

Review

Retroviral Vectors: Post Entry Events and Genomic Alterations

Ali Nowrouzi ^{1,2}, Hanno Glimm ^{1,2}, Christof von Kalle ^{1,2} * and Manfred Schmidt ^{1,2} *

¹ Department of Translational Oncology, German Cancer Research Center, Im Neuenheimer Feld 581, 69120 Heidelberg, Germany; E-Mail: ali.nowrouzi@nct-heidelberg.de (A.N.)

² National Center for Tumor Diseases, Im Neuenheimer Feld 581, 69120 Heidelberg, Germany

* Authors to whom correspondence should be addressed;

E-Mails: Christof.kalle@nct-heidelberg.de (C.v.K.); Manfred.schmidt@nct-heidelberg.de (M.S.);

Tel.: +49-6221-56-6991; +49-6221-42-1600; Fax: +49-6221-56-6930; +49-6221-42-1611.

Received: 9 February 2011; in revised form: 24 March 2011 / Accepted: 5 April 2011 /

Published: 29 April 2011

Abstract: The curative potential of retroviral vectors for somatic gene therapy has been demonstrated impressively in several clinical trials leading to sustained long-term correction of the underlying genetic defect. Preclinical studies and clinical monitoring of gene modified hematopoietic stem and progenitor cells in patients have shown that biologically relevant vector induced side effects, ranging from *in vitro* immortalization to clonal dominance and oncogenesis *in vivo*, accompany therapeutic efficiency of integrating retroviral gene transfer systems. Most importantly, it has been demonstrated that the genotoxic potential is not identical among all retroviral vector systems designed for clinical application. Large scale viral integration site determination has uncovered significant differences in the target site selection of retrovirus subfamilies influencing the propensity for inducing genetic alterations in the host genome. In this review we will summarize recent insights gained on the mechanisms of insertional mutagenesis based on intrinsic target site selection of different retrovirus families. We will also discuss examples of side effects occurring in ongoing human gene therapy trials and future perspectives in the field.

Keywords: retroviral vectors; retroviral integration; gene therapy; insertional mutagenesis; vector genotoxicity; Leukemia; common insertion sites

1. Introduction

The defining feature of retroviral replication is the integration of the reverse transcribed viral DNA into the genome of the host cell [1,2]. Efficient integration of the viral DNA into the host genome is a hallmark of the retroviral life cycle making replication-incompetent retroviral vectors attractive gene transfer vehicles for stable ectopic expression of transgenes in target cells. The understanding of basic principles of the retroviral replication cycle has led to the development of replication incompetent retroviral vectors capable of a single integration event upon infection of target cells in the absence of superinfection [3–6]. Since their first description in the beginning of the 1980s, a variety of vector systems from the two different retroviral subfamilies—*Orthoretrovirinae* and *Spumaretrovirinae*—have been developed, which are broadly applied in basic and clinical research as ectopic gene delivery vehicles [7,8]. The viral integration reaction catalyzed by the viral integrase protein has been extensively analyzed revealing complex interactions with cellular host proteins regulating nuclear import, chromatin tethering and integration into the host genome [9,10].

Being applied in >357 initiated clinical phase I/II gene therapy trials, retroviral based vectors represent the second most commonly used gene delivery vehicles after adenoviral vectors [11]. Stable integration into the host genome and subsequent long-term ectopic expression of therapeutic transgenes underlines the potential of retroviral gene transfer systems for correction of inherited diseases [12].

Gene therapy for inherited diseases has demonstrated that modification and retransplantation of gene corrected (stem) cells cures severe disorders. The success of these gene therapy trials has been accomplished by replacing genetically non-functional genes present in the patients' cells with retroviral vectors constitutively expressing the therapeutical gene after random integration into the host genome. Consequently, the retroviral integration sites (RIS) create unique genetic signatures which can be amplified and sequenced to follow the fate of individual retrovirally 'molecular marked' cells and their clonal progeny in the respective target tissue [13]. Development of technologies for identification and sequencing of retroviral integration sites together with the publication of the human genome sequence have led to precise insights in the global integration pattern of different retrovirus subfamilies [14–20]. These integration site profiling studies uncovered unexpected virus specific integration preferences most likely resulting from different interactions of the viral integration machinery and host factors [21]. Comprehensive studies of retrovirally transduced cells and their fate *in vivo* by large scale integration site analyses of preclinical and clinical samples have provided evidence that the propensity for insertional mutagenesis is, at least in part, influenced by the non-random integration site selection of retroviral vectors [22].

In this chapter we will focus on recent insights in the target site selection of retroviral vectors and the molecular mechanisms underlying retroviral vector induced mutagenesis. An overview of subsequent biological effects of insertional mutagenesis is given based on preclinical and clinical data. We will introduce recent advances in next generation sequencing technologies and their impact on future high-throughput integration site analyses, both for mutation and vector biosafety research, and highlight their potential for a comprehensive clinical monitoring of current and future stem cell gene therapy trials using retroviral based vectors.

2. Trafficking to the Nucleus and the Retroviral Integration Reaction

Pioneering studies identifying the reverse transcriptase as a key component of retroviruses responsible for the conversion of the viral RNA genome into viral DNA [23,24] supported the provirus/protovirus [25,26] hypotheses by Howard Temin and led to the discovery of integrated vector genomes in many different organisms [27–32]. After penetrating the cell membrane, the viral nucleoprotein core particle containing two copies of viral genomic RNA is delivered into the cytoplasm where uncoating of the viral capsid takes place and reverse transcription is initiated [9]. Except for spumaviruses, which seem to have a unique mechanism to reversely transcribe their RNA genome late in the replication cycle [33,34], all other retroviruses initiate reverse transcription directly after endocytosis at the cell membrane [35]. The newly synthesized viral DNA remains in a large nucleoprotein complex called the pre-integration complex (PIC) and is associated with viral and cellular proteins [36–38]. The PIC interacts with the microtubule (MT) network that regulates intracellular trafficking to the nucleus [39–41]. Once at the nuclear membrane, gammaretrovirus based vectors such as MoMLV require the breakdown of the nuclear envelope during cell division to enter the nucleus [42]. In contrast, the PIC components of lentivirus based vectors [43] interact with the nuclear pore complexes allowing entry across the intact nuclear envelope [9]. Similar to yeast retrotransposons whose integration complexes interact with host proteins, interaction of the retroviral PIC with cellular proteins and their karyophilic properties support the tethering of the viral DNA genome to chromatin where the integration reaction is initiated [9,44].

The basic molecular mechanism of retroviral integration has been demonstrated in biochemical studies *in vitro* [45] and has been further elucidated by recent structural work [46,47]. The first steps in the integration reaction are catalyzed by the viral integrase and are initiated by the 3'-prime processing reaction which results in the removal of two nucleotides from each 3'-prime end of the viral DNA [45,48,49]. The exposed 3'-prime hydroxyl groups at each end of the viral DNA are joined to the target DNA and subsequent strand transfer reaction of a pair of processed viral DNA ends leads to a concerted insertion of the viral DNA into the host genome [45,48–50]. The sites of strand transfer on the two target strands are separated by 4–5 base pairs. Repair of this integration intermediate by cellular components [51] results in a direct duplication of 4–5 base pairs flanking the integrated viral DNA [2]. Most DNA sequences can act as integration acceptor sites, however, recent large scale studies on the integration site consensus of retroviruses *in vivo* [15,52,53] and *in vitro* [46,47] have shown that the base composition at retroviral target sites is biased for preferences or avoidances of particular bases supporting physical effects of the primary sequence on chromatin and the integration reaction, respectively.

2.1. Distribution of Retroviral Integration Sites in the Cellular Genome *in vitro* and *in vivo*

Since the discovery that integration is an essential step in the replication cycle of retroviruses, attempts were made to isolate proviruses and map their genomic location [2]. From *in vitro* studies using the purified integrase and recombinant chromatin it was suggested that nucleosome positioning influences the integration reaction [54–57]. Early studies on integration targeting of MLV in cultured cells proposed that integration was favored near DNaseI hypersensitive sites or transcribed regions [58–61]. However, due to the low number of individual insertion sites analyzed and the

formerly unknown human genome sequence, target site selection of retroviruses across the whole genome was considered to be random.

The target site selection of retroviruses and retroviral vectors has gained novel scientific interest, since severe adverse events during the gene therapy trial for X-SCID [62] were linked to the integration of the therapeutic MLV based vector in vicinity of the LMO2 proto-oncogene [63,64]. Activating insertional mutagenesis of LMO2 in combination with acquired somatic mutations was responsible for the development of leukemia in a minority of the treated patients [65,66] and raised serious concerns whether the risk for insertional side effects is dependent on the target site selection and similar across all retroviral based vector systems [22].

With the decoding of the human genome and the development of PCR methodologies for allowing the amplification and sequencing of the genomic junctions between viral and host DNA the exact genomic location of a potentially unrestricted number of viral integrants generated by the acute infection of cultured cells with retroviruses or retroviral vectors became feasible. Analyzing their position in the human genome has uncovered virus specific integration patterns (Table 1) distinguishing most retroviral subfamilies from each other (Figure 1). The first large scale comparative analyses of integration target site selection in the human genome was performed with MLV and HIV-1 based vectors in a human cell line using LM-PCR and subsequent sequencing of the vector cellular junctions [19]. Compelling data demonstrated that RIS from MLV and HIV-1 based vectors showed distinct differences in respect to annotated features of the human genome. MLV vectors revealed a significant clustering around the transcription start site (TSS) of RefSeq genes and in the proximity to CpG-islands whereas HIV-1 based vectors disfavored TSS and CpG islands, but showed preferences for integration inside transcribed regions of RefSeq genes [19]. The preference for the integration of HIV-1 inside RefSeq genes was previously shown in a human lymphoid cell line [20] and also detectable with other vector systems from the lentiviral family pointing to a conserved mechanism of target site selection of lentiviruses [16,67,68]. These insights indicate that a mechanism involving the viral integrase and cellular genes is responsible for tethering the integration complex into particular regions of the genome. However, the investigated integration pattern of avian sarcoma and leukemia virus (ASLV) and prototype foamy virus (PFV) based vectors in human cells showed a rather random like distribution in relation to genes TSS and CpG-islands indicating that integration site selection of these retroviral families is not influenced by particular chromosomal structures and occurs rather by chance [14,15,18]. Based on changed integration patterns of HIV-1 hybrid vectors carrying the integrase gene of MLV it has been suggested that the principal viral determinant of retroviral integration specificity is the viral integrase itself [69]. The interaction of the HIV-1 integration complex with the cellular protein LEDGF/p75 plays a critical role in HIV-1 replication [70,71], protecting the viral integrase from proteosomal degradation [72] and increasing the affinity of the integrase to chromatin [73,74]. The siRNA mediated down regulation of LEDGF/p75 has been shown to influence the normal target site selection of HIV-1 based vectors decreasing its propensity to integrate into transcription units [75–78]. Very recently a comprehensive integration site analyses of HIV-1 based vectors in rodent eye and brain tissue revealed a decreasing preference for integration into transcription units correlating with a cell specific lowered expression level of LEDGF/p75 [79]. More intense investigations of different retroviruses and the functional analyses of cellular proteins interacting with their integration machinery will shed more light on the mechanism that retroviruses

have evolved to maintain host cell genome integrity in accordance with virus specific replication and pathogenesis. The understanding of biochemical principles controlling the tethering of the retroviral integration complex to chromatin together with recent findings showing that such knowledge can be applied to redirect the integration pattern of HIV-1 based vectors are promising for the future development of therapeutic vectors which are designed to integrate into potentially safer sites of the genome [80–83].

