SUPPLEMENTARY METHODS

gRNA library design

gRNAs were designed using the CRISPick tool[1], with compatible BpiI/BbsI overhangs at the 5' end for cloning in a retroviral vector containing a 10x Genomics-compatible capture sequence (CS) for gRNA capture (CS1: GCTTTAAGGCCGGTCCTAGCAA, CS2: GCTCACCTATTAGCGGCTAAGG)[2] 100uM of forward and reverse gRNA oligos were annealed under following conditions; 95°C for 3 minutes, 25°C for 1 minute and holding at 4°C. Annealed gRNAs were pooled, diluted to 0.04uM and ligated into the BpiI-digested vector backbone. This construct was electroporated in Endura electrocompetent cells (Lucigen) according to the manufacturer's protocols. A minimum of 1000 bacterial colonies per gRNA in the library were harvested and DNA was isolated using the Compactprep Plasmid Midi kit (Qiagen) according to manufacturer's directions.

Retroviral work

Retrovirus production was performed by transfection of 150 mm dishes HEK293T cells with 10ug of the retroviral vector of interest and 10ug packaging plasmid pIK6.1MCV.ecopac.UTD, using 36uL Genejuice transfection reagent (Merck Millipore #70767) in 600uL serum-free RPMI-1640 medium. After 24h, medium was replaced with fresh RPMI-1640 supplemented with 10% FBS and retroviral supernatant was subsequently harvested at 48h and 72h post-transfection. Viral supernatant was filtered through a 0.45um filter and aliquots were frozen at -80°C. Viral titer was calculated to determine the optimal transduction conditions to reach around 10% efficiency to avoid introduction of multiple gRNAs in a single cell. Serial dilutions of the virus stock were used to transduce 1 million cells by spinfection and transduction efficiencies were measured by flow cytometry (MACSQuant VYB, Miltenyi). The number of transducing units per mL was determined and used to calculate the optimal volume of virus stock that should be used to reach the optimal MOI.

Data analysis of single-cell CRISPR experiments

CellRanger software (version 5.0.1) was used for alignment of reads to the mouse reference genome (mm10) and unique molecular identifier (UMI) counting. The resulting count matrices were further processed using the Seurat R package v4. We removed cells that expressed less than 200 genes and genes that were expressed in less than 3 cells. Furthermore, cells with high numbers of mitochondrial reads (>15%) and extreme numbers of transcripts (<1500 and >8000) were removed, as these are most likely damaged cells or doublets. Next, the data was normalized using SCTransform. Cell cycle scoring was done with Seurat based on the lists of cell cycle marker genes of Tirosh *et al*. [3] Differential gene expression analysis using the MAST algorithm was performed to identify differentially expressed

marker genes between cells with a gRNA and NT cells. Furthermore this data was analyzed with a number of software packets specifically designed for single cell CRISPR screening, such as Mixscape[4], scMAGeCK[5] (both in RRA and LR mode), SCEPTRE[6] and MUSIC[7]. In an alternative approach, cells with the same perturbation were grouped and a global transcriptional profile per perturbation was determined. DESeq2 (version 1.34.0) was used to compare these pseudo-bulk transcription profiles and identify differentially expressed genes.

As we were mostly interested in the genes of importance in the pro-T cell model, we focused on the five sets identified by Bornschein *et al.*[8], with which we performed gene set enrichment analyses (GSEA) on the ranked list of differentially expressed genes between the cells with a perturbation and the non-perturbed cells. The rank was calculated as: -sign(log2FoldChange)*log(padj). Based on these normalized enrichment scores (NES), an unsupervised hierarchical clustering algorithm was applied to identify perturbations with similar impact on the cells. To determine the regulatory elements involved in each of the perturbations, the pySCENIC software package (version 0.10.0), together with its command line interface (CLI), was used. First a list of adjacencies between transcription factors and its targets was determined with the CLI version on the Flemish Super Computer (VSC). Subsequently, we used the publicly available motif database to predict regulons, after which cellular enrichment for each regulon in each cell was calculated with AUCell. The distributions of these AUCell values for each regulon and each perturbation were then plotted with the stacked violin function in the Scanpy package.

