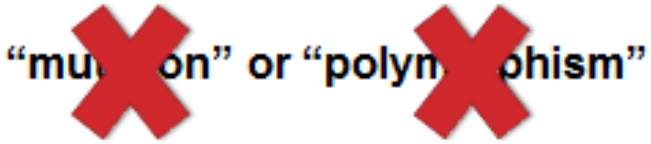
Mutations: variant classification for acquired and constitutional diseases

SOFIE SYMOENS

ANNELIES DHEEDENE

TERMS



“variant” + appropriate modifier

TERMS

Class	Modifier*	Wording report*
1	benign	do not report
2	likely benign°	"diagnosis not confirmed molecularly"
3	uncertain significance	"does not confirm or exclude diagnosis"
4	likely pathogenic°	"consistent with the diagnosis"
5	pathogenic	"the result confirms the diagnosis"

* Based on ACGS guidelines (Wallis et al, 2013)
° >90% certainty of a variant being either benign or disease causing

FIRST PROPOSAL JUNE 1, 2015

CLASSIFICATION CRITERIA

Whether a variant should be categorized as benign/likely benign or pathogenic/likely pathogenic is based on different criteria:

Benign/likely benign: stand-alone (BA1)
strong evidence (BS1-4)
supporting evidence (BP1-7)

Pathogenic/likely pathogenic: very strong evidence (PVS1)
strong evidence (PS1-4)
moderate evidence (PM1-6)
supporting evidence (PP1-5)

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

Class	Modifier	Criteria for classification*
1	benign	(i) 1 stand-alone (BA1) *OR* (ii) ≥ 2 strong (BS1-4)
2	likely benign	(i) 1 strong (BS1-4) and 1 supporting (BP1-7) *OR* (ii) ≥ 2 supporting (BP1-7)
3	uncertain significance	(i) other criteria are not met *OR* (ii) the criteria for benign and pathogenic are contradictory
4	likely pathogenic	(i) 1 very strong (PVS1) *AND* 1 moderate (PM1-6) *OR* (ii) 1 strong (PS1-4) *AND* 1-2 moderate (PM1-6) *OR* (iii) 1 strong (PS1-4) *AND* ≥ 2 supporting (PP1-5) *OR* (iv) ≥ 3 moderate (PM1-6) *OR* (v) 2 moderate (PM1-6) *AND* ≥ 2 supporting (PP1-5) *OR* (vi) 1 moderate (PM1-6) *AND* ≥ 4 supporting (PP1-5)
5	pathogenic	(i) 1 very strong (PVS1) *AND* (a) ≥ 1 strong (PS1-4) *OR* (b) ≥ 2 moderate (PM1-6) *OR* (c) 1 moderate (PM1-6) *AND* 1 supporting (PP1-5) *OR* (d) ≥ 2 supporting (PP1-5) (ii) ≥ 2 strong (PS1-4) *OR* (iii) 1 strong (PS1-4) *AND* (a) ≥ 3 moderate (PM1-6) *OR* (b) 2 moderate (PM1-6) *AND* ≥ 2 supporting (PP1-5) *OR* (c) 1 moderate (PM1-6) *AND* ≥ 4 supporting (PP1-5)

* Based on ACMG guidelines (Richards et al, Genetics in Medicine 2015)

2015 ACMG Guidelines

Transcript Reporting

- A reference transcript should be used and provided in the report

Transcript choice

- Longest known
- Most clinically relevant

Should evaluate all clinically relevant transcripts when there are known, interpretable variants

- Richards et al., Genet Med 2015; 17(5): 405-424

GRCh37 (HG19)

Combination of genomes

- ~20 people, but ~70% from one individual

Not a universal or "gold standard"

- Gaps in sequence
  - Highly repetitive regions
  - Mitochondrial genome
- Long stretches highly specific to one individual
- Allelic diversity limited
- 2 million reference alleles have population frequency of <0.5
- Statistically like an arbitrarily chosen personal genome

GRCh38

Improvements

- More individuals sequenced from diverse ethnicities
- Longer reads (Sanger)
- Haploid genome

Effects

- A more representative genome, better annotation of variation
- Incorrect reference alleles fixed
- More easily define breakpoints of large CNVs if these overlap previous 'missing regions'
- More confidence in CNV and Structural Variation calling –reduced numbers of false negative CNVs and SV
- Better definition of alternative haplotypes- especially important for MHC region and other complex regions. Allows for alignment for first time (until now these sequences were just not included in analyses).
- Mitochondrial genome included

Diagnostic Gene Sequencing Panels: Choosing a Transcript

2019 ACMG Technical Standards

Well-characterised gene

- Adhere to conventions in the field
  - Locus Reference Genomic
  - Stable identifiers
  - No versions or changes
  - Map back to GRCh37 or 38
- Favour transcripts
  - Used in many publications
  - Known biological relevance
- Discrepancies in transcripts should be mitigated by bioinformatic processes

Not well-characterised gene

- Review publications for transcript descriptions to ensure all pathogenic variants would be detected
- Default to all-exon approach across one or more transcripts with largest canonical transcript

Diagnostic Gene Sequencing Panels: Choosing a Transcript 2019 ACMG Technical Standards

Caution against analysis of exons

- in rarely expressed transcripts
- solely predicted by in silico algorithms

If no single transcript covers all exons reported to contain disease-causing variants, use ≥1

Alternate transcripts may have disease-specific consequences

Provide list of transcripts analysed in report

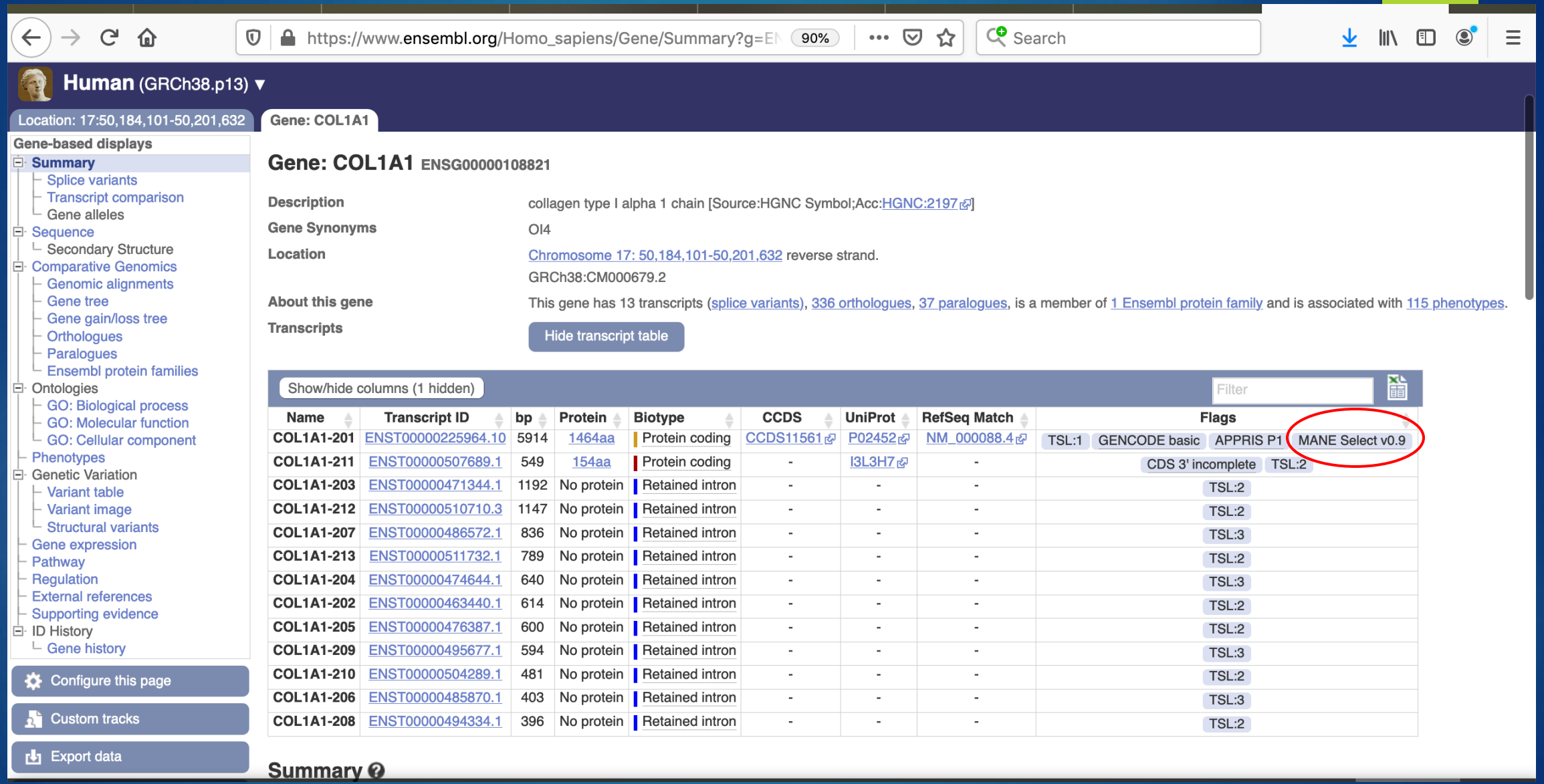
What's new? Matched Annotation from the NCBI and EMBL-EBI

