doi:10.1016/j.ymthe.2004.10.017

METHOD

Real-Time PCR: An Effective Tool for Measuring Transduction Efficiency in Human Hematopoietic Progenitor Cells

Anthony D. Villella,[1,*] Jing Yao,[2] Robert R. Getty,[2] Beth E. Juliar,[3] Constantin Yiannoutsos,[3] Jennifer R. Hartwell,[4] Shanbao Cai,[4] Mohammed A. Sadat,[4] Kenneth Cornetta,[2] David A. Williams,[1,*] and Karen E. Pollok[4,†]

[1]*Division of Experimental Hematology, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA*
[2]*Medical and Molecular Genetics,* [3]*Biostatistics and Data Management Core, Indiana University Cancer Center, and* [4]*Herman B. Wells Center for Pediatric Research, Indiana University, Indianapolis, IN 46202, USA*

**To whom correspondence and reprint requests should be addressed at the Division of Experimental Hematology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA. E-mail: David.Williams@cchmc.org.*

†To whom correspondence and reprint requests should be addressed at the Herman B Wells Center for Pediatric Research, Indianapolis, IN 46202, USA. E-mail: kpollok@iupui.edu.

Available online 2 December 2004

Accurate measurement of gene transfer into hematopoietic progenitor cells is an essential prerequisite for assessing the utility of gene therapy approaches designed to correct hematologic defects. We developed a reliable method to measure transduction efficiency at the level of the progenitor cell with real-time polymerase chain reaction (PCR) analysis of individual progenitor-derived colonies. We hypothesized that this method would demonstrate better sensitivity and specificity than are currently achievable with conventional PCR. An oncoretroviral vector containing the enhanced green fluorescent protein was used to transduce human CD34$^+$ cells derived from bone marrow or granulocyte-colony-stimulating factor-mobilized peripheral blood. Progenitor assays were set up and colonies plucked after visualization by fluorescence microscopy. By analyzing microscopically identified fluorescent samples and nontransduced samples, we calculated an overall sensitivity and specificity of 90.2 and 95.0%, respectively. Real-time PCR had higher specificity and sensitivity than conventional PCR as analyzed by generalized linear models ($P = 0.002$ and $P = 0.019$, respectively). In conclusion, we found real-time PCR to have superior sensitivity and specificity compared to conventional PCR in determining transduction efficiency of hematopoietic progenitor cells.

Key Words: gene therapy, real-time PCR, retrovirus, transduction efficiency, progenitor cell

INTRODUCTION

Determination of gene-transfer efficiency is a critically important element of preclinical and clinical gene therapy trials. In many preclinical settings, a marking gene such as enhanced green fluorescent protein (EGFP) is incorporated along with the "therapeutic gene" into the vector. Transduction efficiency can then be determined easily by flow cytometric analysis. However, expression of a foreign protein such as EGFP is not utilized in clinical trials due to the immunogenic nature of this foreign protein [1]. Because hematopoietic stem and progenitor cells are a common target in human gene therapy protocols, analysis of transduction efficiency into committed progenitor cells, in addition to differentiated cells, can be especially informative to investigators. In the short-term follow-up of patients, transduction of committed progenitor cells does not necessarily reflect whether a hematopoietic stem cell was transduced. However, the presence of transduced progenitors in the long-term follow-up of patients is consistent with transduction of a hematopoietic stem cell. The most common methodology for estimating gene

1525-0016/$30.00

doi:10.1016/j.ymthe.2004.10.017

transfer into hematopoietic progenitor cells in clinical protocols is conventional DNA PCR analysis of individual progenitor-derived colonies. Performing PCR on individual colonies grown in methylcellulose, however, is labor intensive, time consuming, and subject to false positives. Stray cells or DNA, not derived from a colony, can inadvertently be captured while harvesting individual colonies, which can result in an unacceptable rate of false positives due to the sensitivity of PCR. In addition, the presence of methylcellulose can disrupt the polymerase chain reaction, requiring additional labor-intensive handling such as additional washing and DNA extraction prior to running the PCR.

Real-time PCR, because of its high sensitivity, efficiency, and capacity for accurate quantification, has recently been employed to measure gene-transfer efficiency. Real-time PCR utilizes a *Taq* polymerase and a set of primers that cleave a fluorogenic probe bound to the transgene internal to the primers. The resulting fluorescence is measured as the probe is cleaved and is proportional to the amount of amplified DNA, thus allowing for accurate quantification of copy number and, therefore, transduction efficiency. Samples with a greater amount of amplified DNA emit a fluorescent signal that reaches the threshold level at an earlier PCR cycle number (referred to as cycle threshold (C_t)). The computer software calculates the C_t value by generating a graph of cycle number versus mean fluorescence. This allows for a quantitative readout rather than visualization of bands produced by gel electrophoresis. Based on statistical analysis of the differences in C_t values, it is possible to discriminate transduced colonies from false positives caused by stray DNA. In addition, this technology can detect multiple target DNA sequences during the PCR in a 96-well format and thus is capable of rapidly screening many progenitor colonies and controls.

The performance of real-time PCR has been shown to be comparable to that of conventional PCR in the detection of a multitude of different infectious pathogens from human patients [2–5]. When performed with appropriate internal controls, real-time PCR is also capable of accurately measuring vector titer and estimating gene-transfer efficiency in a variety of cell lines [6–8]. To estimate retroviral transduction efficiency into hematopoietic progenitor cells, Gerard *et al.* applied real-time PCR methods to progenitor-derived colonies grown in a methylcellulose culture system [9]. They demonstrated that single colonies could be picked from methylcellulose, placed into a cell lysis solution, added to a prefilled plate of PCR reagents, and run on a real-time PCR machine to determine gene-transfer efficiency. To date, however, no systematic comparison of real-time PCR with conventional DNA PCR for the measurement of transduction efficiency in progenitor colonies has been performed.

We developed a novel method for measuring gene-transfer efficiency into hematopoietic progenitor cells by applying a relatively simple statistical analysis to real-time PCR data generated from individual progenitor-derived colonies. We hypothesized that by coupling statistical analysis with real-time PCR data, we could correct for false positives, resulting in a method with greater specificity compared to conventional PCR. We then compared the sensitivity and specificity of this method to those of conventional PCR by analyzing individual colonies derived from human $CD34^+$ cells transduced with a retroviral vector containing the EGFP sequence. These studies demonstrate that real-time PCR is a valid assay for detection of gene transfer into human hematopoietic progenitor colonies with superior levels of sensitivity and specificity compared to conventional PCR.

