ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	1985348
ToLID	**gdKalVari1**
Species	Kalmusia variispora
Class	Dothideomycetes
Order	Pleosporales

Genome Traits	Expected	Observed
Haploid size (bp)	43,391,403	40,402,156
Haploid Number	4 (source: ancestor)	13
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.6.Q70

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

. Observed Haploid Number is different from Expected

Curator notes

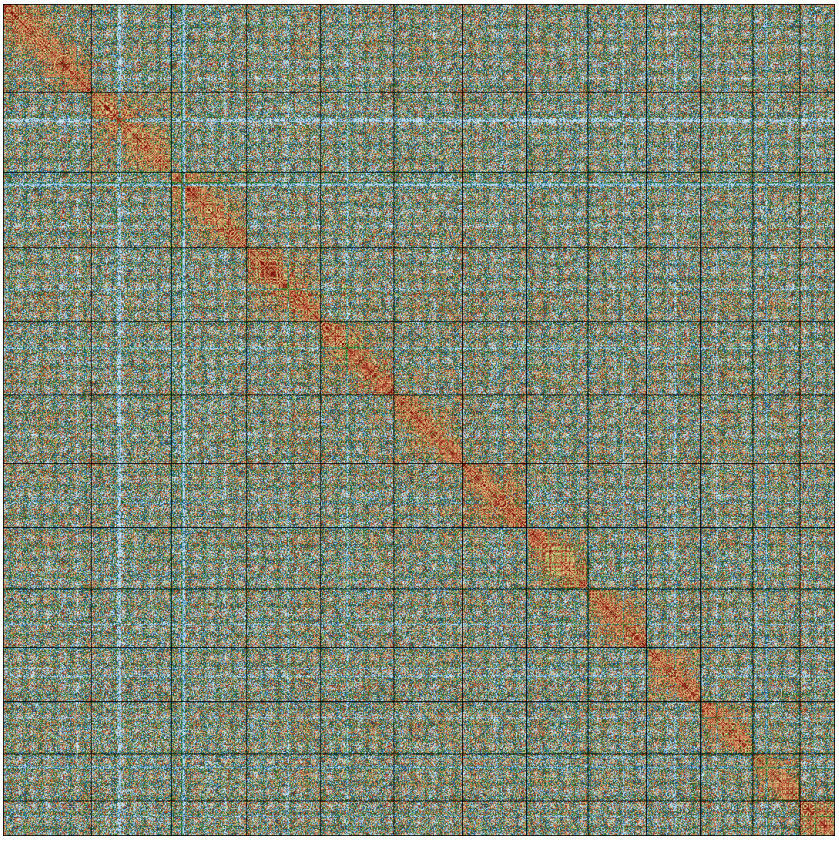
. Interventions/Gb: 99
. Contamination notes: ""
. Other observations: "The assembly of Kalmusia variispora (gdKalVari1) is based on 123X PacBio data and 178X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial PacBio assembly generation with Hifiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. No contigs were found to be contaminants (bacterial, archaeal, or viral). Additionally, 64 regions totaling 2.191 Mb (with the largest being 0.065 Mb) were identified as haplotypic duplications and removed. The mitochondrial genome was assembled using Oatk. Finally, the primary assembly was analyzed and manually improved using Pretext. During manual curation, 1 haplotypic region was removed, totaling 0.017Mb. Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	40,447,405	40,402,156
GC %	50.8	50.83
Gaps/Gbp	49.45	74.25
Total gap bp	200	500
Scaffolds	15	14
Scaffold N50	3,305,911	3,305,911
Scaffold L50	6	6
Scaffold L90	12	11
Contigs	17	17
Contig N50	3,136,180	3,136,180
Contig L50	6	6
Contig L90	12	12
QV	68.834	70.5243
Kmer compl.	99.8136	99.8104
BUSCO sing.	98.4%	98.4%
BUSCO dupl.	0.1%	0.1%
BUSCO frag.	0.0%	0.0%
BUSCO miss.	1.6%	1.6%

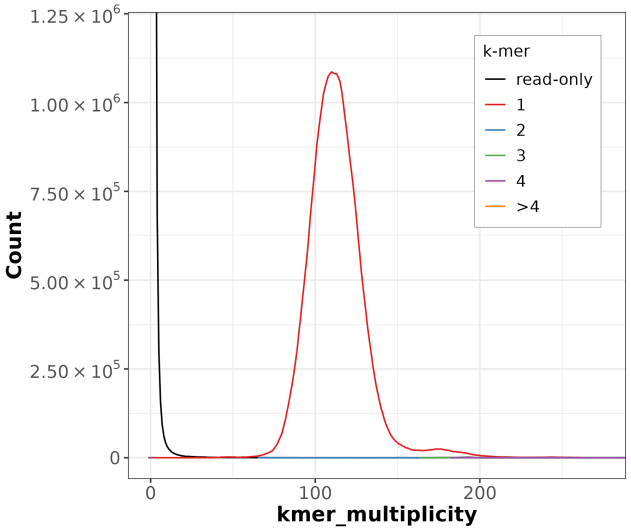
BUSCO: 6.1.0 (euk_genome_min, miniprot) / Lineage: pleosporales_odb12.2 (genomes:76, BUSCOs:5661)

