Hmga2 is a direct target gene of RUNX1 and regulates expansion of myeloid progenitors in mice

Kentson Lam¹, Alexander Muselman¹, Randal Du¹, Yuka Harada²†, Amanda G. Scholl¹,³, Ming Yan¹, Shinobu Matsuura¹*, Stephanie Weng¹, Hironori Harada²†, Dong-Er Zhang¹,³

¹Moores UCSD Cancer Center, University of California, San Diego, La Jolla, CA, USA
²Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima, Japan
³Departments of Pathology and Division of Biological Sciences, University of California, San Diego, La Jolla, CA, USA

Current address:
†Department of Hematology, Juntendo University School of Medicine, Tokyo, Japan
*Department of Medicine, Boston University School of Medicine, Boston, MA, USA

Running title: Hmga2 regulates expansion of myeloid progenitors

Corresponding author:
Dong-Er Zhang, Ph.D.
Moores UCSD Cancer Center
University of California, San Diego
3855 Health Sciences Drive
La Jolla, CA 92093
Tel: (858) 822-5372
Fax: (858) 822-5433
e-mail: d7zhang@ucsd.edu
Key Points

- Loss of RUNX1 by using genetic knockout or dominant-negative approaches leads to up-regulation of its direct target gene *Hmga2* in HSPCs.
- Expansion of myeloid progenitors caused by the loss of RUNX1 is rescued by loss of *Hmga2*, suggesting that RUNX1 functions through *Hmga2*. 
Abstract

RUNX1 is a master transcription factor in hematopoiesis, and mediates the specification and homeostasis of hematopoietic stem and progenitor cells (HSPCs). Disruptions in RUNX1 are well known to lead to hematological disease. In this study, we sought to identify and characterize RUNX1 target genes in HSPCs by performing RUNX1 chromatin immunoprecipitation with high-throughput sequencing (ChIP-seq) using a murine HSPC line and complementing this data with our previously described gene expression profiling of primary wildtype and RUNX1-deficient HSPCs (Lineage-/cKit+/Sca1+). From this analysis, we identified and confirmed that Hmga2, a known oncogene, as a direct target of RUNX1. Hmga2 was strongly up-regulated in RUNX1-deficient HSPCs, and the promoter of Hmga2 was responsive in a cell-type dependent manner upon co-expression of RUNX1. Conditional Runx1 knockout mice exhibit expansion of their HSPCs and myeloid progenitors as hallmark phenotypes. To further validate and establish that Hmga2 plays a role in inducing HSPC expansion, we generated mouse models of HMGA2 and RUNX1 deficiency. Although mice lacking both factors continued to display higher frequencies of HSPCs, the expansion of myeloid progenitors was effectively rescued. The data presented here establish Hmga2 as a transcriptional target of RUNX1 and a critical regulator of myeloid progenitor expansion.
Introduction

RUNX1, also known as AML1 and CBFα2, is a member of the Runt family of proteins. The major function of RUNX1 is to operate as a DNA binding transcription factor. Studies over the past 20 years have established RUNX1 as a critical player in hematopoiesis and specification of hematopoietic stem and progenitor cells (HSPCs), for neither process can occur without RUNX11–3. The importance of RUNX1 is further validated by its prevalence in a variety of hematological diseases and malignancies including myelodysplastic syndrome (MDS), myeloproliferative neoplasms (MPN), and multiple forms for acute myeloid leukemia (AML)4–6. As a transcription factor, RUNX1 binds to DNA regulatory regions in order to guide the expression of its target genes. Most confirmed RUNX1 target genes are mainly expressed in differentiated blood cells5,7–9. Direct targets of RUNX1 in HSPCs, however, have largely remained unexplored. The identification of these genes offers an insightful view into how this master regulator influences HSPC biology. To elucidate RUNX1 target genes in HSPCs, we performed genome-occupancy analysis with chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) using RUNX1 antibodies and a murine HSPC cell line. This data was combined with gene expression signatures from wildtype and RUNX1-deficient HSPCs (Lineage-/cKit+/Sca1+)10. The combination of both sets of data allowed elucidation of Hmga2 as a RUNX1 target gene.

High mobility group AT-hook 2 (HMGA2) is a non-histone chromatin binding protein typically associated with enhancers but lacks its own transcriptional activity11. Its expression is generally highest in stem cells and during embryogenesis12. Aberrant HMGA2 expression has been associated with a variety of mesenchymal tumors including several examples in the hematopoietic system13,14. Hence, HMGA2 is considered to be an oncogene. More recently,
ectopic expression of Hmga2 in transgenic mouse models or bone marrow transplantation with cells expressing Hmga2 have been demonstrated to result in myeloproliferative disease\textsuperscript{15,16}. In these models, the HSPC populations are expanded as well as the myeloid progenitors. All of these studies suggest that HMGA2 serves as a critical regulator of proliferation and survival. Conditional deletion of Runx1 in mice has allowed for the study of its role in HSPCs since mice that are null for Runx1 display embryonic lethality\textsuperscript{1,2,17,18}. Interestingly, one of the most striking phenotypes in Runx1 conditional knockout mice is a marked expansion of their HSPCs and myeloid progenitors\textsuperscript{18,19}. In this study, we identified Hmga2 as a target gene using RUNX1 ChIP-seq and analyzed the role of RUNX1 in regulating Hmga2 expression. Furthermore, using genetic models of RUNX1 and HMGA2 deficiency, we establish that Hmga2 is a crucial regulator of myeloid progenitor expansion.
Methods

Mice

All experimental protocols were approved by the UCSD Institutional Animal Care and Use Committee and all animals were housed at UCSD. Runx1<sup>flox</sup> and Mx1-Cre mice were kindly provided by Dr. Nancy Speck<sup>18</sup> and Hmga2<sup>−/−</sup> mice were kindly provided by Dr. Kiran Chada<sup>20</sup>.

