Department Biologie II

Anthropologie und Humangenetik

Ludwig-Maximilians-Universität München

Nuclear architecture and dynamics
of stably integrated retroviral sequences

Jens Nagel

Nuclear architecture and dynamics of stably integrated retroviral sequences

Dissertation der Fakultät für Biologie

der Ludwig-Maximilians-Universität München (LMU)

vorgelegt von:

Dipl. Biol. Jens Nagel

aus Heidelberg

Gutachter:

PD Dr. Steffen Dietzel

Prof. Dr. Thomas Cremer

Tag der mündlichen Prüfung: 15.12.2011

Table of contents

CONTENTS

4

CONTENTS

6

List of figures

List of tables

1. SUMMARY

Several studies investigated preferences of retroviral integration site selection on the molecular level of the linear genome. The three-dimensional nuclear organization, large-scale chromatin structure and possible dynamics of retroviral integration sites, however, have received little attention. The nucleus of a eukaryotic cell is a highly organized organelle. Retroviruses, to complete their life cycle, have to integrate their genomes into the host cell nucleus. Hence, the question arises if and how integration of foreign material changes the existing nuclear order.

This study investigated the consequences of stably integrated retroviral sequences for nuclear architecture in two independent experimental systems: in HIV infected human astrocytes, HeLa cells and T-lymphocytes as well as in the mouse hematopoietic precursor cell line cloneB transduced with a gamma-retroviral gene transfer vector. In HIV infected cells, integration sites were first mapped via LM-PCR (linker mediated PCR), whereas integration sites in the mouse cell line were previously known. BACs (bacterial artificial chromosomes) covering each integration site were tested and used as probes for fluorescence in situ hybridization (FISH).

Difficulties in visualizing the rather small nuclear HIV and gene transfer vector integrations were solved by modifications of FISH protocols, which allowed proper visualization of fluorescence signals deriving from <10 kb target sequences. Structurally preserved FISH preparations were evaluated by confocal microscopy and digital 3D image analysis. The nuclear 3D radial distribution of the integration loci was determined and compared to the homologous host loci without integration. In both, human and mouse cells, changes in large-scale chromatin structure upon retroviral sequence integration were observed.

Of the nine gene transfer vector integration loci in the mouse hematopoietic precursor cell line cloneB, two were significantly repositioned towards more interior regions of the nucleus. Notably, the more repositioned of these two integration sites is in the immediate vicinity of the protooncogene Evi1 (ecotropic virus integration site 1), a region with frequently observed viral and gene vector integrations. As in most cases upon retroviral integrations at this locus, the Evi1 gene in the cloneB cell line is overexpressed. This upregulation was shown to be essential for cell survival and immortalization of the cell line. Thus, repositioning of the locus is linked to activation

of Evi1 and clonal viability. All other seven gene transfer vector integration loci were slightly but not significantly repositioned towards more interior nuclear regions compared to their homologous loci.

The HIV infected HeLa cells investigated in this thesis were thought to be a clonal cell line, but are declared as a pool of clones here. Several HIV integration sites were found in different clones, whereas the clone harboring integration locus Xq22.1 dominated the population.

Two of five HIV integration loci in HeLa cells, 18q22.3 and Xq22.1, were significantly repositioned, with a more interior nuclear location as the respective loci on the homologous chromosomes. In both cases repositioning was accompanied by microscopically detectable large-scale chromatin decondensation. Repositioning and decondensation were not observed in astrocytes or T-lymphocytes. Therefore, such effects of retroviral integration may be locus or cell type specific. In HeLa cells integration locus 18q22.3 was more decondensed than Xq22.1, correlating with more pronounced repositioning. All five investigated HIV integration sites were transcriptionally active, but did not differ in their location towards SC35 splicing speckles compared to their homolgous regions. Sodium butyrate (NaBut) treatment increased HIV transcription, but did not lead to changes in radial distribution of integration loci compared to untreated cells.

In summary, the data show that stable retroviral integration can lead to alterations of nuclear chromatin organization and has the potential to modulate chromatin structure of the host cell. HIV is known to preferentially integrate within active transcription units with the advantage to ensure transcription of the provirus. Large-scale chromatin reorganization might provide an additional mechanism to increase transcriptional probability.

2. INTRODUCTION

2.1 Motivation and goals of the present work

The nucleus is highly organized and presents the core compartment of a eukaryotic cell. Organizational levels range from the DNA fiber to the nuclear distribution of chromosome territories (see 2.3.2). The nucleus harbors most of the cellular genome and provides the environment for decisive and sophisticated processes like gene expression and replication which are indispensable for life. Even small irritations in a process cascade may provoke serious implications such as apoptosis or malignancy.

Viruses are infectious particles which lack an own metabolism and therefore have to rely on a host cell. Retroviruses integrate their genome within the cellular genome and use the host transcription machinery to complete their life cycle (Alberts, 2011). A successful viral strategy ought to be a compromise between exploitation and survival of the host. The question arises whether the highly and subtle organized nuclear structure is irritated by the infiltration and integration of foreign material like retroviral genomes in any way. As described in the subsequent chapters, it is known that retroviral integration causes changes on the molecular and epigenetic level and affects host cell gene expression properties. Whether these processes lead to rearrangements in large-scale chromatin organization, however, has not yet been investigated.

Although retroviral and especially HIV integration sites have been investigated intensely on the molecular level of the linear genome, the 3D nuclear position of integration sites was not described before. Preferential targets for integration site selection of HIV are active transcription units (discussed in 2.3.3), which are widespread in the nucleus. Hence, the viral DNA could be integrated immediately after entering the nucleus or penetrating further resulting either in a more peripheral or internal position. But since the probability of transcription events declines from gene rich to gene poor regions and these regions are distinctly distributed within the nucleus, preferences for the HIV genome to integrate are conceivable.

Another consideration is repositioning of the integration site. Having integrated into a putative inauspicious region, relocation may be the solution. Moreover, the addressed alterations on the molecular and epigenetic level could lead to alterations in nuclear architecture and therefore benefit repositioning of the integration site.

In this thesis these aspects with the focus on repositioning of the integration locus and changes in the affected chromatin environment were investigated. Experiments were performed, on the one hand, in three HIV infected cell types (astrocytes, HeLa cells and T-lymphocytes), on the other hand, in a mouse hematopoietic precursor cell line transduced with a gamma-retroviral gene transfer vector. Thanks to their capacity to deliver genetic material into target cells, viral gene vectors play an important role in the field of gene therapy. Retroviral vectors integrate stably into the host genome and therefore have the potential to exert enduring therapeutical effects (see 2.4).

2.2 General aspects of HIV

2.2.1 Taxonomy, history and genomic properties

HIV belongs to the retrovirus family and to the genus Lentivirus (Latin: lentus = slow). Lentiviruses are known for long-duration illnesses with a long incubation period. Two types of HIV are known to date: HIV-1 and HIV-2. HIV-2 shows a much lower virulence as well as infectivity compared to HIV-1. Throughout this thesis the term HIV is always referring to HIV-1 except when stated otherwise. HIV is thought to have developed from SIV (simian immunodeficiency virus), the equipollent in African primates. In contrast to humans and HIV, monkeys adapted to SIV and do not suffer any illness. A current study revealed that SIV has been in monkeys for at least 32,000 years, much longer than the previously estimated several hundred years (Worobey et al., 2010). SIV crossed the species barrier from monkey to human during the late 19th or early 20th century (Worobey et al., 2008). The most probable act of transmission is thought to be an accident, where a bushmeat hunter was bitten or cut during the butchering of a monkey and the resulting contact with infected simian blood.

The HIV genome consists of two copies of positive sense single-stranded RNA, each ~9.2 kb in size, which encode the canonical retroviral genes gag, pol and env plus six additional open reading frames with regulatory genes (tat, rev, nef, vif, vpr, and vpu) which makes HIV a member of the complex retroviruses (Frankel and Young, 1998; Watts et al., 2009).
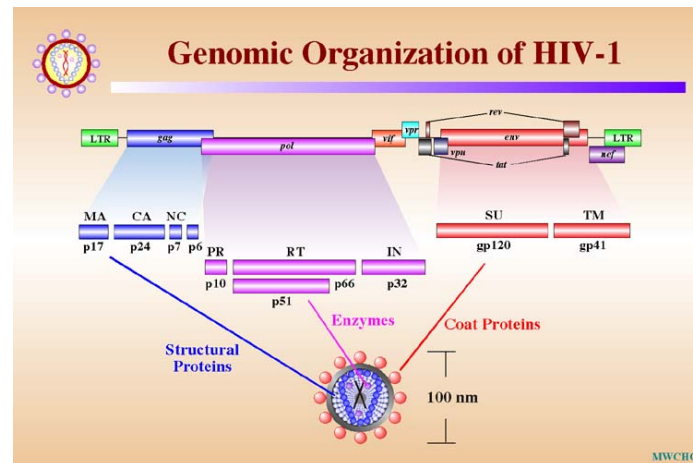
Figure 1: HIV genome
Image taken from Stanford University webpage
(http://www.stanford.edu/group/virus/retro/2005gongishmail/HIV.html)

2.2.2 Life cycle

A brief overview of the HIV life and replication cycle to describe viral behavior in host cells from cell entry until release of viral particles is given here. More detailed accounts are given by (Ciuffi and Bushman, 2006; Klimas et al., 2008; Delelis et al., 2010).

Cell entry

HIV entry into host cells is preferentially mediated by the surface receptor CD4. Therefore, T-lymphocytes (T helper cells), monocytes, macrophages and dendritic cells, which are all cells of the immune system, come into consideration as primary target cells. The viral envelope fuses with the cell membrane and the HIV capsid is released in the cell (Suzuki and Craigie, 2007). Recent studies provide compelling evidence that HIV enters the cell primarily via endocytosis and not by direct fusion of the plasma membrane and release of the viral genome into the cytoplasm (Campbell and Hope, 2008; Miyauchi et al., 2009; Permanyer et al., 2010). In vitro studies revealed that primary fibroblast and osteoblast-like cells were also susceptible for HIV infection suggesting low expression of small amounts of CD4 molecules (Mellert et al., 1990). Furthermore CD4 independent HIV infection of central nervous system-derived cells was described (Harouse et al., 1989; Brack-Werner et al., 1992). Nevertheless, infection of lymphoid cells mediated by CD4 receptor is the most efficient. In addition to direct infection, HIV can hijack dendritic cells and be transmitted to helper T-cells, a process, in which dendritic cells are thought to offer

some kind of reservoir for HIV (Grouard and Clark, 1997; Geijtenbeek et al., 2000; McDonald et al., 2003).

Integration and transcription

In the cytoplasm, the pre-integration complex (PIC) consisting of viral proteins, host proteins and the HIV genome is formed (Miller et al., 1997). Within this complex the viral RNA genome is converted to dsDNA by the viral enzyme reverse transcriptase (reviewed in Herschhorn and Hizi, 2010) while simultaneously the RNA is degraded. Subsequently, the PIC together with the DNA cargo is actively transported through a nuclear pore into the nucleus, where the viral enzyme integrase mediates the integration into the host genome. The integrated viral genome is called provirus and is embedded between 5´ and 3´ long terminal repeat (LTR) sequences generated during the process of reverse transcription. Each LTR is approximately 640 bp in length, which results in an integrated HIV provirus genome size of ~9.7 kb (Ratner et al., 1985). The LTR regions contain all of the essential signals for gene expression such as enhancer, promoter, transcription initiation, transcription terminator and polyadenylation signal.

Gene transcription itself is carried out by the host enzymes RNA polymerase II, poly A synthetase and guanyl transferase. The promoter within the LTR region is boosted by the viral protein Tat (Bannwarth and Gatignol, 2005; Brady and Kashanchi, 2005). The Tat protein is expressed from the small number of initial RNA transcripts and thereupon extensively increases the transcription of the viral genome. Initially the mRNA is spliced by the host splicing machinery and transported to the cytoplasm where it is translated into the proteins Tat and Rev. Rev allows unspliced RNA to exit the nucleus, promoting the production of different viral structural proteins and the whole viral genome. HIV was long believed to insert only one copy of its genome into the host nucleus. Several studies, however, disproved this assumption, revealing multiple infections within one nucleus (Jung et al., 2002; Dang et al., 2004). In this context a certain importance is assigned to the entry pathway since dendritic cells can capture from a few up to a few hundred virions per cell (McDonald et al., 2003; Chen et al., 2005).
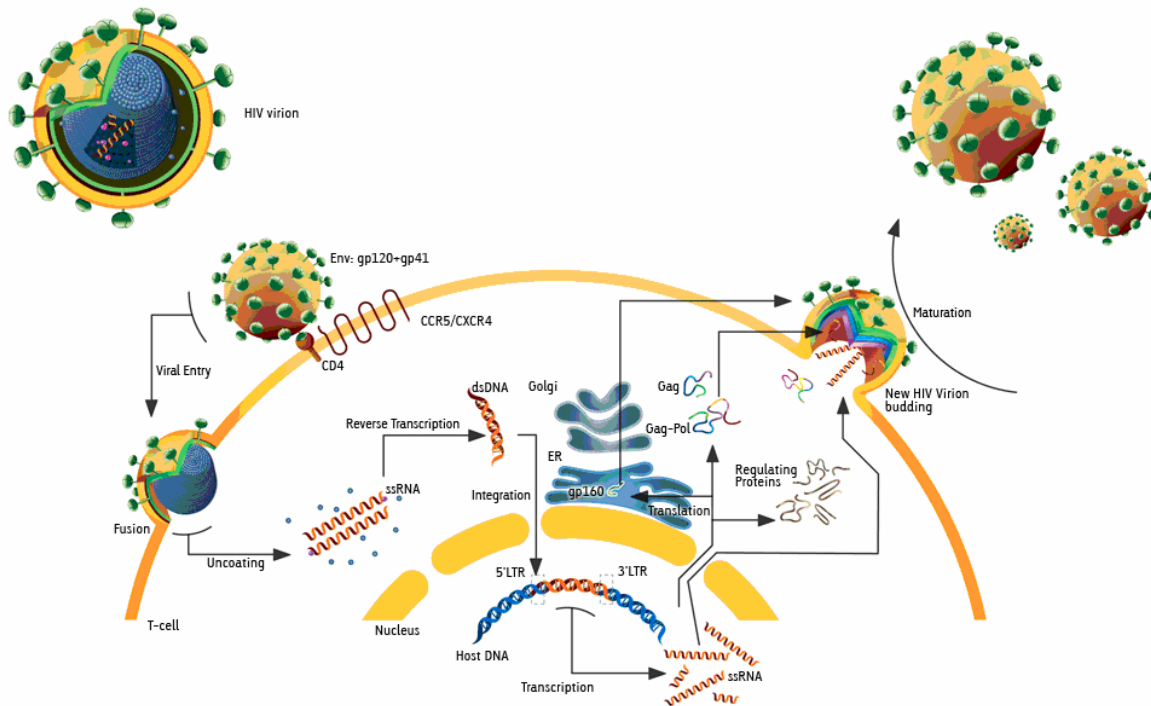
Figure 2: HIV life cycle
Image taken from (Mandell et al., 2004)

Assembly and release

Under assistance of viral glycoproteins generated from the Env polyprotein, the budding of new virus particles (virions), containing the HIV genomic RNA from the cell membrane, starts. A visual summary of the HIV life cycle is given in Figure 2.

2.3 Interplay of HIV and host cell chromatin

2.3.1 Chromatin architecture on the molecular level

Chromatin is the organized and packaged form of DNA in nuclei of eukaryotic cells. Chromatin can be divided into different hierarchical and organizational levels with increasing compaction ratio (Woodcock and Dimitrov, 2001; Woodcock and Ghosh, 2010). Nucleosomes represent the smallest unit. Approximately 146 to 147 base pairs of DNA are wound around a histone octamer building the nucleosome core particle (NCP) which consists of two copies of each of the core histones H2A, H2B, H3 and H4 (Luger et al., 1997; Felsenfeld and Groudine, 2003; Richmond and

Davey, 2003; Woodcock, 2006). Neighboring nucleosomes are connected by linker DNA that varies in length from 10 to 80 base pairs. When DNA is uncondensed this structure is referred to as "beads on a string" under the electron microscope. The linker DNA is attached to the histone H1, which forms by compaction the putative 30 nm fiber (Adkins et al., 2004; Tremethick, 2007). Further compaction leads to the most condensed chromatin form, the mitotic chromosome. Here, the DNA is compacted some 10000- to 20000-fold (Belmont, 2006).

Chromatin in interphase nuclei is subclassified in euchromatin and heterochromatin. Distinct amino acids of the core histones can be modified in various manners such as acetylation, ubiquitylation, methylation and phosphorylation. Depending on the decoration either specific heterochromatic or euchromatic regions are characterized (Li et al., 2007; Campos and Reinberg, 2009). Heterochromatin is preferentially located around nucleoli and at the nuclear periphery where it is associated with the nuclear lamina. Compared to euchromatin, it is more condensed and transcriptionally less active. Methylation of histone H3 at lysine 9 and the heterochromatin protein HP-1 are typical markers and necessary for heterochromatin formation (Grewal and Jia, 2007). Constitutive heterochromatin perpetually remains in the compact silent form and is predominantly occuring in repetitive, gene-poor and late replicating DNA sequences. Regions packaged in facultative heterochromatin are transcriptionally silent but retain the potential to alternate between heterochromatin and euchromatin and therefore become transcriptionally competent. An example is the formation of the inactive X-chromosome (Trojer and Reinberg, 2007). Euchromatin is less condensed with a compaction ratio supposedly below the 30 nm fiber, more accessible and in principle transcriptionally active. Hyperacetylation and -methylation of lysine 4 and lysine 36 of histone H3 are typical indicators for transcriptionally active regions (Li et al., 2007).

2.3.2 Nuclear architecture and inner nuclear dynamics

Intensive research has contributed to see the nucleus as a highly organized and dynamic organelle (Lanctot et al., 2007). In the interphase nucleus chromosomes occupy distinct confined spaces termed chromosome territories which are non randomly distributed within the nucleus (for review see Cremer and Cremer, 2001; Cremer et al., 2006). Depending on the cell type the positioning was linked to size, gene density and morphology (Cremer and Cremer, 2010). In human lymphoblast nuclei, for example, gene dense chromosome territories tend to be located more

18

internal, whereas gene poor ones are more likely to be found towards the periphery and associated with the lamina (Boyle et al., 2001). This gene density dependent distribution was also found in other animals like rodents, cattle, primates and chicken (Habermann et al., 2001; Mayer et al., 2005; Neusser et al., 2007; Köhler et al., 2009). On the other hand, in human fibroblasts a size dependent correlation for the nuclear position of chromosome territories could be shown (Sun et al., 2000; Cremer et al., 2001; Bolzer et al., 2005). The nuclear organization with distinct arrangements is also implemented on chromosomal substructures like chromosome arms (Dietzel et al., 1998), telomeres and centromeres (Molenaar et al., 2003; Weierich et al., 2003). The highly organized structure and compartmentalization does not exclude inner nuclear dynamics. Especially for gene function a prominent role was assigned to dynamics of chromatin and accordingly chromosomal interactions. Studies on nuclear architecture applying molecular methods (3C, 4C) revealed spatial proximity of genes located far apart on the linear chromosomal level or even on different chromosomes (reviewed in de Laat, 2007; Fraser and Bickmore, 2007; Gondor and Ohlsson, 2009). Associations of genes have been described both in *cis* and *trans* upon transcriptional activation (Thompson et al., 2003; Osborne et al., 2004; Simonis et al., 2006; Osborne et al., 2007; Sutherland and Bickmore, 2009) even though the functional basis of such a gene convergence is controversially discussed.

Further examples for inner nuclear dynamics are genes or gene regions looping out from their chromosome territory. This was described for the α-globin gene (Brown et al., 2006) or the gene dense region 11p15.5 (Mahy et al., 2002; Albiez et al., 2006) Furthermore, erythroid genes located on different chromosomes prevalently associate with each other (Brown et al., 2008). Examples for regions looping out upon transcriptional activation are Hox gene loci (Chambeyron and Bickmore, 2004; Morey et al., 2007), genes in the major histocompatibility complex (MHC) (Volpi et al., 2000) or the proneural regulator gene Mash1 (Williams et al., 2002) and several adipogenesis genes (Szczerbal et al., 2009). Protrusions and motions in interphase nuclei can take place both over shorter distances within a range of <2 μm (Abney et al., 1997; Gerlich et al., 2003) and in a long-range manner with movements >2 μm distances (Bornfleth et al., 1999; Chuang et al., 2006).

The radial nuclear position of a given gene may be linked to its transcriptional state (Lanctot et al., 2007; Kumaran et al., 2008; Takizawa et al., 2008b; Zhao et al., 2009). Since gene density increases when approaching the nuclear interior it might

be convenient for a genetic locus to increase its transcriptional potential by moving towards the nuclear center. Repositioning towards the nuclear interior upon transcriptional activation has been reported for transgenes (Tumbar and Belmont, 2001; Dietzel et al., 2004) and several gene regions such as the different nuclear position of the two alleles of the monoallelically expressed GFAP gene in astrocytes (Takizawa et al., 2008a). The active allele generally reveals a more internal position compared to its inactive equivalent. Repositioning for transcriptionally activated genes was shown especially during cell development and differentiation. The β-globin locus in mouse erythroid cells, for instance, departs from the nuclear periphery after its expression is switched on (Ragoczy et al., 2006). The already mentioned Mash1 gene moves away from a peripheral location upon transcriptional upregulation during neural differentiation (Williams et al., 2006). In cell types, where the CFTR gene is transcriptionally inactive, this gene locus is associated with the nuclear periphery and perinuclear heterochromatin. In contrast, in cell types, where the gene is actively transcribed, it is embedded in euchromatin enriched areas more internal in the nucleus (Zink et al., 2004). The influence of a surrounding chromatin milieu was furthermore shown by the correlation of transcriptional silencing and repositioning to a putative transcriptional repressive environment. In Type 1 helper cells the Type 2 helper cell specific transcription factors GATA-3 and c-maf are repositioned to centromeric heterochromatin or the nuclear periphery retaining their transcriptionally inactive status (Hewitt et al., 2004).

But changes in the nuclear distribution of a gene do not necessarily have to alter its transcriptional state or vice versa. In murine ES cells that differentiate to neurons the transcription of the Mash1 gene is upregulated and its localization changes from a peripheral to a more internal position (Williams et al., 2006). Several neighboring genes are transcriptionally affected by this relocation towards the center and some of them remain in their inactive state. Repositioning of several genes during tumorigenesis independent of changes in their activity was also observed (Meaburn and Misteli, 2008). Moreover, results from in-vitro designed systems revealed discrepancies. In one approach, the transcription of a reporter gene was reduced after being tethered to the nuclear lamina (Reddy et al., 2008). On the other hand the nuclear lamina does not have to be a transcriptionally inactive region. Placing an inducible transgene at the periphery to interact with lamin B did not have an effect on its transcriptional property (Kumaran and Spector, 2008).

