Distal elements are critical for human CD34 expression in vivo

Yutaka Okuno, Claudia S. Huettner, Hanna S. Radomska, Victoria Petkova, Hiromi Iwasaki, Koichi Akashi, and Daniel G. Tenen

The elements regulating gene expression in hematopoietic stem cells are still poorly understood. We previously reported that a 141-kilobase (kb) human CD34 transgene confers properly regulated human CD34 expression in transgenic mice. A construct with only the human CD34 promoter and 3' enhancer region is not sufficient, suggesting that critical distal elements are necessary for expression of the human CD34 gene. To further localize such elements, we analyzed deletion constructs of the human CD34 gene and evaluated their function in transgenic mice. Constructs harboring as little as 18 kb of 5' and 26 kb of 3' human CD34 flanking sequence conferred human expression in tissues of transgenic mice with a pattern similar to that of the 141-kb human transgene. In contrast, a construct harboring 10 kb of 5' and 17 kb of 3' human CD34 flanking sequence gave no expression. These data demonstrate that regions between 10 to 18 kb upstream and/or 17 to 26 kb downstream of the human CD34 gene contain critical elements for human CD34 expression in vivo.

Further functional analysis of these regions in transgenic mice will be crucial for understanding CD34 gene expression in hematopoietic stem and progenitor cells. (Blood. 2002;100:4420-4426)

Introduction

Human CD34 is a transmembrane protein expressed on the surface of stem and early hematopoietic cells1-5; expression occurs on a low percentage of bone marrow cells and declines during hematopoietic differentiation. Therefore, human CD34 represents a good model for understanding stem cell gene regulation. Furthermore, identification of human CD34 regulatory elements may allow identification of transcription factors necessary for stem cell expression, as well as facilitate regulated expression of heterologous genes in hematopoietic stem cells.

Initial analysis of the murine CD34 gene demonstrated that the relative conservation with human CD34 decreased from the carboxyl toward the amino terminus of the coding region.6-8 In addition, in contrast to many other hematopoietic genes,9-11 there is no conservation of sequence or regulatory elements in the proximal promoter or 5' untranslated region.12-18 These results suggested that distal elements might play a more important role in CD34 regulation. Indeed, we reported that a human CD34 minigene including a 4.5-kb CD34 5' promoter fragment and a 3' enhancer active in transient transfections12 failed to express human CD34 in transgenic mice.19

In order to delineate the important distal control elements of human CD34, we used the technique of transgenic analysis using large genomic clones, which has been successful in expression of a number of hematopoietic genes.20-22 It has been demonstrated that transgenic mouse models made with yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or P1-based artificial chromosome (PAC) clones bearing large genomic sequences are powerful tools for reproducing proper expression of human genes as well as delineating critical distal elements necessary for proper gene regulation. Therefore, we isolated a human PAC clone including 110 kb of 5' upstream flanking sequence and 26 kb of 3' downstream flanking sequence.19 Transgenic mice harboring a 141-kb fragment from this PAC expressed human CD34 mRNA and protein in nonhematopoietic tissues in a manner similar to that of murine CD34.23 (See the accompanying paper by Radomska et al24 that describes the use of these CD34 regulatory elements to express heterologous genes in early hematopoietic cells, beginning on page 4410.) However, consistent with reports that murine CD34 may not be highly expressed on stem cells functionally defined by transplantation to lethally irradiated recipients,25-27 transgenic mice harboring this PAC express human and murine CD34 differently in the bone marrow. Only human and not murine CD34 targets true repopulating hematopoietic stem cells in these transgenic mice, while common myeloid progenitors (CMP), granulocyte macrophage progenitors (GMP), common lymphoid progenitors (CLP), and thymic T-cell progenitors express both human and murine CD34.28,29 The elements mediating the difference in expression between human and murine CD34 are still unknown. Therefore, in this report, we made a number of transgenic lines with constructs containing various lengths of 5' upstream and 3' downstream human CD34 flanking sequences in order to locate critical cis distal elements.

