

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	1457124
ToLID	icParSeri5
Species	Parapropus sericeus
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	475,491,312	320,578,169
Haploid Number	11 (source: ancestor)	12
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	XY

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 5.7.Q54

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
- . Kmer completeness value is less than 90 for hap1
- . Assembly length loss > 3% for hap1
- . More than 1000 gaps/Gbp for hap1
- . Not 90% of assembly in chromosomes for hap1

Curator notes

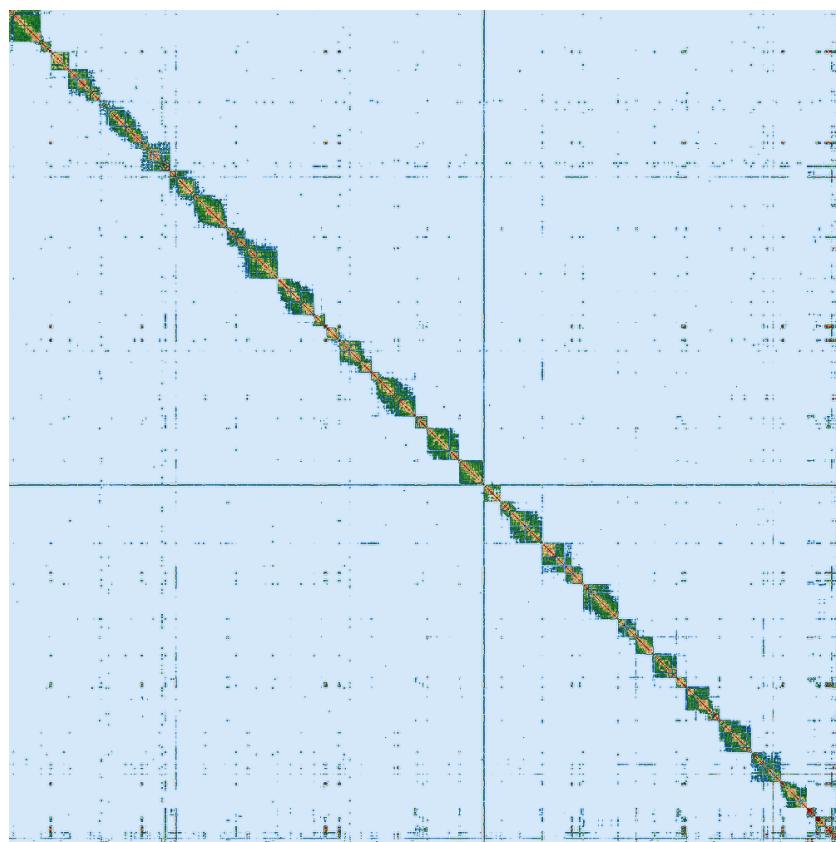
- . Interventions/Gb: 66
- . Contamination notes: "Contaminants were found and removed with the pipeline NCBI FCS GX v Galaxy Version 0.5.5+galaxy1. Scaffolds that are flagged as contaminants by BlobToolKit were confirmed with FCS output table and with contact in the map as non-contaminants."
- . Other observations: "Only one X chromosomes was found with alignment with close relatives Leptodirus hochenwartii. Y chromosome could not be identified"

Quality metrics table

Metrics	Pre-curation hap1	Curated hap1
Total bp	357,343,292	320,578,169
GC %	31.22	31.47
Gaps/Gbp	1,603.5	4,189.31
Total gap bp	114,600	268,600
Scaffolds	182	147
Scaffold N50	16,428,764	16,828,879
Scaffold L50	9	8
Scaffold L90	21	17
Contigs	755	1,490
Contig N50	887,040	403,605
Contig L50	126	224
Contig L90	398	837
QV	54.7745	54.7745
Kmer compl.	67.0712	67.0712
BUSCO sing.	94.2%	94.2%
BUSCO dupl.	1.1%	1.1%
BUSCO frag.	0.9%	0.9%
BUSCO miss.	3.8%	3.8%