2.3.3 HIV integration and host cell chromatin

The pre-integration complex (PIC) is a large nucleoprotein complex carrying the viral cDNA (see 2.2.2). Having entered the nucleus, the PIC immediately begins to interact with the chromosomal DNA. The ubiquitously expressed nuclear protein lens epithelium-derived growth factor (LEDGF)/p75 plays an important role in this process. LEDGF/p75 is a transcriptional regulator, which is tightly associated with chromatin. It contains a nuclear localization signal (NLS) and binds to the viral enzyme integrase in the PIC (Ciuffi and Bushman, 2006; Singh et al., 2006; Engelman and Cherepanov, 2008). The next step is searching an appropriate locus for integration. In this process LEDGF/p75 acts as a targeting factor, but it is not essential for the actual event of integration, which is catalyzed by the viral enzyme integrase (Ciuffi et al., 2005; Shun et al., 2007). A precise target sequence for HIV integration has not been gleaned yet although weak consensus sequences at and near the integration site could be identified (Holman and Coffin, 2005; Wu et al., 2005). HIV integration was first believed to be random (Holmes-Son et al., 2001) but finally proclivities were identified. The preferential but not exclusive locations for HIV to integrate into the host genome are active transcription units (Schröder et al., 2002; Mitchell et al., 2004; Bushman et al., 2005). This was confirmed by investigations of epigenetic properties at the integration site (Wang et al., 2007).

Interestingly, HIV integration site location differs in adult and infant cells (Wellensiek et al., 2009). The frequency of integrations within transcriptionally active genes is higher in cord T-lymphocytes and macrophages than in adult blood cells suggesting a relation to the higher viremia, i.e. more viruses in the bloodstream, and a more rapid progress to AIDS observed in HIV infected infants. Moreover, in patients with low viremia, HIV integrations are predominantly found in the GC-richest isochors. In patients with higher viremia, on the other hand, integrations more often locate in GC-poor isochore regions. Patients with the highest viremia integrations into GC rich regions are often predominant again but not to the same extent as for low viremia levels (Tsyba et al., 2004). Since the GC content correlates with gene density, it is to be expected for the virus to integrate into gene dense and therefore GC rich regions. This might guarantee proper transcription. The observation at higher viremia levels is explained by the authors with the compositional matching of the GC poorness of HIV-1 called `isopycnic integration´. This was also described before for other viruses and is associated with stability of integration and transcription of the

provirus (Rynditch et al., 1998). In the study of (Wang et al., 2007) strong associations with distinct histone modifications like mono-, di- or trimethylated H3K4, acetylated H3K9/K14 and acetylated H4 which all are markers for euchromatin and therefore open and accessible chromatin have been shown.

For the integration of the double-stranded virus genome it is indispensable to create double-strand breaks (DSB) in the host genome. A human cell is confronted with about 10000 DNA lesions per day (Lindahl, 1993), resulting in the evolution of a sophisticated repair machinery in which Atm (ataxia telangiectasia mutated) and Atr (Atm and Rad-related) kinases play a pivotal role. The host cell discerns retroviral integration as an event of DNA damage indicated by the histone variant γ-H2AX (Daniel et al., 2004). Many viruses exploit the cellular damage response machinery (Lilley et al., 2007) and HIV is no exception. ATM is required for efficient HIV infection (Lau et al., 2005) and blocking Atm activity with caffeine leads to inhibition of the HIV replication cycle (Nunnari et al., 2005). After integration of the viral dsDNA genome, chromatin in the 5´LTR region is ordered by arrangement of five nucleosomes (nuc-0 to nuc-4). Thus, viral gene expression is regulated through limited access to the promoter region within the 5´LTR (Easley et al., 2010). Nuc-1 is located at the transcription start site and has to be remodeled for initiation of transcription. This is managed by an interaction of the viral Tat protein with the host SWI/SNF complex. This complex regulates the expression of several endogenous genes by altering their nucleosomal position along the DNA. In vitro, Trichostatin A (TSA) or sodium byturate (NaB) treatment leads to histone acetylation around nuc-1 (Coull et al., 2002). One of the core subunits of the SWI/SNF complex, BRM, is also involved in the regulation of alternative splicing (Batsche et al., 2006) which is discussed in the next chapter.

2.3.4 HIV transcription and the host splicing machinery

HIV dependens on the host cell splicing machinery. More than 40 different mRNAs are created via alternative splicing from the primary RNA transcript (Tazi et al., 2010). For this purpose HIV-1 uses multiple splice donor and splice acceptor sites on the polycistronic pre-mRNA generated by the cellular RNA polymerase II. The SR proteins are serine/arginine rich proteins that are involved in splicing processes. Within the nucleus, SR proteins often localize in nuclear speckles, which occupy discrete regions with little or no DNA. They are dynamic structures, which vary in size, shape, location and number (Spector and Lamond, 2010). Especially for

alternative splicing, SR proteins are essential by promoting the use of nearby splice sites (Graveley, 2000; Lin and Fu, 2007).

In in vitro HIV *infected* cells, the distribution of integrated proviral DNA with respect to SC35 splicing speckles has only been investigated once with the result of a random distribution (Bell et al., 2001). This study was performed in a human osteosarcoma cell line. In cells *transfected* with plasmids containing the HIV genome, random distribution (Berthold and Maldarelli, 1996; Zhang et al., 1996; Boe et al., 1998) but also co-localization of SC35 with HIV DNA was observed (Favaro et al., 1998).

2.3.5 HIV persistence and latency

The highly active antiretroviral therapy (HAART) markedly reduces morbidity of patiens infected with HIV (Palella et al., 1998). Although viremia is decreased to undetectable levels by inhibitors of viral protease, fusion and integration, HIV can persist in cells in a latent state. Latently infected cells harbor per definition replication competent integrated HIV genomes, which are transcriptionally silent but reactivatable (Pace et al., 2011). Pre- and post-integration latency is distinguished. In pre-integration latency, resting CD4$^+$ T cells are infected and therefore provide some reservoir for HIV. Reactivation of viral transcription in these cells can be induced by antigenic stimulation, cytokines, mitogens or phorbol esters (Bisgrove et al., 2005). In post-integration latency on the other hand, active CD4$^+$ cells are infected and subsequently turned into a resting memory state (Dahl et al., 2010).

Several mechanisms and characteristics to induce and maintain viral latency are known. It was shown that integration into heterochromatic structures, where accessibility of transcription factos is hindered, promotes latency (Jordan et al., 2003). Binding of the transcription machinery is also known to be prevented by high levels of CpG methylation of viral promoter sequences (Blazkova et al., 2009). Furthermore, the precise position of the integration site might also be crucial. The immediate vicinity of the LTR promoter region to nearby host promoters can result in transcriptional interference for example by collision of RNA polymerase complexes (Coiras et al., 2009). This can for example explain a latent provirus in transcriptionally active genes. Transcriptional repression of the HIV provirus has been shown to be linked up with distinct epigenetic modifications such as hypoacetylation whereas treatment with deacetylase inhibitors accordingly increases HIV transcription (He and Margolis, 2002).

In (Dieudonne et al., 2009) a scenario of HIV latency with subsequent transcriptional activation of the provirus by phorbol ester (TPA) treatment is described. The investigated Jurkat cell line J-lat A1 was transduced with an HIV-1 vector which contains a reporter gene under the control of the viral 5´LTR. Upon transcriptional activation the original association of the latent provirus with pericentromeric heterochromatin was lost. The nuclear localization of the provirus remained at the periphery.

Efficient viral transcription is regulated by the viral protein Tat (Kao et al., 1987). Therefore, the nuclear presence of Nf-kB and especially Tat, whose activity is mainly regulated by the acetylation of distinct lysine residues, plays an important role in the latent provirus state (Coiras et al., 2009). Inefficient transportation of viral mRNA to the cytoplasm decreases translation of viral proteins, by which transcriptional efficiency is diminished through negative feedback. RNA interference has also been brought in association with HIV latency. Both microRNAs encoded by HIV genome sequences and host microRNAs can modulate viral gene expression (Corbeau, 2008; Kumar and Jeang, 2008). Applying that mechanism the virus can induce its own latency.

2.4 Gene therapy vectors

The use of retrovirally derived gene vectors to deliver genetic material to target cells in terms of gene therapy offers a possibility to profit from retroviruses. The first clinical study using gene transfer is dated to the year 1989 (Rosenberg et al., 1990) where a modified retroviral gene vector was used. Naked DNA and gene vectors based on adenoviruses or retroviruses present the three main groups for delivery of genetic material used in gene therapy (Figure 3).

Retroviral gene vectors provide the advantage of a stable integration into the host genome and hence an enduring therapeutical effect (Kay et al., 2001; Mancheno-Corvo and Martin-Duque, 2006; Edelstein et al., 2007). Although the payload is limited to ca 10 kb, most cDNAs can be accommodated. Lentiviruses became attractive for vector design because of their ability to infect non-dividing cells, an exclusive feature among the retrovirus family. The common method to convert a putative pathogenic retrovirus into a useful gene transfer vector is to delete the viral genes and replace them by the designated expression unit. Therefore, retroviral gene

transfer vectors lack most retroviral protein coding sequences while retaining the viral packaging signal and the 5´ and 3´ terminal repeat sequences (LTRs), which are required for DNA integration (Thiel, 2007; Nolan, 2009).

A major concern using retrovirally derived gene vectors is the unpredictable integration site selection. This can lead to severe consequences associated with insertional mutagenesis by disruption of tumor suppressor genes or activation of nearby proto-oncogenes (Ott et al., 2006; Hacein-Bey-Abina et al., 2008; Howe et al., 2008; Stein et al., 2010). For retroviruses, preferences in the selection of integration sites are known (Bushman et al., 2005), but which factors finally determine the definite genomic position remains to be fully elucidated. Retrovirus-derived vectors have preferences similar to their parental retroviruses (Mitchell et al., 2004; Baum et al., 2006; Daniel and Smith, 2008). The Gammaretrovirus Murine leukaemia virus (MLV) is the most commonly used retrovirus for generating gene transfer vectors (Cotrim and Baum, 2008). MLV based vectors favor the vicinity of transcription start sites for integration site selection. This may alter the structure of the respective host gene, potentially resulting in activation or disruption (Wu et al., 2003).

The mouse hematopoietic precursor cell line cloneB, which was investigated in this thesis, was transduced with the MLV based vector pSF91-GFP (Modlich et al., 2006). Transduction as a technique in molecular biology is the process of inserting DNA into a target cell by a virus or a viral vector (Alberts, 2011). The cloneB cell line was established by transduction of primary murine bone marrow cells followed by limiting dilution assay (Figure 4). Although gene therapy is not fully matured yet and unfortunately, leukemia outbreaks derived from gene transfer vector integration were observed, numerous clinical trials were thoroughly successful and looked promising (Cotrim and Baum, 2008). Nevertheless, further development and improvements still have to be made.
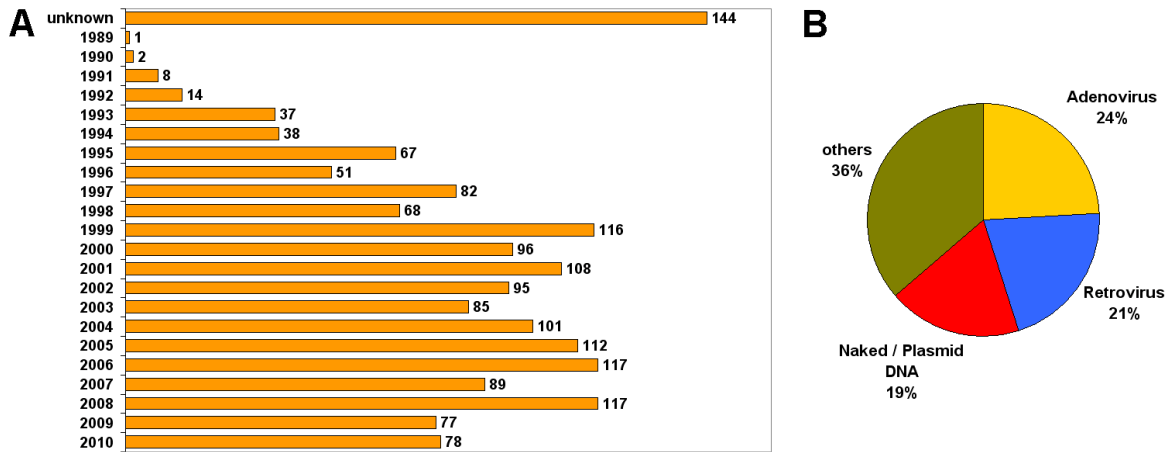
Figure 3: Gene therapy statistics
A – number of gene therapy clinical trials approved worldwide (1989-2010); B – vectors used in gene therapy clinical trials; data from http://www.wiley.com//legacy/wileychi/genmed/clinical/ (state 04/11).
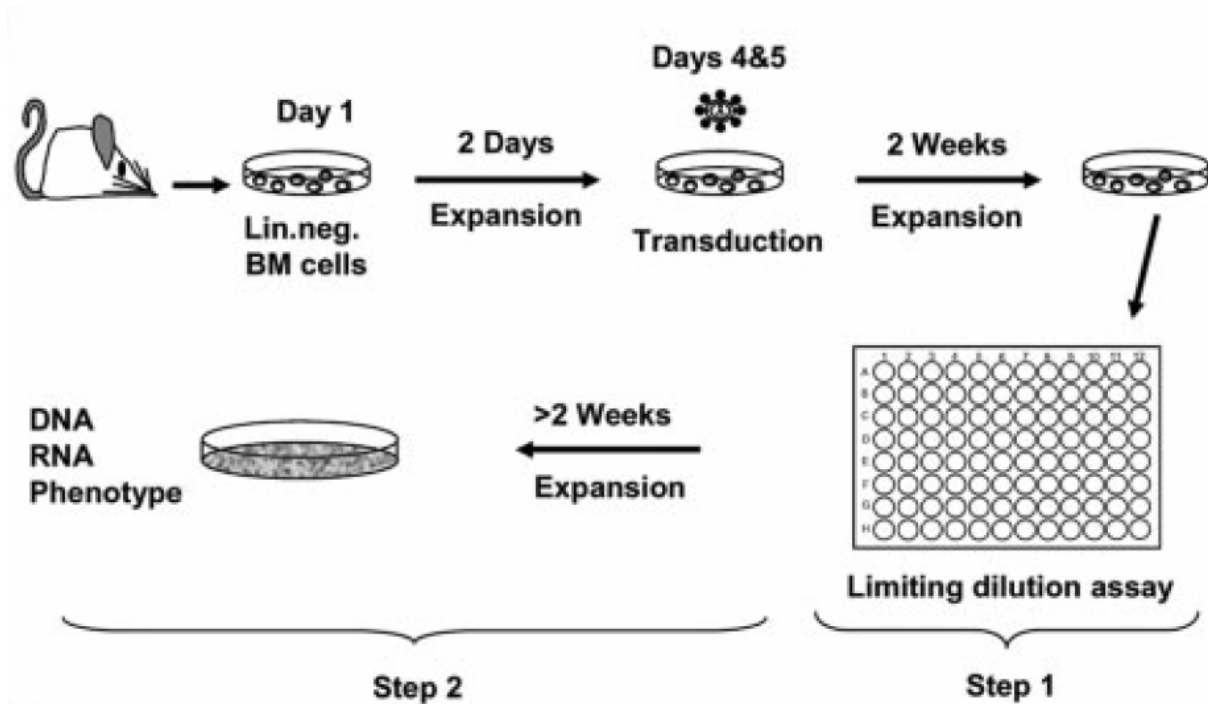
Figure 4: Establishment of the cloneB cell line
Image taken from (Modlich et al., 2006); primary bone marrow cells from mice (C57BL/6J) were isolated, purified using Lin⁻ specific antibodies (Gr1, CD11b, CD45R/B220, CD3e, TER119) and cultivated for two days. Transduction with the pSF91-eGFP vector was performed on day four and repeated on day five. After two weeks of cultivation the cells were diluted and cultivated for additional two weeks. Randomly picked clone were characterized with genotypic and phenotypic parameters by which a clone with stable growth and normal cytology was chosen (cloneB). For further information see (Schambach et al., 2000; Modlich et al., 2006).

2.5 Experimental strategy

The spatial distribution of retroviral integration sites was investigated for two examples: HIV integration sites in astrocytes (TH4-7-5), HeLa cells (LC5-HIV) and T-lymphocytes (KE37/1-IIIB) (see 3.1.1, 3.1.2 and 3.1.3) and gamma-retroviral gene transfer vector integration sites in the mouse hematopoietic precursor cell line cloneB (see 3.1.4). HIV Integration sites were mapped using the LM-PCR technique (see 4.1.1). The gene transfer vector integration sites in the mouse cell line were mapped after the cell line was established in the lab of Christopher Baum (Hannover Medical School, Germany) (Modlich et al., 2006). Besides the five integration sites mentioned in the original publication, four additional ones were identified in the lab of Manuel Grez (Georg Speyer Haus, Frankfurt, Germany), who kindly provided the cell line (for overview of integration sites see 6.2.2).

To determine possible changes in chromatin structure or repositioning upon retroviral integration, the integration site regions were compared to the respective regions on the homologous chromosome within the same nuclei. Therefore, the karyotype of both HIV infected cells and the mouse cell line was analysed, respectively the integration site harboring chromosomes were investigated on homology (see 4.1.4 and 4.7.1). BAC (bacterial artificial chromosomes) clones covering the integration site were used as probes for fluorescence in situ hybridization (FISH). Integration sites and respective loci on the homologous chromosomes were distinguished by co-localization of BAC and HIV or, respectively, gene vector FISH signals. To compare properties of signals originated from FISH probes of equal sizes, the BAC signals were used for the actual location and chromatin analysis. The HIV and gene vector signal was only used to discriminate the homologous BAC signals.

Nuclei for confocal microscopic recordings were selected as follows: the nucleus should show the average shape and morphology of the respective cell type population. Additionally, the HIV or gene vector signal should be visible and co-localizing with the BAC signal used for the respective experiment. Due to the critical size of the HIV and gene vector target sequence (<10 kb), the distinct signal visualization was challanging (discussed in 5.1). Image analysis revealed the position of the integration sites and the homologous regions in a 3D nuclear context (see 4.3 and 4.8.2). Changes in chromatin condensation upon retroviral integration were

determined by measuring and comparing the surface areas of the BAC signals (see 4.4 and 4.8.3). While BAC probe testing, karyotyping and chromatin compaction analysis of the integration site regions in the mouse hematopoietic precursor cell line were part of this thesis, 3D FISH experiments and evaluation were performed by Birgit Groß as a Diploma thesis (Groß, 2009). HIV integration sites were additionally investigated after sodium butyrate (NaBut) treatment, which increased the proviral transcription (see 4.5). Furthermore, the spatial distribution of HIV integration sites and HIV RNA with regard to SC35 splicing speckles was determined (see 4.6).

3. MATERIALS AND METHODS

3.1 Cell material

All HIV infected cells were kindly provided by Ruth Brack-Werner (Helmholtz Zentrum München). The HIV infected astrocytes, HeLa cells and T-lymphocytes were characterized in advance with regard to the number of integrations per nucleus via LightCycler qPCR (Manja Meggendorfer, Helmholtz Zentrum München, personal communication). The results predicted one provirus copy per astrocyte, six copies per T-lymphocyte and 0.6 copies per HeLa cell nucleus. For simplicity in the further course of this thesis, HIV infected cells are referred to as astrocytes, HeLa cells and T-lymphocytes or with the respective abbreviation (Astro, HeLa, TLy).

3.1.1 Astrocytes (TH4-7-5)

TH4-7-5 cells derive from the human astrocytoma cell line 85HG-66, grow adherently and show fibroblast like morphology. HIV-1 infection of 85HG-66 cells was ensured by co-cultivation with chronically HIV$_{IIIB}$-infected KE37/1-IIIB cells for 5 days (Brack-Werner et al., 1992). The successful infection was monitored by indirect immunoperoxidase staining (IPS) using human anti-HIV-1 serum. Cells were cultured in RPMI medium supplemented with 10% FCS, 100 U/ml penicillin and 100 µg/ml streptomycin.

3.1.2 HeLa cells (LC5-HIV)

LC5-HIV cells were established by HIV-1 infection of the cell line L-132 (Mellert et al., 1990), which originally was thought to be derived from embryonic lung tissue but subsequently was identified as a HeLa-derived cell line (ATCC number: CCL-5). HIV-1 infection was accomplished by the same co-cultivation technique just like above. LC5-HIV cells grow adherently and were cultured in RPMI medium supplemented with 10% FCS, 100 U/ml penicillin and 100 µg/ml streptomycin.

3.1.3 T-lymphocytes (KE37/1-IIIB)

KE37/1-IIIB cells derive from chronically HIV-1 (HTLV-III$_B$) infected human T-lymphoma cell line KE37/1 (Popovic et al., 1984). KIII cells grow in suspension and were cultured in RPMI medium supplemented with 10% FCS, 100 U/ml penicillin and 100 µg/ml streptomycin.

3.1.4 Mouse hematopoietic precursor cell line (CloneB)

The mouse haematopoietic precursor cell cloneB was generated by transduction with the LTR-driven gamma-retroviral gene vector pSF91-GFP. The cell line was established in the lab of Christopher Baum (Hannover Medical School, Germany) (Modlich et al., 2006) and kindly provided from Manuel Grez (Georg-Speyer-Haus, Frankfurt, Germany). The cells grow in suspension and were cultivated in Iscove's modified Dulbecco medium (IMDM) including 10 ng/ml m-IL 3 (Murine Interleukin-3), 50 ng/ml m-SCF (Murine Stem Cell Factor), 4 mM L-Glutamin, 10 % FCS and 100 U/ml penicillin and 100 µg/ml streptomycin. It has to be taken into account that the cells tend to differentiate and change their morphology upon longer cultivation (Groß, 2009). Therefore, it is seriously recommended to apply experiments on cells being in early passages as it was done here.

3.2 Cell culture

3.2.1 Safety issues

Living HIV infected cells have to be cultivated in S3 laboratories (reasonable danger for human or environment;(GenTG, 2010)). Only after fixation the cells can be kept in S1 laboratories (no danger for human or environment; (GenTG, 2010)). Hence, the cells had to be cultivated, seeded and fixed under S3 standards in the lab of Ruth Brack-Werner (Institut für Virologie, Helmholtz Zentrum München, Neuherberg, Germany) and then transferred on ice to the Biozentrum in Martinsried in 1xPBS for further treatment.

3.2.2 Cell cultivation

The cell cultivation of the HIV infected cells was kindly accomplished by Manja Meggendorfer, Stephan Kremb and Ingrid Hülsmeyer (AG Ruth Brack-Werner) until seeding and fixation steps.

3.2.3 Sodium butyrate (NaBut) treatment

NaBut treatment was performed by 24h incubation at a final concentration of 0.5 mM immediately prior to fixation. In case of adherently growing cells NaBut incubation was combined with cell attachment to coverslips in quadriperm plates. In suspension growing cells were incubated with NaBut in culture flasks.

