

# ERGA Assembly Report

v24.10.15

Tags: ERGA-BGE

|         |                     |
|---------|---------------------|
| TxID    | 412090              |
| ToLID   | <b>mPipHan1</b>     |
| Species | Pipistrellus hanaki |
| Class   | Mammalia            |
| Order   | Chiroptera          |

| Genome Traits     | Expected              | Observed      |
|-------------------|-----------------------|---------------|
| Haploid size (bp) | 1,711,701,524         | 1,892,181,625 |
| Haploid Number    | 22 (source: ancestor) | 23            |
| Ploidy            | 2 (source: ancestor)  | 2             |
| Sample Sex        | Unknown               | Y             |

## EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 7.7.Q59

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

- . Observed Haploid Number is different from Expected
- . Observed sex is different from Sample sex

### Curator notes

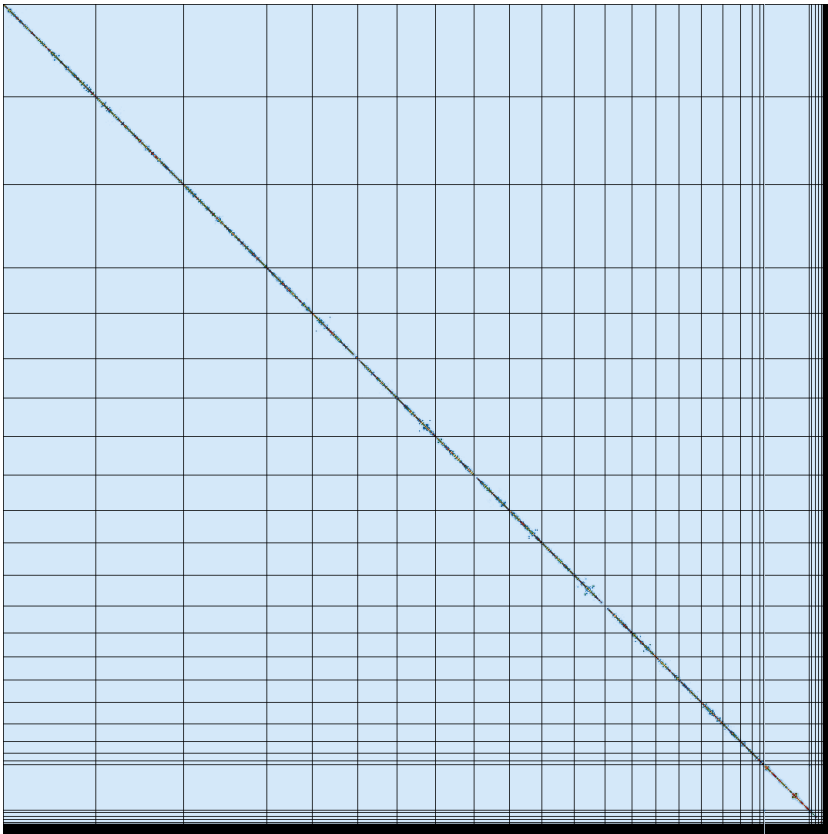
- . Interventions/Gb: 50
- . Contamination notes: ""
- . Other observations: "A misassembled 56 kb scaffold corresponding to mitochondria was deleted and a new mitochondrial sequence was added. "

## Quality metrics table

| Metrics      | Pre-curation<br>collapsed | Curated<br>collapsed |
|--------------|---------------------------|----------------------|
| Total bp     | 1,912,253,451             | 1,892,181,625        |
| GC %         | 42.55                     | 42.55                |
| Gaps/Gbp     | 149.04                    | 161.72               |
| Total gap bp | 28,500                    | 34,100               |
| Scaffolds    | 167                       | 141                  |
| Scaffold N50 | 91,412,170                | 89,115,682           |
| Scaffold L50 | 7                         | 7                    |
| Scaffold L90 | 18                        | 18                   |
| Contigs      | 452                       | 447                  |
| Contig N50   | 45,232,785                | 48,700,701           |
| Contig L50   | 16                        | 16                   |
| Contig L90   | 70                        | 69                   |
| QV           | 59.4962                   | 59.4872              |
| Kmer compl.  | 94.9997                   | 94.6588              |
| BUSCO sing.  | 93.6%                     | 92.7%                |
| BUSCO dupl.  | 1.7%                      | 1.6%                 |
| BUSCO frag.  | 0.5%                      | 0.5%                 |
| BUSCO miss.  | 4.2%                      | 5.2%                 |