Materials and methods

Extraction and purification of PAC clones containing the entire human CD34 genomic sequence

From a human genomic library, 3 clones were isolated using human CD34 specific primers (Genome Systems, St Louis, MO). PAC88L2 included 110 kb of 5' and 24.1 kb of 3' flanking sequence. PAC7H11 included 23 kb of 5' and 95 kb of 3' sequences, and PAC54A19 extended 18 kb 5' and 43 kb 3' of the...
human CD34 coding exons (Figure 1). PAC clones were cultured in LB (Luria-Bertani) medium with 50 μg/mL kanamycin. PAC DNA was extracted with a Plasmid Purification Kit (QIAGEN, Valencia, CA) and resuspended in 1 mM Tris [tris(hydroxymethyl)aminomethane] HCl (pH 7.6), 0.1 mM EDTA (ethylene diamine tetra-acetic acid; 1° TE). The clones were mapped by restriction digests followed by Field Inversion Gel Electrophoresis (FIGE, Biorad, Hercules, CA). In addition, direct sequencing and comparison with the human genome sequence demonstrated that PAC54A19 extended from 18 368 bp upstream to 67 420 bp downstream from the transcription start site8,30 or 42 638 bp downstream of the polyadenylation site.7,8

Transgenic mice

To construct human CD34 transgensics that include 48 kb or 31 kb of 5° upstream genomic regions, 19 the top 2 fragments were derived from PAC88L2, a 160-kb PAC clone extending 110 kb upstream of the transcription start site.19 The 3rd fragment represents Mulf-digested PAC7H11, and the bottom 3 were derived from PAC54A19. Shown on the left are the restriction enzymes used to digest the PAC clones. Restriction sites that cut in the PAC vector (Pvu and Mulf) are indicated by dotted lines. A unique Mulf site is present in PAC88L2 at a position 31 kb upstream of the transcription start site, and a Pvu site is present in PAC54A19 25.7 kb 3° of the polyadenylation site. PAC88L2 extends to position 24.1 kb 3° of the polyadenylation site. Shown as short thick vertical lines are the 8 CD34 coding exons, which encompass 25 kb.8

The limits of the human CD34 sequence in PAC54A19 were determined by DNA sequencing. Shown on the right is the number of independent founder lines derived from each construct and the number that was observed to express human CD34 RNA and protein. Only the last construct failed to express human CD34.

Figure 1. Diagrams of human CD34 genomic fragments used in transgenic mice. The top 2 fragments were derived from PAC88L2, a 160-kb PAC clone extending 110 kb upstream of the transcription start site.19 The 3rd fragment represents Mulf-digested PAC7H11, and the bottom 3 were derived from PAC54A19. Shown on the left are the restriction enzymes used to digest the PAC clones. Restriction sites that cut in the PAC vector (Pvu and Mulf) are indicated by dotted lines. A unique Mulf site is present in PAC88L2 at a position 31 kb upstream of the transcription start site, and a Pvu site is present in PAC54A19 25.7 kb 3° of the polyadenylation site. PAC88L2 extends to position 24.1 kb 3° of the polyadenylation site. Shown as short thick vertical lines are the 8 CD34 coding exons, which encompass 25 kb.8

The limits of the human CD34 sequence in PAC54A19 were determined by DNA sequencing. Shown on the right is the number of independent founder lines derived from each construct and the number that was observed to express human CD34 RNA and protein. Only the last construct failed to express human CD34.

