

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	2936736
ToLID	<b>qqTroKord2</b>
Species	Troglohyphantes kordunlikanus
Class	Arachnida
Order	Araneae

Genome Traits	Expected	Observed
Haploid size (bp)	970,163,739	649,380,381
Haploid Number	12 (source: ancestor)	0
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	NA

## EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 4.4.Q51

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

- . Observed Haploid size (bp) has >20% difference with Expected
- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for hap1
- . BUSCO single copy value is less than 90% for hap1
- . Not 90% of assembly in chromosomes for hap1

### Curator notes

- . Interventions/Gb: NA
- . Contamination notes: "Contaminants were first removed with NCBI FCS and then with a diamond and Blobtoolkit identification. Only the scaffold that were confirmed as contaminant were removed"
- . Other observations: "This is a contig level assembly. Despite the scaffolding the N50 was too low to proceed with the curation."

# Quality metrics table

Metrics	Pre-curation hap1	Curated hap1
Total bp	649,380,381	649,380,381
GC %	31.84	31.84
Gaps/Gbp	0	0
Total gap bp	0	0
Scaffolds	11,362	11,362
Scaffold N50	81,759	81,759
Scaffold L50	2,395	2,395
Scaffold L90	8,042	8,042
Contigs	11,362	11,362
Contig N50	81,759	81,759
Contig L50	2,395	2,395
Contig L90	8,042	8,042
QV	51.6918	51.6918
Kmer compl.	73.1792	73.1792
BUSCO sing.	69.35%	69.35%
BUSCO dupl.	1.83%	1.83%
BUSCO frag.	4.46%	4.46%
BUSCO miss.	24.36%	24.36%

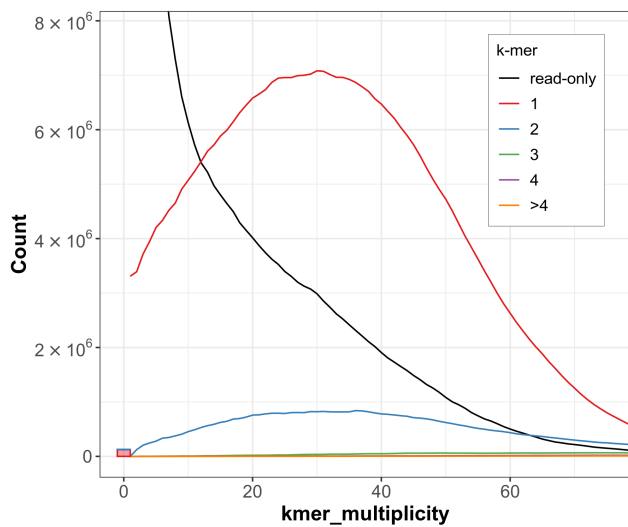
BUSCO: 5.8.0 (euk\_genome\_met, metaeuk) / Lineage: insecta\_odb10 (genomes:10, BUSCOs:2934)

# HiC contact map of curated assembly

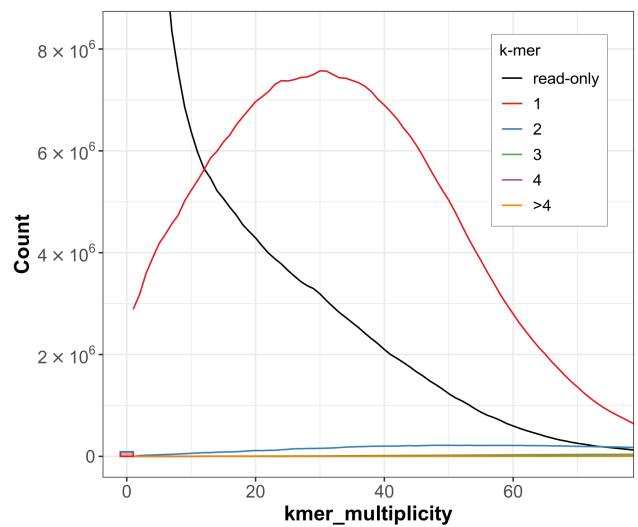
**hap1** HiC PNG is missing!

**hap1** File link is missing!

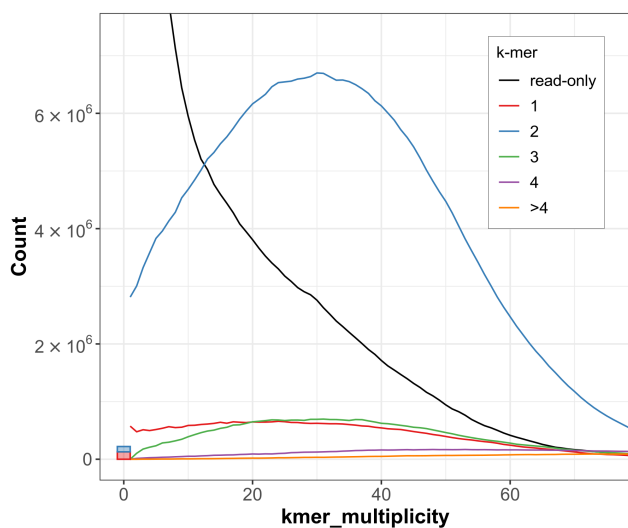
# K-mer spectra of curated assembly



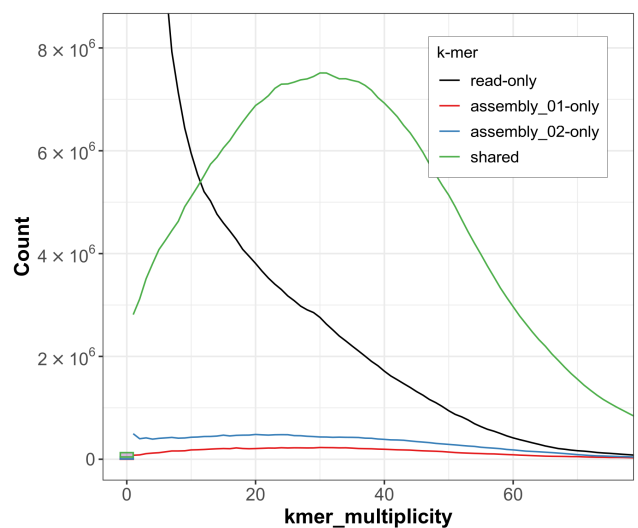
Distribution of k-mer counts per copy numbers found in **assembly\_02** (hapl.)



