

How do I update functional annotation in Apollo?

1. Log into VEuPathDB
2. From the gene record page select “View and update community annotations in Apollo”
3. With a right click on the gene in the Annotated Transcript track open the drop-down menu and select “Create new annotation > gene”
4. Right click on the gene in the User-created Annotations, select “Open Annotation”
5. The annotation panel is on the right-hand side. You can now add functional annotation, i.e. Symbol.
6. Select the tab Gene Product to add a product description with metadata. There are additional tabs to add Gene Ontology terms (the tab GO) and the tab DbXref to add publications.
7. Once you’ve added the functional annotation select from the Status drop-down menu “Finished”.
8. If you used the tab gene product to add a product description with metadata, the new product is visible the following day on the gene record page in the section “Apollo product description”.

ToxoDB Release 54
8 Sept 2021
Toxoplasma Informatics Resources

TGME49_202220 hypothetical protein

Type: protein coding gene
Chromosome: VIIa
Location: TGME49_chrVIIa:3,509,010..3,514,137(-)
Species: *Toxoplasma gondii*
Strain: ME49
Status: Reference Strain

View 1 user comment, or add a comment
View and update community annotations in Apollo

Toxoplasma gondii ME49 File View Help

Select tracks: 3,508,750 3,510,000 3,511,250 3,512,500 3,513,750

User-created Annotations

Annotated Transcripts (UTRs in White when available)

histone deacetylase HDAC5
ToxME49_202220-1
TGME49_202220
hypothetical protein

View Details
View Gene Page
Highlight this gene
Create new annotation > gene
pseudogene
tRNA
snRNA

Toxoplasma gondii ME49 File View Help

Select tracks: 3,508,750 3,510,000 3,511,250 3,512,500 3,513,750

User-created Annotations

TGME49_202220-t26_1-00001

Annotated Transcripts (UTRs in White when available)

histone deacetylase HDAC5
ToxME49_202220-1
TGME49_202220
hypothetical protein

Get Sequence (meta-click)
Get GFF3 (all-meta-click)
Zoom to Base Level
Open Annotation (alt-click)
Close editor

Toxoplasma gondii ME49 [Mar 22, 2021] ucb.40063473

Annotations Tracks Ref Sequence Search Organism Users Groups Admin

Show All Show Visible Only

40d3d0b-4cf9-452f-9f5e-c ID All Types GO: GP: Prov

Reference Sequence All Users All Statuses

Rows 25 1-1 of 1

Name	Seq	Type	Length	Updated
TGME49_202220-t26_1	TGME49_chrVIIa	gene	5,128	Oct 13, 2021
TGME49_202220-t26_1-00001	mRNA	5,128	Oct 13, 2021	

gene: TGME49_202220-t26_1 Link to annotation Close(x)

Details GO Gene Product Provenance DbXref Comment Attributes

Go ID Sync name with transcript Obsolete Annotations Delete

Type gene Status No status select

Name TGME49_202220-t26_1

Symbol ISAP1

Aliases (separated)

Description

Location 3509010 - 3514137 strand(-) Partial: ☐ 3' ☐ 5'

Ref Sequence TGME49_chrVIIa

Owner ucb.40063473

Created Oct 13, 2021 10:43 AM

Updated Oct 13, 2021 10:43 AM

Add new Gene Product to TGME49_202220-t26_1

GO Annotation Guidance

Product IMC suture-associated protein 1 Alternate Evidence ECO:0000314 All ECO Evidence

IDA (ECO:0000314): direct assay evidence used in manual assertion Evidence Code Info

With Prefix ID + Add

Reference PMID 34634933 + Add

Note + Add

Save Cancel

Details GO Gene Product Provenance DbXref Comment Attributes

Go ID Sync name with transcript Obsolete Annotations

Type gene Status No status selected Not Finished Finished Requires Curator Rejected

Name TGME49_202220-t26_1

Symbol ISAP1

TGME49_202220 expand all collapse all

Search section names...

1 Gene models

2 Annotation, curation and identifiers

2 Annotation, curation and identifiers

Apollo Product Description Data sets

IMC suture-associated protein 1

How do I update structural annotation in Apollo?

1. Log into VEuPathDB
2. From the gene record page select "View and update community annotations in Apollo"
3. With a right click on the gene in the Annotated Transcript track open the drop-down menu and select "Create new annotation > gene"
4. Right click on the gene in the User-created Annotations, select "Open Annotation"
5. The annotation panel is on the right-hand side. Select the tab Tracks.
6. Select evidence tracks from Draggable Annotation
7. Drag and drop the evidence into the User-created Annotations area
8. With a right-click on the gene in the User-created Annotations area open the drop-down menu. There are options to merge exons, split and duplicate genes. Choose the modification as appropriate.
9. Once you've finished modifying the gene structure, choose the Annotations tab and select the status finished.

