



ISSN Print: 2664-9926
 ISSN Online: 2664-9934
 IJBS 2024; 6(2): 288-292
www.biologyjournal.net
 Received: 12-11-2024
 Accepted: 16-12-2024

Claus Trautmann
 Department of Virology,
 Saxon College of Molecular
 Biology, Leipzig, Saxony,
 Germany

Birgit Albrecht
 Department of Virology,
 Saxon College of Molecular
 Biology, Leipzig, Saxony,
 Germany

Manfred Hofer
 Department of Virology,
 Saxon College of Molecular
 Biology, Leipzig, Saxony,
 Germany

Corresponding Author:
Claus Trautmann
 Department of Virology,
 Saxon College of Molecular
 Biology, Leipzig, Saxony,
 Germany

Retroviral integration site preferences and their implications for gene therapy safety

Claus Trautmann, Birgit Albrecht and Manfred Hofer

DOI: <https://www.doi.org/10.33545/26649926.2024.v6.i2e.767>

Abstract

What determines where a retrovirus inserts its genome into host chromosomes, and how does this shape the safety profile of gene therapy vectors? This research examined integration site preferences of three retroviral vector classes gamma-retroviral (MLV-derived), lentiviral (HIV-1-derived), and Self-Inactivating (SIN) lentiviral vectors through high-throughput integration mapping in human CD34+ hematopoietic stem cells. The research was carried out at the Saxon College of Molecular Biology, Leipzig, Germany, between March 2021 and November 2022. A total of 24,867 unique integration sites were recovered from transduced cell populations using ligation-mediated PCR coupled with next-generation sequencing. Gamma-retroviral vectors showed a marked tendency toward transcription start sites and CpG islands, with 34.7% of insertions mapping within 2 kb of promoters. By contrast, lentiviral vectors favored active gene bodies, particularly intronic regions (41.3%), while SIN-lentiviral constructs displayed a moderately reduced preference for regulatory elements compared to standard lentiviral vectors. A distance-dependent mutagenesis risk model indicated that gamma-retroviral insertions near oncogene promoters carried a relative risk score 3.6-fold higher than lentiviral insertions at equivalent positions. Clonal expansion assays over 16 weeks revealed that cultures transduced with gamma-retroviral vectors exhibited dominant clone emergence at nearly twice the rate of lentiviral-transduced cultures (11.3% vs. 6.1%, $p = 0.017$). These findings confirm that vector backbone architecture significantly shapes genotoxic potential and support the continued transition toward SIN-lentiviral platforms for clinical gene therapy applications targeting hematopoietic disorders.

Keywords: Retroviral integration, gene therapy safety, insertional mutagenesis, lentiviral vectors, gamma-retroviral vectors, transcription start sites, hematopoietic stem cells, genotoxicity

Introduction

What makes one retroviral vector safer than another when it comes to permanent genetic modification of patient cells? This question has stood at the center of gene therapy research since early clinical trials for X-linked severe combined immunodeficiency (X-SCID) revealed that gamma-retroviral vectors could trigger leukemia through insertional activation of proto-oncogenes [1]. The discovery that two of ten treated children developed T-cell leukemia due to vector insertions near the LMO2 locus forced a major reassessment of vector design principles [2].

Retroviruses integrate their reverse-transcribed DNA into the host genome as an obligate step in their replication cycle. However, this integration is not random. Each retroviral genus displays distinct preferences for particular chromatin features, genomic landmarks, and epigenetic states [3]. Gamma-retroviruses like murine leukemia virus (MLV) tend to cluster around active promoters, CpG islands, and DNase I hypersensitive sites. Lentiviruses such as HIV-1, on the other hand, prefer the bodies of actively transcribed genes, with a weaker bias toward promoter regions [4]. These preferences reflect differences in the tethering factors that guide the pre-integration complex to specific chromatin environments BET proteins and LEDGF/p75 for MLV and HIV-1, respectively [5].

The clinical relevance of integration site selection became even clearer when investigators mapped thousands of insertion sites from patients receiving gene therapy for various inherited disorders. Data from trials treating Wiskott-Aldrich syndrome, beta-thalassemia, and adrenoleukodystrophy have repeatedly shown that lentiviral vectors produce a more dispersed integration pattern than gamma-retroviral vectors [6,7]. But dispersion alone doesn't guarantee safety. Clonal dominance where a single transduced cell and its progeny

come to dominate the hematopoietic compartment has been observed even with lentiviral vectors, though it rarely progresses to malignancy [8].