Transcript set with following attributes

- Match GRCh38 sequence
- 100% identical between RefSeq and Ensembl transcript
- 5'UTR, CDS, 3'UTR

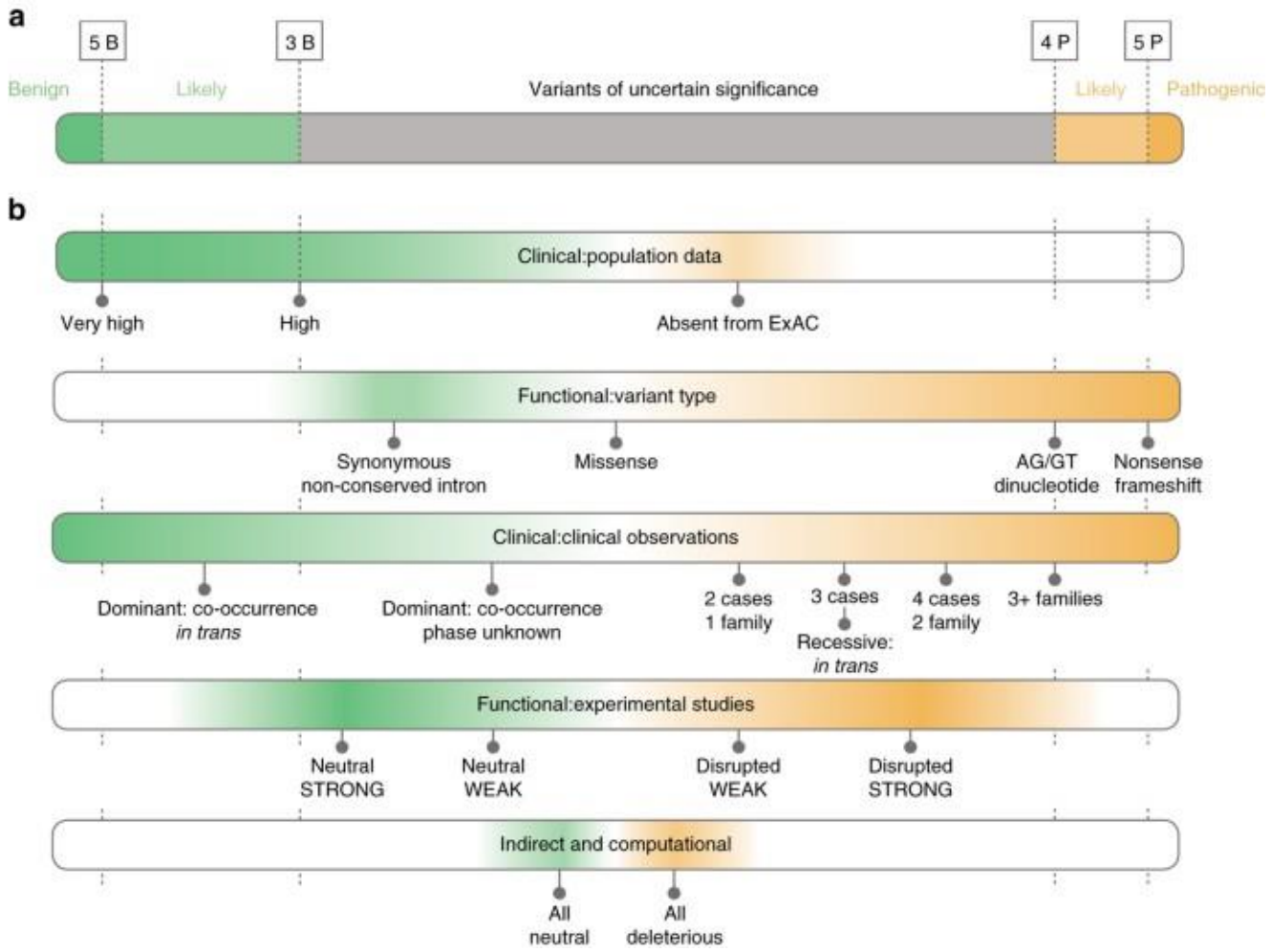
Transcripts should be

- Well supported, expressed, conserved
- Representative of biology at each locus
- BUT, be careful: Most highly supported transcript may not capture biological complexity, by excluding tissue specific or clinically relevant isoforms

