ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	2653846
ToLID	**gdRouPadi1**
Species	Roussoella padinae
Class	Dothideomycetes
Order	Pleosporales

Genome Traits	Expected	Observed
Haploid size (bp)	45,128,588	42,169,467
Haploid Number	4 (source: ancestor)	14
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.6.Q64

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

. Observed Haploid Number is different from Expected

Curator notes

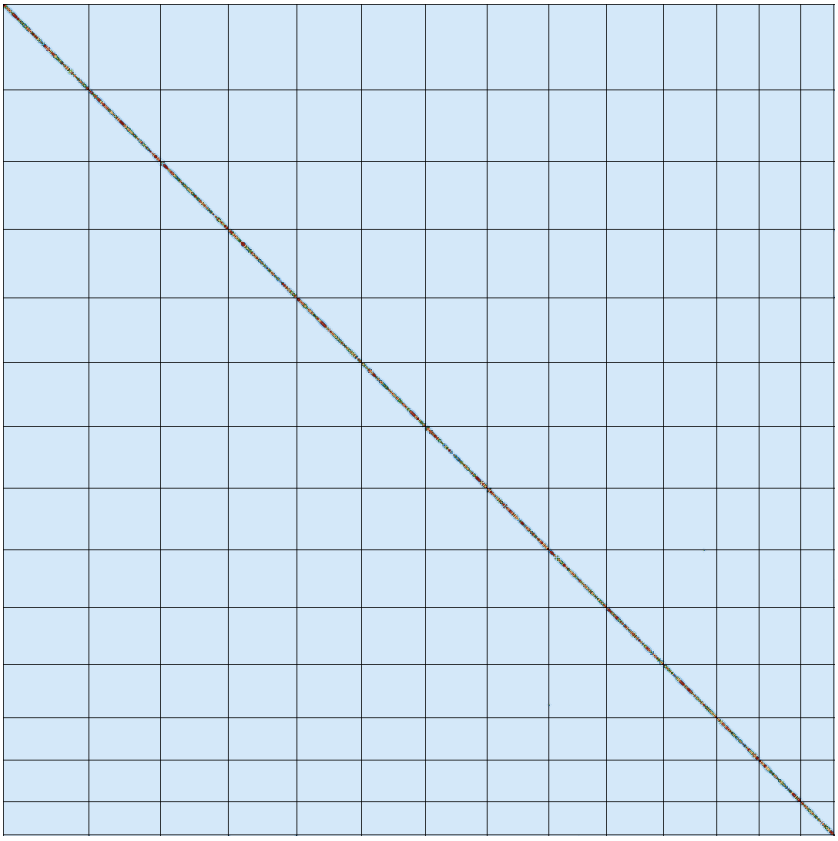
. Interventions/Gb: 0
. Contamination notes: ""
. Other observations: "The assembly of Roussoella padinae (gdRouPadi1) is based on 79.82X PacBio data and 144.14X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial 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, 29 regions totaling 0.734 Mb (with the largest being 0.041 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. 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	42,145,063	42,169,467
GC %	48.77	48.76
Gaps/Gbp	23.73	23.71
Total gap bp	100	100
Scaffolds	15	15
Scaffold N50	3,258,589	3,258,589
Scaffold L50	6	6
Scaffold L90	12	12
Contigs	16	16
Contig N50	3,258,589	3,258,589
Contig L50	6	6
Contig L90	13	13
QV	64.1024	64.3151
Kmer compl.	99.9377	99.9395
BUSCO sing.	97.9%	97.9%
BUSCO dupl.	0.2%	0.2%
BUSCO frag.	0.0%	0.0%
BUSCO miss.	1.8%	1.8%

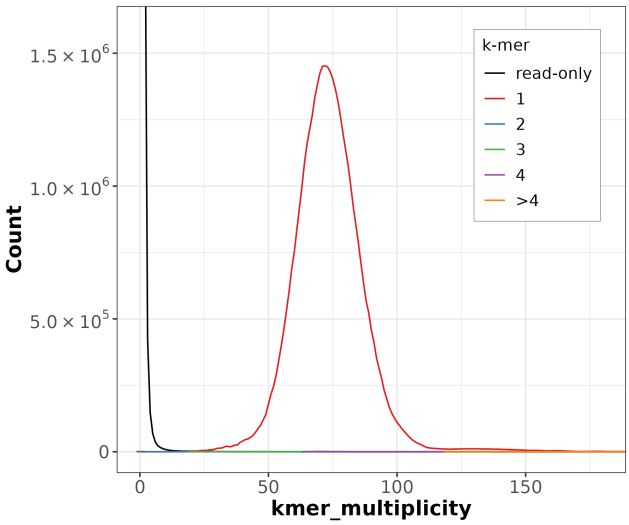
BUSCO: 6.1.0 (euk_genome_min, miniprot) / Lineage: pleosporales_odb12.2 (genomes:76, BUSCOs:5661)

