What is variant curation?

To determine whether a genetic variant is pathogenic, likely pathogenic, uncertain significance (VUS), likely benign, or benign, genetics professionals **perform an assessment of available data** about the variant of interest. Upon reviewing the evidence, they **apply assertion criteria** to ultimately **determine the pathogenicity** of the variant.

Benign	Likely Benign	Variant of Uncertain Significance (VUS)	Likely Pathogenic	Pathogenic
--------	------------------	--	----------------------	------------

To curate small sequence changes, genetics providers and professionals generally utilize criteria outlined in joint guidelines from the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP), which were published in 2015. These criteria include weighted codes, which have various strengths ranging from Supporting to Very Strong. Each code represents a type of evidence that can be applied. Over time, modifications have been applied to these guidelines, and laboratories also can vary in how they apply criteria from these guidelines. Additionally, for some specific genes or diseases, there are ClinGen Variant Curation Expert Panels (VCEPs) which have more refined criteria to determine pathogenicity of variants.

What are the main types of evidence?



Population Frequency

How frequently is the variant seen in healthy populations?



Computational Data

Do in silico tools predict that the variant has an impact on the protein?



Predictive Data

Does the variant affect important domains or cause loss of protein function?



Functional Data*

Do functional assays show a pathogenic effect on the protein?



Proband and Family Data

Has the variant been found in patients with the associated disease and in their family members?

Mass General Brigham VARIANT CURATION TASK FORCE

Each of these types of evidence have specific codes that are assigned. Codes can be applied at **four different levels: Supporting (P), Moderate (M), Strong (S), Very Strong (VS)**. Below, you can find commonly used codes that are associated with each of these pieces of evidence:

	Population Frequency	Computational Data	Predictive Data	Functional Data	Proband and Family Data
Pathogenic	PM2_P	PP3	PVS1, PS1, PM5, PM4	PS3	PS4 (Dominant), PM3 (Recessive), PP1 (Segregations), PM6 (de novo)
Benign	BS1, BA1	BP4	BP7, BP1	BS3	BP2, BP5, BS4

* Functional Data will not be detailed in this handout

How are these types of evidence used in variant curation?

Certain types of evidence provide more support for or against pathogenicity than others. Furthermore, a combination of several types of evidence is required to prove whether a variant is pathogenic. In general:

Examples of Strong Evidence

- If a variant is **found in many patients and family members** with a specific health condition, this can provide up to a strong line of evidence that the variant may be pathogenic.
- ✓ If a variant occurs more frequently in population databases than the prevalence of disease, this is a very strong indicator that the variant may be benign.
- If a variant occurs in a gene for which loss of protein function has been established in literature and by experts as a mechanism of disease, and the variant introduces an early stop codon or causes a frameshift, this may provide strong evidence that the variant is pathogenic.

✓ If an animal model of the variant exhibits a similar phenotype to the associated disease, this provides strong evidence that the variant may be pathogenic. If not, it may provide strong evidence that the variant may be benign.

Examples of Moderate Evidence

- If a few amino acids are predicted to be removed in-frame, this provides a moderate line of evidence that the variant may be pathogenic.
- ✓ Some functional studies provide up to a moderate line of evidence that the variant may be pathogenic.

Examples of Supporting Evidence

- ✓ If computational tools predict that the variant causes/does not cause an impact to the protein, this provides a lower level of support that a variant may be pathogenic/benign.
- ✓ If a variant is absent from population databases, this provides a lower level of support that a variant may be pathogenic.

How much evidence is needed for a variant to be considered Likely Pathogenic or Pathogenic?

To reach **Likely Pathogenic**, a variant must have:

- ✓ 1 Moderate AND ≥4 Supporting
- ✓ 2 Moderate AND ≥2 Supporting
- ✓ 1 Strong AND 1-2 Moderate
- ✓ 1 Very Strong AND 1 Moderate

To reach **Pathogenic**: a variant must have:

- ✓ 1 Strong and either: ≥3 Moderate, 2 Moderate and ≥2 Supporting, 1 Moderate and ≥4 Supporting
- ✓ ≥2 Strong
- ✓ 1 Very Strong and either: ≥1 Strong, ≥2 Moderate, 1 Moderate and 1 Supporting, ≥2 Supporting

Where can I learn more about variant curation guidelines?

- 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 (PMID: 25741868): <u>https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4544753/</u>
 - a. These are the general guidelines for interpretation of small sequence changes.
- 2. Criteria Specification Registry: https://cspec.genome.network/cspec/ui/svi/
 - a. Guidelines for specific genes or diseases, as developed by corresponding Variant Curation Expert Panels, are available through this resource.

Population Frequency & gnomAD database



An important part of variant curation is determining whether or not a variant has been found in the general population and at what **allele frequency (AF)**. If a variant is found at a frequency that is higher than expected for the disease, the variant is most likely not causing disease. The most common way to determine the AF of the variant is by using gnomAD. Below, find information on how to use gnomAD and apply the codes associated with AF.

gnomAD Database

Website: https://gnomad.broadinstitute.org/

Overview: The gnomAD database is a collection of variants that have been found in a multi-ethnic population, which presumably are free of severe childhood onset disease (PMID: 32461654). We use the allele frequency of the highest population (MaxAF) to determine whether a pathogenic or benign code should be applied to the variant. Typically, do not use populations which have less than 2,000 alleles (1,000 individuals) for consideration of MaxAF.

How to search for a variant:

Since variants can differ in nomenclature depending on the build and transcript used, the best way to get to a gnomAD variant is by using Franklin or Varsome (see below). However use the following guide for searching directly on gnomAD.

You can search using the genomic coordinate directly from the first page (for example "<u>1-55051215-G-GA</u>", which includes: "chromosome-genomic coordinate-ref allele-alt allele"). However the easier way to do this is by first going to the gene page, and then scrolling or searching to find the variant of interest. If your variant is not present, that means that it was not identified in the control population.

Please note:

- ✓ It is important to confirm that the build and the transcript that is being used by gnomAD is the same as your variant. If the transcript you were given for the variant is different than the one listed in gnomAD (MANE select transcript), you can instead use the genomic coordinate (g.). However the genomic coordinate will differ based on the genomic build that was used. gnomAD version 2 is on build37 (hg19), gnomAD version 3 and 4 are on build38 (hg38).
- ✓ A variant may be present in one version, but absent in another. To toggle between the different versions of gnomAD, click on the arrow next to the dataset version (Can be found at the top right of the screen) and click on the version you would like to view. Typically, the newest version of gnomAD should be used as it has the biggest sample size.