The screenshots show the following steps:

- Screenshot 1:** The ToxoDB website showing the gene record for *TGME49_223950* (hypothetical protein). It includes details like chromosome (X), location, species (*Toxoplasma gondii*), and strain (ME49).
- Screenshot 2:** The Apollo interface showing the gene record for *TGME49_223950* with a right-click menu open, highlighting "Create new annotation > gene".
- Screenshot 3:** The Apollo interface showing the gene record for *TGME49_223950* with a right-click menu open, highlighting "Open Annotation".
- Screenshot 4:** The Apollo interface showing the gene record for *TGME49_223950* with the "Tracks" tab selected in the annotation panel.
- Screenshot 5:** The Apollo interface showing the gene record for *TGME49_223950* with the "Draggable Annotation" section expanded, showing evidence tracks like "RNA-Seq Evidence for Introns".
- Screenshot 6:** The Apollo interface showing the gene record for *TGME49_223950* with the "Draggable Annotation" section expanded, showing evidence tracks like "RNA-Seq Evidence for Introns".
- Screenshot 7:** The Apollo interface showing the gene record for *TGME49_223950* with the "Draggable Annotation" section expanded, showing evidence tracks like "RNA-Seq Evidence for Introns".
- Screenshot 8:** The Apollo interface showing the gene record for *TGME49_223950* with the "Draggable Annotation" section expanded, showing evidence tracks like "RNA-Seq Evidence for Introns".
- Screenshot 9:** The Apollo interface showing the gene record for *TGME49_223950* with the "Annotations" tab selected, showing the status "Finished" and the "Requires Curator" checkbox.

How do I know if my new gene structure is correct in Apollo?

1. Check if your gene has a Start and Stop site. You can do this by zooming in. Alternatively, right click on the gene to open the menu and select Get sequence. With this option you can recheck if there is a correct Start. You only need to check for correct start and stop sites if your gene does not have annotated UTRs.
2. Check splice sites. In most Eukaryotes the majority of splice sites at the exon/intron boundaries appear as 5'-...exon]GT/AG[exon...-3'. An exclamation mark in Apollo shows non-canonical splice sites.
3. Check if you can annotate UTRs, for example using RNA-Seq data. A useful track that could help with this is the "Combined RNA-Seq plot". This track combines all RNA-Seq plots available for your genome of interest. You can find this track by searching for "combine".
4. Check sequence similarity by running blast. Right click on the gene in the User-created Annotations area, open the drop down menu and select Get sequence. Copy the sequence to the clipboard and run blast (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

The screenshot displays the Apollo genome browser interface with several key components:

- Top Panel:** Shows the reference sequence and user-created annotations for the gene *TGME49*. A right-click context menu is open over a gene model, showing options: "Get Sequence (meta-click)", "Get GFF3 (alt-meta-click)", and "Zoom to Base Level".
- Gene Model:** A track showing the gene structure with exons as boxes and introns as lines. A right-click context menu is also open over the gene model, showing the same three options.
- Sequence Window:** A window titled "Sequence" showing the DNA sequence of the gene. The sequence is:


```
>24a0283c-cf6e-4adc-b66f-4de6928f4e89 [sequence:exon] 245
residues [TGME49_chrX:3440383-3442160] [strand] [peptide]
MGNACKKNTAKTPTRKEADLAERKERQEREKAEAEKAEAEKNAADKAEAEERRAA
EAREERESARKEAEAEAEAEAEAEERLRKEAEKKKAEAEKRAEEQRAAEAEAEQ
RAREEAERRKAEAEAEAEAEAEAEERERQAEALQEEMSPREKYDKLASPEDSASETTMATQPKV
AEHSSAAVTDRTSVVGYTVTPCDMASIDETAKYLSKRCGCDLGDQHDENECPICRHIDLS
APLLN
```
- Annotations Panel:** A panel on the right showing various annotation tracks. The "combine" track is selected, and the "Transcriptomics" track is checked, showing the "TgondilME49 combined RNAseq plot".

Which evidence code should be used when adding GO terms or gene products in Apollo?

