Enforcement of stem-cell dormancy by nucleophosmin mutation is a critical determinant of unrestricted self-renewal during myeloid leukemogenesis

Maria Elena Boggio Merlo,[1*] Maria Mallardo,[1*] Lucilla Luzi,[1*] Giulia De Conti,[1] Chiara Caprioli,[1] Roman Hillje,[1] Mario Faretta,[1] Cecilia Restelli,[1] Andrea Polazzi,[1] Valentina Tabanelli,[2] Angelica Calleri,[2] Stefano Pileri,[2,3] Pier Giuseppe Pelicci[1,4] and Emanuela Colombo[1,4]

[1]Department of Experimental Oncology, European Institute of Oncology (IEO), IRCCS, Milan; [2]Unit of Haemato-Pathology, European Institute of Oncology, Milan; [3]Department of Experimental, Diagnostic and Specialty Medicine, Bologna University School of Medicine, Bologna and [4]Department of Oncology and Haemato-Oncology, University of Milan, Milan, Italy

**MEBM, MM and LL contributed equally as first authors.*

Correspondence: E. Colombo
emanuela.colombo@ieo.it

P.G. Pelicci
piergiuseppe.pelicci@ieo.it.

Received: September 9, 2024.
Accepted: March 6, 2025.
Early view: March 13, 2025.

https://doi.org/10.3324/haematol.2024.286577

Supplementary Materials and Methods

RNA extraction and population-based RNA-seq

RNA was purified using QIAGEN RNeasy kit l from NPMc+/YFP and YFP FACS-sorted LT-HSCs and subjected to microarray analysis. Double stranded cDNA synthesis and related cRNA was performed with Nugen® Pico WTA Systems V2 (NuGEN Technologies, Inc). Hybridization was performed using the Affymetrix GeneChip® Mouse Gene ST 2.0 Arrays. Microarray data were normalized by RMA using Partek Genomic Suite 6.6.

For bulk whole transcriptome sequencing, RNA was purified from FACS-sorted LT-HSCs with PicoPureTM RNA Isolation Kit (ThermoFisher). Sequencing libraries were generated using the SMARTseq protocol based on polyA-enrichment. 50bp single-end sequencing was performed with a HiSeq2000 device (Illumina). Sequences were aligned to the mouse reference genome (NCBI37/mm9) using TopHat2[1]. After alignment, raw gene expression values were obtained with HTSeq [2]. Differential gene expression was then estimated using the edgeR R package, using a TMM normalization[3].

Single-cell library preparation and sequencing

For scRNA-seq libraries, we sorted LT-HSCs from mice bone marrow : 5 mice for Mx and FLT-ITD/Mx and 4 animals for NPMc/Mx and NPMc/FLT-ITD/Mx. The libraries were prepared using the Chromium Single Cell 3'Reagent Kits (v2): Single Cell 3'Library & Gel Bead Kit v2 (PN-120237), Single Cell 3'Chip Kit v2 (PN-120236) and i7 Multiplex Kit (PN-120262) (10x Genomics) and following the Single Cell 3'Reagent Kits (v2) User Guide (manual part no. CG00052 Rev C). Libraries were sequenced on NovaSeq 6000 Sequencing System (Illumina) with an asymmetric paired-end strategy (28 and 91 bp read length for R1 and R2 mate respectively) with a coverage of about 75,000 reads/cell.

ScRNA-seq data analysis

Demultiplexing and sequencing quality control. The binary raw base-call sequencing results of the four samples (Mx, NPMc+/Mx, Flt3-ITD/Mx, and NPMc+/Flt3-ITD/Mx) were demultiplexed and converted to FASTQ files using the 10x Genomics Cell Ranger `mkfastq` pipeline. This program is a wrapper around the bcl2fastq routine and generates two FASTQ files per sample (one for each pair-end sequence, R1 and R2), per lane. Quality control of sequencing is performed on FASTQ files with FASTQC.

Alignment, annotation and creation of genes-by-cells matrix. Cell barcodes and UMIs are extracted for each read in the fastq files with umi_tools extract and the cDNA insert, contained in the R2 reads, are aligned to the mm10/GRCm38 reference genome with STARsolo, which annotates the aligned reads to a gene using the GENCODE gtf file (downloaded from ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M16/ site). Confidently mapped, non-PCR duplicates with valid barcodes and unique molecular identifiers were then used to create the UMI count table of genes-by-cells, using the umi_tools count program.

Pre-processing. All the following scRNA-seq data analyses were carried out using Seurat v4.0, unless otherwise specified. The samples separately obtained from the previous steps, are "merged" in a single seurat object for the second part of the pre-processing route.

To exclude cells that were extreme outliers in terms of library complexity and that might include multiple cells or doublets, we calculated the distribution of genes detected per cell and removed any cells in the top 2% quantile. Further, we removed cells with more than 10% of the transcripts coming from mitochondrial genes and genes expressed by <30 cells (n=16,017). In the final dataset we obtained 10,519 cells (2,481 for Mx, 2,332 for NPMc+/Mx, 2,385 for Flt3-ITD/Mx and 3,321 for NPMc+/Flt3-ITD/Mx). The mean and median numbers of detected genes-per-cell were 2,859.31 and 2,992.00 in Mx, 3,156.01 and 3,190.00 in NPMc+/Mx, 2,880.96 and 2,980.00 in Flt3-ITD/Mx and 2,879.66 and 2,901.00 in NPMc+/Flt3-ITD/Mx sample, respectively. We then normalized the raw data by the total expression of each cell and multiplied by a scale factor of 10,000.

scRNAseq analysis. A standard procedure of scRNA-seq analysis was then followed, for the identification of highly variable genes (FindVariableFeatures, nfeatures = 2000), and linear and non-linear dimensionality reduction by PCA and UMAP with the functions RunPCA and RunUMAP, respectively. To correct for batch effect between samples, we performed data integration applying the FindIntegrationAnchors and

`IntegrateData` commands. The UMAP plots of not-batch-corrected (not shown) and batch-corrected data to jointly visualize our four samples were produced with the `DimPlot` seurat script.

