DOI: 10.1002/jha2.1012

## SHORT REPORT

# Beyond myeloid neoplasms germline guidelines: Validation of the thresholds criteria in the search of germline predisposition variants

Julia Mestre<sup>1,2</sup> I Lorea Chaparro<sup>1</sup> Ana Manzanares<sup>1</sup> Blanca Xicoy<sup>3</sup> Lurdes Zamora<sup>3</sup> Francesc Sole<sup>1,2</sup> I Oriol Calvete<sup>1</sup>

<sup>1</sup>MDS Group, Josep Carreras Leukaemia Research Institute, ICO-Hospital Germans Trias i Pujol, Universitat Autònoma de Barcelona, Badalona, Spain

<sup>2</sup>Facultat de Biociències, Universitat Autònoma de Barcelona, Barcelona, Spain

<sup>3</sup>Department of Hematology, ICO-IJC-Hospital Germans Trias i Pujol, UAB, Badalona, Catalonia, Spain

#### Correspondence

Oriol Calvete and Francesc Sole, MDS Group, Josep Carreras Leukaemia Research Institute, ICO-Hospital Germans Trias i Pujol, Universitat Autònoma de Barcelona, 08916 Badalona, Spain.

Email: ocalvete@carrerasresearch.org; fsole@carrerasresearch.org

#### **Funding information**

José Carreras Leukämie-Stiftung 2020, Grant/Award Numbers: DJCLS01R/2021, 2021SGR00560; (GRC) Generalitat de Catalunya: CERCA Programme/Generalitat de Catalunya; Joan Oró predoctoral programfrom Secretariad' Universitat siRecercadel Departamentde Recercai Universitats delaGeneralitatde Catalunya; Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR), Grant/Award Number: 2023 FI-1 00200; European Social Funding Plus (FSE+); Instituto de Salud Carlos III, Grant/Award Numbers: PI 20/000531, PI23/00007

## Abstract

**Introduction:** Germline predisposition to myeloid neoplasms can be suspected in patients younger than 50 years or when harboring mutations with a variant allele frequency (VAF) higher than 30% for point mutations in specific genes. To investigate the VAF thresholds' accuracy we have explored the prevalence of germline variants below the 30% VAF threshold.

eJHaem

British Society fo

**Methods:** A total of 40 variants with VAF lower than 30% in bone marrow samples of myeloid neoplasm patients were selected and studied in CD3<sup>+</sup> cells.

**Results:** All the selected variants were not found in CD3<sup>+</sup> cells except one variant in the *SF3B1* gene. However, the whole series was found somatic. Selected variants were also evaluated with our previously studied series of 52 variants with VAF higher than 30%.

**Conclusion:** Our study suggests that variants with VAF below 30% are strong somatic candidates but the variants with VAF higher than 30% cannot be considered of germline origin.

#### KEYWORDS

germline guidelines, germline predisposition, myelodysplastic syndromes (MDS), variant allele frequency (VAF)

# 1 | INTRODUCTION

Myelodysplastic syndromes (MDS) are a heterogeneous group of clonal hematopoietic neoplasms characterized by an inefficient

hematopoiesis in the bone marrow (BM). Although MDS have historically been categorized as a *de novo* disease occurring in older adults, the presence of an underlying genetic predisposition could explain around 15% of adult MDS cases [1]. Several germline predisposition to myeloid

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made. © 2024 The Author(s). *eJHaem* published by British Society for Haematology and John Wiley & Sons Ltd.

neoplasms (MN) clinical guidelines have been published with the aim of establishing the criteria to identify candidate patients [2–7]. There is a broad consensus on the use of a 50-age threshold as a criterion to select patients at risk of carrying a hereditary condition predisposing to MN. However, recent studies reported germline variants in patients of all age ranges [8, 9]. Likely, according to the Catalan Health Service (Cat Salut) [2] and the Nordic clinical guidelines [5], a variant with a variant allele frequency (VAF)  $\geq$ 30% or 40%, respectively, in BM might be germline and should therefore be validated in an extra-hematological tissue.

Considering the current discussion on this last criterion, we previously investigated the germline/somatic status of 52 variants with a VAF  $\geq$ 30% found in BM samples in a 168 MN patient cohort using paired germline control DNA [10]. Only 65% of tested variants were annotated as germline while 35% were found to be somatic. Interestingly, both germline and somatic variants were found in BM in all VAF ranges over 30% thus being undistinguishable when using VAF criteria. Consequently, we concluded that the VAF  $\geq$ 30% threshold should not be taken as a reliable criterion to determine the nature of variants.

