haematologica

Journal of The Ferrata Storti Foundation

Latexin regulation by HMGB2 is required for hematopoietic stem cell maintenance

by Cuiping Zhang, Yvonne N Fondufe-Mittendorf, Chi Wang, Jin Chen, Qiang Cheng, Daohong Zhou, Yi Zheng, Hartmut Geiger, and Ying Liang

Latexin regulation by HMGB2 is required for hematopoietic stem cell maintenance

Cuiping Zhang[1], Yvonne N Fondufe-Mittendorf [2], Chi Wang[3], Jin Chen[4], Qiang Cheng[4], Daohong Zhou[5], Yi Zheng[6], Hartmut Geiger[6,7], and Ying Liang[1]

[1]Departments of Toxicology and Cancer Biology, [2]Molecular & Cellular Biochemistry, [3]Cancer Biostatistics, [4]Internal Medicine and Computer Science, University of Kentucky, Lexington, KY 40536.
[5]Dept. of Pharmacodynamic, University of Florida, Gainesville, FL 32610.
[6]Cincinnati Children's Hospital Medical Center, Experimental Hematology and Cancer Biology, Cincinnati, OH 45229.
[7]Institute for Molecular Medicine, University of Ulm, 89081 Ulm, Germany.

Short title: Genetic and epigenetic regulation of latexin transcription

Author of correspondence:

Name: Ying Liang, MD/ PhD

Address: Department of Toxicology and Cancer Biology, Health Sciences Research Bldg. Rm. 340, University of Kentucky, 1095 V.A. Drive, Lexington, KY 40536-0305

Phone number: (859) 323-1729

E-mail: ying.liang@uky.edu

Keywords: latexin, HMGB2, single nucleotide polymorphism, transcription, hematopoietic stem cells

ABSTRACT

Hematopoietic stem cells provide life-long production of blood cells and undergo self-renewal division in order to sustain the stem cell pool. Homeostatic maintenance of hematopoietic stem cell pool and blood cell production is vital for organismal survival. We previously reported that latexin is a negative regulator of hematopoietic stem cells in mice, whose natural variation in the expression is inversely correlated with hematopoietic stem cell number. However, the molecular mechanisms regulating latexin transcription remain largely unknown, and the genetic factors contributing to its natural variation are not clearly defined. Here we discovered a chromatin protein, high-mobility group protein B2, as a novel transcriptional suppressor of latexin by using DNA pull-down and mass spectrometry. High-mobility group protein B2 knockdown increases latexin expression at transcript and protein levels, and decreases hematopoietic stem cell number and regeneration capacity in vivo. Concomitant blockage of latexin activation significantly reverses these phenotypic changes, suggesting that latexin is one of the downstream targets and functional mediators of high-mobility group protein B2. We further identified a functional single nucleotide polymorphism, rs31528793, in the latexin promoter that binds to high-mobility group protein B2 and affects the promoter activity. G allelic variation in rs31528793 associates with the higher latexin expression and lower hematopoietic stem cell number, whereas C allele indicates the lower latexin expression and higher stem cell number. This study, for the first time, reveal that latexin transcription is regulated by both trans-acting (high-mobility group protein B2) and cis-acting (single nucleotide polymorphism rs31528793) factors. It uncovers the functional role of naturally occurring genetic variants, in combination with epigenetic regulator, in determining differential gene expression and phenotypic diversity in hematopoietic stem cell population.

INTRODUCTION

Stem cells are key to the homeostatic maintenance of mature and functional cells in a variety of tissues and organs. They have the unique ability to perpetuate themselves through self-renewal and to replenish dying or damaged cells through multi-lineage differentiation. The balance between self-renewal and differentiation is critical for tissue homeostasis, and a disrupted balance could lead to serious problems such as tissue degeneration and development of cancer(1). Probably the best-studied adult stem cells are hematopoietic stem cells, which are responsible for life-long production of all hematopoietic lineages(2-4). The total number of HSCs is kept constant under steady-state conditions, but it also changes in response to stress or injury. The flexibility of stem cells to adapt to physiological needs is achieved by precise regulation of self-renewal and differentiation. Many molecules and signaling pathways have been found to be involved in this process(3). However, identification of the collection of genes contributing to critical stem cell functions is far from complete.

HSC number and function exhibit natural variation among humans as well as among different mouse strains(5-8). The natural variation is largely attributed to DNA variants in the genome that function as regulatory elements to control gene expression(9). The genetic diversity is a powerful but underused tool for unraveling the critical gene networks in stem cell regulation. Using genome-wide association studies, increasing numbers of gene regulatory variants have been strongly implicated in hematologic phenotypes and diseases in humans, such as fetal hemoglobin-associated genetic variants in patients with sickle cell disease (SCD) and β-thalassemia(10, 11). Recently, several reports have revealed an important yet previously unrecognized role of genetic variants in regulating epigenetics(9, 12-15). For example, DNA variations may affect the recruitment and binding affinity of transcription factors, which in turn lead to histone tail modifications. Such variations in the epigenetic environment result in significant variations in gene expression, which collectively manifest as a phenotypic trait. Despite these advances, the precise molecular mechanisms underlying the association between the genetic variants and hematopoietic phenotypes remain largely unknown.

Inbred mouse strains provide a model system for exploring the myriad of regulatory gene network contributing to hematologic diversity(10, 16). We carried out a comparative

study of two inbred strains, C57BL/6 (B6) and DBA/2 (D2), in which we documented large natural variation in a number of stem cell traits(5-8). One of the most significant traits is the natural size of the HSC population; that is, young B6 mice have 3- to 8-fold fewer stem cells in bone marrow than D2 mice, depending on the assay used for stem cell quantification. We further identified *Lxn* as the regulatory gene whose expression is negatively correlated with HSC number(17, 18). *Lxn* regulates HSCs in a cell-autonomous manner through concerted mechanisms of decreased self-renewal and increased apoptosis. Even though we identified several genetic variants that might be associated with the differential expression of *Lxn* in B6 and D2 stem cells, direct evidence is lacking about how these variants regulate *Lxn* transcription and whether they have any functional effects.

In this study, we, for the first time, reported that a chromatin protein, HMGB2, binds to *Lxn* promoter, and plays an important role in the transcriptional regulation of *Lxn*. Knockdown of HMGB2 increases *Lxn* expression at both transcript and protein levels, suggesting a suppressive role of HMGB2 in *Lxn* transcription. HMGB2 knockdown decreases the number of functional HSCs by promoting apoptosis and reducing proliferation. Concomitant knockdown of *Lxn* reverses these functional effects, suggesting that *Lxn* is one of the downstream targets of HMGB2. Moreover, we discovered that a functional polymorphism, SNP rs31528793, is associated with the differential expression of *Lxn* in different mouse strains, including B6 and D2. This study, for the first time, reveal the genetic and epigenetic regulation of *Lxn* transcription, suggesting that both trans- and cis-elements (HMGB2 and SNP, respectively) contribute to the differential gene expression and phenotypic diversity in HSC population.

METHODS

Luciferase reporter assay
Lxn promoter activity and HMGB2 transcription activity were measured by luciferase reporter assay with a Tropix TR717 luminometer using a dual luciferase assay kit.

Identification of *Lxn* promoter binding proteins
Lxn promoter binding proteins were isolated by µMACSTM FactorFinder Kit (Miltenyi Biotec Inc. Auburn, CA). The high purity double-strand DNA oligonucleotides containing SNP

rs31528793 was used as the DNA bait for protein pull-down. The associated proteins were determined by mass spectrometry at the Mass Spectrometry & Proteomics Facility at Ohio State University.

Protein-DNA binding assays

(1) Chromatin immunoprecipitation (ChIP): ChIP assay was performed on LK (Lin- c-KIT+) cells using ChIP assay kit (Sigma Aldrich, #CHP1) with HMGB2 polyclonal antibodies (ab67282), H2A.X antibody (ab11175), or Rabbit IgG control (ab171870) (Abcam, Cambridge, MA). HMGB2 binding affinity was determined by SYBR green quantitative PCR.

(2) Electrophoretic mobility shift assay (EMSA): EMSAs were performed in 293T cells transduced with HMGB2 lentivirus using the LightShift™ Chemiluminescent EMSA Kit (Thermo Scientific™).

Gene knockdown and expression measurement

EML or LSK cells were transduced with HMGB2 shRNA (MSH027321-LVRU6GP, GeneCopoeia), Lxn Mission shRNA (Sigma-Aldrich) virus. Gene expression was measured by real-time PCR with commercially available primer/probe mix for *Hmgb2* or *Lxn* in ABI PRISM 7700 (Applied Biosystems, Foster City, CA). Protein expression was measured by Western blot with anti-Hmgb2 antibody (ab67282), goat polyclonal anti-Lxn antibodies (ab59521, Abcam), or mouse monoclonal anti-β-actin antibody (A5441, Sigma).

Immunostaining and flow cytometry

(1) HSCs and HPCs: Young (8-12 week) female C57BL/6, DBA2, 129X1/SvJ, A/J and CD45.1 mice (Jackson Laboratories, Bar Harbor, ME) were used for HSCs/HPCs isolation. HSCs/HPCs were defined as Lin-, Sca-1+ (clone E13-161.7) and c-KIT+ (LSK) cells. Long-term HSCs (LT-HSCs) were identified as LSK plus CD34 and FLT3 negative cells. **(2) Cell cycle** was analyzed by BrdU incorporation using BrdU Flow Kit. **(3) Apoptosis** was evaluated by Annexin V staining. **(4) Active caspase 3 analysis** was analyzed using PE Active Caspase-3 Apoptosis Kit. All kits are from BD Pharmingen™. Flow cytometry was performed on a FacsAria II (Becton Dickinson) and the data were analyzed with FlowJo software (Tree star, Ashland, OR).

Functional analysis of HSCs and HPCs

(1) Colony forming cell (CFC) assay was performed in complete MethoCult media (Stem Cell Technologies, Vancouver, Canada), and colony was counted on day14. **(2) Cobblestone area forming cell (CAFC) assay** was performed as described previously. The most primitive HSCs showed cobblestones at day 35 of culture, and its frequency was calculated by using L-Calc Limiting Dilution Analysis Software (Stem Cell Technologies, Vancouver, Canada). **(3) In vivo transplantation assay,** $3x10^5$ transduced cells (GFP+ cells) plus $2x10^5$ competitor B6.SJL/BoyJ BM cells were injected into B6.SJL/BoyJ mice after 24 hours of transduction, and GFP+ chimerism in PB and BM was measured at 16 weeks post transplantation.

Statistical analysis

Data were examined for homogeneity of variances (F-test), then analyzed by student's t-test or One-way ANOVA using Tukey's test. Differences were considered significant at $P<0.05$. All statistical analyses were conducted with Graphpad Prism7.

All animal work and experiments were performed under the guideline of approved IRB and Ethics Committee, IBC, ACUC protocols at the University of Kentucky.

