

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

| | |
|---------|-----------------------------|
| TxID | 3139675 |
| ToLID | qdTypCoec2 |
| Species | <i>Typhloglomeris coeca</i> |
| Class | Diplopoda |
| Order | Glomerida |

| Genome Traits | Expected | Observed |
|-------------------|-----------------------|-------------|
| Haploid size (bp) | 326,504,482 | 325,456,575 |
| Haploid Number | 12 (source: ancestor) | 9 |
| Ploidy | 2 (source: ancestor) | 2 |
| Sample Sex | unknown | unknown |

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.7.Q69

Obtained EBP quality metric for hap2: 7.7.Q71

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
- . Assembly length loss > 3% for hap1
- . Assembly length loss > 3% for hap2

Curator notes

- . Interventions/Gb: 1
 - . Contamination notes: "Repeated attempts to run BlobTools were unsuccessful, so we used an alternative tool called FCS-GX (Foreign Contamination Screen - Genome eXclusion), to identify contaminated scaffolds. In this species, hap1 contained 496 contaminated scaffolds with diverse bacterial sequences. Of these, seven scaffolds exceeded 1 Mbp. Hap2 contained 149 contaminated scaffolds, two of which were also longer than 1 Mbp. All contaminated scaffolds from both haps were flagged as contaminants during a preliminary curation step using pretextview, and were subsequently removed from the analysis in the pretextview file shared for review."
 - . Other observations: "This assembly was generated using Hifiasm with ONT data. Given the high quality of both haplotypes, curation was performed in diploid mode. We are sharing the curated haplotypes extracted from the diploid curation map.
- Intentionally, we selected the aligned SUPER scaffolds with the fewest gaps to construct hap1. Only minor modifications were made to the yahs-scaffolded assemblies.

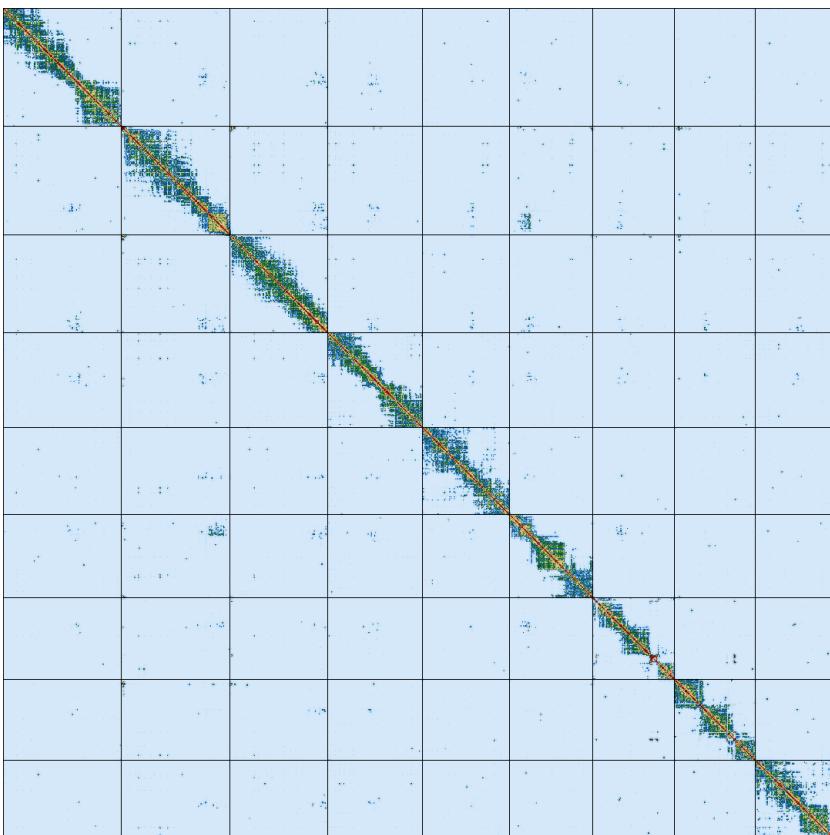
No sex chromosome could be identified. The rev1 assembly addresses the reviewers comments: 1- rearrangement of SUPER_9 of HAP2 to fix a large assembly error, 2- ensure that the SUPER sscaffolds of HAP1 and HAP2 follow the same order."

