HIV-1 integration site preferences in pluripotent cells

Janet L. Markman[1], Robert M. Silvers[1], Abibatou M.N. Ndoye[1], Kyla R. Geary[1], David Alvarado[1], Johanna A. Smith[1], Rene Daniel[1,2,3]

[1]*Division of Infectious Diseases - Center for Human Virology, Thomas Jefferson University, Philadelphia,* [2]*Center for Stem Cell Biology and Regenerative Medicine, Thomas Jefferson University, Philadelphia,* [3]*Kimmel Cancer Center, Immunology Program, Thomas Jefferson University, Philadelphia*

TABLE OF CONTENTS

1. ABSTRACT

HIV-1-based vectors are widely used in gene therapy. In somatic cells, these vectors mainly integrate within genes. However, no distinct integration site preferences have been observed with regard to large chromosomal regions. The recent emergence of induced pluripotent stem (iPS) cells, similar to embryonic stem (ES) cells, has raised questions about where integration occurs in these cells. In this work we investigated the integration site preferences of HIV-1-based vectors in a pluripotent, ES-like cell line. We show that approximately 30% of the integrations occur in the vicinity of telomeres. We have analyzed integration sites in various somatic cells, as reported by us and other groups, and observed that this integration pattern is unique to the analyzed pluripotent cell line. We conclude that pluripotent cells may contain distinct cellular cofactors that participate in integration targeting and that are not present in somatic cells.

2. INTRODUCTION

Integration of viral DNA into host DNA is an essential step in the replication cycle of retroviral vectors (1-3). Gene therapy trials currently use three classes of retroviral vectors. These are MLV (murine leukemia virus)-based, HIV (human immunodeficiency virus)-based, and ASV (avian sarcoma virus)-based vectors. Currently gene therapy trials most commonly utilize MLV-based vectors. This is so because MLV-based vectors are thoroughly characterized, as their genomes are relatively simple. However, HIV-based vectors made tremendous gains in recent years and are slowly replacing MLV-based vectors in widespread use. This is mainly due to the development of simple packaging systems and the removal of all HIV-1- genes which are not necessary for gene transfer (4). In addition, these HIV-based vectors possess a crucial advantage over the MLV-based vectors: unlike the latter, they can efficiently integrate in the

genome of nondividing cells (5, 6). HIV-based vectors are thus more versatile than those constructed from MLV. Finally, some laboratories utilize ASV-based vectors (7, 8). However, these are not yet fully characterized and also integrate in nondividing cells less efficiently than HIV-based vectors (9).

It has long been observed that retroviral vectors of all classes can integrate at virtually any site of the human genome (1-3). This is due to nonspecific binding of cellular DNA by retroviral integrases (1-3). However, integration site selection by retroviral vectors is not a random process and the MLV-based vectors were shown to preferentially integrate in promoter regions (10, 11). This unfortunate fact is also responsible for the disastrous outcome of a gene therapy trial, where an MLV-based vector integrated in the vicinity of a protooncogene, and activated its expression (12). In contrast, HIV-1-based vectors prefer to integrate in genes, but do not specifically prefer the promoter region (13). Finally, ASV-based vectors show a weak preference for genes [40% of integrations occur in genes, as opposed to 33%, which would be a result of random integration, (10, 14)]. The above described integration site preferences were found in all cell types examined to date. So far, it is not known what factors drive integration site selection, with an exception of the HIV-1-based vectors, where integration into genes is directed by the cellular protein LEDGF/p75, which binds to the integrase protein and tethers it to chromatin (15-20). However, additional factors likely exist, which control integration site selection. The above-cited studies focused on integration in somatic cells which originate from stem cells. However, stem cells possess unique characteristics that distinguish them from somatic cells.

Stem cells are undifferentiated cells that are capable of self-renewal and differentiation. Stem cells compose only a small minority of a tissue's cells. A subspecies of stem cells are embryonic stem (ES) cells. They are pluripotent cells, which can be obtained from early stage embryos (blastocyst) and can differentiate into all three primary germ layers [pluripotency, (21, 22)]. ES cells thus represent an important, very early stage of organism development.

ES cells are characterized by the expression of embryonic stem cell transcriptional factors (ESTFs). The most important of these are Oct4, Nanog and Sox2. Animal and other studies have suggested that ESTFs are crucial for self-renewal and pluripotency of ES cells (21). This hypothesis was recently confirmed by reprogramming adult somatic cells into induced pluripotent stem (iPS) cells, which possess ES cell properties (23).

iPS cells were developed with the long-term objective to gain a new therapeutic tool. Since they are pluripotent, they could hypothetically be used to develop into any desired cell type for the replacement of cells that are lacking in patients, from whose cells the iPS cells originated [e.g. insulin-producing cells for the treatment of diabetes, or replacement cells for treating genetic diseases, such as ADA-SCID and Huntington Disease,

(24)]. These cells are also attractive targets for gene transfer.

We speculated that integration site selection in pluripotent stem cells may have unique characteristics, since the protein content of these cells is different from somatic cells. As a model, we have chosen NCCIT cells, which are derived from a human embryonal carcinoma and have features common to ES and iPS cells, including the ability to differentiate into derivatives of all three germ layers (22, 25, 26). We demonstrate that in these cells, integration by HIV-1-based vectors occurs in a unique pattern, which is characterized by an increased frequency of integration in the vicinity of chromosomal ends.

3. MATERIALS AND METHODS

3.1. Cells

NCCIT cells (CRL-2073^{TM}) were purchased from the American Type Culture Collection (ATCC, Manassas, VA) and were maintained in RPMI-1640 medium supplemented with 10% fetal bovine serum, non-essential amino acids, and Pen/Strep.

HIV-1 based vectors:

Vesicular stomatitis virus (VSV) G-pseudotyped HIV-1 based vectors were prepared as described by Naldini and colleagues (27).

