Supplemental Figure 1.

(**A)** Rounds of replating (R) of MLL-AF9 transformed LSK colonies from initiation experiments (S.E.M of triplicate experiments, p=0.0033). *, p<0.05; **, p<0.01; ***, p<0.001; (**B**) Spleen weights and (**C**) Liver weights from moribund mice that were from the primary initiation experiment, data from *Msi2*$^{f/f}$ n=10 and *Msi2*$^{\Delta/\Delta}$

n=9 mice. Weights from spleen (**D**) and liver (**E**) from moribund mice that were from the primary maintenance experiment, *Msi2*$^{f/f}$ n=10 and *Msi2*$^{\Delta/\Delta}$ n=9 mice. *p<0.05, **p<0.01, ***p<0.001.

Supplemental Figure 2.

(A) Frequency of F480 and CD115 of leukemic cells from mice of indicated genotypes *Msi2*$^{f/f}$ n=8 and *Msi2*$^{\Delta/\Delta}$ n=10 from maintenance LSK experiments based on staining in Figure 3C. (**B**) Representative images of Wright-Giemsa staining of cytospins from bone marrow isolated leukemia cells taken from the initiation experiment. (**C**) Cell cycle status based on Hoescht staining in c-kit gated cells from initiation *Msi2*$^{f/f}$ n=9 and *Msi2*$^{\Delta/\Delta}$ n=10 and maintenance experiments *Msi2*$^{f/f}$ n=8 *Msi2*$^{\Delta/\Delta}$ n=10. (**D**) Apoptosis was calculated from the sub-G1 phase from samples in (C). (**E**) Colony assays from initiation LSK experiment *Msi2*$^{f/f}$ n=6 and *Msi2*$^{\Delta/\Delta}$ n=6 leukemic mice. (**F**) Cell counts from colonies analyzed in (**E**). (**G**) Representative flow plot of MSI2 intracellular staining showing complete and incomplete deletion of Msi2. Mean frequency of MSI2 high population in c-Kit^{high} cells of the indicated genotypes from mice from (**H**) LSK initiation, *Msi2*$^{f/f}$ n=19 and *Msi2*$^{\Delta/\Delta}$ n=18 (**I**) GMP initiation *Msi2*$^{f/f}$ n=6 and *Msi2*$^{\Delta/\Delta}$ n=10 and (**J**) maintenance experiment *Msi2*$^{f/f}$ n=6 and *Msi2*$^{\Delta/\Delta}$ n=8.

Supplemental Figure 3.

Survival analysis of the secondary transplants of leukemia cells from different donors from primary transplanted mice that were combined in Figure 3E (**A-D**) LSK initiation experiment and Figure 3F (**E-H**) maintenance experiment.

p<0.01, ****, p<0.0001. (A**) *Msi2*$^{f/f}$ n=10 and *Msi2*$^{\Delta/\Delta}$ n=8. (**B**) *Msi2*$^{f/f}$ n=10 and *Msi2*$^{\Delta/\Delta}$ n=9. (**C**) A combination of three different donors *Msi2*$^{f/f}$ n=11 and *Msi2*$^{\Delta/\Delta}$ n=9. (**D**) *Msi2*$^{f/f}$ n=5 and *Msi2*$^{\Delta/\Delta}$ n=5. (**E**) *Msi2*$^{f/f}$ n=11 and *Msi2*$^{\Delta/\Delta}$ n=9. (**F)** *Msi2*$^{f/f}$ n=10 and *Msi2*$^{\Delta/\Delta}$ n=8. (**G)** *Msi2*$^{f/f}$ n=5 and *Msi2*$^{\Delta/\Delta}$ n=10. (**H**) Frequency of MSI2 positive cells in the c-Kit^{high} cells in Supplemental Figure 3F obtained by intracellular analysis of MSI2^{High} cells in the c-Kit^{High} compartment *Msi2*$^{f/f}$ n=3 and *Msi2*$^{\Delta/\Delta}$ n=5.

Supplemental Figure 4.

Msi2 controls growth factor signaling in LSCs. (**A**) Representative flow cytometric gating on c-Kit^{High} in leukemia cells isolated from the bone marrow. (**B**) Representative histograms and levels of pERK gated in (**A**) after serum starvation and stimulation with cytokines (IL3, IL6 and SCF) for indicated time points (minutes) and then stained for intracellular phosphorylated ERK (pERK). (**C**) Fold of pERK Median fluorescence intensity (MFI) after cytokine stimulation. Mean and S.E.M from leukemia cells of genotype *Msi2*$^{f/f}$ n=5 and *Msi2*$^{\Delta/\Delta}$ n=6 mice from two independent experiments are shown. (**D**) Fold stimulation of pS6 and (**E**) pSTAT5 after cytokine stimulation. Mean and S.E.M from leukemia cells from *Msi2*$^{f/f}$ n=5 and *Msi2*$^{\Delta/\Delta}$ n=5 mice from two independent experiments are shown *, p<0.05; ***, p<0.001.

Supplemental Figure 5.