RESULTS

HMGB2 binds to *Lxn* promoter and suppresses its activity

The transcriptional regulation of the *Lxn* gene remains largely unknown. We used two criteria to identify the potential promoter in the upstream regulatory region of *Lxn*. First, we looked for the regions containing SNPs because the natural variation of *Lxn* expression is mainly caused by genetic variants. Secondly, we and others have shown that promoter hypermethylation is involved in the down-regulation of *Lxn* in several types of cancer cells, including leukemia stem cells[19-24]. This prompted us to search for regions enriched with CG dinucleotides (CpG island). We thus analyzed the mouse *Lxn* upstream putative promoter sequence (http:// www.methprimer.com) and identified a CG-enriched region that contains a SNP rs31528793 (**Fig.1A**). Using the NCBI SNP database (https://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs31528793) search, we confirmed the presence of this polymorphism. This region spans from the canonical 5' promoter (-333 nucleotide, nt) to the transcription start site (+1 nt), and extends through the first exon (+27

6

nt). We next amplified and sequenced this region, confirming the existence of this SNP (data not shown). To determine whether it has promoter activity, we performed *in vitro* luciferase reporter assay and found that this sequence in *Lxn* upstream regulatory region had a strong promoter activity (**Fig. 1B**).

We next performed in silico analyses to search for the potential transcription factors in the *Lxn* promoter region by using the transcription factor prediction program, TRANSFAC (www.cbrc.jp/research/db/TFSEARCH.html). Results from this analysis showed that SNP rs31528793 falls within the consensus binding motif for the transcription factors, *Adr-1* and *Ets-1*. Our previous microarray data showed that only *Ets-1* was expressed in HSCs (data not shown). We thus only evaluated the binding of *Ets-1* to *Lxn* promoter with chromatin immunoprecipitation (ChIP) and found that *Ets-1* did not bind to *Lxn* promoter (data not shown). These results prompted us to use the "DNA-pull down" and mass spectrometry to directly look for proteins that bind to this region. We found 15 candidate binding proteins (Table 1). HMGB2 and H2A.X are of particular interest because of the involvement of their family members in the regulation of stem cell function and differentiation as the chromatin modifiers(25-34). We thus focused on these two proteins and examined their role in *Lxn* transcriptional regulation.

We performed ChIP-qPCR assay to measure the binding and occupancy of HMGB2 and H2A.X to the *Lxn* promoter. The result showed that HMGB2, but not H2A.X, were significantly enriched at the *Lxn* promoter region in comparison to IgG control and to the control region at the downstream 500 bp_s of the promoter (**Fig.1C and Fig.1D**). This result supports the idea of the specific binding of HMGB2 to *Lxn* promoter. We next performed luciferase assay, and found that HMGB2 suppressed *Lxn* promoter activity (**Fig.1E**). Altogether, to our knowledge, this is the first time, HMGB2 has been identified as a novel transcription suppressor of *Lxn* gene.

HMGB2 knockdown increases *Lxn* expression and changes the function of a HSC cell line.

We next examined the regulatory role of HMGB2 in *Lxn* transcription and its effects on EML cells. EML is the only known hematopoietic cell line with both lympho and myelo-erythroid differentiation potential and is considered to represent HSCs(35-37). Knockdown

HMGB2 in EML cells significantly increased *Lxn* expression at both mRNA and protein level (**Fig. 2A and 2B** respectively), reinforcing the finding of transcriptional suppression of HMGB2 on *Lxn* expression. We previously reported *Lxn* as a negative regulator of HSC number(17), we thus hypothesized that knockdown of HMGB2 would decrease EML cell number via upregulation of *Lxn*. We therefore monitored the growth of EML cells with HMGB2 knockdown for 2 weeks, and found that HMGB2 knockdown led to a dramatic decrease in the cell number compared to control group (**Fig. 2C**). Since HMGB2 is a chromatin binding protein, its effect on EML number may not act solely through *Lxn* upregulation. We simultaneously knocked down *Lxn* in HMGB2 shRNA-transduced cells, and determined whether blocking *Lxn* up-regulation could attenuate HMGB2-induced growth inhibition. We confirmed that the expression of *Lxn* mRNA and protein in HMGB2-shRNA transduced EML cells was reduced by the co-transduction with *Lxn*-shRNA (**Fig. 2D and 2E**). **Figure 2F** showed that HMGB2 knockdown significantly decreased EML cell number, and the concomitant *Lxn* knockdown reversed this change, resulting in an increase of the cell number to the level comparable to control group. These data imply that *Lxn* is one of the downstream transcriptional targets of HMGB2, and HMGB2 suppresses *Lxn* transcription which in turn increases EML number.

We previously reported that *Lxn* negatively regulates HSC function through increasing apoptosis and decreasing proliferation(23, 38, 39), we hypothesized that HMGB2 inhibition may have similar effects on EML cells. Indeed, we found that knocking down HMGB2 significantly increased the percentage of apoptotic EML cells (**Fig. 3A**). The change was further confirmed by the increased proportion of active caspase 3 positive cells (**Fig. 3B**). Moreover, knocking down HMGB2 significantly decreased the percentage of cells in the S phase in the cell cycle (**Fig. 3C).** The concomitant increase in apoptosis and decrease in proliferation by HMGB2 inhibition may contribute to the decreased cell number (**Fig. 2B**). We next tested whether apoptosis and proliferation could be rescued by blocking *Lxn* upregulation in HMGB2 knockdown condition (in the same way as **Fig. 2E and 2F**). The results showed that *Lxn* knockdown on the top of HMGB2 knockdown decreased apoptosis (**Fig. 3D)** and the proportion of active caspase 3 positive cells (**Fig. 3E**), and restored their changes to the level of control group. However, the cell cycle changes were not fully restored (**Fig. 3F),** suggesting that *Lxn* may be a major player in regulating apoptosis, and other downstream targets of HMGB2 might be involved in cell cycle regulation that counteract *Lxn* function. Overall, these data suggest that HMGB2 positively regulates HSC

function via the suppression of *Lxn* expression.

HMGB2 knockdown increases *Lxn* expression in bone marrow HSCs and decreases their number and regeneration function

Because of the observed effect of HMGB2 on *Lxn* expression and EML function, we next asked whether HMGB2 plays a similar role in primary HSCs. We knocked down HMGB2 in bone marrow lineage- Sca-1+ c-Kit+ (LSK) cells (**Fig. 4A, left panel**), which are enriched with HSC and hematopoietic progenitor cells (HPCs), and then determined the effect of HMGB2 knockdown on *Lxn* expression and HSC and HPC cell numbers, apoptosis and cell cycling. We found that knockdown HMGB2 in LSK cells also led to a significant increase in *Lxn* expression at both transcript and protein levels (**Fig. 4A, middle and right panels**). We next performed in vitro short-term colony forming cell (CFC) and long-term cobblestone area forming cell (CAFC) assays to determine functional HPCs and HSCs respectively. The result showed that the numbers of HPCs and HSCs in HMGB2-knockdown cells were nearly 2-fold lower than those in control cells (**Fig. 4B and 4C**). Moreover, HMGB2 knockdown led to increase in apoptosis (**Fig. 4D)** and the proportion of active caspase 3 positive cells (**Fig. 4E),** and decrease in proliferation in LSK cells (**Fig. 4F**), similar to those seen in the EML cells. These data confirm that HMGB2 also regulates *Lxn* transcription in primary HSCs, and thereby affects the number and clonality of HSCs and HPCs.

We next performed the more stringent transplantation experiment to determine the effect of HMGB2 inhibition on HSC regeneration capacity *in vivo*. Donor cells are LSK cells that were transduced with either HMGB2 shRNA or control shRNA. They were next transplanted into the myeloablated recipient mice with helper cells, and blood and bone marrow regeneration were examined at 16 weeks post-transplantation (**Fig. 5A**). The results showed that HMGB2 knockdown resulted in significant decreases in the regeneration of blood cells, bone marrow HSC/HPC-enriched LSK cells, and the most primitive long-term HSCs with unlimited self-renewal capacity (**Fig. 5B-5D**). These results suggest that HMGB2 inhibition impairs HSC regeneration functionality. Altogether, our data obtained from the EML cell line and primary HSCs strongly support the idea that the HMGB2 suppresses *Lxn* expression, which in turn affects HSC and HPC number and function.

SNP rs31528793 affects *Lxn* promoter activity and HSC number

The level of a given mRNA transcript is controlled by trans-acting factors and/or cis-acting modulators. Our data suggest that HMGB2 might act as a trans-acting modulator to regulate *Lxn* transcription. Since we previously reported that several SNPs identified by us may contribute to the natural variation of *Lxn* expression(17), we next asked whether any of these SNPs is associated with *Lxn* expression as a cis-acting regulator. SNP rs31528793 is the only genetic variant in the *Lxn* promoter region (**Fig. 1A**), we thus asked whether it affects the *Lxn* promoter activity. We made a G to C mutation in the luciferase reporter construct containing the *Lxn* promoter sequence and performed the luciferase reporter assay. **Fig.6A** showed that the G to C change decreased the promoter activity by more than 2-fold, suggesting a potential suppressive role of this polymorphism in *Lxn* transcription (**Fig. 6A, left two columns**). Since HMGB2 binds to this region, we next examined whether G/C variant affects HMGB2 binding. The result showed that HMGB2 further suppresses *Lxn* promoter activity, and C allele still causes nearly 2-fold decrease of the promoter activity. We next performed the EMSA assay and further confirmed the interaction of HMGB2 with the *Lxn* promoter containing SNP rs31528793 (**Fig. 6B**). These results indicate that SNP rs31528793 influences *Lxn* promoter activity, with the G allele conferring a high activity, while the C allele is associated with a low activity. Therefore, the genetic variants of the *Lxn* promoter add another layer of regulatory mechanism of *Lxn* transcription.

We previously reported that Lxn is differentially expressed in HSCs of C57BL/6 (B6) and DBA2 (D2) mice, and its expression level is inversely correlated with HSC number(17). It is known that B6 mice carry G allele whereas D2 mice have C allele. We therefore hypothesized that the G allele is associated with the higher promoter activity, high *Lxn* expression and low HSC numbers, whereas the C allele has the opposite effect. Next, we examined *Lxn* expression and HSC numbers in B6, D2 and the other two mouse strains, 129X1/SvJ and A/J that carry G allele at the SNP rs31528793 position (http://www.informatics.jax.org/snp/rs31528793). We found that D2 mouse strain had the lowest expression of Lxn and highest HSC number, whereas all other three strains showed higher *Lxn* expression and lower HSC number **(Fig. 6C and 6D),** suggesting that G/C allelic variant could be indicative of *Lxn* expression level and HSC number variation. It is noted that *Lxn* expression level varies in strains carrying G allele suggesting that other SNPs outside of the *Lxn* promoter region may contribute to such variation.