Chromatin immunoprecipitation (ChIP) and sequencing

ChIP procedure and antibody was previously described<sup>21</sup>. Approximately 5x10<sup>7</sup> EML cells were used to perform two replicates. High-throughput sequencing was conducted on a Genome Analyzer II (Illumina, San Diego, USA) resulting in 2.7 x 10<sup>7</sup> and 2.6 x 10<sup>7</sup> reads. Sequences were aligned to the reference mouse genome (version mm9). Peak calling was performed using the MACS algorithm<sup>22</sup>. Peaks with a false discovery rate of less than 1% were associated with the closest transcriptional start site using PeakAnalyzer<sup>23</sup>. De novo motif finding was performed using the MEME software suite (<http://meme.nbcr.net>)<sup>24</sup>. Primers are provided in Table S1.

RT-qPCR

RNA was extracted using the TRIzol method (Life Technologies, Carlsbad, USA) or RNeasy Micro Plus kit (Qiagen, Venlo, Netherlands). cDNA was synthesized using Superscript III (Life Technologies). Primers are provided in Table S1.
Luciferase assays

Cells were transfected with PEI (Sigma, St. Louis, USA), Nucleofection (Amaxa, Basal, Switzerland), or electroporation (Bio-Rad, Hercules, USA) with 1-3 μg luciferase construct, 1-1.5 μg of pCMV5-empty or a combination of pCMV5-RUNX1^{25} and pCMV5-CBFβ^{26} expression constructs, and 50-500 ng of Renilla (pRL-null from Promega, Madison, USA) luciferase construct as an internal transfection control for 293T, NIH3T3, K562, and U937 cells. Jurkat cells were transfected by electroporation with these constructs in the same ratios with a total DNA amount of 20-25 μg. Cells were collected 24 hours after transfection and lysates were prepared. Manufacturer’s protocol for the Dual-Luciferase Reporter Assay System (Promega) and Monolight 3010 Luminometer (BD Biosciences, San Jose, USA) were followed for ascertaining luciferase activity. Firefly luciferase activity was normalized to Renilla luciferase activity.

For additional Methods, please see Supplemental Methods.
Results

*Hmga2 is a direct RUNX1 target gene in HSPCs*

To gain further insight into which genes are directly regulated by RUNX1 in HSPCs, we performed RUNX1 ChIP-seq in the murine HSPC-like EML cell line. EML cells rely on stem cell factor for survival and have the capability of differentiating into the erythroid, myeloid, and lymphoid lineages. Their multipotency and capacity to be grown in large cultures make the EML cell line an ideal cell system for ChIP-seq studies. Peak calling revealed 6370 peaks with a FDR of less than 1%. Interestingly, the location of the majority of peaks (73.3%) lay in regions more than 5 kilobases (kb) away from a transcriptional start site (TSS) (Figure 1A). Only a small percentage (10.8%) of peaks was in promoter regions, defined as 3 kb upstream of a TSS (Figure 1B). De novo motif analyses using the MEME suite revealed the most enriched motif to be TGTGGT, which is the known RUNX consensus motif (Figure 1C). Other top motifs correspond to the ETS and GATA family of transcription factors, both of which are also important players in hematopoiesis and further confirm that RUNX1 often works with these factors to exert its functions.

We previously reported the gene expression signatures of HSPCs (Lineage- /cKit+/Sca1+ or LSK cells) from wildtype, Runx1 conditional knockout (Δ/Δ) and RUNX1(41-214)-transplanted mice. RUNX1(41-214) (RUNX1SF) is primarily a dominant-negative competitor of endogenous RUNX proteins. Of the 59 genes that were commonly differentially expressed in Runx1Δ/Δ and RUNX1SF HSPCs, 27 have RUNX1 ChIP-seq peaks associated with their loci (Figure 2A, Table 1). We also performed comparisons of our ChIP-seq data with other recently published RUNX1 ChIP-seq data (Figure S1), which demonstrated further overlap of genes. The combination of these data and comparisons has revealed a list of high potential RUNX1 target genes. Within this gene list, we decided first to focus our studies on HMGA2. Based on
the published information of HMGA2, increased expression of HMGA2 is likely to result in similar cellular phenotypes as observed in RUNX1 deficiency. HMGA2 has been reported to play important roles in regulating cellular proliferation and cancer development. Our gene expression profiling data obtained from three biologically independent sets of RNA prepared from LSK cells indicate that Hmga2 expression is significantly increased (Figure 2B). Using RT-qPCR studies, we demonstrated that Hmga2 transcript levels were 6- and 15-fold higher in Runx1Δ/Δ and RUNX1SF HSPCs, respectively, relative to wildtype controls (Figure 2C). In addition, Hmga2 is up-regulated in Runx1Δ/Δ granulocyte-macrophage progenitors (GMPs) and common myeloid progenitors (CMPs) relative to wildtype controls (Figure S2). Increase in Hmga2 transcripts also correlated with increase in HMGA2 protein levels in Runx1Δ/Δ HSPCs, GMPs, and CMPs (Figure 2D-F, S3).

