MAYO CLINIC

CENTER FOR INDIVIDUALIZED MEDICINE

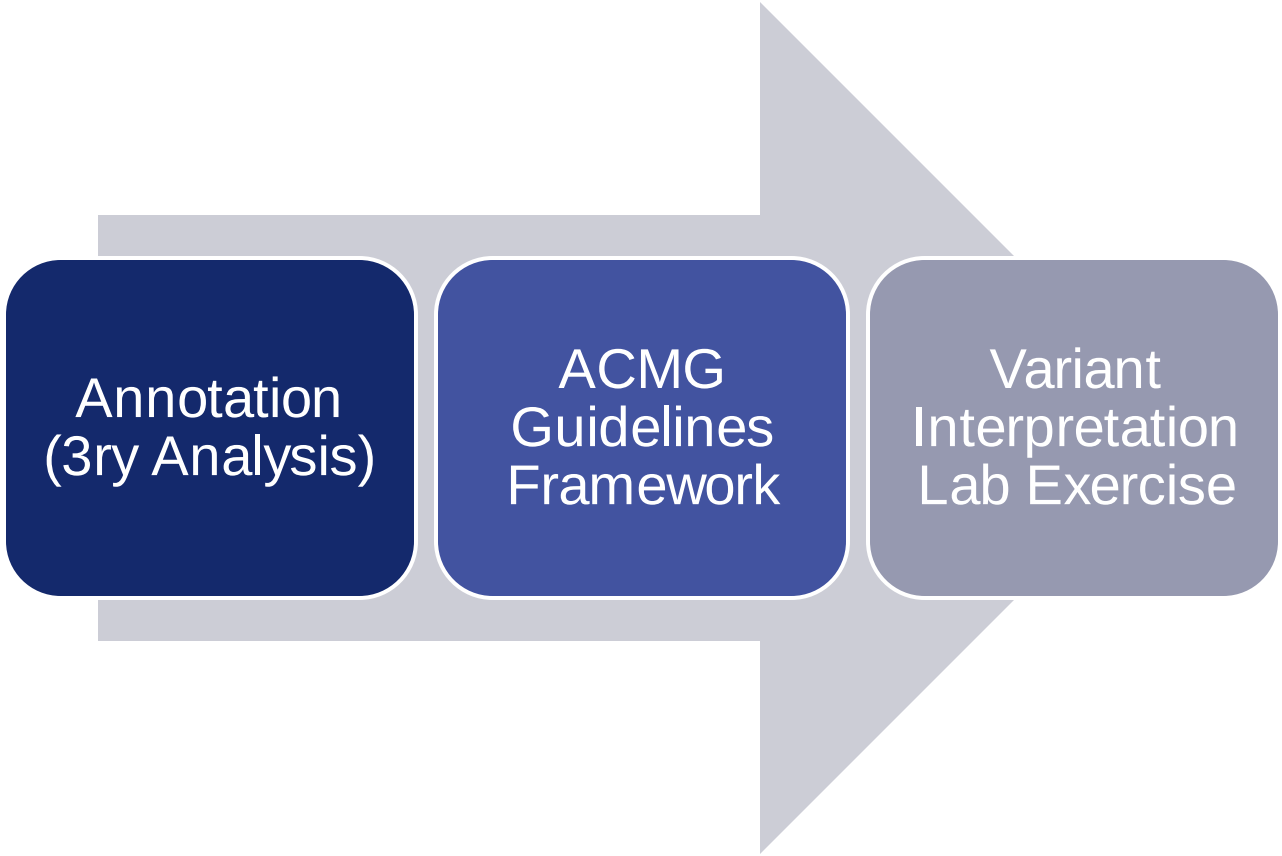
Clinical Variant Interpretation

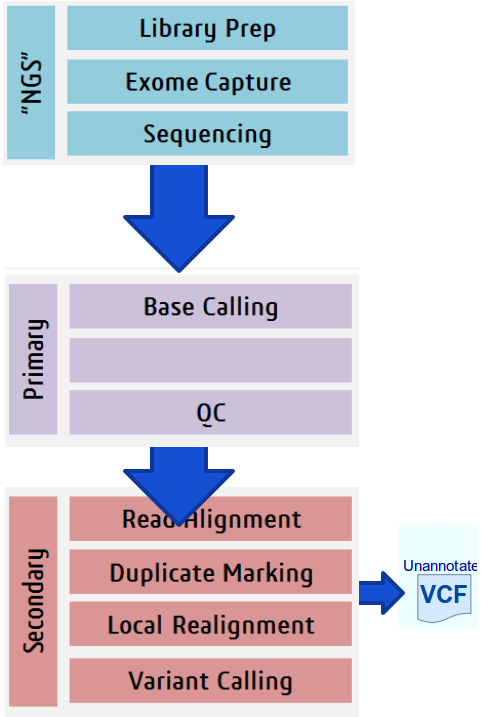
June 22, 2021

Joel A. Morales-Rosado, MD
Post Doctoral Research Fellow
Translational Omics Program
Quantitative Health Sciences Department

Objectives:

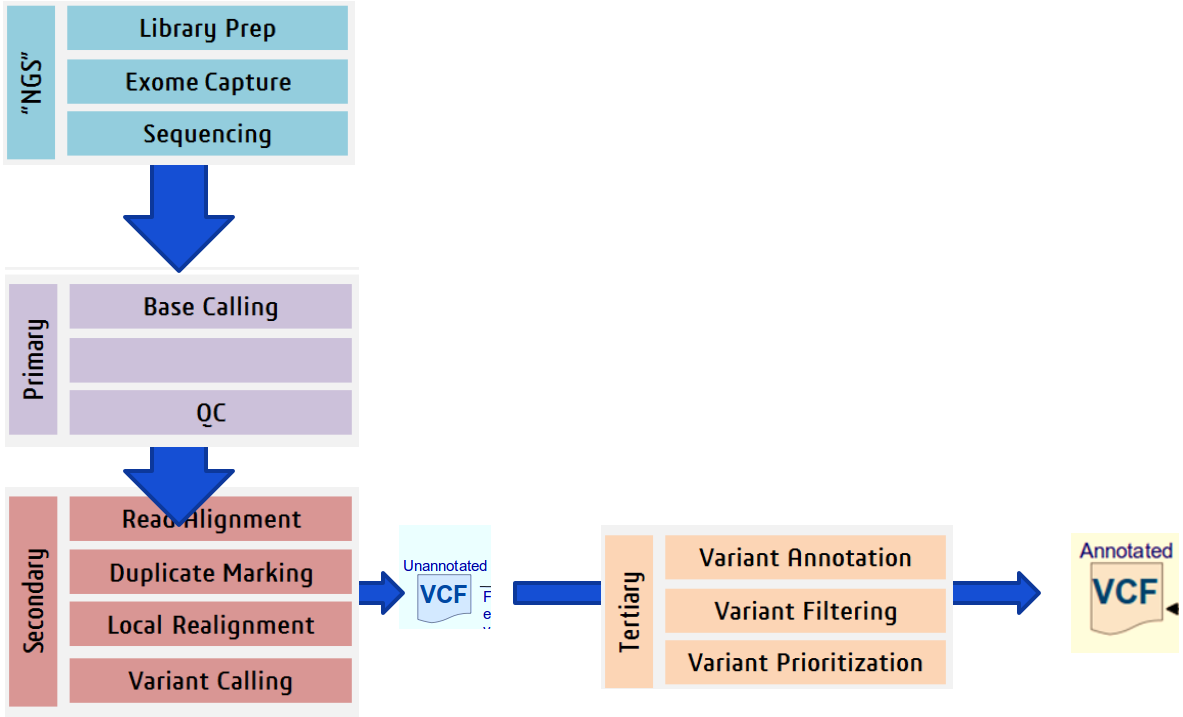
- Overview of the current framework for analysis and interpretation of sequence variants for monogenic disorders
- Overview of key available resources and their utility with variant interpretation.

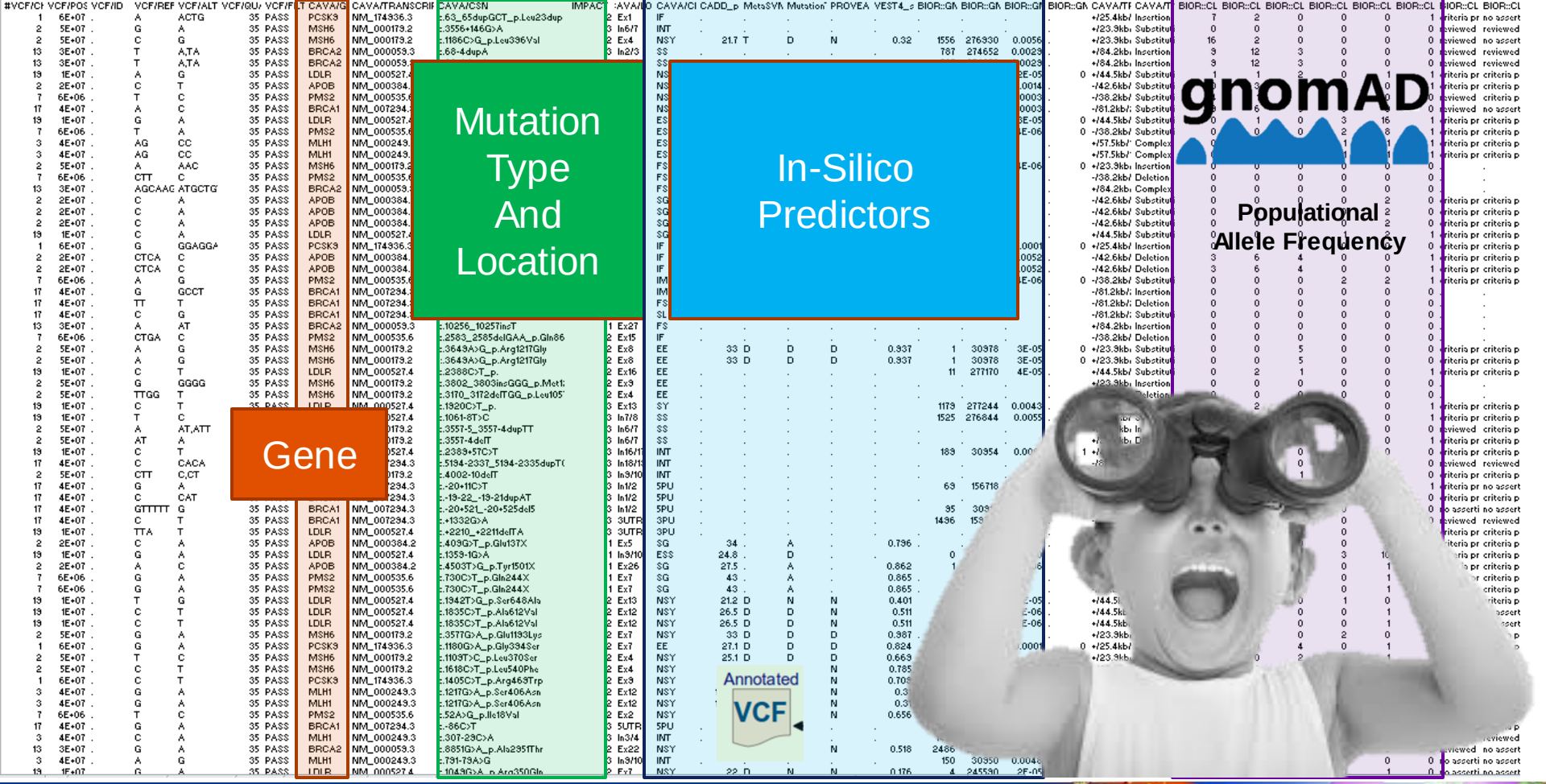

Annotation (3ry Analysis)
ACMG Guidelines Framework
Variant Interpretation Lab Exercise




```
##FILTER=<ID=HETEROPLASMY,Description="Called heterozygous genotype on mitochondrial contig">
##FORMAT=<ID=GP,Number=G,Type=Float,Description="Estimated Genotype Probability">
##FILTER=<ID=IMP,Description="Set if true: IMP==1">
##FILTER=<ID=BOOSTED,Description="Set if true: BOOSTED==1">
##FILTER=<ID=LOWDP,Description="Set if GQ>20 and 10<=DP<=20">
##FILTER=<ID=LOWQ,Description="Set if GQ<=20 or DP<10">
##FILTER=<ID=NOTVALIDATED,Description="Set if variant falls outside of analytic range">
##FORMAT=<ID=GL,Number=G,Type=Float,Description="Genotype likelihoods">
##FORMAT=<ID=VAR_TYPE,Number=.,Type=String,Description="Variant type: SNV, INSERTION, DELETION, SUBSTITUTION, MNV, COMPLEX">
##FORMAT=<ID=VAR_CONTEXT,Number=.,Type=String,Description="Variant genomic context: STR-expansion, STR-contraction, STR-proximal">
##FORMAT=<ID=STR_MAX_LEN,Number=1,Type=Integer,Description="Maximum observed STR sequence length">
##FORMAT=<ID=STR_PERIOD,Number=1,Type=Integer,Description="Repetition period for STR variants">
##FORMAT=<ID=STR_TIMES,Number=1,Type=Float,Description="Number of repetition for STR variants">
##pipeline=helix-v2.6.1
```

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	PC-TA537BFRZC6B332GA9D'
chr1	55033879	.	A	ACTG	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47805173	.	G	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47793163	.	C	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr13	32319070	.	T	A,TA	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11113686	.	A	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	21011802	.	C	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr7	5977709	.	T	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43094795	.	A	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11102787	.	G	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr7	6003794	.	T	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr3	37028782	.	AG	CC	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47798826	.	A	AAC	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr7	5987451	.	CTT	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr13	32340378	.	AGCAAC	ATGCTG	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	21038086	.	C	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11123663	.	C	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr1	55033930	.	G	GGAGGA	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	21010226	.	CTCA	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr7	6003018	.	A	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43124034	.	G	GCCT	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43124037	.	TT	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43045679	.	C	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr13	32338763	.	A	AT	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr7	5973402	.	CTGA	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47806206	.	A	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11128084	.	C	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47806452	.	G	GGGG	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47801152	.	TTGG	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11120166	.	C	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11111506	.	T	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47805601	.	A	AT,ATT	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47805601	.	AT	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11128142	.	C	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43053469	.	C	CACA	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47806751	.	CTT	C,CT	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43125260	.	G	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43124135	.	C	CAT	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43124745	.	GTTTTT	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr17	43044346	.	C	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47806983	.	A	AGTTC	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr13	11133511	.	TTA	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	21038086	.	C	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr13	11113534	.	G	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	21012365	.	A	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr7	5937393	.	G	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr19	11120188	.	T	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr13	11116388	.	C	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47805638	.	G	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr1	55057514	.	G	A	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47793032	.	T	C	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47793601	.	C	T	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV
chr2	47806206	.	A	G	35	PASS	.	GT:AD:DP:GQ:VAR_TYPE	0/1:500,500:1000:99:SNV


"NGS"
Library Prep
Exome Capture
Sequencing
Primary
Base Calling
QC
Secondary
Read Alignment
Duplicate Marking
Local Realignment
Variant Calling
Unannotated
VCF
Tertiary
Variant Annotation
Variant Filtering
Variant Prioritization
Annotated
VCF


Mutation Type And Location
In-Silico Predictors
gnomAD
Populational Allele Frequency
Gene
Annotated VCF

Class 2

"NGS"
- Library Prep
- Exome Capture
- Sequencing

Primary
- Base Calling
- QC

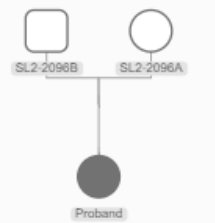
Secondary
- Read Alignment
- Duplicate Marking
- Local Realignment
- Variant Calling

Unannotated VCF

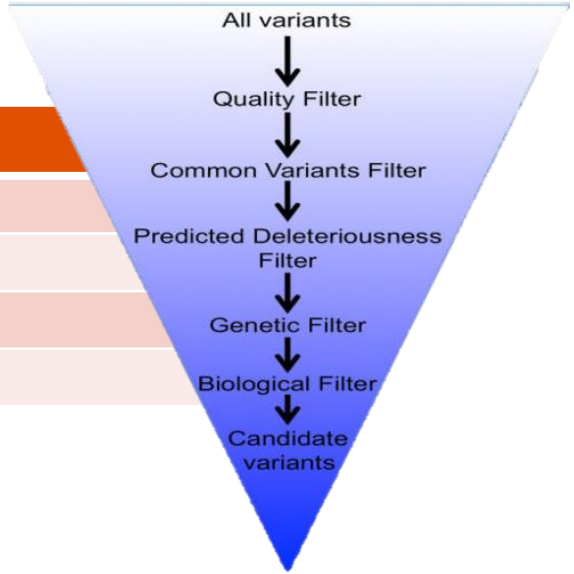
Tertiary
- Variant Annotation
- Variant Filtering
- Variant Prioritization

Annotated VCF

WES Trio



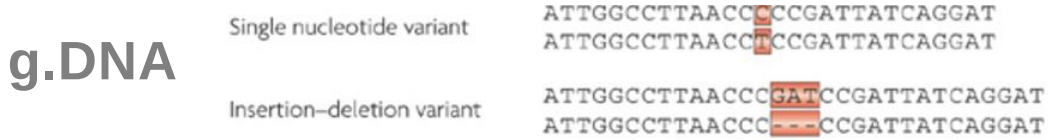
Filter	Variants
None	125,746
DP>10, QUAL>20	25,541
MAF<0.01	1,263
Coding regions	551

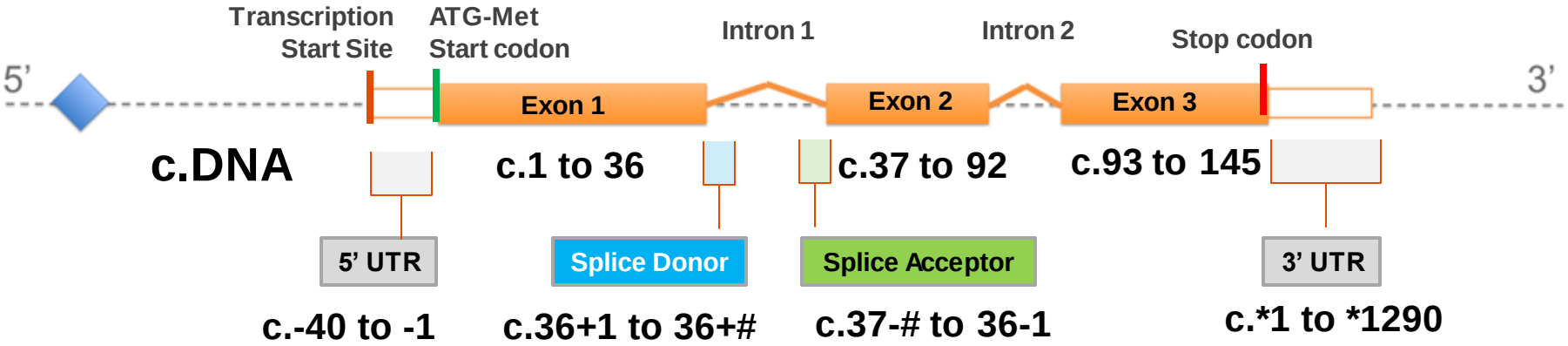


Experimentally Defined Genomic Features:
~Mutation/Variant Impact or Type – Most impactful?

