DDX41 coordinates RNA splicing and transcriptional elongation to prevent DNA replication stress in hematopoietic cells

Satoru Shinriki1,2,15, Mayumi Hirayama2,1,15, Akiko Nagamachi2, Akihiko Yokoyama3, Takeshi Kawamura4, Akinori Kanai5, Hidehiko Kawai7, Junichi Iwakiri7, Manabu Maeshiro1,9, Saruul Tungalag1, Masayoshi Tasaki10, Mitsuharu Ueda11, Kazuhiro Tomizawa12, Naoyuki Kataoka13, Takashi Ideue14, Yutaka Suzuki6, Kiyoshi Asai8, Tokio Inaba3 and Hirota Matsui1

Myeloid malignancies with DDX41 mutations are often associated with bone marrow failure and cytopenia before overt disease manifestation. However, the mechanisms underlying these specific conditions remain elusive. Here, we demonstrate that loss of DDX41 function impairs efficient RNA splicing, resulting in DNA replication stress with excess R-loop formation. Mechanistically, DDX41 binds to the 5′ splice site (5′SS) of coding RNA and coordinates RNA splicing and transcriptional elongation; loss of DDX41 prevents splicing-coupled transient pausing of RNA polymerase II at 5′SS, causing aberrant R-loop formation and transcription-replication collisions. Although the degree of DNA replication stress acquired in S phase is small, cells undergo mitosis with under-replicated DNA being remained, resulting in micronuclei formation and significant DNA damage, thus leading to impaired cell proliferation and genomic instability. These processes may be responsible for disease phenotypes associated with DDX41 mutations.

INTRODUCTION

DDX41 mutation occurs in various hematopoietic malignancies, most frequently in acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS) [1–4]. DDX41 encodes a DEAD-box-type RNA helicase that mainly localizes in the nucleus. Proposed biological functions of nuclear DDX41 include R-loop resolution [5, 6], small nucleolar RNA processing [7] and ribosomal RNA (rRNA) processing [8]. DDX41 was also found in the spliceosome [9–12]. Notably, it has been shown that some individuals with a germline DDX41 variant in one allele later develop hematopoietic malignancies by acquiring a somatic mutation in the other allele. Most germline DDX41 variants are frameshift, nonsense, or missense mutations that occur in the entire coding region without any hotspots, which suggests that these germline variants lose expression or function. On the other hand, somatic mutations are highly concentrated in the R525H mutation located within the helicase domain where DDX41 interacts with ATP. Indeed, our previous study revealed reduced ATPase activity for the helicase domain with the R525H mutation [8]. In addition, compound heterozygous mutations combining a germline variant and the somatic R525H mutation are observed in human AML/MDS, whereas homozygous Ddx41 knockout mice are lethal [13]. Collectively, somatic mutants are considered to be functionally hypomorphic, and the acquisition of somatic mutations to cells with germline variants would be expected to further reduce the activity of DDX41 to the extent that it is not completely lost. The average age of disease onset for patients with a germline DDX41 variant is the 60s, which is comparable to that of patients without the variant [2–4, 14, 15]. However, individuals with a heterozygous germline DDX41 variant present unexplained cytopenia of one or
more hematopoietic lineages before the development of hematopoietic malignancies at the rate of 40–66% [1, 16]. Patients with DDX41 mutations often exhibit hypoplastic bone marrow, which is relatively characteristic of MDS/AML with this mutation. In addition, DDX41 mutations in MDS/AML are not necessarily mutually exclusive with those in genes encoding typical MDS-related RNA splicing factors [17–19], suggesting unique pathological implications of DDX41 mutations somewhat different from those of other MDS-related RNA splicing factors. Here, we demonstrate that DDX41 mainly binds to 5’ splice sites (SS) of coding RNA and is involved in the formation of activated spliceosomes. DDX41 was responsible for interaction between the
DDX41 is involved in RNA splicing by binding to 5' SS but does not play a major role in SS recognition. A Relative CLIP-seq signals at 5'SS and 3'SS on coding RNA. Vertical axis: ratio of CLIP sample signal divided by that of input RNA from same cells. Blue and red lines in top panels indicate relative signal enrichment of CLIP reads from cells expressing Myc-tagged WT DDX41; green and orange lines in bottom panels indicate reads from cells expressing Myc-tagged R525H mutant DDX41. B Quantification of RNA splicing changes in K562 cells expressing shDDX41#1, shDDX41#2, or SRSF2 P95R. We placed splicing events into groups according to rMATS: (1) skipped exon (SE), (2) alternative 5' SS (A5SS), (3) alternative 3' SS (A3SS), (4) mutually exclusive exons (MXE), and (5) retained intron (RI). Cumulative number of events in each cell group with an inclusion level difference (ILD) > 0.1 or <0.1 and a false discovery rate (FDR) <0.05 are shown. C Distribution of RNA splicing events in K562 cells expressing shDDX41#1, shDDX41#2, or SRSF2 P95R compared with control K562 cells. Splicing events were categorized as in B. D Changes in RNA splicing events for SE in DDX41-knockdown cells and SRSF2 P95R-expressing cells. We included splicing events with 10% minimum change of absolute percent spliced-in index (PSI, which indicates rate of incorporation of specific exons into transcript of a gene) (delta PSI ≥0.1) and average reads ≥5; those with FDR < 0.05 with ILD <0.1 or >0.1 in each group were considered significant and plotted with red or blue dots, respectively. Gray dots are not significant. E Overlap of RNA splicing events among DDX41-knockdown cells and SRSF2 P95R-expressing cells. All significant RNA splicing events (SE, MXE, RI, A5SS, and A3SS) in each cell type were summed, and event overlap among DDX41-knockdown cells and SRSF2 P95R-expressing cells is shown. F Subcellular distribution of poly(A)-tailed RNA in DDX41-knockdown cells. Scale bars: 20 μm.

