ERGA Assembly Report

v24.02.07_beta

Tags: ERGA-BGE

ToLID	**mRupRup1**
Species	**Rupicapra rupicapra**
Class	Mammalia
Order	Artiodactyla

Genome Traits	Expected	Observed
Haploid size (bp)	2,693,478,497	2,622,988,171
Haploid Number	29 (source: direct)	29
Ploidy	2 (source: direct)	2
Sample Sex	XY	XY

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 8.8.Q55

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

Curator notes

. Interventions/Gb: 2
. Contamination notes: "found some possible gut bacteria -Veillonellaceae- fragments in three small size scaffolds (scaffold_52, scaffold_55 and scaffold_58) which were subsequently removed from the assembly"
. Other observations: "Yahs did a good initial job. We found a PAR shared by SUPER_X and SUPER_Y and left it only attached to Chr_X. The HiC map above includes the three contaminated scaffolds."

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	2,623,465,591	2,622,988,171
GC %	41.99	41.98
Gaps/Gbp	17.15	18.3
Total gap bp	9,000	9,600
Scaffolds	80	75
Scaffold N50	100,920,001	100,920,001
Scaffold L50	10	10
Scaffold L90	24	24
Contigs	125	123
Contig N50	77,662,214	77,662,214
Contig L50	13	13
Contig L90	36	36
QV	55.4348	55.4492
Kmer compl.	97.0448	97.0442
BUSCO sing.	93.6%	93.6%
BUSCO dupl.	4.1%	4.1%
BUSCO frag.	1.1%	1.1%
BUSCO miss.	1.2%	1.2%

