Ferrata Storti Foundation

How I curate: applying American Society of Hematology-Clinical Genome Resource Myeloid Malignancy Variant Curation Expert Panel rules for *RUNX1* variant curation for germline predisposition to myeloid malignancies

Haematologica 2020
Volume 105(4):870-887

David Wu,[1,*] Xi Luo,[2*] Simone Feurstein,[3] Chimene Kesserwan,[4] Shruthi Mohan,[5] Daniel E. Pineda-Alvarez,[6] and Lucy A. Godley[3,7] on behalf of the collaborative group of the American Society of Hematology - Clinical Genome Resource Myeloid Malignancy Variant Curation Expert Panel[#]

[1]Department of Laboratory Medicine, University of Washington, Seattle, WA; [2]Department of Pediatrics/Hematology-Oncology, Baylor College of Medicine, Houston, TX; [3]Section of Hematology/Oncology, Department of Medicine, and The University of Chicago Comprehensive Cancer Center, Chicago, IL; [4]Albert Einstein College of Medicine, Department of Pathology, New York, NY; [5]Department of Genetics, University of North Carolina School of Medicine, Chapel Hill, NC; [6]Invitae, San Francisco, CA and [7]Department of Human Genetics, The University of Chicago, Chicago, IL, USA

[*]*DW and XL contributed equally to this work.*

[#]*Members of this Clinical Genome Resource Variant Curation Expert Panel are listed in the Acknowledgments.*

ABSTRACT

The broad use of next-generation sequencing and microarray platforms in research and clinical laboratories has led to an increasing appreciation of the role of germline mutations in genes involved in hematopoiesis and lineage differentiation that contribute to myeloid neoplasms. Despite implementation of the American College of Medical Genetics and Genomics and Association for Molecular Pathology 2015 guidelines for sequence variant interpretation, the number of variants deposited in ClinVar, a genomic repository of genotype and phenotype data, and classified as having uncertain significance or being discordantly classified among clinical laboratories remains elevated and contributes to indeterminate or inconsistent patient care. In 2018, the American Society of Hematology and the Clinical Genome Resource co-sponsored the Myeloid Malignancy Variant Curation Expert Panel to develop rules for classifying gene variants associated with germline predisposition to myeloid neoplasia. Herein, we demonstrate application of our rules developed for the *RUNX1* gene to variants in six examples to show how we would classify them within the proposed framework.

Correspondence:

DAVID WU
dwu2@uw.edu

LUCY A. GODLEY
lgodley@medicine.bsd.uchicago.edu

Received: July 16, 2019.

Accepted: October 21, 2019.

Pre-published: March 12, 2020

doi:10.3324/haematol.2018.214221

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/105/4/870

©2020 Ferrata Storti Foundation

Material published in Haematologica is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode. Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.

Introduction

Germline mutations in genes involved in hematopoiesis and lineage differentiation predispose patients to myeloid neoplasia with or without thrombocytopenia. The broad adoption of next-generation sequencing and microarrays in the clinical laboratory has expanded our knowledge of germline contribution to myeloid neoplasia. Drazer *et al.* reported that in six of 24 patients with myeloid neoplasia, presumed somatic variants in *DDX41*, *GATA2* and *TP53* were of germline origin.[1] Similarly, Churpek *et al.* showed that 29% of acute myeloid leukemia (AML)/myelodysplastic syndrome (MDS) kindreds with a positive family history carried a variant in one of 12 genes associated with germline predisposition to hematopoietic malignancies, including *FANCA*, *GATA2*, *RUNX1*, and *SBDS*.[2] To date, more than 65 genes have been associated with a predisposition to hematologic malignancies.[3] Recognizing the contribution of germline variation toward myeloid neoplasia, the 'WHO classification of Tumors of Hematopoietic and

Lymphoid Tissues' incorporated the classification of myeloid neoplasia with germline predisposition in their 2016 revised edition.[4,5]

In parallel, clinical laboratories are increasingly offering broad next-generation sequencing-based tests for patients with myeloid neoplasia for somatic testing, and will readily detect germline variants, if present in a patient. While there is increased clinical awareness of the potential for these germline variants to contribute to a patient's disease, there are often insufficient data in the literature to definitively classify whether a detected variant is contributing to the patient's phenotype.[6,7] For example, familial platelet disorder with predisposition to AML (FPD/AML) is an autosomal dominant disorder in which germline mutations in *RUNX1* result in thrombocytopenia, platelet functional and/or ultrastructural defects, and/or susceptibility to hematologic malignancies commonly including MDS, AML, and other malignancies[8-11] (Table 1). ClinVar (*https://www.ncbi.nlm.nih.gov/clinvar/*) is a database repository of clinically actionable genomic variants[12,13] that currently lists 325 germline *RUNX1* variants deposited by clinical laboratories. More than half of these variants are currently reported as being of uncertain significance.

Worldwide, most clinical laboratories follow the 2015 American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) guidelines for sequence variant interpretation.[14] In this framework, germline variants are classified using a five-tier system: benign (BEN), likely benign (LBEN), variant of uncertain significance (VUS), likely pathogenic (LPATH) and pathogenic (PATH). During sequence variant interpretation, laboratories systematically review the supporting criteria of a genomic variant, such as: minor allele frequencies (MAF), computational predictions, functional experiments and segregation with disease in order to determine the five-tier classification.[14-16]

Although the ACMG/AMP guidelines provide a comprehensive framework for sequence variant interpretation, the high rate of VUS and curation discrepancies continue to be an impediment to accurate clinical annotation and interpretation of genomic variants.[6,7] To encourage genomic and phenotypic data sharing, and engage experts in consensus-driven variant interpretation, the Clinical Genome Resource (ClinGen) convened Variant Curation Expert Panels (VCEP) to develop gene- and disease-specific modifications of the original guidelines and provide expert-reviewed variant classification for depositing into ClinVar (*Online Supplementary Figure S1*).[17] In 2018, the American Society of Hematology (ASH) sponsored a ClinGen Myeloid Malignancy Variant Curation Expert Panel (MM-VCEP), composed of 34 international members, who started working on gene- and disease-specific rules for *RUNX1* as the first of several genes conferring predisposition to myeloid malignancies (*Online Supplementary Figure S1A*). After designing, modifying and testing the preliminary *RUNX1* rules on 52 pilot variants, which improved classification in 33% VUS or variants with conflicting interpretations (CONF), MM-VCEP-specified ACMG/AMP rules were approved by the ClinGen oversight committee and efforts to curate variants to ClinVar using the Variant Curation Interface have commenced (*Online Supplementary Figure S1B*).[18] This pilot effort resulted in one variant being upgraded to PATH, two variants being upgraded to LPATH, and three variants being downgraded to LBEN. ClinGen's website contains the MM-VCEP variant classification recommendations and any subsequent modifications to these codes over time (*https://www.clinicalgenome.org/affiliation/10034/*).

Herein, we demonstrate the application of *RUNX1*-specific rules (Table 2) to classify nine representative *RUNX1* variants in six examples (Table 3) while reviewing phenotypic criteria for FPD/AML and summarizing molecular and functional roles of *RUNX1*.

Example 1. Early nonsense variants, (p.Arg204Ter) (PATH with PVS1, PM2, PS4_supporting, and PP1)

A 50-year old female with new pancytopenia was referred to a hematology service. A bone marrow biopsy showed hypocellularity with severe trilineage dysplasia and 12% blasts, diagnostic of MDS with excess blasts (MDS-EB-2). Further investigation showed pathogenic variants in *RUNX1* (NM_001754:c.610C>T, (p.Arg204Ter)), *BCOR* and *ASXL1* with a normal karyotype. The medical history was positive for thrombocytopenia (baseline 70-120x10^9/L) and a propensity to excessive bleeding after tooth extractions. The family history was positive for two sons with persistent thrombocytopenia (baseline 50-100x10^9/L) not otherwise explained and a granddaughter with thrombocytopenia and MDS with monosomy 7 (Figure 1). During the initial assessment, an increase in lactate dehydrogenase and the peripheral blast count were noted. A second marrow biopsy confirmed transformation into AML with 40% blasts. The patient underwent induction chemotherapy without achieving remission and clofarabine bridging for unrelated stem cell transplantation. During conditioning, the patient developed sepsis with Gram-negative bacteria and died shortly afterwards. Since she had a remarkable personal and family history pointing towards a germline predisposition syndrome, a skin biopsy was performed at the time of the diagnosis of MDS, and DNA testing from cultured skin

Table 1. Clinical phenotypes of RUNX1 familial platelet disorder and hereditary malignancies.

Clinical and laboratory features	Details	Life-time risk
Hematologic malignancy	Commonly AML or MDS; less frequently T-ALL; and rarely mixed MPN/MDS such as CMML, as well as B-ALL, and hairy cell leukemia	~44%
Thrombocytopenia	Mild to moderate thrombocytopenia with normal platelet size, in the absence of other causes	Most patients
Platelet functional and/or ultrastructural defects	Includes impaired platelet aggregation (particularly in response to collagen and epinephrine) and platelet alpha or dense granule secretion defects	Not known

Adapted from Table 2 from Luo and Feurstein, *et al.*[18] AML: acute myeloid leukemia; ALL: acute lymphoblastic leukemia; MPN: myeloproliferative neoplasms; MDS: myelodysplastic syndrome, CMML: chronic myelomonocytic leukemia.

Table 2. Clinical Genome Resource Myeloid Malignancy Variant Curation Expert Panel-approved rules for *RUNX1* variant interpretation.

