SUPPLEMENTAL INFORMATION

GRO-Seq sample preparation, run-on and library construction

Cells were pelleted, washed twice with ice-cold PBS and kept on ice. Lysis buffer (10 mM Tris-HCl, pH 7.4, 0.5% NP-40, 10% glycerol, 3 mM $CaCl_2$, 2mM $MgCl_2$, 1mM DTT, 1X protease inhibitor cocktail and SUPERase-In) was added and cells were pipetted up and down through 1 ml tips to release nuclei. The nuclei were washed twice with lysis buffer and resuspended in 100ul freezing buffer (50 mM Tris-HCl, pH 8.3, 40% glycerol, 5 mM $MgCl_2$, 0.1mM EDTA, and 4 units of SUPERase-In per mL). 8 x 10^6 nuclei were used for each library. Nuclear run-on and library construction were performed as previously described (Franco et al., 2015). Linearized DNA was amplified by PCR with indexed primers containing barcodes, then sequenced on NextSeq500.

Analysis of GRO-seq Data

The GRO-seq data were analyzed using the groHMM tool (Chae et al., 2015) and the approaches described below. Scripts and other related information about the analyses can be obtained by contacting the corresponding author (X.B.).

Quality Control and Read Alignment.

Quality control for the GRO-seq data was performed using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). GRO-seq reads were aligned to the human reference genome (hg19), using the BWA aligner (Li and Durbin, 2010). Only uniquely mapping reads were considered for the downstream analyses. After combining the replicates,

uniquely mappable reads for each strand were converted into bigWig files using the writeWiggle function in the groHMM tool for visualization in UCSC genome browser (Chae et al., 2015).

Differential Expression Analyses.

The differential expression analyses were performed using the Bioconductor package called edgeR (Robinson et al., 2010). A window of +300bp to +13kb from the transcription start site (TSS) was used for the analysis. This window was chosen in order to exclude the reads from RNA polymerase that are engaged at the promoter region and to allow enough time for the elongation of newly initiated Pol II (Hah et al., 2011). Also, this ensured that only the gene body activity level estimates were taken into account to measure only the productively elongating RNA polymerase. In order to capture significant changes in gene body activity between the two conditions, only the significantly regulated genes with fold change greater than 2 were considered for further analyses.

Metagene Analyses.

Metagene analyses were performed to show the distribution of average Gro-seq read densities ±4 kb surrounding the transcription start site (TSS) using the metagene functions in the groHMM tool. Metagene analyses were also performed for the up regulated and down regulated genes to show the difference in the read density distributions in the two conditions.

Gene Ontology Analyses

Gene ontology analyses were performed using the DAVID (Database for Annotation, Visualization, and Integrated Discovery) tool.

Gene Set Enrichment Analysis (GSEA)

GSEA was performed with GSEA software (http://www.broadinstitute.org/gsea/) using the following parameters: permutation, phenotype; metric, Signal2Noise; metric, weighted. Pre-ranked gene lists were based on edgeR differential expression analysis. Dysregulated genes shared between NELF-B and NELF-E knockdown were used to generate pre-ranked lists.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Affiliated to Fig. 1

(A) Typical morphologies of cells produced under erythroid (Ery) and granulocytic (Gra) differentiation. Cells at different days of differentiation were cytospun and stained with May-Grunwald-Giemsa (MGG). Cells at later stages of erythroid differentiation were stained for benzidine for hemoglobin expression.

(B) Western blot of NELF subunits during erythroid differentiation of human $CD34^+$ HSPCs. GAPDH serves as a loading control.

(C) Quantification of western blot in Figure 1A by imageJ.

(D) Quantification of western blot in Figure 1C by imageJ.

(E) Quantification of western blot in Figure 1D by imageJ.

In B, C and D, protein level is normalized to GAPDH and presented as fold change relative to day0 (d0) samples (n=3, mean ± SEM, * $p<0.05$, ** $p<0.01$).

Figure S2. Affiliated with Figure 2.

(A) Correlation analysis of GRO-seq data from biological replicates of d0 and d3 samples.

(B) Enriched GO categories ($p<0.005$) for genes upregulaed (left) and downregulated (right) in d3 granulocytes.

