

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

TxID	3229215
ToLID	<b>icSteAust9</b>
Species	Stenosis austini
Class	Insecta
Order	Coleoptera

Genome Traits	Expected	Observed
Haploid size (bp)	244,271,591	238,605,915
Haploid Number	10 (source: ancestor)	10
Ploidy	2 (source: ancestor)	2
Sample Sex	XX	XX

## EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q74

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

- . Kmer completeness value is less than 90 for collapsed

### Curator notes

- . Interventions/Gb:
- . Contamination notes: "FCS-GX found no contaminants. Tiara found a single bacterial fragment that was removed."
- . Other observations: "Sample is recorded as female, and the X chromosome was identified from synteny with Dailognatha quadricolis (icDaiQuad1.1). Assembly was performed with hifiiasm using HiC data with the --h1/--h2 parameters. The collapsed p\_ctg was selected based on N50/N90 and L50/L90 stats."

## Quality metrics table

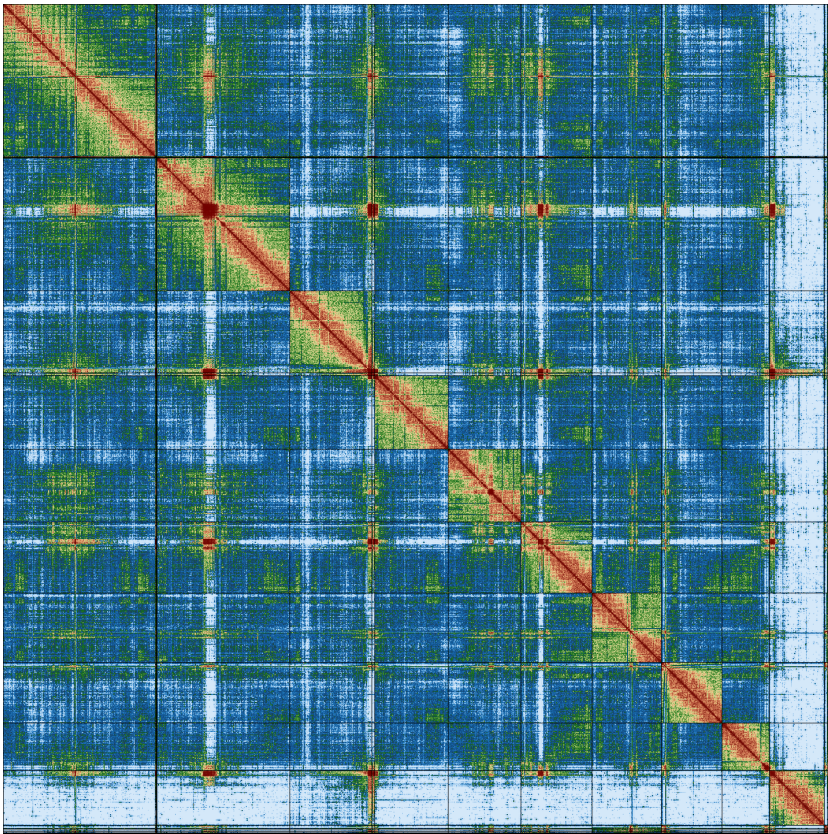
Metrics	Pre-curation collapsed	Curated collapsed
Total bp	240,587,923	238,605,915
GC %	34.47	34.48
Gaps/Gbp	29.1	88.01
Total gap bp	1,400	4,200
Scaffolds	46	30
Scaffold N50	21,101,840	21,101,840
Scaffold L50	6	4
Scaffold L90	12	9
Contigs	53	51
Contig N50	16,021,784	16,021,784
Contig L50	6	6
Contig L90	14	14
QV	72.2782	74.3038
Kmer compl.	79.2986	79.2301
BUSCO sing.	97.8%	98.5%
BUSCO dupl.	0.6%	0.6%
BUSCO frag.	0.4%	0.0%
BUSCO miss.	1.2%	0.9%

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

BUSCO: 5.8.3 (euk\_genome\_min, miniprot) / Lineage: endopterygota\_odb10 (genomes:56, BUSCOs:2124)

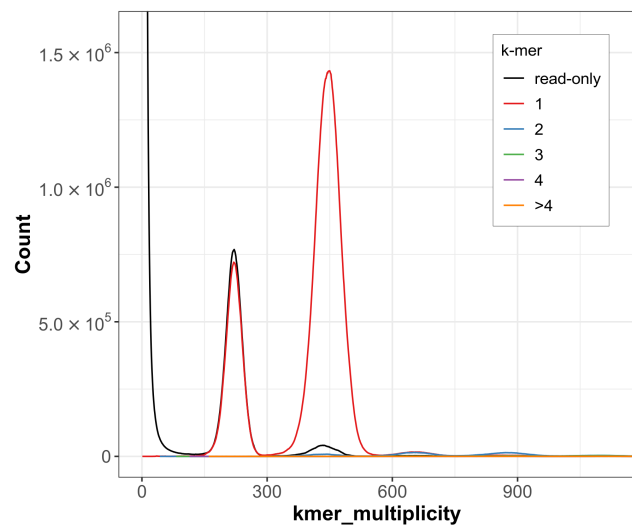
BUSCO: 6.0.0 (euk\_genome\_min, miniprot) / Lineage: endopterygota\_odb10 (genomes:56, BUSCOs:2124)