ACMG/ AMP CC	Original ACMG/AMP rule summary	Specification	Stand alone	Very strong	Strong	Moderate	Supporting	Comments
BA1	Allele frequency is >5% in ESP, 1000G, or ExAC.	Disease-specific	MAF ≥ 0.0015 (0.15%)	na	na	na	na	(1) The variant is present in any general continental population dataset with a minimum number of 2,000 alleles and variant present in ≥5 alleles.
BS1	Allele frequency is greater than expected for disorder.	Disease-specific	na	na	MAF between 0.00015 (0.015%) and 0.0015 (0.15%)	na	na	(1) The variant is present in any general continental population dataset with a minimum number of 2,000 alleles and variant present in ≥5 alleles. (2) Variant can be classified as likely benign based on BS1 alone if there is no contradictory evidence supporting pathogenicity.
BS2	Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age.	na						FPD/AML patients display incomplete penetrance and the average age of onset of hematologic malignancies is 33 years.
BS3	Well-established *in vitro* or *in vivo* functional studies show no damaging effect on protein function or splicing.	Gene-specific, strength	na	na	(1) Transactivation assays demonstrating normal transactivation (80-115% of wt) AND (2) data from a secondary assay demonstrating normal function.	na	Transactivation assays demonstrating normal transactivation (80-115% of wt).	See PS3 (1) and (2)
BS4	Lack of segregation in affected members of a family.	General rec	na	na	Applied when seen in ≥2 informative meioses.	na	na	This code should only be applied for genotype-positive, phenotype-negative (with sufficient laboratory evidence) family members.
BP1	Missense variant in a gene for which primarily truncating variants are known to cause disease.	na						FPD/AML is caused by both pathogenic missense and truncating variants.
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.	General rec	na	na	na	na	Per original ACMG/AMP guidelines.	BP2 can also be applied if the variant is detected in a homozygous state.
BP3	In-frame deletions/insertions in a repetitive region without a known function.	na						*RUNX1* does not contain a repetitive region without known function.
BP4	Multiple lines of computational evidence suggest no impact on gene or gene product.	General rec	na	na	na	na	Per original ACMG/AMP guidelines.	BP4 should be applied for missense variants if all of the following apply: (1) REVEL score <0.15, (2) SSF and MES predict either an increase in the canonical splice

continued on the next page

continued from the previous page

ACMG/ AMP CC	Original ACMG/AMP rule summary	Specification	Stand alone	Very strong	Strong	Moderate	Supporting	Comments
								site score or a decrease of the canonical splice site score by no more than 10%, and (3) no putative cryptic splice sites are created. BP4 should also be applied for synonymous, intronic and non-coding variants for which SSF and MES predict either an increase in the canonical splice site score or a decrease of the canonical splice site score by no more than 10% and no putative cryptic splice sites are created.
BP5	Variant found in a case with an alternate molecular basis for disease.	na						In rare circumstances, a patient can carry two variants in genes predisposing to hematologic malignancies.
BP6	Reputable source recently reported variants as benign, but data are not available for laboratories to perform independent evaluations.	na						According to SVI recommendations.
BP7	A synonymous 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.	General rec	na	na	na	na	Per original ACMG/AMP guidelines. BP7 cannot be applied in combination with PP3.	Also applicable to intronic/non-coding variants at or beyond positions +7/-21 for which (1) SSF and MES predict either an increase in the canonical splice site score or a decrease of the canonical splice site score by no more than 10% and no putative cryptic splice sites are created AND (2) evolutionary conservation prediction algorithms predict the site as not conserved (e.g. PhyloP score <0.1 or the variant is the reference nucleotide in one primate and/or three mammal species.).
PVS1	Null variant in a gene for which LOF is a known mechanism of disease.	Gene-specific, strength	na		Per modified *RUNX1* PVS1 decision tree for SNV, indels and CNV and table of splicing effects.		na	*RUNX1* LOF variants are a common mechanism of disease in FPD/AML. Three major isoforms (A, B, C) are expressed by use of two promotors and alternative splicing. C-terminal variants not predicted to undergo NMD are classified as PVS1_strong, deletions of exons 2 and 3, presumably only affecting RUNX1 isoform 1C, meet PVS1_moderate.
PS1	Same AA change as a previously established pathogenic variant regardless of nucleotide change.	Strength	na	na	Same AA change as a previously established pathogenic variant regardless of nucleotide change.	Same AA change as a previously established likely pathogenic variant regardless of nucleotide change.	na	(1) RNA data or agreement in splicing predictors show no splicing effects (SSF and MES predict either increase in canonical splice site score or decrease of canonical splice score by no more than 10% and no putative splice site are created). (2) The previously established PATH/LPATH variant must be asserted pathogenic/likely pathogenic based on MM-VCEP rules for *RUNX1* before this rule can be applied.
PS2	De novo (maternity and paternity confirmed) in a patient with the disease and no family history.	Disease-specific, strength	na	na	na	≥ 2 proven *de novo* occurrences (maternity	1 proven *de novo* occurrence (maternity	(1) No family history is defined as: absence of the variant and any of the *RUNX1*-phenotypic criteria in first- and/or second-degree relatives. (2) The proband must exhibit at least one phenotypic FPD/AML criterion. (3) The

continued on the next page

continued from the previous page

ACMG/ AMP CC	Original ACMG/AMP rule summary	Specification	Stand alone	Very strong	Strong	Moderate	Supporting	Comments
						and paternity confirmed) in patients with the *RUNX1*-phenotype.	and paternity confirmed) in a patient with the *RUNX1*-phenotype.	maximum allowable strength by combining PS2 and PM6 criteria is to apply one moderate or two supporting rules.
PS3	Well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect.	Gene-specific, strength	na	na	Transactivation assays demonstrating altered transactivation (<20% of wt, and/or reduced to levels similar to well-established pathogenic variants such as Arg201Gln or Arg166Gln) AND data from a secondary assay demonstrating altered function. PS3 cannot be applied if the variant meets PVS1. If the variant meets criteria for PVS1_strong and PS3, we recommend either applying PVS1_strong and PS3_moderate or upgrading PVS1_strong to PVS1 without applying PS3.	Transactivation assays demonstrating altered transactivation (<20% of wt, and/or reduced to levels similar to well-established pathogenic variants such as Arg201Gln or Arg166Gln) OR ≥2 secondary assays demonstrating altered function.	Transactivation assays demonstrating enhanced transactivation (>115% of wt).	(1) Transactivation assays should include wt and known pathogenic controls, as well as co-expression with *CBFβ*. Promoter sequences of *CSF1R* (M-CSF-R), *PF4*, *C-FMS* and *GZMB*, containing consensus *RUNX1* binding sites have been used for transactivation assays. (2) The following secondary assays have been performed: EMSA and yeast hybrid assays (decreased DNA-binding affinity), co-IP, FRET and affinity assays (diminished heterodimerization ability with CBFβ), IF and WB with cell fractionation (abnormal cellular localization), colony-forming assays (reduced colony-forming potential), xenotransplantation experiments (abnormal function of mutant RUNX1 *in vivo)*. (3) PS3 can also be applied for evidence of very low or abnormal mRNA/protein expression of the variant allele as a functional consequence of a null variant or incorrect mRNA/protein products.
PS4	The prevalence of the variant in affected individuals is significantly increased compared to the prevalence in controls.	Disease-specific, strength	na	na	≥ 4 probands meeting *RUNX1*-phenotypic criteria.	2-3 probands meeting *RUNX1*-phenotypic criteria.	1 proband meeting *RUNX1*-phenotypic criteria.	The affected individual has to fit at least one of the *RUNX*1-phenotypic criteria AND variant has to be either absent from gnomAD (overall population) or only present once.
PM1	Located in a mutational hot spot and/or critical and well-established functional domain without benign variation.	Gene-specific, strength	na	na	na	Variant affecting one of the following 13 hotspot residues: Arg107, Lys110, Ala134, Arg162, Arg166, Ser167, Arg169, Gly170, Lys194, Thr196, Asp198, Arg201,	Variant affecting one of the other AA residues 105-204 within the RHD.	The RHD (AA 77-204) has been established to be a highly conserved DNA-binding domain without any benign variation in ClinVar. No germline pathogenic variants have been reported in residues in the region (AA 77-104) to date. The AA range under PM1_supporting may be expanded in the future to other parts of the protein if more evidence emerges.

continued on the next page

continued from the previous page

ACMG/ AMP CC	Original ACMG/AMP rule summary	Specification	Stand alone	Very strong	Strong	Moderate	Supporting	Comments
						Arg204.		
PM2	Absent from controls.	General rec	na	na	na	Per original ACMG/AMP guidelines.	na	Variant must be completely absent from all population databases. The mean coverage of *RUNX1* in the population database used should be at least 20x.
PM3	For recessive disorders, detected *in trans* with a pathogenic variant.	na						FPD/AML is inherited in an autosomal dominant manner.
PM4	Protein length changes due to in-frame deletions/insertions in a non-repeat region or stop-loss variants.	Gene-specific, strength	na	na	na	In-frame deletion/insertion impacting at least one of the 13 hotspot residues Arg107, Lys110, Ala134, Arg162, Arg166, Ser167, Arg169, Gly170, Lys194, Thr196, Asp198, Arg201, Arg204	Other in-frame deletion/ insertion impacting residues 105-204 within the RHD.	see PM1
PM5	Missense change at AA residue where a different missense change determined to be pathogenic has been seen before.	Strength	na	na	Missense change at the same residue where ≥2 different missense changes have previously been determined to be pathogenic. PM5_strong cannot be applied together with PM1.	Missense change at the same residue where a different missense change has previously been determined to be pathogenic.	Missense change at the same residue where a different missense change has previously been determined to be likely pathogenic.	see PS1
PM6	Assumed *de novo* (but without confirmation of maternity and paternity) in a patient with the disease and no family history.	Disease-specific, strength	na	na	na	≥4 assumed *de novo* occurrences (without confirmation of maternity and paternity) in patients with the *RUNX1*-	2 or 3 assumed *de novo* occurrences (without confirmation of maternity and paternity) in patients with the *RUNX1*-	see PS2

continued on the next page

continued from the previous page

ACMG/ AMP CC	Original ACMG/AMP rule summary	Specification	Stand alone	Very strong	Strong	Moderate	Supporting	Comments
						phenotype.	phenotype.	
PP1	Co-segregation with disease in multiple affected family members.	Disease-specific, strength	na	na	≥ 7 meioses observed within one or across multiple families.	5 or 6 meioses observed within one or across multiple families.	3 or 4 meioses observed within one or across multiple families.	(1) Affected individuals show at least one of the *RUNX1*-specific phenotypic criteria. (2) Only genotype and phenotype positive individuals and obligate carriers are counted. (3) Demonstration of co-segregation in multiple families is not required since many *RUNX1* variants are unique and only occur in one family.
PP2	Missense variant in a gene that has a low rate of benign missense variation and where missense variants are a common mechanism of disease.	na						Missense constraint z-score for *RUNX1* is <3.09.
PP3	Multiple lines of computational evidence support a deleterious effect on the gene or gene product.	General rec	na	na	na	na	Per original ACMG/AMP guidelines.	(1) PP3 should be applied for missense variants with a REVEL score of >0.75 (2) PP3 should be applied for missense or synonymous variants if the variant alters the last three bases of an exon preceding a donor splice site or the first three bases of an exon following a splice acceptor site and the predicted decrease in the score of the canonical splice site (measured by both MES and SSF) is at least 75% regardless of the predicted creation/presence of a putative cryptic splice site. (3) PP3 should also be applied for intronic variants (in introns 4-8) located in reference to exons at positions +3 to +5 for splice donor sites or -3 to -5 for splice acceptor sites for which the predicted decrease in the score is at least 75% (measured by both MES and SSF) regardless of the predicted creation/presence of a putative cryptic splice site. (4) PP3 cannot be applied for canonical splice site variants.
PP4	Patient's phenotype or family history is highly specific for a disease with a single genetic etiology.	na						FPD/AML does not exhibit a highly specific phenotype and there is substantial genetic heterogeneity.
PP5	Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent analysis.	na						According to SVI recommendations.

