

ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

TxID	3163510
ToLID	idEumHispl
Species	Eumerus hispidus
Class	Insecta
Order	Diptera

Genome Traits	Expected	Observed
Haploid size (bp)	533,676,914	672,541,521
Haploid Number	6 (source: ancestor)	7
Ploidy	3 (source: ancestor)	2
Sample Sex	M	M

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 7.7.Q65

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 Ploidy is different from Expected

Curator notes

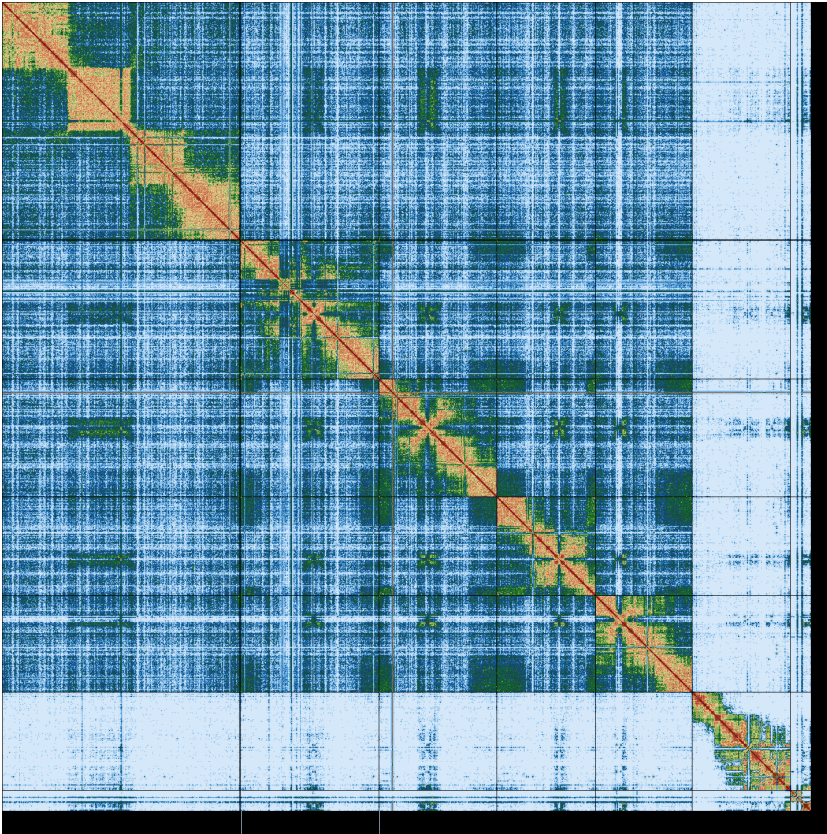
- . Interventions/Gb: 59
- . Contamination notes: "Contamination report for assembly labelled hap1; Total length of scaffolds removed: 1,008,721 (0.1 %);Scaffolds removed: 23 (6.9 %);Largest scaffold removed: (201,185);FCS-GX contaminant species (number of scaffolds; total length of scaffolds):Manihot esculenta, plants (5; 92,971);Hevea brasiliensis, plants (3; 237,937);Asaia bogorensis, a-proteobacteria (3; 135,041);Asaia krungthepensis, a-proteobacteria (2; 119,779);Kozakia baliensis, a-proteobacteria (2; 86,105);Drosophila suzukii associated hytrosavirus 1, viruses (1; 120,106);Asaia lannensis, a-proteobacteria (1; 68,758);Neoasaia chiangmaiensis, a-proteobacteria (1; 36,138);Swaminathania salitolerans, a-proteobacteria (1; 33,799);Asaia astilbis, a-proteobacteria (1; 25,523); Mitochondrion (3; 52,564)"
- . Other observations: "Assembly was Hi-C phased; The exact order and orientation of the contigs on chromosome X (75.4 - 79.2 Mbp) are unknown. The X and Y chromosome was identified by read coverage."

Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	672,552,723	672,541,521
GC %	38.2	38.18
Gaps/Gbp	71.37	160.58
Total gap bp	4,800	17,500
Scaffolds	335	266
Scaffold N50	79,633,640	95,008,139
Scaffold L50	3	3
Scaffold L90	10	6
Contigs	383	374
Contig N50	12,617,015	11,498,168
Contig L50	15	17
Contig L90	48	52
QV	66.3	65.8
Kmer compl.	99.43	99.40
BUSCO sing.	98.8%	98.8%
BUSCO dupl.	0.6%	0.5%
BUSCO frag.	0.3%	0.3%
BUSCO miss.	0.4%	0.4%

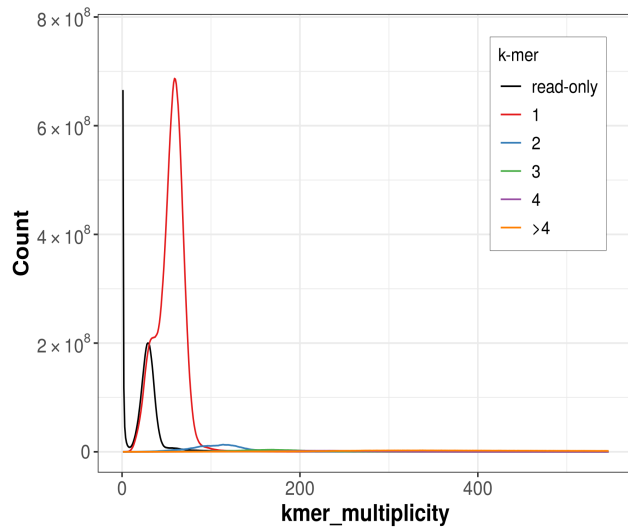
BUSCO 6.0.0 Lineage: insecta_odb10 (genomes:75, BUSCOs:1367)

HiC contact map of curated assembly

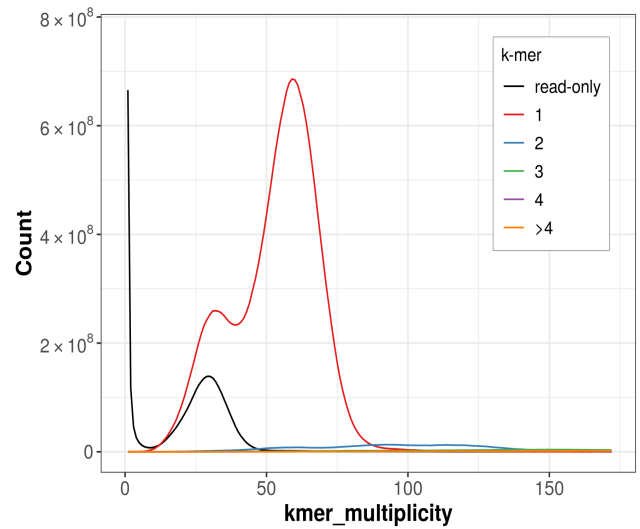


pri [\[LINK\]](#)

