ClinGen
Clinical Genome Resource

Defining the clinical relevance of genes & variants for precision medicine and research...

Scoring Intragenic CNVs

Daniel E. Pineda-Alvarez, MD, FACMG
Medical Director of Oncology
Invitae
daniel.pineda@invitae.com

(On behalf of the ACMG/ClinGen Copy Number Variant Interpretation Working Group)

Today’s Attendance URL and QR code:

https://tinyurl.com/AttendanceFeb62020

Reminder: Questions

- Type your questions into the “Q&A” box
- Monitored by committee members
  - Will verbally answer as many as time allows (at the end)
  - May receive a typed answer through the Q&A box from a committee member
  - Unanswered questions will be saved for consideration during the March 12 Q&A webinar

Today's Focus

Section 1: Initial assessment of genomic content			
Evidence type	**Evidence**	**Suggested points/case**	**Max score**
Copy-number loss content	**1A**. Contains protein-coding or other known functionally important elements.	0 (Continue evaluation)	0
	1B. Does NOT contain protein-coding or any known functionally important elements.	−0.60	−0.60
Section 2: Overlap with established/predicted haploinsufficiency (HI) or established benign genes/genomic regions *(Skip to section 3 if your copy-number loss DOES NOT overlap these types of genes/regions)*			
Overlap with ESTABLISHED HI genes or genomic regions and consideration of reason for referral	**2A**. Complete overlap of an established HI gene/genomic region.	1.00	1.00
	2B. Partial overlap of an established HI genomic region • The observed CNV does NOT contain the known causative gene or critical region for this established HI genomic region OR • Unclear if known causative gene or critical region is affected OR • No specific causative gene or critical region has been established for this HI genomic region	0 (Continue evaluation)	0
	2C. Partial overlap with the 5′ end of an established HI gene (3′ end of the gene not involved)…	See categories below	
	2C-1. …and coding sequence is involved	0.90 *(range: 0.45 to 1.00)*	1.00
	2C-2. …and only the 5′ UTR is involved	0 *(range: 0 to 0.45)*	0.45
	2D. Partial overlap with the 3′ end of an established HI gene (5′ end of the gene not involved)…	See categories below	
	2D-1. …and only the 3′ untranslated region is involved.	0 (Continue evaluation)	0
	2D-2. …and only the last exon is involved. Other established pathogenic variants have been reported in this exon.	0.90 *(range: 0.45 to 0.90)*	0.90
	2D-3. …and only the last exon is involved. No other established pathogenic variants have been reported in this exon.	0.30 *(range: 0 to 0.45)*	0.45
	2D-4. …and it includes other exons in addition to the last exon. Nonsense-mediated decay is expected to occur.	0.90 *(range: 0.45 to 1.00)*	1.00
	2E. Both breakpoints are within the same gene (intragenic CNV; gene-level sequence variant).	See ClinGen SVI working group PVS1 specifications • PVS1 = 0.90 *(Range: 0.45 to 0.90)* • PVS1_Strong = 0.45 *(Range: 0.30 to 0.90)* • PVS1_Moderate *or* PM4 (in-frame indels) = 0.30 *(Range: 0.15 to 0.45)* • PVS1_Supporting = 0.15 *(Range: 0 to 0.30)*	See categories at left

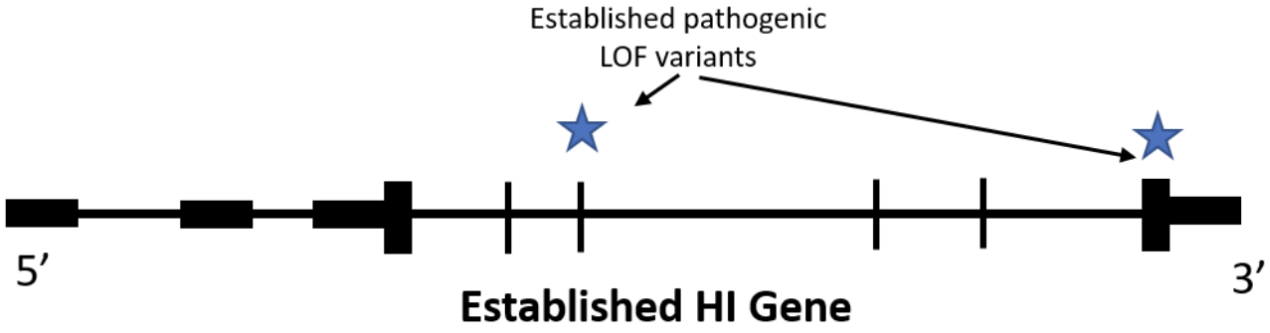
Today’s Focus

Overlap with ESTABLISHED HI gene(s)	**2H**. HI gene fully contained within observed copy-number gain.	0 (Continue evaluation)	0
Breakpoint(s) within ESTABLISHED HI genes	**2I**. Both breakpoints are within the same gene (gene-level sequence variant, possibly resulting in loss of function [LOF]).	See ClinGen SVI working group PVS1 specifications • PVS1 = 0.90 *(Range: 0.45 to 0.90)* • PVS1_Strong = 0.45 *(Range: 0.30 to 0.90)* • N/A = 0 (Continue evaluation)	

Where to begin?

- Assess whether a CNV overlaps an established haploinsufficiency (HI) gene
- Assess whether the testing assay can define if the maximum CNV interval is:
  - Overlapping with a HI gene and extending beyond that gene OR
  - Both breakpoints are within a HI gene
- Determine the ”type” of CNV and its impact to the gene’s open-reading frame
  - Frame-shift
    - Leading to non-sense mediated decay (NMD)
    - Non-NMD
  - In-frame
    - Generally Non-NMD
    - Essential residues being deleted?

Types of Intragenic CNVs



Category 2C-1 — Overlap with 5' end, coding sequence involved

Category 2C-2 — Only 5' UTR is involved

Category 2D-1 — Only 3' UTR is involved

Category 2D-2 — Only the last exon is involved; other established pathogenic variants have been reported in this exon

Category 2D-4 — Overlap with 3' end; includes other exons in addition to the last exon; nonsense-mediated decay expected to occur

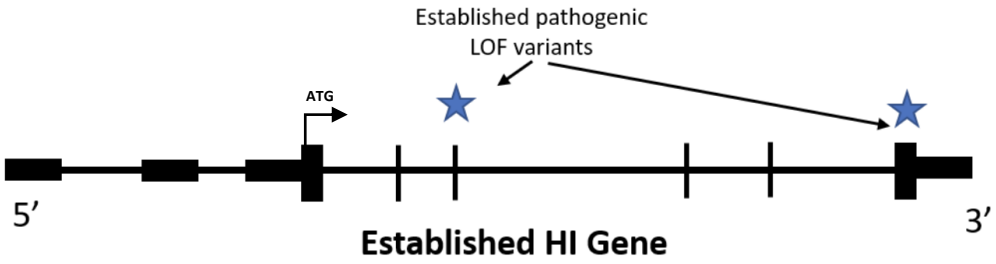
Category 2E — Both breakpoints are within the HI gene

Category 2H — HI gene fully contained within observed copy number gain

Category 2I — Both breakpoints are within the HI gene

Category 2J or 2K — One breakpoint within the HI gene

5’ and 3’-end Gene Deletions: Absent of Protein Product or Truncated Protein Product



Category 2C-1 Overlap with 5’ end, coding sequence involved

Category 2C-2 Only 5’ UTR is involved

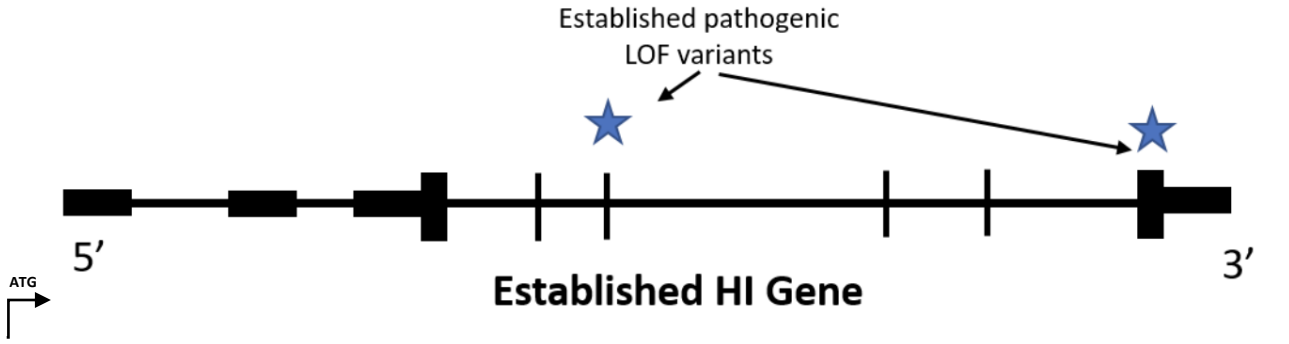
Category 2D-1 Only 3’ UTR is involved

Category 2D-2 Only the last exon is involved; other established pathogenic variants have been reported in this exon