Adapted from Table 1 of Luo and Feurstein *et al.*[18] ACMG: American College of Medical Genetics; AMP: Association for Molecular Pathology; CC: criteria code; , ESP: Exome Sequencing Project; 1000G: 1000 Genome project; ExAc: Exome Aggregation Consortium; MAF: minor allele frequency; na: not applicable; FPD/AML: familial platelet disorder with predisposition to acute myeloid leukemia; rec: recommendation; SSF: splice site finder; MES: MaxEntScan; SVI: ClinGen Sequence Variant Interpretation Working Group; LOF: loss-of-function; SNV: single-nucleotide variant; CNV: copy number variant; NMD: nonsense-mediated decay; AA: amino acid; MM-VCEP: Myeloid Malignancy Variant Curation Expert Panel; wt: wildtype; EMSA: electrophoretic mobility shift assay; IP: immunoprecipitation; FRET: fluorescence resonance energy transfer; IF: immunofluorescence; WB: western blot; gnomAD: Genome Aggregation Database; RHD: Runt homology domain.

fibroblasts later confirmed the germline, nonsense *RUNX1* variant. Her two sons and granddaughter also tested positive for the *RUNX1* variant (Figure 1).

Similar to this case, most patients with FPD/AML will have a characteristic phenotype (Table 1) including mild to moderate thrombocytopenia with normal platelet size, platelet α or dense granule secretion defects and impaired platelet aggregation, particularly in response to collagen and epinephrine as well as a predisposition to hematologic malignancies. Although there is variability in disease onset in FPD/AML,[3] development of a hematologic malignancy is common with a lifetime risk of ~44%: AML and MDS are common, other malignancies occur less frequently (Table 1).[19-23] FPD/AML has a high but incomplete

Table 3. Summary of *RUNX1* variant examples with application of the Myeloid Malignancy Variant Curation Expert Panel criteria.

Example No.	*RUNX1* variant NM_001754 (isoform C)	ClinVar Assertion	Criteria	MM-VCEP classification
1	c.610C>T (p.Arg204Ter)	PATH	PVS1, PM2, PS4_supporting, PP1	PATH
2	c.314A>C p.(His105Pro)	VUS	PM2, PP3, PS4_supporting, PM1_supporting, PM5_supporting	LPATH
	c.315C>A p.(His105Gln)		PS3, PM2, PP3, PM1_supporting	LPATH
3	c.253C>A p.(His85Asn)	CONF: OMIM: PATH Invitae: VUS	BS1, BS3, PP3	LBEN
4	c.508+3delA	PATH	PS3, PP1_strong, PM2, PP3, PS4_supporting	PATH
	c.444C>T p.(Thr148=)	Illumina: VUS Invitae: LBEN	BP4, BP7	LBEN
	c.1257G>A p.(Val419=)	VUS		VUS
5	Copy number variant, deletion of exon 2		PS4, PP1_strong, PM2, PVS1_moderate	PATH
6	c.1118C>A (p.Ser373Ter)		PVS1_strong, PM2, PS4_supporting	LPATH

The five-tier ClinVar classification: PATH (pathogenic), LPATH (likely pathogenic), VUS (variant of uncertain significance), LBEN (likely benign), BEN (benign); CONF (conflicting interpretations in ClinVar); criteria from Luo and Feurstein *et al.*[18]

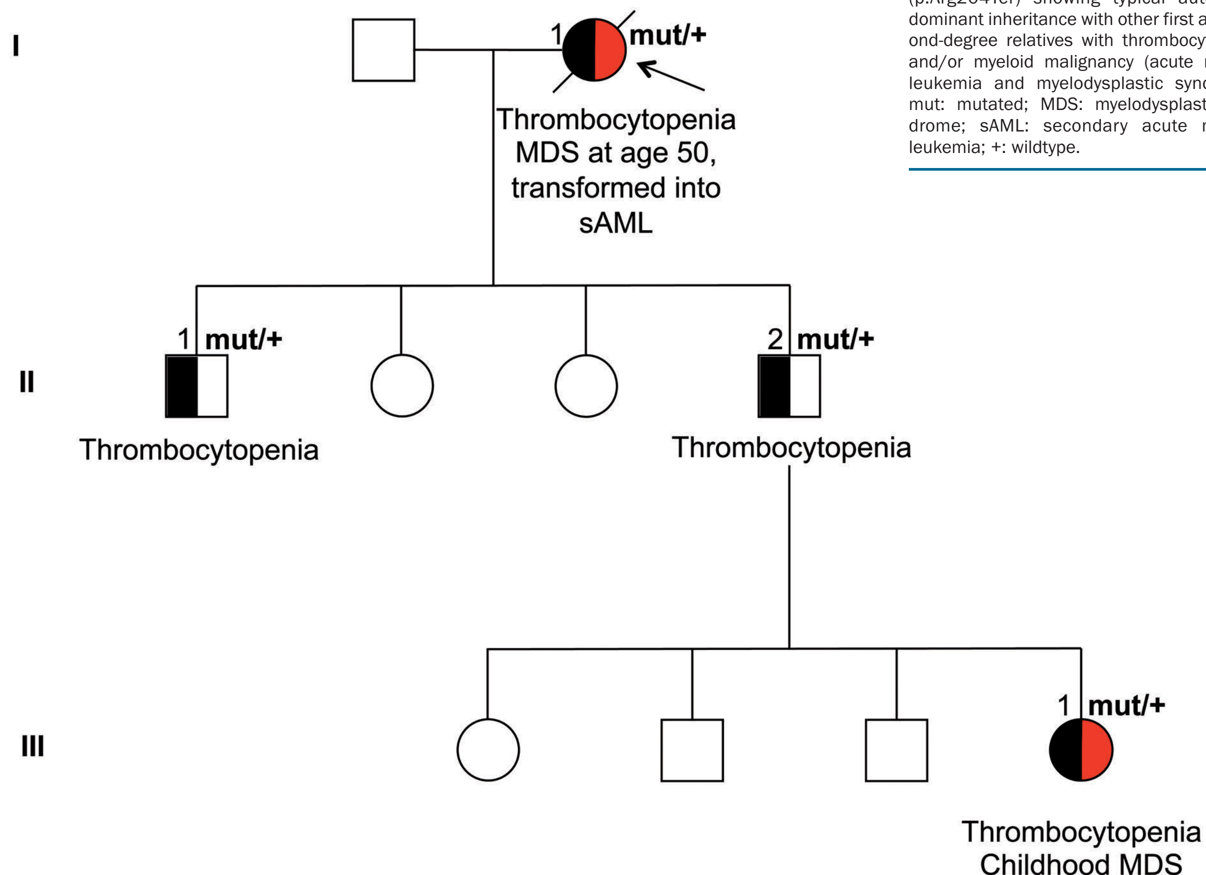


Figure 1. Family pedigree of a patient with acute myeloid leukemia, example 1. Germline *RUNX1* NM_001754:c.610C>T, (p.Arg204Ter) showing typical autosomal dominant inheritance with other first and second-degree relatives with thrombocytopenia and/or myeloid malignancy (acute myeloid leukemia and myelodysplastic syndrome). mut: mutated; MDS: myelodysplastic syndrome; sAML: secondary acute myeloid leukemia; +: wildtype.

A

1 77 204 270 398 438 480

RUNX1c

RHD TAD ID

RHD Runt homology domain (RHD) binds DNA and CBFβ

NLS Nuclear localization signal (NLS)

TAD Transactivation domain (TAD) required to activate transcription

ID Inhibitory domain (ID)

VWRPY VWRPY motif binds a transcriptional repressor

RUNX1 amino acid numbering uses transcript isoform C (NM_001754) as the preferred one for annotation according to MM-VCEP

B

ClinVar non-VUS

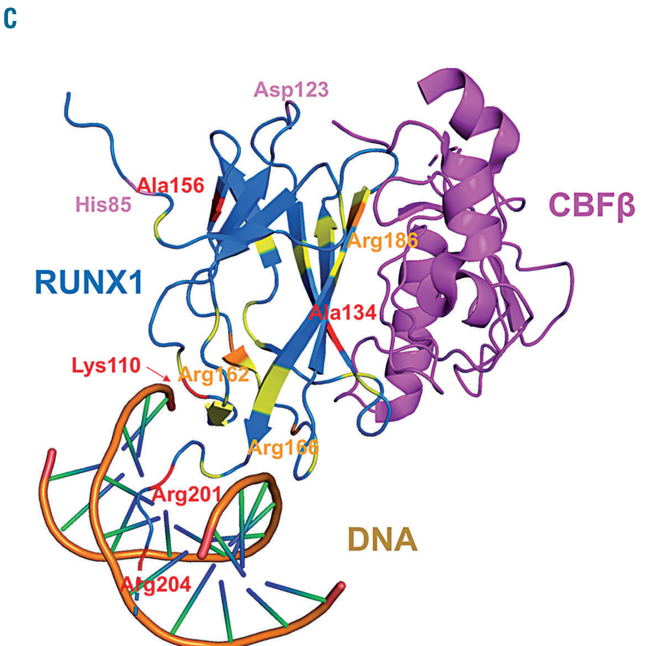
ClinVar VUS

MM-VCEP curated

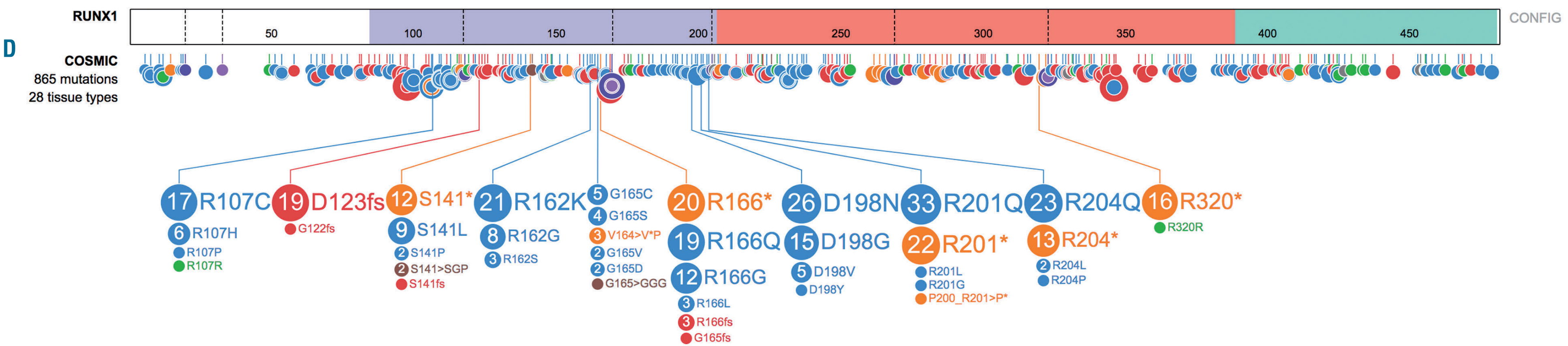
PATH (red), *LPATH* (orange), *VUS* (grey), *LBEN* (blue), *BEN* (purple), *CONF* (green)

