SCUOLA NORMALE SUPERIORE – PISA

Transcription silencing and sub-nuclear positioning of the HIV-1 provirus

Thesis submitted for the degree of Doctor Philosophiae

(Perfezionamento in Genetica Molecolare e Biotecnologie)

Candidate:
Mariacarolina Dieudonné

Supervisor:
Alessandro Marcello

Academic Year 2008/2009

List of figures

Abstract

The human immunodeficiency virus (HIV-1) is a retrovirus that integrates into host cell's chromatin for gene expression and replication. As integrated provirus HIV-1 is able to persist for long periods of time during antiretroviral therapy in quiescent memory T cells reservoirs. Understanding how these reservoirs are established, maintained, and reactivated is essential for developing methods to target and eventually eradicate HIV-1 infection.

Latency is likely established and maintained by numerous blocks at multiple steps in the HIV-1 gene expression pathway, which potentially complicates eradication strategies that aim at the purging of HIV-1 reservoirs from the infected patient. Recently, it has been proposed that the spatial distribution of genes within the nucleus contributes to transcriptional control allowing optimal gene expression as well as constitutive or regulated gene repression. Hence, the position of the provirus within the nucleus and its long-range interaction with other genomic regions could be another unexplored level of HIV-1 transcription control.

In order to gain insight in the conformation of chromatin at the site of HIV-1 integration we exploited seven different cell lines carrying a single latent provirus, which represent well-characterized models of HIV-1 latency. In the silenced state, the provirus was consistently found at the nuclear periphery. After induction of transcription the location of the transcribing provirus remained peripheral. Furthermore, in the J-lat A1 cell line, chromatin conformation studies revealed that the proviral vector is associated to a pericentromeric region of chromosome 12 (Ch12q12) located at the periphery of

the nucleus. Even though the location of the provirus did not change in transcriptionally active cells, the association between these two loci was lost. These results reveal a new mechanism of transcriptional silencing involved in HIV-1 post-transcriptional latency and reinforce the notion that gene transcription can occur also at the nuclear periphery.

Candidate's publications during the Ph.D. training:

Marcello A., Dhir, S, **Dieudonné**. Nuclear positional control of HIV-1 transcription in 4D. Nucleus 2010. [Epub ahead of print].

Dieudonné M, Maiuri P, Biancotto C, Knezevich A, Kula A, Lusic M, Marcello A. Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery. EMBO J. 2009 May 28. [Epub ahead of print].

Introduction

HIV-1: genome and structure

The acquired immunodeficiency syndrome (AIDS) is a life threatening disease caused by the human immunodeficiency virus type 1 (HIV-1). Following 25 years since its discovery the virus has killed more than 25 million people worldwide and it remains a major threat to public health and a challenge for drug development (Ho and Bieniasz, 2008). HIV-1 became the most studied virus in history and many years of research only partially unraveled the complexity of its life cycle. Studies on HIV-1 led to critical discoveries in antiviral drug development but also gave rise to new concepts in viral and cellular biology. However, the plasticity of viral sequences, the establishment of a latent state, and uncovered function of some auxiliary genes still remain a challenge in finding a cure for HIV-1 infection (Richman et al., 2009).

Genome organization

HIV-1 is a retrovirus, possessing an RNA genome that is replicated via a DNA intermediate. The full HIV-1 genome is encoded on one long strand of RNA and contains approximately nine thousand nucleotides. HIV-1 has several major genes encoding for structural proteins that are found in all retroviruses and several nonstructural (accessory) genes that are unique for HIV-1. The gag gene provides structural elements of the virus and the pol provides the replication enzymes. The Env gene codes for glycoproteins which are exposed on the surface of the viral envelope.

Accessory proteins (Tat, Rev, Vif, Vpr and Vpu) act at various stages of the various life cycle.

gag (Group-specific Antigen): codes for MA (p17), a matrix protein; CA (p24), the capsid protein; p6 and NC (p7) which build the nucleocapsid, and two spacer proteins (SP1,SP2).

pol (Polimerases): codes for viral enzymes such as reverse transcriptase (RT), intergrase (IN) and protease (PR).

env (Envelope): codes for gp160, the precursor to gp120 and gp41, proteins embedded in the viral envelope which enables the virus to attach to and fuse with the target cell.

tat (Trans-Activator of Transcription), *rev* (Regulator of Virion), *nef* (Negative Regulatory Factor), *vif* (Viral infectivity factor), *vpr* (Viral Protein R), *vpu* (Viral Protein U). The Figure 1 shows the structure of the HIV-1 genome.

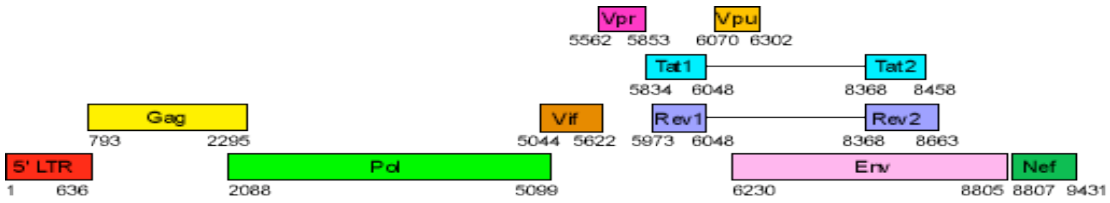


Figure 1. Structure of the HIV-1 genome.

The open reading frames for various polypetides are shown as rectangles. http://bioquest.org/bioinformatics/edgridbeloit/levels.html

Virion structure

HIV-1 is composed of two copies of a single-stranded RNA enclosed by a proteic core (capsid), which is surrounded by the plasma membrane (envelope) of host-cell origin in which the viral glycoproteins are inserted. All structural proteins, which are components of the HIV-1 virion, are derived from the Gag polyprotein. Gag is co-translationally myristoylated, which allows anchoring of this polyprotein to the membrane.

Traditionally two forms of HIV-1 particles, the immature form and the mature form, have been observed. The immature virion is a roughly spherical shell of radically extended uncleaved and multimerized Gag molecules (Figure 2B). During particle maturation, viral protease (PR) is activated and cleaves Gag generating a set of new proteins termed matrix protein (MA), capsid protein (CA), nucleo capsid (NC), spacer proteins 1 and 2 (SP1 and SP2, respectively) and p6 (Figure 2A). These newly processed proteins then reassemble to form the distinct layers of the mature virion: MA remains associated with the inner viral membrane (the 'matrix' layer), NC coats the viral RNA genome (the 'nucleocapsid' layer), and CA assembles into the conical capsid that surrounds the nucleocapsid and its associated enzymes: reverse transcriptase (RT) and integrase (IN) (Figure 2C).

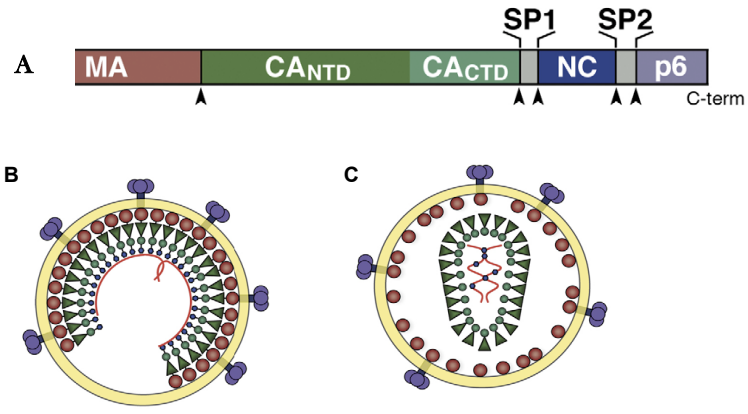


Figure 2. HIV-1 virion and structure of Gag polyprotein.

A) Domain structure showing the localization of MA, CA_{NTD} (N-terminal domain of the capsid protein) CA_{CTD} (C-terminal domain of the capsid protein) SP1, NC, SP2 and p6. Schematic models of the immature (B) and mature (C) HIV-1 virions (Ganser-Pornillos et al., 2008).

HIV-1 life cycle

Human immunodeficiency virus (HIV-1) infectious cycle begins with the interaction of viral envelope glycoproteins gp120/gp41 with CCR4 or CXCR5 co-receptors, which are found in cells of the immune system such as CD4+ T cells and cells of the monocyte – macrophage lineage (Berger et al., 1999). CCR5 and CXCR4 co-receptors provide a critical function for virus entry. Macrophage-tropic viruses (R5 viruses) binds CCR5 and T cell tropic viruses (X4 viruses) binds CXCR4. After virus adsorption, the viral and cell membranes fuse together, and the viral 'core' is released into the cytoplasm, where the virion-associated reverse transcriptase is activated and begins synthesizing viral cDNA. The trigger for the beginning of DNA synthesis is unknown. However, it appears that initiation of reverse transcription is strictly linked to the uncoating process and it is possible that the exposure of the retrotranscription complex to the significant concentration of deoxiribonucleotides in the cytoplasmic environment is what allows reverse transcriptase to begin to act. Disassembly of the lentiviral core is characterized by dissociation of the Gag encoded capsid protein, CA (Bukrinsky et al., 1993). Finishing of reverse transcription originates the HIV-1 preintegration complex (PIC). The PIC form by integrase (IN), MA, reverse transcriptase (RT), NC, Vpr and viral retrotranscribed cDNA is delivered in to the nucleus (Farnet and Haseltine, 1991). In contrast to most retroviruses, HIV-1 is able to infect nondividing cells such as

differentiated macrophages, thus PIC needs to be translocated into the nucleus, where, the HIV-1 integrase, catalyses the integration of the viral cDNA into the host-cell genome. Integrated provirus is transcribed by the host – cell RNA polymerase II (RNAPII) producing a copy of the viral RNA genome that can be used to infect other cells (Haseltine, 1991). Transcription of the integrated viral cDNA leads to the production of genomic and messanger RNA (mRNA) molecules that are transported to the cell cytoplasm. Translation of HIV-1 mRNAs leads to the production of different proteins. Assembly and budding creates new virions through the infected cell membrane. Virus infectivity is acquired after particle maturation due to the actions of protease (Figure 3) (Monini et al., 2004).

To reach the nuclear membrane, the particles must travel through the cytoplasm. HIV-1 transport has been shown to exploit the cellular cytoskeleton. In particular, initial movements at the cell periphery occur in association with actin (Bukrinskaya et al., 1998), while subsequent translocation to the nucleus takes place along the microtubule network, likely by interaction of PIC with the dynein-dependent motor complex (McDonald et al., 2002). After reaching the nuclear envelope PIC is translocated through the nuclear pore, most likely by relying on the cellular nuclear import machinery.

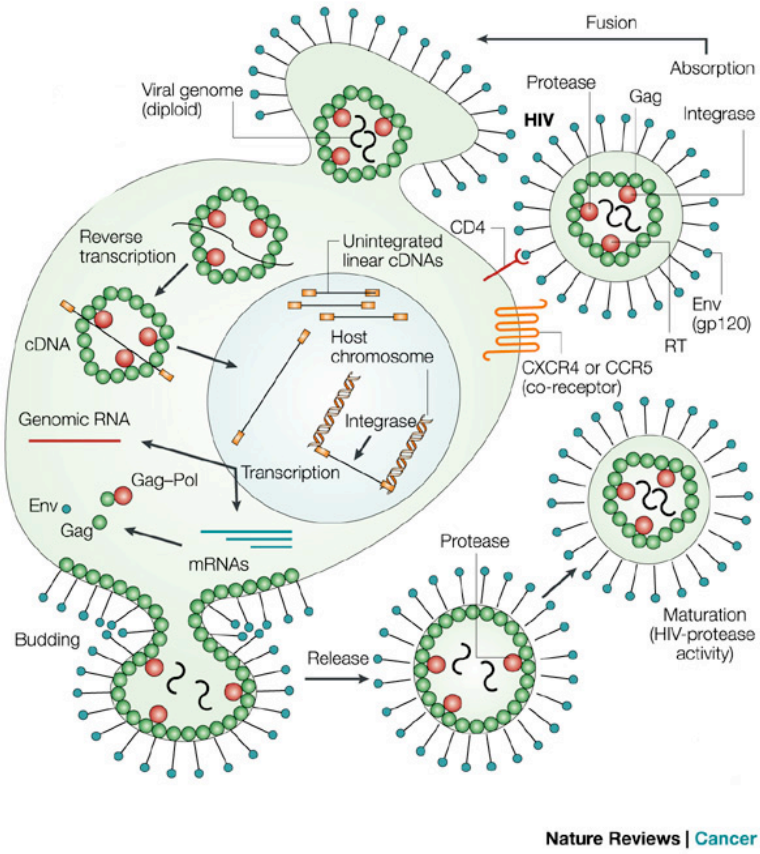


Figure 3. HIV-1 Life cycle.

After infection of a cell, a nucleoprotein reverse transcription complex containing genomic viral RNA associated with virion proteins is deposited in the cytoplasm. Among the virion proteins are enzymes that catalyse the synthesis of complementary DNA (reverse transcriptase; RT) and that catalyse the integration of the viral cDNA into genomic cellular DNA (integrase). Following reverse transcription, the viral cDNA in the complex is transported to the nucleus to form the integrated provirus. The virus-encoded protein Tat and its associated cellular partner, cyclin T, promote the expression of genomic and subgenomic viral transcripts, which are exported from the nucleus by the action of the virus protein Rev and its cellular co-factor, CRM1. Genomic viral RNAs are transported to sites of virus assembly for incorporation into progeny virions. Subgenomic RNAs act as templates for structural virion proteins and regulatory/accessory proteins that perpetuate the infection and have various roles in virus–host interplay (Stevenson, 2003).

Like all retroviruses, HIV-1 integrates into the host chromatin. The process of integration of the linear viral DNA is carried out by the viral integrase protein (IN) and cellular factors like for example LEDGF/p75 (Llano et al., 2004). It has been shown that LEDGF/p75 acts as a chromatin docking factor for IN. LEDGF/p75 tethers HIV-1 integrase to chromatin, protects it from degradation, and strongly influences the genome-wide pattern of HIV-1 integration. Depleting the protein from cells and/or over-expressing its integrase-binding domain blocks viral replication (reviewed in (Poeschla, 2008)). Other proteins have been proposed as positive regulators of HIV-1 integration. HMG I (Y) dramatically stimulates integration reactions in vitro, probably by inducing changes in DNA structure (Farnet and Haseltine, 1991). Similarly, the integrase interactor 1 (Ini1), a subunit of the SWI/SNF chromatin remodelling complex, has been proposed to stimulate the in vitro DNA-joining ability of integrase (Kalpana et al., 1994).

For a long time integration of HIV-1 and more in general of all retroviruses was believed to occur randomly into the host chromatin. Recent reports have challenged this notion showing a bias for HIV-1 integration into transcriptionally active genes (Han et al., 2004). In addition, contrary to other retroviruses that integrate into promoters, HIV-1 viral genomes have been shown to reside within the introns of active host genes (Lassen et al., 2004). Once integrated into host genome, HIV-1 can undergo active transcription allowing continuous rounds of infection.

In an infectious state, the integrated provirus is actively transcribed from its promoter located in the U3 region of the HIV-1 LTR. Once the HIV-1 pre-mRNA transcript is produced, it can be spliced in alternative ways to yield three classes of mRNA in the

nucleus: unspliced mRNA (9 kb), singly spliced mRNA (4 kb) and fully spliced mRNA (2 kb). All three classes of mRNAs must be exported to the cytoplasm and translated into viral proteins for the viral life cycle to proceed. The fully spliced mRNAs, which are the first viral transcripts that appear after infection, are exported to the cytoplasm. They follow the same pathway of cellular RNA (Cullen, 1998) leading to expression of the regulatory proteins Nef, Tat and Rev. The negative factor (Nef) is a modulator of cellular signaling pathways that optimizes the cellular environment for virus replication. The transcriptional activator (Tat) up-regulates the synthesis of full length viral mRNA by elongation of RNAPII. The regulator of viral gene expression (Rev) acts at post-transcriptional level and plays a main role in Rev-dependent export of unspliced and incompletely spliced HIV-1 transcripts from the nucleus to the cytoplasm leading to the export of genomic RNA (Dorman and Lever, 2000).

After two months, in patients on Highly Active Anti-Retroviral Treatment (HAART) the plasma levels of genomic RNA falls below the limit of detection. Therefore, it was initially assumed that prolonged treatment might lead to eradication of the virus in these patients (Ho, 1998). Unfortunately, it is now clear that long-live reservoirs of HIV-1 can persist for years in the presence of HAART. This reservoir is thought to consist mainly of latently infected resting memory CD4+ T cells (Chun and Fauci, 1999) (Pierson et al., 2000). Both host transcription factors and the viral Tat trans-activator have been proposed as limiting factors for transcriptional reactivation of latent HIV-1. However, since HIV-1 is found integrated into the genome of resting memory T cells, it has been proposed that the chromatin environment at the viral integration site may play a role in the transcriptional silencing of the HIV-1 genome (Jordan et al., 2003) (Jordan et al., 2001). Indeed, integrated HIV-1 has nucleosomes positioned in its 5' LTR that are

remodeled by deacetylase inhibitors, cytokines and Tat (Van Lint et al., 1996) (Lusic et al., 2003) (Marcello et al., 2004). Histones play an important role in regulating HIV-1 transcription since they integrate signals for repression, like the heterochromatin marker H3K9 trimethylation, Suv39H1 and HP1γ, and reactivation, like histone acetylation (Lusic et al., 2003). It is not clear for instance if also the three-dimensional (3D) nuclear architecture of the nucleus could be implicated in HIV-1 provirus regulation, as it has been proposed for cellular genes (Misteli, 2007).

Packaging of the genome RNA (gRNA) into a particle is a highly specific process that achieves selection of a single RNA species from the total capped polyadenylated mRNA in the infected cell (Kaye and Lever, 1999). Two copies of genomic RNA are encapsidated, usually linked at their 5' end through the dimer linkage site (Darlix et al., 1990). Dimerization is associated with encapsidation (Greatorex and Lever, 1998), but dimer stabilization occurs post-capture by the Gag protein involving the NC. Viral RNAs to be specifically packaged are identified by the presence of an RNA sequence named the packaging signal (ψ) placed within the 5'leader (Berkowitz et al., 1995; McBride and Panganiban, 1996). However, the major determinants of packaging of HIV-1 reside in a short RNA segment placed between the 3' splice donor and the Gag initiation codon (Lever et al., 1989). Deletion of the 5'leader sequence reduces the incorporation of gRNA into budding virions with little effect on other viral functions. The site of the packaging signal allows for the specificity of encapsidation that is limited to the unspliced HIV-1 mRNA.

The level of Gag may play a major role in determining the fate of the translated gRNA. Newly formed Gag molecules bind the gRNA via NC resulting in the inhibition of

translation and exposing the structural elements required to chaperone gRNA dimerization and the packaging process (DIS/SL3). Selective recognition of the gRNA is probably achieved bi Gag-NC molecules binding to SL1, SL2 and Sl3 of the Ψ signal (Brasey et al., 2003; Clever et al., 1995).

To a first approximation, all the information necessary for retroviral particle assembly resides in the Gag polypeptide. For example, Gag alone can form extracellular virus-like particles in the absence of other viral proteins (Gheysen et al., 1989), and Gag molecules can spontaneously assemble into spherical, immature virus-like particles in vitro (Campbell and Rein, 1999). All the viral proteins necessary for virion assembly and RNA genomes are transported to the plasma membranes close to lipid raft membrane domains, where the building of new virions begins. The Gag-Pol precursor binds to plasma membrane through the myristol group of the MA domain. The resulting virions bud from the plasma membrane but they are still incomplete. Their maturation is finished by viral protease (PR) that cleaves Gag and Gag-Pol. Further; Gag and Pol precursors are processed to originate the single core proteins, marix and the viral enzymes. The proteolytic activity ends when the virion is already detached from plasma membrane and results in the formation of mature infectious viruses.

HIV-1 promoter: LTR

The U3 region of the HIV-1 LTR functions as the viral promoter. The viral LTR promoter has a structure typical of promoters activated by cellular RNA polymerase II. It contains several binding sites for general transcription factors. The HIV-1 promoter contains the core promoter sequence (TATA box) and enhancer elements placed

upstream and downstream of TATA box. Immediately upstream of the TATA box are two tandem NF-κB binding sites and three tandem SP–1 binding sites. Immediately downstream of the start of transcription is the transactivation response element (TAR) (Figure 4). Three tandem-binding sites for the constitutively expressed Sp1 transcription factor are necessary for basal levels of LTR-directed RNA synthesis. Mutation of individual or pairs of Sp1 sites has little, if any, effect on the basal or Tat-transactivated levels of expression (Harrich et al., 1990). However, the mutation of all three Sp1 sites markedly reduces the response to Tat (Berkhout and Jeang, 1992). Two tandemly arranged binding sites (κB sites) are recognized by the dimeric transcription factors composed of several combinations of members of the Rel/NF-κB family of polypeptides (Baeuerle and Baltimore, 1996). The predominant complex that binds to the LTR κB sites in activated cells is NF-κB (p50/p65 heterodimer). Moreover, the arrangement of the transcription factor binding sites in the LTR may vary in different HIV-1 subtypes but the specific contribution of these changes has not yet been thoroughly investigated.

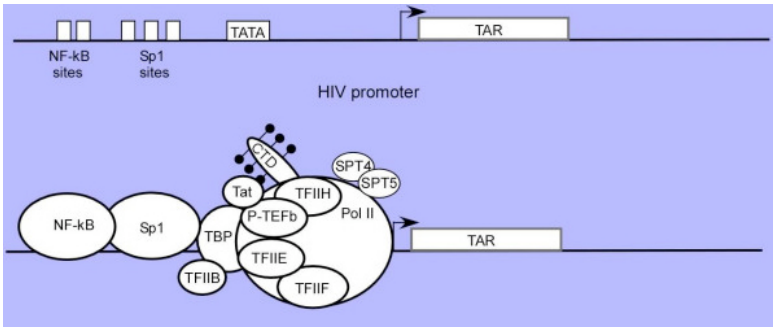


Figure 4. Structure of the HIV-1 viral promoter.

The HIV-1 promoter is comprised of a series of transcription control elements including NF-kB, Sp1, TATA box, RNA initiation site and the downstream TAR RNA enhancer element. In the presence of Tat, a complex interaction between activators which include NF-kB and/or Sp1 bind to the upstream

control region and interact with transcription factors which include, but may not be limited to, TBP, TFIIH, P-TEFb and RNA Pol II. Tat and P-TEFb facilitate the binding of TBP to the complex, setting the stage for binding of other basal transcription factors and assembly of the preinitiation complex. In the initiation complex, although both TFIIH and P-TEFb are present, the Pol II CTD is phosphorylated primarily by TFIIH at Ser5 (black) (Brady and Kashanchi, 2005).

Regulation of the viral gene expression

The regulation of human immunodeficiency virus type 1 (HIV-1) gene expression involves a complex event of pathological significance, which recapitulates general concepts of cellular transcription with some peculiarities. After the integration of HIV-1 in the host genome viral transcription depends of many cellular factors and viral Tat trans-activator.

Transcription Factors

HIV-1 transcription in T lymphocytes is regulated by DNA-binding proteins which recognize sites in the regulatory region of the viral long terminal repeat (LTR). Stimulation of T cells results in increased expression of HIV-1, mediated by the inducible transcription factors (Bielinska et al., 1989). Two of the key transcription factors (TFs), upon which successful transcription depends, are NF-κB and NFAT.

Nuclear factor-kappa B: NF-κB

NF-κB is a members of the Re1 family of transcriptional activator proteins. This transcription factor is ubiquitously expressed and highly inducible. It plays an important

role in the innate/adaptive immunity, and cellular survival through the induction of genetic networks (Barnes and Karin, 1997) (Karin, 1999).

The Re1 protein family has been divided into two groups based on differences in their structures, functions, and modes of synthesis (Baeuerle and Henkel, 1994). The first group consists of p50 (NF-KB1) and p52 (NF-KB2), which are synthesized as precursor proteins of 105 (p105) and 100 (p100) kDa, respectively. Initially quiescent, this first class of proteins contains a long chain of repeats that inhibits their function. p105, is the precursor to one of NF-kB's monomers. Upon activation, the repeat chain is cleaved and P105 becomes p50. The second group generated by proteolytic processing, has a so-called Rel homology region (RHR) that contains one or more transcriptional activator region on the C-terminus side, dimerization domains and a nuclear localization signal. This group, which includes p65 (Re1A), Rel (c-Rel), RelB, and the Drosophila Rel proteins dorsal and Dif, are not synthesized as precursors.

Members of both groups of Rel proteins can form homo or heterodimers; such as, p50-p65 heterodimer (termed also NF-κB), which is the most abudant and biologically active (Huxford et al., 1998). NF-κB is a cytoplasmic complex whose nuclear translocation is controlled by its association with a family of inhibitory proteins, termed I-κBs (Baeuerle and Baltimore, 1996). Together, p50 and p65 dimerize around a 10 base pair region known as a -κB site, forming the NF-κB transcription factor. In unstimulated cells, the p65 subunit of NF-κB is retained in the cytoplasm through its interaction with inhibitor proteins. An inducing signal leads to the phosphorylation of I-κB and p105 (Henkel et al., 1993). This phosphorylation is thought to be the signal for I-κB degradation and p105 processing (to become p50), both of which generate active

NF-κB dimeric complexes that translocate to the nucleus and activate genes containing Rel protein-binding sites (κB sites) (Figure 5). In other words, activation of NF-κB results in the targeted proteolysis of I-κB, releasing NF-κB to enter the nucleus and bind to specific sequences in target promoters. The genomic actions of NF-κB are influenced by the stimulus applied and the promoter context/chromatin structure in which it binds (Tian and Brasier, 2003).

The subunit composition of the ReI complex influences its subcellular localization, transactivation potential, and mode of regulation. In addition to the heterodimer, p50/p65, homodimers of the Rel proteins exist also. The p50 homodimers do not induce transcription. They are thought to be used as post-induction repressors, competing with NF-κB following transcription activation subsequent to a viral invasion (Verma et al., 1995) (Baeuerle and Baltimore, 1996).

In the case of HIV-1, the enhancer region spanning -104 to -80 contains two consensus NF-κB binding sites that are identical with the κB site found on the promoter of the immunoglubulin (Ig) κ light-chain, also know as Ig κB site. The upstream and downstream κB sites on the HIV-1 LTR are referred to as Core II and Core I respectively (Figure 4) (Cron et al., 2000). The sequences of the κB sites and the four nucleotides spacer are highly conserved on most isolated of HIV-1 (Jeeninga et al., 2000) (Rodriguez et al., 2007).

The reactivation of silent HIV-1 provirus through NF-κB occur by the induction of multiples changes at the promotor of latent proviruses, including the recruiment of TFIIH and RNAPII (Kim et al., 2006) and changes in the local chromatin structure at

the HIV-1 LTR (Gerritsen et al., 1997) by recruiting histone acetyltransferases, such as CBP and p300 and chromatin remodeling factors (Lusic et al., 2003).

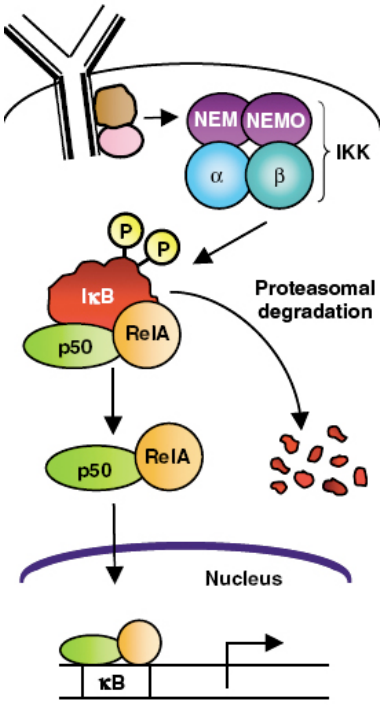


Figure 5. NF-κB pathway

Two pathways lead to NF-κB activation. In the classical NF-kB pathway, NF-kB dimers such as p50/RelA are maintained in the cytoplasm by interaction with an independent IkB molecule (often IkBa). Pro-inflammatory stimuli and genotoxic stress leads to IKKβ- and IKK-dependent phosphorylation of IκB, which results in proteasomal degradation and subsequent release of the NF-κB dimers. Activation of this pathway leads to increased transcription of genes in three functional classes. The alternative pathway works independently of IKKβ and IKK (Gilmore, 2006).

NF-κB activators:

TPA (12-O-tetradecanoyl phorbol acetate)

Phorbol is a natural, plant-derived organic compound. It is a member of the tigliane family of diterpenes. The most common phorbol ester is 12-O-tetradecanoylphorbol-13-acetate (TPA), also called phorbol-12-myristate-13-acetate (PMA), which is a potent tumor promoter able to activate the signal transduction of enzyme protein kinase C (PKC). The effects of TPA on PKC result from its similarity to one of the natural

activators of classic PKC isoforms, diacylglycerol. Therefore, TPA mimic T-cell mitogen activation and activate the NF-κB pathway.

Dinter and collaborators demonstrated that the HIV-1 enhancer is activated specifically by TPA in several non-lymphoid cell types, and that this transcriptional regulation can be reproduced in a cell-free system. In vitro transcription experiments revealed a 6-fold activation of the HIV-1 promoter in nuclear extracts prepared from TPA induced HeLa tk- cells (Dinter et al., 1987). Other studies done by Verdin, et al. have been shown that the *5'* region of integrated HIV-1 DNA contains an array of precisely positioned nucleosomes. These nucleosomes define two large nucleosome-free regions corresponding to the promoter region (nt 200-452) and to the primer binding site region (nt 610-720). A nucleosome located in the R-U5 region in basal conditions was specifically disrupted following TPA or TNF-α treatment independent of transcription and independent of DNA replication (Verdin et al., 1993). Thus, TPA is used as a strong activator of HIV-1 transcription contributing to our understanding of HIV-1 regulation and ultimately AIDS pathogenesis.

Tumor Necrosis Factor: TNF

NF-κB activation is stimulated by a wide variety of both intra- and extracellular stimuli, including inflammatory cytokines, viral and bacterial products, growth factors, and prooncogenic signals (Baldwin, 2001). Among the proinflammatory cytokines, tumor necrosis factor alpha (TNFα) is one of the most potent activators of NF-κB in cell types possessing TNFα receptors (Baud and Karin, 2001). TNFα is a potent proinflammatory cytokine that plays an important role in immunity and inflammation, and in the control of cell proliferation, differentiation and apoptosis. TNFα is

produced primarily from macrophages and functions in the early phases of infection by activating and recruiting other immune cells through the production of chemokines and other proinflammatory cytokines (Tracey and Cerami, 1994). Overproduction of TNFα is also thought to contribute to the pathophysiology of several diseases including septic shock, cancer, AIDS, diabetes, and rheumatoid arthritis (Tracey and Cerami, 1994).

HIV-1 Tat protein amplifies the activity of tumor necrosis factor TNFα, stimulating HIV-1 replication through activation of NF-kB. In HeLa cells stably transfected with the HIV-1 tat gene (HeLa-tat cells), expression of the Tat protein enhanced TNFα-induced activation of NF-kappa. As we mentioned before, in an uninduced state, cellular I-kB proteins interact with NF-kB dimers to mask their nuclear location sequence, retaining the complex of NF-kB and I-kB in the cytoplasm. TNFα, like a variety of other inducers, can stimulate proteolytic degradation of I-kB by the proteasome, allowing the translocation of NF-kB into the nucleus and subsequent binding to its DNA motifs (Baeuerle and Baltimore, 1988) (Blank et al., 1992).

For TNFα-induced IKK and NF-κB activation, two TNF effector molecules are essential, the death domain kinase receptor interacting protein (RIP) and TRAF2. In response to TNFα treatment, IKK is recruited to the TNFα receptor complex through TRAF2, and its activation requires RIP (Liu, 2005).

Tat trans-activator

The *trans*-activator protein (Tat) is an unusual transcription factor because it interacts with a *cis*-acting RNA enhancer element, TAR, present at the 5' end of all viral

transcripts. The function of Tat has been described for the first time by Sodroski *et al* who noted that synthesis of reporter genes placed under the control of the LTR increased 200- to 300-fold in the cells which have been previously infected by HIV-1 (Sodroski et al., 1985). They reasoned that the induction, or "transactivation" of transcription was due to the presence of a novel *trans* – activating factor, which they named Tat.

