Supplemental Materials for

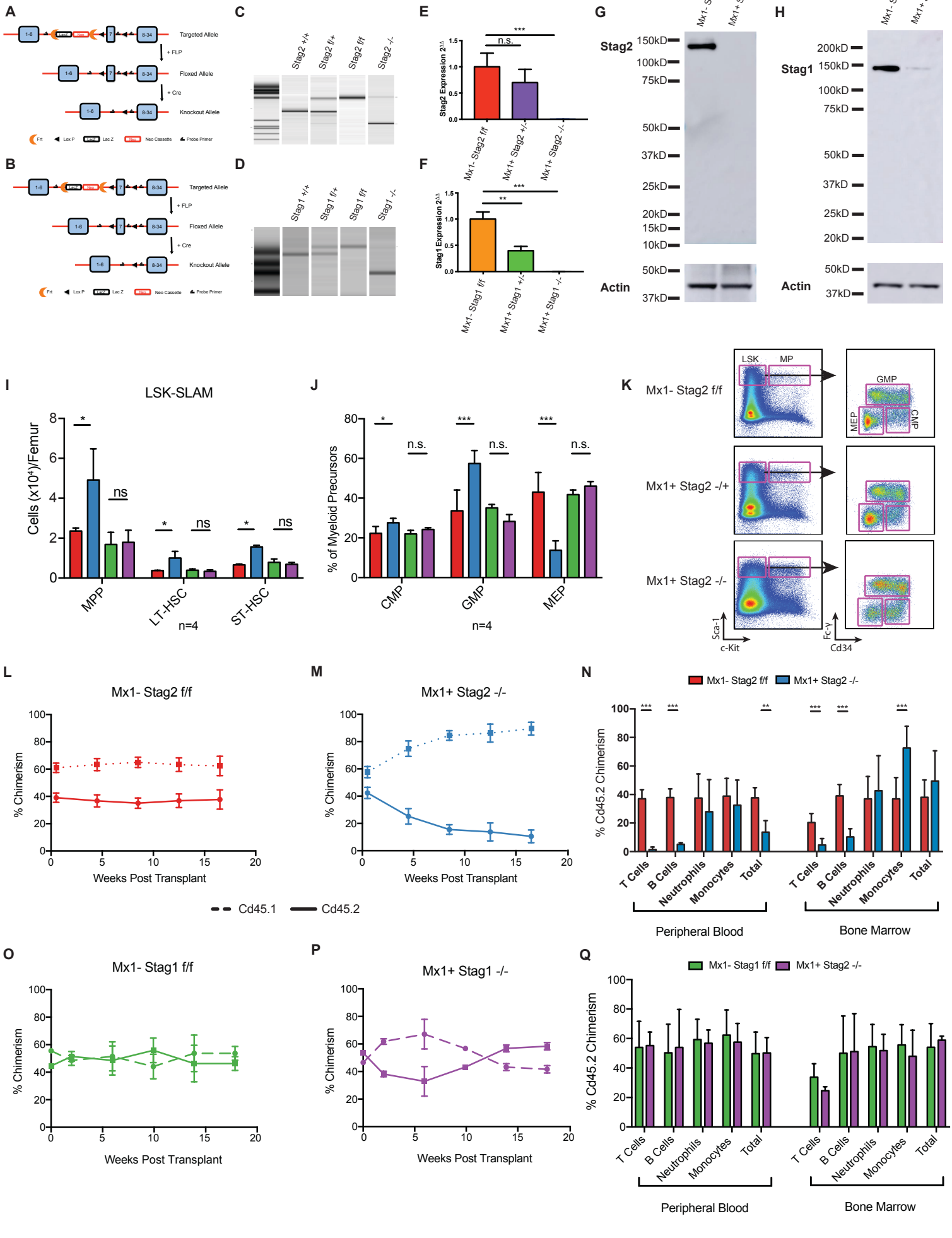
Stag2 regulates hematopoietic differentiation and self-renewal through alterations in gene expression and topological control

Authors: Aaron D. Viny[1,2,3], Robert L. Bowman[1,3], Yu Liu[4], Vincent-Philippe Lavallée[5], Shira E. Eisman[1,3], Wenbin Xiao[1,6], Benjamin H. Durham[1,6], Anastasia Navitski[1,3], Jane Park[3], Stephanie Braunstein[1,3], Besmira Alija[1,3], Abdul Karzai[1,3], Isabelle S. Csete[1,3], Elham Azizi[5], Matthew Witkin[3], Timour Baslan[7], Christopher J. Ott[8], Dana Pe'er[5], Job Dekker[4,9], Richard Koche[3], Ross L. Levine[1,2,3*]

