

ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

TxID	43206
ToLID	daReiLigul
Species	Reichardia ligulata
Class	Magnoliopsida
Order	Asterales

Genome Traits	Expected	Observed
Haploid size (bp)	1,632,064,239	1,480,617,540
Haploid Number	8 (source: direct)	8
Ploidy	2 (source: direct)	2
Sample Sex	H	H

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.8.Q60

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

Curator notes

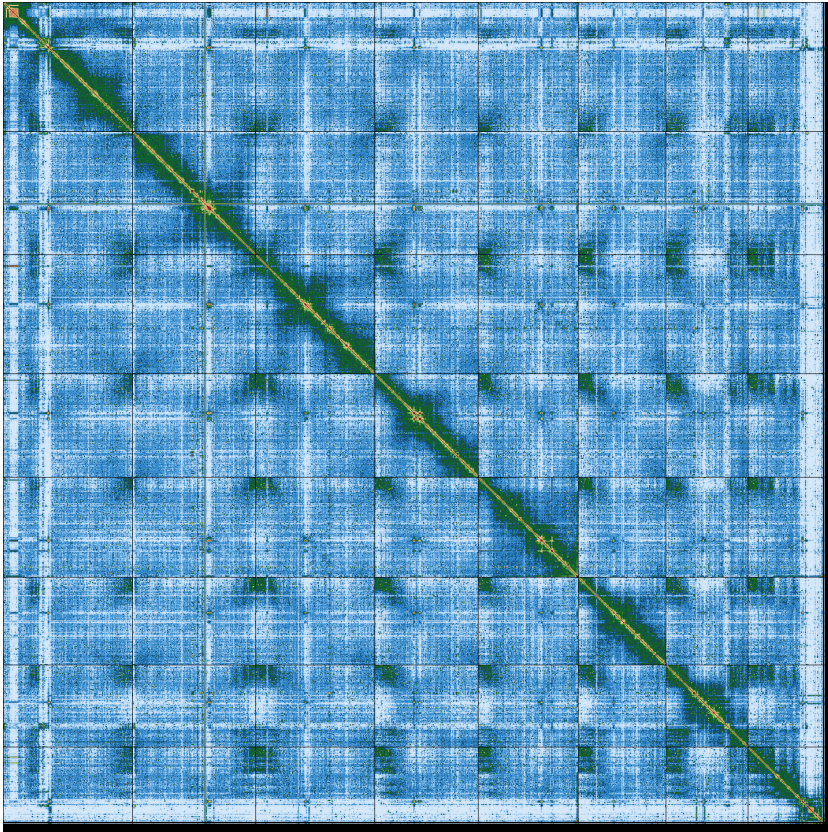
. Interventions/Gb: 130
. Contamination notes: "Contamination report for assembly labelled hap1; Total length of scaffolds removed: 38,332,766 (2.5 %); Scaffolds removed: 861 (80.5 %); Largest scaffold removed: (181,833); FCS-GX contaminant species (number of scaffolds; total length of scaffolds):Pseudomonas coleopterorum, g-proteobacteria (24; 1,352,433);Pseudomonas sp. UBA6276, g-proteobacteria (5; 274,552);Candidatus Portiera aleyrodidarum, g-proteobacteria (2; 60,286);Pseudomonas peli, g-proteobacteria (2; 69,465);Pseudomonas profundus, g-proteobacteria (1; 81,368);Brevundimonas sp. UBA6550, a-proteobacteria (1; 37,834);Pseudomonas cyclaminis, g-proteobacteria (1; 35,464);Hortaea werneckii, ascomycetes (1; 24,606);Mitochondrion (28; 1,705,995);Plastid (796; 34,690,763)"
. Other observations: "This genome has been assembled using PacBio and HiC data and phased. The exact order and orientation of the contigs on chromosome 8 (102 - end Mbp) are unknown."

Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	1,516,984,751	1,480,617,540
GC %	36.84	36.8
Gaps/Gbp	350.04	393.08
Total gap bp	53,100	64,800
Scaffolds	1,069	161
Scaffold N50	167,912,353	184,447,029
Scaffold L50	4	4
Scaffold L90	8	8
Contigs	1,600	743
Contig N50	4,145,258	4,249,769
Contig L50	114	111
Contig L90	389	363
QV	58.5	60.4
Kmer compl.	99.18	99.14
BUSCO sing.	91.4%	91.5%
BUSCO dupl.	4.2%	4.2%
BUSCO frag.	0.6%	0.5%
BUSCO miss.	3.8%	3.8%

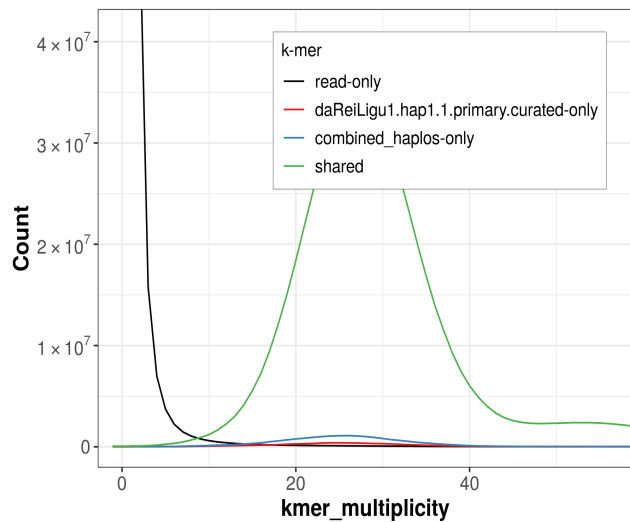
Warning: BUSCO versions or lineage datasets are not the same across results

HiC contact map of curated assembly

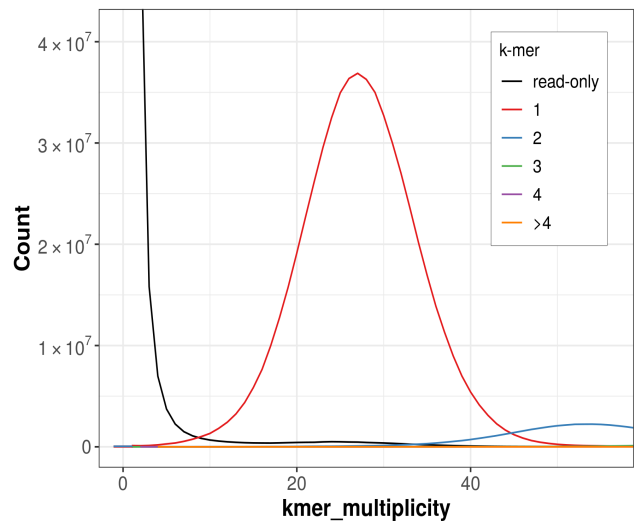


pri [\[LINK\]](#)

