NEOPLASMA accepted, ahead of print manuscript
Cite article as https://doi.org/10.4149/neo_2026_260622N205

**Running title:** NUDT21 in MDS leukemic transformation

# *NUDT21* knockdown modulates alternative polyadenylation and promotes leukemic transformation in myelodysplastic syndromes

Shuo Li[1], Jiakai Bian[1], Peichun Li[1], Wenjie Xu[1], Jingyi Feng[2,3], Zhuanghui Hao[1], Ju Deng[1], Dongmei Guan[1], Zhifang Xu[2,3], Xiuhua Chen[2,3], Hongwei Wang[1,3,4,*], Fanggang Ren[2,3,*]

[1]Shanxi Medical University, Taiyuan, Shanxi, China; [2]Laboratory of Hematology, Second Hospital of Shanxi Medical University, Taiyuan, Shanxi, China; [3]The Key Laboratory of Molecular Diagnosis and Treatment of Hematological Diseases of Shanxi Province, Taiyuan, Shanxi, China; [4]Shanxi University of Chinese Medicine, Jinzhong, Shanxi, China

*Correspondence: wanghw68@hotmail.com; burton@sxmu.edu.cn

**Received June 22, 2026 / Accepted October 2, 2026**

Progression of myelodysplastic neoplasms (MDS) to secondary acute myeloid leukemia (sAML) involves complex clonal evolution; however, the role of alternative polyadenylation (APA) regulators such as *NUDT21* remains unclear. This study aimed to investigate the effects of *NUDT21* knockdown on gene expression profiles in MDS cells and their correlation with clinicopathological features in MDS.

*NUDT21* was knocked down via shRNA in MDS cell lines (SKM-1, MUTZ-1), followed by transcriptome sequencing, functional assays, and clinical validation in 40 MDS patients. Rescue experiments were performed by overexpressing *DDX60* in *NUDT21* knockdown cells and treating cells with the *STAT3* inhibitor Stattic to validate the downstream mechanism.

Data showed that knockdown of *NUDT21* altered the expression of multiple downstream genes and related pathways. The expressions of *DDX60*, *SRPK3*, and *XAF1* were positively correlated with the expression of *NUDT21*. *NUDT21* overexpression was associated with the increased sAML transformation rate in MDS patients. *NUDT21* was associated with the transformation of MDS into sAML, potentially by increasing the expression level of *DDX60* and activating the *JAK2/STAT3* pathway. Rescue experiments confirmed that *DDX60* overexpression restored *JAK2/STAT3* phosphorylation and reversed the phenotypic changes induced by *NUDT21* knockdown, while *STAT3* inhibition abrogated these effects, further supporting the role of the *DDX60-JAK2/STAT3* axis in *NUDT21*-mediated MDS progression.

In conclusion, *NUDT21* appears to be associated with disease progression and leukemic transformation in MDS.

**Key words:** MDS; AML; *NUDT21*; transcriptome sequencing

Myelodysplastic neoplasms (MDS) are a group of hematological neoplasms arising from the clonal expansion of mutated pluripotent hematopoietic progenitor cells, characterized by aberrant differentiation, morphological abnormalities, and ineffective hematopoiesis. According to the 5th



Accepted manuscript

edition of the WHO classification, MDS is categorized into distinct subtypes based on blast percentage, cytogenetic abnormalities, and molecular features, with MDS with increased blasts (MDS-IB) exhibiting a particularly high risk of leukemic transformation [1]. Although MDS predominantly affects middle-aged and elderly individuals, recent studies indicate a trend toward younger age at diagnosis [2-4]. A low percentage of blasts cells in the bone marrow helps distinguish MDS from acute myeloid leukemia (AML) [5]. Although they are clinically distinct disease entities, the boundary between MDS and AML is not always clear [6, 7]. Both diseases share numerous genetic mutations, from clonal hematopoiesis of indeterminate potential (CHIP) through to overt MDS/secondary AML (sAML), suggesting they may represent a disease continuum [8, 9]. However, the prognosis and overall survival of patients with MDS are generally better than those of patients with sAML, highlighting the important clinical significance of preventing MDS progression to sAML [10].

*NUDT21* is a highly conserved component of the cleavage factor I complex (CFIm), involved in the early stages of eukaryotic precursor mRNA (pre-mRNA) processing [11]. In normal cells, *NUDT21* dimerizes with CFIm68 to binds to UGUA motifs upstream of polyadenylation sites, and stabilizes the stem-loop structure, thereby inhibiting the use of proximal polyA sites [12-14]. Under physiological conditions, *NUDT21* promotes the use of distal polyadenylation signals or suppresses the use of proximal signals [15, 16]. Retaining distal polyadenylation signals typically results in longer 3′UTRs, which contain stronger and more conserved distal cis-acting compared with proximal elements [11, 16]. Alterations in 3′UTR length or structure remodel the binding of trans-acting regulators such as miRNAs, modulating RNA-binding protein activity, mRNA stability, translation efficiency, and ultimately protein expression,and further reshaping cellular phenotypes. This mechanism defines the core regulatory function of *NUDT21* in alternative polyadenylation (APA) [17-20].

Polyadenylation governs pre-mRNA 3′-end processing and polyA tail formation, which is indispensable for mRNA stability and cellular homeostasis [21-23]. More than 70% of human genes contain multiple APA sites, enabling the generation of mRNA transcripts of varying lengths [24, 25]. The 3′UTR contains binding sites for various RNA regulatory elements that modulate gene expression post-transcriptionally [26, 27]. Therefore, the length of the 3′UTR significantly affects the accessibility of these regulatory elements [28]. As a cleavage and polyadenylation factor essential for processing the mRNA 3′UTR, *NUDT21* plays a central role in APA regulation [29]. Although *NUDT21* dysregulation has been implicated in various solid tumors, its mechanistic role in MDS remains unclear [30].

Our previous work confirmed that *NUDT21* participates in APA and can regulate the biological functions of MDS cell lines through APA mechanisms, thereby promoting leukemic transformation

Accepted manuscript

[31]. This study aims to further elucidate the downstream targets of *NUDT21* and their correlation with APA through bioinformatics analysis, in order to identify key factors driving MDS transformation.

## Patients and Methods

**Cell culture.** The human MDS cell line SKM-1 was purchased from China Otto Bio (Shenzhen, China), and MUTZ-1 was purchased from China Fenghui Bio (Jiangsu, China). Both cell lines were cultured in RPMI-1640 (Gibco, USA), and all media were prepared with 10% fetal bovine serum (BI, IL) and 1% penicillin (SEVEN BIOTECH, China). All the cells were cultured at 37 °C in a 5% $CO_2$ constant temperature incubator (Fisher, USA). All cells were verified using STR to confirm the absence of mycoplasma infection. All functional experiments (including apoptosis, cell cycle, flow cytometry, transwell migration, colony formation, qRT-PCR, and Western blot) were performed in three independent biological replicates, with each biological replicate consisting of at least two technical replicates.

**Clinical samples.** Bone marrow aspirate samples were collected from newly diagnosed MDS patients (n=40) and normal controls (n=8) in the Second Hospital of Shanxi Medical University from January 2023 to June 2024 and stored at -20 °C for RNA extraction. The healthy controls were healthy individuals with no history of major diseases. AML transformation was defined as progression from MDS to acute myeloid leukemia with ≥ 20% blasts in bone marrow or peripheral blood, according to WHO 2022 criteria. All samples were approved by the Ethics Committee of the Second Hospital of Shanxi Medical University and an informed consent waiver was obtained from the IRB for this study (Approval No.: [2024] YX No. 363). And all the methods adopted in this research are carried out in accordance with relevant guidelines and regulations. For classification of patients into *NUDT21* expression subgroups, the median *NUDT21* expression level (measured by qRT-PCR, normalized to GAPDH) across all 40 MDS patients was used as the cutoff. The median was chosen as it is the most commonly used and statistically robust method for exploratory biomarker studies with continuous variables and limited sample size, avoiding the arbitrary selection of cutoffs that could inflate type I error. Patients with expression values above the median were assigned to the $NUDT21^{high}$ group (n=20), and those with values below or equal to the median were assigned to the $NUDT21^{low}$ group (n=20).

**qRT-PCR.** Total RNA was extracted using TRIzol reagent (Takara, Japan), and cDNA was synthesized using a cDNA synthesis kit (Takara, Japan). Analysis was performed by ABI 7500 real-time PCR instrument (USA) using SYBR Green Mix reagent (SEVEN BIOTECH, China). The cycling

Accepted manuscript

conditions were as follows: 95 °C pre-denaturation for 30 s, 95 °C pre-denaturation for 5 s, 60 °C pre-denaturation for 40 s, and 40 cycles of amplification. The expression levels were analyzed using the $2^{-\Delta\Delta}$Ct method. The primer sequences are shown below: *NUDT21*-F: AGATTTCAGCGCATGAGGGAA; *NUDT21*-R: GCAGCAGTAACACATGGGGT; *COL13A1*-F: GGAGACGGCTATTTTGGGACG; *COL13A1*-R: TCCTTGAGTGGAGCTTCCATT; *SRPK3*-F: GATGAGGCGCAAACGGAAAC; *SRPK3*-R: CCCGTCTAGTCTCAAGCCAGA; *SLC6A2*-F: GCCAACTCGCTACAGTTCGT; *SLC6A2*-R: TACCTCCGCCGTACATCTGG; *DDX60*-F: TGAGGACCGCTTTAATTCTCCA; *DDX60*-R: ACAACATTGACTTTCATTCCCCA; *RYR2*-F: ACAACAGAAGCTATGCTTGGC; *RYR2*-R: GAGGAGTGTTCGATGACCACC; *XAF1*-F: GCTCCACGAGTCCTACTGTG; *XAF1*-R: GTTCACTGCGACAGACATCTC; *JAK2*-F: GGCCTCAGATGTTTGGAGCT; *JAK2*-R: CGCATAAATTCCGCTGGTGG; *STAT3*-F: AGCAGCACCTTCAGGATGTC; *STAT3*-R: GCATCTTCTGCCTGGTCACT; *GAPDH*-F: ACAACTTTGGTATCGTGGAAGG; *GAPDH*-R: GCCATCACGCCACAGTTTC; *DDX60*-3′UTR-F: CAGTGTTTCCTTGCGTGAGC; *DDX60*-3′UTR-R: TCAGTCACATACAACATTCTCGCA

**Western blot.** RIPA protein lysate was mixed with protease inhibitor (SEVEN BIOTECH, China) at the appropriate ratio according to the manufacturer's instructions for cell lysis. Protein concentration was detected using a BCA kit (SEVEN BIOTECH, China), electrophoretic separation was performed on a 12.5% SDS-PAGE gel (SEVEN BIOTECH, China), 5% skimmed milk powder was used to seal the film, and a monoclonal antibody was added and incubated at 4 °C overnight. HRP-coupled secondary antibody (Immunoway, USA) was used at 1:3000 dilution at room temperature action for 1.5 h, and the signal was detected using ECL-HRP substrate. The following antibodies were used: anti-NUDT21 (Abcam, UK, #ab221994), anti-JAK2 (Biodragon, China, #RM8013), anti-phospho-JAK2 (Biodragon, China, #RM4972), anti-STAT3 (Cell Signaling, US, #12640S), anti-phospho-STAT3 (Cell Signaling, US, #9145S) and anti-β-actin (Proteintech, US, #20536-1-AP) and HRP goat anti-IgG (H+L) (Immunoway, #RS0002).

**Flow cytometry.** The cell cycle was analyzed using the kit (Bioss, Beijing, China). After collecting the cells, they were rinsed twice with PBS, overnight with 95% ethanol at 4 °C, labeled with propidium iodide and RNaseA solution, incubated at room temperature and in the dark for 30 min, and detected by Navious (Beckman, USA) flow cytometer. The data were analyzed using the Modfit software. The gating strategy for cell cycle analysis was based on FSC/SSC to exclude debris and doublets, and cell cycle distribution was analyzed on the gated single-cell population.

