ERGA Assembly Report
v24.10.15

Tags: ERGA-BGE

TxID	470652
ToLID	**iqPlaAffi1**
Species	Platycleis affinis
Class	Insecta
Order	Orthoptera

Genome Traits	Expected	Observed
Haploid size (bp)	5,236,719,231	5,448,474,875
Haploid Number	15 (source: ancestor)	14
Ploidy	2 (source: ancestor)	2
Sample Sex	XX	XX

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 8.8.Q73
Obtained EBP quality metric for hap2: 8.8.Q72

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 hap1
. Kmer completeness value is less than 90 for hap2

Curator notes

. Interventions/Gb:
. Contamination notes: "463,602 bp contained in 5 sequences were considered contaminats by FCS-GX and removed from hap1. Additionally, 215,540 bp contained in 3 sequences were also considered contaminants and removed from hap2. The longest removed contigs for hap1 and hap2 were 113,622 and 87,503 bp, respectively."
. Other observations: "This genome assembly was produced by assembling ONT data with Hifiasm, phasing with Hi-C data, and then scaffolding with Yahs. Given the good quality of both haplotypes, curation was made in diploid mode and we are planning to submit both haplotypes to the ENA. Curation made 11 cuts in contigs, 5 breaks at gaps and 22 joins. The X chromosome was identified by WGA against iqPlaFalx. The mitogenome was successfully assembled with FOAM."

Quality metrics table

Metrics	Pre-curation hap1	Pre-curation hap2	Curated hap1	Curated hap2
Total bp	5,454,009,557	5,428,956,239	5,448,474,875	5,433,815,179
GC %	39.28	39.29	39.28	39.28
Gaps/Gbp	9.9	9.95	8.99	13.99
Total gap bp	10,800	10,800	9,800	15,200
Scaffolds	140	73	118	81
Scaffold N50	651,131,869	640,708,031	643,586,180	640,708,031
Scaffold L50	4	4	4	4
Scaffold L90	11	11	11	11
Contigs	194	127	167	157
Contig N50	204,990,438	146,597,035	211,379,782	156,705,242
Contig L50	9	10	8	11
Contig L90	32	36	28	44
QV	72.6626	72.3099	73.086	72.2699
Kmer compl.	77.8108	77.7795	77.8764	77.7093
BUSCO sing.	94.6%	95.0%	95.0%	94.7%
BUSCO dupl.	4.0%	3.6%	3.8%	3.9%
BUSCO frag.	0.3%	0.3%	0.3%	0.4%
BUSCO miss.	1.0%	1.0%	1.0%	1.0%

BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: insecta_odb12 (genomes:79, BUSCOs:3114)