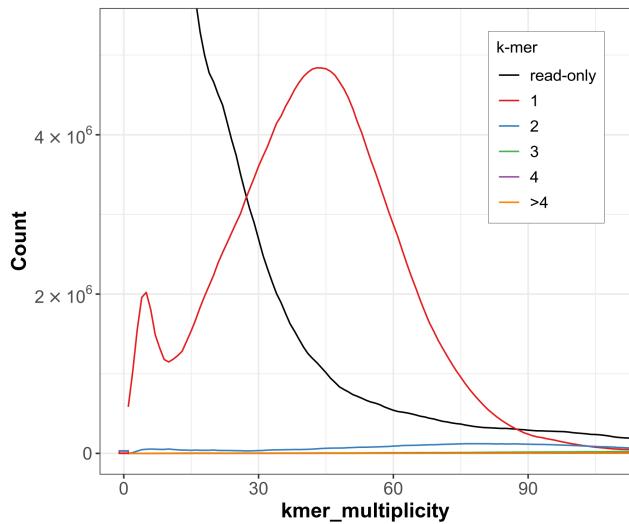
BUSCO: 5.8.0 (euk_genome_min, miniprot) / Lineage: insecta_odb10 (genomes:75, BUSCOs:1367)

HiC contact map of curated assembly

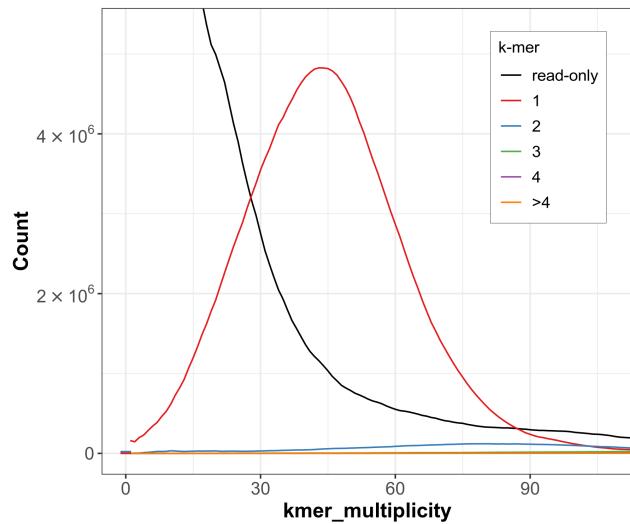


hap1 [\[LINK\]](#)

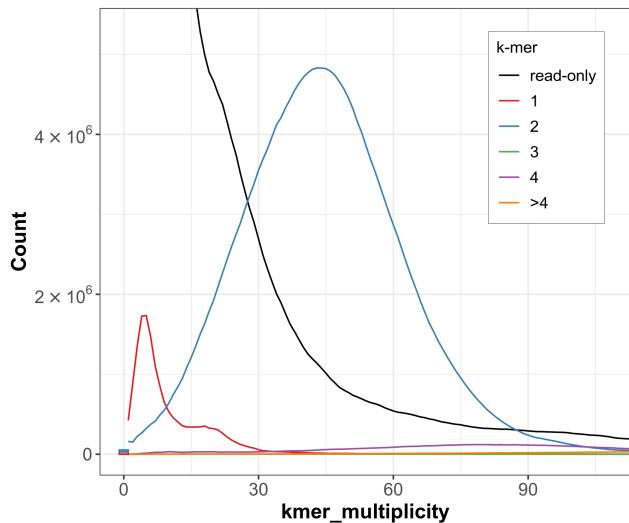
K-mer spectra of curated assembly



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.)

