

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	3082394
ToLID	qqLeiRupe1
Species	Leiobunum rupestre
Class	Arachnida
Order	Opiliones

Genome Traits	Expected	Observed
Haploid size (bp)	792,557,719	788,741,017
Haploid Number	10 (source: ancestor)	11
Ploidy	2 (source: ancestor)	2
Sample Sex	unknown	unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.7.Q60

Obtained EBP quality metric for hap2: 7.7.Q59

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 hap2

Curator notes

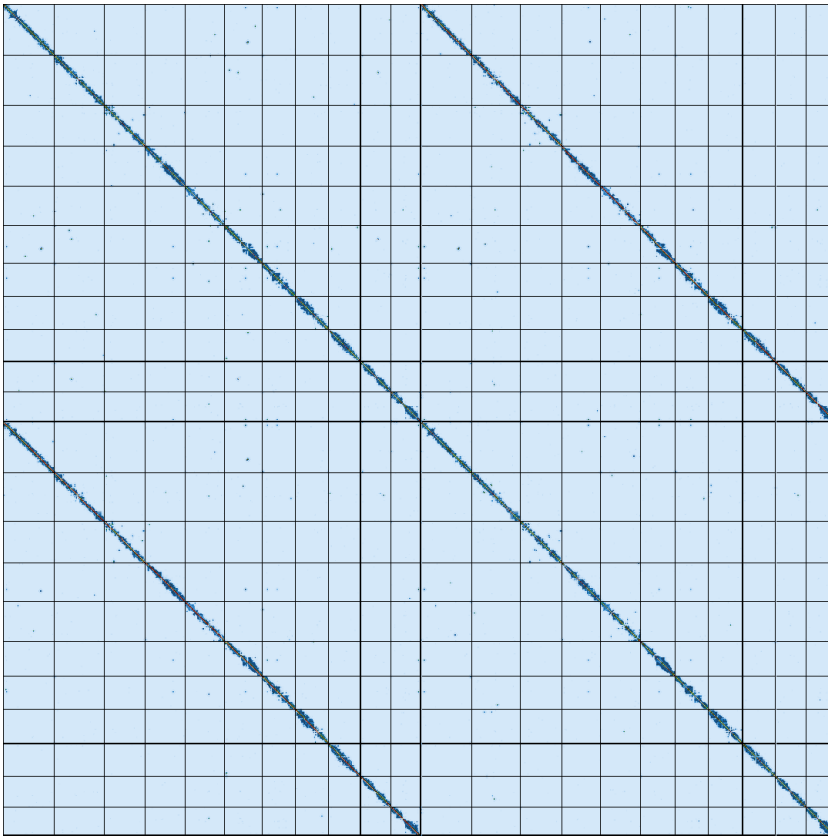
- . Interventions/Gb: 36
- . Contamination notes: "FCS-GX found 1 bacterial contaminant in hap1 that was removed during curation. The curated diploid map is inside a shared folder. Please load the savestate_1 for the normal resolution pretext (without .HR.), before you start the review. In it you will see the SUPERS for each HAP side by side. Note that we just caught an extra FalseDup inside H1_SUPER_6 that actually is duplicated in H1_SUPER_1. Also there are signals of an inversion in SUPER_3 but we think is well contigged with the ONT. As the Hi-C comes from a different individual (qqLeiRupe9) we suspect that is due to some structural variant in a third haplotype found in that individual."
- . Other observations: "Sequencing was done by UNIBA, two different specimens were used: for ONT (qqLeiRupe1) and for Hi-C (qqLeiRupe9). Data was transfered to CNAG, where we assembled the genome. Our hifiiasm assembly produced 2 haplotypes that were curated jointly. We assembled them into 11 super-scaffolds what is consitent with the karyotype of the species (<https://www.doi.org/10.14411/eje.2020.032>). The sex chromosomes, likely homomorphic, were not identified."