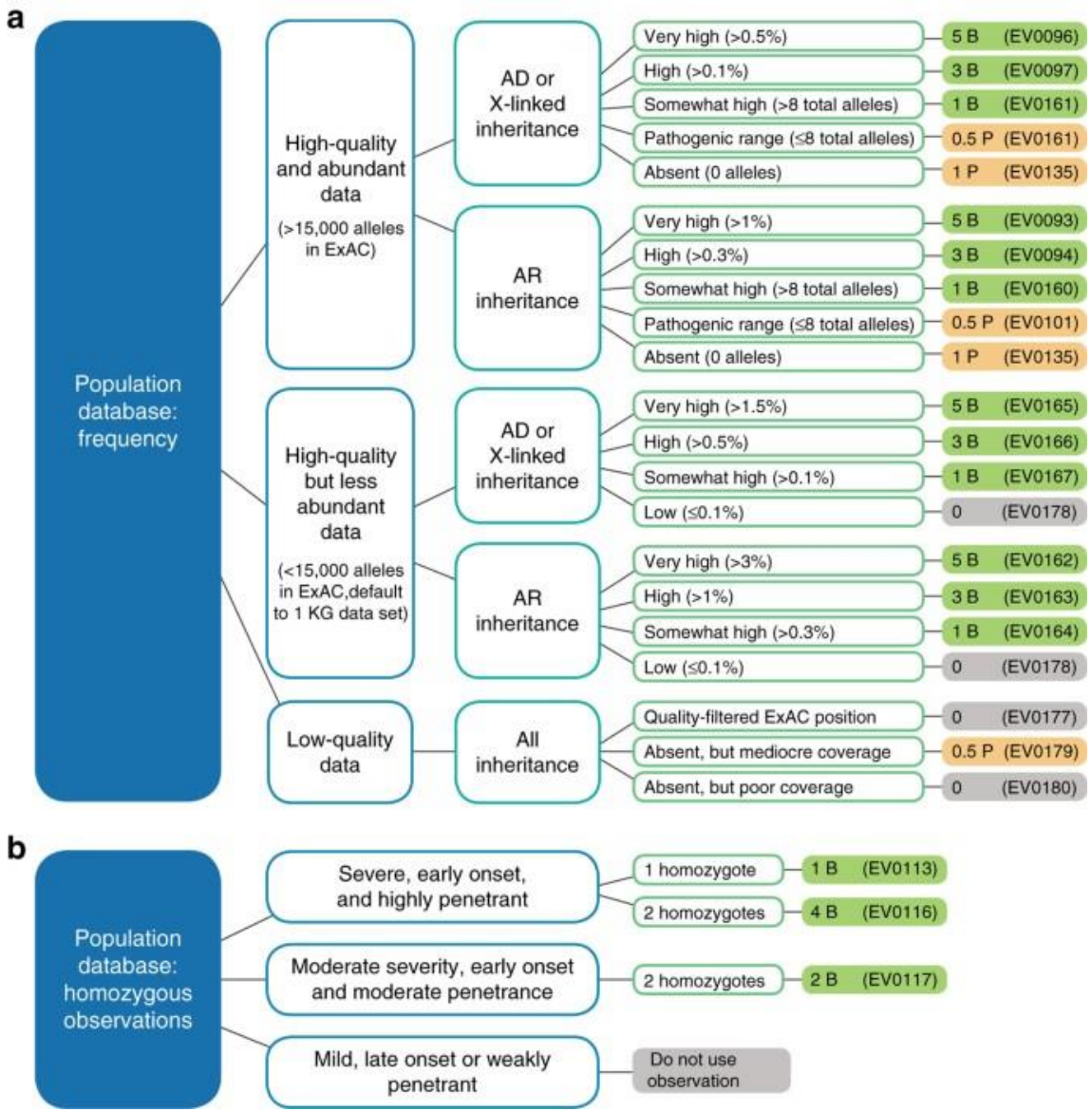

https://www.ensembl.org/Homo_sapiens/Gene/Summary?g=EN
90%
Search
Human (GRCh38.p13)
Location: 17:50,184,101-50,201,632
Gene: COL1A1
Gene-based displays
Summary
Splice variants
Transcript comparison
Gene alleles
Sequence
Secondary Structure
Comparative Genomics
Genomic alignments
Gene tree
Gene gain/loss tree
Orthologues
Paralogues
Ensembl protein families
Ontologies
GO: Biological process
GO: Molecular function
GO: Cellular component
Phenotypes
Genetic Variation
Variant table
Variant image
Structural variants
Gene expression
Pathway
Regulation
External references
Supporting evidence
ID History
Gene history
Configure this page
Custom tracks
Export data
Gene: COL1A1 ENSG00000108821
Description
collagen type I alpha 1 chain [Source:HGNC Symbol;Acc:HGNC:2197]
Gene Synonyms
OI4
Location
Chromosome 17: 50,184,101-50,201,632 reverse strand.
GRCh38:CM000679.2
About this gene
This gene has 13 transcripts (splice variants), 336 orthologues, 37 paralogues, is a member of 1 Ensembl protein family and is associated with 115 phenotypes.
Transcripts
Hide transcript table
Show/hide columns (1 hidden)
Filter
Name
Transcript ID
bp
Protein
Biotype
CCDS
UniProt
RefSeq Match
Flags
COL1A1-201 ENST00000225964.10 5914 1464aa Protein coding CCDS11561 P02452 NM_000088.4 TSL:1 GENCODE basic APPRIS P1 MANE Select v0.9
COL1A1-211 ENST00000507689.1 549 154aa Protein coding - I3L3H7 - CDS 3' incomplete TSL:2
COL1A1-203 ENST00000471344.1 1192 No protein Retained intron - - - TSL:2
COL1A1-212 ENST00000510710.3 1147 No protein Retained intron - - - TSL:2
COL1A1-207 ENST00000486572.1 836 No protein Retained intron - - - TSL:3
COL1A1-213 ENST00000511732.1 789 No protein Retained intron - - - TSL:2
COL1A1-204 ENST00000474644.1 640 No protein Retained intron - - - TSL:3
COL1A1-202 ENST00000463440.1 614 No protein Retained intron - - - TSL:2
COL1A1-205 ENST00000476387.1 600 No protein Retained intron - - - TSL:2
COL1A1-209 ENST00000495677.1 594 No protein Retained intron - - - TSL:3
COL1A1-210 ENST00000504289.1 481 No protein Retained intron - - - TSL:2
COL1A1-206 ENST00000485870.1 403 No protein Retained intron - - - TSL:3
COL1A1-208 ENST00000494334.1 396 No protein Retained intron - - - TSL:2
Summary

Variant classification anno 2021

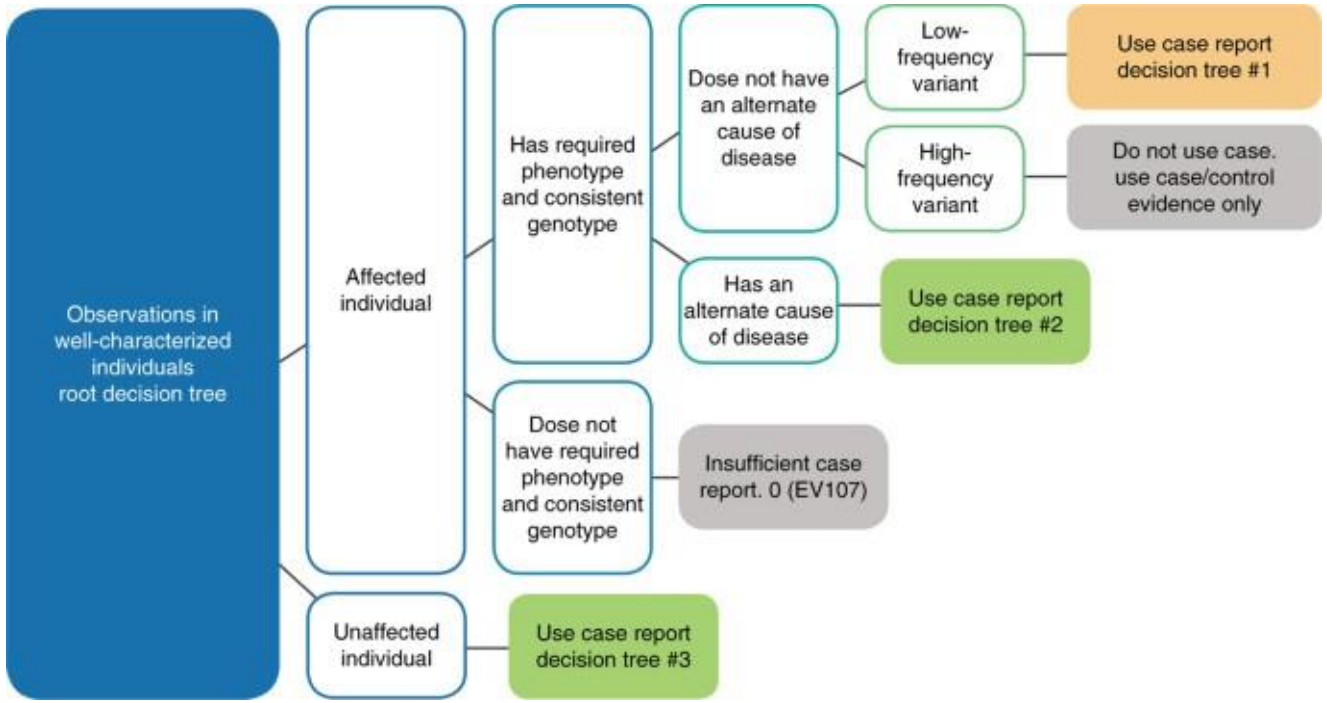
- Current VCT is based on ACMG/AMP guidelines 2015 (link)
- New classification tools are based on
  - ACGS Best Practice Guidelines for Variant Classification 2017 – 2018 – 2019 – 2020 (link)
  - Recommendations for Interpreting the Loss of Function PVS1 ACMG/AMP Variant Criterion 2018 (link)
  - Modeling the ACMG/AMP Variant Classification Guidelines as a Bayesian Classification Framework 2018 (link)
  - Sherloc: a Comprehensive Refinement of the ACMG–AMP Variant Classification Criteria 2017 (link)



Sherloc: a Comprehensive Refinement of the ACMG–AMP Variant Classification Criteria 2017



Population data



Genotype phenotype correlation patient

PVS1 – (Very Strong) null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multi-exon deletion) in a gene where LOF is a known mechanism of disease

- Potential over-classification of variants
  - Genes without loss-of-function disease mechanism
  - Variants in the last exon: NMD?
  - Variants located in biological less relevant transcript
  - Splice variants without effect on protein

SPECIAL ARTICLE

WILEY

Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion

Ahmad N. Abou Tayoun[1,2*] | Tina Pesaran[3] | Marina T. DiStefano[4] | Andrea Oza[4] | Heidi L. Rehm[4,5,6] | Leslie G. Biesecker[7] | Steven M. Harrison[4] | ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI)

- Clinical validity of gene?
- Have LoF variants been reported?

PVS1:
Is LoF disease mechanism?

- Abou Tayoun et al, 2018, Hum Mut

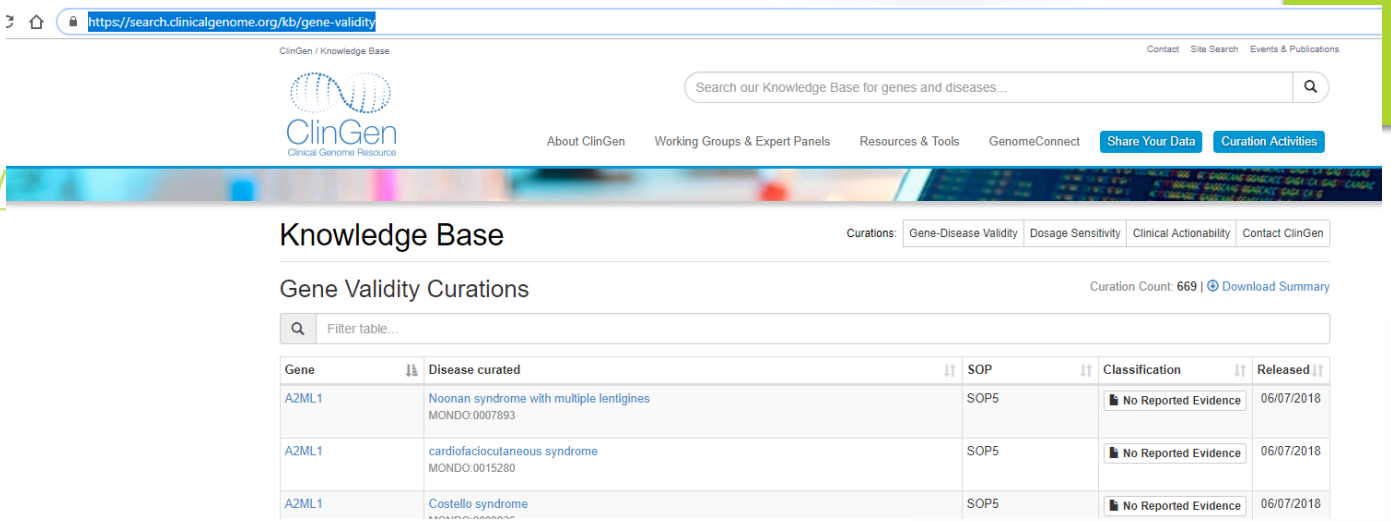
TABLE 1 Criteria for LoF disease mechanism