Bulk RNA-sequencing

The Maxwell Simply RNA purification kit (Promega) was used to extract RNA from sorted pro-T cells, from which 500ng of purified RNA was used to construct NGS libraires using the Lexogen Quantseq RNA sample prep kit (Illumina). Libraries were subjected to 51bp single-end sequencing on a HiSeq4000 (Illumina). The reads were cleaned with the fastq-mcf software package (version 1.04.807) and FastQC (version 0.11.9) was used for quality control. The reads were subsequently aligned to the mouse reference genome (mm10/GRCm38) using HISAT2 (version 2.1.0) and further processed with the SAMtools software (version 1.11). The number of reads per gene was counted with the HTSeq package. The R (version 4.1.1) package DESeq2 (version 1.34.0) was used for differential gene expression analysis and the SCENIC software package (version 1.3.1) was used to determine regulons, *i.e.*, small gene regulatory networks based on co-expression patterns and transcription factor motifs.

Intracellular staining

Intracellular stainings were performed using the FOXP3 Transcription Factor Staining buffer set (Invitrogen 00-5523-00). Briefly, 100 000 cells were seeded in a well of a 96-well plate, stained with fixable viability dye, washed and resuspended in fixation/permeabilization buffer for fixation for 45 minutes. Cells were subsequently stained with primary and secondary antibody in triplicate, and data were acquired on a MACSQuant VYB (Miltenyi). Primary antibodies used were anti-Bcl11b (Abcam ab18465, 1/200 dilution) and anti-PU.1 (Abcam ab227835, 1/100 dilution), with secondary antibodies Goat anti-Rat IgG Secondary Antibody in Alexa Fluor 555 (Thermofisher A21434, 1/1000 dilution) and Donkey anti-Rabbit IgG Secondary Antibody in Alexa Fluor 568 (Thermofisher A10042, 1/1000 dilution). Fixable viability dyes eFluor520 (eBioscience 65-0867-14, 1/1000 dilution) and eFluor780 (eBioscience 65-0865-14, 1/1000 dilution) were used to discriminate living cells.

Western blot

Cell lysates were prepared in cold Cell Lysis Buffer (Cell signaling) containing protease inhibitor (Complete – EDTA-free, Roche) and 5 mM Na_3VO_4. Protein concentration was determined with Protein Assay Dye Reagent Concentrate (Bio-Rad) using the Victor X4 plate reader (Perkin Elmer) according to the manufacturer's directions. Proteins were separated on NuPAGE NOVEX Bis-Tris 4 - 12% gels (Life Technologies) and transferred to PVDF membranes. Western blot analysis was performed using primary antibodies targeting P65 (Cell signaling 8242S, 1/1000 dilution) and phospho-P65 (Cell signaling 3033S, 1/1000 dilution). Western blot detection was performed with secondary antibodies conjugated with horseradish peroxidase (GE Healthcare). Images were acquired using the VILBER FUSION FX6 imager (Analis), bands were quantified using ImageJ software and normalized to the intensity of the loading control beta-actin.

REFERENCES

1. Doench JG, Fusi N, Sullender M, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol 2016;34(2):184–191.
2. Replogle JM, Norman TM, Xu A, et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. Nat Biotechnol 2020;38(8):954–961.
3. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science (1979) 2016;352(6282):189–196.
4. Papalexi E, Mimitou EP, Butler AW, et al. Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens. Nat Genet 2021;53(3):322–331.
5. Yang L, Zhu Y, Yu H, et al. scMAGeCK links genotypes with multiple phenotypes in single-cell CRISPR screens. Genome Biol 2020;21(1):19.
6. Barry T, Wang X, Morris JA, Roeder K, Katsevich E. SCEPTRE improves calibration and sensitivity in single-cell CRISPR screen analysis. Genome Biol 2021;22(1):344.

7. Duan B, Zhou C, Zhu C, et al. Model-based understanding of single-cell CRISPR screening. Nat Commun 2019;10(1):2233.

8. Bornschein S, Demeyer S, Stirparo R, et al. Defining the molecular basis of oncogenic cooperation between TAL1 expression and Pten deletion in T-ALL using a novel pro-T-cell model system. Leukemia 2018;32(4):941–951.

