

SUPPLEMENTARY MATERIAL FOR A. CHAMPHEKAR ET AL.

A. Supplementary Figures S1-S8 and Figure Legends

B. Supplementary Methods and References

C. Supplementary Tables S1-S6

Table S1: DESeq analysis

Table S2: PU.1 binding sites and analysis

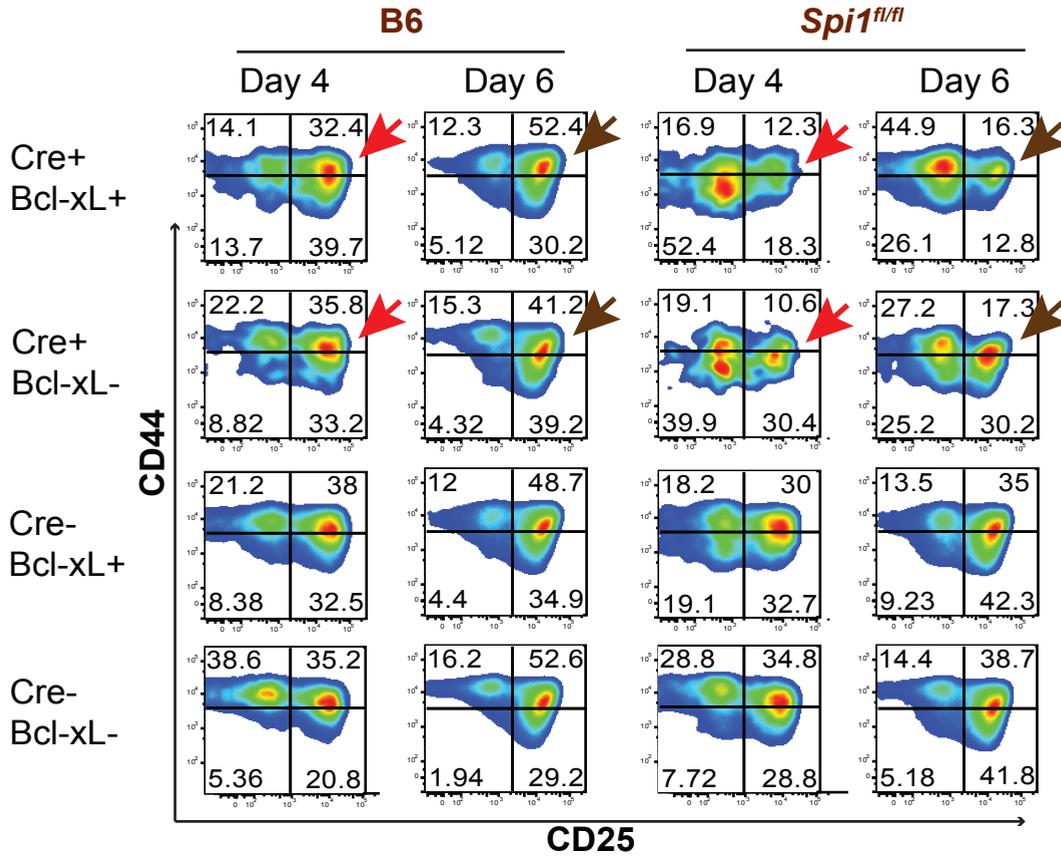
Table S3: PCA

Table S4: GSEA

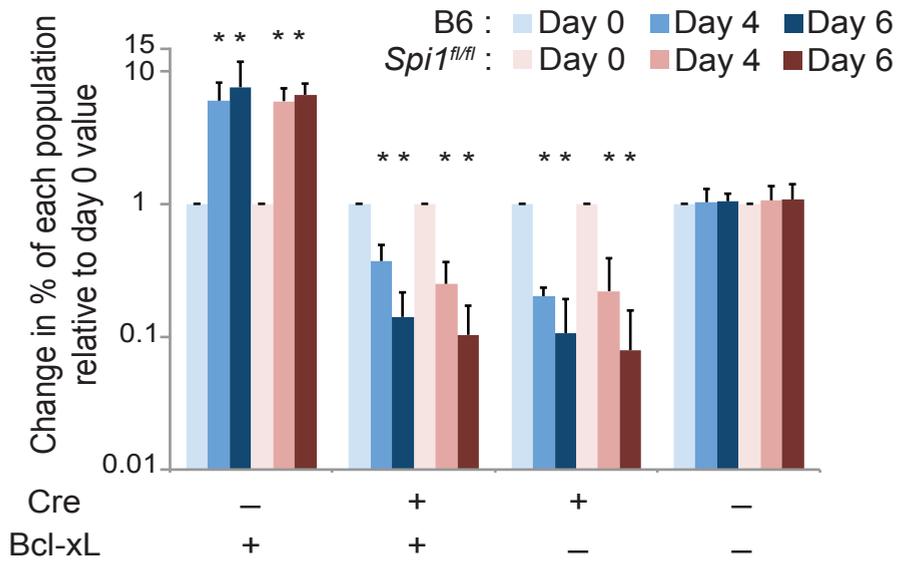
Table S5: GO/KEGG analysis

Table S6: List of qPCR primers

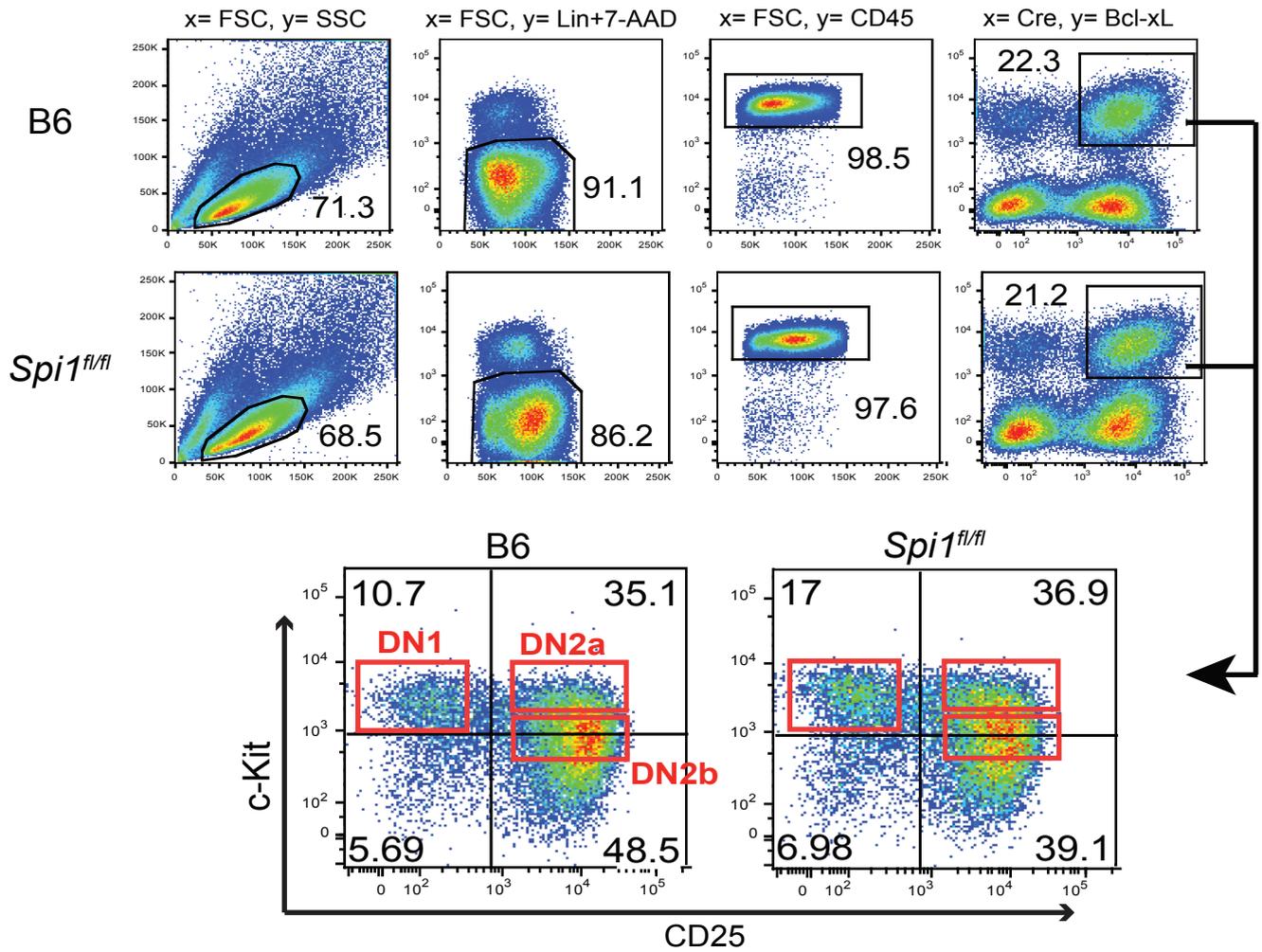
A.



