HYPOTHES.IS ANNOTATION OVERVIEW

This protocol is to assist individuals in learning how to use the web-assisted annotation tool, Hypothes.is. Hypothes.is uses a W3C standard for web annotation, in which the annotation acts as a "layer" over the article. In this way, additional information and context surrounding selected text can be "linked" to any web article, much like adding a sticky note to a book page, but with web annotations these notes are significantly more stable, transparent, directed (or linked) to specific text, and searchable.

By annotation, we are referring to the process of noting important information from scientific articles that is consistent with the evidence of interest to ClinGen. This includes phenotypes, genotyping methods, demographics, and most importantly a genetic variant. In addition, Hypothes.is allows these notes, or annotations, to be "tagged" to create searchable identifiers that can facilitate ClinGen expert panels, or yourself, in identifying the most relevant information for curation.

Why is this annotation important? Assessing the strength of evidence for a gene and/or variant requires review of all information for cases reported. Biomedical literature, while vast, has been unstructured. The lack of structure makes searching for variant of interests, for instance, extremely difficult and time consuming. Through the process of annotation, the data published will become structured and easier for ClinGen expert panels to search, locate, and assess the data. This will ultimately lead to increased data transparency, improved interpretations of variant pathogenicity which will lead to improved diagnositics.

Some Important Notes:
The ability to view another user's annotations in the context of the article (or the "linked" data) is dependent on one's personal access to the webpage. Thus, for scientific articles with restricted access, some users may only be able to see the annotated data on the Hypothes.is dashboard and not on the article itself.

Links to external sources will be outlined to guide curators to more useful training sites. An Appendix has also been provided with templates for each type of annotation report at the end of this overview. These templates can be copied and pasted into the annotation report to standardize and expedite curation. Specific annotation protocols will be provided to annotators and will contain the specific tags used and an example annotation.

If you have any questions, please feel free to email the ClinGen Community Curation Working Group at volunteer@clinicalgenome.org.

Getting Started:

1. Sign up for a hypothes.is account here: https://web.hypothes.is/start/
    a. This will include downloading the application to a web browser.
    b. The preferred web browser is Google Chrome. Firefox is also available.
    c. **Note:** The app does NOT currently work with Apple Safari
2. Join your gene specific annotation group
    a. This link is contained within your specific gene protocol

How to Annotate:

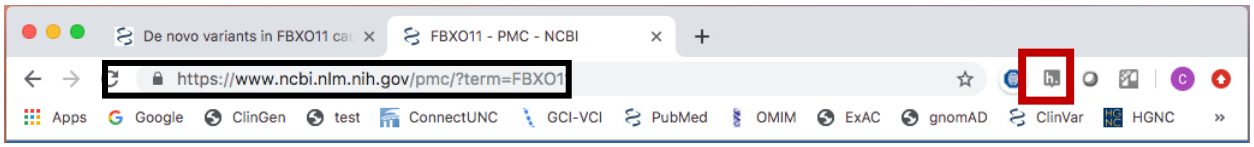
1. Choose an article of interest.
    a. Use of PubMed Central is strongly encouraged as it prioritizes returning open access, freely available articles.

i. As noted above, the ability of additional members to observe the “linked” annotation(s) is dependent on their personal access. However, in the absence of article access, any individual of a group can view the annotated data using the Hypothes.is dashboard.

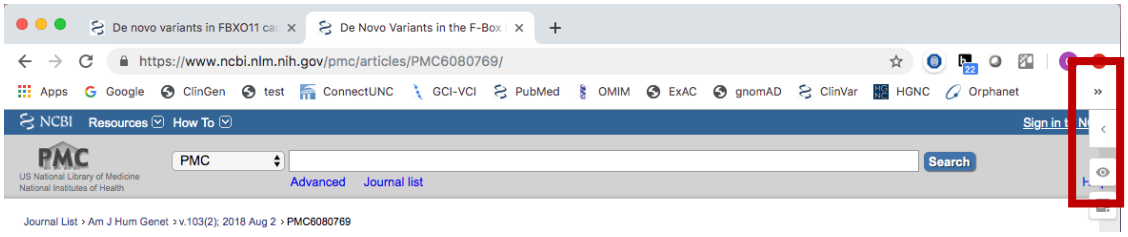
b. The preferred article version for annotation is HTML.
   i. Example for PMID:31638223
      1. For HTML choose “Article” link (red arrow below).
      2. The other possible links should NOT be used as they could alter the stability of the Hypothes.is annotations.

