

Oncogenic DEAD-box ATPase DDX41 utilizes Cdc2-like kinase 3-dependent and -independent mechanisms to establish transcript ensembles

Corresponding Author: Professor Emery Bresnick

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Referee report for: "Oncogenic DEAD-box ATPase DDX41 utilizes Cdc2-like kinase 3-dependent and -independent mechanisms to establish transcript ensembles"

By Kim et al.

March 13th, 2025

Summary:

The authors established a genetic rescue system in which mutant proteins of interest are expressed in immortalized murine myeloid progenitors of *Ddx41*^{+/-} and/or *Clk3*^{-/-} or human *DDX41* variants, and analyzed their effect on the transcriptome upon the expression of different DDX41 mutants in various cell types and differentiation states of murine and human cultured cells.

DDX41 expression in *Ddx41*^{+/-} cells resulted in many differentially expressed transcripts and altered transcript usage. The expression of the G173R and R525H mutants caused differential expressions patterns that partially overlapped with wild-type DDX41 expression, with the somatic mutant R525H showing a larger overlap compared to the germline mutant G173R. *Clk3* was one of the seven genes in which both differential transcript usage (DTU) and differential expression (DE) were altered; specifically, intron 3 retention was induced by wild-type DDX41 expression but not by the G173R mutant and to a lesser extent by the R525H mutant. This effect was also observed in human skin fibroblasts and iPSCs expressing the K331del mutant, and differential splicing in murine cells was validated by RT-qPCR. Expression of chimeric proteins indicated that the DDX41 DEAD-box domain is essential for its activity. The regulation of DDX41 affected CLK3 in differentiated cells but not in the undifferentiated cells. Further analysis demonstrated that altered *Ddx41* and *Clk3* expression affects granulocytic and monocytic differentiation, and that some differential expression caused by DDX41 are dependent on CLK3, while some others are not, varying by differentiation state.

This manuscript provides valuable data on the function of the poorly characterized helicase DDX41 and its physiologically relevant mutants, utilizing extensive genetic and transcriptomic experiments and analyses. It also suggests a potential regulatory mechanism for the unique alternative splicing mode of the CLK genes, in which DDX41 involvement has not been previously described. However, a large portion of the manuscript consists of phrases merely explaining the numerical differences in the obtained data. It does not discuss a possible mechanism in how the helicase interacts with the RNA to induce differential splicing, with reference to the papers cited in the fourth paragraph of the introduction. Additionally, while the study demonstrates the effects of the mutants on myeloid differentiation, there is a lack of discussion on how the differentiation is influenced by the differential splicing, and vice versa. Modifying the structure of the main text and the main figures, as well as incorporating additional discussions on published results regarding mechanisms and physiological implications, would enrich the manuscript.

Major comments:

1. A large part of the main text explains the graphically obvious numerical differences which make the main text difficult to read, such as in Fig. 2b explained in lines 153-156. The number and percentages can be moved to the source data or written directly in the figures, and the main text can be shortened to only explaining the conclusion, such as deleting the

entire sentence in lines 153-156 because it is explained in lines 151-153. This can be applied to other figures such as Fig. 3i and lines 237-239 (the exact p-values should be listed in the figure itself or in the legend) and many others.

2. Although the direction of regulation is discussed for DTE in Fig. 2d, the direction of regulation in DTU for the mutants is not shown nor discussed. It is also confusing that the name of the transcripts with the three-digit numbers after the gene names, instead of only gene names, are shown in the DTE analysis.

3. The figures are too crowded; many of them can be moved to supplementary figures. For example, in Fig. 3b and c, only the relevant isoforms (-205 and -207) can be shown for the exons table and the heatmap, and the rest of them can be in the supplementary figure to be clear that the retention of intron 3 is the important difference. Fig 3f -RT (right side) proves that genomic DNA is not contaminated and is not overestimating retention, and Fig. 3h/Fig. 6e proves that the fractionation was effective; they are essential experiments but are not essential to be shown in the main figure.

4. In Fig. 6k, the blotting of the Tubulin standard is not evenly blotted and is not functioning as a standard. It should be redone so that there is no signal bias between the two ends of the signals, and that the intensity of the signals should be similar enough between the samples.

5. At lines 320-321, it is explained as "Four transcripts were analyzed to test if CLK3 mediates DDX41-dependent splicing" but in the same paragraph, including explanations of Fig7 a-c, it only explains changes in transcript levels under DDX41 expression, not involving CLK3. It is understandable that this data is required to explain Fig. 7h results, but because CLK3 is not relevant at this point, it is better to be moved to supplementary figures and restructure the section to center it with the double mutant and refer to these four genes when Fig. 7h is explained.

6. While the manuscript provides important differences between the germline variants G173R and K331del against the somatic variant R525H in expression and splicing changes, the manuscript does not discuss how their differences are relevant to oncogenesis. It seems contradictory that while R525H rescues more splicing than G173R, the latter is allowed to exist as a germline mutation, and why germline mutations do not show up as somatic mutations.

7. As mentioned in the summary, there is no discussion on how the helicase activity is relevant to the changes in the transcriptome profile, such as unwinding double strands or acting against the spliceosomal complexes. The effect of the mutations on the helicase activity is also not discussed.

Minor comments:

- The title on the main text "Oncogenic DEAD-box ATPase DDX41 establishes transcript ensembles via CLK3-dependent and -independent mechanisms" is different from the one shown on the Supplementary Information and the manuscript tracking system "Oncogenic DEAD-box ATPase DDX41 utilizes Cdc2-like kinase 3-dependent and -independent mechanisms to establish transcript ensembles".

- On line 209, Citation is required for the statement "As intron retention increases the probability of transcript accumulation in the nucleus".

