Title: **Stag1 and Stag2 regulate cell fate decisions in hematopoiesis through non-redundant topological control**

Authors: Aaron D. Viny[1,2,3,4], Robert L. Bowman[1,3,4], Yu Liu[5], Vincent-Philippe Lavallée[6], Shira E. Eisman[1,3,4], Wenbin Xiao[7], Benjamin H. Durham[1,4,7], Anastasia Navitski[1,3,4], Jane Park[3], Stephanie Braunstein[1,3,4], Elham Azizi[6], Matthew Witkin[3], Timour Baslan[8], Christopher J. Ott[9], Dana Pe'er, Job Dekker[5,10], Richard Koche[3,4], Ross L. Levine[1,2,3,4,*]

Affiliations:

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[2]Department of Medicine, Leukemia Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[3]Center for Epigenetics Research, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[4]Center for Hematologic Malignancies, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[5]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA

[6]Center for Computational and Systems Biology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[7]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[8]Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[9]Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, MA, USA

[10]Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, MD, USA

*Correspondence to:
Ross L. Levine, MD
Human Oncology and Pathogenesis Program
Leukemia Service, Department of Medicine
Center for Epigenetics Research
Center for Hematologic Malignancies
Memorial Sloan Kettering Cancer Center
1275 York Ave, Box 20
New York, NY 10065
leviner@mskcc.org

Abstract: Transcriptional regulators, including the cohesin complex member STAG2, are recurrently mutated in cancer. The role of STAG2 in gene regulation, hematopoiesis, and tumor suppression remains unresolved. We show Stag2 deletion in hematopoietic stem/progenitor cells (HSPC) results in altered hematopoietic function, increased self-renewal, and impaired differentiation. ChIP-sequencing revealed that while Stag2 and Stag1 can bind the same loci, a component of Stag2 binding sites are unoccupied by Stag1 even in Stag2-deficient HSPCs. While concurrent loss of Stag2 and Stag1 abrogated hematopoiesis, Stag2 loss alone decreased chromatin accessibility and transcription of lineage-specification genes, including Ebf1 and Pax5, leading to blunted HSPC commitment to the B-cell lineage. Our data illustrate a role for Stag2 in transformation and transcriptional dysregulation distinct from its shared role with Stag1 in chromosomal segregation.

One Sentence Summary: Stag1 rescues topologically associated domains in the absence of Stag2, but cannot restore chromatin architecture required for hematopoietic lineage commitment

Main text

Cell-type specific transcriptional programs are regulated in part by the activity of tissue-specific transcription factors and enhancer elements within structurally defined topologically associating domains (TADs) (*1-3*). The genes which contribute to transcriptional regulation, including

5   members of the cohesin complex, are frequently mutated in human cancers, including leukemias(*4-9*). Cohesin is a tripartite ring comprised of three structural proteins SMC1A, SMC3, and RAD21; this core complex binds to either STAG1 or STAG2. *STAG2,* which is an X-linked gene, is the most commonly mutated cohesin complex member in human cancer, with recurrent nonsense, frameshift, and missense mutations in Ewing's Sarcoma (40-60%)(*9*), bladder cancer

10   (20-35%)(*7, 8*), glioblastoma (4-5%)(*7*), myelodysplastic syndrome (MDS)(5-20%)(*5, 6*), and acute myeloid leukemia (AML)(2-12%)(*5*). This is in contrast to SMC1A, SMC3, and RAD21, which are more commonly mutated in myeloid malignancies than in a broad spectrum of human cancers. The cohesin complex is essential in pleiotropic cellular and gene regulatory functions, including chromosome segregation(*10, 11*) and in the formation of DNA loops(*3, 12-14*) which

15   regulate gene expression(*13, 15, 16*). However, the high mutational frequency and unique mutational spectrum of *STAG2* in human cancers(*5-7*) suggests a distinct role in homeostasis and tumor suppression(*17, 18*) which has not been delineated.

Results

20   **Loss of Stag2 alters hematopoietic stem cell function.**

In order to elucidate the role of *Stag1* and *Stag2* in hematopoietic function/transformation, we generated *Stag1* and *Stag2* conditional knockout alleles. We induced somatic deletion of *Stag1* or *Stag2* in hematopoietic cells through *Mx1*-Cre mediated excision (Figure S1A-H). Following

25   polyinosinic:polycytadinic (PIPC)-mediated deletion, *Mx1*-cre *Stag2* $^{-/-}$ mice compared to cre-negative *Stag2*$^{fl/fl}$ (referred henceforth as KO and WT respectively) manifested an increase in hematopoietic stem cells (HSC), including a significant (2.2-fold; p<0.01) increase in the

proportion of LSK (Lineage$^-$Sca1$^+$cKit$^+$) cells (Figure 1A). *Stag2* loss, but not Stag1 loss, induced an expansion of multipotent progenitors (MPP: LSK$^+$Cd150$^-$Cd48$^+$Cd127$^-$; p<0.04), short term HSCs (ST-HSC: LSK$^+$Cd150$^-$Cd48$^-$; p<0.001), and long-term HSCs (LT-HSC: LSK$^+$Cd150$^+$Cd48$^-$; p<0.02) (Figure S2A). We also observed a significant increase in granulocyte macrophage progenitors (GMP: Lineage$^-$cKit$^+$Sca1$^-$Cd34$^+$Fcγ^+; p<0.001; Figure S2B-C) in the *Stag2*-deficient hematopoietic compartment.

We next investigated the role of Stag2 and Stag1 in hematopoietic self-renewal. *Stag2*-deficient hematopoietic cells displayed increased serial replating capacity in methylcellulose culture assays (≥7 platings), whereas WT and *Stag1* KO cells were not able to form colonies beyond 3 platings (Figure 1B). We next performed competitive transplantation assays to assess the impact of Stag2/Stag1 loss on self-renewal *in vivo*. We observed reduced *Stag2*-deficient derived chimerism in the peripheral blood of mice reconstituted with Stag2 KO and WT competitor cells (p<0.001; Figure S2D-F). However, within the bone marrow of transplant recipient mice, we found a significant increase in *Stag2* KO derived chimerism, which was most significant in the LSK compartment (p<0.001) including in LT-HSCs (p<0.001; Figure 1C) but not in lineage positive cells (p≤0.37). *Stag1* loss did not impact *in vivo* self-renewal or differentiation output (Figure 1D; S2G-I). These data suggest that *Stag2* loss induces both an increase in self-renewal and reduced differentiation capacity in HSCs, which are critical features of hematopoietic transformation.

To determine if Stag2 loss was associated with the expansion of hematopoietic cells with clonal cytogenetic alterations, we performed karyotype analysis in *Stag2* and *Stag1* WT and KO bone marrow cells. Chromosomal aberrations by metaphase karyotyping were not detected in either *Stag2*- or *Stag1*-deficient hematopoietic cells (n=7; Figure S3A-B). We did observe an increase in the proportion of cells with premature sister chromatid separation in *Stag2* KO bone marrow (Figure S3C); however, consistent with metaphase cytogenetics, we did not observe clonal copy

number alterations on low-pass whole genome sequencing (Figure S3D-E). This data suggested that the increase in hematopoietic self-renewal induced by *Stag2* loss is not due to genomic instability and resultant clonal evolution.

Stag2 and Stag1 co-deletion is synthetic lethal.

The observation that *Stag2*-deficient HSCs displayed increased self-renewal is in contrast to the lethality observed with loss of the core cohesin component *Smc3* in hematopoietic cells(*19*). We next sought to evaluate potential compensatory function among the Stag paralogs. We found *Stag1* transcript and protein expression was significantly increased in *Stag2* KO hematopoietic cells, whereas *Stag2* expression was not altered in *Stag1* KO hematopoietic cells (Figure S4A-D). Analysis of the AML TCGA dataset showed that AML patients with STAG2 mutations expressed higher levels of *STAG1* (p<0.006; Figure S4E), suggesting compensatory regulation. We hypothesized that Stag1 and Stag2 may have redundant roles with respect to chromatid segregation such that Stag2 loss does not fully abrogate cohesin function due to residual Stag1 function. *In vitro* studies have recently nominated STAG1 as a synthetic lethal target in *STAG2* mutant cell lines(*20*); however, this has not been assessed *in vivo*. We generated mice with hematopoietic cells retaining one or zero copies of *Stag1* on the background of a *Stag2* (hemizygous male) KO genotype. $Stag2^{-/y}/Stag1^{+/-}$ mice developed rapid, progressive thrombocytopenia (p<0.001) and impaired survival (p<0.001) due to bone marrow failure (median survival of 27.7 weeks). Complete loss of *Stag1* and *Stag2* in hematopoietic cells induced rapid lethality (median survival 0.7 weeks; Figure 1E; Figure S4F). *Stag1/Stag2* deficient mice were characterized by pancytopenia and bone marrow aplasia, consistent with redundancy in hematopoietic function and in alignment with the absolute requirement for core cohesin complex members (Figure 1F). Similar to mice with bi-allelic deletion of *Smc3*, we observed marked chromosomal alterations on metaphase spreads of the *Stag1/Stag2* deleted bone marrow (Figure

S4G). Collectively, these data suggest a role for Stag2 in gene regulation and tumor suppression that is not shared with Stag1, whereas they have a shared, redundant role in chromosomal segregation.

Stag2 alters transcriptional lineage commitment.