RESULTS

DDX41 is an RNA splicing factor that binds mainly to the 5' SS

To investigate the role of DDX41 in RNA biogenesis, we analyzed RNAs with which DDX41 may interact, by performing ultraviolet (UV) crosslinking, immunoprecipitation (IP), and sequencing (CLIP-seq) [20, 21]. We crosslinked K562 cells expressing Myc-tagged wild-type (WT) or R525H mutant DDX41, most frequent somatic mutant found in myeloid malignancies, via UV light, and sequenced RNAs co-precipitated with Myc-DDX41 (Supplementary Fig. S1A).

First, we mapped the sequenced reads to ribosomal DNA (rDNA) and found that 60.0–65.1% CLIP-seq reads were uniquely mapped to rDNA, especially to the 18S and 28S rRNA regions, after removing duplicate reads (Supplementary Fig. S1B, C). Therefore, DDX41 likely interacts primarily with mature rRNA and may play a role in ribosome biogenesis, as we suggested previously [8]. We then mapped the remaining sequence reads to human genome and found that 24.7–28.7% CLIP-seq reads per sample were uniquely mapped to coding genes (Supplementary Fig. S1B), indicating that DDX41 also binds to coding RNA. DDX41 preferentially bound to 5’S, and somewhat less to 3’S, on coding RNAs (Fig. 1A). Mutant R525H DDX41 also bound to rRNA and coding RNA in the same manner as did WT DDX41 (Fig. 1A and Supplementary Fig. S1C), which suggests that R525H mutant retains comparable RNA-binding activity to WT DDX41, although the mutant has reduced ATPase activity [8].

Because DDX41 bound to or near SS on coding RNAs, we analyzed splicing changes using rMATS [22] by comparing RNA sequencing (RNA-seq) data for K562 cells with suppressed DDX41 expression (DDX41-knockdown cells) by DDX41-specific short hairpin RNAs (shRNA) (shDDX41#1 or shDDX41#2) with data for control cells expressing scrambled shRNA (shScramble). As a reference, we included RNA-seq data for cells expressing SRSF2 P95R, one of the RNA splicing-related mutants found in MDS, in the analysis. Similar trends of RNA splicing changes were observed in both DDX41-knockdown cells and SRSF2 P95R-expressing cells (Fig. 1B, C). However, RNA splicing changes in DDX41-knockdown cells were less frequent than those of SRSF2 P95R-expressing cells (Fig. 1B) and were either increased or decreased for each event type, with no particular trend toward either direction (Fig. 1D and Supplementary Fig. S1D), which differed from those of SRSF2 P95R-expressing cells in which skipped exon (SE) was relatively suppressed [23].

Figure 1E shows that only 202 of 703 and 593 differentially spliced events in shDDX41#1- and shDDX41#2-expressing cells, respectively, overlapped. There was also little overlap when aggregated across all shared events that could affect the same exons (Supplementary Fig. S1E). No relevant sequence features existed in differentially spliced exons and 5’SS and 3’SS (Supplementary Fig. S1F, G), although we found that the DDX41 deficiency induced splicing changes in genes partially similar to those seen in myeloid malignancies with mutations in RNA splicing factors accompanied by impaired mRNA production (Supplementary Fig. S1H, I, J, K).

These results suggest that DDX41 may be involved in RNA splicing and mRNA synthesis by interacting with 5’SS, but play smaller roles in determining RNA splicing position and inclusion/exclusion of exons. Nevertheless, RNA fluorescence in situ hybridization with oligo(dt) probes demonstrated speckle-like signals in DDX41-knockdown cell nuclei (Fig. 1F), suggesting that the loss of DDX41 caused impaired maturation or export of mRNA due to a defect in mRNA synthesis.
Fig. 2 Interaction of DDX41 with RNA splicing-related proteins. A Major GOs of proteins interacted with FLAG-DDX41. Nuclear proteins immunoprecipitated with FLAG-DDX41 were categorized via GO analysis. Top 15 GO terms for CC (cellular component) and BP (biological process) categories are shown, with the number of genes indicated by circle sizes and adjusted p values indicated by red to blue colors. B Schematic diagram showing NTC involvement in RNA splicing. The factors within NTC are incorporated into or excluded from the complex depending on the splicing steps, in which core NTC components (PRP19 and CDC5L) occur throughout NTC after incorporation of the complex into the spliceosome; CWC25 and Yju2 are incorporated before the B* complex and excluded before the C complex, and CWC27 is excluded before the C complex. C Interaction of DDX41 with RNA splicing process-specific components in the NTC. Myc-tagged NTC components (Yju2, CWC22, CWC25, CWC27, PRP19, and CDC5L) were expressed with FLAG-tagged DDX41 in HEK293FT cells, and DDX41-interacting proteins were immunoprecipitated with an anti-FLAG antibody. Precipitated proteins were probed with anti-FLAG, anti-Myc, or anti-β-Actin antibody. Left and right panels indicate input and immunoprecipitated samples, respectively. D Non-RNA-mediated interaction of DDX41 with PRP19. FLAG-tagged DDX41 and Myc-tagged PRP19 were expressed in HEK293FT cells, and FLAG-DDX41 was immunoprecipitated with anti-FLAG antibody. Precipitated samples were then treated with 20 μg/ml RNase A for 30 min at 37 °C before being probed with anti-FLAG or anti-Myc antibody.
DDX41 prevents impaired DNA replication and mitotic abnormalities

To clarify the biological significance of DDX41, we suppressed its expression in different cell lines (Fig. 3A). DDX41 knockdown with small interfering RNAs (siRNAs) (siDDX41#1 and siDDX41#2 for HeLa) or shRNAs (shDDX41#1 and shDDX41#2 for K562 and THP-1) significantly suppressed cell proliferation in all cell lines (Fig. 3B). We also observed increased apoptosis for DDX41-knockdown cells (Fig. 3C). These observations are consistent with recent reports including non-mammalian studies...
Functional inhibition of DDX41 induces mild DNA replication stress that leads to delayed G2-M transition