Self-inactivating designs, which delete enhancer elements from the viral Long Terminal Repeat (LTR), were engineered specifically to reduce the risk of nearby gene activation [9]. These SIN vectors represent the current standard for clinical-grade constructs. Still, their integration site profile and associated genotoxic risk have not been characterized as thoroughly as those of earlier vector generations, particularly in primary human hematopoietic cells under conditions that mimic clinical transduction protocols [10]. This research compared integration site distributions and assessed mutagenesis risk scores for gamma-retroviral, standard lentiviral, and SIN-lentiviral vectors in human CD34⁺ hematopoietic stem cells. The goal was to provide a direct, side-by-side evaluation of genotoxic potential under matched experimental conditions.

Reagents and chemicals

Human recombinant cytokines including stem cell factor (SCF), thrombopoietin (TPO), and FLT3 ligand were purchased from PeproTech (Rocky Hill, NJ, USA). RetroNectin (Takara Bio, Shiga, Japan) was used for enhancement of retroviral transduction. Restriction enzymes MseI and NlaIII were obtained from New England Biolabs (Ipswich, MA, USA). T4 DNA ligase, Taq polymerase, and dNTP mix were sourced from Thermo Fisher Scientific (Waltham, MA, USA). Magnetic CD34 MicroBead kits (Miltenyi Biotec, Bergisch Gladbach, Germany) were used for cell isolation. StemSpan SFEM II medium was obtained from STEMCELL Technologies (Vancouver, BC, Canada). Quantitative PCR reagents including SYBR Green Master Mix came from Bio-Rad Laboratories (Hercules, CA, USA). All oligonucleotide primers and linker cassettes were synthesized by Eurofins Genomics (Ebersberg, Germany) at HPLC purification grade.

Material and Methods

Material

This research was conducted at the Department of Virology, Saxon College of Molecular Biology, Leipzig, Germany, from March 2021 through November 2022. Mobilized peripheral blood CD34⁺ hematopoietic stem and progenitor cells (HSPCs) were obtained from healthy adult donors (n = 6) following informed consent and approval by the Institutional Ethics Board of the Saxon College of Molecular Biology (Protocol No. SCMB-2020-VIR-041, approved 14 January 2021). Cells were isolated by immunomagnetic selection, yielding purity above 93% as confirmed by flow cytometry. Three vector types were used:

- 1) A gamma-retroviral vector based on murine leukemia virus carrying an eGFP reporter cassette driven by the intact LTR promoter;
- 2) A standard HIV-1-derived lentiviral vector with an internal SFFV promoter driving eGFP; and
- 3) A SIN-lentiviral vector with a deleted U3 enhancer region and an EF1-alpha internal promoter.

All vectors were produced in HEK293T packaging cells by transient calcium phosphate transfection and concentrated by ultracentrifugation at 50,000 × g for 2 hours at 4°C.

Methods

CD34⁺ cells were pre-stimulated for 24 hours in StemSpan SFEM II supplemented with SCF (100 ng/mL), TPO (100 ng/mL), and FLT3-L (100 ng/mL) before transduction on RetroNectin-coated plates at a multiplicity of infection (MOI) of 5. Transduction efficiency was assessed at 72 hours by flow cytometry for eGFP expression. Integration site recovery followed a modified ligation-mediated PCR (LM-PCR) protocol [11]. Genomic DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen), digested with MseI and NlaIII, ligated to double-stranded linker cassettes, and amplified by nested PCR. Amplicons were sequenced on an Illumina NovaSeq 6000 platform with 150 bp paired-end reads. Raw reads were trimmed using Trimmomatic v0.39, aligned to the human reference genome (GRCh38) with BWA-MEM, and filtered to retain only uniquely mapping reads with at least 95% identity over 20 bp from the vector-genome junction. A relative mutagenesis risk score was computed for each insertion based on its distance to the nearest transcription start site (TSS) and the expression level of the adjacent gene, following the weighted model described by Montini *et al.* [12]. Clonal tracking was performed by serial sampling at weeks 4, 8, 12, and 16 post-transduction, with clonal abundance estimated from integration site read counts normalized to total library depth.