Table 1. Integration pattern of different retroviral vectors.

Virus	Genes	TSS	CpG-Rich Islands	Reference
HIV	+	–	–	[19,20]
SIV	+	–	–	[67]
MLV	±	+	+	[18,19,67]
ALV	±	–	–	[17]
ASLV	±	–	–	[18]
PFV	–	±	±	[14,15]
EIAV	+	–	±	[16,68]
HTLV-1	±	–	–	[84,85]

(+) strong preference; (–) no preference; (±) weak preferences.

The correlation of target site selection of HIV-1 based vectors to gene expression profiling data of virus infected cells has revealed that transcriptional activity of chromatin influences the genome accessibility of the retroviral integration complex [18,86]. Genes showing transcriptional activity have a higher propensity to be targeted by the HIV-1 integration complex [18]. Most likely, the transcriptional activity of chromosomal regions also plays a role in the expression of viral genes for wild-type retroviruses as well as transgene expression from retroviral vectors. Influences on the transgene expression level based on chromosomal regions have been demonstrated for HIV-1 based vectors with higher transgene expression from integrated vectors to be located on more active open chromatin [87]. From the evolutionary standpoint and pathogenesis of retroviral families it is interesting to note that ASV compared to HIV-1 seems to favor integration into more condensed chromatin *in vitro* similar to properties found at heterochromatin [88]; however, it remains elusive what mechanism controls these differences and if such differences might explain the different pathology of these retroviruses.

Recently, a genome wide analysis of >4000 transcription factor binding sites (TFBS) in the vicinity of MLV integration sites identified TFBS as differential genomic determinants of retroviral target site selection in the human genome. Gamma-retroviral vectors integrate in genomic regions enriched in cell-type specific subsets of TFBS suggesting that interaction of transcription factors with the viral LTR enhancer may synergize with the integrase in tethering retroviral pre-integration complexes to transcriptionally active regulatory regions [89]. A comparative analysis of gamma-retroviral integration sites between infused gene corrected mature lymphocytes (peripheral blood lymphocytes, PBL) and single infusion of hematopoietic stem/progenitor cells (HSC) of patients from the adenosine deaminase-severe combined immunodeficiency (ADA-SCID) clinical trial have revealed that the insertional profile of MLV based vectors is cell-specific according to the genetic/chromatin state of the

target cell [90,91] indicating that retroviruses have developed different strategies to interact with the host chromatin which seem to influence integration site selection substantially [89,90].

Figure 1. Integration site selection of retroviral based vectors in the cellular genome. **(A)** Analysis of individual vectors to target sites upstream and downstream of the transcription start site (TSS) at a 2.5 kb window size relative to the random control, which was set arbitrarily to 1 show that murine leukemia virus (MLV) based vectors have the strongest preference to integrate in close proximity to TSS. Prototype foamy virus (PFV) based vectors integrate at a rate approximately three- to four-fold higher than the expected random value, while HIV-based vectors avoid these regions; **(B)** The preference to integrate into transcribed regions of genes is strongest for HIV based vectors followed by MLV vectors, which show weaker preferences, while PFV vectors do not show any preferences for target site selection in genes. Their integration preference towards genes is similar to what would be expected if target site selection would be random; **(C)** Frequency of retroviral vectors to integrate in the vicinity of CpG islands matched to the random control, which was set arbitrarily to 1 is dramatically increased with MLV based vectors. HIV and PFV vectors do not show any strong preferences for integration near to CpG islands (modified from Nowrouzi *et al.* [11]).

A.

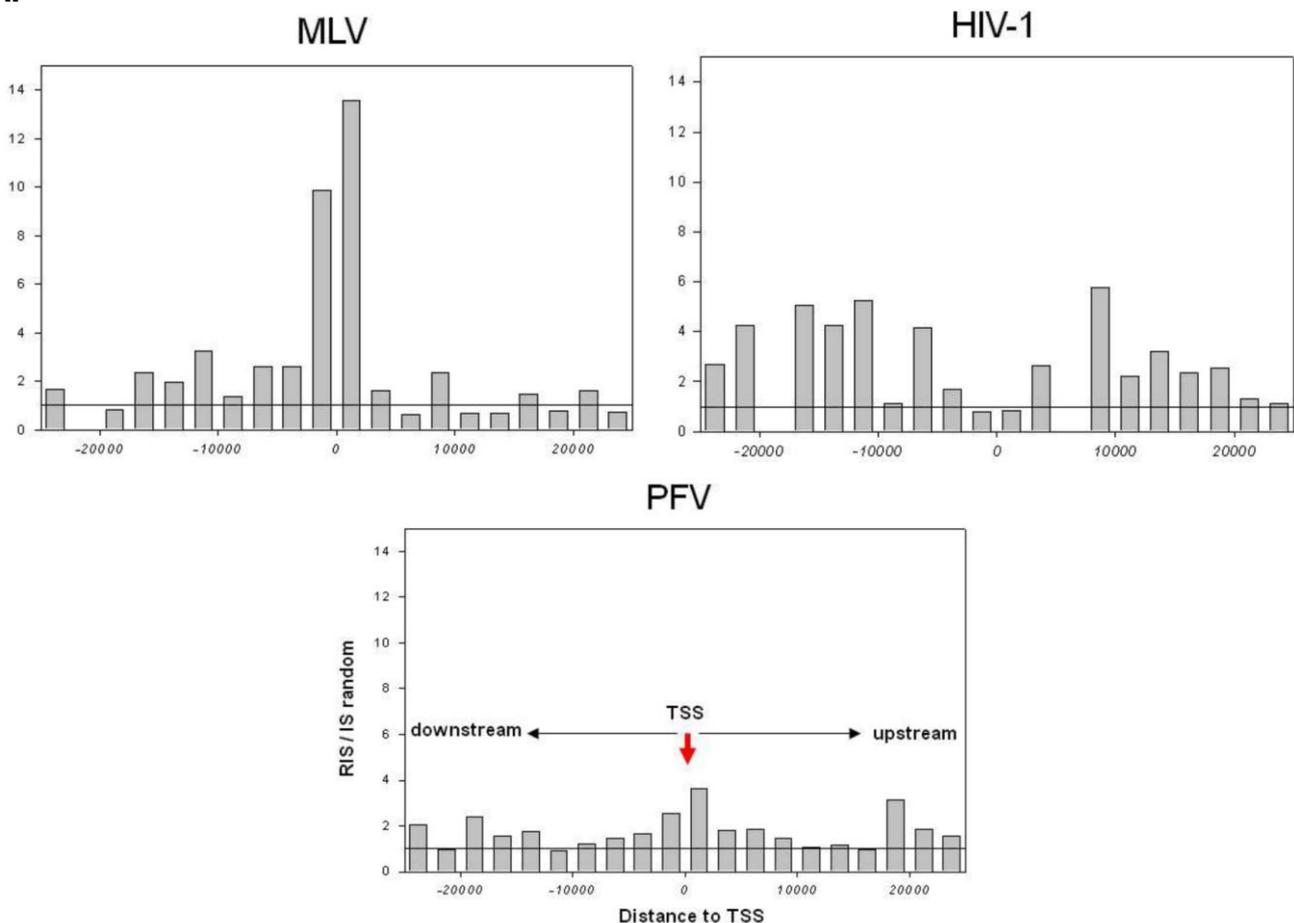
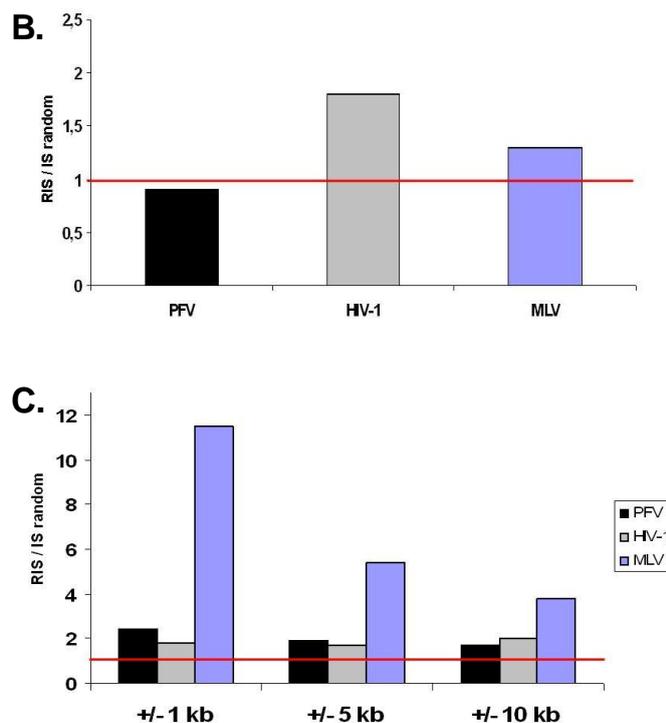


Figure 1. Cont.



To what degree does the different integration pattern of retroviruses and their vector derivatives influence the likelihood for insertional mutagenesis? It is generally recognized that known preferences of gammaretroviral vectors to target regulatory regions of genes increase the risk for genetic alterations of neighboring genes [22,92]. In contrast, it has been shown that independent of the integration pattern also vector design and dose influence the likelihood of vector induced cellular transformation [93,94]. Virus specific transcriptional enhancers within the LTR have been identified as one of the major determinants of genotoxicity independent of retrovirus subtypes [93,94]. The deletion of enhancer/promoter elements within the retroviral LTR, termed self inactivating (SIN) LTR, significantly decrease the propensity to cellular transformation demonstrated in an *in vitro* immortalization assay [95]. However, at the same time it was elegantly shown in a tumor prone mouse model [96] that there is a significant higher load needed with LV vectors containing gammaretroviral enhancer elements to trigger oncogenesis when compared to MLV-based vectors [93,97]. The authors suggest that the higher vector load required for oncogenesis by LV vectors is most likely explainable by a higher preference of MLV-vectors to integrate near cancer promoting genes pointing to a role of intrinsic viral integration patterns inducing genotoxicity. Initial studies using PFV-vectors with rather random like integration patterns with respect to genes and regulatory regions of genes in a canine preclinical model have not resulted in any detectable insertional induced side effects [98], arguing that random and uniform integration is safer in terms of genetic alterations of the host genome.

