Green fluorescent protein transgene driven by Kit regulatory sequences is expressed in hematopoietic stem cells

Francesco Cerisoli,[1] Letizia Cassinelli,[2] Giuseppe Lamorte,[2] Stefania Citterio,[2] Francesca Bertolotti,[1] Maria Cristina Magli,[1] and Sergio Ottolenghi[2]

[1]Institute of Biomedical Technologies, National Council of Research, Pisa, Italy, and [2]Department of Biotechnology and Bioscience, University of Milano-Bicocca, Milan, Italy

ABSTRACT

Background

The transcriptional regulation of stem cell genes is still poorly understood. Kit, encoding the stem cell factor receptor, is a pivotal molecule for multiple types of stem/ progenitor cells. We previously generated mouse lines expressing transgenic green fluorescent protein under the control of Kit promoter/first intron regulatory elements, and we demonstrated expression in hematopoietic progenitors.

Design and Methods

In the present work we investigated whether the transgene is also expressed in hematopoietic stem cells of adult bone marrow and fetal liver. To this purpose, we tested, in long-term repopulating assays, cell fractions expressing different levels of green fluorescent protein within Kit-positive or SLAM-selected populations.

Results

The experiments demonstrated transgene expression in both fetal and adult hematopoietic stem cells and indicated that the transgene is transcribed at distinctly lower levels in hematopoietic stem cells than in pluripotent and committed progenitors.

Conclusions

These results, together with previous data, show that a limited subset of DNA sequences drives gene expression in a number of stem cell types (hematopoietic stem cells, primordial germ cells, cardiac stem cells). Additionally, our results might help to further improve high level purification of hematopoietic stem cells for experimental purposes. Finally, as the Kit/green fluorescent protein transgene is expressed in multiple stem cell types, our transgenic model provides a powerful *in vivo* system to track these cells during development and tissue regeneration.

Key words: Kit, hematopoietic stem cells, hematopoietic progenitor cells, transgenic mouse.

Citation: Cerisoli F, Cassinelli L, Lamorte G, Citterio S, Bertolotti F, Magli MC, Ottolenghi S. Green fluorescent protein transgene driven by Kit regulatory sequences is expressed in hematopoietic stem cells. Haematologica 2009; 94:318-325. doi:10.3324/haematol.13689

Acknowledgments: we are particularly grateful to Prof. Luca Cionini and his collaborators for the irradiation of the animals, to Dr. Esther Haak for the bone marrow lineage depletion protocol and to Dr. G. Castoldi for mir221 and 222 evaluation. We are indebted to the Constem Consortium for allowing access to their FACS equipment.

Funding: this work was supported by grants from Progetto Nazionale Cellule Staminali, grant CS 73 (MCM and SO), MIUR-PNR grant RBNE01R4MJ-002 (MCM), Fondazione CARISBO grant 2005.9025 (MCM), Cofin MIUR 2005 (SO), and Fondazione CARIPLO 2005 (SO).

Manuscript received July 29, 2008. Revised version arrived September 24, 2008. Manuscript accepted October 17, 2008.

Correspondence:
Maria Cristina Magli and Sergio Ottolenghi, MCM, Istituto di Tecnologie Biomediche-CNR, Via Moruzzi 1, 56124 Pisa, Italy/Dipartimento di Biotecnologie e Bioscienze, Università degli Studi Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy.
E-mail: mariacristina.magli@itb.cnr.it/ sergio.ottolenghi@unimib.it

The online version of this article contains a supplementary appendix.

Introduction

In the last years, great emphasis has been placed on the isolation, characterization and potential therapeutic uses of stem cells. There is, therefore, considerable interest in understanding the regulatory transcriptional and signaling networks which underlie stem cell properties, such as pluri-multipotency and self-renewal, as well as specificity of commitment. Recently, transcriptional networks of multipotent embryonic stem cells have started to be analyzed by chromatin immunoprecipitation assays.[1-3] However, this approach cannot easily be extended to the study of the rare and not always prospectively identifiable tissue stem cells. For the latter stem cells, hierarchical networks of transcription factors and signaling molecules have been defined by gene ablation experiments, but to date the number of stem cell genes whose regulation has been approached at the transcriptional level remains relatively small.

Kit encodes the membrane receptor of stem cell factor (SCF), a diffusible ligand. It is expressed in several stem cell types, such as primordial germ cells, neural stem cells of the eye, hematopoietic stem cells (HSC), and cardiac stem cells, and several types of early committed progenitors. In these cells, the function of Kit is important or even essential, as demonstrated by genetic mutants and functional studies.[4-11] The expression of Kit in multiple stem cell types suggests that common transcriptional programs might regulate this gene in different stem cells. In previous work, we showed that a mouse Kit transgene is able to efficiently drive green fluorescent protein (GFP) expression in primordial germ cells and in immature hematopoietic cells, such as CFU-mix and BFU-E progenitor cells.[12] In the present work, we investigated whether the transgene is expressed in fetal liver and adult bone marrow HSC. Our results show that the transgene is indeed expressed in a Kit^+ cell type, capable of long-term hematologic reconstitution upon transplantation into irradiated recipients. The expression level of the transgene in HSC is distinctly lower than in other progenitors, a fact that might be exploited to further improve the purification of HSC. In conclusion, our transgene contains the major regulatory sequences required for Kit (Kit/GFP) expression in several types of stem and progenitor cells, thus providing a powerful model to study common stem cell transcriptional programs.

Design and Methods

Mice

Transgenic *Kit/GFP* mice, in which the *GFP*-encoding sequence was placed under the control of *Kit* regulatory elements (construct 3) have been previously described.[12] *Kit/GFP* founder transgenic mice (originally in a BDF1 background) were initially bred to BDF1 mice (for the first generation), and then interbred among themselves. Two different lines were tested, with similar results. Transgenic and wild type (BDF1) mice were kept according to institutional regulations.Transgenic mice were genotyped by polymerase chain reaction using 5'-ACATGAAGCAGCACGACTTC-3' and 5'-TTGTGGCGGATCTTGAAGTT-3' primers in the GFP sequence. For timed experiments on embryos, the morning when the vaginal plug was detected was considered to correspond to day 0.5 post-coitum (E0.5).