Follow PVS1 Flowchart if:
Clinical validity classification of gene is STRONG or DEFINITIVE
AND
3 or more LOF variants are Pathogenic without PVS1 AND > 10% of variants associated with the phenotype are LOF (must be across more than 1 exon)*
Decrease final strength by *one* level (i.e. VeryStrong to Strong) if:
Clinical validity classification of gene is at least MODERATE
AND
2 or more LOF variants have been previously associated with the phenotype (must be across more than 1 exon)*
AND
Null mouse model recapitulates disease phenotype
Decrease final strength by *two* levels (i.e. VeryStrong to Moderate) if:
Clinical validity classification is at least MODERATE
AND EITHER
2 or more LOF variants have been previously associated with the phenotype (must be across more than 1 exon)*
OR
Null mouse model recapitulates disease phenotype
If there is no evidence that LOF variants cause disease, PVS1 should not be applied at any strength level.

*With the exception of single-exon genes.

PVS1: Is LoF disease mechanism?

https://search.clinicalgenome.org/kb/gene-validity



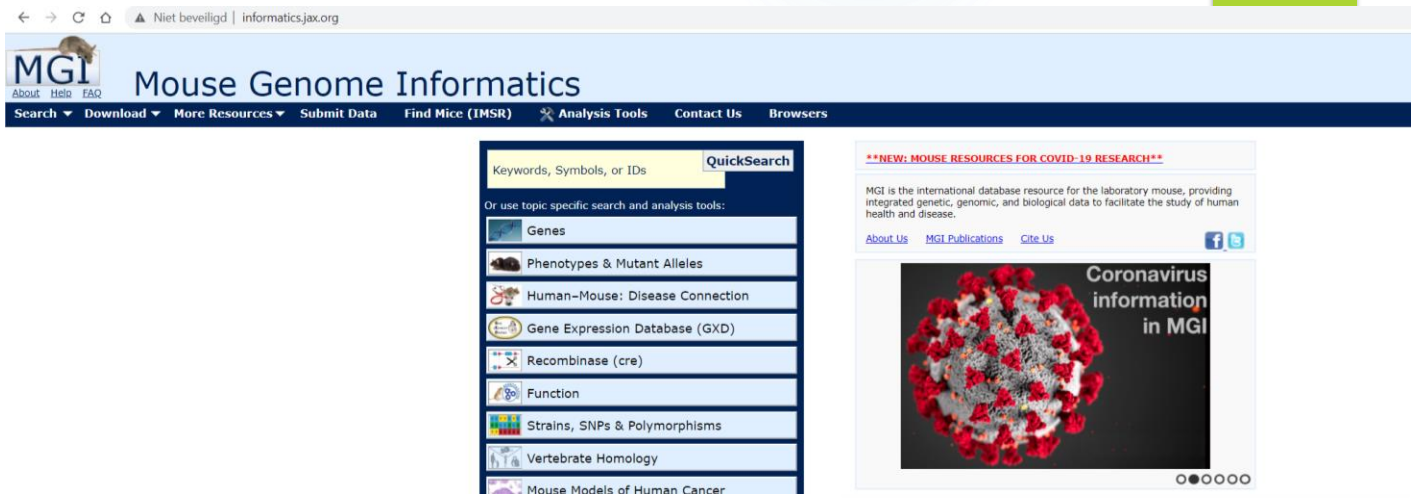
Strande et al 2017 AJHM

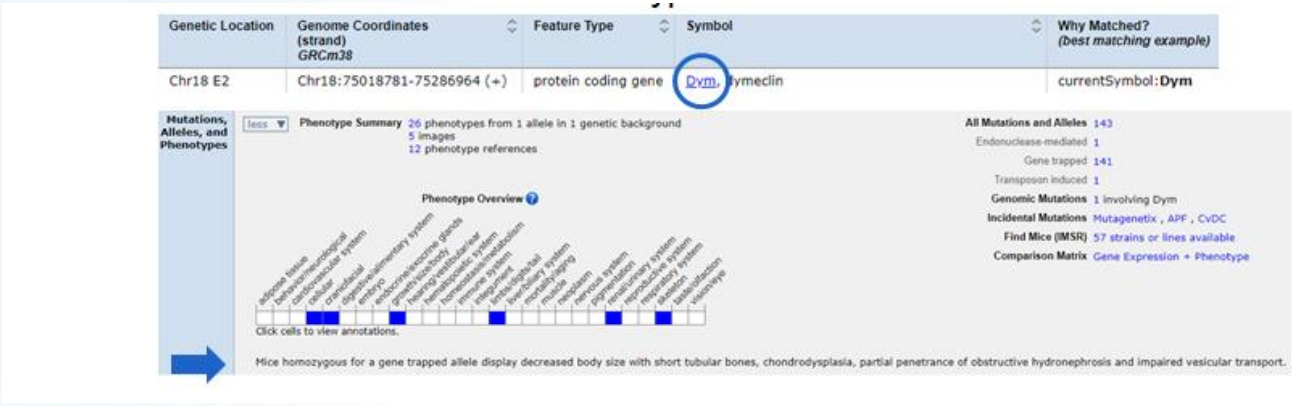
Evidence Level		Evidence Description
Supportive Evidence	DEFINITIVE	The role of this gene in this particular disease has been repeatedly demonstrated in both the research and clinical diagnostic settings, and has been upheld over time (in general, at least 3 years). No convincing evidence has emerged that contradicts the role of the gene in the specified disease.
	STRONG	The role of this gene in disease has been independently demonstrated typically in at least two separate studies providing **strong** supporting evidence for this gene's role in disease, usually including both of the following types of evidence: • Strong variant-level evidence demonstrating numerous unrelated probands with variants that provide convincing evidence for disease causality[1] **as well as** • Compelling gene-level evidence from different types of supporting experimental data[2]. In addition, no convincing evidence has emerged that contradicts the role of the gene in the noted disease.
	MODERATE	There is **moderate** evidence to support a causal role for this gene in this disease, typically including both of the following types of evidence: • Several probands with variants that provide convincing evidence for disease causality[1] • Moderate experimental data[2] supporting the gene-disease association The role of this gene in disease may not have been independently reported, but no convincing evidence has emerged that contradicts the role of the gene in the noted disease.
	LIMITED	There is **limited** evidence to support a causal role for this gene in this disease, such as: • Fewer than three observations of variants that provide convincing evidence for disease causality[1] OR • Variants have been observed in probands, but none have sufficient evidence for disease causality. • Limited experimental data[2] supporting the gene-disease association The role of this gene in disease may not have been independently reported, but no convincing evidence has emerged that contradicts the role of the gene in the noted disease.
		Evidence for a causal role in disease has not been reported. These genes might be

Abou Tayoun et al, 2018, Hum Mut

PVS1:
Is LoF disease mechanism?

