

ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	3139642
ToLID	qdBraStyg1
Species	Brachydesmus stygivagus
Class	Diplopoda
Order	Polydesmida

Genome Traits	Expected	Observed
Haploid size (bp)	291,521,673	340,315,033
Haploid Number	12 (source: ancestor)	5
Ploidy	2 (source: ancestor)	2
Sample Sex	unknown	unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q64

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
- . Assembly length loss > 3% for pri
- . Not 90% of assembly in chromosomes for pri

Curator notes

- . Interventions/Gb: 1029
- . Contamination notes: "BlobTools identified more than 300 small scaffolds as potential contaminants, primarily corresponding to bacteria, with additional matches to fungi and chordates. With the exception of a couple, these contaminated scaffolds were removed from the assembly provided for review. The elimination of these putative contaminants did not seem to affect the BUSCO scores"
- . Other observations: " This extensively modified version of our curated assembly was made possible thanks to the expertise of the EAR reviewers, who recognized that many of the SUPER scaffolds were, in fact, chromosome arms. For the most part, the reviewers subsequently assembled these themselves. The heterochromatic regions between the purported chromosome arms were assembled with a low degree of confidence due to the inherent difficulties of assembling repetitive regions. This second review of the assembly addresses a few more issues that were spotted by the main reviewer"

Quality metrics table

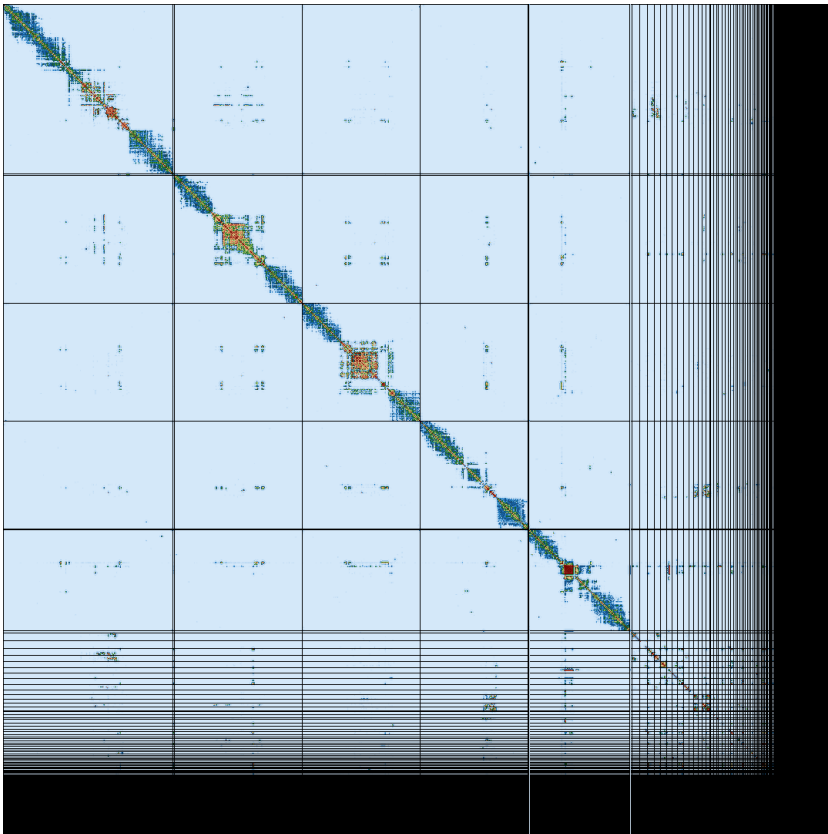
Metrics	Pre-curation pri	Curated pri
Total bp	357,806,435	340,315,033
GC %	32.79	32.4
Gaps/Gbp	433.2	808.07
Total gap bp	31,000	55,000
Scaffolds	685	280
Scaffold N50	3,203,691	48,226,045
Scaffold L50	23	3
Scaffold L90	176	39
Contigs	840	555
Contig N50	1,566,716	1,596,000
Contig L50	60	58
Contig L90	279	242
QV	63.7115	64.5587
Kmer compl.	86.2459	86.0456
BUSCO sing.	92.4%	96.3%
BUSCO dupl.	1.4%	1.0%
BUSCO frag.	2.8%	0.2%
BUSCO miss.	3.4%	2.5%

