

Research Article

Integration of HIV in the Human Genome: Which Sites Are Preferential? A Genetic and Statistical Assessment

Juliana Gonçalves,¹ Elsa Moreira,² Inês J. Sequeira,² António S. Rodrigues,¹ José Rueff,¹ and Aldina Brás¹

¹Centre for Toxicogenomics and Human Health (ToxOmics), Genetics, Oncology and Human Toxicology, Nova Medical School/Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Rua Câmara Pestana 6, 1150-008 Lisbon, Portugal

²CMA, Department of Mathematics, Faculty of Sciences and Technology, Universidade Nova de Lisboa, Campus Caparica, 2829-516 Caparica, Portugal

Correspondence should be addressed to Aldina Brás; aldina.bras@nms.unl.pt

Received 11 February 2016; Accepted 24 April 2016

Academic Editor: Margarita Hadzopoulou-Cladaras

Copyright © 2016 Juliana Gonçalves et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chromosomal fragile sites (FSs) are loci where gaps and breaks may occur and are preferential integration targets for some viruses, for example, Hepatitis B, Epstein-Barr virus, HPV16, HPV18, and MLV vectors. However, the integration of the human immunodeficiency virus (HIV) in Giemsa bands and in FSs is not yet completely clear. This study aimed to assess the integration preferences of HIV in FSs and in Giemsa bands using an *in silico* study. HIV integration positions from Jurkat cells were used and two nonparametric tests were applied to compare HIV integration in dark versus light bands and in FS versus non-FS (NFSs). The results show that light bands are preferential targets for integration of HIV-1 in Jurkat cells and also that it integrates with equal intensity in FSs and in NFSs. The data indicates that HIV displays different preferences for FSs compared to other viruses. The aim was to develop and apply an approach to predict the conditions and constraints of HIV insertion in the human genome which seems to adequately complement empirical data.

1. Introduction

Giemsa staining has long been used for identifying individual human chromosomes. Giemsa dark and light bands are generally thought to correspond to GC-poor and GC-rich regions, respectively. Giemsa light bands are gene-rich and contain most housekeeping genes as well as a large number of CpG islands, whereas Giemsa dark bands are gene-poor and preferentially contain tissue-specific genes. Hence, light bands are transcriptionally more active when compared to dark bands and also have an open chromatin configuration which together with the high content in GC can have an important role in provirus integration [1]. Giemsa bands are also related to functional nuclear processes such as replication. For example, DNA replication timing during cell cycle differs between both; light bands are early-replicating,

whereas dark bands are late replicating. Giemsa bands are also related to chromatin structures as the chromatin in dark bands is more condensed than in light bands during both metaphase and interphase [2]. Another difference between these two Giemsa bands is that the DNA of Giemsa dark bands are located at the nucleus periphery [3] while the DNA of Giemsa light bands is in the interior of the nucleus [4].

FSs are hereditary loci of human chromosomes susceptible to occurrence of breaks, gaps, or rearrangements when under stress conditions or treated with specific chemical agents [5–7]. According to the frequency of their distribution in the human population, FSs can be divided in two distinct groups: common fragile sites (CFSs) present in all individuals and rare fragile sites (RFSs) that are present in less than 5% of the population, and these two groups can also be subdivided according to the inducing agent [8–10]. Both types of FSs have

the capacity to form secondary structures that can interfere with elongation in replication [11] or even cause failure in chromatin condensation [12]. FSs are also involved in sister chromatid exchanges [7] translocations and deletions [13] and in intrachromosomal gene amplifications [14]. CFSs are very unstable regions because they contain sequences of high flexibility [15] and are regions of late replication [5] and also correspond to transition regions in replication timing [16]. Several authors have shown that fragile sites are preferential integration targets for some viruses, for example, Epstein-Barr virus [17] and human papillomaviruses HPV16 and HPV18 [18]. Recently, Christiansen et al. showed that transcriptionally active regions and FSs are the preferred targets for chromosomal HPV integration in cervical carcinogenesis [19].