3.2. Infection

Viral particles with a titer of 5×10^6 particles per mL were applied to each 60 mm cell culture dish at a multiplicity of infection (MOI) of 5 or 20. The viral particles were supplemented with 10 microgram/mL DEAE dextran to enhance infection. The media was changed after a 16 hour incubation period. Four days post-infection the cultured cells were trypsinized and genomic DNA extraction was performed using the QIAamp® DNA mini kit (cat. #51306; QIAGEN, Hilden, Germany). The genomic DNA was eluted with 200 microliters of AE buffer provided with the kit. Concentration of the genomic DNA sample was determined using ultraviolet (UV) spectroscopy. Approximately 1 microgram of DNA was electrophoresed on a 0.6% agarose gel to determine the overall size and purity of the extracted DNA.

3.3. Construction of Genome Walker^{TM} (GW) libraries

The following protocol is adapted from the Clontech Genome Walker^{TM} Universal Kit (cat. # 638904); Clontech/TaKaRa Bio, Mountain View, CA). Approximately 3 micrograms of extracted genomic DNA was digested with four blunt-cutting enzymes (DraI, StuI, MscI, and MslI). A total volume of 100 microliters per reaction included 3 micrograms genomic DNA, 80 units of combined restriction enzymes (20 units each), 10 microliters of 10X restriction enzyme buffer 4, and nuclease-free water. The reactions were gently mixed and incubated at 37°C overnight.

The following day the digested genomic DNA was extracted and purified via phenol chloroform extraction and ethanol precipitation. The pellets were

dissolved in 20 microliters of nuclease-free water. To confirm digestion, 1 microgram of each sample was electrophoresed on a 0.6% agarose gel.

Following electrophoresis, 4 microliters of digested, purified genomic DNA, 1.9 microliters of GenomeWalker® Adaptor (25 µM), 1.6 microliters of 10X ligation buffer, and 0.7 microliters of T4 DNA ligase enzyme (6 units/ microliter) were mixed in a 0.2 mL PCR tube and incubated overnight at 16°C in a PCR thermal cycler. The reaction was stopped the following day by incubating the samples at 70°C for five minutes. Subsequently, 72 microliters of TE buffer (10mM Tris, 1 mM EDTA, pH7.5) was added to each reaction and the samples were gently vortexed for 15 seconds.

Following ligation, PCR was used to clone integration sites. First-round PCR was run using an outer adaptor primer (AP1 from the GW kit) and a custom primer, GagR2. AP1 is a forward primer located in the adaptor and GagR2 is a reverse primer located immediately downstream of the 3' end of the U5 LTR of the virus. This primer was chosen so the U3 LTR end of the virus or an internal viral fragment would not be cloned. Second-round PCR (nested) was run using an inner adaptor primer (AP2 from the GW kit) and a second custom primer, U3RU2. The U3RU2 primer is a reverse primer located in the viral LTR immediately downstream of the 5' end of the U5 LTR. Nested PCR will increase the integration site copies in number and specificity. The primer sequences are as follows: AP1, 5'-GTA ATA CGA CTC ACT ATA GGG C-3'; GagR2, 5'-TTT TGG CGT ACT CAC CAG TCG-3'; AP2, 5'-ACT ATA GGG CAC GCG TGG T-3'; and U3RU2, 5'-TGA GGG ATC TCT AGT TAC CAG AGT-3'.

Each first-round PCR reaction contained the following: 35.8 microliters nuclease-free water, 1 microliter each of AP1 and GagR2 (10 µM) primers, 3 microliters DNA template, 5 microliters 10X PCR buffer, 4 microliters dNTPs (10 mM each), 0.2 microliters TaKaRa Ex Taq DNA polymerase (cat. # RR001A; TaKaRa Bio Company, Madison, WI). The PCR parameters were as follows: a pre-denaturation step for 5 minutes at 94°C, 30 cycles of 94°C for 1 minute, 55°C for 45 seconds, 72°C for 3 minutes, and a final extension step at 72°C for 5 minutes.

Second-round PCR was performed using 1 microliter of the first-round PCR products as the DNA template. All components remained the same except the volume of the nuclease-free water was increased to 37.8 microliters to reach a final volume of 50 microliters and the primers were exchanged for AP2 and U3RU2. The cycling parameters remained the same. To confirm the presence of PCR product, 10 microliters of each sample were electrophoresed on a 1.5% agarose gel.

3.4. Cloning using the Invitrogen™ TOPO TA cloning® kit

Each cloning reaction contained 4 microliters of PCR product, 1 microliter of TOPO® vector (cat # K452001; Invitrogen, Carlsbad, CA), and 1 microliters salt solution in a 0.2 mL PCR tube. The components were mixed gently and incubated at room temperature for 30 minutes.

Two microliters of the TOPO® cloning reaction were added to one vial of thawed OneShot® Chemically Competent *E. coli* cells provided with the kit. The vial contents were gently mixed and incubated on ice for 20 minutes. Following incubation, the competent cells were subjected to heat shock for 30 seconds in a 42°C water bath and then immediately returned to ice. To each vial, 250 microliters of S.O.C. media was added and the samples were placed in a 37°C shaking incubator for 1 hour.

The transformed bacterial cultures were plated on warmed Miller LB agar plates (supplemented with 100 micrograms/mL ampicillin) after 40 microliters of X-gal (40 mg/mL in DMF) was added. After the cultures were spread evenly across the plate, the plates were incubated at 37°C overnight. Non-blue colonies were picked off each plate using a pipette tip and transferred to a 5mL round-bottom tube containing 3 mL of Miller LB broth supplemented with 100 micrograms/mL ampicillin for expansion. These tubes were placed in a 37°C shaking incubator overnight.

The bacterial cultures were transferred to 1.7 mL microcentrifuge tubes and were centrifuged at 6800x g to pellet the cells. The supernatant was removed and plasmid DNA extraction was performed on each sample using a QIAGEN QIAprep® Spin Miniprep Kit (cat # 27106; QIAGEN, Hilden, Germany). The samples were eluted with 50 microliters nuclease-free water and the concentration was measured using a UV spectrophotometer.

