Supplemental information (Methods, Tables and Figures)

Supplemental Methods

Cell lines

HeLa, 293T, NB4 and U937 cells were obtained from American Type Culture Collection (ATCC) or DSMZ. HeLa, NB4 and U937 were maintained in RPMI-1640 medium, whereas 293T was maintained in Dulbecco’s Modified Eagle’s Medium (DMEM), both supplemented with 10% fetal bovine serum.

Fusion transcript identification

The Illumina TruSight RNA Pan-Cancer Panel (targeting 1,385 cancer genes) was used to identify the NUP98-JADE2 fusion in the patient. Libraries were sized on QIAxcel Advanced using the QIAxcel DNA screening kit (Qiagen) and quantified by real-time PCR. Sequencing was performed using the MiSeq Reagent Kit v3 at 76 bp paired-end on an Illumina MiSeq system. Data were analyzed by the RNA fusion analysis module (version 1.0.0.351) on Local Run Manager v2.0 (Illumina).

Reverse transcription-polymerase chain reaction (RT-PCR) and quantitative RT-PCR

Total RNA was reverse transcribed into cDNA using the SuperScript III First-Strand Synthesis System (Thermo Fisher Scientific). NUP98-JADE2 fusion transcript was amplified with primers NUP98-F: 5’-GTTCCTCCAGCAGCATCAA-3’ and JADE2-R: 5’-GGCTTCTTCTCGTTCTGTCG-3’ (287-bp product). The reciprocal fusion was detected with primers JADE2-F: 5’-GGAGAGAGATTGCGATGT-3’ and NUP98-R: 5’-CCAGCCCCATCAAGAGATGT-3’ (237-bp product). GAPDH was amplified with primers GAPDH-F: 5’-CCACCCATGCAAATTTCCATGGCA-3’ and GAPDH-R: 5’-ATCTAGACGCGCTAGGTCCACC-3’ (598-bp product). The GeneAmp Fast PCR Master Mix (Thermo Fisher Scientific) was used for the amplification.

Analysis of HOXA and CDK6 mRNA expression was performed using the TB Green Premix Ex Taq II (Takara Bio) with the following primers: HOXA5-F: 5’-CAGATCTACCCCTGGATGCG-3’ and HOXA5-R: 5’-TTCAATCCCTCTCTCCTGAGG-3’; HOXA6-F: 5’-CTTGGATGAGCGTGAGTGAAC-3’ and HOXA6-R: 5’-TGTCAGTTAGCGTTGAAG-3’; HOXA9-F: 5’-CCTCCTCTCTCCTCGCCAACATA-3’ and HOXA9-R: 5’-CAGTTCCAGGGTCTGTGGT-3’; HOXA10-F: 5’-CCTCCTGAGAGGCAGAAG-3’ and HOXA10-R: 5’-AAACTCCTTCTCCAGCTCCA-3’; CDK6-F: 5’-GCCTTGCCCGCATCTATAGT-3’ and CDK6-R: 5’-TATGCAGCAACTACCTCACAAG-3’. Each sample was measured in triplicate and expression levels were determined by 2^ΔΔCt. GAPDH was used for normalization.
**Targeted next-generation sequencing**

Amplicon sequencing using paired diagnostic and remission bone marrow samples from the patient was performed to identify somatic mutations. Libraries were prepared from genomic DNA using the unique molecular identifier (UMI)-based QIAseq Targeted Human Myeloid Neoplasms Panel (Qiagen) and sequenced on an Illumina NextSeq 500 system. The panel covers the complete coding region of 141 myeloid-related genes (Table S1). The read processing, alignment (hg19 as the reference), calling and annotation of single nucleotide variants/small indels were performed with the UMI-based caller smCounter2 run on the GeneGlobe (Qiagen).\(^1\) Variant filtering was performed based on the method previously described by the German AML Cooperative Group.\(^2\) As our routine testing on a commercial myeloid DNA reference standard (Horizon) containing 22 variants in 19 genes with this method consistently identified variants with a variant allele frequency (VAF) of 5%, this was chosen as a cut-off for variant filtering. Variants with a population frequency of \(\geq 0.1\%\) in the 1000 Genomes Project (Phase 3), Genome Aggregation Database (v2.1.1) or dbSNP (Build 154) were excluded from the analysis.

