

RESEARCH ARTICLE

Tumor Markers and Signatures

Attenuated cell cycle and DNA damage response transcriptome signatures and overrepresented cell adhesion processes imply accelerated progression in patients with lower-risk myelodysplastic neoplasms

Monika Kaisrlikova¹  | David Kundrat¹  | Pavla Koralkova^{2,3}  |
 Iva Trsova^{1,4}  | Zuzana Lenertova^{1,5}  | Hana Votavova¹  |
 Michaela Dostalova Merkerova¹  | Zdenek Krejcik¹  | Jitka Vesela¹  |
 Martin Vostry¹  | Radka Simeckova¹ | Marketa Stastna Markova¹  |
 Marie Lauermannova¹ | Anna Jonasova⁶  | Jaroslav Cermak¹ |
 Vladimir Divoky^{2,3}  | Monika Belickova^{1,7} 

¹Institute of Hematology and Blood Transfusion, Prague, Czech Republic

²Department of Biology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

³Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

⁴Faculty of Science, Charles University, Prague, Czech Republic

⁵First Faculty of Medicine, Charles University, Prague, Czech Republic

⁶First Department of Medicine, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic

⁷Institute of Clinical and Experimental Hematology, First Faculty of Medicine, Charles University, Prague, Czech Republic

Correspondence

Monika Belickova, Institute of Hematology and Blood Transfusion, U Nemocnice 1, 128 20 Prague 2, Czech Republic.
 Email: monika.belickova@uhkt.cz

Vladimir Divoky, Department of Biology, Faculty of Medicine and Dentistry Palacky University, Hnevotinska 3, 779 00 Olomouc, Czech Republic
 Email: vladimir.divoky@upol.cz

Funding information

Ministry of Health of the Czech Republic, Grant/Award Numbers: UHKT,00023736, NU-21-03-00565; Next Generation EU, Programme EXCELES, Grant/Award Number: LX22NPO5102; European Biobanking and Biomolecular Research Infrastructure, Grant/Award Number: LM2023033; Charles University, Grant/Award Number: UNCE/

Abstract

Patients with myelodysplastic neoplasms (MDS) are classified according to the risk of acute myeloid leukemia transformation. Some lower-risk MDS patients (LR-MDS) progress rapidly despite expected good prognosis. Using diagnostic samples, we aimed to uncover the mechanisms of this accelerated progression at the transcriptome level. RNA-seq was performed on CD34+ ribodepleted RNA samples from 53 LR-MDS patients without accelerated progression (stMDS) and 8 who progressed within 20 months (prMDS); 845 genes were differentially expressed ($\log_{2}FCI > 1$, $FDR < 0.01$) between these groups. stMDS CD34+ cells exhibited transcriptional signatures of actively cycling, megakaryocyte/erythrocyte lineage-primed progenitors, with upregulation of cell cycle checkpoints and stress pathways, which presumably form a tumor-suppressing barrier. Conversely, cell cycle, DNA damage response (DDR) and energy metabolism-related pathways were downregulated in prMDS samples, whereas cell adhesion processes were

Vladimir Divoky and Monika Belickova contributed equally to our study.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *International Journal of Cancer* published by John Wiley & Sons Ltd on behalf of UICC.

MED/016; UP Young Researcher Grant Competition, Grant/Award Number: JG_2023_016

upregulated. Also, prMDS samples showed high levels of aberrant splicing and global lncRNA expression that may contribute to the attenuation of DDR pathways. We observed overexpression of multiple oncogenes and diminished differentiation in prMDS; the expression of *ZEB1* and *NEK3*, genes not previously associated with MDS prognosis, might serve as potential biomarkers for LR-MDS progression. Our 19-gene DDR signature showed a significant predictive power for LR-MDS progression. In validation samples (stMDS = 3, prMDS = 4), the key markers and signatures retained their significance. Collectively, accelerated progression of LR-MDS appears to be associated with transcriptome patterns of a quiescent-like cell state, reduced lineage differentiation and suppressed DDR, inherent to CD34⁺ cells. The attenuation of DDR-related gene-expression signature may refine risk assessment in LR-MDS patients.

KEYWORDS

accelerated progression, lower-risk, MDS, transcriptome, tumorigenesis barrier

What's new?

The mechanisms of transformation from myelodysplastic syndrome to acute myeloid leukemia are still unclear, with some lower-risk MDS patients progressing rapidly despite a good prognosis. This comprehensive analysis of the CD34⁺ cell transcriptome identified several molecular signatures underlying accelerated progression in lower-risk MDS patients. Early progression (within 20 months from diagnosis) appeared to be associated with decreased cell cycle and metabolic activity, downregulated DNA damage response and dysregulated cell adhesion gene expression. The suppression of CD34⁺ cell intrinsic DNA damage response seemed to be interconnected with decreased lineage differentiation, and may predict progression in patients with lower-risk MDS.

1 | INTRODUCTION

Myelodysplastic neoplasms (MDS) are a heterogeneous group of disorders characterized by clonal hematopoiesis. Patients with MDS are classified according to their risk of transformation to acute myeloid leukemia (AML) using international prognostic scoring systems. The mechanisms of disease progression are still unclear, and some lower-risk MDS patients (LR-MDS) progress rapidly despite a good prognosis.¹ Identification of patients at risk of rapid progression and early initiation of effective treatment may significantly improve patient outcomes.

In the era of microarrays, a few classifiers and gene signatures connected to a poor prognosis have been introduced for MDS; however, these have not been implemented in clinical practice.^{2,3} Currently, despite the massive availability of the RNAseq technique, only a limited number of studies have been published on the MDS transcriptome, especially those focusing on classification and risk prediction.^{4,5} The heterogeneity of MDS makes these studies difficult.

Recently, we have proposed that DNA damage response (DDR) activation forms an intrinsic antitumor barrier in LR-MDS CD34⁺ cells that counteracts cellular transformation, and this barrier may be disrupted by *RUNX1* mutations.⁶ Here, as an extension, we aimed to analyze the transcriptome of LR-MDS CD34⁺ cells regardless of their driver mutations, to identify the molecular mechanisms underlying rapid progression. Using RNAseq, we demonstrated that CD34⁺ cells from diagnosis of LR-MDS with accelerated progression (within

20 months from diagnosis; progressive MDS; prMDS) exhibited distinct pro-malignant gene expression signatures and attenuation of potential intrinsic tumor-suppressing barrier compared to CD34⁺ cells of stable LR-MDS patients (stMDS). For later confirmation of our findings, we used a set of validation samples. When confirmed by further studies, these signatures may allow to predict the probability of progression for patients with LR-MDS.

2 | METHODS

2.1 | Patients and samples

We enrolled 61 LR-MDS patients according to the IPSS-R⁷ (0–3.5 points). Eight patients (13%) progressed within 20 months, while stable MDS patients were monitored at least 20 months from diagnosis. The progression was defined according to the revised criteria of the International Working Group.⁸ The median age of the cohort was 65 years (range, 23–85 years). The median follow-up period was 52 months (range, 9–208 months). Three patients underwent hematopoietic stem cell (HSC) transplantation (HSCT) after progression and were followed until the date of HSCT for the purposes of this study. The characteristics of the patients are summarized in Table S1A,B. The overall survival curves are depicted in Figure S1. Samples of seven healthy individuals were used as healthy controls (median age was 45 years; range, 29–69 years).

Validation sample set consisted of 7 LR-MDS patients (median age was 68, range, 54–75 years, Table S2), for whom full data became available after the completion of the discovery cohort analyses.

2.2 | Sampling

Bone marrow (BM) CD34+ cells from diagnostic samples were isolated by magnetic separation on an autoMACS Separator (Miltenyi Biotec, Bergisch Gladbach, Germany). RNA was isolated by acid guanidinium thiocyanate-phenol-chloroform extraction and measured with Qubit 3.0 fluorometer (Life Technologies, Carlsbad, CA). RNA quality was checked using an Agilent 2100 bioanalyzer (Agilent Technologies, Palo Alto, CA). Only high-quality RNA samples with an RNA integrity number of at least 7.5 were used. The samples were ribodepleted prior to library preparation using the RiboCop rRNA Depletion Kit (Lexogen, Wien, Austria).

Somatic mutations from BM or peripheral blood were detected by TruSight Myeloid Sequencing Panel (Illumina, San Diego, CA).⁶

2.3 | RNA sequencing and data analysis

The library was prepared with the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA) according to the manufacturers' instructions and sequenced on HiSeq 2500 and NovaSeq (Illumina). The sequencing coverage and quality statistics for each sample are summarized in Tables S3A (discovery cohort) and S3B (validation sample set). The data analysis pipeline has been described previously.⁶ Briefly, the reads were mapped to the human genome GRCh38.p13. R software 4.0.2 packages (described in Ref. 6) and GraphPad Prism 7 software (GraphPad Software, La Jolla, CA) were used for analyses and data visualization. Analysis of protein–protein interactions was performed using the online tool String 11.0; gene set enrichment analysis (GSEA) was performed in GSEA software 3.0. Gene Ontology (GO) and Reactome Pathways were used for functional annotation. To remove redundant GO terms, the REVIGO online tool was used.⁹ Differential alternative splicing events were detected by rMATS 4.1.2. Mutation analysis from RNAseq data is described in the Supplementary Methods in Data S1.

2.4 | Immunohistochemistry

BM formalin-fixed paraffin-embedded sections from 9 LR-MS patients (highlighted in Figure S2A) were stained with CD34, ZEB1 and IDH2 antibodies detailed in the Supplementary Methods in Data S1.

2.5 | Statistical analysis

Unpaired *t* tests, Fisher's exact tests and Mann–Whitney tests were used to compare groups to identify differences between two

groups and Cox proportional hazard regression was used for multivariate analysis. Kaplan–Meier curves were generated in GraphPad Prism 7 software. The level of statistical significance was set at 0.05 unless indicated otherwise in the text. Data were assumed to be non-normal (tested by Shapiro–Wilk test). The description of statistical approaches used for gene expression signatures and the cumulative incidence analysis are detailed in the Supplementary Methods in Data S1. Before statistical analysis of validation sample set, batch effect between discovery and validation sample sets was treated using limma package in R software. When assessing differential gene expression of candidate markers, these were first tested in the discovery cohort and then combined with validation samples into groups of all stMDS and prMDS analyzed cases for stronger statistical power of the boxplot constructions. Also, all candidate gene expression correlations were first tested in the discovery cohort but are presented in combination with validation samples for stronger associations. Whenever validation cohort samples are present in figures, they are differentiated from discovery cohort by color or shape. Alternative splicing analysis in rMATS, GSEA, gene signature analysis and cumulative incidence analysis were performed only in the discovery cohort.

