Review Article   **Open Access**

Identification of Hematopoietic Stem Cell Engraftment Genes in Gene Therapy Studies

John M Powers[1] and Grant D Trobridge[1,2]*

[1]Department of Pharmaceutical Sciences, Washington State University, Pullman, Washington, USA
[2]School of Molecular Biosciences, Washington State University, Pullman, Washington, USA

Abstract

Hematopoietic stem cell (HSC) therapy using replication-incompetent retroviral vectors is a promising approach to provide life-long correction for genetic defects. HSC gene therapy clinical studies have resulted in functional cures for several diseases, but in some studies clonal expansion or leukemia has occurred. This is due to the dyregulation of endogenous host gene expression from vector provirus insertional mutagenesis. Insertional mutagenesis screens using replicating retroviruses have been used extensively to identify genes that influence oncogenesis. However, retroviral mutagenesis screens can also be used to determine the role of genes in biological processes such as stem cell engraftment. The aim of this review is to describe the potential for vector insertion site data from gene therapy studies to provide novel insights into mechanisms of HSC engraftment. In HSC gene therapy studies dysregulation of host genes by replication-incompetent vector proviruses may lead to enrichment of repopulating clones with vector integrants near genes that influence engraftment. Thus, data from HSC gene therapy studies can be used to identify novel candidate engraftment genes. As HSC gene therapy use continues to expand, the vector insertion site data collected will be of great interest to help identify novel engraftment genes and may ultimately lead to new therapies to improve engraftment.

Keywords: Engraftment; Gene therapy; Hematopoietic stem cell; Viral vector; Insertional mutagenesis

Introduction

Gene therapy using hematopoietic stem cells (HSC) has enormous potential to treat diseases of the hematopoietic system including immune diseases. In this approach, HSCs are collected from a patient, gene-modified *ex vivo* using integrating retroviral vectors, and then infused into a patient. To date retroviral vectors have been the only effective gene delivery system for HSC gene therapy. This is primarily due to the ability of retroviral vectors to efficiently integrate into the genome, thereby allowing efficient transmission of therapeutic transgenes to all HSC-derived cells via mitosis. Gene delivery to HSCs using integrating vectors thus allows for efficient delivery to HSC-derived mature hematopoietic cells.

Retroviral vectors have been used successfully in HSC gene therapy clinical trials for several genetic diseases including X-linked severe combined immunodeficiency (SCID-X1) [1,2], adenosine deaminase deficiency (SCID-ADA) [3,4], chronic granulomatous disease (CGD) [5], and adrenoleukodystrophy (ALD) [6]. HSC gene therapy also has the potential to treat acquired diseases of the hematopoietic system such as human immunodeficiency virus infection and acquired immunodeficiency syndrome (HIV/AIDS) [7]. While recent clinical studies have shown promise, the use of retroviral vectors for gene therapy has drawbacks. Gene therapy using HSCs with integrating retroviral vectors can dysregulate cellular genes near the provirus integration site leading to adverse side effects including leukemia [8-10].

Previous human clinical studies have documented the impact of vector-mediated dysregulation of host genes. In both the French and United Kingdom SCID-X1 studies vector-mediated gene dysregulation resulted in the development of leukemia [8-10]. In a CGD study conducted by Ott and colleagues, proviral insertion sites led to the clonal expansion of gene-modified cells over time [5,11]. In this CGD study the vector provirus provided the gene-modified HSCs with a

survival advantage, leading to the clonal dominance of a small subset of gene-modified cells in the patient. In the above SCID-X1 and CGD studies, the ability to determine where the provirus had inserted into the genome allowed for the identification of nearby genes that were dysregulated, leading to clonal expansion. The integrated provirus can thus be used as a molecular tag to identify dysregulated genes in gene therapy studies.

Gene-modified HSCs that are infused into patients undergo various selective pressures during the process of stem cell engraftment. First, the cells must home to the stem cell niche and resist apoptosis during this process. Once in the bone marrow, HSCs begin the production of all hematopoietic cell lineages which involves survival, stem cell self-renewal, proliferation and differentiation. Together, these processes are referred to as engraftment [12], and many genes could potentially provide a selective advantage to repopulating cells if dysregulated. The gene-modified cells that are infused into a patient are a polyclonal population, where different cells have vector proviruses integrated at different chromosomal locations. There may be millions of clones that are infused into a patient and this polyclonal population of cells is, in essence, a library of clones with many different unique integration sites. If a clone has a vector integrant near a gene that may influence the efficiency of engraftment, that clone has a selective

advantage and may be over-represented when engrafted cells are analyzed (Figure 1). Thus, pre-clinical and clinical HSC gene therapy studies provide an opportunity to identify genes near vector proviruses in over-represented clones. These genes may have conferred an increased survival and proliferation advantage to the infused cells due to dysregulation mediated by the integrated provirus.

