

Apollo Tips

- 1. My Apollo window has frozen. What do I do?
- 2. Some data tracks are very slow to load.
- 3. I used to see tracks and now they are gone.
- 4. What is the meaning of the lower case letters in the nucleotide sequence?
- 5. Is it possible to know which organism is the source for a particular protein alignment (in the Protein Alignment track)?
- 6. Where is the track metadata?
- 7. What does the warning symbol mean?
- 8. The track I've viewing says 'Max height reached' but I want to see more rows. How do I do that?
- 9. How do I see genes that need secondary review?
- 10. I want to drag an element from an evidence track to the user-created annotation track but I can't!
- 11. Comment/Renaming-related questions
 - 11.1. Where do I add my comment/rename?
 - 11.2. The comment or status I typed in hasn't saved. How do I make sure it saves?
 - 11.3. The deleted gene I'm reviewing needs to stay deleted. How do I mark this in Apollo?
 - 11.4. Check the names of the genes in the user-created annotations track to maintain AGI history.
- 12. What do those clear/all white gene models mean? Should we keep them?
- 13. How does secondary review work?
- 14. Apollo automatically changed the ORF when I dragged the gene model to the user created annotations band. How do I fix it?

1. My Apollo window has frozen. What do I do?

Try reloading your browser window.

- For Chrome: Shift Command R
- For Firefox: Ctrl+R
- For Safari: shift+reload

2. Some data tracks are very slow to load.

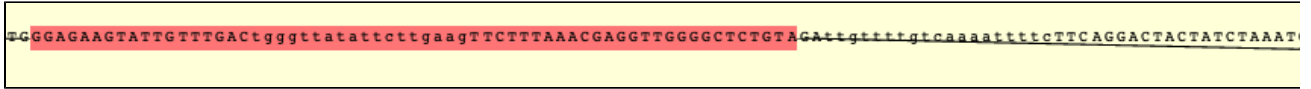
The way JBrowse works within Apollo, the first time your load up a track it might take a while. But once loaded, any successive loads will be quicker.

3. I used to see tracks and now they are gone.

Close the track and open it again. Yes, this is the standard 'turn it off and turn it on again' solution.

4. What is the meaning of the lower case letters in the nucleotide sequence?

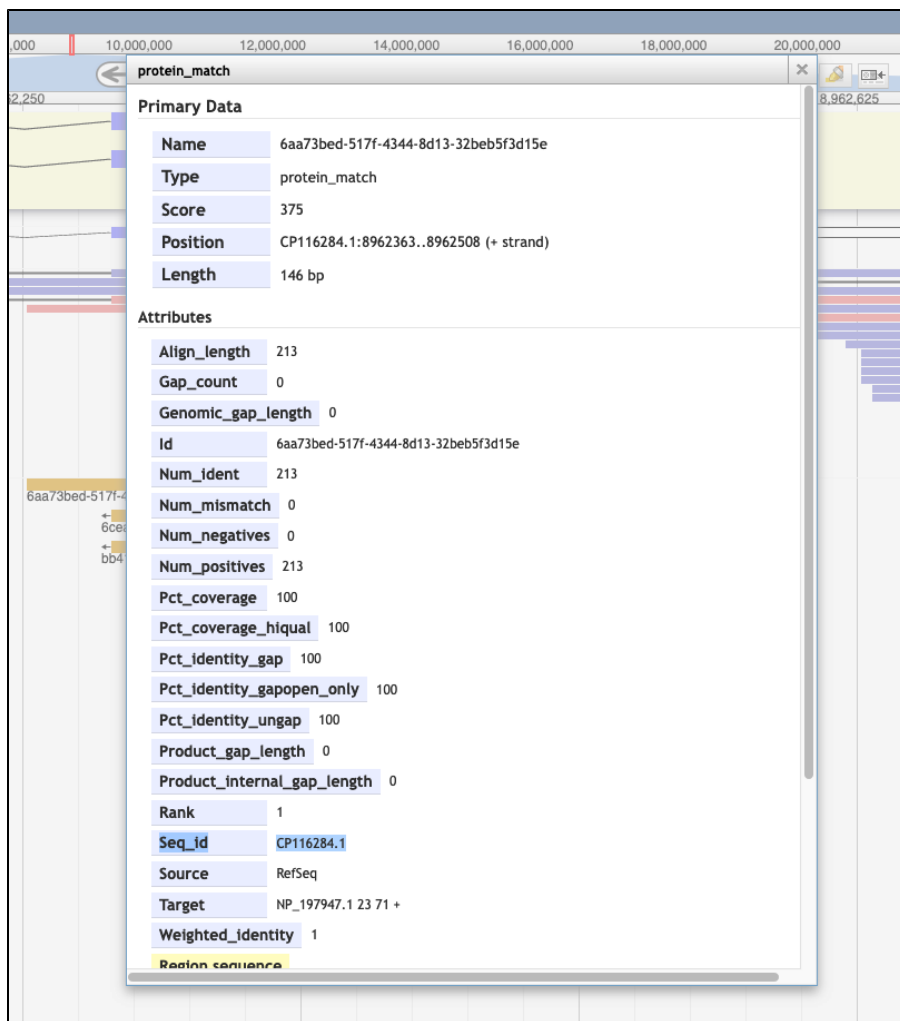
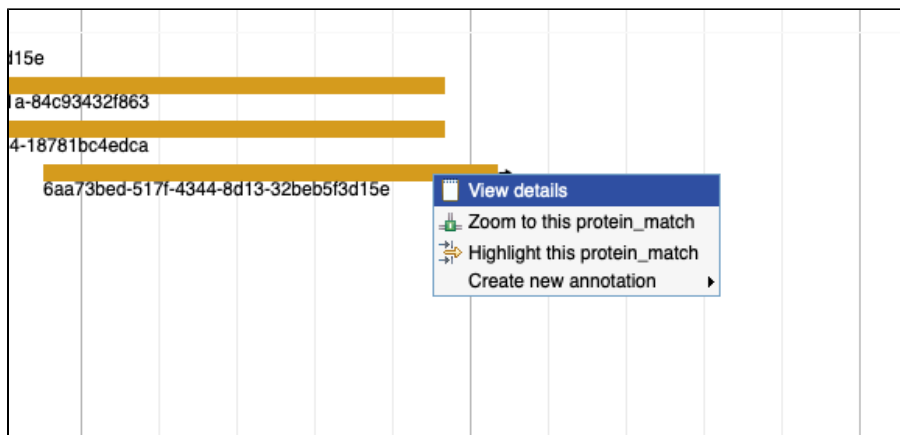
These indicate regions of low complexity. They have no impact/implication with respect to annotation.



```
TGGGAGAAGTATTGTTTGACTggggttatattcttgaagTTCTTTAAACGAGGTTGGGGCTCTGTACattgttttg+caaaattttcTTCAGGACTACTATCTAAAT
```