Upon examination of the Hmga2 gene locus, ChIP-seq revealed RUNX1 binding regions in the promoter, upstream, downstream, and third intron regions (Figure 2G). All except the promoter region were determined to be significant by MACS. These three significant RUNX1 binding regions are in agreement with recently published RUNX1 ChIP-seq data using murine HSPC lines. The binding of RUNX1 to these regions were confirmed using ChIP-qPCR (Figure 2H). The locations of these regions relative to the Hmga2 TSS suggest that they may serve as enhancers and/or silencers for the regulation of Hmga2 expression. Together, our combined gene expression and ChIP analyses show that RUNX1 directly binds to the Hmga2 locus and influences its expression.

*The Hmga2 promoter is regulated by RUNX1 in a cell type-dependent manner and is independent of canonical RUNX binding sites*
ChIP studies indicate that RUNX1 binds to the previously reported *Hmga2* promoter region [base pairs (bp) -800 to +197 relative to the TSS]\(^{34}\). To study how RUNX1 regulates *Hmga2* expression, a DNA fragment containing this region was inserted into the multiple cloning site of the pGLX promoterless luciferase reporter vector. The pGLX vector is a modified version of pGL2, which does not contain any RUNX consensus binding sites\(^{26}\). The cloned *Hmga2* promoter DNA fragment has two RUNX consensus sites at bp -363 and -213 and a polypyrimidine/polypurine tract spanning bp -84 to -25, which gives the promoter much of its activity (Figure 3A)\(^{34}\). To examine whether RUNX1 affects *Hmga2* promoter activity, we conducted transient transfection assays in non-hematopoietic cell lines (NIH3T3 and 293T) and hematopoietic cell lines (K562, Jurkat, and U937). In NIH3T3 and 293T cells, co-expression of the full length promoter with RUNX1 and its partner CBFβ reduces luciferase activity to 72% and 47%, respectively (Figure 3B). These results suggest that RUNX1 acts as a transcriptional repressor on the *Hmga2* promoter in these two adherent cell lines. Interestingly, in the K562 and Jurkat leukemia cell lines, RUNX1 increases luciferase activity by 3- and 1.6-fold, respectively (Figure 3C). Since these cell lines represent erythromyeloid and lymphoid lineages, respectively, we also conducted experiments using U937 cells, which more closely represent monocyte/myeloid cells (Figure 3C). In this myeloid representative cell line, RUNX1 acts to suppress luciferase activity to 31%. Thus, RUNX1 regulates *Hmga2* promoter activity in all of the cell lines tested, but acts on the promoter in a positive or negative manner depending on the cell and/or tissue type. Most importantly, RUNX1 acts as a transcriptional repressor of *Hmga2* in myeloid cell types.

Since the *Hmga2* promoter contains two RUNX consensus sites, we sought to ascertain whether these sites mediate RUNX1 interaction with the promoter by utilizing promoter constructs with these regions truncated or mutated. In NIH3T3 cells, mutation of either or both of the presumptive RUNX1 binding sites in the promoter does not abrogate the repression of
luciferase upon co-expression of RUNX1 and CBFβ (Figure 3D). Constructs harboring truncations of the promoter at the 5’ or 3’ end also failed to affect RUNX1-mediated repression (Figure S4). The same mutation constructs were used in U937 cells. The co-expression of RUNX1 and CBFβ was still able to repress luciferase activity with the three mutation constructs (Figure 3E). These results suggest that RUNX1 does not regulate the Hmga2 promoter through these two sites and likely exerts its effect by interacting with other intermediary factors that bind to the promoter. Given the ability of RUNX1 to interact with a variety of other transcriptional regulators5, this possibility seems likely regarding the regulation of Hmga2.

Runx1 intronic and intergenic binding regions mediate Hmga2 transcription

Significant RUNX1 binding regions were discovered in intronic and intergenic regions of the Hmga2 locus (Figure 2H, S5). We sought to determine how these regions affect the transcription of Hmga2 and whether they contribute to the effect of RUNX1 on the Hmga2 promoter. The three regions were cloned downstream of the luciferase gene in the Hmga2 promoter-luciferase construct (Figure 4A). The upstream binding element adds significant basal luciferase activity relative to the full length promoter alone in all cell lines tested (Figure 4B-C). Interestingly in NIH3T3 cells, addition of the upstream element induces enough transcriptional activation to counteract the repressive effect due to co-expression of RUNX1 and CBFβ (Figure 4B). For example, we no longer observe significant repression in the Promoter + Upstream case compared to Promoter alone. The same trends are observed using the same constructs in U937 cells (Figure 4C). Accordingly, the upstream element continues to result in heightened transcriptional activity in K562 cells (Figure S6). Physical interaction between the upstream element and the promoter was confirmed by 3C-qPCR assay (Figure 4D). The increase in transcriptional activity induced by the upstream element, and confirmation of physical interaction
between the upstream element and the promoter suggests that it may potentially act as an enhancer element.

The downstream element was also tested in NIH3T3 and U937 cells. In both cell lines, the effect of co-expression of RUNX1 and CBFβ remained intact and did not differ significantly from the promoter element by itself. In U937 cells, the downstream element also added significantly more luciferase activity compared to promoter alone.

