Article
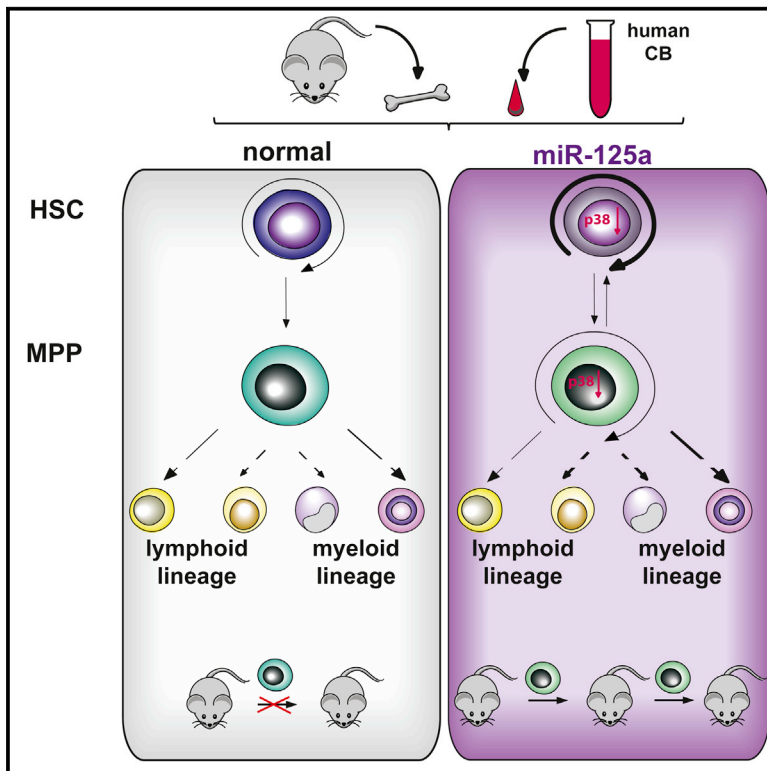
Ectopic miR-125a Expression Induces Long-Term Repopulating Stem Cell Capacity in Mouse and Human Hematopoietic Progenitors

Graphical Abstract

Authors

Edyta E. Wojtowicz, Eric R. Lechman, Karin G. Hermans, ...,
Leonid V. Bystrykh, John E. Dick, Gerald de Haan

Correspondence

jdick@uhnres.utoronto.ca (J.E.D.),
g.de.haan@umcg.nl (G.d.H.)

In Brief

Wojtowicz et al. report that enforced expression of miR-125a confers enhanced long-term self-renewal potential to multipotent hematopoietic progenitors (MPPs). Quantitative proteomics reveals a miR-125a target network that normally functions to restrain self-renewal in more committed progenitors. These results suggest that enhanced MPPs may be used to augment limited HSC sources.

Highlights

- Enforced expression of miR-125 increases hematopoietic stem frequency in vivo

- miR-125 induces stem cell potential in murine and human progenitor cells

- miR-125 represses, among others, targets of the MAP kinase signaling pathway

- miR-125 function and targets are conserved in human and mouse

CellPress

Cell Stem Cell

Article

CellPress

Ectopic miR-125a Expression Induces Long-Term Repopulating Stem Cell Capacity in Mouse and Human Hematopoietic Progenitors

Edyta E. Wojtowicz,[1,9] Eric R. Lechman,[2,9] Karin G. Hermans,[2,9] Erwin M. Schoof,[2] Erno Wienholds,[2] Ruth Isserlin,[3] Peter A. van Veelen,[4] Mathilde J.C. Broekhuis,[1] George M.C. Janssen,[4] Aaron Trotman-Grant,[2] Stephanie M. Dobson,[2,5] Gabriela Krivdova,[2,5] Jantje Elzinga,[2] James Kennedy,[2] Olga I. Gan,[2] Ankit Sinha,[6] Vladimir Ignatchenko,[6] Thomas Kislinger,[6] Bertien Dethmers-Ausema,[1] Ellen Weersing,[1] Mir Farshid Alemdehy,[7] Hans W.J. de Looper,[7] Gary D. Bader,[3] Martha Ritsema,[1] Stefan J. Erkeland,[8] Leonid V. Bystrykh,[1,10] John E. Dick,[2,5,10,*] and Gerald de Haan[1,10,*]

[1]Laboratory of Ageing Biology and Stem Cells, European Research Institute for the Biology of Ageing, University Medical Centre Groningen, University of Groningen, Antonius Deusinglaan 1, 9700 AV Groningen, the Netherlands
[2]Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada
[3]The Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada
[4]Departments of Immunohematology and Blood Transfusion, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands
[5]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
[6]Department of Medical Biophysics, University of Toronto, Toronto, ON M5G 1L7, Canada
[7]Department of Hematology, Erasmus University Medical Center Cancer Institute, Wytemaweg 80, 3015 CN Rotterdam, the Netherlands
[8]Department of Immunology, Erasmus University Medical Center, Wytemaweg 80, 3015CN Rotterdam, the Netherlands
[9]Co-first author
[10]Co-senior author
*Correspondence: jdick@uhnres.utoronto.ca (J.E.D.), g.de.haan@umcg.nl (G.d.H.)
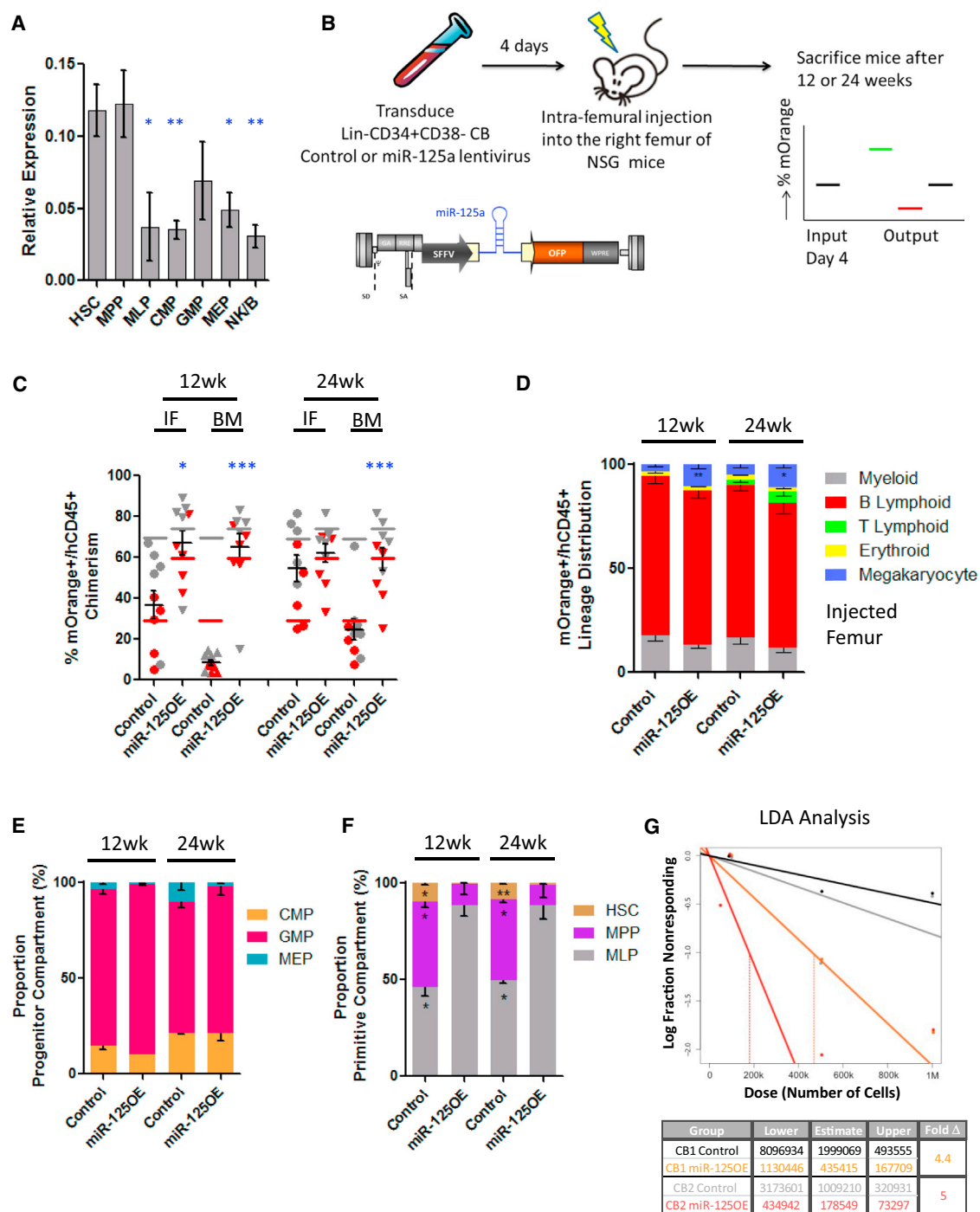http://dx.doi.org/10.1016/j.stem.2016.06.008

SUMMARY

Umbilical cord blood (CB) is a convenient and broadly used source of hematopoietic stem cells (HSCs) for allogeneic stem cell transplantation. However, limiting numbers of HSCs remain a major constraint for its clinical application. Although one feasible option would be to expand HSCs to improve therapeutic outcome, available protocols and the molecular mechanisms governing the self-renewal of HSCs are unclear. Here, we show that ectopic expression of a single microRNA (miRNA), miR-125a, in purified murine and human multipotent progenitors (MPPs) resulted in increased self-renewal and robust long-term multi-lineage re-population in transplanted recipient mice. Using quantitative proteomics and western blot analysis, we identified a restricted set of miR-125a targets involved in conferring long-term repopulating capacity to MPPs in humans and mice. Our findings offer the innovative potential to use MPPs with enhanced self-renewal activity to augment limited sources of HSCs to improve clinical protocols.

INTRODUCTION

Allogeneic stem cell transplantation is an established treatment for malignant diseases such as leukemia, myelodysplasia, and myelofibrosis as well as non-malignant disorders such as severe aplastic anemia and congenital immune or enzyme deficiencies, with >40,000 transplants undertaken worldwide (Aljurf et al., 2014; Pineault and Abu-Khader, 2015). Despite this success, ~67% of patients who require allogeneic transplantation are not able to undergo this procedure due to the lack of matched donors (Oran and Shpall, 2012). Umbilical cord blood (CB) banks have the potential to alleviate these obstacles due to broader representation of histocompatibility leukocyte antigen (HLA) types, less restriction on HLA matching, reduced incidence of graft versus host disease (GvHD), and rapid access compared to registries. However, widespread use of CB is impaired due to insufficient hematopoietic stem cell (HSC) numbers in single CB units, resulting in extended times to full engraftment (Ballen et al., 2013). Transplants combining two unrelated CB units have attempted to overcome the shortfall in HSC number, although improvements are modest (Barker et al., 2001; Wagner et al., 2014). Multiple efforts have been undertaken to attempt HSC expansion, but these have largely been empirical and rarely successful (Metcalf, 2008; Ogawa, 1993; Sauvageau et al., 2004). Recently, promising HSC expansion protocols have employed small molecules such as prostaglandin E2, SR1, UM171, or UM729, although clinical efficacy remains elusive (Boitano et al., 2010; Fares et al., 2014; North et al., 2007).

The unique self-renewal properties of HSCs are thought to be controlled by a combination of intrinsic and extrinsic signaling networks that finely balance the equilibrium among quiescence, proliferation, and differentiation. Generating increased numbers of HSC could come from controlling self-renewal or inducing self-renewal in the more numerous progenitor compartment. To achieve the latter, numerous signaling pathways would have to be appropriately fine-tuned. Small, non-coding RNAs such as microRNAs (miRNAs) act as post-transcriptional regulators of gene expression, and one miRNA can target hundreds of transcripts comprising an entire network of regulated genes,

Figure 1. Enforced Expression of miR-125a Gives a Proliferative Advantage to HSCs

(A) RT-PCR expression levels of miR-125a in flow-sorted populations from the human Lin⁻ CB hierarchy. Cells were sorted to high purity according to the scheme outlined in Figure S1A. RNU48 was used as an endogenous control, and miR-125a levels were normalized to RNU48 levels.

(B) The scheme describing the experimental system.

(C) Human mOrange⁺ (mO⁺) chimerism at 12 and 24 weeks post-transplantation with pre-sorted CD34⁺CD38⁻ umbilical cord blood cells transduced with a miR-125 lentiviral vector (miR-125OE) (n = 10) or an empty control vector (n = 10) in injected femur (IF) and distant bone marrow (BM) sites. Red and gray symbols represent replicate experiments using two distinct human CB pools. Color-matched bars represent input levels of mOrange⁺ cells at time of transplant for each experiment.

(D) Lineage distribution of mO⁺ BM grafts 12 and 24 weeks post-transplantation.

(E) Frequency of CMP, GMP, and MEP populations recovered from total BM within the CD34⁺CD38⁺ committed progenitor compartment.

(F) Proportion of HSC, MPP, and MLP from total BM within the CD34⁺CD38⁻ primitive compartment 12 and 24 weeks post-transplant.

(legend continued on next page)

positioning miRNAs as pivotal regulators of cellular homeostasis (Bartel, 2009; Filipowicz et al., 2008). miRNA profiling of hematopoietic stem and progenitor cells (HSPCs) has identified several miRNAs as differentially expressed and functionally able to govern HSC properties (Gerrits et al., 2012; Lechman et al., 2012; O'Connell et al., 2008, 2010). The family of miR-125 is particularly interesting, as they are highly expressed in murine HSCs (Gerrits et al., 2012; Guo and Scadden, 2010; O'Connell et al., 2010). This family consists of three paralogs (miR-125a, miR-125b1, and miR-125b2) sharing an identical seed sequence and exerting similar effects on HSPCs (Wojtowicz et al., 2014). MicroRNA-125a expression levels significantly decrease upon hematopoietic differentiation in mice (Gerrits et al., 2012; O'Connell et al., 2010), implying a correlation between the level of miR-125 activity and differentiation status. Ectopic expression of miR-125 (a, b1, or b2) in murine bone marrow (BM) cells results in an increased self-renewal, a proliferative advantage to HSCs and the initiation of myeloproliferation or leukemia, depending on miR-125 expression level (O'Connell et al., 2010; Wojtowicz et al., 2014). Targeted downregulation of miR-125a and miR-125b caused reduced HSPC self-renewal in vitro (Guo and Scadden, 2010) and reduced proliferation and blood cell output in vivo (Ooi et al., 2010; Wojtowicz et al., 2014), confirming a pivotal role for miR-125 family members in the molecular control of HSC self-renewal.

In the current study, we explored the potential of a single miRNA, miR-125a, to confer long-term repopulating ability to non-stem cell populations. Here, we report that enforced expression of miR-125a endows multipotent progenitors (MPPs) with enhanced self-renewal potential, offering an innovative means to use MPPs to augment limited sources of HSCs to improve clinical outcomes.

RESULTS

Enforced Expression of miR-125a Increases Human HSC Self-Renewal

Since miR-125a is highly expressed within mouse long-term HSCs (LT-HSCs) and has been shown to increase their self-renewal activity (Gerrits et al., 2012; Guo and Scadden, 2010; O'Connell et al., 2010; Ooi et al., 2010; Wojtowicz et al., 2014), we wanted to evaluate if these attributes are evolutionarily conserved within human hematopoiesis. We sorted immunophenotypically defined HSC and progenitor cell subsets from human-lineage-depleted CB (Figure S1A) (Doulatov et al., 2010; Notta et al., 2011) to obtain a detailed expression profile of mature miR-125a. We observed that the highest expression levels of miR-125a were restricted to the HSC-enriched $CD34^+CD38^-CD45RA^-CD90^+CD49f^+$ and $CD34^+CD38^-CD45RA^-CD90^-CD49f^-$ (MPP) fractions and become significantly downregulated in multi-lymphoid progenitors (MLPs) and committed $CD34^+CD38^+$ progenitor fractions (common myeloid progenitor [CMP], granulocyte monocyte progenitor [GMP], and megakaryocytic-erythroid progenitor [MEP]) (Figure 1A). These data demonstrate

an inverse correlation between differentiation state and miR-125a expression levels.

