

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

| | |
|---------|-------------------------|
| TxID | 40178 |
| ToLID | bAleGra1 |
| Species | <i>Alectoris graeca</i> |
| Class | Aves |
| Order | Galliformes |

| Genome Traits | Expected | Observed |
|-------------------|----------------------|---------------|
| Haploid size (bp) | 1,015,215,291 | 1,120,883,897 |
| Haploid Number | 39 (source: direct) | 40 |
| Ploidy | 2 (source: ancestor) | 2 |
| Sample Sex | ZZ | ZZ |

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.7.Q65

Obtained EBP quality metric for hap2: 7.7.Q66

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

- . Observed Haploid Number is different from Expected

Curator notes

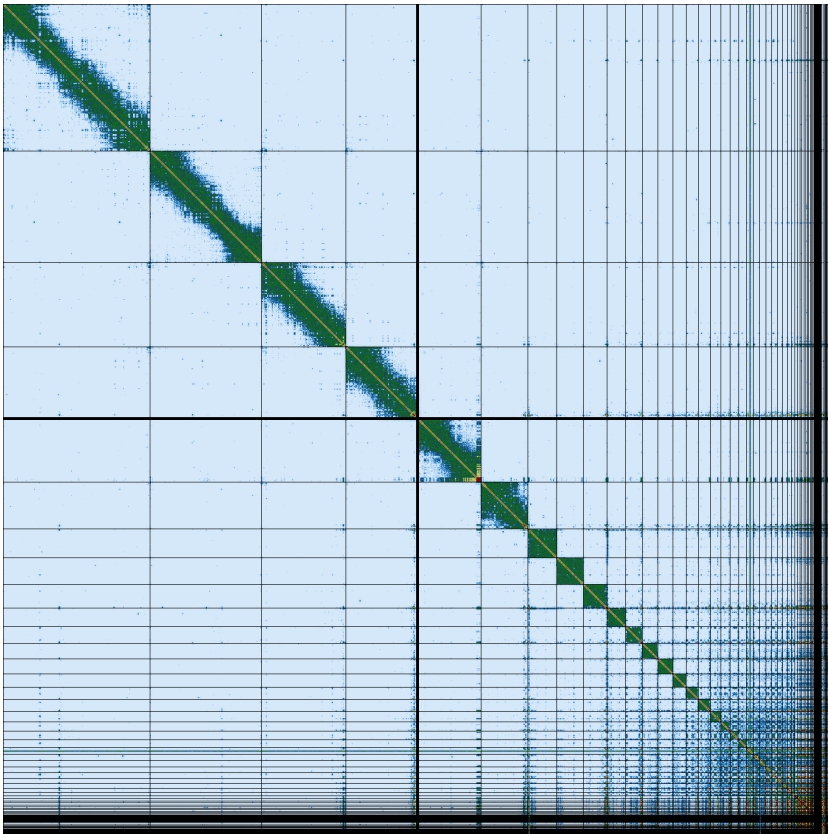
- . Interventions/Gb: None
- . Contamination notes: "No presence of contaminants. Mitochondrial genome was removed from the assembly"
- . Other observations: "Z chromomomes are chromosomes 5 in both haplotypes, curation was done on combined haplotypes, large repeat region in chr5 (77M-84M) in both haplotypes - contig placement are ambiguous; there should be 8 more microchromosomes; purging was manually done to avoid removing of large fractions of microchromosomes, guided by purge_dups dups.bed file"

