

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	2744210
ToLID	ilHipMadel
Species	Hipparchia maderensis
Class	Insecta
Order	Lepidoptera

Genome Traits	Expected	Observed
Haploid size (bp)	375,997,580	380,796,767
Haploid Number	29 (source: ancestor)	30
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	ZZ

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.7.Q67

Obtained EBP quality metric for hap2: 7.7.Q64

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

- . Observed Haploid Number is different from Expected
- . Observed sex is different from Sample sex
- . Kmer completeness value is less than 90 for hap1
- . Kmer completeness value is less than 90 for hap2

Curator notes

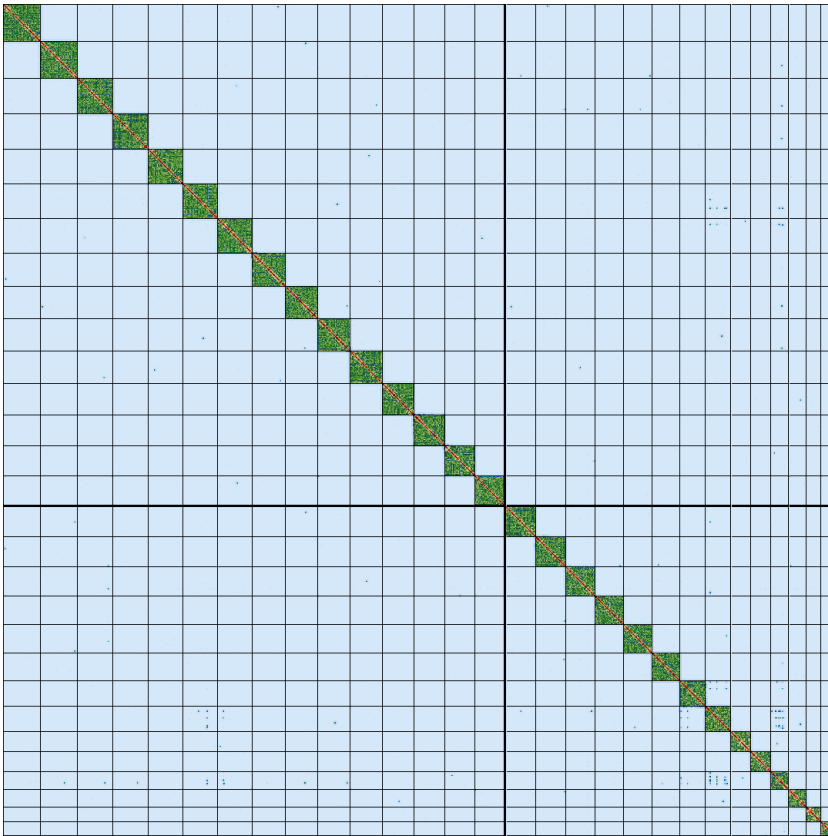
- . Interventions/Gb: 1
- . Contamination notes: "After running blobtoolkit in both haplotypes some pseudomonas contamination was found. One contaminant sequence was removed from hap1 and 71 sequences from hap2."
- . Other observations: "Both haplotypes were curated in this case because none of them required a lot of curation. The mitochondrion has been assembled using FOAM into a single circular contig of 15220 bp."