3 | RESULTS

3.1 | Differentially expressed genes between stMDS and prMDS samples

First, using RNAseq on CD34+ ribodepleted RNA, we aimed to uncover possible transcriptome differences between stMDS and prMDS, who progressed within 20 months. The arbitrary 20 months cut-off was used for the separation of patients with accelerated progression, abnormal for LR-MDS group. We hypothesized that there are pre-existing molecular differences in CD34+ cells at diagnosis that might drive LR-MDS progression, and that these differences (transcriptional signatures) become apparent phenotypically within this 20-month period in the patients with such accelerated progression. prMDS showed a higher level of consistency in their expression profiles (Figure S2A), and in the principal component analysis (PCA), they formed a marginal group of LR-MDS; transcriptome of some prMDS and stMDS samples partially overlapped (Figure S2B). To confirm that we dichotomized the patients based purely on MDS progression without other unrecognized causes of disease severity, we evaluated expression of two components of WNT signaling, known to be associated with differential outcome in MDS/AML. We observed significant downregulation of *LEF1*, a marker of poor outcome in MDS,¹⁰ and significant upregulation of *LRP6*, a marker of AML leukemia stem cells (LSCs) activity¹¹ in prMDS vs stMDS (Figure S3). The differential expression analysis (DEA) showed 378 significantly ($\log_{2}FCI > 1$, $FDR < 0.01$) upregulated and 467 significantly downregulated genes in stMDS (Figure 1A). Protein-coding genes (PCGs) included 369 upregulated and 328 downregulated genes in stMDS. We further analyzed those 369 upregulated PCGs that may include PCGs that form the barrier protecting cells from progression. In total, 237 GO Biological

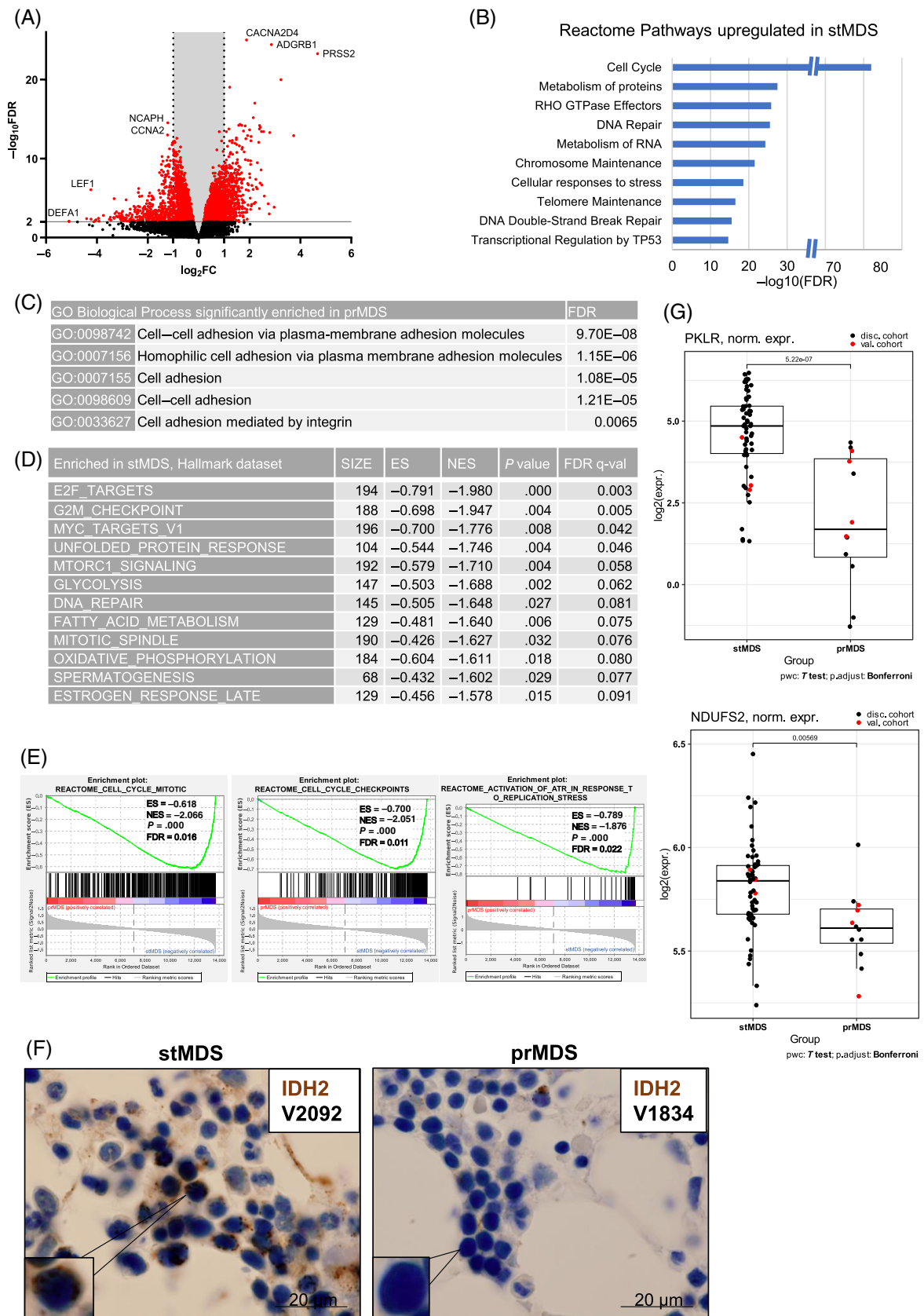


FIGURE 1 Legend on next page.

processes (BP) were enriched (not shown). Using REVIGO (settings: small) to reduce the list of GO terms, the list was shortened to 70 terms. The most enriched terms were associated with nucleosome assembly, regulation of gene expression, regulation of megakaryocyte differentiation, regulation of immune system processes, response to stress and stimulus, export and secretion from cells, cellular transport, telomere capping, signaling, DNA repair, cell adhesion, apoptosis and protein modification.

In REACTOME, 119 pathways were enriched in stMDS. The most significantly upregulated pathways were connected to the cell cycle, DNA repair, stress responses and chromosome and telomere maintenance (Figure 1B). Moreover, gene expression data showed dysregulation of critical signaling pathways mediated by NOTCH, RHO GTPases, β -catenin and WNT and the MAPK family. The pathways of regulation of rRNA expression, splicing and RNA metabolism were also dysregulated.

Next, we analyzed genes upregulated in prMDS. Although 328 PCGs were upregulated, multiprotein analysis in String showed only a few over-represented pathways. Five GO BP terms significantly enriched in prMDS were related to cell adhesion (Figure 1C). No pathway was upregulated according to the RACTOME database.

A set of validation samples (4 prMDS and 3 stMDS) showed the same pattern in PCA analysis (Figure S4A) and in DEA of *LEF1* and *LRP6* (Figure S4B).

3.2 | Downregulation of cell-cycle and energy metabolism related pathways in prMDS BM CD34+ cells

In GSEA, we ran the analysis through Curated Gene Sets (C2) and Hallmark Gene Sets (H); in stMDS, 485 and 12 gene sets were enriched at $FDR < 0.1$, respectively. We observed significant over-representation of processes related to cell cycle progression and energy metabolism in stMDS CD34+ cells (Figure 1D). Active cell cycling leads to replication stress, and ATR-dependent checkpoint activation is essential for cell-cycle progression to maintain the integrity of replication forks.¹² The gene sets for cell cycle, cell cycle checkpoints and activation of ATR in response to replication stress were all significantly overexpressed in stMDS and under-

represented in prMDS BM CD34+ cells (Figure 1E). These data, together with overall decreased metabolism signatures in prMDS CD34+ cells when compared to stMDS CD34+ cells (Figures 1D, F, G and S5), are consistent with slow cycling, more quiescent phenotype of prMDS CD34+ cells.¹³

3.3 | BM CD34+ cells of prMDS have attenuated DNA damage checkpoint and DDR pathways

DDR is closely interconnected with cell cycle progression and DNA damage checkpoints are activated to regulate various DNA-repair mechanisms; therefore, most DNA repair pathways are attenuated in non-dividing cells.¹⁴ We compared differential expression profiles of the DNA damage checkpoint and DNA repair pathways between stMDS and prMDS. GSEA analyses revealed significant under-representation for G2/M DNA damage checkpoint and DNA repair pathways in prMDS BM CD34+ cells compared to those of stMDS (Figures 2A and S6A). Then, we analyzed individual subcategories for single-strand break repair (SSR) and double-strand break repair (DSR) pathways; although stMDS expression patterns for these DNA repair gene sets were highly heterogeneous in the patients, prMDS patients often shared the same profiles of largely downregulated genes (Figures 2B and S6B–E). The downregulation of DNA repair genes in prMDS were primarily seen in mismatch repair/nucleotide excision repair (MMR/NER) genes and Fanconi anemia (FA) genes (Figures 2B and S6E). These data are consistent with the less proliferative, more quiescent-like cell state in prMDS, as MMR/NER SSR and FA DSR are primarily employed by proliferating HSCs for repair.¹³

We questioned whether the crucial regulatory networks of cell cycle progression, checkpoint activation and DNA repair signature in stMDS BM CD34+ cells significantly differ from those of healthy individuals. Thus, we performed DEA between healthy control ($n = 7$; median 45 years, range, 29–69 years; average 49 years) and selected age-matched stMDS (stMDS-age) ($n = 13$; median 53 years, range, 23–69 years; average 49 years) samples. We observed 70 up- and 209 downregulated genes in stMDS-age ($FDR < 0.05$). Only the splicing pathways were upregulated, and the immune pathways were downregulated (Figure S7A). In GSEA, no gene set for cycling,

FIGURE 1 Expression analysis of prMDS vs stMDS. (A) Volcano plot of differentially expressed genes (PCGs and lncRNAs) between CD34+ BM cells of prMDS and stMDS. The red points indicate significantly dysregulated genes. The most upregulated genes in prMDS (ie, downregulated in stMDS) are toward the right; the most downregulated genes in prMDS (ie, upregulated in stMDS) are toward the left. x-axis: logFC, logarithm of fold changes; y-axis: $-\log_{10}$ of FDR value; FDR, false discovery rate, red points: $FDR < 0.01$. (B) The most upregulated Reactome pathways in stMDS. (C) GO Biological Process terms significantly enriched ($FDR < 0.01$) in prMDS. (D) Gene sets enriched ($FDR < 0.1$) in stMDS from Hallmark dataset by GSEA. ES, enrichment score; NES, normalized enrichment score; P , P -value. (E) Selected pathways enriched ($FDR < 0.1$) in stMDS from C2 gene sets by GSEA. (F) IHC detection of IDH2 protein expression (brown) with hematoxylin nuclear counterstain (blue) in BM trephine biopsies of LR-MDS patients. Left: A section from stMDS patient no. V2092 with *IDH2* expression above the median showing nucleated cells with medium to strong mitochondrial staining for IDH2 (details seen in the inset). Right: Rare IHC staining for mitochondrial IDH2 from prMDS patients no. V1834 with *IDH2* expression below the median. *IDH2* transcription levels and detailed evaluation of IDH2 protein expression in individual samples are described in Figure S5B. (G) Differential expression of *PKLR* (left, a glycolytic marker) and *NDUFS2* (a core subunit of mitochondrial complex I) in stMDS vs prMDS samples. For boxplots generated from RNAseq data, we used the combined values of all samples, from the initial discovery cohort ($n = 61$, black dots) and the validation sample set ($n = 7$, red dots); P -value (shown in the graphs) is counted for all samples. Student's t test with Bonferroni correction.

