blood advances

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodadvances@hematology.org

Double mutant DNMT3A AML: a unique subtype experiencing increased DNA damage and poor prognosis

Tracking no: ADV-2024-014698R1

Emma Boertjes (Erasmus University Medical Center, Netherlands) Sanne Massaar (ErasmusMC, Netherlands) Annelieke Zeilemaker (Erasmus MC, Netherlands) Jolinda Konijnenburg (Erasmus MC, Netherlands) Melissa Rijken (Erasmus MC, Netherlands) François Kavelaars (Erasmus MC, Netherlands) Tim Grob (Erasmus University Medical Center, Netherlands) Jurjen Versluis (Erasmus MC, Netherlands) Bob Löwenberg (Department of Hematology, Erasmus MC, Rotterdam, The Netherlands, Netherlands) Peter Valk (Erasmus University Medical Center, Netherlands) Mathijs Sanders (Erasmus University Medical Center, Netherlands)

Abstract:
Mutation of DNMT3A, encoding a de novo methyltransferase essential for cytosine methylation, is a common early event in clonal hematopoiesis (CH) and adult acute myeloid leukemia (AML). Spontaneous deamination of methylated cytosines damages DNA, which is repaired by the base excision repair (BER) enzymes MBD4 and TDG. Congenital MBD4-deficiency has been linked to early-onset CH and AML, and is marked by exceedingly high levels of DNA damage and mutation of DNMT3A. Strikingly, wildtype (WT) DNMT3A binds TDG, thereby potentiating its repair activity. Since TDG is the only remaining BER enzyme in MBD4-deficient AML patients capable of repairing methylation damage, we investigated whether mutant DNMT3A negatively affects the repair function of TDG. We found that, whereas WT DNMT3A stimulates TDG function, mutant DNMT3A impairs TDG-mediated repair of DNA damage in vitro. In light of this finding and to extrapolate our observations to the broader AML patient population, we investigate here the genetic profiles and survival outcomes of AML patients with single (SM) versus double mutant (DM) DNMT3A. DM DNMT3A AML patients show a characteristic driver mutation landscape and reduced overall survival when compared to SM DNMT3A AML patients. Importantly, whole-genome sequencing showed a trend for increased DNA damage in primary DM DNMT3A AML samples, especially when DNMT3A mutations are located at the DNMT3A-TDG interaction interface.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: P.J.M.V. and M.A.S. designed the research; P.J.M.V., B.L. provided patient information and materials; E.L.B., A.Z., M.R. and F.G.K. performed experiments; E.L.B, S.M. and M.A.S and J.K. analyzed data; T.G. and J.V. provided statistical advice; E.L.B and S.M. prepared the figures. E.L.B., S.M., P.J.M.V. and M.A.S. wrote the manuscript. All authors approved the final version of the manuscript.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Data will be made publicly available in the European genome-phenome Archive.

Clinical trial registration information (if any):