Accepted manuscript

Cell apoptosis was analyzed using the kit (KeyGEN BioTEC, Jiangsu, China). After collecting the cells, they were rinsed twice with PBS. The cells were labeled with 7-AAD and Annexin-V and incubated at room temperature and in the dark for 15 min. Detection and analysis were performed using the Navious (Beckman, USA) flow cytometer. Apoptosis was analyzed on the gated viable cell population based on FSC/SSC; early apoptotic cells were identified as Annexin V-positive/7-AAD-negative, and late apoptotic cells as Annexin V-positive/7-AAD-positive.

We also selected the cell surface antigens CD11b (Beckman, USA, #IM2581U) and CD117 (Beckman, USA, #C41144) for flow cytometry analysis. After collecting the cells, antibodies were incubated in a flow tube for 2 h at room temperature in the dark. Appropriate isotype-matched controls (Beckman, USA) were used to determine background fluorescence and set the gating boundaries for positive populations. A minimum of 10,000 events were acquired per sample. The gating strategy was as follows: viable cells were first selected based on forward scatter (FSC) and side scatter (SSC) to exclude debris and doublets, and subsequent analysis was performed on the gated viable cell population. FlowJo software (v10.8, BD Biosciences, USA) was used for data analysis. Detected and analyzed using a Navious flow cytometer from Beckman, USA.

Cytokines were detected using a kit (UNI-MEDICA 7 cytokine kit, #15.06.002.02). The cell supernatant and sample buffer to be tested were collected in a flow tube, then the vortexed capture microspheres mixture and tested antibodies (IL-2, IL-4, IL-6, IL-10, IL-17, IFN-γ, TNF-α) were added, and incubated at room temperature in a dark place for 2 h. Then the fluorescence detection reagent was added to each tube, and incubated at room temperature in a dark place for 30 minutes. Wash buffer was added to each experimental tube prior to acquisition on the Navios flow cytometer. Data were analyzed using the instrument-specific software with gating on the microsphere population based on FSC/SSC.

**Migration of cells.** The cells were collected and the cell density was adjusted to $2.5 \times 10^5$/ml using the basic medium. 500 μl complete medium was added to the lower chamber of the 24-well plate, and 200 μl cells were added to the transwell chamber, and the cells were incubated at 37 °C with 5% $CO_2$ for 48 h. After rinsing with PBS, the cells were stained with 0.1% crystal violet and observed under a microscope after rinsing with PBS.

**Clone formation assay.** The cells were collected and the cell density was adjusted to $5 \times 10^3$/ml with complete medium. 300 μl of the cell suspension was added to 3 ml of thawed methylcellulose semi-solid medium. After the bubbles disappeared completely, the methylcellulose semi-solid medium was added to a 24-well plate, 1 ml per well, and incubated in a 37 °C, 5% $CO_2$ constant temperature incubator. Colony formation was observed microscopically after 14 days.

Accepted manuscript

**RNA isolation and library construction.** Total RNA was extracted from cells or bone marrow blood using TRIzol reagent (Takara). RNA purity was assessed using a BioTek Cytation3 (USA) multifunctional enzyme-labeling instrument. The samples were purified using magnetic bead sorting. Library construction was performed using a Bioanalyzer (model 2100; Agilent, USA). Transcriptome sequencing and analysis were conducted by OE Biotech Co. Ltd. (Shanghai, China).

**RNA sequencing.** Total RNA was extracted from cells using TRIzol reagent (Takara, Japan) following the manufacturer's instructions. RNA purity and concentration were assessed using a BioTek Cytation3 multifunctional microplate reader (USA), and RNA integrity was evaluated using an Agilent 2100 Bioanalyzer (USA) with RNA integrity numbers (RIN) ≥ 7.0 required for library construction. mRNA was enriched from 1 μg of total RNA using oligo(dT)-coated magnetic beads (Thermo Fisher, USA), and cDNA was synthesized using the cDNA Synthesis Kit (Takara, Japan) according to the manufacturer's protocol. After end-repair, A-tailing, adapter ligation, and PCR amplification, the resulting cDNA libraries were quantified and qualified using an Agilent 2100 Bioanalyzer. Paired-end sequencing (125 bp or 150 bp) was performed on the Illumina NovaSeq 6000 platform (Illumina, USA). Each experimental condition was assessed in three independent biological replicates.

Raw sequencing reads were processed to remove adapter sequences and low-quality bases using Trimmomatic (v0.39). Clean reads were aligned to the human reference genome (GRCh38) using HISAT2 (v2.2.1). The mapping rate for each sample was > 90%. Gene-level read counts were quantified using featureCounts (v2.0.1) against the GENCODE annotation (v43). These raw read counts were used as input for subsequent differential expression analysis. All raw RNA-seq data have been deposited in the Gene Expression Omnibus (GSE280172) under accession number.

**Gene level quantification, differential gene screening, functional enrichment, and cluster analysis.**

Differential expression analysis was performed using the DESeq2 R package (v1.36.0). Raw read counts obtained from featureCounts (section 9) were used as input for DESeq2. The data were normalized using the estimateSizeFactors function, and p-values and fold changes were calculated using the DESeq2 Wald test. Genes with |fold change| ≥ 2 and FDR < 0.05 were selected for GO (http://geneontology.org/) and KEGG (http://www.genome.jp/kegg/) [32-34] enrichment analysis to predict the downstream pathways that these differential genes might affect. Principal component analysis (PCA) was performed to confirm overall sample clustering and to exclude batch effects between sequencing runs (data not shown). For multiple testing, the false discovery rate (FDR) was controlled using the Benjamini-Hochberg procedure, and genes with FDR < 0.05 were considered

Accepted manuscript

statistically significant. In addition, unsupervised hierarchical clustering of differential genes was performed and their expression patterns in different samples were visualized using heatmaps, volcano maps, and clustered heat maps using the https://cloud.oebiotech.cn/ online analysis website.

**Transcript-level quantification, differential gene screening, functional enrichment, and cluster analysis.** Transcript-level quantification was performed using eXpress software to obtain read counts for protein-coding transcripts. Differential transcript analysis was then performed using the DESeq2 R package (v1.36.0), with the same criteria as gene-level analysis: $|\text{fold change}| \geq 2$ and FDR < 0.05 by the Benjamini-Hochberg procedure. GO and KEGG enrichment analyses of differential transcripts were performed to determine the biological functions or pathways that were mainly affected by the differential transcripts. At the same time, unsupervised hierarchical clustering of differential transcripts was performed, and the expression patterns of differential transcripts among different samples were shown in the form of heat maps.

**Protein interaction network analysis.** We performed BLAST (e-value less than $10^{-10}$) between species gene sequences and protein sequences using annotations from the STRING database (https://cn.string-db.org/), after which we identified interrelationships between genes. Through this analysis, we found the interactions of differentially expressed genes and filtered out the top 300 scoring genes. Subsequently, a ring diagram representing the top 30 protein-protein interaction (PPI) relationships was plotted.

**Rescue experiments.** The double-stranded DNA fragment encoding the full-length coding sequence of human *DDX60* was designed, synthesized, and cloned into the empty vector. The recombinant plasmid was transformed into competent Escherichia coli DH5α cells for amplification. Positive clones were selected through monoclonal colony screening, expanded, and subjected to plasmid extraction. The construct was verified by agarose gel electrophoresis and DNA sequencing. For transfection, *NUDT21* knockdown SKM-1 and MUTZ-1 cells were transfected with the *DDX60*-overexpression plasmid or empty vector using LipoFiter 3.0 transfection reagent (Hanbio Biotechnology) in serum-free and antibiotic-free DMEM according to the manufacturer's instructions. Stable cell lines were established by puromycin selection. Transfection efficiency was confirmed by qRT-PCR.

The *STAT3* inhibitor Stattic (MCE) were incubated with serial dilutions for 48 h. The half-maximal inhibitory concentration (IC50) values were determined by measuring absorbance using the CCK-8 assay, and the IC50 concentrations were selected as the experimental drug concentrations.

**Statistical analyses.** All statistical analyses were performed using GraphPad Prism 8.3.0 and R 4.2.1. Comparisons between two groups were analyzed using the Mann- Whitney U test or unpaired two-

Accepted manuscript

tailed Student's t-test, as appropriate. Categorical variables were compared using the $\chi^2$ test or Fisher's exact test.

For RNA-seq data, differential expression was analyzed using DESeq2, with multiple comparisons controlled by the Benjamini- Hochberg false discovery rate (FDR) procedure; genes with |fold change| $\geq 2$ and FDR $< 0.05$ were considered significant. For clinical correlation analyses, p-values are reported as nominal without multiplicity adjustment, given the exploratory design and limited pre-specified hypotheses. Effect sizes for key comparisons are reported as Cohen's d (continuous variables) or odds ratios (OR) with 95% confidence intervals (CI) for binary outcomes. Given the pilot nature of this study, no formal a priori power calculation was performed; instead, 95% CIs and effect sizes are provided to facilitate interpretation of the precision and clinical significance of our findings.

All tests were two- sided, and $p < 0.05$ was considered statistically significant. Significance is denoted as: $^{ns}p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

To assess the independent prognostic value of *NUDT21* for sAML transformation, multivariate logistic regression analysis was performed with sAML transformation as the dependent variable and *NUDT21* expression (high vs. low), IPSS-M risk category, blast percentage, and cytogenetics as covariates. Results are presented as adjusted odds ratios (OR) with 95% confidence intervals (CI).

## Results

**Construction of *NUDT21* knockdown cell model.** A lentiviral plasmid for *NUDT21* knockdown was constructed. Lentivirus was packaged in HEK293T cells and used to infect MDS cell lines SKM-1 and MUTZ-1, establishing stable *NUDT21*-knockdown cell models. Successful knockdown was confirmed by qRT-PCR ($p \leq 0.001$ for MUTZ-1; $p \leq 0.001$ for SKM-1) (Figure 1A) and Western blot ($p \leq 0.001$ for MUTZ-1; $p \leq 0.001$ for SKM-1) (Figure 1B).

***NUDT21* can affect the biological function of MDS cell lines.** Apoptosis was measured by flow cytometry using Annexin V as a marker of early apoptosis. Knockdown of *NUDT21* led to a significant increase in apoptosis in both SKM-1 ($p=0.017$) and MUTZ-1 cells ($p \leq 0.001$) (Figure 1C). Cell cycle analysis revealed that *NUDT21* knockdown induced G1 phase arrest ($p=0.023$ for MUTZ-1; $p \leq 0.001$ for SKM-1). No obvious change was observed in G2 phase ($p=0.521$ for MUTZ-1; $p=0.253$ for SKM-1), while the S phase decreased ($p=0.268$ for MUTZ-1; $p \leq 0.001$ for SKM-1) (Figure 1D).

In addition, the expression of CD117 and CD11b in the two cell lines was detected.CD117 is a membrane receptor tyrosine kinase that is closely related to the proliferation of cancer cells. CD117 drives individual development and balance by binding to stem cell factor, forming dimers, activating

Accepted manuscript

tyrosine kinase activity, and influencing signal transduction molecules to regulate cell cycle and differentiation patterns. However, CD11b is a molecular marker of differentiation and is often detected together with CD117 to indicate cell proliferation and differentiation. Our results showed that the expression of CD117 was significantly increased in MUTZ-1 ($p \leq 0.001$) and SKM-1 cells ($p \leq 0.001$) after *NUDT21* knockdown. The expression of CD11b was increased in MUTZ-1 cells ($p=0.0058$) and decreased in SKM-1 cells ($p \leq 0.001$) (Figure 1E). Although CD117 upregulation appeared to contrast with the reduced proliferation observed after *NUDT21* knockdown, CD117 expression in MDS cells is not necessarily a direct correlate of proliferative activity; rather, it is frequently associated with a more undifferentiated or stem-like state. Thus, the observed increase in CD117 may reflect differentiation arrest or altered progenitor characteristics following *NUDT21* knockdown.