Figure 2. Overview of RUNX1 protein domains, ClinVar-deposited and MM-VCEP curated pilot and COSMIC variants. (A) Schematic of the RUNX1 protein showing its key domains, including the Runt homology domain (RHD); annotation according to MM-VCEP is based on *RUNX1* transcript isoform C (NM_001754). (B) Schematic of ClinVar-deposited *RUNX1* variants in a one-dimensional line plot showing the current ClinVar five-tier classification; data plotted using *proteinpaint* from St. Jude.org with data from ClinVar (accessed 6/2019); RHD (purple), and ID (green) are highlighted; numbers indicate amino acid; dotted lines indicate exon boundaries. Some ClinVar-deposited variants (in 5' and 3' untranslated regions and large deletions) are not shown. Lollipops representing the variants are colored as follows: PATH (red), LPATH (orange), VUS (grey), LBEN (blue), BEN (purple), CONF (green). MM-VCEP pilot variants[18] are shown in the third row. (*continued on next page*)

C



D



(*continued from previous page*) (C) Three-dimensional structure of the RHD of RUNX1 (blue) complexed to CBFβ (pink) and DNA (orange).[82] ClinVar *RUNX1* variants in this domain (codons 81 to 204 only) are shown using PyMOL, version 2.3.0, as follows: yellow = VUS (n=20), pink = CONF (n=2), orange = LPATH (n=3), red = PATH (n=5); BEN and LBEN variants (n=0). Non-missense variants are not shown. (D) One-dimensional line plot of somatic variants deposited with the Catalog of Somatic Mutations in Cancer (COSMIC: data release 3 Nov. 2018, version 87; *https://cancer.sanger.ac.uk/cosmic*). MM-VCEP: Myeloid Malignancy Variant Curation Expert Panel; PATH: pathogenic; LPATH: likely pathogenic; VUS: variant of uncertain significance; LBEN: likely benign; BEN: benign; CONF: conflicting interpretations in ClinVar.

penetrance, with several affected individuals reported to have normal platelet counts or function.[19,24] The nonsense mutation in this patient (p.Arg204Ter) is predicted to lead to nonsense-mediated decay of the *RUNX1* mRNA transcript. *RUNX1* is located on the long arm of chromosome 21 and is translated into three major isoforms, designated RUNX1A, RUNX1B, and RUNX1C, by using two different promoters and alternative splicing. All transcripts are expressed during hematopoietic differentiation and/or maintenance of normal bone marrow function.[25-32] For variant annotation, the MM-VCEP utilizes the longest isoform, RUNX1C (NM_001754), as the default transcript, which includes all key domains such as the 128 amino acid (AA) long RUNT homology domain for DNA binding (RHD, AA 77-204), transactivation domain (TAD), inhibitory domain (ID) and the transcriptional repressor binding motif (VWRPY) and is most often used by clinical laboratories for *RUNX1* variant curation (Figure 2A). Germline variants have been reported throughout the gene in ClinVar with the majority currently classified as VUS (Figure 2B). The RUNX1 protein heterodimerizes through its RHD with CBFβ to form a master hematopoietic transcription factor (Figure 2C), which is essential for proliferation and differentiation of hematopoietic stem and progenitor cells, especially in the case of megakaryocytic differentiation.[33-35] Somatic mutations commonly occur in *RUNX1* (Figure 2D). According to *RUNX1*-specific criteria (Table 2),[18] the MM-VCEP applied the following codes (Table 3): PVS1 (nonsense variant predicted to undergo nonsense-mediated decay), PM2 (absence in all population databases), PS4_supporting (one proband meeting at least one of the *RUNX1* phenotype criteria), PP1 (co-segregation with disease in the family, three meioses) and arrived at a consensus classification of PATH.

Example 2. Missense variants, p.(His105Pro) (LPATH with PM2, PP3, PS4_supporting, PM1_supporting, and PM5_supporting) and p.(His105Gln) (LPATH with PS3, PM2, PP3, and PM1_supporting)

Missense mutations in *RUNX1* commonly occur in the RHD in somatic and germline contexts.[36-38] Of 325 *RUNX1* ClinVar variants 122 (37.5%) are missense, and currently in ClinVar, none in the RHD has been classified as BEN or LBEN (Figure 2C). When a novel missense variant is identified which has not been established as PATH or LPATH, it can be difficult to know whether the given change will affect protein function and explain the patient's phenotype. For example, two *RUNX1* missense variants in the RHD (NM_001754:c.314A>C, p.(His105Pro); and NM_001754:c.315C>A, p.(His105Gln)) were considered during the pilot variant analysis. The former was initially classified as a VUS in ClinVar (Figure 2B), but subsequently revised to LPATH upon MM-VCEP review (Figures 2B, 3A and 4). The conclusion of the LPATH assertion is based on the codes applied for this variant: PM2, PP3, PS4_supporting, PM1_supporting, and PM5_supporting (Table 3). Since the variant is completely absent from population databases, the MAF code PM2 is applied. For *in-silico* evaluation of missense variants, the MM-VCEP recommends using REVEL, a meta-predictor that combines 13 individual tools with high sensitivity and specificity, which has demonstrated the highest performance compared with individual tools or other ensemble methods.[39-41] The computational prediction code PP3 is applicable to the p.His105Pro variant due to a high REVEL score of 0.953 (MM-VCEP defined >0.75 as the cutoff). The ClinVar submitter (SCV000807773.1) provided us with the patient's clinical data from their laboratory and the proband met at least one of the *RUNX1* phenotype criteria (Table 1) which qualified for PS4_supporting. This example emphasizes the critical value of sharing internal laboratory data. There is only one meiosis in this family which is lower than the three required for the segregation code PP1. The MM-VCEP defined 13 residues in the RHD as the mutational hotspots for the PM1 code. In addition, variants in other parts of the RHD (AA 105-204) can have a reduced strength-level resulting in application of PM1_supporting. The last code PM5_supporting is applied on the p.(His105Pro) variant, because a different missense change p.(His105Gln) at the same residue has been classified as LPATH by the MM-VCEP (Table 3).

The codes PM2, PP3, PM1_supporting are also applicable to the p.(His105Gln) variant for the same reasons described. Furthermore, a strong pathogenic code PS3 is applied which contributes a significant weight to the final assertion. Transactivation assays of the p.(His105Gln) variant demonstrate altered transactivation (<20% of wildtype) and secondary assays also indicate altered DNA binding and functional consequences in a mouse model[42,43] manifested by disturbed myeloid differentiation and induction of a blast crisis or accelerated phase–like phenotype in mice.[42] These variants highlight the importance of evaluating similar variants and the critical benefit of functional studies showing that variants whose clinical significance were initially uncertain can be subsequently clarified to provide more definitive clinical classification and minimize reporting of VUS. Moreover, these variants demonstrate the value of leveraging the information on one variant to help classify another and data sharing between laboratories (Table 3).

Example 3. Missense variant, p.His85Asn (LBEN with BS1, BS3, and PP3)

RUNX1 NM_001754:c.253C>A, p.(His85Asn) is a missense mutation located within the RHD, but not within the mutational hotspot region (AA 105-204), with conflicting interpretations of pathogenicity in ClinVar (Figure 2B, C). Specifically, this variant had three submissions in ClinVar with two being PATH (submitted in 2002) and one being a VUS (submitted in 2018). The two 2002 submissions are from OMIM, which cited evidence from individual literature sources without a systematic curation process. Osato *et al.* reported an adult patient with AML carrying this variant.[44] However, the germline nature of the variant was not definitively determined. This variant has also been reported in an infant diagnosed with transient myeloproliferative disorder and Down syndrome whose phenotype does not meet any of the *RUNX1* phenotype criteria.[45] After analysis and curation by the MM-VCEP using the *RUNX1*-specific classification rules,[18] this variant was re-classified as LBEN, meeting codes BS1 and BS3, despite meeting PP3 (Table 3). According to the penetrance, prevalence and genetic and allelic heterogeneity of *RUNX*1, MM-VCEP refined the *RUNX1* specific MAF threshold for application of BS1 to 0.00015 (0.015%). The highest MAF of the p.His85Asn variant is 0.00043 (8 out of 18,768 alleles) from the East Asian subpopulation in the Genome Aggregation Database (gnomAD) which is higher than the *RUNX1*-

specific BS1 cutoff. Experimental studies have shown that this missense change displays normal transactivation activities (80-114% of wildtype) and does not affect DNA binding, heterodimerization with CBFβ or subcellular localization of the RUNX1 protein.[44,46] Therefore, the strong benign functional evidence code, BS3, is applied. Although this variant disrupts KMT2A binding, which impairs proper H3K4 histone methylation, this is not a qualified functional assay based on the MM-VCEP *RUNX1*-specific PS3/BS3 rule. Moreover, another well-established BEN variant p.(Leu56Ser) also impairs KMT2A binding.[46] Likely due to the location of the p.(His85Asn) variant within the RHD (Figure 2B), the REVEL score (0.852) of this variant is higher than the MM-VCEP defined 0.75 cutoff,[18] which results in the variant meeting a conflicting PP3 code. However, combining the BS1, BS3 and PP3 codes, a final assertion of LBEN is made based on a Bayesian classification framework.[15] Given that His85 is located away from binding interfaces in the three-dimensional structure (Figure 3B), it seems reasonable that variants at this position are LBEN.

