# The clinical and genomic landscape of patients with *DDX41* variants identified during diagnostic sequencing

Anna Maierhofer,<sup>1</sup> Nikita Mehta,<sup>2</sup> Ryan A. Chisholm,<sup>3</sup> Stephan Hutter,<sup>1</sup> Constance Baer,<sup>1</sup> Niroshan Nadarajah,<sup>1</sup> Christian Pohlkamp,<sup>1</sup> Ella R. Thompson,<sup>4,5</sup> Paul A. James,<sup>4</sup> Wolfgang Kern,<sup>1</sup> Claudia Haferlach,<sup>1</sup> Manja Meggendorfer,<sup>1</sup> Torsten Haferlach,<sup>1</sup> and Piers Blomberv<sup>4-6</sup>

<sup>1</sup>Munich Leukemia Laboratory, Munich, Germany; <sup>2</sup>Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY; <sup>3</sup>Department of Biological Sciences, National University of Singapore, Singapore; <sup>4</sup>The Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, Australia; <sup>5</sup>Clinical Haematology, Peter MacCallum Cancer Centre, Royal Melbourne Hospital, Melbourne, Australia; and <sup>6</sup>Torsten Haferlach Leukaemiediagnostik Stiftung, Munich, Germany

#### Key Points

- Observation of a somatic DDX41 variant with a germ line DDX41 variant can be integrated into existing frameworks to improve curation.
- Diverse types of genomic lesions can be responsible for pathogenic variants in DDX41, including synonymous variants and exon-level deletions.

Deleterious germ line variants in DDX41 are a common cause of genetic predisposition to hematologic malignancies, particularly myelodysplastic neoplasms (MDS) and acute myeloid leukemia (AML). Targeted next-generation sequencing was performed in a large cohort of sequentially recruited patients with myeloid malignancy, covering DDX41 as well as 30 other genes frequently mutated in myeloid malignancy. Whole genome transcriptome sequencing data was analyzed on a separate cohort of patients with a range of hematologic malignancies to investigate the spectrum of cancer predisposition. Altogether, 5737 patients with myeloid malignancies were studied, with 152 different DDX41 variants detected. Multiple novel variants were detected, including synonymous variants affecting splicing as demonstrated by RNA-sequencing. The presence of a somatic DDX41 variant was highly associated with DDX41 germ line variants in patients with MDS and AML, and we developed a statistical approach to incorporate the co-occurrence of a somatic DDX41 variant into germ line variant classification at a very strong level (as per the American College of Medical Genetics and Genomics/Association for Molecular Pathology guidelines). Using this approach, the MDS cohort contained 108 of 2865 (3.8%) patients with germ line likely pathogenic/pathogenic (LP/P) variants, and the AML cohort 106 of 2157 (4.9%). DDX41 LP/P variants were markedly enriched in patients with AML and MDS compared with those in patients with myeloproliferative neoplasms, B-cell neoplasm, and T- or B-cell acute lymphoblastic leukemia. In summary, we have developed a framework to enhance DDX41 variant curation as well as highlighted the importance of assessment of all types of genomic variants (including synonymous and multiexon deletions) to fully detect the landscape of possible clinically relevant DDX41 variants.

#### Introduction

Deleterious germ line variants in the DEAD-box helicase 41 (DDX41) gene are the most frequent monogenic cause of germ line predisposition to hematologic malignancy, being present in ~3% of patients with myelodysplastic neoplasms (MDS) or acute myeloid leukemia (AML).<sup>1</sup> The potential clinical relevance of the detection of deleterious germ line DDX41 variants in myeloid malignancy

Submitted 7 August 2023; accepted 16 October 2023; prepublished online on *Blood Advances* First Edition 24 October 2023. https://doi.org/10.1182/bloodadvances.2023011389.

The data sets used and/or analyzed during this study are available on reasonable request from the corresponding author, Piers Blombery (piers.blombery@petermac.org).

The full-text version of this article contains a data supplement.

<sup>© 2023</sup> by The American Society of Hematology. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

includes impact on disease prognostication,<sup>2</sup> choice of allogeneic transplant donor,<sup>3</sup> and implications for family members who may be at risk through predictive testing. The clinical importance of detecting variants in *DDX41* has led to its rapid inclusion as a gene on many targeted sequencing panels used in routine diagnostic sequencing in patients with myeloid malignancy throughout the world.

Although numerous large cohorts of patients with DDX41 have been recently described, it is unlikely that the entire landscape of genomic variation in DDX41 has been documented, particularly given the presence of variants restricted to specific ancestral populations.<sup>4,5</sup> In addition, variants in *DDX41* present challenges to the diagnostic laboratory in terms of variant curation, given their incomplete penetrance, multiple founder mutations, and relatively poorly understood biology. We aim to describe the clinical and genomic landscape of a large cohort of patients with DDX41 variants identified during clinical sequencing in the diagnostic laboratory context, and in doing so, we describe (1) multiple novel recurrent pathogenic germ line and somatic variants, (2) a statistical rationale for incorporating somatic DDX41 variants as a criterion for pathogenicity in germ line DDX41 variant assessment, and (3) the hematologic disease risk spectrum of patients harboring DDX41 variants.

#### **Methods**

#### **Patient cohort**

A total of 5737 patients were analyzed by next-generation sequencing for *DDX41* as well as genes recurrently somatically mutated in myeloid diseases (supplemental Table 1). Cytomorphological information was available for all patients: 2157 patients were diagnosed with AML, 2865 with MDS, and 715 with classical myeloproliferative neoplasm (MPN; all with *JAK2*, *CALR*, or *MPL* driver variants).

