ClinGen General Sequence Variant Curation Process

Standard Operating Procedure

Version 3.1

April 2022

(Note: Yellow highlights indicate updates from V2.0 to V3.0; green highlight indicates update from V3.0 to V3.1; blue highlight indicates update from V3.1 to V3.2)

The Clinical Genome Resource

ClinGen Variant Curation SOP Committee

Contents

1. INTRODUCTION 1

1.1. USING THIS PROTOCOL 2

2. GETTING STARTED 3

2.1. Logging in and affiliations 3

2.2. Summary of the variant curation process 3

3. GENERAL NAVIGATION 4

3.1 Landing Page 4

3.2 Dashboard View 4

4. GENERAL ORGANIZATION OF EVIDENCE IN THE VCI 5

4.1. Evidence View 5

4.2. Interpretation View 5

4.3. Evidence Tabs 6

4.4. Additional information 6

5. SELECTING A VARIANT FOR CURATION IN THE VCI 6

5.1. Starting a new variant curation 6

5.2 Finding an existing variant curation record 7

5.3 Getting to the Interpretation View 7

6. ASSOCIATING A VARIANT WITH A DISEASE AND MODE OF INHERITANCE 9

6.1 Associating a variant with a disease using a Mondo ID 9

6.2. Free text option for associating a variant with a disease 10

6.3. Adding a mode of inheritance 11

6.4. Linking to a specification document 11

7. FINDING EVIDENCE 12

7.1 How to do a literature search 12

7.2. Curated Literature Evidence 15

8. DOCUMENTING EVALUATION OF EVIDENCE 16

8.1. Criteria Evaluation Labels 16

8.2. Modified Strength of ACMG/AMP Criteria 16

8.3. Summary of steps for criteria evaluation 17

9. EVIDENCE TABS 18

9.1. BASIC INFORMATION TAB 18

9.2. POPULATION TAB 22
9.3. VARIANT TYPE TAB 25
Missense sub-tab 26
Loss of Function (LOF) sub-tab 34
Silent & Intron sub-tab 37
In-frame Indel sub-tab 37
9.4. EXPERIMENTAL TAB 38
9.5. CASE/SEGREGATION TAB 41
9.6. GENE-CENTRIC TAB 54
10. EVALUATION SUMMARY 54
11. CLINVAR SUBMISSION 55
References 56
Resources 58
Appendix I - Computational Tools 59
Appendix II - PVS1 flowchart, SVI recommendations 60
Appendix III - Combining Criteria to Classify Sequence Variants 61

FREQUENTLY USED TERMS

Affiliation: The term that the Variant Curation Interface (VCI) uses for a group of people represented in ClinGen resources (e.g. ClinGen Working Groups, ClinGen Expert Panels, Research Labs, Clinical Labs, etc.) that collectively edit and score/evaluate evidence, and work on and approve classifications of the same variants.

Criterion: A specific rule, based on evidence type and described in the ACMG/AMP variant interpretation guidelines, used for classification of the clinical significance of a variant. Each criterion has an "evidence code" in the ACMG/AMP guidelines (e.g. PVS1), and the two terms are used interchangeably throughout this document.

Variant Classification: The process of aggregating evidence and assessing a variant's clinical significance. Also, a term used to describe the clinical significance of a variant based on assessment using the ACMG/AMP guidelines; Pathogenic, Likely pathogenic, Uncertain Significance, Likely benign, Benign.

Variant Curation Expert Panel (VCEP): A group of experts and biocurators tasked with providing specifications to the ACMG/AMP guidelines for a gene or group of related genes, classifying variants according to these specifications, and publishing the interpretations on ClinGen's Evidence Repository and NCBI's ClinVar database, both of which are publicly available.

Variant Curation Interface (VCI): An publicly available, online platform, developed by ClinGen, which allows users to comprehensively identify, annotate, and share relevant evidence, and to apply evidence codes for variant classification based on the ACMG/AMP variant classification guidelines.

1. INTRODUCTION

The ClinGen General Sequence Variant Curation standard operating procedure (SOP) is designed to provide guidance on variant classification using ClinGen approved processes and tools. Standardized assertion criteria to classify clinical sequence variants associated with Mendelian disorders into a five-tier nomenclature system were developed in "Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology"[1] (ACMG/AMP) and include pathogenic (P), likely pathogenic (LP), uncertain significance (VUS), likely benign (LB), and benign (B). This SOP is intended to guide variant curators through the process of curating and scoring evidence for the ACMG/AMP assertion criteria (population data, computational and predictive analysis, functional criteria, and allelic and co-segregation data) using the ClinGen Variant Curation Interface (VCI). Information from publicly available resources and internal laboratory data is curated and scored with respect to a variant-disease relationship. A classification for each variant is assigned based on assigning ACMG/AMP evidence codes and the appropriate strength of the evidence in the assertion categories. This SOP provides a general overview of the best practices to follow for curating a variant-disease relationship and subsequently assigning a classification including general recommendations from the ClinGen Sequence Variant Interpretation Working Group (SVI), which can be found on the SVI webpage of the ClinGen website (https://www.clinicalgenome.org/working-groups/sequence-variant-interpretation/). This SOP serves as a general guide to the process of variant classification and is intended to complement disease-specific Variant Curation Expert Panel (VCEP) specifications when available. Please note, if applicable disease or gene-specific specifications exist, those should be used in conjunction with this SOP. Please check the ClinGen SVI webpage for a listing of all available VCEP specifications, as well as with the VCEP with which you may be working, for the most up-to-date disease-specific guideline specifications that may be in progress.

1.1. USING THIS PROTOCOL

When using this protocol, please also refer to the Variant Curation Interface Help Document for further details (https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help)

This protocol is designed to be used in conjunction with various publicly available tools and databases, including:

- The ClinGen Variant Curation Interface (VCI)
    - https://curation.clinicalgenome.org/
    - For instructions on VCI registration and login, see the VCI Help Document, https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help
- Standardized Text for ClinGen Variant Curation Expert Panels
    - This document provides standardized summary text for each ACMG/AMP code to facilitate data entry in the VCI. Additional information and the document can be accessed here: https://clinicalgenome.org/docs/standardized-text-for-clingen-variant-curation-expert-panels/
- ClinVar
    - https://www.ncbi.nlm.nih.gov/clinvar/
    - No registration or login is required for ClinVar.
- The ClinGen Allele Registry
    - http://reg.clinicalgenome.org/redmine/projects/registry/genboree_registry/landing
    - To create a login for the ClinGen Allele Registry, please use the following site: http://reg.clinicalgenome.org/redmine/projects/clingen_users/register_clingen_user/cg_users/new
- Mondo
    - To find a Mondo code for a disease entity, go to Mondo:
        - https://www.ebi.ac.uk/ols/ontologies/mondo

When familiarizing yourself with this protocol, you may want to log in to the demo version of the VCI (https://curation-test.clinicalgenome.org/) so that you can follow along with the descriptions and screenshots provided.

2. GETTING STARTED

2.1. Logging in and affiliations

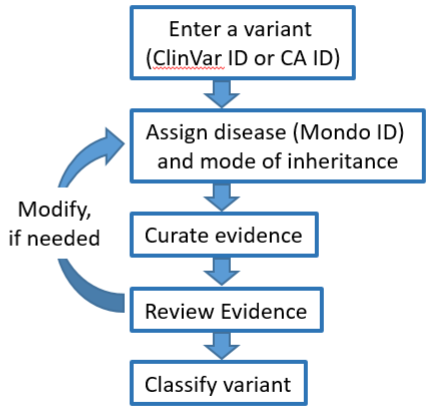
Once you have received a VCI login, you may begin variant classification in the VCI. See the VCI Help Document (https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help) for information on how to obtain a login.

Curators who are working as part of a ClinGen VCEP must select the appropriate affiliation after logging in. For more information on affiliations and how to join an affiliation, please refer to the VCI Help Document (https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help).

2.2. Summary of the variant curation process

A summary of the variant curation process is shown in Figure 1. Each of these steps in explained in detail below.

Figure 1. Variant Curation process in the VCI



3. GENERAL NAVIGATION

3.1 Landing Page

When you first login you will be on the landing page, which contains general information about ClinGen and the VCI and GCI. To access your curations head to the dashboard.

3.2 Dashboard View

Please refer to Figure 2.

Navigation Bar
Available from all pages, this is where you initiate variant (A) and gene curations (B), navigate to the dashboard (C), access the help documents (D) and sign out (E).

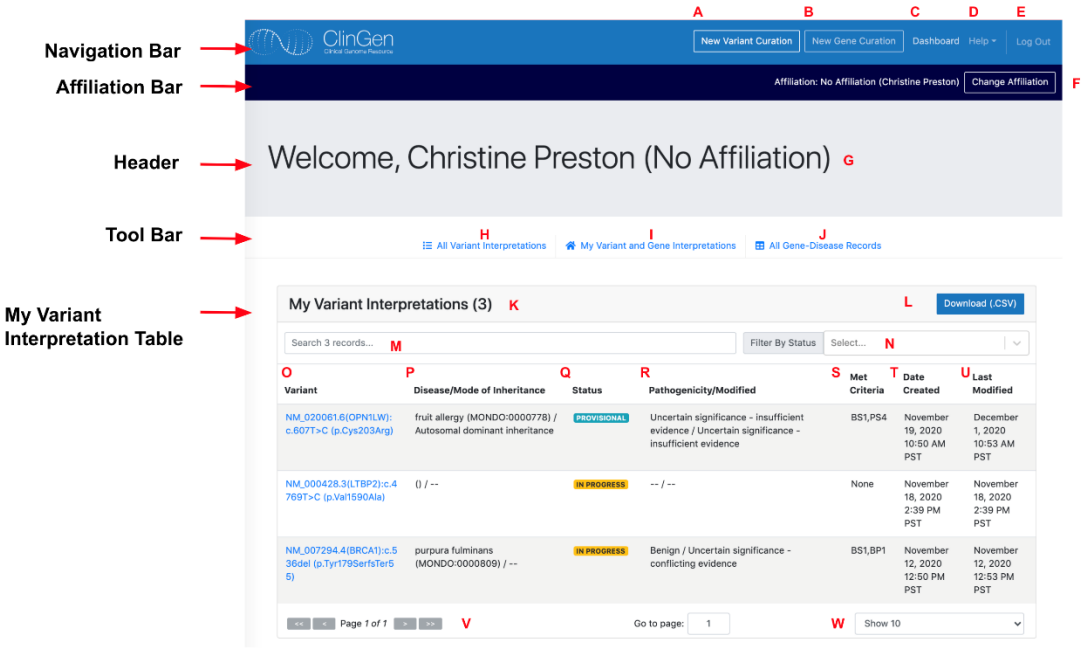
Affiliation Bar
Available from all pages, this is where you can change the affiliation you are curating under (F).

Header
Indicates what account and affiliation you are logged in under.

Tool Bar
● All Variant Interpretations (H) – This list contains all the classifications curated to date, along with their status, creator, date created and date last edited.
● My Variant and Gene Interpretations (I)
● All Gene-Disease Records (J) – This list contains all the Gene-Disease Records curated to date, along with their status, creator, date created and date last edited.

My Variant Interpretations Table
● This table indicates all VCI entries for the user or logged in affiliation. The total number of curations is listed at the top (K).
● The table (or filtered table) can be downloaded as a .csv format (L)
● The table can be filtered on any text in any text field (M), and on the status (N).
● The table displays the variant title (O), the associated disease and mode of inheritance (P), Status (Q), the calculated and modified pathogenicity (R), all criteria which are met in the variant's curation record (S), the date created (T) and the date last modified (U).
● There are pagination features (V) and users can expand the number of curations displayed at once (W).

Figure 2. Dashboard view.



4. GENERAL ORGANIZATION OF EVIDENCE IN THE VCI

4.1. Evidence View

The Evidence View displays evidence from external sources, such as databases (e.g. ExAC, gnomAD, PAGE, 1000 Genomes, ESP, ClinVar), and predictive algorithms (e.g. REVEL, Polyphen, dbNSFP). In addition, the VCI accumulates evidence that a curator enters when in the Interpretation View.

4.2. Interpretation View

The Interpretation View allows you to enter data and record your evaluation of the evidence according to the ACMG/AMP criteria (evidence codes)[1]. Your assessment will be editable by all members of the affiliation under which you are working. Any criteria applied will not be viewable to other affiliations until the variant classification has been approved by your VCEP affiliation. However, any data entered from publications (PMIDs) will be saved for any VCI user to view immediately.

4.3. Evidence Tabs

In both the Evidence View and the Interpretation View, information on the selected variant is organized into 6 tabs based on the ACMG criteria: Basic Information, Population, Variant Type, Experimental, Case/Segregation, and Gene-centric. The Variant Type tab contains subtabs (Missense, Loss of Function, Silent & Intron, In-frame indel) that allow the curator to look at the appropriate evidence, such as in silico predictors, and evaluate the appropriate criteria according to the variant type. Click between the tabs and subtabs to view the different types of information and criteria available (please see the following sections on each tab for further details).

4.4. Additional information

Throughout the VCI, you will find information buttons, , which will provide useful details when you hover over them.

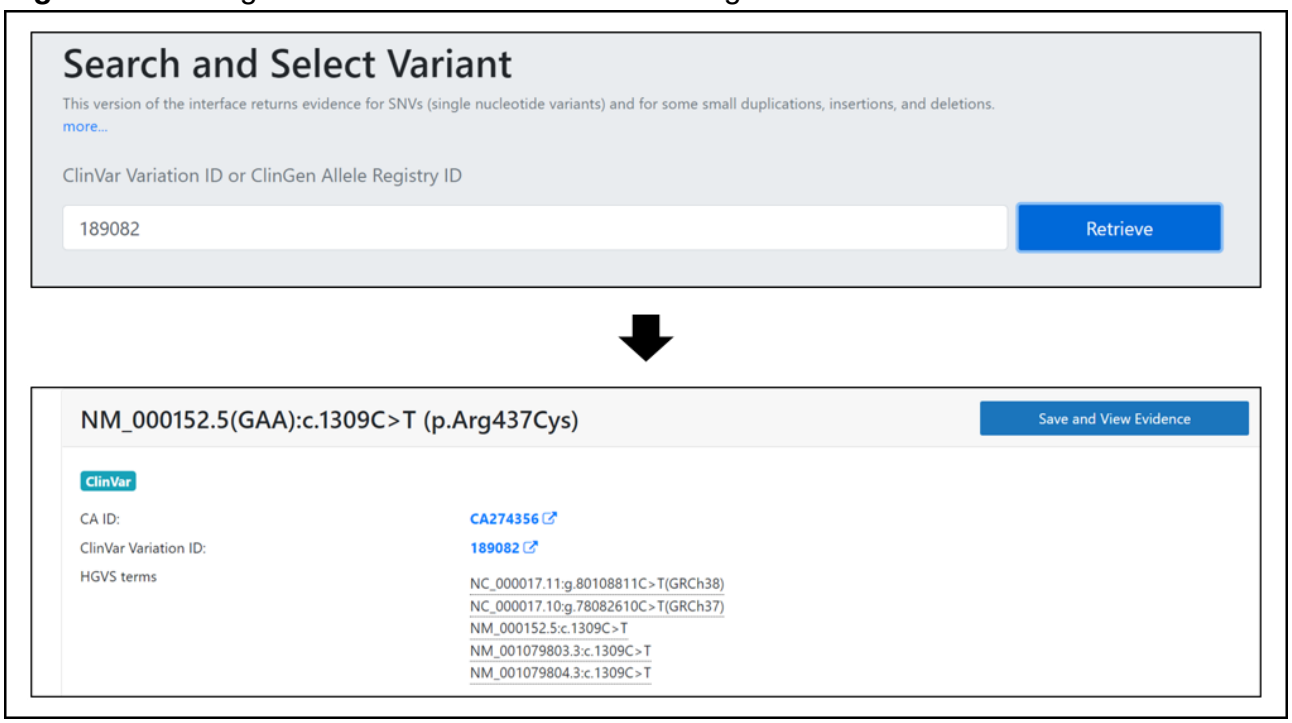
5. SELECTING A VARIANT FOR CURATION IN THE VCI

5.1. Starting a new variant curation

To start a record on a new variant, click “New Variant Curation” (in the blue banner at the top of the page). Enter the ClinVar ID, if available. If the variant is not in ClinVar, search the ClinGen Allele Registry (CAR) to obtain the ClinGen Allele Registry ID (CA ID). The CAR can be searched using various search terms including the HGVS nomenclature for the variant, ClinVar ID, and with partial information. You may have to register the variant in the CAR in order to obtain a CA ID. Once you have the variant, click “retrieve”. Check that the retrieved variant is correct, then click “Save and View Evidence” (Figure 3). You will now be in the “Evidence View” for the selected variant (see next section for details on the Evidence View).

When a new record is created in the VCI, the HGVS nomenclature is created by ClinVar or the CAR. Both systems use the MANE Select transcript (Matched Annotation from NCBI and EMBL-EBI). If no MANE Select transcript is available for the CAR, the HGVS nomenclature is based on the canonical transcript i.e. the transcript with the longest translation with no stop codons or, if there is no translation, the longest non-protein-coding transcript. If a single canonical transcript is not discernible, the HGVS nomenclature is based on the GRCh38 genomic coordinates. For ClinVar, if no MANE Select transcript is available, the transcript is based on the transcript selected by the first submitter.

Figure 3. Starting a new variant curation – selecting the variant



5.2 Finding an existing variant curation record

You may use the dashboard to search to find a variant record that already exists in the VCI (See Figure 2). The dashboard table for the affiliation under which you are working can be filtered on any text in any text field (Figure 2, M), and on the status (Figure 2, N). Note that any time that you are working on a record, you can return to the dashboard via the dashboard link on the top right.

5.3 Getting to the Interpretation View

To get to the Interpretation View from the Evidence view, click on “Interpretation +” (Figure 4). You will be presented with a submission policy agreement (Figure 4). ***Please read the submission policy statement carefully and ensure that you are familiar with the definition of Protected Health Information (PHI) before proceeding; PHI must not be entered into the VCI.*** Clicking “agree” takes you to the Interpretation View (Figure 5).