Figure 1

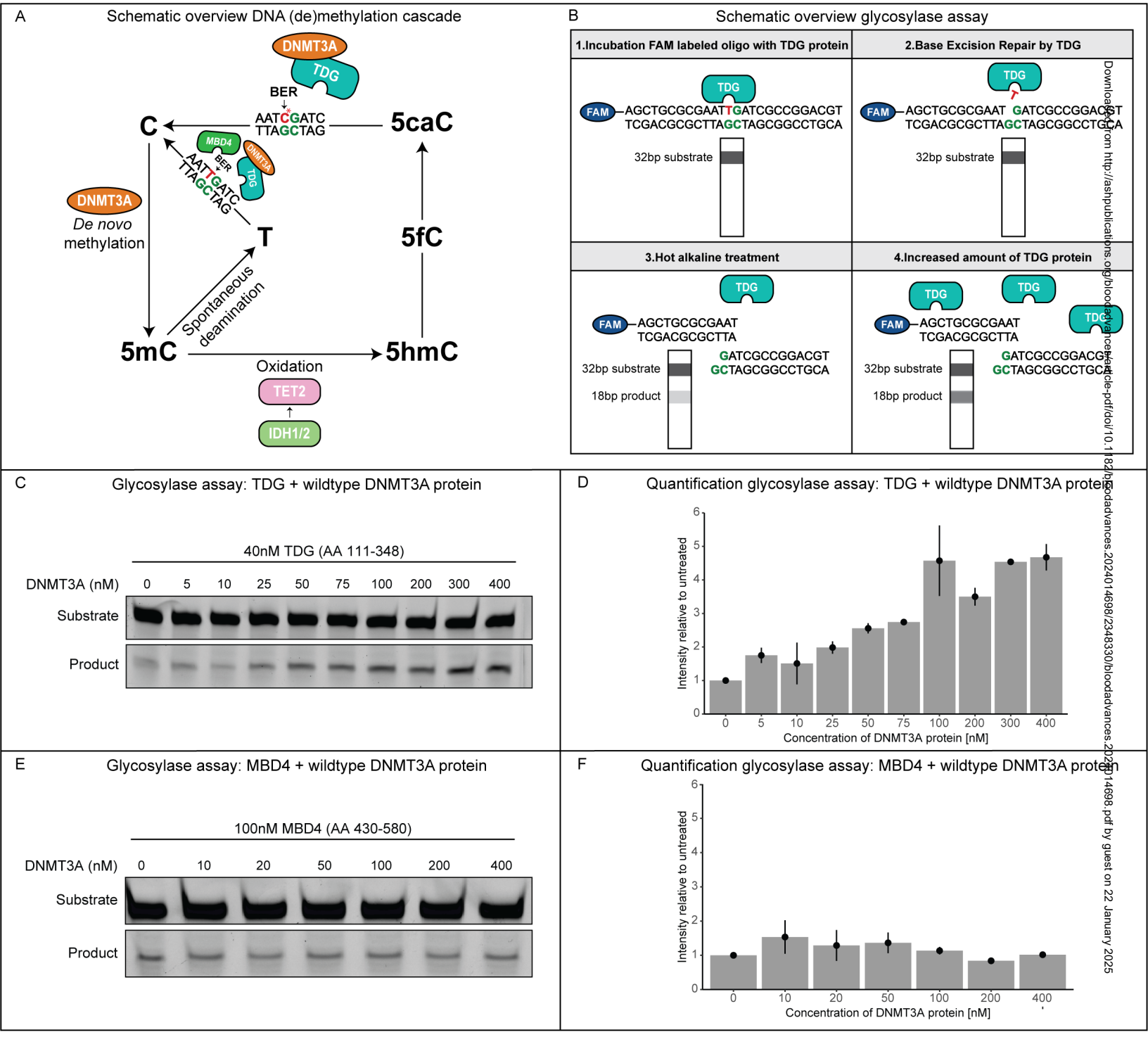


Figure 2

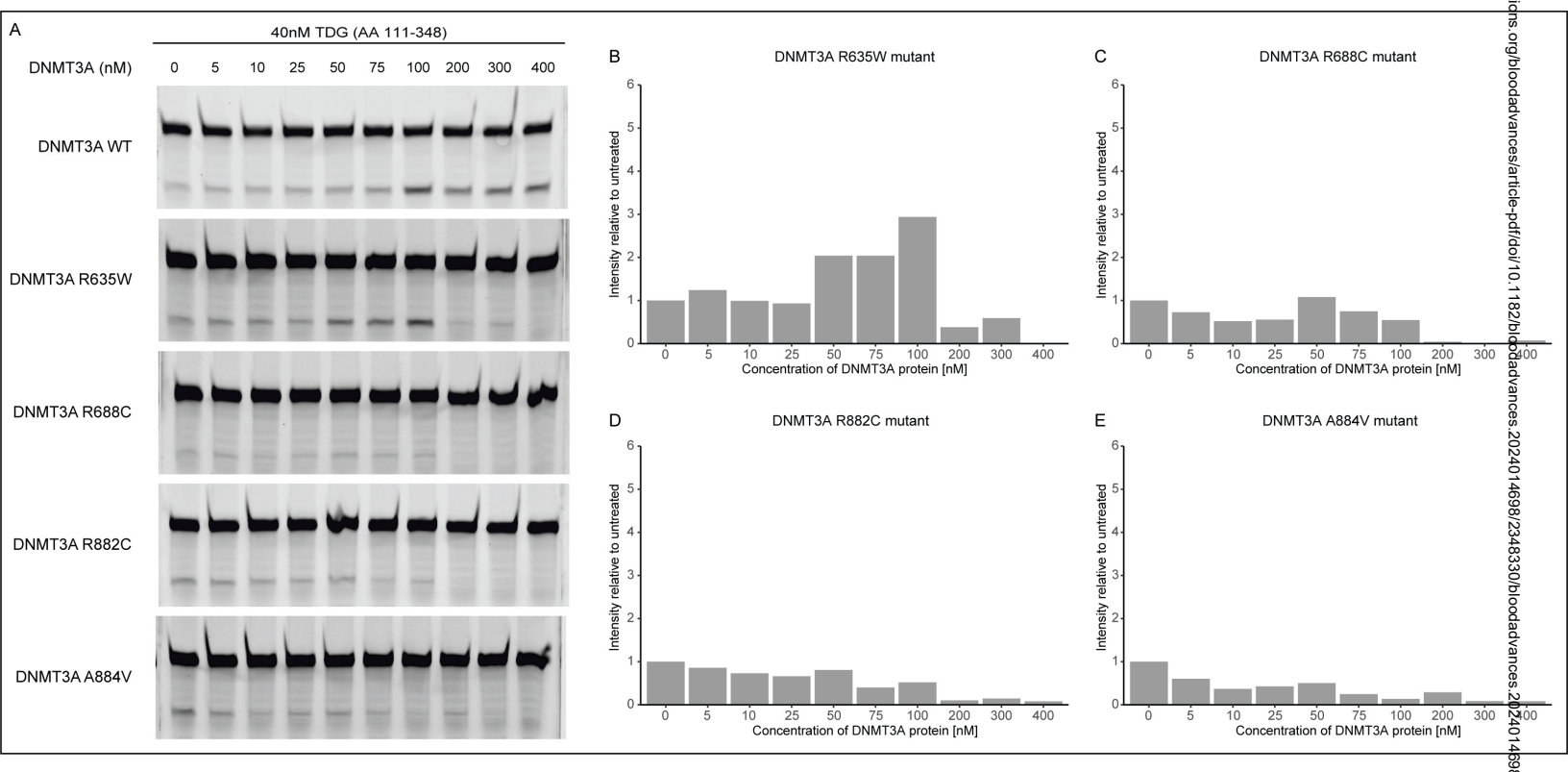


Figure 3

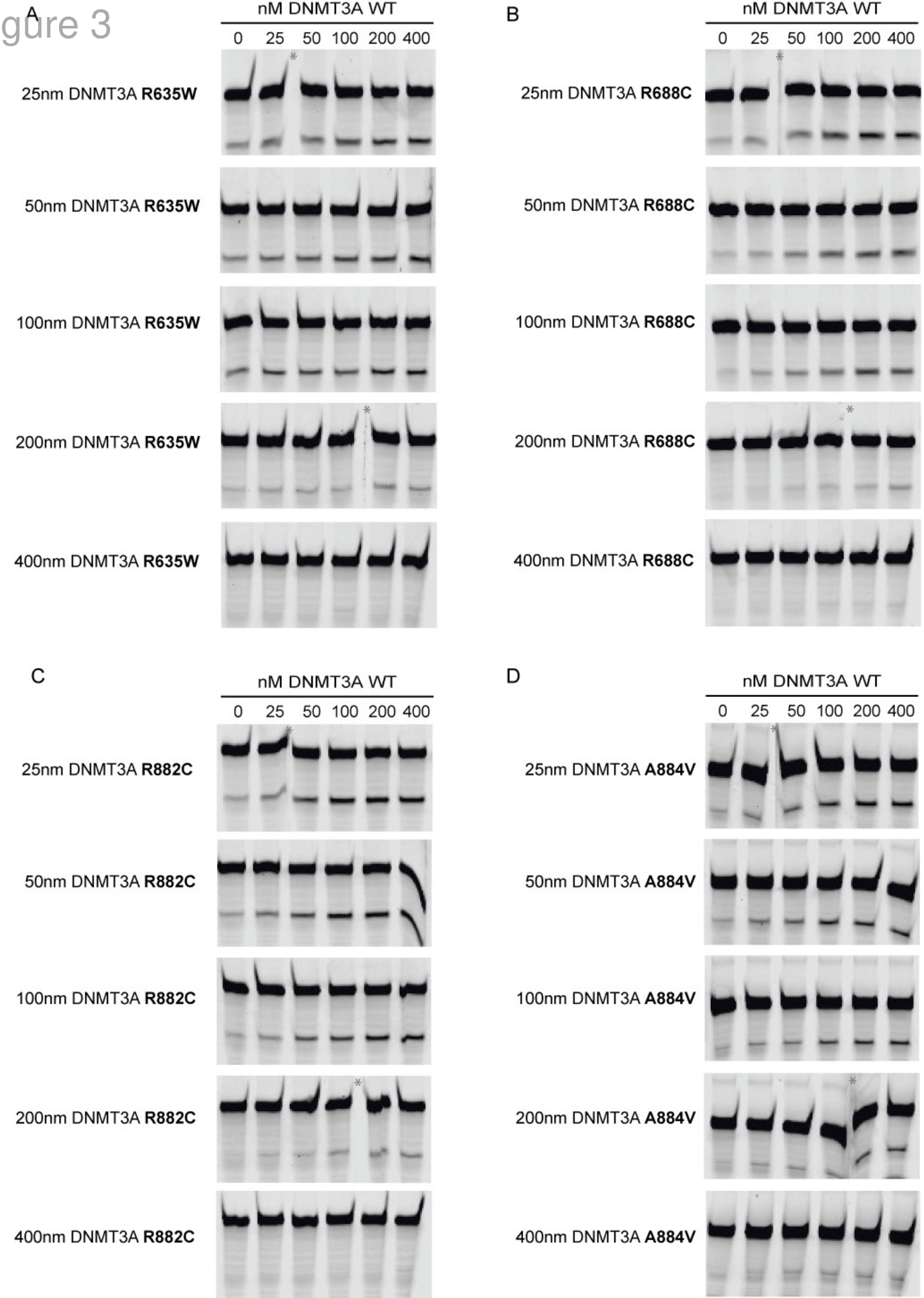


Figure 4

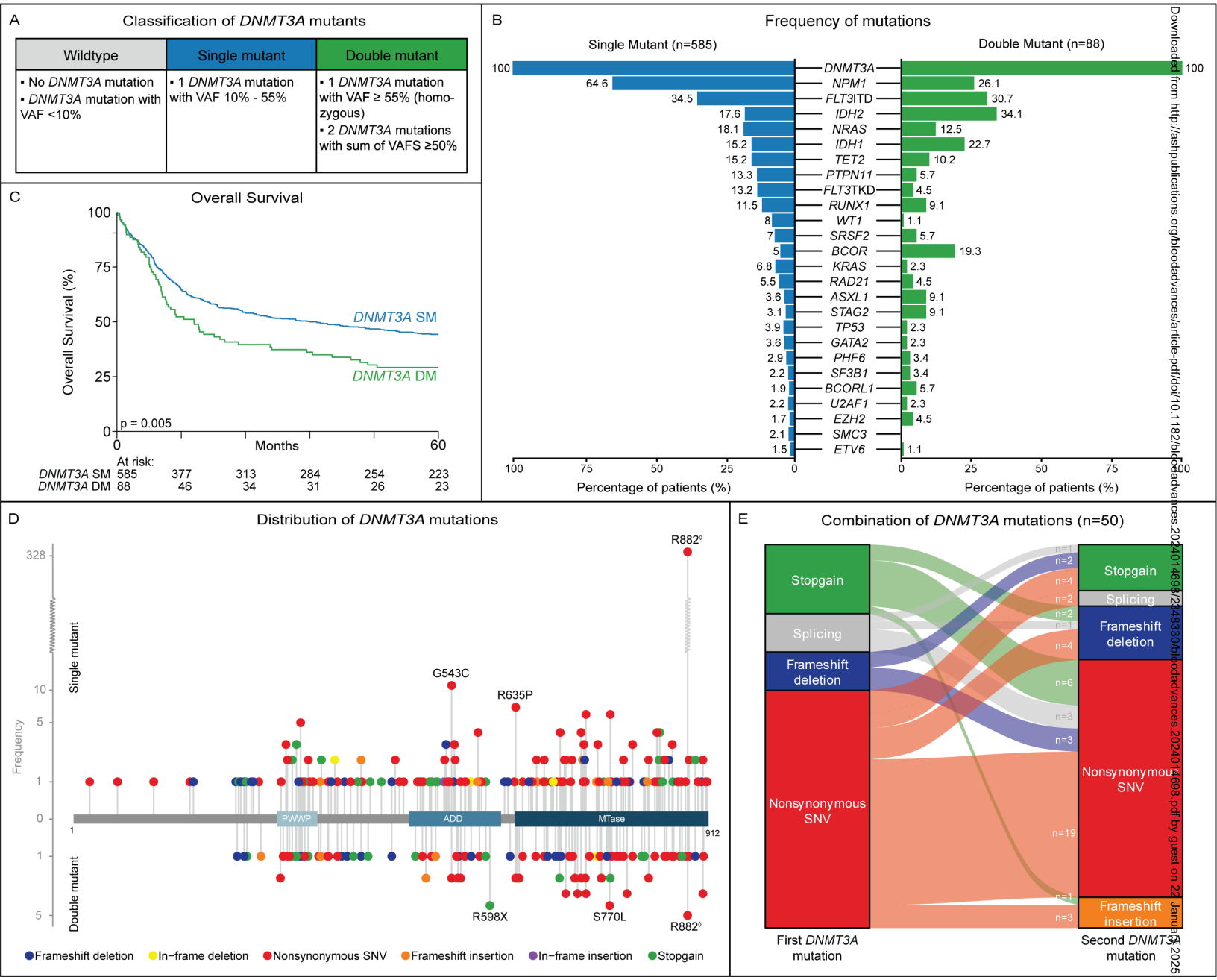
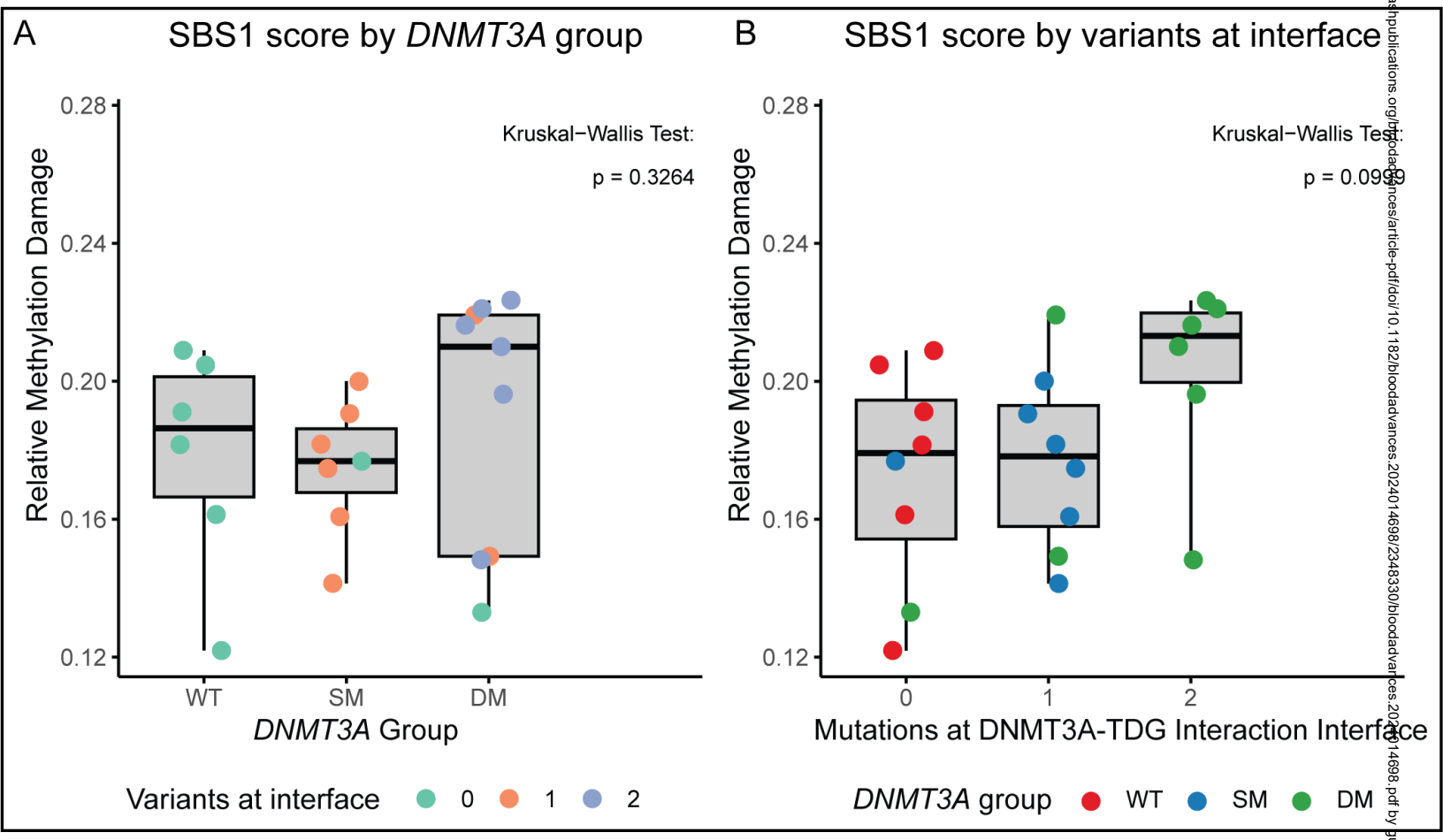


Figure 5



Double mutant *DNMT3A* AML: a unique subtype experiencing increased DNA damage and poor prognosis

Emma L. Boertjes[1*], Sanne Massaar[1*], Annelieke Zeilemaker[1], Jolinda Konijnenburg[1], Melissa Rijken[1], François G. Kavelaars[1], Tim Grob[1], Jurjen Versluis[1], Bob Löwenberg[1], Peter J.M. Valk[1**] and Mathijs A. Sanders[1**]

* These authors contributed equally
** These authors contributed equally

Affiliations
[1] Department of Hematology, Erasmus MC Cancer Institute and Erasmus University Medical Center, Rotterdam, The Netherlands

Corresponding author
Dr. Mathijs Sanders
Erasmus University Medical Center Rotterdam
Department of Hematology, Nc806
Wytemaweg 80
3015 CN Rotterdam Z-H
The Netherlands
Phone: +31.10.7043769
Fax: +31.10.7044745
E-mail: m.sanders@erasmusmc.nl

Data sharing section
Whole genome sequencing data are available at EGA under project accession number EGAS00001007966

Short title: Double mutant *DNMT3A* AML: a unique subtype

Manuscript length
Text word count: 4006
Abstract word count: 214
Number of figures: 4
Number of tables: 1
Number of references: 59

Key points

1. Wildtype DNMT3A enhances TDG-mediated repair, whereas mutant DNMT3A impairs this process.
2. Double mutant *DNMT3A* AML exhibits increased DNA methylation damage and holds significant prognostic value.

Abstract

Mutation of *DNMT3A*, encoding a *de novo* methyltransferase essential for cytosine methylation, is a common early event in clonal hematopoiesis (CH) and adult acute myeloid leukemia (AML). Spontaneous deamination of methylated cytosines damages DNA, which is repaired by the base excision repair (BER) enzymes MBD4 and TDG. Congenital MBD4-deficiency has been linked to early-onset CH and AML, and is marked by exceedingly high levels of DNA damage and mutation of *DNMT3A.* Strikingly, wildtype (WT) DNMT3A binds TDG, thereby potentiating its repair activity. Since TDG is the only remaining BER enzyme in MBD4-deficient AML patients capable of repairing methylation damage, we investigated whether mutant DNMT3A negatively affects the repair function of TDG. We found that, whereas WT DNMT3A stimulates TDG function, mutant DNMT3A impairs TDG-mediated repair of DNA damage *in vitro*. In light of this finding and to extrapolate our observations to the broader AML patient population, we investigate here the genetic profiles and survival outcomes of AML patients with single (SM) versus double mutant (DM) *DNMT3A*. DM *DNMT3A* AML patients show a characteristic driver mutation landscape and reduced overall survival when compared to SM *DNMT3A* AML patients. Importantly, whole-genome sequencing showed a trend for increased DNA damage in primary DM *DNMT3A* AML samples, especially when *DNMT3A* mutations are located at the DNMT3A-TDG interaction interface.

Introduction

At least one mutation in a known driver gene can be identified by targeted next generation sequencing (NGS) in >95% of acute myeloid leukemia (AML) patients, with *DNMT3A* being one of the most frequently mutated genes.[1] Large-scale genomic studies have identified mutations in *DNMT3A* as the most prevalent early event (~20%) in adult AML. The high prevalence of *DNMT3A* mutations in myeloid neoplasms, indicates that these mutations are important in driving leukemogenesis. *DNMT3A* mutations also represent the most frequent genetic aberration driving clonal hematopoiesis (CH) in healthy individuals.[2-5] In CH, *DNMT3A*-mutant hematopoietic stem and progenitor cells (HSPCs) gain a competitive advantage over their normal counterpart. Their expansion is assumed to predispose them to acquiring additional cooperating leukemogenic mutations. The precise effect on disease pathogenesis of the different types of *DNMT3A* mutations remains to be uncovered.[6]

DNMT3A encodes an enzyme crucial for *de novo* DNA methylation. The enzyme contains three functional domains: a Pro-Trp-Trp-Pro (PWWP), an Atrx-DNMT3-DNMT3L (ADD) and a catalytic C-terminal (MTase) domain.[7-10] The PWWP domain targets heterochromatin by recognizing H3K36me2/3, while the ADD domain is required for the recognition and binding to histone tails lacking K4 methylation, thereby facilitating enzymatic interactions with various proteins.[2,11] The highly conserved MTase domain catalyzes 5-cytosine methylation, and is thereby essential for establishing DNA methylation and the control of gene expression.[12,13] During human *de novo* DNA methylation, DNMT3A complexes with DNMT3L to form a tetramer composed of two DNMT3A proteins flanked by DNMT3L.[14,15] The tetramer binds DNA and the catalytic interface catalyzes 5-cytosine methylation to establish methylation patterns.