**Targeted copy number variation (CNV) analysis**

The CNVkit\(^3\) and quandico\(^4\) tools were used to identify CNVs in the 141 genes from the amplicon sequencing data. Twelve DNA samples from healthy individuals (6 males and 6 females) were used as controls to build the reference data for comparison. CNVkit analysis (v0.9.6) was run with DNAnexus using the amplicon sequencing setting, whereas quandico was run with the GeneGlobe (Qiagen). Genomic regions with a log\(_2\) ratio of \(< -0.3\) or \(> 0.3\) detected by CNVkit (these ratios are roughly equivalent to single-copy loss and single-copy gain in diploid leukemic cells accounting for \(~40-50\%\) of total cell population in the samples) and a Q score of \(\geq 50\) in quandico analysis were considered CNVs.

**Plasmid constructs**

Full-length cDNAs of \(NUP98-JADE2\) and \(JADE2\) were cloned into the pCMV-Myc (Takara Bio), pCMV-HA (Takara Bio) and LEGO-iG2\(^5\) expression vectors using the In-Fusion HD Cloning Kit (Takara Bio). For pCMV-Myc and pCMV-HA, the C-terminal Tag vectors were used. The identity of the cDNAs was confirmed by direct sequencing and restriction analysis.

**Immunofluorescence studies**

HeLa cells were transfected with pCMV-HA and/or pCMV-Myc vectors as indicated using Lipofectamine 2000 (Thermo Fisher Scientific). After 48 hours of transfection, cells were fixed with
3.7% formaldehyde and permeabilized with 0.5% Triton X-100. Samples were then blocked with 5% bovine serum albumin and incubated with a rabbit anti-HA (C29F4) or a mouse anti-Myc (9B11) monoclonal antibody (Cell Signaling Technology) at room temperature for 1 hour. After successive washing, samples were incubated with an Alexa Fluor 488- (A21206) or Alexa Fluor 594-conjugated (A11005) secondary antibody (Thermo Fisher Scientific) for another 1 hour. Cells were counterstained with DAPI and examined by the Zeiss Imager M1 fluorescence microscope.

**Co-immunoprecipitation (Co-IP) studies**

Co-IP assays were performed with the Pierce MS-Compatible Magnetic IP Kit (Thermo Fisher Scientific) following the manufacturer’s protocol. Briefly, after 48 hours of transfection, cell lysates were prepared and incubated with an anti-Myc monoclonal antibody (9E10, Thermo Fisher Scientific) at 4°C overnight. After repeated washing, samples were eluted and analyzed by immunoblotting with anti-Myc (9B11), anti-HA (C29F4) and anti-GAPDH (Ab9485, Abcam) antibodies. Twenty micrograms of cell lysates were analyzed in the input.

**Luciferase reporter assays**

The Cignal RARE Reporter Assay Kit (Qiagen) was used to study the transcriptional changes in response to JADE2 perturbation. The RARE reporter contains a mixture of RARE-responsive firefly luciferase construct (tandem repeats of the DR5 retinoic acid-response element) and constitutively expressing *Renilla* luciferase construct for normalization. Briefly, cells were transiently co-transfected with the RARE reporter and expression plasmids/small interfering RNAs as indicated by Lipofectamine 2000. Luciferase activities were measured after 48 hours of the transfection using the Dual-Glo Luciferase Assay System (Promega). To study the effect of ATRA, cells were treated with 1µM of ATRA or DMSO (as vehicle control) for 6 hours before luciferase measurement. Firefly luciferase activity in cell lysates was normalized to *Renilla* luciferase activity for calculations. In all studies, parallel experiments using the negative control reporter were performed to adjust for any non-specific effects.

