

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	1869821
ToLID	ilEarVern1
Species	Earias vernana
Class	Insecta
Order	Lepidoptera

Genome Traits	Expected	Observed
Haploid size (bp)	327,502,376	366,537,864
Haploid Number	31 (source: ancestor)	32
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	ZZ

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.7.Q73

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

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

Curator notes

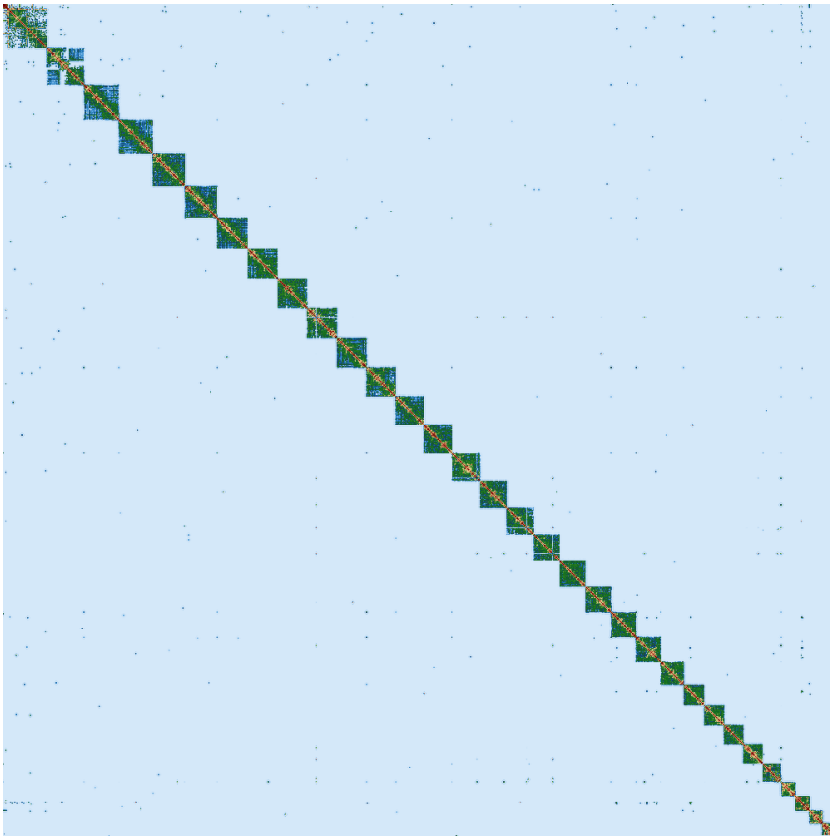
- . Interventions/Gb: 7
- . Contamination notes: "Curation made 1 cut in a contig, 1 break at a gap and 5 joins"
- . Other observations: "None"

Quality metrics table

Metrics	Pre-curation hap1	Curated hap1
Total bp	366,537,064	366,537,864
GC %	35.89	35.89
Gaps/Gbp	68.21	79.12
Total gap bp	5,000	5,800
Scaffolds	44	41
Scaffold N50	12,485,495	12,485,495
Scaffold L50	13	13
Scaffold L90	27	27
Contigs	69	70
Contig N50	11,881,358	11,881,358
Contig L50	14	14
Contig L90	28	28
QV	73.9497	73.9497
Kmer compl.	84.1542	84.1542
BUSCO sing.	98.3%	98.3%
BUSCO dupl.	0.5%	0.5%
BUSCO frag.	0.7%	0.7%
BUSCO miss.	0.5%	0.5%

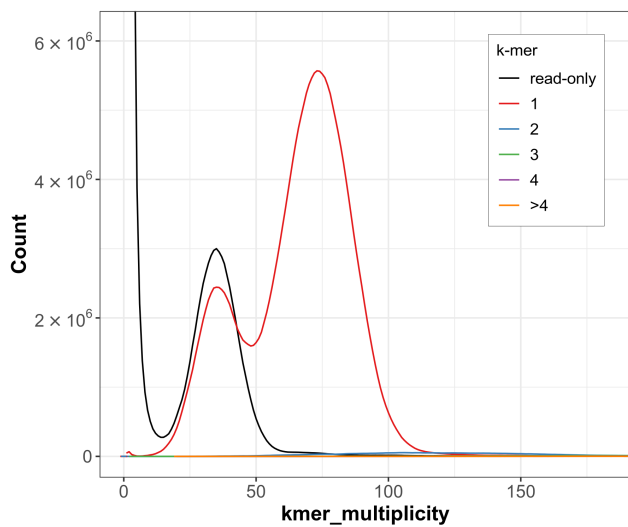
BUSCO: 5.8.0 (euk_genome_min, miniprot) / Lineage: lepidoptera_odb10 (genomes:16, BUSCOs:5286)