The human immunodeficiency virus (HIV) is a retrovirus whose stable integration in the human genome is essential for completing its life cycle [20, 21]. The virus binds to the membrane receptors of host cells to enter the cytoplasm [22]. The RNA genome of HIV is converted into DNA by the reverse transcriptase (RT) enzyme [23] that is transported to the nucleus. Viral integration into the DNA occurs in three steps, (i) processing [24]; (ii) joining [25]; and (iii) postintegration repair [26]. After integration, transcription occurs followed by translation in the host cytoplasm. There are two different types of viruses: HIV-1, identified first in 1983 [27], and HIV-2, later discovered in 1986 [28]. Viral integration in human cells can affect gene expression, leading to molecular and epigenetic alterations, and can even activate oncogenes [22]. Thus knowledge of viral integration sites is important to understand their biological effects. Schröder et al. concluded that integration sites of HIV are not randomly distributed in the human genome, but in regional hotspots [29]. The same group also found that integration sites are related to gene-rich sites which can allow a more efficient expression of the viral genome. Moreover, according to Debyser et al. [30] each retroviral family integrates near a unique and specific subset of genomic features. HIV integration site selection is related to the pathway and efficiency of nuclear translocations [31].

Retroviruses can be used as vectors in gene therapy since they can integrate stably in the host genome [30, 32] and they have the capacity to introduce genetic material in target cells [33]. Additionally, analysis of the integration process is important in HIV-induced disease. Therefore, it is important to understand the integration preferences of HIV since knowledge of the integration sites in the human genome can help, for example, to choose gene-delivery vectors [29].

The integration preferences of HIV in light or dark bands and in FSs are not clear. Thus, our main objective was to study the HIV integration preferences in Giemsa bands and in FSs using bioinformatics and statistical analyses. More specifically we aimed to understand the integration preferences of HIV-1 isolated from Jurkat T cells in Giemsa bands and FSs by the use of an *in silico* approach based on statistical analysis which may complement laboratorial studies and predict HIV constraints and preferences of integration in the human genome.

2. Methods

2.1. Data

2.1.1. HIV Integration Sites. For the HIV-1 isolated from Jurkat T cells Wang et al. have supplied the exact position of the integration sites [34].

2.1.2. Giemsa Bands and FSs. The positions of Giemsa bands used in this study were obtained *in silico* by Niimura and Gojobori [2], available from the Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, in <http://yosniimura.net/research/coordinates.html>.

Regarding FSs, the human genome was divided in FRs and NFRs, according to their positions. A list of FSs was obtained from Mrasek et al. [6] and completed with FSs from Lukusa and Fryns [8]. Two consecutive bands associated with FSs were grouped to form a fragile region (FR) and a region between two FRs was considered a NFR [35, 36]. The Y chromosome was not considered because it does not have well defined FSs; there is only a possibility of existing one FS [37].

2.2. Statistical Analysis. Integration sites of HIV-1 were colocalized with Giemsa dark bands and classified in two groups: *yes* if they colocalized and *no* if they did not. Some integrations occurring in the centromeres and short arm of the acrocentric chromosomes were excluded from the total. Thus for HIV-1 we have integrations in 24 chromosomes in a total of 42912 integration sites. Then a measure was calculated to determine in which type of band the virus integrates. The measure is designated as *integration intensity number* and is given by

$$\begin{aligned} i_{\text{dark band}} &= \frac{n_{\text{yes}}}{l_{\text{dark band}}}; \\ i_{\text{light band}} &= \frac{n_{\text{no}}}{l_{\text{light band}}}, \end{aligned} \quad (1)$$

where n_{yes} represent the number of viral integrations in dark bands, $l_{\text{dark band}}$ is the length of the dark bands, n_{no} is the number of viral integrations in light bands, and $l_{\text{light band}}$ is the length of the light bands.

After the calculations of the measure, one pair (x, y) was obtained for each chromosome, where x represents the measurement value in dark bands and y the measurement value in light bands. To compare dark bands with light bands graphical representations for the measure were constructed where each point represents one chromosome. In order to see the preferences of integration of the virus, in each graphic the line $y = x$ was represented that allows the visualization of the number of chromosomes in which $y > x$ and $y < x$, in other words, if the virus preferentially integrated in dark or in light bands. In order to statistically verify the results obtained graphically two nonparametric tests were applied that enable us to compare two dependent samples, the Sign and the Wilcoxon tests [38]. In both tests a significance level of 1% was used to test our hypothesis.