# HiC contact map of curated assembly



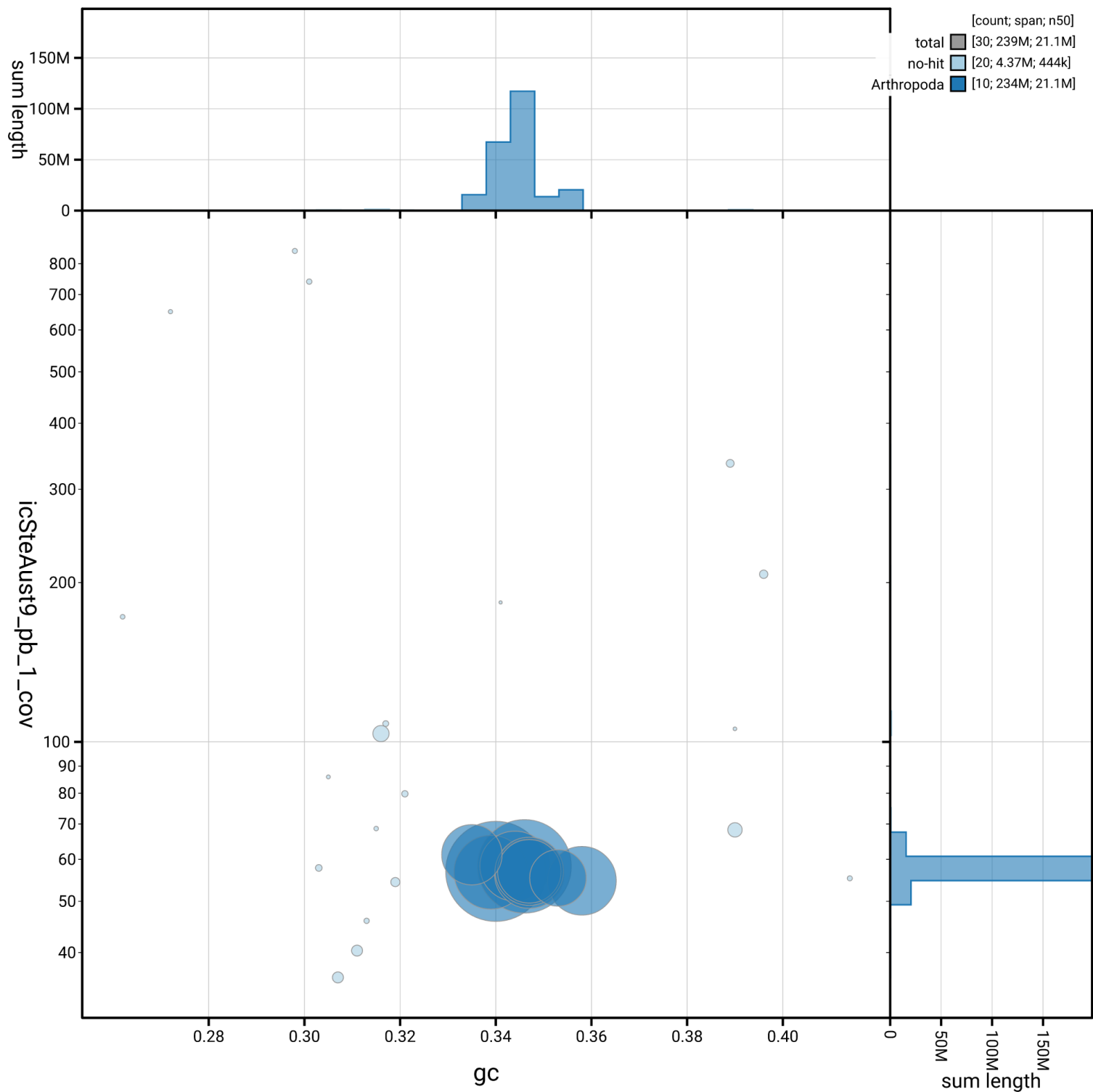
collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly



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

# Post-curation contamination screening



**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	PacBio HiFi	Arima HiC v2
Coverage	54x	375x

# Assembly pipeline

- **hifiasm-hic**
  - |\_ *ver*: 0.25.0-r726
  - |\_ *key param*: --h1
  - |\_ *key param*: --h2
- **purge\_dups**
  - |\_ *ver*: 1.2.5
  - |\_ *key param*: NA
- **tiara**
  - |\_ *ver*: 1.0.3
  - |\_ *key param*: NA
- **sanger-tol/blobtoolkit**
  - |\_ *ver*: 0.8.0
  - |\_ *key param*: NA
- **YaHS**
  - |\_ *ver*: 1.2a.2
  - |\_ *key param*: NA
- **NBIS/Earth-Biogenome-Project-pilot**
  - |\_ *ver*: 99c8937
  - |\_ *key param*: NA

# Curation pipeline

- **sanger-tol/curationpretext**
  - |\_ *ver*: 1.4.2
  - |\_ *key param*: NA
- **sanger-tol/agp-tpf-utils**
  - |\_ *ver*: 35ea45d
  - |\_ *key param*: NA

Submitter: Guilherme B. Dias

Affiliation: SciLifeLab

Date and time: 2025-09-15 20:04:18 CEST