(C) Genome browser captures of GRO-seq for granulocytic primary granule genes. Scale bar and gene diagram are depicted above the captures. Yellow shades highlight promoter peaks.

Figure S3. Affiliated with Figure 4.

(A) Correlation analysis of GRO-seq data from biological replicates of cells transfected with a scramble siRNA (ctrl), or siRNA targeting *NELF-B* or -*E*.

(B) Metagene analysis showing Pol II occupancy measured by GRO-seq on both sense and anti-sense strands in human $CD34^+$ HSPCs transfected with control-siRNA (ctrl) or siRNA targeting *NELF-B* (NB-KD).

(C) Cumulative distribution function analysis to compare PI distribution in cells transfected with a control-siRNA (ctrl), or siRNA targeting *NELF-B* (NB-KD).

(D) Metagene analysis showing Pol II occupancy measured by GRO-seq on genes upregulated ("up") and downregulated ("down") shared by NELF-B and NELF-D knockdown cells.

(E) Boxplot analysis to compare pausing index between upregulated ("up") genes and downregulated ("down") genes upon NELF depletion.

Figure S4. Affiliated with Figure 5.

(A) Genome browser captures of GRO-seq for granulocytic primary granule genes to compare transcription activity among HSPCs (d0), 3-day differentiated granulocytes

(Gra_d3) and NELF-depleted HSPCs (NB-KD_d0 and NE-KD_d0). Scale bar and gene diagram are depicted above the captures. Yellow shades highlight promoter peaks.

(B) Genome browser captures of GRO-seq for *JAK2* and *MAP2K1*.

(C) Quantitative RT-PCR of nascent RNA to compare expression of progenitor markers and granulocyte markers between control HSPCs and cells depleted with NELF-E. Gene expression is normalized to *β-actin* and presented as fold change relative to control cells (n=3, mean ± SEM, all with $p<0.05$).

(D) Venn Diagram showing the overlap of upregulated ("up") and downregulated genes ("down") between NELF-KD HSPCs and d3-granulocytes.