Correspondence to: leviner@mskcc.org

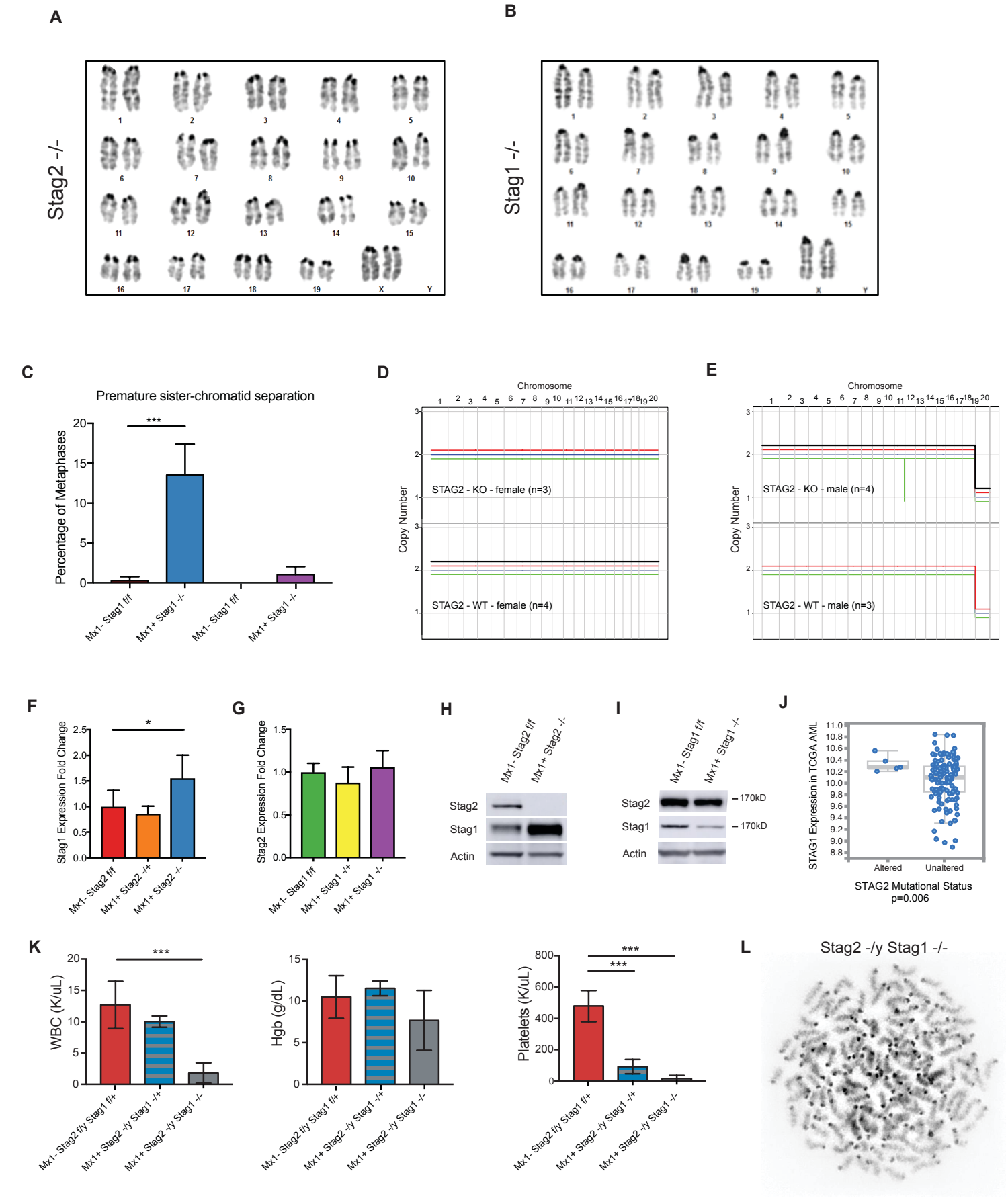
This PDF file includes:

Supplemental Fig. 1



Supplemental Fig 1. (Related to Figure 1). Schematic representing the conditional knockout allele for **A)** Stag2 and **B)** Stag1 with each containing LoxP sites flanking exon 7. Digital PCR of blood genotyped for **C)** Stag2 or **D)** Stag1 from WT, floxed, and excised genotypes. **E)** RT-PCR measuring gene expression of Stag2 or **F)** Stag1 in floxed, heterozygous, and homozygous deletion. **G)** Full length western blot for Stag2 and **H)** Stag1 in floxed and excised bone marrow. **I)** Stag2 KO, but not Stag1 KO mice have increased hematopoietic stem cells as enumerated for each of the SLAM populations MPP, LT-HSC, and ST-HSC. **J)** Myeloid Progenitors ($Lin^-Kit^+Sca1^-$) show expanded granulocyte-macrophage precursors (GMP) at the expense of megakaryocyte-erythroid precursors (MEP). **K)** Representative flow cytometry scatter-plots of Stag2 floxed, heterozygous, and KO bone marrow (Parent gate is Lin^- live singlets) showing increased LSK. Myeloid progenitors are gated by Cd34 and Fc-γ revealing increased granulocyte-macrophage precursors (GMP) and reduced megakaryocyte-erythroid progenitors (MEP). Competitive bone marrow transplantation of **L)** Stag2 WT or **M)** KO bone marrow mixed 1:1 with Cd45.1 normal marrow. Mice were injected with PIPC following engraftment at week 2. Flow cytometry of peripheral blood measured Cd45.2 chimerism every 4 weeks and **N)** at 16 weeks in the bone marrow. Competitive bone marrow transplantation of **O)** Stag1 WT or **P)** KO bone marrow mixed 1:1 with Cd45.1 normal marrow shows no difference in chimerism in peripheral blood **Q)** or in the bone marrow at 16 weeks.