We hypothesized that *Stag2* loss induced alterations in transcriptional output which dysregulate hematopoietic function and promote transformation. To investigate transcriptional mechanisms regulating self-renewal and differentiation in the context of *Stag2* loss, we performed RNA-sequencing on *Stag2* KO and WT LSK cells. We identified 186 genes which were significantly downregulated and 42 genes which were significantly upregulated in *Stag2* KO compared to WT cells (Figure S5A). We observed reduced expression of *Ebf1, Vbrep3, Cd19, Bank1*, and *Pax5,* suggesting a loss of B-cell lineage commitment(*21, 22*) in *Stag2*-deficient stem/progenitor cells. We observed reduced expression of *Vwf(23), Hba-a1, Hba-a2, Hbb-b1(24), Fcgr2b(25), Ccr2(26), Spib(27), S100a8, and S100a9(28)*, genes associated with myeloid/erythroid lineage commitment. Gene set enrichment analysis (GSEA) revealed that Stag2 deficient LSKs showed negative enrichment of genes associated with cell lineage specification (Figure S5B). We further found increased expression of the Hoxb4-target *Hemgn(29)*, as well as of *Itgb3*(*30*), a Periostin-binding integrin shown to mediate HSC self-renewal. Coupled with downregulation of *Flt3(31)* and *Selp*(*32*) (P-selectin), these findings indicate an LSK pool that is transcriptionally skewed towards increased self-renewal.

Given the heterogeneity of the LSK compartment, we sought to determine if the transcriptional alterations observed in bulk LSKs were due to a population wide shift in expression or a result of differential expression within a subset of LSKs. We performed single cell RNA-sequencing (scRNAseq) of ~24,000 Lin⁻ hematopoietic stem/progenitor cells (HSPC) from Stag2 KO (n=3)

and WT (n=3) mice (cohort annotations in Figure S5C-E). We found that within the HSC population, representing ~8% of cells, that the compartment distribution was skewed with KO HSCs having enriched expression of cell cycle gene signatures ($p<10^{-63}$; Figure 1G, S5F top panel). Concomitantly, these HSC showed decreased expression of early lymphoid commitment signatures ($p<10^{-14}$; Figure 1G, S5F bottom panel) even at the earliest HSC states, consistent with negative enrichment for cell specification. We expanded upon these findings by comparing bulk LSK differentially expressed genes to gene expression data from dormant and active HSCs separately isolated through genetic lineage tracing studies(*33*). We observed that genes that were decreased in expression in *Stag2*-deficient mice were most highly expressed in slow-cycling quiescent HSCs, while the upregulated genes were most highly expressed in the proliferative MPP1 fraction (LSK, Cd34$^+$, Flt3$^-$) (Figure S5G). Similar findings were observed when comparing our data to single cell RNA-seq data, such that the most downregulated genes in *Stag2*-deficient LSKs were most highly expressed in the most quiescent HSCs (Figure S5H). Collectively, these analyses highlight a transcriptional signature associated with decreased lineage commitment, decreased quiescence, and enhanced self-renewal in the setting of STAG2 loss.

Stag2 loss impairs differentiation and induces dysplasia.

To determine the consequences of these transcriptional alterations *in vivo*, we assessed the phenotype of mice with *Stag2*-deficient hematopoiesis over time. Hematopoietic-specific *Stag2* loss resulted in progressive leukopenia (p<0.02) and thrombocytopenia (p<0.01; Figure S6A). We observed an expansion in myeloid cells and a significant reduction in B lymphocytes (p<0.001; Figure 2A) in the peripheral blood, which was not seen with *Stag1* deletion. *Stag2* loss, but not *Stag1* loss, induced morphologic alterations including an increase in immature myeloid cells, an expansion of small hypolobated megakaryocytes, and nuclear:cytoplasmic dyssynchrony in the erythroid lineage (Figure 2B). These abnormalities are features of human myelodysplasia—

consistent with the observation that 5-15% of MDS patients have loss-of-function mutations in *STAG2*(5, 6).

To determine at which level of differentiation multi-lineage dysplasia emerges, we analyzed lineage composition through analysis of scRNAseq data (Figure 2C-D). Supporting the reduction in B cell output observed in the peripheral blood, analysis of scRNAseq data suggested a significant decrease in B cells and in B cell progenitors (Figure 2E). Although the paucity of lymphoid cells precluded further transcriptional tracing, we observed population shifts in erythroid and granulocyte development consistent with observed histologic myelodysplasia. Flow cytometric analysis of erythroblast maturation revealed attenuation at Stage III of erythropoiesis(*34*) (Figure 2F, S6B-C), a stage that is demarcated by condensed cell size as measured by forward scatter (FSC) and decreased CD44 expression relative to prior stages of erythroid differentiation. *Stag2*-deficient erythroid cells were also characterized by morphologic dysplasia; immunohistochemical analysis demonstrated a marked decrease in the proportion of Ter-119[+] erythroid cells (Figure 2G). By comparison, *Stag1* KO mice did not manifest hematopoietic alterations, morphologic abnormalities, or defects in erythroid differentiation (Figure S6A-C). Transcriptional analysis of the erythroid lineage in *Stag2* deficient cells further supported impaired terminal maturation (p<10[-44]; Figure 2H). Genes defining this principle component define the progression from common myeloid progenitor to megakaryocyte erythroid progenitor to erythroid A (Figure S6D-E). Overall, the most significant transcriptional difference identified in the scRNAseq was impaired granulocytic terminal maturation (p<10[-92]; Figure 2I). Genes defining this principle component define the progression from common myeloid progenitor to granulocyte macrophage progenitor to granulocyte (Figure S6F-G).

Given the alterations in transcriptional output and differentiation potential, we hypothesized that *Stag2* loss led to altered chromatin accessibility at differentially expressed loci. We therefore

performed RNA sequencing and ATAC-sequencing on sorted *Stag2* WT and KO GMP and erythroid progenitor (CFU-E: Lineage⁻cKit⁺Sca1⁻Cd34⁻Fcγ⁻Cd105⁺) cells. Compared to WT GMPs, *Stag2*-deficient GMPs had enrichment for a stem-cell gene expression signature associated with increased *Evi1* expression and negative enrichment of myeloid cell development signatures by GSEA (Figure 2J). Similarly, *Stag2*-deficient CFU-Es had increased expression of genes normally expressed in HSCs and in the megakaryocyte lineage, consistent with the increased self-renewal and impaired differentiation seen with Stag2 loss (Figure 2K). Integrated analysis of RNA-seq and ATAC-seq revealed that a significant subset of the downregulated genes also had reduced chromatin accessibility. Both GMPs and CFU-E shared loss of accessibility and expression for genes with either CTCF/CTCFL (BORIS) or PU.1(ETS) motifs (Figure 2L-M). CTCF is a well-described cohesin binding partner(*35*) that cooperates with cohesin to mediate enhancer/promoter associations(*12*). Notably, PU.1 and its target genes have a critical role in determining hematopoietic lineage commitment(*36, 37*).

Stag1/2 possess shared and independent chromatin binding.

It has been shown using *in vitro* model systems that cohesin is essential for genomic cis-interactions and that rapid depletion of cohesin ring members result in complete loss of interphase architecture(*3, 14, 38*). However, these model systems do not reflect a physiologic loss/partial loss of a cohesin complex member akin to the genomic alterations observed in either human cancers or the germline cohesinopathy, Cornelia de Lange syndrome(*39*). As the complete loss of cohesin structural loop components may not be compatible with a viable cell division event, the absence of transcriptional changes observed using *in vitro* systems of cohesin degradation(*14*) may not reflect the pathophysiology of Stag2 loss of function. We next sought to determine how *Stag2*-loss affected chromatin topology in primary cells using Hi-C chromosome conformation capture in hematopoietic progenitor cells. *Stag2* KO cells did not show alterations in topologically

associating domain (TAD) boundaries, in global insulation, and did not show differences in the spatially segregated genomic regions known as A/B compartments comparted to WT hematopoietic progenitor cells(*40*) (Figure 3A-C).

5   Given the lack of changes in global structure, we sought to determine if Stag2 loss would affect the ability of Ctcf of the remaining structural cohesin components Smc1a and Smc3 to localize to canonical cohesion/CTCF binding sites. Chromatin immunoprecipitation sequencing revealed that the occupancy of these three proteins was largely unaffected by the presence or absence of Stag2 in HSPCs (Figure S7A-C). We hypothesized that complete loss of Stag2 might not alter

10   higher order DNA topology due to redundancy with Stag1 at CTCF bound sites. We therefore performed ChIP-seq to delineate Stag2 and Stag1 occupancy in HSPCs from *Stag2* WT and KO mice. We identified a set of peaks where Stag2 and Stag1 were bound in WT mice and where Stag1 occupancy increased in the absence of Stag2 (n=2,493; Figure 3D, S7D-E). Conversely, we identified a second set of peaks where Stag2, but not Stag1, was bound in WT HSPCs and

15   lost in *Stag2* KO cells without compensatory Stag1 occupancy (n=10,809; Figure 3E, S7F-G). Previous work has demonstrated CTCF-dependent and independent localization of the cohesin complex(*41, 42*) which in epithelial cell lines have preferential binding sites for Stag1 and Stag2(*43*). We hypothesized that altered regulation of Stag2-bound loci in WT cells which are not bound by Stag1 in WT or KO cells, were key to the altered differentiation and self-renewal seen

20   in Stag2 KO mice. We found that loci with both Stag2 and Stag1 binding had the highest enrichment for CTCF binding (Figure 3E-F).

Loss of chromatin insulation impairs transcriptional output.

25   Stag1 was able to bind a subset of loci previously occupied by Stag2 or both Stag1 and Stag2 in WT cells; this included TAD boundaries determined by Hi-C (Figure 3F). By contrast, when we

focused on Stag2-only bound sites, we observed locus-specific alterations in short-range interactions (as calculation by the insulation metric; Figure S6B). We quantified these alterations through insulation scores (IS) as a measure of cumulative local contacts and insulation score changes (ISC) that reflect local chromosomal organization changes. To examine whether loss of Stag2 binding affects local chromosomal organization, we compared ISC in WT and *Stag2* KO HSPCs and identified 71,815 ISC bins that showed local chromosomal contact changes (10kb bins with insulation score change less than -0.1 or greater than 0.1, Figure S8A, left panel). We then overlapped ISC bins with Stag1/2 co-bound and Stag2-specific peaks. Annotation analysis of both Stag1/Stag2 common and Stag2-specific peaks showed enrichment of Stag2-specific ISC peaks within promoter regions (Figure S8A, right panel), suggesting that ISC seen in *Stag2*-deficient cells contribute to altered gene expression.