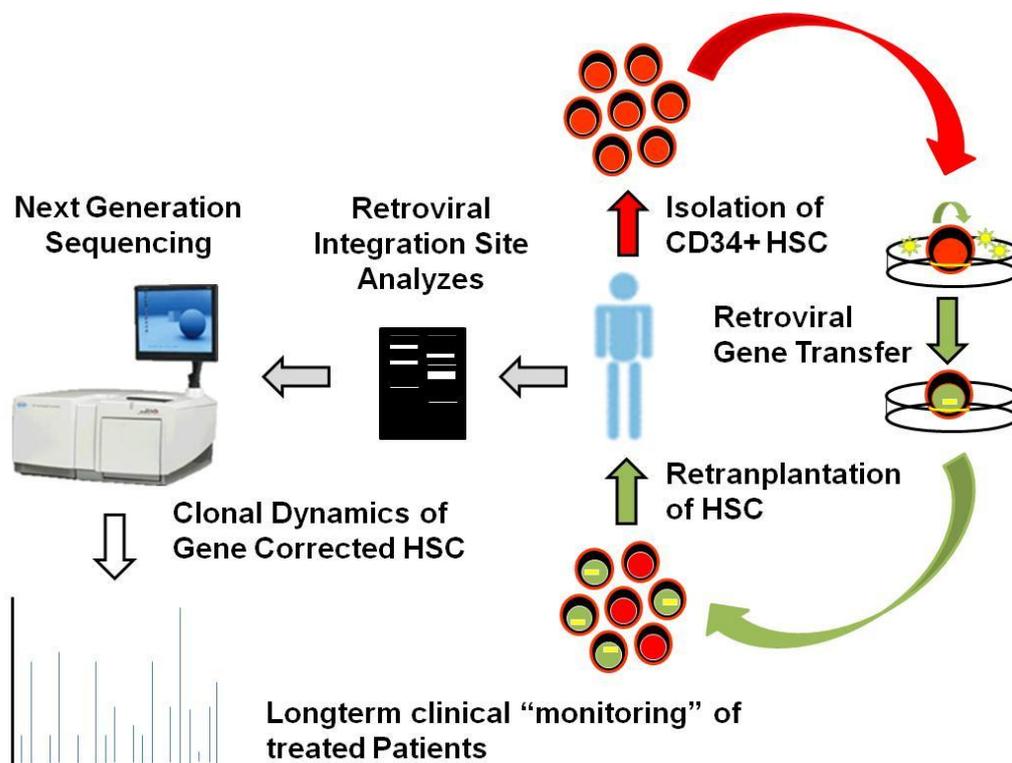
Taken together, many large scale integration site distribution studies have been conducted and contributed to a better understanding of factors which may influence integration site preferences of retroviral vectors. Potential genetically safer vector systems, based on their integration target site selection, have been evaluated by several groups showing that deletion of viral genotoxic elements and

the use of physiological (cellular) promoters significantly reduce the risk for insertional mutagenesis in future gene therapy trials.

2.2. Next Generation Sequencing and Unbiased Retrieval of Vector Integration Sites

For characterization of proviral cellular junctions two fundamental PCR methods were developed: inverse PCR [99] and ligation-mediated PCR [100–102]. Although further development in these methodologies led to improvements in sensitivity, the invention of the linear amplification-mediated PCR (LAM-PCR) [103,104] enabled the characterization of integration sites at the single-cell level for the first time, allowing monitoring of retrovirally gene-modified hematopoiesis directly in limited amounts of peripheral blood leukocytes and bone marrow cells [105]. Combined with next generation sequencing platforms and strategies for saturated genomic access [106–108] LAM-PCR has proven to be highly efficient for retrieval of whole insertional inventories in clinical and preclinical samples [109] (Figure 2) and analyzing the clonality of the hematopoietic repopulation after transplantation in humans [66,110–114].

Figure 2. Analyzing clonal dynamics and insertional induced side effects by retroviral integration site analyzes and next generation sequencing.



The principle of LAM-PCR has been described previously in detail [103]. The first step is the pre-amplification of the vector-genome junctions by a linear amplification step with biotinylated primers hybridizing at one end of the integrated vector. The following steps are carried out on a semisolid streptavidin phase in order to capture DNA strands with an incorporated biotinylated vector primer. After double strand synthesis, a restriction digest, and the ligation of a linker cassette on the genomic end of the fragment, two exponential PCRs with nested arranged vector- and linker cassette

primers are carried out in order to amplify the fragments consisting of linker cassette-, genomic-, and vector-sequence. The generated fragments are sequenced and the integration loci are determined. Application of novel next-generation sequencing technologies for integration site analysis by several groups have led to a marked increase in the retrieval of integration site information. We have recently applied 454 pyrosequencing and annotation of restriction motives in the human genome to define the genomic accessibility to viral integration sites and developed a new modified protocol for unbiased retrieval of insertion sites based on the non-restrictive LAM-PCR (nrLAM-PCR) [106,107]. The nrLAM-PCR allows the genome-wide identification of retroviral integration sites in a single reaction, circumventing the detection bias accompanied by methods dependent on restriction enzymes [106,107]. Multiplex barcoding [115] of samples together with downstream bioinformatical evaluation of up to 1 million 400 bp long reads in one sequencing run enables analyzing over hundred thousand RIS in multiple samples at the same time [107]. This is efficient to analyze whole insertional inventories of patients, clonal dynamics and integration patterns of any kind of integrating genetic elements in parallel. Most importantly, LAM-PCR or nrLAM-PCR combined with next generation sequencing platforms (*i.e.*, 454, Solexa, ABI) allow a semi quantitative retrieval of insertion sites representing unique molecular identities of particular marked clones in a given sample, highly valuable to monitor clonal dynamics in hematopoiesis and many other biological systems [110,114,116–118]. Several commercial sequencing platforms are available, each having its advantages and disadvantages. The platforms from Solexa (Genome Analyzer) and Applied Biosystems (SOLID) generate very high numbers of sequence reads, but are limited in individual sequence length. The pyrosequencing platform from 454 Life Sciences (Titanium system; Roche Diagnostics) allows the generation of up to 5×10^5 sequence reads of ~400 base pairs (bp) in length in a single sequencing run sufficient to chromosomal map amplicons with vector cellular junctions without using complicated algorithms to process short sequence reads.

3. Side Effects in Clinical and Preclinical Gene Therapy Studies

First molecular insights in the oncogenic potential of integrated wild type retroviruses were obtained from bursal lymphomas in chicken. The majority of the identified tumor cells contained a provirus integrated in the vicinity of the proto-oncogene *c-myc* that was overexpressed by the viral promoters/enhancers [119–122]. These genetic alterations caused by retroviruses, termed proviral insertional mutagenesis, have been identified in many types of retrovirus induced tumors [123,124]. Thereby, a variety of genes could be identified that modulate growth and differentiation and significantly contributed to tumor formation. Despite these insights obtained from wild-type retroviruses and replication-competent retroviral vectors, the risk of insertional mutagenesis (Figure 3) and cellular transformation with replication-incompetent vectors specifically developed for clinical purposes was considered to be rather low [125].

However, after first cases of severe adverse events in a minority of patients in the X-SCID gene therapy trial were reported (see next chapter), the first malignant transformation that developed as a result of gene transfer using replication-deficient retroviral vectors designed for clinical use in a murine model was observed in 2002 [126]. In a murine model system the serial transplantation of bone marrow cells marked with gammaretroviral vectors led to tumor development in secondary and tertiary

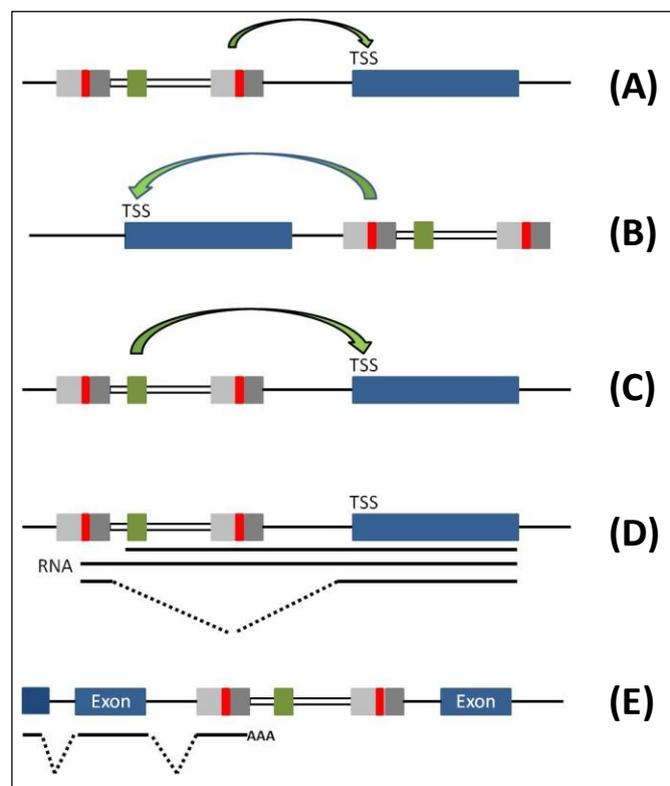
recipients [126]. All secondary transplanted animals showed alterations in their hematopoiesis after 22 weeks, and six out of ten animals developed acute myeloid leukemia. Molecular analyses of the malignant clone in respect to the RIS using LM-PCR revealed that in all developed tumors derived from secondary and tertiary recipients an insertion in the *Evi1* gene locus triggering malignant transformation was detected [126]. Although it was shown that the promoter and enhancer elements in the proviral LTR caused an overexpression of *Evi1*, a synergistic effect of the low-affinity nerve growth factor receptor (LNGFR) remains unclear [126]. In terms of clinical safety it is important to note that vector dose has an impact on leukemogenesis as shown in a murine study where Leukemia not only occurred due to the growth advantage of single clones but also correlated with high vector doses using MLV based vectors [127]. Soon after, gene marking studies in a non-human primate model provided evidence for clonal dominance due to an insertional effect of MLV based vectors in vicinity of the *Mds1/Evi* or *Prdm16* gene. This insertional event led to a growth advantage compared to other marked cell clones without causing malignant transformation [128]. Since the majority of these cell clones contribute to long-term hematopoiesis, a higher engraftment or survival probability as a result of insertional mutagenesis was discussed [22,128]. Vector induced side effects with possible roles in hematopoietic activity have also been reported in a murine gene marking study in which proviral integration within or nearby particular cellular genes promoted the growth of single transduced cells and contributed to their clonal expansion *in vivo* [129]. The first retroviral vector induced acute myeloid leukemia in a non-human primate model was described after retroviral gene transfer in hematopoietic precursor cells. The treated animal died five years after gene therapy due to a myeloid sarcoma caused by two insertions in the *Bcl2-A1* and *Cdw91* genes. In this study the clone harboring these two insertions was both detectable in the blood where it became dominant one year after transplantation as well as in the tumor, strengthening a cooperative functional role of these insertions in tumor development [130].