Alterations upon NaBut treatment were determined for virus particles in the cell supernatant and HIV transcriptional activity, each compared to untreated cells. Virus production was measured with a p24-ELISA test (HIV-1 p24-Antigen Capture ELISA). P24 is a HIV core protein and therefore correlating with the virus particle count. 10^5 cells were placed in a 12-well plate, 24h later NaBut was added (final concentration 0.5mM), another 24h later the supernatant was taken and Triton was added at a final concentration of 0.5% (further proceedings see protocol).

HIV transcriptional activity was determined via qPCR using the LightCycler 480 (Roche) and the Light Cycler 480 SYBR Green I Master PCR kit. HIV RNA levels were compared to RNA polymerase II RNA, a house keeping gene showing constant transcriptional activity. Cellular RNA was extracted with the RNeasy RNA Mini Kit, 1 µg RNA was transcribed into cDNA applying the Superscript First Strand Synthesis System for RT-PCR and 2 µl were used for the PCR (Primer sequences see 6.5).

Both p24-ELISA and LightCycler qPCR were kindly performed by Manja Meggendorfer (Helmholtz Zentrum München).

3.2.4 Seeding cells on coverslips

- **In suspension growing cells (T-lymphocytes, cloneB):**
  The suspension should show an adequate cell density having ~80% confluency when seeded on coverslips subsequently (if necessary the suspension was centrifuged and solved in less medium, increasing the cell concentration). Poly-L-Lysin coated coverslips (26x76mm) were laid into quadriperm plates and covered with 4 ml cell suspension. After 1 h at 37°C in the incubator, the coverslips were rinsed in 0.3x PBS for 45 sec. directly followed by fixation in 4% PFA/0.3x PBS (see 3.2.5).

- **Adherently growing cells (HeLa cells, Astrocytes,):**
  Cells were trypsinized and seeded on sterile coverslips (26x76mm) in a quadriperm plate. The cells grew over night until they showed ~80% confluency and then were fixed in 4% PFA/1x PBS (see 3.2.5).

3.2.5 Cell fixation

Adherently growing cells for DNA FISH:

Cells on coverslips were fixed by incubating the slide in 4% PFA/1xPBS (made immediately before use) for 10 min. at RT followed by washing 3x 5 min. in PBST and stored in 1x PBS at 4°C until subsequent treatment for 3D FISH (see 3.10.1).

MATERIALS AND METHODS

In suspension growing cells for DNA FISH:

Cells on coverslips were fixed in 4% PFA/0.3x PBS (freshly made immediately before use) for 10 min. at RT. The slides were washed 3x 5 min. in PBST and stored in 1x PBS at 4°C until subsequent treatment for 3D FISH (see 3.10.1).

For RNA FISH:

Distilled water in all buffers and solutions used in RNA FISH treatment was DEPC (Diethylpyrocarbonate) treated inactivating RNase.

Cells on coverslips were rinsed in 1x PBS, permeabilized 5 min. in 0.5% Triton in CSK buffer (incl. 10mM VRC) on ice, 45 sec. incubated in 0.3x PBS (only for in suspension growing cells), fixed in 4% PFA/10% acetic acid/1x PBS (or 4% PFA/10% acetic acid /0.3x PBS for in suspension growing cells, respectively) for 10 min., rinsed in 1x PBS and stored in 70% EtOH at 4°C until subsequent treatment for RNA FISH (see 3.10.6). In contrast to DNA FISH, where fixed cells can be stored in formamide for several months, it is recommended, for best results, to continue with the RNA FISH protocol within the next days.

3.3 Metaphase spreads

3.3.1 Fixation of in suspension growing cells for metaphase spreads:

- Per 1 ml cell suspension 10 µl colcemide (10µg/ml) were added (to increase the amount of metaphases, the cells should be in a well growing state).
- After 60 min incubation at 37°C (5%CO_2) the suspension was centrifuged for 10 min at 1000 rpm in a 50 ml falcon tube.
- The supernatant was removed except for 3 ml in which the pellet was resuspended.
- Carefully and slowly the falcon tube was filled up with pre-warmed (37°C) 0.56% KCl solution and incubated for 16 min. in a 37°C water bath.
- 10 min. centrifugation at 1000 rpm (immediately before centrifugation a few drops fixative were added).
- The supernatant was removed except for 3 ml in which the pellet was carefully resuspended.
- Carefully and slowly the falcon tube was filled up with fixative solution (methanol : acetic acid 3:1; -20°C) and stored at -20°C for at least 1h.

- After 10 min centrifugation (1000 rpm) the supernatant was removed except for 5 ml to resuspend the pellet.

3.3.2 Fixation of adherently growing cells for metaphase spreads:

- Cells grew in a T75 flask until 70-80% confluency. Per 1 ml medium 10 µl colcemide (10µg/ml) were added.

- After 60 min incubation at 37°C (5%CO_2) less adherent cells (particularly cells in mitosis) were removed by thorough beats with the hand against the flask.

- The cell suspension including the detached mitotic cells was transferred to a 50 ml falcon tube and centrifuged 10 min. at 1000 rpm (immediately before centrifugation a few drops fixative were added).

- The supernatant was removed except for 3 ml in which the pellet was carefully resuspended.

- Carefully and slowly the falcon tube was filled up with fixative solution (-20°C) and stored at -20°C for at least 1h.

- After 10 min centrifugation (1000 rpm) the supernatant was removed except for 5 ml to resuspend the pellet.

3.3.3 Metaphases dropping

- Microscopic slides were placed into a metal box floating in a 55°C water bath

- The cell suspension was dropped with a Pasteur pipette on the slides (two drops per slide); the water bath cover was closed immediately creating a special humid climate which allows the metaphases to spread (takes about 40 sec.)

- The metaphase spreading could now be observed and judged under the phase contrast microscope.

➔ Troubleshooting:
  - o Cell density to less: cell suspension was centrifuged (10 min, 1000 rpm), pellet was resuspended in less fixative increasing the concentration.
  - o Too much dirt on the slide: cell suspension was centrifuged (10 min. 1000 rpm), pellet was resuspended in fresh fixative and repeated if necessary (mark that a certain amount of dirt is also removed by subsequent pepsin treatment).

3.3.4 Pepsin treatment and aging

- 50 µl pepsin (10%) were mixed in 100 ml 0.01 N HCL (37°C)

- Metaphase slides were incubated in pepsin solution for 10 min. at 37°C

- 3x 5 min. 1x PBST washing step

- After an ascending ethanol series (70% and 90% at RT; 100% at -20°C; each for 5 min.) the metaphase slides were air dried.

- 1h incubation at 60°C

- Metaphase slides were stored with desiccant in a plastic box at -20°C.

3.4 Linker-mediated PCR (LM-PCR)

3.4.1 General considerations

Linker-mediated PCR (LM-PCR, also ligation-mediated PCR), first described in (Pfeifer et al., 1989), is a commonly used technique for the retrieval of a known sequence in an unknown genomic environment e.g. integration sites of viruses in the host genome (Schmidt et al., 2001; Kustikova et al., 2008). A requirement is that the integrated DNA sequence is known or at least its the flanking regions. Also a linker has to be designed. There are tons of different protocols generally based on the same principle with more or less variations indicated in the scheme below.
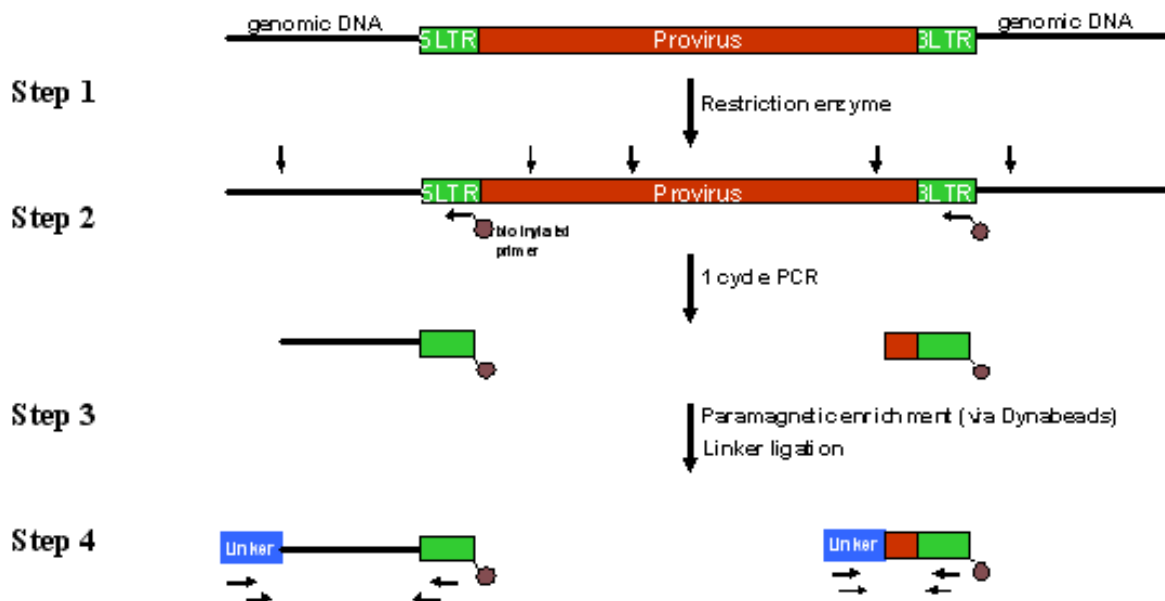
Figure 5: General principle of LM-PCR

Step 1: Digest

The genomic DNA isolated from a virus infected cell line is digested with a restriction enzyme which should fulfill several conditions: on the one hand it should cut with high frequency, on the other hand it should not have a recognition site in the 5´LTR since this is where the primer for the subsequent PCRs bind (recognition sites are indicated by perpendicular arrows in Figure 5).

Step 2: One cycle PCR

The first primer (in executed protocol (3.4.2) termed "Bio-LMPCR-1") has a biotin tag and binds in the proviral 5´LTR with orientation outwards the provirus towards the genomic DNA. After one cycle of a PCR with the biotinylated primer the emerging DNA fragments are marked with biotin and range ideally from the primer to the closest restriction site. In case that the restriction enzyme also cuts within the provirus (right branch in Figure 5), a second DNA fragment is marked with biotin as 5´LTR and 3´LTR have identical sequences (this fact could be used as some kind of positive control).

Strep 3: paramagnetic enrichment and linker ligation

After a paramagnetic enrichment via dynabeads a linker is ligated to the biotynilated DNA fragments.

Step 4: nested PCR and sequencing

With primers for the linker and the 5´LTR sequences, a nested PCR is accomplished. After a gel run of the PCR products, bands are cut out and directly sequenced. The sequencing result is blasted and shows ideally the host genomic region adjacent to the proviral 5´LTR.

3.4.2 Executed LM-PCR protocol

Materials:

- 3 sequence specific primer (stock solution: 10 pmol/µl) thereof one outer biotinylated primer (-> primer extension) and two linker specific primers (stock solution: 10 pmol/µl)
- Linker
- Dynabeads kilobaseBINDER Kit
- Quiageb PCR purification Kit
- Magnetic extractor (MPC)
- Rapid DNA Ligation Kit
- Quiaquick Gel Extraction Kit

MATERIALS AND METHODS

Sequences:

Linker +: 5'-GACCCGGGAGATCTGAATTCAGTGGCACAGCAGTTAGG**T** -3`

Linker -: 5'-Pho-CCTAACTGCTGTGCCACTGAATTCAGATCTCCCG - 3'

Linker specific primer for nested PCRs:

LSP1 (outer): 5' - GACCCGGGAGATCTGAATTC - 3'

LSP2 (inner): 5' - AGTGGCACAGCAGTTAGG - 3'

Provirus specific (HIV) primers:

Bio-LMPCR-1: 5´- TAACTTCTCTGGCTCAACTGGTAC - 3´

LMPCR-2 (outer): 5´ - TGGCCCTGGTGTGTAGTTCTG - 3´

LMPCR-3 (inner): 5´- AGGATATCTTGTCTTCGTTGGGAG - 3´

Method:

Linker-Annealing:

In a tube each 25 µl 100 µM Linker+ and 100 µM Linker-, incubation for 5 min. at 95°C, transfered to 70 °C Hotblock, switched of and let cool down to RT, 1 h at RT, transfered to 4°C fridge in a small Styrofoam box and incubated for 12h. Aliquots were stored at -20°C. Before use aliquots were slowly thawed on ice.

Digest of genomic DNA:

1µg DNA in 50 µl volume was digested with 5 U enzyme over night.

Following restriction enzymes were used all provided from NEB company:

AvrI, BamHI, BglI, DraI, EcoRI, HindIII, MseI, NheI, NlaIII, NotI, SacI, SacII, SmaI, SpeI, SphI, NotI, PstI, PvuII

Digest reactions were carried out as recommended in the NEB protocol. As control extracted DNA which already gave a positive result in a prior LM-PCR experiment was digested with enzyme MseI* and treated together with the other probes as described below (kindly provided from Manja Meggendorfer).

EtOH precipitation

Per tube: 100µl EtOH

   15µl 3M NaAC

   1µl glycogen (=20µg)

1h at -80°C, 30 min. centrifugation at 4°C, washed with 150 µl 70% EtOH by additional 15 min. centrifugation at 4°C. Pellet was air dried under the hood and dissolved in 37 µl ddH$_2$O.

Primer-Extension:

35.5 µl digested DNA

0.5 µl Taq polymerase (=2.5 U)

5 µl polymerase buffer (10x)

5 µl MgCl$_2$ (25mM)

2.5 µl Bio-LMPCR-1 (0.1 pmol/µl)

1 µl dNTPs (je 10 mM)

PCR program:

95°C 5 min

58°C 30 min

72°C 15 min

PCR products were purified with the Qiaquick PCR purification Kit and eluted in 42 µl ddH$_2$O.

Dynabead protocol:

20 µl Dynabeads were immobilized (see protocol). In the last step (6.) dynabeads were dissolved in 40 µl binding solution. 30 µl purified extension mix were added, 3 h incubation on shaker (750 U/min) at RT, washing steps see protocol (9.-11). Dissolving in 14 µl ddH$_2$O

Ligation:

0.5 µl T4 Ligase (NEB) = 200U

2 µl 10x Puffer (NEB)

14 µl DNA (Dynabeads)

1 µl Linker

3 µl H$_2$O

→ Incubation in the PCR tube at 16°C on shaker over night (alternatively, the Rapid DNA Ligation Kit from Roche could be used which only takes 3h of incubation). Washing step at MPC with 100µl dd H$_2$O, dissolving in 10 µl ddH$_2$O

MATERIALS AND METHODS

PCR experiments:

PCR1:

Reagent	Stock conc.	Final conc.	Vol per 50 µl
PCR buffer	10x	1x	5 µl
MgCl$_2$	25 mM	2,5 mM	5 µl
dNTP mix	10 mM	200 µM	1 µl
LSP1 (outer)	10 µM	500 nM	2.5 µl
LMPCR-2 (outer)	10 µM	500 nM	2.5 µl
Template DNA (ligation)	-	-	2 µl
Taq Pol	5 U/µl	0.1 U/µl	1 µl
ddH$_2$O	-	-	31 µl

Number of cycles	Reaction	Temperature	Time
1	Initial denaturation	95°C	5 min.
30	Denaturation	95°C	1 min.
	Annealing	58°C	45 sec.
	Extension	72°C	1.5 min
1	Final extension	72°C	10 min

PCR2:

See PCR1, though 1 µl PCR product from PCR1, 2.5 µl 10 pmol/µl LSP2 (inner),

2.5 µl 10 pmol/µl LMPCR-3 (inner), 33 cycles

Separation via gel

5 µl of each PCR result was analyzed on a 1.5% gel (see 3.5.1). The remaining 45 µl
of the lanes with promising bands were subsequently run on a preparative gel (see
3.5.2). Bands were extracted with the Qiaquick Gel Extraction Kit and directly
sequenced.

Following integration sites were mapped with this protocol:

**Astro 18q22.1, HeLa 6p12.3, HeLa 15q21.3, HeLa 11q22.3, HeLa 18q22.3, HeLa
Xq22.1, TLy 2q11.2**

3.4.3 Universal genome walker kit

Another possibility for integration site mapping is using the universal genome walker
kit (Clontech) in combination with the Advantage 2 PCR Kit (Clontech). It is also
some kind of LM-PCR but without a biotynilated primer and accordingly no
purification with dynabeads. The kit provides several positive such as negative
controls and is very easy to handle.

Following restriction enzymes were used:

- provided in the kit: DraI, EcoRV, PvuII, StuI

- additionally used enzymes (NEB company): BSaAI, FspI, HincII, NaeI, PshAI

38

Digest reactions were carried out as recommended in the protocol from the Kit or NEB company.

The following integration sites were mapped with this Kit (some were already mapped with the protocol from the previous chapter and confirmed with this protocol): **Astro 18q22.1, HeLa 16p13.3, TLy 2q11.2, TLy 2q14.2.**

3.5 Gel electrophoresis

3.5.1 Analytic gel

PCR results were visualized using 1.5% agarose gels. Therefore 1.5 g agarose were solved in 100 ml 1x TAE buffer by heating the solution in a microwave. After brief cooling the transparent solution was poured in the appropriate device forming the analytic gel by cooling down to room temperature.

The gel run at 150 mA and 80 V until the lowest marker band had reached two-thirds of the whole distance. The gel was stained for 10 min. in EtBr solution, briefly rinsed in distilled water and recorded under UV light.

3.5.2 Preparative gel

Depending on the length of the bands to be extracted between 0.8% (to separate larger fragments) and 2% (to separate shorter fragments) gels were used. Due to larger volumes of approximately 50 µl of the probes a larger comb was used to ensure that the whole volume could be put in one gel pocket. The running and staining conditions were the same as in 3.5.1. The bands of interest were cut out under UV light.

3.6 Sequencing of LM-PCR products and BAC clone selection

The extracted bands from the preparative gels were purified with the Quiaquick gel extraction kit. The purified DNA was sent for sequencing either to Sequiserve (Vaterstetten, Germany) or to GATC (Konstanz, Germany). To obtain the genomic position of the integration sites, sequencing results were blasted on the NCBI web page. In case of positive blast hits appropriate BAC (bacterial artificial chromosome) clones were chosen, that cover the genomic region of the integration site. BAC clones were selected using the "cytoview" display of the genome browser on the

ensemble web page and ordered from BAC PAC Resources (http://bacpac.chori.org/).

3.7 BAC clone preparation

Delivered as E.coli LB stocks, BAC clones were isolated via the method of alkaline lysis (Birnboim, 1983) using the Plasmid Miniprep Kit from Peqlab. The isolated DNA was eluted in 100 µl ddH$_2$O. Prior to labeling the BAC DNA was amplified using the illustra GenomiPhi V2 DNA amplification kit (Dean et al., 2001). This rolling cycle amplification method allows resulting DNA amounts of 5-7 µg from a minimum input of 10 ng DNA. The average amplification length of the fragments is about 10 kb. The use of exonuclease-resistant random hexamer primers in combination with the Phi29 DNA polymerase with its high accuracy and processivity and its excellent strand displacement activity are the characteristics of this effective procedure.

Correct chromosomal positions of the BACs were verified in 2D FISH experiments on human lymphocytes or mouse erythroleukemia (MEL) cell metaphase spreads with the respective chromosomal paint probe. If any BAC did not hybridize to the expected position, the BAC was rejected and a new one was ordered. This was twice the case.

3.8 pNL4-3 Plasmid isolation

Detecting the nuclear HIV DNA was implemented by using the pUC-based pNL4-3 plasmid (Adachi et al., 1986) as FISH probe. The pNL4-3 plasmid with 14833 bp in size contains the whole 9.7 kb HIV-1 proviral genome (plasmid map see 6.6). The plasmid was isolated from E.coli with the NecleoBond Kit from Macherey-Nagel. As probe the complete plasmid DNA was used.

3.9 RNase digest

For RNAse digest fixed cells on coverslips stored in 1x PBS were incubated in RNase A solution (200 µg/ml in 1xPBS) for 2h at 37°C. As control an additional coverslip with alike cells was incubated under same conditions in 1x PBS. The slides were washed 2x 5 min. in 1 x PBS and stored in it until further treatment.

3.10 Fluorescent in situ hybridisation (FISH)

FISH is a commonly used technique for visualizing nuclear RNA or DNA. Depending on the type of probe it is possible to visualize either chromosomal subregions (e.g. BAC / centromeric / telomeric probes) or whole chromosomes (chromosomal paint probes). A distinction can be drawn between 2D FISH on metaphase spreads, 3D FISH on structurally preserved nuclei and RNA FISH (for further details concerning the FISH technique see (Cremer et al., 2008))

3.10.1 Cell treatment in preparation for 3D FISH

Fixed cells on coverslips have to undergo a special treatment, amongst others allowing access of the FISH probe to the nuclear target DNA and preservation of the 3D nuclear morphology (protocol adapted and modified from (Solovei et al., 2002)):

- 20 min. incubation in 0.5% Triton X-100/PBS
- 1h incubation in 20% glycerol/1x PBS (o/n incubation also possible)
- Frozen in liquid nitrogen, thawed at RT and put back into 20% glycerol (five times)
- 7 min. incubation in 0.1 N HCl
- 3x 5 min. 1x PBST washing step
- 2x 3 min. 2x SSC washing step
- Storage in 50% formamide/2x SSC at 4°C until use (at least 48h).

3.10.2 Probe providing and labeling

FISH probe	Provided by
pNL4-3 plasmid (HIV)	AG Brack-Werner
pSF91-GFP (gene transfer vector);	AG Grez
BACs	BAC PAC Resources
Chromosome paints	AG Cremer
β-actin (Fitc labeled)	Christian Lanctot

Probes were either labeled via DOP-PCR or nick translation (protocols adapted and modified from (Müller et al., 2007)).

a) Label DOP PCR

One possibility for amplifying an unknown DNA sequence is the **D**egenerate **O**ligonucleotide **P**rimed PCR (Telenius et al., 1992). In this work DOP-PCR was used for labeling chromosome paint probes that were prior amplified via primary/secondary

MATERIALS AND METHODS

DOP PCR. For a master mix, the following reagents were mixed together and stored at -20°C:

Reagent	Amount	Final concentration
Cetus-II-buffer (10x)	100 µl	1x
$MgCl_2$ (25mM)	80 µl	2 mM
6 MW primer (100 µM)	20 µl	2 µM
ACG mix (2 mM each)	50 µl	100 µM
dTTP (1 mM)	80 µl	80 µM
Bio-/dig-/dnp-dUTP or TexasRed-dUTP	20 µl or 40 µl	20 µM or 40 µM
ddH_20	To 970 µl	-

For the labeling reaction, the following reagents are mixed together on ice in a 0.5 ml PCR tube:

Material	Amount
Master Mix	48 µl
DNA	2 µl
Taq polymerase (5 U/µl)	0.5 µl

PCR program:

Number of cycles	Reaction	Temperature	Time
1	Initial denaturation	94°C	3 min.
30	Denaturation	94°C	1 min.
	Annealing	56°C	1 min.
	Extension	72°C	1.5 min.
1	Final extension	72°C	8 min.

