

# ERGA Assembly Report

v24.04.03\_beta

Tags: ERGA-BGE

TxID	270691
ToLID	<b>wcHelEuro1</b>
Species	Helobdella europaea
Class	Clitellata
Order	Rhynchobdellida

Genome Traits	Expected	Observed
Haploid size (bp)	207,706,424	243,872,151
Haploid Number	9 (source: ancestor)	18
Ploidy	2 (source: ancestor)	2
Sample Sex	H	H

## EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 5.7.Q54

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

- . Observed Haploid Number is different from Expected
- . More than 1000 gaps/Gbp for pri

### Curator notes

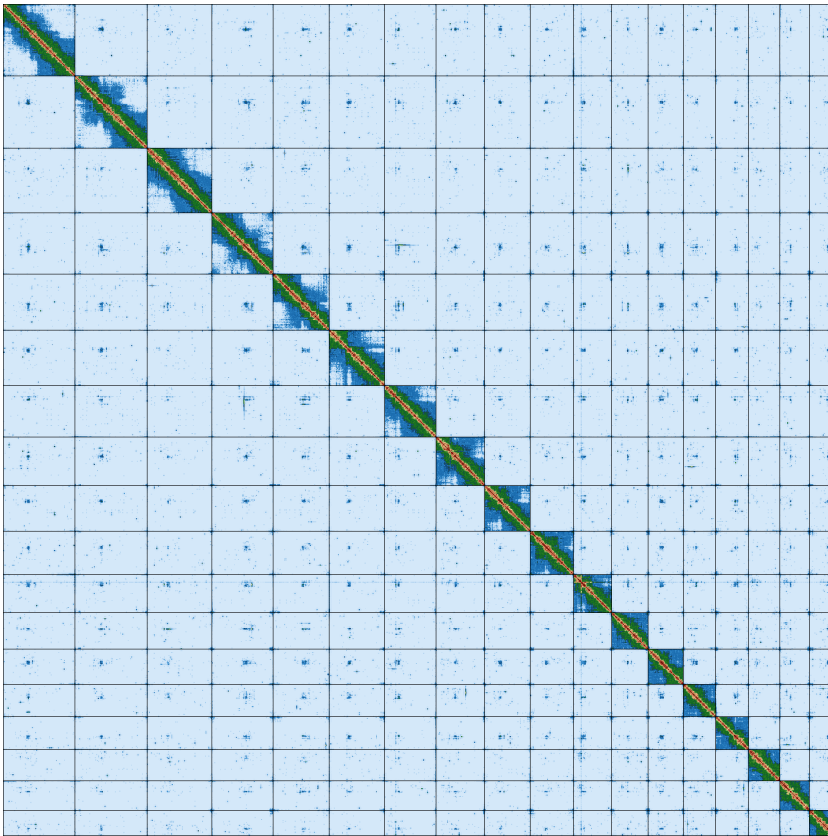
- . Interventions/Gb: 457
- . Contamination notes: "Contamination report for assembly labelled primary; Total length of scaffolds removed: 27,517 (0.0 %);Scaffolds removed: 2 (1.2 %);Largest scaffold removed: (17,478);FCS-GX contaminant species (number of scaffolds; total length of scaffolds):Microbulbifer hydrolyticus, g-proteobacteria (1; 17,478);Mitochondrion (1; 10,039);FCS-Adaptor (4; 402)"
- . Other observations: "Hi-C was from a separate individual (wcHelEuro3);"

# Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	243,960,360	243,872,151
GC %	31.26	31.26
Gaps/Gbp	4,980.32	5,597.2
Total gap bp	121,500	165,102
Scaffolds	162	29
Scaffold N50	14,408,463	14,980,787
Scaffold L50	7	7
Scaffold L90	16	16
Contigs	1,377	1,394
Contig N50	300,589	293,996
Contig L50	244	247
Contig L90	826	840
QV	54.1	54.2
Kmer compl.	99.65	99.67
BUSCO sing.	90.1%	90.1%
BUSCO dupl.	1.4%	1.4%
BUSCO frag.	1.6%	1.6%
BUSCO miss.	6.9%	6.9%

