ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	75933
ToLID	**fUmbKra5**
Species	Umbra krameri
Class	Actinopteri
Order	Esociformes

Genome Traits	Expected	Observed
Haploid size (bp)	1,677,207,767	1,727,840,106
Haploid Number	10 (source: ancestor)	22
Ploidy	2 (source: ancestor)	2
Sample Sex	NA	NA

EBP metrics summary and curation notes

Obtained EBP quality metric for hap1: 7.7.Q65

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

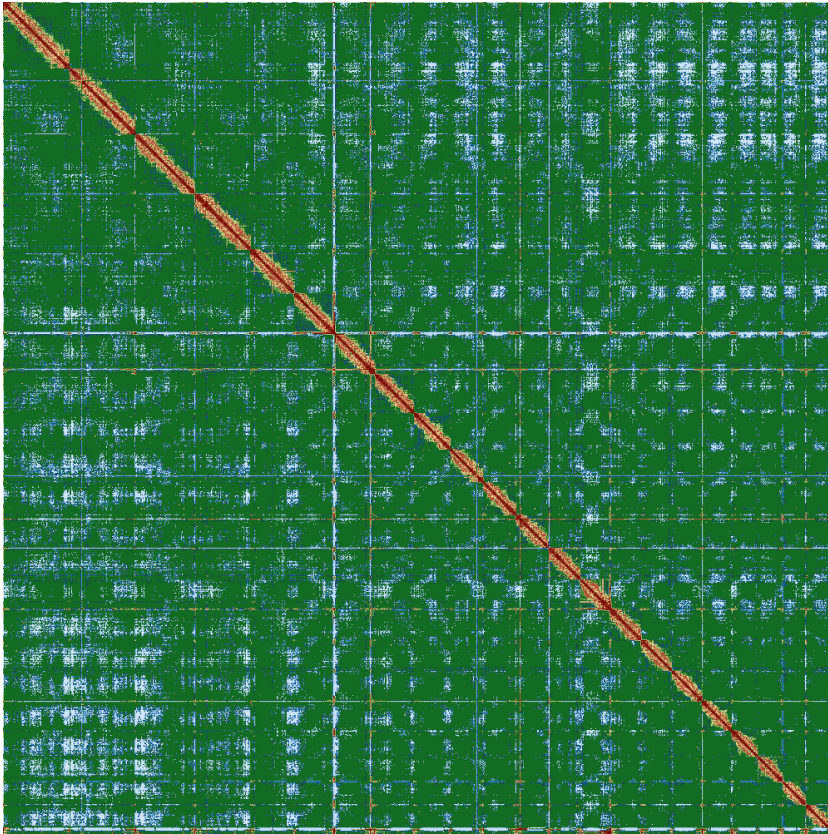. Observed Haploid Number is different from Expected

Curator notes

. Interventions/Gb: 35
. Contamination notes: "Contaminants were removed using NCB FCS and blobtools information. The main contaminant was Arthropoda, scaffold were removed if no contact was present in the map. The big contamination seen in the blob, is relative to two of our species chromsome, SUPER_2, SUPER_5 and SUPER_17. Those scaffolds were not removed since the contact on the map was good and constant in those 3 scaffold. "
. Other observations: "None"

Quality metrics table

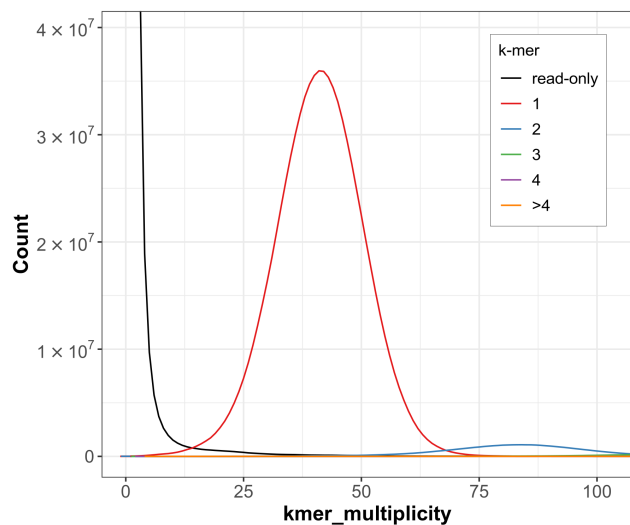| Metrics	Pre-curation hap1	Curated hap1
Total bp	1,729,746,145	1,727,840,106
GC %	40.41	40.4
Gaps/Gbp	79.78	96.65
Total gap bp	27,600	33,400
Scaffolds	237	167
Scaffold N50	75,584,379	75,584,379
Scaffold L50	9	9
Scaffold L90	19	19
Contigs	375	334
Contig N50	23,186,746	23,186,746
Contig L50	25	25
Contig L90	83	83
QV	65.4433	65.5402
Kmer compl.	98.9727	98.9722
BUSCO sing.	97.6%	97.6%
BUSCO dupl.	0.7%	0.7%
BUSCO frag.	1.2%	1.2%
BUSCO miss.	0.5%	0.5%