			- Scarci		About learn	Stats Policies	Publications Blog Char	ngelog Downloa	ids Forum Conta	ict riel
We are aware of	an issue in the gno	omAD v4.0 exomes where well	covered variants have lo	ower than expected allele	numbers. This issu	e will be fixed in	the upcoming v4.1 release.	For more informa	tion, see our write-u	up here,
FBN	1 fibrillin 1						Dataset gno	mAD v4.0.0 - gr	nomAD SVs v4.0 👻	0
				Constraint	Variant co-o	ccurrence 😧				
Genome build GRCh38 / hg38 Ensembl gene ID ENSG0000166147.15 MANE Select transcript @ ENST00000316623.10 / NM_000138.5 Ensembl canonical transcript @ ENST00000316623.10 Other transcripts ENST0000037463.6, ENST00000558230.1, and 10 more Region 15.48408313-48645721			Category	Expected SNVs	Observed SNV	Vs Constraint metrics				
						Z = 1.87				
			Synonymo	us <u>130/</u>	1183	o/e = 0.91 (0.86 - 0.95)	0 1			
			Missense	3771.9	2396	Z = 8.18 o/e = 0.64 (0.61 - 0.66)	01			
Externa	I resources Ensem	ibl, UCSC Browser, and more		pLoF	352.7	27	pLI = <u>1</u> o/e = <u>0.08</u> (<u>0.06</u> - <mark>0.11)</mark>	0 •1		
				Constraint	metrics based on N	MANE Select trans	script (ENST00000316623.10)).		
and the second sec		1. 101	and the second second	-						
tots Only variants i he table below sho elect transcript (or Variant ID	located in or with ws the HGVS cor non-canonical tri <u>Source</u>	nin 75 base pairs of a coding reequence and VEP annotati anscript if no MANE Select tr HGVS Consequence	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u>	To see variants in UTRs ost severe consequence noted with ⁺ . To see con	or introns, use the e across all transcr isequences in a sp <u>Clinical</u>	e region view. ipts in this gene ecific transcript, Flags	c. Cases where the most se use the transcript view. Allele Count	vere consequen <u>Allele</u> Number	ce occurs in a non-	-MANE
Note Only variants he table below sho elect transcript (or Variant ID	located in or with was the HGVS cor non-canonical tra Source	nin 75 base pairs of a coding nsequence and VEP annotation anscript if no MANE Select to HGVS Consequence	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u>	To see variants in UTRs ost severe consequence noted with †. To see con LoF. Curation	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> Significance	e region view. ipts in this gene ecific transcript, Elags	c. Cases where the most se use the transcript view. <u>Allele Count</u>	vere consequent <u>Allele</u> <u>Number</u>	ce occurs in a non Allele Erequency	-MANE Nu Hon
only variants he table below sho elect transcript (or Variant.ID 15-48411050-G-A	located in or with was the HGVS cor non-canonical tra Source	in 75 base pairs of a coding isequence and VEP annotati anscript if no MANE Select to HGVS Consequence p.Asp2852Asp	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u> • synonymous	To see variants in UTRs ost severe consequence noted with †. To see con <u>koF</u> <u>Curation</u>	or introns, use the e across all transcr isequences in a sp <u>Clinical</u> Significance	e region view. ipts in this gene ecific transcript, Elags	c. Cases where the most se use the transcript view. <u>Allele Count</u> 1	Allele Number 833040	ce occurs in a non Allele Frequency 1.20e-6	-MANE <u>Nu</u> Hon
Mariants Only variants he table below sho elect transcript (or Variant.ID 15-48411050-G-A 15-48411051-T-G	located in or with was the HGVS cor non-canonical tra- Source	in 75 base pairs of a coding sequence and VEP annotation anscript if no MANE Select to HGVS Consequence p.Asp2852Asp p.Asp2852Ala	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u> • synonymous • missense	To see variants in UTRs ost severe consequence noted with ±. To see con LoF. Surration	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> Significance	e region view. ipts in this gene ecific transcript, <u>Flags</u>	c. Cases where the most se use the transcript view. <u>Allele Count</u> 1	Allele Allele Number 833040 152214	ce occurs in a non Allele Frequency 1.20e-6 6.57e-6	-MANE <u>Nu</u> Hon
Area Only variants the table below sho elect transcript (or Variant.ID 15-48411050-G-A 15-48411051-T-G 15-48411051-T-C	located in or with wws the HGVS cor non-canonical tra Source	in 75 base pairs of a coding sequence and VEP annotation anscript if no MANE Select to HGVS Consequence p.Asp2852Asp p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense	To see variants in UTRs ost severe consequence noted with To see con LoF Suration	or introns, use the e across all transcr isequences in a sp <u>Clinical</u> Significance	e region view. ipts in this gene ecific transcript. <u>Flags</u>	c. Cases where the most se use the transcript view. <u>Allele Count</u> 1 1 6	Allele Number 833040 152214 1461622	Ce occurs in a non Allele Fresuency 1.20e-6 6.57e-6 4.11e-6 4.11e-6	-MANE <u>Nu</u> Hon
Anter Only variants he table below sho elect transcript (or Variant.ID 15-48411050-G-A 15-48411051-T-G 15-48411051-T-C 15-48411053-T-C	located in or with ways the HGVS corn non-canonical tra- Sourse	in 75 base pairs of a coding sequence and VEP annotation anscript if no MANE Select to HGVS Consequence p.Asp2852Asp p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala p.Asp2852Asp	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense • missense • missense	To see variants in UTRs ost severe consequence noted with +. To see con LoF Suration	or introns, use the e across all transcr isequences in a sp <u>Clinical</u> significance erfain significance	e region view. ipts in this gene ecific transcript. <u>Flags</u>	e. Cases where the most se use the transcript view. <u>Allele Count</u> 1 1 6 1	Allele Number 833040 152214 1461622 833050	Allele Fresuency 1.20e-6 6.57e-6 4.11e-6 1.20e-6	-MANE <u>Ny</u> Hon
