

# ERGA Assembly Report

v24.04.03\_beta

Tags: ERGA-BGE

TxID	498071
ToLID	<b>drZizLotu1</b>
Species	Ziziphus lotus
Class	Magnoliopsida
Order	Rosales

Genome Traits	Expected	Observed
Haploid size (bp)	307,042,740	293,294,843
Haploid Number	24 (source: direct)	12
Ploidy	2 (source: ancestor)	2
Sample Sex	H	H

## EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 7.7.Q62

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

- . Observed Haploid Number is different from Expected
- . Assembly length loss > 3% for pri

### Curator notes

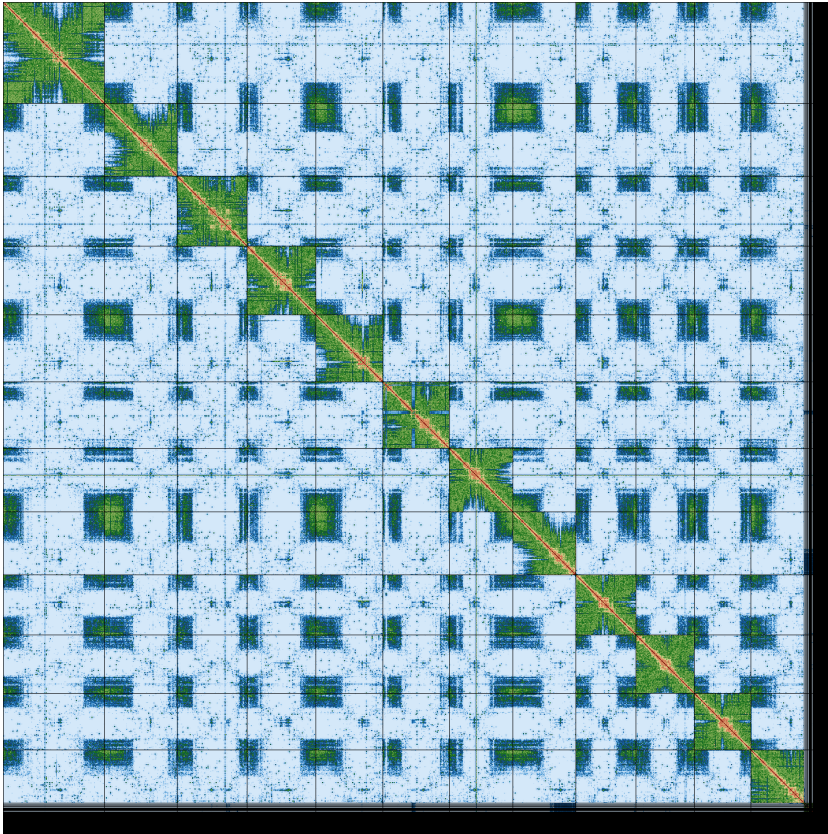
- . Interventions/Gb: 22
- . Contamination notes: "Contamination report for assembly labelled hap1; Total length of scaffolds removed: 200,777,297 (40.4 %); Scaffolds removed: 3555 (95.2 %); Largest scaffold removed: (341,751); FCS-GX contaminant species (number of scaffolds; total length of scaffolds):; Brevundimonas sp. UBA6550, a-proteobacteria (4; 167,575); Delftia acidovorans, b-proteobacteria (1; 8,026); Homo sapiens, primates (1; 7,632); Mitochondrion (1621; 85,342,075); Plastid (1928; 115,251,989)"
- . Other observations: "Assembly was Hi-C phased; "

# Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	496,389,500	293,294,843
GC %	36.18	33.33
Gaps/Gbp	16.12	57.96
Total gap bp	800	2,900
Scaffolds	3,735	143
Scaffold N50	20,141,524	23,263,395
Scaffold L50	10	6
Scaffold L90	2,349	12
Contigs	3,743	160
Contig N50	12,725,994	20,141,524
Contig L50	12	6
Contig L90	2,357	15
QV	51.4	62.6
Kmer compl.	99.47	99.24
BUSCO sing.	97.8%	98.3%
BUSCO dupl.	1.4%	0.8%
BUSCO frag.	0.2%	0.2%
BUSCO miss.	0.6%	0.7%