These data suggested that the role of Stag2 in gene regulation and in hematopoiesis might be due to regulation of loci bound by factors other than CTCF. To determine the impact of *Stag2* loss on chromatin structure and transcriptional output we performed RNA-seq and ATAC-seq in WT and Stag2 KO HSPCs. Similar to gene expression/ATAC-seq data in sorted hematopoietic subsets, the majority of genes with significantly altered chromatin accessibility and gene expression in the absence of *Stag2* had decreased accessibility and reduced transcription (Figure 4A). Several of these peaks were in proximity to critical B-cell and myeloid lineage defining transcription factors, including *Ebf1, Pax5*, and *Cebpb*. To uncover transcription factor networks underlying these changes, we performed motif enrichment on peaks with concomitant reduction in accessibility/gene expression. We observed enrichment for Gata, Ebf, and E2A targets (Figure 4B), but not for Cebpb motifs. The genes associated with loss of insulation by ISC significantly overlapped with those loci with decreased accessibility and decreased expression, with *Ebf1* and *Pax5* among the genes with the greatest magnitude of insulation change (Figure 4C, S8B-D). We therefore performed motif analysis on Stag1/2 common and Stag2-specific sites by dichotomizing

those that were bound and not bound by CTCF. As expected, loci with shared Stag1/Stag2 occupancy and CTCF binding sites were enriched for CTCF and CTCFL (BORIS) binding motifs. By contrast, Stag2-specific loci (not bound by Stag1) were enriched for key lineage-specific transcription factor motifs including ERG, IRF8, and PU.1 (Figure 4D). These transcription factors are essential for myeloid and lymphoid development (*44, 45*) and attenuated function of these transcription factors has transformative potential in murine(*37*) and human hematopoietic cells(*46-48*).

Altered Ebf1 chromatin structure results in B-cell blockade.

We next sought to determine if the altered expression of hematopoietic transcription factors was due to alterations in the population distribution in HSPC or was due to altered expression of these factors in purified hematopoietic populations, including in multipotent progenitors antecedent to full lineage commitment. We measured *PU.1* and *Ebf1* expression and B cell differentiation in MPPs and in common lymphoid progenitors (CLP: LSK$^+$Cd150$^-$Cd48$^+$ Cd127$^+$) from *Stag2* KO and WT mice*.* We found that *Ebf1* expression was reduced/absent in purified *Stag2* KO MPPs (p<0.04, Log$_2$FC=-1.8) and CLPs (not detectible in KO), while there was no significant difference in *PU.1* expression in either cell type (Figure 4E). Consistent with these data, B cell maturation was blocked from the pro-B cell to early pre-B cell stage(*49*) in *Stag2* KO mice, with almost no late pre-B cells (Figure 5A). These findings were further bolstered by *in vitro* colony assays in B cell cytokine enriched methylcellulose where *Stag2* KO HSPCs failed to generate B cell colonies (Figure 5B). We next investigated if the same was true in Stag2-mutant MDS patients by enumerating B cells and B cell progenitors in STAG2-mutant MDS patient bone marrow samples compared to a cohort of non-MDS individuals with cytopenias (Figure 5C, S8E). Immature B cells (CD34+CD19+) were decreased by 1.7-fold (p<0.01) in STAG2-mutant MDS patients, demonstrating that altered B-cell lineage commitment is observed in STAG2-mutant MDS.

Given that PU.1 can regulate myeloid and lymphoid lineage commitment(*37*) and the enrichment for PU.1 motif/targets in our epigenomic/transcriptional analysis, we investigated the role of PU.1 in Stag2-deficient alterations in B-cell lineage commitment and self-renewal. PU.1 silencing reduced B-lineage output derived from WT hematopoietic cells, similar to the effects of Stag2 deletion (Figure 5D). However, the reduced B-cell lineage output of *Stag2* KO cells was not further attenuated by PU.1 silencing as assessed by flow cytometric analysis of B220 expression and B-lineage colony formation (Figure 5D). By contrast, overexpression of PU.1 was not sufficient to restore B220 expression in B cell methylcellulose cultures (Figure 5D) or to attenuate the increased expression of the stem/progenitor marker cKit seen on *Stag2*-deficient hematopoietic cells (Figure 5E). Moreover, overexpression of PU.1 failed to attenuate the capacity of *Stag2*-deficient hematopoietic cells to serially replate (Figure S8F-H).

This led us to hypothesize that key transcription factors downstream of PU.1 which are bound by Stag have locus specific alterations in conformation in the absence of Stag2 that lead to alterations in target gene expression with functional import. *Ebf1* has four Stag2 bound sites in hematopoietic progenitors, all of which are not bound by Stag1 in WT or KO cells (Figure 5F). As noted above, we observed near-complete abrogation of *Ebf1* transcription from the onset of B-lineage commitment (Figure 4E). We observed a significant reduction in local insulation throughout the *Ebf1* locus (Figure 5G) including loss of cis-interactions at three Stag2 binding sites (Figure 5H). Consistent with this hypothesis, expression of *Ebf1* in Stag2 KO progenitors (Figure S8I) fully restored B-lineage marker expression and B-cell colony formation *in vitro* (Figure 5I). Moreover, *Ebf1*-expressing Stag2 KO progenitors had decreased cKit expression (Figure 5J) and *Ebf1* restoration abrogated the serial replating capacity of *Stag2*-deficient cells (Figure 5K). Taken together, these data demonstrate a key role for Stag2-mediated regulation of Ebf1 expression in modulating the balance between hematopoietic differentiation and self-renewal.

Discussion

The cohesin complex is essential in pleiotropic cellular and gene regulatory functions, including in chromosome segregation(*10, 11*) and in maintaining DNA interactions(*3, 14*) which regulate gene expression(*13, 15, 16*). Here we demonstrate a non-redundant role of Stag2 and Stag1 in hematopoiesis, with a specific role for Stag2 in regulating the balance between hematopoietic differentiation and self-renewal. In addition, we show that Stag2 bound loci that are not bound by Stag1 undergo locus specific changes in local interactions that, in the absence of Stag2, leads to altered expression of key loci governing hematopoietic function. Taken together, these data illustrate a critical link between Stag2, DNA interactions, and gene regulation that contributes to hematopoietic transformation.

Stag2 and Stag1 have both shared and distinct roles(*43*). These paralogs have a redundant role in chromatid separation such the presence of either protein within the cohesin complex is adequate for intact sister chromatid alignment. In contrast to deletion of core cohesin ring components, which abrogates the survival of hematopoietic cells(*19*), somatic deletion of either *Stag1* or *Stag2* was compatible with hematopoietic stem cell viability, whereas loss of both genes led to hematopoietic failure and defective chromosomal segregation. Our *in vivo* studies underscore previous *in vitro* dependency studies(*20*) and suggest that therapies which directly or indirectly alter Stag1 function will induce a specific dependency in Stag2 mutant cancer cells. However, increased self-renewal, impaired lineage commitment, and progressive hematopoietic alterations were observed with Stag2 deletion but not with Stag1 loss. These studies illustrate that Stag2 has a key role in hematopoiesis and in hematopoietic transformation that is not shared by Stag1. Future studies will illustrate how Stag2 loss of function cooperates with other alleles to

promote leukemic transformation and how co-occurring disease alleles cooperate with the chromatin architecture alterations induced by *Stag2* loss.

Our data suggest that the role of Stag2 in regulating hematopoiesis and in hematopoietic transformation is Stag1-independent. Key hematopoietic regulators which show differential expression, accessibility, and insulation in the absence of Stag2 are bound by Stag2 and not by Stag1. This includes several key genes essential for lineage commitment to lymphoid fate, including the PU.1 targets *Ebf1* and *Pax5*. Previous studies demonstrated that PU.1-mediated regulation of *Ebf1* is a key event in B-cell lineage commitment(*50, 51*), and PU.1 has been defined to be a pioneer factor with nucleosome remodeling capacity. Given these data and the known role of Ebf1 in B cell commitment/differentiation (*22*), we show that PU.1 restoration cannot rescue the impact of Stag2 loss on hematopoietic self-renewal and lineage commitment without the permissive chromatin context regulated by Stag2. By contrast, restoration of Ebf1 expression restores differentiation and abrogates the aberrant self-renewal of Stag2 deficient cells.

Our data illustrate a key role for Stag2 in the balance between self-renewal and differentiation, a critical feature of transformation across different tumor types. Moreover, the observation that reactivation of the Stag2-target Ebf1 can reverse these alterations suggests a key role for Stag2-regulated transcription factor networks in tumor suppression, which is reversible when key networks are restored in Stag2 mutant cells. We hypothesize that Stag2 loss-of-function induces lineage specific alterations in transcription factor expression and function in a spectrum of malignant contexts. Most importantly, our studies identify a key pathophysiologic role for mutations in the factors that govern gene regulatory architecture in malignant transformation. Our data point to the reversibility of transcriptional dysregulation which mediates aberrant self-renewal and differentiation in cancer cells and suggest that therapies that reverse these alterations may have therapeutic benefit in cohesin-mutant cancers.