This review covers the potential of HSC gene therapy studies to identify genes that play a role in engraftment. The use of retroviral mutagenesis screens to identify dysregulated genes involved in cancer has provided an enormous wealth of data [13]. These screens have been used to identify genes that have an effect on the development and progression of leukemia by analyzing replicating virus insertion sites to identify nearby genes that contributed to tumorigenesis and leukemic development [14]. However it is clear that non-replicating viruses can also perturb nearby genes causing genotoxicity. Thus, HSC gene therapy studies are *de facto* mutagenesis screens where a library of vector-mutagenized cells are infused into patients and clones with a selective advantage to engraft can become over-represented. Although the goal of clinical gene therapy is to develop cures for life-threatening diseases, the data obtained from patient samples can also provide information into the role of genes in hematopoietic processes. Analysis of retroviral integration sites in preclinical and clinical HSC gene therapy studies has the potential to identify novel genes involved in engraftment, and also other hematopoietic processes. Identifying novel engraftment genes can improve our understanding of this complex process, and also identify new therapeutic targets to enhance engraftment.

Selective Pressure During Transplantation in Gene Therapy Studies

HSCs are commonly harvested from the peripheral blood after mobilization. In order to mobilize HSCs from the bone marrow into the peripheral blood, patients receive recombinant human granulocyte-colony stimulating factor (G-CSF). The patient's peripheral blood is collected and enriched for HSCs using the CD34+ marker. HSCs are then cultured *ex vivo* and exposed to viral vectors. The *ex vivo* culture period varies between studies, but can be for approximately 1-4 days. During this time, vector proviruses integrate into the host genome, leading to a polyclonal population of HSCs that possess numerous proviral insertion sites. This highly polyclonal population of repopulating cells with vector proviruses at many integration sites is in essence a library where there is the potential to dysregulate a wide variety of genes. Some proviral integration sites may become over-represented during *ex vivo*

Figure 1: Selective pressure for HSCs to engraft enriches for clones with proviral integration sites that confer an engraftment advantage. After harvesting patient HSCs the cells are transduced with retroviral vectors, leading to a polyclonal population of cells with numerous different proviral insertion sites. Following transfusion of the cells into the patient, cells with insertions near genes that confer a competitive engraftment advantage (red, purple clones) will become enriched. Provirus vector integration sites in the purple and red cells are thus over-represented.

culture due to a proliferative/survival advantage of clone(s) with this provirus.

Prior to the infusion of gene-modified HSCs, patients may be treated with chemotherapy agents or irradiation to help enhance the engraftment efficiency. Gene-modified HSCs are re-infused into the patient intravenously. The cells migrate into the bone marrow before finally residing in the sinusoids and perivascular tissue [15,16]. Both homing and hematopoiesis are integral aspects of engraftment. Cells that have reached the stem cell niche through homing will begin producing mature myeloid and lymphoid cells from each blood lineage. Hematopoiesis continues through the action of long-term HSCs, which are capable of self-renewal for life-long generation of the patient's mature blood cells.

When HSCs are infused into the patient intravenously, the cells must travel from the peripheral blood into the bone marrow, eventually reaching their niche to repopulate the blood system. This process, known as homing, is a multistep process that relies on the action and interactions of various chemokines, cytokines and other proteins. Examples include stromal derived factor 1 (SDF-1) and CXCR4, adhesion molecules such as very late antigen 4 and 5 (VLA-4/5), lymphocyte function associated antigen 1 (LFA-1), and α4ß1 integrin interaction with vascular cell adhesion protein 1 (VCAM-1) [12,17-19]. Circulating HSCs roll and tether to the blood vessel walls through the action of E-and P-selectins and VCAM-1. Tethered HSCs extravasate through the bone marrow endothelium before lodging into the bone marrow stem cell niche (Figure 2) [15,20-23]. The entire process is thought to occur within a matter of hours following infusion [12]. HSCs with vector provirus insertions near genes that enhance homing are more likely to engraft and thus these clones may become over-represented during this process.

After reaching the bone marrow and lodging in the perivascular region, HSCs begin the process of repopulating the patient's blood system. During the process of proliferation, some of the daughter cells produced by the infused HSCs remain as quiescent HSCs, while others self-renew or become committed to either the myeloid or lymphoid system as progenitor cells [24]. As gene-modified daughter cells divide, they begin to produce all of the cellular subsets of each lineage, with all progeny carrying the transgene of interest. For HSCs that harbor proviral integrations near genes involved in stem cell renewal or expansion, dysregulation may provide the HSCs with an engraftment advantage. HSCs that have vectors integrated near genes that provide a selective advantage during these processes of self-renewal or expansion will be more likely to engraft, repopulate, and persist in the patient long-term. Examples of such genes include RUNX1 [25-27], globin transcription factor 2 (GATA2) [28,29], spleen focus forming virus proviral integration oncogene (Spi-1), the transcription factor PU.1 [30,31], as well as homeobox A (HOXA) [32,33].