To better understand the roles of DDX41, we analyzed cell phenotypes after acute inhibition of DDX41, by using recently identified DDX41 inhibitors, 1-(4-ethyl-N-hydroxy-1-phenyl-2,3-dihydroquinolin-4(1H)-imine) and [4-amino-6-(2-amino-1,6-dimethylpyrimidin-4(1H)-ylidene)amino]-1-methyl-2-phenylquinolinium chloride hydrochloride (DDX41inh-1 and DDX41inh-2, respectively), which were demonstrated to inhibit ATPase activity of DDX41 [33]. These inhibitors suppressed cell growth of HeLa and K562 in a concentration-dependent manner (Supplementary Fig. S4A). Because DDX41inh-2 showed a greater anti-proliferative effect than DDX41inh-1, as reflected by their inhibitory effects on ATPase activity of DDX41 [33], we used DDX41inh-2 for subsequent analyses. We confirmed that DDX41inh-2 suppressed proliferation of both cell lines after 48–72 h of culture in a time-dependent manner (Fig. 4A). Given that DDX41 exerts RNA helicase activities in an ATPase-dependent manner [34], the helicase activity of DDX41 likely plays an important role in cell proliferation.

In accord with the knockdown experiments, DDX41inh-2 inhibited BrdU incorporation (Fig. 4B). However, the reduction in BrdU incorporation was modest at 48 h after addition of the inhibitor, when cell growth was markedly inhibited, and as long as 72 h was needed before a marked decrease in BrdU incorporation. This finding supports our hypothesis that problems resulting from DNA replication defects, rather than impaired replication itself, were responsible for the marked cell cycle changes.

Next, we investigated DNA replication in the presence of DDX41inh-2 after release from the G1/S boundary in cell cycle-synchronized HeLa cells (Fig. 4C). DDX41 inhibition subtly reduced γH2AX-positive micronuclei, a hallmark of genomic instability, which were frequently observed in DDX41-knockdown K562 cells (Fig. 3E and Supplementary Fig. S3C). These cellular changes associated with DDX41 inhibition were similar to mild replication stress observed with low concentrations of aphidicolin (APH), a DNA polymerase α inhibitor (Supplementary Fig. S4C, D) [35, 36]; APH treatment at 10 nM induced a slight delay in S-phase progression, as did the DDX41 inhibitor (Fig. 4E and Supplementary Fig. S4E).

We found that DDX41 inhibition for 6 h during S phase increased single-stranded DNA to the same extent as did 10 nM APH (Fig. 4G). Enhanced and prolonged Chk1 activation in S phase was also seen in DDX41inh-2-treated cells (Fig. 4H). These data indicated that mild DNA replication stress was induced by DDX41 inhibition. Nevertheless, an apparent increase in γH2AX signal was not observed until cells completed mitosis (Fig. 4H and Supplementary Fig. S4F), suggesting the requirement for mitosis in DDX41-inhibited cells to manifest marked DNA damage.

We also tested how DNA replication stress by DDX41 inhibition affected mitosis by arresting synchronized cells at G2 phase in the presence of DDX41inh-2, followed by its wash-off (Fig. 4I and Supplementary Fig. S4G). Cell cycle analysis showed that cells pretreated with DDX41inh-2 in S/G2 had a delayed return from 4N to 2N (Fig. 4J) and delayed G2-M transition (Fig. 4K), whereas DDX41 inhibition initiated from the end of G2 did not cause these delays (Supplementary Fig. S4H, I). These data indicated that mild DNA replication stress induced by DDX41 inhibition delayed G2-M transition, consistent with our findings showing G2 phase accumulation for asynchronized DDX41-knockdown cells (Fig. 3I).

DNA replication stress by DDX41 inhibition induces mitotic abnormalities and leaves DNA damage in post-mitotic cells

We then analyzed how DNA replication stress by DDX41 inhibition affected mitosis. We analyzed mitosis after 8 h of exposure to DDX41inh-2 during S phase without CDK1 inhibition by RO-3306...
Although the extent of lagging chromosome and multipolar mitosis was comparable for DDX41inh-2-treated cells and DMSO-treated cells, DDX41inh-2 treatment increased mitotic DNA bridges and ultrafine bridges positive for Blooms syndrome protein (BLM), which strongly suggested an increase in under-replicated DNA in S phase (Fig. 5A) [37].

