Title: State of genomics and epigenomics research in the perspective of HIV cure.
Authors: Ciuffi A, Telenti A
Journal: Current opinion in HIV and AIDS
Year: 2013 May
Volume: 8
Issue: 3
Pages: 176-81
DOI: 10.1097/COH.0b013e32835f7340

State of genomics and epigenomics research in the perspective of HIV cure

Angela Ciuffi* and Amalio Telenti*

Institute of Microbiology, University Hospital of Lausanne, University of Lausanne, Switzerland

*Correspondence: AC (angela.ciuffi@chuv.ch), and AT (Amalio.telenti@chuv.ch), Institute of Microbiology, CHUV, 1011 Lausanne, Switzerland. Phone +41795560751.

ABSTRACT

Purpose of review. One of the seven key scientific priorities indentified in the road map on HIV cure research is to "determine host mechanisms that control HIV replication in the absence of therapy". This review summarizes recent work in genomics and in epigenetic control of viral replication that is relevant for this mission.

Recent findings. New technologies allow the joint analysis of host and viral transcripts. They identify the patterns of antisense transcription of the viral genome, and its possible role in gene regulation. High throughput studies facilitate the assessment of features of integration at the genome level. Integration site, orientation and host genomic context modulate transcription and should also be assessed at the level of single cells. The various models of latency in primary cells can be followed using dynamic study designs to acquire transcriptome and proteome data of the process of entry, maintenance and reactivation of latency. Dynamic studies can also be applied to the study of transcription factors and chromatin modifications in latency and upon reactivation.

Summary. The convergence of primary cell models of latency, new high throughput quantitative technologies applied to the study of time series, and the identification of compounds that reactivate viral transcription brings unprecedented precision to the study of viral latency.

Keywords: deep sequencing, antisense transcription, integration, epigenetic regulation.

Introduction

The global scientific strategy of the International HIV Society "Towards an HIV cure" [1] identified the potential of genomics for the understanding of mechanisms of latency. This includes the analysis of the steps in the establishment of viral transcriptional silencing as well as the steps that allow reactivation of viral transcription: the site of viral DNA integration, the chromatin environment of the integrated provirus and the availability of transcription factors. These components can be investigated through genomic, transcriptomic and epigenomic analyses.

Currently, two models are used to describe how HIV establishes latency in resting CD4+ T cells [2, 3]. The first model suggests that activated cells are first infected by HIV, and even though the majority of them are productively infected and die within a few days, a minority of them revert to a resting memory state, following the natural biology of CD4+ T cells. Although the mechanisms leading to cell memory are yet unclear, it is accompanied by cellular changes at the transcriptome and proteome levels (that are yet to be completely described), that in turn impact the viral transcription, ultimately repressing it. The second model suggests that, even if poorly efficient, HIV is able to directly infect resting CD4+ T cells. Initial studies on HIV latency used different models based on cell lines, as primary cell cultures were difficult and short-lived. However, in the recent years, multiple models using primary cells - based on different CD4+ T cell populations - have been developed that yield sufficient amounts of cells, thereby allowing investigation that recapitulate events that may occur *in vivo* [4].

Latency models using T cell lines and primary CD4+ T cells are also used for the screening and assessment of molecules that promote viral transcription [5-8]. That multiple compounds reactivate viral transcription from latently infected cells *in vitro* may reflect the various mechanisms involved in viral transcriptional control [6, 9]. These include epigenetic regulation (such as histone acetylation and DNA methylation) and immune modulation (such as T cell receptor engagement and protein kinase C signaling) [6, 9]. Some of the agents are moving forward to clinical assessment [1, 9, 10].

This review will present new aspects and opportunities in the use of novel technology for the analysis of contribution of host factors to the viral life cycle in productive infection, and of viral integration site and host factors (including epigenetic control of the viral promoter) in the establishment of and reactivation from latency (**Figure 1**).

