Molecular Therapy
Methods & Clinical Development
Original Article

# DNA Barcoding in Nonhuman Primates Reveals Important Limitations in Retrovirus Integration Site Analysis

Jennifer E. Adair,[1,2,5] Mark R. Enstrom,[1,5] Kevin G. Haworth,[1] Lauren E. Schefter,[1] Reza Shahbazi,[1] Daniel R. Humphrys,[1] Shaina Porter,[3] Kenric Tam,[3] Matthew H. Porteus,[3,4] and Hans-Peter Kiem[1,2]

[1]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; [2]School of Medicine, University of Washington, Seattle, WA 98195, USA; [3]Department of Pediatrics, Stanford University, Stanford, CA 94305, USA; [4]Institute of Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, CA 94305, USA

***In vivo*** **tracking of retrovirus-tagged blood stem and progenitor cells is used to study hematopoiesis. Two techniques are used most frequently: sequencing the locus of retrovirus insertion, termed integration site analysis, or retrovirus DNA barcode sequencing. Of these, integration site analysis is currently the only available technique for monitoring clonal pools in patients treated with retrovirus-modified blood cells. A key question is how these two techniques compare in their ability to detect and quantify clonal contributions. In this study, we assessed both methods simultaneously in a clinically relevant nonhuman primate model of autologous, myeloablative transplantation. Our data demonstrate that both methods track abundant clones; however, DNA barcode sequencing is at least 5-fold more efficient than integration site analysis. Using computational simulation to identify the sources of low efficiency, we identify sampling depth as the major factor. We show that the sampling required for integration site analysis to achieve minimal coverage of the true clonal pool is likely prohibitive, especially in cases of low gene-modified cell engraftment. We also show that early subsampling of different blood cell lineages adds value to clone tracking information in terms of safety and hematopoietic biology. Our analysis demonstrates DNA barcode sequencing as a useful guide to maximize integration site analysis interpretation in gene therapy patients.**

## INTRODUCTION

Understanding the biology of hematopoiesis after transplant is critical to improving the efficacy and safety of hematopoietic stem cell (HSC)-based therapies such as gene therapy and gene editing. Infusion of retrovirally transduced $CD34^+$ cells into autologous patients is the current strategy applied in gene therapy clinical trials. Insertional mutagenesis in patients treated with gamma-retrovirus-transduced $CD34^+$ hematopoietic cells prompted recommendations for clonal analysis of gene-modified cells in patients for safety monitoring (Guidance for Industry: Gene Therapy Clinical Trials—Observing Subjects for Delayed Adverse Events).[1,2] Thus, clone tracking following gene therapy has contributed largely to our understanding of hematopoiesis after autologous transplant.

The primary method used for clone tracking in patients is retrovirus integration site analysis (ISA).[3] Various techniques are utilized by different laboratories to sequence the genomic locus of provirus insertion as a heritable, clone-specific signature.[4] Generally, ISA requires fragmentation of target cell genomic DNA (gDNA) and ligation of known oligonucleotide sequences to the resulting gDNA fragments for primer seeding. Multiple rounds of PCR amplification are performed and the products are sequenced. The method of gDNA fragmentation and/or template used can introduce bias into ISA (reviewed in Bystrykh et al.[5] and Schmidt et al.[6]). Moreover, genomic alignment of highly variable sequence reads is semiquantitative at best, and it is limited by the available annotated genome sequence for the model tested. This method does permit analysis of vector integration patterns and preferences as well as information regarding vector-driven clonal expansion. Despite the caveats, ISA data from preclinical models and gene therapy patients have largely been the basis for interpretation of hematopoietic biology after transplantation.[7–15]

Another method for tracking retrovirus-tagged cells is DNA barcode sequencing (DBS). DBS tracks a unique, small oligonucleotide encoded within the integrated proviral element as the clone-specific signature.[16] This method does not require fragmentation of gDNA or multiple rounds of exponential amplification. DBS avoids sequencing bias with standardized, coded fragments and negates

Received 21 February 2020; accepted 24 March 2020;
https://doi.org/10.1016/j.omtm.2020.03.021.

[5]These authors contributed equally to this work.

**Correspondence:** Jennifer E. Adair, Clinical Research Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Mail Stop D1-100, Seattle, WA 98109-1024, USA.
**E-mail:** jadair@fredhutch.org
**Correspondence:** Hans-Peter Kiem, Clinical Research Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Mail Stop D1-100, Seattle, WA 98109-1024, USA.
**E-mail:** hkiem@fredhutch.org

 

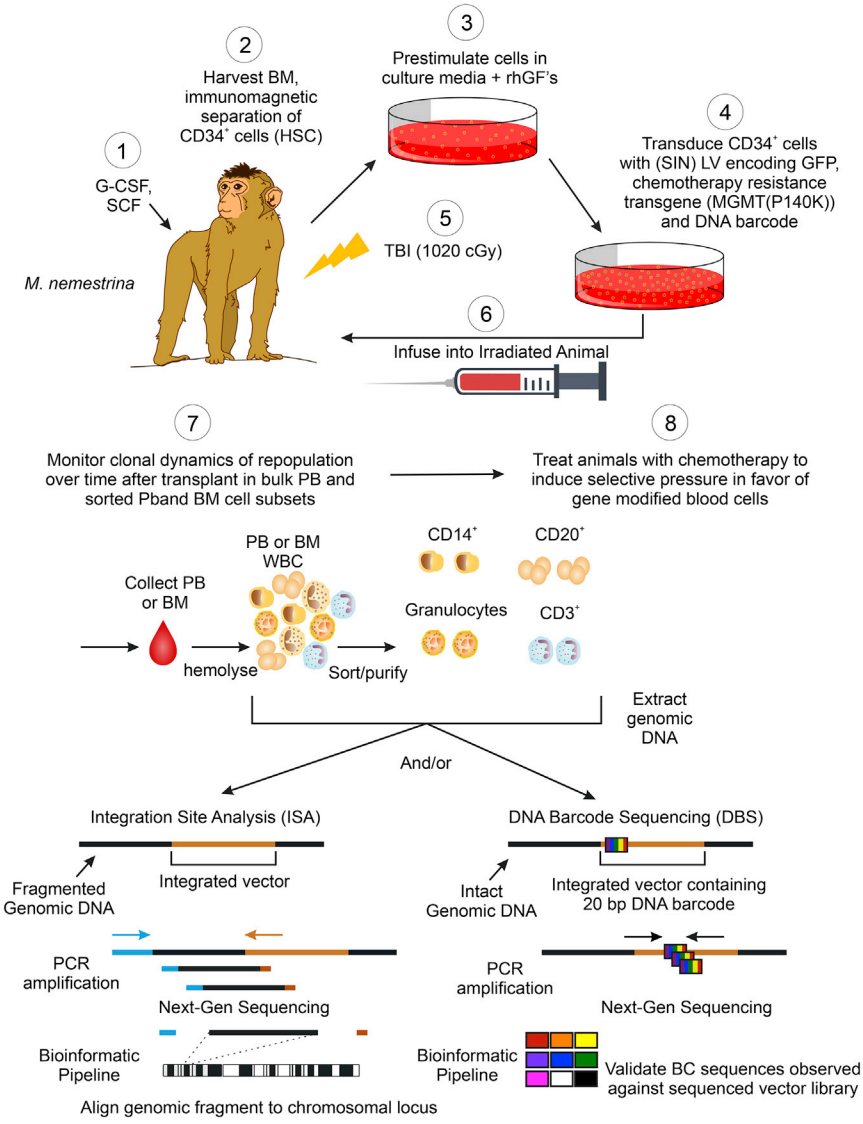


**Figure 1. Experimental Outline and Clonal Tracking Methods**

Treatment schematic for pigtail macaques: (1) Animals were "primed" with daily G-CSF and SCF administration to stimulate CD34$^+$ cell production in the bone marrow. (2) CD34$^+$ cells were purified from harvested bone marrow by immunomagnetic bead-based separation. (3) CD34$^+$ hematopoietic stem and progenitor cells (HSPCs) were cultured overnight in HSPC-supportive media with recombinant human growth factors (rhGFs). (4) Cells were transduced with a SIN LV encoding two transgenes, *GFP* and *MGMT* (*P140K*), and a 20-bp DNA barcode flanked by primer seeding sequences. The DNA-barcoded LV vector library was validated by next-generation (Next-Gen) sequencing to contain ~0.9 million different barcodes. (5) During transduction cultures each animal received 1,020 cGy total body irradiation (TBI). (6) Approximately 24 h after TBI was completed, transduced cells were washed and formulated in saline containing autologous serum for intravenous infusion. (7) Peripheral blood (PB) and bone marrow (BM) were collected at various time points after transplant for clonal analysis. Red blood cells were removed by ammonium chloride lysis, and resulting white blood cell populations were either submitted in bulk or were further sort purified by either density gradient centrifugation, fluorescence activated cell sorting, or by immunomagnetic bead-based methods. gDNA was extracted from resulting cell populations and subjected to either ISA or DBS. (8) At 1 year after transplant, animals were treated with chemotherapy to induce selection in favor of gene-modified blood cells.

genomic alignment. However, reported barcode libraries are limited to a few hundred thousand unique barcodes, insufficient for reconstitution of a large animal or patient. Moreover, barcode identification must be stringent.

Currently, ISA is the only method for tracking clones in patients treated with retrovirus-mediated gene therapy targeting CD34$^+$ cells, as DNA-barcoded retroviruses are, to date, not approved for use in humans. Therefore, we sought to compare these two clone tracking techniques directly using the same, barcoded, retrovirus vector in a clinically relevant large animal model. We previously demonstrated long-term hematopoietic reconstitution of pigtail macaques (*Macaca nemestrina*) after transplantation of autologous lentivirus (LV) gene-modified CD34$^+$ cells isolated from primed bone marrow (BM) in the myeloablative setting for a variety of gene therapy approaches.[17–22] We developed a next-generation ISA approach termed modified genomic sequencing (MGS)-PCR, which uses acoustic shearing to randomly fragment DNA and reduce bias associated with restriction enzyme-type fragmentation strategies.[12,17,19,23–25] For the present study, we developed a barcoded LV with a validated complexity large enough to reliably reconstitute a nonhuman primate and simultaneously performed ISA and DBS clone tracking from the same provirus in two transplanted animals to directly compare these techniques.

## RESULTS

Two clone tracking methods are compared in this study: ISA and DBS (Figure 1; see Materials and Methods).