DISCUSSION

Lxn plays an important role in regulating HSC function(17, 38) It was originally identified via the natural variation of HSC numbers between B6 and D2 inbred mouse strains in which B6 mice have fewer HSCs than D2 mice at a young age. The expression of *Lxn* is inversely correlated to the size of HSC population, that is, its level in B6 is higher than that in D2 cells. Therefore, *Lxn* is a negative regulator of HSC number and its mode of action is primarily through increasing HSC apoptosis and decreasing HSC regeneration capability and proliferation. However, nothing is known about how *Lxn* is transcriptionally regulated in HSCs and other stem cells and why it is differentially expressed in different inbred mouse strains.

Here, we identified a *Lxn* upstream regulatory sequence with a strong promoter activity. More importantly, the SNP (rs31528793) in this region significantly affects its promoter activity and the G allele carried in B6, 129X1/SvJ and A/J mouse strains confers the promoter a stronger activity than the C allele in D2 strain. Genetic variants have been recently identified to play an important role in transcriptional regulation and thereby resulting in gene expression and phenotypical variation(9, 13-15, 40). We thus proposed that the G/C containing promoter might be involved in *Lxn* transcriptional regulation. Using DNA pull-down and mass spectrometry, we, for the first time, identified a chromatin binding protein, HMGB2, as a novel transcriptional suppressor of *Lxn* expression. HMGB2 binding was validated by ChIP-qPCR assay in which the endogenous HMGB2 demonstrated a stronger affinity to the *Lxn* specific promoter sequence. HMGB2 knockdown increases *Lxn* expression and decreases HSC numbers in both HSC cell line and bone marrow-derived primary LSK cells. This effect was abrogated when the increased level of *Lxn* was blocked, indicating that *Lxn* is one of the downstream targets and functional mediators of HMGB2 in HSCs. Altogether, these results suggest that both cis- and trans-factors are involved in the regulation of *Lxn* transcription (**Fig.6E**). In trans-regulating mode, HMGB2 acts as a suppressor for *Lxn* transcription. In cis-regulating mode, G allele at SNP rs31528793 is associated with stronger promoter activity, a high level of *Lxn* expression, and a small size

of HSC pool. In contrast, the C allelic variant attenuates these effects and *Lxn* transcription is less responsive to HMGB2, which leads to a lower *Lxn* expression and an increased stem cell number. Therefore, our work not only identified HMGB2 as a novel transcription regulator of *Lxn*, but also provides a potential functional meaning of SNP rs31528793 in contributing to natural variations in *Lxn* expression and HSC number. Despite these findings, how H2MGB2 regulates *Lxn* transcription requires further investigation. We cannot exclude the possibility that HMGB2 directly regulates *Lxn* transcription as the transcription factor. But it is also likely that HMGB2 acts as a chromatin adaptor or modifier to recruit other transcription factors for the initiation of transcription (**Fig. 6E**). This mode of action was shown in the GFI1b transcription during erythroid differentiation process in which the binding of HMGB2 to GFI1b promoter enhances the binding of other factors, such as Oct-1, GATA-1 and NF-Y, which collectively activates Gfi1b transcription(32). In addition, the relationship of HMGB2 binding site to SNP rs31528793, and how they coordinately or independently regulate *Lxn* transcription requires further investigation. However, our current data more support the independent regulatory mechanism because of the following observations. First, as shown in **figure 6A**, the suppression extent of HMGB2 on G-containing promoter (ratio of "G+ HMGB2" to "G" is 0.47) is similar to that on C-containing promoter (ratio of "C+ HMGB2" to "C" is 0.48). These data suggest that suppression of HMGB2 on *Lxn* promoter activity is independent of allelic variant. Secondly, EMSA result in **figure 6B** also shows the similar intensity of shifted bands, suggesting that G/C variant does not affect HMGB2 binding. Lastly, to further confirm binding of HMGB2 to the *Lxn* promoter and determine the effect of SNP rs31528793 on HMGB2 binding *in vivo*, we performed ChIP-qPCR assay on bone marrow cells of C57BL/6 and DBA/2 mice which naturally carry the SNP. The binding affinity of HMGB2 was quantitatively measured by real-time PCR with primers spanning the promoter sequence containing G/C SNP (same as **Fig. 1C**). We did not detect any difference in the binding affinity, suggesting that G/C allelic variant does not cause differential binding of HMGB2 to *Lxn* promoter (data not shown). Altogether, these evidences strongly suggest that HMGB2 and SNP rs31528793 acts independently to regulate Lxn transcription.

HMGB2 is a member of the high mobility group family proteins. It is a non-histone chromatin-binding protein that remodel chromatin architecture, therefore affecting gene expression. HMGB2 has been shown to play an important role in maintaining stem cell population in a tissue-specific manner. For example, in the nervous system, HMGB2

deletion leads to the increased neural stem/progenitor cells by increasing their proliferation(41). However, in articular cartilage, loss of HMGB2 reduces the regeneration capacity of mesenchymal stem cells by increasing apoptosis(42). Similarly, knockdown of HMGB2 decreased the number of muscle stem (satellite) cells by inhibiting proliferation and stimulating differentiation, thereby leading to the impaired muscle regeneration(43). Our study showed that the functional effects of HMGB2 on HSCs and blood system are similar to those in mesenchymal and muscle stem cells. Knockdown of HMGB2 decreased HSC number and blood regeneration by increasing apoptosis and decreasing proliferation. These effects are mediated at least in part via the upregulation of *Lxn*, which is a negative regulator of HSC function. HMGB2 has also been shown to play an important role in cellular senescence and aging(44, 45). It binds to the chromosome loci of key senescence-associated secretory phenotype (SASP) genes and prevents their incorporation into transcriptionally repressive heterochromatin environment during senescence, thereby inducing SASP gene expression. Since we found that *Lxn* is one of the transcriptional targets of HMGB2, whether *Lxn* is involved in senescence and aging remains very interesting to be determined. Our unpublished data show that *Lxn* expression increases with aging, and old HSCs with *Lxn* depletion have the increased regeneration capacity that is comparable to young HSCs. Thus, inhibition of *Lxn* may rejuvenate old HSCs.

Natural genetic variation is associated with a variety of hematologic phenotypes in humans. Genome-wide association studies have revealed DNA variants that are implicated in hematologic traits such as fetal hemoglobin levels, hematocrit, cell counts and sizes of different types of blood cells, as well as in disease susceptibility(10). One of the best examples of the functional effect of genetic variation is a regulatory SNP that causes the blood disorder α-thalassemia. This SNP creates a new transcriptional promoter that interferes with normal transcription of α globin genes and leads to disease development(11). However, very few genes underlying the vast majority of these DNA variants have been uncovered and very little is known about how they contribute to the phenotypic diversity in the population(9). *Lxn* is the first stem cell regulatory gene reported that accounts for the natural diversity of HSC function(17). Here, we also for the first time discovered that SNP rs31528793 is one of the DNA variants that are associated with the differential expression of *Lxn* in mouse. The *Lxn* gene is evolutionarily conserved. Since it is identified by the genetic diversity that arises through natural selection, it may physiologically regulate a function in other natural populations, such as humans. In fact, our preliminary data have indicated that

a negative correlation between *Lxn* level and the number of hematopoietic stem and progenitor cells also exists in healthy humans (unpublished data). Therefore, *Lxn* may be involved in human hematopoiesis and there might be polymorphisms in human genome that are functionally similar to mouse SNP rs31528793. Interestingly, a recent report has shown that a SNP rs6441224, in *Lxn* promoter is associated with its expression level in human(46). So it would be very interesting to determine whether HMGB2 binds to this SNP-containing promoter region and regulates human *Lxn* and HSC function. They would become very useful genetic markers for screening of transplantation donors with a larger stem cell reservoir or for prediction of better recovery of cancer patients from the therapy-induced bone marrow and stem cell suppression.

Author Contributions

Cuiping Zhang performed majority of the experiments; Yvonne N Fondufe-Mittendorf helped with ChIP assay; Chi Wang provided support for statistical analysis; Jin Chen and Qiang Cheng provided bioinformatic analyses of the promoter sequences; Daohong Zhou, Yi Zheng and Hartmut Geiger provided insight to the work and manuscript; Ying Liang designed the experiments and wrote the manuscript.

Acknowledgements

Conflicts of interest

The authors have no conflicts of interest to declare.

REFERENCES

1.   Warr MR, Pietras EM, Passegue E. Mechanisms controlling hematopoietic stem cell functions during normal hematopoiesis and hematological malignancies. Wiley Interdiscip Rev Syst Biol Med. 2011;3(6):681-701.

2.   Ema H, Morita Y, Suda T. Heterogeneity and hierarchy of hematopoietic stem cells. Exp Hematol. 2014;42(2):74-82.

3.   Rossi L, Lin KK, Boles NC, et al. Less is more: unveiling the functional core of hematopoietic stem cells through knockout mice. Cell Stem Cell. 2012;11(3):302-317.

4.   Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. Cell. 2008;132(4):631-644.

5.   Henckaerts E, Langer JC, Snoeck HW. Quantitative genetic variation in the hematopoietic stem cell and progenitor cell compartment and in lifespan are closely linked at multiple loci in BXD recombinant inbred mice. Blood. 2004;104(2):374-379.

6.   de Haan G, Bystrykh LV, Weersing E, et al. A genetic and genomic analysis identifies a cluster of genes associated with hematopoietic cell turnover. Blood. 2002;100(6):2056-2062.

7.   Geiger H, True JM, de Haan G, et al. Age- and stage-specific regulation patterns in the hematopoietic stem cell hierarchy. Blood. 2001;98(10):2966-2972.

8.   de Haan G, Nijhof W, Van Zant G. Mouse strain-dependent changes in frequency and proliferation of hematopoietic stem cells during aging: correlation between lifespan and cycling activity. Blood. 1997;89(5):1543-1550.

9.   Furey TS, Sethupathy P. Genetics. Genetics driving epigenetics. Science. 2013;342(6159):705-706.

10.   Sankaran VG, Orkin SH. Genome-wide association studies of hematologic phenotypes: a window into human hematopoiesis. Curr Opin Genet Dev. 2013;23(3):339-344.

11.   De Gobbi M, Viprakasit V, Hughes JR, et al. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. Science. 2006;312(5777):1215-1217.

12.   McVicker G, van de Geijn B, Degner JF, et al. Identification of genetic variants that affect histone modifications in human cells. Science. 2013;342(6159):747-749.

13.   Kilpinen H, Waszak SM, Gschwind AR, et al. Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. Science. 2013;342(6159):744-747.

14.   Kasowski M, Kyriazopoulou-Panagiotopoulou S, Grubert F, et al. Extensive variation in chromatin states across humans. Science. 2013;342(6159):750-752.

15.   Heinz S, Romanoski CE, Benner C, et al. Effect of natural genetic variation on enhancer selection and function. Nature. 2013;503(7477):487-492.

16.   Van Zant G, Liang Y. Natural genetic diversity as a means to uncover stem cell regulatory pathways. Ann N Y Acad Sci. 2009;1176:170-177.

17.   Liang Y, Jansen M, Aronow B, et al. The quantitative trait gene latexin influences the size of the hematopoietic stem cell population in mice. Nat Genet. 2007;39(2):178-188.

