

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

|         |                              |
|---------|------------------------------|
| TxID    | 457762                       |
| ToLID   | <b>ddAlyUtril</b>            |
| Species | <i>Alyssoides utriculata</i> |
| Class   | Magnoliopsida                |
| Order   | Brassicales                  |

| Genome Traits     | Expected             | Observed      |
|-------------------|----------------------|---------------|
| Haploid size (bp) | 1,071,941,611        | 1,018,829,908 |
| Haploid Number    | 4 (source: direct)   | 8             |
| Ploidy            | 2 (source: ancestor) | 2             |
| Sample Sex        | NA                   | NA            |

## EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 8.8.Q64

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

- . Observed Haploid Number is different from Expected

## Curator notes

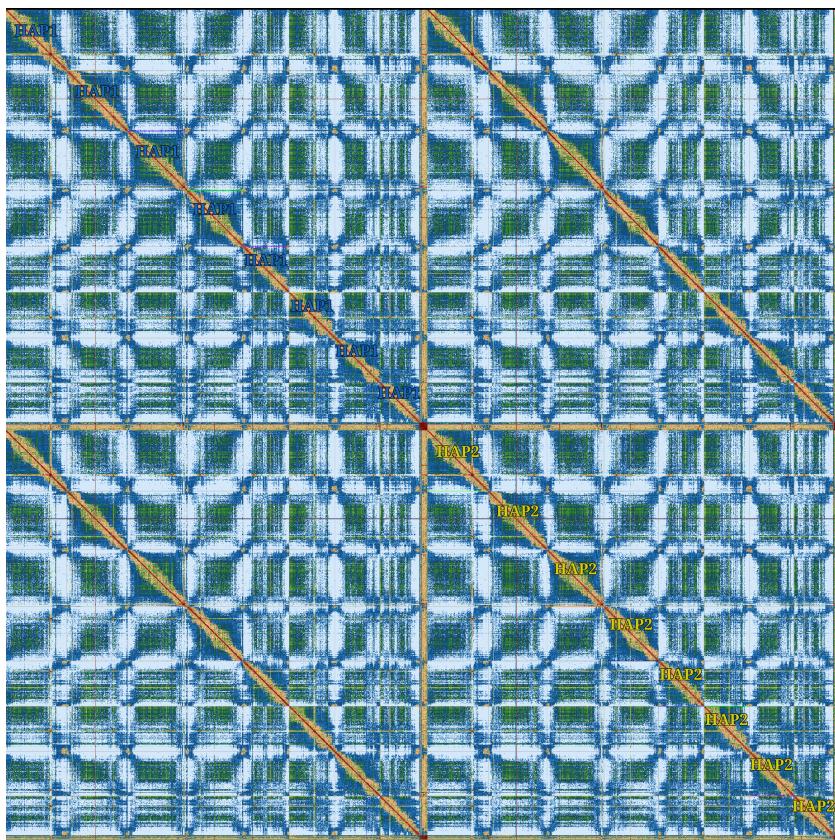
- . Interventions/Gb: 6
- . Contamination notes: "2 mycoplasma sequences removed"
- . Other observations: "Not many interventions to make after yahs. Need to reorder the chromosomes based on either size or synteny, not sure yet"

## Quality metrics table

| Metrics      | Pre-curation<br>hap1 | Curated<br>hap1 |
|--------------|----------------------|-----------------|
| Total bp     | 1,028,428,052        | 1,018,829,908   |
| GC %         | 38.96                | 38.95           |
| Gaps/Gbp     | 83.62                | 83.43           |
| Total gap bp | 17,200               | 17,000          |
| Scaffolds    | 726                  | 413             |
| Scaffold N50 | 685,186,311          | 135,958,589     |
| Scaffold L50 | 1                    | 4               |
| Scaffold L90 | 4                    | 8               |
| Contigs      | 812                  | 498             |
| Contig N50   | 108,093,000          | 108,093,000     |
| Contig L50   | 5                    | 5               |
| Contig L90   | 10                   | 10              |
| QV           | 62.527               | 64.0782         |
| Kmer compl.  | 98.0999              | 98.0773         |
| BUSCO sing.  | 92.6%                | 92.5%           |
| BUSCO dupl.  | 4.1%                 | 4.0%            |
| BUSCO frag.  | 1.5%                 | 1.5%            |
| BUSCO miss.  | 1.8%                 | 1.9%            |