Quality metrics table

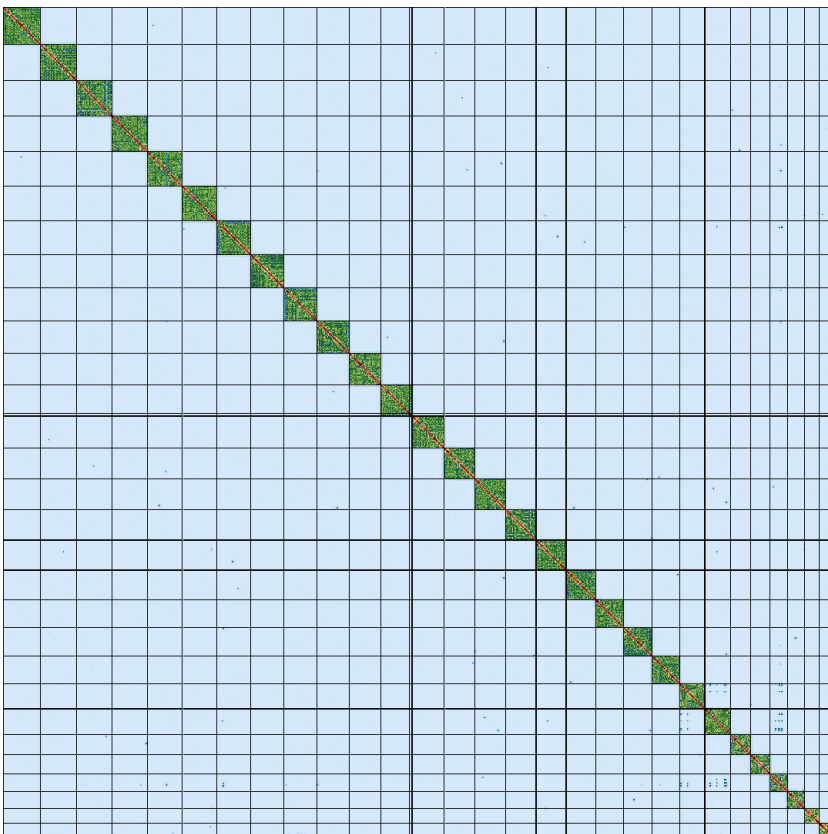
Metrics	Pre-curation hap1	Pre-curation hap2	Curated hap1	Curated hap2
Total bp	379,018,945	385,662,067	378,174,314	380,796,767
GC %	37.23	37.41	37.18	37.21
Gaps/Gbp	2.64	5.19	2.64	5.25
Total gap bp	200	400	200	400
Scaffolds	40	130	39	59
Scaffold N50	14,000,976	14,138,762	14,000,976	14,138,762
Scaffold L50	13	13	13	13
Scaffold L90	25	25	25	25
Contigs	41	132	40	61
Contig N50	14,000,976	14,059,432	14,000,976	14,059,432
Contig L50	13	13	13	13
Contig L90	25	26	25	26
QV	67.5704	62.1845	67.7944	64.2122
Kmer compl.	79.5217	79.5762	79.5202	79.5706
BUSCO sing.	98.4%	98.1%	98.3%	98.1%
BUSCO dupl.	0.3%	0.5%	0.3%	0.5%
BUSCO frag.	0.2%	0.3%	0.3%	0.3%
BUSCO miss.	1.1%	1.1%	1.1%	1.1%

BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: lepidoptera_odb10 (genomes:16, BUSCOs:5286)

HiC contact map of curated assembly

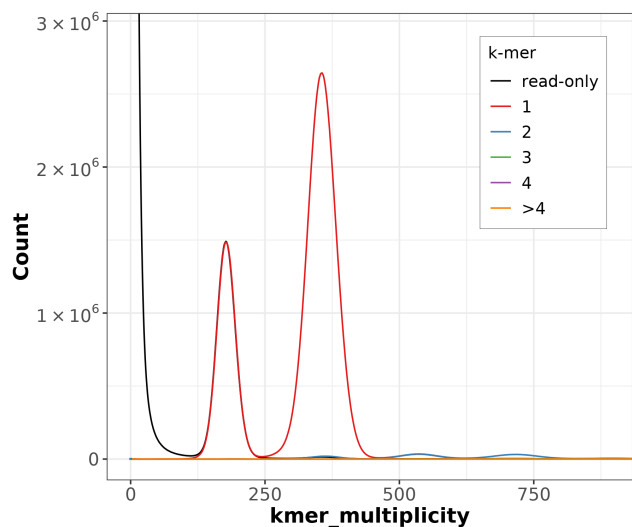


hap1 [\[LINK\]](#)