Finally, the constructs containing the intron element with the Hmga2 promoter were used for promoter-luciferase assays. In contrast to the upstream element, the intron element significantly represses basal luciferase activity relative to the full length promoter in all of the cell lines tested (Figure 4B-C). In NIH3T3 and U937 cells, the promoter and intron construct continues to result in transcriptional repression when co-expressed with RUNX1 and CBFβ (Figure 4B). In K562 cells, addition of the intron element induces enough transcriptional repression to counteract the activation due to co-expression of RUNX1 and CBFβ (Figure S6). Similar to the case of the Promoter + Upstream element in NIH3T3 cells where the trend of repression is still observed but no longer significant, in K562 cells the activation due to co-expression of RUNX1 and CBFβ is still observed but no longer significant. These results suggest that the upstream and intron regions are enhancer and silencer elements, respectively.

**HMGA2 does not modulate the effect of RUNX1 deficiency on HSPC expansion**

Our current studies indicate that Hmga2 expression is directly regulated by RUNX1 and that Hmga2 expression is significantly increased in HSPCs in the absence of RUNX1 or in the presence of RUNX1SF (Figure 2). One of the most dramatic phenotypes of Runx1Δα mice and RUNX1SF mice is expansion of their HSPC populations.10,17,18. HMGA2 is generally thought to be a proliferation-inducing factor. Transgenic mice expressing Hmga2 and over-expression of
Hmga2 via retroviral transduction and bone marrow transplantation have both been recently shown to induce myeloproliferative disease in mice\textsuperscript{15,16}. Since Hmga2 is a target gene of RUNX1, up-regulation of Hmga2 in RUNX1-deficient HSPCs may contribute to their expansion. To examine the role of HMGA2 in HSPC expansion due to loss of RUNX1 function, we generated littermates of wild type, Hmga2\textsuperscript{-/-}, Runx1\textsuperscript{Δ/Δ}, and Hmga2 and Runx1\textsuperscript{Δ/Δ} double knockout (Double KO) mice for further studies. The bone marrow compartments of these four genotypes of mice were analyzed by staining with cell surface markers followed by flow cytometry. As expected, RUNX1 deficiency led to a significant expansion of LSK, LT-HSC, and ST-HSC populations (Figure 5A-D). However, lack of HMGA2 did not have any effects on HSPC frequency. The Double KO also displayed HSPC expansion, suggesting that loss of HMGA2 does not affect HSPC expansion due to the deficiency of RUNX1.

In addition, the common lymphoid progenitor (CLP) populations were assessed. Although we observed a trend toward higher frequencies of CLPs in the absence of RUNX1, this difference was not significant (Figure 5E). In addition, neither Hmga2\textsuperscript{-/-} nor Double KO mice exhibited any significant differences in CLP frequency.

The HSPCs in Runx1\textsuperscript{Δ/Δ} mice have been described to have lower levels of apoptosis overall\textsuperscript{35}. In addition, ectopic expression of Hmga2 has also been demonstrated to lead to lower levels of apoptosis\textsuperscript{36}. To ascertain whether Hmga2 has any role in mediating apoptosis in the context of RUNX1 loss-of-function, we examined the frequency of apoptosis in LSK cells by using Annexin V staining (Figure 5F). Hmga2\textsuperscript{-/-} and Double KO mice exhibited higher levels while Runx1\textsuperscript{Δ/Δ} mice generally exhibited lower levels of Annexin V-positive cells, demonstrating that Hmga2 plays an important role in regulating apoptosis in this population and that it is necessary for the decreased apoptosis associated with loss of RUNX1.
Hmga2 contributes to myeloid progenitor expansion caused by the loss of RUNX1

Another major phenotype of Runx1ΔΔ mice is their expansion of the myeloid progenitors, specifically the GMP and CMP populations18,19. Intriguingly, the role of HMGA2 in inducing proliferation in the hematopoietic system has primarily been biased towards the myeloid lineage15,16. To test the involvement of HMGA2 in regulating myeloid progenitors in RUNX1-deficient mice, we analyzed the bone marrow compartments of wildtype, Hmga2−/−, Runx1ΔΔ and Double KO mice and focused on their myeloid progenitor cells (Figure 6A). As expected, Runx1ΔΔ mice displayed an expansion of the GMP population compared to wildtype (1.41% of total bone marrow versus 0.59%, p < 0.001). Hmga2−/− mice, on the other hand, had significantly fewer GMPs compared to wildtype (0.25%, p < 0.001). In Double KO mice, the expansion of the GMPs is significantly rescued compared to Runx1ΔΔ mice (0.57%, p < 0.001). The CMP populations displayed a similar trend while megakaryocyte-erythrocyte progenitors (MEP) were not affected (Figure 6B-C). Representative flow cytometry charts and averages for each genotype are shown (Figure 6D). Thus, the loss of Hmga2 in a RUNX1-deficient genetic background brings the GMP and CMP frequency back down to wildtype levels. Together these results suggest that Hmga2 contributes to the myeloid progenitor expansion due to the loss of RUNX1 function.