Dynamic aspects of the viral life cycle

HIV-1 is fully dependent of the cellular machinery to complete the replication cycle. The infected cell undergoes a profound modification of its physiology. Although the interaction between HIV-1 and the host cell has been extensively studied [11], previous analyses did not assess all relevant steps of viral replication in a dynamic perspective. Transcriptome analyses used microarray technology in cross-sectional experiments, generally at the completion of the viral replication cycle (24-48 hours) [12]. The transcriptional status of genes contributes to preferential integration of proviruses [13]; however, there is limited data on how the viral integration contributes to host transcription at genome-wide level. Analyses have also been hampered by the

heterogeneity of the infectious system, where the transcriptome profile reflects contribution by the infected and uninfected cells. Recent studies have approached this problem by magnetic sorting of cells infected in vitro identified by a marker recombinant protein that is expressed during the late-phase of viral replication cycle [14].

Progress in deep-sequence technologies now allow the assessment of viral and host transcripts jointly. Lefebvre et al. [15] infected SupT1 cells to analyze cellular and viral transcriptomes by serial analysis of gene expression followed by highthroughput sequencing. At 24 hours post-infection, read mapping resulted in 33 to 44 million tags aligning with the human transcriptome and 0.23 to 0.25 million tags aligning with the genome of the HIV-1 vector. Thus, at peak infection, 1 transcript in 143 is of viral origin (0.7%). Analysis also identified a small component of antisense viral transcription. Antisense transcription from the 3'LTR has been described in infected primary cells, including monocyte-derived cells and activated T lymphocytes [16].

Kobayashi-Ishihara et al. [17] identified a 2.6 kb RNA as the major form of HIV-1 antisense RNAs. This transcript corresponds to the previously reported antisense protein (ASP) mRNA. The expression of this antisense RNA suppressed HIV-1 replication and its knockdown enhanced HIV-1 gene expression and replication. Schopman et al. [18] detected numerous small RNAs encoded by the HIV-1 genome. Most sequences have positive polarity and could correspond to viral miRNAs. A small portion of the viral small RNAs have negative polarity and they could represent viral siRNAs. Importantly, the identified viral siRNAs potently repress HIV-1 production. The authors proposed that these short virally encoded miRNAs and siRNAs modulate cellular and/or viral gene expression.

More recently, the complete dynamic profile of viral replication intermediates, host small RNAs and mRNAs has been captured and modeled [19]. A significant proportion of host genes are modulated in concordance with key steps of viral cycle. Analyses underscored the features of the successful viral replication occurring despite a profound perturbation of the cell at the transcriptional level. This type of work – currently performed in cell lines - represents a referential resource that can be contrasted across cellular systems and viral strains, and that can be extended to study the establishment of and reactivation from viral latency.

Spatial features of HIV integration

HIV inserts its viral genome non-randomly into the host DNA, strongly favoring active transcription units, thereby potentially promoting efficient viral transcription and productive infection [13, 20]. However, the location of the viral genome may also hamper viral transcription and thereby promote latency. To date, the site of viral integration has been shown to contribute to viral repression mostly by the following mechanisms: (i) the chromatin environment may disfavor viral transcription: viral genomes integrated into centromeric alphoid repeats or in gene deserts (although representing a minority of integration sites compared to active transcription units) are more prone to heterochromatin formation and thus more likely to be associated with viral transcriptional repression, as shown in a cell line model of latency and in a model of direct infection of resting primary CD4+ T cells [21-24]; and (ii) viral genome integration in active transcription units imply that

transcription of cellular genes flanking the proviral genome may lead to transcriptional interference through promoter occlusion or collision, thereby potentially resulting in viral transcription repression: highly expressed genes hosting integrated viral genomes lead to reduction of viral transcription in a cell line model of latency and in resting primary CD4+ T cells [23, 25]. Analysis of viral transcription performed upon targeted insertion of the viral genome in one actively transcribed gene (*HPRT*) suggested that viral transcription was inhibited when viral and cellular transcription were in convergent orientations whereas it was enhanced when viral and cellular transcription were in the same orientation [26]. In contrast, in a primary CD4+ T cell model of latency, viral integration in the same orientation as cellular hosting genes was enriched in latent cells, suggesting that repression of viral transcription was preferentially promoted upon promoter occlusion in this system [27]. These data argue for a contribution of integration site location and surrounding host genomic features to viral transcription efficiency; however to which extent each of these factors influences latency remains to be clarified. Diversity in integration site may also modulate the efficacy of reactivating agents used in *in vitro* models of latency [6, 9, 28].

