

ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

TxID	1971494
ToLID	qdTypLobi1
Species	<i>Typhloius lobifer</i>
Class	Diplopoda
Order	Julida

Genome Traits	Expected	Observed
Haploid size (bp)	1,081,278,584	937,744,401
Haploid Number	12 (source: ancestor)	7
Ploidy	2 (source: ancestor)	2
Sample Sex	U	U

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q64

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

- . Observed Haploid Number is different from Expected
- . Assembly length loss > 3% for pri
- . More than 1000 gaps/Gbp for pri
- . Not 90% of assembly in chromosomes for pri

Curator notes

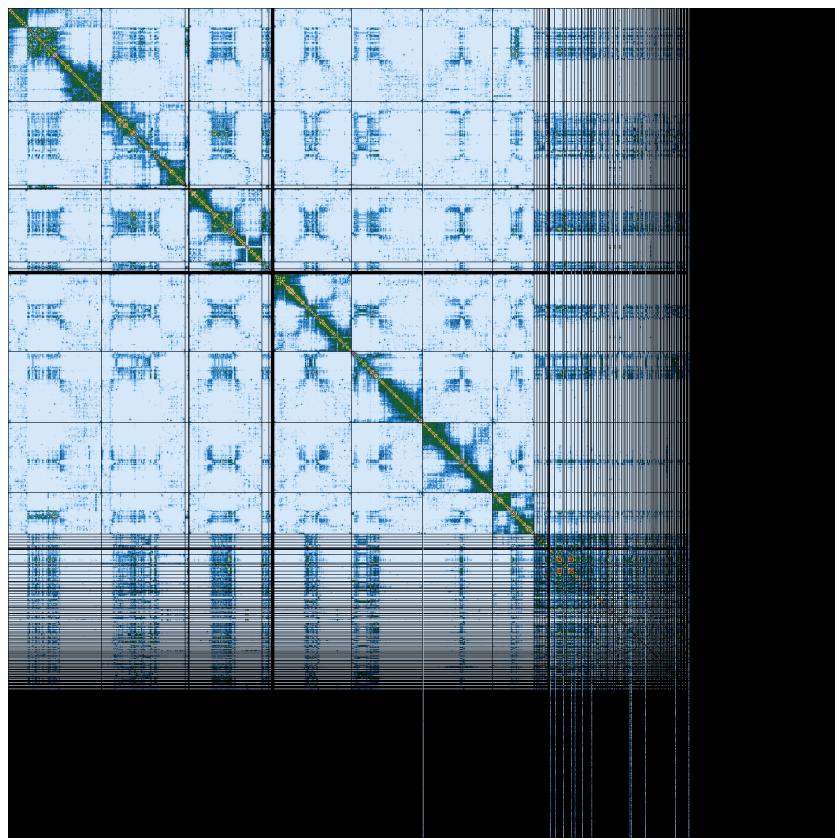
- . Interventions/Gb: 322
- . Contamination notes: "Contamination report for assembly labelled hap1; Total length of scaffolds removed: 797,121 (0.1 %); Scaffolds removed: 37 (1.6 %); Largest scaffold removed: (71,029); FCS-GX contaminant species (number of scaffolds; total length of scaffolds): Mitochondrion (37; 797,121)"
- . Other observations: "Assembly was Hi-C phased; original contigs were filtered from bacterial contamination based on tiara results; for hap1 there were 8 rounds of yahs with read remapping; for hap2 there were 3 rounds of yahs with read remapping; The order and orientation of scaffolds within the following regions is uncertain: Chromosome 1 between approximately 20.97-76.37Mb; Chromosome 2 between approximately 24.36-58.04Mb; Chromosome 3 between approximately 27.29-50.90Mb; Chromosome 4 between approximately 38.17-49.74Mb; Chromosome 5 from the start to approximately 32.83Mb; Chromosome 6 between approximately 42.17-47.45Mb; Chromosome 7 between approximately 20.53-29.70Mb; Inversions between haplotypes were observed in the following locations: Chromosome 1 between approximately 98.76-104.39Mb; Chromosome 2 between approximately 70.2-100.2Mb; Chromosome 3 between approximately 12.06-26.96Mb"

Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	1,430,896,336	937,744,401
GC %	38.12	38.15
Gaps/Gbp	0	1,313.79
Total gap bp	0	137,974
Scaffolds	3,370	2,088
Scaffold N50	1,601,000	78,863,615
Scaffold L50	258	6
Scaffold L90	1,192	383
Contigs	3,370	3,320
Contig N50	1,601,000	1,257,506
Contig L50	258	213
Contig L90	1,192	1,203
QV	64.6	64.6
Kmer compl.	98.37	96.87
BUSCO sing.	96.6%	98.1%
BUSCO dupl.	0.9%	0.7%
BUSCO frag.	1.2%	0.0%
BUSCO miss.	1.3%	1.2%

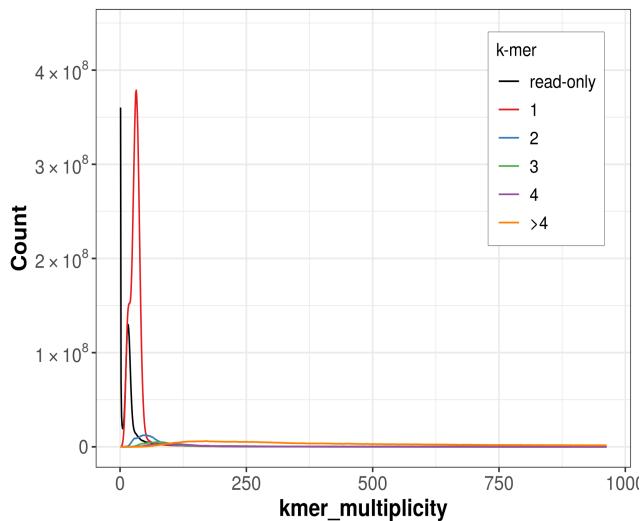
Warning: BUSCO versions or lineage datasets are not the same across results

HiC contact map of curated assembly

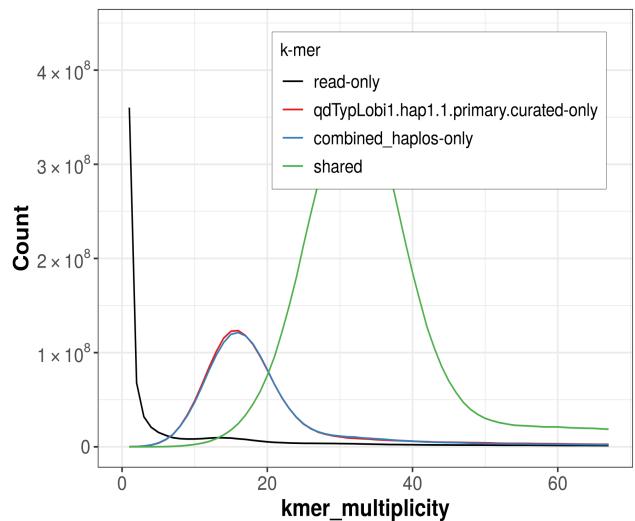


pri [LINK]