- On line 216, there is an extra period after "versus", which is only required if it is abbreviated as "vs."

- On line 318, it is explained that there are "515 transcripts in DTE analysis that were not regulated by pathogenic variants", but 515 is a number that includes those overlapping with mutant-regulated genes according to Fig. 2c, and Ybx1 and Ddx5 is altered in R525H so "that were not regulated by pathogenic variants" is incorrect.

- On line 410, m6A should be superscripted as m⁶A.

- For Fig. 7b, the consequences of different isoforms should be explained (e.g. exon skipping that causes lack of residues in the DNA-binding domain for Ybx1), and isoform diagram should be shown in supplementary material.

Reviewer #2

(Remarks to the Author)

Kim et al established a genetic rescue system with Ddx41 +/- myeloid progenitors to study the global splicing perturbations upon DDX41 pathogenic variants. As the DDX41 pathological mutations play very important roles in MDS and AML, the authors' work represents important contributions to the field. However, there are several major issues yet to be addressed for the publication at Nature Communications:

1) The authors relied on DTUrtle-DRIMseq-SURF for alternative splicing/transcript isoform study upon DDX41 splicing global changes. To avoid potential tool-specific biases in these global discoveries, The authors should independently run their splicing analysis with alternative splicing tools, including MISO (<https://github.com/yarden/MISO>) and QUANTAS(https://zhanglab.c2b2.columbia.edu/index.php/Quantas_Documentation). It would be convincing if the splicing alternation phenotypes could be observed in other splicing tools (e.g. CLK3 intron retentions)

2) The authors compared the splicing rescue of DDX41 WT with G173R, R525H and then K331 del variants. The underlying logic and background why picking these pathological variants is missing. Furthermore, how the current global splicing profiling between DDX41 WT and pathological variants contribute to the patient pathological outcome. This is an important "for what" question that the authors need to address thoroughly.

3) The authors found DDX41 utilized CLK3-dependent and CLK3-independent mechanisms to regulate transcript isoforms. The current data provided by the authors are very limited. A detailed comparison between these two types of regulation targets would help the study and enhance the paper impact.

4) Throughout the manuscript, the authors mentioned a few times that DDX41 regulates splicing of a few other targets besides CLK3 including YTHDF2 which is a well-known factor regulating m6A-dependent mRNA decay (PMID: 28637692; 32492408 and 24284625). Luo et al established a state-of-art deep learning approach in predicting m6A in human transcripts (<https://doi.org/10.1038/s41467-022-30209-7>, the predicted human m6A sites could be downloaded at <https://zenodo.org/records/4734266>). Those m6A sites are of high quality as being experimentally validated independently

(Figure 3 of <https://doi.org/10.1038/s41467-022-30209-7> and Figure 3 of <https://doi.org/10.1038/s41467-023-39897-1>) It is necessary that the authors do a comprehensive check if the transcripts which have changed mRNA level upon DDX41 rescue also have enriched m6A sites in them. It would rule out or prove that there is a crosstalk of m6A regulation and splicing regulation of DDX41 upon its downstream targets.

Reviewer #3

(Remarks to the Author)

Kim et. al. use murine HoxB8 immortalized cell lines and human iPSC lines that are heterozygous for DEAD-box helicase 41 (DDX41) to study the role of DDX41 in alternative splicing. Using retroviral overexpression of wild type Ddx41, the authors identify transcripts with differential expression and/or usage when Ddx41 expression is corrected. In particular, they identify an intron retention event in the dual specificity kinase, CLK3, that leads to increased expression of Clk3 mRNA in undifferentiated and differentiated murine HSPCs. The authors then generate double mutant Clk3^{-/-} Ddx41^{+/-} HoxB8 immortalized cell lines and use their retroviral correction system to evaluate the role of Ddx41 and Clk3 on the splicing of 4 additional genes in undifferentiated and differentiated cells, concluding there are Clk3-dependent and Clk3-independent splicing events associated with Ddx41. The experiments were thorough and well executed, however it seems a missed opportunity to primarily focus on splicing rather than functional consequences including differentiation and additional gene expression changes associated with hemizygous or mutant Ddx41 expression in their elegant model systems.

Major concerns:

1. It is interesting that the authors find that the addition of a somatic variant to Ddx41^{+/-} cells causes fewer splicing changes than correcting Ddx41 haploinsufficiency. The authors need to show that correction of Ddx41 expression with wild type or somatic or germline variants in Ddx41^{+/-} cells with retroviral transduction leads to levels of wild type or mutant Ddx41 expression that are comparable to what is observed in patients. Overexpression, particularly of the DDX41 variants, may affect the assessment of the effects of these variants on RNA splicing.
2. The effects of Ddx41 correction vs variant complementation in Ddx41^{+/-} cells on differentiation are not presented in a manner that is independent of the splicing changes. It would be helpful to see how the different variants affect the behavior of immortalized human and murine HSPCs before diving into the splicing changes.
3. The differential transcript and isoform usage analyses provide a unique opportunity to identify downstream consequences of genes with splicing changes. The authors nicely show the intron 3 retention event is associated with increased expression of Clk3. Are similar changes observed in the other 6 genes highlighted in Supplementary Figure 2?
4. Differential transcript (Figure 1b) and isoform (Figure 1c) usage are shown for several genes in Figure 1. A summary table should be included to note whether the different transcripts or isoforms are expected to cause decreased gene expression (via NMD or generation of a transcript that is not translated) or expression of an altered form of the protein. The percentage of the alternative splicing events that lead to gene expression or protein isoform changes caused by Ddx41 correction vs variant complementation of Ddx41^{+/-} cells should also be presented.
5. CLK3 acts on serine/threonine and tyrosine-containing substrates including serine- and arginine-rich (SR) proteins of the spliceosomal complex. The authors should evaluate for differential localization or phosphorylation of SR proteins in Ddx41^{+/-} vs Ddx41 corrected HoxB8 cells in undifferentiated and differentiated states.
6. The authors should clearly show the effect of Clk3^{-/-} on differentiation of wild type and Ddx41^{+/-} Hoxb8 cells.
7. The authors show the effects of DDX41 on CLK3 intron 3 retention are specific to the DDX41 DEAD box. How many of the differential isoform events identified in Figure 1 are specific to the DDX41 DEAD box? Is it all of them?
8. The authors do not show whether Clk3 overexpression affects differentiation of the different model systems, or the splicing changes observed in Ythdf2 and Khdc4. Is complete loss of this gene well tolerated as cells differentiate or is this associated with a differentiation block?
9. Are any of the 4 target genes studied in Figure 7 predicted to be targets of Clk3 phosphorylation?