The size of the DNA fragments was checked with 2µl of the PCR reaction on a 1% agarose gel. In case of the smear not being between ~200 bp and 1.5 kb, an additional DNase digest as follows for 20 min. was performed.

Material	Amount
DNA (Label DOP-PCR)	50 µl
10x NT buffer	10 µl
ddH_2O	30 µl
DNase (1:1000 in H_2O)	10 µl

The reaction was stopped by freezing down the tube to -20°C and stored at this temperature until use.

b) Nick translation

Incorporation of labeled nucleotides in BAC probes and the pNL4-3 probe was accomplished by the use of nick translation. DNase (2000 U/ml) was diluted 1:250 in ice cold ddH$_2$O and kept on ice until use.

Reagent	Amount	Final concentration
DNA	~ 1µg	~ 0,02 µg/µl
NT buffer (10x)	5 µl	1x
β-Mercaptoethanol (100 mM)	5 µl	10mM
dNTP mix	5 µl	50 µM dATP/dCTP/dGTP each; 10 µM dTTP
Modified dUTP (1 mM)	2.5 µl of Bio/dig/dnp-dUTP or 5 µl of fluorochrom labeled dUTP	50 µM of bio/dig/dnp-dUTP or 100 µM of fluorochrom labeled dUTP
ddH$_2$O	to 48 µl	
DNase I (8 U/ml)	1 µl	0.16 U/ml
DNA polymerase I (10 U/µl)	1 µl	0.2 U/µl

The reaction mixture was incubated for 90 min. at 15°C in a water bath. A subsequent DNase digest was generally added since the length of the DNA fragments in previous nick translations had not shown the optimal range of 300-1000 bp. For this purpose 1 µl of diluted (freshly made) DNase I was added and incubated for additional 30 min. at RT.

Applied labeling:

Probe	Labeling
pNL4-3 (HIV probe)	Dig-dUTP Texas Red-dUTP
BAC probes / CT probes	Bio-dUTP Cy3-dUTP Dig-dUTP DNP-dUTP TAMRA-dUTP Texas Red-dUTP

Generally, due to fluorescence signal quality, BAC and CT probe labeling did not cause major circumstances in contrast to the HIV probe (see discussion). Therefore HIV probe labeling was the decisive element. Having a working HIV probe the labeling scheme for the respective BAC and CT probes was chosen.

3.10.3 Hybridization mixture (HM)

Depending on the type of probe and labeling, the table below shows the general composition of a 12 µl HM:

Material	Amount
DNA (labeled)	pNL4-3: 50 µl BAC (hapten labeled): 25 µl BAC (direct labeled): 50 µl CT* : 25 µl
Cot1-DNA (1 µg/µL)	per µl pNL4-3: 1 µl per µl BAC: 1 µl per µl CT: 0.5 µl
Salmon sperm DNA (10 mg/mL)	2 µl
Nick translation stop mix	1 µl

* If a DNase digest was performed (see 3.10.2), 50 µl of labeled CT DNA were used (because of the minor DNA concentration) and accordingly 0.25 µl Cot1-DNA per µl CT-DNA.

For further steps the HM is protected from light if direct labeled probes are used.

- Components were mixed by vortexing
- 2.5 vol. of ethanol abs. were added
- 30 min. incubation at -80°C
- 30 min. centrifugation at 14000rpm (preferentially at 4°C)
- Supernatant was discarded and the pellet was air dried in a vacuum centrifuge for 5-10 min.
- 6 µl of 100 % formamide were added
- The pellet was dissolved on a heat shaker at 38°C for at least 1h
- 6 µl HM master mix were added and the HM was put back on the shaker for additional 10 min.
- HM was stored at -20°C until use.

3.10.4 2D FISH procedure

2D FISH experiments were used on the one hand to check the correct chromosomal position of the BAC probes, on the other hand to confirm the mapped HIV integration sites. The latter was critical since the integrated HIV DNA is only about 9.7 kb in size. Therefore two different protocols were applied. In both cases the metaphase spreads stored at -20°C had to be thawed at RT.

a) 2D FISH to verify BAC positions

- HM was denatured 5 min. at 80°C in a water bath and briefly put on ice.

- 2 µl of the HM were put on the metaphase spreads and covered with a coverslip (8x8 mm) whose edges were then sealed with fixogum.

- Once fixogum was completely dried the slide was placed on a hot block at 72°C for 100 sec. denaturing the genomic DNA.

- Hybridization was ensured by incubation of the slide in a metal box floating in a 37°C water bath over night.

- For stringent washing steps and antibody detection see 3.10.5.

b) 2D FISH to confirm HIV integration sites

- HM was denatured 7 min. at 72°C in a water bath and afterwards incubated at 37°C for 1h allowing the ssDNA to pre anneal.

- Meanwhile the metaphases were denatured in 70% formamide/2x SSC (pH 7.0) at 72°C for 1.5 min. (70% formamide/2x SSC (pH 7.0) is stored at – 20°C and can be reused up to 4 times).

- After an ascending ethanol series (70% and 90% at -20°C; 100% at RT; each for 3 min.) the metaphase slide was air dried.

- 5µl of HM were put on the metaphase spreads and covered with a coverslip (12x12 mm) whose edges were then sealed with fixogum.

- Hybridization was ensured by incubation of the slide in a metal box floating in a 37°C water bath for 48h.

- For stringent washing steps and antibody detection see 3.10.5.

3.10.5 3D FISH procedure and antibody detection

By dint of 3D FISH it was possible to visualize HIV integration sites and their chromosomal neighborhood in structurally preserved interphase nuclei.

Fixed cells seeded on 76x26 mm coverslips were stored in 50% FA / 2xSSC, for 3D FISH an appropriate piece (app. 12x12 mm) was cut of using a diamond slicer.

- HM was denatured 5 min. at 80°C in a water bath and briefly put on ice.

- 5µl of HM are put on a slide, the coverslip of app. 12x 12 mm in size (if larger the amount of HM had to be increased) carrying the cells was placed on it and the edges were sealed with fixogum.

- Once fixogum was completely dried the slide was placed on a hot block at 72°C for 100 sec. denaturing the genomic DNA.

- Hybridization was ensured by incubation of the slide in a metal box swimming in a 37°C water bath for 48h.

- Having removed the fixogum with forceps the coverslip (3D FISH) / metaphase slide (2D FISH) was immediately put into washing buffer avoiding the cells to dry out.

- Stringent washing steps:
    - 3x 3 min. 2x SSC (37°C)
    - 3x 5 min. 0.1 SSC (62°C)
    - 1 min. 4x SSC (37°C)

- Blocking: 20min 4% BSA/4x SSC (37°C)

- 1. antibody: diluted in 4% BSA/4x SSC; incubation for 60 min. in humidified chamber at 37°C

- Washing steps: 3x 5 min 4x SSC (37°C)

- 2. antibody: diluted in 4% BSA/4x SSC; incubation for 60 min. in humidified chamber at 37°C

- Washing steps: 3x 5 min 4x SSC (37°C)

- Counterstaining: DAPI (500 µg/ml) diluted 1:100 in 4x SSC (37°C)

  -> duration: 2 min. (2D FISH) or 10 min. (3D FISH)

- 2D FISH slide was subsequently rinsed with dH_2O and air dried (lightproof). The hybridized area was mounted in vectashield, covered with a coverslip whose edges were sealed with nail polish. For 3D FISH the slide was briefly rinsed in 4x SSC, mounted in vectashield and placed on a microscope slide by sealing the edges with nail polish.

3.10.6 RNA FISH

RNA FISH was used to visualize actively transcribed HIV RNA. The major difference between RNA and DNA FISH is that in the latter case only the probe is denatured but not the nuclear DNA. Hence, only the single stranded RNA in the nucleus is hybridized by the probe. The protocol was adapted and modified after (http://www.singerlab.org/protocols). For this purpose cells were fixed in a special kind of way and stored in 70% EtOH (see 3.2.5).

- First the cells were rehydrated by incubating in 50% FA/2x SSC for at least 30 min. (not longer than 2h)

- The composition of the HM was the same as for DNA FISH (see 3.10.3) except for the addition of VRC (*V*anadyl ribonucleoside complex) to a final concentration of 10 mM.

- After denaturing (5 min. at 80°C in water bath) the HM was incubated 30 min. at 37°C in the water bath.

- An adequate amount HM was put on a slide, the coverslip carrying the rehydrated cells was placed on it and the edges were sealed with fixogum.

- Hybridization was ensured by incubatiing the slide in a metal box swimming in a 37°C water bath for 48h.

- Stringent washing steps:
    - 2x 3 min. 2x SSC (37°C)
    - 3x 5 min. 50% FA/2x SSC (42°C)
    - 2x 5 min. 4x SSC (37°C)

- Blocking: 10 min. 4% BSA/4x SSC (37°C)

- Antibody detection see 3.10.5 except for the addition of VRC to a final concentration of 10 mM in each antibody dilution.

3.11 Immunocytochemistry in combination with FISH

This method was used to visualize distinct DNA or RNA FISH signals together with SC35 splicing speckles. The SC35 antibody had to be diluted in PBS buffer whereas the antibodies for FISH signal detection were diluted in SSC buffer. Therefore, the buffer had to be changed which was the only crucial step in the protocol. When the cells were in 4x SSC washing buffer after the last (in principle the second) antibody incubation for FISH signal detection, the protocol was continued as follows (in case of RNA FISH VRC was added to a final concentration of 10 mM in each antibody dilution):

- Washing step from FISH detection:
    - 2x 5 min. 2x SSC (37°C)

- Change buffer:
    - 2x 5 min. 1x PBS (37°C)

- Blocking: 10 min. 4% BSA/1x PBS (37°C)

- 1. antibody: MαSC35 diluted 1:100 in 1x PBS (37°C, 1h)

- Washing step:

- – 3x 5 min. 1x PBS
- 2. antibody: diluted in 1x PBS (37°C, 1h)
- Washing step
  - – 3x 5 min. 1x PBS (37°C)
- Counterstaining: DAPI (500 µg/ml) diluted 1:100 in 1x PBS (37°C, 10 min.)

3.12 Microscopy

3.12.1 Phase contrast microscope

The phase contrast microscope was used in the cell culture for checking the cell growth and density to estimate in which dilution they have to be seeded on coverslips. For 2D FISH experiments metaphase slides where checked under the phase contrast microscope for suitable areas for hybridization. For technical properties see 6.13.

3.12.2 Epifluorescence microscope

Results of 2D FISH, 3D FISH and of other fluorescence labeling experiments were checked under the epifluorescence microscope. Images were recorded using the Metaview software with a CCD camera and the 63x oil objective in adequate zoom. For technical properties see 6.13.

3.12.3 Confocal laser scanning microscope (CLSM)

The confocal microscope was used to record nuclei for 3D reconstructions and investigations on nuclear architecture. The scanning process was accomplished in the sequential scan mode to avoid Bleed-through.

The table below displays the main scanning parameters.

Voxel size (x,y,z)	80 nm x 80 nm x 240 nm
Line Average	3
Image size	256 x 256 (TLy, HeLa) / 512 x 512 (astro) pixels
Pinhole	1 airy unit

The filter ranges were adjusted to avoid light emitted from other channels. The photomultiplier settings (gain and offset) were chosen such that both over- and underexposure was prevented, covering the whole grey value scale of 0-255 (8-bit

image format). The laser excitation power was adjusted to highest possible value not causing obvious photobleaching. For technical properties see 6.13.

3.13 Image processing, evaluation and statistics

3.13.1 Chromatic shift correction

For each scanning process, microscopic files were exported as 8-bit tiff-series. Additionally the info files which contain all scanning parameters (e.g. voxel size, laser settings, microscope settings, etc.) were saved in the same main folder (folder name "raw series"). The program ImageJ 1.37c was used for image processing. The first processing that had to be accomplished was the correction of the chromatic shift. Chromatic aberration is a distortion in optical systems and arises from the different refraction properties of light with different wavelengths. Shift correction was performed with a special plugin (Shift corrector by Boris Joffe, LMU München, Germany) while synchronistically in this process the tiff-series were converted into image stacks resulting in one stack for each channel. The outcoming images were saved in an extra folder ("shifted stacks").

3.13.2 Deconvolution

Only in case of splicing speckle experiments the images were deconvolved directly after shift correction and prior to further processing such as threshold defining. Because of the optical properties of a microscopic system out-of-focus light from above and below causes blur and creates a distorted image. In case of heavy blur fine structures are hardly illustrated. A possibility to reduce blur is the mathematical process of deconvolution by which the blurred light is calculated back to its origin resulting in a sharper image. Deconvolution was helpful to improve illustration of SC35 splicing speckles. Images of experiments, where the distance of BAC signals to the nuclear surface was measured, were not deconvolved. Threshold for both BAC and DAPI signal, which represents the nuclear surface, can be defined appropriately in "raw" images without deconvolution. In SC35 experiments all images including BAC, HIV and DAPI channel were deconvolved. Deconvolution was implemented using the Huygens Essential 3.3.0p0 program running under Linux. The deconvolution parameters were generated from a measured point spread function (PSF). More details concerning the deconvolution process see (Albiez, 2007). The

deconvolved images were treated further as described in the following three chapters.

3.13.3 Preliminary work on image files

The images of the BAC channel showed two (or three; depending on the number of chromosomes) distinct signals: one that co-localizes with the HIV signal and one (or two, respectively) on the homologous chromosome(s). In advance to further image analysis one had to distinguish between BAC signal co-localizing with HIV and not. Therefore the following steps were proceeded with the BAC channel images in ImageJ: First the whole image stack was duplicated (image -> duplicate). In one of the images the BAC signal that co-localizes with the HIV was cut out, in the other image the BAC signal(s) on the homologous chromosome(s) was (were) cut out resulting in one image stack for each BAC signal. Both images were saved separately.

3.13.4 Threshold defining

For subsequent image analysis one has to define signal and background for each stack by setting a threshold. In advance all images were modified with the Gaussian blur filter (process -> filters -> Gaussian blur; sigma(radius) was set on 1) to reduce pixel irritations and to create a slight soft-focus effect. Afterwards all images were normalized with the plugin stack normalizer (available on ImageJ homepage) assuring that the whole grey value range from 0-255 was covered. The modified shifted image stacks were again saved in an extra folder ("bearbeitet"). Finally the thresholds for each image stack could be defined and inscribed. It is important to note that now only true signals should be found in the images. All background signals whose grey values were above the threshold had to be cut out otherwise it would irritate the measurement and give false positive results

3.13.5 Enhanced absolute distance to surface (eADS)

The 3D measurement program was eADS developed by T. Thormeyer (Thormeyer, 2005; Albiez et al., 2006) who enhanced and improved the existing ADS program (absolute distance to surface) developed by Dr. J. von Hase (Kirchhoff Institut, Heidelberg). With the new eADS program (which is by now an application in EDMT: enhanced distance measurement tool) it is possible to define, voxel by voxel, the shortest distance in 3 dimensions from a signal to a reference signal. If the nuclear

position of a BAC signal, i.e. the distance to the nuclear surface, was determined, the counterstain was used as reference channel. Figure 6 illustrates the measurement procedure. Green and red circles indicate BAC signals for which the shortest distance to the nuclear surface in 3D should be determined. It is now measured from the respective BAC signal, voxel by voxel, in defined shells in 3 dimensions until the first contact with any voxel of the reference channel. In principle, as reference channel almost any signal can be chosen such as centromeres, telomeres or nucleoli. In this thesis besides the radial nuclear location of BAC signals also the relation to the BAC harboring chromosome territory and to splicing speckles was determined. In these cases either the chromosome territory or SC35 splicing speckles were used as reference channels. The resulting evaluation file can be opened in Excel by using the function "import external data". For illustration adequate graphs can be generated.
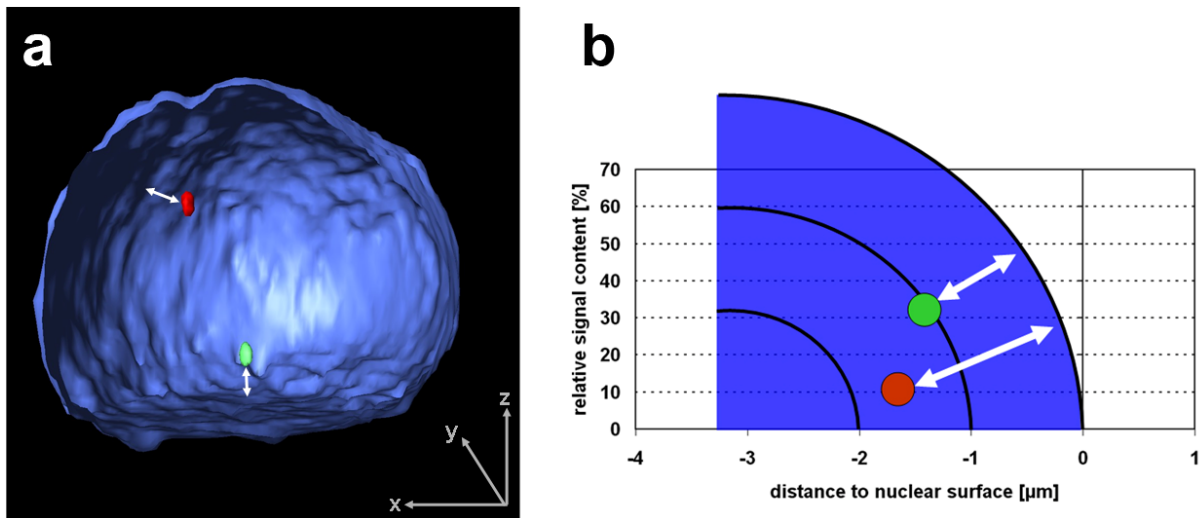
Figure 6: eADS performance illustration
a – 3D Amira reconstruction; view inside a nucleus; nuclear surface in blue, BAC signals in red and green; white arrows indicate shortest distance to nuclear surface.
b – corresponding excel evaluation graph from eADS measurements; with a quarter of the nucleus the scale of the x-axis is exemplified; shortest distance of the BAC signals (red, green) to the nuclear surface is indicated by arrows; negative values indicate location within the reference signal, here the nucleus; positive values indicate location outside reference signal.

Creating the mask (using ImageJ):

The DAPI counterstain reveals less and also not stained areas like nucleoli or accumulations of open chromatin. If a threshold is set it may happen that these areas is omitted. So, if now a signal is positioned nearer to such an internal nuclear relief than to the nuclear surface, eADS measures the wrong distance leading to a false

51

positive result. By using a nuclear mask as reference channel the addressed scenario can be avoided:

- A threshold for the counterstain channel was defined and applied (Image -> Adjust -> Threshold (set threshold) -> Apply). In doing so one had to mind that in the ImageJ options background color is white and foreground color is black.
- The omitted holes were filled (Process -> Binary -> Fill Holes).
- The image was inverted (Edit -> invert).
- The resulting mask was saved as "m_Kern(n).tif" in the same folder as the modified shifted image stacks (folder name "bearbeitet").

The folder "bearbeitet" now contained all processed shifted stacks and the masks for each nucleus.

Generating the parameter file (using EDMT filemaker):

eADS needs a parameter file that caries all important information like for example the thresholds for each image stack, the storage paths for the input and output files or the voxel size. Here, one also defines the reference channel and which channels are measured against that. To receive the chromatin distribution, the DAPI channel as well was measured against the reference channel (mask).

Running the eADS program:

EDMT was started, Ultra Rapid Distance Measurement was chosen and the parameter file was loaded. The output file with the results was imported in Microsoft Office Excel, where graphs with the BAC signal and DAPI distribution were created. For further details and troubleshooting see (Thormeyer, 2005; Albiez, 2007; Kupper, 2007; Kupper et al., 2007).

3.13.6 Determination of surface area with ImageJ

The surface area of BAC signals was measured with the object counter 3D described in (Bolte and Cordelieres, 2006) and downloaded from http://imagejdocu.tudor.lu/doku.php?id=plugin:analysis:3d_object_counter:start. Blur in images was reduced using the Gaussian filter (Sigma (radius)=1) in advance. Subjective influence was minimized by normalizing the BAC signals with the stack normalizer plugin and by applying a constant threshold of 50 (Astrocytes, HeLa cells), 100 (T-lymphocytes) or 130 (cloneB cells). The number of surface pixels was plotted in Excel graphs.

3.13.7 Statistical evaluation

A putative statistical significance was ascertained by using an appropriate significant test. Generally, the paired t-test was used, a pair consisting of the values for the integration site and the homologous site (averaged when two were present: HeLa18q22.3, Astro18q22.1) from the same nucleus. For radial nuclear positioning, the median values were used. Some distributions were not normally distributed, thus the Wilcoxon signed rank (WSR) test for paired samples instead had to be applied (radial distributions of MMUXA1.1, HeLa18q22.3, HeLaXq22.1, Astro18q22.1, surface pixels HeLa18q22.3). Distributions in untreated and sodium butyrate treated cells were compared with the Mann‑Whitney rank sum test. All calculations were performed with SigmaStat 3.5.

3.14 Determination of gene density, GC content and other genomic properties

Except for GC content the genomic properties were directly read of the map viewer on the NCBI homepage (http://www.ncbi.nlm.nih.gov/) after having blasted the respective integration site sequences. From there the sequences surrounding the integration sites were downloaded and submitted to the Repeatmasker homepage (http://www.repeatmasker.org/) to obtain the GC content. The maximum sequence length possible to upload to the Repeatmasker homepage is approximately 5 Mb. Hence, properties of the 10 Mb window were computed by two uploading steps. Total genomic GC content was taken from (Venter et al., 2001) (human genome) and from (Waterston et al., 2002) (mouse genome). Genomic properties were determined in a 0.5, 2 and 10 Mb window around each integration site.

4. RESULTS

4.1 HIV Integration site mapping and karyotype analysis

HIV integration sites were investigated in astrocytes, HeLa cells and T-lymphocytes (for details about cell types and infection see 3.1). The first step was to identify the genomic position of the integrated provirus via LM-PCR (see 3.4). For each identified integration site a BAC clone was ordered that covers the respective integration site. After testing the BAC clones for correct chromosomal position on metaphase spreads, 3D FISH experiments were performed.

4.1.1 Mapping of the HIV integration sites via LM-PCR

The extracted DNA from each cell type was digested using varying restriction enzymes (see 3.4). After linker ligation distinct DNA fragments were amplified. One primer binds in the 5´LTR of the provirus, the other primer in the linker. The amplified DNA fragment in between revealed the genomic host DNA adjacent to the proviral 5´LTR. The analytic gels of the LM-PCR products revealed distinct and bright bands (e.g. both lanes in Figure 7c), weak bands with DNA smear (e.g. Figure 7e lane four and six) or lanes with DNA smear (e.g. Figure 7a lane one).

To allow sequencing of the PCR product, bands had to be distinct and separate from each other and contain a certain amount of DNA. Therefore, only the assays from lanes which satisfy these parameters were subsequently put on a preparative gel (method see 3.5.2). In such a case the complete remaining volume (approximately 45µl) was applied. Bands of interest were cut out and sent for sequencing (see 3.6). An example of a preparative gel before and after extraction of bands is given in Figure 8.

