

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	3086108
ToLID	<b>icSpeNova8</b>
Species	Spelaeobates novaki
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	178,524,440	247,701,476
Haploid Number	11 (source: ancestor)	13
Ploidy	2 (source: ancestor)	2
Sample Sex	unknown	unknown

## EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q62

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
- . Kmer completeness value is less than 90 for pri
- . Assembly length loss > 3% for pri
- . Not 90% of assembly in chromosomes for pri

### Curator notes

- . Interventions/Gb: 140
- . Contamination notes: "Used FCS-GX to remove 1679 contaminant sequences. Dysgonomonas was a prominent cobiont"
- . Other observations: "First round of curation made 3 cuts in contigs, 6 breaks at gaps and 13 joins. A second round made 8 breaks at gaps and 5 joins. Hi-C library was not very good. Relied a lot on the mq10 map for determining chromosome assignment and the mq0 map for visualizing repeats."

# Quality metrics table

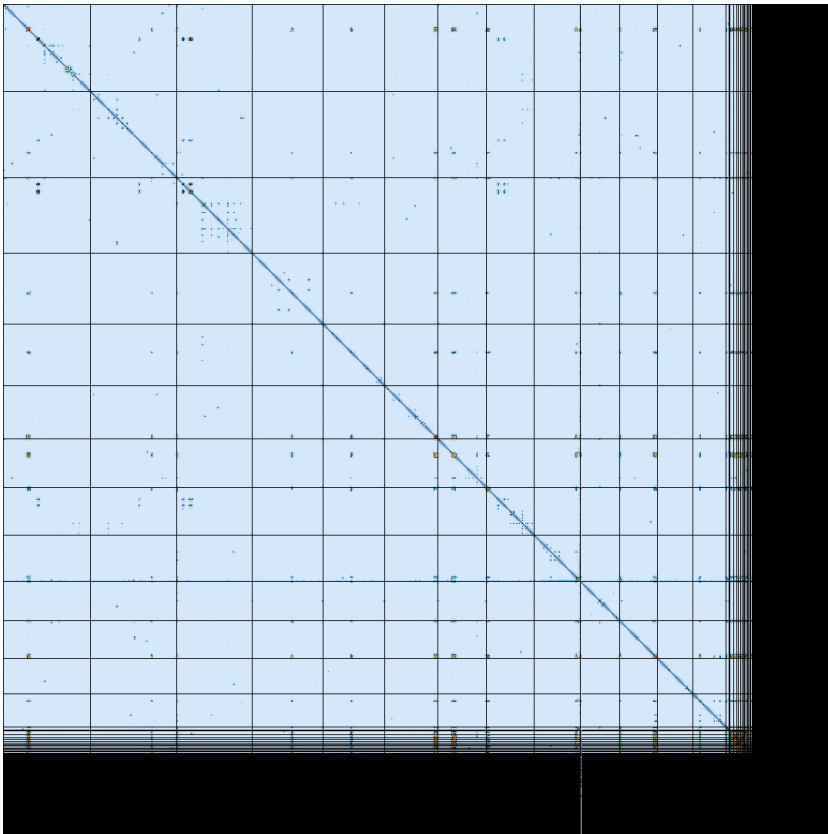
Metrics	Pre-curation pri	Curated pri
Total bp	485,711,178	247,701,476
GC %	36.31	31.91
Gaps/Gbp	1,210.6	294.71
Total gap bp	117,600	14,600
Scaffolds	2,237	552
Scaffold N50	5,227,499	15,758,168
Scaffold L50	17	6
Scaffold L90	670	29
Contigs	2,825	625
Contig N50	2,313,496	5,965,395
Contig L50	41	13
Contig L90	1,038	64
QV	57.2829	62.2597
Kmer compl.	10.5669	70.6668
BUSCO sing.	98.8%	97.8%
BUSCO dupl.	0.5%	0.3%
BUSCO frag.	0.2%	0.5%
BUSCO miss.	0.5%	1.4%

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

BUSCO: 6.0.0 (euk\_genome\_min, miniprot) / Lineage: arthropoda\_odb10 (genomes:90, BUSCOs:1013)