DDX41 inhibition during S phase led to reduced IdU incorporation (Fig. 5B and Supplementary Fig. S5A) and subsequent G2 arrest in daughter cells (Fig. 5C, D), similar to the results in asynchronized DDX41-knockdown cells (Fig. 3F). We also found that DDX41 inhibition in S phase increased the number of nuclear γ-H2AX foci in G1 daughter cells (Fig. 5E and Supplementary Fig. S5A).
Fig. 4 DDX41 inhibition induces mild DNA replication stress in S phase. A Suppression of HeLa and K562 cell proliferation by DDX41 inhibition (50 μM DDX41inh-2 treatment). Cell number was counted by using trypan blue. Values are mean ± SD of triplicate samples; two-tailed unpaired Student’s t test. B Reduced BrdU incorporation by 50 μM DDX41inh-2 treatment in HeLa cells. C Schematic diagram of cell cycle synchronization, drug treatment, and IdU incorporation in HeLa cells. D Reduced IdU incorporation in S phase by 50 μM DDX41inh-2 treatment in HeLa cells. *p < 0.0001, two-tailed unpaired Student’s t test. Results of the 8-h treatment are not shown because S-G2 transition occurred (see Supplementary Fig. 5AB). E Delayed S-phase progression in HeLa cells by 50 μM DDX41inh-2 treatment, analyzed by representative histograms. (Right) DNA synthesis rate as estimated by median fluorescence intensity (MFI) of PI. Values are mean ± SD of triplicate samples; *p < 0.0001; †p < 0.005, two-tailed Welch’s t test. F Slowed replication fork progression by DDX41 inhibition. (Left) Experimental scheme of dual labeling with DNA analogs and representative images of DNA fibers. Thymidine analogues were visualized via immunofluorescence (CldU, green; IdU, red). (Right) Fork progression speed was calculated for each sample. Bars represent mean ± SD; n = 139 and 134 for DMSO and DDX41inh-2 group, respectively; two-tailed Welch’s t test. G Increase in single-stranded DNA by treatment with 50 μM DDX41inh-2 or 10 nM APH. Scans were: 0 μM DMSO, 10 μM APH, respectively; two-tailed Welch’s t test. H Increase in DNA damage-related signals by DDX41 inhibition in HeLa cells. Protein extracts obtained from cell cycle synchronized HeLa cells were probed with antibodies indicated at left. The time indicated are the hours after release from G1/S arrest. I Schematic diagram of mitosis assessment in HeLa cells after 50 μM DDX41inh-2 treatment during S phase. RO-3306, a CDK1 inhibitor, was used to induce G2 arrest. J, K Delayed mitosis in HeLa cells after DDX41 inhibition in S phase. HeLa cells were treated as indicated in I. Flow cytometry analysis of cell cycle change (J) and proportion of cells at M phase (K). Cells at G2 or M were determined as those negative or positive for pH3 with PI signal corresponding to 4N, respectively. Right panel in K indicates proportion of pH3-positive M phase cells 1 h after RO-3306 removal. Bars indicate means, error bars indicate SD of triplicate samples; two-tailed Welch’s t test.
Loss of the 5′SS peak in DDX41-knockdown cells did not depend on the difference in gene expression levels between DDX41-knockdown and the control cells, because a similar pattern at 5′SS was observed even when we performed the same analysis for the exons of genes expressing above the median in DDX41-knockdown cells (Supplementary Fig. S7D). These observations, along with our CLIP-seq data (Fig. 1A) demonstrating selective binding of DDX41 to 5′SS, support a unique role for DDX41 around 5′SS.

We next studied whether observed signal accumulation of Pol II at 5′SS was related to RNA splicing. We selected constitutively skipped exons in adequately transcribed genes in control cells for analysis (Supplementary Fig. S7E). We found no clear peaks at 5′SS and 3′SS on such skipped exons (Fig. 7K), although a signal fluctuation occurred because of the small number of exons included in analysis. These observations indicated that Pol II globally paused at 5′SS in association with RNA splicing similar to previous report [50], but this phenomenon no longer took

Fig. 5  Mild replication stress by DDX41 inhibition triggers mitotic abnormalities and affects cell cycle progression of daughter cells.

A Increased DNA bridges and ultrafine DNA bridges in mitotic HeLa cells treated with DDX41inh-2 during S phase. See Supplementary Fig. SSA for schematic. (Left) Representative images of abnormal mitosis. (Right) Quantitative result of abnormal mitosis. Bars indicate means; error bars, SD of triplicate samples; two-tailed Welch’s t test. n.s., not significant. Scale bars: 10 μm. B Reduced IdU incorporation in cells that had been treated with DDX41inh-2 in S phase and had undergone mitosis. See Supplementary Fig. SSA for schematic. Bars indicate means; error bars, SD of triplicate samples; two-tailed Welch’s t test. C, D Cell cycle arrest at G2 phase in HeLa cells that had been treated with DDX41inh-2 in S phase and had undergone mitosis. Cells were treated as in Fig. 4I. Cell cycle status 18 h after removal of DDX41inh-2 and RO-3306 was identified by PI staining (C). Cells were double-stained with PI and anti-pHH3 antibody to distinguish mitotic cells from cells at G2 (D). E Increase in γ-H2AX foci in G1 HeLa cells that had been treated with DDX41inh-2 in S phase and had undergone mitosis. Cells were double-stained with anti-γ-H2AX antibody and DAPI. See Supplementary Fig. S5B for schematic. Bars represent means ± SD; n = 151 and 195 for DMSO and DDX41inh-2, respectively; two-tailed Welch’s t test. Scale bars: 10 μm. F Increase in γ-H2AX signals primarily occurred at G1 in HeLa cells after DDX41 knockdown. Cells were stained with anti-γ-H2AX and anti-pHH3 antibodies and PI. MFI of γ-H2AX in each cell cycle phase was analyzed with flow cytometry. Bars indicate means; error bars, SD of triplicate samples; two-tailed Student’s t test.
DDX41 expression was decreased, even though RNA splicing changes occurred only in a subset of exons (Fig. 1). This result may be attributed to the reduced interaction of PRP19 with S2a and pSP II (Fig. 7L), especially with pS2 Pol II, which was supported by findings from other groups that RNA splicing depends on the interaction of PRP19 to Pol II CTD [51–54]. All our data suggest that DDX41 coordinates RNA splicing and transcriptional elongation at 5'SS, thereby inhibiting aberrant R-loop accumulation and subsequent DNA replication stress.