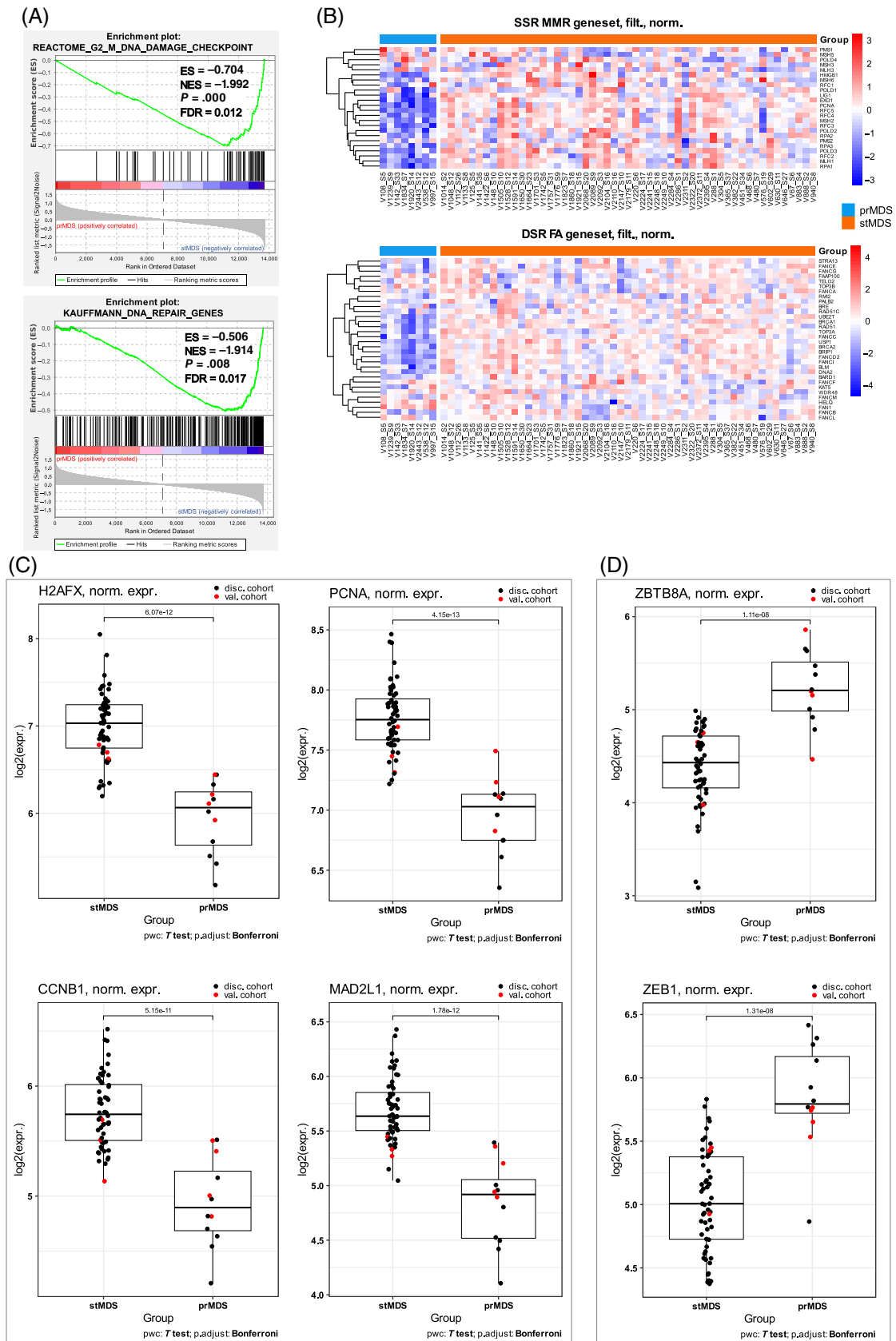


FIGURE 2 Legend on next page.

metabolism and DDR was significant (FDR < 0.1) in C2 and H datasets (Figure S7B). These data support the hypothesis that features that lead to development (and protect against progression) of LR-MDS manifest largely as a consequence of intrinsic molecular changes in cells that occur in processes of normal HSC ageing.^{15,16} The unavailability of BM CD34⁺ cells from age-matched healthy controls for prMDS precluded such comparison, but data from healthy controls and stMDS-age imply that premalignant expression signatures of prMDS are distinct from normal HSC ageing.

prMDS CD34⁺ cells revealed broadly repressed markers of multiple subcategories of DDR among the top 50 downregulated genes when compared to stMDS (Figure S8, Table S4A). *H2AFX* gene, encoding H2AX, a critical chromatin marker of DDR, was significantly downregulated in prMDS vs stMDS (Figure 2C). Among the significantly downregulated DDR genes in prMDS were also typical markers of DNA repair, such as *PCNA* and *AUNIP*, or mitotic phase-related genes, such as *CCNB1*, *MAD2L1*, *NCAPH* and *EXO1* (Figures 2C and S9). In addition, among significantly upregulated genes in prMDS compared to stMDS were genes encoding suppressors of DDR pathways, such as *ZBTB8A* and *ZEB1* (Figure 2D). *ZBTB8A* (*BOZF1*) expression is increased in several cancers where it functions as oncoprotein that stimulates progression through inhibiting p21 pathway.¹⁷ *ZEB1*, a known regulator of DDR in cancer,¹⁸ can downregulate p53 activity to suppress DDR.¹⁹ This led us to hypothesize that the mechanisms of attenuated DDR signaling in prMDS are associated with decreased cycling and other altered intrinsic pathways that are inherent to CD34⁺ cells and that these mechanisms could be also governed by aberrant splicing and epigenetic mediators, such as long noncoding RNAs (lncRNAs).

3.4 | Aberrant splicing and lncRNA expression are increased in prMDS BM CD34⁺ cells and target DDR pathways

To fully address the transcriptomic differences between stMDS and prMDS, we focused on aberrant splicing and differential expression of lncRNAs. Differential alternative splicing analysis revealed 1356 significant events (FDR < 0.05, absolute inclusion level difference >10%). PCA showed a great diversity in alternative splicing events (Figure 3A). We observed a higher number of alternative splicing events in prMDS (977 vs 378 events) (Figures 3B and S10) and found increased use of alternative 3' and 5' splice sites (A3SS, A5SS), but the

highest increase was seen in the retention of introns (RI) (Figures 3B and S10). In contrast, skipped exon (SE) events and usage of mutually exclusive exons (MXE) were decreased in prMDS compared to stMDS. Across all aberrantly spliced genes, several key hematopoietic regulators, such as *EZH2*, *MDM2*, *PDK1* and *DNMT3A*, were present. The top 10 alternatively spliced genes according to the absolute value of the inclusion level difference are listed in Table S5.

Genes aberrantly spliced in prMDS were part of pathways that were downregulated in this group according to DEA. These were metabolic processes including: nucleic acid, amino acid and protein metabolism; cellular response to DNA damage and stress; gene expression; histone modification; regulation of mRNA splicing; and protein transport (Figure 3C). Cellular pathways from GO BPs associated with aberrantly spliced genes in stMDS did not contain cellular responses to stress and responses to DNA damage (Figure 3D); these GO BPs were specifically targeted in prMDS. As aberrant splicing impairs the functionality of multiple cellular processes during carcinogenesis, including tumor-inhibitory pathways^{20,21}; these data are consistent with the proposed impaired tumor-suppressing mechanisms in prMDS.

Furthermore, a total of 629 lncRNAs were upregulated and 75 were downregulated (FDR < 0.05) in prMDS (Figure 3E). Anti-sense RNAs and lincRNAs were the most upregulated groups of lncRNAs, while processed pseudogenes were the most downregulated group. The expression of lncRNAs was heterogeneous in stMDS throughout the patients, but we observed a greater uniformity and mainly upregulation in lncRNA expression in prMDS (Figure S11). In accordance with previously described lncRNAs with prognostic significance in MDS,²²⁻²⁴ we observed upregulation of *MEG3* and *KIAA0125* and downregulation of *TCL6* in prMDS. Multiple other lncRNAs upregulated in prMDS compared to stMDS have a known function in cancer (Table S4B). Among them, *LINC00340* (*CASC15*) correlated positively with *ZEB1* (Figure 3F) and inversely with *CDKN1A* expression (Figure S12) in our LR-MDS patient cohort, in agreement with *LINC00340* proposed role in other cancer types.²⁵ *MALAT1* was published as the most highly expressed oncogenic lncRNA in multiple myeloma, where it directly binds to *PARP1*, a central protein in the base excision repair pathway²⁶; in our study, *MALAT1* correlated inversely with *PARP1* expression (Figure 3F). *ZEB1-AS1* showed moderate positive correlation with *ZEB1* (Figure S12, see also Table S4B for relevant references). Overall, these data suggest that upregulated global lncRNA expression in prMDS samples may contribute to the DDR pathway attenuation.

FIGURE 2 Attenuated DNA damage checkpoints and DDR pathways in prMDS CD34⁺ cells. (A) GSEA identified G2/M DNA damage checkpoint and DNA repair gene sets as significantly enriched in stMDS compared to prMDS. ES, enrichment score; NES, normalized enrichment score; *P*, *P*-value; FDR, false discovery rate. (B) Heatmap representations of the differential expression of genes in selected DNA repair gene categories: single-strand break mismatch repair (SSR MMR), double-strand break repair—Fanconi anemia pathway (DSR FA). Red indicates upregulation, blue indicates downregulation of gene expression, and the color intensity indicates the level of differential expression. Columns in the heatmaps represent individual samples. (C) Boxplots depicting expression of four genes of the 19-gene DDR signature in stMDS and prMDS cohorts. (D) Boxplots depicting expression of *ZBTB8A* and *ZEB1* in stMDS and prMDS. For boxplots generated from RNAseq data, we used the combined values of all samples, from the initial discovery cohort (*n* = 61, black dots) and the validation sample set (*n* = 7, red dots); *P*-values (shown in the graphs) were counted for all samples. Student's *t* test with Bonferroni correction.