In addition, Hmga2 regulates the decreased apoptosis observed in the myeloid progenitors of Runx1ΔΔ mice since Double KO myeloid progenitors continue to show increased frequencies of Annexin V-positive cells (Figure 6E-G).
Discussion

RUNX1 has been implicated in a variety of blood-related diseases and neoplasms\textsuperscript{5}. Many of these diseases originate from mutations in \textit{RUNX1} that occur at the stage of the HSPC and result in the loss of normal RUNX1 function. As a transcription factor, the primary biological activity of RUNX1 is to control the expression of its target genes. The goal of this study was to determine which genes RUNX1 directly targets at the level of the HSPC and whether those genes have any functions in regulating HSPCs. To this end, we used a combination of differential gene expression analyses and genome-wide transcription factor occupancy to identify prospective RUNX1 direct targets. For gene expression analyses, two models of RUNX1 deficiency were employed—an inducible conditional knockout model for RUNX1, and a bone marrow transplantation model employing RUNX1SF as a dominant-negative regulator of endogenous RUNX proteins. To elucidate DNA occupancy of RUNX1 on a genome-wide scale, ChIP-seq was conducted on a HSPC-like cell line using RUNX1 antibody. The result was a condensed list of high potential direct target genes that are regulated by RUNX1 at the level of the HSPC (Table 1). The prospect of modulating the activity of these genes provides potential candidates for creating therapies against diseases caused by the loss of RUNX1 function, and offers an intriguing look into the mechanisms of how perturbed RUNX1 function may result in disease.

All of these genes have diverse roles in mediating various functions both inside and outside of the hematological system. Briefly, \textit{Csf2rb} codes for the common beta chain subunit found in the GM-CSF, IL-3, and IL-5 receptors and is important in regulating specific cytokine responses\textsuperscript{37}. \textit{Gzmb} and \textit{Igf2r} have both been implicated in regulating cytotoxic T-cell-mediated apoptosis\textsuperscript{38}. \textit{Lcp2} is necessary for T-cell development\textsuperscript{39}. Another subset of genes are involved in cell-to-cell interactions in blood cells like \textit{Alcam}\textsuperscript{40} and \textit{Jam3}\textsuperscript{41}, and may help to mediate interactions between HSPCs and the stem cell niche. These are just a few examples of genes
that have blood-specific functions. Others like Krt80\textsuperscript{42} and Tjp1\textsuperscript{43} have described functions primarily in non-hematopoietic cells. Some genes like Fhdc1 and Zcchc18 have biological roles that are relatively unstudied or unknown. The diversity of functions in this list further confirms that RUNX1 is a master transcription factor and that disrupting RUNX1 function can have a variety of consequences resulting from disruption of its target genes. Ultimately, our analysis exposed the oncogene Hmga2 for further study. HMGA2 has known roles in mediating cellular proliferation, and a recent study described the role of HMGA2 in both fetal and adult HSPCs, but interactions between RUNX1 and HMGA2 have remained largely unexplored\textsuperscript{13,36}. The connection of these two genes in regulating various hematopoietic processes is a highly interesting avenue for further investigation.

Promoter-luciferase assays demonstrated that Hmga2 is directly regulated by RUNX1, but the regulation is cell type-specific as shown by contrasting results in a variety of cell lines. Cell lines like K562 and Jurkat cells express different sets of transcriptional regulators and co-factors that may not be present in adherent cell lines like NIH3T3 and 293T cells, or more myeloid-like lines like U937 cells. Various members of the ETS and GATA families offer just a few examples of transcription factors that primarily function in the hematopoietic system. These hematopoietic-specific transcription factors may collaborate with RUNX1 to result in cell-specific regulation of Hmga2. Furthermore, the factors cooperating with RUNX1 in regulating Hmga2 expression can differ in HSPCs, differentiated hematopoietic cells, or leukemia cells. Identification of which partner factor(s) RUNX1 interacts with in these various contexts is a topic of ongoing study.

For the first time, distal regulatory regions in and around the Hmga2 locus were demonstrated to have effects on Hmga2 expression. Interestingly, the effects exerted by the upstream and intron regions were constant and did not depend on the cell type in which they were tested. When these regions were added to the promoter-luciferase assays, the upstream
and intron regions were associated with transcriptional activation and repression, respectively, but did not affect RUNX1-mediated transcription specific to each cell line tested. Hence, the upstream and intron regions most likely serve as enhancer and silencers of \textit{Hmga2}, respectively. Another possibility is that RUNX1 binding to these regulatory elements contributes to the regulation of endogenous \textit{Hmga2} gene expression in HSPCs. Under conditions of transient transfection, however, the relatively higher levels of exogenous RUNX1 expression is sufficient for binding with its collaborating factors at the \textit{Hmga2} proximal promoter region which may overwhelm its effects from the binding to these other regulatory regions.

In addition to demonstrating that \textit{Hmga2} is a transcriptional target of RUNX1, we showed that it contributes to GMP and CMP expansion resulting from loss of RUNX1 function. Importantly, the GMP population has often been associated with harboring leukemia stem cells (LSCs) in a variety of different leukemias\textsuperscript{44–46}. A recent study focused specifically on the RUNX1-ETO (RE) fusion protein described the ability of RE to induce expansion of the GMP population and that the GMP RUNX1-ETO-expressing population can induce a leukemia-like state in an in vivo mouse model\textsuperscript{46}. Loss of RUNX1 in various conditional knockout models has also been shown to expand both the GMP and CMP populations and induce altered hematopoietic states\textsuperscript{18,19}. Hence, loss of RUNX1 function either through mutation or by involvement in chromosomal translocation leads to an increase of cells capable of eventually transitioning to leukemia. Without RUNX1 acting as a tumor suppressor, \textit{Hmga2} is allowed to become up-regulated to induce a pre-malignant state. Limiting the expansion of GMPs by decreasing or regulating the amounts of HMGA2 may provide another method of controlling leukemia progression through keeping the number of LSCs in check.