**Lentiviral packaging and transduction**

The Lenti-X Packaging System (Takara Bio) was used to prepare NUP98-JADE2 and control lentiviruses with the LEGO-iG2 constructs (co-expressed green fluorescent protein (GFP)) according to the manufacturer’s protocol. U937 and NB4 cells were transduced in 6-well plates coated with 45µg/ml of RetroNectin (Takara Bio) with a multiplicity of infection (MOI) of 20 and 40, respectively. Ninety-
six hours post transduction, cells were treated with ATRA or DMSO as indicated and CD11b expression on GFP-positive cells was measured by a BD FACSCalibur flow cytometry.

**RNA interference**

5×10⁶ NB4 cells were transiently transfected with 500nM siRNA by electroporation using the Bio-Rad Gene Pulser and 0.4cm cuvettes (300V and 950μF). The JADE2 siRNAs (siRNA#1:ID_s23596, siRNA#2:ID_s23597) and negative control siRNA were obtained from Ambion. The ON-TARGETplus Smartpool JADE2 siRNAs (Dharmacon) were used to independently validate JADE2 target genes identified by RNA-seq and study the effects of JADE2 knockdown on RARE-mediated transcriptional activity.

**Whole transcriptome sequencing**

Whole transcriptome sequencing was performed to identify gene expression changes upon JADE2 knockdown. Total RNA was extracted from NB4 cells transfected with negative control siRNA, JADE2 siRNA#1 and JADE2 siRNA#2 for 24 hours. Messenger RNA was purified from the total RNA and libraries were prepared using the NEBNext RNA Library preparation kit for Illumina sequencing. After removal of low-quality reads, paired-end clean reads were aligned to the reference genome (hg19) using STAR (v2.5). HTSeq (v0.6.1) was then used to count the read numbers mapped of each gene. FPKM (Fragments Per Kilobase of transcript per Million mapped reads) of each gene was calculated based on the length of the gene and reads count mapped to this gene. All the library preparation and data analysis steps were carried out by Novogene.

Validation of selected JADE2 target genes was done by quantitative RT-PCR using TaqMan Gene Expression Assays (CEBPE: Hs00357657_m1; BTG2: Hs00198887_m1; FOS: Hs99999140_m1; IL1B: Hs01555410_m1; S100A9: Hs00610058_m1; NCF2: Hs01084940_m1; TOMM20: Hs03276810_g1) (Thermo Fisher Scientific). GAPDH was used for normalization.
References

1. Xu C, Gu X, Padmanabhan R, et al. smCounter2: an accurate low-frequency variant caller for targeted sequencing data with unique molecular identifiers. *Bioinformatics*. 2019;35(8):1299-1309.

2. Metzeler KH, Herold T, Rothenberg-Thurley M, et al; AMLCG Study Group. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. *Blood*. 2016;128(5):686-698.

3. Talevich E, Shain AH, Botton T, et al. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput Biol*. 2016;12(4):e1004873.

4. Reinecke F, Satya RV, DiCarlo J. Quantitative analysis of differences in copy numbers using read depth obtained from PCR-enriched samples and controls. *BMC Bioinformatics*. 2015;16:17.

5. Weber K, Bartsch U, Stocking C, et al. A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis. *Mol Ther*. 2008;16(4):698-706.
**Supplemental Table 1.** The 141 genes covered in the panel.