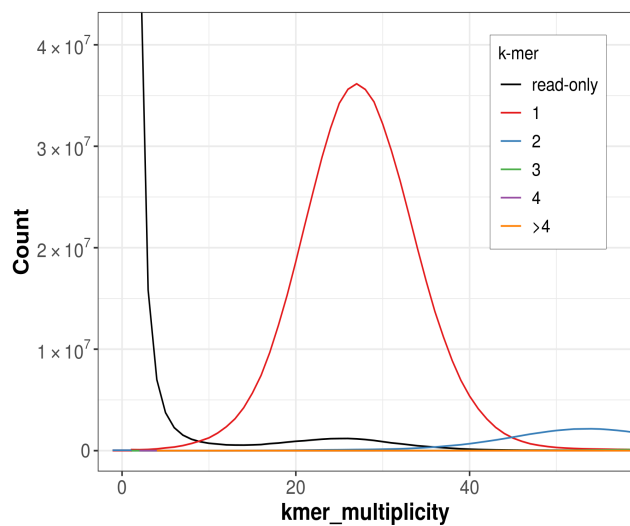
K-mer spectra of curated assembly



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

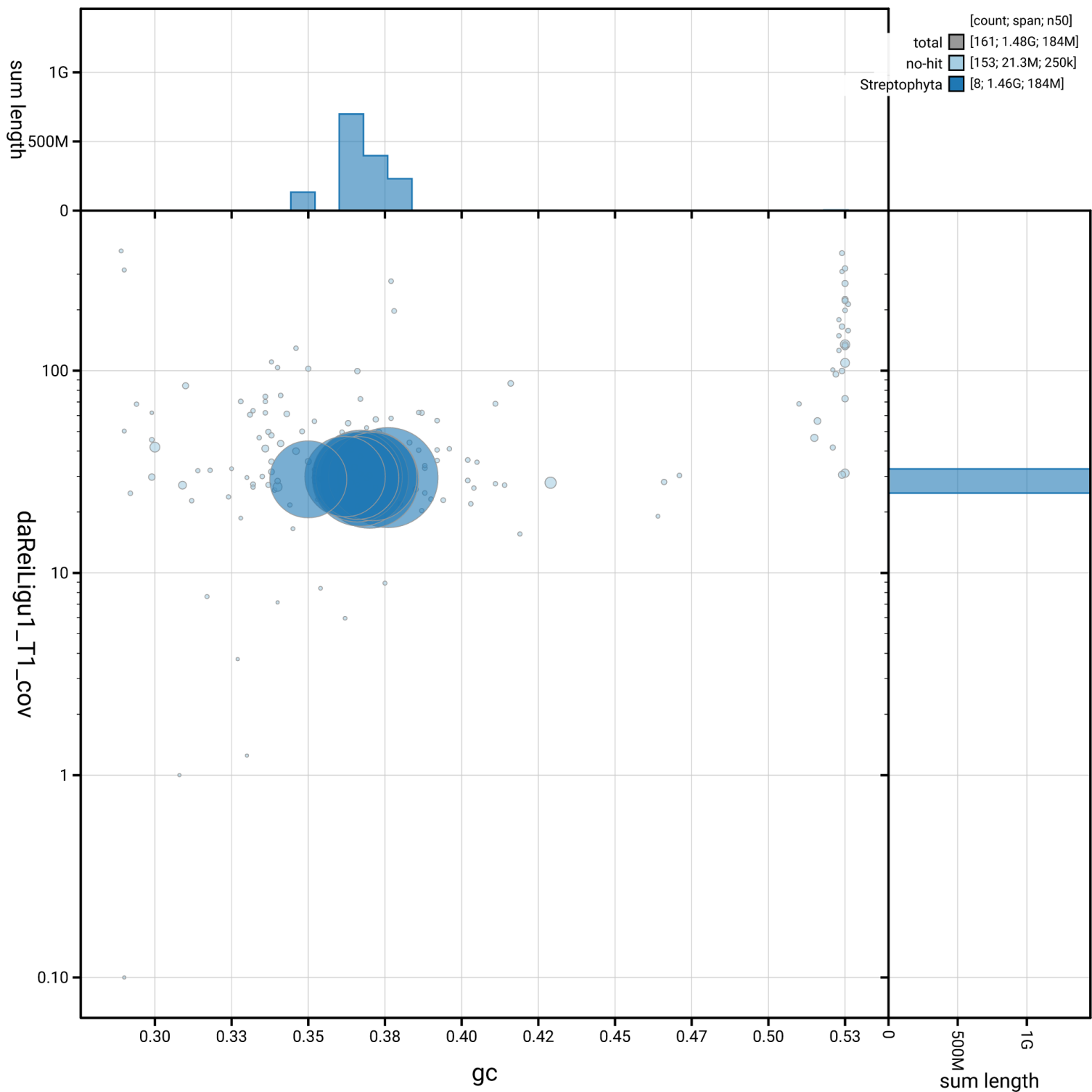


Distribution of k-mer counts per copy numbers found in asm



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

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	28x	120x

Assembly pipeline

- **hifiasm**
 - |_ *ver*: 0.19.8-r603
 - |_ *key param*: --h1/--h2
 - |_ *key param*: --hom-cov 28
- **yahs**
 - |_ *ver*: 1.2.2
 - |_ *key param*: NA

Curation pipeline

- **hifiasm**
 - |_ *ver*: 0.19.8-r603
 - |_ *key param*: --h1/--h2
 - |_ *key param*: --hom-cov 28
- **yahs**
 - |_ *ver*: 1.2.2
 - |_ *key param*: NA
- **TreeVal**
 - |_ *ver*: 1.2.1
 - |_ *key param*: NA

Submitter: Michael Paulini

Affiliation: WSI

Date and time: 2025-02-12 10:29:17 CET