Transplantation of genetically corrected autologous hematopoietic stem cells is an attractive approach for the cure of sickle-cell disease and β-thalassemia. Here, we infected human cord blood cells with a self-inactivating lentiviral vector encoding an anti-sickling β^{A-T87Q}-globin transgene and analyzed the transduced progeny produced over a 6-month period after transplantation of the infected cells directly into sublethally irradiated NOD/LtSz-scid/scid mice. Approximately half of the human erythroid and myeloid progenitors regenerated in the mice containing the transgene, and erythroid cells derived in vitro from these in vivo–regenerated cells produced high levels of β^{A-T87Q}-globin protein. Linker-mediated PCR analysis identified multiple transgene-positive clones in all mice analyzed with 2.1 ± 0.1 integrated proviral copies per cell. Genomic sequencing of vector-containing fragments showed that 86% of the proviral inserts had occurred within genes, including several genes implicated in human leukemia. These findings indicate effective transduction of very primitive human cord blood cells with a candidate therapeutic lentiviral vector resulting in the long-term and robust, erythroid-specific production of therapeutically relevant levels of β-globin protein. However, the frequency of proviral integration within genes that regulate hematopoiesis points to a need for additional safety modifications.
High-level \(\beta\)-globin expression and preferred intragenic integration after lentiviral transduction of human cord blood stem cells

Suzan Imren,1 Mary E. Fabry,2 Karen A. Westerman,3 Robert Pawliuk,3 Patrick Tang,1 Patricia M. Rosten,1 Ronald L. Nagel,2,4 Philippe Leboulch,3,5 Connie J. Eaves,1,6 and R. Keith Humphries1,7

1Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, British Columbia, Canada. 2Department of Medicine (Division of Hematology) and Physiology and Biophysics, Albert Einstein College of Medicine, Bronx, New York, USA. 3Genetics Division, Department of Medicine and Harvard Medical School, Brigham and Women’s Hospital, Harvard New Research Building, Boston, Massachusetts, USA. 4Department of Physiology and Biophysics, Albert Einstein College of Medicine, Bronx, New York, USA. 5Department of Medicine and Harvard Medical School, Brigham and Women’s Hospital, Boston, Massachusetts, USA. 6Departments of Medical Genetics and 7Medicine, University of British Columbia, Vancouver, British Columbia, Canada.

Transplantation of genetically corrected autologous hematopoietic stem cells is an attractive approach for the cure of sickle-cell disease and \(\beta\)-thalassemia. Here, we infected human cord blood cells with a self-inactivating lentiviral vector encoding an anti-sickling \(\beta^{ATM\\_G297Q}\)-globin transgene and analyzed the transduced progeny produced over a 6-month period after transplantation of the infected cells directly into sublethally irradiated NOD/LtSz-scid/scid mice. Approximately half of the human erythroid and myeloid progenitors regenerated in the mice containing the transgene, and erythroid cells derived in vitro from these in vivo–regenerated cells produced high levels of \(\beta^{ATM\\_G297Q}\)-globin protein. Linker-mediated PCR analysis identified multiple transgene-positive clones in all mice analyzed with 2.1 ± 0.1 integrated proviral copies per cell. Genomic sequencing of vector-containing fragments showed that 86% of the proviral inserts had occurred within genes, including several genes implicated in human leukemia. These findings indicate effective transduction of very primitive human cord blood cells with a candidate therapeutic lentiviral vector resulting in the long-term and robust, erythroid-specific production of therapeutically relevant levels of \(\beta\)-globin protein. However, the frequency of proviral integration within genes that regulate hematopoiesis points to a need for additional safety modifications.

Introduction
Sickle-cell disease (SCD) and \(\beta\)-thalessemia are life-threatening inherited disorders of red cell malfunction of major worldwide importance. Both are caused by alterations to the \(\beta\)-globin gene, which is required for normal hemoglobin production in adults. Currently, allogeneic bone marrow transplantation is the only therapy that can produce permanent cures in patients with these disorders (1, 2). However, this approach is limited by lack of suitably matched donors and the significant morbidity and mortality associated with the treatment.

Transplantation of genetically corrected autologous hematopoietic stem cells (HSCs) could circumvent both of these drawbacks, and the development of this approach has therefore been a goal of basic and preclinical studies for many years. These investigations have led to a number of important advances including the identification of the locus control region (LCR) of the \(\beta\)-globin gene cluster (3–6) and \(\beta\)-globin expression and preferred silencing (33). Lentiviral vectors may also be less prone to silencing (34, 35). Thus, a number of groups have been evaluating the potential of lentiviral vectors as potentially superior vehicles for globin gene transfer applications (36–41). We have focused on the development of a lentiviral vector containing an anti-sickling human \(\beta^{ATM\\_G297Q}\)-globin expression cassette that includes a 2.7-kb region of the LCR (comprising the DNase I-hypersensitive 2, 3, and 4 elements) (37, 38). Evaluations of the therapeutic utility of this vector in murine models of SCD (37) and \(\beta\)-thalassemia (38) have, in both cases, demonstrated pancellular expression of the LCR in promoting high-level, erythroid-specific, and position-independent expression of the \(\beta\)-globin gene (7–10).