SUPPLEMENTARY FIGURES

Figure S1.

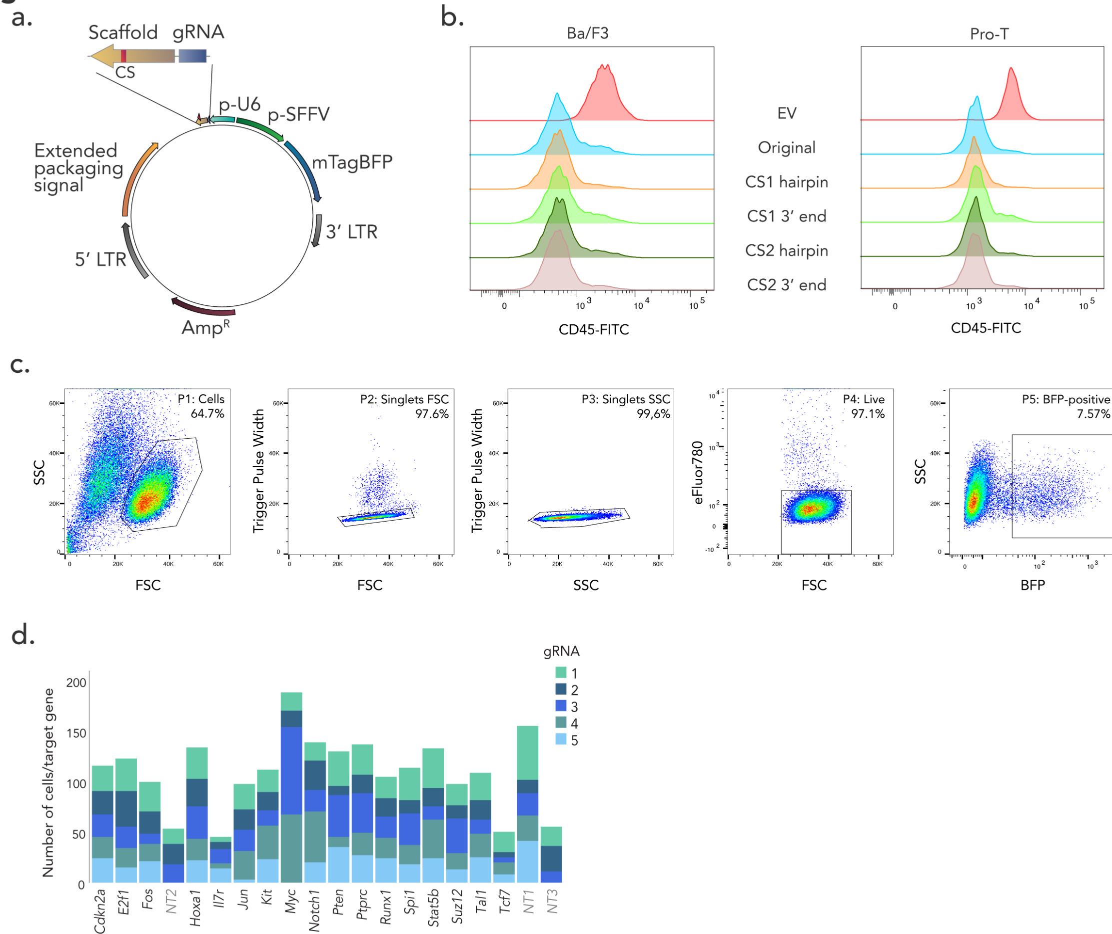

Fig. S1. Experimental details of the pilot CRISPR screen

a. Vector graphic of the retroviral gRNA vector with capture sequence (CS) in the hairpin of the gRNA scaffold. This capture sequence is required for gRNA capture by the 10X Genomics 3' feature barcoding chemistry for single-cell CRISPR screening. Expression of the gRNA is controlled by U6 promoter and expression of mTagBFP fluorescent marker by the spleen focus-forming virus (SFFV) promoter sequence. (LTR= long terminal repeat, Amp^R= ampicillin resistance marker)

b. CD45 protein levels were measured by flow cytometry in Ba/F3 and pro-T cells expressing a *Ptprc*-targeting gRNA to validate the efficiency of four different capture sequence (CS) vector configurations, compared to the original vector without capture sequence and empty vector (EV).

