ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	187398
ToLID	**mLepCas1**
Species	Lepus castroviejoi
Class	Mammalia
Order	Lagomorpha

Genome Traits	Expected	Observed
Haploid size (bp)	2,745,042,277	2,919,066,766
Haploid Number	24 (source: ancestor)	24
Ploidy	2 (source: ancestor)	2
Sample Sex	XY	XY

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.8.Q69

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

Curator notes

. Interventions/Gb: 166
. Contamination notes: "Some contaminants were found with the pipeline NCBI FCS GX v Galaxy Version 0.5.5+galaxy1, and some scaffolds were flagged as contaminants by BlobToolKit. The one flagged in the present blob plot are confirmed as non contaminat based on the HiC contact map and the NCBI FCS pipeline"
. Other observations: "We were able to reconstruct the sex chromosome. We confirmed the assignation of the X and Y flag with an alignment with a close species, Lepus timidus (GCA_040893245.2) "

Quality metrics table

| Metrics	Pre-curation hap1	Curated hap1
Total bp	2,953,087,076	2,919,066,766
GC %	43.84	43.78
Gaps/Gbp	69.76	81.88
Total gap bp	41,200	47,800
Scaffolds	236	212
Scaffold N50	142,000,722	137,255,412
Scaffold L50	9	9
Scaffold L90	19	20
Contigs	442	451
Contig N50	74,601,310	58,008,274
Contig L50	14	16
Contig L90	55	61
QV	69.0173	69.1112
Kmer compl.	98.5651	98.173
BUSCO sing.	94.9%	93.8%
BUSCO dupl.	3.3%	3.6%
BUSCO frag.	0.9%	0.9%
BUSCO miss.	1.0%	1.7%

BUSCO: 5.8.0 (euk_genome_min, miniprot) / Lineage: mammalia_odb10 (genomes:24, BUSCOs:9226)
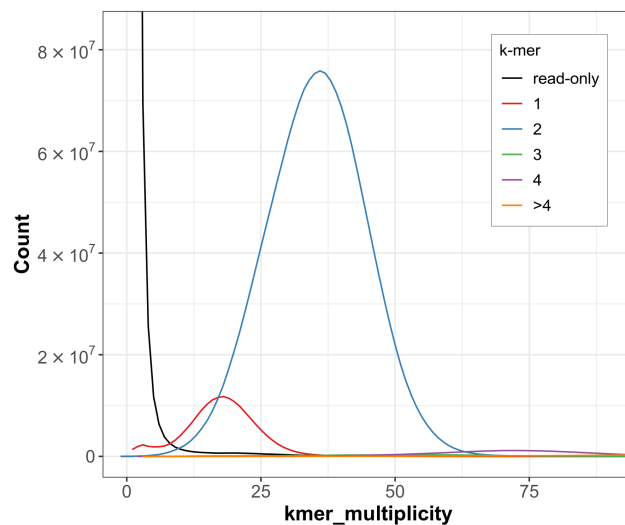
HiC contact map of curated assembly
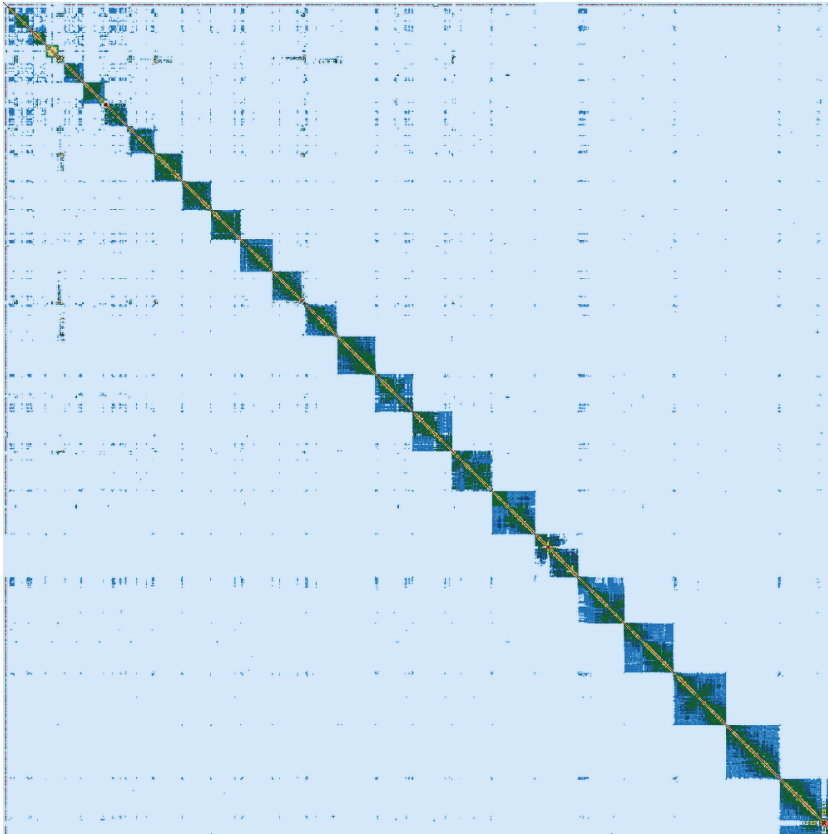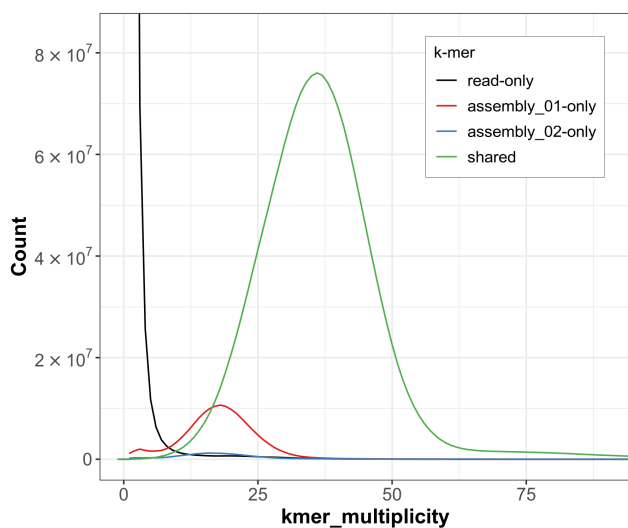

