POSTER H019

Best Practices for the Detection, Interpretation, and Reporting of *In Cis* Variants in AML

Lauren Petersen[1], Rachana Sainger[1], Paulina Gomez[1], Joshua Wemmer[2], Bradley Patay[2], & Jeffrey E. Miller[2] | [1]LabPMM, San Diego, CA, [2]Invivoscribe, San Diego, CA

Background

- The detection, interpretation, and reporting of complex delins variants is a challenge in the field of clinical genomics[1,2,3].
- Bioinformatics pipelines have limitations in detecting *in cis* variants and it is often the responsibility of variant assessment scientists to manually identify and assign their final nomenclature.
- The Human Genome Variation Society (HGVS) Variant Nomenclature Committee recommends that *in cis* variants be described as a single delins when two variants on the same allele are separated by fewer than 2 nucleotides[4]. However, other nearby *in cis* variants may be included in the final classification if that has a significant impact on the overall protein function.
- Defining a workflow that best identifies *in cis* variants, utilizes correct HGVS nomenclature, and results in accurate oncogenic classification is pertinent for clinical diagnostics and reporting for patients with acute myeloid leukemia (AML)[5].

Methods

- We developed a manual workflow for the identification and interpretation of *in cis* variants in AML.
- A total of 41 samples with *in cis* that were previously detected by LabPMM's CAP/CLIA-certified MyAML® and MyMRD® gene panels were used to optimize and test this workflow.
- Our workflow focused on identifying variants with similar VRF, confirming they were on the same allele by viewing the BAM file, and only reporting those that resulted in a new, clinically significant protein change.
- Mutalyzer (https://mutalyzer.nl/) was used to obtain correct HGVS nomenclature. For comparison, the bioinformatics tool VarGrouper[6] was assessed in a subset of samples.

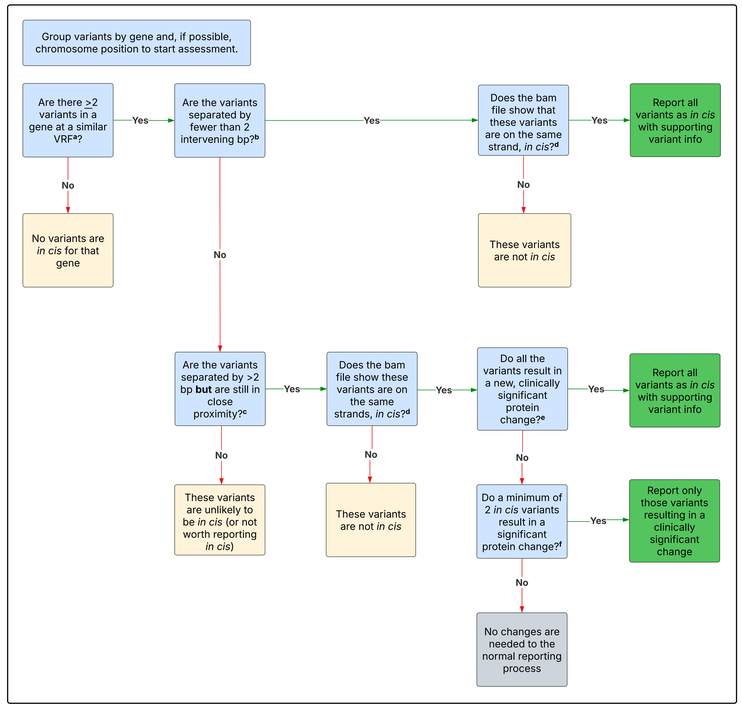


Figure 1. The step-wise process taken for manually identifying and reporting *in cis* variants.
a Similar VRF is defined as <1.0% between variants. If VRF is >1% further analysis is required to determine reportability. It is also recommended to have a lower limit for *in cis* variants to avoid potential artifacts.
b This would strictly be following HGVS nomenclature guidelines, regardless of whether the combination of variants affect two codons separately and/or does not strictly change the protein change.
c Close proximity is defined as <50 bp apart.
d Things to check in the bam file include ensuring there is no strand bias, ensuring that the variants always fall in the same strand, and checking for noise in the reads.
e It is not worth reporting variants as *in cis* if there is no new protein change unless they are separated by fewer than 2 intervening nucleotides (following HGVS guidelines). It is also not worth reporting variants *in cis* if they are more than 2 intervening nucleotides apart AND there is no change in clinical significance; for example, when two VUS missense variants result in a new protein change that would still be classified as a VUS missense variant.
f This would be a case such as when a 3rd variant that is upstream or downstream of other variants *in cis* does not contribute to a protein change. Only include variants that contribute to a clinically significant protein change. One caveat would be in cases where there are 3 variants and the one in the middle does not affect the overall protein change but could still be included in the *in cis* comment for reporting purposes.