Minor concerns:

1. Please include a schematic with the locations of the germline and somatic Ddx41 variants included in this study.
2. Are G173R germline variants commonly seen with DDX41 hemizygous loss or would it be more relevant to study this variant in combination with a somatic variant?
3. Please explain/show how TPMs were used to calculate the ratio of each isoform to the total transcript in Figure 3c and 3d.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Thank you for responding to the points in detail. The authors have addressed most of my concerns.

I do, however, have a few additional comments:

- Excessive reporting of numerical values in the main text

My initial concern regarding the overuse of numerical values (percentages, counts, and statistical results) is not limited to the examples in Fig. 2b (lines 153–156) and Fig. 3i (lines 237–239), but applies more broadly to other figure–text correspondences as well. In many cases, these values simply restate increases or decreases already described, serving only to emphasize statistical significance. This makes the main text unnecessarily dense. Specific examples include:

Fig. 2f and p-values at line 163 (should be moved to legend or source data)

Fig. 6e and fold change/p-values at line 296

Fig. 6f and fold change/p-values at lines 302–303

Fig. 6g (left) and fold change/p-values at line 307

Supplementary Fig. 10b and p-values at lines 373–374 (should be in legend or figure)

Additionally, p-values reported in the main text as “ $p < \dots$ ” should instead be given as exact values. Since all are significant, reiterating them in the main text only adds visual clutter without improving clarity.

- Discussion of mutation effects

In response to major comment 7, the authors have added important information on how mutations affect the ATPase activity of the helicase. While this is valuable, the discussion would benefit from extending beyond the physiological effects. Specifically, the authors should elaborate on molecular–mechanistic aspects: how mutations alter interactions with proteins, RNAs, or DNAs, and how these altered interactions mechanistically connect to observed splicing changes, in addition to discussing cell-level or transcriptome-wide effects. For instance, structural considerations (e.g., <https://www.nature.com/articles/srep34756>) and comparisons across multiple DEAD-box helicases—not just DDX3X—would strengthen the discussion.

Reviewer #2

(Remarks to the Author)

The authors have addressed my comments well. I would recommend the prompt acceptance of this manuscript, and its timely publishing at Nature Communications would benefit the research community.

Reviewer #3

(Remarks to the Author)

The authors have addressed all of my major concerns. It would be helpful to include a statement summarizing the functional differences between the DDX41 mutations selected for this study in the introduction to bolster the rationale for their inclusion in this manuscript.

Minor comment: I think “ransition” should be “transition” in line 475.

Open Access This Peer Review File 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 license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Reviewers

Reviewer 1:

Summary:

The authors established a genetic rescue system in which mutant proteins of interest are expressed in immortalized murine myeloid progenitors of *Ddx41*^{+/-} and/or *Clk3*^{-/-} or human *DDX41* variants, and analyzed their effect on the transcriptome upon the expression of different DDX41 mutants in various cell types and differentiation states of murine and human cultured cells. DDX41 expression in *Ddx41*^{+/-} cells resulted in many differentially expressed transcripts and altered transcript usage. The expression of the G173R and R525H mutants caused differential expressions patterns that partially overlapped with wild-type DDX41 expression, with the somatic mutant R525H showing a larger overlap compared to the germline mutant G173R. *Clk3* was one of the seven genes in which both differential transcript usage (DTU) and differential expression (DE) were altered; specifically, intron 3 retention was induced by wild-type DDX41 expression but not by the G173R mutant and to a lesser extent by the R525H mutant. This effect was also observed in human skin fibroblasts and iPSCs expressing the K331del mutant, and differential splicing in murine cells was validated by RT-qPCR. Expression of chimeric proteins indicated that the DDX41 DEAD-box domain is essential for its activity. The regulation of DDX41 affected CLK3 in differentiated cells but not in the undifferentiated cells. Further analysis demonstrated that altered *Ddx41* and *Clk3* expression affects granulocytic and monocytic differentiation, and that some differential expression caused by DDX41 are dependent on CLK3, while some others are not, varying by differentiation state.

This manuscript provides valuable data on the function of the poorly characterized helicase DDX41 and its physiologically relevant mutants, utilizing extensive genetic and transcriptomic experiments and analyses. It also suggests a potential regulatory mechanism for the unique alternative splicing mode of the CLK genes, in which DDX41 involvement has not been previously described. However, a large portion of the manuscript consists of phrases merely explaining the numerical differences in the obtained data. It does not discuss a possible mechanism in how the helicase interacts with the RNA to induce differential splicing, with reference to the papers cited in the fourth paragraph of the introduction. Additionally, while the study demonstrates the effects of the mutants on myeloid differentiation, there is a lack of discussion on how the differentiation is influenced by the differential splicing, and vice versa. Modifying the structure of the main text and the main figures, as well as incorporating additional discussions on published results regarding mechanisms and physiological implications, would enrich the manuscript.