Human gene therapy using retroviral vectors is a specialized form of therapy that is mainly applied to patients for whom no therapeutic alternatives are available. Therefore, the benefit of gene therapy always has to be opposed to its potential risks. The field has suffered from a lot of initial hype without taking into account that any new kind of specialized therapy may also be accompanied by side effects. The success of human gene therapy for hematopoietic diseases has been enabled by the greater efficiency in performing *ex vivo* transduction of human CD34⁺ cells using retroviral vectors and reinfusion of a high number (up to 10⁸–10⁹) of gene-corrected cells. In retrospect, it is now evident that improvement in effectiveness for curing otherwise challenging lethal diseases is accompanied by a higher risk of biologically relevant side effects of this type of gene therapy. Up until now most gene therapy trials in the hematopoietic system involve the use of first generation MLV derived retroviral vectors. Classical MLV based vectors carrying the full LTR are now based on their integration site selection and vector design considered to have a higher potential for genotoxic events as compared to lentiviral based vectors (see below).

Monitoring the *in vivo* fate of gene corrected CD34⁺ cells retransplanted into patients by high-throughput integration site analysis in five independent clinical gene therapy studies have shown that distribution of gammeretroviral vectors *in vivo* is skewed. Apart from subtle effects, clonal dominance and leukemogenesis influencing the growth and differentiation of clones when inserted near to particular genes and loci have been reported in minority of treated patients [64–66,111–114,131]. We

have recently described that in addition to overt carcinogenesis, even as a single vector copy, many insertion locations more subtly influence the biological fate of a cell clone *in vivo* [111,112,123]. With LAM-PCR technology specifically developed for this purpose, it has been identified that up to 40% of circulating cells carry insertions in a rather small set of frequently affected common insertion sites (CIS) [111,112]. Such CIS are almost always in the direct vicinity of known and novel genes likely to be involved in cellular growth, survival and self-renewal processes of immature progenitor and stem cells.

Figure 3. Retroviral vectors may induce mutations in multiple ways by integration of the retrovirus in the host genome. **(A, B)** Mutagenic proviral insertions in most reported cases induce an activation of neighboring genes by enhancer elements present within the wildtype LTR. Such “enhancer insertions” can induce gene activation from distances up to 100 kb. **(C, D)** In case of SIN-type retroviral vectors strong internal promoters driving transgene expression may induce deregulation of genes in close proximity similar to so-called “promotor insertions” which result in viral-host gene-fusion transcripts. **(E)** Genotoxic side effects resulting from retroviral integration sites leading to inactivation of cellular genes may be induced by viral insertion within a host gene leading to truncated non functional transcripts.



In the gene therapy trial for X-SCID, two to six years after successful correction of the otherwise lethal disease the development of T-cell leukemia in five of 19 patients has been observed [65,66]. In four of these patients the leukemic clone harbored a vector integration in the *LMO2* proto-oncogene inducing its overexpression. Together with acquired mutations these events have led to expansion of these clones and the development of leukemia [65,66]. Although under dispute, the combinatorial effect of the *IL2RG* transgene in the development of the lymphoproliferative disease in the X-SCID

trial has been discussed [132,133] and recent work from Copeland and colleagues provides supportive evidence that *IL2RG* and *LMO2* do cooperate in leukemia induction [134]. At the same time skewing of RIS was observed in the healthy patients causing subtle insertional effects without any signs of malignant transformation [111,112].

Evidence for vector induced effects influencing hematopoietic activity have been gained in the gene therapy for the correction of X-CGD [113,117]. Here, in two adult patients, retroviral gene therapy resulted in the restoration of oxidative antimicrobial activity in phagocytes after gene transfer. Integration site analysis revealed a clonal dominance triggered by insertion sites in the genes *MDS1-EVII*, *PRDM16* or *SETBP1* (Figure 4A), resulting in an expansion of gene corrected myelopoiesis [113,117]. In the follow up of the clinical trial a substantial gene transfer in neutrophil cells had produced a high number of functional phagocytes, however, after the initial resolution of bacterial and fungal infections, both subjects showed silencing of transgene expression due to methylation of the viral promoter, and myelodysplasia with monosomy 7 as a result of insertional activation of ecotropic viral integration site 1 (*EVII*). It has been recently suggested that the overexpression of *Evi1* disrupts normal centrosome duplication leading to genomic instability, monosomy 7 and clonal progression toward myelodysplasia [117].

From these severe adverse events it is tempting to reason that the use of gammaretroviral vectors is *per se* oncogenic. However, given that severe events have occurred in a minority of treated patients in the X-SCID trial, the clinical follow up of the ADA-SCID [131] trial using the same vector backbone as in the X-SCID trial has up to now not been accompanied by any adverse events up to 8 years post therapy. This may indicate that the risk for vector induced side effects may also be dependent on the disease and the clinical protocol. We note that clear evidence for this discussion has not been reported so far.

In the most recent gene therapy trial for Wiskott-Aldrich Syndrome (WAS) an MLV-based vector backbone expressing the WAS protein was used for transduction of autologous CD34⁺ HSC which were transfused into two patients. WAS is a primary immunodeficiency disorder associated with thrombocytopenia, eczema and autoimmunity and is alternatively treatable with haploidentical BM transplantation [114]. Here, monitoring the clonal contribution of gene corrected cells to hematopoietic regeneration three years post transfusion by a comprehensive integration site analysis using next generation sequencing technologies has identified over 10,000 unique clones contributing to short- and long-term hematopoieses. This genome-wide insertion site analysis demonstrated that vector integration targeted multiple genes controlling growth development and immunological responses in a persistently polyclonal hematopoiesis. However, many of the previously observed CIS, which occasionally triggered adverse events, were detectable in both subjects to a similar degree suggesting vector induced skewing similar to what was observed in previous gene therapy trials using the MLV backbones (Figure 4B, C). In fact, it was further shown that in sorted cell populations of the lymphoid or myeloid fraction lineage specific CIS could be identified, which in previous gene therapy trials had triggered either lymphoid or myeloid proliferation suggesting that insertional activation of these genes programs cell fate in the hematopoietic system [114]. In total, 9/10 patients have been treated successfully in this trial. Very recently it has been reported that one patient has developed a T-cell leukemia similar to the patients in the X-SCID trial. Whether vector induced effects have played a role

First proof that switching the integration pattern of the therapeutical retroviral vector improves genotoxic safety has been provided by the first LV-based gene therapy trial for the treatment of the cerebral form of X-chromosomal linked adrenoleukodystrophy (X-ALD), a demyelinating disease of the central nervous system caused by mutations in the *ABCD1* gene [110]. Progress of X-ALD is treatable by haploidentical BM transplantation, but with patients where no matching donor is found gene therapy is a highly promising therapeutic option [135]. In 2006, two patients with no sibling BM donors were treated with HIV-1 based vectors, which show distinct differences in target site selection and efficiency in transducing quiescent cells compared to MLV-based vectors. Mobilized hematopoietic precursor cells were harvested from two seven year-old patients, transduced *ex vivo* with lentiviral SIN-vectors encoding the *ABCD1* transgene and reinfused into the patients. This treatment proved efficient to arrest the progression of the disease in both patients through the constitutive expression of the functional *ABCD1* therapeutic transgene. More importantly, no signs of vector induced side effects have been reported so far in two treated patients continuously being monitored. Up to now, the clonal dynamics of transduced and reinfused CD34⁺ HSC show transduction of multipotent HSC and no signs of enrichment near to CIS genes previously being identified as integration hot spots of MLV-based vectors *in vivo* [110]. The improved clinical safety of SIN-type LV-based vectors compared to classical MLV-based vectors has previously been proposed based on evidence gained in sensitive mouse models designed to measure genotoxic safety [93,136]. Deduced with murine genotoxicity assays, although at low rate, vector induced genetic alterations have also been reported with target site selection of LV-based vectors [93,94,137]. Such insertional alterations in the regulation of the cellular genome by LV-vectors have recently been reported in the gene therapy trial for human β -thalassaemia [118]. Clinical benefit of lentiviral β -globin gene transfer has been achieved in an adult patient 33 months post treatment by clonal dominance initiated by vector induced activation of the *HMGGA2* gene. This particular clonal dominance has been proposed to accompany clinical efficacy for β -thalassaemia, a challenging hematopoietic disease, and the most common form of severe thalassaemia in southeast Asian countries and their diasporas [118].

New Strategies for Vector Biosafety in Gene Therapy

The risk of side effects in future gene therapy trials will be dependent on the choice of the most suitable gene transfer vector dependent on the individual disease and the propensity for genomic alterations driven by vector design and integration pattern. The usage of SIN-type vectors with weak or tissue-specific promoters or non-integrating vector systems—if appropriate—are considered to be safer than classical MLV-vectors. The use of AAV- or Adenovirus-vectors predominantly persisting episomally in post-mitotic tissue (*i.e.*, retina, brain, muscle or liver) are promising for particular diseases [138]. However, insertional mutagenesis caused by rare but detectable integration events with AAV-vectors in the liver has also been reported [139,140] underlining the need for highly sensitive strategies to detect such rare insertions in a clinical setting. In postmitotic tissues, dilution of unintegrated episomal vector forms does not occur thus allowing the use of integrase deficient lentiviral vector systems [141]. Mutations in the core domain of the viral integrase prevent integration, thereby reducing the risk of insertional mutagenesis. Recently, the long-term functional correction of

retinal degeneration in a well-established rodent model for ocular gene therapy was reported under the use of an integrase-deficient HIV-1 vector without apparent side effects [142].

Alternative modification of the viral integrase for targeted integration into desired potentially safe chromosomal regions [143], together with significant improvements in the efficiency of homologous recombination or gene disruption using novel Zinc-Finger-Nucleases (ZFN) will in future define alternative gene delivery approaches [144]. Other non-viral integrating vector systems based on sleeping beauty or piggybac transposons have been identified to have a potentially safer integration pattern *in vitro* [145]. However, given that any integrating vector system by chance, if combined with genotoxic vector elements, may induce genetic alterations of the cellular genome also account for transposon based vectors which are also used to identify novel cancer genes in murine models [146].

With significant improvements in the controlled genetic modification of particular loci in targeted integration approaches and a variety of vectors with alternative integration patterns available it becomes evident that the advance of highly sensitive technologies for a comprehensive genomic screening is required to monitor safety on a genetic level. Whole genome next generation sequencing technologies to evaluate genomic stability and strategies for an unbiased retrieval of integrated sites [106] and induced DNA double-strand breaks will be essential for dissecting the genetic specificity and safety of novel targeted gene editing protocols. In this context, recent pioneering advances in the field of cellular reprogramming and the availability of sources of induced pluripotent cells (iPS) [147–149] or lineage specific progenitor cells [150–152] generated by integrating vector systems expressing particular pluripotency of lineage regulating transcription factors have made prediagnostic genomic safety screening of hundreds of clones feasible for potential future cell based therapies [153]. New vector systems and the potential use of new induced clinical applicable cell sources will greatly benefit from vector biosafety models established in the gene therapy field [154,155].