### Development and Validation of a Barcoded LV with Sufficient Complexity to Transplant a Large Animal

To permit tracking by both methods from the same retrovirus, we generated a high-complexity barcoded LV library similar to that previously described by Porter et al.[26] The LV vector backbone used encoded a bicistronic cassette, including a spleen focus-forming virus (SFFV) promoter driving expression of a synthetic P140K mutant methylguanine methyltransferase (*P140K* [*synMGMT*]) gene for chemotherapy-induced selection *in vivo*[27]

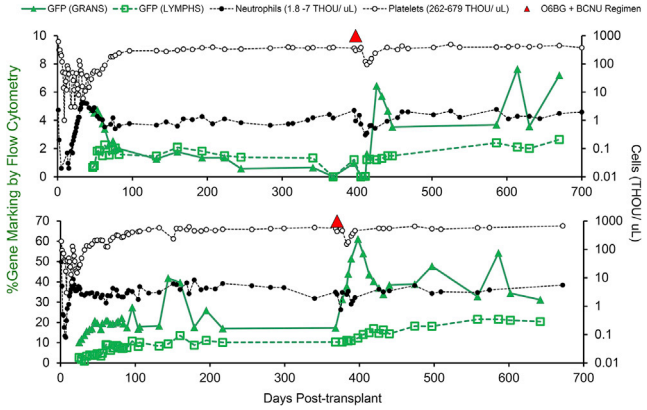


**Figure 2. Monkeys Display Typical Hematopoietic Reconstitution and Long-Term Engraftment of Gene-Modified Blood Cells after Transplantation**

Each graph represents hematopoietic parameters observed during the duration of follow-up after myeloablative transplant for animal Z08103 (top) and animal Z09132 (bottom). All PB cell counts were determined by an automated hematology analyzer. Animals received chemotherapy (O6BG/BCNU) at time points indicated by inverted red triangles. Green lines represent levels of gene-modified PB cells determined by GFP transgene expression. A complete list of samples collected during follow-up and resulting clonal analyses is provided in Table S3.

and a human phosphoglycerate kinase (PGK) promoter driving expression of an enhanced green fluorescent protein (*GFP*) gene for ease of identifying gene-modified cells (Figure S1A). Synthetic barcode fragments were ligated into this vector upstream of the bicistronic cassette and downstream of the proviral 5′ long terminal repeat (LTR) (Figure S1A). A total of 24 independent ligation reactions were completed to achieve an estimated barcode diversity of >1 million unique barcodes in the plasmid library. We sequenced the resulting plasmid library and identified 1,108,011 unique barcodes using stringent filtering to reduce the risk of false-positive barcode calling (Figures S1 and S2; Table S1). This transfer vector was transiently transfected into 293T cells alongside LV helper plasmids to produce viral particles. Some of the resulting virus preparation was subjected to viral RNA isolation, and the resulting virus genomes were sequenced by reverse-transcriptase PCR. Another portion of the resulting virus preparation was used to transduce K562 cells at a multiplicity of infection (MOI) of 1 infectious unit (IU) per cell. Cells were then lysed and gDNA was collected and DBS analysis performed. A total of 846,063 barcodes were identified in this barcoded LV preparation, of which 209,711 (17%) were new barcodes not identified in the transfer plasmid library. The total number of unique barcodes found in the transfer plasmid and LV prep libraries combined was 1,224,935, meaning that at least this many unique barcodes were present in the initial transfer plasmid library. However, to determine whether bias present in the transfer plasmid library was carried through the viral particle preparation, we analyzed the overlap between these two datasets (Table S2).

We observed no overlap between the 1,000 most abundant barcode sequences in the transfer plasmid library and the 100 most abundant barcodes in the LV preparation library, suggesting that any bias seen in the plasmid barcode library was not conserved in the creation of the virus preparation library. This comparison also provided a basis for determining whether bias in the LV preparation library was conserved following primary hematopoietic cell transduction and autologous transplantation.

### Engraftment and Persistence of Barcoded LV-Transduced Blood Cells Permit Comparative Clonal Analysis

Two monkeys were transplanted with autologous, LV-transduced $CD34^+$ cells and analyzed in this study using the schema depicted in Figure 1. A limited subset of the ISA data generated for these two monkeys was previously published in Radtke et al.[28] To facilitate calculations of engraftment, the maximum number of possible clones transplanted was calculated for each animal. To calculate the maximum, transplanted cells were counted and assayed for transduction by flow cytometry for GFP expression. Each gene-modified cell transplanted was assumed to represent a unique clone signature. For animal Z08103, $5.76 \times 10^6$ total $CD34^+$ cells were transplanted, of which 53.1% were $GFP^+$ cells. Thus, the maximum possible number of unique clones was equivalent to $3.06 \times 10^6$. For animal Z09132, a total of $8 \times 10^6$ cells were infused containing 20.4% $GFP^+$ cells, equivalent to $1.63 \times 10^6$ possible clones (Figure S3).

Our previous study estimated that 1 in 20 transplanted $CD34^+$ cells is repopulating blood cells in this model.[28] Thus, the estimated total number of transduced, repopulating cells transplanted was 152,928 and 81,600 in animals Z08103 and Z09132, respectively. Given that some barcodes showed abundance over others in the LV preparation library, we estimated the frequency of transduced repopulating cells with redundant barcodes. For animal Z08103, the frequency of repopulating cells with a barcode identical to one other cell (two of a kind) was 5.8%. This frequency dropped to 0.24% for three cells having the same barcode, and 0.008% for four cells having the same barcode. For animal Z09132, frequencies for two, three, or four cells having the same barcode were 3.2%, 0.07%, and 0.001%, respectively. Five of a kind was never observed for either animal.

Hematologic parameters and gene-modified cell levels were longitudinally tracked for both animals (Figure 2; Figures S4 and S5). Hematopoietic recovery (defined as the number of days to a stable absolute neutrophil count [ANC] $\geq$500/μL and a platelet count $\geq$20,000/μL, and with an increasing trend over 7 consecutive days after infusion of gene-modified cells [defined as day 0]) occurred at 42 days for animal Z08103 and 22 days for animal Z09132 (Figure 2).

At ~1 year after transplant, both animals received a combination of bis-chloroethylnitrosourea (BCNU) chemotherapy and the wild-type MGMT inhibitor *O*6-benzylguanine (O6BG) to induce selection in favor of *P140K* (*synMGMT*) gene-modified cells *in vivo*. Moderate myelosuppression was observed after chemotherapy with a concomitant doubling in the percentage of peripheral blood (PB) white blood cells (WBCs) expressing GFP in each animal (Figure 2). At approximately 2 years after transplant, 3.5% of PB WBCs were $GFP^+$ in

**Table 1. Comprehensive Sequence Reads and Clone Numbers Collected by a Combination of ISA and DBS after Transplant of LV Gene-Modified, Autologous CD34⁺ Cells over Years of Follow-Up**

| Overview | | | | | ISA | | | DBS | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Animal | Cell Dose Received ($10^6$ $CD34^+$/kg) | Transduction Efficiency in Infused Cell Product (flow Cytometry) | Follow-Up Range Included in Analysis (Months) | Gene-Modified PBL at Longest Follow-Up Time Point | Potential IS Reads | IS-Associated Reads | Total No. Unique Clones | Potential Barcode Reads | Barcode-Associated Reads | Total No. Unique Clones |
| Z08103 | 1.5 | 53% | 25 | 3.5% | 12,767,949 | 2,610,833 | 1,960 | 173,879,178 | 160,096,178 | 41,935 |
| Z09132 | 1.8 | 20.4% | 24 | 29.7% | 13,626,716 | 1,839,263 | 3,538 | 96,975,070 | 76,937,082 | 19,853 |

ISA, integration site analysis; DBS, DNA barcode sequencing; NA, not applicable; PBL, peripheral blood leukocytes.

animal Z08103 and 29.7% of PB WBCs were $GFP^+$ in animal Z09132 (Table 1; Figure 2). A total of 14–16 blood or BM samples were collected from each animal during the 2-year follow-up period (Figure 2; Table S3), and these were further subdivided into a total of 102 gDNA samples for analysis by ISA and DBS, respectively (Table S3). The total number of samples per animal was 23 and 24 for ISA and 26 and 29 for DBS from animals Z09132 and Z08103, respectively.

### DBS Consistently Suggests Higher Numbers of Clones

As shown in Table 1, for both animals DBS resulted in higher numbers of sequence reads, clone-associated reads after filtering, and unique numbers of clones. The higher number of sequence reads was not unexpected given the technical differences inherent to each of these tracking techniques: DBS requires only PCR to amplify short barcode reads of uniform length from gDNA, whereas ISA entails gDNA fragmentation, linker ligation, and PCR to amplify highly variable intervening gDNA sequences, prior to next-generation sequencing (NGS).

Using the filtering criteria shown in Figure S1B, for DBS, $\geq$80% of reads were found to be attributed to a barcode signature (i.e., barcode-associated reads), whereas <20% of ISA-generated reads were attributable to an integration site (IS) signature. Altogether, DBS identified tens of thousands of unique clones in both animals, whereas ISA detected less than 5,600 unique clones across both animals.

For both techniques, false positives are of concern; however, this is more likely the case for DBS owing to the random and degenerate nature of the 20-bp synthetic barcode fragment used as the clone signature. Thus, to call barcodes "real" in our dataset, they had to match exact flanking anchor sequences and length requirements, and they were subjected to stringent, empirically evaluated sequence error thresholds (Figure S6). We observed that the threshold for sequence errors was higher ($\leq$5 bp) in nonhuman primate samples compared to *in vitro* libraries (<3 bp). This is not unexpected, as the total numbers of barcodes and sequence similarities across barcodes were much lower *in vivo* than *in vitro* (Table S2). Importantly, nearly all detected barcodes *in vivo* overlapped with barcodes detected in the initial plasmid or LV vector libraries, with the majority mapping back to the initial transfer plasmid library.

Another means to validate barcode sequences is to cross-reference them with identified IS. We designed primers specific to 14 of the 26 most abundant IS clones in animal Z08103 and performed PCR and Sanger sequencing to identify the corresponding LV barcode and clone rank observed by DBS (Table S4). Clonal abundance was calculated by dividing the number of clone-specific sequence reads by the total number of IS- or barcode-associated sequence reads for the sample. All 14 identified barcodes corresponded to DBS clones ranking from 1 to 37 in abundance, suggesting that contributions of more abundant clones are not hindered by the use of ISA.

