

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

| | |
|---------|-----------------------------|
| TxID | 3229210 |
| ToLID | ilPseRect3 |
| Species | Pseudoterpna rectistrigaria |
| Class | Insecta |
| Order | Lepidoptera |

| Genome Traits | Expected | Observed |
|-------------------|-----------------------|-------------|
| Haploid size (bp) | 328,767,194 | 346,830,827 |
| Haploid Number | 31 (source: ancestor) | 31 |
| Ploidy | 2 (source: ancestor) | 2 |
| Sample Sex | ZW | ZW |

EBP metrics summary and curation notes

Obtained EBP quality metric for : 7.7.Q71

Obtained EBP quality metric for : 7.7.Q70

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

- . Kmer completeness value is less than 90 for
- . Kmer completeness value is less than 90 for
- . Assembly length loss > 3% for

Curator notes

- . Interventions/Gb: 1
- . Contamination notes: "No contamination was detected by Blobtoolkit, the no-hit sequences correspond to the W chromosome, due to its high repetitiveness."
- . Other observations: "This assembly was obtained with Hifiasm from ONT data. Given the high quality of both haplotypes, curation was done in diploid mode. Chromosome W was broken between the 2 assemblies and we basically joined the two contigs to hap1, obtaining a complete chromosome. Mitogenome was assembled with FOAM. "

Quality metrics table

| Metrics | Pre-curation <hap1> | Pre-curation <hap2> | Curated <hap1> | Curated <hap2> |
|--------------|------------------------|------------------------|-------------------|-------------------|
| Total bp | 338,235,196 | 338,235,196 | 346,830,827 | 316,793,520 |
| GC % | 36.24 | 36.24 | 36.32 | 36.21 |
| Gaps/Gbp | 0 | 0 | 2.88 | 3.16 |
| Total gap bp | 0 | 0 | 200 | 200 |
| Scaffolds | 37 | 37 | 32 | 35 |
| Scaffold N50 | 11,937,579 | 11,937,579 | 11,937,579 | 11,740,834 |
| Scaffold L50 | 13 | 13 | 14 | 13 |
| Scaffold L90 | 27 | 27 | 27 | 25 |
| Contigs | 37 | 37 | 33 | 36 |
| Contig N50 | 11,937,579 | 11,937,579 | 11,662,666 | 11,740,834 |
| Contig L50 | 13 | 13 | 14 | 13 |
| Contig L90 | 27 | 27 | 28 | 25 |
| QV | 71.7319 | 71.0842 | 71.9967 | 70.833 |
| Kmer compl. | 86.3709 | 82.3098 | 86.4981 | 82.0504 |
| BUSCO sing. | 98.6% | 96.1% | 98.2% | 94.2% |
| BUSCO dupl. | 0.5% | 0.1% | 0.3% | 0.2% |
| BUSCO frag. | 0.4% | 0.4% | 0.3% | 0.2% |
| BUSCO miss. | 0.5% | 3.4% | 1.2% | 5.4% |

