HHS Public Access

Author manuscript

Gene Ther. Author manuscript; available in PMC 2010 June 01.

Published in final edited form as:

Gene Ther. 2009 December ; 16(12): 1452–1464. doi:10.1038/gt.2009.96.

Retroviral vector integration in post-transplant hematopoiesis in mice conditioned with either submyeloablative or ablative irradiation

Mohammed A. Sadat[1,*], **Sara Dirscherl**[1,*], **Lakshmi Sastry**[2,*], **Jessica Dantzer**[2,3], **Nancy Pech**[1], **Samantha Griffin**[2], **Troy Hawkins**[2], **Yiqiang Zhao**[2,3], **Cecilia N. Barese**[1], **Scott Cross**[2], **Attilio Orazi**[4], **Caroline An**[4], **W. Scott Goebel**[1], **Mervin C. Yoder**[1,5], **Xiaoman Li**[3,8,9], **Manuel Grez**[6], **Kenneth Cornetta**[2,7,8], **Sean D. Mooney**[2,3], and **Mary C. Dinauer**[1,2,7]

[1]Herman B Wells Center for Pediatric Research, Department of Pediatrics, Riley Hospital for Children, Indiana University School of Medicine, Indianapolis, IN USA

[2]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN USA

[3]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN USA

[4]Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN USA

[5]Department of Biochemistry & Molecular Biology, Indiana University School of Medicine, Indianapolis, IN USA

[6]Molecular Virology, Georg-Speyer-Haus, Frankfurt, Germany

[7]Department of Microbiology/Immunology, Indiana University School of Medicine, Indianapolis, IN USA

[8]Department of Medicine; Indiana University School of Medicine, Indianapolis, IN 46202 USA

Abstract

X-linked chronic granulomatous disease (X-CGD) is an inherited immunodeficiency with absent phagocyte NADPH oxidase activity caused by defects in the gene encoding gp91^{phox}. Here we evaluated strategies for less intensive conditioning for gene therapy of genetic blood disorders without selective advantage for gene correction, such as might be used in a human X-CGD protocol. We compared submyeloablative with ablative irradiation as conditioning in murine X-CGD, examining engraftment, oxidase activity and vector integration in mice transplanted with

Correspondence should be addressed to M.C.D., Mary C. Dinauer, MD, PhD, Cancer Research Institute R4, Indiana University School of Medicine, 1044 West Walnut Street, Room 402A, Indianapolis, IN 46202 USA, Tel: 317-274-8645, Fax: 317-274-8679, mdinauer@iupui.edu.

*M.A.S., S.D. AND L.S. contributed equally to this work.

[9]current address: Burnett School of Biomedical Science, University of Central Florida, Orlando, FL.

The online version of the article contains supplemental material.

marrow transduced with a gamma-retroviral vector for gp91^{phox} expression. The frequency of oxidase-positive neutrophils in the donor population was unexpectedly higher in many 300 cGy-conditioned mice compared to lethally irradiated recipients, as was the fraction of vector-marked donor secondary CFU-S12. Vector integration sites in marrow, spleen, and secondary CFU-S12 DNA from primary recipients were enriched for cancer-associated genes, including *Evi1*, and integrations in or near cancer-associated genes were more frequent in marrow and secondary CFU-S12 from 300 cGy-conditioned mice compared to fully ablated mice. These findings support the concept that vector integration can confer a selection bias, and suggest that intensity of the conditioning regimen may further influence the effects of vector integration on clonal selection in post-transplant engraftment and hematopoiesis.

Introduction

Inherited diseases of the hematopoietic system can be treated with autologous hematopoietic stem cells (HSCs) genetically modified with gamma-retroviral vectors. Encouraging results have accrued in recent clinical trials of gene therapy for X-linked Severe Combined Immune Deficiency (SCID),[1,2] adenosine-deaminase-deficient SCID,[3] and chronic granulomatous disease, an inherited defect in microbial killing.[4] However, it is now recognized that vector integration can disrupt cellular gene function, including activating effects mediated by potent viral enhancers in gamma-retroviral long-terminal repeats (LTRs). Moreover, five patients in X-linked SCID trials developed T cell leukemia in association with vector insertion near the LMO2-proto-oncogene in four of the cases.[5–8] Although leukemogenesis in X-SCID is likely influenced by additional disease-specific factors,[9,10] data emerging from clinical trials as well as murine and non-human primate transplant models suggest that integrations near proto-oncogenes can lead to nonmalignant clonal expansion due to enhanced "cell fitness".[4,11–16] With acquisition of additional mutations over time, these integrations may be a risk factor for malignant progression.[7,8,17] Hence, there is intense interest in further characterizing the impact of vector integration on hematopoiesis following gene transfer.[9,18]

Chronic granulomatous disease (CGD) is an inherited disorder of innate immunity in which phagocytic leukocytes are unable to generate microbicidal oxidants due to mutations in any one of four genes that encode essential subunits of the phagocyte NADPH oxidase.[19] Approximately two-thirds of CGD results from defects in the X-linked gene encoding gp91^{phox}, a subunit of flavocytochrome b_{558} and the redox center of the oxidase. CGD is a candidate disease for treatment using genetically modified HSCs,[19] and in murine models of X-linked and p47^{phox}-deficient CGD, gamma-retroviral-mediated gene transfer into HSC can correct neutrophil NADPH oxidase activity and improve the defects in host defense.[19–21] Since there is no intrinsic selection for gene-corrected cells in CGD, the development of regimens that permit sufficient levels of engraftment yet decrease the intensity of pre-transplant conditioning is also under investigation. Submyeloablative doses of irradiation or chemotherapy are effective at enabling long-term engraftment of transduced HSC in murine CGD, even though higher donor cell doses are needed and engraftment occurs in a more competitive host environment compared to fully ablated marrow.[22–25]

A recent clinical trial in two adult X-CGD patients incorporated partially ablative busulfan conditioning prior to the infusion of autologous CD34+ cells transduced with the SF71gp91^{phox} gamma-retroviral vector for expression of human gp91^{phox}.4 A bicistronic version, which is derived from the Spleen Focus Forming Virus (SFFV), showed high and relatively sustained expression of gp91^{phox} following transplantation of transduced cells in ablated X-CGD mice.26 In the clinical trial, high initial levels of oxidase-positive peripheral blood neutrophils (12–31%) were seen in the first five months, along with resolution of chronic infections. This was followed by an unexpected further increase in the fraction of gene-corrected neutrophils, although not associated with an increase in total neutrophil numbers. Analysis of vector integration sites revealed sequential expansion of myeloid clones harboring activating insertions in three growth-promoting genes, *MDS1/EVI1*, *PRDM16*, and *SETBP1*. Both patients subsequently developed myelodysplasia with Monosomy 7. (M. Grez, personal communication)

In the current study, we sought to compare submyeloablative with ablative irradiation as conditioning prior to transplantation of SF71gp91^{phox}-transduced marrow in murine X-CGD. We examined long-term engraftment, correction of peripheral blood neutrophil NADPH oxidase activity , and, because of an unexpectedly high frequency of oxidase-positive neutrophils in some recipients, vector integration sites. We also conducted serial transplants and monitored for clonal hematopoiesis and the development of leukemia. We report for the first time that the intensity of host conditioning impacts on subsequent behavior of hematopoietic clones following gammaretroviral transduction, transplantation, and engraftment, which may be related to a competitive advantage associated with vector integration.

Materials and Methods

Retroviral transduction and sex-mismatched transplantation of X-CGD murine BM

Two independent experiments were performed using 6 – 10 week old X-CGD male donors and 8 – 10-week-old X-CGD female recipients. Bone marrow harvested from mice treated three days previously with 5-fluorouracil was transduced with ecotropic envelope-packaged SF71gp91^{phox} 4 (Fig. 1A), and injected intravenously into paired cohorts of X-CGD mice as shown in Fig. 1B (for further details, see previous studies21,26,27 and Supplemental Materials and Methods). One cohort was conditioned with 300 cGy and transplanted with 8 × 10^6 or 5 × 10^6 cells per recipient in the first and second experiments, respectively. The cohort was transplanted with 2 × 10^6 cells per recipient following lethal irradiation using a split dose of 1100 cGy in the first experiment or single dose of 950 cGy in the second experiment (Fig. 1B). Neutrophil NADPH oxidase activity, peripheral blood (PB) counts and differential, and donor chimerism using FISH for the Y chromosome were monitored post-transplant as described.23 Recipients were followed for ≥ 6 months, except for 4 primary recipients (2 in each conditioning group) sacrificed at 5 months in the first experiment. For serial transplants, BM was injected intravenously (2 × 10^6 cells per mouse) into 3 – 5 X-CGD female recipients irradiated with 1100 cGy given as a split dose. In the first experiment, BM from individual mice was transplanted into multiple secondary recipients, whereas in the second experiment, marrow was pooled from 4 to 5 mice in the

same cohort for secondary transplantation. Most tertiary transplants used pooled BM. Necroscopy was performed on all mice, with gross examination of visceral organs, spleen weight, and histology of spleen and BM (liver and thymus, if indicated). Ill mice and/or mice with an abnormal PB counts had marrow, spleen, and thymus immunostained for CD3, B220, MPO, lysozyme and CAF. Animal experiments were approved by the IUSM Institutional Animal Care and Use Committee.

Southern blot analysis

Southern blot analysis of genomic DNA from spleen, secondary CFU-S12, and non-adherent nucleated BM cells, which are 70 – 80% neutrophils, was performed similar to as described.21,27 DNA was digested with Bgl II to analyze provirus marking or with BamHI to generate junctional fragments (Fig. 1A). Southern blots were probed with a random-primed cDNA from human gp91^{phox} 21 or a 460 bp PCR-amplified product of *SRY*, a Y chromosome gene. Genomic DNA from C57Bl/6J male mice served as template for primers Sry-8353 (5'-CGTGGTGAGAGGCACAAGTTG-3') and Sry-8812 (5'-GTGGTGGTGGTGGTCATA-3').

