

Clinical Variant Interpretation

GMS Workshop - 18th Nov 2021

Nicky Whiffin: nwhiffin@well.ox.ac.uk

Structure of the sessions

Session 1 - Variant annotation and inheritance

Session 2 - Allele frequency and penetrance

Session 3 - Variant interpretation with ACMG

~ 40-60 mins each including an introductory talk and a practical session

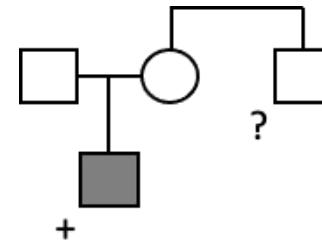
Session 1: Variant annotation and inheritance

Scene setting - the value of a genetic diagnosis

1 Accurate diagnosis



2 Familial screening



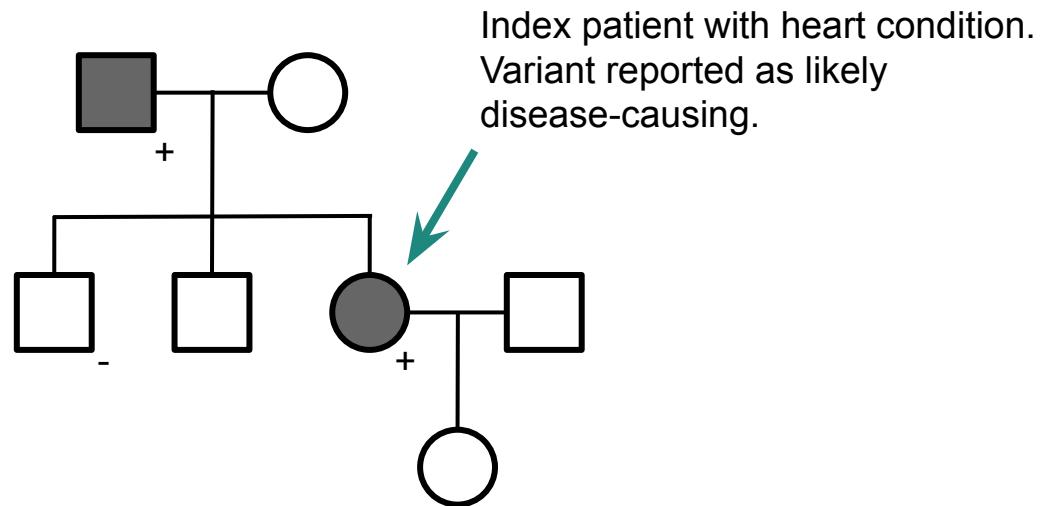
3 Inform prognosis



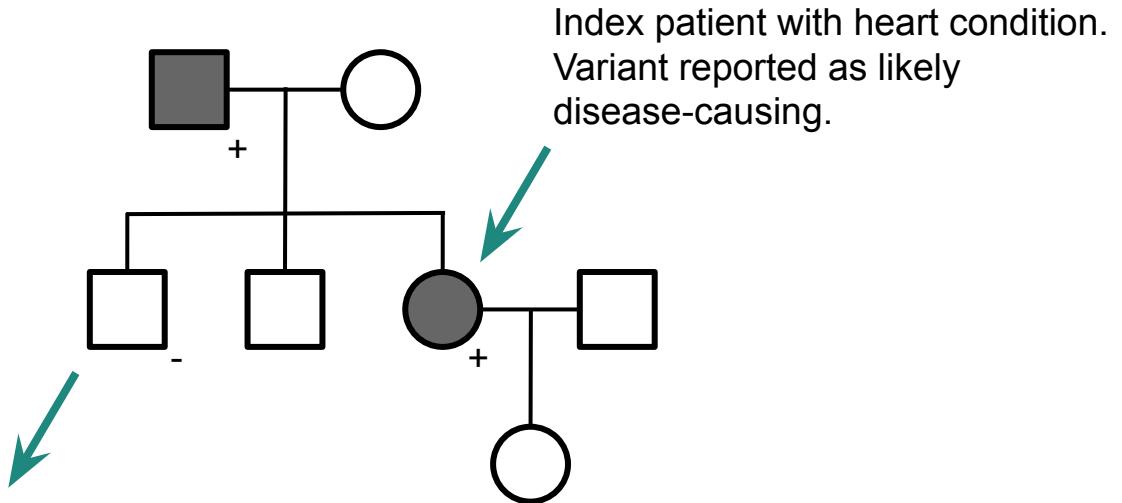
4 Personalised treatment



... and the importance of not getting it wrong



... and the importance of not getting it wrong



The scale of the problem

>1 million variants per person

~1% are protein-coding

Assume one disease causing variant (normally)

Which one?



A forensic investigation - for each variant

- Have we seen it before?
- What do we already know?
- What can we predict (similar variants)?
- What new data can we add?
 - From this family
 - Experiments

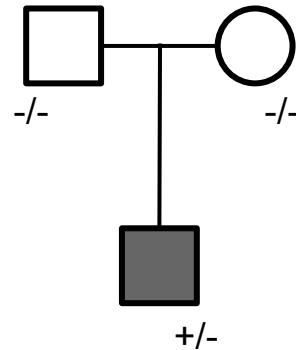
Clue 1: modes of inheritance

The way a disease tracks in a family (if at all) can help limit variants

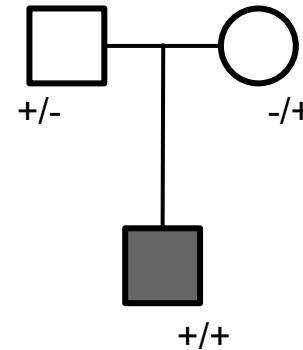
Clue 1: modes of inheritance

The way a disease tracks in a family (if at all) can help limit variants

E.g. trios



De novo

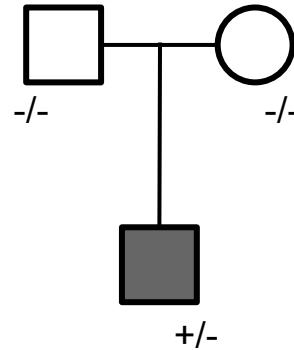


Recessive

Clue 1: modes of inheritance

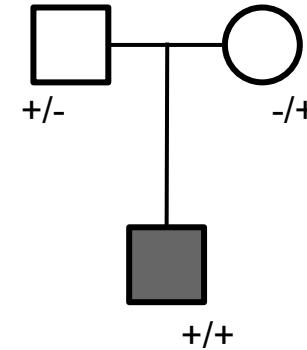
The way a disease tracks in a family (if at all) can help limit variants

E.g. trios



De novo

~70 per individual - high prior

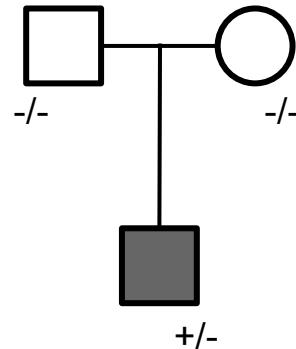


Recessive

Clue 1: modes of inheritance

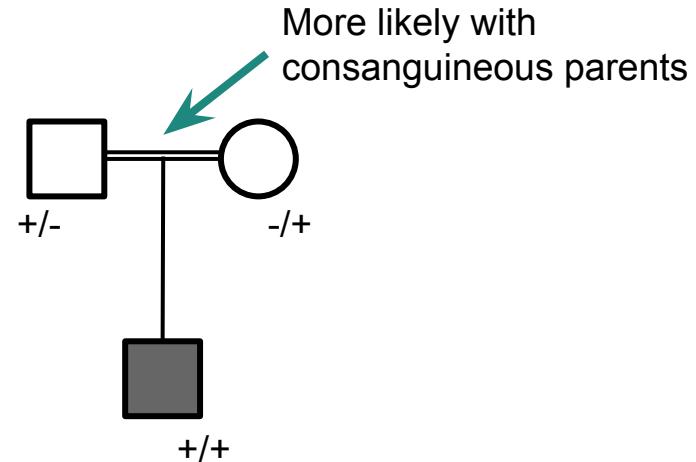
The way a disease tracks in a family (if at all) can help limit variants

E.g. trios



De novo

~70 per individual - high prior

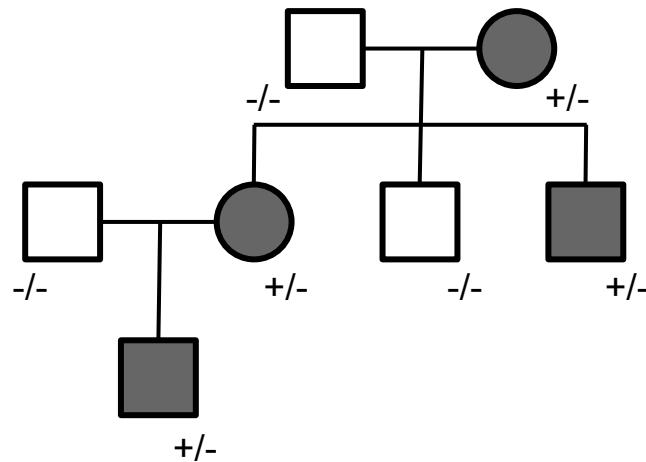


Recessive

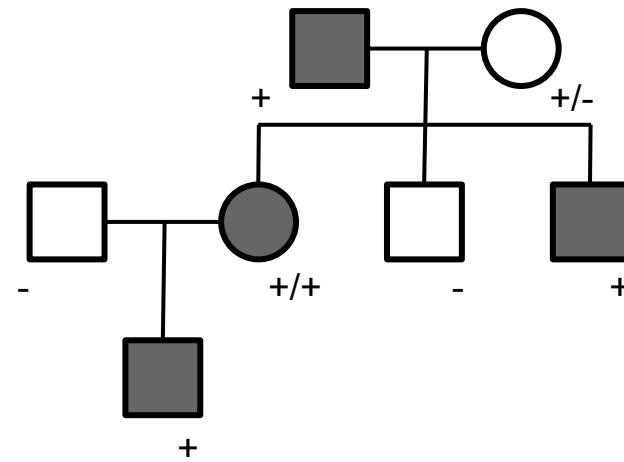
Clue 1: modes of inheritance

The way a disease tracks in a family (if at all) can help limit variants

E.g. larger families



Autosomal dominant

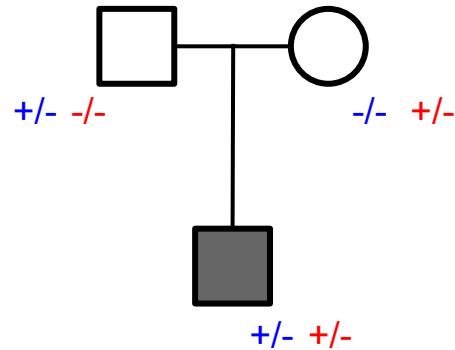


X-linked recessive

Clue 1: modes of inheritance

The way a disease tracks in a family (if at all) can help limit variants

A note on compound heterozygotes

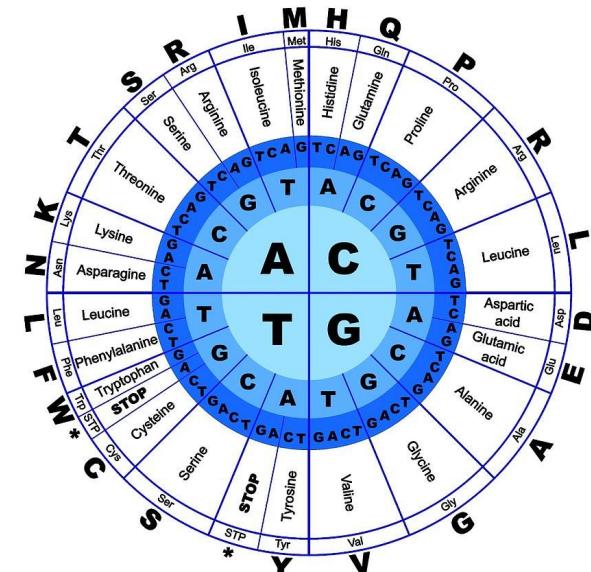


Recessive

Not all variants are equal

- ‘silent’ - no effect on protein sequence or expression - synonymous
- Change a single amino acid - missense / nonsynonymous
- ‘loss-of-function’ / ‘truncating’ / ‘PTV’ - generally lead to NMD
 - Stop-gained / nonsense
 - Essential splice site
 - Frameshift
- Regulatory - (partial) changes in protein levels
- Start-loss / Stop-loss