HSC clones that have vector proviral insertions that dysregulate genes involved with proliferation or survival have a selective advantage at all stages of engraftment. In order for infused cells to engraft and repopulate the patient's blood system they must make it to the bone marrow without undergoing apoptosis. Dysregulation of genes that confer a survival advantage by inhibiting apoptosis, such as MCL1, could benefit HSCs prior to reaching and after lodging in the bone marrow niche [34,35]. Clones with dysregulated genes that provide a proliferative advantage to HSCs, such as CCND2, have been over-represented in gene therapy studies.
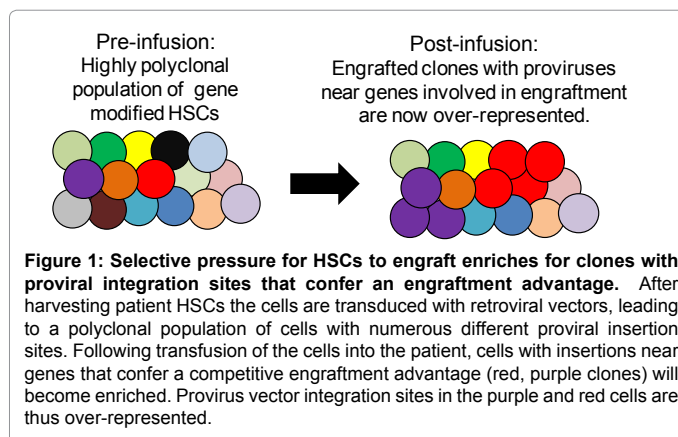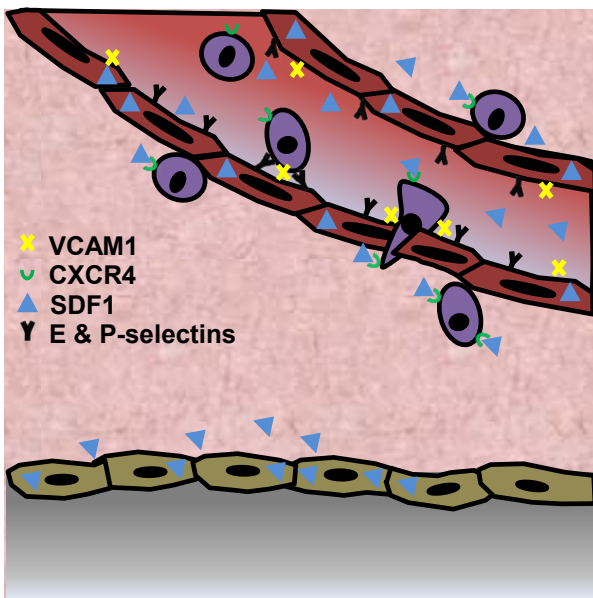
Figure 2: Infused HSCs must home to their bone marrow niches before they can begin the process of hematopoiesis. After infusion of HSCs into the peripheral blood, shown as purple circles, HSCs begin the process of homing to the marrow. E- and P-selectins and VCAM1 on the vessel walls tether circulating HSCs and allow for rolling on the vessel wall to occur. This is followed by extravasation of the HSCs through the extracelluar matrix into the bone marrow. The release of SDF-1 from osteoblasts and epithelial tissues in the bone marrow binding to the HSCs CXCR4 receptors is important. After reaching the bone marrow, HSCs then migrate to the perivascular regions and begin the process of hematopoiesis.

Retroviral Genotoxicity

Integrated vector proviruses have the potential to dysregulate the expression of nearby host cell genes flanking the integration site [36]. Depending on the integration site of the provirus, vector-mediated genotoxicity can lead to gene over-expression, inactivation, or production of novel gene transcripts (Figure 3). Transcriptionally active LTR regions with strong promoters or enhancers are important in the development of genotoxicity. Integrating replication-competent retroviruses are well known for their potential to activate nearby genes leading to oncogenesis. However, it was previously believed that replication-incompetent viral vectors might not mediate significant genotoxicity. Unfortunately, clinical studies have shown that replication-incompetent vectors still cause genotoxicity, in some cases leading to clonal expansion and leukemia.

Identification of Retroviral Integration Sites and Nearby Dysregulated Genes

In order to identify integration sites, genomic DNA is extracted from the bone marrow or the peripheral blood of patients that have received gene-modified HSCs. After isolation of the DNA, the amplification of provirus LTR-chromosome junctions is commonly conducted using ligation-mediated PCR (LM-PCR) or linear-amplification-mediated PCR (LAM-PCR) [37]. LM-PCR utilizes frequent cutting restriction enzymes that cut genomic DNA into small fragments. Some of these fragments contain an LTR-chromosome junction. Following digestion, these fragments are then ligated to linkers and PCR amplified. LAM-PCR employs linear amplification of LTR-chromosome junctions followed by double-stranded DNA (dsDNA) synthesis. The dsDNA sequences are then digested with restriction enzymes, followed by

linker ligation to the sequence and nested PCR. Non-restrictive linear-amplification-mediated PCR (nrLAM-PCR) has been developed which avoids restriction digest bias of recovered integration sites [38].

Alternative non-PCR methods, such as shuttle vector rescue, also exist [39]. In shuttle vector rescue, integrated vector proviruses contain a bacterial origin of replication and a selection gene. Peripheral blood DNA from patients is digested with restriction enzymes or randomly sheared, ligated, and then transformed into bacteria which are grown as colonies. These plasmids contain an LTR-chromosome junction that can be sequenced with an LTR specific primer. Shuttle vector rescue avoids PCR-based skewing of obtained integration sites.

The availability of the human genome sequence, as well as the genomes of other model organisms such as mice and macaques has allowed for rapid identification of genes near vector proviruses in clinical and preclinical studies. Following sequencing of the LTR-chromosome junction, sequence reads can be aligned to the human genome using the BLAST-like alignment tool (BLAT) [40]. Genes and oncogenes located close to the vector integration site can be identified based on the annotation of the human genome. Thus, through the combination of LTR-chromosome junction amplification, next-generation sequencing, and bioinformatics, vector proviruses serve as ideal molecular tags to identify nearby genes.