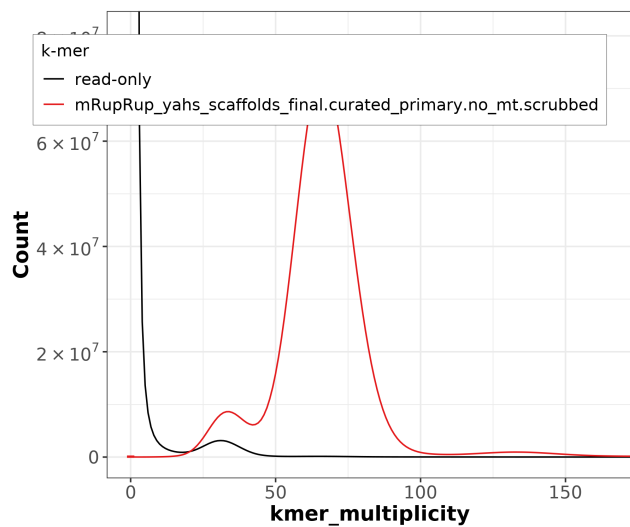BUSCO 5.4.0 Lineage: vertebrata_odb10 (genomes:67, BUSCOs:3354)
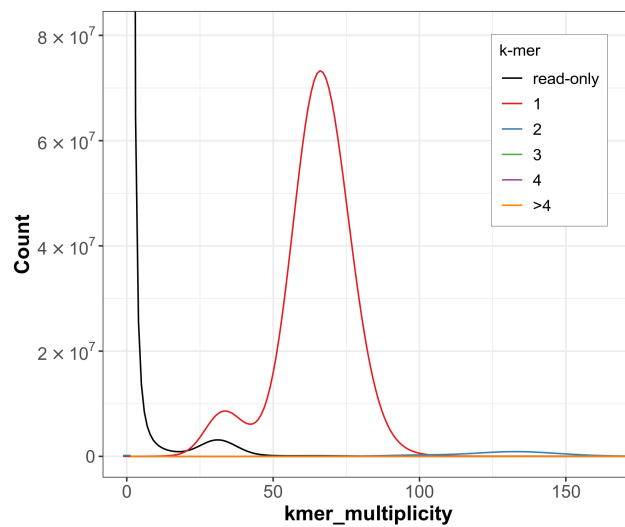
HiC contact map of curated assembly
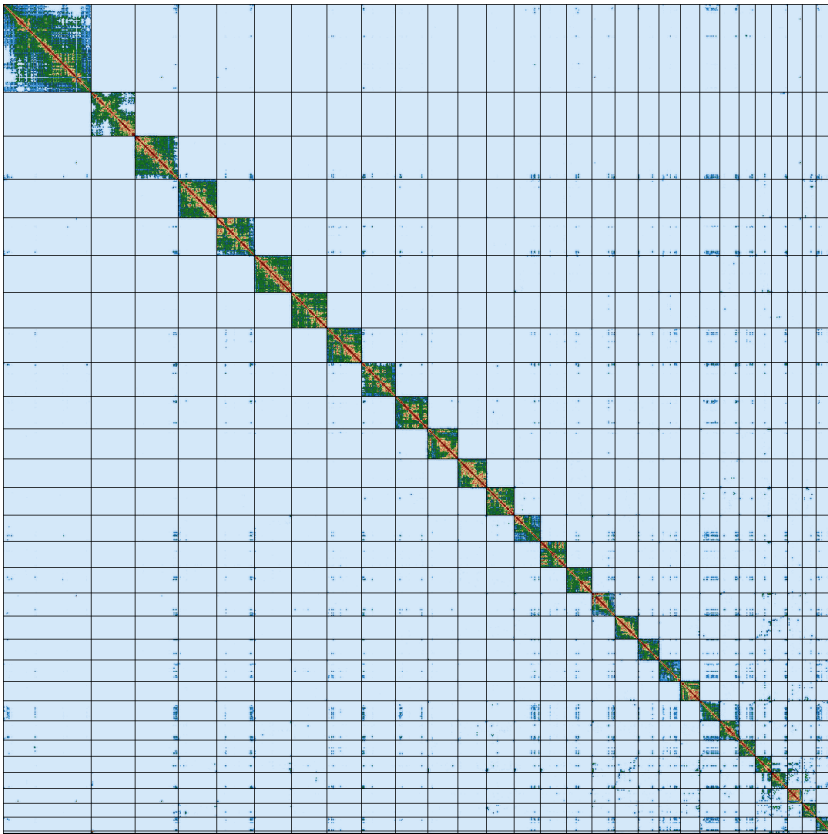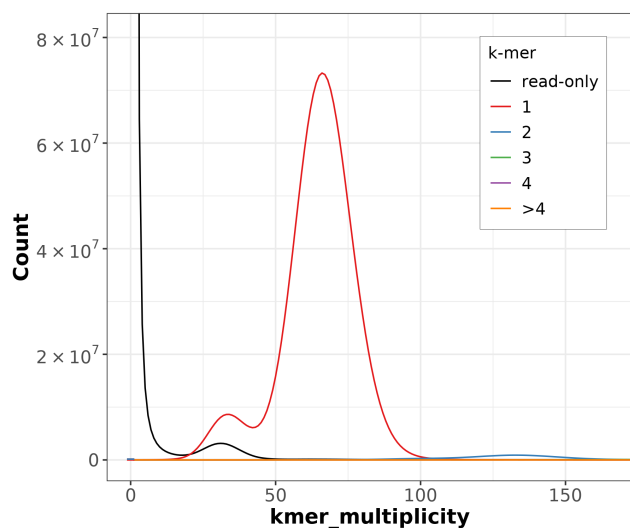

collapsed [LINK]

