

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	3054492
ToLID	ibLobTrog1
Species	Loboptera troglobia
Class	Insecta
Order	Blattodea

Genome Traits	Expected	Observed
Haploid size (bp)	2,370,803,843	2,585,382,996
Haploid Number	17 (source: ancestor)	14
Ploidy	2 (source: ancestor)	2
Sample Sex	X0	X0

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.8.Q43

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

- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for collapsed

Curator notes

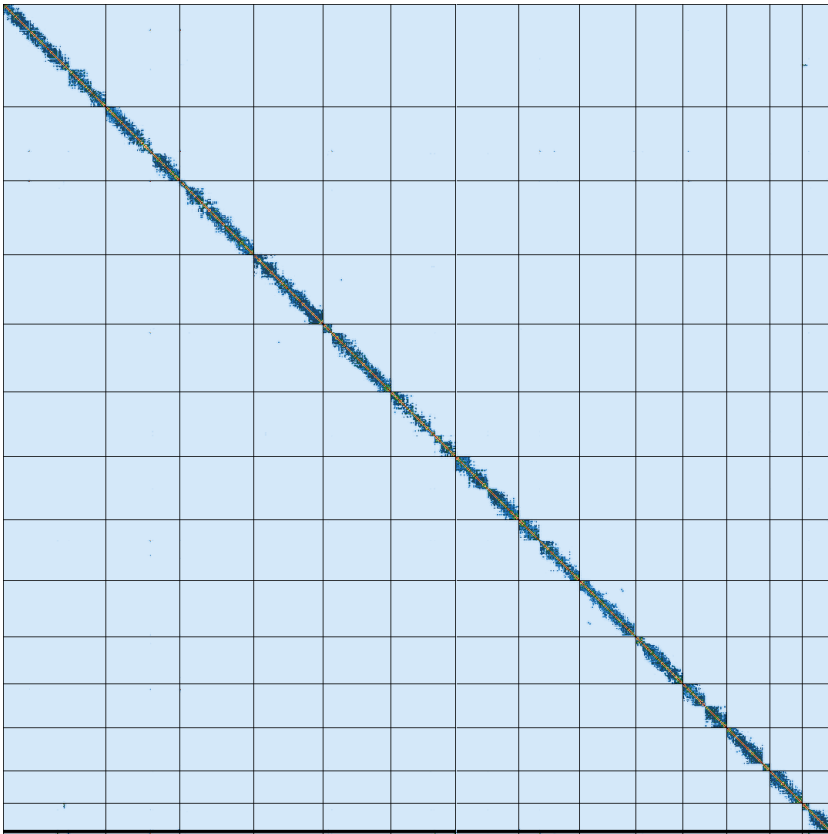
- . Interventions/Gb: 13
- . Contamination notes: "Based on the blobtools report results (using -buscoreregions species-) we removed 26 smaller scaffolds that matched mostly bacterial sequences (largely putative endosymbiont fragments). Some of the scaffolds also contained viral and fungal sequences. The scaffolds in question are: scaffold_36,scaffold_40,scaffold_42,scaffold_45,scaffold_48,scaffold_49,scaffold_50,scaffold_52,scaffold_53,scaffold_57,scaffold_59,scaffold_61,scaffold_64,scaffold_67,scaffold_68,scaffold_69,scaffold_70,scaffold_71,scaffold_74,scaffold_78,scaffold_81,scaffold_83,scaffold_87,scaffold_91,scaffold_93 and scaffold_109. These scaffolds were not removed from the pretext file but we will do so once the review process is complete. They have, however, already been removed from the blobtools plot and busco, merquary and gfastas included in the report"
- . Other observations: "For this specimen the ONT data derived from the heteromorphic sex with the absense of a second sex chromosome (male X0). The number of putative chromosomes seems to match that of another Loboptera genus cockroach (Loboptera subterranea). For a number of the super scaffolds there seem to be weak contacts between the the two arms of the chromosomes, something we have observed in other Loboptera genus cockroaches as well"

