ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	2732611
ToLID	**gdParFlab1**
Species	Parathyridaria flabelliae
Class	Dothideomycetes
Order	Pleosporales

Genome Traits	Expected	Observed
Haploid size (bp)	44,569,452	41,755,988
Haploid Number	4 (source: ancestor)	15
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.6.Q68

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

. Observed Haploid Number is different from Expected

Curator notes

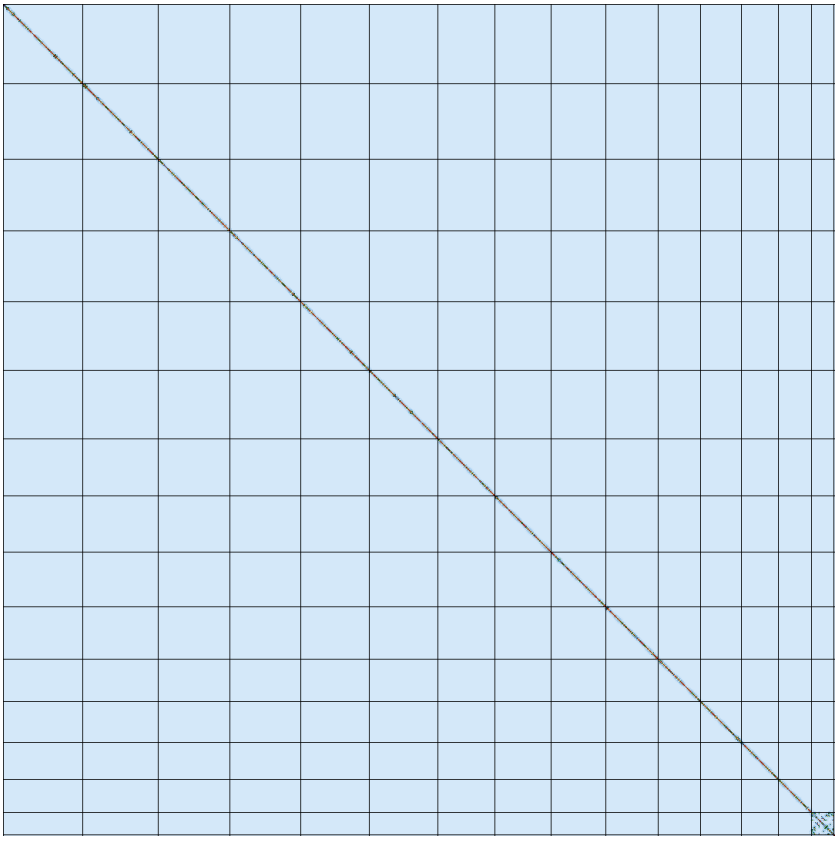
. Interventions/Gb: 0
. Contamination notes: ""
. Other observations: "The assembly of Parathyridaria flabelliae (gdParFlab1) is based on 101.02X PacBio data and 153.04X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: 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, 59 regions totaling 1.903 Mb (with the largest being 0.056 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 contaminant sequence was removed, totaling 0.097Mb. Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. Chromosome 15 shows a rather different Hi-C contact pattern compared with the other chromosomes. It appears to contain numerous repetitive regions, but the sequencing coverage and the absence of gaps does not indicate any obvious assembly issue "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	41,802,868	41,755,988
GC %	50.18	50.21
Gaps/Gbp	0	0
Total gap bp	0	0
Scaffolds	16	16
Scaffold N50	3,421,887	3,421,887
Scaffold L50	6	6
Scaffold L90	13	13
Contigs	16	16
Contig N50	3,421,887	3,421,887
Contig L50	6	6
Contig L90	13	13
QV	66.9592	68.6165
Kmer compl.	99.8921	99.8919
BUSCO sing.	98.3%	98.0%
BUSCO dupl.	0.1%	0.2%
BUSCO frag.	0.2%	0.2%
BUSCO miss.	1.4%	1.6%

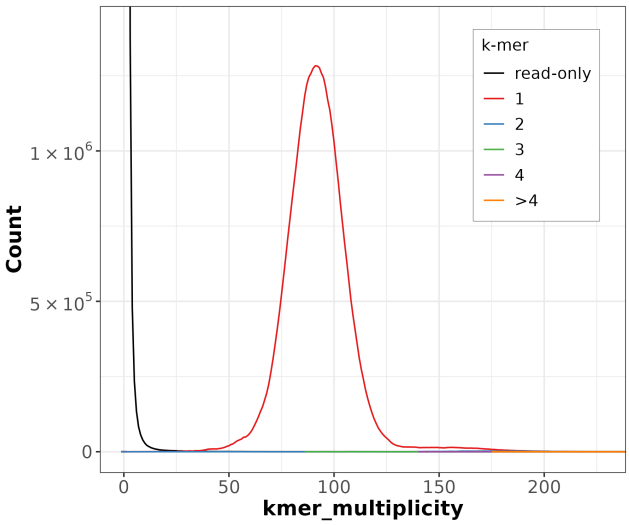
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 6.1.0 (euk_genome_min, miniprot) / Lineage: pleosporales_odb12 (genomes:76, BUSCOs:6066)
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: pleosporales_odb12 (genomes:76, BUSCOs:6066)