Efforts to incorporate these LCR elements into oncoretroviral vectors have yielded some promising results in both mouse (11–14) and human targets (15, 16). However, the generation of high titers of intact oncoretroviral vectors encoding sufficient portions of the LCR to achieve maximal \(\beta\)-globin transgene expression in primary cells has proven difficult (11, 15–20). In addition, obtaining clinically relevant yields of transduced HSCs remains a challenge with these vectors. This is due to the decline in HSC activity that occurs during the several days of growth factor stimulation in culture required for their efficient transduction (21–23).

Lentiviral vectors can also infect primitive primary human hematopoietic cells at high efficiencies (50% and more) but with much shorter in vitro–transduction protocols (24–32). Also, the rev-responsive element in the vector facilitates the export of unspliced transcripts into the cytoplasm, thereby promoting higher protein expression (33). Lentiviral vectors may also be less prone to silencing (34, 35). Thus, a number of groups have been evaluating the potential of lentiviral vectors as potentially superior vehicles for globin gene transfer applications (36–41). We have focused on the development of a lentiviral vector containing an anti-sickling human \(\beta^{ATM\\_G297Q}\)-globin expression cassette that includes a 2.7-kb region of the LCR (comprising the DNase I-hypersensitive 2, 3, and 4 elements) (37, 38). Evaluations of the therapeutic utility of this vector in murine models of SCD (37) and \(\beta\)-thalassemia (38) have, in both cases, demonstrated pancellular expression of the LCR in promoting high-level, erythroid-specific, and position-independent expression of the \(\beta\)-globin gene (7–10).

Nonstandard abbreviations used: BFU-E, burst forming unit–erythroid; Epo, erythropoietin; FL, Flt-3 ligand; HSC, hematopoietic stem cell; LCR, locus control region; LM-PCR, linker-mediated PCR; LTR, long terminal repeat; NOD/SCID, NOD/LtSz-scid/scid; NOD/SCID, NOD/Scid mice; NOD/LtSz-scid/scid mice; SCD, sickle-cell disease; SF, Steel factor; TAE, Tris-acetate EDTA; Tpo, thrombopoietin.

Conflict of interest: C.J. Eaves and R.K. Humphries are paid consultants of StemCell Technologies Inc., suppliers of some of the culture reagents used in this study; R. Pawliuk, K.A. Westerman, and P. Leboulch are officers of Genentech Pharmaceutical Inc.

Citation for this article: J Clin Invest. 114:953–962 (2004). doi:10.1172/JCI200421838.
the transduced β-globin gene with nearly complete correction of the disease phenotype in recipients of transduced HSCs containing an average of 3 proviral integrations per cell.

The present study was undertaken to test the ability of this vector to transduce HSCs of human origin and to measure the level of β-globin transgene expression achieved in their erythroid progeny. The results show efficient transduction of HSCs in normal cord blood as detected by their ability to repopulate sublethally irradiated immunodeficient (NOD/LtSz-scid/scid, or NOD/SCID) mice for at least 6 months. In addition, production of clinically relevant levels of β-globin protein in the erythroid progeny of the transduced HSCs has been demonstrated. Additional evidence of the preferred intragenic integration of such vectors has also been obtained including insertions into genes implicated in human leukemia.

**Results**

Efficient lentiviral transfer of a β\textsubscript{A-T87Q}-globin transgene into primitive human cord blood cells. As an initial test of the ability of our safety-modified β\textsubscript{A-T87Q}-globin lentivirus (Figure 1) to transduce primitive human hematopoietic cells, the proportion of vector-containing hematopoietic colonies was determined on cells plated immediately after a 16-hour exposure of a CD34\textsuperscript{+} cell–enriched cord blood cell suspension to virus following a prior 48-hour period of stimulation in vitro with the following growth factors alone: Flt-3 ligand (FL), Steel factor (SF), thrombopoietin (Tpo), and hyper–IL-6 (hIL-6). PCR analysis of individual colonies (both granulopoietic and erythroid) using primer pairs specific for the β\textsubscript{A-T87Q}-globin transgene and for GAPDH showed that the DNA from 15 of 40 GAPDH\textsuperscript{+} colonies analyzed (38%) also contained the β\textsubscript{A-T87Q}-globin transgene. To determine the efficiency of transducing more primitive cells with multiple-lineage in vivo–repopulating ability, 26 NOD/SCID and 3 NOD/LtSz-scid/scid (NOD/SCID-\textit{nu/nu}) mice were transplanted with infected cord blood cells. In these experiments, CD34\textsuperscript{+} cell–enriched suspensions (\(n = 4\); including experiments 1 and 2) or CD34\textsuperscript{+}CD38\textsuperscript{−} cells isolated by FACS (\(n = 1\)) were first stimulated with growth factors for 18 hours, exposed to β\textsubscript{A-T87Q}-globin lentivirus for the next 6 hours, and then injected into the mice. FACS analysis of serial femoral marrow aspirates obtained from these mice confirmed that the transplants had regenerated persistent populations of hematopoietic (CD45\textsuperscript{+}/71\textsuperscript{−}) cells. Figure 2A shows representative data from two experiments. The variation in the number of human cells regenerated in these two experiments

![Figure 1](image1.png)