Trajectory analysis. To reveal the underlying dynamics of the dormant-to-active biological transition in our samples and to identify the genes that were significantly associated with this process, we performed a trajectory analysis with Monocle2. This method allows to infer a temporal dimension, also known as "pseudotime", by leveraging patterns of variation within the data and to reconstruct an ordered progression of cells through different biological states. For the trajectory analysis, the merged seurat object of the four samples was transformed in a monocle2 object (cds, CellDataSet) and then the functions `reduceDimension,` with method = 'DDRTree' and `orderCells,` was applied, using the dormant MolO list signature (n=27) as ordering gene set. Finally, we plot the resulted trajectory with the `plot_cell_trajectory` function, that charts the "minimum spanning tree" on cells, either coloring the cells by pseudotime or MolO genes signature average expression level. This plot maps cells in low-dimensional space by ICA (Independent Component Analysis); component 1 and 2 allows to visualize how cells progress along different functional states. The algorithm assigns to each cell a pseudo-time value as a metric of their position in the virtual timeline progression.

Identification of groups and subgroups of cells. We classified our single cell dataset in three groups (Gr1, Gr2 and Gr3) based on their assigned pseudo-time value. We arbitrarily set three ranges of the pseudo-time (Gr1 <1.5, Gr2 between 1.5 and 4.3 and Gr3 >= 4.3; based on the density plot profiles in Figure S4E) that segregate the most evident differences in the profiles. These intervals separated cells with different levels of expression of both the dormant and active signatures (see below).

We further categorized cells based on their level of expression (Normalized Mean Expression; NME) of dormant (Do28) and active (Act152) gene lists. These thresholds were set as the 99th percentile of the NME distribution of Mx cells in Group 3 and 1, where the expression level of dormant and active signatures, respectively, reaches their minimum (Figure 5B, panel a and Figure 5C, panel a; minimum NME for dormant signature = 2.098, min NME for active signature = 0.953). Finally, we have applied the thresholds to classify all our cells as dormant HSC (dHSC, cells with NME above the dormant threshold), active HSC (aHSC, cells with NME above the active threshold), dormant/active HSC (cells with NME above both thresholds) and transient HSC (TrHSC, cells with NME below both thresholds) (Table S3C).

Single cell RNAseq differential expression analysis.

To characterize the transcriptional profile of the 4 experimental samples and the groups defined on the pseudo-time trajectory, we performed a differential expression analysis by a multifactorial approach. We used the `differentialGeneTest` Monocle's main differential analysis routine, that accepts a CellDataSet and two model formulae as input, and specifies generalized lineage models as implemented by the VGAM package. As covariates we applied *pseudotimes* – in terms of differences of expression between cells belonging to the different pseudotime groups, Gr1, Gr2 and Gr3 - and *samples*. In particular, we separately studied, pairwise, the following contrasts: NPMc+/Mx vs Mx, Flt3-ITD/Mx vs Mx, NPMc+/Flt3-ITD/Mx vs Mx and NPMc+/Flt3-ITD/Mx vs Flt3-ITD/Mx. In turn, each of these comparisons were challenged for differences in Gr2 vs Gr1, Gr3 vs Gr2 and Gr3 vs Gr1 groups of cells. The list of the significant differentially regulated dormant and active genes, regulated in at least one of the contrasts, is reported in Table S2B.

Gene signatures

The WikiPathways lists of genes for Mus Musculus were downloaded using the Virtuoso SPARQL Query Editor for WikiPathways at http://sparql.wikipathways.org/.

We manually curated two gene signatures for dormant (Do28) and for active (Act152) HSCs, considering both the bulk RNAseq and the single cell RNAseq raw data available in Cabezas *et al.* [4]. In particular, we performed GSEA of the bulk RNAseq dataset using as query the list of genes upregulated in dormant or active HSCs at single cell level, respectively. Leading Edge genes of the enrichment profile represent our dormant (Do28; n=28) or active (Act152; n=152) HSC gene set and are listed in table S2A.

Survival and differential expression analysis on TCGA AML patients

Clinical, cytogenetic, mutational and RNA-seq data from the TCGA-LAML cohort were downloaded from cBioPortal database (http://www.cbioportal.org/) [5] and 173 out of 200 patients were selected based on availability of gene expression data. RSEM expression values were pre-processed with non-paranormal transformation using the R package huge[6,7]. Patients were reclassified into risk classes according to European Leukemia Net guidelines [7] based on their cytogenetic and mutational profile; *FLT3-ITD* mutant patients were always considered at intermediate risk.

To investigate the impact of *TGFB1* expression on survival and relapse, we performed ROC analysis with the R package pROC [8] censoring patients at 17.1 months (the median overall survival of the analyzed cohort) and using *TGFB1* expression as predictor. Best threshold for overall survival prediction was derived according to Youden index method and used to divide patients into high and low *TGFB1* expression. For the two groups, overall survival probability was estimated by Kaplan-Meier method and compared by Logrank test, while cumulative incidence of relapse was tested by Gray's method. For multivariate models, Cox Proportional Hazard Models were performed including total number of white blood cells, age, ELN risk class and *TGFB1* status at diagnosis as covariates and using death by any cause and relapse as competing risks. Survival analyses were performed using R packages survminer and cmprsk. Clinical and molecular characteristics of patients in the two groups were compared by Pearson's Chi-squared test for categorical variables and Kruskal-Wallis rank sum test for numerical variables. To examine differentially expressed genes in *TGFB1* high vs low expressors, RNA-seq raw read counts were downloaded from the NCI Genomic Data Commons (https://gdc.cancer.gov/) [9], matched by barcode and then analyzed with the R package DESeq2[10] at gene level. The resulting DEGs list (adjusted p-value (q-value) < 0.1) was used to perform GSEAs.

1. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013; **14**(4): R36.

2. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015; **31**(2): 166-9.

3. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; **26**(1): 139-40.

4. Cabezas-Wallscheid N, Buettner F, Sommerkamp P, et al. Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell* 2017; **169**(5): 807-23 e19.

5. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012; **2**(5): 401-4.

6. Zhao T, Liu H, Roeder K, Lafferty J, Wasserman L. The huge Package for High-dimensional Undirected Graph Estimation in R. *J Mach Learn Res* 2012; **13**: 1059-62.

7. Dohner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017; **129**(4): 424-47.

8. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12**: 77.

9. Grossman RL, Heath AP, Ferretti V, et al. Toward a Shared Vision for Cancer Genomic Data. *N Engl J Med* 2016; **375**(12): 1109-12.

10. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**(12): 550.

Supplementary tables legends

Supplemental Table 1: *Global RNA expression analyses of WT, NPMc+, Flt3-ITD and NPMc+/Flt3-ITD LT-HSCs.* **A**: DEGs analysis by Affymetrix GeneChip in NPMc+/YFP+ or YFP+ derived LT_HSCs. **B**: DEGs analysis by bulk RNA-seq of NPMc+/Flt3-ITD/Mx; Flt3-ITD/Mx and Mx derived LT_HSCs

Supplemental Table S2: relationships between groups of cells: samples, clusters and pseudo-Groups. Number of cells in each sub-group.

Supplemental Table 3: *Pseudotime trajectories of Mx, NPMc+/Mx, Flt3-ITD/Mx and NPMc+/Flt3-ITD/Mx LT-HSCs*. **A**: signatures genes lists used to characterize the pseudo-time trajectory. **B**: list of genes showing a significant variation (qval<=0.01) among samples and along the pseudo-time either in the dormant (Do28+MolO) or in the active (Act152) genes lists (Sig_Act, n=152 and Sig_Do, n=28), respectively). **C**: Numbers and percentages of dHSCs, aHSCs , trHSCs and daHSCs, in groups and samples

Supplemental Table 4: *Analysis of human TCGA AML dataset*. **A.** Cox Proportional Hazard Models for multivariate analysis of overall survival. HR: hazard ratio; CI: confidence interval. **B.** Cox Proportional Hazard Models for multivariate analysis of cumulative incidence of relapse. HR: hazard ratio; CI: confidence interval. **C.** Clinical and molecular characteristics of AML patients in the TCGA cohort according to TGFb1 level of expression.

Supplemental Table 5: genes differentially expressed in TGFb1 High AML versus TGFb1 low AML

Supplementary figures legends

Figure S1. Analyses of different HSPCs population upon NPMc+ expression. A. *In vivo* BrdU assay: BM was collected 12 hours after *in vivo* BrdU administration and analysed by FACS to evaluate the percentage of BrdU positive cells in different HSPCs population (LSKs, LT-HSCs, ST- HSCs, MPPs) (3-5 mice per cohort; graph representative of 1 of 2 independent experiments; Unpaired Student's t test has been applied, *=p<0.05)**. B.** Percentage of Caspase-3 positive cells in the indicated HSPC subpopulations (3 mice per cohort). Mean ± s.d. values are shown. **C.** Representative FACS gating schemes for the analysis of different BM populations. Upper panel: gating of the LSK (c-Kit+, Sca-1+, Lin-) population within lineage negative cells; middle panel: gating of CD48-/CD150+ population within the LSK cells; lower panel: gating of the CD34- population within the CD48-/CD150+ cells. **D.** Left panels: percentages of LSK/CD48-/CD150+ (upper) and LSK/CD48-/CD150+/CD34- BM subpopulations as gated in C. Right panels: numbers (per million of BM-MNCs) of the same populations (3-5 mice per cohort; graph representative of 1 of 2 independent experiments; *=p<0.05).

Figure S2. **Mx-CRE mediated NPMc+ expression in the PB and in the LKS compartment A.** Representative images of the immunofluorescence analysis performed on Cytospin® preparations of PB WBCs derived from pIpC treated Mx, NPMc+ and NPMc+/FLT3-ITD mice. In red, rabbit polyclonal antibody against NPMc+; in blue, DAPI staining of nuclei. NPMc+ cytoplasmic localization is confirmed by merging red and blue staining. Scale bar 100 um. **B**. Histograms represent the percentage of NPMc+ positive WBCs in the total WBCs population (data represent the pool of four independent experiments). **C**. Plots represent the NPMc+ dependent fluorescence signal intensity/cell, quantified by the A.M.I.CO. computational platform in Mx and NPMc+/Mx LKS samples. DNA content was evaluated using DAPI staining intensity (data represent the pool of two independent experiments). **D.** Kaplan Meyer survival curve for Mx, Flt3-ITD/Mx and NPMc+/Flt3-ITD/Mx mice (5 animals per group, p-value calculated using the log- rank test).