We also examined the cytokine levels in SKM1 and MUTZ1 cell culture supernatants after *NUDT21* knockdown. Results showed that *NUDT21* knockdown decreased the expression of IL-4 ($p=0.004$), IL-6 ($p \leq 0.001$), IL-10 ($p=0.0039$) and IL-17 ($p=0.0155$) and increased the expression of IFN-γ ($p=0.0069$) in SKM1. *NUDT21* knockdown in MUTZ1 also decreased the expression of IL-4 ($p \leq 0.001$), IL-6 ($p \leq 0.001$), IL-10 ($p \leq 0.001$) and IL-17 ($p=0.006$) and increased the expression of IFN-γ ($p=0.0355$) (Figure 2A).

Transwell assays were used to evaluate the effect of *NUDT21* on cell migration. *NUDT21* knockdown significantly reduced migration of SKM-1 ($p=0.0038$) and MUTZ-1 cells ($p \leq 0.001$) toward high- serum conditions (Figure 2B). Colony formation assays in methylcellulose semi-solid medium showed that *NUDT21* knockdown significantly reduced both the number and size of colonies (Figure 2C).

Combined with the apoptosis results, these findings suggest that *NUDT21* expression may be associated with the progression from MDS to sAML.

**RNA transcriptome profiling of *NUDT21* knockdown cell lines.** Transcriptome sequencing was performed to explore the role of *NUDT21* in MDS transformation. Compared with control cells, *NUDT21* knockdown cells exhibited a large number of differentially expressed genes (DEGs). In MUTZ- 1- KD vs. MUTZ- 1, 511 DEGs were identified (333 up and 178 down regulated). In SKM- 1- KD vs. SKM- 1; 5,054 DEGs were detected (1,691 up and 3,363 down regulated).

In MUTZ- 1, the top 10 significantly up- regulated genes ($|\text{Fold Change}| \geq 2$, $\text{FDR} < 0.05$) were *MYL1*, *RALYL*, *PDE4C*, *SLITRK6*, *VIPR1*, *TMPRSS3*, *SYT4*, *CFAP69*, *MAT1A*, and *UBASH3A*; the top 10 down- regulated genes were *XAF1*, *CYSLTR1*, *FCRL4*, *FOXL1*, *CYP2W1*, *DDX60*, *IFI44L*, *TNFRSF13B*, *RYR2*, and *CCL22* (Figure 3A). In SKM- 1, the top 10 up- regulated genes were

Accepted manuscript

*TMEM178A*, *SRPK3*, *FOXL1*, *EPHB3*, *SLC6A2*, *CSMD2*, *DPP4*, *COL13A1*, *PRSS56*, and INHBB; the top 10 down- regulated genes were *SLC6A15*, *GPR174*, *ZNF835, XIRP2*, *SERTM2*, *PPP4R3C*, *SEMA3E*, *PIEZO2*, *NTS*, and *LMO3* (Figure 3B). Most of these genes are implicated in immunity or tumor progression. Six tumor-related DEGs (*COL13A1, SLC6A2, SRPK3, DDX60, RYR2, XAF1*) with the most significant differences were selected for validation in patient bone marrow samples by qRT-PCR. Among these, *DDX60* was prioritized for subsequent mechanistic analysis because it was consistently downregulated upon *NUDT21* knockdown in both SKM- 1 and MUTZ- 1 cells, and because its known function as an interferon- stimulated gene involved in inflammatory signaling made it a biologically plausible candidate for further investigation. In contrast, *XAF1* showed a positive correlation with *NUDT21* that was inconsistent with its established tumor- suppressive role, and *SRPK3* lacked a clear link to the *JAK2/STAT3* pathway, making *DDX60* the most suitable choice for functional studies.

In addition, *NUDT21* is involved in the early stages of eukaryotic pre-mRNA processing and is associated with the 3′UTRs length of a variety of genes. The length of 3′UTRs of *NUDT21* was also examined in both cell lines. The results showed that 3,094 genes became shorter in 3′UTRs and 775 genes became longer in MUTZ-1 knockdown *NUDT21*. There were 4,002 genes with shorter 3′UTRs and 595 genes with longer 3′UTRs after SKM-1 knockdown *NUDT21* (Figure 3C).

**GO enrichment analysis of differentially expressed genes.** GO enrichment analysis was performed to characterize the functions of the DEGs. The top 10 enriched terms in biological process (BP), cellular component (CC), and molecular function (MF) categories are displayed using enrichment and chord plots (Figures 4A, 4B). The left side shows the 10 genes with the largest |logFC| in each category, the right side shows the category composition, and connecting lines indicate gene- category relationships. The outer heatmap displays logFC values.

Circle diagrams illustrate the distribution of DEGs across GO categories (Figures 4C, 4D). The outermost circle shows enriched terms with a coordinate scale for gene counts; the second circle indicates the number of background genes and the Q or *P* value; the third circle shows the ratio of up- /down- regulated genes; the innermost circle displays the RichFactor (foreground/background gene ratio).

A more detailed subclassification within BP, CC, and MF (including 64 sub- entries such as cell adhesion) was applied to all genes, up- regulated DEGs, and down- regulated DEGs. In both cell lines, the most prominent changes occurred in BP categories such as biological regulation, cellular process, metabolic process, and regulation of biological process. Terms related to organelle, membrane (CC),

Accepted manuscript

and binding (MF) were also significantly affected (Figures 3E-3H). These results suggest that *NUDT21* is associated with alterations in multiple biological processes, providing candidate pathways for further mechanistic investigation.

**KEGG pathway analysis.** KEGG enrichment circle diagrams display the 20 terms with the smallest *Q* or p-values (Figures 5A, 5B). The layout is similar to the GO circles, showing gene counts, *Q/P* values, up- /down- regulation ratios, and RichFactor values.

Analysis based on the four major KEGG categories (Metabolism, Human Diseases, Organismal Systems, Environmental Information Processing) revealed that *NUDT21* knockdown was associated with the most pronounced changes in genes belonging to Human Diseases, Environmental Information Processing, and Organismal Systems (Figures 5C-5F).

Bubble diagrams of the top 20 enriched KEGG pathways (screened for pathways containing > 2 DEGs, ranked by -log10 p-value) are shown. Bubble size represents the number of DEGs, and color indicates enrichment significance (blue-white-yellow-red, with red representing smaller p-values). In MUTZ-1, pathways such as parathyroid hormone synthesis and cAMP signaling were among the most enriched (Figures 5G, 5H). In SKM-1, pathways including cytokine-cytokine receptor interaction and cancer pathways showed prominent enrichment (Figures 5I, 5J).

**Predictive analysis of transcription factors and their target genes.** Comparison of transcription factor (TF) distributions between all genes and DEGs identified TF families with notable differences in ratio. The bar plot shows TF families on the x-axis and gene counts on the y-axis. Dark blue represents all genes regulated by a TF family; orange indicates DEGs regulated by the family; pink shows up-regulated DEGs; light blue shows down regulated DEGs (Figures 6A, 6B).

Target genes of differential TF families were extracted based on TF-target relationships. The *zf- C2H2, STAT, bHLH, ETS*, and Forkhead families showed significant differences in both MDS cell lines, suggesting that *NUDT21* may function through these TF families (Figures 6C, 6D).

We also mapped the top 10 TFs with the largest |log2FC| to their top 10 target genes (largest |log2FC|). The graph displays (from left to right) TF family, differential TF, and differential target gene, with connecting lines indicating relationships (Figures 6E, 6F). Results indicate strong associations between *NUDT21* and *TBX21*, *STAT1*, and *MAF*, which are linked to immune response and cancer, implying that *NUDT21* may participate in MDS transformation via these TFs.

**Protein- protein interaction network analysis (PPI).** A circular PPI network of the top 30 interactions was generated using the STRING database (https://cn.string-db.org/). Red nodes denote

Accepted manuscript

up-regulated DEGs, blue nodes denote down- regulated DEGs, and node size reflects the number of interactions (Figures 6G, 6H).

**Relationship between *NUDT21* expression and clinicopathological features in MDS.** Bone marrow samples were collected from de novo MDS patients (2023-2024) at the Second Hospital of Shanxi Medical University. *NUDT21* expression was detected, and IPSS-M scores were calculated using the online tool (https://www.mds-risk-model.com/). No significant differences were found in age, sex, laboratory parameters, or IPSS-M score between the $NUDT21^{high}$ and $NUDT21^{low}$ groups.

Patients were categorized into transformed sAML, MDS, and normal control (n=8) groups. qRT-PCR showed that *NUDT21* expression was significantly higher in the sAML group than in the MDS and normal groups (p=0.005 for MDS, p=0.005 for normal). Expression in the MDS group was slightly higher than in normal controls, but the difference was not statistically significant (p=0.087) (Figure 7A). Notably, high *NUDT21* expression was associated with a significantly increased sAML transformation rate (50% vs. 15%; OR=5.67, 95% CI: 1.28-25.10; p=0.018). To determine whether *NUDT21* expression independently predicted sAML transformation, we performed multivariate logistic regression analysis adjusting for IPSS-M risk category, blast percentage, and cytogenetics. After adjustment for these established prognostic factors, high *NUDT21* expression remained significantly associated with increased risk of sAML transformation (adjusted OR=5.09, 95% CI: 1.084-23.909, p=0.039) (Table 1). These findings suggest that high *NUDT21* expression in bone marrow hematopoietic cells is associated with increased risk of sAML transformation and poor prognosis.

To validate sequencing results, we examined the expression of six selected DEGs-*COL13A1* (p=0.409), *SRPK3* (p=0.015), *SLC6A2* (p=0.085), *DDX60* (p=0.019), *RYR2* (p=0.35), and *XAF1* ($p \leq 0.001$) (Figure 7B) in patient samples stratified by *NUDT21* expression (20 high vs. 20 low). Consistent with sequencing data, *DDX60* ($p \leq 0.001$), *SRPK3*, and *XAF1* levels were significantly higher in the $NUDT21^{high}$ group. Among the six genes validated in patient samples, *DDX60* showed the strongest positive correlation with *NUDT21* expression, further supporting its selection for subsequent mechanistic studies.

The median follow- up duration was 489 days (range: 6-1,791 days). During the follow- up period, 13 of 40 patients (32.5%) progressed to sAML, with a median time-to-transformation of 452 days (range: 33-1,282 days). Patients who did not experience the event of interest (AML transformation or death) were censored at the time of last follow- up. Overall survival was defined as the time from MDS diagnosis to death from any cause or last follow- up.

Accepted manuscript

To evaluate the prognostic impact of *NUDT21* expression on patient survival, we performed Kaplan-Meier survival analysis using the 40 MDS patients stratified by *NUDT21* expression levels (high vs. low, based on median cutoff). Patients with high *NUDT21* expression exhibited significantly shorter overall survival compared to those with low *NUDT21* expression (p=0.04) (Figure 7C). These findings indicate that high *NUDT21* expression is associated with poor prognosis in MDS patients.