RNA extraction and Northern blot analysis

RNA from tissues of transgenic mice were homogenized with 4 M guanidine thiocyanate solution and put onto a cushion of cesium trifluoroacetate (Amersham Pharmacia Biotech, Piscataway, NJ) diluted with 1° TE to adjust the density to 1.60. After ultracentrifugation, RNAs were dissolved in autoclaved water. From each tissue, 10 μg of RNA was electrophoresed in 1° agarose 2.2 M formaldehyde gels. RNAs were transferred to nylon membranes (Biotrans Plus) with 20× SSC and immobilized with a UV cross-linker (Stratalinker). For human CD34 Northern blots, a 0.5 kb HindIII-BamHI fragment from the 3° untranslated region of the human CD34 cDNA was used as a probe.19 For detection of murine CD34 RNA, a 500-bp fragment derived from the 3° region of the murine CD34 cDNA, which was not conserved between human and murine CD34, was amplified by PCR using primers 5’-AGAAGAGATGGGTGGAGGAT-3’ and 5’-GGATCCCAAGAAGATGGATGTA-3’ were used to amplify a fragment 7 kb 3° downstream from exon 8, the most 3° exon of human CD34; 5’-ATGGGGTTAGAGCCTGAAGGTGTT-3’ and 5’-TGGATTCGATTATTTTGGTTTCCC-3’ to amplify a 500-bp genomic fragment located 38 kb 5° of the transcription start site; and GTGCTTTTACGGAGGAGGTTTTA-3’ and 5’-TAAAGGCTCAAGGGGTGGACCTC-3’ to amplify a genomic DNA fragment located −11 kb 5° of the transcription start site. PCRs were performed with Taq DNA polymerase (BRL, Rockville, MD), with cycles of 95° C for 30 sec; 58° C for 30 sec; and 72°C for 2 minutes for 35 cycle after 5 minutes of denaturation at 95°C.

FACS analysis of bone marrow cells from transgenic mice

Bone marrow cells were flushed from femurs and tibias of 2- to 3-month-old transgenic mice, suspended in phosphate buffered saline (PBS) with 2% fetal bovine serum, and filtered through 40 μm nylon mesh (Becton Dickinson, Franklin Lakes, NJ). One million cells of this single cell suspension were stained with phycoerythrin (PE)-conjugated anti-human CD34 antibody or fluorescein...
isothiocyanate (FITC)-conjugated anti–human CD34 antibody (HPCA-2, Becton Dickinson) and FITC-conjugated anti–mouse CD34, Gr-1, Sca-1 (BD Pharmering, San Diego, CA), B220, Mac-1 (Caltag Laboratories, San Francisco, CA) antibodies, or PE-conjugated anti–c-kit, CD4, CD8, TER119 antibodies (BD Pharmering). Cells were washed with PBS twice and resuspended in PBS with 2 μg/mL propidium iodide (PI, Sigma, St Louis, MO). Fluorescence-activated cell-sorter (FACS) analysis was done with a FACScan (Becton Dickinson, San Jose, CA).

**Quantitation of human and murine RNA expression by real-time PCR**

Multiplex PCR with amplification of 18S RNA in the same tube for quantitation of human and murine CD34 expression, TaqMan analysis, and subsequent calculations were performed with an ABI Prism 7700 sequence detection system (Perkin Elmer, Foster City, CA), which detects the signal from a fluorogenic internal probe. For each sample, 100 ng cDNA was subjected to PCR with primers 5'-AAAACATAACACCTTAGCACCTTGGAA-3' and 5'-GAATTGACTGTCGTTTCGTAAGC-3' for human CD34 expression, according to protocols provided by the manufacturer of the Taqman system (ABI, Foster City, CA). The sequence of the double-labeled oligonucleotide used as probe was FAM-CCCTGTGTCTCAACA TGGCAA TGAGGCC-TAMRA. Amplification of 18S RNA was performed in the same reaction tubes as an internal standard with an alternatively labeled probe to distinguish its product from that derived from CD34 RNA (multiplex PCR). For murine CD34, the primers were 5'-TCTCTCTCGCTCGTGAAGCCA-3' and 5'-CCTGGGCAACCTCACTTCC-3', and the sequence of the double-labeled oligonucleotide used as probe was FAM-AAAAGGAGAATCAATGCTGTAAGCC-TAMRA. Experiments were performed in quadruplicates for each standard and bone marrow sample.