Figure S3. NPMc+ induces the expression of HSC and LSC related signatures A-F. Gene expression microarray data were used to identify gene sets enriched in NPMc+/YFP LT-HSCs compared to YFP LT-HSCs. GSEAs show a significant correlation of the NPMc+ LT-HSC gene expression profile with: i) genes upregulated (**A** left panel) and downregulated (**A** right panel and **B**)

in human hematopoietic stem cells; ii) genes upregulated in NPMc+ and MLL human AMLs (**C**; **D-E**); iii) genes up regulated in AML stem cells (**F**). **G-H**. RNAseq data were used to identify enriched gene sets in Flt3-ITD/Mx vs Mx LT-HSCs (left panels) or NPMc+/Flt3-ITD/Mx vs Flt3- ITD/Mx LT-HSCs (right panels). GSEAs show significant correlations with genes upregulated in human hematopoietic stem cells (**G**), and genes upregulated in AML stem cells (**H**). Normalized enrichment score (NES), p value (p) and false discovery rate (q) are indicated in each panel.

Figure S4. Single cell RNAseq analysis. A. UMAP of 10519 LT-HSCs, derived from our 4 mouse models, after filtering and integration. UMAP representation annotated by genotype (**A**), by cluster **(B**) and by pseudotime group (**C**). **D**. scatter plots showing normalized mean expression levels by pseudotime of gene lists reported in Table S2, in the pool of the four samples. Do28 (upper left panel, dHSC related signature); Act166 (upper right panel, aHSC related signature); DNArep41 (lower left panel, genes annotated to the DNA replication pathway from WikiPathways); G1S58 (lower right panel, genes annotated to G1 to S cell cycle transition from WikiPathways). Colored lines depict the polynomial fit of the normalized expression values. **E.** density plot displaying single cell distribution along the trajectory. Groups1-3 were defined as described in materials and methods section (pseudotime <1.5; > 1.5 and <4.3; >4.3, respectively).

Figure S5. Pharmacological inhibition of TGFβ pathway affects AML growth *in vivo*. A. Kaplan Meyer survival curve of mice transplanted with a NPMc+/*Flt3-ITD* murine AML and treated with the LY364947 TGFβR-I inhibitor or vehicle every other day for 10 days (10 animals/group). **B**. Graphs represent the percentage of control (left panel) or TGFβRI inhibited (right panel) blasts (CD45.2+) in the PB of secondary transplanted mice (n=5 animals per group) **C.** Kaplan Meier survival curve of mice transplanted with TGFβRI inhibited blasts or control blasts (n=5 animals). The p value was calculated with the log rank test. **D.** ROC analysis to identify best *TGF*β*1* expression cutoff-predictor by Youden index (red); AUC: area under the curve (blue). **E-G.** RNAseq data from patients (TCGA data set) were used to identify enriched gene signatures in TGFβ1 high versus TGFβ1 low groups. GSEAs show a significant correlation with: i) HALLMARK TGFβ pathway genes; ii) genes upregulated (**F** left panel) and downregulated (**F** right panel) in NPMc+ human AMLs ; iii) HSCs dormant (Do28+MolO; Table S2A)) gene signatures (**G** left panel). No correlation was found with the HSCs active (Act166; Table S2A) gene signature (**G** right panel). Normalized enrichment score (NES), p value (p) and false discovery rate (q) are indicated in each panel.

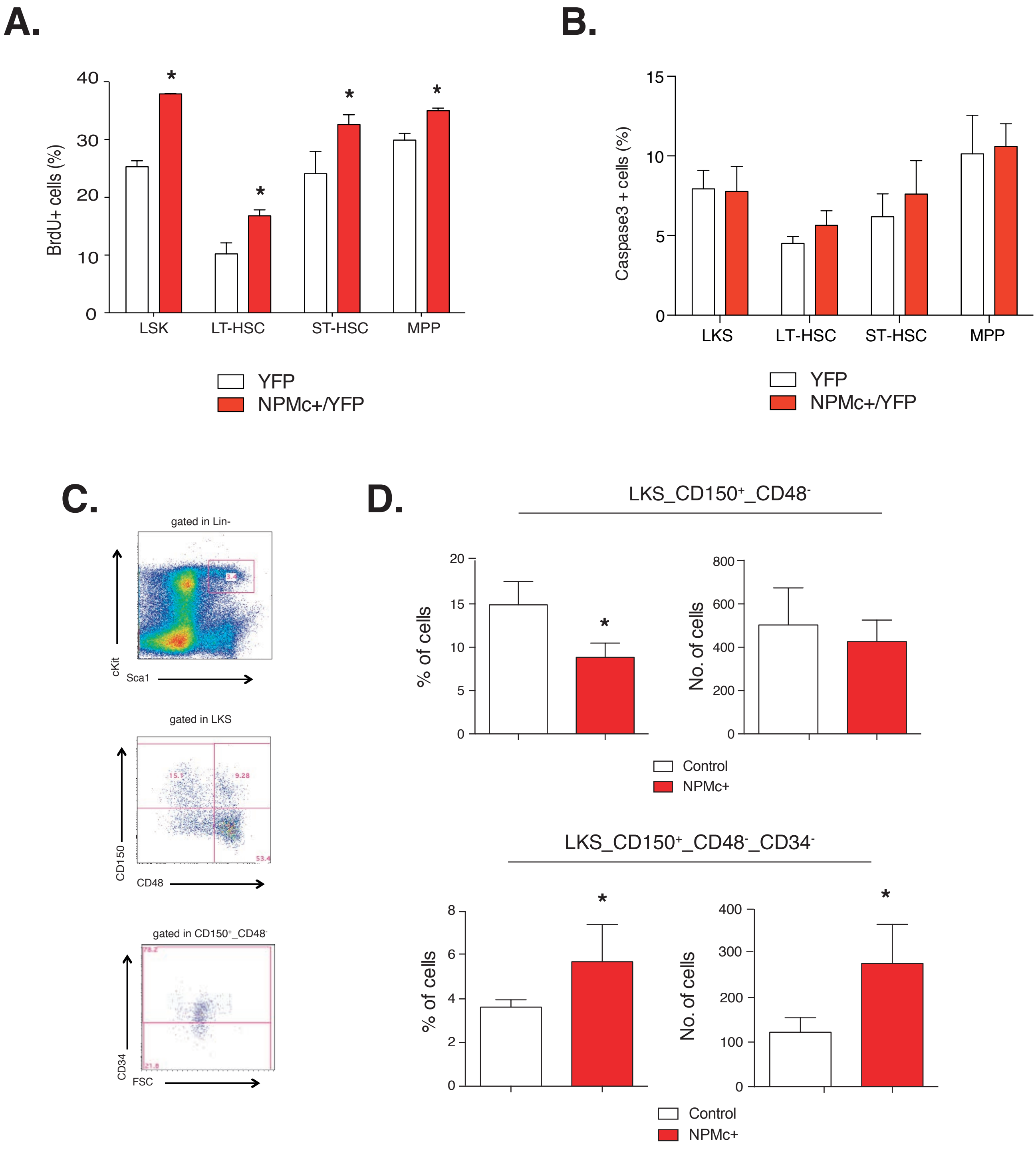
A.
BrdU+ cells (%)
40
30
20
10
0
*
*
*
*
LSK
LT-HSC
ST-HSC
MPP
YFP
NPMc+/YFP
B.
Caspase3 + cells (%)
15
10
5
0
LKS
LT-HSC
ST-HSC
MPP
YFP
NPMc+/YFP
C.
gated in Lin-
3.4
cKit
Sca1
gated in LKS
15.1
9.28
53.4
CD150
CD48
gated in CD150+_CD48-
78.2
21.8
CD34
FSC
D.
LKS_CD150+_CD48-
% of cells
20
15
10
5
0
*
No. of cells
800
600
400
200
0
Control
NPMc+
LKS_CD150+_CD48-_CD34-
% of cells
8
6
4
2
0
*
No. of cells
400
300
200
100
0
*
Control
NPMc+

