

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	2707950
ToLID	<b>qmNipGamm6</b>
Species	<i>Niphargus gammariformis</i>
Class	Malacostraca
Order	Amphipoda

Genome Traits	Expected	Observed
Haploid size (bp)	4,398,769,668	5,526,162,814
Haploid Number	13 (source: ancestor)	25
Ploidy	2 (source: ancestor)	2
Sample Sex	unknown	unknown

## EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.8.Q63

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

- . Observed Haploid size (bp) has >20% difference with Expected
- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for pri

## Curator notes

- . Interventions/Gb: 150
- . Contamination notes: "No presence of contaminants."
- . Other observations: "BUSCO score is not great, but I've seen it multiple times with amphipods getting low BUSCO scores. Lots of interventions, this one took me a while to finish."

# Quality metrics table

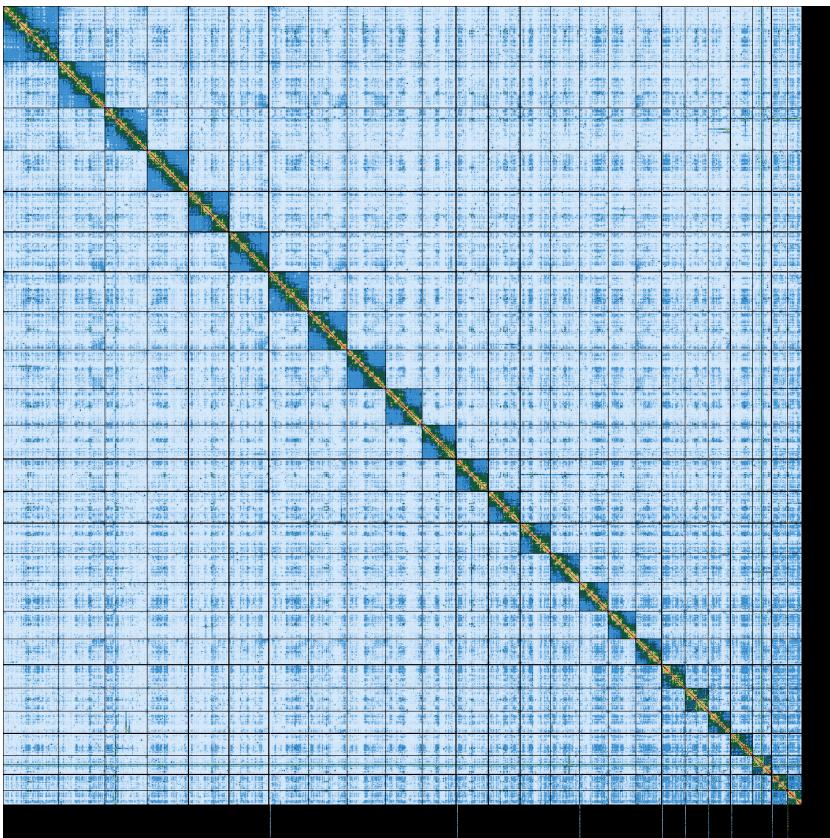
Metrics	Pre-curation pri	Curated pri
Total bp	5,584,042,361	5,526,162,814
GC %	44.02	44.03
Gaps/Gbp	634.49	727.81
Total gap bp	708,600	804,400
Scaffolds	2,739	2,181
Scaffold N50	45,827,757	242,103,678
Scaffold L50	35	10
Scaffold L90	221	23
Contigs	6,282	6,203
Contig N50	2,052,718	2,074,702
Contig L50	829	816
Contig L90	2,867	2,814
QV	22.9	63.0
Kmer compl.	97.23	77.29
BUSCO sing.	91.3%	91.0%
BUSCO dupl.	1.6%	1.7%
BUSCO frag.	1.0%	0.7%
BUSCO miss.	6.1%	6.6%

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

BUSCO: 5.8.2 (euk\_genome\_min, miniprot) / Lineage: metazoa\_odb10 (genomes:65, BUSCOs:954)

