

Genome Annotation Report

Report generated: 2025-12-17

The *Arabidopsis thaliana* genome assembly accession NA was annotated by unknown (contact info unknown). This annotation includes 264 transcribed mRNAs from 264 genes. The average transcript length is 1758.89bp, with an average of 1 coding transcripts per gene and 5 exons per transcript. The annotation file is available at https://github.com/Gaius-Augustus/BRAKER/blob/master/example/results/test3_4/braker.gtf.

Summary

Number of genes:	264
Number of CDSs:	264
Number of transcripts:	264
Mean transcript length:	1758.89
Mean transcripts per gene:	1
Mean exons per transcript:	5
BUSCO summary:	C:1.1%[S:1.0%,D:0.1%],F:0.2%,M:98.7%,n:1614
BUSCO lineage dataset:	embryophyta_odb10
BUSCO mode:	proteins
OMArk completeness:	S:1.10%,D:0.09%[U:0.08%,E:0.01%],M:98.81%
OMArk lineage:	Brassicaceae
OMArk consistency:	A:97.57%[P:0.00%,F:2.08%],I:0.35%[P:0.00%,F:0.00%],C:0.00%[P:0.00%,F:0.00%],U:2.08%

Metadata

Project ID:	NA
Study name:	Annotation reporting tool testing
Project description:	Data for testing the annotation reporting tool

Contact name:	unknown
Contact ID:	contact info unknown
Contact email:	email address unknown
Taxon ID:	3702
Scientific name:	<i>Arabidopsis thaliana</i>
Assembly accession:	assembly accession NA
Assembly seqcol digest:	NA
Assembly aliases:	BRAKER3-test-genome.fa
Evidence ID:	NA
Evidence type:	NA
Evidence version or date of retrieval:	NA
Evidence source:	NA
Annotation tools:	BRAKER3
Annotation tool versions:	unknown
Annotation protocol:	NA
Annotation file label:	BRAKER3-test-annotation-file.gtf
Annotation file URL:	https://github.com/Gaius-Augustus/BRAKER/blob/master/example/results/test3_4/ braker.gtf
Annotation file type:	GTF
Annotation file checksum:	unknown

AGAT Full Stats

Counts

AGAT stats calculated from: annotated transcripts (without isoforms)

Genes: 264

Transcripts:	264
Exons:	1312
CDS:	264
Introns:	1048
Single exon genes:	74
Single exon transcripts:	74

Mean Ratios

Transcripts per gene:	1
Exons per transcript:	5
Exons per CDS:	5
CDS per transcript:	1
Introns per transcript:	4

Mean Lengths (bp)

Gene:	1758.89
Transcript:	1758.89
Exon:	248.53
CDS:	1235.14
Intron:	131.94

Median Lengths (bp)

Gene:	1587.5
Transcript:	1587.5
Exon:	134.5

CDS: 134.5
Intron: 92

Total Lengths (bp)

Genes: 464346
Transcripts: 464346
Exons: 326076
CDS: 326076
Introns: 138270

Longest Features (bp)

Gene: 8111
Transcript: 8111
Exon: 2748
CDS: 4575
Intron: 2081

Shortest Features (bp)

Gene: 201
mRNA: 201

BUSCO Full Stats

Version: 5.8.3
Lineage dataset: embryophyta_odb10

Mode:	proteins
One line summary:	C:1.1%[S:1.0%,D:0.1%],F:0.2%,M:98.7%,n:1614
Complete BUSCOs:	1.1%
Single-copy:	1%
Duplicated:	0.1%
Fragmented:	0.2%
Missing:	98.7%
Total markers:	1614

OMArk Full Stats

Assessment information

OMAmer version:	2.1.0
OMAmer database version:	2.0.3

Completeness Assessment

Clade:	Brassicaceae
Total HOGs:	17996
Single-copy:	1.1%
Total duplicated:	0.09%
Duplicated (expected):	0.01%
Duplicated (unexpected):	0.08%
Missing:	98.81%

Consistency Assessment

Total proteins: 288
Consistent: 97.57%
 Partial consistent: 0%
 Fragmented consistent: 2.08%
Inconsistent: 0.35%
 Partial inconsistent: 0%
 Fragmented inconsistent: 0%
Contaminants: 0%
 Partial contaminants: 0%
 Fragmented contaminants: 0%
Unknown: 2.08%

Detected Species

Clade	Taxon ID	Protein count	Proteome %
Arabidopsis thaliana	3702	282	97.92