During the last years several *in vitro* and *in vivo* systems have been developed assessing the genotoxic potential of different integrating vectors. Transplantation of gammaretroviral transduced hematopoietic cells in mice with knocked-out *Cdkn2a* tumor suppressor gene leads to accelerated tumor growth in case a cellular proto-oncogene is virally activated [96,156]. Such conducted insertional mutagenesis screens have resulted in the establishment of the most extensive database of murine genes with oncogenic potential, the so-called mouse “*Retrovirally Tagged Cancer Gene Database*” (RTCGD) [157]. Recently also sleeping beauty transposon systems have been applied in a similar manner [158,159]. Du and colleagues showed that identification of protooncogenes is feasible by retroviral gene transfer in cell culture. Transduction of murine bone marrow cells with replication deficient retroviruses expressing marker genes resulted in immortalized cell lines, many of which contained integrations in the *Mds1/Evi1* and *Prdm16* genes [160,161]. Vector integration into the *Evi1* locus and overexpression of *Evi1* appeared to be sufficient for a cell to be immortalized [161,162]. Modlich and colleagues adapted and improved this method so that the results could be analyzed in a quantitative manner [95].

The tumor prone mouse model in which the tumor suppressor gene *Arf* as well as the *Il2rg* gene had been knocked-out originally developed by Lund *et al.* was suitable to perform comparative tests on the genotoxic potential of different gene transfer vectors [96,97]. Mice that had received hematopoietic precursor cell transplants transduced with gammaretroviral vectors developed tumors much more rapidly than mice that had received cells harboring lentiviral SIN-vectors with an internal human

phosphoglycerate kinase (hPGK) promoter [97]. Later, it was demonstrated that gammaretroviral vectors with SIN-LTRs show a significant reduction in genotoxicity, but when combined with strong (viral) internal promoters still have the ability to induce leukemia in transplanted mice [93]. Given that in follow-up of the first LV-vector gene therapy trial for X-ALD no signs of clonal dominance or side effects have been observed up to the present it seems reasonable that such mouse models realistically modulate clinical safety of integrating vector systems. Insertion of insulator elements in the LTRs aims to reduce the influence of the integrated vector on the surrounding host genome and *vice versa* [163,164]. However, evidence for increased genotoxic safety with the use of insulators [165] needs to be further provided in common *in vitro* and *in vivo* models for vector safety assessments, especially after insertional mutagenesis mediated clonal dominance has been reported in the lentiviral β -thalassaemia gene therapy trial in which the cHS4 chromatin insulator was implanted in the LTR of the therapeutic vector [118].

Improvement of genotoxic safety using alternative and modified vector systems together with progress in targeted integration approaches and new cell sorting abilities provide a tremendous potential for future gene and cell therapy with increased safety. This will further improve the treatment of patients with otherwise not curable and often lethal inherited diseases.

4. Conclusion and Future Perspective

Large scale genome wide investigations of retroviral target site selection have shown that different retroviral subfamilies have developed different mechanisms to integrate their DNA into the host genome. Such differences result in three main integration patterns with MLV based vectors showing strong preferences for regulatory regions, LV-based vectors integrating preferentially inside RefSeq genes and PFV as well as ASLV having close to random insertional preferences. Towards safe clinical application of retroviral vectors uncovered target site preferences have proven to play substantial roles in the likelihood for insertional mutagenesis and cellular transformation. Knowledge gained from comprehensive integration site analysis of gene therapy patients and preclinical animal models to assess genotoxic safety of retroviral vectors are currently being actively transferred into the development of new and safer clinical protocols for the treatment of hematological and neurodegenerative diseases. The vision of using vectors with targeted synthetic integrases and improving the frequency of homologous recombination with ZFN or Meganucleases in order to correct defective genes or insert therapeutic genes into potentially “safe harbors” is rapidly evolving to become clinically feasible. In the era of whole genome sequencing projects aiming to understand progress of cancer and other diseases on a genetic level, the treatment of many new diseases in which the pathogenesis is dependent on a gene defect may become treatable with novel safe retroviral vectors imposing minimal genomic side effects for the patients.

Acknowledgements

The authors acknowledge all active and former members of the laboratory as well as collaborators for their contributions. This research has been supported in parts by the Deutsche Forschungsgemeinschaft DFG, grant SCHM 2134/1-1 and 2134/1-2 (SPP1230), by a grant of the Tumor Center Heidelberg/Mannheim, by the Bundesministerium für Bildung und Forschung BMBF, grant 01GU0809 (iGene), by the 6th and 7th Framework Programs of the European Commission,

Contracts LSHB-CT-2004-005242-CONCERT, LSHB-CT-2006-018933-CLINIGENE and HEALTH-F5-2009-222878-PERSIST and by the Initiative and Networking Fund of the Helmholtz Association within the Helmholtz Alliance on Immunotherapy of Cancer.

References and Notes

1. Vogt, P.K. Historical introduction to the general properties of retroviruses. In *Retroviruses*; Coffin, J.M., Hughes, S.H., Varmus, E.H., Eds.; ColdSpring Harbor Laboratory Press: Cold Spring harbor, New York, NY, USA, 1997.
2. Brown, P.O. Integration. In *Retroviruses*; Coffin, J.M., Hughes, S.H., Varmus, E.H., Eds.; ColdSpring Harbor Laboratory Press: Cold Spring harbor, New York, NY, USA, 1997.
3. Eglitis, M.A.; Kantoff, P.; Gilboa, E.; Anderson, W.F. Gene expression in mice after high efficiency retroviral-mediated gene transfer. *Science* **1985**, *230*, 1395–1398.
4. Dick, J.E.; Magli, M.C.; Huszar, D.; Phillips, R.A.; Bernstein, A. Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/Wv mice. *Cell* **1985**, *42*, 71–79.
5. Williams, D.A.; Lemischka, I.R.; Nathan, D.G.; Mulligan, R.C. Introduction of new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature* **1984**, *310*, 476–80.
6. Mann, R.; Mulligan, R.C.; Baltimore, D. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **1983**, *33*, 153–159.
7. Miller, A.D. Development and applications of retroviral vectors. In *Retroviruses*; Coffin, J.M., Hughes, S.H., Varmus, E.H., Eds.; ColdSpring Harbor Laboratory Press: Cold Spring harbor, New York, NY, USA, 1997.
8. Anderson, W.F. Prospects for human gene therapy. *Science* **1984**, *226*, 401–409.
9. Suzuki, Y.; Craigie, R. The road to chromatin - nuclear entry of retroviruses. *Nat. Rev. Microbiol.* **2007**, *5*, 187–196.
10. Turlure, F.; Devroe, E.; Silver, P.A.; Engelman, A. Human cell proteins and human immunodeficiency virus DNA integration. *Front. Biosci.* **2004**, *9*, 3187–3208.
11. Edelstein, M.L.; Abedi, M.R.; Wixon, J. Gene therapy clinical trials worldwide to 2007—An update. *J. Gene Med.* **2007**, *9*, 833–842.
12. Verma, I.M.; Weitzman, M.D. Gene therapy: twenty-first century medicine. *Annu. Rev. Biochem.* **2005**, *74*, 711–738.
13. von Kalle, C.; Fehse, B.; Layh-Schmitt, G.; Schmidt, M.; Kelly, P.; Baum, C. Stem cell clonality and genotoxicity in hematopoietic cells: Gene activation side effects should be avoidable. *Semin. Hematol.* **2004**, *41*, 303–318.
14. Trobridge, G.D.; Miller, D.G.; Jacobs, M.A.; Allen, J.M.; Kiem, H.P.; Kaul, R.; Russell, D.W. Foamy virus vector integration sites in normal human cells. *Proc. Nat. Acad. Sci. U. S. A.* **2006**, *103*, 1498–1503.
15. Nowrouzi, A.; Dittrich, M.; Klanke, C.; Heinkelein, M.; Rammling, M.; Dandekar, T.; von Kalle, C.; Rethwilm, A. Genome-wide mapping of foamy virus vector integrations into a human cell line. *J. Gen. Virol.* **2006**, *87*, 1339–1347.

16. Hacker, C.V.; Vink, C.A.; Wardell, T.W.; Lee, S.; Treasure, P.; Kingsman, S.M.; Mitrophanous, K.A.; Miskin, J.E. The integration profile of EIAV-based vectors. *Mol. Ther.* **2006**, *14*, 536–545.
17. Narezkina, A.; Taganov, K.D.; Litwin, S.; Stoyanova, R.; Hayashi, J.; Seeger, C.; Skalka, A.M.; Katz, R.A. Genome-wide analyses of avian sarcoma virus integration sites. *J. Virol.* **2004**, *78*, 11656–11663.
18. Mitchell, R.S.; Beitzel, B.F.; Schroder, A.R.; Shinn, P.; Chen, H.; Berry, C.C.; Ecker, J.R.; Bushman, F.D. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.* **2004**, *2*, E234.
19. Wu, X.; Li, Y.; Crise, B.; Burgess, S.M. Transcription start regions in the human genome are favored targets for MLV integration. *Science* **2003**, *300*, 1749–1751.
20. Schroder, A.R.; Shinn, P.; Chen, H.; Berry, C.; Ecker, J.R.; Bushman, F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **2002**, *110*, 521–529.
21. Bushman, F.; Lewinski, M.; Ciuffi, A.; Barr, S.; Leipzig, J.; Hannenhalli, S.; Hoffmann, C. Genome-wide analysis of retroviral DNA integration. *Nat. Rev. Microbiol.* **2005**, *3*, 848–858.
22. Nienhuis, A.W.; Dunbar, C.E.; Sorrentino, B.P. Genotoxicity of retroviral integration in hematopoietic cells. *Mol. Ther.* **2006**, *13*, 1031–1049.
23. Temin, H.M.; Mizutani, S. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **1970**, *226*, 1211–1213.
24. Baltimore, D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **1970**, *226*, 1209–1211.
25. Temin, H.M. The DNA provirus hypothesis. *Science* **1976**, *192*, 1075–1080.
26. Temin, H.M. The provirus hypothesis: speculations on the significance of RNA-directed DNA synthesis for normal development and for carcinogenesis. *J. Nat. Canc. Ins.* **1971**, *46*, 3–7.
27. Temin, H.M. Origin of retroviruses from cellular moveable genetic elements. *Cell* **1980**, *21*, 599–600.
28. Cohen, J.C.; Shank, P.R.; Morris, V.L.; Cardiff, R.; Varmus, H.E. Integration of the DNA of mouse mammary tumor virus in virus-infected normal and neoplastic tissue of the mouse. *Cell* **1979**, *16*, 333–345.
29. Steffen, D.; Weinberg, R.A. The integrated genome of murine leukemia virus. *Cell* **1978**, *15*, 1003–1010.
30. Shank, P.R.; Hughes, S.H.; Kung, H.J.; Majors, J.E.; Quintrell, N.; Guntaka, R.V.; Bishop, J.M.; Varmus, H.E. Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. *Cell* **1978**, *15*, 1383–1395.
31. Hughes, S.H.; Shank, P.R.; Spector, D.H.; Kung, H.J.; Bishop, J.M.; Varmus, H.E.; Vogt, P.K.; Breitman, M.L. Proviruses of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. *Cell* **1978**, *15*, 1397–1410.
32. Battula, N.; Temin, H.M. Sites of integration of infectious DNA of avian reticuloendotheliosis viruses in different avian cellular DNAs. *Cell* **1978**, *13*, 387–398.
33. Moebes, A.; Enssle, J.; Bieniasz, P.D.; Heinkelein, M.; Lindemann, D.; Bock, M.; McClure, M. O.; Rethwilm, A. Human foamy virus reverse transcription that occurs late in the viral replication cycle. *J. Virol.* **1997**, *71*, 7305–7311.