c. Annotations can be performed on PDF versions of articles, however the article PDF must have a stable URL to stabilize the text links (i.e. www.XXXX.com).
   i. PDFs that open in a new web browser window generally have a newly generated URL each time it is selected, and this can result in losing the annotation link to the text. However, the annotation(s) will still be visible in the hypothes.is dashboard.
   ii. To check for a stable URL, copy the URL (black box in figure below), then open a new browser window and paste in the copied URL. If the PDF article appears, it is stable. If the article does not, then the URL is NOT stable.

2. To activate the Hypothes.is extension, click the “h.” icon button in the web browser (red box in image below).



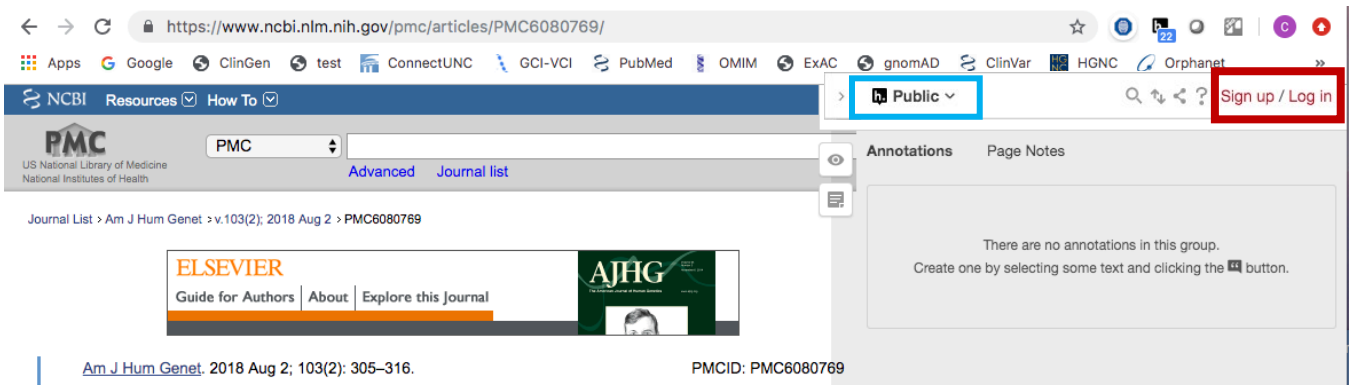
3. Once activated, an arrow in the right corner of the browser window will appear (see red box below).