(**A**) Unsupervised clustering of the *Msi2*$^{\Delta/\Delta}$ MLL-AF9 LSC signature cluster on ECOG1900 AML patients. Heat map and unsupervised clustering identifies nine

groups when (177) genes differentially expressed in the *Msi2*$^{f/f}$ and *Msi2*$^{\Delta/\Delta}$ LSC RNA-sequencing data is overlapped with RNA profiling from AML patients (n=363 patients). Columns represent individual patients and rows are genes Log TPM expression normalized by row means. Color bar above gene rows are displayed as follows: *MSI2* expression above one standard deviation of the median in patients were labeled as "High," below as "Low" and between as "Mid," Cluster numbers separated by colors and mutations as shown. Cluster-4 patients are further subdivided with shaded grey for normal bone marrow controls. Patients in cluster-5 and Cluster-7 were isolated based on their differential MSI2 levels see Figure 4. (**B**) Cluster 5 **(C)** Cluster 7 from Figure 4E, Circos plots indicating relative frequency and pairwise co-occurrence of mutations in patients. The length of the arc corresponds to the frequency of mutations in the first gene, and the width of the ribbon corresponds to the percentage of patients who also had a mutation in the second gene. The number next to the gene represents the number of patients.

Supplemental Figure 6.

IGV generated plot showing reads on the *Ikzf2* locus from RNA sequencing data obtained from LSCs of *Msi2*$^{f/f}$ n=2 and *Msi2*$^{\Delta/\Delta}$ n=2. Sashimi plot shows the splicing of *Ikzf2.*

Supplemental Figure 7

(A) mRNA stability of *Ikzf2*, *Hoxa9* and c-*Myc* in control and MSI2$^{O/E}$ cells. Data of two independent experiments are shown. **(B)** qPCR of *Ikzf2*, *Hoxa9* and c-*Myc*

in c-Kit^{high} cells of indicated genotypes *Msi2*$^{f/f}$ n=3 and *Msi2*$^{\Delta/\Delta}$ n=4**. (C)** Immunofluorescence staining of c-MYC in sorted *Msi2*$^{f/f}$ and *Msi2*$^{\Delta/\Delta}$ LSCs. Cells were stained with c-MYC–specific Ab (left) and DAPI (middle); a merged image is shown (right). Bars, 20 μm. Fluorescence intensity was quantified and normalized to the control (three independent experiments; 184 and 418 cells were quantified per group in total). Means and SEM are shown. **(D)** Representative flow cytometric plot for intracellular staining of MSI2 and IKZF2 in c-Kit^{High} cells (left panel). Frequency of IKZF2^{High} cells in MLL-AF9 leukemic cells averaged from *Msi2*$^{f/f}$ n=4 and *Msi2*$^{\Delta/\Delta}$ n=4 (right panel). Means and S.E.M for B-D, *, p<0.05; **, p<0.01; ***, p<0.001. **(E)** Immunoblot of lysates from leukemic bone marrow cells from moribund mice of indicated genotypes *Msi2*$^{f/f}$ n=2 and *Msi2*$^{\Delta/\Delta}$ n=3. Mean frequency of **(F)** MSI2, **(G)** IKZF2 and **(H)** Myc positive cells after acute deletion of Msi2 at 68hrs after 600nM 4-OHT treatment in *Msi2*$^{f/f\ Cre\text{-}ER^{-/+}}$ leukemic cells. Data shown are from intracellular staining of *Msi2*$^{f/f\ Cre-}$ n=3 and *Msi2*$^{f/f\ Cre+}$ n=4. Means and S.E.M *, p<0.05; **, p<0.01; ***, p<0.001.

Supplemental Figure 8

(A) qPCR of *Ikzf2* in leukemic cells from *Msi2*$^{f/f}$ n=3 and *Msi2*$^{\Delta/\Delta}$ n=5 from the maintenance secondary transplant #2 where all *Msi2*$^{\Delta/\Delta}$ mice formed leukemia. (**B**) Representative flow plot showing IKZF2 high cells by intracellular staining. Mean frequency of IKZF2 high cells in *Msi2*$^{f/f}$ n=4 and *Msi2*$^{\Delta/\Delta}$ n=5 from secondary maintenance transplants (Supplemental Figure 3D and 3F) where all *Msi2*$^{\Delta/\Delta}$ formed leukemia. (**C**) Mean frequency of IKZF2 high cells in *Msi2*$^{f/f}$ n=7 and *Msi2*$^{\Delta/\Delta}$ n=3 from secondary maintenance transplants (Supplemental Figure

3E and 3G) where majority of *Msi2*$^{\Delta/\Delta}$ did not form leukemia.

Supplemental Table 1. HITS-CLIP fold change rank list. K562 UV-cross linked and RNA-IP rank list is a fold change between FLAG-MSI2 overexpression and empty vector transduced cells. Negative Fold change represents background binding of transcripts that have a FKPM>1 1 read on the RNA-seq.

Supplemental Table 2. Summary of deletion status of primary Msi2 deficient leukemia. Samples that had more than 70% deletion in the gated GFP+ leukemic cells were considered deleted based on intracellular flow cytometry and FKPM>1 read from the RNA-seq.

Supplemental Table 3. Summary of differentially expressed genes in *Msi2*$^{f/f}$ and *Msi2*$^{\Delta/\Delta}$ LSCs and Msi2 LSC signature. List of downregulated genes and upregulated genes ranked on log2-Fold change= >0.75, p<0.05, FDR<.05. List of differentially expressed genes that have human homologs >6 mean log intensity).