Figure S1

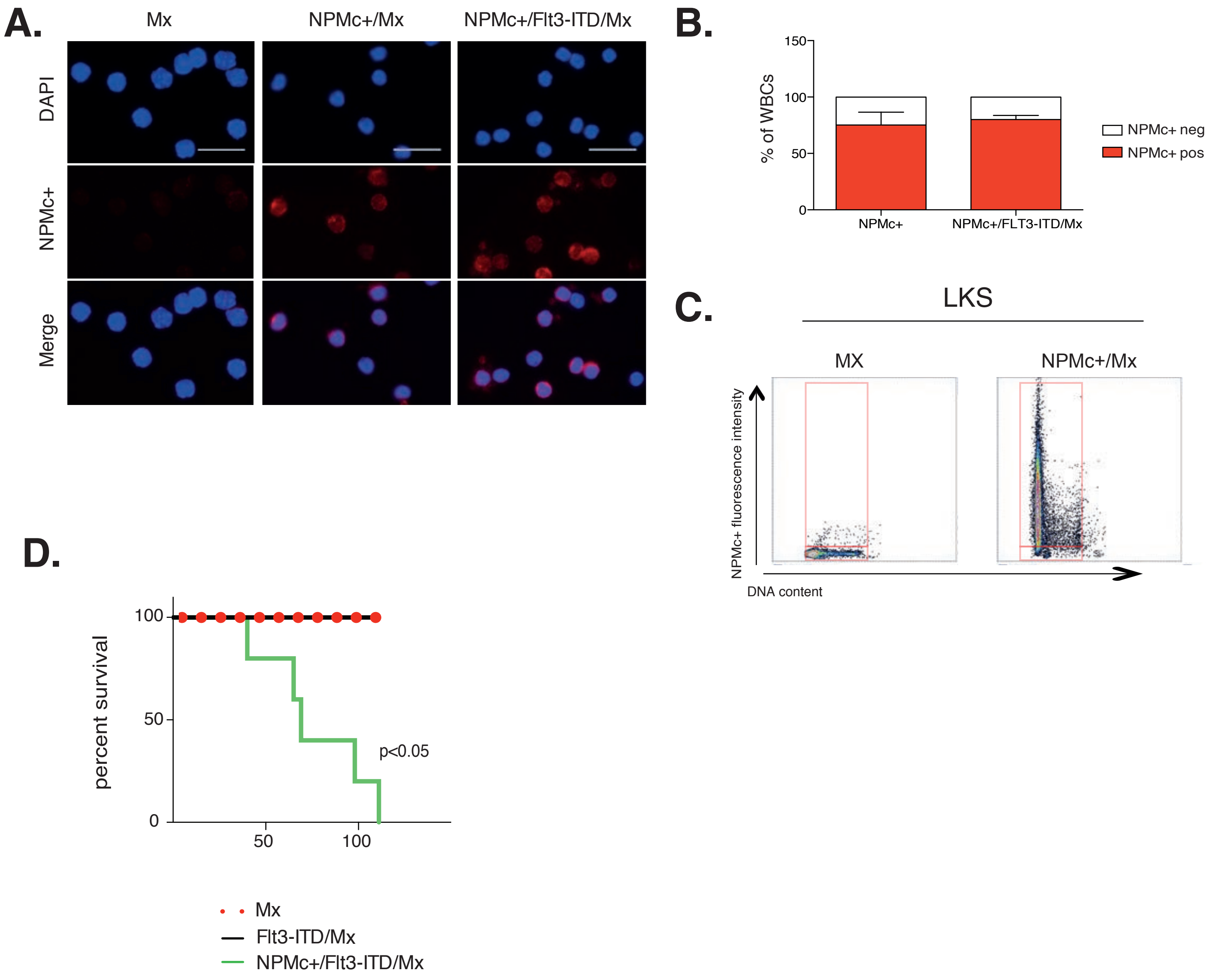
A.
Mx
NPMc+/Mx
NPMc+/Flt3-ITD/Mx
DAPI
NPMc+
Merge
B.
150
100
50
0
% of WBCs
NPMc+
NPMc+/FLT3-ITD/Mx
□ NPMc+ neg
🟥 NPMc+ pos
C.
LKS
MX
NPMc+/Mx
NPMc+ fluorescence intensity
DNA content
D.
100
50
0
percent survival
50
100
p<0.05
• • Mx
— Flt3-ITD/Mx
— NPMc+/Flt3-ITD/Mx