Major comments:

1. A large part of the main text explains the graphically obvious numerical differences which make the main text difficult to read, such as in Fig. 2b explained in lines 153-156. The number and percentages can be moved to the source data or written directly in the figures, and the main text can be shortened to only explaining the conclusion, such as deleting the entire sentence in lines 153-156 because it is explained in lines 151-153. This can be applied to other figures such as Fig. 3i and lines 237-239 (the exact p-values should be listed in the figure itself or in the legend) and many others.

Response: Thank you for your suggestion. We moved the “% of increased or decreased transcripts and p-values” into the figure legends. (lines 897-899 and 921-926)

2. Although the direction of regulation is discussed for DTE in Fig. 2d, the direction of regulation in DTU for the mutants is not shown nor discussed. It is also confusing that the name of the transcripts with the three-digit numbers after the gene names, instead of only gene names, are shown in the DTE analysis.

Response: We added the direction of the proportional change for each transcript by DDX41 and variants in DTU analysis (Supplementary Fig.2, lines 172-174). The proportional expression of commonly regulated transcripts by DDX41 variants changed in the same direction as DDX41.

We quantified expression of each transcript (DTE), which provides a more informative analysis than strictly relying on consolidated differential expression of genes (DE). Three-digit labeling is essential to distinguish each transcript isoform.

3. The figures are too crowded; many of them can be moved to supplementary figures. For example, in Fig. 3b and c, only the relevant isoforms (-205 and -207) can be shown for the exons table and the heatmap, and the rest of them can be in the supplementary figure to be clear that the retention of intron 3 is the important difference. Fig 3f -RT (right side) proves that genomic DNA is not contaminated and is not overestimating retention, and Fig. 3h/Fig. 6e proves that the fractionation was effective; they are essential experiments but are not essential to be shown in the main figure.

Response: We removed isoform genomic structure diagrams, except for the most relevant isoforms and the fully spliced transcripts (Fig.3b and c). We also removed the data in which reverse transcriptase was omitted (-RT) from the main figure (Fig.3f). This data was incorporated into Supplementary Fig.5a-5c. Western blots of laminB/tubulin in Fig. 3h and Fig. 6e were moved to Supplementary Fig.5d and 8a.

4. In Fig. 6k, the blotting of the Tubulin standard is not evenly blotted and is not functioning as a standard. It should be redone so that there is no signal bias between

the two ends of the signals, and that the intensity of the signals should be similar enough between the samples.

Response: We blotted β -actin as a loading control and incorporated this data, since β -actin is more effective as a loading control in this context.

5. At lines 320-321, it is explained as “Four transcripts were analyzed to test if CLK3 mediates DDX41-dependent splicing” but in the same paragraph, including explanations of Fig7 a-c, it only explains changes in transcript levels under DDX41 expression, not involving CLK3. It is understandable that this data is required to explain Fig. 7h results, but because CLK3 is not relevant at this point, it is better to be moved to supplementary figures and restructure the section to center it with the double mutant and refer to these four genes when Fig. 7h is explained.

Response: Fig.7a-c were moved into Supplementary Fig.10d and e.

6. While the manuscript provides important differences between the germline variants G173R and K331del against the somatic variant R525H in expression and splicing changes, the manuscript does not discuss how their differences are relevant to oncogenesis. It seems contradictory that while R525H rescues more splicing than G173R, the latter is allowed to exist as a germline mutation, and why germline mutations do not show up as somatic mutations.

Response: Our results from genome-wide transcript analysis revealed this surprising finding. The importance of helicase activity for transcript isoform profiles was assessed by R525H (in the helicase domain) variant regulated profiles. Despite expression of variant proteins at comparable levels, R525H was modestly defective relative to severe defects detected with DEAD-box domain (G173R and K331del) germline variants. Myeloid malignancy involving *DDX41* germline variants involves a late age of onset and is often accompanied by somatic mutation. Given this scenario and lethality of homozygous knockout *Ddx41* mice, one can predict that heterozygosity of a severely defective germline variant would impact DDX41-regulated processes by ~50%, assuming there is not excess DDX41 in the cell. Acquiring a modestly defective somatic variant would further derail DDX41-regulated activities, avoiding cell death that would result from complete DDX41 loss or two severely defective alleles. While there are hotspots for germline mutations and somatic mutations, mutations can occur as either germline or somatic mutations. R525H is a common somatic mutation, yet there are cases where it appears as a germline mutation in patients. We incorporated text on lines 431-438 to articulate this important issue more clearly.

7. As mentioned in the summary, there is no discussion on how the helicase activity is relevant to the changes in the transcriptome profile, such as unwinding double strands

or acting against the spliceosomal complexes. The effect of the mutations on the helicase activity is also not discussed.

Response: R525H has been reported to exhibit a lower ATPase activity compared to DDX41. K_{cat} values for DDX41 and R525H proteins were 954 ± 45.5 and $293 \pm 23.6/s^{-1}$, respectively. There is no evidence to suggest that the DEAD box-localized variants, K331del and G173R would impact ATPase activity. Although DEAD-box helicases bind and remodel DNA/RNA and protein complexes in an ATP-dependent fashion, whether diminished helicase activity impacts DDX41-dependent alternative splicing has not been elucidated. We demonstrated a modest influence of R525H on transcript isoforms, considerably less severe than for the DEAD-box domain variants. We incorporated text on lines 419-424 to articulate this important issue more clearly.