While different types of mutations, such as missense, frameshift and splice-site, occur within *DNMT3A*, the missense mutation at amino acid R882 is most prevalent.[2,6,16] The effect on methylation activity varies among the different *DNMT3A* mutations and is therefore associated with epigenetic heterogeneity.[17-19] Several studies suggest that missense mutations in *DNMT3A* have a dominant-negative effect, especially the R882 hotspot mutation. However, no gross methylation differences are observed in *DNMT3A* mutant AML, with only focal hypomethylation detected.[19-24] Furthermore, in AML, *DNMT3A* mutations often co-occur with mutations in *NPM1*, *FLT3* or *IDH1/2*, but little is known about the functional relationships between these co-mutations and different types of *DNMT3A* mutations.[1,25,26] In addition, the impact of mutant *DNMT3A* on treatment outcome in AML remains controversial.[16,27] For instance, the effect of monoallelic versus biallelic mutated *DNMT3A* on outcome has been sparsely investigated.[28,29] Although, limited by sample size, patients with biallelic mutated *DNMT3A* AML may have an inferior prognosis.[16,28,30]

Spontaneous deamination of 5-methylcytosine (5mC) is a natural threat to our genomic stability,[31] thereby replacing 5mC with thymine resulting in a C-to-T-transition mutation.[32] The specific base substitution and sequence context of this mutational process is represented by a mutational signature, termed SBS1.[33] Increased DNA methylation damage is signified by an enrichment of SBS1 in the genomic mutation imprint.[33] TET enzymes catalyze a stepwise hydroxylation process in which 5mC is converted to 5-hydroxymethylcytosin (5hmC). 5hmC is subsequently hydroxylated to form 5-formylcytosine (5fC) followed by 5-carboxylcytosine (5caC).[34,35] During spontaneous deamination the DNA glycosylases Methyl-CpG-binding domain 4 (MBD4) and Thymine DNA glycosylase (TDG) replace via BER the 5mC derivatives with cytosine. In the hydroxylation pathway, only TDG is involved in this process (**Figure 1A**).[36] Of note, mutations in *TET2,* another frequent somatic gene mutation in AML, are often loss-of-function and associated with DNA hypermethylation.[17,37-39] Furthermore, mutant isocitrate dehydrogenase 1 or 2 (IDH1 or IDH2) form the oncometabolite D-2-hydroxyglutarate (D-2-HG)

which inhibits TET2 activity, also leading to increased DNA methylation.[40] A rare germline deficiency of MBD4, previously identified by our lab, predisposes to accelerated CH and early-onset AML.[41] These patients share a common path to AML, marked by biallelic *DNMT3A* mutation followed by the acquisition of a mutation in either *IDH1* or *IDH2*. This conserved path to AML is likely caused by impaired DNA methylation pathways, resulting in selective pressure on methylated CpG-loci and, subsequent mutations within cooperating cancer genes.

The excessive amount of DNA methylation damage in MBD4-deficient AML patients, indicates that TDG does not compensate for the complete loss of MBD4.[41] DNMT3A and TDG are binding partners, affecting the enzymatic activity of both proteins, indicating that DNA methylation may be partially linked to BER (**Figure 1A**).[42] Since driver mutations in leukemia are often caused by C-to-T mutations, we set out to examine whether mutant DNMT3A impairs TDG repair activity, thereby increasing the risk of C-to-T mutations and the acquisition of additional driver mutations.

Methods

Patients and cell samples

Bone marrow or peripheral blood samples of 2,545 patients with a confirmed diagnosis of high risk MDS or AML were included. All patients were enrolled in the consecutive Dutch-Belgian Hemato-Oncology Cooperative group (HOVON) and Swiss Group for Clinical Cancer Research (SAKK) trials HO04, HO29, HO42, HO42A, HO43, HO92, HO102, HO103 and HO132.[43-50] In these studies patients with newly diagnosed AML were treated with intensive therapy according to trial protocol. Patients had given written informed consent in accordance with the Declaration of Helsinki.

Targeted Next Generation Sequencing (NGS)

Mutations at diagnosis were determined by NGS with the Illumina Trusight Myeloid Sequencing panel according to manufacturer instructions (Illumina, San Diego, CA). This sequencing panel covers 54 frequently mutated genes in myeloid malignancies.[51] *CEBPA* mutation detection was carried out using a custom NGS amplicon panel previously designed by our lab.[52] Detection of *FLT3* internal tandem duplication (ITD) was done as described previously.[53,54] Definitions of DNMT3A DM: either a single *DNMT3A* mutation with a variant allele frequency (VAF) exceeding 55% or the presence of two *DNMT3A* mutations with a combined VAF above 50%, DNMT3A SM: *DNMT3A* mutation with a VAF ranging from 10 to 55%, DNMT3A WT: absence of a *DNMT3A* mutation or a *DNMT3A* mutation with a VAF below 10% (considered as CH).

Cytogenetics

Cytogenetic testing was carried out at the local reference centers using standard protocols. All data was centrally peer-reviewed by clinical genetics laboratory specialists. All clonal and numerical chromosomal abnormalities were reported in accordance with the International System for Human Cytogenetic Nomenclature and the ELN 2022 recommendations.

Expression and purification of recombinant DNMT3A, MBD4 and TDG proteins

The pET28-MHL hexahistidine-tagged DNMT3A wildtype and mutants, MBD4 residues 430-580 (pET28, Addgene) and TDG residues 111-348 (pET28, kindly provided by Hashimoto et al.)[55] were expressed in Escherichia coli BL21(DE3) Gold and BL21(DE3) pLysS cells (Life Technologies). Single colonies of BL21(DE3) containing pET28 expression vector with the desired insert were cultured for 16 hours at 37°C in Luria Broth medium, supplemented with 50 µg/ml kanamycin. The cultures were subsequently diluted 25 times in Luria Broth medium without selection marker and grown at 37°C until the OD600 reached 0.5-0.8. The cells were induced by adding 0.5 mM isopropyl b-D-1-thiogalactopyranoside and incubated for another 4h at 37°C. Subsequently, cell cultures were centrifuged at 38.000 x g for 30 minutes at 4°C and

pellets were suspended in lysis buffer (pH 7.4, containing 20 mM sodium phosphate, 500 mM NaCl and 10 mM imidazole, supplemented with 1mg/ml lysozyme, 200 µg/ml DNAse, 1x Sigma fast protease inhibitor EDTA-free, Sigma-Aldrich) to minimize protein degradation and incubated on ice for 30 minutes. The cells were sonicated on ice for 6 minutes: 10 sec on, 20 sec off, amplitude 60%, (Branson digital sonifier). The resulting lysate was clarified by centrifugation at 38.000xg for 30 min at 4°C. The hexahistidine-tagged proteins were isolated from the lysate using Ni-NTA Superflow (Qiagen) and an Econo-chromatography column (2.5 x 10 cm, Biorad). To get rid of non-specific interacting proteins, the Ni-NTA resin was washed twice with cold washing buffer (pH 7.4, 20 mM sodium phosphate, 500 mM NaCl and 10mM imidazole). The hexahistidine-tagged proteins were eluted from the Ni-NTA group on the matrix with cold elution buffer (pH 7.4, containing 20 mM sodium phosphate, 500 mM NaCl and 500 mM imidazole). Directly after the proteins were eluted in 500 mM imidazole containing buffer, the suspensions were dialyzed in a slide-a-lyzer G2 dialysis cassette, gamma irradiated with pore size 10K MWCO (Thermo Scientific), against 300x volume dialysis buffer (50 mM Tris-HCl, pH 7.6 and 150 mM NaCl), for 2 hours at 4°C. Dialysis buffer was exchanged for two additional times and dialyzed for 2 hours and 16 hours respectively, at 4°C, prior to use or storage in small aliquots. The protein concentrations were quantified using a Qubit 3.0 fluorometer (Life Technologies) and Qubit protein assay kit, according to the manufacturer's protocol. Subsequently, proteins were verified by SDS-PAGE, using a Bis-Tris XCell SureLock™ Mini-Cell system (Thermo Fisher). The samples were loaded onto a precasted NuPage Novex 4-12% bis-tris gel (thickness 1,5mm, Thermo Fisher) and were run in 1x MOPS (Thermo Fisher) at 200V for 90 minutes. The blots were incubated with appropriate antibodies: α-His H-15 (Santa Cruz), α-DNMT3A (Abcam), α-TDG (Abcam), α-MBD4 (Abcam) in blocking buffer (5% BSA, 0,1% Tween and 1x PBS). Finally, the proteins were visualized with the use of a Li-Cor Odyssey 3.0 (Westburg) and corresponding Li-Cor Odyssey software.

MBD4 and TDG glycosylase activity and bandshift assays

MBD4 and TDG glycosylase activity and bandshift assays were performed as described by Hashimoto et al.[55] See **Figure 1B** for a schematic visualization of the glycosylase activity assays and see supplementary methods for detailed description.