Distribution of k-mer counts per copy numbers found in **assembly\_01** (hapl.)

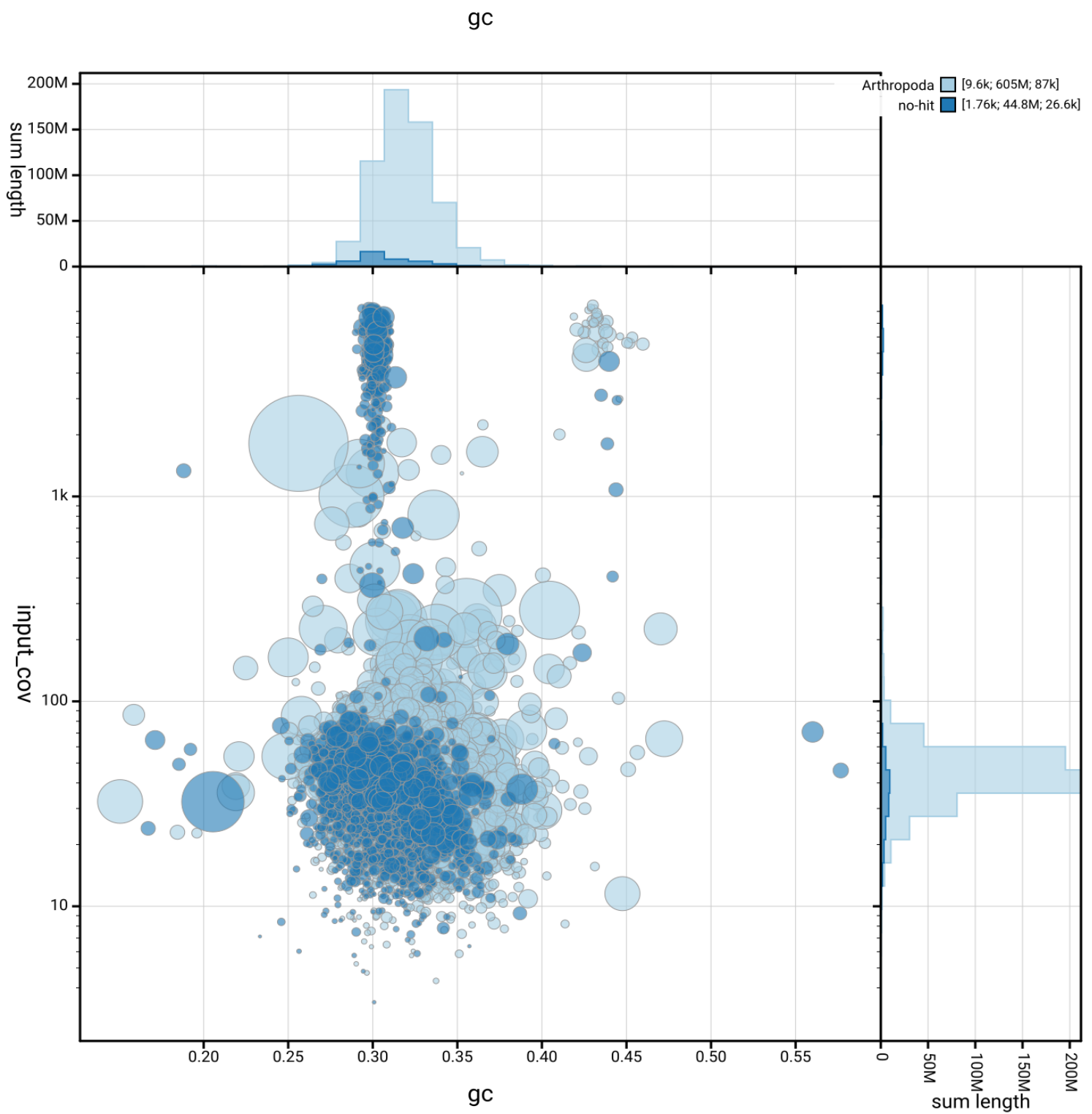


Distribution of k-mer counts per copy numbers found in **asm** (dipl.)



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

# Post-curation contamination screening



**hap1.** 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	HiFi	HiC
Coverage	29.13x	51.48x

## Assembly pipeline

- **Hifiasm**
  - |\_ *ver*: Galaxy Version 0.25.0+galaxy0
  - |\_ *key param*: VGP Galaxy workflow VGP3 v 0.3.4
- **purge\_dups**
  - |\_ *ver*: Galaxy Version 1.2.6+galaxy0
  - |\_ *key param*: VGP Galaxy workflow VGP6 v 0.10.4
- **YaHS**
  - |\_ *ver*: Galaxy Version 1.2a.2+galaxy3
  - |\_ *key param*: VGP Galaxy workflow v 3.3

## Curation pipeline

Invalid pipeline data format

Submitter: Marcella Sozzoni  
Affiliation: UNIFI

Date and time: 2026-02-20 11:14:21 CET