This example highlights the value of functional studies in the context of the MM-VCEP variant curation and shows that substantive corrections of variant annotation may occur upon application of ClinGen MM-VCEP rules.[18] Implicit in this process is the expectation that as knowledge about FPD/AML improves with more functional or family data becoming available, the MM-VCEP rules are subject to revision so that annotation of clinical variants will become more accurate (Figure 4). ClinVar variant classifications such as VUS or those with conflicting interpretations may thus evolve to more diagnostic certainty.[47,48]

Example 4. Synonymous/intronic/non-coding variants, c.508+3delA (PATH with PS3, PP1_strong, PM2, PP3, PS4_supporting), p.Thr148= (LBEN with BP4, BP7) and p.Val419= (VUS with no codes)

RUNX1 variants affecting canonical splice positions ±1 or 2 at intron-exon boundaries are expected to disrupt

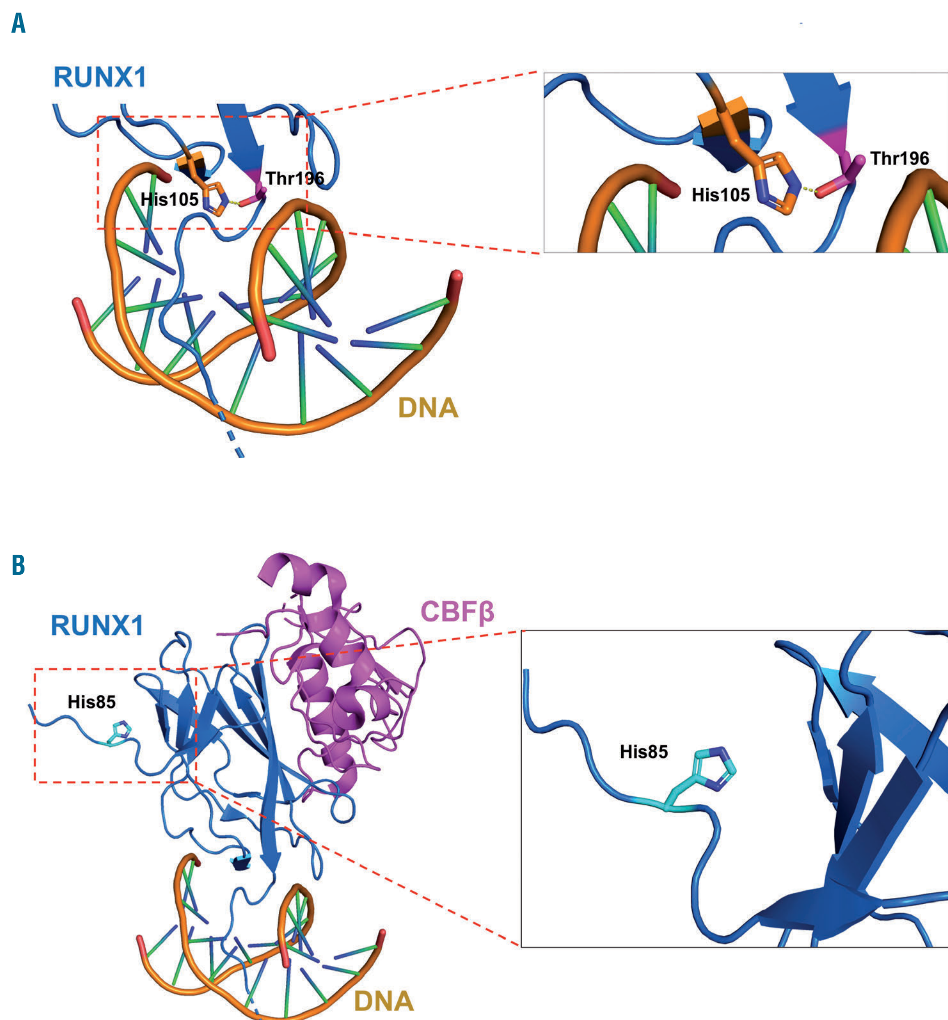


Figure 3. Three-dimensional structure of RUNX1 missense variants His105 and His85 considered as examples 2 and 3. (A) RUNX1 His105 is important functionally due to its location and thus involvement in DNA binding and close interaction with Thr196 by hydrogen bonding. Thr196 is a hotspot residue known to be critical. This structure-function relationship further supports classification of His105 variants as likely pathogenic (LPATH). (B) RUNX1 His85 is located close to the start of the Runt homology domain in a linker region, and is located far from the DNA binding surface. It is not involved in the core β-barrel structure and does not show any interactions, further supporting its classification as likely benign (LBEN). Structure of RUNX1 complexed to DNA and CBFβ (*https://www.rcsb.org/structure/1H9D*)[82] and plotted using PyMOL version 2.3.0.

splicing, leading to protein dysfunction (see *Online Supplementary Table S3* of reference by Luo and Feurstein[18]). All of the three canonical splicing site variants in the pilot set were classified as PATH or LPATH. More challenging, however, is the consideration of synonymous/intronic/non-coding variants which may result in cryptic splice site activation, and/or enhancement or repression of adjacent canonical splice sites. For example, the intronic NM_001754:c.508+3delA variant has been reported in a single family with disease segregation (8 meioses, PP1_strong). Several family members were diagnosed with thrombocytopenia, aspirin-like platelet aggregation defects, and dense granule abnormalities.[49] This variant is absent from population databases (PM2) and both splicing predictors (MaxEntScan and Splice-SiteFinder)[50,51] predict a significant decrease in the score of the canonical splice site (PP3). Moreover, experimental reverse transcriptase polymerase chain reaction studies (PS3), using RNA derived from two affected family members, were performed and indicate the creation of a novel cryptic splice site 23 nucleotides upstream of the normal splice site resulting in a frameshift p.(Arg162fs*177), and the transcript is predicted to undergo nonsense-mediated decay.[49] Combining all of these codes, a final assertion of PATH is given by the MM-VCEP (Table 3).

BP7 is a benign code specifically designed to evaluate synonymous/intronic/non-coding variants in the ACMG/AMP framework. BP7 can be applied if computational evidence suggests no impact on splicing, and the nucleotide is not conserved. The ClinVar variant with conflicting interpretations in ClinVar, NM_001754:c.444C>T, p.(Thr148=), has been classified as LBEN by the MM-VCEP using BP7 and the benign *in silico* prediction code, BP4 (Table 3). This nucleotide change is predicted to have no impact on splicing and it is also not conserved (phyloPscore: -4.3832, below the MM-VCEP-specified threshold of <0.1[18]). Clinical data from seven individuals with this variant were acquired from the original ClinVar submitter (SCV000761123.1) and revealed that none of the probands met any of the *RUNX1* phenotypic criteria.[18] Currently, only two *RUNX1* variants have been reported to display an abnormal splicing effect as demonstrated by RNA assays.[11,49] The potential effects of other splicing variants rely solely on *in silico* predictions. Although there is robust effort in consideration of algorithms to predict the effects of splicing variants, these algorithms require further evaluation. Indeed, we know of only limited experimental data within the *RUNX1* gene specifically to test these tools. Accordingly, the synonymous variant, NM_001754:c.1257G>A, p.(Val419=) is predicted to create alternative splice acceptor sites, but is not expected to abolish any existing consensus sites, as it is too far away from either end of the exon. Due to this *in silico* prediction result, none of the PP3/BP4 and BP7 codes can be assigned, and the classification of this variant remains a VUS. Further resolution of the significance of this variant could be obtained through parental testing, and/or RNA-sequencing data.

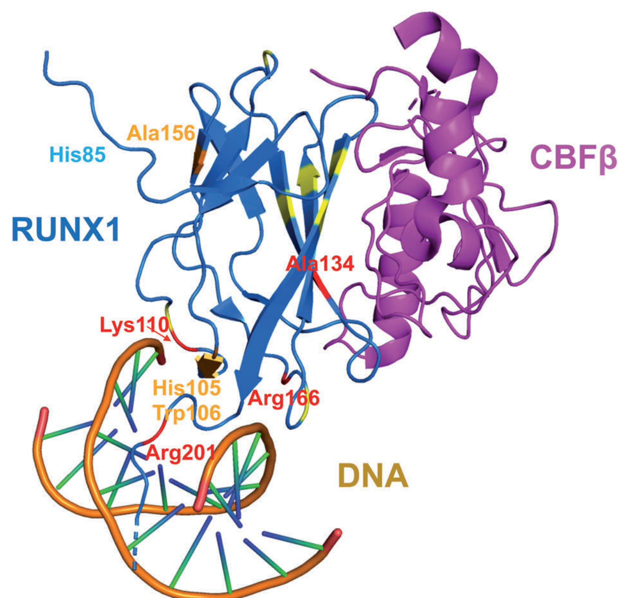


Figure 4. Three-dimensional structure of MM-VCEP-classified RUNX1 missense variants in the RHD. Variants in the RUNX1 RHD (blue) are shown highlighting the PATH, LPATH, and LBEN missense variants curated according to ClinGen MM-VCEP rules. PATH (red, n=4), LPATH (orange, n=4) and LBEN (cyan, n=1) variants are found proximate to key interaction domains of RUNX1 with DNA or its binding partner *CBFβ* (pink). VUS missense variants are shown in yellow (n=7). Non-missense variants and variants outside the RHD are not shown. (PyMOL version 2.3.0; structural data *https://www.rcsb.org/structure/1H9D*[82]). MM-VCEP: Myeloid Malignancy Variant Curation Expert Panel; RHD: Runt homology domain; PATH: pathogenic; LPATH: likely pathogenic; VUS: variant of uncertain significance; LBEN: likely benign; BEN: benign; CONF: conflicting interpretations in ClinVar.

