

ERGA Assembly Report

v24.10.15

Tags: ATLASEa[INVALID TAG]

TxID	104760
ToLID	jaPorPori1
Species	Porites porites
Class	Anthozoa
Order	Scleractinia

Genome Traits	Expected	Observed
Haploid size (bp)	520,537,908	601,828,541
Haploid Number	6 (source: ancestor)	13
Ploidy	2 (source: ancestor)	2
Sample Sex	Unknown	Unknown

EBP metrics summary and curation notes

Obtained EBP quality metric for collapsed: 6.7.Q46

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

- . Observed Haploid Number is different from Expected
- . Kmer completeness value is less than 90 for collapsed
- . BUSCO duplicated value is more than 5% for collapsed
- . Assembly length loss > 3% for collapsed
- . Not 90% of assembly in chromosomes for collapsed

Curator notes

- . Interventions/Gb: 125
- . Contamination notes: ""
- . Other observations: "The assembly of *Porites porites* (jaPorPori1) is based on 28X PacBio data and 136X Arima Hi-C data generated as part of the ATLASEa programme (<https://www.atlasea.fr>). The assembly process included the following steps: initial PacBio+HiC assembly generation with Hifiiasm, removal of contaminant sequences using Context, and Hi-C-based scaffolding with YaHS of each haplotype. In total, 22 contigs were identified as contaminants (bacterial, archaeal, or viral), totaling 0.631 Mb (with the largest being 0.066 Mb). Mitochondrial genome was assembled using OATK. Finally, Both haplotypes were analyzed and manually improved using Pretext. During manual curation, 3199 haplotypic regions and 26 contaminant sequences were removed, totaling 243.063Mb and 0.942Mb, respectively (with the largest being 2.277Mb and 0.121Mb). Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. "

Quality metrics table

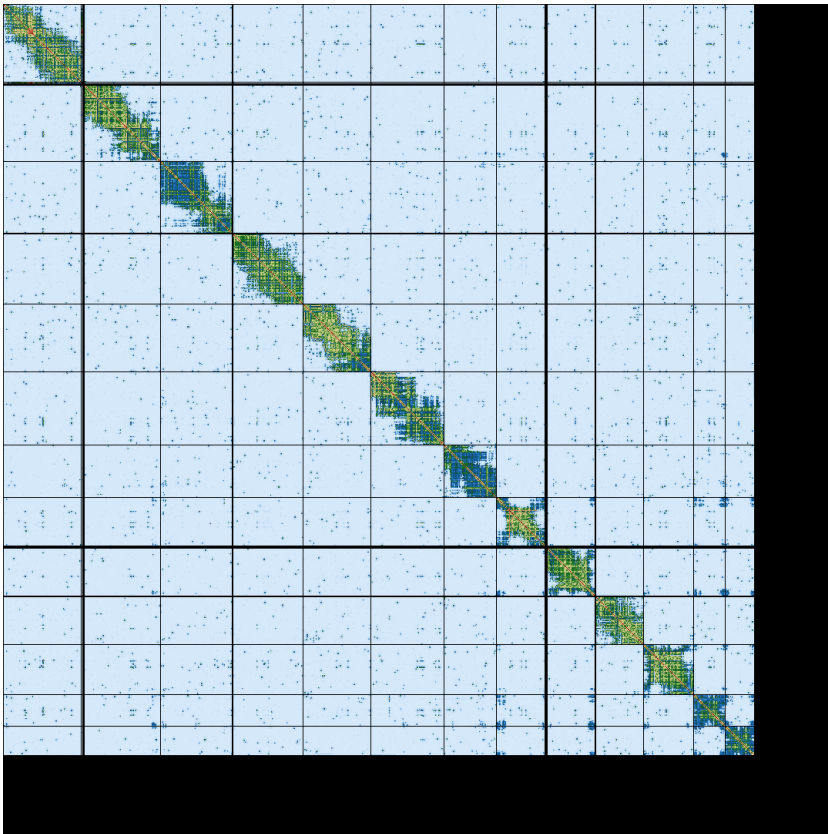
Metrics	Pre-curation collapsed	Curated collapsed
Total bp	710,424,897	601,828,541
GC %	38.95	38.93
Gaps/Gbp	335.01	701.2
Total gap bp	23,800	46,600
Scaffolds	549	1,209
Scaffold N50	34,947,368	48,753,877
Scaffold L50	7	6
Scaffold L90	40	18
Contigs	787	1,631
Contig N50	5,066,575	3,923,073
Contig L50	40	44
Contig L90	168	233
QV	47.7388	46.6269
Kmer compl.	77.7901	73.3827
BUSCO sing.	72.3%	92.8%
BUSCO dupl.	27.5%	5.1%
BUSCO frag.	0.1%	1.3%
BUSCO miss.	0.1%	0.8%