Quality metrics table

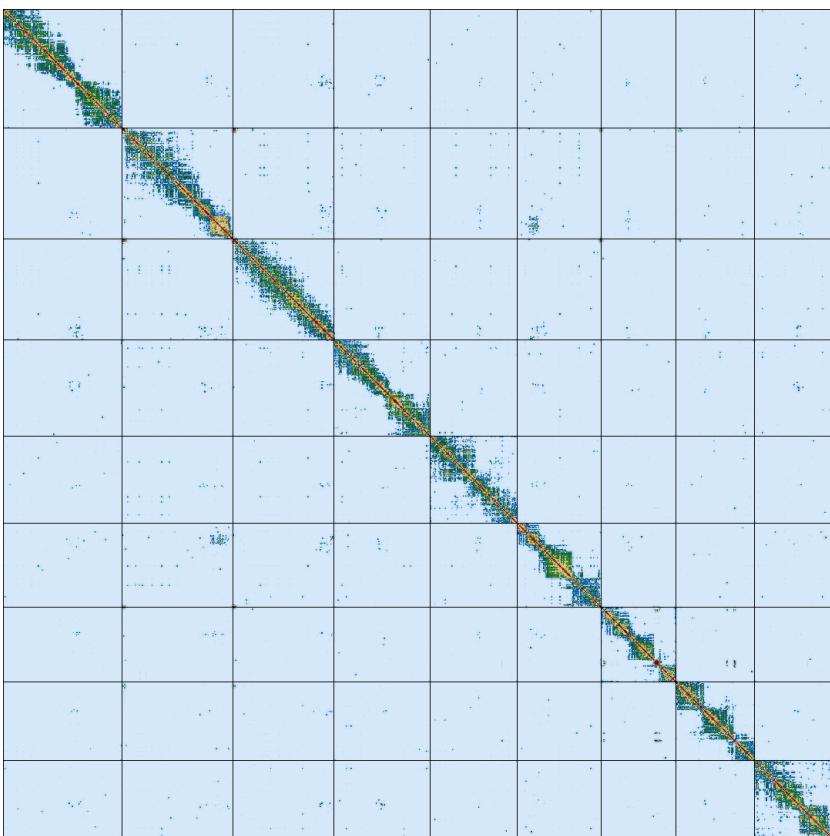
| Metrics | Pre-curation hap1 | Pre-curation hap2 | Curated hap1 | Curated hap2 |
|--------------|----------------------|----------------------|-----------------|-----------------|
| Total bp | 396,695,769 | 350,105,123 | 325,456,575 | 319,913,225 |
| GC % | 39.37 | 38.75 | 38.81 | 38.78 |
| Gaps/Gbp | 22.69 | 51.41 | 6.15 | 53.14 |
| Total gap bp | 1,800 | 3,600 | 400 | 3,400 |
| Scaffolds | 763 | 188 | 9 | 9 |
| Scaffold N50 | 32,385,137 | 33,396,051 | 37,203,801 | 36,917,534 |
| Scaffold L50 | 6 | 5 | 4 | 4 |
| Scaffold L90 | 60 | 9 | 8 | 8 |
| Contigs | 772 | 206 | 11 | 26 |
| Contig N50 | 31,870,558 | 32,668,977 | 33,015,150 | 28,239,313 |
| Contig L50 | 6 | 5 | 5 | 5 |
| Contig L90 | 67 | 16 | 9 | 13 |
| QV | 51.0894 | 57.395 | 69.1425 | 71.1675 |
| Kmer compl. | 85.87 | 84.6058 | 81.2215 | 80.9319 |
| BUSCO sing. | 97.7% | 97.6% | 97.8% | 97.5% |
| BUSCO dupl. | 0.3% | 0.2% | 0.2% | 0.2% |
| BUSCO frag. | 0.4% | 0.5% | 0.3% | 0.4% |
| BUSCO miss. | 1.6% | 1.7% | 1.7% | 1.9% |

BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

HiC contact map of curated assembly

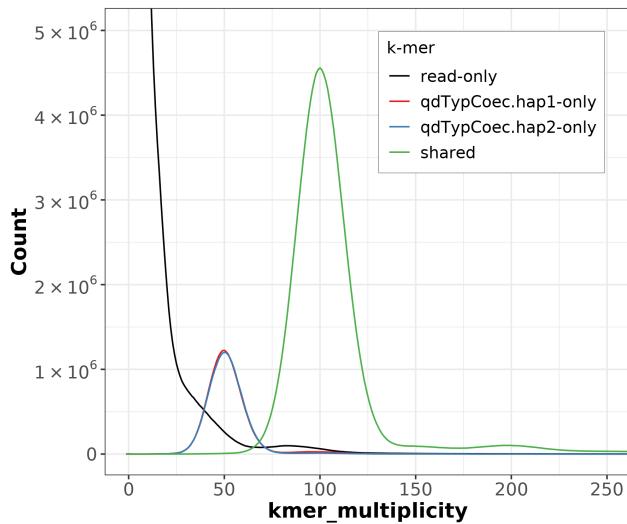


hap1 [\[LINK\]](#)