Extracted DNA was sequenced by the Sidney Kimmel Nucleic Acid Facility at the Kimmel Cancer Center at Thomas Jefferson University (Philadelphia, PA). Each sequencing reaction contained 0.4 micrograms of plasmid DNA, 3.2 pmol of M13 forward primer (sequence: 5'-GTA AAA CGA CGG CCA G-3'), and nuclease-free water to 12 microliters. The facility uses a 3730 DNA analyzer and BigDye® Terminator Cycle Sequencing kits (Applied Biosystems, Foster City, CA).

3.5. Integration site analysis

Sequences were analyzed using the BLAT program (University of California, Santa Cruz, CA; Human Genome Project Working Draft February 2009 Freeze; http://www.genome.ucsc.edu/cgi-bin/hgBlat). Sequences were considered true integration sites if the genomic portion was positioned between a recognizable adaptor sequence and the correct viral LTR end. Also, the genomic sequence must have matched with ≥ 98% identity, must not have any other high probability matches with almost equal length, and must be present in one specific site in the genome. Length away from the centromere was calculated for each site as follows using the values for the corresponding chromosome. If the integration site was on the p arm: p arm location ÷ entire length of p arm − 1. If the integration site were on the q arm: (q arm location − entire
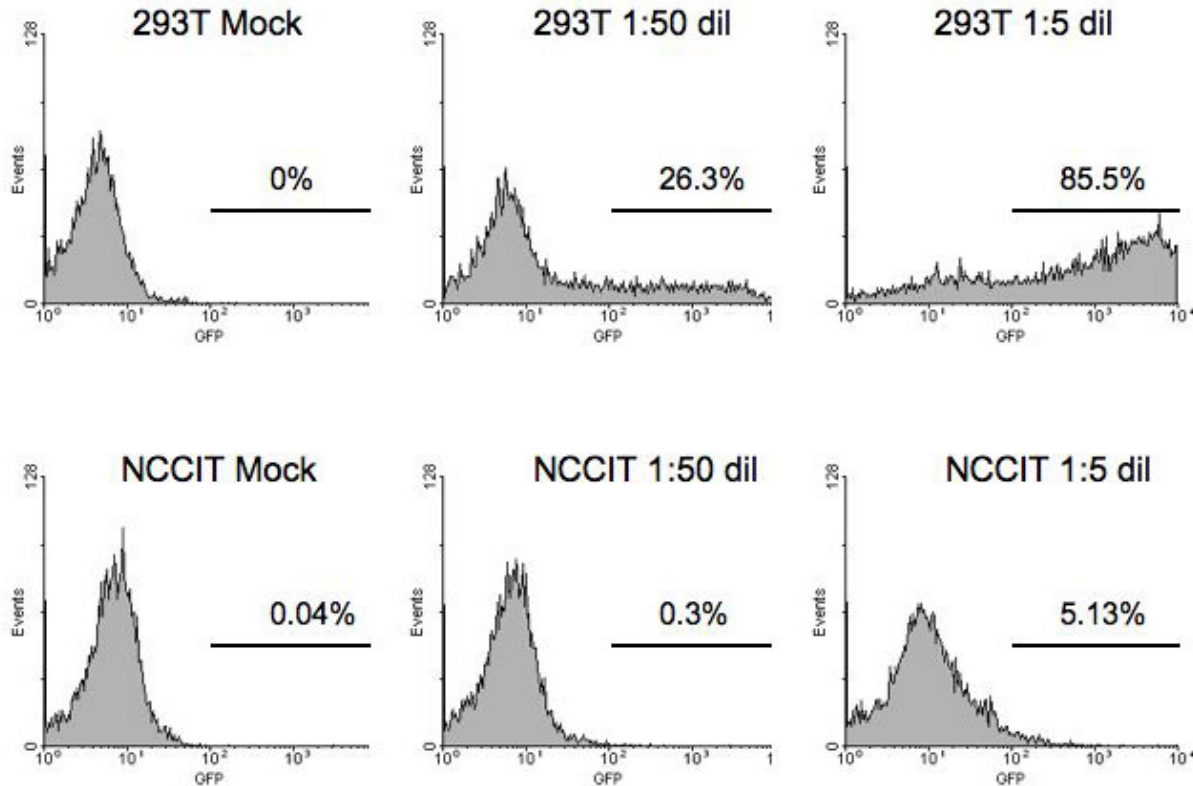
Figure 1. Comparison of the expression of a vector-transduced marker in NCCIT and 293T cells. NCCIT and 293T cells were infected as described with an HIV-1-based vector carrying the EGFP marker, at two different dilutions of the vector, as indicated. Expression of the EGFP gene was measured by flow cytometry 2 days post infection.

length of p arm) ÷ entire length of q arm. Additional features of our sequences were analyzed as described (28).

In addition to the integration site sequences generated in our study, other integration site sequences were analyzed using the above described method. Schroder and colleagues produced the sequences for integration in a T cell line (13) and Shun and colleagues produced sequences for integration in a murine embryonic fibroblast (MEF) cell line and a MEF LEDGF/p75 deficient cell line (19). Integration sites in 293T cells were analyzed by us (28).

3.6. Statistics

All statistical analyses were performed using the Fisher one-tailed test.

4. RESULTS

Several factors complicate analysis of integration sites in ES-like cells. First, these cells need to be grown on a feeder layer of somatic cells (23). This fact results in an undesirable background due to infection of the feeder layer cells. Second, it has been shown that transduction efficiency of stem cells is low, due to blocks prior to integration (29). In addition, retroviral expression is known to be suppressed in stem cells by epigenetic mechanisms (30-32). NCCIT cells, unlike ES or iPS cells do not need a feeder layer, which simplifies the analysis. However, we have observed that expression of a vector-transduced marker in NCCIT cells is much lower than that in somatic cells (Figure 1). This is probably due to a dual effect of low transduction efficiency and suppressed expression. However, this finding again emphasizes the similarities between NCCIT and other types of pluripotent cells. To overcome the low transduction efficiency, we have infected NCCIT cells at a high m.o.i. (5-20). We have identified 98 integration sites in these cells and analyzed them as described below.