To study the function of miR-125a in human HSPCs, we generated a lentiviral vector with miR-125a (miR-125OE) under a constitutive SFFV promoter using an orange fluorescent protein (OFP) reporter (Figure 1B). To determine the impact of increased miR-125a expression (Figure S1B) on human HSPC function, $CD34^+CD38^-$ human CB HSPCs transduced with miR-125OE and control vector transduced cells were analyzed in xenotransplantation assays (Figure 1B). At 12 and 24 weeks after transplantation, we observed significant increases in human chimerism within distant BM sites in the miR-125aOE group compared to controls; chimerism levels in the injected femur (IF) were significantly increased only at 12 weeks (Figure 1C). Mice transplanted with miR-125OE cells also exhibited white BM and splenomegaly (Figure S1C). Furthermore, within enlarged spleens, we found increased proportions of $CD19^+CD10^+CD20^-$ B lymphoid cells, suggesting a partial B cell differentiation block at the pro-B cell stage (Figure S1D). Examination of the lineage distribution of OFP^+ cells in the BM revealed that all lineages were present, with significant expansion of the $CD41^+$ megakaryocyte population (Figure 1D) and an increased $CD3^+$ T cell population. At the progenitor level (Figures 1E and S1E), we observed a trend for increased GMPs and decreased MEP frequency in the BM, whereas the frequency of the multi-lymphoid progenitor fraction (MLP) was significantly increased at the expense of immunophenotypic HSCs and MPPs in the primitive compartment (Figures 1F and S1E). These observations are in line with previous reports in murine HSPCs pointing to conservation of function across species (Gerrits et al., 2012; Guo and Scadden, 2010; O'Connell et al., 2010).

The self-renewal potential of transduced human HSC within primary engrafted mice was assessed using quantitative secondary transplantation assays. Recipients transplanted with miR-125OE cells showed higher chimerism levels compared to controls, and the HSC frequency increased by 4.4- and 5-fold in miR-125OE recipients in independent experiments (Figures 1G and S1F), similar to published murine studies (Gerrits et al., 2012; Guo and Scadden, 2010; O'Connell et al., 2010).

Enforced Expression of miR-125a Confers a Stem Cell-like Program to Human $CD34^+CD38^+$ Committed Progenitors and MPPs

Given the robust increase in HSC frequency upon miR-125a expression, we asked whether more committed cell populations might also be endowed with enhanced self-renewal. We flow sorted $CD34^+CD38^-$ compartment sub-populations (HSC, MPP, and MLP) and $CD34^+CD38^+$ committed progenitors, transduced them with miR-125OE or control vectors, and transplanted cells into immuno-deficient mice measuring repopulation after 12 weeks. miR-125a overexpression in $CD34^+CD38^+$ progenitors produced a substantial graft that disseminated to distant BM locations after 12 weeks, whereas control transduced $CD34^+CD38^+$ cells generated a barely discernable

---

(G) Secondary transplantations. BM cells were harvested from two primary recipients and flow sorted for $hCD45^+mO^+$ cells. Stem cell frequencies were evaluated by limiting dilution transplantation into secondary mice. Confidence intervals were calculated using ELDA software.

All data reflect mean values ± SEM. Statistical significance was assessed using a non-parametric Mann-Whitney test, where *p < 0.05, **p < 0.01, and ***p < 0.001.

CellPress

graft (in one mouse recipient only) in the IF site (Figure 2A). As expected, only control transduced cells from sorted HSC (CD34$^+$CD38$^-$CD45RA$^-$CD90$^+$CD49f$^+$) generated a disseminating graft in recipient mice (Figure 2B). Similarly, sorted HSCs and MPPs (CD34$^+$CD38$^-$CD45RA$^-$CD90$^-$CD49f$^-$) transduced with miR-125OE generated robust engraftment within the IFs and distant BM locations (Figure 2B). Sorted and transduced MLP (CD34$^+$CD38$^-$CD45RA$^+$) were unable to generate any detectable graft (Figure 2B). In all cases, xenografts generated by CD34$^+$CD38$^+$ and MPP transduced with miR-125OE showed multi-lineage repopulation (Figures 2C and 2D). To evaluate if the grafts derived from CD34$^+$CD38$^+$ committed progenitors and MPP were durable, we transplanted these cells into secondary recipients. Whereas control transduced progenitors did not generate a graft in secondary recipient mice, the CD34$^+$CD38$^+$ and MPP cells that overexpressed miR-125a generated a multi-lineage graft for 20 weeks (five out of seven and six out of ten mice, respectively; Figures 2E and 2F). Furthermore, MPP secondary grafts remained multi-lineage (Figure 2G), and in no case did we observe myeloproliferative disease or leukemia in mice engrafted with miR-125OE human CB cells. Finally, to evaluate the long-term self-renewal potential of miR-125a expressing MPP, we transplanted the secondary MPP grafts into tertiary recipients for another 10 weeks. After a total of 30 weeks in mice, we observed small but measurable multi-lineage grafts both from CD34$^+$CD38$^-$ CB progenitors and MPP-expressing miR-125a (Figures S2A–S2D).
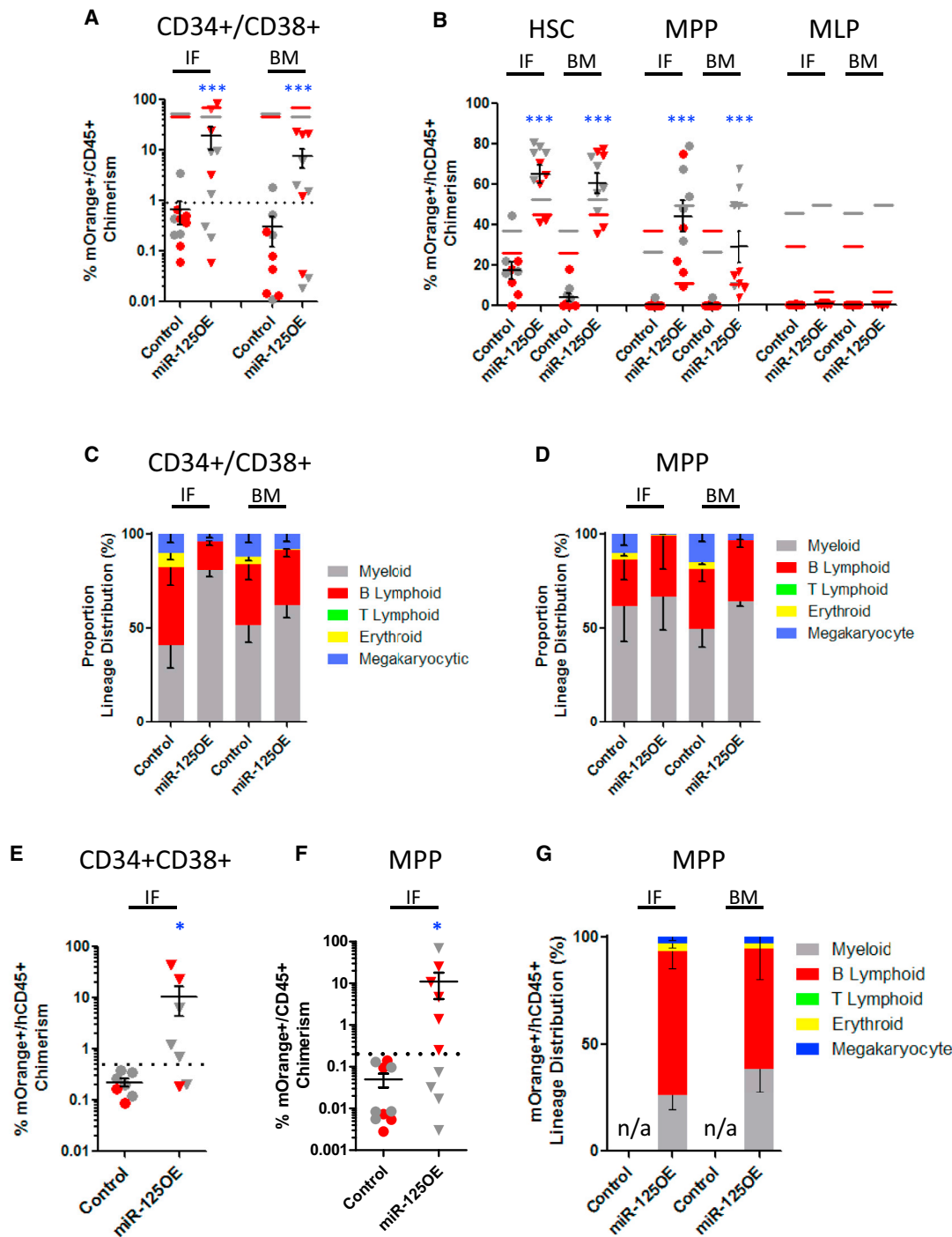
To distinguish between rare contaminating HSCs or enhanced self-renewal of MPPs as the primary mechanism for the observed MPP phenotype in secondary and tertiary mice, we used a modified splinkerette PCR approach to assess the clonality of secondary miR-125OE MPP grafts. Our data demonstrate that miR-125OE MPP BM grafts were oligoclonal, with no lentivirus insertions near or within strong oncogenes (Table S1), suggesting that the enhanced self-renewal observed in miR-125OE MPP mice was not attributable to rare contaminating HSCs. Given that our data strongly suggests that MPPs are frequently endowed with enhanced self-renewal potential, we should be able to model this phenomenon at the clonal level. Using in vitro long-term culture initiating cell (LTC-IC) assays that were established with presorted transduced MPP (Table S2), we found that miR-125OE increased the LTC-IC frequency of transduced MPPs by ~4-fold compared to controls (1/11.9 versus 1/44.7, respectively, p = 0.0000000975), with a concomitant ~9-fold increase in colonies/LTC-IC (6.6 ± 2.1 versus 54.4 ± 11.6, p = 0.0041). Together, these data strongly suggest that the enhancement of self-renewal by enforced expression of miR-125 occurs not only in HSCs but also in MPPs and in an as-yet-unidentified population within the CD34$^+$38$^+$ committed progenitor compartment.

miR-125a Expression Prolongs Self-Renewal Activity in Murine MPPs

To assess whether this enhanced self-renewal induced by miR-125a is conserved in mice, we overexpressed miR-125a (Figure 3A) in Lin$^-$Sca-1$^+$c-Kit$^+$ cells, depleted from CD150$^+$CD48$^-$ cells (hereafter referred as progenitors) and in LT-HSCs (Lin$^-$Sca-1$^+$c-Kit$^+$CD150$^+$CD48$^-$) (Figure S3A). We confirmed enforced overexpression by qPCR (Figure S3B) and performed a

series of in vivo HSC assays. As controls, we included cells transduced with a vector expressing only GFP and a vector expressing a mutant form of miR-125a, where the fourth base of the seed sequence was modified (for a seed 5′-CCC**T**GAC-3′ miR-125$^{mut4T→C}$) (Figure 3A). The mutant construct was as effectively expressed and processed as non-mutated miR-125a (Figure S3B). We competitively transplanted transduced cells in limiting dilution to lethally irradiated recipients together with a high dose of 2 × 10^6 fresh BM cells (Figure 3A) and monitored GFP$^+$ chimerism over time. As expected, miR-125OE in LT-HSCs strongly increased their competitive reconstitution activity compared to controls, similar to what we and others have reported (Figure 3B) (Gerrits et al., 2012; Guo and Scadden, 2010; O'Connell et al., 2010; Ooi et al., 2010; Wojtowicz et al., 2014) and similar to our results using human cells. Control transduced progenitors (and the mutant construct; data not shown) were unable to provide long-term repopulation (Figure 3C). By contrast, miR-125OE progenitors generated robust, high levels of multi-lineage reconstitution up to 16 weeks post-transplantation (Figure 3C). Limiting dilution experiments demonstrated a 10-fold expansion of stem cell pool size (Figures 3D, 3E, and S4A). Importantly, miR-125OE progenitors sustained long-term blood cell production with an estimated frequency similar to control LT-HSCs (Figures 3D, 3E, and S4A). Using cellular barcoding (for gene transfer efficiency and cell doses, see Figure S4B), we observed a significant increase in the number of clones contributing to myeloid engraftment upon miR-125a overexpression, indicating that induction of stem cell activity in progenitors was not a rare event, and likely not from contamination of rare LT-HSC (Figure 3F). At week 20 post-transplantation, we observed a reduced number of clones compared to week 4, but overall, the numbers of contributing clones were comparable in all groups (Figure 3F). Similar to our human CB data, mouse 125OE progenitors were able to sustain multi-lineage potential with myeloid skewing represented by increased numbers of Gr-1$^+$ cells (Figure 3G) and increased numbers of GMPs, at the expense of MEPs (Figure 3H).

To assess the potential of miR-125a to induce self-renewal, we serially transplanted 10^7 BM cells from primary recipients into lethally irradiated secondary recipients. While the repopulating potential of control LT-HSCs in secondary recipients was restricted to short-term engraftment (up to 6 weeks, most probably due to the high dose of competitors used in the first round of transplantation), the secondary recipients transplanted with miR-125OE LT-HSCs or progenitors were highly reconstituted, with an increased frequency of myeloid cells 20 weeks post-transplantation in peripheral blood (Figures 4C and 4D). No lethality was observed in primary recipients. However, in secondary mice, we observed ~20% mortality, starting at week 32, in mice transplanted with either miR-125OE LT-HSC or progenitors. Necropsy revealed increased white blood cell (WBC) counts, splenomegaly, pale bones, and blast-like cells in the blood. Fluorescence-activated cell sorting (FACS) analysis of BM and spleen revealed dominance of EGFP$^+$ cells of myeloid origin (Gr-1$^+$ and/or Mac-1$^+$, a MPN-like phenotype, as previously described; Gerrits et al., 2012; data not shown). In tertiary recipients, mortality further increased to 40% for 125OE progenitors, and 60% for 125OE LT-HSCs. Surviving mice in the tertiary transplantation showed no engraftment (Figure 4E). In

Figure 2. miR-125a-Enforced Expression Improves the Repopulating Abilities of MPPs and CD34+CD38+ Committed Progenitors

(A) Human mO+ chimerism at 12 weeks post-transplantation with pre-sorted CD34+CD38+ committed progenitor compartment from human cord blood transduced with control or miR-125a expressing lentivectors. Red and gray symbols represent replicate experiments using two distinct human CB pools. Color-matched bars represent input levels of mOrange+ cells at time of transplant for each experiment.

(B) Human mO+ chimerism at 12 weeks post-transplantation with pre-sorted HSC, MPP, and MLP human cord blood populations transduced with control or miR-125a expressing lentivectors.
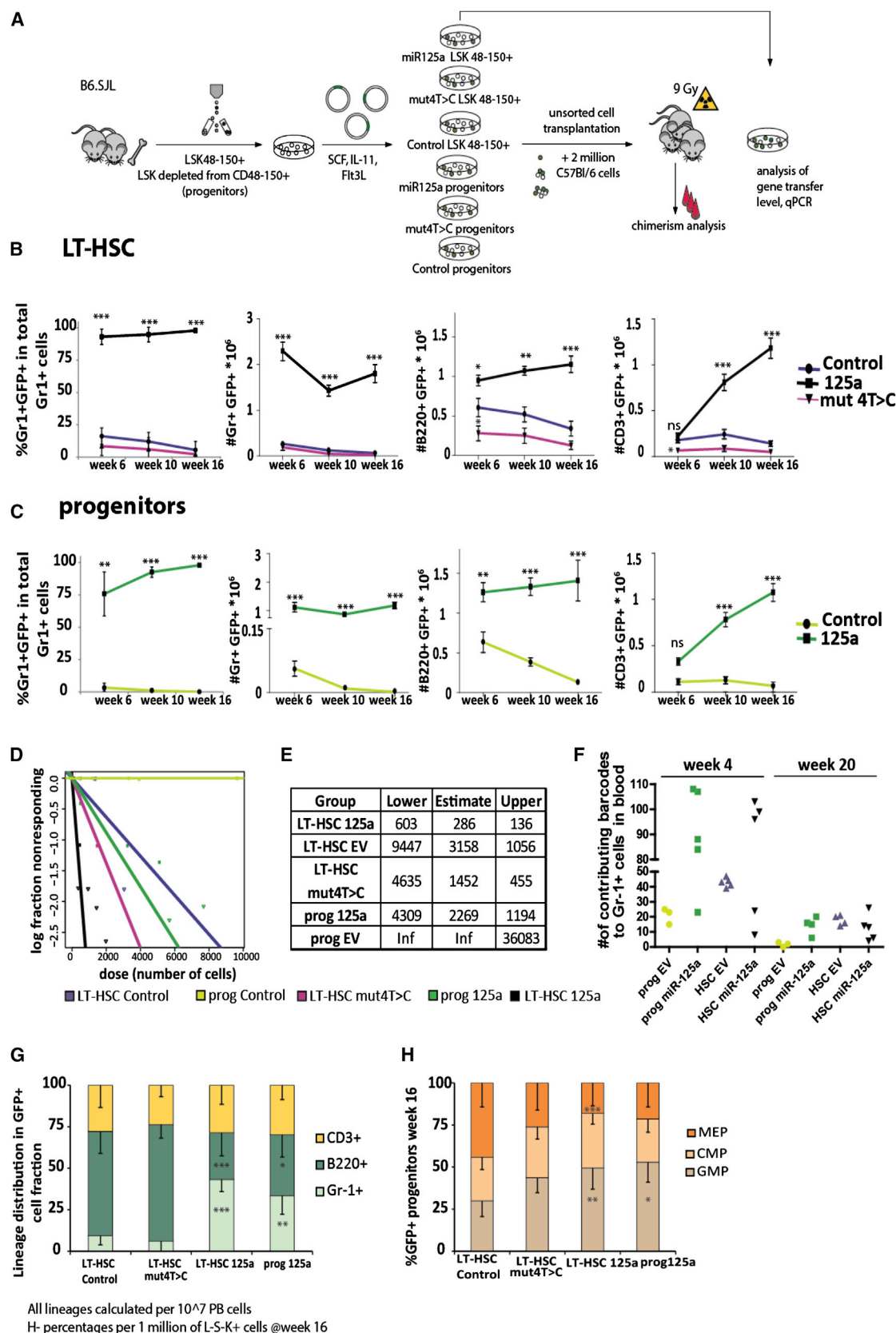
(C) Lineage distribution of mO+ BM grafts from transplanted CD34+CD38+ CB cells 12 weeks post-transplantation.

