

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	48712
ToLID	qlOrcRanz
Species	Orchesella ranzii
Class	Collembola
Order	Entomobryomorpha

Genome Traits	Expected	Observed
Haploid size (bp)	244,559,092	257,022,770
Haploid Number	5 (source: ancestor)	6
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	NA

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 6.7.Q61

Obtained EBP quality metric for hap2: 6.7.Q62

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

- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for hap1
- . Kmer completeness value is less than 90 for hap2

Curator notes

- . Interventions/Gb: None
- . Contamination notes: "No contaminants were found"
- . Other observations: "None"

Quality metrics table

Metrics	Pre-curation hap1	Pre-curation hap2	Curated hap1	Curated hap2
Total bp	253,445,765	262,411,437	254,309,122	257,022,770
GC %	36.07	36.11	36.06	36.08
Gaps/Gbp	382.72	438.24	640.95	614.73
Total gap bp	19,400	23,000	32,600	31,600
Scaffolds	119	98	33	35
Scaffold N50	37,959,712	40,257,473	41,544,551	41,617,015
Scaffold L50	3	3	3	3
Scaffold L90	6	6	6	6
Contigs	216	213	196	193
Contig N50	3,944,353	3,719,052	3,897,879	3,719,052
Contig L50	15	19	16	18
Contig L90	66	80	76	83
QV	61.385	61.6257	61.7509	62.3783
Kmer compl.	83.2072	83.9955	83.6314	83.8139
BUSCO sing.	91.0%	90.0%	91.9%	91.7%
BUSCO dupl.	4.0%	4.8%	3.6%	4.4%
BUSCO frag.	1.9%	2.0%	1.6%	1.2%
BUSCO miss.	3.1%	3.2%	2.9%	2.7%

Warning! BUSCO versions or lineage datasets are not the same across results:

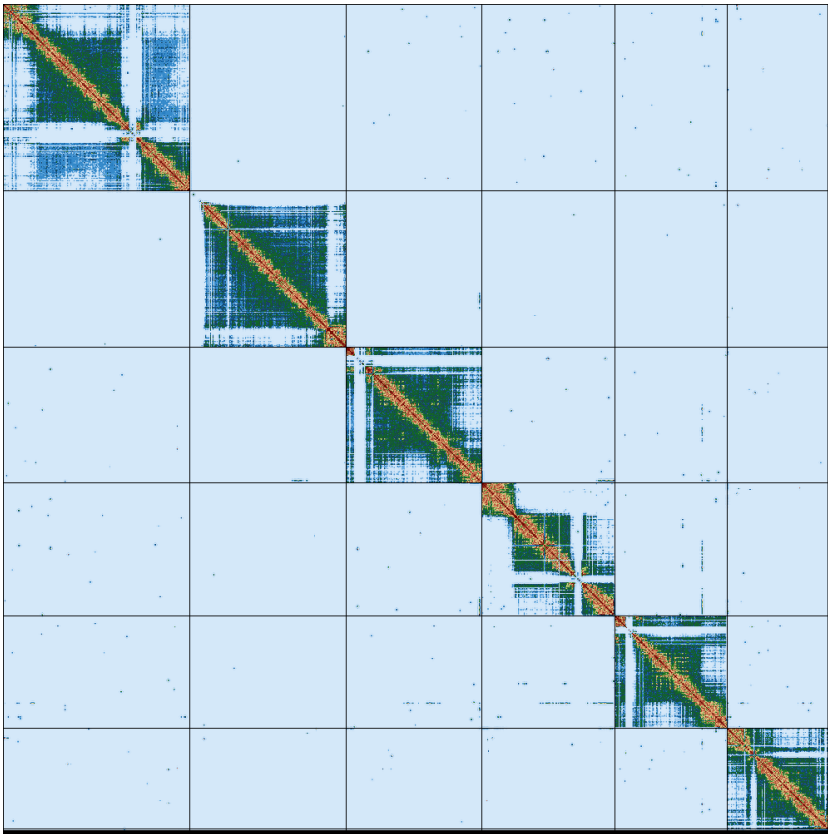
BUSCO: 5.8.0 (euk_genome_met, metaeuk) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

