3' UTR-truncated HMGA2 overexpression induces non-malignant *in vivo* expansion of hematopoietic stem cells in non-human primates

Melissa A. Bonner,[1,7] Antonio Morales-Hernández,[1,7] Sheng Zhou,[2] Zhijun Ma,[3] Jose Condori,[2] Yong-Dong Wang,[4] Soghra Fatima,[5] Lance E. Palmer,[1] Laura J. Janke,[6] Stephanie Fowler,[1] Brian P. Sorrentino,[1] and Shannon McKinney-Freeman[1]

[1]Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA; [2]Experimental Cell Therapeutics Lab, St. Jude Children's Research Hospital, Memphis, TN 38105, USA; [3]Department of Bone Marrow Transplant and Cell Therapy, St. Jude Children's Research Hospital, Memphis, TN 38105, USA; [4]Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA; [5]Immunology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA; [6]Veterinary Pathology Core, St. Jude Children's Research Hospital, Memphis, TN 38105, USA

Vector-mediated mutagenesis remains a major safety concern for many gene therapy clinical protocols. Indeed, lentiviral-based gene therapy treatments of hematologic disease can result in oligoclonal blood reconstitution in the transduced cell graft. Specifically, clonal expansion of hematopoietic stem cells (HSCs) highly expressing HMGA2, a chromatin architectural factor found in many human cancers, is reported in patients undergoing gene therapy for hematologic diseases, raising concerns about the safety of these integrations. Here, we show for the first time *in vivo* multilineage and multiclonal expansion of non-human primate HSCs expressing a 3' UTR-truncated version of HMGA2 without evidence of any hematologic malignancy >7 years post-transplantation, which is significantly longer than most non-human gene therapy pre-clinical studies. This expansion is accompanied by an increase in HSC survival, cell cycle activation of downstream progenitors, and changes in gene expression led by the upregulation of *IGF2BP2*, a mRNA binding regulator of survival and proliferation. Thus, we conclude that prolonged ectopic expression of HMGA2 in hematopoietic progenitors is not sufficient to drive hematologic malignancy and is not an acute safety concern in lentiviral-based gene therapy clinical protocols.

INTRODUCTION

Hematopoietic stem cells (HSCs) are the target cell for gene therapy of many hematologic and non-hematologic diseases because of their lifelong self-renewal and differentiation potential, which ensures long-term clinical benefits.[1–6] Lentiviral vectors represent the favored approach for *ex vivo* gene transfer into HSCs and have been used routinely in clinical gene therapy protocols.[7] However, the safety of lentiviral genomic integration remains a concern.[7] Indeed, hematopoietic clonal expansion has been observed in several clinical protocols employing lentiviral vectors. However, the clinical significance of these expansions and their propensity to convert to hematologic malignancy remains unclear.[7–9]

Disruption of HMGA2 via vector integration has been a safety concern in gene therapy for years. *Hmga2* encodes an architectural transcription factor whose chromosomal rearrangement or truncation results in tumors in patients and transgenic mice.[10–15] Upregulation of a 3' UTR-truncated *HMGA2* caused by intronic integration of a lentiviral vector resulted in HSC clonal expansion in a beta-thalassemia gene therapy trial.[16–18] Numerous clones containing intronic vector integrations in *HMGA2* have also been seen in gamma-retroviral and lentiviral SCID-X1 gene therapy.[19] Moreover, cases of clonal hematopoiesis have been linked to overexpression of proteins functionally downstream of *HMGA2*.[20] These findings are not surprising given that ectopic *Hmga2* can expand murine HSCs and knockdown or overexpression of *HMGA2* perturbs the repopulating and differentiation potential of human HSCs.[21–23] However, a single corrected and markedly expanded clone persisted without malignant transformation in an early ADA-SCID gene therapy trial,[24] suggesting that clonal expansion does not irrevocably progress to malignancy. Thus, the clinical prognosis of gene therapy patients with *HMGA2*-associated hematopoietic clonal expansions is unclear.

Here, we provide evidence that persistent expression of a 3' UTR-truncated *HMGA2* cDNA, which causes a massive overexpression of this protein, leads to the competitive oligoclonal expansion of long-term HSCs in transplanted pigtailed macaques (*M. nemestrina*) without perturbed cellular differentiation or evidence of malignancy >7 years post-transplantation.
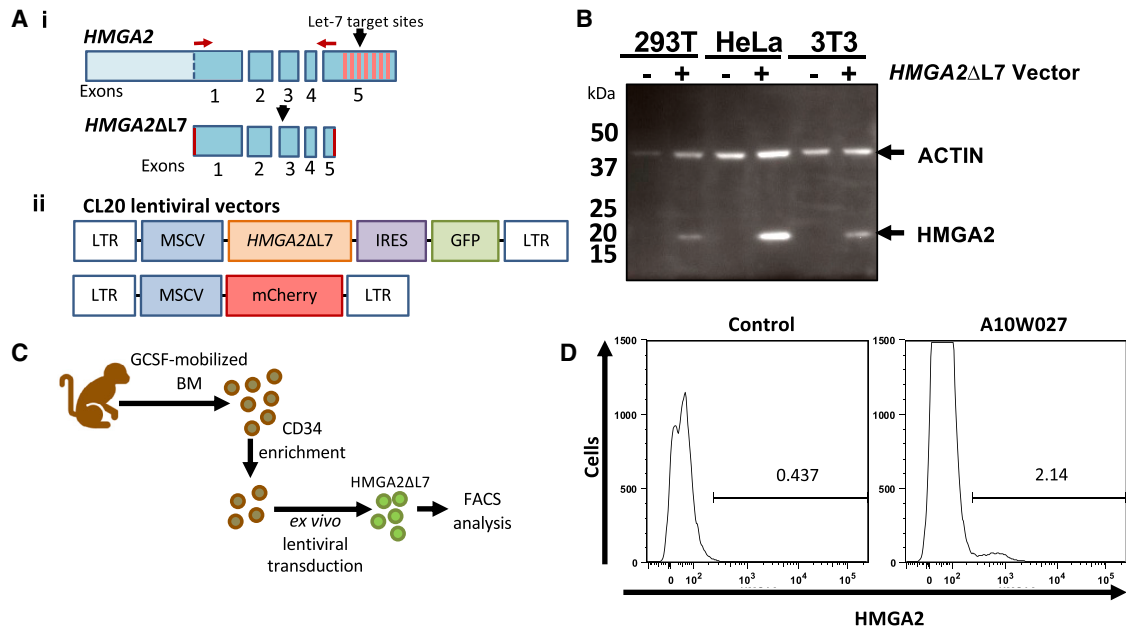
Figure 1. Design and validation of the lentiviral vector expressing 3' UTR-truncated *HMGA2* (*HMGA2Δ7*)
(Ai) *HMGA2* schematic. Let-7 miRNA target sites shown in yellow. Red arrows indicate primer binding sites used to amplify *HMGA2Δ7*. (Aii) Lentiviral vector schematics. (B) Western blot of cells transduced with *HMGA2Δ7*-GFP vector or control vector. (C) Schematics of non-human primate CD34⁺ cell isolation and transduction. (D) Intracellular *HMGA2* expression in non-human primate CD34⁺ cells by flow cytometry.