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

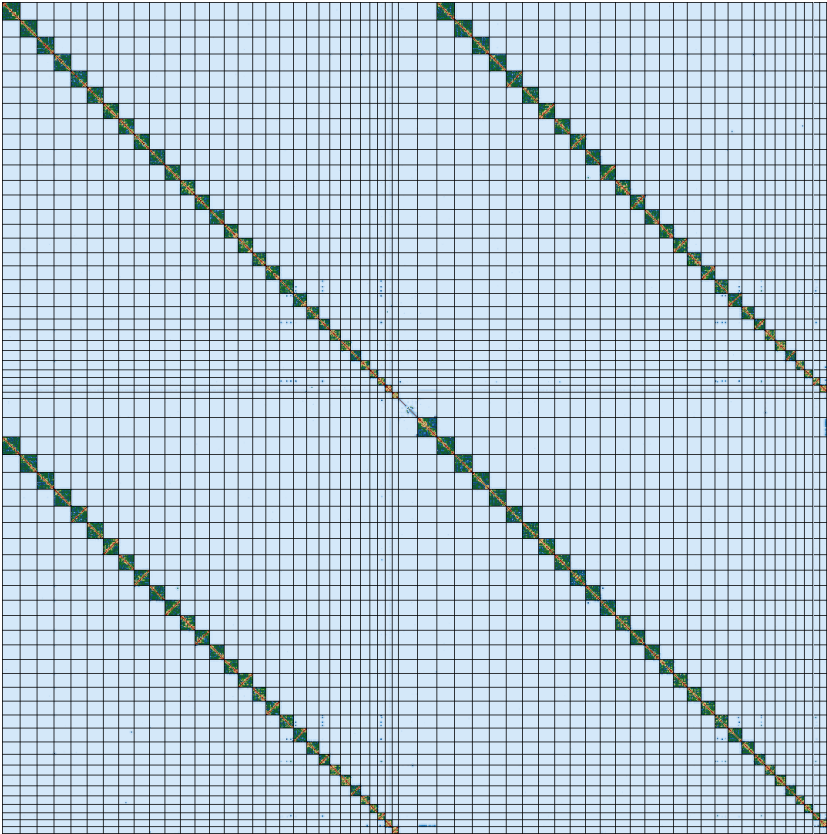
BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: insecta_odb10 (genomes:75, BUSCOs:1367)

BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: insecta_odb10 (genomes:75, BUSCOs:1367)

BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: lepidoptera_odb10 (genomes:16, BUSCOs:5286)

BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: lepidoptera_odb10 (genomes:16, BUSCOs:5286)

HiC contact map of curated assembly

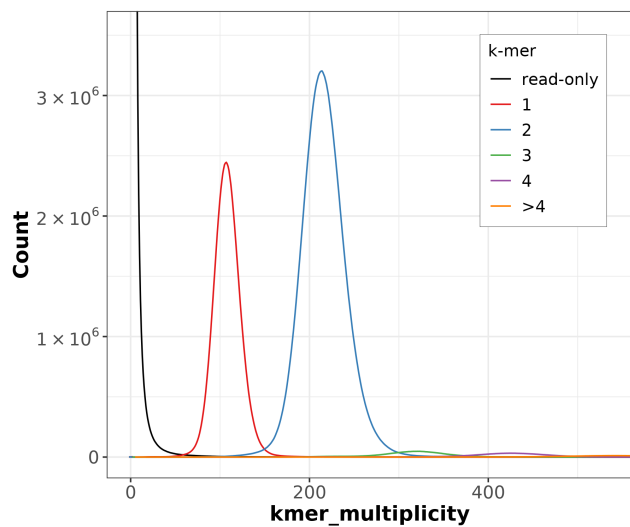


[\[LINK\]](#)

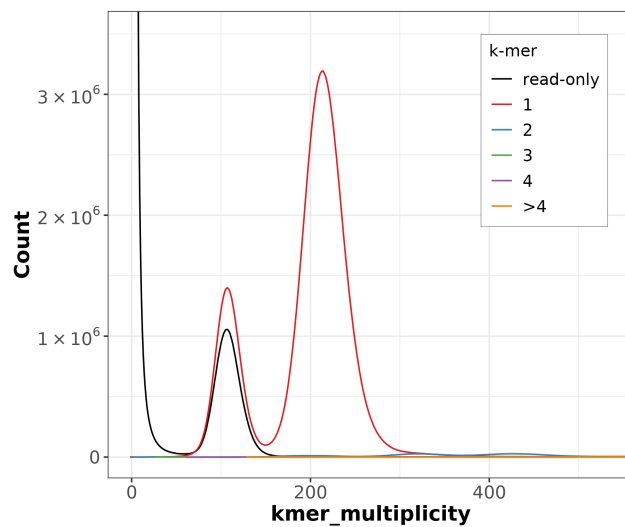
HiC PNG is missing!

File link is missing!

