ERGA Assembly Report
v24.09.10

Tags: ERGA-BGE

TxID	166086
ToLID	**rHemHip1**
Species	Hemorrhois hippocrepis
Class	Lepidosauria
Order	Squamata

Genome Traits	Expected	Observed
Haploid size (bp)	1,677,182,640	1,810,544,746
Haploid Number	18 (source: ancestor)	18
Ploidy	3 (source: ancestor)	2
Sample Sex	ZW	ZW

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.8.Q50

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

. Observed Ploidy is different from Expected

Curator notes

. Interventions/Gb: 5
. Contamination notes: "No contaminants detected with blobtoolkit INSDC pipeline in the precurated assembly."
. Other observations: "Curation made 4 cuts in contigs, 5 breaks at gaps and 0 joins. The W contigs were found in the purge_dups' hap.fa and placed back into the nextdenovo asseembly before yahs scaffolding. The W was assembled into 1 contig of 11.9 Mb and 33 unlocalized scaffolds totalling 96,233,673 bp. Consistently with the cytogenetic information we identified 10 michrochromosomes (SUPER_7-SUPER_16). Although there are 4 scaffolds with clear contcats with them is not possible to place them into a unique SUPER. The link provided for the pretextmap points to a folder where reviewers can add additional save_states if required. Finally, the mitogenome has been assembled independently with FOAM into a 21,416 bp circular contig with high accuracy."

Quality metrics table

| Metrics	Pre-curation collapsed	Curated collapsed
Total bp	1,810,545,746	1,810,544,746
GC %	41.21	41.21
Gaps/Gbp	24.3	21.54
Total gap bp	8,800	7,800
Scaffolds	112	121
Scaffold N50	153,613,959	153,613,959
Scaffold L50	4	4
Scaffold L90	11	14
Contigs	156	160
Contig N50	47,232,710	47,232,710
Contig L50	12	12
Contig L90	40	40
QV	50.2312	50.2312
Kmer compl.	95.6393	95.6393
BUSCO sing.	96.5%	96.5%
BUSCO dupl.	1.0%	1.0%
BUSCO frag.	0.7%	0.7%
BUSCO miss.	1.8%	1.8%

