

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	307678
ToLID	<b>tnTetMela7</b>
Species	tetrastemma melanocephalum
Class	Enopla
Order	Monostilifera

Genome Traits	Expected	Observed
Haploid size (bp)	483,581,228	524,234,113
Haploid Number	2 (source: ancestor)	16
Ploidy	2 (source: ancestor)	2
Sample Sex	unknown	NA

## 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
- . Observed sex is different from Sample sex
- . Kmer completeness value is less than 90 for pri
- . Assembly length loss > 3% for pri
- . More than 1000 gaps/Gbp for pri

### Curator notes

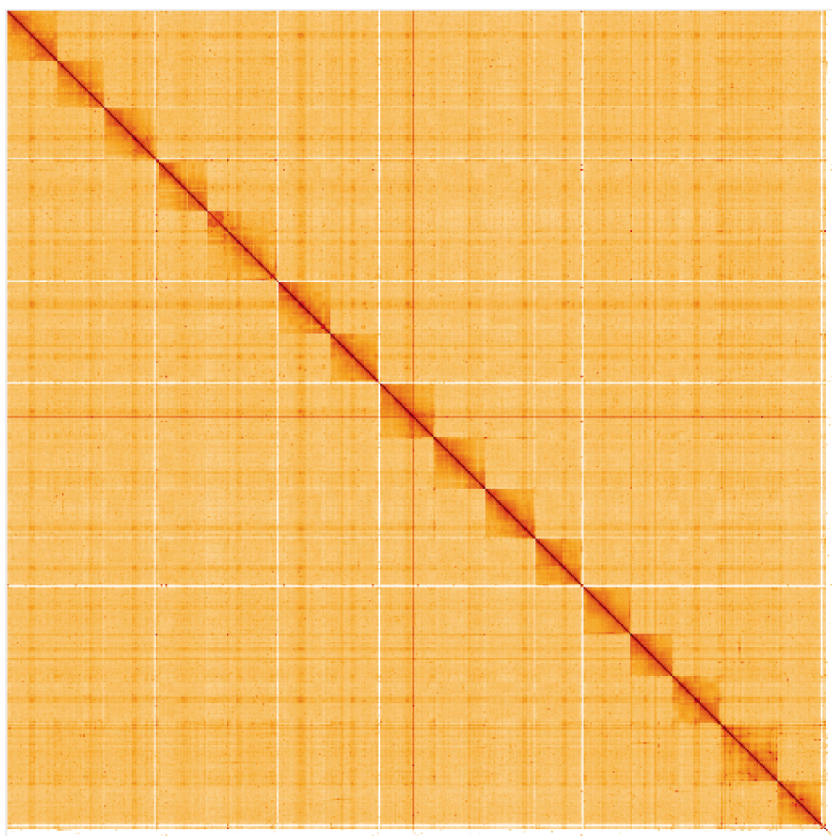
- . Interventions/Gb: many
- . Contamination notes: "There were about 170 proteobacteria sequences totalling 3Mb identified by FCS-GX and then a further 21 scaffolds removed by blobtools corresponding to various fungi and bacteria totalling a further 3Mb. The swissprot databases seems to classify a number of scaffolds as arthropod, likely due to the lack of sequences in this lineage."
- . Other observations: "There were a lot of retained haplotigs in this assembly, likely due to the high heterozygosity rate of 4%. Hi-C was from another individual, so the hifi+hic assembly was not possible. Just over 30Mb of sequence was manually removed as haplotig."

# Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	563,854,870	524,234,113
GC %	44.69	44.71
Gaps/Gbp	858.38	1,085.39
Total gap bp	96,800	101,100
Scaffolds	391	141
Scaffold N50	34,110,936	32,051,925
Scaffold L50	7	8
Scaffold L90	14	15
Contigs	875	709
Contig N50	1,676,826	1,515,124
Contig L50	108	106
Contig L90	336	332
QV	59.6832	60.852
Kmer compl.	64.0939	61.0453
BUSCO sing.	90.7%	93.8%
BUSCO dupl.	5.8%	2.2%
BUSCO frag.	1.0%	0.9%
BUSCO miss.	2.6%	3.1%

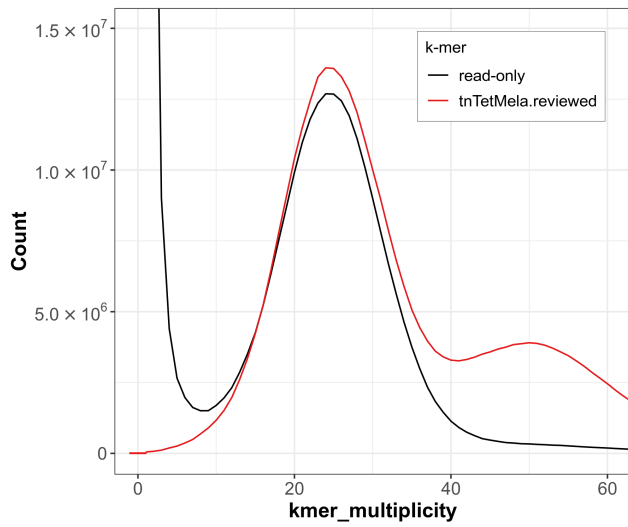
BUSCO: 6.0.0 (euk\_genome\_min, miniprot) / Lineage: lophotrochozoa\_odb12 (genomes:75, BUSCOs:1252)

# HiC contact map of curated assembly

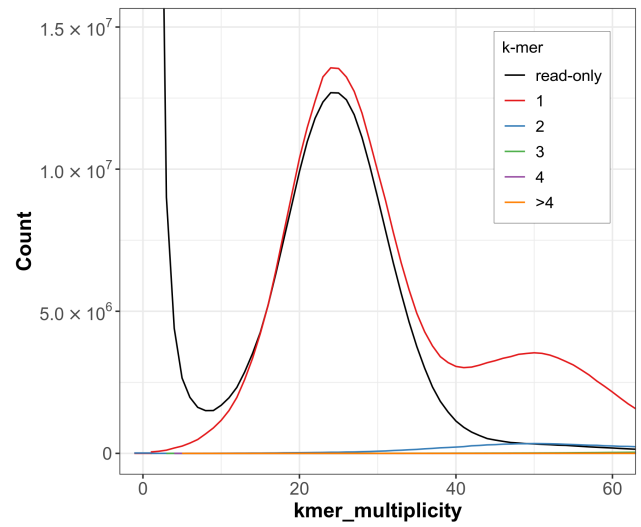


pri [\[LINK\]](#)