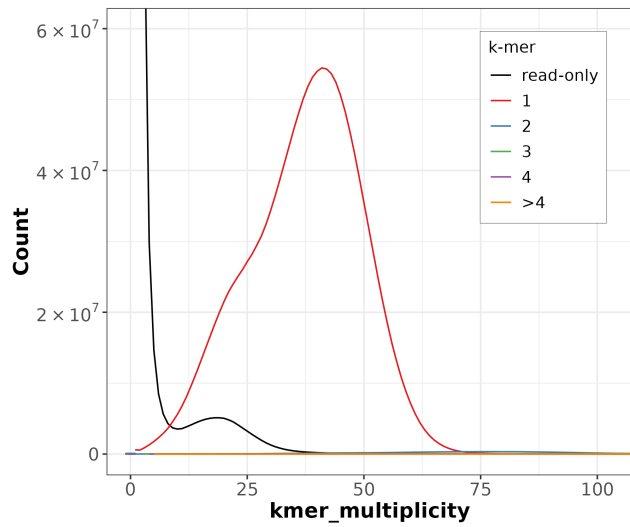
BUSCO: 5.4.3 (euk\_genome\_met, metaeuk) / Lineage: laurasiatheria\_odb10 (genomes:52, BUSCOs:12234)

# HiC contact map of curated assembly

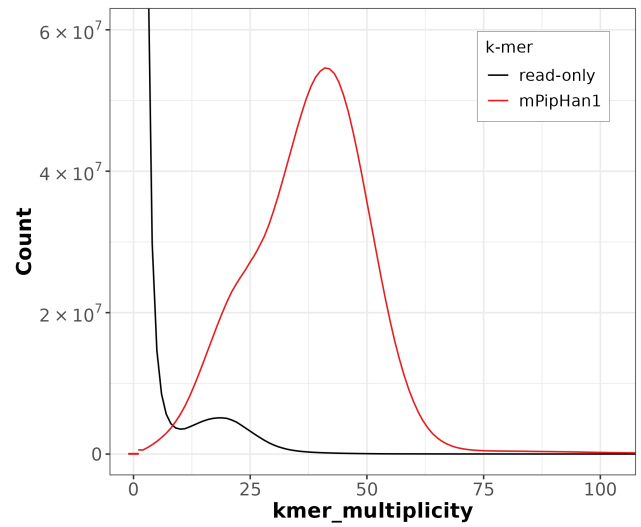


collapsed [\[LINK\]](#)