RESULTS

Determination of the Minimal Number of Provirus-positive Cells Detected by Real-Time PCR

To determine if the amplification of proviral sequences by real-time PCR was in the detection range required for analysis of progenitor colonies, we performed a serial dilution experiment using a human erythroid leukemic (HEL) cell line, HEL-EGFP-1, that contains ~1 copy of the MFG-EGFP provirus per genome (see Materials and Methods). With the assumption that a typical progenitor colony will contain at least 50 cells, we diluted HEL-EGFP-1 cells serially into nontransduced HEL cells and transferred them to a 96-well plate, resulting in cell concentrations ranging from 10 to 10,000 cells/well in triplicate. We then took one-fifth of each DNA sample for amplification and sequence detection by real-time PCR. These volumes were identical to those used for analysis of

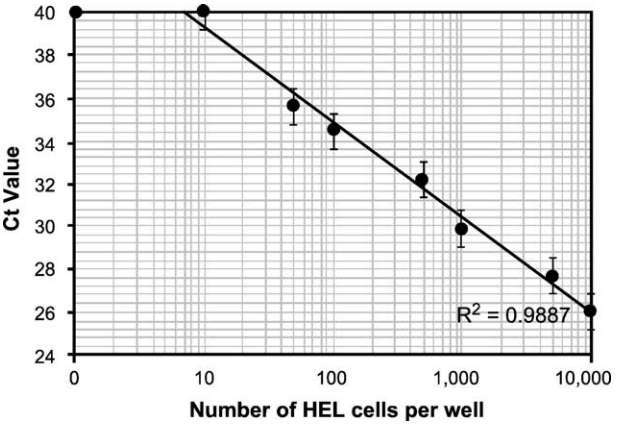

FIG. 1. Determination of the minimal number of HEL-EGFP-1 cells detectable by real-time PCR. The HEL-EGFP-1 cells contain ~1 copy of the MFG-EGFP provirus per genome. Samples were prepared by serial dilution of HEL-EGFP-1 cells into nontransduced HEL cells and were analyzed by real-time PCR. The data points represent the mean C_t values with standard deviation. C_t value is plotted against the number of single-copy HEL cells on a logarithmic scale. Linear regression analysis is shown as a solid line. This is representative of three independent experiments.

DNA from progenitor colonies. The smallest number of HEL-EGFP-1 cells detected by real-time PCR was 50 cells/well (10 cells/reaction) (Fig. 1), which is in the range needed for detection of transduced progenitor colonies. Nontransduced cells and culture medium-only samples were uniformly negative. In addition, to determine if the addition of methylcellulose negatively impacted the real-time polymerase chain reaction, we performed experiments in which 3 μl of methylcellulose was added to the real-time PCR containing HEL-EGFP-1 DNA. Three microliters of methylcellulose is typically the volume of methylcellulose harvested when plucking colonies (see Materials and Methods). In two separate experiments, the addition of the methylcellulose to the samples did not change the C_t values compared to samples not spiked with methylcellulose (data not shown).

Distribution of C_t Values From Real-Time PCR Analysis of Fluorescent (F), Nonfluorescent (NF), and Nontransduced (NT) Progenitor Colonies

Prestimulated human $CD34^+$ cells derived from bone marrow (BM) or granulocyte-colony-stimulating factor (G-CSF)-mobilized peripheral blood (MPB) were transduced on Retronectin-coated plates with the MFG-EGFP retrovirus as described under Materials and Methods. NT cells were mock transduced on Retronectin-coated plates and not exposed to virus. Since the frequency of committed progenitor cells can vary from donor to donor, aliquots of cells were always plated in complete methylcellulose at 250, 500, and 1000 cells/ml so that plucking of progenitor colonies was performed from plates with ~30 colonies or less. We found this decreased the chances of contamination from stray DNA or from adjacent colonies since the colonies were well separated and could be easily plucked. We performed fluorescence microscopy and real-time PCR on progenitor colonies and compared them in 11 independent experiments (Table 1 and Fig. 2). In the plates containing transduced progenitor colonies, the number of F versus NF colonies was first determined by fluorescence microscopy. As a control, analysis of the NT progenitor colonies by fluorescence microscopy indicated that all colonies were negative for EGFP-mediated fluorescence. We plucked the colonies and analyzed them by real-time PCR as described under Materials and Methods using probe/primer sets specific to the packaging signal of the MFG-EGFP provirus or to the housekeeping gene apolipoprotein B (ApoB). We used the ApoB C_t as a means of sorting out samples that did or did not have amplifiable DNA in the sample. In all experiments (Table 1), the mean C_t value for F colonies was lower than for NF and NT, indicating the presence of the MFG-EGFP provirus in the F colonies. Fig. 2 displays the distribution of the F and NT C_t values from all individual experiments (n = 11). In all but one experiment (BM009), we observed a separation of provirus-positive samples (clustered at lower C_t values) from provirus-negative samples (clustered near C_t of 40) within each donor (Fig. 2). Since BM009 lacked the typical expected C_t cluster of values near 40, it was not incorporated into the determination of the C_t cutoff. The choice of the C_t cutoff for provirus presence or absence could be optimally and consistently obtained by using the Moloney murine leukemia virus (MMLV) C_t values derived from NT colonies, which typically cluster at C_t = 40. For instance, ~87% of the NT colonies (127 of 146 NT colonies) analyzed by the MMLV probe/primer set were at C_t = 40. We set the C_t cutoff for detection of the MMLV sequence at 35.9; a simple formula for the determination of the C_t cutoff for provirus presence versus absence is outlined under Materials and Methods. In addition, we compared the ratio of MMLV to ApoB as a measurement of provirus absence or presence to using the MMLV C_t alone. However, the ratio did not allow for