Quality metrics table

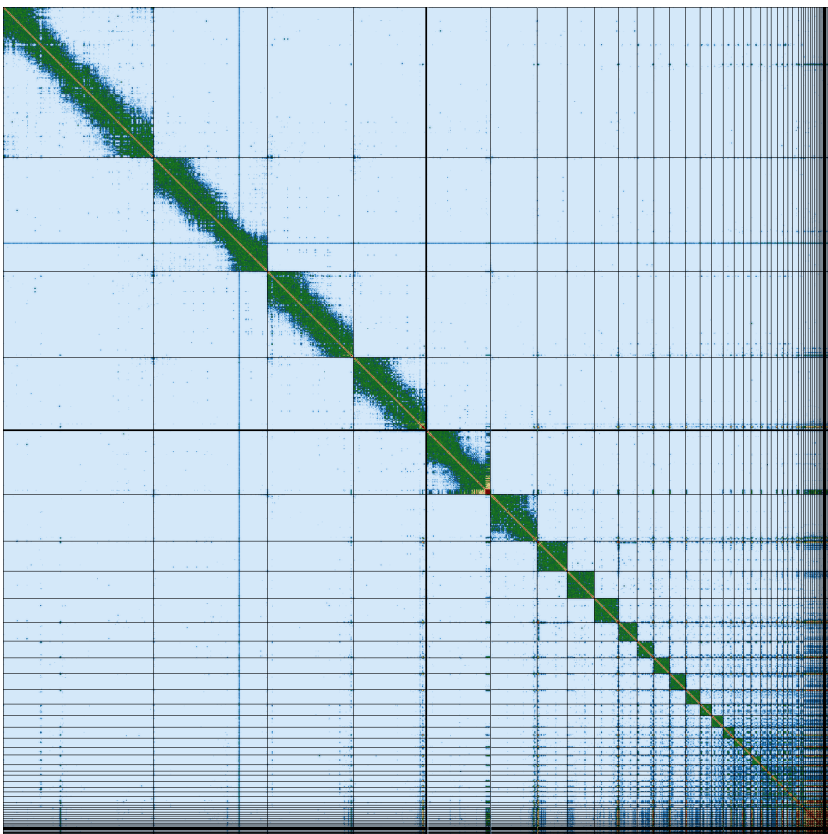
| Metrics | Pre-curation hap1 | Pre-curation hap2 | Curated hap1 | Curated hap2 |
|--------------|-------------------|-------------------|---------------|---------------|
| Total bp | 1,120,904,364 | 1,103,332,262 | 1,120,883,897 | 1,097,305,545 |
| GC % | 42.08 | 41.94 | 42.07 | 41.93 |
| Gaps/Gbp | 157.02 | 125.08 | 177.54 | 151.28 |
| Total gap bp | 17,600 | 13,800 | 22,700 | 19,600 |
| Scaffolds | 324 | 174 | 288 | 125 |
| Scaffold N50 | 78,812,343 | 95,332,562 | 84,007,930 | 95,332,562 |
| Scaffold L50 | 5 | 4 | 5 | 4 |
| Scaffold L90 | 23 | 22 | 22 | 20 |
| Contigs | 500 | 312 | 487 | 291 |
| Contig N50 | 22,503,344 | 24,970,722 | 22,503,344 | 24,970,722 |
| Contig L50 | 11 | 13 | 11 | 13 |
| Contig L90 | 71 | 68 | 72 | 67 |
| QV | 65.8427 | 66.8877 | 65.8647 | 66.9485 |
| Kmer compl. | 97.8383 | 97.6985 | 97.837 | 97.6795 |
| BUSCO sing. | 98.9% | 98.7% | 99.0% | 98.7% |
| BUSCO dupl. | 0.1% | 0.1% | 0.1% | 0.1% |
| BUSCO frag. | 0.1% | 0.1% | 0.1% | 0.1% |
| BUSCO miss. | 0.8% | 1.1% | 0.8% | 1.1% |

BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: aves_odb10 (genomes:62, BUSCOs:8338)

HiC contact map of curated assembly

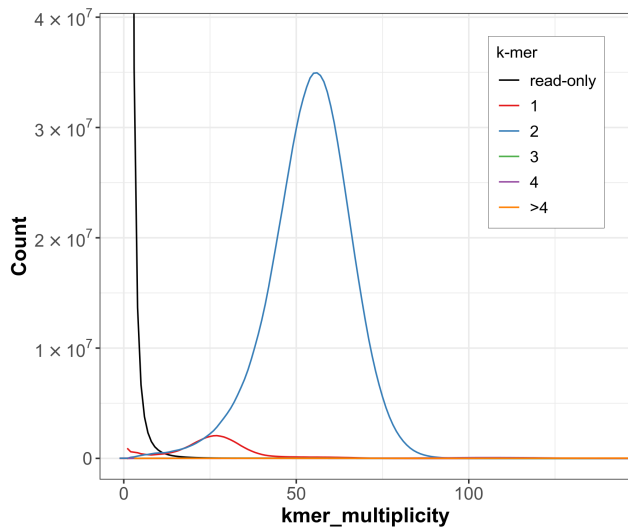


hap1 [\[LINK\]](#)