Overlap with 3’ end; includes other exons in addition to the last exon; nonsense-mediated decay expected to occur Category 2D-4

	2C. Partial overlap with the 5’ end of an established HI gene (3’ end of the gene not involved)…	See categories below	
	2C-1. …and coding sequence is involved	0.90 *(range: 0.45 to 1.00)*	1.00
	2C-2. …and only the 5’ UTR is involved	0 *(range: 0 to 0.45)*	0.45
	2D. Partial overlap with the 3’ end of an established HI gene (5’ end of the gene not involved)…	See categories below	
	2D-1. …and only the 3’ untranslated region is involved.	0 (Continue evaluation)	0
	2D-2. …and only the last exon is involved. Other established pathogenic variants have been reported in this exon.	0.90 *(range: 0.45 to 0.90)*	0.90
	2D-3. …and only the last exon is involved. No other established pathogenic variants have been reported in this exon.	0.30 *(range: 0 to 0.45)*	0.45
	2D-4. …and it includes other exons in addition to the last exon. Nonsense-mediated decay is expected to occur.	0.90 *(range: 0.45 to 1.00)*	1.00

Section 2C-1: 5’ of the Gene including Initiation codon



Category 2C-1 | Overlap with 5’ end, coding sequence involved

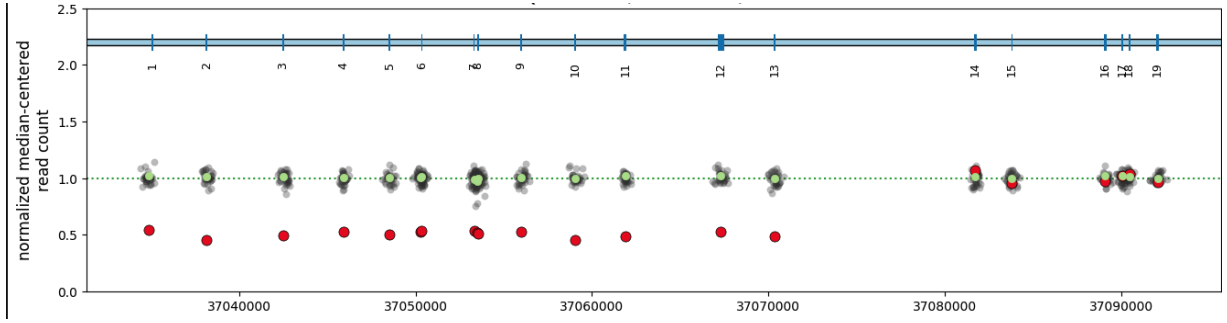
- Generally these deletions results in loss of the transcriptional start of the gene
- Absent gene product lead to HI

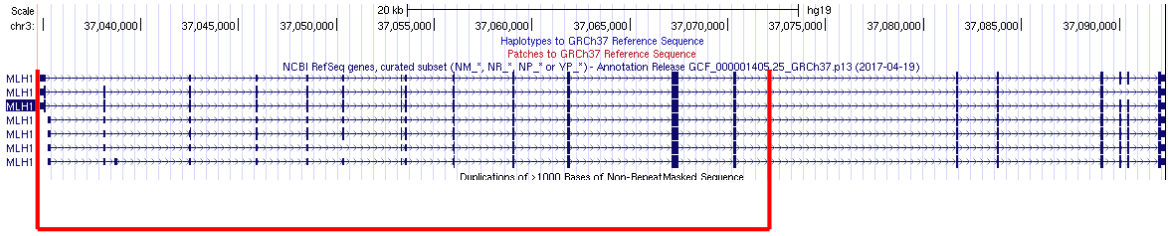
- Default score: 1.0 (Pathogenic)
- Ranges: 0.45 -1.0
- Special considerations:
  - Initiator-codon rescue with and in-frame ATG (alternative isoform or transcript).
  - Use less weight if there is a known alternative isoform

Example: 5’-End deletion in MLH1

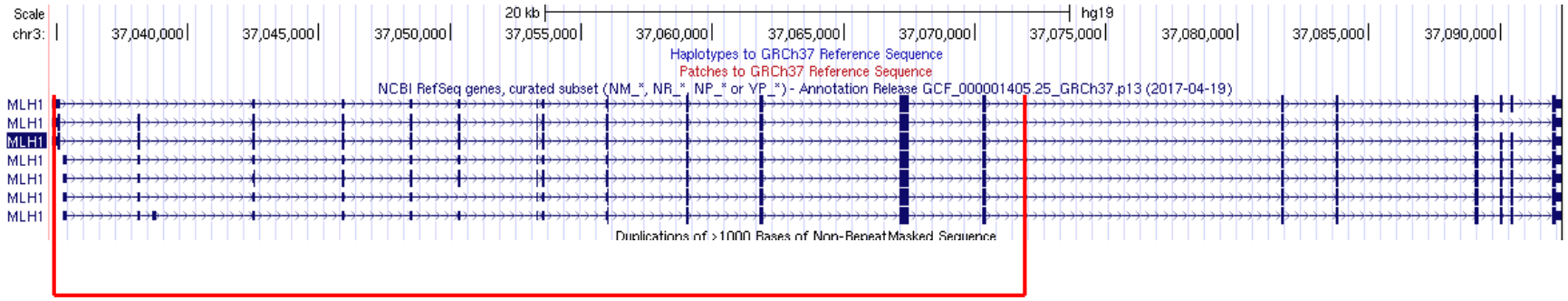
- Patient diagnosed with colorectal cancer in the 40s, adenocarcinoma had loss of MLH1 and PMS2 in IHC.
- Family history: Several maternal relatives with Lynch syndrome (LS) related cancers
- Variant reported in the literature in individuals with LS

MLH1 NM_000249.3 deletion of exons 1-13





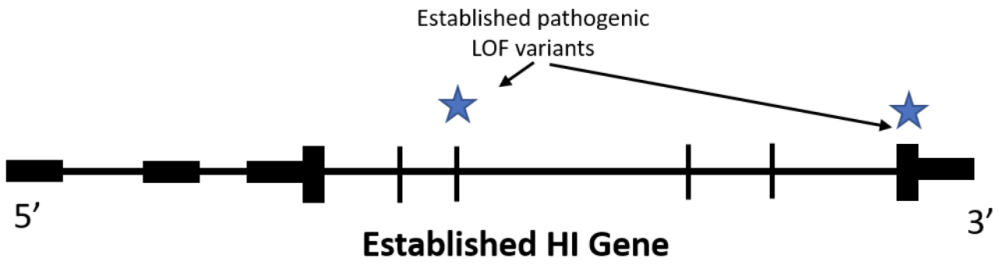
Section 1: Initial assessment of genomic content			
Evidence type	**Evidence**	**Suggested points/case**	**Max score**
Copy-number loss content	**1A**. Contains protein-coding or other known functionally important elements.	0 (Continue evaluation)	0
	1B. Does NOT contain protein-coding or any known functionally important elements.	−0.60	−0.60
Section 2: Overlap with established/predicted haploinsufficiency (HI) or established benign genes/genomic regions *(Skip to section 3 if your copy-number loss DOES NOT overlap these types of genes/regions)*			
Overlap with ESTABLISHED HI genes or genomic regions and consideration of reason for referral	**2A**. Complete overlap of an established HI gene/genomic region.	1.00	1.00
	2B. Partial overlap of an established HI genomic region • The observed CNV does NOT contain the known causative gene or critical region for this established HI genomic region OR • Unclear if known causative gene or critical region is affected OR • No specific causative gene or critical region has been established for this HI genomic region	0 (Continue evaluation)	0
	2C. Partial overlap with the 5′ end of an established HI gene (3′ end of the gene not involved)…	See categories below	
	2C-1. …and coding sequence is involved	0.90 *(range: 0.45 to 1.00)*	1.00
	2C-2. …and only the 5′ UTR is involved	0 *(range: 0 to 0.45)*	0.45
	2D. Partial overlap with the 3′ end of an established HI gene (5′ end of the gene not involved)…	See categories below	



Pathogenic: **1.0** score

- Expected to result in absent protein product
- No alternate initiation codon after exon 13.