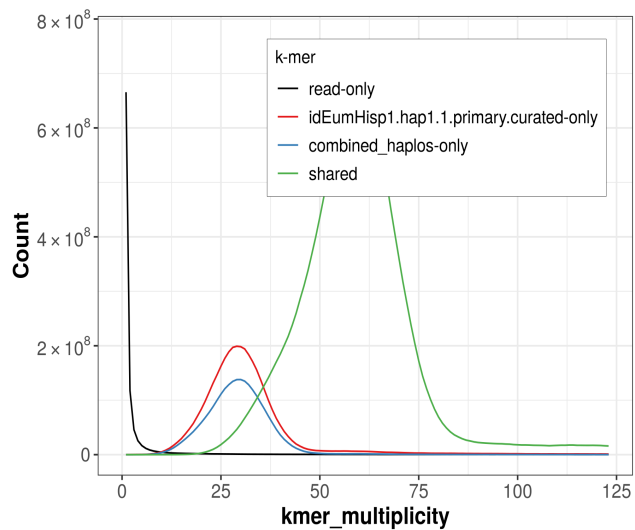
K-mer spectra of curated assembly



Distribution of k-mer counts per copy numbers found in asm

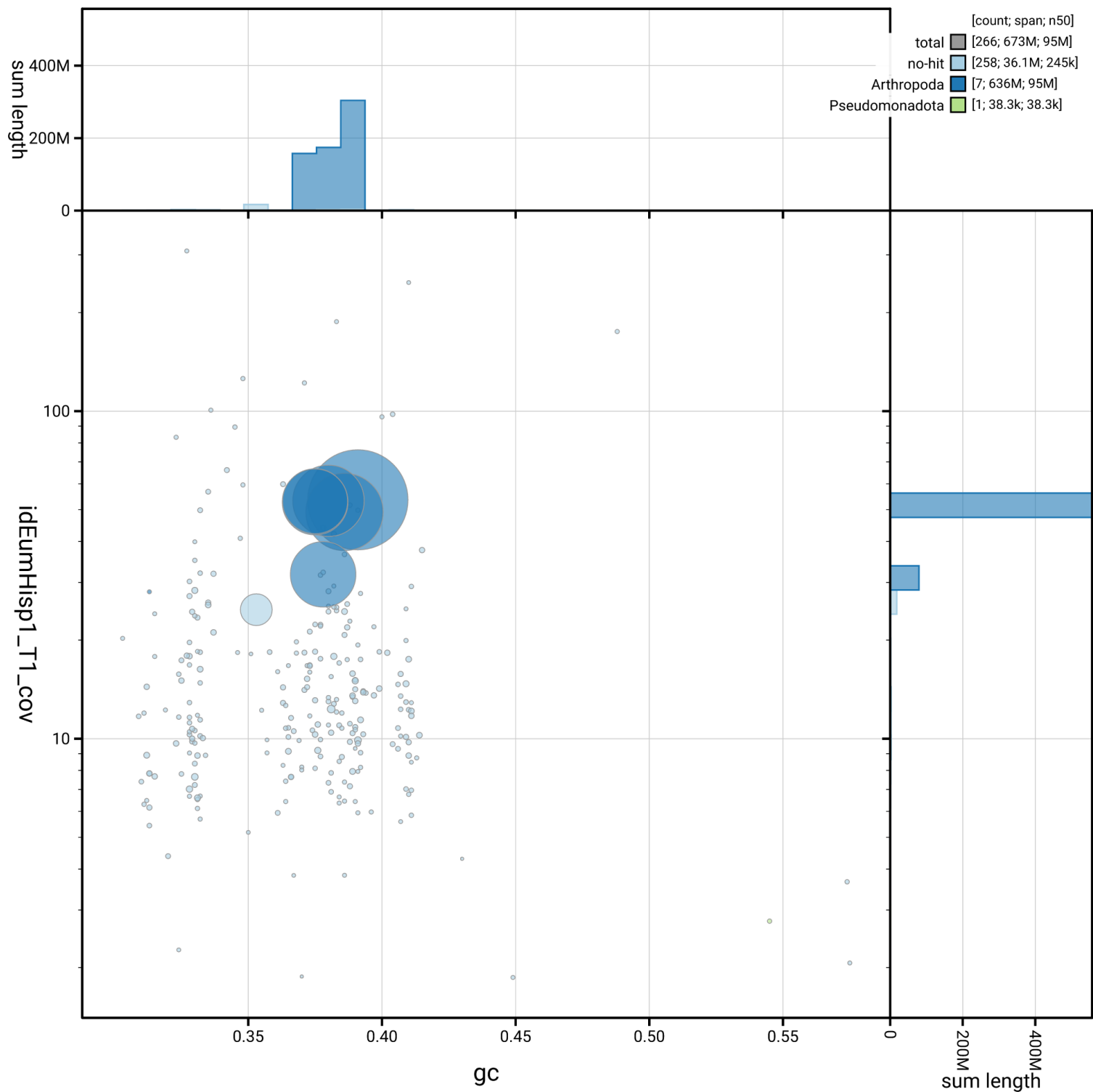


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



pri. 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	PacBio HiFi	Arima v2
Coverage	58x	436x

Assembly pipeline

- **hifiasm-hic**
 - |_ *ver*: 0.19.8-r603
 - |_ *key param*: --h1/--h2
- **yahs**
 - |_ *ver*: 1.2.2
 - |_ *key param*: NA

Curation pipeline

- **hifiasm-hic**
 - |_ *ver*: 0.19.8-r603
 - |_ *key param*: --h1/--h2
- **yahs**
 - |_ *ver*: 1.2.2
 - |_ *key param*: NA
- **TreeVal**
 - |_ *ver*: 1.2.1
 - |_ *key param*: NA

Submitter: Daniel Zilov

Affiliation: WSI

Date and time: 2025-09-16 14:37:44 CEST