## **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.

PG<mark>GGAGAAGTATTGTTTGACtgggttatattcttgaagTTCTTTAAACGAGGTTGGGGCTCTGTA<del>GAttgttttgtcaaaatttccTTCAGGACTACT</del>AATCTAAAT</mark>

## 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.

a-84c93432f863 4-18781bc4edca 6aa73bed-517f-4344-8d13-32beb5f3d15e
4-18781bc4edca 6aa73bed-517f-4344-8d13-32beb5f3d15e
6aa73bed-517f-4344-8d13-32beb5f3d15e Uiew details ↓ Zoom to this protein_match ↓ Highlight this protein_match Create new annotation

0	10.0	00,000 1	2,000,000	14,000,000	16,000,000	18,000,000	20,000.	000
	æ	protein_match	_,,			,,	×	
50	-	Primary Data						8,962,625
	_	, , , , , , , , , , , , , , , , , , ,						
	_	Name	6aa73be	d-517f-4344-8d13-32	beb5f3d15e			
		Туре	protein_	match				
	-	Score	375					
-		Position	CP11628	4.1:8962363896250	8 (+ strand)			
i		Length	146 bp					
	-	Attributes						
		Align lengt	<b>213</b>					
		Gap count	0					
		Genomic ga	n length 0					
		Id	6aa73bed-	517f-4344-8d13-32beb	5f3d15e			
	d-517f-4 €- 6cea €- bb4	Num ident	213	5171 1511 0015 52505	5150150			
		Num misma	atch 0					
		Num negatives 0						
		<sup>4</sup> Num positives 213						
		Pct coveras	e 100					
		Pct coverage	coverage highai 100					
		Pct identity	/ gap 100					
		Pct identity	/ gapopen o	nly 100				
		Pct identity	ungap 100	)				
		Product ga	plength 0					
		Product_int	ernal_gap_le	ength 0				
		Rank	1					
		Seq_id	CP116284.	1				
		Source	RefSeq					
		Target	NP_19794	7.1 23 71 +				
		Weighted_id	dentity 1					
		Region segu	ience					

### 6. Where is the track metadata?

Gnomon Mo	odels	Remove
Col-CC_Ger	nomic_Annotations_Data	
AT-Col-CC-I	Liftoff-from-TAIR10.1	
Transcriptor	neReconstructoR models	Remove

Coverage Tracks, RNAseq and Long Read				
RNAseq combined_recapped_filtered coverage				
Long Read Alignments coverage plus strand	Remove			
Long Read Alignments coverage minus strand	Remove			

0/1

## Merged, Capped, and Filtered RNAseq Reads RNAseq capped\_merged\_filtered reads

# Protein Evidence 05 PFAM domains Remove PFAM domains - Liftoff Remove PANTHER families Remove PANTHER families - Liftoff Remove PANTHER families - Liftoff Remove Portein alignments chained Remove

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

TRR CAGE PAT plaNET	0/6
CAGE forward	Remove
CAGE reverse	Remove
PAT forward	Remove
PAT reverse	Remove
plaNET forward	Remove
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_Ann otations_Data	Result of NCBI Eukaryotic Annotation Pipeline		
AT-Col-CC-Liftoff- from-TAIR10.1	v11 models mapped to v12 reference using Liftoff		
TranscriptomeReco nstructoR models	Method:		
	(1) Assembled expression evidence: ONT-DRS (ERR3764345 - ERR3764351), CAGE-Seq (SRR10045003 - SRR10045005), PAT-Seq (SRR7160296, SRR7160297, SRR7160299), plaNET-Seq (SRR9117170 - SRR9117173)		
(2) Aligned all the datasets to Col-CC genome			
	(3) Built TRR based annotation using alignment (bam) files		
	Output:		
	<ul> <li>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.</li> <li>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.</li> <li>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.</li> <li>Single Method Confidence (SMC) genes - novel genes annotated only from plaNET-Seq data.</li> </ul>		
Gnomon Models	One of the outputs of the annotation pipeline. These are a superset of the final set of annotated models.		
	"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)		
Coverage Tracks, RNAseq and Long Read			
RNAseq capped_merged_fil tered 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_fil tered 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 CoI-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 IncRNAs)
plaNET reverse	TranscriptomeReconstructoR models plaNET evidence, reverse strand (evidence supporting transcription of mRNAs and IncRNAs)
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 CoI-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 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	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
stages/conditions of those parts	
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.