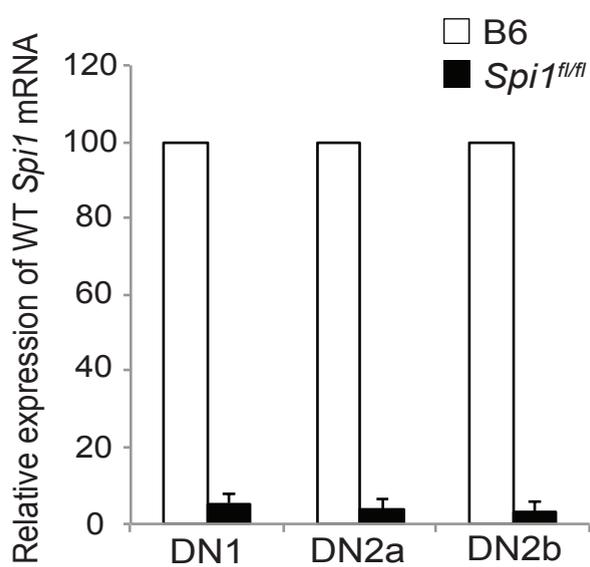
B.



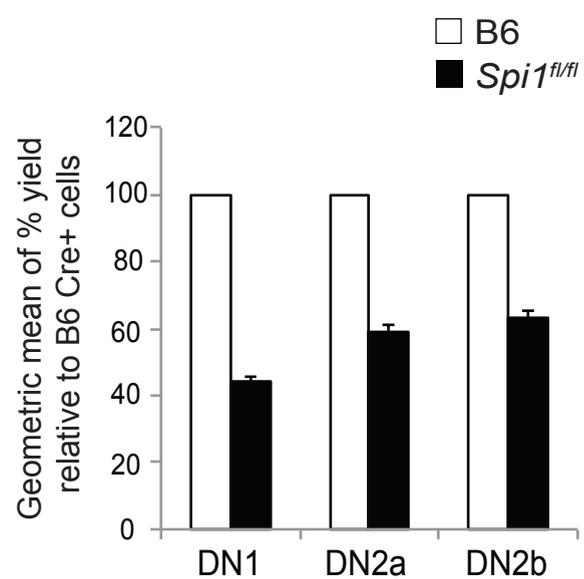
A



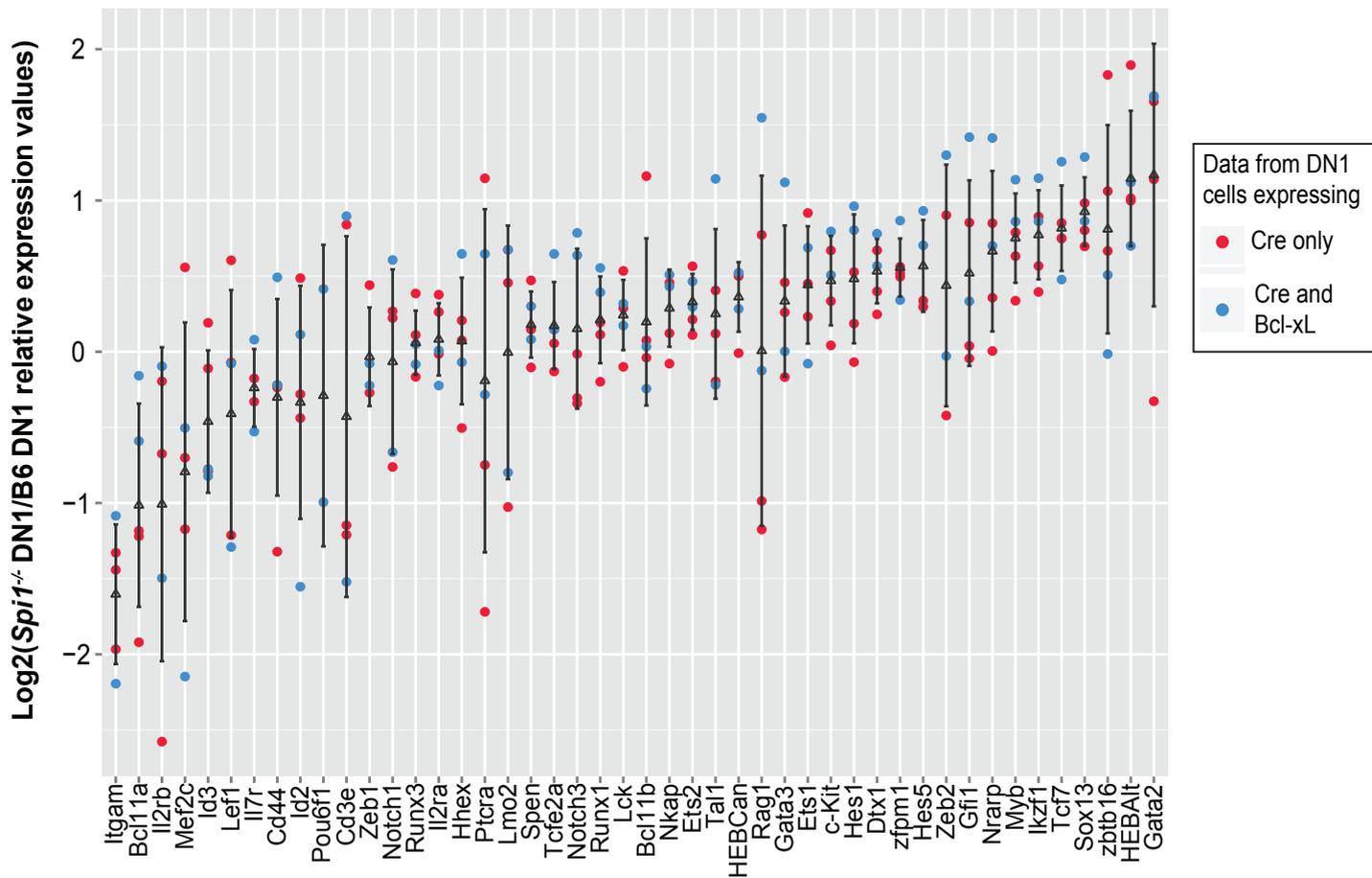
B



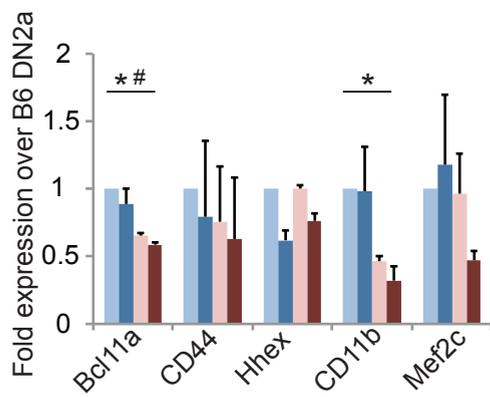
C



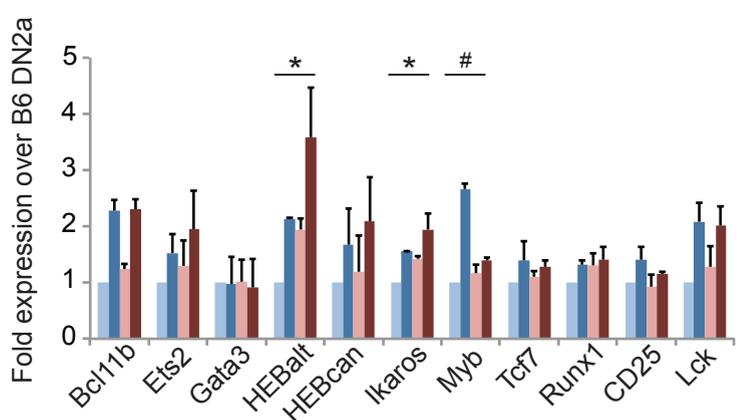
A



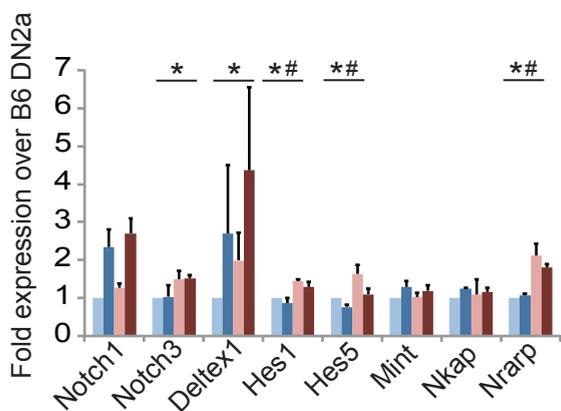