1.	MSH2	NM_000251.2	c.1590A>G_p.Glu530=
2.	BRCA1	NM_007294.3	c.736T>G_p.Leu246Val
3.	BRCA2	NM_000059.3	c.6474delT_p.Gln2159AsnfsTer9
4.	PMS2	NM_000535.6	c.730C>T_p.Gln244X
5.	PCSK9	NM_174936.3	c.524-1G>A
6.	MLH1	NM_000249.3	c.307-29C>A

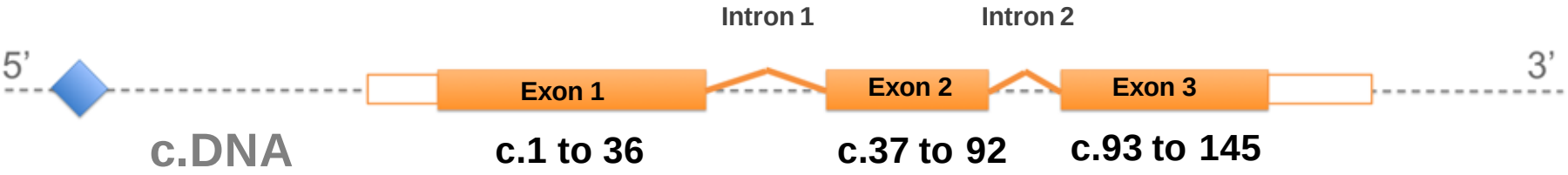
Experimentally Defined Genomic Features:
~Mutation/Variant Impact or Type – cDNA Definitions





MLH1 NM_000249.3 c.307-29C>A

Experimentally Defined Genomic Features:
~Mutation/Variant Impact or Type – Protein Level



Protein Consequences:

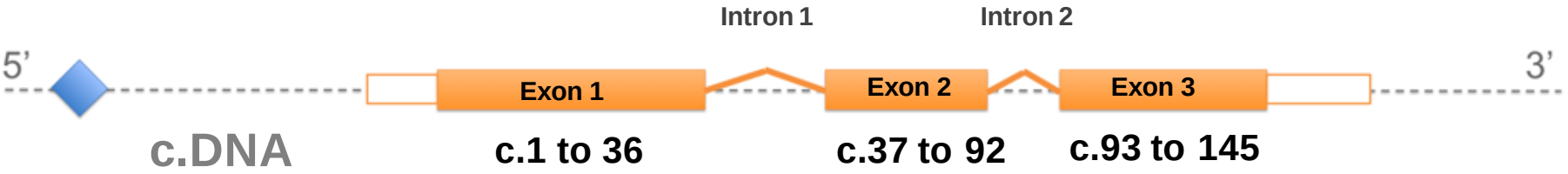
	No mutation	Point mutations			
		Silent	Nonsense	Missense	
				conservative	non-conservative
DNA level	TTC	TTT	ATC	TCC	TGC
mRNA level	AAG	AAA	UAG	AGG	ACG
protein level	Lys	Lys	STOP	Arg	Thr

MSH2
c.1590A>G
p.Glu530=
"synonymous"

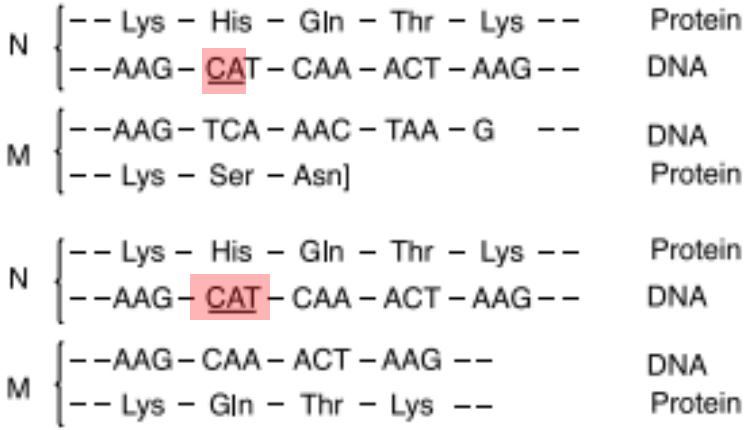
PMS2
c.730C>T
p.Gln244X
p.Gln244Ter

BRCA1
c.736T>G
p.Leu246Val

Experimentally Defined Genomic Features:
~Mutation/Variant Impact or Type – Protein Level



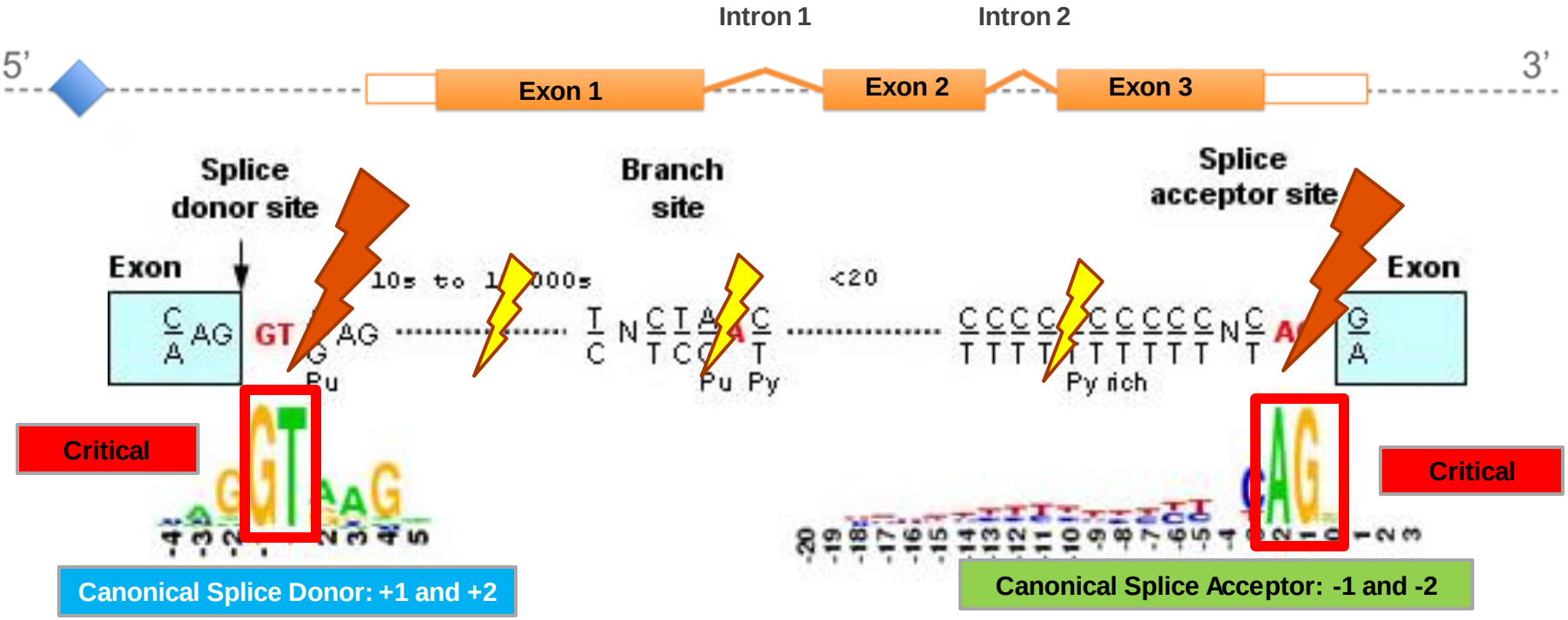
Protein Consequences



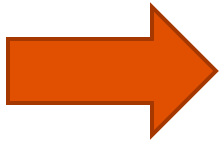
Frameshift
c.6474delT
p.Gln2159AsnfsTer9

In-frame deletion
c.6639_6641delTGA
p.Asp2213del

Experimentally Defined Genomic Features:
~Mutation/Variant Impact or Type – Splicing Disruption



Annotation of single gDNA positions or VCF files with public web resources

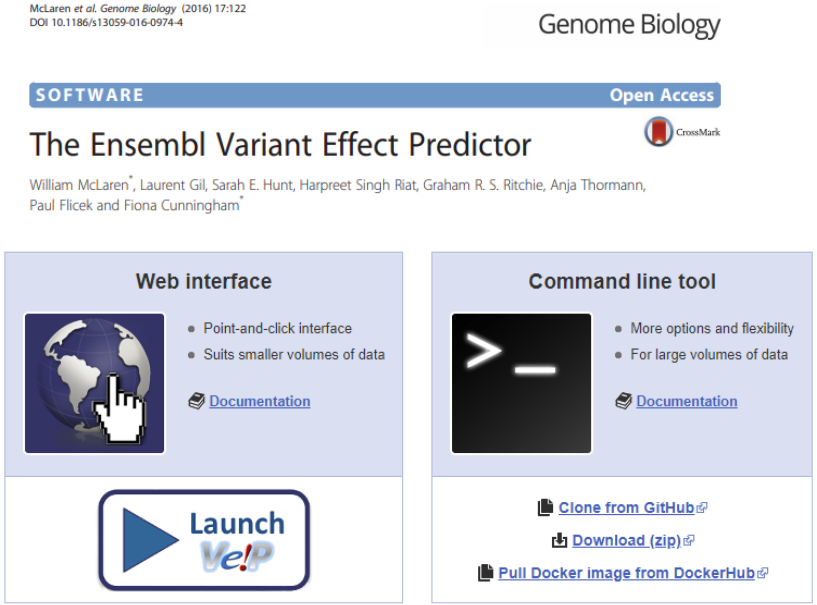


#CHROM	POS	ID	REF	ALT	QUAL	FILTER
chr1	55033873	.	A	ACTG	35	PASS
chr2	47805173	.	G	A	35	PASS
chr2	47793163	.	C	G	35	PASS
chr13	32319070	.	T	A,TA	35	PASS
chr19	11113686	.	A	G	35	PASS
chr2	21011802	.	C	T	35	PASS
chr7	5977703	.	T	C	35	PASS
chr17	43094735	.	A	C	35	PASS
chr19	11102787	.	G	A	35	PASS
chr7	6003794	.	T	A	35	PASS
chr3	37028782	.	AG	CC	35	PASS
chr2	47798826	.	A	AAC	35	PASS
chr7	5987451	.	CTT	C	35	PASS
chr13	32340378	.	AGCAAG	ATGCTG	35	PASS
chr2	21038086	.	C	A	35	PASS

Important: Always ask which genome build are the coordinates.

Experimentally Defined Genomic Features:
VEP: Variant Effect Predictor

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



https://useast.ensembl.org/info/docs/tools/vep/index.html

Experimentally Defined Genomic Features:
VEP: Variant Effect Predictor

Variant Effect Predictor

New job

Recent jobs

Refresh

Show/hide columns (1 hidden)

Analysis	Jobs	Submitted at
Variant Effect Predictor	VEP analysis of pasted data in Homo_sapiens Done [View results]	06/02/2020, 18:59 (GMT)
Variant Effect Predictor	VEP analysis of pasted data in Homo_sapiens Done [View results]	24/01/2020, 21:53 (GMT)
Variant Effect Predictor	VEP analysis of pasted data in Homo_sapiens Done [View results]	24/01/2020, 21:43 (GMT)

Experimentally Defined Genomic Features:
VEP: Variant Effect Predictor

Species: Human (Homo sapiens)

Assembly: GRCh38.p13 *(If you are looking for VEP for Human GRCh37, please go to GRCh37 website.)*

Name for this job (optional):

Input data:

Either paste data:

```
17 43047665 . C T . . .
```

Examples: Ensembl default, VCF, Variant identifiers, HGVS notations, SPDI

Or upload file: Choose File No file chosen

Or provide file URL:

Transcript database to use:

- Ensembl/GENCODE transcripts
- Ensembl/GENCODE basic transcripts
- RefSeq transcripts
- Ensembl/GENCODE and RefSeq transcripts

Identifiers ⊞ *Additional identifiers for genes, transcripts and variants*

Variants and frequency data ⊞ *Co-located variants and frequency data*

Additional annotations ⊞ *Addtional transcript, protein and regulatory annotations*

Predictions ⊞ *Variant predictions, e.g. SIFT, PolyPhen*

Filtering options ⊞ *Pre-filter results by frequency or consequence type*

Advanced options ⊞ *Settings to optimise VEP*

https://useast.ensembl.org/info/docs/tools/vep/index.html

Experimentally Defined Genomic Features:
VEP: Variant Effect Predictor

Species: Human (Homo sapiens)

Assembly: GRCh38.p13 *(If you are looking for VEP for Human GRCh37, please go to GRCh37 website.)*

Name for this job (optional):

Input data:

Either paste data:

```
17 43047665 . C T . . .
```

Examples: Ensembl default, VCF, Variant identifiers, HGVS notations, SPDI

Or upload file: Choose File No file chosen

Or provide file URL:

Transcript database to use:

- Ensembl/GENCODE transcripts
- Ensembl/GENCODE basic transcripts
- RefSeq transcripts
- Ensembl/GENCODE and RefSeq transcripts

Identifiers *Additional identifiers for genes, transcripts and variants*

Variants and frequency data *Co-located variants and frequency data*

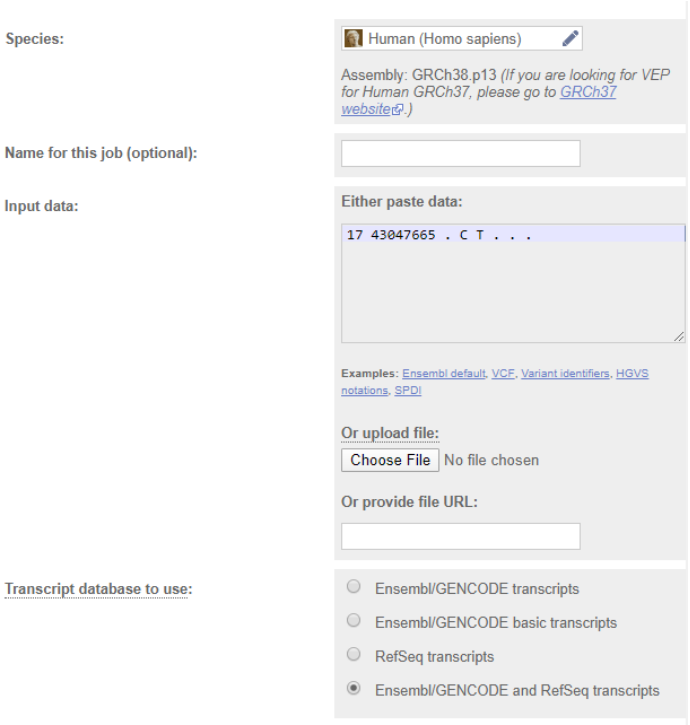
Analysis	Jobs	Submitted at
Variant Effect Predictor	VEP analysis of pasted data in Homo_sapiens Queued	26/02/2020, 04:29 (GMT)

Filtering options *Pre-filter results by frequency or consequence type*

Advanced options *Settings to optimise VEP*

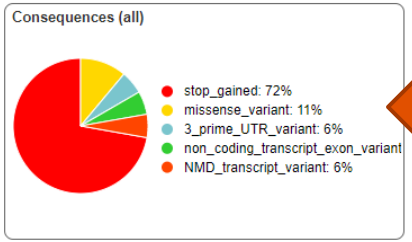
https://useast.ensembl.org/info/docs/tools/vep/index.html

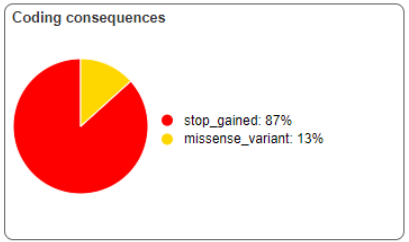
Experimentally Defined Genomic Features:
VEP: Variant Effect Predictor



BRCA1
c.5445G>A
p.Trp1815X

Category	Count
Variants processed	1
Variants filtered out	0
Novel / existing variants	0 (0.0) / 1 (100.0)
Overlapped genes	2
Overlapped transcripts	17
Overlapped regulatory features	0





https://useast.ensembl.org/info/docs/tools/vep/index.html

VEP: Variant Effect Predictor

Uploaded variant	Location	Allele	Consequence	Symbol	Exon	HGVSc	HGVSp	cDNA position	CDS position	Protein position	Amino acids	Codons	Existing variant	Feature strand	MANE
.	17:43047665-43047665	T	stop_gained	BRCA1	21/22	ENST00000352993.7:c.2019G>A	ENSP00000312236.5:p.Trp673Ter	2138	2019	673	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	22/23	ENST00000357654.9:c.5445G>A	ENSP00000350283.3:p.Trp1815Ter	5558	5445	1815	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	NM_007294.4
.	17:43047665-43047665	T	3_prime_UTR_variant, NMD_transcript_variant	BRCA1	22/23	ENST00000461221.5:c.*5228G>A	-	5546	-	-	-	-	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	missense_variant	BRCA1	21/22	ENST00000468300.5:c.2059G>A	ENSP00000417148.1:p.Asp687Asn	2253	2059	687	D/N	GAC/AAC	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	23/24	ENST00000471181.7:c.5508G>A	ENSP00000418960.2:p.Trp1836Ter	5740	5508	1836	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	22/23	ENST00000491747.6:c.2133G>A	ENSP00000420705.2:p.Trp711Ter	2232	2133	711	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	21/22	ENST00000493795.5:c.5304G>A	ENSP00000418775.1:p.Trp1768Ter	5536	5304	1768	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	7/8	ENST00000586385.5:c.375G>A	ENSP00000465818.1:p.Trp125Ter	519	375	125	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	10/11	ENST00000591534.5:c.918G>A	ENSP00000467329.1:p.Trp306Ter	1020	918	306	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	4/5	ENST00000591849.5:c.144G>A	ENSP00000465347.1:p.Trp48Ter	301	144	48	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	14/15	ENST00000644379.1:c.1832G>A	ENSP00000496570.1:p.Trp611Ter	1832	1833	611	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-
.	17:43047665-43047665	T	stop_gained	BRCA1	22/23	NM_007294.4:c.5445G>A	NP_009225.1:p.Trp1815Ter	5558	5445	1815	W/*	TGG/TGA	rs397509284, CM042679, COSV58785802	-1	-

Gene: BRCA1 ENSG00000012048

Description	BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100]
Gene Synonyms	BRCC1, FANCS, PPP1R53, RNF53
Location	Chromosome 17: 43,044,295-43,170,245 reverse strand.

MAYO CLINIC

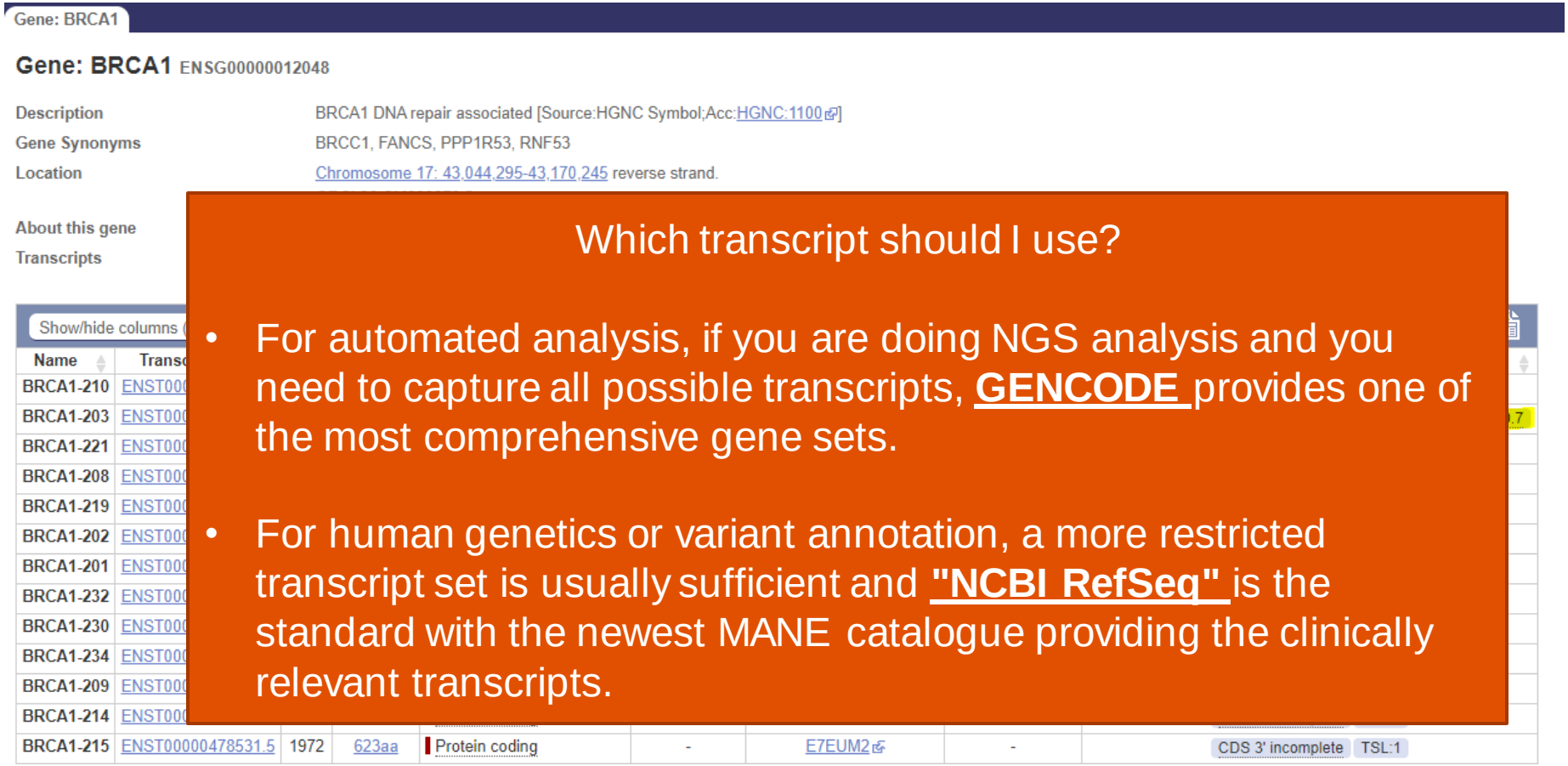
Gene: BRCA1

Gene: BRCA1 ENSG00000012048

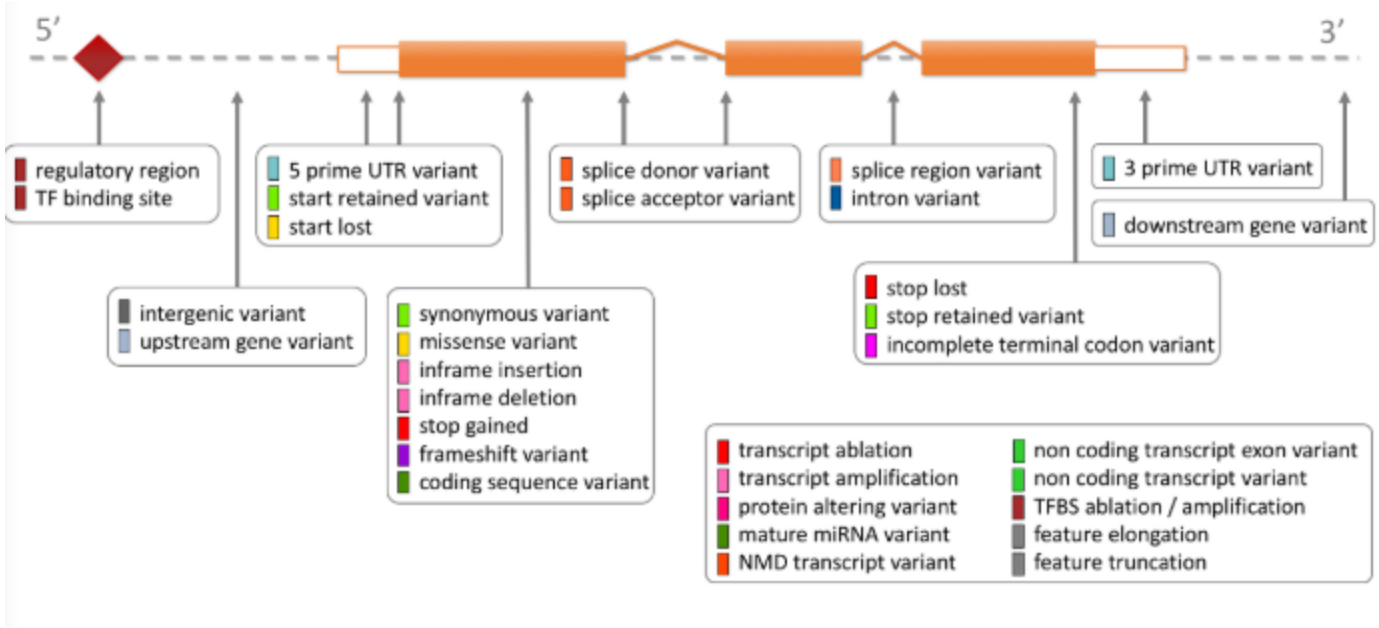
Description	BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100]
Gene Synonyms	BRCC1, FANCS, PPP1R53, RNF53
Location	Chromosome 17: 43,044,295-43,170,245 reverse strand. GRCh38:CM000679.2
About this gene	This gene has 34 transcripts (splice variants), 233 orthologues, is a member of 1 Ensembl protein family and is associated with 80 phenotypes.
Transcripts	Hide transcript table