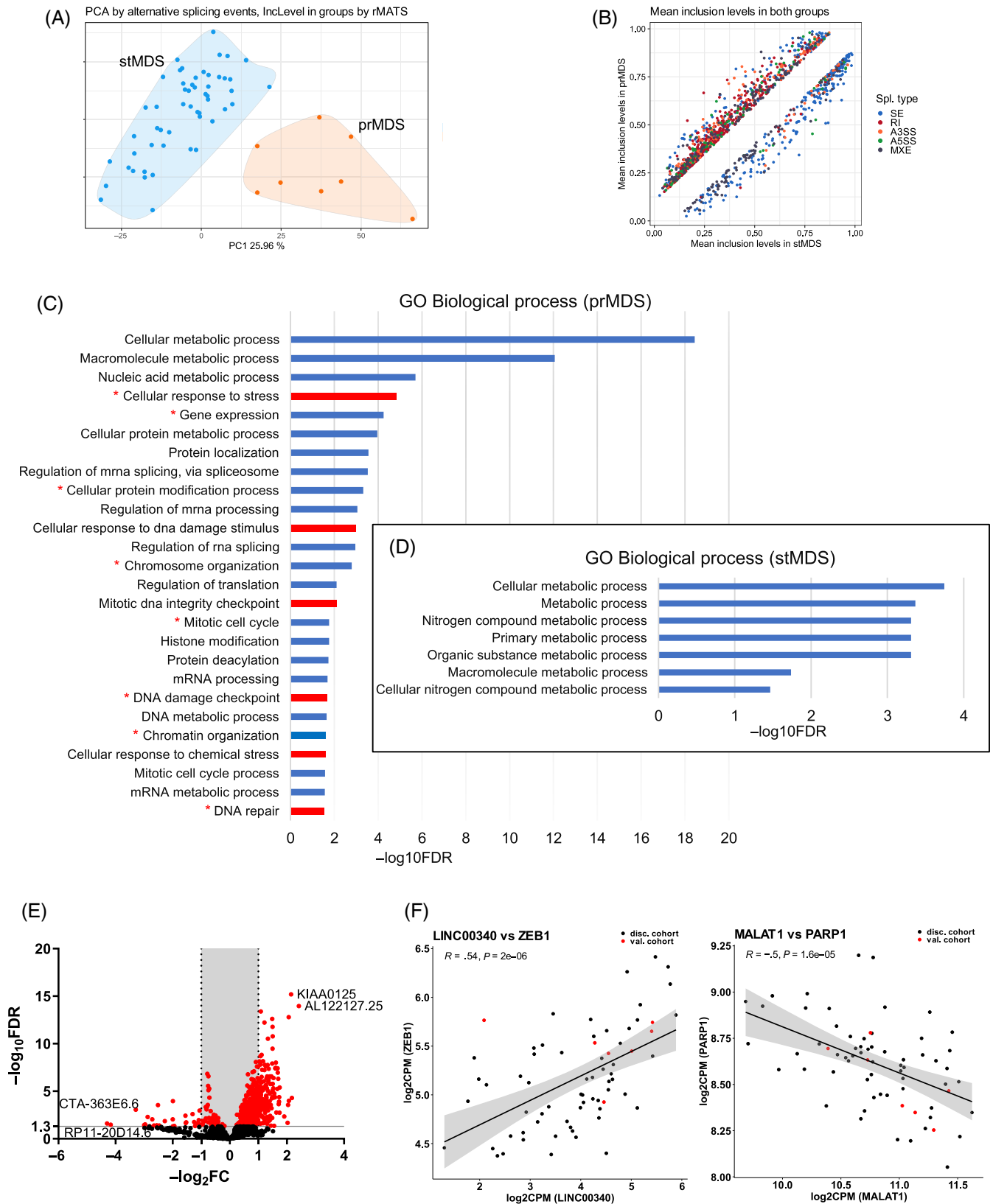


FIGURE 3 Legend on next page.

3.5 | Gene expression signatures reveal overexpression of multiple oncogenic PCGs and diminished differentiation in prMDS BM CD34+ cells

The heatmap of the top 50 upregulated genes in prMDS vs stMDS samples, ranked by GSEA (Figure S8), included several PCGs, previously identified as markers of leukemic progression and/or poor MDS/AML survival, such as *SPNS2*, *TEC*, *MN1* and others (Figure 4A, see Table S4C for the list of these PCGs and relevant references). One of the genes, *NEK3* (Figure 4A, see also Figure S4B for *NEK3* in validation sample set), has not previously been linked to MDS/AML progression, but its overexpression is associated with poor prognosis in patients with gastric cancer and the encoded protein plays an important role in cell migration, cell proliferation and cell viability.²⁷ Other PCGs, such as *RPAP2* and *LRP6* belong to the key molecules of AML LSCs.^{11,28}

We further investigated the commitment to specific hematopoietic lineages as previously described,⁴ where a decrease in the expression of genes specifically expressed in erythroblasts and megakaryocyte/erythrocyte progenitors and upregulation of genes related to immature progenitor cells was associated with shorter survival and leukemic transformation in myelodysplasia samples. Indeed, prMDS had significantly reduced expression of genes of more mature hematopoietic lineages (Figures 4B and S13, Table S6).

These data collectively suggest reduction in markers for mature hematopoietic lineages and expression of LSC phenotype in BM CD34+ cells of prMDS vs stMDS.

3.6 | ZEB1 is significantly overexpressed in prMDS BM CD34+ cells and anticorrelates with CDH1 expression

Next, we addressed the overrepresented GO terms associated with cell–matrix or cell–cell adhesion in prMDS (Figure 1C). A key factor which controls alterations in cell–cell adhesion and its connection with DDR in tumorigenesis is *ZEB1*.¹⁸ As mentioned, we observed significant upregulation of *ZEB1* in prMDS compared to stMDS (Figure 2D, see also Figure S4B for *ZEB1* in validation sample set), which is consistent with a previous study that showed *ZEB1* expression significantly associated with poor

overall survival of AML.²⁹ We examined whether differences in *ZEB1* expression in LR-MDS BM CD34+ cells are reflected at its protein level. Immunohistochemistry (IHC) staining indicated correlation of CD34+ cell transcripts of *ZEB1* and *ZEB1* nuclear protein levels in LR-MDS BM CD34+ cells in the most examined samples (Figure 4C, Table S7). In one biopsy from prMDS patient with *ZEB1* transcript level above median, *ZEB1* immunoreactivity was absent in CD34+ cells but was positive in multiple nucleated cells of granulocytic and monocytic lineage (Figure S14). This, together with above mentioned coregulations of *ZEB1* with some lncRNAs underscores a complex regulation of *ZEB1* expression in LR-MDS.

ZEB1 plays a pivotal role in transcriptional repression of *CDH1*, the gene encoding E-cadherin³⁰ and we observed significantly downregulated *CDH1* in prMDS (Figure 4D). There was a significant moderate negative correlation of expression levels of *ZEB1* and *CDH1*, suggesting their functional relationship in LR-MDS (Figure 4D). Another PCG associated with cell adhesion, migration and invasiveness of cancer cells including AML, *SPNS2*,³¹ was significantly upregulated in prMDS (Figure 4A, see also Figure S4B for *SPNS2* in validation sample set); in fact, it was the most upregulated PCG among the top 50 upregulated genes in the prMDS vs stMDS samples ranked by GSEA (Figure S8, Table S4C). These data confirmed dysregulated cell adhesion gene expression in prMDS.

3.7 | DDR gene expression signature for LR-MDS patient stratification

To propose a gene expression signature predictive of progression of LR-MDS patients, we first identified two groups of genes from the expression profiles of the top 50 up- and downregulated genes in GSEA (Figure S8). From the top 50 upregulated genes in prMDS, we selected genes with known function in leukemic progression and/or poor MDS/AML survival (13 genes, Table S4C) and tested them independently for predictive power on progression (Supplementary Methods in Data S1). The cumulative incidence of MDS progression of the 61 patients (discovery cohort) revealed significant curve separations ($P < .05$) for 11 genes with the best

FIGURE 3 Alternative splicing events and lncRNA expression in LR-MDS CD34+ cells. (A) PCA based on the inclusion levels of the differential splicing events separated the stMDS and prMDS samples. (B) Visualization of alternative splicing events in prMDS and stMDS. The higher number of alternative splicing events was observed in prMDS. The highest difference was detected in the retention of introns (RI, red dots). SE, skipped exon; RI, retention of introns; A3SS, A5SS, alternative 3' and 5' splice sites; MXE, mutually exclusive exons. (C,D) Cellular pathways from GO Biological Processes associated with aberrantly spliced genes in prMDS (C) and in stMDS (D). GO terms related to cellular responses to stress and responses to DNA damage are depicted in red color. The red asterisks indicate GO terms significantly downregulated in prMDS according to DEA. FDR, false discovery rate. (E) Volcano plot of differentially expressed lncRNAs between CD34+ BM cells of prMDS and stMDS. The red points indicate significantly dysregulated genes. The most upregulated lncRNAs in prMDS (ie, downregulated in stMDS) are toward the right; the most downregulated genes in prMDS (ie, upregulated in stMDS) are toward the left. x-axis: $-\log_{10}FC$, logarithm of foldchanges; y-axis: $-\log_{10}$ of FDR value; FDR, false discovery rate, red points: FDR < 0.05. (F) Correlation curves depicting a significant moderate positive correlation of lncRNA *LINC00340* and *ZEB1* expression levels and a significant moderate negative correlation of lncRNA *MALAT1* and *PARP1*. Pearson correlation analysis combined values of both cohorts, from the initial discovery cohort ($n = 61$, black dots) and the validation sample set ($n = 7$, red dots). CPM, counts per million; P , P -value; R , correlation coefficient.