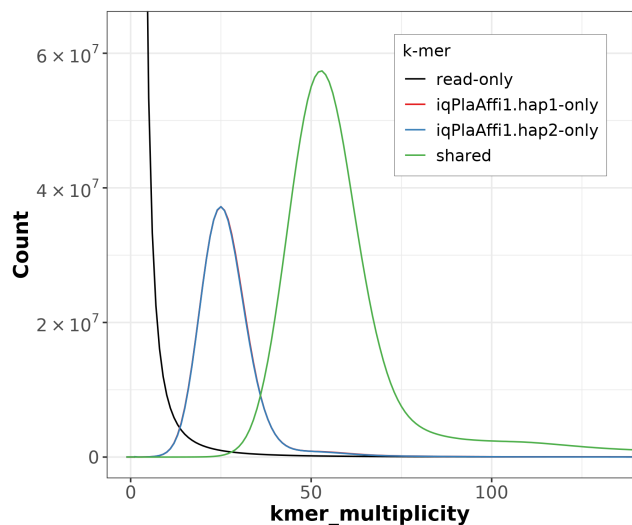# HiC contact map of curated assembly
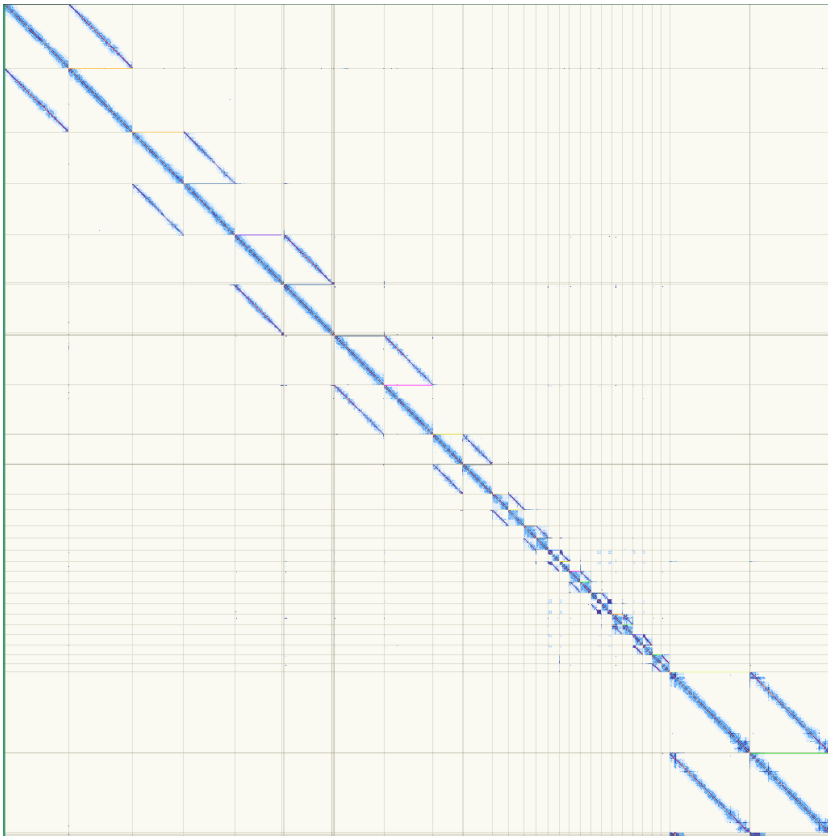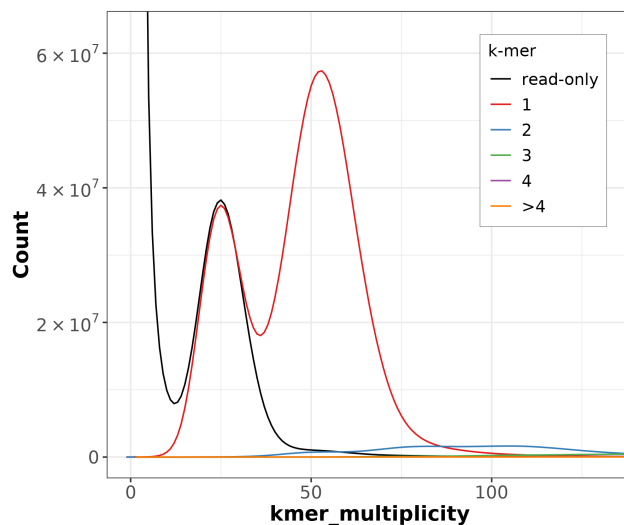

hap1 [LINK]

hap2 HiC PNG is missing!

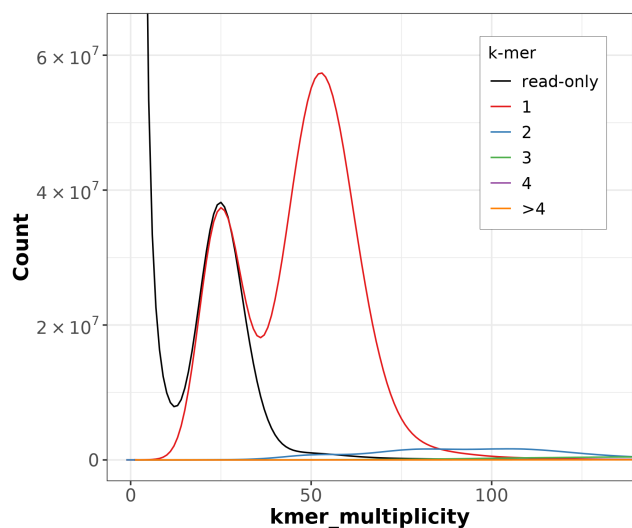**hap2** File link is missing!
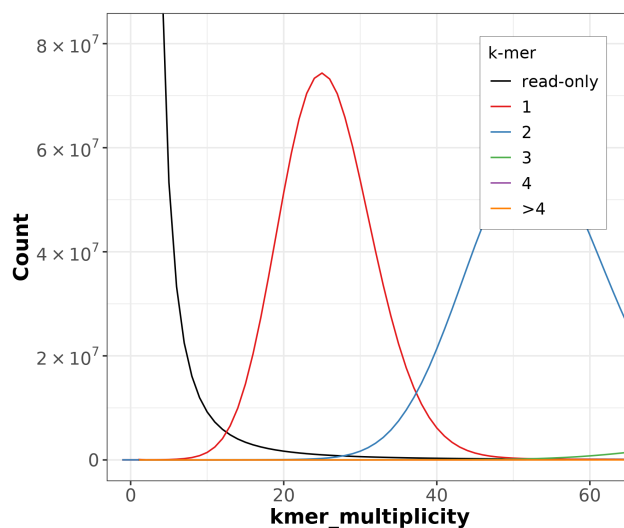
K-mer spectra of curated assembly