Transcription of the HIV-1 provirus is characterized by an early, Tat-independent and a late, Tat-dependent phase. Transcription from the HIV-1 LTR is increased several hundred-fold in the presence of Tat and the ability of Tat to activate transcription is an essential step in the HIV-1 life cycle.

The Tat – TAR RNA interaction

Tat function is mediated by the TAR RNA target element encoded within the LTR. TAR is located downstream of the initiation site for transcription (nucleotides +1 and +59). TAR RNA sequence forms a highly stable, nuclease – resistant, stem – loop structure. The stem – loop structure is located between +19 and +43 nucleotides which is the minimal sequence for transactivation. Because of its location in the R domain of the LTR, TAR is present in both viral RNA and DNA. Important structural features in the HIV-1 TAR are the stem, the pyrimidine – rich bulge, and the loop.

Tat binds TAR. Mutations that destabilize the TAR stem - loop by disturbing base – pairing abolish Tat – stimulated transcription. Tat forms a complex with TAR and recognizes the pyrimidine – rich bulge (+23 to +25) below the apex of the stem – loop. Tat interacts with the first nucleotide (U23) in the bulge and with two nucleotide pairs

on either side of bulge (G26 to C39 and A27 to U38 above the bulge, and A22 to U40 and G21 to C41 below the bulge).

Tat-mediated transcriptional activation

The interaction of Tat with TAR permits activation of HIV-1 transcription by promoting the assembly of transcriptionally active complexes at the LTR by multiple protein–protein interactions. Over the last few years, a number of cellular proteins have been reported to interact with Tat and to mediate or modulate its activity. These include general transcription factors, among which are TBP, TAFII250, TFIIB, TFIIH (Kashanchi et al., 1994) (Parada and Roeder, 1996), RNA polymerase II (Wu-Baer et al., 1995); transcription factor Sp1; the cyclin subunit of the positive transcription elongation factor complex (P-TEFb), cyclin T1, and different transcriptional co-activators that possess histone acetyltransferase activity (Figure 4). These interactions with P-TEFb and HATs appear to be of particular importance in the transcriptional activation of the viral promoter.

P-TEFb

Transcription elongation is up regulated by the positive transcription elongation factor b (P-TEFb), which consists of one component of the cyclin T family and of the cyclin-dependent kinase Cdk9. P-TEFb enables transition from abortive to productive transcription elongation by phosphorylating CTD in RNAPII and negative transcription elongation factors.

Although the HIV-1 LTR contains DNA binding sites for several transcription factors, in the absence of Tat, there is no expression of viral genes. However, there is

production of short, non-polyadenylated RNAs that include TAR RNA. The instability of synthesizing full-length viral transcripts is caused by the low processivity of RNAPII, which is overcome by P-TEFb. Thus, P-TEFb is crucial for efficient transcriptional elongation of HIV-1.

During the pre-initiation of HIV-1 transcription the CTD of RNA polymerase II (RNAPII) is phosphorylated by CDK-7, a component of TFIIH initiation factor (Figure 6A). CTD phosphorylation is an early step associated with the clearance of the promoter. Following clearance of the promoter, the phosphorylated RNAPII is able to transcribe through the TAR region. When the TAR RNA stem-loop structure is synthesized, the Tat protein binds to it and by interacting with cyclin T1, activates P-TEFb (Figure 6A). Then, CDK9 is recruited in close proximity to RNAPII and N-TEF, which become hyperphosphorylated allowing productive elongation (Figure 6B) and, as recently reported, also stimulating subsequent rounds of transcription complex assembly at the HIV-1 promoter and re-initiation (Raha et al., 2005). This molecular event is associated with increased transcriptional processivity (Figure 6B). The formation of the P-TEFb-TAR-Tat tripartite complex is an essential step towards the assembly of the processive RNAPII machinery at the LTR promoter (Bieniasz et al., 1998)

Recently have been indentified "hot-spot" genes that are frequently targeted in latent but not productively HIV-1 infected cells. One of these genes is the gene encoding bromodomain containing protein 4 (BDR4), which binds to the P-TEFb complex (Bisgrove et al., 2007). In many cell types P-TEFb exists in two complexes: large and small. P-TEFb is sequestered in an inactive ribonucleoprotein large complex which

additionally contains HEXIM1, heterogeneous nuclear ribonucleoproteins (hnRNPs) and the 7SK small nuclear RNA (snRNA) (Figure 6C). In response to stimuli such as stress, UV light or actinomycin D 7SK RNA, as well as HEXIM proteins and hnRNPs are released and P-TEFb is activated forming the small complex (Figure 6D) (Barrandon et al., 2007) (Nguyen et al., 2001). A fraction of this free P-TEFb is bound to BDR4, which may tether P-TEFb to actively transcribe genes. In vitro experiments have shown that the partitioning of P-TEFb between the active and inactive complexes changes rapidly in response to stress signals, which disrupt the 7SK ribonucleoprotein complex and promote the formation of BRD4/P-TEFb complexes (Figure 6D).

In vitro experiments demonstrated that HIV-1 Tat protein competes with Hexim1 for binding to the cyclin T and that Tat expression releases P-TEFb from sequestration by Hexim1 and 7SK. Thus, Tat can access the active P-TEFb complex by competing with Hexim1 and can substitute for Hexim1 in the absence of stress signals. The carboxyl terminus of BRD4 mediates its interaction with P-TEFb and targets the site within P-TEFb bound by Hexim1 and Tat. The HIV-1 transactivator Tat and BRD4 competed for binding to P-TEFb, and the carboxyl-terminal BRD4 peptide suppressed HIV-1 transcriptional activation by Tat at low concentrations (Bisgrove et al., 2007).

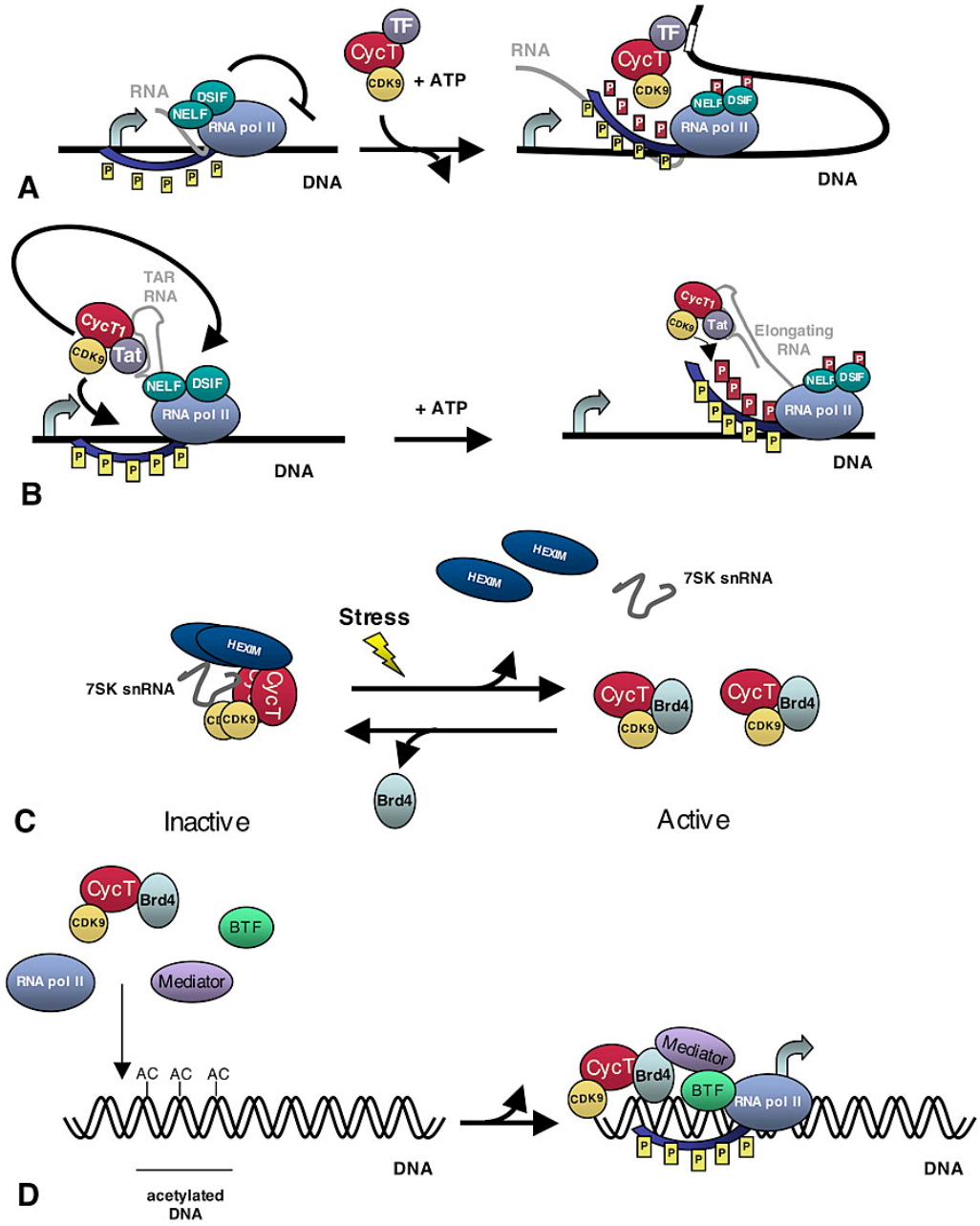


Figure 6. P-TEFb activity.

(A) P-TEFb is recruited near paused RNAPII via specific transcription factors and/or activators (TF), such as myc or NF-kappaB. P-TEFb then phosphorylates negative elongation factors, NELF and DSIF, as well as the C-terminal domain of RNAPII (red *P*), which allows for productive elongation (yellow *P* represent phosphorylation by TFIIH). Both DSIF and NELF remain bound to RNAPII and may play other roles during elongation. (B) Tat recruits P-TEFb to the TAR RNA structure that forms at the *5'* end of the nascent HIV RNA. Then, CDK9 phosphorylates RNAPII as well as NELF and DSIF to allow

for productive elongation (C) 7sk/HEXIM binds and sequesters P-TEFb in an inactive form in the nucleus. Stress signals, induce release of P-TEFb from it's inhibitory complex. Release of active P-TEFb allows for its association with Brd4 and perhaps other transcription factors which target P-TEFb to specific gene targets. (D) Brd4 recruits active P-TEFb to acetylated DNA during transcription initiation. The Mediator may also facilitate recruitment of P-TEFb via its association with Brd4 (Marshall and Grana, 2006).

Histone Acetyltransferase: HATs

HATs are enzymes that acetylate histone proteins by transferring an acetyl coenzyme A (acetyl-CoA) onto the ε-amino group of specific lysine residues present in the amino-terminal tails of each of the core histones, resulting in the neutralization of a single positive charge (Allfrey et al., 1964). Acetylation has long been linked with events both in chromatin synthesis and assembly, as well as the establishment of transcriptionally active chromatin. Acetylation is an energy-intensive, dynamic phenomenon, the steady-state balance of which is mediated by the opposing activities of HATs and deacetylase enzymes. Each of these enzymes generally belongs to one of two categories (Brownell and Allis, 1996) (Garcea and Alberts, 1980). Type A are located in nucleus, this kind of HATs are involved in the acetylation of all four nucleosomal histones and have long been thought to promote transcription related acetylation. In contrast B type HATs, located at the cytoplasm, are believed to have an housekeeping role in the cell, acetylating newly synthesized free histones in the cytoplasm for transport into the nucleus, where they may be deacetylated and incorporated in to chromatin (Allis et al., 1985) (Guo et al., 1995).

The LTR acts as a very strong promoter when analyzed as naked DNA *in vitro*, but when it is integrated into the cellular genome is almost silent. Therefore, chromatin

conformation imposes inhibition onto the integrated promoter. Experiments performed both *in vivo* (Verdin et al., 1993) (El Kharroubi et al., 1998) and *in vitro* using the HIV-1 promoter reconstituted into chromatin (Sheridan et al., 1997) have shown that, independent from the integration site, nucleosomes at the *5*' LTR are precisely positioned with respect to the *cis*-acting regulatory elements. In the transcriptionally silent provirus, these nucleosomes define two large nucleosome-free areas, one associated with the promoter/enhancer in the U3 region and the other spanning the primer-binding site immediately downstream of the *5*' LTR. These two open regions are separated by a single nucleosome called nuc-1 that is specifically and rapidly destabilized during transcriptional activation. Complexes containing HAT activity facilitate transcriptional activation by modulating nucleosomal repression of specific promoters. This event destabilizes the histone-DNA interaction. The role of acetylation in the processing and deposition of displaced histones, could be related with not only the formation of nucleosomes from newly synthesized histones but also reconstitution of nucleosome core particles, including the reassembly of the higher order structure of chromatin (Csordas, 1990).

The HAT proteins responsible for the TAR-dependent Tat transactivation include the transcriptional co-activators p300, the highly homologous cAMP-response element binding protein (CREB) binding protein (CBP), the p300/CBP-associated factor (P/CAF), the general control non-derepressible-*5* (GCN*5*) factor, the TIP60 protein, and the general transcription factor TAFII250. Lusic *et al.* has observed acetylation of histones H3 and H4 at distinct nucleosomal regions and the recruitment of the above mentioned HATs to the viral LTR promoter upon HIV-1 activation either with Tat or phorbol esters (Lusic et al., 2003).

The Tat-TAR-P-TEFb interaction increases the processivity of RNAPol II giving support that Tat plays role in transcription elongation. The initiation model quickly lost support when Kao *et al.* reported that in the absence of Tat the majority of RNA polymerases initiating transcription stall near the promoter (Kao et al., 1987). On the other hand it is also clear that optimal Tat transactivation of HIV-1 gene expression requires upstream transcription co-factors. Along these lines it has been reported that Tat physically interacts with the pre-initiation complex including transcription factors such as Sp1 (Jeang et al., 1993), TATA binding protein (TBP) (Majello et al., 1998), cylinE/cdk2 (Deng et al., 2002), TFIIH, Tip60 (Kamine et al., 1996) and RNAPII (Cujec et al., 1997). The evidence of Tat interaction with histone acetyltransferases supports the role of Tat in chromatin remodeling before the onset of viral transcription. It is possible therefore that Tat facilitates chromatin remodeling, pre-initiation complex assembly and transcription elongation in a sequential manner that leads to high levels of HIV-1 transcription. All of these activities of Tat are probably exerted by the transient formation of large subnuclear complexes at the site of HIV-1 transcription and must therefore be spatially and temporally regulated (Molle et al., 2007). It has been demonstrated (Marcello et al., 2003) that cyclin T1 and p300 interact with the promyleocytic leukemia (PML) protein within specific subnuclear compartments that are coincident with PML nuclear bodies. This observation suggests that PML bodies could regulate Tat activity by modulating the availability of essential factors to the transcriptional machinery.

Latency

The major problem in curing AIDS is a reservoir of memory CD4+T cells, in which HIV-1 resides in a latent form, allowing the virus avoid host immune responses and antiretroviral drugs. Because latency represents a barrier to curing HIV-1 infection, it is crucial to understand the molecular mechanisms which are involved. It is becoming clear that HIV-1 latency is a complex multifactorial phenomenon that ultimately results from the profound differences between resting and activated CD4+T cells (Lassen et al., 2004) (Marcello, 2006).

There are two forms of HIV-1 latency, in the first one, full length viral DNA molecules; the final product of reverse transcription, remain unintegrated in resting CD4*T cells. In this case, when the infected cells are activated by signals from the microenvironment of the lymphoid tissues this extrachromosomal viral DNA integrates into the host genome and produces viral progeny (Bukrinsky et al., 1991) (Zack et al., 1990). Even though preintegration latency is quantitatively dominant in untreated individuals, this labile form of latency decays rapidly when the new infection of resting CD4*T cells is halted with HAART (Pierson et al., 2002) (Zack et al., 1990), making it less of a clinical concern comparing to the stable post integration form. The post integration latency, which represent the second kind, occur when infected CD4*T cells lymphoblast revert to a resting memory state after the integration of the viral genome into the host genome.

Activation of gene expression relies on the combined activity of a series of cellular factors that respond to different external stimuli, and on the function of the viral regulatory protein Tat. In consequence transcriptional activation is a result of both chromatin remodeling and the recruitment of elongation competent RNA polymerase

II complexes onto the integrated promoter, two events that require the coordinate, but transient assembly of different protein complexes. Today many mechanisms playing roles in the regulation of viral expression are well described and they involve interplay between host cell and HIV-1 proteins. Between the most important are:

Latency is achieved and maintained by several mechanisms that predominantly operate at transcription level (Lassen et al., 2004) (Contreras et al., 2006) such as:

Heterochromatin formation at the site of provirus integration, which can allow or not the accessibility of the integrated provirus to the transcriptional machinery (Jordan et al., 2001) (Jordan et al., 2003) (du Chene et al., 2007).

The absence in resting CD4+T cells of the active forms of the host transcription factors (TFs), that are necessary for HIV-1 gene expression (Nabel and Baltimore, 1987) (Marcello et al., 2003) (Karn, 1999).

Transcription interference (TI) from the endogenous gene harboring the provirus (Han et al., 2008) (Lenasi et al., 2008).

The presence of transcriptional repressors (He and Margolis, 2002)

The premature termination of HIV-1 transcripts due to the absence of the viral protein Tat and tat associated host factors (Kao et al., 1987).

All these mechanisms show a general picture of the potential transcriptional blocks in HIV-1 latency (Figure 7).

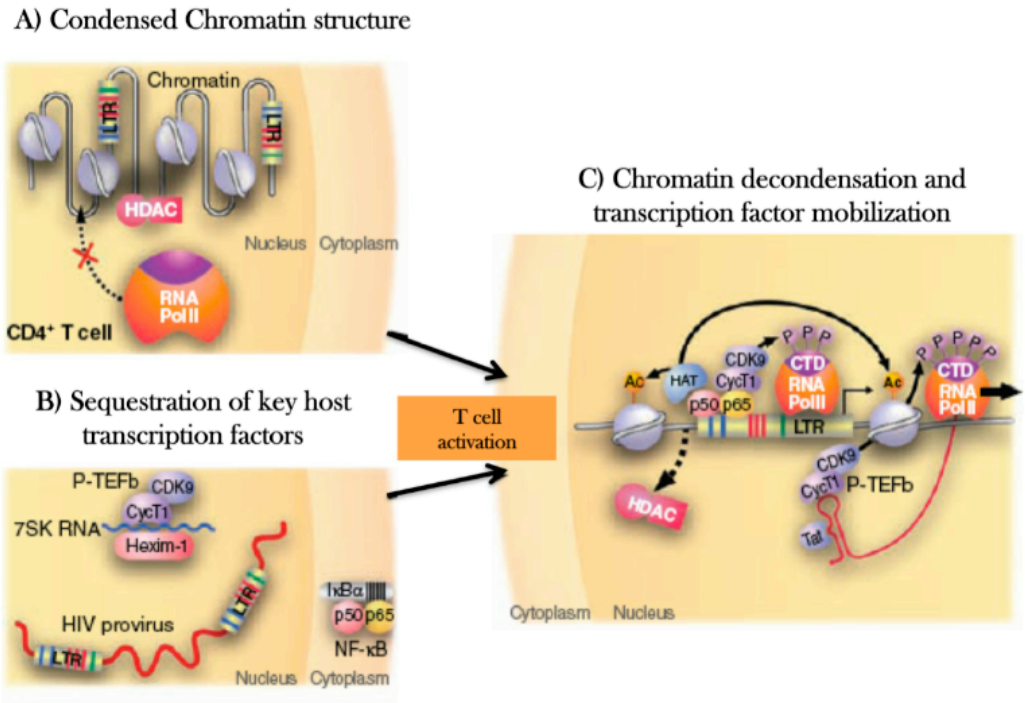


Figure 7. Potential transcriptional blocks in HIV-1 latency.

A) Proviral latency is maintained in part, by the action of several transcriptions factor that recruit HDCAs and other complexes to the HIV-1 long terminal repeat (LRT) promoter, which results in histones modifications within chromatin at the HIV-1 promoter that limit the ability of RNA polymerase to initiate transcription. B) Key cellular factors that are required for robust HIV-1 transcription, such as NF-κB or the P-TEFb-cyclin complex, are sequester in resting CD4+Tcells by cellular regulatory complexes (e.g. inhibitor of nuclear factor κB (IkB)). Release and mobilization of these factors is required for proviral expression. C) When histones acetyltransferases (HATs) substitute the effects of HDACs, coactivators such as NF-κB cab recruit RNA polymerases complexes. Production of Tat allows the recruiment of P-TEFb, mediating and explosive increase in transcription and the escape of provirus from latency (Richman et al., 2009).

In addition, there are also mechanisms that block post-transcriptional activity such RNA interference that possibly binds as miRNA to HIV-1 RNA which is in turn not translated into viral proteins (Figure 8) (Lassen et al., 2006) (Huang et al., 2007).

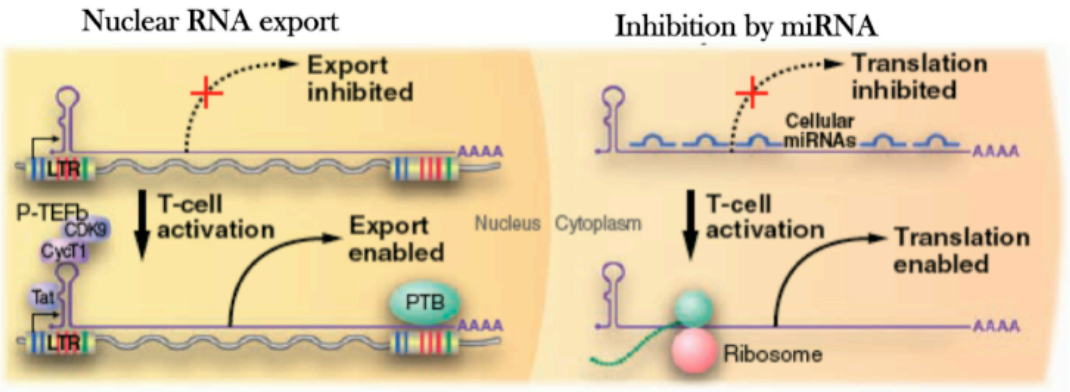


Figure 8. Post-transcriptional blocks of HIV-1 latency

In the left panel (nucleus), the initial wave of Tat production may be further restricted by inefficient export of multiple spliced HIV-1 mRNAs, relieved upon cellular activation by enhanced ecpression of PTB. In the right panel (cytoplasm) cellular miRNAs that bind HIV-1 mRNAs may also restrict translation of early expressed HIV-1 mRNAs and so reduce tat production (Richman et al., 2009).

Transcription Interference

Integration of HIV-1 in the host genome could lead to occurrences of a phenomenon called transcriptional interference (TI), where an ongoing transcription from a host promoter would prevent pre-initiation complex assembly on the 5'LTR, thus interfering with the viral transcription or viceversa. (Lenasi et al., 2008) (Han et al., 2004).

TI is a direct consequence of the nature of HIV-1 integration sites in vivo. There are several potential molecular mechanisms for TI. Transcriptional activation of the

upstream promoter reduces transcription from the downstream promoter. Typically, this occurs when upstream transcription fails to terminate, so that the polymerase 'reads through' into the downstream gene, thereby interfering with initiation at the downstream promoter (Lassen et al., 2004). Footprinting studies showed that upstream read-through transcription complexes impaired the binding of Sp1 to the downstream HIV-1 promoter. This promoter occlusion mechanism has been directly demonstrated in pol I transcription systems (Henderson et al., 1989) and the effect could be blocked by the insertion of a strong transcriptional terminator between the two promoters. In addition, TI affects the gene expression from the HIV-1 long terminal repeat (LTR) in a system in which tandem HIV-1 promoters were integrated into the genome of HeLa cells (Greger et al., 1998), altered levels of transcription can result from changes in supercoiling caused by the transcription of linked genes (Dunaway and Ostrander, 1993), also in the case of two adjacent genes, competition for trans-acting factors might lead to TI. The spreading of epigenetic changes, such as methylation or histone modification, might also be involved (Eszterhas et al., 2002). For HIV-1, promoter occlusion remains a probable mechanism for TI, given that the observed distance between the HIV-1 LTR and the relevant host gene promoter is, on average, 30 kb (2–429 kb) (Han et al., 2004).

Because HIV-1 genomes are found predominantly in the introns of active host genes both in activated and resting CD4+T cells, it is likely that the virus has evolved to function in this environment. The orientation of the HIV-1 genomes with respect to the host genes is not biased (Han et al., 2004). However, strong TI can occur between two adjacent transcription units, oriented in convergent, divergent or in tandem order (Eszterhas et al., 2002).

There are no known transcriptional terminators that would protect the HIV-1 LTR, and HIV-1 sequences can be found in the primary transcripts of host genes (Han et al., 2004). Therefore, TI is a virtually inevitable consequence of the nature of HIV-1 integration sites and might contribute to HIV-1 latency. In activated CD4+T cells, HIV-1 gene expression might be efficient because the concentration of crucial host transcription factors is high enough to overcome TI. In consequence, TI might become more important as the cell reverts to a resting state and crucial transcription factors become limiting. However, it has been also observed coo expression of the endogenous genes and HIV-1 at the same loci.

Chromatin environment

The levels of viral expression are influenced by the chromatin environment of cellular DNA at the integration site of the provirus. Recently the first in vivo analysis of integration sites (Han et al., 2004) demonstrate that in purified resting CD4+T cells from patients on HAART, the cell population in which latent HIV-1 has been most extensively characterized in vivo, viral integration into transcriptional units was strongly favored (93%) (Han et al., 2004). In contrast, data presented by Lewinski, using latent cells generated in vitro, suggest that three chromosomal features correlated with inducible expression: centromeric heterochromatin, gene deserts, and highly active host transcription units, all these situations may contribute to the latent population (Lewinski et al., 2005).

Even in the case that actively transcribed genes could favor HIV-1 integration, there are several factors that could form repressive chromatin structure on the HIV-1 LTR which encompasses the viral HIV-1 promoter. In the transcriptionally silent state, two

nucleosomes are found in the HIV-1 LTR. Nuc-(0) spans the region from positions −415 to −255 and Nuc-(+1) from positions +1 to +155 with respect to the transcription start site of HIV-1 genome. The region covered by and between these two nucleosomes contains recognition elements for the sequence-specific host TFs (Steger and Workman, 1997). The most important ones, for transcription of the HIV-1 genome are: NF-κB, nuclear factor of activated T cell (NFAT), upstream stimulating factor (USF), Ets1 (a protein that binds ETS - a winged helix-turn-helix domain), lymphoid enhancer binding factor-1 (LEF-1), stimulatory protein 1 (Sp1) and leader binding protein-1 (LBP-1) (Figure 9). The availability of certain cellular proteins directs the binding of chromatin remodeling complexes and consequently the formation of the preinitiation complex.

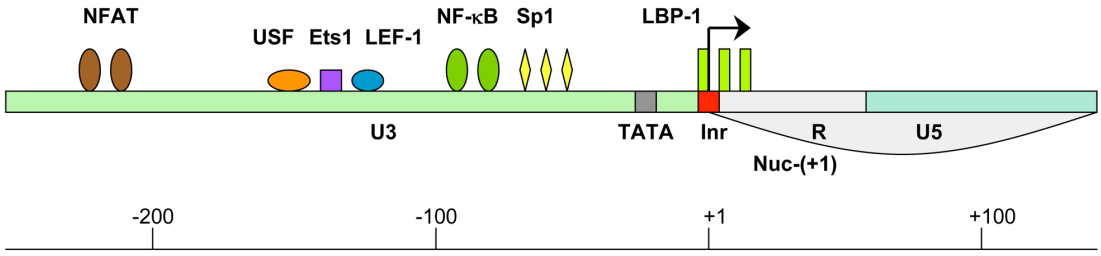


Figure 9. Binding sites for the critical TFs within the HIV-1 LTR

LTR is composed of three regions: 3′ untranslated region (U3) (green), transcription regulatory region (R) (grey) and 5′ untranslated region (U5) (blue). The following host TFs which bind to the HIV-1 LTR are designated: NFAT (brown), USF (orange), Ets1 (purple), LEF-1 (blue), NF-kB (green), Sp1 (yellow) and LBP-1 (yellow-green). TATA box and initiator element are marked in grey and red, respectively. Nucleosome 1 location is from +1 to +155 (Nuc-(+1)). Arrow marks the transcription start site (Contreras et al. 2009).

NF-κB plays a crucial role in the repression/de-repression of local chromatin structure surrounding the HIV-1 LTR. When the levels of free NF-κB in the cell are too low, the

same binding sites are occupied by p50:p50 homodimer which binds the histone deacetylase 1 (HDAC1). HDAC1 promotes deacetylation of surrounding histones, compaction of chromatin and consequently suppression of gene expression (Zhong et al., 2002). Williams et al. demonstrated that the recruitment of NF-κB to the HIV-1 LTR relieves this repressive chromatin environment by removing HDAC1 (Williams et al., 2006b). In this situation, HATs bind to the LTR and cause histone acetylation inducing the relaxation of chromatin structure and activation of gene transcription (Kurdistani and Grunstein, 2003).

Chromatin remodeling

The importance of chromatin structure in the regulation of eukaryotic gene transcription has become much more widely accepted over the past few years. It has been clear for a decade that histones contribute to the regulation of transcription both in vitro and in vivo (Clark-Adams et al., 1988) (Han and Grunstein, 1988). More recent studies have led to the striking observation that several proteins complexes involved in transcription regulation, can function, at least in part by modifying histones or altering chromatin structure (Armstrong and Emerson, 1998) (Kingston and Narlikar, 1999). While it is clear that many of these proteins complexes have functions in addition to chromatin modification, they illustrate the importance of chromatin structure as a part of transcription regulation mechanism.

The most widely characterized chromatin-modifying complexes studied to date can be classified in two major groups, based on their modes of action as follows: (i) histones acetyltransferase (HAT) and histone deacetylase (HDCA) complexes, which regulate

the transcriptional activity of genes by determining the level of acetylation of the amino terminal domais of nucleosomal histones associated with them, and (ii) ATP dependent complexes, which use the energy of ATP hydrolysis to locally disrupt or alter the association of histones with DNA.

Histone modifications

The flexible N-terminal tails of the four histones (H2A, H2B, H3 and H4) undergo a range of post-translational modifications, including acetylation, methylation, phosphorylation and ubiquitination. Histone modifications are indicators of active or repressed chromatin, and the proposed "histone code" hypothesis suggest that combination of specific histone modifications create a complex, functional hierarchy for chromatin regulation. Specific histone modifications are responsible for the compartmentalization of the genome into distinct domains, such as transcriptionally silent heterochromatin and transcriptionally active euchromatin (Martin and Zhang, 2005). The ability of the histone code to dictate the chromatin environment allows it to regulate nuclear processes, such as replication, transcription, DNA repair, and chromosome condensation (Kouzarides, 2007).

Next to DNA methylation, histone acetylation and histone methylation are the most well-characterized epigenetic marks. Euchromatin is characterized by a high level of histone acetylation, which is mediated by histone acetyl transferases (HATs). Conversely, histone deacetylases (HDACs) have the ability to remove this epigenetic mark, which leads to transcriptional repression (Bartova et al., 2008). At least two general and distinct roles have been proposed for the acetylation of the amino-terminal histone tails. The first is based on circumstantial evidence that has for some time linked

acetylation with events that apparently require the attenuation of DNA-histone contacts within chromatin. The second proposed role suggests that acetylation affects the interaction of the amino-terminal tails with non-histone chromatin proteins that in turn modulate chromatin structure. A natural presumption of the first proposal is that acetylation weakens histone-DNA interactions within nucleosomes and thus affects directly the higher-order chromatin structure and ultimately leads to chromatin de-repression (Csordas, 1990). Indeed, hyperacetylation has long been considered a 'hallmark' of transcriptionally active chromatin (Turner and O'Neill, 1995) (Wolffe, 1994) and a wealth of mostly correlative evidence strongly supports the notion that acetylation potentiates transcriptional activity in chromatin (Hebbes et al., 1988). Numerous *in vivo* and *in vitro* experiments have indicated that transcription of some genes is influenced by the acetylation of particular lysine residues in specific histones (Thompson et al., 1994) (Durrin et al., 1991) and, conversely, that the transcriptional silencing of specific loci is associated with reduced nucleosomal acetylation.