BUSCO: 5.8.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb10 (genomes:26, BUSCOs:3640)

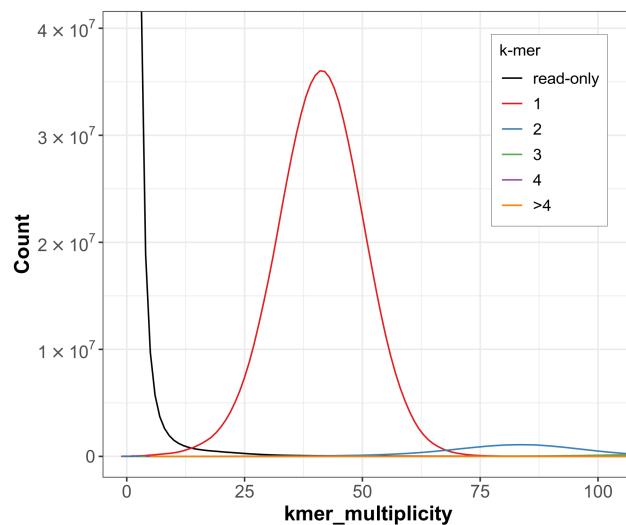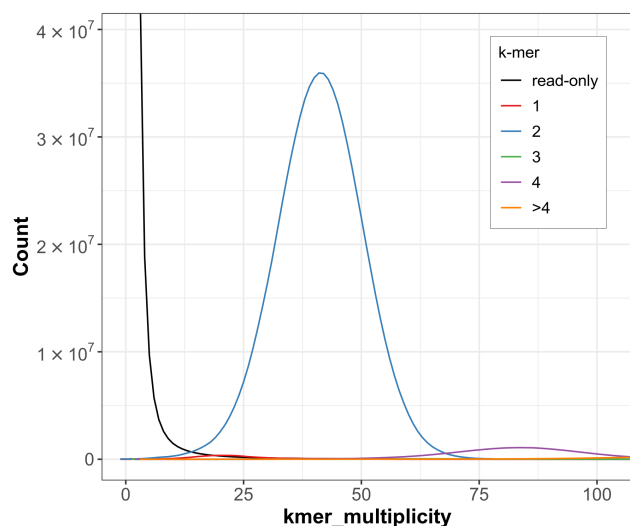# HiC contact map of curated assembly

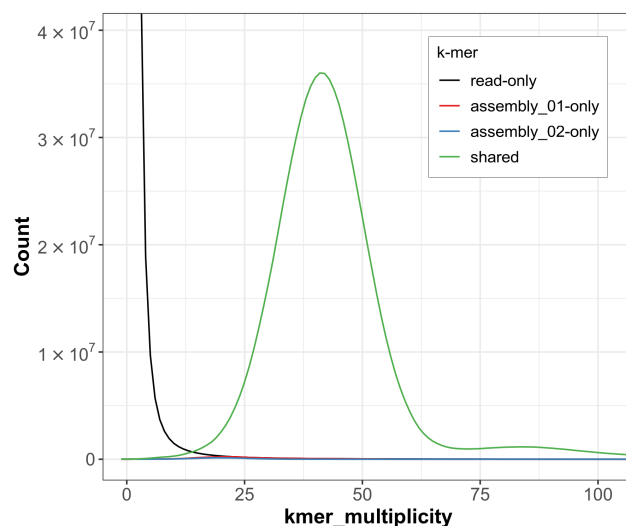

hap1 [LINK]

K-mer spectra of curated assembly

Distribution of k-mer counts per copy numbers found in **assembly_02** (hapl.)

