ERGA Assembly Report

v24.10.15

Tags: ATLASea[INVALID TAG]

TxID	3738056
ToLID	**nxAniSpea1**
Species	Anisakidae sp. A-QB95YQNDS
Class	Chromadorea
Order	Rhabditida

Genome Traits	Expected	Observed
Haploid size (bp)	153,776,468	150,238,396
Haploid Number	7 (source: ancestor)	39
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.6.Q51

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

. Observed Haploid Number is different from Expected

. Kmer completeness value is less than 90 for collapsed

Curator notes

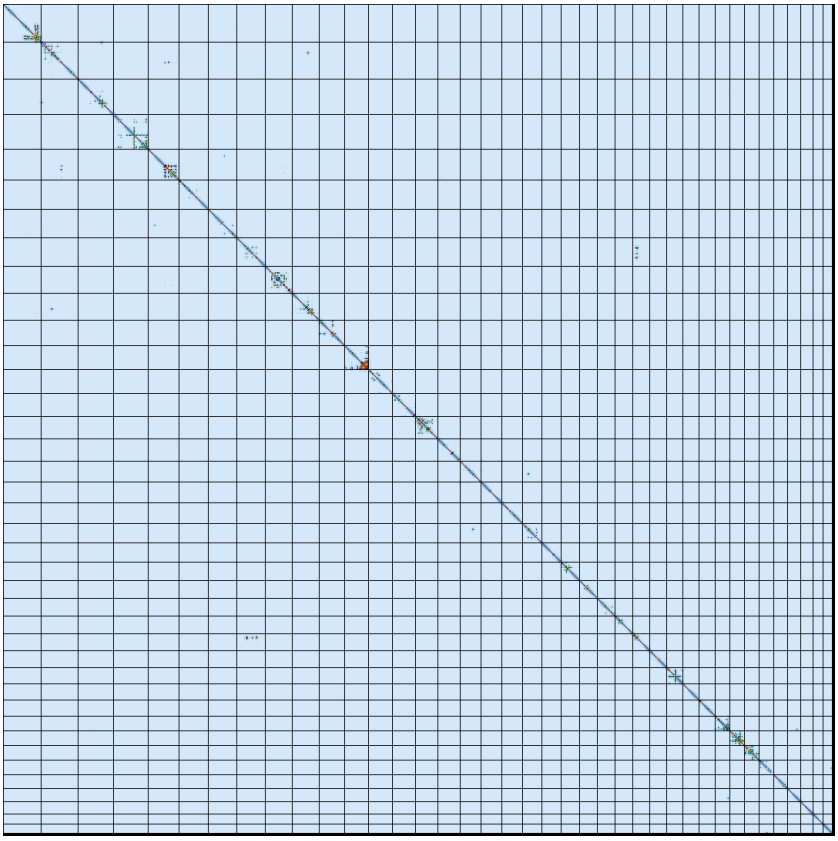
. Interventions/Gb: 93
. Contamination notes: ""
. Other observations: "The assembly of Anisakidae sp. 1 (nxAniSpea1) is based on 73X PacBio data and 79X Arima Hi-C data generated as part of the ATLASea programme (https://www.atlasea.fr). The assembly process included the following steps: initial PacBio assembly generation with Hifiasm, removal of contaminant sequences using Context, removal of haplotypic duplications using purge_dups, and Hi-C-based scaffolding with YaHS. In total, 11 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 0.245 Mb (with the largest being 0.038 Mb). Additionally, 51 regions totaling 1.966 Mb (with the largest being 0.276 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. During manual curation, 3 haplotypic regions were removed, totaling 1.477Mb (with the largest being 0.682Mb). 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	151,780,933	150,238,396
GC %	35.95	35.94
Gaps/Gbp	138.36	179.71
Total gap bp	2,100	3,600
Scaffolds	57	48
Scaffold N50	4,095,666	4,095,666
Scaffold L50	15	15
Scaffold L90	33	33
Contigs	78	75
Contig N50	3,752,000	3,552,733
Contig L50	16	16
Contig L90	38	38
QV	51.11	51.1088
Kmer compl.	74.2507	74.095
BUSCO sing.	97.3%	97.0%
BUSCO dupl.	1.0%	0.2%
BUSCO frag.	0.4%	1.3%
BUSCO miss.	1.3%	1.5%

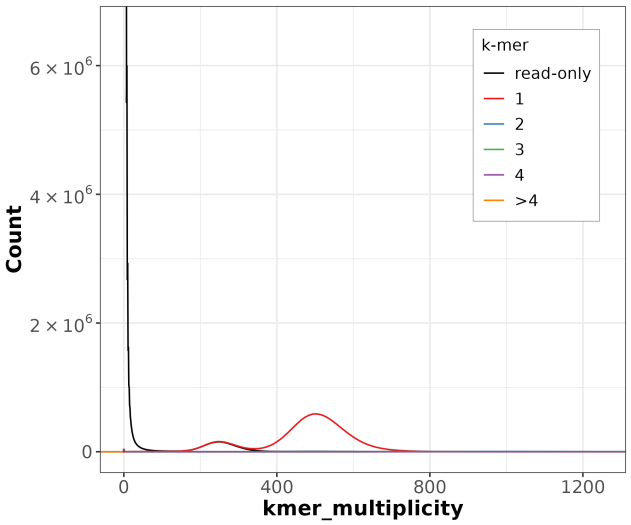
Warning! BUSCO versions or lineage datasets are not the same across results:
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: chromadorea_odb12 (genomes:48, BUSCOs:1126)
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: chromadorea_odb12 (genomes:48, BUSCOs:1126)