**Figure 1**

Schematic diagram of the self-inactivating β\textsubscript{A-T87Q}-globin lentiviral vector used in this study. Features include the HIV LTR; the packaging signal (ψ\textsuperscript{+}); the central purine tract (cPPT); the rev-responsive element (RRE); the β-globin promoter, from SnaBl to the cap site (β\textsubscript{p}); the 3′ β-globin enhancer (to the downstream AvrII site); the 372-bp IVS2 deletion (indicated by the triangle); the β\textsubscript{A-T87Q} mutation (ACA Thr \([\text{Thr}^\text{87}]/\text{Gln}\text{+}\) to CAG Gln \([\text{Thr}^\text{87}]/\text{Gln}\text{+}\)); and the DNase I–hypersensitive sites, SmaI to XbaI (HS2), SacI to PvuII (HS3), and Stul to SpeI (HS4) of the β-globin LCR; 2 stop codons in the ψ\textsuperscript{+} packaging signal; the 400-bp deletion in U3 of the right-hand HIV LTR; and the rabbit β-globin polyA signal. The locations of the β\textsubscript{A-T87Q}-globin transgene-specific primers in exons 2 and 3 (E2, E3) are indicated by arrowheads.
is probably due to differences in the number of CD34+ cells transplanted per mouse \((1.65 \times 10^5 \text{ versus } 4.0 \times 10^5)\) and also to differences in frequency of repopulating cells in the two pools of cord blood used. As expected (42), by 16–20 weeks after the transplant, the regenerated human cells were predominantly B-lymphoid \((\text{CD19/20}^+)\), with a minority of mature granulopoietic \((\text{CD15}^+)\) elements, in almost all \((15/19)\) of the mice. Overall, from five experiments, 19 of 29 recipients were monitored for a minimum of 18 and up to 24 weeks, and 75% of these maintained significant levels of human lymphoid and myeloid cells \((35\% \pm 9\% \text{ total human cells; } 23\% \pm 6\% \text{ human B-lymphoid cells, and } 7\% \pm 3\% \text{ mature human granulopoietic cells}).

Figure 2B shows the results of the PCR analysis of the regenerated human CFCs isolated from 15 NOD/SCID and 3 NOD/SCID-nu/nu mice at different time points after the transplant \((5–40 \text{ colonies per mouse per time point indicated})\). In all but 1 of the 15
High levels of $\beta^{A\text{-T87Q}}$-globin protein are produced in the erythroid progeny of transduced human cord blood cells. To investigate the levels of $\beta^{A\text{-T87Q}}$-globin protein produced by erythroid cells derived from transduced human cord blood cells, we transferred aliquots of cord blood cells immediately after virus exposure to culture conditions designed to promote erythroid differentiation and differentiation over a period of 14 days (see Methods). In the two such experiments performed, more than 85% of the cells at the time of harvest had phenotypic features of maturing glycosphosphatidylinositol (GPI)-expressing erythroblasts (Figure 3, A and B). HPLC analysis of lysates of these cells showed that $\beta^{A\text{-T87Q}}$-globin protein composed 48% and 59% of all $\beta$-globin chains (and 24% and 41% of all $\beta$-like globin chains) in the two experiments (Figure 3C). The ability of transduced HSCs to generate erythroid progeny that would also produce this level of $\beta^{A\text{-T87Q}}$-globin protein was then investigated. Direct assessment of in vivo–generated human erythroblasts is, unfortunately, not possible in NOD/SCID mice repopulated with human hematopoietic cells because of the failure of these hosts to support terminal human erythropoiesis beyond 3–4 weeks after irradiation (43). Therefore, we used the FACS to isolate the human (CD45/71$^-$) cells harvested from the suspension cultures shown $\beta^{A\text{-T87Q}}$-globin protein to represent 35% of all $\beta$-globin chains and 19% of all $\beta$-like globin chains, with values as high as 59% and 31%, respectively (Table 1). Figure 3E shows a representative HPLC profile resulting from a culture initiated with cells present in a mouse receiving a transplant of transduced cord blood cells 11 weeks before. Similar levels of $\beta^{A\text{-T87Q}}$-globin protein (37%) were also detected in lysates of single or 5–10 pooled erythroid colonies produced from regener- ated burst forming unit–erythroids (BFU-Es; Table 1).

RT-PCR analysis of RNA from erythroblasts generated in vitro from transduced cord blood cells confirmed the presence of $\beta^{A\text{-T87Q}}$-globin transcripts in these cells. A strong signal could be detected when the cDNA prepared from these extracts was diluted up to 32 times (Figure 3F, top right panel). In contrast, similar analysis of RNA extracts prepared from human $\beta$-lymphoid cells isolated by FACS from 2 mice reinfused with transduced cells showed no evidence of $\beta^{A\text{-T87Q}}$-globin transcripts in the undiluted cDNA (Figure 3F, top left panel). This demonstrates the erythroid specificity of expression of the $\beta^{A\text{-T87Q}}$-globin transgene achieved in human cells transduced with this vector, as previously observed in murine cells (37).

Polyclonal reconstitution of NOD/SCID mice transplanted with $\beta^{A\text{-T87Q}}$-globin lentiviral vector–infected cord blood cells. To investigate the number and stability of transduced clones obtained in NOD/SCID mice transplanted with $\beta^{A\text{-T87Q}}$-globin lentiviral vector–infected cord blood cells, and to obtain an estimate of the number of copies present in each repopulating cell, we used “bubble” linker-mediated PCR (LM-PCR). A total of 121 human CFC–derived colonies from 11 mice that had been conﬁrmed to contain the $\beta^{A\text{-T87Q}}$-globin transgene were analyzed individually. These included CFCs from 2 mice that were assessed at different time points after receiving transplants. Agarose gel electrophoresis of the LM-PCR products showed that the cells in each colony contained, on average ($\pm$ SEM), 2.1 $\pm$ 0.1 copies of provirus. Multiple unique integration patterns were frequently detected in the different CFCs harvested from a single mouse, indicating the presence of multiple clones transduced by the $\beta^{A\text{-T87Q}}$-globin lentivirus contributing to the regener- ated human cell population present (Figure 4, A–C). Analysis of colonies generated from sequential marrow aspirates indicated dynamic changes in clonal prevalence during the first 6 months after the transplant in both mice followed in this way. Representative data from one of these mice is shown in Figure 4, D and E.