Figure 4. Beginning an interpretation in the VCI

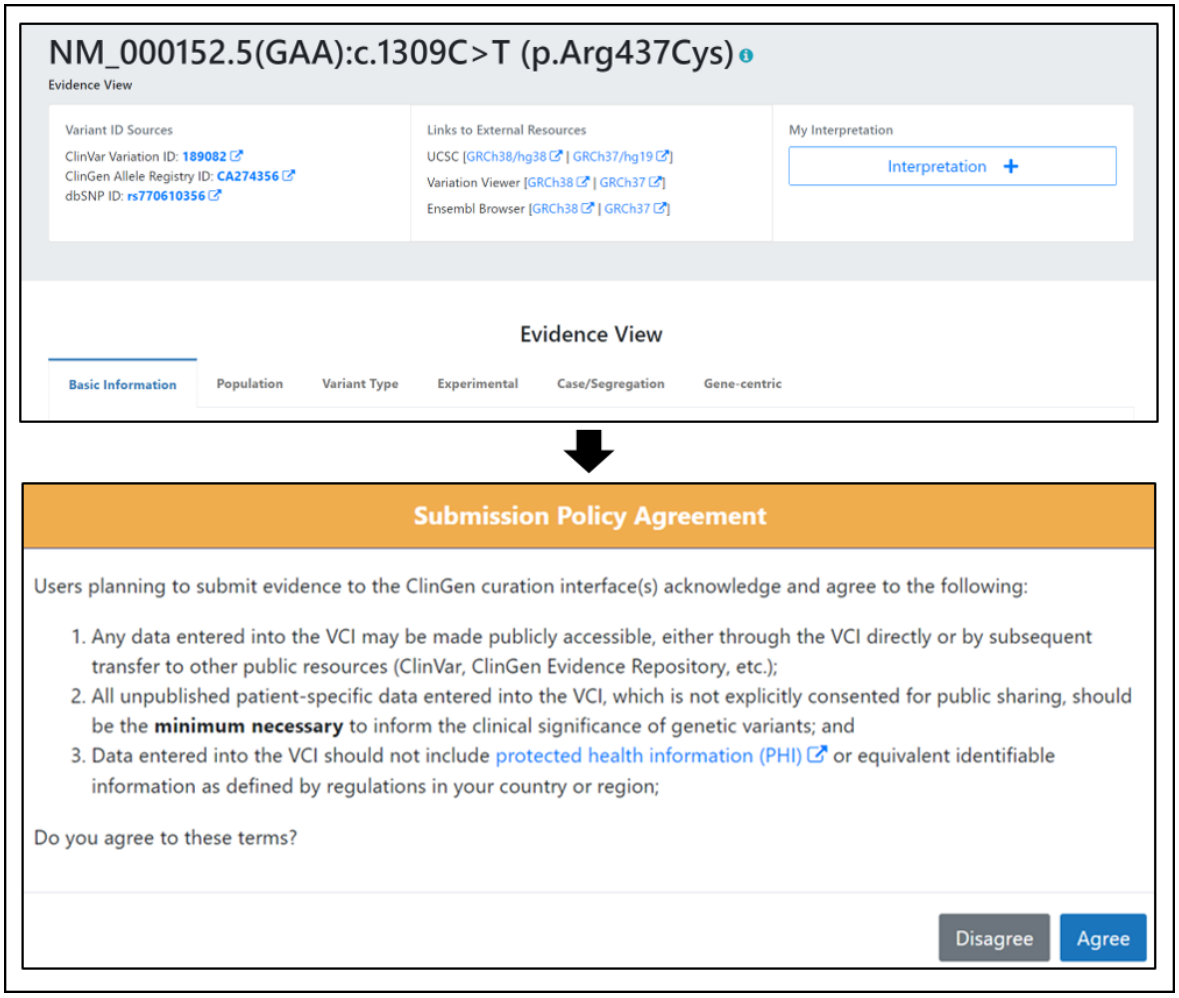
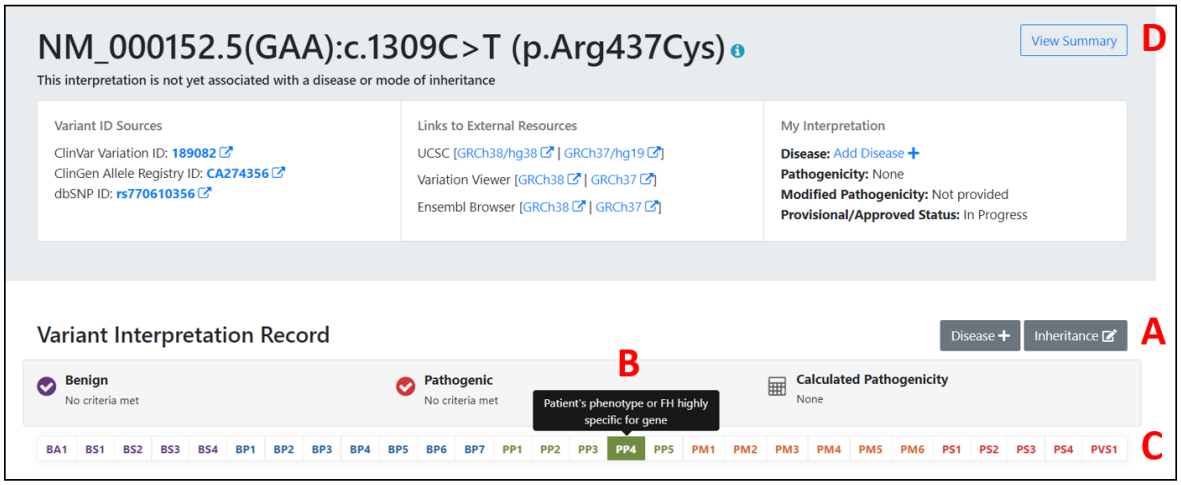


Figure 5. The Interpretation View in the VCI



Once you are in the Interpretation View, the following will appear:

A. “Disease +” and “Inheritance +” buttons for associating a Disease and Mode of Inheritance with the variant.
B. Interpretation Progress bar that indicates the number of benign and pathogenic criteria met and the calculated pathogenicity.
C. Criteria bar - Scroll over individual criteria codes to see a description for each criterion, and click on individual criteria codes to access those fields in the VCI. Criteria are colored if an individual criterion is “Met” and grey if “not met” (see next section, “Evaluating Criteria”)
D. “View Summary” button to view a Summary of all the evaluations and a free text field to enter a written summary.

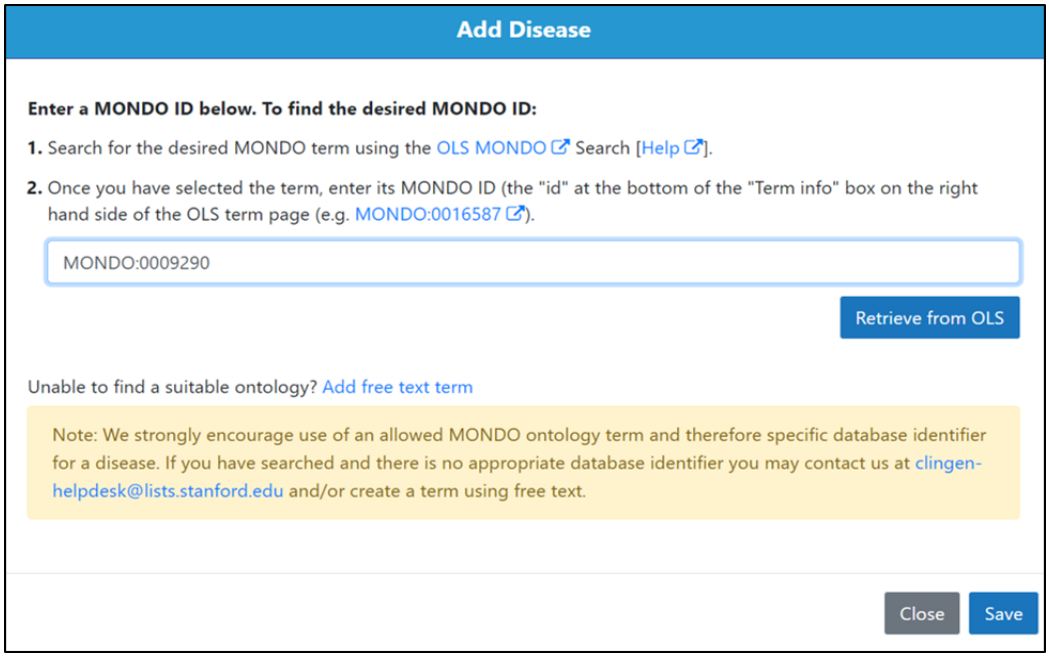
6. ASSOCIATING A VARIANT WITH A DISEASE AND MODE OF INHERITANCE

It is possible to classify a variant without entering a disease and/or inheritance pattern because sometimes the best disease term is not known until the evidence is gathered. However, a disease and mode of inheritance must be associated with all approved variant classifications from ClinGen VCEPs in order to satisfy the FDA’s genetic database requirements. Therefore, a disease and mode of inheritance upon which the classification is based should be associated with a variant before approving the classification. Of note, if necessary, it is possible to change the disease and mode of inheritance at any point before the classification is approved (see Figure 1).

6.1 Associating a variant with a disease using a Mondo ID

When working with an expert panel, the curator should discuss which disease terms are appropriate for the gene and should be based upon those diseases that are chosen during gene-disease validity curation. To associate a variant with a disease, click either the “Disease +” button or the “Add Disease +” link (shown in Figure 5). An “Add Disease” box will pop up (Figure 6). Enter the desired Mondo ID.

Figure 6. Associating a variant with a disease using a Mondo ID.



After entering the Mondo ID and selecting "Retrieve from OLS," the term name and description of the disease (if one exists) will be returned. Select "Save" if this is the desired term. Now you will see the disease term in the "My Interpretation" section under the variant title at the top of the page, and in the "All interpretations for this variant in the Variant Curation Interface (VCI)" section.

Use of a Mondo disease identifier is highly recommended. You can find a link for searching Mondo, and a link for help on searching Mondo, in the "Add Disease" box. If you cannot find an appropriate Mondo ID, please feel free to contact us at clingen-helpdesk@lists.stanford.edu and we will be happy to assist.

6.2. Free text option for associating a variant with a disease

If there is no appropriate Mondo term for the disease of interest, a free text term may be entered as a last resort. Click on "Add free text term" and enter a term (up to 100 characters in length). You must also provide either a set of HPO terms (preferred), a definition for the term you are entering, or both. Please remember that if someone else enters a different phrase for the same disease, the interface will not be able to determine that they are the same disease. Therefore it is important to consult with your VCEP to ensure that curators are using the same term, and that a plan is in place to request an appropriate Mondo ID for the disease. For guidance on naming disease entities, and Mondo IDs, please consult the ClinGen Lumping and Splitting Working Group.

6.3. Adding a mode of inheritance

You can add the mode of inheritance by clicking the “Inheritance” button (show in Figure 5), and selecting the appropriate mode of inheritance from the drop down menu. You also have the option to “select an adjective”. The list of adjectives displayed will depend on the selected mode of inheritance.

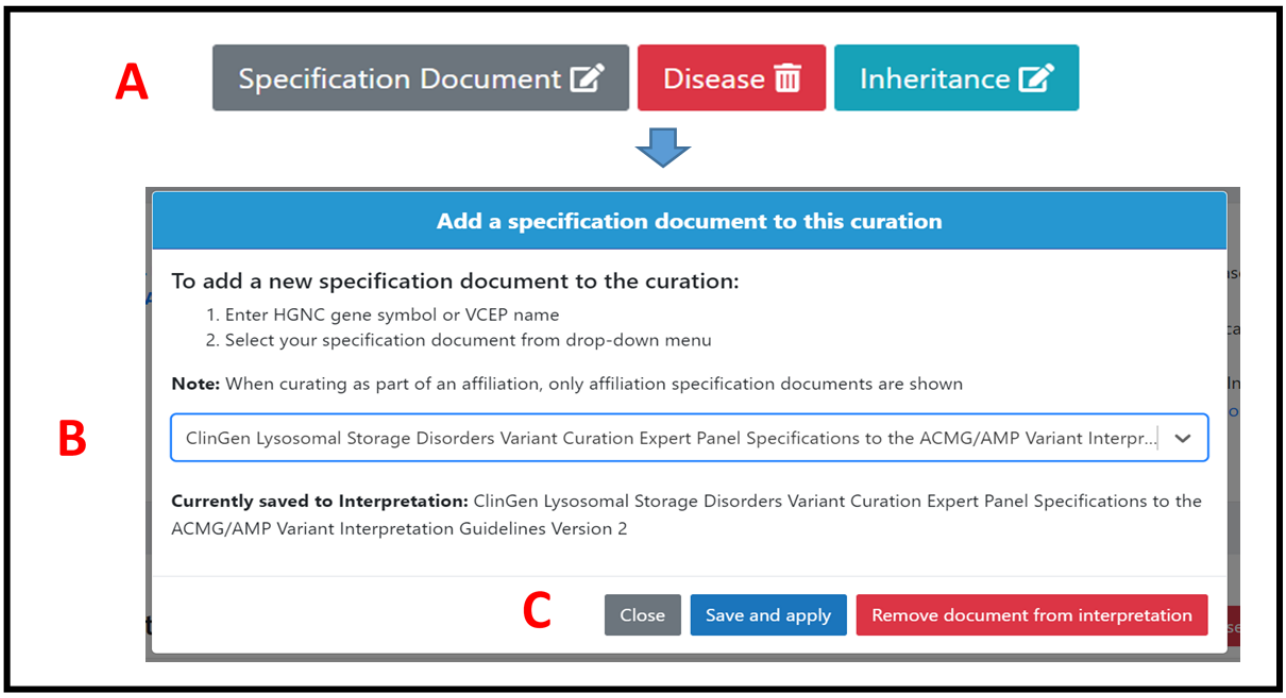
Now that you have started a new variant record (or found a record that you want to edit), and have added a disease and mode of inheritance, you are ready to begin adding data. Please note that the disease and mode of inheritance can be changed, if needed, until the variant classification has been approved.

6.4. Linking to a specification document

Curators have the option to link to the specification document that was used to classify the variant. In order to do this, click on the “Specification document” button located to the left of the “Disease” button, and then click the dropdown to view the documents (Fig 7, A, B).

- Curators who are logged in under an approved VCEP affiliation will be shown documents specific to that affiliation. Select the appropriate document, and then hit “Save and Apply” (Fig 7. C).
- Curators who are logged in under a VCEP affiliation that is not yet approved, and those who are logged in as an individual, will have the ability to search all specification documents in the Criteria Specification Registry (https://cspec.genome.network/cspec/ui/svi/)

Figure 7. Linking to a specification document



7. FINDING EVIDENCE

7.1 How to do a literature search

An important first step for any variant curation is a comprehensive literature search. This will identify articles that allow the curator to apply ACMG criteria during the curation, including data on functional studies, case reports (de novo, segregation, phenotype, co-occurrence), molecular characterization, case-control data etc. Please keep in mind that the evidence collected may be both for and against pathogenicity. In order to perform a literature search, the biocurator must understand variant nomenclature, and know which resources are most useful in finding appropriate information.

Comprehensive literature searches will use a combination of search tools and databases. Examples of such resources that are publicly available, or provide a version that is free of charge, can be found in Table 1. Each has different strengths. For example, PubMed, while useful for scientific literature, will only identify a paper where the variant is contained in the title or abstract. Google has the capability to identify variants found within the text of an article, as well as in supplemental tables; however, it will also find non-genetics related links. Google Scholar is limited to published academic literature, and is often better than Google at finding variants within a paper and within tables.

Table 1. Examples of publicly available databases and search tools.

Example Databases
Human Gene Mutation Database* (HGMD, http://www.hgmd.cf.ac.uk/ac/index.php)
Gene-specific databases e.g. Leiden Open Variation Database (LOVD, https://www.lovd.nl/)
ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/)
ClinGen Allele Registry (http://reg.clinicalgenome.org/redmine/projects/registry/genboree_registry/landing)
Example Search tools
PubMed (https://pubmed.ncbi.nlm.nih.gov/)
Google
Google Scholar (https://scholar.google.com/)
LitVar** (https://www.ncbi.nlm.nih.gov/CBBresearch/Lu/Demo/LitVar/#!?query=)
Mastermind*** (https://mastermind.genomenon.com/)

*HGMD is a commercial database that offers a less up-to-date version freely available to registered users from academic institutions or non-profit organizations.
**LitVar is a web-based tool for searching and retrieving variant-relevant information from the biomedical literature.
*** Mastermind is a commercial search engine that offers a basic plan free-of-charge.

To perform a comprehensive search, it is important to build an appropriate search term. Begin with the gene name and several versions of the variant nomenclature. A variant may have multiple identifiers for several reasons – strict Human Genome Variation Society (HGVS) format is not always followed, the Genome Reference Consortium human (GRCh) build has been updated (this is done periodically), alternate or historical gene names are used, an alternate transcript or non-standard nomenclature is used, etc. Some alternate transcripts and nomenclature can be found under the 'Basic Information' tab in the VCI and in the CAR entry for the variant; this is discussed in more detail under this section in the SOP. Use of quotation marks, parentheses, and the AND/OR Boolean operators assist with collection of a broad, but not overwhelming, output. Examples of search strings that can be used for specific types of variants can be found in Table 2.

Search output can range from few (or zero) to hundreds of articles. The list will most likely include multiple links from the same source, including those that may have already been viewed, such as ClinVar or gnomAD, so the original output list can be reduced. While all

links should be explored, not all will contain relevant data. For example, an article may mention the gene of interest, and also the variant of interest but found in a different gene, or may describe the variant of interest but the paper may discuss downstream care of patients with the disease.

Table 2. Examples of search strings for different types of variants.

Variant Type	Lit search string 'HGNC gene symbol' AND ("variant codon nomenclature" OR "variant protein nomenclature" OR "variant GRCh location")
Nonsense	PTEN AND ("1003C>T" OR "1003 C>T" OR "Arg335Ter" OR "R335X" OR "Arg335STOP" OR "89720852")
Missense	NF1 AND ("277T>C" OR "277 T>C" OR "Cys93Arg" OR "C93R" OR "29486100")
Frameshift	MYBPC3 AND ("1028delC" OR "1028del" OR "Thr343MetfsX7" OR "T343MfsX7" OR "Thr343Metfs*" OR "T343Mfs*" OR "47367820")
Intronic	FBN1 AND ("1148-2A>C" OR "1148-2 A>C" OR "IVS10-2A>C" OR "IVS10-2 A>C" OR "48808561")
In-frame indel	MSH6 AND ("2157_2159delTAC" OR "2157delTAC" OR "2157del3" OR "Thr720del" OR "T720del" OR "48027279")
Different amino acid change at same residue	Variant of interest: TP53 His179Asn TP53 AND ("His179*" OR "H179*") – wouldn't include "NOT His179Asn" because one article might discuss both variants.

When reviewing variant specific literature, it is important to avoid double counting of probands. The same patient or family may be presented in multiple publications, and while this is not always readily apparent, there are certain clues to suggest it. Things to look for include author overlap between articles, use of the same study cohort, highly aligned specific clinical details, etc. If unsure, it is best to contact the authors and/or discuss inclusion of this data with the Expert Panel (if applicable).