To further investigate the applicability and limitations of VAF thresholds, we have extended our previous studies by exploring the prevalence of germline variants below the threshold of 30% of VAF in BM tissue.

#### 2 | METHODS

Retrospective data from next-generation sequencing of a cohort of 198 MN patients was evaluated (Figure 1A). Only missense, stop gain, and splicing variants from BM with a VAF between 10% and 30% and a minor allele frequency lower than 0.03 were considered. Variants with a VAF below 10% were considered as clonal hematopoiesis of indeterminate potential and were excluded from the analysis. A final selection of 40 variants from 27 MDS and 6 myelodysplastic/myeloproliferative neoplasm patients were validated by Sanger sequencing in germline tissue (Table 1 and Table S1). CD3<sup>+</sup> T lymphocytes obtained from peripheral blood samples of the respective patients were used as a reliable source of germline control DNA [5, 7]. Variants were considered germline if they were found in both BM and CD3<sup>+</sup> at a similar VAF value, while they were considered somatic if they were not present in the CD3<sup>+</sup> sample or showed a residual VAF (below 10%).

## 3 | RESULTS

The selected 40 variants with a VAF < 30% in BM were evaluated in CD3<sup>+</sup> T lymphocyte samples. Of those, 39 variants were not found in CD3<sup>+</sup> T lymphocytes (Table 1). The variant p.K666N in the *SF3B1* gene (VAF 29% in BM) was also found in the CD3<sup>+</sup> sample (VAF 19.06%) suggesting a germline origin. Information regarding the purity of the CD3<sup>+</sup> selection, the presence of cross-contamination with tumoral

cells, or another germline sample from this patient were not available. Alternatively, the presence of the variant was evaluated in hair follicle DNA from both healthy parents to confirm its germline nature. The variant was not present in either parent, suggesting that the variant was not inherited. The p.K666 mutation has been widely reported as somatic in COSMIC [11]. Therefore, this variant was considered acquired in the proband although the lineage where it occurred could not be determined.

The average VAF of the variants in BM and CD3<sup>+</sup> T lymphocyte samples were 20% and 2.5%, respectively, and were significantly different (p < 0.0005; Figure 1B). At this point, the 100% of studied variants with VAF < 30% were considered somatic. 85% of variants were annotated as missense, while splicing and stop gain variants were found at 8% and 7%, respectively.

Variants were grouped into functionally relevant pathways and represented according to their clonal size based on VAF. The mutations in genes involved in transcription and RNA splicing functions showed high VAFs (22.3% and 21.4% respectively). Variants involved in signaling were frequent and present across all VAF ranges, while all epigenetic regulators (DNA methylation and chromatin modification) showed a VAF below 20% (Figure 1C).

Finally, the series of 40 variants of the present study (VAF < 30%) was evaluated together with the series of 52 variants (VAF  $\geq$  30%) taken from the study by Calvete et al. [10] to discuss the prevalence of somatic and germline variants across all VAF ranges. Considering the whole series of 92 validated variants, the average VAF of confirmed somatic and germline variants in both BM and CD3<sup>+</sup> tissues were compared. Significant differences were found in both analyses, but they were particularly significant when comparing average VAF values between both types of variants in the CD3<sup>+</sup> tissue (p < 0.00001; Figure 1D).

In addition, the total number of germline and somatic variants were represented in 5% VAF intervals (Figure 1E). Somatic variants were more frequent at low VAF ranges: 21% of somatic variants were occurring in the 25–30% VAF interval, while 19% of variants were found in both 10–15% and 15–20% intervals. Above the 30% VAF threshold, the prevalence of somatic variants was equal to or below 8%. On the other hand, 100% of germline variants were occurring at VAF ranges over 30%. Positive trends were observed for germline variants as VAF increased, while opposite negative trends were observed for somatic variants. Indeed, the number of germline variants increased at higher VAF intervals but only when crossing the 30% VAF threshold, whereas somatic variants showed a negative trend as VAF values were higher (Figure 1E).

Finally, a correlation study was performed including all the variants from both series to assess how the CD3<sup>+</sup> sample purity was influencing the VAF of the variants observed in this tissue. Only 15 confirmed somatic variants with information regarding the purity and VAF in the CD3<sup>+</sup> sample were considered for the study. A determination coefficient (R2) of 0.03 was observed, which suggests that VAF observed in CD3<sup>+</sup> of somatic variants was not explained by the level of infiltration of tumoral cells in the control sample.