Ligation-mediated (LM)-PCR and insertion site analysis

Approximately 250 ng genomic DNA was used as a template for LM-PCR, performed using Tsp5091-digested DNA as described.11,28,29 For further details, see Supplemental Materials and Methods. Retroviral-genomic junction sequences were submitted to SeqMap, a web-based tool for mapping retroviral integration sites (RIS).30 RIS were initially identified by mapping to the UCSC (mm6 assembly) and Ensembl (version38.35) genome databases. RefSeq transcripts within 300 kb of the BLAT31 hit were identified and ranked by distance to the transcription start site using the Mouse Genome Informatics (MGI) identifier for each transcript. Nearby gene assignments were then checked by hand to ensure correct annotation.

Annotations for cancer-associated genes (CAG) were determined using the Retroviral Tagged Cancer Genome Database (RTCGD)32 for common insertion sites (CIS) and the Sanger Institute Human Cancer Gene Census Database.33 Significance for CAG enrichment was calculated using the hypergeometric distribution as performed with the phyper function in the R statistical analysis platform (http://www.gnome.org/projects/gnumeric/doc/gnumeric-R.PHYPER.shtml).

Sequences in a transcriptome stem cell database (SCDb) 34 were mapped to mouse Refseq using BLAST, with an E-value threshold set to E-50, and compared to recovered RIS using the hypergeometric distribution, as above.

Functional analysis of genes located near RIS was performed by identifying over-represented Gene Ontology (GO)35 terms in a data set relative to the annotation set of the entire genome by the MGI Mouse-Genome-Database (revision 1.40). Significantly over-represented terms were found by applying the hypergeometric distribution to all terms annotated to genes in the group. The probability of a GO term *X* being annotated to a gene in the group is found by:

$$f(k;N,m,n)=\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, \quad \text{(Eq. 1)}$$

Where *k* is the number of genes in the cluster annotated with *X*, *N* is the number of annotated genes in MGI, *m* is the number of genes in MGI annotated with *X*, and *n* is the number of annotated genes in the group. GO terms were only considered if they were assigned to more than two genes in the group ($n > 2$). To calculate a p-value for over-representation of a term, we find this probability for annotation of *k* or more genes in the group:

$$p_{hg}(X)=\sum_{i=k}^{n} f(i;N,m,n). \quad \text{(Eq. 2)}$$

Significant differences in GO term representation between two sets of genes were analyzed using an unpaired, one-tailed Student's T test.

Additional Statistical Analysis

The Student's T test and the Mann-Whitney non-parametric test were performed using Instat 3 for the MacIntosh (GraphPad Software, La Jolla, CA).

Website URLs

USCS-Genome-Browser, http://genome.ucsc.edu. ENSEMBL-database, http://www.ensembl.org.Mouse Retrovirus Tagged Cancer Gene Database, http://rtcgd.ncifcrf.gov (accessed September 2007). Sanger Institute Human Cancer Gene Census Database, http://www.sanger.ac.uk/genetics/CGP/Census (accessed February 2007). National Center for Biotechnology Information. BLAST: Basic Local Alignment Search Tool, http://www.ncbi.nlm.nih.gov/BLAST. Gene Ontology Consortium, http://www.geneontology.org (accessed December 2008). Mouse Genome Informatics Database, http://www.informatics.jax.org.

Results

Donor chimerism and expression of gp91^{phox} following transplantation of SF71gp91^{phox}-transduced marrow into 300 cGy and 950–1100 cGy-conditioned X-CGD mice

SF71gp91^{phox}-transduced BM from male X-CGD mice was infused into paired cohorts of female X-CGD mice receiving either 300 cGy or 950–1100 cGy (Fig. 1), and as is typical in submyeloablative regimens, 2.5 – 4-fold more cells were transplanted in the 300 cGy group in order to achieve donor chimerism of ≈50% in the recipients.23 Two independent experiments were performed, with primary recipients generally followed for at least 6 months, and up to 14 months, prior to sacrifice and additional studies. A total of 24 primary, 40 secondary, and 46 tertiary recipients were studied (Table 1).

Chimerism for male donor leukocytes in primary recipients conditioned with 300 cGy was similar in Experiments I and II, and was 49 ± 21% and 47.5 ± 11% two months after

transplant and 63 ± 13% and 58 ± 8 % by 4–5 months post transplant, respectively (N = 4 –8 at each time point). These levels are in agreement with results in 300 cGy-conditioned mice transplanted with a similar number of BM cells transduced with an MSCV-based retroviral vector.23 In contrast, high dose irradiation (≥ 800 cGy) in mouse transplant models for gene therapy produces donor chimerism of over 90%,36 which we confirmed in a previous study. 23

Expression of human gp91^{phox} and murine p22^{phox} in X-CGD neutrophils following SF71gp91^{phox} gene transfer was approximately half of that found in human neutrophils (Fig. S1A). Superoxide production, when corrected in proportion to the fraction of oxidase-positive cells, was comparable to wild type murine neutrophils (Fig. S1B). This level of reconstitution is similar to that reported using a bicistronic SFFV vector expressing gp91^{phox} and NGFR.26

As X-CGD neutrophils otherwise lack NADPH oxidase activity, reconstitution of superoxide production was used to monitor for expression of vector-derived gp91^{phox} (Fig. 2). While donor chimerism in the recipients was similar (see above), there were different levels of oxidase-positive neutrophils between the two experiments, likely representing subtle differences in the transduction or transplant; we have observed similar variability previously.21–24,26,27 Fluctuation in the level of vector-expressing cells over time was also seen in some animals, particularly in Experiment I. Again, this has been observed in previous studies.21–24,26,27 and likely represents cycling of different hematopoietic clones.

In the 950–1100 cGy cohort, the fraction of oxidase-positive neutrophils was 41 ± 11% and 18 ± 5% in mice studied at 5 – 6 months in the first and second experiments, respectively. For the 300 cGy cohort, 60 ± 14% and 14 ± 9 % of neutrophils were oxidase positive at 5–6 months. As neutrophils in the 950–1100 cGy cohort are expected to be almost all donor-derived,23,36 the percentage of NADPH oxidase-positive neutrophils in this group directly reflects the frequency of vector-transduced donor cells expressing gp91^{phox}. In the 300 cGy cohort, even though only ≈ 60% of neutrophils were of donor origin, the percentage of oxidase-positive neutrophils was often similar to or higher than the lethally irradiated cohort (Fig. 2). This indicates that the frequency of vector-expressing neutrophils within the donor cell population in these instances was higher than for animals receiving lethal irradiation. These unexpected results suggested a potential selective advantage for repopulation with gene-modified cells in the setting of submyeloablative irradiation.

Vector marking of CFU-S12 derived from long-term primary recipients

Since reconstitution of NADPH oxidase activity reflects only transcriptionally active vector, we evaluated vector marking by Southern blot. BM obtained from primary recipients at 8 – 11 months post-transplant was infused into lethally irradiated recipients to generate secondary CFU-S12. Individual CFU-S12 represent a clonal population of multipotential myeloid cells that are derived from long term repopulating cells by 6 months post-transplant, 37 and vector marking of secondary CFU-S12 is a well-established method to assess gene transfer efficiency in primitive hematopoietic progenitors.38 In most animals, the percentage of donor-derived neutrophils that were oxidase-positive was considerably less than CFU-S12 vector marking (Fig. 3a), consistent with silencing of some integrants. This result also

suggests that the decline in oxidase-positive neutrophils observed in some mice at late time points (Fig. 2) was due to vector silencing rather than a decrease in vector-marked hematopoietic cells.

Interestingly, a consistently higher fraction of vector-marked donor CFU-S12 were recovered from 300 cGy-conditioned animals compared with the lethally irradiated cohort (Fig. 3a). In aggregate, 67 ± 12% of male secondary CFU-S12 present in the marrow of 300 cGy-conditioned recipients were vector-positive, compared to only 36 ± 20% of CFU-S12 from lethally irradiated mice (Fig. 3b). This almost two-fold difference was statistically significant ($p < 0.02$; unpaired T test), suggesting that engraftment of SF71gp91^{phox}-positive long-term repopulating cells was favored in the submyeloablated cohort. The number of vector integrants in individual secondary CFU-S12 was determined by junctional fragment analysis (Fig. 3c). For mice conditioned with lethal irradiation or with 300 cGy, the mean number of integrants per CFU-S12 was 1.29 ± 0.27 and 1.82 ± 1.05, respectively ($p = 0.18$; Mann-Whitney nonparametric test). Thus, the increased frequency of vector-positive CFU-S12 in the 300 cGy cohort does not appear to be secondary to differences in the number of integrants per CFU-S12.

Serial transplantation

For both sublethally and lethally irradiated primary recipients, we performed secondary and tertiary transplants into lethally irradiated recipients. The level of donor chimerism was monitored in recipients of marrow derived from 300 cGy-conditioned mice, and was generally maintained, with 68.2 ± 18.9% (N =18), and 55.7 ± 29.9% (N =26) male leukocytes detected in secondary and tertiary recipients, respectively. The percentages of NADPH oxidase-positive neutrophils (data not shown) were also generally similar to the primary recipient donors. We saw no evidence of a systematic decline in the frequency of NADPH-oxidase-positive neutrophils with serial transplantation, as previously observed in studies using other gamma-retroviral vectors for expression of gp91^{phox}. 21,26,27

Mice were monitored for hematologic malignancies and no abnormalities found in primary and secondary transplant recipients (data not shown). However, 4 of 46 tertiary transplant recipients developed T cell leukemia/lymphoma associated with splenic and/or thymic enlargement (Table S1). One mouse expired before adequate samples could be obtained, and two animals had malignant cells that appeared to be of host origin (Fig. S2 and data not shown). However, leukemic spleen DNA was vector positive in a tertiary recipient transplanted with cells derived from mouse A4 (300 cGy group in the first experiment), (Fig. S2). Insertion site analysis from this animal is described below.