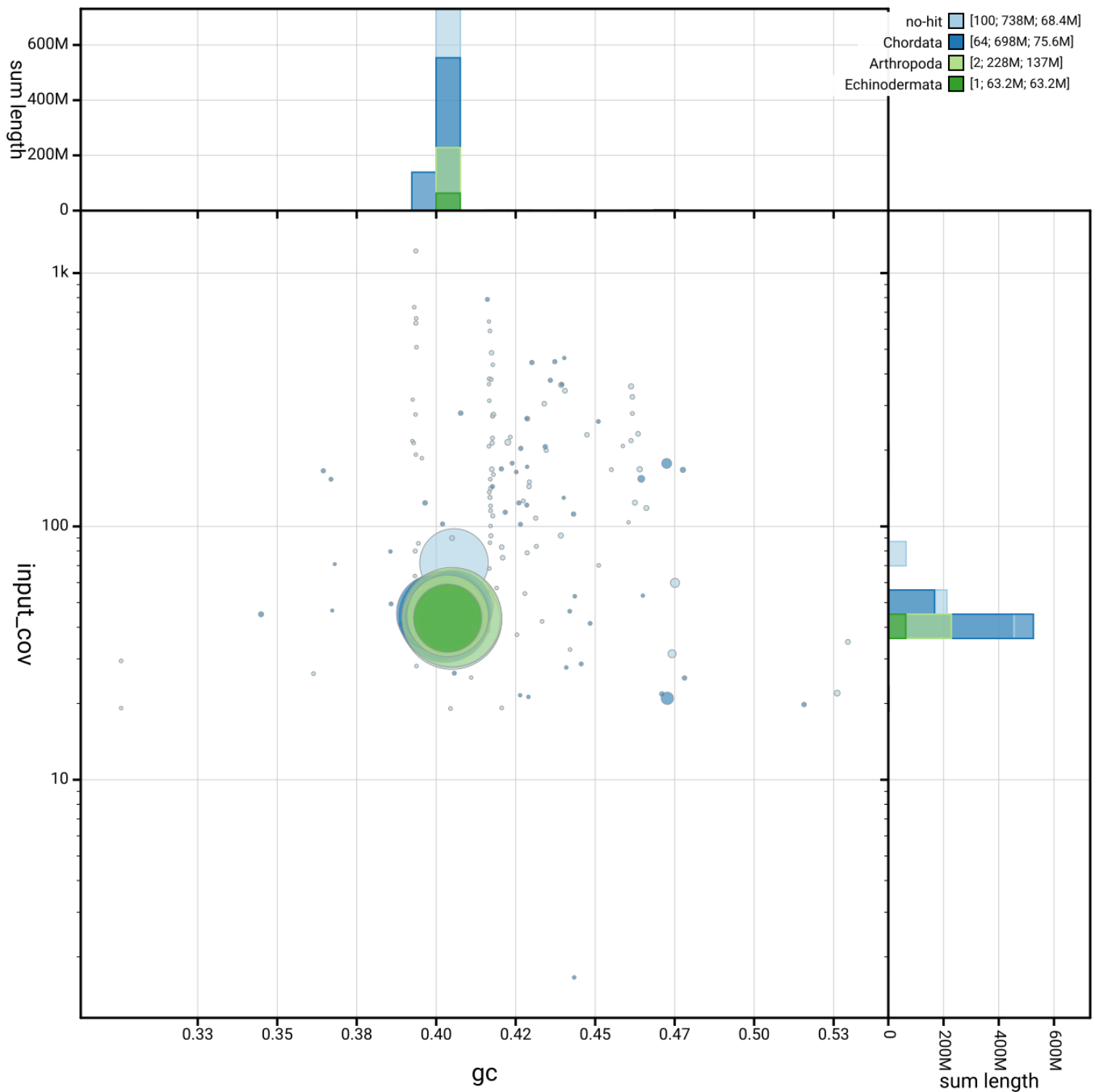Distribution of k-mer counts per copy numbers found in **assembly_01** (hapl.)

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

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

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	HiFi	HiC
Coverage	30.01X	76.45X

Assembly pipeline

- **Hifiasm**
    |_ *ver:* Galaxy Version 0.25.0+galaxy0
    |_ *key param:* VGP Galaxy workflow VGP4 v 0.4
- **YaHS**
    |_ *ver:* Galaxy Version 1.2a.2+galaxy3
    |_ *key param:* VGP Galaxy workflow VGP8 v 3.1

Curation pipeline

- **PretextMap**
    |_ *ver:* Galaxy Version 0.1.9+galaxy1
    |_ *key param:* VGP workflow VGP PretextMap generation v2
- **PretextViewAI**
    |_ *ver:* 1.0.3
    |_ *key param:* NA

Submitter: Marcella Sozzoni
Affiliation: UNIFI

Date and time: 2026-01-19 16:43:17 CET