***NUDT21* promoted the progression of MDS by activating *JAK2/STAT3* pathway through *DDX60*.**

In order to further explore the molecular mechanism of *NUDT21* inducing transformation in MDS patients, *DDX60* in transcriptome sequencing was further analyzed. *DDX60* itself is a potent interferon-stimulated gene because it can amplify innate immune signals during proinflammatory processes. This means that once the initial *DDX60* is induced in small amounts, it activates the *JAK2/STAT3* pathway by inducing the release of more proinflammatory factors, which in turn strongly induces more *DDX60* expression, forming a self-reinforcing positive feedback loop that leads to a sustained low-level inflammatory state that promotes immunosuppression, tumor cell survival, and metastasis. This was consistent with the fact that the expression of pro-inflammatory cytokines IL-6 and IL-17 was significantly decreased after *NUDT21* knockdown. Therefore, the expression levels of tumor-related genes including *DDX60*, *JAK2*, *STAT3* were verified by qRT-PCR in the cell model, and the expression levels of *JAK2*, *STAT3* after phosphorylation were verified by Western blot.

qRT-PCR results showed that the expressions of *COL13A1* (p=0.0044 for MUTZ-1; p=0.0293 for SKM-1), *SRPK3* (p=0.0109 for MUTZ-1; p=0.0296 for SKM-1), *SLC6A2* (p=0.048 for MUTZ-1; p=0.0183 for SKM-1), *DDX60* (p=0.0385 for MUTZ-1; $p \leq 0.001$ for SKM-1), *XAF1* ($p \leq 0.001$ for MUTZ-1; p=0.008 for SKM-1) and *RYR2* (p=0.005 for MUTZ-1; p=0.0289 for SKM-1) were significantly decreased in SKM-1 and MUTZ-1 cells after *NUDT21* knockdown. Meanwhile, *JAK2* expression was significantly decreased in SKM-1 cells (p=0.008), but not in MUTZ-1 cells (p=*0.2986*). *STAT3* expression was significantly decreased in both cell types ($p \leq 0.001$ for MUTZ-1; p=0.0041 for SKM-1) (Figure 7D). Western blot results showed that the expression of JAK2 (p=0.0049 for MUTZ-1; p=0.0084 for SKM-1) and STAT3 ($p \leq 0.001$ for MUTZ-1; $p \leq 0.001$ for SKM-1) was significantly decreased after NUDT21 knockdown in SKM-1 and MUTZ-1 cells, and the expression of JAK2 ($p \leq 0.001$ for MUTZ-1; $p \leq 0.001$ for SKM-1) and STAT3 ($p \leq 0.001$ for MUTZ-1; $p \leq 0.001$ for SKM-1) was also significantly decreased after phosphorylation (Figure 7E).

**Rescue experiments.** To validate APA changes of *DDX60*, we designed two qPCR primer sets targeting the CDS (total mRNA) and distal 3‘UTR (long isoform), calculated the 3’UTR/CDS ratio,

Accepted manuscript

and normalized to control. *NUDT21* knockdown significantly reduced this ratio, confirming a shift from distal to proximal poly(A) site usage ($p \leq 0.001$ for MUTZ-1; $p=0.0046$ for SKM-1) (Figure 8A). To verify the mechanistic role of *DDX60* in *NUDT21*-mediated effects, we performed rescue experiments by overexpressing *DDX60* in *NUDT21* knockdown MDS cells. qRT-PCR results confirmed successful construction of the cell models. To further verify the involvement of the *JAK2/STAT3* pathway, we used the *STAT3* inhibitor Stattic and determined the treatment concentration based on its IC50 in SKM-1 and MUTZ-1 cells (Figure 8B). qRT-PCR and Western blot results showed that the expression of *JAK2* ($p=0.0132$ for MUTZ-1; $p \leq 0.001$ for SKM-1), *STAT3* ($p=0.0013$ for MUTZ-1; $p=0.0011$ for SKM-1) and their phosphorylated forms were significantly up-regulated in SKM-1 and MUTZ-1 cells after restoring *DDX60* expression. However, the expression of *STAT3* and its phosphorylated form was significantly decreased after the addition of Stattic ($p \leq 0.001$ for MUTZ-1; $p=0.0015$ for SKM-1) (Figures 8B, 8C). Functional assays demonstrated that restoration of *DDX60* expression significantly enhanced the migration and colony formation abilities of SKM-1 ($p \leq 0.001$) and MUTZ-1 cells ($p \leq 0.001$), while reducing apoptosis ($p \leq 0.001$ for MUTZ-1; $p=0.0035$ for SKM-1). In contrast, Stattic treatment decreased cell migration ($p \leq 0.001$ for MUTZ-1; $p \leq 0.001$ for SKM-1) and colony formation and increased apoptosis ($p=0.0042$ for MUTZ-1; $p=0.0025$ for SKM-1) (Figures 8E-8G). Collectively, these rescue experiments demonstrate that *NUDT21* promotes cell proliferation and immunosuppression by upregulating *DDX60* expression and sustaining activation of the *JAK2*/*STAT3* pathway and its phosphorylation, thereby maintaining an inflammatory microenvironment favorable for tumor cell survival.

## Discussion

During mRNA maturation relies on three core post-transcriptional modifications, splicing, capping, and polyadenylation. Polyadenylation triggers pre-mRNA 3′ end cleavage and polyA tail attachment, fundamentally governing mRNA stability and cellular biological functions [35]. APA adds considerable complexity to the transcriptome by affecting RNA localization, stability, translation efficiency, and function, thereby playing a significant role in disease development.

*NUDT21* is a highly conserved component of the CFIm, participating in the early stages of eukaryotic pre-mRNA processing. Changes in 3′UTR length or structure can modify mRNA-destabilizing elements, further regulating RNA-binding protein activity, mRNA stability, translation efficiency, and downstream protein expression [36, 37]. Therefore, *NUDT21* acts as a core APA regulator, and its

Accepted manuscript

aberrant expression is closely correlated with the progression of hematological malignancies, including MDS and AML [38, 39].

Our previous study confirmed aberrant *NUDT21* expression in MDS patients, and we successfully constructed MDS cell models for functional exploration [40]. High-risk MDS harbors heterogeneous clonal subsets. Under selective pressure during disease progression, subclones with AML-like phenotypes expand clonally and drive leukemic transformation toward sAML. These observations are consistent with reports that APA dysregulation promotes leukemogenesis, highlighting the importance of *NUDT21* in MDS-sAML progression.

MDS predominantly affects middle-aged and elderly populations, although recent studies suggest a trend toward younger onset. The boundary between MDS and AML remains ambiguous, and the updated WHO 2022 classification further subdivides MDS with increased blasts (MDS-IB) into MDS-IB1 and MDS-IB2 based on blast percentage, with MDS-IB2 exhibiting a high propensity for transformation to sAML.

In this study, we found that *NUDT21* expression in sAML patients was significantly higher than in MDS patients. Transcriptomic analysis further linked *NUDT21* to multiple signaling pathways involved in MDS progression, including TNF, Wnt, and NF-κB signaling pathways. For example, chronic activation of the Wnt pathway is considered fundamental to MDS pathogenesis and a key driver of sAML transformation. Similarly, NF-κB signaling in the bone marrow microenvironment plays a crucial role in MDS progression [41, 42]. Accordingly, *NUDT21*-high subclones with AML-like phenotypes obtain proliferative advantages via activation of the above signaling pathways during MDS progression, outcompeting other subclones and driving phenotypic transformation into sAML. This hypothesis provides a new perspective on the heterogeneity and transformation mechanisms of MDS.

As an important APA regulator, *NUDT21* has been implicated in the pathogenesis of various diseases, including MDS and multiple cancers. *DDX60*, as an important interferon-stimulated gene, is aberrantly up-regulated in the inflammatory microenvironment of MDS, leading to the increased expression of pro-inflammatory cytokines such as IL-6 and IL-17. These inflammatory cytokines are secreted extracellularly and bind to surface receptors on MDS cells via autocrine or paracrine pathways, thereby activating the canonical *JAK2/STAT3* signaling cascade. This forms a positive feedback loop that sustains chronic inflammation, induces immunosuppression, and enhances the survival of malignant MDS cells.

Our experimental results demonstrated that *NUDT21* knockdown produced significant functional effects on both SKM-1 and MUTZ-1 cells. Specifically, *NUDT21* depletion significantly increased

Accepted manuscript

apoptosis and induced G1 phase arrest, accompanied by a reduced proportion of cells in G2 phase. Moreover, *NUDT21* knockdown significantly decreased cell migration ability and impaired colony formation capacity, indicating that *NUDT21* is required for maintaining MDS cell proliferation and clonogenic potential. These phenotypic changes were accompanied by shifts in differentiation marker expression, manifested as upregulation of CD117 and altered CD11b levels. Mechanistically, these functional changes were accompanied by reduced expression of *DDX60* and downregulation of the *JAK2/STAT3* pathway, including decreased phosphorylation levels of JAK2 and STAT3, suggesting that the tumor-promoting role of NUDT21 is mediated at least in part through this signaling axis. Collectively, these findings indicate that *NUDT21* plays a critical role in supporting MDS cell survival, proliferation, and migration capabilities, with the underlying mechanism potentially involving regulation of the *DDX60-JAK2/STAT3* pathway.

To further validate the causal role of the *DDX60-JAK2/STAT3* axis, we performed rescue experiments by overexpressing *DDX60* in *NUDT21*-knockdown MDS cells. *DDX60* overexpression rescues the expression and phosphorylation of *JAK2/STAT3* in NUDT21-knockdown MDS cells, reverses defective cell migration and colony formation, and alleviates cell apoptosis. Conversely, treatment with the *STAT3* inhibitor Stattic abrogated these effects, decreasing migration and colony formation ability and increasing apoptosis. These results confirm that *NUDT21* promotes MDS progression through the *DDX60-JAK2/STAT3* axis. In addition, to directly validate the APA changes of *DDX60*, we performed qRT-PCR using primers targeting the CDS region and the distal 3′UTR region. *NUDT21* knockdown led to a significant reduction in the distal 3′UTR/CDS amplification ratio, confirming 3′UTR shortening of *DDX60* upon *NUDT21* loss. MEME motif analysis further identified conserved UGUA motifs within the DDX60 3′UTR, which are known NUDT21 binding sites, and IGV visualization confirmed reduced read coverage in this region upon NUDT21 knockdown, providing direct sequencing evidence supporting the APA- mediated regulation of DDX60 (Supplementary Figure S1). These findings provide direct experimental evidence linking *NUDT21*-mediated APA dysregulation to *DDX60* 3′UTR shortening and subsequent *JAK2/STAT3* pathway activation. We noted that JAK2 responses differed between the two cell lines: in MUTZ- 1 cells, JAK2 protein decreased without corresponding mRNA changes, whereas SKM- 1 cells showed concordant reductions at both levels. This likely reflects cell- line- specific differences in post- translational regulation, as JAK2 is subject to rapid phosphorylation- dependent ubiquitin- proteasome degradation. Importantly, STAT3 phosphorylation was consistently reduced upon NUDT21 knockdown in both lines, suggesting that

Accepted manuscript

STAT3 is the primary downstream effector of NUDT21- mediated APA dysregulation, while JAK2 may serve as a context- dependent amplifier of STAT3 signaling.

Therefore, therapeutic strategies aimed at modulating *NUDT21* expression warrant further investigation. Multivariate logistic regression analysis revealed that high *NUDT21* expression remained significantly associated with sAML transformation after adjusting for IPSS-M risk, blast percentage, and cytogenetics. Consistent with this, Kaplan-Meier analysis revealed that patients with high *NUDT21* expression had significantly shorter overall survival. These findings suggest that *NUDT21* may serve as an independent predictor of leukemic transformation and poor prognosis in MDS.

Several limitations should be acknowledged. Our cohort was derived from a single center with a relatively small sample size, limited follow-up time, and lacked an independent external validation cohort. Consequently, the adjusted odds ratio for AML transformation (OR=5.09) had a wide confidence interval (1.084-23.909), reflecting substantial uncertainty in the effect estimate, which is expected given the limited sample size and biological heterogeneity among MDS patients. Nevertheless, the lower bound of the CI remains above 1, and the survival analysis yielded a consistent significant result (log- rank p=0.039), suggesting that the association is unlikely to be driven by chance alone. Furthermore, although we validated the APA changes of *DDX60* through 3′UTR targeted qRT-PCR, further validation using 3′RACE or long-read sequencing would strengthen our conclusions. We also acknowledge the lack of in vivo validation, although we have initiated preliminary animal experiments. Future work will focus on establishing xenograft mouse models to assess tumor progression and validate the *DDX60-JAK2/STAT3* axis as a therapeutic target. These studies are essential for translating our findings into clinical applications.

In summary, we report specific gene profiles in MDS cell lines after *NUDT21* knockdown. *NUDT21* expression was associated with disease progression in MDS patients, and its downstream target *DDX60* may contribute to leukemic transformation potentially through the *JAK2/STAT3* pathway. These findings suggest that *NUDT21* may serve as a potential prognostic biomarker for MDS progression, warranting further validation in larger independent cohorts.