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

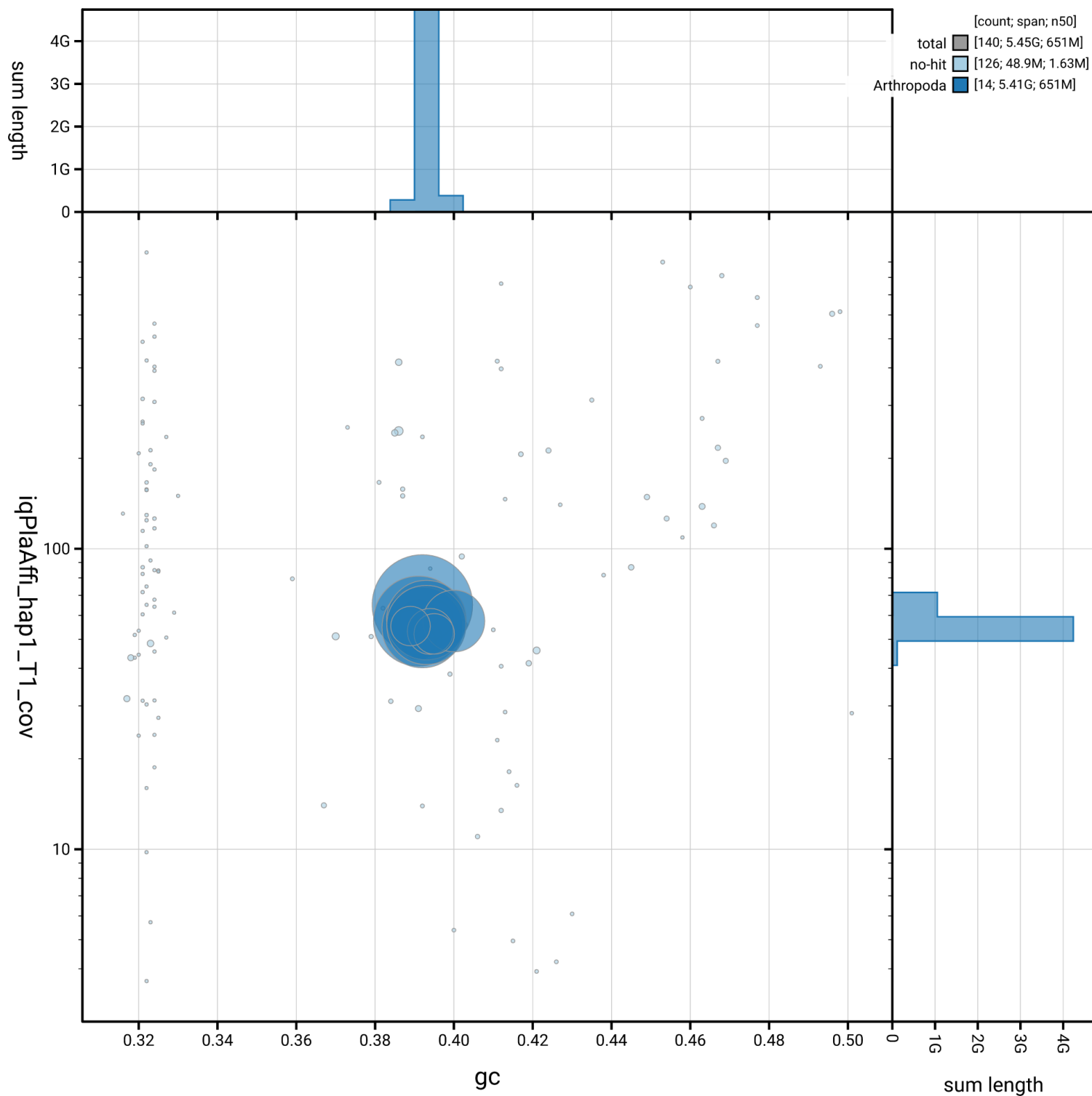Distribution of k-mer counts per copy numbers found in **hap2** (hapl.)