B



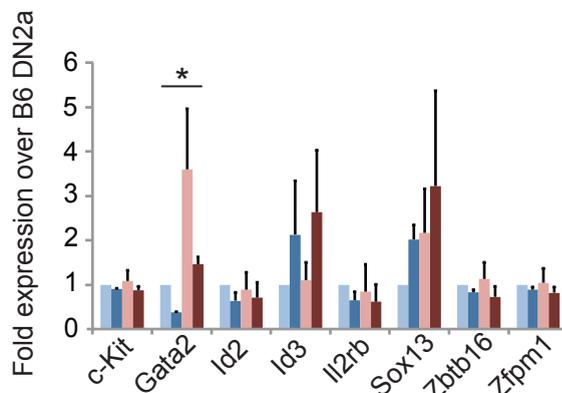
C



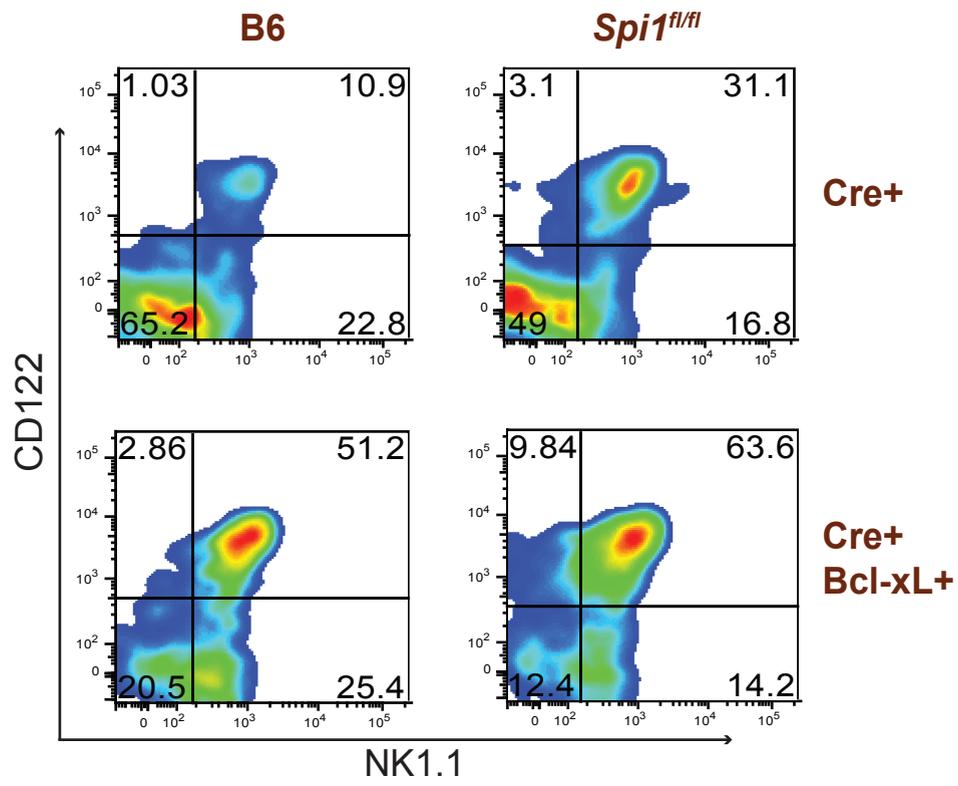
D



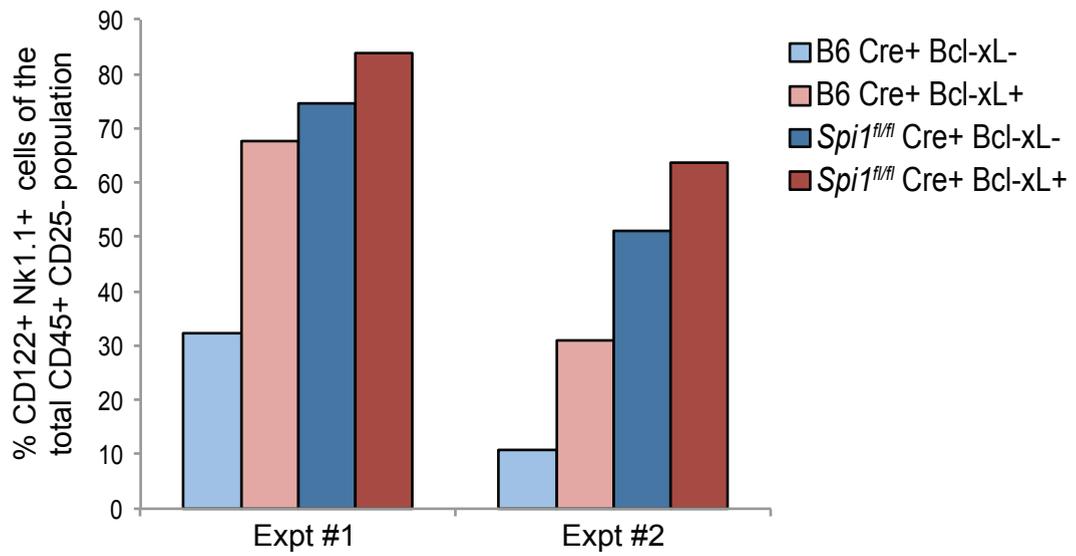
E



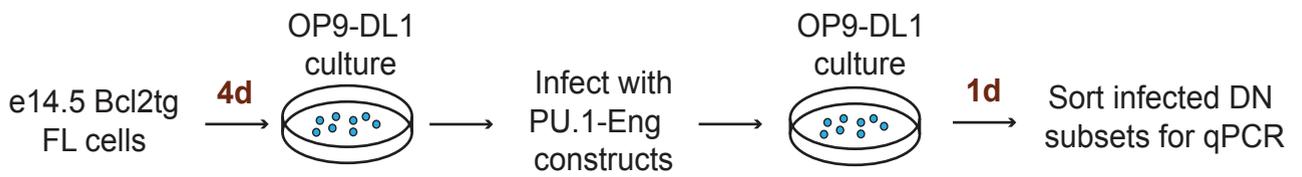
A



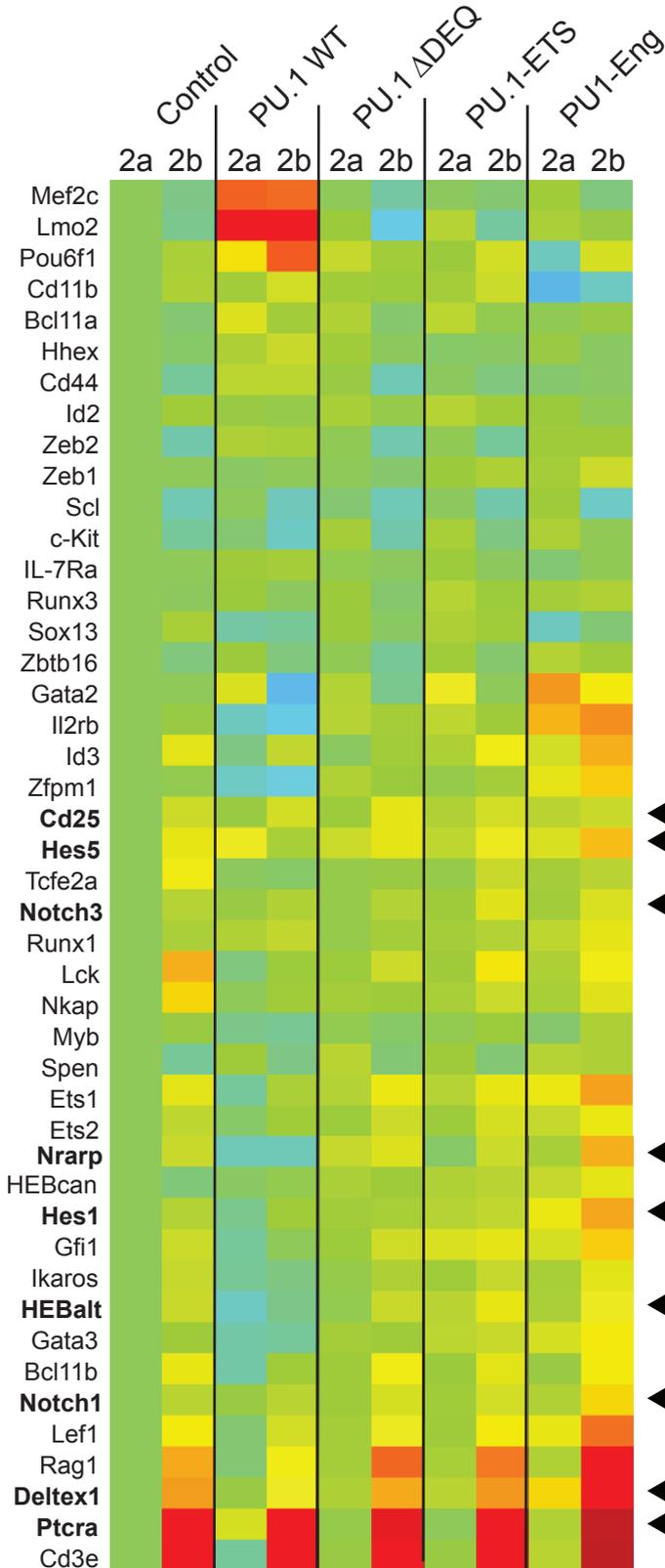
B



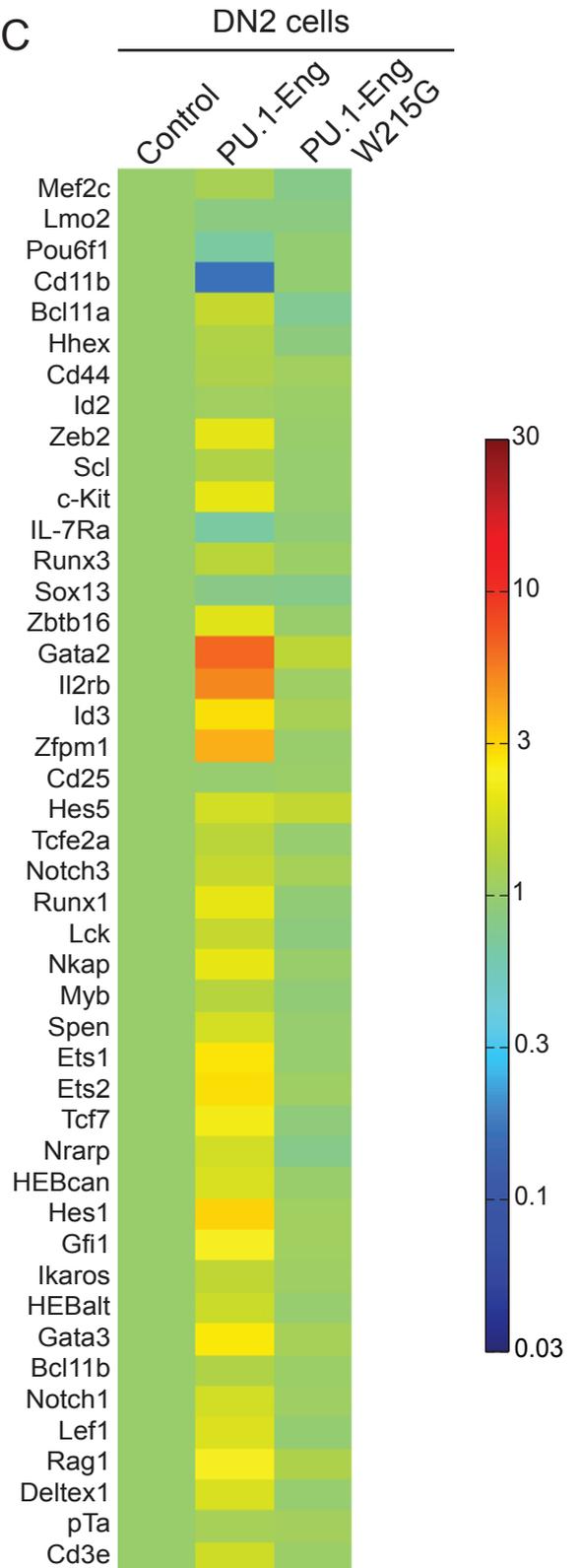