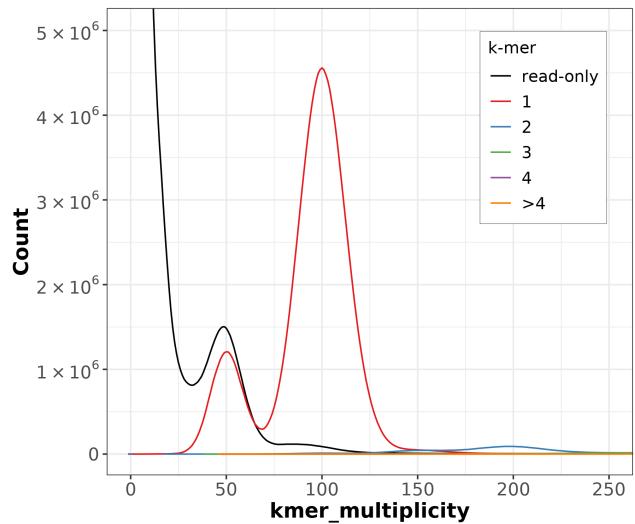


hap2 [\[LINK\]](#)

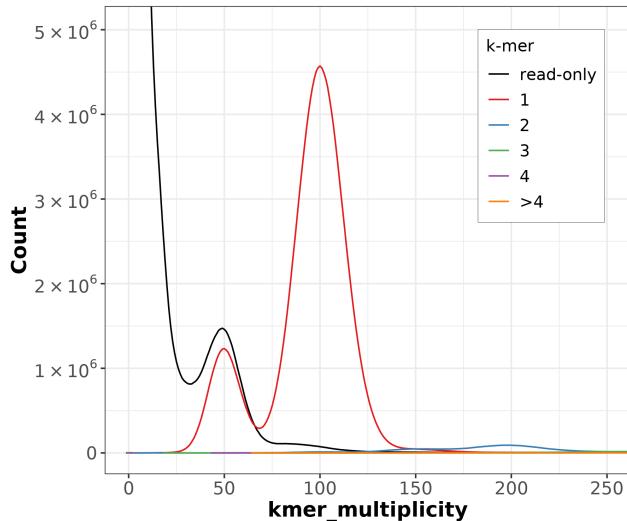
K-mer spectra of curated assembly



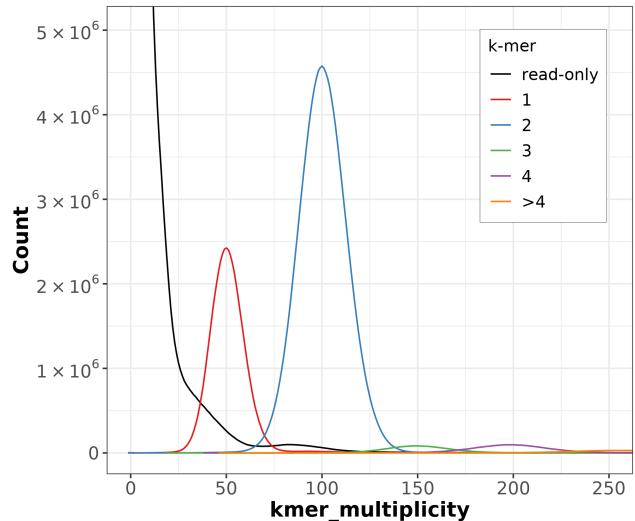
Distribution of k-mer counts coloured by their presence in reads/assemblies