We first analyzed the distribution of integration sites among individual chromosomes. As shown in Figure 2, vector integration sites were distributed broadly in the genome, and we did not detect any particular differences when we compared the pattern to that described for the human SupT1 line (13), or to mouse embryonic fibroblasts (MEFs), both LEDGF/p75-proficient and -deficient ((19), Figure 2). We then analyzed the frequency of integration in genes. We found that 80.6% of all integrations occurred in genes. This finding indicates that the LEDGF/p75-dependent targeting into genes still occurs in NCCIT cells. However, during the analysis, we observed a high frequency of integration events in the vicinity of chromosomal ends (Figure 3). As shown in Figure 3, we found that almost 30% of all integration sites were present in the 1/10 of the chromosome that is closest to the
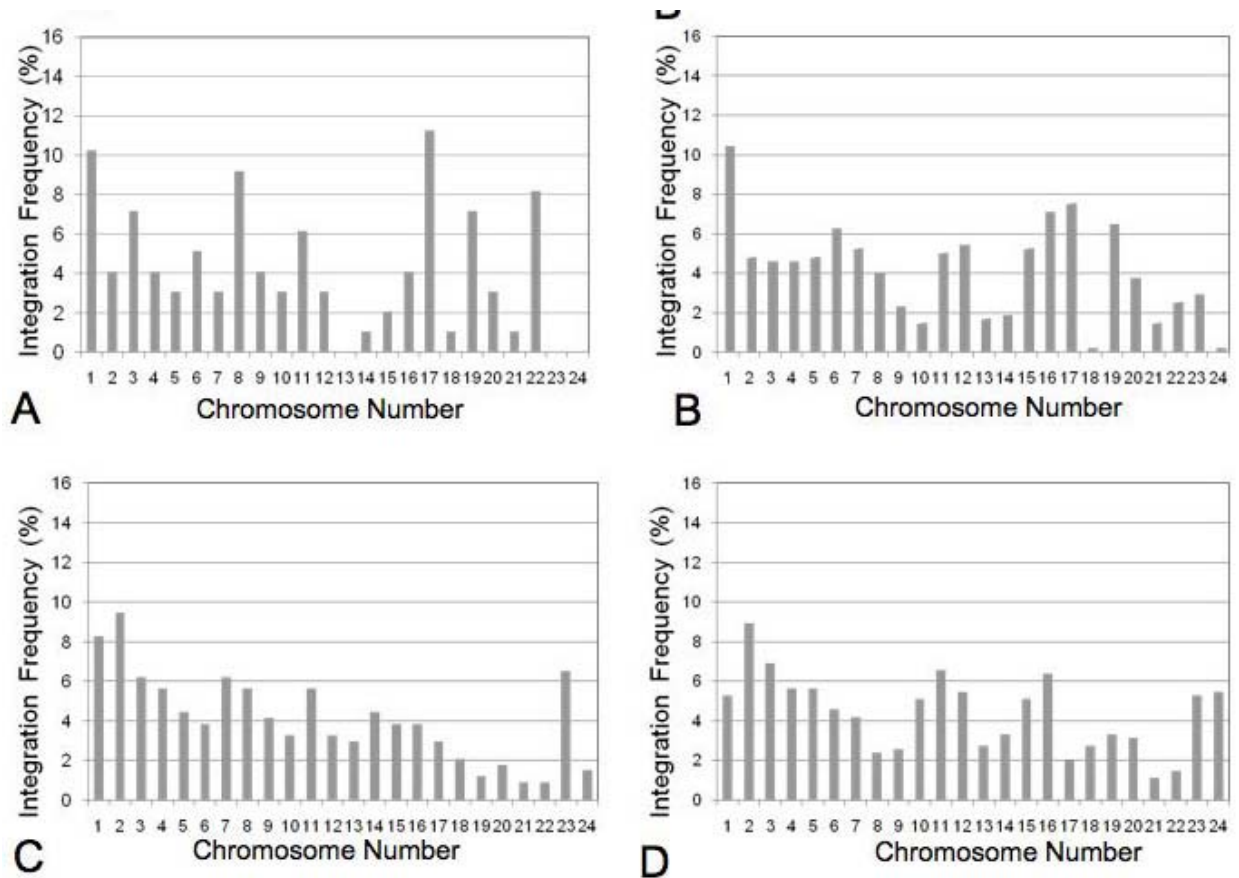
Figure 2. Comparison of the chromosomal distribution of integration sites between NCCIT, SupT1 and MEF cells. Integration sites were analyzed as described in the Results and Methods and integration sites were then assigned to individual chromosomes (x-axis). 23 = X chromosome, 24 = Y chromosome. y-axis – percentage of integration sites per chromosome. A. NCCIT cells. B. SupT1 cells [Schroder *et al.* (13)]. C. LEDGF/p75-proficient MEFs [Shunt *et al.* (19)]. D. LEDGF/p75-deficient MEFs [Shun *et al.* (19)].

telomere. To determine if this feature can be found in other cell types, we have analyzed integration sites in the SupT1 line (13). We found that only 14.9% of integration sites are present in this chromosomal segment (Figure 4). The difference was statistically significant, with p=0.005 (Fisher one-tailed test).

One possible explanation for the differences is that these could be due to the difference in methods used to clone integration sites. Although both labs employed LM-PCR, differences in primers or enzymes used to digest genomic PCR could affect the results. Thus, we compared NCCIT integration sites to those we cloned from 293T cells, using identical methodology. Figure 5 shows that the integration pattern in 293T cells resembled that of SupT1 cells (Figure 4) and was still different from NCCIT cells. The difference in the fraction of integration sites in the last 1/10 of the chromosome between NCCIT and 293T cells was again statistically significant (p=0.0489).