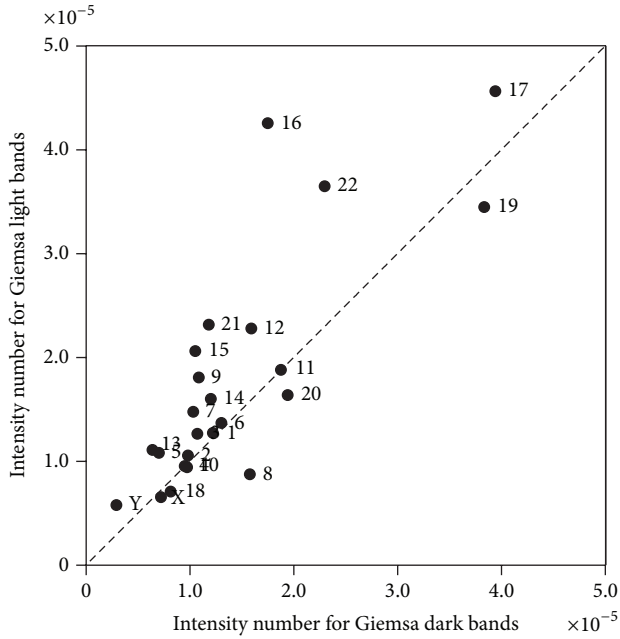


FIGURE 1: Integration of HIV-1 isolated from Jurkat T cells in Giemsa dark bands versus Giemsa light bands. Graphical representation of the results for the *intensity number*. Each point represents a chromosome whose coordinates are the (x, y) pairs obtained in the measure calculations. The statistical analysis indicates that the virus integrates preferentially in Giemsa light bands (Wilcoxon test, T_{obs} : 59; T_{critic} : 69).

The same methodologies were applied in FSs, also classifying the integrations in two groups: *yes* if they colocalized with a FR and *no* if they colocalized with a NFR. Integrations of HIV-1 for the 23 chromosomes were obtained, in a total of 44150 integrations sites.

3. Results and Discussion

3.1. Preferential Integration of HIV-1 in Giemsa Light Bands. The *intensity number* for the HIV-1 isolated from Jurkat T cells was calculated and the graphical representation is presented in Figure 1 for Giemsa bands. For the Wilcoxon test, T_{obs} (59) was lower than T_{critic} (69), so the hypothesis of equal intensity at 1% level can be clearly rejected. It means that the intensity of integration is not equal in the two types of bands and it can be concluded that HIV-1 isolated from Jurkat T cells integrates with more intensity in Giemsa light bands.

Our *in silico* results obtained for Giemsa bands indicate that HIV-1 isolated from Jurkat T cells integrates preferentially in light bands, which have a high content in GC [2]. These results are in line with previous studies which reveal that HIV favours integration in transcriptionally active units [34, 39, 40] which are associated with regions of high GC content and high gene density [24]. When integrated in transcriptionally active regions this ensures viral gene transcription [41, 42]. The virus has a limited time to replicate [24] since T cells infected with HIV have a very short half-time [43]. Thus the virus has to integrate in regions that accelerate

its transcription. Regions that are more transcriptionally active allow an efficient maintenance of the replication cycle of the virus since they permit a higher provirus transcription [44], increasing viral gene expression [34].

Elleder et al. also found that HIV integrates preferentially in Giemsa light bands and in regions with open chromatin which favours integration [44], since the IN enzyme of HIV-1 interacts with components of the chromatin remodeling complex [45]. The interaction with LEDGF/p75 accounts for the karyophilic properties and chromosomal targeting of HIV-1 IN [46]. Nevertheless, one must note that virus integration in transcriptionally active regions with an open chromatin conformation, that is, light bands, could also be a by-product of the integration of HIV-1 integrase with components of the chromatin remodeling complex, which could impact on our interpretation of HIV integration site preference.

Marini et al. also revealed that the cellular genes that are targeted by HIV-1 are enriched in open chromatin marks associated with the nuclear complex pore that are constituted by nucleoporins which participate in HIV-1 transcriptional regulation [47]. This report supports our results for the integration in Giemsa light bands that have a less condensed chromatin than Giemsa dark bands. Moreover, the same group of authors also found that the areas of open chromatin that are targeted by the HIV-1 preintegration complex are those proximal to the nuclear pore. HIV-1 could also have a preference for Giemsa light bands because they have active genes [48]; thus the preference for active gene regions may have been developed to favour HIV gene expression after integration [29]. Another feature that supports our results is that light bands possess high levels of histone acetylation, namely, of histones H3 and H4, which enable access of transcription factors [48, 49]. In this regard Wang et al. demonstrated that the frequency of integration of HIV is associated with epigenetic modifications including H3 and H4 acetylation [34].