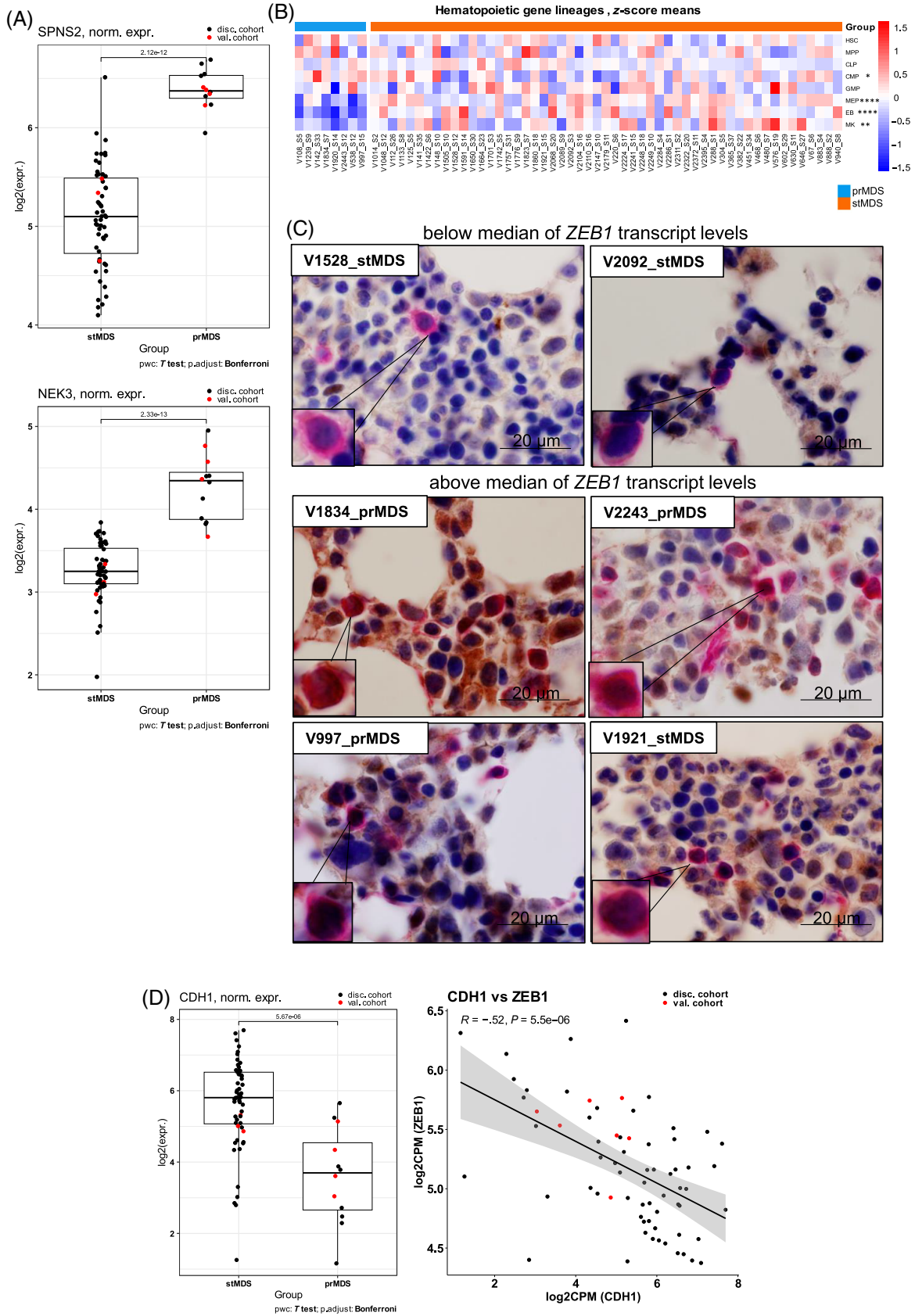


FIGURE 4 Legend on next page.

P-values for *RPAP2*, *MN1* and *DOCK1* (Figure 5A). We also tested other oncogenes among the top 50 upregulated genes in prMDS not previously associated with MDS/AML for their potential biomarker role in predicting LR-MDS progression. The best significance of curve separation was calculated for *NEK3* gene, encoding for NIMA-related kinase 3 (Figure 5B). Expression of *NEK3* was also an independent prognostic factor ($P < .001$) for progression-free disease in multivariate analysis of important clinical and genetic factors (Table S8).

As described above, downregulation of DDR genes expression was the main characteristic feature distinguishing prMDS from stMDS in our cohort. Therefore, we chose 19 DDR genes (Table S4A) from the top 50 downregulated genes in prMDS (Figure S8) and tested them individually for the cumulative incidence of progression; significant curve separation ($P < .05$) was achieved for 17 genes (Figure S15). The entire 19 DDR gene set contained mainly cell cycle-related genes, reflecting non-cycling status of prMDS BM CD34+ cells, not known to be mutated in MDS and AML. To assess whether an expression-based DDR signature could predict LR-MDS patients' progression, we calculated probability of a progression-free disease (with mortality before MDS progression as a competing risk, using regularized Cox regression analysis as described in Supplementary Methods in Data S1) based on expression of these 19 DDR-associated genes. Indeed, our analysis revealed that LR-MDS patients with high DDR gene expression had a significantly higher probability of progression-free disease ($P < .001$) compared to those with low DDR gene expression (Figure 5C). Thus, DDR gene expression signature appears to be a significant predictor of LR-MDS progression in our cohort.

3.8 | Correlations of gene expression profiles suggest a key role for ZEB1 in LR-MDS transcriptional regulatory network

We asked whether ZEB1, a known master regulator (repressor or transcriptional activator, depending on cell context³² of key processes

dysregulated in prMDS, ie, DDR, EMT and quiescence^{18,33,34}) might act as a transcriptional regulator in LR-MDS. We searched for overlap between the ZEB1 target genes in the ChIP-seq datasets (ENCODE transcription factor dataset³⁵) and both, DDR signature genes downregulated in prMDS (Table S4A) and genes upregulated in prMDS with known roles in leukemic progression and/or poor MDS/AML survival (Table S4C). Among the 19 DDR-associated genes selected for the "DDR signature," 14 genes are listed as putative ZEB1 targets. All were found to be significantly anticorrelated with ZEB1 expression in our LR-MDS samples (anticorrelations with *H2AFX* and *H2AZ1* are shown in Figure S16). Four of these genes, *CCNB1*, *MAD2L1*, *NCAPH* and *EXO1*, which are related to mitotic phase and whose Pearson correlation coefficient (*R*) exceeded $-.64$ (Figure 5D), have been previously described to have ZEB1 binding sites in their putative promoter sequences in breast cancer cells, but in which ZEB1 serves as a transcriptional activator of these genes.³⁶

Among the genes predicting poor prognosis in MDS/AML that were positively correlated with ZEB1 expression was *ANGPT1*, encoding angiotensin-converting enzyme 1 (Figure 5E). *ANGPT1* was shown to be a direct target positively regulated by ZEB1 in triple-negative breast cancer cells,³⁷ but its potential transactivation by ZEB1 in MDS or AML progenitors has not been investigated. We also observed significant positive correlation of expression between ZEB1 and *LRP6* (Figure 5E). *LRP6*, encoding the coreceptor of Wnt/ β -catenin signaling,¹¹ is a ZEB1 target according to publicly available (ChIP)-seq datasets from the ENCODE project.

According to the ENCODE project datasets, *NEK3* is also among the ZEB1 target genes. Indeed, a significant moderate positive correlation between ZEB1 and *NEK3* was observed in LR-MDS CD34+ cells (Figure S16), but such a possible relationship in LR-MDS progenitors requires further functional analyses.

4 | DISCUSSION

MDS is a very heterogeneous disease. So far, several transcriptional studies have been performed to understand the

FIGURE 4 Differences in cellular states of CD34+ cells from prMDS and stMDS, defined by transcriptional programs. (A) Differential expression of *SPNS2* and *NEK3* in stMDS vs prMDS samples. For boxplots generated from RNAseq data, we used the combined values of all samples, from the initial discovery cohort ($n = 61$, black dots) and the validation sample set ($n = 7$, red dots); *P*-value (shown in the graphs) is counted for all samples. Student's *t* test with Bonferroni correction. (B) Heatmap of *z* score means from expression values of genes related to hematopoietic lineages. The asterisks flag indicates the level of significance from the *t* test, the values are listed in Table S6, and a heatmap for all related genes is provided in Figure S13. CLP, common lymphoid progenitor; CMP, common myeloid progenitor; EB, erythroblast; GMP, granulocyte monocyte progenitor; HSC, hematopoietic stem cell; MEP, megakaryocyte erythrocyte progenitor; MK, megakaryocyte; MPP, multipotent progenitor. (C) Representative IHC double staining of CD34/ZEB1 protein in BM trephine biopsies of LR-MDS patients. Upper panels: Individual sections from stMDS patients no. V1528 and V2092 with ZEB1 transcript levels below the median (see also Table S7) demonstrated membranous/cytoplasmic staining for CD34 (red) with absent/rare nuclear/cytoplasmic ZEB1 staining (brown, see the insets). Bottom panels: Specific membranous and diffuse cytoplasmic CD34 staining (red) with moderate to strong ZEB1 protein co-expression (brown, see the insets) in individual sections from prMDS patients no. V1834, V2243 and V997 and stMDS patient no. V1921, all with ZEB1 transcript levels above the median. Multiple other nucleated, non-CD34+ cells with hematoxylin nuclear counterstains (blue), showed both nuclear and/or cytoplasmic staining for ZEB1. ZEB1 transcript levels and detailed evaluation of ZEB1 protein expression in individual samples are described in Table S7. (D) Differential expression of *CDH1* gene in stMDS and prMDS patients and Pearson correlation analysis of the expression levels of *CDH1* and its repressor ZEB1. For boxplots and correlation analysis, we combined values of both cohorts, from the initial discovery cohort ($n = 61$, black dots) and the validation sample set ($n = 7$, red dots). Student's *t* test with Bonferroni correction, *P*-values shown in the graphs; CPM, counts per million; *R*, correlation coefficient.

