

ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

TxID	2712997
ToLID	icDenFora10
Species	Dendarus foraminosus
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	557,593,036	593,990,249
Haploid Number	10 (source: ancestor)	11
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q62

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

- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for collapsed
- . Not 90% of assembly in chromosomes for collapsed

Curator notes

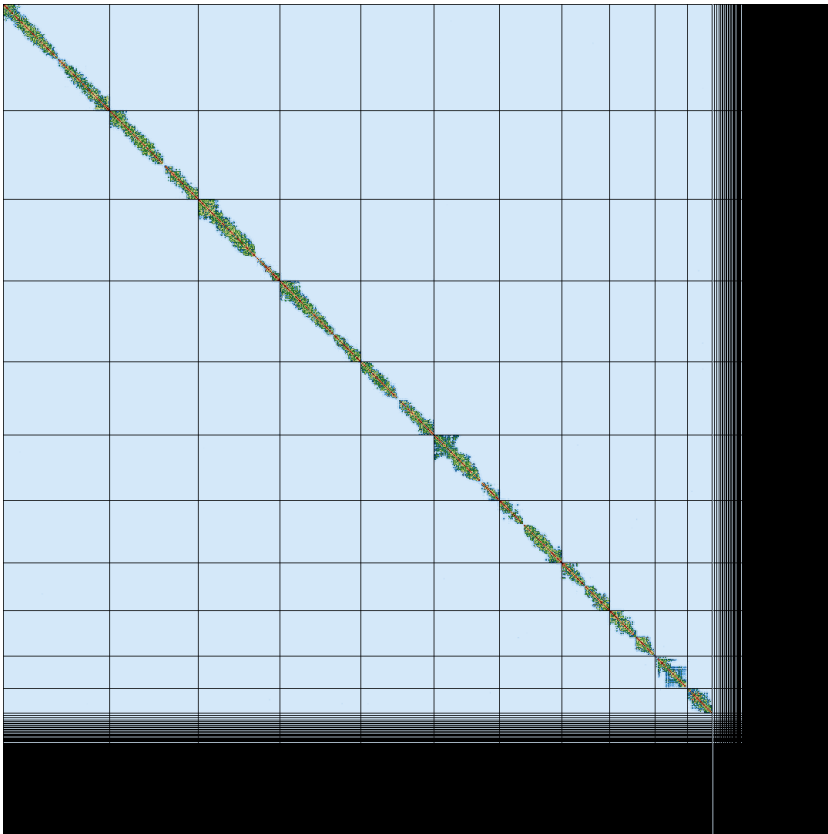
- . Interventions/Gb: 4
- . Contamination notes: "some bacterial contigs have been removed "
- . Other observations: "scaffolds (9,11) with low coverage can be (X,Y), telomeric repeat pattern is TCGGG. "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	593,969,168	593,990,249
GC %	35.92	35.92
Gaps/Gbp	16.84	25.25
Total gap bp	1,000	2,000
Scaffolds	419	417
Scaffold N50	41,134,000	51,873,200
Scaffold L50	6	5
Scaffold L90	52	49
Contigs	429	432
Contig N50	24,404,000	24,404,000
Contig L50	9	9
Contig L90	61	63
QV	62.2928	62.293
Kmer compl.	66.781	66.7811
BUSCO sing.	98.0%	98.0%
BUSCO dupl.	1.6%	1.6%
BUSCO frag.	0.4%	0.4%
BUSCO miss.	0.0%	0.0%

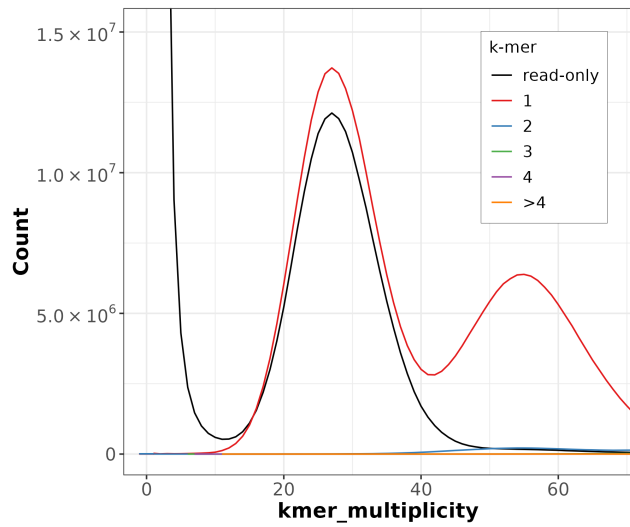
BUSCO 5.4.3 Lineage: eukaryota_odb10 (genomes:70, BUSCOs:255)

