ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	2996
ToLID	**uoHydFoet1**
Species	Hydrurus foetidus
Class	Chrysophyceae
Order	Hydrurales

Genome Traits	Expected	Observed
Haploid size (bp)	188,808,469	208,258,979
Haploid Number	4 (source: ancestor)	40
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 5.6.Q58

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 pri
. BUSCO single copy value is less than 90% for pri
. Assembly length loss > 3% for pri
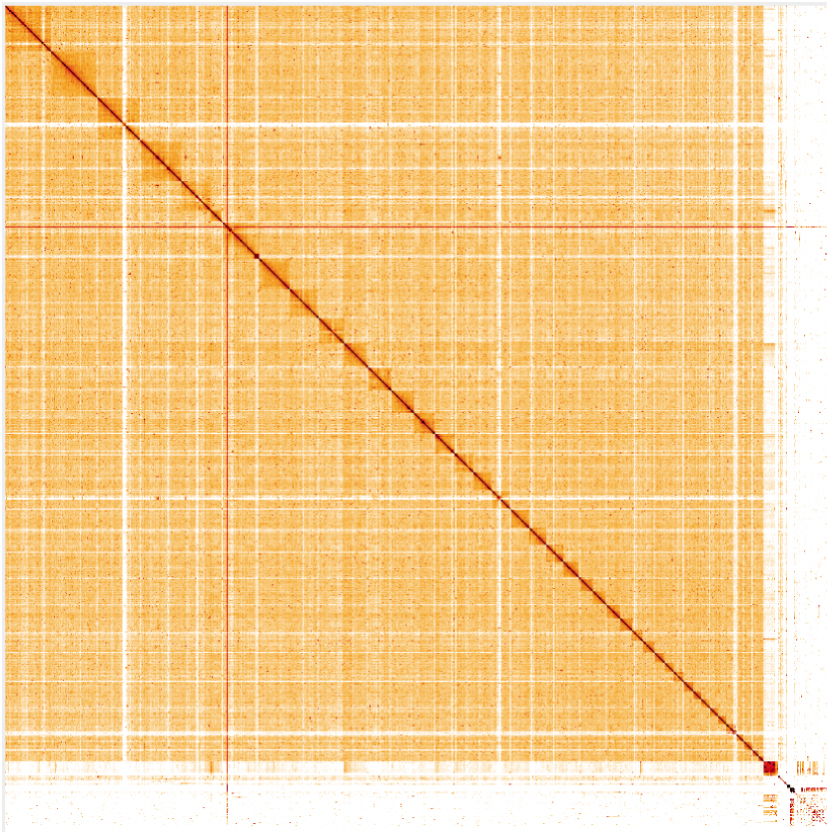. More than 1000 gaps/Gbp for pri

Curator notes

. Interventions/Gb: Many
. Contamination notes: "71 bacterial sequences removed"
. Other observations: "This was a real nightmare. Purging within hifiasm and purge-dups completely failed, so I removed a lot of haplotigs manually (about 2/3 of the genome). There seem to be 40 chromosomes and estimates in GoaT range from 2 to 60+. HiC map was also not of great quality"

Quality metrics table

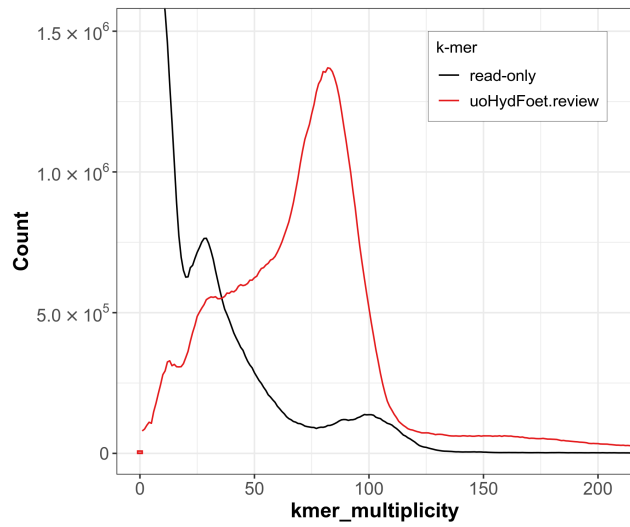| Metrics	Pre-curation pri	Curated pri
Total bp	571,637,880	208,258,979
GC %	43.71	43.44
Gaps/Gbp	1,978.53	2,765.79
Total gap bp	226,200	90,029
Scaffolds	681	390
Scaffold N50	5,370,286	5,213,484
Scaffold L50	35	13
Scaffold L90	124	38
Contigs	1,812	966
Contig N50	700,182	804,214
Contig L50	236	76
Contig L90	867	317
QV	59.8696	58.8157
Kmer compl.		75.599
BUSCO sing.	1.1%	87.5%
BUSCO dupl.	92.8%	4.3%
BUSCO frag.	2.9%	3.3%
BUSCO miss.	3.2%	4.9%

BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: stramenopiles_odb12 (genomes:55, BUSCOs:697)
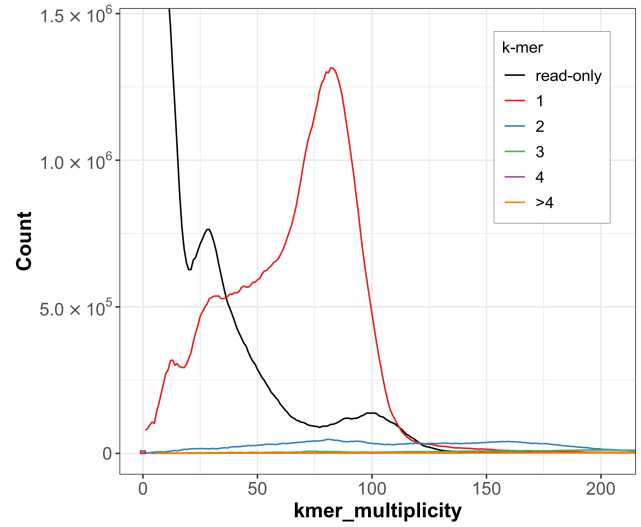
HiC contact map of curated assembly