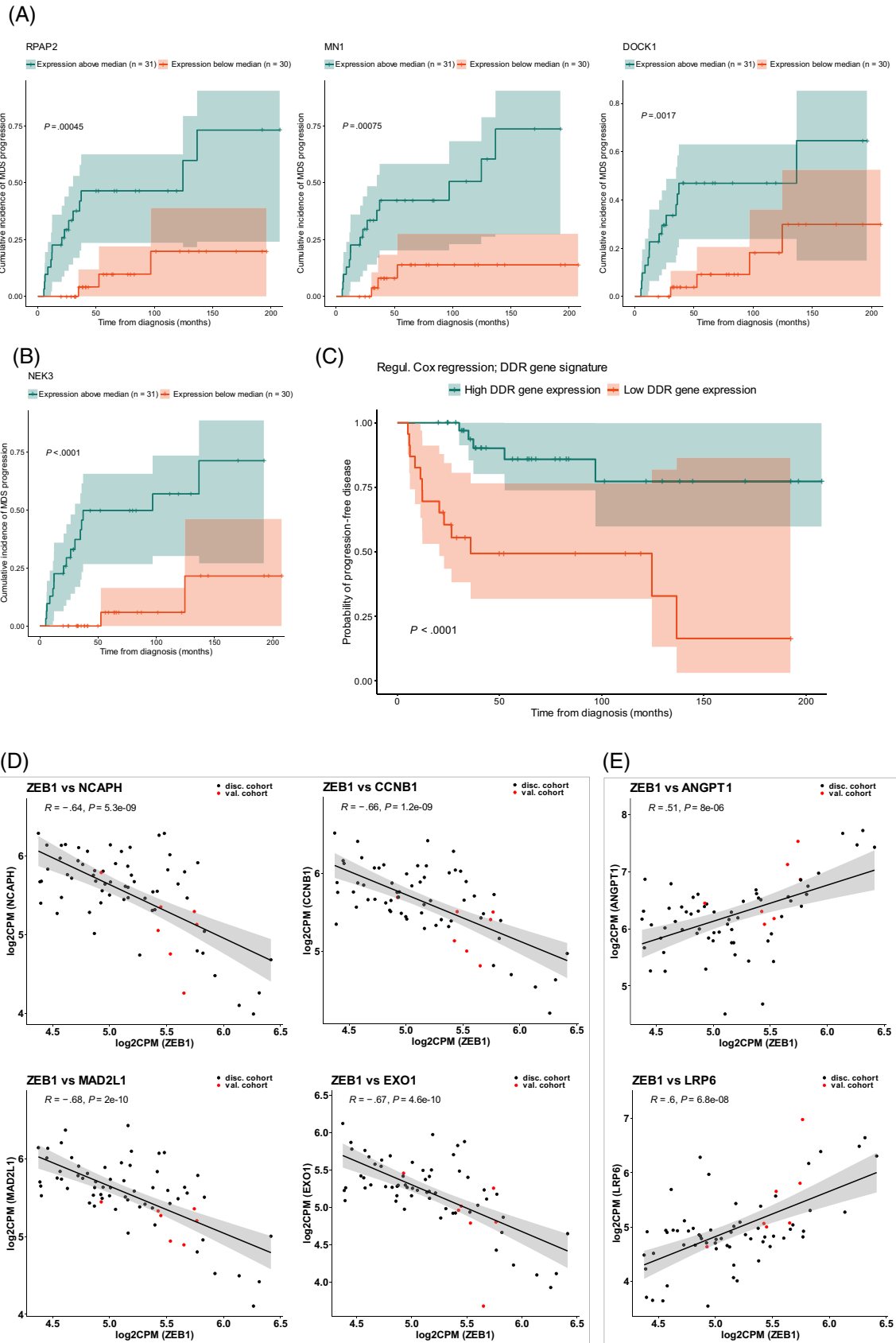


FIGURE 5 Legend on next page.

pathogenesis of MDS and predict the outcome.^{4,5} However, to our knowledge, there is no transcriptomic study targeting exclusively only LR-MDS to study the slight but important differences associated with accelerated progression. Here, we provide a comprehensive transcriptome analysis of LR-MDS samples and identify the differences between patients with stable disease and those at risk of rapid progression. We have previously demonstrated that *RUNX1*-mutated LR-MDS patients are at risk of rapid progression and that *RUNX1* mutations presumably disturb anti-tumor cellular defense pathways and contribute to disease progression.⁶ The work presented here builds on the previous study and examines the differences regardless of driver mutations.

For the purpose of this study, we stratified our cohort of patients into stMDS and prMDS groups; the cut-off point for accelerated progression and inclusion in the prMDS group was set to 20 months from diagnosis. Our initial analyses (transcriptome consistency between prMDS samples, PCA, overlap with recently published markers of WNT signaling) suggested that the division of LR-MDS patients based on such a threshold of accelerated progression might be associated with distinct gene expression signatures of BM CD34+ cells. Our initial discovery cohort consisted of 61 LR-MDS patients, the validation sample set included 7 samples; a higher number of analyzed patients was precluded by unavailability of additional diagnostic CD34+ samples.

In the discovery cohort, we found 697 significantly dysregulated PCGs between stMDS and prMDS. These results show a large heterogeneity within LR-MDS patients, even though they represent one group for treatment decisions, and are consistent with the previously published heterogeneity of progenitor subpopulations within LR-MDS BM CD34+ cells.³⁸ In stMDS, we found transcriptional signatures connected to actively cycling CD34+ cells poised for erythro-megakaryocytic differentiation, with high metabolic activity and with the intrinsic ability to activate stress defense pathways. In contrast, only GO BPs for cell adhesion pathways were upregulated in prMDS. In addition, we found that several oncogenes known to be involved in leukemia progression and/or poor survival in MDS/AML are significantly upregulated in prMDS BM CD34+ cells. Interestingly, the *NEK3* gene, not previously associated with MDS/AML, revealed the best prognostic value for predicting LR-MDS progression. *NEK3* is a member of the serine/threonine NEK kinase family and its overexpression is

associated with progression and poor prognosis in patients with breast cancer and gastric cancer,²⁷ but its role in myeloid leukemogenesis remains to be clarified.

The most characteristic transcriptomic signature of prMDS CD34+ cells distinguishing them from stMDS cells was the expression pattern of quiescent-like or slowly proliferating progenitors with LSC markers, and the signatures of processes that are associated with or regulate non-cycling malignant stem cells: that is, suppressed DDR and dysregulated adhesion pathways.³⁹ It is likely that the aforementioned changes in the cellular composition of CD34+ subpopulations between prMDS vs stMDS CD34+ cells (ie, their differentiation status) determine to some extent the differences in gene expression signatures.³⁸ We propose that prMDS CD34+ cells bear cell-autonomous and perhaps also noncell-autonomous defects to establish a tumor-suppressing DDR barrier; these defects could be due to both genetic and epigenetic changes,⁴⁰ including the interplay with high levels of aberrant splicing.^{41–43} Nevertheless, none of the 22 DDR genes selected for sequence analysis were mutated in prMDS; the four mutant *TP53* samples were among the stMDS cases. Abnormal splicing in MDS is mainly related to the presence of mutated splicing factors,^{41,42,44} and there are limited data regarding their clinical implications. In a study by Yang et al,⁴⁵ almost 30% of genes were aberrantly spliced in MDS samples compared to healthy controls. The aberrantly spliced genes were related to cell proliferation, cell adhesion and protein degradation, and a higher degree of aberrantly spliced genes correlated with shorter OS and time to leukemic transformation. We revealed that stMDS and prMDS significantly differ in their alternatively spliced genes; genes abnormally spliced in prMDS were enriched in specific GO terms, involving DDR-related pathways. Several key hematopoietic regulators were also aberrantly spliced. The highest difference between stMDS and prMDS was seen in the retention of introns, which was significantly higher in prMDS. RI, frequent in multiple cancers (including MDS), has been described as a mechanism of tumor suppressor inactivation.^{21,43}

Furthermore, we observed an overall upregulation of lncRNA expression in prMDS CD34+ cells. A published lncRNA-based risk-scoring system for MDS patients, independent of the patients' mutational profiles, demonstrated the potential of lncRNA profiling to improve current MDS risk stratification.⁴⁶ In our study, the

FIGURE 5 Dysregulated expression of DDR genes and genes previously associated with myeloid malignancies is significantly associated with accelerated progression in LR-MDS. (A) Cumulative incidence of LR-MDS progression with death as a competing risk according to the expression of three genes with a known function in leukemia stem cell function/leukemic progression and/or poor MDS/AML survival (Table S4C) from the top 50 upregulated genes in prMDS (Figure S8). (B) Cumulative incidence of LR-MDS progression with death as a competing risk according to the *NEK3* expression. (C) Probability of a progression-free disease with mortality prior to MDS progression as a competing risk based on DDR gene expression signature. DDR gene expression signature appears to be a significant predictor of LR-MDS progression. The signature contains expression levels of 19 DDR genes (Table S4A) from the top 50 downregulated genes in prMDS (Figure S8). (D) *ZEB1* transcriptional correlations with its putative targets, using a combined set of patients from the discovery cohort ($n = 61$, black dots) and validation samples ($n = 7$, red dots). Pearson correlation analysis revealed a strong negative correlation of *ZEB1* with the expression of four DDR genes *NCAPH*, *CNNB1*, *MAD2L1* and *EXO1* and (E) a strong positive correlation with the expression of *ANGPT1* and *LRP6* genes. CPM, counts per million; *P*, *P*-value; *R*, correlation coefficient.

most significantly upregulated lncRNA in prMDS was *KIAA0125*; its higher expression has previously been associated with concordant upregulation of HSC-associated genes at higher-risk MDS and has emerged as an independent unfavorable prognostic marker for OS and LFS in MDS.²² Indeed, lncRNAs belong to important epigenetic regulators of tumor-promoting and tumor-restraining pathways and contribute to the regulation of expression of multiple oncogenes and tumor-suppressor genes.⁴⁷ Because epigenetic silencing of DDR genes and pathways is a common cause of DDR deficiency in cancer,³⁸ we correlated two of the significantly overexpressed oncogenic lncRNAs in prMDS, *LINC00340* (*CASC15*) and *MALAT1*, with their potentially interacting components of DDR pathway (*ZEB1* and *CDKN1A* in the case of *LINC00340*, and *PARP1* in the case of *MALAT1*), and confirmed earlier findings.^{25,26} These data suggest that lncRNA expression in prMDS may trigger silencing of some DDR factors, thus likely contributing to the overall suppression of DDR expression in prMDS.

To further explain the mechanisms of disease progression in prMDS with attenuated DDR pathways, we addressed the possible impact of overrepresented GO terms associated with cell-matrix or cell-cell adhesion in prMDS. First, we focused on *ZEB1* gene expression. *ZEB1* controls alterations in cell-cell adhesion and E-cadherin expression; thus, it is a key factor of the first steps in epithelial-mesenchymal transition (EMT). Moreover, its activation prevents the induction of DDR (and thus the tumorigenesis barrier) in the early steps of tumorigenesis.¹⁸ Indeed, we observed a significant upregulation of *ZEB1* in prMDS, consistent with previously published *ZEB1* expression significantly associated with poor overall survival of AML.²⁹ *ZEB1* is a transcriptional repressor of *CDH1*, the gene encoding E-cadherin,³⁰ and we observed significantly downregulated *CDH1* in prMDS. Disruption of E-cadherin mediated cell-cell adhesion is frequently observed in hematological malignancies.^{48,49} An earlier report on CD34+ cell populations from MDS patients has shown that the promoter of *CDH1* is often hypermethylated, resulting in decreased protein expression.⁵⁰ Earlier analyses also interconnected overexpression of *ZEB1* with downregulation of H2AX; H2AX silencing in HCT116 human colon cancer cells promoted mesenchymal-like characteristics including upregulation of EMT transcription factors, including *ZEB1*.⁵¹ Expression of histone H2A variants, participants in DDR pathways, was largely suppressed in prMDS, and significantly anticorrelated with *ZEB1* expression. In addition, an association between the downregulation of DDR repair genes and the dysregulated expression of EMT transcription factors, including *ZEB1*, potentially counteracting the mesenchymal to epithelial transition, has been described during an early phase of reprogramming mouse embryonic fibroblasts.⁵² All these data collectively suggest a close coregulation between suppressed DDR gene expression with a transcriptional program of cell-cell adhesion and mesenchymal cell fate promoting leukemogenesis/progression in prMDS.