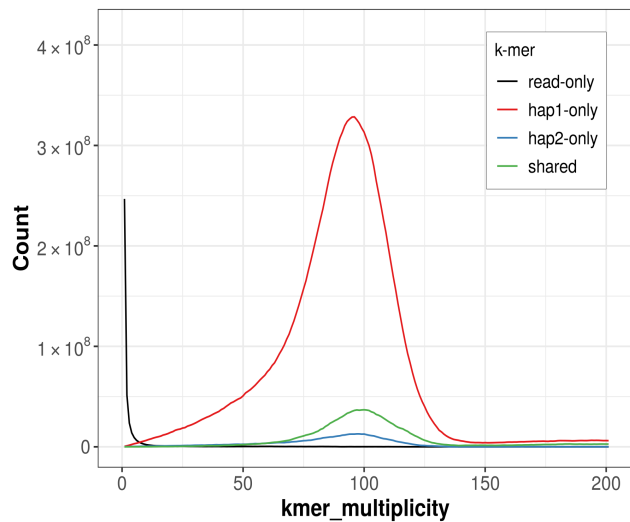
BUSCO 6.0.0 Lineage: metazoa\_odb10 (genomes:65, BUSCOs:954)

# HiC contact map of curated assembly

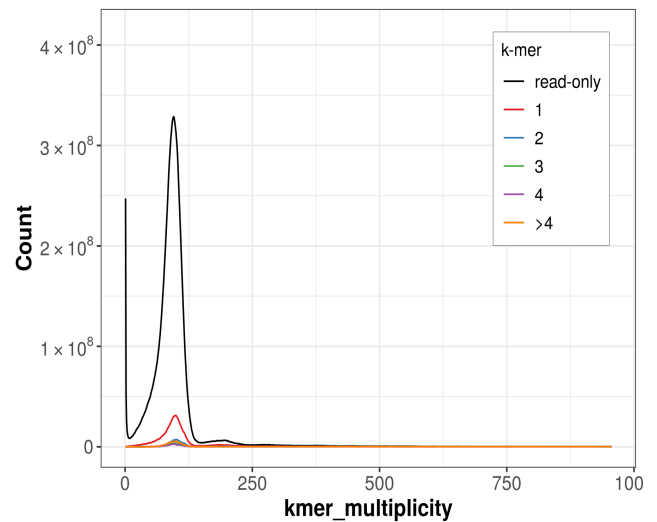


pri [\[LINK\]](#)

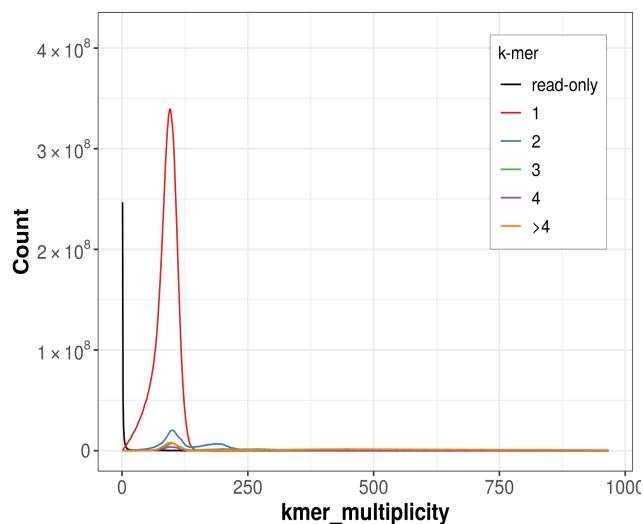
# K-mer spectra of curated assembly



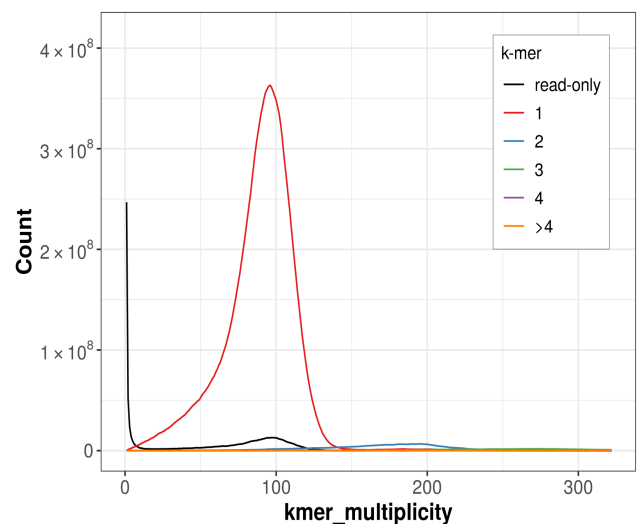
Distribution of k-mer counts coloured by their presence in reads/assemblies



Distribution of k-mer counts per copy numbers found in **hap2** (hap1.)

