

ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

| | |
|---------|-------------------|
| TxID | 1230023 |
| ToLID | xgCatTriz1 |
| Species | Cattania trizona |
| Class | Gastropoda |
| Order | Stylommatophora |

| Genome Traits | Expected | Observed |
|-------------------|-----------------------|---------------|
| Haploid size (bp) | 1,918,048,373 | 2,627,864,951 |
| Haploid Number | 25 (source: ancestor) | 0 |
| Ploidy | 2 (source: ancestor) | 2 |
| Sample Sex | H | |

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 5.5.Q55

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
- . BUSCO single copy value is less than 90% for pri
- . BUSCO duplicated value is more than 5% for pri
- . Not 90% of assembly in chromosomes for pri

Curator notes

- . Interventions/Gb: None
- . Contamination notes: "Contamination report for assembly labelled hap1: Total length of scaffolds removed: 13,447,483 (0.5 %); Scaffolds removed: 192 (0.5 %); Largest scaffold removed: (1,073,004); FCS-GX contaminant species (number of scaffolds; total length of scaffolds): Brevundimonas sp. UBA6550, a-proteobacteria (44; 2,592,848); Microbacterium sp. SCN 69-37, high GC Gram+ (31; 1,212,203); Microbacterium laevaniformans, high GC Gram+ (23; 878,836); Delftia acidovorans, b-proteobacteria (19; 5,781,184); Bosea thiooxidans, a-proteobacteria (9; 229,828); Buchnera aphidicola, g-proteobacteria (8; 522,861); Mesorhizobium denitrificans, a-proteobacteria (6; 275,371); Bosea spartocytisi, a-proteobacteria (4; 111,859); Bosea vaviloviae, a-proteobacteria (3; 78,435); Bosea robiniae, a-proteobacteria (2; 66,471); Agrobacterium cavarae, a-proteobacteria (2; 68,328); Brevundimonas bullata, a-proteobacteria (2; 68,602); Bosea psychrotolerans, a-proteobacteria (2; 53,947); Bosea lupini, a-proteobacteria (2; 42,954); Brucella pituitosa, a-proteobacteria (1; 22,456); Mesorhizobium escarrei, a-proteobacteria (1; 48,819); Mesorhizobium

australicum, a-proteobacteria (1; 54,453); Cupriavidus basilensis, b-proteobacteria (1; 362,638); Alkalispirochaeta americana, spirochetes (1; 27,040); Phyllobacterium endophyticum, a-proteobacteria (1; 26,875); Brevundimonas mediterranea, a-proteobacteria (1; 26,848); Mesorhizobium ciceri, a-proteobacteria (1; 15,795); Mesorhizobium camelthorni, a-proteobacteria (1; 19,173); Bradyrhizobium stylosanthis, a-proteobacteria (1; 25,919); Pseudorhizobium banfieldiae, a-proteobacteria (1; 72,374); Delftia tsuruhatensis, b-proteobacteria (1; 149,166); Devosia sp. 66-14, a-proteobacteria (1; 21,402); Aphidius gossypii, insects (1; 20,038); Chelatavorans intermedius, a-proteobacteria (1; 21,766); Microbacterium ginsengisoli, high GC Gram+ (1; 55,042); Bosea caraganae, a-proteobacteria (1; 41,744); Bosea lathyri, a-proteobacteria (1; 43,352); Rhizobium oryzihabitans, a-proteobacteria (1; 24,141); Brevundimonas sp. UBA7616, a-proteobacteria (1; 33,913); Mesorhizobium carmichaelinearum, a-proteobacteria (1; 22,888); Brevundimonas goettingensis, a-proteobacteria (1; 22,747); Mesorhizobium zhangyense, a-proteobacteria (1; 12,588); Bartonella quintana, a-proteobacteria (1; 13,717); Nitratireductor kimnyeongensis, a-proteobacteria (1; 35,698); Mesorhizobium amorphae, a-proteobacteria (1; 13,640); Brevundimonas sp. UBA7838, a-proteobacteria (1; 49,910); Brevundimonas sp. UBA4553, a-proteobacteria (1; 15,078); Mesorhizobium delmotii, a-proteobacteria (1; 29,073); Aminobacter aganoensis, a-proteobacteria (1; 32,220); Brevundimonas pondensis, a-proteobacteria (1; 43,125); Brucella grignonensis, a-proteobacteria (1; 12,153); Bosea minatitlanensis, a-proteobacteria (1; 19,804); Mitochondrion (2; 28,002); FCS-Adaptor (3; 123); Barcodes (11; 159)"

. Other observations: "Hifiasm assembly run in Hi-C phasing mode; Hi-C was not good enough to scaffold well; submitting as contig level assembly"