#### **Targeted next-generation sequencing**

DNA was isolated from bone marrow (n = 5092) or peripheral blood (n = 645); sequencing was performed on NovaSeq 6000 instruments after Nextera Flex library preparation (Illumina, San Diego, CA) and xGen hybridization capture according to the manufacturer's protocol (IDT Inc, Coralville, IA). Samples were sequenced to a minimum depth of 400× (target depth 1500×) across the cohort. Reads were mapped to the hg19 human reference genome with Isaac,<sup>6</sup> and single-nucleotide variants and small insertions/deletions were called with Pisces<sup>7</sup> and Pindel<sup>8</sup> (for *FLT3*-ITD; available via BaseSpace, Illumina), with a limit of detection of 3%. Structural variants/large insertion-deletion variants were called with Genome Rearrangement IDentification Software Suite (GRIDSS).<sup>9</sup>

#### WGS/RNA-sequencing

Whole genome sequencing (WGS) was performed on a separate cohort of 1407 patients whose samples had been sent to the Munich Leukemia Laboratory (MLL) between 2006 and 2021 for diagnostic workup. The respective diagnosis was established based on cytomorphology, immunophenotyping, cytogenetics, and molecular genetics following World Health Organization guidelines. All patients gave their written informed consent for scientific evaluations. The study was approved by the internal review board and adhered to the tenets of the Declaration of Helsinki. Retrospective DNA samples from bone marrow and peripheral blood, at diagnosis or treatment-naïve stages, were collected from all patients, and DNA and total RNA were extracted using the MagNA Pure 96 instrument and the MagNAPure96 DNA and Viral NA LV Kit and MagNA Pure 96 Cellular RNA LV Kit, respectively (Roche Life-Science, Mannheim, Germany). Samples were sequenced to a median depth of 106× across the cohort. RNA-sequencing was performed on 2 samples to analyze splicing effects. Library preparation and analytical procedures were performed as previously described<sup>10-13</sup>; briefly, stranded RNA-sequencing libraries were constructed from ribosomal RNA-depleted RNA using the TruSeq Total Stranded RNA kit (Illumina, San Diego, CA) and sequenced on a NovaSeg 6000 system with a median of 50 million paired reads per sample, with reads aligned by Spliced Transcripts Alignment to a Reference (version 2.5.0a) to the human reference genome (hg19).

### Variant classification and interpretation/germ line testing

All coding exons of DDX41 (NM 016222.4) were sequenced, and variants were called in coding regions as well as flanking intronic sites (within 8 base pairs). Population variation and cancer or genetic disease databases were used in addition to literature review to assist with variant interpretation including the Genome Aggregation Database (gnomAD v2), the Catalogue of Somatic Mutations in Cancer (cancer.sanger.ac.uk), and ClinVar (ncbi.nlm.nih.gov/clinvar). DDX41 variants with >150 alleles (0.06% global population allele frequency) in gnomAD were excluded. Further analysis was restricted to protein altering variants, canonical splice-acceptor/ donor site variants, and synonymous variants or variants in the flanking intronic sites between  $\pm 3$  and  $\pm 8$  that were predicted to affect splicing by SpliceAI (threshold >0.2).<sup>14</sup> Variants with a variant allele frequency (VAF) higher than 35% for larger insertion/deletions and 40% for single nucleotide substitutions were presumed to be of germ line origin. When available, germ line origin was confirmed by analyzing nails and/or oral mucosa or by analyzing remission samples. Confirmed and presumed germ line variants in DDX41 were classified for pathogenicity according to the guidelines from the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP)<sup>15</sup> with appropriate modifications recommended by the Clinical Genome Sequence Variant Interpretation Working Group (ClinGen SVI).<sup>16</sup> Details about specific use of ACMG/AMP criteria for assigning pathogenicity of DDX41 germ line variants are listed in supplemental Table 2.

#### **Statistical analyses**

Statistical analyses were performed with Statistical Package for the Social Sciences version 19.0.0 (Chicago, Statistical Package for the Social Sciences Inc); the reported *P* values are 2-sided. Dichotomous variables were compared between different groups with the use of the  $\chi^2$  or Fisher exact test and continuous variables by Student *t* test.

#### Ethics approval and consent to participate

All patients gave written informed consent for the use of data for scientific evaluations. The study was approved by the internal review board of MLL and by the Bavarian Ethics Committee of the Bavarian Chamber of Physicians with the number 05117. The study adhered to the tenets of the Declaration of Helsinki.

#### Results

#### Characterization and frequency of DDX41 variants

All coding exons (including 8 base pair-flanking intronic sequences) of *DDX41* were sequenced in a cohort of 5737 sequentially recruited patients with samples referred for molecular investigation of AML (n = 2157), MDS (n = 2865), and MPN (n = 715) to the MLL in Germany, which receives referrals from a predominantly Northwestern European population. Using a relatively permissive approach to variant filtering at the first instance (including gnomAD < 150 alleles [0.06% global population frequency] and/or synonymous/flanking intronic regions variants affecting splicing), 152 different *DDX41* variants were detected in 294 patients (5.1%) over the entire cohort, with variants detected in 137 of 2157 patients (6.4%) with AML, 145 of 2865 patients with MDS (5.1%), and 12 of 715 patients (1.7%) with classical MPN (*JAK2*-, *CALR*-, or *MPL*-mutated; Figure 1; supplemental Table 3).

DDX41 variants were observed in patients in 4 main patterns. The first pattern was the presence of multiple DDX41 variants, with at least 1 variant generally detected at >35-40% VAF and another variant detected at <35-40%. This first pattern occurred in 181 of 294 patients (61.6%) and is most consistent with a somatic DDX41 variant occurring in the context of a germ line DDX41 variant (herein referred to as DDX41<sup>G/S</sup>). The second pattern observed was 2 DDX41 variants detected at <35-40% VAF (herein referred to as DDX41<sup>S/S</sup>). This is most consistent with the presence of 2 somatic variants, and 11 of 294 patients (3.7%) were observed with this pattern. The final 2 patterns observed were patients with only putative germ line variant(s) without an observed putative somatic DDX41 variant (n = 82 of 294, 27.9%), referred to as DDX41<sup>G</sup>, and those with only 1 putative somatic variant without a putative germ line variant (n = 20 of 294; 6.8%), referred to as DDX41<sup>S</sup>. Notably, the DDX41 variant pattern differed with disease subtype, with the group with DDX41<sup>G</sup> being highly enriched in patients with MPN (12 of 12) compared with with AML (30 of 137) or MDS (40 of 145; Figure 1).