Over-Represented Gene Classes Near Proviruses in HSC Gene Therapy Studies

Proviral vector integration occurs throughout the genome, but different viral vector types have different integration site preferences. HIV based lentiviral vectors favor active genes, while murine leukemia virus vectors (MLV) favor transcription start sites [41,42]. Gamma-retroviruses, such as MLV, have a strong preference for integration sites involving previously identified common integration sites (CISs) in the retroviral tagged cancer gene database (RTCGD) [43]. The RTCGD is composed of retroviral integration site data acquired from mouse tumors from a variety of different studies and tumor types [44]. The RTCGD allows researchers to identify candidate cancer genes dysregulated by proviruses that may play a role in human cancer development and progression [13,45].
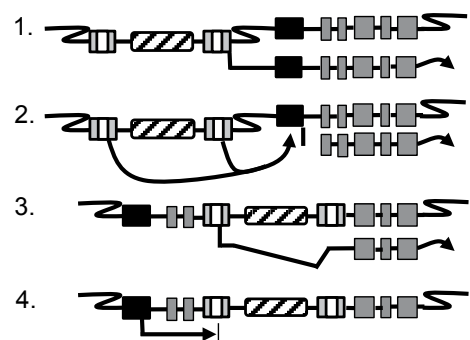
Figure 3: Mechanisms of insertional mutagenesis. (1) 3' proviral LTRs can drive over-expression of nearby genes. (2) Enhancers in the LTRs can activate nearby promoters leading to increased transcription. (3) Proviral insertion within a host gene and transcription from the 5' LTR can lead to the creation of novel gene transcripts. (4) Premature polyadenylation of host cell gene transcripts can be caused by proviral insertion within a gene. Black boxes represent the host gene promoter and grey squares represent the exons. Grey boxes containing white rectangles represent proviral LTRs and striped rectangles are used to show proviral transgenes.

HSC gene therapy trials utilizing MLV and lentiviral vectors have shown that proviral insertions are observed in specific classes of genes [46-50]. Both MLV and lentiviral vector proviruses are over-represented near genes involved in the establishment and/or maintenance of chromatin architecture, signal transduction, and cell cycle [51]. Lentiviral vector proviruses were also over-represented near genes involved in chromatin remodeling and phosphorylation. Many of the genes identified in retroviral mutagenesis screens are linked in gene networks involved in cellular regulatory process such as apoptosis, signal transduction, and transcriptional regulation [51]. The over-representation of vector provirus near genes involved in such processes is likely due to the survival and proliferative advantages that such mutations could confer to HSCs. For example, in a retrospective study of vector integration sites in rhesus macaques that had received autologous MLV transduced hematopoietic repopulating cells, the MDS/EVI1 site was identified as a hot spot of vector insertion [52]. It is likely that vector provirus dysregulation of this locus provided the infused cells with the potential for increased survival, proliferation, or both. Studies have shown that the overexpression of EVI1/MDS1 has the potential to delay or inhibit the myeloid differentiation of HSCs, while increasing the proliferation of HSCs and progenitor cells [53]. This has also been reported for mouse and monkey HSCs [46,54,55]. Proviral integration leading to the dysregulation of the EVI1/MDS1 gene complex can lead to the over-expression of either or both genes, inhibiting cellular differentiation. Dysregulation of this locus has been shown to be involved in clonal expansion and leukemic development, with integration sites likely providing a survival or proliferation advantage to transduced HSCs. Extended culture of macaque HSCs revealed an increase in HSC clones with integration sites in or near the EVI/MDS1 locus compared to other infused clones [55]. Thus, analysis of vector provirus integration sites can provide evidence for dysregulated genes in the absence of adverse events. This data demonstrates the potential of preclinical gene therapy studies to identify genes involved in engraftment and hematopoiesis pathways, as well as their role in gene networks related to these processes.

Engraftment Genes Identified in Retroviral Mutagenesis Screens

Retroviral mutagenesis screens have played an important role in determining genes involved in hematopoiesis. Forward retroviral mutagenesis screens in hematopoietic cells have been highly successful in identifying genes involved in migration, proliferation, and expansion. Identified genes include *Rac2, Jak/Stat*, and *Notch* [44,56]. Notch expression is important in embryonic development, and throughout life for tissue homeostasis, [57]. Dysregulated expression of Notch can affect HSC cell differentiation and lead to skewed differentiation of hematopoietic lineages [58]. A study of murine tumor retroviral insertion sites by Suzuki and colleagues identified Notch as a CIS [59]. Based on the role of Notch family genes in hematopoietic differentiation, dysregulation by proviral insertional mutagenesis has the potential to enhance hematopoietic repopulation.

The role of GATA proteins, especially GATA-1 and GATA-2, is also well established in HSC biology. Both are highly expressed in erythroid precursors. As cells differentiate the GATA-2 level decreases while GATA-1 expression is maintained at high levels [60]. GATA-2 expression is essential for HSC maintenance, survival, and proliferation [29,61,62]. Since GATA-2 expression levels are important in HSC proliferation and differentiation, dysregulation of GATA-2 expression would be expected to enhance engraftment following transplantation of gene-modified cells. GATA2 has in fact been identified as a CIS [28].

