ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	1920007
ToLID	**xbDonVari1.1**
Species	Donax variegatus
Class	Bivalvia
Order	Cardiida

Genome Traits	Expected	Observed
Haploid size (bp)	1,586,383,048	1,635,800,218
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.Q67

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

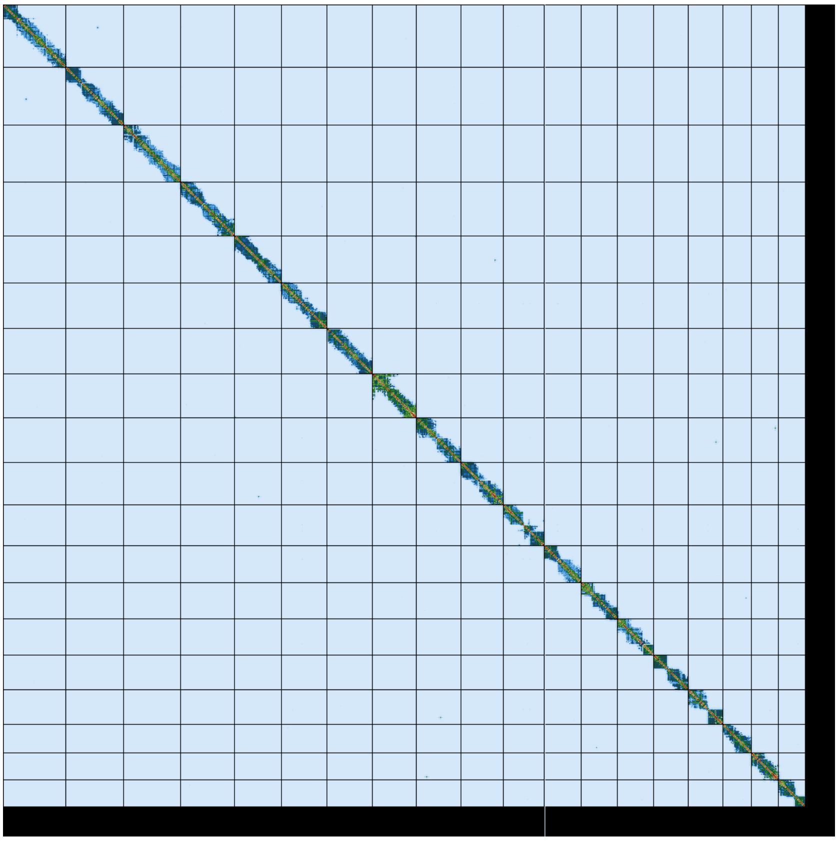
. Interventions/Gb: 9
. Contamination notes: "82 bacterian contigs removed after assembly "
. Other observations: ""

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	1,643,562,596	1,635,800,218
GC %	35.01	35.01
Gaps/Gbp	16.43	22.01
Total gap bp	2,700	4,600
Scaffolds	431	420
Scaffold N50	86,499,100	86,499,100
Scaffold L50	9	9
Scaffold L90	19	18
Contigs	458	456
Contig N50	39,658,000	39,597,000
Contig L50	14	15
Contig L90	36	37
QV	48.9088	67.5913
Kmer compl.	61.4815	61.7099
BUSCO sing.	96.1%	96.1%
BUSCO dupl.	0.0%	0.0%
BUSCO frag.	1.2%	1.2%
BUSCO miss.	2.7%	2.7%

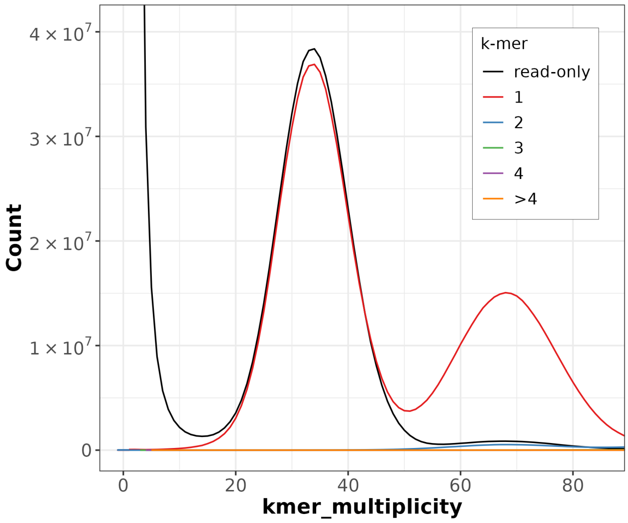
BUSCO: 5.4.3 (euk_genome_met, metaeuk) / Lineage: eukaryota_odb10 (genomes:70, BUSCOs:255)

