ERGA Assembly Report

v24.10.15

Tags: Other[INVALID TAG]

TxID	30501
ToLID	**sCarTau1_hap2**
Species	Carcharias taurus
Class	Chondrichthyes
Order	Lamniformes

Genome Traits	Expected	Observed
Haploid size (bp)	4,296,075,824	4,482,201,460
Haploid Number	39 (source: ancestor)	40
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.8.Q61

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

. Observed Haploid Number is different from Expected

Curator notes

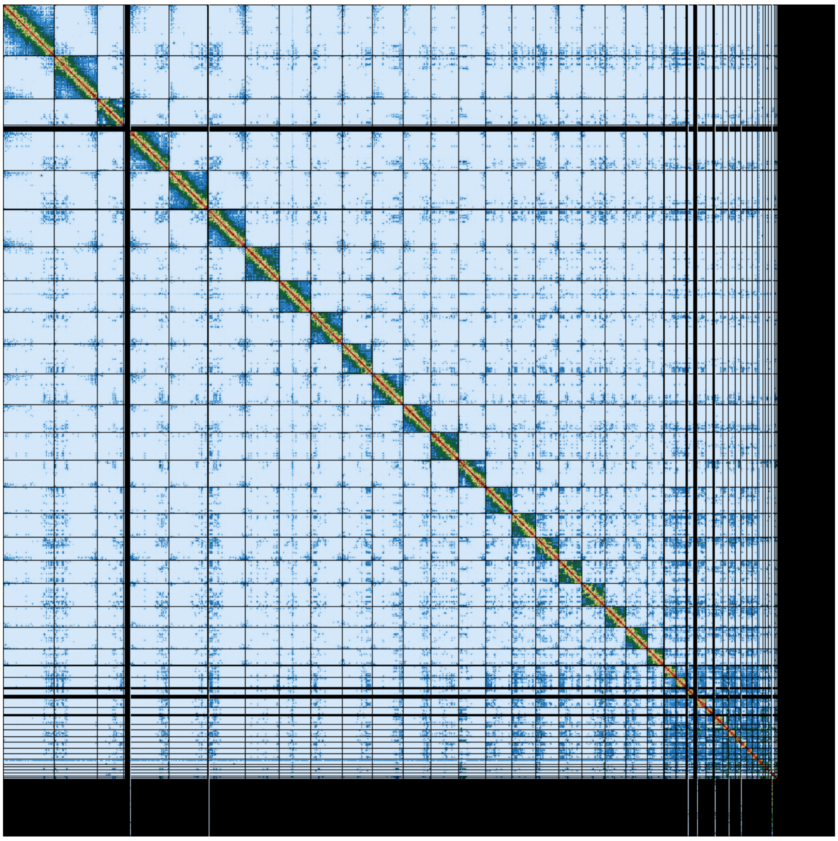
. Interventions/Gb: 24
. Contamination notes: ""
. Other observations: "The assembly of Carcharias taurus (sCarTau1) was generated as part of the Vertebrate Genomes Project (https://vertebrategenomesproject.org/). Both haplotypes were analyzed and manually improved using Pretext "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	4,481,244,989	4,482,201,460
GC %	45.14	45.15
Gaps/Gbp	328.48	342.47
Total gap bp	15,561,678	10,159,298
Scaffolds	1,231	1,170
Scaffold N50	147,306,798	148,959,525
Scaffold L50	12	12
Scaffold L90	53	36
Contigs	2,703	2,705
Contig N50	7,473,854	7,251,361
Contig L50	156	162
Contig L90	816	834
QV	61.0859	61.1035
Kmer compl.	96.4001	96.3376
BUSCO sing.	92.6%	96.0%
BUSCO dupl.	2.0%	2.0%
BUSCO frag.	2.6%	0.7%
BUSCO miss.	2.8%	1.3%

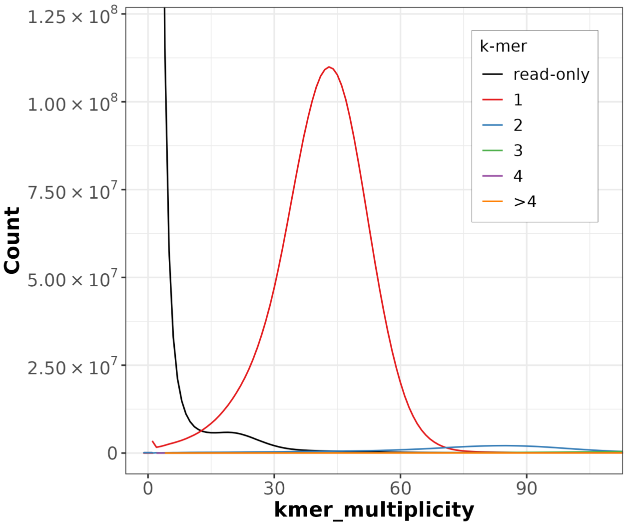
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 5.8.2 (euk_genome_met, metaeuk) / Lineage: vertebrata_odb12 (genomes:119, BUSCOs:3390)
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: vertebrata_odb12 (genomes:119, BUSCOs:3390)