Engraftment Genes Identified in Gene Therapy Studies

Replication-competent retroviruses cause insertional mutagenesis, leading to their common use in mutagenesis screens [13,63]. Although replication-incompetent vectors are capable of providing only a single-hit genetic modification via provirus integration, they can still be utilized to identify dysregulated genes. Deichmann and colleagues investigated integration site data from five clinical gene therapy trials and three pre-clinical trials [64]. This retrospective analysis showed that transplanted gene-modified HSCs had very similar integration sites and dysregulated genes. The most frequent CISs were insertions that would dysregulate genes leading to clonal expansion or leukemic development, such as LMO2 and MDS1/EVI1. Thus, the same dysregulated genes are often observed in multiple gene therapy studies.

The CGD study by Ott and colleagues revealed that the dysregulation of PRDM16 and EVI1/MDS1 caused clonal expansion [5]. The expression of PRDM16 has been shown to be involved in HSC maintenance and renewal. Cells lacking expression of PRDM16 exhibit increased cell death so overexpression is expected to lead to an over-representation of clones with proviral insertions near PRDM16. PRDM16 may be in a gene network involving MDS1/EVI1, GATA2, and other genes that affect HSCs [65,66]. Thus, dysregulation of the PRDM16 gene locus could have an effect on the signaling pathways for other genes involved in normal hematopoiesis, expanding the effects of dysregulation of the PRDM16 gene. During the French SCID-X1 study the dysregulation of LMO2 likely led to the proliferation of common lymphoid progenitor cells. Over time, dysregulation of LMO2 led to the expansion of the lymphoid hematopoietic lineage. LMO2 is expressed only in the earliest stages of lymphopoiesis, with the continued expression in mature T-lymphocytes leading to the development of lymphoblastic leukemias. In the French SCID-X1 study, the dysregulated expression of LMO2 ultimately resulted in lymphoblastic leukemia [67,68].

With the large proviral integration site data sets that gene therapy trials can provide, the ability to quickly and efficiently analyze the integration profiles sites should help to identify candidate engraftment genes. One such utility is the QuickMap utility provided by the gene therapy safety group (GTSG) [69]. The QuickMap utility relies on cancer gene lists provided by the Catalogue of Somatic Mutations in Cancer (COSMIC) [70] as well as the RTCGD. The QuickMap utility is able to rapidly analyze sequence data from LTR-chromosome junctions to determine the proviral integration site. Once the integration site is known, it can identify if the vector provirus is within a gene including known oncogenes, within a CpG island, or in a repetitive DNA sequence. Further, the software compares the integration site data to a randomly generated data set of one million integrations as a control. This database has been utilized previously with *ex vivo* transduced human cells to explore the effect of chemoselection of HSCs on integration site patterns [71]. Within the analyzed data, two of the sixteen CISs identified, STAT5B and TNRC6C were previously identified as CISs.

One limitation of the QuickMap utility is the submission limit of fifty-thousand sequences per analysis, although this limit can be temporarily increased by contacting the GTSG. As many gene therapy studies now use next generation sequencing where sequence reads can be in the hundreds of thousands to millions this restriction may limit

future use of QuickMap. With the increasing number of pre-clinical and clinical HSC gene therapy trials, the development of new utilities to efficiently analyze millions of integration site sequence reads from next generation sequencing may aid in the discovery of additional CISs in HSCs. These CISs may in turn identify novel engraftment genes.

A Preclinical Gene Therapy Study Identifies Candidate Engraftment Genes

A study by Kiem and colleagues of three baboons that received baboon hematopoietic repopulating cells exposed to a gammaretroviral vector revealed a CIS of 664 base pairs in a CpG island that existed between zinc finger protein 91 (ZFP91) and leupaxin (LPXN) [72]. It was hypothesized that the CIS between ZFP91 and LPXN lead to the dysregulation of one or both genes, providing the HSCs with an engraftment advantage. Thus these two genes may play a role in normal engraftment pathways. This study suggests that other HSC gene therapy trials may identify CISs that are near genes including micro-RNAs previously not associated with engraftment. Previous studies regarding the roles of genes involved in the engraftment process have been utilized to improve HSC transplantation [73]. These identified genes could serve as targets for novel small molecule drugs to increase the gene expression of the identified targets prior to HSC infusion. These drugs could be of benefit to patients receiving any type of HSC transplantion, and may be of significant value in the field of cord blood transplantation where low cell numbers and low engraftment limit clinical use [74].

Use of Retroviral Vectors to Identify Genes Involved in Other Biological Processes

The ability of retroviral vectors to dysregulate genes can be exploited to better understand many other biological processes. If a library of cells mutagenized with retroviral vectors is placed under any selective pressure, those clones with integrants near genes that provide a selective advantage will be enriched. For example, it should be possible to analyze over-represented genes in specific lineages of hematopoietic repopulating cells. If a set of genes is overrepresented near vector proviruses in myeloid but not lymphoid repopulating cells those genes are candidates for affecting myeloid differentiation and expansion. There are many possible uses of this technology. Replication-incompetent vectors have been used to identify genes involved in liver cancer [75] and we are using this approach to study the development of acute myeloid leukemia (GDT unpublished data). Further, retroviral integration sites could provide insight into genes that play a role in the metastasis of solid tumor cells to the bone marrow, such as in prostate cancer [76]. Analysis of gene expression in cancer cells that have metastasized to the bone marrow could provide insight into genes that helped them engraft in the bone marrow. Identification of genes that assist in homing and engraftment would be potential molecular targets to reduce the likelihood of metastasis to the bone marrow. Thus, the data obtained from mutagenesis screens using replication-incompetent vectors should provide useful information about the physiological role of genes and their interactions in gene networks for many biological processes including cancer.