RESULTS

3' UTR-truncated *HMGA2* yields multiclonal CD34⁺ cell expansion in non-human primates

HMGA2 is an architectural non-histone chromatin protein that regulates gene expression by altering chromatin structure.[25–28] *HMGA2* contains seven let-7-micro RNA (miRNA) binding sites in the 3' UTR, which negatively regulates HMGA2 translation.[29] To overexpress HMGA2, we amplified the protein coding region devoid of let-7 binding sites (HMGA2ΔL7) from human HEK293T cells by reverse transcription-PCR. The PCR product was subcloned into the Cl20c lentiviral vector so that expression of HMGA2ΔL7 would be driven by an murine stem cell virus (MSCV) promoter, which has good activity in HSCs (Figure 1A). An internal ribosome entry site (IRES)-GFP expression cassette was also included to allow for tracking of HMGA2ΔL7-expressing cells by flow cytometry. A control vector was constructed in which mCherry cDNA is also driven by the MSCV promoter (Figure 1A). Both vectors were produced transiently in HEK293T cells with a titer of 2×10^8 tu/mL and 5×10^8 tu/mL after concentration, respectively. HEK293T, HeLa, and 3T3 cells transduced with the *HMGA2*ΔL7 vector expressed the expected 19 kDa HMGA2 protein (Figure 1B). Pigtailed macaque CD34⁺ cells transduced with the HMGA2ΔL7 vector also expressed HMGA2 (Figure 1D). GFP expression was comparable to HMGA2 (data not shown). Thus, the lentiviral HMGA2 vector can be produced with sufficient titer to transduce CD34⁺ cells from pigtailed macaque.

To test whether HMGA2 expression expands HSCs in large animals, we isolated CD34⁺ cells from granulocyte-colony stimulating factor

(GCSF)-mobilized bone marrow (BM) of pigtailed macaques. Cells were transduced with a lentiviral vector containing *HMGA2*ΔL7-GFP or the control lentiviral vector (containing mCherry only). A mix of transduced cells by the two different vectors was transplanted into the same animal after lethal irradiation (Figure 2A). Two animals (A10W016 and A10W027) were transplanted. The transduction efficiency in the total graft for animal A10W016 was 42.9% for GFP and 19.1% for mCherry, and that for animal A10W027 was 6% for GFP and 7.9% for mCherry. See Table S1 for transplant details. Blood and platelets were supplied as needed until absolute neutrophil counts (ANCs) reached 500/mL and platelet counts reached 50,000/mL (Figure S1A). Six days post-transplant, peripheral blood (PB) chimerism was 4%–6% for GFP and mCherry in both animals, despite higher contribution of GFP-marked cells in A10W016's graft (Figure 2Bi). White blood cell (WBC) chimerism decreased thereafter, and by 3 months post-transplant, PB chimerism was 2.9% GFP⁺ and 1.1% mCherry⁺ for animal A10W016 and 2.7% GFP⁺ and 3.2% mCherry⁺ for animal A10W027 (Figure 2Bi), suggesting equivalent and low HSCs and progenitor cell transduction with both vectors. Thereafter, GFP chimerism slowly but progressively increased, while mCherry chimerism decreased further in both animals (Figure 2Bi). By day 200, the GFP marking in WBCs reached 8.9% and 4.5% in A10W016 and A10W027, respectively, while the mCherry marking decreased to 0.9% and 1.5%, respectively (Figure 2Bi).

PB GFP chimerism increased over 20 months to 39% for A10W016 and 41% for A10W027, while the mCherry-marked cells decreased
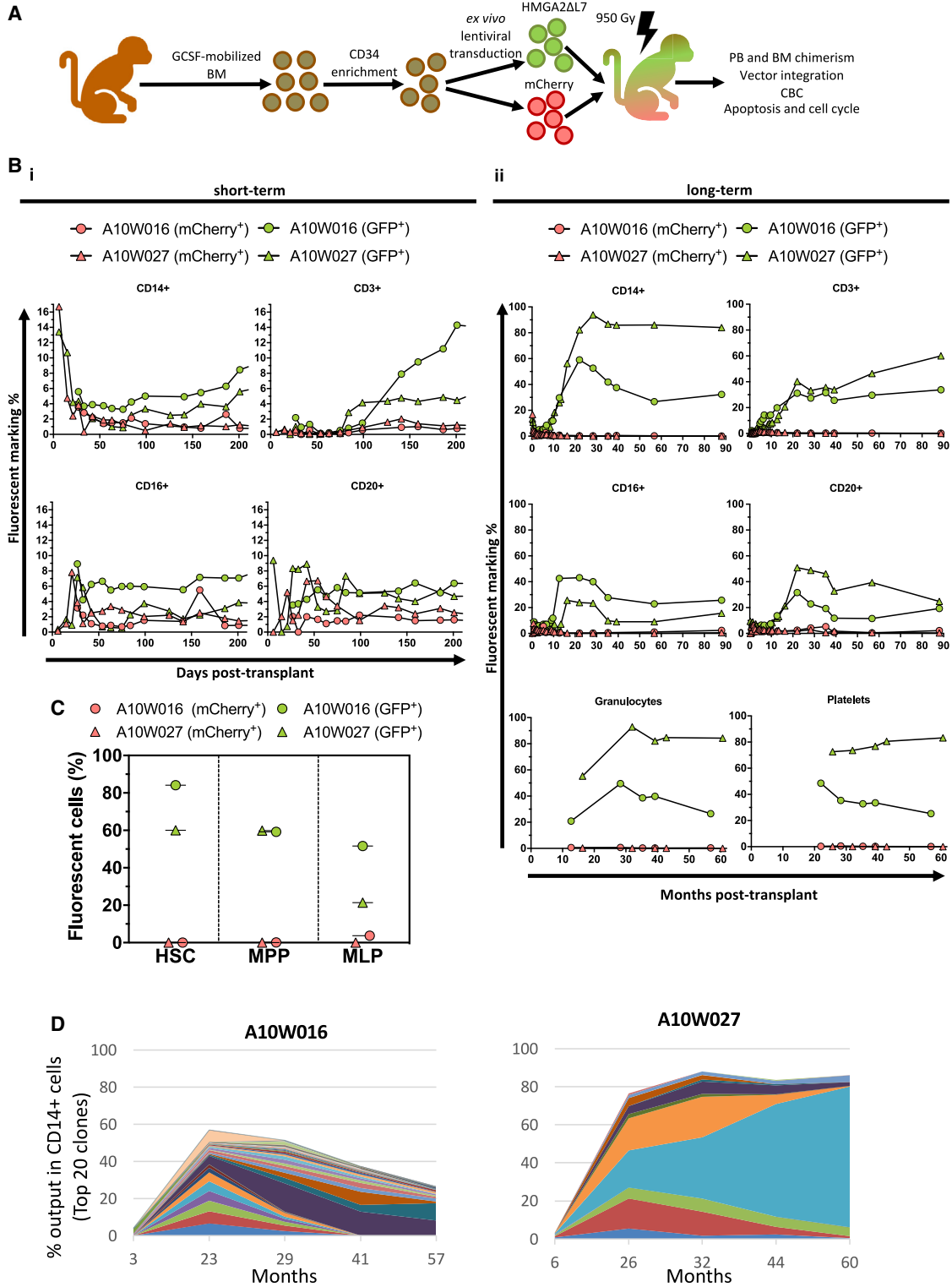
(legend on next page)