Show/hide columns (1 hidden) Filter

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq Match	Flags
BRCA1-210	ENST00000471181.7	7270	1884aa	Protein coding	CCDS11456	P38398	-	TSL:1 GENCODE basic APPRIS P4
BRCA1-203	ENST00000357654.9	7088	1863aa	Protein coding	CCDS11453	P38398	NM_007294.4	TSL:1 GENCODE basic APPRIS ALT2 MANE Select v0.7
BRCA1-221	ENST00000493795.5	5732	1816aa	Protein coding	CCDS11459	P38398	-	TSL:5 GENCODE basic
BRCA1-208	ENST00000468300.5	3273	699aa	Protein coding	CCDS11455	P38398	-	TSL:1 GENCODE basic
BRCA1-219	ENST00000491747.6	2379	759aa	Protein coding	CCDS11454	A0A024R1V0 P38398	-	TSL:5 GENCODE basic
BRCA1-202	ENST00000354071.7	4497	1399aa	Protein coding	-	Q5YLB2	-	TSL:1 GENCODE basic
BRCA1-201	ENST00000352993.7	3668	721aa	Protein coding	-	A0A024R1Z8 P38398	-	TSL:5 GENCODE basic
BRCA1-232	ENST00000644379.1	2571	659aa	Protein coding	-	A0A2R8Y7V5	-	CDS 5' incomplete
BRCA1-230	ENST00000634433.1	2534	798aa	Protein coding	-	A0A0U1RRA9	-	CDS 3' incomplete TSL:5
BRCA1-234	ENST00000652672.1	2291	601aa	Protein coding	-	A0A494C182	-	CDS 3' incomplete
BRCA1-209	ENST00000470026.5	2108	649aa	Protein coding	-	E7EWN5	-	CDS 3' incomplete TSL:1
BRCA1-214	ENST00000477152.5	1980	622aa	Protein coding	-	E9PH68	-	CDS 3' incomplete TSL:1
BRCA1-215	ENST00000478531.5	1972	623aa	Protein coding	-	E7EUM2	-	CDS 3' incomplete TSL:1

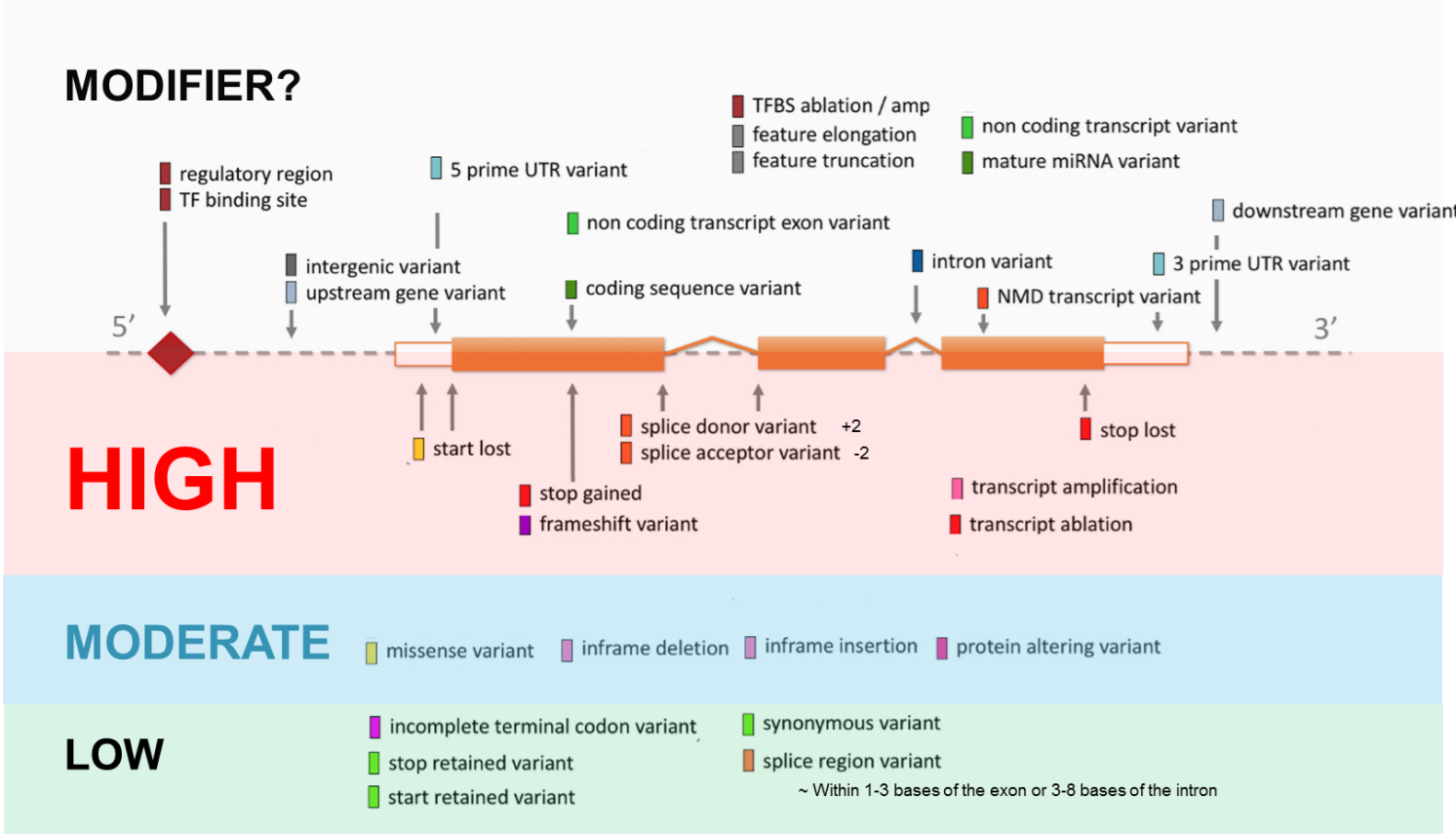

Gene: BRCA1
Gene: BRCA1 ENSG00000012048
Description
BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100]
Gene Synonyms
BRCC1, FANCS, PPP1R53, RNF53
Location
Chromosome 17: 43,044,295-43,170,245 reverse strand.
About this gene
Transcripts
Show/hide columns
Name
BRCA1-210
BRCA1-203
BRCA1-221
BRCA1-208
BRCA1-219
BRCA1-202
BRCA1-201
BRCA1-232
BRCA1-230
BRCA1-234
BRCA1-209
BRCA1-214
BRCA1-215
ENST00000478531.5
1972
623aa
Protein coding
E7EUM2
CDS 3' incomplete
TSL:1
Which transcript should I use?
For automated analysis, if you are doing NGS analysis and you need to capture all possible transcripts, GENCODE provides one of the most comprehensive gene sets.
For human genetics or variant annotation, a more restricted transcript set is usually sufficient and "NCBI RefSeq" is the standard with the newest MANE catalogue providing the clinically relevant transcripts.

Variant Severity: Variable definitions but helps prioritize



https://m.ensembl.org/info/genome/variation/prediction/predicted_data.html

Variant Severity: Variable definitions but helps prioritize



Where is it? Which exon?
Regulatory elements nearby?
Visualization is key!

#CHROM	POS	ID	REF	ALT	QUAL	FILTER
chr1	55039879	.	A	ACTG	35	PASS
chr2	47805173	.	G	A	35	PASS
chr2	47799169	.	C	G	35	PASS
chr13	32319070	.	T	A,TA	35	PASS
chr19	11113686	.	A	G	35	PASS
chr2	21011802	.	C	T	35	PASS
chr7	5977709	.	T	C	35	PASS
chr17	43094795	.	A	C	35	PASS
chr19	11102787	.	G	A	35	PASS
chr7	6003794	.	T	A	35	PASS
chr3	37028782	.	AG	CC	35	PASS
chr2	47798826	.	A	AAC	35	PASS
chr7	5987451	.	CTT	C	35	PASS
chr13	32340378	.	AGCAAC	ATGCTG	35	PASS
chr2	21038086	.	C	A	35	PASS

Experimentally Defined Genomic Features:
UCSC Genome Browser - Visualization

UNIVERSITY OF CALIFORNIA SANTA CRUZ Genomics Institute

UCSC

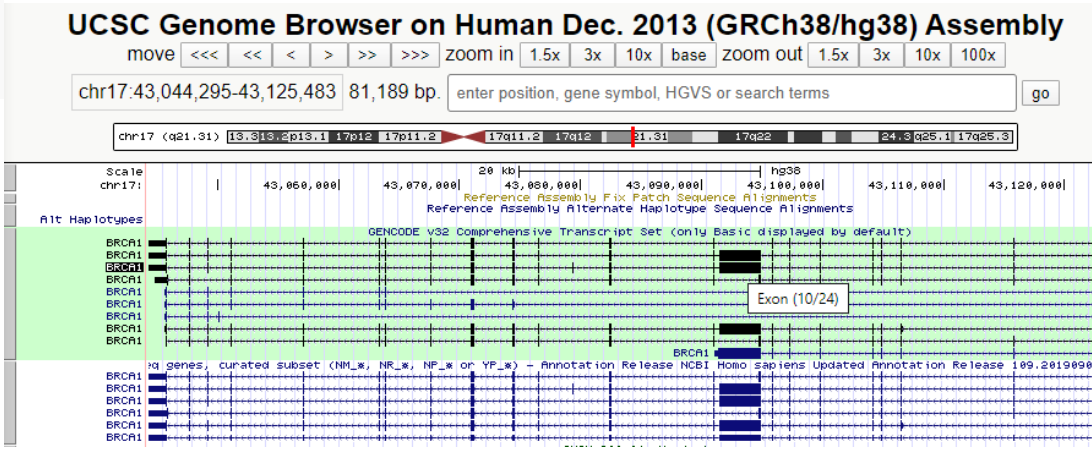
D756–D761 *Nucleic Acids Research, 2020, Vol. 48, Database issue* *Published online 6 November 2019*
doi: 10.1093/nar/gkz1012

UCSC Genome Browser enters 20th year

Christopher M. Lee[1,*], Galt P. Barber[1], Jonathan Casper[1], Hiram Clawson[1], Mark Diekhans[1], Jairo Navarro Gonzalez[1], Angie S. Hinrichs[1], Brian T. Lee[1], Luis R. Nassar[1], Conner C. Powell[1], Brian J. Raney[1], Kate R. Rosenbloom[1], Daniel Schmelter[1], Matthew L. Speir[1], Ann S. Zweig[1], David Haussler[1,2], Maximilian Haeussler[1], Robert M. Kuhn[1] and W. James Kent[1]

[1]Genomics Institute, University of California Santa Cruz, Santa Cruz, CA 95064, USA and [2]Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, CA 95064, USA

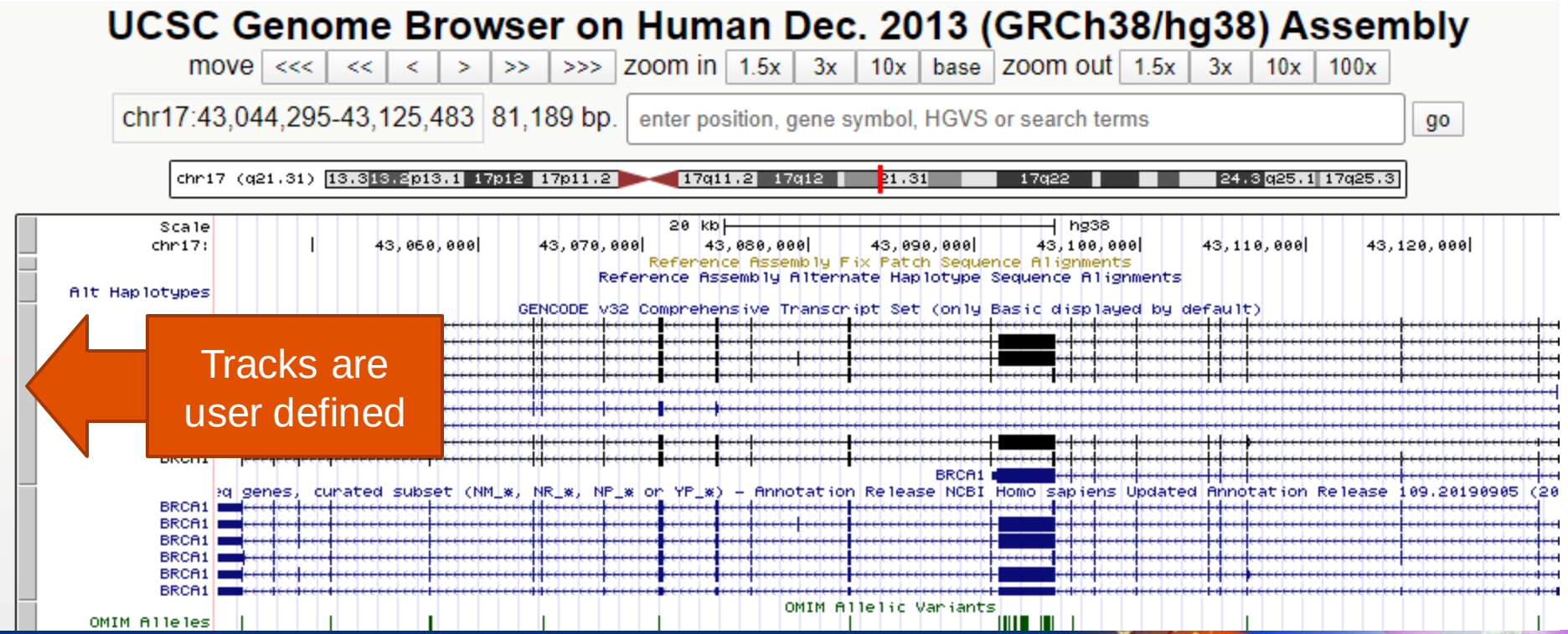
Received September 11, 2019; Revised October 16, 2019; Editorial Decision October 17, 2019; Accepted October 25, 2019



- Web-based viewer for genome sequence data and annotations.
- Steadily added data and software features to the website since first coming online in July 2000, and currently hosts 206 assemblies from 105 species