**Colony assay**

Bone marrow cells were isolated from femurs and tibias from 4-month-old mice, suspended in PBS with 2% fetal bovine serum, and stained with PE-conjugated anti–human CD34 antibody (HPCA-2) and FITC-conjugated anti–mouse CD34 antibody (Pharmering). Cells were washed with PBS twice and resuspended in PBS with 2 μg/mL PI. Pl-Positive cells of human CD34+/mouse CD34-, human CD34+/mouse CD34+, human CD34-/mouse CD34+, and human CD34-/mouse CD34- populations were sorted using a high-speed cell sorter (Moffo-MLS, Cytomation, Fort Collins, CO). Each cell population was then inoculated in Methocult H4100 (Stem Cell Technologies, Vancouver, BC) supplemented with 30% fetal calf serum, 1% bovine serum albumin, 20 ng/mL stem cell factor, 20 ng/mL interleukin-3 (IL-3), 10 ng/mL IL-11, 10 ng/mL granulocyte-macrophage colony-stimulating factor, 1 U/mL erythropoietin, 10 ng/mL thrombopoietin (R&D Systems, Minneapolis, MN), 2 mM L-glutamine (Stem Cell Technologies), 100 U/mL penicillin, and 100 μg/mL streptomycin (GIBCO BRL, Carlsbad, CA). Colony numbers were counted after 5-6 days of culture.

**Results**

**Generation of transgenic mice with constructs harboring human CD34 genomic sequences of various lengths**

In order to characterize the regulatory elements of the human CD34 gene, we generated several constructs that included, in addition to all exons and introns, various lengths of 5' upstream and 3' downstream flanking sequences of the hCD34 gene (Figure 1). PAC7H11 and PAC54A19 included all exons and introns of human CD34 gene, 23 kb and 18 kb 5' upstream flanking sequence, and 95 kb and 43 kb 3' downstream sequence, respectively. Constructs with 48 kb and 31 kb of 5' upstream flanking sequences were derived from PAC88L2 DNA using different restriction enzyme digests. The construct with 10 kb of 5' upstream and 17 kb of 3' downstream flanking sequences was derived from PAC54A19 DNA following digestion with XhoI. We used FIGE to analyze these fragments, which can be greater than 100 kb in size, and to separate them from the PAC vector DNA. The yields of transgenic mice with DNA fragments greater than 100 kb were low. The number of transgenic mouse founders was inversely correlated with the size of the fragment (Figure 1).

To confirm the structure of the human CD34 transgene in transgenic mice, we performed Southern blot analysis of the human CD34 transgene with a probe bearing all of exon 1 and intron 1, as well as PCR corresponding to 4 different regions of the human CD34 genomic sequence, as described as “Materials and methods.” The 5-kb EcoRI fragment detected by Southern blot with the exon 1/intron 1 probe was intact in all transgenic mouse founders (data not shown). The PCR corresponding to the 3' human CD34 genomic sequence was consistent with the transgenes including 5 kb of sequences downstream of the transcription termination site. Thus, the transgenic mice include all human CD34 exons and introns. In order to check the status of the 5' upstream flanking sequence, 2 different regions were amplified by PCR. All transgenic lines except for the line with only 10 kb of 5' upstream flanking sequence included the region corresponding to the sequence 11 kb upstream (data not shown). Only one line, containing a construct including 48 kb of 5' upstream flanking sequence, was positive for PCR using primers corresponding to 38 kb upstream of the transcription start site. These results are consistent with the transgenes harboring the structures shown in Figure 1.

**Transgenic mice with human CD34 constructs ranging from 18 to 48 kb of 5' upstream flanking sequences express human CD34 RNA**

In order to delineate the important cis regulatory elements in human CD34, we assessed human CD34 RNA expression of each founder by Northern blot analysis. All constructs including 18 to 48 kb of 5' upstream sequence confer human CD34 expression in various tissues. Figure 2 demonstrates the pattern of human and murine CD34 expression in a transgene with PAC54A19, which includes 18 kb of 5' upstream and 43 kb of 3' downstream flanking genomic sequence. High levels of RNA were detected in heart, brain, kidney, spleen, and lung, similar to that observed with a larger, 141-kb construct.23,24,29 The expression pattern of human CD34 was very similar to that of murine CD34, with the exception of lung. This expression pattern was observed in all transgenic lines shown in Figure 1, with the exception of the XhoI-digested PAC54A19 construct, containing 10 kb of upstream and 17 kb of downstream flanking sequences. Therefore, these results indicate that elements between 18 kb upstream and 26 kb downstream are sufficient for directing human CD34 expression in vivo.