(D) Lineage distribution of mO+ BM grafts from transplanted MPP 12 weeks post-transplantation.

(E) Human mO+ chimerism at 8 weeks post-secondary transplant from CD34+CD38+ progenitors from primary mouse recipients.

(F) Human mO+ chimerism at 8 weeks post-secondary transplant from MPP from primary mouse recipients.

(G) Lineage distribution of mO+ BM grafts from MPP secondary recipient mice.

(legend on next page)

addition, mice transplanted with 125OE progenitors exhibited an increased fraction of erythroid progenitors (Figures S5A and S5B). These data suggest that both LT-HSCs and progenitors overexpressing miR-125a remained functionally active for significantly longer periods of time compared to control cells (Figure 4F), and the cell type distribution of GFP$^+$ cells in secondary recipients 20 weeks post-transplantation showed myeloid skewing in mice transplanted with LT-HSC overexpressing miR-125a.

We postulated that the myeloproliferative disease observed within 125OE mice could be due to a combination of miR-125-induced intrinsic changes and replicative stress (Figure 4E). To explore the role of replicative stress, we transplanted control and miR-125OE LT-HSCs into non-irradiated W41 recipients (Figure 4G). During the 11-month follow-up period, chimerism levels in 125OE LT-HSCs increased from 26% at transplantation reaching upward of 90% after 9 weeks post transplantation and remained very high. In contrast, the contribution of control LT-HSCs was stable over time, starting at 30% and slightly decreasing to 20% at week 45 (Figure 4H). In this non-irradiated model, we observed only ~40% mortality in mice transplanted with miR-125OE HSC, a 50% reduction compared to serial transplantation, where replicative stress is augmented (Figure 4I). These data suggest that replicative stress, due to irradiation and accompanying cytokine storm, contributes to the development of myeloproliferative disease in 125OE murine cells.

As recently reported, miR-125a overexpression leads to changes in expression of several HSC markers, including Sca-1, c-Kit, and CD150, precluding standard LT-HSC isolation from repopulated mice (O'Connell et al., 2010; Wojtowicz et al., 2014). To determine which cell population contained secondary repopulating potential, we harvested BM cells from reconstituted primary recipients and sub-fractionated GFP$^+$ Lin$^-$ cells into four populations (Figures 4J and 4K). Sub-fractionated cells from primary recipients were transplanted with competitor cells into secondary recipients (see Figures 4J, 4K, and S5C for transplanted cell doses). Repopulating activity in miR-125-overexpressing cells was exclusively observed in L$^-$S$^-$K$^+$ and L$^-$S$^+$K^{mid} cells derived from LT-HSC and in L$^-$S$^-$K$^+$ cells if progenitors were used (Figures 4J and 4K). Taken together, these data show that miR-125a overexpression interferes with markers expression and cells with the repopulating potential are present in L$^-$S$^-$K$^+$ and/or L$^-$S$^+$K^{mid} cell populations upon miR-125a overexpression.

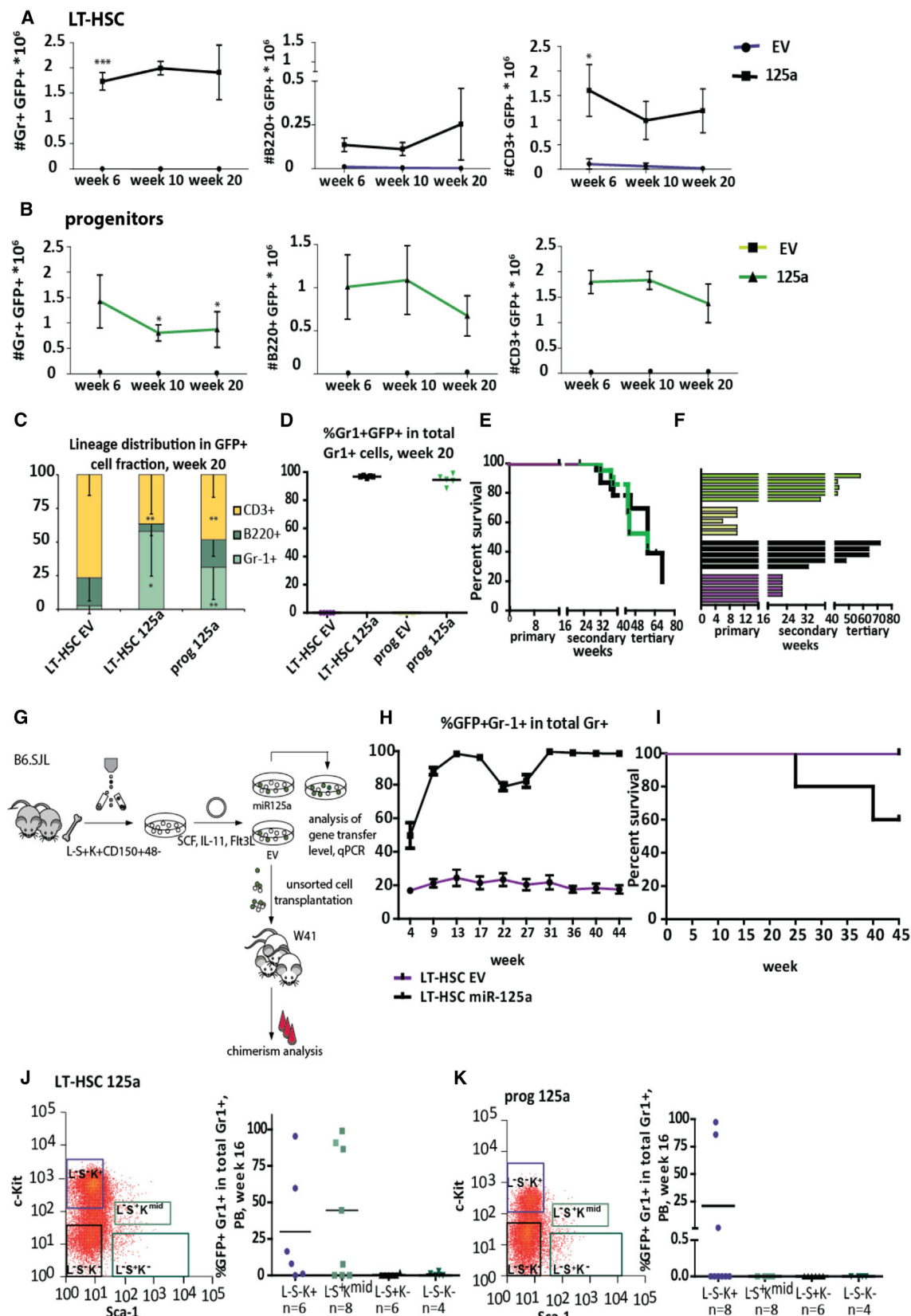## SILAC Reveals a miR-125a-Specific Network of Targets

To gain more insight in the molecular network of miR-125a targets, we performed stable isotope labeling by amino acids in cell culture (SILAC) proteomics in the 32D murine hematopoietic progenitor cell line (Ong et al., 2002; Ward et al., 1999). To do this, we expressed either miR-125a, miR-125$^{mut4T \to C}$ or a control vector and cultured cells in media with isotope-labeled amino acids (Figure 5A). We detected a total of 4,789 expressed proteins in duplicate experiments when analyzing fold-change expression ratios, comparing cells expressing the control vector with those expressing miR-125a and control vector with 125$^{mut4T \to C}$ samples. As the induction of a long-term repopulating program in HSPCs was largely seed sequence dependent, we excluded proteins that were similarly affected by both miR-125a and 125$^{mut4T \to C}$. This filtered out previously reported miR-125 targets such as Bak 1, Traf 6, Klf13, and Bmf (Guo and Scadden, 2010; O'Connell et al., 2010; Ooi et al., 2010) (Figure 5B). We identified 217 proteins that were differentially expressed upon miR-125a overexpression. Of these, 11 genes contained a 3' UTR target sequence that is perfectly complementary to the seed sequence of miR-125 (Figure 5C); ten targets showed decreased protein levels upon miR-125a overexpression in our cell system and were therefore considered to be direct miR-125 targets. These targets include p38 kinase, a negative HSC self-renewal regulator (Ito et al., 2006), and its inhibition leads to ex vivo HSC expansion (Wang et al., 2011), three protein tyrosine phosphatases (Ptpn1, Ptpn7, and Smek1, all known to suppress cytokine signaling and negatively regulate MAP kinases), a ubiquitin ligase (Syvn1, shown to block differentiation of embryonic stem cells; Yagishita et al., 2005), and Vps4b and M6pr, which are involved in transport of protein to vesicles and their lysosomal degradation. In addition, we identified interleukin-16 (IL-16) and Txnrd1, the latter gene playing a role in protection against oxidative stress (Peng et al., 2014).

To gain further insight into the mechanism whereby miR-125a was exerting its biological effects on human HSPCs, we performed label-free protein mass spectrometry on 100,000 sorted mO$^+$CD34$^+$ human cord blood (huCB) HSPCs transduced with miR-125a or control vector. We were able to reliably quantify 6,687 protein groups (Figure S5D), and 517 of them showed a significant change (p < 0.05) between 125OE and control samples (Table S3). Gene set enrichment analysis (GSEA) of the proteomics dataset identified pathways and leading edge genes directly targeted by miR-125a. In post-analysis, predicted

---

Figure 3. miR-125a Overexpression in Mouse LT-HSC and Progenitors Increases the Self-Renewal Potential

(A) Experimental setup.

(B) Chimerism levels in Gr-1$^+$ cells in primary recipients transplanted with LT-HSCs transduced with miR-125$^{mut4T \to C}$ (n = 7), LT-HSCs with control vector (n = 11), or LT-HSCs with miR-125a (n = 16). Panels also show the absolute cell numbers of B220$^+$, CD3ε^+, or Gr-1$^+$ cells in 10^6 viable peripheral blood cells, assessed by FACS at indicated time points.

(C) Chimerism levels in recipients transplanted with progenitors overexpressing control vector (n = 17) or miR-125a (n = 9), measured by FACS at indicated time points.

(D) Stem cell frequencies in mice transplanted with virally transduced LT-HSCs or progenitor cell populations, measured by limiting dilution. Confidence intervals were calculated using ELDA software. The colors indicate the various experimental groups and are the same throughout all figures.

(E) The table containing the upper and lower confidence interval and estimated frequency of repopulating cells, calculated using ELDA.

(F) The number of detected barcodes in peripheral blood in Gr-1$^+$ cells at 4 or 20 weeks post-transplantation.

(G) Lineage distribution in GFP$^+$ peripheral blood cells.

(H) Progenitor (L$^-$S$^-$K$^+$) distribution Lin$^-$GFP$^+$ BM cells.

Data in (E)–(H) reflect mean values ± SD. Statistical significance was assessed using a non-parametric Mann-Whitney test, where *p < 0.05, **p < 0.01, and ***p < 0.001.

(legend on next page)

miR-125a targets were correlated with proteomic-modulated pathways (Figure 5D). The most significant pathways centered on JUN kinase activation, FAS signaling, stress-activated kinase signaling, Toll-like receptor signaling, interferon gamma signaling, and the NoRC complex, an epigenetic pathway involved in heterochromatin formation (Figures 5D and 5E). These data suggest that miR-125 exerts its effects on HSPCs by downregulating several genes within multiple pathways. Our data also confirm that P38^{MAPK14} and PTPN1 are miR-125a targets in both murine and human HSPCs.

To validate identified targets, we performed nanofluidic proteomic analysis and confirmed the suppression of p38, Ptpn1 and Ptpn7, and Vps4b in cell lines used for SILAC (Figure S6A). To assess the downregulation of miR-125a targets in human CD34$^+$ and mouse LT-HSCs and progenitors overexpressing miR-125a, we measured the expression of p38, Ptpn1, and Ptpn7, as well as Smek1 transcripts, by qPCR. We confirmed that all tested targets showed decreased expression compared to controls in both mouse and human cells (Figures 6A, 6B, and 6E). We were also able to verify the repression of these targets at the protein level by western blot or immunofluorescence in transduced human CD34$^+$ CB cells or sorted and transduced mouse LT-HSCs and progenitors (Figures 6C, 6D, and 6F). Together, these data suggest that the endowment of an HSC-like self-renewal program in hematopoietic progenitors that normally lack self-renewal capacity upon miR-125a overexpression is linked to repression of genes controlling the p38/mitogen-activated protein kinase (MAPK) pathway as well as several other pathways. Furthermore, our results reinforce the notion that miR-125a exerts similar functions in mouse and human cells.