18.   de Haan G. Latexin is a newly discovered regulator of hematopoietic stem cells. Nat Genet. 2007;39(2):141-142.

19.   Scadden DT. Nice Neighborhood: Emerging Concepts of the Stem Cell Niche. Cell. 2014;157(1):41-50.

20.   Muthusamy V, Premi S, Soper C, et al. The hematopoietic stem cell regulatory gene latexin has tumor-suppressive properties in malignant melanoma. J Invest Dermatol. 2013;133(7):1827-1833.

21.   Abd Elmageed ZY, Moroz K, Kandil E. Clinical significance of CD146 and latexin during different stages of thyroid cancer. Mol Cell Biochem. 2013;381(1-2):95-103.

22.   Mitsunaga K, Kikuchi J, Wada T, el al. Latexin regulates the abundance of multiple cellular proteins in hematopoietic stem cells. J Cell Physiol. 2012;227(3):1138-1147.

23.   Liu Y, Howard D, Rector K, et al. Latexin is down-regulated in hematopoietic malignancies and restoration of expression inhibits lymphoma growth. PLoS One. 2012;7(9):e44979.

24.   Li Y, Basang Z, Ding H, et al. Latexin expression is downregulated in human gastric carcinomas and exhibits tumor suppressor potential. BMC Cancer. 2011;11:121.

25.   Jin C, Zang C, Wei G, et al. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet. 2009;41(8):941-945.

26.   Creyghton MP, Markoulaki S, Levine SS, et al. H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. Cell. 2008;135(4):649-661.

27.   Andang M, Hjerling-Leffler J, Moliner A, et al. Histone H2AX-dependent GABA(A) receptor regulation of stem cell proliferation. Nature. 2008;451(7177):460-464.

28.   Lee MG, Villa R, Trojer P, et al. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science. 2007;318(5849):447-450.

29.   Gevry N, Chan HM, Laflamme L, et al. p21 transcription is regulated by differential localization of histone H2A.Z. Genes Dev. 2007;21(15):1869-1881.

30.   Pusterla T, de Marchis F, Palumbo R, et al. High mobility group B2 is secreted by myeloid cells and has mitogenic and chemoattractant activities similar to high mobility group B1. Autoimmunity. 2009;42(4):308-310.

31.   Pfannkuche K, Summer H, Li O, et al. The high mobility group protein HMGA2: a co-regulator of chromatin structure and pluripotency in stem cells? Stem Cell Rev. 2009;5(3):224-230.

32.   Laurent B, Randrianarison-Huetz V, Marechal V, et al. High-mobility group protein HMGB2 regulates human erythroid differentiation through trans-activation of GFI1B transcription. Blood. 2010;115(3):687-695.

33.   Nishino J, Kim I, Chada K, et al. Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. Cell. 2008;135(2):227-239.

34.   Yanai H, Ban T, Wang Z, et al. HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses. Nature. 2009;462(7269):99-103.

35.   Ye ZJ, Kluger Y, Lian Z, et al. Two types of precursor cells in a multipotential hematopoietic cell line. Proc Natl Acad Sci USA. 2005;102(51):18461-18466.

36.   Lemieux ME, Cheng Z, Zhou Q, et al. Inactivation of a single copy of Crebbp selectively alters pre-mRNA processing in mouse hematopoietic stem cells. PloS One. 2011;6(8):e24153.

37.   Kutlesa S, Zayas J, Valle A, et al. T-cell differentiation of multipotent hematopoietic cell line EML in the OP9-DL1 coculture system. Exp Hematol. 2009;37(8):909-923.

38.   Liu Y, Zhang C, Li Z, et al. Latexin Inactivation Enhances Survival and Long-Term Engraftment of Hematopoietic Stem Cells and Expands the Entire Hematopoietic System in Mice. Stem Cell Rep. 2017;8(4):991-1004.

39.   You Y, Wen R, Pathak R, et al. Latexin sensitizes leukemogenic cells to gamma-irradiation-induced cell-cycle arrest and cell death through Rps3 pathway. Cell Death Dis. 2014;5:e1493.

40.   Consortium GT, Laboratory DA. Coordinating Center -Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC, et al. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204-213.

41.   Abraham AB, Bronstein R, Reddy AS, et al. Aberrant neural stem cell proliferation and increased adult neurogenesis in mice lacking chromatin protein HMGB2. PLoS One. 2013;8(12):e84838.

42.   Taniguchi N, Carames B, Ronfani L, et al. Aging-related loss of the chromatin protein HMGB2 in articular cartilage is linked to reduced cellularity and osteoarthritis. Proc Natl Acad Sci U S A. 2009;106(4):1181-1186.

43.   Zhou X, Li M, Huang H, et al. HMGB2 regulates satellite-cell-mediated skeletal muscle regeneration through IGF2BP2. J Cell Sci. 2016;129(22):4305-4316.

44.   Aird KM, Iwasaki O, Kossenkov AV, et al. HMGB2 orchestrates the chromatin landscape of senescence-associated secretory phenotype gene loci. J Cell Biol. 2016;215(3):325-334.

45.   Nacarelli T, Liu P, Zhang R. Epigenetic Basis of Cellular Senescence and Its Implications in Aging. Genes (Basel). 2017;24;8(12).

46.   Kloth M, Goering W, Ribarska T, et al. The SNP rs6441224 influences transcriptional activity and prognostically relevant hypermethylation of RARRES1 in prostate cancer. Int J Cancer. 2012;131(6):E897-904.

TABLE1. *Lxn* promoter binging protein

Lxn promoter sequence containing SNP rs31528793
Histone H2A type 1- F (H2A1F)
Histone H2A type 2- A (H2A2A)
Protein S100-A9 (S10A9)
Protein S100-A8 (S10A8)
Histone H2A.X **(H2A.X)**
High mobility group protein B2 **(HMGB2)**
Histone H2B type 1- H (H2B1H)
Histone H2AV (H2AV)
Histone H1.3 (H1.3)
Peptidyl-prolyl cis-trans isomerase A (PPIA)
Eosinophil cationic protein 1 precursor (ECP1)
Myeloperoxidase precursor (PERM)
Histone H1.1 (H11)
Histone H1.5 (H15)
Coronin-1A (COF1)

Double-strand DNA oligonucleotides containing single nucleotide polymorphism (SNP rs31528793) were used as "bait" to capture associated proteins from bone marrow cell lysate of C57BL/6 mouse. Proteins binding to *Lxn* promoter sequences were isolated by µMACS™ FactorFinder Kit (Miltenyi Biotec Inc. Auburn, CA) and identified by Mass Spectrometry. Proteins with a Mascot score of 100 or higher with a minimum of two unique peptides from one protein having a –b or –y ion sequence tag of five residues or better were considered significant.

FIGURE LEGENDS

Figure 1. HMGB2 suppresses *Lxn* promoter activity.

(A) *Lxn* promoter sequence spans from 333 nucleotides upstream of the transcription start site (+1) of *Lxn* gene to 27 nucleotides into the first exon. The chromosomal positions for *Lxn* gene and SNP rs31528793 are indicated. **(B)** *Lxn* promoter sequence has strong promoter activity. Luciferase activity was determined in HEK cells transduced with luciferase reporter construct containing either *Lxn* promoter sequence (Lxn-PGL3) or control vector (PGL3). **(C)** HMGB2 specifically binds to *Lxn* promoter sequence. ChIP assay was performed with an HMGB2 polyclonal antibody (HMGB2) or IgG control (IgG). The genomic sequence in the 500 base pairs downstream of the *Lxn* promoter region were used as the negative sequence control to determine the HMGB2 binding specificity (Negative control). *Lxn* promoter sequence was amplified and quantified by real-time PCR (top panel). The fold enrichment of HMGB2 in the *Lxn* promoter was quantified by normalization to either IgG control (bottom left panel) or negative sequence control (bottom right panel). **(D)** H2A.X does not bind to *Lxn* promoter. ChIP assay was performed with an H2A.X polyclonal antibody (H2A.X) or IgG control (IgG). The genomic sequence in the 500 base pairs downstream of the *Lxn* promoter region were used as the negative sequence control to determine the H2A.X binding specificity (Negative control). *Lxn* promoter sequence was amplified and quantified by real-time PCR. **(E)** HMGB2 suppresses *Lxn* promoter activity. Luciferase activity was determined in HEK cells transduced with luciferase reporter construct containing either *Lxn* promoter sequence (Lxn-PGL3) or control vector with scramble sequence (Scramble-PGL3) without or with HMGB2 plasmid (Scramble-PGL3 + HMGB2, and Lxn-PGL3 + HMGB2). The data were the average of three independent experiments with triplicates in each experiment (n=9). *p<0.05; **p<0.01; ***p<0.001.

Figure 2. HMGB2 knockdown increases *Lxn* expression and decreases the number of HSC cell line.

(A) Knockdown of HMGB2 in EML cells increases *Lxn* mRNA expression. EML cells were infected by control lentivirus (Con) or HMGB2 knockdown shRNA (HMGB2 KD). HMGB2 and *Lxn* mRNA levels were measured by quantitative real-time PCR. GAPDH was the endogenous control for mRNA expression normalization. **(B)** Knockdown of HMGB2 in EML cells increases *Lxn* protein expression. EML cells were infected by control lentivirus (Con) or HMGB2 knockdown shRNA (HMGB2 KD). HMGB2 and *Lxn* protein levels were measured

by Western blot. Actin was the normalization control. Top panel is the representative Western blot out of three independent experiments. The bottom panel is quantification of intensity of HMGB2 (left) and Lxn (right) proteins. **(C)** HMGB2 knockdown decreases EML cell number. EML cells infected with empty (Con) or HMGB2 shRNA (HMGB2 KD) were cultured for 12 days and counted at different time-points. **(D)** *Lxn* mRNA was decreased in HMGB2-knockdown EML cells with simultaneous knockdown of *Lxn.* HMGB2-knockdown EML cells (HMGB2 KD) were co-transfected with *Lxn* shRNA lentiviral vector (HMGB2 KD + Lxn KD). *Lxn* mRNA and protein was measured by real-time PCR and Western blot. **(E)** *Lxn* protein was decreased in HMGB2-knockdown EML cells with simultaneous knockdown of *Lxn.* HMGB2-knockdown EML cells (HMGB2 KD) were co-transfected with *Lxn* shRNA lentiviral vector (HMGB2 KD + Lxn KD). HMGB2 and *Lxn* protein levels were measured by Western blot. Actin was the normalization control. Top panel is the representative Western blot out of three independent experiments. The bottom panel is quantification of intensity of HMGB2 (left) and Lxn (right) proteins. **(F)** *Lxn* knockdown restores the number of HMGB2-knockdown EML cells to a level comparable to control group. The data shown are EML cell number at day 3 of cell culture. All data were the average of three independent experiments with triplicates in each experiment (n=9). *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

Figure 3. HMGB2 knockdown increases apoptosis and decreases proliferation.
(A) HMGB2 knockdown increases apoptosis of EML cells. Representative FACS plots (left panel) and quantification (right panel) of Annexin V+ and 7AAD- apoptotic EML cells transduced with control-shRNA (Con) or HMGB2 knockdown shRNA lentivirus (HMGB2 KD). **(B)** HMGB2 knockdown increases the percentage of active caspase 3 positive EML cells. Representative flow cytometry profile (left panel) and quantification (right panel) of active caspase 3 immnuofluorescence signal in EML cells. **(C)** HMGB2 knockdown decreases proliferation of EML cells. Representative FACS plots showing the G0/G1 (BrdU- and 7AAD-), S (BrdU+), and G2/M (BrdU- and 7AAD+) phages of cell cycle in EML cells (left panel). The right panel shows the frequencies of each phase. **(D)** *Lxn* knockdown (HMGB2 KD + Lxn KD) restores the percentage of apoptotic (Annexin V+) HMGB2-knockdown EML cells (HMGB2 KD) to a level comparable to control group (Con). **(E)** *Lxn* knockdown (HMGB2 KD + Lxn KD) restores the percentage of active caspase-3 positive HMGB2-knockdown EML cells (HMGB2 KD) to a level comparable to control group (Con). **(F)** *Lxn* knockdown (HMGB2 KD + Lxn KD) did not restore the cell cycle status of HMGB2-knockdown EML cells (HMGB2 KD) to control group level (Con). Data presented as the

average ± SD of 6 measurements from 2 independent experiments. *p<0.05; **p<0.01; ***p<0.001.