# K-mer spectra of curated assembly

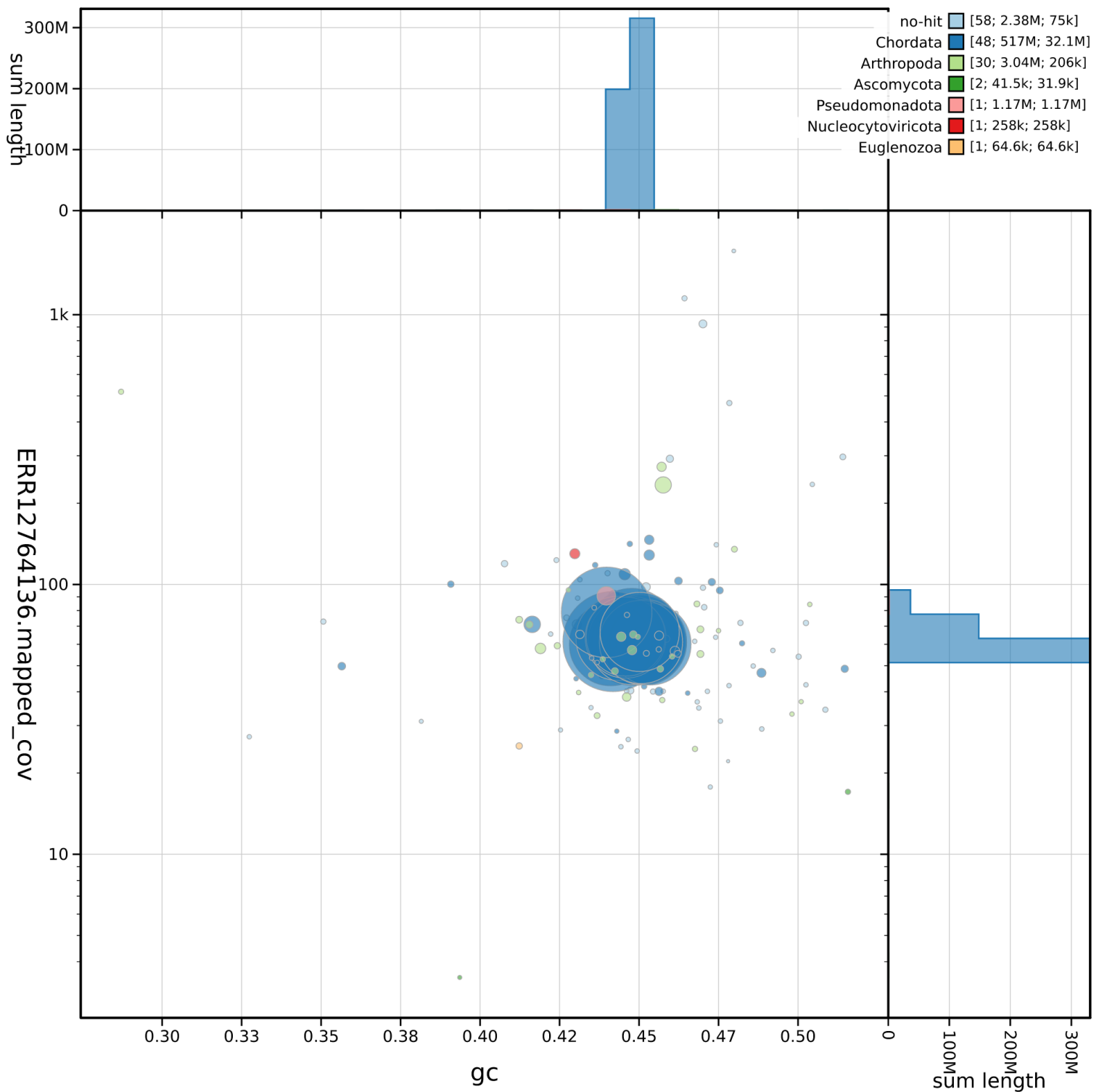


Distribution of k-mer counts coloured by their presence in reads/assemblies



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

# 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	Hi-C
Coverage	49	190

# Assembly pipeline

- **hifiasm**
  - |\_ *ver*: 0.25.0-r726
  - |\_ *key param*: -l 3
  - |\_ *key param*: --primary
- **purge-dups**
  - |\_ *ver*: 1.2.5
  - |\_ *key param*: NA
- **yahs**
  - |\_ *ver*: 1.2a.1
  - |\_ *key param*: NA
- **fcs-gx**
  - |\_ *ver*: 0.5.0
  - |\_ *key param*: NA
- **blobtools**
  - |\_ *ver*: 4.0.7
  - |\_ *key param*: NA

# Curation pipeline

- **PretextViewAI**
  - |\_ *ver*: 1.0.0
  - |\_ *key param*: NA
- **HiGlass**
  - |\_ *ver*: 1.8.0
  - |\_ *key param*: NA

Submitter: Tom Brown

Affiliation: IZW

Date and time: 2025-10-10 09:51:51 CEST