

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	233204
ToLID	<b>drUlmLaev1</b>
Species	<i>Ulmus laevis</i>
Class	Magnoliopsida
Order	Rosales

Genome Traits	Expected	Observed
Haploid size (bp)	1,414,578,037	3,502,430,588
Haploid Number	14 (source: direct)	14
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	NA

## EBP metrics summary and curation notes

Obtained EBP quality metric for Hapl: 6.8.Q63

Obtained EBP quality metric for Hap2: 6.8.Q66

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

- . Observed Haploid size (bp) has >20% difference with Expected
- . Kmer completeness value is less than 90 for Hapl
- . Kmer completeness value is less than 90 for Hap2
- . Not 90% of assembly in chromosomes for Hapl
- . Not 90% of assembly in chromosomes for Hap2

## Curator notes

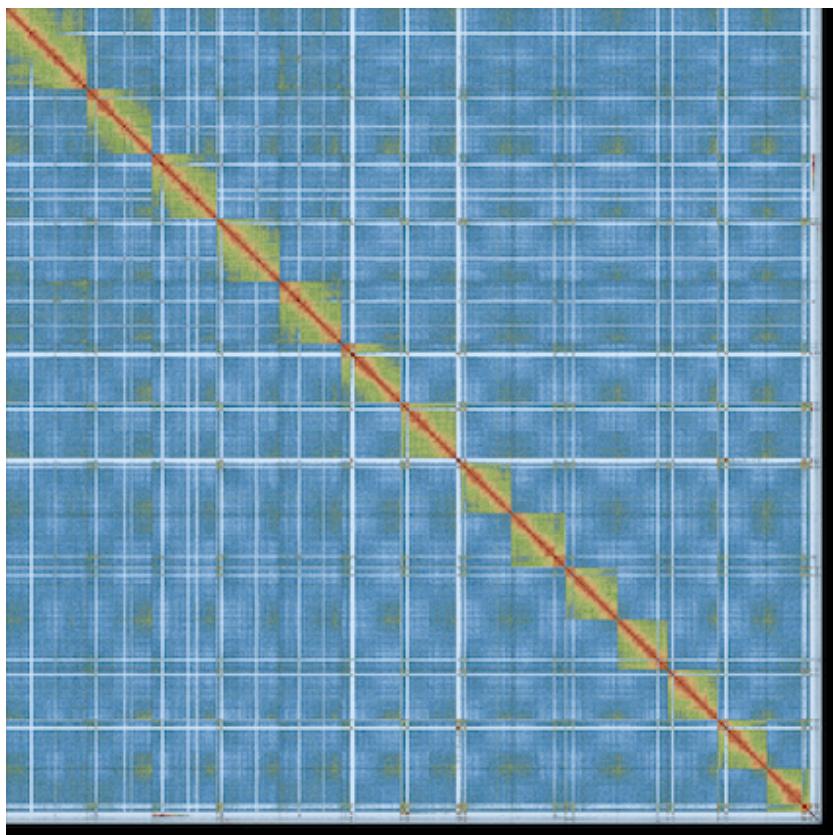
- . Interventions/Gb: 42
- . Contamination notes: "FCS-GX flagged bacteria (*Walczuchella*) and fungi (*Aureobasidium*). Blobtools flagged many possible contaminants (as show in the plot), but only a few were confirmed."
- . Other observations: "mitochondria and chloroplast were assembled but not included in the current assemblies. Compleasm was used instead of BUSCO."

