

| | |
|---|----|
| Community Annotation Users Manual | 1 |
| Registration: | 1 |
| To Login:..... | 1 |
| To change your password or other personal information:..... | 2 |
| Adding an annotation:..... | 2 |
| Gene Symbol: | 3 |
| Synonyms: | 4 |
| Gene name: | 4 |
| Add a GO term: | 4 |
| Add a Fungal Anatomy Term: | 7 |
| Add Additional Citation: | 7 |
| Transcripts: | 8 |
| New Transcript..... | 10 |
| Description | 10 |
| Saving an annotation | 10 |
| Comments..... | 10 |
| Editing and Searching for Annotations | 11 |
| Contact information | 11 |

Community Annotation Users Manual

If you have any questions, please feel free to contact Heather Hood (hmhood@ebs.ogi.edu). Full contact information is at the end of the manual.

Registration:

You must be registered in order to submit annotations. From the Community Annotation home page <http://www.broad.mit.edu/annotation/genome/neurospora/CAHome.html>, click on **request** in the first sentence on this page (Figure 1A).

To Login:

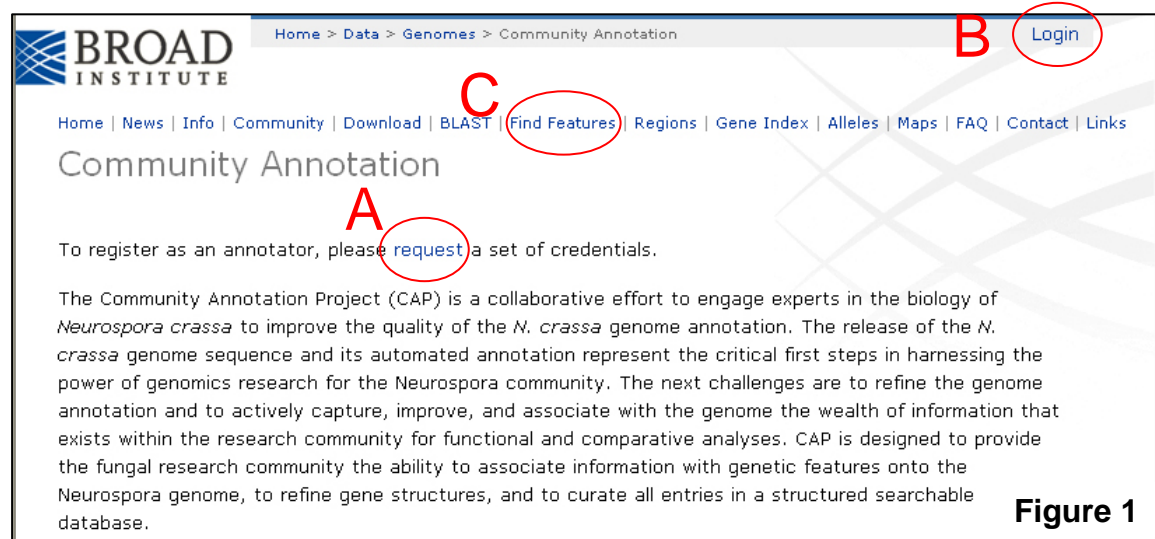
Login is located on the top right hand corner of all web pages on the *Neurospora* site. Click on **login** (Figure 1B), fill in your username and password, and click on submit. After submitting your information, you will be returned to the page you were visiting prior to login.

******IMPORTANT****:** You will not be logged out of a session unless you click **logout**, which is located in the same place as the **login** field. If you are using a shared computer, it is important for you to **logout** when you are finished.

If you forget your password or user name, please contact Heather Hood, and she will be able to retrieve your user name and/or reset your password.