### DDX41 inhibition causes R-loop accumulation and DNA damage in primary hematopoietic cells

We developed Ddx41^{R525H} conditional knock-in (cKI) mice that express a Ddx41 mutant corresponding to human DDX41 R525H in a tamoxifen-inducible manner (Supplementary Fig. S8A, B) and examined its phenotypes of immature hematopoietic cells cultured ex vivo. Because our previous reports showed that R525H mutation reduced ATPase activity of DDX41 [8], we expected that induction of this mutant would reduce Ddx41 function. Lineage marker-negative/c-Kit-positive (Lin^-/c-Kit^+) primary hematopoietic cells infected with lentivirus expressing shDDX41#1, shDDX41#2, or shScramble were treated with ActD or DMSO for 6 h. Cells were stained with S9.6 antibody and DAPI. Scale bars: 20 μm. (Right) Nuclear S9.6 signal intensity. Bars represent means ± SD; n = 2645, 1628, 1490, 2357, 2095 and 2401 nuclei for shScramble/ActD^−, shScramble/ActD^+, shDDX41#1/ActD^−, shDDX41#1/ActD^+, shDDX41#2/ActD^−, and shDDX41#2/ActD^+, respectively; two-tailed Welch's t test. A Increase in nuclear S9.6 signals by DDX41 inhibition. HeLa cells were treated with 50 μM DDX41inh-2 or DMSO as in Fig. 4C. Six hours after release from the G1/S boundary, cells were stained with S9.6 antibody and DAPI. Scale bars: 10 μm. (Right) Nuclear S9.6 signal intensity. Bars represent means ± SD; n = 187 and 407 for DMSO and DDX41inh-2, respectively; two-tailed Welch’s t test. B Increase in nuclear S9.6 signals in S phase by DDX41 inhibition. HeLa cells were treated with 50 μM DDX41inh-2 or DMSO as in Fig. 4C. Six hours after release from the G1/S boundary, cells were stained with S9.6 antibody and DAPI. Scale bars: 10 μm. (Right) Nuclear S9.6 signal intensity. Bars represent means ± SD; n = 187 and 407 for DMSO and DDX41inh-2, respectively; two-tailed Welch’s t test. C Reduced γ-H2AX signals by enforced expression of RNase H1 in DDX41-knockdown cells. K562 cells were infected with lentivirus expressing shDDX41#1, shDDX41#2, or shScramble. Five days later, cells were transfected with FLAG-tagged nuclear-localizing RNase H1-expressing vector or an empty vector; 2 days later, cells were stained with anti-γ-H2AX and anti-FLAG antibodies and DAPI. Scale bars: 20 μm. (Right) γ-H2AX signal intensity for cells negative and positive for RNase H1 signal. Bars represent means ± SD; n = 394, 214, 342, 232 and 612 for shScramble, shDDX41#1/WT-RNase H1^−, shDDX41#1/RNase H1^+, shDDX41#1/D210N-RNase H1^−, and shDDX41#1/D210N-RNase H1^+, respectively; two-tailed Welch’s t test.
immature bone marrow cells isolated from heterozygous (Ddx41<sup>R525H/WT</sup>) and WT (Ddx41<sup>WT/WT</sup>) mice were cultured in the presence of cytokines to support stem/progenitor cell growth, followed by induction of R525H mutation in Ddx41<sup>R525H/WT</sup> cells by adding (Z)-4-hydroxytamoxifen (4-OHT). Induction of the mutant was confirmed by direct sequencing (Supplementary Fig. S8C). As in our previous study [8], induction of R525H mutation inhibited cell proliferation (Fig. 8A). Moreover, these cells showed increased R-loop formation, phosphorylation of replication protein A 32 (RPA32) at serine 4 and 8 residues—a marker of fork collapse [55].
— and hY2AX signals (Fig. 8B, C), which confirmed that increased R-loop and DNA damage from DDX41 inhibition observed in cancer cell lines also occurred in primary immature hematopoietic cells. Ddx41−R525H/WT cells further showed increased micronucleus formation and abnormal nuclear morphology (Fig. 8D, E), which suggested involvement of loss of Ddx41 function in genomic instability. Finally, we observed R-loops in cultured hematopoietic cells from Ddx41 heterozygous (Ddx41WT/KO) mice. Increased R-loop formation was observed in Ddx41−WT/KO cells, which was further enhanced by the induction of R525H mutation (Ddx41−R525H/KO) (Supplementary Fig. 8D). These observations suggest that a reduction of Ddx41 expression level causes certain dose-dependent disturbance and that R525H mutation is functionally hypomorphic.

In conclusion, our study identified a process by which loss of DDX41 expression or its helicase activity caused splicosomal dysfunction and impaired transcriptional elongation, thus leading to DNA replication stress that resulted in genomic instability. Our data suggest a model in which DDX41 regulates Pol II pausing at 5′SS with the NTC, and Pol II waits there for RNA splicing to continue, but if DDX41 is deficient, transcriptional elongation machinery may proceed without slowing down at 5′SS (Fig. 8F) or may terminate elongation and dissociate from chromatin. Although DNA replication defects in the absence of DDX41 are not necessarily severe, persistence and gradual accumulation of minor replication stress beyond mitosis would cause hematopoiesis failure that is often observed in bone marrow of patients with DDX41 mutation.

DISCUSSION

Mutation or aberrant expression of genes encoding an RNA helicase occurs in various malignancies [56–58], which suggests that abnormal RNA recombination is closely associated with tumorigenesis. In hematopoietic malignancies, thus far, only DDX41, encoding a DEAD-box-type RNA helicase, and DHX15 and DHX34, encoding DEAH-box-type helicases, were reproducibly mutated in AML/MDS [31, 59, 60]. Characterization of these gene products would provide an excellent model for clarifying biological roles of RNA helicases and its involvement in leukemogenesis.