DISCUSSION

Our study provides key insights into the powerful role that miRNAs can play in the regulation of HSC expansion/self-renewal properties. Previous work demonstrated expansion of the HSC pool upon miR-125 expression (Ooi et al., 2010). Here, we establish that miR-125a can endow repopulation potential to progenitor populations and potently increase their self-renewal capacity. Our findings highlight the possibility of generating MPP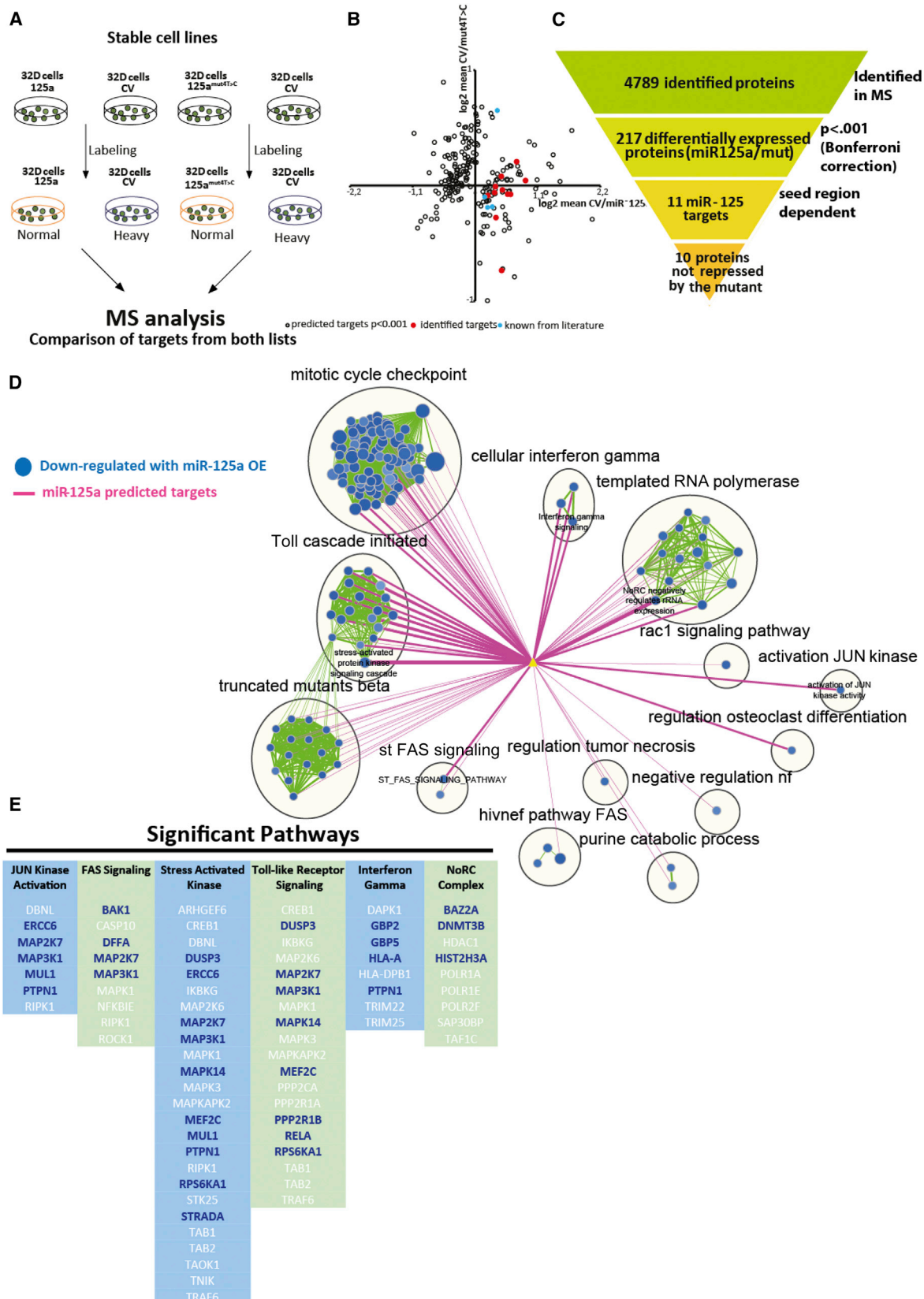s with enhanced self-renewal to augment limited sources of HLA-matched CB-derived HSCs to improve transplant outcomes. As CB-transplant-related mortality is influenced by low numbers of infused HSCs, rapid in vitro expansion of CB progenitors could increase the transplanted dose of stem cells, thus accelerating engraftment and reducing complications. Although it may be possible to therapeutically employ miR-125a to confer MPPs with self-renewal capacity, it is more likely that small-molecule regulation of several critical miR-125a targets will be required to fully translate our work. In the future, combinatorial approaches combining miR-125a modified MPPs with existing small-molecule compounds like PGE2, SR1, and UM171 (Boitano et al., 2010; Fares et al., 2014; North et al., 2007) may support expansion of undersized HLA-matched CB units, allowing them to become a selected source of cells for transplantation.

The development of myeloproliferative disease/leukemia upon miR-125a overexpression in murine cells in secondary and tertiary recipients is in agreement with previous reports (Bousquet et al., 2010; Gerrits et al., 2012; O'Connell et al., 2010) and has been shown to be dependent upon sustained expression of miR-125a (Guo et al., 2012) and dosage (O'Connell et al., 2008, 2010). Our data suggest the miR-125a-induced myeloproliferative disease occurs, in part, as a function of replicative stress. We surmise that with careful titration of miR-125 levels, it may be possible to exploit the beneficial self-renewal effects of miR-125 without incurring the malignant side effects. Interestingly, overt malignancy was never observed in human xenografts through three serial transplantations (30 weeks total), suggesting that the inherent susceptibility of murine cells to oncogenic transformation may contribute to the development of (pre-) malignancy (Balmain and Harris, 2000; Rangarajan et al., 2004). We cannot completely discount the possibility that vector integration contributed to malignant transformation. However, we deem this unlikely, considering the reproducibility of the acquired phenotypes and the absence of any known strong oncogenes in insertion site analysis. In addition, our data on cellular barcoding confirm that mice transplanted with LT-HSCs or progenitors overexpressing miR-125a had polyclonal repopulation similar to control LT-HSC, excluding the possibility that engraftment results from clonal dominance of a single clone.
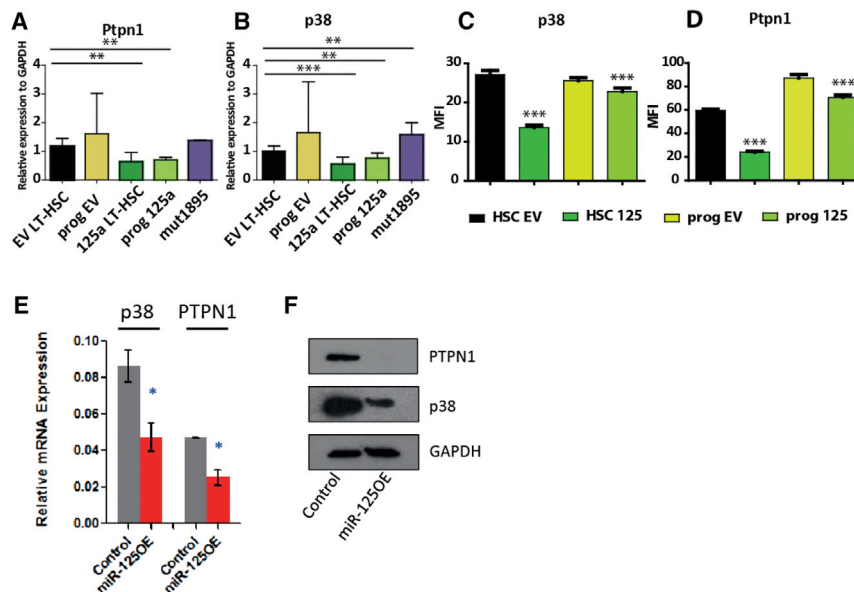
We found similar behavior between control HSCs and miR-125a-transduced MPPs regarding their long-term multi-lineage

Figure 4. Enforced miR-125a Ectopic Expression Sustains Repopulation Potential to MPPs in Serial Transplantation

(A) The number of donor-derived GFP$^+$ B220$^+$, CD3ε$^+$, or Gr-1$^+$ cells per 10^6 peripheral blood cells at indicated time points in secondary recipients transplanted with LT-HSCs transduced with control vector (n = 5) or miR-125a (n = 5).

(B) The number of donor-derived GFP$^+$ B220$^+$, CD3ε$^+$, or Gr-1$^+$ cells per 10^6 peripheral blood cells at indicated time points in secondary recipients transplanted with progenitors transduced with control vector (n = 3) or miR-125a (n = 5). Data reflect mean values ± SEM. Statistical significance assessed using non-parametric Mann-Whitney test in case of Gr-1$^+$ cells.

(C) Peripheral blood lineage distribution of GFP$^+$ cells in secondary recipients.

(D) Peripheral blood Gr-1$^+$ chimerism levels in secondary recipients 20 weeks post-transplantation.

(E) The lifespan of transduced cells (plotted mice were considered positive if the percentage of Gr-1$^+$ was >1%; brakes indicate serial transplantations).

(F) The survival statistics of primary, secondary, and tertiary recipients transplanted with LT-HSCs or progenitors overexpressing miR-125a. Interruptions of the x axis indicate time points of serial transplantations.

(G) Representative FACS plot of the Lin$^-$Sca-1$^+$ c-kit$^+$ compartment in BM cells from primary recipients transplanted with LT-HSCs overexpressing miR-125a. Right panel shows repopulation potential of indicated cell population assessed in secondary recipients. Donor cells were obtained from three mice.

(H) Same as (G), but for progenitors overexpressing miR-125a.

(I) Experimental setup for the second model, non-conditioned W41 mice used as recipients.

(J) The kinetics of chimerism changes in Gr-1$^+$ cells in the course of experiment.

(K) The survival curve of non-irradiated W41 recipients transplanted with control or miR-125a-overexpressing LT-HSCs.

All data reflect mean values ± SD, except for (K), which shows mean values ± SEM. Statistical significance was assessed using a non-parametric Mann-Whitney test, where *p < 0.05, **p < 0.01, and ***p < 0.001.

(legend on next page)

Figure 6. Target Validation

(A–D) Validation of selected targets in mouse transduced LT-HSC or progenitors at the transcript level (A and B) and protein level (C and D).

(E and F) Repression of miR-125a targets in human CD34$^+$ cells at the transcript level (E) and protein level (F).

All data reflect mean values ± SEM. Statistical significance was assessed using a non-parametric Mann-Whitney test, where *p < 0.05, **p < 0.01, and ***p < 0.001.

proliferative disorders that progressed to leukemia (Helgason et al., 1998; Zhang et al., 2006). We speculate that PTPN1 downregulation may similarly activate AKT in HSPCs, resulting in myeloid skewing and a block in B cell differentiation. However, the exact direct and indirect contribution of each individual miR-125a target in moderating self-renewal activity within progenitors requires additional studies. Multiple targets within several pathways are probably required to fully recapitulate the effects of miR125a on MPPs.

Irrespective of the precise mechanism of action, our findings provide clear evidence that miR-125a confers hematopoietic stem cell potential to cells that are normally devoid of this activity. Our results open the prospect of searching for other regulators that have this capability laying the foundation for utilizing progenitors in clinically relevant therapeutics. Molecular identification and functional evaluation of the full complement of miR-125a targets may allow the dissection of targets inhibiting self-renewal and distinguishing them from targets that may be closely aligned with malignancy. Furthermore, small-molecule screening and candidate identification on hematopoietic progenitor populations using induction of miR-125a as an operational readout could circumvent the requirement for virally enforced miRNA expression and accelerate the clinical application of miR-125a for CB transplantation.

repopulating ability, self-renewal potential, and lineage distribution in transplanted recipients, both in mouse and human systems. This similarity points to a common mechanism whereby miR-125a functions to govern stem cell properties. Using proteomic analysis, we uncovered a subset of the relevant miR-125a-targeted pathways. Our study predicts that several of our identified targets should normally act to repress self-renewal in cells downstream of HSCs. Of the ten identified targets in murine cells, four are involved in the MAPK signaling pathway. The best-known and most-studied candidate protein is p38. Chemical inhibitors of p38 activity have previously shown efficacy in ex vivo mouse and human HSC expansion protocols (Baudet et al., 2012; Wang et al., 2011). Another candidate target is PTPN1, of which little is known in the hematopoietic system. Loss of PTPN1 leads to increased numbers of pro- and pre-B cells in the BM (Dubé et al., 2005), similar to previous reports showing increased numbers of pro-B cells upon miR-125b overexpression (Chaudhuri et al., 2012) and correlating with our human CB xenotransplantation data. Moreover, deletion of PTPN1 in fibroblasts leads to suppression of Erk signaling (Dubé et al., 2004) and increased insulin growth factor-1 (IGF-1)-induced AKT protein kinase B activity (AKT/PKB) (Buckley et al., 2002). Constitutive activation of AKT in HSPCs has been shown to deplete B lymphocytes and induce myeloid skewing and the development of myelo-

EXPERIMENTAL PROCEDURES

Cell Sorting

Human Lin$^-$ CB cells were thawed and stained with CD34-APC-Cy7 and CD38-PE-Cy7 (BD). For HSC, MPP, and MLP sorts, Lin$^-$ CB cells were stained

---

Figure 5. Molecular Mechanism of Sustained Self-Renewal Potential in Human and Mouse MPPs

(A) SILAC experimental setup. 32D cells were transduced with control vector, miR-125a, or 125$^{mut4T \rightarrow C}$. Cells were cultured in medium containing normal (miR-125a OE or miR-125$^{mut4T \rightarrow C}$ samples) or heavy labeled amino acids (empty vector [EV]), as indicated.

(B) Graphical representation of 217 differentially expressed proteins. Blue symbols indicate previously reported targets of the miR-125 family. Red symbols indicate the ten differentially expressed seed-sequence-containing proteins in miR-125a and 125$^{mut4T \rightarrow C}$ samples.

(C) Overview of filtering criteria.

(D) Functional enrichment map for protein MS-based expression revealing miR-125a modulated pathways in human CD34$^+$ CB cells. Only the subset of pathways that were found to have a significant overlap with miR-125a targets is shown. Blue nodes (circles) represent gene sets enriched in proteins downregulated in CD34$^+$ human cord blood cells overexpressing miR-125a. Green line (edge) width between nodes corresponds to the number of shared proteins. Predicted miR-125a targets (yellow triangle) are connected to enriched pathways by pink edges and edge width is proportional to the overlap significance (Mann-Whitney [lesser] proteomics p = < 0.01).

(E) Top six pathways significantly downregulated by miR-125a expression. Columns represent individual pathways, with the protein names listed vertically underneath. Protein names in blue are predicted miR-125a targets.

with CD34, CD38, CD49f, CD45RA, CD90, CD7, and CD10. Murine BM cells were isolated by crushing and then stained with a cocktail of antibodies against lineage markers (Ter119, Gr-1, Mac-1, CD3, and B220), c-Kit, Sca-1, CD48, and CD150 and sorted for LT-HSCs (Lin$^-$Sca-1$^+$c-Kit$^+$CD150$^+$48$^-$) and progenitors (Lin$^-$Sca-1$^+$c-Kit$^+$ depleted from CD150$^+$48$^-$ cells).

Lentiviral and Retroviral Constructs and Transduction

For the human part of the study, the lentiviral vector for ectopic miRNA expression (gain of function) has been described previously (Gentner et al., 2010). In the murine part of the study, we used a retroviral vector, miR-125a, and mut $^{4T \rightarrow C}$ cells were cloned as previously described (Gerrits et al., 2012; Wojtowicz et al., 2014).

NSG Repopulation Assay

Male NSG mice (NOD.Cg-*Prkdc*scid*Il2rg*tm1Wjl/SzJ; Jackson Laboratory) were sublethally irradiated (225 cGy) 24 hr before intrafemoral injection. Competitive transplantation was performed using the equivalent of 400 HSCs, 400 MPPs, and 1,000 MLP cells after overnight culture in low-cytokine conditions. Limiting dilution analysis of human CD34$^+$/CD38$^-$ CB cell was performed a secondary mouse was scored as positive if it had >0.5% hCD45$^+$, mOrange$^+$ BM engraftment 8 weeks post transplantation. HSC frequency was estimated using ELDA software (http://bioinf.wehi.edu.au/software/elda/; Hu and Smyth, 2009).

C57BL/6 (B6) mice were purchased from Harlan Laboratories. C57BL/6.SJL mice were bred at the Central Animal Facility of University Medical Centre of Groningen (Groningen, the Netherlands). All animal experiments were approved by the University of Groningen Animal Care Committee.

Lentiviral Insertion Site Analysis by Modified Splinkerette PCR

To determine the lentiviral integration sites we used a splinkerette PCR approach (Uren et al., 2009) that we modified to be non-restrictive and have increased sensitivity for low DNA input amounts (see Supplemental Experimental Procedures). For identified integration sites in MPP-engrafted mice, see Table S1.

Cellular Barcoding

Cells were isolated and transduced as described above (see also Verovskaya et al., 2013). We transplanted 15,760 GFP$^+$ progenitor cells per mouse and 4,300 GFP$^+$ LT-HSCs overexpressing miR-125a, 39,800 GFP$^+$ progenitors, or 16,200 GFP$^+$ LT-HSCs for controls.

Hematopoietic Cell Transplantation and Blood Analysis

For limiting dilution experiments, cells were injected in given doses (Figure S3A) into lethally irradiated (9 Gy) female B6 recipients together with 2 × 10^6 fresh BM B6 cells. Blood samples were taken every 4–6 weeks and stained with antibodies against CD45.1 and CD45.2 to determine the donor chimerism and CD3ϵ, Gr-1, and B220 for FACS analysis. HSC frequency was estimated using ELDA software (http://bioinf.wehi.edu.au/software/elda/).

Serial BM Transplantation

Primary recipients were sacrificed at 16 weeks post-transplantation for BM collection prior to secondary transplantation. 10^7 whole BM cells were transplanted into lethally irradiated secondary recipients. For tertiary recipients, donors were sacrificed 24 weeks post-secondary transplantation, and 5 × 10^6 whole BM cells were injected together with 5 × 10^5 fresh competitor cells from B6.

qPCR

RNA was extracted from 25,000 to 30,000 GFP$^+$ cells using the miRNeasy Kit (QIAGEN), and commercially available TaqMan probes for miRNA qPCR or primers (Table S4) were used according to the manufacturer's protocol.