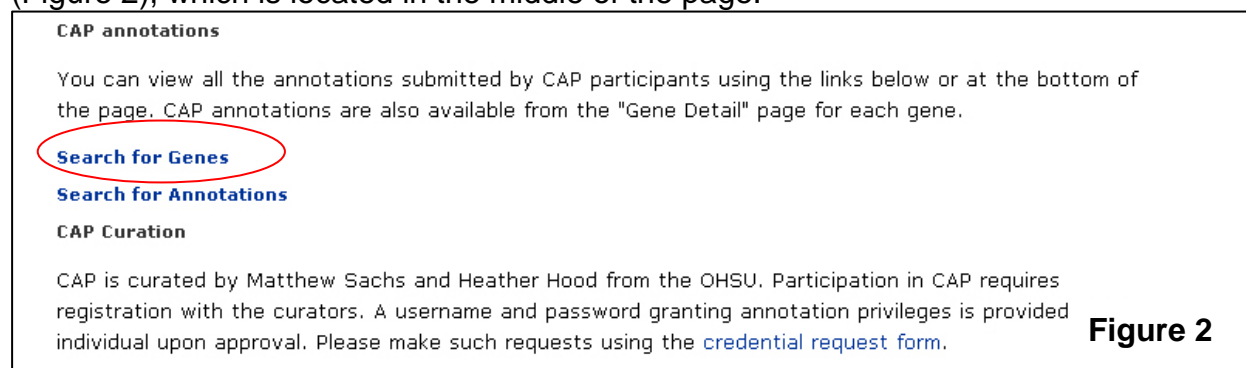
To change your password or other personal information:

On the Community Annotation homepage, once you are logged in, your name in the **CAP Collaborators** list will be in hypertext. Click on your name and you will be redirected to a page where you can change your personal information including your password.



Adding an annotation:

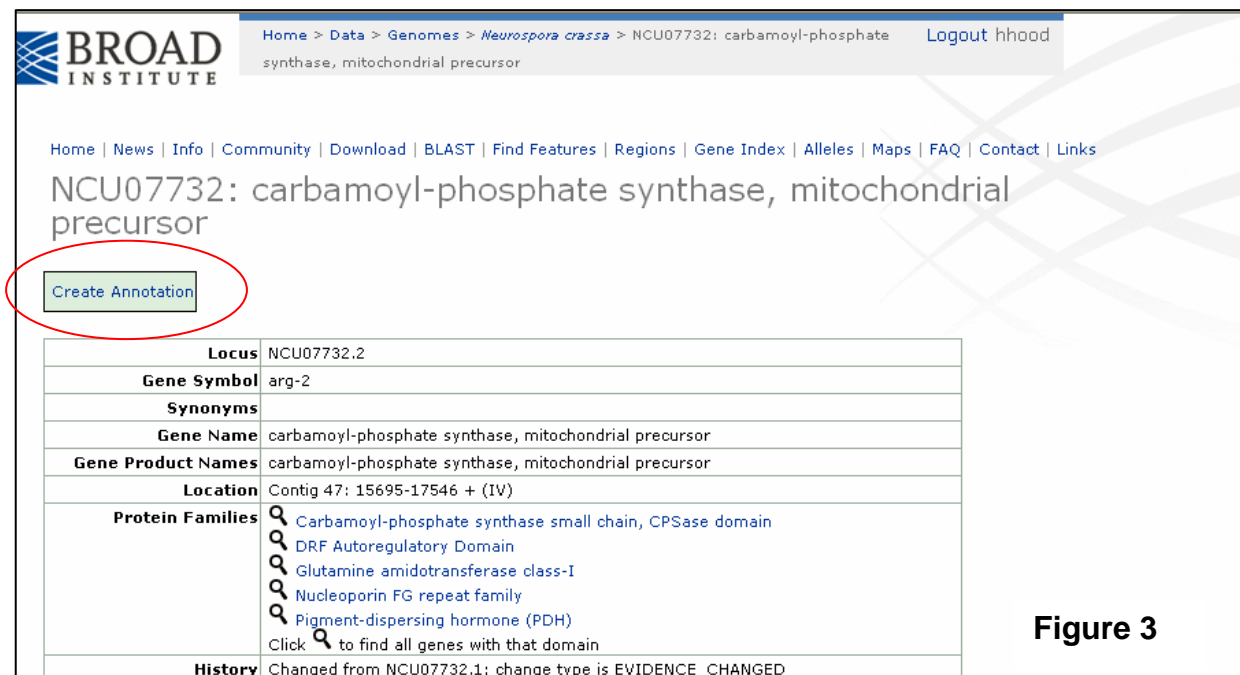
If you know which NCU identifier you want to annotate, you can retrieve its gene detail page by using the **Find Features** (Figure 1C) at the top of any page and by searching by using NCUxxxxx. Also, on the community annotation page, you can use **Search for genes** (Figure 2), which is located in the middle of the page.