**Preferential intragenic integration of $\beta^{A\text{-T87Q}}$-globin lentiviral vector into human HSCs.** The products of the LM-PCR include genomic sequences flanking the integrated provirus, thus allowing the site of proviral integration to be ascertained by DNA sequencing. Using this approach, we analyzed the sequences of 48 gel-puriﬁed fragments that contained the expected vector-related long terminal repeat (LTR) sequence linked to a ﬂanking genomic sequence. Of these 48 fragments, 35 were unique and were obtained from 33 individual colonies generated in vitro from human CFCs produced in the mice. Of 48 fragments, 13 (27%) were repeats; that is, the same sequence was observed in at least 2 colonies generated from the same mouse. For 7 of the 35 inserts (20%), no match was found in the human genome database. Of the remaining 28 inserts, 24 mapped within identiﬁed genes (88%) and 4 were repetitive elements that could not be mapped. The chromosomal distribution and names of the mapped genes are shown in Figure 5 and in Table 2, respectively. Notably, several integrations occurred in known oncogenes (e.g., MLL and NUP214, both implicated in acute myeloid leukemias), genes involved in signal transduction pathways (e.g., SCAP2), or genes with tumor suppressor activity (e.g., CGBP). Interestingly, the majority of intragenic integration events occurred in introns (21/24 or 88%).

**Table 1**

Expression of $\beta^{A\text{-T87Q}}$-globin protein in human erythroid cells produced ex vivo from the in vivo progeny of transduced cord blood cells transplanted into NOD/SCID mice

<table>
<thead>
<tr>
<th>Wk after transplant</th>
<th>% $\beta^{A\text{-T87Q}}$</th>
<th>% $\beta^{A\text{-T87Q}}$ + % $\beta^A$</th>
<th>$\beta$-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ery susp cult</td>
<td>BFU-E</td>
<td>Ery susp cult</td>
<td>BFU-E</td>
</tr>
<tr>
<td>8</td>
<td>34</td>
<td>55</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>54</td>
<td>33</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>47</td>
<td>31</td>
<td>13</td>
</tr>
<tr>
<td>11</td>
<td>46</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>ND</td>
<td>25</td>
</tr>
<tr>
<td>16</td>
<td>33</td>
<td>41</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>29</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>24</td>
<td>10</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>24</td>
<td>59</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>Mean</td>
<td>35 $\pm$ 5</td>
<td>37 $\pm$ 9</td>
<td>19 $\pm$ 4</td>
</tr>
</tbody>
</table>

A total of 1–10 pooled BFU-E–derived colonies per sample was used. Ery susp cult, erythroid suspension culture; ND, not done.
Discussion

Historically, the development of gene therapy approaches to SCD and β-thalassemia has been plagued by problems inherent in the design of stable vectors encoding human β-globin cassettes sufficient to direct the permanently very high levels of expression needed for the correction of these diseases. More recently, the risks of insertional mutagenesis have highlighted the need to achieve these goals with minimal copies of proviral integrants per cell (44). Here, we present the first evidence that a recently described self-inactivating lentiviral vector can be used to efficiently engineer the sustained production of anticipated therapeutic levels of an anti-sickling βA-T87Q-globin protein in the erythroid progeny of primary human hematopoietic cells with in vivo–repopulating ability. We focused on normal cord blood as a target population for these initial studies because of the high frequency of NOD/SCID mouse–repopulating cells in cord blood (45, 46) and their reported ease of transduction after brief in vitro exposure to virus (25, 28, 29, 32, 47–49). The demonstrated levels of anti-sickling βA-T87Q-globin protein in these cells constitute a significant advance and now underscore the need to extend these studies to adult HSCs.

Many groups have now shown that human HSCs with NOD/SCID mouse–repopulating activity can be efficiently transduced after very brief exposure to lentiviral vectors in vitro without affecting their subsequent hematopoietic potential or ability to express the acquired transgene (25, 31, 32, 48, 50–52). However, most of these findings have been limited to vectors carrying a ubiquitously expressed reporter or drug resistance gene. It remained to be demonstrated that such models could be extended effectively to vectors encoding a therapeutic gene and any additional sequences required for regulating their level and cell-specific expression. The findings presented here are the first to document this achievement with a vector carrying a β-globin gene. At least 20% of the CFCs regenerated in mice that had been repopulated with multiple clones of human cord blood cells exposed to concentrated preparations of this virus (0.5 × 10⁹ to 1 × 10⁹ infectious units/ml) had the vector transgene, and values of more than 50% were noted in several mice. These findings indicate that the βA-T87Q-globin lentivirus can be used to transduce human cord blood cells with long-term NOD/SCID mouse–repopulating activity with the same efficiency as has been reported for simpler constructs (25, 27, 28, 31, 32, 48, 53).