We also asked whether the expression of human CD34 was dependent on copy number. While the number of transgenic lines was relatively low, the level of human CD34 RNA in heart tissue of the 4 transgenic founders with PAC54A19, which includes 18 kb of 5' upstream and 26-43 kb of 3' downstream flanking genomic sequences, appeared to be gene copy number dependent (Figure 2B, ○) and similar to that of a transgenic mouse including 141 kb of human CD34 genomic DNA, PAC88L2 (Figure 2B, ●).
we measured human CD34 surface expression by fluorescence-activated cytometry. Transgenic mice made with all of the constructs except for the one with 10 kb of upstream sequence expressed human CD34 protein on the surface of 2% to 4% of bone marrow cells (Figure 3). Of the 2% to 4% of bone marrow cells that expressed human CD34, 42% to 79% also expressed early markers such as murine CD34, Sca-1, and c-Kit (Figure 3). Conversely, the erythroid differentiation marker TER119 was not coexpressed with the human CD34/H11001 cell population. Gr-1, Mac-1, and CD8 antigens were expressed in a smaller fraction of the human CD34/H11001 cells, indicating that these cells might represent progenitors. We did observe significant coexpression of human CD34 with B220, consistent with the findings of others and expression of human CD34 in pro-B cells.29,33,34 The expression patterns of a second founder line made with this same construct (MluI-digested PAC54A19) were almost identical (data not shown).

In order to compare the expression of different constructs in the bone marrow, we measured expression of human CD34 antigen from transgenic mice made with human CD34 genomic DNA constructs harboring different amounts of 5′ flanking sequence. As shown in Figure 4A, constructs with 48, 31, 23, or 18 kb of 5′ upstream and either 26 kb or 95 kb of 3′ downstream flanking sequence confer expression of human CD34 antigen in bone marrow cells of transgenic mice. Therefore, sequences critical for in vivo expression of the human CD34 transgene are located from 18 kb to 26 kb. A similar pattern was observed with the 70-kb PvuI-digested PAC54A19 fragment (Figure 4B). Most of the human CD34+ cells also expressed c-Kit, and in addition, 26% to 65% of the human CD34+ cells coexpressed murine CD34 or Sca-1, again demonstrating coexpression of human CD34 with 3 markers associated with hematopoietic stem cell and/or progenitor populations. We also compared the levels of human CD34 RNA in

Figure 2. A human CD34 transgene containing 18 kb of 5′ upstream and 43 kb of 3′ downstream sequences (PAC54A19) expresses RNA in a manner similar to the endogenous murine CD34 gene in transgenic mice. (A) Northern blot analysis of 10 μg RNA isolated from tissues of nontransgenic (−) and transgenic (+) littermates. The same blot was hybridized with a probe specific for human CD34 (3′ untranslated region), for murine CD34 (3′ untranslated region), and subsequently to GAPDH as described in “Materials and methods.” A similar pattern of expression was observed in all 4 founder lines derived from PAC54A19, including those digested with PvuI, as well as in the founder lines of the larger (−48, −31, and −23 kb) constructs shown in Figure 1. (B) Human CD34 RNA expression in transgenic mice is copy number–dependent. The expression of human CD34 RNA in heart from the 4 independent PAC54A19 founders (○) was determined by ImageQuant (Molecular Dynamics, Sunnyvale, CA) and standardized to GAPDH expression. The gene copy number was determined by Southern blot analysis. ● represents one transgenic founder containing 141 kb of human CD34 genomic DNA (PAC88L2), as previously reported.19

