

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	439822
ToLID	<b>kaMicSqual</b>
Species	<i>Microcosmus squamiger</i>
Class	Asciidiacea
Order	Stolidobranchia

Genome Traits	Expected	Observed
Haploid size (bp)	192,802,479	188,886,422
Haploid Number	16 (source: ancestor)	10
Ploidy	2 (source: ancestor)	2
Sample Sex	unknown	unknown

## EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q61

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

## Curator notes

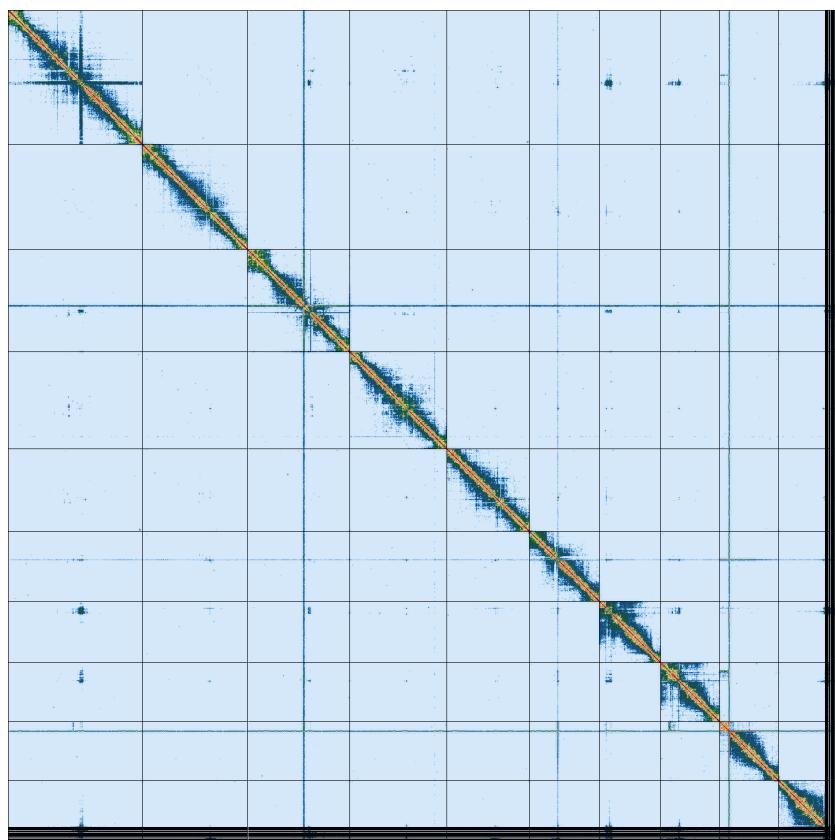
- . Interventions/Gb: 91
- . Contamination notes: "No contaminations were detected with FCS-GX."
- . Other observations: "Reads were assembled with hifiasm (yield: 218M, N50: 5.3M, 666 contigs). FCS-GX detected no contaminations. Purge\_dups removed 427 contigs (26Mb). HiC data is of a decent quality with 8.8M cis-reads pairs above 40Kb. The scaffolding was done in the manual curation steps."

## Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	192,689,173	188,886,422
GC %	43.3	43.43
Gaps/Gbp	0	402.36
Total gap bp	0	15,200
Scaffolds	239	108
Scaffold N50	6,661,360	22,013,698
Scaffold L50	11	4
Scaffold L90	37	9
Contigs	239	184
Contig N50	6,661,360	5,359,238
Contig L50	11	12
Contig L90	37	39
QV	60.6774	61.1407
Kmer compl.	66.7366	66.5603
BUSCO sing.	96.1%	96.3%
BUSCO dupl.	0.3%	0.3%
BUSCO frag.	1.6%	1.6%
BUSCO miss.	1.9%	1.8%

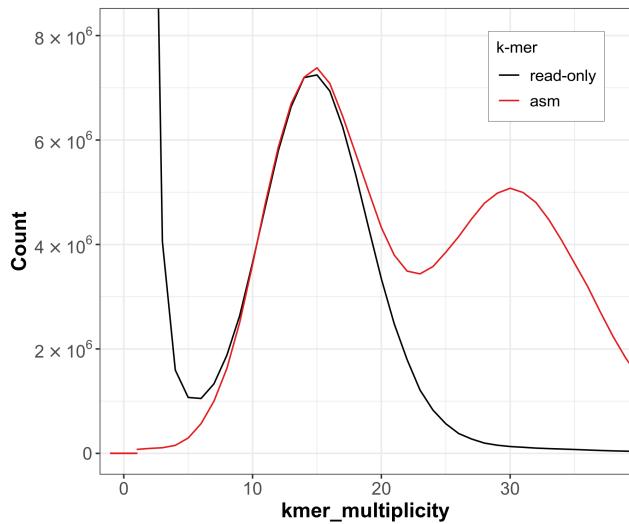
BUSCO: 5.8.3 (euk\_genome\_min, miniprot) / Lineage: metazoa\_odb12 (genomes:206, BUSCOs:672)