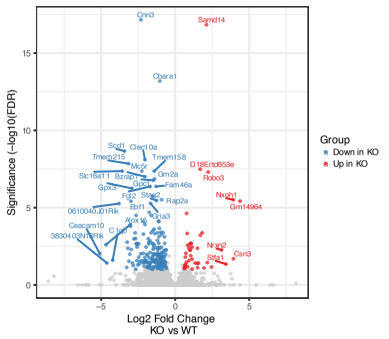
Supplemental Fig. 2



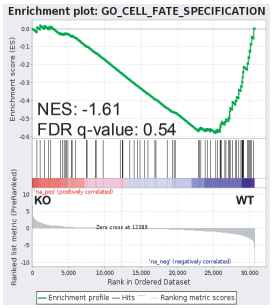
Supplemental Fig 2. (Related to Figure 1) A) Cytogenetic analysis of Stag2 and **B)** Stag1 KO mice. Representative metaphase spreads depicted for each genotype. All metaphase karyotypes were 40,XX [20] for Stag2 WT (n=4), Stag2 KO (n=3), Stag1 KO (n=3), and 40,XY [20] for Stag2 WT (n=3) and Stag2 KO (n=4). One of three Stag1 WT samples had a single tetraploid metaphase (40,XX[19], 80,XXXX [1]), and normal 40,XX [20] in the remaining two samples. **C)** Morphologic analysis of 100 metaphase figures for the presence of premature sister chromatid separation shows that more Stag2 KO cells had premature sister chromatid separation (mean=13.5%, $p<0.001$). **D)** Low depth whole genome sequencing of Stag2 WT and KO bone marrow for copy number shows no alterations genome wide in Stag2 WT females (n=4), Stag2 KO females (n=3), or **E)** Stag2 WT males (n=3). One of four Stag2 KO males shows a small 2Mb deletion at chr11qE2, and 3 have no alterations genome wide. **F-G)** RT-PCR measuring gene expression of (F) Stag1 or (G) Stag2 in floxed, heterozygous, and homozygous deletion of the opposing Stag gene. Stag1 expression increases in Stag2 KO bone marrow (student's t test, $p<0.03$), but Stag2 expression does not change with Stag1 deletion ($p<0.65$). Western blot for Stag2 and Stag1 in **H)** Stag2 and **I)** Stag1 floxed and excised bone marrow. **J)** Patients with acute myeloid leukemia from The Cancer Genome Atlas with STAG2 mutations have higher levels of STAG1 expression compared to patients without cohesin mutations ($p<0.006$). **K)** Peripheral blood counts of Stag2 WT and KO mice with heterozygous or homozygous co-deletion of Stag1 taken at 7 days following PIPC (at the time Stag2/Stag1 KO mice were moribund) reveal decreased leukocytes in Stag2/Stag1 KO ($p<0.001$) and platelets in both Stag2 KO/Stag1$^{-/+}$ ($p<0.001$) and Stag2/Stag1 KO ($p<0.001$). **L)** Representative metaphase spreads depicted for male Stag2/Stag1 KO with chromosomal catastrophe.