BUSCO: 5.8.0 (euk_genome_met, metaeuk) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

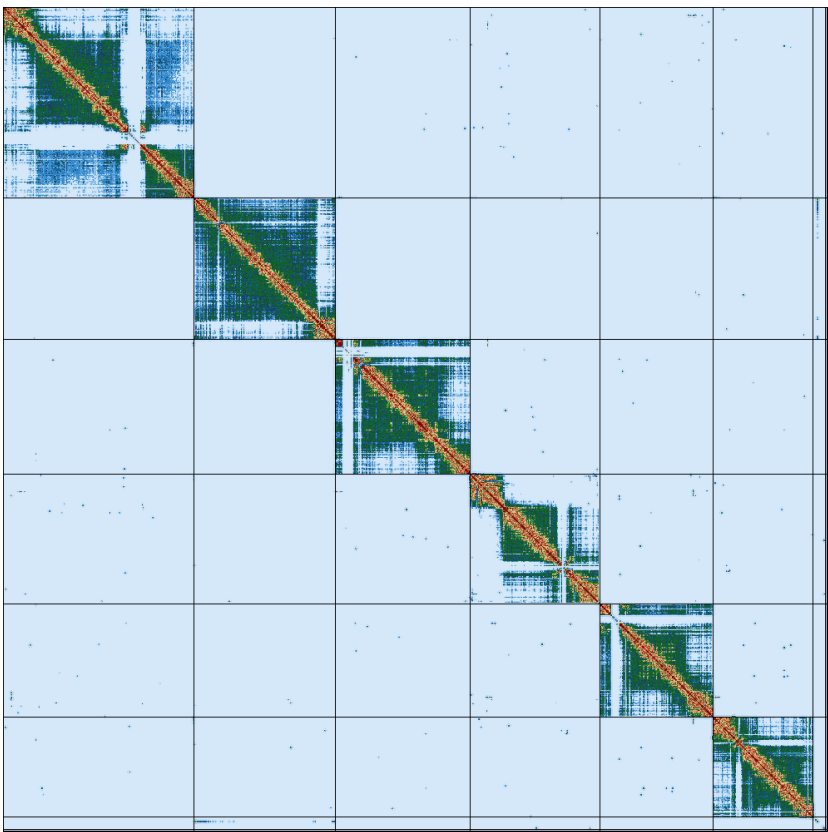
BUSCO: 5.4.4 (euk_genome_met, metaeuk) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

BUSCO: 5.8.0 (euk_genome_met, metaeuk) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

HiC contact map of curated assembly

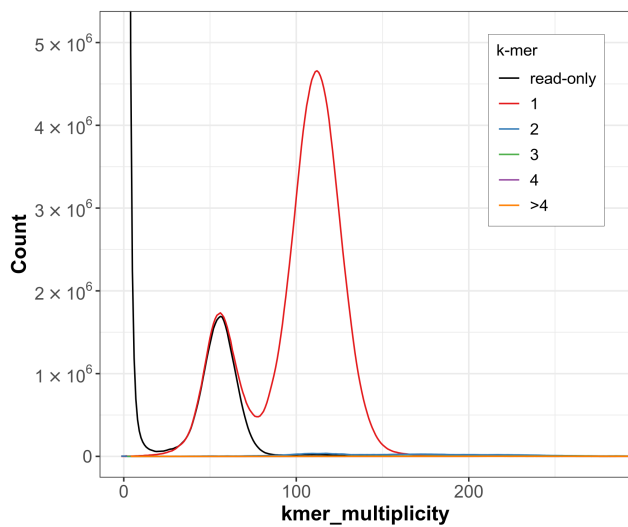


hap1 [\[LINK\]](#)