| Gene 1 | Gene 2 | Gene 3 | Gene 4 | Gene 5 | Gene 6 | Gene 7 | Gene 8 | Gene 9 | Gene 10 | Gene 11 | Gene 12 |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|----------|---------|
| ABL1   | BRCA2  | CTCF   | FAS    | JAK1   | LUC7L2 | NTRK3  | PTPN11 | SRP72  | U2AF2    |          |         |
| ADA    | BRINP3 | CUX1   | FBXW7  | JAK2   | MAP2K1 | OR13H1 | RAD21  | SRSF2  | WAS      |          |         |
| ANKRD26| C1orf97| DAXX   | FLRT2  | JAK3   | MLH1   | OR8B12 | RB1    | STAG2  | WRN      |          |         |
| ASXL1  | CALR   | DDX41  | FLT3   | KAT6A  | MPL    | P2RY2  | RELN   | STAT3  | WT1      |          |         |
| ASXL2  | CARD11 | DNM2   | GATA1  | KCN44  | MSH2   | PAX5   | RUNX1  | STXB2  | XPO1     |          |         |
| ATM    | CBL    | DNMT1  | GATA2  | KCNK13 | MSH6   | PCDHB1 | SAXO2  | SUZ12  | ZRSR2    |          |         |
| ATRX   | CBLB   | DNMT3A | GJB3   | KDM6A  | MYC    | PDGFRA | SETBP1 | TAL1   |          |          |         |
| BCL6   | CBLC   | EED    | GNAS   | KDR    | MYD88  | PHF6   | SF1    | TERC   |          |          |         |
| BCOR   | CDKN2A | EGFR   | HNRNPK | KIT    | NBN    | PML    | SF3A1  | TERT   |          |          |         |
| BCOI   | CEBPA  | ELANE  | HRAS   | KLHDC8B| NF1    | PMS2   | SF3B1  | TET2   |          |          |         |
| BCR    | CHEK2  | EP300  | IDH1   | KLH6L  | NOTCH1 | PRAMEF2| SH2B3  | TNFRSF13B|          |          |         |
| BIRC3  | CREBBP | ETNK1  | IDH2   | KMT2A  | NPA1   | PRF1   | SH2D1A | TP53   |          |          |         |
| BLM    | CRLF2  | ETVI   | IKZF1  | KMT2C  | NPM1   | PRPF40B| SMARCB1| TPMT   |          |          |         |
| BRAF   | CSF1R  | EZH2   | IKZF3  | KRAS   | NRAS   | PRPF8  | SMCA1  | TUBA3C |          |          |         |
| BRCA1  | CSF3R  | FAM47A | IL7R   | LRRC4  | NSD1   | PTEN   | SMC3   | U2AF1  |          |          |         |
Supplemental Table 2. The list of genes consistently altered in JADE2-knocked down NB4 cells.

| Downregulated genes | Upregulated genes |
|---------------------|-------------------|
| Gene               | Mean fc*           | Gene           | Mean fc*     |
| NCF2               | 0.309             | NDUFA7         | 2.176        |
| IL1B                | 0.332             | AL158147.2     | 1.869        |
| S100A9             | 0.379             | MSRB3          | 1.817        |
| BTG2               | 0.384             | RPL36A         | 1.720        |
| SERPINE1           | 0.398             | TOMM20         | 1.678        |
| CSAR1              | 0.399             | CCND3          | 1.631        |
| NLRP12             | 0.428             | BMD1           | 1.592        |
| NK7                | 0.442             | HIFX           | 1.589        |
| OSGIN1             | 0.461             | BOD1           | 1.586        |
| HMOX2              | 0.462             | SPG20060       | 1.577        |
| IFGAM              | 0.479             | ISOC1          | 1.564        |
| PKR3R5             | 0.483             | CIT            | 1.534        |
| SHT1C1             | 0.488             | FRAO43         | 1.524        |
| ICAM1              | 0.498             | BANF1          | 1.519        |
| KLF4               | 0.520             | ERCC5          | 1.512        |
| NCF1               | 0.523             | GOPC           | 1.512        |
| SIGLEC7            | 0.526             | ACBD7          | 1.509        |
| ALOX5AP            | 0.531             | MEGF8B         | 1.501        |
| GLPR1              | 0.533             | BFXAP          | 1.487        |
| RA88B              | 0.549             | CCDC136        | 1.480        |
| IRF8               | 0.551             | CCDC78         | 1.447        |
| IL1RAP             | 0.551             | PRR22          | 1.441        |
| CRISPLD2           | 0.554             | APJS2          | 1.425        |
| TLR1               | 0.562             | GSE1           | 1.410        |
| FOS                | 0.566             | TAP1           | 1.404        |
| IL13RA1            | 0.568             | PPF53B         | 1.397        |
| SMIM14             | 0.571             | WDR96          | 1.387        |
| SLC4A4A2           | 0.573             | POLR3P         | 1.379        |
| IRAK2              | 0.575             | SMIM19         | 1.374        |
| GPR84              | 0.582             | MAMDC4         | 1.373        |
| CSF1R              | 0.586             | SIF7A2         | 1.369        |
| PERP               | 0.586             | PDCD6          | 1.363        |
| MSIA4A             | 0.589             | FAM1356B       | 1.350        |
| C10orf94           | 0.594             |                 |              |
| MAM9               | 0.602             |                 |              |
| KIAA0533            | 0.605             |                 |              |
| TGFB2R2            | 0.607             |                 |              |
| FGGR2A             | 0.609             |                 |              |
| LMRT1              | 0.627             |                 |              |
| DRAI               | 0.635             |                 |              |
| SLC12A6            | 0.638             |                 |              |
| MMP14              | 0.638             |                 |              |
| FOSL2              | 0.644             |                 |              |
| THBD               | 0.649             |                 |              |
| GALNT3             | 0.659             |                 |              |
| GCNT2              | 0.664             |                 |              |
| MILR1              | 0.665             |                 |              |
| CREG1              | 0.672             |                 |              |
| LAMC1              | 0.678             |                 |              |
| DNAJB9             | 0.679             |                 |              |
| PCMTD1             | 0.696             |                 |              |
| ETS3               | 0.701             |                 |              |
| CYTL1              | 0.717             |                 |              |
| RNF225             | 0.723             |                 |              |
| GRN                | 0.724             |                 |              |
| PIK3AP1            | 0.760             |                 |              |