c. Gating strategy for FACS sorting of pro-T cells transduced with the gRNA library. First, cells were discriminated from debris based on forward (FSC) versus side (SSC) scatter (P1), with

subsequent gating for singlets (P2 and P3) and viable cells based on eFluor780 fixable viability dye (P4). Finally, cells with expression of the BFP fluorescent marker were selected (P5).

d. Stacked bar plot displaying the number of cells per gRNA per target gene at day 3 of the experiment. NT controls are indicated in grey font.

Figure S2.

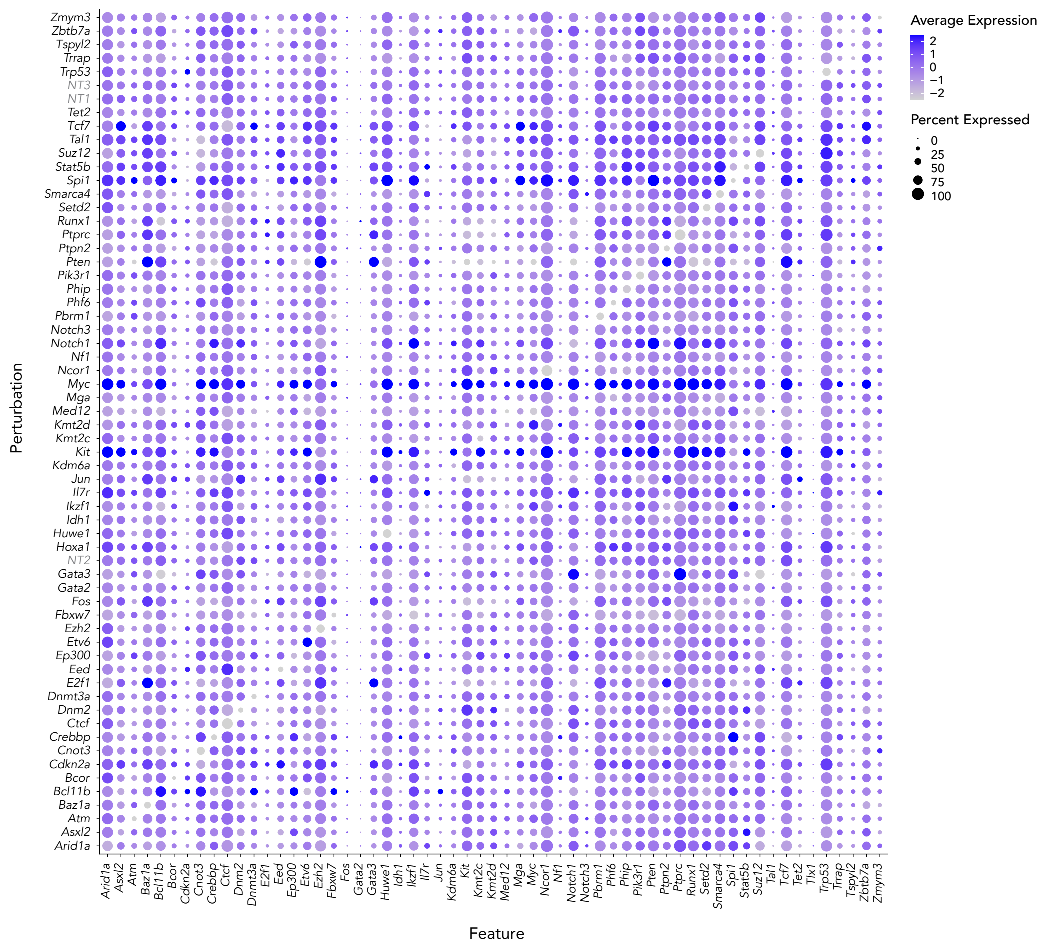

Fig. S2. Gene expression levels per perturbation

Data from both single-cell CRISPR screens were merged, and the average expression level of each target gene was displayed in a dot plot to visualize perturbation efficiency in the cells with different gRNAs.