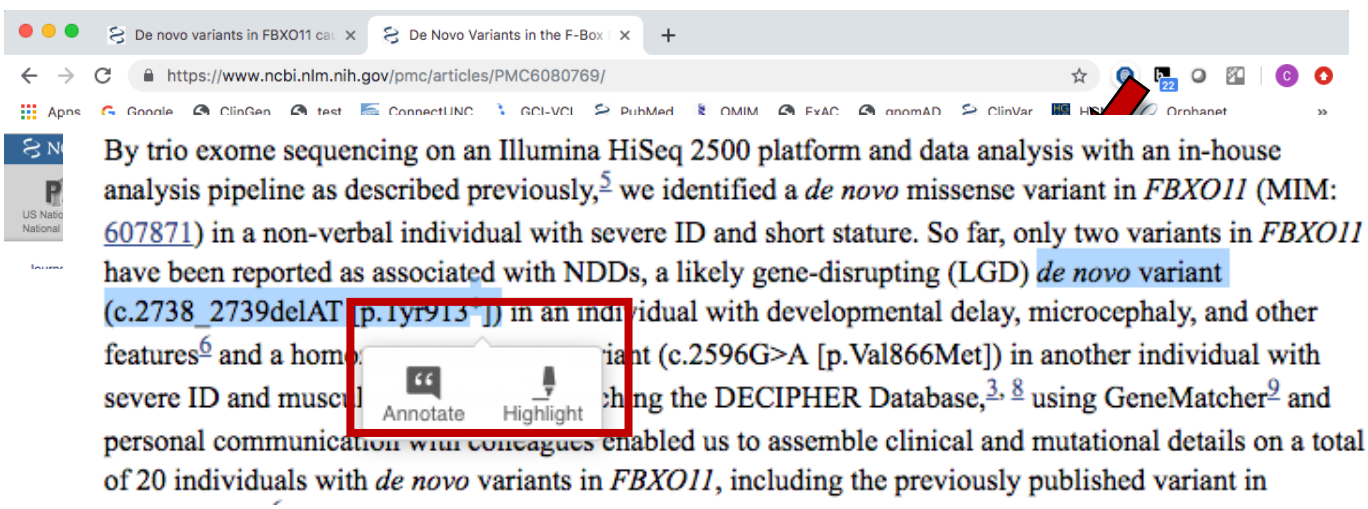
a. Alternatively, the entire annotation module may open (see image below, gray box).

4. Make sure to Log in (see red box above).

5. Choose the Annotation Group for posting the annotations.
   a. Use the \/ arrow next to the "Public" group in the figure above (blue box above).
   b. Choose the appropriate annotation group (see red arrow below, as an example).
      i. If a member of several groups, make sure to select the appropriate group based on

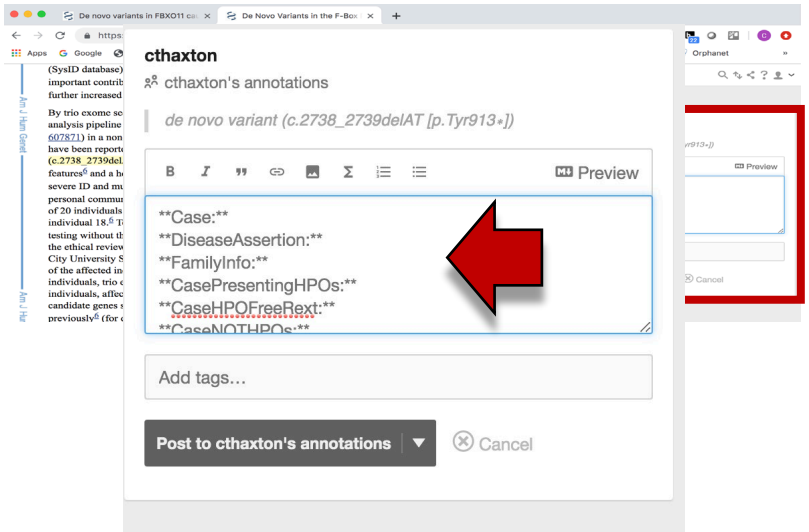


the gene of interest. The hyperlink to the group is located in the gene-specific tagging protocol.



6. After assessing that an article contains a case of interest (i.e. an individual asserted to have a variant in your gene of interest), use the cursor to select/highlight the text indicating the case (blue highlighted text below).
   a. An annotation module box should appear (see picture below, red box).
   b. To add curation information about the case, select "Annotate."
      i. A "link" between the text and the annotation has now been created!
   c. Choosing "Highlight" will simply highlight text, and no annotation layer will be formed.
      i. **Note:** Highlights can only be seen by the original user, not by the group.

7. Once “Annotate” is chosen an annotation box will pop up (see image below, red box). This is where specific data about the case can be typed, or annotated, within the text box, and tags can be added based on the gene specific protocol.