Here, we discovered that inhibition of DDX41 led to DNA replication stress with increased R-loop formation. Although the degree of the stress was modest, cells underwent mitosis with under-replicated DNA, resulting in genomic instability and growth inhibition. Hematopoietic stem cells (HSCs) have relatively longer cell cycles and divide at a low rate, once every 40 weeks [61]. Therefore, most HSCs remain quiescent in a steady state and have fewer opportunities to enter S phase. However, HSCs become more likely to enter the cell cycle when they are exposed to inflammatory environments or when they are aged. Importantly, cycling aged HSCs in mice have higher levels of replication stress associated with cell cycle defects; such replication stress persists even after the cells re-establish quiescence [62]. Therefore, in the presence of DDX41 mutations, even weak replication stress can accumulate in aged HSCs. In addition, considering the sequential acquisition of DDX41 mutations observed in myeloid malignancies, hematopoietic cells with a heterozygous germline DDX41 variant would have to wait until they acquire somatic mutations in other alleles to develop a myeloid malignancy, because DNA replication stress acquired in cells with a heterozygous variant may be limited even in aged HSCs. This was also suggested from the phenotypes observed in Ddx41 genetically modified mice [7]. These ideas may explain why carriers of germline DDX41 variants generally develop myeloid malignancy at older ages.

We also showed that DNA damage caused by loss of DDX41 function was dependent on transcription-related R-loop accumulation. Although impaired RNA splicing can lead to R-loop formation [63], its process remains poorly understood. Accumulating evidences implicate the interdependence of RNA splicing and transcriptional elongation [64]. Specifically, a recent study demonstrated that the upstream 5′SS remained associated to the transcription machinery during intron synthesis [50], consistent with our observation. Furthermore, Pol II paused at 5′SS resumed elongation after completion of RNA splicing [65]. Interestingly, our ChIP-seq analysis with antibody against Pol II showed that Pol II likely ignores pausing at 5′SS when not enough DDX41 is available. Taken together, we suggest that Pol II continues aberrant elongation without waiting for RNA splicing to finish, or alternatively, terminates elongation at 5′SS when RNA splicing is delayed by DDX41 deficiency. This model might be linked to a new perspective that the main obstacle to replication fork progression is the elongating Pol II engaged in R-loop [66]. DDX41 depletion induced mild splicing changes without any specific pattern; one possible explanation for this is that DDX41 regulates efficient splicing in the late step rather than decision of SS. However, further sequencing analysis of clinical specimens with DDX41 mutation is clearly needed for its confirmation, given that INT53 intron retention observed in DDX41-knockdown cells was reported to be highly exclusive to SRSF2 mutation [67]. Since only a subset of R-loops can be hotspots for DNA damage [68],

SS with the NTC, and Pol II waits there for RNA splicing to

Fig. 7 Changes in gene expression and distribution of Pol II by DDX41 knockdown. A Dependence of cell lines on DDX41. We used DepMap portal (https://depmap.org/portal/). For density distributions for CRISPR and RNAi data, smaller scores indicate that DDX41 is essential for cell line survival; –1 was comparable to the median of all pan-essential genes. B Genes co-dependent with DDX41 include those related to RNA splicing and Pol II-mediated transcription. Top co-dependent genes with DDX41 (with q values <0.05) identified in CRISPR screening were subjected to GO analysis; results were visualized via gProfiler (upper). Representative GO terms related to RNA splicing (red) and Pol II-mediated transcription (blue) were numbered (lower). Table S1 gives the complete list. C Gene expression changes by DDX41 knockdown. An hierarchical clustering of 1341 genes that showed expression changes with p < 0.05 in common in shDDX41#1- and shDDX41#2-expressing DDX41-knockdown cells compared with shScramble-expressing control cells was visualized. D Representative gene sets associated with RNA splicing and transcriptional elongation negatively enriched in shDDX41#1- and shDDX41#2-expressing cells. ES, enrichment score; NES, nominal enrichment score. E Representative gene sets associated with transcriptional elongation and RNA processing negatively enriched in DDX41 low-expressing AML cases. The 451 AML cases presented in the article by Tyner et al. [48] were divided into three groups according to the expression level of DDX41, and the transcriptome differences between groups with DDX41 expression levels below mean −SD (DDX41 low) and above mean +SD (DDX41 high) were examined. F Direct interaction of DDX41 with Pol II in HEK293FT cells. Protein extracts from cells expressing FLAG-DDX41 were immunoprecipitated with anti-FLAG antibody or control IgG and then probed with anti-FLAG and anti-Pol II (pS2 and pS5) antibodies.

Average distribution of total Pol II from transcription start site (TSS) to transcription end site (TES) of all RefSeq transcripts visualized with Ngsplot. G Changes in Pol II expression in DDX41-knockdown HEK293FT cells. L J Average distribution of Pol II around exon/intron boundaries. Distribution of Pol II around 5′SS (I) and 3′SS (J) of genes expressing above the median in control cells are shown. K Average distribution of Pol II around exon/intron boundaries of constitutively spliced exons. Exons that met requirements of coverage >20 and average PSI > 0.5 were selected, and average distributions of Pol II up to 1000 bases upstream and downstream of the 5′SS (left panel) and 3′SS (right panel) of exons were visualized with Ngsplot. L Reduced interaction of PRP19 with pS2- and pS5-modified Pol II. Protein extracts from DDX41-knockdown HEK293FT expressing Myc-tagged PRP19 were immunoprecipitated with anti-Myc antibody or control IgG. The samples were probed with anti-Pol II (pS2 and pS5), anti-Myc and anti-DDX41 antibodies.