×	🗞 Arabidopsis thaliana (Arabidopsis, 🗸		🛔 tberardini@gma 🕞		
≔	Annotations Tracks Ref Sequence				
	Show All Show Visible Only				
	Annotation Name	All Types 🗸 🖌 GO	GP T		
		Prov			
	Reference Sequence All Users	✓ All Statuses No Status Assigned Any Status Assigned updated, secondary review requested			
	Rows 25 v 🕑 🕚 1-	updated, no secondary review	w needed		
	Name	under secondary review			
	XM_990006302 (A11G27752) XM_990005110 (AT5G49140) XM_990005110 (AT5G49140)a XM_990005083 () XM_990005083 (A11G28450) XM_990005165 (AT5G32460) XM_990005161 (AT5G32440)	under primary Fevlew unable to update Assigned NOT updated, secc Assigned NOT secondary rev Assigned NOT updated, no s Assigned NOT under second Assigned NOT under primary Assigned NOT unable to upd	ndary review requested iew complete, accepted acondary review needed ary review review ate		

#### 10. I want to drag an element from an evidence track to the usercreated 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/renaming (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 Nam		All Types	GO GP	т
		Par typoo	Prov	
Reference Sequ	ence All	Users v	All Statuses ~	
Rows 25 ~	•	1-50 of 164	<b>()</b>	
Name		Seq	Type Length	Updated
AT1G64795.1	top 2: Click in th	CP116280.1	gene 270	Jul 17, 2023 🔹
AT1G55710.1	and to save stat	CP116280.1	gene 219	Jul 17, 2023 <
AT1G24390.1	aner to save stat	US CP116280.1	gene 795	Jul 17, 2023 <
AT1G79620.1		CP116284.1 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	aene 1.606	Jul 17. 2023 🖪
gene: AT1G55710.1		Link to anno	tation	Close(x)
Details GO Gene	Product Provenance	DbXref Comment	Attributes	
		No status selected	iow requested	
⊖ Go 🚯 II	Sync name w	secondary review comp	lete, accepted	🛍 Delete
Type gape	Statue	updated, no secondary	review needed	
Type gene	Status	under secondary review	I	
Name AT1G5	5710.1	under primary review		
		unable to update		
Symbol		no update needed		
Aliases (' ' separa	ted) Step	1: Update status	s in this panel	
Description				
Description				_
Annotations Tracks	Ref Sequence			
Show All	Show Visible Only			
Appotation Nam	Show Visible Only	All Times	GO GP	T
Annotation Nam	Show Visible Only	All Types	v GO GP	T
Annotation Nam	Show Visible Only	All Types	<ul><li>✓ GO GP</li><li>Prov</li></ul>	T
Show All Annotation Nam Reference Sequ	Show Visible Only	All Types	GO GP Prov All Statuses	T
Show Ali Annotation Nam Reference Sequ	Show Visible Only	All Types	GO     GP       Prov        All Statuses	T
Show Ali Annotation Nam Reference Sequ Rows 25 ~	Show Visible Only le Dib ence All	All Types		Y
Reference Sequeration Name	Show Visible Only le Dib ence All	All Types	GO     GP       Prov     All Statuses       Image: Status of the status of th	▼ Updated
Reference Sequeration Name Rows 25 v Name AT1G55710.1 AT1G24390.1	Show Visible Only	All Types	GO     GP       Prov        All Statuses     ~       Image: State of the	▼ Updated Jul 17, 2023 ∢ Jul 17, 2023 ∢
Reference Sequeration Name Rows 25 - Name AT1G55710.1 AT1G24390.1 GKIF01155911.1	Show Visible Only	All Types	GO     GP       Prov        All Statuses     ✓       Type     Length       gene     219       gene     795       gene     2,432	▼ Updated Jul 17, 2023 < Jul 17, 2023 < Jul 17, 2023 <
Show All           Annotation Nam           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G524390.1           GKIF01155911.1           AT1G79620.1	Show Visible Only	All Types	GO     GP       Prov        All Statuses     ✓       Type     Length       gene     219       gene     2,432       gene     4,931	▼ Updated Jul 17, 2023 < Jul 17, 2023 < Jul 17, 2023 < Jul 17, 2023 <
Show All           Annotation Name           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G79620.1           AT1G56060.2	Show Visible Only e Diversion of the only ence All rece All rece All rece Save comment	All Types	GO     GP       Prov        All Statuses     ✓       Type        Type     Length       gene     219       gene     295       gene     4,931       gene     552	▼ Updated Jul 17, 2023 < Jul 17, 2023 <
Show All           Annotation Nam           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G76620.1           AT1G256060.2           AT1G17230.1	Show Visible Only e Diversion of the only ence All rep 2. Click in the panel to save comment	All Types Users 1-50 of 163 Seq CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1	GO     GP       Prov        All Statuses     ✓       Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,517       gene     4,517	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ←
