ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	2744207
ToLID	**ilGlaPaph6**
Species	Glaucopsyche paphos
Class	Insecta
Order	Lepidoptera

Genome Traits	Expected	Observed
Haploid size (bp)	521,721,393	510,089,785
Haploid Number	23 (source: ancestor)	23
Ploidy	2 (source: ancestor)	2
Sample Sex	ZZ	ZZ

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.7.Q64
Obtained EBP quality metric for hap2: 7.7.Q64

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

. Kmer completeness value is less than 90 for hap1
. Kmer completeness value is less than 90 for hap2

Curator notes

. Interventions/Gb: 0
. Contamination notes: "Wolbachia was assembled (present in both haps) and removed (labeled in pretextview) before running Blobtools, which detected no presence of any further contaminants. Mitochondrial genome has not been assembled, yet. Not detected in the assembly."
. Other observations: "Haplotypes were curated separately. Pre-curation hap1 had some haplotigs an minor corrections (Curation made 7 cuts in contigs, 3 breaks at gaps and 4 joins). Hap2 did not require any corrections. Just a few unlocs made. After curation, a new hap1 was made with best (fewest gaps) SUPER per chromosome chosen. The new hap2 assembly was given the more gappy scaffolds. Assignment of the Z was based on alignment to FR990042.1
https://www.ebi.ac.uk/ena/browser/view/GCA_905404095.1. "

Quality metrics table

Metrics	Pre-curation hap1	Pre-curation hap2	Curated hap1	Curated hap2
Total bp	516,444,632	507,012,600	510,089,785	505,963,101
GC %	35.89	35.84	35.85	35.89
Gaps/Gbp	56.15	78.89	27.45	110.68
Total gap bp	5,800	8,000	2,800	11,200
Scaffolds	42	31	32	40
Scaffold N50	21,475,491	21,505,828	21,718,161	21,475,491
Scaffold L50	9	10	10	10
Scaffold L90	19	20	20	20
Contigs	71	71	46	96
Contig N50	17,799,701	12,384,290	20,270,075	10,873,703
Contig L50	11	13	11	16
Contig L90	31	38	23	49
QV	64.8564	64.7011	64.6016	64.9782
Kmer compl.	80.7313	79.9426	80.2513	79.4864
BUSCO sing.	98.0%	97.9%	98.7%	97.8%
BUSCO dupl.	1.3%	0.2%	0.5%	0.3%
BUSCO frag.	0.4%	0.4%	0.4%	0.4%
BUSCO miss.	0.3%	1.5%	0.4%	1.5%

BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)
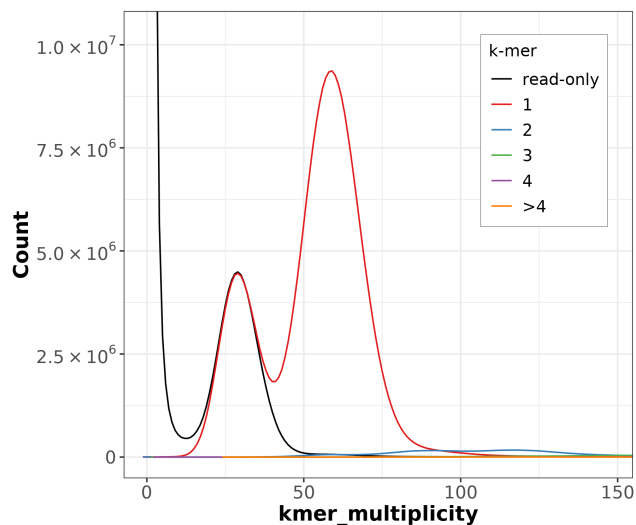
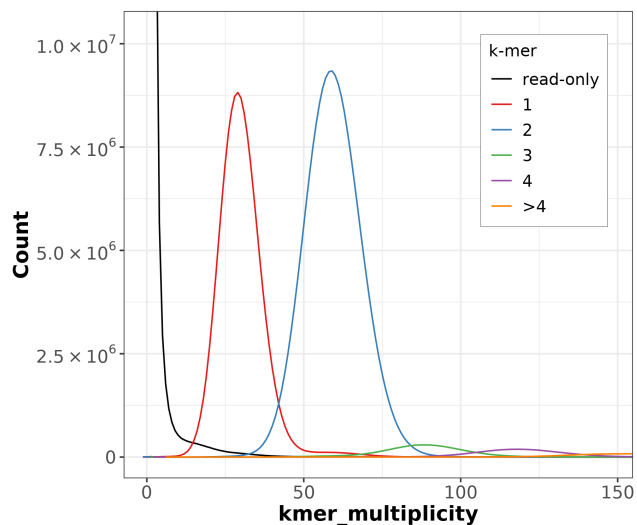# HiC contact map of curated assembly