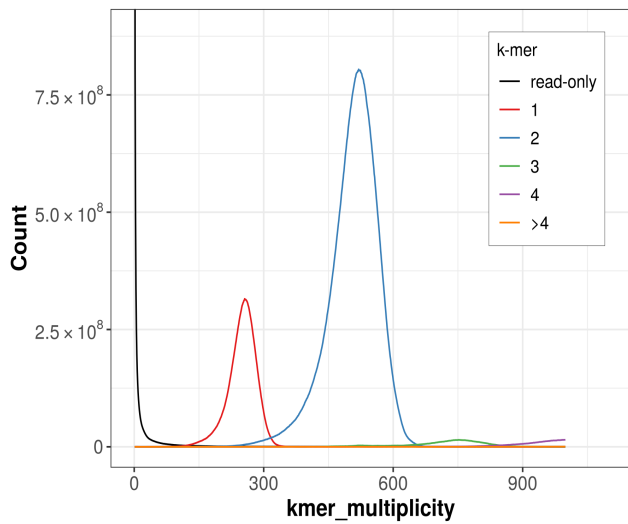
BUSCO 6.0.0 Lineage: eudicots\_odb10 (genomes:31, BUSCOs:2326)

# HiC contact map of curated assembly

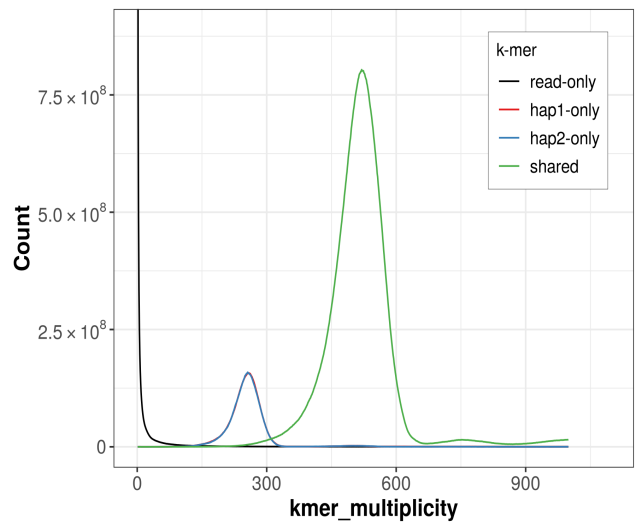


pri [\[LINK\]](#)

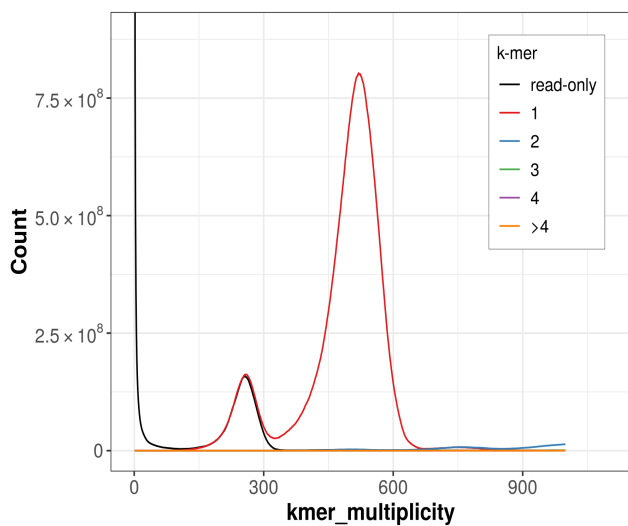
# K-mer spectra of curated assembly



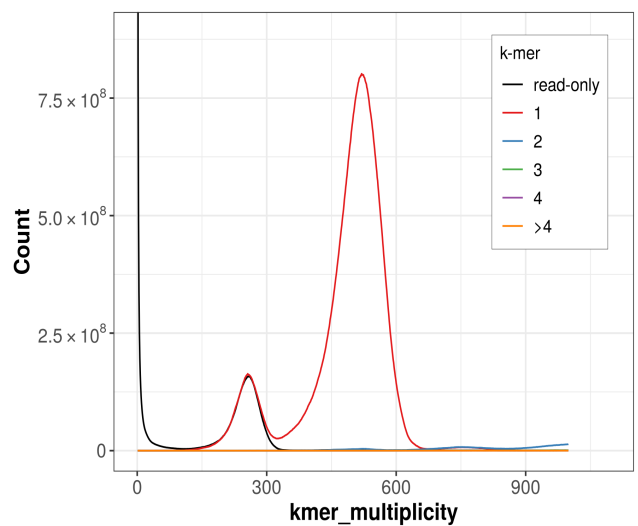
Distribution of k-mer counts per copy numbers found in asm (dipl.)