Section 2C-2: 5’ of the Gene including Initiation codon



Category 2C-2 Only 5’ UTR is involved

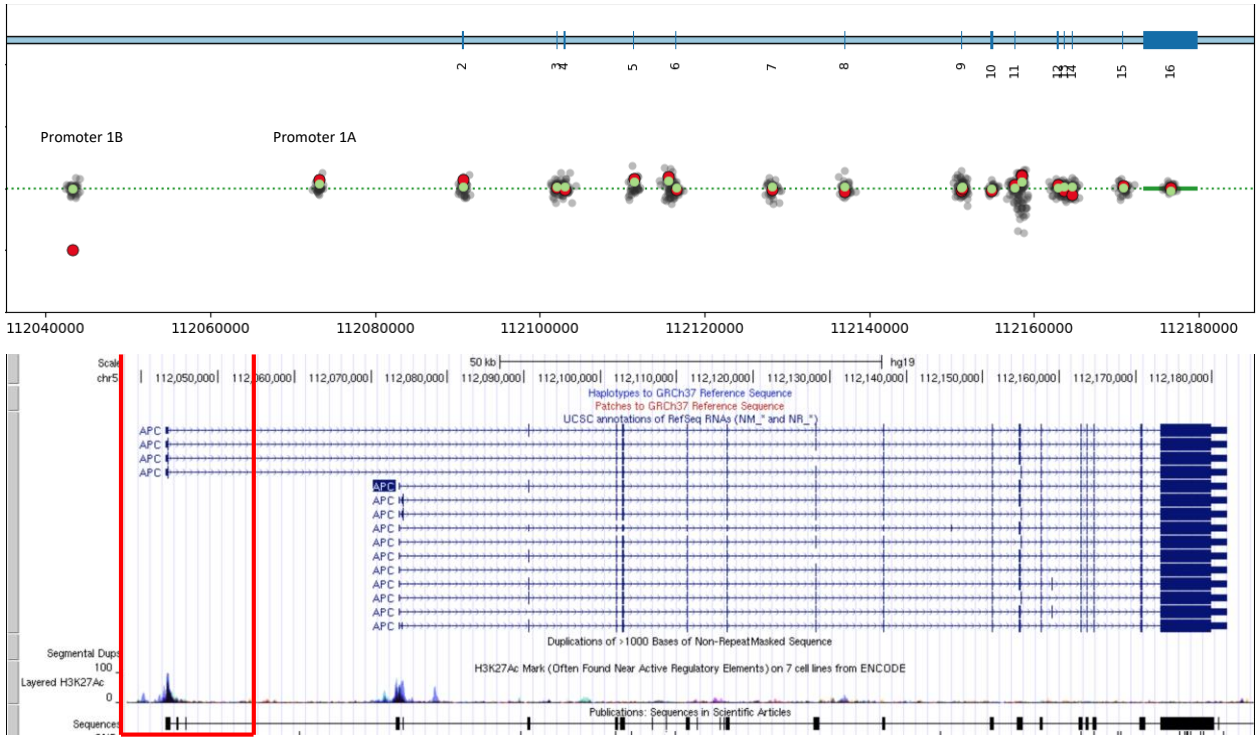
- These are tricky!
- These are scored with less weight and should be VUS in absence of other evidence

- Default score: 0 (VUS)
- Ranges: 0-0.45
- Special considerations:
  - If there is a known, well documented promoter in the deleted 5’UTR region or even upstream of the gene, the score may be upgraded to account to the loss of function (HI).

Example: Deletion of Promoter in *APC*

- Haploinsufficiency of the APC gene is known to cause Familial Adenomatous Polyposis (FAP)
- Deletion of Promoter 1B of APC
  - Deletion of this region leads to reduced expression of APC (PMID: 21643010, 23725351, 25243319)
  - The variant has been observed in more than 4 unrelated cases with FAP (Internal database)
  - Deletion of this region has been seen to segregate with FAP in at least 11 affected individuals of 7-generation family with FAP (PMID:21643010)

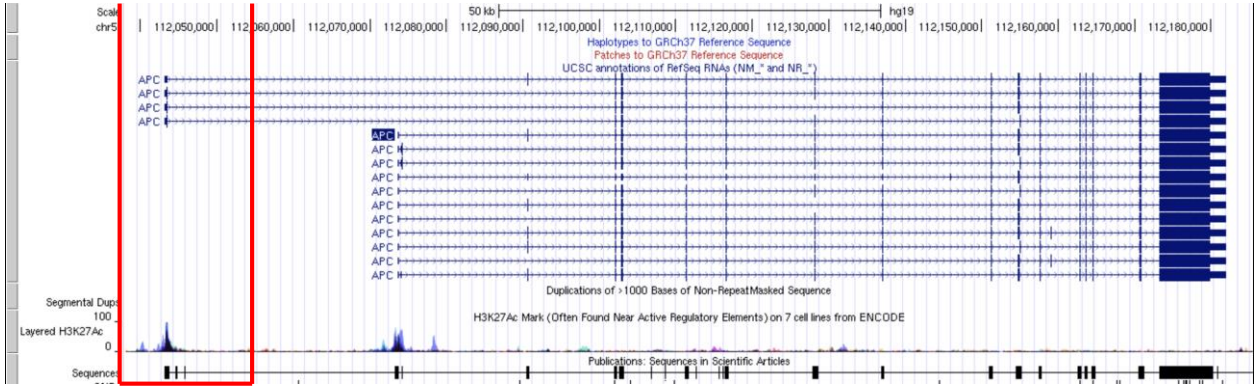
APC deletion of promoter 1B



Pathogenic: **1.0** score

- Combination of deletion of promoter, case reports and segregation

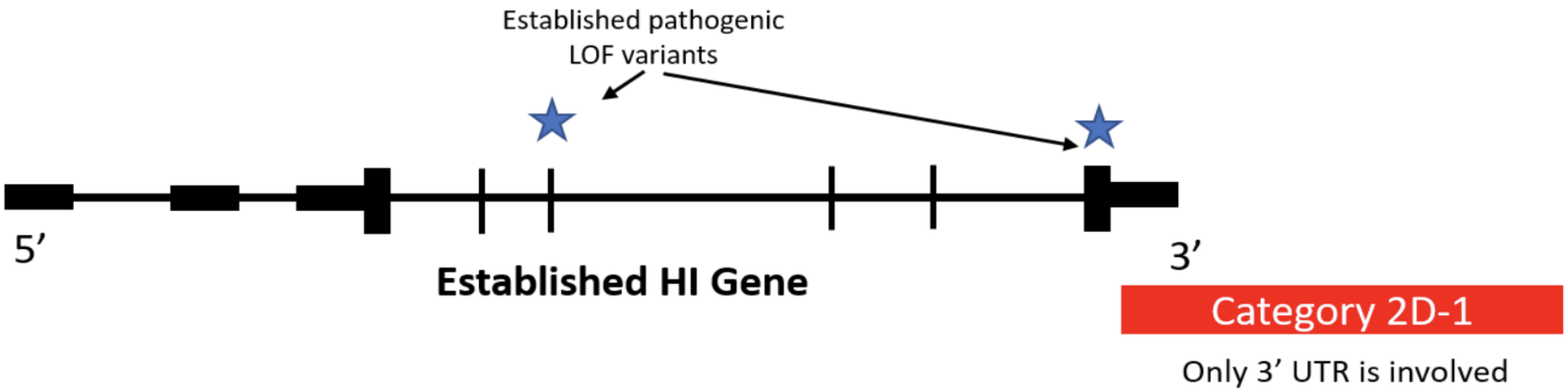
Section 1: Initial assessment of genomic content			
Evidence type	**Evidence**	**Suggested points/case**	**Max score**
Copy-number loss content	**1A**. Contains protein-coding or other known functionally important elements.	0 (Continue evaluation)	0
	1B. Does NOT contain protein-coding or any known functionally important elements.	−0.60	−0.60
Section 2: Overlap with established/predicted haploinsufficiency (HI) or established benign genes/genomic regions *(Skip to section 3 if your copy-number loss DOES NOT overlap these types of genes/regions)*			
Overlap with ESTABLISHED HI genes or genomic regions and consideration of reason for referral	**2A**. Complete overlap of an established HI gene/genomic region.	1.00	1.00
	2B. Partial overlap of an established HI genomic region • The observed CNV does NOT contain the known causative gene or critical region for this established HI genomic region OR • Unclear if known causative gene or critical region is affected OR • No specific causative gene or critical region has been established for this HI genomic region	0 (Continue evaluation)	0
	2C. Partial overlap with the 5′ end of an established HI gene (3′ end of the gene not involved)…	See categories below	
	2C-1. …and coding sequence is involved	0.90 *(range: 0.45 to 1.00)*	1.00
	2C-2. …and only the 5′ UTR is involved	0 *(range: 0 to 0.45)*	0.45
Section 3: Evaluation of gene number			
Number of protein-coding RefSeq genes wholly or partially included in the copy-number loss	**3A**. 0–24 genes	0	0
Individual case evidence—unknown inheritance	**4E**. Reported proband has a highly specific phenotype consistent with the gene/genomic region, but the inheritance of the variant is unknown.	0.10 points each *(range: 0 to 0.15)*	0.30 (total)
Individual case evidence—segregation among similarly affected family members	**4F**. 3–4 observed segregations	0.15	0.45
	4G. 5–6 observed segregations	0.30	
	4H. 7 or more observed segregations	0.45	



