ERGA Assembly Report
v24.10.15

Tags: ERGA-BGE

TxID	207645
ToLID	**iyBomRudr4**
Species	Bombus ruderatus
Class	Insecta
Order	Hymenoptera

Genome Traits	Expected	Observed
Haploid size (bp)	244,749,913	262,740,167
Haploid Number	19 (source: ancestor)	19
Ploidy	1 (source: ancestor)	1
Sample Sex	X	X

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q71

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

. Assembly length loss > 3% for pri

Curator notes

. Interventions/Gb: 27
. Contamination notes: "No contaminations was detected with FCS-GX."
. Other observations: "PacBio reads were subsampled to 60X. Reads were assembled with hifiasm (yield: 306M, N50: 6.4Mb). FCS-GX detected no contaminations. Purge_dups removed 285 contigs (31.3Mb). HiC data is of a decent quality 41.7M UU read pairs with 15% +40Kb length information. Yahs was used to scaffold the contigs (yield: 275Mb, sN50: 15.7Mb). The curation revealed many large coverage drops (from 60X to 5-7X) at adjacent contigs within many chromosomes. Therefore all shrapnel contigs with the same repeat motif and coverage drops were marked as haplotiogs (removed contigs: 143). The current map still has coverage drops (from 60X to 10-15X). After discussing this issue in the assembly_curation slack channel, I did not trimm back these motifs. Apparently those satellites/heterochromatin regions can have up 75% lower coverage in Bombus species."

Quality metrics table

| Metrics	Pre-curation pri	Curated pri
Total bp	275,421,291	262,740,167
GC %	37.53	37.89
Gaps/Gbp	130.71	152.24
Total gap bp	7,200	5,900
Scaffolds	248	35
Scaffold N50	15,784,770	15,784,570
Scaffold L50	8	8
Scaffold L90	17	16
Contigs	284	75
Contig N50	9,300,000	9,401,705
Contig L50	13	12
Contig L90	44	37
QV	70.5123	71.8543
Kmer compl.	99.6941	99.6683
BUSCO sing.	98.3%	98.3%
BUSCO dupl.	0.2%	0.2%
BUSCO frag.	0.4%	0.4%
BUSCO miss.	1.0%	1.0%