(Figure 2Bii). These values were stable for >80 months post-transplant (Figure 2Bii). Increased marking was seen in all blood lineages (CD3[+] T cells, CD14[+] myeloid cells, CD16[+] NK cells, CD20[+] B cells, platelets) (Figure 2B) as well as mononucleated cells and red cells in the BM (Figure S1B), suggesting expansion of multipotent HSCs. HSC expansion was further reflected by the presence of GFP label in BM cells, where the most primitive hematopoietic cells displayed 60% and 80% GFP chimerism (HSC: CD34[+]CD38[−]CD90[+]CD45RA[−]) for animals A10W016 and A10W027, respectively, at the latest time point of 74 months (Figure 2C). BM hematopoietic progenitors (multipotent progenitors [MPPs]: CD34[+]CD38[−]CD90[−]CD45RA[−] and multi-lymphoid progenitors [MLPs]: CD34[+]CD38[−]CD90[−/+]CD45RA[+]) also show increased GFP marking (Figure 2C). HMGA2 protein was detectable in sorted GFP[+] WBCs but not GFP[−] WBCs (Figure S1C).

Vector integration site (VIS) analysis over time and in multiple mature lineages allows tracing of cellular clonal dynamics and the differentiation of transduced cells. GFP[+] and mCherry[+] cells were sorted from PB and the genomic DNA subjected to VIS analysis. The VIS in the mCherry[+] subpopulation in A10W027 was not analyzed because of limited cell number. More than 15 different clones contributing to the blood were detected in all the lineages of both animals (Figure 2D; Table S2), indicating that the HMGA2ΔL7-driven expansion was multiclonal. The oligoclonal nature of expanded HSCs demonstrates that the expansion here was independent of vector insertion and that overexpression of HMGA2ΔL7 alone is sufficient. Remarkably, one of the clones (PRRC2A) in animal A10W027 contributes to >50% of the cells in the main PB lineages (Figure 2D; Table S2). This signifies the extraordinary repopulating potential acquired by HMGA2ΔL7 overexpression.

3' UTR-truncated HMGA2 expression enhances survival and quiescence in primate CD34[+] cells

To explore the cellular mechanisms contributing to the gradual expansion of GFP[+] PB cells in animals A10W016 and A10W027, BM from each animal was analyzed for apoptosis and cell cycle 74 months post-transplant. We compared the percentage of AnnexinV[+] cells expressing HMGA2ΔL7 (GFP[+]) and control cells (mCherry[+]). We observed a reduction (>2-fold) in apoptosis in all hematopoietic progenitors expressing HMGA2ΔL7 (Figure 3A). The cell cycle distribution of GFP[+] and mCherry[+] HSCs and MPPs was similar in transplanted animals (Figure 3B). This also holds true for human CD34[+] cells (Figures 3C–3E). However, a smaller proportion of GFP[+] MLPs were in G0 relative to mCherry[+] MLPs, suggesting greater cell cycle activity (Figure 3B). Thus, an increase in the survival of immature hematopoietic progenitors and enhanced proliferation of mature progenitors likely contribute to the gradual and selective

advantage of HMGA2ΔL7-expressing cells over time in transplanted pigtailed macaques.

3' UTR-truncated HMGA2 expression perturbs the gene expression profile of primate CD34[+] cells without malignancy

HMGA2 alters transcription of multiple genes.[30–32] Our observed HSC expansion could be due to dysregulation of multiple genes and/or independent of gene regulation. Affymetrix array analysis of RNA from BM CD34[+]GFP[+] and CD34[+]GFP[−] cells isolated 22–26 weeks post-transplant revealed 39 transcripts upregulated and three transcripts downregulated in the CD34[+]GFP[+] subpopulation compared to the CD34[+]GFP[−] subpopulation in both animals (Table S3). As expected, HMGA2 was the highest differentially expressed gene in both animals. The only other gene upregulated by >2-fold was IGF2BP2 (Figure 4A; Table S3), which is also differentially expressed in HSCs and progenitors compared to differentiated cells in humans and mice (Figure S2).[33] qRT-PCR analysis showed that the p16INK4a was not altered in animal A10W016 and was slightly reduced (67% of CD34[+]GFP[−] cells) in animal A10W027 (Figure 4B).

Complete blood counts, including total WBC and red blood cell (RBC) counts and PB lineage distribution, were in the normal range (Table S4), demonstrating lack of any detectable hematopoietic abnormality. BM smears also were normal (Figure S3A). Despite IGF2BP overexpression being linked to changes in hemoglobin expression patterns,[34] hemoglobin was also normal in recipients of HMGA2ΔL7-expressing cells (Figure S3B).

DISCUSSION

HMGA2 high expression has been linked to expansion of single corrected clones that persisted without malignant transformation for years in multiple gene therapy trial.[19,21–24] These observations suggest that clonal expansion does not irrevocably progress to malignancy. In contrast, there are concerns that HMGA2 overexpression may increase the risk of expanding clones with pre-existing oncogenic mutations, as in clonal hematopoiesis of undetermined potential.[35] Indeed, HMGA2 is considered a putative oncogen.[10–13] Chromosomal translocations with breakpoints in HMGA2 that eliminate the 3' UTR let-7 miRNA binding sites cause multiple tumor types.[36–38] High expression of HMGA2 in cancers of epithelial origin associates with increased aggressiveness[39,40] and is seen in some leukemias.[41] Chromosomal translocation of HMGA2 has also been seen in two cases of paroxysmal nocturnal hemoglobinuria.[42] Six distinct transgenic mouse models have been made with different promoters driving expression of HMGA2.[43] Interestingly, these mice never develop hematologic malignancies, even after three rounds of serial BM transplantation.[44] A key difference between these transgenic mice and our study is that

Figure 2. 3' UTR-truncated HMGA2 expression enhances CD34[+] cell *in vivo* expansion in primates