Flow cytometric analysis and cell sorting

Flow cytometric and cell sorting were performed by a MoFlo (DAKO-Cytomation) flow cytometer/cell sorter with Summit 4.3 and FACSAria software (Becton Dickinson, RD). Total adult bone marrow or fetal liver cells from *Kit-GFP* transgenic animals and from non-transgenic controls (obtained by disaggregation in phosphate-buffered saline) were incubated with appropriate labeled antibodies from BD-Pharmingen: phycoerythrin (PE)-labeled anti-mouse LY 6A/E (Sca-1) (D7); peridin chlorophyll-cyanin 5.5-labeled anti-mouse CD11b (Mac-1) (M1/70); allophycocyanin (APC)-labeled anti-mouse CD117 (Kit) (2B8); PE-labeled anti-mouse CD48 (HM48-1); APC-labeled anti-mouse Gr-1 (Ly-6G and Ly-6C) (RB6-8C5); biotin anti-mouse T-cell receptor beta chain (H57-597); and biotin anti-mouse CD45R/B220 (RA3-6B2).The PE-labeled anti-mouse CD41 (MWREG30) was from Santa Cruz. The PE anti-mouse CD71, (clone R 17127) and APC anti-mouse CD150 (SC15-12F12.2) were from Biolegend (San Diego, CA). All the analyses were performed with the Moflo flowcytometer cell sorter, except the one shown in Figure 1 for which a Cyan ADP (Dako) was used. Background fluorescence levels (autofluorescence) were defined by comparison to levels in appropriate non-transgenic control cells, stained with isotype control fluorochrome conjugate. For the analysis of the side population,[13] bone marrow cells (3×10^6/mL) were incubated with Hoechst 33342 (Sigma), 5 μg/mL for 90 min at 37°C, either in the absence (experimental samples) or presence (control samples) of 50 μM verapamil. The cells were then washed four times with fresh culture medium (1% fetal calf serum in Dulbecco's modified Eagle's medium), further incubated with labeled antibodies for 30 min at 4°C, washed again at 4°C and analyzed by Moflo (DAKO) with UV (350 nm), 630 nm and 488 nm lasers. Data from at least 1.5×10^6 live cells were acquired. Cell sorting of bone marrow or disaggregated fetal liver cells from heterozygous transgenic embryos (identified by their bright GFP fluorescence) was performed after APC-labeled anti-Kit staining, and Kit^+ cells were fractionated on the basis of levels of GFP fluorescence. In order to sort signaling lymphocytic activation molecule (SLAM)-positive HSC, bone marrow cells were pre-enriched with a double density Ficoll gradient[14] adapted for mouse cells, in which undifferentiated mononuclear cells are layered between 1.069 g/mL and 1.075 g/mL by centrifugation.

In vitro *colony assays*

The culture conditions for the *in vitro* colony assay have been described elsewhere.[12,15] Briefly, bone marrow cells were flushed from femora and tibiae in Iscove's modified Dulbecco's medium (IMDM) with 2% fetal calf serum and cultured at a maximum density of 5×10^4 cells/mL in IMDM containing α-thioglycerol, methyl cellulose, bovine serum albumin, 5% fetal calf serum, iron-saturated transferrin, lecithin, oleic acid, cholesterol, 2 u/mL human recombinant erythropoietin, 10 ng/mL murine stem cell

factor, and 10 ng/mL murine interleukin-3. Cultures were scored both in bright field and fluorescence at 2, 7, and 12 days with an inverted microscope by using standard criteria.

Bone marrow transplantation

BDF1 female recipient mice, 8-12 weeks old, were lethally irradiated with 830 rads (x-ray source). Two hours after irradiation, they were injected intravenously with different numbers of sorted fetal liver or bone marrow cells, and 5×10^4 to 2×10^5 (as specified in individual experiments) unsorted wild type bone marrow cells as support, resuspended in IMDM. Controls were injected with cell-free suspending medium. Engrafted animals were killed and analyzed 4-6 months after transplantation. For secondary transplants, 10^7 cells from primary reconstituted mice were injected intravenously into lethally irradiated secondary BDF1 recipients.

Results

The transgenic Kit/GFP gene is expressed in hematopoietic stem cells of adult bone marrow and fetal liver

The *Kit* transgene consists of 6.7 Kb of mouse promoter sequences, followed by the *GFP* gene and by 8 Kb of sequences from the 5' part of the first intron of the gene.[12] In addition to the promoter, only approximately 4 Kb at the 5' part of the intron are sufficient to drive expression of the transgene; however, mouse lines including all 8 Kb of the intron show more accurately regulated transcription, and we therefore studied one of these lines in detail (mouse line 3 described by Caims *et al.*) (another line obtained with the same construct from a different founder gave similar results).

In previous research, both E11.5 fetal liver and adult bone marrow cells FACS-sorted for GFP (or GFP and Kit) expression were found to be highly enriched for CFU-mix and BFU-E; however, we did not ascertain GFP expression in HSC.[12] To address this question, we selected cell populations with different GFP fluorescence levels from adult bone marrow and E12.5 fetal liver (Figures 1A and 1B), and tested each of them by *in vitro* colony assays,[12] as well as by long-term *in vivo* repopulation studies in lethally irradiated mice (HSC assays) (Figure1). For long-term repopulation studies, we transplanted FACS-sorted cell fractions at various concentrations (ranging from 100 to 50,000 cells) together with a fixed number (2×10^5 cells) of non-transgenic total bone marrow cells as a support. After at least 4 months, we assessed hematologic reconstitution in two ways (Figures 1C-D). First, bone marrow from the recipient mice was plated in methylcellulose cultures, and the proportion of GFP-expressing (versus non-expressing) CFU-mix colonies was evaluated (Figure 1C). Previous work had shown that over 90% of CFU-mix colonies from the transgenic line used here express GFP, thus the proportion of fluorescent CFU-mix colonies is expected to faithfully reflect the level of reconstitution of the bone marrow by *GFP* transgenic versus non-transgenic co-transplanted cells. Second, we analyzed the GFP expression (by FACS) of bone marrow from the transplanted mice. In general a high level of reconstitution, as measured by colony assays (over 90% of fluorescent colonies), was paralleled (Figure 1D) by a high level of fluorescent cells in the bone marrow of reconstituted mice (25-35%, note that most differentiated bone marrow cells do not express Kit/GFP) (see below and *Online Supplementary Figures S1 and S2*).

