

# ERGA Assembly Report

v24.04.03\_beta

Tags: ERGA-BGE

TxID	171369
ToLID	<b>qcEupGros1</b>
Species	Eupolybothrus grossipes
Class	Chilopoda
Order	Lithobiomorpha

Genome Traits	Expected	Observed
Haploid size (bp)	6,577,503,001	6,577,898,879
Haploid Number	20 (source: ancestor)	14
Ploidy	2 (source: ancestor)	2
Sample Sex	F	F

## EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.8.Q65

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

- . Observed Haploid Number is different from Expected

### Curator notes

- . Interventions/Gb: 168
- . Contamination notes: "Contamination report for assembly labelled hap1; Total length of scaffolds removed: 665,591 (0.0 %); Scaffolds removed: 34 (1.6 %); Largest scaffold removed: (57,401); FCS-GX contaminant species (number of scaffolds; total length of scaffolds):; Rickettsia endosymbiont of Cimex lectularius, a-proteobacteria (2; 44,092); Rickettsia endosymbiont of Bradysia coprophila, a-proteobacteria (1; 57,401); Besnoitia besnoiti, alveolates (1; 9,369); Mitochondrion (30; 554,729); FCS-Adaptor (1; 38)"
- . Other observations: "Hifiiasm assembly run in Hi-C phasing mode; The exact order and orientation of the contigs on chromosome 1 (813,203 - 818,222 Kbp) are unknown. The exact order and orientation of the contigs on chromosome 2 (22,884 - 28,304 & 89,329 - 103,1c81 Kbp) are unknown. The exact order and orientation of the contigs on chromosome 3 (548,024 - 553,243 & 665,056 - 669,473 Kbp) are unknown. The exact order and orientation of the contigs on chromosome 4 (359,929 - 371,974 & 389,840 - 400,880 & 663,049 - 668,667 Kbp) are unknown. The exact order and orientation of the contigs on chromosome 7 (275,015 - 280,435 Kbp) are unknown. The exact order and orientation of the contigs on chromosome 8 (50,988 - 55,806 Kbp) are unknown. The exact order and orientation of the contigs on chromosome 9 (271,804 - 275,216 Kbp) are unknown. The exact order and orientation of the contigs on chromosome 10 (64,638 - 75,679 Kbp) are

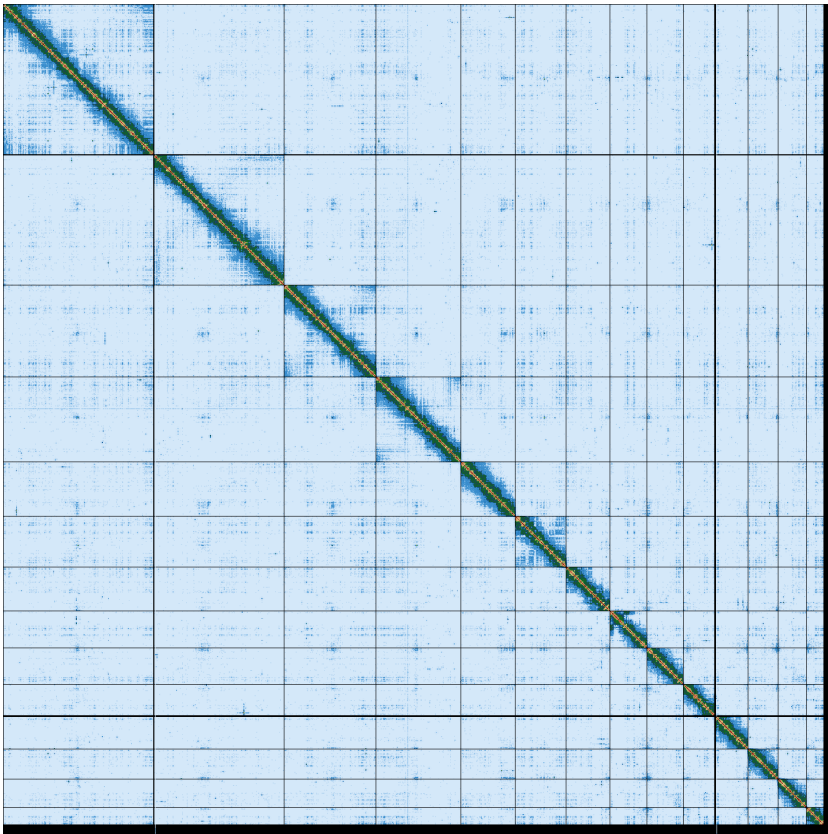
unknown. The exact order and orientation of the contigs on chromosome 13 (176,451 - 181, 871 & g402,312 - 406,943 Kbp) are unknown"

# Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	6,709,770,180	6,577,898,879
GC %	40.61	40.63
Gaps/Gbp	543.68	575.26
Total gap bp	364,800	419,038
Scaffolds	2,106	1,445
Scaffold N50	653,174,762	669,113,179
Scaffold L50	4	4
Scaffold L90	13	12
Contigs	5,754	5,229
Contig N50	3,308,181	3,235,202
Contig L50	604	609
Contig L90	2,143	2,108
QV	65.4	65.4
Kmer compl.	99.09	99.08
BUSCO sing.	96.8%	96.9%
BUSCO dupl.	1.6%	1.5%
BUSCO frag.	0.2%	0.4%
BUSCO miss.	1.4%	1.2%

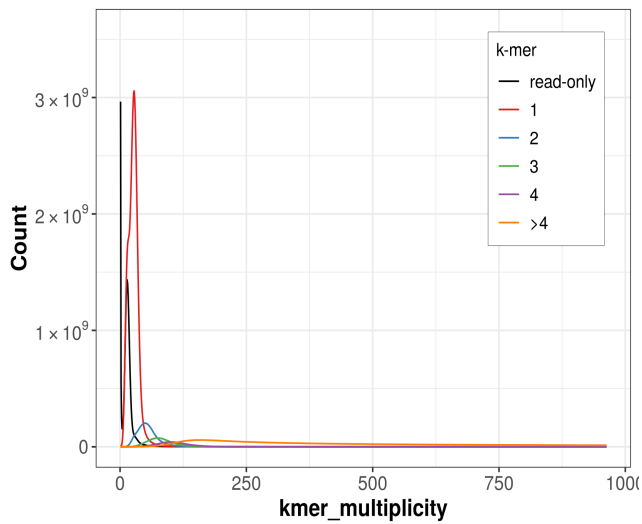
BUSCO 6.0.0 Lineage: arthropoda\_odb10 (genomes:90, BUSCOs:1013)

# HiC contact map of curated assembly

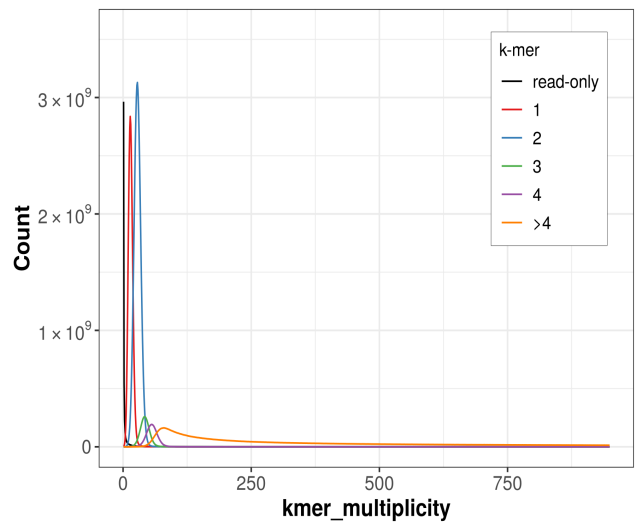


pri [\[LINK\]](#)

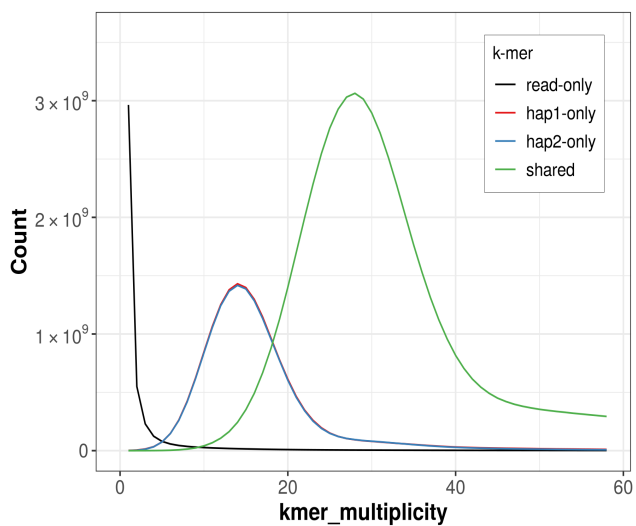
# K-mer spectra of curated assembly



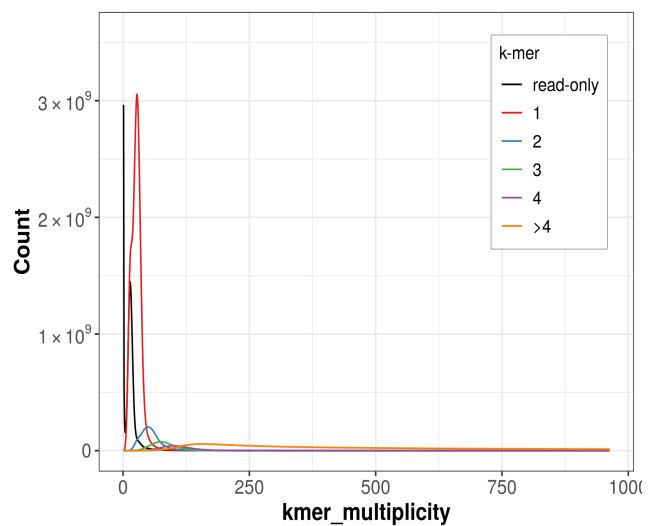
Distribution of k-mer counts per copy numbers found in **hap1** (hap1.)