Figure S3.

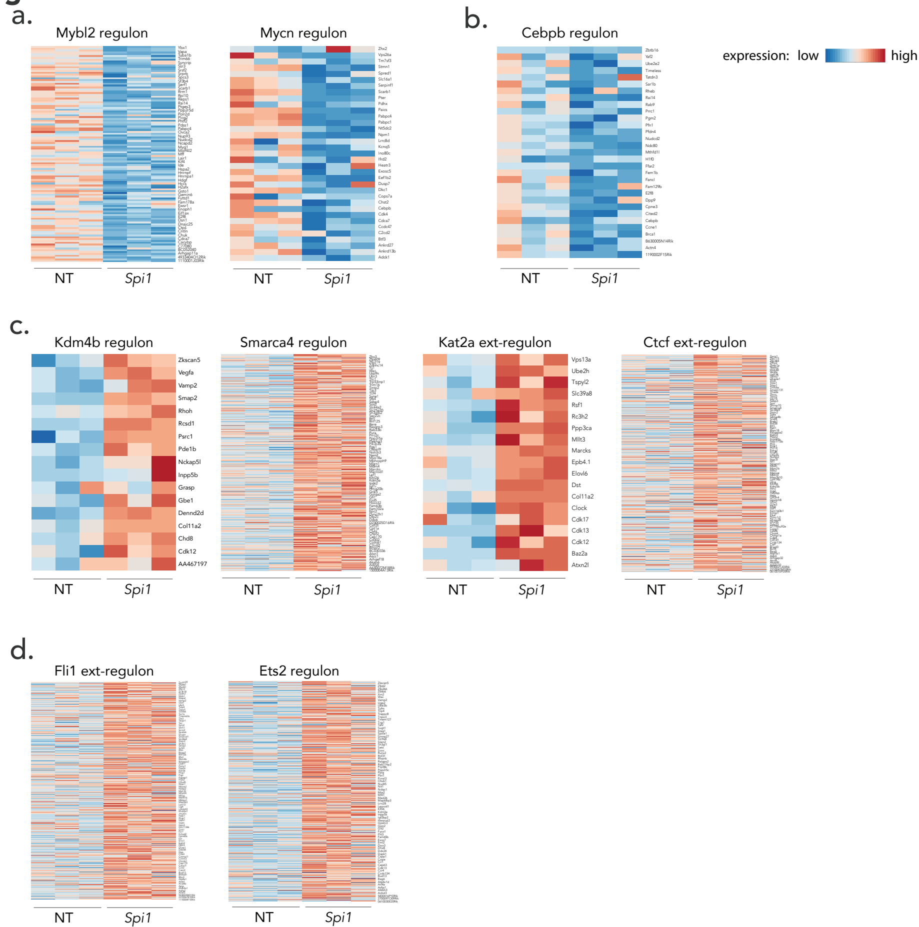

Fig. S3. SCENIC analysis of bulk RNA-seq after *Spi1* perturbation

a. SCENIC heatmaps for MYC regulons in cells with *Spi1* inactivation *vs* NT control cells.

b. SCENIC heatmaps for Cebpb regulon in cells with *Spi1* inactivation *vs* NT control cells.

c. SCENIC heatmaps for epigenetic regulons in cells with *Spi1* inactivation *vs* NT control cells.

d. SCENIC heatmaps for Ets factor regulons in cells with *Spi1* inactivation *vs* NT control cells.

Figure S4.