HiC contact map of curated assembly

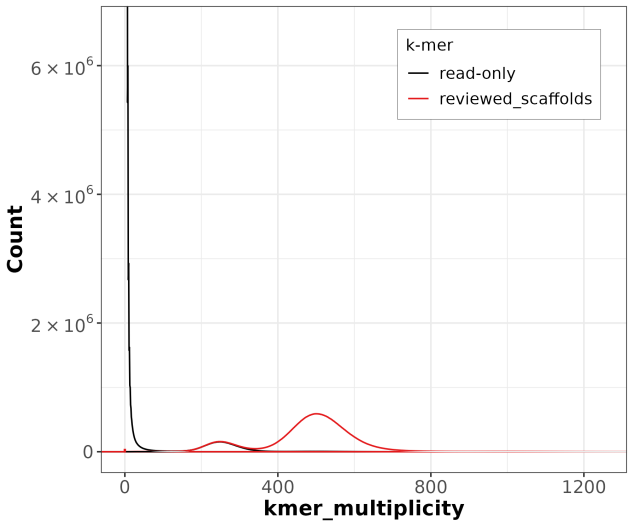


collapsed [LINK]

K-mer spectra of curated assembly

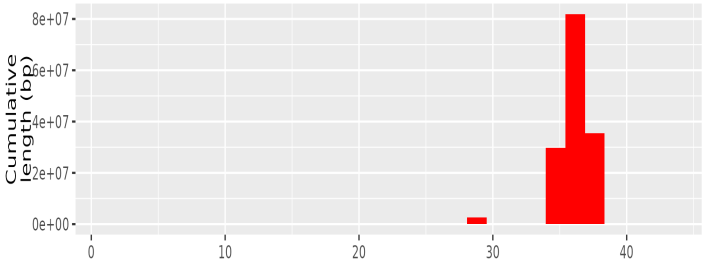


Distribution of k-mer counts per copy numbers found in asm

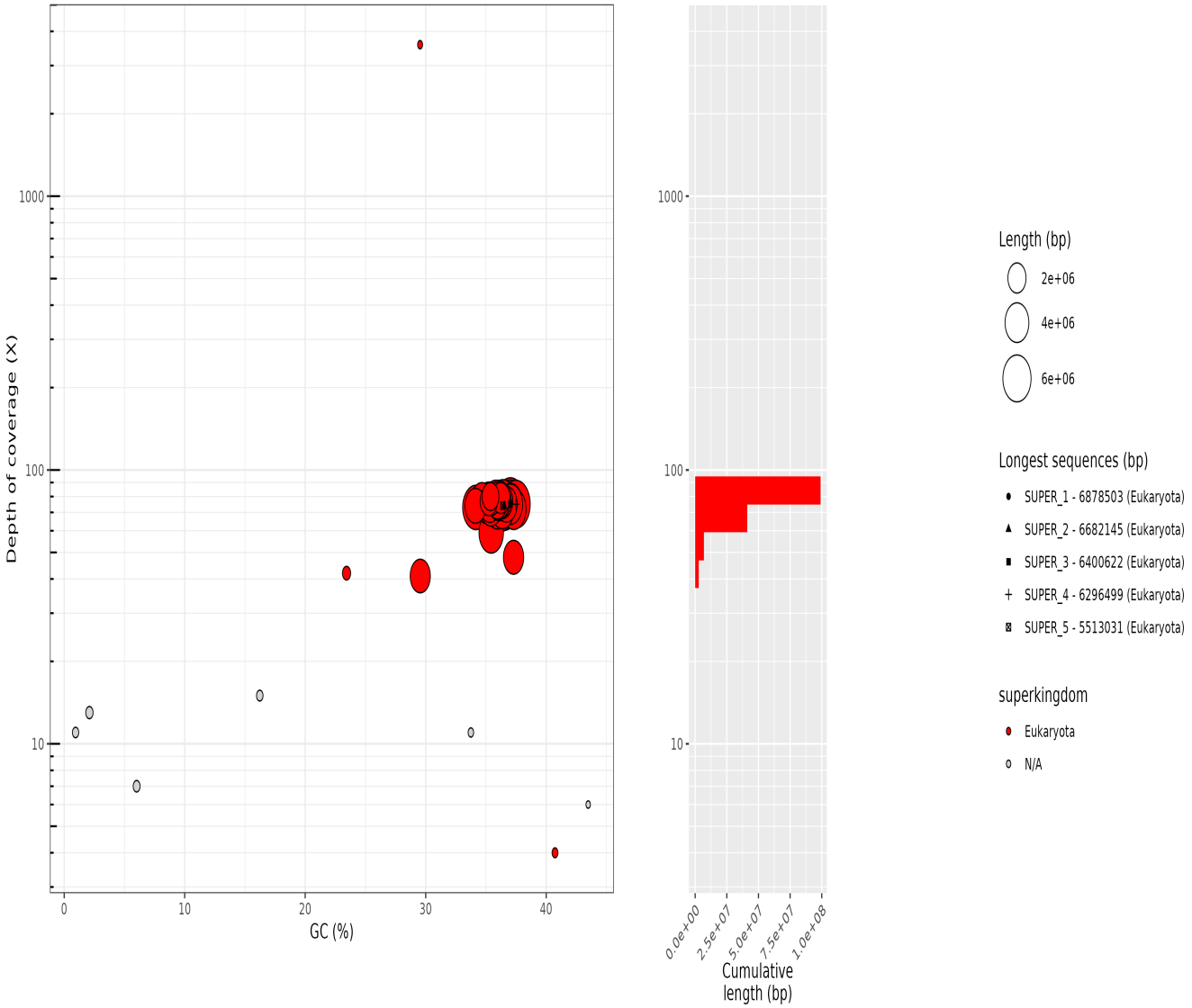


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



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	73	79

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: Jean-Marc Aury
Affiliation: Genoscope

Date and time: 2026-08-01 00:08:39 CEST