# Quality metrics table

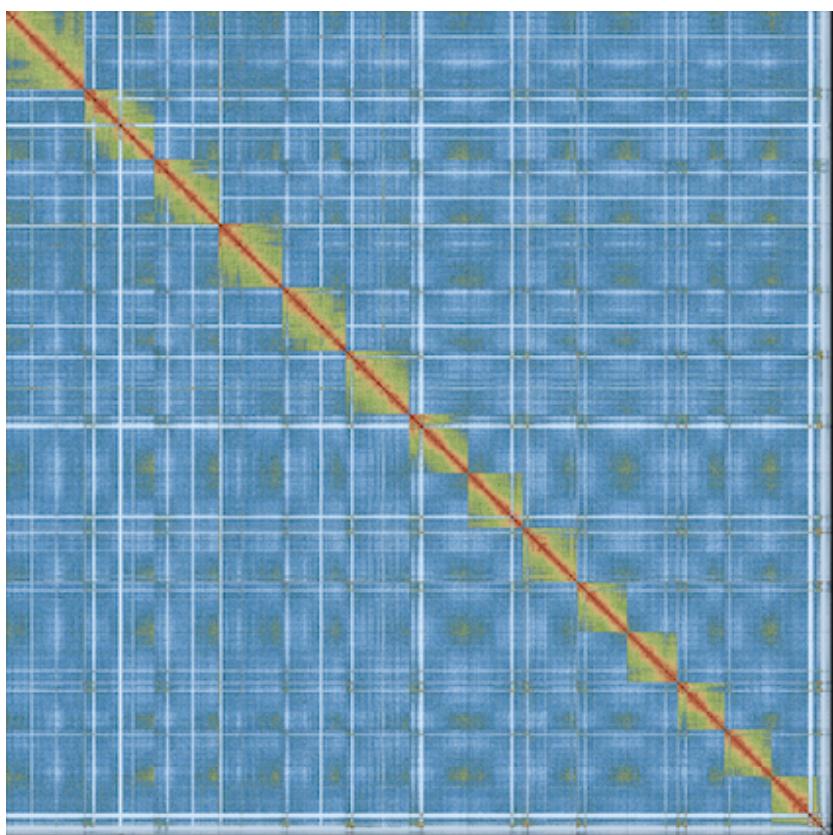
Metrics	Pre-curation Hap1	Pre-curation Hap2	Curated Hap1	Curated Hap2
Total bp	3,453,219,982	3,561,778,274	3,488,832,032	3,502,430,588
GC %	35.13	35.25	35.13	35.28
Gaps/Gbp	187.07	167.33	183.44	178.16
Total gap bp	129,200	119,200	128,000	124,800
Scaffolds	1,468	372	1,404	350
Scaffold N50	121,859,695	127,974,074	116,772,628	124,035,533
Scaffold L50	12	11	13	13
Scaffold L90	24	24	26	26
Contigs	2,760	1,564	2,684	1,598
Contig N50	4,933,302	4,607,366	5,033,854	4,497,321
Contig L50	215	228	209	234
Contig L90	740	746	721	758
QV	62.9663	66.7338	63.0292	66.7262
Kmer compl.	56.3751	58.645	57.008	58.0438
BUSCO sing.	91.86%	91.69%	94.54%	94.51%
BUSCO dupl.	2.80%	4.42%	2.11%	2.28%
BUSCO frag.	0.32%	0.28%	0.32%	0.28%
BUSCO miss.	4.96%	3.54%	2.97%	2.85%

BUSCO: 0.2.7 (genome, miniprot) / Lineage: rosales\_odb12 (genomes:0, BUSCOs:9329)

# HiC contact map of curated assembly

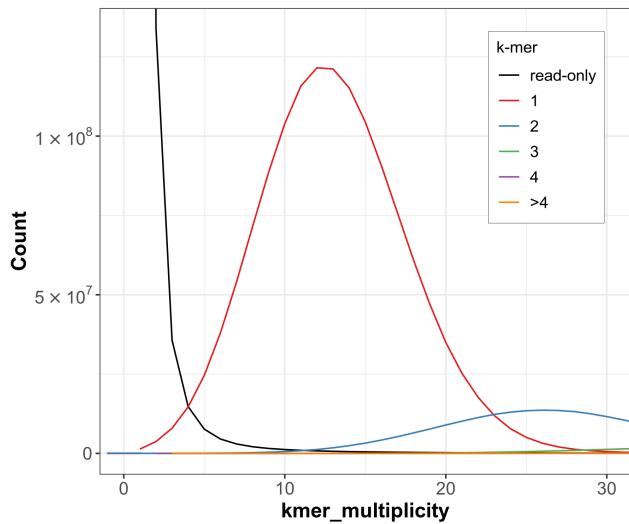


Hap1 [\[LINK\]](#)