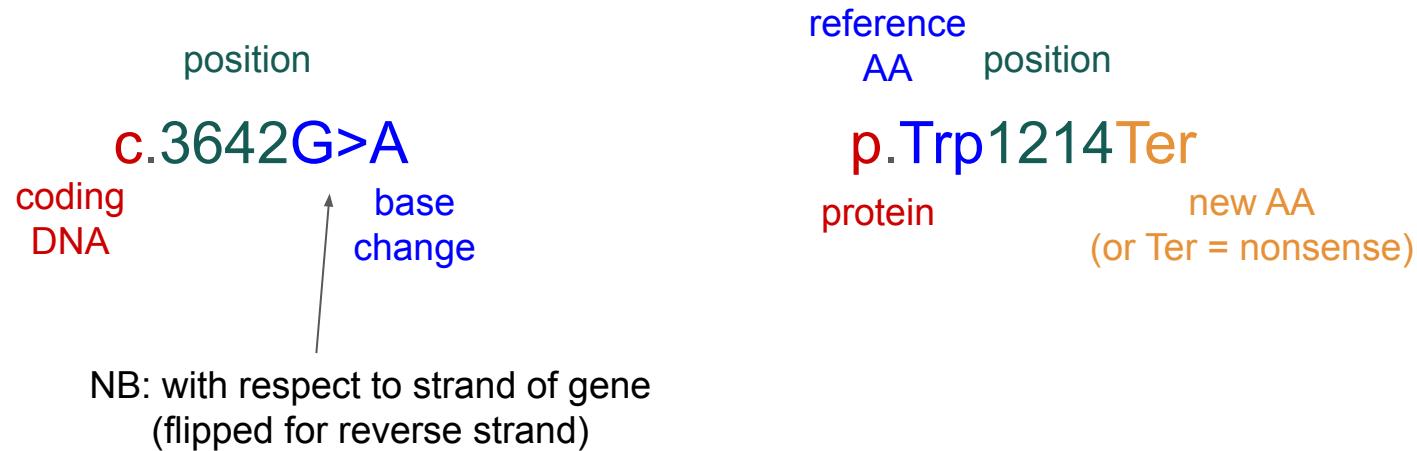
Impact is gene / region dependent



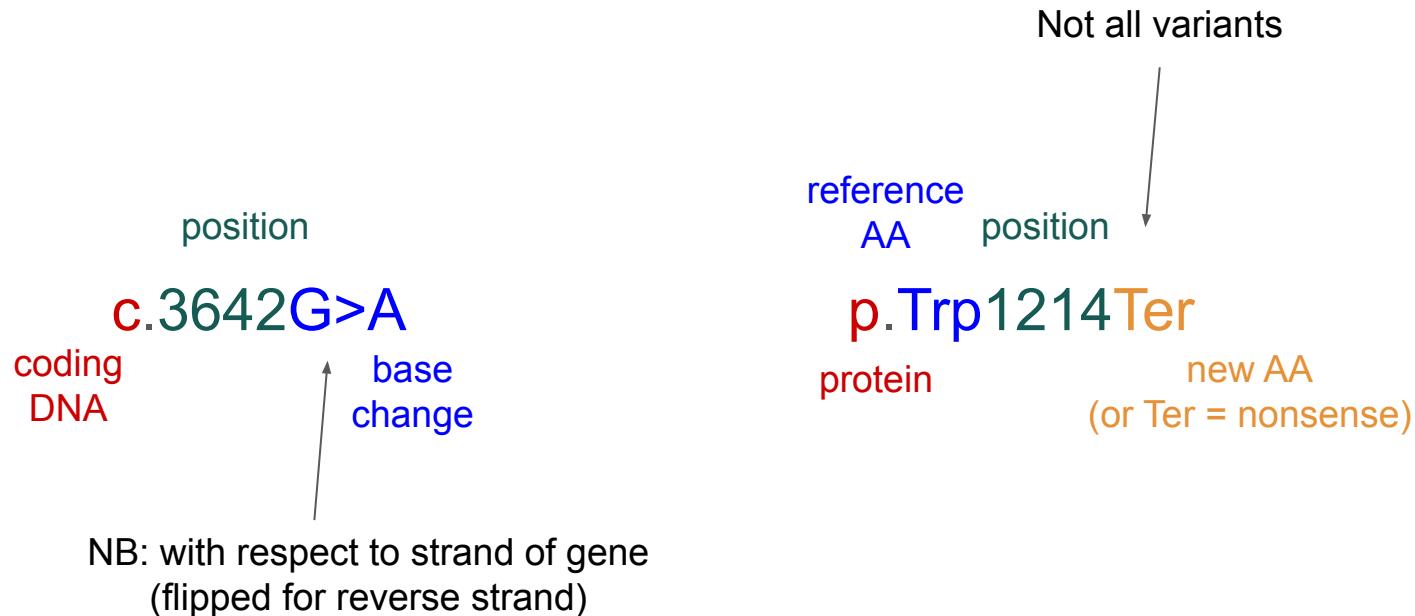
Variant naming nomenclature - HGVS

position	reference AA	position
c.3642G>A	p.Trp	1214Ter
coding DNA	protein	new AA (or Ter = nonsense)

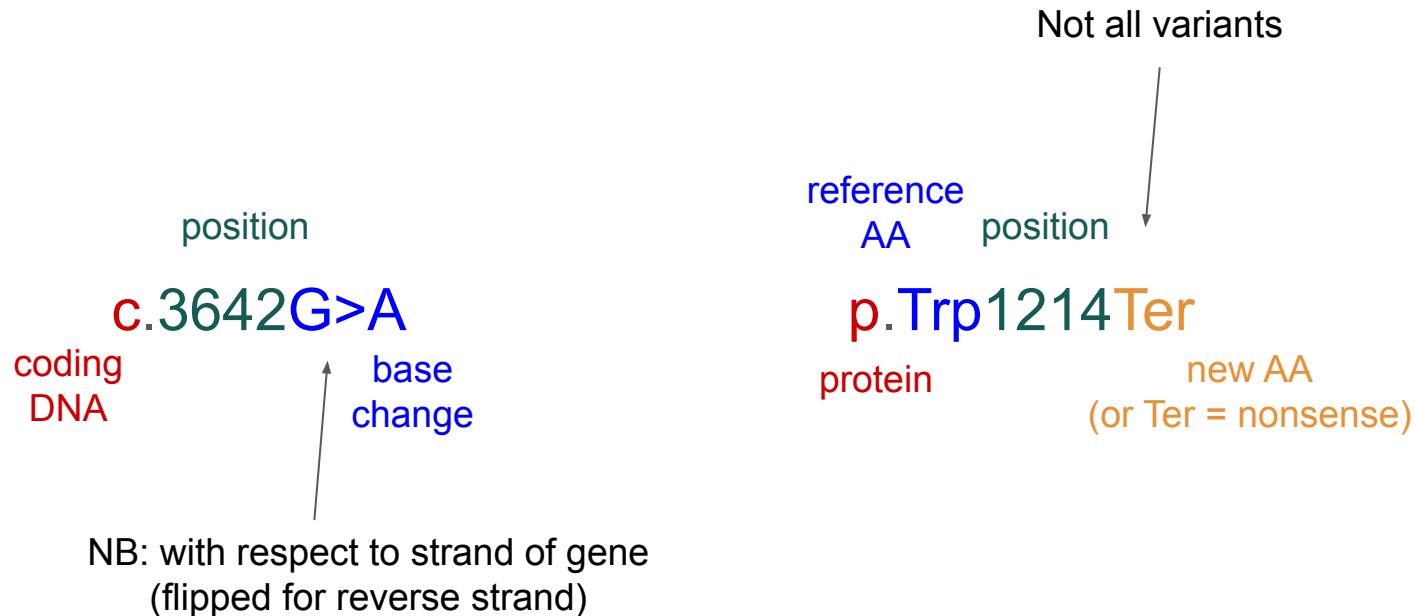
Variant naming nomenclature - HGVS



Variant naming nomenclature - HGVS



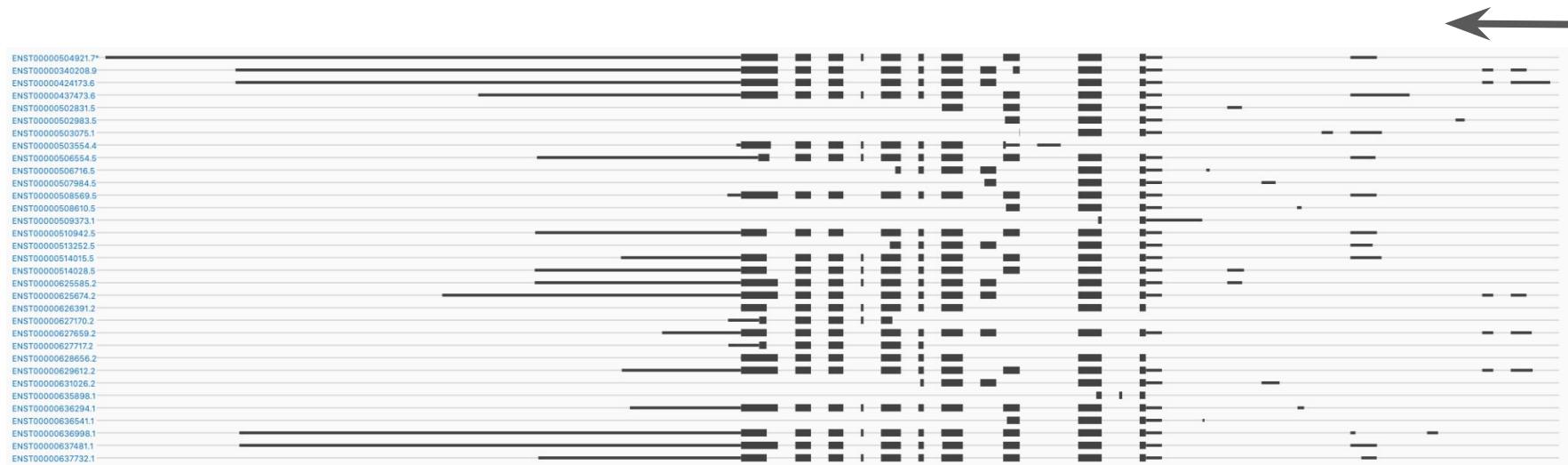
Variant naming nomenclature - HGVS



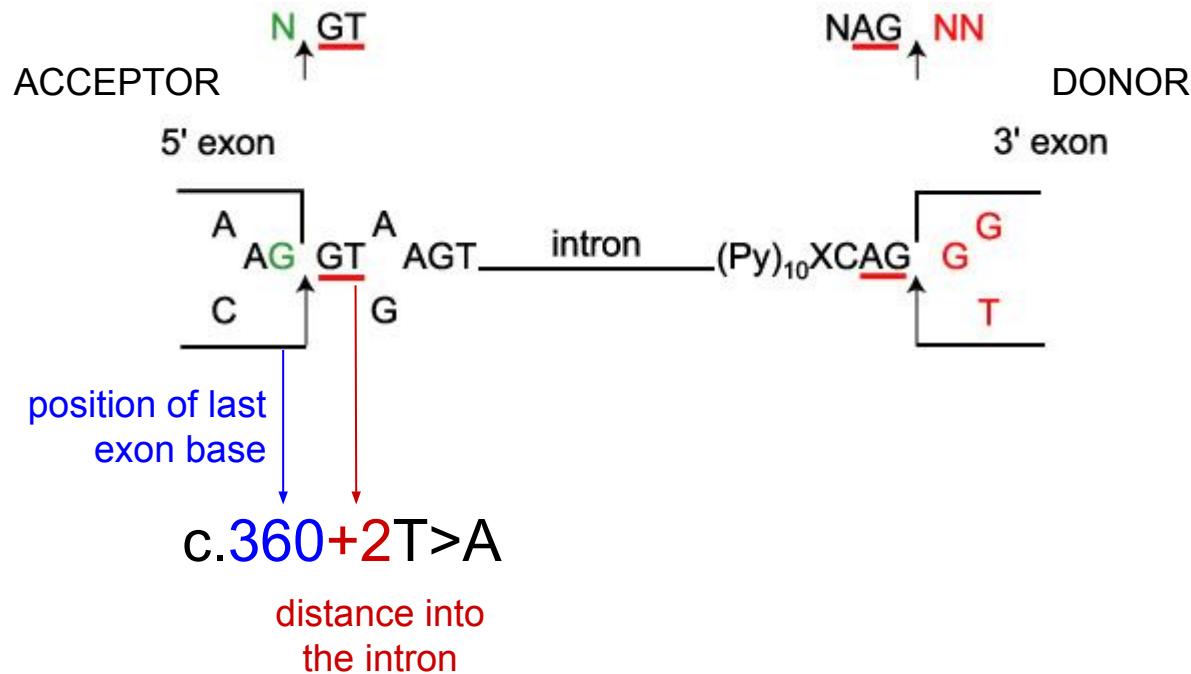
Transcript dependent - each variant can have multiple descriptors. Should always also quote transcript.