Summary

Pre-clinical and clinical trials utilizing HSCs with retroviral vectors have yielded important information regarding the effects of retroviral insertional mutagenesis on host genes. Through the use of annotated genomes for humans and model organisms, retroviral insertion sites in gene therapy trials can be mapped to the genome to determine nearby potentially dysregulated genes. Advances in bioinformatics, and the creation of cancer gene databases, such as the RTCGD, have been instrumental in identifying CISs and thus dysregulated genes.

As the number of HSC gene therapy trials increases more data regarding the role of genes in biological processes will be obtained. The data from these studies can be mined to identify genes that provide a competitive engraftment advantage to infused HSCs. Studies without observed abnormal hematopoiesis following engraftment still have the potential to identify genes that have an effect on hematopoiesis and engraftment. Novel engraftment genes might be targeted with small molecule drugs to increase the engraftment efficiency of infused HSCs. Therefore, HSC gene therapy trials carry the potential to improve HSC transplantation by providing data that identifies genes and gene networks involved in engraftment and hematopoietic pathways.

References

1. Gaspar HB, Parsley KL, Howe S, King D, Gilmour KC, et al. (2004) Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. Lancet 364: 2181-2187.

2. Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, Gross F, Yvon E, et al. (2000) Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science 288: 669-672.

3. Aiuti A, Slavin S, Aker M, Ficara F, Deola S, et al. (2002) Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. Science 296: 2410-2413.

4. Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, et al. (2009) Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med 360: 447-458.

5. Ott MG, Schmidt M, Schwarzwaelder K, Stein S, Siler U, et al. (2006) Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat Med 12: 401-409.

6. Cartier N, Hacein-Bey-Abina S, Bartholomae CC, Veres G, Schmidt M, et al. (2009) Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. Science 326: 818-823.

7. Scherer LJ, Rossi JJ (2011) Ex vivo gene therapy for HIV-1 treatment. Hum Mol Genet 20: R100-R107.

8. Hacein-Bey-Abina S, von Kalle C, Schmidt M, Le Deist F, Wulffraat N, et al. (2003) A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med 348: 255-256.

9. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, et al. (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 302: 415-419.

10. Howe SJ, Mansour MR, Schwarzwaelder K, Bartholomae C, Hubank M, et al. (2008) Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest 118: 3143-3150.

11. Stein S, Ott MG, Schultze-Strasser S, Jauch A, Burwinkel B, et al. (2010) Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nat Med 16: 198-204.

12. Lapidot T, Dar A, Kollet O (2005) How do stem cells find their way home? Blood 106: 1901-1910.

13. Mikkers H, Berns A (2003) Retroviral insertional mutagenesis: tagging cancer pathways. Adv Cancer Res 88: 53-99.

14. Uren AG, Kool J, Berns A, van Lohuizen M (2005) Retroviral insertional mutagenesis: past, present and future. Oncogene 24: 7656-7672.

15. Lo Celso C, Scadden DT (2011) The haematopoietic stem cell niche at a glance. J Cell Sci 124: 3529-3535.

16. Ding L, Morrison SJ (2013) Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. Nature 495: 231-235.

17. Williams DA, Zheng Y, Cancelas JA (2008) Rho GTPases and regulation of hematopoietic stem cell localization. Methods Enzymol 439: 365-393.

18. Peled A, Grabovsky V, Habler L, Sandbank J, Arenzana-Seisdedos F, et al. (1999) The chemokine SDF-1 stimulates integrin-mediated arrest of CD34(+) cells on vascular endothelium under shear flow. J Clin Invest 104: 1199-1211.

19. Papayannopoulou T, Craddock C, Nakamoto B, Priestley GV, Wolf NS (1995) The VLA4/VCAM-1 adhesion pathway defines contrasting mechanisms of lodgement of transplanted murine hemopoietic progenitors between bone marrow and spleen. Proc Natl Acad Sci U S A 92: 9647-9651.

20. Quesenberry PJ, Becker PS (1998) Stem cell homing: rolling, crawling, and nesting. Proc Natl Acad Sci U S A 95: 15155-15157.

21. Frenette PS, Subbarao S, Mazo IB, von Andrian UH, Wagner DD (1998) Endothelial selectins and vascular cell adhesion molecule-1 promote hematopoietic progenitor homing to bone marrow. Proc Natl Acad Sci U S A 95: 14423-14428.

22. Shirvaikar N, Marquez-Curtis LA, Janowska-Wieczorek A (2012) Hematopoietic Stem Cell Mobilization and Homing after Transplantation: The Role of MMP-2, MMP-9, and MT1-MMP. Biochem Res Int 2012: 685267.

23. Kiel MJ, Morrison SJ (2008) Uncertainty in the niches that maintain haematopoietic stem cells. Nat Rev Immunol 8: 290-301.

24. Orkin SH, Zon LI (2008) Hematopoiesis: an evolving paradigm for stem cell biology. Cell 132: 631-644.

25. Wotton S, Stewart M, Blyth K, Vaillant F, Kilbey A, et al. (2002) Proviral insertion indicates a dominant oncogenic role for Runx1/AML-1 in T-cell lymphoma. Cancer Res 62: 7181-7185.