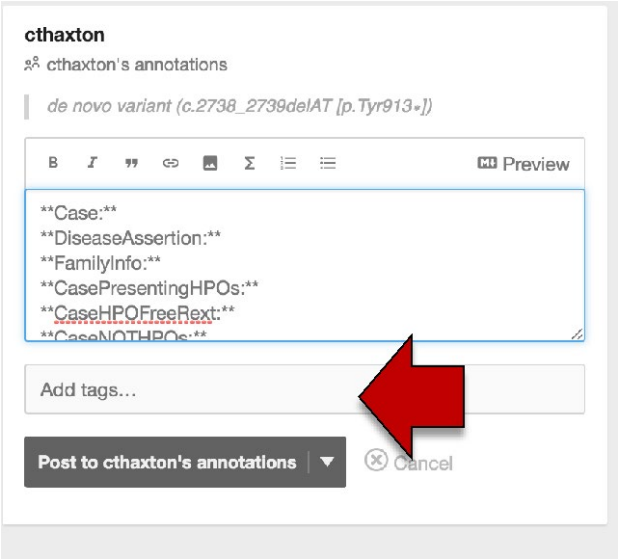


8. To assist curators, templates for each type of annotation report are included in the Appendix. Simply copy and paste them into the Annotation Text Box (see image below, red arrow).

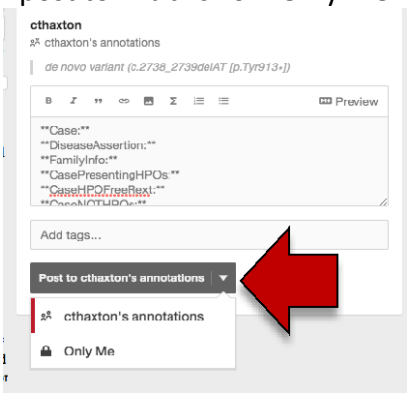
    a. A curator can now enter in the appropriate information based on what is published/asserted in the article for each evidence category (e.g. Case, DiseaseAssertion, etc).

9. To add “Tags” simply begin to type text into the “Add Tags” section, right above the post button (see image below, red arrow). Tags to use are outlined in the Gene Specific Annotation Protocol on Table 2.

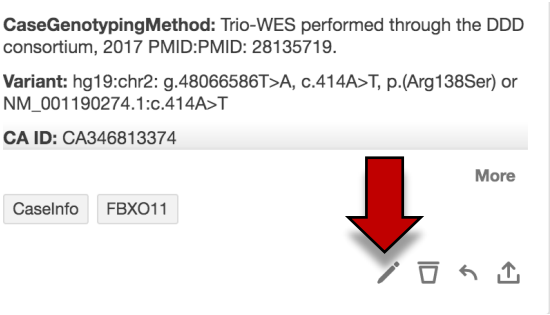
    a. **Note:** Commonly used tags will appear as selection candidates after they are used once. Use caution for tags that are similar, and double check that the correct tag has been chosen before posting an annotation. If an error has been made, just delete the tag.



10. Once all applicable information is added to the annotation, choose the “post to…” button (see image below, red arrow). This stabilizes the annotations and posts it to the group Hypothes.is board.
    a. **Note:** make sure when posting you are choosing the group annotation page that applies for this annotation. Do NOT post to “Public” or “Only Me”.



    b. If you have inadvertently posted to “Public” or “Only Me” you can edit the annotation by clicking the “Pencil icon” at the bottom of the annotation (see image below, red arrow).