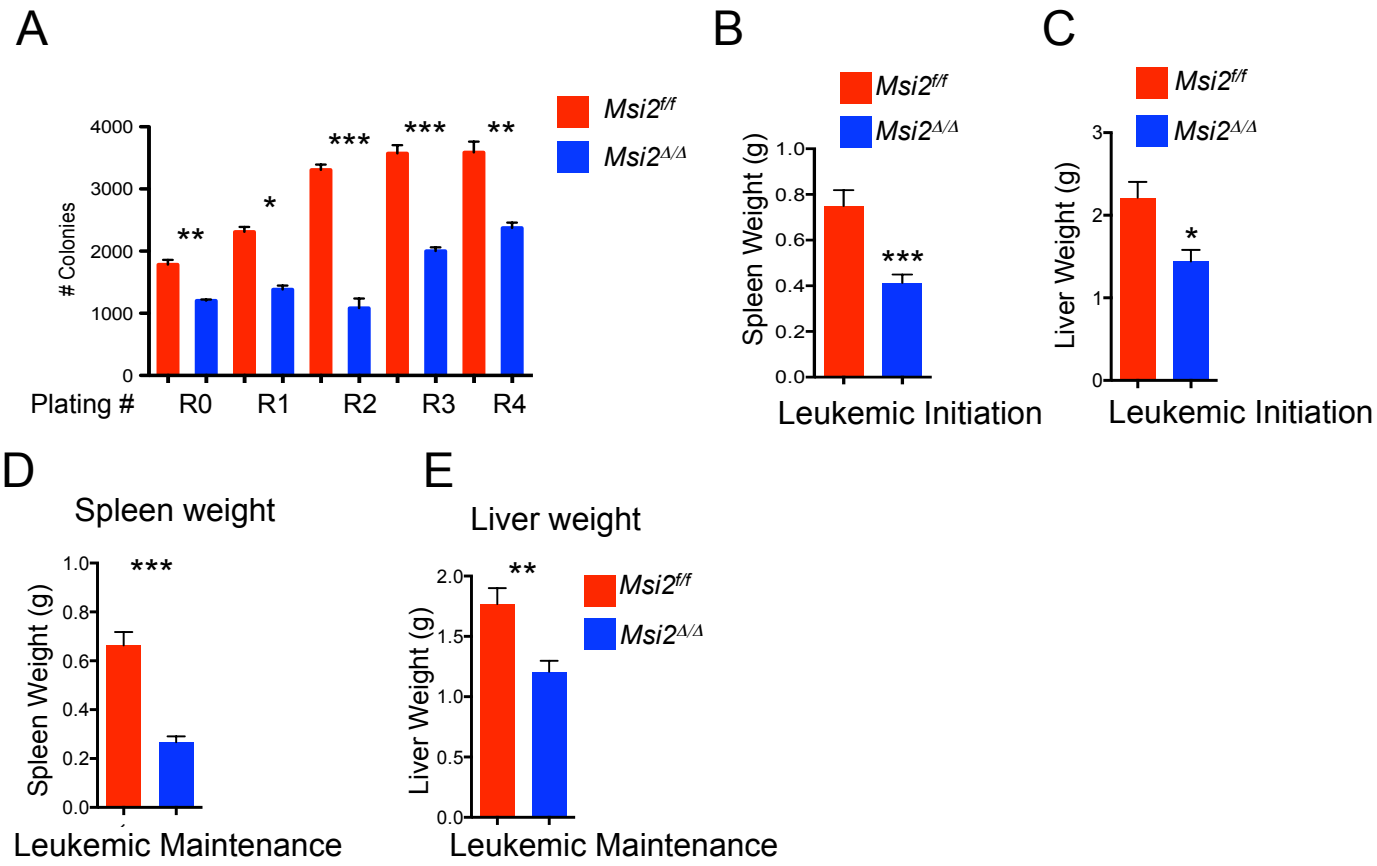
Supplemental Table 4. Gene sets enrichment analysis in *Msi2*$^{\Delta/\Delta}$ LSC that are downregulated.

Supplemental Table 5. Gene sets enrichment analysis in *Msi2*$^{\Delta/\Delta}$ LSC that are upregulated.