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	2,589,487,305	2,585,382,996
GC %	34.21	34.19
Gaps/Gbp	254.1	261.86
Total gap bp	131,600	135,400
Scaffolds	113	69
Scaffold N50	146,328,501	199,875,071
Scaffold L50	7	6
Scaffold L90	17	12
Contigs	771	746
Contig N50	9,620,444	9,620,444
Contig L50	80	80
Contig L90	277	276
QV	43.9298	43.9559
Kmer compl.	84.0456	83.8892
BUSCO sing.	97.0%	97.0%
BUSCO dupl.	1.5%	1.5%
BUSCO frag.	0.6%	0.6%
BUSCO miss.	0.9%	0.9%

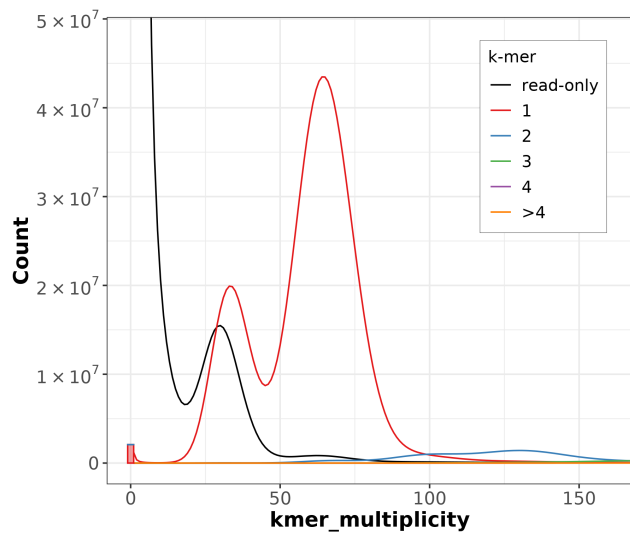
BUSCO: 5.4.0 (euk_genome_met, metaeuk) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

HiC contact map of curated assembly

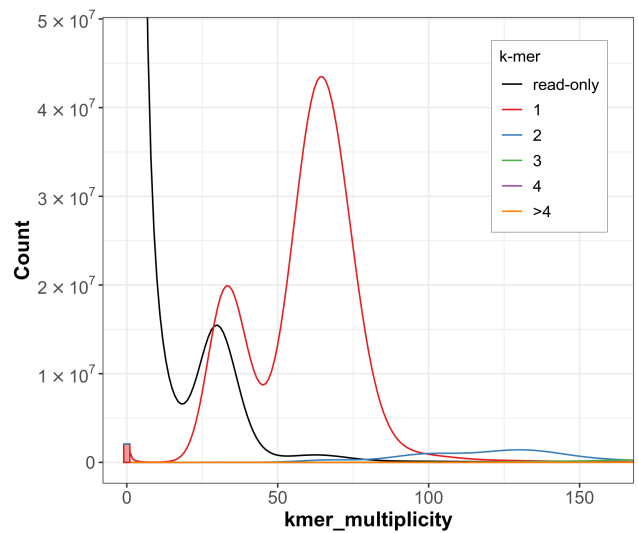


collapsed [\[LINK\]](#)