HiC contact map of curated assembly

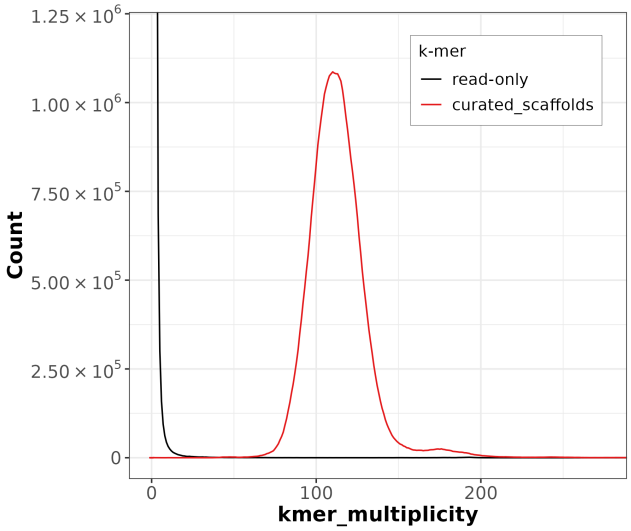


collapsed [LINK]

K-mer spectra of curated assembly

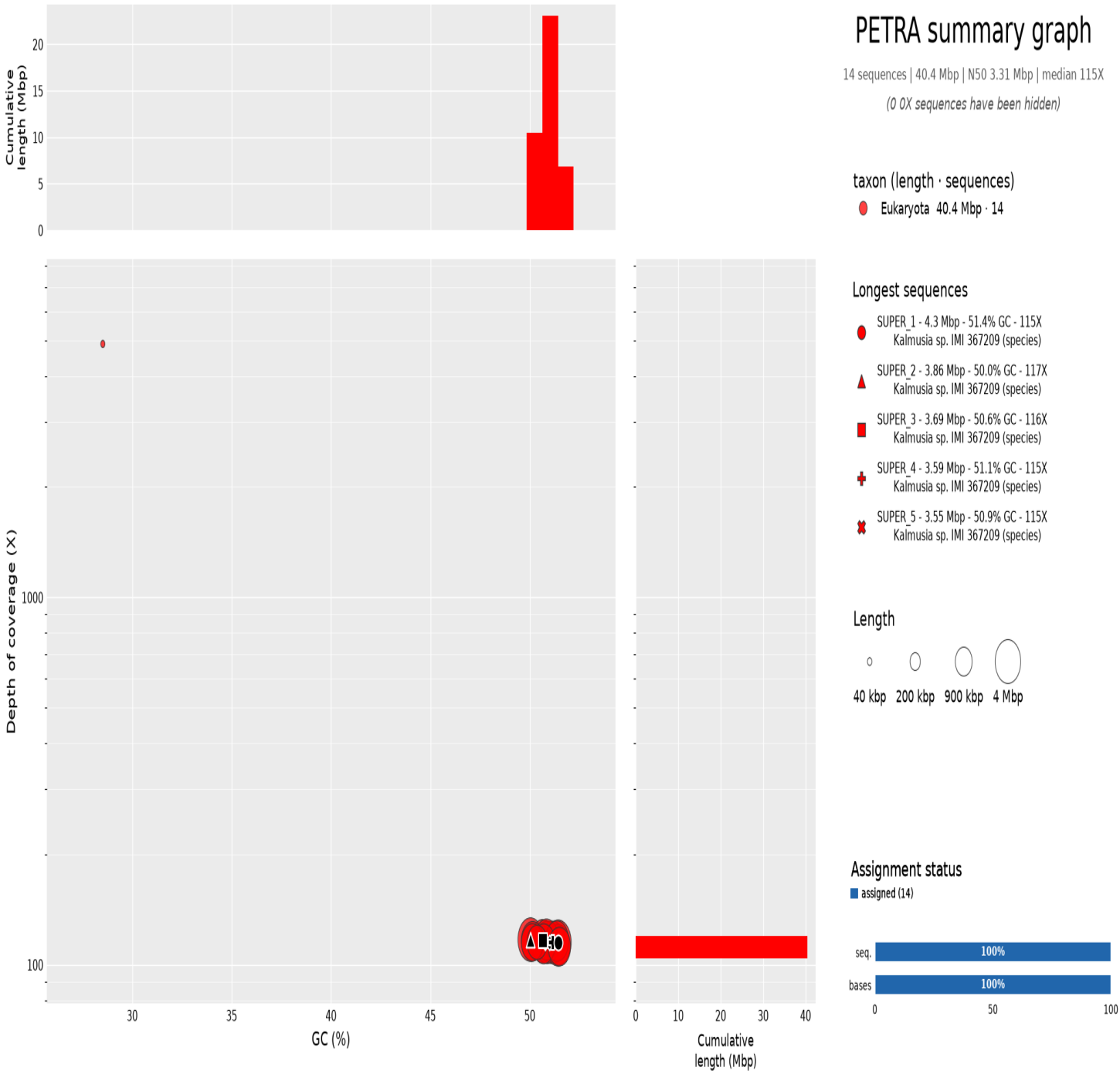


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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	123	178

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:* 0.2.4
  - *key param:* NA
- **PretextView**
  - *ver:* 1.0.7
  - *key param:* NA

Submitter: Jean-Marc Aury
Affiliation: Genoscope

Date and time: 2026-10-23 18:46:52 CEST