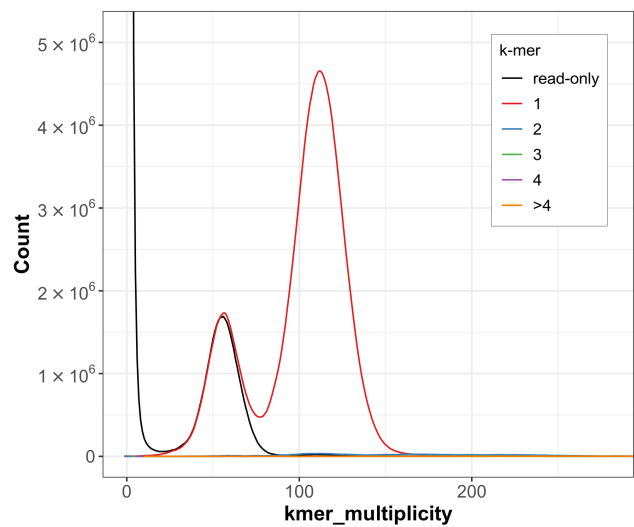


hap2 [\[LINK\]](#)

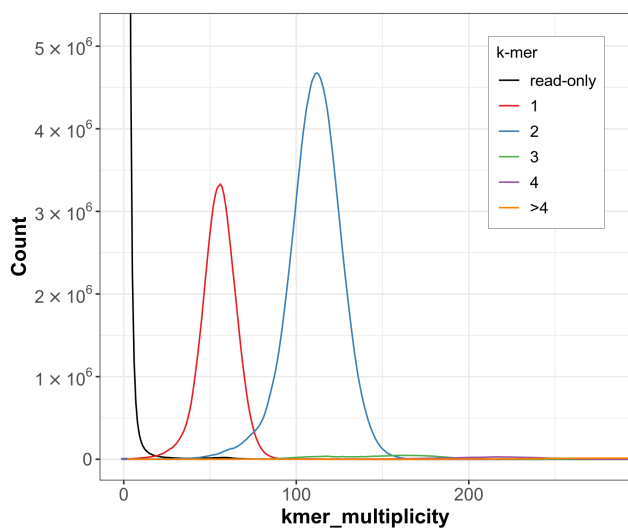
K-mer spectra of curated assembly



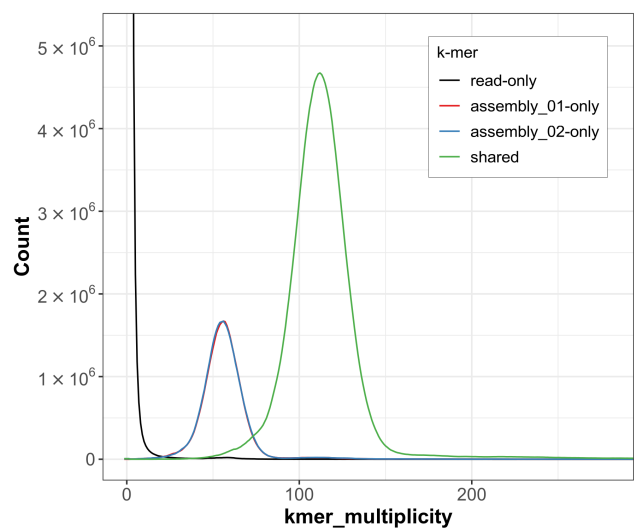
Distribution of k-mer counts per copy numbers found in **assembly_02** (hapl.)



