

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	502525
ToLID	ddvioUlig1_hapl
Species	Viola uliginosa
Class	Magnoliopsida
Order	Malpighiales

Genome Traits	Expected	Observed
Haploid size (bp)	314,524,744	597,719,963
Haploid Number	10 (source: direct)	10
Ploidy	2 (source: direct)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q64

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

- . Observed Haploid size (bp) has >20% difference with Expected
- . BUSCO single copy value is less than 90% for collapsed
- . BUSCO duplicated value is more than 5% for collapsed

Curator notes

- . Interventions/Gb: 300
- . Contamination notes: ""
- . Other observations: ""

Quality metrics table

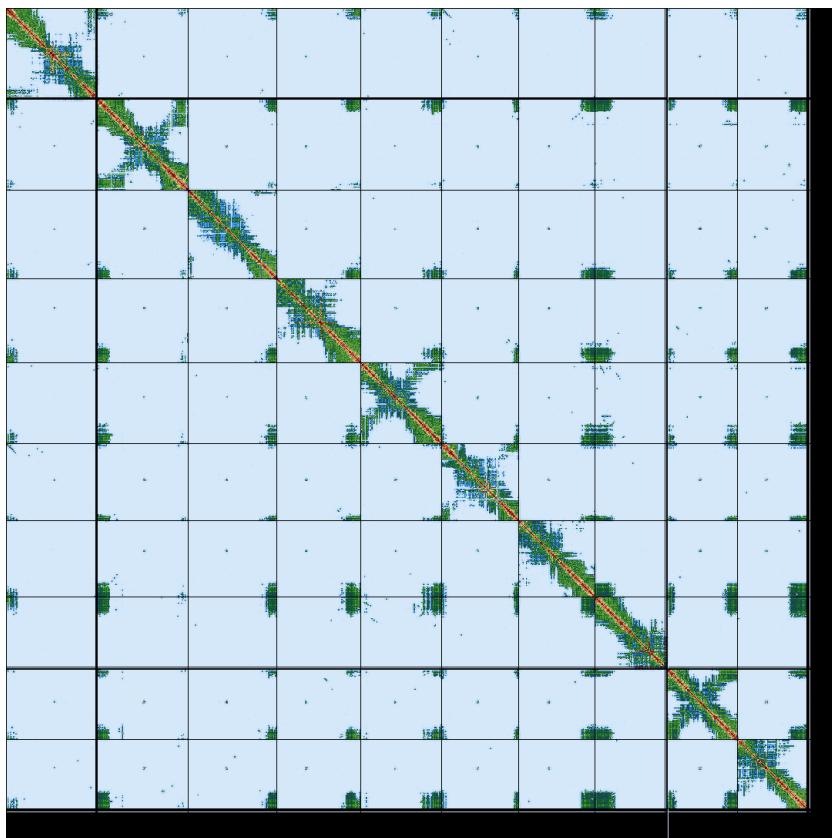
Metrics	Pre-curation collapsed	Curated collapsed
Total bp	607,309,980	597,719,963
GC %	39.56	39.65
Gaps/Gbp	0	321.22
Total gap bp	0	20,900
Scaffolds	700	409
Scaffold N50	10,831,798	58,021,911
Scaffold L50	16	5
Scaffold L90	55	10
Contigs	700	601
Contig N50	10,831,798	4,898,686
Contig L50	16	39
Contig L90	55	131
QV	61.8594	64.1038
Kmer compl.	95.0546	94.6532
BUSCO sing.	15.1%	12.3%
BUSCO dupl.	72.7%	79.8%
BUSCO frag.	1.4%	0.3%
BUSCO miss.	10.9%	7.5%

Warning! BUSCO versions or lineage datasets are not the same across results:

BUSCO: 5.8.2 (euk_genome_met, metaeuk) / Lineage: malpighiales_odb12 (genomes:6, BUSCOS:6134)

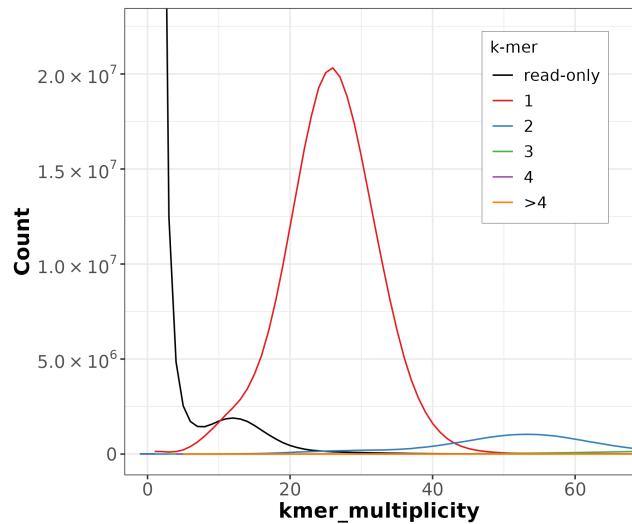
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: malpighiales_odb12 (genomes:6, BUSCOS:6134)

