

ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

TxID	3076427
ToLID	xgMelWern1
Species	<i>Meledella wernerii</i>
Class	Gastropoda
Order	Stylommatophora

Genome Traits	Expected	Observed
Haploid size (bp)	2,690,215,957	2,861,714,789
Haploid Number	8 (source: ancestor)	33
Ploidy	2 (source: ancestor)	2
Sample Sex	U	U

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q60

The following metrics were automatically flagged as below EBP recommended standards or different from expected:

- . Observed Haploid Number is different from Expected
- . BUSCO single copy value is less than 90% for pri
- . BUSCO duplicated value is more than 5% for pri

Curator notes

- . Interventions/Gb: 198
- . Contamination notes: "Contamination report for assembly labelled hap1; Total length of scaffolds removed: 3,013,722 (0.1 %); Scaffolds removed: 35 (5.0 %); Largest scaffold removed: (621,921); FCS-GX contaminant species (number of scaffolds; total length of scaffolds): Rhodoferax koreense, b-proteobacteria (10; 584,234); Delftia acidovorans, b-proteobacteria (3; 42,647); Aegilops tauschii, plants (2; 344,373); Candidatus Vicinibacter proximus, CFB group bacteria (1; 621,921); Spiroplasma endosymbiont of Danaus chrysippus, mycoplasmas (1; 576,126) Chitinophaga sp., CFB group bacteria (1; 99,714); Parafilimonas sp., CFB group bacteria (1; 85,157); Verminephrobacter aporrectodeae, b-proteobacteria (1; 70,053); Niastella yeongjuensis, CFB group bacteria (1; 50,837); Ginsengibacter hankyongi, CFB group bacteria (1; 50,722); Acidovorax avenae, b-proteobacteria (1; 49,680); Elizabethkingia miricola, CFB group bacteria (1; 47,212); Variovorax boronicumulans, b-proteobacteria (1; 40,758); Longitalea luteola, CFB group bacteria (1; 40,434); Pseudomonas reactans, g-proteobacteria (1; 35,923); Sulfurimicrobium lacus, b-proteobacteria (1; 34,827); Limnohabitans parvus, b-proteobacteria (1; 32,480); Pseudorhodoferax aquiterrae, b-proteobacteria (1; 3,000); Adhaeribacter aerolatus,

CFB group bacteria (1; 2,000); Pedobacter segetis, CFB group bacteria (1; 2,000); Polaromonas jejuensis, b-proteobacteria (1; 2,000); Aureibaculum luteum, CFB group bacteria (1; 1,000); Mitochondrion (1; 196,624)"

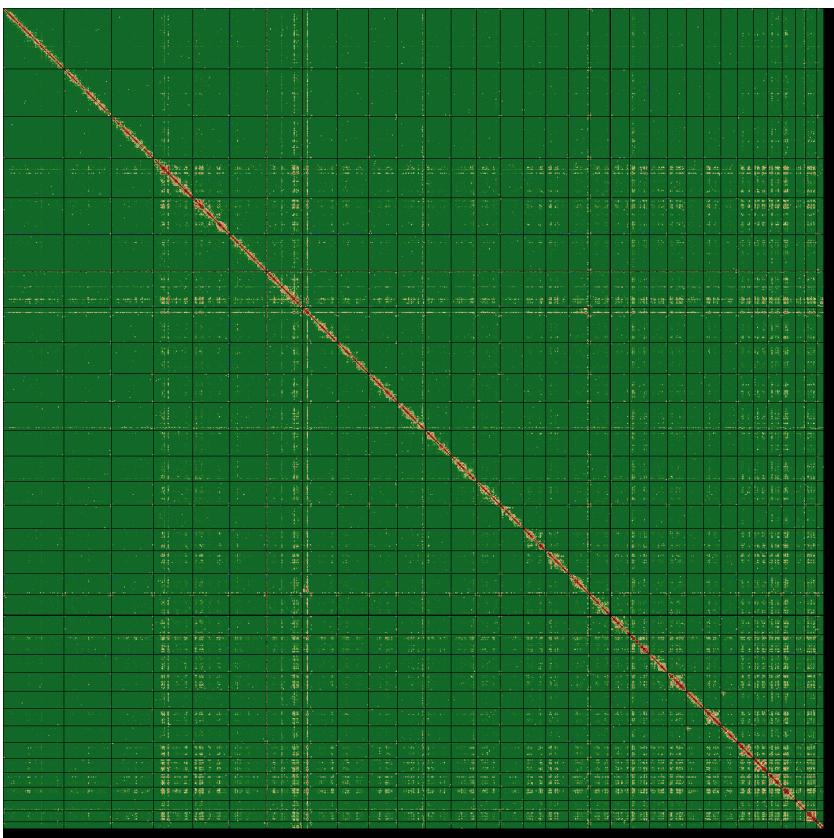
. Other observations: "Assembly was Hi-C phased; The exact order and orientation of the contigs on chromosome 16 (48.3-50.3 Mbp) and chromosome 21 (19.2-28.7 Mbp) is unknown."

Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	2,863,696,462	2,861,714,789
GC %	40.52	40.52
Gaps/Gbp	470.37	498.3
Total gap bp	134,700	150,600
Scaffolds	697	598
Scaffold N50	68,378,180	97,597,643
Scaffold L50	14	11
Scaffold L90	38	27
Contigs	2,044	2,024
Contig N50	3,896,000	3,893,402
Contig L50	221	223
Contig L90	760	769
QV	60.9	60.9
Kmer compl.	99.40	99.31
BUSCO sing.	75.6%	75.3%
BUSCO dupl.	19.8%	20.0%
BUSCO frag.	0.5%	0.5%
BUSCO miss.	4.2%	4.2%

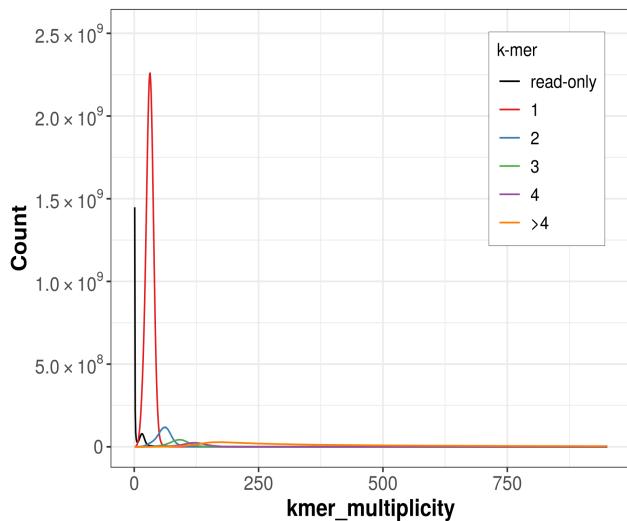
BUSCO 6.0.0 Lineage: mollusca_odb10 (genomes:7, BUSCOs:5295)

HiC contact map of curated assembly

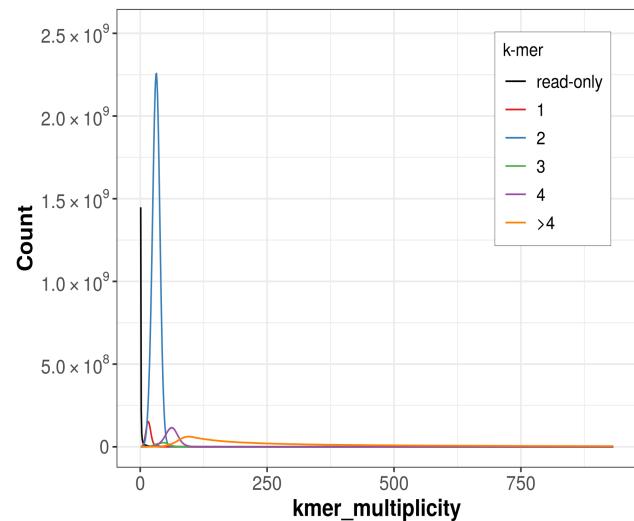


pri [LINK]

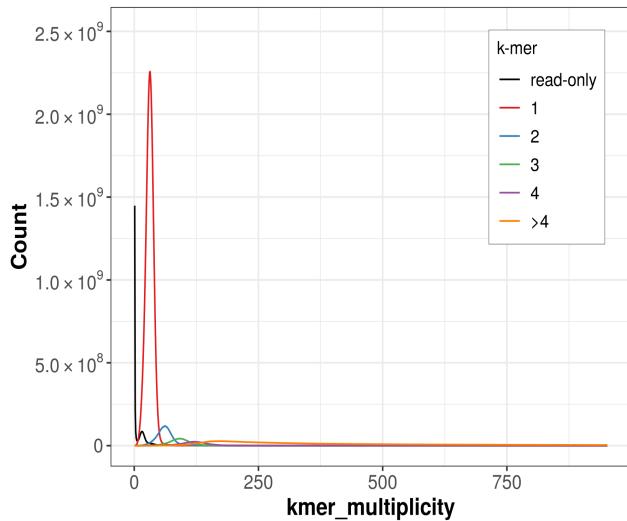
K-mer spectra of curated assembly



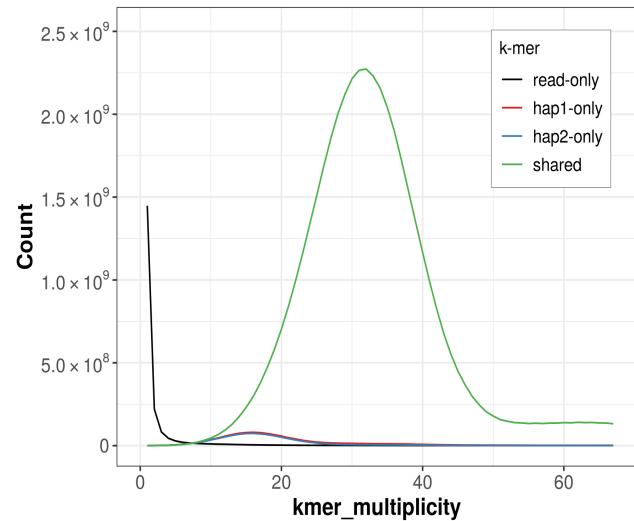
Distribution of k-mer counts per copy numbers found in **hap1** (hapl.)