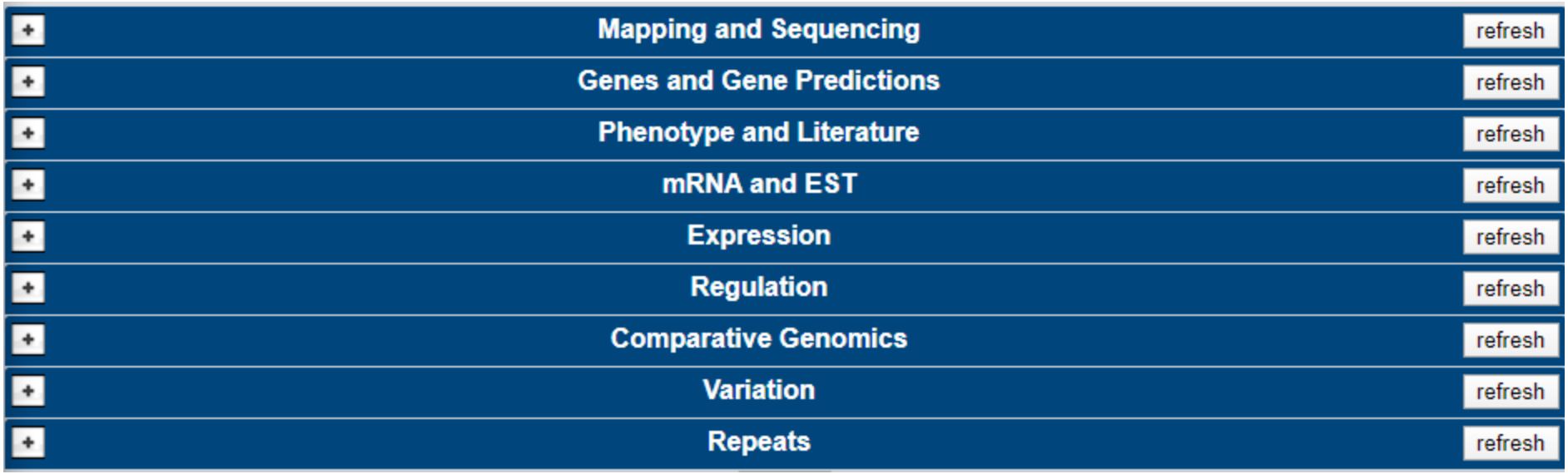
https://genome.ucsc.edu/index.html

Experimentally Defined Genomic Features:
UCSC Genome Browser - Visualization

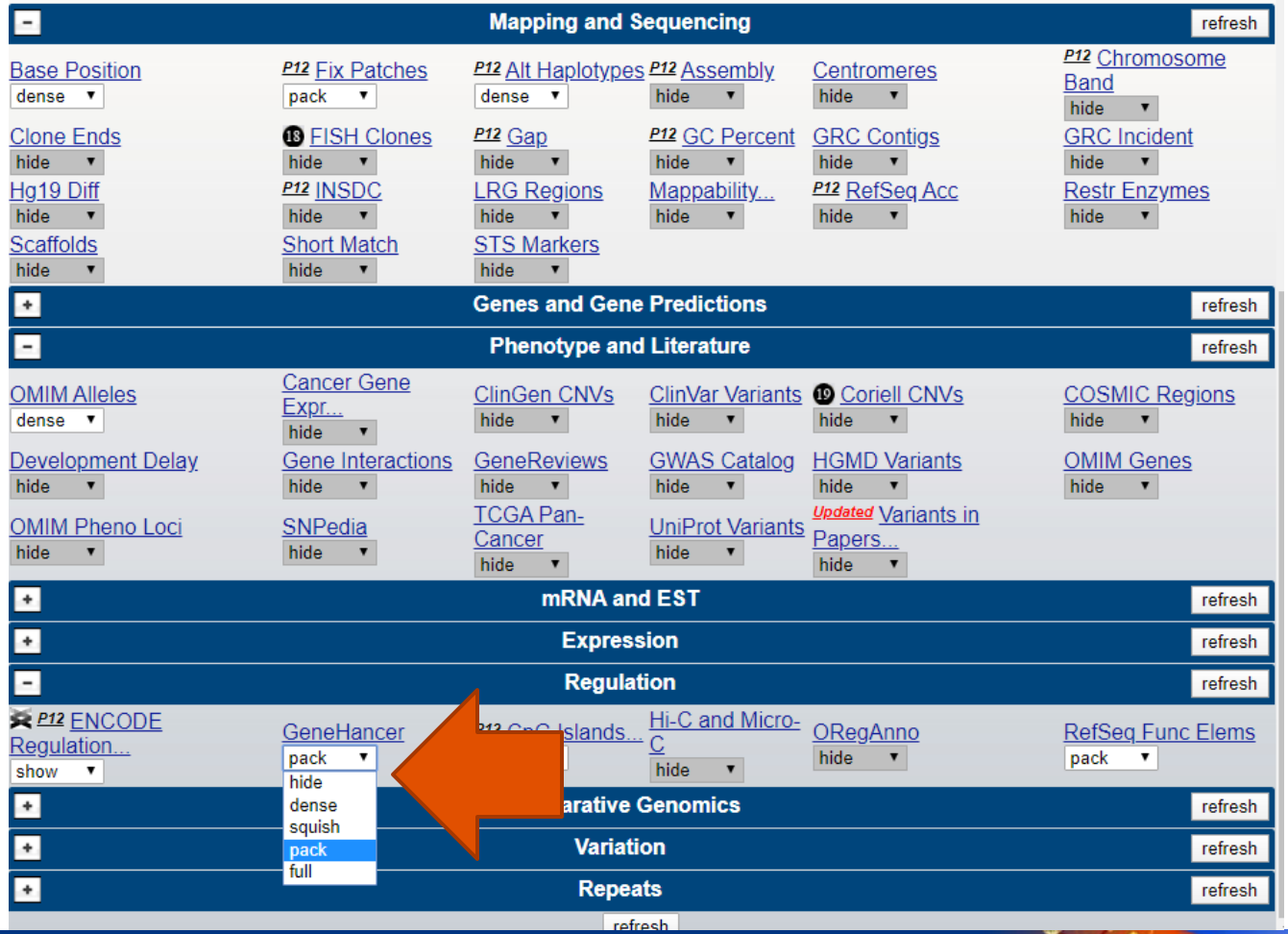


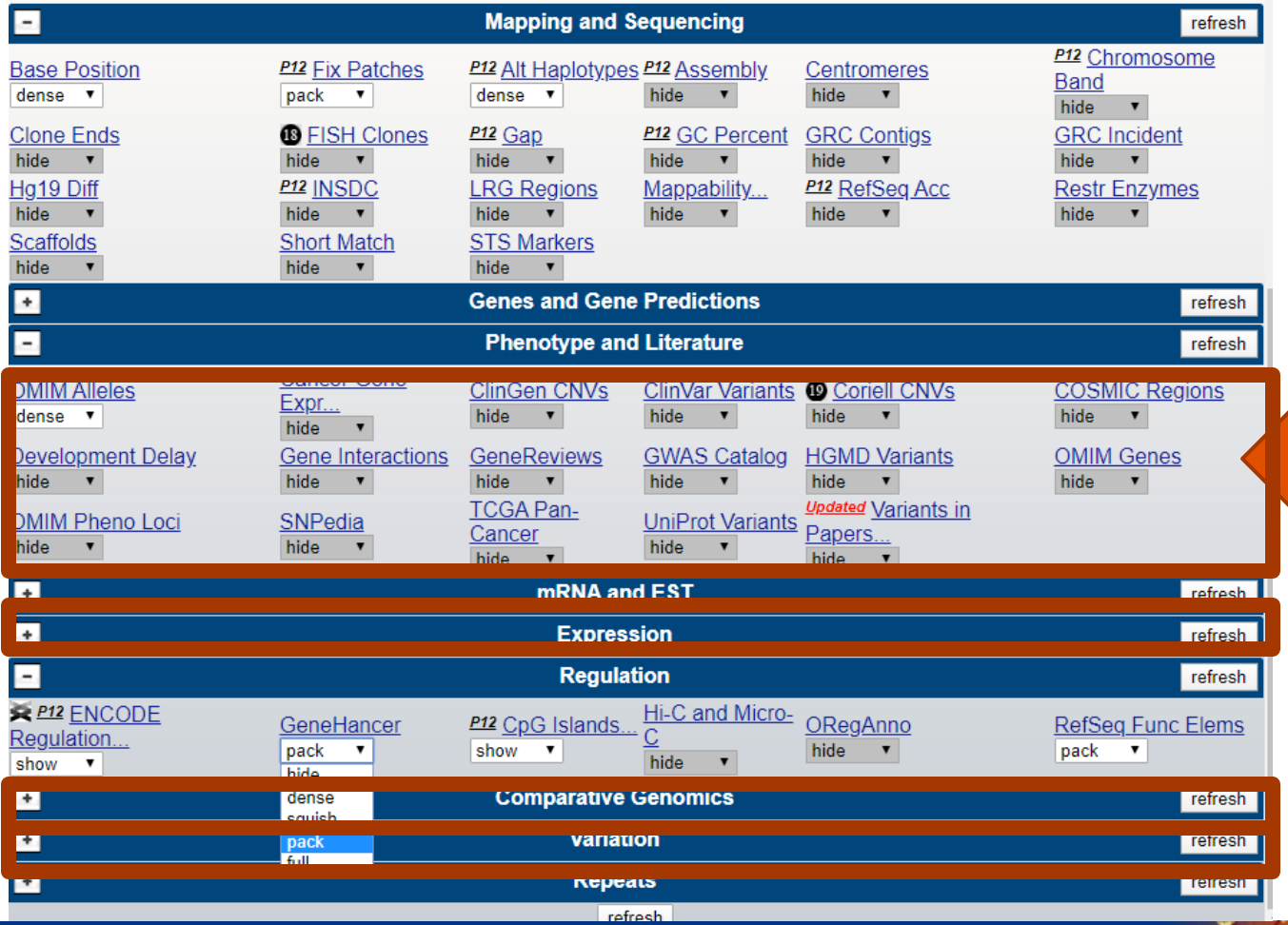
https://genome.ucsc.edu/index.html

Experimentally Defined Genomic Features:
UCSC Genome Browser – Tracks Setup



https://genome.ucsc.edu/index.html

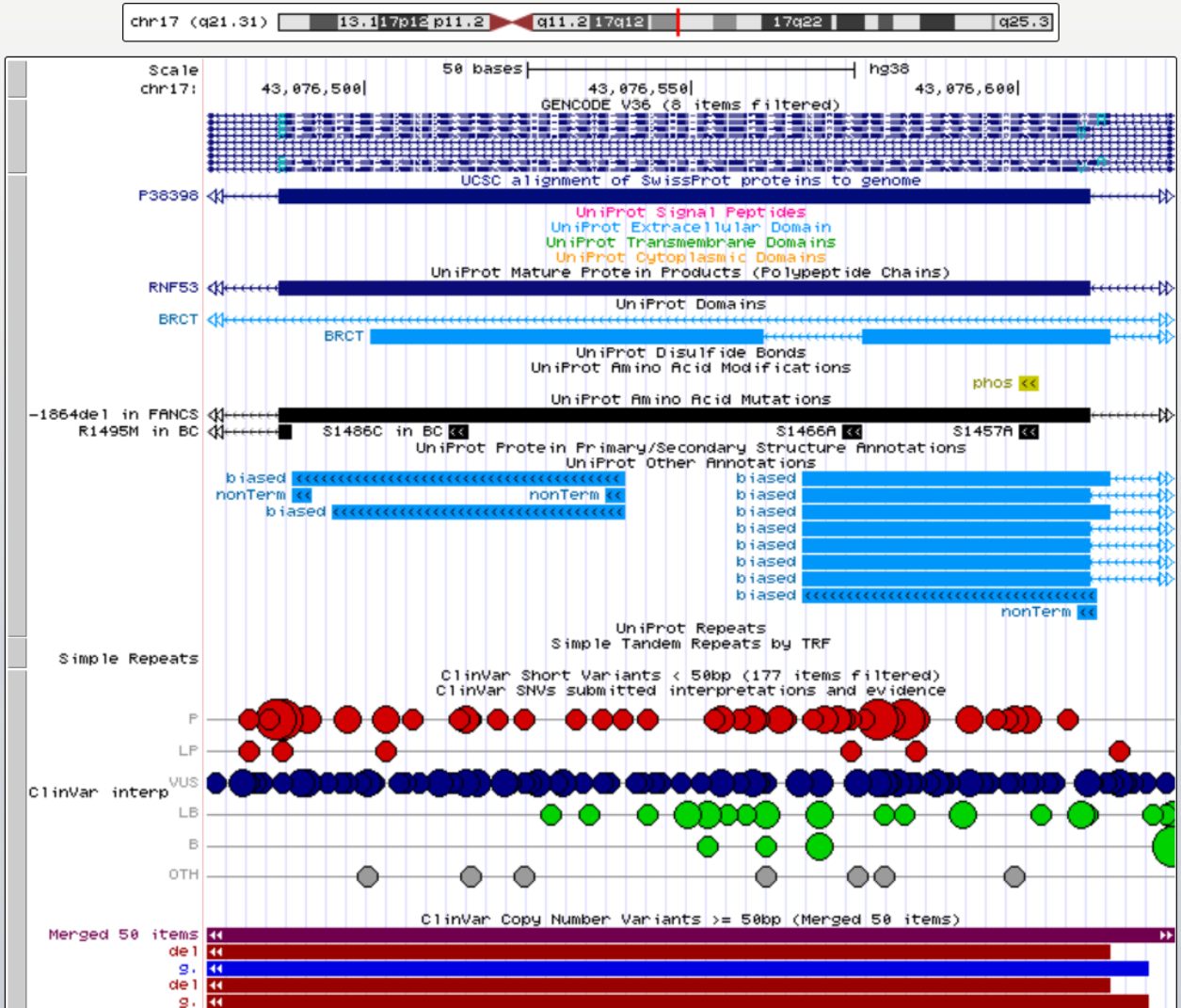

Mapping and Sequencing
refresh
Base Position
dense
P12 Fix Patches
pack
P12 Alt Haplotypes
dense
P12 Assembly
hide
Centromeres
hide
P12 Chromosome Band
hide
Clone Ends
hide
FISH Clones
hide
P12 Gap
hide
P12 GC Percent
hide
GRC Contigs
hide
GRC Incident
hide
Hg19 Diff
hide
P12 INSDC
hide
LRG Regions
hide
Mappability...
hide
P12 RefSeq Acc
hide
Restr Enzymes
hide
Scaffolds
hide
Short Match
hide
STS Markers
hide
Genes and Gene Predictions
refresh
Phenotype and Literature
refresh
OMIM Alleles
dense
Cancer Gene Expr...
hide
ClinGen CNVs
hide
ClinVar Variants
hide
Coriell CNVs
hide
COSMIC Regions
hide
Development Delay
hide
Gene Interactions
hide
GeneReviews
hide
GWAS Catalog
hide
HGMD Variants
hide
OMIM Genes
hide
OMIM Pheno Loci
hide
SNPedia
hide
TCGA Pan-Cancer
hide
UniProt Variants
hide
Updated Variants in Papers...
hide
mRNA and EST
refresh
Expression
refresh
Regulation
refresh
P12 ENCODE Regulation...
show
GeneHancer
pack
hide
dense
squish
pack
full
Hi-C and Micro-C
hide
ORegAnno
hide
RefSeq Func Elems
pack
refresh
Variation
refresh
Repeats
refresh



Part 2

Biodatabases for gene and variant curation.

https://genome.ucsc.edu/index.html


chr17 (q21.31)
13.1 17p12 p11.2
q11.2 17q12
17q22
q25.3
Scale
50 bases
hg38
chr17:
43,076,500
43,076,550
43,076,600
GENCODE V36 (8 items filtered)
UCSC alignment of SwissProt proteins to genome
P38398
UniProt Signal Peptides
UniProt Extracellular Domain
UniProt Transmembrane Domains
UniProt Cytoplasmic Domains
UniProt Mature Protein Products (Polypeptide Chains)
RNF53
UniProt Domains
BRCT
BRCT
UniProt Disulfide Bonds
UniProt Amino Acid Modifications
phos
UniProt Amino Acid Mutations
-1864del in FANCS
R1495M in BC
S1486C in BC
S1466A
S1457A
UniProt Protein Primary/Secondary Structure Annotations
UniProt Other Annotations
biased
nonTerm
nonTerm
biased
biased
biased
biased
biased
biased
biased
biased
biased
nonTerm
UniProt Repeats
Simple Tandem Repeats by TRF
Simple Repeats
ClinVar Short Variants < 50bp (177 items filtered)
ClinVar SNVs submitted interpretations and evidence
P
LP
VUS
ClinVar interp
LB
B
OTH
ClinVar Copy Number Variants >= 50bp (Merged 50 items)
Merged 50 items
del
g.
del
g.

Variant Interpretation: Rationale

- Is a previously published variant associated with a disease phenotype pathogenic?
- Are all variants observed in a control population benign?
- What evidence to we use to ultimately classify a variant?
- How do we ensure consistency among clinicians, clinical laboratories, and researchers?

15-20 years ago…

- If you thought a gene may be implicated in a specific disease:
  - You could screen a cohort of patients and look for variants in said gene
  - If you identified a variant in a patient and did not find it in 50 control samples (100 alleles!!!) you could deem this as a pathogenic variant
- Does this make sense statistically?

Common framework and criteria for germline variant classification

ACMG

American College of Medical Genetics and Genomics

Translating Genes Into Health®

AMP ASSOCIATION FOR MOLECULAR PATHOLOGY

cap

Common framework and criteria for germline variant classification

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[1], Nazneen Aziz, PhD[2,16], Sherri Bale, PhD[3], David Bick, MD[4], Soma Das, PhD[5], Julie Gastier-Foster, PhD[6,7,8], Wayne W. Grody, MD, PhD[9,10,11], Madhuri Hegde, PhD[12], Elaine Lyon, PhD[13], Elaine Spector, PhD[14], Karl Voelkerding, MD[13] and Heidi L. Rehm, PhD[15]; on behalf of the ACMG Laboratory Quality Assurance Committee

We interpret by sorting variants into categories

***Variant:** An alteration in the normal sequence of a gene:

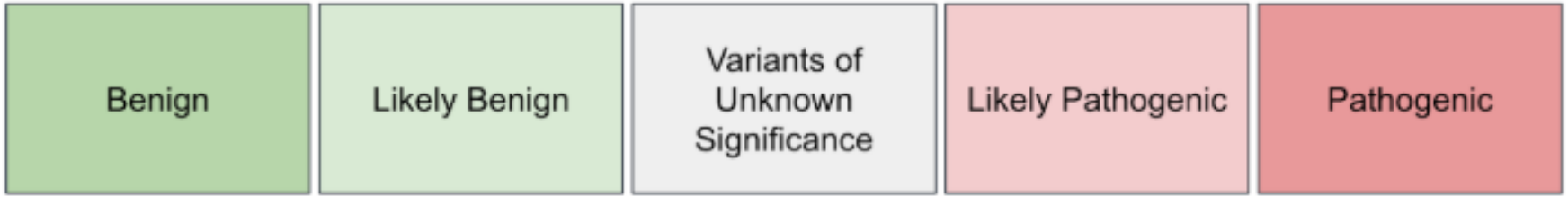
Benign	Variants of Unknown Significance	Pathogenic
A variant that does not appear to have a *deleterious* effect often associated with a "normal" or no human phenotype.	A variant whose association with disease risk is unknown.	A variant which is proven to be deleterious to protein or gene function and is associated with a particular human phenotype. Or disease

Caution: **A deleterious variant is not always pathogenic or disease causing.**

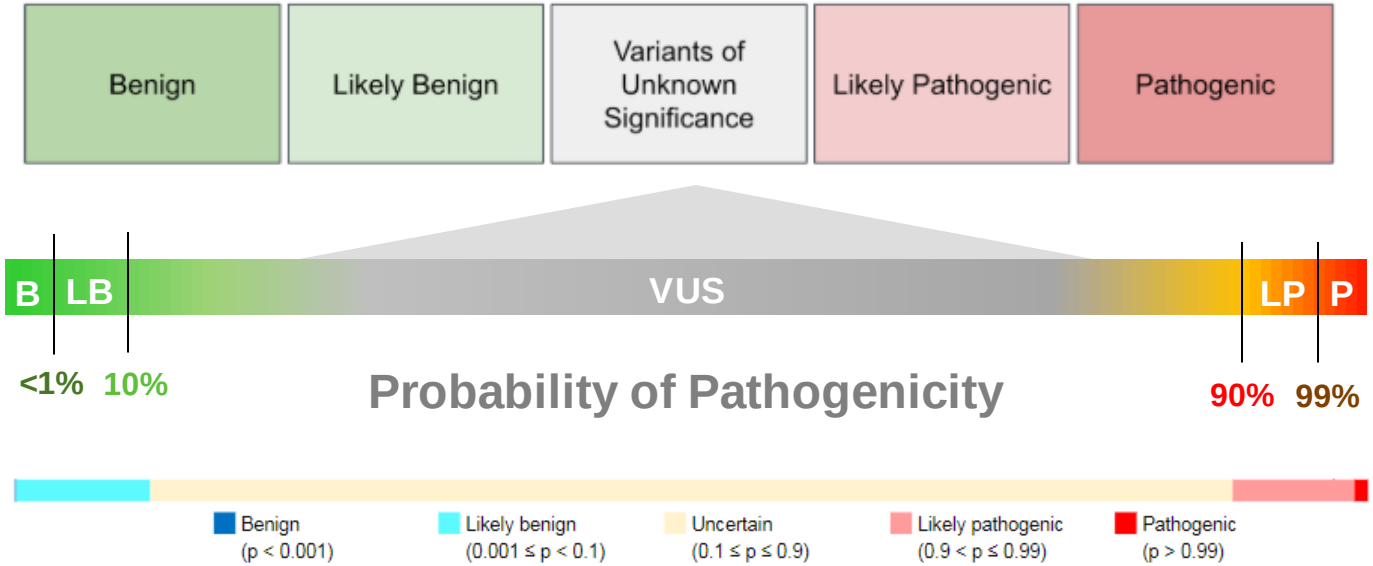
2015 ACMG Guidelines

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

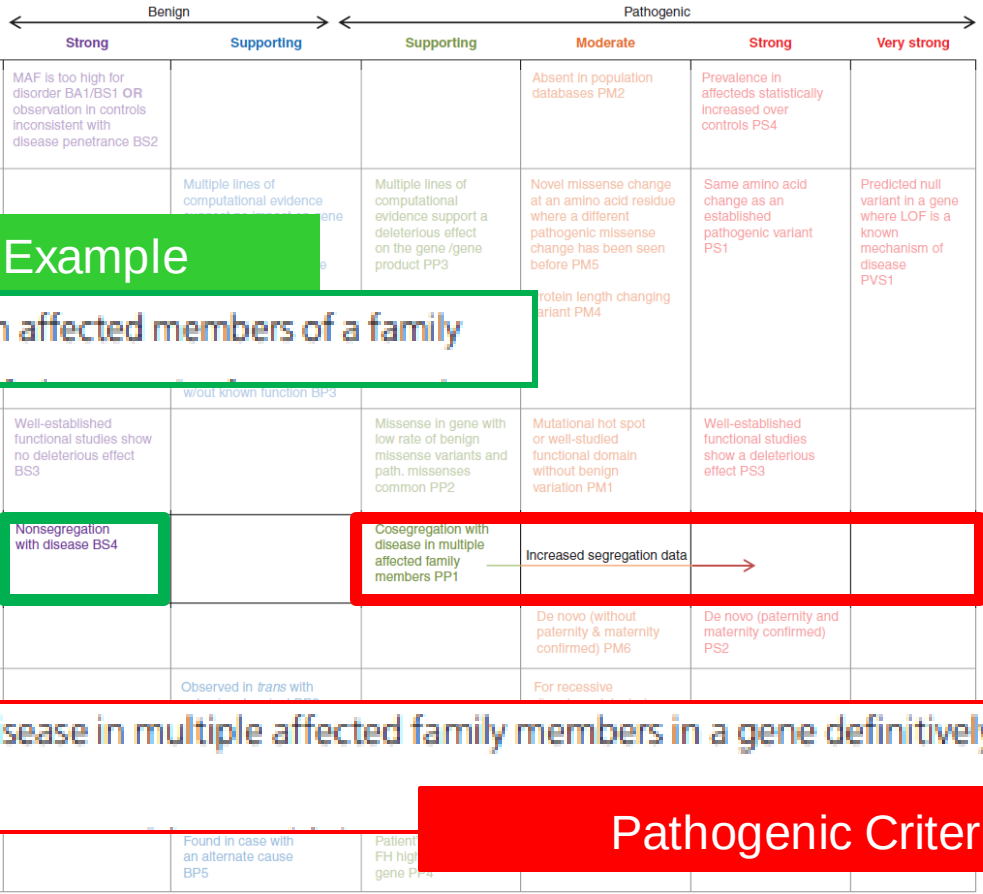
Sue Richards, PhD[1], Nazneen Aziz, PhD[2,16], Sherri Bale, PhD[3], David Bick, MD[4], Soma Das, PhD[5], Julie Gastier-Foster, PhD[6,7,8], Wayne W. Grody, MD, PhD[9,10,11], Madhuri Hegde, PhD[12], Elaine Lyon, PhD[13], Elaine Spector, PhD[14], Karl Voelkerding, MD[13] and Heidi L. Rehm, PhD[15]; on behalf of the ACMG Laboratory Quality Assurance Committee



The rules proposed to classify sequence variants follows is a heuristic system for variant classification that is compatible with a formal, quantitative, naive Bayesian classifier.