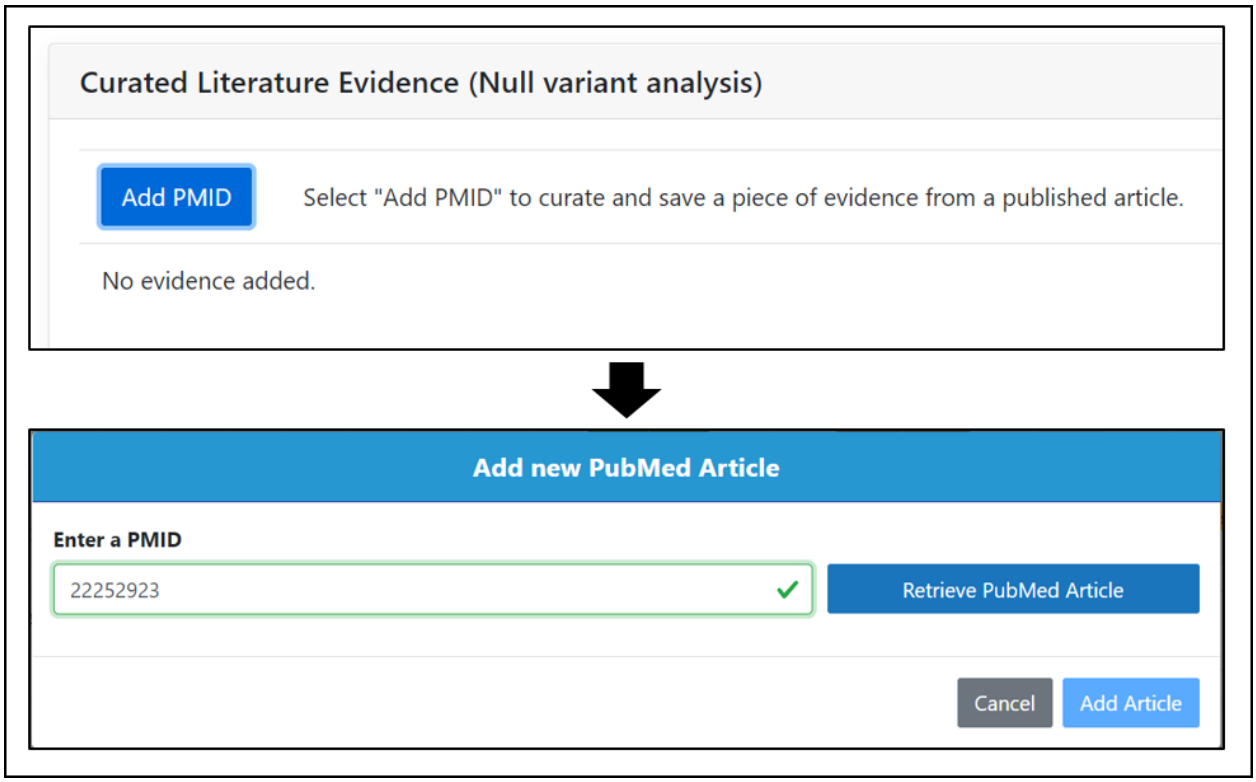
Once it has been verified that the specific variant of interest is included in the publications, this literature can be collected and reviewed, and the data can be entered into the appropriate text boxes, with PMID, in the VCI as described below.

7.2. Curated Literature Evidence

The Population, Variant Type, and Experimental tabs each contain a subsection entitled “Curated Literature Evidence”. In this section, there is an opportunity to add PMIDs to document published data that facilitate the criteria evaluation. Clicking on the “Add PMID” button provides a pop up box (Figure 8) where a PMID may be entered. The VCI will automatically retrieve the details from PubMed. Double check that the correct article has been retrieved and add the article. Once the article is added, a free text ‘Evidence’ box will appear where the curator can document and save relevant data from the article.

In the Case/Segregation tab, in addition to published literature (PMIDs), VCEPs may choose to use evidence from other sources, such as a Clinical Laboratory, Clinic, Research Laboratory, Database, or Other. See the Case/Segregation tab for further details.

Figure 8. Adding evidence from a PMID



8. DOCUMENTING EVALUATION OF EVIDENCE

8.1. Criteria Evaluation Labels

For each criterion, there are at least three evaluation label choices: Met, Not Met, and Not Evaluated. It is important to understand how these labels are used, and how the information considered when applying each label is captured and published in the final classification. To that end, the following definitions will be used to define each label:

Met: This label will be used if evidence meets specified rules for a given criterion. Curators are encouraged to add an explanation for usage that begins with the criterion code in the free text explanation box next to each criterion.
Please see the "Standardized text for ClinGen VCEPs" document (https://clinicalgenome.org/docs/standardized-text-for-clingen-variant-curation-expert-panels/) for guidance on language for text boxes in the VCI. The curator should include all relevant PMIDs and data, as outlined in the standardized text document. All explanation notes and PMIDs (with corresponding PMID notes) will be captured and published to the ClinGen Evidence Repository (ERepo) (https://erepo.clinicalgenome.org/evrepo/) in association with Met codes.

Not Met: This label is to be used if evidence is evaluated and determined not to meet the criterion. Again, curators are encouraged to add an explanation for not applying the code beginning with the criterion code and including relevant PMIDs, as outlined in the "Standardized text for ClinGen VCEPs" document (https://clinicalgenome.org/docs/standardized-text-for-clingen-variant-curation-expert-panels/). All explanations will be captured and published to the ClinGen ERepo.

Not Evaluated: This label is the default label and is currently intended to be used if there is no evidence to evaluate OR if the criterion is not applicable for the variant. It is helpful if the curator documents whether the label was actively chosen as opposed to the code not yet being considered; however, Not Evaluated codes are currently not captured and therefore any added notes are not published to the ClinGen Evidence Repository, nor are any relevant captured PMIDs.

8.2. Modified Strength of ACMG/AMP Criteria

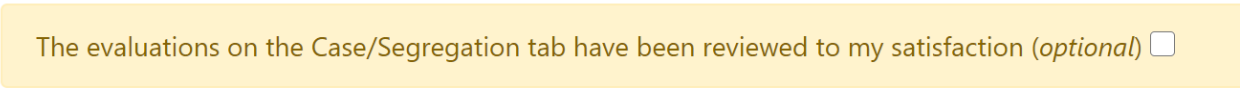
The ACMG/AMP guidelines for variant interpretation include the option of using professional judgement to move criteria listed as one weight to another weight. VCEPs are free to utilize these strength modifications in their specifications. These weight modifications are especially important for quantitative evidence types such as co-segregation with disease in

affected family members (PP1), which can increase in strength with increasing segregation data. Alternatively, strength of given criteria can be downgraded in certain instances. For example, a well-established functional study supportive of a damaging role has a strong level of strength (PS3); however, it may be lessened to moderate or supporting with weaker or less definitive types of functional studies. The SVI has recommended standardized nomenclature to address these changes (https://clinicalgenome.org/site/assets/files/3459/svi_criteria_nomenclature_recommendation_v1.pdf). For strength-modified evidence, the original criterion code is followed by an underscore and the new level of strength (e.g. PP1_Strong, PP1_Moderate, PP1).

8.3. Summary of steps for criteria evaluation

1. Identify, document, and examine evidence associated with the criterion being evaluated.
2. Based on the evidence, select an evaluation for each criterion from the pull-down menu.
3. Enter a text explanation to support your criteria evaluation selection. VCEPs may have a standardized text document for this purpose.
4. Select “Save”.
5. The “Save” button will now change to “Update”. If you would like to change an evaluation, make the edits and be sure to click “Update” afterwards.
   Note: When 2 (or more) criteria are mutually exclusive and cannot be “Met” at the same time, the interface will not allow “Met” to be selected for more than one of the criteria (e.g. PP3 and BP4 cannot both be “Met” for the same variant).
6. If you have evaluated all the evidence on a particular tab to your satisfaction, you can click the checkbox at the bottom of the page (for the Variant Type tab, this means you have evaluated any relevant subtabs to your satisfaction) and a check will appear on the tab for your reference (Figure 9). This checkbox will remain regardless of which tab you are on in the interface and can be unchecked as well. Note that checking this box is optional.

Figure 9.



Once all evidence has been evaluated, a provisional classification can be made for the variant. The VCI will aggregate the selected criteria evaluations using the criteria combining rules outlined by the ACMG/AMP guidelines and will provide a calculated pathogenicity; either Pathogenic, Likely Pathogenic, Likely Benign, Benign, or Uncertain Significance. This calculated classification, saved provisional and approved interpretation(s), criteria meeting

an evaluation strength, criteria evaluated as “Not met” and criteria ‘Not evaluated” can be found by accessing the **evaluation summary** via the ‘View Summary’ button shown in Figure 10. For further details on the evaluation summary, see the “Evaluation Summary” section.

Figure 10. The ‘View Summary’ button in the VCI

Any VCEP modifications to the original ACMG/AMP combining rules must be manually applied by overriding the classification with an explanation. This can be done based on expert opinion/clinical judgment, must have the approval of the VCEP or approving body, and the rationale must be clearly documented in the ‘explain reason(s) for change’ free text box located below the modified pathogenicity classification selection box on the summary page. This explanation is required, and must also be included in the evidence summary for public display.

9. EVIDENCE TABS

9.1. BASIC INFORMATION TAB

This tab provides high-level information about nomenclature and ClinVar assertions (https://www.ncbi.nlm.nih.gov/clinvar/) of the variant chosen for classification. Each section is detailed below.

All interpretations for this variant in the Variant Curation Interface (VCI): Classifications marked as “Approved” may be viewed by any user within the interface; those marked as “In Progress” or “Provisional” are viewable only by the members of the submitting affiliation.

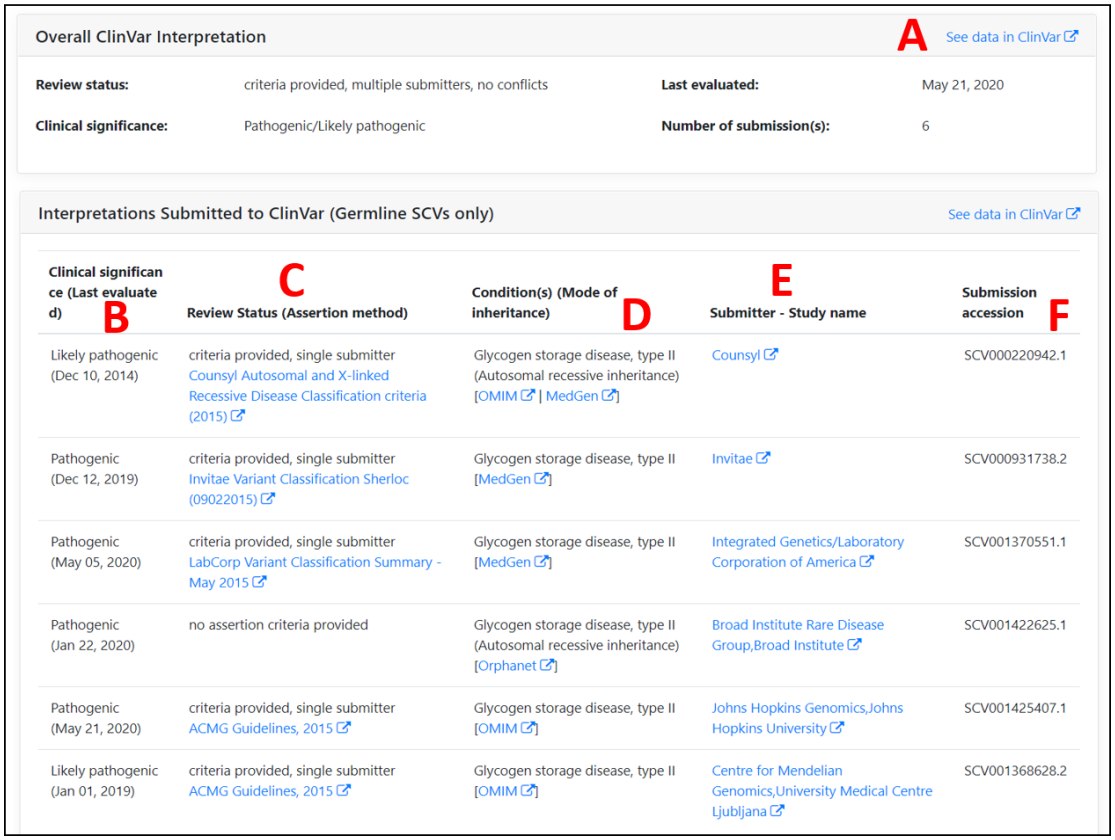
ClinVar interpretation: The review status of the overall variant record in ClinVar is provided for the curator, along with the aggregated clinical significance, total number of submissions and the date last evaluated. This helps to alert the curator if the variant is expert-reviewed and to the concordance of the interpretation (i.e. if there are any conflicts). As published

literature and information about a variant can change often, it is important to keep the date last evaluated in mind. In addition, this section may change over time as you classify the variant, as new data may be submitted to ClinVar.

The ClinVar entry can be accessed directly by clicking the “See data in ClinVar” arrow (A, Figure 11). The ClinVar entry provides access to a submitter’s summary evidence for the variant, which provides a short description of their rationale for classification. This can be found on the Summary Evidence Tab in ClinVar, in the Description column. Many ClinVar entries have links to publications that contain relevant information to the interpretation of the variant.

Information is also provided about each individual ClinVar submission, including the clinical significance and the date it was last evaluated (B), the review status and assertion method (C), condition and mode of inheritance (D), the submitter or if a study name is associated with the submission (E), and the submission accession (SCV - submission to ClinVar) number (F). For more information about ClinVar terminology, please see Harrison et al 2016[2].

Figure 11. The Basic Information tab in the VCI



Please keep in mind that variants can have multiple interpretations based on the condition, mode of inheritance, and whether the variant was reported as somatic or germline. Therefore, it is critical to crosscheck information in column D with the condition and mode of inheritance for the variant-disease association of interest. In general, ClinVar is best used to identify sources of published or unpublished data, and review other laboratory's rationales for classification. If a classification in a ClinVar submission is discordant with the classification arrived at by the VCEP, it may be useful to reach out to the submitting laboratory to discuss the basis for classification.

Transcript Information

Below the ClinVar submission section is information regarding transcripts (Figure 12). A sequence variant can fall on multiple transcripts and may have differing molecular consequences depending on the transcript evaluated. Generally, although a primary transcript is chosen for evaluation, it is critical to evaluate the effect of the variant on all transcripts when proposing a variant classification, especially if the biologically and disease-relevant transcript is uncertain. This information is easily viewable under the Protein change and Molecular Consequence columns of this section. Most often, the longest transcript is chosen as the primary transcript, unless the variant is predicted to have a more severe impact on another transcript (e.g., if a variant is predicted to result in an intronic change on the longest transcript and a nonsense change on a shorter transcript) or a different transcript has been reported as primary or most biologically relevant. It is important to keep in mind the disease of interest when performing transcript evaluation. For more information on choosing transcripts for variant interpretation, please see DiStefano et al 2018[3].

ClinVar primary transcript: The ClinVar primary transcript is a RefSeq transcript designated by ClinVar. The choice of the transcript is largely driven by which transcript(s) have previously been submitted to ClinVar for that particular variant.

RefSeq Transcripts: RefSeq is a transcript annotation and curation effort headed by the National Center for Biotechnology Information (NCBI). These transcripts are most commonly used by clinical labs, are independent of the genome build, and many are supported by manual literature curation. Transcripts can be preceded by multiple prefixes:

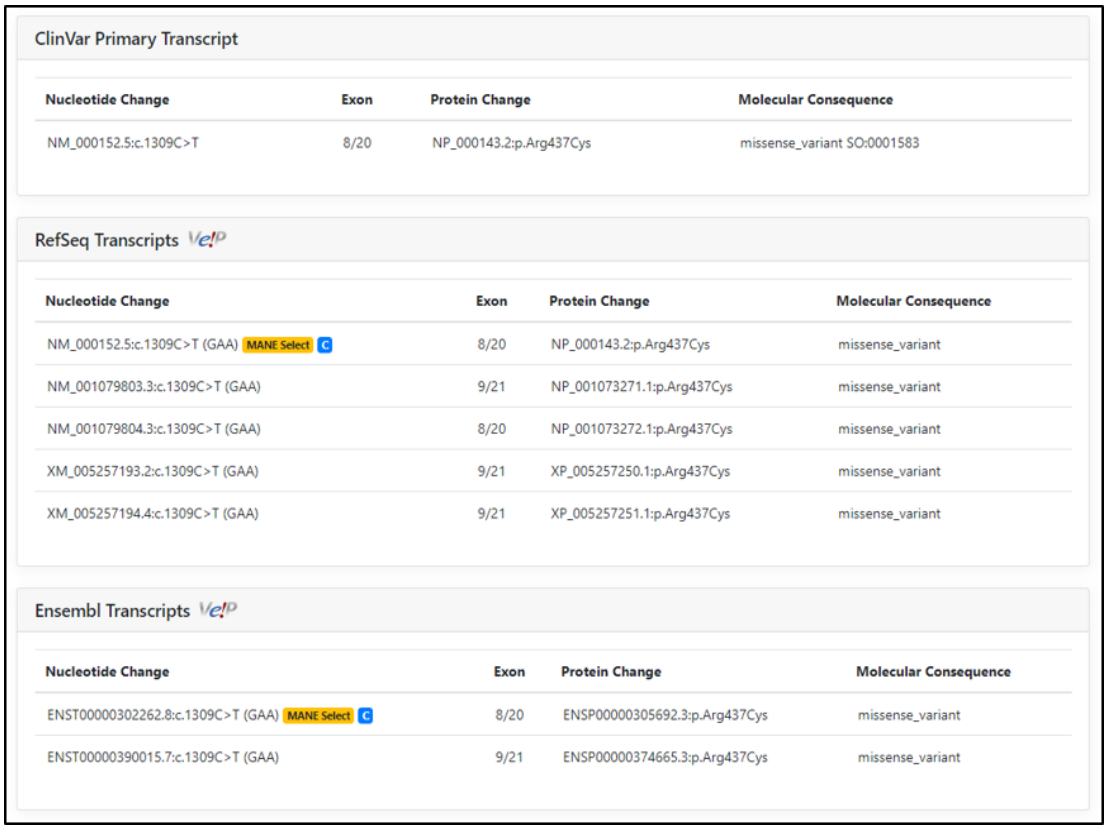
- NM, NR, NP: mRNA, non-coding RNA, and protein sequences, respectively that have been curated by the RefSeq team and are supported by some evidence, whether it is published literature or GenBank cDNA or EST data.
- XM, XR, XP: mRNA, non-coding RNA, and protein sequences, respectively that are predicted, but not confirmed with curation or evidence.

Ensembl Transcripts: Ensembl is a human genome annotation effort headed by the European Bioinformatics Institute (EBI) that includes a transcript annotation and curation

effort. These transcripts are used by gnomAD, are dependent upon the genome build, and are supported by computational and/or manual curation.

Within the VCI, the MANE Select transcript for the gene is indicated for both RefSeq and Ensembl (Figure 12). MANE (Matched Annotation from NCBI and EMBL-EBI) is a collaboration between NCBI and EMBL-EMBL to generate identically matched transcript sequences for human genes. The MANE Select transcript set represents one primary high-quality transcript per protein-coding gene, supported by experimental data. For a small number of genes, a MANE Plus Clinical transcript may also be annotated, which means that there is at least one additional transcript with well-established pathogenic variants not found in the MANE Select transcript set. The canonical transcript for the gene, defined here as the transcript with the longest coding sequence if the gene has translated transcripts, or the longest cDNA if it does not, is also shown for both RefSeq and Ensembl.