Quality metrics table

| Metrics | Pre-curation pri | Curated pri |
|--------------|---------------------|----------------|
| Total bp | 2,652,108,606 | 2,627,864,951 |
| GC % | 42.54 | 42.42 |
| Gaps/Gbp | 0 | 1.14 |
| Total gap bp | 0 | 123 |
| Scaffolds | 37,504 | 37,104 |
| Scaffold N50 | 111,993 | 112,113 |
| Scaffold L50 | 6,457 | 6,402 |
| Scaffold L90 | 24,055 | 23,790 |
| Contigs | 37,504 | 37,107 |
| Contig N50 | 111,993 | 112,113 |
| Contig L50 | 6,457 | 6,402 |
| Contig L90 | 24,055 | 23,790 |
| QV | 55.3 | 55.3 |
| Kmer compl. | 98.20 | 97.72 |
| BUSCO sing. | 71.4% | 71.4% |
| BUSCO dupl. | 12.2% | 12.1% |
| BUSCO frag. | 5.0% | 5.0% |
| BUSCO miss. | 11.4% | 11.4% |

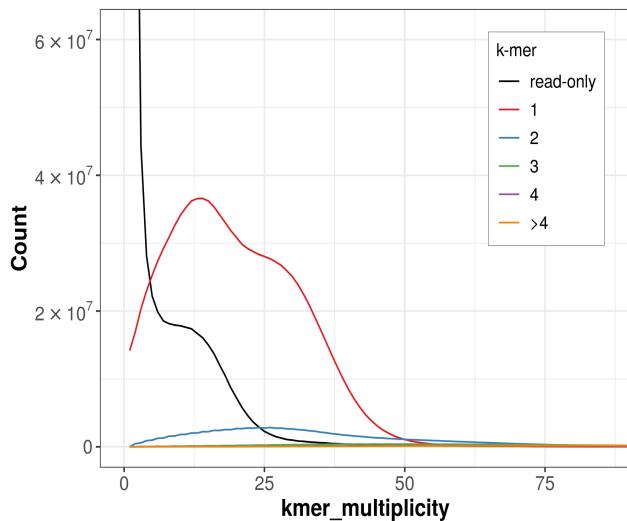
BUSCO 6.0.0 Lineage: mollusca_odb10 (genomes:7, BUSCOs:5295)

HiC contact map of curated assembly

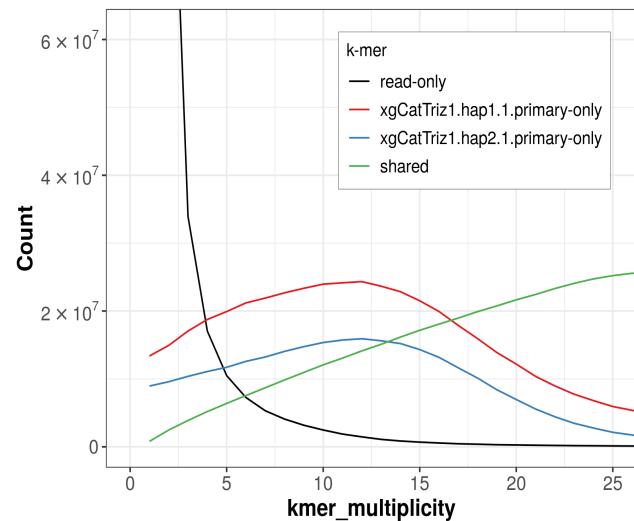
pri HiC PNG is missing!

pri File link is missing!

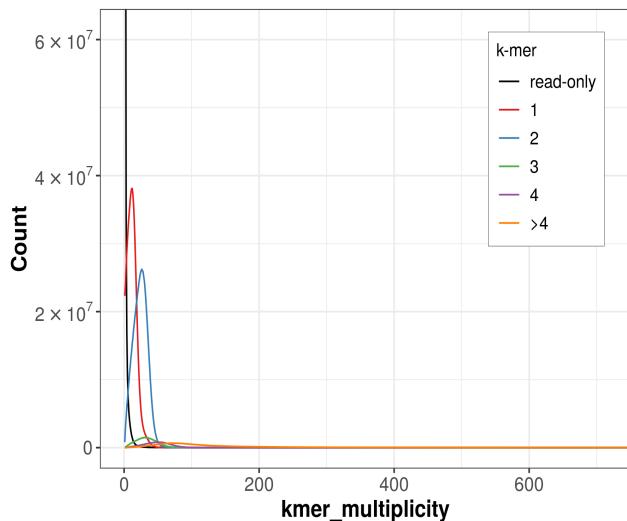
K-mer spectra of curated assembly



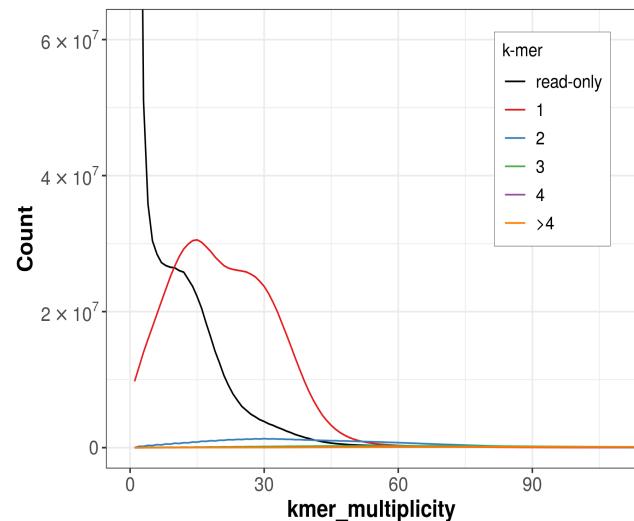
Distribution of k-mer counts per copy numbers found in (hapl.)