We observed that in NCCIT cells, chromosomal ends appear to be preferred targets for integration of HIV-1-based vectors, whereas in somatic cells integration sites

appear to be distributed evenly along the whole length of the chromosomes (see above). As noted above, one of the major factors controlling HIV-1 integration site selection is LEDGF/p75. One possible explanation for our data is that a cellular factor, present in somatic cells, but missing in stem cells, targets integration away from chromosomal ends in somatic cells by binding and controlling cellular distribution of LEDGF/p75. If so, a loss of LEDGF/p75 in somatic cells should result in redistribution of integration sites toward chromosomal ends. To test this hypothesis, we examined integration sites in LEDGF/p75-deficient MEFs that were presented in Shun *et al.* As shown in Figure 6, LEDGF/p75 deficiency resulted in a redistribution of integration sites, with some increase in frequency of integration in the last 1/10 of chromosomes. Since this finding appeared to be consistent with our hypothesis, we have further analyzed the distribution of integration sites, and examined integration sites in p and q arms separately. We observed that in LEDGF/p75-deficient MEFs, HIV-1-based vectors integrate preferentially near p arm ends, whereas in NCCIT cells, near the ends of q arms (Figure 7). Thus, we conclude that our results do not support the original hypothesis, and the integration pattern in NCCIT
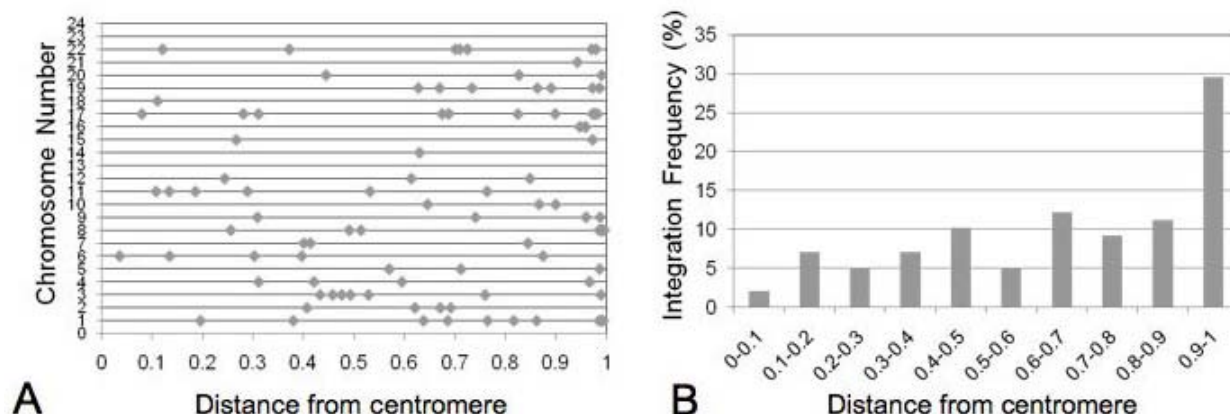
Figure 3. Distribution of integration sites along chromosomes in NCCIT cells. A. Distribution of integration sites along individual chromosomes. Individual chromosomes were divided into segments, each containing 0.1 of chromosome DNA, 0 signifies the position of the centromere, whereas 1 signifies the telomere. 23 = X, 24 = Y. B. Distribution of integration sites in each chromosomal segment. Integration sites in corresponding segments of individual chromosomes were pooled and a fraction of integration sites in a given chromosome segment was calculated (y axis).
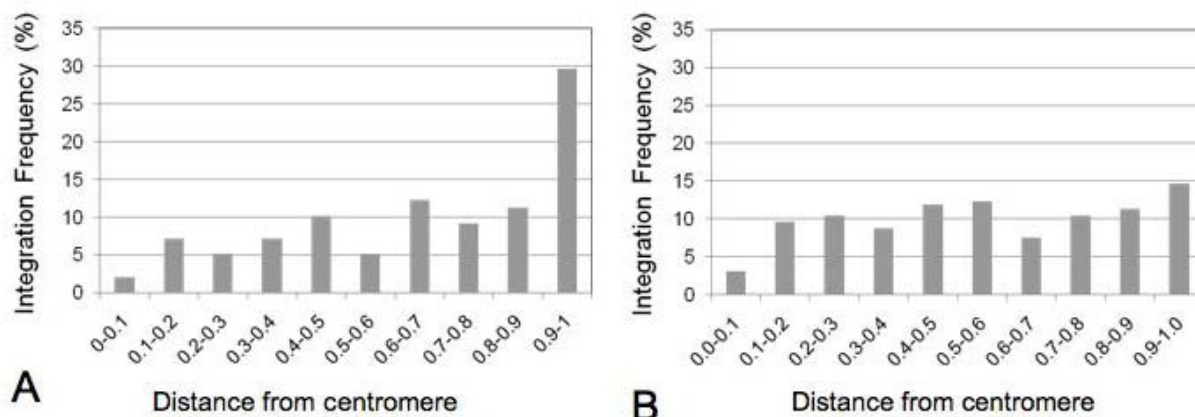

Figure 4. Comparison of the distribution of integration sites along chromosomes of NCCIT and SupT1 cells. A. NCCIT cells. B. SupT1 cells. Integration sites were analyzed as outlined in Figure 3B.

cells is not likely due to a LEDGF/p75-binding cellular factor that is missing in these cells.

Finally, we examined the distribution of integration sites in NCCIT cells with respect to sequence features of human DNA (Table 1). First, we examined the frequency of integration events in long interspersed nuclear elements (LINEs), which constitute a high fraction of the human genome. Next, we examined integrations in short interspersed nuclear elements (SINEs), and LTR repeat elements (human endogenous retroviruses, HERVs). We also compared the frequency of integration in these sequences with results reported by Schroder *et al.*, who first reported the preferences of HIV-1 and HIV-1-based vectors for integration in other elements of the genome (13). Our data indicate that the fraction of integration sites in SINEs and LINEs of NCCIT cells largely corresponded to the frequency of integration in these elements in somatic cells (Table 1). We also noted a high number of integration sites in Alu elements (a type of SINE), which occur frequently in genes, consistent with the published results for somatic

cells (13). Finally, we have analyzed the frequency of integration in the vicinity of CpG islands and upstream and downstream of genes (Table 2). Again we did not observe significant differences between NCCIT and somatic cells.