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

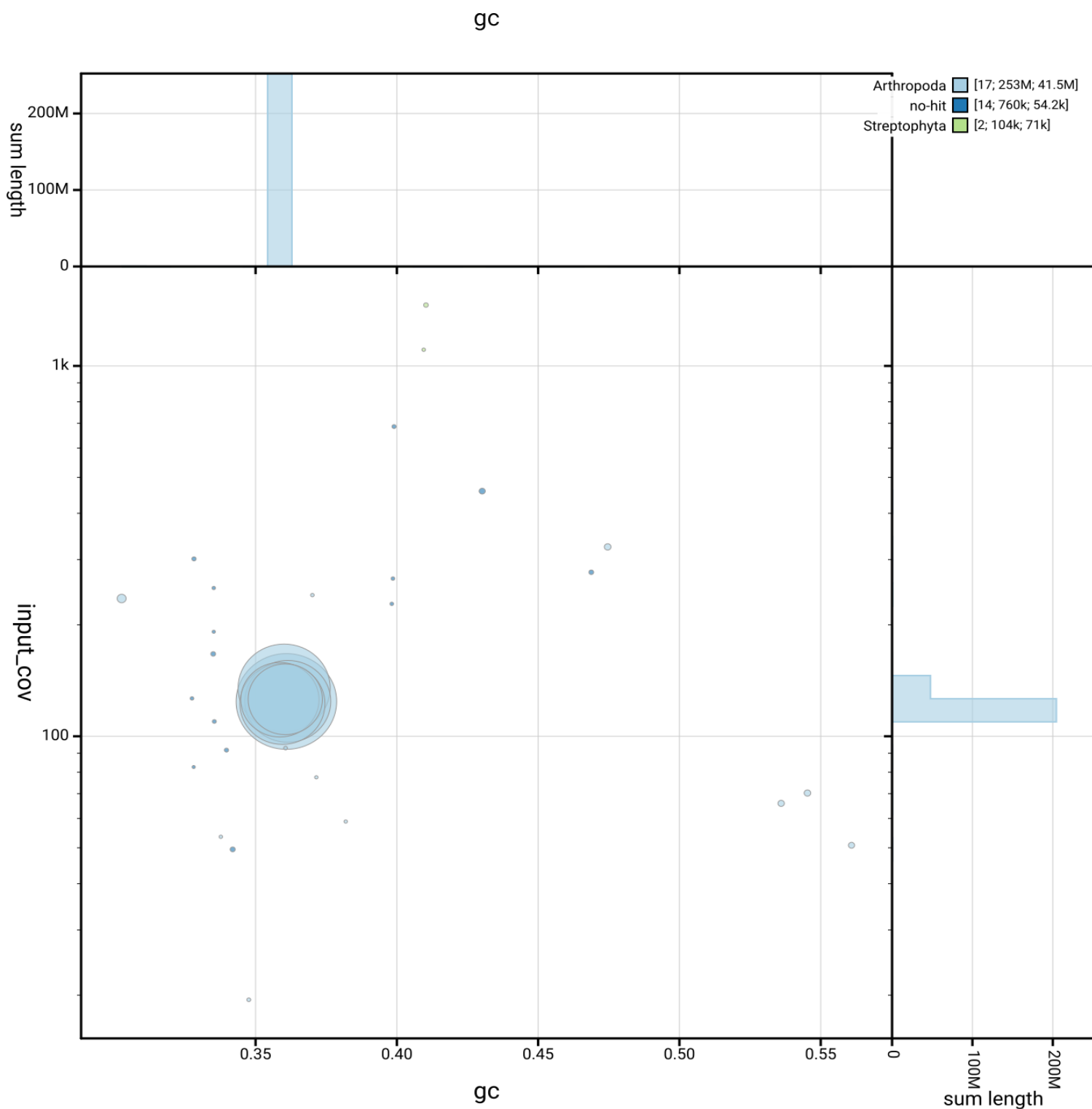


Distribution of k-mer counts per copy numbers found in **asm** (dipl.)