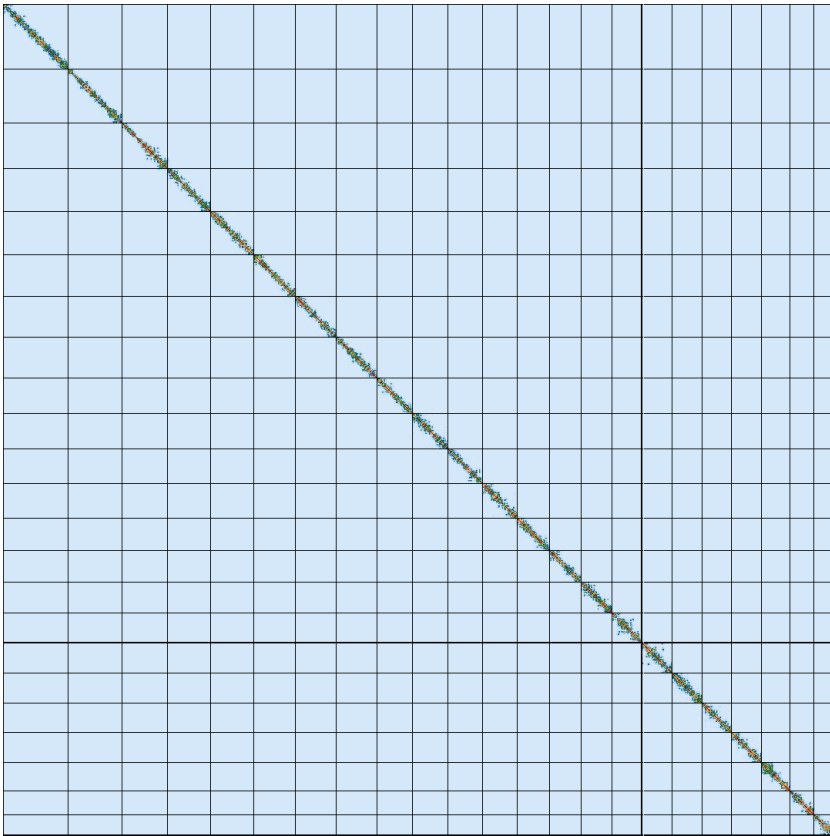
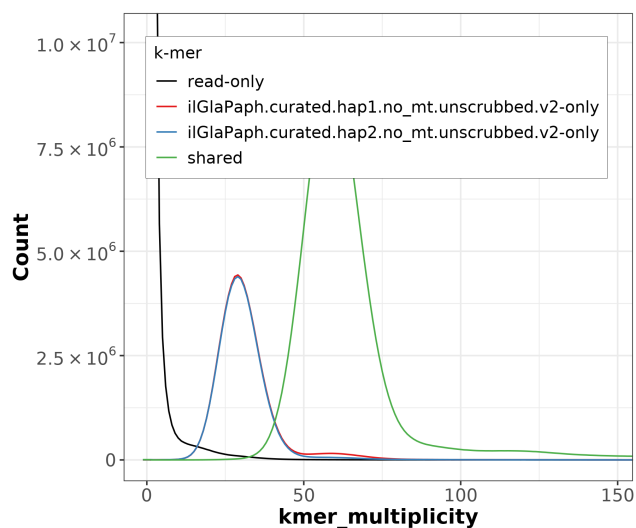**hap1** [LINK]
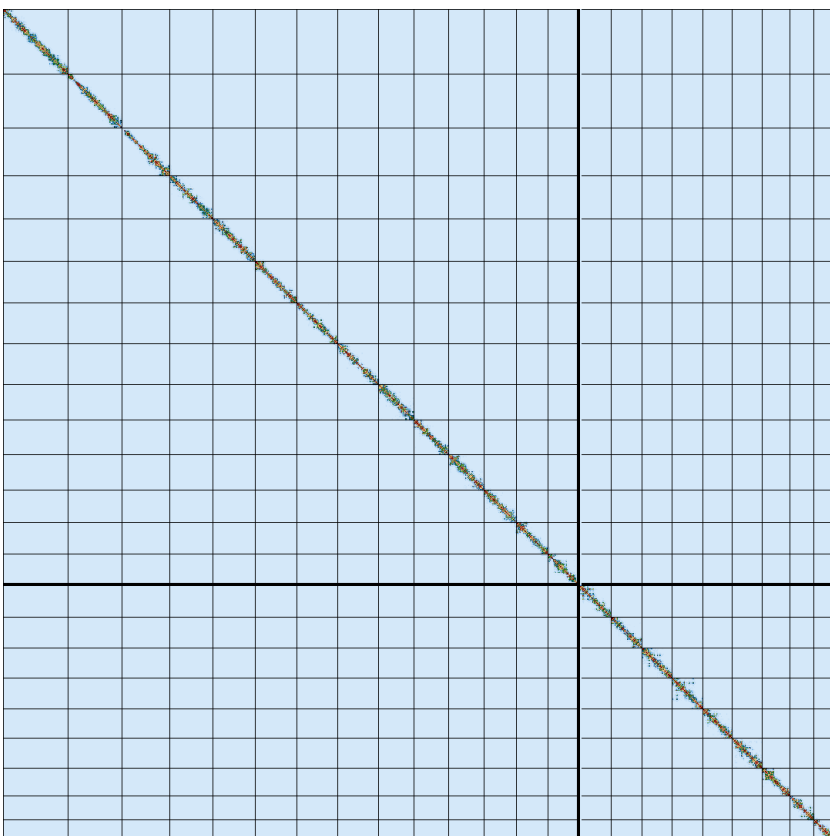
hap2 [LINK]