Label-free Quantitative Proteomics Analysis

The raw files were analyzed using MaxQuant version 1.5.2.8 (Cox and Mann, 2008) and standard settings. For differentially expressed proteins, see Table S3. For sample preparation details, see Supplemental Experimental Procedures.

Quantitative Mass Spectrometry: SILAC

The murine 32D cell line was used for to identify miR-125a targets. After transduction, GFP$^+$ cells were FACS sorted and expanded prior to stable isotope labeling as previously reported (Meenhuis et al., 2011). A nano-high performance liquid chromatography (HPLC) system consisting of Easy-nLC 1000 and Q-Exactive (Thermo Fisher Scientific) was used.

Statistical Analysis

Unless otherwise stated, mean ± SEM values are shown in the graphs. For pairwise comparison, a Mann-Whitney U test was applied. Statistical analysis was performed using PRISM 6 (GraphPad Software). For proteomics data analysis, Student's t test with Bonferroni correction was applied.

Western Blot Analysis

For immunoblotting, proteins were transferred to PVDF membranes, incubated with the specific antibody (anti-p38, Cell Signaling #9212 1:1,000; anti-PTPN1, LifeSpan BioSciences #LS-C61924 1:500; and mouse anti-GAPDH, 1:10,000, Sigma) followed by peroxidase-conjugated secondary antibodies. Bands were visualized on Amersham Hyperfilm (GE Healthcare).

Nanofluidic Proteomic Analysis

The expression levels of p38 in transduced GFP$^+$ 32D cells used for SILAC or in LT-HSCs and progenitors overexpressing miR-125a were investigated by a size-based Simple western immunoassay, using a WES device (Protein Simple).

Immunofluorescence

To measure protein expression levels in transduced GFP+ LT-HSC or progenitors (Lin-Sca-1+c-Kit+ depleted from CD150$^+$ 48- cells), cells were stained with antibodies directed against p38 (1:50, Cell Signaling, #9212) and Ptpn1 (1:100, Millipore, clone# ABS40). We used rabbit-specific IgG conjugated with Alexa-633 (Molecular Probes, Invitrogen) as a secondary antibody. For nuclear staining we treated cells with DAPI (Sigma).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2016.06.008.

AUTHOR CONTRIBUTIONS

E.E.W., K.G.H., E.R.L., L.V.B., J.E.D., and G.d.H. designed research; E.E.W., M.J.C.B., E.W., and M.R. performed mouse experiments; K.G.H., E.R.L., A.T.-G., S.M.D., G.K., J.E., J.K., E.W., and O.I.G. performed human experiments; P.A.v.V., G.M.C.J., M.F.A., H.W.J.d.L., S.J.E., G.K., R.I., E.M.S., and G.D.B. performed proteomics and analyzed MS data; E.E.W., E.R.L., and L.V.B. analyzed and interpreted data; and E.E.W., E.R.L., L.V.B, J.E.D., and G.d.H. wrote the manuscript.

ACKNOWLEDGMENTS

CellPress

REFERENCES

Aljurf, M., Rizzo, J.D., Mohty, M., Hussain, F., Madrigal, A., Pasquini, M.C., Passweg, J., Chaudhri, N., Ghavamzadeh, A., Solh, H.E., et al. (2014). Challenges and opportunities for HSCT outcome registries: perspective from international HSCT registries experts. Bone Marrow Transplant. *49*, 1016–1021.

Ballen, K.K., Gluckman, E., and Broxmeyer, H.E. (2013). Umbilical cord blood transplantation: the first 25 years and beyond. Blood *122*, 491–498.

Balmain, A., and Harris, C.C. (2000). Carcinogenesis in mouse and human cells: parallels and paradoxes. Carcinogenesis *21*, 371–377.

Barker, J.N., Weisdorf, D.J., and Wagner, J.E. (2001). Creation of a double chimera after the transplantation of umbilical-cord blood from two partially matched unrelated donors. N. Engl. J. Med. *344*, 1870–1871.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215–233.

Baudet, A., Karlsson, C., Safaee Talkhoncheh, M., Galeev, R., Magnusson, M., and Larsson, J. (2012). RNAi screen identifies MAPK14 as a druggable suppressor of human hematopoietic stem cell expansion. Blood *119*, 6255–6258.

Boitano, A.E., Wang, J., Romeo, R., Bouchez, L.C., Parker, A.E., Sutton, S.E., Walker, J.R., Flaveny, C.A., Perdew, G.H., Denison, M.S., et al. (2010). Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. Science *329*, 1345–1348.

Bousquet, M., Harris, M.H., Zhou, B., and Lodish, H.F. (2010). MicroRNA miR-125b causes leukemia. Proc. Natl. Acad. Sci. USA *107*, 21558–21563.

Buckley, D.A., Cheng, A., Kiely, P.A., Tremblay, M.L., and O'Connor, R. (2002). Regulation of insulin-like growth factor type I (IGF-I) receptor kinase activity by protein tyrosine phosphatase 1B (PTP-1B) and enhanced IGF-I-mediated suppression of apoptosis and motility in PTP-1B-deficient fibroblasts. Mol. Cell. Biol. *22*, 1998–2010.

Chaudhuri, A.A., So, A.Y., Mehta, A., Minisandram, A., Sinha, N., Jonsson, V.D., Rao, D.S., O'Connell, R.M., and Baltimore, D. (2012). Oncomir miR-125b regulates hematopoiesis by targeting the gene Lin28A. Proc. Natl. Acad. Sci. USA *109*, 4233–4238.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Doulatov, S., Notta, F., Eppert, K., Nguyen, L.T., Ohashi, P.S., and Dick, J.E. (2010). Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. Nat. Immunol. *11*, 585–593.

Dubé, N., Cheng, A., and Tremblay, M.L. (2004). The role of protein tyrosine phosphatase 1B in Ras signaling. Proc. Natl. Acad. Sci. USA *101*, 1834–1839.

Dubé, N., Bourdeau, A., Heinonen, K.M., Cheng, A., Loy, A.L., and Tremblay, M.L. (2005). Genetic ablation of protein tyrosine phosphatase 1B accelerates lymphomagenesis of p53-null mice through the regulation of B-cell development. Cancer Res. *65*, 10088–10095.

Fares, I., Chagraoui, J., Gareau, Y., Gingras, S., Ruel, R., Mayotte, N., Csaszar, E., Knapp, D.J., Miller, P., Ngom, M., et al. (2014). Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. Science *345*, 1509–1512.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat. Rev. Genet. *9*, 102–114.

Gentner, B., Visigalli, I., Hiramatsu, H., Lechman, E., Ungari, S., Giustacchini, A., Schira, G., Amendola, M., Quattrini, A., Martino, S., et al. (2010). Identification of hematopoietic stem cell-specific miRNAs enables gene therapy of globoid cell leukodystrophy. Sci. Transl. Med. *2*, 58ra84.

Gerrits, A., Walasek, M.A., Olthof, S., Weersing, E., Ritsema, M., Zwart, E., van Os, R., Bystrykh, L.V., and de Haan, G. (2012). Genetic screen identifies microRNA cluster 99b/let-7e/125a as a regulator of primitive hematopoietic cells. Blood *119*, 377–387.

Guo, S., and Scadden, D.T. (2010). A microRNA regulating adult hematopoietic stem cells. Cell Cycle *9*, 3637–3638.

Guo, S., Bai, H., Megyola, C.M., Halene, S., Krause, D.S., Scadden, D.T., and Lu, J. (2012). Complex oncogene dependence in microRNA-125a-induced myeloproliferative neoplasms. Proc. Natl. Acad. Sci. USA *109*, 16636–16641.

Helgason, C.D., Damen, J.E., Rosten, P., Grewal, R., Sorensen, P., Chappel, S.M., Borowski, A., Jirik, F., Krystal, G., and Humphries, R.K. (1998). Targeted disruption of SHIP leads to hemopoietic perturbations, lung pathology, and a shortened life span. Genes Dev. *12*, 1610–1620.

Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J Immunol. Methods. *347*, 70–78.

Ito, K., Hirao, A., Arai, F., Takubo, K., Matsuoka, S., Miyamoto, K., Ohmura, M., Naka, K., Hosokawa, K., Ikeda, Y., and Suda, T. (2006). Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. Nat. Med. *12*, 446–451.

Lechman, E.R., Gentner, B., van Galen, P., Giustacchini, A., Saini, M., Boccalatte, F.E., Hiramatsu, H., Restuccia, U., Bachi, A., Voisin, V., et al. (2012). Attenuation of miR-126 activity expands HSC in vivo without exhaustion. Cell Stem Cell *11*, 799–811.

Meenhuis, A., van Veelen, P.A., de Looper, H., van Boxtel, N., van den Berge, I.J., Sun, S.M., Taskesen, E., Stern, P., de Ru, A.H., van Adrichem, A.J., et al. (2011). MiR-17/20/93/106 promote hematopoietic cell expansion by targeting sequestosome 1-regulated pathways in mice. Blood *118*, 916–925.

Metcalf, D. (2008). Hematopoietic cytokines. Blood *111*, 485–491.

North, T.E., Goessling, W., Walkley, C.R., Lengerke, C., Kopani, K.R., Lord, A.M., Weber, G.J., Bowman, T.V., Jang, I.H., Grosser, T., et al. (2007). Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. Nature *447*, 1007–1011.

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.

O'Connell, R.M., Rao, D.S., Chaudhuri, A.A., Boldin, M.P., Taganov, K.D., Nicoll, J., Paquette, R.L., and Baltimore, D. (2008). Sustained expression of microRNA-155 in hematopoietic stem cells causes a myeloproliferative disorder. J. Exp. Med. *205*, 585–594.

O'Connell, R.M., Chaudhuri, A.A., Rao, D.S., Gibson, W.S., Balazs, A.B., and Baltimore, D. (2010). MicroRNAs enriched in hematopoietic stem cells differentially regulate long-term hematopoietic output. Proc. Natl. Acad. Sci. USA *107*, 14235–14240.

Ogawa, M. (1993). Differentiation and proliferation of hematopoietic stem cells. Blood *81*, 2844–2853.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol. Cell. Proteomics *1*, 376–386.

Ooi, A.G., Sahoo, D., Adorno, M., Wang, Y., Weissman, I.L., and Park, C.Y. (2010). MicroRNA-125b expands hematopoietic stem cells and enriches for the lymphoid-balanced and lymphoid-biased subsets. Proc. Natl. Acad. Sci. USA *107*, 21505–21510.

Oran, B., and Shpall, E. (2012). Umbilical cord blood transplantation: a maturing technology. Hematology (Am Soc Hematol Educ Program) *2012*, 215–222.

Peng, X., Mandal, P.K., Kaminskyy, V.O., Lindqvist, A., Conrad, M., and Arnér, E.S. (2014). Sec-containing TrxR1 is essential for self-sufficiency of cells by control of glucose-derived H2O2. Cell Death Dis. *5*, e1235.

Pineault, N., and Abu-Khader, A. (2015). Advances in umbilical cord blood stem cell expansion and clinical translation. Exp. Hematol. 43, 498–513.

Rangarajan, A., Hong, S.J., Gifford, A., and Weinberg, R.A. (2004). Species- and cell type-specific requirements for cellular transformation. Cancer Cell 6, 171–183.

Sauvageau, G., Iscove, N.N., and Humphries, R.K. (2004). In vitro and in vivo expansion of hematopoietic stem cells. Oncogene 23, 7223–7232.

Uren, A.G., Mikkers, H., Kool, J., van der Weyden, L., Lund, A.H., Wilson, C.H., Rance, R., Jonkers, J., van Lohuizen, M., Berns, A., and Adams, D.J. (2009). A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nat. Protoc. 4, 789–798.

Verovskaya, E., Broekhuis, M.J., Zwart, E., Ritsema, M., van Os, R., de Haan, G., and Bystrykh, L.V. (2013). Heterogeneity of young and aged murine hematopoietic stem cells revealed by quantitative clonal analysis using cellular barcoding. Blood 122, 523–532.

Wagner, J.E., Jr., Eapen, M., Carter, S., Wang, Y., Schultz, K.R., Wall, D.A., Bunin, N., Delaney, C., Haut, P., Margolis, D., et al.; Blood and Marrow Transplant Clinical Trials Network (2014). One-unit versus two-unit cord-blood transplantation for hematologic cancers. N. Engl. J. Med. 371, 1685–1694.

Wang, Y., Kellner, J., Liu, L., and Zhou, D. (2011). Inhibition of p38 mitogen-activated protein kinase promotes ex vivo hematopoietic stem cell expansion. Stem Cells Dev. 20, 1143–1152.

Ward, A.C., Smith, L., de Koning, J.P., van Aesch, Y., and Touw, I.P. (1999). Multiple signals mediate proliferation, differentiation, and survival from the granulocyte colony-stimulating factor receptor in myeloid 32D cells. J. Biol. Chem. 274, 14956–14962.

Wojtowicz, E.E., Walasek, M.A., Broekhuis, M.J., Weersing, E., Ritsema, M., Ausema, A., Bystrykh, L.V., and de Haan, G. (2014). MicroRNA-125 family members exert a similar role in the regulation of murine hematopoiesis. Exp. Hematol. 42, 909-18.e1.

Yagishita, N., Ohneda, K., Amano, T., Yamasaki, S., Sugiura, A., Tsuchimochi, K., Shin, H., Kawahara, K., Ohneda, O., Ohta, T., et al. (2005). Essential role of synoviolin in embryogenesis. J. Biol. Chem. 280, 7909–7916.

Zhang, J., Grindley, J.C., Yin, T., Jayasinghe, S., He, X.C., Ross, J.T., Haug, J.S., Rupp, D., Porter-Westpfahl, K.S., Wiedemann, L.M., et al. (2006). PTEN maintains haematopoietic stem cells and acts in lineage choice and leukaemia prevention. Nature 441, 518–522.