Warning! BUSCO versions or lineage datasets are not the same across results:

BUSCO: 5.5.0 (euk_genome_met, metaeuk) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

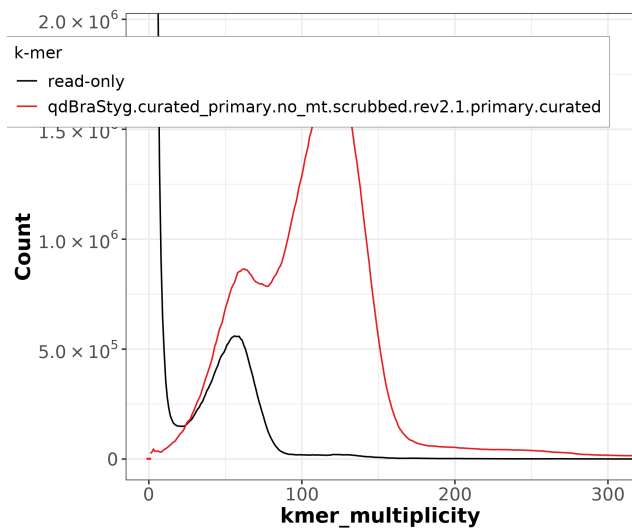
BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: arthropoda_odb10 (genomes:90, BUSCOs:1013)

HiC contact map of curated assembly

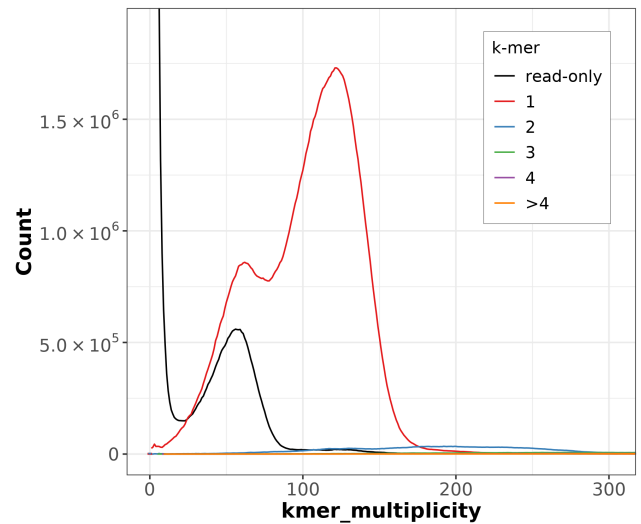


pri [\[LINK\]](#)