HiC contact map of curated assembly

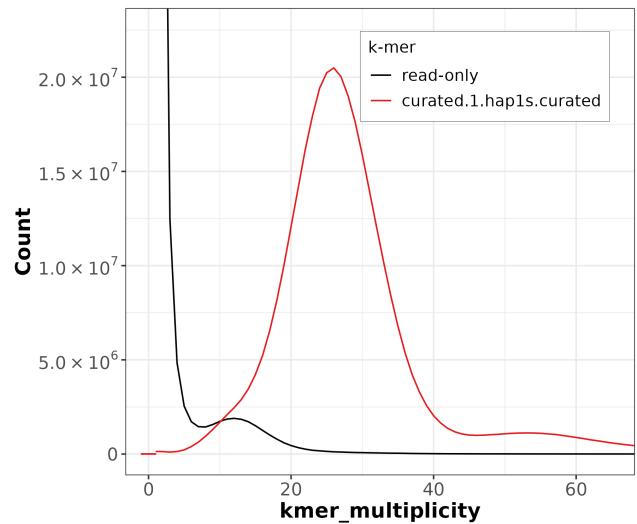


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

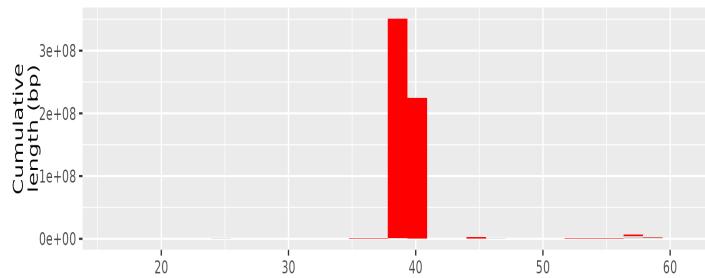


Distribution of k-mer counts per copy numbers found in asm

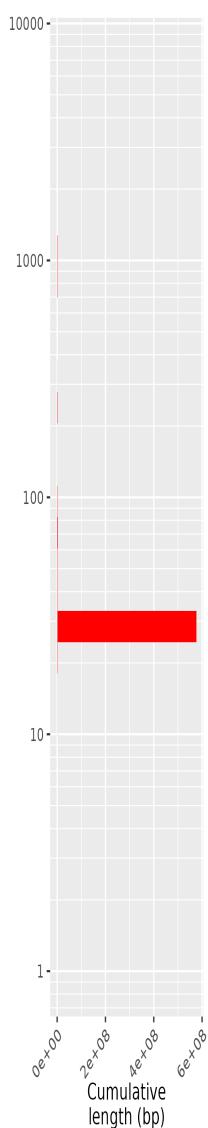
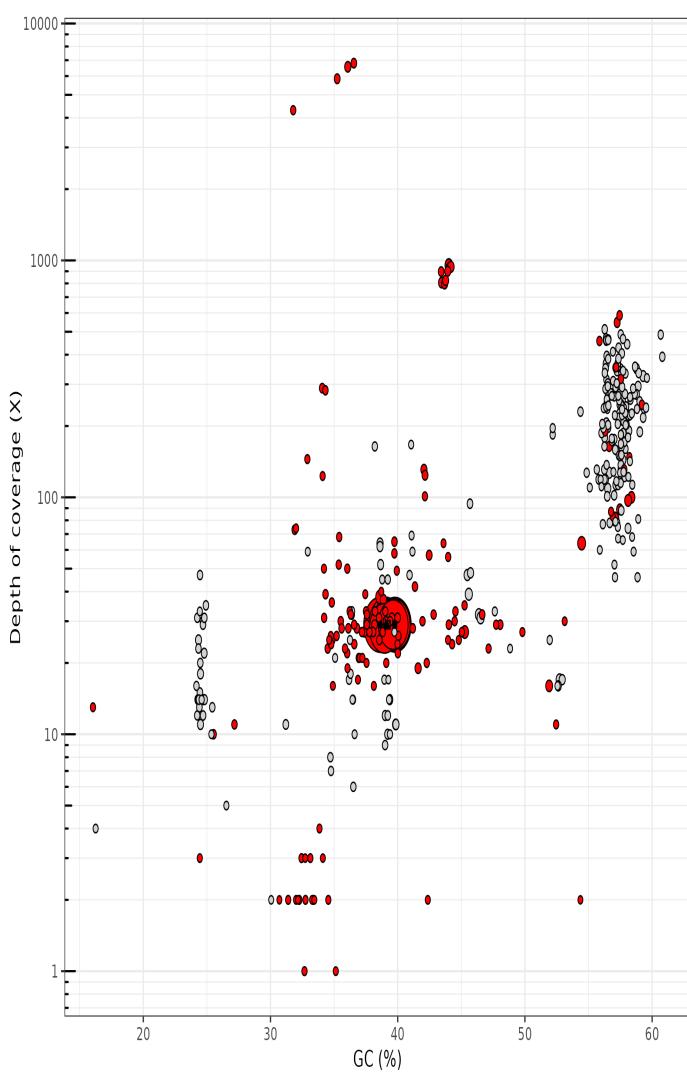


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



Length (bp)

- 2e+07
- 4e+07
- 6e+07

superkingdom

- Eukaryota
- N/A

Longest sequences (bp)

- ddVioUlig1_2 - 65141501 (Eukaryota)
- ▲ ddVioUlig1_1 - 64247586 (Eukaryota)
- ddVioUlig1_3 - 63712635 (Eukaryota)
- + ddVioUlig1_4 - 60474592 (Eukaryota)
- ▣ ddVioUlig1_5 - 58021911 (Eukaryota)

collapsed. 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	Long reads	Arima
Coverage	57	347

Assembly pipeline

```
- Hifiasm
  |_ ver: 0.19.5-r593
  |_ key param: NA
- purge_dups
  |_ ver: 1.2.5
  |_ key param: NA
- YaHS
  |_ ver: 1.2
  |_ key param: NA
```

Curation pipeline

```
- PretextMap
  |_ ver: 0.1.9
  |_ key param: NA
- PretextView
  |_ ver: 0.2.5
  |_ key param: NA
```

Submitter: Lola Demirdjian
Affiliation: Genoscope

Date and time: 2025-12-02 23:45:34 CET