The pattern of GFP expression in adult bone marrow is rather complex. Over 90% of Kit^+ cells are also GFP^+; within the Kit^+ population, the majority of the cells show high levels of fluorescence (GFP^{high} cells), whereas a smaller proportion (GFP^{int} cells) show lower levels of fluorescence, and are poorly separated from the minor Kit^+/GFP^- popu-

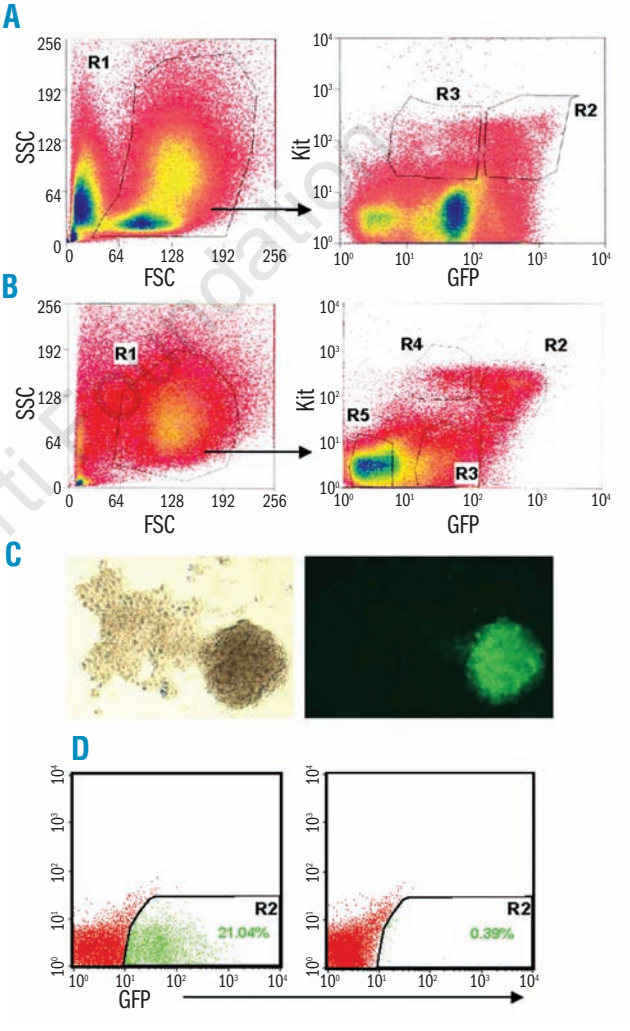

Figure 1. Experimental strategy for the evaluation of the HSC content of different fetal liver and bone marrow fractions. Bone marrow (panel A) and fetal liver (panel B) cell fractions of different GFP fluorescence were transplanted into irradiated mice. Reconstitution was evaluated by the proportion of fluorescent (*i.e.* transgenic) versus non-fluorescent (*i.e.* non-transgenic, from co-transplanted supporting bone marrow) CFU-mix type colony (panel C) obtained from *in vitro* clonogenic assays of the bone marrow of transplanted mice. Note the intense fluorescence of the CFU-mix type colony, versus the complete absence of fluorescence of the adjacent mature myeloid colony. FACS analysis of the same marrow (panel D) shows excellent agreement of the results with those obtained by colony assays. The negative bone marrow sample (panel D, right side) did not yield any fluorescent CFU-mix, whereas the positive sample (left) yielded 87% fluorescent CFU-mix type colonies.

lation. Within the Kit$^-$/lineage marker (Lin)$^+$ population, the analysis of GFP fluorescence (*Online Supplementary Figure S1* and *data not shown*) showed that a large proportion of the Kit$^-$/Lin$^+$ cells are either completely GFP$^-$ or show much lower levels of GFP expression than Kit$^+$ cells. In particular, 60-90% of Kit$^-$/CD71 or Ter119$^+$ cells (erythroid cells) are GFP$^-$, as well as 90-95% of lymphoid cells (B220 and T-cell receptor positive cells, *not shown*); myeloid cells (Gr-1$^+$ and CD11b$^+$) express GFP at low levels. Interestingly, and unexpectedly, Kit$^+$/CD71$^+$ or Kit$^+$/Gr-1$^+$ double positive cells express very high levels of GFP, even higher than those shown in Kit$^+$/CD71$^-$ or Kit$^+$/Gr-1$^-$ cells (*Online Supplementary Figure S1*, panels viii and xii, compare the green line, *i.e.* Kit$^+$/lin$^+$ double positive, with the red and blue lines, *i.e.* Kit single positive and Kit$^-$/lin$^+$ cells, respectively). Similarly to bone marrow, in fetal liver (*Online Supplementary Figure S2*) Kit$^+$ cells are divided into GFP^{high} and GFP^{int} cells (with very few GFP$^-$ cells); Kit$^-$/Lin$^+$ cells (largely erythroid, *i.e.* CD71$^+$ or Ter119$^+$ cells) are mostly (90%) GFP$^-$, with a small proportion of weakly GFP$^+$ cells (panel viii).

For colony assays and transplantation experiments, we FACS-sorted Kit$^+$ cells based on their levels of GFP expression. Kit$^-$/GFP$^+$ cells were tested as well. Table 1 shows that most bone marrow progenitors, as assayed *in vitro* (CFU-mix, BFU-E), are contained in the GFP^{high} fraction. In contrast, HSC are highly enriched in the bone marrow Kit$^+$/GFP^{int} fraction, as demonstrated by the hematologic reconstitution of most transplanted mice (Table 2). Bone marrow cells expressing intermediate levels of GFP efficiently reconstituted a total of 16 out of 23 transplanted animals (Table 2, Kit$^+$/GFP^{int} and GFP^{int} fractions). In contrast, only four out of 22 recipients were repopulated by the GFP^{high} fractions (Table 2, Kit$^+$/GFP^{high} and GFP^{high} cells). Furthermore, whereas 2×10^3-1×10^4 bone marrow Kit$^+$/GFP^{int} cells were sufficient to reconstitute four out of seven transplanted mice at a high level (35-50% of fluorescent CFU-mix), a comparable level of repopulation could be obtained only with 5×10^4 bone marrow Kit$^+$/GFP^{high} cells (Table 2 and Figure 2). As expected, mice that received Kit$^-$/GFP$^+$ cells did not show any reconstitution with fluorescent cells. To further confirm that the bone marrow reconstitution was really due to long-term repopulating cells, bone marrow from primary recipients was transplanted into secondary irradiated recipients. The results presented in Table 2 show that only bone marrow cells derived from GFP^{int} transplanted mice were able to fully repopulate three out of four secondary recipients, whereas GFP^{high} cells did not show reconstituting ability either in primary or secondary transplants. We also obtained similar conclusions by transplantation of fractionated E12.5 fetal liver cells (Figure 1B). Again, Kit$^+$/GFP^{high} fractions contained significantly more colony-

Table 1. Hematopoietic precursors in the sorted cell fractions.