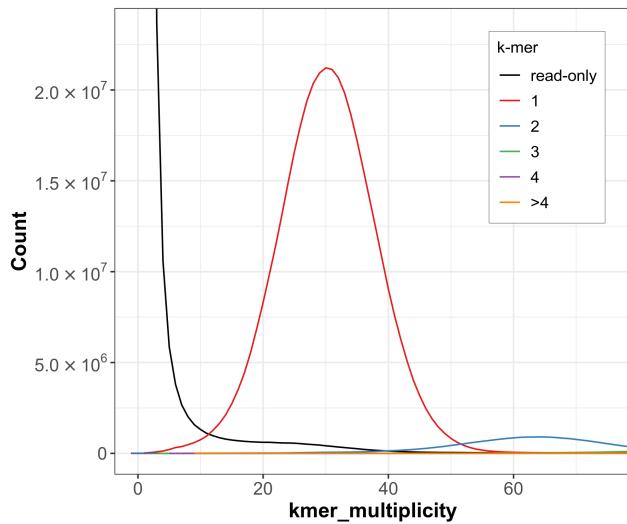
BUSCO: 5.8.2 (euk\_genome\_min, miniprot) / Lineage: eudicotyledons\_odb12 (genomes:76, BUSCOs:2805)

# HiC contact map of curated assembly

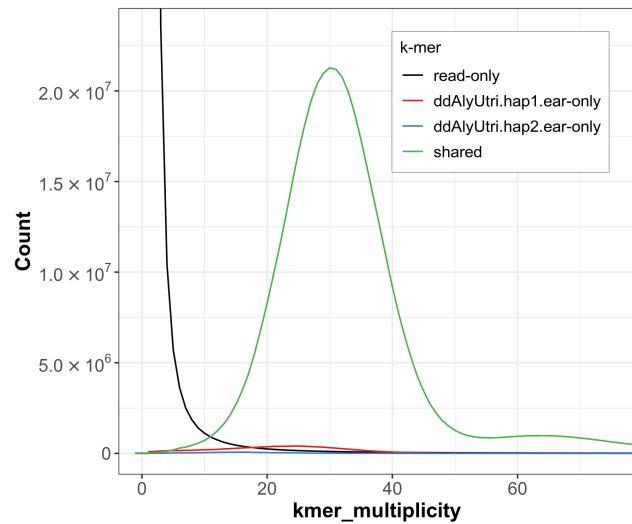


hap1 [LINK]

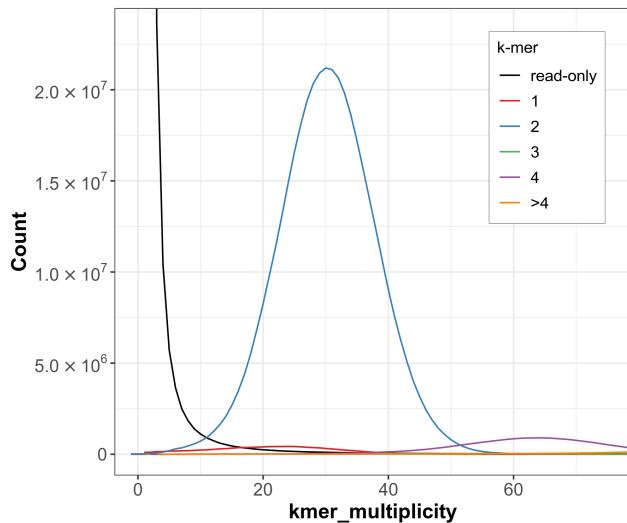
# K-mer spectra of curated assembly



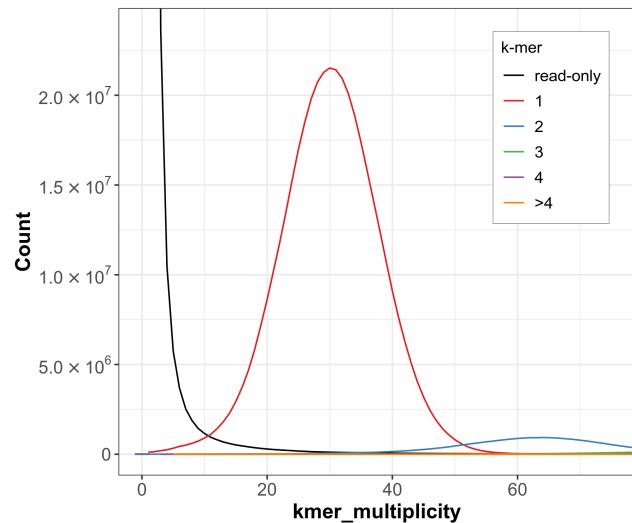
Distribution of k-mer counts per copy numbers found in **hap2** (hapl.)