Minor comments:

- The title on the main text “Oncogenic DEAD-box ATPase DDX41 establishes transcript ensembles via CLK3-dependent and -independent mechanisms” is different from the one shown on the Supplementary Information and the manuscript tracking system “Oncogenic DEAD-box ATPase DDX41 utilizes Cdc2-like kinase 3-dependent and -independent mechanisms to establish transcript ensembles”.

Response: We corrected the title in the Supplementary information.

- On line 209, Citation is required for the statement “As intron retention increases the probability of transcript accumulation in the nucleus”.

Response: We incorporated the citation: Kensuke Ninomiya, Naoyuki Kataoka, Masatoshi Hagiwara, J Cell Biol (2011) 195 (1): 27–40. Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor

- On line 216, there is an extra period after “versus”, which is only required if it is abbreviated as “vs.”

Response: We corrected it.

- On line 318, it is explained that there are “515 transcripts in DTE analysis that were not regulated by pathogenic variants”, but 515 is a number that includes those overlapping with mutant-regulated genes according to Fig. 2c, and Ybx1 and Ddx5 is altered in R525H so “that were not regulated by pathogenic variants” is incorrect.

Response: We removed the paragraph and rewrote it to describe the new data, incorporating the new RNA-seq results from DDX41/CLK3 double-mutant cells.

- On line 410, m6A should be superscripted as m⁶A.

Response: We corrected it.

- For Fig. 7b, the consequences of different isoforms should be explained (e.g. exon skipping that causes lack of residues in the DNA-binding domain for Ybx1), and isoform diagram should be shown in supplementary material.

Response: We incorporated text describing the consequences of the isoforms on lines 350-359 and integrated an isoform diagram into Supplementary Table 3.

We thank the reviewer for insightful comments and constructive criticism.

Reviewer 2:

Kim et al established a genetic rescue system with Ddx41+/- myeloid progenitors to study the global splicing perturbations upon DDX41 pathogenic variants. As the DDX41 pathological mutations play very important roles in MDS and AML, the authors' work represents important contributions to the field. However, there are several major issues yet to be addressed for the publication at Nature Communications:

1) The authors relied on DTUrtle-DRIMseq-SURF for alternative splicing/transcript isoform study upon DDX41 splicing global changes. To avoid potential tool-specific biases in these global discoveries, The authors should independently run their splicing analysis with alternative splicing tools, including MISO (<https://github.com/yarden/MISO>) and QUANTAS(https://zhanglab.c2b2.columbia.edu/index.php/Quantas_Documentation). It would be convincing if the splicing alternation phenotypes could be observed in other splicing tools (e.g. CLK3 intron retentions).

Response: Thank you for the important recommendation. We conducted new analyses with these tools. MISO and QUANTAS require and provide alternative splicing event reference files for mouse and human genomes. For mice, however, they provide a version older than mm39, which we used for our analysis (e.g., MISO provides mm9 and mm10). With MISO, we were able to create and apply newer versions of these files, but QUANTAS is limited. To match the version of the genome annotation file used in our analysis, we used SUPPA² as an alternative method and applied its reference file to MISO analysis.

MISO and SUPPA tools consider distinct splicing events (Supplementary Fig.3a and b) and analyze transcript regulation events at different levels. SURF analyzes transcript level, while MISO and SUPPA analyze the event level. Using SUPPA, 359 splicing events with a Δ PSI change of 10% or more ($p < 0.05$) were DDX41-regulated, 136 by G173R, and 244 by R525H (Supplementary Fig.3c; 62% and 32% reduction, respectively). Using MISO, 162 splicing events were DDX41-regulated, 61 by G173R, and 139 by R525H (Δ PSI change $> 10\%$; Bayes factor > 10 ; Supplementary Fig.3f; 62% and 14% reduction, respectively). With all tools, G173R defects were more severe than R525H, and R525H was less active than DDX41 in regulating splicing. These important additional comparisons were incorporated and discussed on lines 183-191 and 214-216) and supplementary Fig.3.

2) The authors compared the splicing rescue of DDX41 WT with G173R, R525H and then K331 del variants. The underlying logic and background why picking these pathological variants is missing. Furthermore, how the current global splicing profiling between DDX41 WT and pathological variants contribute to the patient pathological

outcome. This is an important “for what” question that the authors need to address thoroughly.

Response: G173R is a recurrent germline variant, and R525H is the common somatic variant. The discovery of K331del as a human pathogenic germline variant was reported in our prior study³. This variant has a reduced capacity to regulate mRNA levels and surface marker expression upon myeloid differentiation. These attributes provided rationale for our effort to analyze functional properties of these mutants.

Compared to DDX41, G173R regulated 90% fewer transcripts. Since pre-mRNA splicing is important for protein and organismal diversity, the degree of disruption of splicing patterns is an important molecular metric almost certainly informs pathogenicity. However, further studies are required to establish definitive genotype-phenotype relationships. We described these important considerations on lines 424-430.

3) The authors found DDX41 utilized CLK3-dependent and CLK3-independent mechanisms to regulate transcript isoforms. The current data provided by the authors are very limited. A detailed comparison between these two types of regulation targets would help the study and enhance the paper impact.

Response: We appreciate this important recommendation. We conducted new experimentation to extend our prior RNA-seq data that identified four transcripts as being regulated by CLK3-DDX41-regulated alternative splicing. To deeply investigate the genetic relationship between DDX41 and CLK3, we engineered *Clk3*^{-/-}:*Ddx41*^{+/-} and *Clk3*^{+/-}:*Ddx41*^{+/-} cells and used a new rescue system to test whether CLK3 is required for global DDX41 transcript-regulatory activities. We conducted RNA-seq analysis in this system, which demonstrated most DDX41-regulated transcripts require CLK3. A small number of CLK3-independent transcripts were also detected (Supplementary Fig.10a). Importantly, DDX41 regulated transcripts from 19 genes with functional roles in mRNA splicing and mRNA transport. These new results were incorporated as Fig.7d and e and Supplementary Fig.10a-c and described on lines 341-349 and 368-374.