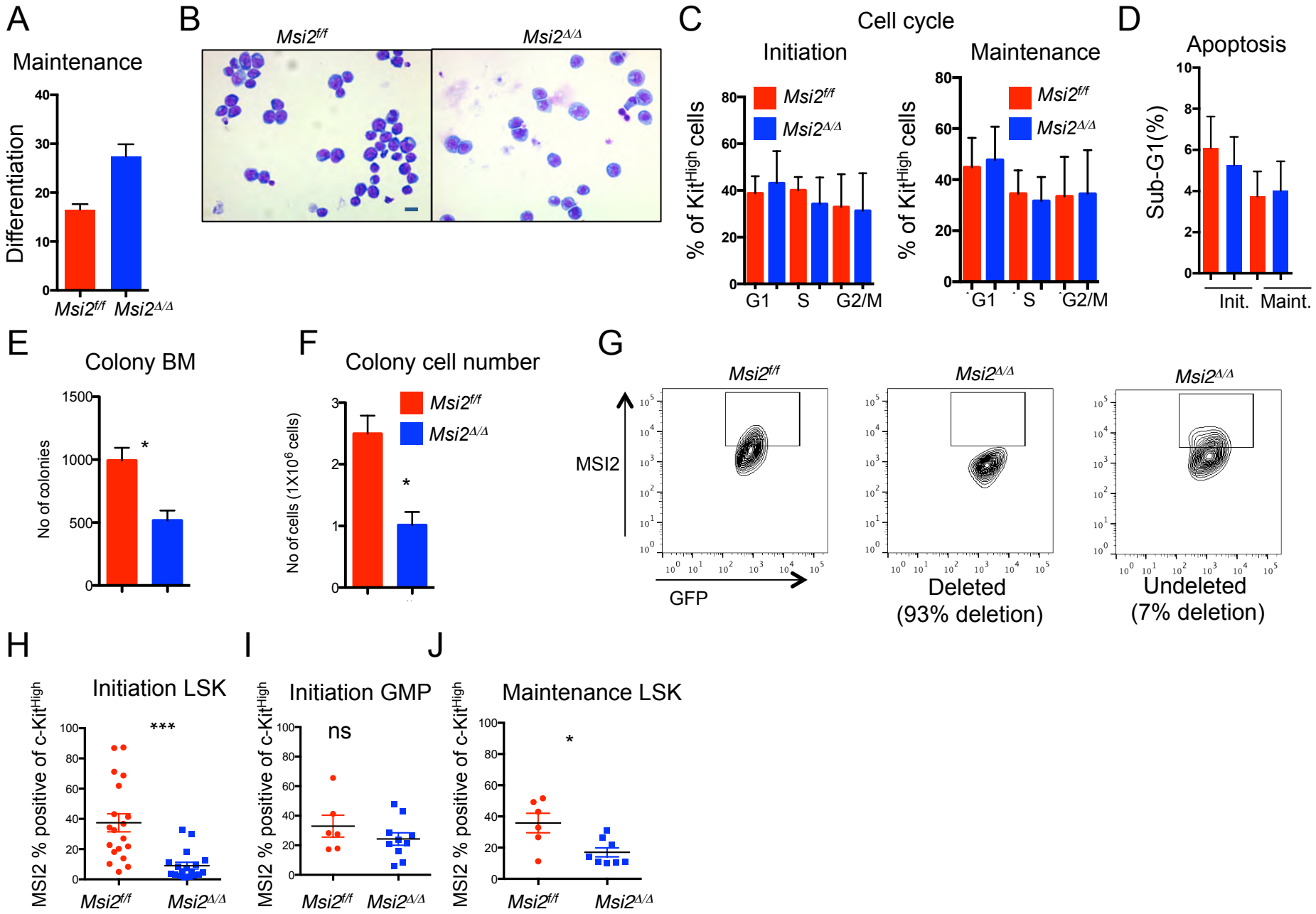
Supplemental Table 6. Msi2 deficient LSK and LSC gene sets enrichment analysis that overlaps with the HITS-CLIP gene sets (Figure 5B).

Supplemental Table 7. List of Top 1097 HITS-CLIP MSI2 bound genes and MLL-AF9 previously annotated gene sets.

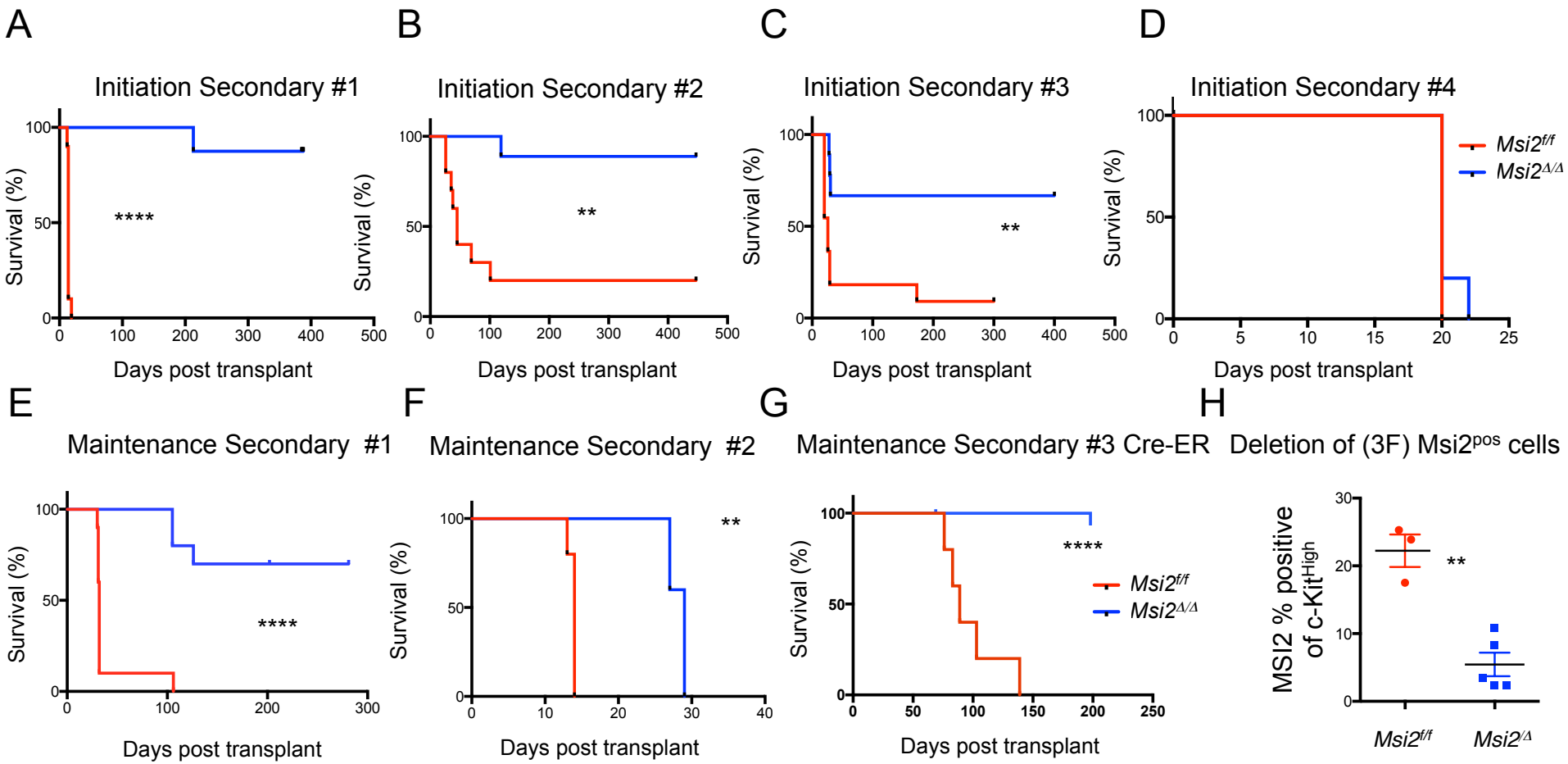
Supplemental Table 8. List of the overlap of HITS-CLIP MSI2 bound genes with MLL-AF9 transcriptional targets for RNA-IP prioritization.



