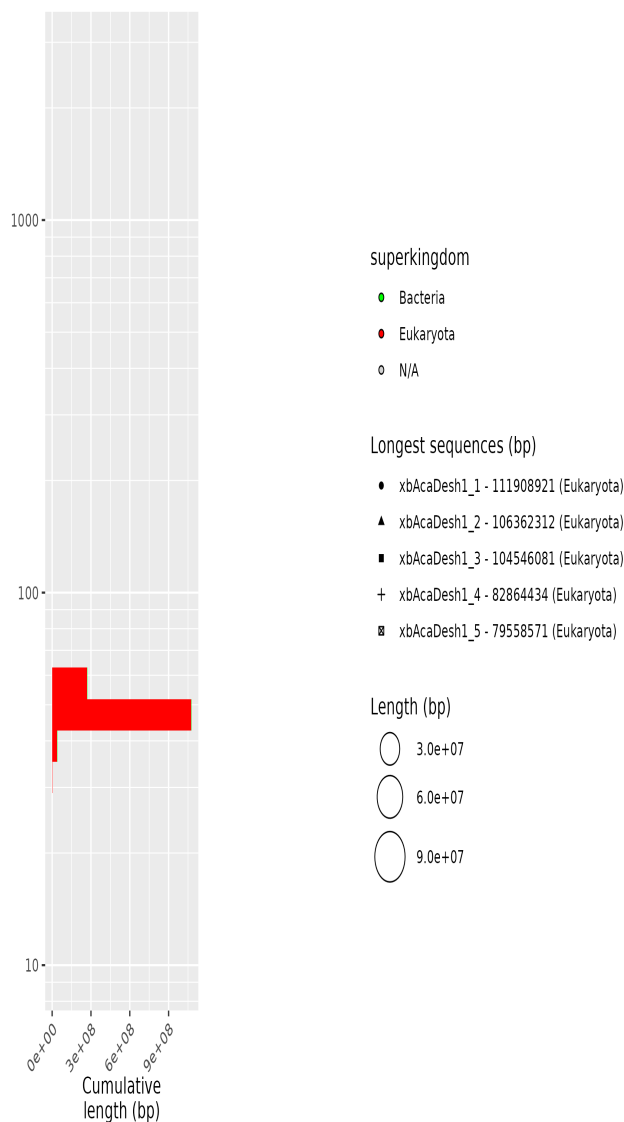
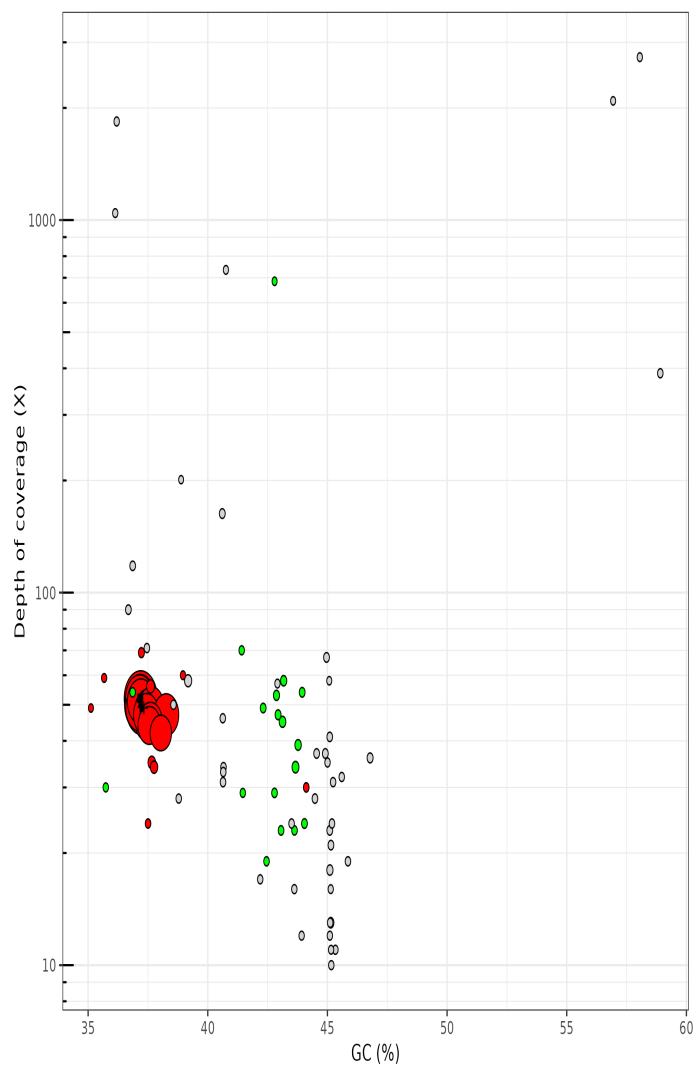
Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph

(2 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	59	162

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: 2026-08-07 11:03:53 CEST