Annotation - transcripts

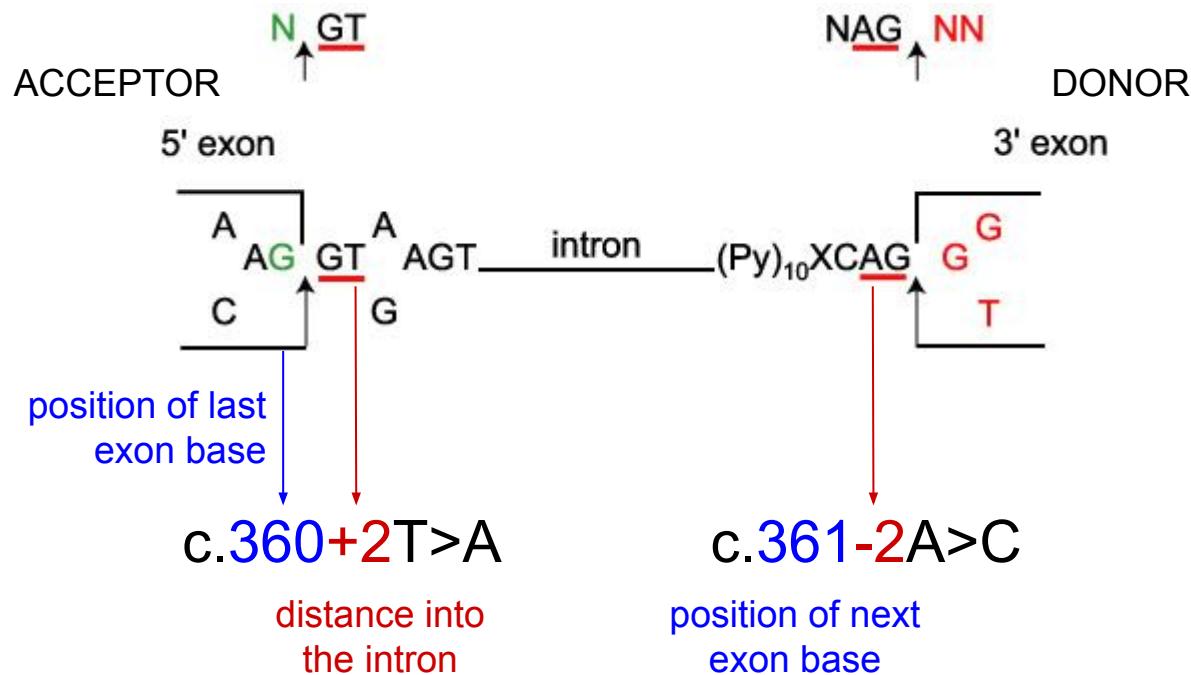
E.g. *MEF2C*



Variant naming nomenclature - intronic variants



Variant naming nomenclature - intronic variants



MANE transcripts

Matched Annotation from NCBI and EMBL-EBI (MANE)

- [What is MANE?](#)
- [Rationale](#)
- [MANE Select](#)
- [MANE Select Methodology](#)
 - [Choosing the transcript](#)
 - [Matching transcript ends](#)
- [Salient features of MANE Select transcripts](#)
- [Manual curation of MANE data](#)
- [Accessing MANE Select data](#)
- [Contact information](#)

What is MANE?

Matched Annotation from NCBI and EMBL-EBI (MANE) is a collaboration between the [National Center for Biotechnology Information](#) (NCBI) and the [European Molecular Biology Laboratories-European Bioinformatics Institute](#) (EMBL-EBI). The goal of this project is to provide a minimal set of matching RefSeq and Ensembl transcripts of human protein-coding genes, where the transcripts from a matched pair are identical (5' UTR, coding region and 3' UTR), but retain their respective identifiers. The MANE transcript set is classified into two groups:

1. MANE Select: One high-quality representative transcript per protein-coding gene that is well-supported by experimental data and represents the biology of the gene.
2. MANE Plus Clinical: Transcripts chosen to supplement MANE Select when needed for clinical variant reporting.

Watch the [MANE webinar on YouTube!](#)

Annotation - Ensembl VEP

Ensembl Variant Effect Predictor (VEP)

VEP determines the effect of your variants (SNPs, insertions, deletions, CNVs or structural variants) on genes, transcripts, and protein sequence, as well as regulatory regions.

Simply input the coordinates of your variants and the nucleotide changes to find out the:

- **Genes and Transcripts** affected by the variants
- **Location** of the variants (e.g. upstream of a transcript, in coding sequence, in non-coding RNA, in regulatory regions)
- **Consequence** of your variants on the protein sequence (e.g. stop gained, missense, stop lost, frameshift), see [variant consequences](#)
- **Known variants** that match yours, and associated minor allele frequencies from the 1000 Genomes Project
- SIFT and PolyPhen-2 scores for changes to protein sequence
- ... And more! See [data types](#), [versions](#).

★ [What's new in release 104?](#)

VEP interfaces

Web interface



- Point-and-click interface
- Suits smaller volumes of data

[Documentation](#)

 **Launch VEP**

Command line tool



- More options and flexibility
- For large volumes of data

[Documentation](#)

 [Clone from GitHub](#)  [Download \(zip\)](#)  [Pull Docker image from DockerHub](#)

REST API

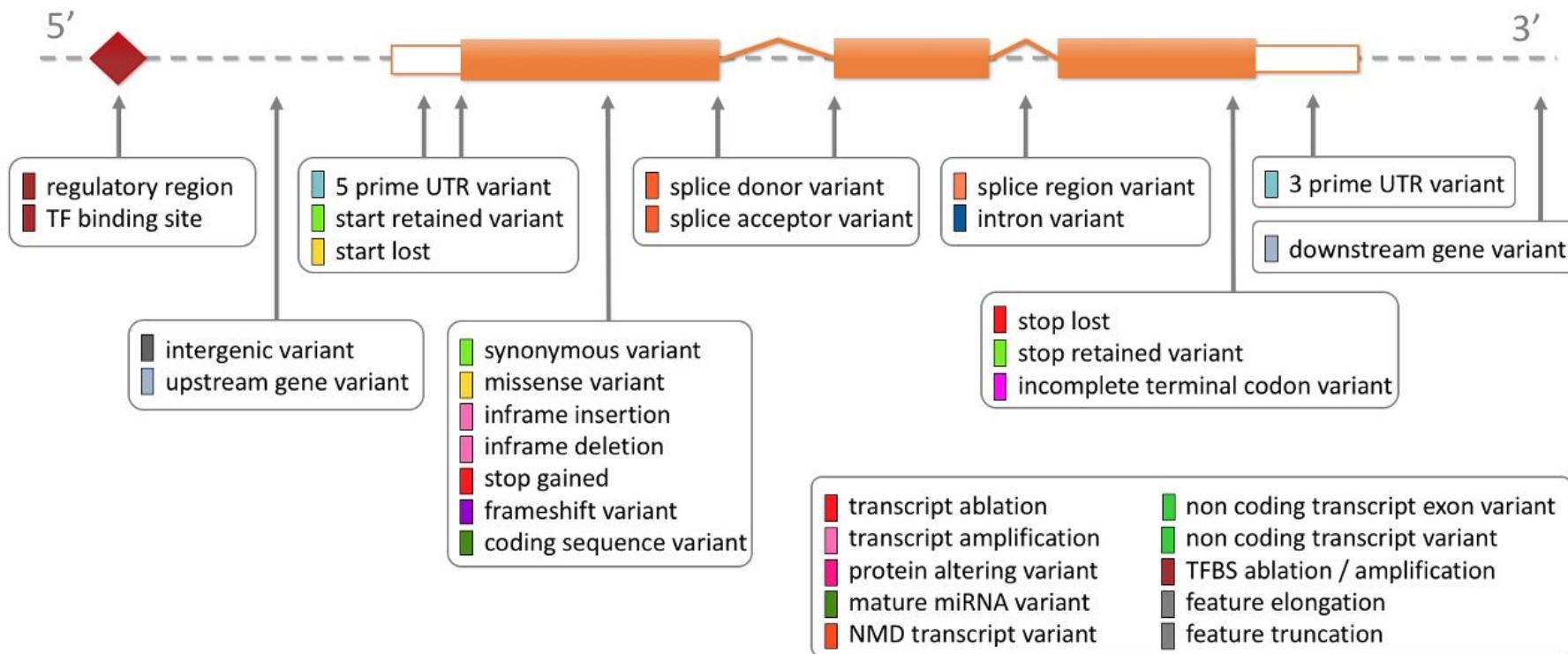


- Language-independent API
- Simple URL-based queries

[Documentation](#)

 [VEP REST API](#)

Annotation - Ensembl VEP



Practical 1: Annotating variants with VEP

- Get to grips with HGVS
(<https://varnomen.hgvs.org/recommendations/general/>)
- VEP annotation - understand output

Session 2: Allele frequency and penetrance

gnomAD

Browser demo

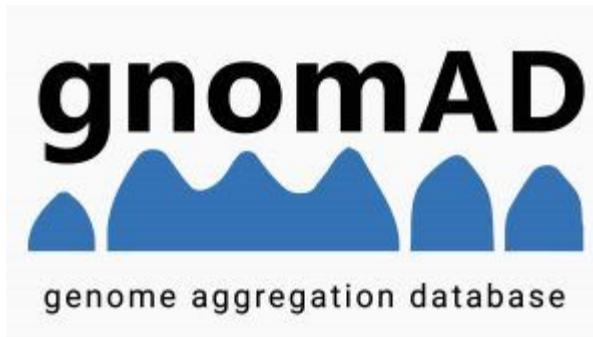
Allele frequency is one of the strongest predictors of benignity, but is NOT an indication of pathogenicity

Allele frequency is one of the strongest predictors of benignity, but is NOT an indication of pathogenicity

But how common is too common?

Population reference datasets

NOT disease-free controls - expect disease at ~ same frequency as population



~125k exomes in v2.1.1

~76k genomes in v3.1



~500k individuals

Let's start with an example

Single nucleotide variant: 14-23886482-G-C(GRCh37) Copy variant ID Dataset gnomAD v2.1.1 ?

Filter	Exomes	Genomes	Total	External Resources	missense
Allele Count	34	No variant	34	<ul style="list-style-type: none">dbSNP (rs397516214)UCSCClinVar (43015)ClinGen Allele Registry (CA14936)	1. MYH7
Allele Number	251430		251430		1. ENST00000355349.3
Allele Frequency	0.0001352		0.0001352		Ensembl canonical transcript for MYH7
Popmax Filtering AF ? (95% confidence)	0.0008165				HGVSp: p.Leu1467Val
Number of homozygotes	0		0		Domains: PF01576 (Pfam), and 2 more
Mean depth of coverage	96.0	32.6			Polyphen: probably_damaging
Feedback					
Report an issue with this variant					

Population Frequencies ?

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
South Asian	34	30616	0	0.001111
African/African American	0	16252	0	0.000
Latino/Admixed American	0	34590	0	0.000
Ashkenazi Jewish	0	10078	0	0.000
East Asian	0	18394	0	0.000
European (Finnish)	0	21642	0	0.000
European (non-Finnish)	0	113718	0	0.000
Other	0	6140	0	0.000
XX	5	115534	0	0.00004328
XY	29	135896	0	0.0002134
Total	34	251430	0	0.0001352

Could this variant
cause HCM?

A generic approach

Allele frequency (AF) = number of alleles with variant / total number of alleles

- Dominant disease - < 0.1%
- Recessive disease - < 1%

But surely the frequency (prevalence) of disease is important?