The fate of animal A16 (300 cGy cohort, Experiment II) was of interest given the increasing percentage of NADPH oxidase-positive neutrophils in the primary recipient (Fig. 2). However, there was no histologic evidence of leukemic transformation in this animal, or in secondary and tertiary recipients.

Retroviral integration sites in primary recipients of SF71gp91^{phox}-transduced marrow

Retroviral integration sites (RIS) in sublethally vs lethally irradiated recipients were compared using ligation-mediated PCR (LM-PCR).29 In contrast to previous studies that

utilized only spleen as the source of DNA for RIS analysis, we also evaluated non-adherent bone marrow cells and secondary CFU-S12 as sources enriched for myeloid progenitors. Sequences of LM-PCR amplified products were analyzed to identify genes within 300 kb of the integration site (Table 2; Table S2). Identification of RIS relative to the number of bands on LM-PCR was 58 ± 14% and 78 ± 21% for the 300 cGy and lethally irradiated cohort, respectively. For the CFU-S12 samples, 65 – 72% of the junctional fragments were recovered and mapped. These compare favorably to published recovery rates.11,12 RIS clustered within +/− 5kb of the transcription start sites (TSS) and within RefSeq genes (Fig. S3; Table 2, consistent with previous studies of gamma-retrovirus integration.11,12,39–41

In both 300 cGy and 950–1100 cGy cohorts, approximately half of RIS were found only in spleen, ≈ one-third found only in marrow, and the remaining detected in both marrow and spleen (Fig. 4a). Interestingly, the mean number of insertion sites recovered from marrow and spleen of each primary recipient was ≈ 9 for both cohorts (Table 2), although donor chimerism in 300 cGy-recipients was approximately 60% and vector copy number in secondary CFU-S12 was similar in both irradiation groups. This result parallels the increased frequency of vector-marked donor CFU-S12 in the 300 cGy-conditioned cohort. We also identified a relatively greater number and diversity of insertion sites in secondary CFU-S12 from the 300 cGy cohort compared to the 950 –1100 cGy group (Table 2; Table S3).

Retroviral integration sites from primary recipients favor genes associated with cancer, especially in 300 cGy-irradiated recipients

We determined the percentage of RIS that were in or near cancer-associated genes (CAG), which represent 1.5% of genes in the murine genome (543 listed CAG genes/ 35,603 total genes), in mice receiving submyeloablative or ablative irradiation. CAG included genes listed as Common Insertion Sites (CIS) in the Retrovirus Tagged Cancer Gene (RTCGD) database32 and/or in the Human Cancer Gene Census.33 Previous studies found that gamma-retrovirus integration sites in transduced murine40 or human42 primitive hematopoietic cells were enriched for proto-oncogenes, accounting for ≈ 8% of insertion sites identified, suggesting a bias occurs even prior to transplantation. Integration near proto-oncogenes can also be associated with a selective advantage in post-transplant hematopoiesis.9,18 A recent study using gamma-retroviral vectors reported that 15% of RIS from lethally irradiated primary recipients involved CAG.12 In the current study, 11/71 (15%) of RIS from lethally irradiated recipient mice were similarly associated with CAG (Table 2, Table 3). However, in the 300 cGy cohort, 24% (23/96) of insertions involved CAG (Table 2, Table 3), suggesting an even greater bias in this group ($p = 0.061$ compared to lethally irradiated recipients; Fig. 4b).

We next performed a subset analysis. As shown in Fig. 4b and Table 2, 8/30 (27%) of RIS identified only in BM and 6/14 (45%) recovered from both BM and spleen from 300 cGy-conditioned mice involved CAG, compared to 3/21 and 2/15, respectively, in lethally irradiated mice ($p = 0.08$ and $p = 0.015$). A strong enrichment for CAG ($p \leq 0.02$) was also seen in secondary CFU-S12 from mice conditioned with submyeloablative irradiation (Fig. 4c). Only 1 of 14 insertions identified in secondary CFU-S12 from lethally irradiated mice

involved a CAG, compared to 10 of 26 (38%) in secondary CFU-S12 from the 300 cGy cohort, which included 5 that were only detected in CFU-S12 and not in BM or spleen (Table 3). Since we analyzed secondary CFU-S12 at 8 –11 months post-transplant, these results suggest that the content of long-term repopulating cells harboring vector in or near CAGs is enriched in recipients conditioned with submyeloablative irradiation compared to high dose irradiation.

Hot spots for SF71gp91^{phox} integration in murine X-CGD gene therapy model

A total of 15 insertions involved 5 genes for which similarly located but independent RIS were identified in two or more primary recipients (Table 4). All of these recurrent insertion sites or "hot spots" are also Common Insertion Sites in the murine RTCGD, and 12 of the 15 were from the 300 cGy cohort. One hot spot was *Evi1*, with four of the five independent insertions recovered from 300 cGy-conditioned mice. This included a site identified in two primary recipients from Experiment 2, one of which exhibited an increasing frequency of NADPH oxidase-positive neutrophils post-transplant (mouse A16, Fig. 2); the majority of secondary CFU-S12 from this animal were also positive for the *Evi1* insertion. Hematopoiesis in a second mouse (A4) also appeared to be dominated by an *Evi1*-positive clone that additionally harbored an insertion in *BC057627* (also known as *Zc3h4*), a poorly characterized protein with a zinc finger domain, as this clone accounted for 11/12 vector-positive secondary CFU-S12.

Overlap with Stem Cell Database and Ontology Analysis of SF71gp91^{phox} integration sites

Previous studies found that RIS recovered from hematopoietic cells following gamma-retrovirus transduction are enriched for genes expressed in primitive hematopoietic cells.12,42 We confirmed this observation, finding that 67% of RIS from primary recipients are present in an HSC transcriptome database34 compared to 45% of genes in the MGI database ($p = 2.6E-09$). There was no significant difference for this enrichment in the 300 cGy and 950–1100 cGy-irradiated recipients. RIS-associated genes were also functionally classified using GO criteria. Integration-associated genes in both the 950–1100 cGy- and 300 cGy-treated cohorts are enriched for functional terms related to transcription (Fig. S4a), as observed in several previous studies.12,42 Other annotations showed trends for over-representation in one but not the other cohort (Fig. S4b–d), including terms for phosphorylation and protein modification in RIS recovered from 300 cGy mice.

Retroviral integration sites in serially transplanted mice

The number of RIS in serial transplant recipients decreased from an average of 8 – 9 in marrow and/or spleen DNA of primary animals to 2 to 4 RIS per animal (Table S4). This suggests a decrease in hematopoietically active clones with serial transplantation, as previously reported.11 Also as previously observed,11 there was disappearance of RIS and appearance of new ones, with only 16 of 171 RIS detected in primary recipients present among the 81 RIS identified in recipients of serial transplantation. CAG-associated RIS not found in primary recipients were seen in serial transplants both from the 950–1100 cohort (2 of 21 RIS) and the 300 cGy cohort (5 of 60 RIS). RIS shared in primary and serial transplants included two *Evi1* insertions and one *Gse1* insertion. For tertiary recipients,

approximately half of RIS were also identified in the secondary recipients. We did not note a further enrichment upon serial transplantation for RIS involving genes associated with cancer, although detection of an increase in the 950–1100 cGy cohort may have been limited by the number of RIS recovered (Table 2). Analysis of GO annotations show a general trend for enrichment of genes related to transcription, similar to that found in the combined tissue data in primary transplants (Fig. S5).

In the tertiary recipient that developed vector-positive leukemia/lymphoma (Fig. S2), RIS were identified in or near *Bloc1s3*, *Madd*, *Rab38* and *Chysy1* (Table S4), none of which are cancer-associated genes. Although a clone with a pair of insertions involving *Evi1/MDS1* and *Bc057627* was prominent in hematopoiesis in the primary recipient (mouse A4, 300 cGy cohort), secondary and tertiary recipients with these insertions appeared healthy and had normal peripheral counts and spleen size.

Discussion

This study was designed to evaluate the outcome of HSC gene therapy in X-CGD mice conditioned with either 300 cGy or ablative irradiation prior to transplantation with SF71gp91^{phox}-transduced marrow. Both regimens were effective in achieving long term expression of gp91^{phox}, confirming previous studies on the efficacy of reduced intensity conditioning regimens for gene therapy of murine X-CGD. 22–25 We also observed the occurrence of higher than expected frequencies of oxidase-positive neutrophils in 300 cGy-conditioned recipients, which we had not seen in submyeloablative conditioning studies using an MSCV-based vector.22–25 Thus, we also compared vector marking and integration sites between the two conditioning regimens, finding differences that suggest that the intensity of irradiation conditioning can influence selection for vector-containing donor cells in post-transplant engraftment and hematopoiesis. This novel finding may have important implications for approaches used to achieve engraftment of genetically corrected human HSC in clinical trials.