Moreover, the number of methyl groups added to a single lysine or arginine residue can vary. Lysine residues can be mono-, di-, or trimethylated, and arginine residues can be monomethylated and symmetrically or asymmetrically dimethylated. Importantly, the precise methylation status (mono-, di-, or trimethylation) can influence the transcriptional status of genes (Schneider et al., 2004). These modifications could directly affect the structure of chromatin; e.g., by neutralizing the positive charge of histones. Notably different methylated states of the same amino acid residue provide additional hierarchical levels of regulation to epigenetic inheritance of chromatin domains. For example, acetylation of histones H3 an H4, and H3 methylation at Lys4 (H3 Lys4-Me), have been predominantly correlated with euchromatin and gene activity,

whereas methylation at Lys9 (H3 Lys9-Me) correlates with transcriptional silent chromatin (Litt et al., 2001). H3 Lys4 dimethylation (H3 Lys4-diMe) is associated with permissive chromatin that is either active or potentially active, and H3 Lys4 trimethylation (H3 Lys4-triMe) is linked with transcriptional activity. Conversely, H3 Lys9 di- and trimethylation (H3/K9-diMe and H3/K9-triMe) mark facultative and constitutive heterochromatin, respectively in mammals. Methylated H3/K9 can be "read out" by the Heterochromatin Protein 1 (HP1), a structural component of condensed chromatin that specifically recognizes and binds to the methylated form of H3/K9 (Bannister et al., 2001) (Lachner et al., 2001). Lack of K9 methylation in heterochromatin can affect the heterochromatin organization. Lanchner suggest that; HP1 itself interacts with the enzyme that methylates H3/K9, forming a positive feedback loop that would allow heterochromatin to spread over large chromosomal regions until further spreading is prevented by a boundary element. Furthermore, active promoters and enhancers are found to be associated with H3/K4 methylation and H3/K9 monomethylation. Moreover, active promoters can be identified by the presence of further activating marks that are linked with transcriptional elongation downstream in the transcribed region of the gene (Barski et al., 2007), in addition, at genes that are not regulated at the level of transcription initiation (polymerase II can be found at these genes) but rather elongation, H3/K4 methylation is enriched at the promoter. Thus, some activating marks are not only found on genes that are transcribed, but also on genes that are poised for transcription. Additionally, modifications considered to be repressing, such as H3/K9 dimethylation, can be found not only in heterochromatin but also on certain active genes and repressive and activated marks as H3/K27 trimethylation and H3/K4 trimethylation, respectively, can

coexist (Azuara et al., 2006) (Bernstein et al., 2006). All these findings suggest that the histone "code" is more complex than initially expected and that not single modifications, but rather the combination of modifications is an indicator for the transcriptional state.

Changes in chromatin structure that include modifications of histones may also have a role in the positioning of chromosomes. In experiments done in yeast, by Taddei and collaborators, the treatment of cells with trichostatin A, an inhibitor of histone deacetylases that increases the acetylation level of histones, results in large-scale movement of centromeric and pericentromeric chromatin to the nuclear periphery. After drug removal, these changes in localization are rapidly reversed (Taddei et al., 2001).

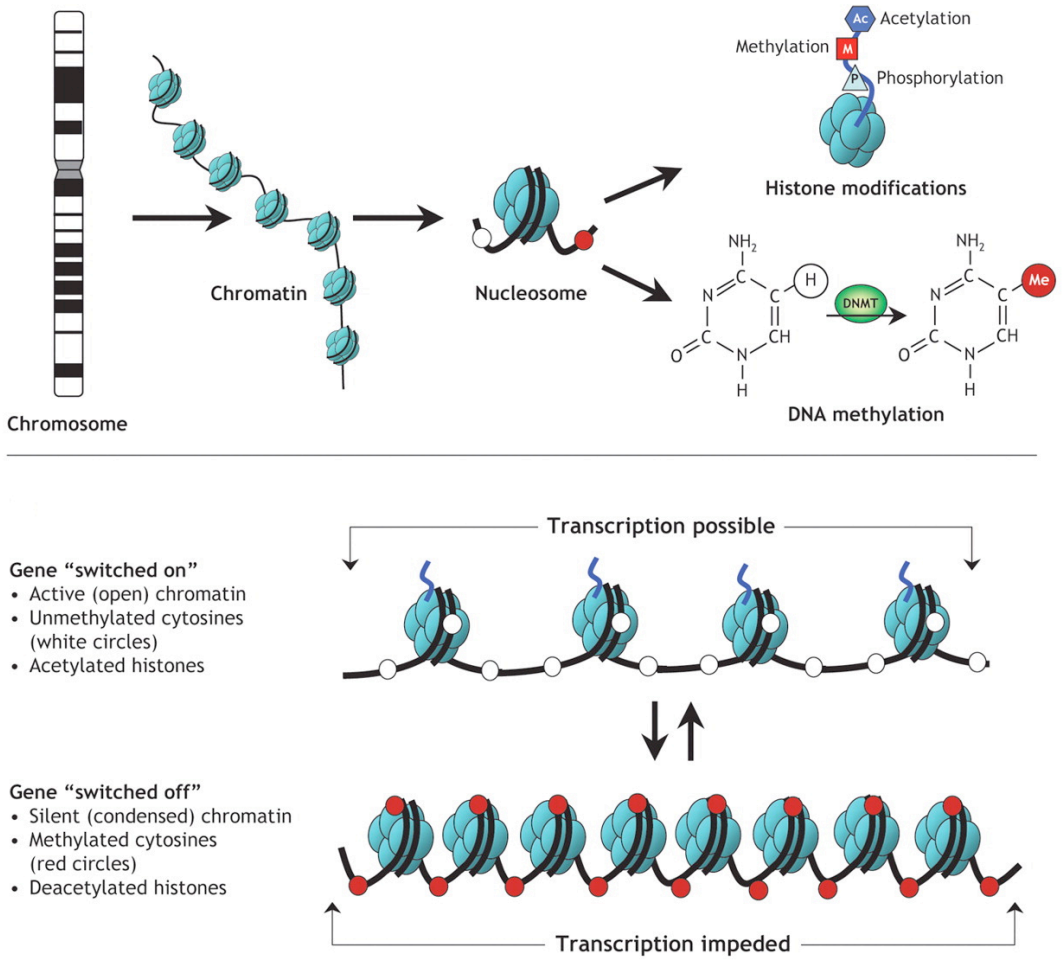


Figure 10. Epigenetic modifications.

Strands of DNA are wrapped around histone octamers, forming nucleosomes, which to be organized into chromatin, the building block of a chromosome. Reversible and site-specific histone modifications occur at multiple sites through acetylation, methylation and phosphorylation. DNA methylation occurs at 5-position of cytosine residues in a reaction catalyzed by DNA methyltransferases (DNMTs). Together, these modifications provide a unique epigenetic signature that regulates chromatin organization and gene expression. (B) Schematic of the reversible changes in chromatin organization that influence gene expression: genes are expressed (switched on) when the chromatin is open (active), and they are inactivated (switched off) when the chromatin is condensed (silent). White circles = unmethylated cytosines; red circles = methylated cytosines (http://cnx.org/content/m26565/latest/).

Chromatin remodeling complex

The second mayor class of chromatin regulators are the protein complexes that utilizes ATP hydrolysis to alter the histone-DNA contacts; because of this, they are generally referred to as chromatin remodeling complexes (Flaus and Owen-Hughes, 2004) (Saha et al., 2006). The consequences of remodeling include transient unwrapping of the end DNA from histone octamers, forming the DNA loop, or moving nucleosomes to different translational positions (sliding), all of which change the accessibility of nucleosomal DNA to transcription factors. Chromatin remodeling complex comprise an ATPase subunit along with other polypeptides that are responsible for the regulation, efficiency, and functional specificity of each complex. The ATPase subunit belongs to the Snf2 super family of proteins, which includes the SWI/Snf group and the imitation SWI (ISWI) group. The SWI2/Snf2 group includes yeast SWI/Snf (ySWI/Snf), yeast RSC, the Drosophila Brm (Brahma) complex, and the hBrm (human Brm) and BRG1 (Brm/SWI2 Related Gene-1, hBRG1 in humans) complexes (Vignali et al., 2000). Moreover, there are many other proteins that are closely related to the ATPase subunits of chromatin remodeling complexes. To remodel chromatin, remodeling complexes recognize and bind to their substrate. The SWI/Snf, NURF, and RSC complexes act as ATP-dependent motors that track along the DNA strands and pull them away from the histone octamer cores. During this shift of histone-DNA contact points, the DNA presumably becomes accessible to the transcriptional machinery (Figure 11). These complexes serve to maintain chromatin in a repressive configuration by dissociating other chromatin-associated proteins from the DNA, such as the TATA-binding protein (Redner et al., 1999).

The exact consequence of remodeling is dependent on the exact context of nucleosomes at a given promoter and can lead to either (i) activation of transcription or (ii) repression (Vignali et al., 2000). In addition to catalyzing nucleosome mobility, chromatin-remodeling factors can enhance the access of DNA-binding factors and nucleases to DNA packaged into chromatin. These activities are also consistent with the ability of remodeling factors to disrupt histone-DNA interactions in the nucleosome (Redner et al., 1999).

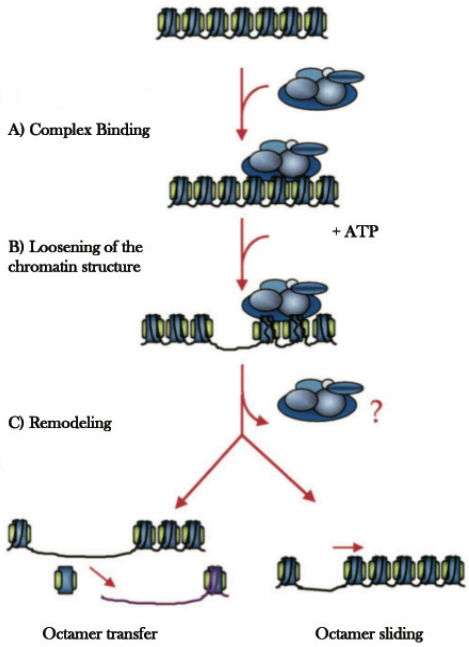


Figure 11. Chromatin remodeling complexes.

The binding of the remodeling complex to chromatin is ATP independent (A). Upon ATP addition, the conformation of nucleosomes changes as a consequence of the alteration of histone-DNA interactions (B). This disruption results in remodeling of the chromatin (C), which might occur while the complex is still bound or might persist after it is released from the chromatin (indicated by the question mark). Remodeling that occurs may result in transfer of histone octamers to different DNA segments in *trans* or in sliding of the octamers in *cis* (i.e., to a different position in the same DNA molecule) (Vignali et al., 2000).

Nuclear architecture

The nucleus is a complex organelle with an internal structure and component organization that is not fully characterized. Understanding the molecular details of nuclei organization including: how the nuclear compartments are formed and maintained, how the synthesis, processing, assembly and transport of molecules are coordinated and regulated and how the chromosomes are arranged remains one key question to be address in understanding the contribution of spatial positioning to genome function. Within the nucleus, protein and RNA-protein complexes can move to and from chromatin through the interchromatic nucleoplasmic space by a diffusion mechanism until they encounter an appropriate binding site. In addition, many nuclear factors are able to form or interact with different types of subnuclear domains, which are either attached to chromatin or located in the interchromatic space.

Chromatin

Chromatin is the name for genomic DNA packaged into highly ordered structures and is composed of structural subunits called nucleosomes. Nucleosomes consist of 146 base pair of DNA, wrapped around an octamer of histones represented by two copies each of histone (H) 2A, H2B, H3 and H4. All the four core histones consist of a predominantly α-helical, evolutionary conserved globular domain, necessary for octamer assembly and DNA binding, and unstructured flexible amino terminal tail rich in lysine residues dispensable for nucleosome assembly but heavily post-translationally modified (Marzio and Giacca, 1999).

Although histones are proteins very conserved between them, the nuclesome is a very dynamic structure. The histone amino termini extend from the nucleosome, where they serve as targets for characteristic covalent posttranslational modifications, including acetylation, phosphorylation, methylation and ubiquitination. The enormous combinatorial potential of these modifications can be read out by proteins that bind to specific modifications. This has provided the basis for the so-called histone "code" hypothesis (Strahl and Allis, 2000) (Rice and Allis, 2001). Although it is still unproven whether these modifications form a true "code," it is now well established that they are involved in the regulation of gene expression (Turner 2007).

Euchromatin

Cytogenetic analysis of chromatin identified the major chromatin domains defined as euchromatin and heterochromatin. Euchromatin consist in a lightly packed form of chromatin, rich in gene concentration, and often under active transcription. The unfolded structure of euchromatin allows gene regulatory proteins and RNA polymerase complexes to bind to the DNA sequence, which can then initiate the transcription process (Misteli, 2005). Many 'euchromatic' regions of the genome probably contain a mixture of transcribed or potentially active loci interspersed within transiently silenced genes and areas of established facultative heterochromatin. Actively transcribed regions are often described as having an 'open' chromatin structure, comprising nucleosomes that are loosely or irregularly packed compared with the tightly packed (Arney and Fisher, 2004). The link between open chromatin conformation and gene activity was described as long ago as 1976 (Weintraub and Groudine, 1976). We now know that chromatin-remodelling factors, which can slide or

reposition nucleosomes on DNA templates, have an important role in establishing and maintaining euchromatic domains (Havas et al., 2001). There is also evidence for non-genic transcription playing a role in opening up chromatin for gene expression, presumably with the aid of transcription-linked remodelling complexes (reviewed by Cook, 2003) (Cook, 2003).

Heterochromatin

In contrast with euchromatin, heterochromatin is often enriched in inactive and silenced genome regions, which remains deeply stained and highly condensed during cell division. Heterochromatin plays several functions, from gene regulation to the protection of the integrity of chromosomes; all of these roles can be attributed to the dense packing of DNA, which makes it less accessible to factors binding DNA or their associated co-factors. However, the extend of heterochromatin of specific regions may differ in different individuals or tissue types and may be determined by a complex process involving factors responsible for chromatin remodeling (Thiagalingam et al., 2003). Heterochromatin can be divided in two kinds, constitutive and facultative heterochromatin. The first one includes regions which often comprise repetitive DNA (such as satellite sequences surrounding centromeres) and are generally thought to be 'gene poor'. Such areas can exert a strong repressive effect on gene transcription, such as positioning of a gene within or close to satellite repeats (Schotta et al., 2003). The hallmarks of constitutive heterochromatin (Lachner et al., 2001) include trimethylation at Lys9 of histone H3, a small number of methylation at H3 Lys4 and trimethylation of H4 Lys20 (Kourmouli et al., 2004) (Schotta et al., 2004). Heterochromatin-associated proteins such as HP1α and HP1β, in human cells HP1α is significantly enriched at

centromeres (Gilbert et al., 2003). HP1 interactions are generally thought to contribute to the stable formation of condensed chromatin structures. The termed facultative heterochromatin describes a previously permissive chromatin environment that is subject to transcriptional silencing. Such an environment is often confused with constitutive heterochromatin, but the two are distinct. The precise nature of the chromatin structures found within facultative heterochromatin has not been fully defined, and different genes might employ a variety of silencing mechanisms. DNA methylation and characteristic histone modifications typify the repressive environment of facultative heterochromatin. DNA methylation is associated with silenced genes, particularly at promoters and some other regulatory elements (Attwood et al., 2002). Histone modifications that occur within regions of facultative heterochromatin include dimethylation at H3 Lys9 and hypoacetylation of lysine residues within H3 and H4, commonly around promoter regions. Methylation of H3 Lys27 has also been proposed to be a feature of facultative heterochromatin and is primarily associated with transcriptional silencing by Polycomb group proteins (PcG) as well as with X inactivation (Arney and Fisher, 2004).

While the two types of chromatin can easily be defined based on their morphology, it is becoming increasingly clear that the correlation of large-scale chromatin morphology and function is an oversimplification. An example was demonstrated by Bickmore and colleagues (Gilbert et al., 2004), which found that, open chromatin fibers correlate with regions of highest gene density, but not with gene expression since inactive genes can be in domains of open chromatin, and active genes in regions of low gene density can be embedded in compact chromatin fibers.

Chromosome territories

The largest unit of organization of the eukaryotic genome is the chromosome. Today is well accepted that organization of chromatin is not random and whole chromosomes reside in specific locations within the nucleus called chromosome territories (CT) (Cremer and Cremer, 2006), which have preferred positions with respect to the centre or periphery of the nucleus and with respect to the other (Figure 13).

Evidence for the existence of CTs in nuclei was first originated in mammalian cells by Cremer's brothers (Cremer et al., 1982) whose microirradiated Chinese Hamster nuclei with a UV laser, and visualized the sites of damage at the following metaphase. Instead of obtaining a scattered signal across several chromosomes, Cremer's and collaborators observed discrete labelling in parts of only a few chromosomes, readily indicating that chromatin must be highly compartmentalized within the nucleus (Figure 12). This idea was later confirmed by fluorescence in situ hybridization (FISH) using whole chromosome paint probes (Cremer et al., 1988) (Lichter et al., 1988). This has since become a standard tool that has enabled extensive study of chromosome organization.

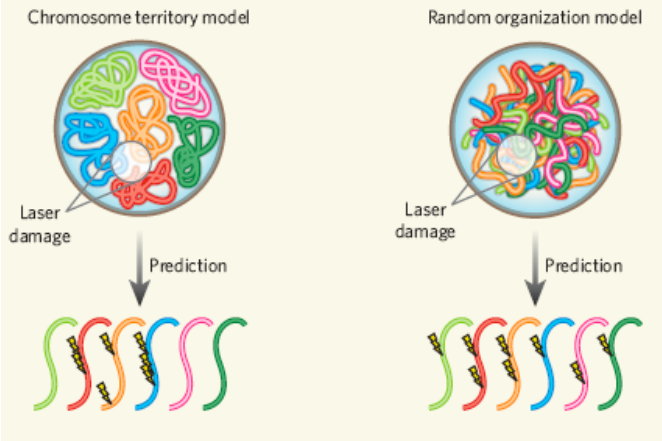


Figure 12. Chromosome territories

Thomas and Christoph Cremer used a microlaser to induce local genome damage, and predicted that inflicting DNA damage within a small volume of the nucleus would yield different results depending on how chromosomes were arranged. If chromosomes occupied distinct territories (a, left panel), localized damage would affect only a small subset of chromosomes, whereas if the chromatin fibres of each chromosome were randomly distributed throughout the nucleus (a, right panel), many chromosomes would be damaged (Meaburn and Misteli, 2007).

A correlation has been observed between CT location and human chromosome size (Bolzer et al., 2005), in which smaller chromosomes are generally situated towards the interior and larger chromosomes towards the periphery of the nucleus (Sun et al., 2000). In addition, there is strong evidence that CTs with similar DNA content, but with very different gene densities, occupy distinct nuclear positions. A striking example comes from the distribution of human chromosome 18 and 19. Although both chromosomes have a similar DNA content (85 and 67 Mb, respectively), the gene-poor chromosome 18 territories were typically found at the nuclear periphery, whereas the gene-rich chromosome 19 territories were located in the nuclear interior (Croft et al., 1999). However gene density or chromosome size alone clearly cannot explain the

position of a chromosome given that the position differs between cell types and tissues where these properties are unchanged (Parada et al., 2004).

Genome organization

Radial positioning is not limited to entire chromosomes, but has also been reported for single genes, with some genes localizing preferentially towards the center and other localizing towards the periphery of the nucleus (Roix et al., 2003). A simple way to assess the position of a genome region within the nucleus is by determining its distances from the nuclear periphery. A general correlation between transcriptional silencing and localization toward the nuclear edge has long been suggested based on the observation that early-replicating and presumably transcriptional active R bands are generally found toward the centre of the nucleus, whereas late replicating, inactive G bands are often located toward the periphery (Gilbert et al., 2004) (Ferreira et al., 1997) (Sadoni et al., 1999). In some cases, position seems to ensure maximal expression of a gene. In others, it ensures a heritable state of repression or correlates with a developmentally determined program of tissue-specific gene expression. However, recent studies have been shown that active genes can reside and be transcribed within "heterochromatin" regions. A more detailed analysis of nuclear organization reveals specialized subcompartments for gene expression and silencing both at the nuclear periphery and in the interior (Sexton et al., 2007). Furthermore, some recent studies have been shown that radial position within the nucleus can influence the expression of some, but not all genes (Finlan et al., 2008).

Genes can move away from the CT (Figure 13). The first study to describe movements of individual loci in human cells was done by Jonathan R. Chubb and collaborators in

2002 (Chubb et al., 2002). They showed that different loci display different levels of motion. Thus, genomic regions associated with the nucleolus and nuclear periphery display significantly more constrained movements than other. The results obtained in this studies, strongly support the hypothesis previously suggested by Abney in 1977 (Abney et al., 1997), that regulating the density of attachments sites, chromatin dynamics can be strongly controlled, in consequence, nuclear periphery and nucleolus are probable nuclear substrates for the attachment of chromatin, suggesting a role for these compartments in maintaining the three dimensional organization of chromatin in the nucleus.

The gene relocation occurs either in domains of constitutively high gene expression (Mahy et al., 2002) or, in some situation, when gene expression is induced (Volpi et al., 2000). For example, the IgH locus is preferentially associated with the nuclear periphery in B cell progenitors where it is silent, but it moves toward the interior when it becomes active in B cell precursors (Ragoczy et al., 2006) (Kosak et al., 2002). Similarly, the CD4 locus repositions from the periphery to the nuclear interior during T cell differentiation, and Hox1b and Hox9 become internalized roughly concomitantly with their transcription activation (Chambeyron and Bickmore, 2004; Kim et al., 2004). In addition, in many cases, no repositioning occurs upon a change of gene activity (Dieudonne et al., 2009) (Zink et al., 2004). A consequence of this relocation is that the 'looping out' of *a* given locus allow the interaction with other chromosomes. This finding, together with the observation of extensive intermingling of DNA from different chromosomes at the boundary of, or just outside of, chromosome territories suggests that chromosome territories might be more dynamic than previously thought. On the other hand, the radial position of a gene is generally not directly related

to its activity as indicated by the fact that in most cells the two alleles are positioned differently yet their functional properties appear to be similar (Roix et al., 2003).

In addition to the ordered organization of chromosomes, protein components of the cell nucleus are able to form or interact with different types of subnuclear domains, which are either attached to chromatin or located in the interchromatic space, often forming subnuclear compartments including nucleoli, splicing factor compartments, or PML bodies.

These nonrandom organization patterns of the genetic material and many of the components involved in gene expressing point towards a role for spatial organization in control of gene expression. Some evidence shows that the major aspects of nuclear organization change as the physiological state of cells changes, for example, during stem cell differentiation and in development (Wade and Kikyo, 2002); this fact make a link between spatial genome organization and genome expression. Despite this evident link, it is not still clear whether changes in nuclear architecture lead to changes in gene expression patterns or vice versa. Instead, is clear that describing the patterns of spatial genome organization and understanding the constraints and rules that lead to these organization patterns are a critical step in understanding fundamental biological processes such as, differentiation, genome reprogramming and gene regulation.

Transcription factories

Our understanding of gene regulation has moved from an initial notion of a one-dimensional array of regulatory elements to an appreciation that contiguous sequences can have a deep effect in gene expression, introducing in this way a second dimension

onto the scene. The picture that is emerging is that nuclear processes may rely of a constant flow of molecules between dynamics compartments created by relatively immobile binding or assembly sites. Transcriptional active genes have been proposed to associate with discrete nuclear sites of nascent RNA production and concentrated transcriptional components such as RNA polymerase, these compartments are termed transcription factories. As the number of pol II transcription sites (estimated to be from 500-10,000/cell) has always been estimated to be less than the number of active polymerases (approximately 20,000 to 100,000 per HeLa nucleus), one would expect that each site might be associated with more than one transcription unit, estimated at ~30 such units per site in HeLa cells (Iborra et al., 1996) (Jackson et al., 1998). Therefore, transcription factories are believed to be dynamic structures containing spatially contiguous DNA templates, RNA polymerase II, newly synthesized mRNA and associated factors that would probably transcribe multiple genes at the same time.

Analogously, ribosomal RNA genes, which are transcribed by RNA polymerase I within the nucleolus represent large specialized transcription centers. The compartmentalization of transcription has the obvious advantage of concentrating the needed factors to ensure efficient interactions among the components of the transcription machinery (Cook, 1999). One possibility is that different transcription factories contain distinct sets of transcription components and thus create distinct transcription environments.

Transcriptionally active genes have been proposed to loop out from their chromosome territory and associate with transcription factories. Work of Osborne et al. showed that widely separated genes are recruited to shared RNAP II sites if transcriptionally active

or reposition away when inactive (Osborne et al., 2004). They have postulated that genes with transcriptional potential often move out to existing transcription factories when activated rather than assemble their own transcription site *de novo*. Transcription units might be associated through promoters and/or enhancers, with loops forming a 'cloud' around the factory. Transcription factories are highly dynamic structures and involved transient interaction of factors with chromatin. Many transcription factors undergo rapid exchange between chromatin and nucleoplasm and for instance the glucocorticoid- and estrogen-receptor transcriptional co-activators bind to their specific response elements in the promoter with residence times of only a few seconds (McNally et al., 2000). Figure 13 shows a general picture of the nuclear organization at the chromatin level.

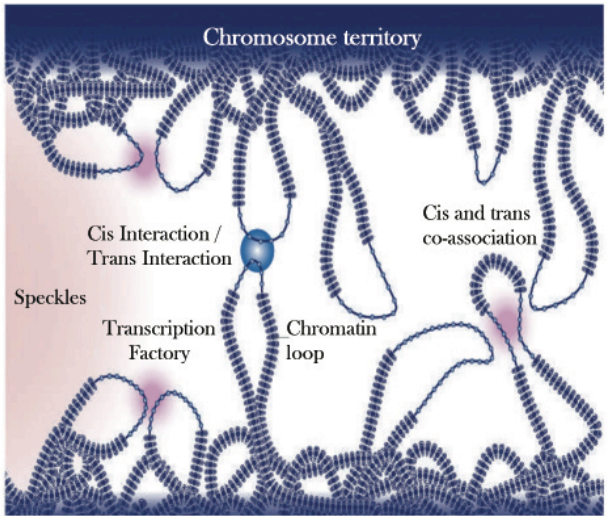


Figure 13. Nuclear Organization.

Active genes on decondensed chromatin loops that extend outside chromosome territories can colocalize both in cis and in trans at sites in the nucleus with local concentrations of Pol II (namely transcription factories; dark pink) and adjacent to splicing-factor enriched speckles (pale pink). Interactions can also occur between regulatory elements and/or gene loci and lead to coregulation in trans (blue circle) (Fraser and Bickmore, 2007).

New Techniques: Imaging and Chromosome Conformation Capture

Analysis of chromosome conformation is complicated by technical limitations. In the last years methodological advances have provided new opportunities to explore the overall spatial organization of chromosomes and to investigate the physical properties at high resolution. There are two different ways, which are complementary, to study spatial conformation of chromosomes: by characterization of the surroundings of a gives locus, using Chromosome conformation capture (3C) or by microscope techniques. Advances in the specific fluorescent labeling of chromatin in fixed and living human cells in combination with three-dimensional (3D) and 4D (space plus time) fluorescence microscopy and image analysis have opened the way for detailed studies of the dynamic, higher-order architecture of chromatin in the human cell nucleus and its potential role in gene regulation. Several features of this architecture are now well established: 1. Chromosomes occupy distinct territories in the cell nucleus with preferred nuclear locations, although there is no evidence of a rigid suprachromosomal order. 2. Chromosome territories in turn contain distinct chromosome arm domains and smaller chromatin foci or domains with diameters of some 300 to 800 nm and a DNA content in the order of 1 Mbp. 3. Gene-dense, early replicating and gene-poor, middle-to-late-replicating chromatin domains exhibit different higher-order nuclear patterns that persist through all stages of interphase (Cremer et al., 2000). Additionally, the develop of a high-throughput methodology as 3C and 4C (Circular Chromosome Conformation Capture) has allowed the analysis of genome-wide interaction frequencies providing information about general nuclear organization as well as physical properties and conformation of chromosomes.

Fluorescence imaging techniques

Microscopic observation allowed the discovery of the cell as a structural unit of tissues and organisms and provided a tool to uncover the subcellular structure and organization. Additionally, the use of fluorophore tags revolutionized the way microscopy is used in biology. It is now possible to visualize the dynamics of proteins and organelles in single living cells and give kinetic measures to cellular processes. The most popular fluorescent label used is the green fluorescense protein (GFP) from the jellyfish *Aequorea victoria*. GFP and its variants allow the monitoring in time and space of processes in living cells and organisms like gene expression, protein localization and dynamics, protein-protein interactions, cell division, chromosome replication and organization, intracellular transport pathways, organelle inheritance and biogenesis. There is now an entire toolbox of fluorescent proteins with different chromatic and structural properties (Figure 14). The technical revolution coming from the use of GFP resulted that Martin Chalfie, Osamu Shimomura and Roger Y. Tsien were awarded the 2008 Nobel Prize in Chemistry *"for the discovery and development of the green fluorescent protein, GFP".*

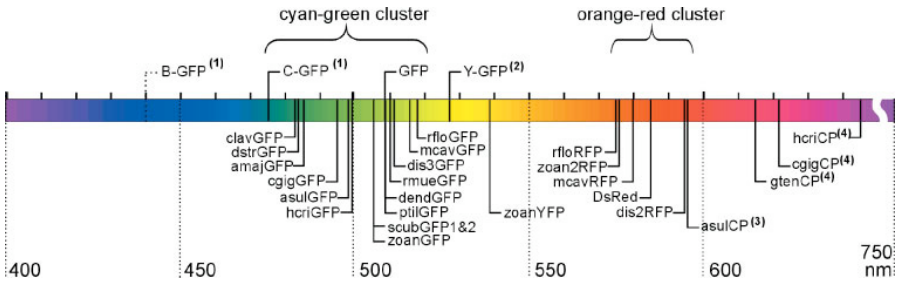


Figure 14. Spectral properties of variants of the GFP family (Matz et al., 2002).

Combinations of GFP variants can be visualized in the same cell and specific pairs can be measured *in vivo.* This can be achieved by fluorescence resonance energy transfer (FRET) (reviewed in (Phair and Misteli, 2001)). By this technique the interaction between two proteins or DNA and protein can be studied. FRET uses two fluorophores which label two molecules under study, a donor and an acceptor. Excitation of the donor by an energy source triggers an energy transfer to the acceptor if they are within a 1-10nm proximity to each other. The acceptor in turn emits light at its given wavelength. On the other hand, imaging techniques, such as fluorescence recovery after photobleaching (FRAP), fluorescence loss in photobleaching (FLIP), inverse-FRAP (i-FRAP) and photoactivation can provide insights into the kinetics and trafficking of subsets of the studied proteins and their accessibility to different cellular compartments.

MS2-based tagging of RNA

Labeling either by fusion with a fluorescent protein or with a fluorescent-dye moiety allows the dynamic study of a protein of interest inside the cell in vivo. In the case of nucleic acids, intra-molecular tagging approaches are required because the target DNA or RNA are typically not accessible for tagging in vivo.

Labeling of chromosomes in vivo can be achieved by fluorescent tagging of chromosome binding factors. Bacterial operator-repressor systems, such as lac operator (lacO) (Robinett et al., 1996) (Janicki et al., 2004) and the bacterial tetracycline operator/repressor (*tet*O/TetR) (Tsukamoto et al., 2000) (Matzke et al., 2005) combined with fluorescent proteins offer a unique opportunity to visualize

fluorescence-tagged loci in nuclei of living, unfixed cells. The operator repeats are integrated into the genome as a transgene array, which then specifically binds the respective nuclear-localized repressor protein that is fused with a fluorescent protein such as GFP. The tagged loci appear as bright fluorescent dots when viewed with appropriate filters under a fluorescence microscope. Using a fluorescence microscope equipped with a motorized z axis and image-processing software, it is possible to make optical sections through nuclei and reconstruct them in three dimensions to determine spatial relationships among fluorescence tagged loci.