Example 5. Copy number variants, deletion of exon 2 (PATH with PVS1_moderate, PM2, PS4, and PP1_strong)

Not infrequently, patients with FPD/AML have been reported to have copy number variants resulting in intragenic deletions of *RUNX1*.[52] As part of our pilot cohort, we evaluated several probands with copy number variants which at a minimum include *RUNX1* exon 2 deletion; data from two cases are shown (Figure 5). The analysis of copy number variants by using next-generation sequencing and/or single nucleotide polymorphism microarrays is particularly challenging because the breakpoints are often not captured in the sequenced regions or the microarray resolution defines only a range for the chromosomal location of the breakpoint, respectively, and thus the nucleotide level breakpoint may remain unknown. It can, therefore, be difficult to know the effect of partial gene/exon deletions, such as if the deletion is in- or out-of-frame, the latter of which may also lead to the introduction of a premature stop codon. Nevertheless, partial or whole gene deletion of *RUNX1* is expected to result in haploinsufficiency of the RUNX1 protein. Although MM-VCEP rules did not include recommendations for the formal classification of copy number variants,[18] several points should be noted. First, evaluation of the reference *Database of Genomic Variants* (*http://dgv.tcag.ca/dgv/app/home)* shows that copy number variants affecting *RUNX1* do not appear to be frequent.[18] Second, annotation of the specific breakpoints of these intragenic deletions may not always be possible, given that whole genome sequencing is not typically performed. Since contiguous exon deletion is a common pathogenic disease mechanism, it is imperative that laboratories performing germline testing for *RUNX1* use concurrent microarray testing, develop appropriate next-generation sequencing bioinformatics pipelines, or use alternate molecular techniques, such as quantitative polymerase chain reaction and multiplex ligation-dependent probe amplification, to screen for and exclude copy loss,

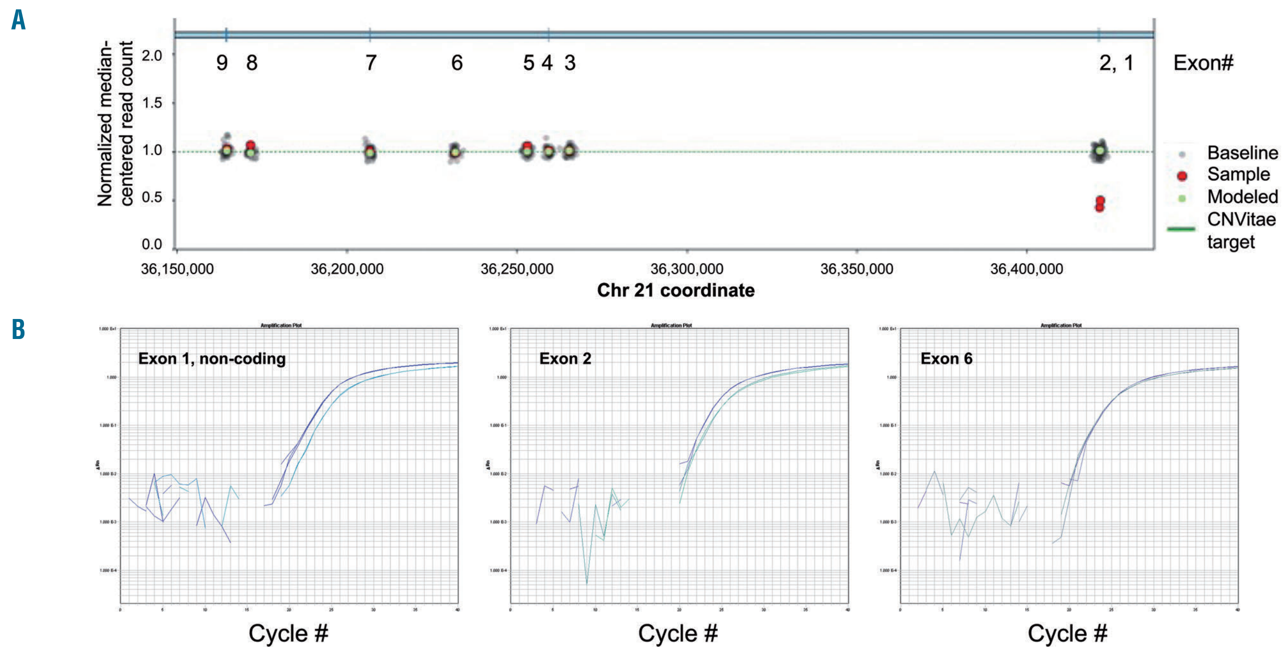


Figure 5. Copy number variant, deletion of exons of *RUNX1*. Testing for germline variants should include evaluation for copy number variants (CNV). (A; B) Clinical data from two different patients showing copy loss of *RUNX1* exons 1 and 2; one proband was identified by next-generation sequencing (A) and the other by single nucleotide polymorphism array analysis (data not shown), and confirmed by quantitative polymerase chain reaction (qPCR) (B). For the latter patient, qPCR confirmation showed heterozygous deletion of exons 1 and 2 with no loss in other exons (exon 6 only shown). CNV are difficult for annotation as breakpoints of the deletion may not be captured and thus, whether the deletion is in- or out-of-frame may not be known without whole genome sequencing. Testing of a germline sample (e.g. fibroblasts) is preferred to blood or bone marrow for CNV evaluation, as somatic copy loss may also occur in the tumor context.

Table 4. Variant details for example 6 from genome and exome sequencing, RNA-sequencing and karyotype analysis.

Structural	*KMT2A*, NM_005933, 11q23 (chr11:118347001-118353900) partial tandem duplication exons 4-8, predicted to be in-frame
SNVs and indels	*IDH1*, NM_005896:c.394C>T, p.(Arg132Cys), VAF ~50% *PHF6*, NM_001015877:c.860G>T, p.(Gly287Val), VAF ~ 95% *RUNX1*, NM_001754:c.1118C>A, (p.Ser373Ter), VAF ~50%
Copy number variants	*PALB2*, heterozygous exon 7 copy loss, NM_02467, chr16p12.2(23637272_23637800)x1 for hg19

Chr: chromosome; VAF: variant allele frequency.

as sequencing for single nucleotide variants and indels alone is insufficient for comprehensive germline evaluation. For the two cases herein (Figure 5), although we do not know the specific breakpoints for each, the common deletion of at least exon 2 allows us to apply the following codes: PVS1_moderate, PM2, PS4 (4 probands: 3 with chronic thrombocytopenia and 1 with AML), PP1_strong (7 meioses) to arrive at a PATH classification (Table 3).

Example 6. Late truncation variant, (p.Ser373Ter) (LPATH with PVS1_strong, PM2, and PS4_supporting)

A 14-year old male with a non-contributory family history presented with malaise, poor appetite, night sweats, and intermittent fever of about 1-month duration, thrombocytopenia ($27x10^9$/L), and subsequent bone marrow biopsy showed AML (PS4_supporting). After whole genome and exome sequencing on paired tumor and germline samples, along with RNA-sequencing (directed for recurrent fusion identification), his leukemia sample was shown to be *RUNX1*-mutated (NM_001754:c.1118C>A, (p.Ser373Ter)), hypodiploid with Y chromosome loss, without chromosomal fusions, and loss of heterozygosity, but had additional mutations including an intragenic heterozygous deletion of one copy of exon 7 of *PALB2* (Table 4).

Similar to Example 1, this *RUNX1* (p.Ser373Ter) variant is a nonsense mutation; however, it is not predicted to undergo nonsense-mediated decay, but rather is expected to generate a truncated protein without part of the TAD, ID and the VWRPY motif (Figure 2A). From a computational and predictive perspective, a PVS1_strong code is assigned following the PVS1 decision tree for null or truncating variants in *RUNX1*.[18] The variant is absent from the gnomAD and other population databases (with confirmed >20x sequencing coverage at this position in gnomAD). Given the variant's absence from population databases and adequate sequencing coverage of the region, a PM2 code is assigned. Although no additional evidence for the other categories (functional, segregation, *de novo* and allelic data) are available, this variant can be classified as LPATH (PVS1_strong, PM2, PS4_supporting). It is of interest to note that the somatic alterations reported in the diagnostic leukemia sample included partial tandem duplication of *KMT2A* and a single nucleotide varia-

tion in *PHF6*. Alterations in these two genes have been reported as cooperating events seen in leukemias from patients with germline *RUNX1* mutations.[37,53]

Importantly, if consideration is given to the mutations found in the leukemic cells in isolation, one cannot determine the germline or somatic origin of the variants reported. This is the case for most of the 'tumor-only' analyses being performed in many clinical laboratories. Without paired analysis of true germline tissue (e.g. cultured skin fibroblasts), such studies cannot definitively identify germline variants. In this case, the *KMT2A* partial tandem duplication and single nucleotide variation in *PHF6* and *RUNX1* could be tumor-drivers in the AML. However, given the sequencing data, including the variant allele frequency, both the *RUNX1* mutation and the *PALB2* exon 7 intragenic deletion could be germline variants. A detected variant allele frequency approaching 50% or 100% in the tumor may indicate potential germline origin[1] with either an intact wildtype allele or loss of heterozygosity, respectively. However, a high variant allele frequency cannot reliably serve as a proxy for testing of a true germline source. Therefore, if there is concern that a variant could be constitutional, testing of true germline material is critical.[1]

Discussion

Kindreds with FPD/AML were first reported by Luddy *et al.* in 1978[54] and phenotypically well-described as having a bleeding diathesis and myeloid neoplasia by Dowton *et al.* in 1985.[55,56] Subsequent linkage analysis identified *RUNX1* as the candidate gene at chromosome 21q22,[11] and mutations were detected in FPD/AML families in 1999.[11] Since these initial early reports, routine clinical testing for *RUNX1* gene mutations is now commonplace for the evaluation of somatic and germline disease in patients with myeloid neoplasms and thrombocytopenia.

In general, *RUNX1* variants include single nucleotide variations and indels, such as missense, nonsense, frameshift, and splice site variants, and copy number variations such as whole-gene and intragenic deletions. *RUNX1* is also frequently mutated somatically in AML and often the partner of various translocations resulting in gene fusions, such as t(8;21)(q22;q22) *RUNX1-RUNX1T1*.[57,58] To date, fusions of *RUNX1* have not been reported in the germline context, and most germline *RUNX1* variants are unique,[24] although some have been rarely seen in unrelated families. Given the limited data on rare variants, the clinical annotation of new variants remains challenging. The MM-VCEP was convened by ASH/ClinGen (*Online Supplementary Figure S1*) to develop rules for curating gene variant causing predisposition to myeloid neoplasia (Table 2). In this review, we describe the classification of six variant examples (Table 3) using the gene- and disease-specific rule modifications of the original ACMG/AMP 2015 framework.[14] Several points should be made.

First, it is critical to ensure that genomic testing intended to assess a germline predisposition is performed on a definitive germline sample because malignant hematologic diseases involve the peripheral blood and bone marrow, and somatic variants in these diseases can confound variant interpretation if an inappropriate sample is used. Here, in keeping with our MM-VCEP rules, cultured skin fibroblasts (gold standard, albeit invasive, costly, and time-consuming), cultured bone marrow mesenchymal stromal cells or DNA from hair roots are appropriate sources.[59,60] Alternatively, confirmation of the germline nature of a variant can be achieved by demonstrating its presence in two or more related individuals. The possibility of sample contamination by malignant cells is significant and consequently, peripheral blood, bone marrow, saliva, buccal swabs, DNA from paraffin blocks and even fingernails, which can contain monocytes, are inappropriate samples for germline testing. In some institutions, laboratories may accept T cells, enriched via flow cytometry sorting or column-based magnetic cell separation, as a germline sample for testing. It is important to recognize that some somatic alterations may occur early in hematopoietic stem and progenitor cells with multilineage potential to differentiate into T cells,[61] as recent single-cell studies have confirmed.[62-65] Thus, if T cells are used, the possibility that a detected variant may be somatic should still be considered. Once a variant is confirmed to be germline in a proband, however, additional testing for the known variant in related family members can be performed on any tissue source.