Because we had found that prior overnight exposure to growth factors can increase the proportion of NOD/SCID mouse–repopulating cells that are ultimately transduced (data not shown), we adopted this approach. No perturbation of the normal pattern...
of differentiation was observed in mice repopulated with the transduced cells. Also, multiple-lineage hematopoesis was stable for as long as 4–6 months after transplant in the NOD/SCID hosts and for at least 13 months after transplant in the one NOD/SCID-\nu/\nu mouse that could be followed. Taken together, these findings indicate that the transduction protocol used had no significant or consistent biological effect on cord blood HSCs.

We did not attempt to document any changes in NOD/SCID mouse–repopulating cell activity that may have been incurred by the use of an overnight infection protocol, nor did we try to determine how these might have affected an ultimate net gain or loss in the absolute yield of transduced HSCs (32, 54). Clearly such data would be important to design a clinical protocol. The latter would also require re-examination of the timing and growth factor treatment needed to optimize the transduction of NOD/SCID mouse–repopulating cells in adult sources of human HSCs, since these are present at much lower frequencies than in cord blood (45, 55). Adult sources of HSCs are also known to be qualitatively as well as quantitatively different from cord blood HSCs in terms of the rate at which they can be activated into G1 and the rate at which HSC function may be lost in vitro (16, 42, 56–59).

We took advantage of the high gene transfer to human HSCs obtained in many experiments and the resultant high degree of repopulation of some mice with transduced human cells to investigate the clonal dynamics of this process. To identify different clones by the detection of unique proviral integration events, we analyzed hematopoietic colonies produced in vitro from in vivo–regenerated progenitors using a bubble LM-PCR method. Bubble LM-PCR was chosen because it is simpler to implement than other methods for detecting proviral inserts in small numbers of cells (fewer than $10^3$). However, we do not yet know the sensitivity of this protocol, particularly where multiple integrations may be encountered and hence the number of proviral copies per cell may be underestimated. Nevertheless, this analysis allowed multiple clones (CFCs with different inserts) to be identified in individual animals, providing strong evidence of their polyclonal reconstitution by transduced human HSCs. CFCs with different patterns of integration were also seen in a single repopulated mouse examined over a 2-month interval (i.e., 3 and 5 months after transplant). This is similar to what has been observed in large animals (60).

### Table 2

<table>
<thead>
<tr>
<th>Accession no.</th>
<th>Chromosomal band</th>
<th>Ref Seq gene name</th>
<th>Biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_144621</td>
<td>1p35.1</td>
<td>BOZF1</td>
<td>Unknown</td>
</tr>
<tr>
<td>NM_006367</td>
<td>1p34.2</td>
<td>CAP1</td>
<td>Signal transduction</td>
</tr>
<tr>
<td>NT_022517.90</td>
<td>3p26.1</td>
<td>Predicted gene</td>
<td>Signal transduction</td>
</tr>
<tr>
<td>NM-014744</td>
<td>3p24.3</td>
<td>TBC1D5</td>
<td>Cell growth and maintenance</td>
</tr>
<tr>
<td>BC017432</td>
<td>4q21.3</td>
<td>MLLT2</td>
<td>Unknown</td>
</tr>
<tr>
<td>AF285118</td>
<td>6p25.1</td>
<td>CGI-203 like gene</td>
<td>Signal transduction and metabolism</td>
</tr>
<tr>
<td>NM_009390</td>
<td>7p15.2</td>
<td>SCAP2</td>
<td>Unknown</td>
</tr>
<tr>
<td>NM_018846</td>
<td>7q21.11</td>
<td>SBB126</td>
<td>Development and cell adhesion and metabolism</td>
</tr>
<tr>
<td>NM_001163</td>
<td>9q21.11</td>
<td>APBA1</td>
<td>Unknown</td>
</tr>
<tr>
<td>NM_005085</td>
<td>9q34.13</td>
<td>NUP214</td>
<td>Cell growth and maintenance</td>
</tr>
<tr>
<td>NM_006795</td>
<td>11q13.1</td>
<td>EHD1</td>
<td>Unknown</td>
</tr>
<tr>
<td>NM_005933/3</td>
<td>11q23.3</td>
<td>MLL and DDX6*</td>
<td>Transcription and cell growth and maintenance</td>
</tr>
<tr>
<td>NM_004397</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NG_001019</td>
<td>14q32.33</td>
<td>IGHG3</td>
<td>Immune response</td>
</tr>
<tr>
<td>NM_174916</td>
<td>15q15.2</td>
<td>UBR1</td>
<td>Metabolism</td>
</tr>
<tr>
<td>NM_182619</td>
<td>16q22.1</td>
<td>LOC348174</td>
<td>Unknown</td>
</tr>
<tr>
<td>NM_00102909</td>
<td>17q21.31</td>
<td>KIAA0553</td>
<td>Unknown</td>
</tr>
<tr>
<td>NM_001433</td>
<td>17q23.3</td>
<td>ERN1</td>
<td>Metabolism</td>
</tr>
<tr>
<td>NT_024871.61</td>
<td>17q25.3</td>
<td>Predicted gene</td>
<td>–</td>
</tr>
<tr>
<td>NM_014593</td>
<td>18q21.1</td>
<td>CGBP</td>
<td>Transcription</td>
</tr>
<tr>
<td>NM_052847</td>
<td>19p13.3</td>
<td>GNG7</td>
<td>Signal transduction</td>
</tr>
<tr>
<td>NM_080665</td>
<td>19p13.2*</td>
<td>MGC19004</td>
<td>Unknown</td>
</tr>
<tr>
<td>NM_153358</td>
<td>19p13.2*</td>
<td>FLJ90396</td>
<td>Transcription</td>
</tr>
<tr>
<td>NM_003370</td>
<td>19q13.2–q13.3</td>
<td>VASP</td>
<td>Cell motility</td>
</tr>
<tr>
<td>NM_003868</td>
<td>Xp11.4</td>
<td>CASK</td>
<td>Metabolism</td>
</tr>
</tbody>
</table>