Whole-Genome Sequencing

Whole-genome sequencing (WGS) was performed on DNA of 22 bone marrow aspirates of *de novo* AML patients (6 *DNMT3A* WT, 7 *DNMT3A* SM and 9 *DNMT3A* DM). Libraries were generated using the Illumina DNA PCR-Free library Prep, tagmentation kit following manufacturer's protocol (Illumina, San Diego, CA). Quality control and quantification of generated libraries were done using Qubit Fluorometric Quantitation 3.0 (Thermofisher, Waltham, MA), taqman qPCR analysis and Agilent 2100 Bioanalyzer platforms (Agilent, Santa Clara, CA). The final WGS libraries were analyzed by 2x151 cycles paired-end sequencing on a NovaSeq 6000 system (Illumina, San Diego, CA) with an average coverage of 35x for AML samples and 20x for matched controls. Genomic sequences were aligned against the hg19 reference genome using bwa-mem2 (https://github.com/bwa-mem2/bwa-mem2). The calling of variants, structural variations and copy number changes was performed using CGPWGS (https://github.com/cancerit/dockstore-cgpwgs), using the hg19 reference database, settings and filtering strategies previously described in detail.[56]

AlphaFold

AlphaFold v2.2.0 was used to predict the interaction of wildtype and mutant forms of DNMT3A with TDG.[57] The top ranked relaxed multimer model of each predicted structure was visualized in ChimeraX and interacting residues were quantified using "interfaces areaCutOff0."

Statistical analysis

For patient characteristics, statistical associations between categorical variables were assessed using the Fisher's exact test and by the Mann-Whitney U test for continuous variables. Mutations counts derived from WGS were tallied with MutationalPatterns.[58] Enrichment of mutational signatures SBS1, SBS5 and SBS18 within the tallied mutation counts were determined with sigfit (https://github.com/kgori/sigfit) using default settings. The Kruskal-Wallis test was used to statistically compare SBS1-enrichment scores representing the amount of methylation-linked DNA damage that occurs in each AML. Overall survival analyses were performed using the Kaplan-Meier method and survival differences assessed using the log-rank test. A multivariable Cox proportional-hazards model was used to perform a univariate survival analysis of *DNMT3A* SM versus *DNMT3A* DM patients. The model was adjusted for the covariates: age, WBC at diagnosis, cycles of chemotherapy required to achieve CR, mutational status of *NPM1, FLT3*-ITD, myelodysplasia-related gene mutations (*ASXL1, BCOR, EZH2, RUNX1, SF3B1, SRSF2, STAG2, U2AF1* and *ZRSR2*) and *TP53* at diagnosis, and HOVON trial. Allogeneic transplantation was included as time-dependent variable. Values of $p<0.05$ were considered as statistically significant. Statistical analyses were executed with STATA Statistical Software, release 18.0 (College Station, TX) and with R 4.3.1 (Boston, MA).

Results

Wildtype DNMT3A potentiates TDG repair activity, but not MBD4 repair activity

Wildtype DNMT3A and TDG are known to interact, thereby increasing the repair activity of the latter.[42] We therefore hypothesized that mutant DNMT3A negatively impacts the function of the TDG-DNMT3A complex. To assess the impact of mutant DNMT3A on TDG function, we first performed glycosylase activity assays with wildtype DNMT3A and TDG protein. In these assays, a TDG construct including both the catalytic domain and a N-terminal regulatory region ($TDG^{AA111-348}$), was incubated with a double-stranded FAM-labeled 32-bp oligonucleotide harboring a G-T mismatch. Functional $TDG^{AA111-348}$ identifies and excises the nucleotide mismatch in this assay. The subsequent hot alkaline treatment facilitates a double-stranded DNA break at the position of the excised base (product, **Figure 1B**). To verify that recombinant full-length wildtype DNMT3A enhanced TDG repair activity, 40nM $TDG^{AA111-348}$ was incubated with increasing concentrations of wildtype DNMT3A. This led to elevated product formation in a dose-dependent matter, thereby confirming the enhanced effect of recombinant wildtype DNMT3A on TDG repair activity (**Figure 1C-D**, **S1A,C**). Increasing the concentration of wildtype DNMT3A did not impact MBD4 glycosylase activity (**Figure 1E-F**, **S1B-C**).

Mutant DNMT3A impairs TDG repair activity

Next, we hypothesized that mutant DNMT3A could negatively impact the capacity of TDG to repair DNA damage. Four DNMT3A mutant forms commonly found in CH in healthy as well as in MBD4-deficient individuals (DNMT3A R635W, R688C, R882C and A884W)[41] were added to glycosylase assays in increasing concentrations to determine the effect on DNA damage repair. All full-length mutant DNMT3A proteins showed no TDG stimulation at low concentrations, and even inhibition of TDG function at higher concentrations compared to wildtype DNMT3A (**Figure 2**, **S2**). However, differences in rates of inhibition were apparent between the different DNMT3A mutants. DNMT3A R635W closely mirrored the pattern observed with wildtype DNMT3A, but in contrast to wildtype DNMT3A, showed decreased TDG activity at high concentrations (**Figure 2A-B**, **S2**). In approximately 80-85% of CH and AML cases *DNMT3A* mutations are heterozygous. To mimic this setting, mutant forms of recombinant DNMT3A were co-titrated with wildtype DNMT3A in the glycosylase assays. All DNMT3A mutants clearly suppressed the capacity of TDG to repair DNA damage compared to wildtype DNMT3A alone. Furthermore, at high concentrations, DNMT3A mutants imposed complete inhibition of TDG, irrespective of the concentration of wildtype DNMT3A present (**Figure 3**). Band shift assays confirmed decreased

binding affinity of the TDG to DNA in the presence of mutant DNMT3A when compared to the wildtype counterpart (**Figure S3**). We were interested in whether these DNMT3A mutations are positioned at the interaction interfaces of DNMT3A and TDG. AlphaFold 2.0 was used to predict the protein interactions between TDG and both wildtype and mutant forms of DNMT3A. The predictions indicated that all mutations, except for DNMT3A R635W, were likely to disturb the interaction surface between TDG and DNMT3A (**Figure S4**). Of note, the R635W mutation was the only mutation leading to less inhibition of TDG repair capacity in our glycosylase assays compared to the other assessed mutations (**Figure 2A-B**)**.**

Correlation of *DNMT3A* mutational status and patient outcomes

We investigated how DNMT3A SM versus DM status impacts the mutational AML landscape and whether it correlates with survival. Within our cohort, 1,872 individuals were classified as *DNMT3A* wildtype (*DNMT3A* WT), 585 patients were classified as *DNMT3A* single mutant (*DNMT3A* SM) and 88 patients were identified as *DNMT3A* double mutant (*DNMT3A* DM) (system of classification, **Figure 4A**). Among the 88 *DNMT3A* DM patients, 38 patients harbored a homozygous *DNMT3A* mutation (**Table S1**). One homozygous *DNMT3A* DM patient had a chromosome 2p deletion that encompasses the *DNMT3A* gene. All other DM patients had a combination of 2 different mutations with relatively high VAFS (**Table S2**). *DNMT3A* DM patients were significantly older compared to *DNMT3A* SM patients (median age 59 and 55 years, respectively, p=0.006), achieved less frequently complete remission (CR) after the first cycle of chemotherapy (69.6% vs. 48.9%; p= <0.001) and were less likely to receive consolidation therapy or allogeneic transplantation in CR (**Table S3**). Compared to *DNMT3A* SM, *DNMT3A* DM patients acquired significantly higher frequencies of mutations in *ASXL1* (3.6% vs. 9.1%; p=0.042), *BCOR* (5.0% vs. 19.3% p<0.001), *BCORL1* (1.9% vs. 5.7%; p=0.046), *IDH2* (17.6% vs. 34.1%; p=0.001), *IKZF1* (0.5% vs. 4.5%; p=0.007) and *STAG2* (3.1% vs. 9.1%; p=0.013), whereas the frequencies of *FLT3*-TKD (13.2% vs. 4.5%; p=0.021), and mutations in *NPM1* (64.6% vs. 26.1%; p<0.001) and *WT1* (8.0% vs. 1.1%; p=0.014) were more prevalent in *DNMT3A* SM (**Figure 4B**, **Table S4**). Of note, a 10-fold enrichment of the *IDH2* R172 hotspot mutation was detected in *DNMT3A* DM compared to *DNMT3A* SM patients (**Figure S5**). In both *DNMT3A* DM and SM mutants, *DNMT3A* mutations were distributed throughout the entire gene, with R882 mutations being the most prevalent mutation in *DNMT3A* SM patients, accounting for 328 out of 585 mutations (**Figure 4D**, **Table S1**). Surprisingly, the DNMT3A R882 hotspot mutation was only detected in 5 out of 142 mutations in the *DNMT3A* DM cohort (**Figure 4D**, **Table S1**). Within the *DNMT3A* DM subset with compound heterozygous *DNMT3A* mutations (n=50), the presence of two nonsynonymous single nucleotide variants (SNV) occurred most frequent. Notably, within the subset of *DNMT3A* DM patients with two loss-of-function (LOF) mutations, we did not observe the combination of two nonsense, two splicing or two frameshift mutations. Yet, combinations of two different LOF mutations were commonly observed (**Figure 4E**). In univariate survival analysis, *DNMT3A* DM patients had significantly poorer overall survival compared to *DNMT3A* SM patients (**Figure 4C**, independently validated using BEAT-AML data**, Figure S6**). After adjusting for covariates known to have a prognostic impact, *DNMT3A* DM status was considered an independent prognostic marker and associated with poor overall survival compared to the *DNMT3A* SM status (HR: 1.40 95% CI [1.05–1.86]; p=0.021) (**Table S5A-B**).