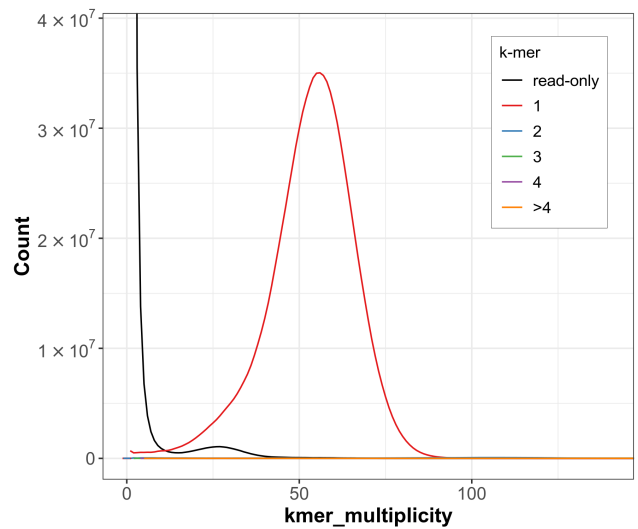


hap2 [\[LINK\]](#)

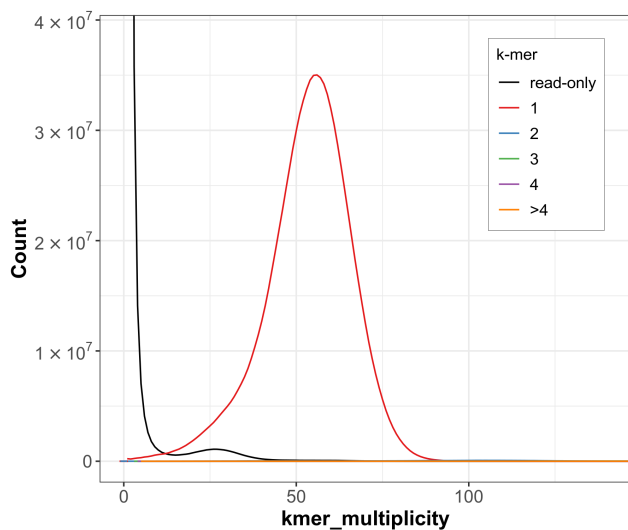
K-mer spectra of curated assembly



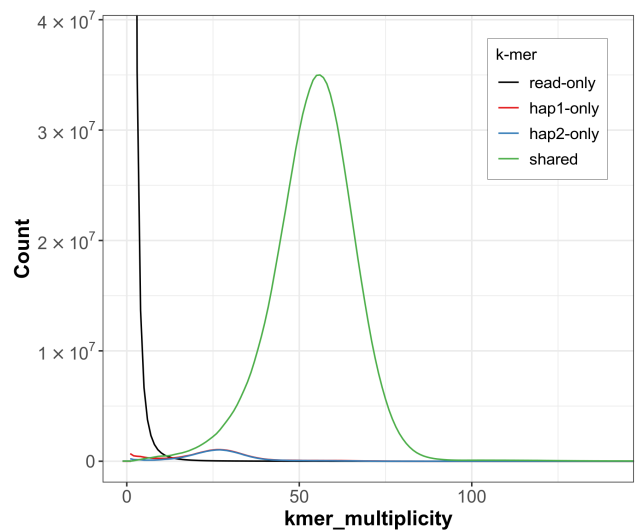
Distribution of k-mer counts per copy numbers found in asm (dipl.)