NB: Assuming dominant

A disease-specific statistical approach

Allele frequency \leq disease prevalence

The issue of penetrance

The proportion of individuals with a genetic variant that have/develop the disease

- Measured on variant or gene level
- Difficult to measure
 - Thresholds of disease (penetrance vs expressivity)
 - Family based measures are over-estimates
 - Biobanks likely underestimate
 - Variants are often very rare / unique to individual families

NB: Assuming dominant

A disease-specific statistical approach

Allele frequency \leq Disease prevalence

Disease prevalence = Allele frequency x Penetrance

NB: Assuming dominant

A disease-specific statistical approach

Allele frequency \leq Disease prevalence

$$\text{Disease prevalence} = \sum_{\text{All variants}} (\text{Allele frequency} \times \text{Penetrance})$$

NB: Assuming dominant

A disease-specific statistical approach

Allele frequency \leq Disease prevalence

$$\text{Disease prevalence} = \sum_{\text{All variants}} (\text{Allele frequency} \times \text{Penetrance})$$

$$\text{Max AF} = \frac{\text{Disease prevalence} \times \text{heterogeneity}}{\text{Penetrance}}$$

heterogeneity = maximum proportion of disease attributable to single variant

NB: Assuming dominant

A disease-specific statistical approach

MYBPC3:c.1504C>T
causes 2.2% (1.6-3.0%)
of European HCM cases

$$(0.5 \times 1/500) \quad 0.03$$

Max AF = Disease prevalence x heterogeneity

Penetrance

0.5

NB: Assuming dominant

A disease-specific statistical approach

6x10⁻⁵

(0.5 x 1/500)

0.03

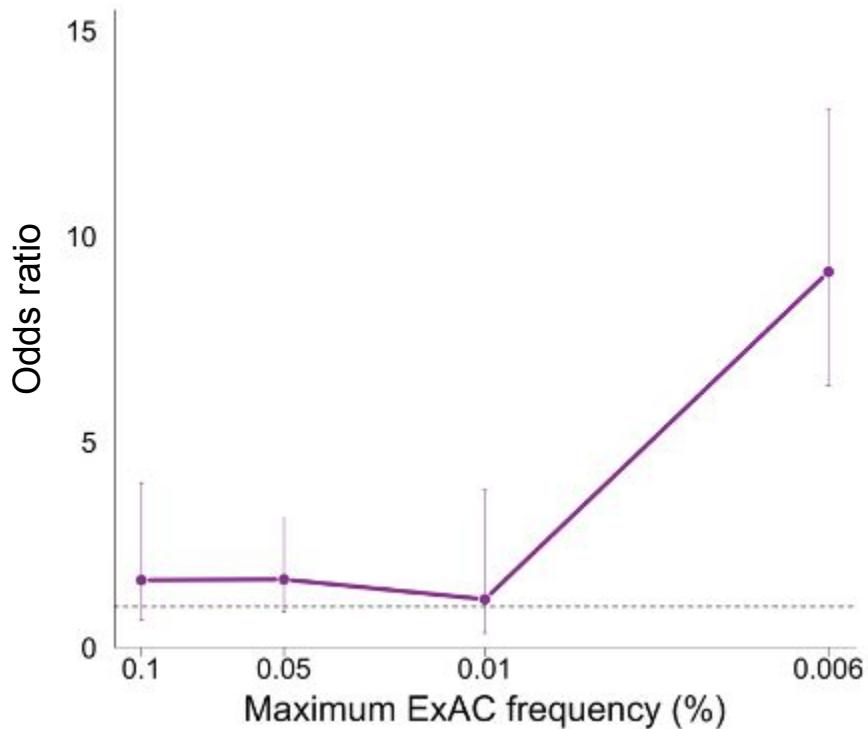
Max AF = Disease prevalence x heterogeneity

Penetrance

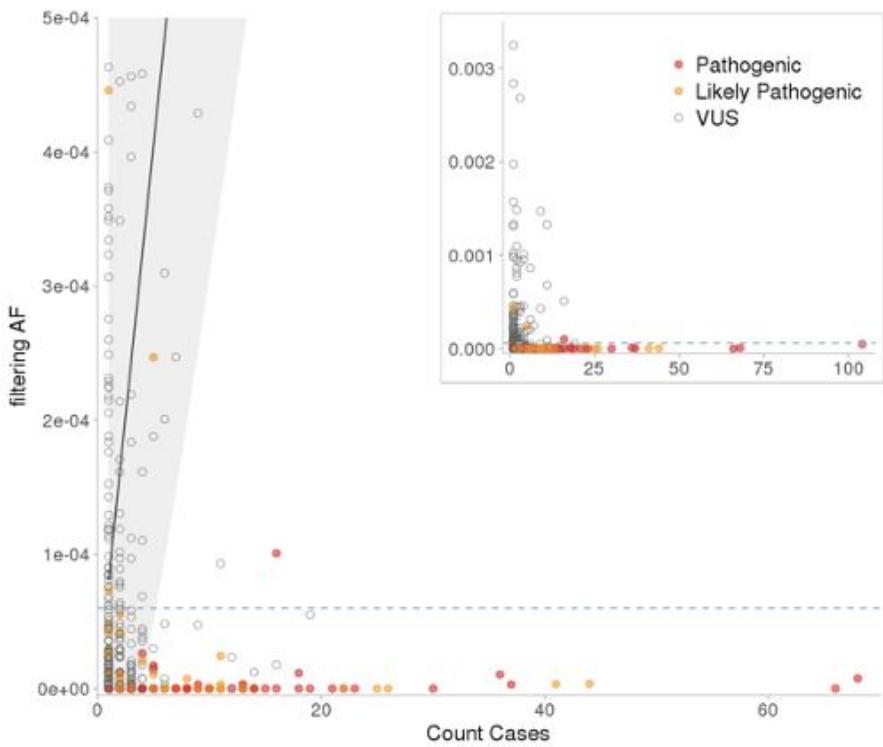
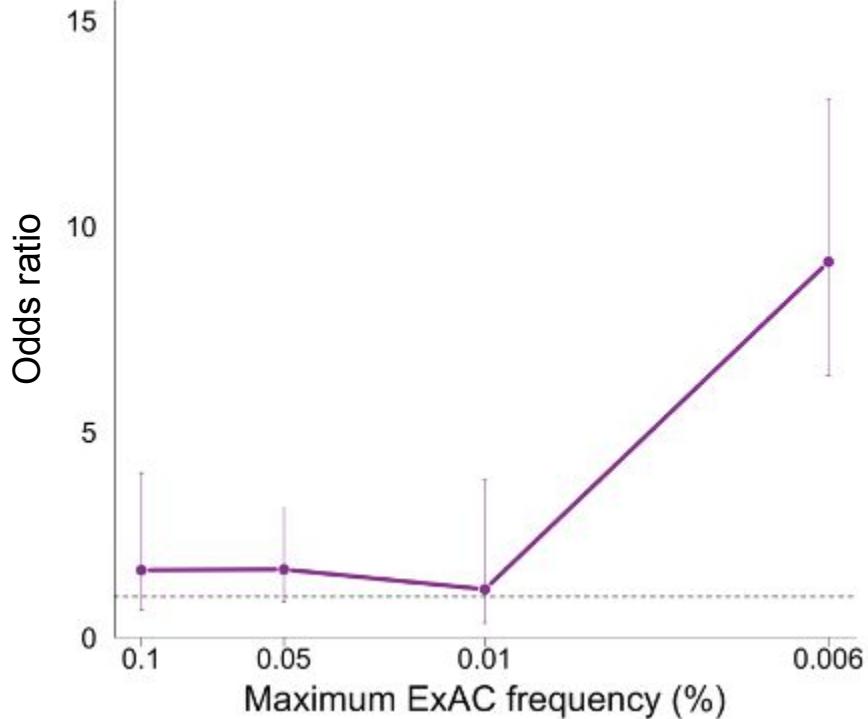
0.5

MYBPC3:c.1504C>T
causes 2.2% (1.6-3.0%)
of European HCM cases

Stricter filtering enriches for disease-causing variants



Stricter filtering enriches for disease-causing variants



Heading back to that example

Single nucleotide variant: 14-23886482-G-C(GRCh37) Copy variant ID Dataset gnomAD v2.1.1 ?

Filter	Exomes	Genomes	Total
Allele Count	34	No variant	34
Allele Number	251430		251430
Allele Frequency	0.0001352		0.0001352
Popmax Filtering AF (95% confidence)	0.0008165		
Number of homozygotes	0		0
Mean depth of coverage	96.0	32.6	

External Resources

- dbSNP (rs397516214)
- UCSC
- ClinVar (43015)
- ClinGen Allele Registry (CA14936)

missense

1. **MYH7**

1. ENST00000355349.3
Ensembl canonical transcript for MYH7
HGVS: p.Leu1467Val
Domains: PF01576 (Pfam), and 2 more
Polyphen: probably_damaging
SIFT: deleterious

Feedback

Report an issue with this variant

Population Frequencies ?

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
South Asian	34	30616	0	0.001111
African/African American	0	16252	0	0.000
Latino/Admixed American	0	34590	0	0.000
Ashkenazi Jewish	0	10078	0	0.000
East Asian	0	18394	0	0.000
European (Finnish)	0	21642	0	0.000
European (non-Finnish)	0	113718	0	0.000
Other	0	6140	0	0.000
XX	5	115534	0	0.00004328
XY	29	135896	0	0.0002134
Total	34	251430	0	0.0001352

> 0.00006
Too common?

Heading back to that example

Single nucleotide variant: 14-23886482-G-C(GRCh37) Copy variant ID Dataset gnomAD v2.1.1 ?

Filter	Exomes	Genomes	Total
Allele Count	34	No variant	34
Allele Number	251430		251430
Allele Frequency	0.0001352		0.0001352
Popmax Filtering AF (95% confidence)	0.0008165		
Number of homozygotes	0		0
Mean depth of coverage	96.0	32.6	

External Resources

- dbSNP (rs397516214)
- UCSC
- ClinVar (43015)
- ClinGen Allele Registry (CA14936)

missense

1. **MYH7**

1. ENST00000355349.3
Ensembl canonical transcript for MYH7
HGVS: p.Leu1467Val
Domains: PF01576 (Pfam), and 2 more
Polyphen: probably_damaging
SIFT: deleterious

Feedback

Report an issue with this variant

Population Frequencies

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
South Asian				0.001111
African/African American				0.000
Latino/Admixed American				0.000
Ashkenazi Jewish				0.000
East Asian				0.000
European (Finnish)				0.000
European (non-Finnish)				0.000
Other				0.000
XX				0.0004328
XY				0.0002134
Total				0.0001352

NM_000257.4(MYH7):c.4399C>G (p.Leu1467Val)

Interpretation: Conflicting interpretations of pathogenicity
Benign(1);Uncertain significance(5)

Review status: ★☆☆☆ criteria provided, conflicting interpretations

Submissions: 7 (Most recent: Jul 20, 2021)

Last evaluated: Feb 22, 2021

Accession: VCV000043015.10

Variation ID: 43015

Description: single nucleotide variant

Variant details

NM_000257.4(MYH7):c.4399C>G (p.Leu1467Val)

> 0.00006
Too common?

Heading back to that example

Single nucleotide variant: 14-23886482-G-C(GRCh37) Copy variant ID Dataset gnomAD v2.1.1 ?

Filter	Exomes	Genomes	Total
Allele Count	34	No variant	34
Allele Number	251430		251430
Allele Frequency	0.0001352		0.0001352
Popmax Filtering AF ? (95% confidence)	0.0008165		
Number of homozygotes	0		0
Mean depth of coverage	96.0	32.6	

External Resources

- dbSNP (rs397516214)
- UCSC
- ClinVar (43015)
- ClinGen Allele Registry (CA14936)

missense

1. **MYH7**

1. [ENST00000355349.3](#)
Ensembl canonical transcript for MYH7
HGVS: p.Leu1467Val
Domains: [PF01576 \(Pfam\)](#), and 2 more
Polyphen: probably_damaging
SIFT: deleterious

Feedback

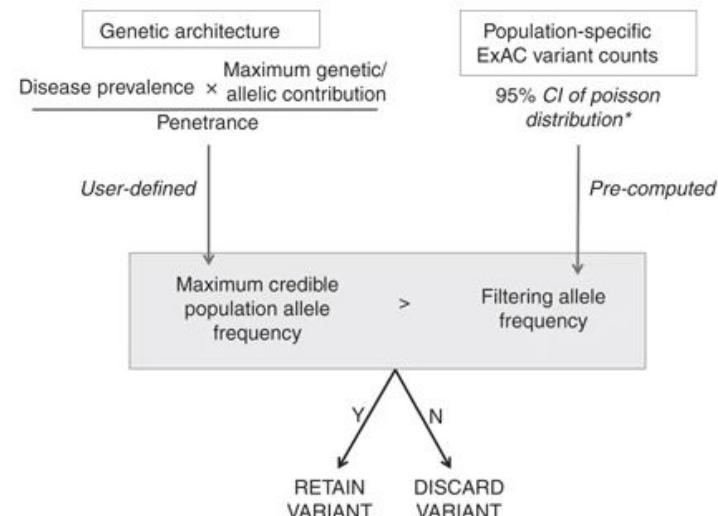
[Report an issue with this variant](#)

Population Frequencies ?