(A) Transplant schematic. (B) GFP and mCherry marking in WBCs in PB of transplanted animals. Cells were first gated on the viable lineage[+] cells, and % of GFP[+] or mCherry[+] cells within the gates is graphed. (C) GFP and mCherry marking in BM hematopoietic progenitors in transplanted animals 74 months post-transplant. Cells are first gated as viable GFP[+] or mCherry[+] cells, and normalized frequency of each progenitor compartment (HSC: CD34[+]CD38[−]CD90[+]CD45RA[−], MPP: CD34[+]CD38[−]CD90[−]CD45RA[−], and MLP: CD34[+]CD38[−]CD90[−/+]CD45RA[+]) within the gates is graphed. (D) VIS in CD14[+] cells analyzed at indicated time points. % of total read counts corresponding to the top 20 VISs within the total reads count is shown. Data represent individual values.
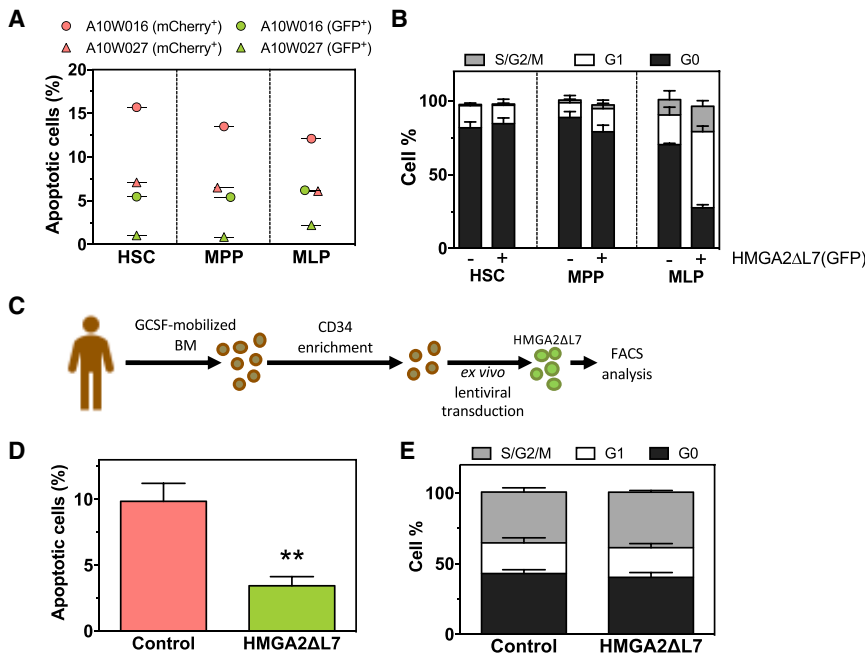
Figure 3. 3' UTR-truncated *HMGA2* expression perturbs survival and cell cycle in CD34⁺ cells

(A) Apoptosis in BM hematopoietic progenitors 74 months post-transplant. Cells are first gated as GFP⁺ or mCherry⁺ cells, and the apoptotic cell (DAPI⁻AnnexinV⁺) frequency of each progenitor compartment (HSC: CD34⁺CD38⁻CD90⁺CD45RA⁻, MPP: CD34⁺CD38⁻CD90⁻CD45RA⁻, and MLP: CD34⁺CD38⁻CD90⁻/⁺CD45RA⁺) within the gates is graphed. Data represent individual values. (B) Cell cycle status of BM hematopoietic progenitors 74 months post-transplant. Cells are first gated as GFP⁻ or GFP⁺ cells, and the cell cycle status (G0: DAPI⁻Ki67⁺, G1: DAPI⁻Ki67⁺, and S/M/G2: DAPI⁺Ki67⁺) of each progenitor compartment (HSC: CD34⁺CD38⁻CD90⁺CD45RA⁻, MPP: CD34⁺CD38⁻CD90⁻CD45RA⁻, and MLP: CD34⁺CD38⁻CD90⁻/⁺CD45RA⁺) within the gates is graphed. (C) Schematics of human CD34⁺ cell isolation and transduction. (D) Apoptosis in human hematopoietic progenitors 5 days post-transduction with control or HMGA2Δ7 vector. Cells are first gated as GFP⁺ or mCherry⁺ cells, and the apoptotic cell (DAPI⁻AnnexinV⁺) frequency within the gates is graphed. (E) Cell cycle status in human hematopoietic progenitors 5 days post-transduction with control or HMGA2Δ7 vector. Cells are first gated as GFP⁺ or mCherry⁺ cells. and the cell cycle status (G0: DAPI⁻Ki67⁻, G1: DAPI⁻Ki67⁺, and S/M/G2: DAPI⁺Ki67⁺) frequency within the gates is graphed. Data represent mean ± SEM. **p > 0.005.

overexpression of *HMGA2* started in the fetus and could not be downregulated later during development because of a pan-cellular promoter. Additionally, IGF2BP overexpression is also implicated in the malignant transformation of human CD34⁺ cells.[45] However, animals transplanted here with HMGA2ΔL7-expressing cells displayed a normal hematologic profile. HMGA2 directly regulates *IGF2BP2* during myogenesis[46] and embryonic development. IGF2BP2 upregulation has also been associated with the regulation of stem cells during tissue regeneration.[46,47] IGF2BP proteins have been ascribed an "anti-apoptotic role" in multiple cell contexts,[48–50] and IGF2BP is known to boost proliferation of several tissues.[51–53] Thus, upregulated *IGF2BP2* may explain our observed increased survival and proliferation of hematopoietic stem and progenitor cells (HSPCs). Alternatively, HMGA2 binds the DNA replication fork in murine embryonic stem cells and protects against nucleolytic fork collapse, which can cause double strand DNA breaks and trigger apoptosis, a differentiation block, or senescence.[28] Thus, it is possible that HMGA2 stabilizes replication forks in HSC, thereby reducing apoptosis while also enhancing self-renewal, resulting in the slow but progressive competitive expansion of GFP⁺ HSCs in transplanted animals. Future studies will be needed to test these potential molecular mechanisms. HMGA2 expression also suppresses the CDKN2a locus in neural stem cells in young mice.[54]

In sum, here we show for the first time in non-human primates that long-term HSCs dramatically expand *in vivo* when 3' UTR-truncated HMGA2 (HMGA2ΔL7) is overexpressed. This expansion was slow but progressive, taking ∼2 years for cells to expand from 1%–3% to 40%–60%. Once at those levels, expansion stabilized for >7 years, reaching 60%–80% in the most primitive HSPCs. Importantly, expansion was oligoclonal, as multiple GFP-marked clones were observed. Further, expanded HSCs maintained multilineage differentiation potential as evidenced by the progressive increase of GFP⁺ cells and the presence of a multiple identical VIS in all PB lineages. Finally, expansion was accompanied by increased HSPC survival and enhanced proliferation of downstream progenitors, which likely contributed to the competitive advantage of HMGA2ΔL7-expressing cells.