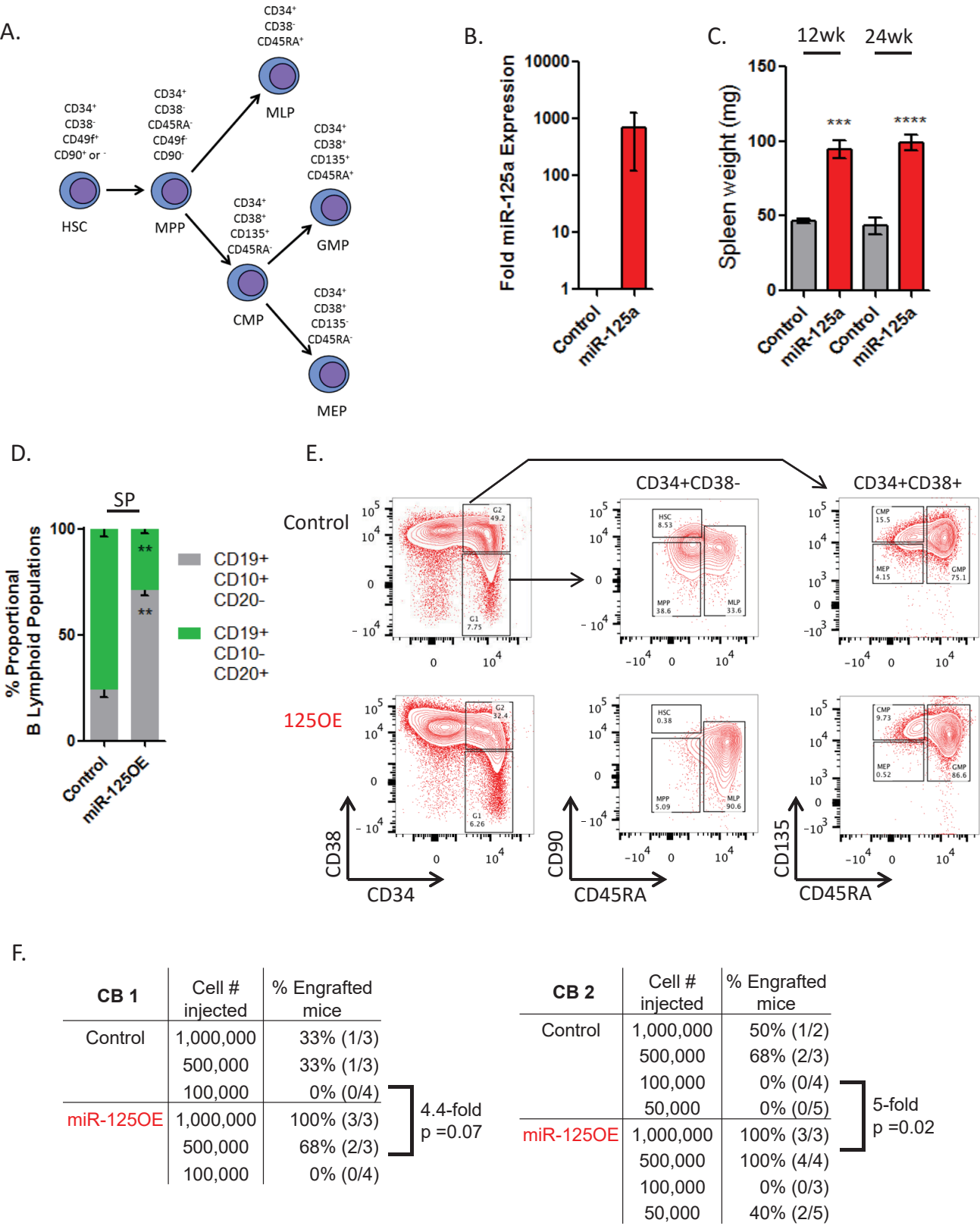**Supplemental Information**

Ectopic miR-125a Expression Induces

Long-Term Repopulating Stem Cell Capacity

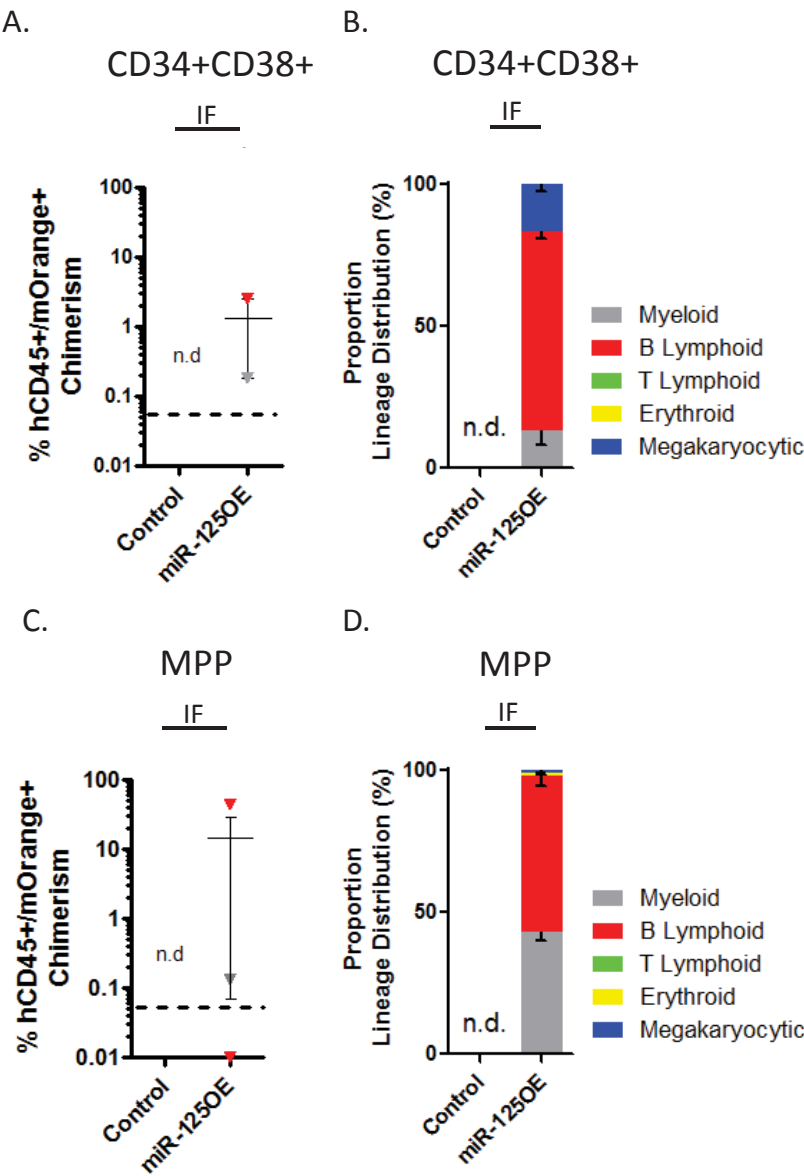**in Mouse and Human Hematopoietic Progenitors**

Edyta E. Wojtowicz, Eric R. Lechman, Karin G. Hermans, Erwin M. Schoof, Erno Wienholds, Ruth Isserlin, Peter A. van Veelen, Mathilde J.C. Broekhuis, George M.C. Janssen, Aaron Trotman-Grant, Stephanie M. Dobson, Gabriela Krivdova, Jantje Elzinga, James Kennedy, Olga I. Gan, Ankit Sinha, Vladimir Ignatchenko, Thomas Kislinger, Bertien Dethmers-Ausema, Ellen Weersing, Mir Farshid Alemdehy, Hans W.J. de Looper, Gary D. Bader, Martha Ritsema, Stefan J. Erkeland, Leonid V. Bystrykh, John E. Dick, and Gerald de Haan
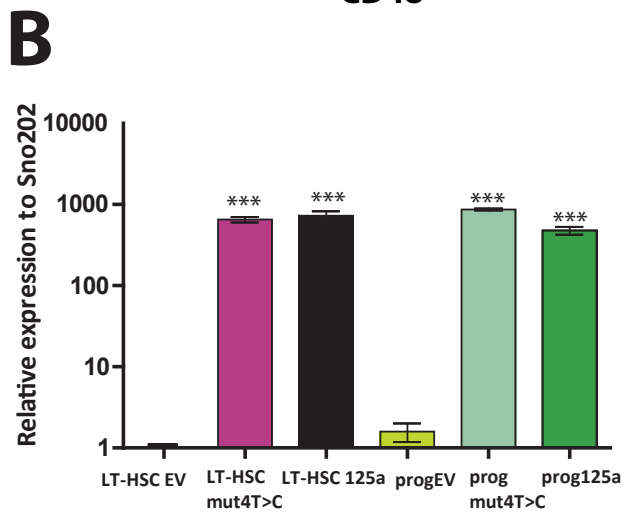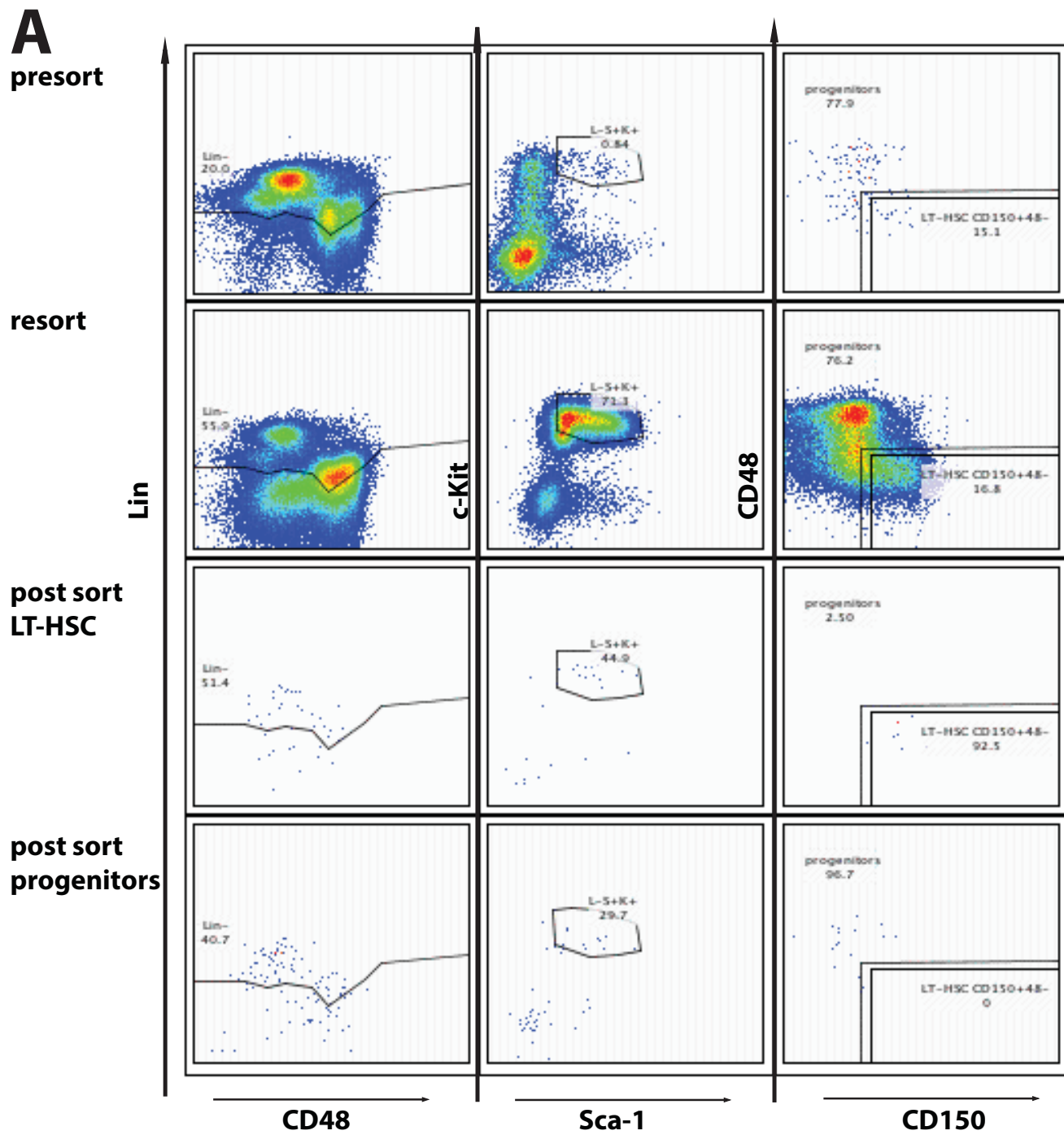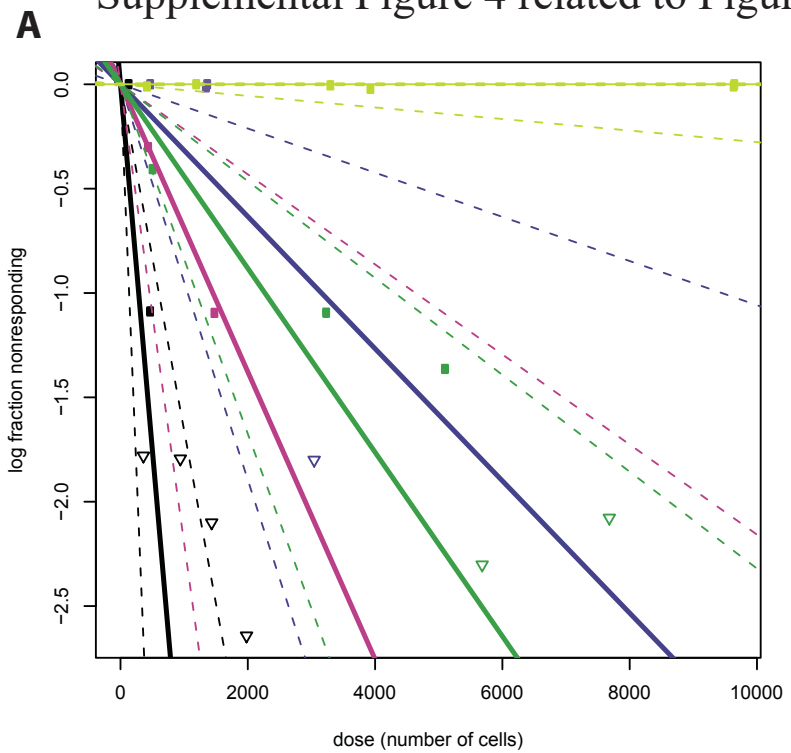
Supplemental Figure 1, related to Figure 1

A.

B.

C.

D.

E.

F.

CB 1	Cell # injected	% Engrafted mice
Control	1,000,000	33% (1/3)
	500,000	33% (1/3)
	100,000	0% (0/4)
miR-125OE	1,000,000	100% (3/3)
	500,000	68% (2/3)
	100,000	0% (0/4)

4.4-fold
p =0.07

CB 2	Cell # injected	% Engrafted mice
Control	1,000,000	50% (1/2)
	500,000	68% (2/3)
	100,000	0% (0/4)
	50,000	0% (0/5)
miR-125OE	1,000,000	100% (3/3)
	500,000	100% (4/4)
	100,000	0% (0/3)
	50,000	40% (2/5)

5-fold
p =0.02

Supplemental Figure 2 related to Figure 2

A.

CD34+CD38+

IF

B.

CD34+CD38+

IF

- Myeloid
- B Lymphoid
- T Lymphoid
- Erythroid
- Megakaryocytic

C.

MPP

IF

D.

MPP

IF

- Myeloid
- B Lymphoid
- T Lymphoid
- Erythroid
- Megakaryocytic

Supplemental Figure 3 related to Figure 3

Supplemental Figure 4 related to Figure 3

A

Group LT−HSC125a
Group LT−HSCEV
Group LT−HSCmut4T>C
Group prog125a
Group progEV

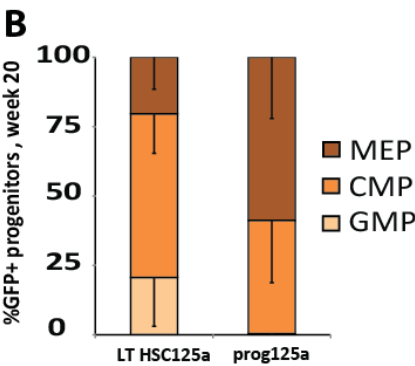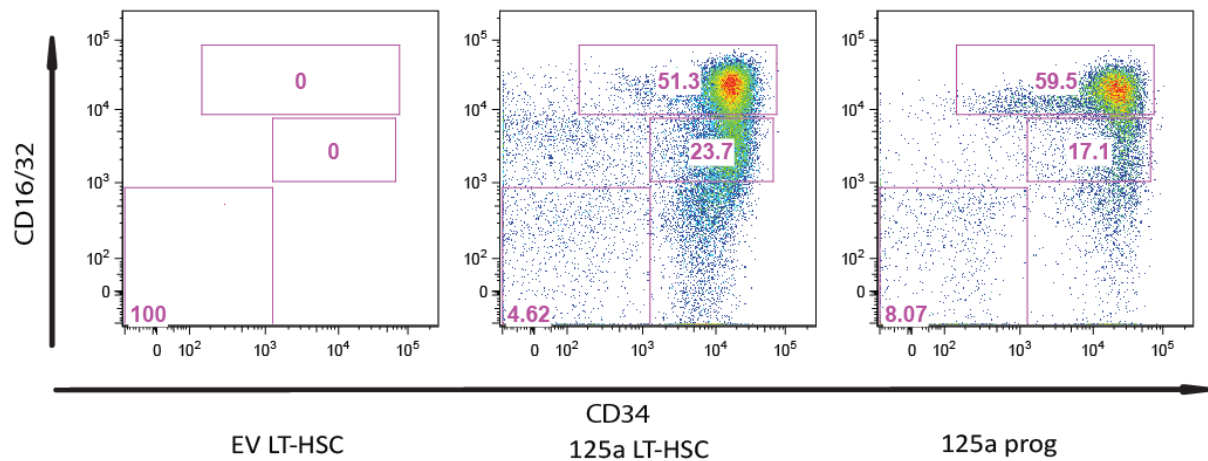| Counter	Dose	Tested	Response	Group
1	531	3	1	prog125a
2	5125	4	3	prog125a
3	7680	4	4	prog125a
4	5682	5	5	prog125a
5	3259	3	2	prog125a
6	1369	1	0	LT-HSCEV
7	494	3	0	LT-HSCEV
8	1390	2	0	LT-HSCEV
9	3043	3	3	LT-HSCEV
10	155	2	0	LT-HSC125a
11	493	3	2	LT-HSC125a
12	360	3	3	LT-HSC125a
13	1432	4	4	LT-HSC125a
14	1980	7	7	LT-HSC125a
15	940	3	3	LT-HSC125a
16	1217	3	0	progEV
17	3954	4	0	progEV
18	9671	4	0	progEV
19	9660	4	0	progEV
20	443	3	0	progEV
21	3325	3	0	progEV
22	462	4	1	LT-HSCmut4T>C
23	1503	3	2	LT-HSCmut4T>C

Limiting dilution data entered.

