ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	3073150
ToLID	**icAdeBosn1,icAdeBosn2,icAdeBosn3**
Species	Adelopsella bosnica
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	768,177,425	946,343,750
Haploid Number	11 (source: ancestor)	14
Ploidy	2 (source: ancestor)	2
Sample Sex	unknown	unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for : 6.7.Q57

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

. Assembly length loss > 3% for
. Not 90% of assembly in chromosomes for

Curator notes

. Interventions/Gb: 186
. Contamination notes: "Removed 541 scaffolds identified by Blobtools as contaminants. We are running blobtools again on the decontaminated curated assembly again, especially to correct the axis and figure legends that have the wrong species tolid! It will be included in any updated EAR."
. Other observations: "Curation made 35 cuts in contigs, 48 breaks at gaps and 103 joins. Removed 14 haplotigs (3 Mb). Centromeric regions were mostly unlocalized (order and orientation uncertain) using strong contact weith localized centromeric sequence and contacts with non-centromeric flanking sequence."

Quality metrics table

| Metrics	Pre-curation <pri>	Curated <pri>
Total bp	991,324,255	946,343,750
GC %	30.48	30.2
Gaps/Gbp	388.37	456.49
Total gap bp	77,000	86,400
Scaffolds	2,271	1,698
Scaffold N50	36,347,139	50,659,766
Scaffold L50	10	7
Scaffold L90	473	259
Contigs	2,656	2,130
Contig N50	2,300,463	2,223,258
Contig L50	126	126
Contig L90	826	659
QV	57.5578	57.8464
Kmer compl.	94.0162	93.2434
BUSCO sing.	94.4%	94.5%
BUSCO dupl.	4.8%	4.7%
BUSCO frag.	0.4%	0.4%
BUSCO miss.	0.4%	0.4%

BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: insecta_odb10 (genomes:75, BUSCOs:1367)
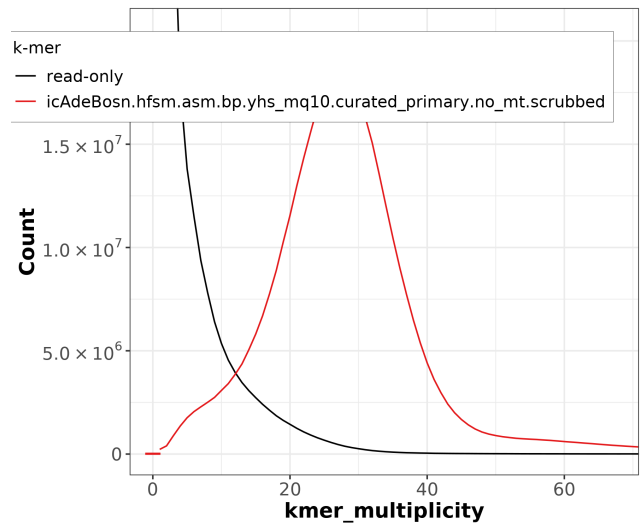
HiC contact map of curated assembly
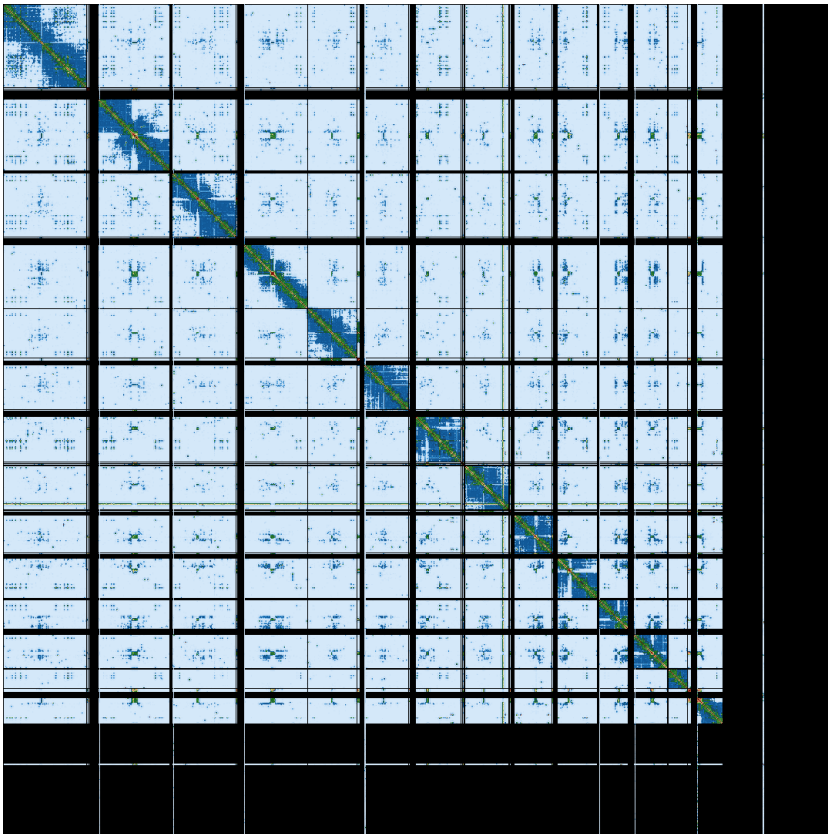
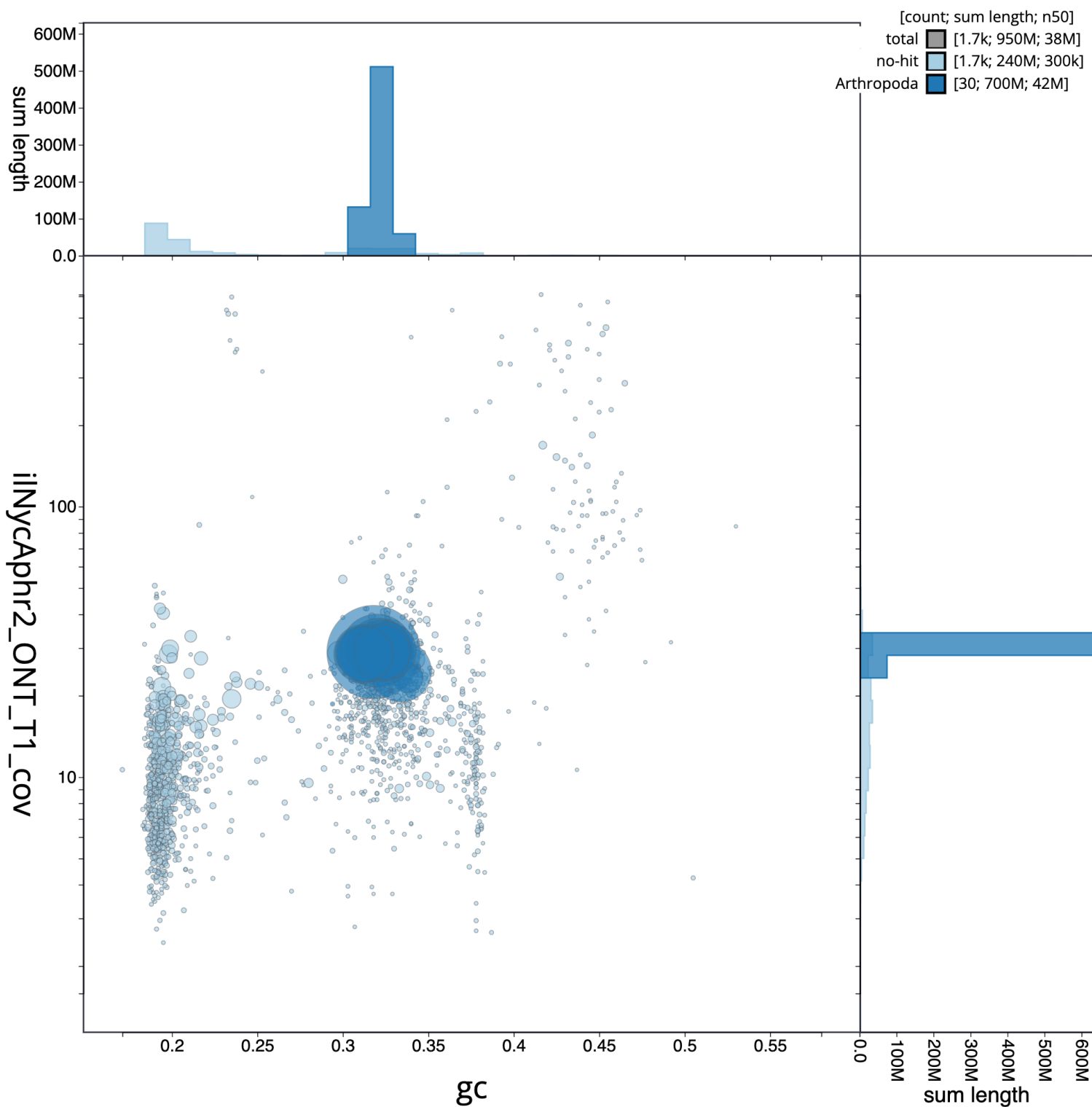
[LINK]