Distribution of k-mer counts per copy numbers found in **asm** (diplo.)

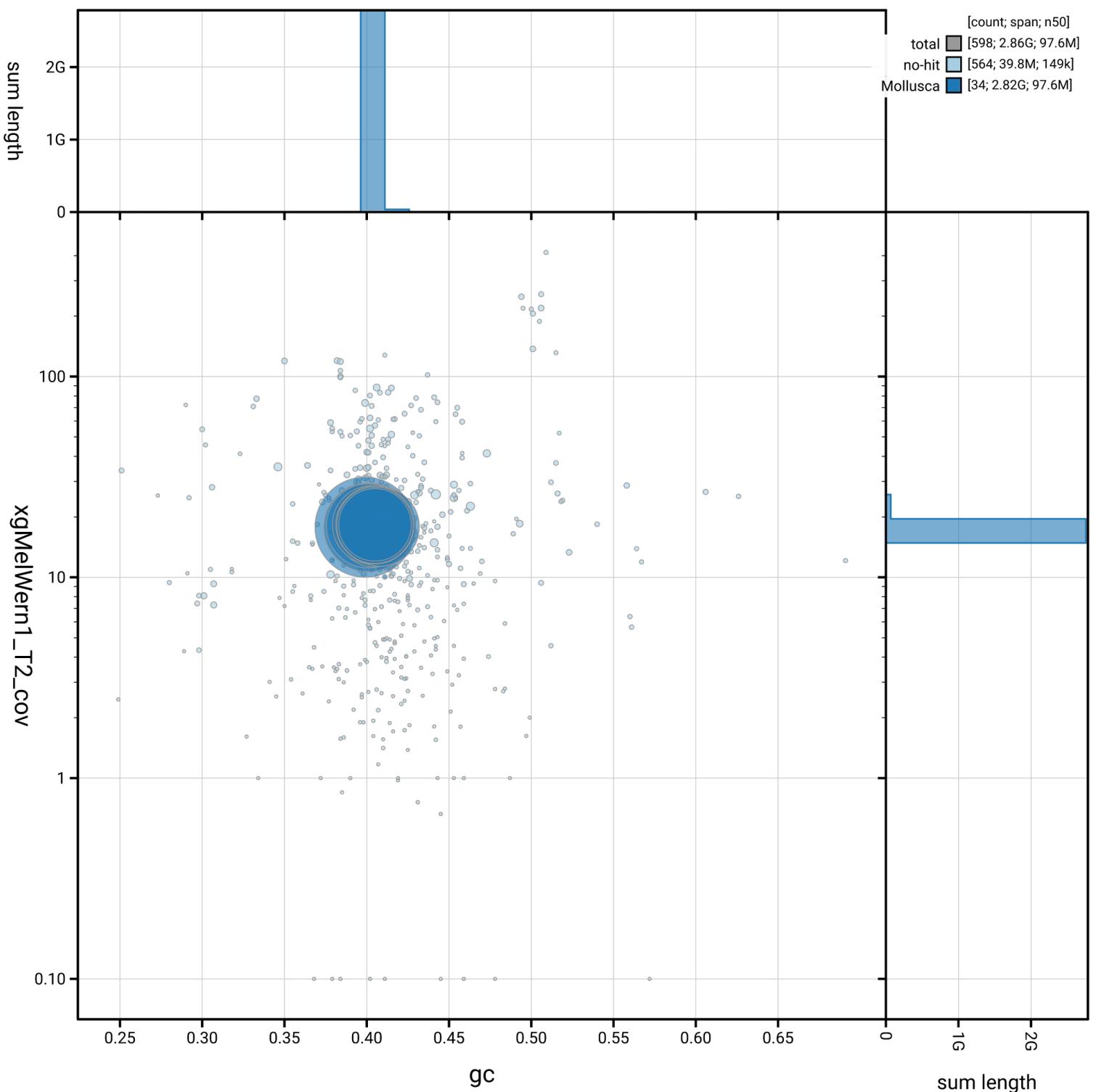


Distribution of k-mer counts per copy numbers found in **hap2** (hapl.)



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

Post-curation contamination screening



pri. 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 v2
Coverage	31x	236x

Assembly pipeline

```
- hifiasm
  |_ ver: 0.25.0-r726
  |_ key param: --h1/--h2
- yahs
  |_ ver: 1.2.2
  |_ key param: NA
```

Curation pipeline

```
- hifiasm
  |_ ver: 0.25.0-r726
  |_ key param: --h1/--h2
- yahs
  |_ ver: 1.2.2
  |_ key param: NA
- TreeVal
  |_ ver: 1.2.2
  |_ key param: NA
```

Submitter: Michael Paulini

Affiliation: WSI

Date and time: 2025-11-21 16:35:44 CET