Figure 12. Transcript information in the VCI; the MANE Select transcript and canonical transcript (C) are labeled.



Variant Genomic Context: This section lists the HGVS notation for the genomic coordinates on the two most recent genome builds, GRCh37 and GRCh38. Genome builds can differ across resources such as population databases, so it is critical to double check the genome build when curating a variant.

9.2. POPULATION TAB

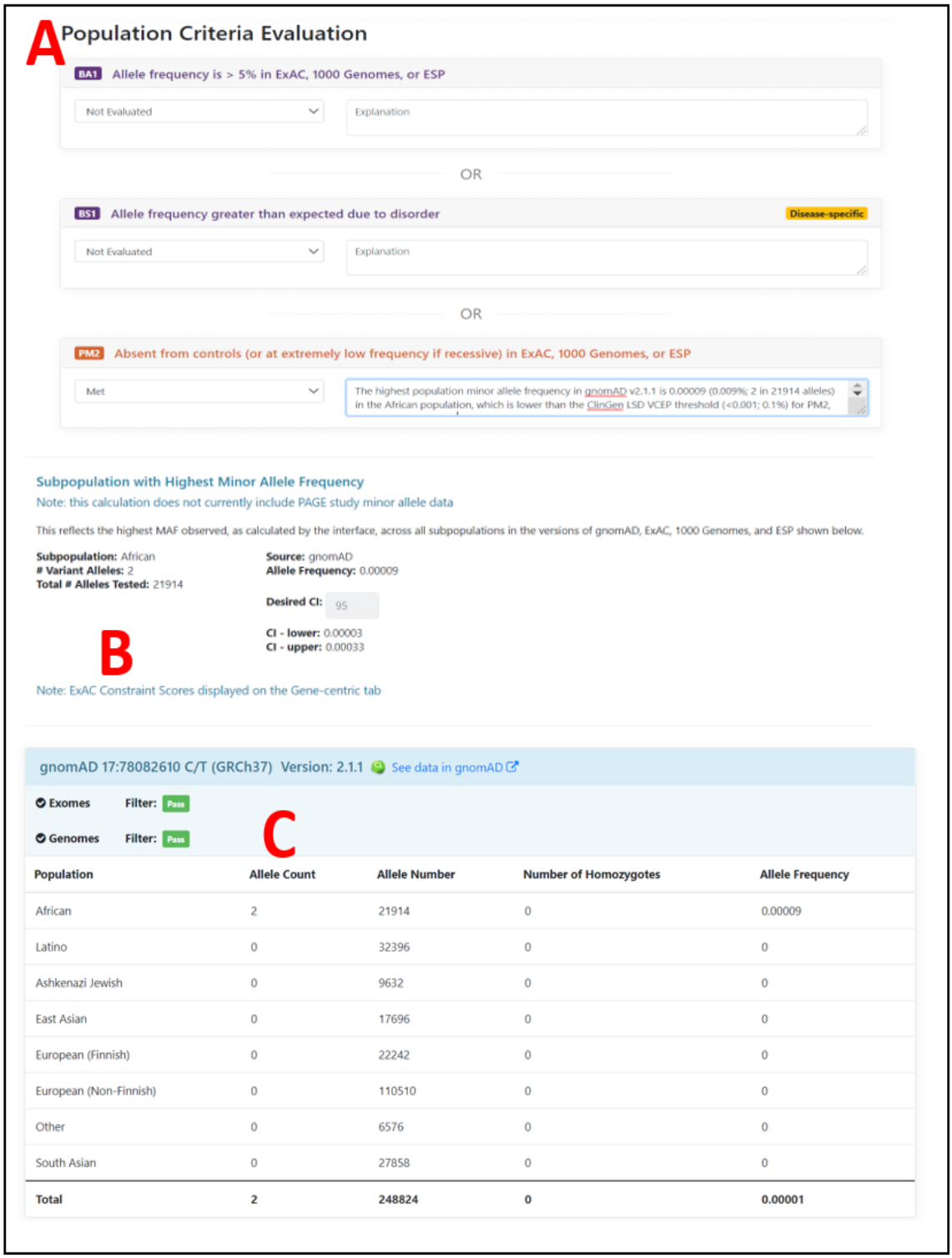
ACMG/AMP criteria codes: BA1, BS1, PM2

Background: This tab displays population frequency data, if available, from the Genome Aggregation Database (gnomAD), Exome Aggregation Consortium (ExAC), Population Architecture using Genomics and Epidemiology (PAGE), 1000 Genomes (1000G) and the Exome Sequencing Project (ESP), including minor allele frequency (MAF) and genotype data (Figure 13). For example, for the variant displayed below, the highest MAF (0.0009) is found in the African subpopulation in gnomAD (Figure 13, C). These data are displayed in the "Highest Minor Allele Frequency" (Figure 13, B) section along with subpopulation, source, total number of variant alleles, and total number of alleles tested in that subpopulation. The VCI defaults to showing the upper and lower bounds of the 95% confidence interval to help users determine the level of confidence in the computed MAF.

Limitations:

● The VCI pulls in data for all short genetic variants for gnomAD v2.1.1 and ExAC. For PAGE, 1000 Genomes and ESP, data is currently only auto-populated for single nucleotide substitution variants, and not deletion or insertion variants. If assessing a deletion or insertion variant, please manually search the population frequency resources using either the dbSNP ID (provided in the Variant ID Sources section at the top of the page) or the genomic positions (provided in the Genomic section on the Basic Information tab). The ClinGen Allele Registry also provides direct links to the population frequencies and is faster than some of the other searches. Curators can view variants identified in ExAC and gnomAD in the region (+/- 30 bp) of the deletion/insertion by clicking the links provided in the Population tab.

Figure 13. Population tab in the VCI



Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
African	2	21914	0	0.00009
Latino	0	32396	0	0
Ashkenazi Jewish	0	9632	0	0
East Asian	0	17696	0	0
European (Finnish)	0	22242	0	0
European (Non-Finnish)	0	110510	0	0
Other	0	6576	0	0
South Asian	0	27858	0	0
Total	**2**	**248824**	**0**	**0.00001**

Population Tab Criteria Evaluations:

General instructions:

● Please confirm variant information is correct by checking the genomic coordinates and/or dbSNP IDs.

● As criteria codes BA1, BS1, and PM2 are mutually exclusive, only one of the three criteria are allowed to be "Met."

BA1: Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.

● The default BA1 threshold (5%) is displayed in the MAF cutoff field.

● This threshold is often modified in a gene/disease-specific manner by the appropriate VCEPs, to account for variability of disease prevalence and genetic heterogeneity. Some VCEPs have also defined a set of variants (exclusion variants) whose frequency surpasses the designated BA1 threshold in one or more population data sets, but have a plausible argument for pathogenicity and should be exempted from this rule. For example, BTD NM_000060.4:c.1330G>C [p.Asp444His] has a MAF >5% in gnomAD (0.05558 in the Finnish population) but is classified as Pathogenic (ACMG/AMP criteria also applied: 15 PS3; PM3_Strong; PP3; PP4); therefore, the BA1 criterion is not applicable. Please check with the VCEP with which you are working for any specifications of their BA1 rule.

● The SVI has provided an updated definition of this criterion which states that “Allele frequency is >0.05 in any general continental population dataset of at least 2000 observed alleles and found in a gene without a gene- or variant-specific BA1 modification”. They have also provided an exclusion set of nine variants (exclusion variants defined above) with an allele frequency over 5% where BA1 does not apply[4] .

● The SVI has provided a mechanism for identifying further alleles that should be excluded from this BA1 criteria to be shared publicly through a submission form located on the ClinGen webpage (https://www.clinicalgenome.org/working-groups/sequencevariant-interpretation/)

BS1: Allele frequency is greater than expected for disorder.

● This is disease-specific. Please consult with the specific VCEP (if applicable) for details.

● The incidence of the disorder being analyzed can be found using the resources located in the Resource Section of this SOP (such as the Genetic Testing Registry). If the variant is found more frequently in the population (or at least one subpopulation in ExAC or gnomAD) than the disease occurs, this is considered strong evidence of a benign impact.

PM2: Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.

● Although the original code states “absent”, many VCEPs have modified this code in a disease-specific manner. Please consult with the specific VCEP (if applicable) for details.

● **The SVI recommends using PM2 at no higher than the supporting level** (https://clinicalgenome.org/site/assets/files/5182/pm2_-_svi_recommendation_-_approved_sept2020.pdf)

● As with BS1, incidence of the disorder being analyzed can be found using available resources. If the variant is found at a frequency below the expected carrier frequency of a given autosomal recessive condition, this is considered moderate evidence of a pathogenic impact.

● Since coverage information is not currently displayed (for variants present or absent in population databases), please use the link-out to navigate to the population database to determine if the region has adequate coverage (typically >20X coverage). If the variant is present in a population database, a ‘filter’ button will appear above the pre-populated data. If coverage of the region is adequate, this button will be a green ‘Pass’ (Figure 12, C).

Evidence Explanation Notes:

● In the Explanation field (Figure 13, A), please list the frequency, source of data used, and the allele counts. For example, “Variant identified in 0.009% (2/21914) of African chromosomes in gnomAD” for the variant shown in Figure 12. Although the date you are working on the variant is maintained in the system, it is also helpful to specify in your comment the date on which this information is captured.

9.3. VARIANT TYPE TAB

Under the parent Variant Type tab are four sub-tabs separated out by variant type; Missense, Loss of Function, Silent & Intron, and In-frame Indel (Figure 14) with each sub-tab containing the ACMG/AMP criteria specific for that variant type. For example, criteria PS1 and PM5 (missense change in a codon with other pathogenic missense changes) are only found on the Missense tab, while PVS1 (LOF variant) is only found on the Loss of Function tab. Thus, **curators only need to complete the Variant Type sub-tab specific for their variant**. However, please note that PP3, which is on the missense subtab, can also be used for non-canonical +/-1 or 2 splicing variants and inframe deletions and insertions.

Figure 14. Sub-tabs found under the Variant Type tab in the VCI

Missense sub-tab

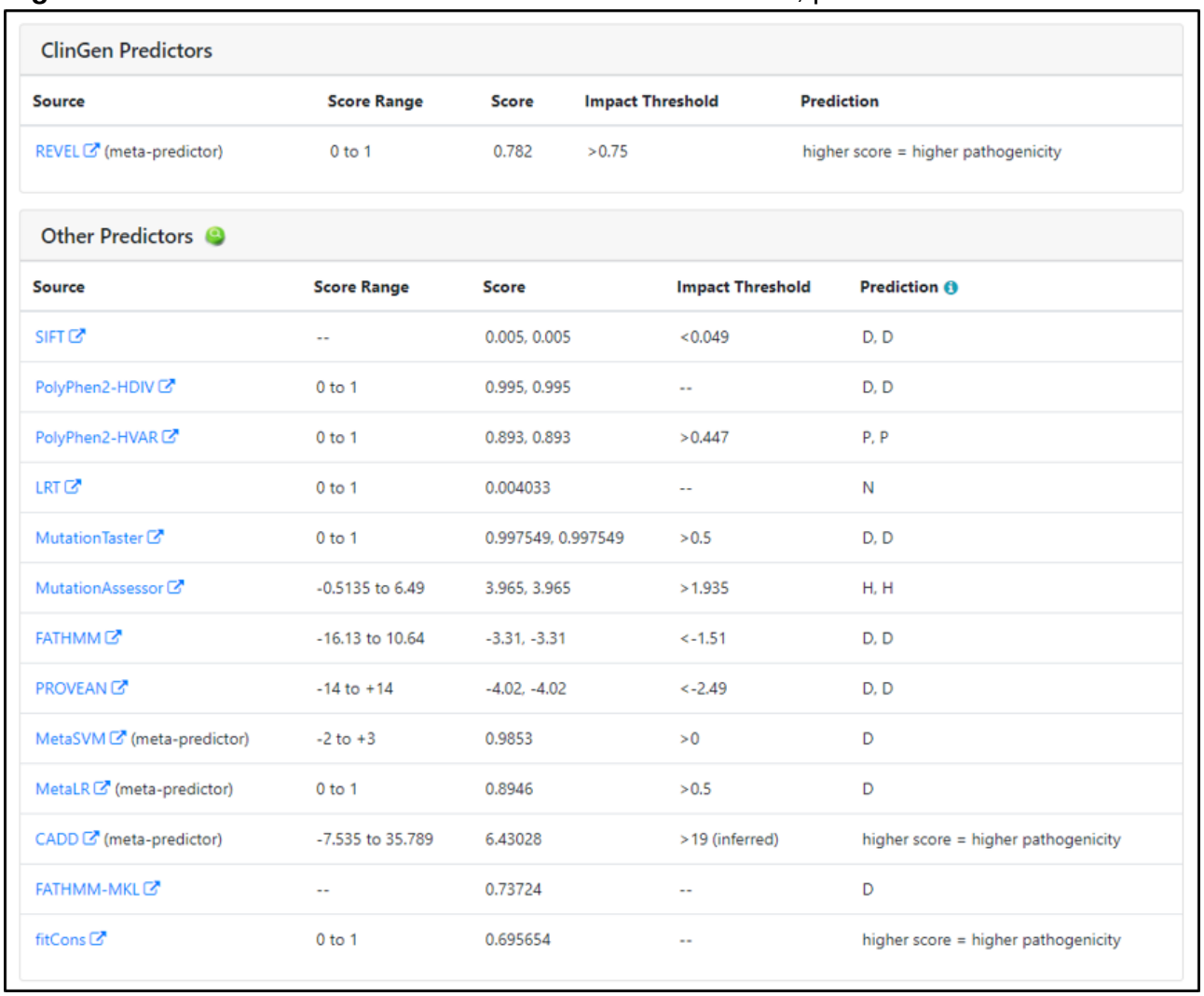
Functional, Conservation, and Splicing Predictors Section
ACMG/AMP criteria codes: PP3, BP4, PP2, BP1

Background: The VCI provides data from a large number of computational predictors to allow flexibility for a given VCEP. These tools vary in predictive power based on the gene and other factors. Follow the specifications from your VCEP regarding which tools to use for which variant types, score thresholds, and so on.

For **missense variants**, multiple prediction scores are displayed and separated out by type. The first type, "ClinGen Predictors" (Figure 15), displays results from the meta-predictor REVEL[5] , which predicts the pathogenicity of missense variants based on a combination of scores from 13 individual tools: MutPred, FATHMM v2.3, VEST 3.0, Polyphen-2, SIFT, PROVEAN, MutationAssessor, MutationTaster, LRT, GERP++, SiPhy, phyloP, and phastCons (please see Appendix I, Computational Tools, for links to each individual tool). The REVEL score can range from 0 to 1, with 1 most suggestive of pathogenicity. As many of the individual tools incorporated by REVEL are also displayed in the VCI (Other Predictors), the score from REVEL is most useful in assessing whether, in aggregate, the tools are suggestive of a pathogenic or benign impact. In general, a score > 0.75 is considered evidence of pathogenicity, though some VCEPs have specified different cutoffs for PP3 and BP4 or recommended the combination of REVEL and other predictors.

Below "ClinGen Predictors", "Other Predictors" (Figure 15) displays results from 13 individual prediction tools (SIFT, PolyPhen2-HDIV, PolyPhen2-HVAR, LRT, MutationTaster, MutationAssessor, FATHMM, PROVEAN, MetSVM, MetLR, CADD, FATHMM-MKL, fitCons) and includes both the quantitative score and prediction classification from each tool. While there is some overlap between these 13 prediction tools and the 13 tools aggregated in REVEL, they are not identical. Informational links for these 13 individual tools can also be found in Appendix I, Computational Tools. As some prediction tools assess the impact of the missense change on all transcripts, these tools may display multiple quantitative scores and prediction classifications for a single variant. To determine the prediction classification text from each single letter abbreviation displayed, hover over the information button (Table 3).

Figure 15. Predictors section of the Missense tab in the VCI, part 1



ClinGen Predictors

Source	Score Range	Score	Impact Threshold	Prediction
REVEL (meta-predictor)	0 to 1	0.782	>0.75	higher score = higher pathogenicity

Other Predictors

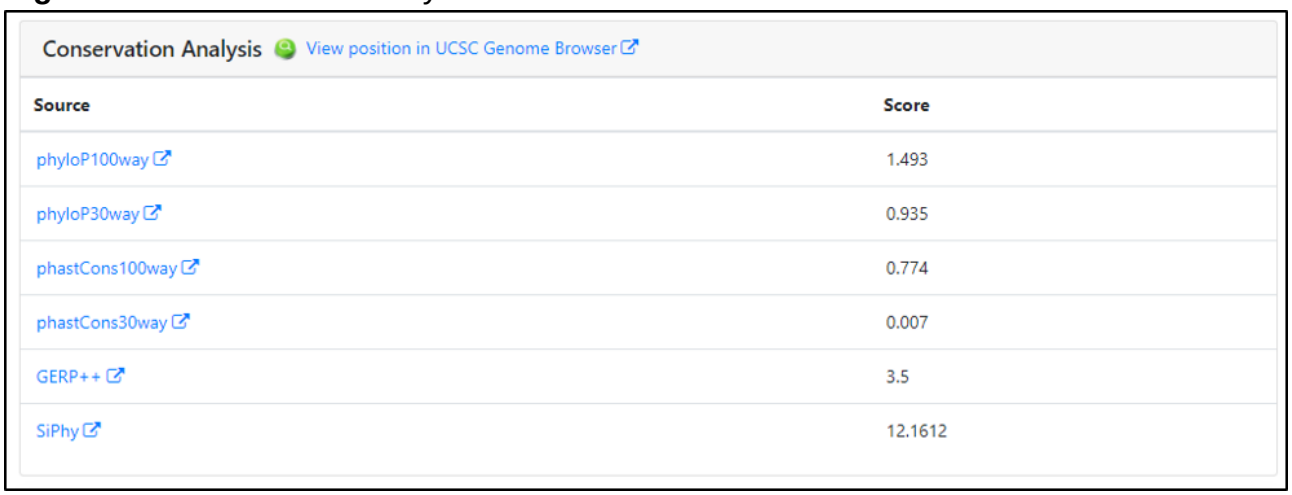
Source	Score Range	Score	Impact Threshold	Prediction
SIFT	--	0.005, 0.005	<0.049	D, D
PolyPhen2-HDIV	0 to 1	0.995, 0.995	--	D, D
PolyPhen2-HVAR	0 to 1	0.893, 0.893	>0.447	P, P
LRT	0 to 1	0.004033	--	N
MutationTaster	0 to 1	0.997549, 0.997549	>0.5	D, D
MutationAssessor	-0.5135 to 6.49	3.965, 3.965	>1.935	H, H
FATHMM	-16.13 to 10.64	-3.31, -3.31	<-1.51	D, D
PROVEAN	-14 to +14	-4.02, -4.02	<-2.49	D, D
MetaSVM (meta-predictor)	-2 to +3	0.9853	>0	D
MetaLR (meta-predictor)	0 to 1	0.8946	>0.5	D
CADD (meta-predictor)	-7.535 to 35.789	6.43028	>19 (inferred)	higher score = higher pathogenicity
FATHMM-MKL	--	0.73724	--	D
fitCons	0 to 1	0.695654	--	higher score = higher pathogenicity

Table 3. Prediction classification text letter code key.