For integration site analysis, we chose LM-PCR rather than the more sensitive linear amplification-mediated PCR (LAM-PCR) in order to recover sites that were relatively abundant in the population. There are several points regarding the approach taken in our and similar studies. First, not all RIS are equally recoverable, due to intrinsic limitations in the LM-PCR assay41 and to sequence-dependent differences in cloning efficiency and sequence amplification. Second, hematopoiesis in the murine transplant model is oligoclonal43 and active clones vary over time. For example, serial studies of secondary CFU-S12 find 3 – 9 independent clones present at any one time, with new clones emerging every few months while others disappear.44,45 Third, previous studies in the mouse have largely focused on analysis of spleen,11,12,40 which is predominantly lymphoid cells. While we evaluated spleen, we additionally analyzed non-adherent BM, which are primarily neutrophils and myeloid progenitor cells. Furthermore, our analysis of secondary CFU-S12 allowed for characterization of integrations in clonal populations of primitive multipotential myeloid cells.

We saw both similarities and important differences in SF71gp91^{phox} RIS recovered from primary recipients conditioned with either 300 cGy or 950 –1110 cGy. Similarities include 1) the relative distribution of unique sites identified in spleen, marrow, or both; 2) the percentage of insertions near the TSS and in or near RefSeq genes; 3) vector copy number in secondary CFU-S12; 4) an increased frequency of integrations in or near cancer-associated genes relative to their occurrence in the genome; and 5) a similar enrichment for genes present in an HSC transcriptome database and an over-representation of certain gene ontology categories, most notably terms associated with transcription. We observed three notable differences between the two conditioning cohorts. First, almost twice as many donor CFU-S12 were vector-positive in 300 cGy-conditioned mice. Second, there was a significantly higher frequency of RIS associated with cancer-associated genes in marrow and in multipotential myeloid cells (CFU-S12) from mice conditioned with 300 cGy compared to lethally irradiated mice. Third, the majority of insertions seen in more than one primary recipient, most notably *Evi1*, were found in the 300 cGy -conditioned cohort. Although we did not track peripheral blood insertion sites in this study, we speculate that the increased fraction of oxidase-positive neutrophils in many 300 cGy-conditioned recipients relative to the 950–1100 cGy cohort also reflects these differences.

In aggregate, our results confirm previous studies in mice, non-human primates, and clinical trials, showing that gamma-retroviral integration sites within engrafted HSCs represent a biased population.4,9,11–15,18 Additionally, our results suggest for the first time that the transplant conditioning regimen may further bias this trend. In partially ablated recipients, where the marrow is a more "competitive" environment for repopulation,25,46 insertions with potential activating effects on proliferation and/or survival may confer an advantage for successful HSC engraftment. Although vector insertion near a CAG may also promote clonal expansion of progenitor cells, it is noteworthy that the relatively greater enrichment in 300 cGy-conditioned mice compared to 950–1100 cGy is present in samples obtained at more than 6 months post-transplant, suggesting that these insertions conferred an advantage at the time of HSC engraftment.

It seems unlikely that the 2.5- to 4-fold higher cell dose used for transplantation of 300 cGy recipients can account for the increased frequency of vector marking and the relatively greater fraction of integration sites associated with cancer genes. HSC represent ≈ 0.01% of mouse BM cells47 and transplantation of 2×10^6 transduced cells corresponds to transplantation of approximately 200 HSC. While the precise number of HSC after gene transfer is hard to estimate, as 5-FU treatment will enrich for HSC in marrow progenitors while ex vivo transduction can decrease their number, the number of HSC infused in our model appears to be considerable, given the number of clones actively contributing to hematopoiesis post-transplant. Moreover, HSC do not appear to be limiting as new clones continue to emerge after serial transplantation. Most importantly, the 2.5- to 4-fold fold increase in cell dose administered to the 300 cGy group may explain the relatively greater diversity of integration sites in this cohort, but it would not be predicted to alter the rate of gene transfer into the population. Therefore, the higher cell dose would not account for the higher frequency of vector-positive donor cells and of integration sites involving cancer-associated genes.

The small degree of overlap of RIS in serial transplants with those in primary recipients suggests that reconstitution of hematopoiesis involved activation of previously dormant vector-containing HSC. Unlike prior studies,11,12 we did not detect a further enrichment upon serial transplantation for RIS involving genes associated with cancer (Table 2) or signal transduction (Fig. S3). Approximately 10% of the RIS identified in our study are present in a database of 280 RIS identified in primary and secondary recipient mice in these studies,12 which also used SFFV-based vectors. These differences between laboratories may reflect differences in vectors or experimental conditions.

Despite the occurrence of 15% or more of vector insertions in or near cancer-associated genes, the incidence of leukemia/lymphoma was low in our study, even in serial transplants, a procedure that can promote vector insertion-related leukemic progression.28. Only 4 malignant cases were identified, all in tertiary recipient mice; leukemic tissue was vector-positive in one case whereas two appeared to be host-derived. This is consistent with observations indicating that vector-induced leukemia is uncommon in the time frame of the mouse transplant model unless there are multiple insertions.28,48 It is difficult to compare the overall incidence of leukemia with other murine gene therapy studies, due to differences in retroviral vector backbones, transduction protocols, conditioning regimens and mouse strains. The incidence has varied significantly from study to study, and is further confounded by a variable incidence of vector-negative leukemia which has been reported to be as high as 6 out of 40 primary recipients in one study.49 This variability has been speculated to reflect effects of irradiation along with possible activation of endogenous retroviruses (see 49).

In the current study, 5 of 184 independent provirus insertions in primary recipients occurred in or near *Evi1* and two mice (A16 and A4) had evidence of expansion or dominance of an *Evi1*-positive clone. Insertions in *Evi1* are often over-represented in gamma-retrovirus-mediated gene transfer studies of murine hematopoietic cells either in vitro 50,51 or post-transplantation.11,12 Insertions in *EVI1* or the adjacent *MDS1* were also enriched in a non-human primate gene therapy study, accounting for 14 of 702 vector integration sites.13 Although *MDS-EVI1* activation has been implicated as a cooperating event in murine and human myeloid leukemias,52 we did not observe leukemia in either primary or serial transplant recipients harboring *Evi1* insertions.

The SF71gp91^{phox} vector in this study was also used in a recent clinical trial, where two patients exhibited myeloid expansion of clones with activating insertions involving *MDS/EVI1*, *PRDM16* and *SETBP1* that accounted for more than 50% of granulopoiesis.4 Both patients subsequently developed MDS with Monosomy 7, involving a *MDS/EVI1*-insertion positive clone.(M. Grez, personal communication) RIS recovered in our mouse transplant model using this vector were clearly more diverse, and domination of myelopoiesis with clones harboring integrations involving Evi1 was uncommon. The differences between these findings and the clinical trial could be related to many factors, including differences in the transduced cell population, the conditioning regimen and murine vs primate hematopoiesis, illustrating the challenges and limitations in predicting potential outcomes in the clinical setting using the murine model.

Finally, additional studies are required to determine if the findings noted with the SF71gp91^{phox} vector are applicable to other gamma-retroviral vectors or to self-inactivating lentiviral vectors, where insertions in or near cancer-associated genes appear to be much less frequent.9 Of note, in our previous studies using nonmyeloablative conditioning in combination with an MSCV-based vector,22–24 we did not observe an increased fraction of oxidase-corrected neutrophils in sub-myeloablated cohorts relative to recipients receiving high dose radiation, suggesting that the potent enhancer in SF71gp91^{phox}4 perhaps contributed to the biases seen in the current study.

In summary, our data confirm that certain integrations in HSCs and their progeny are over represented in long-term primary transplant recipients, and further suggest that this bias may be influenced by the marrow environment at the time of transplantation. Our data highlights the importance of modeling all aspects of a transplant approach in the murine system, as reduced intensity conditioning of a host for transplantation would have been predicted to result in less morbidity to the host, when in fact, in the context of gene replacement therapy in HSC, a potential deleterious effect was noted. The analysis of vector integration sites in the setting of partial ablation may thus be a useful approach to augment other assays aimed at examining the potential for vector-induced alterations in hematopoiesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Robert Getty for assistance with LM-PCR and Shari Upchurch, Maureena Lewis, Melody Warman, and Catherine Matthews for help with preparing the manuscript. This work was supported by National Institutes of Health Grants P01 HL53586 (MCD, MY, KC), K22 LM009135 (SM), K08 HL075253 (SG), RO1 HG004359 (XL), T32 CA111198 Cancer Biology Training Program (SD), T32 HL007910 Basic Science Studies on Gene Therapy of Blood Disease (TH). and the Riley Children's Foundation (MCD, MY, WSG).

M.A.S., S.D., L.S. and M.C.D. designed, performed, and analyzed experiments, and helped draft the paper and figures; J.D. and T. H. analyzed data and prepared figures; N.P., S.G. and S.C. performed and analyzed experiments; Y.Z. and X.L. analyzed data, C.B. helped with experimental design and developed a critical procedure; A.O. and C.A. helped with analysis of leukemic tissue; W.S.G. and M.C.Y. helped with interpretation of data and paper preparation; M.G. provided critical reagents and helped with data interpretation and preparation of the paper; K.C. and S. M. helped design, analyze and interpret the experiments and preparation of the paper; M.C.D. oversaw this entire project including experimental design, analysis, interpretation of the data, and preparation of the paper. Conflict-of-interest disclosure: The authors declare no competing financial interests.