Finally, we also considered re-capturing barcodes as a validation of realness. Re-capturing was defined here as repeated detection across multiple samples (*in vivo* or *in vitro*), or representation by multiple (>1) sequence reads in the same sample. In animal Z09132, 289 of the total 19,853 barcodes detected were never re-captured. In animal Z08103, 356 of the total 41,935 barcodes detected were never re-captured. We did observe abundant DBS clones *in vivo* that were not found in the *in vitro* library dataset, with the most prevalent being ranked 53rd in abundance across all DBS clones detected in both animals over time.

Altogether, these observations suggest that the influence of false positives in our DBS data is low and lend support to the interpretation that initial library estimates were under-representative of the true diversity.

### DBS Is Not Perfect, but It Is Markedly More Efficient Than ISA

Since DBS captured so many more clones than ISA, we wanted to know whether DBS was providing the full picture, or if this was still an under-representation of true clonal diversity *in vivo*. Thus, we wanted to determine whether we could calculate the capture efficiency of each method. We defined the capture efficiency as the likelihood that an individual clone would be detected by either method using an unsorted blood sample as the gDNA source for analysis. To perform this calculation, we needed to simulate the dataset. However, to believe such a calculation, we first needed to confirm whether simulation could emulate observed data for each method, accounting not only for numbers of unique clones detected, but also observed abundance distributions.

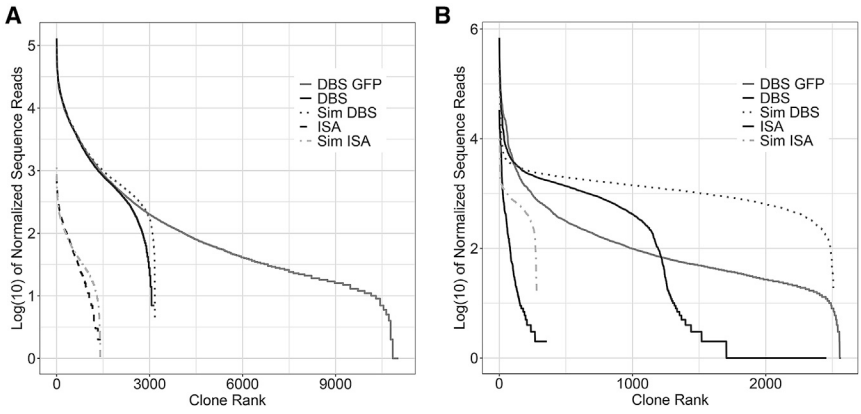


**Figure 3. Simulation Successfully Emulates Experimentally Observed Data Patterns**

(A and B) Graphs show both experimentally observed data (solid lines) and simulated data (dotted lines) for animals Z09132 (A) and Z08103 (B). The y axes represent the logarithmic$_{10}$ frequency of clone-specific sequence reads normalized to the total number of sequence reads for the sample, as a function of the clone rank from most abundant to least abundant on the x axis. Observed datasets include GFP-sorted blood samples (100% genetically modified; GFP) versus unsorted blood samples. Simulated (Sim) samples were generated based on method of detection (DBS or ISA). For simulations, clonal populations were estimated to be the largest number of unique clones identified on a single day after transplant, regardless of the number of sample re-tests. This was set as the clone population. The observed frequency of GFP$^+$ blood cells in the animal on the same day of sampling was used to estimate the fraction of gene-modified versus unmodified cells in simulated, unsorted samples. The "sample" function of R was then used to randomly select $0.5 \times 10^6$ blood cells for analysis by ISA or DBS. Both ISA and DBS efficiency were simulated as a random chance to capture each clone, and the chance of capture was varied from 0% to 100% for each process. A doubling of each captured clone sequence was simulated for each cycle of PCR, and NGS was simulated as a random chance of capturing each cell out of the simulated PCR reaction products.

To generate experimental data for this calculation, we compared the clones identified from a sample of PB WBCs flow sorted for GFP expression (i.e., all cells are positive for genetic modification) at >100 days after transplant to clones identified in the same number of unsorted PB WBCs from the same animal at the nearest time point. For animal Z09132, DBS detected a total of 11,026 unique clones in 500,000 GFP-sorted WBCs at 118 days after transplant. From unsorted PB WBCs containing 11.3% GFP$^+$ cells at 104 days after transplant, we detected a total of 3,152 clones by DBS. We assumed the day +104 time point contains 11.3% GFP$^+$ cells from an initial sample of 500,000 cells, for a total of 56,600 gene-marked cells. We also assumed this time point was composed of a random sampling of cells from the same clonal abundance distribution as the day +118 GFP-sorted time point.

We calculated the DBS capture efficiency as an equal and fixed probability of capturing each of the 56,600 gene-marked cells in the day +104 time point, such that the total number of unique clones detected would match the experimentally observed value of 3,152, and also the clonal abundance distribution curve would match the experimental value (Figure 3A). From these data we estimated the DBS capture efficiency to be 43%. Applying the same test and clonal distribution data for ISA, efficiency of clone detection was calculated as 8% for this method (Figure 3A).

To validate our findings, we used computation to randomly sample the experimentally identified clone population from the day +118, GFP-sorted test assuming 500,000 input genomes and 11.3% gene marking (i.e., a total of 56,600 modified genomes) at different capture efficiencies. The capture efficiency that resulted in the same number of total clones identified with similar abundance contributions for DBS was 43% (Figure S3A, dotted versus solid lines). Importantly, conditions for this simulation that matched experimentally observed data included PCR as a perfect process wherein each barcode would be duplicated during each PCR cycle. This assumption causes some granularity at low sequence counts that is not seen in real data, suggesting that PCR is probabilistic, but it also indicates that the DBS process compromises efficiency considerably less than does the ISA process. Since the only steps prior to DBS processing include cell isolation and gDNA extraction, these steps are more likely to influence capture efficiency. However, applying the same test and simulation data for ISA, efficiency of clone detection was calculated as 8% for this method. For ISA, the steps prior to PCR processing include cells isolation, gDNA extraction, and fragmentation. This suggests that sampling and fragmentation are likely the greatest contributors to low efficiency of ISA.

For animal Z08103, we repeated the same procedures. At day 145 after transplant, PB was sorted for GFP$^+$ cells. Given the lower gene marking in this animal at this time point (1.74%), the resulting sorted sample was prohibitively small to perform both DBS and ISA. Since DBS was observed to be more efficient than ISA in clone detection from animal Z09132, DBS analysis was prioritized. A total of 225,000 cell genomes were submitted for DBS, which detected a total of 2,565 unique clones. We then collected a larger sample of unsorted PB WBCs on day 160 after transplant and performed both DBS and ISA. In this sample, DBS detected 2,455 clones and ISA detected 269 clones from 500,000 genomes. Simulation of this dataset determined the efficiency of DBS to be 44% and ISA to be 4%, although the simulation did not align with observed clone contributions as well as for animal Z09132 (Figure 3B). These data suggest that additional variables alter efficiency of clone capture in circumstances of low levels of genetic modification *in vivo* for both ISA and DBS, but ISA capture efficiency is more susceptible to compromise.

To verify these capture efficiency estimates for each technique, we also simulated an entire *in vivo* blood sample containing a variable level of gene-modified clones (11.3% or 1.74%), with the same

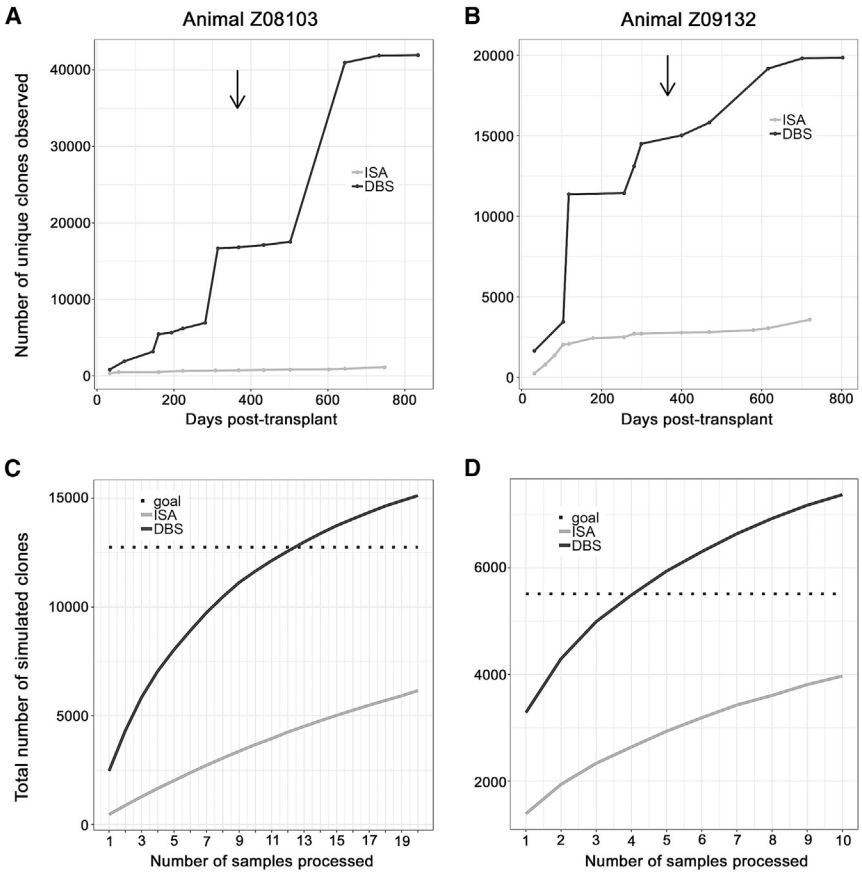


**Figure 4. Simulated Sampling Depth Estimates 6- to 10-Fold the Number of Analyses Required for ISA to Detect the Lowest Limit o Clones Observed by DBS.**
(A and B) Graphs represent the observed cumulative number of unique clones detected as a function of time (x axis) and chemotherapy (↓) for each analysis method (ISA, gray; DBS, black) for animals Z08103 (A) and Z09132 (B). (C and D) Graphs represent simulations to estimate the total number of clones detected as a function of the number of sample tests performed for both DBS (black) and ISA (dashed) for animals Z08103 (C) and Z09132 (D). Horizontal lines in each plot represent the threshold detection for 50% of all clones observed for any given time point in our study.

distribution of clone abundance as in our observed data. We randomly sampled these simulated blood samples at an input efficiency of 43% and 8%, respectively, since the simulation resulting in these efficiency levels most closely resembled observed data. Simulated sampling in this study was consistent with observed data, confirming efficiency calculations using base input data for each clone tracking method (Figure S7).

