

ERGA Assembly Report

v24.04.03_beta

Tags: ERGA-BGE

TxID	270879
ToLID	ijBorWest1
Species	Boreus westwoodi
Class	Insecta
Order	Mecoptera

Genome Traits	Expected	Observed
Haploid size (bp)	529,682,445	657,981,536
Haploid Number	13 (source: ancestor)	24
Ploidy	2 (source: ancestor)	2
Sample Sex	F	F

EBP metrics summary and curation notes

Obtained EBP quality metric for pri: 6.7.Q54

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
- . More than 1000 gaps/Gbp for pri
- . Not 90% of assembly in chromosomes for pri

Curator notes

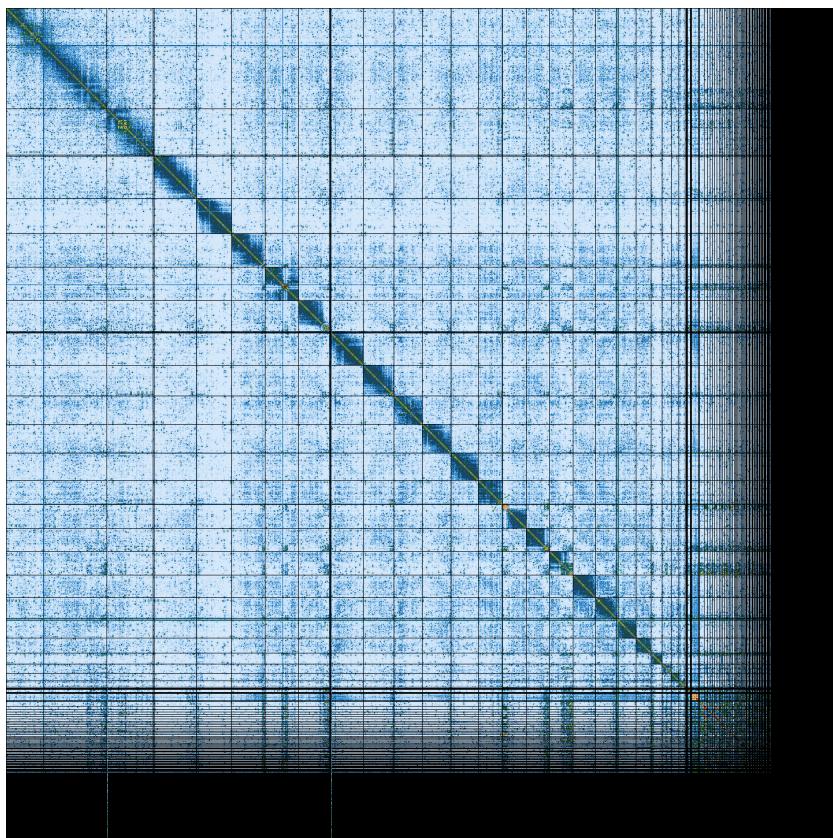
- . Interventions/Gb: 216
- . Contamination notes: "Contamination report for assembly labelled primary; Total length of scaffolds removed: 15 (0.0 %); Scaffolds removed: 0 (0.0 %); Largest scaffold removed: (15); FCS-GX contaminant species (number of scaffolds; total length of scaffolds):Barcodes (1; 15)"
- . Other observations: "hifiasm run in non-phasing mode; Hi-C is from 3 different individuals, ijBorWest2, ijBorWest4, ijBorWest5."

Quality metrics table

Metrics	Pre-curation pri	Curated pri
Total bp	667,801,017	657,981,536
GC %	33.59	33.6
Gaps/Gbp	1,758.01	1,759.93
Total gap bp	117,400	130,900
Scaffolds	764	739
Scaffold N50	17,691,776	21,986,458
Scaffold L50	12	11
Scaffold L90	94	69
Contigs	1,938	1,897
Contig N50	1,007,420	1,012,485
Contig L50	197	184
Contig L90	828	800
QV	54.0	54.0
Kmer compl.	99.24	65.72
BUSCO sing.	93.4%	95.2%
BUSCO dupl.	4.7%	3.3%
BUSCO frag.	0.4%	0.0%
BUSCO miss.	1.5%	1.5%