3.2. FSs Are Not Preferential Integration Targets for HIV-1. Figure 2 shows the result of *intensity number* for the HIV-1 virus isolated from Jurkat T cells in FRs. For the Sign test a nonsignificant p value of 0.202 was obtained. The Wilcoxon test resulted in a T_{obs} of 82 and a T_{critic} of 62 which leads us to accept the hypothesis of equal integration intensity. Thus, HIV-1 isolated from Jurkat T cells integrates with equal intensity in FRs and NFRs.

In the methodology followed, two consecutive bands associated with FSs were grouped to form a FR and a region between two FRs was considered a NFR. We verified that HIV-1 isolated from Jurkat cells integrates with equal intensity in FRs and in NFRs. Other known viruses, such as Hepatitis B, Epstein-Barr virus [17], HPV16, and HPV18 [18], integrate more in FSs. These differences between HIV and the other viruses may be explained by the different phases of the cell cycle in which the virus enters in human cells. Pyeon et al. verified that progression in cell cycle through mitosis is critical to HPV infection [50] while HIV can infect nondividing cells [24, 31]. The fact that HIV does not have a tendency to integrate in FRs could be related to the

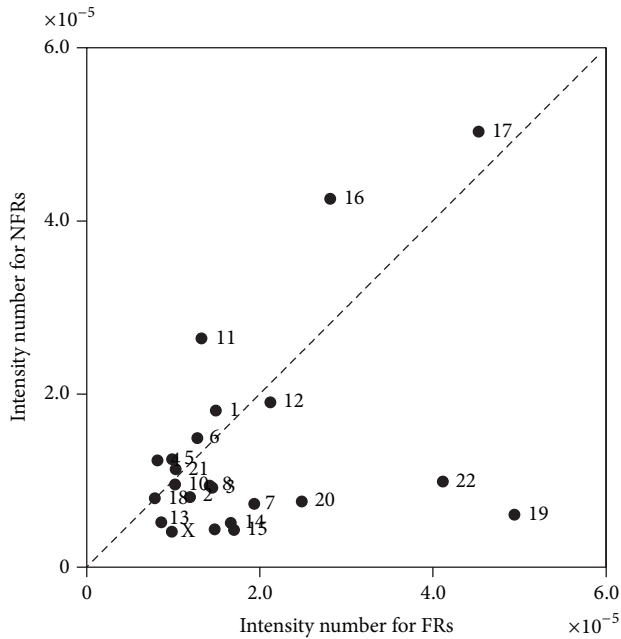


FIGURE 2: Integration of HIV-1 isolated from Jurkat T cells in FRs versus NFRs. Graphical representation of the results for the *intensity number*. Each point represents a chromosome whose coordinates are the (x, y) pairs obtained in the measure calculations. The statistical analysis indicates that the virus integrates with equal intensity in FRs and in NFRs (Sign test, p value: 0.202, and Wilcoxon test, T_{obs} : 82; T_{critic} : 62).

structure of the FSs which are vulnerable to DNA breaks [51], which may not facilitate viral integration or the conclusion of viral replication. Moreover, FSs have the tendency to form secondary structures, which interfere with replication [11] thus hampering viral integration and replication. Genomic instability is another characteristic of FSs [7] which does not favour viral integration as it needs to integrate its genome stably in hosts to complete its life cycle [21]. HIV integrates more in regions with GC content which is not the case of FSs that are predominantly constituted by AT [15, 52].

In both graphics for Giemsa bands and for FSs we found that there were some chromosomes that differ from the rest, for example, chromosome 17. This result is in agreement with the data by Soto et al. [53]. This chromosome is rich in protein coding genes and has a high CG content which indicates a high gene density, besides being rich in SINEs [54].