5. DISCUSSION

In this study, we demonstrated that in a model pluripotent stem cell line, integration sites of HIV-1-based vectors are distributed in a unique pattern, with a plurality of integration events occurring near chromosomal ends. We further demonstrate that integration still occurs predominantly in genes, indicating that LEDGF/p75 still plays a major role in targeting in this cell type.

Integration site selection by retroviral vectors, including HIV-1-based, has recently attracted tremendous attention, due to adverse events occurring in clinical trials (12, 33). Thus, it is necessary to identify factors that influence integration site selection, with the objective of eventually controlling where vector DNA integrates. Our results show that in addition to LEDGF/p75, there are

Targeting integration

Table 1. Data comparison with previous studies[1]

Chromosomal Feature	Schroder *et al* (2002) [13]	Percentage in 293T Integration Sites	Percentage in NCCIT Integration Sites vs. Schroder	Percentage in NCCIT Integration Sites vs. 293T
Genes	69%	72.4% (p=0.792)	80.6% (p=0.0126)	80.6% (p=0.1069)
LINE	17%	7.8% (p=0.998)	10.2% (p=0.0579)	10.2% (p=0.3486)
Alu (SINE)	15.9%	21.6% (p=0.0912)	16.3% (p=0.5023)	16.3% (p=0.2142)
MIR (SINE)	0.7%	0% (p=1.000)	2.0% (p=0.2410)	2.0% (p=0.2487)
DNA Elements	2.2%	5.2% (p=0.088)	8.2% (p=0.0070)	8.2% (p=0.3913)
LTR Elements (HERV)	3.7%	5.2% (p=0.291)	4.1% (p=0.5032)	4.1% (p=0.3737)
Satellite	0.4%	0% (p=0.670)	0% (p=7095)	0% (p=1.000)
Total Cloned Integration Sites	524	116	98	98

[1]p Values are a comparison of each population from this study against data from Schroder *et al.* (2002) [13] or Silvers *et al.* (2010)[28], using Fisher's exact one-tailed test.

Table 2. Integration into various chromosomal regions

	Percentage in 293T Integration Sites	Percentage in NCCIT Integration Sites
Within +/- 1 kb of CpG Islands	1.7%	3.1% (p=0.4208)[1]
Within 5 kb upstream of genes	6%	21.1% (p=0.1289)[1]
Within 5 kb downstream of genes	3.4%	5.3% (p=0.6110)[1]
Total Cloned Integration Sites	116	98

[1]p values are a comparison of 293T and NCCIT integration sites, using Fisher's exact one-tailed test
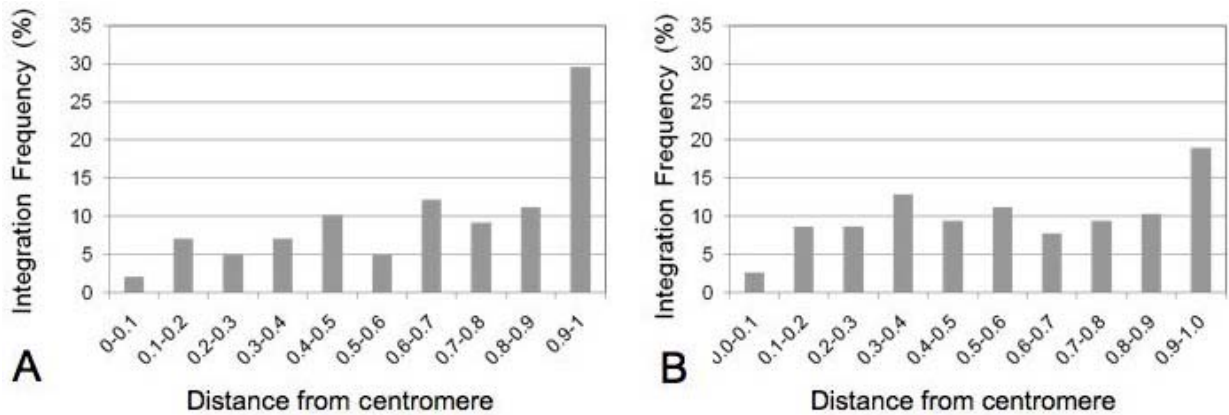

Figure 5. Comparison of the distribution of integration sites along chromosomes of NCCIT and 293T cells. A. NCCIT cells. B. – 293T cells. Integration sites were analyzed as outlined in Figure 3B.
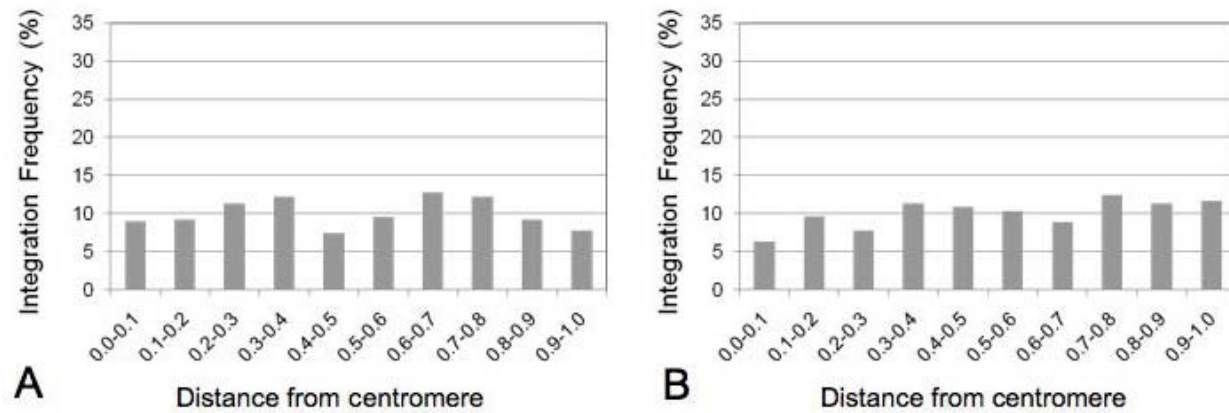