5. Is it possible to know which organism is the source for a particular protein alignment (in the Protein Alignment track)?

By right clicking on the alignment, and then selecting 'View details.' You'll see the GenBank accession in the id section and can search for that at [NCBI](#).



6. Where is the track metadata?

<input type="checkbox"/>	Gnomon Models	Remove
<input checked="" type="checkbox"/>	Col-CC_Genomic_Annotations_Data	
<input checked="" type="checkbox"/>	AT-Col-CC-Liftoff-from-TAIR10.1	
<input type="checkbox"/>	TranscriptomeReconstructoR models	Remove

Coverage Tracks, RNAseq and Long Read		0/3	<input type="checkbox"/>
<input type="checkbox"/>	RNAseq combined_recapped_filtered coverage		
<input type="checkbox"/>	Long Read Alignments coverage plus strand	Remove	
<input type="checkbox"/>	Long Read Alignments coverage minus strand	Remove	

Merged, Capped, and Filtered RNAseq Reads		0/1	<input type="checkbox"/>
<input type="checkbox"/>	RNAseq capped_merged_filtered reads		

Protein Evidence		0/5	<input type="checkbox"/>
<input type="checkbox"/>	PFAM domains	Remove	
<input type="checkbox"/>	PFAM domains - Liftoff	Remove	
<input type="checkbox"/>	PANTHER families	Remove	
<input type="checkbox"/>	PANTHER families - Liftoff	Remove	
<input type="checkbox"/>	Protein alignments chained	Remove	

RNA-Seq tracks (Col-0)/Flower	0/4	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Leaf	0/3	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Pollen	0/5	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Root	0/2	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Seed	0/5	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Seedling	0/25	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Shoot apex	0/4	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Silique (fruit)	0/5	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Stem	0/4	<input type="checkbox"/>
RNA-Seq tracks (Col-0)/Whole flower	0/5	<input type="checkbox"/>
RNA-Seq tracks (Landsberg)	0/9	<input type="checkbox"/>

TRR CAGE PAT plaNET		0/6	<input type="checkbox"/>
<input type="checkbox"/>	CAGE forward	Remove	
<input type="checkbox"/>	CAGE reverse	Remove	
<input type="checkbox"/>	PAT forward	Remove	
<input type="checkbox"/>	PAT reverse	Remove	
<input type="checkbox"/>	plaNET forward	Remove	
<input type="checkbox"/>	plaNET reverse	Remove	

Transcript Evidence

2/5

☐ Known_Reference_Sequences

Remove

☒ Col-CC Same Species Combined Track

☐ Model_Reference_Sequences

Remove

☒ TSA_alignments

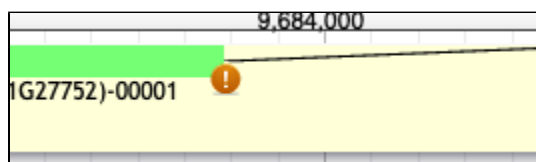
Remove

☐ [Long Read Alignments](#)