In addition to the multiple ways by which HIV integration site location and orientation can contribute to latency, various CD4+ T populations have a different capacity to support the reservoir of HIV infection. These may result in significant loss of homogeneity in the population of latently infected cells, emphasizing the interest of examining cell-to-cell heterogeneity in single cell assays. There are now different techniques that make use of microfluidic platforms and highly multiplexed assays that allow the analysis of expression of 20 – 800 genes at the single cell level [29, 30]. Progress in single cell analysis could allow a more precise evaluation of the influence of viral genome integration site on viral transcription.

Transcriptome and proteome modifications during latency and reactivation

Some of the recent applications of transcriptome analysis aimed at the characterization of the pattern of host expression after control of viral replication *in vivo* [12, 31, 32]. Similar approach was applied to the characterization of host expression profile in elite controllers. The transcriptome profile in CD4+ T cells of successfully treated individuals becomes similar to that of elite controllers. Detailed analysis of specific cell subsets may however demonstrated residual abnormalities in gene regulation among individuals receiving antiretroviral treatment compared with healthy blood donors [33]. These analyses however capture population effects that reflect overall consequences of immune activation and the contribution of transcriptomes of non-infected cells.

Comparison of resting and activated CD4+ T cell transcriptomes identified differentially expressed and alternatively spliced genes [34-36]. Of note, these studies used resting cells, and then stimulated them to activated cells; however, no study describes the contrary, i.e. activated to resting. Assessing transcriptome changes from activated infected cells to resting cells is needed to better characterize the sequence of events leading to viral transcriptional silencing.

Resting and activated CD4+ T cells also differ in their proteome composition. Although no large-scale proteome study was performed so far to compare CD4+ T cells in these two states, the differential expression or localization of specific proteins has been reported. More particularly, the pool of available transcription factors

in the nucleus regulating viral transcription varies between resting and activated CD4+ T cells [4, 37]. Examples include (i) the nuclear availability of NF-κB in activated cells while sequestered in the cytoplasm of quiescent CD4+ T cells, (ii) the high amount of CBF-1, an inhibitor of LTR-driven viral transcription, in resting cells while reduced amounts were detected in activated cells [37, 38], (iii) the degradation of cyclinT1 and the absence of T186 phosphorylation on cdk9, the units composing pTEFb kinase, in resting cells while abundantly present in activated cells [39].

Epigenetic modifications of the HIV promoter

As the viral genome is integrated into the host DNA, it is chromatinized and subject of cellular epigenetic regulation. Chromatin organization at the viral LTR promoter has been investigated, identifying precisely positioned nucleosomes [40, 41]: nuc-0 at position -415 to -255 (compared to Transcriptional Start Site, TSS, +1), a nucleosome-free region or with poorly positioned nucleosomes (-255 to -3/+10), nuc-1 at +10 to +155, a nucleosome-free region (+155 to +265), and nuc-2 at +265- to +409. Nucleosome-free regions present binding sites for multiple transcriptional regulators, while nuc-1 is positioned just after the TSS, blocking viral transcription, and thus needed to be displaced or disrupted to allow successful transcriptional elongation. Thus, the identity of bound transcriptional regulators as well as histone post-translational modifications and DNA methylation will dictate viral gene expression [4].

Binding of available specific transcription factors recruit chromatin modifying enzymes that will shape the chromatin environment at the viral LTR and participate in regulation of viral transcription activity. Studies using chromatin immunoprecipitation (ChIP) showed that histone post-translational modifications, specifically acetylation and methylation, are correlated with viral transcriptional activity [4, 41-43]. DNA methylation in the CpG islands surrounding the viral transcription start site is associated with viral gene silencing [3, 44, 45]. DNA methylation is currently described as being the last step of repression of gene expression, contributing to a compact chromatin structure and decreasing the efficiency of some reactivating agents. These data confirmed that specific transcription factors, histone modifications and DNA methylation levels at the viral promoter determine the efficiency of viral gene expression. However, most of these studies compared activated vs resting cells at one single time point focusing on specific histone post-translational modifications, without sequential evaluation of chromatin modifications.