Figure 6. Comparison of the distribution of integration sites along chromosomes of LEDGF/p75-proficient and –deficient MEFs. A. LEDGF/p75-proficient cells. B. LEDGF/p75-deficient cells. Integration sites as published in Shun *et al.* (19).
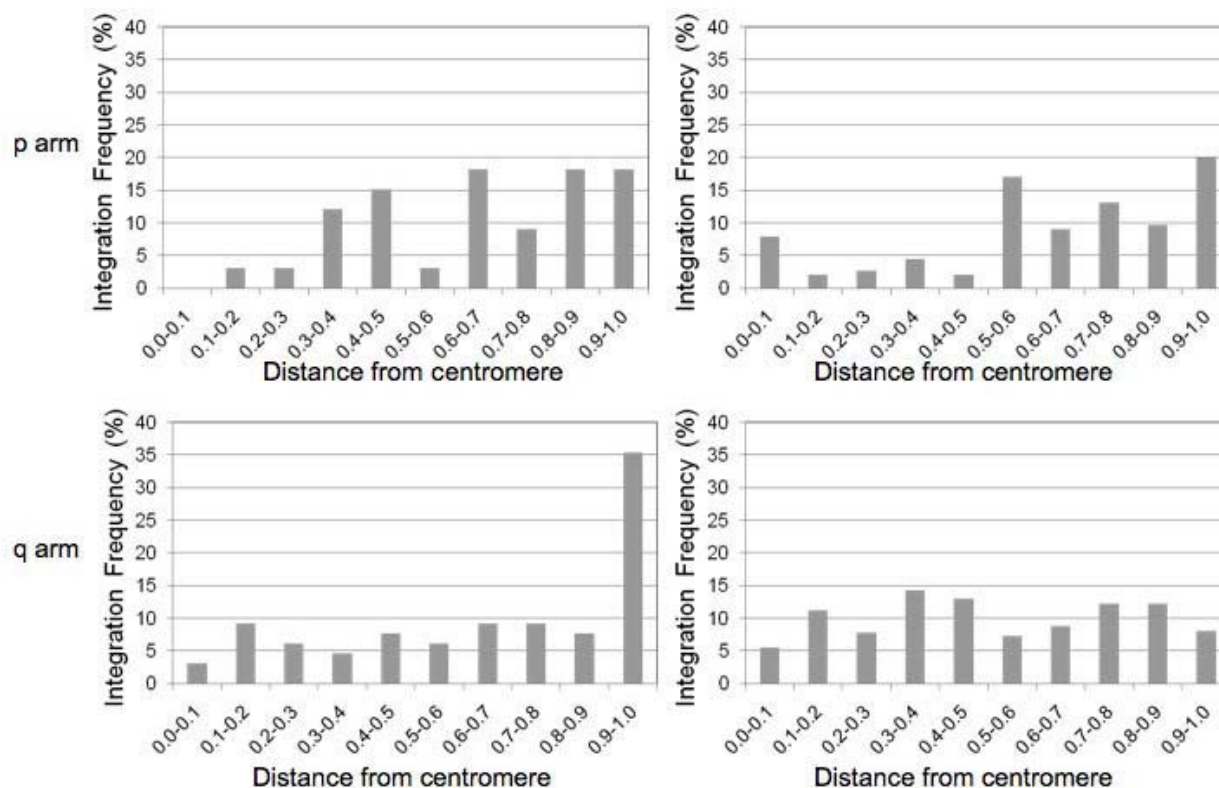
Figure 7. Comparison of the distribution of integration sites in the p and q arms of NCCIT and LEDGF/p75 –deficient MEFs. Left graphs – NCCIT cells, right graphs – LEDGF/p75-deficient cells. Top graphs – distribution of integration sites in p arms, bottom graphs – distribution of integration sites in q arms.

unknown factors that influence targeting toward large chromosomal segments, irrespective of whether integration occurs in genes.

What could be the reason for the observed accumulation of integration sites in the vicinity of chromosomal ends in pluripotent cells? These cells are defined by the expression of stem cell transcription factors (see the Introduction), but they possess other unique features that distinguish them from somatic cells. These features include widespread DNA demethylation, a lack of lamin A/C and expression of other, stem cell specific genes (21). Moreover, mutations in lamin A/C were associated with the relocalization of chromosomes within cells, with telomeres moving into the periphery of the nucleus (34). It will be intriguing to determine which of these factors affects integration site selection in these cells. We note that understanding how integration site selection occurs is necessary to achieve targeting of retroviral vectors to predetermined, "safe" chromosomal regions, away from genes. This will lead to increasing the safety of clinical gene therapy trials. Finally, we note that our data indicate that there are differences in integration targeting between somatic and pluripotent cells. This is a crucial observation, since a recently developed type of pluripotent cells, iPS cells, is an attractive target for gene therapy and gene transfer. It is thus possible that newly developed methods of integration targeting will

have to be tailored to individual cell types, which are the subject of gene transfer.

7. REFERENCES

1. Coffin, J. M., Hughes, S. H., and Varmus, H. E., *Retroviruses*, Cold Spring Harbor Laboratory Press, Plainview, NY (1997)

2. Flint, S. E., LW; Racaniello, VR; Skalka, AM, *Principles of Virology*, ASM Press, Washington, DC (2004)

3. Skalka, A., Retroviral DNA Integration, *52* (1999)

4. Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L., and Trono, D., Multiply attenuated lentiviral vector achieves efficient gene delivery *in vivo*, *Nat Biotechnol, 15*, 871 (1997)

5. Lewis, P., Hensel, M., and Emerman, M., Human immunodeficiency virus infection of cells arrested in the cell cycle, *EMBO J, 11*, 3053 (1992)

6. Weinberg, J. B., Matthews, T. J., Cullen, B. R., and Malim, M. H., Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes, *J Exp Med, 174*, 1477 (1991)

7. Barsov, E. V., and Hughes, S. H., Gene transfer into mammalian cells by a Rous sarcoma virus-based retroviral vector with the host range of the amphotropic murine leukemia virus, *J Virol, 70*, 3922 (1996)