TRACK NAME	Description of data
Col-CC_Genomic_Annotations_Data	Result of NCBI Eukaryotic Annotation Pipeline
AT-Col-CC-Liftoff-from-TAIR10.1	v11 models mapped to v12 reference using Liftoff
TranscriptomeReconstructoR models	<p>Method:</p> <p>(1) Assembled expression evidence: ONT-DRS (ERR3764345 - ERR3764351), CAGE-Seq (SRR10045003 - SRR10045005), PAT-Seq (SRR7160296, SRR7160297, SRR7160299), plaNET-Seq (SRR9117170 - SRR9117173)</p> <p>(2) Aligned all the datasets to Col-CC genome</p> <p>(3) Built TRR based annotation using alignment (bam) files</p> <p>Output:</p> <ul style="list-style-type: none"> High Confidence (HC) genes - full length mRNA picked from ONT which have both proper start site from CAGE and end site from PAT-Seq coinciding with ONT data. Medium Confidence (MC) genes - full length mRNA picked from ONT with either start site from CAGE or end site from PAT-Seq not coinciding with ONT data. Low Confidence (LC) genes - full length mRNA picked from ONT with both start site from CAGE and end site from PAT-Seq not coinciding with ONT data. Single Method Confidence (SMC) genes - novel genes annotated only from plaNET-Seq data.
Gnomon Models	<p>One of the outputs of the annotation pipeline. These are a superset of the final set of annotated models.</p> <p>"Gnomon annotation of the genomic sequence. Sequence identifiers are provided as accession.version for the genomic sequences and Gnomon identifiers for the Gnomon models:gene.XXX for genes, GNOMON.XXX.m for transcripts and GNOMON.XXX.p for proteins. These identifiers are NOT universally unique. They are unique per annotation release only." (from NCBI documentation)</p>
Coverage Tracks, RNAseq and Long Read	
RNAseq capped_merged_filtered coverage	Coverage track version of the RNAseq capped_merged_filtered reads.
Long Read Alignments coverage, plus strand	Coverage track of the long reads, plus strand ONLY.
Long Read Alignments coverage, minus strand	Coverage track of the long reads, minus strand ONLY.
Merged, Capped, and Filtered RNAseq Reads	
RNAseq capped_merged_filtered reads	RNAseq reads from the 62 RNAseq experiments after filtering for overly long (>5Kbp) inserts, and successive capping and merging to try to reduce the overabundance of highly expressed genes. File was then filtered to remove overly long (>12Kbp) introns.
Protein Evidence	

PFAM domains	Results from an INTERPROSCAN run on the proteins from the V12 prediction to get the PFAM domain information, converted to absolute position on the Col-CC assembly.
PFAM domains - Liftoff	Results from an INTERPROSCAN run on the proteins from the Araport11 release to get the PFAM domain information, converted to absolute position on the Col-CC assembly (using the Liftoff file that converted Araport11 coordinates to Col-CC coordinates).
PANTHER families	Results from an INTERPROSCAN run on the proteins from the V12 prediction to get the PANTHER family information, converted to absolute position on the Col-CC assembly.
PANTHER families - Liftoff	Results from an INTERPROSCAN run on the proteins from the Araport11 release to get the PANTHER family information, converted to absolute position on the Col-CC assembly (using the Liftoff file that converted Araport11 coordinates to Col-CC coordinates).
Protein alignments chained	Alignments of Arabidopsis thaliana and other Brassicaceae proteins, including Araport 11 annotated proteins, to the genomic sequence(s). These alignments may have been used as evidence for gene prediction by the NCBI annotation pipeline. Pieces of the same protein have been connected together for easier visualization.
TRR CAGE PAT plaNET	
CAGE forward	TranscriptomeReconstructoR models CAGE evidence, forward strand (evidence supporting the start site of transcription)
CAGE reverse	TranscriptomeReconstructoR models CAGE evidence, reverse strand (evidence supporting the start site of transcription)
PAT forward	TranscriptomeReconstructoR models PAT evidence, forward strand (evidence supporting the end site of transcription)
PAT reverse	TranscriptomeReconstructoR models PAT evidence, reverse strand (evidence supporting the end site of transcription)
plaNET forward	TranscriptomeReconstructoR models plaNET evidence, forward strand (evidence supporting transcription of mRNAs and lncRNAs)
plaNET reverse	TranscriptomeReconstructoR models plaNET evidence, reverse strand (evidence supporting transcription of mRNAs and lncRNAs)
Transcript Evidence	
Known Reference Sequences	"Alignments of the annotated Known RefSeq transcripts (identified with accessions prefixed with NM_ and NR_) to the genome." (from NCBI documentation) These were NOT used in generating the Col-CC annotation. They are alignments of the annotated transcripts to the genome and can provide additional insight into the predicted gene structures independent of the prediction.
Model Reference Sequences	"Alignments of the annotated Model RefSeq transcripts (identified with accessions prefixed with XM_ and XR_) to the genome." (from NCBI documentation) These were NOT used in generating the Col-CC annotation. They are alignments of the annotated transcripts to the genome and can provide additional insight into the predicted gene structures independent of the prediction.
Col-CC Same Species Combined	Alignments of same-species cDNAs, ESTs and TSAs to the genomic sequence(s). cDNAs and ESTs alignments (not TSAs) may have been used as evidence for gene prediction by the NCBI annotation pipeline. The TSA alignment track is a subset of the Col-CC Same Species track. Pieces of the same transcript have been connected together for easier visualization.
TSA alignment	Alignments of transcripts assembled from RNA-Seq reads, and submitted to GenBank (see accessions DAHAIV01, GGJX01, GJRK01 and GKIF01). These were not used as evidence for gene prediction by the NCBI annotation pipeline.
RNA seq tracks from various plant parts and growth stages/conditions of those parts	Name is based on the GenBank record, for example, SRR1019221. You can link to that record using this base URL for more information on the experiment: https://www.ncbi.nlm.nih.gov/sra/SRR1019221
Long Read alignments	Alignments of individual IsoSeq reads in SRA. These alignments may have been used as evidence for gene prediction by the NCBI annotation pipeline. Right clicking on the read itself will allow you to 'View Details' and see the ID of the SRA entry for the experiment. Using the id (e.g., SRR11031292), you can go to the full GenBank record for the experiment. https://www.ncbi.nlm.nih.gov/sra/?term=SRR11031292