Once you are on the gene detail page for the NCU identifier you wish to annotate, click the **Create Annotation** button (Figure 3). If you are not logged in, this button is not visible. This will open a new page where you add information to your gene of interest. Detailed information follows.

Original information is shown in white fields and community annotation information is in green fields. The fields that you can edit have a pencil icon next them (Figure 4) or will be in hypertext.

If you are adding information to different fields for the same gene, you do *not* have to start a new annotation for each field. Save what you are adding at each step and simply continue to the next field that you want to edit.



The screenshot shows the Broad Institute website interface. At the top, there is a navigation bar with the Broad Institute logo and a breadcrumb trail: Home > Data > Genomes > *Neurospora crassa* > NCU07732: carbamoyl-phosphate synthase, mitochondrial precursor. A 'Logout hhood' link is also present. Below the navigation bar, there is a secondary navigation bar with links: Home | News | Info | Community | Download | BLAST | Find Features | Regions | Gene Index | Alleles | Maps | FAQ | Contact | Links. The main heading is 'NCU07732: carbamoyl-phosphate synthase, mitochondrial precursor'. Below this heading, a 'Create Annotation' button is circled in red. The main content area is a table with the following fields:

| | |
|---------------------------|--|
| Locus | NCU07732.2 |
| Gene Symbol | arg-2 |
| Synonyms | |
| Gene Name | carbamoyl-phosphate synthase, mitochondrial precursor |
| Gene Product Names | carbamoyl-phosphate synthase, mitochondrial precursor |
| Location | Contig 47: 15695-17546 + (IV) |
| Protein Families | <ul style="list-style-type: none"> Carbamoyl-phosphate synthase small chain, CPSase domain DRF Autoregulatory Domain Glutamine amidotransferase class-I Nucleoporin FG repeat family Pigment-dispersing hormone (PDH) |
| History | Changed from NCU07732.1; change type is EVIDENCE_CHANGED |

Figure 3





Gene Symbol:

To add or correct a gene symbol, click on the pencil icon (Figure 4A), which will open a field for you to insert information. When you are finished, click on the diskette icon to save this part of your annotation (Figure 4B).

N. crassa gene symbols are generally three lower case letters and a number separated by a hyphen (e.g., *arg-2*). Please see the article by David Perkins <http://www.fgsc.net/fgn46/perkins.htm> for the guidelines concerning proper *N. crassa* gene nomenclature.

Viewing Annotation: (CA-1085) arginine-2 of Gene NCU07732

Viewing CA-1085 [Done](#)

| | |
|--------------------|--|
| Locus | NCU07732.2 |
| Gene Symbol | arg-2 |
| | arg-2   A |
| Synonyms | |
| |  |
| Gene Name | carbamoyl-phosphate synthase, mitochondrial precursor |
| | arginine-2  |

Gene Symbol arg-2

arg-2 

B

Figure 4

Synonyms:

Alternative names for genes go in this field (i.e., mutant alleles identified by different research groups led to different names for the same gene). For example, *ccg-2* and *bli-7* are alleles of *eas* and are inserted into the **Synonyms** field. Click on the pencil icon to open this field for editing, insert your information, and click on the diskette icon to save. To add more than one synonym, each needs to be separated by a space.