Results

Figure 2. Out of 41 total samples, a total of 2 *in cis* variants per gene was the most common observation and one sample with 8 *CBL in cis* variants detected across within a 68 bp region. The average differences in VRF among *in cis* variants was 0.69% (+/- 1.04%). In most cases clinically significant variants *in cis* will be within 5 bp of each other. The average distance between *in cis* variants is 1.9 bp (+/- 1.39).

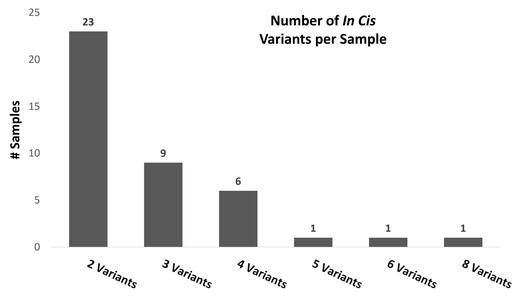
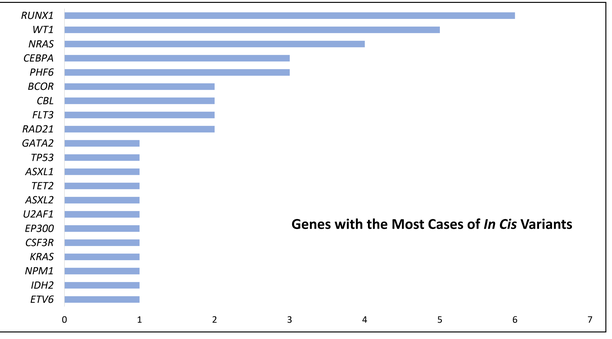


Figure 3. *RUNX1* had the most cases of *in cis* variants (6), followed by *WT1, NRAS, CEBPA*, and *PHF6* with 5, 4, 3, and 3 instances, respectively.

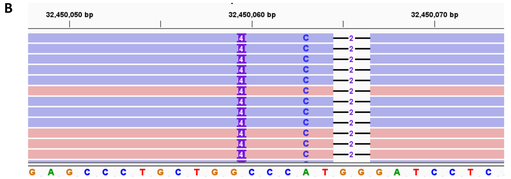


A

Example: How to assess variants *in cis*

Chr	Pos	Ref	Alt	VRF	Gene	Transcript	HGVSc	HGVSp	Classification
11	32417934	.	CAGGCACACG...	9.72	WT1	NM_024426	c.1099-3_1117dup	p.G373Afs*19	LabPMM: Oncogenic
11	32450059	.	CCAC	12.18	WT1	NM_024426	c.752_753insGTGG	p.Q252Wfs*8	LabPMM: Oncogenic
11	32450063	A	C	13.08	WT1	NM_024426	c.749T>G	p.M250R	LabPMM: VUS
11	32450065	GG	.	12.61	WT1	NM_024426	c.746_747del	p.P249Hfs*9	LabPMM: Oncogenic
11	32456619	GCC	.	1.36	WT1	NM_024426	c.271_273delGGC	p.G91del	LabPMM: VUS

B



C

Alteration of all 3 variants:
- Ref (5'→3'): CCC.ATG.GGC.CAG.CAG (Pro.Met.Gly.Gln.Gln)
  - Delete CC, substitute T>G, and insert GTGG where the underlining of the sequence above ends
- Alt (5'→3'): CAG.GGG.GTG.GCC (Gln.Gly.Val.Ala)

New alteration with all 3 variants:
- Mutalyzer input:
  - NM_024426:c.746_752delinsAGGGGGTGG
- **Overall protein change is p.P249Qfs*38**

D

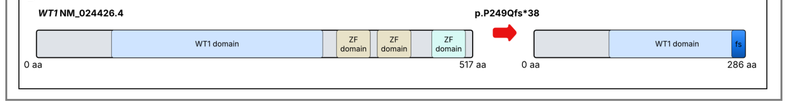


Figure 4. How to assess a sample with three *WT1* variants *in cis*. **A)** The variants were sorted by chromosome position and observed at similar VRF and in close proximity to each other. **B)** The bam file confirmed all three variants were on the same allele. **C)** The combination of all three variants results in a new, clinically significant protein change. **D)** The frameshift variant interrupts the WT1 domain and results in the loss of all ZF domains.