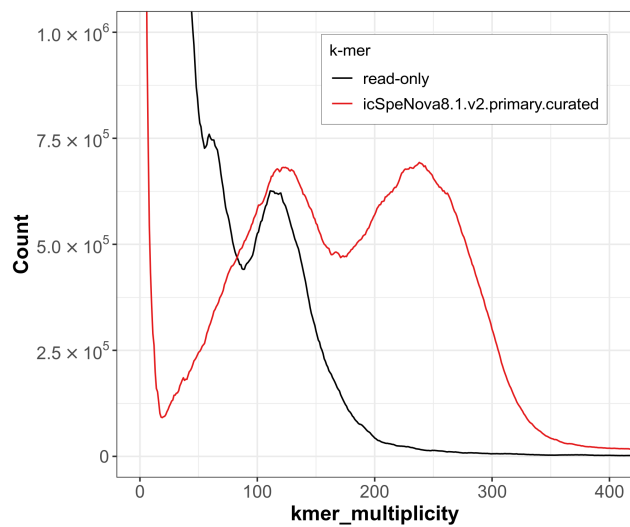
BUSCO: 6.0.0 (euk\_genome\_min, miniprot) / Lineage: arthropoda\_odb12 (genomes:76, BUSCOs:1667)

# HiC contact map of curated assembly

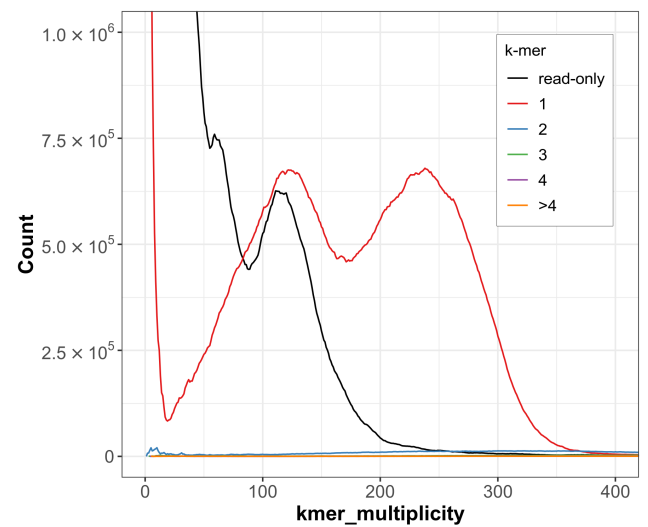


pri [\[LINK\]](#)