Nonhematologic germ line samples were available for 53 patients, and germ line origin was confirmed in 51 of 181 patients exhibiting the  $DDX41^{G/S}$  pattern and in 2 of 82 patients exhibiting the  $DDX41^G$  pattern by testing of nails, buccal swabs, or remission samples. In all 51  $DDX41^{G/S}$  cases, suspected somatic variants were confirmed to be absent from the germ line. For further analyses, those patients harboring a variant with a VAF greater than 35-40% from whom a nonhematologic specimen could not be obtained were considered to have a putative germ line variant. Hereafter, the term germ line refers to presumed and confirmed germ line variants.

Multiple recurrent germ line variants were observed in patients with  $DDX41^{G/S}$  and  $DDX41^G$  in our cohort, including the start-loss variant p.(Met1?) (n = 52), p.(Gln41\*) (n = 22), p.(Asp140Glyfs\*2) (n = 15), and p.(Lys331del) (n = 13). All of these germ line variants are reported to be common in the European population, consistent with the population from which patients were referred; conversely, the p.(Ala500Cysfs\*9) variant, which is most commonly described in

populations from Korea and Japan, was not detected in our cohort.1,4,17 One relatively recurrent variant in our cohort was the c.571G>A variant, which we detected in 8 unrelated individuals (also described in ClinVar VCV001330726.9). Although this variant is predicted to result in a missense mutation at the protein level (p.(Ala191Thr)), the variant affected the last nucleotide of exon 6, and RNA-sequencing demonstrated retention of intron 6 with a predicted stop codon after 11 residues (Figure 2A; SpliceAl, donor loss,  $\Delta$ score = 0.62). In addition, 4 synonymous variants were detected that were predicted to alter splicing. One synonymous variant (c.1230G>A, p.(Gln410Gln)) detected at ~50% VAF was strongly predicted to alter splicing (SpliceAl, donor loss,  $\Delta$  score = 0.82) and was observed in a patient with DDX41<sup>G/S</sup> pattern, with RNA-sequencing again demonstrating intron retention (Figure 2B). Finally, structural variant calling (GRIDSS) performed on the cohort identified 2 patients with novel multiexon (NC\_000005.9:g.176941255\_176942766del deletions and NC\_000005.9:g.176937300\_176939456del) (Figure 2C-D).<sup>18</sup>

## Pathogenicity classification of *DDX41* germ line variants

Germ line variants in the clinical laboratory context are typically classified using guidelines from the ACMG and AMP published in 2015.<sup>15</sup> These guidelines provide a general framework for aggregating evidence surrounding an individual variant to arrive at a likelihood that a variant is causative of a given phenotype. Using the ACMG/AMP guidelines, variants are classified as pathogenic (P), likely pathogenic (LP), or variant of uncertain significance (VUS) based on the strength of accumulated evidence. We classified the 105 different *DDX41* germ line variants identified in our cohort using the ACMG/AMP framework with standard modifications (supplemental Table 2). Using this approach, 30 of 105 variants were classified as LP or P, and 75 of 105 were classified as VUS (supplemental Table 3). Twenty-seven of 44 truncating variants were classified as LP/P, whereas only 3 of 61 nontruncating variants were classified as such.

Given the possible propensity for ACMG/AMP criteria to classify disease-causing variants in *DDX41* as VUSs because of incomplete penetrance, late-onset disease, the presence of multiple founder mutations, and a relative lack of understanding of the functional domains of *DDX41*, we investigated the effect of incorporating the observation of a somatic variant in *DDX41* as an independent criterion for the pathogenicity of a given co-observed germ line variant. For patients with AML/MDS, the observation of a somatic *DDX41* variant in our cohort was highly nonrandomly associated with the presence of a germ line variant (181/251 vs 31/4771 without), with an odds ratio of 395.4 (95% confidence interval, 252.6-618.9).

A previously described assessment of the ACMG/AMP 2015 criteria, by the ClinGen SVI, using a Bayesian probability framework<sup>19</sup> has proposed various categorization of strength (from "supporting" to "very strong") based on calculated odds of pathogenicity (OddsPath). Around the same time, the ClinGen Germline/Somatic Variant Subcommittee released guidance indicating that tumor features, such as second mutation of alternate allele, can be considered at least supporting evidence of pathogenicity but may be modified with increasing metadata. Based on this framework and guidance, the PP4 evidence code (patient's



Figure 1. Distribution of the identified variants along the *DDX41* gene in patients with MDS, AML, or MPN. Germ line variants were classified as LP/P or VUS based on our modified criteria (incorporating PP4\_very strong and the points-based rule system).



Figure 2. Genomic alterations in DDX41. (A) RNA-sequencing demonstrated retention of intron 6 with a predicted stop codon after 11 residues in a patient with variant c.571G>A, p.(Ala191Thr). (B) RNA-sequencing showed intron retention in the patient with the synonymous variant c.1230G>A, p.(Gln410Gln). Structural variant calling (GRIDSS) identified 2 patients with novel multiexon deletions: NC\_00005.9:g.176937300\_176939456del (C) and NC\_00005.9:g.176941255\_176942766del (D).

phenotype or family history is highly specific for a disease with a single genetic etiology) may be used in a modified manner after the assessment of a variant's somatic origin and alignment with Knudson two-hit hypothesis.<sup>20</sup> Given the strong nonrandom association of *DDX41* somatic variants occurring with deleterious germ line *DDX41* variants and the current understanding of the pathobiology of AML/MDS development in patients, we explored using the presence of a somatic *DDX41* variant as evidence for the pathogenicity of a given co-observed germ line variant. Regarding the strength of evidence applicable, we were able to show through proof, under minimal assumptions about the independence of variants at the population level (supplemental Material), that

$$\Pr(X_g | M_g \cap M_s \cap C) = \frac{1}{\Pr(M_g \cap M_s | C)} \begin{cases} 1 - (1) \\ -\Pr(M_g \cap M_s | C) \end{pmatrix} \frac{1 - \Pr(X_g | M_g) \Pr(M_g) \Pr(M_s)}{1 - \Pr(M_g) \Pr(M_s)} \end{cases}$$

in which

- M<sub>q</sub> is the event of having any germ line variant in DDX41;
- X<sub>g</sub> is the event of having a loss-of-function (deleterious) germ line variant in *DDX41*;
- M<sub>s</sub> is the event of having any somatic variant in DDX41;
- X<sub>s</sub> is the event of having a loss-of-function (deleterious) somatic variant in DDX41;
- C is the event of having AML/MDS; and

• Z is the event of having AML/MDS that is unrelated to the presence of *DDX41* variants.