Quality metrics table

Metrics	Pre-curation hap1	Pre-curation hap2	Curated hap1	Curated hap2
Total bp	806,831,472	820,196,170	788,741,017	781,314,626
GC %	36.33	36.29	36.32	36.3
Gaps/Gbp	109.07	106.07	117.91	121.59
Total gap bp	17,600	17,400	18,600	19,000
Scaffolds	37	27	31	21
Scaffold N50	66,210,421	63,372,409	74,207,850	73,188,438
Scaffold L50	6	5	5	5
Scaffold L90	11	11	10	10
Contigs	125	114	124	116
Contig N50	13,213,833	13,739,694	12,592,226	12,857,947
Contig L50	18	23	19	23
Contig L90	64	61	64	62
QV	61.0992	59.5274	60.8851	59.7151
Kmer compl.	90.6565	90.8925	90.4461	90.1649
BUSCO sing.	94.2%	94.6%	95.7%	96.8%
BUSCO dupl.	2.9%	2.2%	0.7%	0.2%
BUSCO frag.	0.7%	0.7%	0.7%	0.7%
BUSCO miss.	2.2%	2.5%	2.9%	2.2%

BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: arthropoda_odb12 (genomes:76, BUSCOs:1667)

HiC contact map of curated assembly

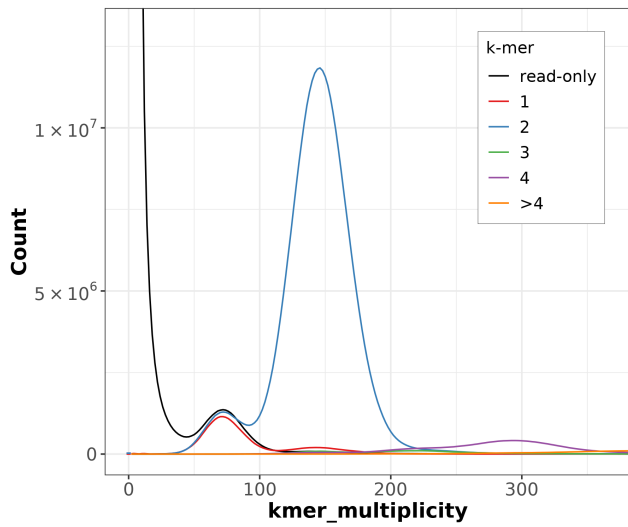


hap1 [\[LINK\]](#)

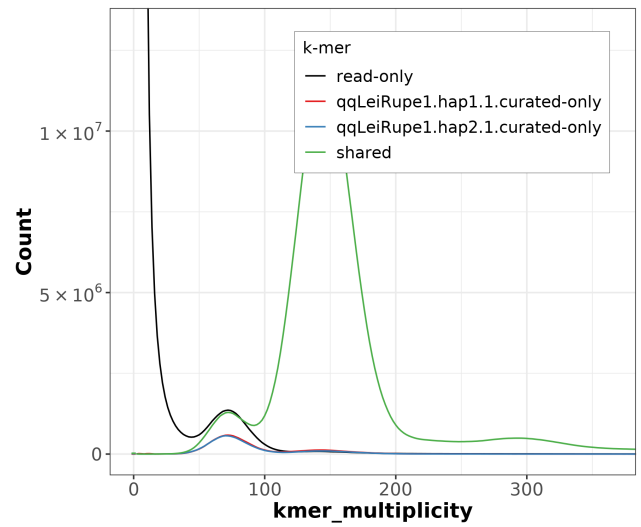
hap2 HiC PNG is missing!

hap2 File link is missing!