K-mer spectra of curated assembly
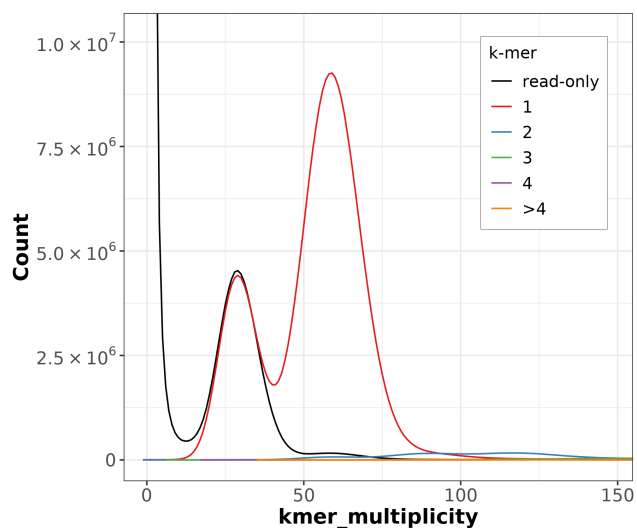

Distribution of k-mer counts per copy numbers found in **hap1** (hapl.)

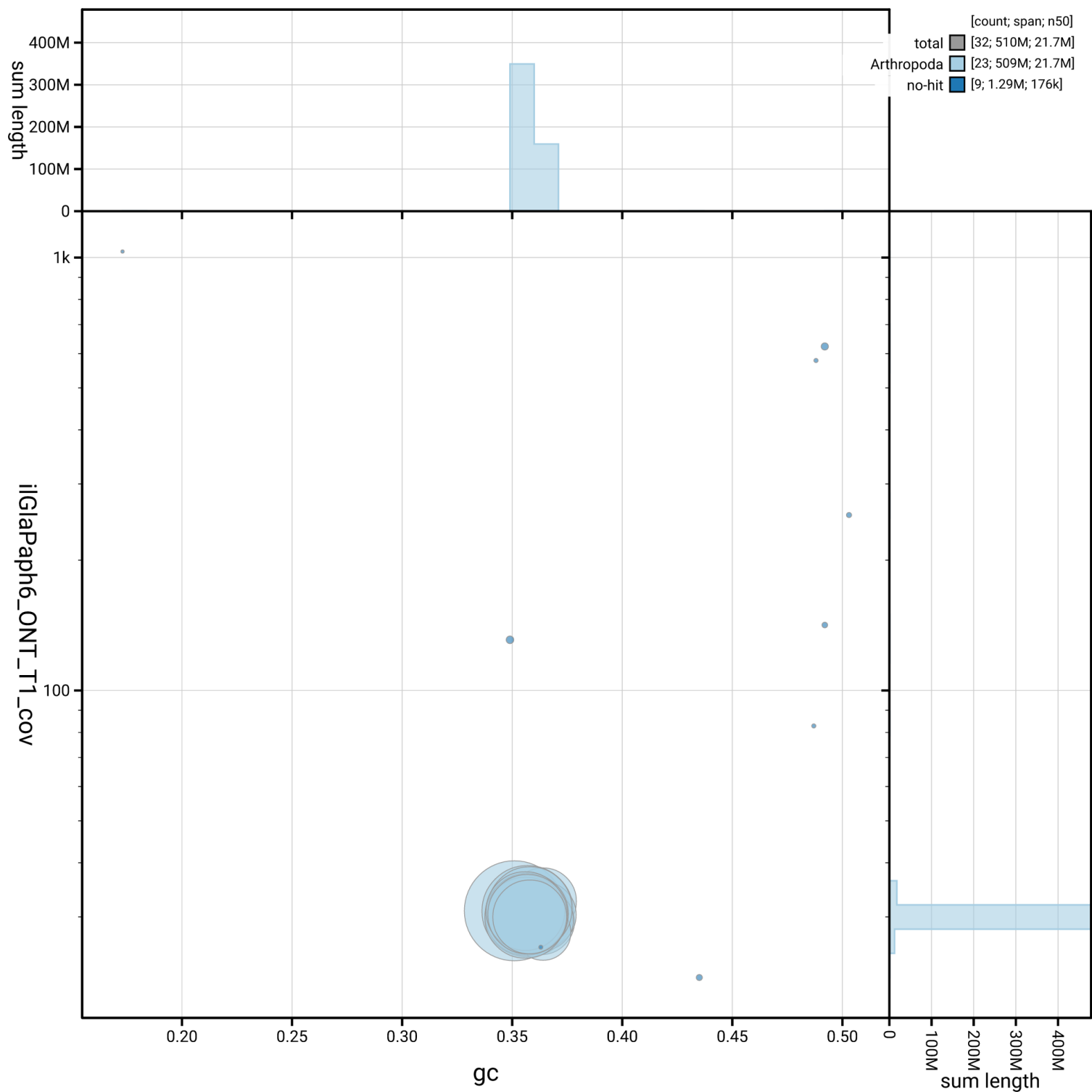Distribution of k-mer counts per copy numbers found in asm (dipl.)