**FIGURE 1** Evaluation of studied variants regarding their VAF and its relationship with the pathophysiology of the disease. (A) Workflow on patient series. (B) Average VAF of the validated variants depending on the studied tissue. (C) Representation of the clonal size of somatic variants based on the average VAF found in BM tissue and according to the functional category of the gene. (D) Average VAF of germline and somatic variants depending on the tissue considering both series. (E) Number of somatic and germline variants per VAF interval considering both Calvete et al., [10] and our present series. Trend lines regarding the variation of the number of variants throughout VAF intervals are also shown. BM, bone marrow; CC, cohesin complex; CM, chromatin modification; VAF, variant allele frequency. \*\*\*p < 0.0005.

1023

	Patient								VAF BM	VAF CD3 <sup>+</sup>	CD3 Purity	
Series	₽	Diagnosis	Age	Karyotype	IPSS-R	Gene	Variant ID	AA change	(%)	(%)	(%)	Pathogenicity
WES	1161/14	MDS-MLD	77	46, XX, del(5)(q11q32)[5]/46, XX[15]	Int	CSNK1A1	ND	p.S27T	11.62	0.00	98.00	VUS
	1405	MDS-EB2	73	45, X, -Y[5]/46, XY[15]	Int	IDH2	rs121913503	p.R172K	17.15	0.00	96.20	Ъ
	1536	MDS-RS- MLD	53	47, XX, t(1;5)(q12;q11.2), +der(5)t(1;5)[10]/46, XX[10]	Int	SF3B1	rs377023736	p.K666N	29.00	19.48*	XK	ď
	986	MDS-MLD	47	45, XX, -7[7]/46, XX[8]	Int	TET2	ND	p.L1151R	23.40	0.00	94.39	LP
tNGS	1417	MDS/MPN	79	46, XY[20]	NA	NRAS	rs121913237	p.G12D	15.00	1.10	NK	д.
						PTPN11	rs397507510	p.D61N	18.70	0.49	NK	Ъ
	668	MDS-EB1	60	46, XY[30]	NK	STAG2	ŊŊ	p.C869R	15.32	0.00	86.80	LP
						SRSF2	rs751713049	р.Р95Н	27.63	2.24	86.80	Ъ
	254	MDS-MLD	70	46, XY[20]	Very low	JAK2	rs77375493	p.V617F	27.91	0.00	95.70	Ъ
	877	MDS-EB2	76	46, XX[7]	Int	TP53	rs751477326	p.P177R	13.08	1.39	99.10	Ъ
	591	MDS 5q-	65	46, XX, del(5q)(q12q33)[12]/46, XX[3]	Low	SF3B1	rs559063155	p.K700E	18.92	0.00	93.80	ъ
	766	MDS-MLD	XK	46, XY, del(20)(q11.2)[5]/46, XX[15]	NK	SRSF2	rs751713049	p.P95H	28.40	8.72	99.00	Ъ
Sanger	1600	MDS/MPN	84	46, XX[20]	NA	TET2	ND	p.G1361R	10.30	NK	98.10	LP
	1401	MDS/MPN	61	46, XX[20]	NA	NRAS	rs121913250	p.G12C	20.20	NK	98.90	Ъ
	801	MDS-MLD	80	46, XY[20]	Low	IDH2	rs121913502	p.R140Q	19.00	NK	99.55	Ъ
	1530	MDS-RS- MLD	82	46, XY[20]	Very low	SF3B1	rs559063155	p.K700E	28.20	XX	96.30	Ъ
	1626	MDS-EB2	79	48, XY, +20, del(20)(q11.2q13.3)x2, +21[20]	Very high	BCOR	QN	c.4976+1G>A	24.50	XK	97.50	LP
	1448	MDS-RS-	73	46, XX[20]	Very low	TET2	ND	p.A1379V	13.20	NK	99.50	VUS

 TABLE 1
 Characteristics of patients and overview of variants with VAF < 30% detected in this cohort.</th>

(Continues)

MLD

TP5
ZRS
SETI
CEB
RUN
TET
DNN
DNN
TP5
SF3I
ASX
DNN
TP5
SETI
KRA
TP5
SF3
AEB
CBL
CSN
NRA
CH2

MESTRE ET AL.

TABLE 1 (Continued)

## 4 DISCUSSION

In our study, we explored the prevalence of germline predisposition under the 30% VAF threshold in BM by validating in a control germline tissue all the variants found in the 10–30% VAF range. Significant differences were found between the average VAF observed in BM and the VAF in the CD3<sup>+</sup> samples in our series (Figure 1B) and when considering both Calvete et al. [10] (52 variants with VAF $\geq$ 30%) and our series together (Figure 1D). These results suggest that CD3<sup>+</sup> T lymphocytes can be used as a control sample to discriminate somatic and germline variants. Nevertheless, the p.K666N variant in the *SF3B1* gene could not be proven to be germline despite showing a high VAF in CD3<sup>+</sup> due to lack of data.