The mRNA can be conjugated to fluorescent dyes and directly microinjected to the cells, but this is a rather invasive procedure (Ainger et al., 1993). Several other techniques have been developed which include the use of fluorescently labeled probes for in vivo hybridization, the use of molecular "beacons" which are active only when bound to target RNA or the use of fluorescently tagged RNA-binding proteins (Politz et al., 1998). However, none of those approaches provide information about the behavior of a single RNA molecule in real-time in living cells. An innovative approach was developed inserting a series of RNA aptamers– stem-loops into the transcript of interest by Singer and collaborators (Bertrand et al., 1998). This RNA is then detected by high affinity specific interaction between RNA stem-loops and fluorescently labeled bacteriophage coat protein MS2. Therefore, the RNA detection system is comprised of two elements, a tag protein/peptide fused to the MS2 and a reporter RNA containing MS2- binding sites repeated in tandem. MS2 binds specifically to an RNA stem–loop structure consisting of 19 nucleotides (Figure 15).

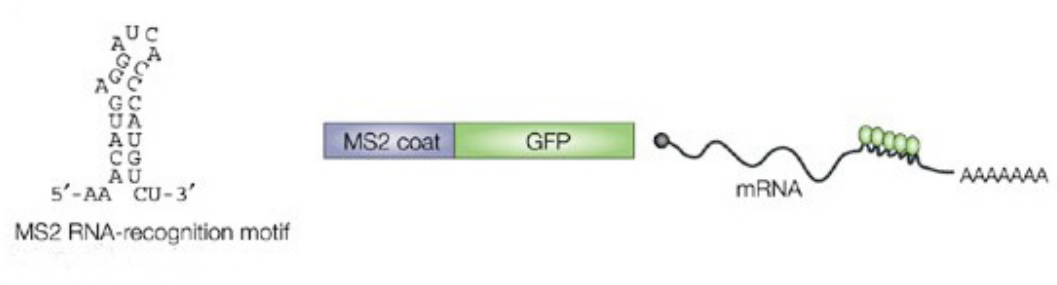


Figure 15. RNA detection MS2-based system.

RNA molecules are detected using a fusion protein that comprises a fluorescent tag, such as green fluorescent protein (GFP) and the MS2 bacteriophage coat protein. MS2 has an extremely high affinity for a 19 nt RNA-recognition motif that is derived from the phage genome (Shav-Tal et al., 2004).

This method allows the detecting of specific RNA in order to study RNA dynamics (i.e. biogenesis, processing, export) in vivo. This technique can be used in broad range of applications such as visualization and kinetics of RNA in single cells by tagging MS2 with a fluorescent mark; or affinity purification and further proteomic studies of endogenously assembled RNA-protein complexes in order to identify new factors involved in RNA regulation. The MS2-based system has been effectively used to tracking specific cytoplasmic and nuclear mRNA in yeast, Dictyostelium, plants, flies and mammalian cells. Using the MS2-system in living mammalian cells, it was possible to demonstrate that cytoplasmic mRNAs exhibit directed, corralled, diffusive and static movements (Fusco et al., 2003). Furthermore, this approach has also been used for the detection and kinetic studies of nuclear events such as transcription sites (Janicki et al., 2004). Visualization and trafficking of single mRNP in the nucleus revealed that diffusion is the primary mechanism by which these molecules translocate from transcription site to nuclear periphery (Shav-Tal et al., 2004). MS2-based approaches allowed the study of the dynamics of RNAPII in vivo (Darzacq et al., 2007).

Photobleaching and photoactivation of fluorescent MS2 proteins used to label nascent messenger RNAs provided sensitive elongation measurements. It was shown that RNAPII elongated at 4.3 kb/min on a synthetic gene. These data were consistent with elongation rates measured on integrated HIV-1 genes (Boireau et al., 2007).

In conclusion, RNA detection by the MS2 system coupled with live cell imaging technologies is a powerful approach which allows the analysis of mRNA biogenesis in single cell in real time. Alternatively, MS2 system coupled with high affinity chromatography and mass spectrometry-based techniques permits analysis of proteome with the identification of new components of RNA-protein complexes in vivo.

3D FISH

Fluorescence in situ hybridization (FISH) with specific DNA probes has become a widely used technique for chromosome analysis and for studies of the chromosomal location of specific DNA segments in metaphase preparations as well as in interphase nuclei. FISH of DNA probes to three-dimensionally preserved cells, is termed 3D-FISH. This technique allows three-dimensional visualization of specific DNA and RNA targets within the nucleus at all stages of the cell cycle. It provides information about the arrangement of chromosome territories and the organization of subchromosomal domains, about the pattern of chromatin density within a chromosome territory, about positions of individual genes and RNA transcripts read from them. Accumulation of such data is necessary for understanding relationships between the spatial organization of the genome and its functioning in the interphase nucleus.

3D-FISH in combination with 3D-microscopy and image reconstruction has become a major tool for studying the higher order chromatin organization in the cell nucleus (Solovei et al., 2002). Comparing with FISH on metaphase chromosomes and conventional interphase cytogenetics, 3D –FISH requires special demands with regard to probe quality, fixation conditions and pretreatment steps of cells in order to optimize the preservation of nuclear architecture and at the same time an efficient probe accessibility. As a result, differently labeled nuclear targets can be delineated simultaneously and their spatial interrelationships can be analyzed on the level of individual nuclei.

Chromosome Conformation Capture: 3C

Important chromosomal activities have been linked with spatial conformation of chromosomes and their structural properties. Local properties of the chromatin fiber influence important process as: gene expression, origin firing, and DNA repair (Woodcock and Dimitrov, 2001). Higher order structural features—such as formation of the 30-nm fiber, chromatin loops, and interchromosomal connections—are important for chromosome morphogenesis and also have roles in gene expression and recombination. Transcription and timing of replication have been related to overall spatial nuclear disposition of different regions and their relationships to the nuclear envelope (Andrulis et al., 1998) (Cimbora and Groudine, 2001). At each of these levels, chromosome organization is highly dynamic, varying both during the cell cycle and among different cell types.

Since few years our understanding of genomic organization in the nuclear space was based mostly on microscope studies that often used FISH to visualize selected parts of the genome. However FISH can analyze only a limited number of DNA loci simultaneously. High-resolution techniques have allowed new insights into the nuclear architecture and its relationship to gene expression. One of the most used approaches is Chromosome Conformation Capture (3C). 3C technique provides a powerful tool to study interchromosomal interactions with high accuracy in molecular terms (Dekker et al., 2002). In 2002 Dekker and collaborators, described this innovative assay that allows the analysis of the overall spatial organization of chromosomes and to investigate their physical properties at high resolution. The technique consist on isolating intact nuclei that are then subjected to formaldehyde fixation, which cross-links proteins to other proteins and to DNA. The overall result is cross-linking of physically touching segments throughout the genome via contacts between their DNA-bound proteins. Cross-linked chromatin is then solubilized and digested with an appropriate restriction enzyme. This is followed by intramolecular ligation of cross-linked fragments (Figure 16). The resulting "3C templates" thus contains a large collection of ligation products. Each ligation products reflects an interaction between two genomic loci and can be detected by quantitative PCR using specific primers: the abundance of each ligation product is a measurement of the frequency with which the two loci interact. Thus, the 3C method can be used to capture and quantify physical interactions between genes and distant elements, both is *cis* and in *trans*.

Circular Chromosome Conformation Capture: 4C

Even if the 3C method provides the opportunity to study chromosomal folding in the nucleus, this method has a limitation, the prior knowledge of the bait and interacting sequence, rendering it less useful for genome wide studies. This limitation has been overcome by exploiting a circular intermediate in a variant of the 3C method, termed 4C (Lomvardas et al., 2006) (Zhao et al., 2006). 4C approach involves a circularization step that enables high-throughput screening of physical interactions between chromosomes without a preconceived idea of one of the interacting partners. An outline of the 4C procedure is given in Figure 26. In short, 4C procedure follows the same steps as in 3C, except additional processing is needed before quantification of the fragments of interest. Therefore, after reversal cross-linked of DNA, the restriction fragments are subjected to another round of restriction digest, this time with a frequent cutter that will result in smaller fragments with restriction ends that differ from the central restriction site (ligation junction). Self-circularization of the DNA fragments is more favored now that they are not bound to other proteins or fragments. Intramolecular ligation occurs to induce the formation of the circular fragments. The pool of circular fragments becomes the 4C library. Primers are designed against the outer restriction sites of the "bait" sequence, which result in the amplification of the small unknown captured fragment. Large-scale sequencing can be used to sequence the 4C library. Custom microarrays can also be made using probes designed against the adjacent upstream and downstream regions of all genomic sites of the restriction enzyme used (Simonis et al., 2006).

3C-carbon copy (5C) and 4C on ChIP

In addition, some other variants of 3C have emerged: 3C-carbon copy (5C) is used to measure networks of millions of chromatin interactions in parallel. As in 3C, cells are treated with formaldehyde to cross-link chromatin interactions. The chromatin is solubilized, digested with a restriction enzyme and ligated at low DNA concentration to promote intra-molecular ligation of cross-linked DNA fragments. Ligation products are subsequently purified to generate a 3C library. The 5C technology then employs highly multiplexed ligation-mediated amplification (LMA) to detect and amplify 3C ligation junctions. The resulting 5C library of ligated primers is analyzed using either microarray detection or ultra-high-throughput DNA sequencing (Dostie and Dekker, 2007). In the same manner, ChIP-loop assays (Simonis et al., 2007), is another techique generated from 3C analysis, the difference is this case from the previous techniques in that the interaction formed between two chromosomal regions is mediated by a bound protein. Like in the 5C methodology, a single DNA site is often considered to interact with multiple other sites. After the cross-linking and digestion, ChIP is performed to pull down the protein bound to the site of interest. Normal 3C procedures are conducted after this step (Simonis et al., 2007).

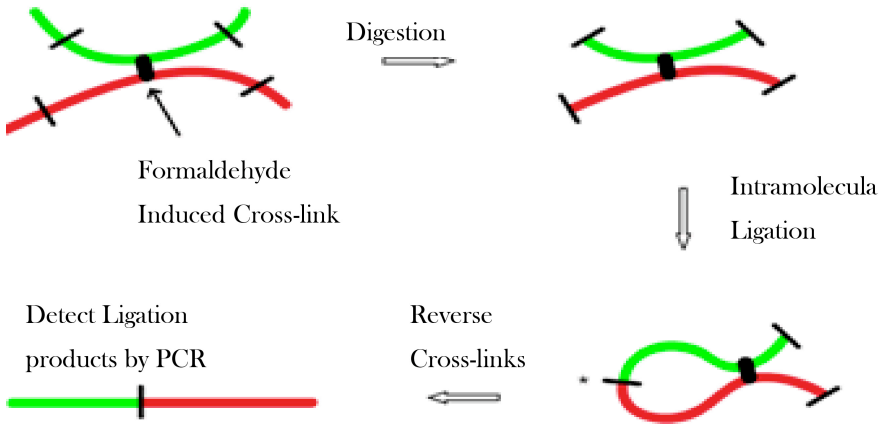


Figure 16. Schematic diagram of 3C assay.

Chromatin is cross-linked, digested with a restriction enzyme and then ligated to obtain a 3C and 4C templates, which contain a collection of ligation products, each of which reflects a physical interaction between two restriction fragments. The blue and red lines represent the bait (known sequence) and the unknown sequence respectively (in the case of 4C). In 4C assays the strategic positioning of the primers allow the amplification of a circular molecule. In 3C approach both sequences are known (red an blue line), the PCR occur with primer positioning toward the interior as a normal PCR reaction.

Results

Sub-nuclear positioning of HIV-1 provirus in latent cells

The position of a specific gene within the nucleus is correlated with its transcriptional status and specialized sub-compartments for gene expression or repression have been described (Sexton et al., 2007). Silent gene loci are often found in close proximity to centromeric heterochromatin, and centromeres are known to be preferentially associated with the nuclear periphery (Brown et al., 1997) (Brown et al., 1999) (Brown et al., 1999) (Kim et al., 2004) (Merkenschlager et al., 2004). The nuclear periphery has been functionally associated with transcriptional silencing in yeast (Andrulis et al., 1998) and Drosophila melanogaster (Marshall et al., 1996) (Pickersgill et al., 2006), although this view has been challenged by the demonstration of regions of active transcription associated with nuclear pores (Casolari et al., 2004) (Taddei et al., 2006). In addition, in mammals, the nuclear periphery has been associated with transcription repression, particularly through associations with the nuclear lamina (Guelen et al., 2008). Tethering of transgenes to the periphery of the nucleus causes transcriptional silencing that can be reverted by repositioning towards the interior of the nucleus (Dietzel et al., 2004) (Finlan et al., 2008) (Reddy et al., 2008). Several endogenous genes have also been shown to behave similarly (Kosak et al., 2002) (Zink et al., 2004) (Chuang et al., 2006; Williams et al., 2006a), although the picture might be more complex with genes being unaffected by their proximity to the periphery (Nielsen et al., 2002) (Zhou et al., 2002) (Hewitt et al., 2004) (Finlan et al., 2008) (Kumaran and Spector, 2008) (Reddy et al., 2008). Changes in gene expression after cellular differentiation, or after various

external stimuli, are often associated with changes in the subnuclear positioning of genes and regulatory sequences that may establish long-range interactions (Spilianakis et al., 2005) (Lomvardas et al., 2006). Following these facts, we decided to exploit a peculiarity of retroviruses that are able to integrate into the host's chromatin. In this respect a retroviral construct carrying a reporter gene can be seen as an ectopic transcription unit that transcribes from its integration site. Thus, by 3D FISH method (a method previously shown to preserve the three-dimensional structure of the nucleus (Solovei et al., 2009)), we investigated the localization of HIV-1 with respect to the nuclear periphery in seven different cell lines which are well-characterized models of HIV-1 latency. The cell lines used in this study are described below.

Jurkat clonal cell lines

Jurkat clonal cell llines generated by Jordan et al. are a convenient model to study HIV-1 post-integrative latency (Jordan et al., 2003). These cell clones were obtained by transduction of the human lymphocytic T cell line Jurkat with an HIV-1 vector containing the viral protein Tat and the green fluorescent protein (GFP) open reading frames under the control of the viral 5' LTR (Jordan et al., 2003). The vector used was derived from the prototype pHR' series where flag-tagged Tat101 (corresponding to the two-exon form of the HIV-1 Tat gene with a C-terminal flag tag) and GFP were under the control of the HIV-1 5'-LTR by using an internal ribosome entry site (LTR-Tat-IRES-EGFP)(Figure 17) (Jordan et al., 2003).

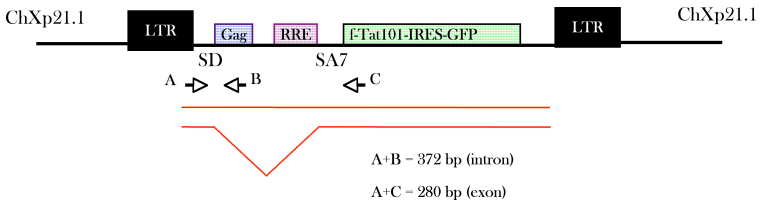


Figure 17. J-lat A1 cellular model

Diagram of the HIV-1 construct integrated in position ChXp21.1 of J-Lat A1. The 5' LTR is followed by, the major splice donor site (SD) and a portion of the gag gene fused to the Rev responsive element (RRE). The Tat acceptor site SA7 precedes a cassette containing the Tat101 gene fused to a flag tag (f-Tat101), an internal ribosome entry site (IRES) and the GFP reporter. Below the pre mRNA and spliced RNA are shown together with the dimensions of the fragment amplified by RT PCR with the indicated primers.

Cells were selected in order to obtain resting lymphocytic cells inducible under different stimuli. Infected Jurkat cells were sorted using differential fluorescence-activity based on GFP expression. Cells that were negative for GFP were isolated and then stimulated with phorbol esters (TPA) to isolate GFP positive cells, which subsequently, were cloned and further characterized (Jordan et al., 2003). In this work, we used four different Jurkat cell line clones derived as described above: J-lat A1, J-lat A2, J-lat H2 and J-lat 8.4. Table 1 shows the main characteristics of all cell lines under study.

HOS clonal cell lines

HOS clonal cell lines are an excellent model to study RNA dynamics in vivo, allowing the visualization of RNA in single living cells (De Marco et al., 2008).

The HIV-Intro MS2-ECFPskl-IRES-TK lentiviral vector (for simplicity: HIV-Intro), carried by HOS cell clones, has been specifically engineered to contain all the elements required for RNA production and, at the same time, to allow the visualization of viral RNA in single living cells. The construct contains: the 5' LTR, the major splice donor (SD1), the packaging signal Ψ, the RRE, the splice acceptor A7 and the 3' LTR that drives 3'-end formation, a reporter of gene expression (ECFP) fused to the peroxisome localization signal Ser-Lys-Leu (skl) and the selectable marker thymidine kinase (TK) of herpes simplex type 1 (Figure 18, Table 1) (De Marco et al., 2008).

The construct carries also an array of 24 binding sites for the MS2 phage coat protein within the intron, necessary for the tracking of viral RNA in vivo upon transfection of Tat and fluorescent tagged-MS2nls.

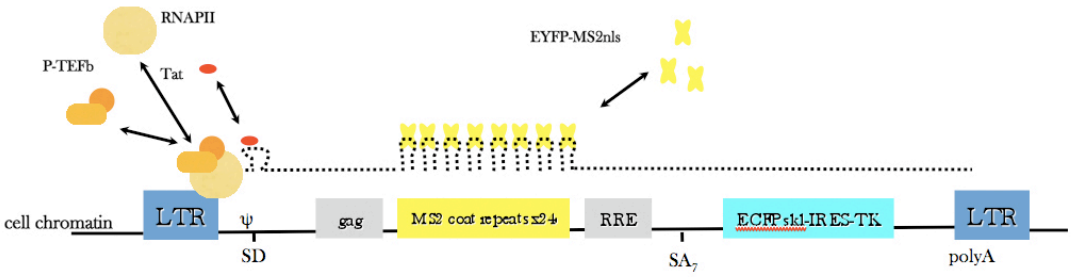


Figure 18. HIV24xMS2introECFPskl-IRES-TK lentiviral vector

Diagram of the HIV24xMS2 construct. The 5' LTR is followed by the major splice donor site (SD) and a portion of the gag gene fused to an array of 24 repeats of the MS2 phage coat protein and the Rev responsive element (RRE). The Tat acceptor site SA7 precedes a cassette containing an internal ribosome entry site (IRES) and a reporter of gene expression (ECFP) fused to the peroxisome localization signal Ser-Lys-Leu (skl) of herpes simplex type 1. The various factors that are recruited at the site of RNA biogenesis are also indicated (De Marco et al., 2008).

HOS_A4 and HOS_B3, used in this study, were generated by transduction of Osteosarcoma HOS_143b cells (which are negative for thymidine kinase activity (TK-) with the HIV-Intro vector pseudotyped with the vesicular stomatitis virus G envelope protein (VSV-G). To select for clones that carry an inducible integrated provirus, cells that constitutively expressed high levels of HSV-TK were selected against by treatment with 50μg/ml ganciclovir. Surviving cells, that were either non-transduced, or transduced but with a low level of TK expression, were treated with GST-Tat and briefly selected for inducible HSV-TK expression in hypoxanthine, aminopterin and thymidine (HAT) medium. Clonal populations were obtained by limiting dilutions and colonies were visually scored for low basal level of ECFP expression in the cytoplasm and to be highly inducible by GST-Tat by fluorescence microscopy. The HOS_A4 cell clone showed a robust and homogenous induction of ECFPskl in the cytoplasm upon treatment with GST-Tat (De Marco et al., 2008).

U1 cellular model

The last model used in this study to address HIV-1 positioning was the U1 cell clone generated from promonocytic U937 cell line (Folks et al., 1987). U1 cells contain two copies of integrated HIV-1 provirus, one on the short arm of chromosome X and the other on a rearranged chromosome 6 (Deichmann et al., 1997). U1 cells produce very low levels of HIV-1 mRNA under basal conditions, due to defective viral *trans*-activator Tat mRNAs (Emiliani et al., 1998). U1 latent proviral transcription can be markedly induced by exogenous Tat expression (Emiliani et al., 1998) or exposure to the phorbol ester TPA, which induces NF-κB heterodimers to bind within the LTR U3 region.

Both stimuli cause substantial modification of chromatin structure over the viral 5′LTR promoter (Lusic et al., 2003).

3D FISH analysis was performed in order to establish the sub-nuclear position of HIV-1 during latent state of transcription in seven different clonal cell lines described above. The minimal distance between provirus and the nuclear periphery was measured and box plot analyses are shown in Figures 19A and B. Distances obtained for all clonal cell lines are showed as normalized data, with respect to the diameter of the cells, and as non normalized data (absolutes values), respectively. In Figure 19A the median of radial distribution for all cell lines under study varied between 0.05 and 0.15 (Distance/Diameter). Considering that 0.5 is the nucleus centre, these data strongly suggest that HIV-1 was mostly located at the nuclear periphery. The same situation was observed without the normalization of data. In this case 0 μm is the periphery and the median of absolute values varied between 1 and 1.5 μm. The only exception was the HOS_ B3 clone, in which the provirus appeared more distant from the periphery, with the median around 2 μm and a considerably higher standard deviation. These results indicated that most latent proviruses localize close to the periphery. Statistical analysis confirmed that there were not significant differences among the different cell lines, with the exception of HOS_ B3 clone, as mentioned earlier. This clone harbours two integrated proviruses, one of which appears to be more distant from the periphery in comparison with J-lat A1 cell line, which represent the cell line studied more deeply in this work (K-S test, P = 0,001).

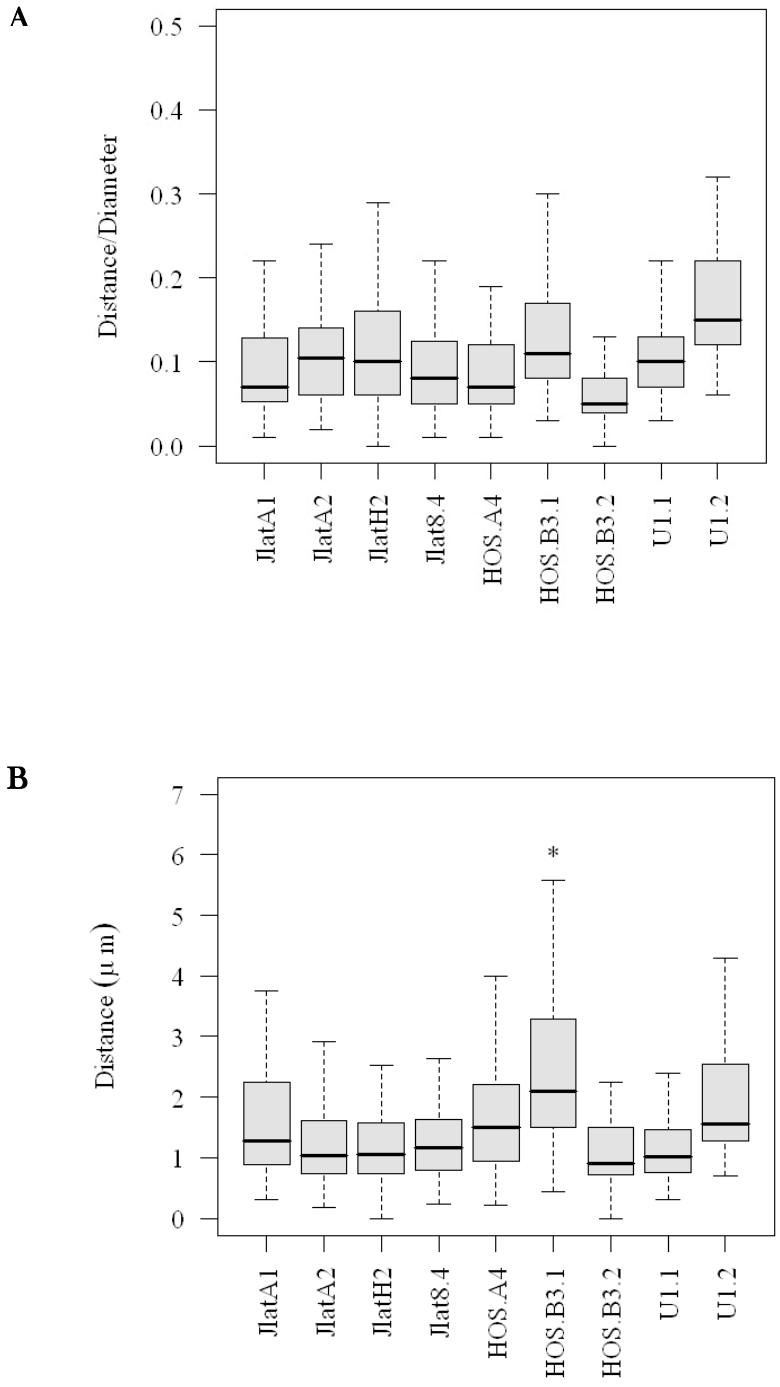


Figure 19. Localization of HIV-1 at the nuclear periphery in different cell clones

Box plot analysis of the distribution of distances from the periphery of the integrated provirus in different cell clones. When two integrations are present, as for U1 and HOS_B3, each of them is shown. Distances are presented as normalized (A) and absolutes values (B).

Table 1. Cell lines used in this study

Cell line	Integrations	Provirus/Vector	ISH probe	NOTES
J-Lat A1	1 (ChX)	Vector EV731	pEV731	(Jordan et al., 2003)
J-Lat A2	1 (Helicase V)	Vector EV731	pEV731	(Jordan et al., 2003)
J-Lat H2	1 (Ch7)	Vector EV731	pEV731	(Jordan et al., 2003)
J-Lat 8.4		HIV-R7/E-/GFP		(Jordan et al., 2003)
U1	2 (ChX active, Ch6 rearranged inactive)	HIV NL4-3	pHIV NL4-3	(Perkins et al., 2008) (Folks et al., 1987)
HOS A4	1 (Ch8 HMBOX1)	HIV24xMS2introECFPskl-IRES-TK	pHIV24xMS2introECFPskl-IRES-TK	(De Marco et al., 2008)
HOS B3	2 (Ch1 SIPA1L2 & Ch 3 DOC1)	HIV24xMS2introECFPskl-IRES-TK	pHIV24xMS2introECFPskl-IRES-TK	Our laboratory

Sub nuclear positioning of HIV-1 provirus in transcribing cells

By 3D FISH analysis, we identified that in different latent cells lines the positioning of the HIV-1 provirus was proximal to the nuclear periphery. The next step was to investigate the possible differences in HIV-1 localization between transcribing and silent provirus. To address this question we took advantage of two well-characterized cell lines J-lat A1 and HOS_A4 cells.

J-lat A1 characterization

J-lat A1 cells, generated as described before, carry a single integrated HIV-1 vector at position ChXp21.1. (Figure 20A) (Jordan et al., 2003). In our hands the integration site was confirmed using PCR analysis with one primer designed on ChX and one primer within the provirus (Table 2). Figure 20B shows the predicted band obtained.

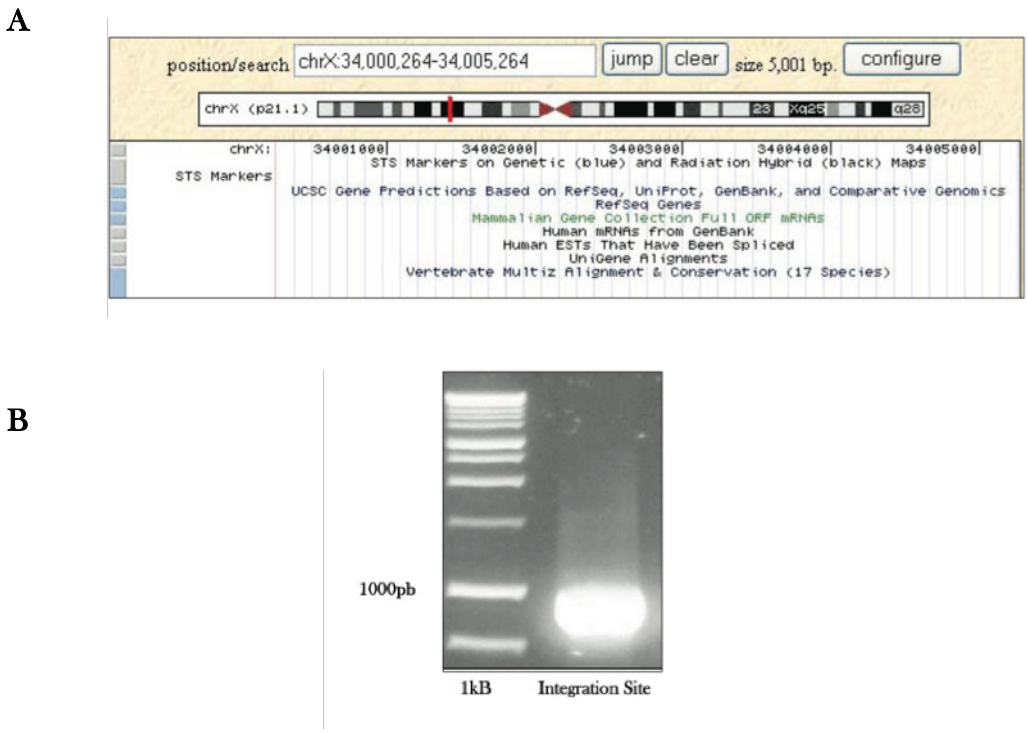


Figure 20. J-Lat A1 Integration site.

A) Integration site of HIV-1 in ChXp21.1of J-lat A1 cell clone (http://genome.ucsc.edu/) B) PCR analysis of integration site in J-Lat A1, the fragment (Integration Site), represents the amplification between the provirus and ChXp21.1. 1Kb molecular marker indicates 1000bp size.

Transcription activation

To confirm that J-lat A1 was in fact inducible for transcription, we treated the cells with TPA and measured the amount of GFP expression. Flow cytometry analysis,

demonstrated that approximately 86% of cells express GFP after 24 hours, indicating that our cellular model present a defined silenced state (off state) and an activated state (on state) (Figure 21).

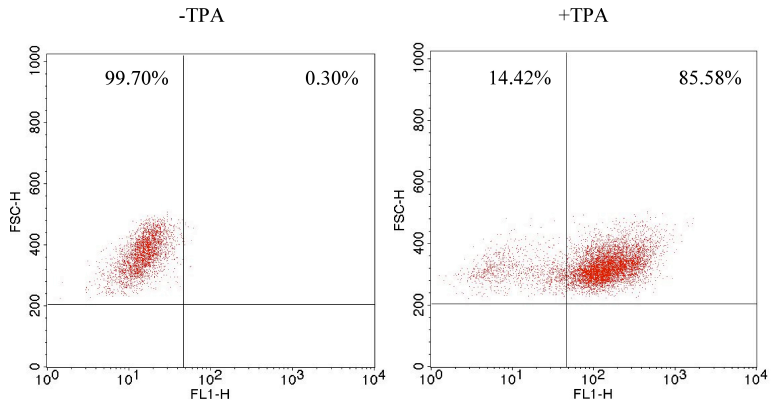


Figure 21. Transcription activation in J-Lat A1 cell line.

Cytofluorimetric analysis of cells was conducted in uninduced cells (left) and after induction with TPA (final concentration 1.6 μM) for 15 hours (right). More than 85% of cells were activated as shown by the expression of the GFP reporter.

The transcription activation of HIV-1 in J-lat A1 cells was also confirmed by the presence of the viral protein Tat and increased production of viral mRNA upon cell induction. The expression of Flag-tagged Tat (f-Tat101) upon TPA induction was confirmed by Western blot. Cells were lysed and the extracts were immunoprecipitated with an anti-flag antibody. Western blot analysis is shown in Figure 22. Moreover, we were able to detected the assembly of Tat:CyclinT1 complex upon TPA treatment. These findings indicate that the treatment with phorbol ester efficiently induces transcription of HIV-1 in J-lat A1 cells.