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

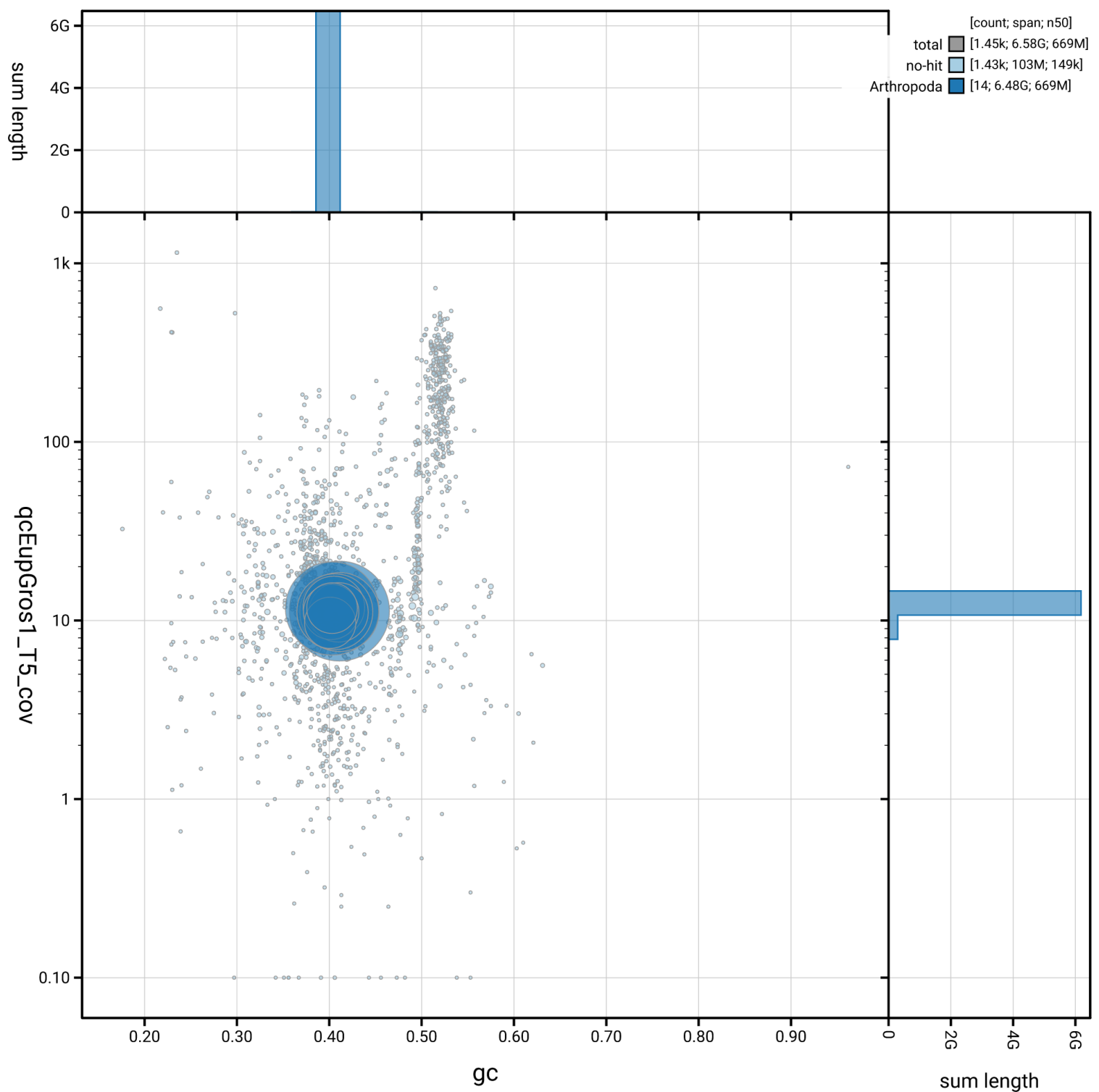


Distribution of k-mer counts coloured by their presence in reads/assemblies



Distribution of k-mer counts per copy numbers found in **hap2** (hap1.)

# 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	27x	26x

# 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*: 1.2.1
  - |\_ *key param*: NA

Submitter: Karen Brooks

Affiliation: WSI

Date and time: 2026-01-13 19:20:46 CET