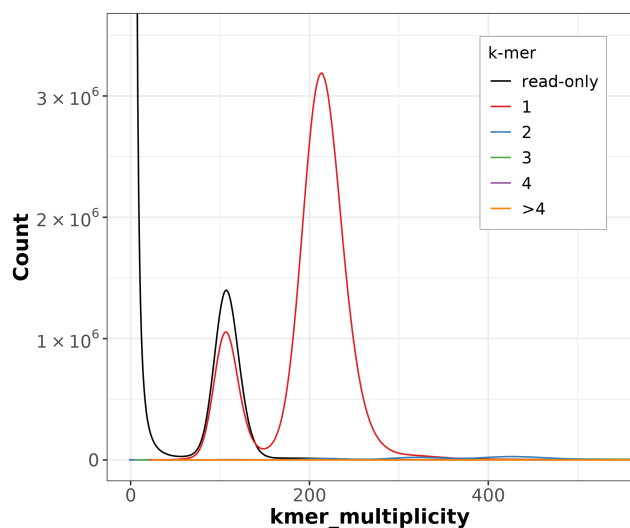
K-mer spectra of curated assembly



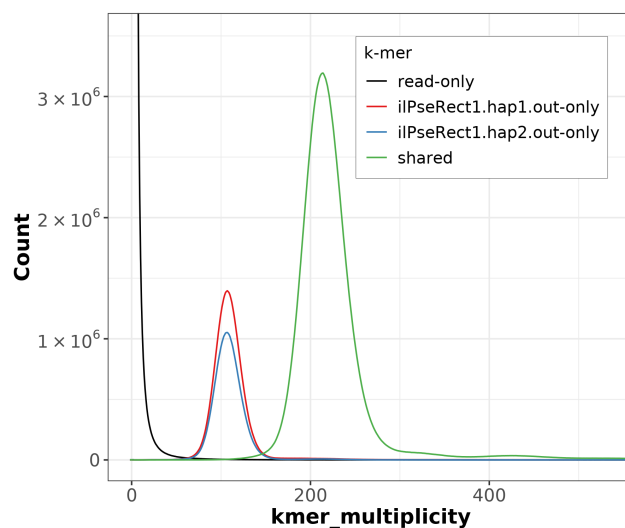
Distribution of k-mer counts per copy numbers found in asm (dipl.)



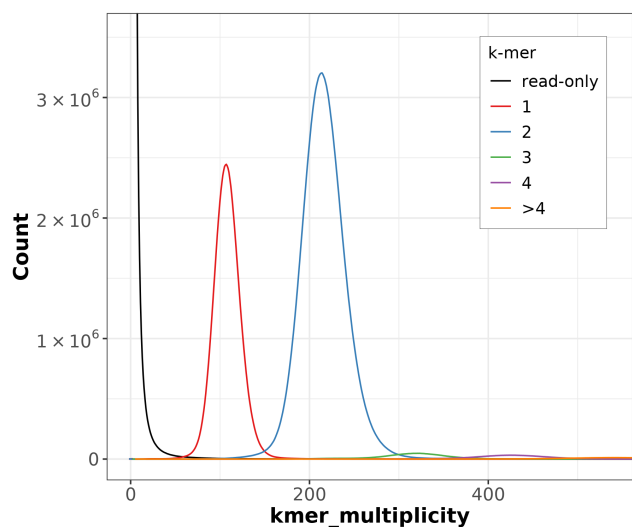
Distribution of k-mer counts per copy numbers found in **hap1** (hap1.)



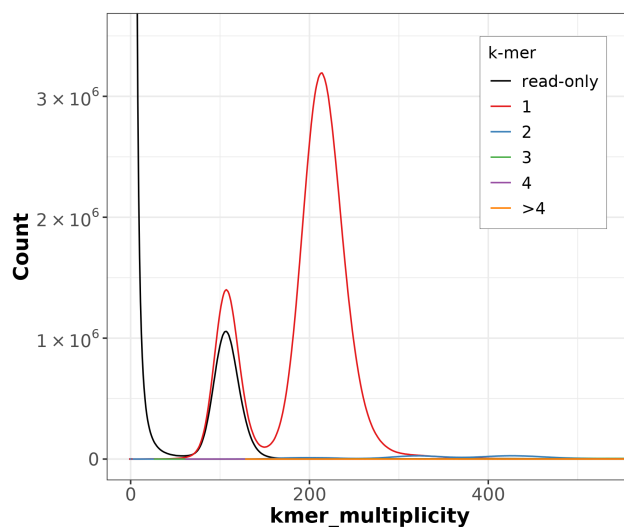
Distribution of k-mer counts per copy numbers found in **hap2** (hap1.)