Taking this one step further, we performed cumulative clone tracking over time and chemotherapy in both animals. Cumulative clone numbers detected by ISA never equaled those detected by DBS for either animal, despite similar sampling frequency and type (Figures 4A and 4B). Notably, repeat analysis of the same sample or subdivision of individual blood samples into multiple assays by lineage sorting prior to gDNA isolation improved the number of clones detected by both methods evaluated herein (Figures 4A and 4B; Table S1).

Altogether, this combination of experimental and simulated data suggests that DBS is at least 5-fold more efficient at clone detection than ISA, and that random sampling with a probabilistic chance of capturing each cell is a reasonable estimation of each clone tracking process. If true, DBS would still be under-representing true clonal diversity, but it would be achieving nearly 50% coverage of the clonal pool, weighted in favor of more abundant clones, with a single test of each sample. ISA would also favor more abundant clones, but it would capture less than 10% of true clonal repertoire.

### How Much Sampling Is Needed to Feel Confident?

Based on our data, without empirically evaluating changes to the DBS or ISA processes themselves, increasing the sampling frequency by collecting and testing additional blood samples and re-processing samples already collected to increase sampling depth are the obvious strategies to improve coverage of the true clonal repertoire. However, the clinical feasibility and safety of additional blood draws, and/or the volume required to extract gDNA amounts that permit re-testing, have to be considered when deferring to these strategies. Again, in this study, we were able to apply computational simulation to aid in determining how much sampling would be needed. To estimate the sampling required for ISA to achieve the same number of cumulative clones as DBS, we assumed a population of n clones as observed in GFP-sorted blood from each animal (i.e., 11,026 for animal Z09132 and 2,565 for animal Z08103), which we then randomly sampled at the calculated efficiency values of 43% and 8% for DBS and ISA, respectively (Figures 4C and 4D). For animal Z09132, with higher gene-modified cell engraftment, a total of 6- to 7-fold the number of sample analyses by ISA (i.e., a total of 138 analyses during 2 years, or 1.3 samples per week) would be required to detect the lowest cumulative number of clones as determined by DBS. For animal Z08103, with much lower levels of gene-modified blood cells, as much as 10-fold the number of sample analyses (i.e., 240 analyses during 2 years, or 2.3 samples per week) would be required to detect the lower limit of cumulative clones detected by DBS.

For these two animals, this translates to a total of 138 and 240 sample analyses (where analyses include everything downstream of gDNA isolation) required for ISA to achieve the lowest number of

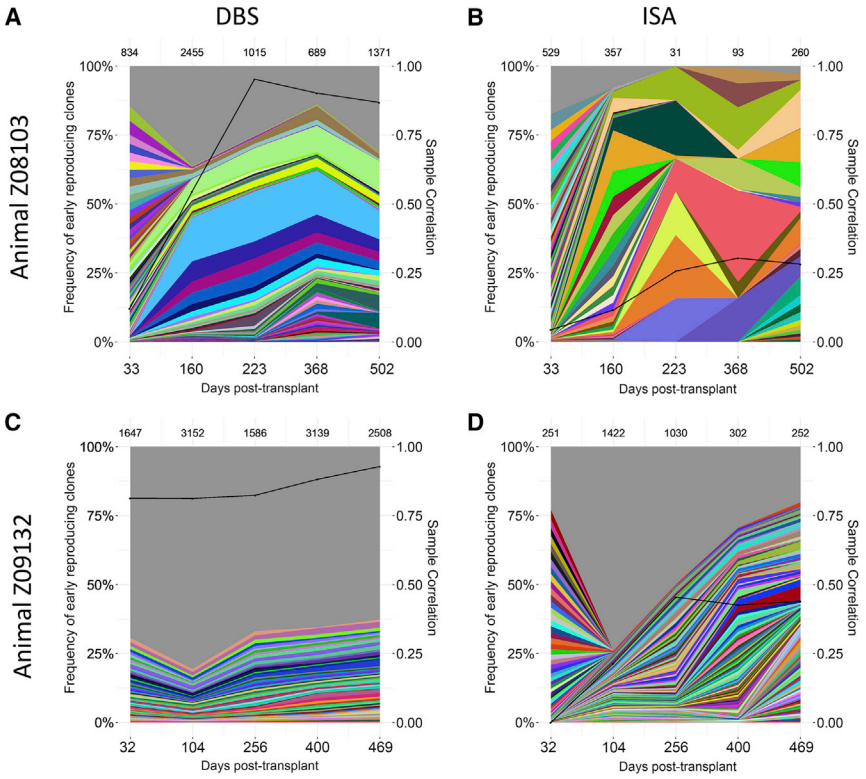


**Figure 5. DBS Suggests Clonal Stability while ISA Suggests Clonal Succession.**

(A–D) DBS (A and C) and ISA (B and D) for animals Z08103 (A and B) and Z09132 (C and D). All graphs represent the contribution (% frequency) of all clones observed in total white blood cells over time after transplant (x axis). Each color band represents a unique clone demonstrating ≥0.5% frequency of sequence reads in a given sample. All other clones are grouped into the gray band at the top of each graph. The total number of unique clones identified at each time point is listed at the top of the plot. Black lines overlaying each plot represent the diversity correlation from sample to sample within each plot.

cumulative clones detected by DBS. If individual blood samples are analyzed only once, this would require blood collection up to 2.3 times per week for ISA, which is not feasible. In this study we sampled exclusively for clone tracking less than once per month. If the same numbers of collected blood samples were re-analyzed, each blood sample would require sufficient gDNA for 6–10 repeats, or 18–30 μg of starting gDNA, which is reasonable to collect from a 10-mL blood sample.

### DBS versus ISA over Time and Chemotherapy: Stability, Not Succession

The pattern of hematopoietic reconstitution following autologous gene therapy has been proposed to be clonal succession by studies using ISA,[9,10] but a more recent study implementing DBS found hematopoietic reconstitution to be better explained by clonal stability.[29] When DBS is compared directly to ISA, we observe a clear pattern of clonal stability over time and chemotherapy-induced clonal selection in both monkeys (Figure 5).

Of note, this analysis represented all clones contributing ≥1% to the clonal pool observed for each sample. As noted above, ISA does favor more abundant clones, and multiple variations on the ISA method have been used by different groups. To determine whether this may influence clonal reconstitution patterns in our model, we repeated this analysis with only the 25 most abundant clones over all time points and samples identified in each animal during the course of study (Figure 6). We combined all datasets across all time points and sample types and identified the top 25 clone signatures represented for each animal as a function of the number of sequence reads corresponding to each clone. We then asked a simple "yes" or "no" question as to whether the clone signature was identified at any frequency in any of the samples analyzed for that animal for each clone tracking method. These data confirmed that ISA fails to re-capture even abundant clone signatures over time, suggesting a false interpretation of clonal succession. However, in animal Z08103 with lower levels of gene-modified cells, DBS detection of the most abundant clones was also intermittent, but with much greater reliability than for ISA. This was validated by calculating the Shannon diversity indices for each animal, which account for re-capture and abundance (Figure S8).

### Tracking HSC Clone Contributions Provides Added Value for Safety

We previously observed early and persistent engraftment of stem cell-like clones in these two animals by ISA alone with a much smaller dataset.[28] We wanted to assess whether DBS validated this observation *in vivo*, and whether these HSC-like clones were detected in the time points with low clone detection by ISA. To do this, we defined stem cell-like clones as those clone signatures present in more than one mature blood cell lineage, one of which must be short-lived granulocytes, and present at more than one time point of analysis, one of which must be more than 1 year after transplant. We analyzed all available DBS and ISA datasets (Figure 7). Not surprisingly, DBS identified 3- to 6-fold more HSC-like clone signatures than did ISA. Interestingly, DBS robustly confirmed the early and persistent contribution of these clones to hematopoiesis after transplant and chemotherapy (Figures 7A and 7C). The same was true for animal Z09132 when ISA was performed (Figure 7D), while additional ISA data in animal Z08103 suggested fluctuations in clonal contributions (Figure 7B). Notably, the small numbers of clones identified in this animal at days 223 and 368 after transplant that suggested clonal restriction in whole blood (Figure 5B) were found to be of HSC origin based on the above stem cell clone definition

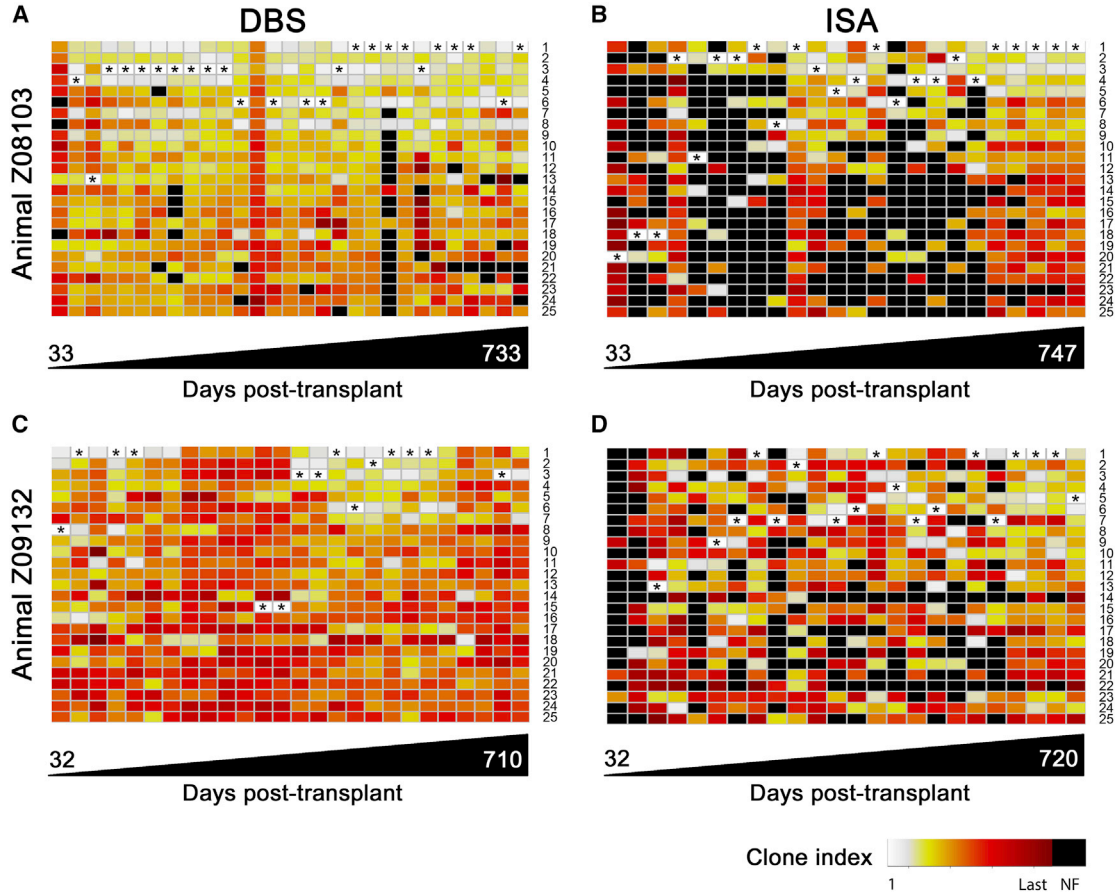


**Figure 6. Comparative Analysis of Top Clones over All Samples and Time Points Confirms ISA Bias against Prevailing Clonal Stability Model of Hematopoiesis**

