ERGA Assembly Report

v24.10.15

Tags: Other[INVALID TAG]

TxID	8153
ToLID	**fHapBur1_hap2**
Species	Haplochromis burtoni
Class	Actinopteri
Order	Cichliformes

Genome Traits	Expected	Observed
Haploid size (bp)	827,591,195	881,617,941
Haploid Number	19 (source: ancestor)	20
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q59

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

. Observed Haploid Number is different from Expected

Curator notes

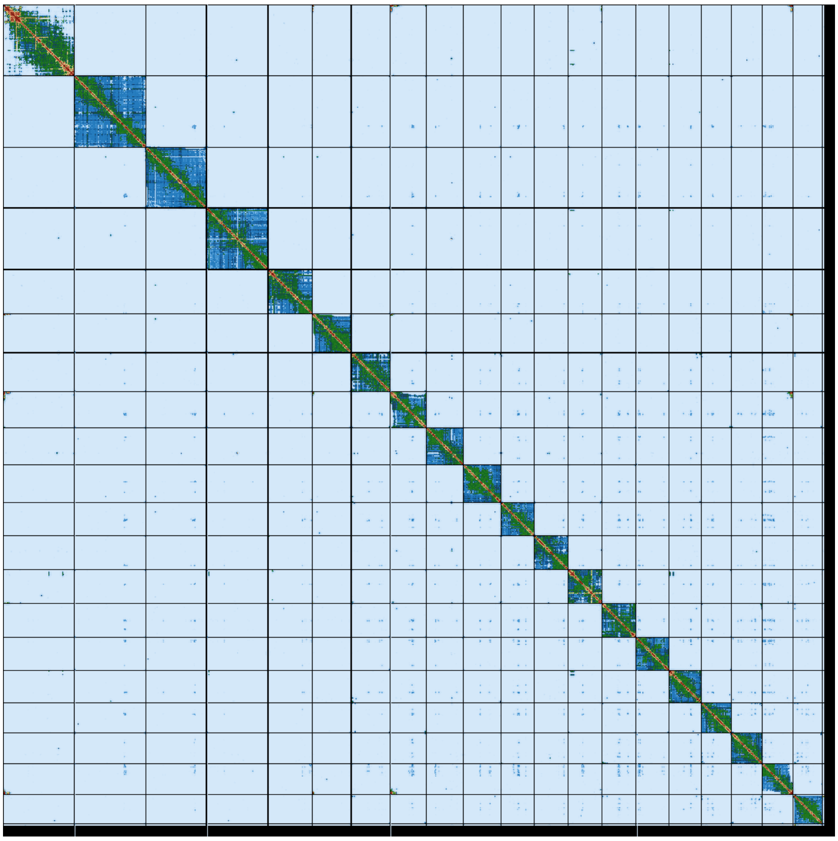
. Interventions/Gb: 116
. Contamination notes: ""
. Other observations: "The assembly of Astatotilapia burtoni (fHapBur1) was generated as part of the Vertebrate Genomes Project (https://vertebrategenomesproject.org/). Both haplotypes were analyzed and manually improved using Pretext "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	902,615,402	881,617,941
GC %	41.07	41.03
Gaps/Gbp	172.83	239.33
Total gap bp	31,200	42,200
Scaffolds	231	219
Scaffold N50	46,594,217	39,607,863
Scaffold L50	7	8
Scaffold L90	17	18
Contigs	387	430
Contig N50	9,134,839	8,396,100
Contig L50	28	34
Contig L90	101	117
QV	58.8809	59.1715
Kmer compl.	95.5461	94.0898
BUSCO sing.	96.6%	97.3%
BUSCO dupl.	0.8%	0.9%
BUSCO frag.	0.8%	0.1%
BUSCO miss.	1.9%	1.6%

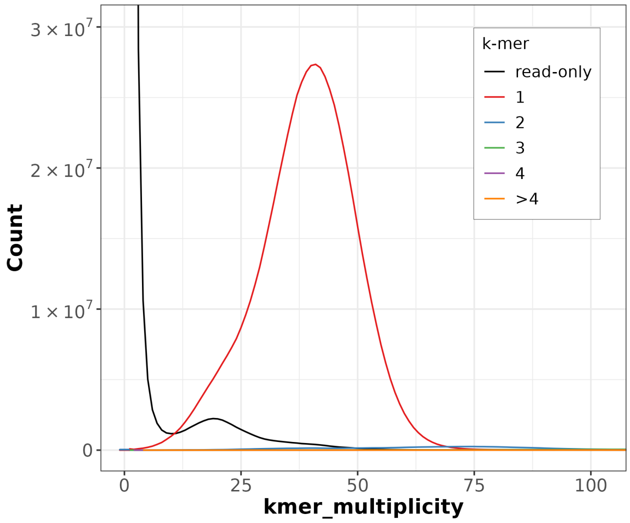
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 5.8.2 (euk_genome_met, metaeuk) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

