ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	1849190
ToLID	**icAntGang1**
Species	Anthroherpon ganglbaueri
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	289,632,189	245,273,758
Haploid Number	11 (source: ancestor)	11
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	NA

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 5.7.Q64

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

. Assembly length loss > 3% for hap1
. More than 1000 gaps/Gbp for hap1
. Not 90% of assembly in chromosomes for hap1

Curator notes

. Interventions/Gb: 66
. Contamination notes: "Protobacteria contaminants were found in two scaffolds and removed with the pipeline NCBI FCS GX v Galaxy Version 0.5.5+galaxy1. Scaffolds that are flagged as contaminants by BlobToolKit were confirmed with FCS output table and with contact in the map as non-contaminants. Contact for scaffolds was also confirmed using HiGlass"
. Other observations: "Only one X chromosomes was found with alignement with close relatives Leptodirus hochenwartii and Catops nigricans"
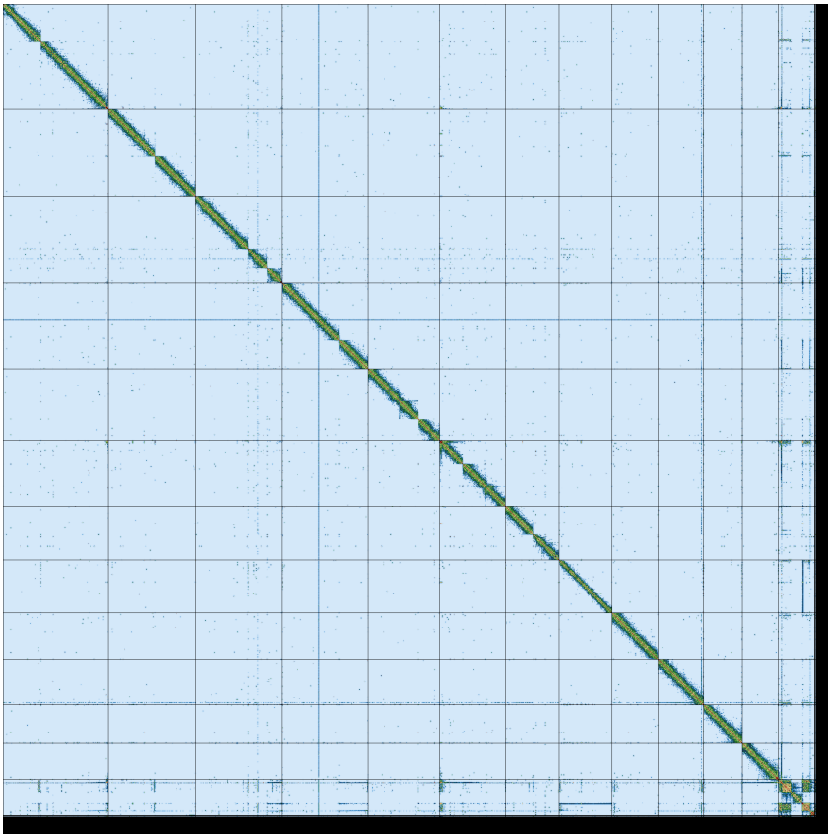
Quality metrics table

Metrics	Pre-curation hap1	Curated hap1
Total bp	332,096,506	245,273,758
GC %	32.02	32.53
Gaps/Gbp	1,165.32	1,777.61
Total gap bp	77,400	87,200
Scaffolds	1,498	157
Scaffold N50	22,404,328	21,074,117
Scaffold L50	6	5
Scaffold L90	609	12
Contigs	1,885	593
Contig N50	1,050,880	954,296
Contig L50	89	93
Contig L90	952	273
QV	60.1544	64.9846
Kmer compl.	91.1711	91.6898
BUSCO sing.	96.6%	97.7%
BUSCO dupl.	1.0%	1.1%
BUSCO frag.	0.7%	0.1%
BUSCO miss.	1.7%	1.2%

Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 5.8.0 (euk_genome_met, metaeuk) / Lineage: insecta_odb10 (genomes:75, BUSCOs:1367)
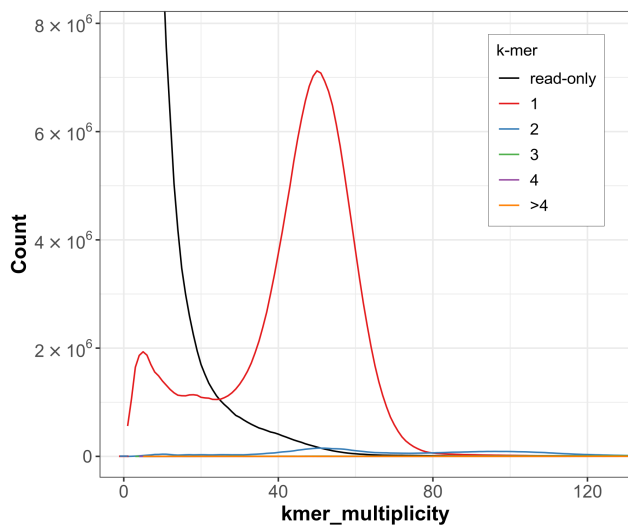BUSCO: 5.8.0 (euk_genome_min, miniprot) / Lineage: insecta_odb10 (genomes:75, BUSCOs:1367)
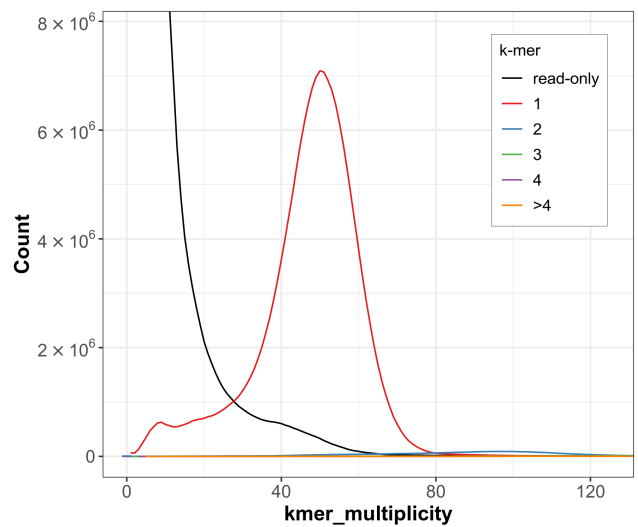
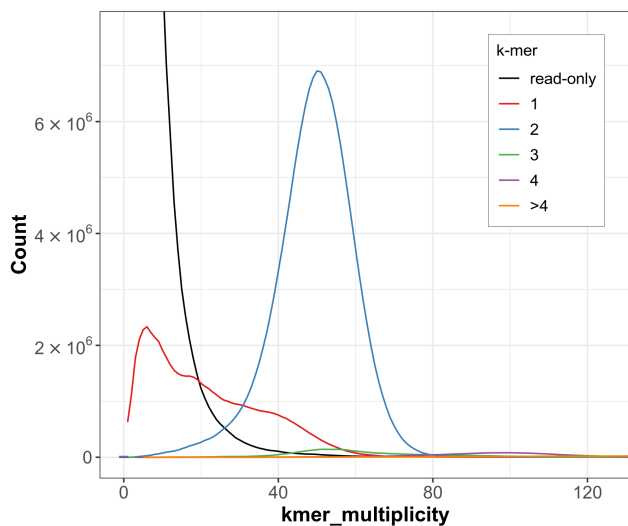# HiC contact map of curated assembly

hap1 [LINK]
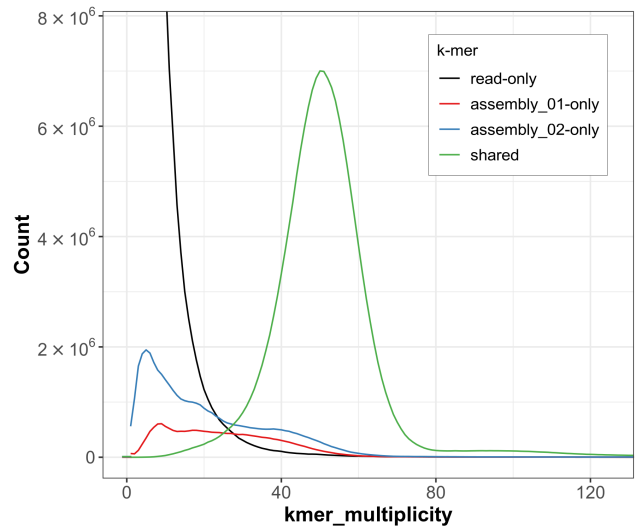
K-mer spectra of curated assembly

Distribution of k-mer counts per copy numbers found in **assembly_02** (hapl.)