(A–D) DBS (A and C) and ISA (B and D) for animals Z08103 (A and B) and Z09132 (C and D). All heatmaps represent the $\log_2$(rank) contribution (% frequency) of the 25 most abundant clones observed for the whole study (rows) at each time point of analysis after transplant (columns). Black represents clone not found (NF) in that sample. The highest ranking clone in each sample is indicated by the asterisk in the corresponding cell of the heatmap.

(same time points, Figure 7B). Thus, lineage sorting at an earlier time point of analysis (i.e., before perceived clonal restriction) validated that these clones were likely of stem cell origin. Taken together with DBS data and hematologic data, the likelihood of insertional mutagenesis can be inferred to be low, despite the low numbers of highly abundant clones tracked by ISA in these samples.

## DISCUSSION

In this study, we directly compared ISA, the only available method for tracking gene-modified blood cells in gene therapy patients, to DBS, a more recent clone tracking approach with limited application in large animals and with no human application to date. Our data demonstrate that both methods track abundant clones at similar frequencies; however, DBS is significantly more reliable and efficient than ISA. The ability for ISA to effectively measure the clonal repertoire is significantly impacted by repertoire size, relative to DBS. Moreover, we demonstrate that patterns of hematopoietic reconstitution after transplant and patterns of hematopoietic reconstitution after selective pressure differ significantly based on the clone tracking method applied. Importantly, our DBS observations align closely with our previously published work demonstrating early and consistent HSC contributions after autologous transplant.[28] Collectively, these data suggest that DBS is a superior methodology when the study goal is to describe hematopoietic biology in a preclinical system. In gene therapy patients where ISA is the only tracking method available, collecting sufficient blood volume for repeat analysis of the same sample to increase depth of detection and subsampling based on lineage sorting of blood early after reconstitution can significantly improve the quality of data obtained. This also suggests that careful consideration is required in studies where ISA is solely used for interpretation of hematopoietic biology with limited datasets or sampling.

DBS to track blood stem cells *in vivo* was first demonstrated in mice in 2010.[30] Since then, numerous studies have applied this technique to study hematopoietic biology in mice by transplanting barcoded murine and human blood stem and progenitor cells.[16,31–33] However, the

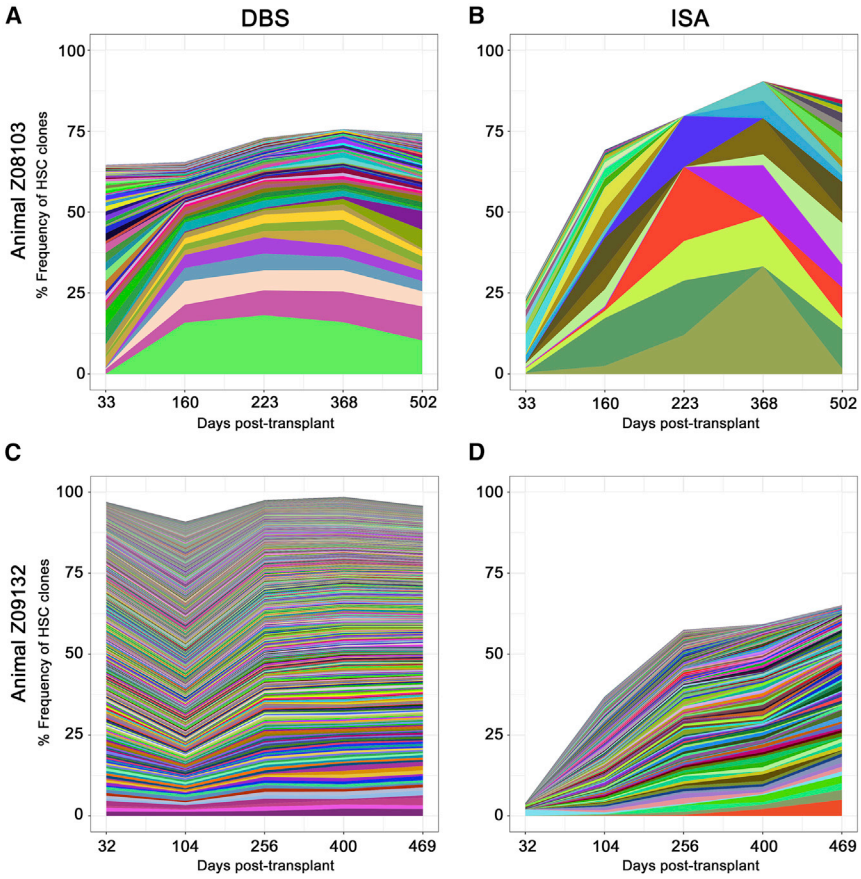


**Figure 7. DBS Confirms Early and Persistent Engraftment of HSC Clones**

(A–D) DBS (A and C) and ISA (B and D) for animals Z08103 (A and B) and Z09132 (C and D). All graphs represent the contribution (% frequency) of all HSC clones observed in total white blood cells over time after transplant (x axis). Clone signatures were demonstrated to be HSC when they were identified in at least one short-term and one long-term repopulating lineage at more than one time point, one of which had to be later than 1 year after transplant. Each color band represents a unique HSC clone demonstrating ≥0.5% frequency of sequence reads in a given sample. All other clones are grouped into the white band at the top of each graph.

differences between human and murine hematopoiesis (e.g., different cell surface marker HSC phenotypes), as well as the confounding nature of xenogeneic transplant studies, limit extrapolation of these observations to clinically relevant human hematopoiesis.

Nonhuman primate models have been suggested to more closely relate to human hematopoiesis.[12,19,34] In 2014, two reports separately applied ISA or DBS clone tracking strategies in transplanted rhesus macaques (*Macaca mulatta*), each reporting significant variations in the onset and persistence of hematopoietic repopulating clones.[10,11] The study using ISA alone identified between 846 and 5,758 total clones in each of four animals transplanted with either LV- or simian immunodeficiency virus (SIV)-based vector-transduced cells, which followed a bi-phasic reconstitution during the first year, then sequential expansion of HSC-like clones. This study reported that the cumulative number of clones detected was directly proportional to the cell dose infused at the time of transplant. In a side-by-side study, more comprehensive sampling within the first 9.5 months after transplant was reported by Wu et al.[11] utilizing DBS alone in three rhesus macaques. This study reported no correlation between cumulative clone numbers detected and transplanted cell dose. This report also identified multi-lineage clones, but not until 2–3 months after infusion. A follow-up to this study reported that cumulative clone numbers and output were stable within 6 months after transplant during as much as 49 months of follow-up.[29]

There are notable differences between these rhesus studies and our present study. First, different species were used (rhesus versus pigtail macaques), as well as different sources of $CD34^+$ cells: granulocyte colony-stimulating factor (G-CSF)-mobilized leukapheresis-derived $CD34^+$ cells in rhesus macaques versus G-CSF- and stem cell factor (SCF)-primed, marrow-derived $CD34^+$ cells in pigtail macaques. The fraction of true repopulating HSCs/cells isolated declines in mobilized leukapheresis products compared to BM products, with the highest proportion of true HSCs found in cord blood.[35] Thus, the lower number of clones transplanted and tracked in the rhesus studies could reflect a lower repopulating cell content in mobilized leukapheresis products. Moreover, neither of the rhesus studies induced selection of gene-modified clones *in vivo*.

Additionally, there are technical differences between ISA methods: the rhesus ISA study used a restriction enzyme-based method for gDNA fragmentation, whereas we employed acoustic shearing. Our method allows for quality control of clone abundance by capturing different fragment sizes of the same genomic sequence from single-end sequencing, whereas the rhesus method relied on paired-end sequencing of resulting genomic fragments. The paired-end approach can reduce bias, but it also displays lower sequencing efficiency, resulting in lower clone numbers. In this study, we verified abundant clonal signatures with both ISA and DBS clone tracking methods by cross-sequencing, as well as re-capture, supporting the "realness" of clones detected. Specifically, the rhesus study sampled animals between 7 and 12 times after transplantation, during as many as 140 months, with very few samples (three or fewer time points) collected within the first year. Thus, our data suggest that the bi-phasic reconstitution and clonal succession observed in the rhesus ISA study may have resulted from under-sampling.

Similar to the rhesus DBS studies, we observed stability of clone numbers and abundance in both animals within 6 months after

transplant, which were maintained through follow-up. This observation aligns with a recent publication in patients treated with gene therapy for Wiskott-Aldrich syndrome wherein ISA demonstrated clonal stability after transplant.[36] Interestingly, this work was a follow-up to an earlier publication by the same group with the same patient cohort that described different hematopoietic kinetics, namely bi-phasic reconstitution, when fewer samples were analyzed.[9] Indeed, the more recent work tracked more than 140,000 clones by ISA, whereas the earlier report tracked only 89,000. This supports our findings that sampling depth in ISA studies limits hematopoietic interpretation and demonstrates the significant investment in clone tracking required in gene therapy patients to make interpretations of hematopoietic biology.

Most importantly, our studies suggest that collection of clone data by ISA from lineage-sorted blood samples early in clinical studies can aid in interpreting clonal dynamics. Indeed, while there is much effort being put toward identifying potentially unsafe clone signatures as early as possible after gene therapy, the clinical implications of gaining this information are unclear. In this study, we demonstrate that early subsampling of clones from a lineage-sorted blood sample adds significant value in clone tracking data for both analysis of safety and interpretations of hematopoietic biology after transplantation.

