

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	1342031
ToLID	<b>icLarUrsu1</b>
Species	Larinus ursus
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	1,613,542,317	1,753,491,222
Haploid Number	19 (source: ancestor)	21
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	NA

## EBP metrics summary and curation notes

Obtained EBP quality metric for Pri: 7.7.Q62

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

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

### Curator notes

- . Interventions/Gb: None
- . Contamination notes: "Blobtoolkit detected no contamination."
- . Other observations: "This assembly was obtained with Hifiasm from ONT data.

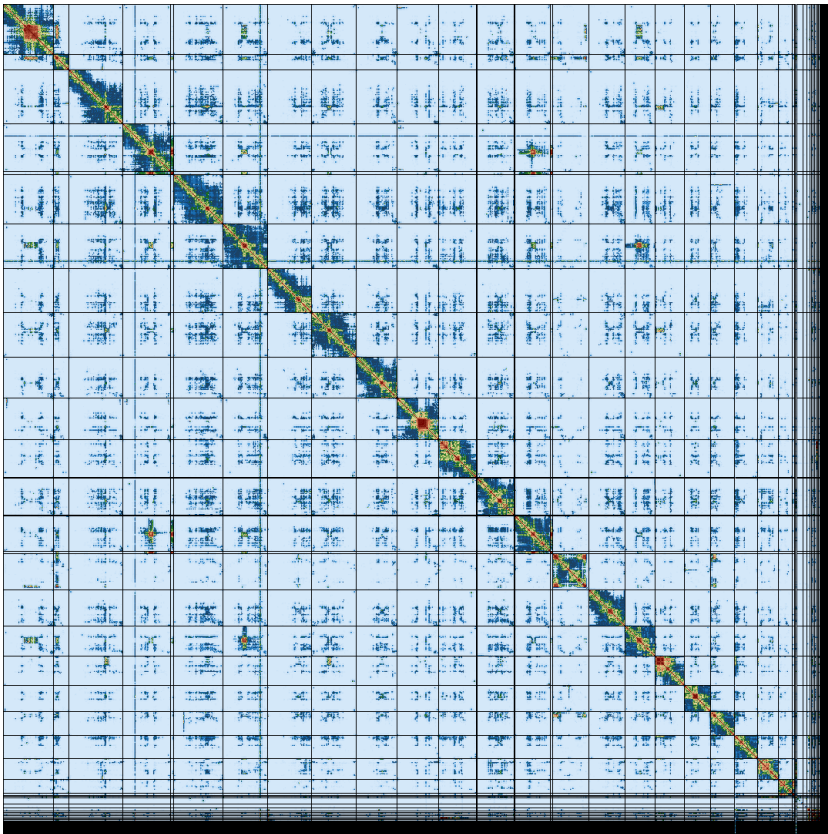
Mitogenome was successfully assembled with FOAM. Curation made 10 cuts in contigs, 3 breaks at gaps and 15 joins. Four haplotig sequences, adding up a total of 25.2 Mb of sequence, were removed during curaton."

# Quality metrics table

Metrics	Pre-curation Pri	Curated Pri
Total bp	1,778,658,397	1,753,491,222
GC %	35.31	35.32
Gaps/Gbp	20.8	27.94
Total gap bp	7,400	9,800
Scaffolds	179	173
Scaffold N50	86,794,555	86,231,057
Scaffold L50	9	9
Scaffold L90	22	20
Contigs	216	222
Contig N50	48,742,161	45,976,650
Contig L50	14	15
Contig L90	42	45
QV	62.4669	62.4982
Kmer compl.	79.8695	79.6503
BUSCO sing.	98.5%	98.7%
BUSCO dupl.	1.5%	1.3%
BUSCO frag.	0.0%	0.0%
BUSCO miss.	0.0%	0.0%