Tavtigian SV et al., Genet Med. 2018 Sep;20(9):1054-1060
Epub 2018 Jan 4. PMID: 29300386



Benign Criteria Example

BS4 Lack of segregation in affected members of a family

Segregation Data Criterion

PP1 Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease

Pathogenic Criteria Example

Criteria Nomenclature

- Each discrete criteria (benign “B” or pathogenic “P”) with its corresponding level of evidence is combined and together they yield a pathogenicity classification.

PS4 + PM2 + PP1 + BP4

Pathogenic

Level of Evidence: *Strong*

Benign

Level of Evidence: *Supporting*

Table 5 Rules for combining criteria to classify sequence variants

Pathogenic	(i) 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) (ii) ≥2 Strong (PS1–PS4) *OR* (iii) 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	(i) 1 Very strong (PVS1) *AND* 1 moderate (PM1–PM6) *OR* (ii) 1 Strong (PS1–PS4) *AND* 1–2 moderate (PM1–PM6) *OR* (iii) 1 Strong (PS1–PS4) *AND* ≥2 supporting (PP1–PP5) *OR* (iv) ≥3 Moderate (PM1–PM6) *OR* (v) 2 Moderate (PM1–PM6) *AND* ≥2 supporting (PP1–PP5) *OR* (vi) 1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Benign	(i) 1 Stand-alone (BA1) *OR* (ii) ≥2 Strong (BS1–BS4)
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–

Variant Interpretation Summary Example:
BRCA1 (NM_007294.3) c. 212G>C, p.(Arg71Thr)

SUMMARY

The heterozygous c.212G>C (p.R71T) variant was detected in the BRCA1 gene (NM_007294.3) and involves the last residue of exon 4 of 23.

This variant has been reported in a single affected individual with the associated disease (Harter et al., 2017; PMID: 29053726).

Functional testing has been performed for this variant and supports decreased protein function with a reduced expression of mRNA in transfected HAP1 cells. (Findlay et al, 2018, PMID 30209399).

Another amino acid substitution occurring in the same residue (p.Arg71Gly, p.Arg71Lys) has been determined to contribute to the disease associated with this gene.

The variant detected is absent in a large control population database without reported homozygotes (Karczewski et al., 2020, PMID: 32461654).

Multiple computational predictors suggest a damaging effect on gene or protein function.

*Therefore, c.212G>C (p.R71T) in the BRCA1 gene is classified as **Pathogenic.** Clinical correlation is recommended.*

CONCEPT	ACMG CRITERIAS CODE
Introduction	
PREVIOUSLY REPORTED CASES	PS4_Supporting
FUNCTIONAL TESTING	PS3
RESIDUE LEVEL ANNOTATION	PM5
ALLELE FREQUENCY	PM2
INSILICO PREDICTORS	PP3
Conclusion	

Variant Interpretation Summary Example:

BRCA1 (NM_007294.3) c. 212G>C, p.(Arg71Thr)

Table 5 Rules for combining criteria to classify sequence variants

Pathogenic	(i) 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) (ii) ≥2 Strong (PS1–PS4) *OR* (iii) 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	(i) 1 Very strong (PVS1) *AND* 1 moderate (PM1–PM6) *OR* (ii) 1 Strong (PS1–PS4) *AND* 1–2 moderate (PM1–PM6) *OR* (iii) 1 Strong (PS1–PS4) *AND* ≥2 supporting (PP1–PP5) *OR* (iv) ≥3 Moderate (PM1–PM6) *OR* (v) 2 Moderate (PM1–PM6) *AND* ≥2 supporting (PP1–PP5) *OR* (vi) 1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Benign	(i) 1 Stand-alone (BA1) *OR* (ii) ≥2 Strong (BS1–BS4)
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) *OR* (ii) ≥2 Supporting (BP1–BP7)
Uncertain significance	(i) Other criteria shown above are not met OR (ii) the criteria for benign and pathogenic are contradictory

ACMG CRITERIAS CODE

PS4_Supporting
PS3
PM5
PM2
PP3

Conflicting Evidence Example

- **PS4 + PM2 + BP2 + BP4**

= Variant of Uncertain Significance (VUS)

Likely pathogenic	(i) 1 Very strong (PVS1) *AND* 1 moderate (PM1–PM6) *OR*
	(ii) 1 Strong (PS1–PS4) *AND* 1–2 moderate (PM1–PM6) *OR*
	(iii) 1 Strong (PS1–PS4) *AND* ≥2 supporting (PP1–PP5) *OR*
	(iv) ≥3 Moderate (PM1–PM6) *OR*
	(v) 2 Moderate (PM1–PM6) *AND* ≥2 supporting (PP1–PP5) *OR*
	(vi) 1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Benign	(i) 1 Stand-alone (BA1) *OR*
	(ii) ≥2 Strong (BS1–BS4)
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) *OR*
	(ii) ≥2 Supporting (BP1–BP7)
Uncertain significance	(i) Other criteria shown above are not met OR
	(ii) the criteria for benign and pathogenic are contradictory

https://www.medschool.umaryland.edu/genetic_variant_interpretation_tool1.html/

Variant Interpretation Framework Summary

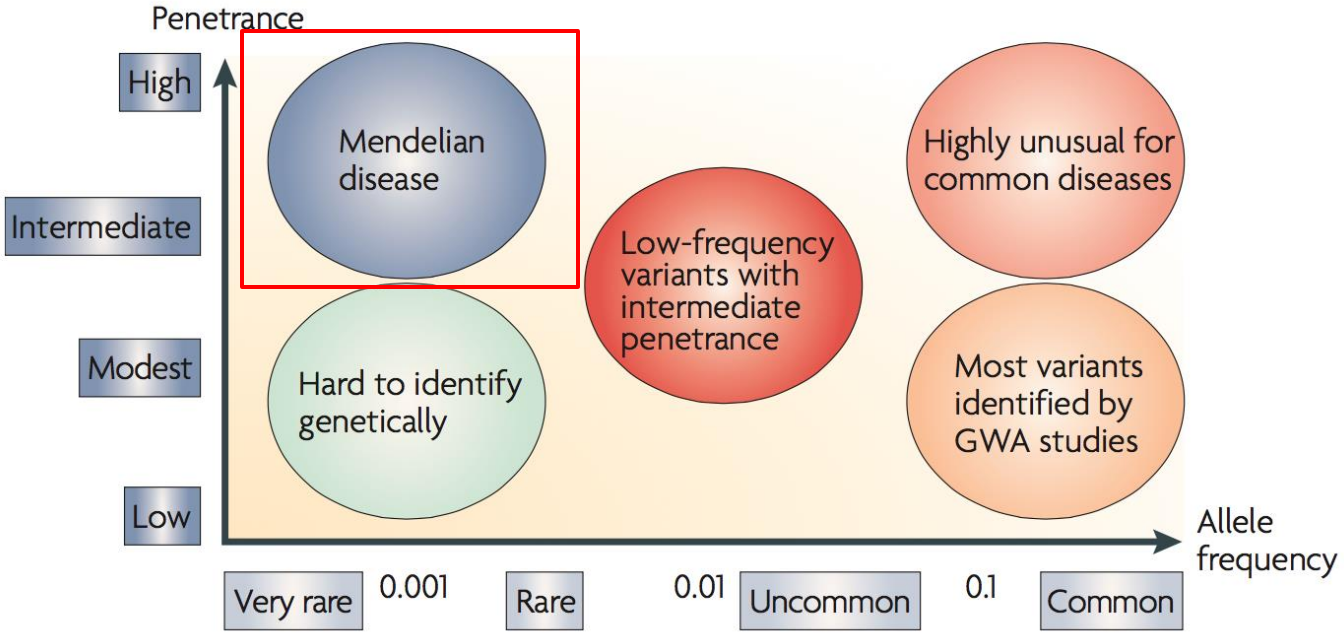
	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very strong
Population data	MAF is too high for disorder BA1/BS1 **OR** observation in controls inconsistent with disease penetrance BS2			Absent in population databases PM2	Prevalence in affecteds statistically increased over controls 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 In-frame indels in repeat w/out known function BP3	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	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and 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			

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2

Allele Frequency (BA1, BS1, PM2)
How common is this variant?



https://gnomad.broadinstitute.org/

Allele Frequency (BA1, BS1, PM2)
How common is this variant?

BA1:

>5% allele frequency in any general continental population of at least 2,000 alleles for a gene without a gene or variant specific recommendation.

BS1:

Allele frequency is greater than expected for disorder (lower than BA1)

PM2:

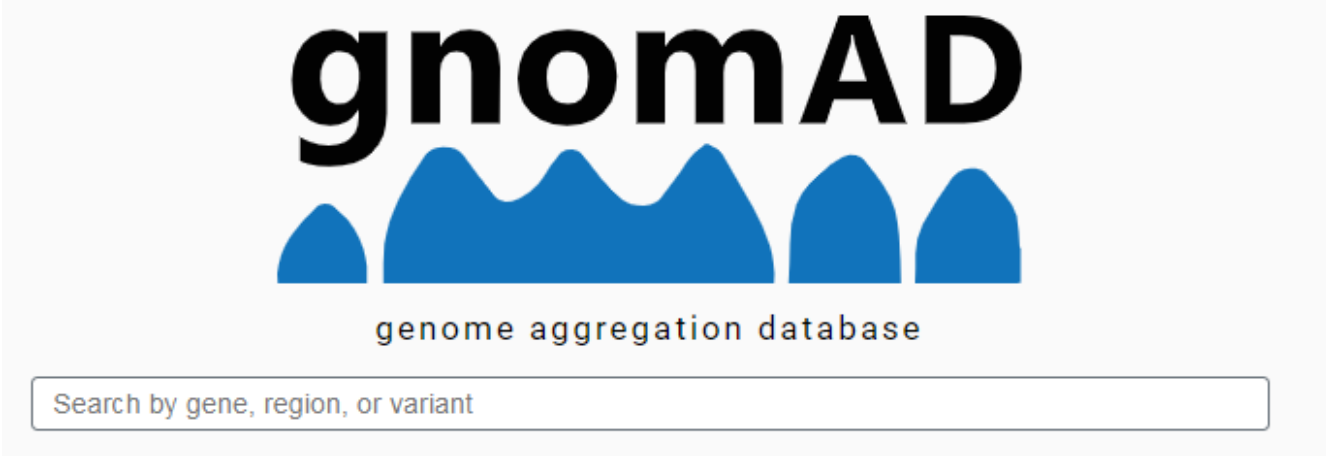
Absent from controls (or at an extremely low frequency if recessive).

https://gnomad.broadinstitute.org/

Exome Aggregation Consortium (ExAC)

gnomAD browser

125,748 exome sequences
15,708 whole-genome sequences
141,456 individuals



http://gnomad.broadinstitute.org/

gnomAD

Single nucleotide variant: 6-51944718-G-A (GRCh37)

	Exomes	Genomes	Total
Filter	Pass	Pass	
Allele Count	2	1	3
Allele Number	251434	31396	282830
Allele Frequency	0.000007954	0.00003185	0.00001061
Popmax Filtering AF (95% confidence)	–	–	
Number of homozygotes	0	0	0

References

- dbSNP (rs727504096)
- UCSC
- ClinVar (177240)

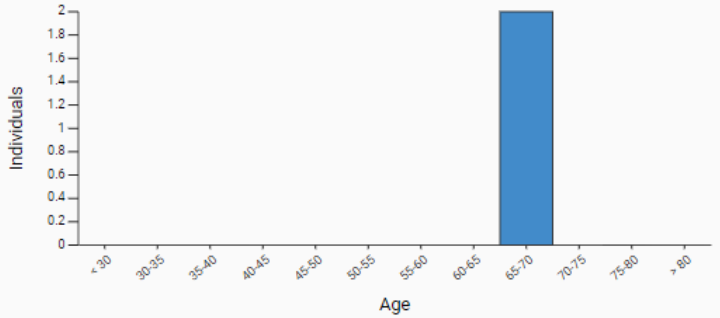
Report

- Report this variant
- Request additional information

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
African	2	24962	0	0.00008012
East Asian	1	19950	0	0.00005013
Latino	0	35438	0	0.000
Ashkenazi Jewish	0	10370	0	0.000
European (Finnish)	0	25122	0	0.000
European (non-Finnish)	0	129146	0	0.000
Other	0	7226	0	0.000
South Asian	0	30616	0	0.000
Male	2	153358	0	0.00001304
Female	1	129472	0	0.000007724
Total	**3**	**282830**	**0**	**0.00001061**

Include: Exomes Genomes

Age Distribution



Heterozygous Variant Carriers

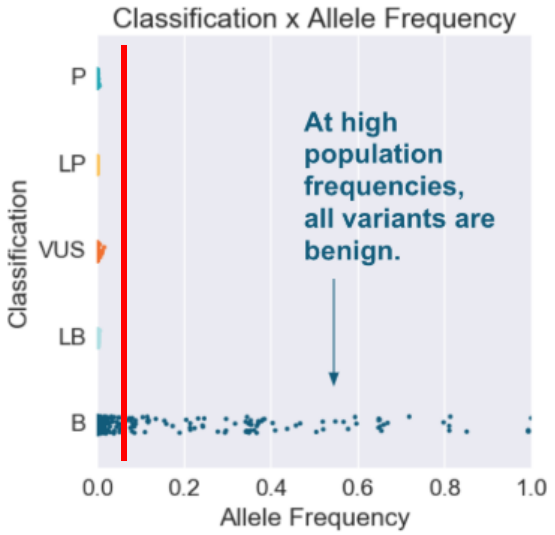
Exomes Genomes

http://gnomad.broadinstitute.org/

Allele Frequency (BA1, BS1, PM2)
How common is this variant?

BA1:

>5% allele frequency in any general continental population of at least 2,000 alleles for a gene without a gene or variant specific recommendation.



https://gnomad.broadinstitute.org/variant/19-11210912-C-T?dataset=gnomad_r2_1

Population Frequencies

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
European (non-Finnish)	14963	129102	879	0.1159
European (Finnish)	2629	24508	157	0.1073
Other	695	7222	35	0.09623
Ashkenazi Jewish	981	10358	45	0.09471
South Asian	2189	30612	94	0.07151
Latino/Admixed American	2317	35436	89	0.06539
African/African-American	950	24920	20	0.03812
East Asian	194	19948	4	0.009725
XX	11122	129108	623	0.08614
XY	13796	152998	700	0.09017
Total	**24918**	**282106**	**1323**	**0.08833**

Allele Frequency (BA1, BS1, PM2)
BA1 Exceptions

BA1:

>5% allele frequency in any general continental population of at least 2,000 alleles for a gene without a gene or variant specific recommendation.

Gene	Variant	Classification	Criteria	ExAC Pop	ExAC Pop MAF	Disease
HFE	NM_000410.3: c.187C>G (p.His63Asp)	P	PS4	NFE	**1.3%**	Hereditary hemochromatosis
HFE	NM_000410.3: c.845G>A (p.Cys282Tyr)	P	PS4, PP3	NFE	**5.1%**	Hereditary hemochromatosis
BTD	NM_000060.4: c.1330G>C (p.Asp444His)	P	PS3, PM3_Strong, PP3, PP4	FIN	**5.3%**	Biotidinase deficiency
GJB2	NM_004004.5: c.109G>A (p.Val37Ile)	P	PS4; PP1_Strong; PM3_VeryStrong; PS3_Moderate	EAS	**7.2%**	Deafness, autosomal recessive

https://www.clinicalgenome.org/site/assets/files/3460/ba1_exception_list_07_30_2018.pdf

BA1

BS1

PM2

http://cardiodb.org/allelefrequencyapp/

Table 2 Comparison of Population Frequency Thresholds from ClinGen Variant Curation Expert Panels[a]

	Criteria	Prevalence	Heterogeneity	Penetrance	Threshold
Cardiomyopathy (AD)	BA1	1:200	10.60%[L]	30%	≥0.001 (0.1%)
	BS1	1:200	2%[A]	30%	≥0.0002 (0.02%)
	PM2	1:500	2%[A]	50%	<0.00004 (0.004%)
RASopathy (AD)	BA1	1:2500	100%	40%	≥0.0005 (0.05%)
	BS1	1:2500	50%[L]	40%	≥0.00025 (0.025%)
	PM2	–	–	–	Absent[R]
CDH1 (AD)	BA1	1:800	100%	30%	≥0.002 (0.2%)
	BS1	1:1250	100%	30%	≥0.001 (0.1%)
	PM2	–	–	–	<0.00001 (0.001%)[R]
Hearing loss (AD)	BA1	1:30	5%[L/A]	80%	≥0.001 (0.1%)
	BS1	1:150	5%[L/A]	80%	≥0.0002 (0.02%)
	PM2	–	–	—	<0.00002 (0.002%)[M]
Hearing loss (AR)	BA1	1:200	7.2%[A]	100%	≥0.005 (0.5%)
	BS1	1:200	4.4%[A]	100%	≥0.003 (0.3%)
	PM2	–	–	–	<0.00007 (0.007%)[M]
PAH (AR)	BA1	1:5000	90%[L]	80%	≥0.015 (1.5%)
	BS1	1:5000	2%[A]	80%	≥0.002 (0.2%)
	PM2	–	–	–	<0.0002 (0.02%)[M]
PTEN[b] (AD)	BA1	–	–	–	≥0.01 (1%)
	BS1	–	–	–	≥0.001 (0.1%)
	PM2	–	–	–	<0.00001 (0.001%)[R]

Allele Frequency (BA1, BS1, PM2)
Other considerations…

What is a control population? Unselected?

For dominant disorders (AD):
Adult-onset disorders could be represented in the gnomAD database in still unaffected probands. Instead of controls the database could be refer to better as "general population".

PM2:
Absent from controls (or at an extremely low frequency if recessive).

PM2_supporting

Demotion of this category to supporting is currently recommended.

https://clinicalgenome.org/site/assets/files/5182/pm2_-_svi_recommendation_-_approved_sept2020.pdf

https://gnomad.broadinstitute.org/

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

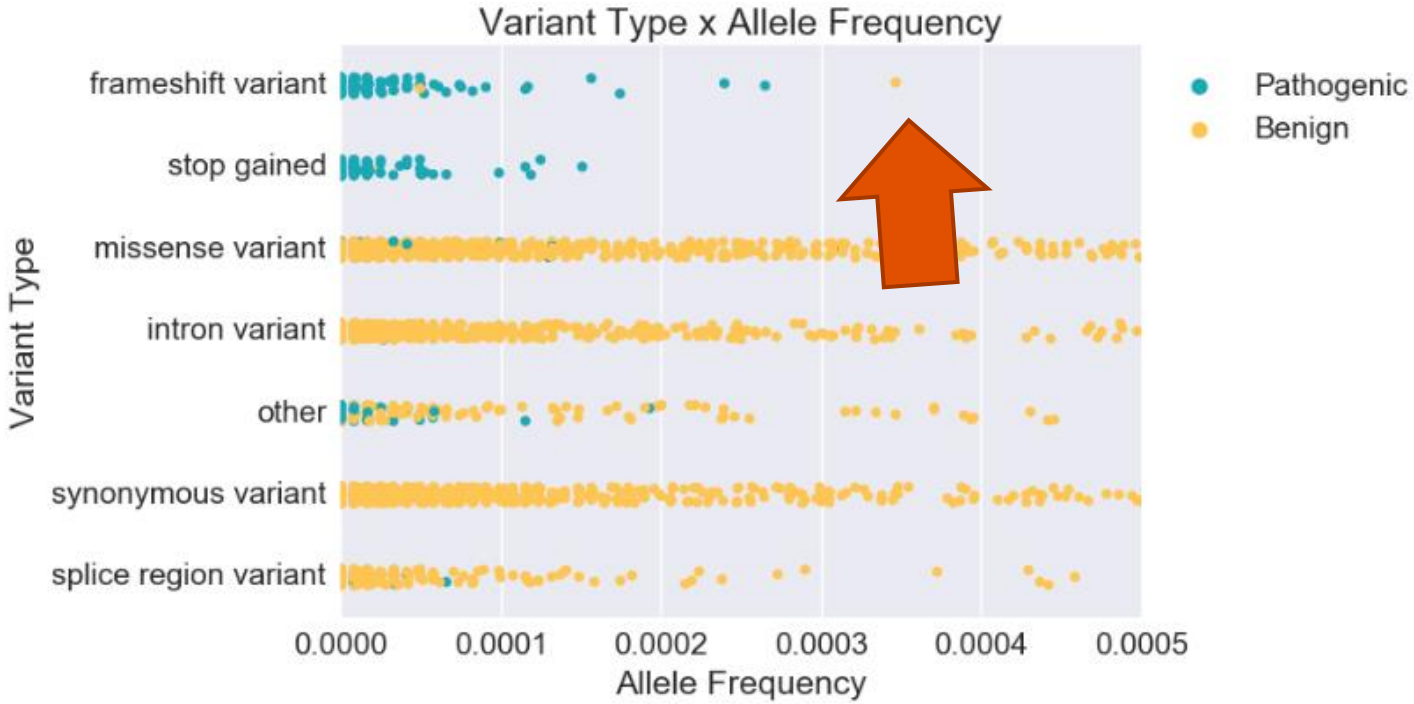
Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1

PVS1

Null variant in a gene where loss of function (LoF) is a known mechanism of disease.