Importantly, the two transplanted pigtailed macaques in our study showed no evidence of hematologic malignancies >7 years post-transplant. Thus, even in this "worse-case scenario" where HMGA2 is massively overexpressed and no longer subject to typical post-transcriptional regulation, HMGA2 overexpression was not oncogenic in non-human primate HSCs. Recently, one case of acute myeloid leukemia (AML) and one case of myelodysplastic syndrome (MDS) have been reported in two participants with sickle cell disease who were treated with lentiviral vector gene therapy; it is currently unclear whether these malignancies are related to vector integration.[55] Yet our study, with truncated HMGA2 overexpression driven by a strong viral promoter in a lentiviral vector and no detected hematopoietic malignancy, suggests that massive HMGA2 overexpression in the context of lentiviral integration is not sufficient to drive malignancy. Although these clones should be monitored in the gene therapy setting, our data suggest that they are benign and unlikely to progress to malignancy.

A

B
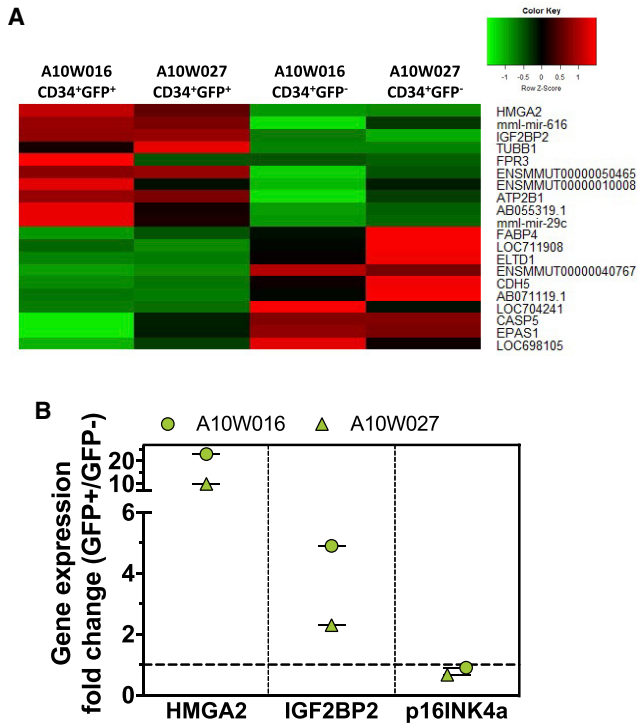
Figure 4. 3′ UTR-truncated *HMGA2* expression modifies gene expression in CD34⁺ non-human primate cells

Total RNA from sorted CD34⁺GFP⁺ or CD34⁺GFP⁻ cells at 22–26 weeks post-transplant were subjected to Affymetrix chip array analysis. (A) Heatmap of top 10 upregulated or downregulated genes for each animal. (B) qRT-PCR analysis for *HMGA2*, *IGF2BP2*, and *p16INK4a* in sorted CD34⁺ cells. The ratio of GFP⁺ to GFP⁻ is shown on the y axis. Data represent individual values.

MATERIALS AND METHODS

Animal usage

Healthy juvenile pigtailed macaques (*M. nemestrina*) weighing between 3 and 8 kg were housed at the St. Jude Children's Research Hospital Animal Research Center (Memphis, TN, USA) and at the Keeling Center for Comparative Medicine and Research at MD Anderson Cancer Center (Houston, TX, USA). All studies and procedures were reviewed and approved by the St. Jude Children's Research Hospital Institutional Animal Care and Use Committee.

Cell culture

HEK293T, HeLa, and NIH 3T3 cells were obtained from the American Type Culture Collection (Manassas, VA, USA). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum, 2 mM L-glutamine (Invitrogen, Waltham, MA, USA), 50 U/mL penicillin, and 50 mg/mL streptomycin (Invitrogen, Waltham, MA, USA), unless otherwise specified, and maintained at 37°C and 5% CO_2.

Lentiviral vector design, production, and titration

Human HMGA2DL7 (*HMGA2* lacking the Let-7 target sites) was amplified from HEK293T cDNA with primers designed to exclude the 3′ UTR Let-7 target sites (forward: TGACGAATTCAGGCAG GATGAGCGCACGCG; reverse: GCGAGCATCGATCTAGTC CTC TTCGGCA). The forward primer contains an EcoRI restriction site (underlined), and the reverse primer contains a ClaI restriction site (underlined). The HMGA2DL7 amplicon (359 bp) was subcloned by using the designed EcoRI and ClaI restriction sites to create pCL20c-MSCV-HMGA2DL7-IRES-GFP. Lentiviral vector was produced with HEK293T cells, using calcium phosphate precipitation to cotransfect the four plasmids: the CL20 vector plasmid (HMGA2DL7 or mCherry), pCAG-VSVG, pCAG4-RTR2, and pCAGkGP1R. The following day, cells were washed once in PBS and viral collection media. Stemline supplemented with 1% human serum albumin (HSA) and 2 mM L-alanyl-L-glutamine (Cellgro, Herndon, VA, USA) was added. Vector-containing supernatant was collected 48 h post-transient transfection. Human HMGA2 has been used previously in murine cells and shown to be functional.[56] Also, human and macaque HMGA2 proteins are nearly identical (only 7/147 amino acids differ, while 55/147 amino acids differ compared with murine HMGA2 protein).

To determine vector titer, 2×10^5 HeLa cells were transduced in 6-well dishes with 6 μg/mL polybrene and various volumes of vector in a final volume of 2 mL/well. The following day, 5 mL of medium was added to each well. Cells were analyzed for GFP or mCherry expression 4–6 days later with flow cytometry.

Western blot

Cells ($1–2 \times 10^6$) were collected, washed with PBS, resuspended in 50 mL of PBS with protease inhibitor (Halt) and then mixed with 50 mL of 2× Laemmli buffer with 5% β-mercaptoethanol. Samples were boiled for 10 min and then loaded onto 4%–12% NuPage gels. Antibodies were against HMGA2 (Abcam, Cambridge, MA, USA) and Actin (SCBT, Dallas, TX, USA).

Flow cytometry

Whole blood or BM samples were treated with BD Pharm Lyse (BD Biosciences, San Jose, CA, USA) for 5 min at room temperature and then spun for 10 min at $300 \times g$ at 4°C. Leukocytes were resuspended in magnetic cell sorting (MACS) buffer and stained with indicated antibodies against cell surface markers at 4°C for 20–40 min. Table S5 details all antibodies used in this article. Samples were analyzed with an LSR Fortessa (BD Biosciences, San Jose, CA, USA). Intracellular HMGA2 staining was achieved with the Foxp3/Transcription Factor Staining Buffer Set (eBioscience, San Diego, CA, USA) according to manufacturer's protocol and HMGA2 antibody (Cell Signaling Technology, Danvers, MA, USA).

CD34⁺ cell enrichment

GCSF-mobilized BM was harvested from two pigtailed macaques. BM was diluted 10-fold with RBC lysis buffer (BD Biosciences, San Jose, CA, USA) and incubated at room temperature for 5 min prior to centrifugation at $300 \times g$ for 10 min at 4°C. Cells were resuspended in MACS buffer (Miltenyi Biotec, Carlsbad, CA, USA) and passed through a 70 mm cell strainer. Cells were counted and resuspended

at 2×10^8 cells/mL in MACS buffer. Cells were incubated with CD34-PE antibody for a total of 45 min on ice with gentle shaking. Antibody-stained cells were then washed, resuspended at 2×10^8 cells/mL, and incubated with a combination of anti-PE microbeads (Miltenyi Biotec, Carlsbad, CA, USA) for 30 min on ice with gentle shaking. Cells were washed, resuspended at 2×10^8 cells/mL, and enriched with a LD column (Miltenyi Biotec, Carlsbad, CA, USA). CD34$^+$ cells were eluted and analyzed by flow cytometry for purity.

