

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	2560939
ToLID	<b>ilAmaRaga7</b>
Species	<i>Amata ragazzii</i>
Class	Insecta
Order	Lepidoptera

Genome Traits	Expected	Observed
Haploid size (bp)	355,561,558	2,040,249,389
Haploid Number	31 (source: ancestor)	0
Ploidy	2 (source: ancestor)	2
Sample Sex	ZW	NA

## EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 5.5.Q57

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: "NA"
- . Other observations: "This is a contig level assembly. The production of the HiC reads failed so there is no scaffolding for this species."

## Quality metrics table

Metrics	Pre-curation hap1	Curated hap1
Total bp	2,040,249,389	2,040,249,389
GC %	36.77	36.77
Gaps/Gbp	0	0
Total gap bp	0	0
Scaffolds	12,156	12,156
Scaffold N50	249,096	249,096
Scaffold L50	2,552	2,552
Scaffold L90	8,042	8,042
Contigs	12,156	12,156
Contig N50	249,096	249,096
Contig L50	2,552	2,552
Contig L90	8,042	8,042
QV	57.4524	57.4524
Kmer compl.	90.8128	90.8128
BUSCO sing.	89.3%	89.3%
BUSCO dupl.	2.4%	2.4%
BUSCO frag.	3.4%	3.4%
BUSCO miss.	4.9%	4.9%

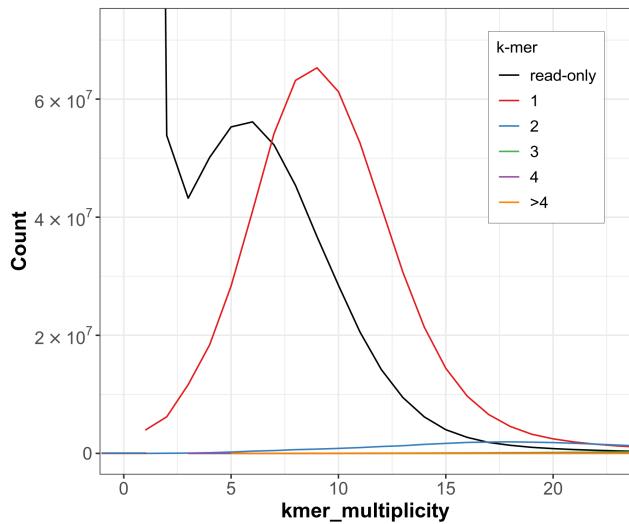
BUSCO: 5.8.0 (euk\_genome\_min, miniprot) / Lineage: lepidoptera\_odb10 (genomes:16, BUSCOs:5286)

# HiC contact map of curated assembly

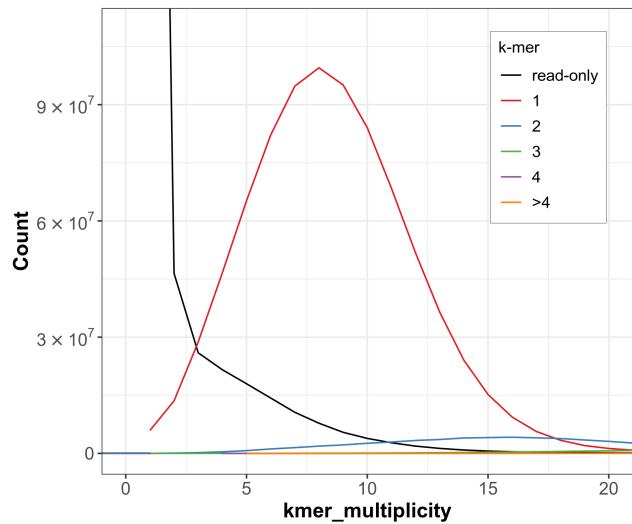
**hap1** HiC PNG is missing!