Figure 4. HMGB2 knockdown increases *Lxn* expression and decreases the number and function of bone marrow HSCs.

(A) HMGB2 knockdown increases *Lxn* expression in hematopoietic stem and progenitor cells. HMGB2 knockdown (HMGB2 KD) in bone marrow LSK cells (left panel) increases *Lxn* mRNA (middle panel) and protein expression (right panel) compared to control group (Con). *Lxn* mRNA and protein levels were measured by quantitative real-time PCR and Western blot. GAPDH was the endogenous control for mRNA expression normalization. Actin was the control for protein normalization. **(B)** HMGB2 knockdown decreases the number of clonogenic and functional hematopoietic progenitor and **(C)** stem cells**.** LSK cells transduced with control and knockdown lentivirus were sorted, and the numbers of progenitor and stem cells were measured by colony forming cell (CFC) and cobblestone area forming cell assay (CAFC) respectively. **(D)** HMGB2 knockdown increases apoptosis (Annexin V+) of LSK cells. **(E)** HMGB2 knockdown increases the proportion of active caspase-3 positive LSK cells. The representative histogram of active capase-3 flow cytometry profile (top panel) and the quantification of positive cell proportion (bottom panel) are shown. **(F)** HMGB2 knockdown decreases proliferation of LSK cells. The apoptosis and proliferation were determined with the same way as in EML cells. Values are the mean ± SD from 3 independent experiments (*p<0.05; **p<0.01; ***p<0.001).

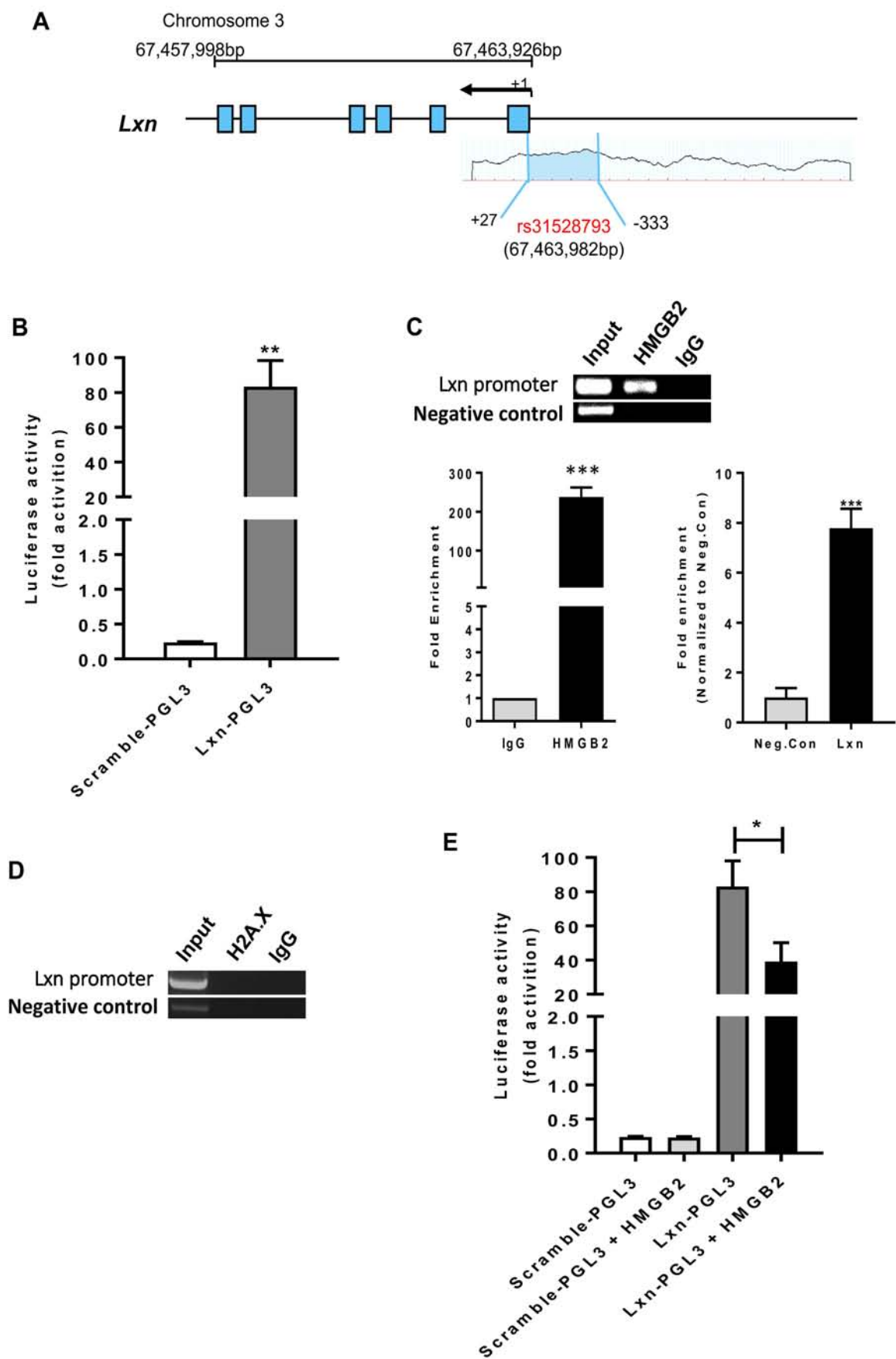**Figure 5. HMGB2 knockdown decreases HSC regeneration capacity.**

(A) Experimental scheme for competitive repopulation assay. Donor cells were 3×10^5 LSK cells transduced with HMGB2 shRNA (HMGB2 KD) or control vector (Con), and transplanted into myeloablated recipient mice along 2×10^5 competitor cells. Donor derived cells were determined by CD45.2 markers in the peripheral blood (PB), bone marrow (BM) LSK cells and long-term HSCs (LT-HSC) at 16 weeks post-transplantation. Long-term HSCs were determined by the markers lineage-Sca-1+c-kit+flk2-CD34-. **(B)** Frequencies of HMGB2 KD or control (CD45.2)-derived leukocytes, **(C)** BM LSK cells, and **(D)** LT-HSCs. Presented data are the average ± SD pooled from 2 independent experiments with 5 recipients in each group per experiment (n=10 per donor group). **p<0.01.
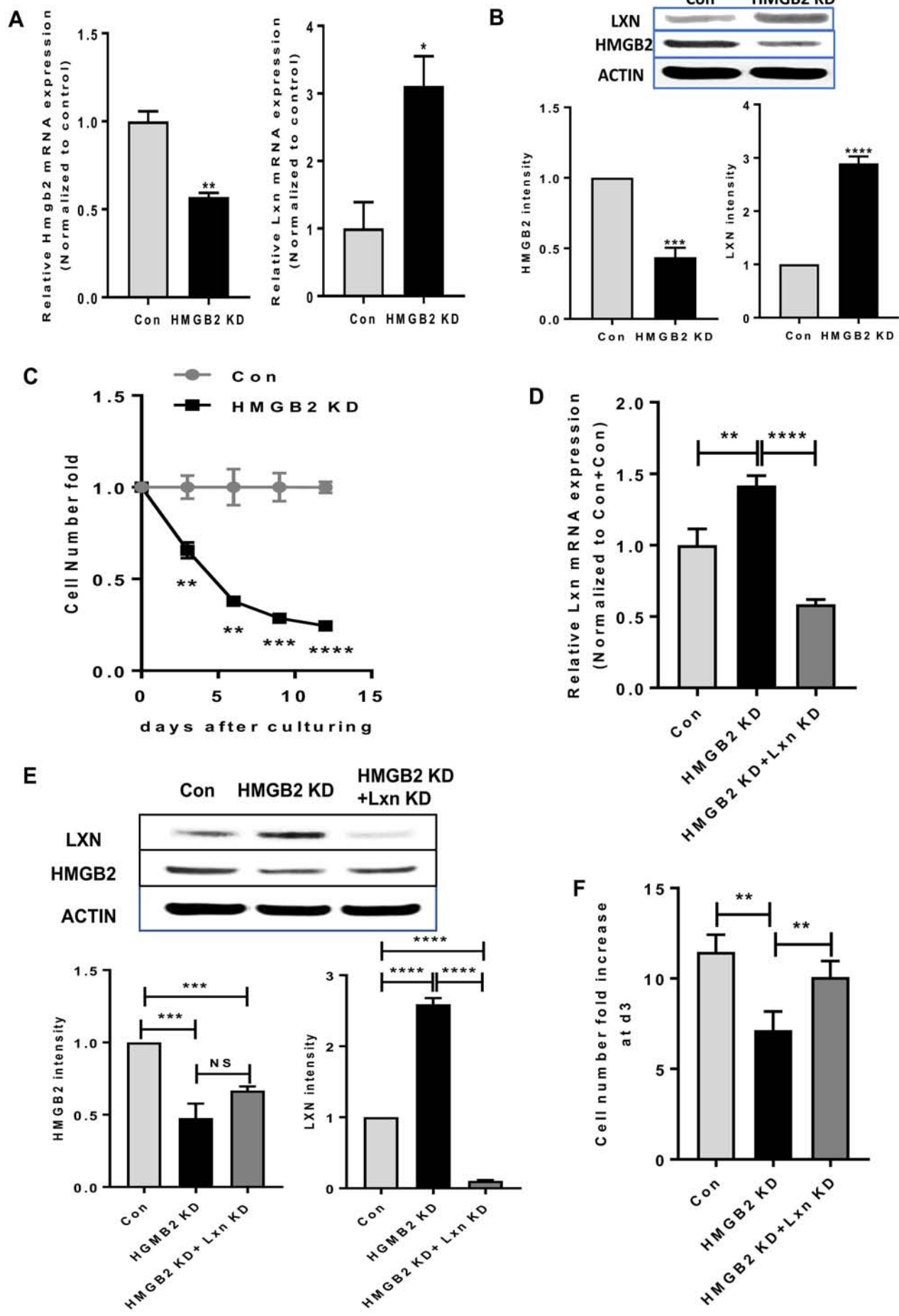
Figure 6. SNP rs31528793 associates with *Lxn* promoter activity and HSC number.