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
South Asian	34	30616	0	0.001111
African/African American	0	16252	0	0.000
Latino/Admixed American	0	34590	0	0.000
Ashkenazi Jewish	0	10078	0	0.000
East Asian	0	18394	0	0.000
European (Finnish)	0	21642	0	0.000
European (non-Finnish)	0	113718	0	0.000
Other	0	6140	0	0.000
XX	5	115534	0	0.00004328
XY	29	135896	0	0.0002134
Total	34	251430	0	0.0001352

Popmax filtering AF (FAF)

- Reference datasets are a sub-sample of the population
- What is the highest AF that is compatible with the observed AC and AN at 95% confidence?
- For each population separately - take maximum
- If Popmax FAF > Max AF = too common



Learn more

Open Access | Published: 18 May 2017

Using high-resolution variant frequencies to empower clinical genome interpretation

[Nicola Whiffin](#), [Eric Minikel](#), [Roddy Walsh](#), [Anne H O'Donnell-Luria](#), [Konrad Karczewski](#), [Alexander Y Ing](#),
[Paul J R Barton](#), [Birgit Funke](#), [Stuart A Cook](#), [Daniel MacArthur](#) & [James S Ware](#) 

Genetics in Medicine **19**, 1151–1158 (2017) | [Cite this article](#)

18k Accesses | **171** Citations | **93** Altmetric | [Metrics](#)

Cautions

- Low coverage - absence may not mean rare
- Ethnicity is important - founder variants

Single nucleotide variant: 14-23885263-T-C(GRCh37) [Copy variant ID](#) Dataset gnomAD v2.1.1 [?](#)

Filter	Exomes	Genomes	Total	External Resources
Allele Count	Pass	Pass	16	<ul style="list-style-type: none">dbSNP (rs769259796)UCSCClinGen Allele Registry (CA44177)
Allele Number	251456	31398	282854	
Allele Frequency	0.00004375	0.0001592	0.00005657	
Popmax Filtering AF ? (95% confidence)	—	—		
Number of homozygotes	0	0	0	
Mean depth of coverage	95.9	32.7		

Population Frequencies [?](#)

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
European (Finnish)	15	25118	0	0.0005972
Other	1	7228	0	0.0001384
African/African American	0	24970	0	0.000
Latino/Admixed American	0	35440	0	0.000
Ashkenazi Jewish	0	10368	0	0.000
East Asian	0	19952	0	0.000
European (non-Finnish)	0	129162	0	0.000
South Asian	0	30616	0	0.000

Practical 2: Allele frequency filtering

- Explore gnomAD and get to grips with AF data

<https://gnomad.broadinstitute.org/>

- Pick a disease of interest

Examples:

- Search literature - disease prevalence and minimum penetrance
- Calculate maximum AF

Session 3: Variant interpretation with ACMG

The overall aim

Categorise into one of five classes

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

Limit VUS - not helpful, but with caution

The overall aim

Categorise into one of five classes

- Pathogenic
 - Likely Pathogenic
 - Uncertain Significance (VUS)
 - Likely Benign
 - Benign
- Clinically actionable

Limit VUS - not helpful, but with caution

The ACMG/AMP guidelines

Published: 05 March 2015

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

[Sue Richards PhD](#) , [Nazneen Aziz PhD](#), [Sherri Bale PhD](#), [David Bick MD](#), [Soma Das PhD](#), [Julie Gastier-Foster PhD](#), [Wayne W. Grody MD, PhD](#), [Madhuri Hegde PhD](#), [Elaine Lyon PhD](#), [Elaine Spector PhD](#), [Karl Voelkerding MD](#) & [Heidi L. Rehm PhD](#) on behalf of ; on behalf of the ACMG Laboratory Quality Assurance Committee

[Genetics in Medicine](#) **17**, 405–423 (2015) | [Cite this article](#)

234k Accesses | **9507** Citations | **295** Altmetric | [Metrics](#)

	Benign			Pathogenic		
	Strong	Supporting	Supporting	Moderate	Strong	Very Strong
Population Data	MAF is too high for disorder <i>BA1/BS1 OR</i> observation in controls inconsistent with disease penetrance <i>BS2</i>			Absent in population databases <i>PM2</i>	Prevalence in affecteds statistically increased over controls <i>PS4</i>	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i> Missense in gene where only truncating cause disease <i>BP1</i> Silent variant with non predicted splice impact <i>BP7</i>	Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i>	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before <i>PM5</i> Protein length changing variant <i>PM4</i>	Same amino acid change as an established pathogenic variant <i>PS1</i>	Predicted null variant in a gene where LOF is a known mechanism of disease <i>PVS1</i>
Functional Data	Well-established functional studies show no deleterious effect <i>BS3</i>		Missense in gene with low rate of benign missense variants and path. missenses common <i>PP2</i>	Mutational hot spot or well-studied functional domain without benign variation <i>PM1</i>	Well-established functional studies show a deleterious effect <i>PS3</i>	
Segregation Data	Non-segregation with disease <i>BS4</i>		Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data →		
De novo Data				<i>De novo</i> (without paternity & maternity confirmed) <i>PM6</i>	<i>De novo</i> (paternity & maternity confirmed) <i>PS2</i>	
Allelic Data		Observed in <i>trans</i> with a dominant variant <i>BP2</i> Observed in <i>cis</i> with a pathogenic variant <i>BP2</i>		For recessive disorders, detected in <i>trans</i> with a pathogenic variant <i>PM3</i>		
Other Database		Reputable source w/out shared data = benign <i>BP6</i>	Reputable source = pathogenic <i>PP5</i>			
Other Data		Found in case with an alternate cause <i>BP5</i>	Patient's phenotype or FH highly specific for gene <i>PP4</i>			

	Benign			Pathogenic		
	Strong	Supporting	Supporting	Moderate	Strong	Very Strong
Population Data	MAF is too high for disorder <i>BA1/BS1 OR</i> observation in controls inconsistent with disease penetrance <i>BS2</i>			Absent in population databases <i>PM2</i>	Prevalence in affecteds statistically increased over controls <i>PS4</i>	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i> Missense in gene where only truncating cause disease <i>BP1</i> Silent variant with non predicted splice impact <i>BP7</i>	Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i>	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before <i>PM5</i> Protein length changing variant <i>PM4</i>	Same amino acid change as an established pathogenic variant <i>PS1</i>	Predicted null variant in a gene where LOF is a known mechanism of disease <i>PVS1</i>
Functional Data	Well-established functional studies show no deleterious effect <i>BS3</i>		Missense in gene with low rate of benign missense variants and path. missenses common <i>PP2</i>	Mutational hot spot or well-studied functional domain without benign variation <i>PM1</i>	Well-established functional studies show a deleterious effect <i>PS3</i>	
Segregation Data	Non-segregation with disease <i>BS4</i>		Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data →		
De novo Data				<i>De novo</i> (without paternity & maternity confirmed) <i>PM6</i>	<i>De novo</i> (paternity & maternity confirmed) <i>PS2</i>	
Allelic Data		Observed in <i>trans</i> with a dominant variant <i>BP2</i> Observed in <i>cis</i> with a pathogenic variant <i>BP2</i>		For recessive disorders, detected in <i>trans</i> with a pathogenic variant <i>PM3</i>		
Other Database		Reputable source w/out shared data = benign <i>BP6</i>	Reputable source = pathogenic <i>PP5</i>			
Other Data		Found in case with an alternate cause <i>BP5</i>	Patient's phenotype or FH highly specific for gene <i>PP4</i>			

Pathogenic

1. 1 Very Strong (PVS1) *AND*
 - a. ≥1 Strong (PS1–PS4) *OR*
 - b. ≥2 Moderate (PM1–PM6) *OR*
 - c. 1 Moderate (PM1–PM6) and 1 Supporting (PP1–PP5) *OR*
 - d. ≥2 Supporting (PP1–PP5)
 2. ≥2 Strong (PS1–PS4) *OR*
 3. 1 Strong (PS1–PS4) *AND*
 - a. ≥3 Moderate (PM1–PM6) *OR*
 - b. 2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR*
 - c. 1 Moderate (PM1–PM6) *AND* ≥4 Supporting (PP1–PP5)

Likely Pathogenic

1. 1 Very Strong (PVS1) *AND* 1 Moderate (PM1–PM6) *OR*
2. 1 Strong (PS1–PS4) *AND* 1–2 Moderate (PM1–PM6) *OR*
3. 1 Strong (PS1–PS4) *AND* ≥2 Supporting (PP1–PP5) *OR*
4. ≥3 Moderate (PM1–PM6) *OR*
5. 2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR*
6. 1 Moderate (PM1–PM6) *AND* ≥4 Supporting (PP1–PP5)

Benign

1. 1 Stand-Alone (BA1) *OR*
2. ≥2 Strong (BS1–BS4)

Likely Benign

1. 1 Strong (BS1–BS4) and 1 Supporting (BP1–BP7) *OR*
2. ≥2 Supporting (BP1–BP7)

A forensic investigation - for each variant

- Have we seen it before?
- What do we already know?
- What can we predict (similar variants)?
- What new data can we add?
 - From this family
 - Experiments

Have we seen it before - in the population?

Strong

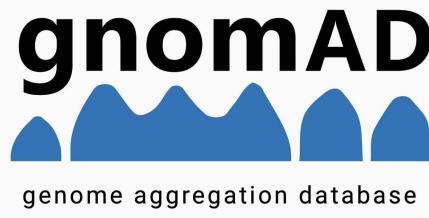
MAF is too high for disorder *BA1/BS1* OR observation in controls inconsistent with disease penetrance *BS2*

Moderate

Absent in population databases *PM2*



controversial



<https://beacon-network.org>



[Search Beacons](#)

[Login](#)

A global search engine for genetic mutations.

GRCh37 ▾

e.g. 1 : 100,000 A>C

Search

Quickstart: [Search for a BRCA2 variant](#)

<https://beacon-network.org>

Beacon Network

A global mutation database

GRCh37 ▾

Quickstart: Search for variants

GRCh37 ▾

14 : 23886482 G > C

Search

Response All None

Found 4

Not Found 37

Not Applicable 41

Access All None

Controlled 1

Public 81

Organization All None

- Aalborg University Hospital
- AMPLab, UC Berkeley
- Australian Genomics Health Alliance
- Autism Speaks
- BC Cancer Agency
- Belgian Medical Genomics Initiative
- BGI
- Bioinformatics Area, Fundacion Progreso ...
- BioReference Laboratories
- Brazilian Initiative on Precision Medicine
- BRCA Exchange
- Broad Institute
- Centre for Genomic Regulation
- Centro Nacional de Analisis Genomico
- Children's Mercy Hospital
- COGR consensus
- Curaverse
- CytoGnomix Inc
- ELIXR
- EMBL European Bioinformatics Institute
- Garvan Institute of Medical Research
- Global Alliance for Genomics and Health
- Google

Log in with Science ID to search controlled access beacons

HGMD Public Hosted by University of California, Santa Cruz Found

RD-Connect Hosted by Centro Nacional de Analisis Genomico Found

UCSC Hosted by University of California, Santa Cruz Found

VariantMatcher Hosted by VariantMatcher Found

Have we seen it before - in cases?

ClinVar Genomic variation as it relates to human health

Search ClinVar

Advanced search

About Access Submit Stats FTP Help

Was this helpful?  

Follow ? Print Download

NM_000256.3(MYBPC3):c.3742_3759dup (p.Gly1248_Cys1253dup)

Cite this record

Interpretation: Pathogenic/Likely pathogenic

Review status: ★★☆☆☆ criteria provided, multiple submitters, no conflicts

Submissions: 10 (Most recent: Oct 22, 2021)

Last evaluated: Jul 16, 2021

Accession: VCV000008603.14

Variation ID: 8603

Description: 18bp duplication

Evidence details

Publications
PubMed (8)

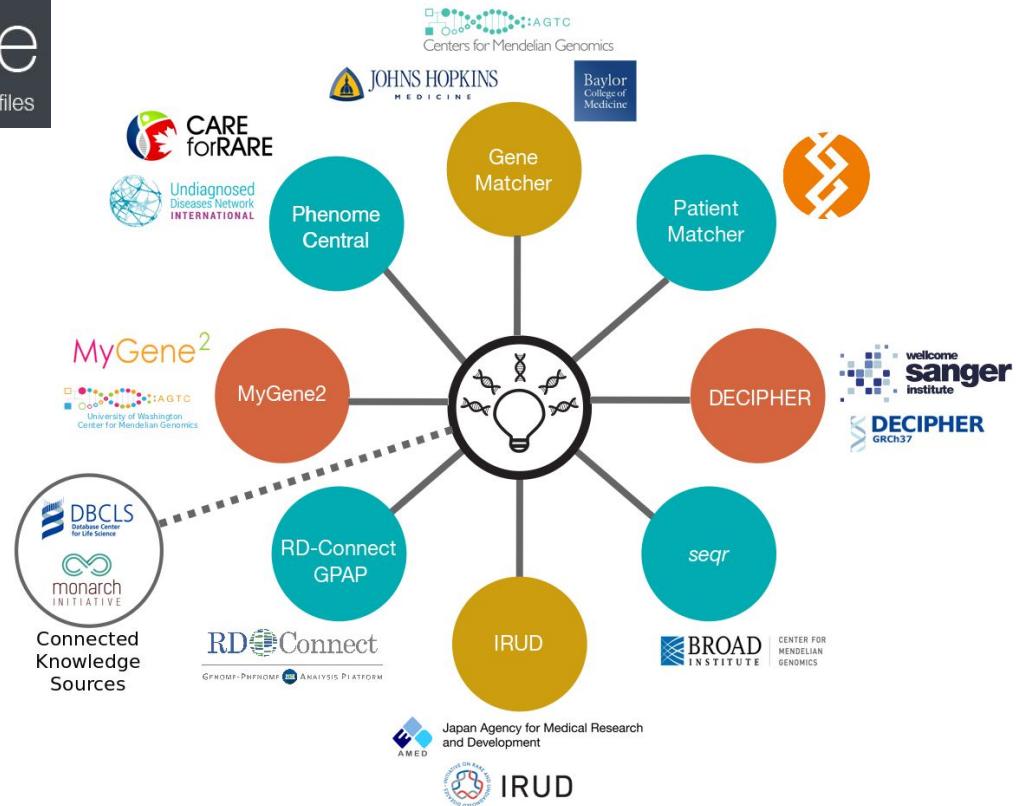
Comment:
The p.Gly1248_Cys1253dup variant in MYBPC3 has been identified in >20 individuals with HCM and segregated with disease in 9 affected family members from 3 families (Watkins 1995, Maron 2001, Helms 2014, Kapplinger 2014; LMM unpublished data). This variant is a duplication of 6 amino acids after position 1253 and is not predicted to alter the protein reading-frame. In-vitro functional studies provide some evidence that the p.Gly1248_Cys1253dup variant may impact protein function (Brown 2002, Helms 2014); however, these types of assays may not accurately represent biological function. In summary, this variant meets our criteria to be classified as pathogenic for HCM in an autosomal dominant manner based upon segregation studies and impact to the protein.