Conjy variants i the table below sho elect transcript (or Variant.ID 15-49411050-G-A 15-48411051-I-G 15-48411051-I-G 15-48411053-I-G 15-48411055-I-C	located in or with was the HGVS cor non-canonical tra- * Source 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	in 75 base pairs of a coding isequence and VEP annotati anscript if no MANE Select to <u>HGVS.Consequence</u> p.Asp2852Asp p.Asp2852Asp p.Asp2852As p.Asp2852Giv p.Lys285TAsn p.Lys285TGiu	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense • missense • missense • missense	To see variants in UTRs ost severe consequence noted with *. To see con LoF Suration	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> <u>Sianificance</u> ertain significance	e region view. ipts in this gene ecific transcript. <u>Flags</u>	e. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73	Allele Number 833040 152214 1461622 833050 1613810	ce occurs in a nom Allele Frsauensy 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 4.52e-5	-MANE <u>Nu</u> Hon
Conjugation of the second seco	located in or with was the HGVS cor non-canonical tra- * Source 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	In 75 base pairs of a coding insequence and VEP annotati anscript if no MANE Select to HGVS Consequence p.Asp2852Asp p.Asp2852Ala p.Asp2852Gly p.Lys2851Asn p.Lys2851Glu p.Asp2850Gly a.Tu-2007p	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u>	To see variants in UTRs ost severe consequence noted with ⁺ . To see con Loff Curstion	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> Significance ertain significance ertain significance	e region view. ipts in this gene ecific transcript. Flags	c. Cases where the most se use the transcript view. Allele Count 1 1 6 1 7 3 1	Allele Number 833040 152214 1461622 833050 1613810 833064 933054	Ce occurs in a nom Allele Frequency 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6	-MANE Nu Han
2012 Only variants he table below sho elect transcript (or Variant.ID 15-48411051-T-G 15-48411051-T-G 15-48411053-T-G 15-48411055-T-C 15-48411055-T-C 15-48411055-T-C	located in or with was the HGVS cor non-canonical tri Source	In 75 base pairs of a coding sequence and VEP annotati anscript if no MANE Select tr HGVS Consequence p.Asp2852Asp p.Asp2852Ala p.Asp2852Gly p.Lys285TAsn p.Lys285TGlu p.Asp2850Gly p.Tyr2849Tyr	exon are shown here. on for each variant's mr anscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense • missense • missense • missense • missense • missense • missense • missense • synonymous	To see variants in UTRs ost severe consequence noted with 1. To see con <u>Lef</u> <u>Curation</u> : Unc	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> <u>Significance</u> ertain significance ertain significance	e region view. ipts in this gene ecific transcript. Flags	c. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 1 4	Allele Number 833040 152214 1461622 833050 1613810 833064 833076	Ce occurs in a non Allele Frequency 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 4.37	-MANE Nu Lion
Conjy variants the table below sho elect transcript (or Variant.ID 15-48411051-T-G 15-48411051-T-G 15-48411053-T-C 15-48411055-T-C 15-48411055-T-C 15-48411055-T-C 15-48411055-T-C 15-48411055-T-C	located in or with was the HGVS cor non-canonical tri Source	in 75 base pairs of a coding isequence and VEP annotation anscript if no MANE Select to HGVS Consequence p.Asp2852Asp p.Asp2852Ala p.Asp2852Gly p.Lys2851Glu p.Asp2850Gly p.Tyr2849Tyr p.Tyr2849Asn	exon are shown here. on for each varian'ts m ranscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense • missense • missense • missense • missense • synonymous • missense	To see variants in UTRs ost severe consequence noted with 1.7 see con <u>LoF</u> <u>Curation</u> Unc Unc	or introns, use the a across all transcr sequences in a sp <u>Clinical</u> <u>significance</u> ertain significance ertain significance	e region view. ipts in this gene ecific transcript. Elags	c. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 1 4 2	Allele Number 833040 152214 1461622 833050 1613810 833064 833056 1461614	Ce occurs in a non Allele Frequency 1.20=6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 1.37e-6 0.01 - 7	-MANE Nu Hon
Variant: Variant: I 5-48411050-G-A 15-48411051-T-G 15-48411051-T-G 15-48411053-T-G 15-48411055-T-C 15-48411055-T-C 15-48411057-T-C 15-48411057-T-C 15-48411057-T-C 15-48411057-T-C	located in or with was the HGVS cor non-canonical tra- Source C C C C C C C C C C C C C C C C C C C	in 75 base pairs of a coding isequence and VEP annotation anscript if no MANE Select to p.Asp2852Asp p.Asp2852Ala p.Asp2852Gly p.Lsy2851Asn p.Lsy2851Glu p.Asp2850Gly p.Tyr2849Tyr p.Tyr2849Asn p.Lsy22850Gly p.Tyr2849Asn p.Lsy22850Gly	exon are shown here. on for each variant's m manscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense • missense • missense • missense • missense • missense • synonymous • missense • missense • missense	To see variants in UTRs ost severe consequence noted with . To see con <u>LOF</u> <u>Suration</u> Unc	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> isomificance ertain significance ertain significance	e region view. ipts in this gene ecific transcript, Elags	e. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 7 1 4 2 1	Allele Number 833040 152214 1461622 833050 1613810 833054 833054 833056 1461614 1461626	Ersturence Fristurence 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 1.37e-6 6.84e-7	-MANE Nu Hon
2012 Only variants he table below sho elect transcript (or Variant.I.D 15-48411050-G-A 15-48411051-T-G 15-48411053-T-G 15-48411053-T-G 15-48411057-T-C 15-48411057-T-G 15-48411057-T-G 15-48411063-T-A 15-48411063-T-A	located in or with ways the HGVS cor- non-canonical tra- Constant	in 75 base pairs of a coding insequence and VEP annotati anscript if no MANE Select to <u>HGVS.Consequence</u> p.Asp2852Asp p.Asp2852Ala p.Asp2852Ala p.Asp2852Glu p.Lys2851Asn p.Lys2851Glu p.Asp2850Gly p.Tyr2849Tyr p.Tyr2849Asn p.Lys2849Asn p.Lys2849Asn p.Lys2849Asn	exon are shown here. on for each variant's m manscript exists) are der <u>VEP Annotation</u> synonymous missense missense missense missense missense missense missense missense missense missense missense missense missense missense missense missense	To see variants in UTRs ost severe consequence noted with +. To see con <u>LoF</u> <u>Suration</u> Unc Unc	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> significance ertain significance ertain significance ertain significance ertain significance	e region view. ipts in this gene ecific transcript, Elags	e. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 1 4 2 1 6 1 6 1 73 1 4 2 1 6	Allele Number 833040 152214 1461622 833050 1613810 833064 833076 1461614 1461626 1613848	Ce occurs in a nom Allele Fristunder 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 1.37e-6 6.84e-7 3.72e-6 6.74e-7 3.72e-6 6.54e-7 3.72e-6 3.72e-7 3.72e-6 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7 3.72e-7	Mane Nu Hon