34. Rethwilm, A. The replication strategy of foamy viruses. *Curr. Topics Microbiol. Immunol.* **2003**, *277*, 1–26.
35. Telesnitsky, A.; Goff, S.P. Reverse transkriptase and the generation of retroviral DNA. In *Retroviruses*; Coffin, J.M., Hughes, S. H., Varmus, E. H., Eds.; ColdSpring Harbor Laboratory Press: Cold Spring harbor, New York, NY, USA, 1997.
36. Bowerman, B.; Brown, P.O.; Bishop, J.M.; Varmus, H.E. A nucleoprotein complex mediates the integration of retroviral DNA. *Gene. Develop.* **1989**, *3*, 469–478.
37. Fujiwara, T.; Mizuuchi, K. Retroviral DNA integration: structure of an integration intermediate. *Cell* **1988**, *54*, 497–504.
38. Brown, P.O.; Bowerman, B.; Varmus, H.E.; Bishop, J.M. Correct integration of retroviral DNA *in vitro*. *Cell* **1987**, *49*, 347–356.
39. Petit, C.; Giron, M.L.; Tobaly-Tapiero, J.; Bittoun, P.; Real, E.; Jacob, Y.; Tordo, N.; De The, H.; Saib, A. Targeting of incoming retroviral Gag to the centrosome involves a direct interaction with the dynein light chain 8. *J. Cell Sci.* **2003**, *116*, 3433–3442.
40. McDonald, D.; Vodicka, M.A.; Lucero, G.; Svitkina, T.M.; Borisy, G.G.; Emerman, M.; Hope, T.J. Visualization of the intracellular behavior of HIV in living cells. *J. Cell Biol.* **2002**, *159*, 441–452.
41. Saib, A.; Puvion-Dutilleul, F.; Schmid, M.; Peries, J.; de The, H. Nuclear targeting of incoming human foamy virus Gag proteins involves a centriolar step. *J. Virol.* **1997**, *71*, 1155–1161.
42. Roe, T.; Reynolds, T.C.; Yu, G.; Brown, P.O. Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* **1993**, *12*, 2099–2108.
43. Naldini, L.; Blomer, U.; Gallay, P.; Ory, D.; Mulligan, R.; Gage, F.H.; Verma, I.M.; Trono, D. *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **1996**, *272*, 263–267.
44. Bushman, F.D. Targeting survival: integration site selection by retroviruses and LTR-retrotransposons. *Cell* **2003**, *115*, 135–138.
45. Craigie, R.; Fujiwara, T.; Bushman, F. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration *in vitro*. *Cell* **1990**, *62*, 829–837.
46. Hare, S.; Gupta, S.S.; Valkov, E.; Engelman, A.; Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **2010**, *464*, 232–236.
47. Maertens, G.N.; Hare, S.; Cherepanov, P. The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **2010**, *468*, 326–329.
48. Katzman, M.; Mack, J.P.; Skalka, A.M.; Leis, J. A covalent complex between retroviral integrase and nicked substrate DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 4695–1699.
49. Engelman, A.; Mizuuchi, K.; Craigie, R. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* **1991**, *67*, 1211–1221.
50. Bushman, F.D.; Fujiwara, T.; Craigie, R. Retroviral DNA integration directed by HIV integration protein *in vitro*. *Science* **1990**, *249*, 1555–1558.
51. Yoder, K.E.; Bushman, F.D. Repair of gaps in retroviral DNA integration intermediates. *J. Virol.* **2000**, *74*, 11191–1200.

52. Holman, A.G.; Coffin, J.M. Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6103–6107.
53. Wu, X.; Li, Y.; Crise, B.; Burgess, S.M.; Munroe, D.J. Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. *J. Virol.* **2005**, *79*, 5211–5214.
54. Pryciak, P.M.; Sil, A.; Varmus, H.E. Retroviral integration into minichromosomes *in vitro*. *EMBO J.* **1992**, *11*, 291–303.
55. Pryciak, P.M.; Varmus, H.E. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **1992**, *69*, 769–780.
56. Muller, H.P.; Varmus, H.E. DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. *EMBO J.* **1994**, *13*, 4704–4714.
57. Pruss, D.; Reeves, R.; Bushman, F.D.; Wolffe, A.P. The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J. Biol. Chem.* **1994**, *269*, 25031–25041.
58. Scherдин, U.; Rhodes, K.; Breindl, M. Transcriptionally active genome regions are preferred targets for retrovirus integration. *J. Virol.* **1990**, *64*, 907–912.
59. Mooslehner, K.; Karls, U.; Harbers, K. Retroviral integration sites in transgenic Mov mice frequently map in the vicinity of transcribed DNA regions. *J. Virol.* **1990**, *64*, 3056–3058.
60. Rohdewohld, H.; Weiher, H.; Reik, W.; Jaenisch, R.; Breindl, M. Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase I-hypersensitive sites. *J. Virol.* **1987**, *61*, 336–343.
61. Vijaya, S.; Steffen, D.L.; Robinson, H.L. Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin. *J. Virol.* **1986**, *60*, 683–692.
62. Hacein-Bey-Abina, S.; Le Deist, F.; Carlier, F.; Bouneaud, C.; Hue, C.; De Villartay, J.P.; Thrasher, A.J.; Wulffraat, N.; Sorensen, R.; Dupuis-Girod, S.; *et al.* Sustained correction of X-linked severe combined immunodeficiency by *ex vivo* gene therapy. *N. Engl. J. Med.* **2002**, *346*, 1185–1193.
63. Hacein-Bey-Abina, S.; von Kalle, C.; Schmidt, M.; Le Deist, F.; Wulffraat, N.; McIntyre, E.; Radford, I.; Villeval, J.L.; Fraser, C.C.; Cavazzana-Calvo, M.; *et al.* A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* **2003**, *348*, 255–256.
64. Hacein-Bey-Abina, S.; Von Kalle, C.; Schmidt, M.; McCormack, M.P.; Wulffraat, N.; Leboulch, P.; Lim, A.; Osborne, C.S.; Pawliuk, R.; Morillon, E.; *et al.* LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **2003**, *302*, 415–419.
65. Hacein-Bey-Abina, S.; Garrigue, A.; Wang, G.P.; Soulier, J.; Lim, A.; Morillon, E.; Clappier, E.; Caccavelli, L.; Delabesse, E.; Beldjord, K.; *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* **2008**, *118*, 3132–3142.
66. Howe, S.J.; Mansour, M.R.; Schwarzwaelder, K.; Bartholomae, C.; Hubank, M.; Kempinski, H.; Brugman, M.H.; Pike-Overzet, K.; Chatters, S.J.; de Ridder, D.; *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.

67. Hematti, P.; Hong, B.K.; Ferguson, C.; Adler, R.; Hanawa, H.; Sellers, S.; Holt, I.E.; Eckfeldt, C.E.; Sharma, Y.; Schmidt, M.; *et al.* Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. *PLoS Biol.* **2004**, *2*, e423.
68. Themis, M.; Waddington, S.N.; Schmidt, M.; von Kalle, C.; Wang, Y.; Al-Allaf, F.; Gregory, L.G.; Nivsarkar, M.; Themis, M.; Holder, M.V.; *et al.* Oncogenesis following delivery of a nonprimate lentiviral gene therapy vector to fetal and neonatal mice. *Mol. Ther.* **2005**, *12*, 763–771.
69. Lewinski, M.K.; Yamashita, M.; Emerman, M.; Ciuffi, A.; Marshall, H.; Crawford, G.; Collins, F.; Shinn, P.; Leipzig, J.; Hannenhalli, S.; *J. Virol.* Retroviral DNA integration: Viral and cellular determinants of target-site selection. *PLoS Pathog.* **2006**, *2*, e60.
70. Engelman, A.; Cherepanov, P. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog.* **2008**, *4*, e1000046.
71. Vandekerckhove, L.; Christ, F.; Van Maele, B.; De Rijck, J.; Gijssbers, R.; Van den Haute, C.; Witvrouw, M.; Debyser, Z. Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J. Virol.* **2006**, *80*, 1886–1896.
72. Llano, M.; Delgado, S.; Vanegas, M.; Poeschla, E.M. Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. *J. Biol. Chem.* **2004**, *279*, 55570–55577.
73. Busschots, K.; Vercammen, J.; Emiliani, S.; Benarous, R.; Engelborghs, Y.; Christ, F.; Debyser, Z. The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J. Biol. Chem.* **2005**, *280*, 17841–17847.
74. Emiliani, S.; Mousnier, A.; Busschots, K.; Maroun, M.; Van Maele, B.; Tempe, D.; Vandekerckhove, L.; Moisant, F.; Ben-Slama, L.; Witvrouw, M.; *J. Virol.* Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J. Biol. Chem.* **2005**, *280*, 25517–25523.
75. Ciuffi, A.; Llano, M.; Poeschla, E.; Hoffmann, C.; Leipzig, J.; Shinn, P.; Ecker, J.R.; Bushman, F. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **2005**, *11*, 1287–1289.
76. Gijssbers, R.; Ronen, K.; Vets, S.; Malani, N.; De Rijck, J.; McNeely, M.; Bushman, F.D.; Debyser, Z. LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. *Mol. Ther.* **2010**, *18*, 552–560.
77. Marshall, H.M.; Ronen, K.; Berry, C.; Llano, M.; Sutherland, H.; Saenz, D.; Bickmore, W.; Poeschla, E.; Bushman, F.D. Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS ONE* **2007**, *2*, e1340.
78. Shun, M.C.; Raghavendra, N.K.; Vandegraaff, N.; Daigle, J.E.; Hughes, S.; Kellam, P.; Cherepanov, P.; Engelman, A. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Gene. Develop.* **2007**, *21*, 1767–1778.
79. Bartholomae, C.C.; Arens, A.; Balaggan, K.S.; Yanez-Munoz, R.J.; Montini, E.; Howe, S.J.; Paruzynski, A.; Korn, B.; Appelt, J.U.; Macneil, A.; *J. Virol.* Lentiviral Vector Integration Profiles Differ in Rodent Postmitotic Tissues. *Mol. Ther.* **2011**, *19*, 703–710.
80. Hendrix, J.; Gijssbers, R.; De Rijck, J.; Voet, A.; Hotta, J.; McNeely, M.; Hofkens, J.; Debyser, Z.; Engelborghs, Y. The transcriptional co-activator LEDGF/p75 displays a dynamic scan-and-lock mechanism for chromatin tethering. *Nucl. Acid. Res.* **2011**, *39*, 1310–1325.