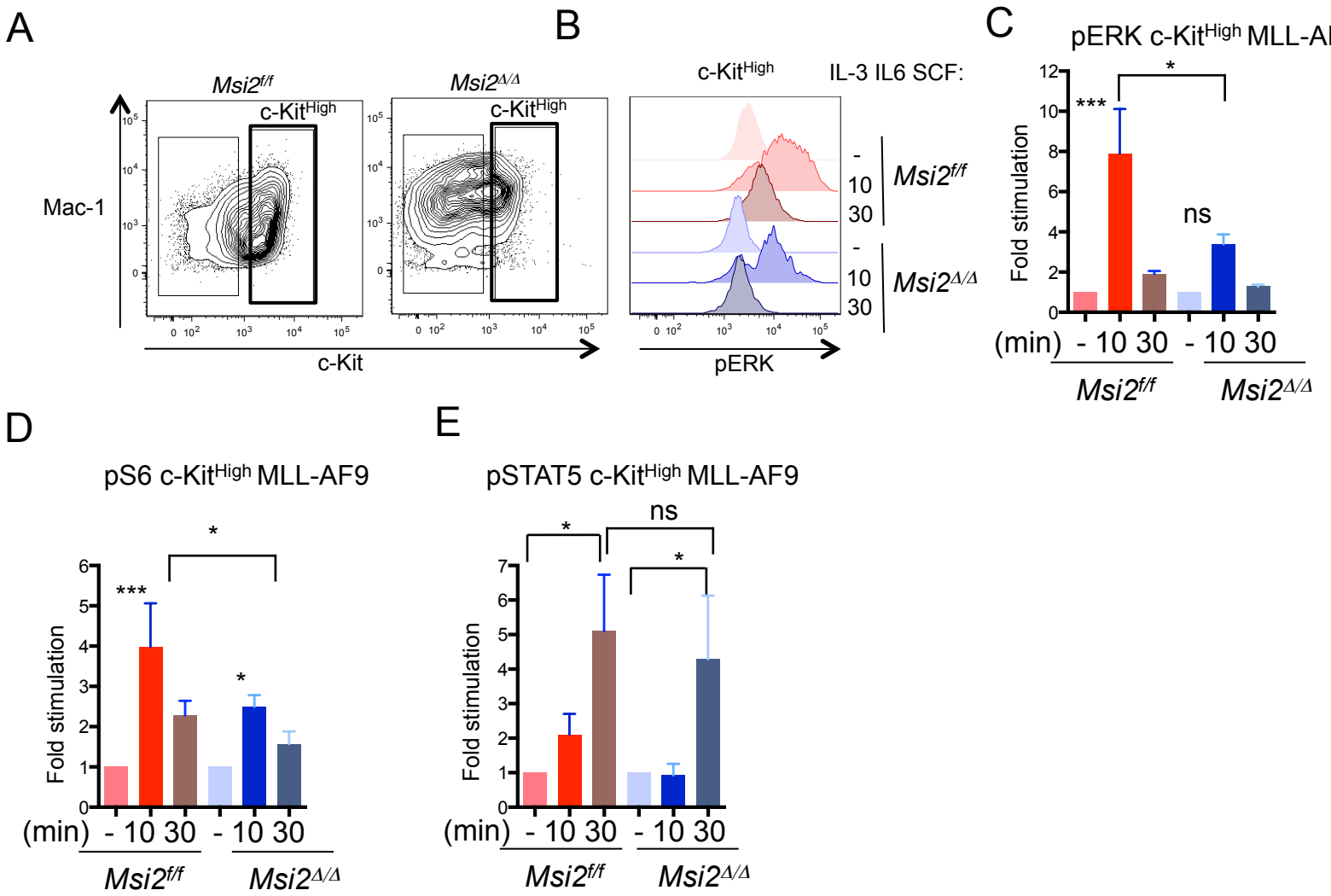
Supplemental Figure 1.



Supplemental Figure 2.