A

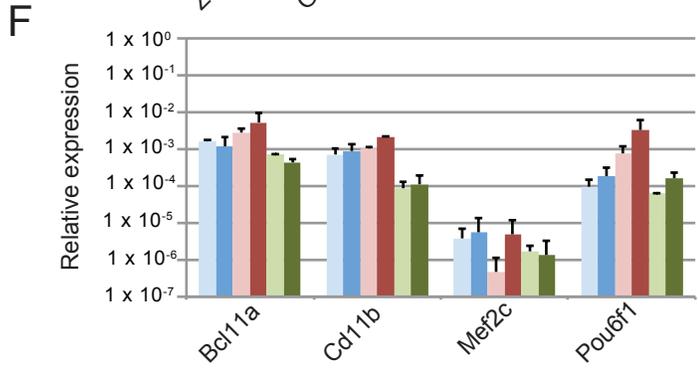
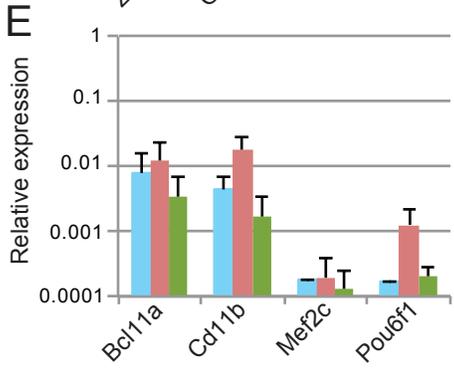
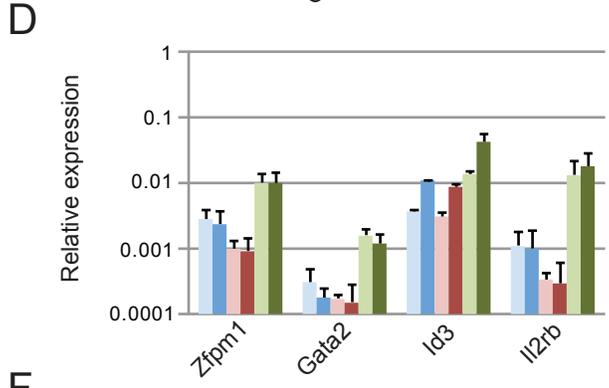
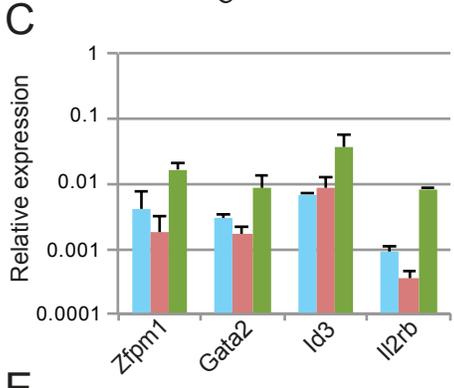
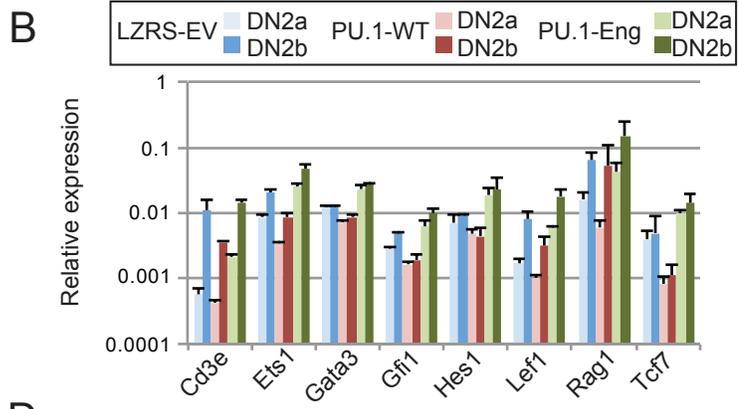
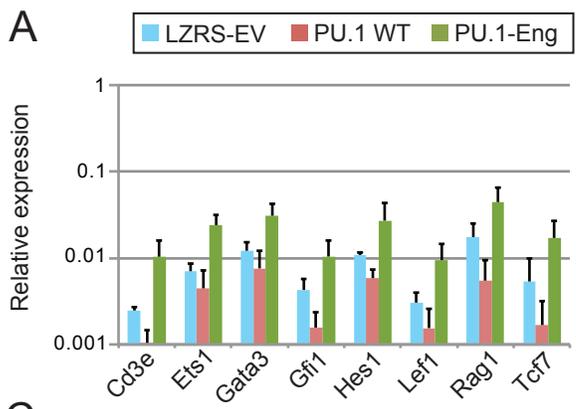