Figure 3. The human CD34 PAC54A19 (18 kb 5′ upstream) is coexpressed with early hematopoietic surface markers in bone marrow and down-regulated with differentiation. Bone marrow cells were stained with anti-human CD34 antibody and a second marker as indicated and subjected to fluorescence cytometry. The vertical axes represent human CD34 fluorescence in all panels. The percent of positive cells in each quadrant is noted. Dead cells were gated out following staining with PI. The data shown were obtained from one of the MluI-digested PAC54A19 founders. A similar pattern was observed in the 3 other PAC54A19 founder lines, as well as in the larger (−48, −31, and −23 kb) constructs shown in Figure 1.
bone marrow with that of the endogenous murine CD34 gene for one of the constructs, MluI/PvuI-digested PAC88L2. Using quantitative real time PCR, we found that the human construct expressed at levels 3-4 times that of the endogenous murine CD34 gene, when corrected for transgene copy number (data not shown).

**Upstream sequences extending from −18 to −10 kb and/or downstream from +17 to +26 are required for human CD34 expression in transgenic mice**

In order to further delineate the regions critical for in vivo expression, we made one additional deletion construct (Figure 1). Digestion of PAC54A19 with XhoI allowed us to purify a transgenic fragment that included up to 10 kb of upstream and 17 kb of downstream human CD34 sequences. In contrast to all of the other constructs shown in Figure 1, deletion of sequences from −18 to −10 and/or from +17 to +26 led to complete loss of RNA expression in multiple transgenic lines (Figure 5). In order to exclude the possibility that this construct was still expressing human CD34 on bone marrow progenitors, we performed fluorescence-activated cytometry. In contrast to what was observed with PAC54A19 and larger constructs (Figures 3 and 4), 3 different transgenic mouse founders with 10 kb of 5′ upstream and 17 kb of 3′ downstream flanking sequence failed to express human CD34 protein in bone marrow cells (Figure 6). These data indicate that the critical elements for human CD34 expression are located between −18 and −10 kb and/or from +17 to +26 kb.

**The human CD34 transgene is expressed in the most immature progenitor cells**

The data in Figures 3 and 4 demonstrate that human CD34 is expressed in cells with an immature phenotype. Therefore, we wanted to test whether human CD34 was selectively expressed on cells that function as immature progenitors. As shown in Figure 3, hematopoietic cells from MluI-digested PAC54A19, which includes 18 kb of upstream and 43 kb of downstream flanking sequences, can be subdivided into 4 populations based on expression of human and murine CD34; human CD34+/murine CD34−; human CD34+/murine CD34+; hCD34+/murine...
CD34; and hCD34/murine CD34. To assess the function of these 4 cell populations, we sorted the cells and performed colony-forming assays with each population. As shown in Table 1, all CD34 cell populations gave rise to colony-forming units in culture (CFU-Cs), whereas no CFU activity was observed in human CD34/murine CD34 cells. The majority of CFU-Cs, as well as CFU-granulocytes, erythrocytes, monocyte/macrophages, and megakaryocytes (CFU-GEMMs, the most immature multipotent progenitor cells), were found in the human CD34/murine CD34 cell population, whereas a relatively small number of GEMMs were observed to be human CD34/murine CD34. In contrast, no GEMM colonies were observed in cells that lacked expression of human CD34. In summary, all very immature CFUs, the CFU-GEMMs, expressed human CD34, whereas some of the more mature CFUs, the CFU-Cs, did not, consistent with expression of human CD34 in the earliest progenitors.

**Discussion**

We and others have studied the regulation of the CD34 gene as a means of understanding gene regulation in hematopoietic stem cells, identifying the promoter region of human and murine CD34,12,16,17,30 a cell type–specific enhancer element at the 3’ end of the human gene,12,19 and an enhancer at the 5’ end of the murine gene.14 DNaseI hypersensitivity assays spanning a 43-kb region extending from 8.5 kb upstream to 9.5 kb downstream of the gene identified a number of CD34–specific hypersensitive sites, including several within 3 kb of the transcription start and termination sites, as well as in introns 1 and 4.13,19 In addition, a number of transcription factors, including c-myb,35,36 ets-2,36 MZF-1,15,37 Sp1 and Sp3,17 and NFY18 were demonstrated to bind to their cognate sites and modulate the activity of the human and/or murine promoters and/or 5’ untranslated regions.