Various cases of \textit{HMGA2} disruption associated with hematological disease have been described and are typically associated with perturbing the myeloid lineages\textsuperscript{47–49}. Notably, some mutations result in truncation of the 3' untranslated region of \textit{HMGA2}. This region of \textit{HMGA2}
contains binding sites for the let-7 family of miRNAs, which target a variety of cellular mediators and is a major regulator of HMGA2 levels\textsuperscript{50,51}. Let-7 family members have been reported to be lower in RE-positive leukemia in patients and mouse models of RE\textsuperscript{52,53}, but various large gene expression studies in RE-positive leukemia patients have reported contrasting levels of HMGA2 when compared to non-RE-positive leukemia patients (The Cancer Genome Atlas). No comprehensive study has been conducted comparing the expression of let-7 family members or HMGA2 in MDS or MPN patients based on RUNX1 mutation status, but such an investigation offers an interesting future direction to further establish the role HMGA2 in the context of RUNX1 loss-of-function. In summary, our study establishes Hmga2 as a target gene of RUNX1 and that Hmga2 mediates myeloid progenitor expansion due to loss of RUNX1 function.

Acknowledgements: We would like to thank Dr. Nancy Speck for providing the Runx1 conditional knockout mice and Dr. Kiran Chada for providing Hmga2 knockout mice. We would also like to thank Feng Yue and Uli Wagner from Dr. Bing Ren’s lab for help with high throughput sequencing and analyses. Finally, we are very grateful to everyone in the Zhang Lab and Dr. Daniel Tenen for all the helpful and enlightening discussions. The work is supported by funding from the National Institutes of Health (DK080665 and DK098808). K.L. is a recipient of a National Research Service Award (F30103106).

Authorship contributions: K.L. designed, performed, analyzed experiments and wrote the manuscript; A.M., R.D., Y.H., A.G.S. M.Y., S.M, and S.W. performed or helped with experiments, H.Y. analyzed data, and D-E.Z. designed and analyzed experiments, and supervised manuscript preparation.

Conflict-of-interest disclosure: The authors declare no competing financial interests.
References


Table 1. Table of 27 common genes differentially expressed in Runx1Δ/Δ and RUNX1SF HSPCs that overlap with RUNX1 genome occupancy.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Runx1Δ/Δ array</th>
<th>RUNX1SF array</th>
<th>Peak location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adamtsl4</td>
<td>ADAMTS-like 4</td>
<td>2.42</td>
<td>3.37</td>
<td>chr3:95511225-95512034</td>
</tr>
<tr>
<td>Alcam</td>
<td>activated leukocyte cell adhesion molecule</td>
<td>-2.05</td>
<td>-4.28</td>
<td>chr16:54016417-54016980</td>
</tr>
<tr>
<td>Atp8a2</td>
<td>ATPase, aminophospholipid transporter, class I, type 8A, member 2</td>
<td>2.05</td>
<td>2.84</td>
<td>chr14:60691216-60691929 Chr14:60748188-60748880</td>
</tr>
<tr>
<td>Bh1</td>
<td>3-hydroxybutyrate dehydrogenase, type 1</td>
<td>-2.35</td>
<td>-5.59</td>
<td>chr16:30967909-30969214 Chr16:30971738-30972608  chr16:31080812-31081818  chr16:31209726-31210547</td>
</tr>
<tr>
<td>Cd72</td>
<td>CD72 antigen</td>
<td>-2.08</td>
<td>-2.32</td>
<td>chr4:43469362-43470612</td>
</tr>
<tr>
<td>Cpa3</td>
<td>carboxypeptidase A3</td>
<td>8.36</td>
<td>13.89</td>
<td>chr3:20141473-20142494 Chr3:20143698-20144216</td>
</tr>
<tr>
<td>Csf2rb</td>
<td>colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage)</td>
<td>2.41</td>
<td>2.74</td>
<td>chr15:78097680-78098791 Chr15:78150945-78151767</td>
</tr>
<tr>
<td>Cyb561</td>
<td>cytochrome b561</td>
<td>-3.78</td>
<td>-3.34</td>
<td>chr11:105807879-105808525 Chr11:105820316-105821213  chr11:105918137-105919089</td>
</tr>
<tr>
<td>Deptor</td>
<td>DEP domain containing MTOR-interacting protein</td>
<td>2.72</td>
<td>3.15</td>
<td>chr15:54963581-54964309 Chr15:55022569-55023361  chr15:55049202-55050131</td>
</tr>
<tr>
<td>Fhdc1</td>
<td>FH2 domain containing 1</td>
<td>3.48</td>
<td>4.5</td>
<td>chr3:84264390-84265426</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Name</td>
<td>Chromosome 1</td>
<td>Chromosome 2</td>
<td>Chromosome 3</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Fscn1</td>
<td>fascin homolog 1, actin bundling protein (Strongylocentrotus purpuratus)</td>
<td>chr3:84295898-84296552</td>
<td>chr3:84307981-84308734</td>
<td>chr5:143581961-143584992</td>
</tr>
<tr>
<td>Gzmb</td>
<td>granzyme B</td>
<td>chr14:56880645-56881698</td>
<td>chr14:56884808-56885570</td>
<td>chr14:56914779-56915975</td>
</tr>
<tr>
<td>Hmga2</td>
<td>high mobility group AT-hook 2</td>
<td>chr10:119760560-119761484</td>
<td>chr10:119883733-119884448</td>
<td>chr10:120053435-120054404</td>
</tr>
<tr>
<td>Igf2r</td>
<td>insulin-like growth factor 2 receptor</td>
<td>chr17:12947269-12947943</td>
<td>chr11:102330569-102332184</td>
<td>chr11:102330569-102332184</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Log2FoldChange</td>
<td>p-value</td>
<td>Chromosome/Position</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------</td>
<td>---------</td>
<td>---------------------------------------------------------</td>
</tr>
<tr>
<td>Map3k6</td>
<td>mitogen-activated protein kinase kinase kinase 6</td>
<td>2.01</td>
<td>2.39</td>
<td>chr4:132793901-132794220</td>
</tr>
<tr>
<td>Pcdh7</td>
<td>protocadherin 7</td>
<td>5.41</td>
<td>9.62</td>
<td>chr5:58108350-58109816</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr5:58126695-58127555</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr5:58232582-58233087</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr5:58248528-58249096</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr5:58422191-58422711</td>
</tr>
<tr>
<td>Podxl</td>
<td>podocalyxin-like</td>
<td>2.81</td>
<td>4.19</td>
<td>chr6:32107983-32109068</td>
</tr>
<tr>
<td>Slc24a3</td>
<td>solute carrier family 24 (sodium/potassium/calcium exchanger), member 3</td>
<td>4.95</td>
<td>7.09</td>
<td>chr2:145050905-145052230</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr2:145087407-145088337</td>
</tr>
<tr>
<td>Spp1</td>
<td>secreted phosphoprotein 1</td>
<td>-2.41</td>
<td>-3.6</td>
<td>chr5:104863923-104864409</td>
</tr>
<tr>
<td>Stx3</td>
<td>syntaxin 3</td>
<td>2.66</td>
<td>3.32</td>
<td>chr19:11881628-11882355</td>
</tr>
<tr>
<td>Tjp1</td>
<td>tight junction protein 1</td>
<td>6.45</td>
<td>6.76</td>
<td>chr7:72518004-72518826</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr7:73323785-73326321</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr7:73436915-73437490</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chr7:73448143-73449672</td>
</tr>
<tr>
<td>Zcchc18</td>
<td>zinc finger, CCHC domain containing 18</td>
<td>-2.5</td>
<td>-10.84</td>
<td>chrX:133520640-133521165</td>
</tr>
</tbody>
</table>
Figure Legends