Figure S2

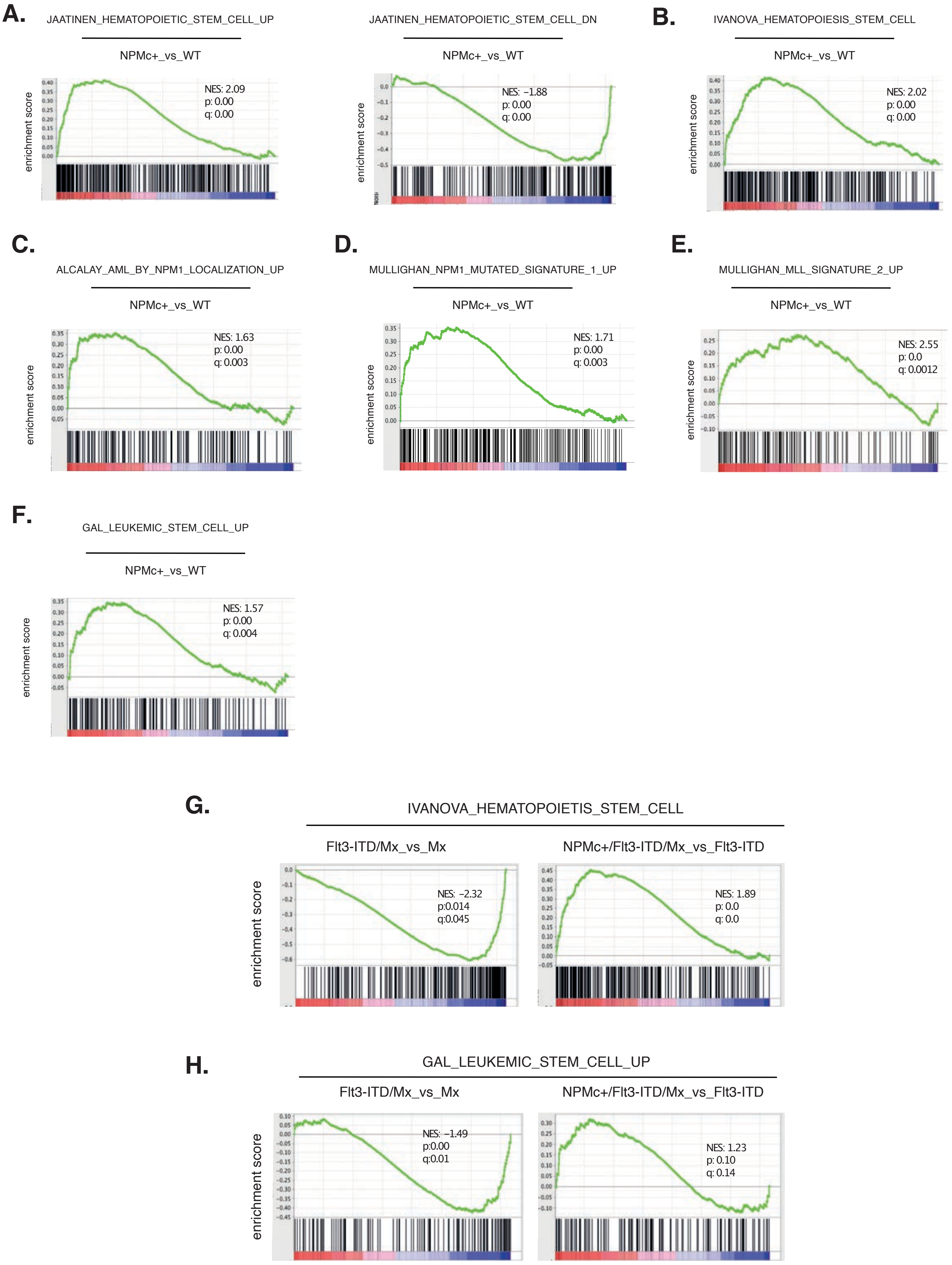
A.
JAATINEN_HEMATOPOIETIC_STEM_CELL_UP
NPMc+_vs_WT
enrichment score
NES: 2.09
p: 0.00
q: 0.00
JAATINEN_HEMATOPOIETIC_STEM_CELL_DN
NPMc+_vs_WT
enrichment score
NES: −1.88
p: 0.00
q: 0.00
B.
IVANOVA_HEMATOPOIESIS_STEM_CELL
NPMc+_vs_WT
enrichment score
NES: 2.02
p: 0.00
q: 0.00
C.
ALCALAY_AML_BY_NPM1_LOCALIZATION_UP
NPMc+_vs_WT
enrichment score
NES: 1.63
p: 0.00
q: 0.003
D.
MULLIGHAN_NPM1_MUTATED_SIGNATURE_1_UP
NPMc+_vs_WT
enrichment score
NES: 1.71
p: 0.00
q: 0.003
E.
MULLIGHAN_MLL_SIGNATURE_2_UP
NPMc+_vs_WT
enrichment score
NES: 2.55
p: 0.0
q: 0.0012
F.
GAL_LEUKEMIC_STEM_CELL_UP
NPMc+_vs_WT
enrichment score
NES: 1.57
p: 0.00
q: 0.004
G.
IVANOVA_HEMATOPOIETIS_STEM_CELL
Flt3-ITD/Mx_vs_Mx
NPMc+/Flt3-ITD/Mx_vs_Flt3-ITD
enrichment score
NES: −2.32
p:0.014
q:0.045
NES: 1.89
p: 0.0
q: 0.0
H.
GAL_LEUKEMIC_STEM_CELL_UP
Flt3-ITD/Mx_vs_Mx
NPMc+/Flt3-ITD/Mx_vs_Flt3-ITD
enrichment score
NES: −1.49
p:0.00
q:0.01
NES: 1.23
p: 0.10
q: 0.14