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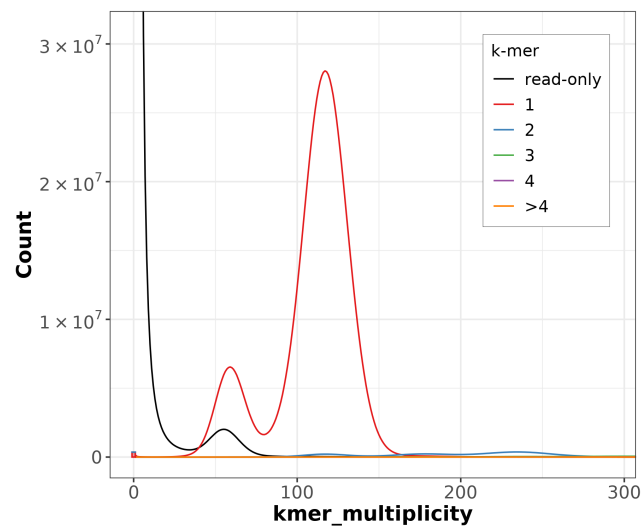
HiC contact map of curated assembly
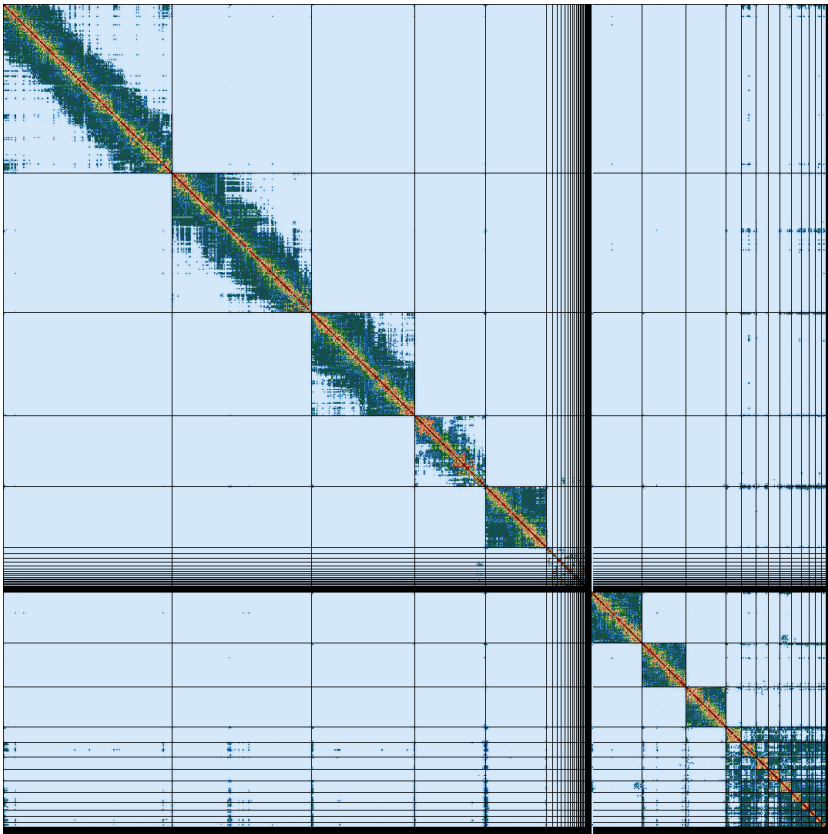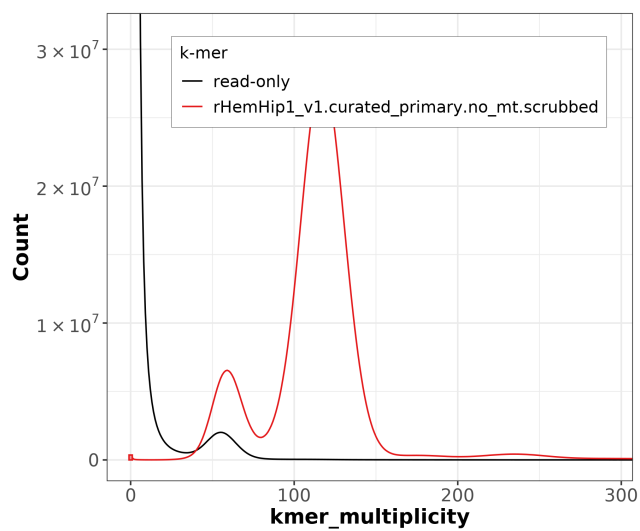

collapsed [LINK]