Fig. S1

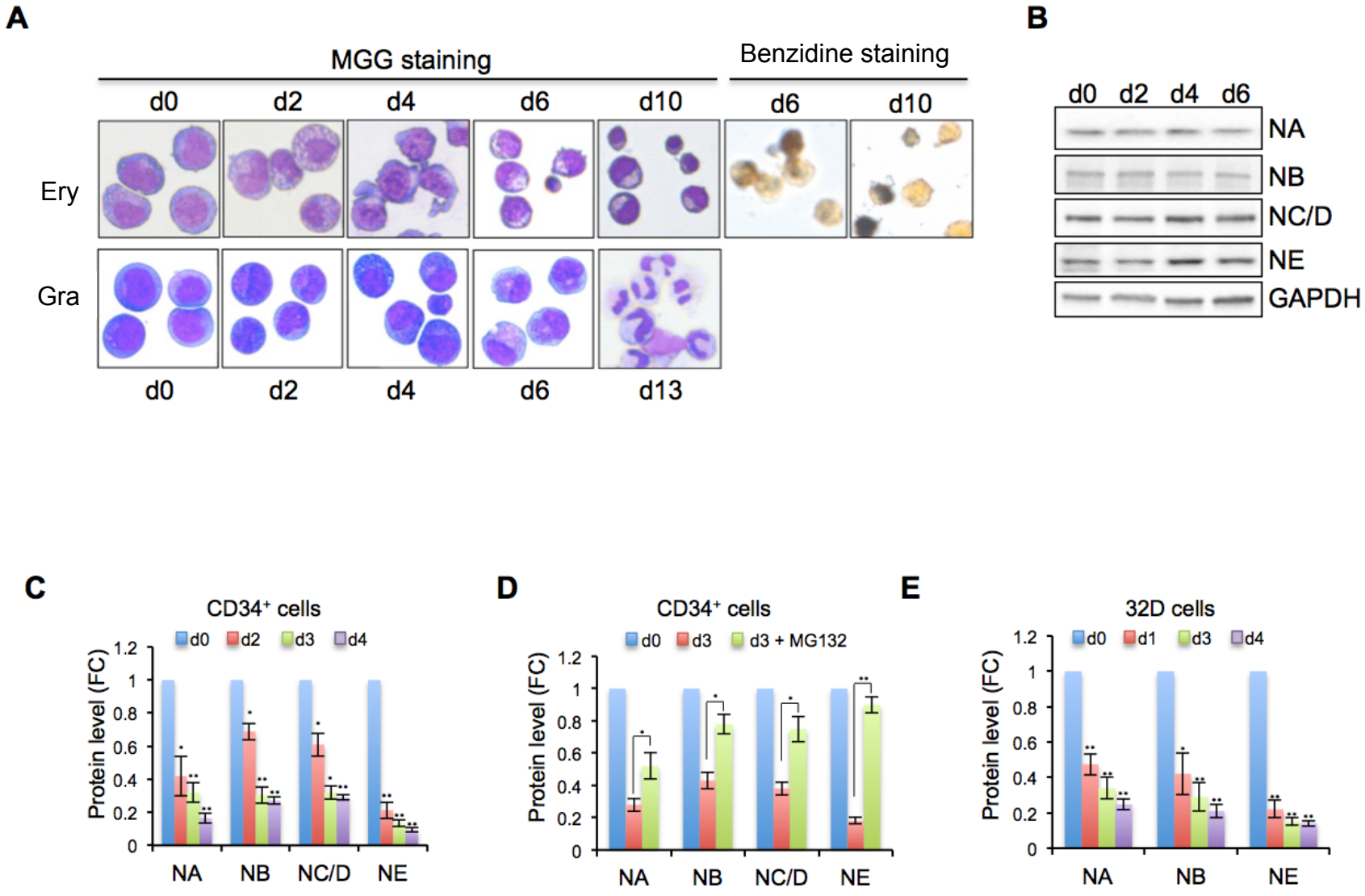

A
MGG staining
Benzidine staining
d0
d2
d4
d6
d10
d6
d10
Ery
Gra
d0
d2
d4
d6
d13
B
d0 d2 d4 d6
NA
NB
NC/D
NE
GAPDH
C
CD34⁺ cells
d0 d2 d3 d4
Protein level (FC)
NA NB NC/D NE
D
CD34⁺ cells
d0 d3 d3 + MG132
Protein level (FC)
NA NB NC/D NE
E
32D cells
d0 d1 d3 d4
Protein level (FC)
NA NB NE