CD34$^+$ cell culture and transduction
CD34+ cells were cultured in X-VIVO 10 medium (Lonza, Portsmouth, NH, USA) supplemented with 1% HSA and 100 ng/mL human stem cell factor (hSCF), human thrombopoietin (hTPO), and human Fms-like tyrosine kinase 3 ligand (hFLT3-L) (CellGenix, Portsmouth, NH, USA). Cells were cultured on tissue culture dishes coated with RetroNectin (TaKaRA Bio, Shiga, Japan). Cells were cultured for one night prior to the first overnight vector exposure in the presence of 4 mg/mL protamine sulfate. After the first vector exposure, cells were washed and cultured in the previously described X-VIVO 10 medium for 8–12 h before a second overnight vector exposure. The multiplicity of infection (MOI) of transduction was between 50 and 100. The next day, cells were washed in fresh medium and resuspended in PlasmaLyte-148 (Baxter Healthcare, Cleveland, MS, USA) with 2% HSA (Sigma-Aldrich, St. Louis, MO, USA) and passed through a 70-mm filter prior to infusion into autologous recipients irradiated with 950 cGy.

Vector insertion site determination
The method for vector insertion site determination has been detailed elsewhere.[57] Genomic DNA from sorted cell populations was isolated with DNeasy Blood & Tissue Kits (QIAGEN, Germantown, MD, USA). A control sample with 19 known and confirmed integration sites was spiked into each sample at 2% of the final amount. DNA samples were subjected to random shearing with a E210 sonicator (Covaris, Woburn, MA, USA) (200 cycles/burst, Intensity 3, 65 s). After shearing, the DNA ends were repaired with NEBNext End Repair Module, and then a poly(A) tail was added with the NEBNext dA-Tailing Module (New England Biolabs, Ipswich, MA, USA). Then, the NEBNext Adaptor for Illumina was added, followed by a round of linear amplification with a biotinylated primer. The amplified product was then enriched using streptavidin Dynal M-270 beads (Thermo Fisher Scientific, Waltham, MA, USA) and amplified further with a final nest PCR. Samples were pooled and run on an E-gel and DNA ranging from ∼250 to 800 bp was extracted with a DNA Gel Extraction Kit (QIAGEN, Germantown, MD, USA) and analyzed with Mi-Seq (Illumina, San Diego, CA, USA). 100–1,000 ng of genomic DNA isolated from sorted subpopulations of PB was used for VIS analysis using a quantitative PCR.

Microarray analysis
Total RNA was extracted from sorted GFP$^+$CD34$^+$ or GFP$^-$CD34$^+$ BM cells and subjected to Cynomolgus/Rhesus Gene 1.0 ST Array (Thermo Fisher Scientific, Waltham, MA, USA) with established protocols.

Statistics
Variables of interest are summarized by mean, standard deviation, median, and range by groups. After checking normality by Shapiro-Wilk test, two-sample t test (one-sample t test if control group being normalized) is applied to groups following normal distribution, or two-sample exact Wilcoxon rank sum test (one-sample Wilcoxon test if control group being normalized) is applied to groups not fulfilling normal assumption, to examine whether there is significant difference between two groups. False discovery rate (FDR) method is employed to adjust for multiple testing problem. All analyses were produced in R-3.5.3.

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2021.04.013.

AUTHOR CONTRIBUTIONS
M.A.B., S.Z., and B.P.S. conceived the study. M.A.B., S.Z., Z.M., S. Fatima, J.C., S. Fowler, and A.M.-H. performed the study and analyzed and organized data. Y.-D.W. performed vector insertion site analysis. L.E.P. performed Affymetrix array data analysis. L.J.J. performed pathological analysis. M.A.B., S.Z., B.P.S., A.M.-H., and S.M.-F. wrote the paper.

DECLARATION OF INTERESTS
The authors declare no competing interests.

REFERENCES
1. Bryder, D., Rossi, D.J., and Weissman, I.L. (2006). Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. Am. J. Pathol. 169, 338–346.

2. Kondo, M., Wagers, A.J., Manz, M.G., Prohaska, S.S., Scherer, D.C., Beilhack, G.F., Shizuru, J.A., and Weissman, I.L. (2003). Biology of hematopoietic stem cells and progenitors: implications for clinical application. Annu. Rev. Immunol. 21, 759–806.

3. Körbling, M., and Estrov, Z. (2003). Adult stem cells for tissue repair - a new therapeutic concept? N. Engl. J. Med. 349, 570–582.

4. Copelan, E.A. (2006). Hematopoietic stem-cell transplantation. N. Engl. J. Med. 354, 1813–1826.

5. Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., and Dick, J.E. (2011). Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science 333, 218–221.

6. Brunstein, C.G., Gutman, J.A., Weisdorf, D.J., Woolfrey, A.E., Defor, T.E., Gooley, T.A., Verneris, M.R., Appelbaum, F.R., Wagner, J.E., and Delaney, C. (2010). Allogeneic hematopoietic cell transplantation for hematologic malignancy: relative risks and benefits of double umbilical cord blood. Blood 116, 4693–4699.

7. Milone, M.C., and O'Doherty, U. (2018). Clinical use of lentiviral vectors. Leukemia 32, 1529–1541.

8. Maldarelli, F., Wu, X., Su, L., Simonetti, F.R., Shao, W., Hill, S., Spindler, J., Ferris, A.L., Mellors, J.W., Kearney, M.F., et al. (2014). HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. Science 345, 179–183.

9. Moiani, A., Paleari, Y., Sartori, D., Mezzadra, R., Miccio, A., Cattoglio, C., Cocchiarella, F., Lidonnici, M.R., Ferrari, G., and Mavilio, F. (2012). Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. J. Clin. Invest. 122, 1653–1666.

10. Sun, M., Song, C.X., Huang, H., Frankenberger, C.A., Sankarasharma, D., Gomes, S., Chen, P., Chen, J., Chada, K.K., He, C., and Rosner, M.R. (2013). HMGA2/TET1/ HOXA9 signaling pathway regulates breast cancer growth and metastasis. Proc. Natl. Acad. Sci. USA 110, 9920–9925.

11. Guo, Z., He, C., Yang, F., Qin, L., Lu, X., and Wu, J. (2019). Long non-coding RNA-NEAT1, a sponge for miR-98-5p, promotes expression of oncogene HMGA2 in prostate cancer. Biosci. Rep. 39, BSR20190635.