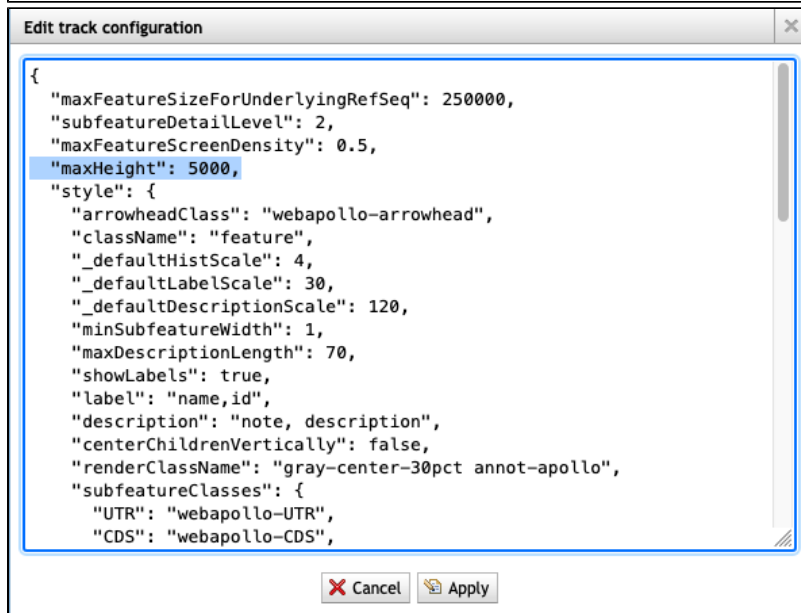
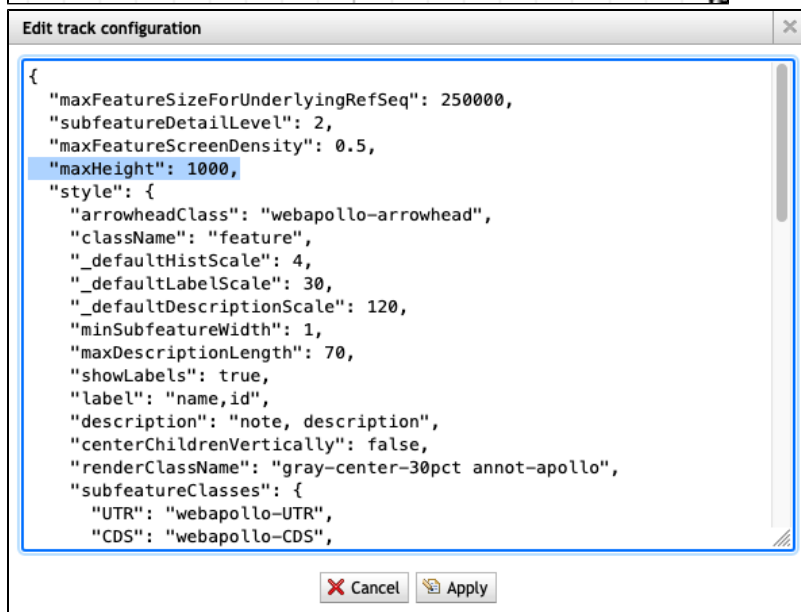
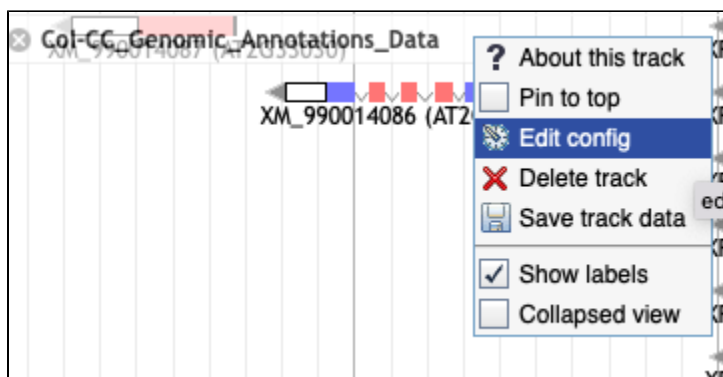
7. What does the warning symbol mean?



This symbol means that the location and sequence of the splice site needs to be investigated and verified because it does not conform to the GT-AG rule. Other possibilities include GC-AG and AT-AC.