Using this approach and a 2 × 2 contingency table derived from our cohort (Table 1; assuming a prior probability of 0.1), we calculated a posterior probability (ie,  $Pr(X_g|M_g \cap M_s \cap C)$ ) of 0.998, supporting the application of PP4 at a very strong level (OddsPath > 350) for the use of an observed somatic *DDX41* variant in the context of a germ line variant.

Incorporating this modified PP4 rule (including any observations of a germ line variant alongside an assumed somatic variant in the peerreviewed literature) as well as incorporating a previously proposed points-based rules system for variant classification,<sup>21</sup> 38 of 44 truncating variants and 30 of 61 missense variants were classified as LP/P (supplemental Table 3). Using the same approach but arbitrarily capping the strength applied to this code as an independent strong level criterion (PP4\_strong; to ensure that further evidence types are also required for a variant to reach an LP/P classification) resulted in 37 of 44 truncating variants and 25 of 61 missense variants being classified as LP/P (supplemental Table 3).

 
 Table 1. Contingency table of association between germ line and somatic DDX41 variants in a cohort of patients with MDS/AML

	Presence of somatic DDX41 variant?		
	Yes	No	
Presence of germ line DDX41 variant?			
Yes	181	70	
No	31	4740	

Using our modified PP4 rule (PP4\_very strong), the MDS cohort contained 108 of 2865 patients (3.8%) with germ line LP/P variants ( $DDX41^{G}$ , n = 20;  $DDX41^{G/S}$ , n = 88), the AML cohort 106 of 2157 (4.9%;  $DDX41^{G}$ , n = 13;  $DDX41^{G/S}$ , n = 93), and the MPN cohort 3 of 715 (0.4%;  $DDX41^{G}$ , n = 3). Using the more conservative approach for the modified PP4 rule (PP4\_strong), the MDS cohort contained 99 of 2865 patients (3.5%) with germ line LP/P variants ( $DDX41^{G}$ , n = 18;  $DDX41^{G/S}$ , n = 81), the AML cohort 104 of 2157 (4.8%) ( $DDX41^{G}$ , n = 11;  $DDX41^{G/S}$ , n = 93), and the MPN cohort 2 of 715 (0.3%) ( $DDX41^{G}$ , n = 2). Variant details are present in supplemental Tables 4 to 6.

### Characteristics of *DDX41* somatic mutations and comutated genes

*DDX41* comutation most commonly occurred in the form of a truncating (nonsense, frameshift, multiexon deletion, start loss, canonical splice site, synonymous/missense variants with splice effect, and splice region variants with splice effect) germ line variant along with a nontruncating acquired variant in patients with DDX41<sup>G/S</sup> (129 of 181) and in the form of a truncating variant plus a nontruncating variant in patients with *DDX41<sup>S/S</sup>* (7 of 11). Two nontruncating variants were observed in 52 of 181 patients with DDX41<sup>G/S</sup> and in 4 of 11 patients with *DDX41<sup>S/S</sup>*. No patients were observed with 2 truncating *DDX41* variants. The non-truncating acquired *DDX41* variants were clustered in the helicase and DEAD-box domains with the most frequently observed

acquired variant being p.(Arg525His) (n = 133) followed by p.(Gly530Asp) (n = 21) (supplemental Table 3).

For all patients with  $DDX41^{G/S}$  in whom the DDX41 variants were close enough to undergo phase analysis (n = 11), variants were confirmed to be present on different alleles, consistent with the current model of biallelic DDX41 dysfunction contributing to leukemogenesis (an example is shown in Figure 3).

Out of 108, 29 patients (26.8%) with MDS carrying a germ line LP/ P *DDX41* variant as per our modified criteria (using somatic second hit as a very strong criteria) had no further genetic variants in other established MDS-related genes. In those with additional mutations, the most frequently mutated genes were *ASXL1* (20.2%), *SRSF2* (17.6%), *DNMT3A* (15.7%), and *TET2* (13.3%). Compared with *DDX41*-wild-type (WT) cases, patients with germ line LP/P *DDX41* variants showed significantly fewer mutations in *TET2* and *SF3B1* (P < .001; Figure 4A).

Out of 106 patients with AML, 30 (28.3%) carrying a germ line LP/ P *DDX41* variant as per our modified criteria harbored no further genetic variants in other established AML-related genes. In those with additional mutations, the most frequently mutated genes were *ASXL1* (19.8%), *SRSF2* (17.9%), *DNMT3A* (17.0%), and *EZH2* (11.4%). In comparison with *DDX41*-WT patients, multiple genes (*NPM1*, *RUNX1*, *FLT3*, *IDH2*, *NRAS*, and *IDH1*) were significantly less often mutated, and *CUX1* was significantly more often mutated in patients carrying a germ line LP/P variant (P < .001;

Figure 3. Phase analysis of sequencing from patient MDS-P124 demonstrating the germ line variant c.1585dup/p.(Thr529Asnfs\*13) and the somatic variant c.1574G>A/p.(Arg525His) demonstrating their presence in different reads/alleles. The germ line variant c.1585dup/p.(Thr529Asnfs\*13) is indicated by the purple bar, and the somatic variant c.1574G>A/p.(Arg525His) is

indicated by a red T.