Second, we should keep an open mind about disease-causing alleles and the type of variants that may be seen and thus, we advocate for a broad testing approach. For example, in some laboratories, non-coding variants are automatically filtered as part of bioinformatic pipelines and may thus be omitted from subsequent review and interpretation. Recently, however, synonymous variants[66] in the *GATA2* gene, another gene predisposing to myeloid malignancy, were reported in addition to the known pathogenic deep intronic variants of an enhancer region of *GATA2*.[67] In *ANKRD26*, variants of the 5' untranslated region cause disease.[68,69] Furthermore, copy number alterations may not be assessed in somatic tumor testing panels. As diagnosticians, it is important to think broadly when analyzing genomic information for germline pathogenic variants. Given that these are rare diseases, we should not inadvertently exclude disease mechanisms and/or specific classes of mutations. For example, in case 4, some variants may remain as VUS until additional functional or familial segregation data become available for reclassification.[47,48]

Third, definitive annotation of variants by one institution will likely remain challenging. However, consistent application of MM-VCEP rules with ClinVar data deposition and thus inter-laboratory correspondence can significantly improve the accuracy and consistency of variant curation. In this regard, examples 2, 3, and 5 show how leveraging shared genomic and phenotypic data can be helpful to clarify VUS. We therefore advocate that clinical variant data be deposited into ClinVar. Specifically, laboratories offering germline testing should modify their test requisition forms to indicate that de-identified phenotype and variant data will be deposited into ClinVar as part of ongoing quality assurance and improvement efforts (*https://www.clinicalgenome.org/share-your-data/laboratories/*).[70,71] Additional details of the ClinVar deposition process are included in *Online Supplementary Figure S1.*

Fourth, *RUNX1* variant curation will improve as more is understood about the disease and gene through functional and family studies. Currently, variant annotation remains a challenging task, because of limited data for

determining the functional effect of a given variant change, despite methods of engineering variants for functional assessment.[72-74] Early studies in Speck's laboratory on *RUNX1* showed the significance of key residues in the RHD of RUNX1 by performing alanine scanning mutagenesis.[75,76] However, these early approaches are limited in that not every combination of nucleotide change was explored. By contrast, recent high-throughput functional genomic methods,[77] known as deep mutational scanning, utilize large-scale approaches to mutate every nucleotide of a gene, permitting one to test the functional consequence of all single nucleotide variations. This has, for example, been recently demonstrated for *BRCA1*.[78] Additionally, systematic mutagenesis of *PTEN* has provided a wealth of functional data to inform the classification of *PTEN* variants,[79] in conjunction with published rules developed by the PTEN-VCEP.[80] In the future, focused functional assays targeting specific VUS[16] and deep mutational scanning of genes should contribute to variant curation to resolve VUS.

Fifth, while functional testing of every given genomic variant is possible, it can be costly and difficult to do for every clinically significant gene. In this regard, family studies can aid in the classification of VUS. By systematically evaluating disease segregation in family members with paired genotyping for a known variant, accurate classification of a given variant can be achieved. For example, a recent study showed that this family-based method for variant classification can resolve a VUS classification more frequently than other traditional approaches can.[7,81] For rare diseases, such as FPD/AML, detailed pedigree and segregation analyses can be incredibly informative, and clinicians should be encouraged to test family members when possible, seeking help from local genetic counselors and/or geneticists as needed.[81] Hematologists and oncologists need to consistently take a detailed family and genetic history.

Summary

RUNX1 germline mutations associated with FPD/AML are key events in myeloid neoplasms, thrombocytopenia and leukemogenesis and represent a model of a germline gene disorder with pathogenic variants predisposing to myeloid and (to a lesser extent) lymphoid malignancies.[36] Providing an accurate clinical and pathologic variant interpretation for genomic variants detected in routine laboratory testing will remain critical for the provision of appropriate clinical care, including genetic counseling for the index patient and their at-risk relatives and donor-selection, in some cases benefiting from stem cell transplantation.

The ClinGen MM-VCEP variant interpretation process requires a detailed understanding of the biological and functional properties of *RUNX1* and disease phenotype. Here, we demonstrate the process for sequence variant interpretation of six variant examples. By introducing and thus standardizing genomic variant interpretation, we hope to improve patients' care, identify VUS that may benefit from directed research and encourage sharing of internal laboratory data to resolve uncertainty. In doing so, the MM-VCEP rules may ensure optimal insurance coverage for appropriate genomic testing and screening of family members, and ensure appropriate reimbursement for clinical laboratories. Overall, the ASH/ClinGen collaboration resulting in the first set of modified criteria for germline *RUNX1* variants should improve clinical care and recommendations for FPD/AML patients.

Acknowledgments

The Variant Curation Expert Panel (VCEP) thanks the Clinical Genome Resource (ClinGen) Sequence Variation Interpretation Working Group as well as the Executive Committee of the Hereditary Cancer Clinical Domain Working Group. The variant curation rules discussed in this publication were generated by the American Society of Hematology (ASH) in collaboration with Baylor College of Medicine and the University of North Carolina, National Institutes of Health (NIH)-funded ClinGen grant award recipients. The NIH supported this work through: U41HG009649 (to XL, and SEP) and U41HG009650 (to SM and JER), and the 2018 NIH/NCI Leukemia SPORE DRP award (P50CA100632-16, project 00007529) (C.D.D.).

The ASH-ClinGen Myeloid Malignancy MM-VCEP collaborative group includes co-authors (DW and LAG as co-chairs; XL, SF, SM, CK and DEP), and others, as follows:[18] Christopher Porter (Emory University, USA), Sarah Jackson (GeneDx, USA), Sioban Keel (University of Washington, USA), Michael Chicka (Prevention Genetics, USA), Anna Brown (Center for Cancer Biology, Australia), Anupriya Agarwal (Oregon Health & Science University, The Knight Cancer Institute, USA), Minjie Luo (Children's Hospital of Philadelphia, USA), Zejuan Li (Houston Methodist Institute for Academic Medicine, USA), Justyne E. Ross (University of North Carolina, USA), Panagiotis Baliakas (Uppsala University, Sweden), Courtney D. DiNardo (UT MD Anderson Cancer Center, USA), Alison Bertuch (Baylor College of Medicine, USA), Nikita Mehta (Mayo Clinic, USA), Thomas Vulliamy (Queen Mary University of London, UK), Ying Wang (BioReference, USA), Kim Nichols (St Jude Children's Research Hospital, USA), Luca Malcovati (University of Pavia & S. Matteo Hospital, Italy), Michael Walsh (Memorial Sloan Kettering Cancer Center, USA), Lesley Rawlings (Centre for Cancer Biology, Australia), Shannon McWeeney (Oregon Health & Science University, USA), Jean Soulier (Hopital Saint-Louis and University de Paris, France), Anna Raimbault (INSERM U1016, Institut Cochin, France), Mark Routbort (UT MD Anderson Cancer Center, USA), Liying Zhang (Memorial Sloan Kettering Cancer Center, USA), Gabriella Ryan (American Society of Hematology, USA), Nancy Speck (Abramson Family Cancer Research Institute, USA), and Sharon E. Plon (Baylor College of Medicine, USA).