TABLE 1: Descriptive statistics for MMLV C_t

Donor	Samples	*N*	Mean	Median	SD
BM001	F	20	32.71	32.02	3.85
	NF	20	39.97	40	0.14
	NT	20	40	40	0
BM005	F	25	31	30.87	4.37
	NF	12	38.06	40	3.51
	NT	21	40	40	0
BM006	F	27	31.23	30.3	4.83
	NF	12	35.52	35.26	4.42
	NT	9	38.26	40	2.82
BM007	F	26	30.19	29.91	2.88
	NF	30	37.88	39.31	2.98
	NT	5	38.7	40	2.91
BM008	F	24	33.23	32.9	3.47
	NF	23	37.49	40	3.53
	NT	5	37.18	40	3.88
BM009	F	24	28.08	28.1	2.15
	NT	24	33.59	33.45	0.51
BM012	F	20	29.98	30.04	1.61
	NF	20	38.55	40	3.14
	NT	20	39.88	40	0.42
MPB02_1	F	12	29.03	29	3.5
	NF	11	39	40	1.65
	NT	6	39.01	40	1.7
MPB02_2	F	20	28.92	29.25	1.94
	NF	22	36.01	37.09	4.17
	NT	14	39.85	40	0.43
MPB04_1	F	20	29.81	30.69	3.12
	NF	20	35.14	34.14	3.41
	NT	20	39.61	40	1.05
MPB04_2	F	20	27.93	28.52	3.11
	NF	19	34.72	34.37	4.13
	NT	20	39.94	40	0.29
Total	F	238	30.29	29.98	3.68
	NF	189	37.23	39.94	3.65
	NT	164	38.72	40	2.49

BM, bone marrow donor sample; F, fluorescent colony; NF, nonfluorescent colony; NT, nontransduced colony; MPB, G-CSF-mobilized peripheral blood sample.

doi:10.1016/j.ymthe.2004.10.017

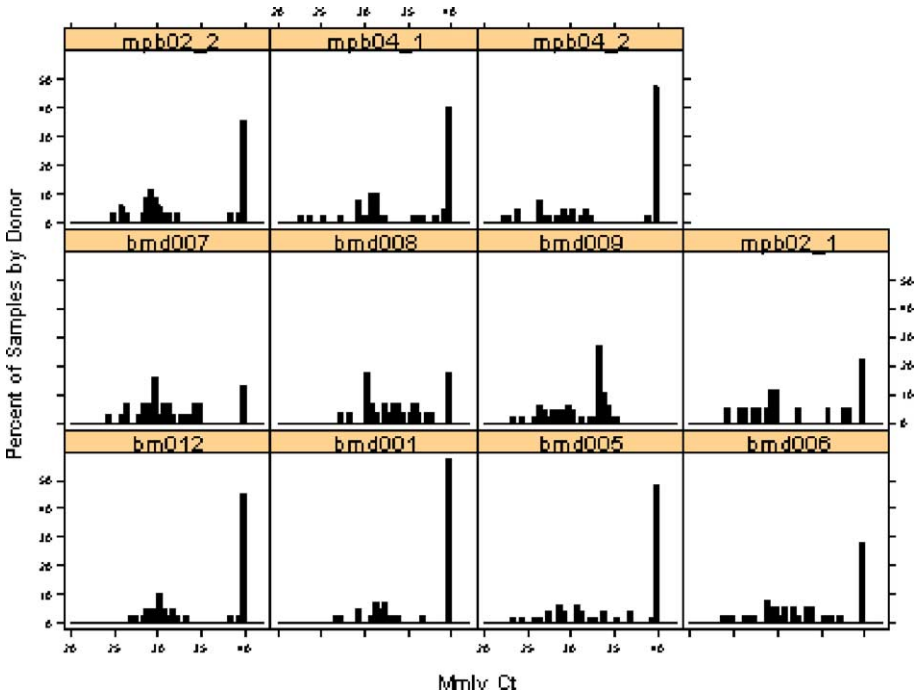

FIG. 2. Distribution of MMLV C_t values for nontransduced (NT) and fluorescent (F) colonies. The bar graphs represent the percentage of F and NT samples present for each C_t value. A C_t value of 40 is maximum and is indicative of the unequivocal absence of MMLV transgene. C_t values less than 40 were determined to be PCR^+ or PCR^- based on statistical analyses of the distribution of the NT C_t values.

a consistent C_t cutoff value since the distinction between the distributions of provirus positive and negative samples was diminished (data not shown).

Correlation of Transduction Efficiency by Flow Cytometry, Fluorescence Microscopy, and Real-time PCR

In 8 of the 11 experiments, we compared transduction efficiencies determined by flow cytometry, fluorescence microscopy, and real-time PCR (Table 2). In these experiments, we first monitored the transduction efficiency of the $CD34^+$-derived cells by flow cytometry and, in parallel, set up progenitor assays. Approximately 10–14 days later, we determined the transduction efficiency of the progenitors by fluorescence microscopy. Then we randomly picked progenitors and determined the transduction efficiency by real-time PCR as described above. The mean transduction efficiency of the bulk population of $CD34^+$-derived cells as determined by flow cytometry was 42%. The mean transduction efficiency of colony-forming units was 39% $EGFP^+$ by fluorescence microscopy and 48% $EGFP^+$ by real-time PCR (Table 2). These differences were not statistically significant (Kruskal–Wallis test). In some of the experiments, transduction efficiency as measured by fluorescence microscopy and flow cytometry may be slightly underestimated because of the presence of transduced colonies that did not express EGFP at detectable levels.

TABLE 2: Comparison of transduction efficiency

Expt	Flow[a]	Microscopy[b]	Real time[c]
MPB001	35%	32%	38%
MPB002	44%	47%	46%
BM001	47%	32%	50%
BM002	56%	39%	50%
BM003	51%	48%	60%
BM004	36%	32%	40%
BM005	40%	43%	50%
BM006	26%	39%	51%
Mean ± SD	42% ± 10%	39 ± 7%	48 ± 7%

[a] Flow cytometry was performed 24–72 h following the second day of transduction.
[b] Fluorescence microscopy was performed on progenitor colonies after 10–14 days of culture.
[c] Real-time PCR was performed on progenitor colonies harvested randomly from culture within 24 h of enumeration via fluorescence microscopy. MPB, G-CSF-mobilized peripheral blood; BM, bone marrow.

Sensitivity and Specificity of Real-Time PCR

We used ten independent experiments (BM009 was excluded) with BM or MPB to calculate the sensitivity and specificity of the real-time PCR method for detection of integrated provirus in progenitor-derived colonies. We performed real-time PCR on individual colonies that were determined by fluorescence microscopy to be F, NF, and NT. As a representative example, individual C_t values derived from colonies and controls generated from BM012 are shown in Fig. 3. We considered NF colonies unreliable and did not use them for determination of specificity because of possible silencing of the EGFP transgene. We chose fluorescence microscopy as a positive control (provirus presence) and used it for the determination of sensitivity; we chose NT samples as a negative control (provirus absence) for determination of specificity.

