ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	3023605
ToLID	**icCarBano7**
Species	Carabus banonii
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	71,013,296	207,791,100
Haploid Number	14 (source: ancestor)	15
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 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

. Kmer completeness value is less than 90 for collapsed
. Assembly length loss > 3% for collapsed

Curator notes

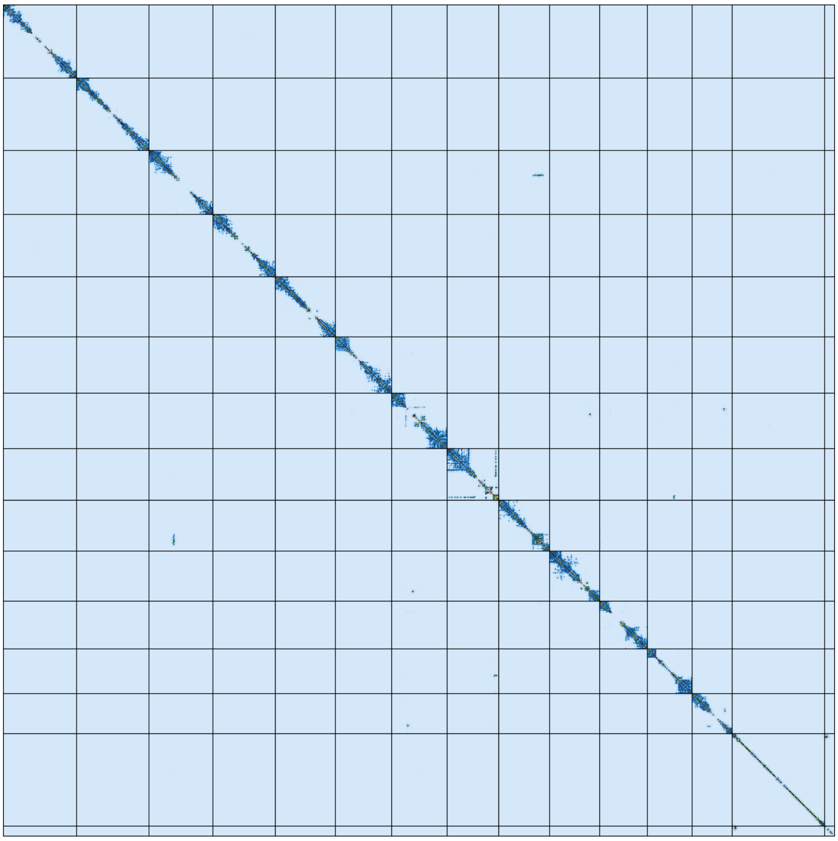
. Interventions/Gb: 426
. Contamination notes: "Only one contig was detected as bacterial contaminant. Some small scaffolds for which no HiC links was found were removed "
. Other observations: "Presence of larges repeated regions (telomeric or centromeric?) at the ends of the scaffolds,therefore need to invert the arms of chromosomes "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	215,432,743	207,791,100
GC %	35.44	35.4
Gaps/Gbp	0	336.88
Total gap bp	0	13,800
Scaffolds	184	18
Scaffold N50	12,129,189	14,906,334
Scaffold L50	7	6
Scaffold L90	17	13
Contigs	184	88
Contig N50	12,129,189	7,491,500
Contig L50	7	12
Contig L90	17	28
QV	56.8752	57.4655
Kmer compl.	81.1251	81.1152
BUSCO sing.	98.4%	98.4%
BUSCO dupl.	0.6%	0.6%
BUSCO frag.	0.6%	0.5%
BUSCO miss.	0.4%	0.5%

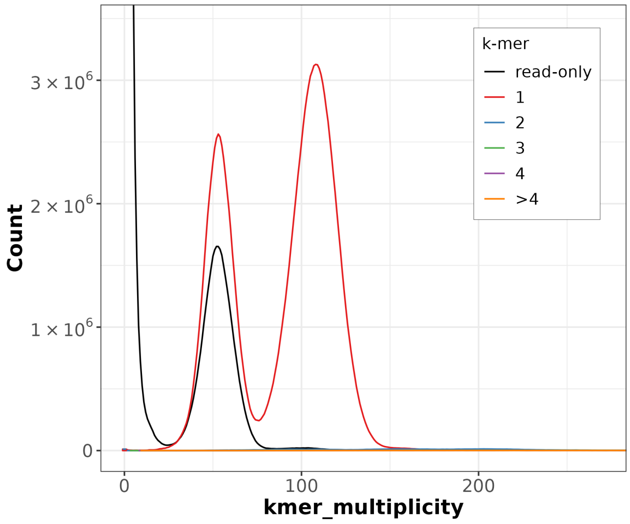
BUSCO: 5.4.3 (euk_genome_met, metaeuk) / Lineage: endopterygota_odb10 (genomes:56, BUSCOs:2124)