\*The integration is in the second intron of the DDX6 gene, which is located in the fifth intron of the MLL gene. \#Genomic locus hit twice in two independent experiments.
The present studies are also important in their demonstration of the high lineage-specific levels of β^A-Thal^2-globin produced in the erythroblast progeny of transduced HSCs. In many examples, these levels were above the estimated values of 20% of total β-globin per erythroid cell required to achieve a therapeutic effect in SCD and β-thalassemia (41, 61), even when all the β-like globin chains present in the cells were taken into consideration. Although the transgene copy number was estimated as 2 copies per cell, no meaningful calculation of the amount of β^A-Thal^2-globin protein produced per transgene copy was possible because the extracts analyzed were obtained from a highly heterogeneous mixture of erythroblasts at different stages of differentiation. Nevertheless, it is clear that the results obtained here represent a significant improvement over the levels of β-globin transgene-derived transcripts and protein recently documented in erythroblasts derived from human cord blood cells transduced with an oncoretroviral vector containing a similar β-globin cassette with a smaller LCR fragment (16).

We found that the majority (24/28, 86%) of the proviral integrations occurred within genes identified in ReSeq, and these included a number of genes that possess cell signaling or regulatory functions or that have been associated with leukemia. This proportion is significantly higher than would be expected (P < 0.0001) from the current estimates that genes occupy approximately 35% of the genome (62, 63). Thus our findings support the concept of preferred intragenic integration of lentiviral vectors into human DNA, as noted by others who reported 67% and 58% intragenic HIV-1 integrations in human lymphoid (64) and HeLa cells (65), respectively. We also observed preferential integration into introns (21/24 = 88% of inserts characterized), as reported by others (64), perhaps reflecting the greater relative size of introns as compared to exons (21/24 = 88% of inserts characterized), as reported by others (64), (21/24 = 88% of inserts characterized), as reported by others (64), as reported by others (64), as reported by others (64). Taken together, these findings refocus attention on the possibility that such integration could result in activation or ablation of the expression of endogenous genes following the transduction of human HSCs, with functional consequences. As such, they point to the need for improvements in vector design aimed at both insulating integration sites from the effects of the provirus and minimizing the integrated proviral copy number required for therapeutic levels of expression.

**Methods**

**Lentiviral vector design and production.** The β^A-Thal^2-globin lentiviral vector used (Figure 1) was based on a vector previously described in detail (37) that was then modified to include the following additional safety features: a 400-bp deletion in U3 plus replacement of U5 in the right LTR with the polyadenylation/termination sequence of the rabbit β-globin gene and mutation of two potential ATG initiation codons in the 3′ portion of the packaging signal (details available upon request). High-titer recombinant virus pseudotyped with vesicular stomatitis virus glycoprotein-G was initially produced by using a standard 4-plasmid packaging system and later by using a recently developed “super-split” 7-plasmid packaging system (66). Recombinant virus made using the safety-modified vector and super-split 7-plasmid packaging system was used in the last two experiments (one of which is experiment 2 presented in Figure 2A), whereas in the first three experiments (including experiment 1 in Figure 2A), the original vector (37) was used. Harvested virus-containing supernatants were concentrated by two rounds of ultracentrifugation approximately 1,000- to 2,000-fold to achieve titers of 0.5 × 10^8 to 1 × 10^9 infectious units/ml. Viral titers were determined using quantitative Southern blot analysis to compare vector DNA levels in cells containing a known number of integrated proviral copies with those in NIH 3T3 cells infected with test virus preparations.

Absence of replication-competent virus in the viral supernatants was verified by a mobilization assay analogous to an assay previously employed for oncoretroviral vectors (67).

**Isolation and transduction of human cord blood cells.** Cord blood was obtained from consenting mothers undergoing cesarean delivery of healthy, full-term infants, and low-density (less than 1.077 g/ml) cells were isolated by centrifugation on Ficoll-Hypaque (Pfizer). CD34^+^ cells—enriched populations (65–98% CD34^+^ cells) were obtained by one of the following methods: (a) removal of lineage marker-positive cells using a column (n = 2), (b) sorting by FACSVantage SE (n = 1; BD); (c) positive (EasySep) selection using magnetic beads (n = 1; StemCell Technologies Inc.). In one experiment, CD34^+^CD38^−^ cells were isolated by FACS as previously described (46). Cells were stimulated overnight for in vivo experiments and for 48 hours for in vitro experiments at densities less than or equal to 2 × 10^6 cells/ml in Iscove’s medium supplemented with 1% BSA, 10 μg/ml bovine pancreatic insulin, and 200 μg/ml human transferrin (BdT; StemCell Technologies Inc.), 10 μmol 2-mercaptoethanol, 2 mM glutamine, 100 ng/ml FL (Immunex Corp.), 100 ng/ml SF, 50 ng/ml Tpo (Genentech Inc.), and 100 ng/ml hIL-6 (provided by S. Rose-John, Christian-Albrechts University, Kiel, Germany). The following day, the cells were pelleted, resuspended in fresh growth factor—supplemented medium with 5 μg/ml protamine sulfate and 0.5 × 10^8 to 5 × 10^9 infectious units/ml (MOI = 9–140, 140 in experiment 1, 9 and 90 in experiment 2), placed in a 24-well plate coated with 2 μg/cm^2 Retronectin (Takara Shuzo Co.) or with 5 μg/cm^2 fibronectin (Sigma-Aldrich), and then incubated at 37°C for 6 hours. In experiment 2, cells were plated in a Retronectin-coated well, which was then preloaded with virus for 2 hours at 4°C.

