ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	202727
ToLID	**fHelDac2**
Species	Helicolenus dactylopterus
Class	Actinopteri
Order	Perciformes

Genome Traits	Expected	Observed
Haploid size (bp)	764,064,405	811,380,064
Haploid Number	24 (source: ancestor)	24
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q50

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

. Kmer completeness value is less than 90 for collapsed

Curator notes

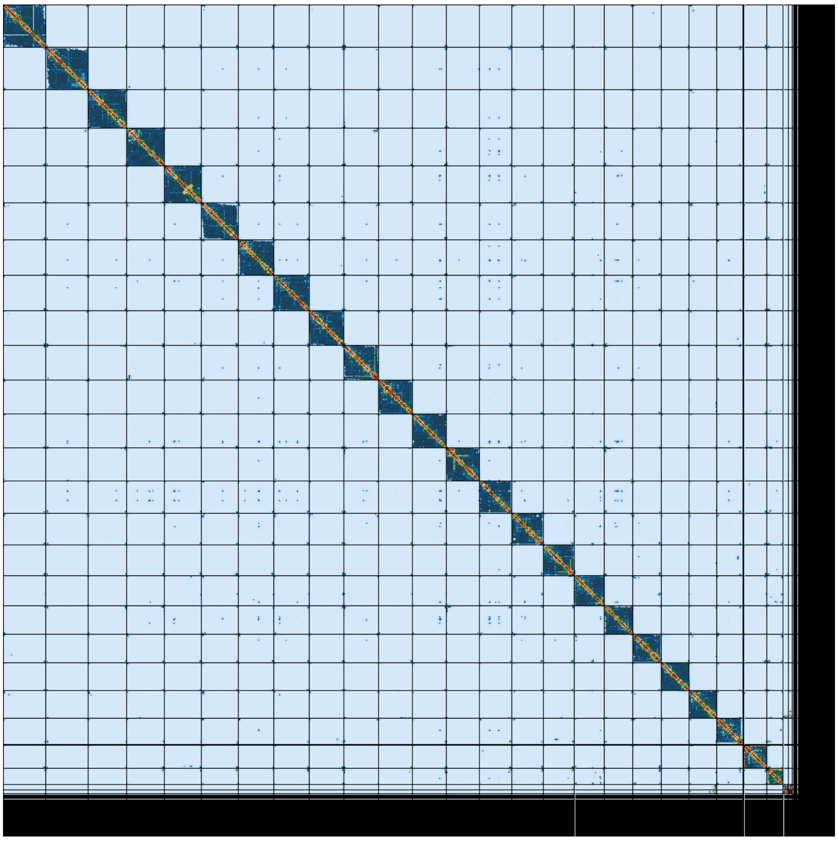
. Interventions/Gb: 5
. Contamination notes: ""
. Other observations: "The assembly of \'Helicolenus dactylopterus\' (fHelDac2) is based on 46X PacBio data and 144X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial PacBio assembly generation with Hifiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. No contigs were found to be contaminants (bacterial, archaeal, or viral). Additionally, 77 regions totaling 7.044 Mb (with the largest being 0.761 Mb) were identified as haplotypic duplications and removed. Mitochondrial genome was assembled using OATK. Finally, the primary assembly was analyzed and manually improved using Pretext. Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. "

Quality metrics table

Metrics	Pre-curation collapsed	Curated collapsed
Total bp	811,386,264	811,380,064
GC %	40.29	40.29
Gaps/Gbp	101.06	105.99
Total gap bp	8,200	9,000
Scaffolds	218	212
Scaffold N50	33,022,138	33,022,138
Scaffold L50	12	12
Scaffold L90	23	23
Contigs	300	298
Contig N50	18,565,825	18,565,825
Contig L50	16	16
Contig L90	61	61
QV	50.0904	50.0925
Kmer compl.	84.1794	84.1794
BUSCO sing.	99.3%	99.3%
BUSCO dupl.	0.3%	0.3%
BUSCO frag.	0.0%	0.0%
BUSCO miss.	0.3%	0.3%

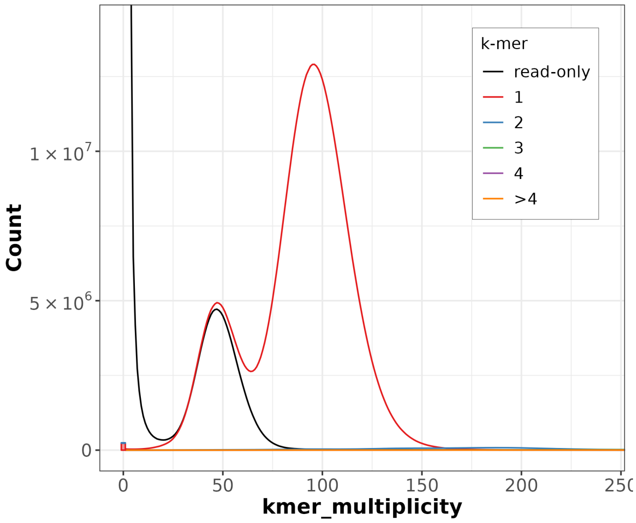
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: actinopterygii_odb12 (genomes:75, BUSCOs:7207)