15

Reference:

1.	E. P. Nora *et al.*, Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-385 (2012).
2.	J. R. Dixon *et al.*, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380 (2012).
3.	S. S. P. Rao *et al.*, Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320 e324 (2017).
4.	N. Cancer Genome Atlas Research *et al.*, Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* **368**, 2059-2074 (2013).
5.	S. Thota *et al.*, Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood* **124**, 1790-1798 (2014).
6.	A. Kon *et al.*, Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat Genet* **45**, 1232-1237 (2013).
7.	M. L. Bailey *et al.*, Glioblastoma cells containing mutations in the cohesin component STAG2 are sensitive to PARP inhibition. *Mol Cancer Ther* **13**, 724-732 (2014).
8.	C. Balbas-Martinez *et al.*, Recurrent inactivation of STAG2 in bladder cancer is not associated with aneuploidy. *Nat Genet* **45**, 1464-1469 (2013).
9.	D. A. Solomon *et al.*, Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* **333**, 1039-1043 (2011).
10.	A. Losada, M. Hirano, T. Hirano, Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. *Genes Dev* **12**, 1986-1997 (1998).
11.	A. Toth *et al.*, Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. *Genes Dev* **13**, 320-333 (1999).
12.	S. S. Rao *et al.*, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).
13.	V. C. Seitan *et al.*, Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res* **23**, 2066-2077 (2013).
14.	W. Schwarzer *et al.*, Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51-56 (2017).
15.	M. H. Kagey *et al.*, Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-435 (2010).
16.	S. Cuartero *et al.*, Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat Immunol*, (2018).
17.	C. Mazumdar *et al.*, Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation. *Cell Stem Cell* **17**, 675-688 (2015).
18.	J. Mullenders *et al.*, Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *J Exp Med* **212**, 1833-1850 (2015).
19.	A. D. Viny *et al.*, Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *J Exp Med* **212**, 1819-1832 (2015).

20. L. Benedetti, M. Cereda, L. Monteverde, N. Desai, F. D. Ciccarelli, Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1. *Oncotarget* **8**, 37619-37632 (2017).

21. J. S. Herman, Sagar, D. Grun, FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods* **15**, 379-386 (2018).

22. S. Zandi *et al.*, EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. *J Immunol* **181**, 3364-3372 (2008).

23. A. Grover *et al.*, Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nat Commun* **7**, 11075 (2016).

24. P. D. Kingsley *et al.*, Ontogeny of erythroid gene expression. *Blood* **121**, e5-e13 (2013).

25. F. Ferrari *et al.*, Genomic expression during human myelopoiesis. *BMC Genomics* **8**, 264 (2007).

26. S. Reid *et al.*, Enhanced myeloid progenitor cell cycling and apoptosis in mice lacking the chemokine receptor, CCR2. *Blood* **93**, 1524-1533 (1999).

27. G. H. Su *et al.*, Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. *EMBO J* **16**, 7118-7129 (1997).

28. E. Lagasse, I. L. Weissman, Mouse MRP8 and MRP14, two intracellular calcium-binding proteins associated with the development of the myeloid lineage. *Blood* **79**, 1907-1915 (1992).

29. J. Jiang *et al.*, Hemgn is a direct transcriptional target of HOXB4 and induces expansion of murine myeloid progenitor cells. *Blood* **116**, 711-719 (2010).

30. T. Umemoto *et al.*, Expression of Integrin beta3 is correlated to the properties of quiescent hemopoietic stem cells possessing the side population phenotype. *J Immunol* **177**, 7733-7739 (2006).

31. J. Adolfsson *et al.*, Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* **15**, 659-669 (2001).

32. C. Sullivan *et al.*, Functional ramifications for the loss of P-selectin expression on hematopoietic and leukemic stem cells. *PLoS One* **6**, e26246 (2011).

33. N. Cabezas-Wallscheid *et al.*, Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell* **169**, 807-823 e819 (2017).

34. K. Chen *et al.*, Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis. *Proc Natl Acad Sci U S A* **106**, 17413-17418 (2009).

35. K. S. Wendt *et al.*, Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796-801 (2008).

36. I. Antony-Debre *et al.*, Pharmacological inhibition of the transcription factor PU.1 in leukemia. *J Clin Invest* **127**, 4297-4313 (2017).

37. B. Will *et al.*, Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia. *Nat Med* **21**, 1172-1181 (2015).

38. J. H. I. Haarhuis *et al.*, The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**, 693-707 e614 (2017).

39. J. Liu, I. D. Krantz, Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet* **76**, 303-314 (2009).

40. E. Lieberman-Aiden *et al.*, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293 (2009).

41. V. Parelho *et al.*, Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422-433 (2008).

42. E. D. Rubio *et al.*, CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* **105**, 8309-8314 (2008).

43. A. Kojic *et al.*, Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. *Nat Struct Mol Biol* **25**, 496-504 (2018).

44. S. R. McKercher *et al.*, Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J* **15**, 5647-5658 (1996).

45. H. Wang *et al.*, A reporter mouse reveals lineage-specific and heterogeneous expression of IRF8 during lymphoid and myeloid cell differentiation. *J Immunol* **193**, 1766-1777 (2014).

46. B. U. Mueller *et al.*, ATRA resolves the differentiation block in t(15;17) acute myeloid leukemia by restoring PU.1 expression. *Blood* **107**, 3330-3338 (2006).

47. R. K. Vangala *et al.*, The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. *Blood* **101**, 270-277 (2003).

48. E. L. Pogosova-Agadjanyan *et al.*, The prognostic significance of IRF8 transcripts in adult patients with acute myeloid leukemia. *PLoS One* **8**, e70812 (2013).

49. J. N. Pucella *et al.*, miR-182 is largely dispensable for adaptive immunity: lack of correlation between expression and function. *J Immunol* **194**, 2635-2642 (2015).

50. K. L. Medina *et al.*, Assembling a gene regulatory network for specification of the B cell fate. *Dev Cell* **7**, 607-617 (2004).

51. E. W. Scott *et al.*, PU.1 functions in a cell-autonomous manner to control the differentiation of multipotential lymphoid-myeloid progenitors. *Immunity* **6**, 437-447 (1997).

52. R. Kuhn, F. Schwenk, M. Aguet, K. Rajewsky, Inducible gene targeting in mice. *Science* **269**, 1427-1429 (1995).

53. M. Lakso *et al.*, Efficient in vivo manipulation of mouse genomic sequences at the zygote stage. *Proc Natl Acad Sci U S A* **93**, 5860-5865 (1996).

54. J. de Boer *et al.*, The E2A-HLF oncogenic fusion protein acts through Lmo2 and Bcl-2 to immortalize hematopoietic progenitors. *Leukemia* **25**, 321-330 (2011).

55. O. Abdel-Wahab *et al.*, Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *The Journal of experimental medicine* **210**, 2641-2659 (2013).

56. T. Baslan *et al.*, Optimizing sparse sequencing of single cells for highly multiplex copy number profiling. *Genome Res* **25**, 714-724 (2015).

57. T. Baslan *et al.*, Genome-wide copy number analysis of single cells. *Nat Protoc* **7**, 1024-1041 (2012).

58. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

59. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

60. E. Azizi *et al.*, Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* **174**, 1293-1308 e1236 (2018).

61.   J. H. Levine *et al.*, Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184-197 (2015).

62.   C. A. de Graaf *et al.*, Haemopedia: An Expression Atlas of Murine Hematopoietic Cells. *Stem Cell Reports* **7**, 571-582 (2016).

63.   E. Del Barrio, J. A. Cuesta-Albertos, C. Matran, J. M. Rodriguez-Rodriguez, Tests of goodness of fit based on the L-2-Wasserstein distance. *Ann Stat* **27**, 1230-1239 (1999).

64.   U. Steidl *et al.*, Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. *Nat Genet* **38**, 1269-1277 (2006).

65.   M. R. Corces *et al.*, Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature Genetics* **48**, 1193-1203 (2016).

66.   H. Belaghzal, J. Dekker, J. H. Gibcus, Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56-65 (2017).

67.   J. M. Belton *et al.*, Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**, 268-276 (2012).

68.   B. R. Lajoie, J. Dekker, N. Kaplan, The Hitchhiker's guide to Hi-C analysis: practical guidelines. *Methods* **72**, 65-75 (2015).

69.   G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).

70.   D. Lara-Astiaso *et al.*, Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943-949 (2014).

and Novartis. He has received honoraria from Roche, Lilly and Amgen for invited lectures and from Gilead for grant reviews. R.L.B, Y.L., B.H.D., S.E.E., A.N., M.W., T.B., C.J.O., and R.K. have nothing to disclose.