12. Klemke, M., Müller, M.H., Wosniok, W., Markowski, D.N., Nimzyk, R., Helmke, B.M., and Bullerdiek, J. (2014). Correlated expression of HMGA2 and PLAG1 in thyroid tumors, uterine leiomyomas and experimental models. PLoS ONE 9, e88126.

13. Di Cello, F., Hillion, J., Hristov, A., Wood, L.J., Mukherjee, M., Schuldenfrei, A., Kowalski, J., Bhattacharya, R., Ashfaq, R., and Resar, L.M. (2008). HMGA2 participates in transformation in human lung cancer. Mol. Cancer Res. 6, 743–750.

14. Fejzo, M.S., Yoon, S.J., Montgomery, K.T., Rein, M.S., Weremowicz, S., Krauter, K.S., Dorman, T.E., Fletcher, J.A., Mao, J.I., Moir, D.T., et al. (1995). Identification of a YAC spanning the translocation breakpoints in uterine leiomyomata, pulmonary chondroid hamartoma, and lipoma: physical mapping of the 12q14-q15 breakpoint region in uterine leiomyomata. Genomics 26, 265–271.

15. Ashar, H.R., Fejzo, M.S., Tkachenko, A., Zhou, X., Fletcher, J.A., Weremowicz, S., Morton, C.C., and Chada, K. (1995). Disruption of the architectural factor HMGI-C: DNA-binding AT hook motifs fused in lipomas to distinct transcriptional regulatory domains. Cell 82, 57–65.

16. Battista, S., Fidanza, V., Fedele, M., Klein-Szanto, A.J., Outwater, E., Brunner, H., Santoro, M., Croce, C.M., and Fusco, A. (1999). The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis. Cancer Res. 59, 4793–4797.

17. Arlotta, P., Tai, A.K., Manfioletti, G., Clifford, C., Jay, G., and Ono, S.J. (2000). Transgenic mice expressing a truncated form of the high mobility group I-C protein develop adiposity and an abnormally high prevalence of lipomas. J. Biol. Chem. 275, 14394–14400.

18. Fedele, M., Battista, S., Manfioletti, G., Croce, C.M., Giancotti, V., and Fusco, A. (2001). Role of the high mobility group A proteins in human lipomas. Carcinogenesis 22, 1583–1591.

19. Wang, G.P., Berry, C.C., Malani, N., Leboulch, P., Fischer, A., Hacein-Bey-Abina, S., Cavazzana-Calvo, M., and Bushman, F.D. (2010). Dynamics of gene-modified progenitor cells analyzed by tracking retroviral integration sites in a human SCID-X1 gene therapy trial. Blood 115, 4356–4366.

20. Espinoza, D.A., Fan, X., Yang, D., Cordes, S.F., Truitt, L.L., Calvo, K.R., Yabe, I.M., Demirci, S., Hope, K.J., Hong, S.G., et al. (2019). Aberrant Clonal Hematopoiesis following Lentiviral Vector Transduction of HSPCs in a Rhesus Macaque. Mol. Ther. 27, 1074–1086.

21. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., et al. (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature 467, 318–322.

22. Ronen, K., Negre, O., Roth, S., Colomb, C., Malani, N., Denaro, M., Brady, T., Fusil, F., Gillet-Legrand, B., Hehir, K., et al. (2011). Distribution of lentiviral vector integration sites in mice following therapeutic gene transfer to treat β-thalassemia. Mol. Ther. 19, 1273–1286.

23. De Ravin, S.S., Wu, X., Moir, S., Anaya-O'Brien, S., Kwatemaa, N., Littel, P., Theobald, N., Choi, U., Su, L., Marquesen, M., et al. (2016). Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. Sci. Transl. Med. 8, 335ra57.

24. Schmidt, M., Carbonaro, D.A., Speckmann, C., Wissler, M., Bohnsack, J., Elder, M., Aronow, B.J., Nolta, J.A., Kohn, D.B., and von Kalle, C. (2003). Clonality analysis after retroviral-mediated gene transfer to CD34+ cells from the cord blood of ADA-deficient SCID neonates. Nat. Med. 9, 463–468.

25. Yu, H., Lim, H.H., Tjokro, N.O., Sathiyanathan, P., Natarajan, S., Chew, T.W., Klonisch, T., Goodman, S.D., Surana, U., and Dröge, P. (2014). Chaperoning HMGA2 protein protects stalled replication forks in stem and cancer cells. Cell Rep. 6, 684–697.

26. Copley, M.R., Babovic, S., Benz, C., Knapp, D.J., Beer, P.A., Kent, D.G., Wohrer, S., Treloar, D.Q., Day, C., Rowe, K., et al. (2013). The Lin28b-let-7-Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. Nat. Cell Biol. 15, 916–925.

27. Krahn, N., Meier, M., To, V., Booy, E.P., McEleney, K., O'Neil, J.D., McKenna, S.A., Patel, T.R., and Stetefeld, J. (2017). Nanoscale Assembly of High-Mobility Group AT-Hook 2 Protein with DNA Replication Fork. Biophys. J. 113, 2609–2620.

28. Zhao, X., Peter, S., Dröge, P., and Yan, J. (2017). Oncofetal HMGA2 effectively curbs unconstrained (+) and (-) DNA supercoiling. Sci. Rep. 7, 8440.

29. Shell, S., Park, S.M., Radjabi, A.R., Schickel, R., Kistner, E.O., Jewell, D.A., Feig, C., Lengyel, E., and Peter, M.E. (2007). Let-7 expression defines two differentiation stages of cancer. Proc. Natl. Acad. Sci. USA 104, 11400–11405.

30. Cesana, M., Guo, M.H., Cacchiarelli, D., Wahlster, L., Barragan, J., Doulatov, S., Vo, L.T., Salvatori, B., Trapnell, C., Clement, K., et al. (2018). A CLK3-HMGA2 Alternative Splicing Axis Impacts Human Hematopoietic Stem Cell Molecular Identity throughout Development. Cell Stem Cell 22, 575–588.e7.

31. Venkatesan, N., Deepa, P., Vasudevan, M., Khetan, V., Reddy, A.M., and Krishnakumar, S. (2014). Integrated Analysis of Dysregulated miRNA-gene Expression in HMGA2-silenced Retinoblastoma Cells. Bioinform. Biol. Insights 8, 177–191.

32. Yu, K.R., Park, S.B., Jung, J.W., Seo, M.S., Hong, I.S., Kim, H.S., Seo, Y., Kang, T.W., Lee, J.Y., Kurtz, A., and Kang, K.S. (2013). HMGA2 regulates in vitro aging and proliferation of human umbilical cord blood-derived stromal cells through the mTOR/p70S6K signaling pathway. Stem Cell Res. (Amst.) 10, 156–165.

33. Bagger, F.O., Kinalis, S., and Rapin, N. (2019). BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. Nucleic Acids Res. 47 (D1), D881–D885.