8. Daniel, R., Katz, R. A., and Skalka, A. M., A role for DNA-PK in retroviral DNA integration, *Science, 284*, 644 (1999)

9. Katz, R. A., Greger, J. G., Darby, K., Boimel, P., Rall, G. F., and Skalka, A. M., Transduction of interphase cells by avian sarcoma virus, *J Virol, 76*, 5422 (2002)

10. Mitchell, R. S., Beitzel, B. F., Schroder, A. R., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R., and Bushman, F. D., Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences, *PLoS Biol, 2*, E234 (2004)

11. Wu, X., Li, Y., Crise, B., and Burgess, S. M., Transcription start regions in the human genome are favored targets for MLV integration, *Science, 300*, 1749 (2003)

12. 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., Sorensen, R., Forster, A., Fraser, P., Cohen, J. I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L. E., Wissler, M., Prinz, C., Rabbitts, T. H., Le Deist, F., Fischer, A., and Cavazzana-Calvo, M., LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1, *Science, 302*, 415 (2003)

13. Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F., HIV-1 integration in the human genome favors active genes and local hotspots, *Cell, 110*, 521 (2002)

14. Narezkina, A., Taganov, K. D., Litwin, S., Stoyanova, R., Hayashi, J., Seeger, C., Skalka, A. M., and Katz, R. A., Genome-wide analyses of avian sarcoma virus integration sites, *J Virol, 78*, 11656 (2004)

15. Cherepanov, P., Devroe, E., Silver, P. A., and Engelman, A., Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase, *J Biol Chem, 279*, 48883 (2004)

16. Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z., HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells, *J Biol Chem, 278*, 372 (2003)

17. Llano, M., Saenz, D. T., Meehan, A., Wongthida, P., Peretz, M., Walker, W. H., Teo, W., and Poeschla, E. M., An essential role for LEDGF/p75 in HIV integration, *Science, 314*, 461 (2006)

18. Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y., LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells, *J Biol Chem, 278*, 33528 (2003)

19. Shun, M. C., Raghavendra, N. K., Vandegraaff, N., Daigle, J. E., Hughes, S., Kellam, P., Cherepanov, P., and Engelman, A., LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration, *Genes Dev, 21*, 1767 (2007)

20. Vanegas, M., Llano, M., Delgado, S., Thompson, D., Peretz, M., and Poeschla, E., Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering, *J Cell Sci, 118*, 1733 (2005)

21. Okita, K., and Yamanaka, S., Intracellular signaling pathways regulating pluripotency of embryonic stem cells, *Curr Stem Cell Res Ther, 1*, 103 (2006)

22. Taranger, C. K., Noer, A., Sorensen, A. L., Hakelien, A. M., Boquest, A. C., and Collas, P., Induction of dedifferentiation, genomewide transcriptional programming, and epigenetic reprogramming by extracts of carcinoma and embryonic stem cells, *Mol Biol Cell, 16*, 5719 (2005)

23. McDevitt, T. C., and Palecek, S. P., Innovation in the culture and derivation of pluripotent human stem cells, *Curr Opin Biotechnol, 19*, 527 (2008)

24. Park, I. H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M. W., Cowan, C., Hochedlinger, K., and Daley, G. Q., Disease-specific induced pluripotent stem cells, *Cell, 134*, 877 (2008)

25. Damjanov, I., Horvat, B., and Gibas, Z., Retinoic acid-induced differentiation of the developmentally pluripotent human germ cell tumor-derived cell line, NCCIT, *Lab Invest, 68*, 220 (1993)

26. Mancino, M., Esposito, C., Watanabe, K., Nagaoka, T., Gonzales, M., Bianco, C., Normanno, N., Salomon, D. S., and Strizzi, L., Neuronal guidance protein Netrin-1 induces differentiation in human embryonal carcinoma cells, *Cancer Res, 69*, 1717 (2009)

27. Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., Verma, I. M., and Trono, D., *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector, *Science, 272*, 263 (1996)

28. Silvers, R. M., Smith, J. A., Schowalter, M., Litwin, S., Liang, Z., Geary, K., and Daniel, R., Modification of integration site preferences of an HIV-1-based vector by

expression of a novel synthetic protein, *Hum Gene Ther, 21*, 337.

29. Santoni de Sio, F. R., Gritti, A., Cascio, P., Neri, M., Sampaolesi, M., Galli, C., Luban, J., and Naldini, L., Lentiviral vector gene transfer is limited by the proteasome at postentry steps in various types of stem cells, *Stem Cells, 26*, 2142 (2008)

30. Cherry, S. R., Biniszkiewicz, D., van Parijs, L., Baltimore, D., and Jaenisch, R., Retroviral expression in embryonic stem cells and hematopoietic stem cells, *Mol Cell Biol, 20*, 7419 (2000)

31. Matsui, T., Leung, D., Miyashita, H., Maksakova, I. A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M. C., and Shinkai, Y., Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET, *Nature, 464*, 927.

32. Rowe, H. M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P. V., Layard-Liesching, H., Verp, S., Marquis, J., Spitz, F., Constam, D. B., and Trono, D., KAP1 controls endogenous retroviruses in embryonic stem cells, *Nature, 463*, 237.

33. Daniel, R., and Smith, J. A., Integration site selection by retroviral vectors: molecular mechanism and clinical consequences, *Hum Gene Ther, 19*, 557 (2008)

34. Taimen, P., Pfleghaar, K., Shimi, T., Moller, D., Ben-Harush, K., Erdos, M. R., Adam, S. A., Herrmann, H., Medalia, O., Collins, F. S., Goldman, A. E., and Goldman, R. D., A progeria mutation reveals functions for lamin A in nuclear assembly, architecture, and chromosome organization, *Proc Natl Acad Sci U S A* (2009)

Send correspondence to: Rene Daniel, Division of Infectious Diseases - Center for Human Virology, Thomas Jefferson University, Philadelphia, Tel: 215-503-5725, Fax: 215-923-1956, E-mail: Rene.Daniel@jefferson.edu