ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	1618146
ToLID	**xgTamMedi1**
Species	Martadoris mediterranea
Class	Gastropoda
Order	Nudibranchia

Genome Traits	Expected	Observed
Haploid size (bp)	1,503,153,486	1,589,375,243
Haploid Number	12 (source: ancestor)	13
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 8.8.Q43

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

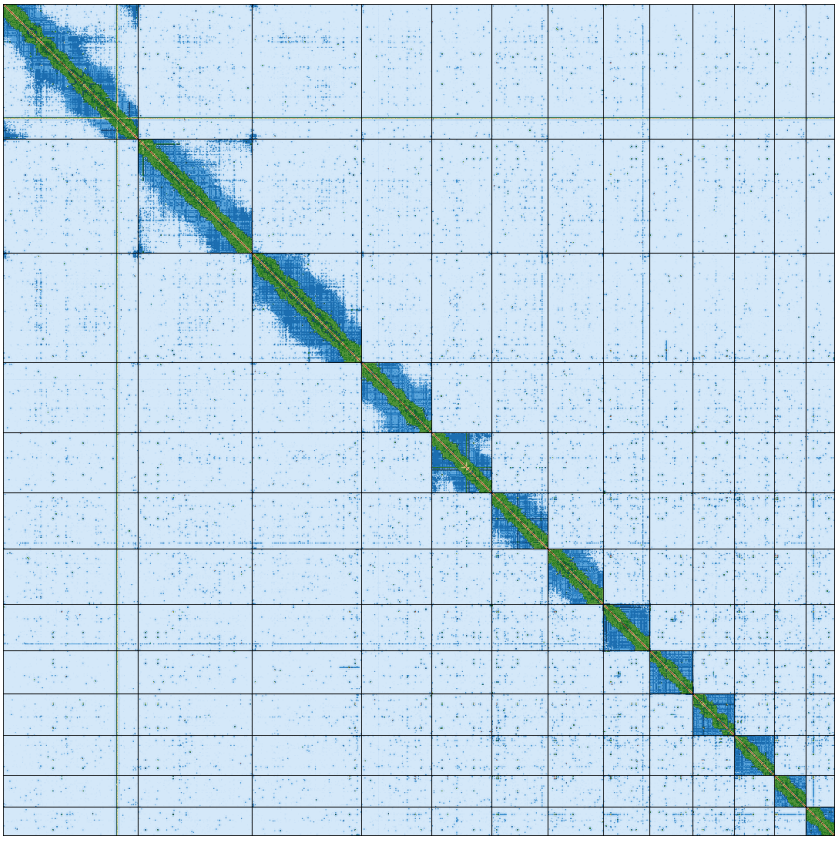
. Interventions/Gb: 6
. Contamination notes: ""
. Other observations: "The assembly of \'Martadoris mediterranea\' (xgTamMedi1) is based on 52X ONT data and 141X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial ONT assembly generation with Hifiasm, removal of contaminant sequences using Context, and Hi-C-based scaffolding with YaHS. In total, 3 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 0.052 Mb (with the largest being 0.025 Mb). Mitochondrial genome was assembled using ptGAUL. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 8 haplotypic regions were removed, totaling 34.748Mb (with the largest being 20.718Mb). 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	1,624,122,359	1,589,375,243
GC %	37.97	37.97
Gaps/Gbp	5.54	8.18
Total gap bp	900	1,900
Scaffolds	39	29
Scaffold N50	134,826,238	134,826,238
Scaffold L50	4	4
Scaffold L90	11	11
Contigs	48	42
Contig N50	108,343,644	108,343,644
Contig L50	5	5
Contig L90	15	14
QV	43.3885	43.3852
Kmer compl.	77.535	76.4736
BUSCO sing.	97.4%	95.2%
BUSCO dupl.	1.1%	0.7%
BUSCO frag.	0.5%	2.1%
BUSCO miss.	1.0%	1.9%

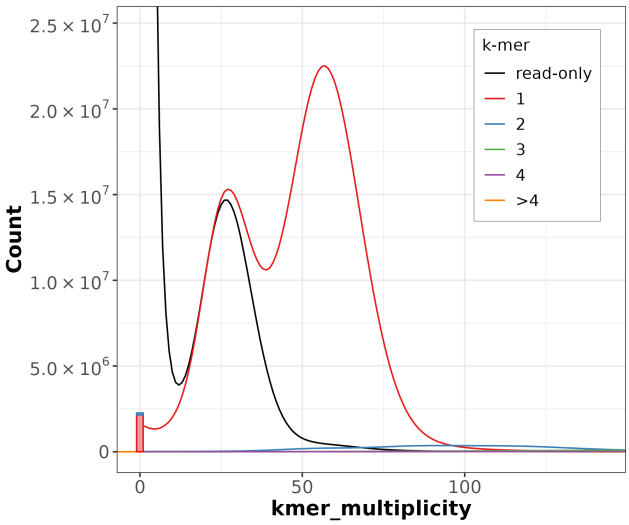
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 6.1.0 (euk_genome_min, miniprot) / Lineage: mollusca_odb12 (genomes:36, BUSCOs:4421)
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: mollusca_odb12 (genomes:36, BUSCOs:4421)