**In vitro CFC assays.** Appropriate aliquots of human cells were plated in 1 ml of methylcellulose-containing medium (Methoclut H4230; Stem-Cell Technologies Inc.) supplemented with 50 ng/ml human SF (purified from human SF cDNA-transfected Cos cell supernatants) and 20 ng/ml each of human IL-3 (Novartis), IL-6 (Cangene), GM-CSF (Novartis), granulocyte-CSF (G-CSF; Novartis), and 3 U/ml erythropoietin (Epo; StemCell Technologies Inc.) (68). The cultures were then incubated for 2 weeks at 37°C, and colonies of terminally differentiating erythroid, myeloid, and mixed erythroid-myeloid cells (from BFU-E, GM-CFU, and granulocyte-erythroid-megakaryocyte-macrophage–CFU [GEMM-CFU], respectively) were identified.

**Animals.** NOD/SCID mice and NOD/SCID-nu/nu mice were bred and maintained in microisolators at our center. Original breeding pairs of NOD/SCID-nu/nu mice were provided by J. Nolta (University of Washington, St. Louis, Missouri, USA), who obtained these mice by serially backcrossing the nude gene onto the NOD/SCID background for seven generations to give a host that supports human hematopoiesis in a fashion that is similar to the NOD/SCID mouse but without the development of endogenous thymomas and the resultant reduced lifespan of the NOD/SCID mouse (69).

At 8–10 weeks of age, mice were irradiated with 350 cGy of 137Cs γ-rays the day prior to being intravenously injected with virus-infected human cells. (These cells were the progeny of an initial aliquot containing 1.7 × 10^9 to 4 × 10^9 CD34^+^ or 7 × 10^9 CD34^+^CD38^−^ cord blood cells per mouse.) Thereafter, mice were given acidified water containing ciprofloxacin (100 mg/l; Bayer AG). Marrow cells were aspirated (70) 3, 7, 11, and 16 weeks after transplantation; the mice were killed 4–8 weeks later, and all cells from both tibiae and femurs were harvested for analysis. For phenotyping studies, cells were stained with human-specific monoclonal antibodies for total hematopoietic cells (CD45 and CD71), progenitors (CD34), B-lymphoid cells (CD19 and CD20), and granulopoietic cells (CD15), as previously described (71). Human multiple-lineage engraftment was defined as the detection of five or more CD19/20/CD34^+^ and five or more CD15^+^ events per 2 × 10^6 live (propidium iodide–negative) events analyzed. In some instances, human CD45/71^−^ cells
were isolated using a 3-laser FACS Vantage SE (BD), as previously described (46), and then plated in methylcellulose cultures to generate hematopoietic colonies or placed in erythroid differentiation cultures.

PCR analysis. Cell lysates were prepared by incubating the cells from individually pulsed colonies at 42°C for 30 minutes in the presence of proteinase K (1 mg/ml). PCR was performed on 2 μl (1/25th) of these columns using primers specific for human GAPDH (5′-ACCGTCAAGG CTGAGAACGG-3′ and 3′-ACGCTACGAGCCGACATC-5′) and for the β°-TETO, globin transgene (5′-GGGCGCTTGGCCG-3′ and 3′-TGTCGATTTCCA-5′) to amplify the expected 100-bp and 600-bp fragments, respectively. After 40 cycles of denaturation (30 seconds at 94°C), annealing (30 seconds at 63.5°C), and extension (60 seconds at 72°C), PCR products were separated on a 1.5% Tris-acetate EDTA (TAE) agarose gel.

RT-PCR analysis. Total RNA was extracted using a commercial kit (Trizol; Gibco BRL, Invitrogen Corp.) and reverse transcribed by random priming using 1 μg of total RNA and SuperScript II Reverse Transcriptase (Invitrogen Corp.) at 42°C for 30 minutes, followed by denaturation at 72°C for 10 minutes and snap cooling to 4°C for 5 minutes. A PCR was performed on undiluted and diluted (1/2–1/4–1/8–1/16–1/32–1/64) cDNA using the same primer sets as those used for PCR to amplify the expected 100-bp fragments for the β°-TETO-globin and GAPDH transcripts. After 40 cycles of denaturation (30 seconds at 94°C), annealing (30 seconds at 63.5°C), and extension (60 seconds at 72°C), the PCR products were separated on a 1.5% TAE agarose gel.