Show All           Annotation Nam           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G76620.1           AT1G568060.2           AT1G52820.1           AT1G52820.1	Show Visible Only e ID ID ID ID IC	All Types 1-50 of 163 Seq CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1	GO         GP           Prov            All Statuses         ✓           Type         Length           gene         219           gene         2,432           gene         4,931           gene         552           gene         4,517           gene         1,253           nene         1,606	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ←
Show All           Annotation Nam           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G76620.1           AT1G56600.2           AT1G17230.1           AT1G52820.1           AT1G52820.1           AT1G90794.1           GKIF01109402.1	Show Visible Only e Proce All e Proce All	All Types 1-50 of 163 Seq CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1 CP116280.1	GO         GP           Prov         Prov           All Statuses         ×           Type         Length           gene         2,193           gene         2,432           gene         2,432           gene         4,931           gene         1,253           gene         1,263           gene         1,606           oene         3,828	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ←
Show All           Annotation Nam           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G76620.1           AT1G56600.2           AT1G52820.1           AT1G52820.1           AT1G52820.1           AT1G52820.1           AT1G55710.1           GKIF01109402.1           gene: AT1G55710.1	Show Visible Only e Proce All Proce All Proceedings All Procedings All Proceedings All Proceedings All Proceedings All Proced	All Types 1-50 of 163 Seq CP116280.1 C	GO         GP           Prov         Prov           All Statuses         ~           Type         Length           gene         2,193           gene         2,432           gene         4,931           gene         1,253           gene         1,263           gene         1,253           gene         3,828	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ←
Show All           Annotation Nam           Reference Sequ           Rows         25 ✓           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G7660.2           AT1G52820.1           AT1G52820.1           AT1G52820.1           AT1G52820.1           GKIF01109402.1           gene: AT1G55710.1	Show Visible Only e Diversion of the only ence All rep 2. Click in the panel to save comment	All Types 1-50 of 163 Seq CP116280.1 C	GO         GP           Prov            All Statuses         ✓           Type         Length           gene         219           gene         2,432           gene         4,931           gene         4,517           gene         1,253           gene         1,606           acene         3,828           tation	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ← Close(x)
Show All           Annotation Nam           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G56060.2           AT1G52820.1           AT1G52820.1           AT1G52820.1           AT1G52820.1           GKIF01109402.1           gene: AT1G55710.1           Details         GO	Show Visible Only e Diversion of the only ence All ence Provenance	All Types All Types 1-50 of 163 Seq CP116280.1 CP	GO     GP       Prov        All Statuses        Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,517       gene     1,253       gene     3.828       tation	▼ Updated Jul 17, 2023 ↓ Jul 17, 2023 ↓
Show All           Annotation Nam           Reference Sequ           Rows         25 ~           Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G56060.2           AT1G52820.1           AT1G52820.1           AT1G52820.1           AT1G52820.1           GKIF01109402.1           gene: AT1G55710.1           Details         GO           Gene	Show Visible Only e Diversity of the only ence All ence All ence All ence All ence All ence All ence All ence All ence All ence Provenance	All Types	GO     GP       Prov        All Statuses        Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,517       gene     1,606       acne     3.828       tation	▼ Updated Jul 17, 2023 ↓ Jul 17, 2023 ↓ Close(x)