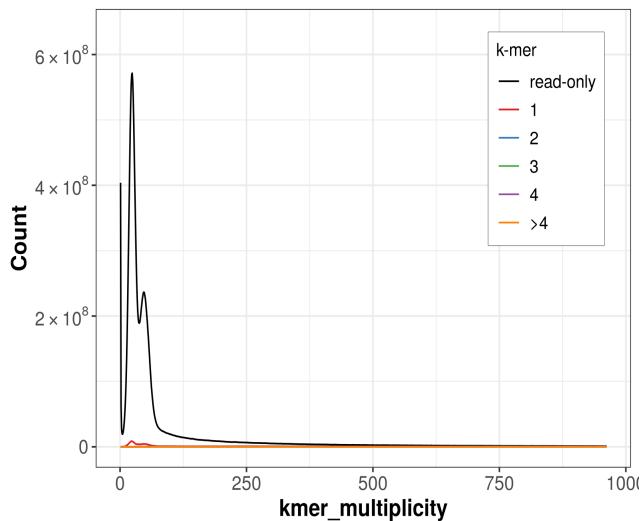
Warning: BUSCO versions or lineage datasets are not the same across results

HiC contact map of curated assembly

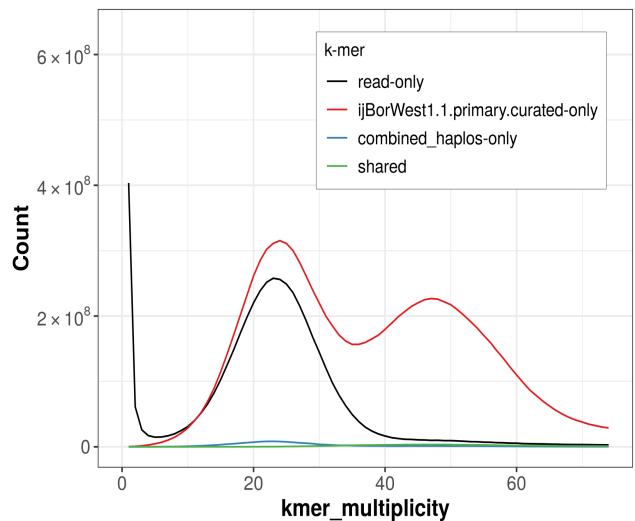


pri [LINK]