4. Conclusion

Our aim was to develop and apply an approach intended to predict the conditions and constraints of HIV-1 insertion in the human genome. At the present stage, our approach seems to adequately predict most of the conditions unravelled by empirical data but is still not exempt from weaknesses. We concluded that HIV-1 isolated from Jurkat T cells integrates with more intensity in Giemsa light bands and with equal intensity in FRs and NFR. Our work is based on statistical analysis which complements laboratorial studies. Other

factors such as the cell cycle phase and the cellular type that act *in vivo* could also influence the integration site selection of the virus. Moreover, the distribution of integration sites may be altered *in vitro* due to repeated cell division and selection for certain clones. Nevertheless, our data are in line with previous reports and may contribute to the understanding of viral integration in HIV disease and gene therapy strategies. The presented *in silico* approach offers promise of useful application.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments

The authors are deeply grateful to Professor Bushman and his team who provided the raw data on integration sites of HIV in Jurkat T cells. This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through Project UID/BIM/00009/2013 (Centre for Toxicogenomics and Human Health (ToxOmics)). The authors would also like to thank the Center for Mathematics and Applications (CMA), FCT/UNL.

References

- [1] M. Costantini, F. Auletta, and G. Bernardi, "The distributions of a "new" and "old" Alu sequences in the human genome: the solution of a "mystery," *Molecular Biology and Evolution*, vol. 29, pp. 421–427, 2012.
- [2] Y. Niimura and T. Gojobori, "In silico chromosome staining: reconstruction of Giemsa bands from the whole human genome sequence," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 2, pp. 797–802, 2002.
- [3] T. Cremer and C. Cremer, "Chromosome territories, nuclear architecture and gene regulation in mammalian cells," *Nature Reviews Genetics*, vol. 2, no. 4, pp. 292–301, 2001.
- [4] K. Küpper, A. Kölbl, D. Biener et al., "Radial chromatin positioning is shaped by local gene density, not by gene expression," *Chromosoma*, vol. 116, no. 3, pp. 285–306, 2007.
- [5] I. Büttel, A. Fechter, and M. Schwab, "Common fragile sites and cancer: targeted cloning by insertional mutagenesis," *Annals of the New York Academy of Sciences*, vol. 1028, pp. 14–27, 2004.
- [6] K. Mrasek, C. Schoder, A.-C. Teichmann et al., "Global screening and extended nomenclature for 230 aphidicolin-inducible fragile sites, including 61 yet unreported ones," *International Journal of Oncology*, vol. 36, no. 4, pp. 929–940, 2010.
- [7] S. G. Durkin and T. W. Glover, "Chromosome fragile sites," *Annual Review of Genetics*, vol. 41, pp. 169–192, 2007.
- [8] T. Lukusa and J. P. Fryns, "Human chromosome fragility," *Biochimica et Biophysica Acta (BBA)—Gene Regulatory Mechanisms*, vol. 1779, no. 1, pp. 3–16, 2008.
- [9] J. Mitsui and S. Tsuji, "Common chromosomal fragile sites: breakages and rearrangements in somatic and germline cells," *Atlas of Genetics and Cytogenetics in Oncology and Haematology*, vol. 15, pp. 1089–1096, 2011.