(A) SNP rs31528793 affects *Lxn* promoter activity. G allele is associated with the higher promoter activity whereas C allele lowers its activity. HMGB2 has transcription suppressor activity in both alleles. The luciferase reporter assay was performed on G or C containing vectors (G or C) and on the vectors co-transfected with HMGB2 plasmids (G+ HMGB2; or C+ HMGB2). The average values of three independent experiments are shown with standard deviation (SD) (n=12). **(B)** HMGB2 binds to G/C containing *Lxn* promoter. EMSA was performed with HMGB2 polyclonal antibody and biotin labelled oligonucleotides containing G or C allele. One out of three independent experiments was representatively shown. **(C)** G/C allele in SNP rs31528793 indicates LXN protein level. Western blot measured LXN protein in bone marrow cells of different mouse strains carrying either G or C allele. One out of three independent experiments was representatively shown. **(D)** G/C allele in SNP rs31528793 indicates HSC number. Long-term HSCs and frequency (%) were determined by the markers lineage-Sca-1+c-kit+flk2-CD34-. Results are the average (± 1 SD) of 9 mice derived from 3 independent experiments. (*p<0.05, **p<0.01). **(E)** Model of transcriptional regulation of *Lxn* by HMGB2 and the effect of SNP rs31528793 on the natural variation of *Lxn* expression and HSC number. HMGB2 binds to the *Lxn* promoter and acts as a transcriptional suppressor. SNP rs31528793 in the *Lxn* promoter sequence causes the differential promoter activity in which G allele is associated with higher *Lxn* expression and lower HSC number whereas C allele is associated with higher *Lxn* expression and lower HSC number. HMGB2 is involved in *Lxn* transcriptional regulation either by itself or by modifying chromatin structure in the *Lxn* promoter region, thereby facilitating access of other transcription factors to the region.
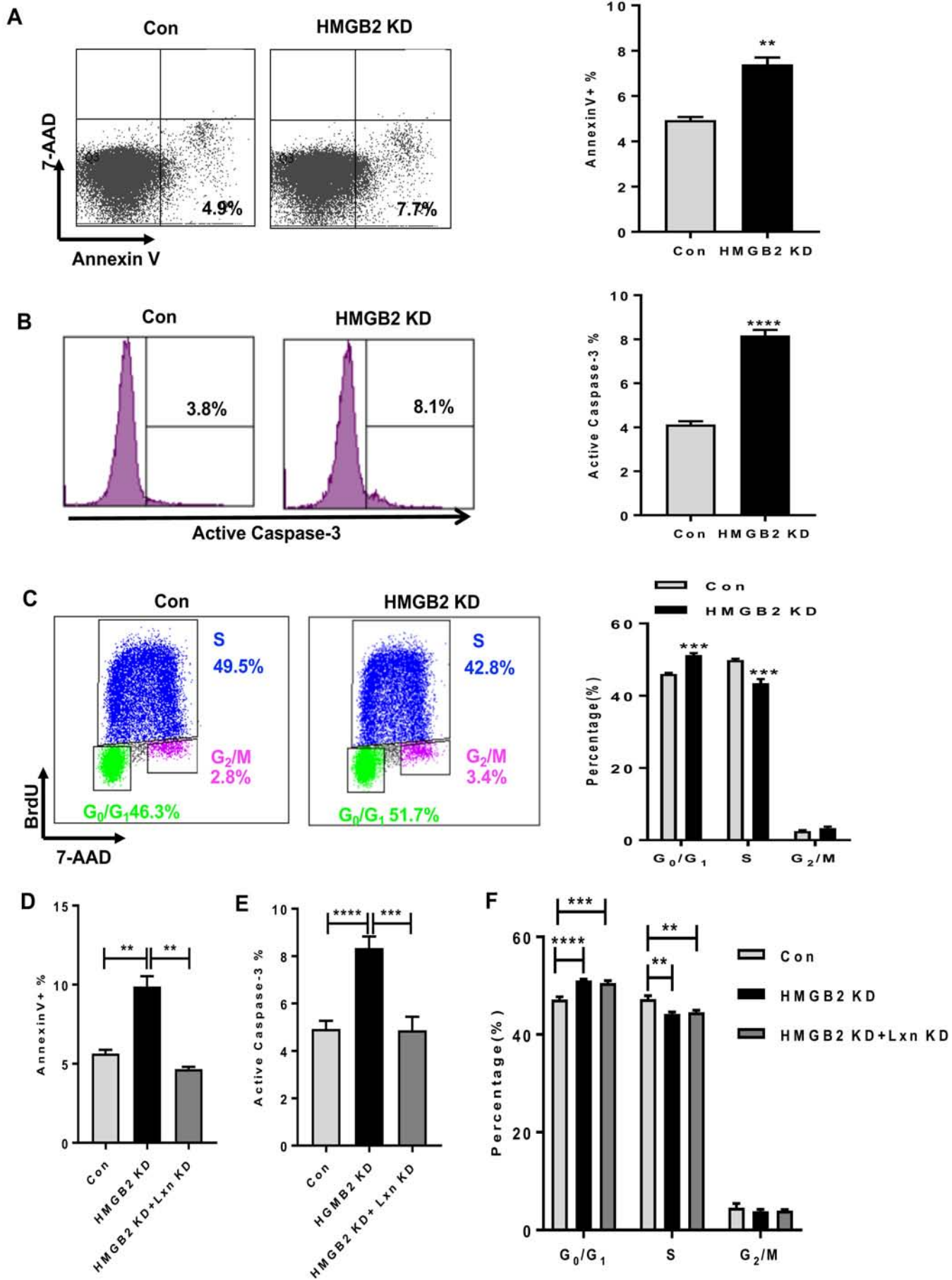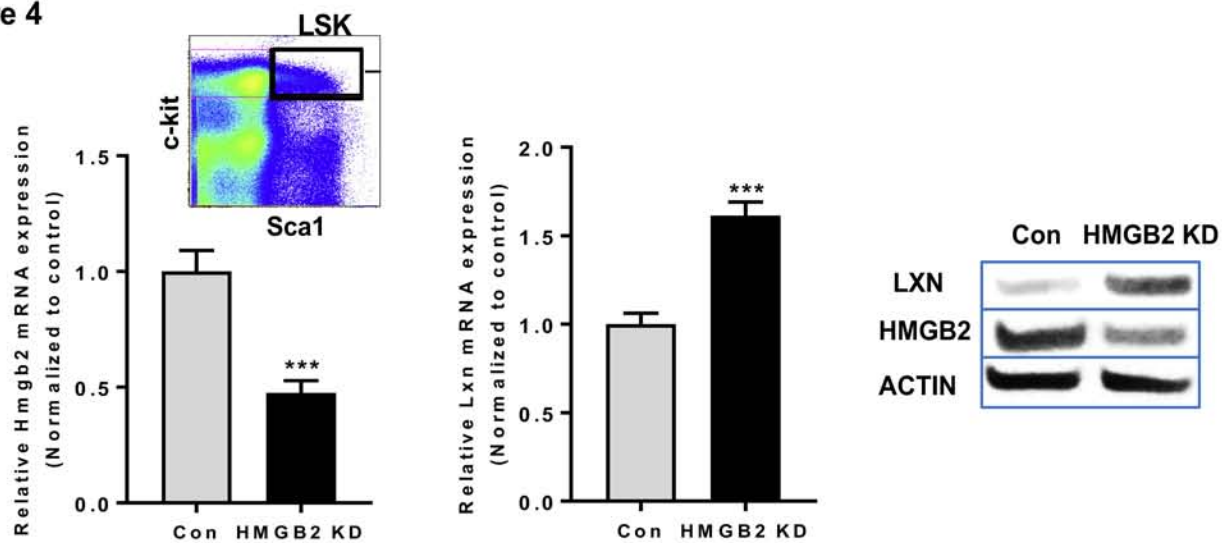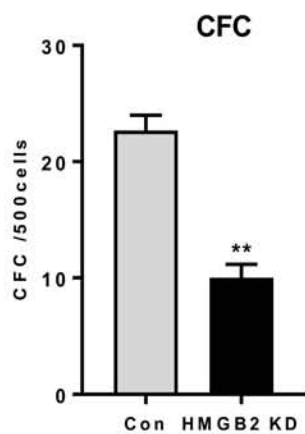
Figure 1

Figure 2

Figure 3
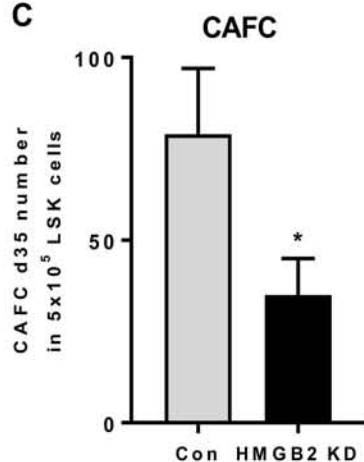
Figure 4

Figure 5

A

Donor
CD45.2
HMGB2 KD or WT
LSK cells (3x10^5)

Competitor
CD45.1
(2x10^5)