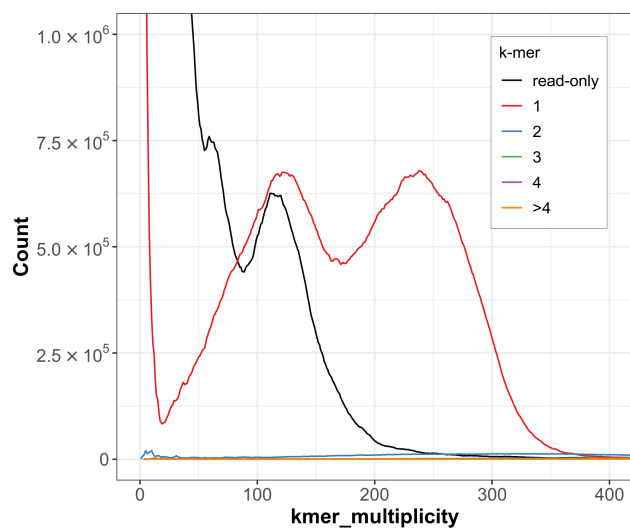
# K-mer spectra of curated assembly



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

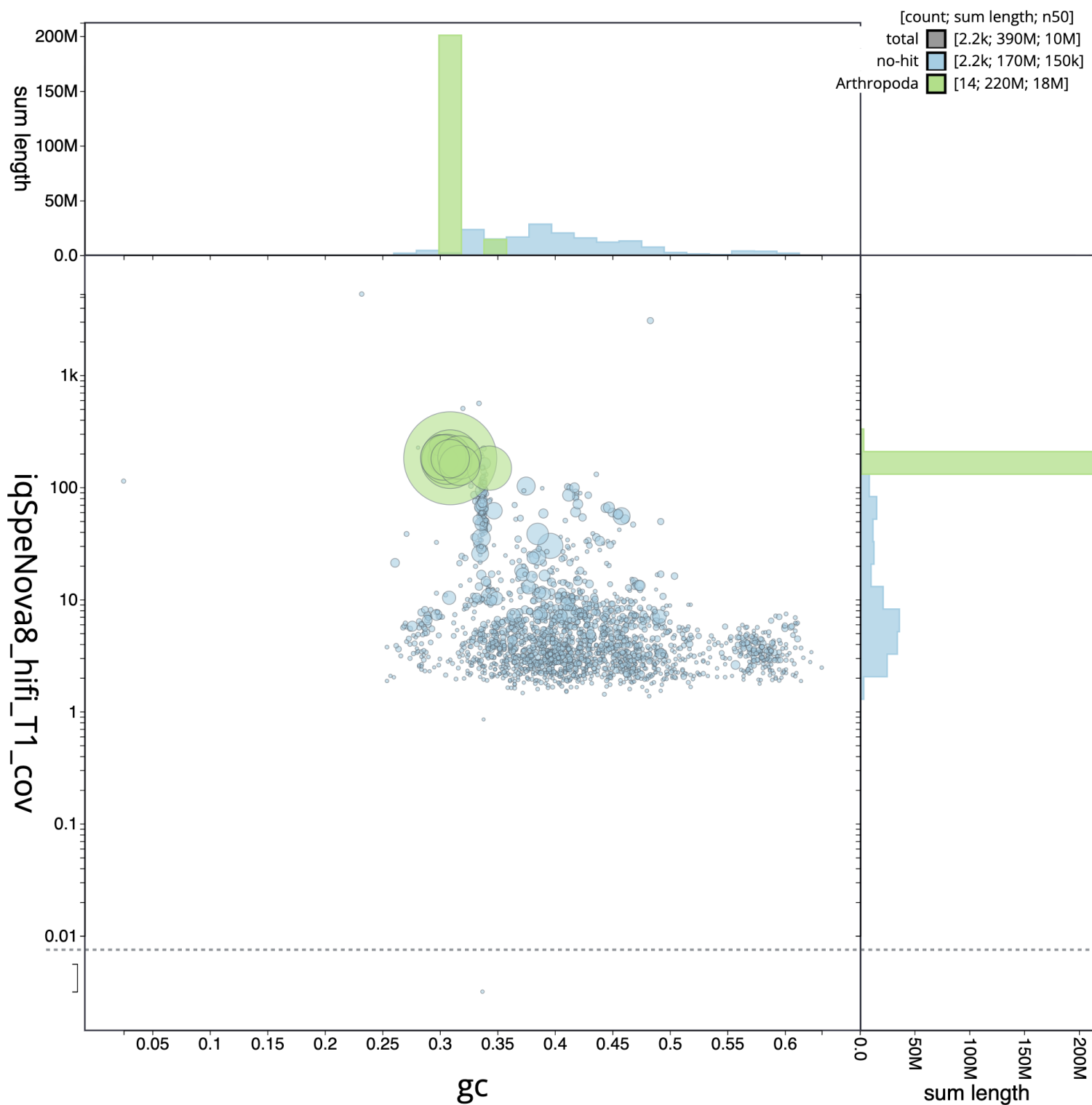


Distribution of k-mer counts per copy numbers found in asm



icSpeNova8.1.v2.primary.curated.spectra-cn.m40  
0.1n.png

# 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 Hi-C
Coverage	195x	455x

# Assembly pipeline

- **hifiasm**
  - |\_ *ver*: 0.24.0-r702
  - |\_ *key param*: NA
- **purge\_dups**
  - |\_ *ver*: 1.2.6
  - |\_ *key param*: NA
- **YaHS**
  - |\_ *ver*: 1.2a
  - |\_ *key param*: NA
- **CLAWS**
  - |\_ *ver*: 2.3
  - |\_ *key param*: NA

# Curation pipeline

- **PretextViewAI**
  - |\_ *ver*: 1.0.5
  - |\_ *key param*: NA
- **GRIT\_Rapid**
  - |\_ *ver*: 2.0
  - |\_ *key param*: NA

Submitter: Tyler Alioto

Affiliation: CNAG

Date and time: 2026-02-26 00:45:46 CET