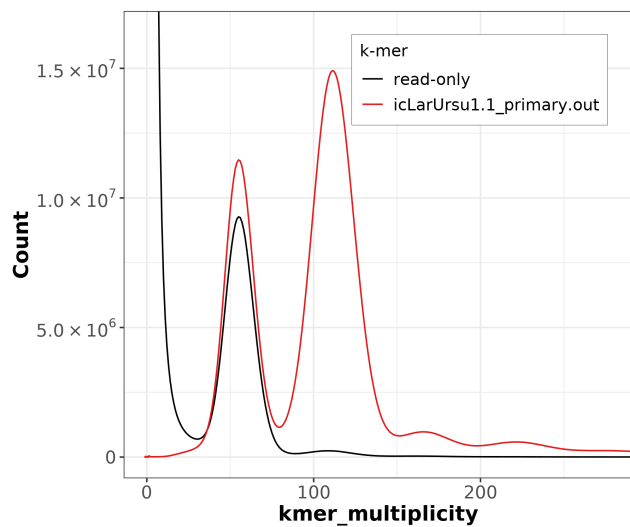
BUSCO: 6.0.0 (euk\_genome\_min, miniprot) / Lineage: insecta\_odb10 (genomes:75, BUSCOs:1367)

# HiC contact map of curated assembly

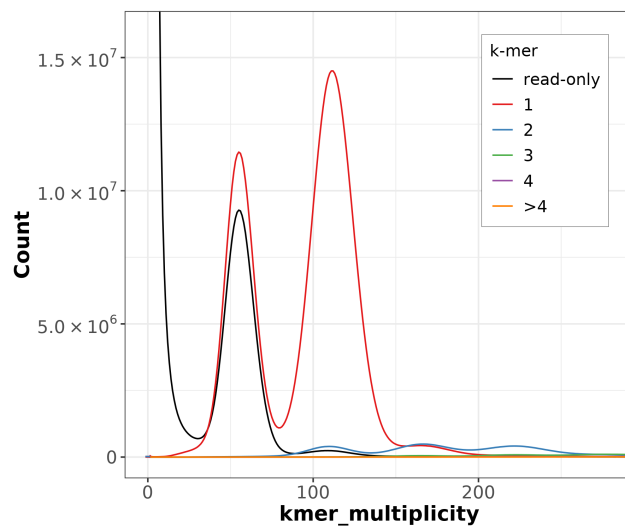


Pri [\[LINK\]](#)