HiC contact map of curated assembly

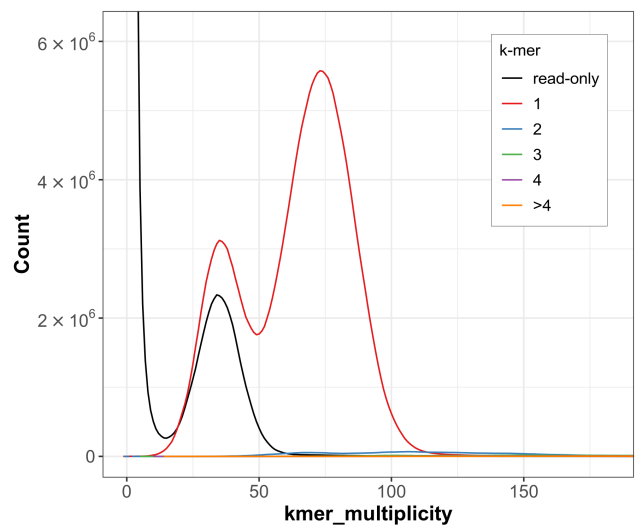


hap1 [\[LINK\]](#)

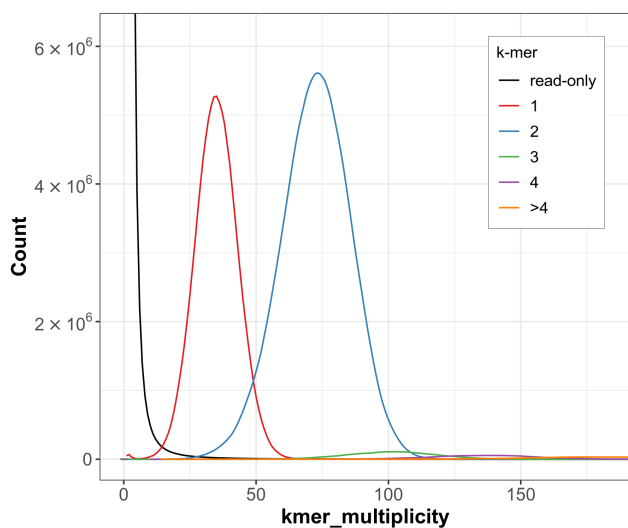
K-mer spectra of curated assembly



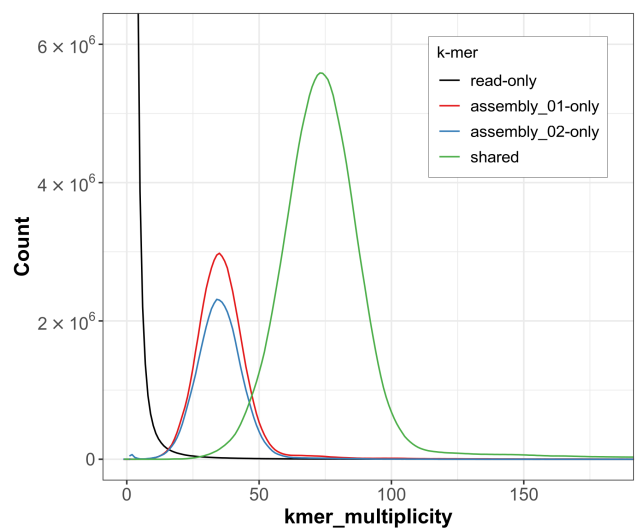
Distribution of k-mer counts per copy numbers found in **assembly_02** (hapl.)