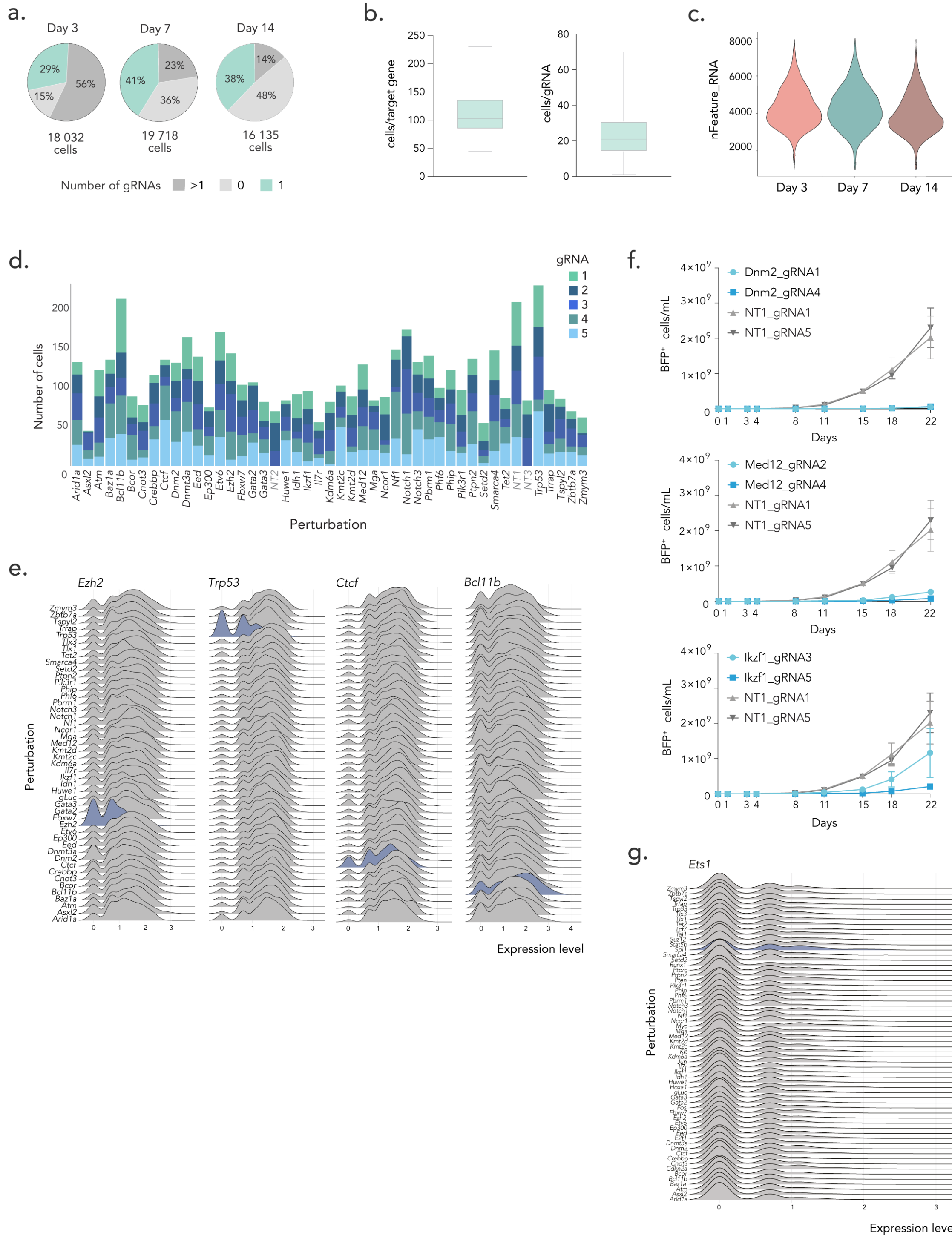

Fig. S4. Experimental details of the second CRISPR screen

a. Pie charts show percentages of cells with zero, one or multiple gRNAs in the second CRISPR screen, with indication of the total number of cells passing filtering parameters at each timepoint.

b. Number of cells per gene and per gRNA. Box plots showing the number of sequenced cells passing filter criteria per gene (left panel) and per gRNA (right panel) in CRISPR screen 2. Plots

show a line at the median and box from Q1 to Q3, with whiskers ranging from min to max values.