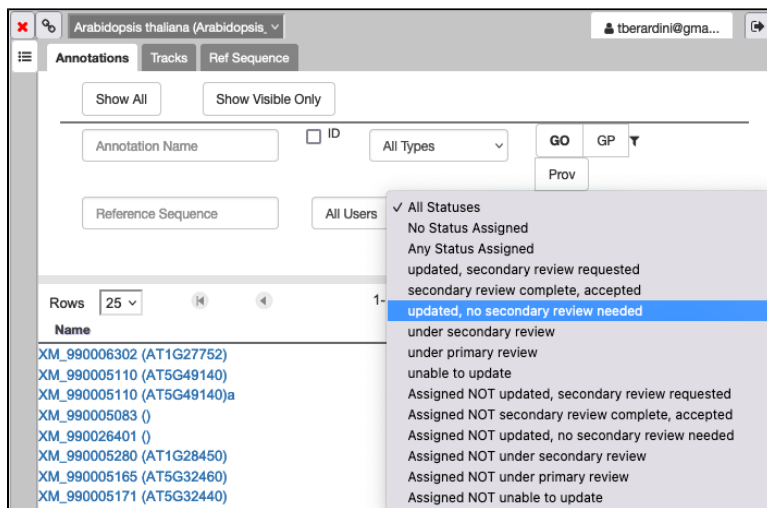
8. The track I've viewing says 'Max height reached' but I want to see more rows. How do I do that?

- Click on Edit config.
- Scroll to the top of the config, find the 'maxHeight' setting.
- Change the value to '5000'
- Click on 'Apply'
- See more rows!



9. How do I see genes that need secondary review?

In the Annotations tab of your right hand panel, click on the dropdown for "All Statuses" and select the status you want to filter the annotations to review.



10. I want to drag an element from an evidence track to the user-created annotation track but I can't!

Right click on the element and select 'Create new annotation' from the menu. Pick the type of element you want to create and it will appear in the yellow track.



11. Comment/Renaming-related questions

11.1. Where do I add my comment/rename?

- If you have already added comments at the gene level, these will apply to all mRNAs under that gene.
- In cases where there is only one mRNA/gene, the comment/rename (DELETE) on the mRNA transitively applies to the gene.
- If there is **more than one mRNA/gene** and you want to **record different actions for individual mRNAs**, please put the specific comments (or DELETE) **on the relevant mRNAs only**.
- If there is **more than one mRNA/gene** and you want to **record the same action for ALL mRNAs/the whole gene**, please put the specific comments (or DELETE) on **any one of the mRNAs**.

11.2. The comment or status I typed in hasn't saved. How do I make sure it saves?

For saving comments and gene status, make sure you click outside of the panel where you created the comment to 'make it stick'.

Annotations Tracks Ref Sequence

Show All Show Visible Only

Annotation Name ☐ ID All Types GO GP Prov

Reference Sequence All Users All Statuses

Rows 25 1-50 of 164

Name	Seq	Type	Length	Updated
AT1G64795.1	CP116280.1	gene	270	Jul 17, 2023
AT1G55710.1	CP116280.1	gene	219	Jul 17, 2023
AT1G24390.1	CP116280.1	gene	795	Jul 17, 2023
GKIF01155911.1	CP116284.1	gene	2,432	Jul 17, 2023
AT1G79620.1	CP116280.1	gene	4,931	Jul 17, 2023
AT1G56060.2	CP116280.1	gene	552	Jul 17, 2023
AT1G17230.1	CP116280.1	gene	4,517	Jul 17, 2023
AT1G52820.1	CP116280.1	gene	1,253	Jul 17, 2023
AT1G09794.1	CP116280.1	gene	1,606	Jul 17, 2023

gene: AT1G55710.1 [Link to annotation](#) [Close\(x\)](#)

Details GO Gene Product Provenance DbXref Comment Attributes

Go ID Sync name w

Type gene Status

Name AT1G55710.1

Symbol

Aliases (' separated)

Description

Step 1: Update status in this panel

Step 2: Click in this panel to save status

No status selected
 updated, secondary review requested
 secondary review complete, accepted
 ✓ updated, no secondary review needed
 under secondary review
 under primary review
 unable to update
 no update needed

Delete

Annotations Tracks Ref Sequence

Show All Show Visible Only

Annotation Name ☐ ID All Types GO GP Prov

Reference Sequence All Users All Statuses

Rows 25 1-50 of 163

Name	Seq	Type	Length	Updated
AT1G55710.1	CP116280.1	gene	219	Jul 17, 2023
AT1G24390.1	CP116280.1	gene	795	Jul 17, 2023
GKIF01155911.1	CP116284.1	gene	2,432	Jul 17, 2023
AT1G79620.1	CP116280.1	gene	4,931	Jul 17, 2023
AT1G56060.2	CP116280.1	gene	552	Jul 17, 2023
AT1G17230.1	CP116280.1	gene	4,517	Jul 17, 2023
AT1G52820.1	CP116280.1	gene	1,253	Jul 17, 2023
AT1G09794.1	CP116280.1	gene	1,606	Jul 17, 2023
GKIF01109402.1	CP116283.1	gene	3,828	Jul 17, 2023

gene: AT1G55710.1 [Link to annotation](#) [Close\(x\)](#)