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

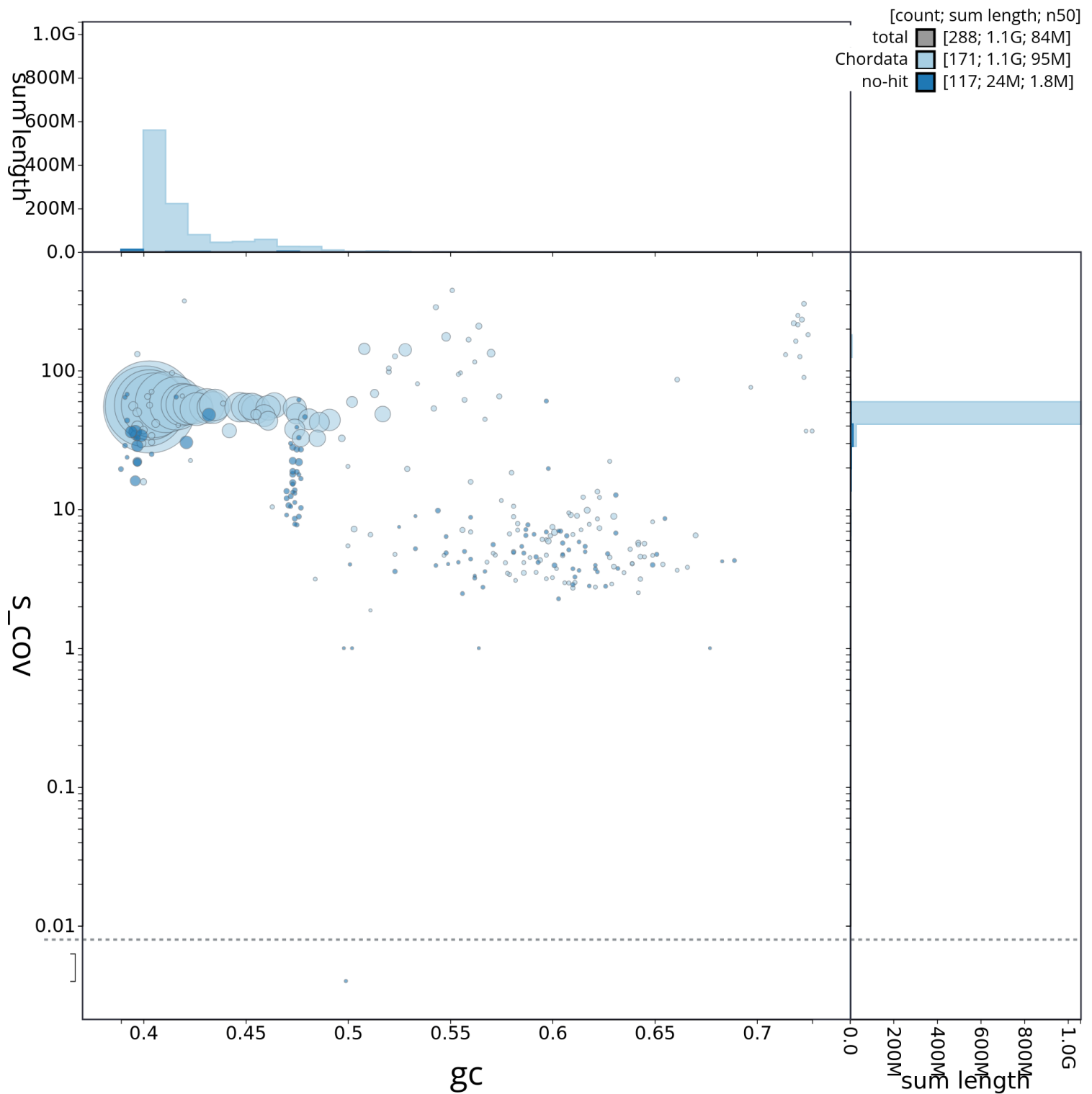


Distribution of k-mer counts per copy numbers found in **hap2** (hapl.)