B

Group name	Cell type	Transduction efficiency	Number of GFP+ cells/mouse	number of injected mice	Number of clones in Gr-1+, week 20	Positive mice in Gr-1+, week 20
A	prog CV	60,00	39840	3	0-1	2
B	HSC CV	80,00	16213	4	14-21	4
C	prog 125a	37,00	15758	4	6-20	4
D	HSC 125a	51,10	4324	5	4-26	5

Supplemental Figure 5 related to Figure 5

A Second transplantation

EV LT-HSC   125a LT-HSC   125a prog

Y-axis: CD16/32 — X-axis: CD34

B

%GFP+ progenitors, week 20

Legend: ■ MEP ■ CMP ■ GMP

X-axis: LT HSC125a, prog125a

C

cell population	#GFP+cells/recipient			
	#L-S+K-	#L-S-K+	#L-S-K-	#L-S+Kmid
LT-HSC 125a	24446	203791	333240	19071
prog 125a	31590	119999	474276	15210

Supplemental Figure 6 related to Figure 6

A

B

C

Supplemental Table 1, related to Figure 2

DNA sample	Insertion ID	Total no. PCR hits	Unique no. PCR hits	% Total	Genomic Location GRCh38	Overlapping Genes
M1	M1-01	6	5	20	1:158727516	-
	M1-02	6	6	24	12:54973789	TESPA1
	M1-03	2	2	8	12:57274635	RP11-123K3.4, R3HDM2
	M1-04	3	3	12	19:7548123	PNPLA6
	M1-05	4	4	16	19:810099	PTBP1
	M1-06	8	5	20	22:42903322	PACSIN2
		29	25	100		
	M2-01	17	10	38	12:14955890	ARHGDIB
	M2-02	3	2	8	3:17402188	TBC1D5
M2	M2-03	21	7	27	6:24809228	RP3-369A17.6, FAM65B
	M2-04	9	7	27	9:137249496	-
		50	26	100		
	M3-01	2	1	4	1:16414881	SPATA21
	M3-02	3	2	8	2:218558029	USP37
	M3-03	9	6	25	20:35087623	TRPC4AP
	M3-04	10	8	33	22:-40396240	RP5-1042K10.14, SGSM3
M3	M3-05	1	1	4	3:138900257	-
	M3-06	5	4	17	4:101193236	PPP3CA
	M3-07	1	1	4	16:14625001	PARN
	M3-08	2	1	4	9:132906944	TSC1
		33	24	100		

Table S1. Identification of Lentiviral Insertions within MPP Engrafted mice.
mOrange+ cells were flow sorted from 3 mice (M1-M3) engrafted for 20 weeks with MPP expressing miR-125a. Lentiviral insertions were identified using a modified Splinkerette PCR protocol. Starting levels of MPP lentiviral marking were ~10% (M2) and ~45% (M1 and M3). Insertion ID denotes a cloned fragment specific to a unique genomic location. Total number of PCR hits denotes the total number of cloned fragments that mapped to a unique genomic location. Unique number of PCR hits represents the total quantity of non-redundant clone fragments that map to that unique genomic location.

Supplemental Table 2   related to Figure 2

Group	Exp No.	Cell No.	Total No. wells	No. (+) wells	Frequency	Interval	Colony/LTC-IC
MPP CTRL	1	2	18	3			
	2	2	20	2			
	3	2	40	0			
	4	2	40	0			
	1	4	5	1			
	2	4	6	1			
	3	4	40	3			
	4	4	39	3			
	Total		90	8	**1 in 44.7**	26.2 - 76.6	**6.6 + 2.1**
MPP 125OE	1	2	19	4			
	2	2	18	5			
	3	2	39	3			
	4	2	40	4			
	1	4	19	5			
	2	4	18	9			
	3	4	40	16			
	4	4	40	16			
	Total		117	46	**1 in 11.9**	9.2 – 15.4	**54.4 + 11.6**
					P = 0.0000000975		**P = 0.0041**

Table S2. LT-CIC assay with miR-125OE MPP.
Either 2 or 4 mO+ MPP (transduced with empty control or miR-125OE virus) were flow sorted onto MS5 stroma and cultured for 4 weeks. After 1 month, positive wells were scored, harvested and plated into methylcellulose. Two weeks later, colonies were enumerated.

Supplemental Table 3 related to Figure 5

Identified by label-free MS, proteins differentially expressed in control or miR-125a OE

$mO^{+}CD34^{+}$ cells

Supplemental Figure Legends

Figure S1

A.) Scheme describing the human hematopoietic hierarchy. B.) PCR analysis of miR-125 expression in human lin-CB cells post-transduction with miR-125OE virus or control. C.) Spleen weights of miR-125OE and control transplanted animals 12 and 24 weeks post-transplantation. D.) Proportion of B-Lymphoid populations in the spleen of mice transplanted with miR-125OE or Control. E.) Representative flow plots describing the gating scheme for recovered human lin-CB from mice xenotransplanted with miR-125aOE cells. F.) Secondary transplantations. Table describing the numbers of transplanted/engrafted mice for each LDA experiment. Briefly, BM cells were harvested from 2 primary recipients and flow sorted for hCD45+mO+ cells. Stem cell frequencies were evaluated by limiting dilution transplantation into secondary mice. Confidence intervals were calculated using ELDA software. All data reflects mean values ± SEM. Statistical significance assessed using a non-parametric Mann-Whitney test where $p < 0.05$ *, $p < 0.01$ ** and $p<0.001$ ***.

Figure S2

A.) Human mO+ chimerism after 10 weeks tertiary transplant (30 weeks total) from $CD34^+CD38^+$ progenitors from secondary mouse recipients. B.) Lineage distribution of mO^+ bone marrow grafts from $CD34^+CD38^+$ progenitors from tertiary recipient mice. C.) Human mO^+ chimerism after 10 weeks tertiary transplant (30 weeks total) from MPP from secondary mouse recipients. D.) Lineage distribution of mO^+ bone marrow grafts from MPP tertiary recipient mice.

Figure S3

A.) Sorting strategy. Lin⁻Sca-1$^+$ c-kit$^+$ cells were first pre-sorted only for L-S+K+ (the last plot in the upper row only shows the frequency of CD150+48- cells within L-S+K+ cell population and was not included in the sort definition during the presort) and then resorted to separate Lin⁻Sca-1$^+$ c-kit$^+$CD150$^+$CD48$^-$ (defined as LT-HSC) from the Lin⁻Sca-1$^+$ c-kit$^+$ population (defined as progenitors). Lowest 2 panels show the purity of sorted populations. B.) Expression levels of miR-125a measured by qPCR.

Figure S4

A.) The table with cell doses, recipient and positive responders used for the limiting dilution with murine cells. B.) Transduction efficiencies, cell doses and number of recipients transplanted with barcoded cells.

Figure S5

A.) Representative FACS plots of the progenitor compartment in secondary recipients 20 weeks post transplantation. B) Progenitors distribution in miR-125a overexpressing LT-HSC or progenitors. C) Table indicating cell doses transplanted into secondary recipients (shown in Fig. 4J and 4K, transplanted cell doses calculated based on bone marrow FACS profiles from primary recipients transplanted with LT-HSC or progenitors overexpressing miR-125a). D) Correlation plots showing the Pearson correlation coefficients for the biological reproducibility of the proteomics analyses in flow sorted CD34$^+$ human CB cells with and without miR-125a overexpression.

Figure S6

A.) Validation of downregulation of chosen targets in 32D cell line used for SILAC as assessed by nanofluidic proteomic analysis. B.) QPCR for Ptpn7 in LT-HSC or progenitors overexpressing control, miR-125a or the seed sequence mutant. All data reflect mean values ± SEM. Statistical significance assessed using a non-parametric Mann-Whitney test where $p < 0.05$ *, $p < 0.01$ ** and $p<0.001$ ***.

Supplemental Table 1

Integration site analysis by splinkerette PCR.

Supplemental Table 2

The frequency of LTC-IC formed by control or miR-125a overexpressing human MPP.

Supplemental Table 3

This table is attached as a separate Excel spread sheet. The list of identified and quantified proteins in label-free MS, 518 proteins were differentially expressed between miR-125OE and CTRL, $p<0.05$ (2nd Sheet).

Supplemental Table 4

Sequences of primers used for qPCR in murine cells.

Supplemental Table 4 related to Figure 6

Species	Primer	Sequence 5`->3`
mouse	p38 qPCR Fwd1	CCAAGCCATGAGGCAAGAAAC
mouse	p38 qPCR Rev1	GGGTCGTGGTACTGAGCAAA
mouse	PTPN1 qPCR Fwd1	GGCTATTTACCAGGACATTCGAC
mouse	PTPN1 qPCR Rev1	TCCGACTGTGGTCAAAAGGG
mouse	PTPN7 qPCR Fwd1	GCAGACTTCTGGGAGATGGT
mouse	PTPN7 qPCR Rev1	TGCTCTTTCATGTCCTGGAT
mouse	Smek1 Fwd1	GGCATTATGTGCCCTCCGT
mouse	Smek1 Rev1	CGCATGCTATCAAGTTTGGGG

Table S4. Primers used for qPCR in murine cells.

Supplemental Materials and Methods

Cord blood samples

All cord blood (CB) samples were obtained with informed consent and procedure has been approved by the institutional review boards of the University Health Network, Trillium and Credit Valley Hospital. CB mononuclear cells (MNC) were obtained by gradient centrifugation with Lymphoprep (StemCell Technologies). Lineage-negative (Lin-) CB cells were obtained by negative selection with the StemSep human progenitor cell enrichment kit (StemCell Technologies) according to the manufacturer's protocol. Lin- cells were stored at -150°C until use.

Cell sorting

Lin- CB cells were thawed by dropwise addition of 50% FBS/50% X-VIVO 10/DNAseI (100µg/ml) and resuspended at 10^7 cells per ml. For CD34/CD38 sort lin- CB cells were stained with CD34-APC-Cy7 and CD38-PE-Cy7 (BD). For HSC, MPP and MLP sorts lin- CB cells were stained with: CD34, CD38, CD49f, CD45RA, CD90, CD7, CD10 cells were sorted on BD FACS Aria II/III (Becton Dickinson) or MoFlo (Beckman Coulter) sorters, consistently yielding >95% purity.

Bone marrow cells were isolated by crushing, and stained with a cocktail of antibodies against lineage markers (Ter119, Gr-1, Mac-1, CD3, B220), c-Kit, Sca-1, CD48 and CD150 and sorted for LT-HSC (Lin-Sca-1+c-Kit+CD150+48-) and progenitors (Lin-Sca-1+c-Kit+ depleted from CD150+48- cells) on MoFlo or Astrios (Beckman Coulter) sorters with a high purity . Cells were transduced as previously described (Verovskaya et al., 2013).

For the analysis of progenitors we have used antibodies: lineage-A700, c-Kit-PE, Sca-1-Pacific Blue (all from BioLegend), CD16.32-PE-Cy7 (eBiosicence), CD34-A647 (BD Pharmingen) and run on LSR-II (BD-Biosciences).

Lentiviral and retroviral constructs and transduction

<u>For the human part of the study the len</u>tiviral vector for ectopic miRNA expression (gain-of-function) has been described previously (Gentner et al., 2010). This vector contains the strong spleen focus forming virus (SFFV) promoter and links miR-125a overexpression with the expression of a transcriptionally linked fluorescence reporter (mOrange). A 239 bp genomic DNA fragment containing pri-miR125a was cloned into the intron of LV.SFFV.intron.mOrange. Sorted cells were plated in X-VIVO10 medium (BioWhittaker) supplemented with 1% BSA, SCF, TPO, FLT-3 and IL-6, and transduced with viral particles at a multiplicity of infection (MOI) of 30 for at least 16 hours. Subsequent expansion was performed in the same medium.

<u>In the murine part of the study we used a retroviral vector, miR</u>-125a and the mut $^{4T>C}$ were cloned as previously described (Gerrits et al., 2012; Wojtowicz et al., 2014b).

NSG repopulation assay

All animal experimental procedures were approved by the University Health Network Animal Care Committee. Male NSG mice (NOD.Cg-*Prkdc*scid*Il2rg*tm1Wjl/SzJ; Jackson laboratory) were sublethally irradiated (225 cGy) 24 hours before intrafemoral injection. Competitive transplantation was performed using the equivalent of 1 x 10^4 CD34+/CD38- cells or 3 x 10^4 CD34+/CD38+ cells after 4 days of culture in low cytokine conditions. The equivalent of 400

HSC, 400 MPP and 1000 MLP cells were transplanted after overnight culture in low cytokine conditions. Limiting dilution anaysis of human CD34+/CD38- CB cell was performed by sorting hCD45+/mOrange+ cells from pooled BM samples of primary mice 12 weeks post-transplantation and injection of different doses into sub-lethally irradiated secondary recipients. A secondary mouse was scored as positive if it had >0.5% hCD45+, mOrange+ BM engraftment 8 weeks post transplantation. HSC frequency was estimated using the ELDA software (http://bioinf.wehi.edu.au/software/elda/, Hu and Smyth, 2009). Secondary transplantation assays for MPP and CD34+/CD38+ cells were performed by taking BM cells from injected femur from primary recipient and injecting all the cells into the right femur of a sub-lethally irradiated secondary recipient.

Mice were euthanized at the indicated time points post transplantation, cells from injected femur, other bones (other femur and both tibia), and spleen were collected, by flushing the bones, for analysis. Cells were stained with human antibodies for CD45, CD41, GlyA, CD33, CD19 and CD3 and analyzed on a LSR II (BD) or Canto II (BD) cytometer. Spleen cells were also stained with CD19, CD20, CD10 and CD45. Data was analyzed with FlowJo software (Tree Star Inc.) For purification of the human lin- compartment from xenotransplanted mice, bone marrow cells from 2-3 mice were combined and processed with the StemSep mouse/human chimera enrichment kit (Stem Cell Technologies) according to the manufacturer's instruction. To deplete human lineage positive cells 50μl/ml human hematopoietic progenitor enrichment antibody cocktail from the StemSep human progenitor cell enrichment kit was added during the anti-biotin incubation step.

C57BL/6 (B6) mice were purchased from Harlan Laboratories (The Netherlands). C57BL/6.SJL mice were bred at Central Animal Facility of University Medical Centre of Groningen

(Groningen, The Netherlands). All animal experiments were approved by the University of Groningen Animal Care Committee.

Lentiviral Insertion Site Analysis by Splinkerette PCR

To determine the lentiviral integration sites we used a splinkerette PCR approach (Uren et al., 2009) that we modified to be non-restrictive and have increased sensitivity for low DNA input amounts. Genomic DNA from mOrange⁺ sorted cell populations was isolated using a Purelink Genomic DNA mini kit (Invitrogen). Approximately 30-150 ng DNA was randomly fragmented by incubation with 0.25 µl dsDNA Fragmentase in 1X Fragmentase reaction buffer (New England Biolabs) at 37°C for 30 min in a final volume of 10 µl, followed by heat-inactivation of the dsDNA Fragmentase at 72°C for 20 min. Subsequently, the fragmented DNA was end-repaired, phosphorylated and 3′-adenylated by addition of an enzyme mix containing 0.25 µl *E. coli* DNA ligase for Fragmentase (New England Biolabs), 1.25 units T4 DNA polymerase (Thermo Scientific), 1 unit T4 Polynucleotide Kinase (Thermo Scientific), 0.5 unit Taq DNA polymerase, 200 µM of each dNTP and 1X rapid ligation buffer (Thermo Scientific) in a final volume of 25 µl, and an incubation series of 11°C for 20 min, 37°C for 20 min and 72°C for 20 min.

The repaired DNA fragments were ligated to a modified splinkerette adaptor in which the long-strand oligonucleotide (SplnkEW2LZ: 5′-GTGATGTCGGCAGTATAGGACAGGATATCAGGTCTGCAGCCTATAGTGAACGAGCACTAG[*]T-3′) has a complementary 3'-overhanging thymine residue that is linked by a phosphorothioate bond and in which the short-strand oligonucleotide (SplnkEW2SP: 5′-[Phos]CTAGTGCTCGTTCACTATAGGCTTAATTTTTTTTTTCAAAAAAA-3′) is 5′

phosphorylated. In total, 10 μl ligation mix containing 20 pmol modified adaptor molecule, 2.5 units T4 DNA ligase (Thermo Scientific) and 1X rapid ligation buffer was added to the repaired DNA, which was subsequently incubated at 22°C for 3 h and 14°C for 9 h and finally inactivated by incubation at 65°C for 15 min. After ligation, the samples were diluted 2.5-fold with milli Q water.