Cell fraction	BFU-E CFUx10^{-4} cells	CFU-Mix CFUx10^{-4} cells	Total myeloid CFUx10^{-4} cells
A. Bone marrow			
Kit$^+$/GFP^{high}	145 (82%)	233 (72%)	330 (37%)
Kit$^+$/GFP^{int}	32 (18%)	91 (28%)	574 (63%)
Kit$^-$/GFP^{int}	–	–	–
Kit$^-$/GFP$^-$	–	–	–

Cell fraction	BFU-E CFUx10^{-4} cells	CFU-Mix CFUx10^{-4} cells	Total myeloid CFUx10^{-4} cells
B. Fetal liver			
Kit$^+$/GFP^{high}	46 (79%)	25 (53%)	134 (64%)
Kit$^+$/GFP^{int}	53 (21%)	96 (47%)	283 (33%)
Kit$^-$/GFP^{int}	–	–	–
Kit$^-$/GFP$^-$	–	–	–

Numbers of different progenitors were evaluated in duplicate plates. The figures shown in the first column represent progenitor number/10^4 cells. Figures between parentheses represent the proportion of each progenitor that is contained in the cell fraction tested, calculated on the basis of numbers of cells present in each sorted fraction (see legend to Figure 2), multiplied by the progenitor numbers/10^4 cells obtained in the present experiment.

Table 2. Long-term reconstitution of bone marrow by transplantation of cell fractions expressing different levels of green fluorescent protein.

Tissue	Fraction	Cells injected	I Reconstituted	II Reconstituted
Bone marrow	Kit$^+$/GFP^{high}	5×10^4	1/2	–
		2×10^4	2/3	–
		10^4	1/6	–
		2×10^3	0/4	–
		2×10^2	0/3	–
	Kit$^+$/GFP^{int}	5×10^4	3/3	–
		2×10^4	3/3	–
		10^4	4/7	–
		2×10^3	3/4	–
		2×10^2	0/3	–
	Kit$^-$/GFP$^-$	10^4	0/1	–
	Kit$^-$/GFP$^+$	5×10^4	0/1	–
		2×10^4	0/2	–
		10^4	0/2	–
Bone marrow	GFP^{high}	10^5	0/2	0/9
		2×10^4	0/2	–
	GFP^{int}	10^5	2/2	3/4
		2×10^4	1/1	–
	GFP$^-$	10^5	0/3	0/2
Fetal liver	Kit$^+$/GFP^{high}	4×10^4	0/1	–
		2×10^4	0/1	–
		10^4	0/1	–
		10^3	0/1	–
	Kit$^+$/GFP^{int}	4×10^4	1/1	–
		2×10^4	1/1	–
		10^4	2/2	–
		10^3	0/2	–
	Kit$^-$/GFP^{int}	4×10^4	0/1	–
		2×10^4	0/1	–
		10^4	0/1	–
		10^3	0/1	–
	Kit$^-$/GFP$^-$	4×10^4	0/2	–

GFP: green fluorescent protein. In most experiments, cells were selected on the basis of their expression of Kit and GFP as indicated. In some experiments, cells were selected only on the basis of their GFP expression level. For fraction identification, see examples in Figure 1.

forming cells than Kit$^+$/GFP^{int} fractions (Table 1). In contrast, upon transplantation Kit$^+$/GFP^{int}, but not Kit$^+$/GFP^{high} cell fractions (Figure 1), efficiently repopulated lethally irradiated mice (Table 2 and Figure 2B).

Long-term repopulating hematopoietic stem cells are contained in the GFP^{int} population

We subsequently tested whether the above results were confirmed in populations highly enriched in HSC. Recently, it was shown that stringent selection for CD150$^+$, CD48$^-$, CD41$^-$ cells (SLAM cells) yields a population containing the large majority of HSC, at a high level of purity.[16,17] We initially selected lineage marker-negative cells from Kit/GFP bone marrow, and from them we purified the HSC-enriched SLAM population (Figure 3 and *Online Supplementary Figure S3*). As shown in Figure 3, lineage marker-negative cells are dramatically enriched, relative to the total population (compare panels a and b), in Kit/GFP^{high} cells and Kit/GFP^{int} cells, which yield two prominent, well identified peaks (Int and High).Upon SLAM purification, the selected population CD150$^+$/CD48$^-$/CD41$^-$ (representing 0.0045 ± 0.0005% of total bone marrow), consisted mostly of GFP^{int} cells (about 58%), with a shoulder (33%) of GFP^{high} cells (Figure 3, compare panel c to panels a and b): a low percentage of SLAM HSC cells (about 9%) were GFP$^-$. We then transplanted limiting dilutions of SLAM HSC sorted cells (*Online Supplementary Figure S3*) on the basis of high, intermediate or null expression of GFP into lethally irradiated recipients. None of the mice transplanted with SLAM