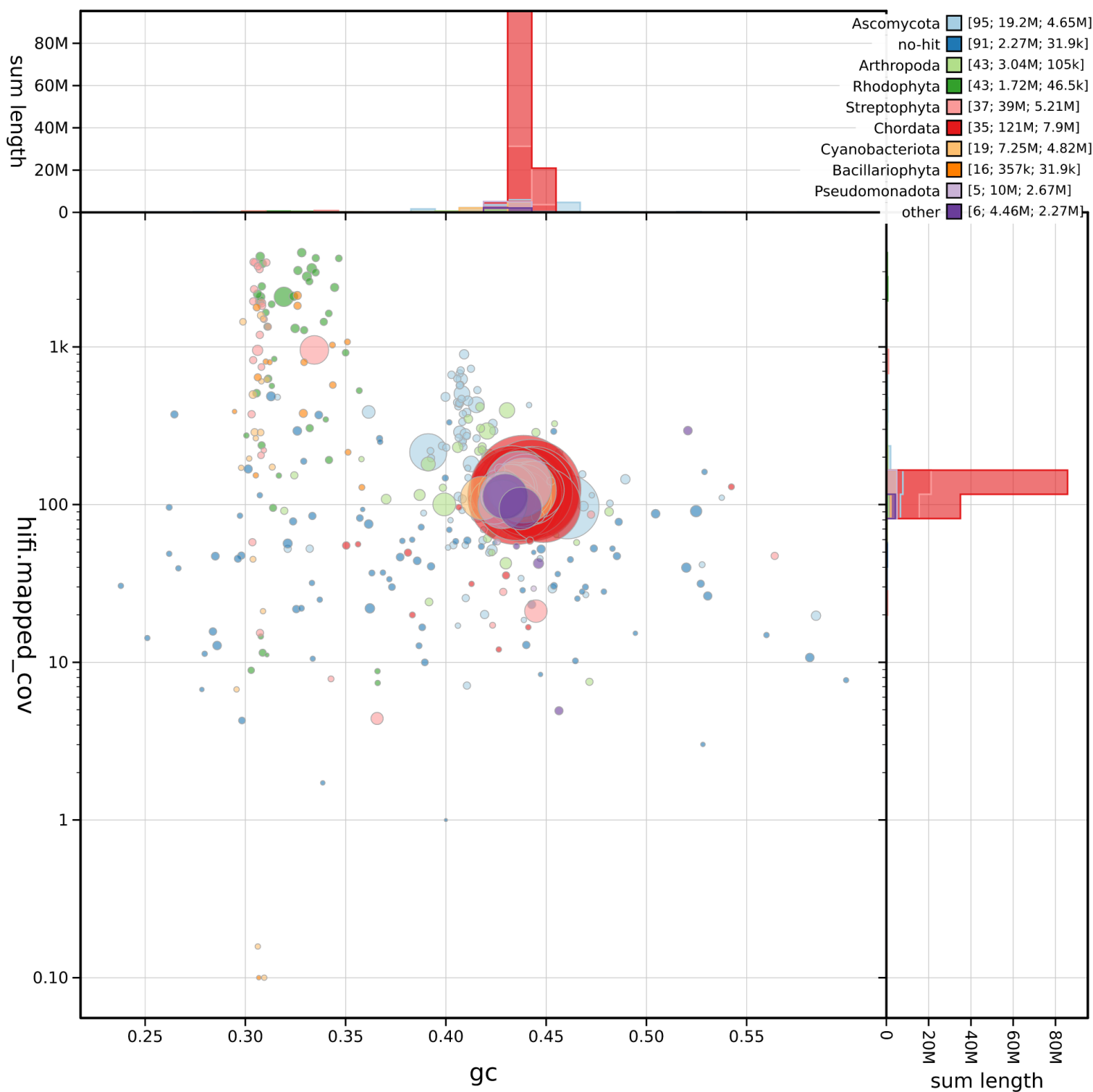**pri** [LINK]

K-mer spectra of curated assembly

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

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

Post-curation contamination screening

Ascomycota		[95; 19.2M; 4.65M]
no-hit		[91; 2.27M; 31.9k]
Arthropoda		[43; 3.04M; 105k]
Rhodophyta		[43; 1.72M; 46.5k]
Streptophyta		[37; 39M; 5.21M]
Chordata		[35; 121M; 7.9M]
Cyanobacteriota		[19; 7.25M; 4.82M]
Bacillariophyta		[16; 357k; 31.9k]
Pseudomonadota		[5; 10M; 2.67M]
other		[6; 4.46M; 2.27M]

pri. 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	PacBio HiFi	HiC
Coverage	80X	135X

Assembly pipeline

- **hifiasm**
    |_ *ver:* 0.25.0-r726
    |_ *key param:* NA
- **fcs-gx**
    |_ *ver:* 0.5.4
    |_ *key param:* NA
- **yahs**
    |_ *ver:* 1.2a1
    |_ *key param:* NA

Curation pipeline

- **PretextView**
    |_ *ver:* 1.0
    |_ *key param:* NA

Submitter: Tom Brown
Affiliation: IZW

Date and time: 2025-11-27 12:15:33 CET