Dependence of cell lines on DDX41. We used DepMap portal (https://depmap.org/portal/). For density distributions for CRISPR and RNAi data, smaller scores indicate that DDX41 is essential for cell line survival; –1 was comparable to the median of all pan-essential genes. B Genes co-dependent with DDX41 include those related to RNA splicing and Pol II-mediated transcription. Top co-dependent genes with DDX41 (with q values <0.05) identified in CRISPR screening were subjected to GO analysis; results were visualized via gProfiler (upper). Representative GO terms related to RNA splicing (red) and Pol II-mediated transcription (blue) were numbered (lower). Table S1 gives the complete list. C Gene expression changes by DDX41 knockdown. An hierarchical clustering of 1341 genes that showed expression changes with p < 0.05 in common in shDDX41#1- and shDDX41#2-expressing DDX41-knockdown cells compared with shScramble-expressing control cells was visualized. D Representative gene sets associated with RNA splicing and transcriptional elongation negatively enriched in shDDX41#1- and shDDX41#2-expressing cells. ES, enrichment score; NES, nominal enrichment score. E Representative gene sets associated with transcriptional elongation and RNA processing negatively enriched in DDX41 low-expressing AML cases. The 451 AML cases presented in the article by Tyner et al. [48] were divided into three groups according to the expression level of DDX41, and the transcriptome differences between groups with DDX41 expression levels below mean −SD (DDX41 low) and above mean +SD (DDX41 high) were examined. F Direct interaction of DDX41 with Pol II in HEK293FT cells. Protein extracts from cells expressing FLAG-DDX41 were immunoprecipitated with anti-FLAG antibody or control IgG and then probed with anti-FLAG and anti-Pol II (pS2 and pS5) antibodies.

Average distribution of total Pol II from transcription start site (TSS) to transcription end site (TES) of all RefSeq transcripts visualized with Ngsplot. G Changes in Pol II expression in DDX41-knockdown HEK293FT cells. L J Average distribution of Pol II around exon/intron boundaries. Distribution of Pol II around 5′SS (I) and 3′SS (J) of genes expressing above the median in control cells are shown. K Average distribution of Pol II around exon/intron boundaries of constitutively spliced exons. Exons that met requirements of coverage >20 and average PSI > 0.5 were selected, and average distributions of Pol II up to 1000 bases upstream and downstream of the 5′SS (left panel) and 3′SS (right panel) of exons were visualized with Ngsplot. L Reduced interaction of PRP19 with pS2- and pS5-modified Pol II. Protein extracts from DDX41-knockdown HEK293FT expressing Myc-tagged PRP19 were immunoprecipitated with anti-Myc antibody or control IgG. The samples were probed with anti-Pol II (pS2 and pS5), anti-Myc and anti-DDX41 antibodies.
regions and types of splicing defect may determine the impact on genomic stability. We conclude that the vulnerability in DNA replication that partially remains in daughter cells would be essential to explain the unique phenotype of DDX41-related myeloid malignancies.

**MATERIALS AND METHODS**

**Cell lines**

K562 and THP-1 cells were cultured in RPMI-1640 (Sigma-Aldrich) with 10% heat-inactivated fetal bovine serum. HEK293FT, HEK293T, HeLa, U2OS and ARPE-19 cells were cultured in Dulbecco’s Modified Eagle Medium (Sigma-
Fig. 8  R-loop accumulation and DNA damage by introduction of R525H mutation in primary hematopoietic progenitor cells in mice. A Reduced proliferation of immature bone marrow cells expressing R525H mutation. Lin−/c-Kit− bone marrow cells isolated from 9-week-old Ddx41R525Hfl+/− or Ddx41WTfl− mice were cultured as in A were stained with anti γ-H2AX antibody (Left) Representative images of Ddx41R525Hfl+/− cell with γ-H2AX-positive micronucleus (arrowhead). Scale bars: 50 μm. F Schematic illustration of how DDX41 deficiency causes impaired hematopoiesis and leukemogenesis. (a) Normal condition: DDX41, together with NTC, coordinates RNA splicing and transcriptional elongation. The Pol II complex transiently slows at 5′ SS and 3′ SS, which leads to increased opportunities for R-loop formation and transcription-replication conflicts. This results in increased under-replicated DNA at end of replication and DNA bridge formation during mitosis that result in genomic instability.

Aldrich) with 10% heat-inactivated fetal bovine serum. All cells were maintained in a 5% CO2 incubator at 37 °C.

DDX41 inhibitors
DDX41inh-1 and DDX41inh-2 were gifts from Axcelead Drug Discovery Partners Inc.

Cell cycle synchronization
HeLa cells were synchronized to the late G2 phase or G1/S boundary by double thymidine block with or without CDK1 inhibitor, respectively (see Supplementary Methods).

Flow cytometry
Cell cycle distribution, incorporation of BrdU or IdU, apoptosis and γ-H2AX expression were analyzed via flow cytometry (see Supplementary Methods).

CLIP-seq and ChiP-seq analyses
CLIP-seq was performed according to a previously reported method (Supplementary Fig. S1A) [20, 21]. ChiP-seq was performed by applying the fractionation-assisted native ChiP method (49). Detailed procedures are described in Supplementary Methods.

Mice
Ddx41fl−/− mice were crossed with C57BL/6-Gt(Rosa)26Sor1m1FLn(wt/Emx1)Artm (ERT2Cre) mice (Model6466, Taconic Biosciences) to allow tamoxifen-inducible tamoxifen excision of floxed regions. Ddx41 heterozygous knockout mice (C57BL/6N Ddx41fl+/−) were purchased from the UC Davis KOMP Repository (MMRRC Stock #047340-UCD). All mice were kept according to guidelines of the Institute of Laboratory Animal Science, Hiroshima University. The Animal Care Committee at the Japanese Foundation for Cancer Research approved all murine studies. Detailed procedures for generating Ddx41R525Hfl−/− mice and experimental procedures with the mice are described in Supplementary Methods.

Statistical analysis
Statistical analysis was performed with R software (version 4.1.2). All statistical details of experiments are given in Supplementary Methods and corresponding figure legends, with p values ≤0.01 being considered statistically significant unless otherwise indicated. Error bars in figures indicate standard deviation (SD).

Additional methods are available in the Supplementary Information.

DATA AVAILABILITY
RNA-seq, ChiP-seq, and CLIP-seq data are available in the sequence read archive (SRA) database (https://www.ncbi.nlm.nih.gov/assembly) (accession ID for CLIP-seq: DRA011992, for RNA-seq: DRA014267, and for ChiP-seq: DRA014255). The datasets generated during the study are available from the corresponding author upon reasonable request.