Predictor	Letter Code	Prediction
SIFT	D	Damaging
	T	Tolerated
PolyPhen2-HDIV PolyPhen2-HVAR	D	probably Damaging
	P	Possibly damaging
	B	Benign
LRT	D	Deleterious
	N	Neutral
	U	Unknown
MutationTaster	A	disease causing Automatic
	D	Disease causing
	P	Polymorphism automatic
	N	polymorphism
MutationAssessor	H	High (predicted functional)
	M	Medium (predicted functional)
	L	Low (predicted non-functional)
	N	Neutral
FATHMM	D	Damaging
	T	Tolerated
PROVEAN	D	Damaging
	N	Neutral
METASVM	D	Damaging
	T	Tolerated
METALR	D	Damaging
	T	Tolerated

The next section, **Conservation Analysis** (Figure 16), contains links to, and displays results from, six conservation analysis tools. The conservation of a nucleotide at a particular position in the genome can give an indication of its importance. Selective constraint can be used to assess functional significance of a variant. There are multiple tools available for conservation analysis. The VCI utilizes several programs from the Phylogenetic Analysis with Space/Time Models (PHAST) software package, including phyloP100way, phyloP20way, phastCons100way and phastCons20way. The score for each is automatically calculated and provided in the VCI. Again, many of the predictors previously described incorporate these conservation scores.

Figure 16. Conservation analysis tools in the VCI



● PhyloP is a suite of programs that measure the calculated p-value of conservation versus acceleration (non-conserved) at each nucleotide across an aligned area, as compared to a model of neutral conservation. The absolute value score range is -14 to 3. A negative value represents faster than expected evolution, while a positive score indicates slower than expected evolution (areas predicted to be conserved). Information and tutorials can be found at the following link: http://compgen.cshl.edu/phast/index.php.

● PhastCons calculates conservation using a Phylogenetic Hidden Markov Model. The score ranges from 0 to 1, with a lower score indicating faster than expected evolution and a higher score indicating conservation among 17 vertebrate species. Additional information and tutorials for phastCons100way and phastCons20 way can be found here: http://compgen.cshl.edu/phast/index.php.

● The Genomic Evolutionary Rate Profiling (GERP++; http://mendel.stanford.edu/SidowLab/downloads/gerp/) program estimates position-specific evolutionary constraint by measuring the number of substitutions expected minus those observed under neutral drift at each position independently. The calculated score varies with the level of conservation; positive scores indicate evolutionary constraint. Thresholds are chosen based on desired sensitivity and specificity, and may be dictated by VCEP; discuss with VCEP if applicable.

● SiPhy, available through the Broad Institute (http://portals.broadinstitute.org/genome_bio/siphy/index.html), uses deeply sequenced phylogeny data and multiple statistical tests to identify conservation patterns.

The last section under “Functional, Conservation, and Splicing Predictors’’ is **Splice Site Predictors**. Splice site predictors utilize bioinformatics tools and splice site consensus sequences to predict variant effects. While the VCI offers dynamic links to four of these tools - MaxEntScan, NNSPLICE, SpliceAI and VarSEAK - it does not yet automatically import

data from these tools. The curator must access and utilize these tools via the links provided if they wish to use this data. Many VCEPs have specified the splice predictor tools that are to be used in variant classification for the gene in question, as described in Limitations below.

- MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) is based on the 'Maximum Entropy Principle', and utilizes probabilistic models of short sequence motifs to account for non-adjacent and adjacent dependencies between nucleotide positions. The score generated is the difference between a reference allele and a variant, and a higher score implies a higher probability of a true splice site.
- NNSPLICE (http://www.fruitfly.org/seq_tools/splice.html) is part of the Berkeley Drosophila Genome Project and utilizes a neural network method trained to recognize 5' and 3' eukaryotic splice sites using a representative data set from D. melanogaster. The score ranges from 0-1, with anything above 0.5 indicative of a possible splice site gain.
- SpliceAI[6] is a deep neural network based on pre-mRNA transcript sequences that predicts splice sites using long-range primary genomic sequence flanking each position as input (+/-50 bp as default; +/-10,000 bp maximum). The user inputs HGVS nomenclature for the variant (https://spliceailookup.broadinstitute.org/). SpliceAI provides a table with delta scores (0-1) for acceptor loss, donor loss, acceptor gain, and donor gain within the designated flanking sequence. The delta score indicates the probability that the variant will alter splicing at the pre-mRNA position indicated.
- VarSEAK's JSI splice site prediction tool (https://varseak.bio/) predicts splicing effects for genetic variants based on canonical splice site sequences (core motif GT for 5' donor splice sites or AG for 3' acceptor splice sites). The user enters the gene name, transcript, and variant. Output includes a graphical representation of the normal and variant sequence with annotated splicing impact, the overall splicing prediction class (1 for no splicing effect, to 5 for splicing effect) and a table with relevant splicing positions including the splice site prediction score, and ENT and delta ENT scores from MaxEntScan.

Limitations:

Splicing in silico tools can be difficult to utilize and the interpretation is often not standardized. More training is necessary than can be provided in this general SOP. Although the VCI provides links to all of these tools, each VCEP will determine the tools most useful for their group. Each Expert Panel should validate which of these tools are able to predict the native sites for each splice donor/acceptor for their gene, specify which tools will be used by their EP, and provide training to their biocurators. Before evaluating a variant using the provided links, determine the tools to be used and receive instruction in usage, thresholds, etc., with the appropriate VCEP.

Functional, Conservation, and Splicing Predictors Section Criteria Evaluations:
PP3: Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.)
BP4: Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.)

● To complete assessment of PP3 and BP4, compare in silico predictors, conservation analysis predictors, and/or splice site predictors to determine if overall the predictors suggest a deleterious effect (PP3), no impact (BP4), or if predictions are mixed/unclear. It is not necessary for all prediction tools to be concordant to apply PP3 or BP4, but expert judgment or VCEP specifications must be applied and a consistent threshold should be used for all the variants in that gene.
● In application of PP3/BP4, a meta-predictor such as REVEL may be used in place of multiple predictors in the in silico analysis of missense variants.
● As the positive predictive value for a tool may vary by the gene, VCEPs may define thresholds based on a specific prediction tool or set of prediction tools for their gene of interest.
● In silico tools to be used should be specified by the VCEP. In particular, splicing results often benefit from discussion with the full EP (or splicing experts within the panel) for proper application, and any abnormal results or questions should result in screenshots of the output and presentation to the EP.

PP2: Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease.

As a general recommendation based on gnomAD data, the SVI Committee suggests a gene with a missense z-score > 3.09 is more likely to be intolerant of missense changes[7]. As this will be gene-specific, each VCEP will decide if PP2 applies for their gene. Please check with your VCEP if appropriate.

BP1: Missense variant in a gene for which primarily truncating variants are known to cause disease.
● To complete assessment of PP2 and BP1, additional gene-centric knowledge is required regarding the variant spectrum of the gene/disease of interest. Currently, the only data displayed in the interface to support assessment of these criteria are the LOF ExAC Constraint Scores displayed on the Gene-centric tab (pLI). pLI is the probability of a gene being LOF-intolerant. Those genes with pLI scores of 0.9 and above have the highest likelihood of LOF intolerance in a heterozygous state. This score is most useful for haploinsufficient genes with pediatric onset conditions. It will not identify recessive disease

genes that tolerate LOF carriers, or late onset disease genes where the variants are passed on to the next generation. ExAC also computes a missense constraint score (z), available on any gene page in ExAC, but this score is not currently displayed in the interface. This z-score represents the deviation of the observed from expected variant count. A positive z-score indicates increased intolerance to variation (fewer variants than expected), and a negative z-score indicates variant tolerance (more variants than expected). GnomAD also computes missense and LOF gene constraint metrics; however, these are based on an observed versus expected score which is calculated slightly differently from the ExAC constraint metrics. This ratio is a continuous measure of gene tolerance to different classes of variant (i.e., missense, LOF) and a LOWER score indicates stronger selection (decreased tolerance) for that variant type.

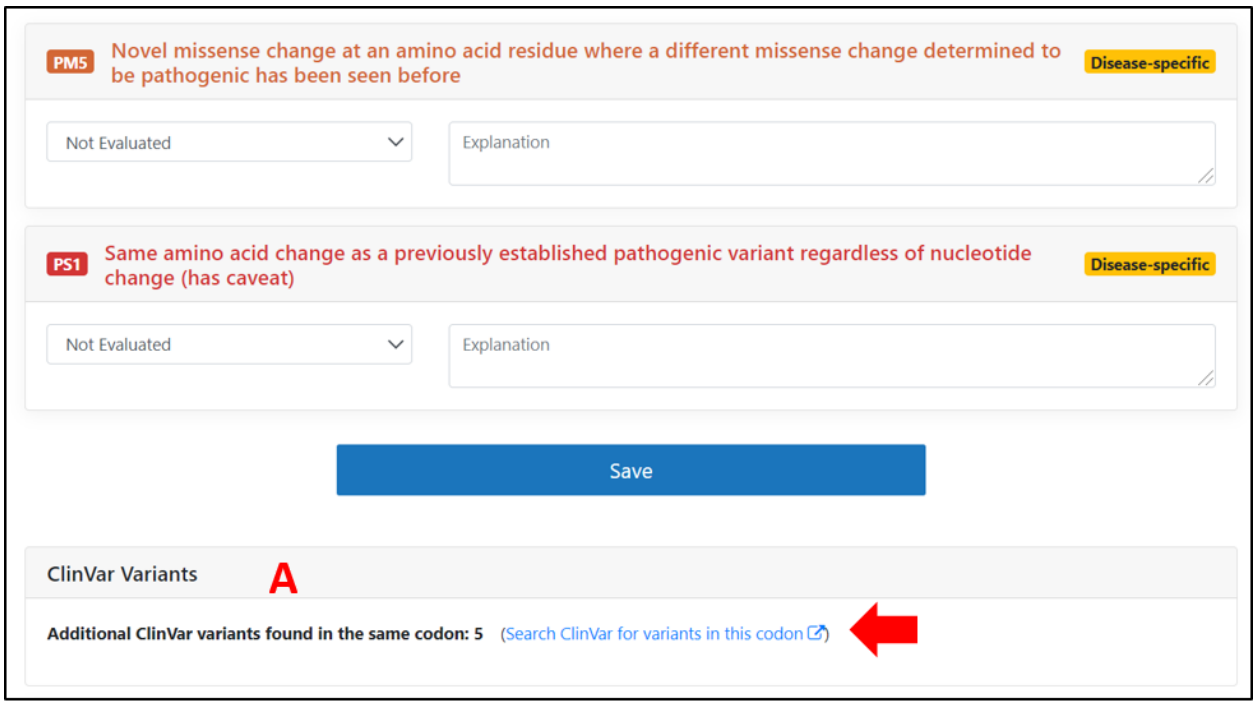
- To thoroughly evaluate these criteria, it is recommended that curators assess the variant spectrum (missense/nonsense, splicing, indels, LOF, copy number variants, etc.) for the gene in ClinVar, the Human Gene Mutation Database (HGMD), ExAC, and gnomAD as described above, and receive guidance from the appropriate VCEP, if available. Please note that HGMD (http://www.hgmd.cf.ac.uk/ac/index.php) has two versions - a public version and a licensed version. The public version is less up-to-date, but freely available to users from academic institutions or non-profit organizations upon registration.
- As this will be gene-specific, each VCEP will decide if BP1 applies for their gene. Please check with the VCEP if appropriate.

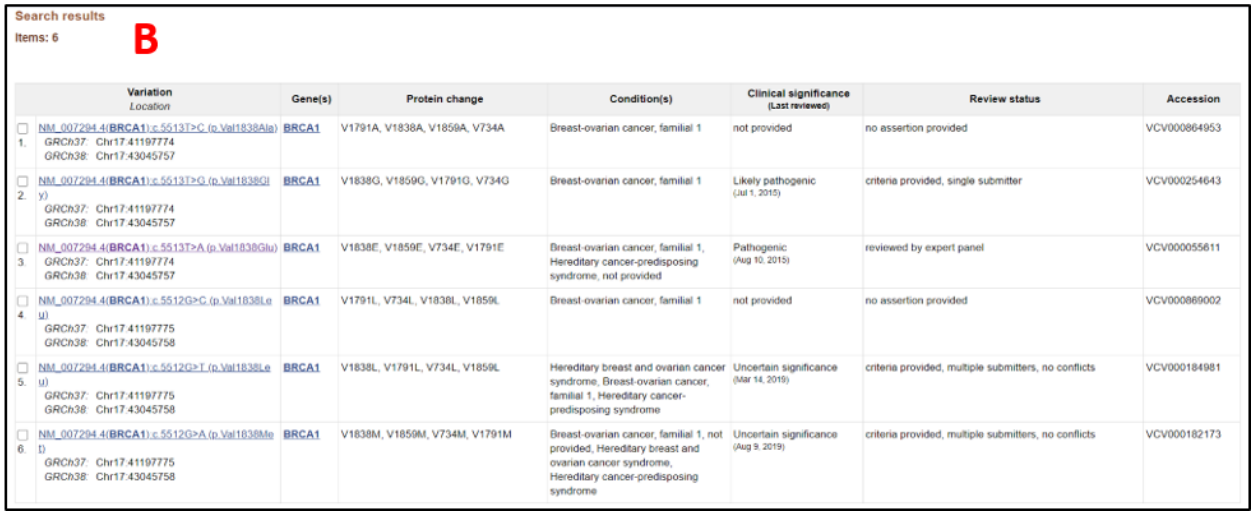
Other Variants in Same Codon Section

ACMG/AMP criteria codes: PS1, PM5

Background: This section provides a link to additional ClinVar variants in the codon of interest (click "Search ClinVar for variants in this codon"). Please note, if the variant under assessment is already in ClinVar, the ClinVar search result will also return that variant. For example, if assessing variant NM_007294.3(BRCA1):c.5513T>A (p.Val1838Glu), the interface informs the curator that 5 additional variants are found in the Val1838 codon in ClinVar (Figure 17, A). The search result screen in ClinVar will display 6 returned variants as the c.5513T>A (p.Val1838Glu) variant is already present in ClinVar (Figure 17, B). Also note that the variants shown are not dependent on their classification or the variant type. Therefore, the curator must assess each of these variants to determine if any of them can be used as evidence for the variant being curated.

Figure 17. A, 'Other Variant in Same Codon' section of Missense tab, and B, ClinVar search result screen.





Other Variants in the Same Codon Criteria Evaluations:

PS1: Same amino acid change as a previously established pathogenic***** variant regardless of nucleotide change.
Example: Val to Leu caused by either G>C or G>T in the same codon.
Caveat: Beware of changes that impact splicing rather than the predicted amino acid/protein change.

PM5: Novel missense change at an amino acid residue where a different missense change determined to be pathogenic* has been seen before.
Example: p.Arg156Cys is pathogenic; now you observe p.Arg156His.
Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level.

***** Note that some VCEPs will consider PS1/PM5 if the previous missense change has a 'pathogenic' OR a 'likely pathogenic' classification.

● The alert to additional ClinVar variants in the codon is a helpful tool to identify these variants; however, assessment of the pathogenicity of the additional variants should not solely depend on pathogenic interpretations in ClinVar. Additionally, as ClinVar may not contain all known variants within a codon, curators should also search other variant databases. Curators must thoroughly evaluate the potential pathogenicity of other variants in the codon and not rely on the variant's database designation, as these may not reflect current variant interpretation practice.
● These criteria are VCEP-specific; please consult the appropriate VCEP if applicable. Some examples of VCEP specifications are describe here:
○ Certain VCEPs do not require the missense change to be novel.
○ Certain VCEPs will consider PS1/PM5 if the previous missense change has a 'pathogenic' OR a 'likely pathogenic' classification. To meet either of these criteria, the other variant in the codon should reach a pathogenic (or likely pathogenic, if appropriate) classification without relying on PS1 or PM5, in order to avoid circular dependencies.
○ Certain VCEPs require that an in silico tool predicts the missense change being interrogated to be more damaging than the known variant or the Granthem distance between the native amino acid and the variant be greater than the Pathogenic change.
○ Certain VCEPs require complete curation for all variants at the relevant codon in order to completely assess whether PS1/PM5 should be applied.

Loss of Function (LOF) sub-tab

ACMG/AMP criteria code: PVS1

Background: Nonsense, frameshift, canonical +/- 1 or 2 splice site, initiation codon changes, and single or multi-exon deletions are all variant types that often (though not always) result in loss of protein function. The SVI has provided specific guidance on the use of this criterion[8]. A flowchart of these recommendations can be found in Appendix II. To apply the PVS1 criteria to these types of variants, the variant should be expected to lead to a truncated mRNA that undergoes nonsense-mediated decay, resulting in complete loss of protein translation, and be in a gene associated with a condition where the disease mechanism is consistent with LOF. The LOF tab in the VCI allows the user to capture any

evidence that supports these criteria but does not specifically address how to evaluate them. The following information can be used to aid in determining whether PVS1, (or one of its strength modifications), can be appropriately applied to a suspected null/LOF variant. It is important to keep in mind that use of this criterion is also disease-specific and should be cross-referenced with the Expert Panel for which curation is being done, although the ClinGen reference is the primary guide.

LOF Sub-tab Criteria Evaluations:

PVS1: Null variant (nonsense, frameshift, canonical +/- 1 or 2 splice sites, initiation codon, single or multi-exon deletion) in a gene where LOF is a known mechanism of disease (has caveats).

Disease Mechanism Note

This type of information is not always well established (particularly for recently identified and/or less-well studied disease genes), but in some instances is published in the literature. It is important to critically evaluate evidence supporting published assertions of LOF as the disease mechanism, since this can be assumed without sufficient supporting data in some publications. If strong genetic or functional evidence is not available, it is possible to estimate how well a gene tolerates LOF variants using the "constraint metrics" provided by ExAC (which are included on the gene-centric tab in the VCI) or gnomAD. These metrics provide a probability score (pLI) for a gene's intolerance of LOF variants (more detail can be found on the ExAC website: http://exac.broadinstitute.org/faq)[9]. For dominant diseases caused by LOF variants the expectation would be that the pLI score should be close to 1 (i.e. the gene is extremely intolerant to LOF variants) and there should be an excess of pathogenic LOF variants relative to other variant types. In gnomAD, an observed and expected variant score (o/e constraint metric) is also used to determine the probability that a given gene is intolerant to LOF variants (https://gnomad.broadinstitute.org/)[10] . The o/e metric differs from that of the pLI. It incorporates a 90% confidence interval, and has an opposite scale - the closer the o/e is to zero, the more likely the gene is LOF-constrained. At this time, these metrics only include nonsense and canonical splice site variants in the calculation; they do not include frameshift variants. Additionally, variant databases (internal or public such as ClinVar) can be used to assess the spectrum of reported pathogenic/likely pathogenic variants in a gene associated with the condition being curated. The Gene-centric tab automatically populates the ExAC constraint scores for a gene (Figure 18), but only provides a link-out to view ClinVar variants in that gene. **This method for evaluating disease mechanism should be used with caution**, as it primarily reflects constraint due to reproductive fitness. It is most useful for pediatric dominant disorders with a severe phenotype, and **should not be used for (most) adult onset Mendelian disorders** (e.g., BRCA1 has a pLI of 0 and an o/e of 0.73). If curating with a VCEP, discuss disease mechanism and whether use of this metric is appropriate, with the Expert Panel. An additional resource to consider is the ClinGen

Dosage Sensitivity Map (https://www.ncbi.nlm.nih.gov/projects/dbvar/clingen/)[11] . Through this resource, genes and genomic regions are evaluated for evidence of haploinsufficiency (HI) and triplosensitivity (TS). While this resource was originally created to assist with the interpretation of copy number variants, those genes with sufficient evidence of HI (i.e., an HI score of "3" per this evaluation system) are thought to have LOF as a disease mechanism.