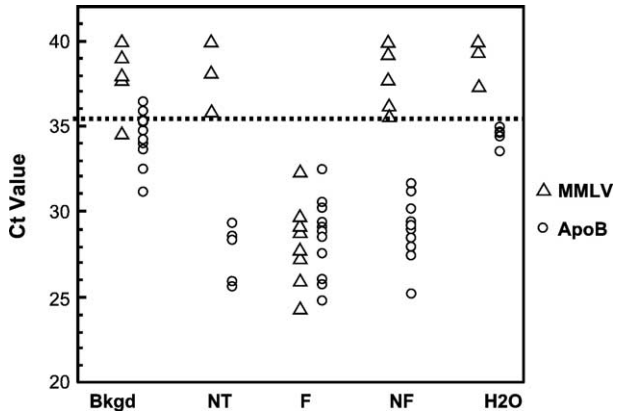

FIG. 3. Scatter plot of C_t values for MMLV and ApoB obtained from real-time PCR analysis of individual progenitor colonies from a representative experiment (BM012). Each colony was analyzed for both MMLV and ApoB sequences. C_t values for MMLV < 35.9 (dotted line) are positive for the transgene, while values ≥35.9 are negative. Bkgd, background methylcellulose samples; NT, nontransduced colonies; F, fluorescent colonies; NF, nonfluorescent colonies; H_2O, buffer control.

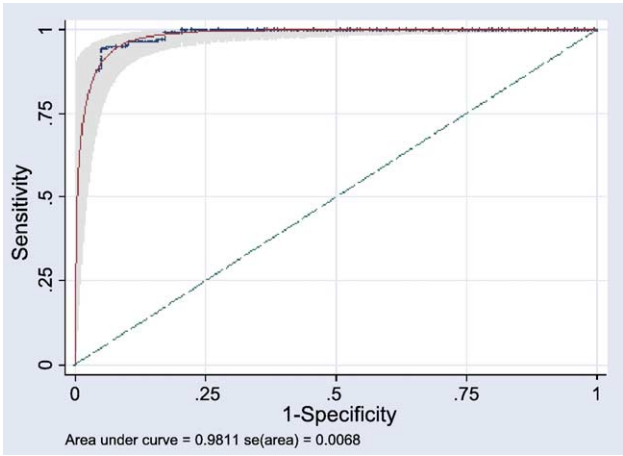

FIG. 4. ROC curve generated from GEE analysis of C_t values. ROC is generated from NT and F data. Each point on the curve corresponds to a different potential C_t cutoff. The area under the curve was 98.1%. A large AUC such as this indicates potential for high sensitivity and specificity.

In Table 3, sensitivity and specificity were calculated using the concordance rates between fluorescence microscopy and real-time PCR with C_t 35.9 as the cutoff point. Percentage specificity was calculated as the number of provirus-negative colonies divided by the total number of nontransduced negative colonies × 100. Percentage sensitivity was calculated as the number of provirus-positive colonies divided by the total number of microscopy-positive colonies × 100. The receiver operator characteristic (ROC) area under the curve was 98.1%; with a C_t cutoff of 35.9; this corresponded to less than 5.5% false positives (Fig. 4). We observed high levels of specificity and sensitivity (95.0 and 90.2%, respectively) indicating that real-time PCR is a valid assay for assessment of transduction efficiency in progenitor colonies (Table 3).

Conventional PCR vs Real-Time PCR

We compared three experiments with 154 samples using conventional PCR and 10 experiments with 354 samples using real-time PCR analysis to assess provirus-positive and provirus-negative colonies (Table 3). We picked separate progenitor colonies from the same plates for analysis by conventional PCR and real-time PCR; we did not split colonies due to low amounts of DNA. In each experiment, we plucked samples of F, NF, NT, and background, prepared them as described under Materials and Methods, and ran them on the ABI Prism 7700 for amplification and detection. We compared these samples to an equivalent number of samples plucked from the same methylcellulose culture plates but transferred to Qiagen plates for DNA isolation and PCR analysis by conventional PCR.

TABLE 3: Real-time versus conventional PCR

Sample	PCR	Number of samples analyzed	Number of samples negative or positive		Specificity[a]	Sensitivity[b]	False positive rate[c]
			Negative	Positive			
NT colonies	Conventional	50	39	11	78.0%	—	—
	Real time	140	133	7	95.0%	—	—
F colonies[d]	Conventional	58	9	49	—	84.5%	—
	Real time	214	21	193	—	90.2%	—
Background samples[e]	Conventional	72	56	16	—	—	22.2%
	Real-time	104	98	6	—	—	5.8%

NT, nontransduced progenitor colony; F, fluorescent progenitor colony.

[a] The percentage specificity was calculated as the number of PCR-negative colonies classified as negative divided by total nontransduced negative colonies × 100 using the C_t 35.9 as the cutoff point.

[b] The percentage sensitivity was calculated as the number of PCR-positive colonies classified as positive divided by total microscopy-positive colonies × 100 using the C_t 35.9 as the cutoff point.

[c] The false positive rate was calculated as the number of PCR-positive background samples divided by the total number of background samples × 100.

[d] Fluorescent colonies were identified by visual examination using the Axiovert 25 inverted fluorescence microscope.

[e] Background samples were methylcellulose culture medium samples collected from the noncolony areas of the plate.

doi:10.1016/j.ymthe.2004.10.017

We compared specificity and sensitivity of the two methods separately using generalized linear models (GEE) analyses. Odds ratios from GEE analyses present the magnitude of the differences. The odds ratio is the ratio of the probability of a correct real-time test divided by the probability of a correct conventional test (see Materials and Methods). In the analysis of data from NT samples, specificity using real-time PCR was significantly higher compared to conventional PCR as analyzed by GEE (P = 0.002). For real-time compared to conventional method, the odds ratio of a provirus-absent sample testing negative was 5.56 (95% confidence interval 1.85 to 16.73). In the analysis of data from F samples, sensitivity using real-time PCR was significantly higher compared to conventional PCR as analyzed by GEE (P = 0.019). For real-time compared to conventional method, the odds ratio of a provirus-present sample testing positive was 1.94 (95% confidence interval 1.11 to 3.37). As shown in Table 3, for real-time PCR, specificity was 95.0% (exact 95% CI 90.0 to 98.0%) and sensitivity was 90.2% (exact 95% CI 85.4 to 93.8%). For conventional PCR, specificity was 78.0% (exact 95% CI 64.0 to 88.5%) and sensitivity was 84.5% (exact 95% CI 72.6 to 92.7%). As hypothesized, the conventional PCR occasionally led to amplification of stray cells or DNA not associated with a colony as evidenced by the presence of bands in 22% of the background samples (Table 3). The PBS controls in those experiments were negative, eliminating the possibility of cross-contamination of reagents. Background samples from real-time PCR experiments clustered around a C_t value of 40 with 5.8% false positives (Table 3).

