Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

Interaction of vectors and parental viruses with the host genome

Irene Gil-Farina and Manfred Schmidt

Viral replication by acquisition of the host cell biology represents a central part of a virus life cycle. Thereby, integration into the host genome constitutes a successful strategy to ensure viral persistence and viruses have developed different mechanisms to integrate and benefit from cell's transcriptional and translational machinery. While lentiviral (e.g. HIV) integration is influenced by the chromatin landscape encountered upon nuclear entry, certain parvoviruses (e.g. AAV) integrate specifically within genomic regions bearing increasingly known sequence motifs. Gene therapy exploits these viral persistence strategies to achieve efficient and safe long-term transgene expression. Here we focus on two widely used vectors and their parental viruses, HIV and AAV, to discuss recent insights into lentiviral vector oncogenicity by alteration of endogenous transcripts as well as the unresolved AAV vectors genotoxic potential.

Address
Department of Translational Oncology, National Center for Tumor Diseases and German Cancer Research Center, 69120 Heidelberg, Germany

Corresponding author: Schmidt, Manfred (manfred.schmidt@nct-heidelberg.de)

Current Opinion in Virology 2016, **21**:35–40

This review comes from a themed issue on **Viral gene therapy vector–host interactions**

Edited by **David Markusic** and **Roland Herzog**

http://dx.doi.org/10.1016/j.coviro.2016.07.004

Introduction

The viral life cycle is organized in attachment, penetration, uncoating, replication, gene expression, assembly and budding. Depending on the virus type, individual steps may start simultaneously or overlap (e.g. replication and gene expression). For most viruses, replication and gene expression take place in the nucleus, where viral genomes persist episomally or integrate into the host genome constituting the so-called proviruses. The latter is a hallmark of lysogenic viruses that persist in the infected cells (latency) without inducing cell lysis. In contrast, viruses mediating cell lysis commonly do not penetrate the nucleus but remain in the cytoplasm and directly produce massive progeny.

Gene therapy makes use of episomal or integrating viral vectors that are chosen according to the target cell type. Episomally persisting vectors leave the host genome physically inert and may provide sustained transgene expression, but are diluted out when the cells divide. Therefore, they may be used on purpose if only transient transgene activity is intended or if post-mitotic tissues will be targeted. Integrated vectors represent an integral part of the host genome and provide, in principle, lifelong transgene expression also in dividing cells. However, integrations may lead to insertional mutagenesis affecting the infected cell by alteration of normal gene expression and regulatory mechanisms that could result in apoptosis or immortalization and malignant transformation, particularly when involving genes controlling cell proliferation or survival. Therefore, the understanding of how viruses select their sites of integration in the host cell genome is essential in order to evaluate their genotoxic potential and, more importantly for the gene therapy field, to develop strategies in order to suppress or modulate specific integration preferences.

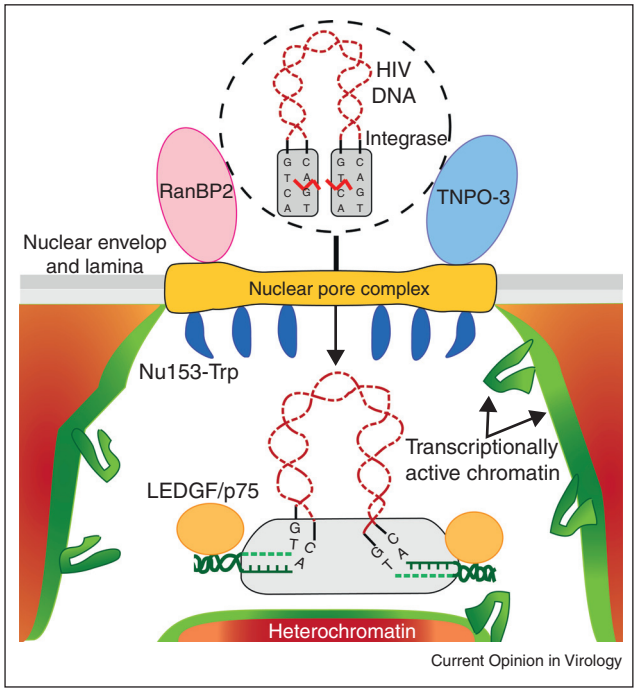
This article reviews studies performed over the last five years providing novel information about how viruses, as well as their derived gene therapy vectors, select integration site location within the cellular genome and the safety concerns associated. We focused on the human immunodeficiency virus (HIV) and the wild-type adeno-associated viruses (wtAAV) as examples bearing unspecific and site-specific integrases, respectively. Similarly, their derived gene therapy vectors, only conserving the parental terminal repeats, are discussed in terms of their integration profiles and genotoxic potential under the light of recent studies.