REFERENCES
1. Churpek JE, Smith-Simmer K. DDX41-Associated Familial Myelodysplastic Syndrome and Acute Myeloid Leukemia. In: Adam MP, Arndt HH, Pagon RA, Wallace SE, Bean LJH, Gripp KW, et al. editors. University of Washington, Seattle: GeneReviews; 2021.
2. Li P, White T, Xie W, Cui W, Peker D, Zeng G, et al. AML with germline DDX41 variants is a clinicopathologically distinct entity with an indolent clinical course and favorable outcome. Leukemia. 2021;35:6664–74.
3. Sébert M, Passet M, Raimbault A, Rahme R, Raffoux E, Sicre de Fontbrune F, et al. Germline DDX41 mutations define a significant entity within adult MDS/AML patients. Blood. 2019;134:1441–4.
4. Yang F, Long N, Anekiparitangan T, Bottomly D, Savage JC, Lee T, et al. Identification and prioritization of myelodysplasia germline variants in a large cohort of adult patients with AML. Blood. 2022;139:1208–11.
5. Mosler T, Conte F, Longo GM, Mikcik I, Kreim N, Möllek MM, et al. R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. Nat Commun. 2021;12:7314.
6. Weinreb JT, Ghazale N, Pradhan K, Gupta V, Ports KS, Tricomi B, et al. Excessive R-loops trigger an inflammatory cascade leading to increased HSPC production. Dev Cell. 2022;56:627–40.
7. Chlon TM, Stepanchick E, Hershberger CE, Daniels NJ, Hueneman KM, Kuenzi S, et al. Unique ethnic features of myelodysplastic syndrome and acute myeloid leukemia patients: an ALFA-FILO study. Blood. 2022;140:756–65.
8. Schmidt C, Grønborg M, Deckert J, Bessonov S, Conrad T, Lührmann R, et al. Mass spectrometry-based relative quantification of proteins in pre-catalytic and catalytically active spliceosomes by metabolic labeling (SLALC), chemical labeling (ITRAQ), and label-free spectral count. RNA. 2014;20:406–20.
9. Bessonov S, Anokhina M, Krausaks A, Golas MM, Sander B, Will CL, et al. Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. RNA. 2010;16:4745–54.
10. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009;136:701–18.
11. Cvitkovic I, Jurica MS. Spliceosome database: a tool for tracking components of the spliceosome. Nucleic Acids Res. 2013;41:D132–41.
12. Schmidt C, Grönborg M, Deckert J, Bessonov S, Conrad T, Lührmann R, et al. Mass spectrometry-based relative quantification of proteins in pre-catalytic and catalytically active spliceosomes by metabolic labeling (SLALC), chemical labeling (ITRAQ), and label-free spectral count. RNA. 2014;20:406–20.
13. Ma J, Mahmud N, Bosland MC, Ross SR. DDX41 is needed for pre- and postnatal hematopoietic stem cell differentiation in mice. Stem Cell Rep. 2022;17:879–93.
14. Alkhateeb HB, Nanaa A, Viswanatha D, Foran JM, Badar T, Sproat L, et al. Genetic features and clinical outcomes of patients with isolated and comutated DDX41-mutated myeloid neoplasms. Blood Adv. 2022;6:528–32.
15. Duployez N, Largeaud L, Duchmann M, Kim R, Plesnier J, Lambert J, et al. Prognostic impact of DDX41 germline mutations in intensively treated acute myeloid leukemia patients: an ALFA-FILO study. Blood. 2022;140:756–68.
16. Choi EJ, Cho YU, Hur EH, Jang S, Kim N, Park HS, et al. Unique ethnic features of myelodysplastic syndrome and acute myeloid leukemia. Haematologica. 2022;107:510–8.
17. Quesada AE, Routbort MJ, D’Nardo CD, Bueso-Ramos CE, Kanagal-Shamanna R, Khoury JD, et al. DDX41 mutations in myeloid neoplasms are associated with male gender, TPS3 mutations and high-risk disease. Am J Hematol. 2019;94:757–66.
for providing materials and reagents. We thank Dr. Takayuki Nojima (Kyushu University, Fukuoka, Japan), Dr. Daisuke Morishita (Kumamoto University), and Dr. Yu Nagayoshi (Kumamoto University) for insightful suggestions and feedback. We thank Dr. Hiroaki Honda (Tokyo Women's Medical University, Tokyo, Japan) and Mr. Norimasa Yamasaki (Hiroshima University, Hiroshima, Japan) for assisting with gene-targeting vector construction and ES cell manipulations.

AUTHOR CONTRIBUTIONS
SS, MH, and HM designed the study, conducted experiments, and wrote the paper. AK, JI, KA, and YS performed sequencing data analysis. AN and T Inaba performed mouse experiments. AY performed DDX41 complex purification and ChIP-seq experiments. TK performed mass spectrometry analysis. HK, LR, MM, ST, MT, NK, and T Ideue conducted other experiments including proteomics analysis and RNA detection. MU, TT and KT reviewed the experimental data and supervised the study.

FUNDING
This study was supported by Japanese Society for the Promotion of Science Grants-in-Aid for Scientific Research No. 16H06279(PAGS), No. 21K08419, and No. 18K08334 (HM); No. 22K17187 (MH); and No. 22K10195 (SS). This work was also partly supported by The Uehara Memorial Foundation, the Friends of Leukemia Research Fund, the Kanae Foundation for the Promotion of Medical Science, the Senshin Medical Research Foundation, the Takeda Science Foundation, and the Foundation for Promotion of Cancer Research.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41375-022-01708-9.

Correspondence and requests for materials should be addressed to Satoru Shinriki or Hirotaka Matsui.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.