B



C

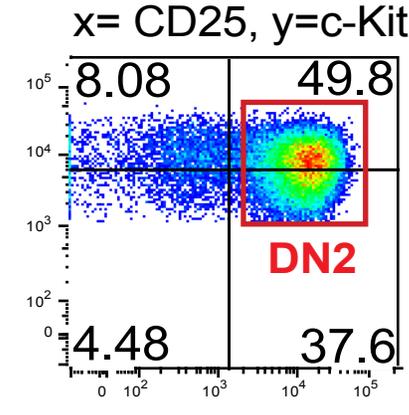
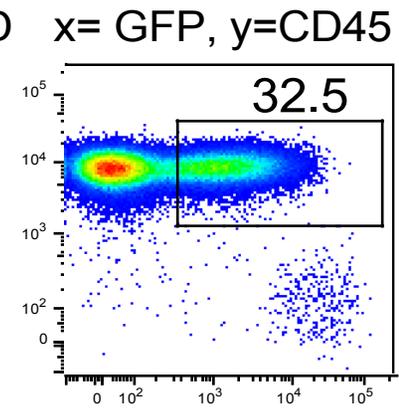
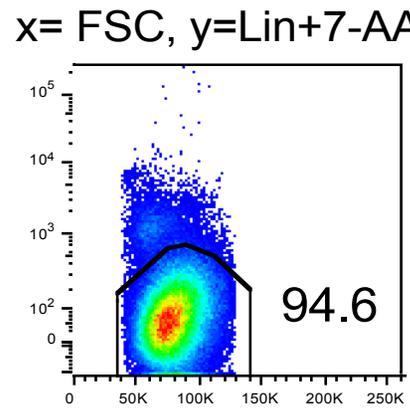
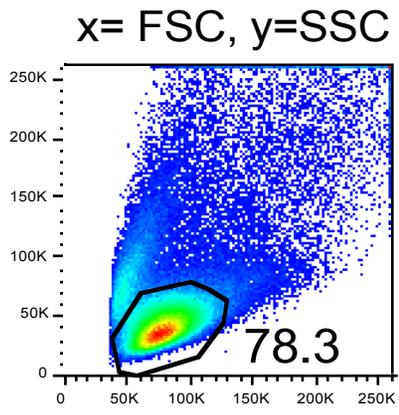




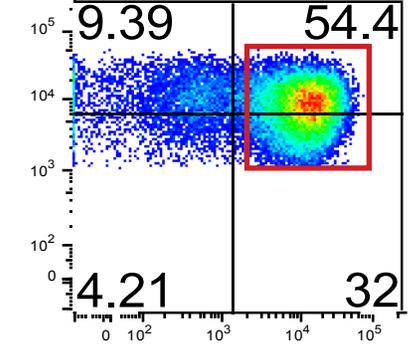
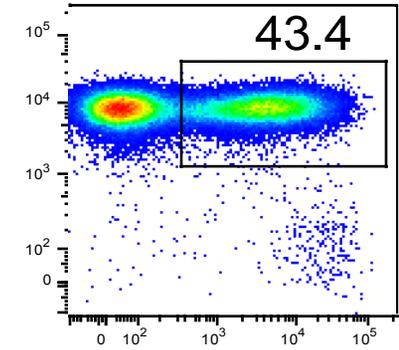
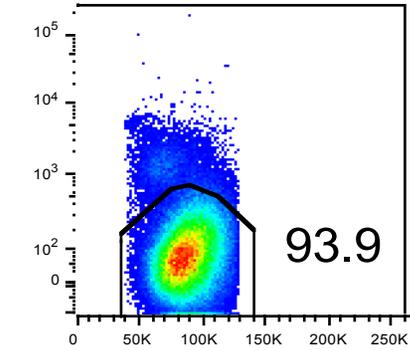
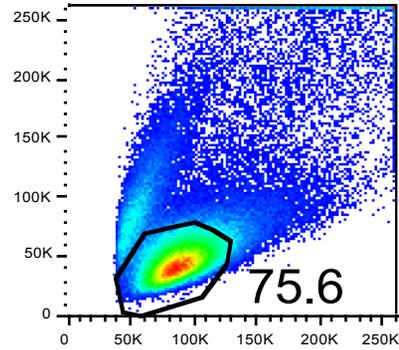
G

Gene name	Average fpkm (2 expts)			Fold effect		EV vs PU.1-ENG (padj)		EV vs PU.1-ETS (padj)	
	EV	PU.1-ENG	PU.1-ETS	PU.1-ENG/EV	PU.1-ETS/EV	Expt 1	Expt 2	Expt 1	Expt 2
Ccnd1	1.2	1.1	0.7	0.90	0.53	1	1	1	1
Ccnd2	46.5	49.6	37.7	1.07	0.81	1	1	1	1
Ccnd3	111.9	132.7	170.7	1.19	1.53	1	1	1	1
Ccne1	15.8	13.3	17.0	0.85	1.08	1	1	1	1
Ccne2	21.0	11.2	22.3	0.54	1.06	1	1	1	1
Cdk1	86.6	107.7	107.1	1.24	1.24	1	1	1	1
Cdk2	41.0	40.1	69.7	0.98	1.70	0.49	1	1	1
Cdk4	351.4	467.1	346.8	1.33	0.99	1	1	1	1
Cdk5	10.1	7.5	8.7	0.74	0.86	1	1	1	1
Cdk6	37.8	35.3	26.0	0.93	0.69	1	1	1	1
Cdkn1a	119.3	88.1	159.7	0.74	1.34	1	1	1	1
Cdkn1b	46.6	28.3	52.6	0.61	1.13	1	1	1	1
Cdkn1c	1.8	2.4	12.2	1.39	6.97	0.04	0.54	1	1
Cdkn2c	4.3	3.7	4.5	0.87	1.05	0.91	1	1	1
Cdkn2d	0.3	1.3	1.3	3.90	3.79	1	1	1	1
Cdkn3	7.7	6.1	9.1	0.79	1.18	1	1	1	1
Rb1	19.1	12.8	19.7	0.67	1.03	1	1	1	1
Trp53	116.2	131.3	130.6	1.13	1.12	1	1	1	1

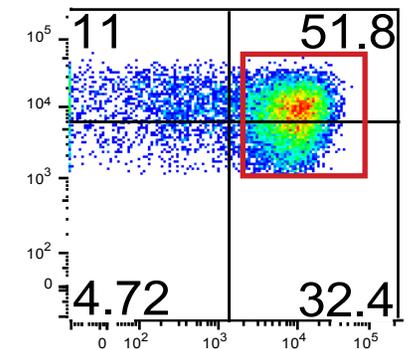
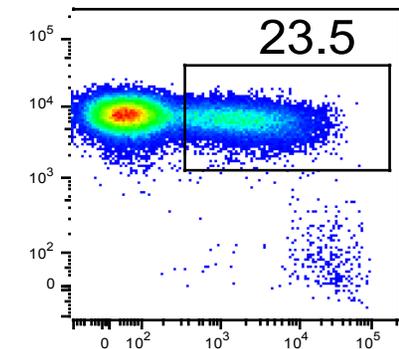
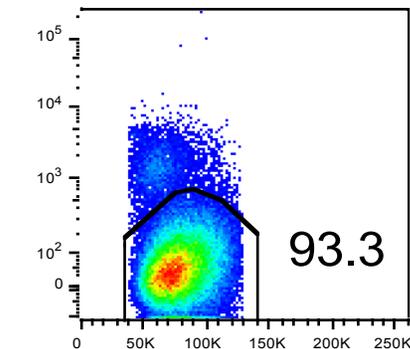
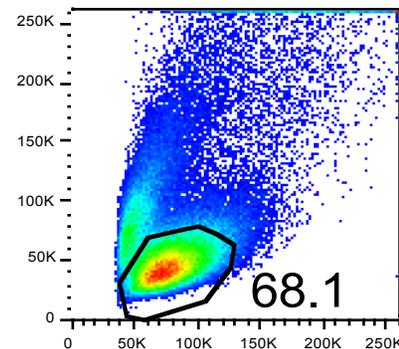
LZRS-EV