- [10] C. H. Freudenreich, "Chromosome fragility: molecular mechanisms and cellular consequences," *Frontiers in Bioscience*, vol. 12, no. 13, pp. 4911–4924, 2007.
- [11] M. Schwartz, E. Zlotorynski, and B. Kerem, "The molecular basis of common and rare fragile sites," *Cancer Letters*, vol. 232, no. 1, pp. 13–26, 2006.
- [12] Y.-H. Wang, "Chromatin structure of human chromosomal fragile sites," *Cancer Letters*, vol. 232, no. 1, pp. 70–78, 2006.
- [13] T. W. Glover and C. K. Stein, "Chromosome breakage and recombination at fragile sites," *American Journal of Human Genetics*, vol. 43, no. 3, pp. 265–273, 1988.
- [14] A. Coquelle, E. Pipiras, F. Toledo, G. Buttin, and M. Debatisse, "Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplifications," *Cell*, vol. 89, no. 2, pp. 215–225, 1997.
- [15] E. Zlotorynski, A. Rahat, J. Skaug et al., "Molecular basis for expression of common and rare fragile sites," *Molecular and Cellular Biology*, vol. 23, no. 20, pp. 7143–7151, 2003.
- [16] Y. Watanabe, T. Ikemura, and H. Sugimura, "Amplicons on human chromosome 11q are located in the early/late-switch regions of replication timing," *Genomics*, vol. 84, no. 5, pp. 796–805, 2004.
- [17] W.-J. Luo, T. Takakuwa, M. F. Ham et al., "Epstein-Barr virus is integrated between REL and BCL-11A in American Burkitt lymphoma cell line (NAB-2)," *Laboratory Investigation*, vol. 84, no. 9, pp. 1193–1199, 2004.
- [18] M. Matovina, I. Sabol, G. Grubišić, N. M. Gašperov, and M. Grce, "Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions," *Gynecologic Oncology*, vol. 113, no. 1, pp. 120–127, 2009.
- [19] I. K. Christiansen, G. K. Sandve, M. Schmitz, M. Dürst, and E. Hovig, "Transcriptionally active regions are the preferred targets for chromosomal HPV integration in cervical carcinogenesis," *PLoS ONE*, vol. 10, no. 3, Article ID e0119566, pp. 1–11, 2015.
- [20] F. Bushman, M. Lewinski, A. Ciuffi et al., "Genome-wide analysis of retroviral DNA integration," *Nature Reviews Microbiology*, vol. 3, no. 11, pp. 848–858, 2005.
- [21] F. Turlure, E. Devroe, P. A. Silver, and A. Engelman, "Human cell proteins and human immunodeficiency virus DNA integration," *Frontiers in Bioscience*, vol. 9, pp. 3187–3208, 2004.
- [22] A. Ciuffi and F. D. Bushman, "Retroviral DNA integration: HIV and the role of LEDGF/p75," *Trends in Genetics*, vol. 22, no. 7, pp. 388–395, 2006.
- [23] A. Herschhorn and A. Hizi, "Retroviral reverse transcriptases," *Cellular and Molecular Life Sciences*, vol. 67, no. 16, pp. 2717–2747, 2010.
- [24] R. Craigie and F. D. Bushman, "HIV DNA integration," *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 7, Article ID a006890, pp. 1–18, 2012.
- [25] R. Daniel and J. A. Smith, "Integration site selection by retroviral vectors: molecular mechanism and clinical consequences," *Human Gene Therapy*, vol. 19, no. 6, pp. 557–568, 2008.
- [26] A. M. Skalka and R. A. Katz, "Retroviral DNA integration and the DNA damage response," *Cell Death and Differentiation*, vol. 12, no. 1, pp. 971–978, 2005.
- [27] F. Barré-Sinoussi, J. C. Chermann, F. Rey et al., "Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)," *Science*, vol. 220, no. 4599, pp. 868–871, 1983.
- [28] F. Clavel, D. Guétard, F. Brun-Vézinet et al., "Isolation of a new human retrovirus from West African patients with AIDS," *Science*, vol. 233, no. 4761, pp. 343–346, 1986.
- [29] A. R. W. Schröder, P. Shinn, H. Chen, C. Berry, J. R. Ecker, and F. Bushman, "HIV-1 integration in the human genome favors active genes and local hotspots," *Cell*, vol. 110, no. 4, pp. 521–529, 2002.
- [30] Z. Debyser, F. Christ, J. De Rijck, and R. Gijssbers, "Host factors for retroviral integration site selection," *Trends in Biochemical Sciences*, vol. 40, no. 2, pp. 108–116, 2015.
- [31] R. W. Wong, J. I. Mamede, and T. J. Hope, "Impact of nucleoporin-mediated chromatin localization and nuclear architecture on HIV integration site selection," *Journal of Virology*, vol. 89, no. 19, pp. 9702–9705, 2015.
- [32] M. A. Kay, J. C. Glorioso, and L. Naldini, "Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics," *Nature Medicine*, vol. 7, no. 1, pp. 33–40, 2001.
- [33] J. Nagel, B. Grob, M. Meggendorfer et al., "Stably integrated and expressed retroviral sequences can influence nuclear location and chromatin condensation of the integration locus," *Chromosoma*, vol. 121, no. 4, pp. 353–367, 2012.
- [34] G. P. Wang, A. Ciuffi, J. Leipzig, C. C. Berry, and F. D. Bushman, "HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications," *Genome Research*, vol. 17, no. 8, pp. 1186–1194, 2007.
- [35] I. J. Sequeira, J. T. Mexia, J. Santiago et al., "Predominance of constitutional chromosomal rearrangements in human chromosomal fragile sites," *Open Journal of Genetics*, vol. 3, no. 2, pp. 8–13, 2013.
- [36] A. Laganà, F. Russo, C. Sismeiro, R. Giugno, A. Pulvirenti, and A. Ferro, "Variability in the incidence of miRNAs and genes in fragile sites and the role of repeats and CpG islands in the distribution of genetic material," *PLoS ONE*, vol. 5, no. 6, Article ID e11166, pp. 1–8, 2010.
- [37] J. J. A. Holden, P. Ridgway, and A. Smith, "A possible fragile-site at Yq12: case report and possible relevance to de novo structural rearrangements of the Y-chromosome," *American Journal of Medical Genetics*, vol. 23, no. 1–2, pp. 545–555, 1986.
- [38] S. Siegel, "O caso de duas amostras relacionadas," in *Estatística Não-Paramétrica para as Ciências do Comportamento*, pp. 75–93, McGraw-Hill, São Paulo, Brazil, 1975.
- [39] C. Berry, S. Hannenhalli, J. Leipzig, and F. D. Bushman, "Selection of target sites for mobile DNA integration in the human genome," *PLoS Computational Biology*, vol. 2, no. 11, pp. 1450–1462, 2006.
- [40] R. S. Mitchell, B. F. Beitzel, A. R. W. Schroder et al., "Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences," *PLoS Biology*, vol. 2, no. 8, article e234, 2004.
- [41] A. Jordan, P. Defechereux, and E. Verdin, "The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation," *The EMBO Journal*, vol. 20, no. 7, pp. 1726–1738, 2001.
- [42] M. K. Lewinski, D. Bisgrove, P. Shinn et al., "Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription," *Journal of Virology*, vol. 79, no. 11, pp. 6610–6619, 2005.
- [43] A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho, "HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time," *Science*, vol. 271, no. 5255, pp. 1582–1586, 1996.
- [44] D. Elleder, A. Pavlíček, J. Pačes, and J. Hejnar, "Preferential integration of human immunodeficiency virus type 1 into genes, cytogenetic R bands and GC-rich DNA regions: Insight from the human genome sequence," *FEBS Letters*, vol. 517, no. 1–3, pp. 285–286, 2002.