5   **Supplementary Materials:**
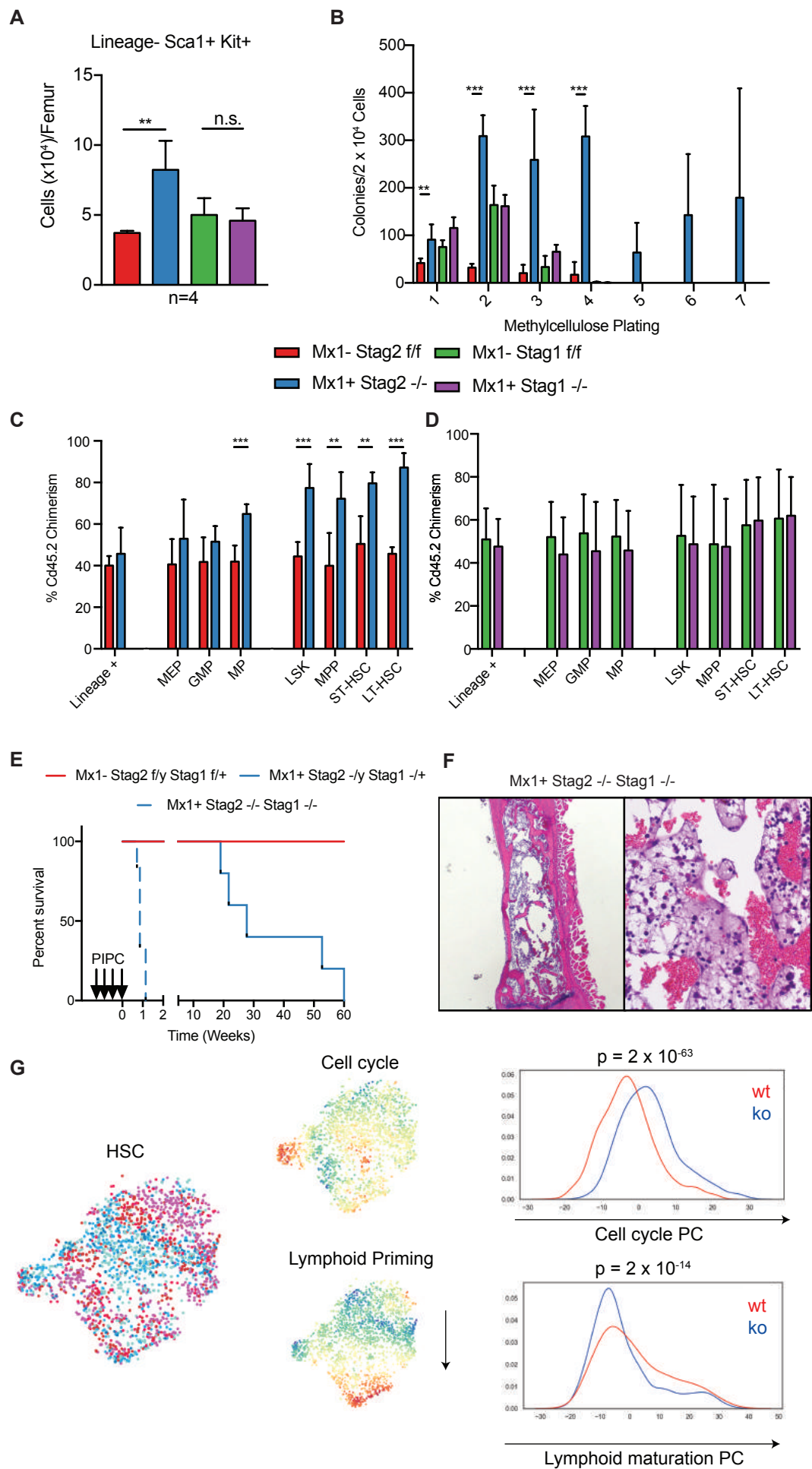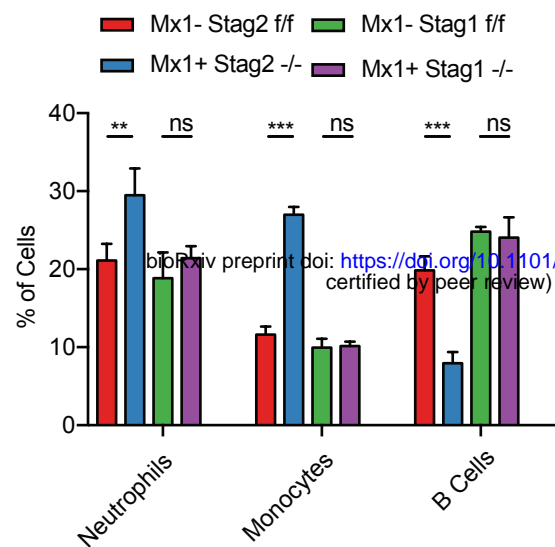
Materials and Methods

Figures S1-S8

Fig. 1

A — Lineage- Sca1+ Kit+

Cells (x10⁴)/Femur

n=4

B

Colonies/2 x 10⁴ Cells

Methylcellulose Plating

- Mx1- Stag2 f/f
- Mx1+ Stag2 -/-
- Mx1- Stag1 f/f
- Mx1+ Stag1 -/-

C — % Cd45.2 Chimerism

Lineage +   MEP GMP MP   LSK MPP ST-HSC LT-HSC

D — % Cd45.2 Chimerism

Lineage +   MEP GMP MP   LSK MPP ST-HSC LT-HSC

E

Percent survival

- Mx1- Stag2 f/y Stag1 f/+
- Mx1+ Stag2 -/y Stag1 -/+
- Mx1+ Stag2 -/- Stag1 -/-

PIPC

Time (Weeks)

F — Mx1+ Stag2 -/- Stag1 -/-

G

HSC

Cell cycle

Lymphoid Priming

p = 2 x 10⁻⁶³

wt / ko

Cell cycle PC

p = 2 x 10⁻¹⁴

wt / ko

Lymphoid maturation PC

Fig 1. Hematopoietic specific loss of Stag2, but not Stag1, results in altered stem cell function.
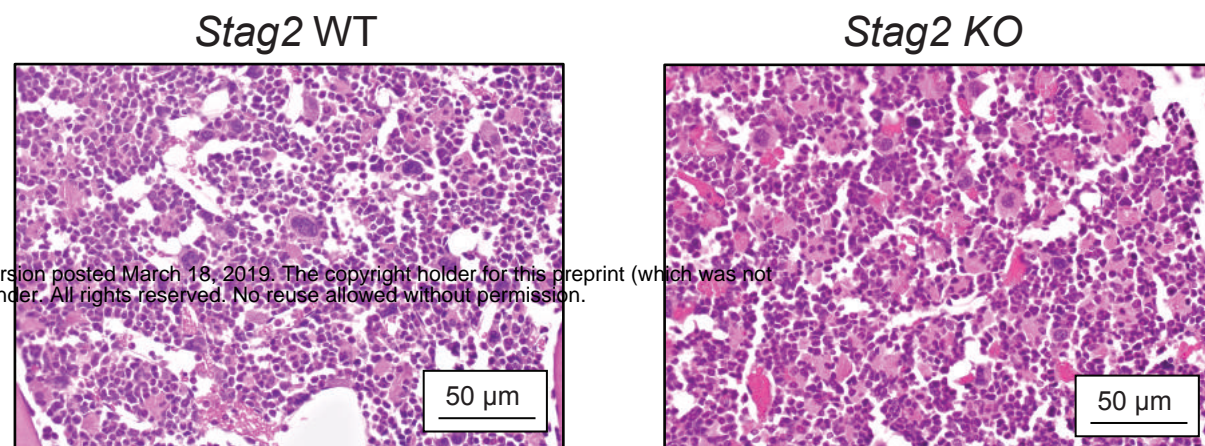
A) Stag2 KO, but not Stag1 KO mice have increased LSK (Lin$^-$Kit$^+$Sca1$^+$) hematopoietic stem cells (Log$_2$FC=2.2; p<0.01**) B)** Whole bone marrow plated in cytokine-enriched methylcellulose reveals that Stag2 KO marrow has increased self-renewal capacity and can serially replate, but WT and Stag1 KO cannot. Competitive bone marrow transplantation of **C)** Stag2 WT (Red)/KO (Blue) or **D)** Stag1 WT (Green)/KO (Purple) bone marrow mixed 1:1 with Cd45.1 normal marrow shows increased chimerism in stem and progenitor populations in the bone marrow at 16 weeks. **E)** Kaplan-Meier curve shows Stag2/Stag1 is lethal with a median survival of 0.7 weeks (p=0.01), while Stag2 KO/Stag1$^{-/+}$ has a lethal phenotype with a median survival of 27.7 weeks (p=0.05). **F)** Hematoxylin and eosin (H&E) staining of Stag2/Stag1 KO bone marrow reveals marked cellular aplasia. **G**) Left: t-SNE projection of library-size normalized and log transformed data for inferred HSC subset (2025 cells). Each dot represents a single cell colored by genetic condition (Stag2 KO: shades of blue, Stag2 WT: shades of red). Middle: t-SNE projection colored by the second and third principal components most correlated with lymphoid priming and cell cycle, respectively. Right: Distribution of cells along lymphoid and cell cycle components; p-values for Mann-Whitney U test are shown.
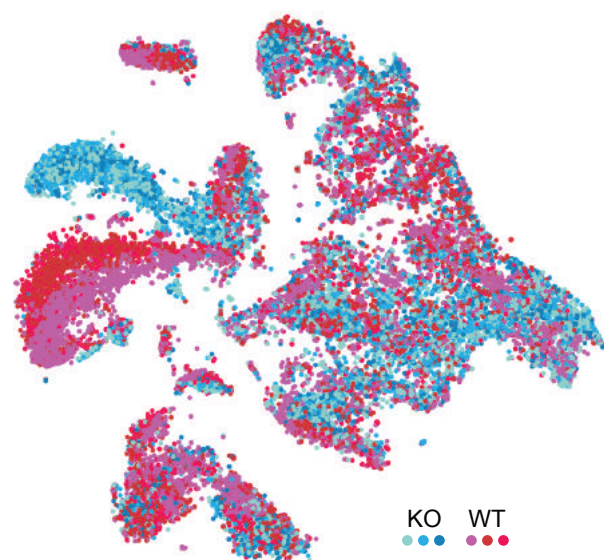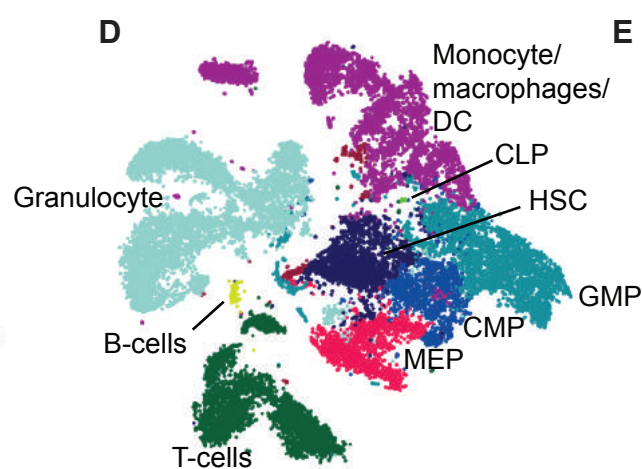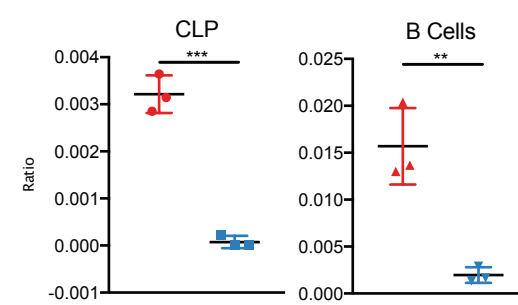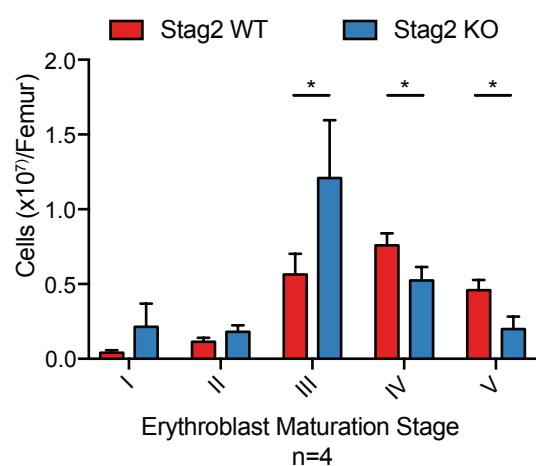
Fig. 2