Sequencing results were blasted on NCBI web page. In case of no blast hit the sequence was either internal of the provirus (see Figure 5) or not assignable (red rectangled bands in Figure 7). For the investigated astrocytes, HeLa cells and T-lymphocytes nine HIV integration sites were identified, seven of them were located within a gene (Table 1). It has to be noted here, that the T-lymphocytes harbor more than the two identified integration sites since in subsequent FISH experiments HIV signals were visible, that did not co-localize with one of the BAC signals (see Figure 23).
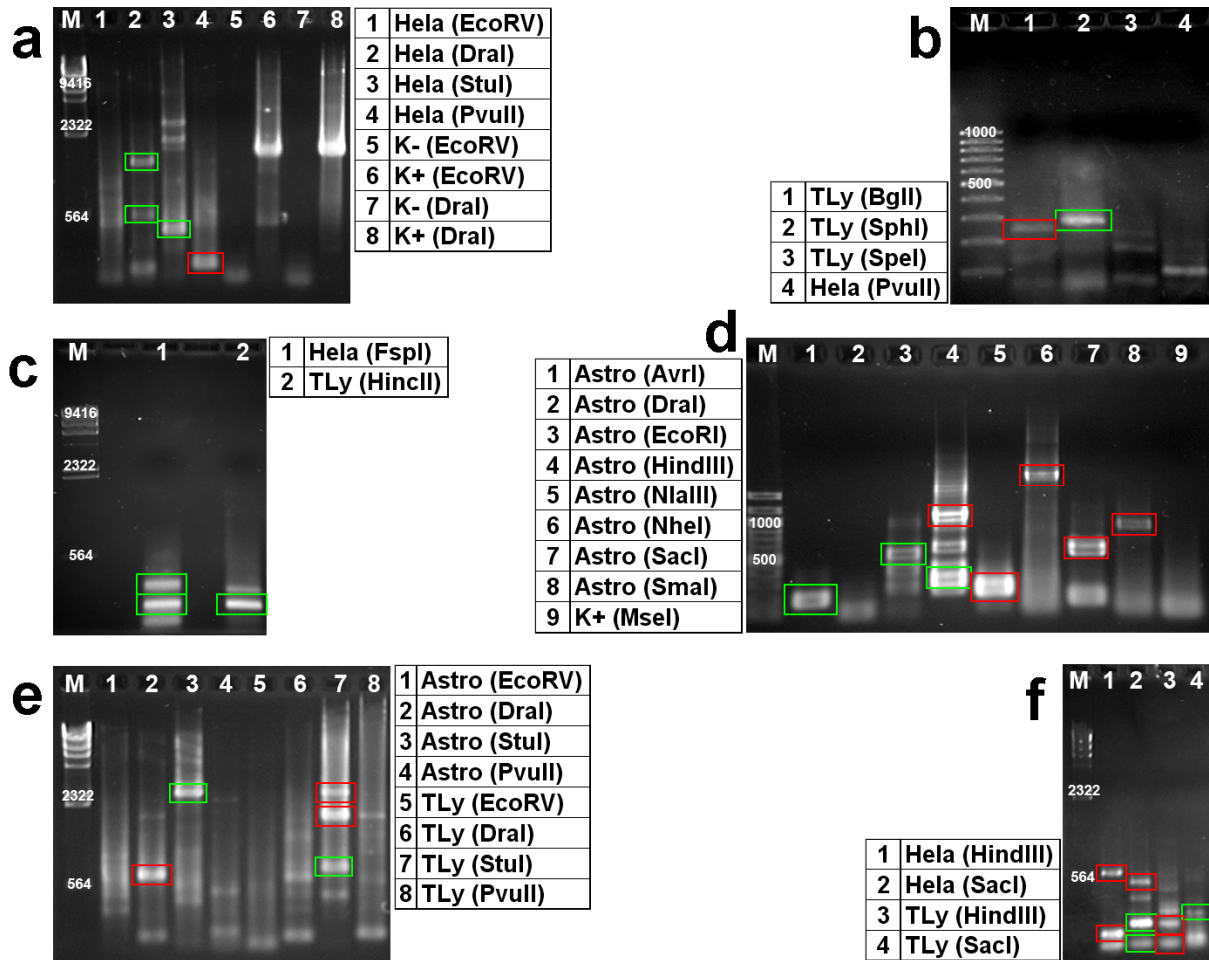
Figure 7: Analytic agarose gel examples of LM-PCR products
M – marker (numbers in bp); K- - negative control (without DNA); K+ - positive control (contain human genomic DNA provided by the Genome Walker Universal Kit; in brackets each used restriction enzyme; Rectangle indicated bands were cut out after a subsequent preparative gel run and sequenced; green rectangles mark bands which revealed integration sites, red rectangles indicate bands with no results in terms of integration site position; 5µl of each PCR product was analysed on the agarose gel.
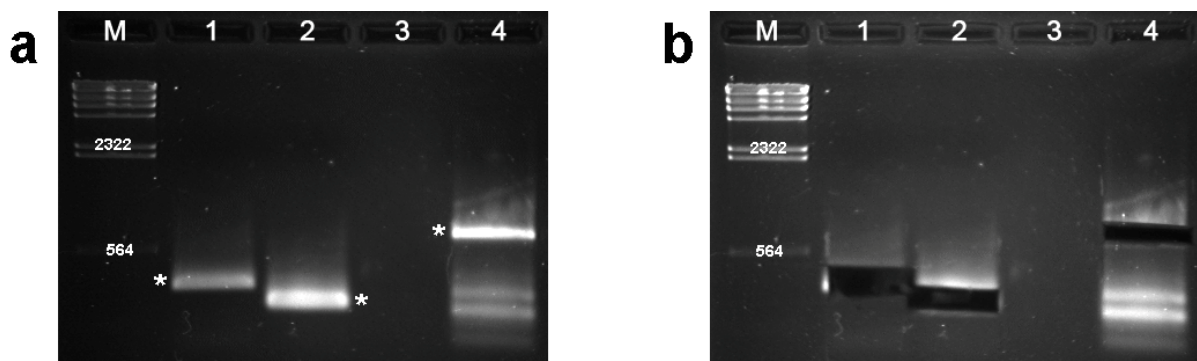
Figure 8: Preparative gel of LM-PCR products
Restriction enzyme assays of astrocyte DNA with HindIII (lane one), NlaIII (lane two) and SacI (lane four), respectively; lane three is empty; a – before bands marked by asterisks were extracted, b – afterwards; M – marker (numbers in bp)

Integration site	Astro 18q22.1	HeLa 6p12.3	HeLa 11q22.3	HeLa 15q21.3	HeLa 16p13.3	HeLa 18q22.3	HeLa Xq22.1	TLy 2q11.2	TLy 2q14.2
Nearest gene	CDH19	SUPT3H	PDGFD	ZNF 280D	AXIN1	NETO1	CSTF2	KIAA 1310	PTPN4
Distance to nearest gene	0	0	250kb	0	0	700kb	0	0	0
Gene ID	28513	8464	80310	54816	8312	81832	1478	55683	5775
sequence	ATAACA ATATTG TATAAT ATATGA AGAAAT	GCAAGT TTTCAGT TACTGTG CTTGATA GGC	AAATC ATTAA GTTGTT AATGG AATTTT AAG	CAGAA AAATG CAAAT AGGGT CATTTT AGTC	GGGCTTT CCCTAG GTTCTGT TCATCAG GTG	CATAGA GCAAG GGGCC AACAGA CCCCCT TA	ACAG CACA ACTG GATA CCAC TGAA TTAAT C	GGTGG GTCTG GACCA CTCCT GGATT CTGGG	TGCCT TCCTC AGACT TTGTT ACAGC ATAGG

Table 1: Genomic positions of HIV integration sites
Overview of HIV integration sites in three different human cell types. HeLa – HeLa cells; TLy – T lymphocytes; Astro – astrocytes. Description behind the cell type defines the chromosomal region of the integration site. A distance of 0 to the nearest gene means the integration is inside the gene; gene ID from NCBI human genome database; sequence – start of sequenced host genome from LM-PCR directly adjacent to 5´LTR of integrated provirus.

4.1.2 Control PCRs confirmed integration site mapping results and absence of tandem copies

To better characterize HIV in general and to confirm mapping results, integration sites were subjected to control PCRs. For this purpose, specific primers were designed, one binding within the provirus 5´LTR and the other 1 kb away within the genomic host DNA (sequence see 6.5). In case of a correct mapping result a distinct band with 1 kb in size should be visible on the gel. As candidates the integration sites HeLa 11q22.3, HeLa 18q22.3, HeLa Xq22.1 and Astro 18q22.1 were chosen. In all lanes, except for the negative controls, the expected 1 kb bands appeared indicating the correct mapping result (Figure 9). It has to be noted, that PCRs were performed in a S1 lab. Therefore, digested DNA assays had to be used and not genomic DNA (see 3.2.1). Correct mapping of integration sites can also be confirmed in subsequent FISH experiments by co-localization of HIV signal and BAC signal, which marks the integration site region. Thus further PCRs were not necessary.

To test whether multiple integrations were arranged consecutively, PCR with specific primers was performed. For this purpose, the primer "GSP2" binding within the 5´LTR from the LM-PCR experiments (see 3.4.2) and a primer binding within the 3´LTR ("GSP3LTR_100") were used. In case of two proviruses integrated directly next to each other with the same orientation, a fragment of approximately 150 bp in size should be amplified as indicated in Figure 10. For each cell type two digestion assays (EcoRI and SacI) of the extracted genomic DNA were tested (Figure 11).
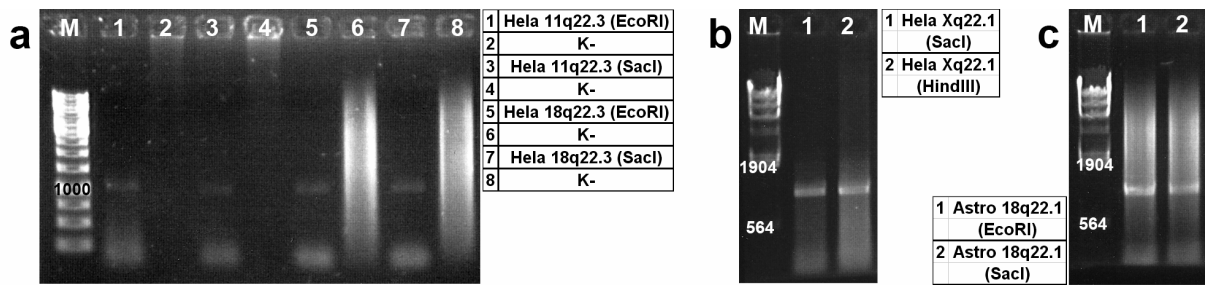
Figure 9: Control PCRs to confirm LM-PCR results
M – marker (numbers in bp); negative controls (K-) contain PCR assay from respective left lane without primers; for each integration site the EcoRI and the SacI restriction enzyme assay were analyzed; a – integration sites HeLa 11q22.3 and HeLa 18q22.3, b – integration site HeLa Xq22.1, c – integration site Astro 18q22.1.
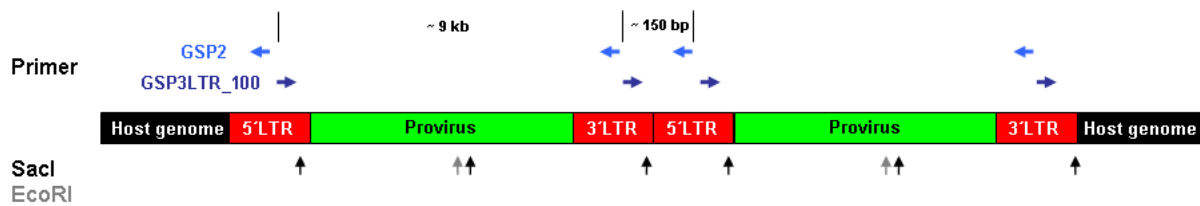
Figure 10: Primer position illustration for control PCRs
Two consecutively into host genome (black bars) integrated proviruses (green bars) flanked by their LTR regions (red bars) are shown; due to identical sequences of the LTRs both primers GSP2 (light blue) and GSP3LTR_100 (dark blue) bind, opposedly orientated, within the 3´LTR and the 5´LTR; primer binding sites indicated by horizontal light blue (GSP2) and dark blue (GSP3LTR_100) arrows; primer sequences see 6.5; putative amplifiable fragment lengths, delineated by vertical black lines, are displayed; restriction enzyme recognition sites indicated by vertical black (SacI) and gray (EcoRI) arrows.
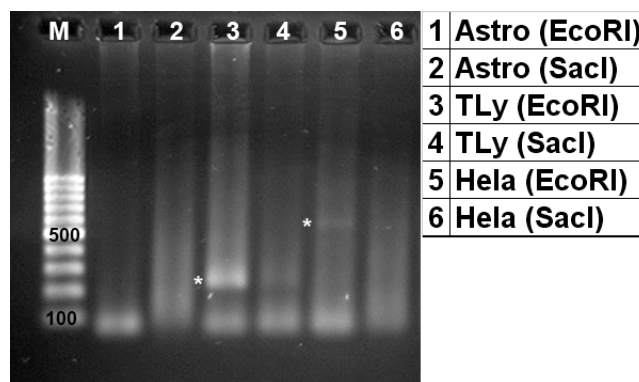
Figure 11: Control PCRs to test multiple integration sites arranged consecutively
1.5% agarose gel; M – marker (numbers in bp); Astro – astrocytes; TLy – T-lymphocytes; HeLa – HeLa cells; EcoRI, SacI – restriction enzymes; asterisks mark bands; 5µl of probe was applied.

Contrary to the EcoRI digestion assay, the SacI assay should never reveal an amplified fragment of 150 bp, since SacI has a recognition site within the LTR region (see vertical black arrows in Figure 10. The predicted bands at approximately 150 bp

did not appear in one the EcoRI assays, but in lane 3 (T-lymphocytes) at 250 bp and in lane 5 (HeLa cells) at 600 bp weak bands were visible (Figure 11). This means that both in T-lymphocytes and in HeLa two HIV integrations could be in immediate vicinity (250 bp and 600 bp far away from each other, respectively). It cannot be said here which integration sites were concerned. For this, PCRs with specific primers are necessary. It should be taken into account that two neighboring proviruses might be orientated in opposite directions which was not considered in the primer design here. The blurry bands shorter than 100 kb in each lane of Figure 11 apparently derive from the primer.

Both EcoRI and SacI have recognition sites within the provirus (perpendicular arrows in Figure 10). The amplification of an additional fragment of ~9 kb in size that covers the whole integrated provirus is consequently not possible and was also never found in one of the lanes in Figure 11.

4.1.3 BAC clones were tested for correct chromosomal position

For each identified integration site appropriate BAC clones were selected which cover the respective genomic region. All BAC clones were tested for correct chromosomal binding by FISH on metaphase spreads. If the provided BAC clones were loacetd differently as expected, allternative clones were selected, which was once the case (HeLa 18q22.3). Finally, the following BAC clones were used for further FISH experiments:

Integration site	Astro 18q22.1	HeLa 6p12.3	HeLa 11q22.3	HeLa 15q21.3	HeLa 16p13.3	HeLa 18q22.3	HeLa Xq22.1	TLy 2q11.2	TLy 2q14.2
BAC clone	RP11-831H17	RP11-818O21	RP11-63H12	RP11-566D24	RP11-517F15	RP11-693I21	RP11-255J06	RP11-67L23	RP11-132N24

Table 2: BAC clones covering HIV integration sites
Overview of BACs which were used for further FISH experiments; HeLa – HeLa cells; TLy – T lymphocytes; Astro – astrocytes. Description behind the cell type defines the chromosomal region of the integration site

4.1.4 Due to chromosomal rearrangements, not all identified integration sites could be evaluated

An essential condition for detecting possible changes in the nuclear position with and without viral integration is the same genomic constitution for the harboring chromosome and its homolog. Since all used cell types are immortalized or cancer derived, chromosomal aberrations are to be expected. Putative chromosomal

aberrations were investigated in 2D FISH experiments with chromosomal paint probes and respective BAC probes for each cell type and integration site.

The results of the 2D FISH mapping revealed several chromosomal abnormalities summarized in Table 3. Only integration site harboring chromosomes were analyzed. HSA18 in astrocytes, HSA2 in T-lymphocytes and HSA18 in HeLa cells were of normal appearance (see Figure 11a,b and Figure 13e). Thus, the respective integration sites 18q22.1 in astrocytes, 2q11.2 and 2q14.2 in T-lymphocytes and 18q22.3 in HeLa cells could be evaluated in 3D FISH experiments with regard to their nuclear architecture. In HeLa cells, among the integration site harboring chromosomes HSA6, HSA11, HSA15 and HSA16 an abnormal constitution was determined. Of each of these chromosomes three or four copies of different sizes were found, often including translocations (see Figure 13a,b,c,d).
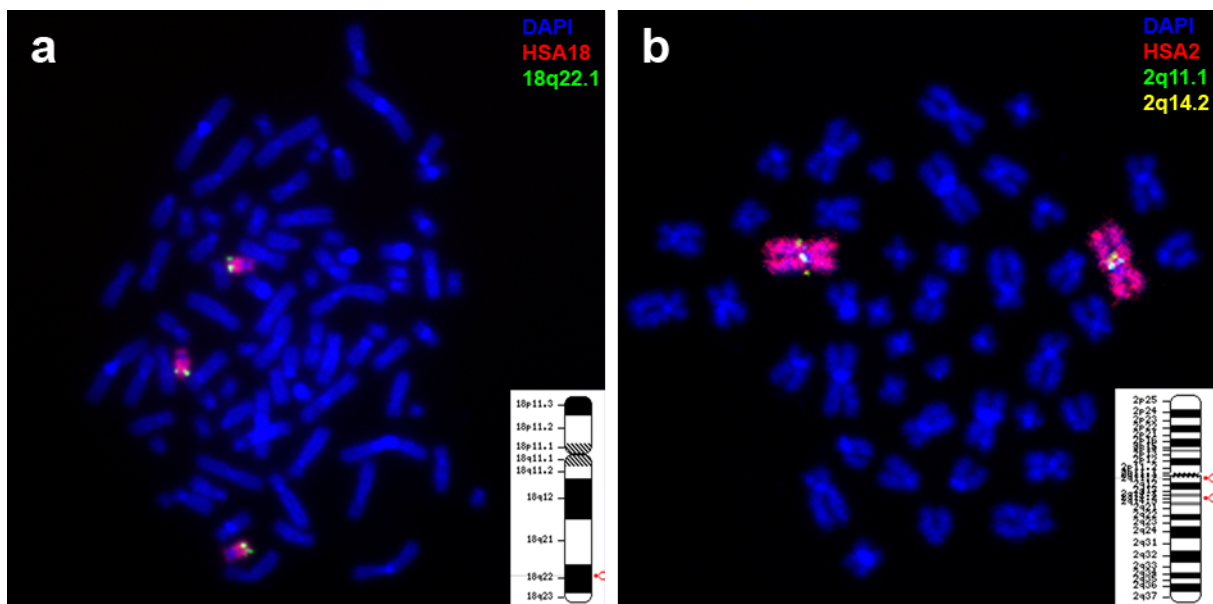
Figure 12: Chromosomal analysis of astrocytes and T-lymphocytes
Overlay of widefield microscopic images (a - astrocytes) and z-projections of confocal image stacks (b - T-lymphocytes); FISH on metaphase spreads; blue – DAPI; red – chromosome paint probe (HSA18 and HSA2); green – BAC probe marking respective HIV integration site (18q22.1 on astrocyte metaphases and 2q11.1 on T-lymphocyte metaphases); yellow – BAC probe marking HIV integration site 2q14.2 on T-lymphocyte metaphases.
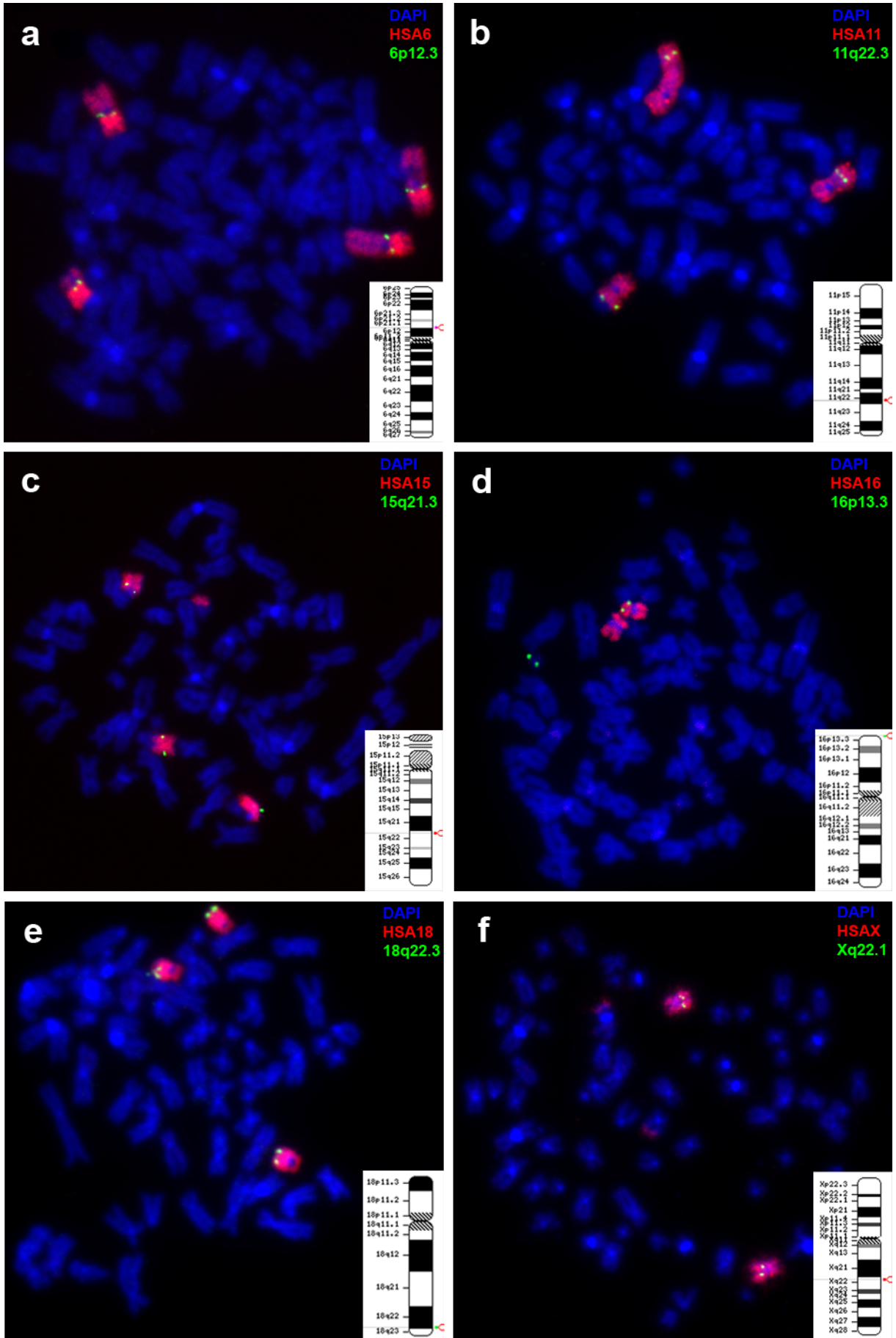
Figure 13 (previous page): chromosomal analysis of HeLa cells
Overlay of widefield microscopic images of FISH on metaphase spreads; blue – DAPI; red – chromosome paint probe (HSA6, HSA11, HSA15, HSA16, HSA18, HSAX); green – BAC probe marking respective HIV integration site (6p12.3, 11q22.3, 15q21.3, 16p13.3, 18q22.3, Xq22.1); red mark in each ideogram indicates respective integration site position; karyotypes are not unconditionally complemented since each image does not show all chromosomes of a metaphase spread or ideogram covers some chromosomes in the image.