Conclusions

Genomic analyses are well suited for the study of viral latency and reactivation as a dynamic process. As such, these approaches respond to several of the priorities of the HIV Cure Research agenda published in 2012 [1] (**Box 1**); in particular the call for detailed understanding of the transcriptional features of the process of latency, and the possibility to use novel technologies. Specific issues where novel technology may be of importance include assessment of (i) the extent of “true” latency, *i.e.* the degree of HIV transcriptional silence, (ii) the role of spatial features of viral integration contributing to transcription, (iii) the identification of candidate biomarkers of latency, and (iv) the characterization of successful activation programs leading to viral expression at population and single cell level. Keypoints are presented in **Box 2**.

Acknowledgements and Conflict of interest

The authors are supported by the Swiss National Science Foundation (grants 31003A_132863/1 and 310030-130699; www.snf.ch) and by a Grand Challenges Exploration award from the Bill and Melinda Gates Foundation. The authors have no conflict of interest related to this work.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

*of special interest

**of outstanding interest

Reference List

[1] International ASSWGoHIVC, Deeks SG, Autran B, Berkhout B, Benkirane M, Cairns S, et al. Towards an HIV cure: a global scientific strategy. Nat Rev Immunol. 2012 Aug;12(8):607-14.

[2] Pace MJ, Agosto L, Graf EH, O'Doherty U. HIV reservoirs and latency models. Virology. 2011 Mar 15;411(2):344-54.

[3] Siliciano RF, Greene WC. HIV Latency. Cold Spring Harbor perspectives in medicine. 2011 Sep;1(1):a007096.

***[4] Hakre S, Chavez L, Shirakawa K, Verdin E. HIV latency: experimental systems and molecular models. FEMS microbiology reviews. 2012 May;36(3):706-16.**

An excellent review of the molecular mechanisms leading to viral repression.

[5] Yang HC, Xing S, Shan L, O'Connell K, Dinoso J, Shen A, et al. Small-molecule screening using a human primary cell model of HIV latency identifies compounds that reverse latency without cellular activation. J Clin Invest. 2009 Nov;119(11):3473-86.

[6] Planelles V, Wolschendorf F, Kutsch O. Facts and fiction: cellular models for high throughput screening for HIV-1 reactivating drugs. Curr HIV Res. 2011 Dec 1;9(8):568-78.

[7] Gallastegui E, Marshall B, Vidal D, Sanchez-Duffhues G, Collado JA, Alvarez-Fernandez C, et al. Combination of biological screening in a cellular model of viral latency and virtual screening identifies novel compounds that reactivate HIV-1. J Virol. 2012 Apr;86(7):3795-808.

[8] Xing S, Bhat S, Shroff NS, Zhang H, Lopez JA, Margolick JB, et al. Novel structurally related compounds reactivate latent HIV-1 in a bcl-2-transduced primary CD4+ T cell model without inducing global T cell activation. J Antimicrob Chemother. 2012 Feb;67(2):398-403.

[9] Smith MZ, Wightman F, Lewin SR. HIV reservoirs and strategies for eradication. Curr HIV/AIDS Rep. 2012 Mar;9(1):5-15.

[10] Archin NM, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Crooks AM, et al. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. Nature. 2012 Jul 26;487(7408):482-5.

[11] Bushman FD, Nabel GJ, Swanstrom R. HIV - from biology to prevention and treatment. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press 2012.

[12] Mehla R, Ayyavoo V. Gene array studies in HIV-1 infection. Curr HIV/AIDS Rep. 2012 Mar;9(1):34-43.

[13] Craigie R, Bushman FD. HIV DNA Integration. Cold Spring Harbor perspectives in medicine. 2012 Jul;2(7):a006890.