Figure 4. Mutation frequency of MDS- and AML-related genes sorted by descending frequency in patients with DDX41-WT. Mutation frequencies in patients with DDX41-WT MDS are illustrated in red and in patients with MDS with germ line LP/P DDX41 variant in blue (A). Mutation frequencies in patients with DDX41-WT AML are illustrated in red and in patients with AML with germ line LP/P DDX41 variant in blue (B). \*P < .001 (Fisher exact test).

Figure 4B). Thus, patients with MDS/AML harboring germ line LP/ P variants as per our modified criteria have a unique pattern of comutation compared with patients with *DDX41*-WT, providing support for this approach to identifying a unique biological entity.

#### Clinicopathological features of patients with MDS/ AML with germ line LP/P DDX41 variants

Characteristics of patients with *DDX41*-WT and patients with germ line LP/P variants in *DDX41* are summarized in Table 2 for patients with MDS and AML. We observed a male predominance in our cohort, with 163 of 215 (75.8%) patients with MDS/AML with germ line LP/P *DDX41* variants being male. In MDS and AML, the median age of patients with germ line LP/P *DDX41* variants was 72 years (MDS: range, 47-92 years; AML: range, 47-93 years). Most patients with MDS had increased blasts at diagnosis, with 40.7% being classified as MDS-EB-2 and 25.0% being classified as MDS-EB-1. In comparison with patients with *DDX41*-WT with MDS, patients with the germ line LP/P *DDX41* variant were significantly more often diagnosed with MDS-EB-2 and significantly less often with MDS-RS-MLD (P < .001; Table 2). In addition, patients with *DDX41*-WT with AML and MDS had statistically significantly higher

	Patients with MDS		Patients with AML					
	Total cohort (n = 2865)	DDX41 germ line LP/P (n = 108)	<i>DDX41</i> -WT (n = 2720)	Р	Total cohort (n = 2157)	DDX41 germ line LP/P (n = 106)	<i>DDX41-</i> WT (n = 2020	Р
Male/female (ratio)	1824/1041 (1.8)	84/24 (3.5)	1713/1007 (1.7)	NS	1243/914 (1.4)	79/27 (2.9)	145/875 (1.3)	<.001
Median age, y (range)	76 (15-96)	72 (47-92)	76 (15-96)	NS	72 (4-97)	72 (47-93)	71.4 (4-97)	NS
Cytomorphological diagnosis (WHO 2017), n (%)	Total cohort (n = 2865)	DDX41 germ line LP/P (n = 108)	<i>DDX41</i> -WT (n = 2720)	Р	Total cohort (n = 2157)	DDX41 germ line LP/P (n = 106)	<i>DDX41</i> -WT (n = 2020)	Р
MDS	337 (11.8)	18 (16.7)	311 (11.4)	NS				
MDS with isolated del(5q)	113 (3.9)	1 (0.9)	111 (4.1)	NS				
MDS-EB-1	798 (27.9)	27 (25.0)	758 (27.9)	NS				
MDS-EB-2	591 (20.6)	44 (40.7)	537 (19.7)	<.001				
MDS-MLD	388 (13.5)	15 (13.9)	371 (13.6)	NS				
MDS-RS-MLD	568 (19.8)	3 (2.8)	563 (20.7)	<.001				
MDS-SLD	10 (0.3)	0	10 (0.4)	NS				
MDS-RS-SLD	46 (1.6)	0	45 (1.7)	NS				
MDS-U	14 (0.5)	0	14 (0.5)	NS				
Acute monoblastic/monocytic leukemia					56 (2.6)	0	56 (2.8)	NS
Acute myelomonocytic leukemia					158 (7.3)	1 (0.9)	153 (7.6)	NS
AML					368 (17.1)	3 (2.8)	360 (17.8)	<.001
AML with CBFB::MYH11					31 (1.4)	0	30 (1.5)	NS
AML with maturation					513 (23.8)	44 (41.5)	462 (22.9)	<.001
AML with minimal differentiation					178 (8.3)	20 (18.9)	156 (7.7)	<.001
AML with myelodysplasia-related changes					277 (12.8)	11 (10.4)	263 (13.0)	NS
AML without maturation					360 (16.7)	12 (11.3)	341 (16.9)	NS
s-AML arising from previous MDS					216 (10.0)	15 (14.2)	199 (9.9)	NS
CBC parameters, n	Total cohort (n = 2865)	DDX41 germ line LP/P (n = 108)	<i>DDX41</i> -WT (n = 2720)	Р	Total cohort (n = 2157)	DDX41 germ line LP/P (n = 106)	<i>DDX41</i> -WT (n = 2020)	Р
Median WBC count, × 10 <sup>3</sup> /µL (range)	3.4 (0.1-117.4); 2160	2.2 (0.8-9.1); 86	3.5 (0.1-117.4); 2045	<.001	5.9 (0.1-493.2); 1778	1.8 (0.6-104.2); 87	6.9 (0.1-493.2); 1671	<.001
Median Hb level, g/dL (range)	9.4 (2.8-17.0); 2089	9.7 (5.5-14.5); 84	9.4 (2.8-17.0); 1977	NS	8.9 (1.5-17.7); 1658	9.6 (2.9-14.6); 83	8.8 (1.5-17.7); 1556	<.001
Median platelet count, × 10 <sup>3</sup> /µL (range)	113.0 (1.0-1,391.0); 2077	96.0 (12.0-336.0); 83	114.0 (1.0-1391.0); 1966	<.001	70.0 (1.0-1,505.0); 1660	64.0 (7.0-256.0); 85	70.0 (1.0-1505.0); 155	6 <.001
Median blasts bone marrow, % (range)	4.5 (0-19.5); 2670	8.0 (0-19.5); 104	4.5 (0-19.5); 2532	<.001	46.0 (6.0-99.5); 1783	32.0 (20.0-89.5); 104	48.0 (6.0-99.5); 1650	<.001

P values are given for significant differences (P<.001) between patients with DDX41-WT and patients with germ line LP/P DDX41 variants. Patients carrying only somatic mutations or a germ line VUS in DDX41 are not listed separately but are included in the total cohort.