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

BUSCO: 6.0.0 (euk_genome_min, miniprot) / Lineage: anthozoa_odb12 (genomes:12, BUSCOs:3649)

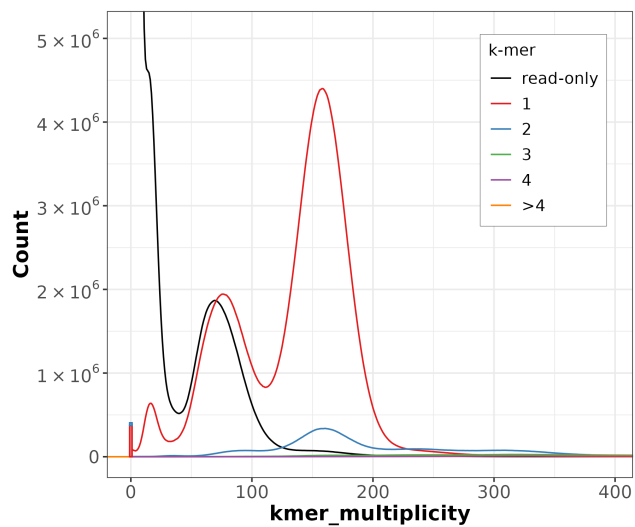
BUSCO: 5.8.2 (euk_genome_min, miniprot) / Lineage: anthozoa_odb12 (genomes:12, BUSCOs:3649)

HiC contact map of curated assembly

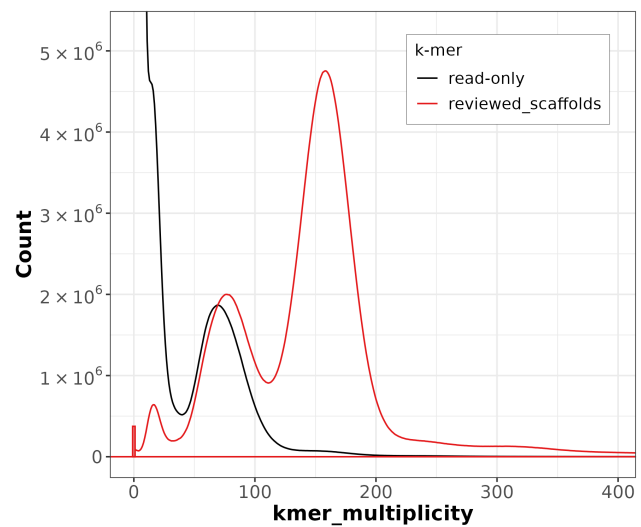


collapsed [\[LINK\]](#)