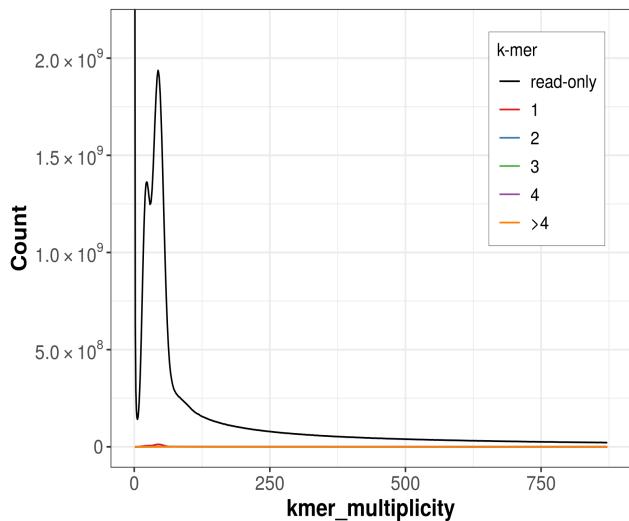
BUSCO: 6.0.0 (euk\_genome\_min, miniprot) / Lineage: metazoa\_odb10 (genomes:65, BUSCOs:954)

# HiC contact map of curated assembly

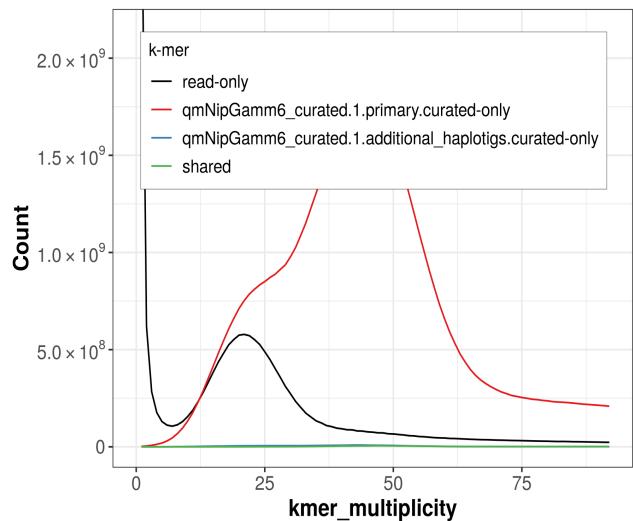


pri [LINK]