K-mer spectra of curated assembly

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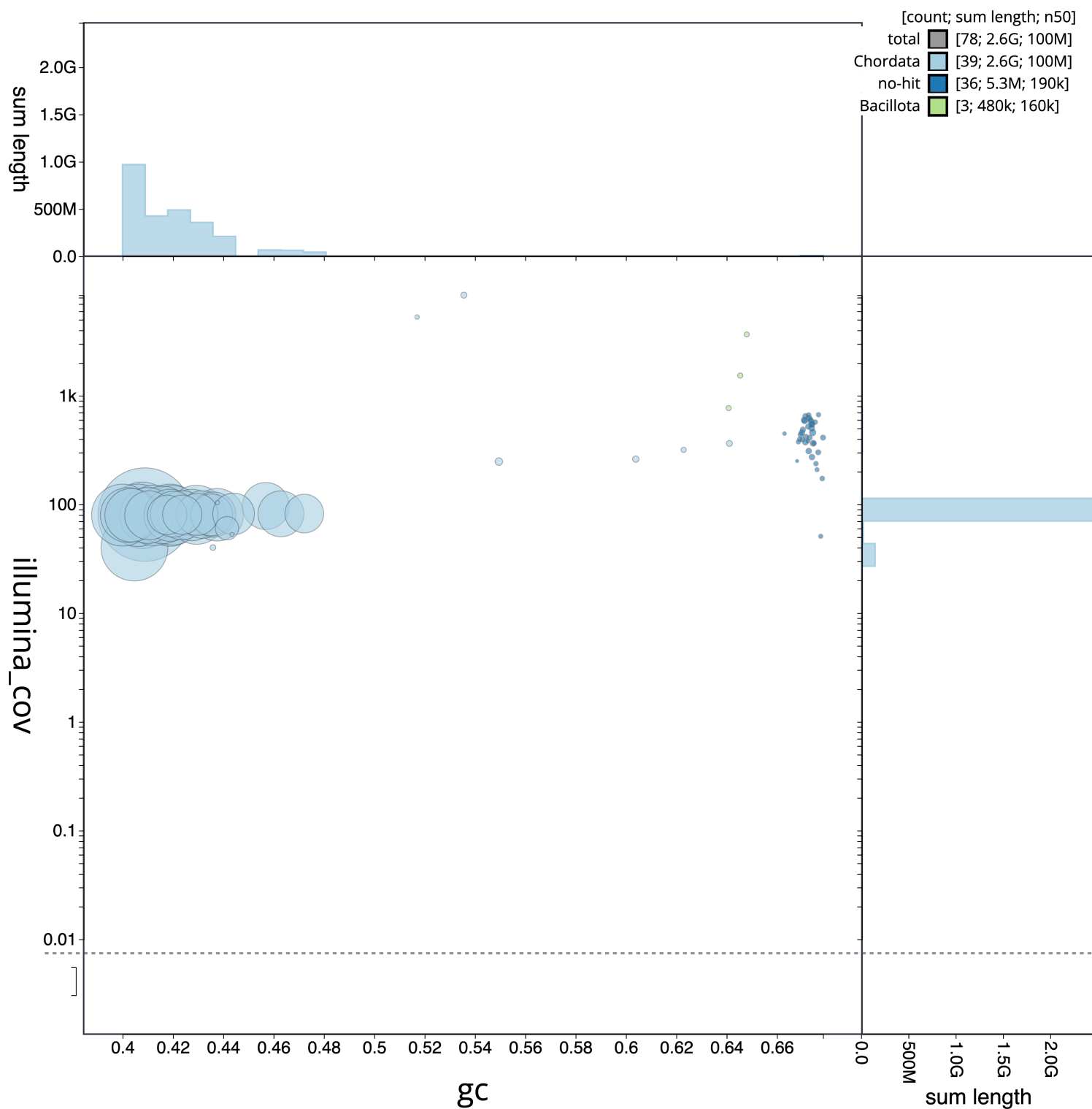their presence in reads/assemblies

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

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

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	ONT	Illumina	OmniC
Coverage	125x	84x	90x

Assembly pipeline

- **Trim_Galore**
  - |_ *ver:* 0.6.7
  - |_ *key param:* "--gzip -q 20"
  - |_ *key param:* "--paired"
  - |_ *key param:* "--retain_unpaired"
- **Filtlong**
  - |_ *ver:* 0.2.1
  - |_ *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

Submitter: Francisco Camara
Affiliation: CNAG Barcelona

Date and time: 2024-02-07 15:39:55 CET