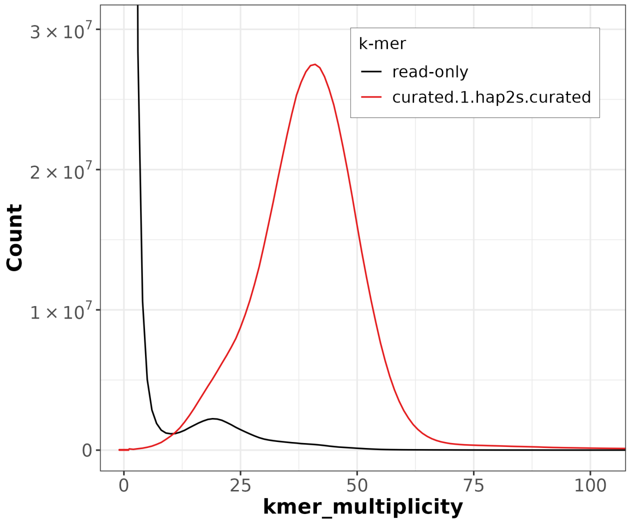


collapsed [LINK]

K-mer spectra of curated assembly

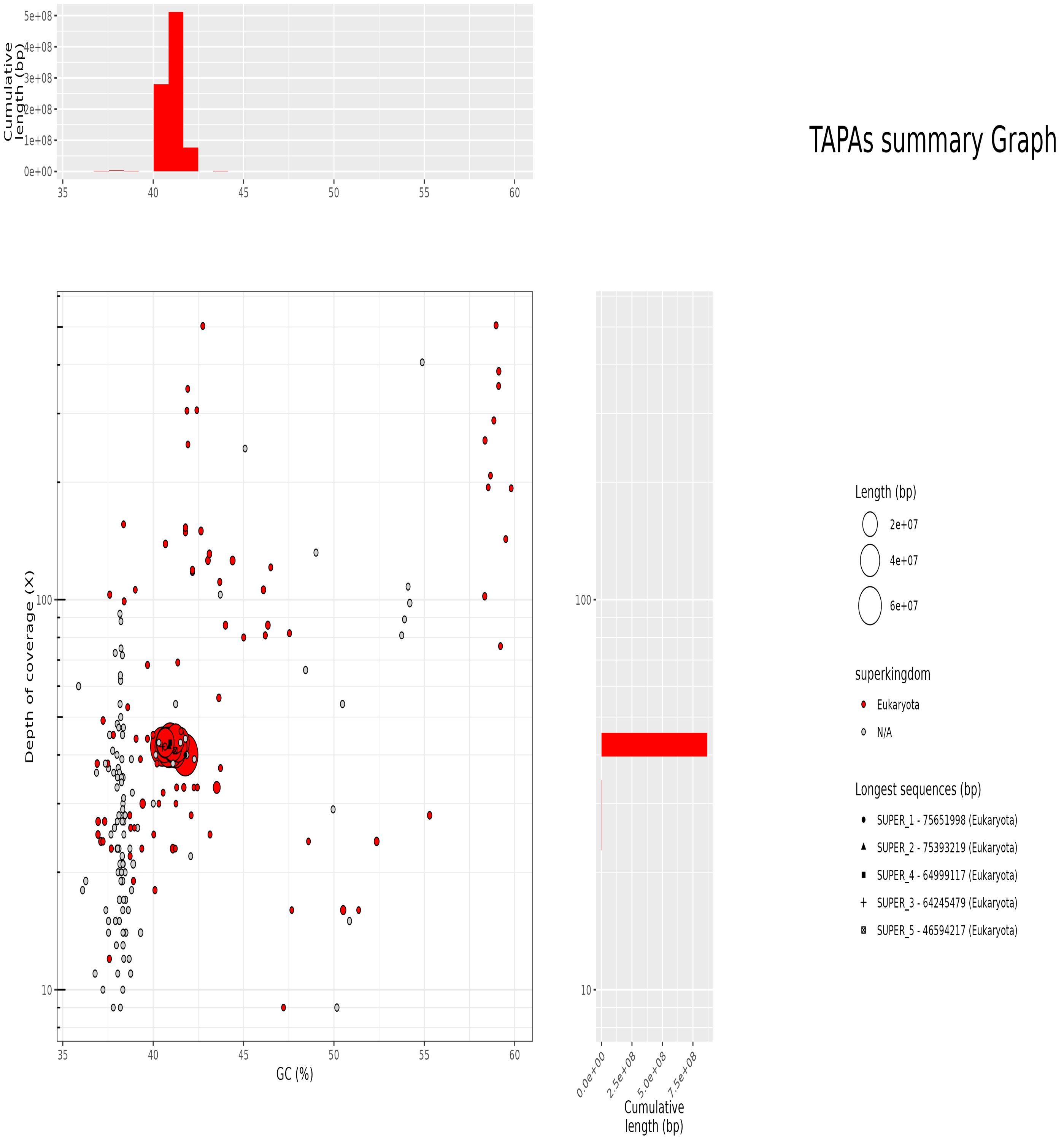


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



collapsed. 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	Long reads	Arima
Coverage	NA	NA

Assembly pipeline

- **Hifiasm**
    |_ *ver:* 0.19.5-r593
    |_ *key param:* NA
- **purge_dups**
    |_ *ver:* 1.2.5
    |_ *key param:* NA
- **YaHS**
    |_ *ver:* 1.2
    |_ *key param:* NA

Curation pipeline

- **PretextMap**
    |_ *ver:* 0.1.9
    |_ *key param:* NA
- **PretextView**
    |_ *ver:* 0.2.5
    |_ *key param:* NA

Submitter: Emilie Teodori
Affiliation: Genoscope

Date and time: 2025-10-28 15:15:01 CET