***[14] Imbeault M, Giguere K, Ouellet M, Tremblay MJ. Exon Level Transcriptomic Profiling of HIV-1-Infected CD4(+) T Cells Reveals Virus-Induced Genes and Host Environment Favorable for Viral Replication. PLoS Pathog. 2012 Aug;8(8):e1002861.**

A first approach to separating infected from noninfected cells for the purpose of transcriptome analysis.

***[15] Lefebvre G, Desfarges S, Uyttebroeck F, Muñoz M, Beerenwinkel N, Rougemont J, et al. Analysis of HIV-1 expression level and sense of transcription by high-throughput sequencing of the infected cell. J Virol. 2011;85:6205-11.**

First use of deep sequencing for the joint analysis of viral and cellular transcriptome. The study highlights the proportion of viral antisense transcripts.

[16] Laverdure S, Gross A, Arpin-Andre C, Clerc I, Beaumelle B, Barbeau B, et al. HIV-1 Antisense Transcription Is Preferentially Activated in Primary Monocyte-Derived Cells. J Virol. 2012 Dec;86(24):13785-9.

[17] Kobayashi-Ishihara M, Yamagishi M, Hara T, Matsuda Y, Takahashi R, Miyake A, et al. HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. Retrovirology. 2012;9:38.

****[18] Schopman NC, Willemsen M, Liu YP, Bradley T, van Kampen A, Baas F, et al. Deep sequencing of virus-infected cells reveals HIV-encoded small RNAs. Nucleic Acids Res. 2012 Jan;40(1):414-27.**

Deep sequencing analysis of viral-encoded miRNAs and siRNAs. The study is completed with functional demonstration of the regulatory potential of these small RNAs.

***[19] Mohammadi P, Desfarges S, Bartha I, Joos B, Zangger N, Muñoz M, et al. 24 hours in the life of HIV-1 in a T cell line. PLoS Pathog.**

The first complete dynamic model of the viral life cycle. It considers the timing of generation of viral replication intermediates, and the concordance between cellular transcription and viral progression.

[20] Desfarges S, Ciuffi A. Retroviral integration site selection. Viruses. 2010 Jan;2(1):111-30.

***[21] Pace MJ, Graf EH, Agosto LM, Mexas AM, Male F, Brady T, et al. Directly Infected Resting CD4+T Cells Can Produce HIV Gag without Spreading Infection in a Model of HIV Latency. PLoS Pathog. 2012 Jul;8(7):e1002818.**

Analysis of integration and viral transcription in HIV-infected naïve CD4+ T cells. While they identify full length transcripts, other transcripts are represented at very low levels.

[22] Jordan A, Bisgrove D, Verdin E. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. The EMBO journal. 2003 Apr 15;22(8):1868-77.

[23] Lewinski MK, Bisgrove D, Shinn P, Chen H, Hoffmann C, Hannenhalli S, et al. Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. J Virol. 2005 Jun;79(11):6610-9.

[24] Brady T, Agosto LM, Malani N, Berry CC, O'Doherty U, Bushman F. HIV integration site distributions in resting and activated CD4+ T cells infected in culture. AIDS. 2009 Jul 31;23(12):1461-71.

[25] Han Y, Lassen K, Monie D, Sedaghat AR, Shimoji S, Liu X, et al. Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. J Virol. 2004 Jun;78(12):6122-33.

[26] Han Y, Lin YB, An W, Xu J, Yang HC, O'Connell K, et al. Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. Cell Host Microbe. 2008 Aug 14;4(2):134-46.

***[27] Shan L, Yang HC, Rabi SA, Bravo HC, Shroff NS, Irizarry RA, et al. Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. J Virol. 2011 Jun;85(11):5384-93.**

In a latency model using primary CD4 + T cells, the authors show enrichment for proviral genomes placed in antisense orientation. This exemplifies a possible mechanism of transcriptional interference.

[28] Miller-Jensen K, Dey SS, Pham N, Foley JE, Arkin AP, Schaffer DV. Chromatin accessibility at the HIV LTR promoter sets a threshold for NF-kappaB mediated viral gene expression. Integrative biology : quantitative biosciences from nano to macro. 2012 Jun;4(6):661-71.