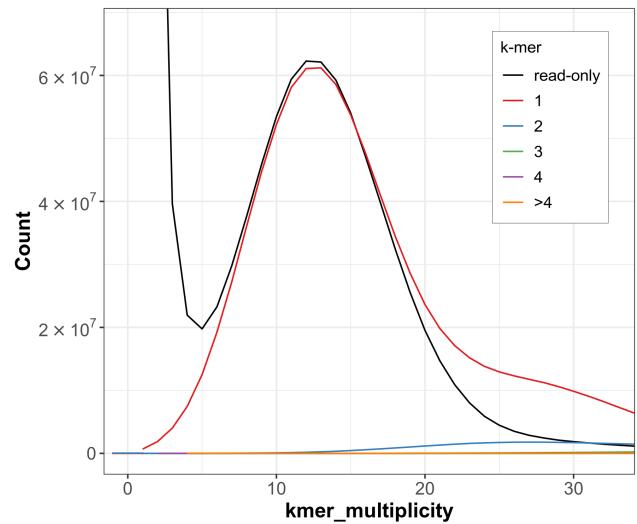


Hap2 [\[LINK\]](#)

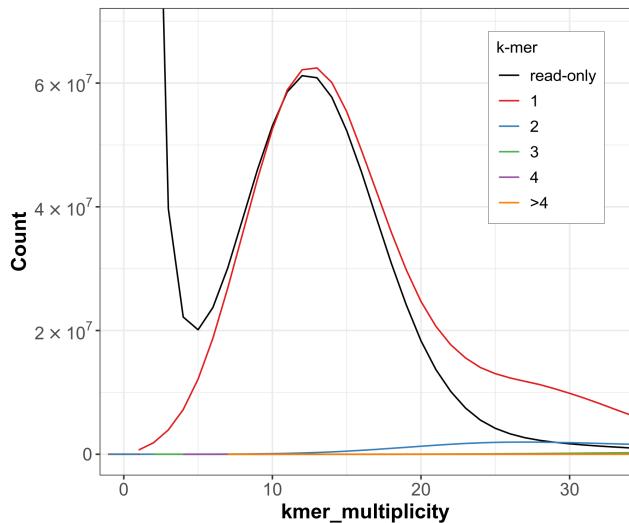
# K-mer spectra of curated assembly



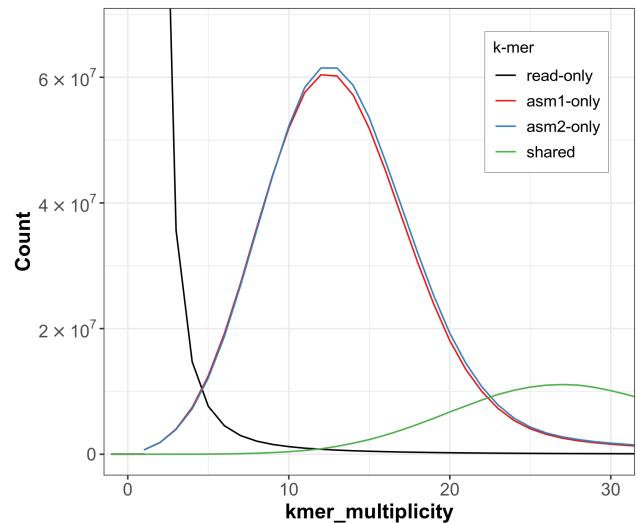
Distribution of k-mer counts per copy numbers found in **asm** (diploid)



Distribution of k-mer counts per copy numbers found in **asm1** (haploid)