26. Brady G, Elgueta Karstegl C, Farrell PJ (2013) Novel function of the unique N-terminal region of RUNX1c in B cell growth regulation. Nucleic Acids Res 41: 1555-1568.

27. Ichikawa M, Goyama S, Asai T, Kawazu M, Nakagawa M, et al. (2008) AML1/Runx1 negatively regulates quiescent hematopoietic stem cells in adult hematopoiesis. J Immunol 180: 4402-4408.

28. Slape C, Hartung H, Lin YW, Bies J, Wolff L, et al. (2007) Retroviral insertional mutagenesis identifies genes that collaborate with NUP98-HOXD13 during leukemic transformation. Cancer Res 67: 5148-5155.

29. Bresnick EH, Katsumura KR, Lee HY, Johnson KD, Perkins AS (2012) Master regulatory GATA transcription factors: mechanistic principles and emerging links to hematologic malignancies. Nucleic Acids Res 40: 5819-5831.

30. Paul R, Schuetze S, Kozak SL, Kabat D (1989) A common site for immortalizing proviral integrations in Friend erythroleukemia: molecular cloning and characterization. J Virol 63: 4958-4961.

31. Ridinger-Saison M, Boeva V, Rimmele P, Kulakovskiy I, Gallais I, et al. (2012) Spi-1/PU.1 activates transcription through clustered DNA occupancy in erythroleukemia. Nucleic Acids Res 40: 8927-8941.

32. Bijl J, Sauvageau M, Thompson A, Sauvageau G (2005) High incidence of proviral integrations in the Hoxa locus in a new model of E2a-PBX1-induced B-cell leukemia. Genes Dev 19: 224-233.

33. Wheadon H, Ramsey JM, Dobbin E, Dickson GJ, Corrigan PM, et al. (2011) Differential Hox expression in murine embryonic stem cell models of normal and malignant hematopoiesis. Stem Cells Dev 20: 1465-1476.

34. Alenzi FQ, Alenazi BQ, Ahmad SY, Salem ML, Al-Jabri AA, et al. (2009) The haemopoietic stem cell: between apoptosis and self renewal. Yale J Biol Med 82: 7-18.

35. Suzuki T, Minehata K, Akagi K, Jenkins NA, Copeland NG (2006) Tumor suppressor gene identification using retroviral insertional mutagenesis in Blm-deficient mice. Embo J 25: 3422-3431.

36. Trobridge GD (2011) Genotoxicity of retroviral hematopoietic stem cell gene therapy. Expert Opin Biol Ther 11: 581-593.

37. Schmidt M, Glimm H, Wissler M, Hoffmann G, Olsson K, et al. (2003) Efficient characterization of retro-, lenti-, and foamyvector-transduced cell populations by high-accuracy insertion site sequencing. Ann N Y Acad Sci 996: 112-121.

38. Paruzynski A, Arens A, Gabriel R, Bartholomae CC, Scholz S, et al. (2010) Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. Nat Protoc 5: 1379-1395.

39. Trobridge GD, Miller DG, Jacobs MA, Allen JM, Kiem HP, et al. (2006) Foamy virus vector integration sites in normal human cells. Proc Natl Acad Sci U S A 103: 1498-1503.

40. Kent WJ (2002) BLAT--the BLAST-like alignment tool. Genome Res 12: 656-664.

41. Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, et al. (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol 2: E234.

42. Bushman F, Lewinski M, Ciuffi A, Barr S, Leipzig J, et al. (2005) Genome-wide analysis of retroviral DNA integration. Nat Rev Microbiol 3: 848-858.

43. Montini E, Cesana D, Schmidt M, Sanvito F, Ponzoni M, et al. (2006) Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat Biotechnol 24: 687-696.

44. Akagi K, Suzuki T, Stephens RM, Jenkins NA, Copeland NG (2004) RTCGD: retroviral tagged cancer gene database. Nucleic Acids Res 32: D523-D527.

45. Gabriel R, Eckenberg R, Paruzynski A, Bartholomae CC, Nowrouzi A, et al. (2009) Comprehensive genomic access to vector integration in clinical gene therapy. Nat Med 15: 1431-1436.

46. Kustikova O, Fehse B, Modlich U, Yang M, Dullmann J, et al. (2005) Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. Science 308: 1171-1174.

47. Deichmann A, Hacein-Bey-Abina S, Schmidt M, Garrigue A, Brugman MH, et al. (2007) Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. J Clin Invest 117: 2225-2232.

48. Cattoglio C, Pellin D, Rizzi E, Maruggi G, Corti G, et al. (2010) High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. Blood 116: 5507-5517.

49. Biasco L, Ambrosi A, Pellin D, Bartholomae C, Brigida I, et al. (2011) Integration profile of retroviral vector in gene therapy treated patients is cell-specific according to gene expression and chromatin conformation of target cell. EMBO Mol Med 3: 89-101.

50. Beard BC, Dickerson D, Beebe K, Gooch C, Fletcher J, et al. (2007) Comparison of HIV-derived lentiviral and MLV-based gammaretroviral vector integration sites in primate repopulating cells. Mol Ther 15: 1356-1365.

51. Cattoglio C, Facchini G, Sartori D, Antonelli A, Miccio A, et al. (2007) Hot spots of retroviral integration in human CD34+ hematopoietic cells. Blood 110: 1770-1778.