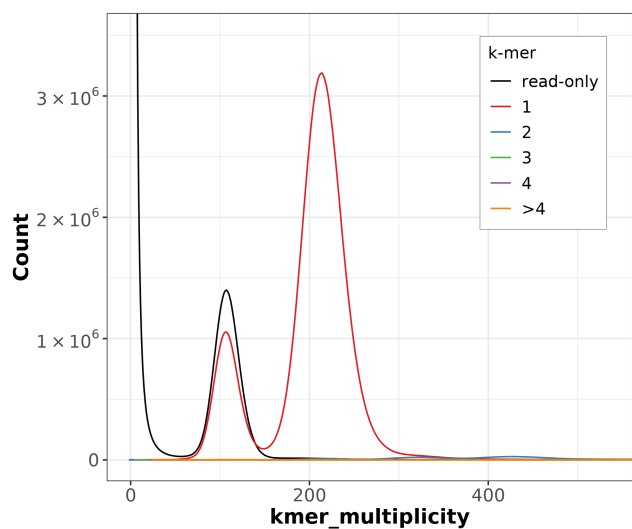
Distribution of k-mer counts coloured by their presence in reads/assemblies



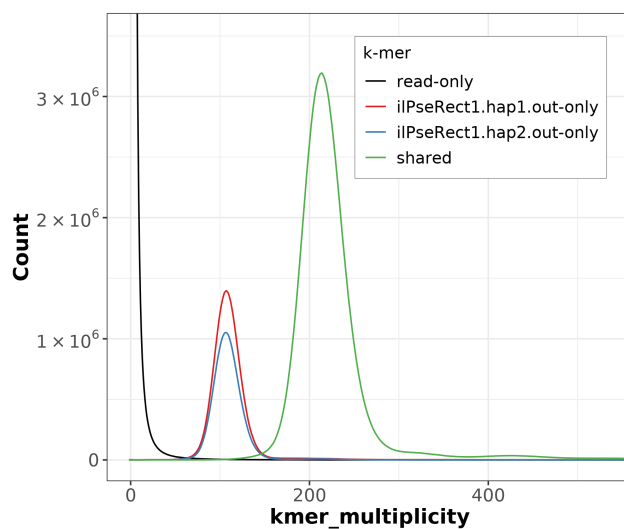
Distribution of k-mer counts per copy numbers found in `asm` (dipl.)