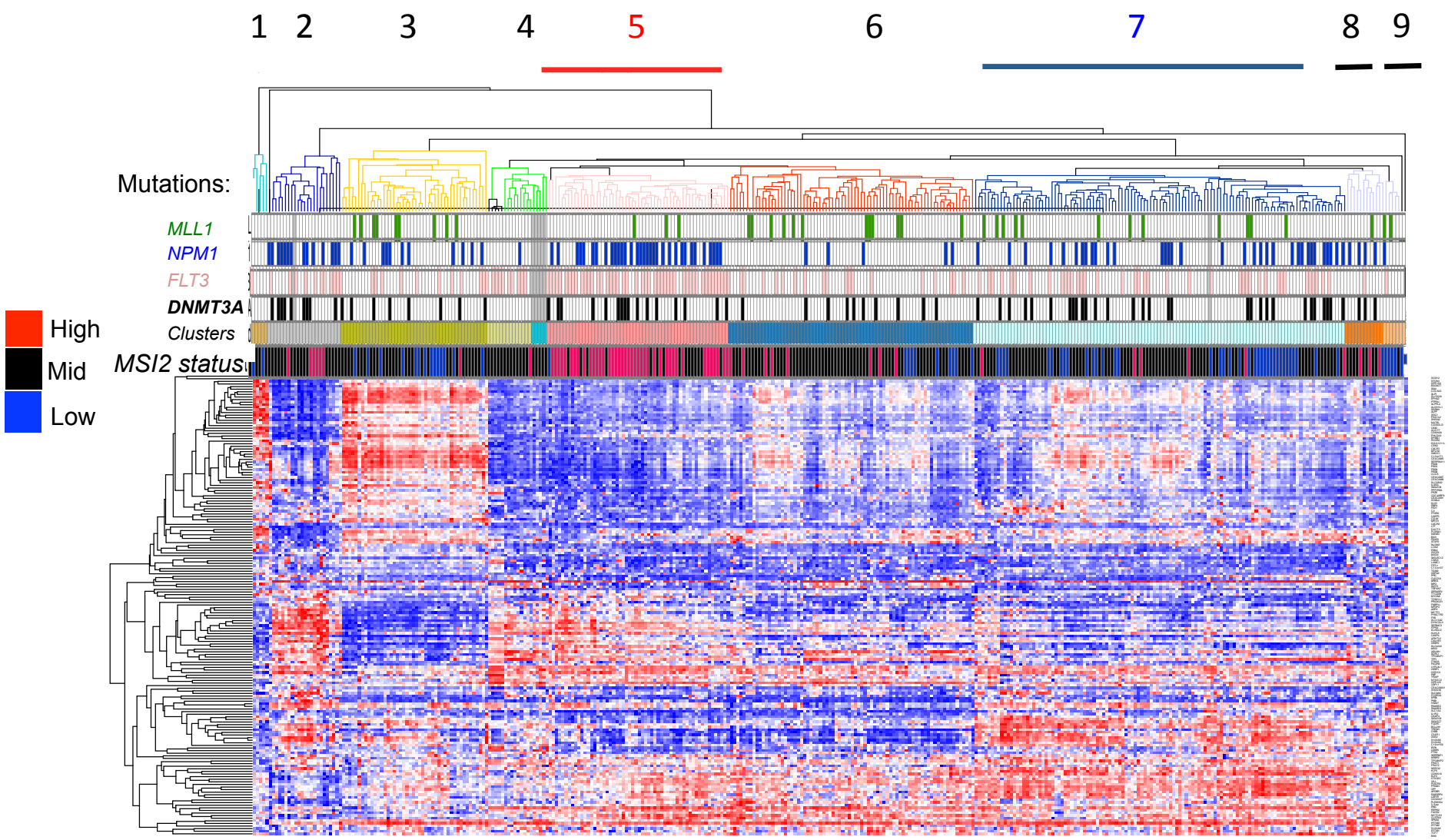


Supplemental Figure 3.



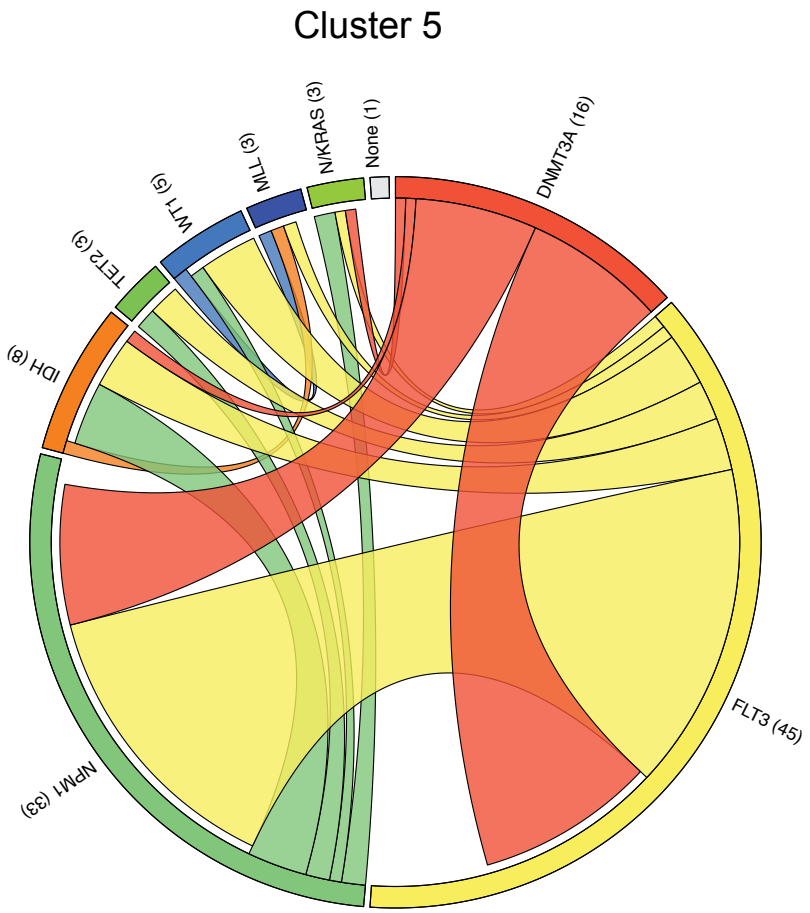
Supplemental Figure 4.