CBC, complete blood count; EB, excess blasts; NS, not significant; RS, ring sideroblasts; s-AML, secondary-acute myeloid leukaemia; SLD, single lineage dysplasia; WBC, white blood cell; WHO, World Health Organization.

Table 2 (continued)								
Cytogenetics, n (%)	Total cohort (n = 2637)	<i>DDX41</i> germ line LP/P (n = 96)	<i>DDX41</i> -WT (n = 2509)	d	Total cohort (n = 1971)	DDX41 germ line LP/P (n = 97)	<i>DDX41</i> -WT (n = 1848)	٩
Normal karyotype	1569 (59.5)	79 (82.3)	1467 (58.5)	<.001	928 (47.1)	73 (75.3)	837 (45.3)	<.001
Aberrant karyotype	1068 (40.5)	17 (17.7)	1042 (41.5)		1043 (52.9)	24 (24.7)	1011 (54.7)	
Trisomy 8	150 (5.7)	1 (1.0)	148 (5.9)		131 (6.6)	5 (5.2)	126 (6.8)	
chr7 aberration	64 (2.4)	0	64 (2.6)		76 (3.9)	1 (1.0)	74 (4.0)	
Loss of Y-chromosome	109 (4.1)	7 (7.3)	100 (4.0)		25 (1.3)	1 (1.0)	23 (1.2)	
t(8;21)(q22;q22.1)	1 (0.0)	0	1 (0.0)		27 (1.4)	0	27 (1.5)	
inv(16)(p13.1q22)/ t(16;16)(p13.1;q22)	0	0	0		48 (2.4)	0	47 (2.5)	
del(5q)	182 (6.9)	3 (3.1)	175 (7.0)		52 (2.6)	0	51 (2.8)	
Complex	240 (9.1)	1 (1.0)	238 (9.5)		308 (15.6)	7 (7.2)	300 (16.2)	
All other	322 (12.2)	5 (5.2)	316 (12.6)		376 (19.1)	10 (10.3)	363 (19.6)	
P values are given for significant diffe are included in the total cohort. CBC, complete blood count; EB, ex	irences ( <i>P</i> < .001) between pati cess blasts; NS, not significant;	ents with <i>DDX41</i> -WT and patie ; RS, ring sideroblasts; s-AML,	ints with germ line LP/P DDX41 secondary-acute myeloid leuka	variants. F emia; SLI	<sup>a</sup> tients carrying only somatic ), single lineage dysplasia; W	mutations or a germ line VUS /BC, white blood cell; WHO	in <i>DDX41</i> are not listed separ , World Health Organization.	ately but

and lower blast counts, respectively, than patients with the germ line LP/P *DDX41* variants. Among patients with AML, most of the patients with the germ line LP/P *DDX41* variant were diagnosed with AML with maturation (41.5%) and AML with minimal differentiation (18.9%), which was significantly more frequently than in patients with *DDX41*-WT (P < .001; Table 2).

Cytogenetic analyses were available for 96 of 108 patients with MDS harboring germ line LP/P *DDX41* variants. A fraction (17.7%) had an abnormal karyotype, including 7 with loss of chromosome Y, 3 with del(5q), and 1 with a complex karyotype (Table 2). In the AML cohort, cytogenetic analyses were performed for 97 of 106 patients harboring germ line LP/P *DDX41* variants, with 24.7% having an abnormal karyotype, including 7 with a complex karyotype (Table 2). Thus, most patients with germ line LP/P *DDX41* variants had a normal karyotype (MDS: 81.4% and AML: 75.3%), which differs significantly from those with *DDX41*-WT, among whom only 58.5% (MDS, *P* < .001) or 45.3% (AML, *P* < .001) had a normal karyotype.

## Spectrum of hematologic malignancy associated with *DDX41* variants

We then investigated the enrichment of DDX41 variants across different hematologic neoplasms using DDX41 variants classified as LP/P by our modified ACMG/AMP criteria. We compared the frequency of DDX41 LP/P variants in patients with a germ line variant (DDX41<sup>G/S</sup> or DDX41<sup>G</sup>) in AML, MDS, and MPN in our cohort with that of all individuals from the gnomAD database (n  $\approx$ 125 000). This analysis revealed a marked enrichment of DDX41 variants in patients with AML and MDS and a smaller enrichment in patients with MPN. We then compared gnomAD DDX41 LP/P variant frequency with (1) 954 patients with B-cell neoplasm (chronic lymphocytic leukemia, n = 317; mantle cell lymphoma, n =94; hairy cell leukemia/hairy cell leukemia variant/splenic marginal zone lymphoma, n = 215; lymphoplasmacytic lymphoma, n = 65; and multiple myeloma, n = 65) and (2) 453 patients with acute lymphoblastic leukemia (B- cell ALL, n = 321 and T- cell ALL, n = 132) detected via WGS. There was a statistically significant enrichment in cases of B-cell neoplasm but not in T- or B-cell ALL. Notably, no patient in the MPN, B-cell neoplasm, or ALL group had an assumed somatic variant observed in conjunction with their germ line LP/P DDX41 variant. The relative incidence of DDX41 variants is shown in Table 3.