[29] Citri A, Pang ZP, Sudhof TC, Wernig M, Malenka RC. Comprehensive qPCR profiling of gene expression in single neuronal cells. Nature protocols. 2012 Jan;7(1):118-27.

[30] Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, et al. Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell. 2011 Feb 4;144(3):439-52.
[31] Rotger M, Dang KK, Fellay J, Heinzen EL, Feng S, Descombes P, et al. Genome-wide mRNA expression correlates of viral control in CD4+ T-cells from HIV-1-infected individuals. PLoS Pathog. 2010;6(2):e1000781.
[32] Rotger M, Dalmau J, Rauch A, McLaren P, Bosinger SE, Marttinez R, et al. Comparative transcriptome analysis of extreme phenotypes of human HIV-1 infection and sooty mangabey and rhesus macaque models of SIV infection. J Clin Invest. 2011;121:2391-400.
[33] Vigneault F, Woods M, Buzon MJ, Li C, Pereyra F, Crosby SD, et al. Transcriptional Profiling of CD4 T Cells Identifies Distinct Subgroups of HIV-1 Elite Controllers. J Virol. 2011 Mar;85(6):3015-9.
[34] Grigoryev YA, Kurian SM, Nakorchevskiy AA, Burke JP, Campbell D, Head SR, et al. Genome-wide analysis of immune activation in human T and B cells reveals distinct classes of alternatively spliced genes. PLoS ONE. 2009;4(11):e7906.
[35] Butte MJ, Lee SJ, Jesneck J, Keir ME, Haining WN, Sharpe AH. CD28 costimulation regulates genome-wide effects on alternative splicing. PLoS ONE. 2012;7(6):e40032.
[36] Kalsotra A, Cooper TA. Functional consequences of developmentally regulated alternative splicing. Nat Rev Genet. 2011 Oct;12(10):715-29.
[37] Colin L, Van Lint C. Molecular control of HIV-1 postintegration latency: implications for the development of new therapeutic strategies. Retrovirology. 2009;6:111.
[38] Tyagi M, Karn J. CBF-1 promotes transcriptional silencing during the establishment of HIV-1 latency. The EMBO journal. 2007 Dec 12;26(24):4985-95.
[39] Budhiraja S, Famiglietti M, Bosque A, Planelles V, Rice AP. Cyclin T1 and CDK9 T-loop phosphorylation are downregulated during establishment of HIV-1 latency in primary resting memory CD4+ T cells. J Virol. 2012 Nov 14.
[40] Verdin E, Paras P, Jr., Van Lint C. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. The EMBO journal. 1993 Aug;12(8):3249-59.
[41] Rafati H, Parra M, Hakre S, Moshkin Y, Verdin E, Mahmoudi T. Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. PLoS Biol. 2011 Nov;9(11):e1001206.
A very important amount of work to show that BAF-containing SWI/SNF chromatin remodeling complexes were not able to displace Nuc-1, thus hampering viral transcription .
[42] Easley R, Carpio L, Dannenberg L, Choi S, Alani D, Van Duyne R, et al. Transcription through the HIV-1 nucleosomes: effects of the PBAF complex in Tat activated transcription. Virology. 2010 Sep 30;405(2):322-33.
[43] Tyagi M, Pearson RJ, Karn J. Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction. J Virol. 2010 Jul;84(13):6425-37.
[44] Blazkova J, Trejbalova K, Gondois-Rey F, Halfon P, Philibert P, Guiguen A, et al. CpG methylation controls reactivation of HIV from latency. PLoS Pathog. 2009 Aug;5(8):e1000554.
[45] Kauder SE, Bosque A, Lindqvist A, Planelles V, Verdin E. Epigenetic regulation of HIV-1 latency by cytosine methylation. PLoS Pathog. 2009 Jun;5(6):e1000495.

Keypoints.