Details GO Gene Product Provenance DbXref Comment Attributes

Comment

Evidence from JBrowse tracks; expression in seedling and meristems

Comment

- Add Canned Comment

Step 1. Add comment in this panel

Step 2. Click in this panel to save comment

11.3. The deleted gene I'm reviewing needs to stay deleted. How do I mark this in Apollo?

1. create a gene model of the gene that should stay deleted by dragging into the user-created annotations

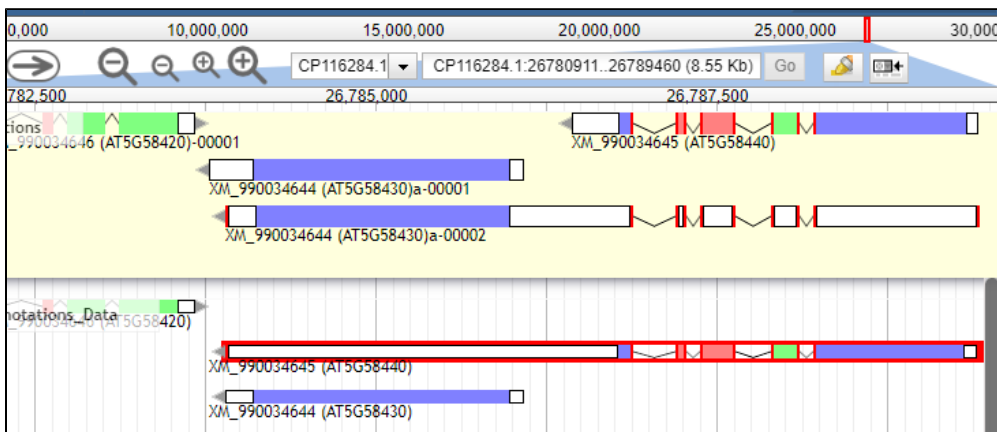
2. rename so that the name **ends with 'DELETE'**
 3. set status (either 'updated, secondary review requested' OR 'updated, no secondary review needed')
- (see Q10 for what to do if there are different actions for different mRNAs of the same gene)

11.4. Check the names of the genes in the user-created annotations track to maintain AGI history.

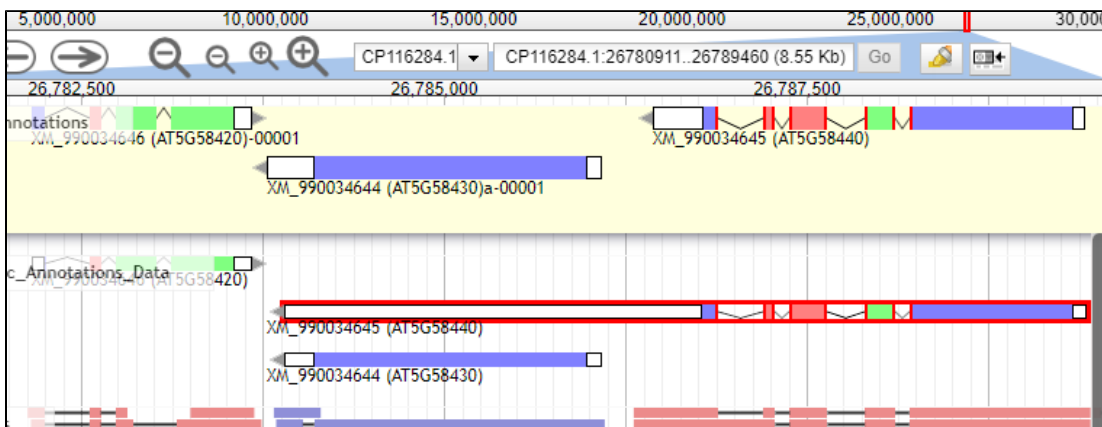
When dragging an annotation from the Col-CC to the user-created annotations, if the gene range is overlapping with the gene on the left, the gene ID will be replaced by the gene on the left. Just want to bring awareness to this small bug so that we don't mistakenly assign gene IDs to the user-created annotation track.

In this example, the AT5G58440 has a mistakenly annotated 3'UTR, that was merged with AT5G58430. When I dragged AT5G58440 from Col-CC track to user-created annotation track, the gene ID was replaced by AT5G58430. So the gene ID needs to be curated in this case. (Thanks, Xiaohui Li, for reporting.)

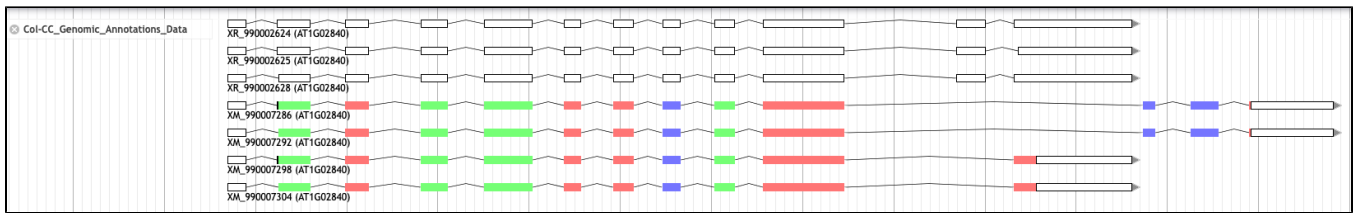
Before: (Dragged longer gene model has AGI = AT5G58430)