**hap1** File link is missing!

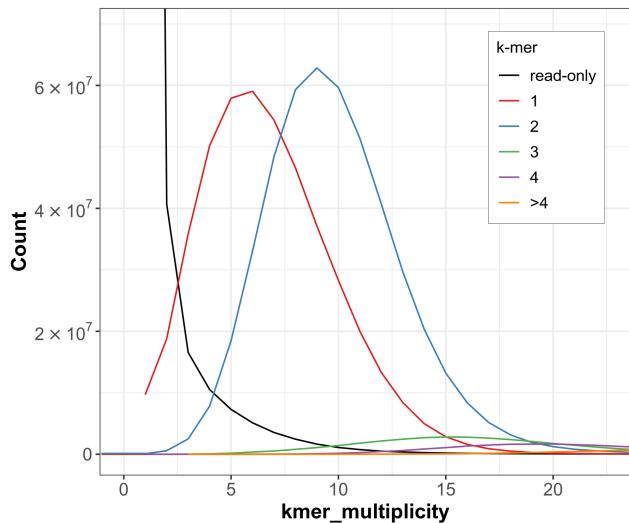
# K-mer spectra of curated assembly



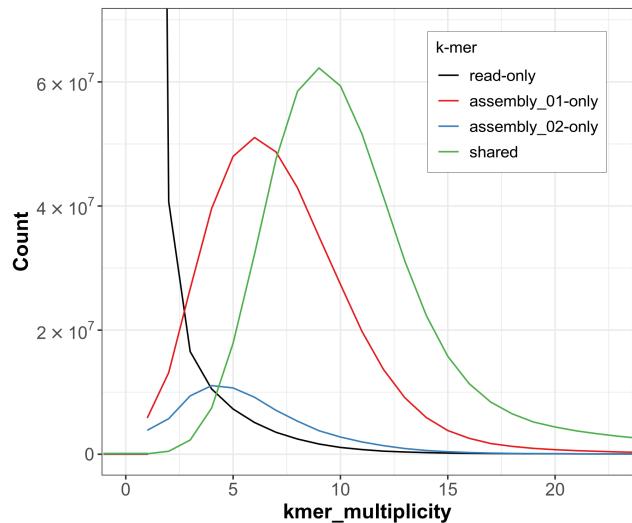
Distribution of k-mer counts per copy numbers found in **assembly\_02** (hapl.)



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

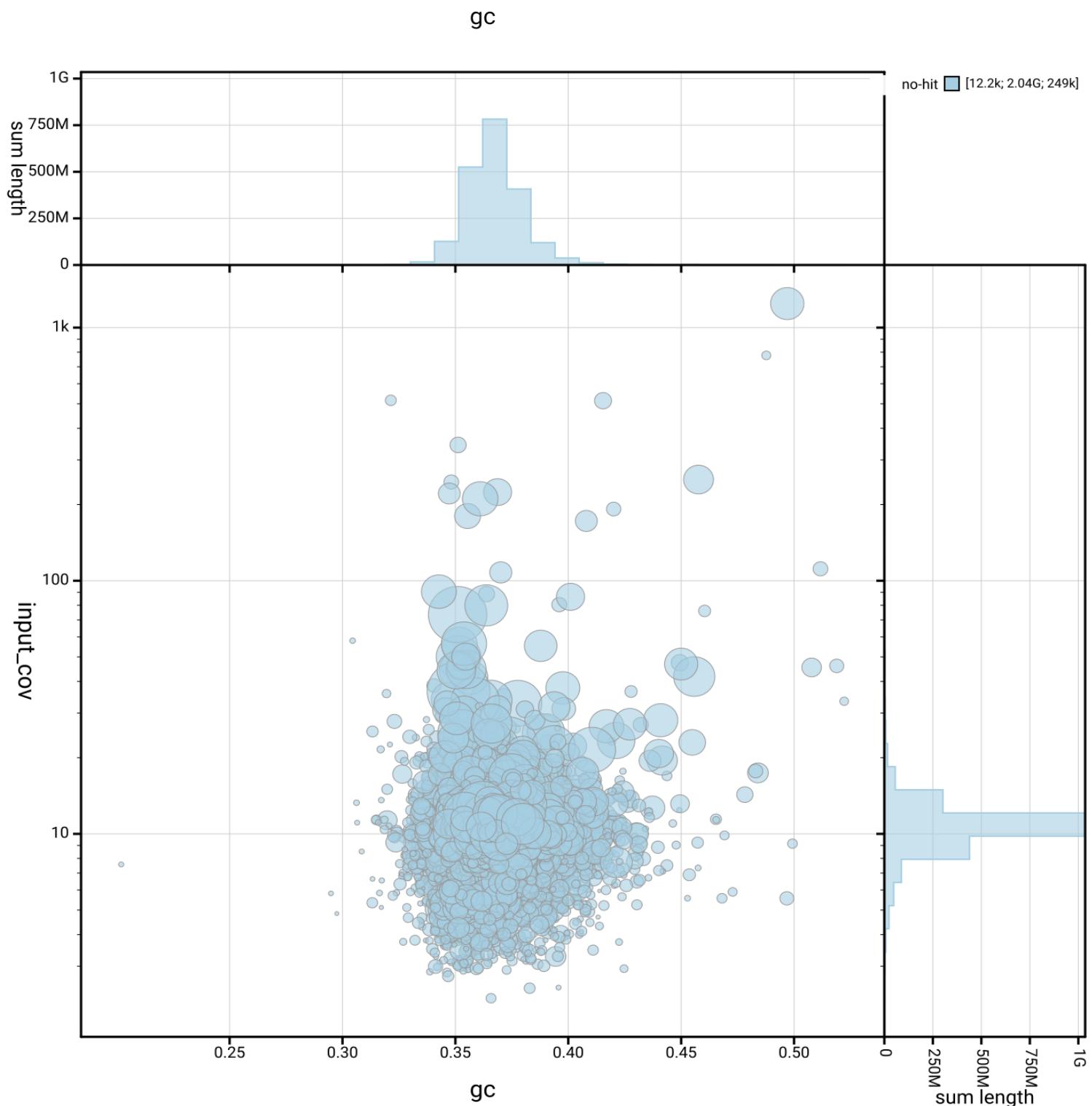


Distribution of k-mer counts per copy numbers found in **asm** (dipl.)



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	28.25x	61.67x	

## 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

## Curation pipeline

Invalid pipeline data format

Submitter: Marcella Sozzoni

Affiliation: UNIFI

Date and time: 2026-02-16 10:21:50 CET