A

Legend: Mx1- Stag2 f/f (red), Mx1+ Stag2 -/- (blue), Mx1- Stag1 f/f (green), Mx1+ Stag1 -/- (purple)

Y-axis: % of Cells

Categories: Neutrophils, Monocytes, B Cells
Significance: Neutrophils ** ns; Monocytes *** ns; B Cells *** ns

B

Stag2 WT — *Stag2 KO* (H&E images, 50 µm scale bars)

C

UMAP plot — KO, WT

D

UMAP with labels: Monocyte/macrophages/DC, CLP, HSC, Granulocyte, GMP, CMP, MEP, B-cells, T-cells

E

CLP *** ; B Cells **
Y-axis: Ratio

F

Legend: Stag2 WT (red), Stag2 KO (blue)
Y-axis: Cells (x10⁷/Femur)
X-axis: Erythroblast Maturation Stage (I, II, III, IV, V) n=4
Significance: III *, IV *, V *

G

Ter119 IHC — Stag2 WT (100x), Stag2 KO (100x)

H

Erythrocytes — Erythroid maturation
p = 3 x 10⁻⁴⁴
X-axis: Ery maturation PC
wt (red), ko (blue)

I

Granulocytes — Granulocyte maturation
p = 8 x 10⁻⁹²
X-axis: Granulocytic maturation PC
wt (red), ko (blue)

J Granulocyte Monocyte Precursors

Enrichment plot: EVI1_05
FDR q-value: 1.0
NES: 1.34

Enrichment plot: BROWN_MYELOID_CELL_DEVELOPMENT_UP
FDR q-value: 0.27
NES: -1.70

K CFU-E Committed Erythroid Progenitors

Enrichment plot: GEORGANTAS_HSC_MARKERS
FDR q-value: 1.0
NES: 1.55

Enrichment plot: RAGHAVACHARI_PLATELET_SPECIFIC_GENES
FDR q-value: 1.0
NES: 1.46

L

TF	Motif	% Target/ % Background	P-value
CTCF		24.3% / 5.76%	1e-38
Sfpi1		49.67% / 23.755	1e-32
BORIS		26.0% / 8.70%	1e-27

M

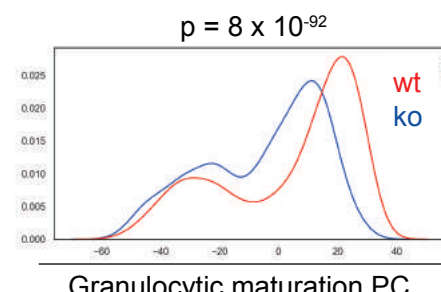| TF	Motif	% Target/ % Background	P-value
PU.1(ETS)		18.06% / 1.72%	1e-9
CTCF		19.4% / 3.05%	1e-7
BORIS		20.83% / 6.23%	1e-6

Fig 2. Stag2 alters transcriptional lineage commitment in stem and progenitor cells.

A) Peripheral blood myeloid skewing with increase in Neutrophils (Gr1$^+$Mac1^{hi}), Monocytes (Gr1$^+$Mac1^{lo}), and decrease in B cells (B220$^+$) in Stag2 KO mice. No statistical differences were observed in Stag1 KO mice. **B)** Hematoxylin and eosin (H&E) staining of sternal bone marrow from Stag2 KO mice show small, hypolobated megakaryocytes, increased immature myeloid cells, and nuclear-to-cytoplasmic dyssynchrony in erythropoiesis. Single cell RNA sequencing was performed on Stag2 WT (n=3; shades of red) and Stag2 KO (n=3; shades of blue) Lin$^-$ HSPC. **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 genetic condition (Stag2 KO: shades of blue, Stag2 WT: shades of red). **D**) t-SNE map colored by inferred cell type, as detailed in Supplemental Figure 4. **E)** Frequency of CLPs (left) and B-cells (right) in Stag2 WT (red) and KO (blue) samples; asterisks indicate statistical significance (student's t test, **p<0.01, ***p<0.001) **F)** Flow cytometric quantification of erythroblast stage in Stag2 WT and KO bone marrow reveals increased erythroblast stage III (p=0.05), but reduction of mature erythroblast stage IV (p=0.02) and stage V (p=0.01). **G)** Immunohistochemical analysis of Stag2 WT and KO bone marrow reveals a marked reduction in Ter-119 expression in Stag2 KO. **H**) Top left: 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 genetic condition (Stag2 KO: blues, Stag2 WT: reds). Top right: t-SNE projection colored by the second principal component most correlated with erythroid maturation. Bottom: Distribution of cells along erythroid maturation component. **I**) Top left: t-SNE projection of library-size normalized and log transformed data for inferred granulocytic subset (6316 cells). Each dot represents a single cell colored by genetic condition (Stag2 KO: blues, Stag2 WT: reds). Top right: t-SNE projection colored by the first principal component most correlated with granulocyte maturation. Bottom: Distribution of cells along granulocyte maturation component. **J)** Gene-set enrichment analysis of GMP RNAseq shows increased expression of genes in the Evi1 gene set and decreased expression of myeloid development genes. **K)** CFU-E