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

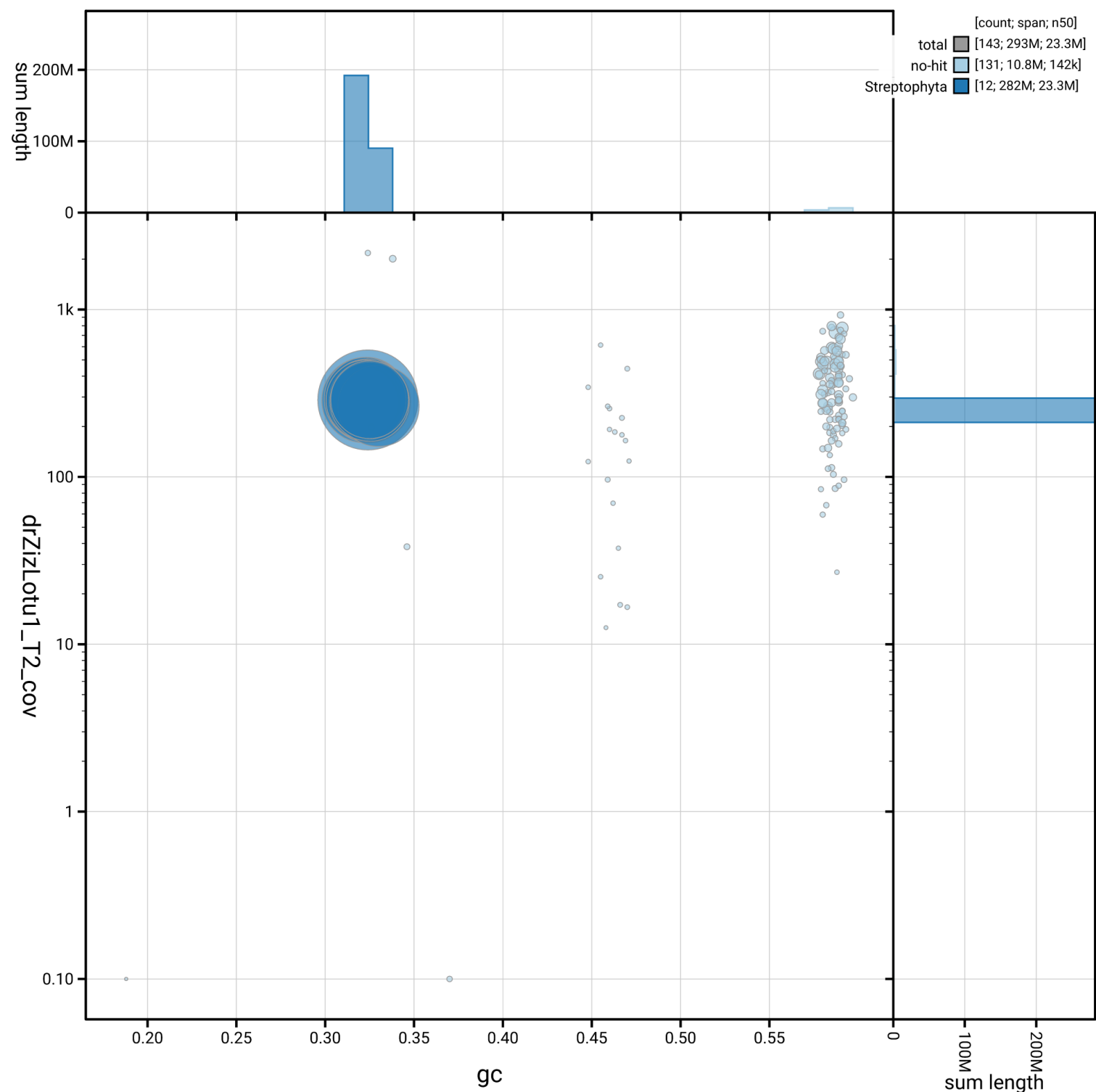


Distribution of k-mer counts per copy numbers found in **hap1** (hap1.)



Distribution of k-mer counts per copy numbers found in **hap2** (hap1.)

# Post-curation contamination screening



**pri.** 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	PacBio HiFi	Arima v2
Coverage	514x	524x

# Assembly pipeline

- **hifiasm**
  - |\_ *ver*: 0.25.0-r726
  - |\_ *key param*: --h1/--h2
- **yahs**
  - |\_ *ver*: 1.2.2
  - |\_ *key param*: NA

# Curation pipeline

- **hifiasm**
  - |\_ *ver*: 0.25.0-r726
  - |\_ *key param*: --h1/--h2
- **yahs**
  - |\_ *ver*: 1.2.2
  - |\_ *key param*: NA
- **TreeVal**
  - |\_ *ver*: 1.2.2
  - |\_ *key param*: NA

Submitter: Michael Paulini

Affiliation: WSI

Date and time: 2025-11-21 16:08:17 CET