The attenuated DDR, alteration of cell adhesion and relative quiescence of prMDS CD34+ cells led us to hypothesize, that *ZEB1* might be an important transcriptional regulator in LR-MDS.

This could be supported by *ZEB1* IHC detection in the patients' biopsies and the high transcriptional association of *ZEB1* with multiple DDR, cell adhesion/migration and cell cycle progression genes, putative *ZEB1* targets proposed in the literature and/or biomedical databases. Indeed, these literature/database data were obtained from human cell types other than CD34+ HSPCs (MDS) and in different cell contexts,^{25,30,32-36,51,52} and the expression data shown here represent only basis for further functional studies. Another regulatory role of *ZEB1* in LR-MDS progression could be its proposed stimulation of *AGPN1*, encoding the angiogenic factor angiopoietin-1. Such coregulation may have functional implications similar to the *ZEB1* role in tumor stromal endothelial cells.⁵³ The expression of the LSC phenotype in prMDS, characterized by *LRP6* gene,¹¹ could result from Wnt/ β -catenin/*ZEB1* coregulation. Wnt/ β -catenin signaling is known to directly induce *ZEB1* transcription,⁵⁴ and *Lrp6* is a downstream target of *ZEB1* in endothelial progenitors.³³ Indeed, our observed coregulation between *ZEB1* and *LRP6* expression in LR-MDS CD34+ cells should be corroborated by further studies.

In conclusion, our data show transcriptional signatures of active cycling and functional DDR signaling in stMDS CD34+ cells, which corresponds to an enhanced intrinsic ability of stMDS CD34+ cells to suppress malignant progression through antitumor barrier activation. This is consistent with current understanding that proliferation in pre-malignant disease states generates replicative stress and upregulation of DNA repair processes, which evokes checkpoint activation, participating in an anti-transformation barrier and thus preventing pathways leading to tumor development.¹² The accelerated progression phenotype of prMDS appears to be a consequence of the downregulation of DNA damage checkpoints/repair, and hence allows grounds for increased genomic instability. This cell-autonomously and perhaps also noncell-autonomously mediated suppression of DDR barrier seems to be interconnected with dedifferentiation and a quiescent-like gene expression program which promotes MDS progression.

Our findings also point to appropriate therapeutic approaches for prMDS patients. We propose that specific therapeutic approaches can be exploited to selectively target malignant progenitors in stMDS and prMDS; we present a DDR gene expression signature with potential implication in prognostic stratification of LR-MDS patients, independent of the mutational status. Previous studies have documented that the quiescent cell state and a transcriptional program resembling that observed for prMDS CD34+ cells represent features of pre-existing failure to respond to azacitidine monotherapy.^{13,55} On the other hand, regimens combining azacitidine with proliferation/myeloid differentiation-inducing growth factors⁵⁶ could represent candidate treatment options for prMDS.

AUTHOR CONTRIBUTIONS

Monika Kaisrlikova: Processed the samples and performed the experiments; Processed and interpreted the data; Drafted the article. **David Kundrat:** Performed the bioinformatic analyses. **Pavla Koralkova:** Processed the samples and performed the experiments; Interpreted the data; Reviewed the article. **Iva Trsova:** Processed the samples and performed the experiments. **Zuzana Lenertova:** Processed the

samples and performed the experiments. **Hana Votavova**: Reviewed the article. **Michaela Dostalova Merkerova**: Processed and interpreted the data. **Zdenek Krejcek**: Processed the samples and performed the experiments. **Jitka Vesela**: Processed the samples and performed the experiments. **Martin Vostry**: Processed the samples and performed the experiments. **Radka Simeckova**: Processed the samples and performed the experiments. **Marketa Stastna Markova**: Responsible for patient selection. **Marie Lauermannova**: Responsible for patient selection. **Anna Jonasova**: Responsible for patient selection. **Jaroslav Cermak**: Responsible for patient selection. **Vladimir Divoky**: Processed and interpreted the data; Drafted the article. **Monika Belickova**: Arranged funding and supervised the project; Reviewed the article. All authors contributed to the article and approved the final version. The work reported in the article has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

The authors would like to thank Sarka Ransdorfova (Institute of Hematology and Blood Transfusion, Prague, Czech Republic), Zuzana Zemanova (General University Hospital, Prague, Czech Republic) for providing cytogenetic data and Patrik Flodr (University Hospital Olomouc, Czech Republic) for evaluation of ZEB1 IHC staining. Graphical abstract was created in BioRender.

FUNDING INFORMATION

This work was supported by Ministry of Health of the Czech Republic (NU21-03-00565), MH CZ-DRO (UHK, 00023736), Charles University UNCE/MED/016 and BBMRI.CZ, reg. no. LM2023033. PK was supported by UP Young Researcher Grant Competition, ID Project JG_2023_016. VD and PK were supported by the European Union - Next Generation EU, Programme EXCELES, ID Project No. LX22NPO5102.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing financial interests.

DATA AVAILABILITY STATEMENT

The raw RNAseq data generated in this study were deposited in the National Centre for Biotechnology Information BioProject (accession number for discovery cohort: PRJNA955117 and for validation sample set: PRJNA1023700). Other data that support the findings of this study are available from the corresponding author upon request.

ETHICS STATEMENT

All patients provided their signed informed consent, the study was approved by the Institutional Scientific Board and the IHBT Ethics Committee (EK 4/AZV CR/06/2017) and the study was conducted in accordance with the ethical standards of the Declaration of Helsinki.

ORCID

Monika Kaisrlikova  <https://orcid.org/0000-0001-8840-7654>

David Kundrat  <https://orcid.org/0000-0003-4770-1229>

Pavla Koralkova  <https://orcid.org/0000-0001-7970-1501>

Iva Trsova  <https://orcid.org/0000-0002-6594-776X>

Zuzana Lenertova  <https://orcid.org/0009-0003-5745-8856>

Hana Votavova  <https://orcid.org/0000-0002-2990-579X>

Michaela Dostalova Merkerova  <https://orcid.org/0000-0002-6345-9180>

Zdenek Krejcek  <https://orcid.org/0000-0002-8321-2644>

Jitka Vesela  <https://orcid.org/0000-0002-1585-6763>

Martin Vostry  <https://orcid.org/0000-0002-7557-2543>

Marketa Stastna Markova  <https://orcid.org/0000-0003-2206-5246>

Anna Jonasova  <https://orcid.org/0000-0003-0246-1524>

Vladimir Divoky  <https://orcid.org/0000-0003-0202-245X>

Monika Belickova  <https://orcid.org/0000-0002-9158-881X>

REFERENCES

- DeZern AE. Lower risk but high risk. *Hematology Am Soc Hematol Educ Program*. 2021;2021(1):428-434. doi:10.1182/HEMATOLOGY.2021000277
- Mills KI, Kohlmann A, Williams PM, et al. Microarray-based classifiers and prognosis models identify subgroups with distinct clinical outcomes and high risk of AML transformation of myelodysplastic syndrome. *Blood*. 2009;114(5):1063-1072. doi:10.1182/BLOOD-2008-10-187203
- Pellagatti A, Benner A, Mills KI, et al. Identification of gene expression-based prognostic markers in the hematopoietic stem cells of patients with myelodysplastic syndromes. *J Clin Oncol*. 2013;31(28):3557-3564. doi:10.1200/JCO.2012.45.5626
- Shiozawa Y, Malcovati L, Galli A, et al. Gene expression and risk of leukemic transformation in myelodysplasia. *Blood*. 2017;130(24):2642-2653. doi:10.1182/BLOOD-2017-05-783050
- Im H, Rao V, Sridhar K, et al. Distinct transcriptomic and exomic abnormalities within myelodysplastic syndrome marrow cells. *Leuk Lymphoma*. 2018;59(12):2952-2962. doi:10.1080/10428194.2018.1452210
- Kaisrlikova M, Vesela J, Kundrat D, et al. RUNX1 mutations contribute to the progression of MDS due to disruption of antitumor cellular defense: a study on patients with lower-risk MDS. *Leukemia*. 2022;36(7):1898-1906. doi:10.1038/s41375-022-01584-3
- Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood*. 2012;120(12):2454-2465. doi:10.1182/blood-2012-03-420489
- Cheson BD, Greenberg PL, Bennett JM, et al. Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. *Blood*. 2006;108(2):419-425. doi:10.1182/BLOOD-2005-10-4149
- Supek F, Bošnjak M, Škunca N, Šmuc T. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One*. 2011;6(7):e21800. doi:10.1371/journal.pone.0021800
- Pellagatti A, Marafioti T, Paterson JC, et al. Marked downregulation of the granulopoiesis regulator LEF1 is associated with disease progression in the myelodysplastic syndromes. *Br J Haematol*. 2009;146(1):86-90. doi:10.1111/j.1365-2141.2009.07720.x
- Sakoda T, Kikushige Y, Miyamoto T, et al. TIM-3 signaling hijacks the canonical Wnt/ β -catenin pathway to maintain cancer stemness in acute myeloid leukemia. *Blood Adv*. 2023;7(10):2053-2065. doi:10.1182/bloodadvances.2022008405
- Takacova S, Slany R, Bartkova J, et al. DNA damage response and inflammatory signaling limit the MLL-ENL-induced leukemogenesis in vivo. *Cancer Cell*. 2012;21(4):517-531. doi:10.1016/J.CCR.2012.01.021

