ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	101491
ToLID	**rTriTgu1**
Species	Trionyx triunguis
Class	NA
Order	Testudines

Genome Traits	Expected	Observed
Haploid size (bp)	2,172,236,689	2,320,326,796
Haploid Number	33 (source: ancestor)	33
Ploidy	2 (source: ancestor)	2
Sample Sex	ZZ	ZZ

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 7.8.Q70

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

Curator notes

. Interventions/Gb: 0
. Contamination notes: "No contaminations could be detected"
. Other observations: "Very good assembly. Indeed purge_dups and Yahs created a worse assembly as is truncated telomeres and the scaffolding cut off pieces. Most of the hifiasm contig assembly contained already full chromosomes with telomere motifs at both ends. Therefore I used the hifiasm assembly and scaffolded it manually with PretextView and the PixelSort option. Haplotigs were removed based on coverage and the purge_dups annotation. I could not identify the Z chromosome, as I could not find a turtle assembly with known sex chromosomes."

Quality metrics table

| Metrics	Pre-curation pri	Curated pri
Total bp	2,330,438,147	2,320,326,796
GC %	45.83	45.8
Gaps/Gbp	0	28.88
Total gap bp	0	13,400
Scaffolds	138	43
Scaffold N50	80,362,655	137,054,899
Scaffold L50	6	6
Scaffold L90	29	23
Contigs	138	110
Contig N50	80,362,655	80,362,655
Contig L50	6	6
Contig L90	29	29
QV	70.2438	70.2902
Kmer compl.	98.0915	98.0429
BUSCO sing.	97.2%	97.2%
BUSCO dupl.	1.2%	1.2%
BUSCO frag.	0.8%	0.8%
BUSCO miss.	0.8%	0.8%

BUSCO: 5.8.3 (euk_genome_min, miniprot) / Lineage: sauropsida_odb12 (genomes:79, BUSCOs:6118)