Gene name:

The **Gene Name** field is where you insert the full gene name (not the gene symbol). Example gene names are *arginine-2* and *easily wettable*. Click on the pencil icon to open the **Gene Name** field, insert the information, and click on the diskette icon to save. Important note: Gene Product Names (e.g., arginine-specific carbamoyl-phosphate synthetase small subunit, mitochondrial precursor) are added as annotations using the **Transcripts** field (see pages 7-8); these are different from gene names.

Add a GO term:

Gene Ontology (GO) terms for biological process, molecular function, and cellular component are added here along with a reference(s). When you click on the **Add a GO term**, you will be redirected to a new page headed **Citation** (Figure 5).

The screenshot shows the 'Citation' page of the Broad Institute's annotation system. At the top is the Broad Institute logo and a navigation bar with links: Home, Data, Genomes, Citation, and Logout. Below this is a secondary navigation bar with links: Home, News, Info, Community, Download, BLAST, Find Features, Regions, Gene Index, Alleles, Maps, FAQ, Contact, and Links. The main heading is 'Citation'. A usage note states: 'Usage note : if you use the PubMed Id field, the Publication field will be populated automatically. To reference a Publication that is not in PubMed, leave the PubMed Id field blank.' The form contains several fields: 'Gene Ontology Term' with a link 'Choose Gene Ontology Term' (labeled A), 'PubMed Id' with an 'Open Entrez PubMed' link (labeled B), 'Publication' with a text area showing '[unspecified]' (labeled C), and 'Evidence Code' with a dropdown menu set to 'IC: Inferred by Curator' and a link '[Guide to Gene Ontology Evidence Codes]' (labeled D). At the bottom are three buttons: 'Validate' (labeled E), 'Save' (labeled F), and 'Cancel'. The footer includes links: Home, Internal Web, Contact Us, Site Map, Search, and a copyright notice: © 2006 Bro.

Figure 5

Click on **Choose Gene Ontology Term** (Figure 5A) to open a search page, which is called **Find Ontology Term** (Figure 6). See this link for general search tips <http://www.broad.mit.edu/annotation/genome/neurospora/TextSearchHelp.html>.

If you already know the GO ID number, you can search using it. Another way to search is to put your query in quotes (e.g., “arginine metabolism” or “DNA damage”). If you just put these phrases in without the quotation marks your search will return everything that contains these two words but not necessarily in the same phrase, thus returning too many GO terms.

Enter your query and click **Find Terms** (Figure 6A). When you find the appropriate GO term and definition that describes the molecular function, biological process, or cellular component, click on that term to select it (Figure 6B). This will direct to a page showing the detailed information about the term you selected (e.g., parent term and child term relationships or more granular terms vs. fine-grained terms) (Figure 6C). Click **select this term** if you are satisfied with it (Figure 6C); you will be redirected to a page similar to Figure 5 (see above) with the GO term filled in.

Next, insert the publication that provides the evidence for the GO term. There are two ways to insert your citation:

- 1) insert the PubMed ID (PMID) which can be found on either the summary or the abstract page of the article in PubMed. You can use the **Open Entrez PubMed** link to search for the article (Figure 5B; this will open a new page).
- 2) You can manually type (or copy and paste) in the citation if it is not in PubMed (Figure 5C; e.g., meeting abstract or dissertation).

If the article is in PubMed, use the PMID because PubMed and annotation databases can then be linked by this identifier.