After: (Gene model to the 3' end after editing was renamed to be AGI = AT5G58440)



12. What do those clear/all white gene models mean? Should we keep them?



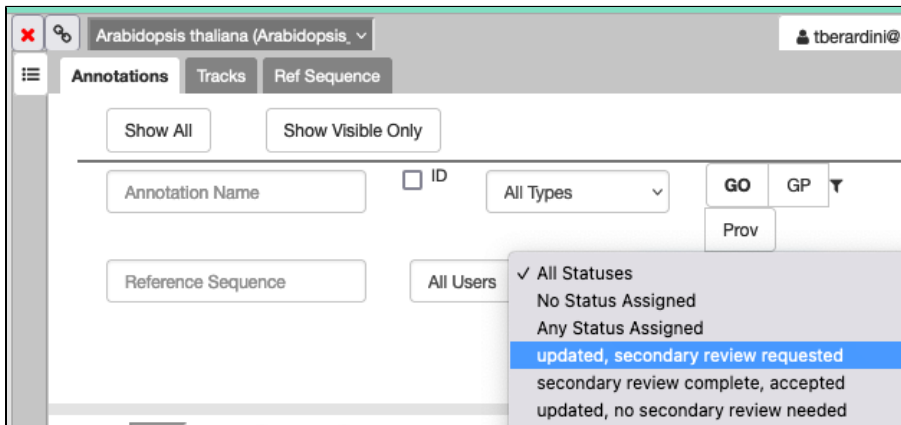
These gene models exist because the transcripts are made BUT they have been predicted to be targeted for Nonsense-Mediated Decay (NMD). No protein product is predicted to be made even though there may be a recognizable open reading frame. The gene models should be retained because they may have regulatory or other functions as RNAs.

NMD review: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7854845/>

In the example above, the three all white gene models have subtle differences that can be seen when zooming in closer.

13. How does secondary review work?

- Go to Annotations tab in Apollo.
- Select status = 'updated, secondary review requested' from the drop down.



- See list of results (620 as of 10/17/23)

Name	Seq	Type	Length	Updated
XM_990017854 (AT1G10970)	CP116280.1	gene	2,037	Oct 17, 2023
XM_990005558 (AT1G30620)	CP116280.1	gene	3,634	Oct 17, 2023
XM_990001009 (AT1G24440)	CP116280.1	gene	1,919	Oct 17, 2023
AT5G39645	CP116284.1	gene	289	Oct 17, 2023
AT3G06895	CP116282.1	gene	163	Oct 17, 2023
AT4G05091	CP116283.1	gene	415	Oct 17, 2023
AT2G27402	CP116281.1	gene	1,593	Oct 17, 2023
AT2G20660	CP116281.1	gene	306	Oct 17, 2023
AT5G04210	CP116284.1	gene	676	Oct 17, 2023
AT5G04220	CP116284.1	gene	3,220	Oct 17, 2023
AT3G10500	CP116282.1	gene	2,960	Oct 17, 2023
AT3G10490	CP116282.1	gene	3,378	Oct 17, 2023
AT5G35067	CP116284.1	gene	466	Oct 17, 2023
AT3G16930	CP116282.1	gene	488	Oct 17, 2023
AT2G35290	CP116281.1	gene	1,112	Oct 17, 2023
AT1G53285	CP116280.1	gene	536	Oct 17, 2023
AT1G43825	CP116280.1	gene	2,839	Oct 17, 2023
AT1G31270	CP116280.1	gene	855	Oct 17, 2023
AT1G08985	CP116280.1	gene	462	Oct 17, 2023
AT5G36490.1	CP116284.1	gene	555	Oct 16, 2023
XM_990029835 (AT4G38320)	CP116283.1	gene	1,831	Oct 16, 2023
AT3G50410.1	CP116282.1	gene	1,224	Oct 16, 2023
AT2G17043.1	CP116281.1	gene	520	Oct 16, 2023

- Click on one of the results. (if it's one that was originally yours, skip!)

- Click on the blue GO button.

Gene ID	Gene Name	Type	Length	Date
XM_990017854 (AT1G10970)	CP116280.1	gene	2,037	Oct 17, 2023
XM_990005558 (AT1G30620)	CP116280.1	gene	3,634	Oct 17, 2023
XM_990001009 (AT1G24440)	CP116280.1	gene	1,919	Oct 17, 2023
AT5G39645	CP116284.1	gene	289	Oct 17, 2023
AT3G06895	CP116282.1	gene	163	Oct 17, 2023
AT4G05091	CP116283.1	gene	415	Oct 17, 2023
AT2G27402	CP116281.1	gene	1,593	Oct 17, 2023

gene: XM_990017854 (AT1G10970) [Link to annotation](#) [Close\(x\)](#)

Details GO Gene Product Provenance DbXref Comment Attributes

☐ Obsolete

Type **gene** Status updated, secondary review

Name XM_990017854 (AT1G10970)

Symbol

Aliases (!' separated)

Description isoforms updated

Location 3669279 - 3671315 strand(-) Partial: ☐ fm ☐ fmax

Ref Sequence CP116280.1

- The genome view will change to that location.
- Set the status to 'under secondary review' so no one else nabs this.

