ERGA Assembly Report
v24.10.15

Tags: ERGA-BGE

TxID	3237715
ToLID	**qqEpiCave3**
Species	Episinus cavernicola
Class	Arachnida
Order	Araneae

Genome Traits	Expected	Observed
Haploid size (bp)	947,168,479	1,401,245,227
Haploid Number	11 (source: ancestor)	0
Ploidy	2 (source: ancestor)	2
Sample Sex	ZW	NA

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 4.4.Q47

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
. Observed sex is different from Sample sex

. 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: "Decontamination was performed using NCBI FCS were maily proteobacteria contamiants were removed. Another step of decontamination was done based on diamond scores and blobtoolkit identification, it this step 1.3Gb of contaminants were removed."
. Other observations: "This is a contig level assembly. Despite the scaffolding the N50 is really low and is not possibile to procede with the curation"

Quality metrics table

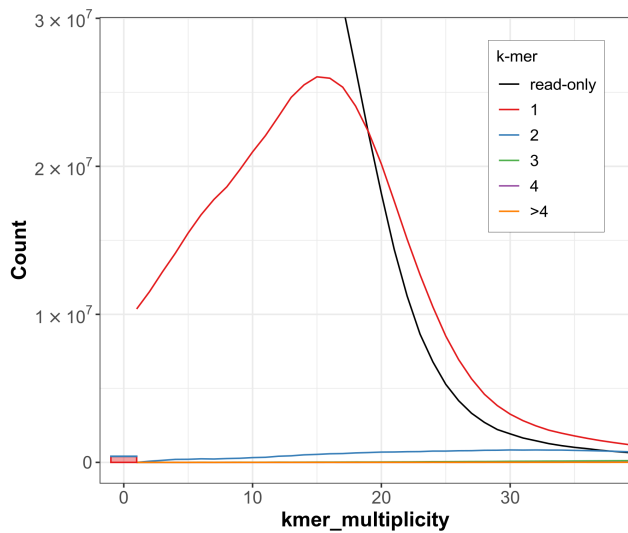| Metrics	Pre-curation hap1	Curated hap1
Total bp	1,401,245,227	1,401,245,227
GC %	30.14	30.14
Gaps/Gbp	0	0
Total gap bp	0	0
Scaffolds	34,513	34,513
Scaffold N50	47,451	47,451
Scaffold L50	7,577	7,577
Scaffold L90	26,147	26,147
Contigs	34,513	34,513
Contig N50	47,451	47,451
Contig L50	7,577	7,577
Contig L90	26,147	26,147
QV	47.2944	47.2944
Kmer compl.	95.8517	95.8517
BUSCO sing.	64.8%	64.8%
BUSCO dupl.	2.5%	2.5%
BUSCO frag.	4.7%	4.7%
BUSCO miss.	28.1%	28.1%

BUSCO: 5.8.0 (euk_genome_min, miniprot) / Lineage: arachnida_odb10 (genomes:10, BUSCOs:2934)
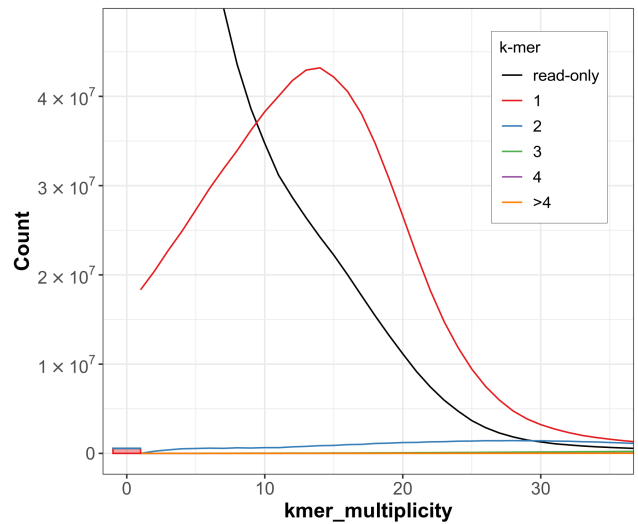
HiC contact map of curated assembly

hap1 HiC PNG is missing!
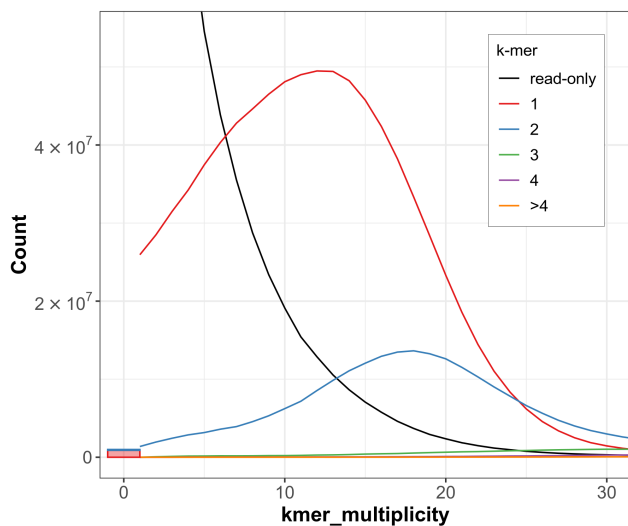
hap1 File link is missing!