A

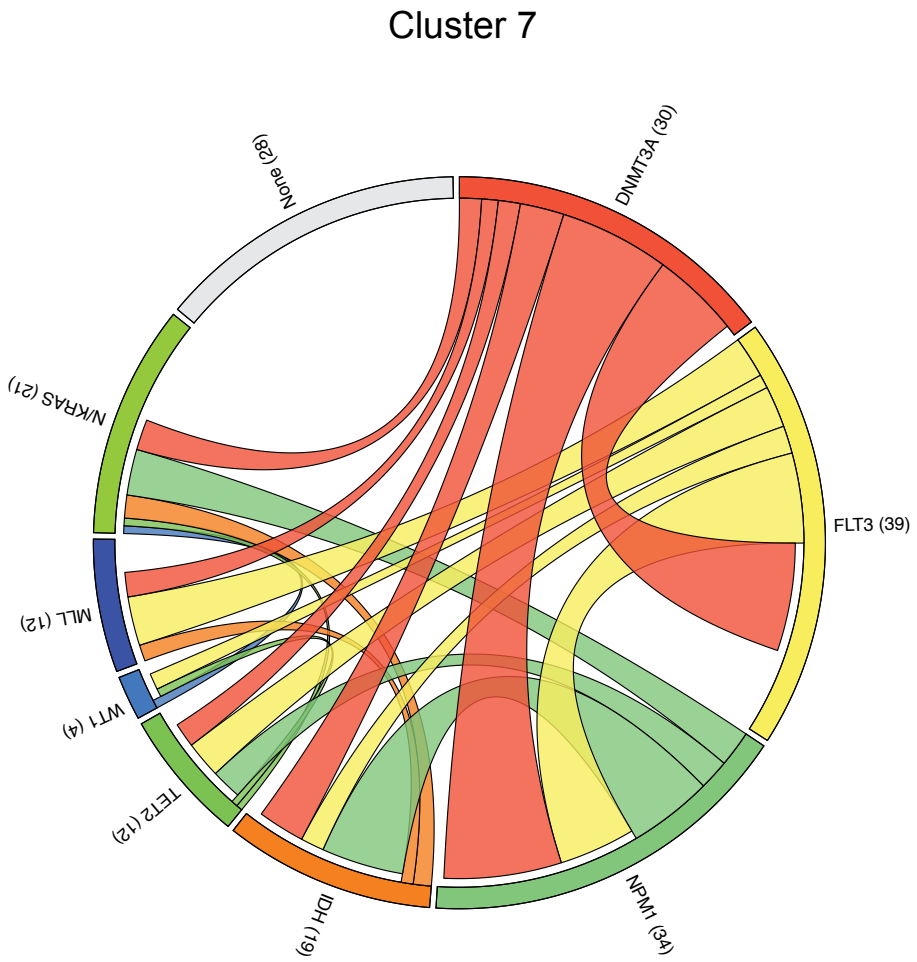


Supplemental Figure 5.