Submitted interpretations and evidence

Interpretation (Last evaluated)	Review status (Assertion criteria)	Condition (Inheritance)	Submitter	Supporting information (See all)
Pathogenic (-)	criteria provided, single submitter (ACMG Guidelines, 2015) Method: clinical testing	Primary familial hypertrophic cardiomyopathy Allele origin: germline	Molecular Diagnostic Laboratory for Inherited Cardiovascular Disease, Montreal Heart Institute Accession: SCV000987509.1 Submitted: (Feb 19, 2019)	Evidence details
Pathogenic (Aug 01, 2020)	criteria provided, single submitter (Invitae Variant Classification Sherloc (09022015)) Method: clinical testing	Hypertrophic cardiomyopathy Allele origin: germline	Invitae Accession: SCV000546482.7 Submitted: (Jan 07, 2021)	Evidence details Publications PubMed (8) Comment: This sequence change inserts 18 nucleotides in exon 33 of the MYBPC3 mRNA (c.3742_3759dup). This leads to the insertion of 6 amino acid residues in ... (more)
Pathogenic (Jun 09, 2015)	criteria provided, single submitter (LMM Criteria) Method: clinical testing	Hypertrophic cardiomyopathy (Autosomal dominant inheritance) Allele origin: germline	Laboratory for Molecular Medicine, Partners HealthCare Personalized Medicine Accession: SCV000203962.4 Submitted: (Mar 21, 2019)	Evidence details Publications PubMed (8) Comment: The p.Gly1248_Cys1253dup variant in MYBPC3 has been identified in >20 individuals with HCM and segregated with disease in 9 affected family members from ... (more)

<https://www.ncbi.nlm.nih.gov/clinvar/>

Have we seen it before - in cases?



Have we seen it before - in cases?



The screenshot shows the PubMed.gov homepage. At the top left is the logo "PubMed.gov". Below it is a search bar with a placeholder "Search" and a green "Search" button to its right. Underneath the search bar is a link "Advanced". A descriptive text block follows, stating: "PubMed® comprises more than 33 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full text content from PubMed Central and publisher web sites." The background features a dark blue header with a faint molecular or network graphic.



The screenshot shows the LitVar search interface. At the top right is the logo "LitVar". Below it is a search bar with the query "Ex: BRAF V600E". Underneath the search bar is a "Try:" section containing the text "Try: A146T, c.436G>A, rs121913527 or CFH R1210C". A detailed description follows: "LitVar [1] allows the search and retrieval of variant relevant information from the biomedical literature and shows key biological relations between a variant and its close related entities (e.g. genes, diseases, and drugs). The LitVar results are automatically extracted (with regular updates) from over 30 million PubMed articles as well as applicable full-text articles in PubMed Central. [Read More](#)".

Functional Data	Well-established functional studies show no deleterious effect <i>BS3</i>			Well-established functional studies show a deleterious effect <i>PS3</i>
------------------------	------------------------------------------------------------------------------	--	--	--------------------------------------------------------------------------

PS4 - statistically increased in cases over controls?

Strong

Prevalence in
affecteds statistically
increased over
controls PS4

- What is the ‘denominator’?
- What is significant?

1/100 vs 0/125,000 -> Fisher’s $P=8\times 10^{-4}$

A forensic investigation - for each variant

- Have we seen it before?
- What do we already know?
- **What can we predict (similar variants)?**
- What new data can we add?
 - From this family
 - Experiments

Returning to those VEP annotations - variant effect

Loss-of-function variants

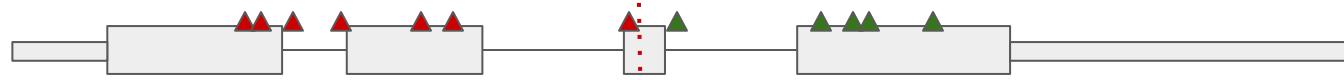
Missense variants

Predicted null
variant in a gene
where LOF is a
known
mechanism of
disease
PVS1

Be careful!

Not all LoF are really LoF

- Frameshift, essential splice, and nonsense variants
- Assumed to lead to nonsense mediated decay (NMD)



- Exons not highly expressed (unconserved)

LOFTEE

<https://github.com/konradjk/loftee>

- Classifies as HC and LC + series of flags
- Removes common LoF errors
 - Terminal exon
 - Unconserved exons
 - Splice-site rescue
 - Ancestral allele
 - etc.

LOFTEE and LoF curations on gnomAD

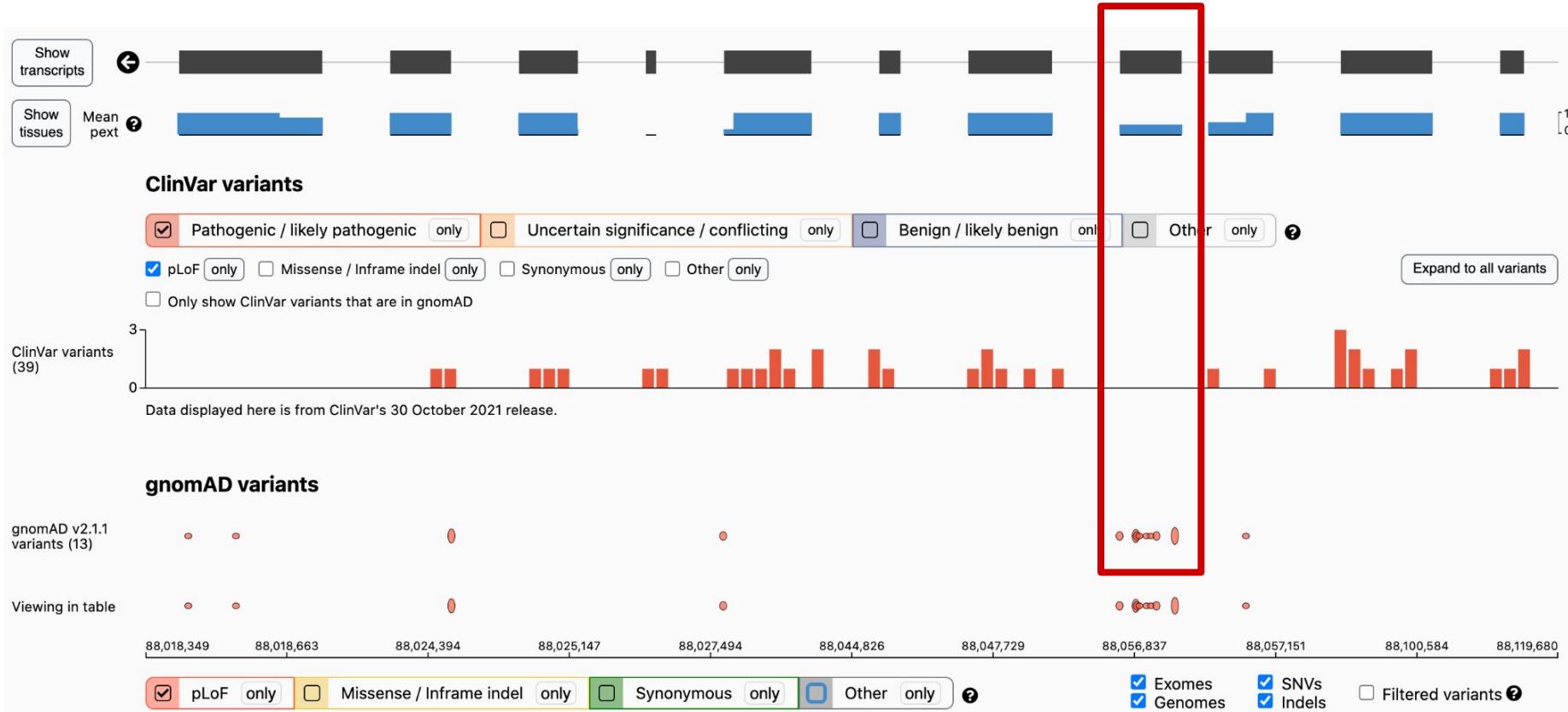
<https://gnomad.broadinstitute.org/>

Variant ID	Source	HGVS Consequence	VEP Annotation	LoF Curation	Clinical Significance	Flags	Allele Count	Allele Number
5-88018444-GTC...	E	p.Met475ThrfsTer3	● frameshift			LC pLoF	1	248652
5-88018550-G-C	E	p.Tyr441Ter	● stop gained	Not LoF	Likely benign		1	249106
5-88024446-CTA...	G	c.995_6_995_2delAT...	● splice acceptor ● splice donor		Likely benign	LC pLoF	1	31350
5-88027522-A-G	E	c.832+2T>C †				pLoF flag	2	244508
5-88056804-C-G	E	c.456+1G>C					2	242156
5-88056840-C-A	E	p.Glu141Ter					7	277704
5-88056843-C-A	E	p.Glu140Ter					2	246504
5-88056848-A-AT	E	p.Ile138AsnfsTer2	● frameshift	Likely not LoF	Uncertain significance		1	245524
5-88056864-C-A	E	p.Glu133Ter	● stop gained	Uncertain			1	243850
5-88056874-T-TG	E	p.Arg130ThrfsTer3	● frameshift	Uncertain			1	248580
5-88056886-A-T	E	p.Tyr125Ter	● stop gained	Uncertain			2	248740
5-88056927-C-A	G	p.Glu112Ter	● stop gained	Uncertain	Likely benign		2	31370
5-88057085-C-A	E	c.318+1G>T	● splice donor	Likely LoF			1	248964

Pext scores highlight lowly expressed exons



Pext scores highlight lowly expressed exons



Returning to those VEP annotations - variant effect

Loss-of-function variants

Predicted null
variant in a gene
where LOF is a
known
mechanism of
disease
PVS1

Be careful!

Missense variants

Missense in gene with
low rate of benign
missense variants and
path. missenses
common *PP2*

Novel missense change
at an amino acid residue
where a different
pathogenic missense
change has been seen
before *PM5*

Same amino acid
change as an
established
pathogenic variant
PS1

Returning to those VEP annotations - variant effect

Loss-of-function variants

Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*

Be careful!

Missense variants

Missense in gene where only truncating cause disease *BP1*

Missense in gene with low rate of benign missense variants and path. missenses common *PP2*

Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5*

Same amino acid change as an established pathogenic variant *PS1*

Is this gene and variant type known to cause disease?



GenCC

The Gene Curation Coalition

A global effort to harmonize gene-level resources.