Figure 18. Example of ExAC constraint scores for TP53 in the VCI

ExAC Constraint Scores

	pLI	pRec	pNull	syn Z	mis Z
All ExAC	0.91222	0.08775	0.00003	-0.04224	1.37892
Non-psych	0.86063	0.13926	0.00011	-0.16776	1.26734
Non-TCGA	0.97165	0.02834	0.00001	0.00886	2.12492

pLI: the probability of being loss-of-function intolerant (intolerant of both heterozygous and homozygous LOF variants)

pRec: the probability of being intolerant of homozygous, but not heterozygous LOF variants

pNull: the probability of being tolerant of both heterozygous and homozygous LOF variants

syn Z: corrected synonymous Z score

mis Z: corrected missense Z score

Assessing Molecular Consequence

Variants that potentially result in LOF should be visualized using a genome browser to determine their location with respect to exon structure and the 3' (C-terminal) end of the gene. The header provides a link out to multiple different genome browsers, and the transcript section in the Basic Information tab notes in which exon the variant is located and total number of exons for all transcripts (RefSeq, Ensembl) e.g. exon 2/20 (see Figure 12). Only those variants that result in premature termination codons > 50 bp upstream of the last exon-exon junction of the transcript are expected to result in nonsense-mediated decay (NMD) and consequently lead to LOF. Variants not predicted to undergo NMD may still impact a critical domain of the protein product or eliminate a large percentage of the protein (particularly if the protein is small and/or the last exon is very large). It is additionally important to consider which transcripts the variant impacts. In some scenarios, a nonsense variant may occur in only one transcript that is not biologically relevant; therefore, it is critical to determine if the LOF variant will impact a biologically relevant transcript.

For variants that may alter mRNA splicing (i.e. canonical splice site variants, internal exon deletions) and/or translation (initiation codon) it is important to consider what impact the variant will have on the predicted protein product. It is possible that such variants could result in in-frame insertions/deletions creating an intact protein of inappropriate size (shorter or larger), but still retain all necessary protein domains and thus have little impact on the

overall protein function. If curating as part of a VCEP, that Expert Panel will provide guidance on how to modify the PVS1 evidence level. Anything that falls outside of the specified rules should be presented to the Expert Panel for discussion and to make a final decision on evidence level.

Silent & Intron sub-tab

ACMG/AMP criteria code: BP7

Background: The major consideration for synonymous and/or intronic variants is whether they impact mRNA splicing. As previously mentioned, in silico splice site predictors can be useful in evaluating whether a variant might impact splicing, but these tools are not currently integrated into the VCI and instead are linked out to the appropriate sites directly within this tab. This tab also provides the option of including any published functional information regarding the splicing impact of synonymous and/or intronic variants. The BP7 criterion also takes into consideration evolutionary conservation of the nucleotide across species. This information can be evaluated using the conservation track provided by most genome browsers (which are linked out on the VCI header, e.g. see Figure 3).

Silent & Intron Subtab Criteria Evaluations:
BP7: A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved
● Some VCEPs do not require lack of a predicted splicing impact. Please check with the appropriate VCEP (if applicable).
● Nucleotide conservation comparison also varies by VCEP. Different VCEPs may specify alternate conservation tools, cutoffs to use, or species to evaluate for presence of other nucleotides.

In-frame Indel sub-tab

ACMG/AMP criteria codes: PM4, BP3

Background: An indel refers to a small insertion or deletion variant and is in-frame if the altered nucleotide number is divisible by three, keeping the same reading frame. Variants that extend or shorten a protein by deletion, insertion, or altering the stop codon to another amino acid can disrupt protein function. When a variant results in loss of the termination codon (stop-loss variant), the protein is extended; if a variant creates a premature termination codon (nonsense variant), the protein is shortened. It is important to understand

whether the domain the variant alters is functionally important and conserved. There are examples of severe disorders that can result from the loss of a single amino acid, e.g. Coffin-Siris syndrome. Insertions/deletions that occur in repetitive regions are more likely to be of little functional impact; therefore, it is important to assess the surrounding sequence for repetitiveness using a genome browser (linked directly within this tab). It can also help to assess population databases, such gnomAD, for high confidence variant calls that indicate the site is multi-allelic, which could indicate that the region is prone to indels that are generally tolerated, depending on the overall allele frequency.

In-frame Indel sub-tab criteria evaluations:
PM4: Protein length changes as a result of in-frame deletions/insertions in a non-repeat region or stop-loss variant
● VCEPs may define cutoffs to apply PM4 for variants that cause premature truncation but not nonsense-mediated decay (similar to PVS1_moderate). Please discuss with specific VCEP if appropriate.

BP3: In-frame deletions/insertions in a repetitive region without a known function
● VCEPs may define if this rule applies, and if so, to what specific areas of their gene; please discuss with specific VCEP if appropriate.

9.4. EXPERIMENTAL TAB

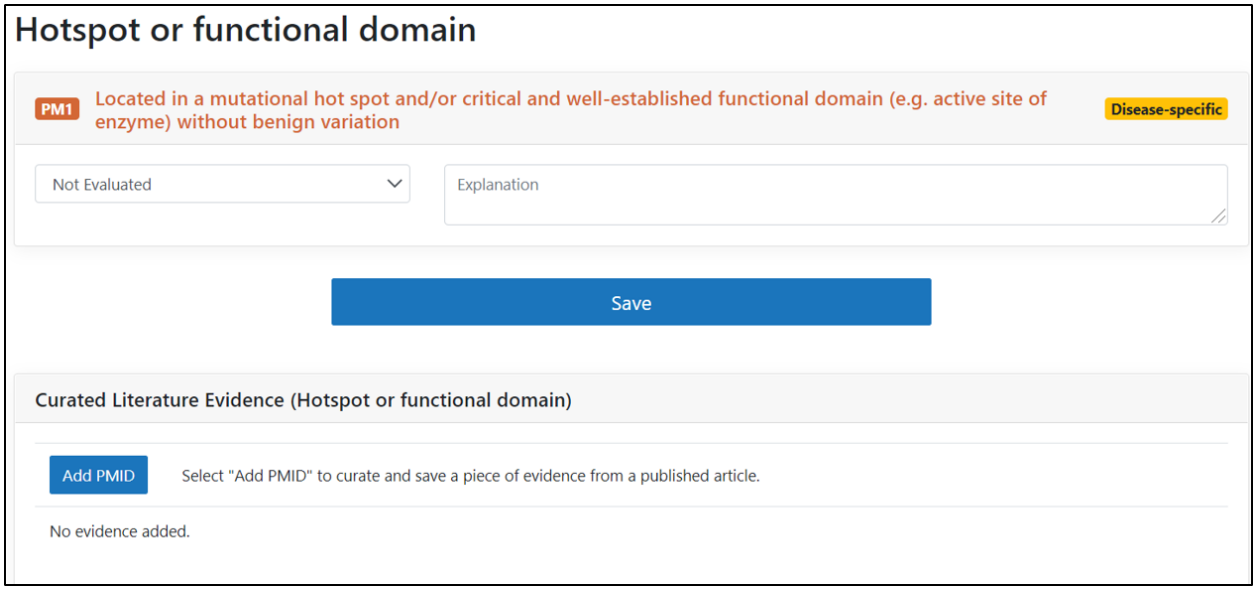
Hotspot or functional domain section
ACMG/AMP criteria codes: PM1

Background: This tab provides a space to evaluate experimental evidence at the variant level (Figure 19). Mutational hotspots and functional domains are disease-specific and evaluated by the appropriate Expert Panel, which should be consulted for guidance.

Limitations:
● Currently, this tab does not automatically pull in data, and curators must manually add the PMID to curate evidence from published articles.

Hotspot or functional domain section Criteria Evaluations:
PM1: Located in a mutational hot spot and/or critical and well-established functional domain (e.g. active site of enzyme) without benign variation.
● Expert Panels determine gene/disease-specific hot spots and functional domains, including specific residues.

Figure 19. Evaluation of experimental evidence.



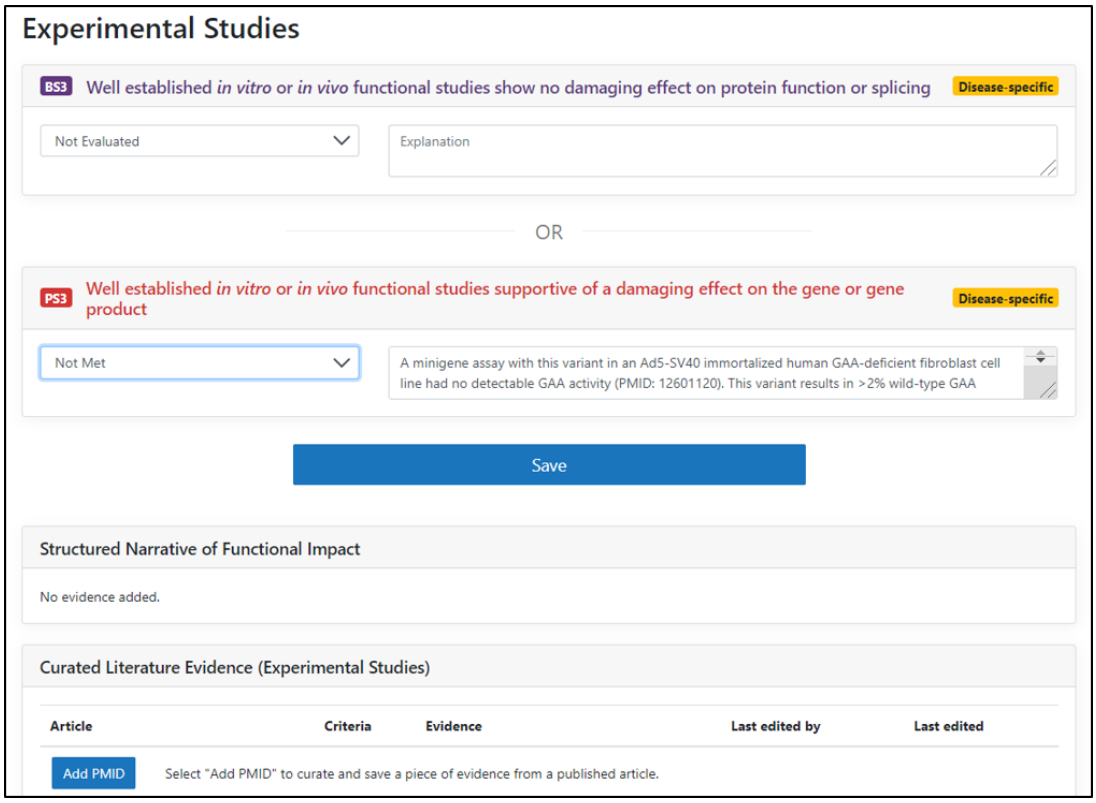
Experimental Studies Section
ACMG/AMP Criteria Codes: PS3, BS3

Background: Functional studies cover a broad range of experimentation. Only validated and approved assays should be evaluated for classification evidence. Appropriate assays are disease-specific and identified by the Expert Panel. Accepted functional assays must meet a high level of rigor and reproducibility as described in the SVI guidance (https://clinicalgenome.org/docs/recommendations-for-application-of-the-functional-evidence-ps3-bs3-criterion-using-the-acmg-amp-sequence-variant-interpretation/)[12]. However, curators may still document all functional data in the free text 'evidence' box, and pull in the appropriate PMID of the article, regardless of whether the evidence code is applied, so that the data is captured for potential utilization at a later date. Curators are encouraged to present results from functional studies to Expert Panel members to decide whether PS3/BS3 applies, particularly if a study has not previously been reviewed by the Expert Panel or if results from functional studies appear discrepant.

Experimental studies section Criteria Evaluations:
BS3: Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing.
PS3: Well established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product.
● In the "Evidence" text box, curators should enter at least the assay and output under evaluation.

● Functional assays should be confirmed by expert review if they are curator identified.
● Functional assays that may not be explicitly specified by VCEPs may still meet criteria for decreased strength modifications for PS3 (i.e. PS3_Moderate or PS3_Supporting). These should be documented and shared with the VCEP before analysis.

Figure 20. Experimental studies section in the VCI



General considerations:

● The existence of a functional assay is not necessarily sufficient for PS3 or BS3 criteria:
  - Exemplary functional assays should demonstrate reproducibility and experimental rigor, including benchmarking with variants of definitive clinical significance as determined by other evidence.
  - The assay should be relevant to disease mechanism and manifestation.
  - Curated evidence should demonstrate variant-level effect on the gene or gene product.

● Applying BS3 criteria should be done with caution, as negative results (equivalent to wild type protein function) may support a benign classification but should be contingent on how thoroughly protein function is assessed. Multiple assays comprehensively probing gene function will more accurately represent clinical significance. VCEPs should have guidelines in place for biocurators to follow regarding the types of studies and results that may be used for BS3 to be applied; discuss with the appropriate VCEP as applicable.

9.5. CASE/SEGREGATION TAB

Assessing the de novo occurrence, frequency and inheritance of a variant is useful in evaluating its pathogenicity. The Case/Segregation tab contains several sections broken down into population data (BS2, PS4), co-segregation data (BS4, PP1), de novo occurrences (PS2, PM6), allelic evidence (BP2, PM3), alternative mechanism of disease (BP5), and specificity of phenotype (PP4) criteria.

Documenting evidence in the Case/Segregation tab
On the case/segregation tab, VCEPs may choose to use evidence from various sources in addition to published data (PMIDs), including a Clinical Laboratory, Clinic, Research Laboratory, Database, or Other source e.g. public database.

Follow these steps to add evidence on the Case/Segregation tab:
1. Select the code for which you want to add evidence. You may do this by clicking the code on the Criteria/Evaluation bar, or simply by scrolling down the fields on the Case/Segregation tab.
2. When you have found the correct code, click on “Select Source”. A drop down menu with a list of evidence sources will appear. Click on the appropriate source (Figure 21). If the data exists in multiple sources, please choose the most detailed source.
3. After selecting the source of evidence from the dropdown, click “Add Evidence”. A box with fields for entering further details on the specific evidence source will appear. When adding free text about evidence sources e.g. the name of a Clinical Lab or Research Lab, you may wish to ensure that all curators in the affiliation are using the same name for the entity for consistency.
4. After choosing the source of evidence, the curator will be presented with a page (referred to as the Case/Segregation evidence page) that includes fields for all of the following ACMG/AMP codes: PP4, PS4, BS2, PP1, BS4, PM6, PS2, PM3, BP2, BP5. The fields highlighted green are the fields that are relevant to the code selected in Step 1. Enter the details from the evidence source in the fields for the appropriate codes (see subsequent sections for further guidance on assessment of each of the codes and data entry). If there is evidence relevant to more than one code (e.g. PP4 and PM3), all of the evidence can be entered at the same time, or you can return later to add more evidence for a specific evidence source. You may also add a Case Label, and check the box for “unaffected” cases, if appropriate (Figure 22). Note that all of the evidence for the variant from the evidence source must be entered on this page e.g. if there are multiple patients with the same variant reported in the same PMID, you will need to include the evidence for all of them in the appropriate fields. Note that ***under no circumstances should any PHI be recorded in the VCI.***

Figure 21. Selecting evidence source on the Case/Segregation tab; Clinical Lab example.