In our analysis, we also witnessed the presence of some residual VAF (between 1% and 8%) in the CD3+ sample of other confirmed somatic variants (Figure S1). This finding suggests that there might be some factors influencing the trustability of the CD3<sup>+</sup> sample as germline control, such as the low efficiency in the CD3<sup>+</sup> isolation leading to the presence of infiltration of tumoral cells in the control sample. Another limitation of CD3<sup>+</sup> sample as germline control could be in cases with an early hematopoietic somatic event (affecting both myeloid and lymphoid lineages) which could lead to false positive germline interpretations. In addition, a growing number of studies of hematopoietic disorders have emphasized the existence of somatic reversion events in multiple lymphocyte subsets [12]. These somatic alterations would hide the prevalence of pathogenic germline variants. All these scenarios compromise the usefulness of CD3<sup>+</sup> T lymphocytes as a reliable source of germline DNA. Therefore, other germline tissues such as skin fibroblasts or hair follicle cells might be more suitable for germline validation studies [4, 5, 13].

No association was found between the gene ontology of the mutated genes and the VAF of the mutations found in these genes. Our results also show that variants involved in transcription correlate with higher VAF values, but owing to the low number of studied variants, no strong associations were found (Figure 1C). Nevertheless, some differences between VAF variability trends could be observed depending on the gene ontology group. Categories such as splicing showed a wide VAF variability, while VAF values in genes related to DNA methylation tend to be more homogenous (Figure 1C). These results might be explained since the VAF of somatic variants directly correlates with the size of the tumoral clone carrying the mutation. Mutations conferring a proliferative advantage or occurring at an early stage are expected to be found in the dominant clones at higher VAFs. Our results agree with Montalban-Bravo et al. [14] studies since mutations with higher VAFs are found in genes that have mainly been described in dominant clones such as SF3B1, BCOR, TP53, and JAK2.

# 5 | CONCLUSION

In summary, our results support that as VAF increases, germline variants become more prevalent while the frequency of somatic variants decreases despite being still present at higher VAF ranges. On the other hand, these results also reveal that the strategy of searching for germline predisposition variants when VAF is above 30% is appropriate since no germline variants were found below that threshold in the tested genes.

Thus, although this guidelines threshold does not allow the discrimination between somatic and germline variants and should not be considered as trustable criteria to identify germline predisposition, it could serve to discard the germline origin of point mutations when VAF falls below the 30% value.

### AUTHOR CONTRIBUTIONS

Julia Mestre, Oriol Calvete, and Francesc Sole designed the study and wrote the manuscript. Blanca Xicoy and Lurdes Zamora recruited the patients, provided the samples, and revised the work. Lorea Chaparro and Ana Manzanares collected the clinical data and processed the samples. Julia Mestre conducted the experiments and performed data analysis. Julia Mestre and Oriol Calvete interpreted the data. Francesc Sole and Oriol Calvete supervised the project.

#### ACKNOWLEDGEMENTS

The authors acknowledge the patients and IJC Collection of Samples circuit for their collaboration. This work was supported in part by two grants from the Instituto de Salud Carlos III PI 20/000531 and PI23/00007 (Co-funded by European Regional Development Fund. ERDF, a way to build Europe); José Carreras Leukämie-Stiftung 2020 (DJCLS 01R/2021); 2021 SGR00560 (GRC) Generalitat de Catalunya; economical support from CERCA Programme/Generalitat de Catalunya. JM was supported by the Joan Oró predoctoral program from Secretaria d'Universitats i Recerca del Departament de Recerca i Universitats de la Generalitat de Catalunya, Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR), (2023 FI-1 00200), co-funded by European Social Funding Plus (FSE+).

### CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

#### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon a reasonable request. The data are not publicly available due to privacy or ethical restrictions.

#### ETHICS STATEMENT

Patient data and related samples have been recorded in a single database. The confidentiality and traceability of the patient data are guaranteed by the NorayBio laboratory management tool (Noray Bioinformatics S.L., IGTP-HUGTiPBiobanc). Only samples of patients who have signed an informed consent have been included in the project.

#### PATIENT CONSENT STATEMENT

The authors have confirmed patient consent statement is not needed for this submission.

## CLINICAL TRIAL REGISTRATION

The authors have confirmed clinical trial registration is not needed for this submission.