References

1. Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, Gross F, Yvon E, Nusbaum P, et al. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science. 2000; 288:669–672. [PubMed: 10784449]
2. Gaspar HB, Parsley KL, Howe S, King D, Gilmour KC, Sinclair J, et al. Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. Lancet. 2004; 364:2181–2187. [PubMed: 15610804]
3. Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, Callegaro L, et al. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med. 2009; 360:447–458. [PubMed: 19179314]

4. Ott MG, Schmidt M, Schwarzwaelder K, Stein S, Siler U, Koehl U, et al. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat Med. 2006; 12:401–409. [PubMed: 16582916]
5. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science. 2003; 302:415–419. [PubMed: 14564000]
6. Baum C. What are the consequences of the fourth case? Mol Ther. 2007; 15:1401–1402. [PubMed: 17646833]
7. Howe SJ, Mansour MR, Schwarzwaelder K, Bartholomae C, Hubank M, Kempski H, et al. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest. 2008; 118:3143–3150. [PubMed: 18688286]
8. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest. 2008; 118:3132–3142. [PubMed: 18688285]
9. Nienhuis AW, Dunbar CE, Sorrentino BP. Genotoxicity of retroviral integration in hematopoietic cells. Mol Ther. 2006; 13:1031–1049. [PubMed: 16624621]
10. Shou Y, Ma Z, Lu T, Sorrentino BP. Unique risk factors for insertional mutagenesis in a mouse model of XSCID gene therapy. Proc Natl Acad Sci U S A. 2006; 103:11730–11735. [PubMed: 16864781]
11. Kustikova O, Fehse B, Modlich U, Yang M, Dullmann J, Kamino K, et al. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. Science. 2005; 308:1171–1174. [PubMed: 15905401]
12. Kustikova OS, Geiger H, Li Z, Brugman MH, Chambers SM, Shaw CA, et al. Retroviral vector insertion sites associated with dominant hematopoietic clones mark "stemness" pathways. Blood. 2007; 109:1897–1907. [PubMed: 17119121]
13. Calmels B, Ferguson C, Laukkanen MO, Adler R, Faulhaber M, Kim HJ, et al. Recurrent retroviral vector integration at the Mds1/Evi1 locus in nonhuman primate hematopoietic cells. Blood. 2005; 106:2530–2533. [PubMed: 15933056]
14. Aiuti A, Cassani B, Andolfi G, Mirolo M, Biasco L, Recchia A, et al. Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. J Clin Invest. 2007; 117:2233–2240. [PubMed: 17671653]
15. Deichmann A, Hacein-Bey-Abina S, Schmidt M, Garrigue A, Brugman MH, Hu J, et al. Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. J Clin Invest. 2007; 117:2225–2232. [PubMed: 17671652]
16. Schwarzwaelder K, Howe SJ, Schmidt M, Brugman MH, Deichmann A, Glimm H, et al. Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution in vivo. J Clin Invest. 2007; 117:2241–2249. [PubMed: 17671654]
17. Seggewiss R, Pittaluga S, Adler RL, Guenaga FJ, Ferguson C, Pilz IH, et al. Acute myeloid leukemia is associated with retroviral gene transfer to hematopoietic progenitor cells in a rhesus macaque. Blood. 2006; 107:3865–3867. [PubMed: 16439674]
18. Bushman FD. Retroviral integration and human gene therapy. J Clin Invest. 2007; 117:2083–2086. [PubMed: 17671645]
19. Stein S, Siler U, Ott MG, Seger R, Grez M. Gene therapy for chronic granulomatous disease. Curr Opin Mol Ther. 2006; 8:415–422. [PubMed: 17078383]
20. Mardiney M 3rd, Malech HL. Enhanced engraftment of hematopoietic progenitor cells in mice treated with granulocyte colony-stimulating factor before low-dose irradiation: implications for gene therapy. Blood. 1996; 87:4049–4056. [PubMed: 8639760]
21. Bjorgvinsdottir H, Ding C, Pech N, Gifford MA, Li LL, Dinauer MC. Retroviral-mediated gene transfer of gp91phox into bone marrow cells rescues defect in host defense against Aspergillus fumigatus in murine X-linked chronic granulomatous disease. Blood. 1997; 89:41–48. [PubMed: 8978275]

22. Goebel WS, Pech NK, Dinauer MC. Stable long-term gene correction with low-dose radiation conditioning in murine X-linked chronic granulomatous disease. Blood Cells Mol Dis. 2004; 33:365–371. [PubMed: 15528159]

23. Barese C, Pech N, Dirscherl S, Meyers JL, Sinn AL, Yoder MC, et al. Granulocyte colony-stimulating factor prior to nonmyeloablative irradiation decreases murine host hematopoietic stem cell function and increases engraftment of donor marrow cells. Stem Cells. 2007; 25:1578–1585. [PubMed: 17347493]

24. Goebel WS, Pech NK, Meyers JL, Srour EF, Yoder MC, Dinauer MC. A murine model of antimetabolite-based, submyeloablative conditioning for bone marrow transplantation: biologic insights and potential applications. Exp Hematol. 2004; 32:1255–1264. [PubMed: 15588950]

25. Goebel WS, Yoder MC, Pech NK, Dinauer MC. Donor chimerism and stem cell function in a murine congenic transplantation model after low-dose radiation conditioning: effects of a retroviral-mediated gene transfer protocol and implications for gene therapy. Exp Hematol. 2002; 30:1324–1332. [PubMed: 12423686]

26. Sadat MA, Pech N, Saulnier S, Leroy BA, Hossle JP, Grez M, et al. Long-term high-level reconstitution of NADPH oxidase activity in murine X-linked chronic granulomatous disease using a bicistronic vector expressing gp91phox and a Delta LNGFR cell surface marker. Hum Gene Ther. 2003; 14:651–666. [PubMed: 12804147]

27. Dinauer MC, Li LL, Bjorgvinsdottir H, Ding C, Pech N. Long-term correction of phagocyte NADPH oxidase activity by retroviral-mediated gene transfer in murine X-linked chronic granulomatous disease. Blood. 1999; 94:914–922. [PubMed: 10419882]

28. Modlich U, Kustikova OS, Schmidt M, Rudolph C, Meyer J, Li Z, et al. Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. Blood. 2005; 105:4235–4246. [PubMed: 15713797]

29. Schmidt M, Hoffmann G, Wissler M, Lemke N, Mussig A, Glimm H, et al. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. Hum Gene Ther. 2001; 12:743–749. [PubMed: 11339891]

30. Peters B, Dirscherl S, Dantzer J, Nowacki J, Cross S, Li X, et al. Automated analysis of viral integration sites in gene therapy research using the SeqMap web resource. Gene Ther. 2008; 15:1294–1298. [PubMed: 18580967]

31. Kent WJ. BLAT--the BLAST-like alignment tool. Genome Res. 2002; 12:656–664. [PubMed: 11932250]

32. Akagi K, Suzuki T, Stephens RM, Jenkins NA, Copeland NG. RTCGD: retroviral tagged cancer gene database. Nucleic Acids Res. 2004; 32:D523–D527. [PubMed: 14681473]

33. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, et al. A census of human cancer genes. Nat Rev Cancer. 2004; 4:177–183. [PubMed: 14993899]

34. Lemischka IMK, Stoeckert C. SCDb: Stem Cell database. 2002Accessed October 2008http://stemcell.mssm.edu/v2/.

35. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25:25–29. [PubMed: 10802651]

36. Kang E, Giri N, Wu T, Sellers S, Kirby M, Hanazono Y, et al. In vivo persistence of retrovirally transduced murine long-term repopulating cells is not limited by expression of foreign gene products in the fully or minimally myeloablated setting. Hum Gene Ther. 2001; 12:1663–1672. [PubMed: 11535169]

37. Jones RJ, Wagner JE, Celano P, Zicha MS, Sharkis SJ. Separation of pluripotent haematopoietic stem cells from spleen colony-forming cells. Nature. 1990; 347:188–189. [PubMed: 2395470]

38. Robbins PB, Skelton DC, Yu XJ, Halene S, Leonard EH, Kohn DB. Consistent, persistent expression from modified retroviral vectors in murine hematopoietic stem cells. Proc Natl Acad Sci U S A. 1998; 95:10182–10187. [PubMed: 9707621]

39. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science. 2003; 300:1749–1751. [PubMed: 12805549]

40. Montini E, Cesana D, Schmidt M, Sanvito F, Ponzoni M, Bartholomae C, et al. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat Biotechnol. 2006; 24:687–696. [PubMed: 16732270]
41. Wang GP, Garrigue A, Ciuffi A, Ronen K, Leipzig J, Berry C, et al. DNA bar coding and pyrosequencing to analyze adverse events in therapeutic gene transfer. Nucleic Acids Res. 2008; 36:e49. [PubMed: 18411205]
42. Cattoglio C, Facchini G, Sartori D, Antonelli A, Miccio A, Cassani B, et al. Hot spots of retroviral integration in human CD34+ hematopoietic cells. Blood. 2007; 110:1770–1778. [PubMed: 17507662]
43. Jordan CT, Lemischka IR. Clonal and systemic analysis of long-term hematopoiesis in the mouse. Genes Dev. 1990; 4:220–232. [PubMed: 1970972]
44. Drize NJ, Keller JR, Chertkov JL. Local clonal analysis of the hematopoietic system shows that multiple small short-living clones maintain life-long hematopoiesis in reconstituted mice. Blood. 1996; 88:2927–2938. [PubMed: 8874189]
45. Drize NJ, Olshanskaya YV, Gerasimova LP, Manakova TE, Samoylina NL, Todria TV, et al. Lifelong hematopoiesis in both reconstituted and sublethally irradiated mice is provided by multiple sequentially recruited stem cells. Exp Hematol. 2001; 29:786–794. [PubMed: 11378275]
46. Wyss BK, Meyers JL, Sinn AL, Cai S, Pollok KE, Goebel WS. A novel competitive repopulation strategy to quantitate engraftment of ex vivo manipulated murine marrow cells in submyeloablated hosts. Exp Hematol. 2008; 36:513–521. [PubMed: 18243491]
47. Weissman IL, Shizuru JA. The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. Blood. 2008; 112:3543–3553. [PubMed: 18948588]
48. Modlich U, Schambach A, Brugman MH, Wicke DC, Knoess S, Li Z, et al. Leukemia induction after a single retroviral vector insertion in Evi1 or Prdm16. Leukemia. 2008; 22:1519–1528. [PubMed: 18496560]
49. Will E, Bailey J, Schuesler T, Modlich U, Balcik B, Burzynski B, et al. Importance of murine study design for testing toxicity of retroviral vectors in support of phase I trials. Mol Ther. 2007; 15:782–791. [PubMed: 17299409]
50. Du Y, Jenkins NA, Copeland NG. Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. Blood. 2005; 106:3932–3939. [PubMed: 16109773]
51. Modlich U, Bohne J, Schmidt M, von Kalle C, Knoss S, Schambach A, et al. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. Blood. 2006; 108:2545–2553. [PubMed: 16825499]
52. Metais JY, Dunbar CE. The MDS1-EVI1 gene complex as a retrovirus integration site: impact on behavior of hematopoietic cells and implications for gene therapy. Mol Ther. 2008; 16:439–449. [PubMed: 18227842]