DISCUSSION

Real-time PCR is an attractive method for measuring gene-transfer efficiency in a large number of samples simultaneously. Despite its increasing use, no published study has compared real-time PCR directly with standard conventional PCR for the detection of transgenes in hematopoietic progenitor cells. One report determined gene transfer into hematopoietic progenitor cells by analyzing individual colonies from a methylcellulose culture system [9]. In this report, the investigator showed that real-time PCR was an effective method for detecting individual transduced progenitor colonies. A high correlation between the real-time PCR fluorescence data and the visualization of bands on an agarose gel produced from the real-time PCRs was observed. The real-time PCR experiments were successfully performed using direct lysates, demonstrating that no DNA purification or extraction was necessary and that the presence of small amounts of methylcellulose did not interfere with the real-time polymerase chain reactions.

Our study has extended this previous report by comparing directly the sensitivity and specificity of conventional PCR with real-time PCR. Our results demonstrate superior sensitivity and specificity of real-time PCR over conventional PCR in measuring gene-transfer efficiency into hematopoietic progenitor cells. Sensitivity may potentially be higher with real-time PCR because of its more objective readout. Real-time PCR data are reported as C_t values, while conventional PCR is relatively more subjective than real-time PCR and dependent on identifying bands on an agarose gel. For smaller colonies, small amounts of DNA may produce faint bands or be below the detection level of conventional PCR yet still be detectable by real-time PCR. Gerard *et al.* showed that several colonies that tested positive by real-time PCR did not produce visible bands when subsequently run on an agarose gel. In our study, we were able to identify as few as 10 transduced cell equivalents per PCR (50 cells/lysate sample) by real-time PCR. Even very small colonies, which are generally defined as a minimum of 50 cells, can easily be identified with real-time PCR.

In our study, real-time PCR was also more specific than conventional PCR in detecting the provirus in progenitor colonies. The procedure of picking colonies from a methylcellulose plate can be technically difficult and the inadvertent capture of stray cells or DNA can result in false positives. To avoid this error, we plated $CD34^+$ cells in methylcellulose at a lower cell density (250–1000 cells/ml), which prevented overcrowding of the plate with colonies, making it technically easier to pluck colonies without contamination. In addition, we routinely plucked progenitor colonies for PCR analysis only if the plates had fewer than 30 colonies per plate. By applying the statistical analysis to the distribution of C_t values for nontransduced colonies to determine the appropriate C_t cutoff value, a more specific discrimination of positive colonies was also possible.

There are several sources for potential error with either method described here. NT colonies may become PCR positive either from cross-contamination within the 96-well plate or from integration of wild-type retroviruses into genomic DNA, which are capable of binding to the universal primers. Cross-contamination can be minimized with careful handling and the use of multichannel pipettes. When utilizing human DNA, it may be advantageous to use transgene-specific primers rather than universal primers to avoid recognition of wild-type viruses. Additionally, very dim or nonexpressing colonies that were successfully transduced with the EGFP retrovirus may be read out as NF, thereby causing an artificially increased false positive rate. Fluorescent colonies may be PCR negative if the transgene is absent but the colony produces some degree of autofluorescence, as can be seen in apoptotic cells. False positives can occur if stray transduced cells or DNA is inadvertently collected in the process of plucking of negative colonies. Finally, distinguishing primer dimers from weakly positive samples can be problematic, especially in real-time PCR, in which a gel is not routinely run.

To avoid high levels of primer dimers routinely observed in negative control samples that do not contain DNA, we utilized nontransduced cells as the negative control.

Because real-time PCR is relatively automated and capable of analyzing many samples simultaneously, it remains an attractive tool for the measurement of transduction efficiency. The technology is especially applicable to human gene therapy clinical trials, in which PCR is performed on a large number of patient samples and repeated regularly. With the described technique of analyzing real-time PCR progenitor colony data, a more accurate assessment of transduction efficiency is possible.

MATERIALS AND METHODS

Isolation of bone marrow and mobilized peripheral blood CD34⁺ cells. BM and G-CSF-mobilized MPB samples were obtained from healthy adult volunteers after obtaining informed consent. for samples collected at the Cincinnati Children's Hospital Medical Center (n = 8), the protocol utilized was approved by the Institutional Review Board of Cincinnati Children's Hospital Medical Center. For samples collected at Indiana University (n = 3), the protocol utilized was approved by the Institutional Review Board at the Indiana University School of Medicine. The bone marrow samples were filtered through a 40-μm cell strainer and then diluted 1:1 with Hanks' balanced solution (GIBCO BRL, Grand Island, NY, USA). Low-density (<1.077 g/cm^3) cells were isolated by centrifugation at 1700 rpm for 30 min (without brake) at room temperature on Ficoll–Hypaque (Sigma, St. Louis, MO, USA) and then washed once in phosphate-buffered saline (PBS). Residual red blood cells were lysed with lysis buffer (0.1 mM EDTA, 10 mM $KHCO_2$, 155 mM NH_4CL in PBS). The cells were then resuspended in magnetic-activated cell sorting (MACS) buffer (5 mM EDTA, 0.5% bovine serum albumin (BSA) in PBS, PH 7.2). CD34⁺ cells were isolated using the CD34 progenitor cell isolation kit and Vario MACS separator (Miltenyi Biotech, inc., auburn, ca, usa) according to the manufacturer's instructions. Following CD34 isolation, the cells were immediately placed into culture as described below. For isolation of MPB CD34⁺ cells, volunteers were treated for 5 days with subcutaneous human G-CSF (Amgen, Thousand Oaks, CA, USA) at 5 μg/kg/day as previously described [10]. On day 5 of treatment, white blood cells were collected by apheresis and CD34⁺ cells were isolated by immunomagnetic methods with the Isolex 300SA cell selection device according to the manufacturer's instructions (Baxter, Irvine, CA, USA). These samples were cryopreserved in aliquots of 3×10^6 cells in 90% fetal calf serum (FCS) + 10% dimethyl sulfoxide (Sigma) and later thawed for use in gene transfer experiments. CD34 purity ranged from 75 to 95%. For determination of purity, isolated cells were stained with an allophycocyanin-conjugated anti-CD34 antibody (clone 581) (BD Biosciences Pharmingen, San Jose, CA, USA) and the percentage CD34⁺ was determined by flow cytometry.