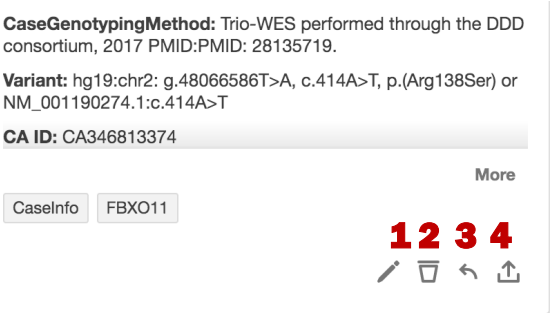


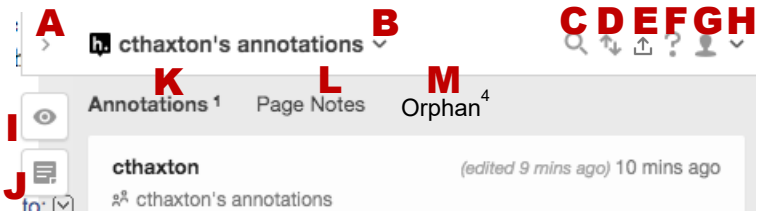
    c. Then choose the appropriate group posting (Step 10a above).

11. If for any reason there is a need to step away from an annotation, it is best practice to post the annotation and return to complete it. If not, then there is a high risk of losing the annotation and having to start again.

Annotation Icons:

The annotation module has 4 icons: (1) a pencil, to enable editing of an annotation; (2) a trash can, to delete an annotation; (3) an arrow, to enable a reply to an annotation; (4) a link icon, generates a hyperlink to enable sharing of a specific annotation.





<u>Other annotation icons are listed:</u>

A: Tab that can increase and decrease the size of the Annotation module on the webpage
B: Chosen Annotation Group
C: Find function, will search the tags and text within the annotations
D: Sort function for the annotations
E: Hyperlink button to enable sharing of the annotations in the article.
F: Help button to direct users to Hypothes.is webpage.
G and H: User and account information.
I: Show Highlights
J: Generate a Page Note

- This creates an unlinked annotation (i.e. it is not linked to text in the document).
    1. The tags used in a Page note are still searchable (e.g. the variant ID).
- Use this function when annotating data found in Supplemental materials, as many times this is an external file.

K: Shortcut to view Annotations. Metric of the # of Annotations on a webpage, indicated by the superscript.
L: Shortcut to view Page Notes. Metric of the # of Page Notes on a webpage, indicated by the superscript.
M: Orphan annotations

- Orphan Annotations represent former linked annotations, that have lost a textual link. These are unlike Page Notes which are created in the absence of a textual link from the beginning.
- Reasons for a lost textual link may be due to updates from the publisher to a webpage (e.g. corrected a sentence).

APPENDIX A: Annotation example and templates

Example Annotation:

See PubMed article PMID:28790153 for publicly available example.

MYH7 variant curation

AJV

CaseAJV: 17 years diagnosis, Australia

DiseaseAssertion: Hypertrophic Cardiomyopathy

FamilyInfo: Father (index case) died awaiting cardiac transplant (carried both variants). Two possibly affected relatives.

CasePresentingHPOs: HP:0001639, HP:0006536

(Hypertrophic cardiomyopathy, Obstructive lung disease)

HPOsFreeText: Maximum left ventricular hypertrophy at 17 mm, Sudden cardiac death event at 17 years, Maximal wall thickness at 22mm,

CaseNotHPOs: N/A

NotHPOsFreeText: N/A

CasePreviousTesting: See Table 1

CaseGenotypingMethod: DNA was isolated from peripheral blood. Most participants underwent testing from the Illumina Cardiomyopathy Sequencing Panel, which includes 46 cardiomyopathy related genes. For others, whole exome sequencing or Sanger squencing was used. After the results were returned, variants were filtered for pathogenicity and rarity.