Pathogenic: **Score = 1.0**

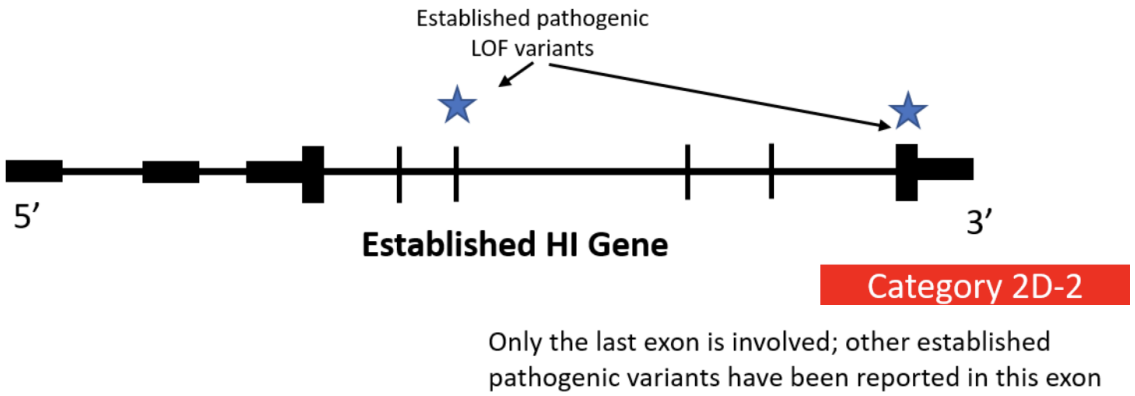
- **Section 1:** Score = 0
- **Section 2:** Score = 0.45
  - **2C-2**: Deletion of 5’UTR encompassing a known promoter characterized by RNA.
- **Section 3:** Score = 0
- **Section 4:** Score = 0.30 – 0.45
  - **4E:** Observed in patients with FAP (>4)
  - **4H:** Over 7 segregations in a family with FAP

Section 2D-1: Deletion of the 3’ UTR



- In general these deletions do not result in loss of function
- Default score: 0 (VUS)
- Continue evaluation and determine in other evidence apply

Section 2D-2 and D-3: 3’ of the Gene including Stop codon



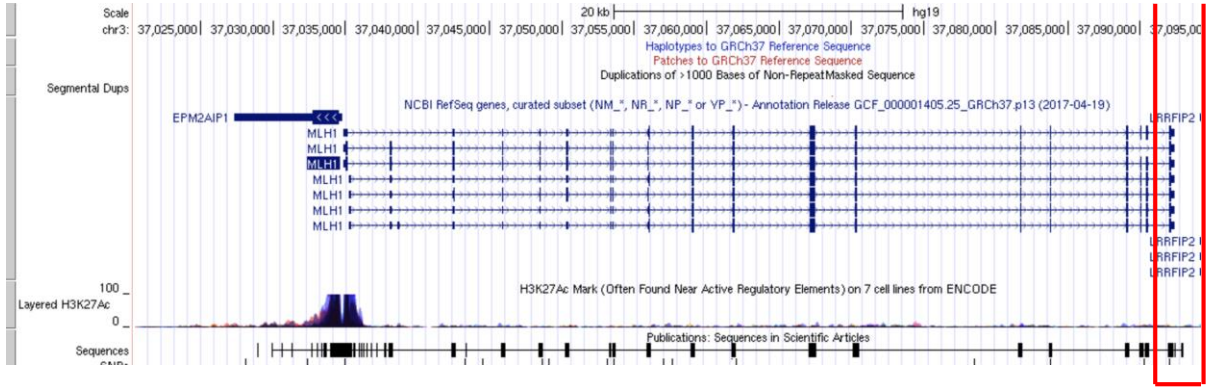
- These deletions do not necessarily lead to nonsense-mediated RNA decay
- **Truncated protein** which may or may not have normal physiologic function

- 2D-2: Other pathogenic variants (SVs) have been reported in the deleted exon
  - Implies that the deleted region is essential for protein function
  - Score: 0.90 (Range: 0.45-0.90)
- 2D-3: NO pathogenic variants (SVs) have been reported in the deleted exon
  - Unclear whether the exon is essential for normal function
  - Score 0.3 (Range: 0-0.45)

Example: Deletion of last exon of MLH1

- MLH1 is associated with Lynch Syndrome
- Deletion is not expected to result in NMD
  - Eliminates the MLH1-PMS2 dimerization domain
  - There are several disease-causing truncating variants and missense variants reported in the literature and internal database (PMID: 12799449, 16338176)
  - The deletion has been observed in at least one individual with MLH1-associated LS with MSI-H and IHC.

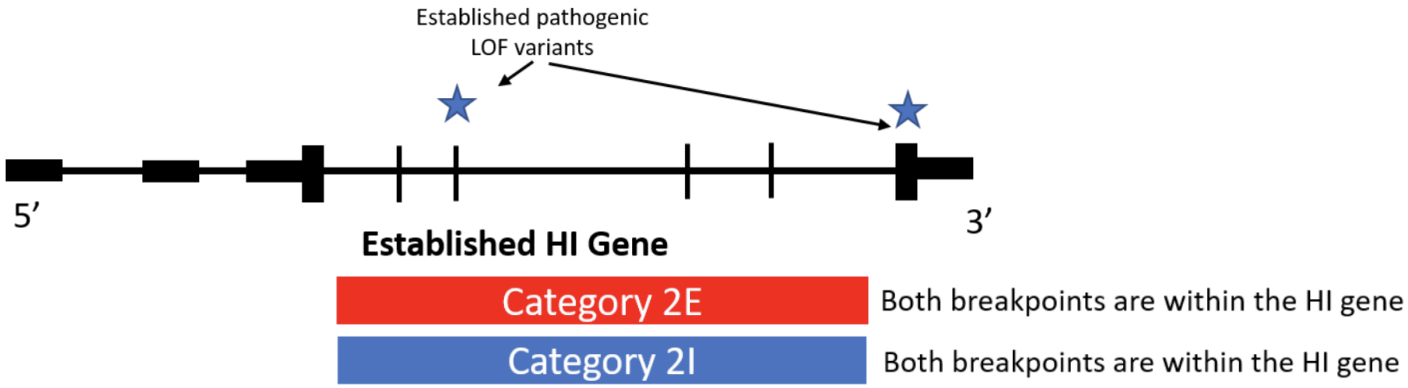
MLH1 deletion of exon 19



Pathogenic: **1.0** score

- **2D-2**: Other pathogenic variants (SVs) have been reported.
  - Score: 0.90
- **4E Case reports:** 0.10

Intragenic CNVs: Bottom line – Does it result in NMD?



2E. Both breakpoints are within the same gene (intragenic CNV; gene-level sequence variant).

See ClinGen SVI working group PVS1 specifications

- PVS1 = 0.90

(Range: 0.45 to 0.90)

- PVS1_Strong = 0.45

(Range: 0.30 to 0.90)

- PVS1_Moderate *or* PM4 (in-frame indels) = 0.30

(Range: 0.15 to 0.45)

- PVS1_Supporting = 0.15

(Range: 0 to 0.30)

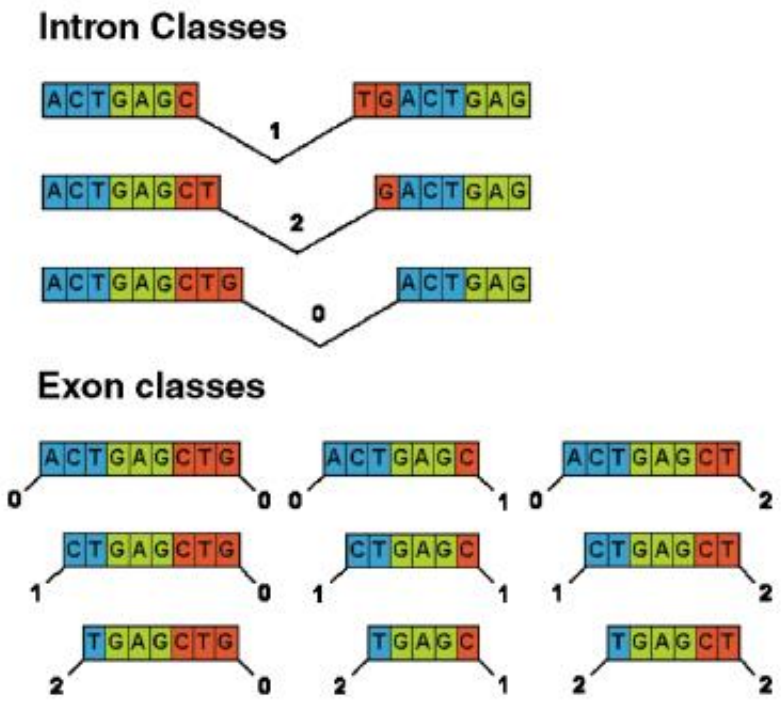
- Intragenic CNVs are scored in the frame-work of the SV interpretation guidelines using modified PVS1 criteria
- The SVI working group put forth a rubric to classify intragenic CNVs (PMID: 30192042) and it is adapted in this metric