34. de Vasconcellos, J.F., Tumburu, L., Byrnes, C., Lee, Y.T., Xu, P.C., Li, M., Rabel, A., Clarke, B.A., Guydosh, N.R., Proia, R.L., and Miller, J.L. (2017). IGF2BP1 overexpression causes fetal-like hemoglobin expression patterns in cultured human adult erythroblasts. Proc. Natl. Acad. Sci. USA 114, E5664–E5672.

35. Steensma, D.P. (2018). Clinical consequences of clonal hematopoiesis of indeterminate potential. Blood Adv. 2, 3404–3410.

36. Rabban, J.T., Dal Cin, P., and Oliva, E. (2006). HMGA2 rearrangement in a case of vulvar aggressive angiomyxoma. Int. J. Gynecol. Pathol. 25, 403–407.

37. Medeiros, F., Erickson-Johnson, M.R., Keeney, G.L., Clayton, A.C., Nascimento, A.G., Wang, X., and Oliveira, A.M. (2007). Frequency and characterization of HMGA2 and HMGA1 rearrangements in mesenchymal tumors of the lower genital tract. Genes Chromosomes Cancer 46, 981–990.

38. Pierantoni, G.M., Santulli, B., Caliendo, I., Pentimalli, F., Chiappetta, G., Zanesi, N., Santoro, M., Bulrich, F., and Fusco, A. (2003). HMGA2 locus rearrangement in a case of acute lymphoblastic leukemia. Int. J. Oncol. 23, 363–367.

39. Wang, Y., Hu, L., Wang, J., Li, X., Sahengbieke, S., Wu, J., and Lai, M. (2018). HMGA2 promotes intestinal tumorigenesis by facilitating MDM2-mediated ubiquitination and degradation of p53. J. Pathol. 246, 508–518.

40. Zhang, H., Tang, Z., Deng, C., He, Y., Wu, F., Liu, O., and Hu, C. (2017). HMGA2 is associated with the aggressiveness of tongue squamous cell carcinoma. Oral Dis. 23, 255–264.

41. Marquis, M., Beaubois, C., Lavallée, V.P., Abrahamowicz, M., Danieli, C., Lemieux, S., Ahmad, I., Wei, A., Ting, S.B., Fleming, S., et al. (2018). High expression of HMGA2 independently predicts poor clinical outcomes in acute myeloid leukemia. Blood Cancer J. 8, 68.

42. Inoue, N., Izui-Sarumaru, T., Murakami, Y., Endo, Y., Nishimura, J., Kurokawa, K., Kuwayama, M., Shime, H., Machii, T., Kanakura, Y., et al. (2006). Molecular basis

of clonal expansion of hematopoiesis in 2 patients with paroxysmal nocturnal hemo-globinuria (PNH). Blood *108*, 4232–4236.

43. Anand, A., and Chada, K. (2000). In vivo modulation of Hmgic reduces obesity. Nat. Genet. *24*, 377–380.

44. Ikeda, K., Mason, P.J., and Bessler, M. (2011). 3'UTR-truncated Hmga2 cDNA causes MPN-like hematopoiesis by conferring a clonal growth advantage at the level of HSC in mice. Blood *117*, 5860–5869.

45. Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., et al. (2018). Recognition of RNA N[6]-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat. Cell Biol. *20*, 285–295.

46. Li, Z., Gilbert, J.A., Zhang, Y., Zhang, M., Qiu, Q., Ramanujan, K., Shavlakadze, T., Eash, J.K., Scaramozza, A., Goddeeris, M.M., et al. (2012). An HMGA2-IGF2BP2 axis regulates myoblast proliferation and myogenesis. Dev. Cell *23*, 1176–1188.

47. Zhou, X., Li, M., Huang, H., Chen, K., Yuan, Z., Zhang, Y., Nie, Y., Chen, H., Zhang, X., Chen, L., et al. (2016). HMGB2 regulates satellite-cell-mediated skeletal muscle regeneration through IGF2BP2. J. Cell Sci. *129*, 4305–4316.

48. Kessler, S.M., Pokorny, J., Zimmer, V., Laggai, S., Lammert, F., Bohle, R.M., and Kiemer, A.K. (2013). IGF2 mRNA binding protein p62/IMP2-2 in hepatocellular car-cinoma: antiapoptotic action is independent of IGF2/PI3K signaling. Am. J. Physiol. Gastrointest. Liver Physiol. *304*, G328–G336.

49. Jiang, T., Li, M., Li, Q., Guo, Z., Sun, X., Zhang, X., Liu, Y., Yao, W., and Xiao, P. (2017). MicroRNA-98-5p Inhibits Cell Proliferation and Induces Cell Apoptosis in Hepatocellular Carcinoma via Targeting IGF2BP1. Oncol. Res. *25*, 1117–1127.

50. Xu, Y., Zheng, Y., Liu, H., and Li, T. (2017). Modulation of IGF2BP1 by long non-cod-ing RNA HCG11 suppresses apoptosis of hepatocellular carcinoma cells via MAPK signaling transduction. Int. J. Oncol. *51*, 791–800.

51. Liu, Y., Dong, N., Miao, J., Li, C., Wang, X., and Ruan, J. (2019). Lin28 promotes dental pulp cell proliferation via upregulation of cyclin-dependent proteins and inter-action with let-7a/IGF2BP2 pathways. Biomed. Pharmacother. *113*, 108742.

52. Wu, X.L., Lu, R.Y., Wang, L.K., Wang, Y.Y., Dai, Y.J., Wang, C.Y., Yang, Y.J., Guo, F., Xue, J., and Yang, D.D. (2018). Long noncoding RNA HOTAIR silencing inhibits in-vasion and proliferation of human colon cancer LoVo cells via regulating IGF2BP2. J. Cell. Biochem. *120*, 1221–1231.

53. Huang, R.S., Zheng, Y.L., Li, C., Ding, C., Xu, C., and Zhao, J. (2018). MicroRNA-485-5p suppresses growth and metastasis in non-small cell lung cancer cells by targeting IGF2BP2. Life Sci. *199*, 104–111.

54. Nishino, J., Kim, I., Chada, K., and Morrison, S.J. (2008). Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. Cell *135*, 227–239.

55. Leonard, A., and Tisdale, J.F. (2021). A pause in gene therapy: Reflecting on the unique challenges of sickle cell disease. Mol. Ther. *29*, 1355–1356.

56. Efanov, A., Zanesi, N., Coppola, V., Nuovo, G., Bolon, B., Wernicle-Jameson, D., Lagana, A., Hansjuerg, A., Pichiorri, F., and Croce, C.M. (2014). Human HMGA2 protein overexpressed in mice induces precursor T-cell lymphoblastic leukemia. Blood Cancer J. *4*, e227.

57. Zhou, S., Fatima, S., Ma, Z., Wang, Y.D., Lu, T., Janke, L.J., Du, Y., and Sorrentino, B.P. (2016). Evaluating the Safety of Retroviral Vectors Based on Insertional Oncogene Activation and Blocked Differentiation in Cultured Thymocytes. Mol. Ther. *24*, 1090–1099.