In conclusion, we used direct comparison to demonstrate that DBS is more efficient and displays a larger dynamic range relative to ISA in tracking individual retrovirus-tagged clones *in vivo*. These studies demonstrate that the sampling frequency required to make biological interpretations about hematopoiesis using ISA alone are likely prohibitive in gene therapy patients, but tracking clonal abundance for safety is still feasible within the dynamic range of ISA and can be greatly increased in value with early, multi-lineage subsampling for clone tracking. These observations suggest that clonal stability is the preferred model of homeostasis after transplant in the nonhuman primate, akin to that recently described in human gene therapy patients.

## MATERIALS AND METHODS

### Animals, Transplantation, and Care

This study was carried out in strict accordance with recommendations in the *Guide for the Care and Use of Laboratory Animals* of the National Institutes of Health. All protocols were approved by the Institutional Animal Care and Use Committees of the Fred Hutchinson Cancer Research Center and the University of Washington. Animals were housed at the Washington National Primate Research Center.

Two pigtail macaques (*M. nemestrina*) were included for clone tracking (Z08103 and Z09132). Both animals were juvenile males. Autologous grafts consisted of LV-transduced, autologous BM-derived $CD34^+$ cells collected after *in vivo* stimulation with daily G-CSF and SCF as previously described.[37] Harvested marrow was enriched for $CD34^+$ cell fractions with a purity of 60%–80% confirmed by flow cytometry.[21] Animals received 1,020 cGy total body irradiation (TBI) prior to infusion of gene-modified cells. Standard supportive care was administered as needed and in consultation with veterinary staff. Hematopoietic recovery was monitored by complete blood counts using an automated hematology analyzer. PB was collected at various time points by venipuncture, and BM was collected by intra-femoral or intra-humoral aspiration.

For chemotherapy treatment, macaques received a total 120 $mg/m^2$ O6BG and 15 $mg/m^2$ BCNU administered as two staggered O6BG doses (O6BG-2×), with BCNU infusion 45–60 min after the first O6BG infusion, as previously described.[17]

### LV Vectors and Transduction

The transfer vector (pRRLsincPPT-SFFV-MGMT9P140K-PGK-GFP-sWPRE) encodes a self-inactivating (SIN) human immunodeficiency virus type 1 (HIV-1)-derived LV genome consisting of a SFFV promoter regulating an enhanced green fluorescent protein (*GFP*) gene and a human PGK promoter regulating a P140K mutated *MGMT* gene, which renders gene-modified cells resistant to a combination of the alkylating agent BCNU and the MGMT inhibitor O6BG.[17] Inclusion of the P140K gene provides a mechanism by which to induce clonal selection *in vivo*. For DBS, we generated a barcoded LV library as shown in detail in Figure S1A using methods similar to that published by Porter et al.[26] Synthetic oligonucleotide cassettes containing a P5 Illumina sequencing adaptor, a 2-bp index sequence, and a 20-bp random nucleotide sequence flanked by identical 6-bp anchor sequences were cloned into the non-coding region of the LV transfer vector described above and shown in Figure S1A upstream of a bicistronic transgene cassette. Use of different and unique restriction enzyme-based sticky-end cloning was used to facilitate ligation without concatemerization or reverse orientation of the oligonucleotides containing the barcodes. The Illumina MiSeq 5′ adaptor sequence was included to facilitate amplification and render 36-bp single-end sequence reads usable. The anchor sequences permit stringent filtering of resulting sequence reads to reduce false positives in the dataset, as shown in Figure S2B. Vesicular stomatitis virus G protein (VSV-G)-pseudotyped viral particles were produced by four-plasmid transient transfection of 293T cells from the American Type Culture Collection (ATCC; 293T ATCC CRL-3216). Supernatants were collected and viral particles were collected by ultracentrifugation. Infectious titers were determined by limiting dilution of purified supernatants on HT1080 cells (ATCC CCL-121) followed by flow cytometry for GFP expression.

K562 cells were purchased from the American Type Culture Collection (ATCC CCL-243) and cultured in Iscove's modified Dulbecco's medium (IMDM) supplemented with 10% heat-inactivated fetal bovine serum (FBS) to maintain a density of $\sim 1 \times 10^6$ cells/mL. For transduction, LV was added to cells in culture at an MOI of 1 IU per cell. Cultures were maintained for 24 h prior to harvest.

For transduction of primary nonhuman primate $CD34^+$ cells, isolated cells were plated into StemSpan serum-free expansion medium (SFEM) containing 1% penicillin/streptomycin (P/S, Life

Technologies) supplemented with recombinant human SCF, thrombopoietin (TPO), and flightless 3 ligand (FLT3-L) at 100 ng/mL each. Cells were prestimulated overnight and transduced with 4 μg/mL protamine sulfate and 1 μg/ml cyclosporine A on flasks coated with CH296 (retronectin) at an MOI of 10 IU per cell. A second LV dose was added after 6–8 h for a final MOI of 20 IU/cell. The next day, cells were collected and prepared for infusion. Cells were pulsed with prostaglandin $E_2$ for 2 h on ice immediately prior to infusion.

#### Gene-Modified Cell Level Measurement

Flow cytometry was used to enumerate gene-modified cells in PB and BM samples. Samples were first hemolyzed with ammonium chloride or subjected to a Ficoll gradient separation whereby mononuclear cells were purified as a distinct layer and granulocytes were isolated by ammonium chloride lysis of the red blood cell fraction. Resulting cells were washed and resuspended in fluorescence-activated cell sorting (FACS) buffer consisting of sterile Dulbecco's modified phosphate-buffered saline (D-PBS), 1% FBS, and 1 mM ethylenediaminetetraacetic acid (EDTA). Cells were incubated for 30 min under low light with various combinations of fluorophore-conjugated anti-human or anti-nonhuman primate antibodies, including $CD3^+$, $CD4^+$, $CD8^+$, $CD14^+$, $CD16^+$, $CD20^+$, $CD34^+$, and $CD45^+$, and then analyzed on a FACSCalibur, FACSCanto, or LSR II device (all from Becton Dickinson). Antibodies from BD Biosciences include anti-human CD3 (clone SP34-2), anti-human CD4 (clone L200), anti-human CD34 (clone 563), and anti-non-human primate CD45 (clone D058-1283). Antibodies from BioLegend include anti-human CD8 (clone RPA-T8) and anti-human CD20 (clone 2H7). Antibodies from eBioscience include anti-human CD14 (clone 61D3). Where multiple antibodies and fluorophore combinations were applied, compensation was facilitated by UltraComp eBeads compensation beads (Invitrogen) according to the manufacturer's specifications. Flow cytometric data were analyzed by FlowJo versions 8.5.2 to 10 (Tree Star) or CellQuest version 3.1f software (Becton Dickinson), with gating to exclude fewer than 3% control cells in the relevant region.

#### Sample Collection for Clone Tracking or Engraftment

Bulk WBCs were isolated by ammonium chloride lysis of red blood cells (RBCs). Granulocytes and mononuclear cells (MNCs) were purified by Ficoll gradient sedimentation (GE Healthcare) per the manufacturer's specifications, followed by ammonium chloride lysis of RBCs for granulocyte purification. Subsets including $CD3^+$, $CD4^+$, $CD8^+$, $CD14^+$, $CD16^+$, $CD20^+$, and $CD34^+$ cells were enriched using either FACS on a FACSAria device (Becton Dickinson) or by immunomagnetic bead-based purification. For FACS, cells resuspended in FACS buffer were labeled with appropriate fluorophore-conjugated anti-human antibodies per the manufacturer's specifications. $Fluorophore^+$ cells from any of these fractions were isolated by FACS on a FACSAria device (Becton Dickinson). Where multiple antibodies and fluorophore combinations were applied, compensation was facilitated by UltraComp eBeads compensation beads (Invitrogen) according to the manufacturer's specifications. Dead cells and debris were excluded via FSC (forward light scatter)/SSC (side light scatter) gating as shown in Figures S3–S5.

For immunomagnetic bead-based separations, antibody-based immunomagnetic microbeads were all supplied by Miltenyi Biotec. Microbead reagents included anti-nonhuman primate CD3 microbeads, CD20 microbeads, and CD14 microbeads. An anti-human CD34-fluorescein isothiocyanate (FITC) antibody (clone 12.8) was used in combination with anti-FITC microbeads for indirect magnetic labeling and separation. All immunomagnetic separations were performed according to the manufacturer's specifications.

For all sorted cell populations, flow cytometric analyses for purity assessment were performed on an LSR II (Becton Dickinson, Franklin Lakes, NJ, USA) and FACSAria II (Becton Dickinson), and analyzed by FlowJo versions 8.5.2 to 10 (Tree Star) or CellQuest version 3.1f software (Becton Dickinson), with gating to exclude fewer than 3% control cells in the relevant region.

#### gDNA Isolation for ISA or DBS

gDNA was extracted from cells using either the QIAamp blood midi kit (QIAGEN, Germantown, MD, USA) for cell yields $>5 \times 10^5$ total cells, or the MasterPure DNA purification kit (Lucigen, Middleton, WI, USA) for small cell yields ($<5 \times 10^5$ total cells). Both kits were used according to the manufacturers' specifications. Where both clone tracking methods were applied, equal gDNA masses were apportioned to each analysis. In some cases insufficient gDNA was available for both methods, in which case one method was prioritized, as listed in Table S3.

#### ISA

Processing of gDNA to amplify integration loci included MGS-PCR[23] methods. Two NGS platforms were used, including single-end Ion Torrent PGM and paired-end Illumina MiSeq. A detailed list of samples is included in the dataset available for download (BioProject: PRJNA517494). A summary of data processing is included in Figure S1B. First, forward and reverse reads were stitched using PEAR with the *-q 30* option to trim sequence reads after two bases with a quality score below 30 were observed.[38] Stitched FASTQ files and raw FASTA files for all sequencing data were filtered using a custom c++ program. Each read was compared to the reference provirus LTR sequence. Reads with <90% match to the LTR sequence were discarded. The LTR sequence was trimmed off of remaining reads. Reads were then compared with vector sequence (as opposed to genomic insertion sequences). Reads with $\geq$80% match to the vector sequence were discarded. The remaining reads were output in FASTA format for alignment.