First Step: Determine the effect of the reading frame of the CNV

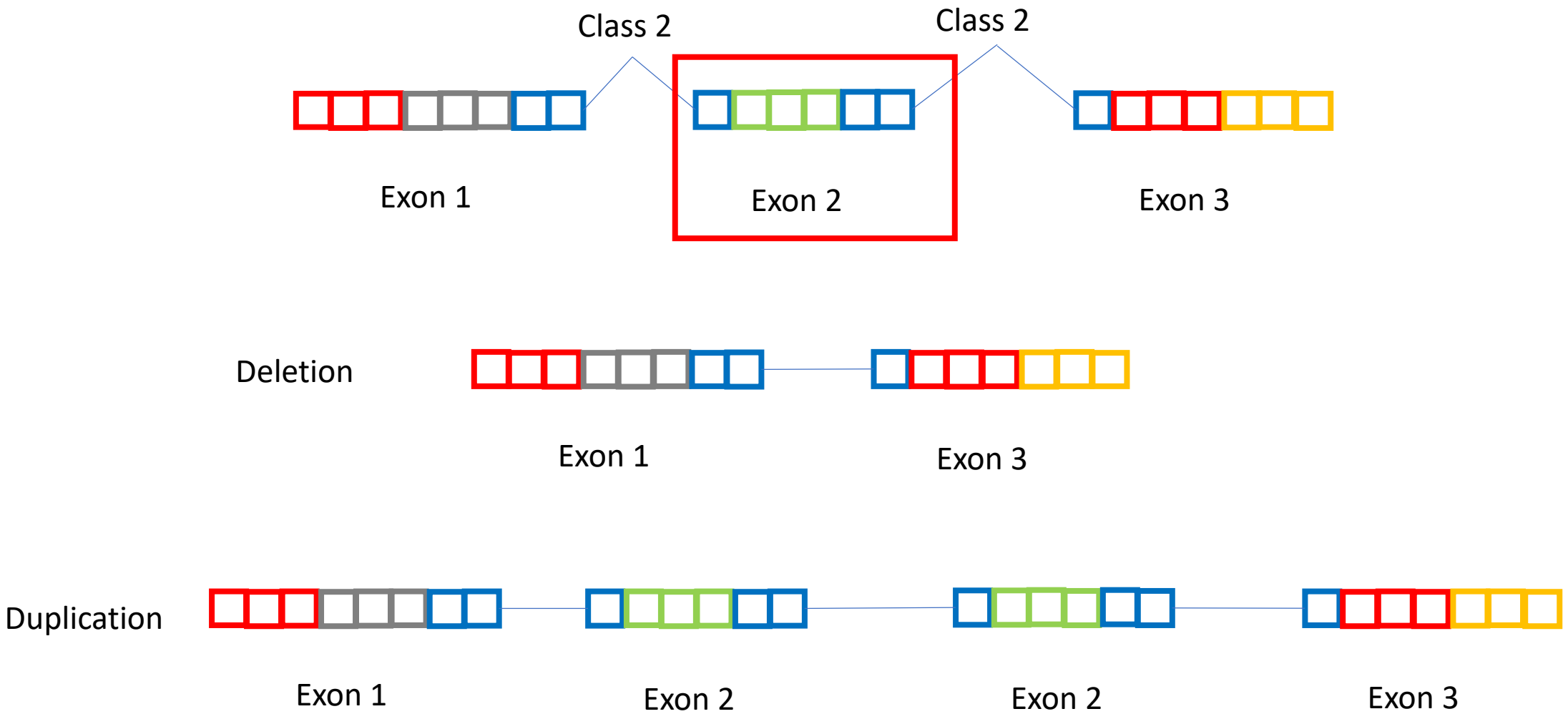
- This applies to CNVs occurring in the 2nd to penultimate exon or within a gene's open reading frame
- Out-frame deletions are predicted to result in NMD and subsequent HI
- In-Frame deletions do not lead to NMD
  - Are the deleted exons essential for the normal gene's function?
- Intragenic duplications are **likely** to occur **in tandem** and it is safe to assume most of these are in-tandem (PMID: 30192042, 31642931)
  - Out-of-frame: May lead to NMD

In-frame vs Frameshift: Intron-exon phase

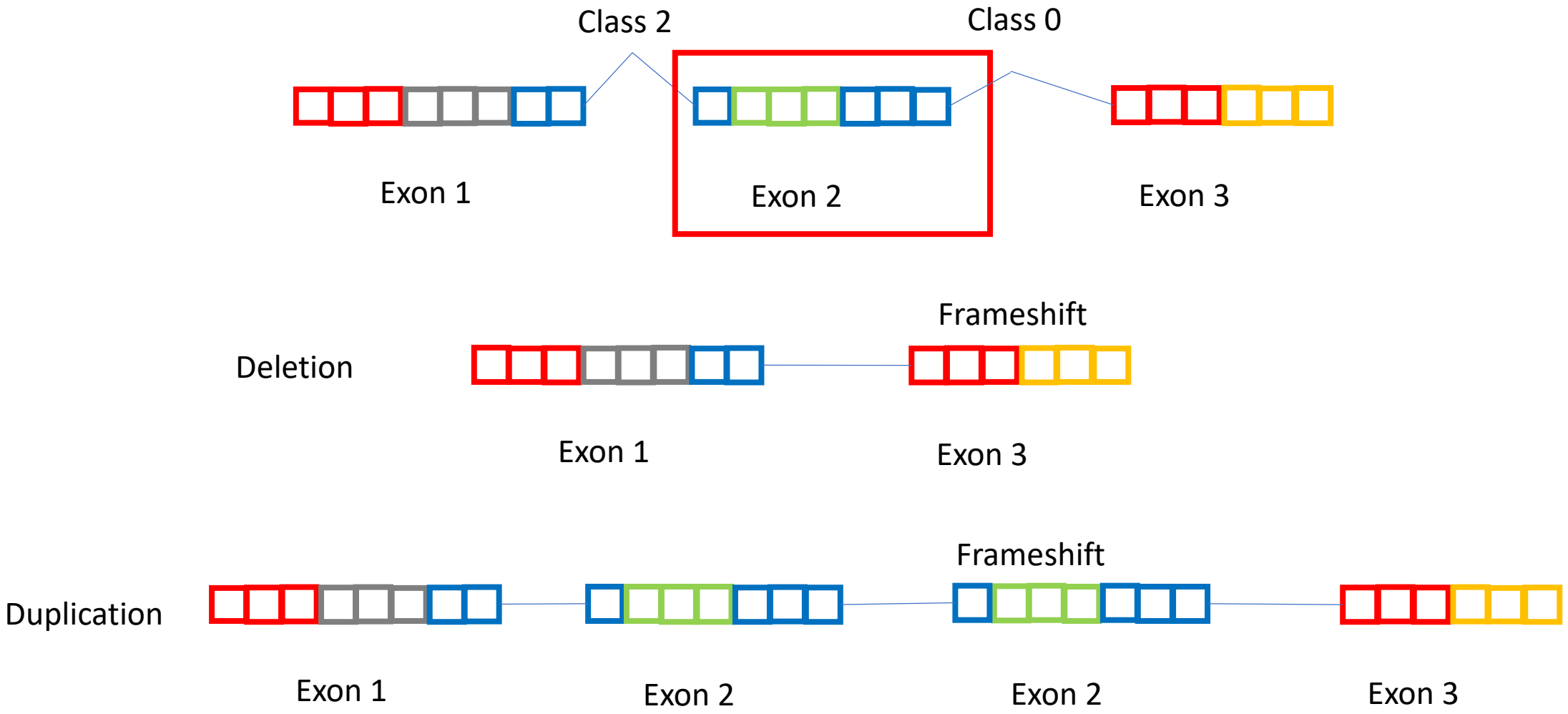
- Determine the intron-exon class by looking how codons are interrupted by introns
- Look at both sides of the CNV
  - If both sides have the same intron class, the CNV is in-frame
  - If either side has a different class, the CNV is out-of-frame
- Other method: take the **coding** nucleotides in the CNV and divide by 3, In-frame CNVs = Integers