81. Silvers, R.M.; Smith, J.A.; Schowalter, M.; Litwin, S.; Liang, Z.; Geary, K.; Daniel, R. Modification of integration site preferences of an HIV-1-based vector by expression of a novel synthetic protein. *Hum. Gene Ther.* **2010**, *21*, 337–349.
82. Meehan, A.M.; Poeschla, E.M. Chromatin tethering and retroviral integration: recent discoveries and parallels with DNA viruses. *Biochim. Biophys. Acta* **2010**, *1799*, 182–191.
83. Ferris, A.L.; Wu, X.; Hughes, C.M.; Stewart, C.; Smith, S.J.; Milne, T.A.; Wang, G.G.; Shun, M.C.; Allis, C.D.; Engelman, A.; Hughes, S.H. Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 3135–3140.
84. Meekings, K.N.; Leipzig, J.; Bushman, F.D.; Taylor, G.P.; Bangham, C.R. HTLV-1 integration into transcriptionally active genomic regions is associated with proviral expression and with HAM/TSP. *PLoS Pathog.* **2008**, *4*, e1000027.
85. Derse, D.; Crise, B.; Li, Y.; Princler, G.; Lum, N.; Stewart, C.; McGrath, C.F.; Hughes, S.H.; Munroe, D.J.; Wu, X. Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *J. Virol.* **2007**, *81*, 6731–6741.
86. Maxfield, L.F.; Fraize, C.D.; Coffin, J.M. Relationship between retroviral DNA-integration-site selection and host cell transcription. *Proc. Nat. Acad. Sci. U. S. A.* **2005**, *102*, 1436–1441.
87. Lewinski, M.K.; Bisgrove, D.; Shinn, P.; Chen, H.; Hoffmann, C.; Hannenhalli, S.; Verdin, E.; Berry, C.C.; Ecker, J.R.; Bushman, F.D. Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *J. Virol.* **2005**, *79*, 6610–6619.
88. Taganov, K.D.; Cuesta, I.; Daniel, R.; Cirillo, L.A.; Katz, R.A.; Zaret, K.S.; Skalka, A.M. Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure *in vitro*. *J. Virol.* **2004**, *78*, 5848–5855.
89. Felice, B.; Cattoglio, C.; Cittaro, D.; Testa, A.; Miccio, A.; Ferrari, G.; Luzi, L.; Recchia, A.; Mavilio, F. Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. *PLoS ONE* **2009**, *4*, e4571.
90. Biasco, L.; Ambrosi, A.; Pellin, D.; Bartholomae, C.; Brigida, I.; Roncarolo, M.G.; Di Serio, C.; von Kalle, C.; Schmidt, M.; Aiuti, A. 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.* **2011**, *3*, 89–101.
91. Baum, C. Parachuting in the epigenome: the biology of gene vector insertion profiles in the context of clinical trials. *EMBO Mol. Med.* **2011**, *3*, 75–77.
92. Baum, C.; Dullmann, J.; Li, Z.; Fehse, B.; Meyer, J.; Williams, D.A.; von Kalle, C. Side effects of retroviral gene transfer into hematopoietic stem cells. *Blood* **2003**, *101*, 2099–2114.
93. Montini, E.; Cesana, D.; Schmidt, M.; Sanvito, F.; Bartholomae, C.C.; Ranzani, M.; Benedicenti, F.; Sergi, L.S.; Ambrosi, A.; Ponzoni, M.; *et al.* The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J. Clin. Invest.* **2009**, *119*, 964–975.
94. Maruggi, G.; Porcellini, S.; Facchini, G.; Perna, S.K.; Cattoglio, C.; Sartori, D.; Ambrosi, A.; Schambach, A.; Baum, C.; Bonini, C.; *et al.* Transcriptional enhancers induce insertional gene deregulation independently from the vector type and design. *Mol. Ther.* **2009**, *17*, 851–856.

95. Modlich, U.; Bohne, J.; Schmidt, M.; von Kalle, C.; Knoss, S.; Schambach, A.; Baum, C. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood* **2006**, *108*, 2545–2553.
96. Lund, A.H.; Turner, G.; Trubetskoy, A.; Verhoeven, E.; Wientjens, E.; Hulsman, D.; Russell, R.; DePinho, R.A.; Lenz, J.; van Lohuizen, M. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. *Nat. Genet.* **2002**, *32*, 160–165.
97. Montini, E.; Cesana, D.; Schmidt, M.; Sanvito, F.; Ponzoni, M.; Bartholomae, C.; Sergi, L.S.; Benedicenti, F.; Ambrosi, A.; Di Serio, C. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat. Biotechnol.* **2006**, *24*, 687–696.
98. Bauer, T.R. Jr.; Allen, J.M.; Hai, M.; Tuschong, L.M.; Khan, I.F.; Olson, E.M.; Adler, R.L.; Burkholder, T.H.; Gu, Y.C.; Russell, D.W.; *J. Virol.* Successful treatment of canine leukocyte adhesion deficiency by foamy virus vectors. *Nat. Med.* **2008**, *14*, 93–97.
99. Li, J.; Shen, H.; Himmel, K.L.; Dupuy, A.J.; Largaespada, D.A.; Nakamura, T.; Shaughnessy, J.D.Jr.; Jenkins, N.A.; Copeland, N.G. Leukaemia disease genes: large-scale cloning and pathway predictions. *Nat. Genet.* **1999**, *23*, 348–353.
100. Pfeifer, G.P.; Steigerwald, S.D.; Mueller, P.R.; Wold, B.; Riggs, A.D. Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* **1989**, *246*, 810–813.
101. Mueller, P.R.; Wold, B. *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **1989**, *246*, 780–786.
102. Silver, J.; Keerikatte, V. Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. *J. Virol.* **1989**, *63*, 1924–1928.
103. Schmidt, M.; Schwarzwaelder, K.; Bartholomae, C.; Zaoui, K.; Ball, C.; Pilz, I.; Braun, S.; Glimm, H.; von Kalle, C. High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nat. Methods* **2007**, *4*, 1051–1057.
104. Schmidt, M.; Hoffmann, G.; Wissler, M.; Lemke, N.; Mussig, A.; Glimm, H.; Williams, D.A.; Ragg, S.; Hesemann, C.U.; von Kalle, C. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. *Hum. Gene. Ther.* **2001**, *12*, 743–749.
105. Schmidt, M.; Carbonaro, D.A.; Speckmann, C.; Wissler, M.; Bohnsack, J.; Elder, M.; Aronow, B. J.; Nolta, J.A.; Kohn, D.B.; von Kalle, C. Clonality analysis after retroviral-mediated gene transfer to CD34+ cells from the cord blood of ADA-deficient SCID neonates. *Nat. Med.* **2003**, *9*, 463–468.
106. Gabriel, R.; Eckenberg, R.; Paruzynski, A.; Bartholomae, C.C.; Nowrouzi, A.; Arens, A.; Howe, S.J.; Recchia, A.; Cattoglio, C.; Wang, W.; *et al.* Comprehensive genomic access to vector integration in clinical gene therapy. *Nat. Med.* **2009**, *15*, 1431–1436.
107. Paruzynski, A.; Arens, A.; Gabriel, R.; Bartholomae, C.C.; Scholz, S.; Wang, W.; Wolf, S.; Glimm, H.; Schmidt, M.; von Kalle, C. Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nat. Protocol.* **2010**, *5*, 1379–1395.
108. Warlich, E.; Kuehle, J.; Cantz, T.; Brugman, M.H.; Maetzig, T.; Galla, M.; Filipczyk, A.A.; Halle, S.; Klump, H.; Scholer, H.R.; *et al.* Lentiviral Vector Design and Imaging Approaches to Visualize the Early Stages of Cellular Reprogramming. *Mol. Ther.* **2011**, *19*, 782–789.

109. Schmidt, M.; Schwarzwaelder, K.; Bartholomae, C.C.; Glimm, H.; von Kalle, C. Detection of retroviral integration sites by linear amplification-mediated PCR and tracking of individual integration clones in different samples. *Meth. Mol. Biol.* **2009**, *506*, 363–372.
110. Cartier, N.; Hacein-Bey-Abina, S.; Bartholomae, C.C.; Veres, G.; Schmidt, M.; Kutschera, I.; Vidaud, M.; Abel, U.; Dal-Cortivo, L.; Caccavelli, L.; *et al.* Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* **2009**, *326*, 818–823.
111. Schwarzwaelder, K.; Howe, S.J.; Schmidt, M.; Brugman, M.H.; Deichmann, A.; Glimm, H.; Schmidt, S.; Prinz, C.; Wissler, M.; King, D.J.; *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.
112. Deichmann, A.; Hacein-Bey-Abina, S.; Schmidt, M.; Garrigue, A.; Brugman, M.H.; Hu, J.; Glimm, H.; Gyapay, G.; Prum, B.; Fraser, C.C.; *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.
113. Ott, M.G.; Schmidt, M.; Schwarzwaelder, K.; Stein, S.; Siler, U.; Koehl, U.; Glimm, H.; Kuhlcke, K.; Schilz, A.; Kunkel, H.; *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.
114. Boztug, K.; Schmidt, M.; Schwarzer, A.; Banerjee, P.P.; Diez, I.A.; Dewey, R.A.; Bohm, M.; Nowrouzi, A.; Ball, C.R.; Glimm, H.; *et al.* Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N. Engl. J. Med.* **2010**, *363*, 1918–1927.
115. Wang, G.P.; Garrigue, A.; Ciuffi, A.; Ronen, K.; Leipzig, J.; Berry, C.; Lagresle-Peyrou, C.; Benjelloun, F.; Hacein-Bey-Abina, S.; Fischer, A.; *et al.* DNA bar coding and pyrosequencing to analyze adverse events in therapeutic gene transfer. *Nucl. Acid. Res.* **2008**, *36*, e49.
116. Wang, G.P.; Berry, C.C.; Malani, N.; Leboulch, P.; Fischer, A.; Hacein-Bey-Abina, S.; Cavazzana-Calvo, M.; Bushman, F.D. Dynamics of gene-modified progenitor cells analyzed by tracking retroviral integration sites in a human SCID-X1 gene therapy trial. *Blood* **2010**, *115*, 4356–4366.
117. Stein, S.; Ott, M.G.; Schultze-Strasser, S.; Jauch, A.; Burwinkel, B.; Kinner, A.; Schmidt, M.; Kramer, A.; Schwable, J.; Glimm, H.; *et al.* Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat. Med.* **2010**, *16*, 198–204.
118. Cavazzana-Calvo, M.; Payen, E.; Negre, O.; Wang, G.; Hehir, K.; Fusil, F.; Down, J.; Denaro, M.; Brady, T.; Westerman, K.; *et al.* Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature* **2010**, *467*, 318–322.
119. Payne, G.S.; Bishop, J.M.; Varmus, H.E. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. *Nature* **1982**, *295*, 209–214.
120. Payne, G.S.; Courtneidge, S.A.; Crittenden, L.B.; Fadly, A.M.; Bishop, J.M.; Varmus, H.E. Analysis of avian leukosis virus DNA and RNA in bursal tumours: viral gene expression is not required for maintenance of the tumor state. *Cell* **1981**, *23*, 311–322.