c. Violin plots displaying the number of detected transcripts in all single cells for each timepoint.
d. Stacked bar plot displaying the number of cells per gRNA per target gene at day 3 of the experiment. NT controls are indicated in grey font.
e. Normalized expression levels of the target genes in cells with the respective gRNA. Similar to the first single-cell CRISPR screen, all cells of the three timepoints were aggregated per perturbation to create ridge plots of the expression level of each target gene. On the top of the graph is indicated for which transcript the expression level is displayed, with the purple ridge highlighting the corresponding perturbation.
f. Proliferation effects of *Dnm2*, *Med12* and *Ikzf1* gRNAs. To confirm the effects observed in the CRISPR screens, individual proliferation assays were performed using two out of the five gRNAs per target gene and compared to two NT gRNAs. The cell density as number of BFP-positive cells (mean with SD) was followed over time.
g. Normalized expression level of *Ets1*. All cells of both CRISPR screens were aggregated per perturbation to create ridge plots of the normalized expression level of *Ets1*. The purple ridge highlights cells with *Spi1* gRNA.

Figure S5.

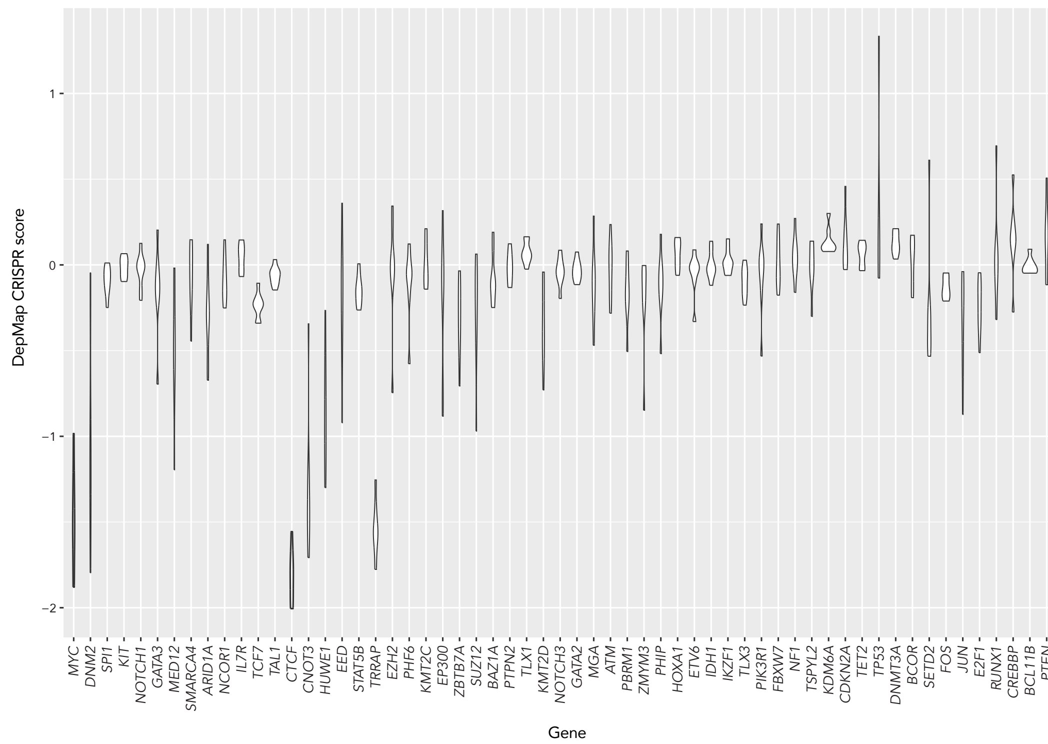

Fig. S5. Violin plots displaying CRISPR scores from the DepMap database for all perturbations in all lymphoid cell lines. Negative scores mean negative impact on cell proliferation, while genes with positive scores cause a proliferative advantage. Genes on x-axis were ordered based on proliferative effects in pro-T cells.

Figure S6.