2012 Only variants he table below sho elect transcript (or Variant.ID 15-48411051-T-G 15-48411051-T-G 15-48411057-T-C 15-48411057-T-C 15-48411057-T-G 15-48411057-T-G 15-48411057-T-G 15-48411063-T-G 15-48411063-T-G 15-48411063-T-G	located in or with ways the HGVS cor- non-canonical tra- Source G G G G G G G G G G	In 75 base pairs of a coding insequence and VEP annotati anscript if no MANE Select to HGVS Consequence p.Asp2852Asp p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala p.Lys2851Glu p.Asp2850Gly p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Tyr p.Lys2848Ile p.Lys2848Ile p.Lys2848Ile	exon are shown here on for each variant's m mascript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense	To see variants in UTRs ost severe consequence noted with F. To see con LoF Suration : Unc Unc	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> <u>Sianificance</u> ertain significance ertain significance ertain significance ertain significance	e region view. ipts in this gene ecific transcript. Elaas	e. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 1 4 2 1 1 6 1 7 3 1 4 2 1 1 6 1 7 3	Allele Number 833040 152214 1461622 833050 1613810 833064 833050 1613810 833066 1461614 1461626 1613848 1461626	Ce occurs in a nom Allele Prosucosy 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 1.37e-6 6.84e-7 3.72e-6 6.84e-7 6.84e-7	<u>Nun</u>
The table below sho elect transcript (or Variant.ID 15-48411051-CG 15-48411051-CG 15-48411051-CG 15-48411053-T-G 15-48411053-T-G 15-48411053-T-G 15-48411053-T-G 15-48411063-T-G 15-48411063-T-G 15-48411063-T-G 15-48411063-T-G	located in or with was the HGVS cor- non-canonical tra- Source C C C C C C C C C C	In 75 base pairs of a coding in requence and VEP annotati anscript if no MANE Select tr HGVS Consequence p.Asp2852Asp p.Asp2852Ala p.Asp2852Gly p.Lys2851Glu p.Asp2850Gly p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Asn p.Lys2848Ile p.Lys2848Ile p.Lys2848Ile p.Lys2848Ile	exon are shown here." on for each varian's m mascript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense • missense • missense • missense • synonymous • missense • synonymous • missense • missense • missense • missense • missense • missense • missense • missense • missense • missense	To see variants in UTRs ost severe consequence noted with 1. To see com Curation : Unc Unc Unc Unc Unc Unc	or introns, use the e across all transcr sequences in a sp <u>Clinical</u> Significance ertain significance ertain significance ertain significance ertain significance ertain significance ertain significance	e region view. ipts in this gene ecific transcript. Flags	2. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 1 4 2 1 1 6 1 1 6 1 1 7 3 1 1 4 2 1 1 6 1 7 3 1 1 7 3 1 7 3 1 7 3 1 1 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 3 1 7 7 3 1 7 3 1 7 3 1 7 7 3 1 7 7 7 7	Allele Number 833040 152214 1461622 833050 1613810 833064 833076 1461614 1461626 1613848 1461626 700914	Ce occurs in a nom Allele Frequency 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 1.37e-6 6.84e-7 3.72e-6 6.84e-7 8.96e-5	Nun Hann
To Service 2014 Se	located in or with ways the HGVS cor- non-canonical tra- Source C C C C C C C C C C	in 75 base pairs of a coding isequence and VEP annotation ascript if no MANE Select to HGVS.Consequence p.Asp2852Ala p.Asp2852Ala p.Asp2852Ala p.Asp2852Gly p.Lys2851Glu p.Asp2850Gly p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Tyr p.Lys2848Thr p.Lys2848Thr p.Lys2848Thr p.Lys2848Thr p.Lys2848Thr p.Lys2848Thr p.Lys2848Thr	exon are shown here. on for each variant's mr anscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense	To see variants in UTRs ost severe consequence noted with 1. To see con <u>Loff</u> <u>Curration</u> Unc Unc Unc Unc Unc	or introns, use the a across all transcr sequences in a sp <u>Clinical</u> significance ertain significance ertain significance ertain significance ertain significance or an action significance ertain significance	e region view. ipts in this gene ecific transcript. Elags	e. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 1 4 2 1 4 2 1 1 6 1 7 3 3	Allele Number 833040 152214 1461622 833050 1613810 833064 833066 1461614 1461626 1613848 1461626 780914 1461608	Allele FISSMARCY 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 1.37e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 8.96e-5 2.05e-6	Nun Hom
Conjugation of the second state of the second	located in or with was the HGVS cor non-canonical tra- Source C C C C C C C C C C	in 75 base pairs of a coding isequence and VEP annotation anscript if no MANE Select to p.Asp2852Asp p.Asp2852Ala p.Asp2852Gily p.Lys2851Gilu p.Asp2850Gily p.Tyr2849Tyr p.Tyr2849Tyr p.Tyr2849Asn p.Lys2846Asp p.Lys2848Arq p.Lys2848Arq p.Lys2848Arq p.Lys2848Arq p.Lys2848Arq p.Lys2848Cin p.Glu2846Glu p.Glu2846Asp	exon are shown here. on for each variant's m ranscript exists) are der <u>VEP Annotation</u> • synonymous • missense • missense	To see variants in UTRs ost severe consequence noted with 1.70 see con <u>LOF</u> <u>Suration</u> Unc Unc Unc Unc	or introns, use the e across all transcriss equences in a sp <u>Clinical</u> ertain significance ertain significance ertain significance ertain significance ertain significance ertain significance w benign	e region view. ipts in this gene ecific transcript. Elags	e. Cases where the most se use the transcript view. Allele Count 1 1 6 1 73 1 4 2 1 1 6 1 7 3 1 4 2 1 1 6 3 4 3 4	Allele Number 833040 152214 1461622 833050 1613810 833054 833054 833056 1461614 1461626 1613848 1461626 700914 1461608 1613846	Erissuence Fristuence 1.20e-6 6.57e-6 4.11e-6 1.20e-6 4.52e-5 1.20e-6 4.80e-6 1.37e-6 6.84e-7 3.72e-6 6.84e-7 8.96e-5 2.05e-6 6.84e-7 3.72e-6 6.84e-7 8.96e-5 2.05e-6 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 6.84e-7 3.72e-6 3.84e-7 3.95e-5 2.05e-6 2.48e-6 3.48e-7 3.48e-6 3.48e-6 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-6 3.48e-7 3.48e-	Nu Hon