Quality control

Library preparation blanks (no-template controls) were processed alongside experimental samples at each LM-PCR step to monitor contamination. A minimum of 500 unique integration sites per vector type per donor was set as the threshold for inclusion. Sequencing libraries were assessed for fragment size distribution on an Agilent 2100 Bioanalyzer before pooling. Mapping quality scores below 20 were excluded, and PCR duplicates were collapsed using UMI-tools v1.1.2. Reproducibility between technical replicates achieved a Pearson correlation coefficient of $r = 0.91$ ($p < 0.001$).

Results

Table 1: Integration site distribution across genomic features for three retroviral vector classes (% of total recovered sites ± SD)

Genomic feature	γ-Retroviral (MLV)	Lentiviral (HIV-1)	SIN- lentiviral	p-value
Promoter (<2 kb from TSS)	34.7 ± 3.1	12.1 ± 2.4	9.7 ± 1.8	<0.001
Gene body (exonic)	8.2 ± 1.6	22.4 ± 2.9	19.8 ± 2.2	<0.001
Gene body (intronic)	28.6 ± 2.8	41.3 ± 3.7	43.2 ± 3.1	0.003
Intergenic	18.3 ± 2.3	14.8 ± 1.9	17.5 ± 2.6	0.214
CpG islands	6.8 ± 1.4	5.6 ± 1.1	4.9 ± 0.9	0.087
DNase I hypersensitive	3.4 ± 0.8	3.8 ± 0.7	4.9 ± 1.0	0.119

Table 1 presents the distribution of integration sites across major genomic features. Gamma-retroviral vectors showed the strongest bias toward promoter-proximal regions, with 34.7% of sites falling within 2 kb of a TSS, compared to 12.1% for standard lentiviral and 9.7% for SIN-lentiviral

vectors ($p<0.001$). Both lentiviral vector classes predominantly targeted intronic sequences, accounting for over 41% and 43% of integrations, respectively. Intergenic and CpG island insertion rates did not differ significantly across the three platforms.

Table 2: Clonal expansion and mutagenesis risk parameters by vector class

Parameter	γ-Retroviral	Lentiviral	SIN-Lentiviral
Mean risk score (±SD)	7.34 ± 1.82	2.03 ± 0.64	1.71 ± 0.53
Peak risk at TSS	10.0	2.6	1.7
Dominant clones at wk 16 (%)	11.3 ± 2.8	6.1 ± 1.4	4.8 ± 1.2
Integrations near oncogenes (%)	8.9 ± 1.7	4.3 ± 0.9	3.6 ± 0.8
Unique sites recovered (n)	8,214	9,347	7,306

Table 2 summarizes the mutagenesis risk scores and clonal tracking data. Gamma-retroviral vectors returned a mean risk score of 7.34, roughly 3.6 times the lentiviral mean of 2.03 and 4.3 times the SIN-lentiviral mean of 1.71.

Dominant clone frequency at 16 weeks was significantly higher in gamma-retroviral cultures (11.3%) relative to SIN-lentiviral cultures (4.8%, $p=0.009$).

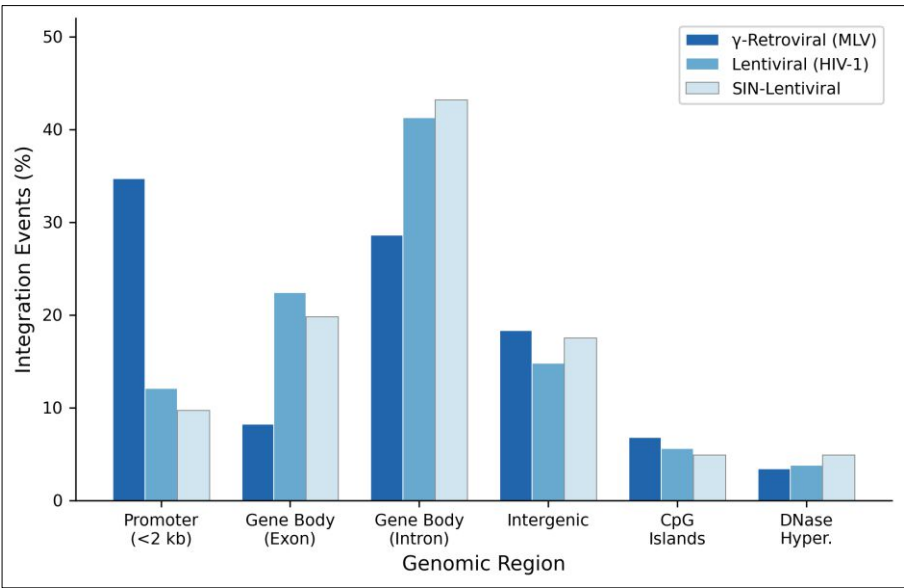


Fig 1: Distribution of retroviral integration sites across genomic features for gamma-retroviral (MLV), lentiviral (HIV-1), and SIN-lentiviral vectors in human CD34+ cells