RNAseq shows enrichment for the Georgantas HSC markers gene set and retention of platelet specific genes. Integration of differentially expressed genes in **L)** GMP and **M)** CFU-E populations with ATAC-sequencing commonly have genes with CTCF/CTCFL (BORIS) motif signatures and PU.1 (ETS) motif signatures with decreased expression and decreased accessibility in both populations.
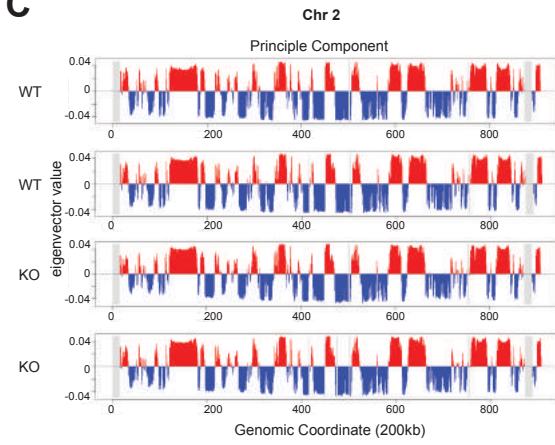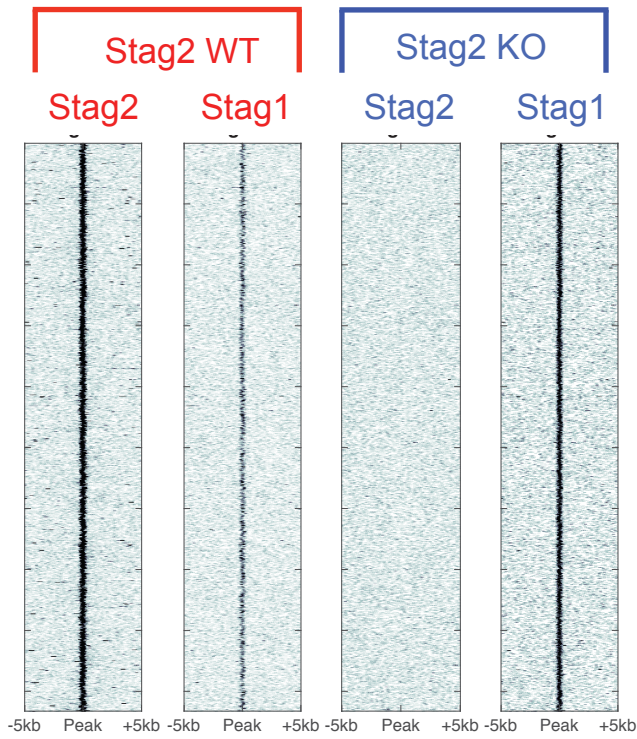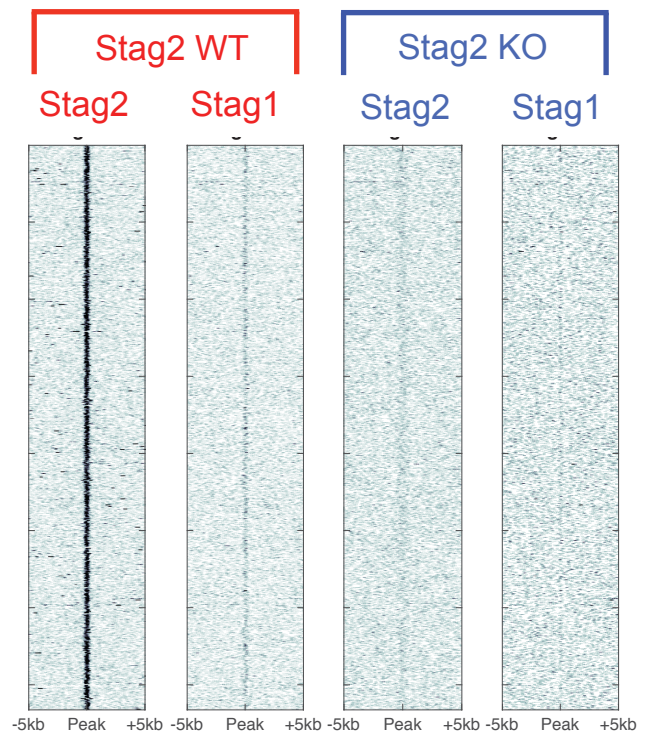
5
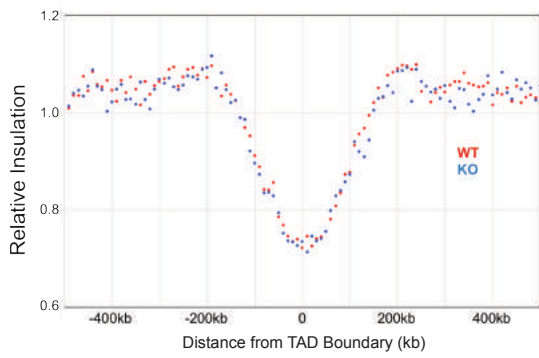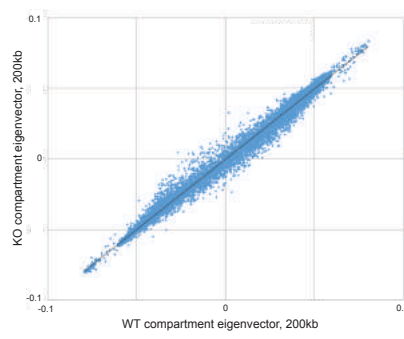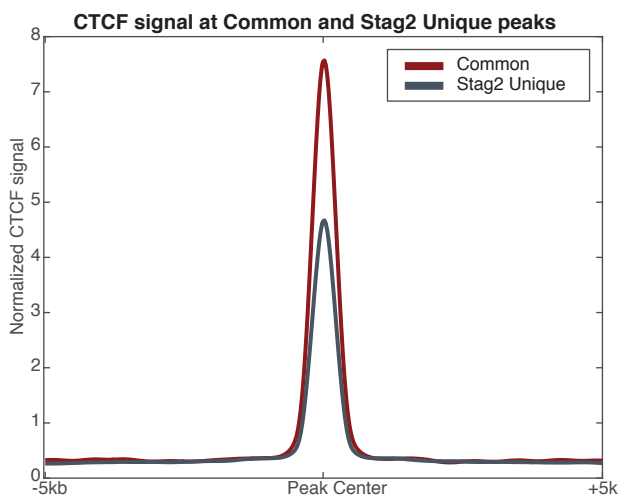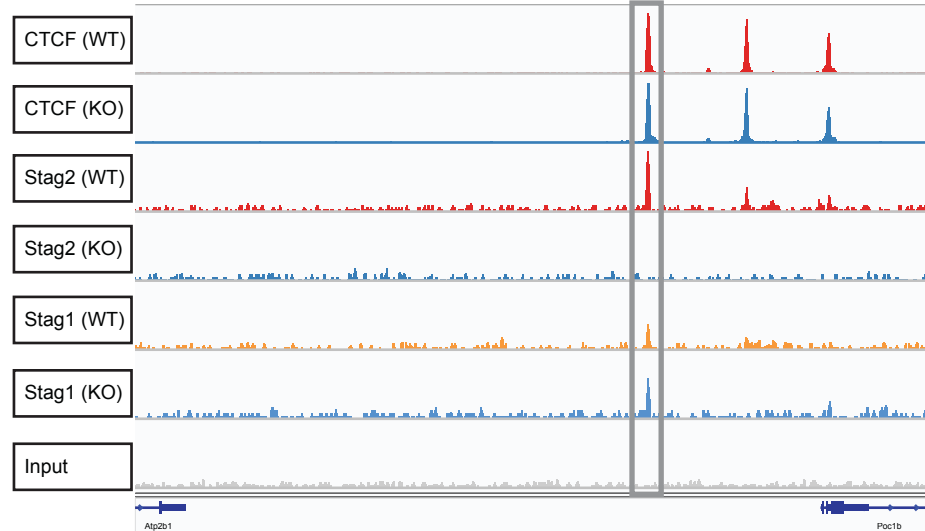
Fig. 3

Fig 3. Stag2 and Stag1 possess shared and independent chromatin binding.

A) Insulation pileup plot from each dot shows the median of the interactions for each 200kb bins located 500kb upstream or downstream of a TAD border, normalized by the median interactions for WT and KO, respectively. **B)** Hi-C insulation scores were normalized by the median interactions for WT and KO, respectively. The diagonal bins were excluded for calculating the median. **C)** A/B compartments of chr2 for each biological replicate. The A compartments, positive PC1 signals, are highlighted in red, while the B compartments, negative PC1 signals, are highlighted in blue. **D)** Chromatin immunoprecipitation and sequencing for Stag1 and Stag2 in Stag2 WT and KO HSPC show a discrete subset of the genome where Stag1 is able to bind Stag2 loci in its absence (common sites) and **E)** discrete loci where Stag1 is unable to bind in the absence of Stag2 (Stag2-specific sites). **F)** Enrichment of Stag1/2-common sites and Stag2-specific sites for CTCF shows strong enrichment in both sets. **G)** IGV track at TAD boundary as measured by Hi-C insulation on chromosome 10 showing CTCF occupancy as well as Stag2 in WT with Stag1 occupancy increased in Stag2 KO HSPC.

Fig. 4

A

B

TF	Motif	% Target/ % Background	P-value
Gata	GAGATAAG	42.8% / 18.8%	1e-36
EBF	GGTCCCTGGGGG	15.6% / 5.68%	1e-16
E2A	GTGCCAGCTG	24.0% / 13.0%	1e-11

C

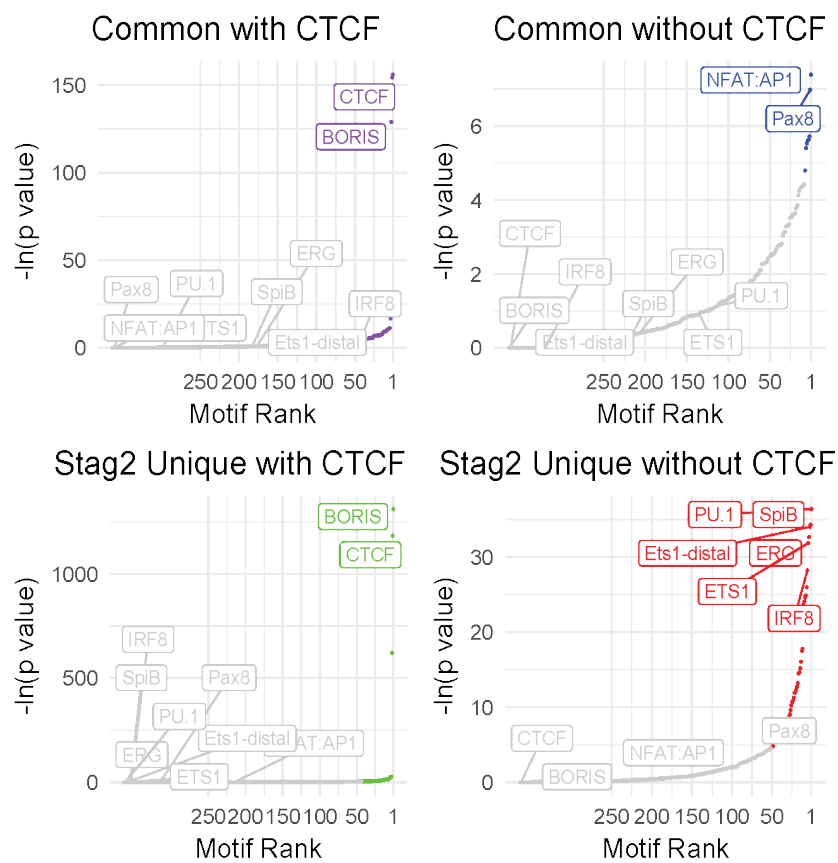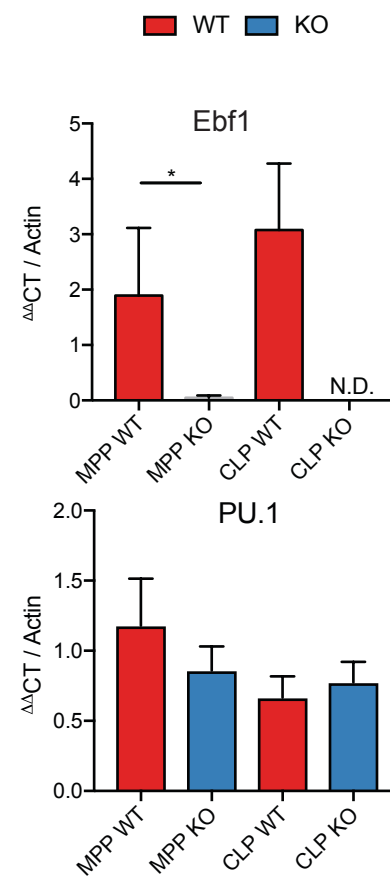| GeneID	ISC	Stag2 WT Peaks
Fgf1	-2.6	2
Pax5	-2.4	2
Lonrf2	-2.2	4
Gfra2	-2.2	2
Fcgr1	-2.1	2
Cpm	-2.0	2
Dtx1	-1.9	3
Ebf1	-1.9	2
0610040J01Rik	-1.9	2
Bach2	-1.4	6