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

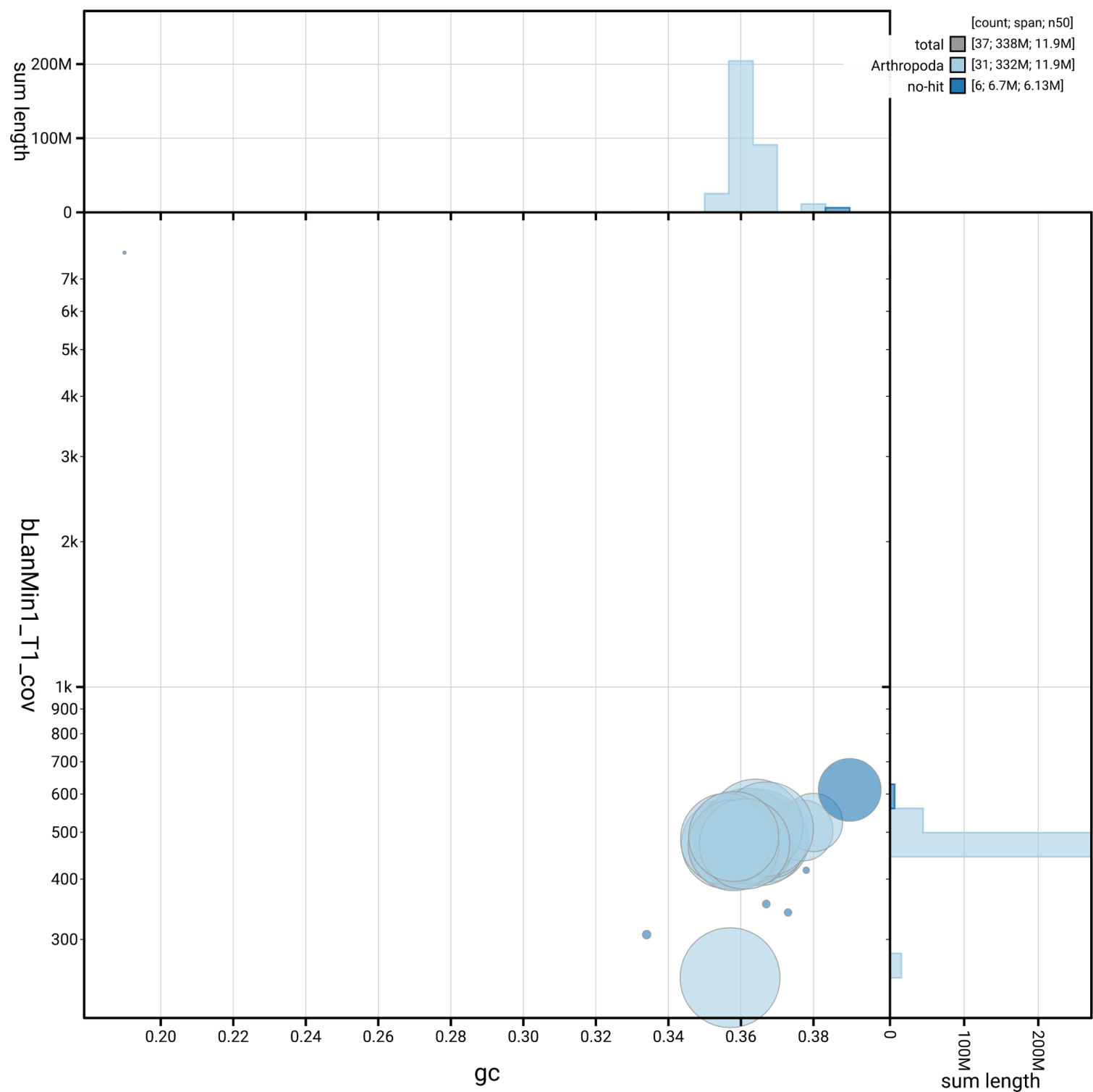


Distribution of k-mer counts per copy numbers found in `hap2` (hapl.)



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

Post-curation contamination screening



. 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 | ONT | Illumina | Omni-C |
|----------|-----|----------|--------|
| Coverage | 100 | 290 | 270 |

Assembly pipeline

- **CLAWS**
 - |_ *ver*: v2.3.0
 - |_ *key param*: NA
- **Filtlong**
 - |_ *ver*: v0.2.1
 - |_ *key param*: -t target_bases 36000000000
- **Hifiasm**
 - |_ *ver*: 0.24.0
 - |_ *key param*: --ont
- **Yahs**
 - |_ *ver*: 1.2a
 - |_ *key param*: -mq 0
 - |_ *key param*: -no-contig-ec
- **FOAM**
 - |_ *ver*: 0.5
 - |_ *key param*: NA

Curation pipeline

- **Blobtoolkit Nextflow pipeline**
 - |_ *ver*: 0.6
 - |_ *key param*: NA
- **PretextViewAI**
 - |_ *ver*: 1.0.3
 - |_ *key param*: NA

Submitter: Jessica Gomez-Garrido

Affiliation: CNAG

Date and time: 2025-09-26 12:16:17 CEST