Acknowledgments: We thank the hematology Department of the Second Hospital of Shanxi Medical University for their support in the experiments.

The study was supported by the National Natural Science Foundation of China (NO. NO.81670126;81500104), The Second Hospital of Shanxi Medical University Doctoral Fund Project (202201-2), the Second Hospital of Shanxi Medical University Scientific and Technological Activities

Accepted manuscript

of Overseas Students Merit Grant Project (2020041), the Special Project of Wu Jieping Medical Foundation (320.6750.2022-8-4).

**Supplementary data are available in the online version of the paper.**

**References**

[1] KHOURY JD, SOLARY E, ABLA O, AKKARI Y, ALAGGIO R et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia 2022; 36: 1703-1719. https://doi.org/10.1038/s41375-022-01613-1

[2] ARBER DA, ORAZI A, HASSERJIAN R, THIELE J, BOROWITZ MJ et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood 2016; 127: 2391-2405. https://doi.org/10.1182/blood-2016-03-643544

[3] CHEN BY, SONG J, HU CL, CHEN SB, ZHANG Q et al. SETD2 deficiency accelerates MDS-associated leukemogenesis via S100a9 in NHD13 mice and predicts poor prognosis in MDS. Blood 2020; 135: 2271-2285. https://doi.org/10.1182/blood.2019001963

[4] GHOBRIAL IM, DETAPPE A, ANDERSON KC, STEENSMA DP. The bone-marrow niche in MDS and MGUS: implications for AML and MM. Nat Rev Clin Oncol 2018; 15: 219-233. https://doi.org/10.1038/nrclinonc.2017.197

[5] HERON M, DOVERN E, BAKKER-JONGES LE, POSTHUMA EFM, BROUWER RE et al. Translating the MDS flow cytometric score into clinical practice. Cytometry B Clin Cytom 2015; 88: 207-209. https://doi.org/10.1002/cyto.b.21208

[6] HOSPITAL MA, VEY N. Myelodysplastic Syndromes: How to Recognize Risk and Avoid Acute Myeloid Leukemia Transformation. Curr Oncol Rep 2020; 22: 4. https://doi.org/10.1007/s11912-020-0869-0

[7] MENSSEN AJ, WALTER MJ. Genetics of progression from MDS to secondary leukemia. Blood 2020; 136: 50-60. https://doi.org/10.1182/blood.2019000942

[8] TANAKA TN, BEJAR R. MDS overlap disorders and diagnostic boundaries. Blood. 2019; 133: 1086-1095. https://doi.org/10.1182/blood-2018-10-844670

[9] SHIOZAWA Y, MALCOVATI L, GALLÌ A, PELLAGATTI A, KARIMI M et al. Gene expression and risk of leukemic transformation in myelodysplasia. Blood 2017; 130: 2642-2653. https://doi.org/10.1182/blood-2017-05-783050

[10] GARCIA-MANERO G. Myelodysplastic Syndromes: 2026 Update on Diagnosis, Risk-Stratification and Management. Am J Hematol 2026; 101: 2393-2411. https://doi.org/10.1002/ajh.70405

[11] GRUBER AJ, ZAVOLAN M. Alternative cleavage and polyadenylation in health and disease. Nat Rev Genet 2019; 20: 599-614. https://doi.org/10.1038/s41576-019-0145-z

[12] ALCOTT CE, YALAMANCHILI HK, JI P, VAN DER HEIJDEN ME, SALTZMAN A et al. Partial loss of CFIm25 causes learning deficits and aberrant neuronal alternative polyadenylation. Elife 2020; 9: e50895. https://doi.org/10.7554/eLife.50895

[13] BRUMBAUGH J, DI STEFANO B, WANG X, BORKENT M, FOROUZMAND E et al. Nudt21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling. Cell 2018; 172: 106-120.e21. https://doi.org/10.1016/j.cell.2017.11.023

Accepted manuscript

[14] BURRI D, ZAVOLAN M. Shortening of 3' UTRs in most cell types composing tumor tissues implicates alternative polyadenylation in protein metabolism. RNA 2021; 27: 1459-1470. https://doi.org/10.1261/rna.078886.121
[15] DHARMALINGAM P, MAHALINGAM R, YALAMANCHILI HK, WENG T, KARMOUTY-QUINTANA H et al. Emerging roles of alternative cleavage and polyadenylation (APA) in human disease. J Cell Physiol 2022; 237: 149-160. https://doi.org/10.1002/jcp.30549
[16] GARAULET DL, ZHANG B, WEI L, LI E, LAI EC. miRNAs and Neural Alternative Polyadenylation Specify the Virgin Behavioral State. Dev Cell 2020; 54: 410-423.e4. https://doi.org/10.1016/j.devcel.2020.06.004
[17] JAFARI NAJAF ABADI MH, SHAFABAKHSH R, ASEMI Z, MIRZAEI HR, SAHEBNASAGH R et al. CFIm25 and alternative polyadenylation: Conflicting roles in cancer. Cancer Lett 2019; 459: 112-121. https://doi.org/10.1016/j.canlet.2019.114430
[18] KUBO T, WADA T, YAMAGUCHI Y, SHIMIZU A, HANDA H. Knock-down of 25 kDa subunit of cleavage factor Im in Hela cells alters alternative polyadenylation within 3'-UTRs. Nucleic Acids Res 2006; 34: 6264-6271. https://doi.org/10.1093/nar/gkl794
[19] LEE SH, SINGH I, TISDALE S, ABDEL-WAHAB O, LESLIE CS et al. Widespread intronic polyadenylation inactivates tumor suppressor genes in leukemia. Nature 2018; 561: 127-131. https://doi.org/10.1038/s41586-018-0465-8
[20] MASAMHA CP, XIA Z, YANG J, ALBRECHT TR, LI M et al. CFIm25 links alternative polyadenylation to glioblastoma tumor suppression. Nature 2014; 510: 412-416. https://doi.org/10.1038/nature13261
[21] Ran Y, Huang S, Shi J, Feng Q, Deng Y et al. CFIm25 regulates human stem cell function independently of its role in mRNA alternative polyadenylation. RNA Biol 2022; 19: 686-702. https://doi.org/10.1080/15476286.2022.2071025
[22] ROUNDTREE IA, EVANS ME, PAN T, HE C. Dynamic RNA Modifications in Gene Expression Regulation. Cell 2017; 169: 1187-1200. https://doi.org/10.1016/j.cell.2017.05.045
[23] SOMMERKAMP P, TRUMPP A. Driving differentiation: targeting APA in AML. Blood 2022; 139: 317-319. https://doi.org/10.1182/blood.2021013814
[24] SUBRAMANIAN A, HALL M, HOU H, MUFTEEV M, YU B et al. Alternative polyadenylation is a determinant of oncogenic Ras function. Sci Adv 2021; 7: eabh0562. https://doi.org/10.1126/sciadv.abh0562
[25] SUN M, DING J, LI D, YANG G, CHENG Z et al. NUDT21 regulates 3'-UTR length and microRNA-mediated gene silencing in hepatocellular carcinoma. Cancer Lett 2017; 410: 158-168. https://doi.org/10.1016/j.canlet.2017.09.026
[26] TIAN B, MANLEY JL. Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol 2017; 18: 18-30. https://doi.org/10.1038/nrm.2016.116
[27] WENG T, HUANG J, WAGNER EJ, KO J, WU M et al. Downregulation of CFIm25 amplifies dermal fibrosis through alternative polyadenylation. J Exp Med 2020; 217: e20181384. https://doi.org/10.1084/jem.20181384
[28] XIANG Y, YE Y, LOU Y, YANG Y, CAI C et al. Comprehensive Characterization of Alternative Polyadenylation in Human Cancer. J Natl Cancer Inst 2018; 110: 379-389. https://doi.org/10.1093/jnci/djx223
[29] XING Y, CHEN L, GU H, YANG C, ZHAO J et al. Downregulation of NUDT21 contributes to cervical cancer progression through alternative polyadenylation. Oncogene 2021; 40: 2051-2064. https://doi.org/10.1038/s41388-021-01693-w

Accepted manuscript

[30] XIONG M, CHEN L, ZHOU L, DING Y, KAZOBINKA G et al. NUDT21 inhibits bladder cancer progression through ANXA2 and LIMK2 by alternative polyadenylation. Theranostics 2019; 9: 7156-7167. https://doi.org/10.7150/thno.36030
[31] XU Z, PLATIG J, LEE S, BOUEIZ A, CHASE R et al. Cigarette smoking-associated isoform switching and 3′ UTR lengthening via alternative polyadenylation. Genomics 2021; 113: 4184-4195. https://doi.org/10.1016/j.ygeno.2021.11.004
[32] KANEHISA M, FURUMICHI M, SATO Y, MATSUURA Y, ISHIGURO-WATANABE M. KEGG: biological systems database as a model of the real world. Nucleic Acids Res 2025; 53: D672-D677. https://doi.org/10.1093/nar/gkae909
[33] KANEHISA M. Toward understanding the origin and evolution of cellular organisms. Protein Sci 2019; 28: 1947-1951. https://doi.org/10.1002/pro.3715
[34] KANEHISA M, GOTO S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 2000; 28: 27-30. https://doi.org/10.1093/nar/28.1.27
[35] TAN S, LI H, ZHANG W, SHAO Y, LIU Y et al. NUDT21 negatively regulates PSMB2 and CXXC5 by alternative polyadenylation and contributes to hepatocellular carcinoma suppression. Oncogene 2018; 37: 4887-4900. https://doi.org/10.1038/s41388-018-0280-6
[36] GHOSH S, ATAMAN M, BAK M, BÖRSCH A, SCHMIDT A et al. CFIm-mediated alternative polyadenylation remodels cellular signaling and miRNA biogenesis. Nucleic Acids Res 2022; 50: 3096-3114. https://doi.org/10.1093/nar/gkac114
[37] CHU Y, ELROD N, WANG C, LI L, CHEN T et al. Nudt21 regulates the alternative polyadenylation of Pak1 and is predictive in the prognosis of glioblastoma patients. Oncogene 2019; 38: 4154-4168. https://doi.org/10.1038/s41388-019-0714-9
[38] MASAMHA CP. The emerging roles of CFIm25 (NUDT21/CPSF5) in human biology and disease. Wiley Interdiscip Rev RNA 2023; 14: e1757. https://doi.org/10.1002/wrna.1757
[39] Sun M, Ding J, Li D, Yang G, Cheng Z et al. NUDT21 regulates 3′-UTR length and microRNA-mediated gene silencing in hepatocellular carcinoma. Cancer Lett 2017; 410: 158-168. https://doi.org/10.1016/j.canlet.2017.09.026
[40] LI S, REN F, LIU XL, ZHANG HY, XU ZF et al. A new pathway for the transformation of AML in MDS: APA mechanism regulated by NUDT21 and RUNX1. Neoplasma 2023; 70: 350-360. https://doi.org/10.4149/neo_2023_230115N27
[41] KRIMPENFORT RA, NETHE M. Canonical Wnt: a safeguard and threat for erythropoiesis. Blood Advances 2021; 5: 3726-3735. https://doi.org/10.1182/bloodadvances.2021004845
[42] CENERI E, DE STEFANO A, CASALIN I, FINELLI C, CURTI A et al. Signaling pathways and bone marrow microenvironment in myelodysplastic neoplasms. Adv Biol Regul 2025; 95: 101071. https://doi.org/10.1016/j.jbior.2024.101071

## Figure Legends

**Figure 1.** *NUDT21* knockdown in MDS cell lines and its effects on cellular functions. A) qRT-PCR and B) Western blot confirming NUDT21 knockdown in SKM-1 and MUTZ-1 cells; densitometric quantification of NUDT21/β-actin ratio is shown below. C) Apoptosis and D) cell cycle distribution analyzed by flow cytometry. E) Expression of CD117 and CD11b detected by flow cytometry. Data are presented as mean±SD from n=3 independent biological replicates. Statistical significance was