http://www.informatics.jax.org/

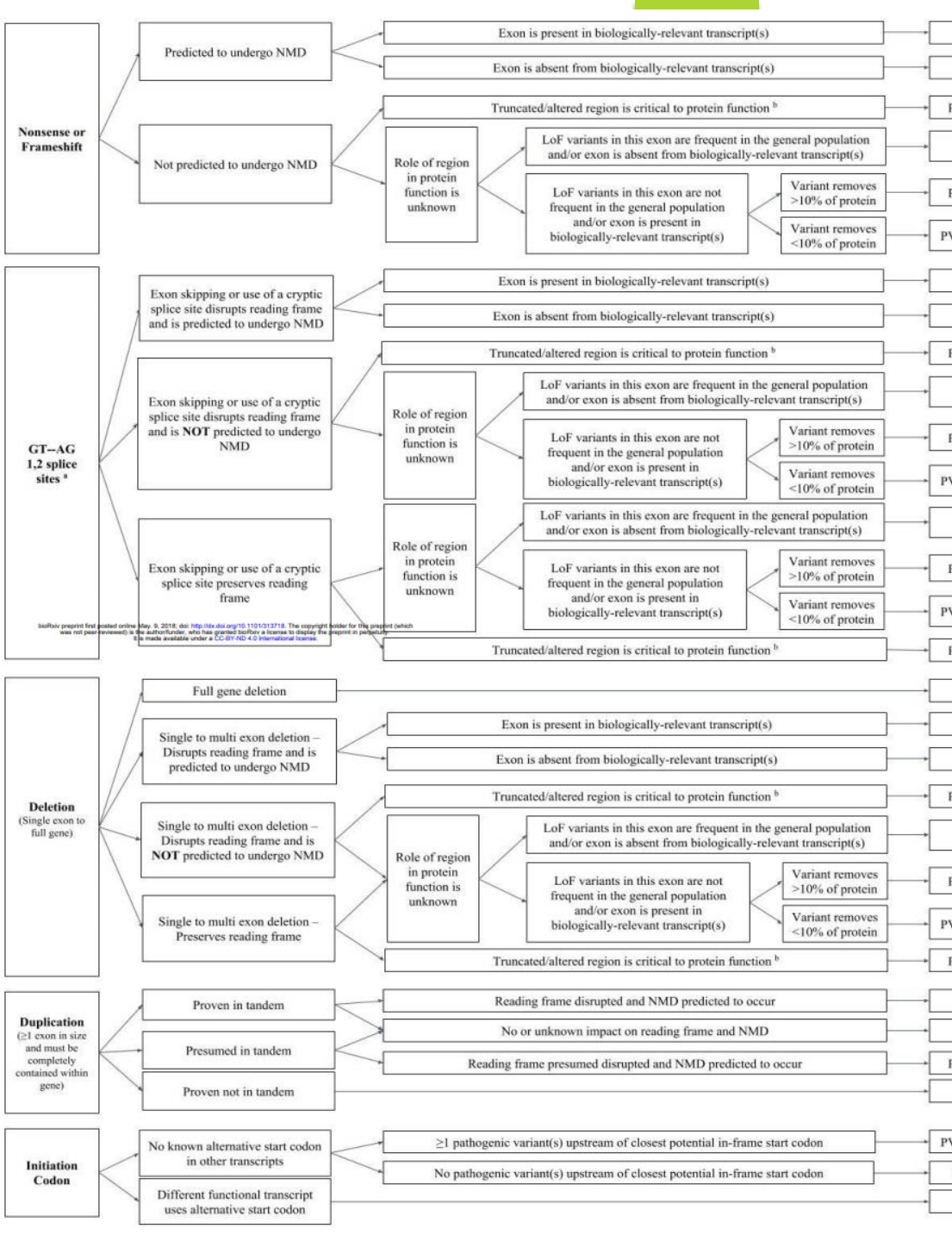


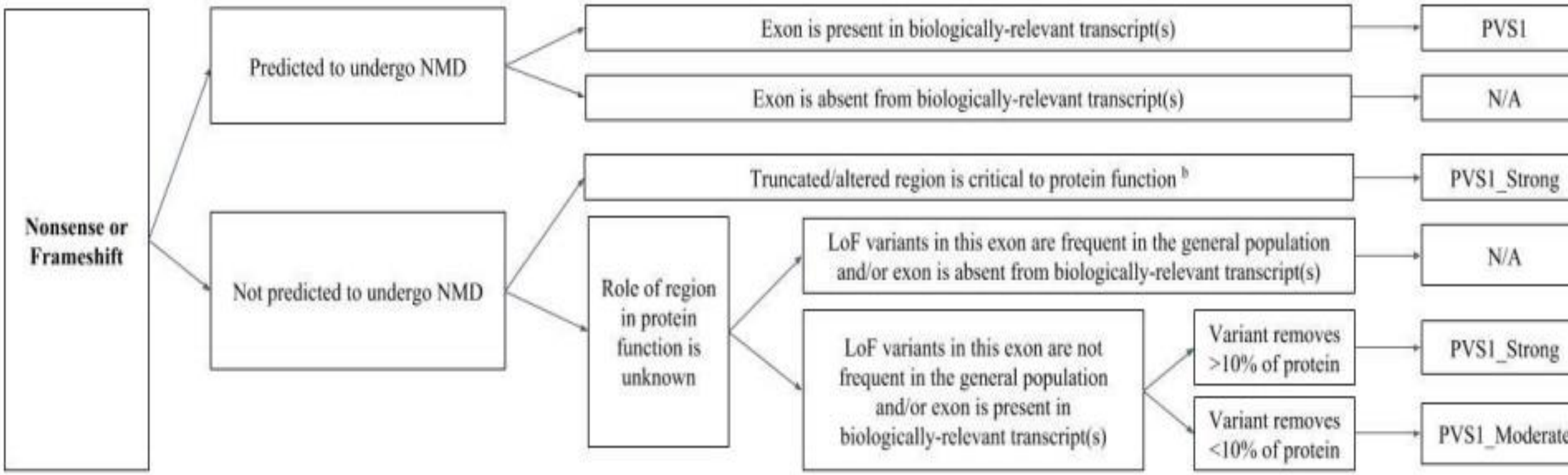


Abou Tayoun et al, 2018, Hum Mut

PVS1: decision tree

- Abou Tayoun et al, 2018, Hum Mut

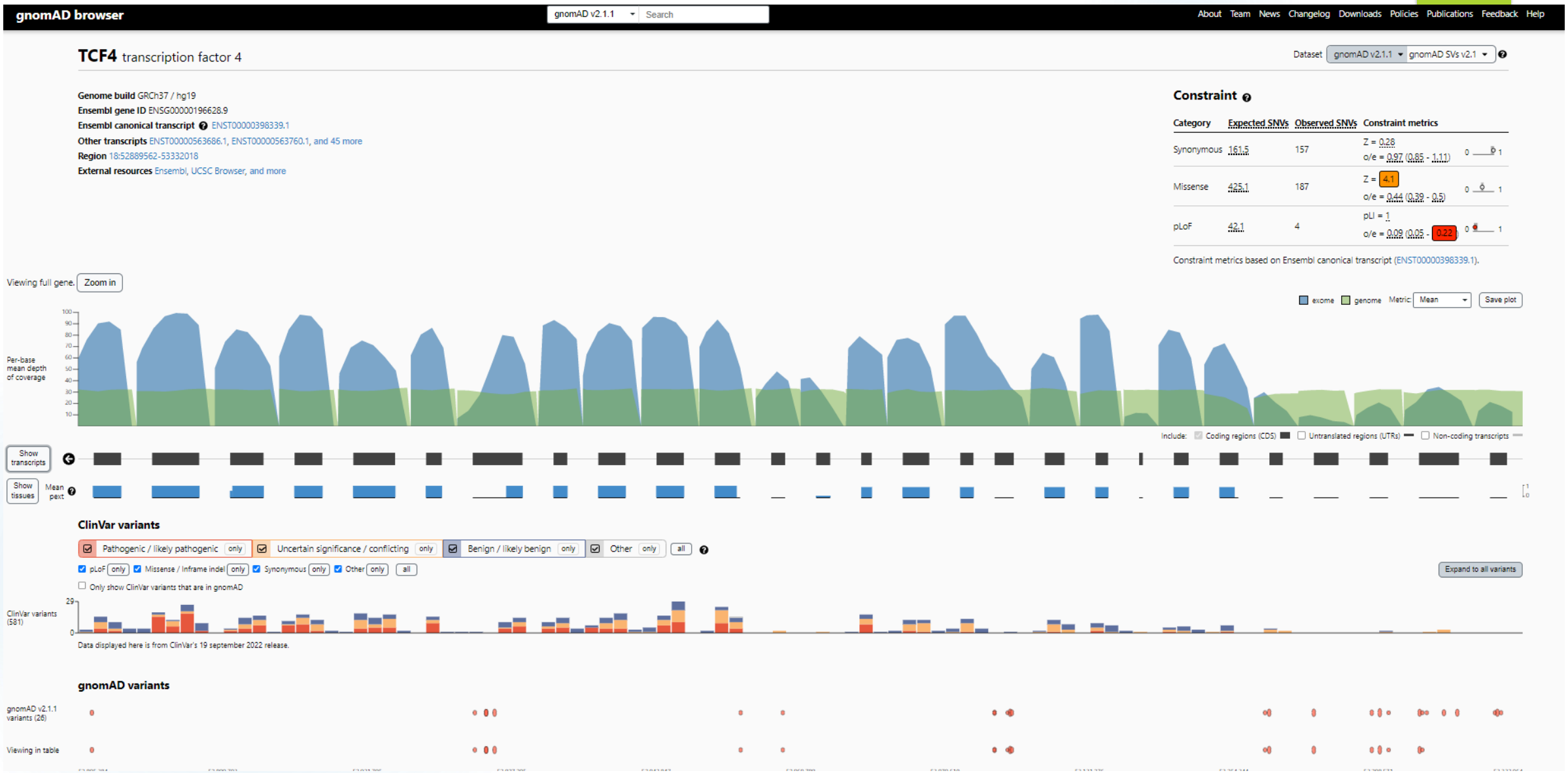




PVS1: decision tree

NMD PREDICTION BASED ON THE PREMATURE TERMINATION CODON NOT OCCURRING IN THE 3′ MOST EXON OR THE 3′ -MOST 50 BP OF THE PENULTIMATE EXON.