HiC contact map of curated assembly

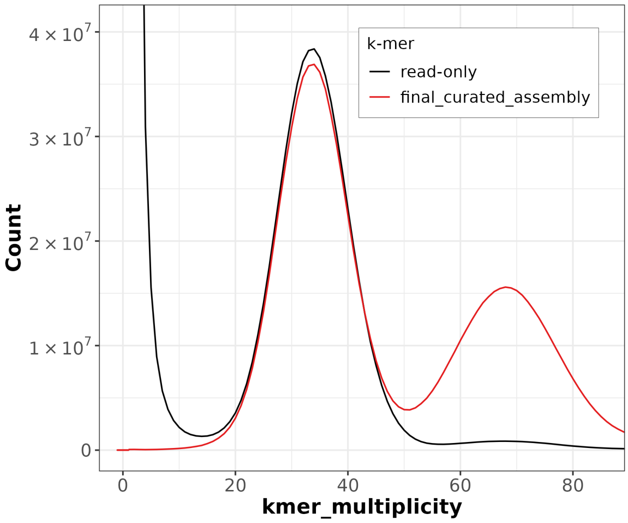


collapsed [LINK]

K-mer spectra of curated assembly

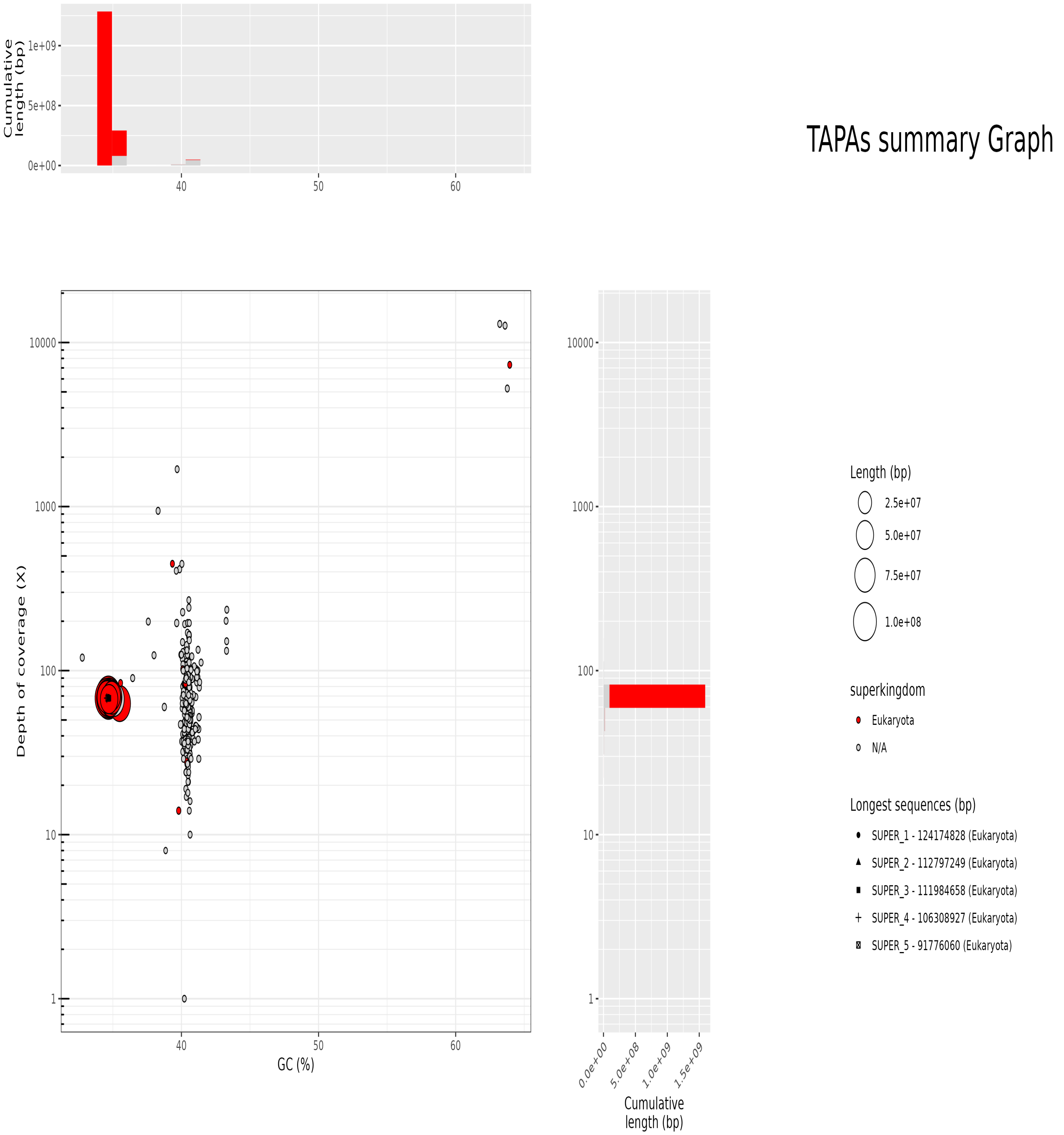


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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	Arima
Coverage	72	28

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: Emilie Teodori
Affiliation: Genoscope

Date and time: 2024-12-06 11:26:35 CET