- New technologies have been employed for the quantitative, dynamic analyses of viral-host interaction: RNA sequencing, proteome analysis, ChIP sequencing.
- The first models of productive and latent infection are generated using cell lines and primary CD4 + T cell.
- There is increasing recognition of the role of antisense transcription of the HIV genome in regulating viral gene expression.
- Integration site, orientation, and surrounding host genetic features influence viral transcription efficiency.
- Single cell analysis will be needed to assess the role of integration features on viral transcription, and other aspects of heterogeneity in the latent cell population.

Figure 1. Schematic overview of tools and technologies for the study of latency. Recent developments in cellular models of latency, and increasing performance of high-throughput studies pave the way to a detailed description of the sequential steps leading to repression of viral transcription.

Figure 2. Recommendations from the International AIDS Society Scientific Working Group on HIV Cure, Towards an HIV Cure on the use of novel technologies in latency research. (A Global Scientific Strategy, Full Recommendations Report", July 19, 2012, www.iasociety.org)

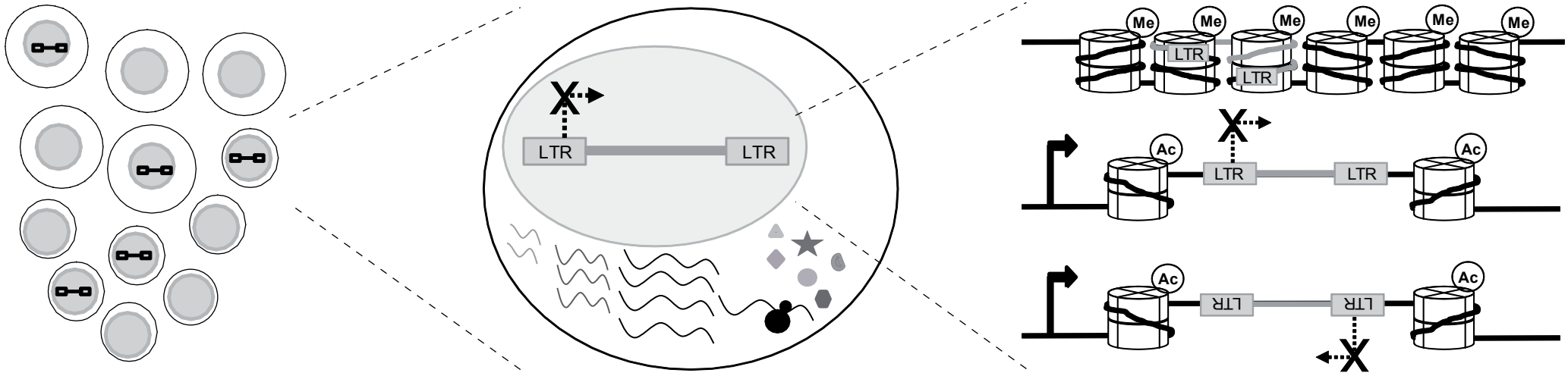
1. Parallel assessment of existing in vitro latency models and patient derived latently infected cells. Some potential approaches include:
    a. Deep-sequencing of latently infected cells – using latently infected cells from *in vitro* models and patient derived cells.
    b. Validation of the phenotype/expression profile of the latently infected cell. This includes testing of "expression modules" (short list of genes associated with the study profile), and development of single cell assays.
2. Characterization of viral transcripts (if any) associated with latency, or with entry/exit from latency – including antisense transcripts with a possible regulatory role.
3. Clinical and biological determinants of reservoir size. This could first be assessed using large well characterized cohort studies comparing reservoir size to a range of clinical parameters. The use of genomics, transcriptomics and other population based approaches should also be considered.
4. Adaptation of identified candidate markers for high throughput screening (FACS, gene expression assays).
5. Identification of host factors that determine the size of the reservoir.

Systems/Technologies:

Cellular models
(cell lines / primary cells / single cell)

High-throughput analyses
(transcriptome / proteome)

Spatial integration / epigenetics
(histone modifications / chromatin organization, host gene environment, provirus orientation)



Output:

Dynamic modeling

Heterogeneity of latently infected cells

Molecular mechanisms

Nature and availability of transcription factors

Antisense transcription

Integration site-specific features