13. Koralkova P, Belickova M, Kundrat D, et al. Low plasma citrate levels and specific transcriptional signatures associated with quiescence of CD34+ progenitors predict azacitidine therapy failure in MDS/AML patients. *Cancers (Basel)*. 2021;13(9):2161. doi:10.3390/cancers13092161
14. Brnzei D, Foiani M. Regulation of DNA repair throughout the cell cycle. *Nat Rev Mol Cell Biol*. 2008;9(4):297-308. doi:10.1038/nrm2351
15. De Haan G, Lazare SS. Aging of hematopoietic stem cells. *Blood*. 2018;131(5):479-487. doi:10.1182/BLOOD-2017-06-746412
16. Heibl S, Stauder R, Pfeilstöcker M. Is myelodysplasia a consequence of normal aging? *Curr Oncol Rep*. 2021;23(12):1-11. doi:10.1007/s11912-021-01136-5
17. Kim MK, Jeon BN, Koh DI, et al. Regulation of the cyclin-dependent kinase inhibitor 1A gene (CDKN1A) by the repressor BOZF1 through inhibition of p53 acetylation and transcription factor Sp1 binding. *J Biol Chem*. 2013;288(10):7053-7064. doi:10.1074/jbc.M112.416297
18. Caramel J, Ligier M, Puisieux A. Pleiotropic roles for ZEB1 in cancer. *Cancer Res*. 2018;78(1):30-35. doi:10.1158/0008-5472.CAN-17-2476
19. Sánchez-Tilló E, Fanlo L, Siles L, et al. The EMT activator ZEB1 promotes tumor growth and determines differential response to chemotherapy in mantle cell lymphoma. *Cell Death Differ*. 2014;21:247-257. doi:10.1038/cdd.2013.123
20. Stanley RF, Abdel-Wahab O. Dysregulation and therapeutic targeting of RNA splicing in cancer. *Nat Cancer*. 2022;3(5):536-546. doi:10.1038/s43018-022-00384-z
21. Jung H, Lee D, Lee J, et al. Intron retention is a widespread mechanism of tumor-suppressor inactivation. *Nat Genet*. 2015;47(11):1242-1248. doi:10.1038/ng.3414
22. Hung SY, Lin CC, Hsu CL, et al. The expression levels of long non-coding RNA KIAA0125 are associated with distinct clinical and biological features in myelodysplastic syndromes. *Br J Haematol*. 2021;192(3):589-598. doi:10.1111/bjh.17231
23. Benetatos L, Hatzimichael E, Dasoula A, et al. CpG methylation analysis of the MEG3 and SNRPN imprinted genes in acute myeloid leukemia and myelodysplastic syndromes. *Leuk Res*. 2010;34(2):148-153. doi:10.1016/J.LEUKRES.2009.06.019
24. Szikszai K, Krejci Z, Klema J, et al. LncRNA profiling reveals that the deregulation of H19, WT1-AS, TCL6, and LEF1-AS1 is associated with higher-risk myelodysplastic syndrome. *Cancers (Basel)*. 2020;12(10):1-21. doi:10.3390/cancers12102726
25. Wu Q, Xiang S, Ma J, et al. Long non-coding RNA CASC15 regulates gastric cancer cell proliferation, migration and epithelial mesenchymal transition by targeting CDKN1A and ZEB1. *Mol Oncol*. 2018;12(6):799-813. doi:10.1002/1878-0261.12187
26. Hu Y, Lin J, Fang H, et al. Targeting the MALAT1/PARP1/LIG3 complex induces DNA damage and apoptosis in multiple myeloma. *Leukemia*. 2018;32(10):2250-2262. doi:10.1038/S41375-018-0104-2
27. Panchal NK, Evan Prince S. The NEK family of serine/threonine kinases as a biomarker for cancer. *Clin Exp Med*. 2023;23(1):17-30. doi:10.1007/s10238-021-00782-0
28. Nachmias B, Khan DH, Voisin V, et al. IPO11 regulates the nuclear import of BZW1/2 and is necessary for AML cells and stem cells. *Leukemia*. 2022;36(5):1283-1295. doi:10.1038/s41375-022-01513-4
29. Stavropoulou V, Kaspar S, Brault L, et al. MLL-AF9 expression in hematopoietic stem cells drives a highly invasive AML expressing EMT-related genes linked to poor outcome. *Cancer Cell*. 2016;30(1):43-58. doi:10.1016/j.ccell.2016.05.011
30. Eger A, Aigner K, Sonderegger S, et al. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Oncogene*. 2005;24(14):2375-2385. doi:10.1038/sj.onc.1208429
31. Yoshino S, Yokoyama T, Sunami Y, et al. Trib1 promotes acute myeloid leukemia progression by modulating the transcriptional programs of Hoxa9. *Blood*. 2021;137(1):75-88. doi:10.1182/blood.2019004586
32. Gheldof A, Hulpiau P, van Roy F, de Craene B, Berx G. Evolutionary functional analysis and molecular regulation of the ZEB transcription factors. *Cell Mol Life Sci*. 2012;69(15):2527-2541. doi:10.1007/s00018-012-0935-3
33. Yu QC, Geng A, Preusch CB, et al. Activation of Wnt/ β -catenin signaling by Zeb1 in endothelial progenitors induces vascular quiescence entry. *Cell Rep*. 2022;41(8):111694. doi:10.1016/J.CELREP.2022.111694
34. Drápela S, Bouchal J, Jolly MK, Culig Z, Souček K. ZEB1: a critical regulator of cell plasticity, DNA damage response, and therapy resistance. *Front Mol Biosci*. 2020;7:36. doi:10.3389/FMOLB.2020.00036
35. Rouillard AD, Gundersen GW, Fernandez NF, et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database (Oxford)*. 2016;2016:baw100. doi:10.1093/DATABASE/BAW100
36. Hu Z, Mao JH, Curtis C, et al. Genome co-amplification upregulates a mitotic gene network activity that predicts outcome and response to mitotic protein inhibitors in breast cancer. *Breast Cancer Res*. 2016;18(1):70. doi:10.1186/S13058-016-0728-Y
37. Maturi V, Enroth S, Heldin CH, Moustakas A. Genome-wide binding of transcription factor ZEB1 in triple-negative breast cancer cells. *J Cell Physiol*. 2018;233(10):7113-7127. doi:10.1002/JCP.26634
38. Will B, Zhou L, Vogler TO, et al. Stem and progenitor cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations. *Blood*. 2012;120(10):2076-2086. doi:10.1182/blood-2011-12-399683
39. Chen Y, Li J, Xu L, Găman MA, Zou Z. The genesis and evolution of acute myeloid leukemia stem cells in the microenvironment: from biology to therapeutic targeting. *Cell Death Discov*. 2022;8(1):397. doi:10.1038/s41420-022-01193-0
40. Knijnenburg TA, Wang L, Zimmermann MT, et al. Genomic and molecular landscape of DNA damage repair deficiency across the cancer genome atlas. *Cell Rep*. 2018;23(1):239-254. doi:10.1016/J.CELREP.2018.03.076
41. Pellagatti A, Armstrong RN, Steeples V, et al. Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. *Blood*. 2018;132(12):1225-1240. doi:10.1182/blood-2018-04-843771
42. Flach J, Jann JC, Knaflic A, et al. Replication stress signaling is a therapeutic target in myelodysplastic syndromes with splicing factor mutations. *Haematologica*. 2021;106(11):2906-2917. doi:10.3324/haematol.2020.254193
43. Zhang Q, Li H, Jin H, Tan H, Zhang J, Sheng S. The global landscape of intron retentions in lung adenocarcinoma. *BMC Med Genomics*. 2014;7(1):1-9. doi:10.1186/1755-8794-7-15
44. Shiozawa Y, Malcovati L, Galli A, et al. Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nat Commun*. 2018;9(1):3649. doi:10.1038/s41467-018-06063-x
45. Yang YT, Chiu YC, Kao CJ, et al. The prognostic significance of global aberrant alternative splicing in patients with myelodysplastic syndrome. *Blood Cancer J*. 2018;8(8):78. doi:10.1038/s41408-018-0115-2
46. Yao CY, Chen CH, Huang HH, et al. A 4-lncRNA scoring system for prognostication of adult myelodysplastic syndromes. *Blood Adv*. 2017;1(19):1505-1516. doi:10.1182/bloodadvances.2017008284
47. Nadhan R, Dhanasekaran DN. Decoding the oncogenic signals from the long non-coding RNAs. *Onco*. 2021;1(2):176-206. doi:10.3390/ONCO1020014/S1
48. Melki JR, Vincent PC, Brown RD, Clark SJ. Hypermethylation of E-cadherin in leukemia. *Blood*. 2000;95(10):3208-3213. doi:10.1182/BLOOD.V95.10.3208
49. Corn PG, Smith BD, Ruckdeschel ES, Douglas D, Baylin SB, Herman JG. E-cadherin expression is silenced by 5' CpG Island

- methylation in acute leukemia. *Clin Cancer Res.* 2000;6(11):4243-4248.
50. Aggerholm A, Holm MS, Guldborg P, Olesen LH, Hokland P. Promoter hypermethylation of p15INK4B, HIC1, CDH1, and ER is frequent in myelodysplastic syndrome and predicts poor prognosis in early-stage patients. *Eur J Haematol.* 2006;76(1):23-32. doi:10.1111/j.1600-0609.2005.00559.x
51. Weyemi U, Redon CE, Choudhuri R, et al. The histone variant H2A.X is a regulator of the epithelial-mesenchymal transition. *Nat Commun.* 2016;7(1):1-12. doi:10.1038/ncomms10711
52. Peñalosa-Ruiz G, Bousgouni V, Gerlach JP, et al. WDR5, BRCA1, and BARD1 co-regulate the DNA damage response and modulate the mesenchymal-to-epithelial transition during early reprogramming. *Stem Cell Rep.* 2019;12(4):743-756. doi:10.1016/J.STEMCR.2019.02.006
53. Fu R, Li Y, Jiang N, et al. Inactivation of endothelial ZEB1 impedes tumor progression and sensitizes tumors to conventional therapies. *J Clin Invest.* 2020;130(3):1252-1270. doi:10.1172/JCI131507
54. Sánchez-Tilló E, de Barrios O, Siles L, Cuatrecasas M, Castells A, Postigo A. β -Catenin/TCF4 complex induces the epithelial-to-mesenchymal transition (EMT)-activator ZEB1 to regulate tumor invasiveness. *Proc Natl Acad Sci U S A.* 2011;108(48):19204-19209. doi:10.1073/PNAS.1108977108/-/DCSUPPLEMENTAL/PNAS.1108977108_SI.PDF
55. Unnikrishnan A, Papaemmanuil E, Beck D, et al. Integrative genomics identifies the molecular basis of resistance to Azacitidine therapy in myelodysplastic syndromes. *Cell Rep.* 2017;20(3):572-585. doi:10.1016/j.celrep.2017.06.067
56. Stopka T, Minařík L, Dusilková N, et al. G-CSF plus azacitidine versus azacitidine alone for patients with high-risk myelodysplastic syndrome: academic, open label, randomized trial. *Blood Cancer J.* 2022; 12(7):105. doi:10.1038/S41408-022-00698-2

SUPPORTING INFORMATION

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

How to cite this article: Kaisrlikova M, Kundrat D, Koralkova P, et al. Attenuated cell cycle and DNA damage response transcriptome signatures and overrepresented cell adhesion processes imply accelerated progression in patients with lower-risk myelodysplastic neoplasms. *Int J Cancer.* 2024; 1-17. doi:10.1002/ijc.34834