Lentiviral LTR and splinkerette-specific ligation fragments were amplified by nested PCR. In the first round of amplification, the PCR samples contained 5 μl ligation product, 0.2 μM adaptor primer 1 (ADAP-1: 5′-GTGATGTCGGCAGTATAGGACA-3′), 0.2 μM 5′ or 3′ LTR-specific primer 1 (5LTR-1: 5′-AGAGCTCCCAGGCTCAG-3′; 3LTR-1: 5′-ATAAAGCTTGCCTTGAGTGC-3′), 200 μM of each dNTP, 25 mM Tricine, 7.0% Glycerol (m/v), 1.6% DMSO (m/v), 2 mM MgCl2, 85 mM NH4Acetate (pH8.7), and 1 unit Taq DNA polymerase in a total volume of 25 μl. Fragments were amplified by a touchdown PCR program (94°C for 120 sec; 10 cycles of 94°C for 20 sec, 68°C for 30 sec with a decrement of 1°C per cycle, and 72°C for 60 sec; 25 cycles of 94°C for 20 sec, 58°C for 30 sec, and 72°C for 60 sec; 72°C for 180 sec) in a GeneAmp PCR system 9700 (Applied Biosystems). After amplification, PCR products were diluted 800-fold with milli Q water. In the second (nested) round of PCR, fragments were amplified under identical cycling conditions with 5 μl diluted first PCR product as template, 0.2 μM adaptor primer 2 (ADAP-2: 5′-GGATATCAGGTCTGCAGCCTATAG-3′), and 0.2 μM 5′ or 3′ LTR-specific primer 2 (5LTR-2: 5′-ATCTGGTCTAACCAGAGAGACC-3′; 3LTR-2: 5′- TGTGTGACTCTGGTAACTAGAGATC -3′) in a total reaction volume of 25 μl. PCR products were size selected on a 1% agarose gel and 300-600 bp fragments were purified using a PureLink Quick Gel Extraction kit (Invitrogen). Purified, size-selected fragments for the 5′ LTR and 3′ LTR PCR reactions were combined and cloned using a TOPO TA Cloning kit

(Invitrogen). Individual transformants were analyzed by standard colony PCR assays and Sanger DNA sequencing. The lentiviral insertions sites were finally extracted using BLAST/BLAT searches in the human genome (Build GRCh38) at Ensembl (www.ensembl.org).

Hematopoietic cell transplantation and blood analysis.

For limiting dilution experiments cells were injected in given doses (Fig. S3A) into lethally irradiated (9 Gy) female B6 recipients together with 2×10^6 fresh bone marrow B6 cells. Due to the fact that not all transduced cells express GFP at the moment of transplantation a small aliquot of cells was kept in culture to test the gene transfer efficiency and sort GFP+ cells for qPCR 3-5 days post-transduction. Blood samples were taken every 4-6 weeks and stained with antibodies against CD45.1 and CD45.2 to determine the donor chimerism and CD3ε, Gr-1, B220 for FACS analysis. HSC frequency was estimated using ELDA software (http://bioinf.wehi.edu.au/software/elda/).

Serial bone marrow transplantation

Primary recipients were sacrificed at 16 weeks post transplantation for bone marrow collection prior secondary transplantation- 10^7 whole bone marrow cells were transplanted into lethally irradiated secondary recipients. For tertiary recipients donors were sacrificed 24 weeks post-secondary transplantation and 5×10^6 whole bone marrow were injected together with 5×10^5 fresh competitor cells from B6.

Cellular barcoding

Cells were isolated and transduced as described above (see also Verovskaya et al., 2013). We have transplanted 15760- GFP+ progenitor cells/mouse and 4300 GFP+ LT-HSC overexpressing miR-125a, 39800 GFP+ progenitors or 16200 GFP+ LT-HSC for controls.

Quantitative PCR

RNA was extracted from 25-30,000 GFP$^+$ cells using miRNeasy Kit (Qiagen, Valencia) and eluted in 30 uL of RNeasy-free water. For miRNA qPCR commercially available TaqMan probes (Life Technologies) were used according to manufacturer`s protocol. For target validation primers listed in Table 4 were used. For the human cells, RNA was extracted from mOrange$^+$ cells using Trizol (LifeTechnologies). Confirmation of miR-125a overexpression in transduced cells was done by Taqman small RNA assays according to manufacturer's instructions (LifeTechnologies). For target validation Taqman gene expression (LifeTechnologies) assays were used according to manufacturer's instructions.

Mass Spectrometry Sample Preparation

Triplicate Lin- CB cell sample pools were thawed and plated in X-VIVO (BioWhittaker) supplemented with 20% BIT 9500 (StemCell Technologies), 2 mM L-glutamine and 100 U/mL penicillin/streptomycin (P/S). Viral particles were resuspended in the same medium and added to the cells at a multiplicity of infection (MOI) of 30. The mixture was then supplemented with stem cell factor (SCF) 100 ng/mL, Flt3 ligand (Flt3-L) 100 ng/mL, thrombopoietin (TPO) 20 ng/mL and interleukin 6 (IL-6) 20 ng/mL, for 4 days. Four days post viral transduction, lin- CB cells were sorted for mO$^+$CD34$^+$ populations. Transduced CD34+ CB cells with miR-125a (125/OE) and control vector (CTRL) were counted and washed twice with ice-cold PBS. 100,000 cells for each experimental condition, in biological triplicate, were subjected to sample preparation similar

to (Kulak et al., 2014). Cells were lysed using 20 µl of lysis buffer (consisting of 6 M Guanidinium Hydrochloride, 10 mM TCEP, 40 mM CAA, 100 mM Tris pH8.5). Samples were boiled at 95°C for 5 minutes, after which they were sonicated on high for 3x 10 seconds in a Bioruptor sonication water bath (Diagenode) at 4°C. Samples were diluted 1:3 with 10% Acetonitrile, 25 mM Tris pH 8.5, LysC (MS grade, Wako) was added in a 1:50 (enzyme to protein) ratio, and samples were incubated at 37°C for 4hrs. Samples were further diluted to 1:10 with 10% Acetonitrile, 25 mM Tris pH 8.5, trypsin (MS grade, Promega) was added in a 1:100 (enzyme to protein) ratio and samples were incubated overnight at 37°C. Enzyme activity was quenched by adding 2% trifluoroacetic acid (TFA) to a final concentration of 1%. Prior to mass spectrometry analysis, the peptides were fractionated using Strong Cation Exchange (SCX) in StageTip format. For each sample, 3 discs of SCX material (3M Empore) were packed in a 200ul tip, and the SCX material activated with 80 µl of 100% Acetonitrile (HPLC grade, Sigma). The tips were equilibrated with 80 µl of 0.2% TFA, after which the samples were loaded using centrifugation at 4,000x rpm. After washing the tips twice with 100µl of 0.2% TFA, two initial fractions were eluted into clean 500µl Eppendorf tubes using 75, and 200 ammonium acetate, 20% Acetonitrile, 0.5% formic acid respectively. The final fraction was eluted using 5% ammonium hydroxide, 80% Acetonitrile. The eluted fractions were frozen on dry ice and concentrated in an Eppendorf Speedvac, and re-constituted in 1% TFA, 2% Acetonitrile for Mass Spectrometry (MS) analysis.

Mass Spectrometry Acquisition

For each SCX fraction, peptides were loaded onto a 2cm C18 trap column (ThermoFisher 164705), connected in-line to a 50cm C18 reverse-phase analytical column (Thermo EasySpray ES803) using 100% Buffer A (0.1% Formic acid in water) at 750bar, using the Thermo EasyLC

1000 µHPLC system in a single-column setup and the column oven operating at 45°C. Peptides were eluted over a 200 minute gradient ranging from 5 to 48% of 100% acetonitrile, 0.1% formic acid at 250 nl/min, and the Orbitrap Fusion (Thermo Fisher Scientific) was run in a 3 second MS-OT, ddMS2-IT-HCD top speed method. Full MS spectra were collected at a resolution of 120,000, with an AGC target of 4×10^5 or maximum injection time of 50ms and a scan range of 400–1500m/z. Ions were isolated in a 1.6m/z window, with an AGC target of 1×10^4 or maximum injection time of 35ms, fragmented with a normalized collision energy of 30 and the resulting MS2 spectra were obtained in the ion trap. Dynamic exclusion was set to 60 seconds, and ions with a charge state <2, >7 or unknown were excluded. MS performance was verified for consistency by running complex cell lysate quality control standards, and chromatography was monitored to check for reproducibility. Each sample was run in technical duplicate and biological triplicate, and the reproducibility of the analyses is depicted in **Supplementary Figure S5**. The mass spectrometry data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD001994.

Label-free Quantitative Proteomics Analysis

The raw files were analyzed using MaxQuant version 1.5.2.8 (Cox and Mann, 2008) and standard settings. Briefly, label-free quantitation (LFQ) was enabled with a requirement of 3 unique peptides per protein, and iBAQ quantitation was also enabled during the search. Variable modifications were set as Oxidation (M), Acetyl (protein N-term), Gln->pyro-Glu and Glu->pyro-Glu. Fixed modifications were set as Carbamidomethyl (C), false discovery rate was set to 1% and "match between runs" was enabled. The resulting protein groups file was processed with an in-house developed tool (PINT), which imputes missing LFQ values with adjusted iBAQ

values. Briefly, the distributions of iBAQ intensities for each sample are adjusted to overlap with the LFQ intensity distributions using median-based adjustment, enabling the direct imputation of missing LFQ values with adjusted iBAQ values for those proteins that did not have LFQ values across all the samples. By simultaneously filtering for reverse hits, contaminants and only those proteins observed in 3 biological replicates (N=3), this resulted in a final list of 6,687 proteins identified and quantified in all samples. This table is included as **Table S5**, and was used as input for downstream analysis with GSEA.

Proteomics Data Analysis

MaxQuant LFQ (Cox and Mann, 2008) intensities were used as a measure of protein expression in 125OE and control samples. The entire protein expression set consisting of 3 biological replicates for each treatment group and corresponding technical replicates (total of 12 samples) was quantile normalized in R (R version 3.1.1) using the normalizer package (version 1.0). The normalized protein expression was further filtered to contain only proteins that had at least two measurements in either treatment or control.

Difference in protein expression between the groups was assessed using a moderated t-test available in the bioconductor limma package (version 3.20.9). P-values were further corrected to control for multiple hypothesis testing using the Benjamini-Hochberg procedure. While adjusted p-values did not reach significance, 518 proteins had significant differential expression with nominal p-value <0.05, of which 209 were upregulated (t>0) and 309 were downregulated (t<0). Proteins and their corresponding t-statistic were used to create a rank file to be used in pathway analysis described below.

Pathway Analysis on the proteomics data

Gene Set Enrichment Analysis (Subramanian et al., 2005) was performed using the protein expression ordered from largest to smallest t statistics with parameters set to 1000 gene-set permutations and gene-sets size between 5 and 500. The gene-sets included for the GSEA analyses were obtained from KEGG, MsigDB-c2, NCI, Biocarta, IOB, Netpath, HumanCyc, Reactome, Panther and Gene Ontology (GO) databases, updated October 26, 2015 (http://download.baderlab.org/EM_Genesets/). An enrichment map (version 2.1.0 of Enrichment Map software (Merico et al., 2010) was generated using Cytoscape 3.3.0 using significantly enriched gene-sets with a nominal p-value <0.05 and FDR <0.2. Similarity between gene-sets was filtered by Jaccard coefficient >0.25. The resulting enrichment map was further annotated using the AutoAnnotate Cytoscape App.

Correlation between miR-125 predicted targets and the proteomics modulated pathway

Four databases were used to create a list of miR-125 predicted targets (DIANA microT, picTar, TargetScan from the miRbase website (http://www.mirbase.org)) and miRanda from the microCosm website (http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/).

Predicted downregulated targets were compared to enriched gene-sets in the enrichment map and significance of overlap was scored using a one-sided (lesser) Mann-Whitney test so that highest significance corresponds to proteins in the overlap showing the highest reduction in expression amplitude .

Correlation of the miR-125 predicted targets and the miR-125 modulated pathways

Four databases were used to create a list of miR-125 predicted targets (DIANA microT, picTar, TargetScan from the miRbase website (http://www.mirbase.org)) and miRanda from the microCosm website (http://www.ebi.ac.uk/enright srv/microcosm/htdocs/targets/v5/). 84

predicted targets were down-regulated in 125OE. This list was compared to enriched gene-sets in the enrichment map and overlap was scored using a one-sided Mann-Whitney test. Overlaps with p-value <0.01 were visualized.

Quantitative Mass Spectrometry (SILAC)

The murine 32D cell line was used for to identify miR-125a targets. After transduction GFP+ cells were FACS-sorted and expanded prior to stable isotope labelling as previously reported (Meenhuis et al., 2011). A nanoHPLC system consisting of Easy-nLC 1000 and Q-Exactive (Thermo Fisher Scientific) was used.

Statistical Analysis

Unless otherwise stated mean ± SEM values are shown in the graphs. For pairwise comparison a Mann-Whitney-U test was applied. Statistical analysis was performed using PRISM 6 (GraphPad Software). For proteomics data analysis student t-test with Bonferroni correction was applied.

Western blot analysis

Human lin-CB cells were thawed and plated in X-VIVO 10 medium (BioWhittaker) supplemented with 1% BSA, SCF, TPO, FLT-3 and IL-6, and transduced with viral particles at a multiplicity of infection (MOI) of 30 for at least 16 hours. Cells were allowed to expand for 5 days and then flow sorted for mO+ hCD34+ cells. Total cellular proteins were extracted with RIPA buffer (20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 5 mM EDTA, 1% Triton X-100) supplemented with protease and phosphatase inhibitors: 1 mM PMSF, 10 mM NaF, 1 mM Na_3VO_4, CompleteMini™ and PhosStop ™ (Roche). Samples were resuspended in the lysis solution and incubated at 4°C for 30 min. Cell lysates were cleared by centrifugation at $10,000 \times g$ for 10 min at 4°C, and the supernatants were collected and assayed for protein

concentration using Lowry assay based method (DC, BioRad). 40-50 micrograms of proteins were run on SDS-PAGE under reducing conditions. For immunoblotting, proteins were transferred to PVDF membranes, incubated with the specific antibody (anti-p38 Cell Signaling #9212 1:1000, and anti-PTPN1 LifeSpan BioSciences #LS-C61924 1:500 and mouse anti-GAPDH 1:10,000, Sigma) followed by peroxidase-conjugated secondary antibodies. Bands were visualized on Amersham Hyperfilm[TM] (GE Healthcare).

Nanofluidic proteomic analysis

The expression levels of p38 in transduced GFP+ 32D cells used for SILAC or in LT-HSC and progenitors overexpressing miR-125a were investigated by a size-based Simple Western immunoassay, using a WES device (Protein Simple). Equal number of cells was sorted and lysed in RIPA buffer, reduced and denatured before loading onto the plate. Plates were prepared according to the manufacturing's procedure, using all reagents from Protein Simple. P38 antibody (Cell Signaling, # 9212) was used at 1:100 dilution, vinculin antibody (SIGMA, # hVIN-1) was used at a 1:150 dilution, Smek1 (Abcam, #ab70635), Ptpn1 (Millipore, # ABS40) and Ptpn7 (Abcam, # 3b10) were used at 1:50 dilution. Data were analyzed using the Simple Western Compass software. Protein expression (area under the curve, AUC) was corrected for vinculin loading control (AUC p38/AUC vinculin).

Immunofluorescence

The protein expression levels in transduced GFP+ LT-HSC (Lin-Sca-1+c-Kit+CD150+48-) or progenitors (Lin-Sca-1+c-Kit+ depleted from CD150+ 48- cells) were analyzed using immunofluorescence. Cells were FACS-purified, transduced with control or miR-125a

overexpressing vectors for 3 days. On day 5 we sorted GFP+ cells (2000-5000/condition) directly on pre-coated slides (Marienfield). Afterwards cells were stained over night at 4^0C according to manufacturer's protocols with primary antibody. Primary antibodies and dilutions: p38 (1:50, Cell Signaling, #9212), Ptpn1 (1:100, Millipore, clone# ABS40). We used rabbit-specific IgG conjugated with Alexa-633 (Molecular Probes, Invitrogen) as a secondary antibody. For nuclear staining we treated cells with DAPI (Sigma). Fluorescent images were obtained using TissueFaxs (TissueGnostics, based on a high-end fully motorized Zeiss AxioObserver Z1 microscope).