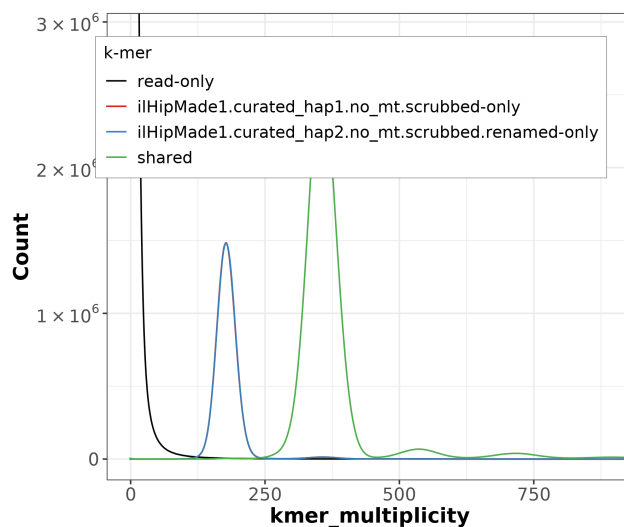


hap2 [\[LINK\]](#)

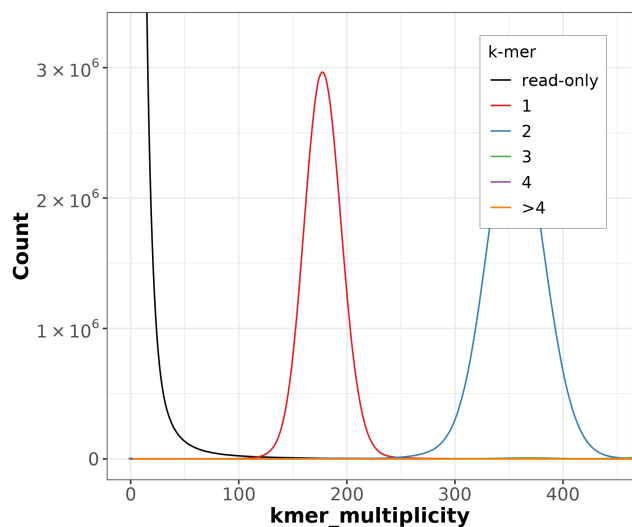
K-mer spectra of curated assembly



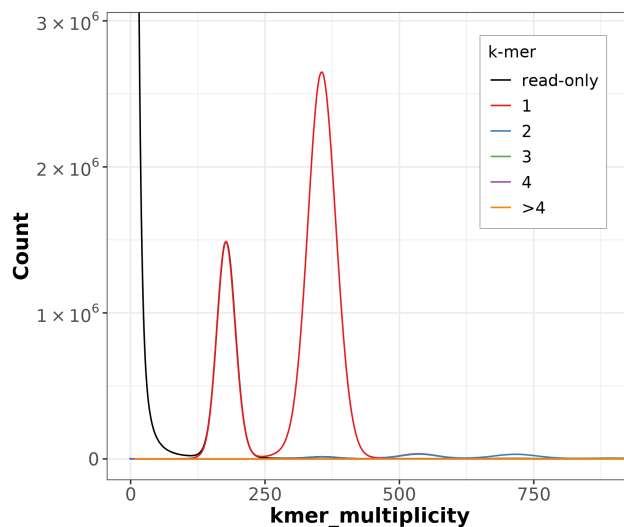
Distribution of k-mer counts per copy numbers found in **curated_hap2 renamed** (hap1.)