Generation of gibbon ape leukemia virus (GALV)- and RD114-pseudotyped MFG-EGFP. The MFG-EGFP retrovirus vector expresses the EGFP under the control of the MMLV long terminal repeat and has been previously described [11]. For generation of the GALV-pseudotyped vector, supernatant from an amphotropic MFG-EGFP clone was used to infect the PG13 packaging cell line (American Type Culture Collection, Manassas, VA, USA) [12]. Pooled supernatant from PG13 cells was then collected and used for transduction. Vector particles pseudotyped with feline endogenous virus (RD114) were derived from the packaging cell line FLYRD18 [13] (kindly provided by Mary Collins, Chester Beatty Laboratories, London, UK). An amphotropic MFG-EGFP clone was used to infect FLYRD18 cells and the resultant supernatant was used for transduction. The titers of the supernatants were calculated by infecting HeLa cells and determining the percentage EGFP expression 48 h after end-point dilution of the supernatant. Titers of supernatants utilized in experiments described here were 2×10^5/ml (GALV) and 5×10^5/ml (RD114).

Transduction of CD34⁺ cells. Isolated CD34⁺ cells were prestimulated with cytokines at a cell density of 4×10^5/ml in Iscove's modified Dulbecco's medium (IMDM) (GIBCO BRL) containing 10% FCS and 1% penicillin–streptomycin (referred to as complete medium—CM). The cytokines used were G-CSF (Amgen), stem cell factor (SCF) (Amgen), and megakaryocyte growth and development factor (Amgen; all at 100 ng/ml), referred to as G/S/M. Non-tissue-culture 24-well plates were coated with 8 μg/cm^2 fibronectin (CH-296) (Takara Shuzo, Otsu, Japan) for 2 h at room temperature as previously described [14,15]. The wells were blocked with 2% BSA (Roche, Indianapolis, IN, USA) for 20 min and washed once with PBS. Cells (2×10^5) suspended in 1 ml CM + G/S/M were added to 1 ml of viral supernatant in each CH-296-coated well. The multiplicity of infection was 1.0 for transductions performed with GALV and 2.5 for transductions with RD114. The cells were transduced for 4 h at 37°C, 5% CO_2 and then the CM + supernatant was replaced with fresh CM + G/S/M for incubation overnight. A second transduction was performed the following day in an identical fashion. BMD003 and BMD009 were transduced with the RD114-pseudotyped vector. All other experiments utilized the GALV-pseudotyped vector. Twenty-four hours after the second transduction, the cells were harvested, counted, and transferred into progenitor assays. An aliquot of the transduced cells was resuspended in CM + G/S/M for an additional 24–72 h and then analyzed by flow cytometry to determine %EGFP⁺ cells.

Progenitor assays. Progenitor assays were performed 24 h after the second day of transduction in complete methylcellulose medium (Methocult GF, H4434; Stem Cell Technologies, Inc.) containing SCF (50 ng/ml), interleukin-3 (10 ng/ml), granulocyte–macrophage-CSF (10 ng/ml), and erythropoietin (3 units/ml). The cells were seeded onto 35-mm dishes in triplicate at concentrations of 250, 500, and 1000 cells/ml and incubated at 37°C in 5% CO_2. After 10–14 days of incubation, colony-forming units–granulocyte–macrophage, burst-forming units–erythroid, and colony-forming units–mixed lineage were enumerated using the Axiovert 25 inverted fluorescence microscope. Direct visualization of fluorescence from colonies was used to determine transduction efficiency.

Real-time PCR. Fifty microliters of crude PCR cell lysing buffer consisting of 1× PCR buffer, 0.5% NP-40 (Sigma), 0.5% Tween 20 (Sigma), and 0.91 mg/ml proteinase K (Boehringer Mannheim, Mannheim, Germany) was delivered to each well of a 96-well optical plate (Applied Biosystems, Foster City, CA, USA) using a repeating pipettor. Under a fluorescence microscope, single progenitor colonies growing in methylcellulose were harvested using a P-10 pipette set at 3.0 μl and transferred directly into the 96-well plate containing the crude cell lysing buffer. Approximately 20 F colonies and 20 NF colonies (each from transduced cultures) and 20 NT colonies were collected and labeled. "Background" methylcellulose samples obtained from noncolony areas of the plates and samples containing cell lysis buffer alone were analyzed as negative controls. After all colonies were transferred, the plate was covered with an optical adhesive (Applied Biosystems) and placed in a thermal cycler (Applied Biosystems) for 1 h at 60°C followed by 15 min at 95°C. Ten microliters of cell lysate from each well was then transferred to each corresponding well of a new 96-well plate previously loaded with 40 μl per well PCR master mix. The preparation of the PCR master mix and the addition of the cell lysates were performed under a hood in a dedicated PCR room and filter barrier pipet tips were used to prevent aerosol contamination. After all the cell lysates were transferred to the fresh 96-well plate, the plate was covered with a new optical adhesive (Applied Biosystems) and placed in the real-time PCR machine (ABI Prism 7700 sequence detection system, PE Applied Biosystems) for amplification and detection.

The ABI Prism 7700 sequence detection system (PE Applied Biosystems) along with a MMLV universal probe/primer set was utilized to detect

doi:10.1016/j.ymthe.2004.10.017

the presence of the integrated MFG-EGFP provirus in individual colonies. This universal probe/primer set has been previously shown to recognize the packaging sequence in numerous MMLV-derived vectors including the MFG vector [8,11]. The forward primer (MMLV-F, 5′-CGCAACCCTGGGAGACGTCCC-3′) and the reverse primer (MMLV-R, 5′-CGTCTCCTACCAGAACCACATATCC-3′) amplified a 134-bp fragment of the packaging signal. Amplification was detected with a FAM-labeled probe (6FAM–5′-CCGTTTTTGTGGCCCGACCTGAGG-3′–TAMRA) internal to the PCR product. To minimize false negatives, the human ApoB gene was used as an internal control for DNA content and quality was analyzed concurrently. The forward primer ApoB-F (5′-TGAAGGTGGAGGACATTCCTCTA-3′) and the reverse primer ApoB-R (5′-CTGGAATTGCGATTTCTGGTAA-3′) were used and amplification was detected with a VIC-labeled internal probe (VIC–5′-CGAGAATCACCCTGCCAGACTTCCGT-3′–TAMRA). Samples that did not contain the ApoB sequence were removed from the analysis. Negative controls included NT colonies, background methylcellulose samples, cell lysis buffer, and injectable water.