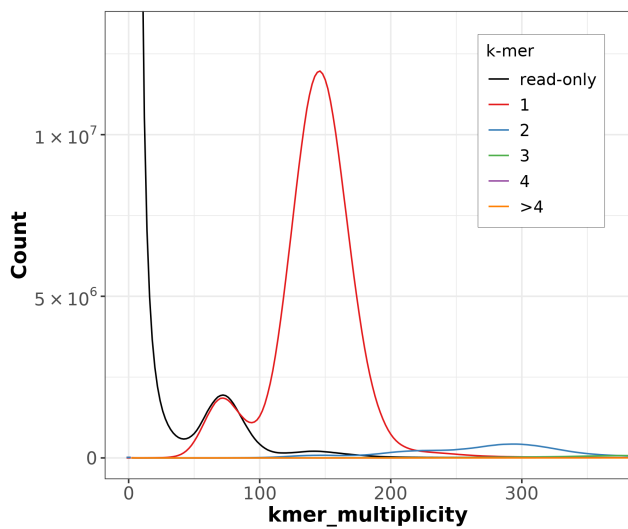
K-mer spectra of curated assembly



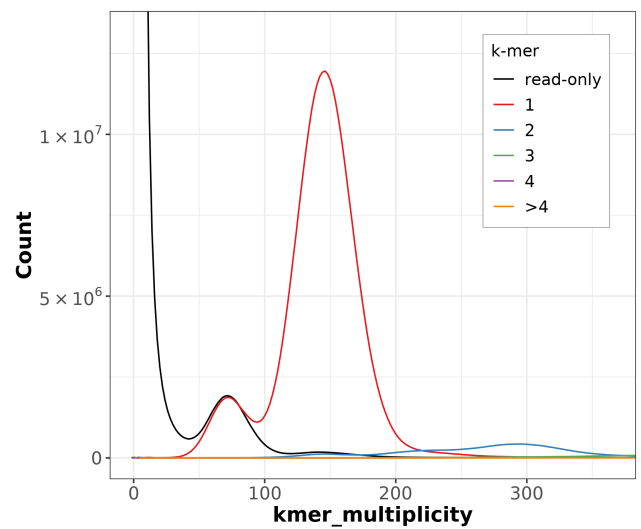
Distribution of k-mer counts per copy numbers found in asm (dipl.)