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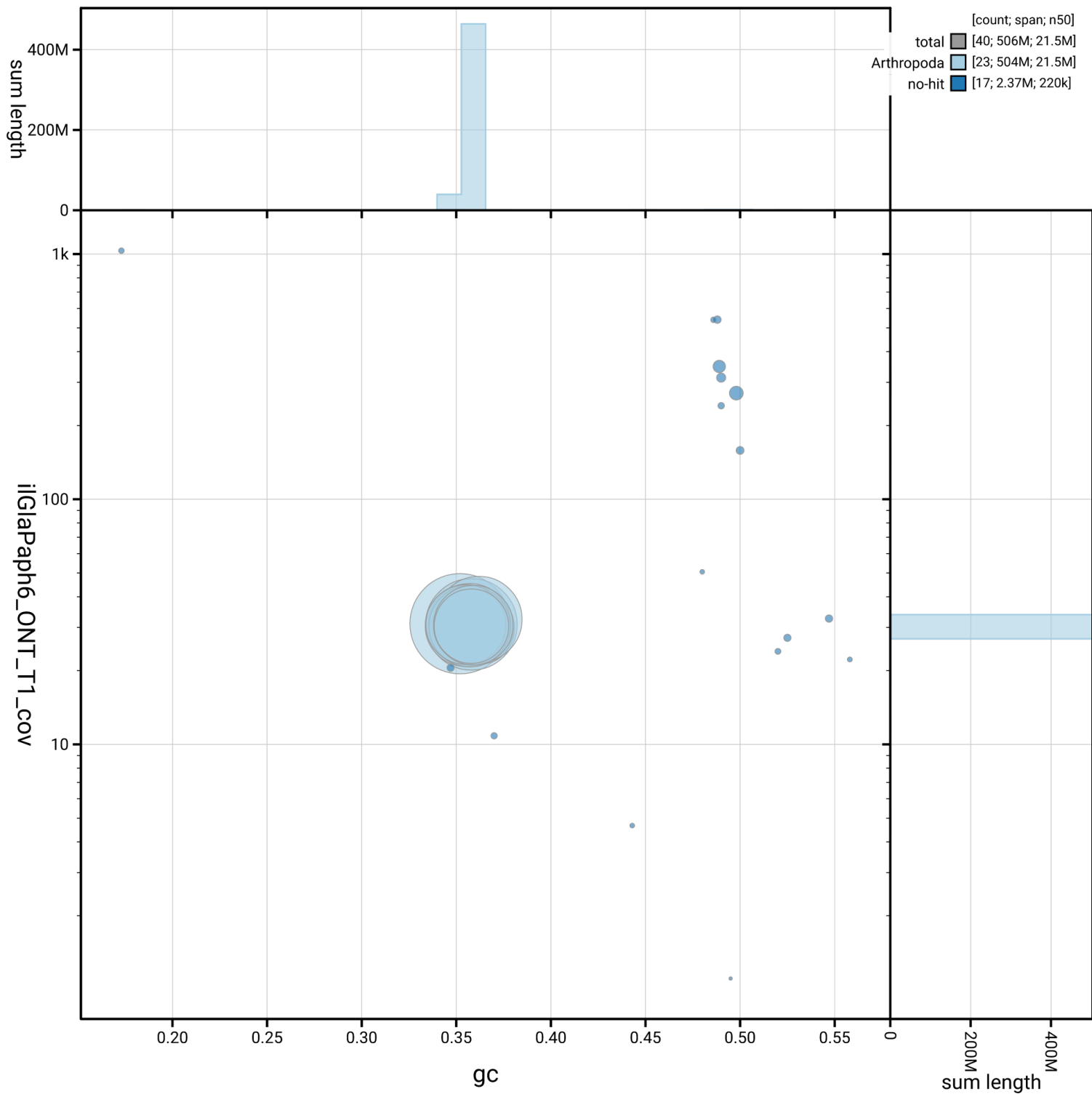

Distribution of k-mer counts per copy numbers found in **hap2** (hapl.)

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.

hap2. 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	OmniC	Illumina WGS
Coverage	72x	36x	39x

Assembly pipeline

- **Filtlong**
  - |_ *ver:* 0.2.1
  - |_ *key param:* --min_length 1000 --min_mean_q 99
- **Hifiasm**
  - |_ *ver:* 0.24.0
  - |_ *key param:* --ont --telo-m AACCT"
- **YaHS**
  - |_ *ver:* 1.2a
  - |_ *key param:* NA
- **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:* 1.0.3
  - |_ *key param:* NA

Submitter: Tyler Alioto
Affiliation: CNAG

Date and time: 2025-09-19 14:17:24 CEST