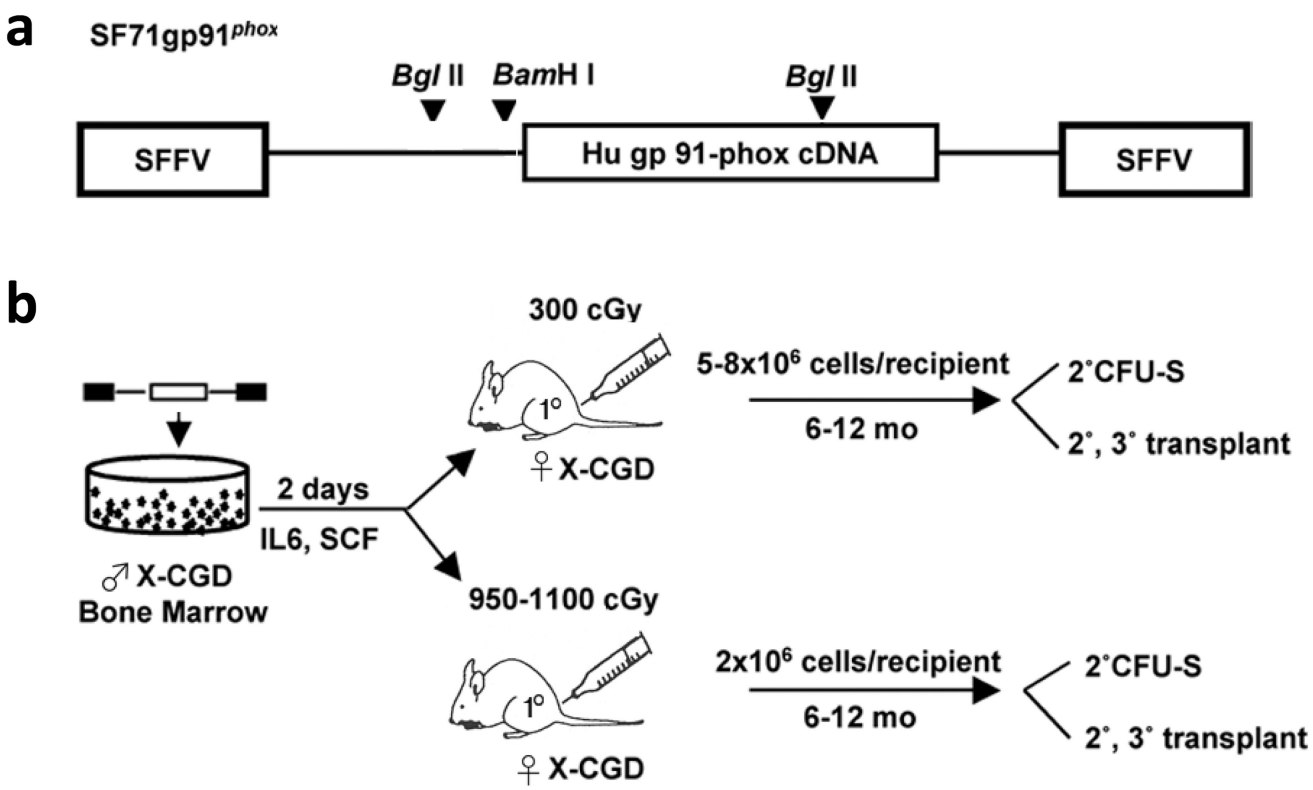

Fig. 1. Experimental design

(a) Schematic diagram of SF71gp91^{phox}, along with restriction sites used for analysis of vector integration. **(b)** Marrow was harvested from 5-FU treated male X-CGD mice and transduced with SF71gp91^{phox} prior to transplantation into female X-CGD mice conditioned with either submyeloablative (300 cGy) or lethal (1100 cGy) irradiation. See Materials and Methods for details of subsequent analyses, which included CFU-S12 and serial transplantation.

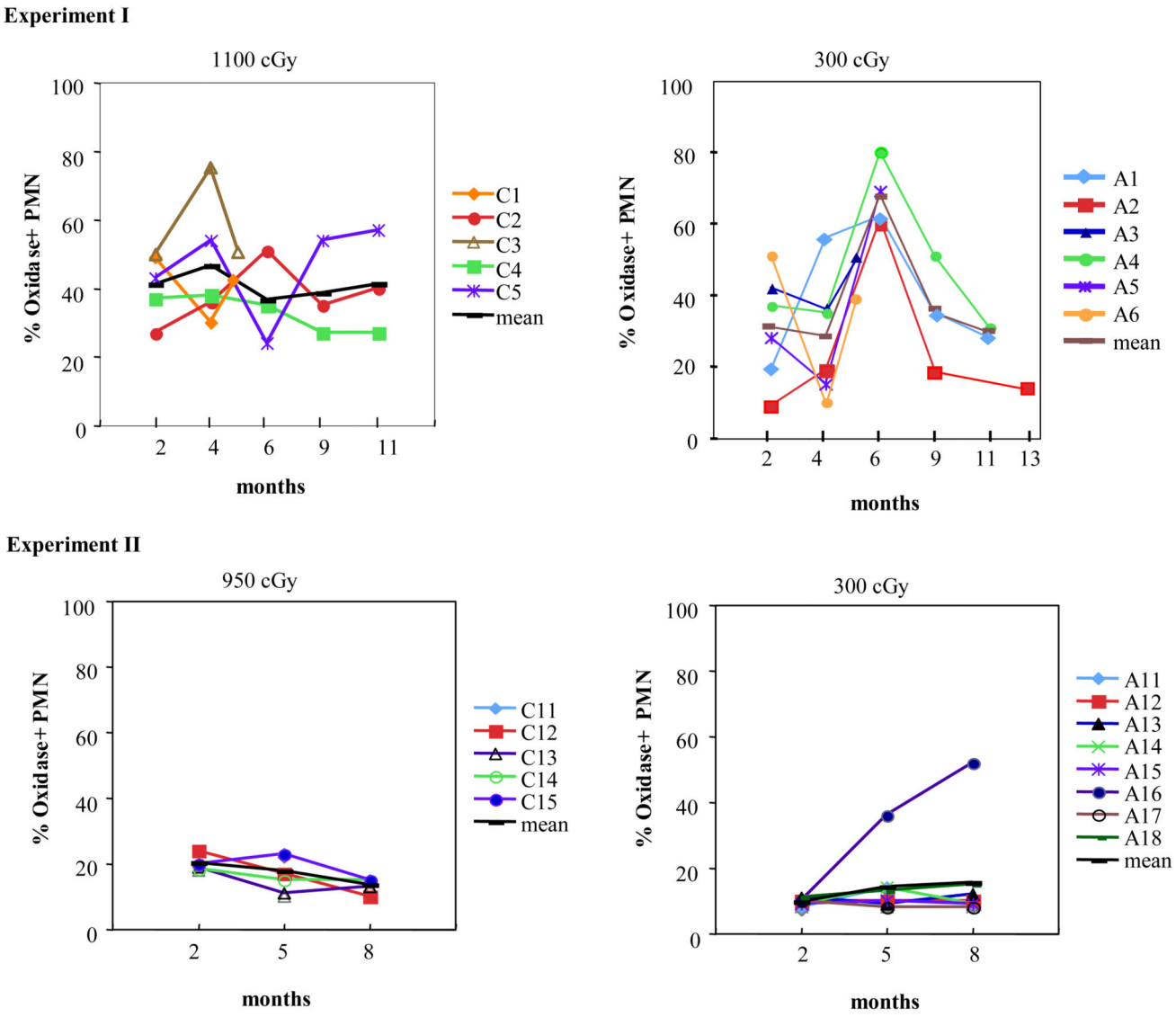

Fig. 2. NADPH oxidase activity in peripheral blood neutrophils in primary recipients following transplantation with SF71gp91^{phox}-transduced marrow

Peripheral blood was sampled at the indicated times following transplantation of SF71gp91^{phox}-transduced marrow into X-CGD mice conditioned with either 300 cGy or 950 – 1100 cGy irradiation. The percentage of neutrophils with NADPH oxidase activity was determined using the DHR 123 assay. Graphs show results for individual mice studied in two independent experiments in the upper and lower panels, respectively.