Accepted manuscript

687 determined by two- tailed Student's t- test. $^{ns}p > 0.05$, *p < 0.05, **p < 0.01, ***p < 0.001. Exact p-
688 values are provided in the main text where the corresponding results are described.
689
690 **Figure 2.** Effects of *NUDT21* knockdown on cytokine secretion, migration, and colony formation in
691 MDS cell lines. A) Cytokine levels in culture supernatants of SKM-1 and MUTZ-1 cells detected by
692 flow cytometry. B) Transwell migration assay and C) colony formation assay in methylcellulose semi-
693 solid medium. Data are presented as mean±SD from n=3 independent biological replicates. Statistical
694 significance was determined by two- tailed Student's t- test. $^{ns}p > 0.05$, *p < 0.05, **p < 0.01, ***p <
695 0.001. Exact p-values are provided in the main text where the corresponding results are described.
696
697 **Figure 3.** Transcriptomic profiling of *NUDT21* knockdown MDS cell lines. A, B) Volcano plots and
698 heatmaps of differentially expressed genes (|fold change| ≥ 2, FDR < 0.05) in MUTZ-1 (A) and SKM-1
699 (B). Red and blue dots/boxes indicate up-regulated and down-regulated genes, respectively; grey
700 indicates non-significant genes. C) Volcano plots showing genes with altered 3′UTR length upon
701 *NUDT21* knockdown in both cell lines (red: longer 3′UTRs; blue: shorter 3′UTRs). Each condition was
702 analyzed using n=3 independent biological replicates.
703
704 **Figure 4.** GO enrichment analysis of differentially expressed genes in *NUDT21* knockdown MDS cell
705 lines. A, B) Chord plots showing the top 10 enriched GO terms in MUTZ-1 (A) and SKM-1 (B), with
706 connecting lines indicating gene-term relationships; outer heatmaps display log2 fold change values. C,
707 D) Circle plots showing the distribution of all genes and DEGs across GO categories in MUTZ-1 (C)
708 and SKM-1 (D); colors indicate up- vs. Down-regulation as shown in the legends. E-H) Bar plots
709 comparing GO enrichment between all genes and DEGs: (E, G) all genes; (F, H) up and down-
710 regulated DEGs. Data are shown for MUTZ-1 (E, F) and SKM-1 (G, H). Each condition was analyzed
711 using n=3 independent biological replicates.
712
713 **Figure 5.** KEGG pathway enrichment analysis of differentially expressed genes in *NUDT21*
714 knockdown MDS cell lines. A, B) Circle plots showing the top 20 enriched KEGG pathways in
715 MUTZ-1 (A) and SKM-1 (B); gene count, Q-value, up-/down-regulation ratio, and RichFactor are
716 displayed as indicated in the legend. C-F) Bar plots showing pathway enrichment for all genes (C, E)
717 and up-/down-regulated DEGs (D, F) in MUTZ-1 (C, D) and SKM-1 (E, F); numbers on the right
718 indicate gene counts. G-J) Bubble plots of the top 20 enriched KEGG pathways for down-regulated (G,



Accepted manuscript

I) and up-regulated (H, J) genes in MUTZ-1 (G, H) and SKM-1 (I, J); bubble size represents gene count, and color indicates significance level. Each condition was analyzed using n=3 independent biological replicates.

**Figure 6.** Transcription factor family distribution and protein-protein interaction network analysis of DEGs in *NUDT21* knockdown MDS cell lines. A, B) Bar plots showing transcription factor (TF) family distributions among all genes and DEGs in MUTZ-1 (A) and SKM-1 (B); dark blue indicates all genes regulated by a TF family, orange indicates DEGs, pink indicates up-regulated DEGs, and light blue indicates down-regulated DEGs. C, D) Bar plots showing the numbers of up and down-regulated target genes of differential TF families in MUTZ-1 (C) and SKM-1 (D). E, F) Mulberry plots showing the relationships between TF families, differential TFs, and their target genes in MUTZ-1 (E) and SKM-1 (F). G, H) Circular protein-protein interaction (PPI) networks of the top 30 interacting DEGs in MUTZ-1 (G) and SKM-1 (H); red nodes indicate up-regulated genes, blue nodes indicate down-regulated genes, and node size reflects the number of interactions. Each condition was analyzed using n=3 independent biological replicates.

**Figure 7.** Validation of target genes and JAK2-STAT3 pathway in MDS patients and cell models. A) *NUDT21* expression in bone marrow samples from healthy controls (n = 8), MDS patients (n = 27), and sAML patients (n=13) detected by qRT-PCR. B) Expression of six selected DEGs (*COL13A1*, *SRPK3*, *SLC6A2*, *DDX60*, *RYR2*, *XAF1*) in *NUDT21*$^{high}$ versus *NUDT21*$^{low}$ MDS patient groups. C) Overall survival curves for *NUDT21*$^{high}$ (n=20) and *NUDT21*$^{low}$ (n=20) groups. D) qRT-PCR validation of DEGs and JAK2/STAT3 expression after *NUDT21* knockdown in SKM-1 and MUTZ-1 cells. E) Western blot analysis of phosphorylated and total JAK2/STAT3 proteins in SKM-1 and MUTZ-1 cells; densitometric quantification is shown below. For cell- based assays, data are presented as mean±SD from n=3 independent biological replicates; statistical significance was determined by two- tailed Student's t- test. For clinical sample comparisons, one- way ANOVA followed by Tukey's post- hoc test; the log- rank test was used for survival analysis. $^{ns}p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Exact p-values are provided in the main text where the corresponding results are described.

**Figure 8. Rescue experiment.** A) qRT-PCR validation of *DDX60* 3′UTR shortening. B) The IC50 concentrations of MUTZ-1 and SKM-1 were measured by CCK-8 assay. C) qRT-PCR showed that the cell model was successfully constructed and Stattic could significantly inhibit the expression of STAT3

Accepted manuscript

D) The phosphorylated and total JAK2/STAT3 proteins in SKM-1 and MUTZ-1 cell models and after Stattic addition were analyzed by Western blot. E) Transwell migration assay and F) colony formation assay in methylcellulose semi-solid medium. G) Cell apoptosis was analyzed by flow cytometry. Data are presented as mean±SD from n=3 independent biological replicates. Statistical significance was determined by one- way ANOVA followed by Tukey's post- hoc test. $^{ns}p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Exact p-values are provided in the main text where the corresponding results are described.

Accepted manuscript

**Table 1.** Relationship between *NUDT21* expression levels and clinicopathological features in myelodysplastic neoplasms Expression level. ($^{ns}p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

| Characterization | *NUDT21* expression level | | | | |
|---|---|---|---|---|---|
| | Number (N=40, 100%) | Low expression (N=20, 50%) | High expression (N=20, 50%) | p-value | Adjusted OR (95% CI) |
| **Genders，** | | | | 1.000 | |
| **N (%)** | | | | | |
| Male | 21 (52.5) | 11 (55) | 10 (50) | | |
| Female | 19 (47.5) | 9 (45) | 10 (50) | | |
| **Age (years), N (%)** | | | | 0.527 | |
| ≥ 60 | 19 (47.5) | 11 (55) | 8 (40) | | |
| < 60 | 21 (52.5) | 9 (45) | 12 (60) | | |
| **Laboratory tests, average (range)** | | | | | |
| Blast (%) | 6.65 (0.3-19) | 5.55 (0.6-19) | 7.75 (0.3-19) | 0.221 | |
| Haemoglobin (g/l) | 75.08 (39-121) | 73.85 (52-121) | 76.3 (39-117) | 0.398 | |
| Platelet ($\times 10^9$/l) | 111.78 (4-922) | 126.2 (4-922) | 97.35 (6-608) | 0.583 | |
| Absolute neutrophil count ($\times 10^9$/l) | 1.58 (0.1-17.78) | 1.27 (0.11-2.5) | 1.88 (0.1-17.78) | 0.265 | |
| **IPSS-M** | | | | 0.239 | |
| Very Low | 2 (5) | 1 (5) | 1 (5) | | |
| Low | 6 (15) | 3 (15) | 3 (15) | | |
| Moderate-Low | 5 (12.5) | 4 (20) | 1 (5) | | |
| Moderate-High | 5 (12.5) | 4 (20) | 1 (5) | | |
| High | 9 (22.5) | 3 (15) | 6 (30) | | |
| Very high | 13 (32.5) | 5 (25) | 8 (40) | | |
| **Whether or not to transform AML** | | | | 0.018 | |
| Yes | 13 (32.5) | 3 (15) | 10 (50) | | |
| No | 27 (67.5) | 17 (85) | 10 (50) | | |
| Multivariate analysis | | | | 0.039 | 5.09 (1.084-23.909) |

Accepted manuscript

Fig. 1 [Download full resolution image]