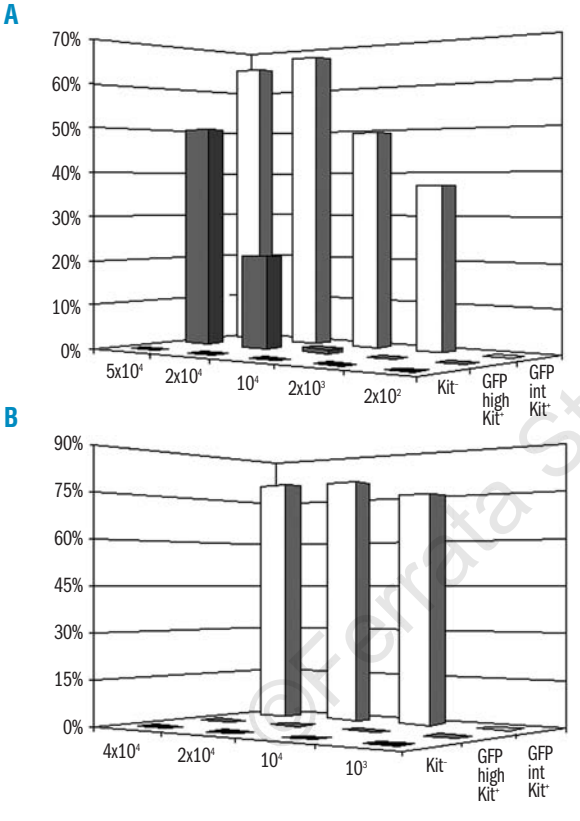

Figure 2. Relative hematopoietic stem cells content of bone marrow and fetal liver cell fractions expressing different levels of GFP, as evaluated by long-term repopulation studies. Bone marrow long-term repopulation (4-5 months) by cell fractions from bone marrow (A) and fetal liver (B). Reconstitution was evaluated by the proportion of GFP$^-$ expressing CFU-mix type colonies obtained from the bone marrow of each transplanted mouse; GFP$^-$ colonies are assumed to derive from non-transgenic stem cells contained in the support bone marrow co-transplanted with the FACS-fractionated transgenic cells. Note that 90-100% of CFU-mix type colonies from untreated Kit/GFP mice express GFP[12] The number of cells used for each transplant, and their level of GFP expression are shown. Fractions used are as exemplified in Figure 1. Kit$^-$ indicates GFP$^+$ or GFP$^-$ cell fractions not expressing Kit; GFP^{int}: Kit$^+$ cells expressing intermediate levels of GFP; GFP^{high}: Kit$^+$ cells expressing high levels of GFP. The data for the transplantation of fetal liver cells refer to an experiment using cells pooled from 16 GFP$^+$ livers (five pregnant females) at E12.5. The data for bone marrow are from four adult donor transgenic mice.

Table 3. Long-term reconstitution of bone marrow by transplantation of SLAM HSC expressing different levels of Kit/GFP.

Fraction (CD150$^+$, 48$^-$, 41$^-$)	Cells injected	Reconstituted	% of reconstitution
GFP^{high}	300	2/3	2.4, 0.5
	100	1/3	1.20
	10	0/3	
GFP^{int}	300	3/3	54, 19.4, 31.4
	100	2/2	14.9, 1
	10	1/3	1.10
GFP$^-$	300	0/3	
	100	0/3	

Reconstitution of transplanted mice was evaluated as percentage of GFP$^+$ cells in bone marrow, by FACS analysis, 4 months after transplantation.

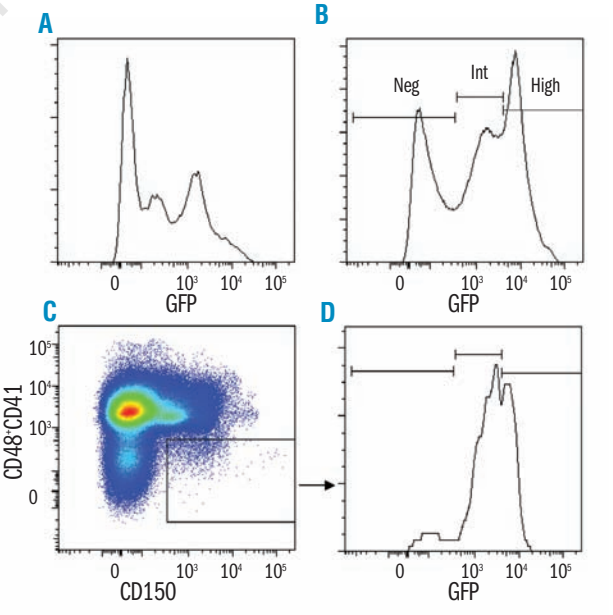

Figure 3. FACS analysis and separation of CD150$^+$CD48$^-$CD41$^-$ lineage-depleted cells on the basis of Kit/GFP expression. Total bone marrow from adult Kit/GFP mice (A) was depleted of lineage-positive cells (B). After lineage depletion, SLAM hematopoietic stem cells were gated as shown in (C); the right panel (D) represents the levels of GFP expression of CD150$^+$CD48$^-$CD41$^-$ hematopoietic stem cells-enriched fractions. GFP-negative (Neg), intermediate (Int) or high cells were gated (gates were set on the general GFP expression level shown in *Online Supplementary Figure S1* and represented 8.92±1.80%, 57.77±4.60%, 33.10±2.52%, respectively. The same gating strategy was used for sorting (see *Online Supplementary Figure S3*).

HSC GFP$^-$ cells was repopulated in the short or long-term after transplantation, while both SLAM HSC GFP^{high} and GFP^{int} cells were capable of short-term repopulation, at all the cell doses used (*data not shown*). In strong contrast, however, after 4 months SLAM HSC GFP^{int} cells showed a much better long-term repopulating ability than that of GFP^{high} cells (Table 3). Taken together, these results demonstrate that the *KitGFP* transgene is expressed in HSC of adult bone marrow and fetal liver. The intermediate level of GFP expression in HSC contrasts with the high level of expression seen in progenitors (Table 1). Interestingly, FACS analysis of the *side population* of adult bone marrow shows that the *tip* of this population, which is known to consist of a high proportion of Kit$^+$/Sca1$^+$ HSC[18] (and SLAM HSC),[19] is enriched in Kit$^+$/Sca1$^+$ GFP^{int} cells, whereas the *base* includes mainly Kit$^+$/Sca1$^-$ Kit/GFP^{high} cells (*Online Supplementary Figure S4*). This result independently confirms, at the phenotypic level, the data obtained by SLAM analysis and transplantation. Similarly, in the hematopoietic fetal liver, Sca1$^+$, CD11b$^+$, Kit$^+$ cells (representing the immunophenotypic properties of fetal liver HSC)[20-22] mostly exhibit intermediate levels of Kit/GFP expression (*Online Supplementary Figure S5*).

Thus, HSC, as defined by both functional and phenotypic criteria, express GFP and, unexpectedly, are Kit/GFP^{int}.

Discussion

In a previous study we generated a *Kit* transgene capable of driving GFP expression in progenitor cells both of the germ and the hematopoietic cell lineages.[12] In this work we report that this transgene is expressed in the large majority of fetal liver and adult bone marrow HSC. The distinct intermediate level of GFP expression detected in these cells may further help their purification for experimental purposes.