Statistical analysis for determination of C_t cutoff value. Data were collected in separate experiments from 11 bone marrow donors. BM009 lacked the typical expected C_t cluster of values near 40; therefore, it was not included in the analysis to determine the C_t cutoff value. EGFP-positive and NT progenitor colonies were analyzed. Since the EGFP transgene was known to be absent in nontransduced samples, these samples served as the background control for determination of the C_t cutoff. Samples with an ApoB C_t of 40 or greater were interpreted as having an inadequate amount of DNA and removed from the analysis. To standardize the C_t cutoff for all experiments, the threshold (level of fluorescence for a given C_t value) was set at the computer default setting. In two experiments, this was adjusted slightly higher to avoid assigning an inappropriately low C_t value to a sample that was clearly negative based on the computer-generated graph. A C_t value of 40 is the default setting on the machine for a negative sample since after 40 cycles of amplification, if DNA still has not been detected, it is assumed absent.

Due to some overlap between the C_t measurements of transgene-present and -absent samples, the cutoff C_t value would unavoidably produce some number of false positives and negatives. Determination of the C_t cutoff was therefore based on the priority of minimizing false positives while maintaining sensitivity greater than 70%. A statistical analysis of the C_t values from NT control samples was used to develop a threshold C_t value-based test to determine which progenitor-derived colonies exposed to retrovirus were PCR positive. An individual colony was determined to be PCR positive if its C_t value for the EGFP transgene was less than 35.9. The threshold of 35.9 was selected based on distributions of the NT negative control C_t values and ROC curve (Fig. 4). The ROC curve is a plot of sensitivity vs false positive rate with area under the curve (AUC) as a measure of test accuracy. Points on the curve correspond to probabilities of gene presence and to all possible diagnostic cut-points for values of sensitivity and false positive rate. Approximately normal distributions at lower C_t values were consistently observed in the gene-present (i.e., EGFP-positive) samples for each donor. However, means varied considerably between donors. In contrast, gene-negative data were typically at or near C_t of 40. The consistency of the NT data (Fig. 2), together with the study priority of minimizing false positives, resulted in the choice of cutoff based on ranked gene-absent data. Data from one experiment (BM009) were very atypical and were therefore not incorporated into the determination of the C_t cutoff. To generate points for the ROC curve, GEE was used since multiple samples from the same donor are correlated. In this analysis, probability of gene presence was predicted by C_t value adjusted for donor. A C_t threshold of 35.9 was found to correspond to less than 5.5% false positives. This analysis was a robust nonparametric rank-based approach appropriate for expected experimental data distributions. Accordingly, an individual colony was determined to be PCR positive if its C_t value generated using the MMLV probe/primer set was less than 35.9.

Simple formula for determination of C_t cutoff. We developed a method for determining a C_t cutoff value based on a false positive rate of less than 5.5%. For a given probe/primer set, the C_t values from a minimum of 62 nontransduced samples should be analyzed. This is based on the statistical analysis of the proportion of C_t values of 40 for nontransduced samples. Data from an experiment without a large percentage of C_t near the default value of 40, such as BM009, must be excluded, since they would not meet the distributional requirements of the real-time method. The following formula can be used to estimate the C_t cutoff from nontransduced test samples:

(1) Rank C_t values from minimum to maximum.
(2) Multiply number of samples by 0.055. Round down to integer value. This is the rank of the sample that will be used for the C_t cutoff.
(3) Obtain cutoff value by rounding down to 1 decimal point.
(4) Positives are < new cut-point and negatives are ≥ new cut-point.

For example, given that the number of test nontransduced samples = 67:

(1) Rank by C_t values (34.71, 35.54, 37.32, 38.93, 39.1, 39.7, remainder 40.00).
(2) $67 \times 0.055 = 3.685$. Round down to 3. Select third from lowest C_t value, 37.32.
(3) Round to 1 decimal to obtain cut-off of 37.3.
(4) Positives have $C_t < 37.3$ and negatives have $C_t \geq 37.3$.

Conventional PCR. DNA was isolated using the QIAamp 96 spin blood kit (Qiagen, Valencia, CA, USA) as per the manufacturer's instructions with minor modifications. Colonies were picked from methylcellulose as described above, but transferred directly to the 96-well block prefilled with 200 μl per well of PBS. Approximately 20 F, 20 NF, and 20 NT colonies were picked per donor sample and transferred and labeled accordingly. Twenty-five microliters of Qiagen protease and 200 μl of buffer AL were added to each sample and the wells were sealed tightly with caps. The samples were mixed thoroughly by shaking and then spun at 3000 rpm briefly. The block was incubated at 70°C and respun to collect any condensate from the caps. Two hundred ten microliters of 100% ethanol was added and then mixed by shaking after reapplying the caps. The QIAamp 96 plate was placed on top of a new block. The mixture from the original block was then transferred to the QIAamp 96 plate and then spun at 6000 rpm for 4 min. Five hundred microliters of buffer AW1 was then added and the samples were respun at 6000 rpm for 2 min. Five hundred microliters of buffer AW2 was added and the samples were respun at 6000 rpm for 2 min. The plate was then placed on top of a microtube collection rack and incubated at 70°C for 10 min. One hundred microliters of Tris buffer (10 mM, pH 8.0) (Sigma) was then added to each well. The samples were incubated at room temperature for 5 min and then spun at 6000 rpm for 4 min. DNA was stored at 4°C. Two PCRs were performed for each sample, one using MMLV primers and a second using the β-globin forward primer (5′-CAATCCAGCTACCATTCTGC-3′) and the β-globin reverse primer (5′-GAATCCAGATGCTCAAGGCC-3′). Thirty microliters of DNA was used for MMLV PCR and 10 μl of DNA for the β-globin PCR. Each sample was amplified with both sets of primers and run on 2% agarose gels separately.