Show All           Annotation Nam           Reference Sequ           Rows         25 ✓           Name           AT1G55710.1           AT1624390.1           GKIF01155911.1           AT1G79620.1           AT1G556060.2           AT1G52820.1           AT1G52820.1           AT1G52820.1           GKIF01109402.1           gene: AT1G55710.1           Details         GO           Comment           Evidence for	Show Visible Only e Diversional Diversion of the original Diversion of	All Types	GO     GP       Prov     Prov       All Statuses     ~       Type     Length       gene     2,193       gene     2,432       gene     4,931       gene     1,253       gene     1,253       gene     1,606       acene     3.828       tation     Langth	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ← Close(x)
Show All           Annotation Nam           Reference Sequ           Rows         25 ✓           Name           AT1G55710.1           AT1624390.1           GKIF01155911.1           AT1G79620.1           AT1G5570.0.1           AT1G5520.1           AT1G52820.1           AT1G52820.1           AT1G52820.1           GKIF01109402.1           gene: AT1G55710.1           Details         GO           Gene           ▲ Comment           Evidence from	Show Visible Only e Diversional Diversion of the original Diversion of	All Types 1-50 of 163 Seq CP116280.1 C	GO     GP       Prov     All Statuses     ~       Type     Length       gene     2,19       gene     2,432       gene     4,931       gene     4,517       gene     1,253       gene     1,606       ocene     3.828	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ← Ju
Show All           Annotation Name           Reference Sequ           Rows         25 √           Name           AT1G55710.1           AT1624390.1           GKIF01155911.1           AT1676620.1           AT1676620.1           AT1675230.1           AT1675280.1           AT1652820.1           AT1652820.1           AT1652820.1           AT1652820.1           AT1652820.1           AT1652820.1           AT1652820.1           AT1652820.1           GKIF01109402.1           gene: AT1655710.1           Details         GO           Gene           ▲ Comment           Evidence from	Show Visible Only e ID ID ID ID IC ID	All Types 1 Users 1-50 of 163 Seq CP116280.1 CP1	GO     GP       Prov     Prov       All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,931       gene     1,606       acene     3.828       tation	▼ Updated Jul 17, 2023 ← Jul 17, 2023 ← Ju
Show All           Annotation Name           Annotation Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G7620.1           AT1G57600.2           AT1G17230.1           AT1G52820.1           AT1G55710.1           GKIF01109402.1           gene: AT1G55710.1           Details         GO           Comment           Evidence from	Show Visible Only e ID ID e ID ID e ID	All Types 1 Users v 1-50 of 163 Seq CP116280.1 CP	GO     GP       Prov        All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,931       gene     4,517       gene     1,253       gene     1,606       acene     3.828       tation	▼ Updated Jul 17, 2023 ↓ Jul 17, 2024 ↓ Jul 17, 2025 ↓ Ju
Show All           Annotation Name           Annotation Name           AT1G55710.1           AT1G24390.1           GKIF01155911.1           AT1G7620.1           AT1G5620.1           AT1G5620.1           AT1G5620.1           AT1G5620.1           AT1G55710.1           GKIF01109402.1           gene: AT1G55710.1           Details         GO           ▲ Comment           Evidence from	Show Visible Only e Proce All Product Provenance Product Provenance Step 1. Add comment in th	All Types Users v 1-50 of 163 Seq CP116280.1 CP11	GO     GP       Prov        All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,931       gene     4,931       gene     4,517       gene     1,253       gene     1,606       acne     3.828       tation       Attributes	▼ Updated Jul 17, 2023 ↓ Jul 17, 2024 ↓ Jul 17, 2025 ↓ Ju