Abou Tayoun et al, 2018, Hum Mut

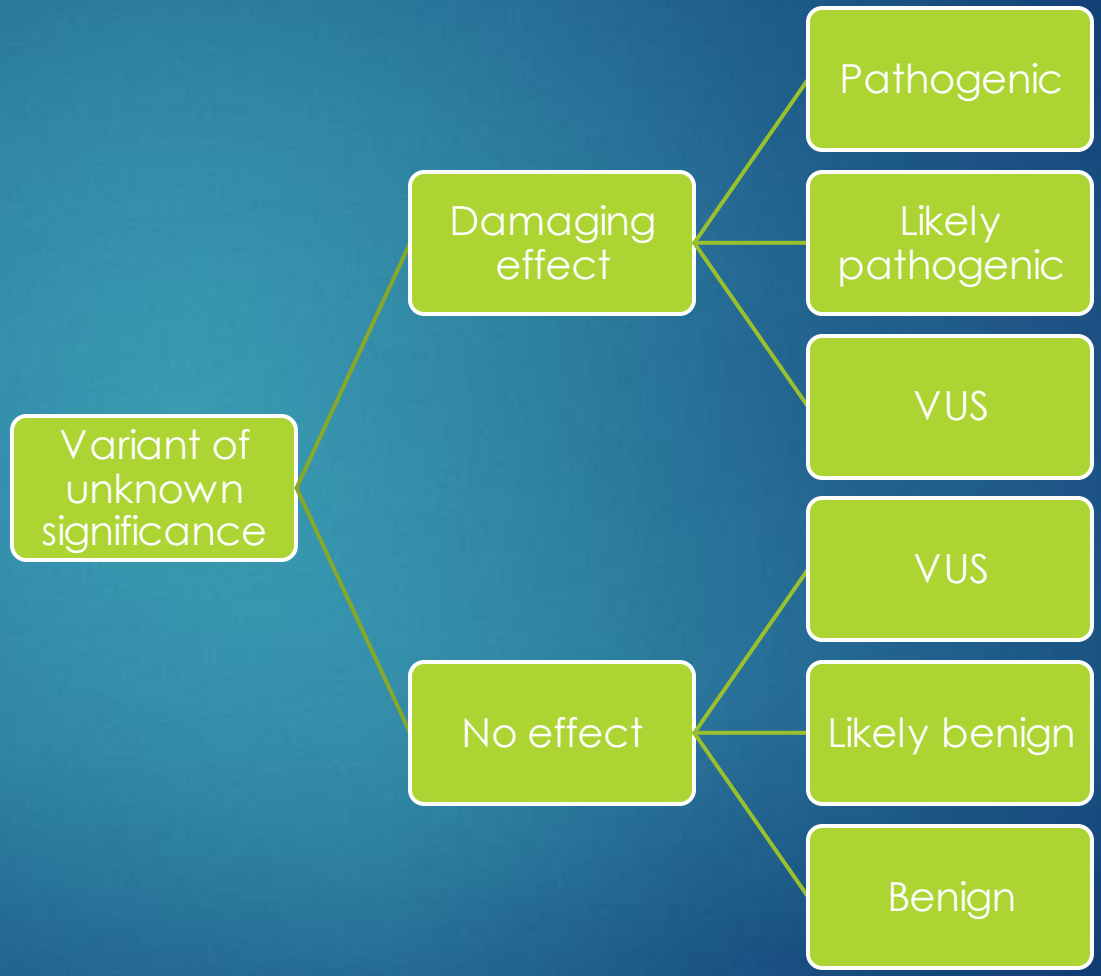


PVS1: Biological relevant transcript?

What can change the class of a variant?