DNA methylation damage increases in *DNMT3A* DM AML, especially when *DNMT3A* mutations are positioned at the DNMT3A-TDG interaction interface

Based on the *in vitro* glycosylase assays, we hypothesized that patients with a complete loss of wildtype DNMT3A might exhibit greater accumulation of DNA damage. To investigate this

hypothesis, we selected 16 patients from our biobank who had either SM (n=7) or DM (n=9) *DNMT3A* mutations that were most likely to result in the complete loss of DNMT3A, along with 6 corresponding controls and performed WGS on their diagnostic bone marrow aspirates. An overview of the selected AML patients is listed in **Table 1**. All selected *DNMT3A* SM and DM cases carried *IDH1/2* mutation. This selection was made to rule out possible biases introduced by *IDH* mutations, which are known to be associated with DNA hypermethylation. The *DNMT3A* WT group contained both *IDH* mutant and *IDH* WT patients. For an overview of all co-mutations, see Supplementary **Table S6A-C**. Alphafold 2.0 was used to model the interaction between DNMT3A and TDG to determine the possible impact of *DNMT3A* mutations positioned at the interaction interface. In general, *DNMT3A* DM patients had a relative higher SBS1 enrichment compared to *DNMT3A* SM and WT patients (**Figure 5A**). In line with our hypothesis, AMLs with two *DNMT3A* mutations located at the TDG-DNMT3A interaction interface tended to have a SBS1 enrichment compared to patients with one or no *DNMT3A* mutation at this interface (**Figure 5B**). One *DNMT3A* DM patient, with a low SBS1 enrichment (DM-2), had a homozygous L637P mutation, which is predicted not to be involved in DNMT3A-TDG interaction, thereby potentially explaining the lower SBS1 enrichment observed. We observed no differences in gene expression between *DNMT3A* SM and DM patients (data not shown) and we therefore believe that *DNMT3A* DM patients forms a unique subtype based on their level of methylation damage and subsequent genetic landscape, rather than a distinct gene expression profile.

Discussion

Spontaneous deamination of 5mC is a naturally occurring mutational process. If this DNA damage is not corrected on time, it will lead to C-to-T transition mutations. To prevent the acquisition of these mutations, life has evolved a DNA repair system involving the two DNA glycosylase enzymes MBD4 and TDG, which safeguard against DNA damage through BER. DNMT3A, following repair, methylates the 5-cytosine position, restoring the original DNA methylation pattern. Li et al, showed that in the process of DNA (de)methylation, DNMT3A and TDG interact, thereby regulating the activity of both proteins and linking DNA repair to methylation.[42]

DNMT3A mutations predicted to impair the function of the TDG-DNMT3A complex, i.e., DNMT3A R688C, R882C and A884W, confer a remarkable decrease in TDG repair activity in our glycosylase assays. This inhibitory effect persists when mutant DNMT3A was co-titrated with wildtype DNMT3A, mimicking a heterozygous situation (the most common situation in CH and AML). In contrast, a single DNMT3A mutant (DNMT3A R635W), predicted to not be present at the interaction interface, only exhibits decreased TDG activity at high concentrations. Interestingly, we observed greater DNA damage in primary AML samples carrying two TDG-DNMT3A interaction altering variants compared to those with one or no interaction altering mutation. Notably, the *DNMT3A* DM patient homozygous for the DNMT3A L637P mutation showed relatively low DNA damage. This missense mutation is proximal to the DNMT3A R635W mutation, which shows only moderate inhibition of TDG in our glycosylase assay. Collectively, these results indicate that mutated DNMT3A retains its ability to bind and enhance TDG function provided that the mutations are positioned outside the interacting interface. It is postulated that wildtype DNMT3A facilitates TDG binding to G-T mismatches and subsequent processing, however, mutant DNMT3A may induce complex conformational changes interfering with the ability of TDG to bind DNA, thereby diminishing its repair function.[42] As support for this notion, our band shift assays confirm a decrease in the TDG binding affinity to DNA in the presence of mutant DNMT3A.

Additionally, Alphafold 2.0 predicted changes in residues of both mutant DNMT3A and TDG resulting in novel interactions. Most forms of mutant DNMT3A are predicted to interact with residues located within the uracil-DNA glycosylase domain of TDG, whereas wildtype DNMT3A is predicted not to interact with this domain (**Figure S7**)**.** Given the significance of this domain for TDG repair capacity, such changes in interaction may further impact TDG repair activity. However, to validate these interactions and potential structural changes, protein crystallography experiments would be necessary to visualize the TDG-DNMT3A protein complex upon mutation of DNMT3A. In addition, the exact mechanisms through which mutant DNMT3A impairs TDG function remain unknown and should be investigated in further detail.

The results of the glycosylase assays representing a heterozygous mutant context, indicate that titration of increasing amounts of recombinant mutant DNMT3A at steady levels of wildtype DNMT3A, inhibit TDG repair capacity. In fact, as the concentration of mutant DNMT3A equals or surpasses that of wildtype DNMT3A, TDG function appears to decline. In this situation, the activity level of TDG becomes even less compared to baseline without addition of any wildtype DNMT3A (**Figure 3**). These results suggest that both wildtype and mutant DNMT3A compete to interact with TDG. When mutant DNMT3A concentration exceeds that of wildtype DNMT3A, a likely greater proportion of TDG interaction sites will be occupied by mutant DNMT3A, leading to a reduction in glycosylase activity.

DNMT3A DM AML with biallelic inactivating mutations, likely lacking functional DNMT3A, were subjected to WGS. The trend of increased DNA damage in these AML patients, although cohort size is limited by biobank availability, suggests that the acquired DNMT3A mutations impair TDG function. Moreover, these patients still had functional MBD4 which remains capable of recognizing and repairing G-T mismatches. As both our initial experiments and available literature did not suggest interactions between MBD4 and DNMT3A, we did not include the results on the impact of mutant DNMT3A on MBD4 repair function.

To assess the clinical impact of *DNMT3A* mutational status, we compared patients carrying single or double *DNMT3A* mutations. We found that both *DNMT3A* SM and *DNMT3A* DM mutations were distributed across the entire gene. However, mutations occurring at R882 (hotspot mutation), which accounts for almost 60% of all *DNMT3A* mutations in AML, were almost exclusively seen in *DNMT3A* SM patients.[59] Previous studies have shown that R882 mutations inhibit wildtype DNMT3A to form active tetramers, thereby altering the process of CpG-methylation which might be another contributing factor of *DNMT3A* mutation in oncogenesis.[20,23,60] Also, *DNMT3A* DM patients were of significantly older age, and in the context of an intensive chemotherapy setting they were less likely to attain an early CR (i.e. after the first cycle of remission induction chemotherapy) and their OS was reduced compared to that of *DNMT3A* SM patients. These findings extend the limited research that has been conducted on this AML subgroup previously and establish DNMT3A DM as a distinct subset of AML.[29] In a multivariable analysis investigating OS, after adjusting for several clinically relevant covariates, *DNMT3A* DM status remained a significant independent risk factor for poorer OS. Taking into account the older age and mutational landscape of *DNMT3A* DM AML patients, it would be of great interest to explore the responses to HMAs in combination with venetoclax or ivosidenib (in *IDH1* mutant AML), considering that these treatments show higher response rates and longer overall survival in unfit AML patients. [61,62] Altogether, our study indicates that *DNMT3A* DM AML, a unique subgroup comprising approximately 4% of the complete AML population, experiences increased DNA damage and that *DNMT3A* mutational status carries significant prognostic value.

Acknowledgements

We thank, all the patients, clinicians, research nurses, data managers and laboratory scientists who provided samples, conducted cytogenetic analyses and performed analyses at the participating centers of the Dutch–Belgian Cooperative Trial Group for Hematology–Oncology (HOVON) and Swiss Group for Clinical Cancer Research (SAKK), including centers in the Netherlands, Belgium, Switzerland, Norway, Sweden, Lithuania, Germany and Denmark. In addition, we thank Eric Bindels for performing next generation sequencing and Remco Hoogenboezem and Dorien Pastoors for their help with data analysis.