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

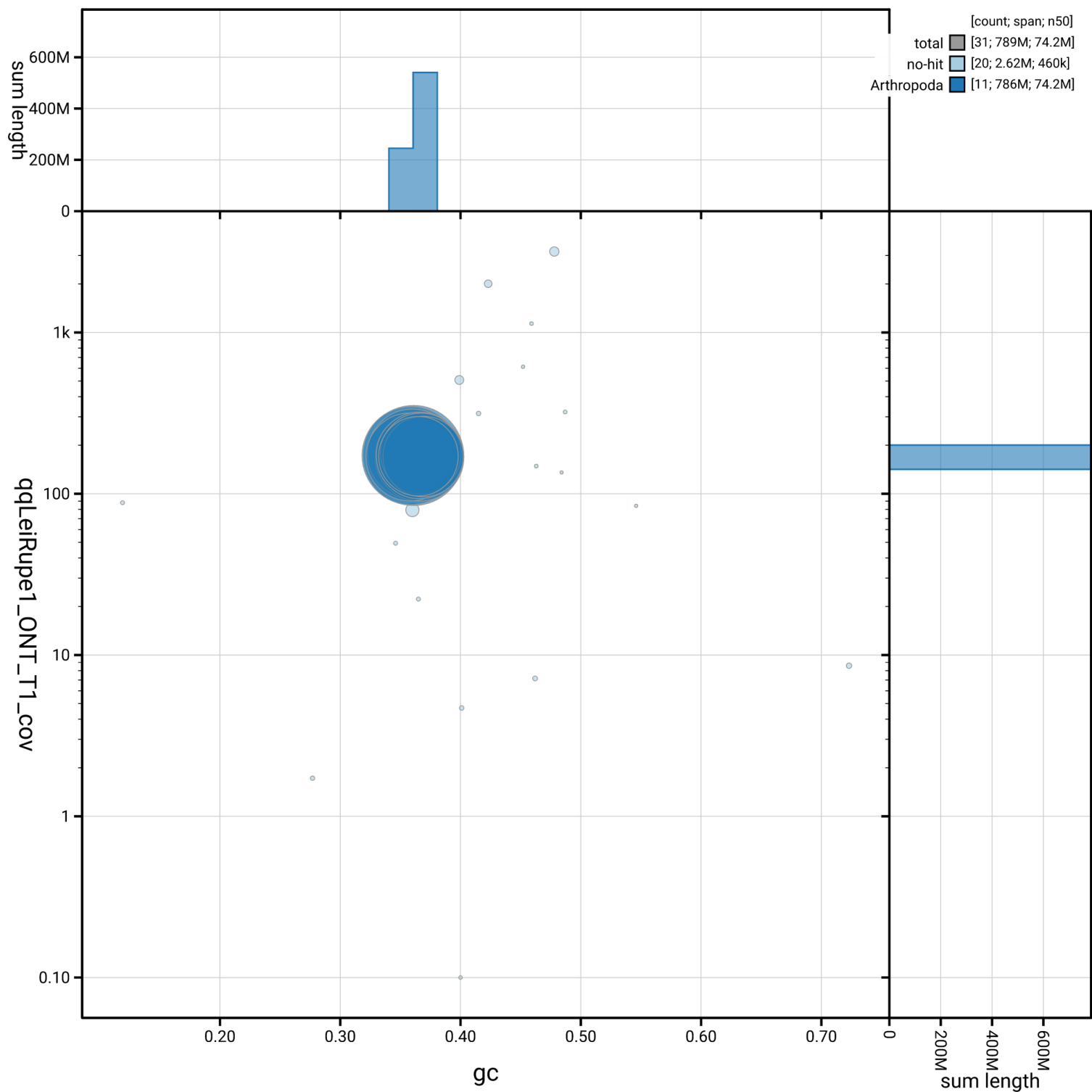


Distribution of k-mer counts per copy numbers found in **hap2** (hapl.)