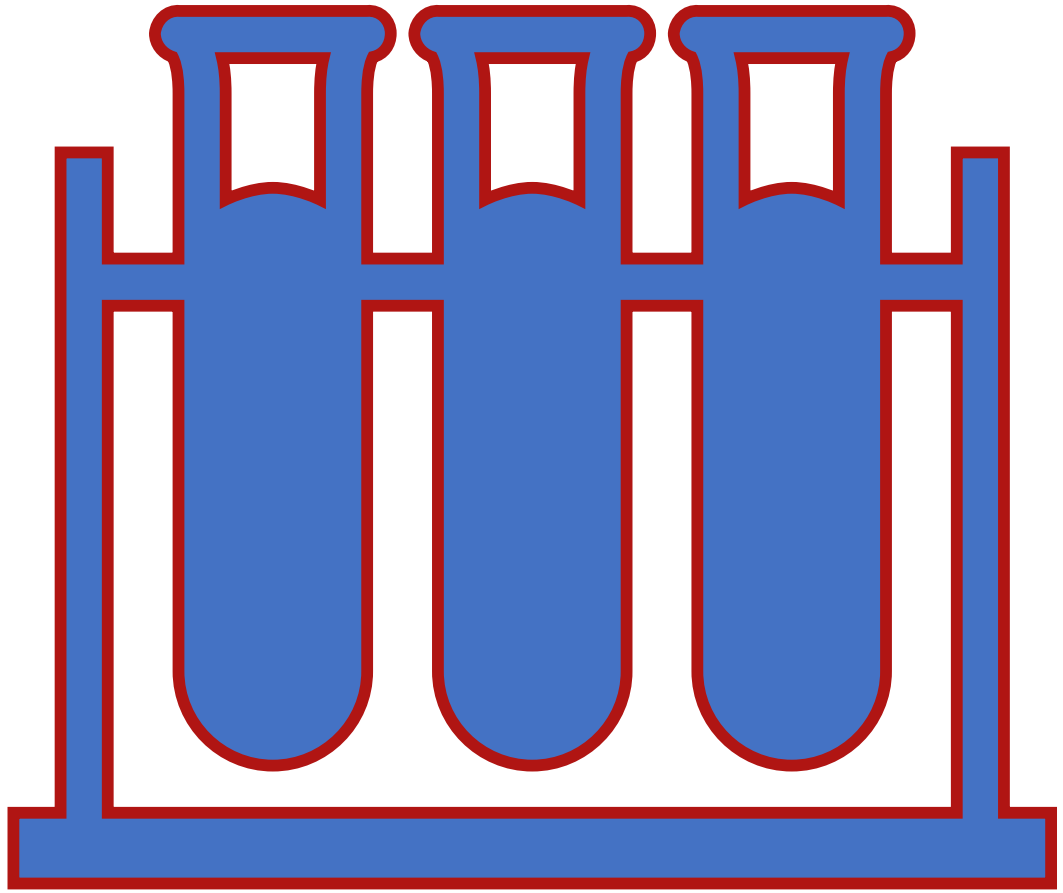
Functional data

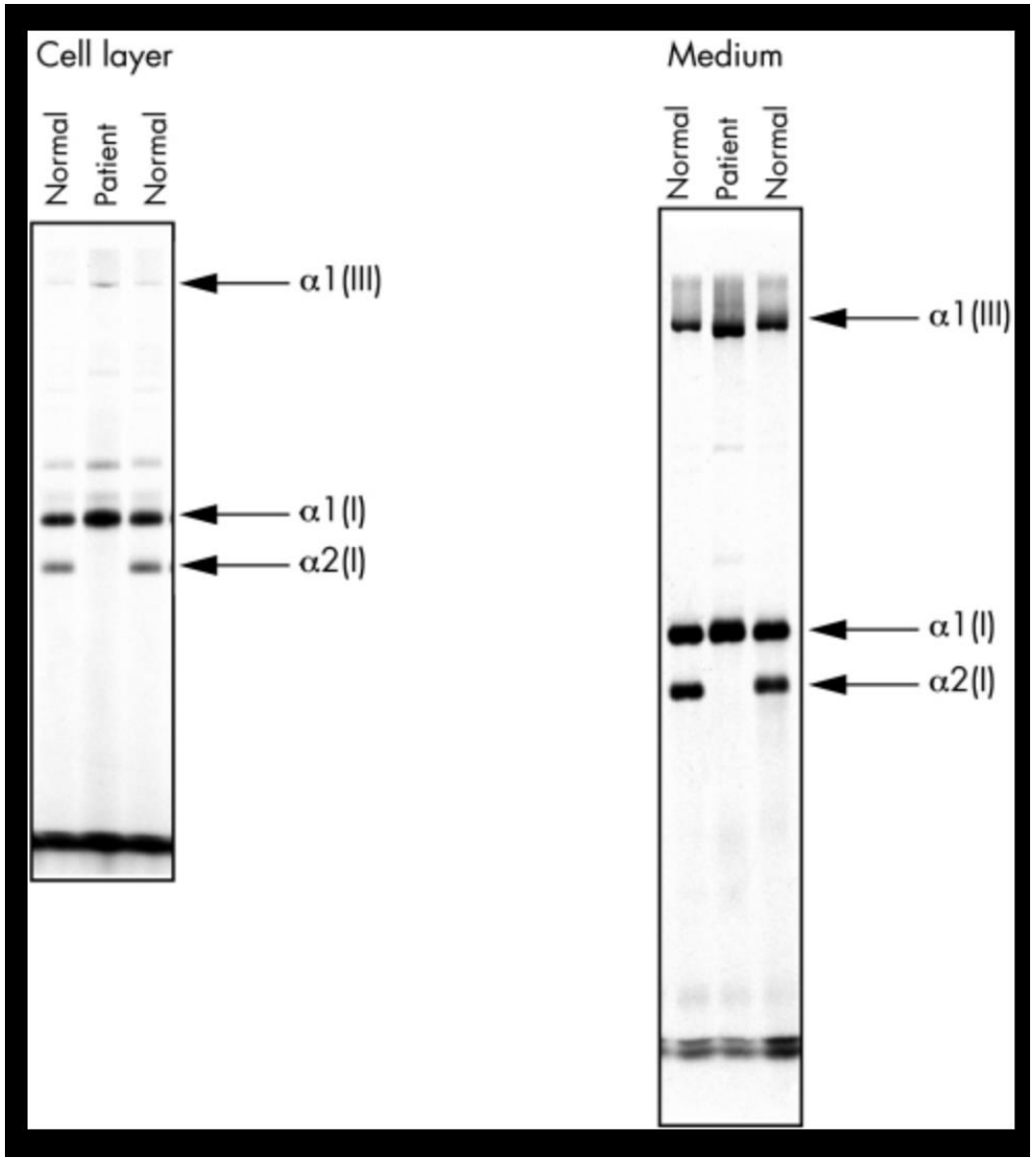
--> when you have well-validated functional data

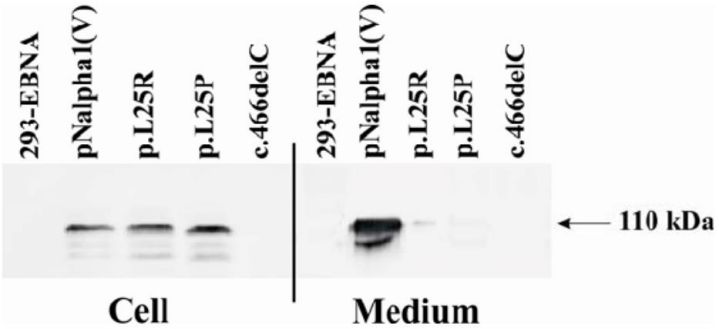


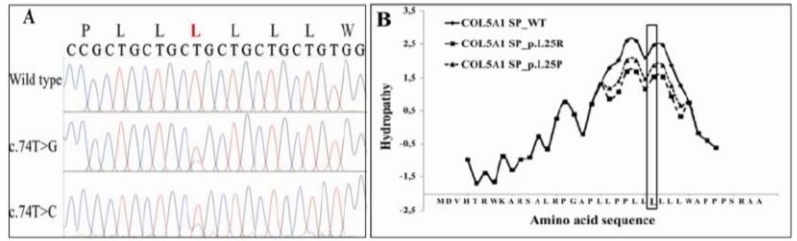
What is a well-validated functional assay?

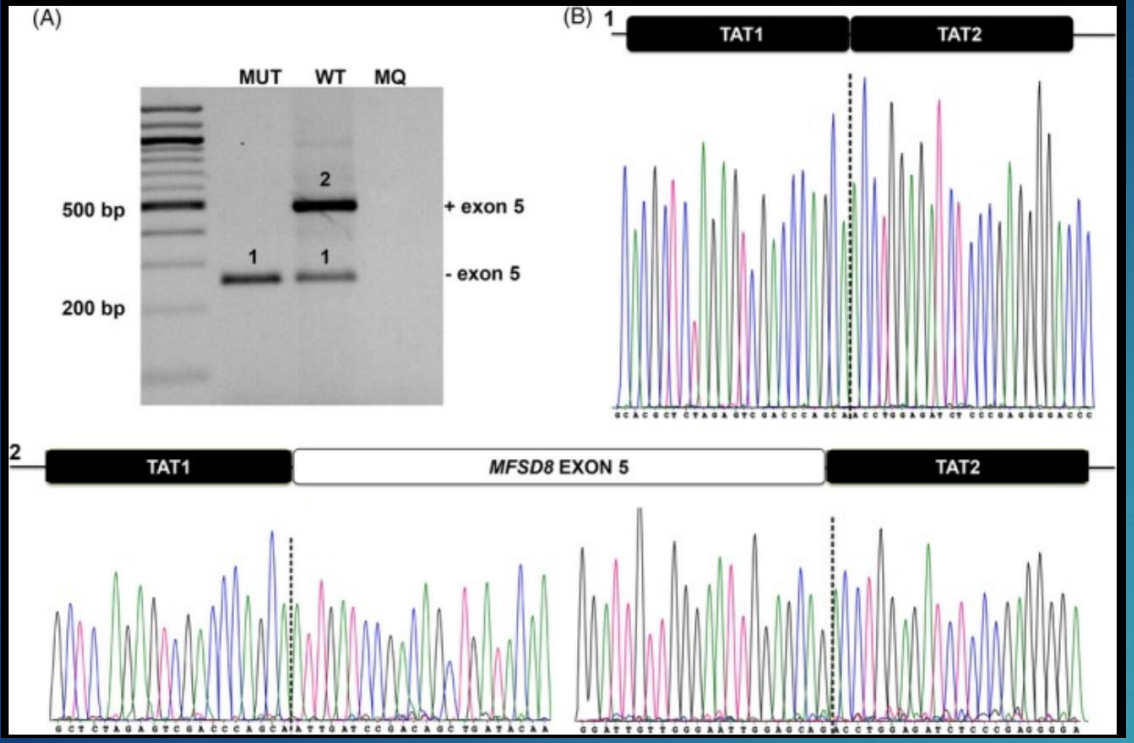
- = Assay that investigates the effect on the gene product
- Activity
  - direct or indirect
- Interaction(s)
  - multimer assembly
- Localization
  - subcellular compartment
- Structure
  - now mostly *in silico*
  - splicing

