

# ERGA Assembly Report

v24.04.03\_beta

Tags: ERGA-BGE

TxID	3163543
ToLID	<b>ihCypFauv1</b>
Species	Cyphopterus fauveli
Class	Insecta
Order	Hemiptera

Genome Traits	Expected	Observed
Haploid size (bp)	1,148,501,624	1,202,388,514
Haploid Number	13 (source: ancestor)	14
Ploidy	2 (source: ancestor)	2
Sample Sex	M	M

## EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 7.7.Q59

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

- . Observed Haploid Number is different from Expected

### Curator notes

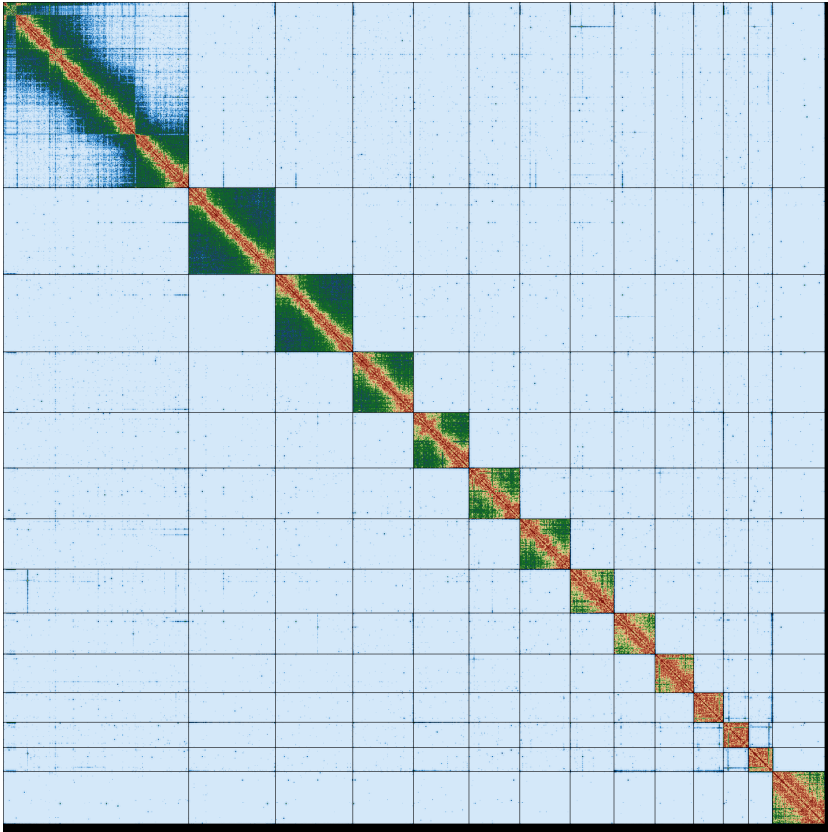
- . Interventions/Gb: 70
- . Contamination notes: "Contamination report for assembly labelled hap1; Total length of scaffolds removed: 1,253,070 (0.1 %); Scaffolds removed: 6 (2.4 %);Largest scaffold removed: (992,013);FCS-GX contaminant species (number of scaffolds; total length of scaffolds); Rickettsia endosymbiont of Cimex lectularius, a-proteobacteria (2; 1,096,646); Nilaparvata lugens yeast-like symbiont, ascomycetes (1; 61,632);Mitochondrion (3; 94,792)"
- . Other observations: "Hifiasm assembly run in Hi-C phasing mode; XO sex chromosome system observed; "

# Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	1,128,364,655	1,202,388,514
GC %	30.23	30.19
Gaps/Gbp	129.39	157.19
Total gap bp	14,600	23,400
Scaffolds	394	198
Scaffold N50	86,774,930	80,806,226
Scaffold L50	4	5
Scaffold L90	11	12
Contigs	540	387
Contig N50	15,213,093	12,755,399
Contig L50	23	27
Contig L90	94	104
QV	59.4	59.3
Kmer compl.	99.70	99.59
BUSCO sing.	95.6%	98.0%
BUSCO dupl.	0.8%	1.5%
BUSCO frag.	0.2%	0.1%
BUSCO miss.	3.4%	0.5%