Results

Table 1. We chose a subset of six clinical samples (S1-S6) with *in cis* variants to analyze and test with VarGrouper.
- Initially, all variants in these six samples were reportable individually, as well in a separate comment describing the overall protein effect with two exceptions; In S3, where the combination of all *FLT3* variants did not result in a new protein change, and in S6 the c.992_997del variant is not reported because it is downstream of the stop codon created by c.989T>G and c.990C>A (see figure 5).

	Gene	Transcript	Chr	Pos	HGVSc	HGVSp	VRF (%)	Ontology	Classification
S1 ✓	***ASXL2***	NM_018263	2	25972883	c.1455_1458del	p.N485Kfs*8	2.93	Frameshift	Oncogenic
	ASXL2	NM_018263	2	25972895	c.1434_1446del	p.T479Ifs*11	2.93	Frameshift	Oncogenic
				***ASXL2 in cis* variant:**	**c.1434_1456delinsAATTAT**	**p.T479Ifs*4**		**Frameshift**	**Oncogenic**
S2 ✓	***CBL***	NM_005188	11	11914882	c.1102_1111del	p.Y368Tfs*21	7.75	Frameshift	Oncogenic
	CBL	NM_005188	11	11914895	c.1115del	p.C372Lfs*20	7.01	Frameshift	Oncogenic
	CBL	NM_005188	11	11914897	c.1117G>C	p.E373Q	8.17	Missense	VUS
	CBL	NM_005188	11	11914900	c.1120A>C	p.M374L	7.24	Missense	VUS
	CBL	NM_005188	11	11914901	c.1121T>G	p.M374R	7.49	Missense	VUS
	CBL	NM_005188	11	11914904	c.1124_1163del	p.G375Vfs*4	7.45	Frameshift	Oncogenic
	CBL	NM_005188	11	11914947	c.1167_1168insA	p.D390Rfs*5	7.41	Frameshift	Oncogenic
	CBL	NM_005188	11	11914949	c.1169_1170insAC	p.D390Efs*3	7.47	Frameshift	Oncogenic
				***CBL in cis* variant:**	**c.1101_1170delins21**	**p.Y368_D390delinsTSAGKRT**		**Inframe delins**	**VUS**
S3 x	***FLT3***	NM_004119	13	28592639	c.2504_2506delATA	p.D835_I836delinsV	0.38	Inframe delins	Oncogenic
	FLT3	NM_004119	13	28592637	c.2508C>G	p.I836M	0.4	Missense	Likely Oncogenic
	FLT3	NM_004119	13	28592634	c.2510_2511ins15	p.M837delinsIPTMLS	0.38	Inframe delins	Oncogenic
				***FLT3 in cis* variant:**	**c.2504_2508delins5**	**p.D835_I836delinsV**		**Inframe delins**	**Oncogenic**
S4 ✓	***NPM1***	NM_002520	5	170837545	c.861_862insAAT	p.L287_W288insN	42.2	Inframe insertion	VUS
	NPM1	NM_002520	5	170837547	c.863G>T	p.W288L	40.75	Missense	VUS
	NPM1	NM_002520	5	170837550	c.866delA	p.Q289Rfs*11	42.61	Frameshift	Likely Oncogenic
	NPM1	NM_002520	5	170837552	c.868_869insCTTC	p.W290Sfs*10	43.36	Frameshift	Oncogenic
	NPM1	NM_002520	5	170837554	c.870G>A	p.W290*	40.34	Nonsense	Oncogenic
	NPM1	NM_002520	5	170837558	c.874_875delAA	p.K292Vfs*6	39.98	Frameshift	VUS
				***NPM1 in cis* variant:**	**c.862_875delins18**	**p.W288Nfs*12**		**Frameshift**	**Oncogenic**
S5 ✓	***RUNX1***	NM_001754	21	36259162	c.329A>G	p.K110R	32.08	Missense	Oncogenic
	RUNX1	NM_001754	21	36259161	c.330G>C	p.K110N	32.26	Missense	Oncogenic
	RUNX1	NM_001754	21	36259160	c.331A>C	p.T111P	31.82	Missense	VUS
				***RUNX1 in cis* variant:**	**c.329_330delinsGC**	**p.K110S**		**Missense**	**VUS**
S6 ✓	***RUNX1***	NM_001754	21	36164886	c.989T>G	p.F330C	2.53	Missense	VUS
	RUNX1	NM_001754	21	36164885	c.990C>A	p.F330L	2.46	Missense	VUS
	RUNX1	NM_001754	21	36164878	c.992_997del	p.S331_P333delinsT	2.37	Inframe delins	VUS
				***RUNX1 in cis* variant:**	**c.989_997delinsGAA**	**p.F330***		**Nonsense**	**Oncogenic**