Supplemental Fig. 3

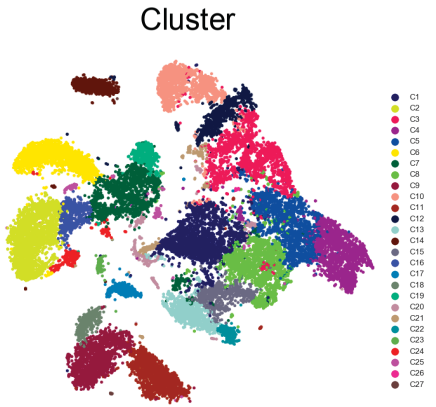
A



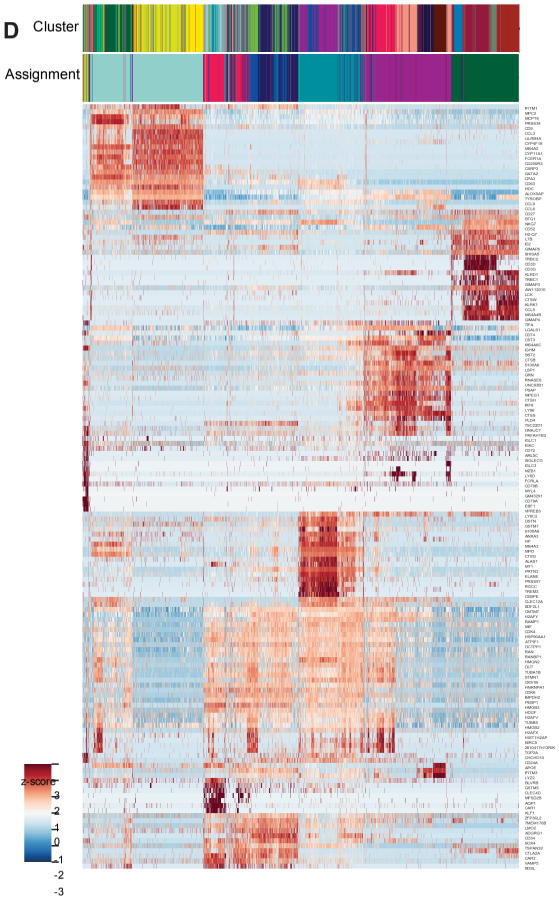
B



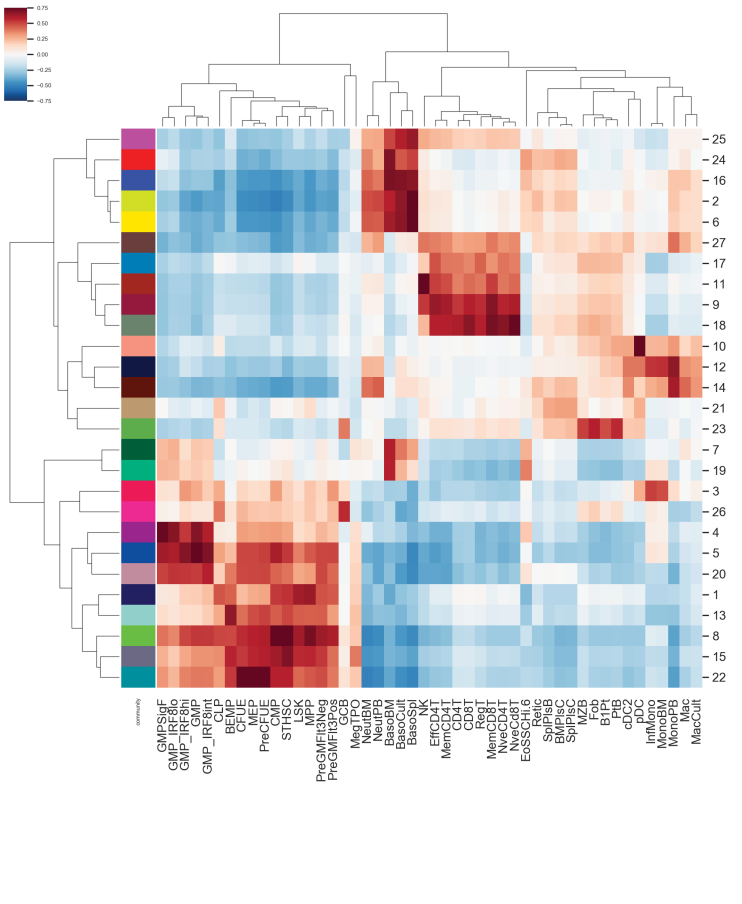
C



D



E

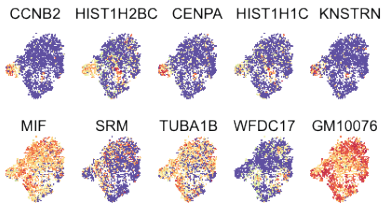


F

Positive Enrichment

Gene Ontology	Name	NES	FDR q-value
GO_0034728	NUCLEOSOME ORGANIZATION	3.9097435	0
GO_0006323	DNA PACKAGING	3.8820088	0
GO_0065004	PROTEIN-DNA COMPLEX ASSEMBLY	3.5663702	0
GO_0006334	NUCLEOSOME ASSEMBLY	3.5006585	0
GO_0031497	CHROMATIN ASSEMBLY	2.8557205	0
GO_0006333	CHROMATIN ASSEMBLY OR DISASSEMBLY	2.6403115	0
GO_0007059	CHROMOSOME SEGREGATION	2.521101	0

Cell cycle

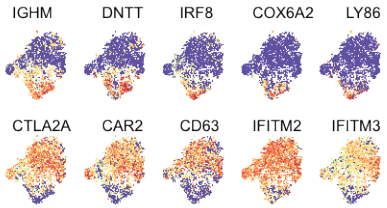


G

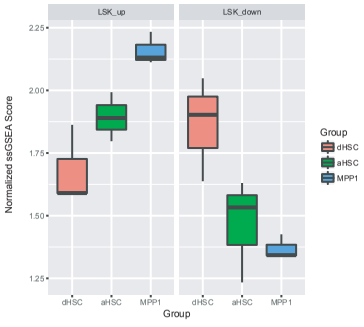
Negative Enrichment

Gene Ontology	Name	NES	FDR q-value
GO_0042113	B CELL ACTIVATION	2.9033923	0
GO_0030098	LYMPHOCYTE DIFFERENTIATION	2.5199957	0
GO_0042110	T CELL ACTIVATION	2.3701916	0
GO_0030217	T CELL DIFFERENTIATION	2.235279	0
GO_0002521	LEUKOCYTE DIFFERENTIATION	2.2332351	0
GO_0044087	REGULATION OF CELLULAR COMPONENT BIOGENESIS	2.0384235	0
GO_0042113	B CELL ACTIVATION	2.9033923	0

