

ERGA Assembly Report

v24.09.10

Tags: ERGA-BGE

TxID	10034
ToLID	mCriCri2
Species	Cricetus cricetus
Class	Mammalia
Order	Rodentia

Genome Traits	Expected	Observed
Haploid size (bp)	3,273,505,721	2,520,162,027
Haploid Number	11 (source: direct)	11
Ploidy	2 (source: ancestor)	2
Sample Sex	XY	XY

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.8.Q54

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

- . Observed Haploid size (bp) has >20% difference with Expected

Curator notes

- . Interventions/Gb: 8 . Contamination notes: "No contaminants detected in the Nextdenovo assembly selected for curation."

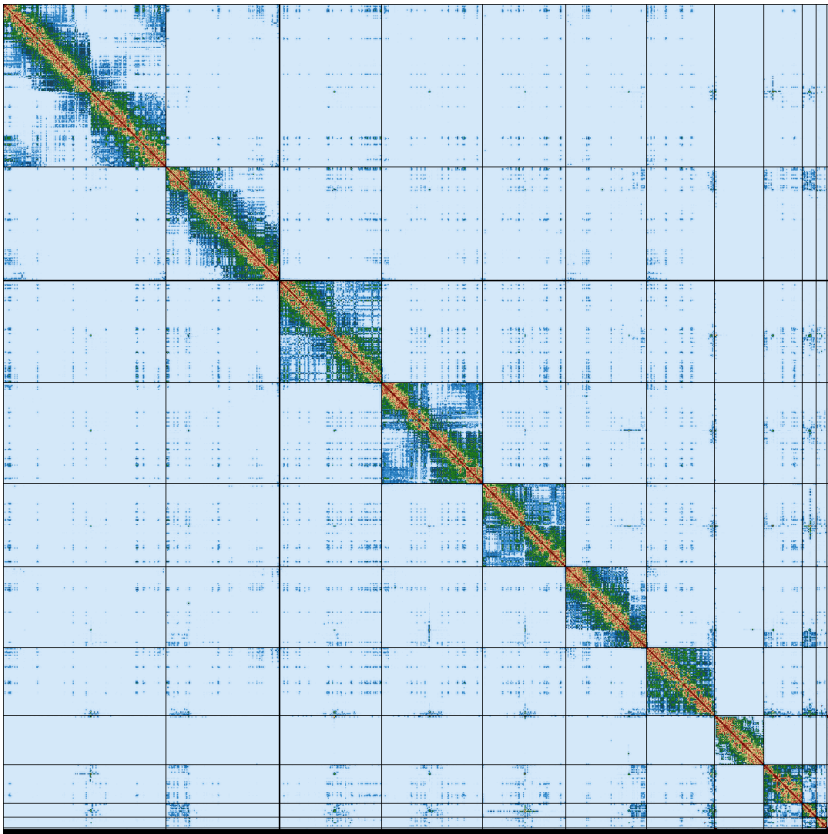
- . Other observations: "Curation made 1 cut in a contig, 4 breaks at gaps and 9 joins. SUPER_Y sequence accounts for 5,195,203 bp. Despite of searching thoroughly for Y sequence in the purged haplotigs removed from our assemblies we did not found enough to scaffold the male Y chromosome. Mitogenome was not successfully assembled. During review we reoriented a contig in SUPER_2 and localized 3 scaffolds. Thus we made just 2 additional breaks at gaps and 3 additional joins."

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	2,520,160,827	2,520,162,027
GC %	41.46	41.46
Gaps/Gbp	49.2	51.58
Total gap bp	24,800	26,000
Scaffolds	54	48
Scaffold N50	250,072,296	305,542,590
Scaffold L50	5	4
Scaffold L90	9	8
Contigs	178	178
Contig N50	47,690,202	47,690,202
Contig L50	17	17
Contig L90	55	55
QV	54.4254	54.4254
Kmer compl.	94.0784	94.0784
BUSCO sing.	95.3%	95.3%
BUSCO dupl.	2.8%	2.8%
BUSCO frag.	0.7%	0.7%
BUSCO miss.	1.2%	1.2%