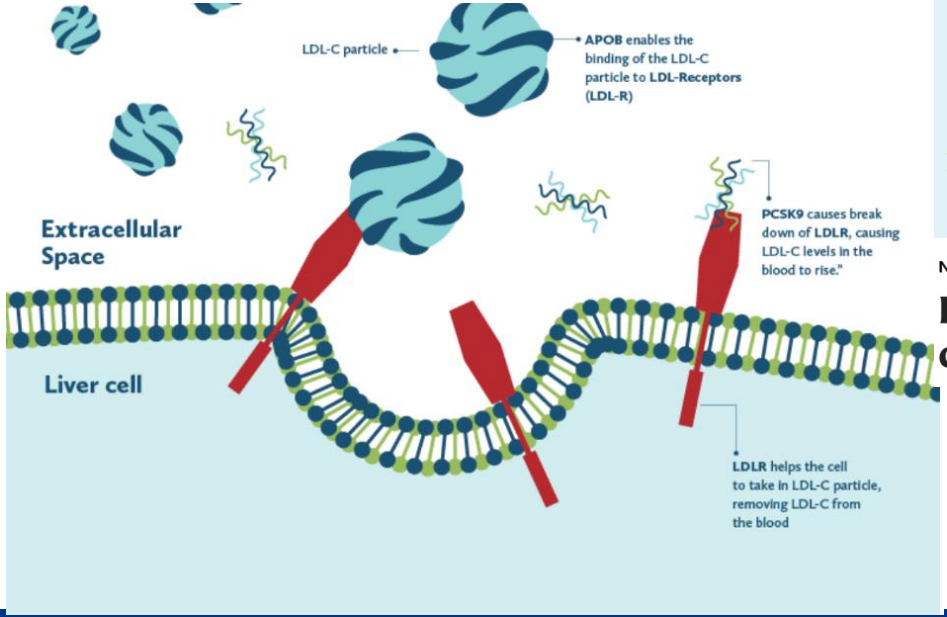
Null Variant Type	Example
Frameshift:	c.2019delA, p.Glu673Aspfs*28
Canonical splice sites:	c.234+2G>A p.?
Nonsense / “stop gain”:	c.1501 A > T, p.K501Ter
Single, multi-exon or gene deletions	

Variant type is a stronger predictor of pathogenicity than population frequency



Loss of Function Criteria (PVS1)
(*PCSK9* – Familial Hypercholesterolemia Example)

Location	Phenotype	Phenotype MIM number	Inheritance
1p32.3	{Low density lipoprotein cholesterol level QTL 1}	603776	AD
	Hypercholesterolemia, familial, 3	603776	AD



NM_174936.4(PCSK9):c.142G>T (p.Glu48Ter)

Interpretation:	**Likely benign**
Review status:	★☆☆☆ criteria provided, single submitter
Submissions:	1 (Most recent: May 19, 2020)
Last evaluated:	Sep 30, 2019
Accession:	VCV000927517.1
Variation ID:	927517
Description:	single nucleotide variant

NEWS IN BRIEF · 06 SEPTEMBER 2019

PCSK9-lowering RNAi contender clears first phase III trial

PVS1

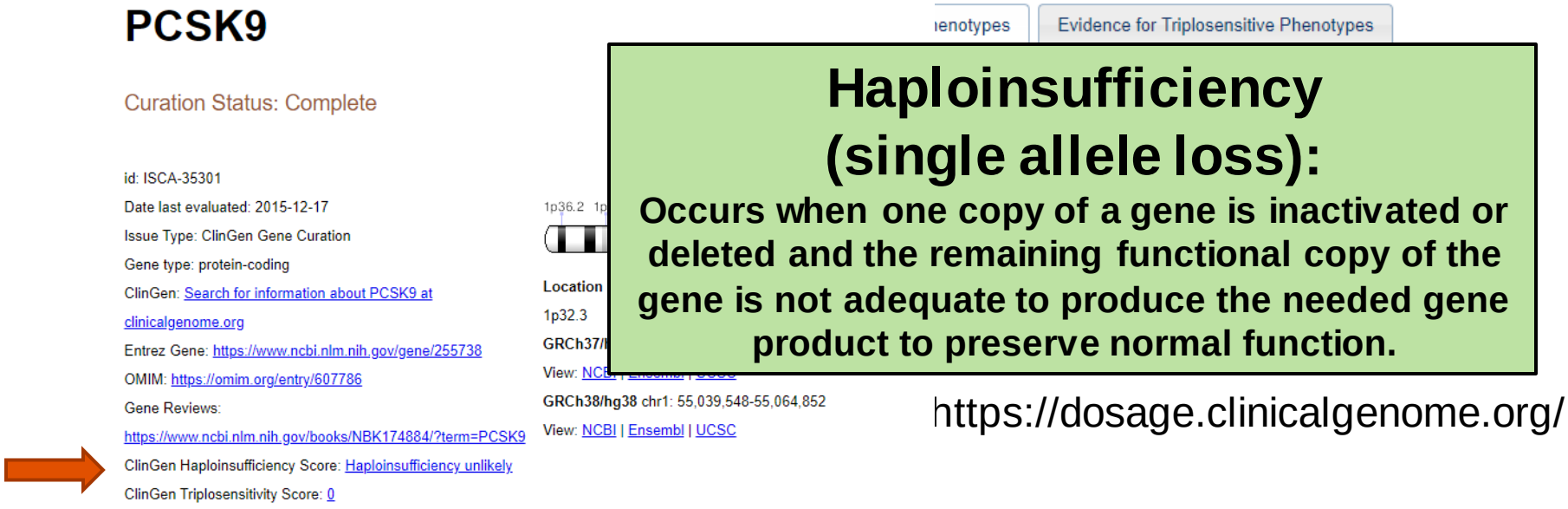
Null variant in a gene where loss of function (LoF) is a known mechanism of disease.

Null Variant Type	Example
Frameshift:	c.2019delA, p.Glu673Aspfs*28
Canonical splice sites:	c.234+2G>A p.?
Nonsense / “stop gain”:	c.1501 A > T, p.K501Ter
Single, multi-exon or gene deletions	

How do I know if loss of function variants cause disease for a specific gene?

How do I know if loss of function variants cause disease?

- (1) ClinGen Dosage Sensitivity Curation Page



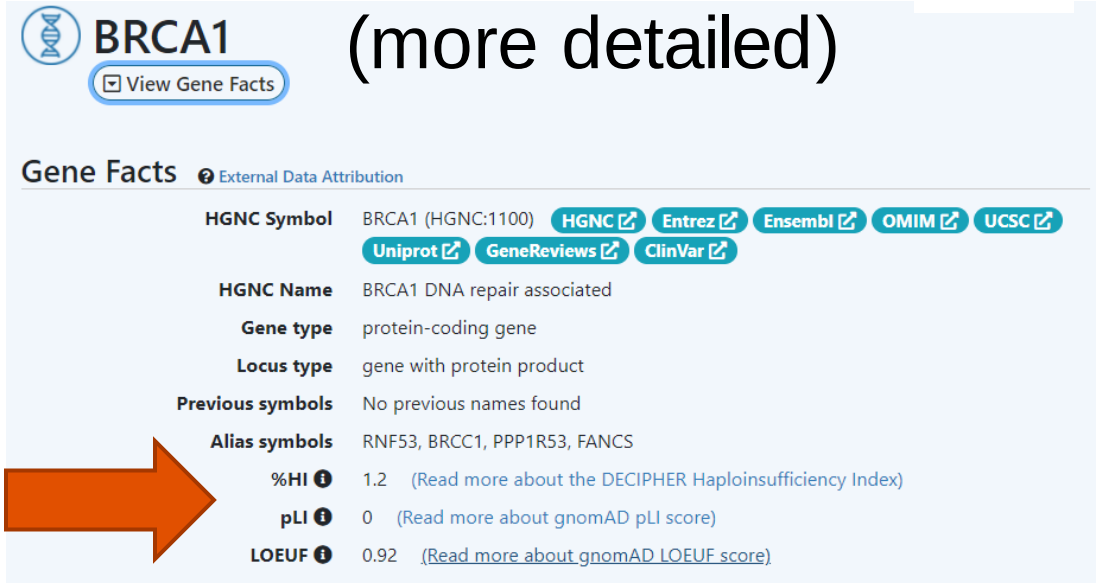
Haploinsufficiency (single allele loss):

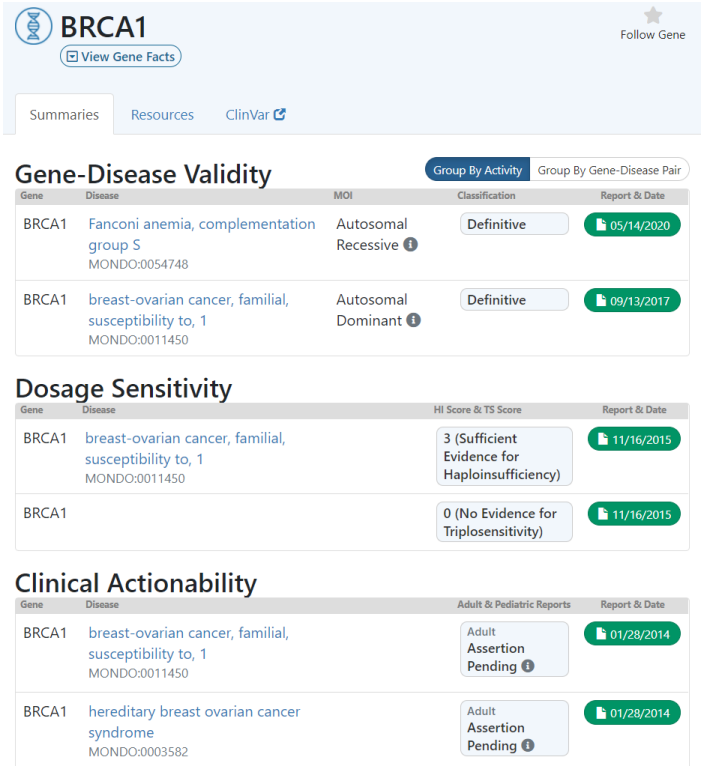
Occurs when one copy of a gene is inactivated or deleted and the remaining functional copy of the gene is not adequate to produce the needed gene product to preserve normal function.

https://dosage.clinicalgenome.org/

How do I know if loss of function variants cause disease?

- (2) ClinGen Search (more detailed)

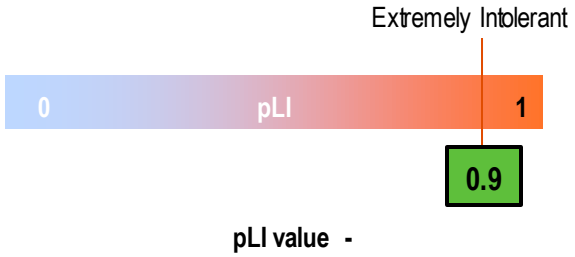




https://search.clinicalgenome.org/kb/genes/HGNC:1100

How do I know if loss of function variants cause disease?

- (3) pLI score - Probability a gene is haploinsufficient - where heterozygous LoFs are not tolerated. >0.9 is a common threshold. Particularly good for autosomal dominant disease.



PTPN11 protein tyrosine phosphatase non-receptor... Dataset gnomAD v2.1.1 gnomAD SVs v2.1

Genome build GRCh37 / hg19
Ensembl gene ID ENSG00000179295.11
Ensembl canonical transcript ENST00000351677.2
Other transcripts
ENST00000392597.1, ENST00000530818.1, ENST00000531326.1
Region 12:112856155-112947717
External resources Ensembl, UCSC Browser, and more

Constraint

Category	Expected SNVs	Observed SNVs	Constraint metrics
Synonymous	123.7	112	Z = 0.82 o/e = 0.91 (0.78 - 1.06)
Missense	331.3	171	Z = 3.13 o/e = 0.52 (0.46 - 0.59)
pLoF	35.2	1	pLI = 1 o/e = 0.03 (0.01 - 0.14)

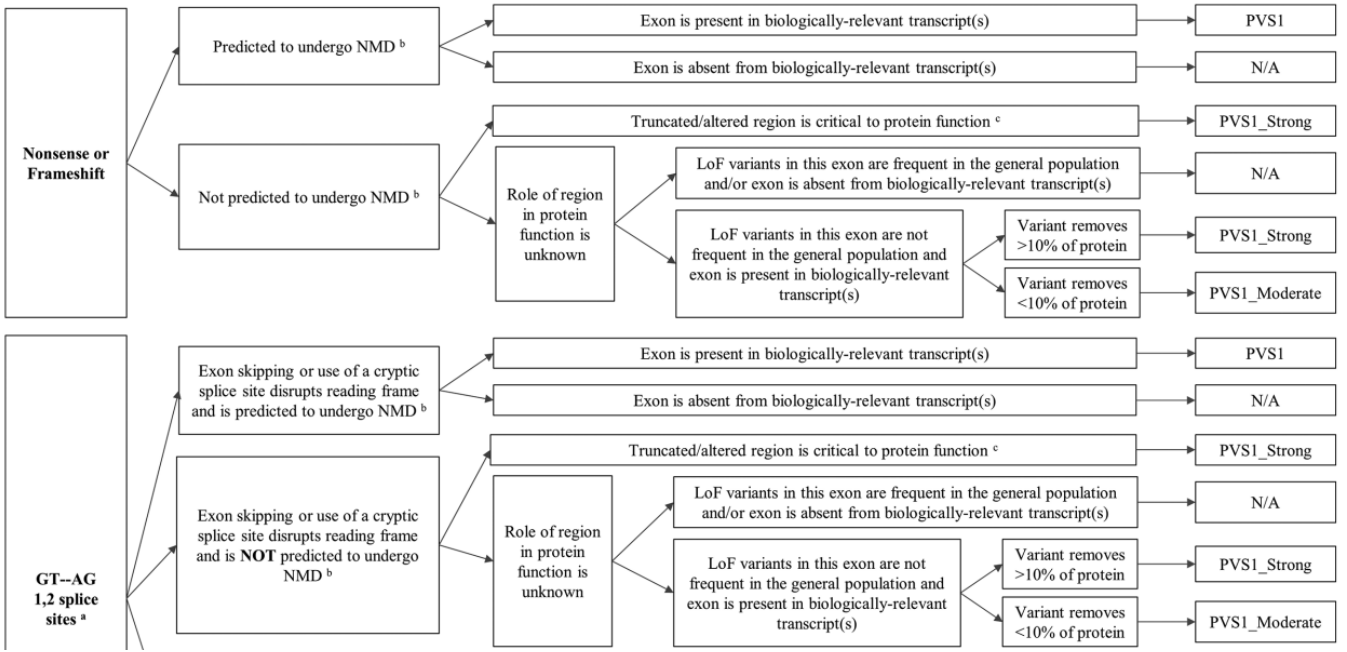
Constraint metrics based on Ensembl canonical transcript (ENST00000351677.2).

https://gnomad.broadinstitute.org/gene/ENSG00000179295?dataset=gnomad_r2_1

Null variants impact is position-dependent

PVS1

Null variant in a gene where loss of function (LoF) is a known mechanism of disease.



NMD: nonsense mediated decay.

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3

In-frame removal or insertion of amino acids

PM4

Protein length changes as a result of in-frame deletions/insertions

NM_000179.3(MSH6):
c.535_546del
p.(Ala179_Ala182del)

- Insertions/deletions that occur in repetitive regions are more likely to be of little functional impact; therefore, it is important to assess the surrounding sequence for repetitiveness using a genome browser.
- It can also help to assess population databases, such gnomAD, for high confidence variant calls that indicate the site is multi-allelic, which could indicate that the region is prone to indels that are generally tolerated, depending on the overall allele frequency.
- It is important to verify the functional impact the deletion or insertion might have. Does it affect the zinc-fingers of a transcription factor? Does it remove important aminoacids in the catalytic site?

Stop loss: Protein extending variants

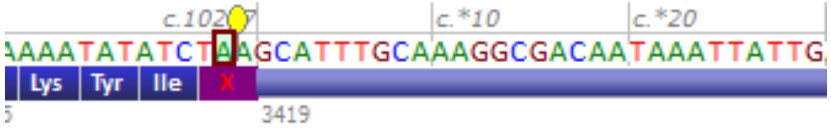
PM4

Protein length changes as a result of in-frame deletions/insertions

PM4

Stop-loss variant

- When a variant results in loss of the termination codon (stop-loss variant), the protein is extended; if a variant creates a premature termination codon (nonsense variant), the protein is shortened.



NM_000059.3(BRCA2):
c.10256_10257insT
p.(*3419Tyrext*18)

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions?		PP3, BP4
	Potential splicing impact?		BP7

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions?		PP3, BP4
	Potential splicing impact?		BP7

For example:
BRCA1
(NM_007294.3)
c.736T>G
p.Leu246Val

- In-silico predictions are primarily useful in amino acid substitutions or **missense variants.**
- Informative also for splicing/intronic variation.
- Do not generalize, missense variant can also have splicing effects, particularly at the start or end of an exon.

Computational Impact Prediction

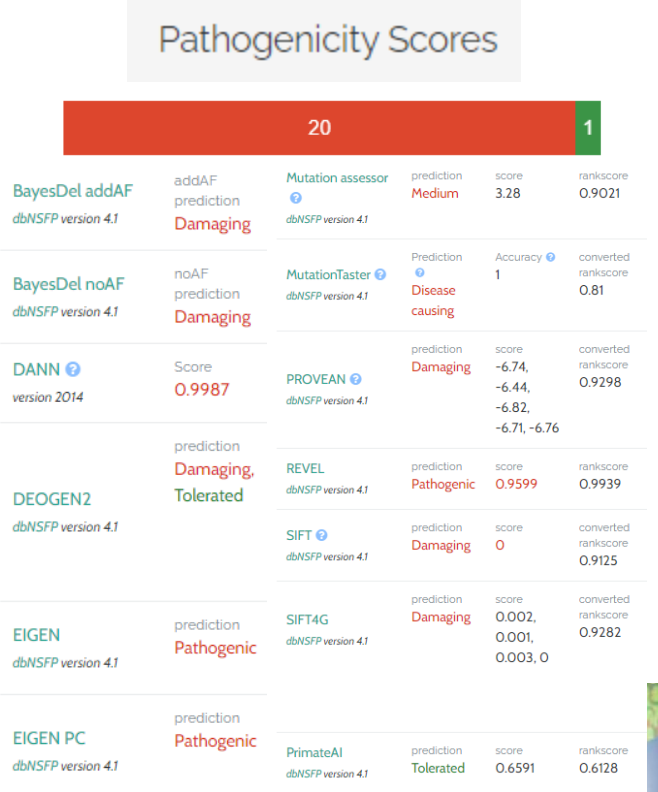
"In silico scores"

PP3

Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.)

BP4

Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.)



Computational Impact Prediction

Considerations

PP3

BP4

- PP3 or BP4 can be used only once in a variant. Many algorithms used the same or very similar training data for their predictions, each algorithm cannot be counted as an independent criterion.
- Consistent threshold for the tool(s) should be used for all the variants in that gene.
- Currently, a **meta-predictor such as REVEL** may be used in place of multiple predictors in the in silico analysis of missense variants.
- Splicing in silico tools can be difficult to utilize and the interpretation is often not standardized.

Richards CS et al. *Gen Med.* 2015;17:405-423

Computational Impact Prediction

Splicing Scores

PP3

BP4

BP7

A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved.