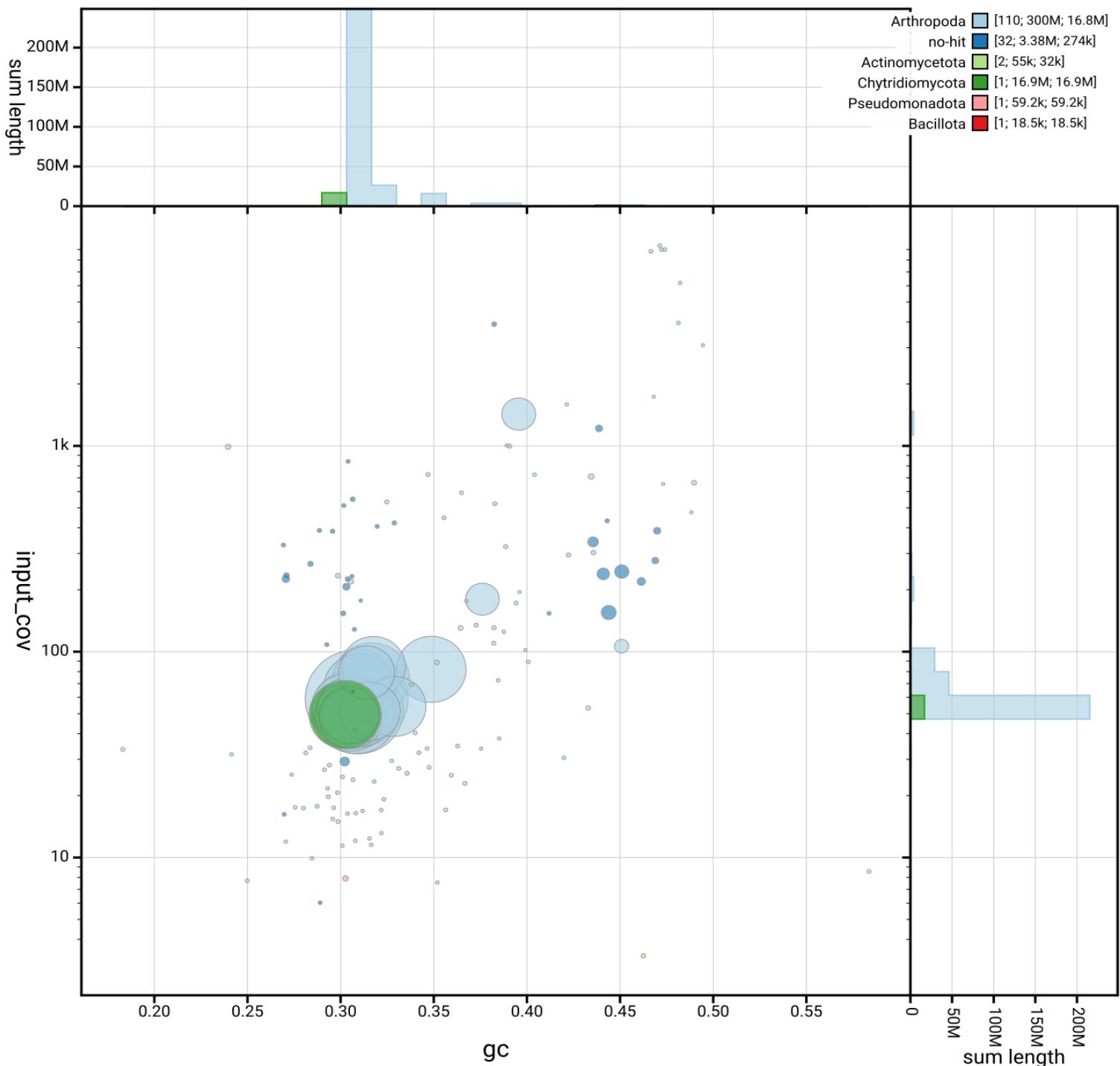


Distribution of k-mer counts per copy numbers found in asm (dipl.)

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

Post-curation contamination screening

gc



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	32.93x	57.75x	

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
- **HiGlass**
 - |_ ver: 1.0
 - |_ key param: NA
- **PretextView**
 - |_ ver: 1.0.0
 - |_ key param: NA

Submitter: Marcella Sozzoni

Affiliation: UNIFI

Date and time: 2026-02-23 09:25:21 CET