Identified genomic fragments were aligned to the rhesus macaque (rhemac8) genome (GenBank: 2701138) to determine the chromosome, locus, and orientation of integration (e.g., Chr14_8020175_+), as the pigtail macaque genome has not been sequenced. rheMac8 was downloaded from the UCSC (University of California Santa Cruz) genome browser (http://genome.ucsc.edu/).[39] Filtered and

trimmed sequence reads were aligned to the reference genome using BLAT (BLAST-like alignment tool) with the following options:

> out = blast8, -tileSize = 11, -stepSize = 5, and -ooc = rh11-2253.ooc.[40]

The rh11-2253.occ file contains a list of 11-mer occurring at least 2,253 times in the genome to be masked by BLAT and was generated using the following command:

> $blat rheMac8.2bit /dev/null /dev/null -tileSize = 11 -stepSize = 5 -makeOoc = rh11-2253.ooc -repMatch = 2253

as recommended by UCSC (http://genome.ucsc.edu/goldenpath/help/blatSpec.html[40] and http://genome.ucsc.edu/FAQ/FAQblat.html#blat6). Resulting blast8 files were parsed using a custom python script. Any read with a BLAT alignment length less than 30 or alignment start greater than 10 was discarded. The top scoring alignment and all alignments within 95% of the top score were saved for each read. The ratio of the best alignment to the second-best alignment is the degree with which the insert can be mapped to one location in the genome (multi-align-ratio). Starting with the highest count reads, reads with matching alignments were combined. Reads with multiple possible alignments were not discarded at this point but were grouped together with other reads with the same alignment(s). For each group of reads with matching alignments, the original FASTA sequence files were read, and then the sequences were aligned with Clustal Omega.[41,42] This alignment was used to build a single consensus sequence for the alignment group. The consensus sequence was then used to search the pool of all sequences that could not be aligned by BLAT, and any sequence with $\geq$90% identity was merged into the group. Starting with the highest count groups and comparing them with the lowest count groups, groups with $\geq$90% sequence similarity were merged. Finally, when comparing all sequence files for one test subject, all groups with exact alignment matches were merged into one clone ID. Clone IDs with exact consensus sequence matches were also merged. Non-uniquely aligned groups (multi-align-ratio $\geq$ 0.9) that had $\geq$90% similar reference sequences were also merged into a single clone ID.[43]

## DBS

DNA barcodes were amplified by PCR using high-fidelity DNA polymerase (Phusion, New England Biolabs). PCR conditions included initial denaturation at 98°C for 2 min, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 62°C for 30 s, and elongation at 72°C for 30 s, with a final elongation step of 72°C for 5 min and a hold at 4°C. Primers (forward, 5′-AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTTCCGATCTTCTAGA-3′; reverse, 5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCC TGCTGAACCGCTCTTCCGATCTAAATGGCGTTACTGCAGCT A-3′) contained Illumina adaptor sequences and 4-bp multiplexing tags, resulting in 250-bp fragments that were size-selected on 1.1% agarose gel and purified using the QIAquick gel extraction kit (QIAGEN) per the manufacturer's specifications. DNA concentrations were checked on a BioAnalyzer (Agilent). Single-end sequencing of DBS libraries was conducted on an Illumina Genome Analyzer II or Illumina MiSeq sequencer. A detailed list of samples is included in the dataset available for download (BioProject: PRJNA517494). Barcodes were characterized using a method that is summarized in detail in Figure S2B and similar to that outlined in Thielecke et al.[44] with the following exceptions: raw barcode reads from FASTQ format files were first filtered for quality. Bases with a Q score <20 were converted to "N." Any read with four or more N bases was discarded. Barcode reads without an exact match of the six fixed nucleotides on either side of the 20-bp variable barcode were discarded. Valid barcode reads were trimmed to 20 bp, and identical barcodes were combined into groups. Barcodes with a count >1 were sorted by count, and then starting with the most abundant barcode, the Hamming distance to each single barcode was calculated. All barcodes within a Hamming distance of 2 were merged. Unmerged single barcodes were put aside. All barcodes with N bases within a Hamming distance of 2 to a valid barcode were merged. Unmerged N barcodes were discarded. Single-count barcodes were added back at this point. Starting with the most abundant barcode, the entire barcode list was searched from the least abundant barcode up, and barcodes within a Hamming distance of 2 were merged. We used a Hamming distance of 2 for these initial comparisons because we wanted to correct small errors first and because our software uses an index table to search the entire barcode library for all barcodes within a Hamming distance of 2 with just four table look-ups. Finally, all barcodes within a Hamming distance of 5 were merged.

## Simulations

Custom R scripts were written to perform all simulations. For redundant barcoding simulations, the following assumptions were made: (1) the absolute barcode library size was estimated to be 1,213,714 barcodes, (2) all 1,213,714 barcodes are available to each cell during transduction, and (3) each cell will randomly take up one barcode. This was simulated as a dice probability where the number of dice rolled was equal to the number of transduced repopulating cells (i.e., 152, 928, or 81,600), and the number of sides on each dice was equal to the number of unique barcodes (i.e., 1,213,714). Each simulation was repeated 100 times, and the mean frequency of obtaining two to five identical barcodes (two, three, four, or five of a kind) was determined.

For simulation to estimate the clonal population of Z08103 and Z09132, we first combined all barcode samples taken on a given day into one pool for each animal and then selected the day when the maximum number of clones was detected as the estimated clonal population. We used %GFP expression at the day of testing to determine the fraction of gene-modified clones in blood versus non-gene-modified blood cells. We then used the "sample" function of R to randomly select 500,000 blood cells for either ISA or DBS processing. We determined how many gene-modified cells would be selected by the random blood sample and also what the starting clone fractions would be. We simulated ISA and DBS efficiency as a random chance

to capture each gene-modified cell and varied the chance of capture from 0% to 100% for each process. We simulated PCR as a doubling of each captured clone for each PCR cycle. We then simulated NGS as a random chance of capturing each cell out of the simulated PCR product.

#### Diversity Measurements

Two methods were used to track diversity of clone populations in this study: sample correlation and Shannon diversity index (H). Sample correlation was calculated for adjacent samples by first selecting the 1,000 most abundant clones for each time point. The clones detected across all time points are combined into one list, and then a matrix is constructed with a column for each time point and the rows are defined as the unique list of all clones. The abundance for clones not detected at any given time point is set to 0. The Pearson correlation of rows is calculated for this matrix. The final correlation for each time point is the mean of the correlation compared to the previous time point and the correlation compared to the subsequent time point. The Shannon diversity index of a sample was calculated as $-\text{sum}(p \cdot \ln(p))$ for each clone, where p is the fraction of the total population of each unique clone.

#### Plot Generation

All plots were created with custom R scripts. Some used ggplot2 (https://ggplot2.org).[45]

#### Data Availability

All sequence data obtained in this study are available for download at https://www.ncbi.nlm.nih.gov/bioproject/ (BioProject: PRJNA517494).

#### Code Availability

All custom R and Python codes are available on GitHub (https://www.github.com/mrenstrom). All other codes are publicly available and are cited in the appropriate methods description.

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtm.2020.03.021.

## AUTHOR CONTRIBUTIONS

J.E.A., M.H.P., and H.-P.K. designed the study. J.E.A., K.G.H., D.R.H., L.E.S., S.P., and K.T. performed sample analyses. J.E.A., K.G.H., D.R.H., M.R.E., L.E.S., S.P., and K.T. analyzed sequence data. M.H.P., S.P., and K.T. generated the barcoded LV vector. M.R.E. generated custom analysis software. J.E.A., M.R.E., R.S., and L.E.S. generated the figures. J.E.A., M.H.P., and H.-P.K. funded the study. J.E.A. and M.R.E. wrote the manuscript. All authors reviewed and edited the final manuscript.

## CONFLICTS OF INTEREST

The authors declare no competing interests.

## ACKNOWLEDGMENTS

This work was supported in part by grants to H.-P.K. from the National Institutes of Health (NIH), including R01 HL116217, R01 HL115128, and U19 HL129902, and by funds to J.E.A. from the Fred Hutchinson Cancer Research Center, the Cuyamaca Foundation, the Hartwell Foundation, and the Bill & Melinda Gates Foundation. M.H.P. was supported by, and thanks, the Laurie Kraus Lacob Faculty Scholar in Pediatric Translational Research and the Amon Carter Foundation for this work. This research was also funded in part through a pilot study award to R.S. from the NIDDK Cooperative Center of Excellence in Hematology Grant U54 DK106829, and the NIH/NCI Cancer Center Support Grant P30 CA015704. J.E.A. receives support as the Fleishcauer Family Endowed Chair in Gene Therapy Translation. H.-P.K. is a Markey Molecular Medicine Investigator and received support as the inaugural recipient of the José Carreras/E. Donnall Thomas Endowed Chair for Cancer Research. We acknowledge the Vector Production Core, Flow Cytometry and Genomics Shared Resources of the Fred Hutchinson Cancer Research Center (FHCRC), Veronica Nelson, Erica Curry, and Kelvin Sze for animal care, and Helen Crawford for manuscript formatting.

## REFERENCES

1. Gennery, A.R., Khawaja, K., Veys, P., Bredius, R.G., Notarangelo, L.D., Mazzolari, E., Fischer, A., Landais, P., Cavazzana-Calvo, M., Friedrich, W., et al. (2004). Treatment of CD40 ligand deficiency by hematopoietic stem cell transplantation: a survey of the European experience, 1993–2002. Blood *103*, 1152–1157.

2. Braun, C.J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., Modlich, U., Beier, R., Göhring, G., Steinemann, D., et al. (2014). Gene therapy for Wiskott-Aldrich syndrome—long-term efficacy and genotoxicity. Sci. Transl. Med. *6*, 227ra33.

3. Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.L., Fraser, C.C., Cavazzana-Calvo, M., and Fischer, A. (2003). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N. Engl. J. Med. *348*, 255–256.

4. Biasco, L., Baricordi, C., and Aiuti, A. (2012). Retroviral integrations in gene therapy trials. Mol. Ther. *20*, 709–716.

5. Bystrykh, L.V., Verovskaya, E., Zwart, E., Broekhuis, M., and de Haan, G. (2012). Counting stem cells: methodological constraints. Nat. Methods *9*, 567–574.

6. Schmidt, M., Glimm, H., Lemke, N., Muessig, A., Speckmann, C., Haas, S., Zickler, P., Hoffmann, G., and Von Kalle, C. (2001). A model for the detection of clonality in marked hematopoietic stem cells. Ann. N Y Acad. Sci. *938*, 146–155.