Determination of the minimal number of cells required for real-time PCR detection. The HEL cell line containing ~1 copy of the MFG-EGFP provirus (HEL-EGFP-1) per genome was generated by transducing HEL cells with a limiting dilution of supernatant containing the GALV-pseudotyped MFG-EGFP retroviral vector to yield a cell population that was ~5% $EGFP^+$. The $EGFP^+$ cells were then sorted by flow cytometry and expanded. Genomic Southern analysis with a MFG-EGFP vector copy number standard indicated that the copy number was ~1 copy per genome equivalent (data not shown). Cells from the HEL-EGFP-1 cell line, which were transduced with a single copy of the EGFP gene using the MFG-GALV vector, were enumerated with the automated particle counter (Coulter Z1;

Beckman/Coulter, Hialeah, FL, USA) and prepared by limiting dilution. The cells were suspended in IMDM in concentrations ranging from 1×10^6 to 1×10^3 cells/ml to allow analysis of the following samples of cell numbers: 10,000, 5000, 1000, 500, 100, 50, and 10. Ten microliters from each concentration was added to 40 μl of crude cell lysing buffer in a 96-well plate (Applied Biosystems). The plate was covered and placed in a thermal cycler (Applied Biosystems) and heated at 60°C for 1 h and 95°C for 15 min. Ten microliters of the cell lysate from each well was then analyzed using the ABI Prism 7700 exactly as performed with colony cell lysates as described above. Nontransduced HEL cells and IMDM were used as negative controls.

Statistical analyses. Statistical analyses were performed using SAS version 8.02 and ROC graphs were created using STATA version 8.2. Statistical tests were performed as two-tailed at the 0.05 significance level. Sensitivity and specificity are presented with exact 95% confidence intervals for real-time and conventional PCR. Real-time and conventional methods were tested for differences in specificity and sensitivity using GEE. GEE analysis was necessary to account for correlation between repeated samples from the same donor [16]. Typically, samples obtained from a single donor will be more similar than samples obtained from different donors. If positive correlations among each donor's samples were ignored, the standard error of the parameter estimate would be too small and type I error would be inflated. Separate comparisons of specificity and sensitivity between the two methods were performed using population-averaged generalized linear models. The correlation structure was assumed to be exchangeable. ROC curve was generated using GEE analysis of the fluorescent and nontransduced sample data sets.

ACKNOWLEDGMENTS

Real-time PCR was performed in the Indiana University Vector Production Facility, which is a NIH-designated National Gene Vector Laboratory (U42 RR11148), and this work was supported, in part, by a Core Centers of Excellence in Molecular Hematology (CCEMH) grant (PHS P50 DK49218) and a core laboratory supporting grant, PHS P01 HL53586. The majority of mobilized peripheral blood and bone marrow harvests and CD34+ cell isolations were performed by the Cell Manipulation Lab and Translational Trials Development and Support Laboratory of the Cincinnati Center of Excellence in Hematology (CCEH) at Cincinnati Children's Hospital Medical Center. We also wish to thank the Translational Research Trials Office of the CCEH and Hoxworth Blood Center at the University of Cincinnati for providing human CD34+ cells. We also acknowledge the staffs at the Riley Hospital Apheresis Facility and the stem cell core laboratory at the Indiana University Medical Center for their expert technical assistance in apheresis and processing of CD34+ cells. This work was also funded in part by an Indiana University Cancer Center pilot project grant (K.E.P.) and a core pilot grant funded through the CCEMH (K.E.P.). We also acknowledge the Riley Children's Foundation for their support of core facilities.

RECEIVED FOR PUBLICATION MAY 13, 2004; ACCEPTED OCTOBER 28, 2004.

REFERENCES

1. Stipecke, R., *et al.* (1999). Immune response to green fluorescent protein: implications for gene therapy. *Gene Ther.* **6:** 1305–1312.
2. Apfalter, P., *et al.* (2003). Comparison of a new quantitative ompA-based real-time PCR TaqMan assay for detection of *Chlamydia pneumoniae* DNA in respiratory specimens with four conventional PCR assays. *J. Clin. Microbiol.* **41:** 592–600.
3. Rantakokko-Jalava, K., and Jalava, J. (2001). Development of conventional and real-time PCR assays for detection of *Legionella* DNA in respiratory specimens. *J. Clin. Microbiol.* **39:** 2904–2910.
4. Ke, D., *et al.* (2000). Development of conventional and real-time PCR assays for the rapid detection of group B streptococci. *Clin. Chem.* **46:** 324–331.
5. Mygind, T., *et al.* (2002). Determination of PCR efficiency in Chelex-100 purified clinical samples and comparison of real-time quantitative PCR and conventional PCR for detection of Chlamydia pneumoniae. *BMC Microbiol.* **2:** 17.
6. Becker, K., *et al.* (1999). Real-time PCR quantitative polymerase chain reaction to assess gene transfer. *Hum. Gene Ther.* **10:** 2559–2566.
7. Klein, D., *et al.* (2000). Accurate estimation of transduction efficiency necessitates a multiplex real-time PCR. *Gene Ther.* **7:** 458–463.
8. Sanburn, N., and Cornetta, K. (1999). Rapid titer determination using quantitative real-time PCR. *Gene Ther.* **6:** 1340–1345.
9. Gerard, C. J., *et al.* (1996). A rapid and quantitative assay to estimate gene transfer into retrovirally transduced hematopoietic stem/progenitor cells using a 96-well format PCR and fluorescent detection system universal for MMLV-based proviruses. *Hum. Gene Ther.* **7:** 343–354.
10. Croop, J. M., *et al.* (2000). Large-scale mobilization and isolation of CD34+ cells from normal donors. *Bone Marrow Transplant.* **26:** 1271–1279.
11. Bierhuizen, M. F. A., *et al.* (1997). Enhanced green fluorescent protein as selectable marker of retroviral-mediated gene transfer in immature hematopoietic bone marrow cells. *Blood* **90:** 3304–3315.
12. Miller, A., *et al.* (1991). Construction and properties of retrovirus packaging cell based on gibbon ape leukemia virus. *J. Virol.* **65:** 2220–2224.
13. Cossett, F. L., *et al.* (1995). High titer packaging cells producing recombinant retroviruses resistant to human serum. *J. Virol.* **69:** 7430–7436.
14. Hanenberg, H., *et al.* (1996). Colocalization of retrovirus and target cells on specific fibronectin fragments increases genetic transduction of mammalian cells. *Nat. Med.* **2:** 876–882.
15. Pollok, K., *et al.* (2001). Differential transduction efficiency of SCID-repopulating cells derived from umbilical cord blood and granulocyte colony-stimulating factor-mobilized peripheral blood. *Hum. Gene Ther.* **12:** 2095–2108.
16. Zeger, S. L., and Liang, K. Y. (1992). An overview of methods for the analysis of longitudinal data. *Stat. Med.* **11:** 1825–1839.