Figure S3

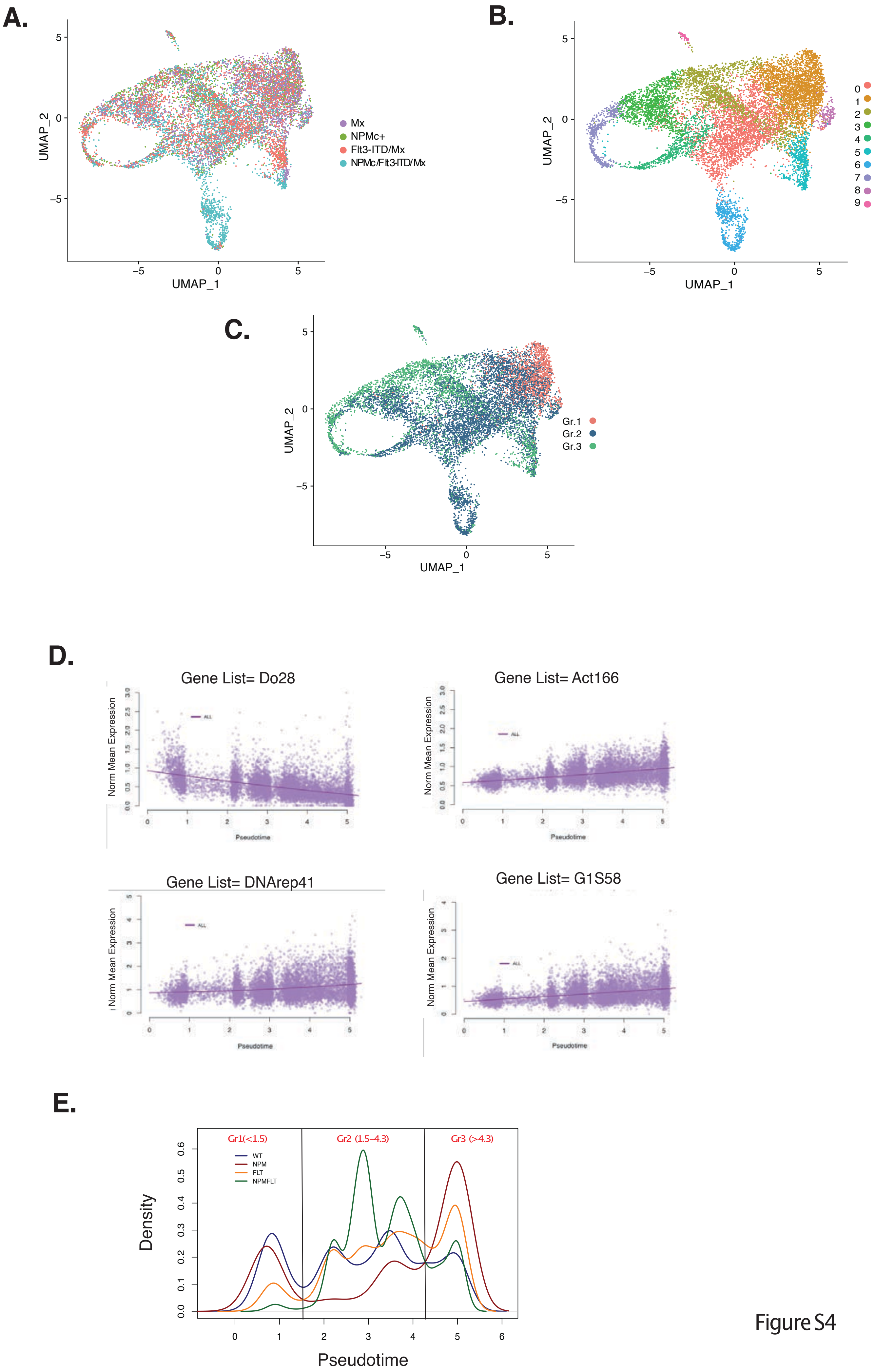
A.
UMAP_2
UMAP_1
Mx
NPMc+
Flt3-ITD/Mx
NPMc/Flt3-ITD/Mx
B.
0
1
2
3
4
5
6
7
8
9
C.
Gr.1
Gr.2
Gr.3
D.
Gene List= Do28
Gene List= Act166
Gene List= DNArep41
Gene List= G1S58
Norm Mean Expression
Pseudotime
ALL
E.
Gr1(<1.5)
Gr2 (1.5-4.3)
Gr3 (>4.3)
WT
NPM
FLT
NPMFLT
Density
Pseudotime