# HiC contact map of curated assembly

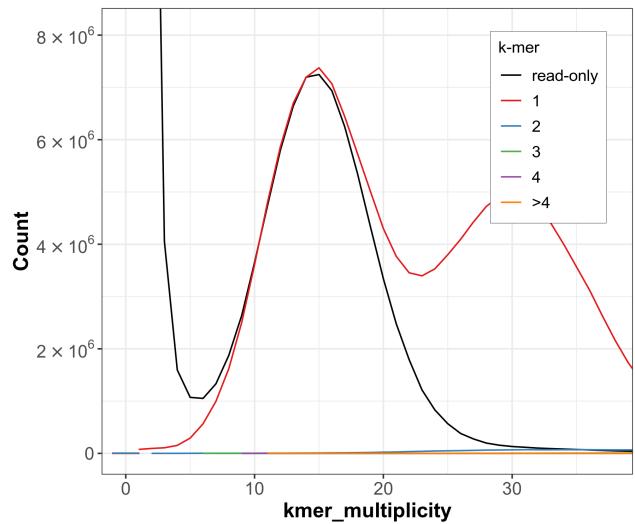


pri [LINK]

# K-mer spectra of curated assembly

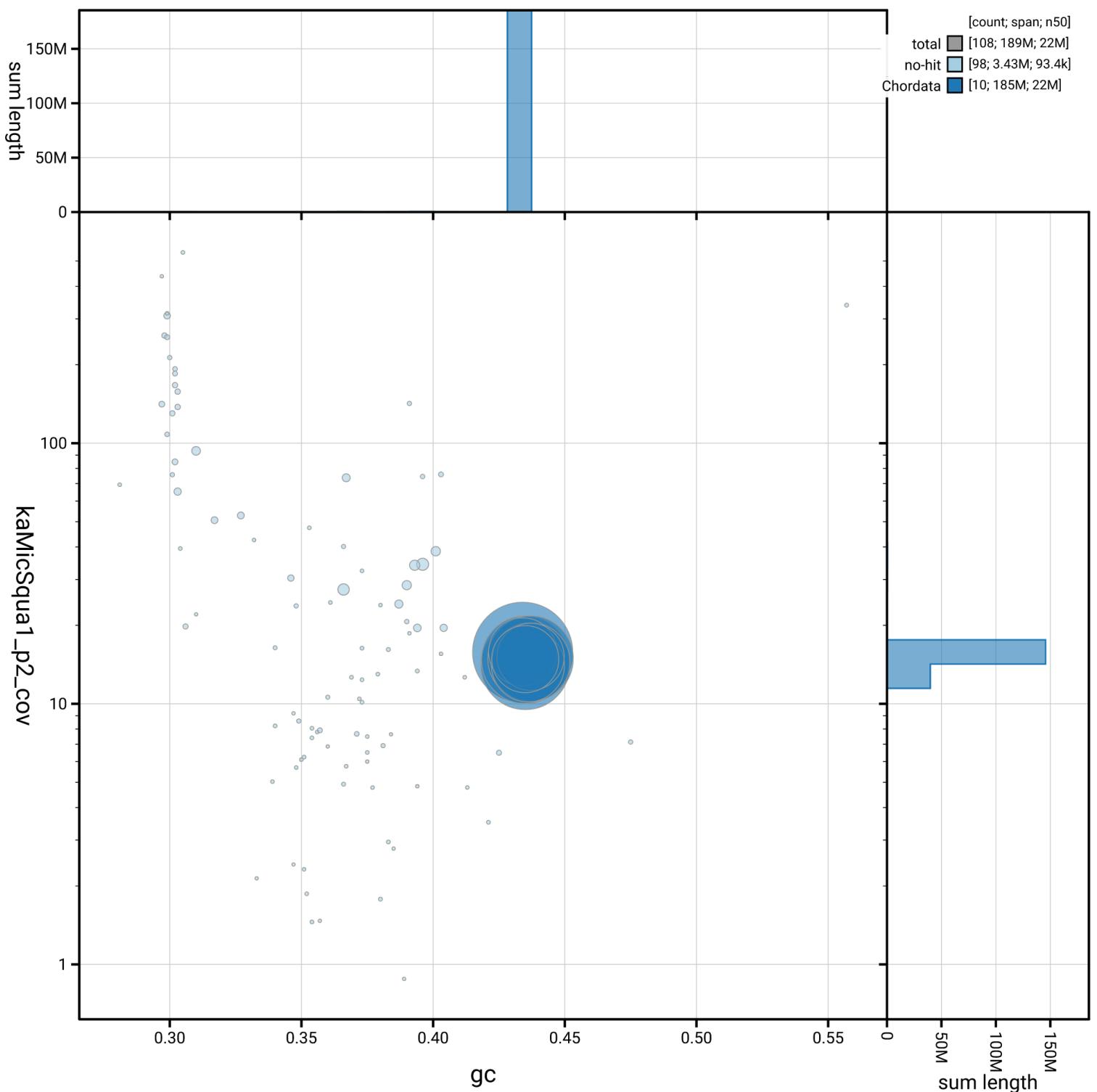


Distribution of k-mer counts coloured by their presence in reads/assemblies



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

# Post-curation contamination screening



**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	HiFi	HiC
Coverage	33x	176x

## Assembly pipeline

- **Hifiasm**
  - |\_ ver: 0.25.0-r726
  - |\_ key param: 13
- **purge\_dups**
  - |\_ ver: 1.2.6
  - |\_ key param: NA

## Curation pipeline

- **GRIT\_Rapid**
  - |\_ ver: 1a3d79a8
  - |\_ key param: NA
- **HiGlass**
  - |\_ ver: 0.10.4
  - |\_ key param: NA

Submitter: Martin Pippel

Affiliation: SciLifeLab

Date and time: 2025-11-13 06:32:18 CET