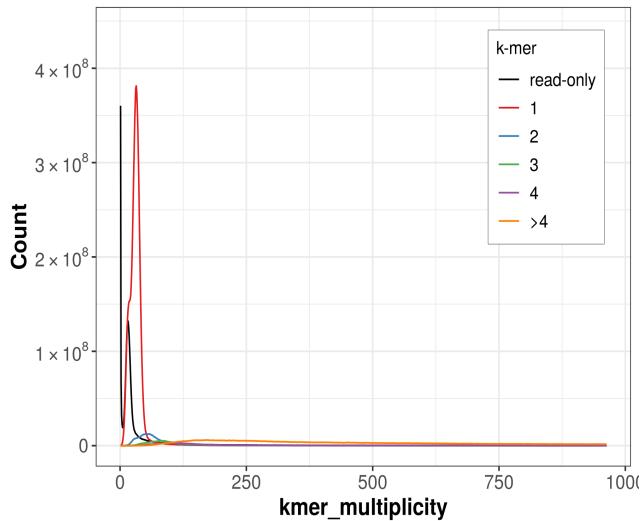
K-mer spectra of curated assembly



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

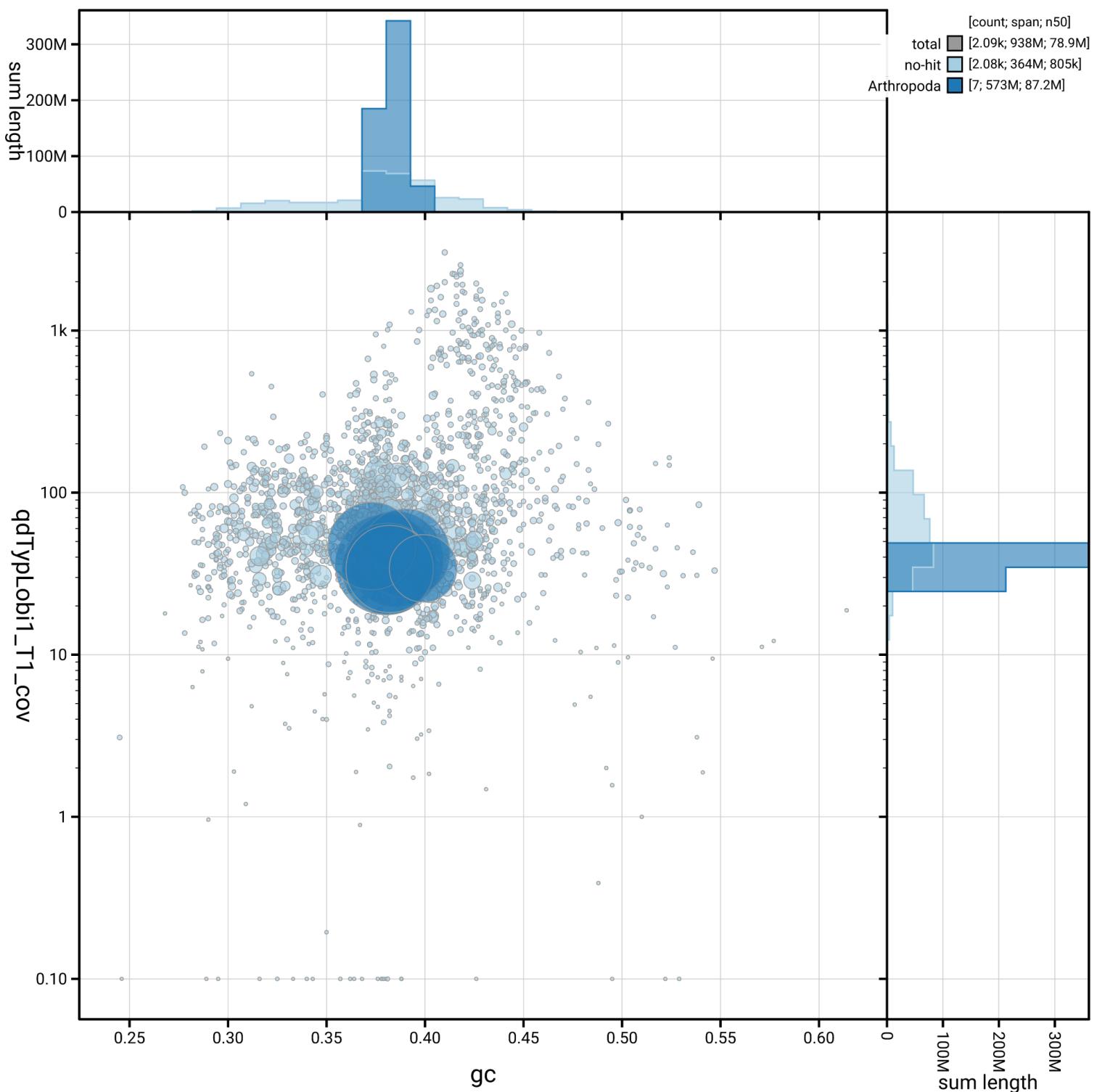


Distribution of k-mer counts coloured by their presence in reads/assemblies



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	PacBio HiFi	Arima v2
Coverage	32x	176x

Assembly pipeline

```
- 0.19.8-r603
  |_ ver: NA
  |_ key param: --h1/--h2
- yahs
  |_ ver: 1.2.2
  |_ key param: NA
```

Curation pipeline

```
- 0.19.8-r603
  |_ ver: NA
  |_ key param: --h1/--h2
- yahs
  |_ ver: 1.2.2
  |_ key param: NA
- TreeVal
  |_ ver: 1.2.2
  |_ key param: NA
```

Submitter: Sarah Pelan

Affiliation: WSI

Date and time: 2026-01-23 14:24:02 CET