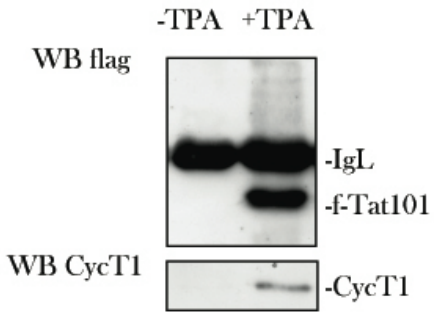


Figure 22. Assembly of Cyclin T1 with Tat on induction of J-lat A1 cells.

Top panel, western blot analysis probing with an anti-flag antibody: IgL immunoglobulin light chains. Bottom panel, western blot probing with an antibody against Cyclin T1.

A time course of nascent viral RNA is shown in Figure 23. Induction of viral mRNA was already detected at 4 hours post induction and accumulates further on with a peak at 8 hours. In non-induced cells the viral mRNA is maintained on time with a very low level of amplification, indicating that there is a basal level of transcription. Furthermore, we used primers for the intronic region, and in this case, the precursor RNA was also produced within 4 h and peaked 8 h after induction. β - Actin mRNA amplification indicated that the amount of RNA was equal in all the samples.

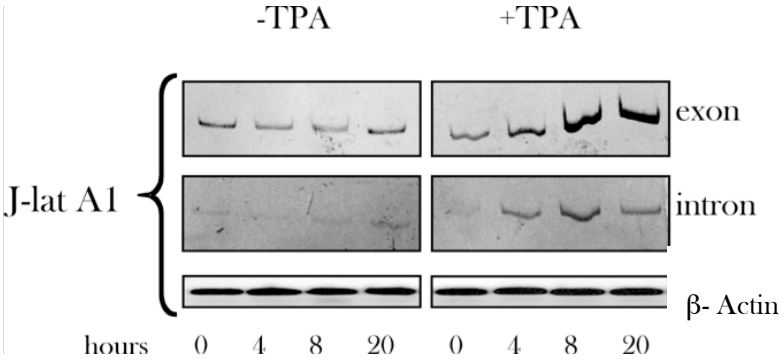


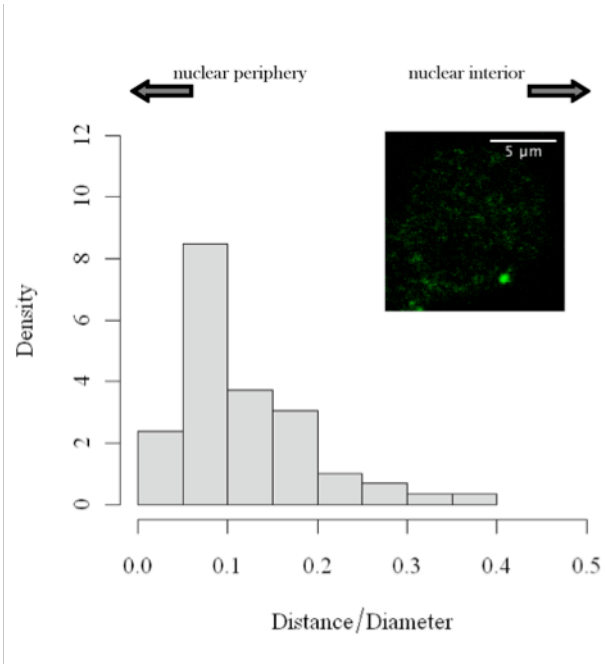
Figure 23. Time course of provirus expression in J-Lat A1 cells.

RT PCR analysis was conducted for the spliced RNA (exon, top panels), the pre-mRNA (intron, middle panels) and the β-actin control (bottom) before (left) and after induction with TPA (right). Samples were analyzed at the indicated time points (Dieudonne et al., 2009).

Localization of HIV-1 provirus in transcribing J-lat A1

Once confirmed that our cellular model presented two well-differentiated states of transcription (off/on), the next step was to identify whether the localization of HIV-1 provirus changes between these two states. To reach this goal we employed a 3D FISH method in cells with and without treatment with TPA and we observed that HIV-1 is frequently localized at the nuclear periphery also in induced cells. In both cases, the greater density of proviruses is found between 0.05 and 0.1 distance/diameter. However, after induction, the position of the provirus seems to move slightly toward the nuclear interior. However, the shift was from a median absolute distance of 1.3 μm in not induced cells to a distance of 1.7 μm in induced cells, and this change was not statistically significant (K-S test) (Figure 24 A and B).

A



B

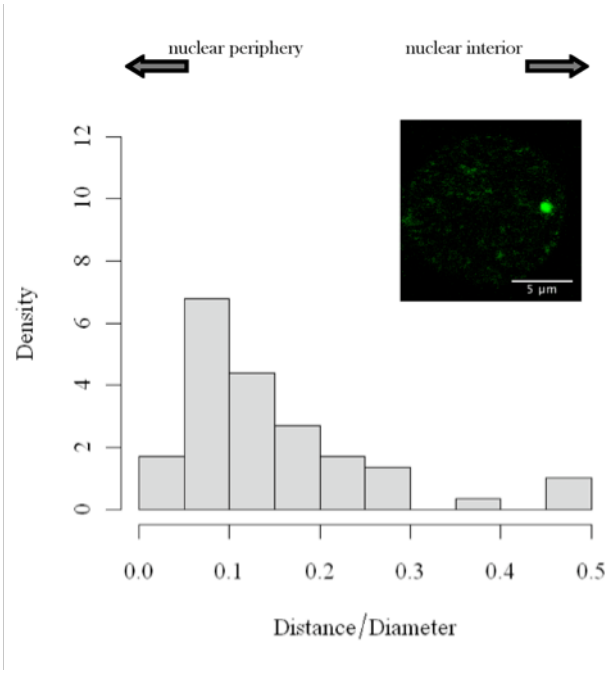


Figure 24. Localization of HIV-1 in transcribing J-lat A1.

A) Localization of HIV-1 provirus in silenced J-lat A1 cells. The abscissa represents the radius of the cell were 0 is the periphery and 0.5 is the centre. Values are normalized for the nuclear diameter. The

density is intended as the frequency of distances between the two loci that fall between a discrete interval divided for the interval amplitude. B) Localization of HIV-1 provirus in transcribing J-lat A1 cells.

RNA In situ Hybridization

So far we showed that the majority of the provirus is found at the nuclear periphery (Figure 24A and B), but since we could establish only indirectly in J-lat A1 that transcriptional activation does not result in changes of sub-nuclear positioning of the gene, we decided to visualize HIV-1 RNA at the site of transcription. We took advantage of HOS_A4 cell line, which allows the visualization of viral RNA in single living cells by the MS2 tagging method. HOS_A4 clone carry a single copy of the lentiviral vector HIV-Intro within the second giant intron of *HMBOX1* (De Marco et al., 2008).

By RNA in situ hybridization we first confirmed that nascent HIV-1 RNA remains at the nuclear periphery upon transcriptional activation (Figure 25A). Interestingly, the distribution of distances from the periphery measured for the provirus in HOS_A4 in the silenced state (median of the absolute values = 1.5 μm) did not differ significantly from those of nascent RNA in the activated state (median of the absolute values = 1.2 μm) indicating that the transcribing provirus did not move from its position (Figure 25B). Furthermore, the major density of proviruses was localized between 0.05 and 0.1 distance/diameter from the nuclear periphery confirming the results obtained for genome positioning shown in the Figure19.

To further confirm that HIV-1 RNA is transcribed at the nuclear periphery, a time course of RNA biogenesis in living cells expressing both Tat and EYFP-MS2nls was

measured (Figure 25C). Once more, our results clearly demonstrated, that active transcription from an integrated provirus was continuously associated to the nuclear periphery also in living cells.

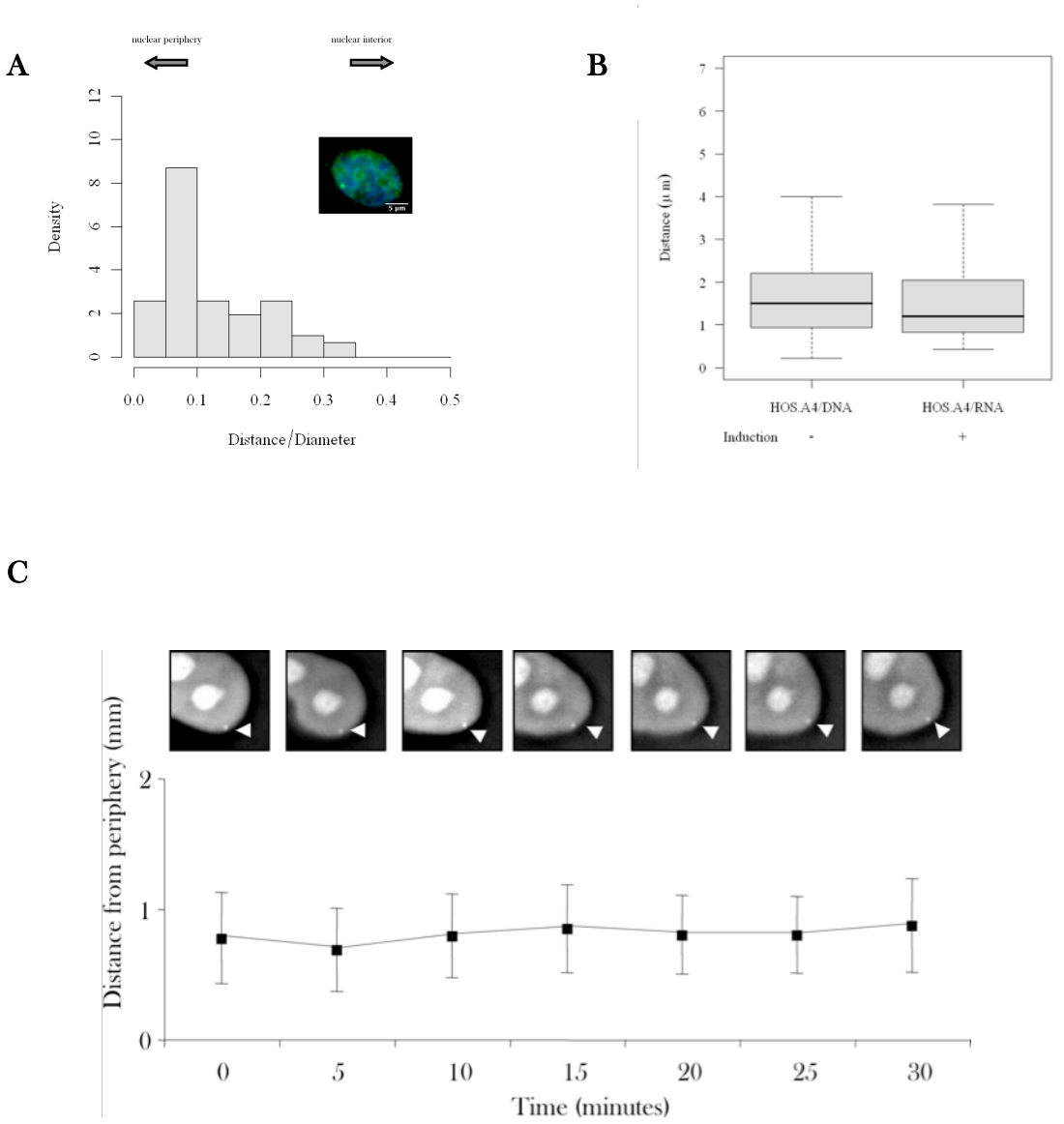


Figure 25. HIV-1 nascent RNA and its positioning at the nucleus in living cells.

A) HOS_A4 cells were analyzed for the localization of the HIV-1 nascent RNA with respect to the nuclear periphery in Tat-activated cells (n = 62). Analysis of distances was conducted as in (Figure 24) Inset: example of the RNA in situ hybridization with an intronic probe that shows localization of nascent HIV-1 RNA at the site of transcription. B) Box plot analysis of the distribution of absolute distances from

the periphery of the provirus in latent HOS_A4 cells (left) compared to the distribution of absolute distances from the periphery of the nascent HIV-1 RNA in induced cells (right). The status of induction is indicated below. The differences in the distribution of the distances to the periphery in the inactive or active state were not statistically significant (K-S test, $p > 0.01$). C) Localization of transcription in single living cells. HOS_A4 cells expressing Tat and EYFP_MS2nls were monitored in time for the localization of nascent RNA. Distances from periphery for 13 cells (± s.d.) are plotted at 5 minutes interval for half an hour. A picture of single nucleus is show at the various time points to illustrate the position of the transcribing locus (white arrowhead).

Identification of genomic sequences interacting with HIV-1 provirus

To establish the possible mechanisms involved in the silencing of HIV-1 in latent clones, we decided to use the Circular Chromosome Conformation Capture (4C) approach employing J-Lat A1 cellular model. Based on the Chromosome Conformation Capture (3C) method (Dekker et al., 2002), 4C approach exploits a circular intermediate to allow the identification of inter-chromosomal interactions without previous knowledge about the interacting sequences (Figure 26) (see introduction).

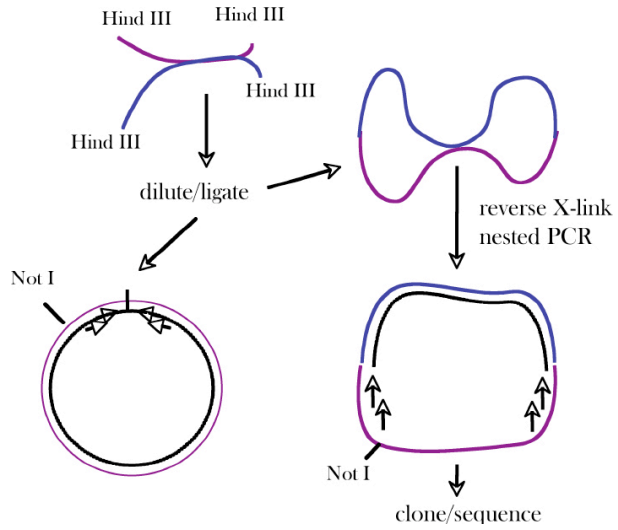


Figure 26. Circular Chromosome Conformation Capture (4C) assays.

A) Diagram of the 4C protocol. The bait (blue line) derived from the HIV-1 provirus, is cross-linked to an unknown genomic locus (violet line). After digestion with Hind III, the reaction is diluted and ligated to generate both intra- and inter-molecular ligations. Intra-molecular ligations could be reduced by Not I treatment. After reversal of cross-linking a nested PCR is performed with primers within the bait pointing outwards.

Therefore, J-Lat A1 cells were either left untreated or incubated with TPA for 8 hours and then cross-linked with formaldehyde. In initial experiments the genome of J-lat A1 was digested with AflII, as shown Figure 27. This enzyme cut twice in HIV-1 provirus (EV731), within the 5' and 3' LTR sequences (positions 518 and 4409 respectively, according to the sequence of the provirus). After digestion the genomic material was diluted, ligated and de-cross linked. In order to amplify the possible sequences interacting with HIV-1 provirus, we designed primers pointing outwards from an AflII fragment, within the EV731 provirus (Figure 26). Due to our lack of knowledge about the size of the fragment to be obtain, we used different programs of PCR amplifications,

different amounts of DNA, and diverse tag polymerases, in order to set up the proper protocol that allows the identification of the possible sequences interacting with the bait in the cellular nucleus. A smear of several bands was evident in agarose gels after PCR reaction, in most of the conditions. Finally the bands were either excised from the gel or cloned as a bulk. However, all clone sequences showed a religated bait. This probably was related to the length of the bait that allowed a high frequency of relegated products. In order to generate smaller bait, we decided to use HindIII restriction enzyme which cut more frequently inside of EV731 (Figure 27).

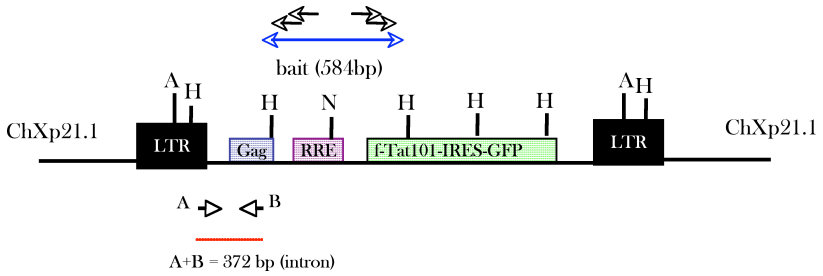


Figure 27. Diagram of the HIV-1 construct integrated in J-Lat A1.

Sites of cleavage by HindIII (H), AflII (A) and Not I (N) are indicated. The position of the primers used in the 4C protocol is indicated by black arrows above the bait indicated by a blue arrow. Below the position of the primers (A+B) for the amplification used as control in 3C and 4C experiments.

Following the same approach, primers pointing outwards from a HindIII fragment (position 1088-1672 according to the sequence of the provirus) within the HIV-1 provirus were used as a bait to amplify the possible DNA sequences interacting with it (Figure 26). As shown in Figure 28, we were able to distinguish a smear of several bands in agarose gels after nested PCR for active and silent cells. The major difference

between these two states consisted in the appearance of bands at approximately 600 bps and 1200 bps only in the induced cells. The identity of the bands was confirmed to be the bait (584 bps) and multimers of this sequence. We hypothesized that this could be due to multiple runs of PCR over the junction on religated templates (Figure 26). To confirm this hypothesis we introduced an additional step in the 4C procedure by cutting with Not I (position 1145 of the provirus), a unique restriction site within the bait fragment (Figure 26 and 27), before the nested PCR reaction. As shown in the second panel in Figure 28 treatment with Not I reduces the products of the religated template. However, despite this modification of the protocol, we were unable to detect host chromatin sequences associated with the provirus in activated cells. Cloning of the PCR products showed only religated templates, primer multimers or boundary sequences of the provirus due to incomplete HindIII digestion. This could be explained by a technical failure to identify all the fragments involved and may be solved by more in depth sequencing or by using microarrays (Simonis et al., 2006). Alternatively, the observation that the intra-molecular ligation by-products of the 4C reaction increase upon induction may indicate that the chromatin loops out from a more compact conformation and therefore is more accessible to enzymatic digestion. Looping of HIV-1 chromatin upon transcriptional activation coupled to high levels of Tat-mediated transcription and local enrichment of nascent RNA, although not formally proven, may well account for a single transcription factory exclusive for HIV-1 transcription (Boireau et al., 2007) (Perkins et al., 2008).

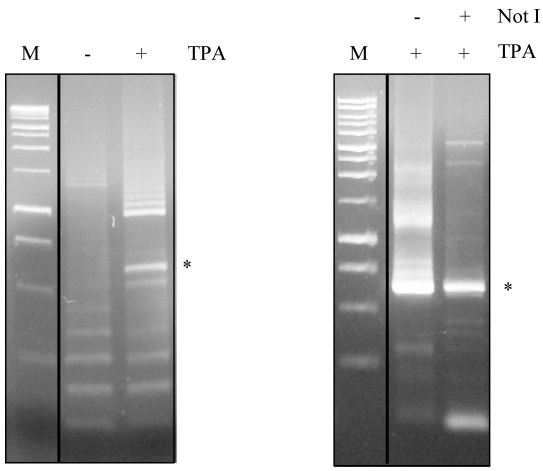


Figure 28. 4C amplified fragments from J-lat A1 cells.

Products of the nested PCR amplification of cells either not induced or induced with TPA were resolved on agarose gels. Molecular weight marker (1kb, M) and position of the major product of intra-ligation (asterisk) are shown. In the second panel 4C amplified fragments from J-lat A1 cells after Not I digestion. To reduce intra-molecular ligation in the TPA-activated cells the cross-linked material was treated with Not I prior to performing the nested PCR. Molecular weight marker (1kb, M) and position of the major product of intra-ligation (asterisk) are shown.

In contrast, in silenced cells after bulk cloning of all the nested PCR products we identified a specific fragment from host chromatin that corresponded to the positions 37.017.830 – 37.017.941 of chromosome 12q12 (Figure 29). To further verify these results we moved to 3C analysis, which in contrast with 4C assays allows studying chromosomal folding in the nucleus only if previous knowledge of the bait and the interacting sequence exists. With a primer (Table 2) mapping within the identified fragment in Ch12, and a primer in the bait, as shown in Figure 30A, a hemi-nested PCR was perform on 3C latent J-Lat A1 samples, confirming the interaction of the provirus with Ch12q12 (Figure 30A, top panels).

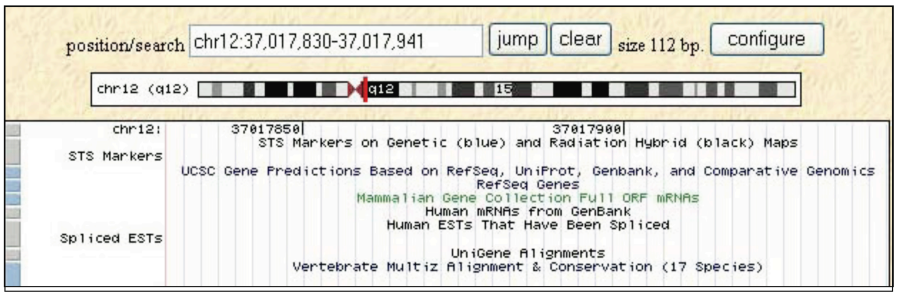


Figure 29. Region of Chromosome 12 found to interact with the provirus by 3C and 4C techniques.

Table 2. Primers used for PCR amplification.

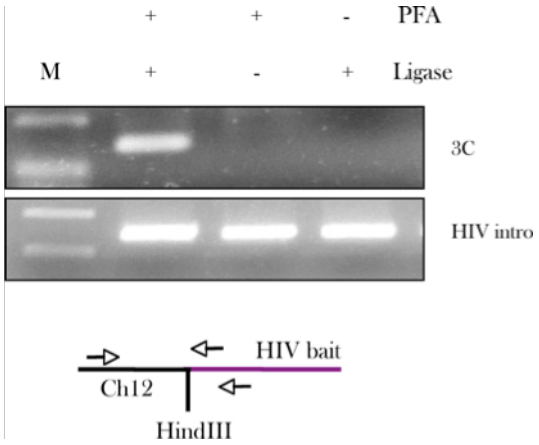
Technique	Primers	
	First Amplification	Second Amplification
4C PCR	5' out A Hind III 5' TGTGCGGTGGTCTTACTTTTG 3'	5' out B Hind III 5' GTTTTGCTCTTCCTCTATCTTGTC 3'
	3' out Hind III 5' AGATTTGGAATCACACGACC 3'	3' out B Hind III 5' GGAGTGGGACAGAGAAATTAAC 3'
3C PCR	5' Chromo 12B 5' GCTTTTCATGACGTTGCTCTCCTC 3'	5' Chromo 12B 5' GCTTTTCATGACGTTGCTCTCCTC 3'
	5'out A Hind III 5' TGTGCGGTGGTCTTACTTTTG 3'	5' out B Hind III 5' GTTTTGCTCTTCCTCTATCTTGTC 3'
RT PCR	Splice Exo: 5' CGAGATCCGTTCACTAATCGAATG 3' A (nuc 1 B177): 5' CGTCTGTTGTGTGACTCTGGTAACT 3' B (Splice Intro): 5' GGATTAACTGCGAATCGTTCTAGC 3'	
PCR Integration Site	pEV731	ChX
	IS ChrX R 5' GTGCCTTGATGGGAATTAAG 3'	Splice Intro (B) 5' GGATTAACTGCGAATCGTTCTAGC 3'

As control of template amplification we used a set of primers that detect the HIV-1 provirus (A+B, bottom panels of Figure 30A). As negative controls we omitted the ligation step and the cross-linked step in 3C assays, and as we were expecting no amplification was observed in both cases.

Having shown that in the off state the provirus interacts physically with a pericentromeric region of Ch12q12, we sought to know whether this interaction exist also in transcriptionally active cells. Therefore, we examined J-lat A1 cells by 3C after the induction with TPA. As shown in Figure 30B, the amount of template that could be amplified by hemi-nested PCR was reduced upon activation of HIV-1 transcription. Serial 2-fold dilutions of the template demonstrated that the 3C product was abundant in latent versus activated cells.

We conclude that in J-lat A1 quiescent cells the provirus integrated in ChXp21.1 interacts physically with Ch12q12. Upon activation of proviral transcription with TPA, this interaction is lost.

A



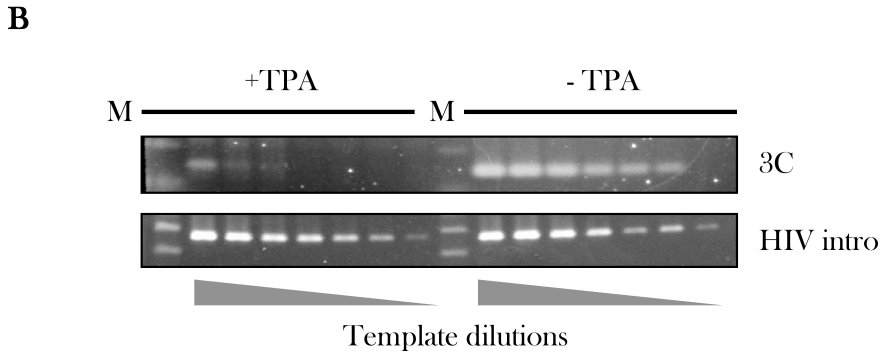


Figure 30. 3C analysis at the site of HIV integration in J-lat A1 cells.

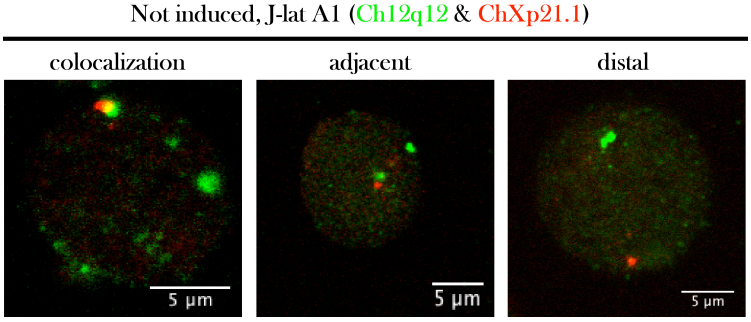
A) 3C analysis of the interaction of the provirus and Ch12q12. To confirm 4C data a 3C analysis was performed by hemi-nested amplification using a primer within the Ch12q12 region (see diagram). Control amplification was performed with primers mapping within the provirus (Table 2., Figure17). B) Loss of the interaction between the provirus and Ch12q12 upon induction of transcription. 3C analysis of the Ch12q12/provirus interaction was performed both in induced (left) and non-induced (right) cells. Two-fold serial dilutions of the template show that in non-induced cells there is more cross-linked material for the interaction.

Mapping of the genomic regions by 3D-FISH

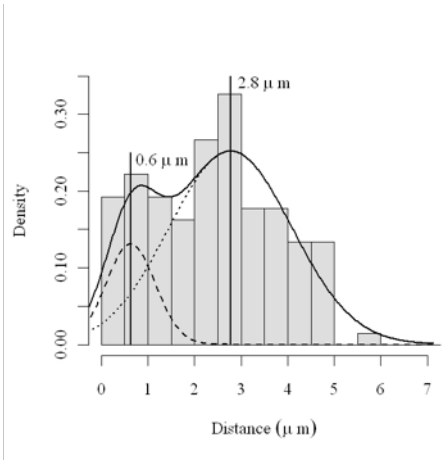
In order to explore in depth the interaction between Ch12q12 and the integrated provirus into ChXp21.1, we performed 3D FISH on J-lat A1 cells (following the protocol explained in detail in Material and Methods). To mark ChXp21.1 we used the BAC RP11-77013 and to mark Ch12q12 we used the BAC RP11-379CZ4. These probes were selected to cover the region where the provirus is integrated (ChX) and the region in Ch12 that we demonstrated by 3C to be physically associated with it. As shown in Figure 31A, single-cell disposition of the two marked regions varied. First, we measured the minimal distance between ChXp21.1 and Ch12q12 (Figure 31B), to look for the interaction detected in 4C, and between Ch12q12 and Ch12q12 as a control

(Figure 31C). More than one hundred cells were analyzed and the results are shown in Figure 31B and C. Both samples show a Gaussian distribution around a 3-μm mean of the minimal distances of interaction. However, for the Ch12/ChX analysis, the data could be fitted with two Gaussians, with the smaller representing a subset of close associations, centered around 0.6 μm (Figure 31B) and the other centered around 2.8 μm. We proposed that all the interactions found to be under or equal to 0.6 μm could be considered adjacent. Within the population of cells tested in the silenced state, 2.9±0.7% showed co-localization of the two regions and 11±4% showed proximity at a distance of <0.6 μm (Figure 31E). Clearly, this could be the fraction of interaction that we could detect by 3C. Therefore, we wanted to explore what happens when HIV-1 transcription is induced by TPA. As shown in Figure 31D, treatment of cells with TPA induces the disappearance of the cluster of close associations of Ch12q12 with ChXp21.1. We found that the Gaussian distribution of the distances between these two chromosomes in activated cells presented a very similar behavior to the Gaussian generated in control samples (Ch12 and Ch12), confirming once more that this interaction is disrupted upon transcription activation. Fraction of cells presenting colocalization and adjacency were significantly reduced to 0.8±0.8% (colocalizing) and 3±2% (adjacent) (Figure 31E).

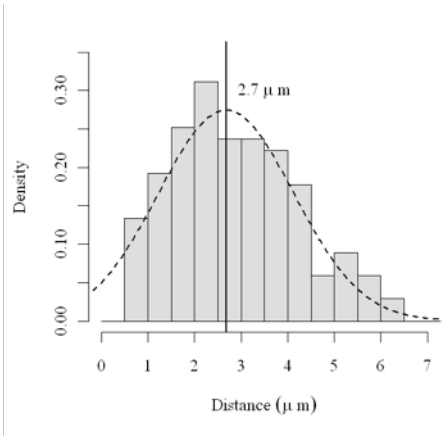
A



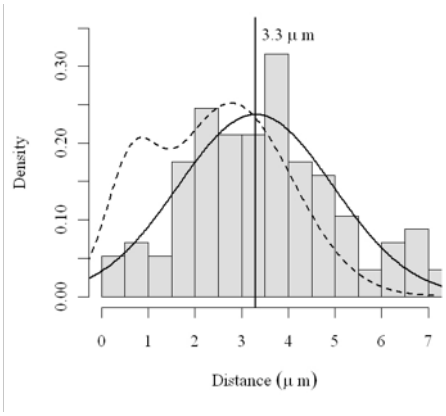
B



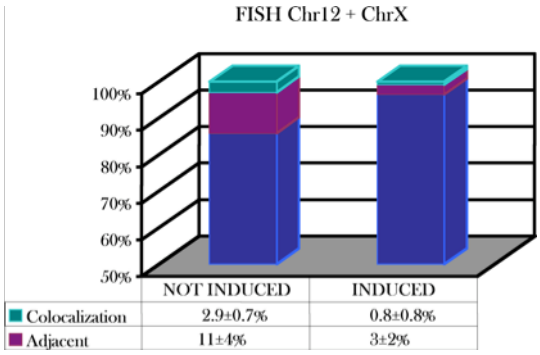
C



D



E



	NOT INDUCED	INDUCED
Colocalization	2.9±0.7%	0.8±0.8%
Adjacent	11±4%	3±2%

Figure 31. Analysis of the interaction between Chr12q12 and ChXp21.1

A) J-lat A1 cells were analyzed in FISH with a probe for Ch12q12 (BAC RP11 379CZ4 green) and a probe for ChXp21.1 (BAC RP11-77013 red). The position of the two loci varied from co-localizing (left panel), to adjacent at ≤ 0.6 μm (central panel), to distal at > 0.6 μm (right panel). B) Quantitative analysis of the distribution of the distances between Ch12q12 and ChXp21.1 in non-induced J-lat A1 cells (n =135). The density is intended as the frequency of distances between the two loci that fall within a discrete interval divided for the interval amplitude. The data were fitted and showed a bi-Gaussian distribution (solid line with the two Gaussian shown as dotted lines) with a subset of nuclei where the two genomic loci were closely associated. The mean values of the two Gaussian are also indicated by vertical lines (0.6 μm and 2.8 μm, respectively). C) As a control, the same analysis as in B was conducted for the distances between Chromosome 12. This analysis showed a Gaussian distribution (dotted line). D) The distribution of minimal distances between Chromosome X and 12 is compared between non-induced (dotted line, n = 135) and induced cells (solid line, n = 114). Histogram shows the distribution of the induced cells. E) Quantitative analysis of the frequency of nuclei showing co-localization (green), proximity at ≤ 0.6 μm (fuchsia) or distance at > 0.6 μm (blue). Numeric values are indicated below. Decrease of the co-localization is significant (χ^2 , p = 0.016).