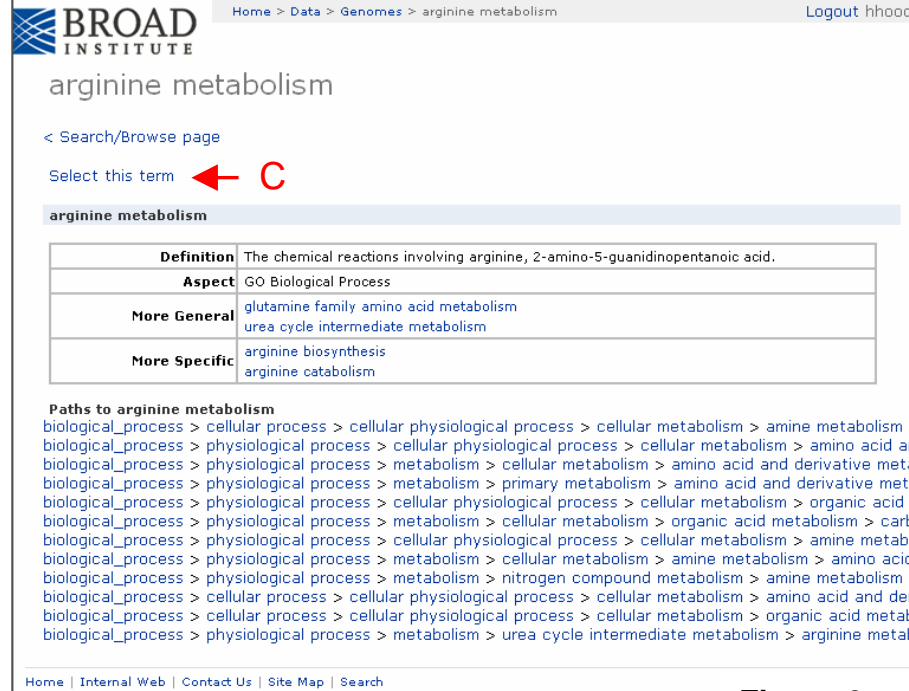
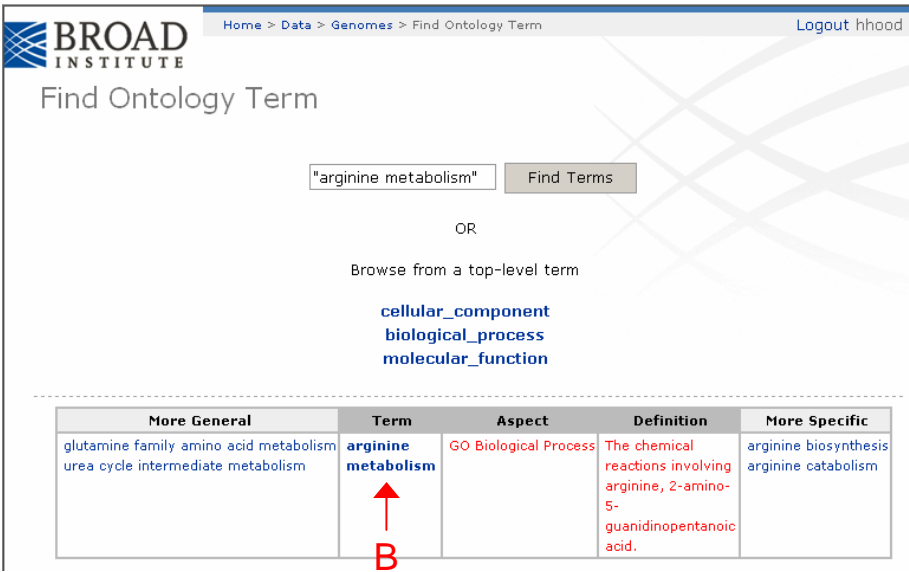
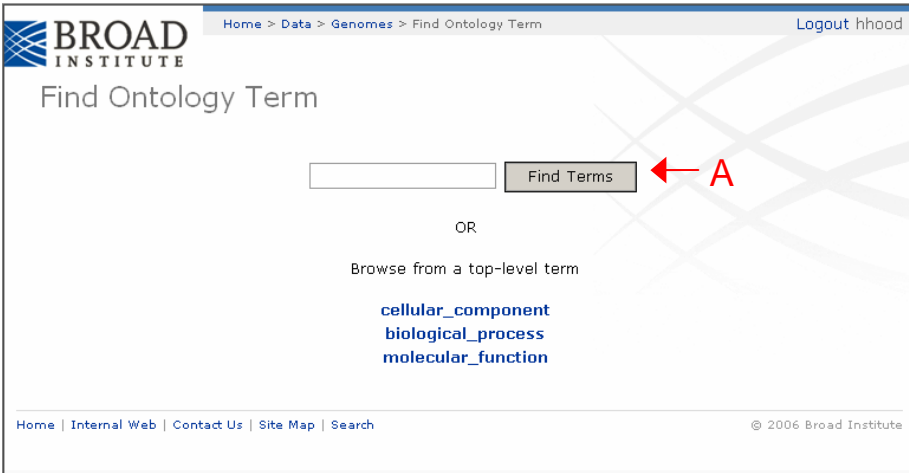