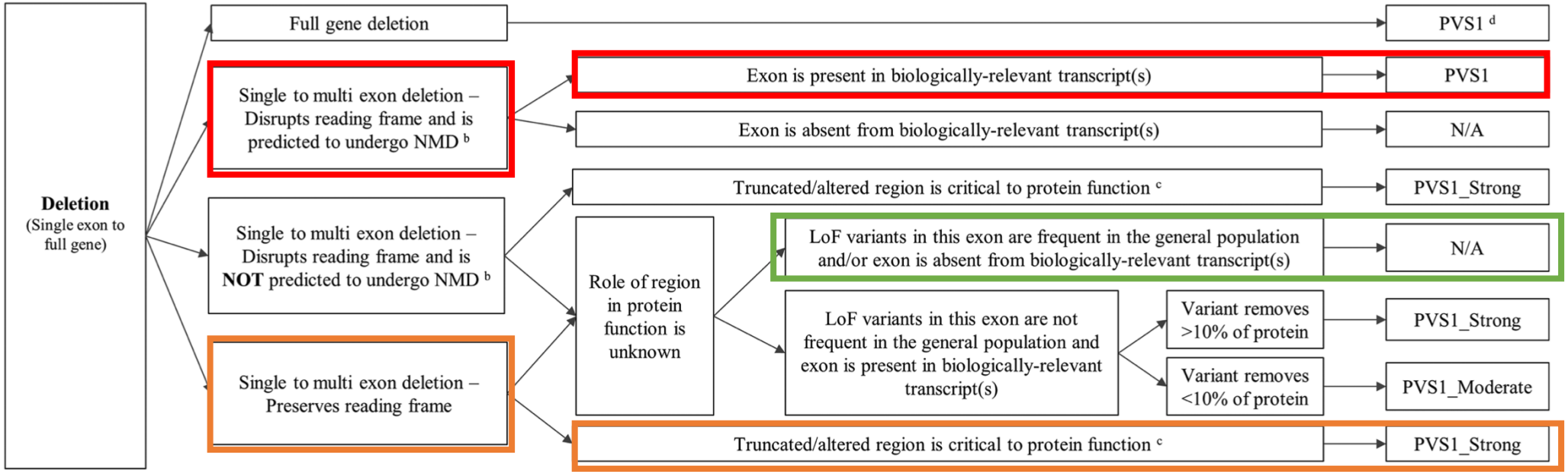
In-frame CNVs



Out-of-frame CNVs



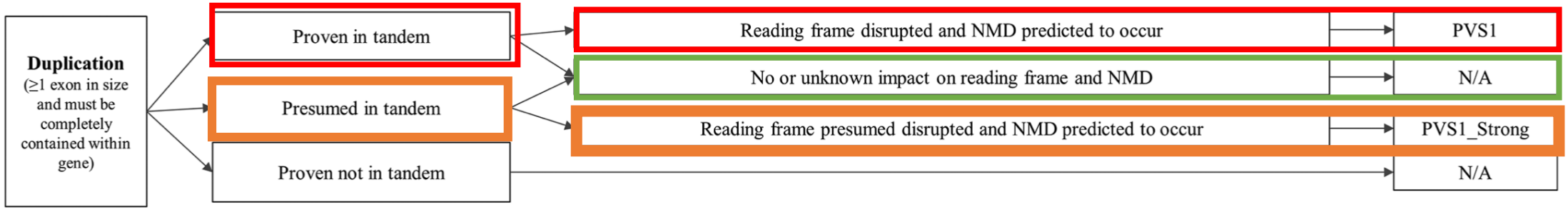
Using the modified PSV1 framework



(Modified from: Tayoun, et al., 2019)

- PVS1 = 0.90
  (Range: 0.45 to 0.90)
- PVS1_Strong = 0.45
  (Range: 0.30 to 0.90)
- PVS1_Moderate *or* PM4 (in-frame indels) = 0.30
  (Range: 0.15 to 0.45)
- PVS1_Supporting = 0.15
  (Range: 0 to 0.30)

Using the modified PSV1 framework



(Modified from: Tayoun, et al., 2019)

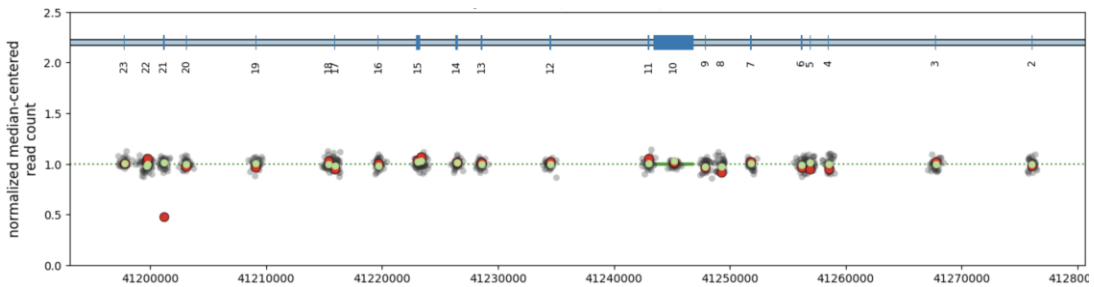
See ClinGen SVI working group PVS1 specifications

- PVS1 = 0.90
  (Range: 0.45 to 0.90)
- PVS1_Strong = 0.45
  (Range: 0.30 to 0.90)
- N/A = 0 (Continue evaluation)

Example: Out-of-frame deletion in BRCA1

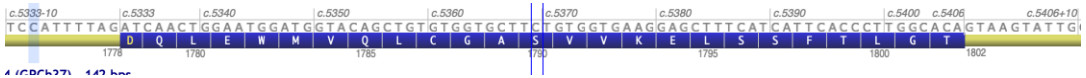
- Haploinsufficiency of BRCA1 is associated with Hereditary Breast and Ovarian Cancer (HBOC)
- This deletion is a well-known Dutch founder variant and seen to co-segregate with HBOC in these families (PMID: 9354803, 19894111, 11462239)

Heterozygous deletion of exon 21 in *BRCA1*



Class 1

Class 0



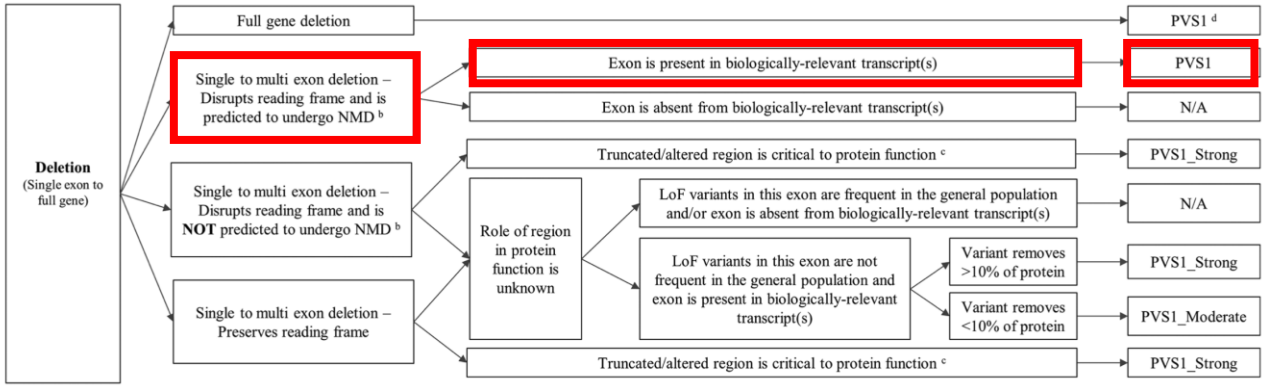
Exon 21

Out-of-frame deletion predicted to result in NMD

Section 1: Initial assessment of genomic content			
Evidence type	**Evidence**	**Suggested points/case**	**Max score**
Copy-number loss content	**1A**. Contains protein-coding or other known functionally important elements.	0 (Continue evaluation)	0
	1B. Does NOT contain protein-coding or any known functionally important elements.	−0.60	−0.60
Section 2: Overlap with established/predicted haploinsufficiency (HI) or established benign genes/genomic regions *(Skip to section 3 if your copy-number loss DOES NOT overlap these types of genes/regions)*			
Overlap with ESTABLISHED HI genes or genomic regions and consideration of reason for referral	**2A**. Complete overlap of an established HI gene/genomic region.	1.00	1.00
	2B. Partial overlap of an established HI genomic region • The observed CNV does NOT contain the known causative gene or critical region for this established HI genomic region OR • Unclear if known causative gene or critical region is affected OR • No specific causative gene or critical region has been established for this HI genomic region	0 (Continue evaluation)	0
	2C. Partial overlap with the 5′ end of an established HI gene (3′ end of the gene not involved)…	See categories below	
	2C-1. …and coding sequence is involved	0.90 *(range: 0.45 to 1.00)*	1.00
	2C-2. …and only the 5′ UTR is involved	0 *(range: 0 to 0.45)*	0.45
	2D. Partial overlap with the 3′ end of an established HI gene (5′ end of the gene not involved)…	See categories below	
	2D-1. …and only the 3′ untranslated region is involved.	0 (Continue evaluation)	0
	2D-2. …and only the last exon is involved. Other established pathogenic variants have been reported in this exon.	0.90 *(range: 0.45 to 0.90)*	0.90
	2D-3. …and only the last exon is involved. No other established pathogenic variants have been reported in this exon.	0.30 *(range: 0 to 0.45)*	0.45
	2D-4. …and it includes other exons in addition to the last exon. Nonsense-mediated decay is expected to occur.	0.90 *(range: 0.45 to 1.00)*	1.00
	2E. Both breakpoints are within the same gene (intragenic CNV; gene-level sequence variant).	See ClinGen SVI working group PVS1 specifications • PVS1 = 0.90 *(Range: 0.45 to 0.90)* • PVS1_Strong = 0.45 *(Range: 0.30 to 0.90)* • PVS1_Moderate *or* PM4 (in-frame indels) = 0.30 *(Range: 0.15 to 0.45)* • PVS1_Supporting = 0.15 *(Range: 0 to 0.30)*	See categories at left