+

Recipients 9Gy
CD45.1

PB, BM LSK and LT-HSC chimerism

B

GFP % of whole PB

Con   HMGB2 KD

**

C

LSK GFP+% in BM

Con   HMGB2 KD

**

D

LT-HSC GFP+% in BM

Con   HMGB2 KD

**

Figure 6

A

Luciferase activity

G, C, G + HMGB2, C + HMGB2

B

HMGB2	−	+	+		−	+	+
Competitive DNA	−	−	+		−	−	+
Biotin-DNA	(G) +	+	+	(C) +	+	+

Shift →

Free probe

C

G · C · G

C57BL/6 · DBA/2 · 129X1/SvJ · A/J

LXN

ACT

D

LT-HSC X10^4 in BM

C57BL/6, DBA/2, 129X1/SvJ, A/J

E

HMGB2 · SNP · Lxn expression · HSC number

G → High → HSC
C → Low → HSC

HMGB2, SNP, Unidentified transcription factors, HSC
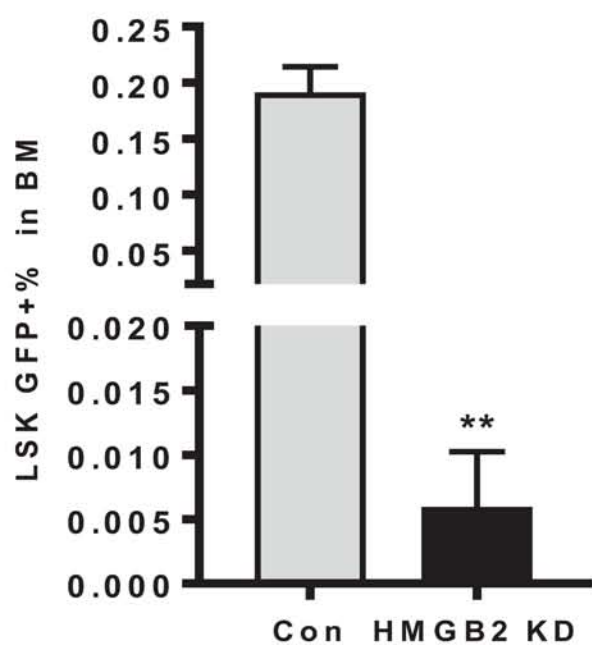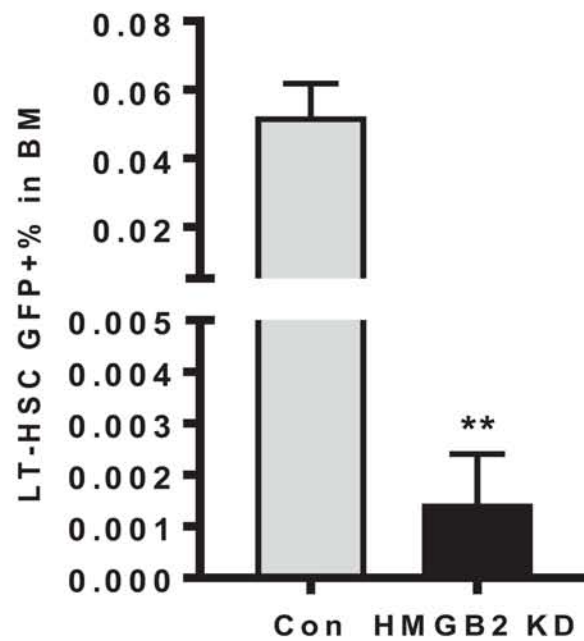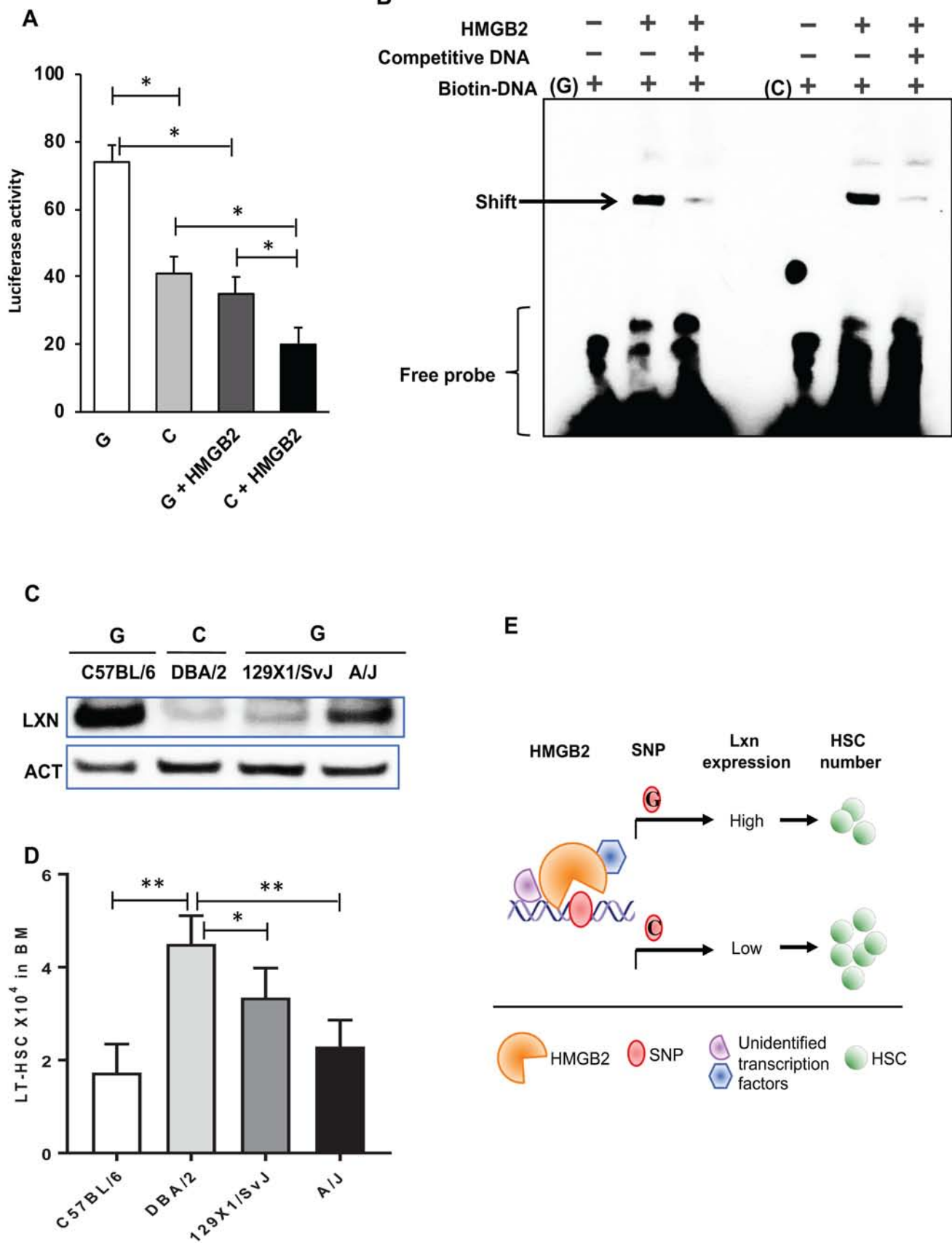
MATERIALS AND METHODS

Animals

Young 8- to 12-week old female C57BL/6, DBA2, 129X1/SvJ, A/J and CD45.1 mice were purchased from Jackson Laboratories (Bar Harbor, ME). Mice were kept in the animal facilities of the University of Kentucky under pathogen-free conditions according to NIH-mandated guidelines for animal welfare. They were fed with acidified water and food *ad libitum*.

Cell culture

293TA and NIH 3T3 cell lines were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS), 80 U/mL penicillin, and 80 mg/mL streptomycin. EML cells was cultured in Iscove's Modified Dulbecco Medium (IMDM, Stemcell technology) supplemented with 20% FBS (Stemcell technology) and 200 ng/mL murine stem cell factor (Peprotech). These cells were incubated in a humidified atmosphere of 5% CO_2 in air at 37°C. FBMD-1 cells were maintained in DMEM with 20% horse serum (HS), 0.1 mM 2-mecaptoethanol, 10 mg/mL hydrocortisone (HC, Sigma), 80 U/mL penicillin, and 80 mg/mL streptomycin, in a 33°C incubator with a humidified atmosphere of 5% CO_2. All cell lines used in this study were purchased from American Type Culture Collection.

Preparation and isolation of stem/progenitor-enriched hematopoietic cells

Bone marrow (BM) cells were harvested after crunching the bilateral iliac, femora and tibia and used for isolation of primitive HSCs. The Lin- cells was sorted using MACS mouse Lineage Cell Depletion Kit (Cat# 130-090-858) following the instructions. The Lin- cells were labelled with lineage antibodies including CD5 (clone 53–7.3), CD8a (clone 53–6.7), CD45R/B220 (clone RA3-6B2), CD11b/MAC-1 (clone M1/70), LY-6G/GR-1 (clone RB6-8C5), and TER119/Ly-76 (clone TER-119)-APC-Cy7, and stem cell markers, Sca-1-PE (clone E13-161.7) and c-KIT-APC (clone 2B8). The viable cells were distinguished by their ability to exclude 7-AAD. All monoclonal antibodies were purchased from BD Pharmingen (San Diego, CA). The LSK (Lin- Sca-1+ c-KIT+) cell sorting and flow cytometry analysis were performed on FACS LSRII (Becton Dickinson Immunocytometry Systems, San Jose, CA).

Generation of lentiviral vectors

The mouse *Lxn* promoter sequence was isolated from the genomic DNA of Lin- cells of C57/BL6 strains by PCR using the following primers: forward (from -341 to -327nt): 5'-TCGACTCGAGTCCAGGTTCTCCCTC-3', and reverse (from +13 to +27nt): 5'-

AGCCAAGCTTCAGCATAGTGGGTGG-3'. The PCR products were digested with XhoI and HindIII and cloned into the pGL3-basic vector (Promega, Madison, WI). The G in pGL-3-Lxn promoter vector was mutated to C using a QuickChangeII sited-directed mutagenesis kit (Stratagene, La Jolla, CA) following manufacturer's instructions. Hmgb2 shRNA plasmids were purchased from GeneCopoeia (Catlog# MSH027321-LVRU6GP). *Lxn* shRNA plasmids were purchased from Sigma-Aldrich (Lxn Mission shRNA). Crimson fluorescence marker replaced the puromycin resistance gene and was used for fluorescent cell sorting. Negative control was the scramble shRNA. Scramble sequence was amplified from mouse Lxn CDS zone using the following primers: forward 5' TCGACTCGAGTGGCGGAGAACTGCATT, and reverse 5' AGCCAAGCTTTAAGTGCTGGACTGGCTT , and digested with XhoI and HindIII and cloned into the PGL3-basic vector. HshRNA and LshRNA plasmids are both lentivirus in this study. The lentiviral vectors were packaged using 293TA cells by co-transfection with Δ8.9 and VSVG plasmids through calcium-phosphate precipitation. After culturing the respective cell lines for 48 h and 72 h, virus supernatant was collected and viral titers were determined.

Luciferase reporter assay

Human embryonic kidney (HEK) cells were transfected with 290ng of luciferase reporter constructs including pGL-3-*Lxn* promoter vector (G Lxn-PGL3) and the vector containing SNP rs31528793 (C Lxn-PGL3), together with 10ng of pRL-null (thymidine kinase promoter-Renilla luciferase) plasmid using FuGene® 6 (Roche, Indianapolis, IN). To test HMGB2 and H2A.X activity, cells were contransfected with pcDNA-HMGB2 and pcDNA-H2A.X plasmids (10ng, 50ng, 100ng, 200ng or 500ng). Cell lysate was made 48 h later and the luciferase activities were measured with a Tropix TR717 luminometer using a dual luciferase assay kit and were normalized to Renilla luciferase values. Except as indicated, all vectors and assay kit were purchased from Promega (Madison MI). All assays were performed 2 to 3 times with 4 to 6 replicates in each time.