*Fold change (fc) was calculated by comparing the FPKM of each gene in the JADE2 and negative siRNA groups. Mean fc was then calculated from the results of the two JADE2 siRNAs (si#1 and si#2) in two independent experiments.
Supplemental Figure 1. NUP98-JADE2 inhibited ATRA-mediated transcriptional activation. 293T cells were co-transfected with the RARE reporter and an increasing amount of pCMV-HA-NUP98-JADE2 (NJ) or the empty pCMV-HA (EV). After 48 hours of the transfection, cells were treated with 1μM of ATRA or DMSO for 6 hours before luciferase measurement. Parallel experiments using the negative control reporter were performed to adjust for any non-specific effects. Firefly luciferase activity in cell lysates was normalized to Renilla luciferase activity, and results are presented as fold induction by comparing to DMSO treatment. Results are expressed as mean plus standard error from 3 independent experiments each performed in triplicate. *** indicates $P<0.001$ by the Mann-Whitney U test. A dose-dependent reduction of fold activation is indicated.
Supplemental Figure 2. NUP98-JADE2 inhibited ATRA-induced U937 cell differentiation. After 96 hours of transduction with NUP98-JADE2 (NJ) or control (EV) lentiviruses, U937 cells were treated with 1µM of ATRA or DMSO (vehicle control) for 3 and 5 days. CD11b expression on GFP-positive cells was measured by flow cytometry. Results are expressed as mean plus standard error from three independent experiments. * and ** indicates $P<0.05$ and $P<0.01$, respectively by paired t test.
Supplemental Figure 3. *JADE2* knockdown repressed RARE-mediated gene transcription. HeLa cells were transfected with the RARE reporter in the presence of the ON-TARGETplus Smartpool *JADE2* siRNAs (*JADE2* kd) or negative control siRNA (Control). After 48 hours of the transfection, cells were treated with (+) or without (-) 1µM of ATRA for 6 hours before luciferase measurement. Parallel experiments using the negative control reporter were performed to adjust for any non-specific effects. Firefly luciferase activity in cell lysates was normalized to *Renilla* luciferase activity. Results are presented as normalized luciferase activity and expressed as mean plus standard error from 3 independent experiments each performed in triplicate. ** indicates *P*<0.01 by the Mann-Whitney U test.
Supplemental Figure 4. NUP98-JADE2 is associated with aberrant HOXA and CDK6 expression. Quantitative RT-PCR analysis of HOXA genes and CDK6 in the diagnostic (AD) and remission (REM) bone marrow (BM) samples from the patient. Whole BM cells were used for the analysis. Fold change (vs REM, ranged 5.5- to 81.7-fold) was calculated using $2^{\Delta\Delta Ct}$. Results are expressed as mean plus standard error from triplicate measurements.