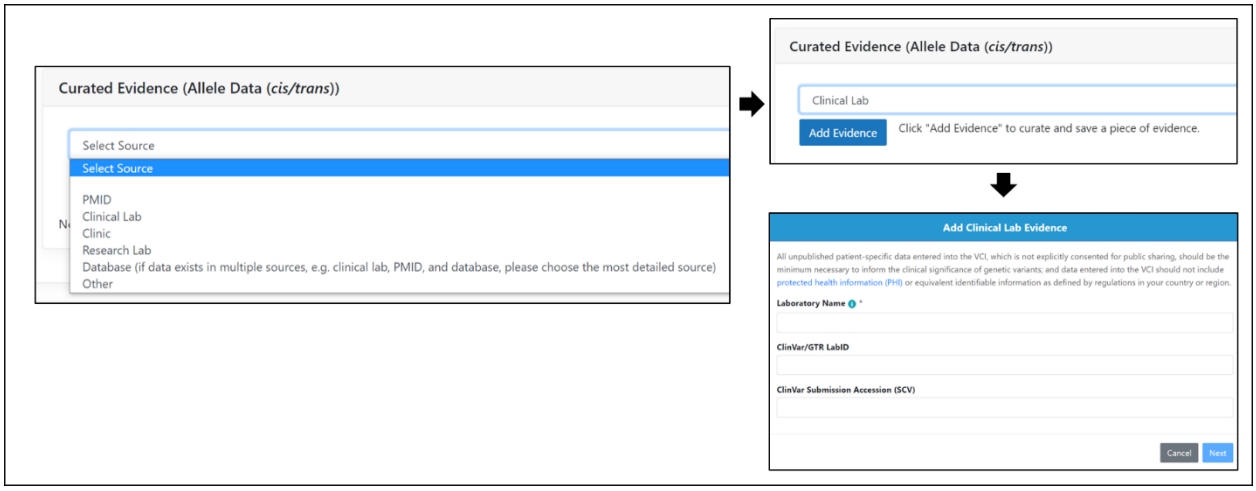
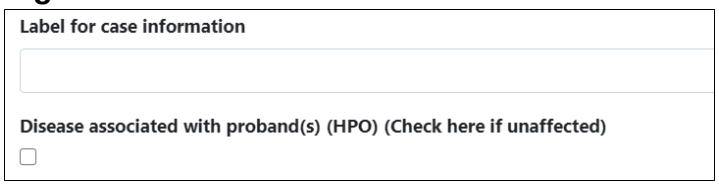
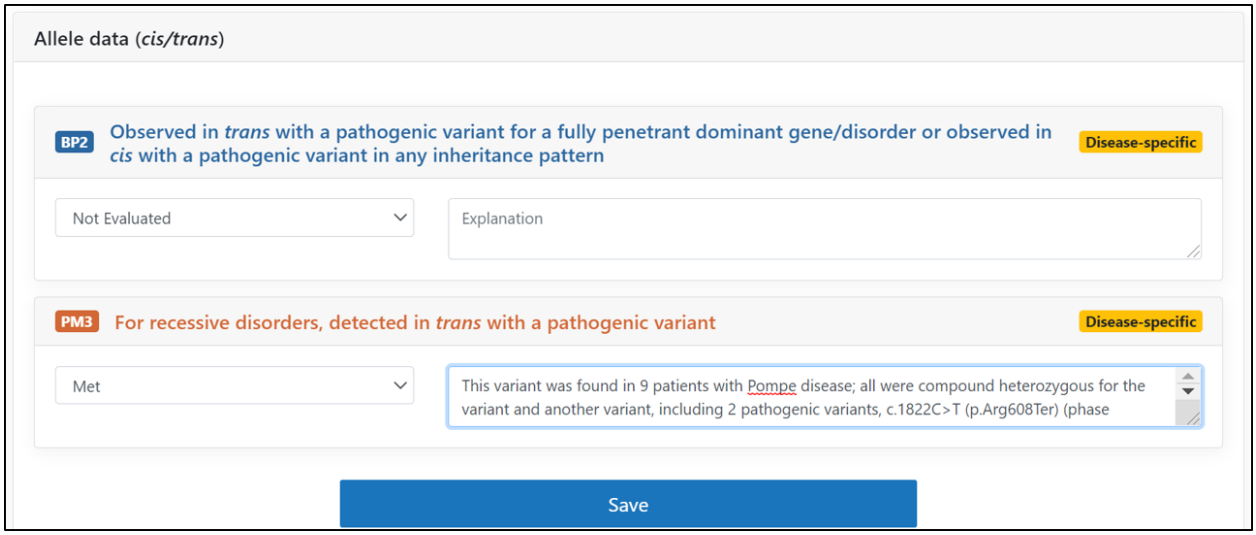


Figure 22. Check box for unaffected cases



5. When you have entered all of the relevant information, click “submit”, at the bottom, right-hand corner of the page.
6. After the data has been saved, you will return to the main Case/Segregation tab. At the top of the tab, you will now see a table which lists all of the evidence sources you have added, the name of the curator who added the evidence, and which codes were applied for the evidence in that source. The “Sum” column will tally the number of probands from different sources of evidence. For each individual code, another table will be presented, showing the details that have been added from each evidence source for that code.
7. The data entered for any source of evidence may be edited by either clicking on the edit or delete icon in the main evidence table at the top of the page, or by clicking the “edit” or “delete” button by the evidence under a specific code.
8. When you have finished adding all of the evidence from various sources, and made an assessment as to the strength of evidence for a particular code, you may enter a summary statement and select the criteria evaluation (e.g. met, not met) on the main Case/Segregation page, and select “save”. For an example, see Figure 23.

Figure 23. Allele data example of criteria evaluation on the main Case/Segregation page



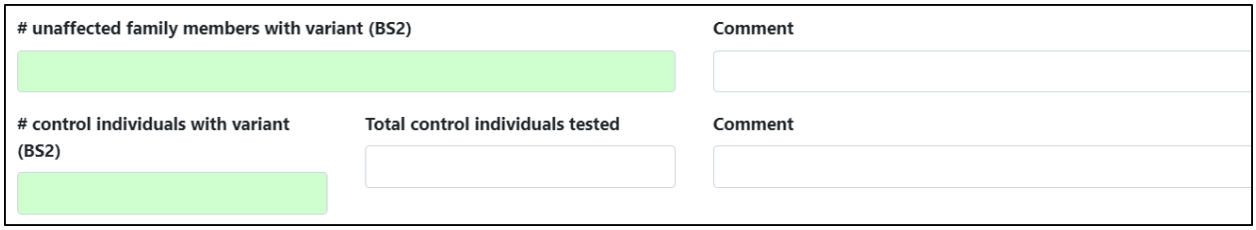
Observed in healthy adults
ACMG/AMP criteria: BS2

Background: This criterion is disease-specific and most applicable if the condition in question is fully penetrant at an early age. Allele frequency can be found in publicly available databases such as gnomAD, or through a thorough literature search. In disease-specific instances, occurrence thresholds may be added; VCEPs may have customized this criterion in alternative ways so should be consulted (if applicable).

Observed in healthy adults criteria evaluations:
BS2: Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age.

Observed in healthy adults (BS2) fields on the Case/Segregation evidence page
After entering the source of information (as described above), complete the BS2 fields (Figure 24) including the number of unaffected family members reported with the variant in this evidence source, the number of control individuals with the variant and any comments, then click “save”. Note that gnomAD does provide the age of individuals with a given variant in the database. For some adult-onset disorders, the VCEP may set an age cut-off to use this criteria.

Figure 24. BS2 fields on the Case/Segregation evidence page



Case-control section
ACMG/AMP criteria: PS4

Note: Although this section has been titled 'Case-control' in the VCI, due to specifications by many VCEPs, this rule is more often used for proband counting.

Background: PS4 can be used for typical case-control studies when a relative risk or odds ratio (OR) > 5.0, and confidence interval not including 1.0, has been calculated to assess whether a variant is likely to be associated with a particular phenotype. In instances of very rare variants where case-control studies are not available, the prior observation of the variant in multiple unrelated patients with the same phenotype (and its absence in population databases) may be used as evidence. The number of observations in probands can be used to determine the application of PS4 or any of its modified strengths (see Table 4 for ClinGen VCEP examples). Number of probands can be gathered from the various evidence sources listed in the VCI including a thorough literature search for PMIDs, through internal laboratory data, and publicly available databases.

Limitations:
● Access to internal data can impact rule application and/or strength.
● The VCI does not currently support the import of spreadsheets containing de-identified internal data such as proband count. Each entry must be logged independently into the "Explanation" free text box; proband count for strength level must be calculated by hand.
● Caution should be exercised to not duplicate counting of one proband reported multiple times in the literature or within both the literature and a clinical testing cohort.

Case control criteria evaluations:
PS4: The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls.

Table 4. Examples of PS4 specifications from different ClinGen VCEPs.

	PS4_Supporting	PS4_Moderate	PS4 (Strong)	PM2 threshold
MYH7 Hypertrophic cardiomyopathy	≥2 probands with consistent phenotypes	≥6 probands with consistent phenotypes	≥15 probands with consistent phenotypes	<0.004%
RASopathy	≥1 probands that meets RASopathy criteria*	≥3 probands that meet RASopathy criteria	≥5 probands that meet RASopathy criteria	Strictly absent
PTEN PTEN Hamartoma Tumor syndrome (PHTS)	>1 proband meets Cleveland Clinic criteria (30 score)**	>2 probands meets Cleveland Clinic criteria (30 score)	>4 probands meets Cleveland Clinic criteria (30 score)	<0.001%
CDH1 Hereditary diffuse gastric cancer (HDGC)	>1 HDGC family meets criteria***	>2 HDGC family meets criteria	>4 HDGC family meets criteria	<0.001%

Table 4 notes - For a more detailed explanation, see:
*https://www.clinicalgenome.org/site/assets/files/8852/clingen_rasopathy_acmg_specifications_v1-1.pdf
** https://www.clinicalgenome.org/working-groups/clinical-domain/hereditary-cancer-clinicaldomain-working-group/pten-variant-curation-expert-panel/announcements/summary-of-acmgamp-classification-rules-specified-for-pten-variant-curation/
*** https://www.clinicalgenome.org/working-groups/clinical-domain/hereditary-cancer-clinicaldomain-working-group/cdh1-variant-curation-expert-panel/

Case-control (PS4) fields on the Case/Segregation evidence page

After entering the source of information, complete the PS4 fields (Figure 25) including the number of probands with the relevant phenotype and any comments, then click “save”.

Figure 25. PS4 fields on the Case/Segregation evidence page

Segregation data section
ACMG/AMP criteria: PP1, BS4

Background: Co-segregation data can be used as evidence for or against variant pathogenicity. The more frequently a variant co-segregates with a disease phenotype, the more likely the variant and the disease locus are linked. This can be determined using a LOD (log of the odds) score, which is a statistical estimate of the proximity of two loci and the likelihood that they will be inherited together. The higher the LOD score, the more likely the loci are linked. For dominant disorders, co-segregation of a variant in affected family members is counted as supportive pathogenic evidence. Conversely, when a variant is not found in an affected family member, this is considered good evidence that the variant is benign. In either case, several issues must be thoroughly investigated. Other variants in linkage disequilibrium with the linked variant must be considered when applying PP1, and thorough phenotypic evaluation of affected non-segregating family members must be done for BS4 to rule out phenocopies. Correct application of these criteria are hampered by issues of penetrance, expressivity, and age of onset of the disorder. Different conditions apply if diagnostic clarity or phenocopy are a concern.

Segregation data criteria evaluations:
PP1: Co-segregation with disease in multiple affected family members in a gene definitively known to cause disease. VCEPs may specify higher levels of evidence to be applied for higher numbers of meioses/families demonstrating segregation with disease (see Table 5 for VCEP examples).
● Please note that only affected individuals are counted as a part of this criterion. Unaffected individuals may complicate the picture due to incomplete or age/gender-related penetrance, variable expressivity, or phenocopies.

BS4: Lack of segregation in affected members of family
● Please note, BS4 non-segregation is meant for affected family members who do not carry the variant of interest (phenotype positive; genotype negative). Individuals who carry the variant but do not have the phenotype (genotype positive; phenotype negative) may be counted under BS2, if applicable for the gene/disease of interest with agreement from the Expert Panel.

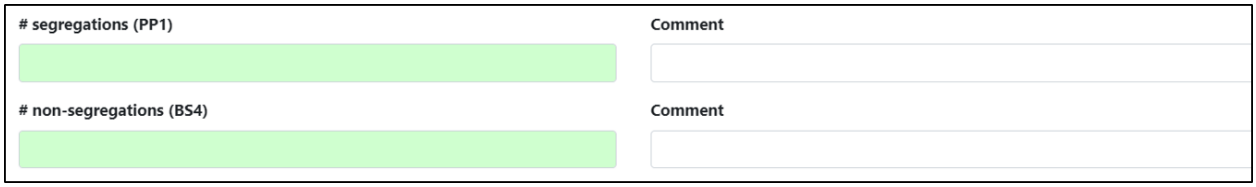
Table 5. Segregation counts from approved VCEPs.

VCEP	PP1 segregation counts
MYH7	PP1_Strong: variant segregate with ≥7 meioses PP1_Moderate: variant segregates with ≥5 meioses PP1: variant segregates with ≥3 meioses
RASopathy	Segregation in more than one family is recommended PP1_Strong: ≥7 informative meioses PP1_Moderate: ≥5 informative meioses PP1: ≥3 informative meioses
PTEN	PP1_Strong: variant segregates with ≥7 meioses observed across ≥ 2 families PP1_Moderate: variant segregates with 5-6 meioses across ≥ 1 family PP1: variant segregates with 3-4 meioses across ≥ 1 family
CDH1	PP1_Strong: variant segregates with ≥7 meioses across ≥ 2 families PP1_Moderate: variant segregates with 5-6 meioses across ≥ 1 family PP1: variant segregates with 3-4 meioses across ≥ 1 family
Hearing Loss	PP1_Strong: Segregation in 3 affected relatives for recessive and 5 affected relatives for dominant PP1_Moderate: Segregation in 2 affected relatives for recessive and 4 affected relatives for dominant PP1: Segregation in 1 affected relative for recessive and 2 affected relatives for dominant

Segregation (PP1 and BS4) fields on the Case/Segregation evidence page

After entering the source of information, complete the PP1 or BS4 fields (Figure 26) including the number of segregations (for PP1), number of non-segregations (for BS4), any comments, and then click “save”, bottom right.

Figure 26. PP1 and BS4 fields in the Case/Segregation evidence Page



De novo occurrence section

ACMG/AMP criteria: PS2, PM6

Background: There are different levels of strength given to de novo data. To be considered as strong evidence, parents must be tested and both maternity and paternity must be confirmed (using techniques such as identity panels or NGS trio analysis), or there must be multiple observations of de novo occurrence, the variant must be in a gene that is associated

with a condition consistent with the phenotype of the patient, and there must be no family history of the disease. The evidence is considered moderate if the second two criteria are met but maternity and paternity have both not been confirmed via identity/parentage testing. Exome trio testing may be accepted as capable of confirming maternity and paternity.

De novo data criteria evaluations:
PS2: De novo (both maternity and paternity confirmed) in a patient with the disease and no family history.
● See Table 6.

PM6: Assumed de novo, but without confirmation of paternity and maternity.
● See Table 6.

Using parental confirmation, phenotypic consistency, and number of de novo observations, the SVI has proposed a point-based system for modified strength levels of these criteria as shown below in Table 6. (https://www.clinicalgenome.org/site/assets/files/3461/svi_proposal_for_de_novo_criteria_v1_1.pdf).

The SVI also provides additional considerations for applying de novo criteria based on inheritance:
● X-linked conditions: if an X-linked variant occurs de novo in an unaffected carrier mother, and family history is consistent - i.e. she has no affected brothers/other male relatives apart from her affected son(s) – de novo criteria may be applied despite the fact that she is unaffected.
● Autosomal recessive conditions: for a de novo occurrence in a gene associated with an autosomal recessive condition without an additional pathogenic/likely pathogenic variant identified, the strength of evidence should be decreased by one level.
● Mosaicism: for cases with apparent germline mosaicism (multiple affected siblings with both parents negative for the variant), paternity/maternity must be confirmed in order for de novo criteria to apply.

Limitations:
● The de novo guidelines PM6 and PS2 are mutually exclusive; only one of the two criteria is allowed to be met, and the higher strength prevails.
● Phenotype must be consistent with the gene. Further phenotypic specifications may be applied by an Expert Panel for criteria application.

Table 6. Points awarded per de novo occurrence
*Maximum allowable value of 1 may contribute to overall score

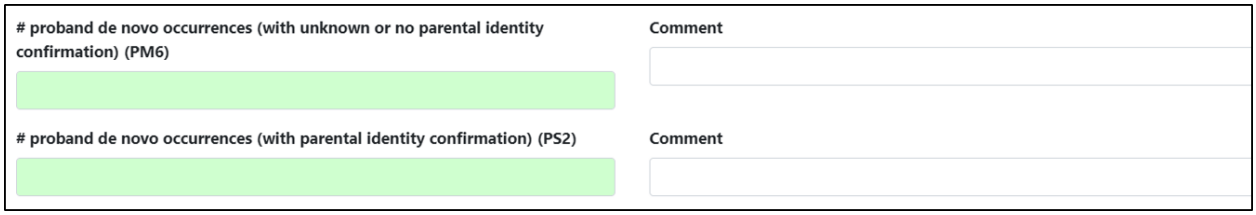
Phenotypic consistency	Points per Proband	
	Confirmed *de novo*	Assumed *de novo*
Phenotype highly specific for gene	2	1
Phenotype consistent with gene but not highly specific	1	0.5
Phenotype consistent with gene but not highly specific and high genetic heterogeneity*	0.5	0.25
Phenotype not consistent with gene	0	0

Recommendation for determining the appropriate ACMG/AMP evidence strength level for de novo occurrence(s)

Supporting (PS2_Supporting or PM6_Supporting)	**Moderate** (PS2_Moderate or PM6)	**Strong** (PS2 or PM6_Strong)	**Very Strong** (PS2_VeryStrong or PM6_VeryStrong)
0.5	1	2	4

De novo (PS2, PM6) fields on the Case/Segregation evidence page
After entering the source of information, complete the relevant fields (Figure 27) including the number of de novo occurrences, then click “save”.

Figure 27. PS2 and PM6 fields on the Case/Segregation evidence page



Allele data (cis/trans) section
ACMG/AMP criteria: BP2, PM3

Background: Phase of the variant is important in assessing its effect. When two variants are found on a single chromosome in the same gene, they are considered to be in cis; variants identified in a gene but on different chromosomes are in trans. Inheritance of a condition is key for applying these rules. For autosomal recessive conditions, PM3 is a method of counting probands (Table 7). Use of these rules for autosomal dominant disorders is dependent upon the effect of homozygosity on the phenotype. For example, two pathogenic variants in trans for an AD disorder could result in embryonic lethality or a severe disease phenotype, in which case BP2 could be applicable (the variant would have to be benign by definition). This is a gene-dependent (and therefore an Expert Panel-level) decision.

Allele data criteria evaluations:
PM3: For recessive disorders, detected in trans with a pathogenic variant
● For recessive disorders, the strength of PM3 may be altered under certain conditions, as recommended by the SVI (see Table 7).

BP2: Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern
● In order to apply BP2, the pathogenic variant must have been observed independently of the variant being assessed. The criteria should not be applied if there is evidence of a synergistic effect.

Table 7. SVI Recommended Scoring, Allele Data.

Classification/Zygosity of other variant [1]	**Points per Proband**	
	Confirmed in *trans*	**Phase unknown**
Pathogenic or Likely pathogenic variant	1.0	0.5 (P) 0.25 (LP)
Homozygous occurrence *(max point 1.0)*	0.5	N/A
Uncertain significance variant *(max point 0.5)*	0.25	0.0

[1]All variants should be sufficiently rare (meet PM2 specification); P - Pathogenic; LP - Likely pathogenic

PM3 Point Table			
PM3_Supporting	**PM3**	**PM3_Strong**	**PM3_VeryStrong**
0.5 points	**1.0 points**	**2.0 points**	**4.0 points**
Examples: - 1 observation w/ P/LP variant but phase unknown - 1 homozygous occurrence - 2 observations of either homozygous occurrence due to consanguinity ***OR*** rare VUS	Examples: - 1 observation w/ P/LP variant confirmed in *trans* - 2 homozygous occurrences - 2 observations w/ P/LP variant but phase unknown		

Allele data (cis/trans) (BP2, PM3) fields on the Case/Segregation evidence page

After entering the source of information, complete the relevant fields (Figure 28) for PM3 or BP2. Note that there are two options for PM3; homozygous occurrences and heterozygous occurrences. Record the number of cases that you wish to count and add any relevant comments. For example, in the compound heterozygotes occurrences field, include details of the other variant, whether the variant were confirmed to be in trans (e.g. by parental testing), and how many points you wish to assign using Table 7, or your VCEP's specifications (if modified from Table 7 for any reason), as a guide. Currently, the VCI adds the number of occurrences but not the number of points and so this must be done by hand and then summarized in the PM3 field on the main Case/Segregation page. After data entry is complete, click "save".

Figure 28. BP2 and PM3 fields on the Case/Segregation evidence page

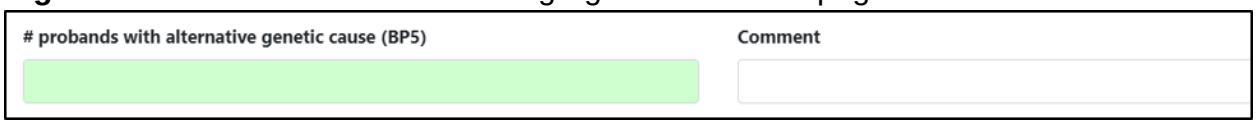
Alternate mechanism for disease section
ACMG/AMP criteria: BP5

Background: If a variant is found in an affected individual with an obvious alternate cause of disease, this is supportive of a benign classification; however, there are exceptions.