#### ORCID

Julia Mestre D https://orcid.org/0000-0002-4809-3897 Lorea Chaparro D https://orcid.org/0009-0008-6806-9342 Blanca Xicoy D https://orcid.org/0000-0002-0295-1307 Lurdes Zamora D https://orcid.org/0000-0003-1713-7110 Francesc Sole D https://orcid.org/0000-0002-3251-2161 Oriol Calvete D https://orcid.org/0000-0002-2623-2876

#### REFERENCES

- Bannon S, DiNardo C. hereditary predispositions to myelodysplastic syndrome. Int J Mol Sci. 2016;17(6):838.
- Bellosillo B, Brunet J, Lazaro C, Montoro MJ, et al. Servei Catala de la Salut Determinacions del perfil genètic de les síndromes hereditàries de càncer en l'adult i pediatria Servei Català de la Salut; 2023 (version 3). Available from: https://hdl.handle.net/11351/8438.3
- Rudelius M, Weinberg OK, Niemeyer CM, Shimamura A, Calvo KR. The International Consensus Classification (ICC) of hematologic neoplasms with germline predisposition, pediatric myelodysplastic syndrome, and juvenile myelomonocytic leukemia. Virchows Archiv [Internet]. 2023;482(1):113–30. Available from: https://doi.org/10. 1007/s00428-022-03447-9
- Greenberg PL, Stone RM, Al-Kali A, Bennett JM, Borate U, Brunner AM, et al. NCCN Guidelines<sup>®</sup> Insights: myelodysplastic syndromes, Version 3.2022. J Natl Compr Canc Netw. 2022 Feb;20(2):106–17.
- Baliakas P, Tesi B, Wartiovaara-Kautto U, Stray-Pedersen A, Friis LS, Dybedal I, et al. Nordic guidelines for germline predisposition to myeloid neoplasms in adults: recommendations for genetic diagnosis, clinical management, and follow-up. Hemasphere. 2019 Dec;3(6):e321.
- Grupo Español de SMD (GESMD). GUIAS ESPAÑOLAS DE SMD Y LMMC. Edición 2022.
- Speight B, Hanson H, Turnbull C, Hardy S, Drummond J, Khorashad J, et al. Germline predisposition to haematological malignancies: best practice consensus guidelines from the UK Cancer Genetics Group (UKCGG), CanGene-CanVar and the NHS England Haematological Oncology Working Group. Br J Haematol. 2023;201(1):25–34.

- Feurstein S, Trottier AM, Estrada-Merly N, Pozsgai M, McNeely K, Drazer MW, et al. Germ line predisposition variants occur in myelodysplastic syndrome patients of all ages. Blood. 2022;140(24):2533–48.
- Kubota Y, Zawit M, Durrani J, Shen W, Bahaj W, Kewan T, et al. Significance of hereditary gene alterations for the pathogenesis of adult bone marrow failure versus myeloid neoplasia. Leukemia. 2022;36(12):2827–34.
- Calvete O, Mestre J, Durmaz A, Gurnari C, Maciejewski JP, Solé F. Are the current guidelines for identification of myelodysplastic syndrome with germline predisposition strong enough? Br J Haematol. 2023;201(1).
- Brian Dalton W, Helmenstine E, Pieterse L, Li B, Gocke CD, Donaldson J, et al. The K666N mutation in SF3B1 is associated with increased progression of MDS and distinct RNA splicing. Blood Adv. 2020;4(7):1192–96.
- Revy P, Kannengiesser C, Fischer A. Somatic genetic rescue in Mendelian haematopoietic diseases. Nat Rev Genet. 2019;20(10):582–98.
- Duncavage EJ, Bagg A, Hasserjian RP, DiNardo CD, Godley LA, lacobucci I, et al. Genomic profiling for clinical decision making in myeloid neoplasms and acute leukemia. Blood. 2022;140(21):2228– 47.
- Montalban-Bravo G, Takahashi K, Patel K, Wang F, Xingzhi S, Nogueras GM, et al. Impact of the number of mutations in survival and response outcomes to hypomethylating agents in patients with myelodysplastic syndromes or myelodysplastic/myeloproliferative neoplasms. Oncotarget. 2018;9(11):9714–27.

#### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mestre J, Chaparro L, Manzanares A, Xicoy B, Zamora L, Sole F, et al. Beyond myeloid neoplasms germline guidelines: Validation of the thresholds criteria in the search of germline predisposition variants. eJHaem. 2024;5:1021–27. https://doi.org/10.1002/jha2.1012

1027