1. IDA is used if the GO term/product has been experimentally verified
2. ISS: GO term/product is inferred from similarity. For GO annotations add GO_REF:0000024 as reference. GO references describe the methods and procedures used in creating such annotations.
3. ISM: GO term/product is inferred from InterPro/Pfam domain. Add GO_REF:000011 in the reference field for GO annotations.
4. IMP: GO term/product is inferred from mutant phenotype

Additional information on evidence codes can be found here:

<http://geneontology.org/docs/guide-go-evidence-codes>

A list of GO references can be found here:

<http://current.geneontology.org/metadata/gorefs/index.html>

The figure displays four screenshots of the Apollo GO Annotation Guidance interface, each illustrating a different evidence code and its corresponding reference field. The interface is divided into two main sections: 'GO Annotation Guidance' and 'Evidence Code Info'.

Screenshot 1: Shows the 'GO Annotation Guidance' section with 'Aspect' set to 'CC' (cellular component) and 'Go Term' set to 'GO:0020011' (apicoplast). The 'Relationship between Gene Product and GO Term' is set to 'part of'. The 'Evidence' field is set to 'ECO:0000314' (IDA). The 'Evidence Code Info' section shows 'With' set to 'Prefix' and 'ID' set to '17449654'. The 'Reference' field is set to 'PMID' and '17449654'. The 'Note' field is empty.

Screenshot 2: Shows the 'GO Annotation Guidance' section with 'Aspect' set to 'BP' (biological process) and 'Go Term' set to 'GO:0016571' (histone methylation). The 'Relationship between Gene Product and GO Term' is set to 'involved in'. The 'Evidence' field is set to 'ECO:0000250' (ISS). The 'Evidence Code Info' section shows 'With' set to 'UniProtKB' and 'Q86X55'. The 'Reference' field is set to 'GO_REF' and '0000024'. The 'Note' field is empty.

Screenshot 3: Shows the 'GO Annotation Guidance' section with 'Aspect' set to 'MF' (molecular function) and 'Go Term' set to 'GO:0003676' (nucleic acid binding). The 'Relationship between Gene Product and GO Term' is set to 'enables'. The 'Evidence' field is set to 'ECO:0000255' (ISM). The 'Evidence Code Info' section shows 'With' set to 'InterPro' and 'IPR003029'. The 'Reference' field is set to 'GO_REF' and '0000011'. The 'Note' field is empty.

Screenshot 4: Shows the 'GO Annotation Guidance' section with 'Aspect' set to 'BP' (biological process) and 'Go Term' set to 'GO:0042493' (response to drug). The 'Relationship between Gene Product and GO Term' is set to 'involved in'. The 'Evidence' field is set to 'ECO:0000315' (IMP). The 'Evidence Code Info' section shows 'With' set to 'Prefix' and 'ID' set to '24352242'. The 'Reference' field is set to 'PMID' and '24352242'. The 'Note' field is empty.

How do I add a new gene model in Apollo based on a blast match?

1. Log into VEuPathDB
2. Select on the right-hand side the tab "Search", choose "Blat protein" or "Blat nucleotide" from the drop-down menu and paste in the sequence. Select the tick box "All genome sequences" if you want to search all chromosomes/contigs of your genome of interest. Click on Search.
3. You can now select the search results. They are shown in yellow.
4. If you want to create a gene model based on the search result, click on Action and select Create annotation.

Step 2: Search

Annotations Tracks Ref Sequence **Search** Organism Users Groups Admin

Blat protein ☒ All genome sequences Search Clear

MEEGKKEKSLSGQSDQGVQSRSLSRVSACVSNLASERRKSSGGNSLRVSLPQTRSASLRFLSVFCCLGF
AAVASVCGKLALDFHSQAPLSRLSLVLVDWLSPLPGSSSPSCFASSEEDGGAQSRDPDRGEEQSRDPER
EERPESEGPSRTEKTDACRGQRGEASPSVARLPSSLLSLHAIWNTNCFLGSCHAETRASLAGDRNAST
LLPSSPLSSSASEVFSRRNSLSVRPPSLCEGVLLPCSLLLLGILRVAPLLMALSNALLHCQFRA

Step 3: Search Results

Annotations Tracks Ref Sequence Search Organism Users Groups Admin

Blat protein ☒ All genome sequences Search Clear

1-50 of 190

ID	Start	End	Strand	Score	Significance	Identity	Action
TGME49_2,613,817	2,614,875	1	685	0	99.43	--	
TGME49_482,522	482,608	-1	26	42	62.07	--	
TGME49_1,812,902	1,812,964	1	22	620	66.67	--	

Step 4: Create Annotation

Annotations Tracks Ref Sequence Search Organism Users Groups Admin

Blat protein ☒ All genome sequences Search Clear

1-50 of 190

ID	Start	End	Strand	Score	Significance	Identity	Action
TGME49_2,613,817	2,614,875	1	685	0	99.43	--	

Save sequence
Create annotation