Wild-type HIV: nuclear architecture as major determinant of integration

The single-stranded positive RNA HIV belongs to the genus *Lentivirus*. As other retroviruses, its genome encodes a reverse transcriptase to synthesize the DNA intermediate and an integrase that catalyzes viral integration [1]. HIV replication and persistence completely relies on its ability to integrate into the host genome [2]. The viral integrase is responsible for this process and, although preferential targeting of gene dense and transcriptional active regions is well known [3], how the integration site (IS) location is selected remains unclear. Nuclear architecture presents a

functional organization where active transcription units are located within the nuclear periphery nearby the nuclear pore complex (NPC) while heterochromatin is mainly found at inner regions or associated to the nuclear lamina [4•]. Taking into account the nuclear architecture and HIV nuclear entry by active translocation through the NPC [5••], the incoming HIV encounters chromatin regions enriched on actively transcribed genes suitable for integration (Figure 1). Recently, HIV recurrently targeted genes and provirus have been topologically localized in the nuclear periphery in association with the NPC and absent at internal nuclear regions and lamin-associated domains [6]. Trp knock-down, that modifies nuclear architecture by allocating silent chromatin to regions nearby the NPC, resulted in a decreased HIV expression without altering its nuclear location [7]. Moreover, alterations in the nuclear import machinery, such as the knock-down of the NPC-associated proteins RanBP2 and Transportin-3, resulted in HIV integration within low gene density regions [8–10]. Therefore, this evidences that chromatin spatial proximity to the NPC strongly influences HIV IS location. Nonetheless, multiple viral–host protein interactions, such as LEDGF/p75 that guides the pre-integration complex through active transcription units, are also involved providing a further fine-tuning [11–15].

Figure 1



Nuclear pore trafficking and nuclear architecture influence HIV integration. The pre-integration complex (PIC), mainly composed by the viral genome and the integrase, interacts with the nuclear pore complex (NPC) and associated proteins (RanBP2, TNPO-3) resulting in an active translocation. Once in the nucleus, the PIC encounters regions of transcriptionally active chromatin in the NPC surrounding area. Here, cellular cofactors (Nu153-Trp, LEDGF/p75) further guide the PIC toward actively transcribed genes and the integrase catalyzes HIV integration by cleaving the viral long-terminal repeats and the cellular genome for later binding them together in a trans-esterification reaction. Finally, cellular repair machinery eliminates viral overhangs and repairs the gaps left in the host genome.

In the last years, next-generation sequencing has enabled the detection of eventual clonal expansions induced by HIV integration events. Maldarelli and colleagues analyzed 2.410 HIV IS from patients peripheral blood mononuclear cells mapping to 985 different genes of which 67% corresponded to single integrations and, remarkably, 33% corresponded to single or multiple integrations that underwent clonal expansion [16]. Complementary, Wagner and co-workers showed that HIV integration into cancer- or cell cycle-related genes induces cell proliferation contributing to viral persistence [17]. Although it might not be the sole mechanism, these evidences support that integration-induced clonal expansion of infected cells plays a role in HIV persistence.

HIV-derived lentiviral vectors: retroviral gene therapy lacking visible genotoxicity

Lentiviral vectors (LV) are replication incompetent HIV-based gene therapy vectors and the last generations consist of an expression cassette flanked by the viral long-terminal repeats (LTR). Second generation LV contain wild-type LTRs and the psi packaging signal, whereas the third generation presents self-inactivating (SIN) LTRs, that is, they bear deletions that reduce their transcriptional activity. LV main advantages over other retroviral vectors, such as gammaretroviruses, are their ability to infect both dividing and non-dividing cells and their potentially safer integration profile [18]. Similarly to HIV, LV have been shown to preferentially integrate within transcription units involved in chromatin modification/remodeling, functions related to the major histocompatibility complex class II, steroid hormone receptors and RNA processing [19,20]. Notably, a recent study demonstrated the direct functional correlation between active transcription and LV integration by microarray and integration analysis performed upon thymine treatment in $CD34^+$ cells [21]. Interestingly, preclinical data have also described LV targeting of satellite DNA upon intraventricular administration into the mouse brain [22].

LV have been successfully used in different clinical trials with positive outcomes and, despite a relative clonal expansion found upon integration in the *HMGA2* gene in one patient treated for beta-thalassemia correction [23], no adverse genotoxic events have been reported so far [24–26]. However, some concerns regarding LV safety still exist as different studies suggested that they may alter the expression of genes flanking the IS or even affect cellular transcripts generating loss-of-function or gain-of-function variants [27,28]. A recent study has addressed these concerns by testing different LV constructs in a