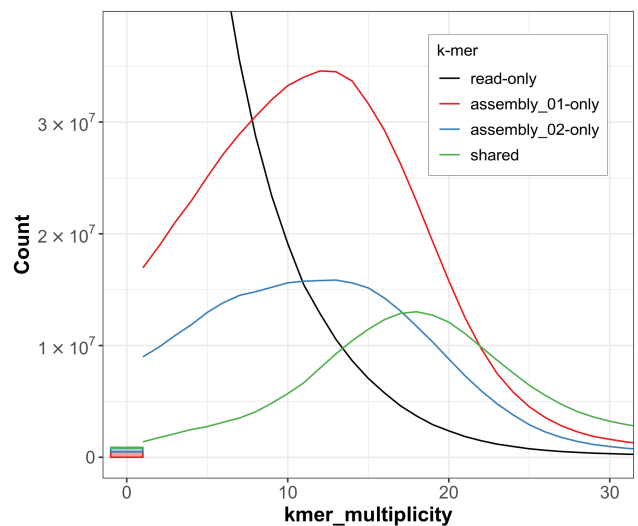# K-mer spectra of curated assembly

Distribution of k-mer counts per copy
numbers found in **assembly_02** (hapl.)

Distribution of k-mer counts per copy
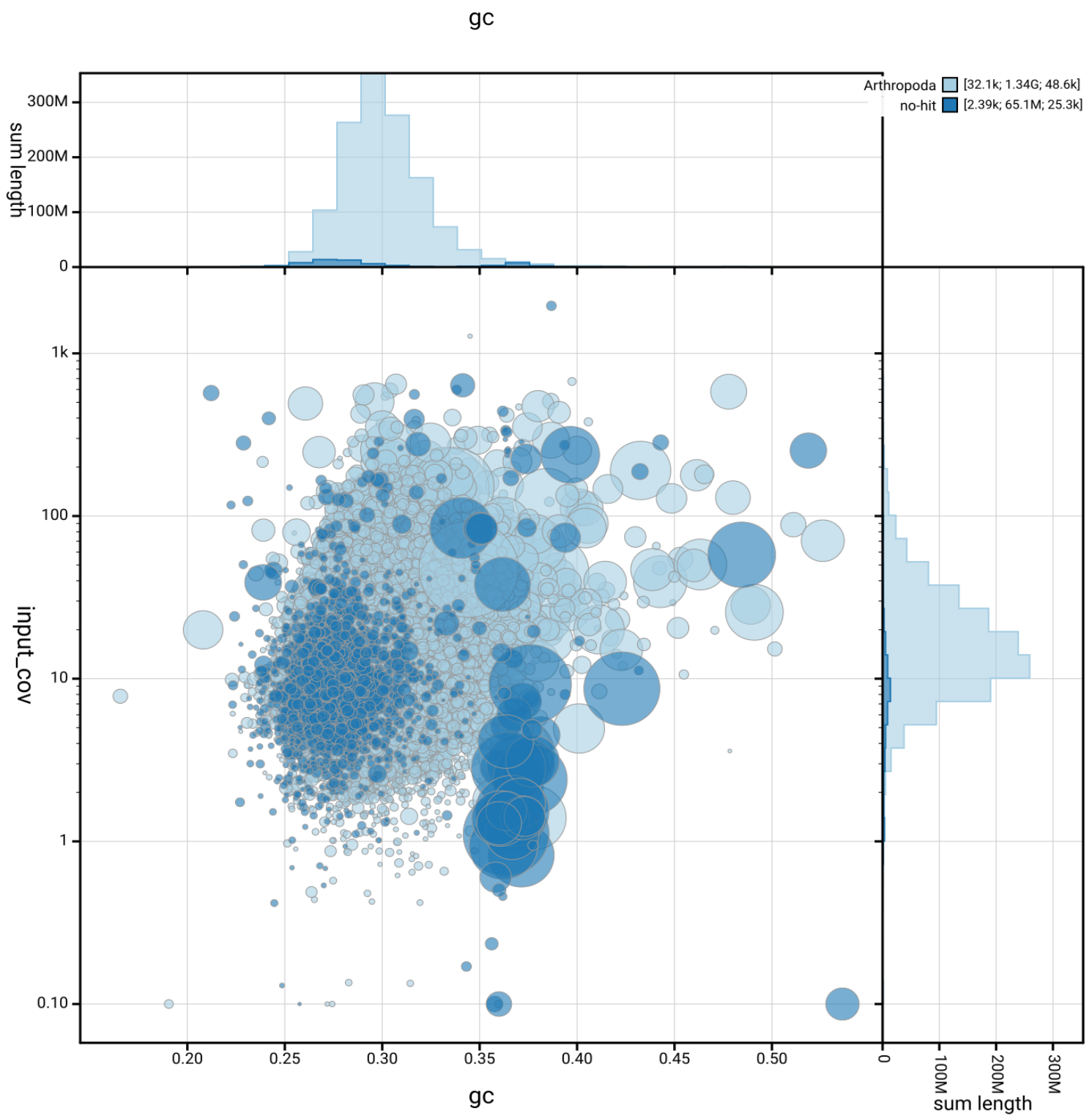numbers found in **assembly_01** (hapl.)

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	24.61x	54.19x

Assembly pipeline

- **Hifiasm**
    - |_ *ver:* Galaxy Version 0.25.0+galaxy0
    - |_ *key param:* VGP Galaxy workflow VGP4 v 0.5
- **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 VGP8 v 3.3

Curation pipeline

Invalid pipeline data format

Submitter: Marcella Sozzoni
Affiliation: UNIFI

Date and time: 2026-02-19 10:40:21 CET