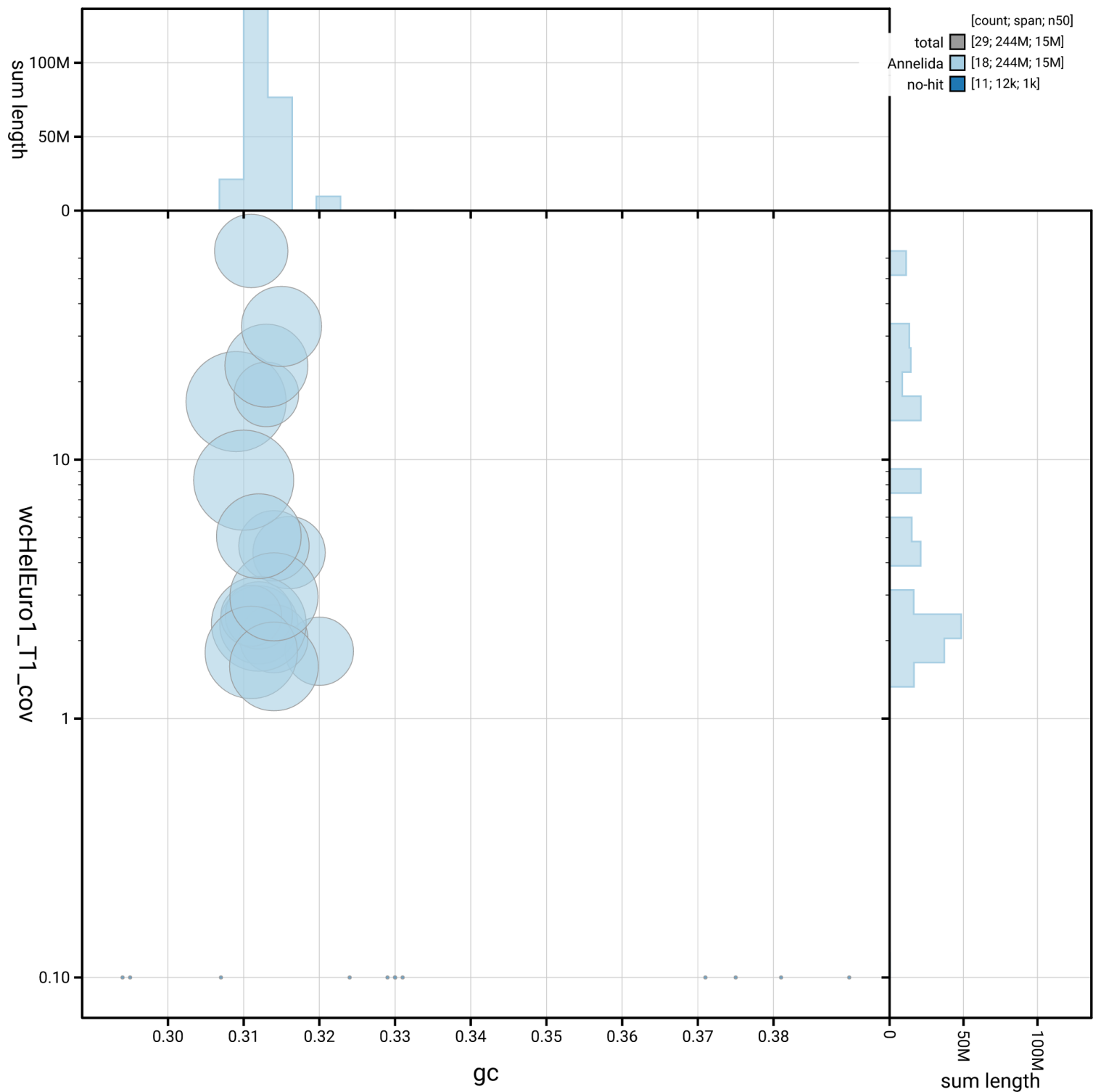


Distribution of k-mer counts per copy numbers found in asm (dipl.)



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

# 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	92x	652x

# Assembly pipeline

- **hifiasm**
  - |\_ *ver*: 0.19.8-r603
  - |\_ *key param*: --primary
- **yahs**
  - |\_ *ver*: 1.2.2
  - |\_ *key param*: NA

# Curation pipeline

- **hifiasm**
  - |\_ *ver*: 0.19.8-r603
  - |\_ *key param*: --primary
- **yahs**
  - |\_ *ver*: 1.2.2
  - |\_ *key param*: NA
- **TreeVal**
  - |\_ *ver*: 1.2.1
  - |\_ *key param*: NA

Submitter: Michael Paulini

Affiliation: WSI

Date and time: 2025-08-18 16:48:20 CEST