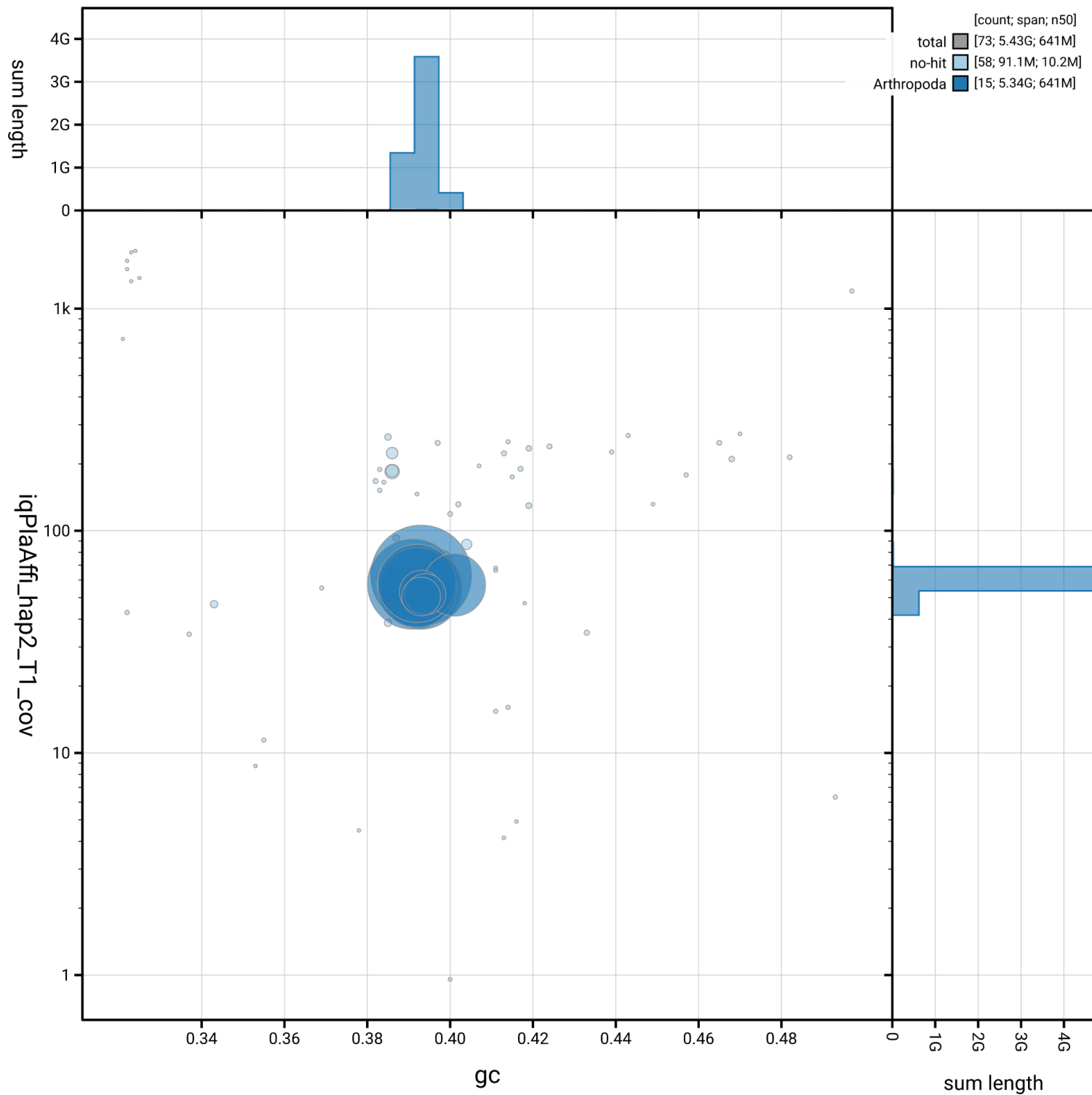Distribution of k-mer counts per copy numbers found in **hap1** (hapl.)

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

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	Arima Hi-C
Coverage	115	58

Assembly pipeline

- **CLAWS**
    |_ *ver:* v3.1
    |_ *key param:* NA
- **Filtlong**
    |_ *ver:* v0.2.1
    |_ *key param:* --minlen 1000
    |_ *key param:* --min_mean_q 80
    |_ *key param:* --target_bases 300000000000
- **Hifiasm**
    |_ *ver:* 0.24.0
    |_ *key param:* --ont
    |_ *key param:* --telo-m TTAGG
    |_ *key param:* -h1 -h2
- **Yahs**
    |_ *ver:* 1.2a
    |_ *key param:* -mq 10
    |_ *key param:* --no-contig-ec
- **FOAM**
    |_ *ver:* 0.5
    |_ *key param:* NA

Curation pipeline

- **FCS-GX**
    |_ *ver:* None
    |_ *key param:* NA
- **PretextViewAI**
    |_ *ver:* 1.0.5
    |_ *key param:* NA

Submitter: Jessica Gomez-Garrido
Affiliation: CNAG

Date and time: 2026-02-18 14:10:49 CET