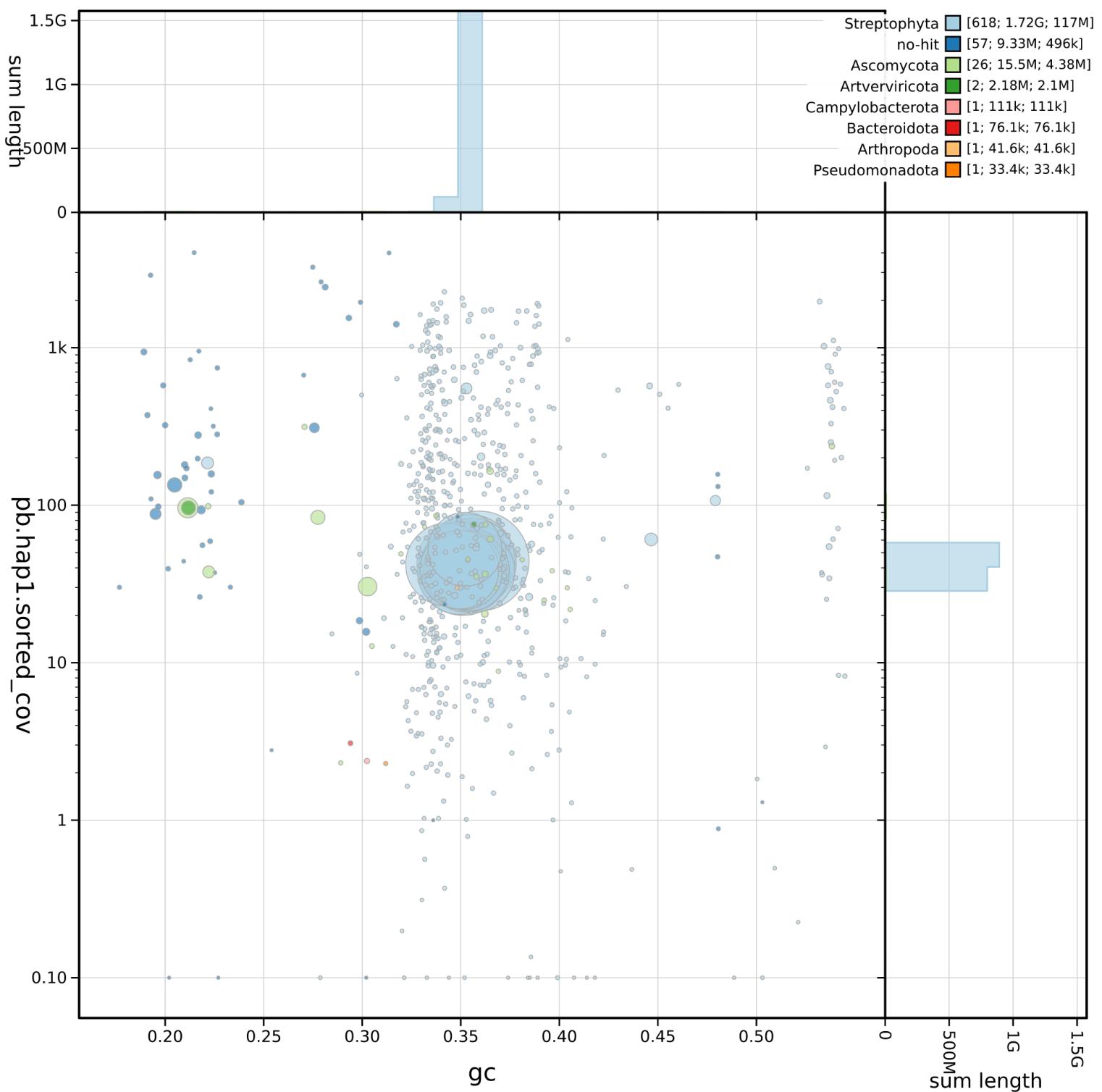


Distribution of k-mer counts per copy numbers found in **asm2** (haploid)