Fig. S2

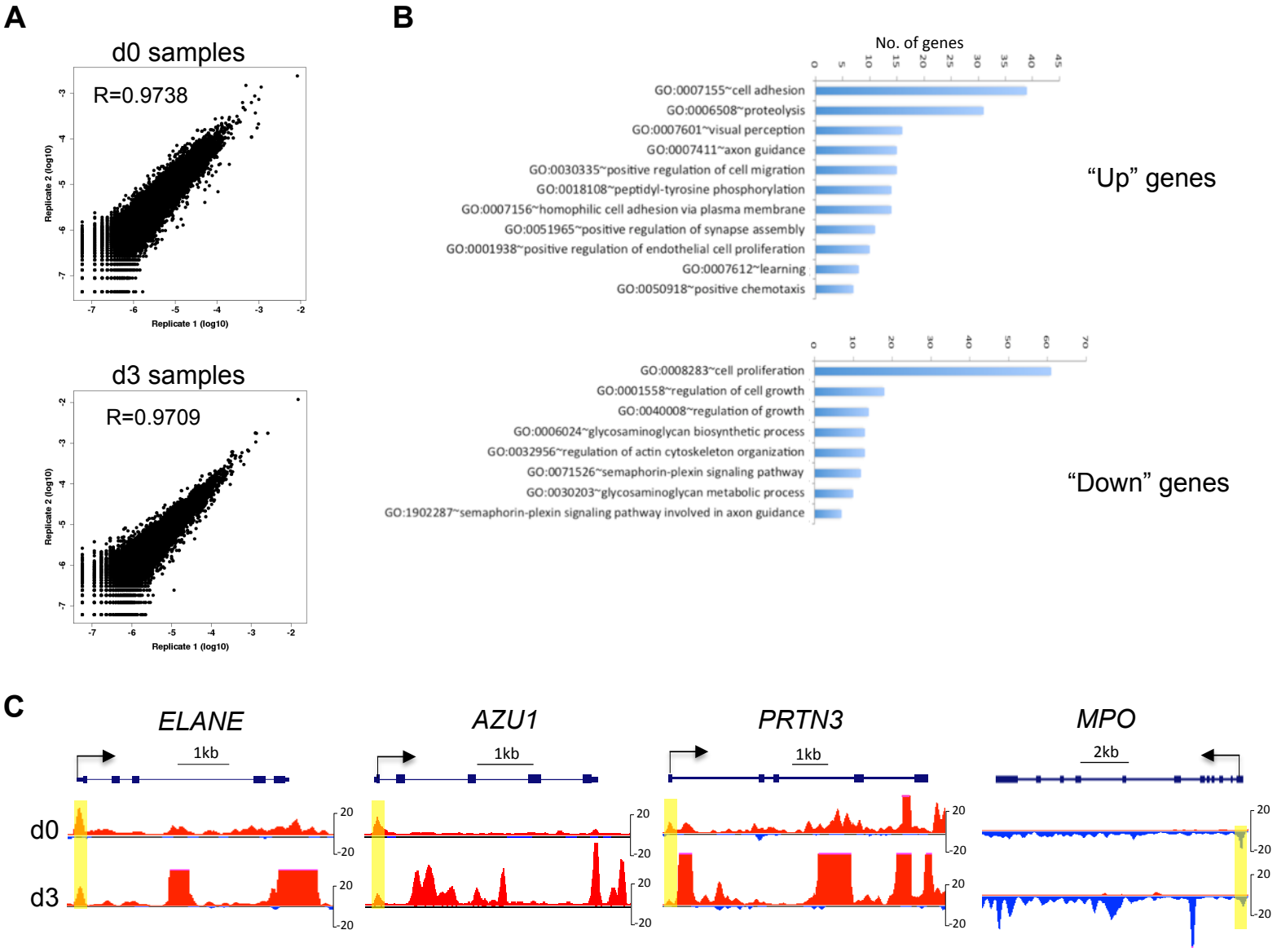

A
d0 samples
R=0.9738
Replicate 2 (log10)
Replicate 1 (log10)
d3 samples
R=0.9709
Replicate 2 (log10)
Replicate 1 (log10)
B
No. of genes
GO:0007155~cell adhesion
GO:0006508~proteolysis
GO:0007601~visual perception
GO:0007411~axon guidance
GO:0030335~positive regulation of cell migration
GO:0018108~peptidyl-tyrosine phosphorylation
GO:0007156~homophilic cell adhesion via plasma membrane
GO:0051965~positive regulation of synapse assembly
GO:0001938~positive regulation of endothelial cell proliferation
GO:0007612~learning
GO:0050918~positive chemotaxis
"Up" genes
GO:0008283~cell proliferation
GO:0001558~regulation of cell growth
GO:0040008~regulation of growth
GO:0006024~glycosaminoglycan biosynthetic process
GO:0032956~regulation of actin cytoskeleton organization
GO:0071526~semaphorin-plexin signaling pathway
GO:0030203~glycosaminoglycan metabolic process
GO:1902287~semaphorin-plexin signaling pathway involved in axon guidance
"Down" genes
C
ELANE
AZU1
PRTN3
MPO
1kb
1kb
1kb
2kb
d0
d3