7. Tacke, M., Ball, C.R., Schmidt, M., Klingenberg, S., Maurer, B., Fessler, S., Eaves, C.J., von Kalle, C., and Glimm, H. (2010). The inherent differentiation program of short-term hematopoietic repopulating cells changes during human ontogeny. Stem Cells Dev. *19*, 621–628.

8. Drize, N., Chertkov, J., Sadovnikova, E., Tiessen, S., and Zander, A. (1997). Long-term maintenance of hematopoiesis in irradiated mice by retrovirally transduced peripheral blood stem cells. Blood *89*, 1811–1817.

9. Biasco, L., Pellin, D., Scala, S., Dionisio, F., Basso-Ricci, L., Leonardelli, L., Scaramuzza, S., Baricordi, C., Ferrua, F., Cicalese, M.P., et al. (2016). In vivo tracking of human hematopoiesis reveals patterns of clonal dynamics during early and steady-state reconstitution phases. Cell Stem Cell *19*, 107–119.

10. Kim, S., Kim, N., Presson, A.P., Metzger, M.E., Bonifacino, A.C., Sehl, M., Chow, S.A., Crooks, G.M., Dunbar, C.E., An, D.S., et al. (2014). Dynamics of HSPC repopulation in nonhuman primates revealed by a decade-long clonal-tracking study. Cell Stem Cell *14*, 473–485.

11. Wu, C., Li, B., Lu, R., Koelle, S.J., Yang, Y., Jares, A., Krouse, A.E., Metzger, M., Liang, F., Loré, K., et al. (2014). Clonal tracking of rhesus macaque hematopoiesis highlights a distinct lineage origin for natural killer cells. Cell Stem Cell *14*, 486–499.

12. Shepherd, B.E., Kiem, H.P., Lansdorp, P.M., Dunbar, C.E., Aubert, G., LaRochelle, A., Seggewiss, R., Guttorp, P., and Abkowitz, J.L. (2007). Hematopoietic stem-cell behavior in nonhuman primates. Blood *110*, 1806–1813.

13. Schmidt, M., Hacein-Bey-Abina, S., Wissler, M., Carlier, F., Lim, A., Prinz, C., Glimm, H., Andre-Schmutz, I., Hue, C., Garrigue, A., et al. (2005). Clonal evidence for the transduction of $CD34^+$ cells with lymphomyeloid differentiation potential and self-renewal capacity in the SCID-X1 gene therapy trial. Blood *105*, 2699–2706.

14. Jordan, C.T., and Lemischka, I.R. (1990). Clonal and systemic analysis of long-term hematopoiesis in the mouse. Genes Dev. *4*, 220–232.

15. Snodgrass, R., and Keller, G. (1987). Clonal fluctuation within the haematopoietic system of mice reconstituted with retrovirus-infected stem cells. EMBO J. *6*, 3955–3960.

16. Lu, R., Neff, N.F., Quake, S.R., and Weissman, I.L. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat. Biotechnol. *29*, 928–933.

17. Beard, B.C., Trobridge, G.D., Ironside, C., McCune, J.S., Adair, J.E., and Kiem, H.P. (2010). Efficient and stable MGMT-mediated selection of long-term repopulating stem cells in nonhuman primates. J. Clin. Invest. *120*, 2345–2354.

18. Chandrasekaran, D., Nakamoto, B., Watts, K.L., Kiem, H.P., and Papayannopoulou, T. (2014). Modeling promising nonmyeloablative conditioning regimens in nonhuman primates. Hum. Gene Ther. *25*, 1013–1022.

19. Kiem, H.P., Sellers, S., Thomasson, B., Morris, J.C., Tisdale, J.F., Horn, P.A., Hematti, P., Adler, R., Kuramoto, K., Calmels, B., et al. (2004). Long-term clinical and molecular follow-up of large animals receiving retrovirally transduced stem and progenitor cells: no progression to clonal hematopoiesis or leukemia. Mol. Ther. *9*, 389–395.

20. Peterson, C.W., Wang, J., Norman, K.K., Norgaard, Z.K., Humbert, O., Tse, C.K., Yan, J.J., Trimble, R.G., Shivak, D.A., Rebar, E.J., et al. (2016). Long-term multilineage engraftment of autologous genome-edited hematopoietic stem cells in nonhuman primates. Blood *127*, 2416–2426.

21. Trobridge, G.D., Beard, B.C., Gooch, C., Wohlfahrt, M., Olsen, P., Fletcher, J., Malik, P., and Kiem, H.P. (2008). Efficient transduction of pigtailed macaque hematopoietic repopulating cells with HIV-based lentiviral vectors. Blood *111*, 5537–5543.

22. Younan, P.M., Polacino, P., Kowalski, J.P., Peterson, C.W., Maurice, N.J., Williams, N.P., Ho, O., Trobridge, G.D., Von Laer, D., Prlic, M., et al. (2013). Positive selection of mC46-expressing $CD4^+$ T cells and maintenance of virus specific immunity in a primate AIDS model. Blood *122*, 179–187.

23. Beard, B.C., Adair, J.E., Trobridge, G.D., and Kiem, H.P. (2014). High-throughput genomic mapping of vector integration sites in gene therapy studies. Methods Mol. Biol. *1185*, 321–344.

24. Kiem, H.P., Baum, C., Bushman, F.D., Byrne, B.J., Carter, B.J., Cavagnaro, J., Malech, H.L., Mendell, J.R., Naldini, L.M., Sorrentino, B.P., et al. (2014). Charting a clear path: the ASGCT Standardized Pathways Conference. Mol. Ther. *22*, 1235–1238.

25. Berry, C.C., Gillet, N.A., Melamed, A., Gormley, N., Bangham, C.R., and Bushman, F.D. (2012). Estimating abundances of retroviral insertion sites from DNA fragment length data. Bioinformatics *28*, 755–762.

26. Porter, S.N., Baker, L.C., Mittelman, D., and Porteus, M.H. (2014). Lentiviral and targeted cellular barcoding reveals ongoing clonal dynamics of cell lines in vitro and in vivo. Genome Biol. *15*, R75.

27. Davis, B.M., Roth, J.C., Liu, L., Xu-Welliver, M., Pegg, A.E., and Gerson, S.L. (1999). Characterization of the P140K, PVP(138-140)MLK, and G156A $O^6$-methylguanine-DNA methyltransferase mutants: implications for drug resistance gene therapy. Hum. Gene Ther. *10*, 2769–2778.

28. Radtke, S., Adair, J.E., Giese, M.A., Chan, Y.Y., Norgaard, Z.K., Enstrom, M., Haworth, K.G., Schefter, L.E., and Kiem, H.P. (2017). A distinct hematopoietic stem cell population for rapid multilineage engraftment in nonhuman primates. Sci. Transl. Med. *9*, eaan1145.

29. Koelle, S.J., Espinoza, D.A., Wu, C., Xu, J., Lu, R., Li, B., Donahue, R.E., and Dunbar, C.E. (2017). Quantitative stability of hematopoietic stem and progenitor cell clonal output in rhesus macaques receiving transplants. Blood *129*, 1448–1457.

30. Gerrits, A., Dykstra, B., Kalmykowa, O.J., Klauke, K., Verovskaya, E., Broekhuis, M.J., de Haan, G., and Bystrykh, L.V. (2010). Cellular barcoding tool for clonal analysis in the hematopoietic system. Blood *115*, 2610–2618.

31. Cheung, A.M., Nguyen, L.V., Carles, A., Beer, P., Miller, P.H., Knapp, D.J., Dhillon, K., Hirst, M., and Eaves, C.J. (2013). Analysis of the clonal growth and differentiation dynamics of primitive barcoded human cord blood cells in NSG mice. Blood *122*, 3129–3137.

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

33. Verovskaya, E., Broekhuis, M.J., Zwart, E., Weersing, E., Ritsema, M., Bosman, L.J., van Poele, T., de Haan, G., and Bystrykh, L.V. (2014). Asymmetry in skeletal distribution of mouse hematopoietic stem cell clones and their equilibration by mobilizing cytokines. J. Exp. Med. *211*, 487–497.

34. Schmidt, M., Zickler, P., Hoffmann, G., Haas, S., Wissler, M., Muessig, A., Tisdale, J.F., Kuramoto, K., Andrews, R.G., Wu, T., et al. (2002). Polyclonal long-term repopulating stem cell clones in a primate model. Blood *100*, 2737–2743.

35. Wang, J.C., Doedens, M., and Dick, J.E. (1997). Primitive human hematopoietic cells are enriched in cord blood compared with adult bone marrow or mobilized peripheral blood as measured by the quantitative in vivo SCID-repopulating cell assay. Blood *89*, 3919–3924.

36. Scala, S., Basso-Ricci, L., Dionisio, F., Pellin, D., Giannelli, S., Salerio, F.A., Leonardelli, L., Cicalese, M.P., Ferrua, F., Aiuti, A., and Biasco, L. (2018). Dynamics of genetically engineered hematopoietic stem and progenitor cells after autologous transplantation in humans. Nat. Med. *24*, 1683–1690.

37. Jung, C.W., Beard, B.C., Morris, J.C., Neff, T., Beebe, K., Storer, B.E., and Kiem, H.P. (2007). Hematopoietic stem cell engraftment: a direct comparison between intramarrow and intravenous injection in nonhuman primates. Exp. Hematol. *35*, 1132–1139.

38. Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614–620.

39. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res. *12*, 996–1006.

40. Kent, W.J. (2002). BLAT—the BLAST-like alignment tool. Genome Res. *12*, 656–664.

41. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

42. Sievers, F., and Higgins, D.G. (2018). Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. *27*, 135–145.

43. Hocum, J.D., Battrell, L.R., Maynard, R., Adair, J.E., Beard, B.C., Rawlings, D.J., Kiem, H.P., Miller, D.G., and Trobridge, G.D. (2015). VISA—Vector Integration Site Analysis server: a web-based server to rapidly identify retroviral integration sites from next-generation sequencing. BMC Bioinformatics *16*, 212.

44. Thielecke, L., Aranyossy, T., Dahl, A., Tiwari, R., Roeder, I., Geiger, H., Fehse, B., Glauche, I., and Cornils, K. (2017). Limitations and challenges of genetic barcode quantification. Sci. Rep. *7*, 43249.

45. Wickham, H. (2016). ggplot2. Elegant Graphics for Data Analysis (Springer International Publishing).