4) Throughout the manuscript, the authors mentioned a few times that DDX41 regulates splicing of a few other targets besides CLK3 including YTHDF2 which is a well-known factor regulating m6A-dependent mRNA decay (PMID: 28637692; 32492408 and 24284625). Luo et al established a state-of-art deep learning approach in predicting m6A in human transcripts (<https://doi.org/10.1038/s41467-022-30209-7>, the predicted human m6A sites could be downloaded at <https://zenodo.org/records/4734266>). Those m6A sites are of high quality as being experimentally validated independently (Figure 3 of <https://doi.org/10.1038/s41467-022-30209-7> and Figure 3 of <https://doi.org/10.1038/s41467-023-39897-1>) It is

necessary that the authors do a comprehensive check if the transcripts which have changed mRNA level upon DDX41 rescue also have enriched m6A sites in them. It would rule out or prove that there is a crosstalk of m6A regulation and splicing regulation of DDX41 upon its downstream targets.

Response: We appreciate the important recommendation. CLK3-dependent regulation of *Ybx1* and *Ythdf2* transcripts has potential to be important. We conducted new analyses to evaluate the predicted m⁶A stringent motif, RRACH (R = A or G, H = A, C, or U), in sequences of genes generating DDX41-regulated transcripts using the genome-wide motif dataset from im6A modeling by Luo et al ⁷. Among the DDX41-regulated transcripts, 72 genes (16%) contain the RRACH motif; 98 genes (22%) contain the RAC motif (prediction probability >0.7). These results suggest that m⁶A has a role in the mechanism by which DDX41 regulates transcripts. This is further supported by DDX41 regulation of *Ybx1* and *Ythdf2*. We incorporated this result into Supplementary Fig.14 and discussed on lines 444-450.

Reviewer 3:

Kim et. al. use murine HoxB8 immortalized cell lines and human iPSC lines that are heterozygous for DEAD-box helicase 41 (DDX41) to study the role of DDX41 in alternative splicing. Using retroviral overexpression of wild type Ddx41, the authors identify transcripts with differential expression and/or usage when Ddx41 expression is corrected. In particular, they identify an intron retention event in the dual specificity kinase, CLK3, that leads to increased expression of Clk3 mRNA in undifferentiated and differentiated murine HSPCs. The authors then generate double mutant Clk3^{-/-} Ddx41^{+/-} HoxB8 immortalized cell lines and use their retroviral correction system to evaluate the role of Ddx41 and Clk3 on the splicing of 4 additional genes in undifferentiated and differentiated cells, concluding there are Clk3-dependent and Clk3-independent splicing events associated with Ddx41. The experiments were thorough and well executed, however it seems a missed opportunity to primarily focus on splicing rather than functional consequences including differentiation and additional gene expression changes associated with hemizygous or mutant Ddx41 expression in their elegant model systems.

Major concerns:

1. It is interesting that the authors find that the addition of a somatic variant to Ddx41^{+/-} cells causes fewer splicing changes than correcting Ddx41 haploinsufficiency. The authors need to show that correction of Ddx41 expression with wild type or somatic or germline variants in Ddx41^{+/-} cells with retroviral transduction leads to levels of wild type or mutant Ddx41 expression that are comparable to what is observed in patients. Overexpression, particularly of the DDX41 variants, may affect the assessment of the effects of these variants on RNA splicing.

Response: Thank you for raising the critical point. Previously, we described DDX41 and variants in hi-*Ddx41*^{+/-} cells are expressed comparably to *Ddx41*^{+/+}.³ To extend these studies, we analyzed endogenous and expressed DDX41 in iPSCs (Fig. 3i). Expressed DDX41 was not overexpressed relative to endogenous DDX41 in hi-cells and human iPSCs. A sentence was added on lines 148-149 and 230 to indicate this.

2. The effects of Ddx41 correction vs variant complementation in Ddx41^{+/-} cells on differentiation are not presented in a manner that is independent of the splicing changes. It would be helpful to see how the different variants affect the behavior of immortalized human and murine HSPCs before diving into the splicing changes.

Response: We demonstrated DDX41 haploinsufficiency increased monocytic and reduced granulocytic cells. DDX41 and R525H expression in hi-*Ddx41*^{+/-} cells reduced monocytic and increased granulocytic cells. K331del and G173R variants lacked or reduced activity - reported in our previous paper.³

In the current paper, we analyzed the global transcript isoform repertoire to elucidate mechanisms underlying these functional differences. While we have not analyzed DDX41 in HSPCs, and our primary focus was to discover global mechanisms in myeloid progenitors, expanding these findings to additional cell types in the hematopoietic hierarchy will be an important extension of the current work.

3. The differential transcript and isoform usage analyses provide a unique opportunity to identify downstream consequences of genes with splicing changes. The authors nicely show the intron 3 retention event is associated with increased expression of *Clk3*. Are similar changes observed in the other 6 genes highlighted in Supplementary Figure 2?