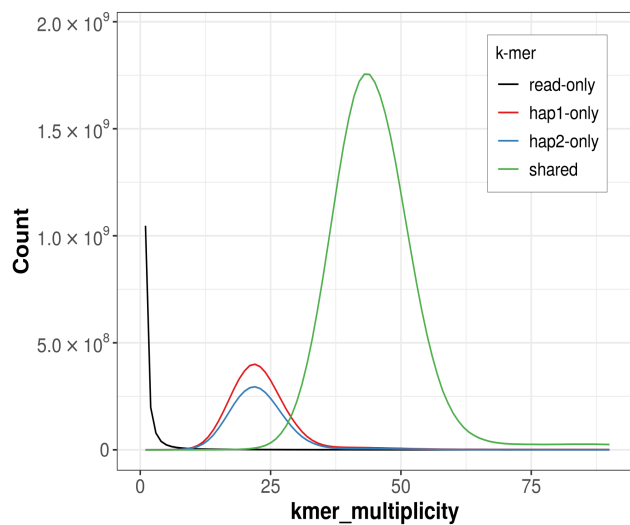
BUSCO 6.0.0 Lineage: insecta\_odb10 (genomes:75, BUSCOs:1367)

# HiC contact map of curated assembly

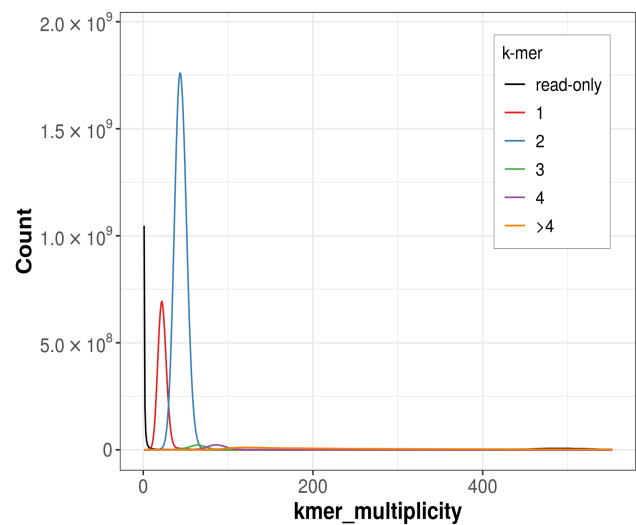


pri [\[LINK\]](#)

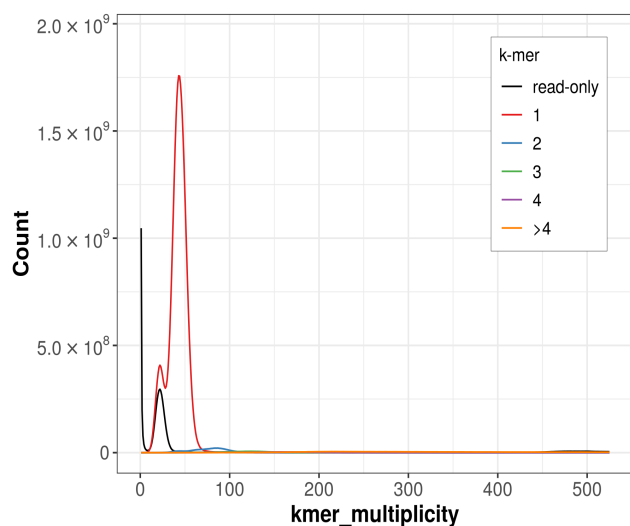
# K-mer spectra of curated assembly



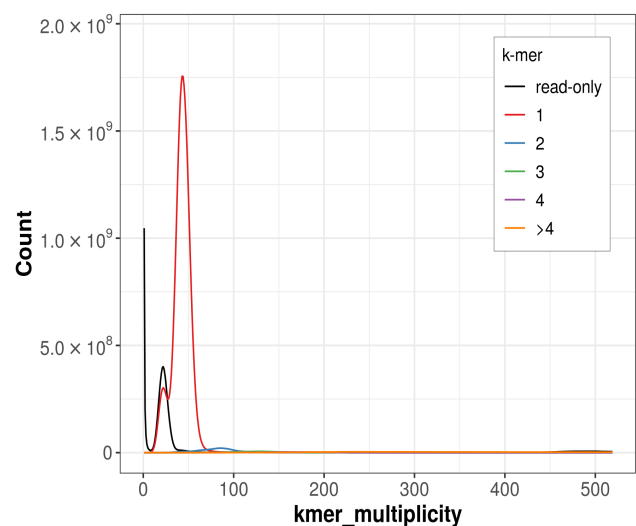
Distribution of k-mer counts coloured by their presence in reads/assemblies