Erythroid differentiation cultures and analysis of globin protein. Cells were cultured for the first 9 days in α-medium plus 15% FCS plus 1 U/ml Epo, 100 ng/ml SF, 40 ng/ml IGF-I (R & D Systems), 10−4 mol freshly dissolved hydrocortisone (Sigma-Aldrich), 10−6 mol 17β-estradiol (Sigma-Aldrich), 1.28 μg/ml iron-saturated transferrin (StemCell Technologies Inc.), and 10−4 mol 2-mercaptoethanol. The cells were then transferred to fresh α-medium plus 15% FCS, 1 U/ml Epo, 1 μg/ml insulin (Sigma-Aldrich), 1.28 μg/ml iron-saturated transferrin, and 10−4 mol 2-mercaptoethanol for another 3–5 days; finally, cells were harvested and lysed using acid-acetone. Globins were analyzed by HPLC using a Vydac large-pore (300 Å) C18 column and a modified acetonitrile/H2O/trifluoroacetic acid gradient as previously described (72). The amount of β°-TETO-globin protein was then calculated, both as a percentage of the total β-globin present (β°-TETO-globin/β°-globin × 100%) and as a percentage of all β-like proteins present (β°-TETO-globin/β°-TETO-globin + β°-globin + γ°-globin + ρ°-globin × 100%).

LM-PCR. Integrated LTR and flanking genomic sequences were amplified and then isolated using a modification of the bubble LM-PCR strategy (73). Aliquots of the cell lysates from transgene-positive colonies (5–10 μl; one-fifth to one-tenth) were digested with HinfI (New England Biolabs), and the fragments were then ligated overnight at room temperature to a double-stranded bubble linker (5′-CTCTCCCTTTCTGAA CTGTAACGGTTCTG TACGAAATCGTCTGCTTCTCCTTG-3′ and 5′-ANTCAAGGAAGGAC GCCTGTCTGCAAGGAGGACAGGACGGAGGAG-3′) prior to performing a first PCR (PCR-A) on 10 μl (one-tenth) of the ligation product using a linker-specific Vectorrete primer (5′-CGAATCGTTCA- GGTTCTGATAGAATCGCTCAT-3′) and an LTR-specific primer (LTR-A: 5′-CAACACACACATTAGGAACGTCACAGGA-3′) and under the following conditions: one cycle of 94°C for 2 minutes, 20 cycles of 94°C for 30 seconds and 65°C for 1 minute, and one cycle of 72°C for 2 minutes. The bubble linker contains a 30-nucleotide nonhomologous sequence in the middle region that prevents binding of the linker primer in the absence of minus strand generated by the LTR-specific primer. A 1-μl aliquot of the PCR-A reaction (one-fifteenth) was then used as a template for a second nested PCR (PCR-B) using an internal LTR-specific primer (LTR-B: 5′-GAGA GCTCCCCAG CTGCAGACG TGGCTAA-3′) and the same linker-specific Vectorrete primer as was used in PCR-A with the following conditions: one cycle of 94°C for 2 minutes, 30 cycles of 94°C for 60 seconds and 72°C for 1 minute, and one cycle of 72°C for 2 minutes. Ten microliters (one-half) of the final PCR-B product were electrophoresed using 2% agarose TAE gel. Individual bands were excised and purified using the Qaqex II Gel Extraction Kit (QIAGEN) for sequencing. Two types of marked fragments were obtained. One was always 750 bp, representing the vector sequence bound by the primer specific to the 3′ LTR and the next HinfI restriction site. The other was of variable length (158 bp or more) depending on the distance to the next HinfI site in the host genome adjacent to the 5′ LTR.

Sequence analysis. Gel-purified DNA bands were sequenced directly from the LTR-B primer with the ABI PRISM Model 377 according to the manufacturer’s instructions. Only sequences containing both complete 5′ LTR and intact HinfI-linked Vectorrete ends were analyzed. Sequencing of fragments with identical lengths sometimes yielded different integration sites. The sensitivity of determining the number of proviral integrations per colony could be increased by further cloning of the gel-isolated fragments. BLAST searches were performed using the University of California Santa Cruz (UCSC) genome project website (http://genome.ucsc.edu) to identify the genomic location of the flankin g sequences. Chromosomal localizations of the mapped genes were determined using the Ensembl map viewer (http://www.ensembl.org).

Acknowledgments This work was supported by grants from the NIH (Program Project grant HL 55435 to C.J. Eaves, M.E. Fabry, R.K. Humphries, P. Leboulic, and R.L. Nagel; HL38055 and HL68962 to R.L. Nagel), the Stem Cell Network (to C.J. Eaves and R.K. Humphries), and the International Cooley’s Anemia Foundation (to R.L. Nagel). P. Tang received a Research Studentship from the Stem Cell Network. The authors thank Lars Palmquist for integration site analysis; Sandy Suzuka, members of the Stem Cell Assay Service of the Terry Fox Laboratory, Sanja Sekulovic, Margaret Hale, and Michael Scherzter for technical help; StemCell Technologies Inc., Immunex Corporation, Genentech, P. Lansdorp, and S. Rose-John for gifts of reagents; and Colleen MacKinnon for assistance in manuscript preparation.

Received for publication April 9, 2004, and accepted in revised form July 20, 2004.

Address correspondence to: R. Keith Humphries, Terry Fox Laboratory, 601 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada. Phone: (604) 877-6070, ext. 3095; Fax: (604) 877-0712; E-mail: khumphri@bccrc.ca.

Connie J. Eaves and R. Keith Humphries contributed equally to this work.