The region of interaction at position Ch12q12 is close to the centromer of chromosome 12, a heterochromatic region that might be involved in the silencing of HIV-1 in trans. Indeed, pericentromeric satellite alphoid DNA repeats are present and can be used as a marker for heterochormatin at that location. Therefore we exploited a DNA probe specific for these α-satellite repeats present in Ch12 (Baldini et al., 1990), in order to map other possible sequences interacting with the region of ChX where HIV-1 was integrated(Baldini et al., 1990). As expected, the distances between the two marked regions varied (Figure 32A). A fraction of the non-induced cells showed colocalization of signals (4±2%), whereas 7±0.7% of cells showed proximity at a distance of 0.6 μm.

As shows in Figure 32B, when cells were induced the proximity was lost. The colocalization was reduced to 1±1% and the fraction showing proximity at a distance of <0.6 μm was also reduced to 0%. We next explored the localization of the provirus (using a specific probe for it), with respect to Ch12 α-satellites (Figure 32C). In this case, 4±2% of non-induced cells showed colocalization of the provirus with the centromer and 4±2% showed proximity at a distance of <0.6 μm (Figure 32D). The same figure shows that, when cells were induced, the colocalzation was reduced to 1±1% and the fraction showing proximity at a distance of <0.6 μm was also reduced to 1±1%.

These results confirm the 3C data since we could demonstrate that a fraction of the population showed interaction among the provirus and the pericentromeric region of chromosome 12.

Interestingly, we also observed that upon TPA treatment the interaction between provirus and Ch12 is disrupted both in 3C and FISH. How this disruption occurs remains to be established. We did not observe an increase of the cell volume in J-Lat A1 upon TPA stimulation as it occurs for example in resting T cells (Branco et al., 2008). Furthermore, although the TPA treatment induces cell-cycle arrest, it remains possible that in the period after TPA stimulation a certain proportion of cells are still capable of completing mitosis thus justifying the disruption of the interaction observed in J-lat A1. Alternatively, active rearrangements of chromatin domains during interphase may be required (Louvet and Percipalle, 2009).

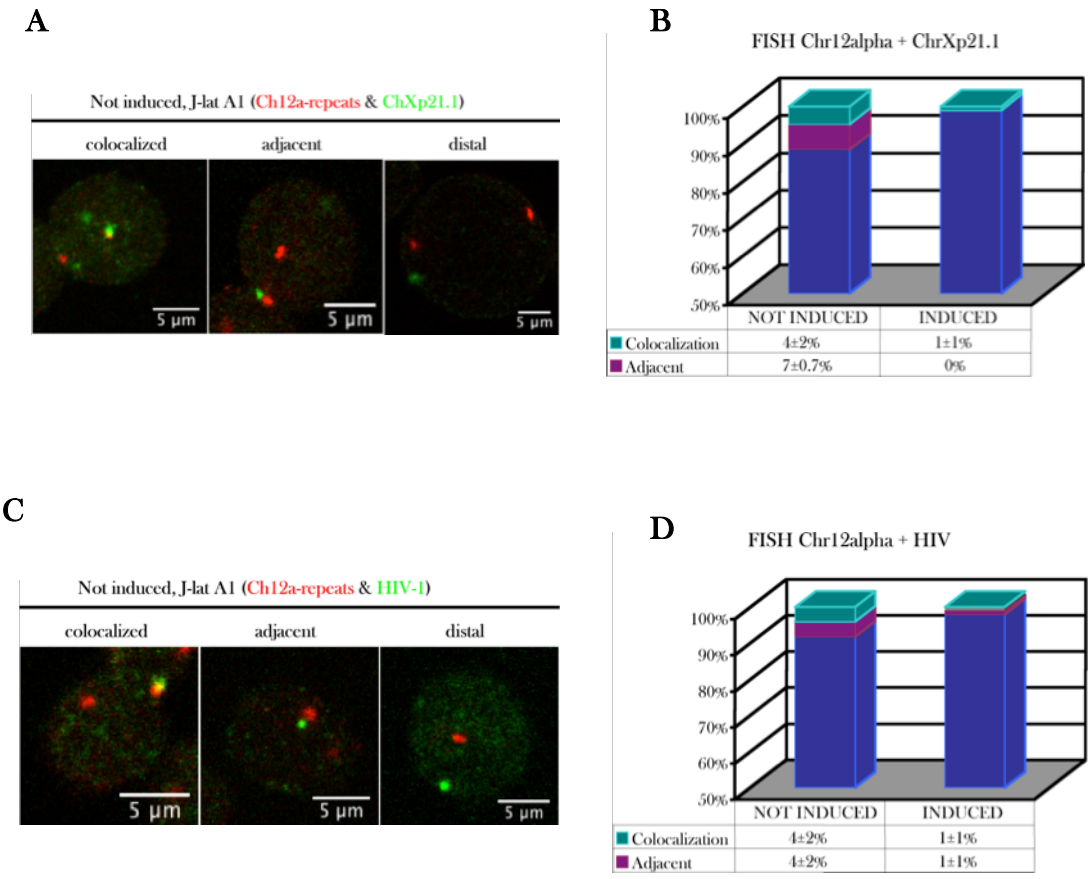


Figure 32. Analysis of the interaction of Chr12 α-repeats with either ChXp21.1 or the provirus.

A) J-lat A1 cells were analyzed in FISH with a probe for Ch12 α-repeats (red) and BAC RP11-77013 (Green) mapping at ChXp21.1. Similarly to Figure 31A, the position of the two loci varied from co-localizing (left panel), to adjacent at ≤0.6 µm (middle panel), to distal at > 0.6 µm (right panel). B) Quantitative analysis, performed like in Figure 31E, is shown in B (non-induced, n = 105; induced, n = 103). Decrease of the co-localization is significant (χ^2, p = 0.016). C) J-lat A1 cells were analyzed in FISH with a probe for Ch12 alpha repeats (red) and the provirus (Green). D) Quantitative analysis for Ch12 alpha repeats and the provirus (non-induced, n = 104; induced, n = 78). Decrease of the co-localization is significant (χ^2 , p = 0.04).

Localization of HIV-1 provirus independent of its interaction with Ch12

A question that arises from our studies so far was whether the localization of HIV-1 provirus changes when the interaction with Ch12 is disrupted (transcriptional activation). In order to answer this question, we decided to choose a subset of cells were the signals for Ch12 pericentromeric α-satellites and the provirus were at a distances of < 0.6 μm (the mean value of the Gaussian fitted for minimal distances between Ch12 and ChX in non-induced cells) and measured their distance from the periphery of the nucleus. The measurements of this subset of population that presented the interaction, interestingly demonstrated that the position of the interaction is consistently associated with the nuclear periphery (Figure 33A) with the mayor amount of proviruses in the range of 0.05 and 0.1 (radial distribution). In other terms, in a certain proportion of cells, the silent provirus is found in close proximity of Ch12 pericentromeric α-satellites close to the nuclear periphery. Then, we compared the localization of this interaction with the position of the HIV-1 independent of Ch12 in silent and transcriptional states. As we showed before, the provirus was associated with the periphery irrespective of its association with Ch12 (Figure 25). The Figure 33B shows a box plot comparing these three different situations; the distribution of the distances is shown as absolutes values where the value 0 represents the periphery and 0.5 the centre of the nucleus.

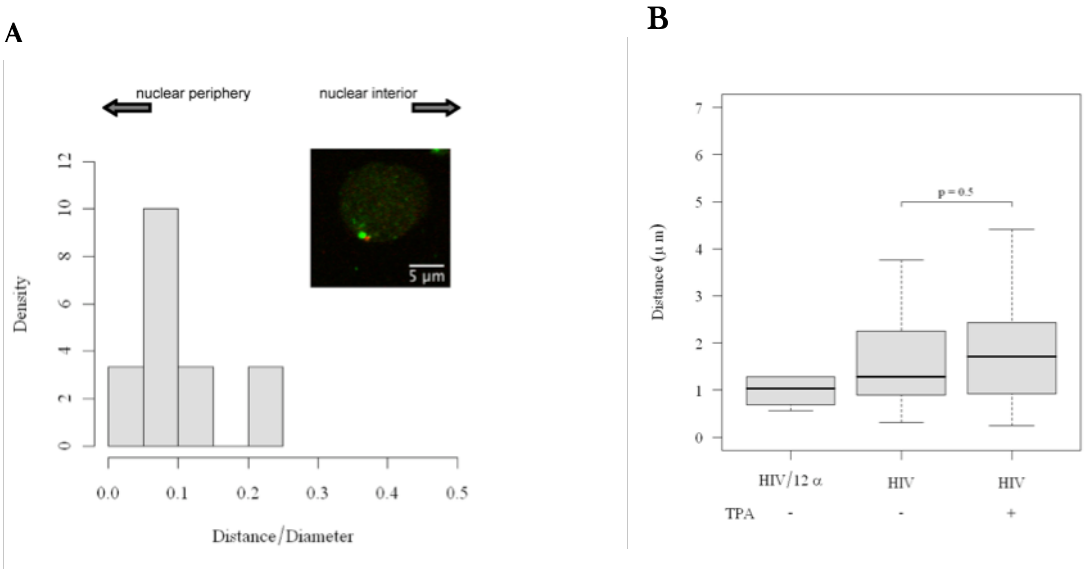


Figure 33. Localization of the interaction Ch12/HIV-1 in the cell nucleus.

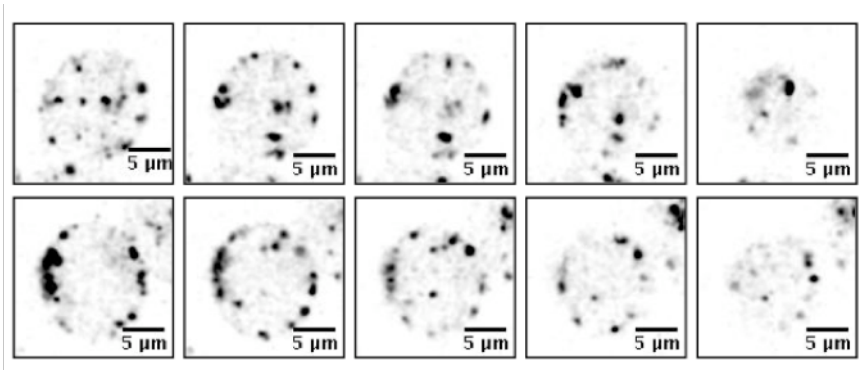
A) The sub-population of J-lat A1 non-induced cells carrying the interaction between the provirus and Ch12 centromeric α-repeats was analyzed for distances of the interaction from the nuclear periphery (n=25). The analyses were done as in Figure 24. B) Box plot analysis of the distribution of absolute distances from the periphery of the subset of interactions of provirus and Ch12 and centromeric α-repeats (median =1 µm). In addition, the distance of the provirus from the periphery, independent of its association with chromosome 12, is shown both in the inactive (median = 1.3 µm) and in active (median = 1.7 µm) states. Status of induction with TPA is indicated below. The differences in the distribution of the distances to the periphery of the provirus in the inactive or active state were not statistically significant (K–S test, P = 0.5).

Heterochromatin distribution

Since heterochromatin is related with silencing, we decided to study the distribution of heterochromatin in J-lat A1. In order to do this we used two different approaches, the first one was a 3D FISH with a probe for pan-α-satellite repeats probe that hybridizes *in situ* exclusively to the centromeric regions of all human chromosomes (Mitchell et al.,

1985) (Figure 34A), the second one was a Immuno-FISH against a heterochromatin marker, histone H4 Lys20 (Figure 34B). Several groups have described the organization of chromatin in lymphocytes. Human resting lymphocytes (G0) show a predominantly peripheral localization of centromers (Weierich et al., 2003), while cycling cells show also centromers in the interior at various proportions through the cell cycle (see: (Solovei et al., 2004) and references therein). In both cases the distribution of pericentromeric heterochromatin in J-lat A1 is characteristic of cycling lymphocytes (Figure 34A and B), with a high, but not exclusive prevalence at the nuclear periphery.

A



B

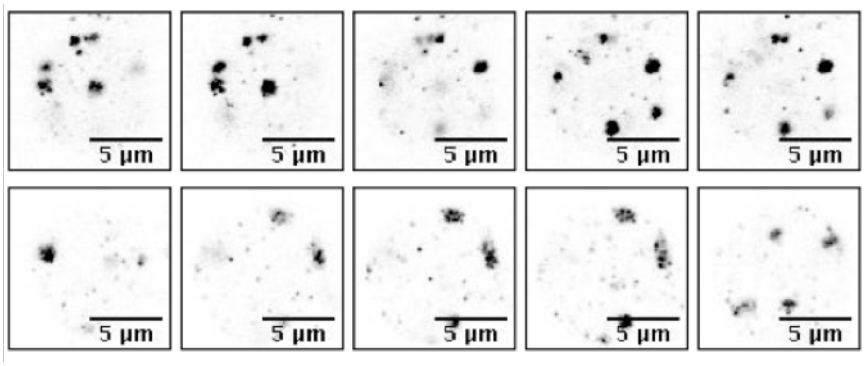
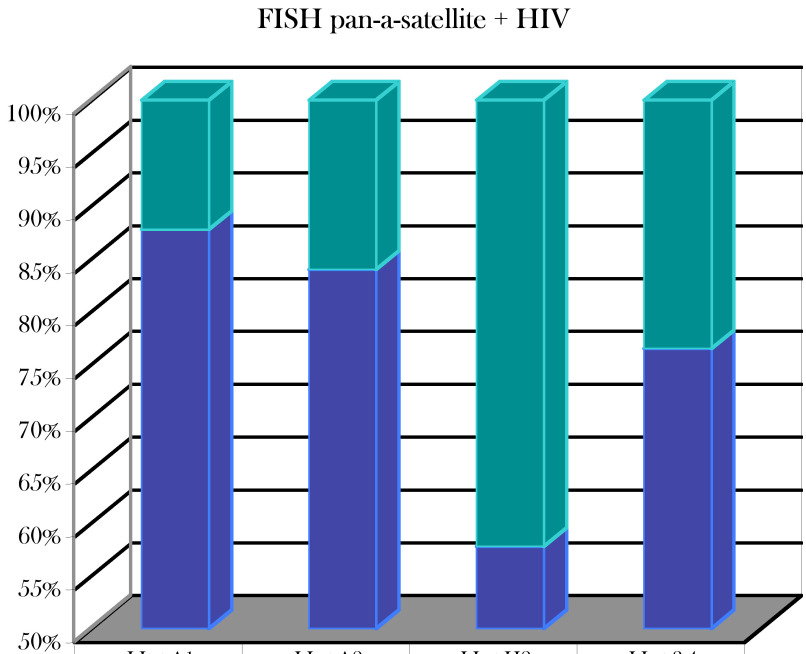


Figure 34. Heterochromatin Distribution

A) Stained with a pan-satellite-repeats probe that hybridizes in situ exclusively to the centromeric regions of all human chromosomes (Mitchell et al., 1985). J-lat A1 cells used in our work present a distribution of centromers characteristic of cycling lymphocytes. B) Distribution of centromeric heterochromatin in J-lat A1. Optical sectioning of two J-lat A1 cells stained with an antibody against tri-methylated H4K20, a marker of centromeric heterochromatin (Schotta et al., 2004).

Our results indicated that approximately 10% of J-lat A1 were not activable with TPA, while around 10% of the cells presented an interaction between Ch12 and HIV-1. The next step was to understand if this subpopulation of cells which present the interaction between Ch12 and the provirus, corresponds to those that were not activated by TPA (Figure 21). To answer this we proceeded with 3D FISH analysis. Different Jurkat cell clones carrying an integrated provirus were analyzed with a probe for pan-α-satellite repeats, and specifics probe for the provirus (Table 1). The analyses of the data are shown in Figure 35. The graphic shows the percentage of interaction between the provirus and pan-α-satellite repeats and the percentage of GFP positive cells following TPA treatment. While different latent cell lines show colocalization of the provirus with a pan-α-satellite probe at various frequencies, the extent of cells not reactivating with TPA did not correlate with the extent of colocalization. For example J-lat H2 cells show 42% of colocalization of the provirus with centromeric heterochromatin while more than 85% of cells could not be reactivated by TPA (nor by TNFα). Hence, a consistent population remains inactive (>85%) of which only a portion can be considered silenced by proximity to pericentromeric heterochromatin. Similar findings were obtained in all Jurkat cell clones under investigation. Most likely the timing of transcriptional activation differs among different cells in the population, possibly reflecting also the chromatin interactions of the provirus.



	J Lat A1	J Lat A2	J Lat H2	J Lat 8.4
Coloc. < 0.6 mm	12,24%	16%	42,22%	23,52%
Distal	87,75%	84%	57,77%	76%
GFP (- TPA)	0,3%	0,44%	0,53%	0,18%
GFP (+ TPA)	86%	1,5%	12%	42%

Figure 35. Relation between heterochromatin interaction and TPA activation.

HIV-1 integration and nuclear periphery

Lamin-associated domains (LADs) represent a chromatin repressive environment tightly associated with the nuclear periphery. Recently Bas Stenseel's group (Guelen et al., 2008) reported a high-resolution map of the interaction sites of the entire human genome with LADs.

In order to further analyze the position of HIV-1 with respect to the nuclear periphery, we decided to study whether there was a correlation between integrations sites and LADs. To do this, we took advantage of two libraries of integrations sites generated by Frederic Bushman and collaborators (Lewinski et al., 2005). One set of integrations

sites was derived from cells that expressed the provirus constitutively, and the other from cells that expressed the provirus poorly, but could be reactivated by TNFα. This cell population represents a relevant particularity, since is the population that persists during antiretroviral therapy (Marcello, 2006). From this analysis done by Somdutta Dhir from the ICGEB bioinformatics group, we observed that in both cases, (integrations in poor and rich genes regions), integration within LADs was equally disfavored comparing to random generated integration sites (Figure 36), showing around 15 % of association with LADs for the two libraries and 40% of association for random integrations. However this analysis was carried out in two different cells types that may differ in the nuclear organization, suggesting that a more clear results can be obtained from the same cell type or better from primary cells from latently infected patients.

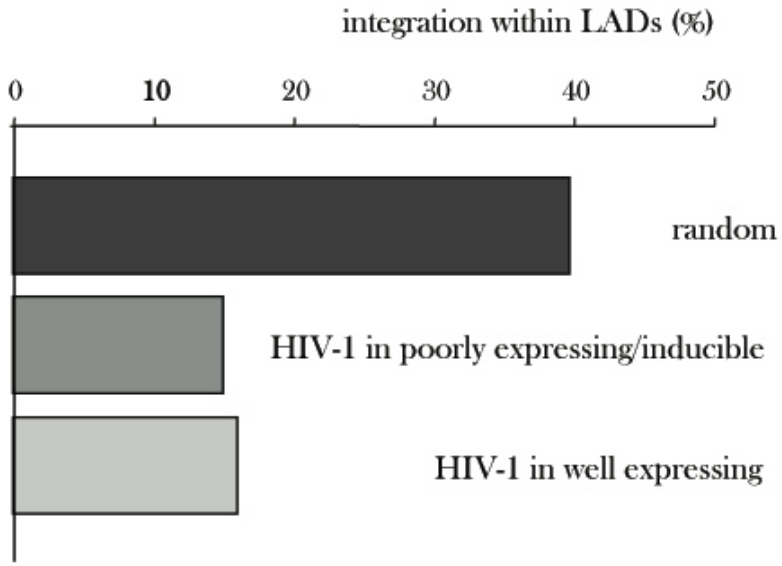


Figure 36. HIV-1 integration and association with LADs .

HIV-1 integration into the LADs is disfavored. The integration sites of (i) cell carrying a proviral vector that is well-expressing (n = 587) and (ii) cells carrying an integrations site that is poorly expressing but inducible (n = 384) derived from Lewinski et al. are compared to random integrations (100 iterations of n = 500) for their association to lamin associated domains (LADs) as defined in Guelen et al.

DISCUSSION

HIV-1 preferentially infects CD4+T lymphocytes and macrophages. After entry into the target cell, the viral RNA is reverse transcribed and integrated into host genome. Upon integration, the proviral DNA is organized in a chromatin structure and is subject to the same control mechanism governing transcription in the cell nucleus. In such a way HIV-1 can both be actively transcribed and persist in a latent form in long-lived infected cells. Viral reservoirs established early during the infection remain unaffected by anti-retroviral therapy for a long time and are able to replenish systemic infection upon interruption of the treatment. Consequently, the major obstacle to HIV-1 eradication is the establishment of a latent infection. Therefore, a better understanding of the basic mechanisms underlying the establishment and the long-term maintenance of HIV-1 latency will provide new insights for innovative therapeutics aimed at viral eradication.

Molecular mechanisms underlying long-term transcriptional control of proviral gene expression are still obscure. Regulation of gene expression is profoundly involved in HIV-1 pathogenesis and the chromatin three-dimensional conformation into the nucleus is one of the main mechanisms controlling gene expression. Therefore the aim of this study was to investigate the chromatin architecture at the site of HIV-1 integration in active and silenced states of transcription. The cellular model chosen for this study involved several clones of a Jurkat lymphoblastoid T cell line (J-Lat clones) carrying an integrated HIV-1 vector with a reporter gene (GFP). These cell lines mimic the latent state of HIV-1 and can be induced to transcribe the silenced provirus.

Gene positioning

Silent gene loci are often found in close proximity to centromeric heterochromatin, and centromeres are known to be preferentially associated with the nuclear periphery (Brown et al., 2001) (Brown et al., 1999) (Merkenschlager et al., 2004). In order to determine the positioning in the nuclear space of HIV-1 and its possible relation with heterochromatin we analyzed different J-Lat cell clones, representing a wide range of provirus latency.

The nuclear periphery has been described as a repressive compartment for transcription in eukaryotic cells (see introduction). Definition of this compartment at the nuclear periphery varies between different reports. Several groups used normalized radial distances, while others used data without normalization, which represent the absolutes values. In order to provide a more accurate interpretation of our results and to compare with other reports in different cell lines, we used both kinds of analysis. The Figures 19 A and B, show the distances measured between the provirus and the periphery of the nucleus for all cell lines as normalized data and as data not normalized, respectively. Our results demonstrated that latent HIV-1 provirus is mostly associated with the nuclear periphery independently of the integration site. In both cases statistical analysis demonstrated that there are not significant differences between the different cell lines and our model system J-Lat A1. The only exception is the HOS-B3 clone, which harbors two integrated proviruses, both within active genes that are transcriptionally competent (data not shown). One provirus appears more distal from the periphery compared with J-lat A1 (K–S test, P1/40.001). However this cell line requires further investigation to explain its internal localization.

HIV-1 is a retrovirus that crosses the nuclear pore to integrate into cellular chromatin. The PIC containing retro-transcribed viral DNA and the viral integrase must cross the nuclear pore and overcome the steep DNA concentration gradient in the nucleus (Fassati, 2006). Consequently integration is more likely to occur in the proximity of the nuclear periphery rather than towards the interior of the nucleus, as long as a favorable integration site is found. This hypothesis has been recently confirmed by the observations that the pre-integration complex is found preferentially at the nuclear periphery in decondensed chromatin (Albanese et al., 2008). In most cells, the virus is transcriptionally active, but in some cells it may be subjected to reversible silencing. This latter feature, although infrequent, is critical for the establishment of long-term viral reservoirs in patients undergoing antiviral therapy. From our work, it appears that certain regions of the periphery provide such an environment required for reversible silencing, since we have observed that in all cell lines analyzed carrying an integrated latent provirus the localization was close to the nuclear periphery. It has been shown that the provirus in latently infected cells is found both in gene poor and in gene-rich regions (reviewed in (Marcello, 2006)) and also within transcribed genes as recently described (De Marco et al., 2008) (Han et al., 2008) (Lenasi et al., 2008). These cases are represented in our panel of cells (see Results for a description of the cell lines). Therefore, it appears that positioning at the nuclear periphery is dominant over the site of integration in determining the latent/inducible state.

We conclude that the silent provirus is peripheral in latent cells. But what happens to the position of the provirus upon transcriptional activation? To answer this question, we took advantage of the J-lat A1 cells.

The J-lat A1 cellular model

J-lat A1 cells carry a transcriptionally silent HIV-1 proviral construct that can be induced by phorbol esters (TPA).

RT PCR analysis of viral transcripts showed that a basal level of transcription was still present in non-induced cells and that the signal increased in time after TPA treatment (Figure 23). These data were confirmed by cytometric analysis, where we were able to observe a very low level of GFP expression in uninduced cells (0.3%), compared with the 86% obtained for cells following TPA treatment (Figure 21).

Immunoprecipitation analysis showed that the expression of the Tat protein and of the associated co-factor Cyclin T1 was present only in induced cells (Figure 22). This is consistent with the availability of sub-optimal levels of Tat in latent cells from the few transcripts observed by RT PCR. Since the J-lat A1 cell clone can be strongly activated by phorbol esters treatment, this cell line represents an excellent model to study HIV-1 transcriptional silencing and reactivation.

Gene positioning and transcription activation

Increasing evidence indicates that gene activation and gene silencing are often associated with repositioning of the locus relative to the nuclear compartments and other genomic loci. Gene activation has been correlated with movement away from the periphery during lymphocyte development (Kosak et al., 2002) and during the neuronal induction of a given locus (Williams et al., 2006a). However, exceptions to this simple model have been found, in which gene position is not necessarily altered by induced

expression (Hewitt et al., 2004) (Meaburn and Misteli, 2008), suggesting that different parts of the nuclear periphery have distinct roles in transcriptional regulation.

Since gene position may be involved in regulation of gene expression we investigated the localization of HIV-1 in transcribing cells. Our results demonstrated that upon transcriptional activation the HIV-1 provirus remained associated with the nuclear periphery (Figure 19B). Although numerous studies have associated the nuclear periphery with silent chromatin (reviewed in (Lanctot et al., 2007)), our analysis clearly demonstrated that the transcribing provirus is localized at the nuclear periphery in most of the cases (active and silent).

Transcriptional activation has been described as a multistep process initiated by sequence-specific proteins binding to regulatory sites of DNA upstream of the transcribed regions of a gene (Thanos and Maniatis, 1995). Formation of the enhanceosome results in the cooperative binding of accessory chromatin-remodeling proteins, nuclear coactivators, kinases, and/or histone acetylases into the nucleoprotein complex. The presence of chromatin-remodeling factors induces formation of an activated preinitiation complex that, in turn, controls DNA-dependent RNA polymerase II activity and subsequent gene expression (Tian and Brasier, 2003). As expected, also HIV-1 transcription leads to local changes that consistently modify chromatin at the site of HIV-1 integration, including histone acetylation, the removal of negative factors, the recruitment of positive factors and the phosphorylation of RNA polymerase (Lusic et al., 2003) (Marcello et al., 2003). In addition, HIV-1 DNA adopts a loop conformation associated with transcriptional activation (Perkins et al., 2008). In our hands all these changes during transcription activation appear to occur at the

integrated provirus without a significant modification of the spatial position of the locus. Furthermore, we could also show continuous transcription at that location in single living cells where nascent RNA could be visualized, providing the first example of naturally transcribing genetic locus, albeit of viral origin, at the nuclear periphery (Figure 25C).

So far HIV-1 is located at the nuclear periphery independently of the transcription state, but which are the mechanisms governing its transcriptional regulation? To address this issue we decided to investigate the surroundings of the provirus in transcriptionally silent and active J-lat A1 cells.

HIV-1 chromatin conformation characterization

Movement of genetic loci has been observed within the nucleus upon transcriptional activation. Specific interactions have also been observed between loci present on different chromosomes. The overall picture that is emerging is that genes have some independence in the nucleus but do no necessarily function in isolation from each other. They are taken to, or find themselves in, different nuclear environments, which are often shared by other genes. In particular multiple active genes and gene cluster are often located together at places in the nucleus that have high local concentrations of transcriptional and mRNA processing machinery.

In the original work by Jordan (2003), HIV-1 post-integrative latency correlated with integration in close proximity to alphoid repetitive DNA in *cis*. However, the integrated provirus was not associated with these repetitive sequences (Jordan et al., 2003). We were interested in the characterization of the provirus surroundings at the nuclear level.

By 4C analysis we found that transcriptionally silent HIV-1 provirus carried in J-lat A1 cells is associated with a pericentromeric region of chromosome 12 (Figure 29). Since heterochromatin and periphery localization have been related with silencing, we were curious to investigated whether this pericentromeric region of Chr12 associated with HIV-1 in the silent cells was also present in transcribing cells. By 3C we were able to confirm the association of these two loci in latent J-lat A1 cells. However, in the case of activated cells, the interaction of the two loci dropped significantly. These results strongly suggest that the *trans* association between the pericentromeric region of Ch12 and HIV-1 is involved in the silencing of HIV-1, at least in this cell line. It is possible that by integrating into a region involved in a specific cellular pathway that requires inter- or intra-chromosomal interaction to regulate transcription, HIV-1 becomes regulated in the same manner as has been proposed for cellular genes (Fraser and Bickmore, 2007) (Misteli, 2007). Moreover, heterochromatin is associated with specialized chromosome structures, such as centromeres and telomeres, which define transcriptionally inactive nuclear domains. The silencing effect of heterochromatin is not restricted to the region packaged into heterochromatin itself but also to neighboring DNA sequences providing an environment of condensed chromatin enriched in specific histone modifications and heterochromatin proteins (Dillon and Festenstein, 2002). Therefore, heterochromatin centered at the chromosome 12 centromere might extend its repression also in *trans* involving closely associated chromatin loci such as ChXp21 harboring the integrated provirus.

An important consideration in the interpretation of the 3C data is the understanding that not all pairs of restriction fragments that provide a positive result (in this case a novel PCR product) are necessarily engaged in a functional interaction in the nucleus.

The reason for this is that loci along a chromatin fiber will also randomly, and quite frequently collide as the result of the inherent flexibility of chromatin (Dekker et al., 2002). Clearly, different fragments can be cross-linked by formaldehyde simply because they are near each other in the nucleus, and presumably can "bump into" each other during the fixation process (Osborne et al., 2004). Therefore fixation conditions are critical in the 3C assay since increased fixation leads to greater cross-linking resulting in the detection of chromatin fragments that may be in proximity in the nucleus but not necessarily engaged in a specific intermolecular functional interaction. For this reason we decided to do two different controls for the interaction between HIV-1 and Ch12 that we detected in latent J-lat A1 by 4C and 3C methods. First, samples without cross-linking and without ligation were processed for 3C. To check the DNA quality we performed PCR using internals primer on HIV-1 (Table 2) (Figure 27) (Figure 30A, bottom panel). In the case of the controls, as we expected, neither of the amplifications resulted positive, indicating that in fact our interaction is specific for this cell type (Figure 30A, top panel).

We conclude that in J-lat A1 quiescent cells, the provirus integrated in ChXp21.1 interacts physically with Ch12q12. Upon activation of proviral DNA with TPA, this interaction is lost, suggesting its possible role in HIV-1 silencing.

3D FISH and heterochromatin associations

Next, we sought to visualize by microscopy techniques how was the disposition of the loci found to be associated by the 3C assays. Using BACs probes to cover the region where the provirus is integrated (ChX) and the region in Ch12 that was physically

associated with it, we were able to observe by 3D FISH that in fact Ch12 and ChX in J-lat A1 were very close between each other in a significant fraction of cells. An accurate statistical analysis of the minimal distance between the two loci demonstrated the presence of two discrete populations, represented by a Gaussian curve fitted for the distances between the provirus and Ch12 (Figure 31B). Through this approach we were able to calculate the frequency of interaction, which resulted in approximately 10% of the associations centered around 0.6 μm. In this calculation, we included both loci that colocalized and loci that were within a distance of 0.6 μm, considering that the mean radius of confinement of a genomic locus in a mammalian nucleus is between 0.5 μm (Spector, 2003) (Lanctot et al., 2007) and 0.3 μm for more constrained peripheral loci (Chubb et al., 2002). In contrast with this result, 3D FISH experiments for induced cells shown a different behavior. Indeed, upon transcriptional activation of the provirus by TPA, the interaction of the two loci decreased significantly, confirming the data obtained by 3C.