Show All           Annotation Name           Reference Sequ           Rows         25 ∨           Name           AT1G55710.1           AT1G55710.1           AT1G55911.1           AT1G7620.1           AT1G56060.2           AT1G7620.1           AT1G5620.1           AT1G56060.2           AT1G17230.1           AT1G50794.1           GKIF01109402.1           gene: AT1G55710.1           Details         GO           Comment           Evidence from	Show Visible Only e ence All ence All ence All ence Product Provenance Product Provenance Step 1. Add comment in th panel	All Types All Types I Users v I-50 of 163 Seq CP116280.1 CP100000000000000000000000000000000000	GO     GP       Prov     All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,517       gene     1,253       gene     1,606       aene     3.828       tation	▼ Updated Jul 17, 2023 ↓ Jul 17, 2023 ↓ Markov All 17, 2023 ↓ Jul 17, 2024 ↓ Jul 17, 2025 ↓ Jul 17, 202
Show All           Annotation Name           Reference Sequ           Rows         25 ∨           Name           AT1G55710.1           AT1G55710.1           AT1G55911.1           AT1G7620.1           AT1G56060.2           AT1G7620.1           AT1G7620.1           AT1G56060.2           AT1G17230.1           AT1G52820.1           AT1G55710.1           GKIF01109402.1           gene: AT1G55710.1           Details           GO           Comment           Evidence from	Show Visible Only e ence All ence All ence All ence Product Provenance Product Provenance Step 1. Add comment in th panel	All Types All Types I Users v I-50 of 163 Seq CP116280.1 CP100000000000000000000000000000000000	GO     GP       Prov     All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     2,432       gene     4,931       gene     552       gene     1,253       gene     1,606       acene     3.828       tation	▼ Updated Jul 17, 2023 ↓ Jul 17, 2023 ↓ Markov All 17, 2023 ↓ Jul 17, 2023 ↓ Markov All 17, 2023 ↓ Jul 17, 2024 ↓ Jul
Show All           Annotation Name           Reference Sequ           Rows         25 ∨           Name           AT1G55710.1           AT1G55710.1           AT1G55911.1           AT1G7620.1           AT1G56060.2           AT1G7620.1           AT1G52820.1           AT1G56060.2           AT1G17230.1           AT1G52820.1           AT1G55710.1           GKIF01109402.1           gene: AT1G55710.1           Details           GO           Comment           Evidence from	Show Visible Only e ence All ence All ence All ence Product Provenance Product Provenance Step 1. Add comment in th panel	All Types All Types I Users 1-50 of 163 Seq CP116280.1 CP16	GO     GP       Prov       All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     2,432       gene     4,931       gene     552       gene     1,253       gene     1,253       gene     1,606       acene     3.828       tation       Attributes	▼ Updated Jul 17, 2023 ↓ Jul 17, 2023 ↓ Markov ↓
Show All         Annotation Name         Reference Seque         Rows       25 √         Name         AT1G55710.1         AT1G55710.1         AT1G55710.1         AT1G79620.1         AT1G56060.2         AT1G17230.1         AT1G5820.1         AT1G5820.1         AT1G56710.1         GKIF01109402.1         gene: AT1G55710.1         Details       GO         GO       Gene         ▲ Comment	Show Visible Only e ence All ence All ence All ence Product Provenance Product Provenance Step 1. Add comment in th panel	All Types All Types I Users	GO     GP       Prov     All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     4,931       gene     4,517       gene     1,253       gene     1,606       acene     3.828       tation	▼ Updated Jul 17, 2023 ↓ Jul 17, 2023 ↓ Markov ↓ Close(x) ment ↓
Show All         Annotation Name         Reference Sequ         Rows       25 √         Name         AT1G55710.1         AT1G55710.1         AT1G55710.1         AT1G79620.1         AT1G5700.1         AT1G582820.1         AT1G582820.1         AT1G59794.1         GKIF01109402.1         gene: AT1G55710.1         Details       GO         GO       Gene         ▲ Comment	Show Visible Only e ence All ence All ence All ence Product Provenance Product Provenance Step 1. Add comment in th panel	All Types All Types I Users 1-50 of 163 Seq CP116280.1 CP16	GO     GP       Prov       All Statuses     ~       Type     Length       gene     219       gene     2,432       gene     2,432       gene     4,931       gene     552       gene     1,253       gene     1,606       acene     3.828       tation       Attributes	▼ Updated Jul 17, 2023 4 Jul 17, 2024 4 Jul 17, 2025 4 Jul 17, 2025 4 Jul 17, 2025 4 Ju