Computational Impact Prediction

A Commonly Used Powerful Splicing Tool

PP3

BP4

BP7

SpliceAI

Deep neural network based on pre-mRNA transcript sequences that predicts splice sites using long-range primary genomic sequence flanking each position as input
(+/-50 bp as default; +/-10,000 bp maximum).

https://spliceailookup.broadinstitute.org/

Δ type	Δ score ?	pre-mRNA position ?
Acceptor Loss	0.00	
Donor Loss	0.72	0 bp
Acceptor Gain	0.00	
Donor Gain	0.01	-47 bp

SpliceAI provides a table with delta scores (0-1) for acceptor loss, donor loss, acceptor gain, and donor gain within the designated flanking sequence. The delta score indicates the probability that the variant will alter splicing at the pre-mRNA position indicated

Scores are not deterministic of biological effect/deleteriousness, they are used as “supporting evidence”

gDNA: Chr6(GRCh37):g.51720765A>G
cDNA: NM_138694.3(PKHD1):c.7837T>C
Protein: p.Trp2613Arg

Polyphen-2: Probably damaging
CADD: 29
M-CAP: Probably
PredictSNP2: Deleterious

Scores agree towards SNV being deleterious

Likelihood of pathogenicity is affected, not determined.

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions?		PP3, BP4
	Potential splicing impact?		BP7

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?		PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level		PP2, BP1 (PM1)

Missense Intolerance / Constraint Metrics

PP2

Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease. (Uses the “missense z-score”)

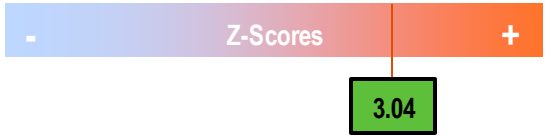
BP4

Missense variant in a gene for which primarily truncating variants are known to cause disease.

PP2

gnomAD Constraint information

Increased constraint!
(intolerance to variation)



Deviation of observed counts from the expected number of variants for a specific gene

PTPN11 protein tyrosine phosphatase non-receptor... Dataset gnomAD v2.1.1 gnomAD SVs v2.1

Genome build GRCh37 / hg19
Ensembl gene ID ENSG00000179295.11
Ensembl canonical transcript ENST00000351677.2
Other transcripts
ENST00000392597.1, ENST00000530818.1, ENST00000531326.1
Region 12:112856155-112947717
External resources Ensembl, UCSC Browser, and more

Constraint

Category	Expected SNVs	Observed SNVs	Constraint metrics
Synonymous	123.7	112	Z = 0.82 o/e = 0.91 (0.78 - 1.06)
Missense	331.3	171	Z = 3.13 o/e = 0.52 (0.46 - 0.59)
pLoF	35.2	1	pLI = 1 o/e = 0.03 (0.01 - 0.14)

Constraint metrics based on Ensembl canonical transcript (ENST00000351677.2).

This is not absolute –
Interpret with caution!
Among the factors that modify this are variable expressivity, penetrance and some specific regions of the gene may be more intolerant than others for missense.

https://www.nature.com/articles/nature19057

Impact Prediction: Computational or Knowledge-based

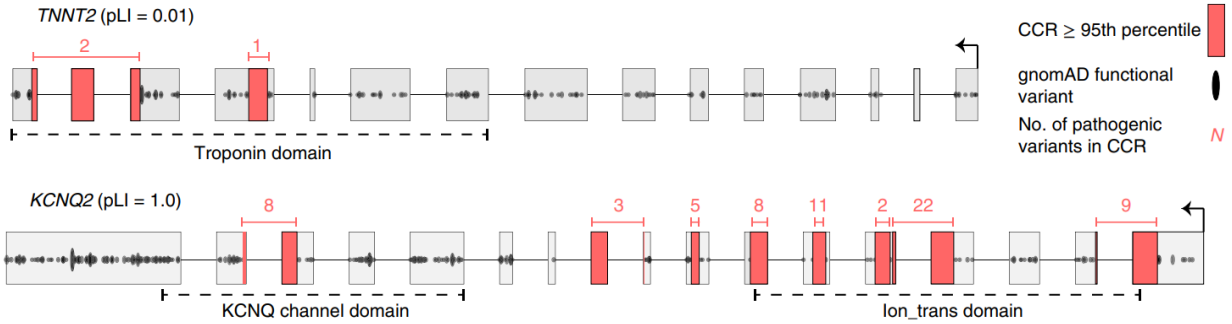
- Gene-wide summary measures of constraint are prone to overstating and understating constraint within specific regions of protein-coding genes

A map of constrained coding regions in the human genome

James M. Havrilla[1,2], Brent S. Pedersen[1,2], Ryan M. Layer[3,4] and Aaron R. Quinlan[1,2,5*]

nature genetics

Regional intolerance correlates with important functional domains



Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?		PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level		PP2, BP1

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?		PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level		PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?		PM1

Mutational Hotspot or Critical Functional Domain

PM1

Located in a mutational hot spot and/or critical and well established functional domain (ex. Active site of an enzyme) without benign variation

- What/who determines the size or location of a mutational hotspot?
- Where can I find an established functional domain for a particular gene/protein product?

Entry	Entry name	Protein names	Gene names	Organism	Length
P43246	MSH2_HUMAN	**DNA mismatch repair protein Msh2**	**MSH2**	Homo sapiens (Human)	934

PTM / Processing

Molecule processing

Feature key	Position(s)	Description	Actions	Graphical view	Length
Initiator methionine		Removed — Combined sources			
Chain (PRO_0000115183)	2 – 934	DNA mismatch repair protein Msh2	Add BLAST		933

Amino acid modifications

Feature key	Position(s)	Description	Actions	Graphical view	Length
Modified residue	2	N-acetylalanine — Combined sources			1
Cross-link	430	Glycyl lysine isopeptide (Lys-Gly) (interchain with G-Cter in SUMO2) — Combined sources			
Modified residue	555	N6-acetyllysine — Combined sources			1
Modified residue	567	N6-acetyllysine — By similarity			1
Modified residue	921	Phosphoserine — Combined sources			1

9 Publications

Mutational Hotspot or Critical Functional Domain

PM1

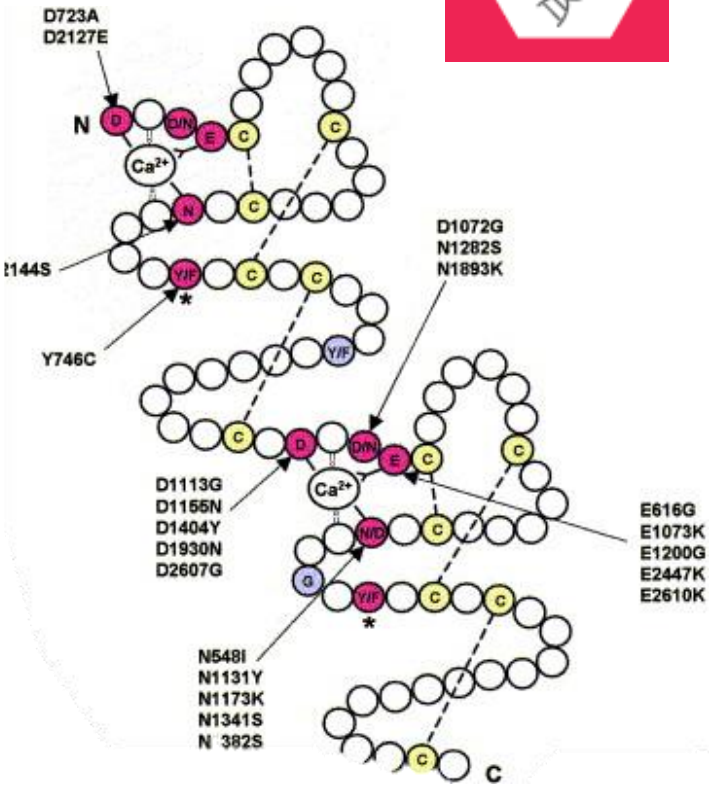
This is a challenging subjective criteria subject of debate in the field and clinical interpretation laboratories. However there are efforts to specify per gene which regions or residues are amenable to the use of this criteria.

Circulation: Genomic and Precision Medicine
Volume 11, Issue 6, June 2018
https://doi.org/10.1161/CIRCGEN.117.002039

ORIGINAL ARTICLE

Tailoring the American College of Medical Genetics and Genomics and the Association for Molecular Pathology Guidelines for the Interpretation of Sequenced Variants in the *FBN1* Gene for Marfan Syndrome

Proposal for a Disease- and Gene-Specific Guideline



UNIPROT

PP2 vs PM1 Primary Evidence
The genic vs. regional level constraints

- One potential area of concern is applying PP2 for all missense variants in a gene when the constraint observed is confined to a specific region of the gene.
- MYH7 has a missense constraint Z score of 6.54 and 3.93 in ExAC and gnomAD, respectively, suggesting significant depletion of missense variation (or missense constraint).
- Recent guidelines specific to this gene specified PP2 as not applicable for this gene by the as recent studies suggest that this high constraint is actually driven by a statistically significant clustering of pathogenic variants in the head region of the protein (Homburger et al., 2016; Walsh et al., 2017).
- Cardiomyopathy VCEP concluded this evidence was most appropriately weighted as Moderate (PM1) through application of the critical domain rule (Kelly et al., 2018). Specification of PM1 has required gene-by-gene curation by disease experts, combined with laboratory data to determine regions intolerant to variation.

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

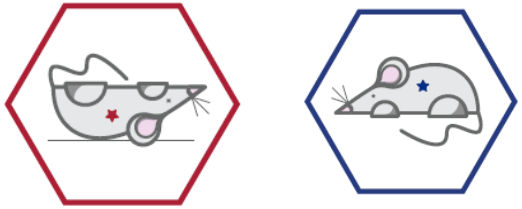
Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?		PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level		PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?		PM1

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?		PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level		PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?		PM1
	(6) Variant effect functionally studied?		PS3, BS3

Functional Evidence:



PS3

Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product.

What defines a "well established" functional study or assay?
How reliable? This is not simple.

BS3

Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing

Functional Evidence:

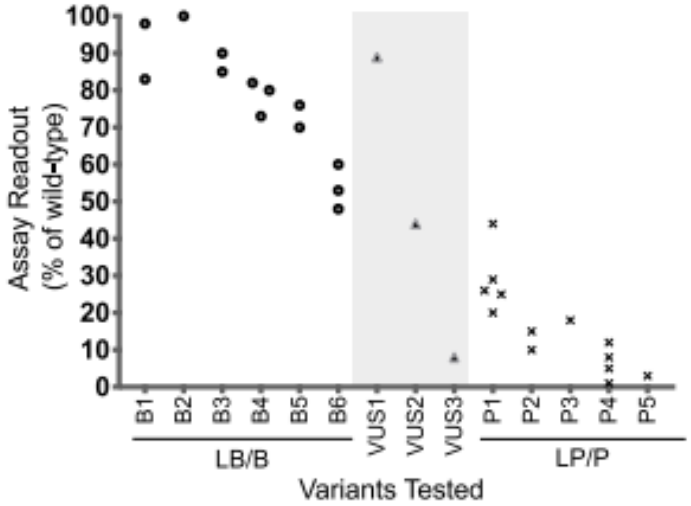
GUIDELINE Open Access

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

Sarah E. Brnich[1], Ahmad N. Abou Tayoun[2], Fergus J. Couch[3], Garry R. Cutting[4], Marc S. Greenblatt[5], Christopher D. Heinen[6], Dona M. Kanavy[1], Xi Luo[7], Shannon M. McNulty[1], Lea M. Starita[8,9], Sean V. Tavtigian[10], Matt W. Wright[11], Steven M. Harrison[12], Leslie G. Biesecker[13], Jonathan S. Berg[1*] and On behalf of the Clinical Genome Resource Sequence Variant Interpretation Working Group

Table 3 Evidence strength equivalent of odds of pathogenicity

Odds of pathogenicity (OddsPath)	Evidence strength equivalent
< 0.053	BS3
< 0.23	BS3_moderate*
< 0.48	BS3_supporting
0.48–2.1	Indeterminate
> 2.1	PS3_supporting
> 4.3	PS3_moderate
> 18.7	PS3
> 350	PS3_very_strong



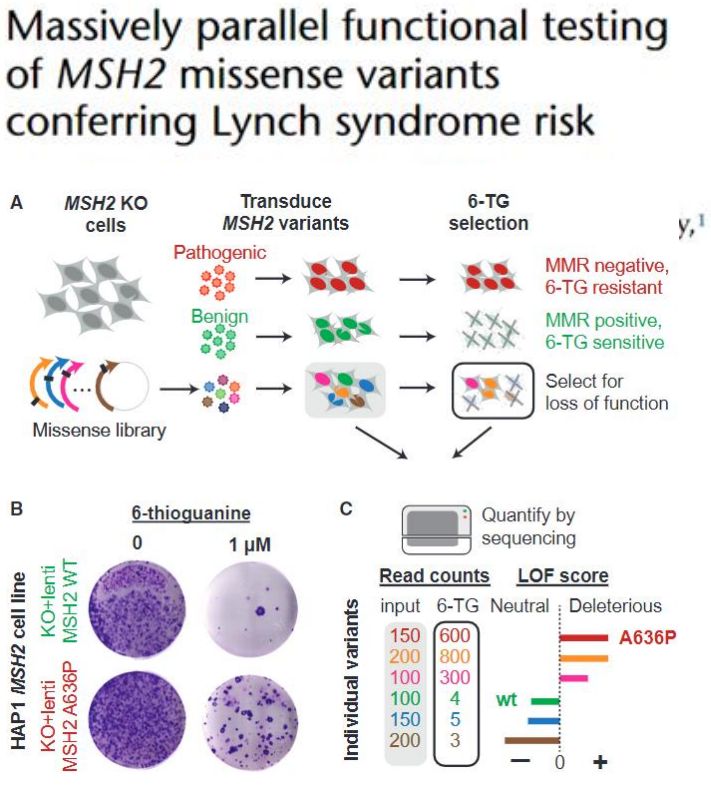
- Most functional evidence under these recommendations is demoted to PS3_supporting and in order to increase to moderate or strong, need to consider appropriate level of controls.
- Always consider if a test or assay is measuring the protein function or one of many.

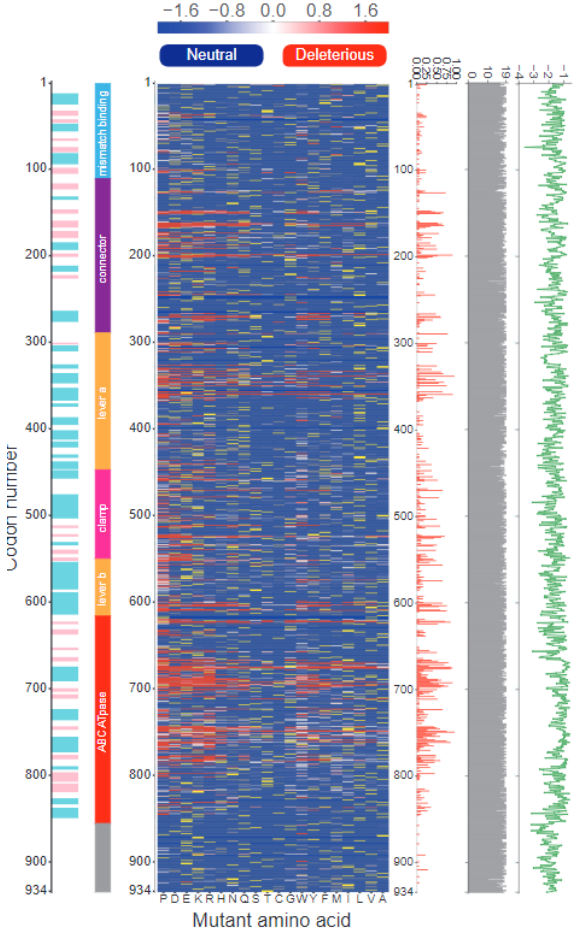
Functional Evidence:

Categories	Evidence Type
Protein Function	Biochemical function Protein interaction Expression
Models	Non-human organism Cell culture model
Cell-based assays	Patient cells Non-patient cells Non-eukaryote? Yeast? Rescue experiments

Functional testing has been performed for this variant and supports decreased protein function with a reduced expression of mRNA in transfected HAP1 cells (Findlay et al, 2018, PMID 30209399).

Functional Evidence: The Future





Variant Interpretation Framework Summary

(10 questions always ask to a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?		PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level		PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?		PM1
	(6) Variant effect functionally studied?		PS3, BS3

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type: Loss of function	PVS1
	(2) Variant Impact/Type: In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1

Knowledge Databases \ Previous Interpretations – ClinVar, HGMD

Variant of Interest: NM_000249.3(MLH1):c.1038G>T (p.**Gln346**His) (p.**Q346**H)

~~**PP5**~~

Reputable source recently **reports variant as pathogenic, but the evidence is not available** to the laboratory to perform an independent evaluation

NM_000249.3(MLH1):c.1038G>T (p.Gln346His)

Interpretation:	**Pathogenic**
Review status:	★★★☆ reviewed by expert panel
Submissions:	1 (Most recent: Dec 18, 2013)
Last evaluated:	Sep 5, 2013
Accession:	VCV000089618.1
Variation ID:	89618
Description:	single nucleotide variant

Knowledge Databases \ Previous Interpretations – ClinVar, HGMD

Variant of Interest: NM_000249.3(MLH1):c.1038G>T (p.**Gln346**His) (p.**Q346H**)

PM5

Missense change at an amino acid residue where a different missense change previously established as pathogenic

c.1037A>G (p.**Gln346**Arg) (p.**Q346**R)

PS1

Same amino acid change as a previously established pathogenic variant regardless of nucleotide change

c.1038G>(p.**Gln346**His) (p.**Q346H**)

Knowledge Databases \ Previous Interpretations – ClinVar, HGMD

Variant of Interest: NM_000249.3(MLH1):c.1038G>T (p.**Gln346**His) (p.**Q346**H)

https://www.ncbi.nlm.nih.gov/clinvar/?term=Q346+%5Bvariant+name%5D+and+MLH1

	Variation *Location*	Gene(s)	Protein change	Condition(s)	Clinical significance (Last reviewed)	Review status
	NM_000249.3(**MLH1**):c.1037A>G (p.Gln346Arg) *GRCh37:* Chr3:37061953 *GRCh38:* Chr3:37020462	MLH1	Q346R, Q248R, Q313R, Q105R, Q5R	Lynch syndrome	Pathogenic (Oct 18, 2018)	reviewed by expert panel
2.	NM_000249.3(**MLH1**):c.1038G>T (p.Gln346His) *GRCh37:* Chr3:37061954 *GRCh38:* Chr3:37020463	MLH1	Q346H, Q1 Q248H, Q3			expert
	NM_000249.3(**MLH1**):c.1038G>C (p.Gln346His) *GRCh37:* Chr3:37061954 *GRCh38:* Chr3:37020463	MLH1	Q346H, Q1 Q248H, Q3			expert

PM5

PS1

PS1 and PM5 is typically used in clinical laboratories if the ClinVar submission has a “review status” of 2 stars or multiple submitters of P and LP interpretations without any conflicts or 3 stars expert panel.

Knowledge Databases \ Previous Interpretations – ClinVar, HGMD

Variant of Interest: NM_000249.3(MLH1):c.1038G>T (p.**Gln346**His) (p.**Q346**H)

PS1

HGMD® Professional 2020.4

CM1812352	CAG-CAC	Gln346His	c.1038G>C	p.Q346H	DM	Colorectal cancer, non-polyposis	Shirts (2018) Am J Hum Genet **103**, 19
CM092210	CAG-CAT	Gln346His	c.1038G>T	p.Q346H	DM	Colorectal cancer, non-polyposis	Tang (2009) Clin Genet **75**, 334 Pagenstecher (2006) *Hum Genet* **119**: 9 [Functional characterisation] Zhu (2013) *Oncol Lett* **5**: 1710 [Additional report] 2 more reference(s)...

- Potential PS1 or PM5 – if there is literature available for the same missense variant or a similar substitution without a ClinVar assertion, carefully review the data for a potential application of PM5 or PS1.
- The variant has to stand on its own merits as P/LP for use of PS1 or PM5.

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type: Loss of function	PVS1
	(2) Variant Impact/Type: In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type: Loss of function	PVS1
	(2) Variant Impact/Type: In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1
	(8) Previously reported cases?	PS4, BS2, BP5

PS4

Case Prevalence (Odds Ratio) (PS4)

High Prevalence or Multiple Unrelated Patients Observed with Variant and Phenotype

- The prevalence of the variant is increased in affected individuals is significantly increased compared with the prevalence in controls.
- Relative risk (RR) or odds ratio (OR) in a case-control study is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0.
- **IF some genetic diseases have a very low prevalence (1: 1,000,000) is this feasible?**

PS4

OR Previously Reported Cases (PS4)

High Prevalence or Multiple Unrelated Patients Observed with Variant and Phenotype

- The prevalence of the variant is increased in affected individuals is significantly increased compared with the prevalence in controls.
- Relative risk (RR) or odds ratio (OR) in a case-control study is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0.
- **IF some genetic diseases have a very low prevalence (1: 1,000,000) is this feasible?**

PS4

Previously Reported Cases (PS4)

High Prevalence or Multiple Unrelated Patients Observed with Variant and Phenotype

- In instances of very rare variants where case-control studies are not feasible or difficult to reach statistical significance…
- The prior observation of the variant in multiple unrelated patients with the same phenotype, and its absence in controls, may be used as strong, moderate or supporting level of evidence.
- How do we define "multiple" ?