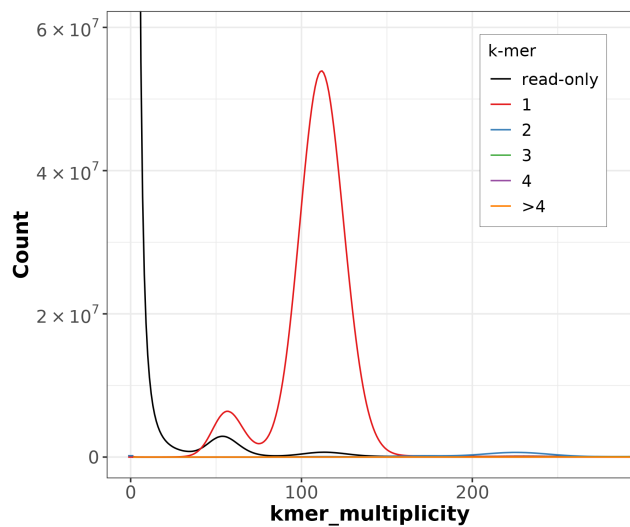
BUSCO 5.4.0 Lineage: vertebrata_odb10 (genomes:67, BUSCOs:3354)

HiC contact map of curated assembly

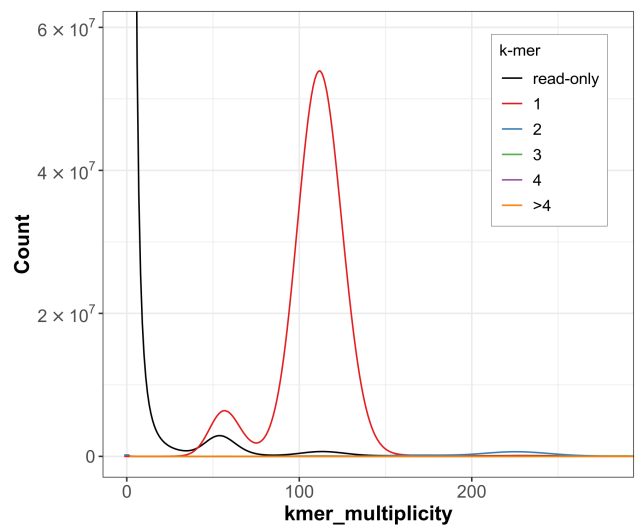


collapsed [\[LINK\]](#)

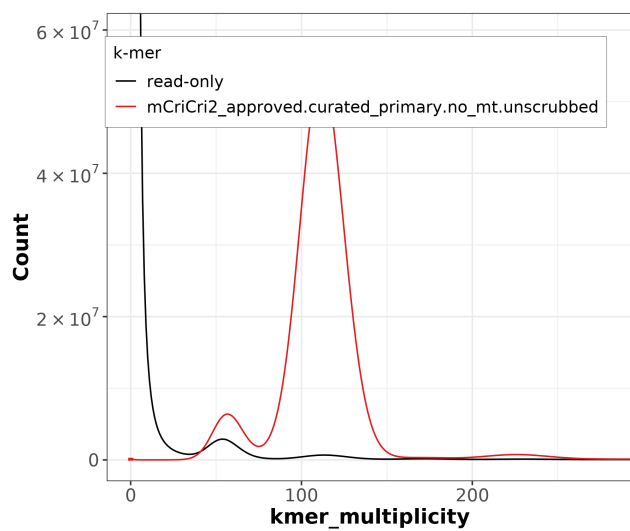
K-mer spectra of curated assembly



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



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



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

Post-curation contamination screening

Data profile

Data	ONT	Illumina	Omnich
Coverage	105x	84x	78x

Assembly pipeline

- **CLAWS pipeline**
 - |_ *ver*: 2.2.0
 - |_ *key param*: NA
- **FOAM pipeline**
 - |_ *ver*: 0.5
 - |_ *key param*: NA
- **Trim_galore**
 - |_ *ver*: 0.6.7
 - |_ *key param*: NA
- **Filtlong**
 - |_ *ver*: 0.2.1
 - |_ *key param*: NA
- **flye**
 - |_ *ver*: 2.9.1
 - |_ *key param*: NA
- **nextdenovo**
 - |_ *ver*: 2.5.0
 - |_ *key param*: NA
- **hypo**
 - |_ *ver*: 1.0.3
 - |_ *key param*: NA
- **purge_dups**
 - |_ *ver*: 1.2.6
 - |_ *key param*: NA
- **YaHS**
 - |_ *ver*: 1.2a
 - |_ *key param*: NA

Curation pipeline

- **PretextView**
 - |_ *ver*: 0.2.5
 - |_ *key param*: NA
- **GRIT_Rapid**
 - |_ *ver*: 2.0
 - |_ *key param*: NA

Date and time: 2024-11-05 09:56:55 CET