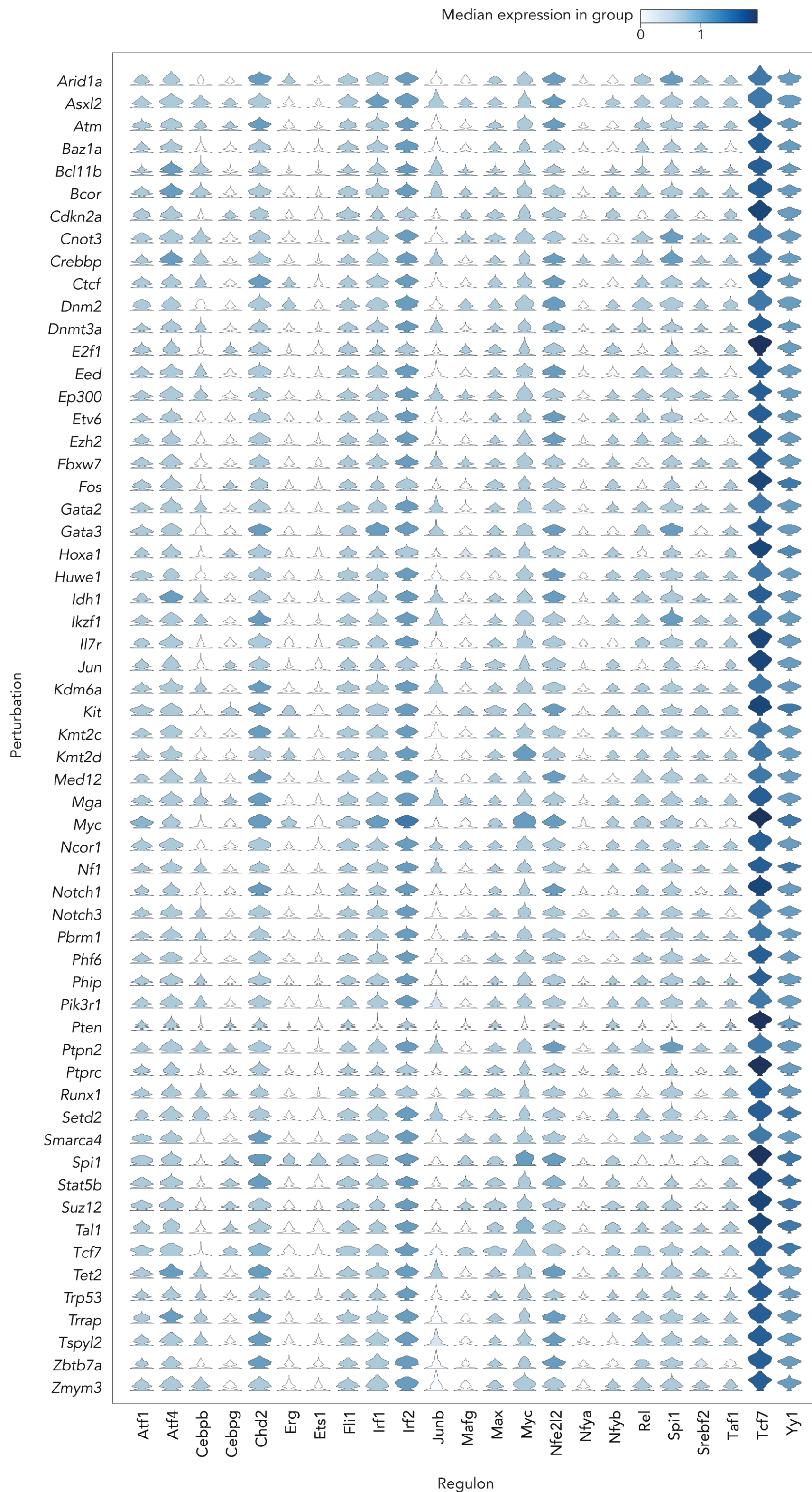

Fig. S6. Enrichment of SCENIC regulons per perturbation.

Stacked violin plot displaying the distribution of AUCell values for each regulon and each perturbation. AUCell values represent enrichment of each predicted regulon in each single cell. The color intensity is a measure for the median expression of each regulon in each perturbation.

SUPPLEMENTARY TABLES

Table S1: List of gRNA sequences used in the first CRISPR screen.

Table S2: Differentially expressed genes in bulk RNA-seq on cells with *Spi1* or *Bcl11b* inactivation, compared to NT cells.

Table S3: List of gRNA sequences used in the second CRISPR screen.