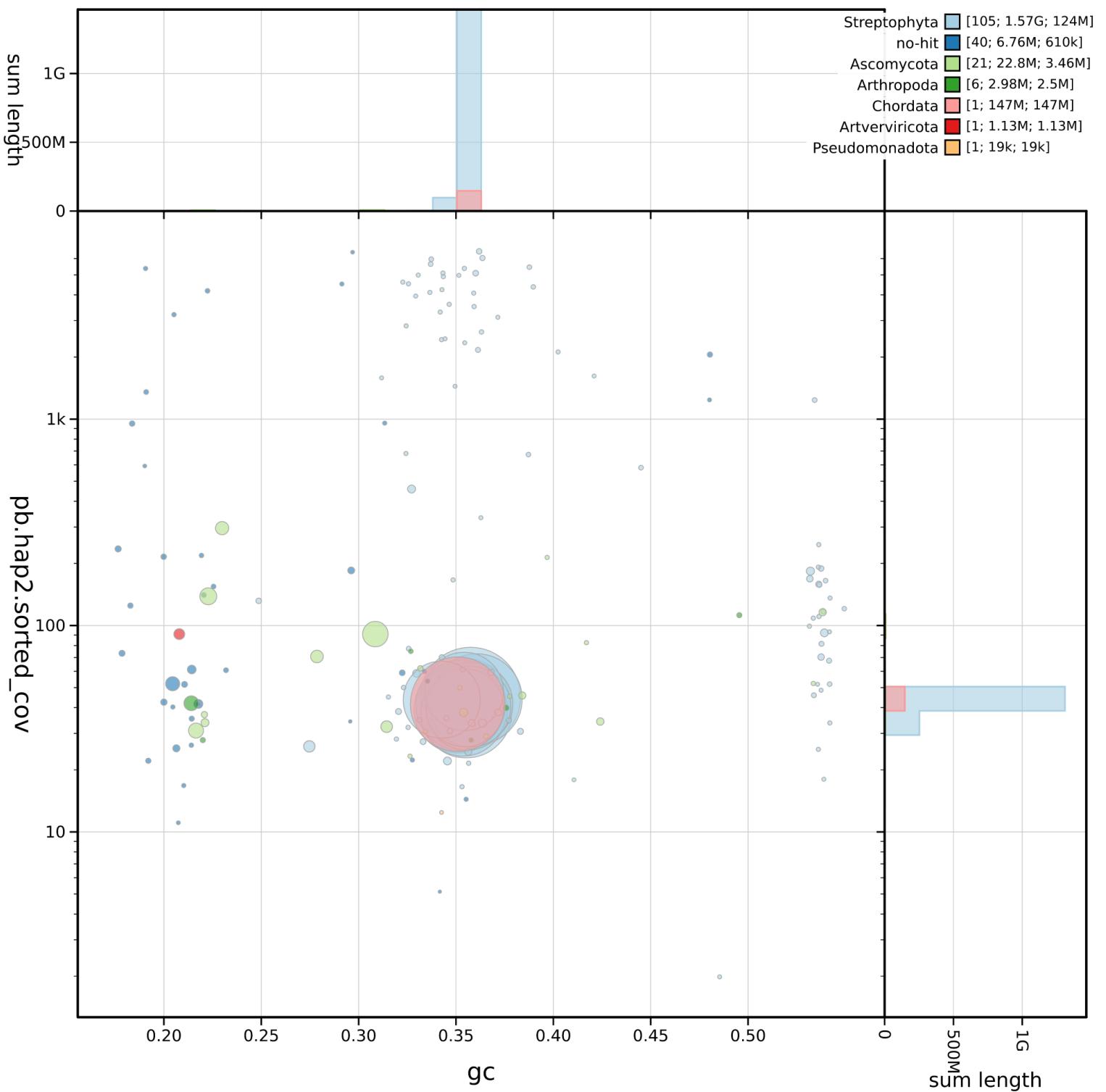


Distribution of k-mer counts coloured by their presence in reads/assemblies

# 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.



**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	HiFi	Hi-C
Coverage	14	54

## Assembly pipeline

```
- hifiasm
  |_ ver: 0.25.0-r726
  |_ key param: --hom-cov 29
  |_ key param: -l 3
  |_ key param: --purge-max 68
  |_ key param: --telo-m TTTAGGG
  |_ key param: --h1 --h2
- yahs
  |_ ver: 1.2.2
  |_ key param: NA
```

## Curation pipeline

```
- FCS-GX
  |_ ver: 0.5.5
  |_ key param: NA
- Blobtoolkit
  |_ ver: 4.5.0
  |_ key param: NA
- Pretext
  |_ ver: 1.0.5
  |_ key param: NA
```

Submitter: Diego De Panis

Affiliation: IZW

Date and time: 2026-02-26 13:58:04 CET