Example: Out-of-frame deletion in BRCA1

- Haploinsufficiency of BRCA1 is associated with Hereditary Breast and Ovarian Cancer (HBOC)
- This deletion is a well-known Dutch founder variant and seen to co-segregate with HBOC in these families (PMID: 9354803, 19894111, 11462239)



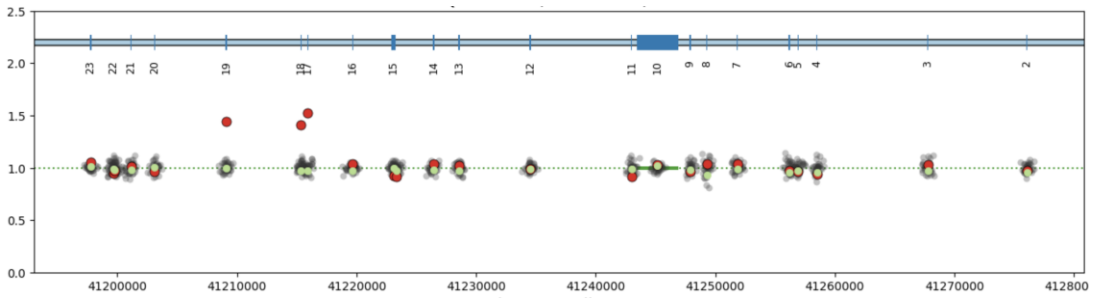
- PVS1 = 0.90
(Range: 0.45 to 0.90)
- PVS1_Strong = 0.45
(Range: 0.30 to 0.90)
- PVS1_Moderate *or* PM4 (in-frame indels) = 0.30
(Range: 0.15 to 0.45)
- PVS1_Supporting = 0.15
(Range: 0 to 0.30)

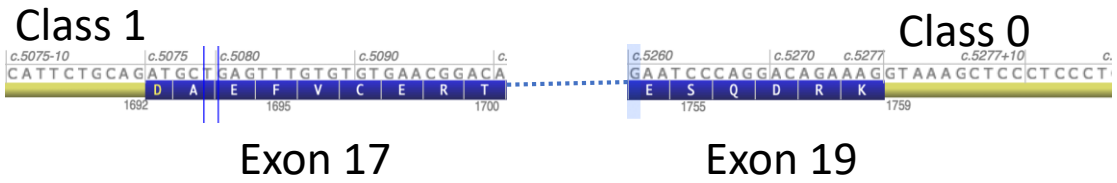
Pathogenic: Score = **1.0**
2E: PVS1 = 0.90 (results in NMD)
4H segregation: additional 0.10

Example: Out-of-frame duplication in BRCA1

- BRCA1 is associated with Hereditary Breast and Ovarian Cancer (HBOC)
- Similar out-of-frame duplications have been reported in individuals with HBOC (PMID: 18330910, 12203994)

Duplication of exons 17-19 of BRCA1

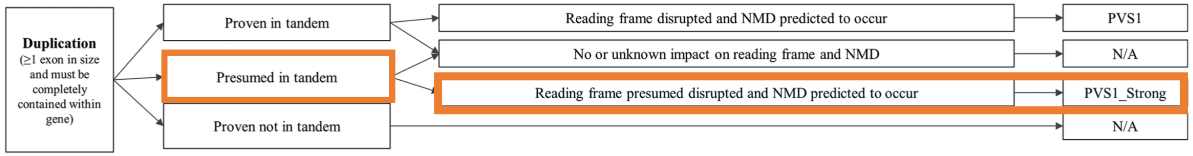




Out-of-frame duplication that results in NMD

Example: Out-of-frame duplication in BRCA1

- BRCA1 is associated with Hereditary Breast and Ovarian Cancer (HBOC)
- Similar out-of-frame duplications have been reported in individuals with HBOC (PMID: 18330910, 12203994)



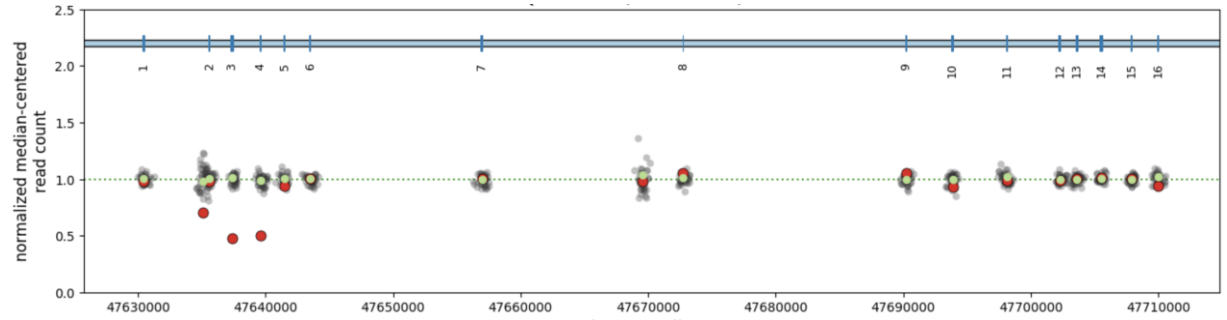
See ClinGen SVI working group PVS1 specifications
- PVS1 = 0.90
(Range: 0.45 to 0.90)
- PVS1_Strong = 0.45
(Range: 0.30 to 0.90)
- N/A = 0 (Continue evaluation)

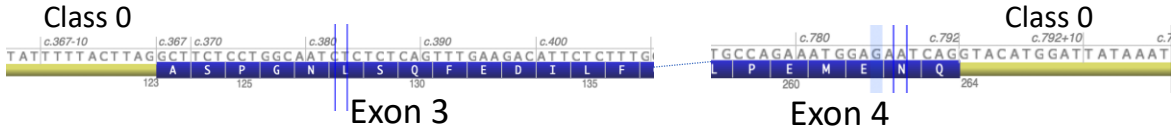
Likely Pathogenic: Score = **0.90**
2E: PVS1_Strong upgraded to 0.90 (most likely in tandem and may result in NMD, similar CNVs in patients with HBOC)

Example: In-frame Deletion of MSH2

- MSH2 is associated with Lynch syndrome (LS)
- In-frame deletion of exons 3 or 4 have been reported as pathogenic in patients with LS along with multiple SVs in those exons
- Deletion disrupts DNA-binding domain needed for MMR
  - Exons are essential for normal gene function
- Observed in at least 3 unrelated patients with LS

Heterozygous deletion of exon 3-4 in *MSH2*

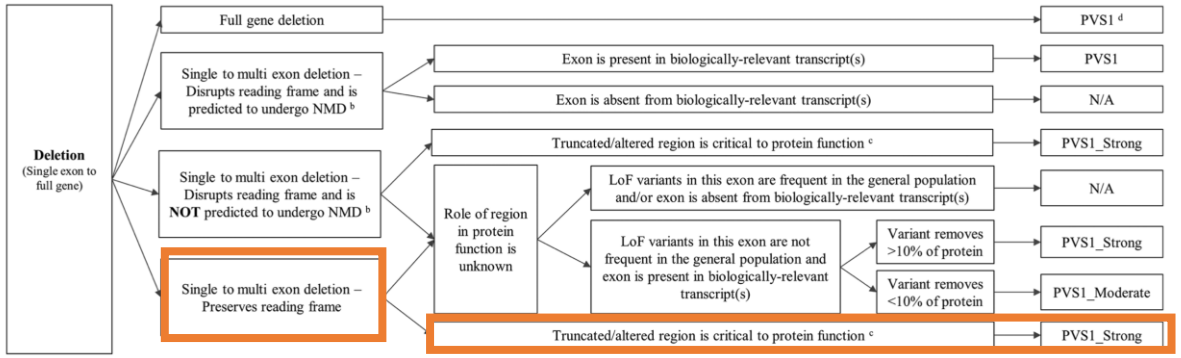




In-frame deletion

Example: In-frame Deletion of MSH2

- *MSH2* is associated with Lynch syndrome (LS)
- In-frame deletion of exons 3 or 4 have been reported as pathogenic in patients with LS along with multiple SVs in those exons
- Deletion disrupts DNA-binding domain needed for MMR (PMID: 18822302, 18383312)
  - Exons are essential for normal gene function
- Observed in at least 3 unrelated patients with LS