The present results, together with other reports,[9,12,23] now indicate that this transgene is expressed in a variety of stem cells, ranging from primordial germ cells to cardiac and hematopoietic stem cells, providing the basis for the elucidation of DNA sequences regulating a stem cell gene in multiple stem cell types.

Transgenic Kit/GFP *is expressed in most hematopoietic Kit$^+$ cells*

The *Kit/GFP* transgene appears to be regulated correctly in early Kit$^+$ hematopoietic progenitors. In fact, over 90% of cells in bone marrow fractions highly enriched in HSC and early progenitor cells, known to be Kit$^+$ (SLAM HSC, CD150$^+$,CD48$^-$,CD41$^-$),[16,17] and side population Kit$^+$/Sca1$^+$ cells[18,19] (Figure 3 and *Online Supplementary Figure S4*), show GFP expression. Similarly, most E12.5 fetal liver Kit$^+$ cells, including Kit$^+$/Sca1$^+$/CD11b$^+$ cells, which have the immunophenotype of fetal HSC,[20-22] are also GFP$^+$ (*Online Supplementary Figure S5*). The small proportion of Kit$^+$/GFP- cells might represent cells in which the transgene is sporadically silenced by adjacent sequences due to its integration in a non-optimal chromatin environment.[24]

A proportion of Kit$^-$/Lin$^+$ cells, from both bone marrow and fetal liver, express GFP. Expression of GFP in these cells may either reflect transcription of the transgene in these cells, or persistence of GFP-mRNA and protein produced in Kit$^+$ progenitors upstream to these cells. Although we cannot strictly rule out the possibility that some transcription of the *Kit/GFP* transgene occurs in Kit$^-$ cells, we favor the alternative possibility for several reasons. First, as exemplified in the erythroid lineage in bone marrow and fetal liver, most (60-90%, respectively) Kit$^-$/CD71$^+$ cells are GFP$^-$, whereas the remainder show moderate GFP levels, much lower than those found in Kit$^+$/CD71$^+$ double positive and Kit$^+$/CD71$^-$ early progenitor cells (*Online Supplementary Figure S1*, panel viii; *Online Supplementary Figure S2*, panel viii, compare Kit$^-$/CD71$^+$ cells (blue line) with Kit$^+$ /CD71$^-$ or Kit$^+$/CD71$^+$ cells, red and green, respectively). Similar data were also obtained in the myeloid lineage, as exemplified in bone marrow Kit$^-$/Gr1$^+$ cells (*Online Supplementary Figure S1*, panel xii) and Kit$^-$/CD11b$^+$, as well as in lymphoid B and T cells (*data not shown*). Second, these data are consistent with the results of our previous *in vitro* colony assays, which showed that colonies were homogeneously highly fluorescent during the early days of culture, whereas at later stages the fluorescence was greatly decreased (in erythroid colonies) or fully suppressed (in myeloid colonies[12] see also Figure 1C). Thus, we suggest that, following its very efficient expression in the majority of early Kit$^+$/Lin$^+$ double positive hematopoietic progenitor cells, the transgene is downregulated in most of the more mature cell types, which progressively dilute, over successive cell divisions, the GFPmRNA and protein produced at high levels in early progenitors. Persistence of some GFP, but not Kit, in differentiated cells, may also be explained by the presence in hematopoietic cells of abundant micro RNA mir 221 and 222 (*data not shown*), which target Kit mRNA[25] but not Kit/GFP mRNA, and might thus cause much faster turnover of Kit-mRNA than of Kit/GFP-mRNA.

Transgenic Kit/GFP *is expressed in most hematopoietic stem cells*

The transplantation experiments presented in Figures 1-3 and Tables 2 and 3 indicate that the *Kit/GFP* transgene is expressed in most bone marrow and fetal liver HSC. Unexpectedly, functional HSC are contained within cell fractions that express the transgene at intermediate levels. This is seen both using cells fractionated on the basis of Kit and Kit/GFP expression, and using cell fractions highly enriched in HSC on the basis of stringent SLAM-selection. In the latter case, most of the SLAM-selected GFP$^+$ population are GFP^{int}, and essentially all the HSC capable of long-term reconstitution are in this fraction (Figure 3, Table 3). Interestingly, cells selected less stringently for CD150, which are expected to contain lower proportions of HSC, show a higher proportion of GFP^{high} versus GFP^{int} cells (*data not shown*). These results are in agreement with the expression of Kit/GFP at intermediate levels also in cells with the immunophenotype[18,20-22] of HSC, such as side population Kit$^+$/Sca1$^+$ bone marrow cells and Kit$^+$/Sca1$^+$/CD11b$^+$ fetal liver cells (*Online Supplementary Figures S4 and S5*). In contrast to HSC, progenitors such as CFU-mix and BFU-E express

Kit/GFP at high levels (Table 1), as also do Kit$^+$/Lin$^+$ cells (*Online Supplementary Figures S1 and S2*). The basis for this surprising result is unknown. One possibility is that HSC are inherently less able than more downstream progenitors to transcribe Kit mRNA; this hypothesis may find some support from studies that showed that cell fractions with very low levels of Kit mRNA and protein could give efficient long-term bone marrow reconstitution.[26-28] An alternative, perhaps more credible, hypothesis is that the *Kit* transgene, though capable of transcription in all types of Kit$^+$ cells, may lack some of the regulatory elements that are required for optimal transcription in HSC, but not in other more downstream progenitors (*see below*). The distant 5' regulatory sequences[29] discussed below are possible candidates for this role.

Finally, the observation that within SLAM-selected bone marrow cells[16,17] almost all cells are GFP$^+$, and the majority of HSC are GFP^{int}, has some potential experimental applications. In fact, the precise gating for SLAM selection is a matter of discussion in HSC purification;[30,31] selection for GFP^{int} in conjunction with SLAM purification might thus help to further enrich in HSC the SLAM population.

The Kit/GFP *transgene includes sequences controlling gene expression in different subsets of stem cells*

The mouse *Kit* gene is a large gene extending for more than 100 Kb, and distant upstream deletions affect *Kit* activity.[32] In spite of the large size of *Kit*, our present work demonstrates that a small subset of sequences, including a 6.7 Kb fragment of the promoter and 8 Kb of the first intron, is sufficient for expression in HSC, early progenitors and more mature Kit$^+$ cells, indicating that a minimum complement of sequences essential for expression in HSC is contained in the transgene.