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

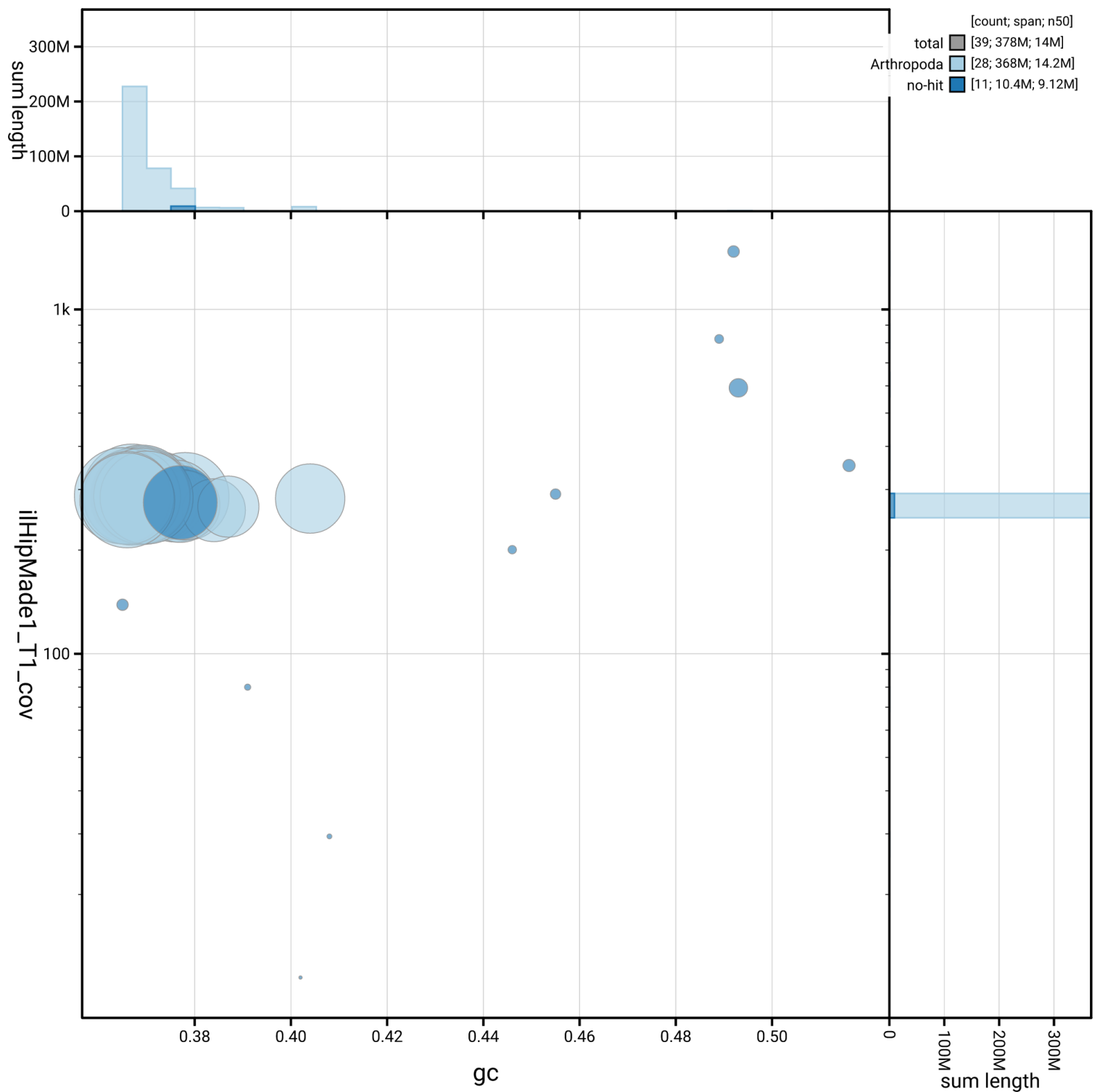


Distribution of k-mer counts per copy numbers found in **asm** (dipl.)