Details GO Gene Product Provenance DbXref Comment Attributes

Type **gene** Status **updated, secondary review requested**

Name XM_990017854 (AT1G10970)

Symbol

Aliases (!' separated)

Description isoforms updated

Location 3669279 - 3671315 strand(-) Partial: ☐ fm ☐ fmax

No status selected

- ✓ updated, secondary review requested
- secondary review complete, accepted
- updated, no secondary review needed
- under secondary review**
- under primary review
- unable to update
- no update needed

- TO FINISH:** Review, edit, comment to track what you did/if you agree and finally change status to 'secondary review completed, accepted'. (or select 'for discussion' if you can't decide)

Details GO Gene Product Provenance DbXref Comment Attributes

Type **gene** Status **secondary review complete, accepted**

Name XM_990017854 (AT1G10970)

Symbol

Aliases (!' separated)

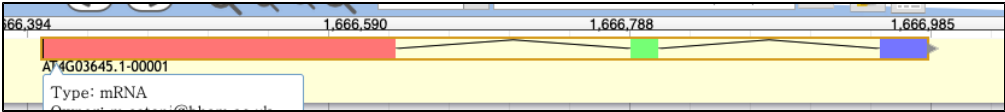
Description isoforms updated

No status selected

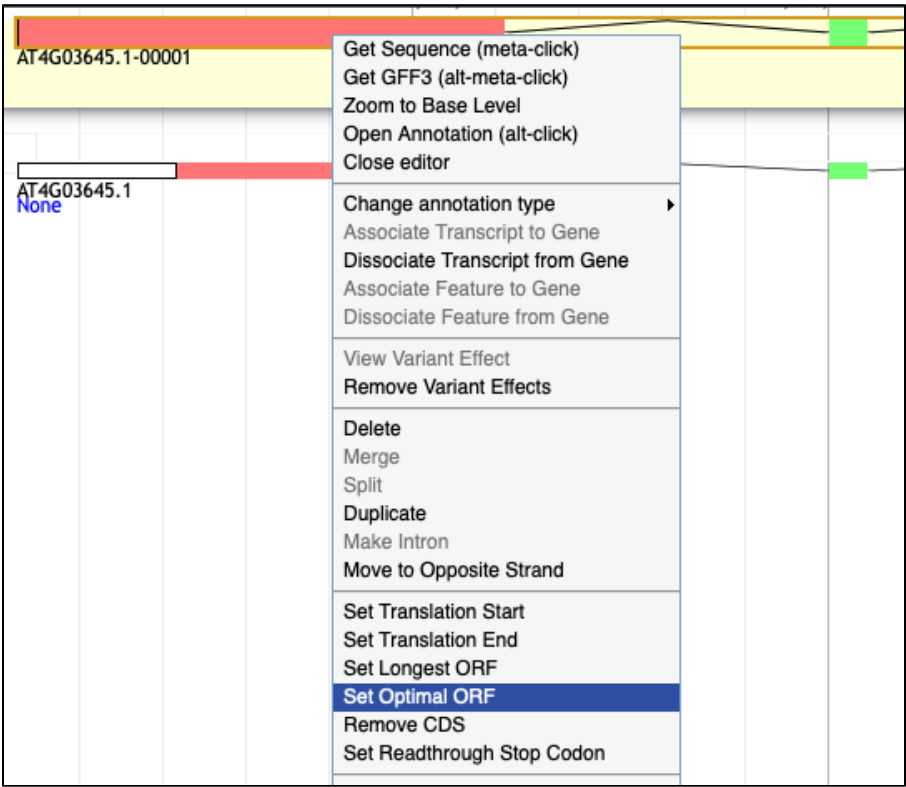
- ✓ updated, secondary review requested
- secondary review complete, accepted**
- updated, no secondary review needed
- under secondary review
- under primary review
- unable to update
- no update needed

14. **Apollo automatically changed the ORF when I dragged the gene model to the user created annotations band. How do I fix it?**

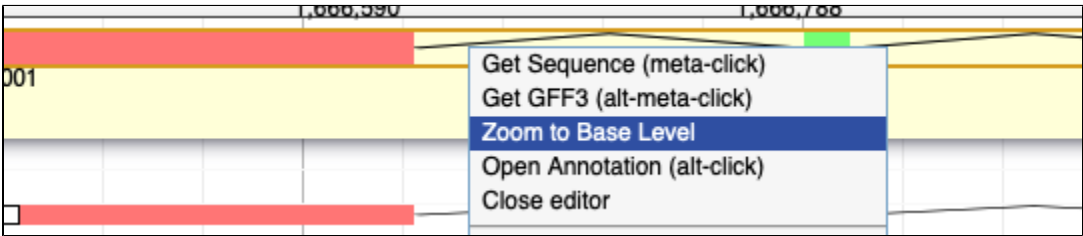
Left click to highlight the whole gene model (in an intron) in the user-created annotations track.



Option A. Right click and select 'Set Optimal ORF'. If that restores the v11 ORF, you are done.



Option B. Right click and select 'Zoom to Base Level'.



Find the ATG where the v11 gene model starts translation in the user created annotations version. Hover the cursor on the A.
Right click and select Set Translation Start. The v11 ORF should be restored.