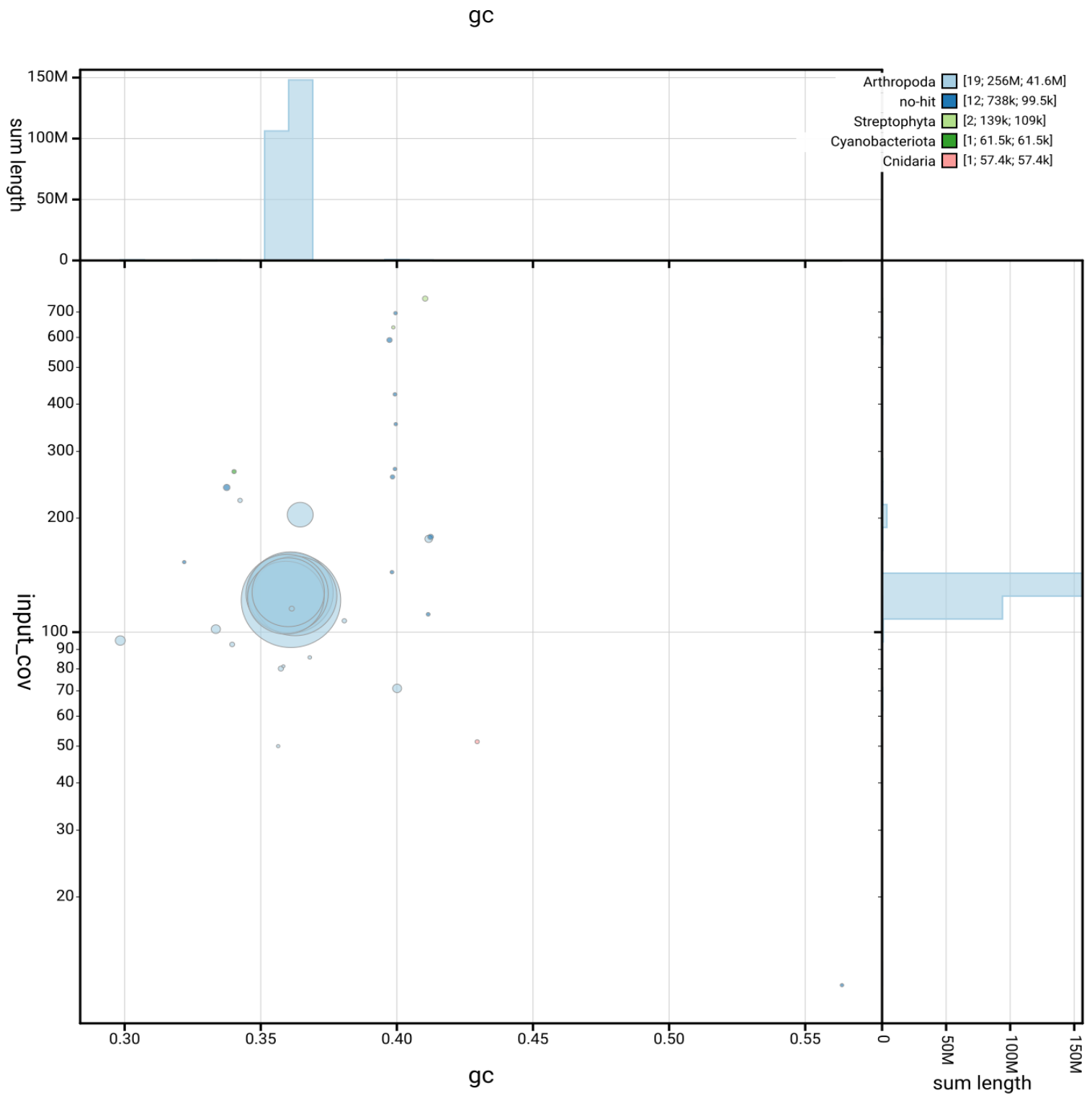


Distribution of k-mer counts coloured by their presence in reads/assemblies

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.



hap2. 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	NA	NA

Assembly pipeline

- **Hifiasm**
 - |_ *ver*: Galaxy Version 0.24.0+galaxy0
 - |_ *key param*: VGP Galaxy workflow VGP3 v 0.2.8
- **purge_dups**
 - |_ *ver*: Galaxy Version 1.2.6+galaxy0
 - |_ *key param*: VGP Galaxy workflow VGP6 v 0.5
- **YaHS**
 - |_ *ver*: Galaxy Version 1.2a.2+galaxy2
 - |_ *key param*: VGP Galaxy workflow VGP8 v 1.4

Curation pipeline

- **PretextMap**
 - |_ *ver*: Galaxy Version 0.1.9+galaxy1
 - |_ *key param*: VGP workflow VGP PretextMap generation v2
- **PretextView**
 - |_ *ver*: 1.0.0
 - |_ *key param*: NA

Submitter: Marcella Sozzoni

Affiliation: UNIFI

Date and time: 2025-06-18 11:11:28 CEST