References

1. Drazer MW, Kadri S, Sukhanova M, et al. Prognostic tumor sequencing panels frequently identify germ line variants associated with hereditary hematopoietic malignancies. Blood Adv. 2018;2(2):146-150.
2. Churpek JE, Pyrtel K, Kanchi KL, et al. Genomic analysis of germ line and somatic variants in familial myelodysplasia/acute myeloid leukemia. Blood. 2015;126(22): 2484-2490.
3. Churpek JE, Bresnick EH. Transcription factor mutations as a cause of familial myeloid neoplasms. J Clin Invest. 2019;129(2):476-488.
4. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127 (20):2391-2405.
5. Swerdlow SH CE, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, 2017.
6. Kohane IS, Masys DR, Altman RB. The incidentalome: a threat to genomic medicine. JAMA. 2006;296(2):212-215.
7. Shirts BH, Pritchard CC, Walsh T. Family-specific variants and the limits of human genetics. Trends Mol Med. 2016;22(11):925-934.
8. Beri-Dexheimer M, Latger-Cannard V, Philippe C, et al. Clinical phenotype of germline RUNX1 haploinsufficiency: from point mutations to large genomic deletions. Eur J Hum Genet. 2008;16(8):1014-1018.
9. Schlegelberger B, Heller PG. RUNX1 deficiency (familial platelet disorder with predisposition to myeloid leukemia, FPDMM). Semin Hematol. 2017;54(2):75-80.
10. Sood R, Kamikubo Y, P. L. Role of RUNX1 in hematological malignancies. Blood. 2017;129(15):2070-2082. Erratum in: Blood. 2018;131 (3):373.
11. Song WJ, Sullivan MG, Legare RD, et al. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. Nat Genet. 1999;23(2):166-175.
12. Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014;42 (Database issue):D980-985.
13. Rehm HL, Berg JS, Brooks LD, et al. ClinGen--the Clinical Genome Resource. N Engl J Med. 2015;372(23):2235-2242.
14. Richards S, Aziz N, Bale S, et al. 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. Genet Med. 2015;17(5):405-424.
15. Tavtigian SV, Greenblatt MS, Harrison SM, et al. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. Genet Med. 2018;20(9): 1054-1060.
16. Brnich SE, Rivera-Munoz EA, Berg JS. Quantifying the potential of functional evidence to reclassify variants of uncertain significance in the categorical and Bayesian interpretation frameworks. Hum Mutat. 2018;39(11):1531-1541.
17. Rivera-Munoz EA, Milko LV, Harrison SM, et al. ClinGen Variant Curation Expert Panel experiences and standardized processes for disease and gene-level specification of the ACMG/AMP guidelines for sequence variant interpretation. Hum Mutat. 2018;39 (11):1614-1622.
18. Luo X, Feurstein S, Mohan S, et al. Gene-specific criteria for germline RUNX1 variant curation: recommendations from the ClinGen Myeloid Malignancy Variant Curation Expert Panel. Blood Adv. 2019;3 (20):2962-2979.
19. Owen CJ, Toze CL, Koochin A, et al. Five new pedigrees with inherited RUNX1 mutations causing familial platelet disorder with propensity to myeloid malignancy. Blood. 2008;112(12):4639-4645.
20. Shiba N, Hasegawa D, Park MJ, et al. CBL mutation in chronic myelomonocytic leukemia secondary to familial platelet disorder with propensity to develop acute myeloid leukemia (FPD/AML). Blood. 2012;119(11):2612-2614.
21. Linden T, Schnittger S, Groll AH, Juergens H, Rossig C. Childhood B-cell precursor acute lymphoblastic leukaemia in a patient with familial thrombocytopenia and RUNX1 mutation. Br J Haematol. 2010;151(5):528-530.
22. Toya T, Yoshimi A, Morioka T, et al. Development of hairy cell leukemia in familial platelet disorder with predisposition to acute myeloid leukemia. Platelets. 2014;25(4):300-302.
23. Nickels EM, Soodalter J, Churpek JE, Godley LA. Recognizing familial myeloid leukemia in adults. Ther Adv Hematol. 2013;4(4):254-269.
24. Latger-Cannard V, Philippe C, Bouquet A, et al. Haematological spectrum and genotype-phenotype correlations in nine unrelated families with RUNX1 mutations from the French network on inherited platelet disorders. Orphanet J Rare Dis. 2016;11:49.
25. Challen GA, Goodell MA. Runx1 isoforms show differential expression patterns during hematopoietic development but have similar functional effects in adult hematopoietic stem cells. Exp Hematol. 2010;38(5):403-416.
26. Komeno Y, Yan M, Matsuura S, et al. Runx1 exon 6-related alternative splicing isoforms differentially regulate hematopoiesis in mice. Blood. 2014;123(24):3760-3769.
27. Brady G, Elgueta Karstegl C, Farrell PJ. Novel function of the unique N-terminal region of RUNX1c in B cell growth regulation. Nucleic Acids Res. 2013;41(3):1555-1568.
28. Lacaud G, Gore L, Kennedy M, et al. Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development in vitro. Blood. 2002;100(2):458-466.
29. Navarro-Montero O, Ayllon V, Lamolda M, et al. RUNX1c regulates hematopoietic differentiation of human pluripotent stem cells possibly in cooperation with proinflammatory signaling. Stem Cells. 2017;35(11):2253-2266.
30. Lorsbach RB, Moore J, Ang SO, et al. Role of RUNX1 in adult hematopoiesis: analysis of RUNX1-IRES-GFP knock-in mice reveals differential lineage expression. Blood. 2004;103(7):2522-2529.
31. Tsuzuki S, Hong D, Gupta R, et al. Isoform-specific potentiation of stem and progenitor cell engraftment by AML1/RUNX1. PLoS Med. 2007;4(5):e172.
32. Tsuzuki S, Seto M. Expansion of functionally defined mouse hematopoietic stem and progenitor cells by a short isoform of RUNX1/AML1. Blood. 2012;119(3):727-735.
33. Elagib KE, Racke FK, Mogass M, et al. RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation. Blood. 2003;101(11):4333-4341.
34. Kuvardina ON, Herglotz J, Kolodziej S, et al. RUNX1 represses the erythroid gene expression program during megakaryocytic differentiation. Blood. 2015;125(23):3570-3579.
35. Li Y, Jin C, Bai H, et al. Human NOTCH4 is a key target of RUNX1 in megakaryocytic differentiation. Blood. 2018;131(2):191-201.
36. Bellissimo DC, Speck NA. RUNX1 Mutations in inherited and sporadic leukemia. Front Cell Dev Biol. 2017;5:111.
37. Gaidzik VI, Bullinger L, Schlenk RF, et al. RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. J Clin Oncol. 2011;29(10):1364-1372.
38. Tang JL, Hou HA, Chen CY, et al. AML1/RUNX1 mutations in 470 adult patients with de novo acute myeloid leukemia: prognostic implication and interaction with other gene alterations. Blood. 2009;114(26):5352-5361.
39. Bean LJH, Hegde MR. Clinical implications and considerations for evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. Genome Med. 2017;9(1):111.
40. Ghosh R, Oak N, Plon SE. Evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. Genome Biol. 2017;18(1):225.
41. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet. 2016;99(4):877-885.
42. Zhao LJ, Wang YY, Li G, et al. Functional features of RUNX1 mutants in acute transformation of chronic myeloid leukemia and their contribution to inducing murine full-blown leukemia. Blood. 2012;119(12):2873-2882.
43. Tsai SC, Shih LY, Liang ST, et al. Biological activities of RUNX1 mutants predict secondary acute leukemia transformation from chronic myelomonocytic leukemia and myelodysplastic syndromes. Clin Cancer Res. 2015;21(15):3541-3551.
44. Osato M, Asou N, Abdalla E, et al. Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2alphaB gene associated with myeloblastic leukemias. Blood. 1999;93(6):1817-1824.
45. Taketani T, Taki T, Takita J, et al. Mutation of the AML1/RUNX1 gene in a transient myeloproliferative disorder patient with Down syndrome. Leukemia. 2002;16(9): 1866-1867.
46. Koh CP, Wang CQ, Ng CE, et al. RUNX1 meets MLL: epigenetic regulation of hematopoiesis by two leukemia genes. Leukemia. 2013;27(9):1793-1802.
47. Mersch J, Brown N, Pirzadeh-Miller S, et al. Prevalence of variant reclassification following hereditary cancer genetic testing. JAMA. 2018;320(12):1266-1274.
48. Slavin TP, Manjarrez S, Pritchard CC, Gray S, Weitzel JN. The effects of genomic germline variant reclassification on clinical cancer care. Oncotarget. 2019;10(4):417-423.
49. Michaud J, Wu F, Osato M, et al. In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. Blood. 2002;99(4):1364-1372.
50. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol. 2004;11(2-3):377-394.
51. Desmet FO, Hamroun D, Lalande M, et al. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37(9):e67.

52. Cavalcante de Andrade Silva M, Krepischi ACV, Kulikowski LD, et al. Deletion of RUNX1 exons 1 and 2 associated with familial platelet disorder with propensity to acute myeloid leukemia. Cancer Genet. 2018;222-223:32-37.
53. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. N Engl J Med. 2016;374(23):2209-2221.
54. Luddy RE, Champion LA, Schwartz AD. A fatal myeloproliferative syndrome in a family with thrombocytopenia and platelet dysfunction. Cancer. 1978;41(5):1959-1963.
55. Dowton SB, Beardsley D, Jamison D, Blattner S, Li FP. Studies of a familial platelet disorder. Blood. 1985;65(3):557-563.
56. Liew E, Owen C. Familial myelodysplastic syndromes: a review of the literature. Haematologica. 2011;96(10):1536-1542.
57. Miyoshi H, Shimizu K, Kozu T, et al. t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. Proc Natl Acad Sci U S A. 1991;88(23):10431-10434.
58. Erickson P, Gao J, Chang KS, et al. Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. Blood. 1992;80(7):1825-1831.
59. Feurstein S, Drazer MW, Godley LA. Genetic predisposition to leukemia and other hematologic malignancies. Semin Oncol. 2016;43(5):598-608.
60. Churpek JE, LA G. How I diagnose and manage individuals at risk for inherited myeloid malignancies. Blood. 2016;128 (14):1800-1813.
61. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014;506(7488):328-333.
62. Arends CM, Galan-Sousa J, Hoyer K, et al. Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. Leukemia. 2018;32(9):1908-1919.
63. Thol F, Klesse S, Kohler L, et al. Acute myeloid leukemia derived from lymphomyeloid clonal hematopoiesis. Leukemia. 2017;31(6):1286-1295.
64. Petti AA, Williams SR, Miller CA, et al. Mutation detection in thousand of acute myeloid leukemia cells using single cell RNA-sequencing. bioRxiv. 2018 October 17. [Epub ahead of Print]
65. Velten L, Story BA, Hernancez-Malmierca P, et al. MutaSeq reveals the transcriptomic consequences of clonal evolution in acute myeloid leukemia. bioRxiv. 2018 December 21 [Epub ahead of Print]
66. Wehr C, Grotius K, Casadei S, et al. A novel disease-causing synonymous exonic mutation in GATA2 affecting RNA splicing. Blood. 2018;132(11):1211-1215.
67. Hsu AP, Johnson KD, Falcone EL, et al. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. Blood. 2013;121(19):3830-38371.
68. Pippucci T, Savoia A, Perrotta S, et al. Mutations in the 5' UTR of ANKRD26, the ankirin repeat domain 26 gene, cause an autosomal-dominant form of inherited thrombocytopenia, THC2. Am J Hum Genet. 2011;88(1):115-120.
69. Noris P, Perrotta S, Seri M, et al. Mutations in ANKRD26 are responsible for a frequent form of inherited thrombocytopenia: analysis of 78 patients from 21 families. Blood. 2011;117(24):6673-6680.
70. Azzariti DR, Riggs ER, Niehaus A, et al. Points to consider for sharing variant-level information from clinical genetic testing with ClinVar. Cold Spring Harb Mol Case Stud. 2018;4(1).
71. Riggs ER, Azzariti DR, Niehaus A, et al. Development of a consent resource for genomic data sharing in the clinical setting. Genet Med. 2019;21(1):81-88.
72. Ma N, Zhang JZ, Itzhaki I, et al. Determining the pathogenicity of a genomic variant of uncertain significance using CRISPR/Cas9 and human-induced pluripotent stem cells. Circulation. 2018;138(23): 2666-2681.
73. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32(4):347-355.
74. Tsai SQ, Joung JK. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. Nat Rev Genet. 2016;17(5):300-312.
75. Zhang L, Li Z, Yan J, et al. Mutagenesis of the Runt domain defines two energetic hot spots for heterodimerization with the core binding factor beta subunit. J Biol Chem. 2003;278(35):33097-33104.
76. Li Z, Yan J, Matheny CJ, et al. Energetic contribution of residues in the Runx1 Runt domain to DNA binding. J Biol Chem. 2003;278(35):33088-33096.
77. Findlay GM, Boyle EA, Hause RJ, Klein JC, Shendure J. Saturation editing of genomic regions by multiplex homology-directed repair. Nature. 2014;513(7516):120-123.
78. Findlay GM, Daza RM, Martin B, et al. Accurate classification of BRCA1 variants with saturation genome editing. Nature. 2018;562(7726):217-222.
79. Matreyek KA, Starita LM, Stephany JJ, et al. Multiplex assessment of protein variant abundance by massively parallel sequencing. Nat Genet. 2018;50(6):874-882.
80. Mester JL, Ghosh R, Pesaran T, et al. Gene-specific criteria for PTEN variant curation: recommendations from the ClinGen PTEN Expert Panel. Hum Mutat. 2018;39(11):1581-1592.
81. Garrett LT, Hickman N, Jacobson A, et al. Family studies for classification of variants of uncertain classification: current laboratory clinical practice and a new web-based educational tool. J Genet Couns. 2016;25(6): 1146-1156.
82. Bravo J, Li Z, Speck NA, Warren AJ. The leukemia-associated AML1 (Runx1)--CBF beta complex functions as a DNA-induced molecular clamp. Nat Struct Biol. 2001;8 (4):371-378.