Taken together, only for five of the nine integration sites the relevant chromosomes were of normal appearance. Compared to the ideograms, BAC signals showed the correct chromosomal positions in all cases. The integration site HeLa Xq22.1 showed a special situation (see 4.1.5). Results were verified by at least ten metaphase spreads per BAC. When performing 3D FISH experiments, for each batch of cells the chromosomal constitution was controlled by FISH on metaphases, since further chromosomal aberrations during cell culture cannot be excluded.

Cell type	Integration site	Chromosomal constitution	Evaluable?
HeLa	6p12.3	Four chromosomes 6 of 3 different sizes	no
HeLa	11q22.3	Three different chromosomes 11	no
HeLa	15q21.3	Three different chromosomes 15 (two "homologous", one larger); additionally one translocated region	no
HeLa	16p13.3	Translocation of one of the two BAC binding regions	no
HeLa	18q22.3	All three chromosomes 18 show normal constitution	yes
HeLa	Xq22.1	Both X chromosomes show normal constitution; additionally two small translocations	? (see 4.1.5)
T-lymphocytes	2q11.2	Both chromosomes 2 show normal constitution	yes
T-lymphocytes	2q14.2	Both chromosomes 2 show normal constitution	yes
Astrocytes	18q22.1	All 3 chromosomes 18 show normal constitution	yes

Table 3: Evaluation of FISH on metaphase spreads

4.1.5 LC5-HIV HeLa cells have no inactive X chromosome

One of the mapped integration sites in the female HeLa cell line was located on the X chromosome. In normal female human cells, one of the two X chromosomes is inactived resulting in a condensed heterochromatic structure, the Barr body, frequently located at the nuclear rim (Barr and Bertram, 1949). Shape and gene positioning of active and inactive X chromosome can differ significantly (Eils et al., 1996; Dietzel et al., 1999). Hence, it had to be clarified whether one of the two X chromosomes in the HIV infected HeLa cell line was inactive. The inactive X

chromosome can be distinguished from the active by specific epigenetic properties, such as accumulation of the histone variants macroH2A or 3meH3K27 (Chow and Heard, 2009).

Figure 14 shows, that the LC5-HIV HeLa cell line does not carry an inactive X chromosome, since neither an accumulation of 3meH3K27 nor of macroH2A histone modifications was detectable. Neither did the cells display a Barr body. In the female control cells an accumulation of both histone modifications and therefore an inactive X chromosome was found (HFb images in Figure 14). Thus, all X chromosomes in this cell line were active.
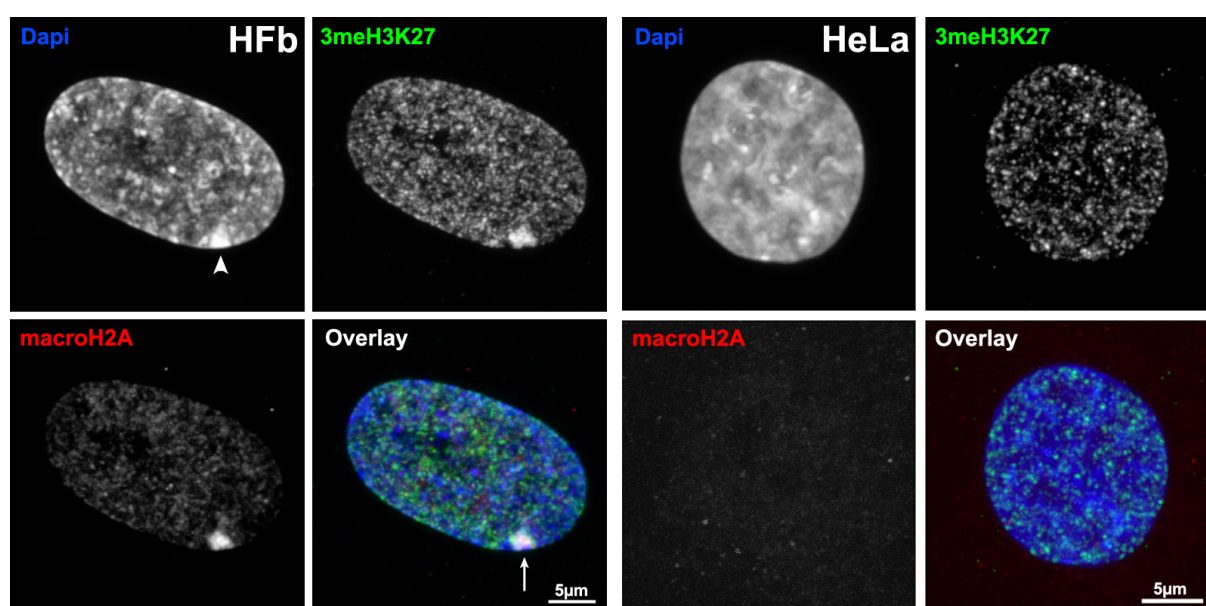
Figure 14: Immunostaining to ckeck for inactivated X chromosome
Z-projections of confocal image stacks; arrow indicates inactivated X chromosome in HFb (human fibroblasts) cells; arrowhead indicates the Barr body in the DAPI channel; arrow in overlay image indicates the inactive X chromosome; of the HeLa cell population 50 cells were checked.

4.1.6   Gene density and GC content in the integration site environment

For further characterization of the genomic environment of the integration sites, gene density and GC content in 0.5, 2 and 10 Mb windows around each integration was determined (Figure 15). An obvious tendency of integration sites towards gene dense or GC-rich regions was not detectable. Integration sites were found both in gene dense regions (HeLa 16q13.3, HeLa Xq22.1, TLy 2q11.2) and in gene poor regions (HeLa 11q22.3, HeLa 18q22.3, Astro 18q22.1). Integration sites also located in regions with both high and low GC content.

Gene density
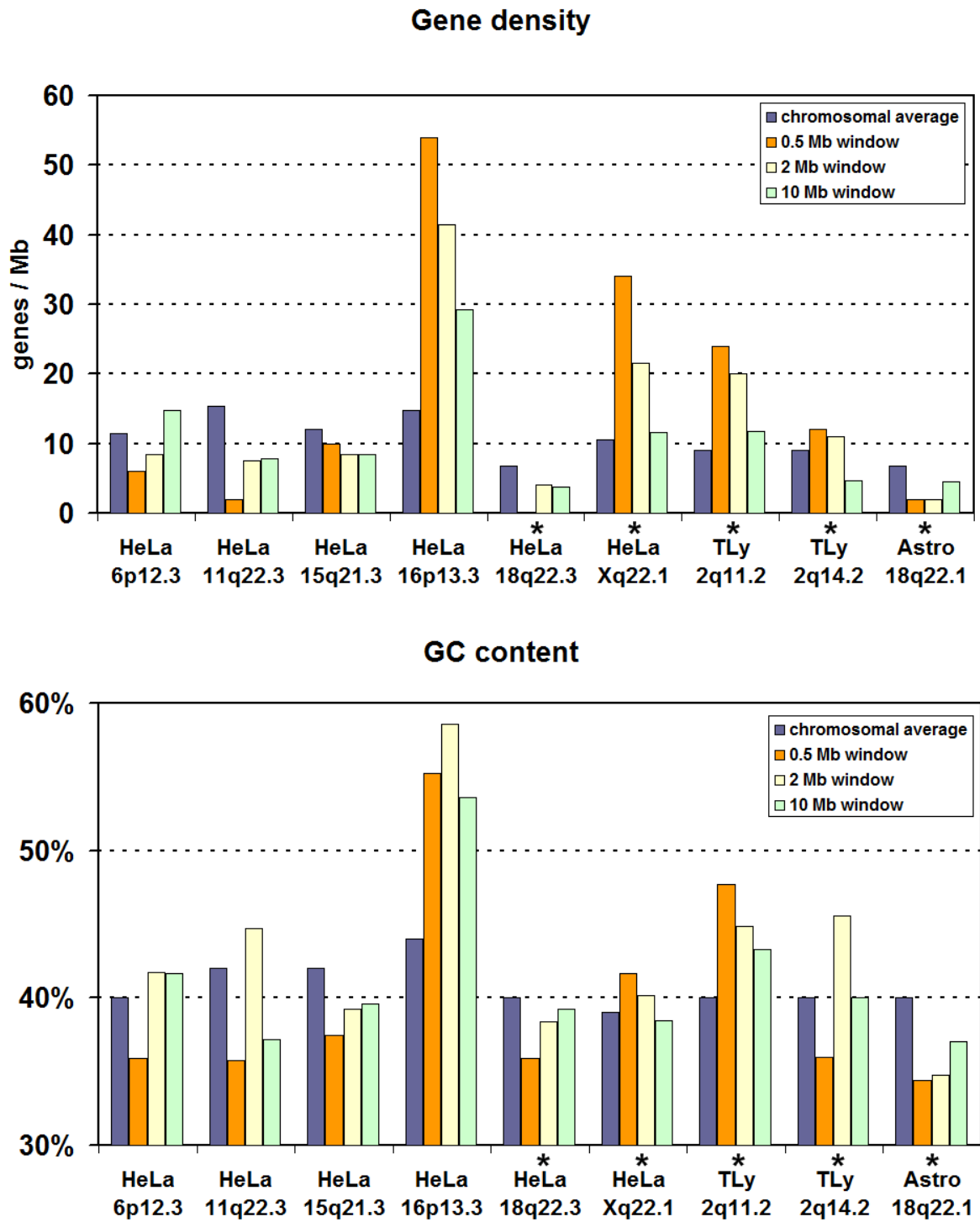
GC content

Figure 15: Gene density and GC content of integration site regions in HIV infected cells
Method see 3.14; evaluable integration sites indicated by asterisks.

4.1.7 2D FISH experiments revealed co-localization of BAC and HIV signal

FISH on metaphase spreads of HeLa cells with BAC marking integration site Xq22.1 and HIV probes showed co-localization of both signals (Figure 16). This was a very challenging experiment, since only in less than 5% of the recorded metaphase spreads a HIV signal was distinctly visible. During recording at the microscope, the HIV signal was not identifiable at the screen. Only after images were normalized concerning the grey value range, the HIV signal could be localized by co-localization with the BAC signal that marks the integration site.

Comparing the BAC and HIV signal in Figure 16b, the difference in size (BAC ~150 kb, HIV ~10 kb) becomes apparent. In consideration of the fact that the putative co-localization of HIV signal and its corresponding BAC signal was finally obvious in 3D FISH experiments, no further 2D FISH experiments of this kind were performed.
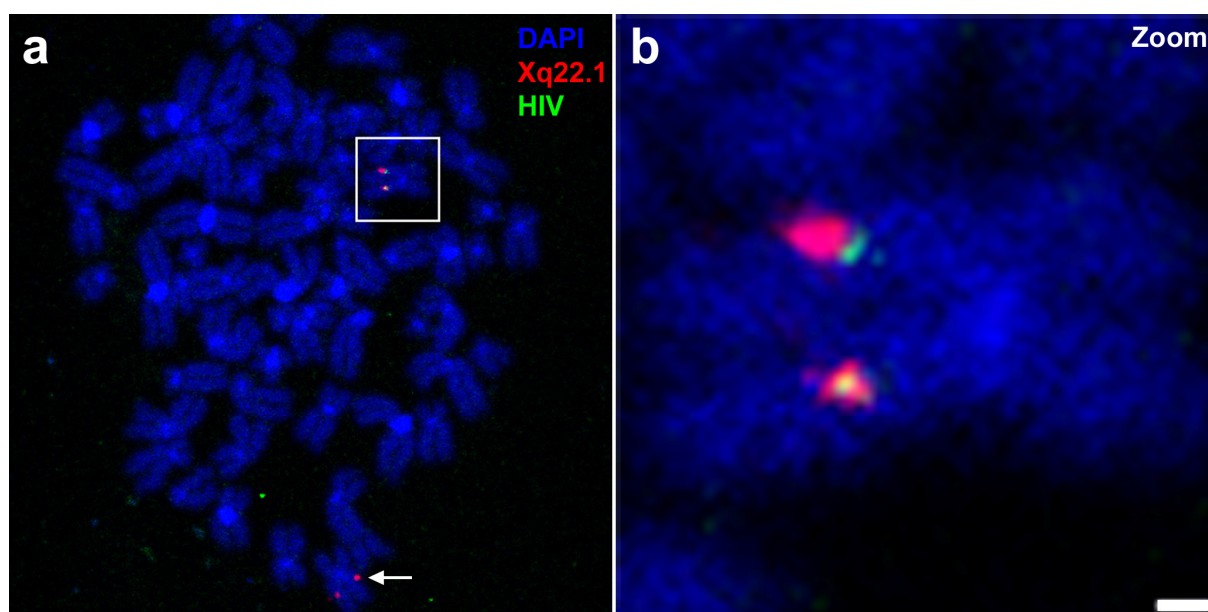
Figure 16: FISH on metaphase spreads of HeLa cells
Projection of three confocal sections; b – Zoom of region marked by square in a; a, b: blue – DAPI, red – BAC marking integration site Xq22.1, green - HIV; homologous X chromosome without HIV integration marked by arrow in a; scale bar 0.5 µm.

4.2 FISH on intact HIV infected interphase nuclei

FISH with probes for HIV, BACs marking the integration sites and chromosomal paint probes was performed on structurally preserved nuclei. Depending on the cell type, nuclei with different counts of HIV FISH signals were found.

The HIV infected astrocytes, HeLa cells and T-lymphocytes were characterized in advance with regard to the number of integrations per nucleus via LightCycler qPCR (Manja Meggendorfer, Helmholtz Zentrum München, personal communication). The results predicted one provirus copy per astrocyte, six copies per T-lymphocyte and 0.6 copies per HeLa cell nucleus.

4.2.1 T-lymphocyte nuclei showed diverse morphologies

As mentioned above HIV infected astrocytes, HeLa cells and T-lymphocytes were investigated in this thesis. The first two show rather uniform fibroblast-like nuclear morphologies throughout a cell population. However, T-lymphocytes showed large variation in nuclear shape from round to strechted or asymmetric (Figure 17). Hence, for objective 3D measurements, nuclei had to be selected. Only the ones were chosen for scanning that show the average morphology of the cell population. An example for such a nucleus is indicated by arrow in Figure 17.
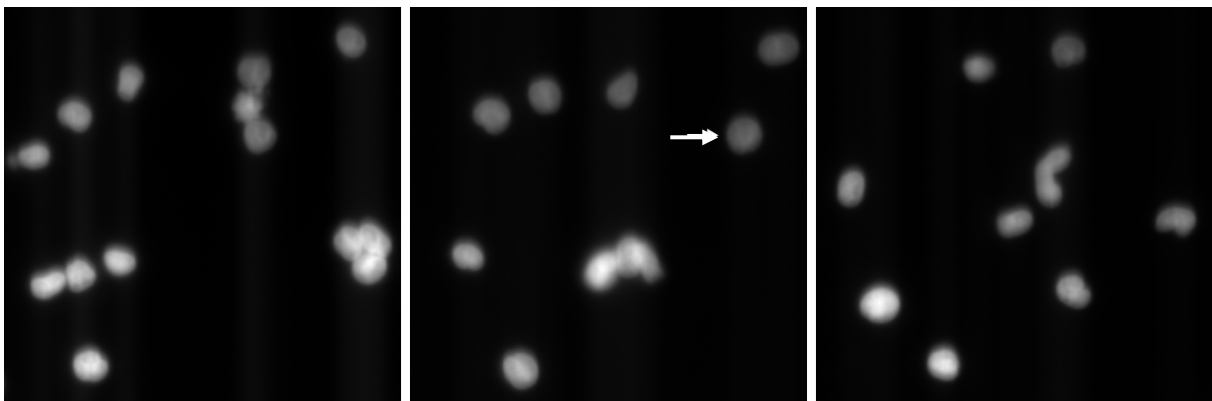
Figure 17: nuclear morphologies of KE37/1-IIIB T-lymphocytes
Three widefield microscopic images; DAPI staining; the three images exemplify the variety of nuclear morphologies among the T-lymphocyte population; the difference in brightness due to different DAPI staining intensities can be neglected, only the nuclear shape and morphology are crucial; arrow marks a nucleus with averaged morphology.

4.2.2 The HIV FISH probe detected HIV DNA and RNA

With a target size of 9.7 kb the HIV provirus is at the lowest detection limit for FISH probes. The HIV signal after 3D FISH on interphase nuclei of all three cell types was surprisingly strong (Figure 18). One would expect an at least 10 times larger BAC signal compared to the HIV signal, due to the lengths of sequences detected by both probes (~150 kb and 9.7 kb, respectively). A potential explanation for the strong HIV signal was that the HIV probe also detects HIV RNA. This assumption was verified by RNase A digest (Figure 19). The HIV signal was clearly smaller after RNase digest (compare HIV signal -RNase and +RNase in Figure 19) indicating the loss of RNA, which otherwise contributed to the FISH signal. DAPI staining and BAC signal quality were not influenced by the treatment. The smaller HIV signal (indicated by arrow in Figure 19) is hardly identifiable without the co-localizing BAC signal which marks the integration site. HIV FISH signal comparison with and without RNase A digest was performed with all three HIV infected cell types. In all cases the HIV signal was smaller with treatment.

An RNA FISH experiment with HeLa cells and T-lymphocytes gives information about the transcriptional levels of the HIV provirus. Probes for HIV and the house-keeping gene β-actin were co-hybridized. β-actin is known to be involved in cell motility, structure and integrity and should therefore be expressed at microscopic detectable levels. Comparing the HIV and β-actin RNA signals in HeLa cells in Figure 20 the strength of the HIV RNA signal and therefore the grade of HIV transcription becomes obvious. In T-lymphocytes a β-actin RNA signal was hardly visible. In summary, it can be concluded that the five investigated HIV integration sites were strongly transcribed.
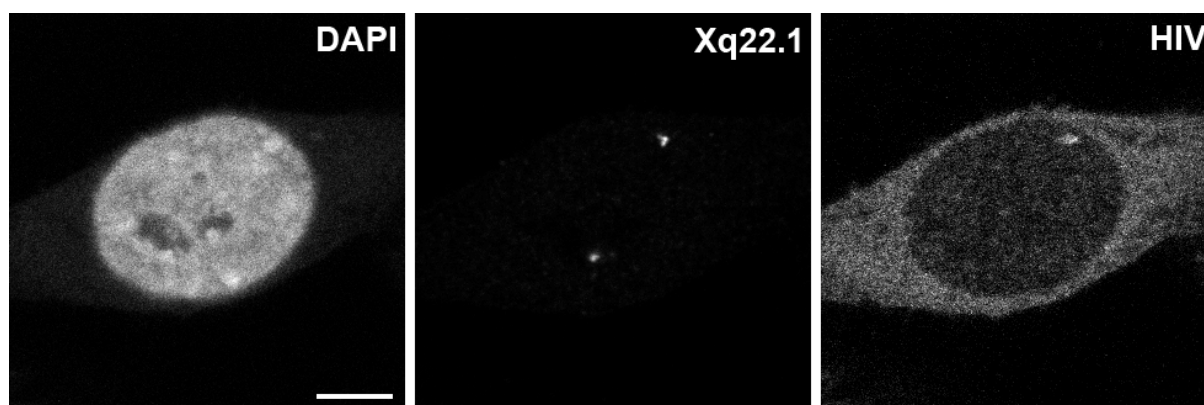
Figure 18: 3D FISH on interphase nucleus
Mid confocal sections; 3D FISH on HeLa cell nucleus with BAC probe (Xq22.1) marking the respective integration site and HIV probe; nuclear DNA stained with DAPI; all images are normalized with the stack normalizer plugin in ImageJ so that the whole gray value range from 0 – 255 in each image is used; scale bar 5 µm.
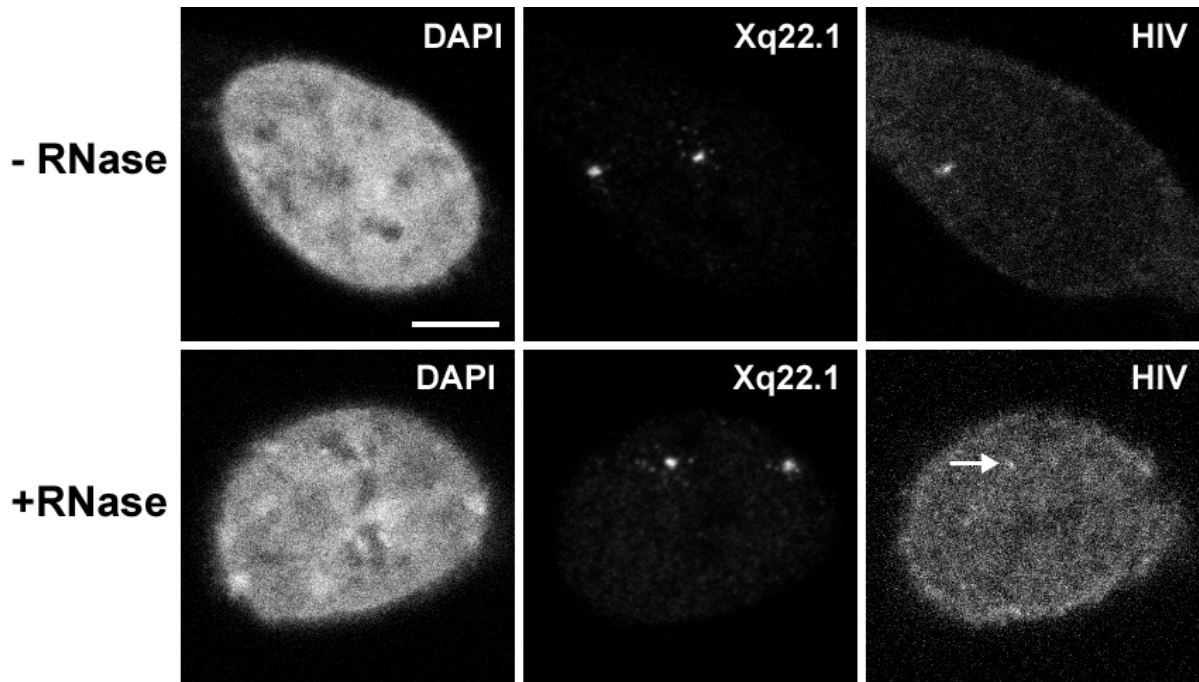
Figure 19: RNase A digest
Mid confocal sections; 3D FISH on HeLa cells with BAC probe (Xq22.1) marking the respective integration site and HIV probe; nuclear DNA stained with DAPI; control (-RNase; upper panel) and RNase treated (+RNase; lower panel); arrow indicates HIV signal; all images are normalized with the stack normalizer plugin in ImageJ so that the whole gray value range from 0 – 255 is used; this increases background signal in the HIV channel after RNase digest; for each cell type (HeLa cells, astrocytes and T-lymphocytes) approximately 30 cells were compared; scale bar 5 µm for both panels.
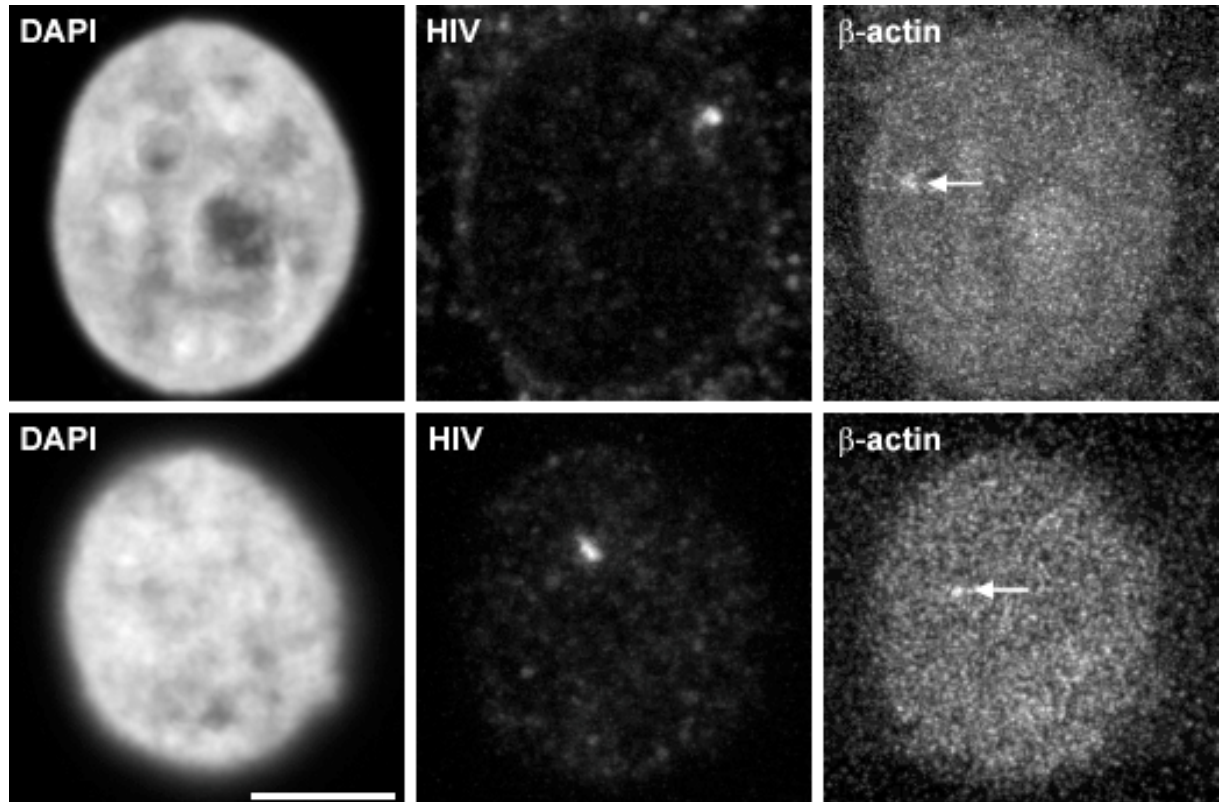
Figure 20: RNA FISH with probes for HIV and β-actin
Z-projections of three mid confocal sections; two different HeLa cell nuclei after RNA FISH with probes for HIV and β-actin; arrows in β-actin images indicate RNA signals; all images are normalized as described for Figure 19; scale bar 5 µm.

