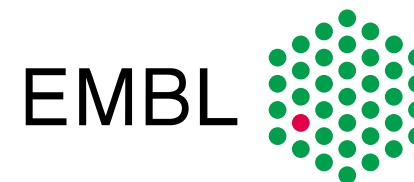


# Genome Annotation Report

Report generated: 2025-12-11



The *Arabidopsis thaliana* genome assembly [assembly accession NA](#) was annotated by [unknown \(contact info unknown\)](#). This annotation includes 264 transcribed mRNAs from 264 genes. The average transcript length is 1758.89 bp, 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](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 bp
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
-------------	----

<b>Study name:</b>	Annotation reporting tool testing
<b>Project description:</b>	Data for testing the annotation reporting tool
<b>Contact name:</b>	unknown
<b>Contact ID:</b>	contact info unknown
<b>Contact email:</b>	email address unknown
<b>Taxon ID:</b>	3702
<b>Scientific name:</b>	Arabidopsis thaliana
<b>Assembly accession:</b>	assembly accession NA
<b>Assembly seqcol digest:</b>	NA
<b>Assembly aliases:</b>	BRAKER3-test-genome.fa
<b>Evidence ID:</b>	NA
<b>Evidence type:</b>	NA
<b>Evidence version or date of retrieval:</b>	NA
<b>Evidence source:</b>	NA
<b>Annotation tools:</b>	BRAKER3
<b>Annotation tool versions:</b>	unknown
<b>Annotation protocol:</b>	NA
<b>Annotation file label:</b>	BRAKER3-test-annotation-file.gtf
<b>Annotation file URL:</b>	<a href="https://github.com/Gaius-Augustus/BRAKER/blob/master/example/results/test3_4/braker.gtf">https://github.com/Gaius-Augustus/BRAKER/blob/master/example/results/test3_4/braker.gtf</a>
<b>Annotation file type:</b>	GTF
<b>Annotation file checksum:</b>	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)**

<b>Gene:</b>	1587.5
<b>Transcript:</b>	1587.5
<b>Exon:</b>	134.5
<b>CDS:</b>	134.5
<b>Intron:</b>	92

### **Total Lengths (bp)**

<b>Genes:</b>	464346
<b>Transcripts:</b>	464346
<b>Exons:</b>	326076
<b>CDS:</b>	326076
<b>Introns:</b>	138270

### **Longest Features (bp)**

<b>Gene:</b>	8111
<b>Transcript:</b>	8111
<b>Exon:</b>	2748
<b>CDS:</b>	4575
<b>Intron:</b>	2081

### **Shortest Features (bp)**

<b>Gene:</b>	201
<b>Transcript:</b>	201

## BUSCO Full Stats

<b>Version:</b>	5.8.3
<b>Lineage dataset:</b>	embryophyta_odb10
<b>Mode:</b>	proteins
<b>One line summary:</b>	C:1.1%[S:1.0%,D:0.1%],F:0.2%,M:98.7%,n:1614
<b>Complete BUSCOs:</b>	1.1%
<b>Single-copy:</b>	1%
<b>Duplicated:</b>	0.1%
<b>Fragmented:</b>	0.2%
<b>Missing:</b>	98.7%
<b>Total markers:</b>	1614

## OMArk Full Stats

### Assessment information

<b>OMAmer version:</b>	2.1.0
<b>OMAmer database version:</b>	2.0.3

### Completeness Assessment

<b>Clade:</b>	Brassicaceae
<b>Total HOGs:</b>	17996
<b>Single-copy:</b>	1.1%
<b>Total duplicated:</b>	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