Variant:NM_000257.3:c.1954A>G (p.Arg652Gly)

ClinVarID:177626
https://www.ncbi.nlm.nih.gov/clinvar/variation/177626/

gnomAD: Not in gnomAD

Multiple Gene Variants:

MYBPC3 Variant

Variant: NM_000256.3:c.2980C>T (p.Leu994Phe)

ClinVarID:180992
https://www.ncbi.nlm.nih.gov/clinvar/variation/180992/

gnomAD: European (Non-Finnish) 1.624e-4, Overall 8.461e-5
https://gnomad.broadinstitute.org/variant/11-47355487-G-A

Less

CaseInfo | Gene:MYH7 | HGNC:7577 | ClinVarID:177626 | FamilyInfo | Germline | InheritancePattern:NoInheritanceAssertion | DiseaseEntity:HCM

Annotation Text Box

Tag(s)

Templates

These strings can be copied and pasted to use as templates for filling in information in the annotations. Please see your specific gene protocol for the appropriate tags to include for each Annotation report (e.g. Case report). **Note:** These templates can be adjusted, and sections can be removed or added based on your specific gene protocol, however please do NOT alter how they are constructed. The CamelCase used is important to standardize the annotations.

Article Information template:

PMID:
Gene:
HGNCID:

Case-Individuals template:
Please see your specific gene protocol for a template that corresponds to their specifications.

Case:
DiseaseAssertion:
FamilyInfo:
CasePresentingHPOs:
CaseHPOFreeText:
CaseNOTHPOs:
CaseNOTHPOFreeText:
CasePreviousTesting:
GenotypingMethod:
SamplingMethod:
PreviouslyPublished:
SupplementalData:

Variant:
ClinVarID: **A curator only needs to include either a ClinVarID or CAID, not both.**
CAID:
gnomAD:
VariantEvidence: Only use if applicable

MultipleGeneVariants: **Only use when more than one genetic variation is noted for a case.**
GeneName:
Variant:
ClinVarID or CAID: **Choose one and annotate as above**
gnomAD:

If more than one genetic variant is noted for a case, use the same 4 categories listed under "MultipleGeneVariants:" (i.e. GeneName, Variant, etc) for each variant listed.

APPENDIX B: Annotation shortcuts

Within the annotation module, there are shortcuts to apply bold, italics, and underlined text. Instructions for use of github syntax can be found here.

Here are some commonly used shortcuts:

To bold text: Apply a two asterisks before and after the word or phrase of interest (e.g. **CaseInfo** will look like **CaseInfo** when posted)

To italicize text: Apply a single asterisk before and after the word or phrase of interest (e.g. *CaseInfo* will look like *CaseInfo* when posted)

To bold and italicize text: Apply three asterisks before and after the word or phrase of interest (e.g. ***CaseInfo*** will look like ***CaseInfo*** when posted).

To italicize a single word in an entirely bolded phrase: Apply a single underscore (_) around the word to italicize, and apply two asterisk before and after the intended phrase for bolding. (e.g. **Variation in _MYH7_ is associated with cardiomyopathy** will look like **Variation in *MYH7* is associated with cardiomyopathy**).

Hypothes.is does NOT have a shortcut to underline at this time.

Use of the # symbol can categorize text as a heading and will result in larger bolded text. The greater the number of # signs used, the smaller the heading. Therefore, a single # denotes the largest heading.

APPENDIX C: Useful Resource links

This appendix contains the links to helpful resources used for annotation and curation. The group specific annotation protocols also contain many of these links.

- PubMed Central
  - https://www.ncbi.nlm.nih.gov/pmc/
  - Preferred search engine to locate articles to annotate
- ClinVar
  - https://www.ncbi.nlm.nih.gov/clinvar/
  - Variant IDs
- ClinGen Allele Registry
  - http://reg.clinicalgenome.org/redmine/projects/registry/genboree_registry/landing
  - Variant IDs
- Human Phenotype Ontology (HPO)
  - https://hpo.jax.org/app/
  - Phenotype IDs (HPOs)
- gnomAD
  - https://gnomad.broadinstitute.org/
  - Population frequencies
- Hugo Gene Nomenclature Committee (HGNC)
  - https://www.genenames.org/
  - Gene IDs
- Online Mendelian Inheritance in Man (OMIM)
  - https://www.omim.org/
  - Useful to get information on the gene(s) of interest for annotation
-