HiC contact map of curated assembly

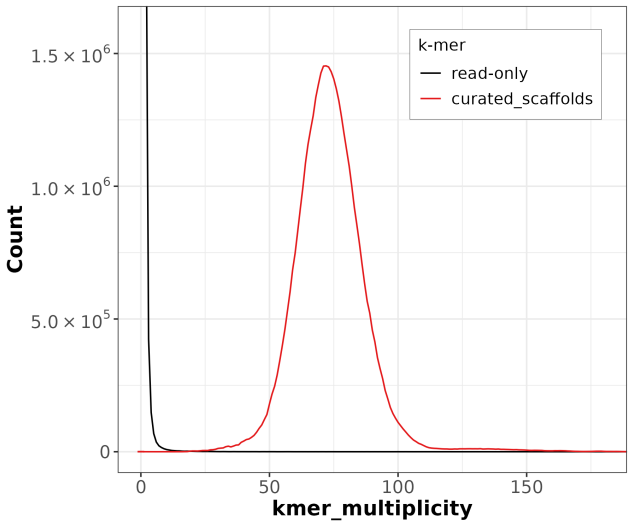


collapsed [LINK]

K-mer spectra of curated assembly

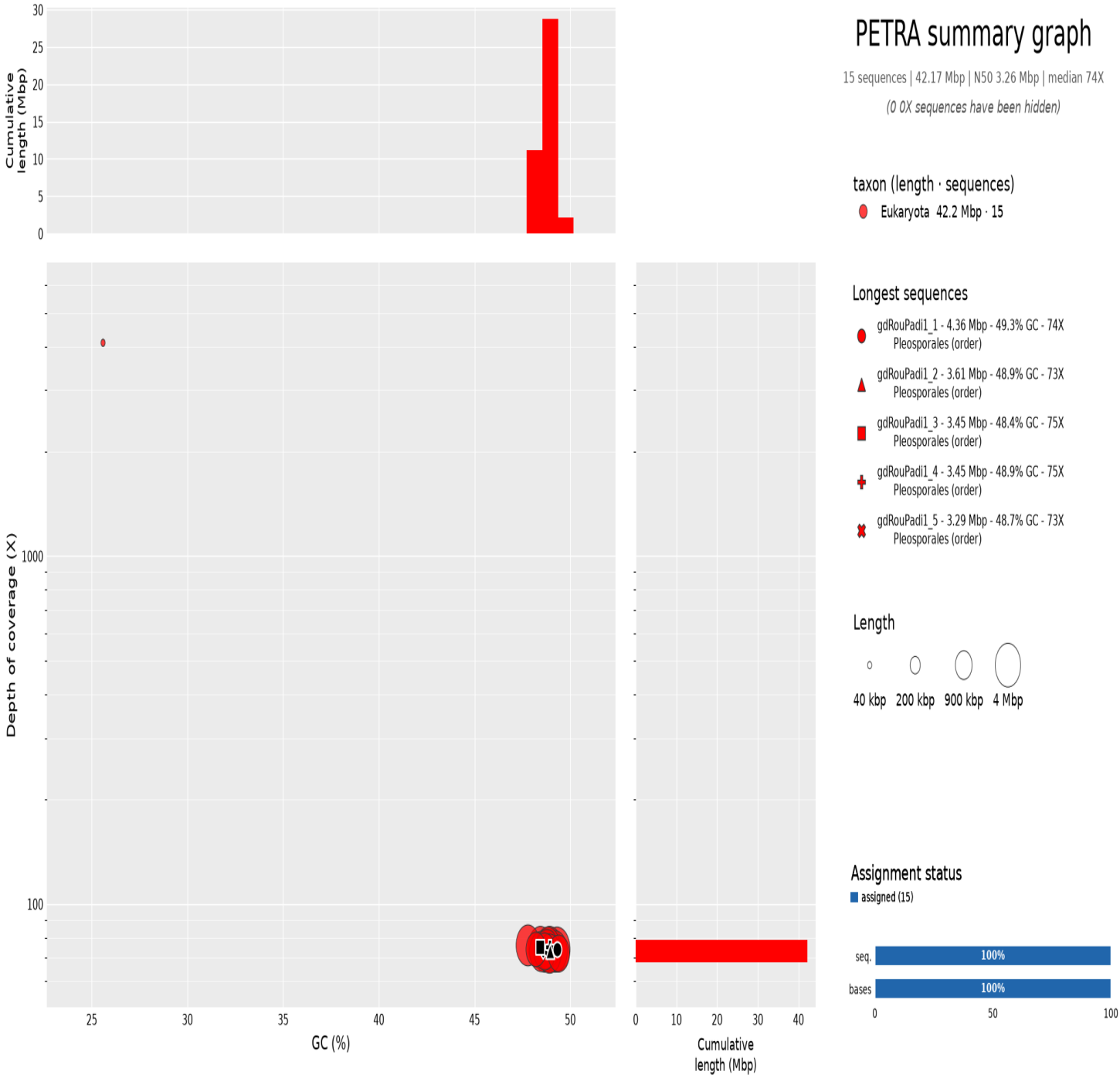


Distribution of k-mer counts per copy numbers found in asm



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	79	144

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: Emilie Teodori
Affiliation: Genoscope

Date and time: 2026-10-19 15:25:02 CEST