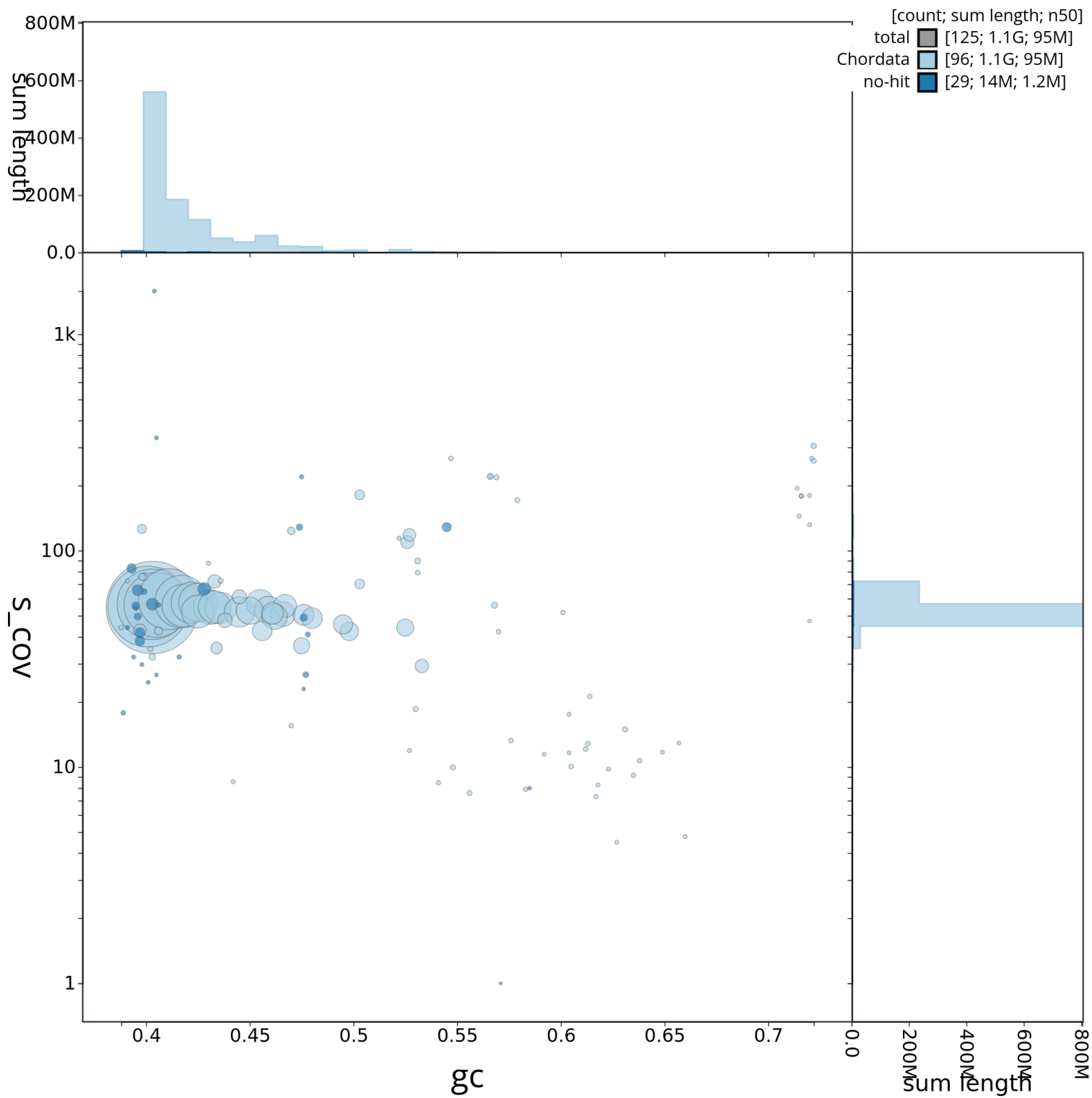


Distribution of k-mer counts coloured by their presence in reads/assemblies

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 | Bionano | OmniC |
|----------|------|---------|-------|
| Coverage | 53x | NA | 143x |

Assembly pipeline

- **Hifiasm**
 - |_ *ver*: 0.24.0-r703
 - |_ *key param*: HiC
 - |_ *key param*: 13
- **purge_dups**
 - |_ *ver*: 1.2.5
 - |_ *key param*: NA
- **YaHS**
 - |_ *ver*: 1.2.2
 - |_ *key param*: NA

Curation pipeline

- **GRIT_Rapid**
 - |_ *ver*: 1a3d79a8
 - |_ *key param*: NA
- **HiGlass**
 - |_ *ver*: 0.10.4
 - |_ *key param*: NA

Submitter: Martin Pippel

Affiliation: SciLifeLab

Date and time: 2025-03-03 13:35:16 CET