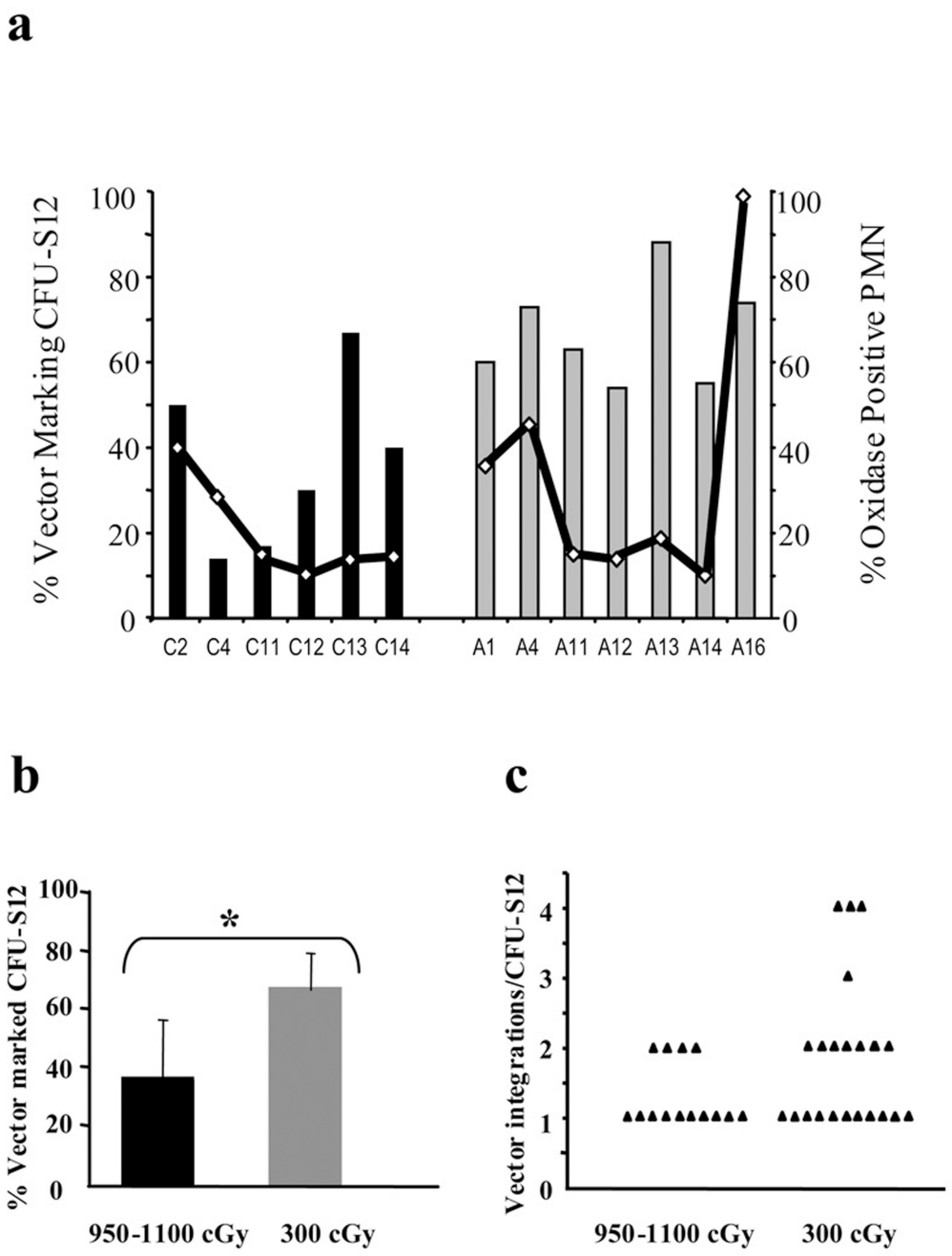

Fig. 3. Vector integration in donor-derived secondary CFU-S12 from long-term primary transplants

Primary recipients were studied at ≥ 8 months post-transplant. Combined data from Experiment I and II is shown. A4, A11, C2 and C4 mice were from the first experiment and the remainder from the second experiment. Genomic DNA was isolated from individual CFU-S12, digested with either BglII or BamHI or to identify proviral integration and junctional fragments, respectively, and analyzed by Southern blot for presence of integrated vector, and, for CFU-S12 from the 300 cGy-conditioned cohort, for Y-chromosome-specific Sry sequences. A total of 94 male CFU-S12 from seven primary recipients conditioned with

300 cGy and 69 CFU-S12 from six primary recipients conditioned with lethal irradiation were analyzed. Black bars indicate results from 950–1100 cGy cohort and gray bars indicate results for the 300 cGy cohort. **(a)** Percentage of vector-marked male CFU-S12 (filled bars) from individual primary recipients. Also shown is the percentage of male neutrophils in the peripheral blood that are NADPH oxidase-positive, an indicator of vector gp91^{phox} expression. The fraction of oxidase-positive neutrophils was determined by the DHR-123 assay and the fraction that were male was calculated based on donor chimerism. **(b)** Each bar represents the mean ± SD of the percentage of vector-marked male CFU-S12 for mice from each conditioning group, as indicated. N = 6–7 primary recipients per group. *P < 0.05. **(c)**. Vector copy number, as determined by junctional fragment number, in individual secondary CFU-S12 from 1100 cGy- and 300 cGy-conditioned recipient mice. For recipients conditioned with lethal irradiation, there were 14 independent clones out of 27 vector-positive CFU-S12 analyzed for junctional fragments, with a mean copy number per CFU-S12 of 1.29 ± 0.27. In 300 cGy-conditioned recipients, there were 22 independent clones out of 55 vector-positive CFU-S12 analyzed for junctional fragments, with a mean copy number of 1.82 ±1.05 (p = 0.18 compared to the 1100 cGy cohort; Mann-Whitney non-parametric test).

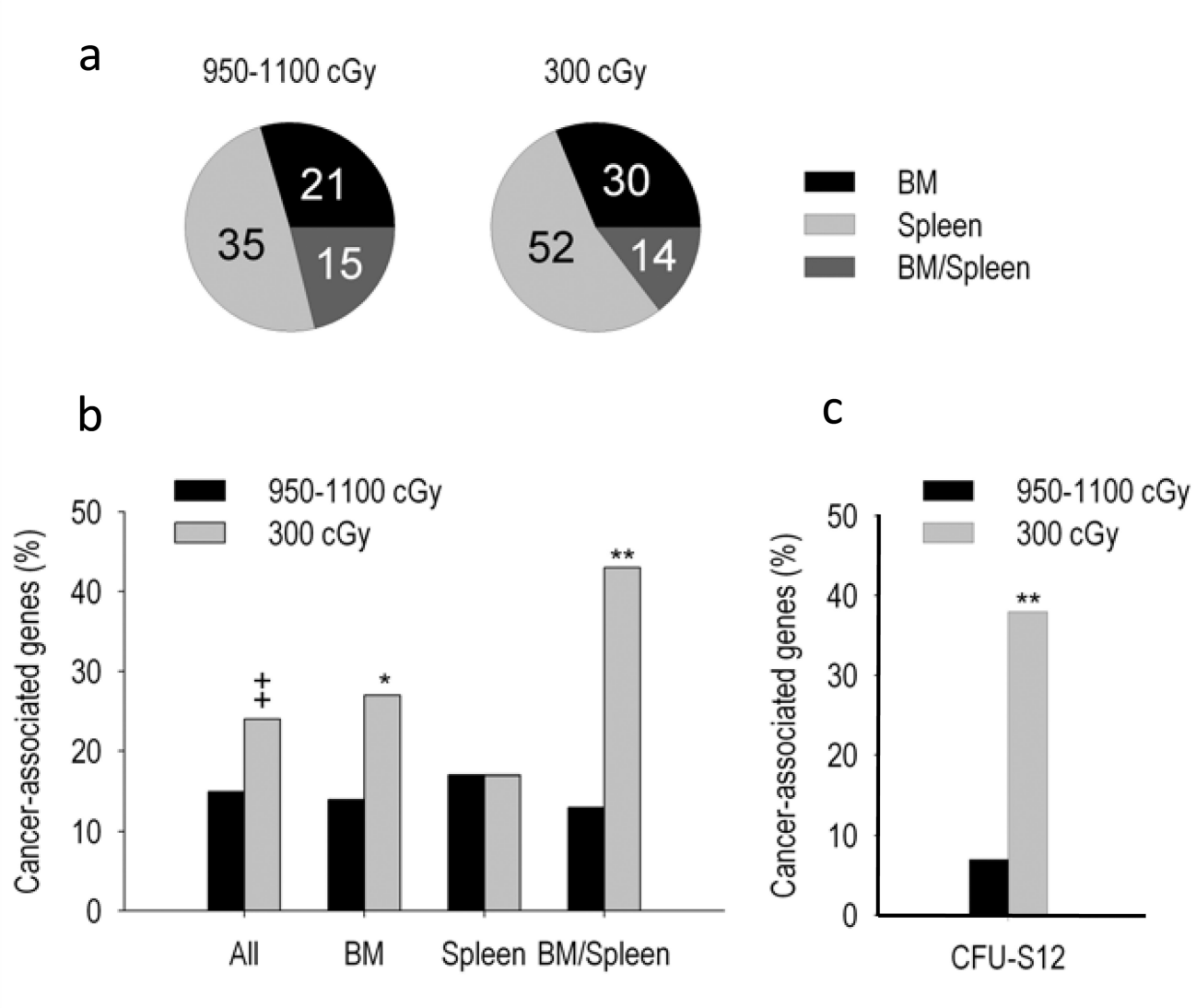

Fig. 4. Distribution of retroviral integration sites and their association with cancer associated genes

(a) Percentage of unique RIS as distributed among different hematopoietic tissues analyzed from primary recipients conditioned with either 950 – 1100 cGy (left) or 300 cGy (right), as indicated; the actual number of insertion sites identified is also shown. **(b, c)** The percentage of cancer-associated genes associated with unique RIS were calculated for primary recipients conditioned with either 950 – 1100 cGy or 300 cGy, as indicated by the black and gray bars, respectively, and also shown as distributed in specific hematopoietic tissues **(b)** and for those recovered from secondary CFU-S12 **(c)**. ‡ $p = 0.061$; * $p = 0.08$; ** $p \leq 0.02$.