# K-mer spectra of curated assembly

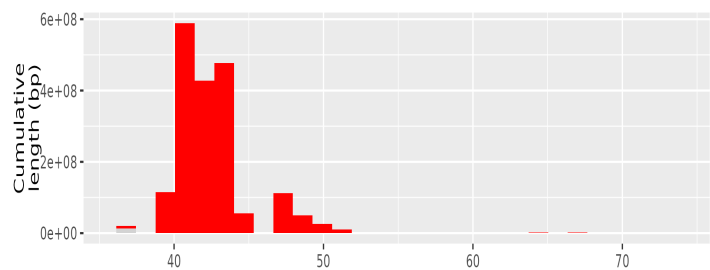


Distribution of k-mer counts per copy numbers found in asm

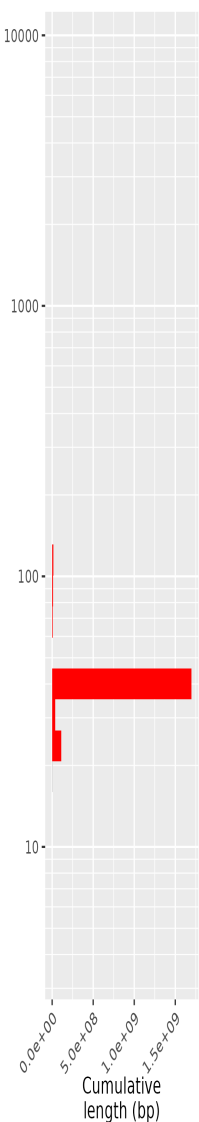
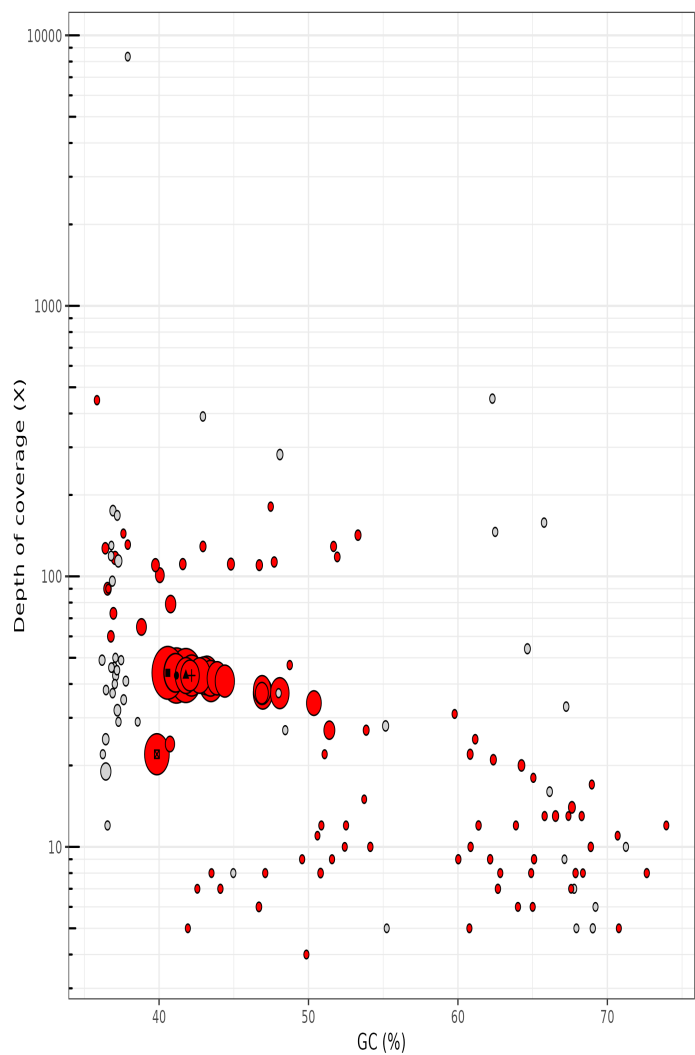


Distribution of k-mer counts coloured by their presence in reads/assemblies

# Post-curation contamination screening



TAPAs summary Graph



- Longest sequences (bp)
- SUPER\_1 - 210776044 (Eukaryota)
  - ▲ SUPER\_2 - 200298861 (Eukaryota)
  - SUPER\_3 - 189558047 (Eukaryota)
  - + SUPER\_4 - 103029890 (Eukaryota)
  - X - 102996988 (Eukaryota)

- superkingdom
- Eukaryota
  - N/A

- Length (bp)
- 5.0e+07
  - 1.0e+08
  - 1.5e+08
  - 2.0e+08

**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 |
|----------|-------------|-------|
| Coverage | 36          | 19    |

## Assembly pipeline

- **Hifiasm**
  - |\_ *ver*: 0.19.5-r593
  - |\_ *key param*: NA
- **purge\_dups**
  - |\_ *ver*: 1.2.5
  - |\_ *key param*: NA
- **YaHS**
  - |\_ *ver*: 1.2
  - |\_ *key param*: NA

## Curation pipeline

- **PretextMap**
  - |\_ *ver*: 0.1.9
  - |\_ *key param*: NA
- **PretextView**
  - |\_ *ver*: 0.2.5
  - |\_ *key param*: NA

Submitter: Sophie Mangenot

Affiliation: Genoscope

Date and time: 2025-01-30 14:43:06 CET