HiC contact map of curated assembly

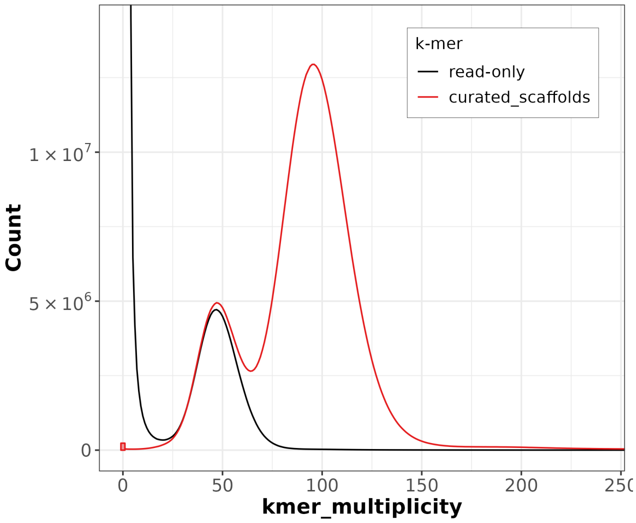


collapsed [LINK]

K-mer spectra of curated assembly

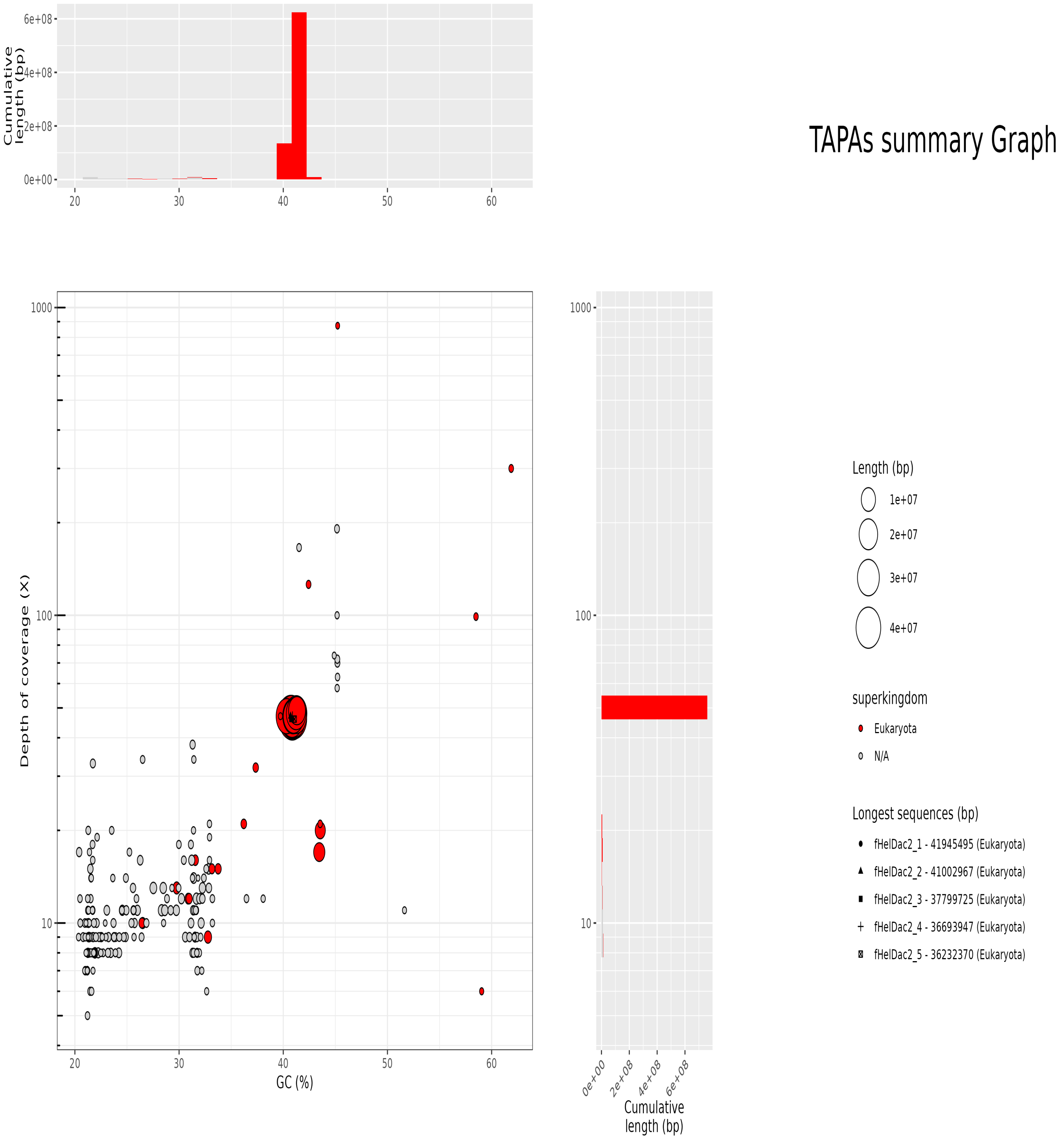


Distribution of k-mer counts per copy numbers found in asm



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	46	144

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: Simone Duprat
Affiliation: Genoscope

Date and time: 2026-01-22 20:17:20 CET