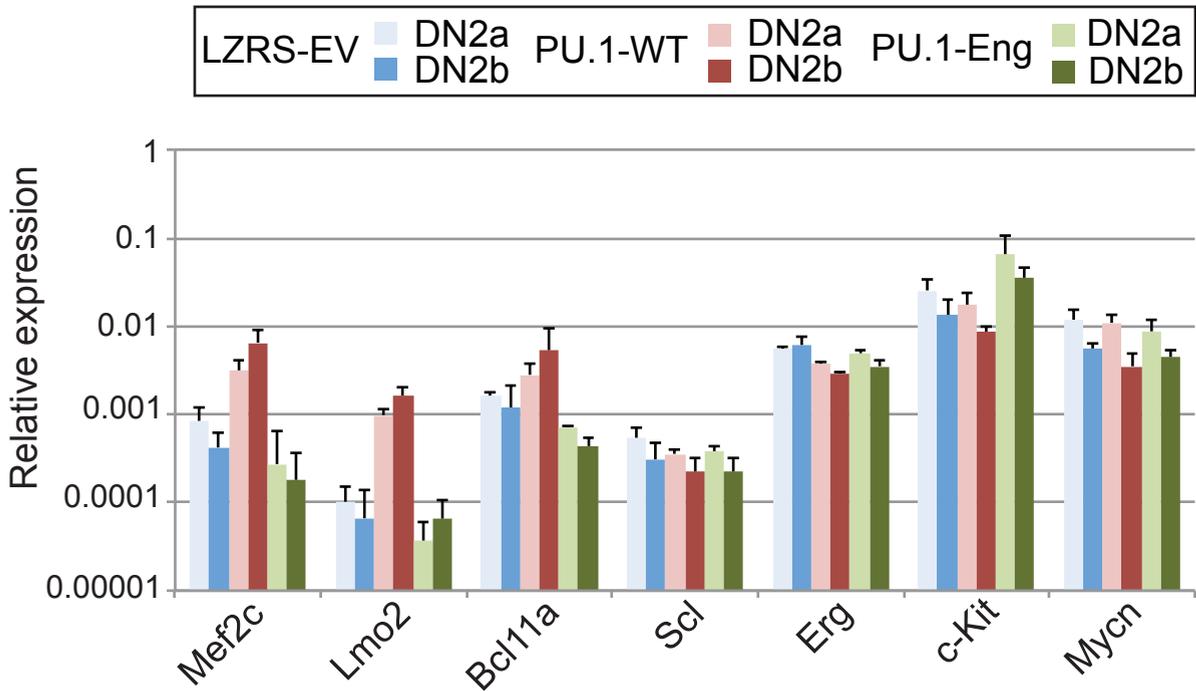
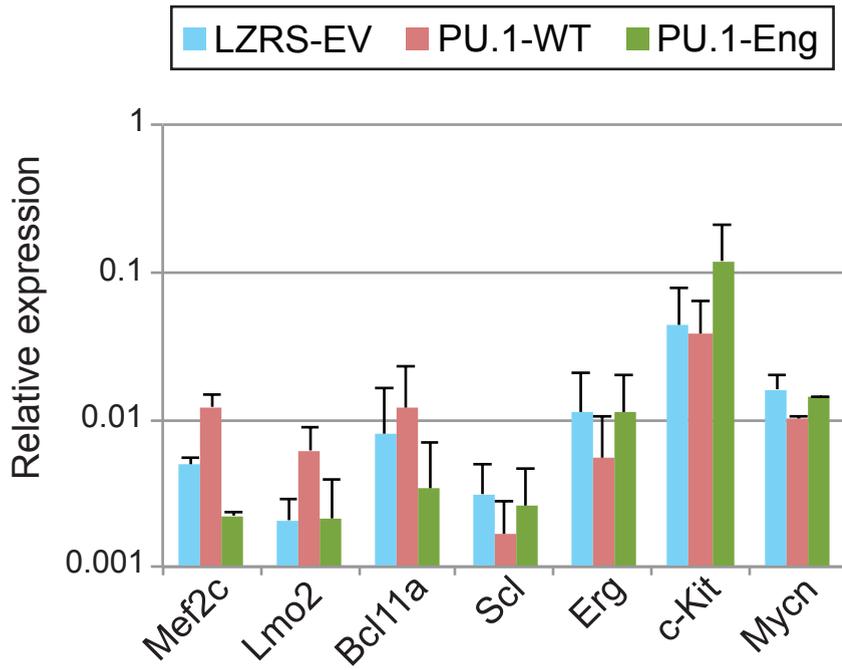


PU.1-ETS



PU.1-Eng





LEGENDS FOR SUPPLEMENTARY FIGURES

Figure S1 – *Deletion of PU.1 in c-Kit⁺ CD27⁺ FLPs results in impaired DN progression and poor survival and recovery of early DN-stage T cells.* E14.5 B6 and *Spi1^{fl/fl}* FLPs were co-infected with Cre- and Bcl-xL-carrying retroviruses or empty vector controls, and cultured on OP9-DL1 cells. **A:** DN progression was assayed on days 4 and 6 of culture by analysis of cells still expressing both retroviruses. FACS plots are representative of results from 2 independent experiments. Red and brown arrows indicate the comparable populations between B6 and *Spi1^{fl/fl}* cultures on day 4 and day 6 respectively. **B:** Plot showing change in the proportion of each double-infected population (GFP⁺ NGFR⁺) relative to its day 0 value. For example, a decline from 50% of the day 0 population to 5% of the day 6 population would be reported as “0.1”. Note the log scale. Each bar represents mean values from 2 independent experiments while error bars indicate range. * Indicates significant increase or decrease of a cell population compared to its day 0 value at $p < 0.05$.

Figure S2 – *Impacts of PU.1 deletion in cells that have already entered the T-cell development pathway.* **A:** *Sorting B6 and Spi1^{fl/fl} DN subsets.* E14.5 B6 and *Spi1^{fl/fl}* FLPs were co-cultured with OP9-DL1 as described in the methods section. Cells were harvested on day 3 of culture, transduced only with Cre or with Cre- and Bcl-xL-carrying retroviruses, and put back in fresh OP9-DL1 co-cultures. FLDN cells were harvested after 2 days and Lin⁻ 7-AAD⁻ CD45⁺ Cre⁺ Bcl-xL⁺ DN subsets were sorted using the gates indicated in the lower panels. FACS plots show the gating strategy and phenotype of retrovirally-transduced FLDN cells on the day of the sort. Numbers in the boxes show percentage in each quadrant or gated population. **B:** *Spi1^{fl/fl} is rapidly deleted by Cre transduction.* Cre mediated deletion of PU.1 in DN subsets sorted in **A** was determined immediately after the sort by measuring the undeleted WT message by qPCR. Bars represent mean expression values while error bars indicate 1 standard deviation. The graph summarizes data from 2 independent experiments. **C:** *PU.1 supports DN proliferation in early T-cells.* Cre⁺ B6 and *Spi1^{fl/fl}* DN1, DN2a and DN2b cells were sorted as in Fig 2F. Equal numbers of sorted Cre⁺ B6 and *Spi1^{fl/fl}* DN subsets were cultured on fresh OP9-DL1 cells, harvested after

5 days and FLDN cells were counted to determine the yield from each seeded population. Cell numbers were normalized to B6 samples for each subset and the geometric mean was plotted as shown in the graph. Error bars represent 1 standard deviation around the geometric mean. Plots are calculated from 3 independent experiments for DN1 and 4 independent experiments for DN2a and DN2b each.

Figure S3 – *Effect of deletion of PU.1 on gene expression in DN1, DN2a and DN2b cells.* B6 and PU.1 KO DN subsets cells were sorted as described in fig S2 and expression of the indicated genes relative to Actin was determined using qPCR. **A:** Gene expression changes in DN1 cells following PU.1 deletion. Each point on the plot represents log₂-transformed ratio of *Spi1*^{-/-} and B6 expression values for a particular gene in DN1 cells from 4-5 independent experiments. The red dots are expression ratios from experiments in which DN1 cells were infected with Cre alone (to delete *Spi1*), while the blue dots represent experiments in which Bcl-xL was co-expressed with Cre to improve the recovery of *Spi1*^{-/-} DN1 cells. Open triangles indicate the mean ratio while the black bars represent one standard deviation. **B-E:** Gene expression changes in B6 and *Spi1*^{-/-} DN2a and 2b subsets sorted as in Fig 3. Actin-normalized expression values averaged from 2 independent experiments are expressed as fold-change relative to B6 DN2a values. The resulting data were plotted for **B:** phase 1 and alternate lineage genes; **C:** T-lineage genes; **D:** components of the Notch-signaling pathway and **E:** alternate lineage genes. Error bars represent one standard deviation. Significant differences at $p < 0.05$ in the DN2a stages are indicated with an asterisk (*) while those in the DN2b cells are indicated by the pound (#) sign. Similar results were obtained in two additional experiments in which Bcl-xL was not used. However, without Bcl-xL the DN1 subsets did not show as complete deletion of PU.1, suggesting selection for some retention of PU.1 activity.