Figure 6

After you have inserted the citation information, select the **Evidence Code** from the drop-down menu indicating the type of data that supports the GO term. To familiarize yourself with the evidence codes, click on the link **Guide to Gene Ontology Evidence Codes** (**Figure 5D**). When you are finished providing the evidence code, click the **Validate** button to check that the proper citation is retrieved. If the correct citation populates the **Publication** field (**Figure 5C**), click the **save** button (**Figure 5F**) and all the information you have inserted for the GO term will be saved. Only one reference can be cited per GO term. If you have another reference that you wish to cite for the same GO term, you will have to repeat the above process.

If you have made a mistake and need to start over, click on the **Cancel** button. None of your Gene Ontology information will be saved in the database.

Please note: the Gene Ontology is a work in progress. If there are terms that should be defined but are not listed, please contact Heather Hood (hmhood@ebs.ogi.edu) and she will assist with updating the ontology.

Add a Fungal Anatomy Term:

To **Add a Fungal Anatomy Term**, the approach is similar to the **Add a GO term**. Clicking on the **Add a Fungal Anatomy Term** will direct you to a search page. Click **Choose Fungal Anatomy Ontology Term**, enter the anatomy term you are searching for (e.g., hyphae) and click **Find Terms**. This will retrieve a similar page as **Figure 6B** (see above). Click on the appropriate term. The next page will display the term, its definition, and the path to the definition (i.e., more general or specific terms). If this term is appropriate, click **select this term**. You will be redirected to the search page with your term listed in the **Fungal Anatomy Ontology Term** field. Next add the citation that supports this term either using a PubMed ID or by manually inserting it into the **Publication** field. Select the appropriate **Evidence Code** from the drop-down menu. When you are finished adding the information, click **validate** to make sure that the correct citation was retrieved from PubMed. When all information is correct, click **save** and your data will be appended to the NCU identifier. The Fungal Anatomy Ontology is also a work in progress. If there are terms that should be defined but are not listed, please contact Heather Hood (hmhood@ebs.ogi.edu) and she will assist with updating the ontology.

If you have made a mistake and need to start over, click on the **Cancel** button. None of your Fungal Anatomy information will be saved in the database.

Add Additional Citation:

Here you can add a citation without having to add Gene Ontology or Fungal Anatomy terms. Click on **Add additional citation** to get to the citation page. Either insert the PubMed ID or manually type (or copy and paste) in the publication information. If you inserted a PubMed ID click on **validate** to make sure the database retrieved the correct citation. When finished, click on **save** and this publication will be added to your annotation.

Transcripts:

This section is used to edit/correct transcript information. Here you can manually change the coordinates of intron/exon boundaries, add/delete exons, add 5' and 3' untranslated regions (UTRS), and add/change the gene product name. To edit the transcript, begin by clicking on the pencil in the **Transcripts** section (Figure 7A)

