

ERGA Assembly Report

v24.09.10

Tags: ERGA-BGE

TxID	3034287
ToLID	xgDerLasi
Species	Deroceras lasithionense
Class	Gastropoda
Order	Stylommatophora

Genome Traits	Expected	Observed
Haploid size (bp)	1,014,304,119	1,318,418,884
Haploid Number	16 (source: ancestor)	31
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q55

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

- . Observed Haploid size (bp) has >20% difference with Expected
- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for collapsed

Curator notes

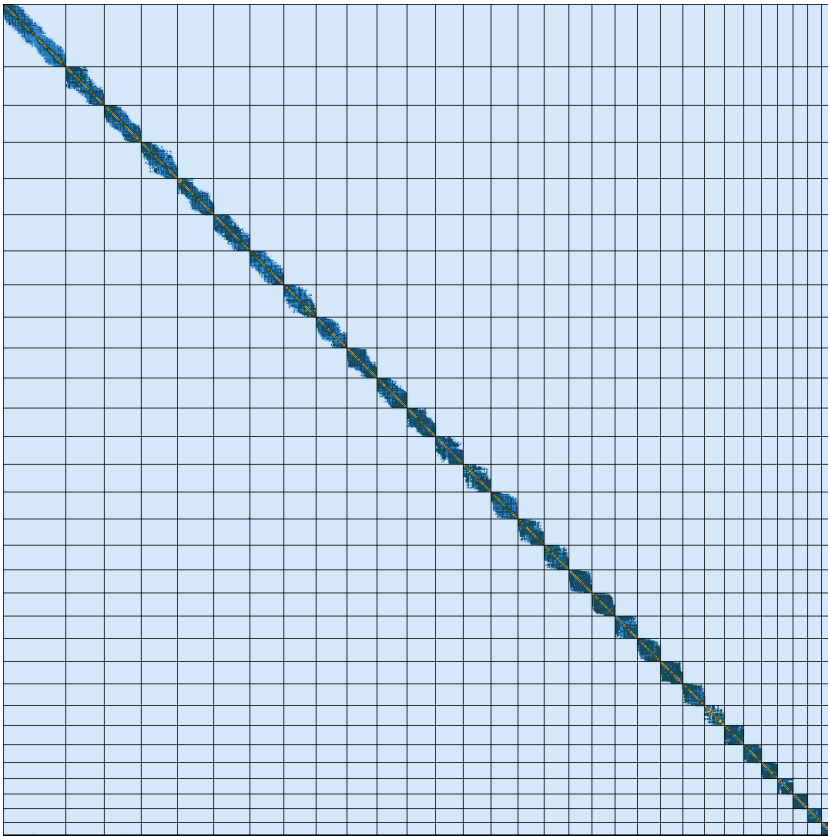
- . Interventions/Gb: 73
- . Contamination notes: "3 bacterial contigs were removed "
- . Other observations: ""

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	1,330,276,266	1,318,418,884
GC %	41.44	41.45
Gaps/Gbp	798.33	811.58
Total gap bp	106,200	109,500
Scaffolds	90	66
Scaffold N50	45,757,421	45,757,421
Scaffold L50	12	12
Scaffold L90	26	26
Contigs	1,152	1,136
Contig N50	2,037,934	2,037,218
Contig L50	190	188
Contig L90	630	628
QV	55.0595	55.0543
Kmer compl.	77.1708	76.6606
BUSCO sing.	98.0%	98.0%
BUSCO dupl.	1.6%	1.6%
BUSCO frag.	0.0%	0.0%
BUSCO miss.	0.4%	0.4%

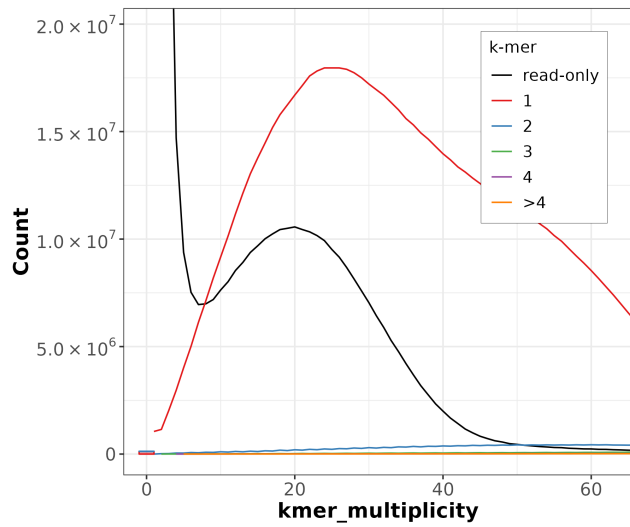
BUSCO 5.4.3 Lineage: eukaryota_odb10 (genomes:70, BUSCOs:255)