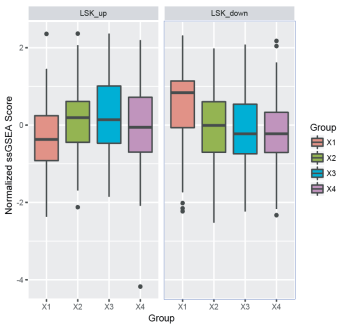
Lymphoid



H

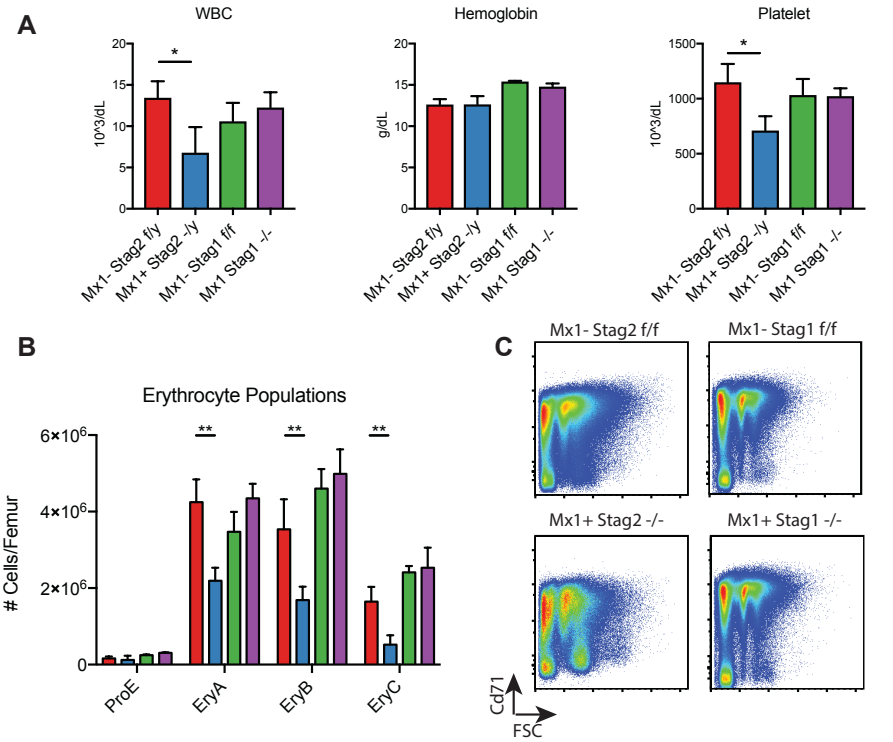


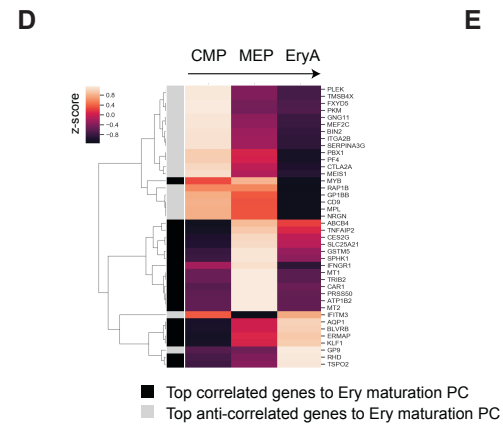
I

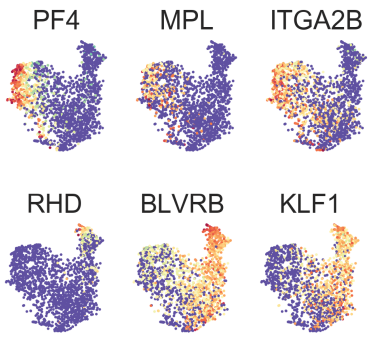


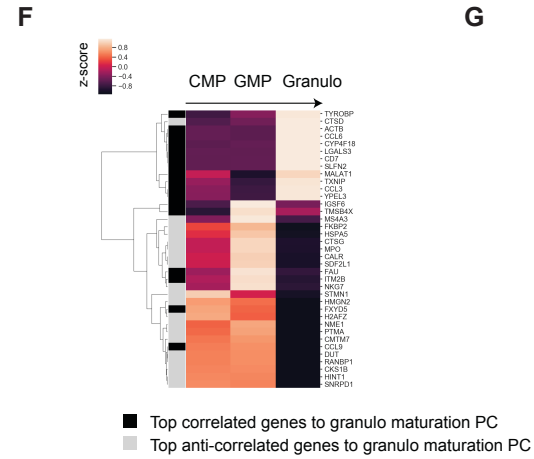
Supplemental Fig 3. (Related to Figure 1) A) Volcano plot for differentially expressed genes by RNA sequencing in LSK cells of Stag2 WT and Stag2 KO. Genes decreased in expression in Stag2 KO in blue (n=186) and genes increased in expression in Stag2 KO in red (n=42). **B)** Gene-set enrichment analysis of LSK RNAseq shows decreased expression of the GO Cell Fate Specification gene set. **C**) t-SNE projection of library-size normalized and log transformed data for complete collection (24,153 cells). Each dot represents a single cell colored by Phenograph clustering (Levine et al., 2015). **D**) Pearson correlation between centroids of Phenograph clusters to standardized bulk RNA-sequencing data from selected sorted mouse hematopoietic cells populations (from Haemopedia-Mouse RNAseq (de Graaf et al., 2016)). **E**) Heatmap of normalized and log transformed expression of top 20 differentially expressed genes per inferred lineage; cells and genes are hierarchically clustered. Differential expression was determined by calculating, for every gene, the Wasserstein distance between normalized and log-transformed expression in cells from inferred lineage and all other cells; top heatmaps show assignment of cells to clusters and inferred lineages labeled in Figure 2D. **F)** Gene set enrichment meeting threshold of NES >2, FDR <0.25 for positive enrichment (top) and **G)** negative enrichment (bottom) with gene ontology. T-SNE fot top 10 genes influencing the positive and negative principle component are shown. **H)** Single cell RNA seq data from Cabezas- Wallscheid et al. Genes increased in Stag2 KO LSK are enriched for genes expressed in active HSC and downregulation of quiescent genes. **I)** Comparisons to single cell RNAseq further refines the decreased expression in stage 1 of the dHSC (X1) compared to stages 2-4 (vs X2 $p=2.8 \times 10^{5}$; vs X3 $p=1.9 \times 10^{5}$; vs X4 $p=2.5 \times 10^{6}$). Statistical comparisons were generated using pairwise Student's t-test.