In general, frequency of random collision is inversely related to the genomic distances between loci, consistent with theoretical predictions (Dekker et al., 2002). Thus, relatively frequent but nonfunctional interactions should always be observed for loci separated by small distances. For sites separated by larger genomic distances, this background signal decreases rapidly. Indeed a way to identify unspecific interaction is by mapping in regions increasingly further away along the site of association, for Dekker and collaborators unspecific interactions remained detectable for sites separated by as much as 150 Kilobases (Dekker et al., 2002). Detection of an interaction between loci of different chromosomes is generally specific, however we wanted to confirm our interaction with other probes. In this way, we could evaluate the specificity of our

association and at the same time establish the regions involved in the HIV-1 silencing. To this end, we decided to use as a probe a pericentromeric satellite alphoid DNA repeats specific for Ch12 (Baldini et al., 1990), which were located approximately 2,3 megabase (2300 Kilobase) from our previously found associated region. Also in this case, we observed that the trans association between ChX and Chr12 was more frequent in silent cells, comparing with induced cells. Finally, we evaluated the localization of the provirus, using a specific probe for it, with respect to the Ch12 α-satellite. These experiments showed once more that the HIV-1: Ch12 α-satellite association was found at a high frequency in silent cells. In addition, the association of these two loci occurs at the nuclear periphery. When cells are activated, HIV-1 remains at this position (Figure 33B) (independent of the transcription state and independent of its association with Ch12) and is not found more associated with Ch12, suggesting that the latter is moving according with the transcription state and that this interaction is actually involved in HIV-1 silencing. It is important to note that peripheral localization of the region of interaction of chromosome territories corresponding to Ch12 and ChX has been also observed by others in both human primary fibroblasts (Bolzer et al., 2005) (Guelen et al., 2008) and lymphocytes (Kim et al., 2004) (Parada et al., 2004).

It has been shown previously that recruitment of chromatin domains to centromeric heterochromatin in *trans* imposes silencing in *D. melanogaster* and differentiating mouse lymphocytes (Csink and Henikoff, 1996) (Dernburg et al., 1996) (Brown et al., 1997) (Brown et al., 1999) (Brown et al., 2001) (Merkenschlager et al., 2004). Heterochromatin markers, such as Suv39H1, HP1 and histone H3Lys9 trimethylation, have an important function in chromatin-mediated repression of integrated HIV-1 gene expression being reversibly associated with HIV-1 in a transcription-dependent manner

(Marban et al., 2007) (du Chene et al., 2007) (Mateescu et al., 2008). Interestingly, reactivation of HIV-1 in J-lat A1 cells by siRNA knockdown of HP1α and HP1γ has been shown (du Chene et al., 2007). All these observations point to an involvement of heterochromatin in the silencing of HIV-1. However, a paradox emerges from our observations since the provirus was found in the proximity of the centromere of chromosome 12 in only 10% of the cells when in fact all cells were silenced. Two explanations could be proposed: either within the same clonal population, the provirus associates with different heterochromatin environments that went undetected in our 4C approach, or two or more states of silencing exist, ranging from one closely associated with heterochromatin, to one possibly more poised for transcription. Other interactions are unlikely since we found that the provirus interact with Ch12 and with regions marked by pan α satellite probe at the same frequency in J-lat A1 cells (Figure 35). Furthermore, to expand this observation, we extended the analysis of the association of the provirus with heterochromatin in all cells tested. By this approach we observed that independently of the integration site, all latent cell lines presented different percentages of association with heterochromatin, any of which show values over 45%, even though all cell clones were totally silent (Figure 35). However, one possibility could be an under estimation of the total heterochromatin, since nuclear heterochromatin is represented not only by centromers, but also by telomeres and other chromosomes regions. Regarding the hypothesis that two or more states of silencing exist, we suppose that silencing is a very dynamic process that could be achieved in different steps. The first one could involve the assembly of a pre-initiation complex with a stalled RNAPII that requires further signals to elongate, and we believe that the association of HIV-1 with Ch12 occurs at this step. We called this state off-off (Figure 37). A second step (called

off) could involve the movement of heterochromatin or certain chromosomes regions far away from the transcription unit, and this state should be more poised for transcription. Indeed, it has been shown that RNAPII elongation complexes initiating from the viral LTR can prematurely terminate transcription in the absence of Tat, also in quiescent lymphocytes from patients (Kao et al., 1987) (Lassen et al., 2004). Furthermore, experiments involving bulk analysis of cells, such as, for example, chromatin immunoprecipitation, detected both HP1β and RNA polymerase on the viral LTR in the latent state (Mateescu et al., 2008). This observation could be explained if, in the ensemble population of cells that is pooled in a ChIP experiment, some cells are in one status and some in the other. Finally, in a third step, the interaction between HIV-1 and Ch12 is probably totally lost and HIV-1 RNA is transcribed, possibly within a single transcription factory, without changing its localization to the nuclear periphery (Figure 37). We called this last step 'on'.

Lamin-associated domains (LADs) represent a chromatin repressive environment tightly associated with the nuclear periphery. Recently Bas Stenseel (Guelen et al., 2008) reported a high-resolution map of the interaction sites of the entire human genome with LADs. In addition, Frederic Bushman and collaborators (Lewinski et al., 2005) have generated two libraries of HIV-1 integrations sites. One set of cells carried a proviral vector that is well expressing; the other cells expressed the provirus poorly, but could be reactivated by TNFα. To establish whether there was a correlation between HIV-1 integrations sites and LADs, we calculated how many integration sites are within LAD domains. From these analyses, we observed that for poorly and constitutively expressed HIV-1, integrations within LADs were equally disfavored, showing approximately 15 % of associations with them, while random generated integration sites

were associated with LADs in 40 % of the cases (Figure 36). Consequently, HIV-1 integration sites are not favored in regions of repressive chromatin in tight conjunction with the nuclear lamina. These results validate the observation that HIV-1 integrates preferentially in regions that are transcriptionally active (Schroder et al., 2002) and not constitutively repressed like LADs.

Nearly 40% of the human genome consists of LADs (Guelen et al., 2008). Compactions of chromatin at the nuclear periphery allow a relatively large fraction of the genome to be located at the lamina. Assuming that the nucleus of a lymphocyte can be approximated to a sphere, the periphery represents 50% of the total nuclear volume (considering that in our 3D FISH experiments the integration sites were within the outer 20% of the nuclear radius). If 40% of random interactions were associated with LADs it results that 10% of the periphery non-associated with LADs could be available for integration of HIV -1 at this location. It would be interesting to evaluate whether this 10% of integrations are random at the periphery or are represented by "hot spots" sites of integration. Nevertheless, the picture that is emerging is that HIV-1 integrates preferentially at the nuclear periphery within active genes that are not associated to the nuclear lamina.

This work is not only providing a new concept about the three-dimensional conformation of chromatin and its implications in HIV-1 transcription and silencing, but also presents an insight in how chromatin organization could modulate the expression levels of a given gene. From our results and others studies (Ragoczy et al., 2006) (Meaburn and Misteli, 2008), it seems that internal positioning is not a requirement for activity and that transcription alone does not always drive the position

of a gene, suggesting that radial positioning is not necessarily linked to gene activity. Evidently, radial gene positioning is affected by multiple components, and changes in the locus positions vary between different genes, ruling out difficulties to deduce universal rules.

To conclude, we propose that HIV-1 integrated close to pericentromeric heterochromatin at the nuclear periphery in quiescent lymphocytes finds a convenient environment to maintain a silent but inducible state. However, it would be important to extend these observations to a wider population of latent cells from infected patients undergoing antiviral therapy to get a general picture of the phenomenon.

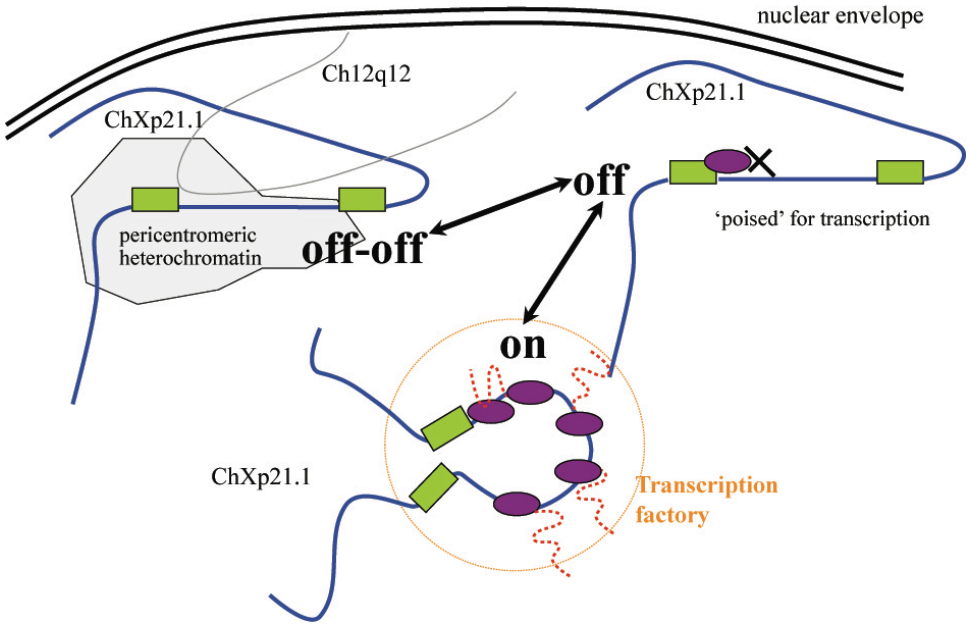


Figure 37. Schematic drawing that summarizes the concepts emerging from the experimental data.

In J-lat A1 cells, the provirus is found integrated in the X chromosome and localized at the nuclear periphery. Although all cells harbour a transcriptionally silent provirus, this could be found spatially associated close to the centromere of Chromosome 12 in a fraction of the nuclei. We identify two states of the silenced provirus: one associated with pericentromeric heterochromatin of chromosome 12 at the nuclear periphery (off-off) and one not associated with chromosome 12 that may be poised for transcription (off). On activation, this interaction is lost and HIV-1 RNA is transcribed, possibly within a transcription factory, without changing its localization to the nuclear periphery (on).

Materials and Methods

Cells and their characterization

J-lat cell lines, derivatives of the Jurkat cell line originally developed in Eric Verdin's laboratory, were obtained through the NIH AIDS Research and Reference Reagent Program. The promonocytic U937 cell line U1, which contains two copies of the integrated provirus (Lusic et al., 2003) (Folks et al., 1987), and the HOS 143b cells (ECACC n. 91112502) carrying an integrated HIV-1 vector (De Marco et al., 2008) were described previously (Table1). Cells were grown in RPMI 1640, 10% FCS and antibiotics. As expected for tumor cell lines in culture, that are frequently subject to chromosomal abnormalities and polyploidy, we observed that J-lat A1 cells were often polyploid with loss of the original Y chromosome (Schneider et al., 1977). Cytofluorimetric analysis was performed on untreated cells or cells treated with TPA and analyzed after 15 hours. RT-PCR was performed to asses viral transcription levels. Cells were harvested at different time points by trypsinisation. Total RNA was extracted using RNAeasy Mini Kit (Qiagen), treated with DNAaseI (Invitrogen) and quantified. Reverse transcription was performed with M-MLV RT (Invitrogen) using random primers. Primers used for PCR amplification are described in Table 2.

Integration site confirmation

To confirm the integration site of HIV-1 in J-lat A1 cells we performance PCR assays. Thus, genome DNA was extracted from J-lat A1 using DNeasy Blood & Tissue Kit (Cat No 69504) following the manufacture's instructions. DNA samples were then

amplified by PCR using primers designed in the region adjacent to the site of proviral integration and on the provirus. The primers used are shown in Table 2.

The samples were amplified using Go taq Polimerase (Promega). The PCR conditions were as follow: 94 °C for 1 min followed by 35 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s followed by one cycle of 72°C for 2 min.

Western blot analysis and antibodies

Whole cell lysates extracts were resolved by 10% or 12% SDS-polyacrylamide gel electrophoresis (SDS-PAGE). For Western blotting, nitrocellulose membrane (Reinforced NC, Whatman) was used and membranes were blocked for 1hr in 5% milk followed by the incubation with the appropriate primary antibody diluted in 3% milk/0.5% Tween-20 for 1-2 hrs. After three washings with TBS 0.5% Tween-20, secondary antibodies conjugated with HRP (DakoCytomation) were diluted in 3% milk/0.5% Tween-20 and incubated for 1 hr. Blots were developed using Immobilon Western Chemiluminescent HRP Substrate (Millipore) according to the manufacturer's instructions.

Immunoblots were performed with the following antibodies: flag (Sigma, F1804, 1:1000) and Cyclin T1 (Santa Cruz, H245).

Fluorescent in situ hybridization (FISH)

3D FISH was performed essentially as described in (Solovei et al., 2006). Briefly, freshly grown J-lat A1 cells were resuspended at 3×10^6 cells/ml in 50% serum and

seeded for 20 minutes on glass coverslips previously treated with polylysine. Cells were then fixed in 4% paraformaldehyde in 0.3× hypotonic PBS for 10 minutes, permeabilized with Triton X-100 for 10 minutes and leaved in PBS with glycerol 20% from 1 hour to O/N. The next day the coverslips with the cells were subjected to three cycles of liquid nitrogen in glycerol, treated with 0.1N HCl for 10 minutes and then incubated in 0,002% of Pepsin in 0,01N HCL at 37 °C for 8 minutes, then washed in PBS followed by RNAse A (200μg/ml) treatment. Coverslips were then equilibrated in 50% formamide, SSC 2× for 1-2 hours.

BAC DNA containing the regions under study, ChXp21.1 and Ch12q12 (RP11-77013 and RP11 379CZ4 respectively) and the DNA probe specific for Ch12 α-satellite repeats (Baldini et al., 1990) were obtained from Dr Mario Rocchi at the University of Bari, Italy. The pan-satellite repeats probe p82H that is specific for the centromeric regions of all human chromosomes (Mitchell et al., 1985) was obtained from Wendy Bickmore. Probes were conjugated to digoxigen-11-dUTP and labeled with α-Digoxigenin Fluorescein (Roche) or Tetramethyl-rhodamine-5-dUTP (Roche). The provirus was conjugated to Biotin-16-dUTP (Roche) and labeled with the TSA Cyanine 5 System (Perkin Elmer). The incorporation of modified dUTP's for all samples was obtained by nick translation (Roche), unincorporated nucleotides were removed by gel filtration through a MicroSpin G-25 Column (GE Healhtcare). The probes were then resuspended into the hybridization mix and applied to the slides. The coverslips were immobilized with rubber cement and leaved to hybridization at 37 °C in a humid chamber for at least 2 days. Washes and antibody staining were performed using standard techniques. Coverslips were then mounted in Vectashield (Vector Labs).

Fluorescent images of fixed cells were captured on a Zeiss LSM 510 META confocal microscope (Carl Zeiss Microimaging, Inc.) with a 63× NA=1.4 Plan-Apochromat oil objective. The pinhole of the microscope was adjusted to get an optical slice of less than 1.0 μm for any wavelength acquired. Samples labeled with Digoxigenin-11-dUTP were excited with the 488 nm line of the Ar laser and its emission was monitored using a custom-made Meta band pass filter between 510 and 563 nm. Biotin/TSA-Cy5 System and Tetramethyl-rhodamine-5-dUTP fluorophores were excited with the 543 nm HeNe laser and their emission collected using a custom made Meta band pass filter between 552 and 670 nm.

Z Stacks of images were analyzed with Zeiss LSM Image Examiner. The shortest distance from the centers of signal intensity between FISH spots (gene locus) or from the spot to the nuclear periphery (defined by a sharp drop in stain) was measured using the Ortho tool of the LSM510 software with the 3D distance display mode. The periphery data were normalized using the maximum diameter of each cell.

In the case of heterochromatin distribution we used pan-α-satellite repeats probe (Mitchell et al., 1985) label with Tetramethyl-rhodamine-5-dUTP (Roche), following the protocol explained before in detail.

Circular chromatin conformation capture (4C)

The 4C protocol was established similarly to published procedures (Zhao et al., 2006).

J-lat A1 cells were maintained in warm fresh medium at 0.5×10^{6} cells/mL. After 1 hour TPA (Fluka) was added to the final concentration of 1.6 μM. After 5-8 hours, 10^{8} cells

were cross-linked on a rocker platform in 40 ml pf PBS with 1% formaldehyde at room temperature for 10 minutes. The reaction was then quenched with 0.125 M glycine and kept on ice. After two washes with cold PBS, the pellet was resuspended in lysis buffer (10 mM Tris pH 8.0, 10 mM NaCl, 0.2% NP40 and freshly added protease inhibitors) and left on ice for 10 minutes. Cell lysis was completed with 10 strokes of Dounce homogenizer (pestle B, Wheaton). Nuclei were pellet at 600g for 8 minutes and washed with 10 ml of 1× restriction buffer (New England Biolabs, NEB2). Nuclei were re-suspended in 712.8 μl of restriction buffer including 0.1% SDS, and were incubated for 1 hour at 37 °C with shaking. Triton-X was then added to a final concentration of 1.8% and further incubated at 37 °C for 1 h before adding 4000 units (U) of HindIII. The reaction was incubated overnight at 37 °C with shaking. The restriction enzyme was inactivated by the addition of 2% SDS (final concentration) and was incubated at 65 °C for 20 minutes. The digestion mixture was diluted in 8 ml of 1× T4 ligase buffer (New England Biolabs) including 1% triton X-100, and was incubated at 37 °C for 1 hour. The digested sample was ligated using 4000 U of T4 DNA ligase (New England Biolabs) for 3 days at 4 °C, followed by incubation at 16 °C for 5 hours and 1 hour at RT with additional enzyme and ATP. The reaction was stopped with 10 mM EDTA. Then 1 μg/ml of RNAse was added and incubated 30' at 37 °C followed by 0.1 mg/ml protein kinase and incubated at 56°C for 30 minutes. Finally, cross-links were reversed by incubation overnight at 65 °C. The DNA was purified by 2× phenol:chlorophorm, 2× chlorophorm extraction and ethanol precipitation.

In some experiments, the purified DNA was digested with a restriction enzyme (Not I) that cuts between the inverse PCR primers in order to eliminate products from self-ligated templates. The DNA samples were then amplified by nested PCR. The high

fidelity AccuPrime Pfx (Invitrogen) was used for first and second round amplification. The PCR conditions were as follow: 94 °C for 2 min followed by 35 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s followed by one cycle of 72°C for 2 min. For the second round the PCR conditions were as follow: 94 °C for 2 min followed by 35 cycles of 94°C for 30 s, 62°C for 30 s, and 72°C for 30s followed by one cycle of 72°C for 2 min. The primers used are shown in Table 2.

The PCR products were either resolved on agarose gels, and the fragments cloned individually, or bulk cloned into Zero Blunt TOPO PCR Cloning Kit (Invitrogen) and analyzed by DNA sequencing.

In order to confirm the interaction, we performance 3C assays following the protocol of 4C technique described previously. With a primer on Ch12q12 (found by 4C) and the primers on the provirus a hemi-nested PCR was done following the same conditions as in 4C PCR. The primers used are shown in Table 2. As positive control we used internal amplification of the provirus using Nuc1-B177 and splice-Intro primers

Fluorescence microscopy, immunofluorescence, RNA ISH

Immunofluorescence

Cells prepared in PBS ($3x10^6$/ml), were plated on cover slide previoly treated with Poly-L- Lysine solution (Sigma P8920). Cells treated as described for 3D FISH at the 20% glycerol step were blocked with 5% horse Serum in PBS for 1 hour at 37°C. As primary antibody we used the rabbit antiserum for trimethyl histone H4 Lys20 (H4K20me3)

(#07-463, Upstate). The antibody was diluted 1:500 in 5% Horse Serum and leaved for 2 hr at 37 °C. Anti-rabbit Alexa 594 (Invitrogen) was used as the secondary antibody. The proteins were crosslink with ethylene glycol bis (succinimidyl succinate) (EGS) 0.5M diluted 1:500 (Sigma E -3257) for 10 min at RT. Optical sectioning of fixed cells were captured on a Zeiss LSM 510 META confocal microscope as described above.

RNA in situ hybridization

RNA ISH on HOS_A4 cells has been conducted as described (De Marco et al., 2008). Cells were plated on glass, transiently transfected with lipofectamine using plasmids expressing the Tat transactivator and EYFP-MS2nls and kept for 24 hours at 37°C. Later on, cells were fixed with 3.7% PFA and permeabilized overnight in 70% ethanol. The amino-allyl thymidine modified oligonucleotide probes listed below were courtesy of E. Bertrand (IGMM, Montpellier, France). Probes were labeled with Cy3 or Cy5 (Cyn monoreactive dye, Amersham) as described (Boireau et al., 2007). For quantitative measurements stacks of 21 planes were acquired at bin = 1 with steps of 0.5 μm in the z-axis using a wide-field Leica DMRI inverted microscope (63x objective, NA 1.3) controlled by Metamorph (Universal Imaging). Digital images were collected using a CoolSnap K CCD camera (Roper scientific). The three dimensional deconvolution and reconstruction was performed with the ImageJ plug-in 'Iterative Deconvolve 3D' (http://rsb.info.nih.gov/ij/). The co-localization analysis was performed with the plugin of P. Bourdoncle in ImageJ. Volume rendering was performed with the 'VolumeJ' plugin (Abramoff and Viergever, 2002).

Probes for RNA FISH

HIV_MS2:

AxGTCGACCTGCAGACAxGGGTGATCCTCAxGTTTTCTAGGCAATxA

HIV_SPLICED:

AxGGTTGGGAGGTGGGxCAGTCGCCGCCCCTCxA

RNA ISH in vivo

For live cell experiments, cells were plated on glass-bottom plates (MatTek), transiently transfected with lipofectamine with plasmids expressing the Tat transactivator and EYFP-MS2nls and analyzed next day at 37°C in a 5% CO_2 humidified atmosphere in a non-fluorescent complete DMEM media. Stacks of 21 planes were acquired at bin = 1 with steps of 0.5 μm in the z-axis using a wide-field Leica DMRI inverted microscope (63× objective, NA 1.3) controlled by Metamorph (Universal Imaging) and equipped with a conventional light source (Hg, 100 W), a filter cube for YFP detection (Leica Microsystems) and an automated shutter control to minimize exposure of samples to light (Sutter). Digital images were collected using a CoolSnap K CCD camera (Roper scientific). The three-dimensional deconvolution and reconstruction was performed with the ImageJ plug-in 'Iterative Deconvolve 3D' (http://rsb.info.nih.gov/ij/).

Statistic analysis

The statistical analysis of the data was performed with R (www.r-project.org/).

The distribution of distances between chromosomes (Figure 31B, 3C, 3D) was analyzed with the R MCLUST package (www.stat.washington.edu/mclust/). The parametric

method Bayesian Information Criterion (BIC) was used as a useful statistical criterion for model selection to determine the number of components in a Gaussian mixture model. Value of BIC allowed the distinction between a single Gaussian (Figure 31C, 3D) and the sum of two Gaussian (Figure 31B). The two Gaussian in Figure 31B were centered at 0.6 μm and 2.8 μm, with a standard deviation of 0.5 μm and 1 μm, respectively. Mean and standard deviation of the curve in Figure 31C were 2.7 μm and 1 μm and in Figure 31D were 3.3 μm and 2 μm.

The differences in spatial associations measured in Figure 31E and Figure 32B and D were analyzed with the χ^2 test and were always significative (significance level set at $p < 0.05$).

The differences between the distribution of distances of the provirus to the periphery in induced and not induced cell (Figure 33B) was analyzed with the Kolmogorov-Smirnov test (K-S test) and were not significative (significance level set at $p < 0.01$). The K-S test is a non-parametric and distribution free statistical test that determines if two datasets differ significantly.

The distributions of distances of the provirus from the periphery for different cell lines compared to the J-lat A1 distribution (Figure 19B) were also analyzed with a K-S test (significance level set at $p < 0.01$).

BIBLIOGRAPHY

Abney, J. R., Cutler, B., Fillbach, M. L., Axelrod, D., and Scalettar, B. A. (1997). Chromatin dynamics in interphase nuclei and its implications for nuclear structure. J Cell Biol *137*, 1459-1468.

Abramoff, M. D., and Viergever, M. A. (2002). Computation and visualization of three-dimensional soft tissue motion in the orbit. IEEE Trans Med Imaging *21*, 296-304.

Ainger, K., Avossa, D., Morgan, F., Hill, S. J., Barry, C., Barbarese, E., and Carson, J. H. (1993). Transport and localization of exogenous myelin basic protein mRNA microinjected into oligodendrocytes. J Cell Biol *123*, 431-441.

Albanese, A., Arosio, D., Terreni, M., and Cereseto, A. (2008). HIV-1 pre-integration complexes selectively target decondensed chromatin in the nuclear periphery. PLoS One *3*, e2413.

Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A *51*, 786-794.

Allis, C. D., Chicoine, L. G., Richman, R., and Schulman, I. G. (1985). Deposition-related histone acetylation in micronuclei of conjugating Tetrahymena. Proc Natl Acad Sci U S A *82*, 8048-8052.

Andrulis, E. D., Neiman, A. M., Zappulla, D. C., and Sternglanz, R. (1998). Perinuclear localization of chromatin facilitates transcriptional silencing. Nature *394*, 592-595.

Armstrong, J. A., and Emerson, B. M. (1998). Transcription of chromatin: these are complex times. Curr Opin Genet Dev *8*, 165-172.

Arney, K. L., and Fisher, A. G. (2004). Epigenetic aspects of differentiation. J Cell Sci *117*, 4355-4363.

Attwood, J. T., Yung, R. L., and Richardson, B. C. (2002). DNA methylation and the regulation of gene transcription. Cell Mol Life Sci *59*, 241-257.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., and Fisher, A. G. (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol *8*, 532-538.

Baeuerle, P. A., and Baltimore, D. (1988). Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-kappa B transcription factor. Cell *53*, 211-217.

Baeuerle, P. A., and Baltimore, D. (1996). NF-kappa B: ten years after. Cell *87*, 13-20.

Baeuerle, P. A., and Henkel, T. (1994). Function and activation of NF-kappa B in the immune system. Annu Rev Immunol *12*, 141-179.

Baldini, A., Rocchi, M., Archidiacono, N., Miller, O. J., and Miller, D. A. (1990). A human alpha satellite DNA subset specific for chromosome 12. Am J Hum Genet *46*, 784-788.

Baldwin, A. S. (2001). Control of oncogenesis and cancer therapy resistance by the transcription factor NF-kappaB. J Clin Invest *107*, 241-246.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Barnes, P. J., and Karin, M. (1997). Nuclear factor-kappaB: a pivotal transcription factor in chronic inflammatory diseases. N Engl J Med *336*, 1066-1071.

Barrandon, C., Bonnet, F., Nguyen, V. T., Labas, V., and Bensaude, O. (2007). The transcription-dependent dissociation of P-TEFb-HEXIM1-7SK RNA relies upon formation of hnRNP-7SK RNA complexes. Mol Cell Biol *27*, 6996-7006.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bartova, E., Krejci, J., Harnicarova, A., Galiova, G., and Kozubek, S. (2008). Histone modifications and nuclear architecture: a review. J Histochem Cytochem *56*, 711-721.

Baud, V., and Karin, M. (2001). Signal transduction by tumor necrosis factor and its relatives. Trends Cell Biol *11*, 372-377.

Berger, E. A., Murphy, P. M., and Farber, J. M. (1999). Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. Annu Rev Immunol *17*, 657-700.

Berkhout, B., and Jeang, K. T. (1992). Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. J Virol *66*, 139-149.

Berkowitz, R. D., Hammarskjold, M. L., Helga-Maria, C., Rekosh, D., and Goff, S. P. (1995). *5'* regions of HIV-1 RNAs are not sufficient for encapsidation: implications for the HIV-1 packaging signal. Virology *212*, 718-723.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., and Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. Mol Cell *2*, 437-445.

Bielinska, A., Krasnow, S., and Nabel, G. J. (1989). NF-kappa B-mediated activation of the human immunodeficiency virus enhancer: site of transcriptional initiation is independent of the TATA box. J Virol *63*, 4097-4100.

Bieniasz, P. D., Grdina, T. A., Bogerd, H. P., and Cullen, B. R. (1998). Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. EMBO J *17*, 7056-7065.

Bisgrove, D. A., Mahmoudi, T., Henklein, P., and Verdin, E. (2007). Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. Proc Natl Acad Sci U S A *104*, 13690-13695.

Blank, V., Kourilsky, P., and Israel, A. (1992). NF-kappa B and related proteins: Rel/dorsal homologies meet ankyrin-like repeats. Trends Biochem Sci *17*, 135-140.

Boireau, S., Maiuri, P., Basyuk, E., de la Mata, M., Knezevich, A., Pradet-Balade, B., Backer, V., Kornblihtt, A., Marcello, A., and Bertrand, E. (2007). The transcriptional cycle of HIV-1 in real-time and live cells. J Cell Biol *179*, 291-304.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M. R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol *3*, e157.

Brady, J., and Kashanchi, F. (2005). Tat gets the "green" light on transcription initiation. Retrovirology *2*, 69.

Branco, M. R., Branco, T., Ramirez, F., and Pombo, A. (2008). Changes in chromosome organization during PHA-activation of resting human lymphocytes measured by cryo-FISH. Chromosome Res *16*, 413-426.

Brasey, A., Lopez-Lastra, M., Ohlmann, T., Beerens, N., Berkhout, B., Darlix, J. L., and Sonenberg, N. (2003). The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. J Virol *77*, 3939-3949.

Brown, K. E., Amoils, S., Horn, J. M., Buckle, V. J., Higgs, D. R., Merkenschlager, M., and Fisher, A. G. (2001). Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. Nat Cell Biol *3*, 602-606.

Brown, K. E., Baxter, J., Graf, D., Merkenschlager, M., and Fisher, A. G. (1999). Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. Mol Cell *3*, 207-217.

Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M., and Fisher, A. G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845-854.

Brownell, J. E., and Allis, C. D. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. Curr Opin Genet Dev *6*, 176-184.

Bukrinskaya, A., Brichacek, B., Mann, A., and Stevenson, M. (1998). Establishment of a functional human immunodeficiency virus type 1 (HIV-1) reverse transcription complex involves the cytoskeleton. J Exp Med *188*, 2113-2125.

Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G., and Stevenson, M. (1993). Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc Natl Acad Sci U S A *90*, 6125-6129.

Bukrinsky, M. I., Stanwick, T. L., Dempsey, M. P., and Stevenson, M. (1991). Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection. Science *254*, 423-427.

Campbell, S., and Rein, A. (1999). In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J Virol *73*, 2270-2279.

Casolari, J. M., Brown, C. R., Komili, S., West, J., Hieronymus, H., and Silver, P. A. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell *117*, 427-439.

Chambeyron, S., and Bickmore, W. A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes Dev *18*, 1119-1130.

Chuang, C. H., Carpenter, A. E., Fuchsova, B., Johnson, T., de Lanerolle, P., and Belmont, A. S. (2006). Long-range directional movement of an interphase chromosome site. Curr Biol *16*, 825-831.

Chubb, J. R., Boyle, S., Perry, P., and Bickmore, W. A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol *12*, 439-445.

Chun, T. W., and Fauci, A. S. (1999). Latent reservoirs of HIV: obstacles to the eradication of virus. Proc Natl Acad Sci U S A *96*, 10958-10961.

Cimbora, D. M., and Groudine, M. (2001). The control of mammalian DNA replication: a brief history of space and timing. Cell *104*, 643-646.

Clark-Adams, C. D., Norris, D., Osley, M. A., Fassler, J. S., and Winston, F. (1988). Changes in histone gene dosage alter transcription in yeast. Genes Dev *2*, 150-159.