- PVS1 = 0.90
(Range: 0.45 to 0.90)
- PVS1_Strong = 0.45
(Range: 0.30 to 0.90)
- PVS1_Moderate *or* PM4 (in-frame indels) = 0.30
(Range: 0.15 to 0.45)
- PVS1_Supporting = 0.15
(Range: 0 to 0.30)

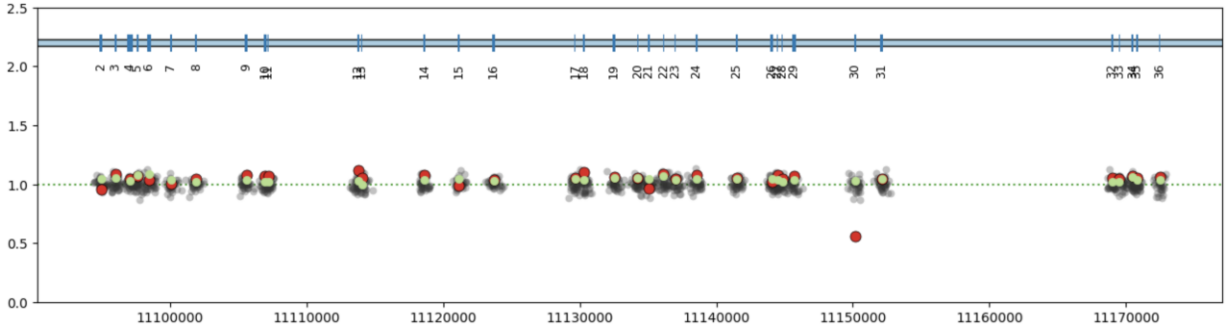
Pathogenic: Score = **1.0**

- **2E:** PVS1_Strong upgraded to 0.90 (deletes key domain)
- **4E case reports:** additional 0.30 (or 0.10 to reach score of 1.0)

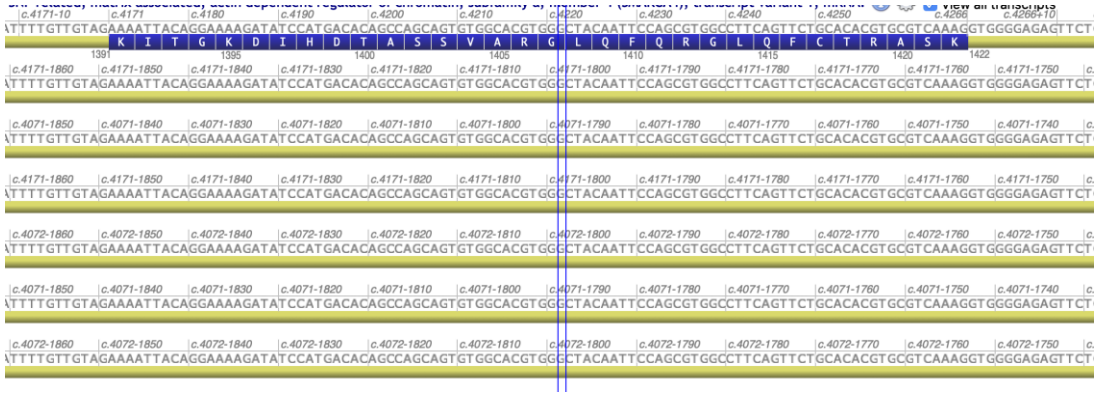
Example In-frame deletion

- SMACA4 HI is associated with hypercalcemic type of small cell carcinoma of the ovaries (SCCOTH)
- Exon 30 is alternatively spliced out of other isoforms of the gene
- Patients carrying this deletion do not appear to be affected with SCCOTH.

Heterozygous deletion of exon 30 in SMARCA4



Exon 30 in NM_001128849



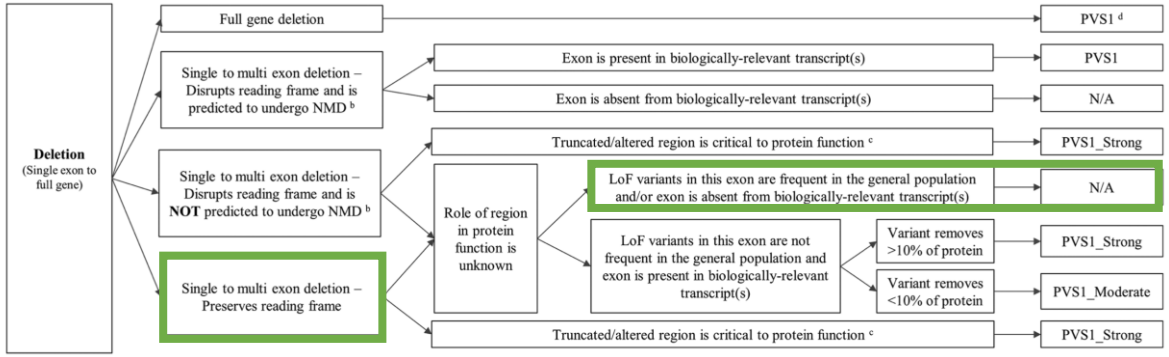
Class 0

Class 0

In-frame deletion of exon 30

Example In-frame deletion

- SMACA4 HI is associated with hypercalcemic type of small cell carcinoma of the ovaries (SCCOTH)
- Exon 30 is alternatively spliced out of other isoforms of the gene
- Patients carrying this deletion do not appear to be affected with SCCOTH.



- PVS1 = 0.90

(Range: 0.45 to 0.90)

- PVS1_Strong = 0.45

(Range: 0.30 to 0.90)

- PVS1_Moderate *or* PM4 (in-frame indels) = 0.30

(Range: 0.15 to 0.45)

- PVS1_Supporting = 0.15

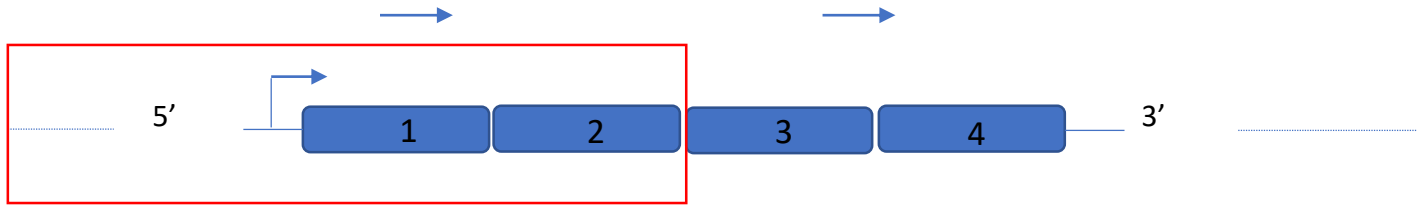
(Range: 0 to 0.30)

Variant of Uncertain Significance: Score = **0**

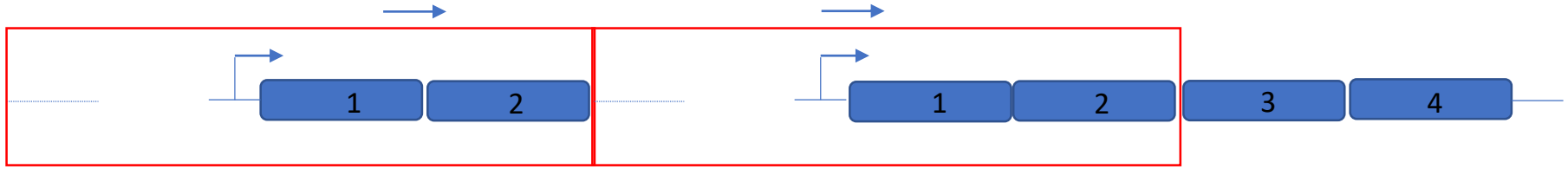
2E: N/A Score = 0.0 (Exon is absent from other biologically relevant transcripts)

Continue with other sections

N- terminal or C-terminal Gene Duplications



Tandem and Direct: Reading frame is reconstituted



2J. One breakpoint is within an established HI gene, patient's phenotype is either inconsistent with what is expected for LOF of that gene OR unknown.	0 (Continue evaluation)	0
2K. One breakpoint is within an established HI gene, patient's phenotype is highly specific and consistent with what is expected for LOF of that gene.	0.45	0.45

Acknowledgements

- Christa Lese Martin
- Daniel Pineda Alvarez
- Swaroop Aradhya
- Erica Andersen
- Athena Cherry
- Sibel Kantarci
- Hutton Kearney
- Ankita Patel
- Gordana Raca
- Deborah Ritter
- Sarah South
- Erik Thorland
- Outside Reviewers (11)
- ClinGen Genomic Variant and Sequence Variant Interpretation Working Groups
- Ronak Patel and ClinGen Baylor team for CNV calculator

ClinGen is primarily funded by the National Human Genome Research Institute (NHGRI), through the following three grants:U41HG006834, U41HG009649, U41HG009650