You can use the "Search variant table" feature located above the "Export variants to CSV" button to find an amino acid p. Or c. (Please note: this variant table may appear on a different place of the page depending on the size of your browser).

If you click on the link to the variant page (blue coordinates in the Variant ID column), you will be able to see population frequencies. We would use the frequency of the most common population (in the case below - "European" at an AF of 0.000005 or 0.0005%).

gnomAD browser gnomAD v4	0.0 • Se	arch		About Team Stats Polici	es Publications B	log Changelog Download	ds Forum Contact Help/FAQ
We are aware of an issue in the gnomAD v4.0 exomes	where well covered	d variants have lowe	er than expected al	lele numbers. This issue will be fixed	in the upcoming v4.	1 release. For more informat	ion, see our write-up <u>here</u>
SNV: 15-4841106	B-T-G(GR	Ch38) 🔍	opy variant ID	Gene page		Dataset gnomAD v4.0	0.0 - 0
Filters Allele Count Allele Number Allele Finguenzy Gramas Filtering AF (95% confidence) Number of homozvates Frastion of individuals with .20x sovernae Genetic Ancestry Group	Exomes 5 1461626 0.000003421 0.000001320 0 1.0 Frequencie	Genomes Peri 1 152222 0.000006569 0 0 1.0 5 ©	Total 6 1613848 0.000003718 0.000001830 0	External Resources • dbSNP (v;75839151) • UCSC • ClinVar (22868) • All of Us Feedback Report an issue with this variant			
Genetic Ancestry Group	A	llele Count	Allele Nun	ber Number o	f Homozygotes	Allele Frequency	-
+ European (non-Finnish)		6	117	9964	0	0.000005085	
 Remaining 		0	6	2482	0	0.000	
 Admixed American 		0	5	9984	0	0.000	
+ European (Finnish)		0	6	4032	0	0.000	
 Middle Eastern 		0		5076	0	0.000	
+ South Aslan		0	9	1024	0	0.000	

gnomAD can also give you other information about the gene or variant:

- ✓ External resources It provides links to ClinVar or UCSC if available
- ✓ Variant Effect Predictor it can show you what effect the variant has on alternative transcripts or in other genes
- ✓ In silico predictions it will give you REVEL scores and other computational predictors
- ✓ Gene constraint scores it can show you if the gene is constrained for loss of function or missense variants (i.e. it has a higher or lower than expected number of those variants)
- ✓ Age distribution: gnomAD will sometimes show you the ages of the individuals who have that variant. This may be helpful to understand the likelihood that those individuals would already have the disease in question. For example, if the individuals are under the age that you would expect to see the disease in question, you may not consider those individuals as helpful to your curation.

How to determine if the allele frequency is high or low: Allowed AF is different for each gene and condition, and depends on a number of things including penetrance, prevalence, mode of inheritance and allelic/genetic heterogeneity (whether or not multiple genes or variants can cause the condition). We have compiled a list of AF cutoffs for different conditions below. If the allele frequency of your variant is *below* the **PM2_Supporting** cutoff, you can apply that code. If the AF is *above* the **BS1 or BA1** cutoff, the respective code can be applied. Please note that BA1 is a stand alone code. Therefore if your variant reached BA1 it is automatically considered benign. (Please note that there are exceptions to this rule such as hypomorphic or low penetrance variants, etc.)

How each factor influences the allele frequency of a gene/disease pair:

 Mode of inheritance: The allowed AF of a recessive condition will be higher than a dominant condition, because unaffected carriers will be present in the database and are expected

- ✓ Penetrance: If a condition has low penetrance, the allowed AF will be higher, as there will be more "healthy" individuals with a pathogenic variant that will not develop disease.
- Prevalence: If a disease is more common, we expect a higher allele frequency compared to a condition that is extremely rare
- ✓ Allelic/genetic heterogeneity: If many genes and variant exist that cause the same condition, the allowed AF should be reduced, as only some of the prevalence of the disease can be attributed to that specific variant.

How to calculate an allowed allele frequency if your gene does not have a VCEP guideline: For autosomal recessive conditions, the carrier frequency of a variant should not exceed the Hardy-Weinberg equilibrium based on the prevalence ($p^2 + 2pq + q^2 = 1$) where q^2 is the prevalence of the condition and 2pq is the carrier frequency (to calculate we set $p^2 = 1$). An online calculator such as <u>https://www.perinatology.com/calculators/Hardy-Weinberg.htm</u> or <u>https://www.cardiodb.org/allelefrequencyapp/</u> can be used to find the maximum allowed carrier frequency if you know the prevalence of the condition.

TIP: Caution should be taken in the case of hypomorphic alleles, which may not cause disease in the homozygous state, however may cause disease when in trans with a severe variant. These may be at higher frequencies in the general population or found in the homozygous state in healthy individuals, but still impact disease risk.

PM2_Supporting/BS1/BA1 Cutoffs

Please note: gnomAD will provide you the allele frequency and not the percentage. To use the below percentages, make sure to convert.

Gene/Disease	PM2_Supporting	BS1	BA1
ACADVL	0.1%	0.35%	0.7%
Brain Malformations		0.0185%	0.0926%
Cardiomyopathy (<i>MYH7</i> , etc.)	0.004%	0.02%	0.1%
CDH1	0.001%	0.1%	0.2%
Cerebral Creatine Deficiency (GAMT)	0.04%	0.1%	0.3%
Cerebral Creatine Deficiency (GATM)	0.0055%	0.01%	0.05%
Cerebral Creatine Deficiency (<i>SLC6A8</i>)	0.002%	0.02%	0.2%
F8	Absent in males	0.003%	0.03%

F9	Absent in males	0.00056%	0.0056%
DICER1	<0.0005%	0.03%	0.3%
BRCA1	absent	0.01%	0.1%
BRCA2	absent	0.02%	0.1%
Epilepsy - SCN1A	0-1 alleles	0.0004%	0.02%
Epilepsy - SCN2A	0-1 alleles	0.0002%	0.01%
Epilepsy - SCN3A	0-1 alleles	0.0002%	0.01%
Epilepsy - SCN8A	0-1 alleles	0.0002%	0.01%
Epilepsy - SCN1B	0-1 alleles	0.01%	0.3%
Familial Hypercholesterolemia (<i>LDLR</i>)	0.02%	0.2%	0.5%
FBN1	0.0005%	0.005%	0.1%
Glaucoma- MYOC	0.01%	0.1%	1%
Hearing Loss (Autosomal Recessive)	0.007%	0.3%	0.5%
Hearing Loss (Autosomal Dominant)	0.002%	0.02%	0.1%
ATM	0.001%	0.05%	0.5%
PALB2	0.0003%	0.01%	0.1%
Hereditary Hemorrhagic Telangiectasia - ENG/ACVRL1	0.004%	0.08%	0.2%
Colorectal Cancer - APC	0.0003%	0.001%	0.1%
LCA/early onset retinal dystrophy	0.02%	0.08%	0.8%
Lysosomal Diseases (GAA)	0.1%	0.5%	1%
RYR1 (Malignant	Not used	0.08%	0.38%

hyperthermia)			
SLC19A3	0.005%	0.05%	0.1%
PDHA1	0.00092%	0.092%	
POLG	0.05%	.5%	1%
ETHE1	0.002%	0.02%	0.1%
mtDNA genes	0.002%	0.5%	1%
Monogenic diabetes (HNF1A, HNF4A, GCK)	0.0003%*	0.003% (0.004% for GCK)	0.01%
RUNX1	absent	0.015%	0.15%
РАН	0.02%	0.2%	1.5%
Platelet disorders (<i>ITGA2B</i> , <i>ITGB3</i>)	0.01%	0.158%	0.24%
PTEN	0.001%	0.0043%	0.056%
Pulmonary Hypertension (<i>BMPR2</i>)	0.01%	0.1%	1%
RASopathy (SHOC2, NRAS, RAF1, SOS1, SOS2, PTPN11, KRAS, MAP2K1, HRAS, MAP2K2, BRAF)	absent	0.025%	0.05%
Rett and Angelman- like disorders (<i>MECP2, CDKL5, FOXG1, UBE3A,</i> <i>SLC9A6, TCF4</i>)	absent	0.008% (0.025% for CDKL5)	0.03% (0.05% for TCF4)
Severe Combined Immunodeficiency (SCID) - <i>ADA</i>	0.017%	0.16%	0.7%
SCID - DLCRE1C	0.003% and no homozygotes	0.078%	0.35%
SCID - IL7R	0.004%	0.13%	0.57%
SCID -JAK3	0.01%	0.1%	0.45%
SCID - RAG1	0.01%	0.19%	0.87%