Alternate mechanism for disease criteria evaluations:
BP5: Variant found in a case with an alternate molecular basis for disease
● Stronger for supporting benign classification in a gene for a dominant disorder as compared to a gene for a recessive disorder. Caution should be used when applying for a recessive disorder as the individual may be a carrier. For example, multiple genes can cause nonsyndromic hearing loss and a diagnosed individual may be a carrier for a pathogenic variant in one gene but harbor two pathogenic variants in trans in another gene causative for their disease.
● In some disorders, having multiple variants can contribute to more severe disease. If a novel variant is seen in a patient with a dominant disorder and a family history of a more mild presentation, this would not be considered supportive of a benign classification for the novel variant.
● In conditions where multigenic inheritance is known to occur, additional variants at a second locus may also be pathogenic. Some EPs may qualify the necessity of multiple observations or specify phenotypic considerations in order for this rule to apply.

Alternate mechanism for disease (BP5) fields on the Case/Segregation evidence page
After entering the source of information, complete the relevant fields (Figure 29) for BP5, then click save at the bottom right.

Figure 29. BP5 fields on the Case/Segregation evidence page



Specificity of phenotype section
ACMG/AMP criteria: PP4

Background: This criterion is only appropriate for application when a phenotype is highly specific to a single gene. For example, all known cases of Pompe disease are caused by variants in the GAA gene, have a specific phenotype, and have a deficiency of activity of the GAA gene product, acid alpha glucosidase. Family history should be consistent with mode of inheritance, testing should be highly sensitive, and the gene should have little benign variation. The PP4 criterion may also be applied if non-genetic confirmatory assays

are available, such as enzyme levels for biochemical disorders. For example, the Phenylalanine Hydroxylase Variant Curation Expert Panel (PAH VCEP) applies PP4 if a case has diagnostic plasma phenylalanine levels >120 umol/L. However, this criterion should not be modified in weight due to multiple observations of the variant. This criterion may not be used for commonly encountered, non-specific phenotypes such as developmental delay or cardiomyopathy. These disorders have multiple genetic causes, and phenotype cannot be distinguished by the gene in which the pathogenic variant is located. Of note, several Expert Panels have wrapped PP4 into other evidence types (PS4, PS2/PM6, PP1) as opposed to having a separate evidence code.

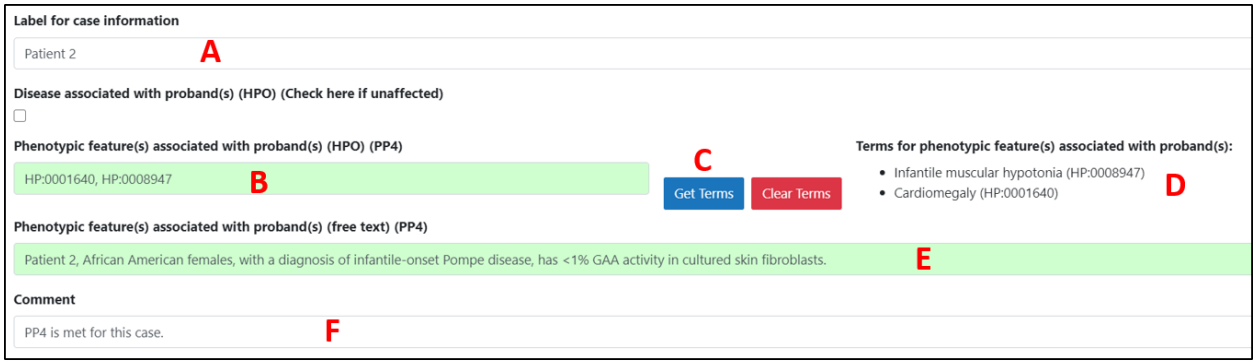
Specificity of phenotype criteria evaluations:

PP4: Patient's phenotype or family history is highly specific for a disease with a single genetic etiology

Specificity of phenotype criteria evaluations (PP4) fields on the Case/Segregation evidence page

The VCI allows curators to add HPO terms for clinical symptoms. To do so, go to https://hpo.jax.org/app/browse to obtain HPO terms. Copy and paste HPO terms into the phenotypic features field (Figure 30, B). Clicking "Get terms" (Figure 30, C) will convert the HPO terms into text, which will be displayed on the right hand side (Figure 30, D). Free text can also be added to describe the patient(s) (Figure 30, E), and additional comments can be included (Figure 30, F). A label for the patient can also be included (Figure 30, A).

Figure 30. PP4 fields on the Case/Segregation evidence page



Reputable source section
ACMG/AMP criteria: BP6, PP5

ClinGen has determined that these rules should not be applied in any context[13]. While they are part of the original 2015 ACMG/AMP Variant Classification guidelines, the SVI is of the opinion that evidence collected for variant interpretation should be limited to primary data, which is becoming more widely available through resources such as ClinVar, and that BP6 and PP5 are commonly misused 10. These criteria have been disabled in the VCI.

9.6. GENE-CENTRIC TAB

The gene-centric tab contains general information about the gene for use with the various guidelines. This includes the ExAC constraint scores as described in the LOF section, an 'other ClinVar variants in same gene' section with a link-out to ClinVar, and multiple link-outs to gene and protein resources including HGNC, Entrez Gene, Ensembl, GeneCards, UniProtKB, InterPro, Protein Data Bank in Europe (PDBe), and Gene Ontologies.

10. EVALUATION SUMMARY

Once all evidence has been curated into the VCI, an evaluation summary can be viewed by clicking the "View Summary" button (see Figure 5, D). Once in the Evaluation Summary, the curator can click "Interpretation" to return to the Interpretation View.

The Evaluation Summary includes the calculated pathogenicity, which is derived from combining the applied ACMG/AMP guidelines (see key found in Appendix III), some of which may have been modified for some VCEPs during their specification processes. At this point, the curator can modify the classification, provided that a reason for the change is documented and the change is consistent with the VCEP's approved methods of variant classification.

The curator will also be responsible for providing a written evidence summary that contains all evidence used to arrive at the classification, including any evidence sources that have been used, as well as all criteria that met an evaluation strength. This summary will be publicly available in the ERepo and in ClinVar.
Please see the "Standardized text for ClinGen VCEPs" document (https://clinicalgenome.org/docs/standardized-text-for-clingen-variant-curation-expert-panels/) for guidance on drafting written summaries.

The first time an Evaluation Summary is saved the status of the interpretation will change to "In Progress". The Interpretation will remain "In Progress" until a Summary is saved as

Provisional. Provisional is used by curators to indicate that their Interpretation is complete and ready for review by experts in the VCEP. Every time a curator updates evidence and/or evaluations they will need to re-click the Save button if they wish to make a Provisional Interpretation based on their new Summary. Different VCEPs may have different work-flows for expert review and approval. Curators should discuss the workflow for the VCEP with the VCEP coordinator. For more details on evaluation summaries and saving provisional and approved classifications see the VCI Help Document (https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help).

11. CLINVAR SUBMISSION

Typically, each VCEP has a designated individual who is responsible for submitting data to ClinVar. Curators can talk with the VCEP coordinator to determine who has this role. For any individuals who have questions about the ClinVar submission process, please contact the Variant Curation Working Group at variantcuration@clinicalgenome.org

References

1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015. 17(5):405-24.

2. Harrison SM, Riggs ER, Maglott DR, Lee JM, Azzariti DR, Niehaus A, Ramos EM, Martin CL, Landrum MJ, Rehm HL. Using ClinVar as a Resource to Support Variant Interpretation. Curr Protoc Hum Genet. 2016. 89:8.16.1–8.16.23.

3. DiStefano MT, Hemphill SE, Cushman BJ, Bowser MJ, Hynes E, Grant AR, Siegert RK, Oza AM, Gonzalez MA, Amr SS, Rehm HL, Abou Tayoun AN. Curating Clinically Relevant Transcripts for the Interpretation of Sequence Variants. J Mol Diagn. 2018. 20(6):789- 801.

4. Ghosh R, Harrison SM, Rehm HL, Plon SE, Biesecker LG; ClinGen Sequence Variant Interpretation Working Group. Updated recommendation for the benign stand-alone ACMG/AMP criterion. Hum Mutat. 2018. 39(11):1525-1530.

5. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016. 99(4):877-885.

6. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, Chow ED, Kanterakis E, Gao H, Kia A, Batzoglou S, Sanders SJ, Farh KK. Predicting Splicing from Primary Sequence with Deep Learning. Cell. 2019. 176(3):535-548.e24.

7. Harrison SM, Biesecker, LG, Rehm HL. Overview of Specifications to the ACMG/AMP Variant Interpretation Guidelines. Curr Protoc Hum Genet. 2019. 103(1):e93.

8. Abou Tayoun AN, Pesaran T, DiStefano MT, Oza A, Rehm HL, Biesecker LG, Harrison SM; ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI). Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat. 2018. 39(11):1517-1524.

9. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, et al; Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016. 536(7616):285-91.

10. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020. 581(7809):434-443.

11. Riggs ER, Church DM, Hanson K, Horner VL, Kaminsky EB, Kuhn RM, Wain KE, Williams ES, Aradhya S, Kearney HM, Ledbetter DH, South ST, Thorland EC, Martin CL. Towards an evidence-based process for the clinical interpretation of copy number variation. Clin Genet. 2012. 81(5):403-12.

12. Brnich SE, Abou Tayoun AN, Couch FJ, Cutting GR, Greenblatt MS, Heinen CD, Kanavy DM, Luo X, McNulty SM, Starita LM, Tavtigian SV, Wright MW, Harrison SM, Biesecker LG, Berg JS; Clinical Genome Resource Sequence Variant Interpretation Working Group. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. Genome Med. 2019. 12(1):3.

13. Biesecker LG, Harrison SM. The ACMG/AMP reputable source criteria for the interpretation of sequence variants. Genet Med. 2018. 20(12):1687-1688.

Resources

ClinGen Resources

ClinGen website: https://www.clinicalgenome.org/
SVI webpage: https://www.clinicalgenome.org/working-groups/sequence-variant-interpretation
CDWG webpage: https://www.clinicalgenome.org/working-groups/clinical-domain/
Variant Curation Interface: https://curation.clinicalgenome.org/
VCI Help document: https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help
ClinGen Allele Registry:
http://reg.clinicalgenome.org/redmine/projects/registry/genboree_registry/landing
ClinGen Evidence Repository: https://erepo.clinicalgenome.org/evrepo/

Other useful resources

ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/
OMIM: https://www.omim.org
HGMD: http://www.hgmd.cf.ac.uk/ac/index.php
Decipher: https://decipher.sanger.ac.uk/
The Monarch Initiative: https://monarchinitiative.org/
Genetic Testing Registry: https://www.ncbi.nlm.nih.gov/gtr/

Sequence databases

RefSeq: https://www.ncbi.nlm.nih.gov/refseq/
Ensembl: http://uswest.ensembl.org/index.html
UCSC genome browser: https://genome.ucsc.edu/cgi-bin/hgGateway
Variation Viewer: https://www.ncbi.nlm.nih.gov/variation/view/overview/

Population databases

Genome Aggregation Database (gnomAD): http://gnomad.broadinstitute.org/
Population Architecture using Genomics and Epidemiology (PAGE):
https://www.genome.gov/27541456/population-architecture-using-genomics-and-epidemiology/
1000 Genomes (1000G): http://grch37.ensembl.org/index.html
Exome Sequencing Project (ESP): http://evs.gs.washington.edu/EVS/
dbSNP: https://www.ncbi.nlm.nih.gov/projects/SNP/

Literature sources

PubMed: https://www.ncbi.nlm.nih.gov/pubmed/
Google Scholar: https://scholar.google.com/
bioRxiv: https://www.biorxiv.org/content/early/2017/06/12/148353
LitVar: https://www.ncbi.nlm.nih.gov/CBBresearch/Lu/Demo/LitVar/

Appendix I - Computational Tools

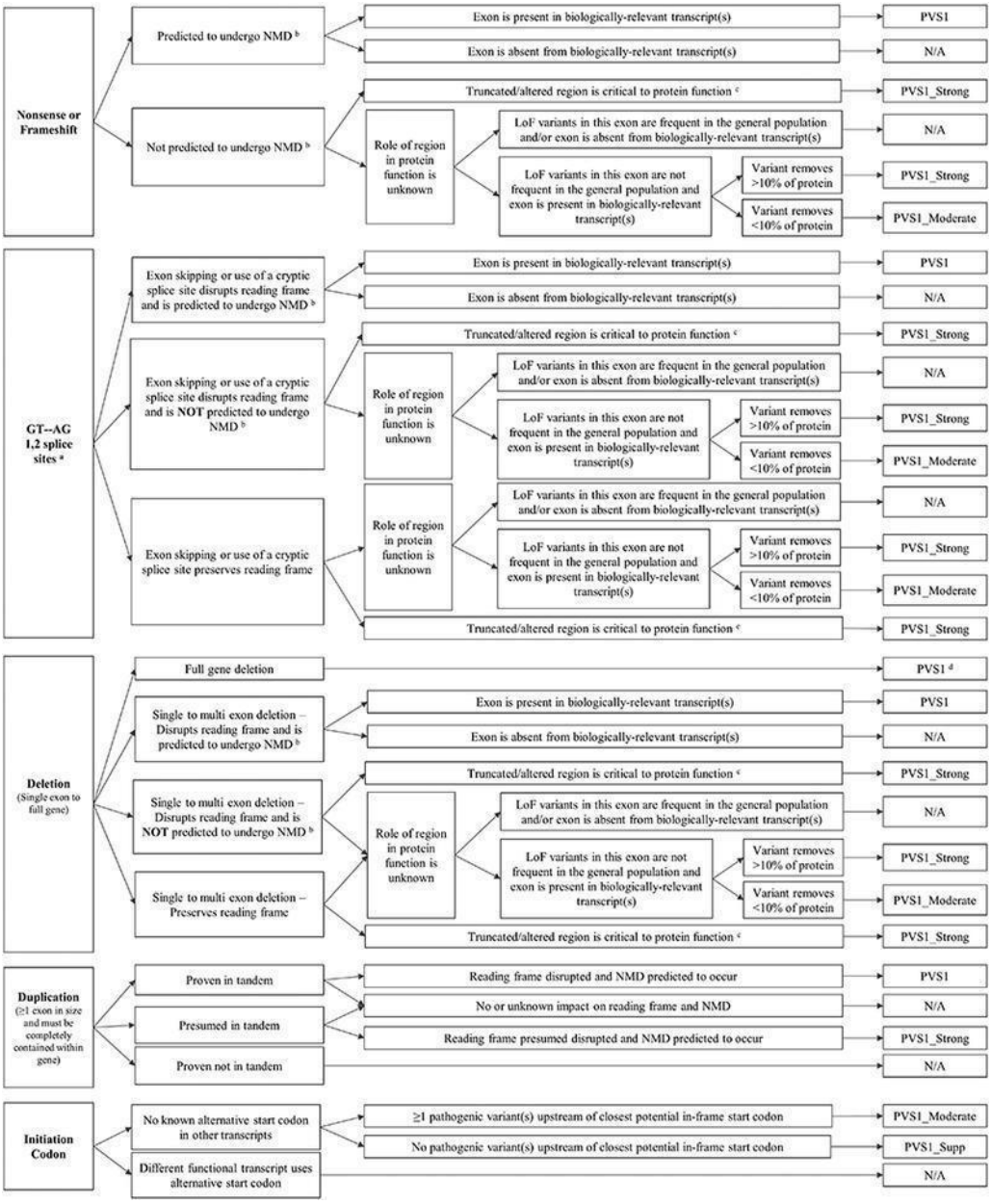
In silico predictors linked in VCI In silico tools are used to predict the effect of a variant on a particular gene. The scores for some of these tools are calculated directly in the VCI, and ranges, as well as prediction meaning are described. For further information, descriptions, and tutorials, please see each individual site.

REVEL: https://sites.google.com/site/revelgenomics/about
SIFT: http://sift.bii.a-star.edu.sg/
PolyPhen2-HDIV/HVAR: http://genetics.bwh.harvard.edu/pph2/
LRT: http://www.genetics.wustl.edu/jflab/lrt_query.html
MutationTaster: http://mutationtaster.org
MutationAssessor: http://mutationassessor.org/r3/
FATHMM: http://fathmm.biocompute.org.uk/
PROVEAN: http://provean.jcvi.org/index.php
MetaSVM, MetaLR: https://sites.google.com/site/jpopgen/dbNSFP
CADD: https://cadd.gs.washington.edu/
FATHMM-MKL: http://fathmm.biocompute.org.uk/fathmmMKL.htm
fitCons: http://compgen.cshl.edu/fitCons/

Conservation analysis tools linked in VCI
phyloP100way, phyloP20way: http://compgen.cshl.edu/phast/index.php
phastCons100way, phastCons20way: http://compgen.cshl.edu/phast/index.php
GERP++: http://mendel.stanford.edu/SidowLab/downloads/gerp/
SiPhy: http://portals.broadinstitute.org/genome_bio/siphy/index.html

Splice site predictors linked in VCI
MaxEntScan: http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html
NNSPLICE: http://www.fruitfly.org/seq_tools/splice.html
varSEAK: https://varseak.bio/
SpliceAI: https://spliceailookup.broadinstitute.org/

Appendix II - PVS1 flowchart, SVI recommendations



For full article, see PMID: 30192042

Appendix III - Combining Criteria to Classify Sequence Variants

Pathogenic

1. 1 Very Strong AND:
    a. ≥ 1 strong, OR
    b. ≥ 2 moderate, OR
    c. 1 moderate and 1 supporting, OR
    d. ≥ 2 supporting
2. ≥ 2 strong
3. 1 strong AND:
    a. ≥ 3 moderate, OR
    b. 2 moderate AND ≥ 2 supporting, OR
    c. 1 moderate AND ≥ 4 supporting

Likely Pathogenic

1. 1 very strong AND 1 moderate, OR
2. 1 very strong AND 1 supporting*, OR
3. 1 strong AND 1-2 moderate, OR
4. 1 strong AND ≥ 2 supporting, OR
5. ≥ 3 moderate, OR
6. 2 moderate AND ≥ 2 supporting, OR
7. 1 moderate AND ≥ 4 supporting

For example, variants meeting PVS1 and PM2_Supporting may be classified as Likely Pathogenic. For further guidance, see: https://clinicalgenome.org/site/assets/files/5182/pm2_-_svi_recommendation_-_approved_sept2020.pdf

Benign

1. 1 stand alone, OR
2. ≥ 2 strong

Likely Benign

1. 1 strong AND 1 supporting, OR
2. ≥ 2 supporting

Uncertain Significance

1. Other criteria shown above are not met, OR
2. The criteria for benign and pathogenic are contradictory

For full article, see PMID: 25741868