52. Calmels B, Ferguson C, Laukkanen MO, Adler R, Faulhaber M, et al. (2005) Recurrent retroviral vector integration at the MDS1-EVI1 locus in non-human primate hematopoietic cells. Blood.

53. Steinleitner K, Rampetsreiter P, Koffel R, Ramanathan G, Mannhalter C, et al. (2012) EVI1 and MDS1/EVI1 expression during primary human hematopoietic progenitor cell differentiation into various myeloid lineages. Anticancer Res 32: 4883-4889.

54. Du Y, Jenkins NA, Copeland NG (2005) Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. Blood 106: 3932-3939.

55. Sellers S, Gomes TJ, Larochelle A, Lopez R, Adler R, et al. (2010) Ex vivo expansion of retrovirally transduced primate CD34+ cells results in overrepresentation of clones with MDS1/EVI1 insertion sites in the myeloid lineage after transplantation. Mol Ther 18: 1633-1639.

56. Kustikova OS, Geiger H, Li Z, Brugman MH, Chambers SM, et al. (2007) Retroviral vector insertion sites associated with dominant hematopoietic clones mark "stemness" pathways. Blood 109: 1897-1907.

57. Bigas A, Espinosa L (2012) Hematopoietic stem cells: to be or Notch to be. Blood 119: 3226-3235.

58. Milner LA, Bigas A, Kopan R, Brashem-Stein C, Bernstein ID, et al. (1996) Inhibition of granulocytic differentiation by mNotch1. Proc Natl Acad Sci U S A 93: 13014-13019.

59. Suzuki T, Shen H, Akagi K, Morse HC, Malley JD, et al. (2002) New genes involved in cancer identified by retroviral tagging. Nat Genet 32: 166-174.

60. Farina SF, Girard LJ, Vanin EF, Nienhuis AW, Bodine DM (1995) Dysregulated expression of GATA-1 following retrovirus-mediated gene transfer into murine hematopoietic stem cells increases erythropoiesis. Blood 86: 4124-4133.

61. Tsai FY, Orkin SH (1997) Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood 89: 3636-3643.

62. Ling KW, Ottersbach K, van Hamburg JP, Oziemlak A, Tsai FY, et al. (2004) GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. J Exp Med 200: 871-882.

63. Du Y, Spence SE, Jenkins NA, Copeland NG (2005) Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. Blood 106: 2498-2505.

64. Deichmann A, Brugman MH, Bartholomae CC, Schwarzwaelder K, Verstegen MM, et al. (2011) Insertion sites in engrafted cells cluster within a limited repertoire of genomic areas after gammaretroviral vector gene therapy. Mol Ther 19: 2031-2039.

65. Aguilo F, Avagyan S, Labar A, Sevilla A, Lee DF, et al. (2011) Prdm16 is a physiologic regulator of hematopoietic stem cells. Blood 117: 5057-5066.

66. Chuikov S, Levi BP, Smith ML, Morrison SJ (2010) Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. Nat Cell Biol 12: 999-1006.

67. Oram SH, Thoms JA, Pridans C, Janes ME, Kinston SJ, et al. (2010) A previously unrecognized promoter of LMO2 forms part of a transcriptional regulatory circuit mediating LMO2 expression in a subset of T-acute lymphoblastic leukaemia patients. Oncogene 29: 5796-5808.

68. Gratzinger D, Zhao S, West R, Rouse RV, Vogel H, et al. (2009) The transcription factor LMO2 is a robust marker of vascular endothelium and vascular neoplasms and selected other entities. Am J Clin Pathol 131: 264-278.

69. Appelt JU, Giordano FA, Ecker M, Roeder I, Grund N, et al. (2009) QuickMap: a public tool for large-scale gene therapy vector insertion site mapping and analysis. Gene Ther 16: 885-893.

70. Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, et al. (2008) The Catalogue of Somatic Mutations in Cancer (COSMIC). Curr Protoc Hum Genet Chapter 10: Unit 10 11.

71. Grund N, Maier P, Giordano FA, Appelt JU, Zucknick M, et al. (2010) Analysis of self-inactivating lentiviral vector integration sites and flanking gene expression in human peripheral blood progenitor cells after alkylator chemotherapy. Hum Gene Ther 21: 943-956.

72. Kiem HP, Ironside C, Beard BC, Trobridge GD (2010) A retroviral vector common integration site between leupaxin and zinc finger protein 91 (ZFP91) observed in baboon hematopoietic repopulating cells. Exp Hematol 38: 819-822.

73. Liles WC, Broxmeyer HE, Rodger E, Wood B, Hubel K, et al. (2003) Mobilization of hematopoietic progenitor cells in healthy volunteers by AMD3100, a CXCR4 antagonist. Blood 102: 2728-2730.

74. Li HW, Sykes M (2012) Emerging concepts in haematopoietic cell transplantation. Nat Rev Immunol 12: 403-416.

75. Ranzani M, Cesana D, Bartholomae CC, Sanvito F, Pala M, et al. (2013) Lentiviral vector-based insertional mutagenesis identifies genes associated with liver cancer. Nat Methods 10: 155-161.

76. Shiozawa Y, Pedersen EA, Havens AM, Jung Y, Mishra A, et al. (2011) Human prostate cancer metastases target the hematopoietic stem cell niche to establish footholds in mouse bone marrow. J Clin Invest 121: 1298-1312.