#### 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 CoI-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)

0,000	10,000,000	15,000,000	20,000,000	25,000,000	30,000
$\rightarrow$ (	<b>ର୍ ବ୍ <del>ପ୍</del></b>	CP116284.1 - CP11628	4.1:2678091126789460 (8	.55 Kb) Go 🇳	
782,500		26,785,000	26,7	87,500	
tions 	AT5G58420)-00001		XM_990034645 (A	T5G58440)	
	XM 990034	4644 (AT5G58430)a-00001			
		134644 (AT5G58430)a-00002			
iotations. Dat	मि 5G58 <mark>420)</mark>				
	XM <mark>_99003</mark> 2 XM_990034	1645 (AT5G58440) 1644 (AT5G58430)			

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

5,000,000	10,000,000	15,000,000	20,000,000	25,000,000	30,000
$\supset \bigcirc \bigcirc$	Q @ 🕀	CP116284.1 - C	P116284.1:2678091126789460	(8.55 Kb) Go 💋	
26,782,500		26,785,000	26	5,787,500	
motations XM_990034646 (AT5	G58420)-00001		XM_990034645	(AT5G58440)	
	XM_990034	1644 (AT5G58430)a-00	001		
c_Annotations_Bata 5	G58 <b>420)</b>				
	VA 990032	6/15 (AT5658/J/0)			
	//003-	045 (AI 2020440)			
	XM_990034	644 (AI 5G58430)			

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

Ol-CC_Genomic_Annotations_Data	XR_990002624 (AT1G02840)		
	XR_990002625 (AT1G02840)		<b>⇒</b>
	XR_990002628 (AT1G02840)		>
	XM_990007286 (AT1G02840)		
	XM_990007292 (AT1G02840)		
	XM_990007298 (AT1G02840)		>
	XM_990007304 (AT1G02840)		

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.

×	Arabidopsis thaliana (Arabidopsis_ ~
≔ ]	Annotations Tracks Ref Sequence
	Show All Show Visible Only
	Annotation Name ID All Types V GO GP T
	Prov
	Reference Sequence       All Users       ✓ All Statuses         No Status Assigned       Any Status Assigned
	updated, secondary review requested
	secondary review complete, accepted updated, no secondary review needed

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

	=				
Rows 25 - (1)	1-50 of 620	▶	H		
Name	Seq	Туре	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	
170000000 4	00440004.4	-	470	0 1 40 0000	

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

• Click on the blue GO button.