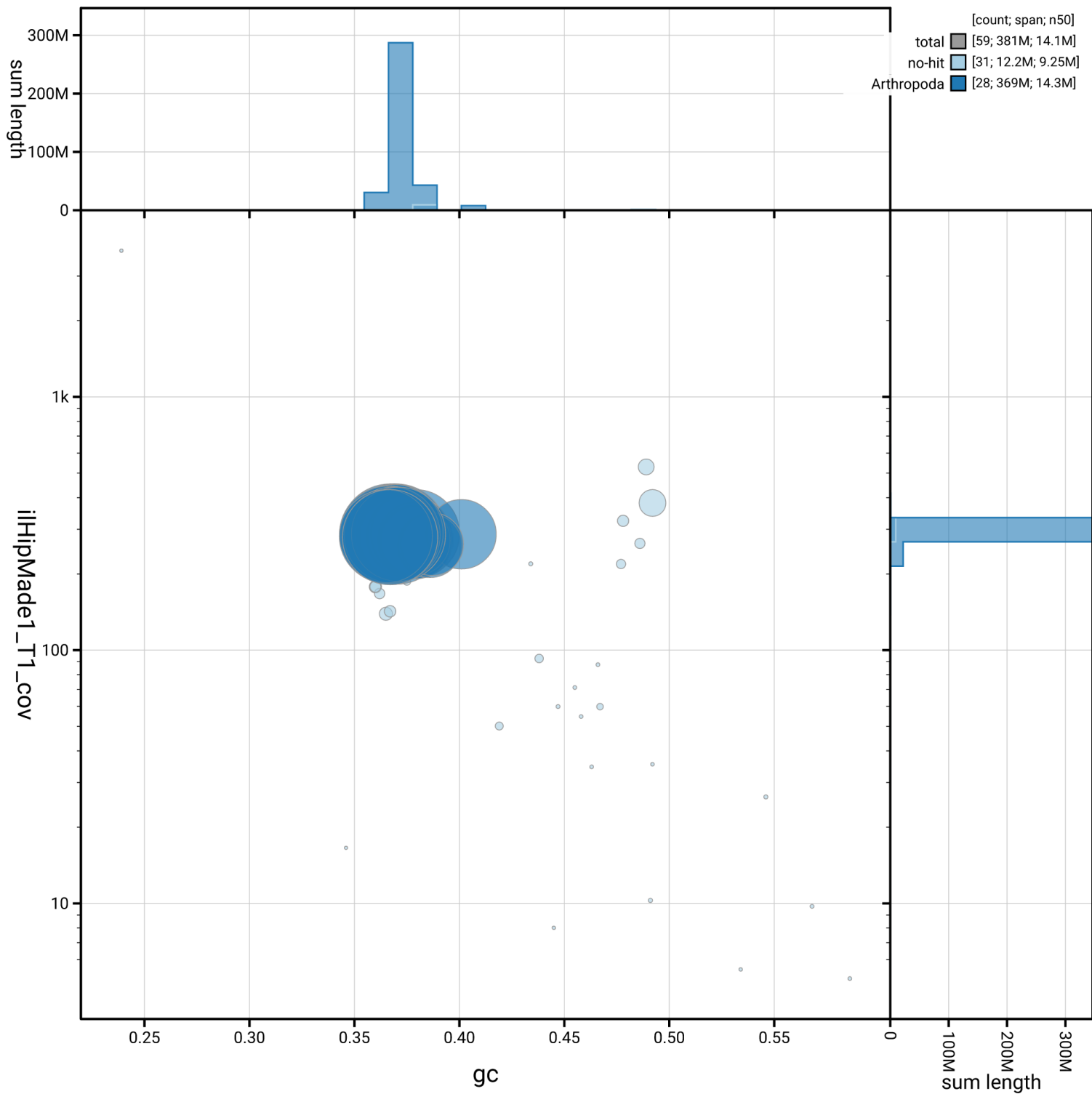


Distribution of k-mer counts per copy numbers found in **curated_hap1** (hap1.)

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	Illumina	OmniC
Coverage	274x	144x	101x

Assembly pipeline

- **CLAWS**
 - |_ *ver*: v2.3.0
 - |_ *key param*: NA
- **Filtlong**
 - |_ *ver*: v0.2.1
 - |_ *key param*: -t target_bases 300000000000
- **Hifiasm**
 - |_ *ver*: 0.24.0
 - |_ *key param*: --ont
 - |_ *key param*: telo-m TAACC
- **Yahs**
 - |_ *ver*: 1.2a
 - |_ *key param*: -mq 10
 - |_ *key param*: -no-contig-ec

Curation pipeline

- **Blobtoolkit Nextflow pipeline**
 - |_ *ver*: 0.6
 - |_ *key param*: NA
- **PretextViewAI**
 - |_ *ver*: 1.0.3
 - |_ *key param*: NA

Submitter: Jessica Gomez-Garrido

Affiliation: CNAG

Date and time: 2025-06-18 15:52:25 CEST