HiC contact map of curated assembly
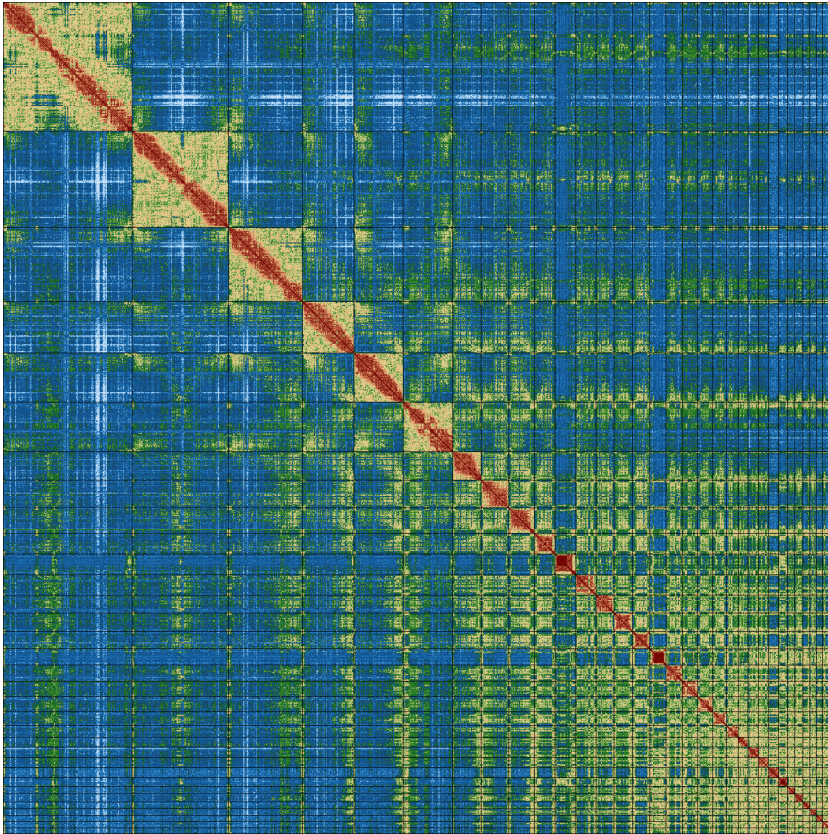
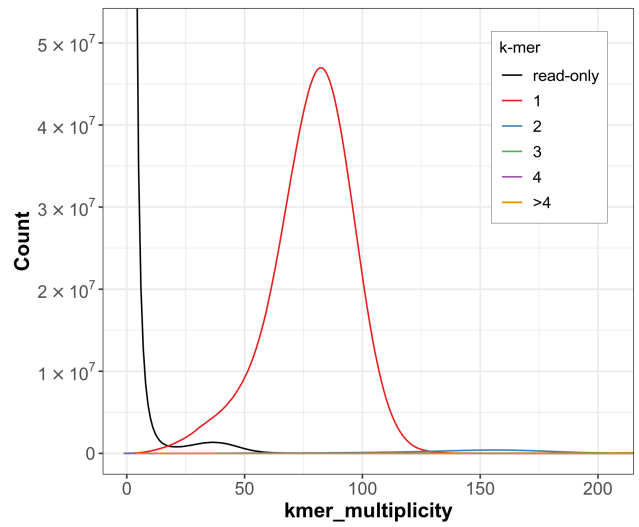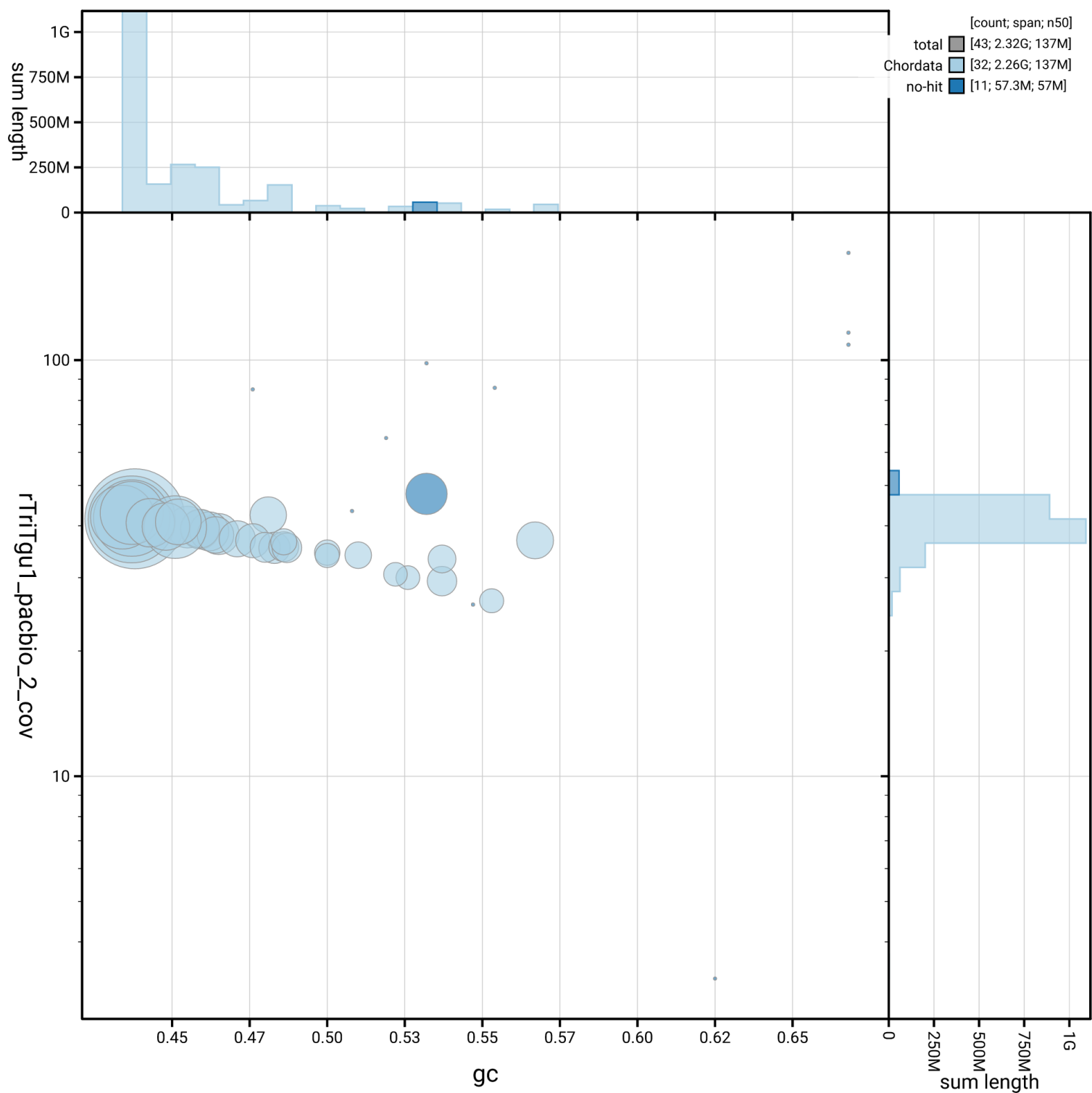**pri** [LINK]

K-mer spectra of curated assembly
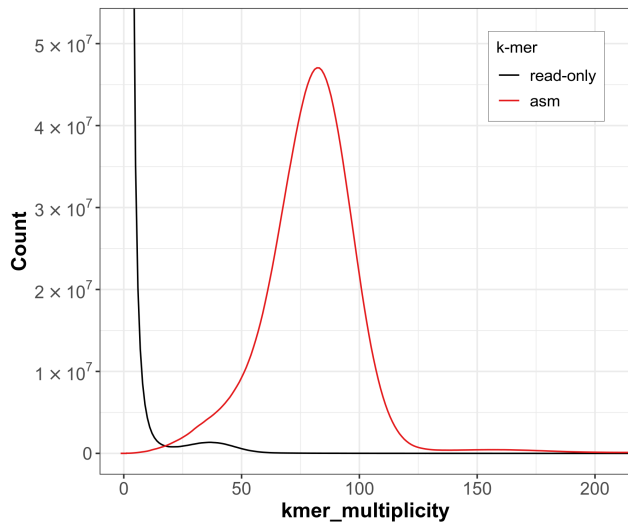
Distribution of k-mer counts coloured by their presence in reads/assemblies

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

Post-curation contamination screening

pri. 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	80x	118x

Assembly pipeline

- **Hifiasm**
  |_ *ver:* 0.25.0-r726
  |_ *key param:* l3
- **purge_dups**
  |_ *ver:* 1.2.6
  |_ *key param:* NA

Curation pipeline

- **GRIT_Rapid**
  |_ *ver:* 1a3d79a8
  |_ *key param:* NA
- **HiGlass**
  |_ *ver:* 0.10.4
  |_ *key param:* NA

Submitter: Martin Pippel
Affiliation: SciLifeLab

Date and time: 2025-10-14 07:54:27 CEST