Table 1

Primary, secondary and tertiary transplant mice[1]

Murine X-CGD Recipients	Conditioning Regimen		Total
	300 cGy	1100 cGy	
Primary	14	10	24
Secondary	20	20	40
Tertiary	26	20	46
Total	**60**	**50**	**110**

[1]The number of mice transplanted in each group is shown.

Table 2

Retroviral insertion sites in recipients of SF71gp91phox-transduced marrow

	950–1100 cGy					300 cGy				
	Total	**(%)**	**BM**	**BM/ Spleen**	**Spleen**	**Total**	**(%)**	**BM**	**BM/ Spleen**	**Spleen**
Primary Transplants[1]										
Total mappable sites	71		21	14	36	96		29	14	53
Mappable sites with no nearby genes	4	**6%**	1	1	2	0	**0%**	0	0	0
Insertions in Refseq Genes	26	**37%**	9	6	11	32	**33%**	9	3	20
Insertions in Refseq Genes + 10kbp[2]	40	**56%**	14	8	18	42	**44%**	11	5	26
Insertions near TSSs (+/− 5kbp)	17	**24%**	7	5	5	21	**22%**	5	4	12
Closest gene is cancer-associated	11	**15%**	3	2	6	23	**24%**	7	7	9
Secondary CFU-S12[3]										
Total mappable sites	**14**		N/A	N/A	N/A	**26**		N/A	N/A	N/A
Mappable sites with no nearby genes	1	**7%**				0	**0%**			
Insertions in Refseq Genes	4	**29%**				4	**15%**			
Insertions in Refseq Genes + 10kbp	7	**50%**				8	**31%**			
Insertions near TSSs (+/− 5kbp)	4	**29%**				4	**15%**			
Closest gene is cancer-associated	**1**	**7%**				**10**	**38%**			
Secondary and Tertiary Transplants[4]										
Total mappable sites	**21**		7	9	5	**60**		17	24	19
Insertions in Refseq Genes	6	**29%**	2	3	1	29	**48%**	9	12	8
Insertions in Refseq Genes + 10kbp	9	**43%**	2	5	2	43	**72%**	14	17	11
Insertions near TSSs (+/− 5kbp)	5	**24%**	1	2	2	20	**33%**	7	10	3
Closest gene is cancer-associated	**4**	**19%**	2	2	0	**7**	**12%**	2	3	2

Numbers reflect unique sites, which in some cases were identified in more than one animal.

[1] Spleen and bone marrow DNA were analyzed in 8 animals in the 950–1100 cGy group and from 11 animals in the 300 cGy group; only spleen DNA was available from 1 animal in the latter group.

[2] including 10-kbp surrounding area

[3] CFU-S12 were analyzed from 6 primary recipients in the 950–1100 cGy group and 7 primary recipients in the 300 cGy group.

[4] spleen and marrow DNA were analyzed from 16 mice in the 950–1100 cGy cohort (8 secondary, 8 tertiary) and 31 mice in the 300 cGy cohort (13 secondary, 18 tertiary).

Table 3

Integrations identified in primary recipients of SF71gp91phox-transduced marrow that are in or near cancer-associated genes

Mouse ID	Gene	Site	Protein Function	Position from TSS	Orientation	Mouse Chromosome	Human Chromosome	Accession Number
300 cGy								
A1[1]	Elmo1	CFU-S12	Rac GEF	upstream (2653)	Same	13	7p14.1	NM_080288
	Evi1	Spleen/BM/CFU-S12	Transcription	upstream (119137)	Inverse	3	3q24-q28	NM_007963
	Prdm16	CFU-S12	Transcription	upstream (105492)	Same	4	1p36.23-p33	NM_027504
A4[1]	Evi1	Spleen/BM/CFU-S12	Transcription	upstream (109351)	Inverse	3	3q24-q28	NM_007963
	Gse1	Spleen	Unknown	upstream (137710)	Same	8	16q24.1	NM_198671
A6[2]	Usp3	Spleen**	Ubiquitin peptidase	upstream (23334)	Inverse	9	15q22.3	NM_144937
A11[1]	Prdm16	Spleen	Transcription	intron 1 (50858)	Inverse	4	1p36.23-p33	NM_027504
A12[1]	Gse1	BM	Unknown	upstream (152466)	Same	8	16q24.1	NM_198671
	Prkcbp1	BM/CFU-S12	Unknown	intron 2 (25132)	Same	2	20q13.12	NM_027230
A13[1]	AB041803	Spleen/BM/CFU-S12	Unknown	downstream (92581)	Same	6		NM_144555
	Ccr1	CFU-S12	Chemokine receptor	downstream (11830)	Same	9	3p21	NM_009912
	Evi1	CFU-S12	Transcription	upstream (121275	inverse	3	3q24-q28	NM_007963
	Mafk	Spleen	Transcription	intron 1 (1593)	Same	5	7p22.3	NM_010757
A14[1]	B4galt5	CFU-S12	Galactosyltransferase	upstream (30927	Inverse	2	20q13.1-q13.2	NM_019835
	Rreb1	BM	Transcription	upstream (73827)	Same	13		NM_026830
	Usp3	Spleen/BM	Ubiquitin peptidase	upstream (22186)	Inverse	9	15q22.3	NM_144937
	Zfp217	BM	Transcription	intron 1 (853)	Inverse	2	20q13.2	NM_001033299
A15	Chd9	BM	Transcription	intron 1 (4284)	Inverse	8	16q12.2	NM_177224
	Evi1	BM	Transcription	upstream (106963)	Inverse	3	3q24-q28	NM_007963
	Sept9	BM	GTP binding	upstream (89216	Inverse	11	17q25	NM_017380
A16[1]	Ela1	Spleen	Elastase	downstream (14444)	Inverse	15	12q13	NM_033612
	Evi1	Spleen/BM/CFU-S12	Transcription	upstream (106963)	Inverse	3	3q24-q28	NM_007963
A17	Arid1a	Spleen	Transcription	upstream (62600)	Same	4	1p35.3	NM_001080819
	Ly6e	Spleen	Receptor	upstream (35122)	Inverse	15	8q24.3	NM_008529
A18	4930474N05Rik	Spleen	Unknown	downstream (225302)	Inverse	14		NM_175008

Mouse ID	Gene	Site	Protein Function	Position from TSS	Orientation	Mouse Chromosome	Human Chromosome	Accession Number
300 cGy								
	Chek2	Spleen	Kinase	downstream (138690)	Same	5	22q12.1	NM_016681
	Rgs3	Spleen/BM	G-protein signaling	downstream (187630)	Inverse	4	9q32	NM_001081650
	Wasf2	BM	Actin binding	upstream (39937)	Same	4	1p36.11-p34.3	NM_153423
Mouse ID	**Gene**	**Site**	**Protein Function**	**Position from TSS**	**Orientation**	**Mouse Chromosome**	**Human Chromosome**	**Accession Number**
950–1100 cGy								
C2[1]	Arhgap4	BM	Rho GAP	upstream (3597)	Same	X	Xq28	NM_138630
	Dym	Spleen	Unknown	intron 13 (165749)	Inverse	18	18q12-q21.1	NM_027727
	Kdr	Spleen	Receptor	downstream (124315)	Same	5	4q11-q12	NM_010612
C4[1]	Gse1	BM/CFU-S12	Unknown	upstream (118913)	Same	8	16q24.1	NM_198671
C5	Erg	Spleen	Transcription	intron 2 (~ 86200)	Same	16	21q22.3	NM_133659
	Evi1	Spleen	Transcription	upstream (15259)	Same	3	3q24-q28	NM_007963
	Lmo2	Spleen/BM	Transcription	intron 1 (646)	Inverse	2	11p13	NM_008505
C11[1]	Gse1	Spleen/BM	Unknown	upstream (123223)	Inverse	8	16q24.1	NM_198671
C12[1]	Ikzf1	BM	Transcription	intron 1 (3388)	Inverse	11	7p13-p11.1	NM_001025597
	Mpl	Spleen	Receptor	upstream (7494)	Inverse	4	1p34	NM_010823
C14[1]	Zfp36	Spleen	Transcription	upstream (478	Same	7	19q13.1	NM_011756

[1] CFU-S12 analysis done

[2] No marrow available

Table 4

Genes associated with more than one independent retroviral insertion

Nearest gene	Mouse ID	Bm/Spleen/CFU-S	Position from TSS[1]	Orientation
Bcl9l	A11	Spleen	2113	Inverse
Bcl9l	A18	Spleen	9774	Same
Evi1	A1	Spleen/BM/CFU-S	119137	Inverse
Evi1	A4	Spleen/BM/CFU-S	109351	Inverse
Evi1	A13	CFU-S	121275	Inverse
Evi1	A15[2]	BM	106963	Inverse
Evi1	A16[2]	Spleen/BM/CFU-S	106963	Inverse
Evi1	C5	Spleen	15259	Same
Gse1	A4	Spleen	137710	Same
Gse1	A12	BM	152466	Same
Gse1	C4	BM/CFU-S	118913	Same
Gse1	C11	Spleen/BM	123223	Inverse
Prdm16	A1	CFU-S	105492	Same
Prdm16	A11	Spleen	50858	Inverse
Usp3	A6[3]	Spleen	23332	Inverse
Usp3	A14	Spleen/BM	22186	Inverse

[1] All insertions 5' to TSS except Prdm16 insertion in A11, which is in intron 1

[2] Same insertion site

[3] Marrow not available