### **Discussion**

The detection and subsequent management of patients with *DDX41* variants is an emerging and pressing clinical issue. Through sequencing a large cohort of sequentially recruited patients in the clinical diagnostic laboratory, we have expanded our current understanding of the landscape of genomic abnormalities in *DDX41*, the hematologic malignancy spectrum associated with *DDX41* variants as well as proposed ACMG/AMP-classified alterations that may aid variant interpretation and subsequent clinical decision-making. We also confirmed the genomic findings of other studies; the most common comutated genes were *ASXL1*, *TP53*, *EZH2*, *CUX1*, and *SRSF2*.<sup>4,22,23</sup> In the study by Makishima et al, patients harboring *DDX41* variants showed an overrepresentation of *GNAS* and *CUX1* mutations and an underrepresentation of *STAG2*, *NRAS*, *NPM1*, *SF3B1*, and *TET2* mutations.<sup>4</sup> With the

Table 3. Comparison of DDX41 P/LF	variants in cohorts of hemato	logic malignancies when	compared with health	v controls (gnomAD)

	DDX41 <sup>G</sup> or DDX41 <sup>G/S</sup> (LP/P)/DDX41-WT	P value (compared with gnomAD)	Proportion of DDX41 <sup>G/S</sup> pattern
Population database (gnomAD)	77/125 000* (0.06%)	N/A	N/A
AML	106/2157 (4.9%)	<.00001 (χ <sup>2</sup> )	93/106 (87.7%)
MDS	108/2865 (3.8%)	<.00001 (χ <sup>2</sup> )	88/108 (81.5%)
MPN	3/741 (0.4%)	.01 (Fisher)	0/3 (0%)
B-cell neoplasm	8/954 (0.84%)	.00001 (Fisher)	0/954 (0%)
Acute lymphoblastic leukemia	1/453 (0.22%)	.25 (Fisher)	0/453 (0%)
N/A, not applicable.			

\*Approximate value.

Approximate value.

exception of *GNAS*, these trends could also be seen in our patients with AML/MDS.

In our cohort, we have expanded the current genomic landscape of DDX41 germ line variants both guantitatively and gualitatively. Firstly, we have described multiple novel single nucleotide variants and small insertion and deletions in the somatic setting, including p.(Ala346Val), p.(Ser405\_Val408del), and p.(Tyr451Thrfs\*10) as well as the nonsense variants p.(Gly72\*), p.(Glu118\*), and p.(Lys187\*) and canonical splice variants c.572-2A>G, c.799-1G>A, and c.799-1G>T in the germ line setting. Interestingly, we detected highly recurrent variants in our population that have not been described as frequently in other large cohorts (eg, c.571G>A), consistent with a relative geographic/ethnic restriction of variants. In addition, our data demonstrate the importance of including assessment of all variant types when testing for DDX41 variants, given our detection of synonymous variants affecting splicing (as demonstrated by RNA-sequencing) as well as noncanonical splice sites predicted to affect splicing and multiexon deletions resulting in predicted protein truncation. These types of variants may be either bioinformatically or technically excluded by current technologies and approaches typically used by diagnostic laboratories. Moreover, future studies should further assess the noncoding/regulatory regions of DDX41 in order to fully understand the spectrum of variants causative of the germ line DDX41 predisposition to a hematologic malignancy.

The management of any patient with a myeloid malignancy associated with the presence of a DDX41 variant is critically dependent on the likelihood that the variant is causal to the observed phenotype. Although the generally accepted framework for this is the ACMG/AMP guidelines published in 2015,<sup>15</sup> these guidelines are currently undergoing a process of gene/clinical contextspecific modifications by variant curation expert panels in order to improve the utility of these guidelines for specific disease and gene contexts, such as has occurred for RUNX1.24 Our data suggest that use of the ACMG/AMP guidelines without modification results in a possible overclassification of variants into the VUS category. Classifying pathogenic variants as VUS has potential significant clinical consequences, including the failure to exclude these variants from sibling allogeneic transplant donors. We have proposed a model accompanied by statistical proof of modifying the ACMG/AMP PP4 criterion for its application at a very strong level based on the highly specific occurrence of DDX41 somatic variants in the context of DDX41 germ line predisposition. This approach could be further developed by potentially integrating the number of times the variant has been observed with a somatic DDX41 variant as a proportion of total observations of the variant in the context of patients with a diagnosis of AML/MDS as determining strength level (eg, greater or less than a threshold percentage resulting in very strong or strong, respectively). In addition, it may also be reasonable to require more than 1 observation of a somatic variant (or specific recurrent somatic variants) in the context of a given germ line variant to further reduce the likelihood of coincidental co-occurrence. The Met155lle variant in this context merits specific mention as it has been observed in our cohort in the context of a somatic DDX41 variant; however, it is more commonly observed without a somatic variant. Given its relatively high population frequency (gnomAD = 0.002%), it may be that it represents a true pathogenic, but low penetrance, allele. Larger studies specifically focusing on this variant would be of value to further understand these complexities.

Although we have modified an existing ACMG/AMP evidence code (PP4), another approach would be to consider this evidence as a new criterion. Of note, other groups have chosen to alter other ACMG/AMP criteria,<sup>23</sup> and there are pragmatic benefits to modifying existing criteria in terms of integration into existing clinical informatics software and frameworks. In addition, PP4 has been previously proposed as an appropriate criterion to alter for this purpose.<sup>20</sup> Our data supports the modification of PP4 (patient's phenotype or family history is highly specific for a disease with a single genetic etiology) at a very strong level, given the specificity of the finding of the second hit for the germ line *DDX41* disease. Finally, given that our model simply requires a  $2 \times 2$  contingency table, this approach could be considered for other gene-disease combinations in which somatic mutations are specific for a clinical phenotype (eg, *SAMD9/SAMD9L* somatic revertant mutations).

Precisely which hematologic neoplasms are predisposed to by the presence of *DDX41* germ line variants is unclear. We used WGS from across a range of hematologic malignancies and showed that AML/MDS is the highest risk subtype of hematologic neoplasm predisposed to by the presence of LP/P *DDX41* variants. Although there was a statistically significant enrichment of cases observed in MPN and B-cell neoplasm, the absolute numbers were significantly lower, and when LP/P *DDX41* variants were observed in the context of MPN and lymphoid malignancy, they were not accompanied by somatic *DDX41* variants in our cohort, suggestive of a different biological basis to AML/MDS, in which *DDX41* somatic variants are observed in most cases with germ line LP/P variants. Given the relatively small numbers in the non-AML/MDS cohorts,

further studies are required before any magnitude of increased risk for these subtypes of hematologic malignancy is understood, and despite the statistical significance, it is not clear what biological role this lower level of enrichment for monoallelic pathogenic variants is playing in relation to the development of these diseases. However, given our observations and the broader literature to date, a re-evaluation of the current description of the phenotype in the Online Mendelian Inheritance in Man (myeloproliferative/lymphoproliferative neoplasms, familial (multiple types), susceptibility to; 616871) database may be appropriate.