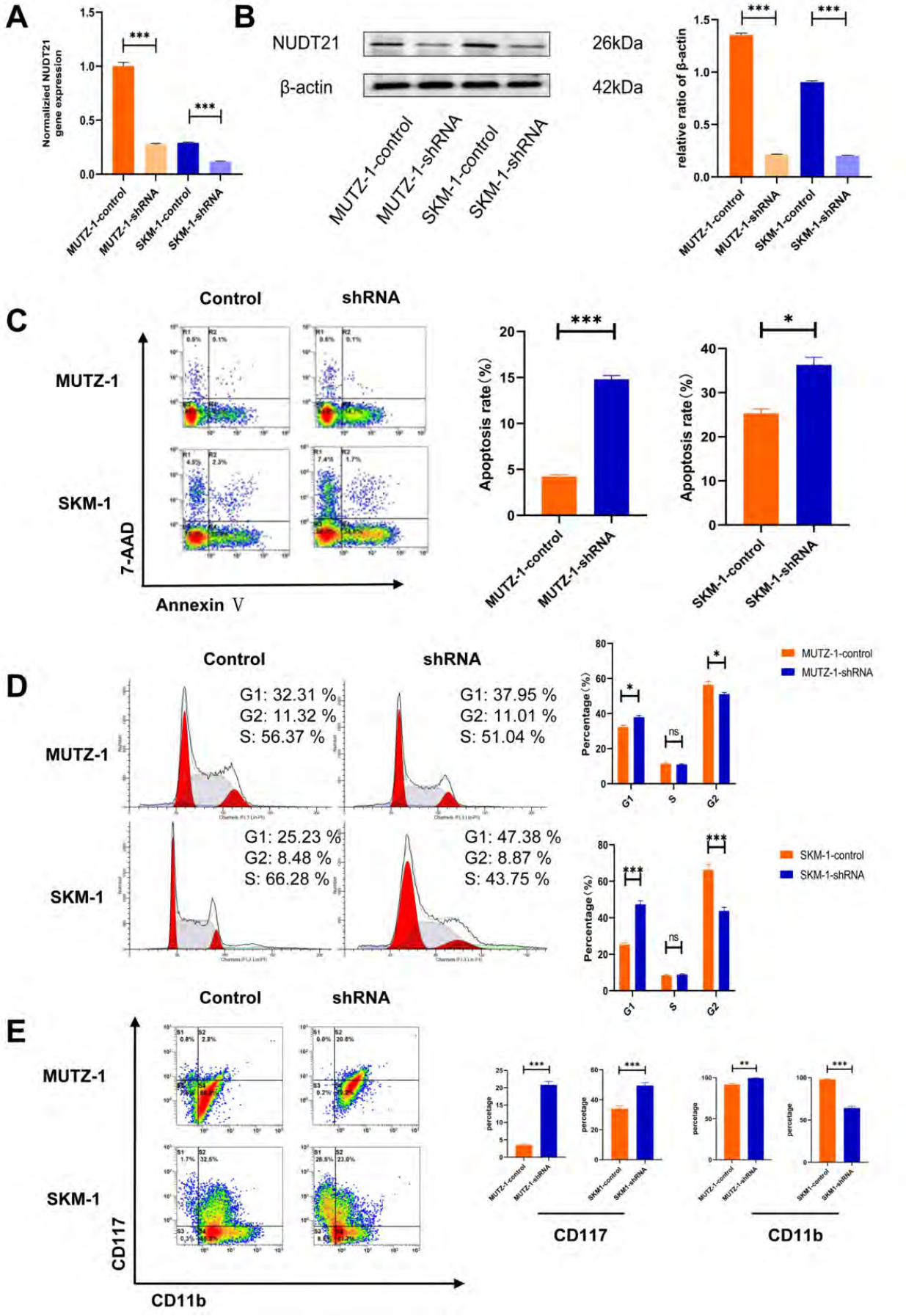


Fig. 2 Download full resolution image

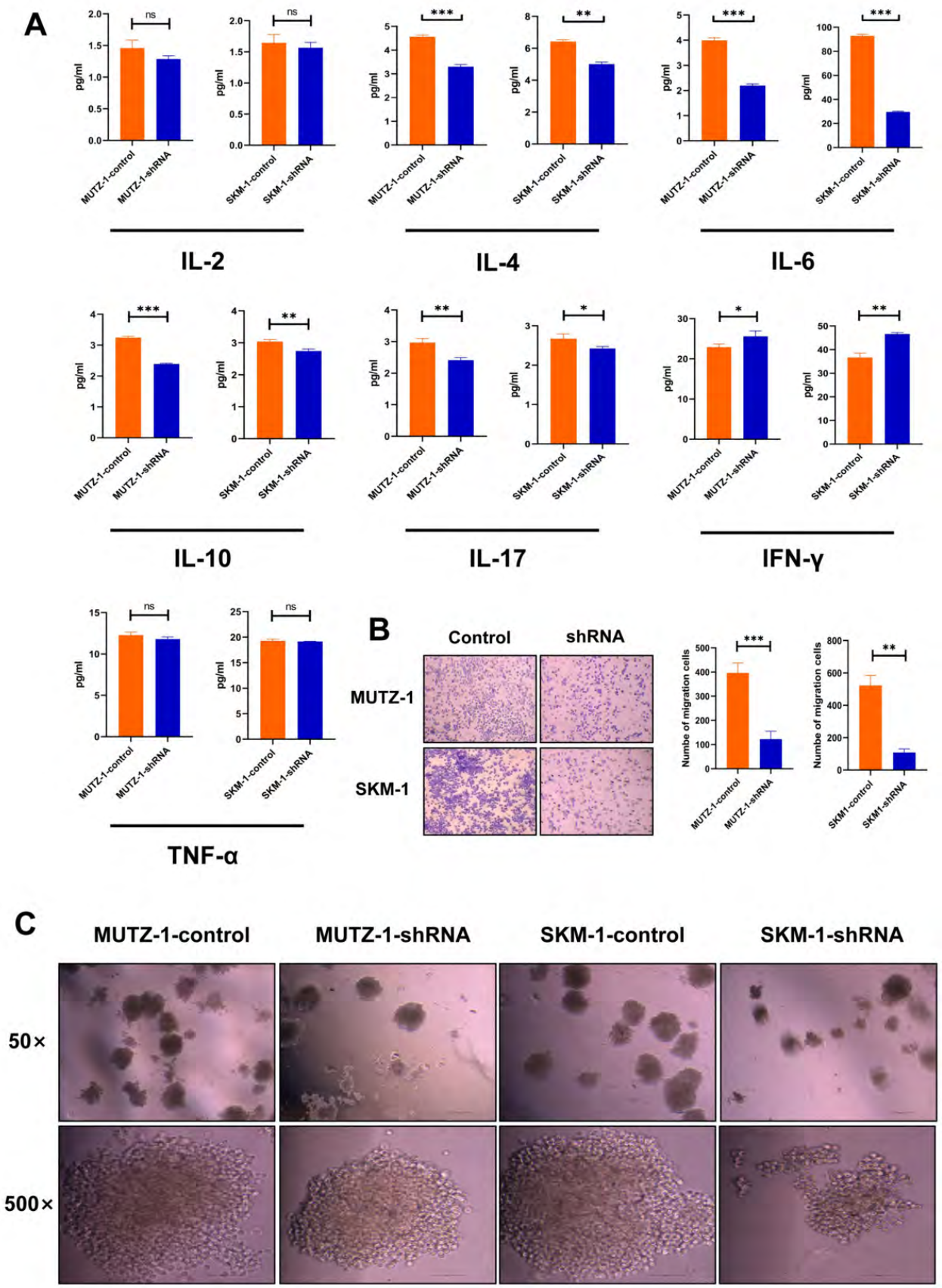


Fig. 3 [Download full resolution image]

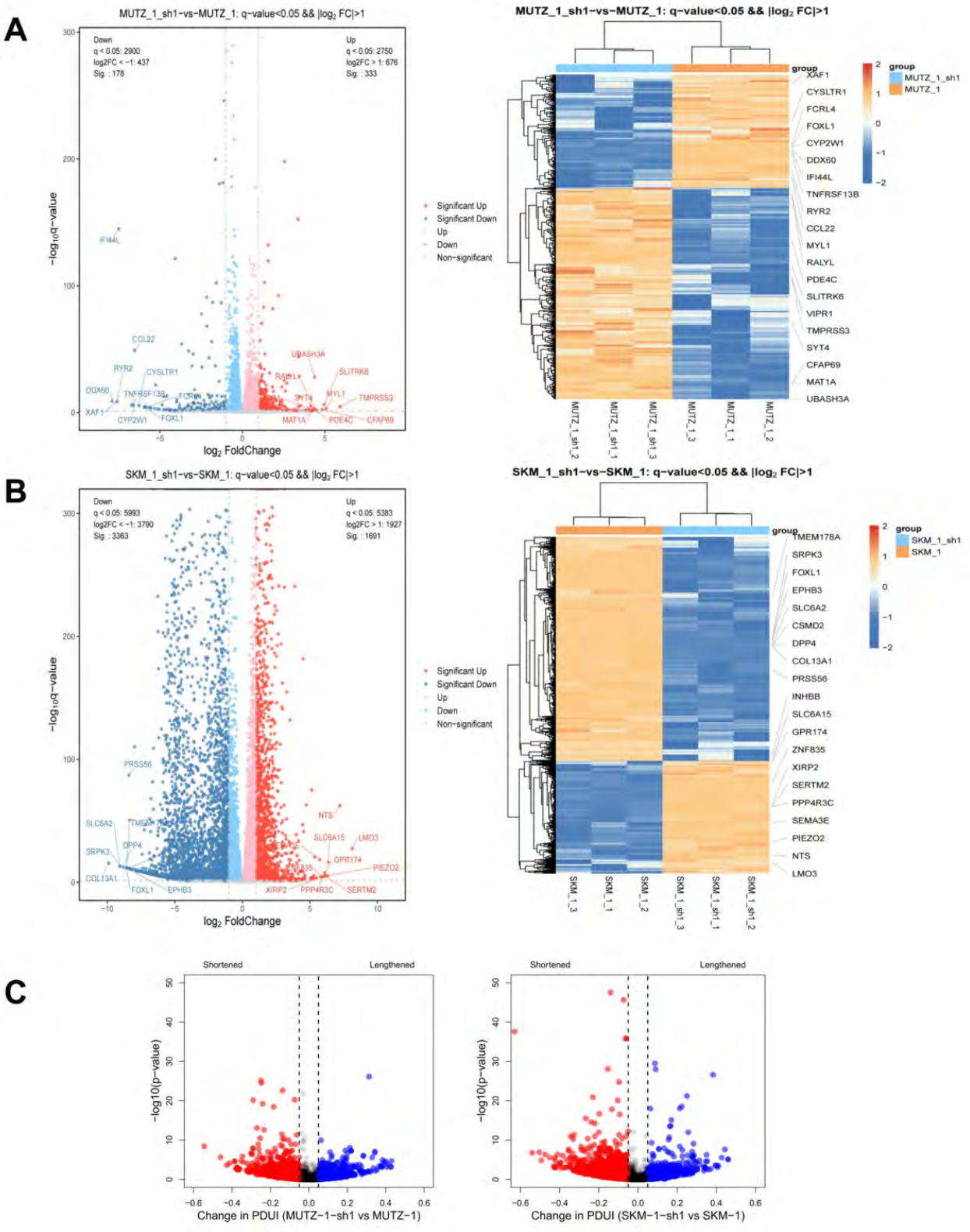


Fig. 4 [Download full resolution image]

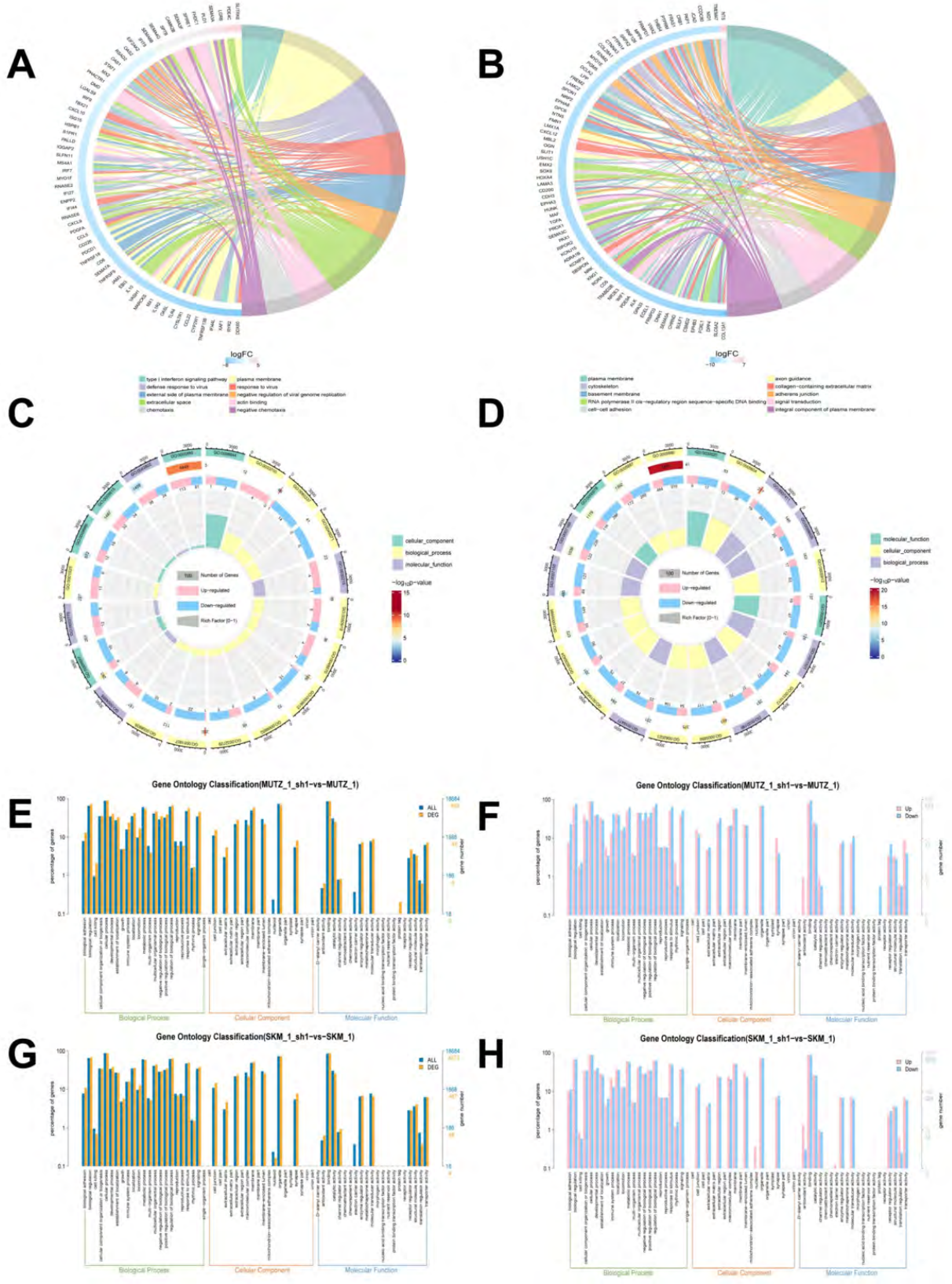


Fig. 5 [Download full resolution image]