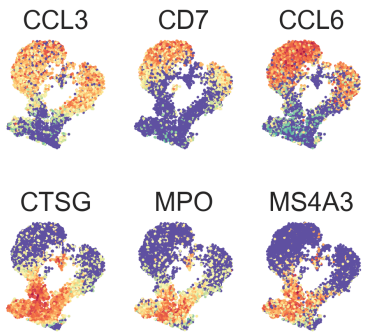
Supplemental Fig. 4









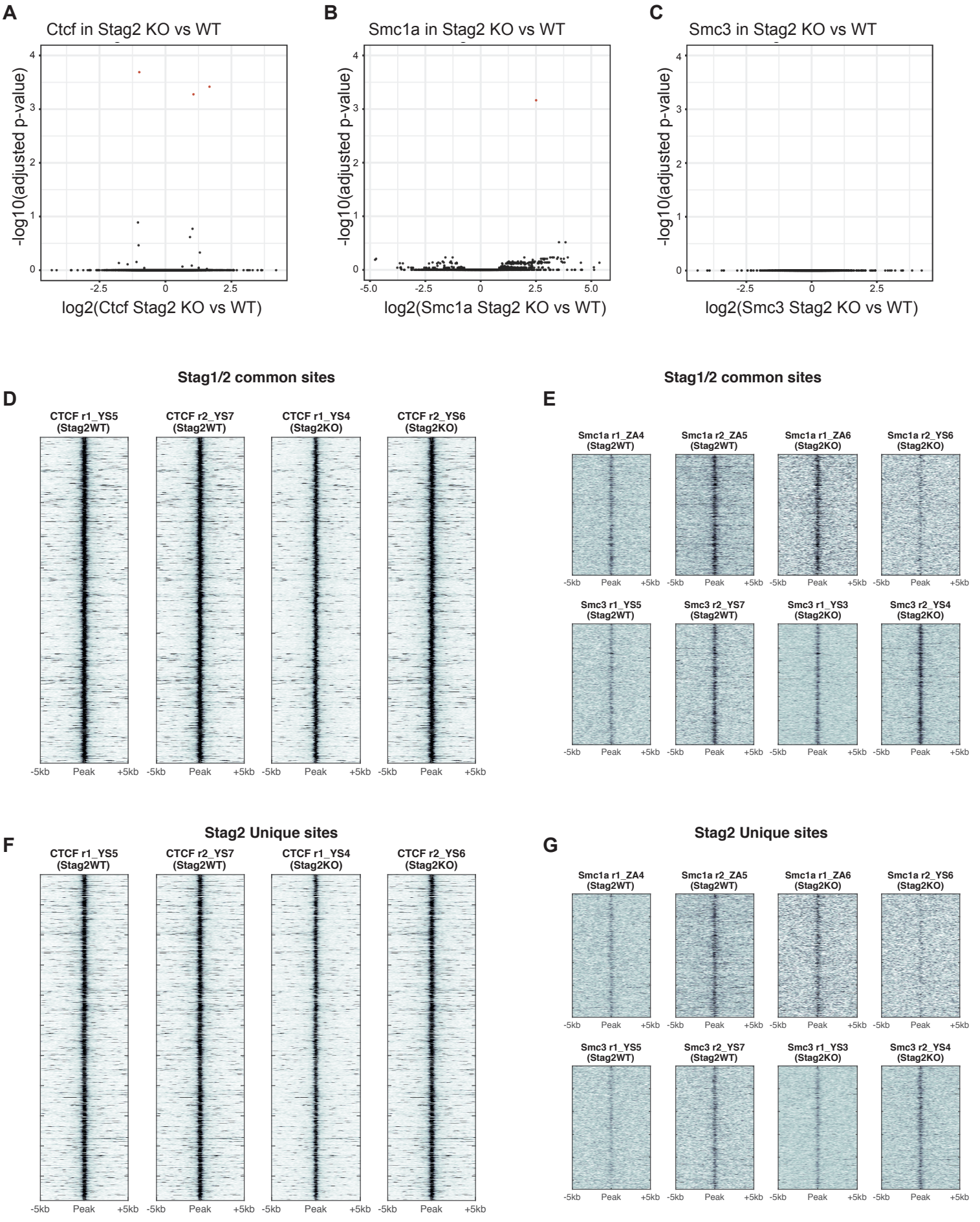


H

TF	Motif	% Target/ % Background	P-value
PU.1		62.5% / 26.6%	1e-15
CEBPB		28.3% / 5.97%	1e-13