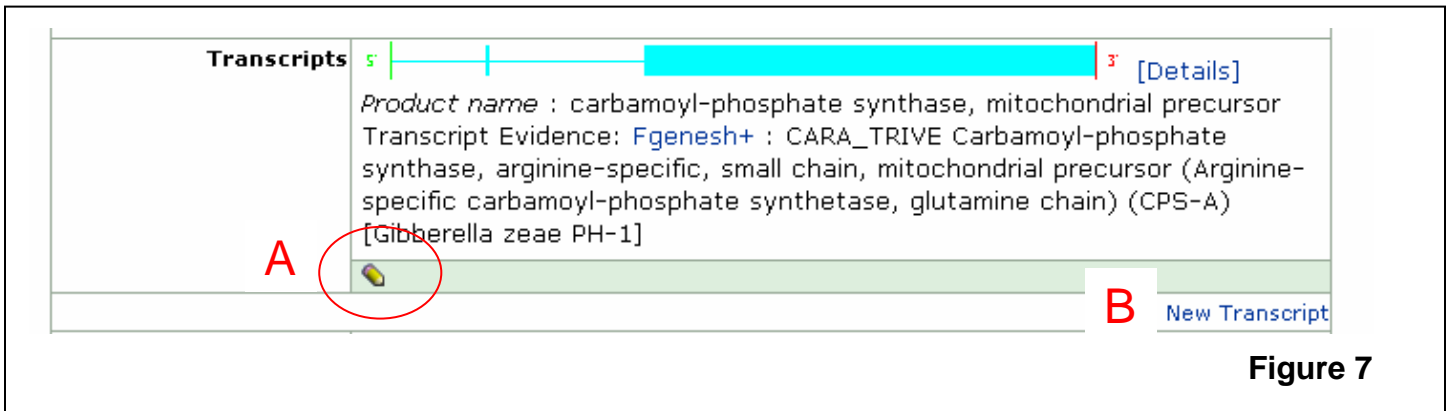


Figure 7

You will be redirected to the **Edit Transcript** page (Figure 8). The original transcript information such as the exon coordinates, coding sequence (CDS) start and stop (CDS Start and CDS stop) is shown at the top of the page (Figure 8A). In the middle of the page, under **Transcript being edited** is where changes are made (Figure 8B). To change the start and stop of an exon, insert the contig coordinates in the boxes. To insert a new exon, click on **New Exon** located under the **start** box(es). A set of two boxes (for the start and stop of the exon) will appear after the other exons, so you may have to shift the coordinates if the exon you are adding is in-between the other exons. To delete an exon, click on the “x” button at the end of the row that contains the exon that may be wrongly predicted.

Helpful Hint: An easy way to find the exon coordinates is to perform a tblastn query using the peptide sequence on the Broad Neurospora site. It is important to check that splice sites are correct.

A very important and useful feature in the transcript editor is the ability to add transcript UTRs. Depending on the structure of the UTR, you can either extend an exon or add a new exon to include the UTR. The way to denote UTRs is to indicate the location of the first nucleotide of the start codon and the final nucleotide of the stop codon. That way, the underlying program is provided with coordinates to start and finish of the conceptual translation.


Product name is editable (Figure 8B). Delete the name that is present (e.g., hypothetical protein) and insert the correct name.

When you are finished, click the **validate** button (Figure 8C). The edited transcript will be automatically checked against a set of rules (e.g., open frame length is a multiple of three and proper start and stop codons have been chosen). If an error message is returned, double-check your entry and adjust coordinates. See the **Transcript Validation Reference**

for more rules

(<http://www.broad.mit.edu/annotation/genome/neurospora/TranscriptCheckReference.html>).

Click the validate button again and if correct, click save.


Home > Data > Genomes > Gene Details > Edit Transcript
Logout hhood

[Home](#) | [News](#) | [Info](#) | [Community](#) | [Download](#) | [BLAST](#) | [Find Features](#) | [Regions](#) | [Gene Index](#) | [Alleles](#) | [Maps](#) | [FAQ](#) | [Contact](#) | [Links](#)

Edit Transcript

Original transcripts:

| Start | Stop | Length | Strand |
|-------|-------|--------|--------|
| 15695 | 15699 | 5 | + |
| 15948 | 15960 | 13 | + |
| 16365 | 17546 | 1182 | + |

CDS Start


15695

CDS Stop

17546

Product name

carbamoyl-phosphate synthase, mitochondrial precursor



Transcript being edited:


| Start | Stop | Length | Strand | |
|------------------------------------|------------------------------------|--------|--------|----------------------------------|
| <input type="text" value="15695"/> | <input type="text" value="15699"/> | 5 | + | <input type="button" value="X"/> |
| <input type="text" value="15948"/> | <input type="text" value="15960"/> | 13 | + | <input type="button" value="X"/> |
| <input type="text" value="16365"/> | <input type="text" value="17546"/> | 1182 | + | <input type="button" value="X"/> |