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

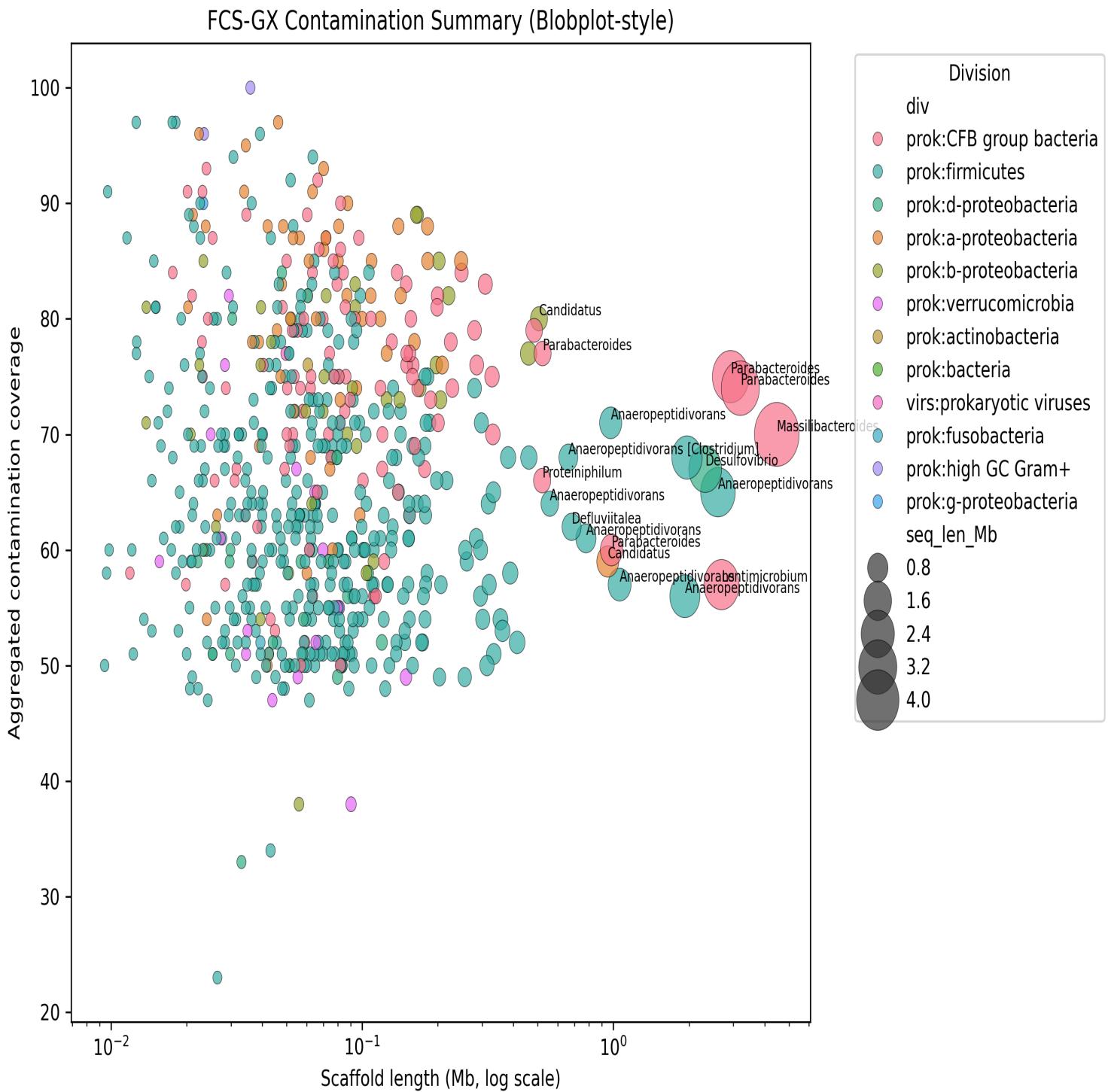


Distribution of k-mer counts per copy numbers found in **hap1** (hapl.)



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.

Data profile

| Data | ONT | Omni-C |
|----------|-----|--------|
| Coverage | 111 | 100 |

Assembly pipeline

- **CLAWS**
 - |_ ver: v3.1.0
 - |_ key param: NA
- **Filtlong**
 - |_ ver: v0.2.1
 - |_ key param: NA
- **Hifiasm**
 - |_ ver: 0.24.0
 - |_ key param: --ont
- **Yahs**
 - |_ ver: 1.2a
 - |_ key param: NA

Curation pipeline

- **FCS-GX (ForeignContamination Screen-Genome eXclusion) pipeline**
 - |_ ver: v0.5.5-0
 - |_ key param: NA
- **PretextViewAI**
 - |_ ver: 1.0.5
 - |_ key param: NA

Submitter: Francisco Camara Ferreira
Affiliation: CNAG

Date and time: 2026-02-19 15:14:52 CET