This finding will allow detailed characterization of the specific sequences responsible for Kit regulation in HSC. Upstream distant sequences (lying more than 100 Kb from the transcription start site) contact the Kit promoter-proximal intron region hypersensitive site 5 (HS5)[12] (corresponding to the fragments present in our construct) in early erythroid progenitors grown *in vitro*, which express Kit.[29] These sequences are then displaced upon induced differentiation and Kit repression.[29] The transcription factor GATA1 may replace the GATA2 factor bound to a sequence in HS5 and may thus be instrumental in these chromatin changes and in repression.[33] These data suggest that upstream sequences are involved in Kit activity,[33] although, unexpectedly, large transgenic constructs including them are very poorly expressed in bone marrow cells.[34]

Notably, the upstream sequences are dispensable in our construct. It is possible that these sequences might be involved in a different level of regulation than the promoter and first intron sequences, perhaps controlling chromatin structure and accessibility of the proximal sequences to the transcriptional machinery, rather than directly affecting their activity. Alternatively, upstream sequences might play specific roles in different subsets of stem/progenitor cells. Addition of fragments of the upstream sequences to the presently defined transgenic sequences will allow their role to be tested in HSC versus downstream progenitors.

A few transgenic constructs have previously been shown to be expressed in HSC and/or early hematopoietic progenitors. These are based on regulatory sequences corresponding to genes encoding transcription factors such as *Scl-Tal1*,[35,36] *Runx1*[37] and possibly *GATA-2*[38] and the membrane proteins Sca1[39] and CD34.[40] All of these genes are coexpressed with Kit; SCL-Tal1 and Gata2 are thus potential regulators of Kit transcription.

A direct comparison between our results and those reported above is difficult, because of the differences in the reporter genes used, in the expression of the transgene in non-HSC populations and in the numbers and type of coinjected supporting cells used for transplantation. Overall, the data presented in Table 2 and Figure 2 show that the enrichment obtained with Kit$^+$/GFP^{int} cells is of a similar order to that reported by Silberstein *et al.* and Ma *et al.*[36,39] Notably, as in our study, in these works the transgene was expressed in a large number of non-HSC/progenitor cells, a fact presumably reflecting the persistence in downstream cells of reporter mRNA/protein accumulated in earlier progenitors.

Expression of the *Scl/Tal1* and *Runx1* transgenes in HSC and other hematopoietic cells requires a subset of Scl/Tal1, GATA-2 and ETS binding sites;[36,37,41] similar sites are also present in the Kit HS regions, although with different reciprocal arrangements than in the *Scl/Tal1* gene. Further comparison and identification of common regulatory motifs between other stem cell genes and *Kit* will be important to define a critical regulatory network of gene expression in HSC.

The *Kit* gene is active in several types of stem cells, and its defects affect multiple cell lineages[4] pointing to Kit being a pleiotropic, though cell-type-restricted, molecule. It would, therefore, be particularly relevant to have a mouse model in which the transgene efficiently recapitulates the activity of the endogenous molecule. The present *Kit/GFP* transgene, in addition to being expressed in HSC, is also active in primordial germ cells and subsets of spermatogonia,[12,23] in early blastocysts grown *in vitro* and in embryonic stem cells (*unpublished data: SO, LC and S. Dolci*), and in a population of *cardiac stem cells*.[9] Thus, Kit might be regulated by subsets of common transcriptional programs in different stem cells. Targets of the responsible transcription factors must lie within the relatively short regions encompassed by our transgene. Finally, transgene expression in HSC/progenitors has enabled monitoring of bone marrow cellular homing to muscle[42] and heart (*FC, unpublished data*) following tissue damage. Thus, our transgenic mouse lines provide a powerful tool to shed light on aspects of cell trafficking and tissue regeneration, and on relationships with HSC.

Authorship and Disclosures

FC performed bone marrow and fetal liver *in vitro* clonogenic assays and *in vivo* long-term repopulation assays, and statistical analysis of the experiments; LC, GL and SC performed cytofluorimetric analysis and sorting; FB helped in transplantation experiments; MCM and SO designed experiments and wrote the paper.

The authors reported no potential conflicts of interest.