Contributions

P.J.M.V. and M.A.S. designed the research; P.J.M.V., B.L. provided patient information and materials; E.L.B., A.Z., M.R. and F.G.K. performed experiments; E.L.B, S.M. and M.A.S and J.K. analyzed data; T.G. and J.V. provided statistical advice; E.L.B and S.M. prepared the figures. E.L.B., S.M., P.J.M.V. and M.A.S. wrote the manuscript. All authors approved the final version of the manuscript.

Funding information

M.A.S and S.M. are supported by a KWF Kankerbestrijding Young Investigator Grant (12797, Bas Mulder Award; Dutch Cancer Foundation).

Disclosure

All authors declare no conflicts of interest.

References

1. 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.
2. Khrabrova DA, Yakubovskaya MG, Gromova ES. AML-Associated Mutations in DNA Methyltransferase DNMT3A. *Biochemistry (Mosc)*. 2021;86(3):307-318.
3. Challen GA. Dominating the Negative: How DNMT3A Mutations Contribute to AML Pathogenesis. *Cell Stem Cell*. 2017;20(1):7-8.
4. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.
5. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487.
6. Venugopal K, Feng Y, Shabashvili D, Guryanova OA. Alterations to DNMT3A in Hematologic Malignancies. *Cancer Res*. 2021;81(2):254-263.
7. Brunetti L, Gundry MC, Goodell MA. DNMT3A in Leukemia. *Cold Spring Harb Perspect Med*. 2017;7(2).
8. Qin S, Min J. Structure and function of the nucleosome-binding PWWP domain. *Trends Biochem Sci*. 2014;39(11):536-547.
9. Dhayalan A, Rajavelu A, Rathert P, et al. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem*. 2010;285(34):26114-26120.
10. Ren W, Gao L, Song J. Structural Basis of DNMT1 and DNMT3A-Mediated DNA Methylation. *Genes (Basel)*. 2018;9(12).
11. Zhang Y, Jurkowska R, Soeroes S, et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res*. 2010;38(13):4246-4253.

12. Gowher H, Jeltsch A. Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J Biol Chem*. 2002;277(23):20409-20414.
13. Challen GA, Sun D, Jeong M, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet*. 2011;44(1):23-31.
14. Holz-Schietinger C, Matje DM, Harrison MF, Reich NO. Oligomerization of DNMT3A controls the mechanism of de novo DNA methylation. *J Biol Chem*. 2011;286(48):41479-41488.
15. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*. 2007;449(7159):248-251.
16. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med*. 2010;363(25):2424-2433.
17. Hajkova H, Markova J, Haskovec C, et al. Decreased DNA methylation in acute myeloid leukemia patients with DNMT3A mutations and prognostic implications of DNA methylation. *Leuk Res*. 2012;36(9):1128-1133.
18. Qu Y, Lennartsson A, Gaidzik VI, et al. Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. *Epigenetics*. 2014;9(8):1108-1119.
19. Chaudry SF, Chevassut TJ. Epigenetic Guardian: A Review of the DNA Methyltransferase DNMT3A in Acute Myeloid Leukaemia and Clonal Haematopoiesis. *Biomed Res Int*. 2017;2017:5473197.
20. Kim SJ, Zhao H, Hardikar S, Singh AK, Goodell MA, Chen T. A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood*. 2013;122(25):4086-4089.
21. Yang X, Wang X, Yang Y, et al. DNMT3A mutation promotes leukemia development through NAM-NAD metabolic reprogramming. *J Transl Med*. 2023;21(1):481.
22. Emperle M, Dukatz M, Kunert S, et al. The DNMT3A R882H mutation does not cause dominant negative effects in purified mixed DNMT3A/R882H complexes. *Sci Rep*. 2018;8(1):13242.
23. Russler-Germain DA, Spencer DH, Young MA, et al. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell*. 2014;25(4):442-454.
24. Ribeiro AF, Pratcorona M, Erpelinck-Verschueren C, et al. Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. *Blood*. 2012;119(24):5824-5831.
25. Kishtagari A, Levine RL, Viny AD. Driver mutations in acute myeloid leukemia. *Curr Opin Hematol*. 2020;27(2):49-57.
26. Zhang X, Wang X, Wang XQD, et al. Dnmt3a loss and Idh2 neomorphic mutations mutually potentiate malignant hematopoiesis. *Blood*. 2020;135(11):845-856.
27. Torabi A, Alonzo TA, Othus M, et al. Characteristics and Prognostic Effects of DNMT3A Co-Mutations. *Blood*. 2022;140(Supplement 1):745-747.
28. Narayanan D, Pozdnyakova O, Hasserjian RP, Patel SS, Weinberg OK. Effect of DNMT3A variant allele frequency and double mutation on clinicopathologic features of patients with de novo AML. *Blood Advances*. 2021;5(11):2539-2549.
29. Chen X, Tian C, Hao Z, et al. The impact of DNMT3A variant allele frequency and two different comutations on patients with de novo cytogenetically normal acute myeloid leukemia. *Cancer Med*. 2023;12(9):10340-10350.
30. Kawashima N, Kubota Y, Bravo-Perez C, et al. Landscape of biallelic DNMT3A mutant myeloid neoplasms. *J Hematol Oncol*. 2024;17(1):87.
31. Ehrlich M, Zhang XY, Inamdar NM. Spontaneous deamination of cytosine and 5-methylcytosine residues in DNA and replacement of 5-methylcytosine residues with cytosine residues. *Mutat Res*. 1990;238(3):277-286.
32. Shen JC, Rideout WM, 3rd, Jones PA. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res*. 1994;22(6):972-976.

33. Koh G, Degasperi A, Zou X, Momen S, Nik-Zainal S. Mutational signatures: emerging concepts, caveats and clinical applications. *Nat Rev Cancer*. 2021;21(10):619-637.
34. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology*. 2013;38(1):23-38.
35. Crawford DJ, Liu MY, Nabel CS, Cao XJ, Garcia BA, Kohli RM. Tet2 Catalyzes Stepwise 5-Methylcytosine Oxidation by an Iterative and de novo Mechanism. *J Am Chem Soc*. 2016;138(3):730-733.
36. Sjolund AB, Senejani AG, Sweasy JB. MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles. *Mutat Res*. 2013;743-744:12-25.
37. Feng Y, Li X, Cassady K, Zou Z, Zhang X. TET2 Function in Hematopoietic Malignancies, Immune Regulation, and DNA Repair. *Front Oncol*. 2019;9:210.
38. McClatchy J, Strogantsev R, Wolfe E, et al. Clonal hematopoiesis related TET2 loss-of-function impedes IL1β-mediated epigenetic reprogramming in hematopoietic stem and progenitor cells. *Nat Commun*. 2023;14(1):8102.
39. Tulstrup M, Soerensen M, Hansen JW, et al. TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. *Nat Commun*. 2021;12(1):6061.
40. Chan SM, Majeti R. Role of DNMT3A, TET2, and IDH1/2 mutations in pre-leukemic stem cells in acute myeloid leukemia. *Int J Hematol*. 2013;98(6):648-657.
41. Sanders MA, Chew E, Flensburg C, et al. MBD4 guards against methylation damage and germ line deficiency predisposes to clonal hematopoiesis and early-onset AML. *Blood*. 2018;132(14):1526-1534.
42. Li YQ, Zhou PZ, Zheng XD, Walsh CP, Xu GL. Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. *Nucleic Acids Res*. 2007;35(2):390-400.
43. Lowenberg B, Boogaerts MA, Daenen SM, et al. Value of different modalities of granulocyte-macrophage colony-stimulating factor applied during or after induction therapy of acute myeloid leukemia. *J Clin Oncol*. 1997;15(12):3496-3506.
44. Lowenberg B, Ossenkoppele GJ, van Putten W, et al. High-dose daunorubicin in older patients with acute myeloid leukemia. *N Engl J Med*. 2009;361(13):1235-1248.
45. Pabst T, Vellenga E, van Putten W, et al. Favorable effect of priming with granulocyte colony-stimulating factor in remission induction of acute myeloid leukemia restricted to dose escalation of cytarabine. *Blood*. 2012;119(23):5367-5373.
46. Lowenberg B, Pabst T, Maertens J, et al. Therapeutic value of clofarabine in younger and middle-aged (18-65 years) adults with newly diagnosed AML. *Blood*. 2017;129(12):1636-1645.
47. Brune MM, Stussi G, Lundberg P, et al. Effects of lenalidomide on the bone marrow microenvironment in acute myeloid leukemia: Translational analysis of the HOVON103 AML/SAKK30/10 Swiss trial cohort. *Ann Hematol*. 2021;100(5):1169-1179.
48. Lowenberg B, Pabst T, Maertens J, et al. Addition of lenalidomide to intensive treatment in younger and middle-aged adults with newly diagnosed AML: the HOVON-SAKK-132 trial. *Blood Adv*. 2021;5(4):1110-1121.
49. Randomized study to assess the added value of Laromustine in combination with standard remission-induction chemotherapy in patients aged 18-65 years with previously untreated acute myeloid leukemia (AML) or myelodysplasia (MDS) (RAEB with IPSS = 1.5) - HOVON 92 AML.
50. Dutch-Belgian Cooperative Trial Group for Hematology-Oncology.
51. Jongen-Lavrencic M, Grob T, Hanekamp D, et al. Molecular Minimal Residual Disease in Acute Myeloid Leukemia. *N Engl J Med*. 2018;378(13):1189-1199.
52. Al Hinai ASA, Grob T, Rijken M, et al. PPM1D mutations appear in complete remission after exposure to chemotherapy without predicting emerging AML relapse. *Leukemia*. 2021;35(9):2693-2697.