121. Neel, B.G.; Hayward, W.S.; Robinson, H.L.; Fang, J.; Astrin, S.M. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. *Cell* **1981**, *23*, 323–334.
122. Hayward, W.S.; Neel, B.G.; Astrin, S.M. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukemia. *Nature* **1981**, *290*, 475–480.
123. Kool, J.; Berns, A. High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. *Nat. Rev. Canc.* **2009**, *9*, 389–399.
124. Rosenberg, N.; Jolicoeur, P. Retroviral pathogenesis. In *Retroviruses*; Coffin, J.M., Hughes, S.H., Varmus, E.H. Eds.; ColdSpring Harbor Laboratory Press: Cold Spring harbor, New York, NY, USA, **1997**, S475–S586.
125. Moolten, F.L.; Cupples, L.A. A model for predicting the risk of cancer consequent to retroviral gene therapy. *Hum. Gene. Ther.* **1992**, *3*, 479–486.
126. Li, Z.; Dullmann, J.; Schiedlmeier, B.; Schmidt, M.; von Kalle, C.; Meyer, J.; Forster, M.; Stocking, C.; Wahlers, A.; Frank, O.; Ostertag, W.; *et al.* Murine leukemia induced by retroviral gene marking. *Science* **2002**, *296*, 497.
127. Modlich, U.; Kustikova, O.S.; Schmidt, M.; Rudolph, C.; Meyer, J.; Li, Z.; Kamino, K.; von Neuhoff, N.; Schlegelberger, B.; Kuehlcke, K.; *et al.* Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. *Blood* **2005**, *105*, 4235–4246.
128. Calmels, B.; Ferguson, C.; Laukkanen, M.O.; Adler, R.; Faulhaber, M.; Kim, H.J.; Sellers, S.; Hematti, P.; Schmidt, M.; von Kalle, C.; *et al.* Recurrent retroviral vector integration at the Mds1/Evi1 locus in nonhuman primate hematopoietic cells. *Blood* **2005**, *106*, 2530–2533.
129. Kustikova, O.; Fehse, B.; Modlich, U.; Yang, M.; Dullmann, J.; Kamino, K.; von Neuhoff, N.; Schlegelberger, B.; Li, Z.; Baum, C. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science* **2005**, *308*, 1171–1174.
130. Seggewiss, R.; Pittaluga, S.; Adler, R.L.; Guenaga, F.J.; Ferguson, C.; Pilz, I.H.; Ryu, B.; Sorrentino, B.P.; Young, W.S.; Donahue, R.E.; *et al.* Acute myeloid leukemia is associated with retroviral gene transfer to hematopoietic progenitor cells in a rhesus macaque. *Blood* **2006**, *107*, 3865–3867.
131. Aiuti, A.; Cattaneo, F.; Galimberti, S.; Benninghoff, U.; Cassani, B.; Callegaro, L.; Scaramuzza, S.; Andolfi, G.; Mirolo, M.; Brigida, I.; *et al.* Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N. Engl. J. Med.* **2009**, *360*, 447–458.
132. Woods, N.B.; Bottero, V.; Schmidt, M.; von Kalle, C.; Verma, I.M. Gene therapy: Therapeutic gene causing lymphoma. *Nature* **2006**, *440*, 1123.
133. Thrasher, A.J.; Gaspar, H.B.; Baum, C.; Modlich, U.; Schambach, A.; Candotti, F.; Otsu, M.; Sorrentino, B.; Scobie, L.; Cameron, E.; *et al.* Gene therapy: X-SCID transgene leukaemogenicity. *Nature* **2006**, *443*, E5–E6; discussion E6–E7.
134. Dave, U.P.; Akagi, K.; Tripathi, R.; Cleveland, S.M.; Thompson, M.A.; Yi, M.; Stephens, R.; Downing, J.R.; Jenkins, N.A.; Copeland, N.G. Murine leukemias with retroviral insertions at Lmo2 are predictive of the leukemias induced in SCID-X1 patients following retroviral gene therapy. *PLoS Genet.* **2009**, *5*, e1000491.

135. Cartier, N.; Aubourg, P. Hematopoietic stem cell transplantation and hematopoietic stem cell gene therapy in X-linked adrenoleukodystrophy. *Brain Pathol.* **2010**, *20*, 857–862.
136. Montini, E.; Cesana, D.; Schmidt, M.; Sanvito, F.; Ponzoni, M.; Bartholomae, C.; Sergi, L.; Benedicenti, F.; Ambrosi, A.; Di Serio, 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.
137. Modlich, U.; Navarro, S.; Zychlinski, D.; Maetzig, T.; Knoess, S.; Brugman, M.H.; Schambach, A.; Charrier, S.; Galy, A.; Thrasher, A.J.; *et al.* Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol. Ther.* **2009**, *17*, 1919–1928.
138. Daya, S.; Berns, K.I. Gene therapy using adeno-associated virus vectors. *Clin. Microbiol. Rev.* **2008**, *21*, 583–593.
139. Russell, D.W. AAV vectors, insertional mutagenesis, and cancer. *Mol. Ther.* **2007**, *15*, 1740–1743.
140. Donsante, A.; Miller, D.G.; Li, Y.; Vogler, C.; Brunt, E.M.; Russell, D.W.; Sands, M.S. AAV vector integration sites in mouse hepatocellular carcinoma. *Science* **2007**, *317*, 477.
141. Wanisch, K.; Yanez-Munoz, R.J. Integration-deficient lentiviral vectors: A slow coming of age. *Mol. Ther.* **2009**, *17*, 1316–1332.
142. Yanez-Munoz, R.J.; Balaggan, K.S.; MacNeil, A.; Howe, S.J.; Schmidt, M.; Smith, A.J.; Buch, P.; MacLaren, R.E.; Anderson, P.N.; Barker, S.E.; *et al.* Effective gene therapy with nonintegrating lentiviral vectors. *Nat. Med.* **2006**, *12*, 348–353.
143. Bushman, F. Targeting retroviral integration? *Mol. Ther.* **2002**, *6*, 570–571.
144. Urnov, F.D.; Rebar, E.J.; Holmes, M.C.; Zhang, H.S.; Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **2010**, *11*, 636–646.
145. Ivics, Z.; Izsvak, Z. The expanding universe of transposon technologies for gene and cell engineering. *Mob. DNA* **2010**, *1*, 25.
146. Copeland, N.G.; Jenkins, N.A. Harnessing transposons for cancer gene discovery. *Nat. Rev. Canc.* **2010**, *10*, 696–706.
147. Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **2007**, *131*, 861–872.
148. Takahashi, K.; Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126*, 663–676.
149. Yamanaka, S.; Blau, H.M. Nuclear reprogramming to a pluripotent state by three approaches. *Nature* **2010**, *465*, 704–712.
150. Vierbuchen, T.; Ostermeier, A.; Pang, Z.P.; Kokubu, Y.; Sudhof, T.C.; Wernig, M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **2010**, *463*, 1035–1041.
151. Szabo, E.; Rampalli, S.; Risueno, R.M.; Schnerch, A.; Mitchell, R.; Fiebig-Comyn, A.; Levadoux-Martin, M.; Bhatia, M. Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* **2010**, *468*, 521–526.

152. Ieda, M.; Fu, J.D.; Delgado-Olguin, P.; Vedantham, V.; Hayashi, Y.; Bruneau, B.G.; Srivastava, D. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **2010**, *142*, 375–386.
153. Papapetrou, E.P.; Lee, G.; Malani, N.; Setty, M.; Riviere, I.; Tirunagari, L.M.; Kadota, K.; Roth, S.L.; Giardina, P.; Viale, A.; *et al.* Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. *Nat. Biotechnol.* **2011**, *29*, 73–78.
154. Yamanaka, S. Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem. Cell* **2007**, *1*, 39–49.
155. Ellis, J.; Baum, C.; Benvenisty, N.; Mostoslavsky, G.; Okano, H.; Stanford, W.L.; Porteus, M.; Sadelain, M. Benefits of utilizing gene-modified iPSCs for clinical applications. *Cell Stem. Cell* **2010**, *7*, 429–430.
156. Suzuki, T.; Shen, H.; Akagi, K.; Morse, H.C.; Malley, J.D.; Naiman, D.Q.; Jenkins, N.A.; Copeland, N.G. New genes involved in cancer identified by retroviral tagging. *Nat. Genet.* **2002**, *32*, 166–174.
157. Akagi, K.; Suzuki, T.; Stephens, R.M.; Jenkins, N.A.; Copeland, N.G. RTCGD: Retroviral tagged cancer gene database. *Nucl. Acid. Res.* **2004**, *32*, D523–D527.
158. Starr, T.K.; Allaei, R.; Silverstein, K.A.; Staggs, R.A.; Sarver, A.L.; Bergemann, T.L.; Gupta, M.; O'Sullivan, M.G.; Matisse, I.; Dupuy, A.J.; *et al.* A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science* **2009**, *323*, 1747–1750.
159. Dupuy, A.J.; Rogers, L.M.; Kim, J.; Nannapaneni, K.; Starr, T.K.; Liu, P.; Largaespada, D.A.; Scheetz, T.E.; Jenkins, N.A.; Copeland, N.G. A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. *Canc. Res.* **2009**, *69*, 8150–8156.
160. Du, Y.; Spence, S.E.; Jenkins, N.A.; Copeland, N.G. Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* **2005**, *106*, 2498–2505.
161. Du, Y.; Jenkins, N.A.; Copeland, N.G. Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. *Blood* **2005**, *106*, 3932–3939.
162. Copeland, N.G.; Jenkins, N.A. Retroviral integration in murine myeloid tumors to identify Evi-1, a novel locus encoding a zinc-finger protein. *Adv. Canc. Res.* **1990**, *54*, 141–157.
163. Neff, T.; Shotkoski, F.; Stamatoyannopoulos, G. Stem cell gene therapy, position effects and chromatin insulators. *Stem Cells* **1997**, *15*, 265–271.
164. Valenzuela, L.; Kamakaka, R.T. Chromatin insulators. *Annu. Rev. Genet.* **2006**, *40*, 107–138.
165. Arumugam, P.I.; Higashimoto, T.; Urbinati, F.; Modlich, U.; Nestheide, S.; Xia, P.; Fox, C.; Corsinotti, A.; Baum, C.; Malik, P. Genotoxic potential of lineage-specific lentivirus vectors carrying the beta-globin locus control region. *Mol. Ther.* **2009**, *17*, 1929–1937.