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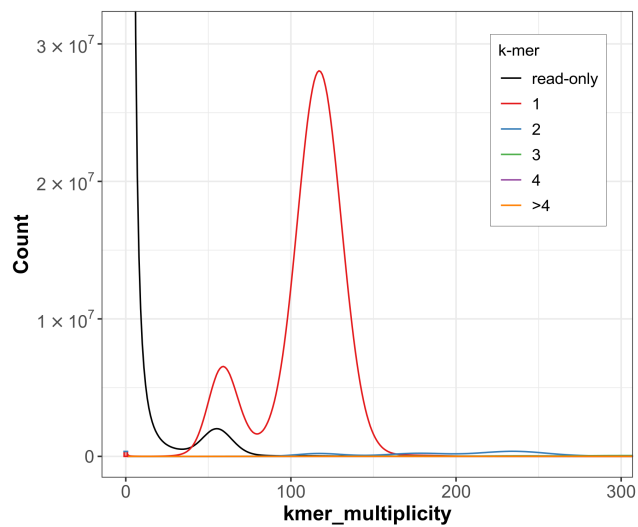
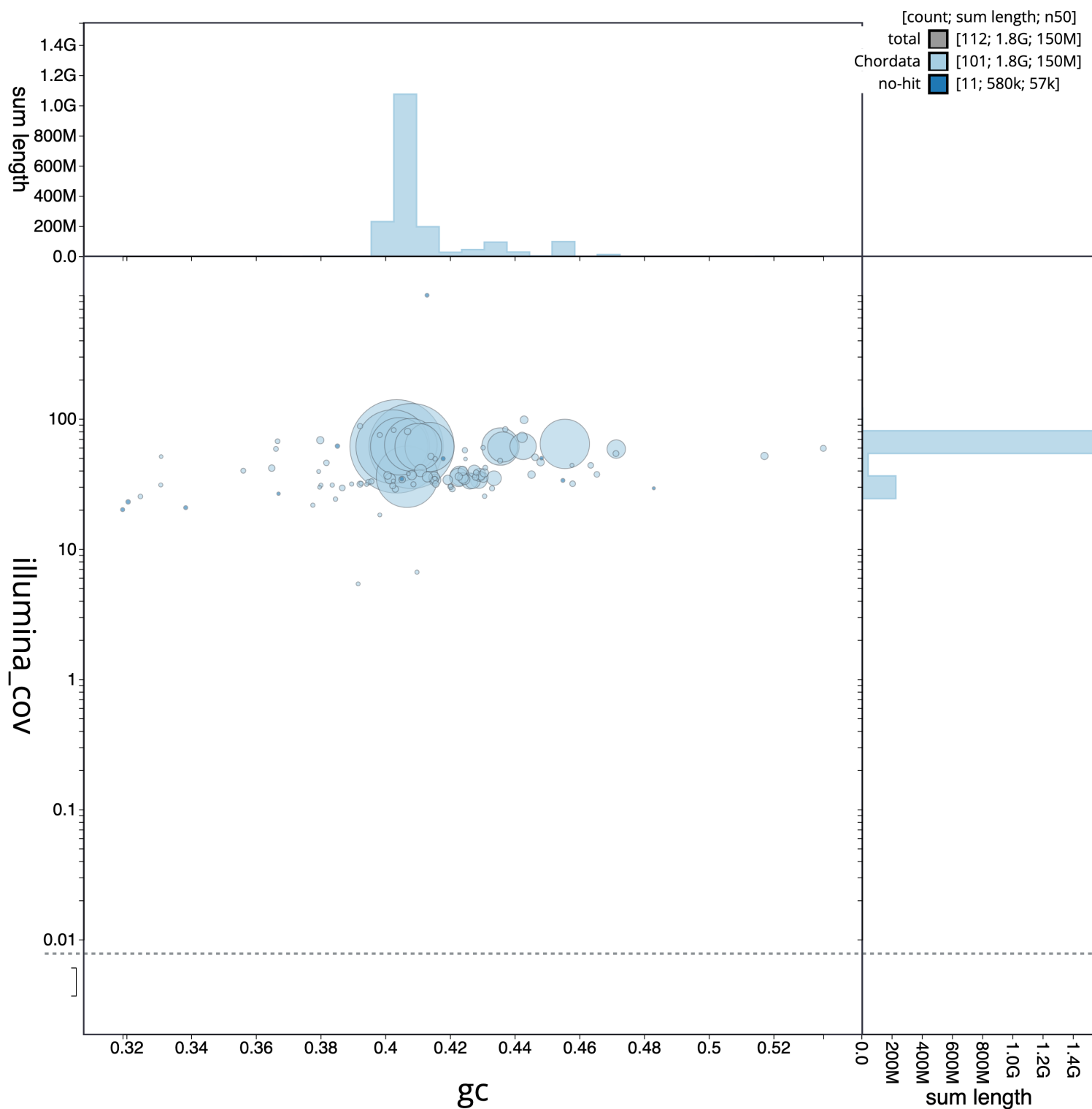Distribution of k-mer counts per copy
numbers found in asm

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

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	84x	60x	53x

Assembly pipeline

- **CLAWS pipeline**
  |_ *ver:* 2.2.0
  |_ *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
- **Blobtoolkit INSDC pipeline**
  |_ *ver:* 1.0
  |_ *key param:* NA
- **FOAM pipeline**
  |_ *ver:* 0.5
  |_ *key param:* NA
- **mitos**
  |_ *ver:* 2.1.3
  |_ *key param:* NA

Curation pipeline

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

Submitter: Fernando Cruz

Affiliation: CNAG Barcelona

Date and time: 2024-11-04 15:37:41 CET