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

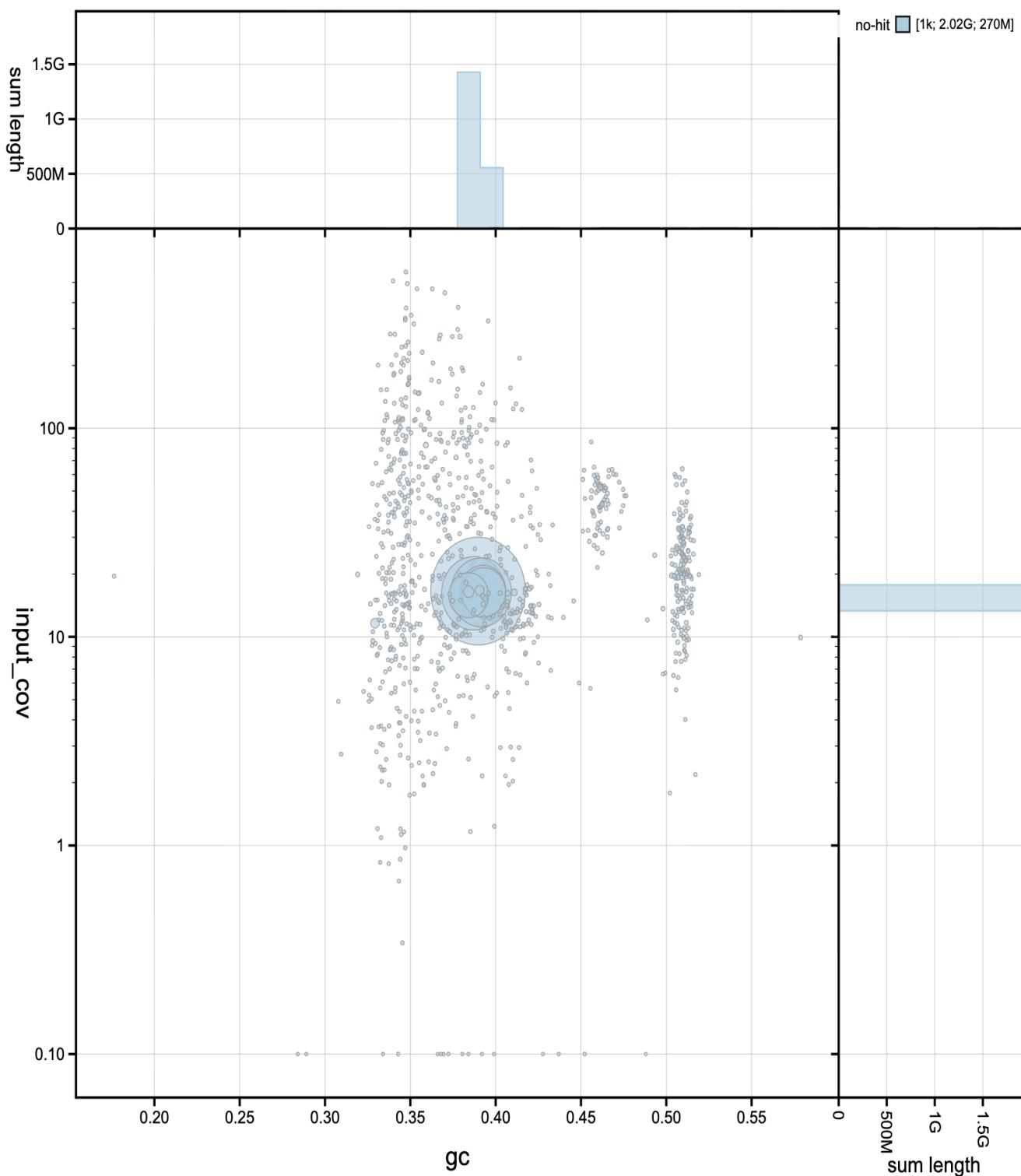


Distribution of k-mer counts per copy numbers found in **asm** (diplo.)



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

## Data profile

|          | Data | HiFi | HiC |
|----------|------|------|-----|
| Coverage | 32   | 130  |     |

## Assembly pipeline

- **HiFiasm**
  - |\_ ver: 0.25.0
  - |\_ key param: --hom-cov 32
- **purge-dups**
  - |\_ ver: 1.2.6
  - |\_ key param: NA
- **yahs**
  - |\_ ver: 1.1a
  - |\_ key param: NA

## Curation pipeline

- **PretextViewAI**
  - |\_ ver: 1.0.0
  - |\_ key param: NA

Submitter: Tom Brown

Affiliation: IZW

Date and time: 2025-10-10 10:52:27 CEST