XM_990017854 (AT1G10970)	CP	116280.1	gene	2,037	Oct 17, 2023	•	
XM_990005558 (AT1G30620)	CP	116280.1	gene	3,634	Oct 17, 2023	•	
XM_990001009 (AT1G24440)	CP	116280.1	gene	1,919	Oct 17, 2023	•	
AT5G39645	CP	116284.1	gene	289	Oct 17, 2023	•	
AT3G06895	CP	116282.1	gene	163	Oct 17, 2023	•	
AT4G05091	CP	116283.1	gene	415	Oct 17, 2023		
AT2G27402	CP	116281.1	gene	1,593	Oct 17, 2023	•	
gene: XM_990017854 (AT1G10970)	Link to annota	ation		Close(x)			
Details GO Gene Product P	ovenance DbXref	Comment At	ttributes				
O Go     ID     ID     ID       Type     gene	name with transcript	Obsolete	1 Ann	otations	间 Delete		
Name XM_990017854 (AT1	à10970)						
Symbol							
Aliases (' ' separated)							
Description isoforms update	Description isoforms updated						
Location 3669279 - 3671315	trand(-)			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.

⊖ Go	() ID	C Sync name v	No status selected	i Delet
Туре	gene	Status	<ul> <li>updated, secondary review requested secondary review complete, accepted</li> </ul>	
Name	XM_9900178	54 (AT1G10970)	updated, no secondary review needed	
			under secondary review	
Symbol			under primary review	
			unable to update	-
Aliases (	(' ' separated)		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)



# 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.

	21 21 21 21		
66.394	1.666.590	1.666.788	1.666.985
	1,		
AT4G03645.1-00001			
Type: mRNA			
Omment an eastern (@b)	a sur a sula		

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

AT4G03645.1-00001	Get Sequence (meta-click) Get GFF3 (alt-meta-click) Zoom to Base Level Open Annotation (alt-click) Close editor	
4T4G03645.1 None	Change annotation type Associate Transcript to Gene Dissociate Transcript from Gene Associate Feature to Gene Dissociate Feature from Gene	•
	View Variant Effect Remove Variant Effects	
	Delete Merge Split Duplicate Make Intron Move to Opposite Strand	
	Set Translation Start Set Translation End Set Longest ORF Set Optimal ORF	
	Remove CDS Set Readthrough Stop Codon	

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

1,000,390	1,000,700	
001	Get Sequence (meta-click) Get GFF3 (alt-meta-click) Zoom to Base Level	
	Open Annotation (alt-click) Close editor	

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.

-	File	View	, Help											
4,00	0,000		6,000,000		8,000,000	)	10,000,0	00	12,00	0,000		- 14	4,000,0	00
			$\rightarrow$	Θ	QÐ	Ð	CP1162	83.1 👻	CP116	283.1:1	6664	181	66656	7 (150 b
	1,66	6,450			1,666	475			1,66	6,500			_	_
					<b>C</b> 0		17 0	•	F D	N 7				e
L	TI		Q R S A I	TL		N S	C F		K T	Q H	H	L	SI	F I H Y
GCTT CGAA	GACA	TT GATAT A A CTATA	AAGTCGCC	ACCTT:	ATTGCCA FAACGG1	AACTC	ATGTTTC. TACAAAG	ATAAAG	GAAAACC CTTTTGG	GTTGT	CCA	GGA	GAGCT	TTCATT
A A	S ) Q C			S R G	I A		M N T E	Y I	SF	G V	G V	G V	S S Q A	E N
				V K	N G	F E	H K	M F	F V	W (		C C T	L	K M
AT4	G0364	5.1-00001	TTURGUGU	RUUTTI	ATTGUUP	AAUTU	Get Seq Get GFF Zoom Ba	uence (n 3 (alt-m ack Out	neta-click eta-click)	;)			GRECT	FTCATI
Data	a						Open An Close ed	notation litor	(alt-click	.)				
							Change Associat Dissociat Associat Dissociat	annotati e Transo te Trans e Featur te Featu	on type cript to Ge cript from re to Gen re from G	ene 1 <b>Gene</b> e àene	•			
							View Var Remove	iant Effe <b>Variant</b>	ect Effects					
							Delete Merge Split Duplicate Make Int Move to	e ron Opposite	e Strand					
							Set Tran Set Tran Set Long Set Optin Remove Set Read	slation S slation E jest ORF mal ORF CDS dthrough	Start End End Stop Co	don				