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

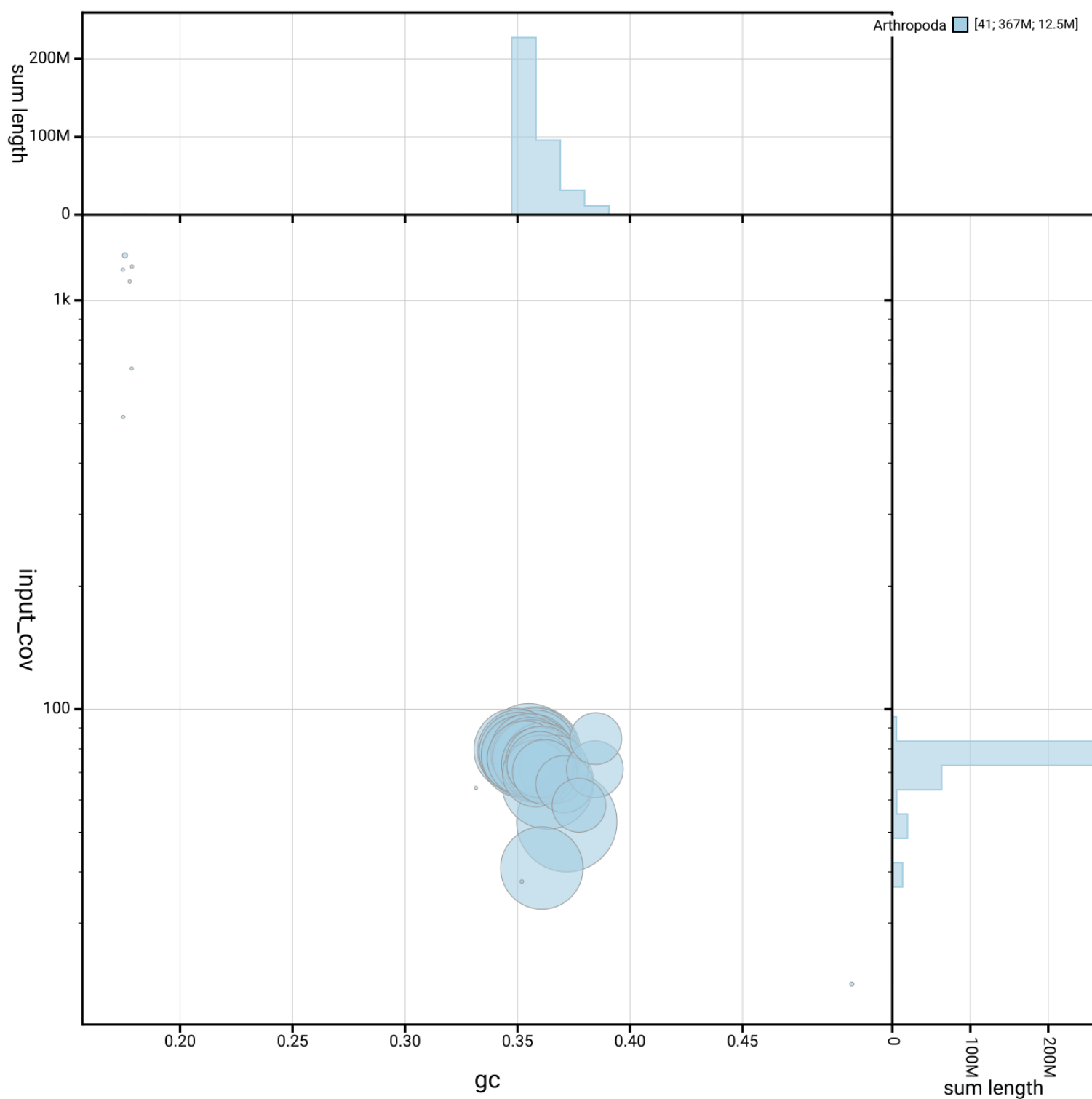


Distribution of k-mer counts per copy numbers found in **asm** (dipl.)



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

Post-curation contamination screening



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	42.44X	54.85X

Assembly pipeline

- **Hifiasm**
 - |_ *ver*: Galaxy Version 0.24.0+galaxy0
 - |_ *key param*: VGP Galaxy workflow VGP3 v 3.12
- **YaHS**
 - |_ *ver*: Galaxy Version 1.2a.2+galaxy2
 - |_ *key param*: VGP Galaxy workflow VGP8 v 3.1

Curation pipeline

- **PretextMap**
 - |_ *ver*: Galaxy Version 0.1.9+galaxy1
 - |_ *key param*: VGP workflow VGP PretextMap generation v2
- **PretextViewAI**
 - |_ *ver*: 1.0.3
 - |_ *key param*: NA

Submitter: Marcella Sozzoni

Affiliation: UNIFI

Date and time: 2025-12-16 14:13:49 CET