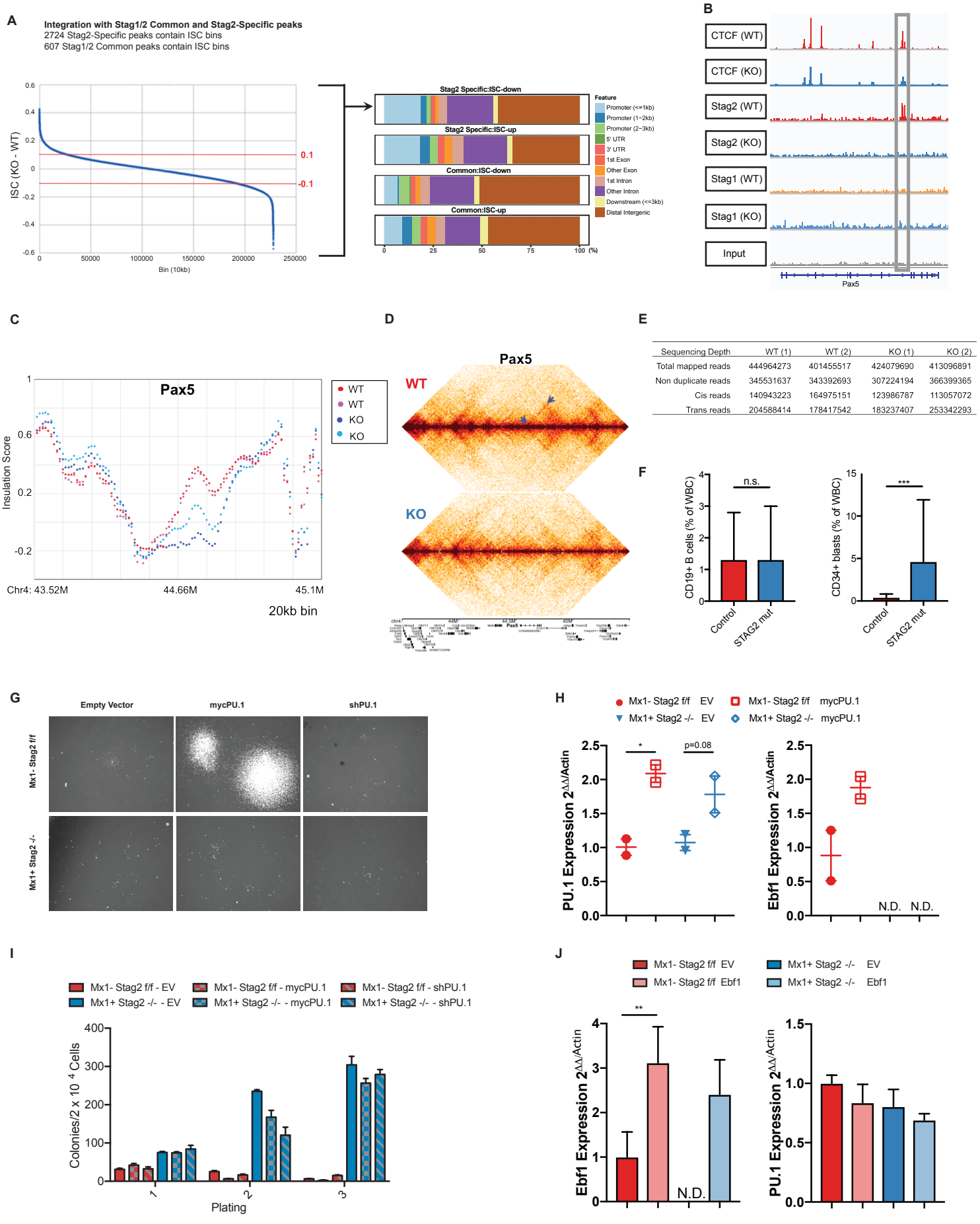
Supplemental Fig 4. (Related to Figure 2) A) Peripheral blood counts of Stag2 and Stag1 WT and KO mice taken at 8 weeks following PIPC treatment reveal decreased leukocytes ($p<0.02$) and platelets ($p<0.01$) in Stag2 KO mice and no changes in Stag1 KO mice. No differences in hemoglobin were seen in either Stag2 or Stag1 KO. **B)** Flow cytometric analysis of erythroid development using Ter-119 and Cd71 showing reduced mature erythroid populations (EryA $p=0.002$; EryB $p=0.007$; EryC $p=0.005$) in Stag2 KO bone marrow. **C)** Representative flow cytometry scatter-plots of Stag2 WT / KO and Stag1 WT / KO bone marrow showing Stag2 KO cells fail to decrease in FSC during maturation (Parent Gate on Ter-119$^+$ Cells). **D**) Heatmap of bulk RNA-sequencing data for cells with progressive erythroid lineage commitment (from GSE60101 (Lara-Astiaso et al., 2014)) showing normalized and standardized expression of genes most correlated or anticorrelated with erythroid maturation component. Correlations were computed using the Pearson method. **E**) t-SNE projection of library-size normalized and log transformed data for inferred MEP subset (1787 cells). Each dot represents a single cell colored by expression of labelled genes. **F**) Heatmap of bulk RNA-sequencing data for cells with progressive granulocyte lineage commitment (from GSE60101 (Lara-Astiaso et al., 2014)) showing normalized and standardized expression of genes most correlated or anticorrelated with granulocyte maturation component. Correlations were computed using the Pearson method. **G**) t-SNE projection of library-size normalized and log transformed data for inferred granulocyte subset (6316 cells). Each dot represents a single cell colored by expression of labelled genes. **H)** HOMER analysis of common loci of decreased accessibility by ATACseq for Smc3 heterozygous loss and Stag2 KO show enrichment for PU.1 and CEBPB motifs.