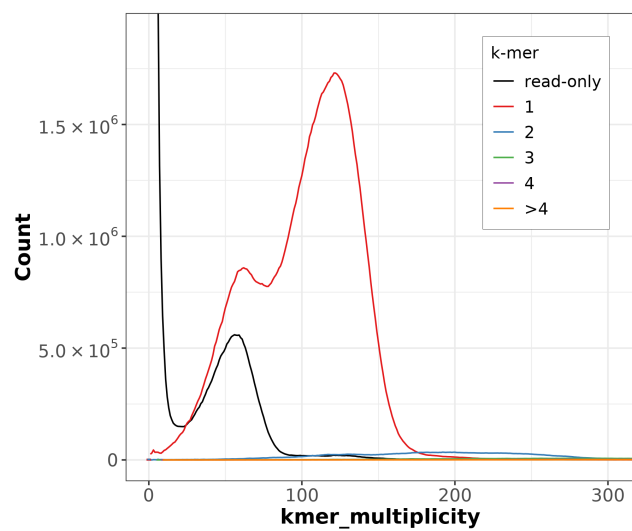
K-mer spectra of curated assembly



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

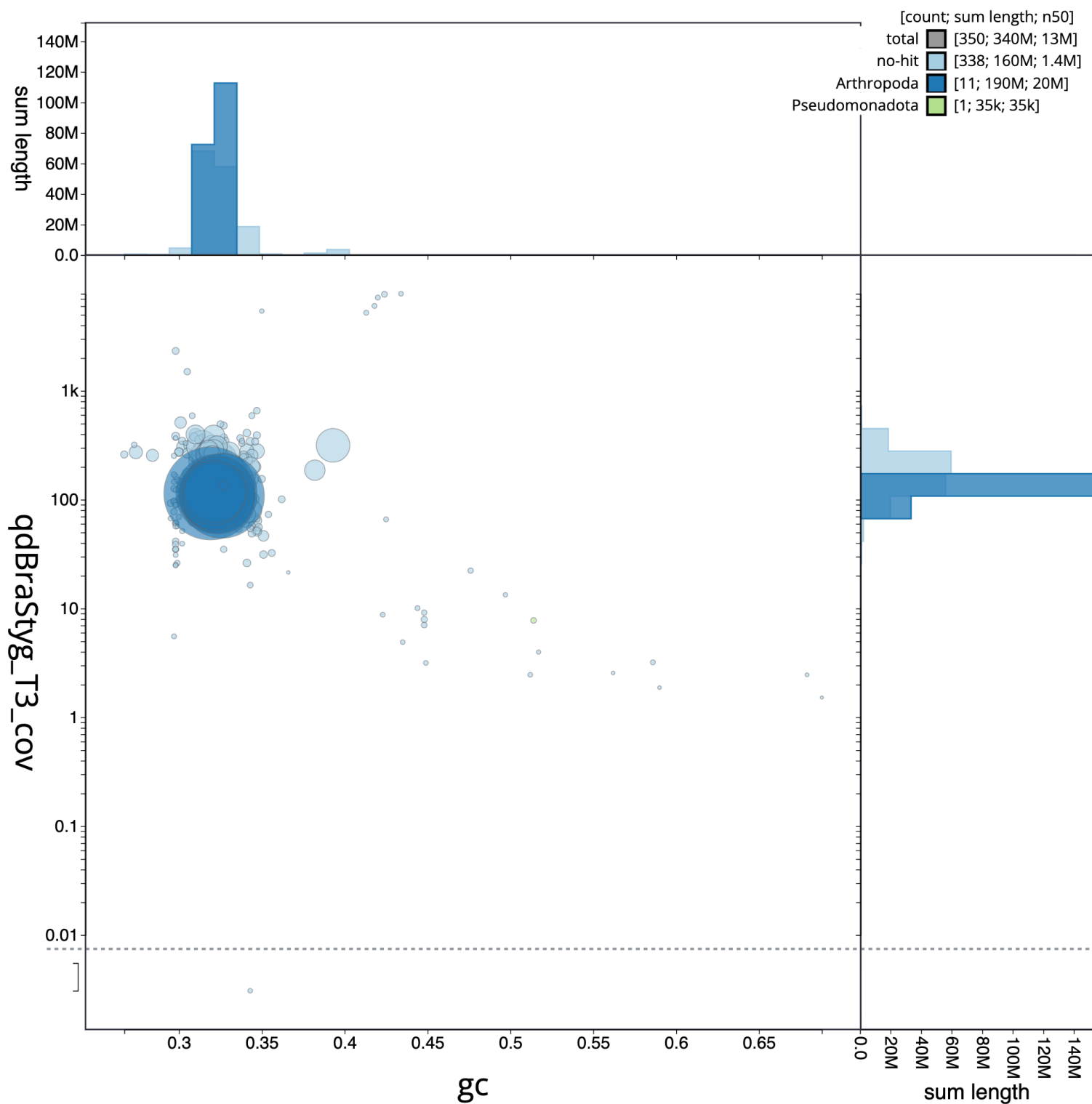


Distribution of k-mer counts per copy numbers found in asm



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

Post-curation contamination screening



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	OmniC
Coverage	100x	105x

Assembly pipeline

- **hifiasm**
 - |_ *ver*: 0.24.0-r702
 - |_ *key param*: NA
- **purge_dups**
 - |_ *ver*: 1.2.6
 - |_ *key param*: NA
- **YaHS**
 - |_ *ver*: 1.2a
 - |_ *key param*: NA
- **CLAWS**
 - |_ *ver*: 3.1
 - |_ *key param*: NA

Curation pipeline

- **PretextViewAI**
 - |_ *ver*: 1.0.5
 - |_ *key param*: NA
- **GRIT_Rapid**
 - |_ *ver*: 2.0
 - |_ *key param*: NA

Submitter: Francisco Camara

Affiliation: CNAG Barcelona

Date and time: 2026-02-20 15:59:40 CET