SCID - RAG2	0.0059% and no homozygotes	0.19%	0.87%
SCID - IL2RG	0.01%	0.25%	1.1%
Thrombosis (SERPINC1)	0.002%	0.02% (be careful of frequent founder variants)	0.2%
TP53	Absent	0.03%	0.1%
VHL	0.00016%	0.0016%	0.016%

Computational Data



In silico prediction tools can be used as part of the variant curation process. Typically, these will only be assigned a supporting weight code (**PP3**). While there are many in silico tools (SIFT, PolyPhen, etc.) that can be used, typically the meta-predictor REVEL is used. The following REVEL scores are used:

*Please note that ClinGen VCEPs sometimes change the REVEL cutoffs. Please look at the VCEP for the disease you are curating to see if this aligns. If there is no VCEP, use the below.



REVEL score

Where to find REVEL scores: REVEL scores can be found on the gnomAD website or on the curation tools detailed below (Franklin and Varsome). If your variant is not present in gnomAD, you can also add the REVEL score track on UCSC, which will give you the REVEL score for each variant at a specific location. See below for an example:



If you do not have the REVEL track present on your view of UCSC, you can add it using the track feature at the bottom of the website. When you scroll down, you will see "REVEL Scores" under the "Phenotype and Literature" category:



Predictive Data (PVS1, Loss of function)

Variants which predict a **loss of function** (such as frameshift, nonsense and splice site variants) can have a very strong code (PVS1) applied. This means that if the variant qualifies for PVS1 at the very strong level, the variant only needs to be found infrequently in the population (PM2_Supporting) to reach Likely Pathogenic. Please see the ClinGen specifications for using these codes here: https://clinicalgenome.org/site/assets/files/5182/pm2_-_svi_recommendation_-_approved_sept2020.pdf. Since these variants can make it to Likely Pathogenic without any other evidence (proband, functional, segregations, etc.), it is important to make sure that the loss of function variant qualifies for PVS1. For a full explanation of using PVS1, please see PMID: 30192042. Below find an overview of how to determine if PVS1 should be applied and at what level.

The following flowcharts can be used if:

- Loss of function is a proven mechanism of disease for the gene
- The exon that the variant is in is in biologically relevant transcripts and is not alternatively spliced (it is present in transcripts that are expressed in the body)

For **nonsense and frameshift variants**, the weight of PVS1 is determined by whether **nonsense mediated decay (NMD)** is predicted to occur and how much of the protein is removed.



Nonsense Mediated Decay is a mechanism used by cells to degrade mRNAs that contain premature stop codons caused by nonsense or frameshift variants. However typically, **stop codons which are in the final exon or the 50bp of the second to last exon** are NOT predicted to have NMD occur.



For splice variants, the weight of PVS1 applied is dependent on whether the removal of the exon that the splice variant impacts is out of frame or in frame.



An exon is **out of frame** (the reading frame is disrupted) if the number of nucleotides in the exon removed is NOT a multiple of 3. In this case, when the exon before and after are spliced together, the reading frame shifts causing frameshift or premature stops in the mRNA. An exon is **in frame** if the number of nucleotides is a multiple of 3. In this case, the exons before and after will splice together in a way that preserves the reading frame. In that case, the level of PVS1 is dependent how much of the protein is removed (less or more than 10%).



Mass General Brigham VARIANT CURATION TASK FORCE



Proband and Family Data



A significant part of variant curation is looking in the literature for individuals with the variant of interest who are affected with the corresponding condition. The number of probands needed to reach a supporting, moderate or strong level of evidence depends on the disease inheritance and specificity. Below find general guidelines on how many probands to count.

Dominant Conditions:

Category 1: Common diseases (ex. Cardiomyopathy, BRCA1/2)

Probands needed	Strength of code
2-5	Supporting (PS4_P)
6-14	Moderate (PS4_M)
15+	Strong (PS4)

Category 2: Rare disease, but the phenotype is not specific (ex. Neurodevelopmental disorder)

Probands needed	Strength of code
1-2	Supporting (PS4_P)
3-4	Moderate (PS4_M)
5+	Strong (PS4)

Category 3: Rare disease with specific phenotypes (ex. Pheochromocytoma)

Probands needed	Strength of code
1	Supporting (PS4_P)
2-3	Moderate (PS4_M)
4+	Strong (PS4)

Recessive Conditions:

For recessive conditions, the points per proband depends on the zygosity of the variant. For homozygous variants, the allotted number of points below is assigned. For compound heterozygotes, the points assigned depend on whether the other variant found is pathogenic and whether it was confirmed to be *in trans*. and, for compound heterozygotes, the pathogenicity of the variant that is found along with your variant. Probands who have the variant in question in the heterozygous state WITHOUT another variant identified do not receive any points. Points are summed across all probands identified in the literature to determine the strength of the code.

Zygosity	Points assigned (per proband)
Homozygous variant in a proband with consanguineous parents	0.25
Homozygous variant	0.5
Compound heterozygous with a pathogenic variant found in trans	1
Compound heterozygous with a pathogenic variant of unknown phase	0.5
Compound heterozygous with a rare VUS in trans	0.25

Total points (summed across all probands found in literature)	Strength of code
0.5	Supporting (PM3_P)
1	Moderate (PM3)
2	Strong (PM3_S)
4	Very Strong (PM3_VS)

Segregations:

Segregations are the number of affected family members with the same variant as the proband and the same phenotype. For more information on how to count segregations please see ClinGen's training tool here: https://clinicalgenome.org/docs/segregation-analysis/

Do	minant	Reces	sive
Total segregations	Strength of code	Total Segregations	Strength of code
3	Supporting (PP1)	1	Supporting (PP1)
5	Moderate (PP1_M)	2	Moderate (PP1_M)
7	Strong (PP1_S)	3	Strong (PP1_S)

Variant Curation Databases

ClinVar

Website: https://www.ncbi.nlm.nih.gov/clinvar/

Overview: ClinVar contains entries for specific variants from various submitters (clinical labs, research labs, providers, etc.) asserting a pathogenicity. For instance, Invitae or GeneDx, may submit a variant as Pathogenic and may include information about why they have classified it that way.