New Exon

CDS Start

CDS Stop

Product name



[Transcript validation reference](#)

Translation:

MLITSQDGPVFQGTAFGANTNISGEAVFTTSLVGYPESMTDPSYRGQILVFTQPLIGNYGVPSENERDEFNLLKYFESPHI
 QCAGIVVSDVATQYSHWTAQSLGEWCASEGIPAISGVDTRAIVTYLREQGSSLARISIGDEYDADEDEGFIDPGQINLV
 KRVSTKAPFVVTPNPAKFHVALIDCGVKENILRSLVSRGASVTVPYNYPIHKVAENFDGVFISNGPGDPHQCQETVYNL
 AKLMETSPIPINGICIGHQLLALAVGAKTIKLYGNRAHNIPALDLTTGQCHITSQNHGYAVDISTLPSEDFKEYFVNLND

Done

Figure 8

New Transcript

This section is used to add a **New Transcript** (Figure 7B), such as an alternatively spliced transcript, in contrast to fixing the existing CDS. The mechanism is the same as editing a transcript (see above section).

Description

This is the last field on the annotation page. Here you can information describing the gene (e.g., Gamma subunit of heterotrimeric GTP binding protein).

Saving an annotation

When you are finished adding all your information to a gene, click the **Done** button at the top of the gene detail page. Your annotation will be immediately appended to the bottom of gene detail page for that NCU and become publicly available. An automated e-mail is also sent to curators to alert them that a new annotation was created. They will double check the submission and may contact you if an annotation needs some clarification.

Comments

After the annotation is saved, you can add a **Comment** from the Community annotations summary on the gene detail page (Figure 9). Comments can be added to either your annotation or to someone else's. Comments can be marked as private so that only you and the curators will be able to view them.





| Community Annotations | | | | | |
|-----------------------|------------|--|---|---------|--------|
| Id | Submitter | Summary | Comments | Curated | Delete |
| CA-1085 | hhood | Edited transcripts : arginine-specific carbamoyl-phosphate synthetase small subunit, mitochondrial precursor  , [unnamed product]  | <div>[enter comment]</div> <div> Private</div> | | |
| CA-2958 | compendium | Added terms : Added 14 publications Edited transcripts : small (glutaminase) subunit of arginine-specific carbamoyl phosphate synthetase  | | | |

Figure 9

Editing and Searching for Annotations

A submitted annotation can be edited at a later time. You can retrieve a previously submitted annotation by going to the Community annotation home page, click on find annotation which will redirect you to the page illustrated below (Figure 10). There are two ways to retrieve annotations: either search by the NCU ID (locus) or by submitter's last name. In the drop down menu next to **Status**, select **All**. This will retrieve curated, uncurated, and deleted annotations. To retrieve the annotation, click on the annotation ID on the return page

BROAD INSTITUTE

Home > Data > Genomes > Find Annotations Logout hhood

Home | News | Info | Community | Download | BLAST | Find Features | Regions | Gene Index | Alleles | Maps | FAQ | Contact | Links

Find Annotations

Search Parameters

Locus:

Submitter's Last Name:

Status:

- Active
- Curated
- Deleted
- All

Home | Internal Web | Contact Us | Site Map | Search © 2006 Broad Institute

Figure10

Contact information

If you have any questions, feedback, or seek additional guidance, please contact Heather Hood. Usually, a quick phone call or e-mail is all that is needed.

Heather M. Hood, Ph.D.
 Department of Environmental and Biomolecular Systems
 OGI School of Science and Engineering
 Oregon Health & Science University
 20000 N.W. Walker Road
 Beaverton, OR 97006-8921
 503 748-2204 phone
 503 748-1464 fax
hmhood@ebs.ogi.edu