HiC contact map of curated assembly

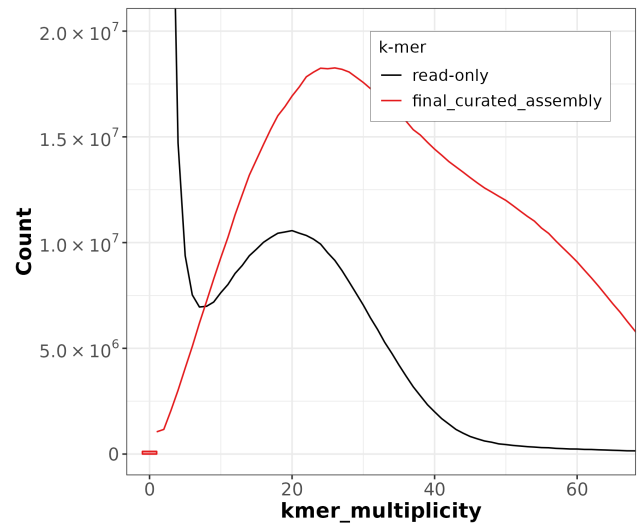


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

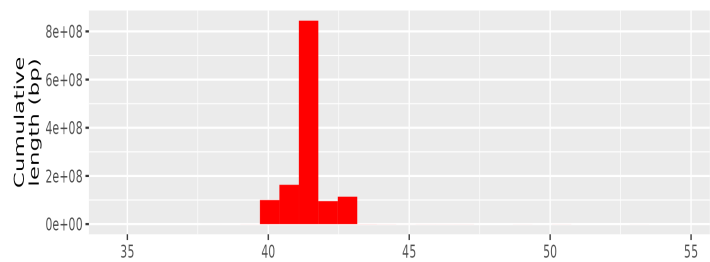


Distribution of k-mer counts per copy numbers found in asm

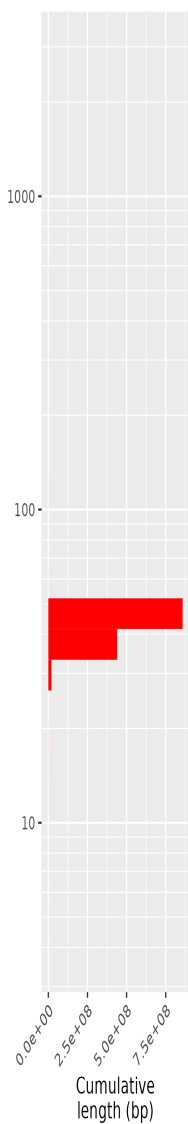
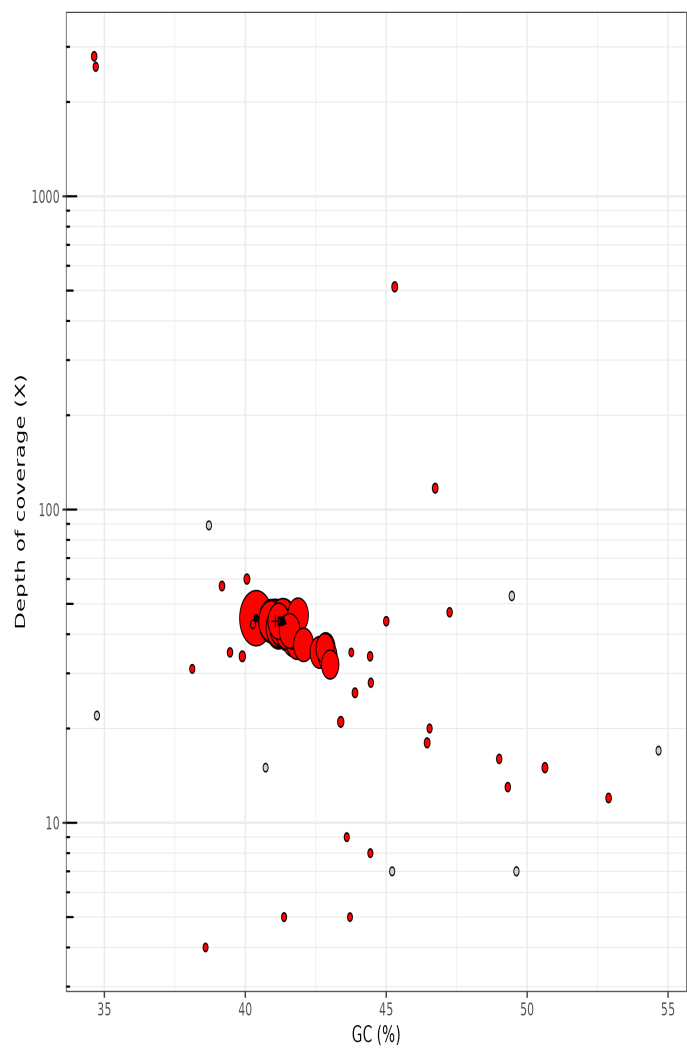


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



Longest sequences (bp)

- xgDerLasj_1 - 99419141 (Eukaryota)
- ▲ xgDerLasj_2 - 61325393 (Eukaryota)
- xgDerLasj_3 - 58383161 (Eukaryota)
- + xgDerLasj_4 - 57976078 (Eukaryota)
- ▣ xgDerLasj_5 - 57252941 (Eukaryota)

superkingdom

- Eukaryota
- N/A

Length (bp)

- 2.5e+07
- 5.0e+07
- 7.5e+07

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	PACBIO Hifi	Arima (4-enz)
Coverage	45	42

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: Caroline Belser

Affiliation: Genoscope

Date and time: 2024-09-17 16:46:39 CEST