53. Engen C, Hellesøy M, Grob T, et al. FLT3-ITD mutations in acute myeloid leukaemia - molecular characteristics, distribution and numerical variation. *Mol Oncol*. 2021;15(9):2300-2317.
54. Grob T, Sanders MA, Vonk CM, et al. Prognostic Value of FLT3-Internal Tandem Duplication Residual Disease in Acute Myeloid Leukemia. *J Clin Oncol*. 2023;41(4):756-765.
55. Hashimoto H, Liu Y, Upadhyay AK, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res*. 2012;40(11):4841-4849.
56. Moore L, Cagan A, Coorens THH, et al. The mutational landscape of human somatic and germline cells. *Nature*. 2021;597(7876):381-386.
57. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589.
58. Manders F, Brandsma AM, de Kanter J, et al. MutationalPatterns: the one stop shop for the analysis of mutational processes. *BMC Genomics*. 2022;23(1):134.
59. Huang YH, Chen CW, Sundaramurthy V, et al. Systematic Profiling of DNMT3A Variants Reveals Protein Instability Mediated by the DCAF8 E3 Ubiquitin Ligase Adaptor. *Cancer Discov*. 2022;12(1):220-235.
60. Holz-Schietinger C, Matje DM, Reich NO. Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *J Biol Chem*. 2012;287(37):30941-30951.
61. DiNardo CD, Jonas BA, Pullarkat V, et al. Azacitidine and Venetoclax in Previously Untreated Acute Myeloid Leukemia. *N Engl J Med*. 2020;383(7):617-629.
62. Montesinos P, Recher C, Vives S, et al. Ivosidenib and Azacitidine in IDH1-Mutated Acute Myeloid Leukemia. *N Engl J Med*. 2022;386(16):1519-1531.

Figure legends

Figure 1. WT DNMT3A potentiates glycosylase activity of TDG but not MBD4. (*A*) Schematic overview of the DNA(de)methylation cascade in humans, including *de novo* methylation, spontaneous deamination of 5mC and base excision repair (BER) of G-T mismatches. (*B*) Experimental set-up of glycosylase assay. A FAM labeled double stranded 32bp oligo, containing a G-T mismatch, is incubated with TDG (AA 111-348). Functional TDG identifies and excises the mismatch. The subsequent hot alkaline treatment facilitates the oligo to break at the site of the excised base, resulting in a 18bp product. More functional TDG results in increased levels of product detected on gel. (*C-D*) Addition of wildtype DNMT3A to the glycosylase assay stimulates TDG activity in a dose-dependent matter. Represented on gel in *C* and quantified relative to non-stimulated TDG in *D*. (*E-F*) Glycosylase activity of MBD4 is not stimulated by addition of WT DNMT3A to the glycosylase assay. Represented on gel in E and quantified relative to non-stimulated MBD4 in *F*.

Figure 2. DNMT3A mutants commonly found in MBD4-deficient patients impair glycosylase activity of TDG. (*A*) Gel image of glycosylase assay with mutant DNMT3A. Addition of mutant DNMT3A does not potentiate TDG activity and even inhibits TDG activity at higher concentrations. (*B-E*) Quantification of each DNMT3A mutant tested in *A*, relative to unstimulated TDG.

Figure 3. Co-titration of increasing amount of recombinant wildtype DNMT3A relative to different mutant counterparts. * indicates samples ran on different gels in the same experiment and imaged at the same time due to logistic limitations.

Figure 4. Patient and mutation characteristics by DNMT3A allelic state. (A) Classification of DNMT3A mutants. (B) Frequency of driver mutations per DNMT3A mutant group (single mutant left, double mutant right). The percentage in which each individual gene occurs is shown behind each bar. (C) Overall Survival of DNMT3A single mutant (blue line) and DNMT3A double mutant (green line). (D) Lollipop plot illustrating the distribution of mutations in the DNMT3A single mutant cohort (upper part) and DNMT3A double mutant cohort (bottom part). The length of a lollipop represents the number of patients that carry a mutation at a specific amino acid. Each point is colored by the mutation type of the most frequent occurring mutation at that specific location. (E) Combination of type of mutations in DNMT3A double mutants with heterozygous mutations.

Figure 5. Methylation damage burden is increased in patients carrying two *DNMT3A* mutations, particularly when these mutations reside at the DNMT3A-TDG interaction surface. (*A*) Boxplot for SBS1 score grouped by *DNMT3A* mutant group. Each point depicts the SBS1 score of an individual patient. The color of each data point represents the quantity of *DNMT3A* mutations situated within the DNMT3A-TDG interface. (*B*) Boxplots for SBS1 score grouped by number of *DNMT3A* mutations located at the DNMT3A-TDG interaction surface. The color of each data point represents the *DNMT3A* mutant group.

Patient ID	Age	Sex	*DNMT3A* status	*DNMT3A* mutation (HGVSc, NM_022552.5)	*DNMT3A* mutation (HGVSp, NP_072046)	VAF	Effect of mutation	Karyotype
WT-1	63	F	WT	-	-		-	46,XX[22]
WT-2	49	F	WT	-	-		-	47,XX,+21[11]/46,XX[8]
WT-3	54	F	WT	-	-		-	46,XX [?]
WT-4	65	F	WT - *IDH* mutant	-	-		-	46,XX[20]
WT-5	65	F	WT - *IDH* mutant	-	-		-	46,XX[20]
WT-6	52	F	WT - *IDH* mutant	-	-		-	46,XX[25]
SM-1	56	F	SM	c.2006_2007insGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCA	p.His821fs	0.377	Frameshift	46,XX[31]
SM-2	49	M	SM	c.2644C>T	p.Arg882Cys	0.469	Missense	46,XY,t(2;17)(q?31;p11),+?add(11)(p10),-20[19]/46,XY,der(11;12)(q10;q10),+del(12)(p11p12)[6]
SM-3	59	F	SM	c.2644C>T	p.Arg882Cys	0.439	Missense	46,XX[21]
SM-4	44	F	SM	c.2711C>T	p.Pro904Leu	0.303	Missense	46,XX[21]
SM-5	69	F	SM	c.2264T>C	p.Phe755Ser	0.445	Missense	46,XX[20]
SM-6	69	M	SM	c.1166delA	p.Asp389fs	0.470	Frameshift	46,XY[20]
SM-7	49	F	SM	c.2576T>A	p.Leu859*	0.464	Stopgain	45,XX,-7[18]/46,XX[2]
DM-1	65	F	DM	1. c.994G>A	1. p.Gly332Arg	0.517	Missense	46,XX,?del(17)(q?)[4]/45,XX,-7[3]/46,XX[23]
				2. c.2578T>C	2. p.Trp860Arg	0.477	Missense	
DM-2	39	F	DM	c.1910T>C	p.Leu637Pro	0.930	Missense	46,XX,del(2)(p2?2p2?4)[7],add(15)(p10)[cp9]/46,XX[16]
DM-3	63	F	DM	c.972delC	p.Thr325fs	0.713	Frameshift	46,XX,t(7;10)(p13;q22),del(12)(q22q24)[12]/46,idem,del(18)(p11)[10]/46,XX[6]
DM-4	63	F	DM	1. c.1154delC	1. p.Pro385fs	0.459	1. Frameshift	47,XX,+8[21]
				2. c.1792C>T	2. p.Arg598*	0.451	2. Stopgain	
DM-5	63	M	DM	1. c.1516_1517insGGGGT	1. p.His506fs	0.384	1. Frameshift	46,XY[24]
				2. c.1919T>C	2. p.Phe640Ser	0.387	2. Missense	
DM-6	49	F	DM	c.2311C>T	p.Arg771*	0.920	Stopgain	46,XX[20]
DM-7	78	F	DM	1. c.2309C>A	1. p.Ser770*	0.426	1. Stopgain	46,XX[20]
				2. c.2043delC	2.p.Met682fs	0.430	2. Frameshift	
DM-8	53	F	DM	c.1656delC	p.Asn552fs	0.831	Frameshift	N.A.
DM-9	67	M	DM	1. c.2311C>T 2. c.2196dupT	1. p.Arg771* 2. p.Glu733fs	0.405 0.411	1. Stopgain 2. Frameshift	N.A.

Table 1. Characteristics of patients selected for whole genome sequencing.