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

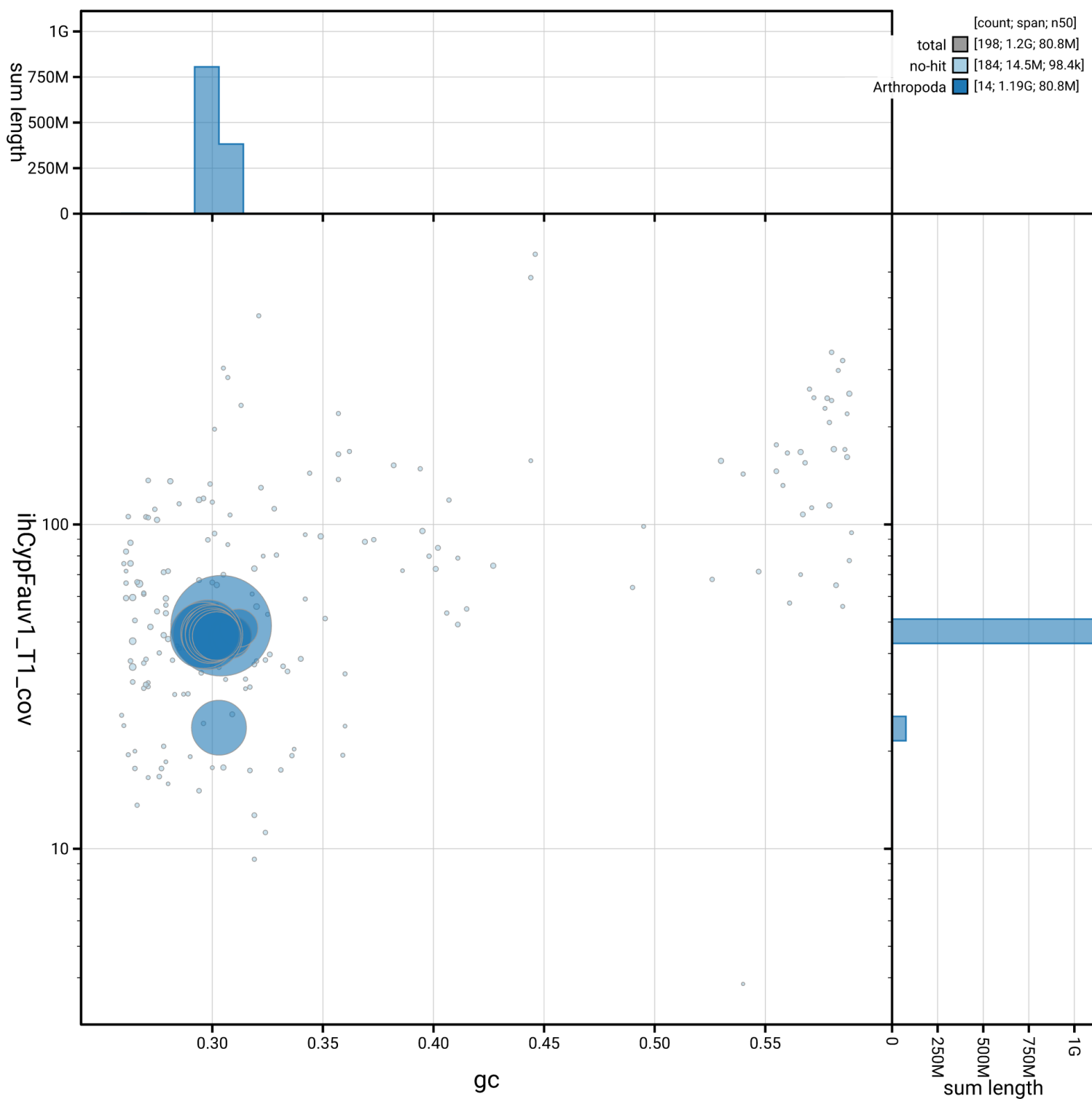


Distribution of k-mer counts per copy numbers found in **hap1** (hapl.)



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

# 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	43x	99x

# Assembly pipeline

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

# Curation pipeline

- **hifiasm**
  - |\_ *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: Thomas Mathers

Affiliation: WSI

Date and time: 2025-09-24 14:34:58 CEST