Response: We identified 7 genes regulated at differential gene expression and differential transcript usage levels by DDX41 in Supplementary Fig.4a (*Cc2d1b*, *Chchd1*, *Clk3*, *Fam133b*, *Gm47283* (lncRNA), *Prpf18*, *Slc25a28*). To extend this data, we conducted new experimentation to quantify mRNA levels of DDX41-regulated *Cc2d1b*, *Chchd1*, *Clk3*, *Fam133b*, *Prpf18* in undifferentiated and differentiated cells by RT-qPCR (Supplementary Fig.4b). DDX41 increased *Slc25a28* and *Prpf18* mRNAs in undifferentiated and differentiated cells. DDX41 did not significantly affect the steady-state protein levels in undifferentiated or differentiated cells (Supplementary Fig.4c). In RNA-seq data, the proportion of *Slc25a28-202* (intron-retained form) increased from 17% to 20%, while *Slc25a28-201* (encoding full-length protein) decreased from 22% to 19%. For *Prpf18* gene, two transcripts (*Prpf18-201* and *Prpf18-204*) encode the full-length protein. The proportion of *Prpf18-201* decreased from 70% to 59%, while that of *Prpf18-204* increased from 26% to 35%. In this context, the small magnitude of the changes might not suffice to alter steady-state protein levels. We described these results on lines 451-460.

4. Differential transcript (Figure 1b) and isoform (Figure 1c) usage are shown for several genes in Figure 1. A summary table should be included to note whether the different transcripts or isoforms are expected to cause decreased gene expression (via NMD or generation of a transcript that is not translated) or expression of an altered form of the protein. The percentage of the alternative splicing events that lead to gene expression or protein isoform changes caused by Ddx41 correction vs variant complementation of Ddx41^{+/-} cells should also be presented.

Response: The expected functional consequences of regulated transcript isoforms and proportional changes were incorporated into Supplementary Table 1.

5. CLK3 acts on serine/threonine and tyrosine-containing substrates including serine- and arginine-rich (SR) proteins of the spliceosomal complex. The authors should

evaluate for differential localization or phosphorylation of SR proteins in *Ddx41*^{+/-} vs *Ddx41* corrected HoxB8 cells in undifferentiated and differentiated states.

Response: We appreciate the important recommendation. We conducted new experimentation to analyze SR protein phosphorylation in *Ddx41*^{+/+} and *Ddx41*^{+/-} cells with DDX41 rescue in undifferentiated and differentiated states using anti-phosphoSR antibody (clone 1H4; Supplementary Fig.12). DDX41 regulated phospho-SR proteins with apparent *M_r* of 92, 76, 47, 37, 30, and 23 kDa (Supplementary Fig.12). The proteins decreased in differentiated *Ddx41*^{+/-} cells, and DDX41 increased their levels; 37, 30, and 23 kDa were unaffected in undifferentiated cells. The 92, 76, and 47 kDa proteins decreased in *Ddx41*^{+/-} cells, and DDX41 increased their levels in undifferentiated cells. Thus, DDX41 regulates SR protein phosphorylation in a context-dependent manner. We incorporated these results into Supplementary Fig.12 and discussed on lines 322-330.

6. The authors should clearly show the effect of *Clk3*^{-/-} on differentiation of wild type and *Ddx41*^{+/-} Hoxb8 cells.

Response: We conducted experimentation to analyze the impact of DDX41 on differentiation. DDX41 increased CD11b⁺CD115⁻ (granulocytic) and decreased CD11b⁺CD115⁺ (monocytic) populations in *Ddx41*^{+/-} and *Clk3*^{-/-};*Ddx41*^{+/-} lines (Supplementary Fig.13; lines 375-378).

7. The authors show the effects of DDX41 on CLK3 intron 3 retention are specific to the DDX41 DEAD box. How many of the differential isoform events identified in Figure 1 are specific to the DDX41 DEAD box? Is it all of them?

Response: G173R variant in DEAD-box domain was severely defective in regulating transcript isoforms. G173R regulated only 10% of transcripts regulated by wild type DDX41 (Fig.2c). This result provide evidence that the intact DEAD-box domain is important for transcript-regulatory activity, and text was added on lines 257-260.

We also demonstrated that a DEAD-box chimera failed to increase the levels of *Ybx1-202*, *Ddx5-208*, *Ythdf2-202*, and *Khdc4-206* transcripts. Thus, the DDX41 DEAD box has unique determinants required to regulate these transcripts (added in Supplementary Fig. 7, lines 274-275).

8. The authors do not show whether *Clk3* overexpression affects differentiation of the different model systems, or the splicing changes observed in *Ythdf2* and *Khdc4*. Is complete loss of this gene well tolerated as cells differentiate or is this associated with a differentiation block?

Response: We conducted experimentation to quantify myeloid differentiation of *Clk3*^{-/-} progenitors by flow cytometry with surface markers CD11b and CD115 in Fig.7a.

Genetic rescue of CLK3 (with expression level comparable to the endogenous level) increased the granulocytic population (CD11b⁺CD115⁻) 1.5-fold ($p < 0.0001$), resembling DDX41 rescue in *Ddx41*^{+/-} progenitors. This suggests that CLK3 also regulates myeloid differentiation. We discovered that *Ythdf2* and *Khdc4* transcripts decreased in *Clk3*^{-/-} cells and increased upon CLK3 rescue in undifferentiated and differentiated conditions. We incorporated this qPCR result into Supplementary Fig.11 and text on lines 331-334 and 395-396).

9. Are any of the 4 target genes studied in Figure 7 predicted to be targets of Clk3 phosphorylation?

Response: We are unaware of links between these four target genes and DDX41 or CLK3 and cannot predict CLK3 substrates. Although CLK3 phosphorylates splicing factors, in principle, it might have a broader cadre of substrates, and this will be important to investigate in the future.

Minor concerns:

1. Please include a schematic with the locations of the germline and somatic Ddx41 variants included in this study.

Response: The locations of the germline and somatic DDX41 variants used in this study were incorporated into Fig.2a.

2. Are G173R germline variants commonly seen with DDX41 hemizygous loss or would it be more relevant to study this variant in combination with a somatic variant?

Response: G173R germline variant patients usually have a somatic variant in DDX41, including R525H.