hap1 [LINK]

K-mer spectra of curated assembly
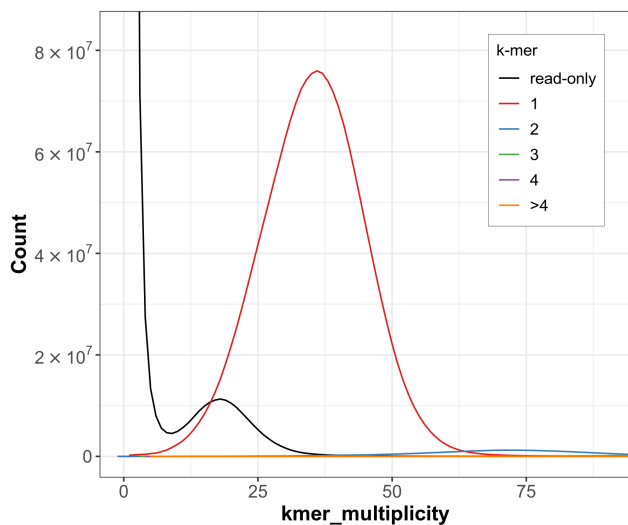
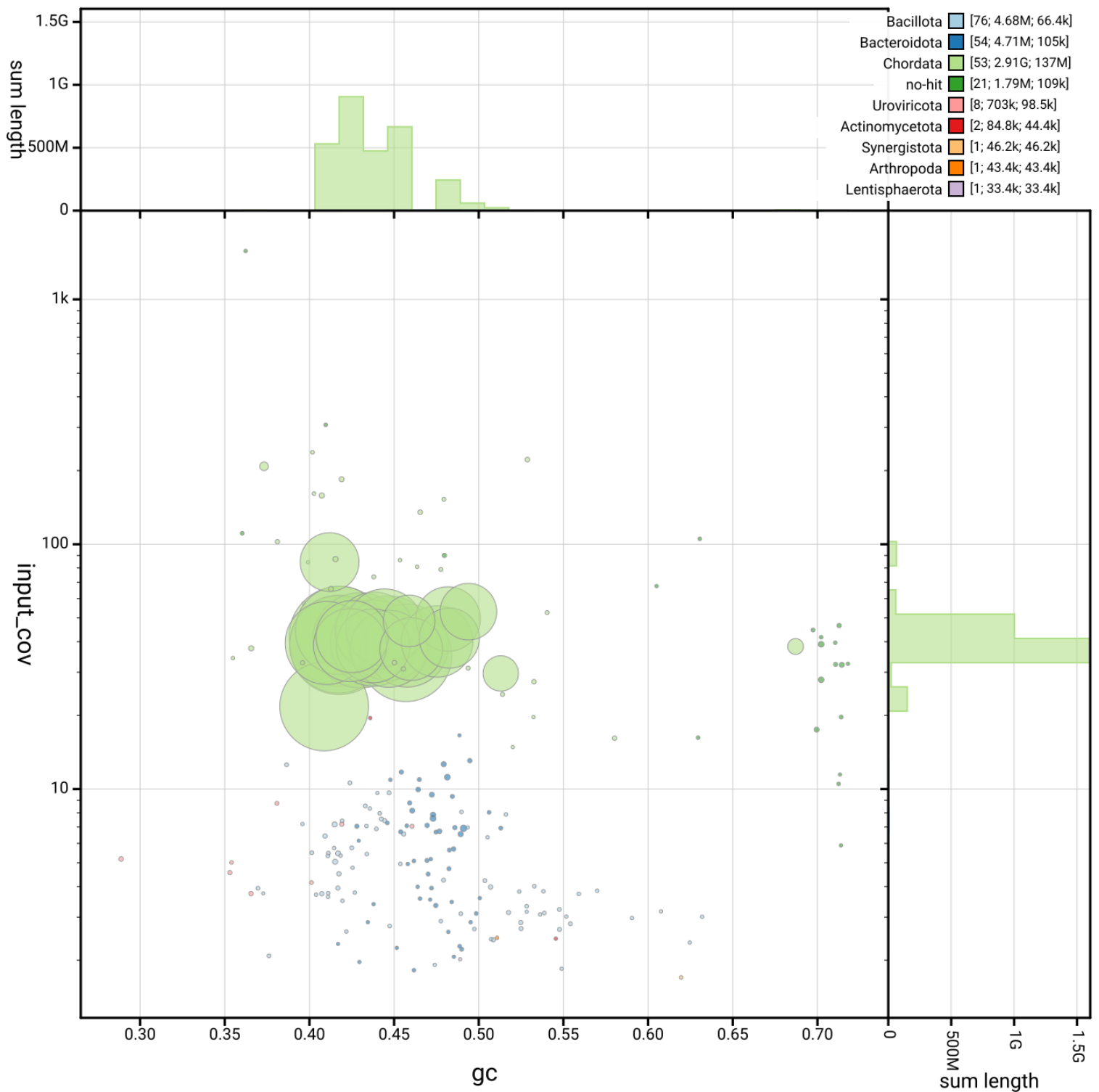Distribution of k-mer counts per copy
numbers found in asm

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

Legend:
- Bacillota [76; 4.68M; 66.4k]
- Bacteroidota [54; 4.71M; 105k]
- Chordata [53; 2.91G; 137M]
- no-hit [21; 1.79M; 109k]
- Uroviricota [8; 703k; 98.5k]
- Actinomycetota [2; 84.8k; 44.4k]
- Synergistota [1; 46.2k; 46.2k]
- Arthropoda [1; 43.4k; 43.4k]
- Lentisphaerota [1; 33.4k; 33.4k]

hap1. 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	HiFi	HiC
Coverage	38.29X	48.80X

Assembly pipeline

- **Hifiasm**
    - |_ *ver:* Galaxy Version 0.24.0+galaxy0
    - |_ *key param:* VGP Galaxy workflow VGP3 v 0.2.8
- **purge_dups**
    - |_ *ver:* Galaxy Version 1.2.6+galaxy0
    - |_ *key param:* VGP Galaxy workflow VGP6 v 0.5
- **YaHS**
    - |_ *ver:* Galaxy Version 1.2a.2+galaxy2
    - |_ *key param:* VGP Galaxy workflow VGP8 v 1.4

Curation pipeline

- **PretextMap**
    - |_ *ver:* Galaxy Version 0.1.9+galaxy1
    - |_ *key param:* VGP workflow VGP PretextMap generation v2
- **PretextView**
    - |_ *ver:* 1.0.0
    - |_ *key param:* NA

Submitter: Marcella Sozzoni
Affiliation: UNIFI

Date and time: 2025-12-11 10:09:59 CET