Fig. S3

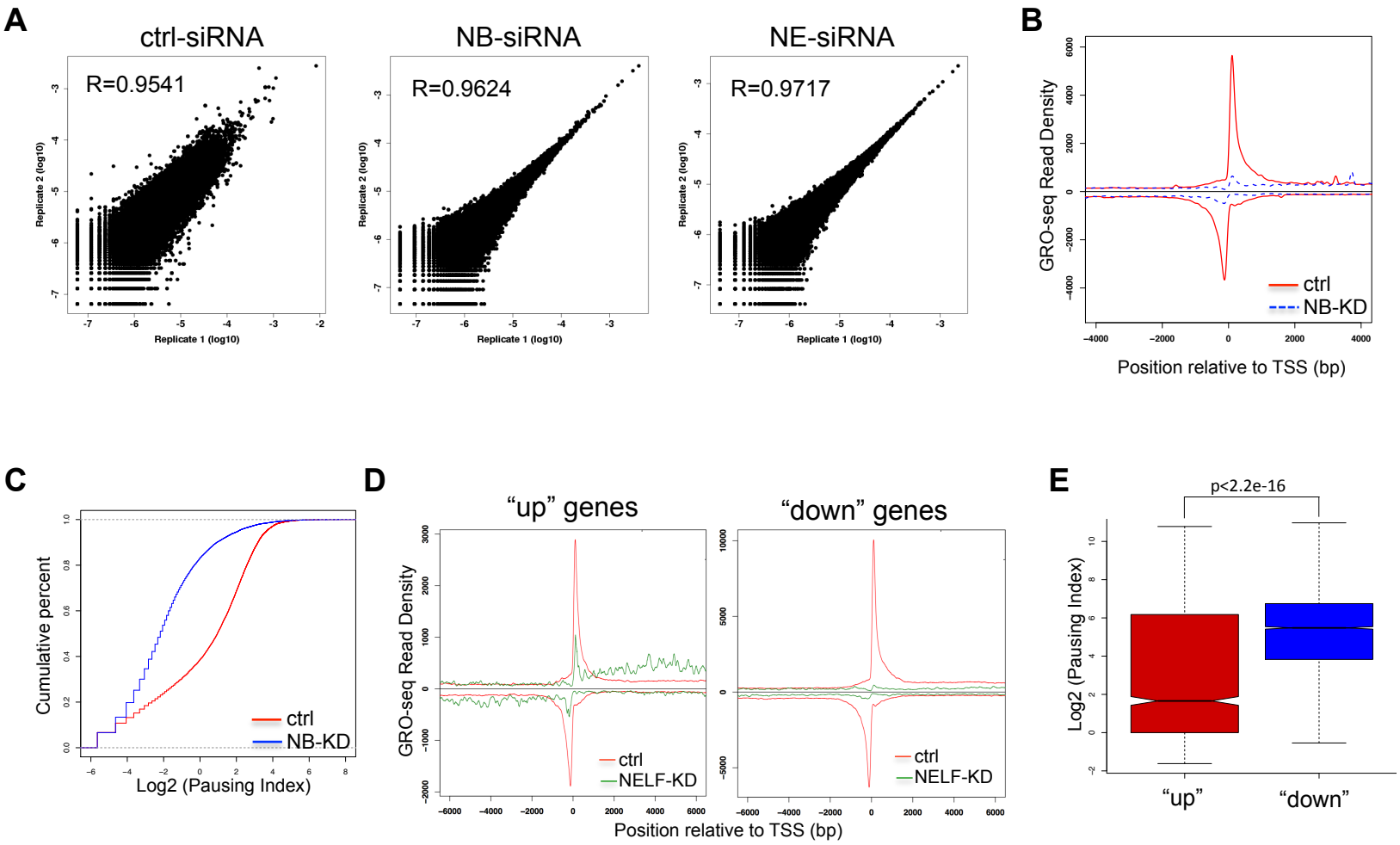
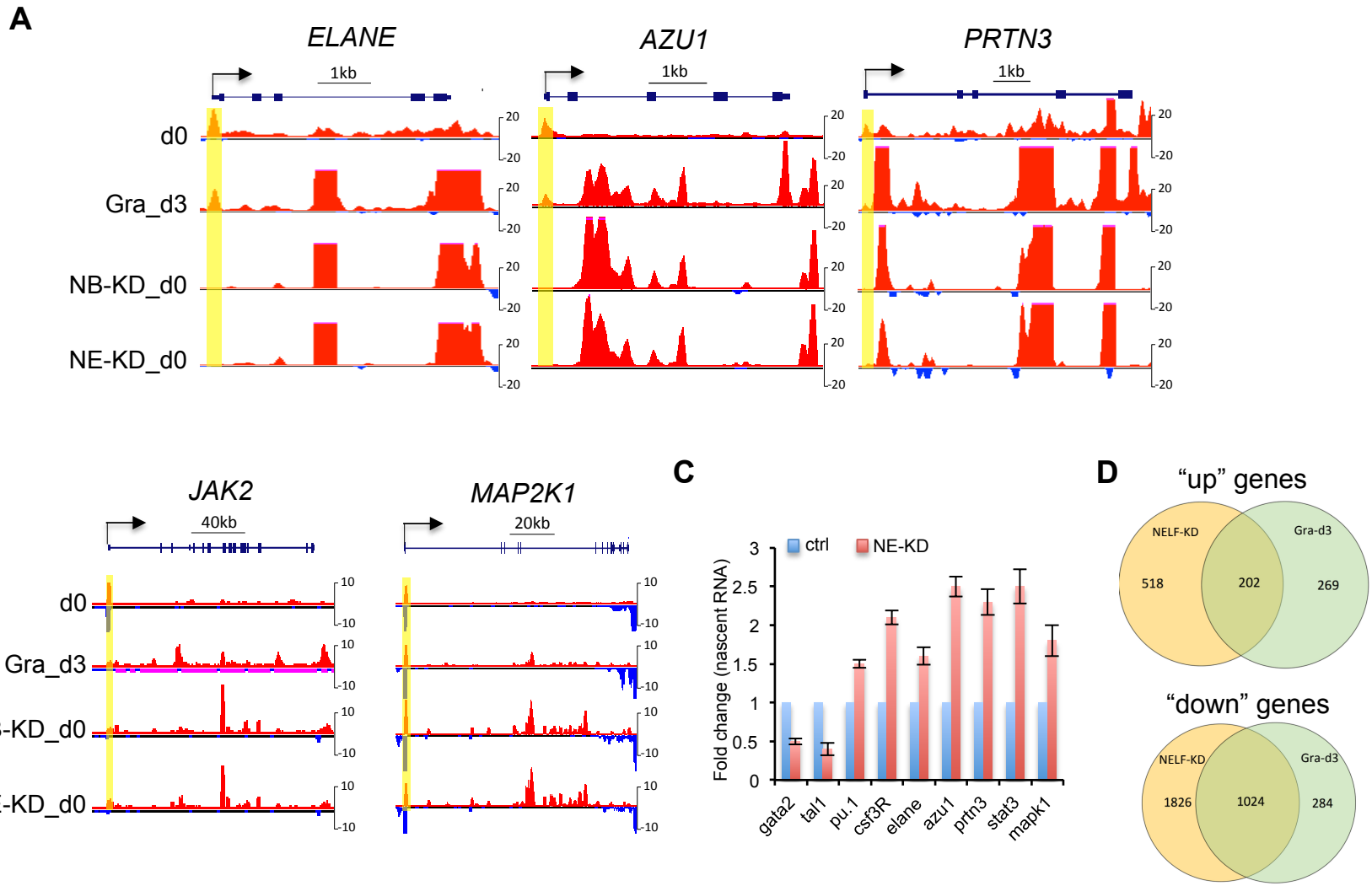

A
ctrl-siRNA
R=0.9541
NB-siRNA
R=0.9624
NE-siRNA
R=0.9717
Replicate 1 (log10)
Replicate 2 (log10)
B
GRO-seq Read Density
ctrl
NB-KD
Position relative to TSS (bp)
C
Cumulative percent
ctrl
NB-KD
Log2 (Pausing Index)
D
"up" genes
"down" genes
GRO-seq Read Density
ctrl
NELF-KD
Position relative to TSS (bp)
E
p<2.2e-16
Log2 (Pausing Index)
"up"
"down"

Fig. S4


A
ELANE
AZU1
PRTN3
1kb
d0
Gra_d3
NB-KD_d0
NE-KD_d0
B
JAK2
MAP2K1
40kb
20kb
C
ctrl
NE-KD
Fold change (nascent RNA)
gata2
tal1
pu.1
csf3R
elane
azu1
prtn3
stat3
mapk1
D
"up" genes
NELF-KD
Gra-d3
518
202
269
"down" genes
1826
1024
284