- VarGrouper is a tool in which many options are available for detection of *in cis* variants such as lower and upper thresholds for VRF and distance between variants. All variants were detected after iterative testing with varying thresholds, with the exception of *FLT3* p.D835_I836delinsV in S3, which was never detected by VarGrouper. ✓ indicates VarGrouper accurately detected the *in cis* variant and the **X** indicates it did not.

Alteration of the 2 *RUNX1* variants c.989T>G and c.990C>A, that result in a clinically significant change (c.992_997del has no impact as shown below):
- Ref (5'→3'): GCG.TTC.AGC.GAC.CGG (Ala.Phe.Ser.Asp.Pro)
  - Substitute T>G, substitute C>A, delete GCGACC
- Alt (5'→3'): GCG.TGA.AGG (Ala.STOP.Arg)

New alteration with 2 of the 3 variants:
- Mutalyzer input:
  - NM_001754:c.989_990delinsGA
- **Overall protein change is p.F330***

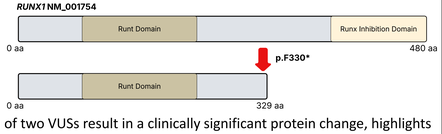


Figure 5. The case of S6, where the combination of two VUSs result in a clinically significant protein change, highlights the necessity of accurately detecting and reporting these variants.

Conclusions & References

- Manual evaluation of *in cis* variants is time-intensive and could lead to human error. Our stepwise process provides an efficient workflow for their detection, interpretation and accurate classification.

1. Lo YC, Narayan R, Nardi V, Lennerz JK. Two In Cis Variants-Two Worlds Apart. Oncologist. 2021 Dec;26(12):997-999. doi: 10.1002/onco.13946. Epub 2021 Sep 2. PMID: 34407276; PMCID: PMC8649014.

2. Wang Q, Pierce-Hoffman E, Cummings BB, Alföldi J, Francioli LC, Gauthier LD, Hill AJ, O'Donnell-Luria AH; Genome Aggregation Database Production Team; Genome Aggregation Database Consortium; Karczewski KJ, MacArthur DG. Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. Nat Commun. 2020 May 27;11(1):2539. doi: 10.1038/s41467-019-12438-5. PMID: 32461613; PMCID: PMC7253413.

3. Degalez F, Jehl F, Muret K, Bernard M, Lecerf F, Lagoutte L, Désert C, Pitel F, Klopp C, Lagarrigue S. Watch Out for a Second SNP: Focus on Multi-Nucleotide Variants in Coding Regions and Rescued Stop-Gained. Front Genet. 2021 Jul 7;12:659287. doi: 10.3389/fgene.2021.659287. PMID: 34306009; PMCID: PMC8293744.

4. Hart RK, Fokkema IFAC, DiStefano M, Hastings R, Laros JFJ, Taylor R, Wagner AH, den Dunnen JT. HGVS Nomenclature 2024: improvements to community engagement, usability, and computability. Genome Med. 2024 Dec 20;16(1):149. doi: 10.1186/s13073-024-01421-5. PMID: 39702242; PMCID: PMC11660784.

5. Döhner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H, Ebert BL, Fenaux P, Godley LA, Hasserjian RP, Larson RA, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood. 2022 Sep 22;140(12):1345-1377. doi: 10.1182/blood.2022016867. PMID: 35797463.

6. Schmidt RJ, Macleay A, Le LP. VarGrouper: A Bioinformatic Tool for Local Haplotyping of Deletion-Insertion Variants from Next-Generation Sequencing Data after Variant Calling. J Mol Diagn. 2019 May;21(3):384-389. doi: 10.1016/j.jmoldx.2018.12.003. Epub 2019 Feb 20. PMID: 30794986.