**Figure 1. RUNX1 genome-wide occupancy in HSPCs as determined by ChIP-seq.**

(A) The 6370 RUNX1 ChIP-seq peaks with a FDR less than 1% are shown in relation to the absolute distance of the TSS of known genes in kb. (B) Distribution of RUNX1 ChIP-seq peaks are grouped based on promoter defined as >3 kb upstream of TSS (10.8%), intron (44.3%), exon (5.7%), and intergenic (39.2%) regions. (C) The 500 peaks with the lowest FDR were submitted to the MEME suite. The top motifs are shown and given with E-values, the associated motif family, and the frequency of the motif found in the 500 peaks.

**Figure 2. Hmga2 is a RUNX1 target gene in HSPCs.**

(A) Venn diagram showing the overlap between the common differentially expressed genes in Runx1<sup>Δ/Δ</sup> and RUNX1SF HSPCs (n=59)<sup>10</sup> and the genes associated with RUNX1 ChIP-seq peaks. (B) Relative expression of Hmga2 in wildtype, Runx1<sup>Δ/Δ</sup> and RUNX1SF HSPCs based on microarray data published in Matsuura et al<sup>10</sup>. (C) Validation of up-regulation of Hmga2 in microarray results by RT-qPCR using cDNA from wildtype, Runx1<sup>Δ/Δ</sup>, and RUNX1SF HSPCs. RNA extraction and RT-qPCR was performed from at least 3 independent batches of mice. HMGA2 protein levels by intracellular flow cytometry are shown in (D) LSK, (E) GMP, and (F) CMP cell populations between wildtype and Runx1<sup>Δ/Δ</sup> mice (n = 5 each). (G) Location of RUNX1 ChIP-seq peaks relative to the Hmga2 gene locus on chromosome 10. IgG control and one RUNX1 ChIP-seq replicate are shown. (H) Confirmation of RUNX1 occupancy regions in relation to the Hmga2 gene locus by ChIP-qPCR. RUNX1 ChIP was compared relative to IgG ChIP and normalized to a negative control region. Data represents 3 replicates of ChIP followed by qPCR.
Figure 3. RUNX1 controls transcription of *Hmga2* via its promoter in a tissue-specific manner.