Supplemental Fig. 5



Supplemental Fig 5. (Related to Figure 3) Chromatin immunoprecipitation and sequencing for **A)** Ctcf **B)** Smc1a and **C)** Smc3 in Stag2 WT (n=2) and KO (n=2) HSPC. Volcano plots show lack of statistically significant differential loci. Heatmaps for **D)** Ctcf and **E)** Smc1a (top row)/Smc3 (bottom row) at Stag2/Stag1 commonly bound sites. Heatmaps for **F)** Ctcf and **G)** Smc1a (top row)/Smc3 (bottom row) at Stag2-uniquely bound sites. No differential occupancy in either Stag1/2 common or Stag2-unique sites were identified.

Supplemental Fig. 6



Supplemental Fig 6. (Related to Figure 5) A) Insulation score changes (ISC: Insulation score KO – Insulation score WT) from Hi-C analysis of Stag2 WT and KO Lin^- bone marrow using windows of 10kB. Stacked bar plot of genomic feature for common and Stag2-specific sites according to gain or loss of insulation. **B)** IGV track of the *Pax5* locus with Stag2 and Ctcf binding at an intergenic locus that is lost in Stag2 KO and not bound by Stag1 either in WT or KO. **C)** ISC plotted across Pax5 for Stag2 WT (n=2; shades of red) and KO (n=2; shades of blue) shows marked loss of insulation. **D)** Contact map of *Pax5* shows Stag2 KO cells lose local cis-interaction at two loci (arrows). **E)** Hi-C read counts for biologic replicates across WT and KO samples. The total valid interaction pairs for each sample are enumerated. No significant difference of sequencing depth was observed between samples. **F)** Enumeration of mature B cells (CD34-CD19+) and CD34+ blasts in STAG2 mutated MDS patients (n=11) compared to controls (n=15). **G)** Methylcellulose colony assay using IL-3, SCF, and IL-6 enriched media for stem cell replating. Stag2 WT and KO marrow were infected with lentivirus containing GFP-tagged empty vector, GFP-mycPU.1, or GFP-shPU.1. B-cell colonies were markedly larger and of higher cell output in Stag2 WT mycPU.1 transfected cells than in any other group. **H)** RT-PCR for *PU.1* and *Ebf1* after infection with GFP-tagged empty vector or GFP-mycPU.1 virus. Both Stag2 WT ($p<0.01$) and Stag2 KO ($P<0.08$) have increased *PU.1* expression in sorted GFP^+ cells. *Ebf1* increases in Stag2 WT cells infected with mycPU.1-GFP but *Ebf1* remains not detectable (N.D.) in Stag2 KO cells infected with either EV or mycPU.1 across 7 technical replicates. **I)** 20,000 GFP^+ cells from each condition were serially replated in M3434 Stemcell Methylcellulose. Overexpression and shRNA manipulation of PU.1 were unable to abrogate serial replating of the Stag2 KO cells. **J)** RT-PCR for *Ebf1* after infection with GFP-empty vector or Ebf1-GFP shows both Stag2 WT ($p<0.02$) and Stag2 KO (Not detectible in GFP-EV) have increased *Ebf1* expression in sorted GFP^+ cells. PU.1 expression did not change. Asterisks indicate statistical significance (student's t test, * $p<0.05$, ** $p<0.01$, *** $p<0.001$)

Table S1 (Related to Figure 5): Patient Characteristics

STAG2 mutant	Age	Gender	diagnosis	karyotype	Mutation
1	74	M	therapy related MDS/MPN, 7% blasts	Normal	STAG2/TET2/GATA2/SRSF2/ZRSR2
2	66	M	therapy related MDS, 6% blasts	normal	STAG2/DNMT3A/RUNX1/BCOR/BCORL1/U2AF1
3	67	M	MDS-EB1, 5% blasts	normal	STAG2/TET2/ASXL1/RUNX1/NRAS/SRSF2
4	68	M	CMML-1, 4% blasts	normal	STAG2/ASXL1/SF3B1/TET2/TP53/SRSF2
5	71	M	MDS-EB1, 6% blasts	normal	STAG2/TET2/RUNX1/NRAS/EZH2/PTPN11/ZRSR2
6	17	F	MDS-MLD, 2% blasts	normal	STAG2/BCOR/GATA2
7	58	F	therapy related MDS, 1% blasts	normal	STAG2/ASXL1/NRAS
8	53	F	MDS-EB1, 5% blasts	normal	STAG2/NPM1/CBL
9	77	M	MDS-MLD, 2% blasts	normal	STAG2/DNMT3A/ETV6/U2AF1
10	79	M	therapy related MDS-EB1, 9% blasts	47,XY,+8[20]	STAG2/TET2/RUNX1/NRAS/EZH2/ZRSR2
11	54	M	MDS-EB1, 8% blasts	normal	STAG2/ASXL1/IDH2/NRAS/SRSF2

Control	Age	Gender	diagnosis	karyotype	Mutation
1	82	F	anemia	normal	negative
2	78	F	anemia	normal	negative
3	81	M	anemia and thrombocytopenia	normal	negative
4	83	F	aplastic anemia	normal	negative
5	72	M	anemia	normal	negative
6	84	M	anemia	normal	negative
7	53	F	thrombocytopenia	normal	negative
8	52	M	anemia and leukopenia	normal	negative
9	64	M	thrombocytopenia	normal	negative
10	49	M	pancytopenia	normal	negative
11	54	F	pancytopenia	normal	negative
12	65	M	anemia and thrombocytopenia	normal	negative
13	72	F	anemia and thrombocytopenia	normal	negative
14	67	M	anemia and leukopenia	normal	negative
15	65	F	pancytopenia	normal	negative