Figure S4 – *PU.1 restricts access to the NK lineage fate in developing thymocytes.* E14.5 FLPs were transduced with Cre and Bcl-xL carrying retroviruses and co-cultured with OP9-DL1 in media supplemented with 5 ng/ml each of IL-7 and Flt3L to initiate T-cell development. Cultures were not supplemented with any other cytokines (such as IL-15) that are known to be important for NK cell survival, proliferation and development. Cells were harvested on day 6 and the fraction of NK cells in each of the indicated populations was determined by flow cytometry. NK

cells were defined by the CD45⁺ CD25⁻ CD122⁺ NK1.1⁺ phenotype. **A:** FACS plots showing NK cell percentage in the total CD45⁺ CD25⁻ population. **B:** Graph showing the proportion of NK cells in the indicated samples from two independent experiments.

Figure S5 – *PU.1* indirectly represses *T*-lineage gene expression. **A:** Bcl2tg FLDN cells were processed as shown in the flowchart and sorted to purify DN2a and DN2b cells in **B**, or DN2 cells in **C**, and used for gene expression analysis by qPCR. **B-C:** Retroviral supernatants used to express various constructs are indicated on top of each lane. Sorted DN stage cells from 2-3 independent experiments were pooled to obtain each set of qPCR data. Average, Actin-normalized log₁₀ expression values from 2 such pooled datasets were used to generate the heatmaps shown in the figure. All values are row normalized to the control DN2a sample in **B** and to the control DN2 sample in **C**. The common scale on the right denotes 30-fold up- (dark red) and 30-fold down-regulation (dark blue) of gene expression. Bold face and arrows in panel B designate Notch-activated target genes, as characterized previously (Del Real and Rothenberg 2013).

Figure S6 – Gene expression changes in response to *PU.1-Eng* expression. Bcl2tg FLPs were cultured on OP9-DL1 and transduced to obtain EV, PU.1-Eng and PU.1 WT expressing DN1, DN2a and DN2b cells as shown in fig 5B. Sorted cells were used to analyze the expression of several genes as shown in the heatmap in fig 5C and D. Here the relative gene expression levels over two independent experiments are shown for DN1 (panels, A, C and E) and DN2a and DN2b cells (panels B, D and F) for **A-B:** T-cell genes **C-D:** alternate lineage genes and **E-F:** phase 1 and alternate lineage genes. **G:** RNAseq experiments and data analysis were performed as described in materials and methods section and the table shows the effect of PU.1-Eng and PU.1-ETS on the expression of selected cell cycle genes from two independent experiments. The last 4 columns show the padj values obtained from comparing the expression of cell cycle genes in indicated samples using the DESeq package. Note that none of the cell cycle genes were significantly differentially expressed (padj ≤ 0.05) in both RNAseq datasets.

Figure S7 – *Sorting DN2 cells for RNA-seq experiments.* OP9-DL1 co-cultures were seeded with e14.5 Bcl2tg FLPs and retrovirally-transduced on day 4 with the constructs indicated on the left. Cells were harvested the next day and Lin- 7-AAD- CD45+ GFP+ c-Kit+ CD25+ DN2 cells were sorted using the gating strategy shown in the figure. GFP is a marker of retroviral infection. Numbers in the boxes show percentage in each quadrant or gated population. PU.1 regulates CD45 expression in myeloid cells (Anderson et al 2001) and here PU.1-Eng, an obligate repressor construct, is seen to downregulate its surface expression on infected FLDN cells.

Figure S8 – *Changes in expression of phase 1 genes in response to PU.1 WT and PU.1-Eng expression.* OP9-DL1 co-cultures were seeded with e14.5 Bcl2tg FLPs and retrovirally-transduced with the indicated constructs after 4 days as shown in fig 5B. DN1, DN2a and DN2b subsets were sorted the next day and expression of the indicated phase 1 genes was determined by qPCR. Bars represent average, Actin-normalized expression of each gene from 2 independent experiments and error bars indicate 1 standard deviation.

SUPPLEMENTARY METHODS

Purification of fetal liver-derived precursors, cell culture and flow cytometry

Fetal liver cells from various crosses indicated in the methods section “Cell cultures” were used to initiate OP9-DL1 cultures as indicated for each experiment. Briefly, male and female mice were co-housed on the day of mating and the time until the next morning was counted as 0.5 days. After 14 days, pregnant females were sacrificed and fetal livers were isolated. Biotin conjugated antibodies against Ter-119, Gr-1, NK1.1, CD19 and F4/80 (eBioscience) were used to stain lineage positive cells, which were then depleted using streptavidin bound magnetic beads (Miltenyi) per manufacturer’s protocol. The enriched fetal liver precursors (FLPs) were either used directly to start OP9-DL1 co-cultures or frozen at -80°C for future use. Frozen FLPs were first thawed in OP9 medium (α -MEM, 20% FBS, 50 μ M β -ME, Pen-Step-Glutamine) supplemented with stem cell factor (SCF, 1 ng/ml) and IL-7 and Flt3L (5 ng/ml each) for a day before starting OP9-DL1 co-cultures, which were without SCF. For some experiments, depleted

fetal liver cells were used to sort c-Kit⁺ CD27⁺ T-cell precursors on the BD FACSAria sorter using 6-color flow cytometry. OP9-DL1 cultures were harvested at the indicated times, stained for DN progression and cytokine receptor expression, and analyzed with the MACSQuant flow cytometer (Miltenyi).

Retroviral infection and sorting

Fetal liver precursors were isolated from gestational day 14.5 embryos as above and co-cultured with OP9-DL1 cells in the presence of IL-7 and Flt3L (5 ng/ml each) for 3-4 days to obtain FLDN1, DN2a and DN2b cells. These cultures were harvested and pre-plated to obtain FLDN cells free of OP9-DL1 stromal cells. Briefly, culture medium was removed and the OP9-DL1 co-cultures were trypsinized for 5-7 min at 37°C. Fresh media with cytokines was added to these cultures and the cells were transferred to 30 cm tissue culture dishes for 35-40 mins at 37°C to allow the OP9-DL1 cells to re-adhere to the plate. The non-adherent FLDN cells were then harvested and used for retroviral infection.

Retroviral supernatants were generated by transfecting Phoenix-Eco cells with the appropriate retroviral constructs using Fugene 6 (Roche). Culture supernatants were collected 48 and 72 hours after transfection, filtered through a 0.45 µm filter and were either frozen at -80°C or used directly to infect FLDN cells. In short, 24-well non tissue culture treated plates were prepared by adding 500 µl of 30-40 µg/ml Retronectin solution/well and incubated at 4°C. The next day, Retronectin was replaced by 500 µl retroviral supernatant and the plates were spun at 2000 x g for 2 hours at 32°C. Retroviral supernatant was then replaced with FLDN cells from OP9-DL1 cultures, incubated at 37°C for 4 hours and transferred back on fresh OP9-DL1 plates. Cells were harvested at indicated timepoints, and the retrovirally transduced (GFP⁺ or NGFR⁺) DN1 (Lin⁻ c-Kit^{hi} CD44^{hi} CD25⁻), DN2a (Lin⁻ c-Kit^{hi} CD44^{hi} CD25⁺) and DN2b (Lin⁻ c-Kit⁺⁺ CD44⁺ CD25⁺) hematopoietic cells (all CD45⁺) were sorted using 6-color flow cytometry. Note that at later stages of infection with PU.1-Eng, the CD45 gate had to be expanded due to the downregulation of direct PU.1 target gene *Ptprc* (CD45)(Anderson et al. 2001).

Determination of Spi1^{fl/fl} deletion efficiency

Efficiency of deletion of *Spi1* exon 5 was assessed at the RNA level (cf. Fig. S2B), by quantitating the level of transcripts in the cells detected with primers spanning exons 4 and 5 (Table S6) and comparing the level in sorted Cre⁺ cells from the floxed genotype with Cre⁺ sorted wildtype cells at the same developmental stage. This assay measured deletion rather than altered transcriptional regulation, because cells with >80% deletion of *Spi1* exon 5 still expressed equal or more RNA than the controls from the 5' exons of the *Spi1* gene (data not shown).

RNA isolation and qPCR

Sorted FLDN populations were used to extract RNA with the RNeasy mini kit (Qiagen). Typically, total RNA was eluted in 30 μ l DEPC treated water and cDNA was synthesized using SuperScript III Reverse Transcriptase (Invitrogen). An appropriate cDNA dilution was then used for qPCR with SYBR[®] Greener[™] reagent (Life Technologies) and expression data was obtained on an ABI 7900 HT fast qPCR machine. Sequences of PCR primers used were from published work (David-Fung et al. 2009; Li et al. 2010; Yui et al. 2010; Del Real and Rothenberg 2013; Scripture-Adams et al. 2014). Each reaction was run in triplicate, and the value for each biological sample was taken as the average of these technical replicates. The resulting data were normalized to *Actin* expression and mean values for each sample type from 2-3 experiments were expressed as a heat map (Matlab) or plotted as log₁₀-scale graphs. For statistical analysis, actin-normalized gene expression values were log₁₀-transformed and two-tailed Student's *t* tests (paired, equal variance) were performed to determine the significance of differences between expression values. Significant differences in expression with a p-value < 0.05 are indicated with an asterisk. Error bars in all plots represent one standard deviation, or range when experiments had only two biological replicates.