D

E

Fig 4. Stag2 loss decreases chromatin insulation and results in impaired transcriptional output. A) Intersection of RNAseq and ATACseq in lineage negative bone marrow plotted at Log_2 fold change of Stag2 KO compared to WT and each point representing a one ATAC peak. The majority of differentially expressed genes lose accessibility, including distinct B-cell regulators (e.g. Ebf1, Pax5, Cd19). **B)** HOMER motif analysis of genes in lower left quadrant of Figure 4A from intersection of RNAseq and ATACseq in lineage negative bone marrow. Genes that are downregulated and lose accessibility are targets of Gata ($p=10^{-36}$), EBF ($p=10^{-16}$), and E2A ($p=10^{-11}$). **C)** Table of top 10 genes with greatest magnitude of ISC and number of WT Stag2 peaks in the gene. **D)** Motif analyses of common and Stag2-specific sites with and without CTCF show common enrichment for CTCF and CTCFL (BORIS), only Stag2-unique sites bind targets of key lineage priming factors PU.1, SpiB, ERG, ETS1, and IRF8, which Stag1 is unable to bind. **E)** Putative expressers of PU.1 and Ebf1 from Stag2 WT and KO bone marrow were sorted and *Ebf1* and *Pu.1* expression was measured by RT-PCR in multipotent progenitors (MPP, p=0.04) and common lymphoid progenitors (CLP, no mRNA detectable in KO). Ebf1 expression was decreased in both populations (MPP p=0.04; CLP not detectable). Pu.1 expression was not statistically different in either population (MPP p=0.13; CLP p=0.32).
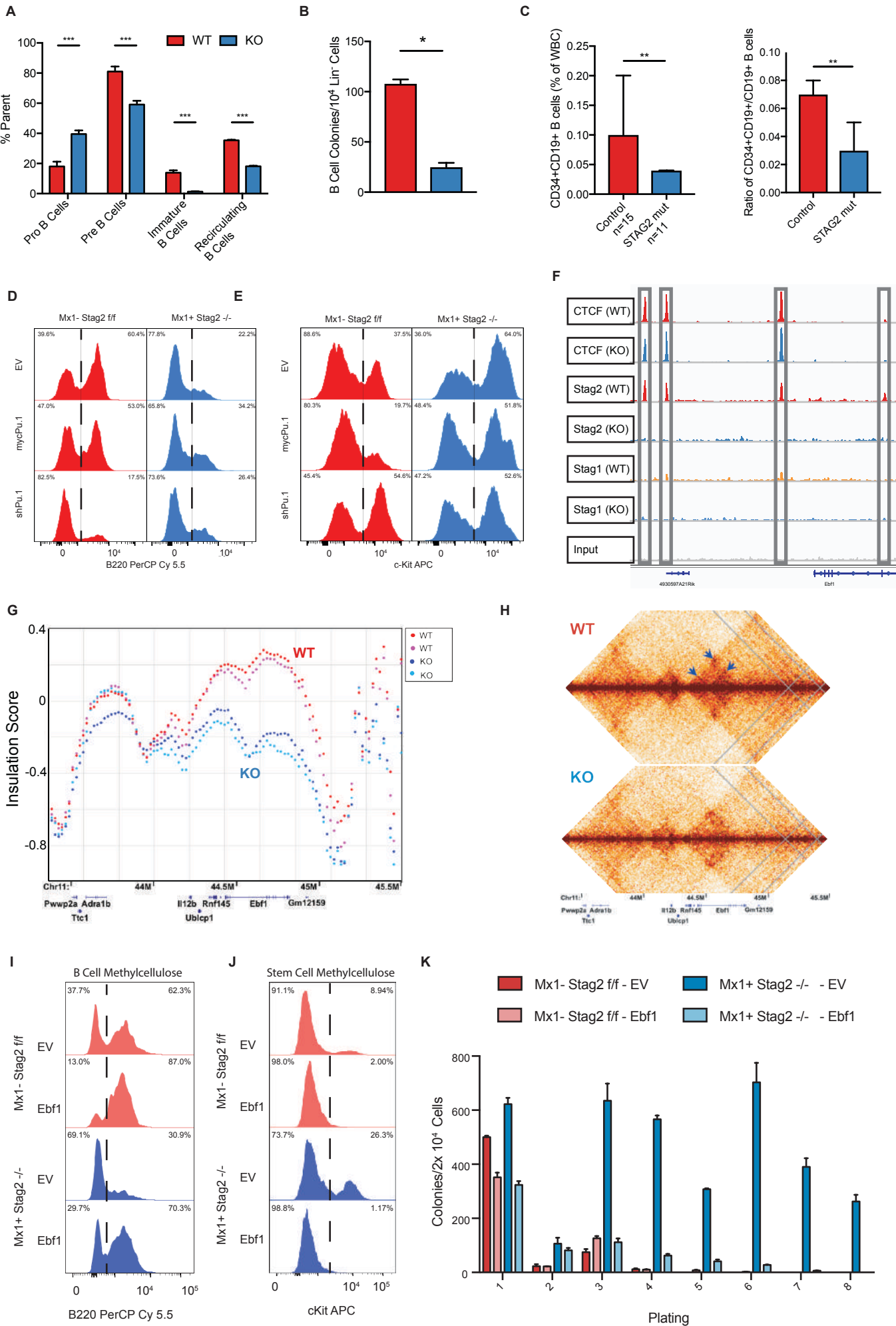
Fig. 5

A

% Parent

■ WT   ■ KO

Pro B Cells, Pre B Cells, Immature B Cells, Recirculating B Cells

***

B

B Cell Colonies/10⁴ Lin⁻ Cells

*

C

CD34+CD19+ B cells (% of WBC)

Control n=15, STAG2 mut n=11

**

Ratio of CD34+CD19+/CD19+ B cells

Control, STAG2 mut

**

D

Mx1- Stag2 f/f   Mx1+ Stag2 -/-

EV   39.6% 60.4%   77.8% 22.2%
mycPu.1   47.0% 53.0%   65.8% 34.2%
shPu.1   82.5% 17.5%   73.6% 26.4%

B220 PerCP Cy 5.5

E

Mx1- Stag2 f/f   Mx1+ Stag2 -/-

EV   88.6% 37.5%   36.0% 64.0%
mycPu.1   80.3% 19.7%   48.4% 51.8%
shPu.1   45.4% 54.6%   47.2% 52.6%

c-Kit APC

F

CTCF (WT)
CTCF (KO)
Stag2 (WT)
Stag2 (KO)
Stag1 (WT)
Stag1 (KO)
Input

4930597A21Rik   Ebf1

G

Insulation Score

WT

KO

■ WT ■ WT ■ KO ■ KO

Chr11:   44M   44.5M   45M   45.5M
Pwwp2a Adra1b   Il12b Rnf145 Ebf1 Gm12159
Ttc1   Ublcp1

H

WT

KO

Chr11:   44M   44.5M   45M   45.5M
Pwwp2a Adra1b   Il12b Rnf145 Ebf1 Gm12159
Ttc1   Ublcp1

I

B Cell Methylcellulose

Mx1- Stag2 f/f
EV   37.7% 62.3%
Ebf1   13.0% 87.0%

Mx1+ Stag2 -/-
EV   69.1% 30.9%
Ebf1   29.7% 70.3%

B220 PerCP Cy 5.5

J

Stem Cell Methylcellulose

Mx1- Stag2 f/f
EV   91.1% 8.94%
Ebf1   98.0% 2.00%

Mx1+ Stag2 -/-
EV   73.7% 26.3%
Ebf1   98.8% 1.17%

cKit APC

K

■ Mx1- Stag2 f/f - EV   ■ Mx1+ Stag2 -/- - EV
■ Mx1- Stag2 f/f - Ebf1   ■ Mx1+ Stag2 -/- - Ebf1

Colonies/2x 10⁴ Cells

Plating   1 2 3 4 5 6 7 8

Fig 5. Altered Ebf1 chromatin structure results in a blockade of B cell development.

A) Enumeration of B lymphocyte stages of maturation in Stag2 WT and KO mice. Bone marrow was analyzed using flow cytometry for pro- (Cd43$^+$) and pre- (Cd43$^-$) B-cells in parent gate B220^{lo}IgM$^-$ showing B-cell development was impeded in the pro-B to pre-B cell stage (p=0.003). Immature (B220^{lo}IgM$^+$) and recirculating B cells (B220^{hi}IgM$^+$) were analyzed as a percentage of live singlets, which were both markedly reduced in Stag2 KO mice (p<0.001 for both populations). **B)** Methylcellulose colony assay using IL-7, SCF, and Flt3-L enriched media shows reduction in the number of B-cell colonies able to be formed by Stag2 KO bone marrow compared to WT (p=0.003). **C)** Enumeration of immature B cells (CD34+CD19+) and ratio compared to mature B cells (CD34-CD19+) show reduced immature B cells and immature:mature ratio in STAG2 mutated MDS patients (n=11) compared to controls (n=15). Stag2 WT and KO HSPC were infected with lentivirus containing GFP-tagged empty vector, GFP-mycPu.1, or GFP-shPu.1; and GFP$^+$ cells were plated in either B cell colony methylcellulose or stem cell methylcellulose. Cells plated in methylcellulose were harvested after 7 days and analyzed by flow cytometry for **D)** mature cell markers including the B marker B220 or **E)** stem cell markers including cKit. **F)** IGV track of the Ebf1 locus with Stag2 and Ctcf binding at 4 distinct sites which are lost in Stag2 KO and not bound by Stag1 either in WT or KO. **G)** ISC plotted across Ebf1 for Stag2 WT (n=2; shades of red) and KO (n=2; shades of blue) shows marked loss of insulation. **H)** Contact map of Ebf1 shows loss of cis-interaction at three loci (arrows). **I)** Lineage negative Stag2 WT and KO marrow were infected with retrovirus containing GFP-tagged empty vector or GFP-pCMV-Ebf1, and GFP$^+$ cells were plated in either B-cell colony methylcellulose or stem cell methylcellulose. As in C-D, cells were harvested after 7 days and analyzed by flow cytometry for **I)** mature cell markers including the B marker B220 or **J)** stem cell markers including cKit. **K)** Cells plated in stem cell methylcellulose were serially replated with exhaustion of Stag2 KO cell overexpressing Ebf1.