B

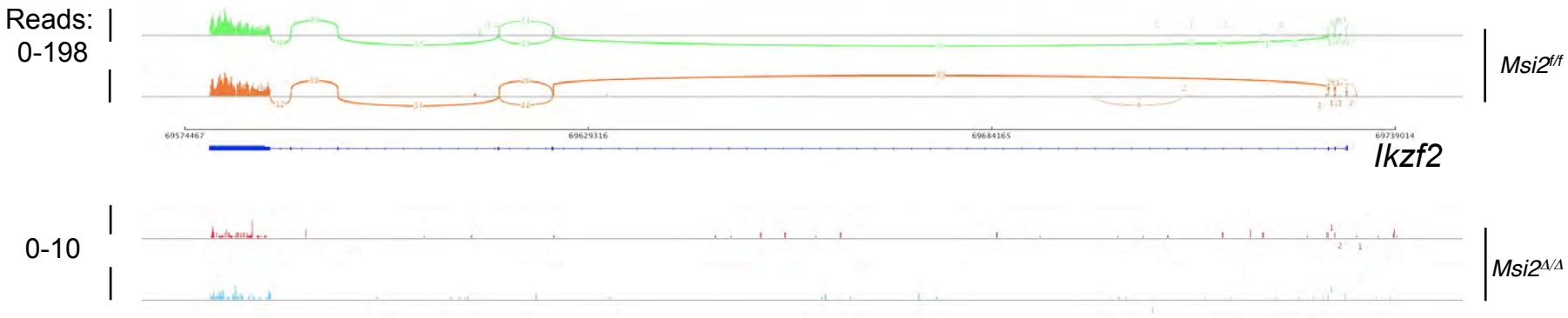


C

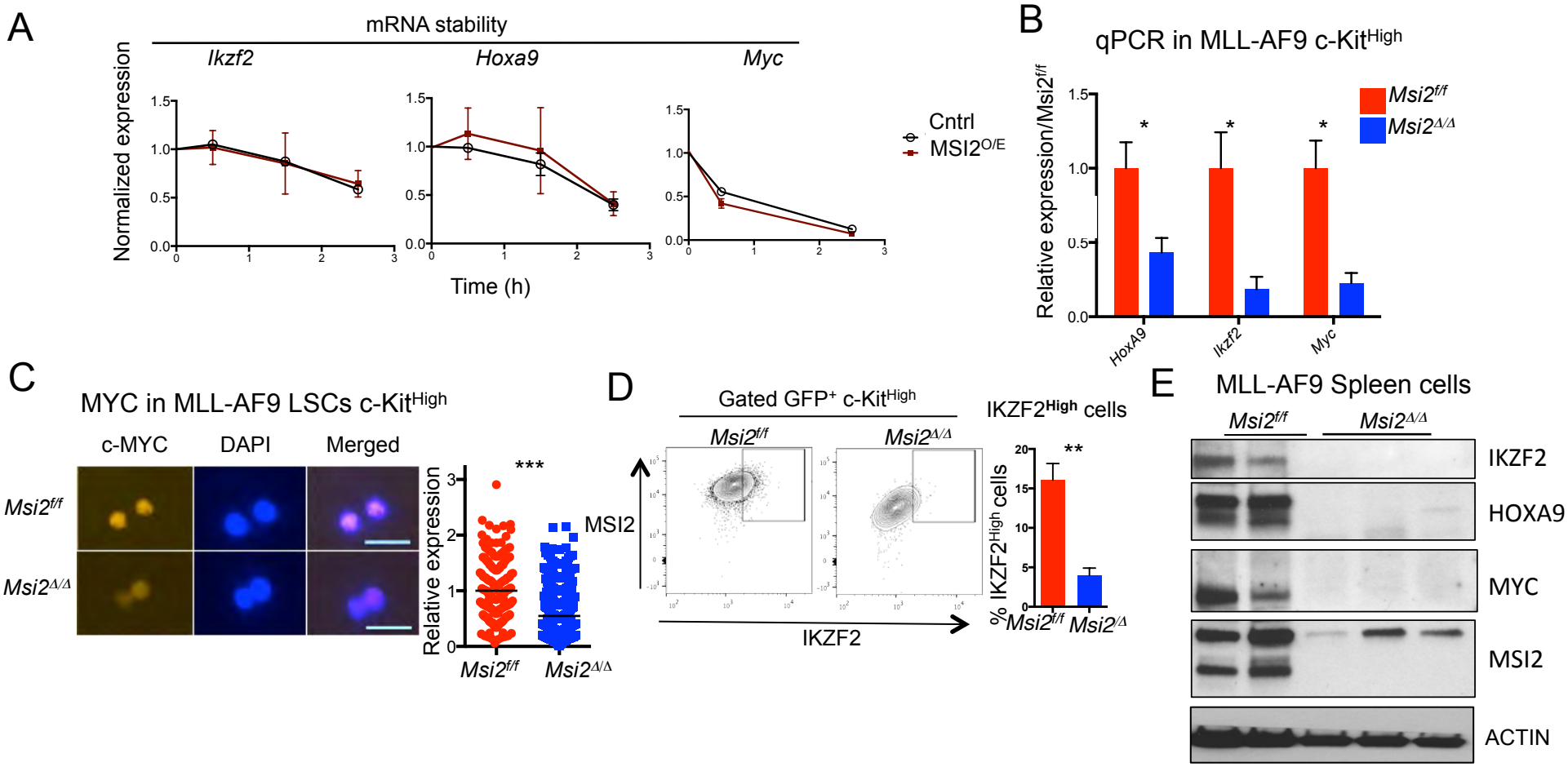


Supplemental Figure 5.

RNA-seq MLL-AF9 LSCs *(Ikzf2)*

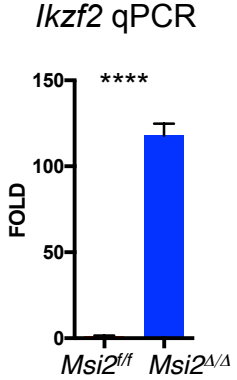


Supplemental Figure 6.



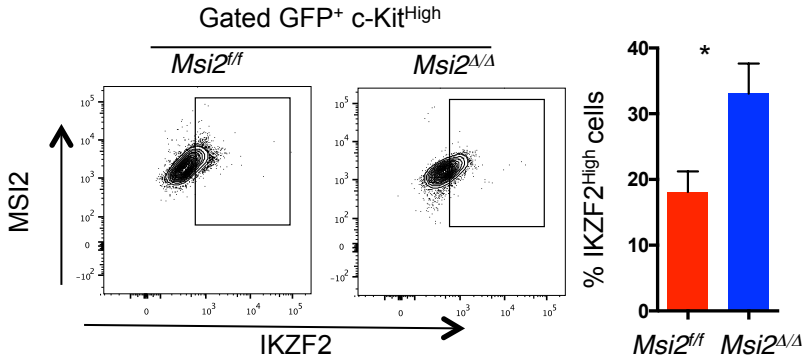
Supplemental Figure 7.

A



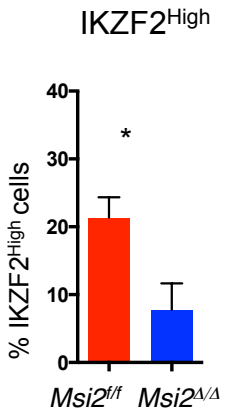
Supplemental Figure 3D and 3F

B



Supplemental Figure 3D and 3F

C



Supplemental Figure 3E and 3G

Supplemental Fig 8.