HiC contact map of curated assembly

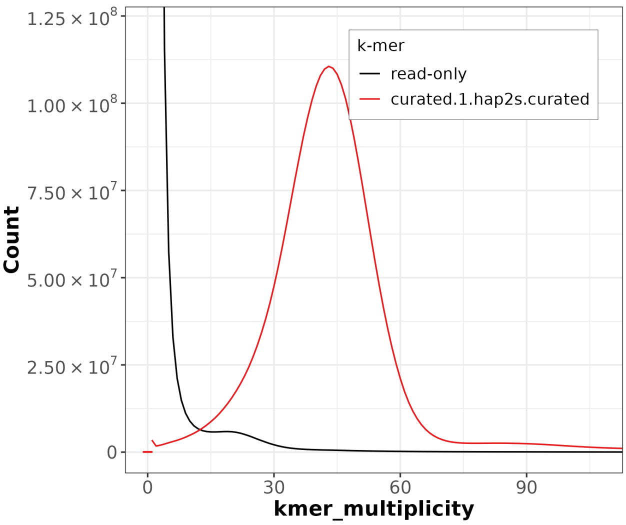


collapsed [LINK]

K-mer spectra of curated assembly

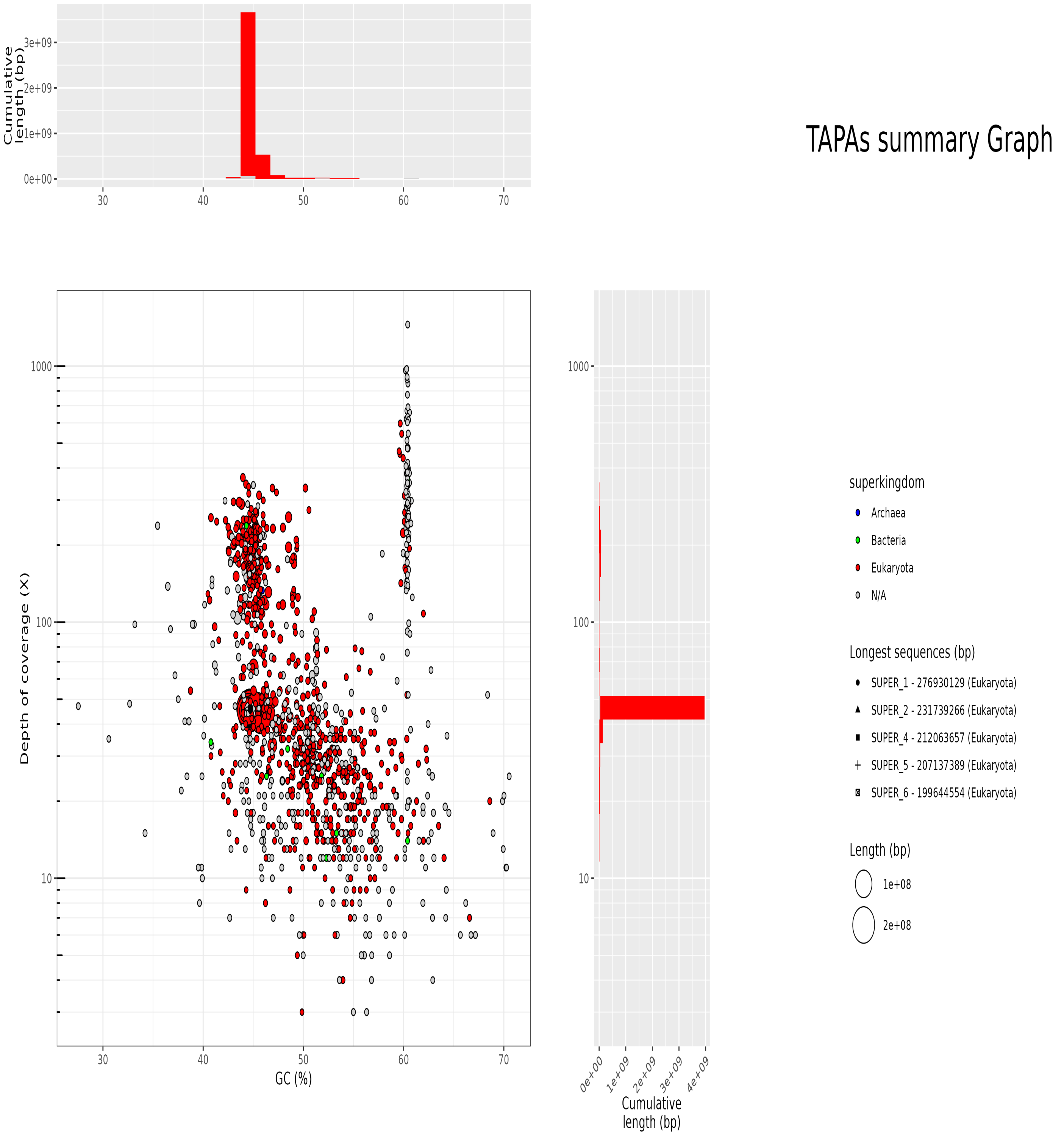


Distribution of k-mer counts per copy numbers found in asm



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: 2025-11-06 09:15:07 CET