Microbeads-magnetic extraction of *Lxn* promoter binding proteins

Lxn promoter binding proteins were isolated by µMACSTM FactorFinder Kit (Miltenyi Biotec Inc. Auburn, CA). In brief, a high purity double-strand DNA oligonucleotides containing SNP rs31528793 (5' GGTAGGCGGGCACCTCCCG(C)GAGGAAAGCTC 3') was modified with biotin and incubated with the bone marrow cell lysate. µMACSTM streptavidin microbeads, which bind to biotinylated DNA with extremely high affinity, were added to capture the DNA (oligoes)-protein complex. The associated proteins were separated by MACS® µ Columns in

which MicroBeads-attached proteins were retained on the strong magnetic field of column. The associated proteins were eluted and analyzed by gel electrophoresis and visualized by silver staining (SilverSNAP Stain Kit II, PIERCE, Rockford, IL).

Identification of binding proteins by mass spectrometry

To identify components of eluted proteins, mass spectrometry was performed on the samples obtained above at the Mass Spectrometry & Proteomics Facility at Ohio State University. Briefly, proteins were digested by trypsin or chymotrypsin. Capillary-liquid chromatography-nanospray tandem mass spectrometry (Nano-LC/MS/MS) was performed on a Thermo Finnigan LTQ mass spectrometer equipped with a nanospray source operated in positive ion mode. The LC system is an UltiMateTM Plus system from LC-Packings A Dionex CO (Sunnyvale, CA) with a Famous autosampler and Switchos column switcher. 5 ul of each sample was first injected on to the trapping column (LC-Packings A Dionex Co, Sunnyvale, CA), and washed with 50 mM acetic acid. The peptides were eluted off the trap onto the column and then into the LTQ system using a gradient of 2-80% acetonitrile over 50 minutes, with a flow rate of 300 nl/min. The MS/MS was acquired by using a nanospray source operated with a spray voltage of 3 KV and a capillary temperature of 200°C. The analysis was programmed for a full scan recorder between 350-2000 Da, and a MS/MS scan to generate product ion spectra to determine amino acid sequence in consecutive instrument scans of the ten most abundant peaks in the spectrum. Sequence information from the MS/MS data was processed by converting the raw data files into a merged file. The resulting mgf files were searched using Mascot Daemon by Matrix Science Version 2.2.1 (Boston, MA) and the database searched against the full SwissProt database version 54.1 (283454 sequences; 104030551 residues). Proteins with a Mascot score of 100 or higher with a minimum of two unique peptides from one protein having a –b or –y ion sequence tag of five residues or better were accepted.

Chromatin immunoprecipitation (ChIP) assay

ChIP assays kit (Sigma Aldrich, #CHP1) was used for this purpose. FACS sorted LK (Lin- c-KIT+) cells cross-linked with 1% formaldehyde were lysed and sonicated. The protein-DNA complex was immunoprecipitated with HMGB2 polyclonal antibodies (Abcam, Catalog# ab67282) or H2A.X antibody (Abcam, Catalog# ab11175). PCR specific for amplifying *Lxn* promoter region or for 500 base pairs downstream of the promoter region was performed on the DNA recovered after reverse cross-linking. The primers for amplifying the *Lxn* promoter

sequence are 5' GGAATGCCCGAGCTTTTCT 3' (forward) and 5' TCCTTCCCTCTTCCTCCTTCA 3' (reverse). The primers for control downstream 500 base pairs are 5' GAGGCAGAGGACTTGAGTTTGGT 3' (forward) and 5' GCTTAATGAGCGTGAACCACC 3' (reverse). Immunoprecipitation using HMGB2 or H2A.X antibody related normal Rabbit IgG (Abcam, Catalog# ab171870) and amplification using the primer for downstream 500 base pairs serve as negative control. SYBR green was used for the quantitative real-time PCR.

Infection of EML cell line and measurement of growth of virally-transduced cells

EML cells were transduced with HMGB2 shRNA or control (Con) viral supernatant supplemented with 200 ng/mL murine stem cell factor (Peprotech) and 8 μg/ml of polybrene for 6 hours. Fresh culture medium replaced viral supernatant and the cells were recovered for another 42 hours. The infected (GFP+) cells were sorted and cultured for 12 days. GFP+ EML cells were counted on a hemacytometer using trypan blue dye exclusion and 100,000 cells were seeded into 24-well non-tissue culture plates (Costar). The cell numbers were subsequently measured at different time-points for 12 days. At each time point, cells were split and maintained at a concentration of 100,000 cells per well. The cumulative cell number was calculated from the cell counts and the dilutions made at each culture split. *Lxn* knockdown vector transduction, Crimson+ cell sorting and cell growth measurement were performed similarly as described above.

Quantitative real-time PCR

To measure the expression of *Hmgb2* and *Lxn* mRNA in virally transduced cells, quantitative real-time PCR was performed. Total RNA was extracted from EML cells or primary bone marrow cells using RNeasy Mini kit (QIAGEN) according to the manufacture's protocol. cDNA was synthesized from total RNA by MultiScribe reverse transcriptase (Applied Biosystems) following the manufacturer's instruction with random primers (Applied Biosystems). In real-time PCR reactions, primer and probe mixes for mice *Hmgb2* and *Lxn* were purchased from Applied Biosystems. TaqMan rodent glyceraldehyde-3-phosphate dehydrogenase (Gapdh) served as an endogenous control for normalization of *Hmgb2* and *Lxn* expression. The PCR amplification was carried out on an ABI PRISM 7500™ real-time PCR machine using the TaqMan reagent standard curve method.

Western blot

Total protein extraction was performed as previously reported[17]. For Western blot, protein lysates were thawed and mixed with running buffer and a reducing agent (Novex, LifeTechnologies) according to the manufacturer's instructions and heated at 95°C for 5 minutes. Samples were then separated on a denaturing 10% bis-Tris PAGE gel (Novex) using the equivalent of 1×10^5 cells per lane at 200 V for 35mins. Following electrophoresis, samples were then transferred onto a PVDF membrane (Millipore) by electro-blotting at 30 V for 1 h, which were subsequently blocked in 5% skimmed milk and probed with rabbit monoclonal anti-Hmgb2 antibody (Abcam, Catalog# ab67282), goat polyclonal anti-Lxn antibodies (Abcam, Catalog# ab59521), and mouse monoclonal anti-β-actin antibody (Sigma, Catalog# A5441), respectively. Primary antibodies were detected using HRP-linked secondary antibodies (Cell Signaling Technology) and Chemiluminescen reagent (Pharmacia Biotech) according to the manufacturer's instructions.

Infection of primary bone marrow cells

Flow cytometry sorted LSK cells from C57/BL6 mice were stimulated with cytokines including 100 ng/mL FMS-like tyrosine kinase-3 ligand, 50 ng/mL mouse stem cell factor, 20 ng/mL interleukin-3 (IL-3), and 20 ng/mL TPO in StemSpan SFEM (STEMCELL Technologies). After 24 hr, the cells were transduced with lentiviral particles encoding either HMGB2 shRNA, or its related scramble control vector, at an MOI of 100 along with 8 μg/ml of polybrene for 6 hr. After 48 hr, the GFP-positive cells were sorted for real-time PCR, western blotting, CAFC assay. In transplantation assay, 3×10^5 transduced cells (GFP+ cells for both groups) plus 2×10^5 competitor B6.SJL/BoyJ BM cells were injected into B6.SJL/BoyJ mice after 24 hurs of transduction, and GFP+ chimerism in PB and BM was measured at 16 weeks post transplantation.

Colony Forming Cell (CFC) assay

CFC assay was performed in MethoCult™ GF M3434 medium (STEMCELL Technologies, Catalog # 03434) according to the manufacturer's protocol. Briefly, the sorted GFP+ bone marrow cells infected with HMGB2 shRNA lentivirus or control virus were seeded into 35 mm culture dishes, and colony formation was observed on days 12.

Cobblestone area forming cell (CAFC) assay

The CAFC assay was carried out as described previously[17]. Briefly, a confluent monolayer of FBMD-1 stromal cells was established in 96-well tissue culture plates (Costar). The sorted

GFP+ bone marrow cells infected with HMGB2 shRNA lentivirus or control virus were seeded on FBMD-1 containing plates at (3-fold) decreasing numbers. Twenty replicate wells were evaluated for each cell number. Individual wells were screened at day 35 for the presence of a cobblestone area, defined as a colony of at least five small, non-refractile cells growing underneath the stromal layer. The most primitive HSCs show cobblestones at day 35. Frequencies of CAFCs were calculated by using L-CaLc Limiting Dilution Analysis Software (Stem Cell Technology).

Cell cycle and apoptosis analysis

Cell cycle was analyzed by BrdU incorporation on sorted transduced GFP+ cells (after HMGB2 shRNA and control vector lentivirus transduction) or GFP+Crimson+ cells (after HMGB2 shRNA and Lxn shRNA with related control vector lentivirus transduction following the protocol provided by BD Pharmingen™ BrdU Flow Kit. Apoptosis was evaluated by flow cytometric determination of Annexin V (BD Pharmingen™) binding on transduced GFP+ EML cells or GFP+ primary LSK cells, or transduced GFP+ Crimson+ EML cells. For live cell isolation purpose, we used 7-aminoactinomycin D(Invitrogen) for dead cell exclusion. For active caspase 3 analysis, sorted transduced GFP+ EML or LSK cells (after HMGB2 shRNA and control vector lentivirus transduction) was analyzed using PE Active Caspase-3 Apoptosis Kit (BD Pharmingen™) following the instructions from the company.

Electrophoretic mobility shift assay

EMSAs were performed using the LightShift™ Chemiluminescent EMSA Kit (Thermo ScientificTM). Briefly, Hmgb2 lentiviral plasmid was transfected into 293T cells using calcium phosphate method. After 48 hs, 293T cells nuclear extracts were prepared using NE-PER Nuclear and Cytoplasmic Extraction Reagents (ThermoFisher Scientific) according to the manufacturer's protocol. EMSA binding reactions included 1× binding buffer, 1μg sperm DNA, 3 μl nuclear extract, and 20 uM biotin-labelled probes. One binding reaction without nuclear extract was as negative control. For competition assays, 4pM unlabeled oligonucleotides were allowed to bind the nuclear extracts (10 min at room temperature) before the addition of labeled probes. Binding reactions were incubated for 20 min at room temperature, size-separated on a 6% DNA retardation gel (ThermoFisher Scientific) in 0.5X TBE buffer, and transferred to a Biodyne B Nylon membrane (ThermoFisher Scientific) in 0.5xTBE buffer at 380mA for 45 minutes at 4 °C, and then crosslinked to the membrane using a TL-2000 UV Translinker (Ultra-Violet Products).Free or protein-bound biotin-labelled probes were detected

using streptavidin–horseradish peroxidase conjugates and chemiluminescent substrate according to the manufacturer's instructions. Probe sequences for promoter regions are listed below. (1) G 5' biotin-GGTAGGCGGGCACCTCCCGGAGGAAAGCTC (2) C 5' biotin - GGTAGGCGGGCACCTCCCCGAGGAAAGCTC

Statistical analysis

Data were examined for homogeneity of variances (F-test), then analyzed by student's t-test or One-way ANOVA using Tukey's test. Differences were considered significant at $P < 0.05$. All statistical analyses were conducted with Graphpad Prism7.