4.2.3 Some cell types showed multiple integration sites

To investigate the nuclear architecture of HIV integration sites, 3D FISH on interphase nuclei was applied. HIV and a BAC marking the respective integration site were used as probes. In some cases the integration site harboring chromosome territory was also stained. Figure 21 and Figure 23 show examples for each investigated integration site.
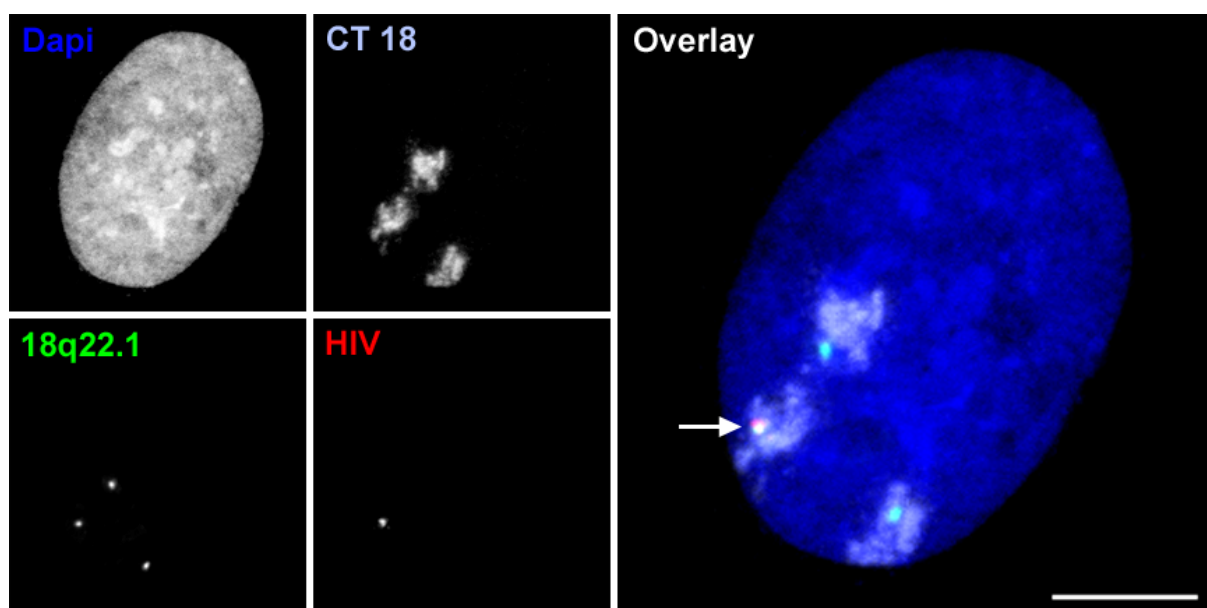
Figure 21: 3D FISH on astrocyte interphase nucleus
Z-projections of confocal image stacks; Overlay image: light blue – chromosome territory 18 (CT18), green – BAC marking integration site (18q22.1), red – HIV signal, blue – DAPI counterstain. Co-localization of HIV and BAC signals indicated by arrow; scale bar 5 µm. To benefit HIV FISH signal quality and to reduce background signal, the brightness/contrast tool in ImageJ was applied for the HIV channel.
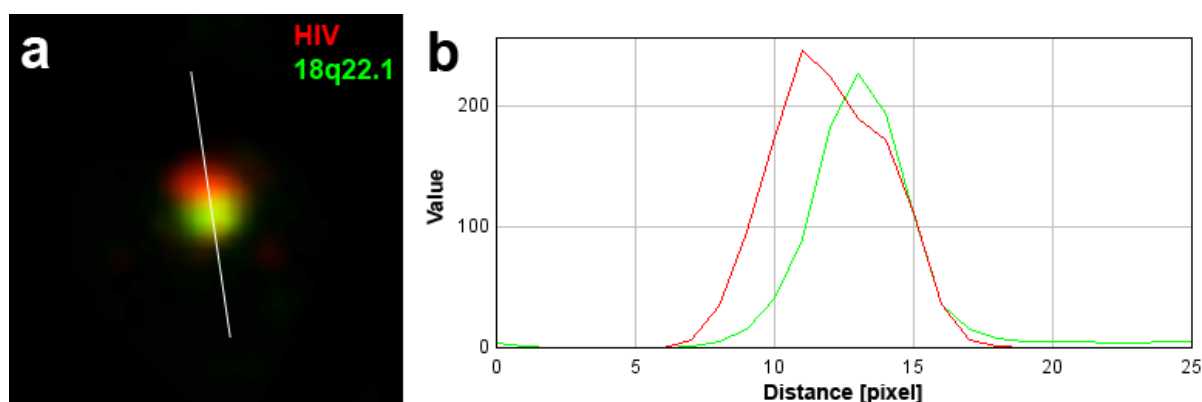
Figure 22: line scan of HIV and co-localizing BAC signal
a: microscopic image; magnification from overlay of HIV (red) and BAC (18q22.1, green) signals from Figure 21; b: line scan analysis from microscopic image in a; red curve – HIV signal intensity, green curve – BAC (18q22.1) signal intensity; value – gray value (0-255); distance in pixel.
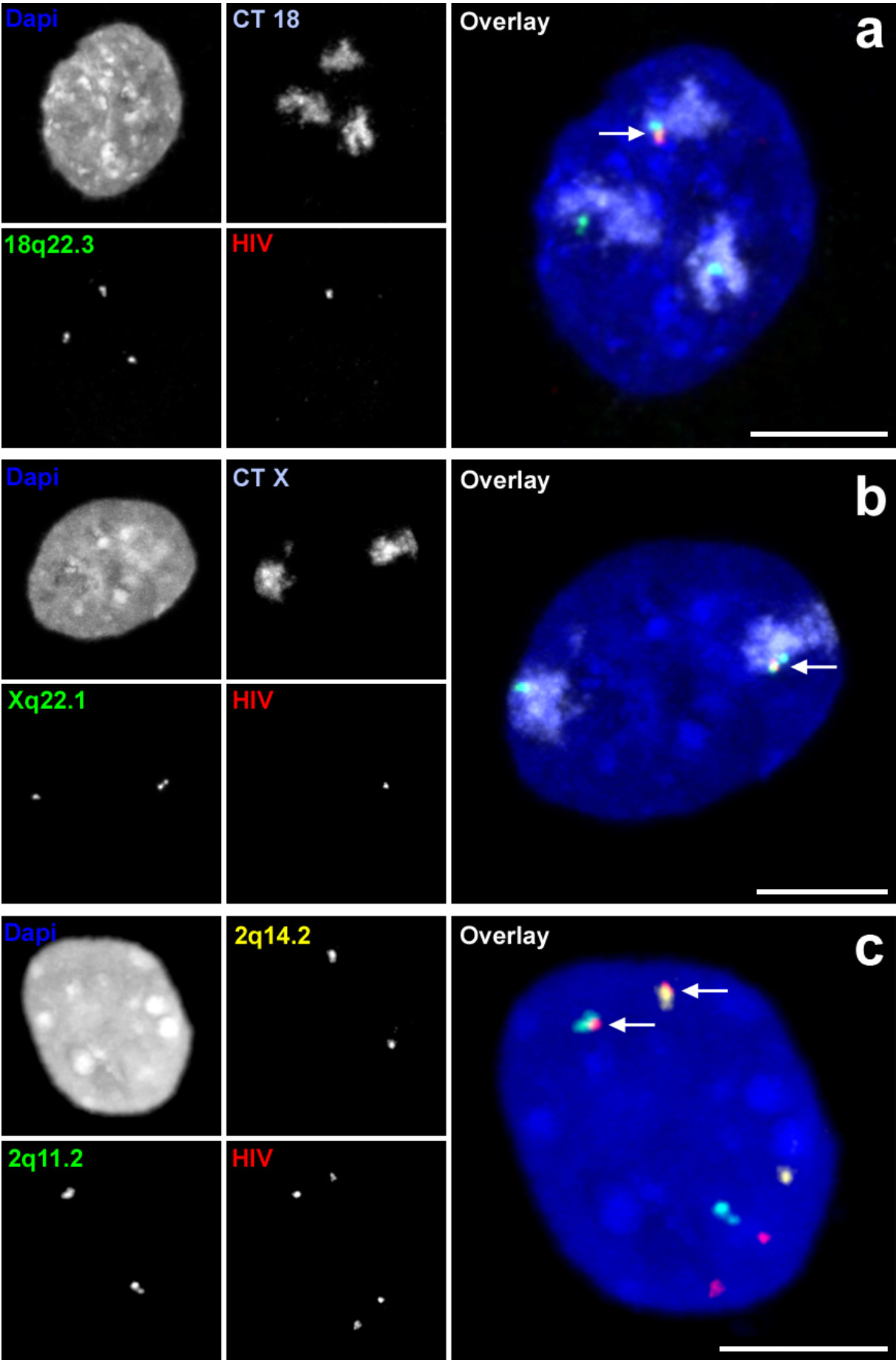
Figure 23: 3D FISH on interphase nuclei of HeLa cells and T-lymphocytes

Figure 23 (previous page): Z-projections of confocal image stacks;
a,b: HeLa cell nuclei; green – BAC signals (18q22.3; Xq22.1); light blue – chromosome territories 18 and X (CT18, CTX); red – HIV signal; Co-localizing BAC and HIV signals are indicated by arrows.
c: T-lymphocyte nucleus; BAC signals marking integration sites in yellow (2q14.2) and green (2q11.2); red – HIV signal; Co-localizing BAC and HIV signals are indicated by arrows; the T-lymphocyte nucleus (c) harbors multiple HIV integration sites (four HIV signals are visible); scale bars 5 µm.
For each nucleus the co-localization of HIV with one of the BAC signals is indicated by arrow in the overlay images; to benefit HIV FISH signal quality and to reduce background signal, the brightness/contrast tool in ImageJ was applied for the HIV channel.

The overlay image in Figure 21 clearly shows the co-localizing HIV and BAC signals (indicated by arrow). This observation was finally confirmed by a line scan analysis of the fluorescence intensities of both signals (Figure 22).

Interestingly, in some T-lymphocyte nuclei up to four HIV FISH signals were visible, indicating multiple integration sites (Figure 23c gives an example for this observation). The two integration sites 2q11.2 and 2q14.2 in the T-lymphocytes were located both on same chromosome 2. This was confirmed by an additional FISH experiment with probes for HIV, both BACs marking the integration sites and a paint probe for chromosome 2 (data not shown). On which chromosome the additional integration sites were located remained undetermined.

Cell populations of astrocytes, HeLa cells and T-lymphocytes were investigated with regard to the number of HIV FISH signals per nucleus. For this purpose only nuclei came into consideration, which were, based on the DAPI staining and nuclear morphology, convenient to be potentially evaluated for 3D measurements. This means the selected nuclei had to show the average morphology of the respective cell population. Of each population between 70 and 80 nuclei were investigated. The BACs for the evaluable integration sites were co-hybridized. In case of the HeLa cells additionally BAC 11q22.3 was co-hybridized. Although this integration site could not be evaluated, due to chromosomal aberrations (4.1.4), it should account for further characterization of the HeLa cells (see next chapter). Some HIV FISH signals did not co-localize with one of the BAC signals; however, they were taken as distinct signals if shape and fluorescence intensity was comparable to those that did co-localize.

Astrocytes never harbored more than one HIV FISH signal. This always co-localized with the BAC signal marking integration site region 18q22.1 (Figure 24a). In the majority of HeLa cell nuclei one HIV signal was found (74%, Figure 24b), very rarely two (3%). In the latter case none HIV signal co-localized with one of the three BAC signals (18q22.3, Xq22.1 and 11q22.3). In T-lymphocytes up to four HIV signals

per nucleus were found, in most cases one or two (33% and 34%, respectively, Figure 24c). This observation is in accordance with results from qPCR experiments mentioned at the beginning of this chapter, whereupon the T-lymphocytes harbor up to six HIV provirus copies and are therefore multiple infected. If two or more HIV signals were visible, they mostly (69%) co-localized with both integration site marking BACs 2q11.2 and 2q14.2.

The qPCR result, which determined one HIV provirus copy per astrocyte nucleus, was also confirmed by the count of HIV FISH signals (Figure 24a). Since the qPCR experiment determined 0.6 provirus copies per nucleus in the HeLa cells, they were characterized in details (see following chapter 4.2.4).

In a certain amount of nuclei (between 5% and 23%, Figure 24) of all three cell populations no HIV signal was found. It was not verifiable whether this was due to nuclei without integrated HIV provirus or due to an insufficient hybridization because of the critical size of the target sequence (<10 kb).
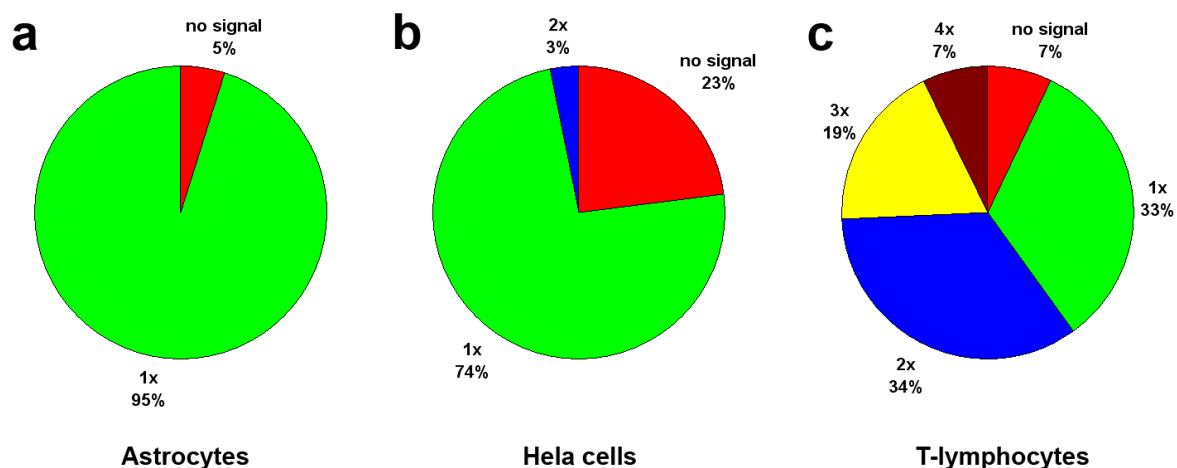
Figure 24: Count of HIV FISH signals per nucleus
Overview of number of HIV signals per nucleus among the cell population of astrocytes (a), HeLa cells (b) and T-lymphocytes (c); between 70 and 80 cells for each population were investigated; nuclei were selected according to their nuclear morphology, which should represent the average of the respective cell population.

4.2.4 The HeLa cell line LC5-HIV is a pool of clones

The cell line LC5-HIV was established by HIV-1 infection of a HeLa-derived cell line followed by sub-cloning via limiting dilution (Mellert et al., 1990). qPCR results revealed 0.6 provirus copies per nucleus (Manja Meggendorfer, Helmholtz Zentrum München, personal communication). Six integration sites were mapped via LM-PCR

(see 4.1.1). Among the HeLa cell population most nuclei harbored one HIV signal (74%, Figure 24b). In nearly one fourth (23%, Figure 24b) no HIV signal was visible. To characterize the HeLa cell population further, in nuclei with one HIV signal, the frequency was determined how often the HIV signal co-localizes with one of the BACs marking the integration sites 18q22.3, Xq22.1 and 11q22.3 (see Figure 25). For this purpose the 74% of the cell population with one HIV signal per nucleus indicated in Figure 24b was evaluated.
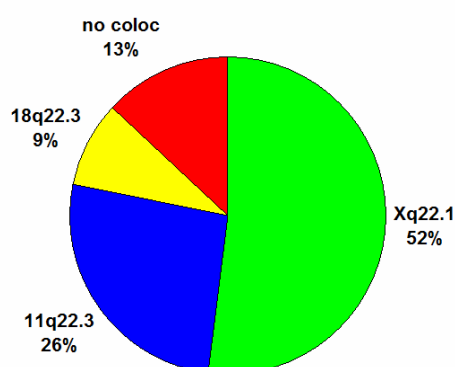
Figure 25: HIV integration site distribution among the HeLa cell population in nuclei with one HIV signal
Evaluation of the 74% of the cell population indicated in Figure 24b with one HIV signal per nucleus. Percentage indicates how often the HIV signal did not (no coloc) or co-localized with one of the BAC signals marking the integration sites Xq22.1, 11q22.3 and 18q22.3; n=80.

Interestingly, in the majority of evaluated nuclei, the HIV signal co-localized with the BAC signal marking integration site region Xq22.1 (52%, Figure 25). Only in 9% the HIV signal co-localized with the BAC 18q22.3, in 26% with the BAC 11q22.3 (Figure 25). In 13% of nuclei the HIV signal did not co-localize with one of the co-hybridized BACs Figure 25. This is explainable with overall six mapped integration sites in the HeLa cells.

Apparently, the LC5-HIV HeLa cells are a pool of different clones, although the cells were supposed to be sub-cloned (Mellert et al., 1990). There is a domination of the integration locus Xq22.1 harboring clone of 40% among the population (52% of cells with one HIV signal per nucleus in Figure 25 equates to 40% of the whole population). Multiple infections can be excluded at least for integration site regions 18q22.3, 11q22.3 and Xq22.1. In 2D FISH experiments as described and shown in 4.1.7 with probes for these regions never more than one HIV signal per metaphase spread was found. Furthermore, qPCR experiments revealed only 0.6 HIV provirus copies per nucleus, although six integration sites were mapped in the HeLa cells

(4.1.1). In 23% of the HeLa cells no HIV signals was found, which suggests that a certain amount of cells are not infected. As mentioned before, this could also be due to the critical target size of the HIV provirus. This seems to be unlikely in case of HeLa cells, since the percentage of HeLa nuclei without HIV signal exceeds the respective percentages determined in astrocytes and T-lymphocytes by far (23% compared to 5%, respectively 7%, Figure 24).

4.2.5 HIV integration sites within a nucleus differed in their transcriptional activities

To characterize further the transcriptional activity of the integration sites, RNA FISH experiments were applied for T-lymphocytes and HeLa cells (method see 3.10.6). Briefly, without denaturing nuclear DNA, FISH with a probe for HIV was performed. Generally, in both cell types the HIV RNA signals were surprisingly strong and distinct indicating high HIV transcription (Figure 28 and Figure 27). This was already determined before by comparing HIV and β-actin RNA signals (see Figure 20). HeLa cells and T-lymphocytes varied in the HIV signal count per nucleus (Figure 26), whereas the determination of the exact number of HIV RNA signals per nucleus was sometimes difficult (exemplified in Figure 28f). A BAC signal, which could validate the HIV signal by co-localization as in DNA FISH experiments, would not have been visible due to the procedure of RNA FISH. Hence, BAC probes were not used here.