PS4

Previously Reported Cases (PS4)

High Prevalence or Multiple Unrelated Patients Observed with Variant and Phenotype

Table 4 Overview of Case-Level Data Specifications: Point Value Thresholds per Strength Level for Proband Count Thresholds per Variant Curation Expert Panel for PS4

		Supporting	Moderate	Strong	Very strong
PS4	Cardiomyopathy	2 probands	6 probands	15 probands	N/A
	RASopathy	1 proband	3 probands	5 probands	N/A
	PTEN	1 point	2 points	4 points	16 points
	CDH1	1 proband	2 probands	4 probands	16 probands
	Hearing loss (AD)	2 probands	6 probands	15 probands	N/A

In most cases, this approach requires PM2 to be applicable. In other words the variant must be absent from population databases or ultra rare.

BS2

- Observed in a healthy adult individual for a recessive (homozygous/compound heterozygous with pathogenic variant), dominant (heterozygous), or X-linked (hemizygous) disorder with full penetrance expected at an early age.
- Benign counterpart to PS4.

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type: Loss of function	PVS1
	(2) Variant Impact/Type: In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1
	(8) Previously reported cases?	PS4, BS2,
	(9) Phenotype specificity	PP4

PP4 Specific Phenotype

- Patient's phenotype or family history is highly specific for a disease with a single mono genetic etiology.
- Disorders with specific biochemical findings are easier to account for in this criterion than those with non-specific phenotype like "neurodevelopmental delays" or "dysmorphologies".
- Functional evidence from patient-derived material best reflects the organismal phenotype and, in general, it would be better to use this evidence to satisfy PP4 (specific phenotype)

PP4 Specific Phenotype

- Examples are:
  - Mitochondrial deficiency demonstrated in electron transport chain activity deficiency from patient fibroblasts.
  - PAH (-Plasma phenylalanine >120umol/L)

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

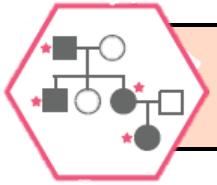
Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type: Loss of function	PVS1
	(2) Variant Impact/Type: In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1
	(8) Previously reported cases?	PS4, BS2, BP5
	(9) Phenotype specificity	PP4

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type — Loss of function	PVS1
	(2) Variant Impact/Type — In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1
	(8) Previously reported cases?	PS4, BS2, BP5
	(9) Phenotype specificity	PP4
	(10) Segregation? De novo?	PP1, BS4. PS2, PM6

Case-Specific Evidence - Segregation Data



PP1

Co-segregation with disease in multiple affected family members of a single family in a gene definitively known to cause the disease.

Note: May be used as stronger evidence with increasing segregation data.

Case-Specific Evidence – De novo ocurrence

De novo (both maternity and paternity confirmed) in a patient with the disease and no family history.

PM6

De novo (maternity and paternity <u>not confirmed)</u> in a patient with the disease and no family history.

Table 1. Points* awarded per *de novo* occurrence

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

*Note that these points are *not* equivalent to the points used to classify a variant per the Tavtigian et al 2020 "Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines"
**Maximum allowable value of 1 may contribute to overall score

Table 2. Recommendation for determining the appropriate ACMG/AMP evidence strength level for *de novo* occurrence(s)

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

https://www.clinicalgenome.org/site/assets/files/3461/svi_proposal_for_de_novo_criteria_v1_1.pdf

Variant Interpretation Framework Summary

((11 questions to always ask from a variant)

Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type: Loss of function	PVS1
	(2) Variant Impact/Type: In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1
	(8) Previously reported cases?	PS4, BS2, BP5
	(9) Phenotype specificity	PP4
	(10) Segregation? De novo?	PP1, BS4. PS2, PM6

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

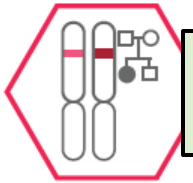
Concept	Questions	ACMG Criteria
Allele Frequency	(1) Common or rare?	BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type — Loss of function	PVS1
	(2) Variant Impact/Type — In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?	PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level	PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?	PM1
	(6) Variant effect functionally studied?	PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar	~~PP5~~, PM5, PS1
	(8) Previously reported cases?	PS4, BS2, BP5
	(9) Phenotype specificity	PP4
	(10) Segregation? De novo?	PP1, BS4. PS2, PM6
	(11) Trans / cis observations	PM3, BP2

Case-Specific Evidence – Phase Evidence

PM6

For recessive disorders, detected in trans with a pathogenic variant.

This requires testing of parents (or offspring) to determine phase.



BP2

Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern.

Variant Interpretation Framework Summary

(11 questions to always ask from a variant)

Concept	Questions		ACMG Criteria
Allele Frequency	(1) Common or rare?		BA1, BS1, PM2
Computational & Predictive Data	(2) Variant Impact/Type	Loss of function	PVS1
		In-frame indel	PM4, BP3
	(3) In-silico predictions? Potential splicing impact?		PP3, BP4 BP7
	(4) Constraint metrics Gene/regional level		PP2, BP1
Functional Knowledge	(5) Residue/Domain? Hotspot?		PM1
	(6) Variant effect functionally studied?		PS3, BS3
Clinical Knowledge (published, or case/sample specific)	(7) Interpretation Databases - ClinVar		~~PP5 ,~~ PM5, PS1
	(8) Previously reported cases?		PS4, BS2, BP5
	(9) Phenotype specificity		PP4
	(10) Segregation? De novo?		PP1, BS4. PS2, PM6
	(11) Trans / cis observations		PM3, BP2

	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 In-frame indels in repeat w/out known function BP3	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	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and 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			

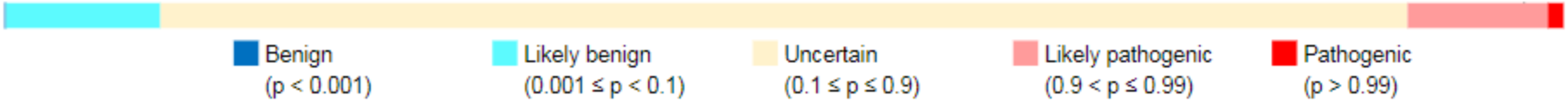
2015 ACMG Guidelines

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

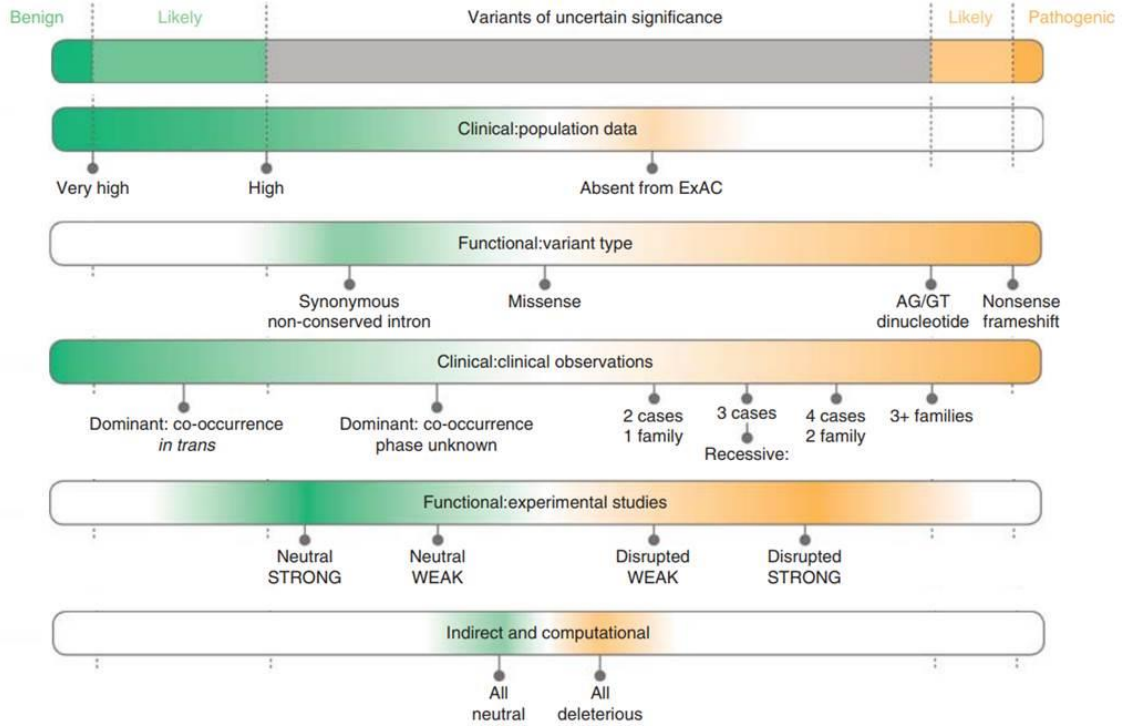
Sue Richards, PhD[1], Nazneen Aziz, PhD[2,16], Sherri Bale, PhD[3], David Bick, MD[4], Soma Das, PhD[5], Julie Gastier-Foster, PhD[6,7,8], Wayne W. Grody, MD, PhD[9,10,11], Madhuri Hegde, PhD[12], Elaine Lyon, PhD[13], Elaine Spector, PhD[14], Karl Voelkerding, MD[13] and Heidi L. Rehm, PhD[15]; on behalf of the ACMG Laboratory Quality Assurance Committee

Benign	Likely Benign	Variants of Unknown Significance	Likely Pathogenic	Pathogenic

A heuristic system for variant classification that is compatible with a formal, quantitative, naive Bayesian classifier.



The ACMG guidelines are not mandatory or the only ones used

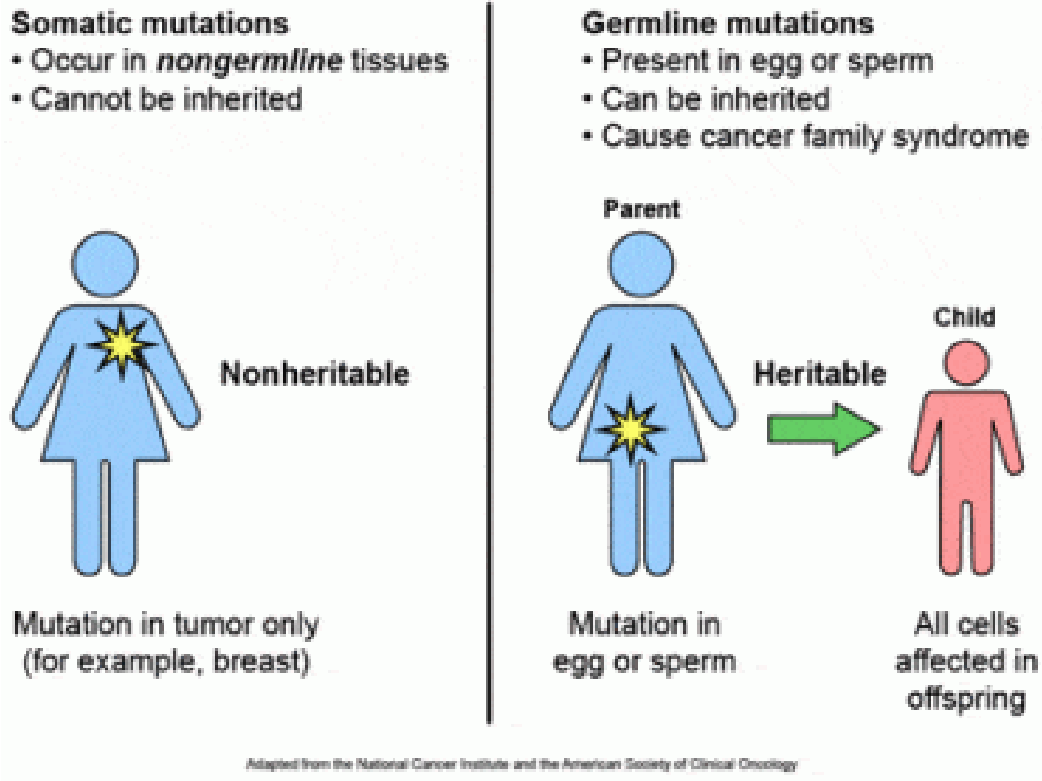


Warning!

Germline and Somatic Classification and Catalogue Differences

OncoKB

COSMIC
Catalogue Of Somatic Mutations In Cancer



OMIM®

human phenotype ontology

ClinVar

Warning!

Germline and Somatic Classification and Catalogue Differences

Categories:

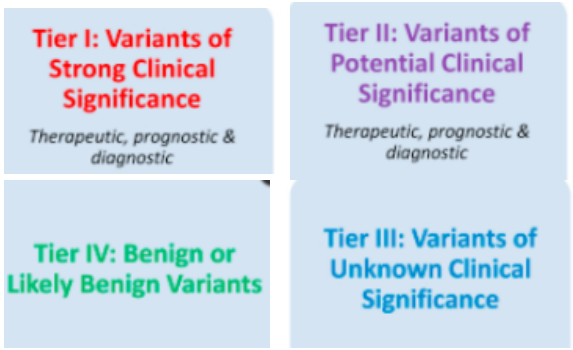
Diagnostic

Prognostic

Therapeutic

Somatic mutations
- Occur in ***nongermline*** tissues
- Cannot be inherited

Germline mutations
- Present in egg or sperm
- Can be inherited
- Cause cancer family syndrome



Categories:

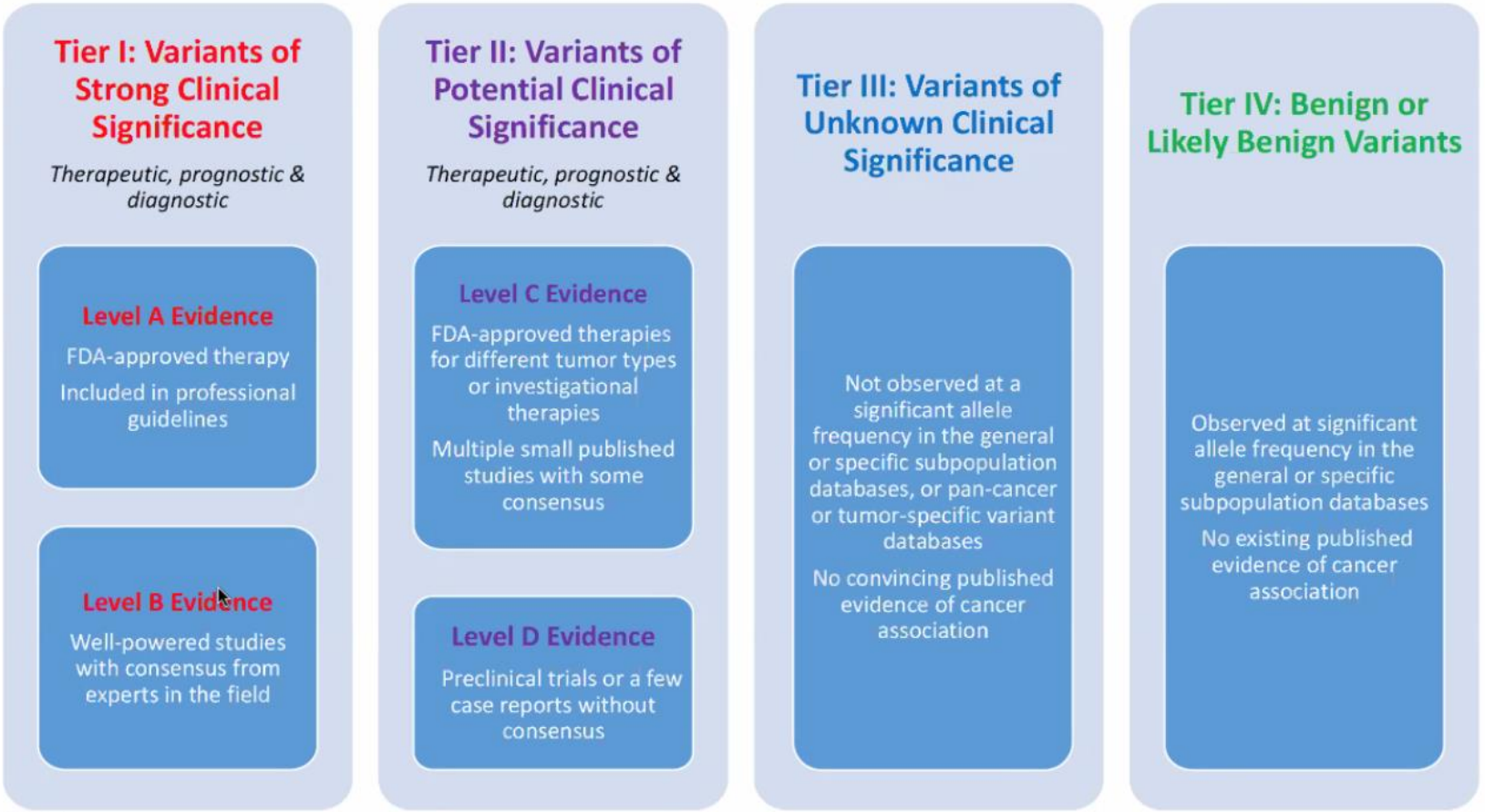
Pathogenic

Likely Pathogenic

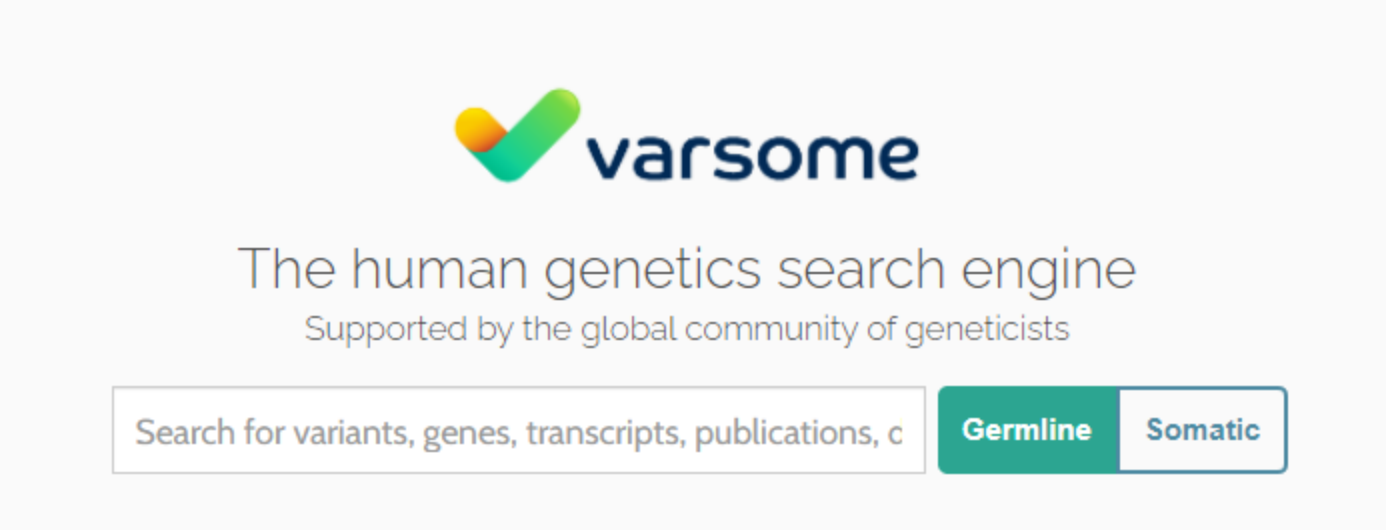
VUS – Variant of Uncertain Significance

Likely Benign

Benign


Tier I: Variants of Strong Clinical Significance
Therapeutic, prognostic & diagnostic
Level A Evidence
FDA-approved therapy
Included in professional guidelines
Level B Evidence
Well-powered studies with consensus from experts in the field
Tier II: Variants of Potential Clinical Significance
Therapeutic, prognostic & diagnostic
Level C Evidence
FDA-approved therapies for different tumor types or investigational therapies
Multiple small published studies with some consensus
Level D Evidence
Preclinical trials or a few case reports without consensus
Tier III: Variants of Unknown Clinical Significance
Not observed at a significant allele frequency in the general or specific subpopulation databases, or pan-cancer or tumor-specific variant databases
No convincing published evidence of cancer association
Tier IV: Benign or Likely Benign Variants
Observed at significant allele frequency in the general or specific subpopulation databases
No existing published evidence of cancer association

Please sign in to Varsome for the lab section



Questions?