K-mer spectra of curated assembly

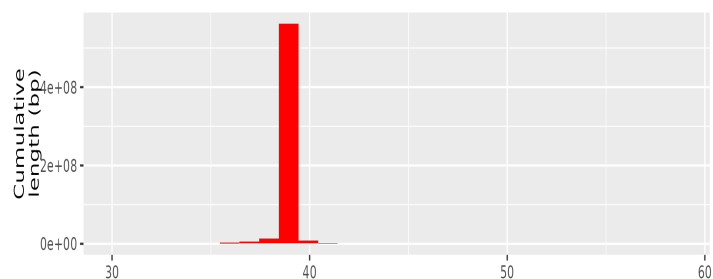


Distribution of k-mer counts per copy numbers found in asm

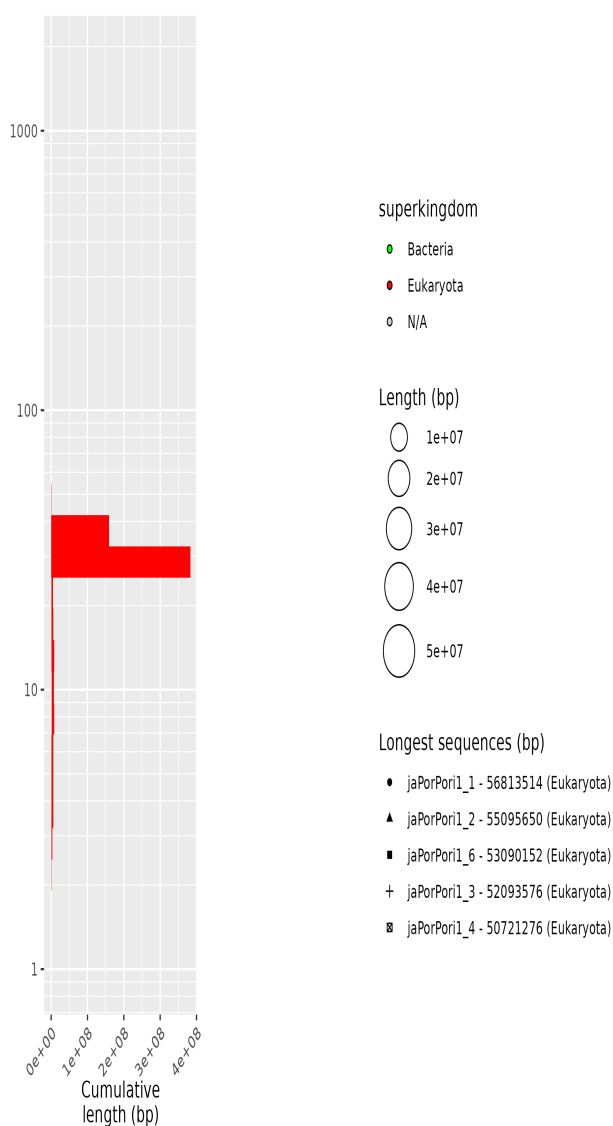
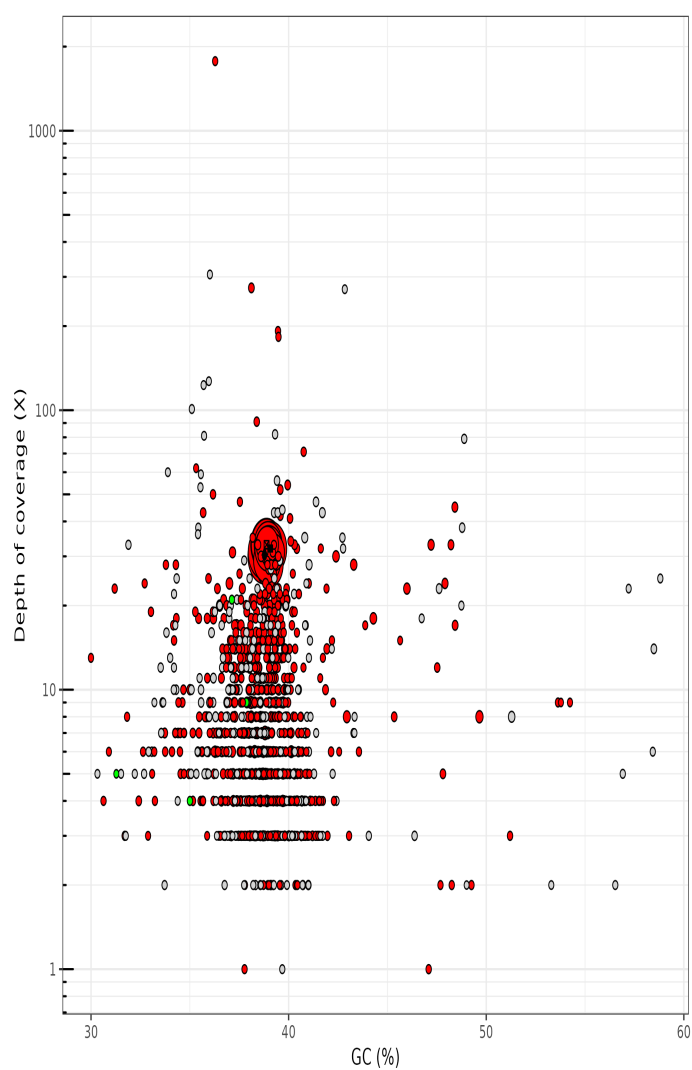


Distribution of k-mer counts coloured by their presence in reads/assemblies

Post-curation contamination screening



TAPAs summary Graph



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	Long reads	Arima
Coverage	32	160

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: Jean-Marc Aury

Affiliation: Genoscope

Date and time: 2026-07-13 05:23:50 CEST