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

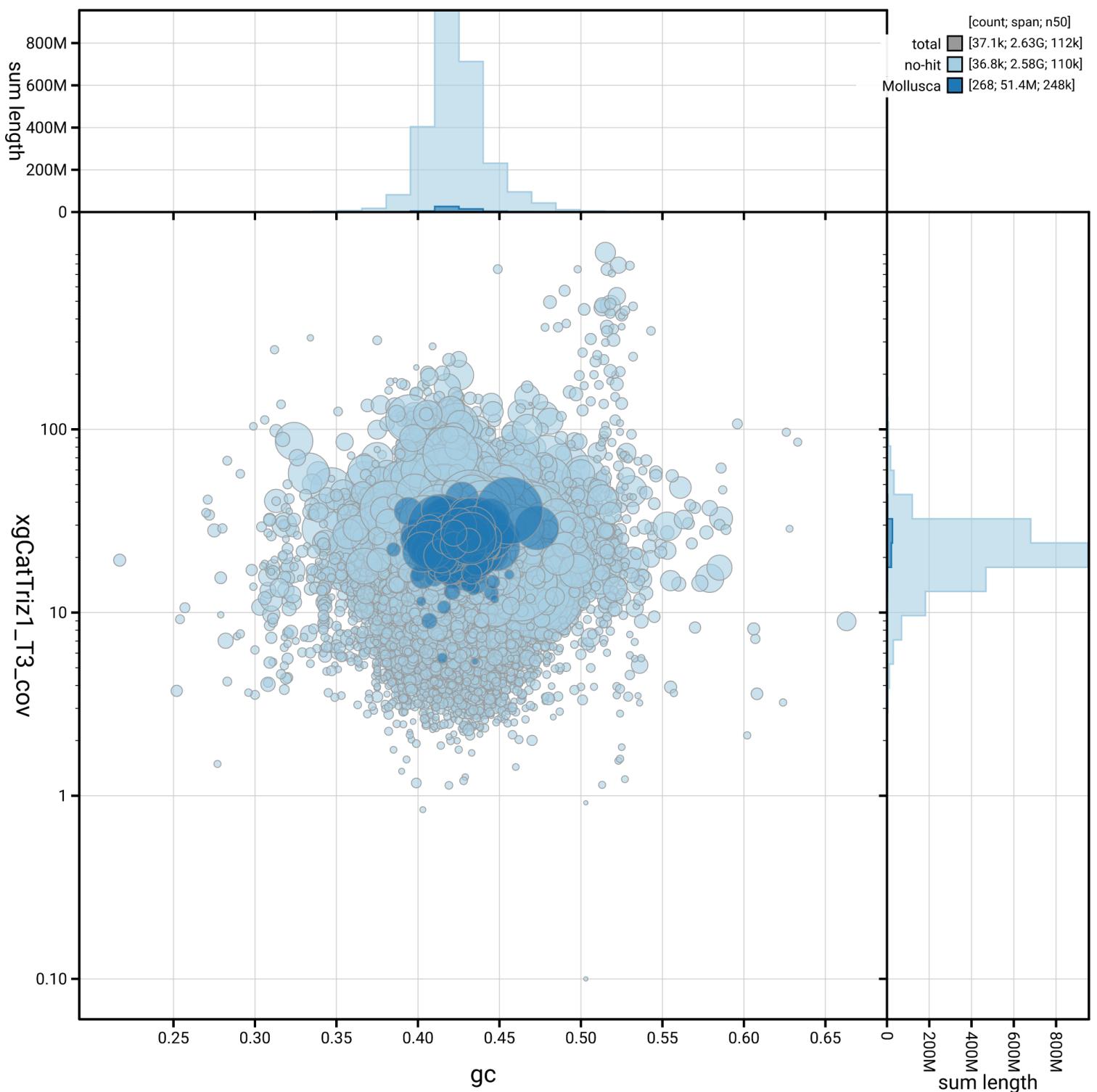


Distribution of k-mer counts per copy numbers found in asm (diplo.)



Distribution of k-mer counts per copy numbers found in **hap2** (hapl.)

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 | Arima v2 |
|----------|-------------|----------|
| Coverage | 30x | 670x |

Assembly pipeline

```
- 0.25.0-r726
|_ ver: NA
|_ key param: --h1/--h2
|_ key param: NA
```

Curation pipeline

```
- hifiasm
|_ ver: 0.25.0-r726
|_ key param: --h1/--h2
- Treeval
|_ ver: 1.2.1
|_ key param: NA
```

Submitter: Shane McCarthy

Affiliation: WSI

Date and time: 2026-02-28 22:04:23 CET