HiC contact map of curated assembly

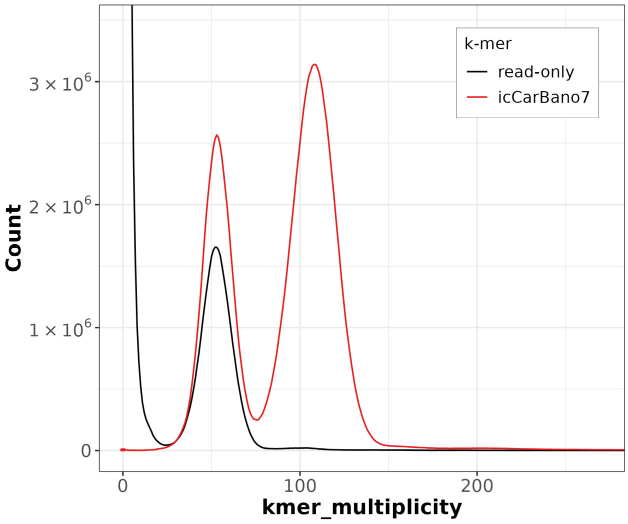


collapsed [LINK]

K-mer spectra of curated assembly

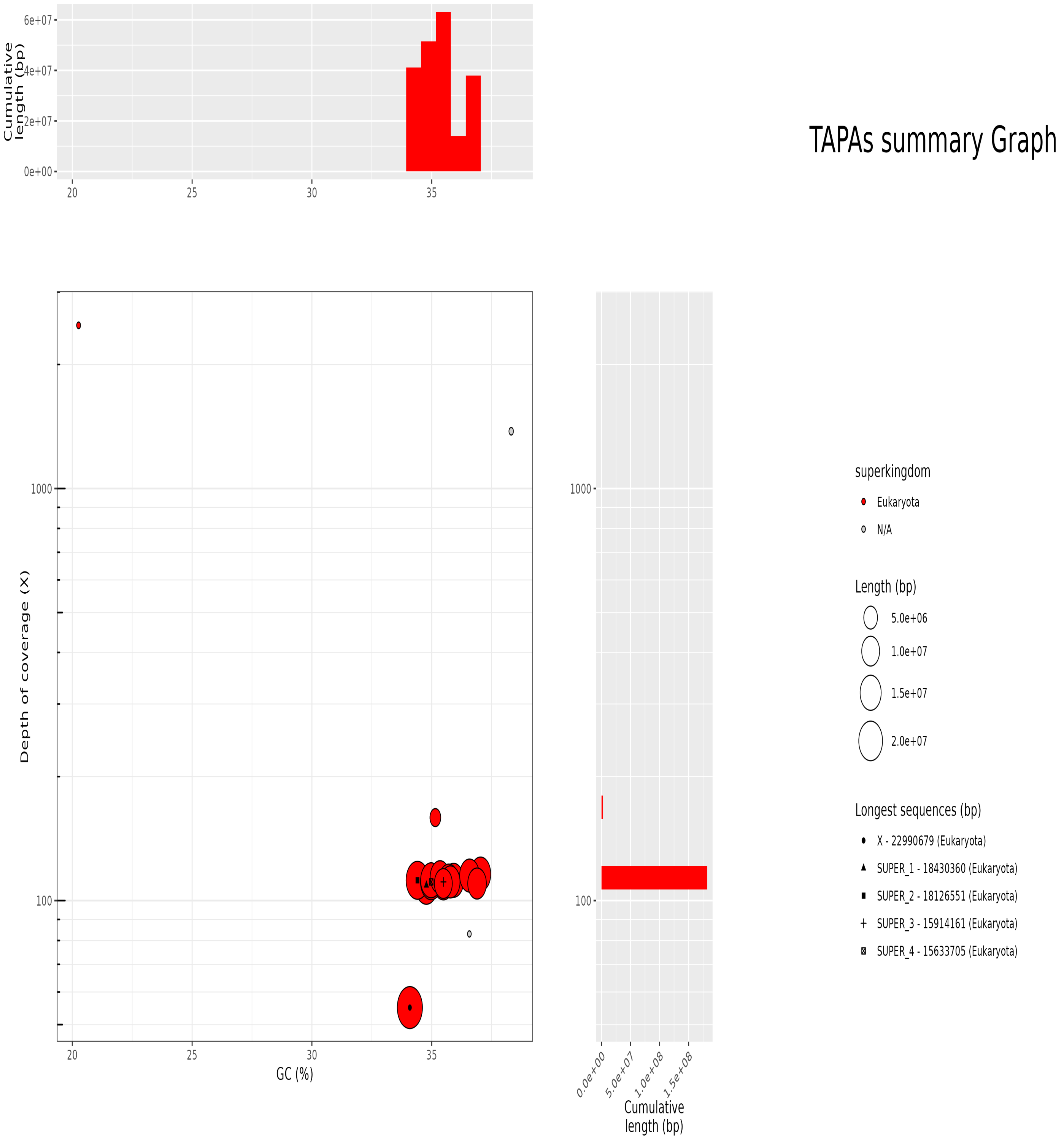


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



collapsed. 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	30	50

Assembly pipeline

- **Hifiasm**
    |_ *ver:* 0.19.5-r593
    |_ *key param:* NA
- **purge_dups**
    |_ *ver:* 1.2.5
    |_ *key param:* NA
- **YaHS**
    |_ *ver:* 1.2
    |_ *key param:* NA

Curation pipeline

- **PretextMap**
    |_ *ver:* 0.1.9
    |_ *key param:* NA
- **PretextView**
    |_ *ver:* 0.2.5
    |_ *key param:* NA

Submitter: Arnaud Couloux
Affiliation: Genoscope

Date and time: 2025-05-21 15:25:47 CEST