However, in most vertebrate genes, several kbs of promoter region often fail to confer proper gene expression,22 as they lack distal regulatory elements located far upstream,28,39 within introns,40 or 3’ of the coding region.31,42 This was the case for human CD34: a minigene construct including 4.5 kb of human CD34 promoter as well as the 3’ enhancer, which included the previously identified DNasel hypersensitive sites described above, failed to express human CD34 mRNA in transgenic mice.19

In recent years, transgenic mice models made with YAC, BAC, or PAC clones bearing large genomic sequences have been shown to be powerful tools for studying human gene regulatory elements in transgenic mice.20,22 With this in mind, we generated transgenic mice with a large PAC clone that included 141 kb encompassing the entire human CD34 genomic region. This PAC expressed high levels of human CD34 mRNA and protein,23,24 demonstrating that other distal elements were critical for human CD34 gene expression in transgenic mice. To identify such critical distal elements, we analyzed a series of deletion constructs of the human CD34 gene in transgenic mice. As little as 18 kb upstream and 26 kb downstream sequences were sufficient for expression of human CD34 DNA and protein in several tissues, including bone marrow, at levels with a specificity similar to that of transgenic mice made with the 141-kb human CD34 genomic clone.29 The expression levels of the 18-kb construct appear to be gene copy number dependent, suggesting that this construct includes all the critical elements needed for human CD34 expression in various tissues, including bone marrow cells. In contrast, further deletion of both an upstream and downstream region eliminated human CD34 expression in all tissues examined. These data strongly suggested that distal critical elements were critical for human CD34 expression in transgenic mice.

**Table 1. Colony assays for each cell population**

<table>
<thead>
<tr>
<th>Cell population</th>
<th>CFU-GEMM</th>
<th>CFU-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>hCD34 mCD34+</td>
<td>9.3 ± 6.1</td>
<td>77.3 ± 2.3</td>
</tr>
<tr>
<td>hCD34 mCD34+</td>
<td>160.0 ± 69.3</td>
<td>1626.7 ± 46.2</td>
</tr>
<tr>
<td>hCD34 mCD34+</td>
<td>0.0 ± 0.0</td>
<td>130.0 ± 43.6</td>
</tr>
<tr>
<td>hCD34 mCD34+</td>
<td>0.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
</tbody>
</table>

Colony numbers are from 10 000 cells of each cell population. Figures are given as the mean ± SEM of 3 different experiments.
cis elements are located between −18 and −10 kb and/or +17 kb to +26 kb downstream of the gene. With both the larger 141-kb PAC88L2 transgene and the smaller PAC54A19 constructs, human CD34 protein was expressed on 2% to 4% of total bone marrow cells, a slightly lower frequency than that of the murine CD34 gene (5% to 7%). Analysis of human CD34 expression in progenitors demonstrated that all CFU-GEMMs expressed human CD34. These data strongly suggested that critical distal cis elements necessary for proper expression of human CD34 in bone marrow cells are contained within the PAC88L2 (141 kb) and PAC54A19 (86 kb) transgenes and that important cis elements are located within the −18 to −10 kb and/or +17 to +26 regions. There may well be additional key elements lying within the introns of the hCD34 gene, although to date we have been unable to detect such elements.

Further delineation of the regulatory regions will require smaller deletions and assessment of function in transgenic animals, as well as testing the ability of smaller regulatory elements to complement the ability of the human CD34 promoter to express human CD34 in vivo. The results of such studies will lead to further insights into the regulation of stem cell genes, including identification of the transcription factors mediating this specificity, as well as potentially provide the means to develop vectors that can target heterologous genes to stem cells.  

Acknowledgments
We thank Maris Fenys for expert animal husbandry and genotyping, other labmates from the Tenen lab for their support, Joel Lawitts for assistance with sequence comparisons, Linda Clayton for thoughtful discussions, and Mary Singleton and Alison Lugay for expert assistance with preparation of the manuscript.

References