Home | GenCC Database | About | Members | News & Articles

<https://thegencc.org/>

111 Genes with classifications based on your filters

Filter by gene symbol... Filter by Submitter ▾

Definitive Strong Moderate Supportive Limited Disputed Refuted Animal No Known

② About GenCC Classifications

Gene	Disease Equivalents	Submitters	Classification Distribution	Details
ABCC9 HGNC:60	7 <small>Disease Equivalents</small>	6 <small>Submitters</small>	1 1 1 4 3 0 0 0 0 0 0 0	Details +
ACTA1 HGNC:129	11 <small>Disease Equivalents</small>	4 <small>Submitters</small>	2 1 0 8 0 0 0 0 0 0 1	Details +
ACTC1 HGNC:143	5 <small>Disease Equivalents</small>	4 <small>Submitters</small>	1 2 1 1 0 0 0 0 0 0 1	Details +
ACTN2 HGNC:164	5 <small>Disease Equivalents</small>	6 <small>Submitters</small>	0 0 5 1 1 0 0 0 0 0 0	Details +
AGK HGNC:21869	2 <small>Disease Equivalents</small>	2 <small>Submitters</small>	1 0 0 2 0 0 0 0 0 0 0	Details +
ALPK3 HGNC:17574	1 <small>Disease Equivalents</small>	1 <small>Submitters</small>	0 1 0 0 0 0 0 0 0 0 0	Details +
ANKRD1 HGNC:15819	3 <small>Disease Equivalents</small>	2 <small>Submitters</small>	0 0 0 1 2 0 0 0 0 0 0	Details +
BAG3 HGNC:939	4 <small>Disease Equivalents</small>	2 <small>Submitters</small>	2 0 0 2 0 0 0 0 0 0 0	Details +
CACNB2 HGNC:1402	4 <small>Disease Equivalents</small>	4 <small>Submitters</small>	0 0 0 0 2 3 0 0 0 1	Details +

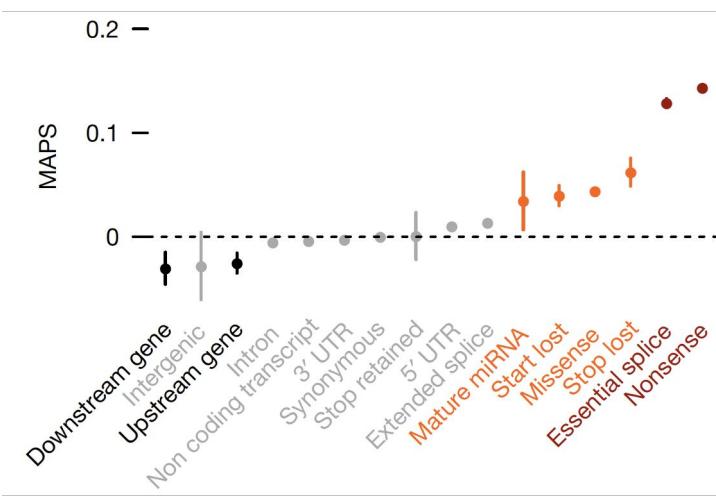
Is this gene and variant type known to cause disease?

Definitive classifications					
Definitive	 MYBPC3 HGNC:7551	left ventricular noncompaction 10 MONDO:0014163	AD ⑦	10/24/2017 Evaluated 03/02/2021 Submitted	 Ambry Genetics Assertion Criteria More Details
Definitive	 MYBPC3 HGNC:7551	hypertrophic cardiomyopathy 4 MONDO:0007268	AD ⑦	10/24/2017 Evaluated 03/02/2021 Submitted	 Ambry Genetics Assertion Criteria More Details
Definitive	 MYBPC3 HGNC:7551	hypertrophic cardiomyopathy 4 MONDO:0007268	AD ⑦	09/05/2017 Evaluated 08/16/2021 Submitted	 ClinGen Public Report Assertion Criteria More Details
Moderate classifications					
Moderate	 MYBPC3 HGNC:7551	left ventricular noncompaction 10 MONDO:0014163	AD ⑦	11/19/2018 Evaluated 03/02/2021 Submitted	 Ambry Genetics Assertion Criteria More Details
Supportive classifications					
Supportive	 MYBPC3 HGNC:7551	Familial isolated dilated cardiomyopathy MONDO:0015470	AD ⑦	09/14/2021 Evaluated 09/14/2021 Submitted	 Orphanet Assertion Criteria More Details
Limited classifications					
Limited	 MYBPC3 HGNC:7551	arrhythmogenic right ventricular cardiomyopathy MONDO:0016587	AD ⑦	08/06/2019 Evaluated 08/16/2021 Submitted	 ClinGen Public Report Assertion Criteria More Details
Limited	 MYBPC3 HGNC:7551	Primary dilated cardiomyopathy MONDO:0005021	AD ⑦	09/04/2020 Evaluated 08/16/2021 Submitted	 ClinGen Public Report Assertion Criteria More Details



An aside - genetic constraint

- We don't expect to see variants causing severe disease in the population
- Purged out due to negative selection
- This is an effect we can measure
 - Calculate 'expected' number of variants - compare to 'observed'
 - Look at shifts in allele frequency spectrum (more very rare variants)
- Need to adjust for mutability

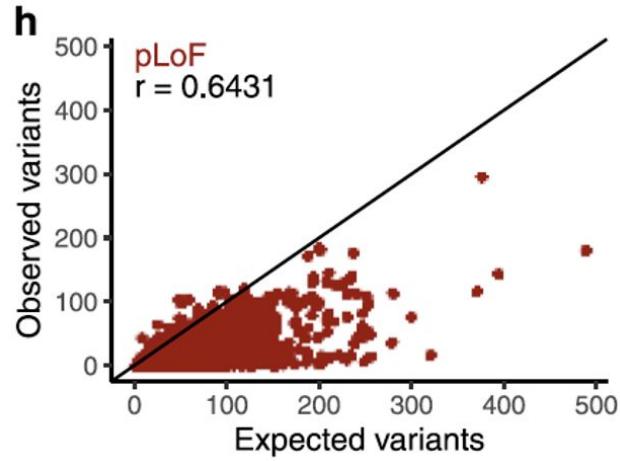
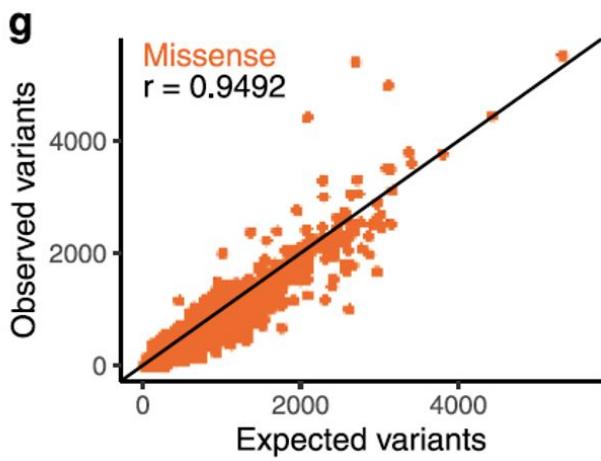
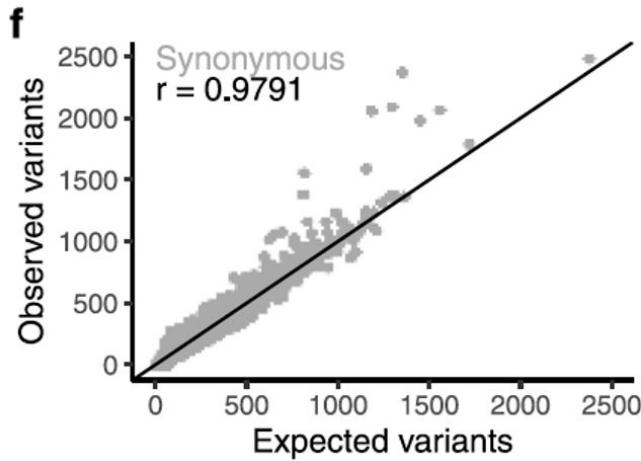


An aside - genetic constraint - gene based



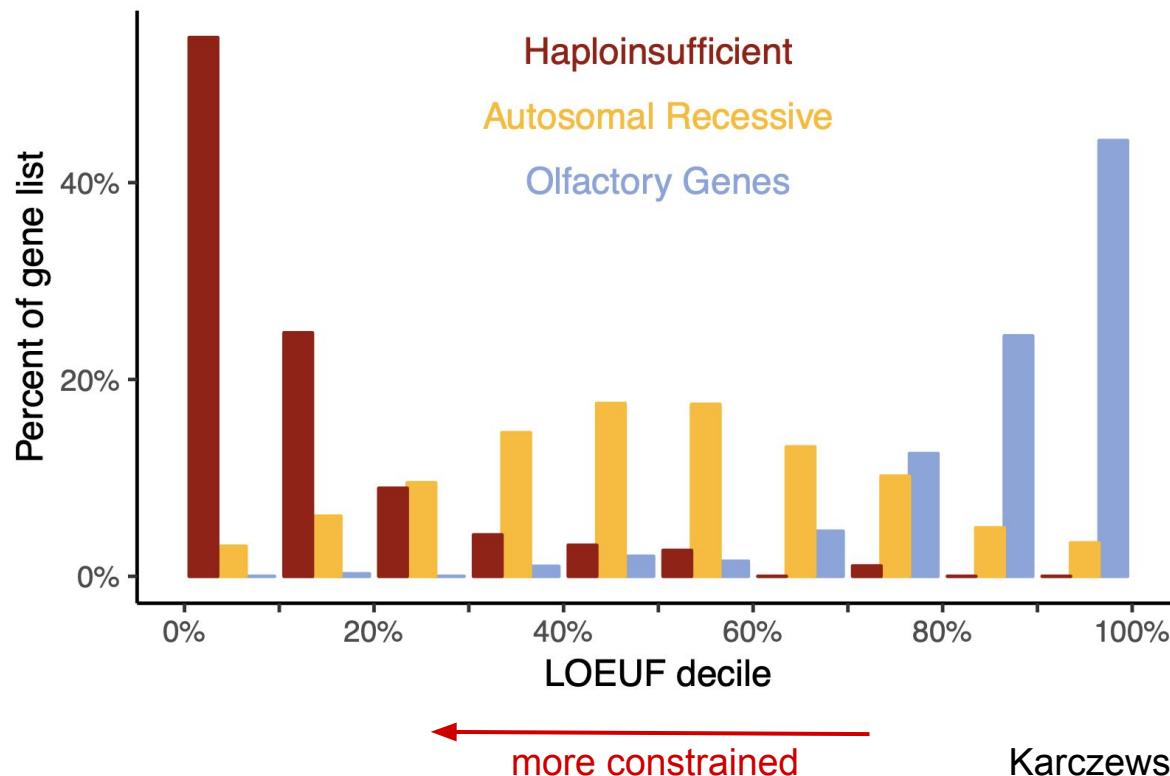
Kaitlin
Samocha

Konrad
Karczewski



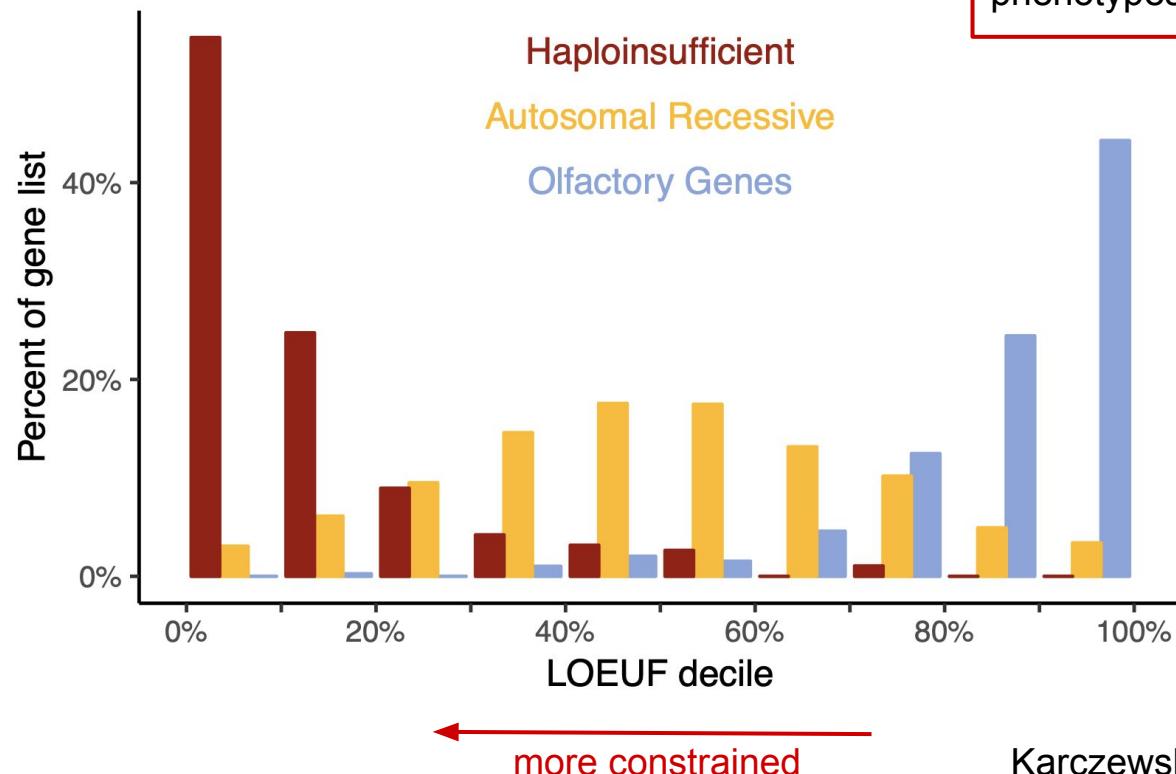
In large enough datasets it is possible to drill down to sub-gene regions

An aside - genetic constraint - LOEUF



An aside - genetic constraint - LOEUF

NB: captures negative selection which acts at level of reproduction - early onset phenotypes