BUSCO: 5.8.3 (euk_genome_min, miniprot) / Lineage: apoidea_odb12 (genomes:38, BUSCOs:6507)
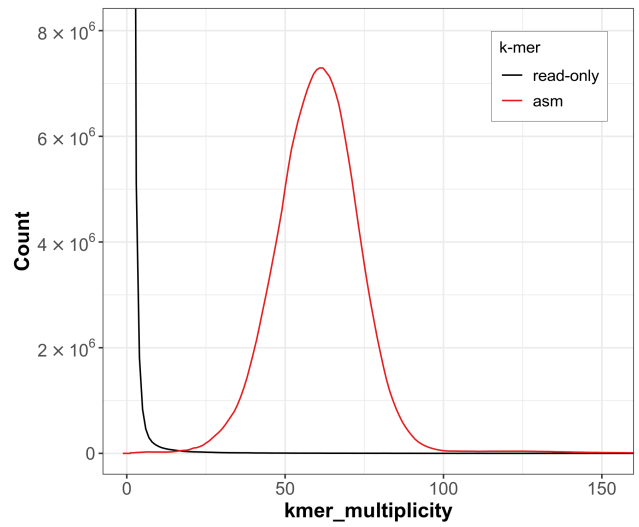
HiC contact map of curated assembly
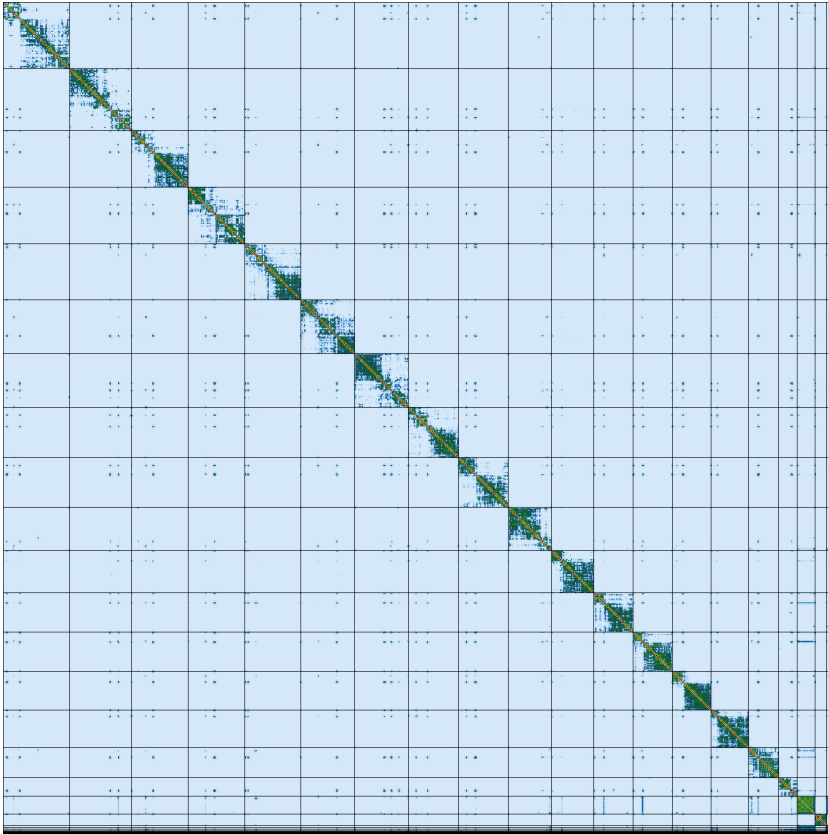
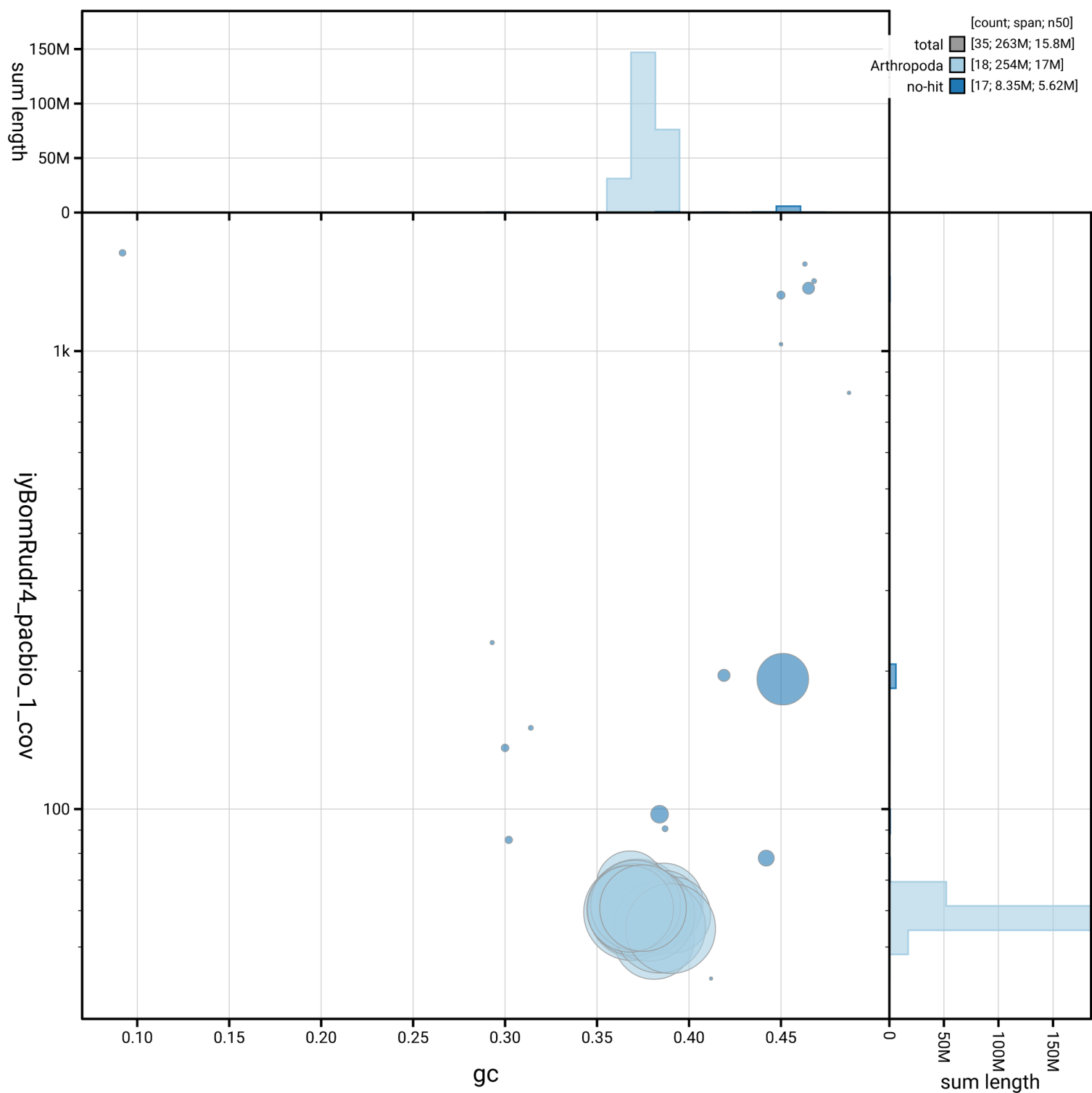**pri** [LINK]

K-mer spectra of curated assembly
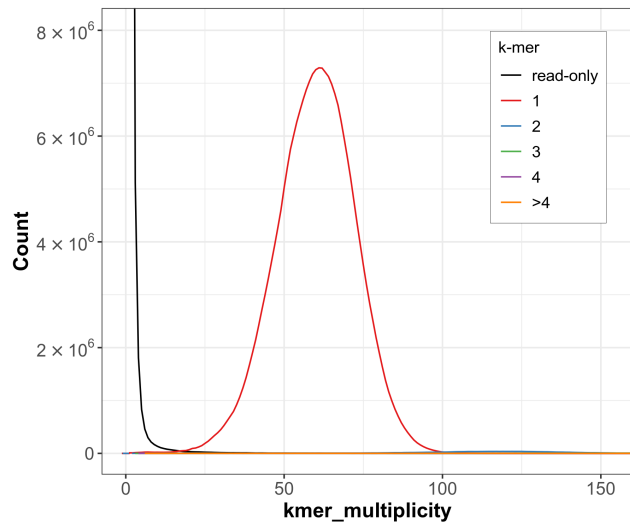
Distribution of k-mer counts per copy numbers found in asm

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

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	389x	133x

Assembly pipeline

- **Hifiasm**
    - |_ *ver:* 0.25.0-r726
    - |_ *key param:* l3
- **purge_dups**
    - |_ *ver:* 1.2.6
    - |_ *key param:* NA
- **YaHS**
    - |_ *ver:* 1.2.2
    - |_ *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: 2025-10-10 09:42:34 CEST