HiC contact map of curated assembly

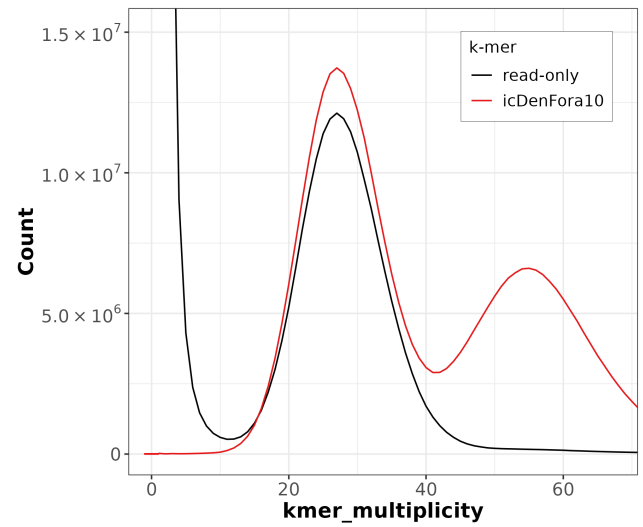


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

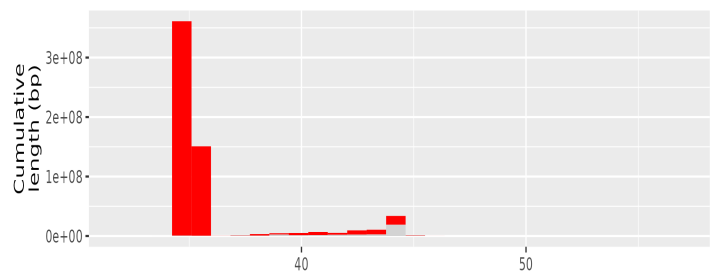


Distribution of k-mer counts per copy numbers found in asm



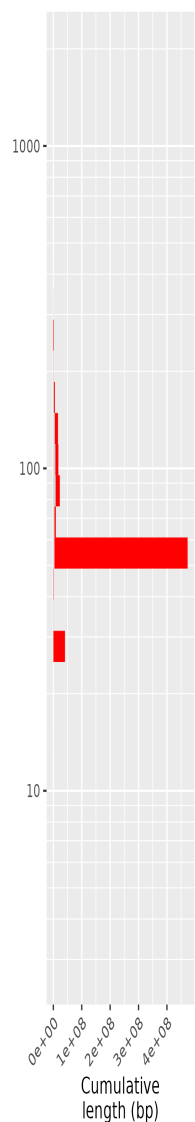
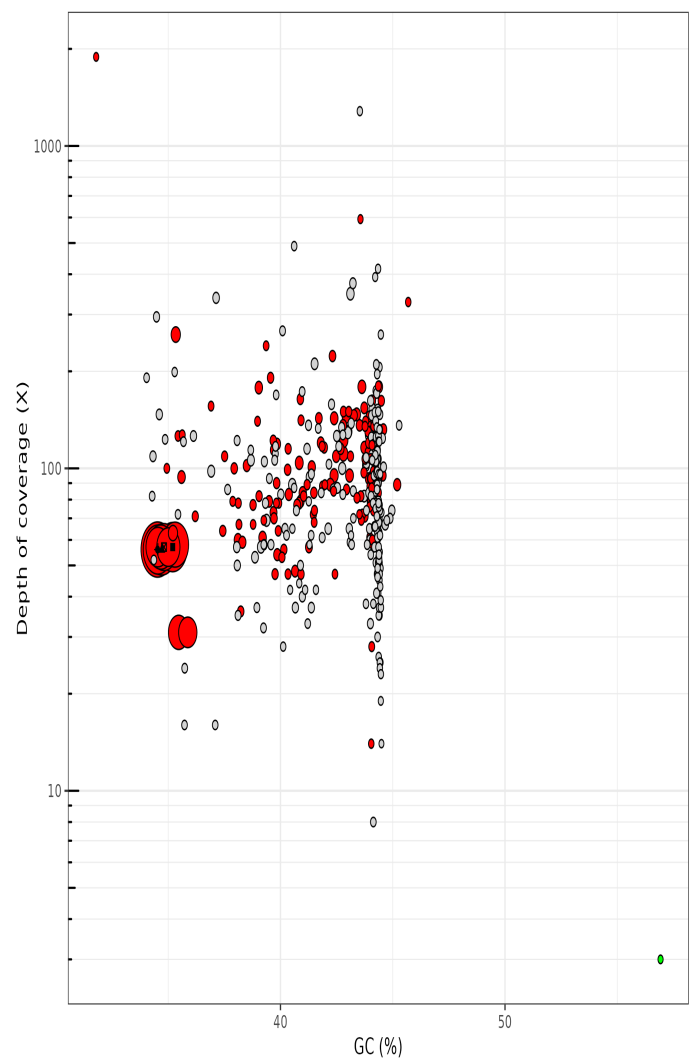
Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph

(2 0X contigs have been hidden)



Longest sequences (bp)

- icDenFora_1 - 76253626 (Eukaryota)
- ▲ icDenFora_2 - 63295100 (Eukaryota)
- icDenFora_6 - 58385200 (Eukaryota)
- + icDenFora_3 - 57732603 (Eukaryota)
- ▣ icDenFora_4 - 51873200 (Eukaryota)

Length (bp)

- 2e+07
- 4e+07
- 6e+07

superkingdom

- Bacteria
- Eukaryota
- N/A

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	Omnic
Coverage	58	82

Assembly pipeline

- **Hifiasm**
 - |_ *ver*: 0.19.5-r593
 - |_ *key param*: NA
- **purge_dups**
 - |_ *ver*: 1.2.5
 - |_ *key param*: NA

Curation 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

Submitter: S.Duprat

Affiliation: Genoscope

Date and time: 2024-07-31 10:45:45 CEST