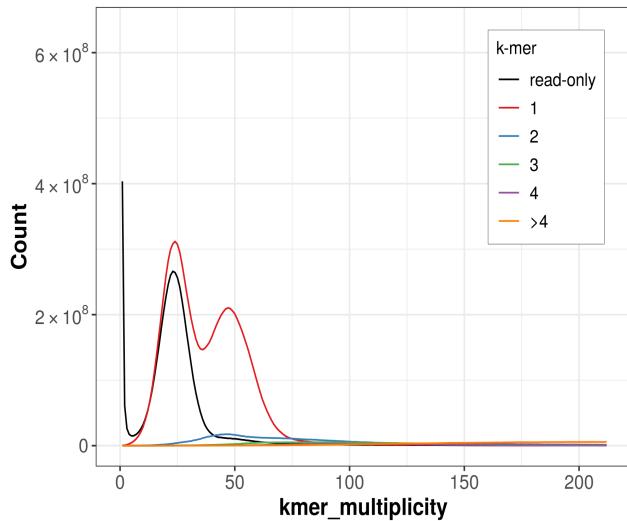
K-mer spectra of curated assembly



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

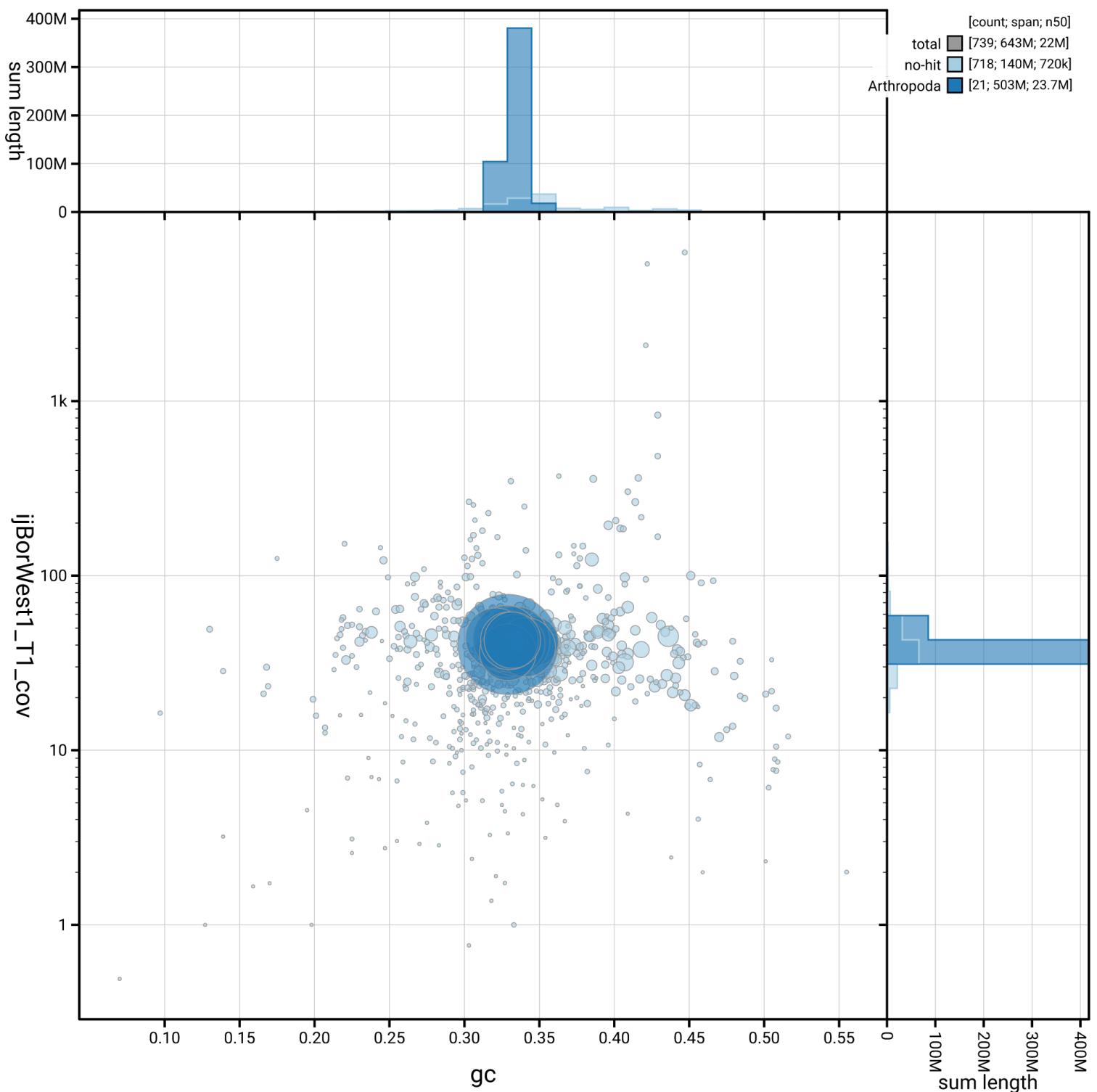


Distribution of k-mer counts coloured by their presence in reads/assemblies



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	PacBio HiFi	Arima v3
Coverage	46x	819x

Assembly pipeline

```
- hifiasm
  |_ ver: 0.19.8-r603
  |_ key param: --primary
- yahs
  |_ ver: 1.2.2
  |_ key param: NA
```

Curation pipeline

```
- hifiasm
  |_ ver: 0.19.8-r603
  |_ key param: --primary
- yahs
  |_ ver: 1.2.2
  |_ key param: NA
- TreeVal
  |_ ver: 1.2.1
  |_ key param: NA
```

Submitter: Dominic Absolon

Affiliation: WSI

Date and time: 2026-02-16 10:50:53 CET