HiC contact map of curated assembly

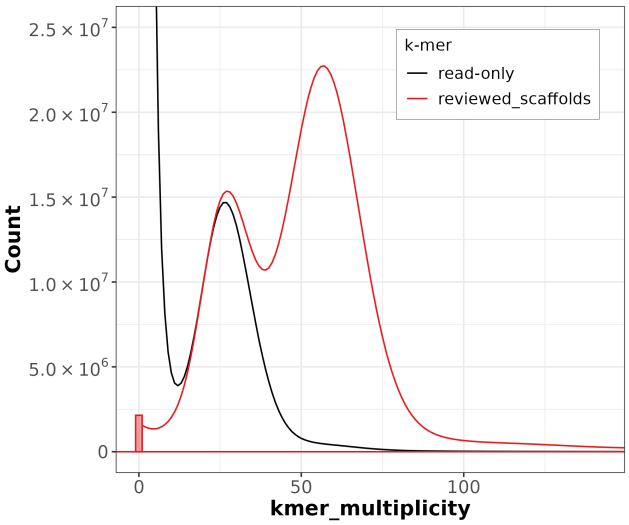


collapsed [LINK]

K-mer spectra of curated assembly

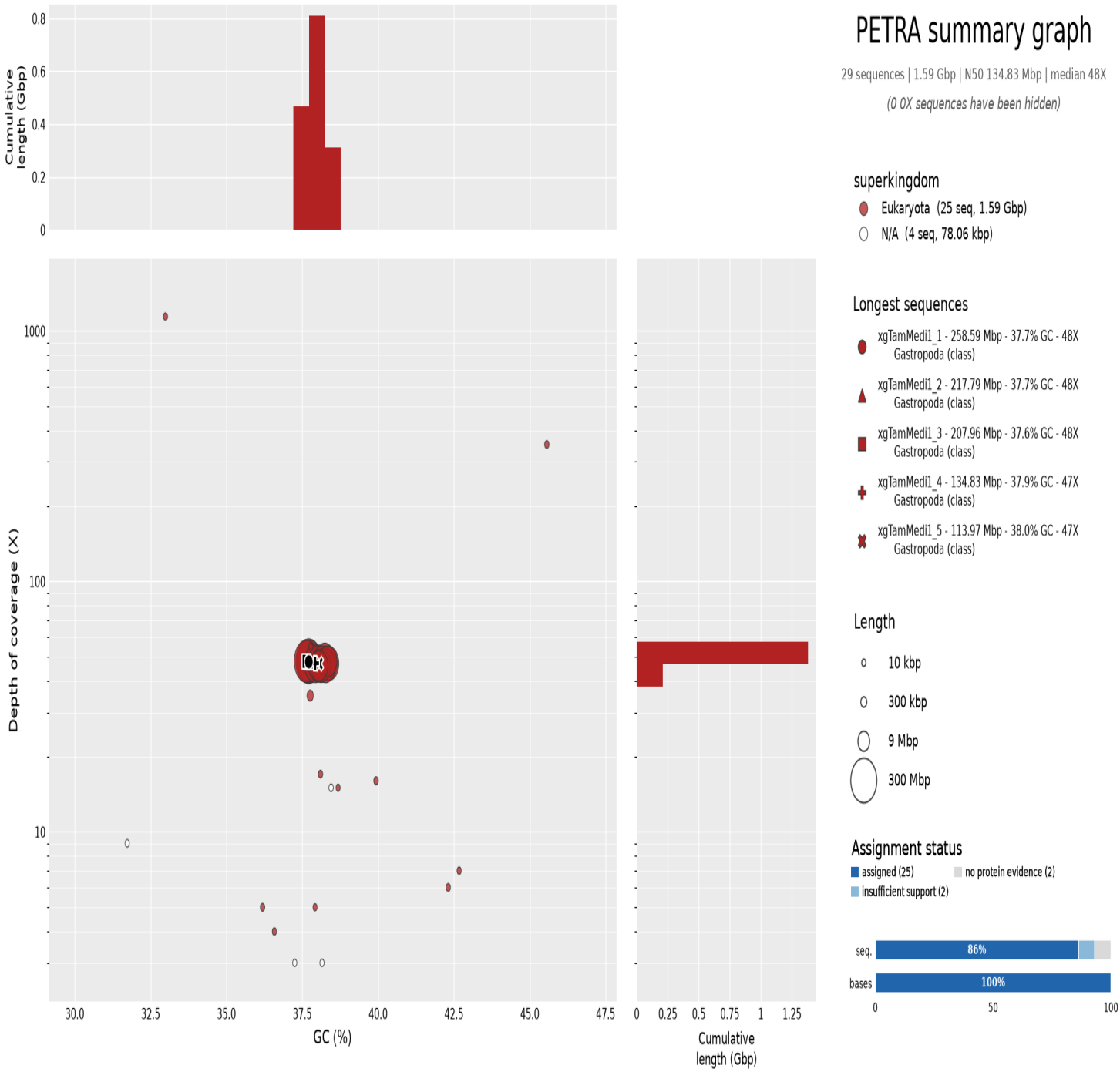


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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	53	143

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:* 0.2.4
  |_ *key param:* NA
- **PretextView**
  |_ *ver:* 1.0.7
  |_ *key param:* NA

Submitter: Simone Duprat
Affiliation: Genoscope

Date and time: 2026-09-22 21:01:09 CEST