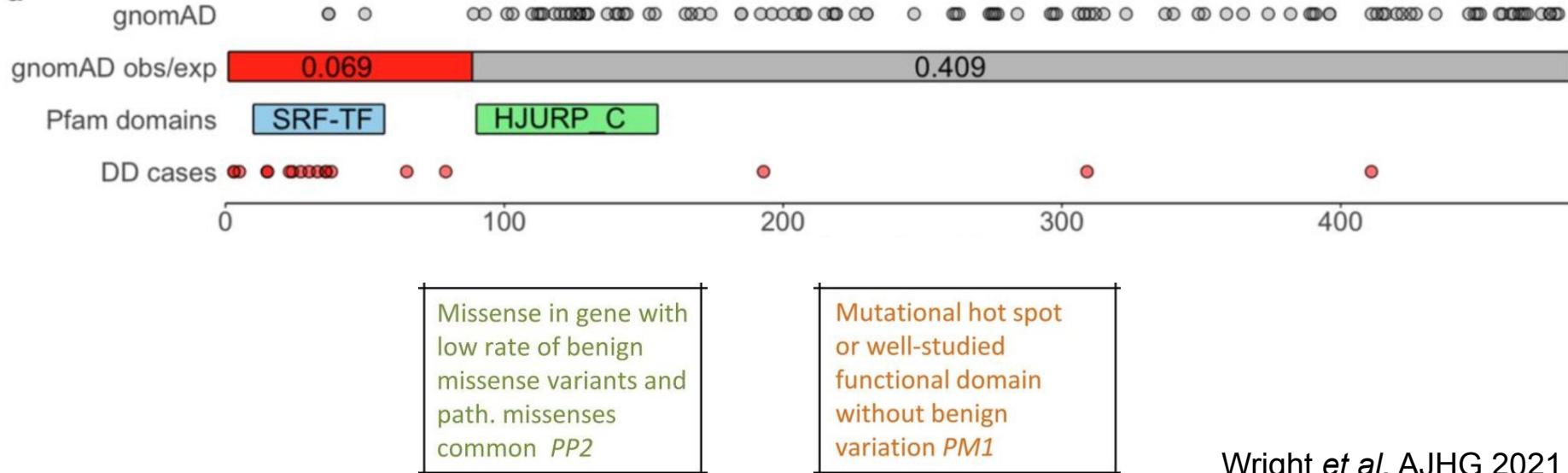


An aside - genetic constraint - regional

Katherine Chao

- MEF2C has a N-terminal DNA binding domain
- Missense variants in this region cause Developmental Disorders

a



In silico prediction tools

- Missense e.g. SIFT, PolyPhen, REVEL
- Splicing e.g. SpliceAI, CADD-splice
- Genome-wide e.g. CADD, ReMM

Often trained on ‘known’ pathogenic and benign variants

Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i>	Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i>
---------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------

A forensic investigation - for each variant

- Have we seen it before?
- What do we already know?
- What can we predict (similar variants)?
- **What new data can we add?**
 - From this family
 - Experiments

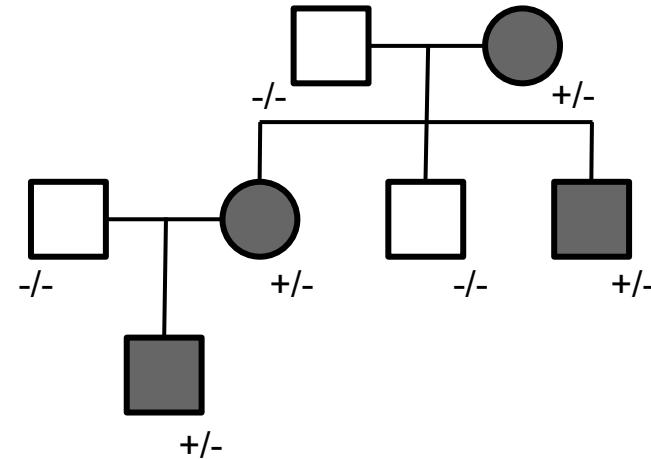
Segregation data

Does the variant track with disease in the family?

Calculate LOD (log of odds) score

Cautions/limitations:

- Late onset phenotypes
- Low penetrance



Segregation Data	Non-segregation with disease <i>BS4</i>	Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data
-------------------------	-----------------------------------------	----------------------------------------------------------------------------	----------------------------

Segregation data

Does the variant track with disease in the family?

Calculate LOD (log of odds) score

Cautions/limitations:

- Late onset phenotypes
- Low penetrance

AJHG
Volume 98, Issue 6, 2 June 2016, Pages 1077-1081

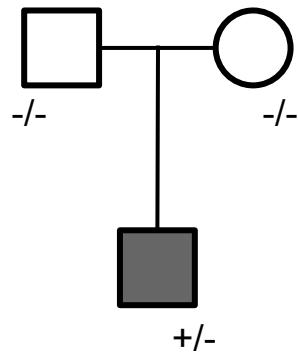


Article
Consideration of Cosegregation in the Pathogenicity Classification of Genomic Variants

Gail P. Jarvik ¹  Brian L. Browning ¹

Segregation Data	Non-segregation with disease <i>BS4</i>	Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data
-------------------------	-----------------------------------------	----------------------------------------------------------------------------	----------------------------

De novo occurrence



De novo

<i>De novo</i> (without paternity & maternity confirmed) PM6	<i>De novo</i> (paternity & maternity confirmed) PS2
--------------------------------------------------------------	------------------------------------------------------

Functional data

Guideline | [Open Access](#) | Published: 31 December 2019

Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework

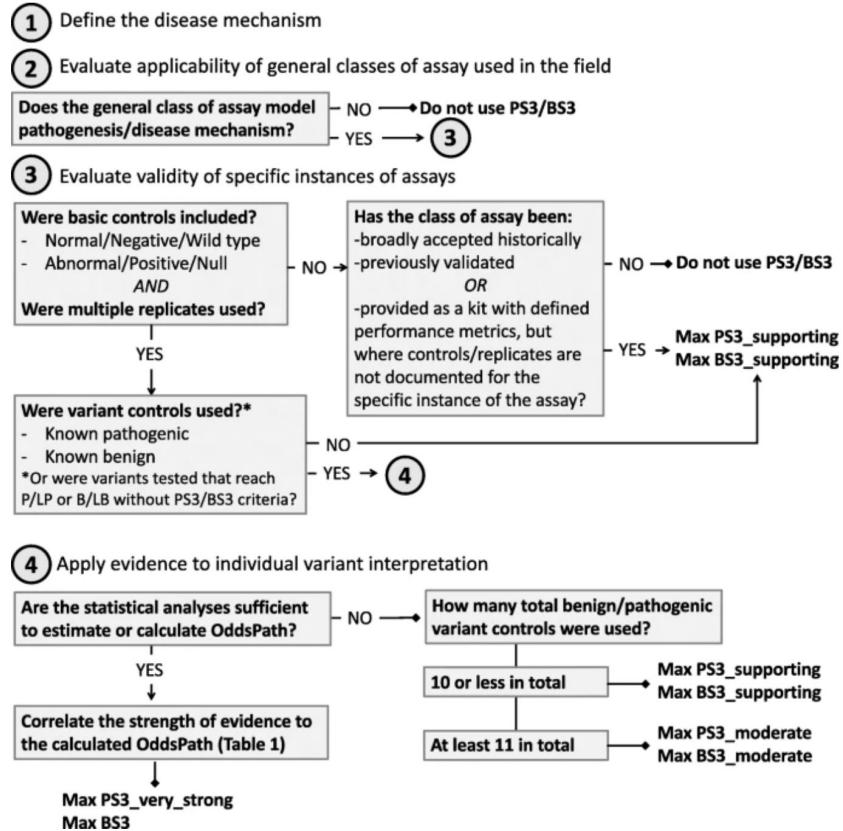
Sarah E. Brnich, Ahmad N. Abou Tayoun, Fergus J. Couch, Garry R. Cutting, Marc S. Greenblatt, Christopher D. Heinen, Dona M. Kanavy, Xi Luo, Shannon M. McNulty, Lea M. Starita, Sean V. Tavtigian, Matt W. Wright, Steven M. Harrison, Leslie G. Biesecker & Jonathan S. Berg On behalf of the Clinical Genome Resource Sequence Variant Interpretation Working Group

[Genome Medicine](#) 12, Article number: 3 (2020) | [Cite this article](#)

17k Accesses | 77 Citations | 2 Altmetric | [Metrics](#)

Functional Data	<p>Well-established functional studies show no deleterious effect BS3</p> <p>Well-established functional studies show a deleterious effect PS3</p>
------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------

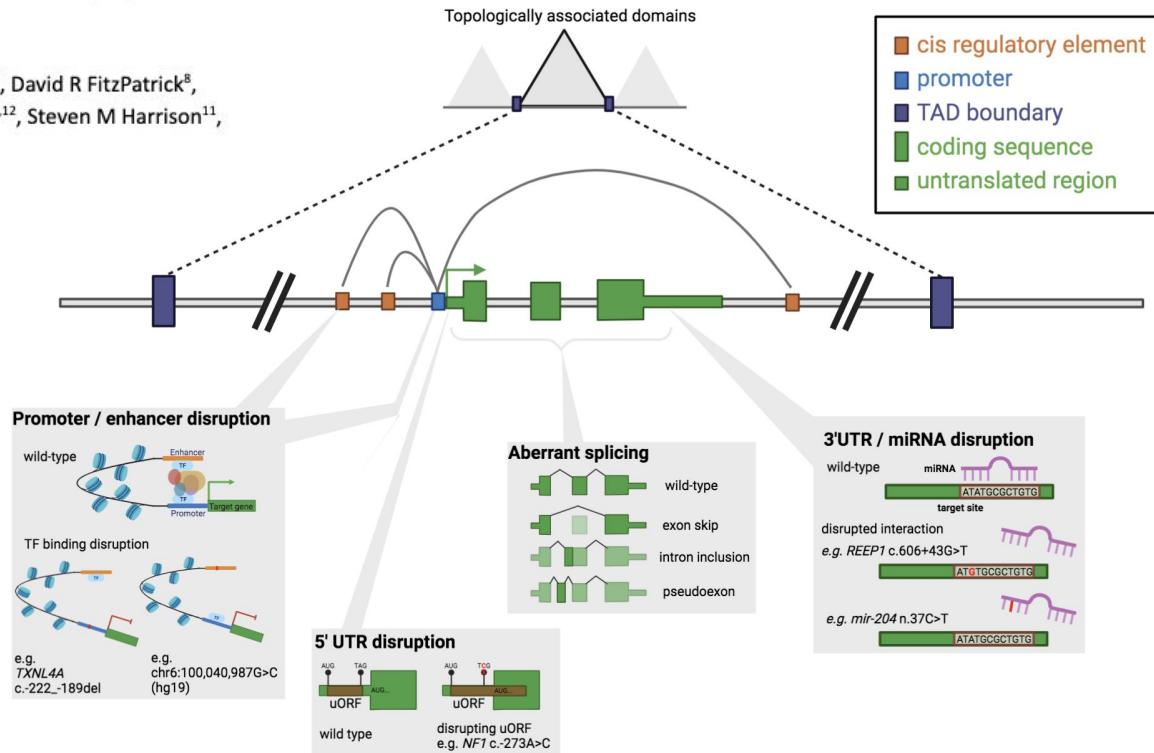
Fig. 2



Creating guidelines for non-coding region variants

Recommendations for clinical interpretation of variants found in non-coding regions of the genome

Jamie M Ellingford^{1,2*}, Joo Wook Ahn³, Diana Baralle^{4,5}, Sian Ellard^{6,7}, David R FitzPatrick⁸, Jenny Lord⁴, Hilary C Martin⁹, William G Newman¹, Jenny C Taylor^{10,12}, Steven M Harrison¹¹, and Nicola Whiffin^{12*}



The future - a Bayesian framework

- Start with a prior probability that is updated using likelihoods
- Allows quantitative measures rather than on/off at different strengths

Article | [Published: 04 January 2018](#)

Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework

[Sean V Tavtigian PhD](#)  [Marc S Greenblatt MD, PhD](#), [Steven M Harrison PhD](#), [Robert L Nussbaum MD](#),
[Snehit A Prabhu PhD](#), [Kenneth M Boucher PhD](#) & [Leslie G Biesecker MD](#) on behalf of the ClinGen
Sequence Variant Interpretation Working Group (ClinGen SVI)

[Genetics in Medicine](#) **20**, 1054–1060 (2018) | [Cite this article](#)

7787 Accesses | **130** Citations | **71** Altmetric | [Metrics](#)

Practical 3: Variant curation with ACMG

- Have a go at interpreting a list of variants using the ACMG guidelines
- Disease = Hypertrophic cardiomyopathy (HCM)

Resources:

- PDF of ACMG guidelines
- VCF of variants annotated with VEP in practical 1

Have a go in pairs - 2 variants each - present findings back to the group