HiC contact map of curated assembly

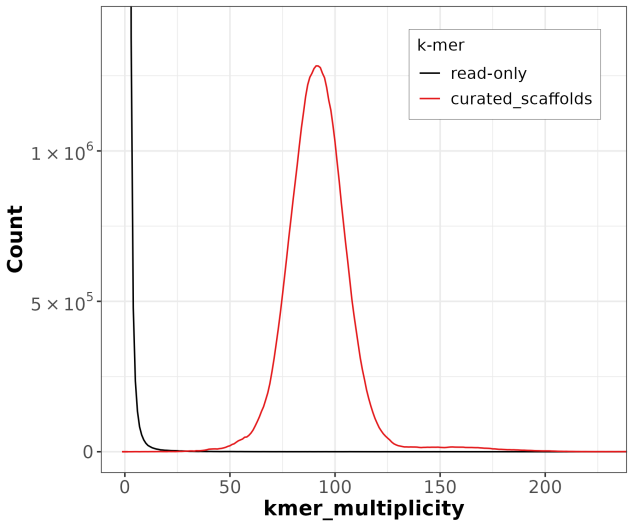


collapsed [LINK]

K-mer spectra of curated assembly

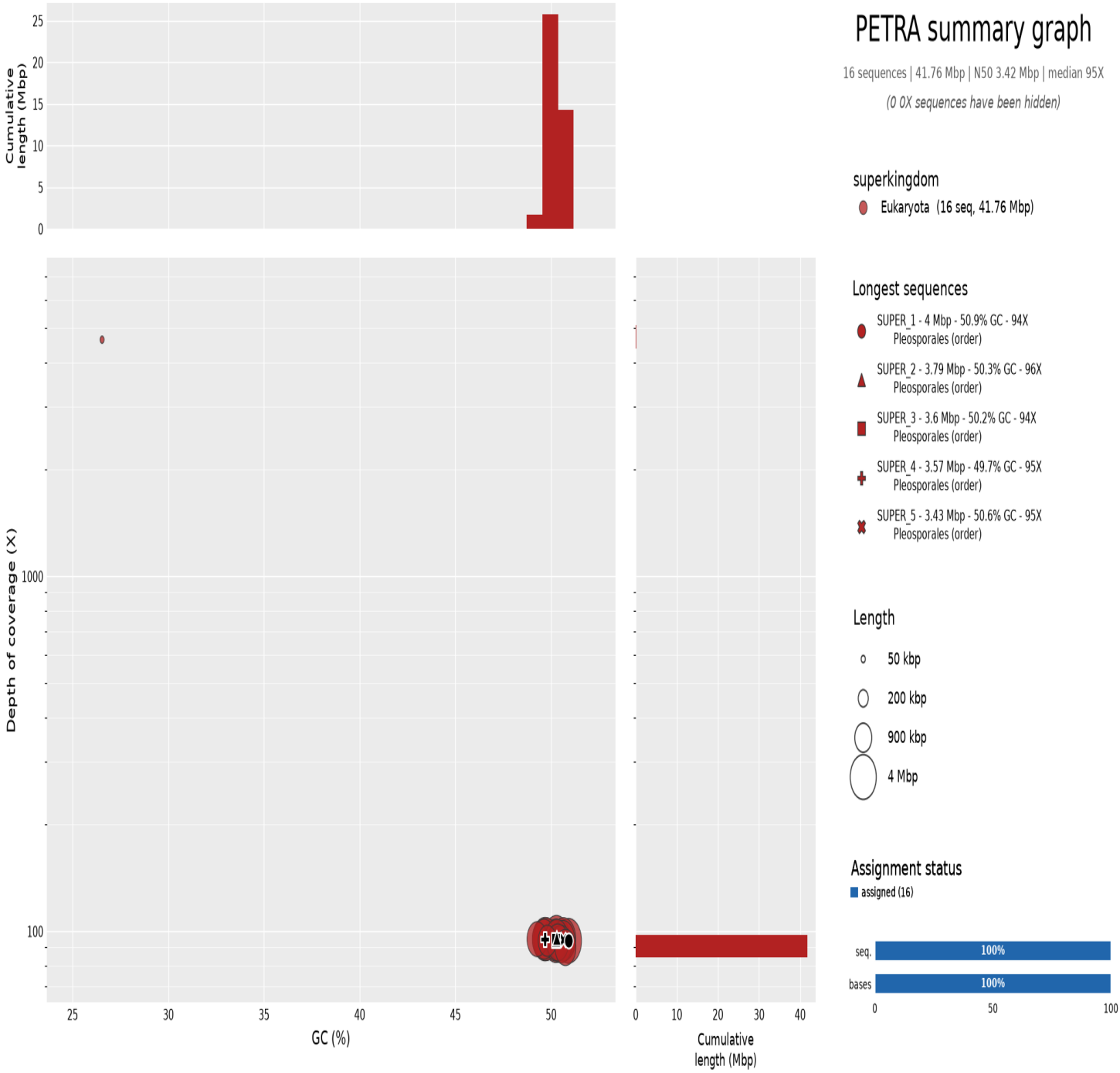


Distribution of k-mer counts per copy numbers found in asm



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	101	153

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: Caroline Menguy
Affiliation: Genoscope

Date and time: 2026-10-05 02:17:04 CEST