Clever, J., Sassetti, C., and Parslow, T. G. (1995). RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. J Virol *69*, 2101-2109.

Contreras, X., Lenasi, T., and Peterlin, B. M. (2006). HIV latency: present knowledge, future directions. Future Virol *1*, 733-745.

Cook, P. R. (1999). The organization of replication and transcription. Science *284*, 1790-1795.

Cook, P. R. (2003). Nongenic transcription, gene regulation and action at a distance. J Cell Sci *116*, 4483-4491.

Cremer, T., and Cremer, C. (2006). Rise, fall and resurrection of chromosome territories: a historical perspective. Part II. Fall and resurrection of chromosome territories during the 1950s to 1980s. Part III. Chromosome territories and the functional nuclear architecture: experiments and models from the 1990s to the present. Eur J Histochem *50*, 223-272.

Cremer, T., Cremer, C., Baumann, H., Luedtke, E. K., Sperling, K., Teuber, V., and Zorn, C. (1982). Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. Hum Genet *60*, 46-56.

Cremer, T., Kreth, G., Koester, H., Fink, R. H., Heintzmann, R., Cremer, M., Solovei, I., Zink, D., and Cremer, C. (2000). Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. Crit Rev Eukaryot Gene Expr *10*, 179-212.

Cremer, T., Lichter, P., Borden, J., Ward, D. C., and Manuelidis, L. (1988). Detection of chromosome aberrations in metaphase and interphase tumor cells by in situ hybridization using chromosome-specific library probes. Hum Genet *80*, 235-246.

Croft, J. A., Bridger, J. M., Boyle, S., Perry, P., Teague, P., and Bickmore, W. A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. J Cell Biol *145*, 1119-1131.

Cron, R. Q., Bartz, S. R., Clausell, A., Bort, S. J., Klebanoff, S. J., and Lewis, D. B. (2000). NFAT1 enhances HIV-1 gene expression in primary human CD4 T cells. Clin Immunol *94*, 179-191.

Csink, A. K., and Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organization in Drosophila. Nature *381*, 529-531.

Csordas, A. (1990). On the biological role of histone acetylation. Biochem J *265*, 23-38.

Cujec, T. P., Cho, H., Maldonado, E., Meyer, J., Reinberg, D., and Peterlin, B. M. (1997). The human immunodeficiency virus transactivator Tat interacts with the RNA polymerase II holoenzyme. Mol Cell Biol *17*, 1817-1823.

Cullen, B. R. (1998). Retroviruses as model systems for the study of nuclear RNA export pathways. Virology *249*, 203-210.

Darlix, J. L., Gabus, C., Nugeyre, M. T., Clavel, F., and Barre-Sinoussi, F. (1990). Cis elements and trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. J Mol Biol *216*, 689-699.

Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S. M., Phair, R. D., and Singer, R. H. (2007). In vivo dynamics of RNA polymerase II transcription. Nat Struct Mol Biol *14*, 796-806.

De Marco, A., Biancotto, C., Knezevich, A., Maiuri, P., Vardabasso, C., and Marcello, A. (2008). Intragenic transcriptional cis-activation of the human immunodeficiency virus 1 does not result in allele-specific inhibition of the endogenous gene. Retrovirology *5*, 98.

Deichmann, M., Bentz, M., and Haas, R. (1997). Ultra-sensitive FISH is a useful tool for studying chronic HIV-1 infection. J Virol Methods *65*, 19-25.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Deng, L., Ammosova, T., Pumfery, A., Kashanchi, F., and Nekhai, S. (2002). HIV-1 Tat interaction with RNA polymerase II C-terminal domain (CTD) and a dynamic association with CDK2 induce CTD phosphorylation and transcription from HIV-1 promoter. J Biol Chem *277*, 33922-33929.

Dernburg, A. F., Broman, K. W., Fung, J. C., Marshall, W. F., Philips, J., Agard, D. A., and Sedat, J. W. (1996). Perturbation of nuclear architecture by long-distance chromosome interactions. Cell *85*, 745-759.

Dietzel, S., Zolghadr, K., Hepperger, C., and Belmont, A. S. (2004). Differential large-scale chromatin compaction and intranuclear positioning of transcribed versus non-transcribed transgene arrays containing beta-globin regulatory sequences. J Cell Sci *117*, 4603-4614.

Dieudonne, M., Maiuri, P., Biancotto, C., Knezevich, A., Kula, A., Lusic, M., and Marcello, A. (2009). Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery. EMBO J.

Dillon, N., and Festenstein, R. (2002). Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. Trends Genet *18*, 252-258.

Dinter, H., Chiu, R., Imagawa, M., Karin, M., and Jones, K. A. (1987). In vitro activation of the HIV-1 enhancer in extracts from cells treated with a phorbol ester tumor promoter. EMBO J *6*, 4067-4071.

Dorman, N., and Lever, A. (2000). Comparison of viral genomic RNA sorting mechanisms in human immunodeficiency virus type 1 (HIV-1), HIV-2, and Moloney murine leukemia virus. J Virol *74*, 11413-11417.

Dostie, J., and Dekker, J. (2007). Mapping networks of physical interactions between genomic elements using 5C technology. Nat Protoc *2*, 988-1002.

du Chene, I., Basyuk, E., Lin, Y. L., Triboulet, R., Knezevich, A., Chable-Bessia, C., Mettling, C., Baillat, V., Reynes, J., Corbeau, P.*, et al.* (2007). Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. EMBO J *26*, 424-435.

Dunaway, M., and Ostrander, E. A. (1993). Local domains of supercoiling activate a eukaryotic promoter in vivo. Nature *361*, 746-748.

Durrin, L. K., Mann, R. K., Kayne, P. S., and Grunstein, M. (1991). Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. Cell *65*, 1023-1031.

El Kharroubi, A., Piras, G., Zensen, R., and Martin, M. A. (1998). Transcriptional activation of the integrated chromatin-associated human immunodeficiency virus type 1 promoter. Mol Cell Biol *18*, 2535-2544.

Emiliani, S., Fischle, W., Ott, M., Van Lint, C., Amella, C. A., and Verdin, E. (1998). Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. J Virol *72*, 1666-1670.

Eszterhas, S. K., Bouhassira, E. E., Martin, D. I., and Fiering, S. (2002). Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. Mol Cell Biol *22*, 469-479.

Farnet, C. M., and Haseltine, W. A. (1991). Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. J Virol *65*, 1910-1915.

Fassati, A. (2006). HIV infection of non-dividing cells: a divisive problem. Retrovirology *3*, 74.

Ferreira, J., Paolella, G., Ramos, C., and Lamond, A. I. (1997). Spatial organization of large-scale chromatin domains in the nucleus: a magnified view of single chromosome territories. J Cell Biol *139*, 1597-1610.

Finlan, L. E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J. R., and Bickmore, W. A. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. PLoS Genet *4*, e1000039.

Flaus, A., and Owen-Hughes, T. (2004). Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? Curr Opin Genet Dev *14*, 165-173.

Folks, T. M., Justement, J., Kinter, A., Dinarello, C. A., and Fauci, A. S. (1987). Cytokine-induced expression of HIV-1 in a chronically infected promonocyte cell line. Science *238*, 800-802.

Fraser, P., and Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. Nature *447*, 413-417.

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S. M., Blanchard, J. M., Singer, R. H., and Bertrand, E. (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. Curr Biol *13*, 161-167.

Ganser-Pornillos, B. K., Yeager, M., and Sundquist, W. I. (2008). The structural biology of HIV assembly. Curr Opin Struct Biol *18*, 203-217.

Garcea, R. L., and Alberts, B. M. (1980). Comparative studies of histone acetylation in nucleosomes, nuclei, and intact cells. Evidence for special factors which modify acetylase action. J Biol Chem *255*, 11454-11463.

Gerritsen, M. E., Williams, A. J., Neish, A. S., Moore, S., Shi, Y., and Collins, T. (1997). CREB-binding protein/p300 are transcriptional coactivators of p65. Proc Natl Acad Sci U S A *94*, 2927-2932.

Gheysen, D., Jacobs, E., de Foresta, F., Thiriart, C., Francotte, M., Thines, D., and De Wilde, M. (1989). Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. Cell *59*, 103-112.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N. P., and Bickmore, W. A. (2004). Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell *118*, 555-566.

Gilbert, N., Boyle, S., Sutherland, H., de Las Heras, J., Allan, J., Jenuwein, T., and Bickmore, W. A. (2003). Formation of facultative heterochromatin in the absence of HP1. EMBO J *22*, 5540-5550.

Gilmore, T. D. (2006). Introduction to NF-kappaB: players, pathways, perspectives. Oncogene *25*, 6680-6684.

Greatorex, J., and Lever, A. (1998). Retroviral RNA dimer linkage. J Gen Virol *79 (Pt 12)*, 2877-2882.

Greger, I. H., Demarchi, F., Giacca, M., and Proudfoot, N. J. (1998). Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. Nucleic Acids Res *26*, 1294-1301.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W., and van Steensel, B. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948-951.

Guo, K., Wang, J., Andres, V., Smith, R. C., and Walsh, K. (1995). MyoD-induced expression of p21 inhibits cyclin-dependent kinase activity upon myocyte terminal differentiation. Mol Cell Biol *15*, 3823-3829.

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. Cell *55*, 1137-1145.

Han, Y., Lassen, K., Monie, D., Sedaghat, A. R., Shimoji, S., Liu, X., Pierson, T. C., Margolick, J. B., Siliciano, R. F., and Siliciano, J. D. (2004). 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 *78*, 6122-6133.

Han, Y., Lin, Y. B., An, W., Xu, J., Yang, H. C., O'Connell, K., Dordai, D., Boeke, J. D., Siliciano, J. D., and Siliciano, R. F. (2008). Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. Cell Host Microbe *4*, 134-146.

Harrich, D., Garcia, J., Mitsuyasu, R., and Gaynor, R. (1990). TAR independent activation of the human immunodeficiency virus in phorbol ester stimulated T lymphocytes. EMBO J *9*, 4417-4423.

Haseltine, W. A. (1991). Molecular biology of the human immunodeficiency virus type 1. FASEB J *5*, 2349-2360.

Havas, K., Whitehouse, I., and Owen-Hughes, T. (2001). ATP-dependent chromatin remodeling activities. Cell Mol Life Sci *58*, 673-682.

He, G., and Margolis, D. M. (2002). Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. Mol Cell Biol *22*, 2965-2973.

Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J *7*, 1395-1402.

Henderson, S. L., Ryan, K., and Sollner-Webb, B. (1989). The promoter-proximal rDNA terminator augments initiation by preventing disruption of the stable transcription complex caused by polymerase read-in. Genes Dev *3*, 212-223.

Henkel, T., Machleidt, T., Alkalay, I., Kronke, M., Ben-Neriah, Y., and Baeuerle, P. A. (1993). Rapid proteolysis of I kappa B-alpha is necessary for activation of transcription factor NF-kappa B. Nature *365*, 182-185.

Hewitt, S. L., High, F. A., Reiner, S. L., Fisher, A. G., and Merkenschlager, M. (2004). Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. Eur J Immunol *34*, 3604-3613.

Ho, D. D. (1998). Toward HIV eradication or remission: the tasks ahead. Science *280*, 1866-1867.

Ho, D. D., and Bieniasz, P. D. (2008). HIV-1 at 25. Cell *133*, 561-565.

Huang, J., Wang, F., Argyris, E., Chen, K., Liang, Z., Tian, H., Huang, W., Squires, K., Verlinghieri, G., and Zhang, H. (2007). Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. Nat Med *13*, 1241-1247.

Huxford, T., Huang, D. B., Malek, S., and Ghosh, G. (1998). The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. Cell *95*, 759-770.

Iborra, F. J., Pombo, A., Jackson, D. A., and Cook, P. R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J Cell Sci *109 (Pt 6)*, 1427-1436.

Jackson, D. A., Iborra, F. J., Manders, E. M., and Cook, P. R. (1998). Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. Mol Biol Cell *9*, 1523-1536.

Janicki, S. M., Tsukamoto, T., Salghetti, S. E., Tansey, W. P., Sachidanandam, R., Prasanth, K. V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R. H., and Spector, D. L. (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683-698.

Jeang, K. T., Chun, R., Lin, N. H., Gatignol, A., Glabe, C. G., and Fan, H. (1993). In vitro and in vivo binding of human immunodeficiency virus type 1 Tat protein and Sp1 transcription factor. J Virol *67*, 6224-6233.

Jeeninga, R. E., Hoogenkamp, M., Armand-Ugon, M., de Baar, M., Verhoef, K., and Berkhout, B. (2000). Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. J Virol *74*, 3740-3751.

Jordan, A., Bisgrove, D., and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J *22*, 1868-1877.

Jordan, A., Defechereux, P., and Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. EMBO J *20*, 1726-1738.

Kalpana, G. V., Marmon, S., Wang, W., Crabtree, G. R., and Goff, S. P. (1994). Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. Science *266*, 2002-2006.

Kamine, J., Elangovan, B., Subramanian, T., Coleman, D., and Chinnadurai, G. (1996). Identification of a cellular protein that specifically interacts with the essential cysteine region of the HIV-1 Tat transactivator. Virology *216*, 357-366.

Kao, S. Y., Calman, A. F., Luciw, P. A., and Peterlin, B. M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature *330*, 489-493.

Karin, M. (1999). The beginning of the end: IkappaB kinase (IKK) and NF-kappaB activation. J Biol Chem *274*, 27339-27342.

Karn, J. (1999). Tackling Tat. J Mol Biol *293*, 235-254.

Kashanchi, F., Piras, G., Radonovich, M. F., Duvall, J. F., Fattaey, A., Chiang, C. M., Roeder, R. G., and Brady, J. N. (1994). Direct interaction of human TFIID with the HIV-1 transactivator tat. Nature *367*, 295-299.

Kaye, J. F., and Lever, A. M. (1999). Human immunodeficiency virus types 1 and 2 differ in the predominant mechanism used for selection of genomic RNA for encapsidation. J Virol *73*, 3023-3031.

Kim, S. H., McQueen, P. G., Lichtman, M. K., Shevach, E. M., Parada, L. A., and Misteli, T. (2004). Spatial genome organization during T-cell differentiation. Cytogenet Genome Res *105*, 292-301.

Kim, Y. K., Bourgeois, C. F., Pearson, R., Tyagi, M., West, M. J., Wong, J., Wu, S. Y., Chiang, C. M., and Karn, J. (2006). Recruitment of TFIIH to the HIV LTR is a rate-limiting step in the emergence of HIV from latency. EMBO J *25*, 3596-3604.

Kingston, R. E., and Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev *13*, 2339-2352.

Kosak, S. T., Skok, J. A., Medina, K. L., Riblet, R., Le Beau, M. M., Fisher, A. G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science *296*, 158-162.

Kourmouli, N., Jeppesen, P., Mahadevhaiah, S., Burgoyne, P., Wu, R., Gilbert, D. M., Bongiorni, S., Prantera, G., Fanti, L., Pimpinelli, S., *et al.* (2004). Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. J Cell Sci *117*, 2491-2501.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kumaran, R. I., and Spector, D. L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. J Cell Biol *180*, 51-65.

Kurdistani, S. K., and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. Nat Rev Mol Cell Biol *4*, 276-284.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G., and Cremer, T. (2007). Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet *8*, 104-115.

Lassen, K., Han, Y., Zhou, Y., Siliciano, J., and Siliciano, R. F. (2004). The multifactorial nature of HIV-1 latency. Trends Mol Med *10*, 525-531.

Lassen, K. G., Ramyar, K. X., Bailey, J. R., Zhou, Y., and Siliciano, R. F. (2006). Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. PLoS Pathog *2*, e68.

Lenasi, T., Contreras, X., and Peterlin, B. M. (2008). Transcriptional interference antagonizes proviral gene expression to promote HIV latency. Cell Host Microbe *4*, 123-133.

Lever, A., Gottlinger, H., Haseltine, W., and Sodroski, J. (1989). Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J Virol *63*, 4085-4087.

Lewinski, M. K., Bisgrove, D., Shinn, P., Chen, H., Hoffmann, C., Hannenhalli, S., Verdin, E., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2005). Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. J Virol *79*, 6610-6619.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D. C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum Genet *80*, 224-234.

Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science *293*, 2453-2455.

Liu, Z. G. (2005). Molecular mechanism of TNF signaling and beyond. Cell Res *15*, 24-27.

Llano, M., Vanegas, M., Fregoso, O., Saenz, D., Chung, S., Peretz, M., and Poeschla, E. M. (2004). LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. J Virol *78*, 9524-9537.

Lomvardas, S., Barnea, G., Pisapia, D. J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.

Louvet, E., and Percipalle, P. (2009). Transcriptional control of gene expression by actin and myosin. Int Rev Cell Mol Biol *272*, 107-147.

Lusic, M., Marcello, A., Cereseto, A., and Giacca, M. (2003). Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter. EMBO J *22*, 6550-6561.

Mahy, N. L., Perry, P. E., and Bickmore, W. A. (2002). Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. J Cell Biol *159*, 753-763.

Majello, B., Napolitano, G., and Lania, L. (1998). Recruitment of the TATA-binding protein to the HIV-1 promoter is a limiting step for Tat transactivation. AIDS *12*, 1957-1964.

Marban, C., Suzanne, S., Dequiedt, F., de Walque, S., Redel, L., Van Lint, C., Aunis, D., and Rohr, O. (2007). Recruitment of chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. EMBO J *26*, 412-423.

Marcello, A. (2006). Latency: the hidden HIV-1 challenge. Retrovirology *3*, 7.

Marcello, A., Ferrari, A., Pellegrini, V., Pegoraro, G., Lusic, M., Beltram, F., and Giacca, M. (2003). Recruitment of human cyclin T1 to nuclear bodies through direct interaction with the PML protein. EMBO J *22*, 2156-2166.

Marcello, A., Lusic, M., Pegoraro, G., Pellegrini, V., Beltram, F., and Giacca, M. (2004). Nuclear organization and the control of HIV-1 transcription. Gene *326*, 1-11.

Marshall, R. M., and Grana, X. (2006). Mechanisms controlling CDK9 activity. Front Biosci *11*, 2598-2613.

Marshall, W. F., Dernburg, A. F., Harmon, B., Agard, D. A., and Sedat, J. W. (1996). Specific interactions of chromatin with the nuclear envelope: positional determination within the nucleus in Drosophila melanogaster. Mol Biol Cell *7*, 825-842.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol *6*, 838-849.

Marzio, G., and Giacca, M. (1999). Chromatin control of HIV-1 gene expression. Genetica *106*, 125-130.

Mateescu, B., Bourachot, B., Rachez, C., Ogryzko, V., and Muchardt, C. (2008). Regulation of an inducible promoter by an HP1beta-HP1gamma switch. EMBO Rep *9*, 267-272.

Matz, M. V., Lukyanov, K. A., and Lukyanov, S. A. (2002). Family of the green fluorescent protein: journey to the end of the rainbow. Bioessays *24*, 953-959.

Matzke, A. J., Huettel, B., van der Winden, J., and Matzke, M. (2005). Use of two-color fluorescence-tagged transgenes to study interphase chromosomes in living plants. Plant Physiol *139*, 1586-1596.

McBride, M. S., and Panganiban, A. T. (1996). The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. J Virol *70*, 2963-2973.

McDonald, D., Vodicka, M. A., Lucero, G., Svitkina, T. M., Borisy, G. G., Emerman, M., and Hope, T. J. (2002). Visualization of the intracellular behavior of HIV in living cells. J Cell Biol *159*, 441-452.

McNally, J. G., Muller, W. G., Walker, D., Wolford, R., and Hager, G. L. (2000). The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. Science *287*, 1262-1265.

Meaburn, K. J., and Misteli, T. (2007). Cell biology: chromosome territories. Nature *445*, 379-781.

Meaburn, K. J., and Misteli, T. (2008). Locus-specific and activity-independent gene repositioning during early tumorigenesis. J Cell Biol *180*, 39-50.

Merkenschlager, M., Amoils, S., Roldan, E., Rahemtulla, A., O'Connor, E., Fisher, A. G., and Brown, K. E. (2004). Centromeric repositioning of coreceptor loci predicts their stable silencing and the CD4/CD8 lineage choice. J Exp Med *200*, 1437-1444.

Misteli, T. (2005). Concepts in nuclear architecture. Bioessays *27*, 477-487.

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell *128*, 787-800.

Mitchell, A. R., Gosden, J. R., and Miller, D. A. (1985). A cloned sequence, p82H, of the alphoid repeated DNA family found at the centromeres of all human chromosomes. Chromosoma *92*, 369-377.

Molle, D., Maiuri, P., Boireau, S., Bertrand, E., Knezevich, A., Marcello, A., and Basyuk, E. (2007). A real-time view of the TAR:Tat:P-TEFb complex at HIV-1 transcription sites. Retrovirology *4*, 36.

Monini, P., Sgadari, C., Toschi, E., Barillari, G., and Ensoli, B. (2004). Antitumour effects of antiretroviral therapy. Nat Rev Cancer *4*, 861-875.

Nabel, G., and Baltimore, D. (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature *326*, 711-713.

Nguyen, V. T., Kiss, T., Michels, A. A., and Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. Nature *414*, 322-325.

Nielsen, J. A., Hudson, L. D., and Armstrong, R. C. (2002). Nuclear organization in differentiating oligodendrocytes. J Cell Sci *115*, 4071-4079.

Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J. A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065-1071.

Parada, C. A., and Roeder, R. G. (1996). Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. Nature *384*, 375-378.

Parada, L. A., Sotiriou, S., and Misteli, T. (2004). Spatial genome organization. Exp Cell Res *296*, 64-70.

Perkins, K. J., Lusic, M., Mitar, I., Giacca, M., and Proudfoot, N. J. (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Mol Cell *29*, 56-68.

Phair, R. D., and Misteli, T. (2001). Kinetic modelling approaches to in vivo imaging. Nat Rev Mol Cell Biol *2*, 898-907.

Pickersgill, H., Kalverda, B., de Wit, E., Talhout, W., Fornerod, M., and van Steensel, B. (2006). Characterization of the Drosophila melanogaster genome at the nuclear lamina. Nat Genet *38*, 1005-1014.

Pierson, T., McArthur, J., and Siliciano, R. F. (2000). Reservoirs for HIV-1: mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. Annu Rev Immunol *18*, 665-708.

Pierson, T. C., Zhou, Y., Kieffer, T. L., Ruff, C. T., Buck, C., and Siliciano, R. F. (2002). Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection. J Virol *76*, 8518-8531.

Poeschla, E. M. (2008). Integrase, LEDGF/p75 and HIV replication. Cell Mol Life Sci *65*, 1403-1424.

Politz, J. C., Browne, E. S., Wolf, D. E., and Pederson, T. (1998). Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. Proc Natl Acad Sci U S A *95*, 6043-6048.

Ragoczy, T., Bender, M. A., Telling, A., Byron, R., and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev *20*, 1447-1457.

Raha, T., Cheng, S. W., and Green, M. R. (2005). HIV-1 Tat stimulates transcription complex assembly through recruitment of TBP in the absence of TAFs. PLoS Biol *3*, e44.

Reddy, K. L., Zullo, J. M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature *452*, 243-247.

Redner, R. L., Wang, J., and Liu, J. M. (1999). Chromatin remodeling and leukemia: new therapeutic paradigms. Blood *94*, 417-428.

Rice, J. C., and Allis, C. D. (2001). Code of silence. Nature *414*, 258-261.

Richman, D. D., Margolis, D. M., Delaney, M., Greene, W. C., Hazuda, D., and Pomerantz, R. J. (2009). The challenge of finding a cure for HIV infection. Science *323*, 1304-1307.

Robinett, C. C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A., and Belmont, A. S. (1996). In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J Cell Biol *135*, 1685-1700.

Rodriguez, M. A., Shen, C., Ratner, D., Paranjape, R. S., Kulkarni, S. S., Chatterjee, R., and Gupta, P. (2007). Genetic and functional characterization of the LTR of HIV-1 subtypes A and C circulating in India. AIDS Res Hum Retroviruses *23*, 1428-1433.

Roix, J. J., McQueen, P. G., Munson, P. J., Parada, L. A., and Misteli, T. (2003). Spatial proximity of translocation-prone gene loci in human lymphomas. Nat Genet *34*, 287-291.

Sadoni, N., Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B. M., and Zink, D. (1999). Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. J Cell Biol *146*, 1211-1226.

Saha, A., Wittmeyer, J., and Cairns, B. R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. Nat Rev Mol Cell Biol *7*, 437-447.

Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol *6*, 73-77.

Schneider, U., Schwenk, H. U., and Bornkamm, G. (1977). Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. Int J Cancer *19*, 621-626.

Schotta, G., Ebert, A., Dorn, R., and Reuter, G. (2003). Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. Semin Cell Dev Biol *14*, 67-75.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev *18*, 1251-1262.

Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell *110*, 521-529.

Sexton, T., Schober, H., Fraser, P., and Gasser, S. M. (2007). Gene regulation through nuclear organization. Nat Struct Mol Biol *14*, 1049-1055.

Shav-Tal, Y., Darzacq, X., Shenoy, S. M., Fusco, D., Janicki, S. M., Spector, D. L., and Singer, R. H. (2004). Dynamics of single mRNPs in nuclei of living cells. Science *304*, 1797-1800.

Sheridan, P. L., Mayall, T. P., Verdin, E., and Jones, K. A. (1997). Histone acetyltransferases regulate HIV-1 enhancer activity in vitro. Genes Dev *11*, 3327-3340.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.

Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nat Methods *4*, 895-901.

Sodroski, J., Rosen, C., Wong-Staal, F., Salahuddin, S. Z., Popovic, M., Arya, S., Gallo, R. C., and Haseltine, W. A. (1985). Trans-acting transcriptional regulation of human T-cell leukemia virus type III long terminal repeat. Science *227*, 171-173.

Solovei, I., Cavallo, A., Schermelleh, L., Jaunin, F., Scasselati, C., Cmarko, D., Cremer, C., Fakan, S., and Cremer, T. (2002). Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). Exp Cell Res *276*, 10-23.

Solovei, I., Kreysing, M., Lanctot, C., Kosem, S., Peichl, L., Cremer, T., Guck, J., and Joffe, B. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell *137*, 356-368.

Solovei, I., Schermelleh, L., During, K., Engelhardt, A., Stein, S., Cremer, C., and Cremer, T. (2004). Differences in centromere positioning of cycling and postmitotic human cell types. Chromosoma *112*, 410-423.

Solovei, I., Walter, J., Cremer, M., Schermelleh, L., and Cremer, T. (2006). FISH on three-dimensional preserved nuclei, In FISH, A Practica Approach, B. Beatty, Mai, S., Squire, J., ed. (Oxford University Press), pp. 119-157.

Spector, D. L. (2003). The dynamics of chromosome organization and gene regulation. Annu Rev Biochem *72*, 573-608.

Spilianakis, C. G., Lalioti, M. D., Town, T., Lee, G. R., and Flavell, R. A. (2005). Interchromosomal associations between alternatively expressed loci. Nature *435*, 637-645.

Steger, D. J., and Workman, J. L. (1997). Stable co-occupancy of transcription factors and histones at the HIV-1 enhancer. EMBO J *16*, 2463-2472.

Stevenson, M. (2003). Dissecting HIV-1 through RNA interference. Nat Rev Immunol *3*, 851-858.

Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Sun, H. B., Shen, J., and Yokota, H. (2000). Size-dependent positioning of human chromosomes in interphase nuclei. Biophys J *79*, 184-190.

Taddei, A., Maison, C., Roche, D., and Almouzni, G. (2001). Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. Nat Cell Biol *3*, 114-120.

Taddei, A., Van Houwe, G., Hediger, F., Kalck, V., Cubizolles, F., Schober, H., and Gasser, S. M. (2006). Nuclear pore association confers optimal expression levels for an inducible yeast gene. Nature *441*, 774-778.

Thanos, D., and Maniatis, T. (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. Cell *83*, 1091-1100.

Thiagalingam, S., Cheng, K. H., Lee, H. J., Mineva, N., Thiagalingam, A., and Ponte, J. F. (2003). Histone deacetylases: unique players in shaping the epigenetic histone code. Ann N Y Acad Sci *983*, 84-100.

Thompson, J. S., Ling, X., and Grunstein, M. (1994). Histone H3 amino terminus is required for telomeric and silent mating locus repression in yeast. Nature *369*, 245-247.

Tian, B., and Brasier, A. R. (2003). Identification of a nuclear factor kappa B-dependent gene network. Recent Prog Horm Res *58*, 95-130.

Tracey, K. J., and Cerami, A. (1994). Tumor necrosis factor: a pleiotropic cytokine and therapeutic target. Annu Rev Med *45*, 491-503.

Tsukamoto, T., Hashiguchi, N., Janicki, S. M., Tumbar, T., Belmont, A. S., and Spector, D. L. (2000). Visualization of gene activity in living cells. Nat Cell Biol *2*, 871-878.

Turner, B. M., and O'Neill, L. P. (1995). Histone acetylation in chromatin and chromosomes. Semin Cell Biol *6*, 229-236.

Van Lint, C., Emiliani, S., Ott, M., and Verdin, E. (1996). Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. EMBO J *15*, 1112-1120.

Verdin, E., Paras, P., Jr., and Van Lint, C. (1993). Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. EMBO J *12*, 3249-3259.

Verma, I. M., Stevenson, J. K., Schwarz, E. M., Van Antwerp, D., and Miyamoto, S. (1995). Rel/NF-kappa B/I kappa B family: intimate tales of association and dissociation. Genes Dev *9*, 2723-2735.

Vignali, M., Hassan, A. H., Neely, K. E., and Workman, J. L. (2000). ATP-dependent chromatin-remodeling complexes. Mol Cell Biol *20*, 1899-1910.

Volpi, E. V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R. D., Goldsworthy, M., Powis, S. H., Ragoussis, J., *et al.* (2000). Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci *113 (Pt 9)*, 1565-1576.

Wade, P. A., and Kikyo, N. (2002). Chromatin remodeling in nuclear cloning. Eur J Biochem *269*, 2284-2287.

Weierich, C., Brero, A., Stein, S., von Hase, J., Cremer, C., Cremer, T., and Solovei, I. (2003). Three-dimensional arrangements of centromeres and telomeres in nuclei of human and murine lymphocytes. Chromosome Res *11*, 485-502.

Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. Science *193*, 848-856.

Williams, R. R., Azuara, V., Perry, P., Sauer, S., Dvorkina, M., Jorgensen, H., Roix, J., McQueen, P., Misteli, T., Merkenschlager, M., and Fisher, A. G. (2006a). Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. J Cell Sci *119*, 132-140.

Williams, S. A., Chen, L. F., Kwon, H., Ruiz-Jarabo, C. M., Verdin, E., and Greene, W. C. (2006b). NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. EMBO J *25*, 139-149.

Wolffe, A. P. (1994). Transcription: in tune with the histones. Cell *77*, 13-16.

Woodcock, C. L., and Dimitrov, S. (2001). Higher-order structure of chromatin and chromosomes. Curr Opin Genet Dev *11*, 130-135.

Wu-Baer, F., Sigman, D., and Gaynor, R. B. (1995). Specific binding of RNA polymerase II to the human immunodeficiency virus trans-activating region RNA is regulated by cellular cofactors and Tat. Proc Natl Acad Sci U S A *92*, 7153-7157.

Zack, J. A., Arrigo, S. J., Weitsman, S. R., Go, A. S., Haislip, A., and Chen, I. S. (1990). HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. Cell *61*, 213-222.

Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K. S., Singh, U.*, et al.* (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet *38*, 1341-1347.

Zhong, H., May, M. J., Jimi, E., and Ghosh, S. (2002). The phosphorylation status of nuclear NF-kappa B determines its association with CBP/p300 or HDAC-1. Mol Cell *9*, 625-636.

Zhou, J., Ermakova, O. V., Riblet, R., Birshtein, B. K., and Schildkraut, C. L. (2002). Replication and subnuclear location dynamics of the immunoglobulin heavy-chain locus in B-lineage cells. Mol Cell Biol *22*, 4876-4889.

Zink, D., Amaral, M. D., Englmann, A., Lang, S., Clarke, L. A., Rudolph, C., Alt, F., Luther, K., Braz, C., Sadoni, N.*, et al.* (2004). Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. J Cell Biol *166*, 815-825.