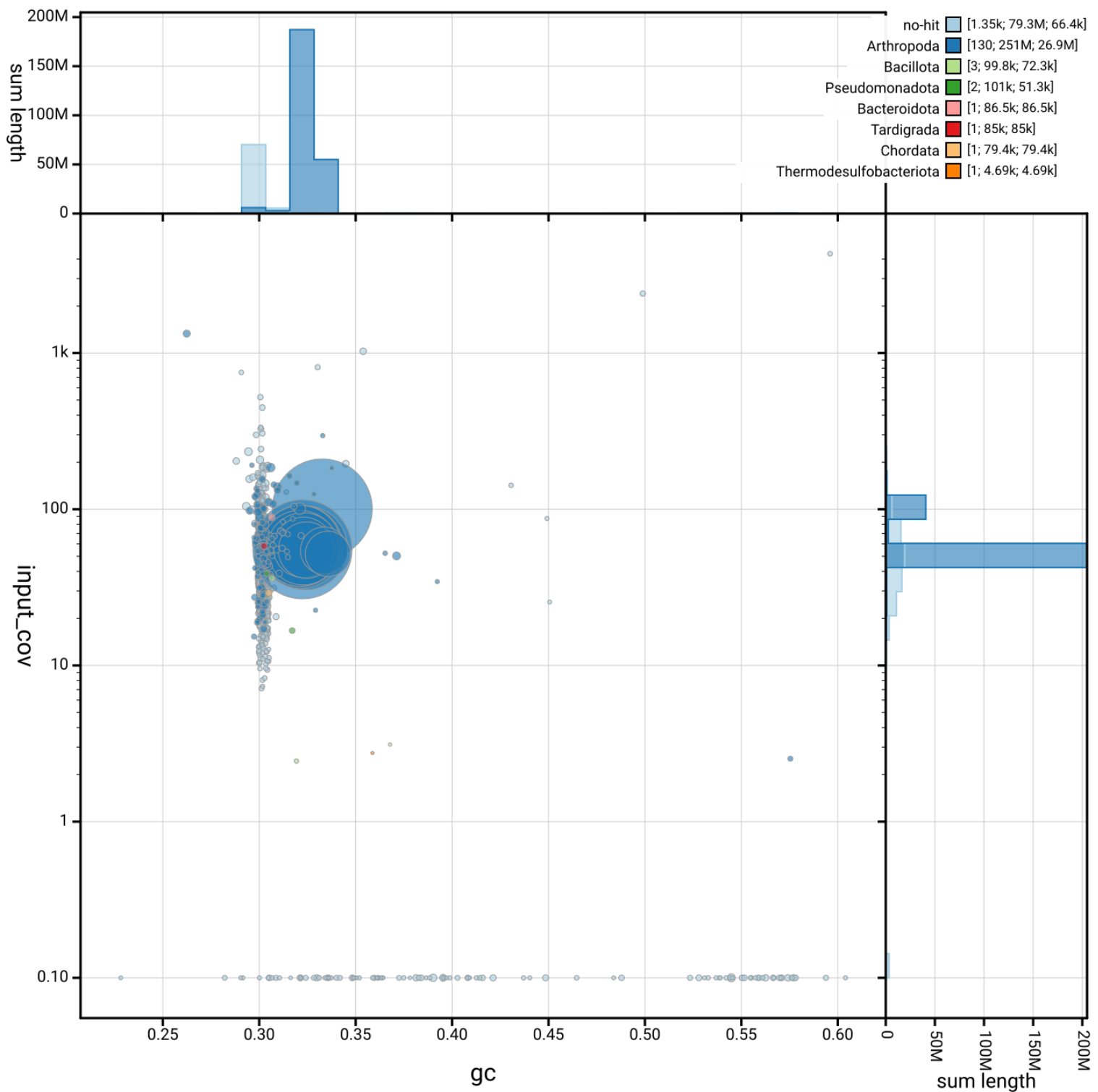Distribution of k-mer counts per copy numbers found in **assembly_01** (hapl.)

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

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

Post-curation contamination screening

hap1. 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	HiFi	HiC
Coverage	58.8x	312x

Assembly pipeline

- **Hifiasm**
    - |_ *ver:* Galaxy Version 0.24.0+galaxy0
    - |_ *key param:* VGP Galaxy workflow VGP3 v 0.2.8
- **purge_dups**
    - |_ *ver:* Galaxy Version 1.2.6+galaxy0
    - |_ *key param:* VGP Galaxy workflow VGP6 v 0.5
- **YaHS**
    - |_ *ver:* Galaxy Version 1.2a.2+galaxy2
    - |_ *key param:* VGP Galaxy workflow VGP8 v 1.4

Curation pipeline

- **PretextMap**
    - |_ *ver:* Galaxy Version 0.1.9+galaxy1
    - |_ *key param:* VGP workflow VGP PretextMap generation v2
- **HiGlass**
    - |_ *ver:* 1.0
    - |_ *key param:* NA
- **PretextView**
    - |_ *ver:* 1.0.0
    - |_ *key param:* NA

Submitter: Marcella Sozzoni
Affiliation: UNIFI

Date and time: 2026-01-19 12:34:20 CET