Figure S4

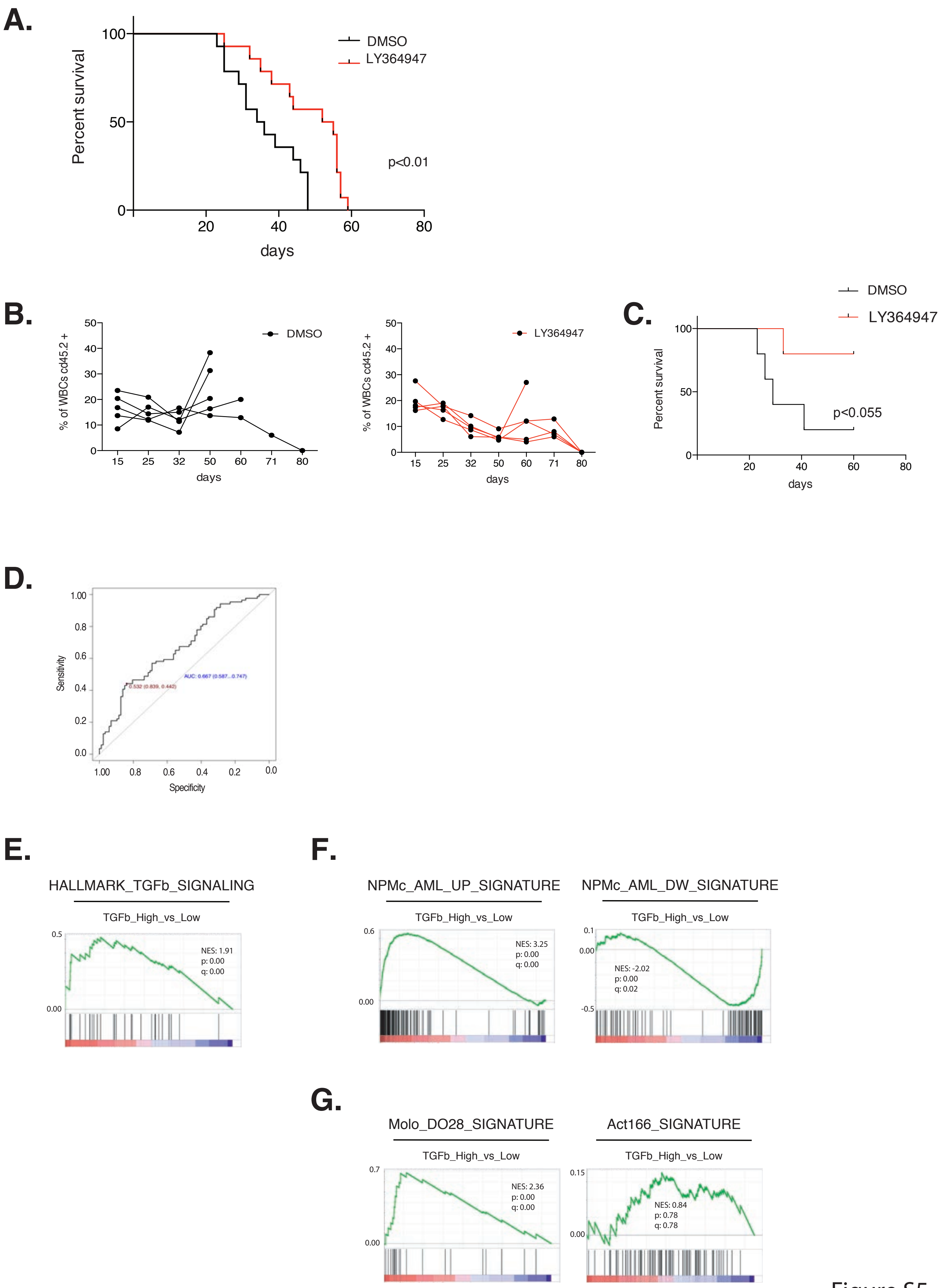
A.
Percent survival
100
50
0
20
40
60
80
days
DMSO
LY364947
p<0.01
B.
% of WBCs cd45.2 +
50
40
30
20
10
0
15
25
32
50
60
71
80
days
DMSO
% of WBCs cd45.2 +
50
40
30
20
10
0
15
25
32
50
60
71
80
days
LY364947
C.
DMSO
LY364947
Percent survival
100
50
0
20
40
60
80
days
p<0.055
D.
Sensitivity
1.00
0.8
0.6
0.4
0.2
0.0
1.00
0.8
0.6
0.4
0.2
0.0
Specificity
AUC: 0.667 (0.587...0.747)
0.532 (0.839, 0.442)
E.
HALLMARK_TGFb_SIGNALING
TGFb_High_vs_Low
0.5
0.00
NES: 1.91
p: 0.00
q: 0.00
F.
NPMc_AML_UP_SIGNATURE
TGFb_High_vs_Low
0.6
0.00
NES: 3.25
p: 0.00
q: 0.00
NPMc_AML_DW_SIGNATURE
TGFb_High_vs_Low
0.1
0.00
-0.5
NES: -2.02
p: 0.00
q: 0.02
G.
Molo_DO28_SIGNATURE
TGFb_High_vs_Low
0.7
0.00
NES: 2.36
p: 0.00
q: 0.00
Act166_SIGNATURE
TGFb_High_vs_Low
0.15
0.00
NES: 0.84
p: 0.78
q: 0.78

Figure S5