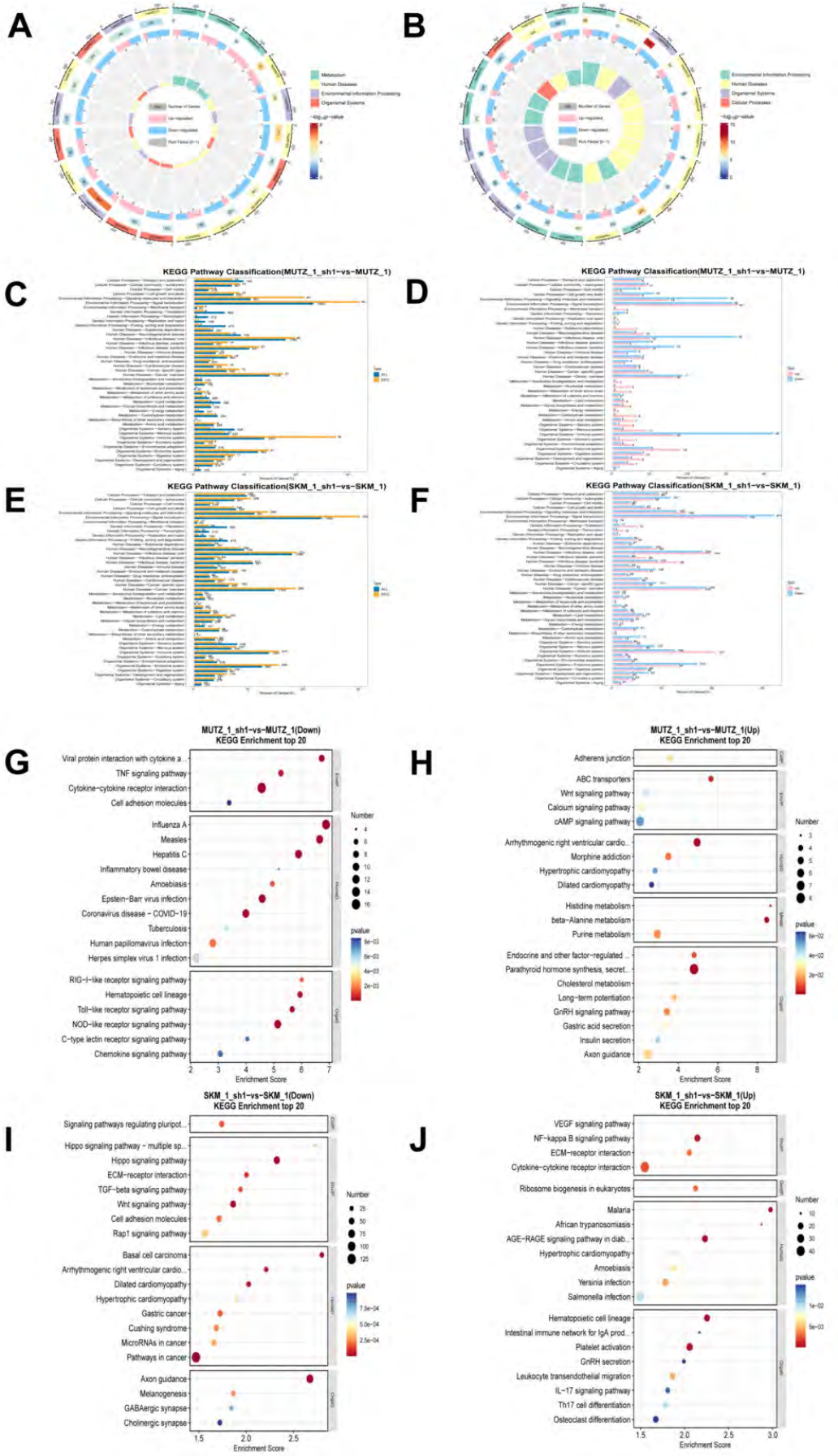


Fig. 6 [Download full resolution image]

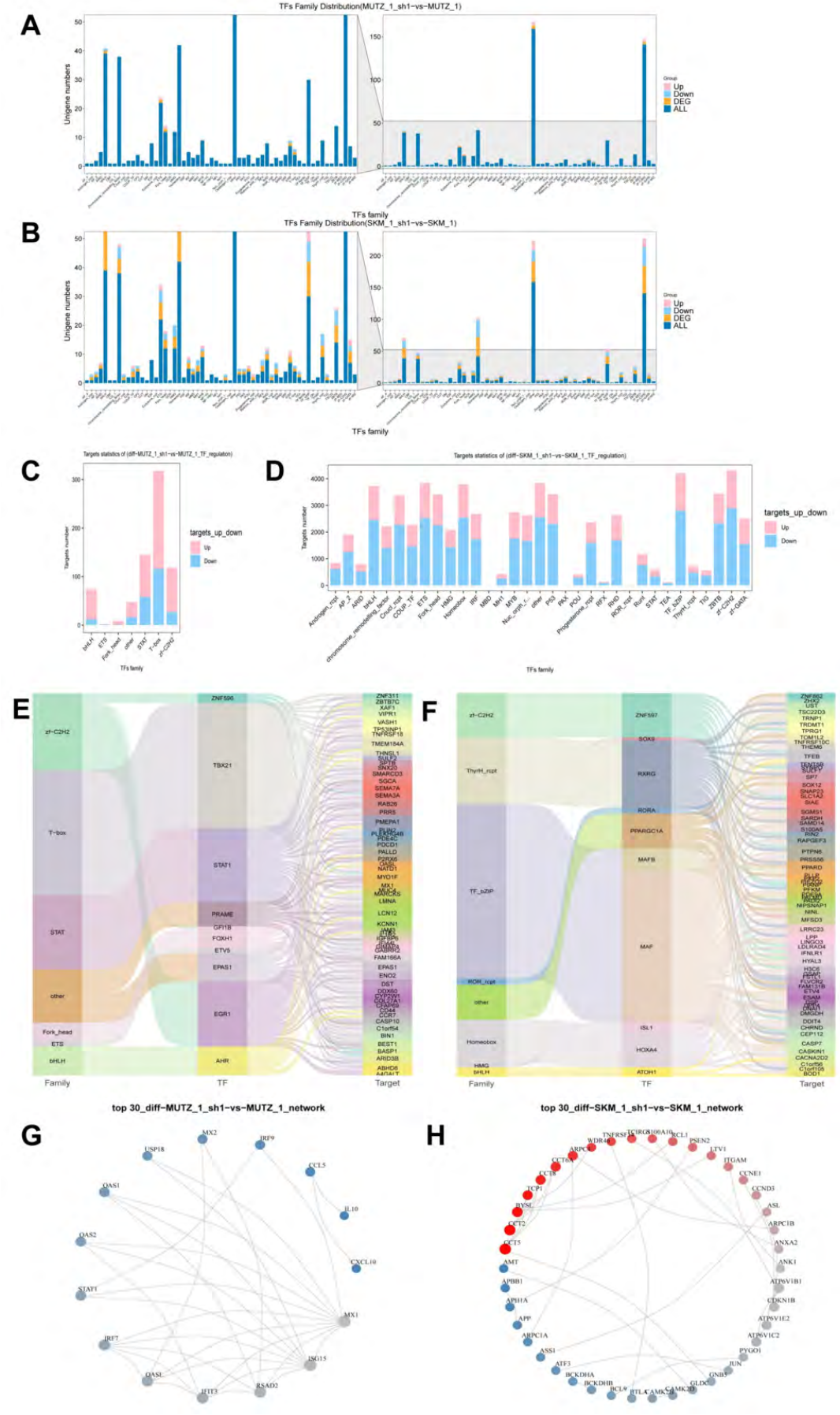


Fig. 7 [Download full resolution image]

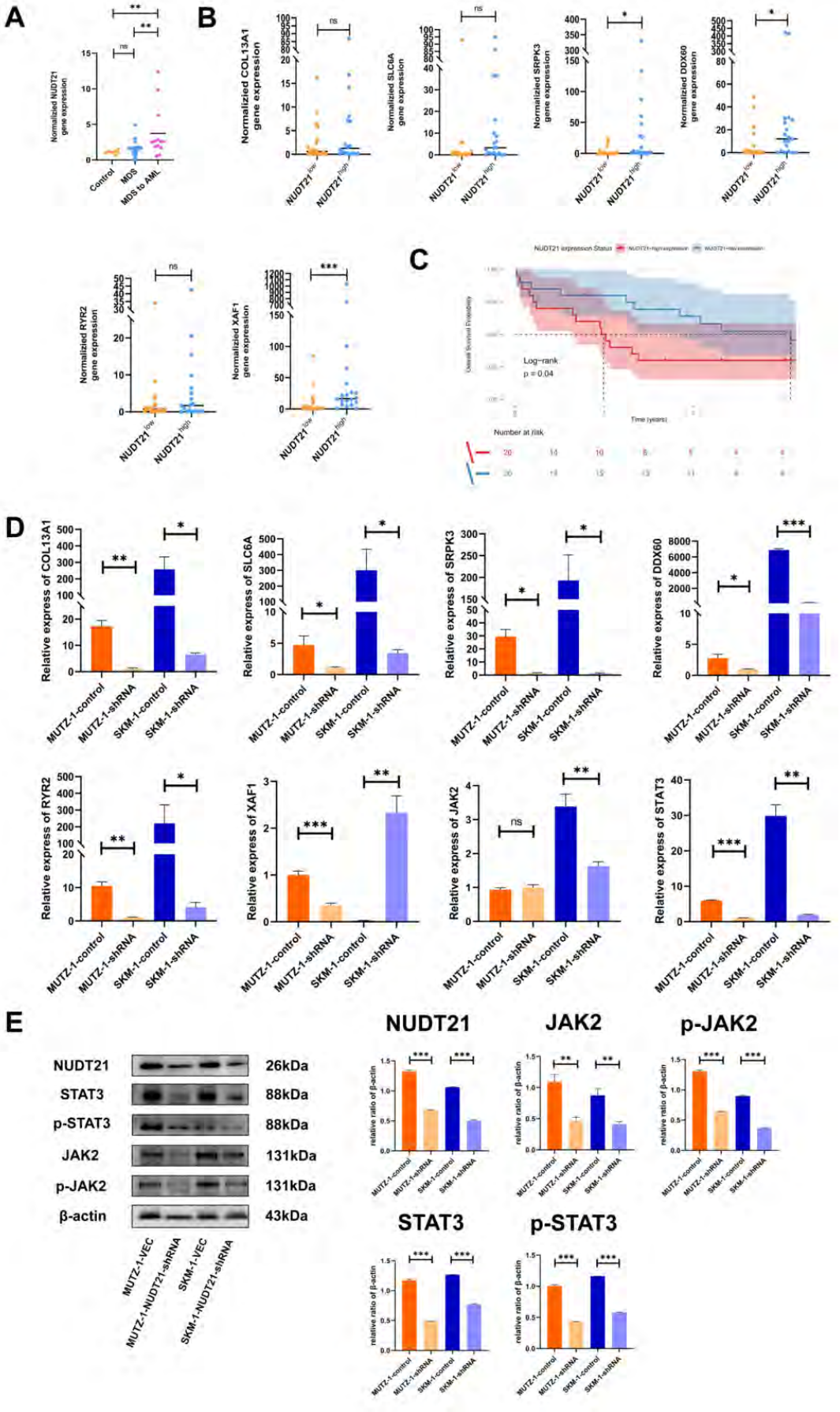


Fig. 8 [Download full resolution image]