Complete methods for transcriptome profiling and analysis of sequencing data

1. Read mapping and differential expression analysis

Total RNA was processed essentially as described (Zhang et al. 2012). cDNA libraries were sequenced with the Illumina HiSeq 2000 sequencer following manufacturer's protocols

(<http://www.illumina.com>) and produced between 11-14 million mapped 50bp single-end reads per experiment. Sequenced reads were mapped onto mouse genome build NCBI37/mm9 using TopHat (TopHat v2.0.6, <http://tophat.cbc.umd.edu>) with settings ‘--library-type fr-unstranded --no-novel-juncs’ and bowtie settings ‘-v 2 -k 40 -m 40’. HTseq (v0.6.0, <http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html>)(Anders et al. 2015) was used to process read alignments produced by TopHat. Read counts were obtained for each gene across all samples using htseq-count with the '-S no -a 10' options and with the Ensembl gene model file Mus_musculus.NCBIM37.66.gtf. The count data was analyzed for differential expression using the DESeq package (v1.16.0, <http://bioconductor.org/packages/release/bioc/html/DESeq.html>) (Anders and Huber 2010). Each replicate experiment was analyzed separately and a set of 168 genes with an adjusted p-value of ≤ 0.1 in both experiments were considered to be differentially expressed (DE); this set was augmented by 36 genes with padj-value < 0.1 in one set and $0.1 < \text{padj} < 0.2$ in the other. Note that the position of *Spi1* (*Sfp1*) in these gene lists as an “upregulated gene” is an artifact of the overexpressed sequences from the exogenous PU.1-Eng or PU.1-ETS.

2. Gene ontology and pathway enrichment analysis

We performed GO term and KEGG pathway over representation analysis on a set of genes that were called DE between the vector control and PU.1-Eng expressing samples at $\text{padj} \leq 0.1$ in at least one experiment. This set of 290 genes was further divided into 148 repressed and 142 PU.1-Eng upregulated genes. These lists were processed in R (R Core Team 2014)(<http://www.R-project.org/>) using the package Goseq (v1.16.2), which calculates enrichment after correcting for transcript length bias (Young et al. 2010). Only the GO terms and KEGG pathways with an adjusted (Benjamini-Hochberg method) p-value of ≤ 0.05 were considered to be enriched in the DE gene set.

3. Analysis of PU.1 binding sites for repressed and activated genes

A set of genes which were differentially expressed at $\text{padj} \leq 0.1$ in at least one set and at $\text{padj} < 0.2$ in both sets were selected if there was a ≥ 2 -fold expression difference in either direction between the PU.1-Eng and the vector control samples. The genes that were up- or down-regulated by PU1-Eng were separated, and compared with each other and with a control set of non-regulated but well-expressed genes in early T cells. The control set consisted of 841 genes from expression clusters 16 and 18 in (Zhang et al. 2012), two clusters of genes that are expressed stably at moderately high levels throughout T cell development, after the removal of 4 genes that were affected by PU.1-Eng. All DN1 and DN2a stage PU.1 peaks in the vicinity of each gene (Supp. Table 5, from (Zhang et al. 2012)) were first ordered by strength of occupancy signals (rpm) and all sites in order of strength were matched to the relevant genes on the three lists. The cumulative occupancy index for each gene was then calculated by summing the peak signals from DN1 and DN2a stages for peaks assigned to that gene. Although the number of sites per gene ranged from zero to 21 (*Haa0*), only values for the top 4 sites from the two stages were summed to minimize background error and reduce outlier effects. The distribution of the sum of binding strength for each group of genes was tested for similarity using a pairwise Z-test for means using the R package BSDA (Kitchens 2002). This score does not take into account gene length, but there was no significant difference in gene length distribution among the three groups of genes (Z scores -1.47 to +1.46, p values 0.14-0.43, Table S2).

4. Indexing developmental state by principal component analysis

We used a set of 173 developmentally regulated transcription factors to assess the developmental state of our RNA-seq samples (Scripture-Adams et al. 2014). The expression values for these genes in 11 wild-type samples spanning DN1/ETP to DP stages and 6 samples from the present study were first log-transformed and normalized so that the mean of the expression values for each gene across all samples was zero. Then, the signals from the wildtype samples were used to calculate the principal components using the `prcomp` function from the R stats package (R Core Team 2014). These values provide a framework within which a perturbed sample can be related to normal development. We then used the principal component loadings for the 173 genes from the normal samples to transform the values from the six experimental samples, so that they could

be plotted on the scales defined by the normal reference ones. The first two principal components for the experimental and reference samples were plotted using the R package ggplot2 (Wickham 2009).

5. *Gene Set Enrichment Analysis (GSEA)*

We used the “GSEAPreranked” tool in order to fully characterize the effects of PU.1-ENG on developmentally regulated and lineage-restricted genes sets. All genes that were expressed at ≥ 1 FPKM in at least 1 sample were included in the analysis. Genes were ranked in descending order of mean log₂ ratios between all pairwise comparisons from both RNA-seq experiments. These ranked lists were used to test a number of gene sets obtained from two previous reports. First, we used 25 gene sets from our previous study (Zhang et al 2012) that were obtained by K-means clustering of all genes expressed during T-cell development. The most important ones included clusters of genes whose expression was a) highest in pre-commitment stages (clusters 3, 7, 9 and 23) b) upregulated as PU.1 expression is silenced (clusters 1, 6, 15 and 25). Some clusters, which showed similar expression patterns, were combined into larger sets that were tested in addition to running each cluster by itself. Additionally, we used data from a report by Kamath et al looking at changes in gene expression in myeloid precursor cells expressing PU.1 at 2% (Blac allele) of WT levels. This dataset was used to identify genes that were repressed by PU.1 (2-, 5-, or 10-fold upregulated in Blac samples compared to WT) and positively regulated (5-, 10-, or 15-fold downregulated in Blac samples compared to WT). This dataset was chosen for two reasons - 1) a decrease in PU.1 dosage was shown to upregulate T-lineage-specific genes in myeloid cells similar to our results showing that PU.1 represses the T-lineage developmental program in early T-cells and 2) several myeloid genes are expressed in early T-cells and we wanted to identify PU.1 functions that are common to both these lineages. GSEAPreranked was run using phenotype permutations and by using the conservative “classic” setting to calculate enrichment score as suggested on the GSEA website (<http://www.broadinstitute.org/gsea/index.jsp>). All gene sets with FDR ≤ 0.01 were considered to be significantly enriched.

SUPPLEMENTARY REFERENCES

- Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* **11**: R106.
- Anders S, Pyl PT, Huber W. 2015. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166-169.
- Anderson KL, Nelson SL, Perkin HB, Smith KA, Klemsz MJ, Torbett BE. 2001. PU.1 is a lineage-specific regulator of tyrosine phosphatase CD45. *JBiolChem* **276**: 7637-7642.
- David-Fung ES, Butler R, Buzi G, Yui MA, Diamond RA, Anderson MK, Rowen L, Rothenberg EV. 2009. Transcription factor expression dynamics of early T-lymphocyte specification and commitment. *Dev Biol* **325**: 444-467.
- Del Real MM, Rothenberg EV. 2013. Architecture of a lymphomyeloid developmental switch controlled by PU.1, Notch and Gata3. *Development* **140**: 1207-1219.
- Kitchens LJ. 2002. *Basic Statistics and Data Analysis*. Thomson/Brooks/Cole.
- Li L, Leid M, Rothenberg EV. 2010. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. *Science* **329**: 89-93.
- R Core Team. 2014. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Scripture-Adams DD, Damle SS, Li L, Elihu KJ, Qin S, Arias AM, Butler RR, 3rd, Champhekar A, Zhang JA, Rothenberg EV. 2014. GATA-3 dose-dependent checkpoints in early T cell commitment. *J Immunol* **193**: 3470-3491.
- Wickham H. 2009. *ggplot2: Elegant graphics for data analysis (Use R!)*. Springer, New York.
- Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* **11**: R14.
- Yui MA, Feng N, Rothenberg EV. 2010. Fine-scale staging of T cell lineage commitment in adult mouse thymus. *J Immunol* **185**: 284-293.
- Zhang JA, Mortazavi A, Williams BA, Wold BJ, Rothenberg EV. 2012. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell* **149**: 467-482.