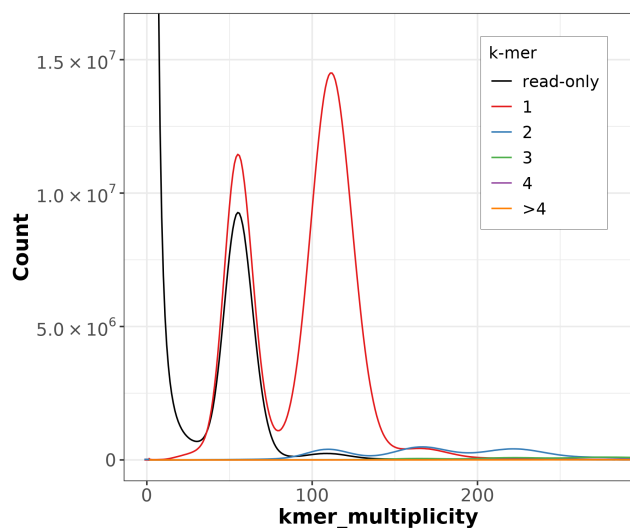
# K-mer spectra of curated assembly



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

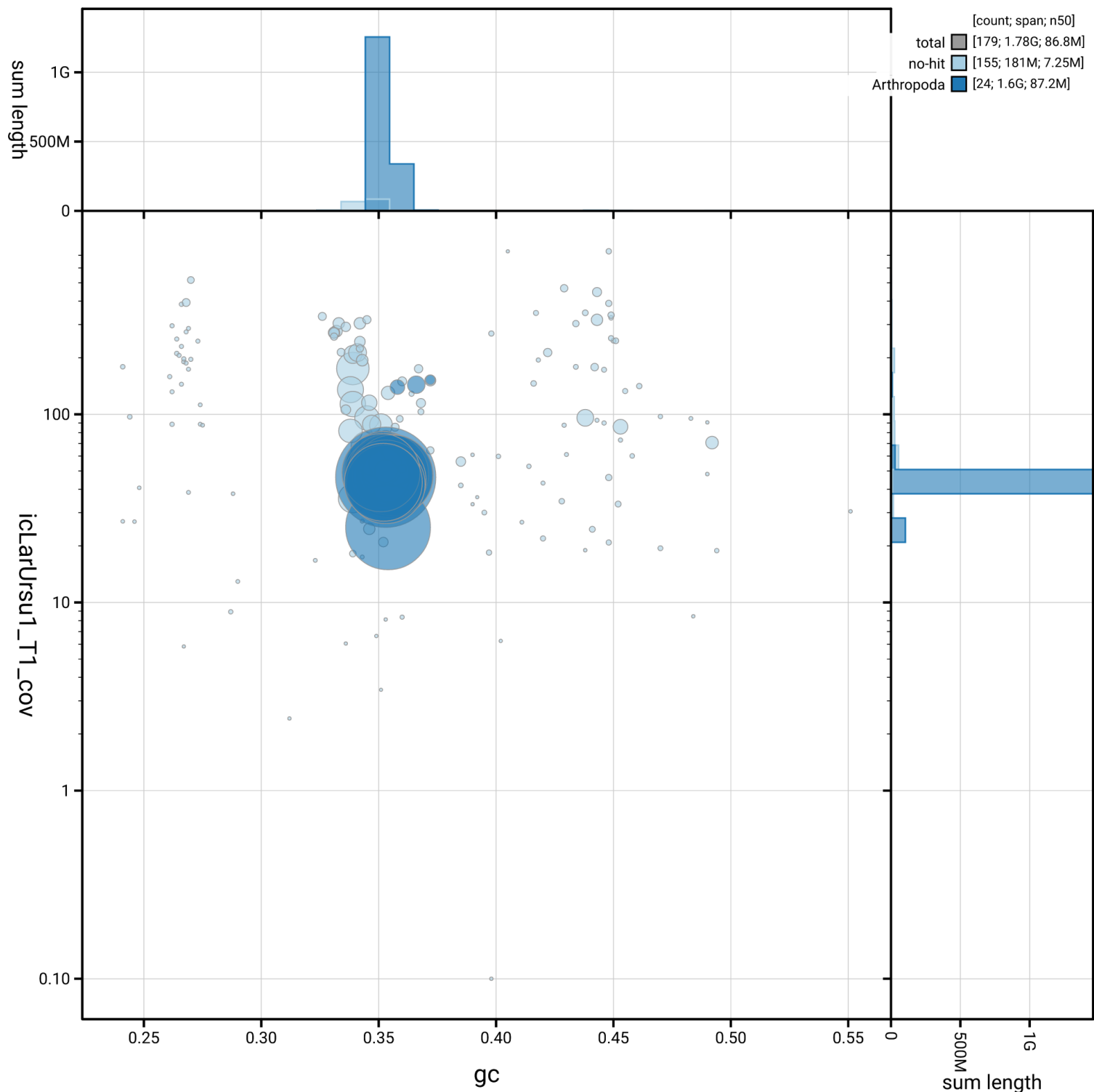


Distribution of k-mer counts per copy numbers found in asm



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

# 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	ONT	Illumina	Omni-C
Coverage	42	83	61

## Assembly pipeline

- **CLAWS**
  - |\_ *ver*: v2.3.0
  - |\_ *key param*: NA
- **Filtlong**
  - |\_ *ver*: v0.2.1
  - |\_ *key param*: --minlen 1000
  - |\_ *key param*: --min\_mean\_q 80
- **Hifiasm**
  - |\_ *ver*: 0.24.0
  - |\_ *key param*: --ont
  - |\_ *key param*: --telo-m AAATAACCC
- **Yahs**
  - |\_ *ver*: 1.2a
  - |\_ *key param*: -mq 10
  - |\_ *key param*: --no-contig-ec
- **FOAM**
  - |\_ *ver*: 0.5
  - |\_ *key param*: NA

## Curation pipeline

- **Blobtoolkit Nextflow pipeline**
  - |\_ *ver*: 0.6
  - |\_ *key param*: NA
- **PretextViewAI**
  - |\_ *ver*: 1.0.5
  - |\_ *key param*: NA

Submitter: Jessica Gomez-Garrido

Affiliation: CNAG

Date and time: 2026-01-12 11:11:35 CET