How to use: Search for a variant of interest using the transcript and c. or by going first to the gene page and then scrolling to find your variant.

Note: Please make sure that the transcript that you are using matches the one used in ClinVar. Alternatively you can use the genomic coordinate.



Classification © (Last evaluated)	Review status @ (Assertion criteria)	Condition Θ	Submitter @	More information 0	~
Uncertain significance (May 16, 2022)	★ ☆ ☆ ☆ (ACMG Guidelines, 2015) Method: clinical testing	Cardiomyopathy Affected status: unknown Allele origin: germline	Color Diagnostics, LLC DBA Color Health Accession: SCV004359495.1 First in CilnVar: Feb 14, 2024 Last updated: Feb 14, 2024		^
	Comment: This missense variant replaces a protein structure and function (inte this variant. This variant has not be general population by the Genome conclusively. Therefore, this varian	sparagine with serine at codon 1933 of ernally defined REVEL score threshold « sen reported in individuals affected with Aggregation Database (gnomAD). The t is classified as a Variant of Uncertain i	the MYH7 protein. Computational prediction o.5. PMID: 27665373). To our knowledge, fu cardiovascular disorders in the literature. Th available evidence is insufficient to determini significance. (less)	suggests that this variant may not i inctional studies have not been rep is variant has not been identified in e the role of this variant in disease	mpact orted for the
Uncertain significance (Dec 06, 2020)	★ ☆ ☆ ☆ (Invitae Variant Classification Sherloc (09022015)) Method: clinical testing	Hypertrophic cardiomyopathy Affected status: unknown Allele origin: germline	Invitae Accession: SCV002117333.3 First in ClinVar: Mar 28, 2022 Last updated: Feb 28, 2024		^
	Comment: This variant has not been reported in the literature in individuals with MYH7-related conditions. This variant is not present in population databases (ExAC no frequency). This sequence change replaces asparagine with serine at codon 1933 of the MYH7 protein (p. Asn1933Ser). The asparagine residue is highly conserved and there is a small physicochemical difference between asparagine and serine. Algorithms developed to predict the effect of missense changes on protein structure and functions are either unavailable or do not agree on the potential impact of this missense change (SIFT: 'Deleterious' PolyPhen-Z' Benign', Align-GVBD: 'Class CO'). In summary, the available evidence is currently insufficient to determine the role of this variant in disease. Therefore, it has been classified as a Variant of Uncertain Significance. Algorithms developed to predict the effect of sequence changes on RNA splicing suggest that this variant may create or strengthen a splice site, but this prediction has not been confirmed by published transcriptional studies, (less)				

You can read through the comments section of the lab submissions to determine if there are any publications of interest or whether or not there are functional or genetic data available.

Why labs may differ in their classification:

- ✓ Different interpretations of the guidelines used for variant interpretation (ACMG)
- ✓ Internal patient data and proprietary databases/tools
- ✓ Timing: Variant classification may change over time due to updated information/more publication data. Make sure to look at the date assessed to determine if a classification may be out of date

Franklin

Website: https://franklin.genoox.com/clinical-db/home

Description: A variant curation tool, which compiles information about the variant from multiple other databases (such as gnomAD/ClinVar).

PLEASE NOTE: The initial automated classification that is given with Franklin should not be taken as the actual classification. Instead, use the tools provided to add your own codes and use Franklin only to provide information from relevant databases.

How to Use: Enter your variant into the search tool. (Suggested nomenclature is "Transcript : c." such as: NM_000368.5:c.359T>C)

Franklin will give you a suggested classification, along with which codes they are using. If you disagree with the codes, you can downgrade or add your own.

Fr	anklin	NM_000257.4	4:c.4258C>T	Q SEARCH		DGE BASE	🝺 MY CASES	
Search M chr14-	Page MYH7:c.4258C 1 H7:c.42 23417598 G>A p.Arg1 Classify Variant	>T 58C>T 420Trp NM_00025 2 Follow 🖾	7.4 dbSNP, <u>rs14521</u> Save Case	<u>3771 UCSC</u> 亿, g <u>r</u> Export Summary	iomAD ප			
<	Franklin ACMG	Classification	₀∬0 Variant Asse	ssment 🚛	Publications ¥	Gene Assessm Suggested Classifica	nent 🔞 Associat	ed (
		Benign	• Likely Benign	VUS	Likely Pathogenic	Pathogenic		

Franklin can tell you:

- ✓ If the gene has ClinGen specific guidelines
- ✓ Population frequency from multiple databases, including gnomAD
- \checkmark ClinVar entries
- \checkmark Summary of proband data that has been submitted to ClinVar
- ✓ Publications associated with variant (not always correct publications may be missed or may not include the variant in question)
- ✓ If there are other variants at this location published in ClinVar
- \checkmark Many in silico predictions

Things to consider with Franklin:

- They use PM2 at the moderate level, which is no longer suggested by ClinGen Guidelines. PM2 should only be used at the supporting level (please see: <u>https://clinicalgenome.org/site/assets/files/5182/pm2 - svi recommendation -</u> <u>approved sept2020.pdf</u>)
- They may overcall PM1 (mutational hotspots). They just look at if there are other reported LP/P variants in the area instead of if there are published functional domains. Example (MYH7 p, NM_000257.4:c.4258C>T/p.Arg1420Trp :

Function	al Data
√ PM1 ■ Moderate	Pathogenic Moderate: Non-truncating non-synonymous variant is located in a mutational hot spot and/or critical and well-established functional domain Close Details
Edit	Non-truncating non-synonymous variant
	Exonic hotspot
	13 pathogenic or likely pathogenic reported variants were found in a 184bp region surrounding this variant in exon 31 within the region 23417502-23417686 without any missense benign variants

The MYH7 gene has a ClinGen-approved mutational hotspot at amino acids 181-937. This would not qualify for PM1 based on the ClinGen rules.

• They may also overcall PP2, as it is just based on statistical analysis of ClinVar reports (which may not be accurate)

✓ PP2 .11 Supporting Edit	Pathogenic Supporting: Missense variant in a gene with low rate of benign missense mutations and for which missense mutation is a common mechanism of a disease Close Details
	SThree times more pathogenic variants than benign variant for curated missense variants in this gene
	High number of pathogenic missense variants in this gene
	Number of pathogenic missense variants: 569 Number of benign missense variants: 32
	Gnomad constraint of missense upper Z-score for gene is greater than 3.09 Gene score: 6.7889

o They may use PP5 and BP6 based on ClinVar reports (which may not be accurate)

ØPP5	Pathogenic Strong: Reputable source recently reports variant as pathogenic, but the evidence is not available
strong Edit	to the laboratory to perform an independent evaluation Close Details
	Reported as pathogenic:
	 By Cliniver CliniVar (Ds. 223176 Number of submissions' 1 pathogenic Number of submissions from 2015 and above: 1 pathogenic Number of submissions before 2015. 0 Latest submissions: 2/2016, psthogenic, SCV000264696, PCV000206825 By UniProt
	© Based on overall submissions rule strength was given pathogenic strong

VarSome

Website: <u>https://varsome.com/</u>

Description: A variant curation tool, which compiles information about the variant from multiple other databases (such as gnomAD/ClinVar).