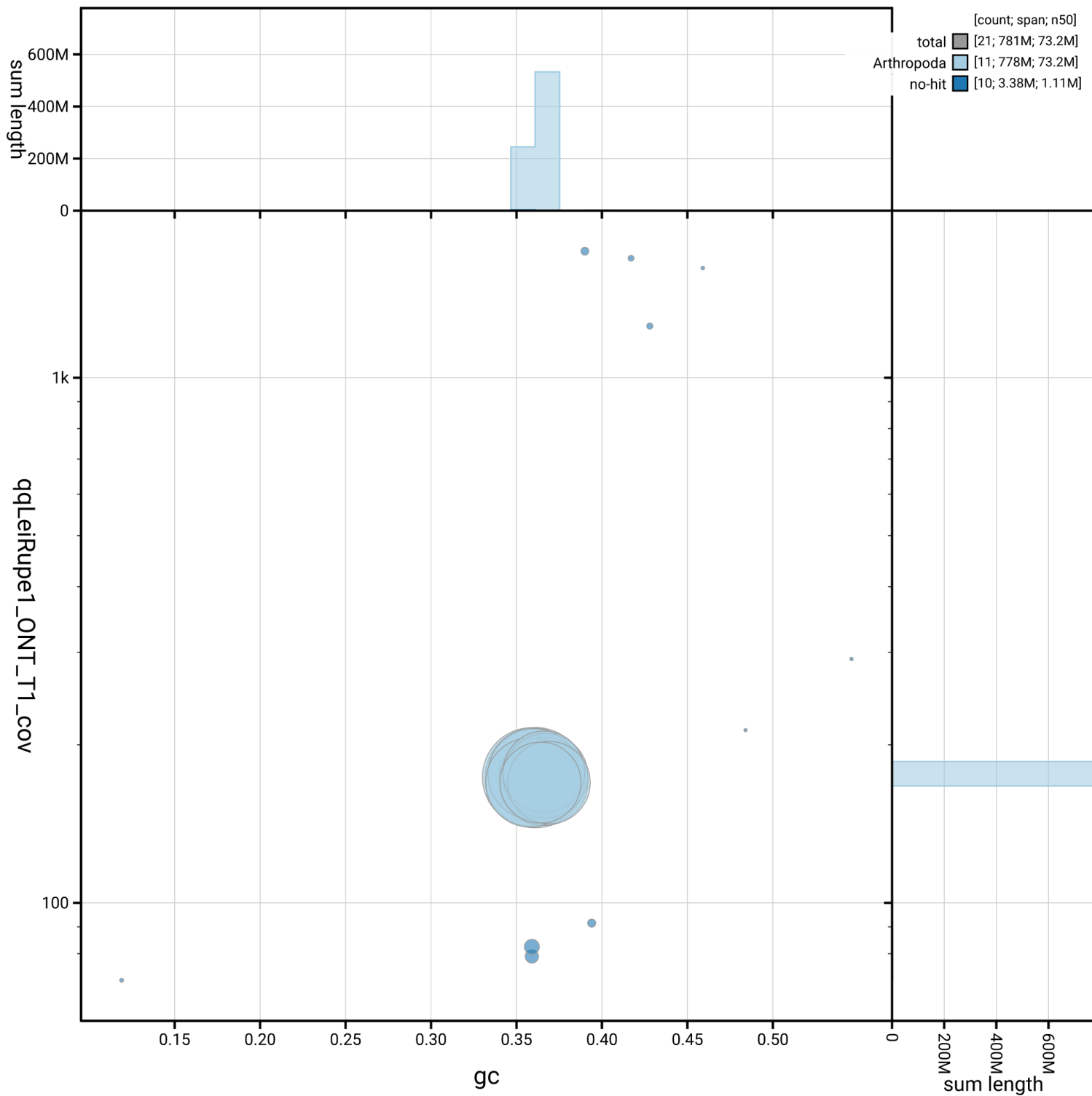


Distribution of k-mer counts per copy numbers found in **hap1** (hapl.)

Post-curation contamination screening



hap1. 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.



hap2. 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	Hi-C
Coverage	323x	53x

Assembly pipeline

- **CLAWS pipeline**
 - |_ ver: 3.1
 - |_ key param: NA
- **Filtlong**
 - |_ ver: 0.2.1
 - |_ key param: --min_length 1000
 - |_ key param: --min_mean_q 80
 - |_ key param: --target_bases 96000000000
- **Hifiasm**
 - |_ ver: 0.24.0
 - |_ key param: --ont
 - |_ key param: --telo-m TTAGG
- **YaHS**
 - |_ ver: 1.2a
 - |_ key param: --no-contig-ec
- **Merqury**
 - |_ ver: 1.3
 - |_ key param: NA
- **BUSCO**
 - |_ ver: 6.0.0
 - |_ key param: NA
- **tidk**
 - |_ ver: 0.2.65
 - |_ key param: NA
- **chromap**
 - |_ ver: 0.3.2
 - |_ key param: NA
- **FOAM pipeline**
 - |_ ver: 0.5
 - |_ key param: NA
- **mitos**
 - |_ ver: 2.1.3
 - |_ key param: NA

Curation pipeline

- **Blobtoolkit DToL Nextflow pipeline**
 - |_ ver: 0.6.0
 - |_ key param: NA
- **FCS-GX**
 - |_ ver: 0.5.5
 - |_ key param: NA
- **PretextViewAI**
 - |_ ver: 1.0.5
 - |_ key param: NA
- **SANGER tola-agp-tpf-utils**
 - |_ ver: 1.2.1

|_ *key param*: NA

Submitter: Fernando Cruz

Affiliation: CNAG

Date and time: 2026-02-09 12:52:27 CET