- [45] G. V. Kalpana, S. Marmon, W. Wang, G. R. Crabtree, and S. P. Goff, "Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5," *Science*, vol. 266, no. 5193, pp. 2002–2006, 1994.
- [46] G. Maertens, P. Cherepanov, W. Pluymers et al., "LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells," *The Journal of Biological Chemistry*, vol. 278, no. 35, pp. 33528–33539, 2003.
- [47] B. Marini, A. Kertesz-Farkas, H. Ali et al., "Nuclear architecture dictates HIV-1 integration site selection," *Nature*, vol. 521, no. 7551, pp. 227–231, 2015.
- [48] G. P. Holmquist and T. Ashley, "Chromosome organization and chromatin modification: influence on genome function and evolution," *Cytogenetic and Genome Research*, vol. 114, no. 2, pp. 96–125, 2006.
- [49] N. Sadoni, S. Langer, C. Fauth et al., "Nuclear organization of mammalian genomes: polar chromosome territories build up functionally distinct higher order compartments," *The Journal of Cell Biology*, vol. 146, no. 6, pp. 1211–1226, 1999.
- [50] D. Pyeon, S. M. Pearce, S. M. Lank, P. Ahlquist, and P. F. Lambert, "Establishment of human papillomavirus infection requires cell cycle progression," *PLoS Pathogens*, vol. 5, no. 2, Article ID e1000318, 9 pages, 2009.
- [51] I. Büttel, A. Fechter, and M. Schwab, "Common fragile sites and cancer: targeted cloning by insertional mutagenesis," *Annual New York Academy of Sciences*, vol. 1028, pp. 14–27, 2004.
- [52] O. Mortusewicz, P. Herr, and T. Helleday, "Early replication fragile sites: where replication-transcription collisions cause genetic instability," *The EMBO Journal*, vol. 32, no. 4, pp. 493–495, 2013.
- [53] M. J. Soto, Á. Peña, and F. G. Vallejo, "A genomic and bioinformatics analysis of the integration of HIV in peripheral blood mononuclear cells," *AIDS Research and Human Retroviruses*, vol. 27, no. 5, pp. 547–555, 2011.
- [54] M. C. Zody, M. Garber, D. J. Adams et al., "DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage," *Nature*, vol. 440, no. 7087, pp. 1045–1049, 2006.