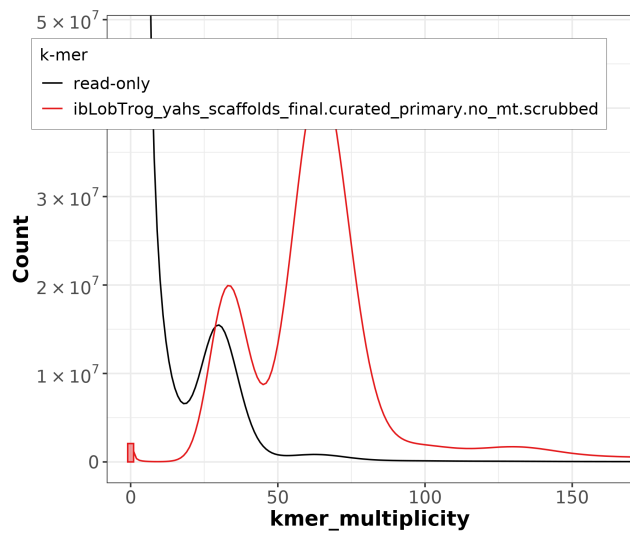
K-mer spectra of curated assembly



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

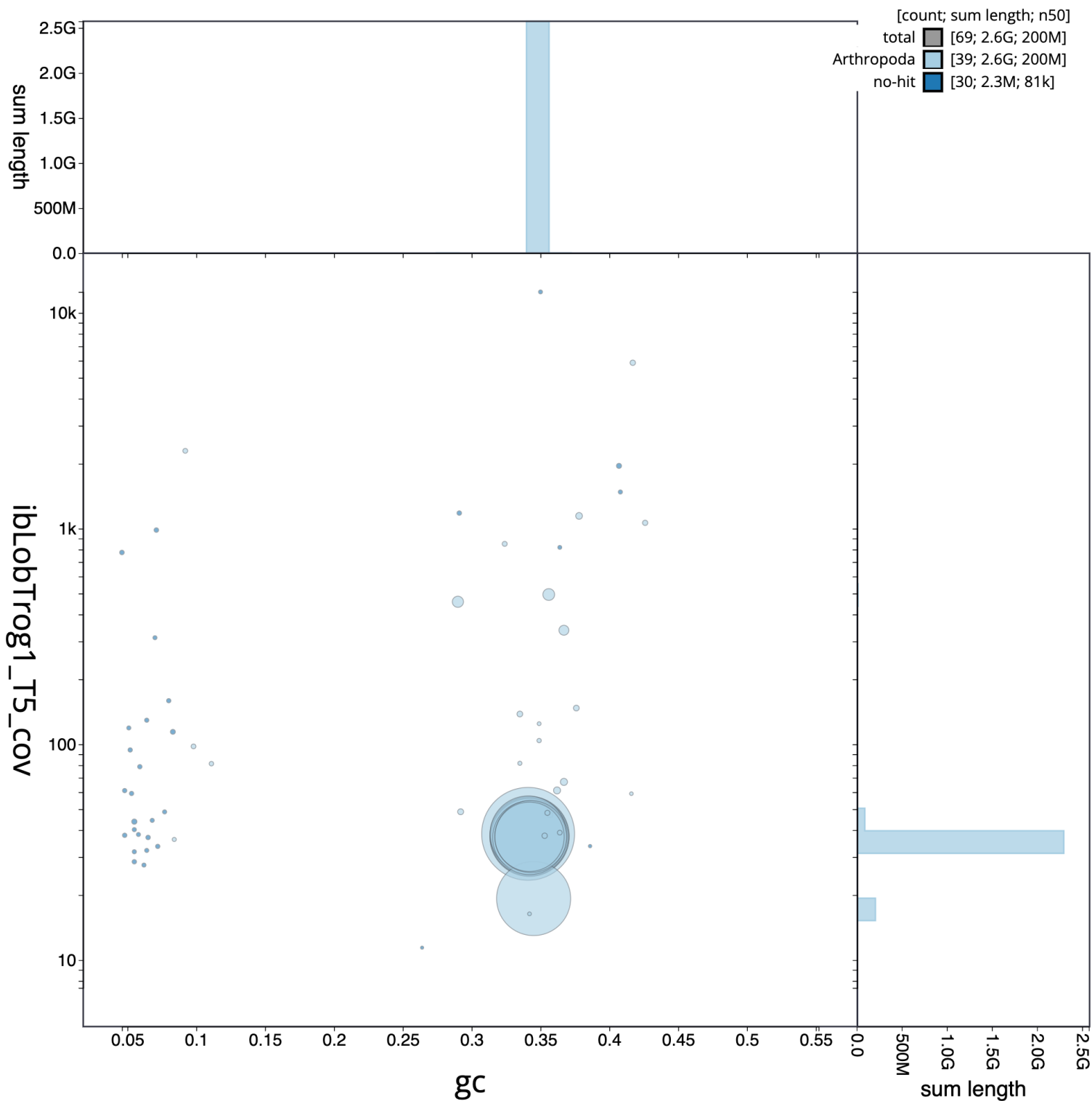


Distribution of k-mer counts per copy numbers found in asm



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	40x	49x	121x

Assembly pipeline

- **Trim_galore**
 - |_ *ver*: 0.6.7
 - |_ *key param*: --gzip
 - |_ *key param*: =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
- **GRIT_Rapid**
 - |_ *ver*: 2.0
 - |_ *key param*: NA

Submitter: Francisco Camara

Affiliation: CNAG Barcelona

Date and time: 2024-12-12 11:39:31 CET