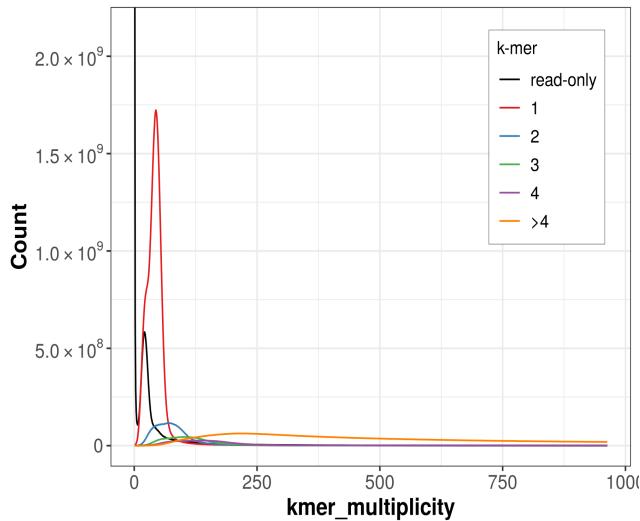
# K-mer spectra of curated assembly



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

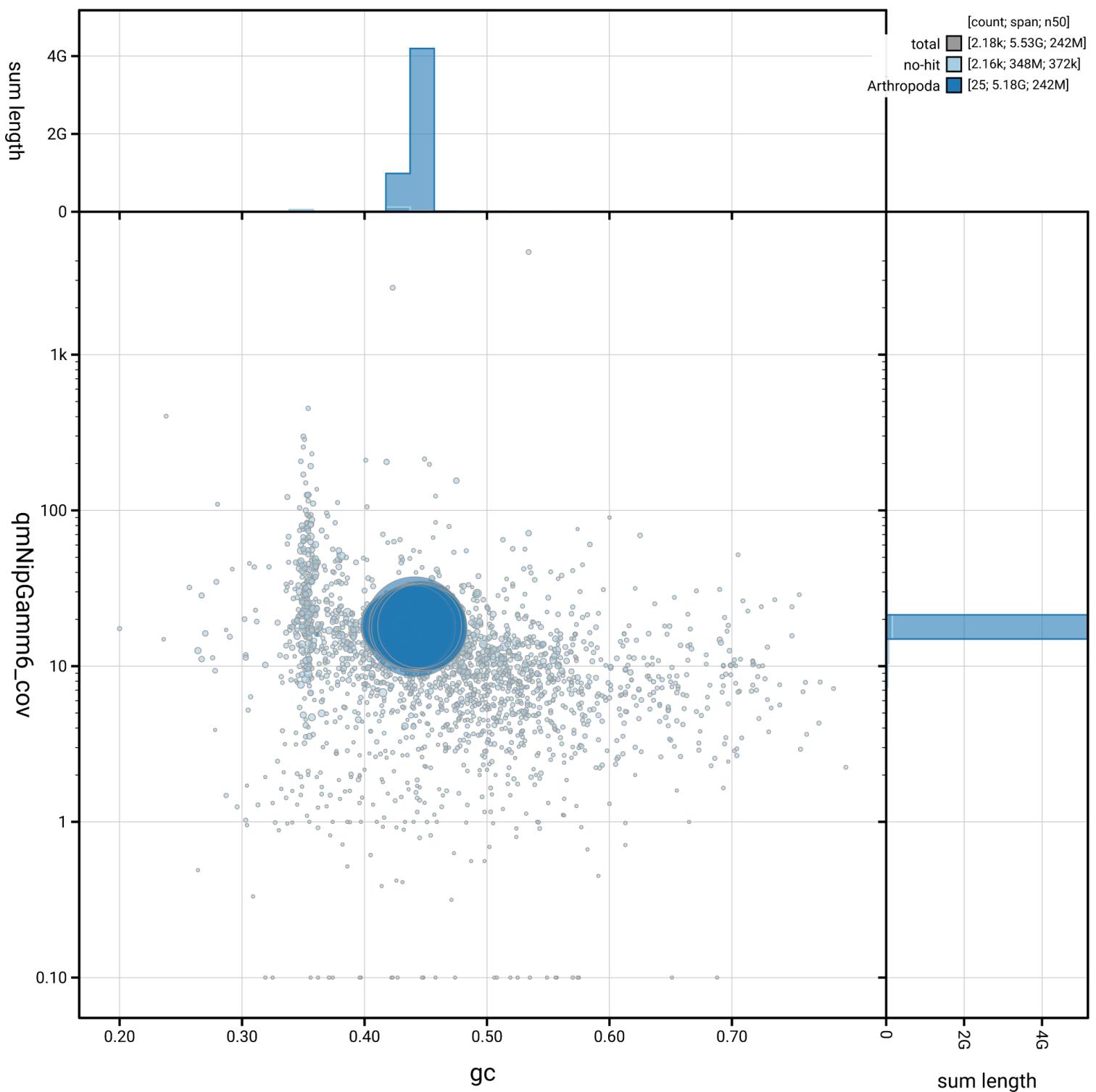


Distribution of k-mer counts coloured by their presence in reads/assemblies



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	OmniC
Coverage	36x	126x	

## Assembly pipeline

- **Hifiasm**
  - |\_ ver: 0.20.0-r639
  - |\_ key param: NA
- **purge\_dups**
  - |\_ ver: 1.2.5
  - |\_ key param: NA
- **YaHS**
  - |\_ ver: 1.2.2
  - |\_ key param: NA

## Curation pipeline

- **sanger-tol/curationpretext**
  - |\_ ver: v1.5.0-g5fd8e6e
  - |\_ key param: NA
- **Nextflow**
  - |\_ ver: 25.04.2
  - |\_ key param: NA
- **pretextmap**
  - |\_ ver: 0.1.9
  - |\_ key param: NA
- **HiGlass**
  - |\_ ver: 0.10.4
  - |\_ key param: NA
- **PretextViewAI**
  - |\_ ver: 1.0.5
  - |\_ key param: NA

Submitter: André E. R. Soares

Affiliation: SciLifeLab

Date and time: 2026-02-19 15:50:46 CET