3. Please explain/show how TPMs were used to calculate the ratio of each isoform to the total transcript in Figure 3c and 3d.

Response: We used TPMs to calculate the ratio of each isoform to total transcript (Percentage of [TPM value of an isoform/Sum of TPM value of all isoforms of the gene]; Fig.3c and 3d). We added text on lines 210-211 to state this more clearly.

- 1 Singh, R. S. *et al.* DDX41 is required for cGAS-STING activation against DNA virus infection. *Cell Rep* **39**, 110856 (2022). <https://doi.org/10.1016/j.celrep.2022.110856>
- 2 Alamancos, G. P., Pages, A., Trincado, J. L., Bellora, N. & Eyraes, E. Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA* **21**, 1521-1531 (2015). <https://doi.org/10.1261/rna.051557.115>
- 3 Kim, J. A. *et al.* Discriminating activities of DEAD-Box Helicase 41 from myeloid malignancy-associated germline variants by genetic rescue. *Leukemia* **37**, 235-239 (2023). <https://doi.org/10.1038/s41375-022-01753-4>
- 4 Kovalenko, A. *et al.* The tumour suppressor CYLD negatively regulates NF-kappaB signalling by deubiquitination. *Nature* **424**, 801-805 (2003). <https://doi.org/10.1038/nature01802>
- 5 van Andel, H., Kocemba, K. A., Spaargaren, M. & Pals, S. T. Aberrant Wnt signaling in multiple myeloma: molecular mechanisms and targeting options. *Leukemia* **33**, 1063-1075 (2019). <https://doi.org/10.1038/s41375-019-0404-1>
- 6 Vekariya, U. *et al.* DNA polymerase theta protects leukemia cells from metabolically induced DNA damage. *Blood* **141**, 2372-2389 (2023). <https://doi.org/10.1182/blood.2022018428>
- 7 Luo, Z., Zhang, J., Fei, J. & Ke, S. Deep learning modeling m(6)A deposition reveals the importance of downstream cis-element sequences. *Nat Commun* **13**, 2720 (2022). <https://doi.org/10.1038/s41467-022-30209-7>

Response to Reviewers

Reviewer #1 (Remarks to the Author):

Thank you for responding to the points in detail. The authors have addressed most of my concerns.

I do, however, have a few additional comments:

- Excessive reporting of numerical values in the main text

My initial concern regarding the overuse of numerical values (percentages, counts, and statistical results) is not limited to the examples in Fig. 2b (lines 153–156) and Fig. 3i (lines 237–239), but applies more broadly to other figure–text correspondences as well. In many cases, these values simply restate increases or decreases already described, serving only to emphasize statistical significance. This makes the main text unnecessarily dense. Specific examples include:

Fig. 2f and p-values at line 163 (should be moved to legend or source data)

Fig. 6e and fold change/p-values at line 296

Fig. 6f and fold change/p-values at lines 302–303

Fig. 6g (left) and fold change/p-values at line 307

Supplementary Fig. 10b and p-values at lines 373–374 (should be in legend or figure)

Additionally, p-values reported in the main text as “ $p < \dots$ ” should instead be given as exact values. Since all are significant, reiterating them in the main text only adds visual clutter without improving clarity.

Response: Thank you for the recommendation. We moved the p-values/fold changes to the figure legends (lines 955-960, 1047-1048, 1048-1049, 1052-1053, legend in supplementary fig.10). In addition, the p-values expressed as ranges have been modified to exact values (lines 219 and 241).

- Discussion of mutation effects

In response to major comment 7, the authors have added important information on how mutations affect the ATPase activity of the helicase. While this is valuable, the discussion would benefit from extending beyond the physiological effects. Specifically, the authors should elaborate on molecular–mechanistic aspects: how mutations alter interactions with proteins, RNAs, or DNAs, and how these altered interactions mechanistically connect to observed splicing changes, in addition to discussing cell-level or transcriptome-wide effects. For instance, structural considerations (e.g., <https://www.nature.com/articles/srep34756>) and comparisons across multiple DEAD-box helicases—not just DDX3X—would strengthen the discussion.

Response: Thank you for the recommendation. R525H has reduced ATPase activity compared to DDX41 and exhibited reduced unwinding activity with duplex DNA, and

normal annealing activity.¹ There is no evidence to suggest that the DEAD box variants, K331del and G173R would impact these activities. Our analysis revealed only modestly reduced R525H activity, considerably less than the DEAD-box variants G173R and K331del. From a molecular perspective, these two mutant residues are in the DEAD domain, which is associated with substrate binding. DEAD-box helicases bind and reassemble DNA/RNA and protein complexes in an ATP-dependent manner, and protein-substrate interaction affinity will be a factor controlling this process. Structural comparisons with other helicase family members that share high sequence similarity with DDX41 may yield additional insights into how DDX41 mutants interact with substrates or other spliceosome proteins. We added this to discussion (lines 417-426).

Reviewer #2 (Remarks to the Author):

The authors have addressed my comments well. I would recommend the prompt acceptance of this manuscript, and its timely publishing at Nature Communications would benefit the research community.

Response: Thanks for your consideration.

Reviewer #3 (Remarks to the Author):

The authors have addressed all of my major concerns. It would be helpful to include a statement summarizing the functional differences between the DDX41 mutations selected for this study in the introduction to bolster the rationale for their inclusion in this manuscript.

Minor comment: I think “ransition” should be “transition” in line 475.

Response: Thank you for your consideration and comment. The typo was corrected.

1 Singh, R. S. *et al.* DDX41 is required for cGAS-STING activation against DNA virus infection. *Cell Rep* **39**, 110856 (2022).
<https://doi.org/10.1016/j.celrep.2022.110856>