How to use: Enter your variant into the search tool. (Suggested nomenclature is "Transcript : c." such as: NM_000368.5:c.359T>C).

Varsome will then give you links and information from many different databases. If you click on "Germline Classification" it will suggest codes to use.

chr14-23417598-G-A (MYH7 p.R	1420W)		-Submit to C	ClinVar 🛛 🗈 Link publication 🖉 🏳 Classif	y 🗈 Share 🗞 API Link 🏠 Favorites
General Information SNV MYH7(NM_000267.4):c.4258C>T p.(Arg1420Trp)	Princip Dell Only available in Premium	Germline Classification	Frequencies exomes: f = 0.000013 (cov: 49.0) genomes: f = 0.0000131 (cov: 31.9)	Conservation Scores phyloP100: 1.774	Structural Vaniants 🗐 🤨
Genes MYH7 @ (+2 more)	Transcripts NM_000257.4 - missense MANE Select	ClinVar Likely Pathogenic ***** 6 7 2	Michigo No data available	In-Silico Predictors PP3: Pathogenic 16 7 11 1	Beacon Network
Community Contributions	Region Browser 🚖	LOVD Only available in Premum	Combress Winster (Dombase No cata available	ClinGen Likely Pathogenic Expert Panel: Cardiomyopathy VCEP	Protein Viewer 📦
Publications Variant: 90 Gene: 740	Expression Data Top: heart_left_ventricle Tissues: 53	Uniprot Variants Pathogenic	сполівств Опар акулованій по Розпонскої	Copacio I Justi Consta suvertantes	

Varsome can tell you:

- ✓ Population frequency from databases (gnomAD)
- ✓ ClinVar entries
- ✓ Publications associated with variant (not always correct publications may be missed or may not include the variant in question)
- ✓ Many in silico predictions

Things to consider with VarSome:

- o May overuse PP5 (reputable source calls Pathogenic)
- They may overcall PM1 (mutational hotspots). They just look at if there are other reported LP/P variants in the area instead of if there are published functional domains (similarly to Franklin above). Example: it uses a Very Strong code just based off of ClinVar Submissions:



Clir/Var classifies this variant as Pathogenic, 2 stars (multiple consistent, reviewed Dec '23, 21 submissions), citing 11 articles (35653365, 34542152, 32894683, 31589614, 31513939 and 6 more). UniProt Variants classifies this variant as Pathogenic, citing 2 articles (16938236 and 15358028),

Show failed criteria

Glossary

Genotype	A score of the type of variants present at a specific location within the genome (i.e. What variants are present in a specific gene in a patient, and how many copies of each are there?) <i>Example: e3/e3 is the most common genotype for APOE.</i>
Phenotype	 The manifestation of the disease of interest in the patient. Example: The patient's phenotype involves ectopia lentis, aortic dilatation, pectus carinatum, and arachnodactyly. They would be a good candidate for genetic testing for Marfan syndrome.
Affected	 A descriptor of an individual with the phenotype. Example: Has this variant been found in individuals affected with cystic fibrosis?
Unaffected	 A descriptor of an individual without the phenotype. For some phenotypes, evaluations may need to be done to confirm whether someone is truly unaffected. <i>Example: Although the patient's mother passed away from Huntington's disease, he was very surprised that the cause of her death was genetic because all three of his mother's siblings were unaffected, and his maternal grandparents died in a car crash when they were younger than the age of onset for Huntington's disease.</i>
Allele Frequency	 The frequency at which a specific genetic variant is seen in the general population or in certain ancestral groups <i>Example: The allele frequency of this benign variant in European populations is 5%.</i>
Proband	 A patient that has the variant of interest and the phenotype associated with that variant or the gene that the variant is in. There can only be one proband within each affected family, regardless of whether there are other individuals with both the variant and phenotype of interest. <i>Example: The proband is a 45 year old woman, assigned female at birth, with a personal history of bilateral breast cancer, and a family history of breast and ovarian cancer. She was identified to have the p.Ala295Thr variant in BRCA2.</i>
Segregations	 Refers to family members of a proband who also have the variant and phenotype of interest. Can also be used as a verb to describe whether the variant has been identified in affected individuals in the family. <i>Example: The proband's mother has a personal history of breast cancer, and her maternal aunt was recently diagnosed with epithelial ovarian cancer. Both of them underwent genetic testing and were identified to have the p.Ala295Thr variant in BRCA2, which had previously been identified in the proband.</i>

	Therefore, this variant segregated with disease in two affected individuals in this family.
in cis	 A descriptor for variants inherited from the same parent of a proband. <i>Example: The two variants in the CFTR gene identified in our patient were inherited from the patient's father. Those variants were therefore inherited</i> in cis.
in trans	 A descriptor for variants inherited from different parents of a proband. Typically, for autosomal recessive conditions, two variants must be inherited <i>in trans</i> to cause the phenotype. <i>Example: One variant in the CFTR gene was inherited from the patient's father, and another variant in the CFTR gene was inherited from the patient's mother; therefore, the variants were inherited in trans.</i>
De novo	 A descriptor for variants that were not identified in either of a patient's parents. A variant is assumed <i>de novo</i> if parentage (confirmation of whether an individual's parents are genetically their parents) has not been confirmed, and it is confirmed <i>de novo</i> if parentage has been confirmed. <i>Example:</i> Both parents were tested for the novel variant identified in our patient, and neither of them have the variant. The lab also confirmed parentage of both parents. This variant was identified de novo in our patient.
Phasing	 A verb to describe the process of determining whether variants are <i>in cis, in trans,</i> or if a variant is <i>de novo.</i> Also a general descriptor for how variants are inherited in reference to a proband's parents. <i>Example: To help us determine the likelihood that this CFTR VUS may be pathogenic, we should determine the phasing of this variant in reference to the other pathogenic variant that our patient has. Are the parents available for testing?</i>

If you have any questions or feedback on this handout, would like to suggest changes or inquire about additional training, please contact the Variant Curation Task Force at variantcurationrequests@mgb.org.

Thank you to the following people for helping in the development of this handout: Lauren Briere, MS, CGC, Helen Chen, MS, CGC, Emma Henricks, MS, CGC, Anna Nagy, MS, CGC, Chelsea Stevens, MS, CGC, the Variant Curation Task Force members