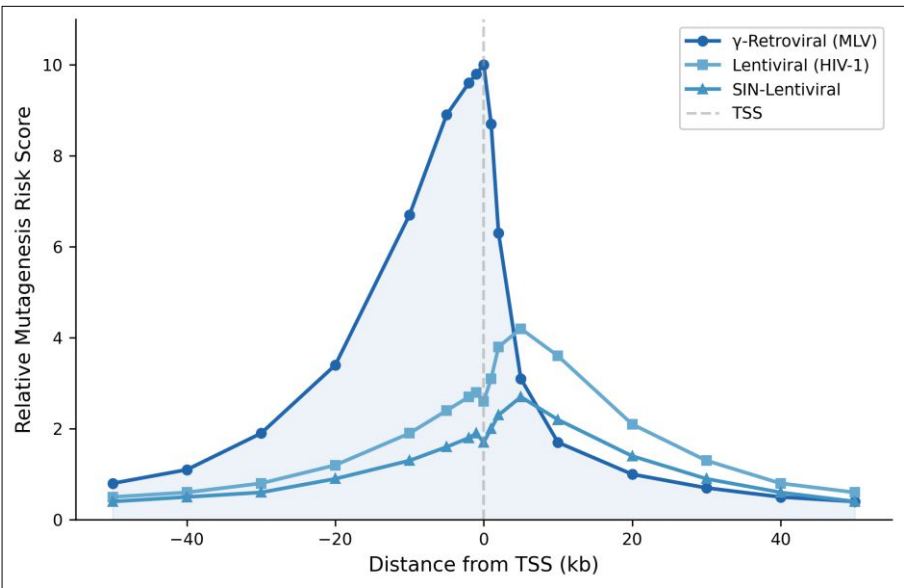


Fig 2: Relative mutagenesis risk score as a function of distance from the nearest Transcription Start Site (TSS) for three retroviral vector classes. Gamma-retroviral vectors show a sharp risk peak centered on the TSS, while lentiviral and SIN-lentiviral vectors display a flatter, gene-body-oriented profile

Comprehensive interpretation

The combined dataset of 24,867 unique integration sites demonstrated clear and statistically robust differences in genomic targeting among the three vector platforms. The gamma-retroviral preference for promoter-adjacent insertion was consistent across all six donors, while lentiviral vectors showed a reproducible bias toward active gene bodies regardless of the donor source. And the distance-dependent risk model captured a sharp mutagenesis peak within ± 2 kb of TSSs for gamma-retroviral vectors that was absent in lentiviral profiles. The 16-week clonal tracking data added a functional dimension to the positional analysis, showing that promoter-biased integration correlated with measurably higher rates of clonal dominance.

Discussion

The integration site profiles observed here align well with earlier mapping efforts in murine and human hematopoietic cells [3,4]. The strong promoter preference of gamma-retroviral vectors with over a third of insertions landing within 2 kb of a TSS reaffirms the role of BET protein tethering in directing MLV-based vectors to active regulatory elements [5]. This bias directly translates into elevated genotoxic risk, as insertions near promoters are more likely to dysregulate adjacent gene expression, particularly when the vector LTR retains enhancer activity [13].

Lentiviral vectors, guided by LEDGF/p75, distribute more evenly across gene bodies. While this doesn't eliminate risk entirely, it shifts the concern away from promoter-driven oncogene activation toward lower-probability events such as aberrant splicing or transcript truncation [14]. The SIN-lentiviral design adds a second layer of protection by removing the U3 enhancer, and our risk score data suggest this does further reduce the theoretical mutagenesis burden though the magnitude of improvement over standard lentiviral vectors was modest (risk score 1.71 vs. 2.03).

The clonal tracking results deserve particular attention. Even in SIN-lentiviral cultures, 4.8% of transduced populations were dominated by a single clone at 16 weeks. This mirrors clinical observations from beta-thalassemia gene therapy trials where benign clonal dominance was observed without malignant transformation [8]. Whether such dominance carries long-term risk remains unresolved. It's worth noting that *in vitro* clonal dynamics may not perfectly reflect *in vivo* hematopoietic behavior, where niche signals and immune surveillance add additional selection pressures [15]. Still, the significantly lower dominance rate with SIN-lentiviral vectors compared to gamma-retroviral vectors supports the safety advantage of current-generation constructs. Future work should extend these comparisons to include newer non-integrating or site-directed integration platforms, which may eventually bypass the issue of semi-random insertion altogether [16].