In summary, we have further expanded the landscape of genomic variation in *DDX41* and our understanding of the predisposition to hematologic malignancy. Our proposed curation framework modifications, along with future refinements, may be of use to diagnostic laboratories currently encountering these variants in daily practice to provide optimal advice to clinical teams managing these patients and families.

#### **Acknowledgments**

The authors acknowledge David Nott (National University of Singapore) for his critical appraisal of the Bayesian model. The authors thank all coworkers at the MLL for their dedicated work and all the physicians who provided samples, cared for patients, and collected data.

This work was partially supported by a Cancer Center Support grant from the National Institutes of Health/National Cancer Institute to Memorial Sloan Kettering Cancer Center (P30 CA008748).

### Authorship

Contribution: P.B. and A.M. designed the study; S.H. and N.N. were responsible for bioinformatic analyses; C.B., C.P., W.K., C.H., M.M., and T.H. were responsible for diagnostics and for clinical data interpretation; P.B., N.M., E.R.T., and A.M. were responsible for *DDX41* variant interpretation; R.A.C. and P.A.J. developed the statistical approach within the established Bayesian framework; P.B. and A.M. drafted the manuscript; and all authors read and contributed to the final version of the manuscript.

Conflict-of-interest disclosure: C.H., W.K., and T.H. declare part ownership of MLL. A.M., S.H., C.B., N.N., C.P., and M.M. are employed by MLL. The remaining authors declare no competing financial interests.

ORCID profiles: N.M., 0000-0003-1465-2934; R.A.C., 0000-0002-9847-1710; S.H., 0000-0003-2432-8898; P.A.J., 0000-0002-4361-4657.

Correspondence: Piers Blombery, Haematology, Peter Mac-Callum Cancer Centre, 305 Grattan St, Melbourne, VIC 3000, Australia; email: piers.blombery@petermac.org.

### References

- 1. Sebert M, Passet M, Raimbault A, et al. Germline DDX41 mutations define a significant entity within adult MDS/AML patients. *Blood.* 2019;134(17): 1441-1444.
- Duployez N, Largeaud L, Duchmann M, et al. Prognostic impact of DDX41 germline mutations in intensively treated acute myeloid leukemia patients: an ALFA-FILO study. *Blood.* 2022;140(7):756-768.
- 3. Gibson CJ, Kim HT, Zhao L, et al. Donor clonal hematopoiesis and recipient outcomes after transplantation. J Clin Oncol. 2022;40(2):189-201.
- 4. Makishima H, Saiki R, Nannya Y, et al. Germ line DDX41 mutations define a unique subtype of myeloid neoplasms. Blood. 2023;141(5):534-549.
- 5. Polprasert C, Schulze I, Sekeres MA, et al. Inherited and somatic defects in DDX41 in myeloid neoplasms. Cancer Cell. 2015;27(5):658-670.
- Raczy C, Petrovski R, Saunders CT, et al. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics*. 2013; 29(16):2041-2043.
- 7. Dunn T, Berry G, Emig-Agius D, et al. Pisces: an accurate and versatile variant caller for somatic and germline next-generation sequencing data. *Bioinformatics*. 2019;35(9):1579-1581.
- 8. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*. 2009;25(21):2865-2871.
- 9. Cameron DL, Schroder J, Penington JS, et al. GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Res.* 2017;27(12):2050-2060.
- 10. Bendig S, Walter W, Meggendorfer M, et al. Whole genome sequencing demonstrates substantial pathophysiological differences of MYC rearrangements in patients with plasma cell myeloma and B-cell lymphoma. *Leuk Lymphoma*. 2021;62(14):3420-3429.
- 11. Haferlach T, Hutter S, Meggendorfer M. Genome sequencing in myeloid cancers. N Engl J Med. 2021;384(25):e106.
- 12. Stengel A, Shahswar R, Haferlach T, et al. Whole transcriptome sequencing detects a large number of novel fusion transcripts in patients with AML and MDS. *Blood Adv.* 2020;4(21):5393-5401.
- 13. Walter W, Shahswar R, Stengel A, et al. Clinical application of whole transcriptome sequencing for the classification of patients with acute lymphoblastic leukemia. *BMC Cancer*. 2021;21(1):886.
- 14. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting splicing from primary sequence with deep learning. *Cell*. 2019;176(3): 535-548.e24.
- 15. 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.

- 16. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat.* 2018;39(11):1517-1524.
- 17. Choi EJ, Cho YU, Hur EH, et al. Unique ethnic features of DDX41 mutations in patients with idiopathic cytopenia of undetermined significance, myelodysplastic syndrome, or acute myeloid leukemia. *Haematologica*. 2022;107(2):510-518.
- Maierhofer A, Fox LC, Hutter S, et al. Detection of novel occult germline multi-exon deletions in patients with DDX41 familial predisposition to myeloid malignancy. *Blood.* 2022;140(suppl 1):8676-8677.
- 19. 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.
- 20. Walsh MF, Ritter DI, Kesserwan C, et al. Integrating somatic variant data and biomarkers for germline variant classification in cancer predisposition genes. *Hum Mutat.* 2018;39(11):1542-1552.
- 21. Tavtigian SV, Harrison SM, Boucher KM, Biesecker LG. Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. *Hum Mutat.* 2020;41(10):1734-1737.
- 22. Qu S, Li B, Qin T, et al. Molecular and clinical features of myeloid neoplasms with somatic DDX41 mutations. Br J Haematol. 2021;192(6):1006-1010.
- 23. Li P, White T, Xie W, et al. AML with germline DDX41 variants is a clinicopathologically distinct entity with an indolent clinical course and favorable outcome. *Leukemia*. 2022;36(3):664-674.
- 24. Feurstein S, Luo X, Shah M, et al. Revision of RUNX1 variant curation rules. Blood Adv. 2022;6(16):4726-4730.