(A) Diagram of promoter-luciferase construct showing RUNX consensus Sites 1 and 2 (*) at bp -363 and -213, respectively. (B) NIH3T3 and 293T cell lines were transfected with the full length *Hmga2* promoter-luciferase construct and luciferase assays were performed 24 hours after transfection. Co-expression was conducted with an empty expression vector (Control) or RUNX1 and CBFβ expression vectors. Error bars indicate standard deviation and are from at least 3 replicates. (C) K562, Jurkat, and U937 cell lines were transfected with the full length *Hmga2* promoter-luciferase construct and luciferase assay was performed 24 hours after transfection. Co-expression was conducted with an empty expression vector (Control) or RUNX1 and CBFβ expression vectors. Error bars indicate standard deviation and are from at least 3 replicates. (D) NIH3T3 cells were transfected with promoter-luciferase constructs with Site1, Site2, or both sites mutated. Luciferase assays were performed 24 hours after transfection. Co-expression was conducted with an empty expression vector (Control) or RUNX1 and CBFβ expression vectors. Error bars indicate standard deviation and are from at least 3 replicates. (E) U937 cells were transfected with the constructs indicated in (D) and luciferase assays were performed 24 hours after transfection. Co-expression was conducted with an empty expression vector (Control) or RUNX1 and CBFβ expression vectors. Error bars indicate standard deviation from at least 3 replicates. (* indicates p<0.05)

Figure 4. RUNX1 distal binding regions exert differential effects on *Hmga2* expression.

(A) Schematic of distal element-promoter luciferase constructs. The distal elements being tested are the RUNX1 binding regions that are downstream, upstream, and in the intron of the *Hmga2* locus. Distal element sequences were cloned downstream of the *Hmga2* promoter in the
promoter-luciferase constructs. (B) These constructs were co-transfected with empty vector
(Control) or with RUNX1 and CBFβ (designated as RUNX1) expression constructs into NIH3T3
cells. Luciferase activity was performed 24 hours after transfection, and error bars indicate
standard deviation from 4 biological replicates. (C) Same experiment as described in (B)
performed using U937 cells. Luciferase activity was performed 24 hours after transfection, and
error bars indicate standard deviation from 3 biological replicates. For (B) and (C), Promoter
alone is designated as “P”, Promoter + Upstream is “P + U”, Promoter + Downstream is “P + D”,
and Promoter + Intron is “P + I”. (D) Diagram and chart showing enrichment of interaction
between the promoter and upstream element (located at approximately -139 kilobases
upstream) as demonstrated by 3C-qPCR assay. Assay was performed using EML cells. Data
shown is average of two independent 3C assays and error bars indicate standard deviation. (*)
indicates p<0.05.

Figure 5. HMGA2 does not modulate the effect of RUNX1 deficiency on HSPC expansion.

Percentages of hematopoietic stem cell populations analyzed by flow cytometry, including (A)
LSK cells, (B) LT-HSCs, and (C) ST-HSCs. For LSK: Wildtype (Hmga2+/+ or +/-, Runx1(fl/fl),
Mx1Cre-, n = 15), Hmga2−/− (Hmga2−/−, Runx1(fl/fl), Mx1Cre-, n = 9 ), Runx1Δ/Δ (Hmga2+/+ or +/-,
Runx1(fl/fl), Mx1Cre+, n = 21), Double KO (Hmga2−/−, Runx1Δ/Δ, Mx1Cre+, n = 8). For LT-HSCs
and ST-HSCs: Wildtype (n = 10), Hmga2−/− (n = 4), Runx1Δ/Δ (n = 10), Double KO (n = 4). (D)
Representative flow cytometry gating of LSK and SLAM populations for each of the four
genotypes of mice and their averages are shown. (E) Percentages of CLP populations were
analyzed by flow cytometry. For CLPs: Wildtype (n = 5), Hmga2−/− (n = 5), Runx1Δ/Δ (n = 10),
Double KO (n = 4). (F) Averages of frequencies of Annexin V-positive/7AAD-negative cells from
the LSK cells of each genotype (n = 4 each) are shown. (*) indicates p<0.05.
Figure 6. Lack of HMGA2 rescues myeloid progenitor expansion due to loss of RUNX1.

Percentages of hematopoietic stem cell populations analyzed by flow cytometry, including (A) GMPs, (B) MEPs, and (C) CMPs. For myeloid progenitor staining, mice used were: Wildtype (Hmga2+/+ or +/-, Runx1(fl/fl), Mx1Cre-, n = 9), Hmga2−/− (Hmga2−/−, Runx1(fl/fl), Mx1Cre-, n = 9), Runx1ΔΔ (Hmga2+/+ or +/-, Runx1(fl/fl), Mx1Cre+, n = 14), and Double KO (Hmga2−/−, Runx1(fl/fl), Mx1Cre+, n = 8). (D) Representative flow cytometry gating of GMP, MEP, and CMP populations on each of the four genotypes of mice and the averages are shown. (E-G) Averages of frequencies of Annexin V-positive/7AAD-negative cells from the (E) GMP, (F) MEP, and (G) CMP gates of each genotype (n = 4 each) are shown. (*) indicates p<0.05.
Figure 1, Lam et al.
Common genes between Runx1Δ/Δ and RUNX1SF

Genes associated with Runx1 ChIP-seq peaks

Figure 2, Lam et al.
Figure 3, Lam et al.
Figure 4, Lam et al.
Figure 5, Lam et al.
Figure 6, Lam et al.
Hmga2 is a direct target gene of RUNX1 and regulates expansion of myeloid progenitors in mice

Kentson Lam, Alexander Muselman, Randal Du, Yuka Harada, Amanda G. Scholl, Ming Yan, Shinobu Matsuura, Stephanie Weng, Hironori Harada and Dong-Er Zhang

Information about reproducing this article in parts or in its entirety may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at: http://www.bloodjournal.org/site/subscriptions/index.xhtml

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include digital object identifier (DOIs) and date of initial publication.

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036. Copyright 2011 by The American Society of Hematology; all rights reserved.