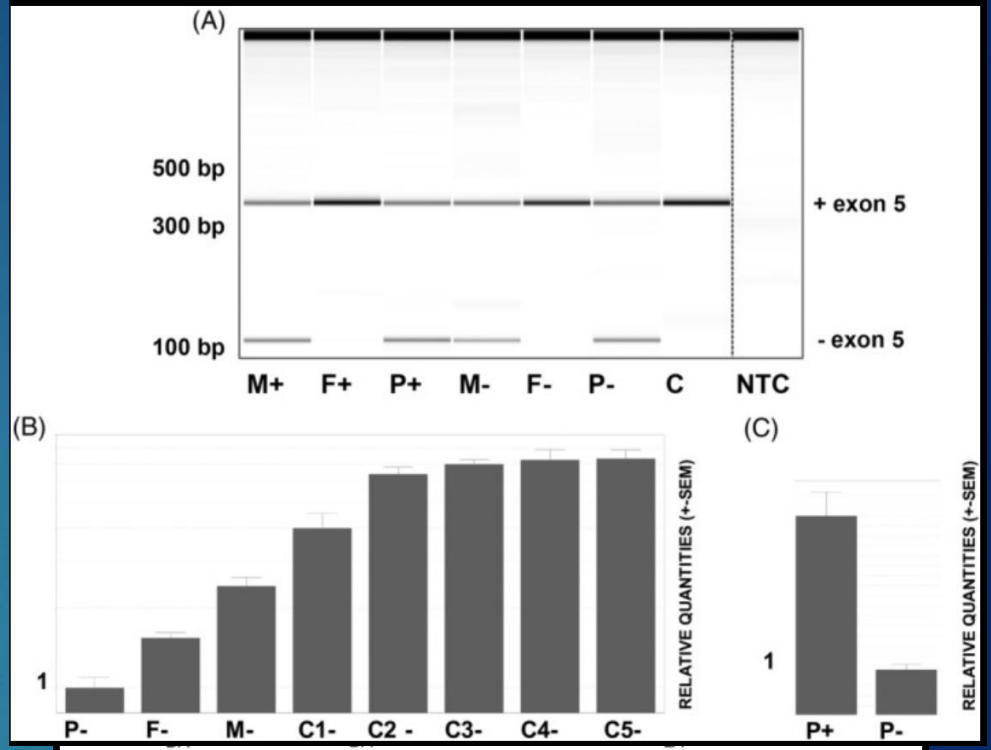


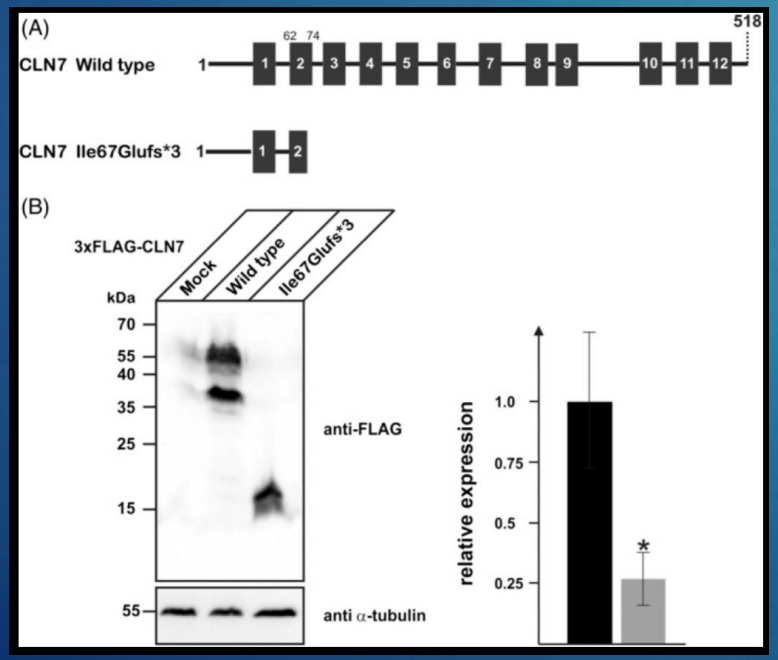

Cell layer
Medium
Normal
Patient
Normal
Normal
Patient
Normal
α1(III)
α1(I)
α2(I)
α1(III)
α1(I)
α2(I)

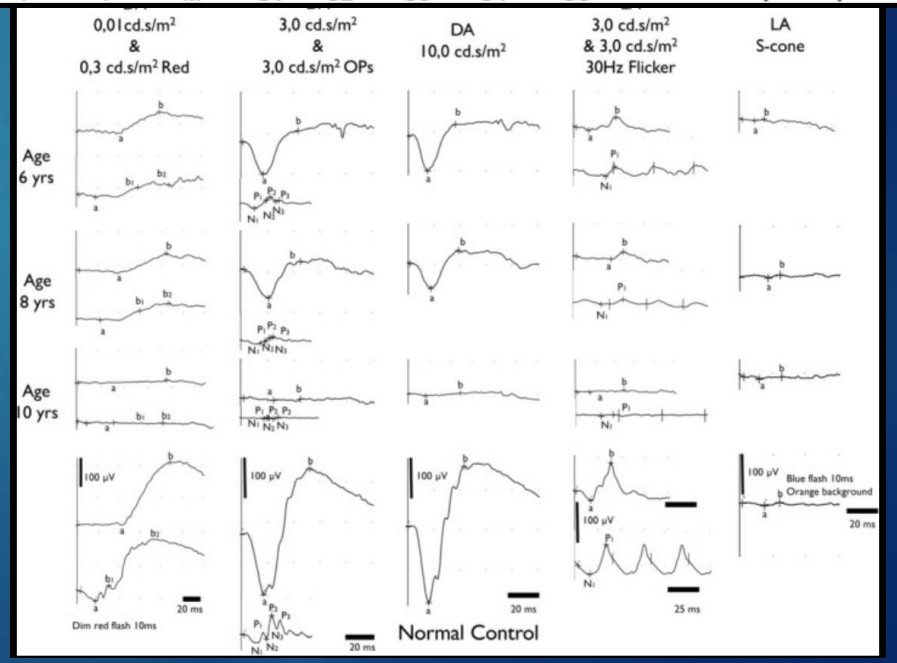

293-EBNA
pNalpha1(V)
p.L25R
p.L25P
c.466delC
293-EBNA
pNalpha1(V)
p.L25R
p.L25P
c.466delC
110 kDa
Cell
Medium


A
P L L L L L L W
CCGCTGCTGCTGCTGCTGCTGTGG
Wild type
c.74T>G
c.74T>C
B
COL5A1 SP_WT
COL5A1 SP_p.L25R
COL5A1 SP_p.L25P
Hydropathy
Amino acid sequence







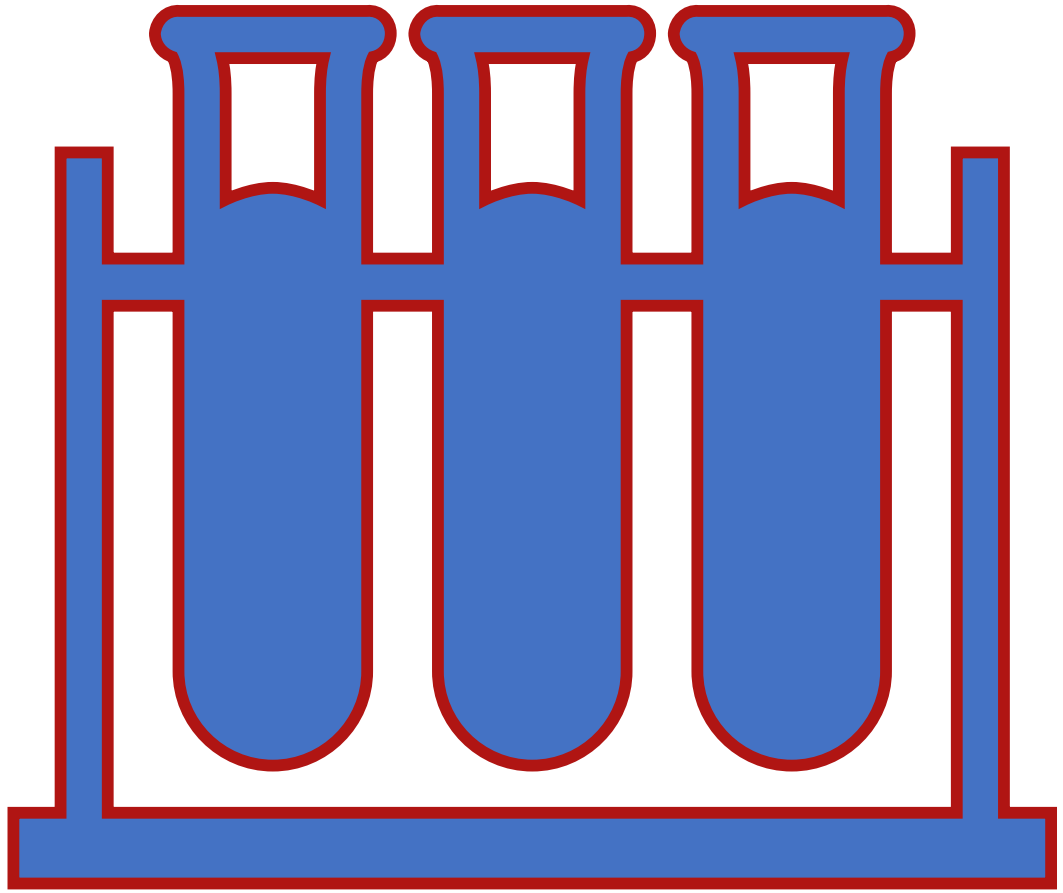


Functional characterization of novel MFSD8 pathogenic variants anticipates neurological involvement in juvenile isolated maculopathy. Bauwens M., et al. Clin Genet. 2020 Mar;97(3):426-436.

What is a well-validated functional assay?

- = Assay that investigates the effect on the gene product

- Compare wild type status with variant status
- Statistical significance ≠ clinical/biological significance
- Compare common variants with rare variants
- Face validity of the assay
- Construct validity of the assay (e.g. genotype)
- *In vivo* assays in proband where possible



CNV classification

ACMG TECHNICAL STANDARDS | Genetics inMedicine

Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen)

Erin Rooney Riggs, MS, CGC[1], Erica F. Andersen, PhD[2,3], Athena M. Cherry, PhD[4], Sibel Kantarci, PhD[5], Hutton Kearney, PhD[6], Ankita Patel, PhD[7], Gordana Raca, MD, PhD[8], Deborah I. Ritter, PhD[9], Sarah T. South, PhD[10], Erik C. Thorland, PhD[6], Daniel Pineda-Alvarez, MD[11], Swaroop Aradhya, PhD[4,11] and Christa Lese Martin, PhD[1]