References

1. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006;125:315-26.
2. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 2006;125:301-13.
3. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 2006;441:349-53.
4. Broudy VC. Stem cell factor and hematopoiesis. Blood 1997;90:1345-64.
5. Ikuta K, Weissman IL. Evidence that hematopoietic stem cells express mouse c-kit but do not depend on steel factor for their generation. Proc Natl Acad Sci USA 1992;89:1502-6.
6. Bernex F, De Sepulveda P, Kress C, Elbaz C, Delouis C, Panthier JJ. Spatial and temporal patterns of c-kit-expressing cells in WlacZ/+ and WlacZ/WlacZ mouse embryos. Development 1996;122:3023-33.
7. Das AV, James J, Zhao X, Rahnenfuhrer J, Ahmad I. Identification of c-Kit receptor as a regulator of adult neural stem cells in the mammalian eye: interactions with Notch signaling. Dev Biol 2004;273: 87-105.
8. Beltrami AP, Barlucchi L, Torella D, Baker M, Limana F, Chimenti S, et al. Adult cardiac stem cells are multipotent and support myocardial regeneration. Cell 2003;114:763-76.
9. Messina E, De Angelis L, Frati G, Morrone S, Chimenti S, Fiordaliso F, et al. Isolation and expansion of adult cardiac stem cells from human and murine heart. Circ Res 2004;95:911-21.
10. Kimura Y, Jones N, Kluppel M, Hirashima M, Tachibana K, Cohn JB, et al. Targeted mutations of the juxtamembrane tyrosines in the Kit receptor tyrosine kinase selectively affect multiple cell lineages. Proc Natl Acad Sci USA 2004;101:6015-20.
11. Dolci S, Williams DE, Ernst MK, Resnick JL, Brannan CI, Lock LF, et al. Requirement for mast cell growth factor for primordial germ cell survival in culture. Nature 1991;352: 809-11.
12. Cairns LA, Moroni E, Levantini E, Giorgetti A, Klinger FG, Ronzoni S, et al. Kit regulatory elements required for expression in developing hematopoietic and germ cell lineages. Blood 2003;102:3954-62.
13. Goodell MA, Brose K, Paradis G, Conner AS, Mulligan RC. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. J Exp Med 1996;183: 1797-806.
14. Schneider E, Pollard H, Lepault F, Guy-Grand D, Minkowski M, Dy M. Histamine-producing cell-stimulating activity. Interleukin 3 and granulocyte-macrophage colony-stimulating factor induce de novo synthesis of histidine decarboxylase in hemopoietic progenitor cells. J Immunol 1987; 139:3710-7.
15. Magli MC, Iscove NN, Odartchenko N. Transient nature of early haematopoietic spleen colonies. Nature 1982;295:527-9.
16. Kiel MJ, Yilmaz OH, Iwashita T, Yilmaz OH, Terhorst C, Morrison SJ. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. Cell 2005;121:1109-21.
17. Kiel MJ, He S, Ashkenazi R, Gentry SN, Teta M, Kushner JA, et al. Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU. Nature 2007; 449:238-42.
18. Camargo FD, Chambers SM, Drew E, McNagny KM, Goodell MA. Hematopoietic stem cells do not engraft with absolute efficiencies. Blood 2006;107:501-7.
19. Weksberg DC, Chambers SM, Boles NC, Goodell MA. CD150 side population cells represent a functionally distinct population of long-term hematopoietic stem cells. Blood 2008; 111:2444-51.
20. Morrison SJ, Hemmati HD, Wandycz AM, Weissman IL. The purification and characterization of fetal liver hematopoietic stem cells. Proc Natl Acad Sci USA 1995;92:10302-6.
21. Sanchez MJ, Holmes A, Miles C, Dzierzak E. Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. Immunity 1996;5: 513-25.
22. Taoudi S, Morrison AM, Inoue H, Gribi R, Ure J, Medvinsky A. Progressive divergence of definitive haematopoietic stem cells from the endothelial compartment does not depend on contact with the foetal liver. Development 2005;132:4179-91.
23. Filipponi D, Hobbs RM, Ottolenghi S, Rossi P, Jannini EA, Pandolfi PP, Dolci S. Repression of kit expression by Plzf in germ cells. Mol Cell Biol 2007;27: 6770-81.
24. Pikaart MJ, Recillas-Targa F, Felsenfeld G. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev 1998;12:2852-62.
25. Felli N, Fontana L, Pelosi E, Botta R, Bonci D, Facchiano F, et al. MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. Proc Natl Acad Sci USA 2005; 102:18081-6.
26. Doi H, Inaba M, Yamamoto Y, Taketani S, Mori SI, Sugihara A, et al. Pluripotent hemopoietic stem cells are c-kit<low. Proc Natl Acad Sci USA 1997;94:2513-7.
27. Klarmann K, Ortiz M, Davies M, Keller JR. Identification of in vitro growth conditions for c-Kit-negative hematopoietic stem cells. Blood 2003;102:3120-8.
28. Kotton DN, Fabian AJ, Mulligan RC. A novel stem-cell population in adult liver with potent hematopoietic-reconstitution activity. Blood 2005; 106:1574-80.
29. Jing H, Vakoc CR, Ying L, Mandat S, Wang H, Zheng X, Blobel GA. Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. Mol Cell 2008; 29:232-42.
30. Kiel MJ, Yilmaz OH, Morrison SJ. CD150- cells are transiently reconstituting multipotent progenitors with little or no stem cell activity. Blood 2008;111:4413-4.
31. Chambers SM, Weksberg DC, Goodell MA. The CD150high compartment is not the exclusive reservoir of LT-HSCs within the bone marrow. Blood 2008;111:4414-5.
32. Berrozpe G, Timokhina I, Yukl S, Tajima Y, Ono M, Zelenetz AD, Besmer P. The W(sh), W(57), and Ph Kit expression mutations define tissue-specific control elements located between -23 and -154 kb upstream of Kit. Blood 1999;94:2658-66.
33. Munugalavadla V, Dore LC, Tan BL, Hong L, Vishnu M, Weiss M J, Kapur R. Repression of c-Kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. Mol Cell Biol 2005;25:6747-59.
34. Berrozpe G, Agosti V, Tucker C, Blanpain C, Manova K, Besmer P. A distant upstream locus control region is critical for expression of the Kit receptor gene in mast cells. Mol Cell Biol 2006;26:5850-60.
35. Sanchez MJ, Bockamp EO, Miller J, Gambardella L, Green AR. Selective rescue of early haematopoietic progenitors in Scl(-/-) mice by expressing Scl under the control of a stem cell enhancer. Development 2001;128: 4815-27.
36. Silberstein L, Sanchez MJ, Socolovsky M, Liu Y, Hoffman G, Kinston S, et al. Transgenic analysis of the stem cell leukemia +19 stem cell enhancer in adult and embryonic hematopoietic and endothelial cells. Stem Cells 2005;23:1378-88.
37. Nottingham WT, Jarratt A, Burgess M, Speck CL, Cheng JF, Prabhakar S, et al. Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. Blood 2007;110:4188-97.
38. Kobayashi-Osaki M, Ohneda O, Suzuki N, Minegishi N, Yokomizo T, Takahashi S, et al. GATA motifs regulate early hematopoietic lineage-specific expression of the Gata2 gene. Mol Cell Biol 2005;25:7005-20.
39. Ma X, Robin C, Ottersbach K, Dzierzak E. The Ly-6A (Sca-1) GFP transgene is expressed in all adult mouse hematopoietic stem cells. Stem Cells 2002;20:514-21.
40. Okuno Y, Huettner CS, Radomska HS, Petkova V, Iwasaki H, Akashi K, et al. Distal elements are critical for human CD34 expression in vivo. Blood 2002; 100:4420-6.
41. Landry JR, Kinston S, Knezevic K, de Bruijn MF, Wilson N, Nottingham WT, et al. Runx genes are direct targets of Scl/Tal1 in the yolk sac and fetal liver. Blood 2008;111:3005-14.
42. Musarò A, Giacinti C, Borsellino G, Dobrowolny G, Pelosi L, Cairns L, et al. Stem cell-mediated muscle regeneration is enhanced by local isoform of insulin-like growth factor 1. Proc Natl Acad Sci USA 2004;101:1206-10.