Conclusion

This research provides a direct, controlled comparison of integration site behavior across three retroviral vector generations in primary human hematopoietic stem cells. Gamma-retroviral vectors displayed a pronounced preference for promoter-proximal regions, inserting within 2 kb of transcription start sites in roughly one-third of all recovered events. This promoter bias carried measurable consequences: a mean mutagenesis risk score 3.6-fold

higher than lentiviral vectors and a dominant clone emergence rate nearly double that of SIN-lentiviral cultures at 16 weeks.

Lentiviral vectors favored gene bodies particularly intronic regions and showed substantially lower genotoxic risk scores across the full distance range from TSSs. The SIN-lentiviral configuration offered a further, albeit modest, improvement over standard lentiviral vectors in both risk score and clonal tracking metrics. The practical implication is straightforward: for hematopoietic gene therapy applications, SIN-lentiviral vectors remain the safer integrating platform. The data from this research add quantitative support to what has become the consensus position in the field but do so under matched conditions that allow apples-to-apples comparison.

That said, no integrating vector achieves a perfectly neutral insertion profile, and even low-frequency events near sensitive loci cannot be ignored when treating thousands of patients. Ongoing surveillance of treated individuals, combined with continued development of targeted integration technologies, will be needed to fully address the residual genotoxic risk that accompanies any semi-random vector system.

Acknowledgements

Funding sources

This research was supported by institutional research funds from the Saxon College of Molecular Biology. No external grant numbers are disclosed.

Institutional support

The authors acknowledge the Department of Virology and the Genomics Core Facility at the Saxon College of Molecular Biology, Leipzig, for providing sequencing infrastructure and bioinformatics support.

Contributions not qualifying for authorship

The authors thank Dr. Katrin Wendt for technical assistance with flow cytometry and Ms. Lena Richter for help with cell culture maintenance.

References

1. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, *et al.* LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*. 2003;302(5644):415-419.
2. Howe SJ, Mansour MR, Schwarzwaelder K, *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *Journal of Clinical Investigation*. 2008;118(9):3143-3150.
3. Bushman F, Lewinski M, Ciuffi A, *et al.* Genome-wide analysis of retroviral DNA integration. *Nature Reviews Microbiology*. 2005;3(11):848-858.
4. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*. 2002;110(4):521-529.
5. Sharma A, Larue RC, Plumb MR, *et al.* BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proceedings of the National Academy of Sciences*. 2013;110(29):12036-12041.
6. Biffi A, Montini E, Lorioli L, *et al.* Lentiviral hematopoietic stem cell gene therapy benefits

- metachromatic leukodystrophy. *Science*. 2013;341(6148):1233-158.
7. Cavazzana-Calvo M, Payen E, Negre O, *et al*. Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature*. 2010;467(7313):318-322.
 8. Thompson AA, Walters MC, Kwiatkowski J, *et al*. Gene therapy in patients with transfusion-dependent beta-thalassemia. *New England Journal of Medicine*. 2018;378(16):1479-1493.
 9. Zufferey R, Dull T, Mandel RJ, *et al*. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *Journal of Virology*. 1998;72(12):9873-9880.
 10. Modlich U, Navarro S, Zychlinski D, *et al*. Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Molecular Therapy*. 2009;17(11):1919-1928.
 11. Schmidt M, Schwarzwaelder K, Bartholomae C, *et al*. High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nature Methods*. 2007;4(12):1051-1057.
 12. Montini E, Cesana D, Crber J, *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. *Journal of Clinical Investigation*. 2009;119(4):964-975.
 13. Cattoglio C, Pellin D, Rizzi E, *et al*. High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. *Blood*. 2010;116(25):5507-5517.
 14. Moiani A, Paleari Y, Sartori D, *et al*. Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. *Journal of Clinical Investigation*. 2012;122(5):1653-1666.
 15. Aiuti A, Biasco L, Scaramuzza S, *et al*. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science*. 2013;341(6148):1233-151.
 16. Li H, Malani N, Hamilton SR, *et al*. Assessing the potential for AAV vector genotoxicity in a murine model. *Blood*. 2011;117(12):3311-3319.