K-mer spectra of curated assembly
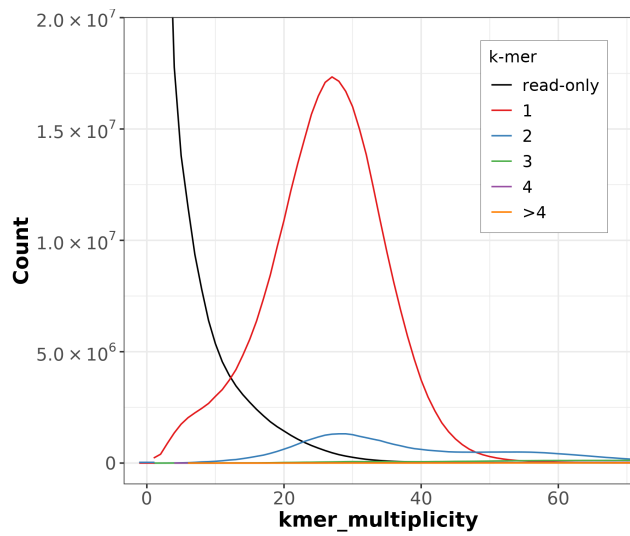
Distribution of k-mer counts per copy numbers found in asm

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

Post-curation contamination screening

. 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
Coverage	24x	106x

Assembly pipeline

- **Filtlong**
    |_ *ver:* 0.2.1
    |_ *key param:* --min_length 1000 --min_mean_q 80
- **Hifiasm**
    |_ *ver:* 0.24.0
    |_ *key param:* --hom-cov 28.4 --n-perturb 30000 --f-perturb 0.3 --min-hist-cnt 5
-N 80 --n-hap 4 --telo-m AATAAT"
- **YaHS**
    |_ *ver:* 1.2a
    |_ *key param:* --no-contig-ec
- **CLAWS pipeline**
    |_ *ver:* 2.3.0
    |_ *key param:* NA
- **Blobtoolkit Nextflow pipeline (latest)**
    |_ *ver:* 0.6
    |_ *key param:* NA
- **FOAM pipeline**
    |_ *ver:* 0.5
    |_ *key param:* NA
- **mitos**
    |_ *ver:* 2.1.3
    |_ *key param:* NA

Curation pipeline

- **GRIT_Rapid**
    |_ *ver:* 2.0
    |_ *key param:* NA
- **PretextViewAI**
    |_ *ver:* 1.0.4
    |_ *key param:* NA

Submitter: Tyler Alioto
Affiliation: CNAG

Date and time: 2025-09-26 16:34:07 CEST