Among HeLa cells, one HIV RNA signal was visible in 67% of nuclei, two signals in 4% of nuclei (Figure 26a). In nearly one third (29%) no HIV RNA signal could be detected, which supports the assumption in the previous, that not all HeLa cells are HIV infected. T-lymphocyte nuclei habrored mostly four or more signals (48%Figure 26b), whereas, as mentioned before, in nuclei with more than four signals the exact count is hardly identifiable without a co-localizing BAC signal. In Figure 28f for example case five RNA signals were counted (indicated by arrowheads).

The HIV RNA signals in HeLa cells generally seem to be larger, even though slightly, than in T-lymphocytes, suggesting stronger HIV transcription in HeLa cells. Noteworthy, both in HeLa cells and T-lymphocytes HIV RNA signals within the same nucleus frequently differed in their sizes and intensities (compare HIV signals marked by arrows and asterisks in Figure 27 and Figure 28). This indicates different transcriptional activities of HIV proviruses within the same nucleus. The possibility of several proviruses being integrated consecutively cannot be fully excluded since final

74

control experiments are missing. This should therefore be taken into account for higher local transcriptional potential.
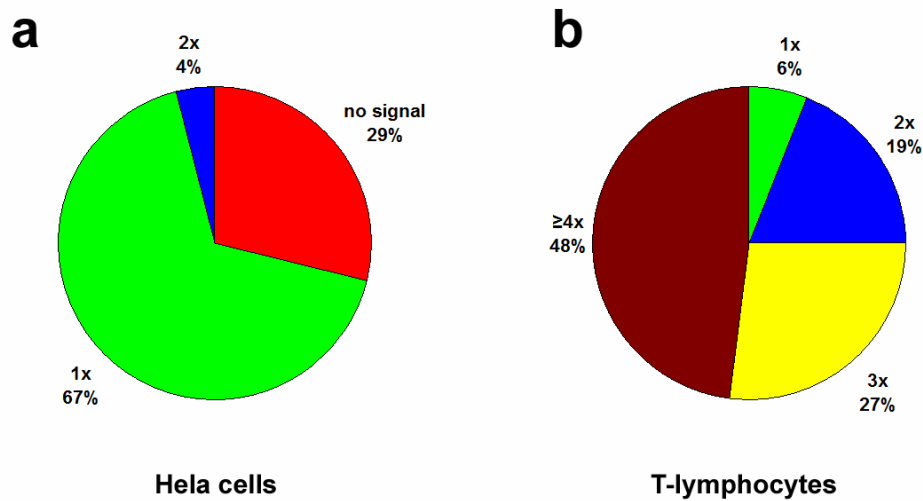
Figure 26: Percentage of different counts of HIV signals per nucleus in HeLa cells (a) and T-lymphocytes (b) after RNA FISH
Of each cell type 100 nuclei were evaluated
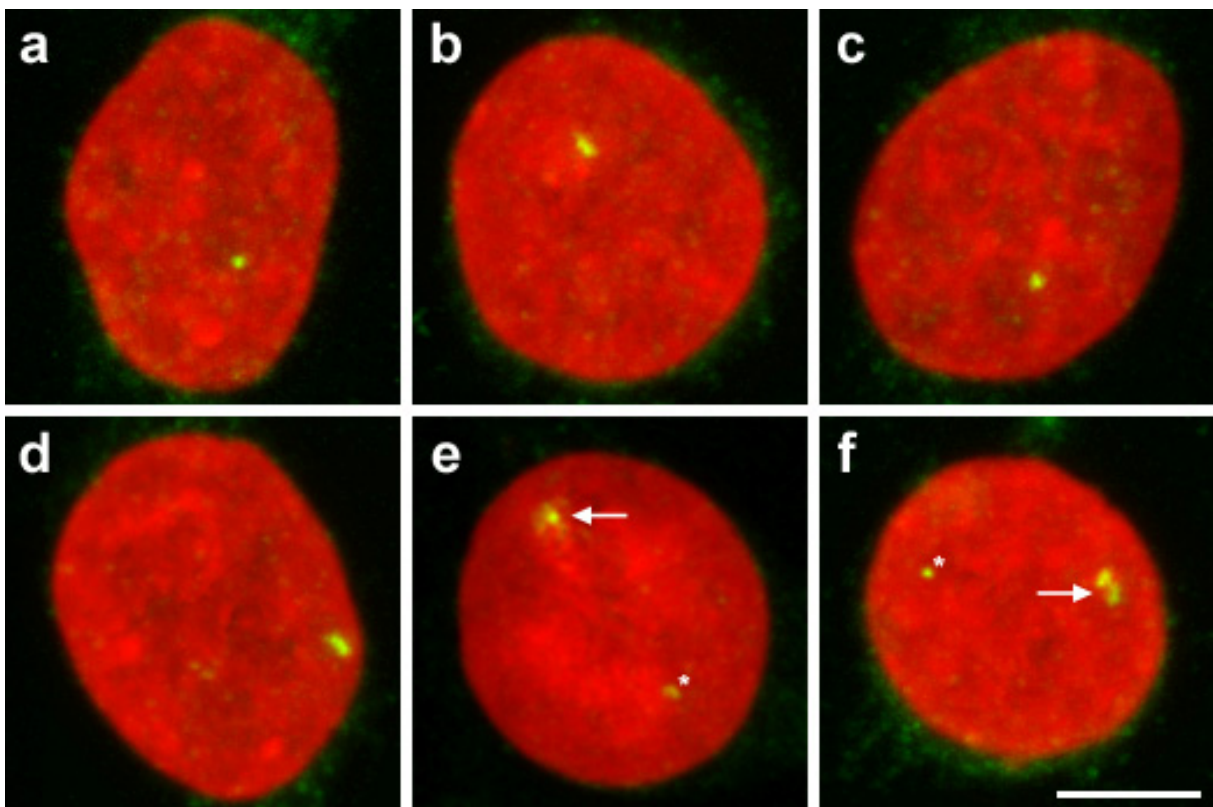
Figure 27: HIV RNA FISH signals in HeLa cell nuclei
Z-projection of confocal image stacks; RNA FISH with HIV probe; red – DAPI, green – HIV RNA; e,f: arrows mark stronger HIV RNA signal, asterisks weaker signals; scale bar 5 µm.

75
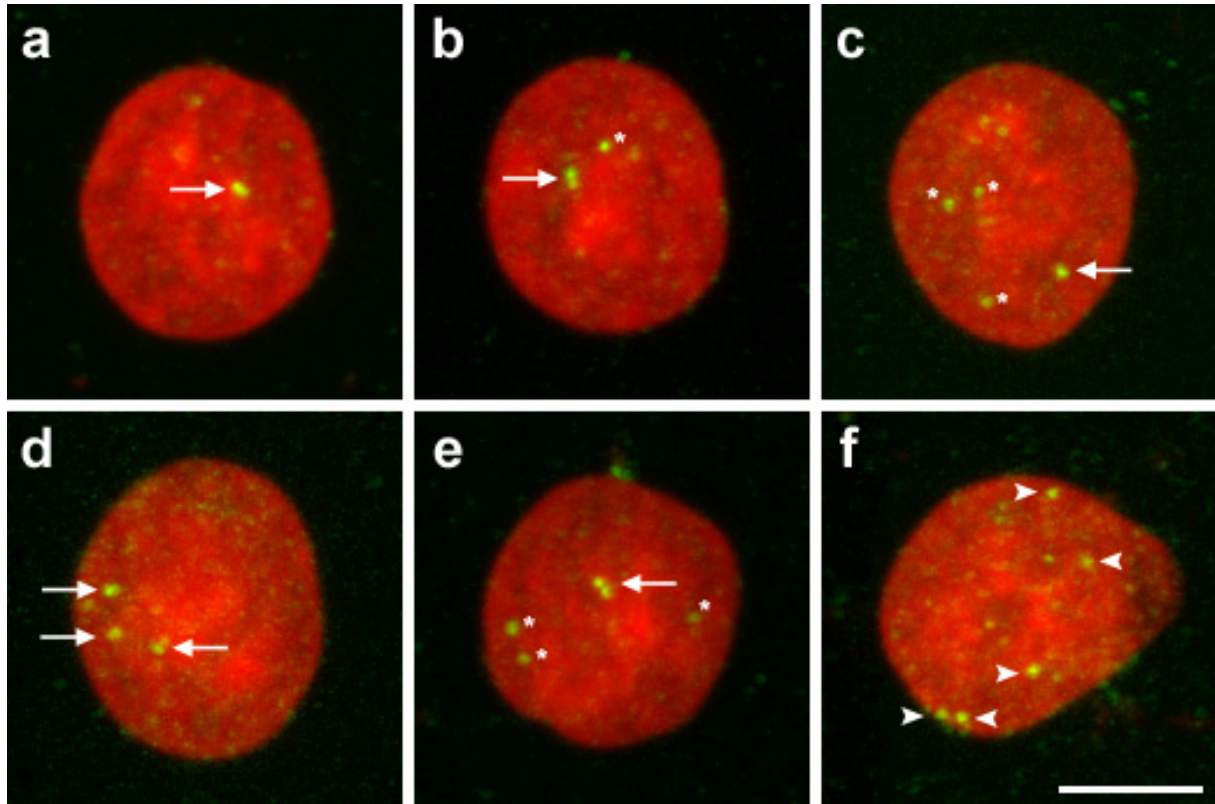
Figure 28: HIV RNA FISH signals in T-lymphocyte nuclei
Z-projections of confocal image stacks; RNA FISH with HIV probe; red – DAPI, green – HIV RNA; a-e: arrows mark strong HIV RNA signals, asterisks mark weaker signals; f: putative HIV RNA signals are indicated by arrowheads; scale bar 5 µm.

4.3 Radial nuclear positioning of HIV integration sites

To determine the 3D nuclear distribution of HIV integration sites, FISH with probes for HIV and BAC probes marking the respective integration site regions was applied. Nuclei were selected according to their morphology and co-localization of HIV and BAC signal. Only nuclei that showed the average morphology of each cell type population were selected. Images were processed as described (see 3.13). Briefly, DAPI and BAC images were normalized, blur was carefully reduced and a threshold was defined. Signals of integration sites and homologous control sites were separated and analyzed independently. For both signals the shortest distance in 3D to the nuclear surface was determined as described (see 3.13). Possible repositioning of the HIV integration locus was investigated by comparing the nuclear distribution of the BAC signal co-localizing with the HIV signal and the respective BAC signal on the homologous chromosome.

In HeLa cells both integration loci (18q22.3 and Xq22.1, green curves; Figure 29) were highly significant (p<0.001 and p=0.002, respectively) located more internal

than their respective regions on the homologous chromosomes. The integration loci in T-lymphocytes (TLy 2q22.2 and TLy 2q14.2) and astrocytes (Astro 18q22.1) did not show significant positional changes (p>0.05). Comparing the mean values of their medians, differences between the distribution curves in Figure 29 become even more obvious (Figure 30). In HeLa cells, the mean value of the medians of the integration locus 18q22.3 distribution is 0.24 µm more internal compared to the control region, in case of integration locus Xq22.1 0.12 µm. The differences for all other integration loci distributions and respective control regions are marginal (between 0 and 0.04 µm, Figure 30).
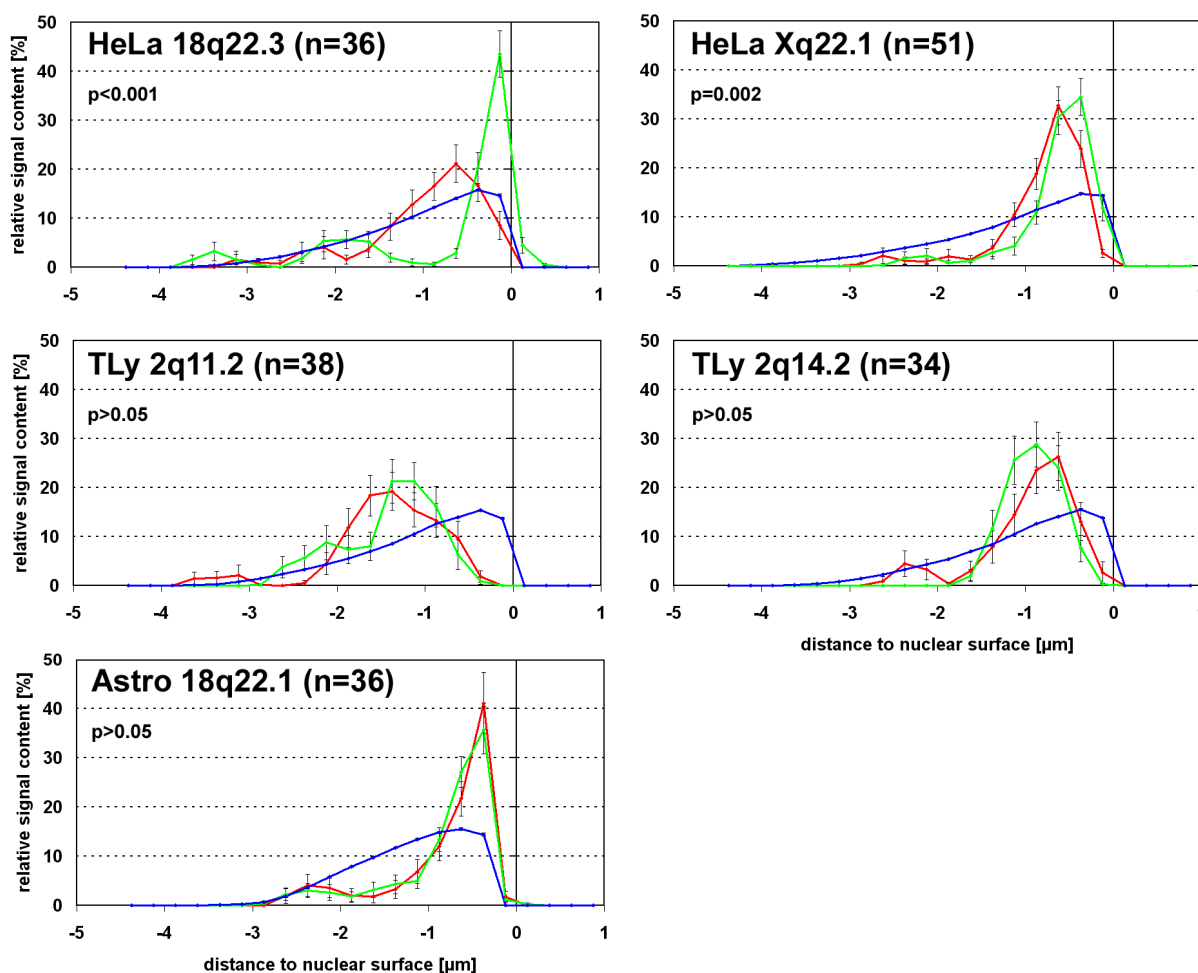
Figure 29: nuclear distribution of HIV integration loci
distance to nuclear surface: negative values – inside nucleus, positive values – outside nucleus; red curves – nuclear distribution of BAC signal co-localizing with HIV signal; green curves – nuclear distribution of BAC signal on homologous chromosome; blue – DNA distribution; HeLa – HeLa cells; TLy – T-lymphocytes; Astro – astrocytes; statistics was performed with medians of green and red curves; n – number of evaluated nuclei.
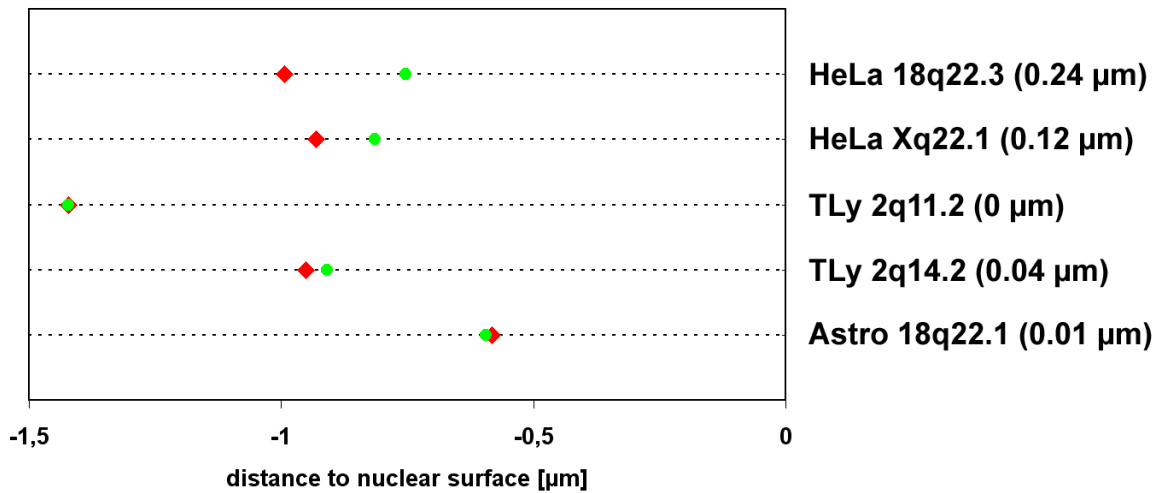
Figure 30: mean values of medians, same dataset as in Figure 29
Mean values of medians of distribution curves of BAC signals co-localizing with HIV (red squares) and respective BAC signals on homologous chromosomes (green circles); differences between values for each integration locus in brackets.

To characterize both repositioned integration loci in HeLa cells (18q22.3 and Xq22.1) further, their distribution within their respective chromosome territories 18 and X was determined. Both integration loci show the same distribution within their harboring chromosome territories as their respective regions on the homologous chromosomes (p>0.05; comparing the medians of red and green curves in Figure 31a and b). A significant difference in the nuclear distribution of the integration site harboring chromosome territories and the homologous ones was not determined (p>0.05; comparing the medians of red and green curves in Figure 31c and d).

Taken together, repositioning of integration loci 18q22.3 and Xq22.1 in HeLa cells was only with regard to the nuclear surface. The distance to the surface of the respective chromosome territory did not change. Neither chanced the integration site harboring chromosome territories 18 and X their nuclear positions compared to their homologues.
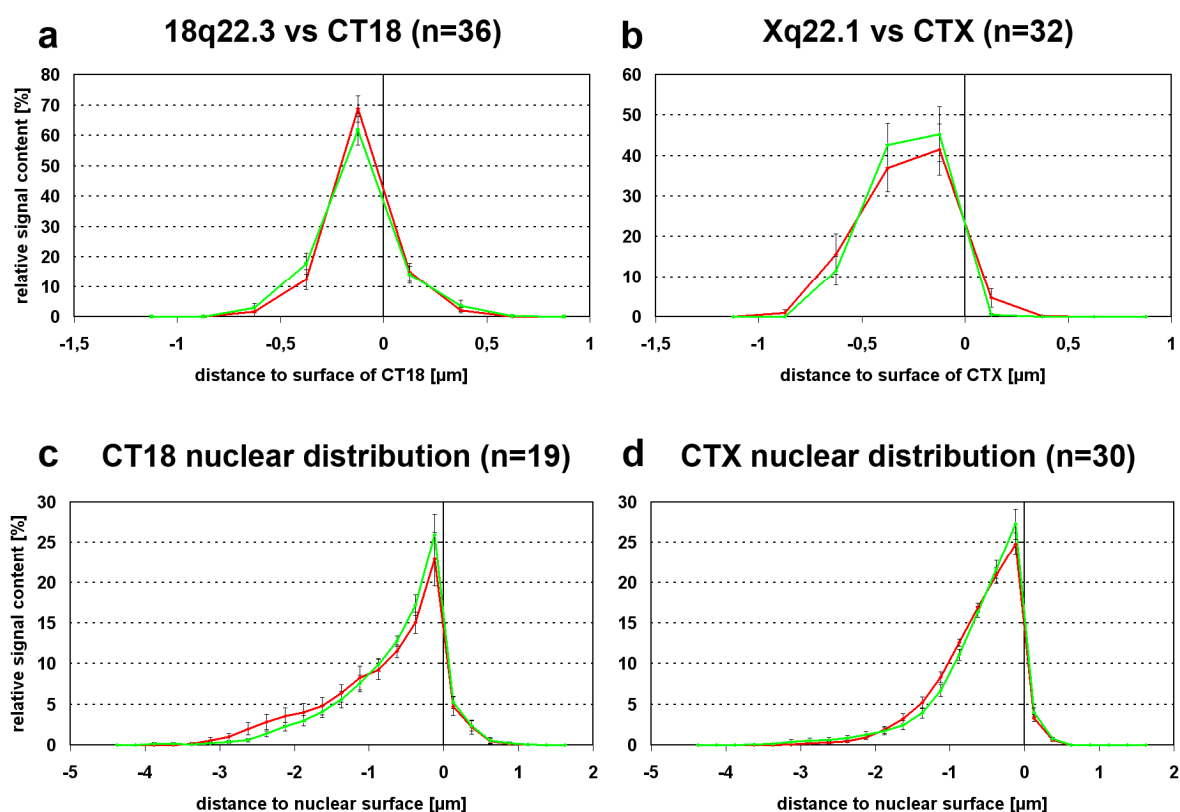
Figure 31: BAC distributions within their respective chromosome territories (a,b) and nuclear distribution of chromosome territories 18 and X (c,d)

Distances of BACs to the chromosome territory surface and distanes of the chromosome territories to the nuclear surface was determined as described in 3.13.

a,b: distribution of BAC signals with regard to the surface of their harboring chromosome territory (CT18 and CTX, respectively); red curves – BAC signals (18q22.3 and Xq22.1) co-localizing with HIV signals; green curves – respective BAC signals on homologous chromosomes;

c,d: nuclear distribution of CT18 and CTX discriminating between integration site harboring (red curves) and homologous chromosome territory (green curves).

x-axis: negative values – inside chromosome territory (a,b) or nucleus (c,d), positive values – outside chromosome territory (a,b) or nucleus (c,d).

4.4 Repositioning was accompanied by chromatin decondensation

Upon visual inspection of BAC signals, in some cases signals of integration sites appeared more decondensed than the BAC signals on the homologous chromosomes. In 2D z-projections of three different HeLa nuclei in Figure 32, the BAC signals co-localizing with the HIV signals (indicated by arrows in the overlay, Figure 32 lower panel) occupy larger areas than their homologues.

This impression was confirmed by 3D surface measurements of the BAC signals (Figure 33) as described in 3.13.6. Briefly, using the 3D object counter plugin in ImageJ the surface area of the BAC signals was determined. To avoid personal bias a fixed threshold was used. In HeLa cells, the BAC signal covering integration site region 18q22.3 showed on average a 1.49 times larger surface (p<0.001, Figure

33) than the respective BAC on the homologous chromosome. For the BAC marking integration site region Xq22.1, the surface difference was smaller (1.14x) but still highly significant (p<0.001, Figure 33). In T-lymphocytes and astrocytes no significant difference in the surface areas could be determined. Interestingly, these observations correspond to the extent of repositioning (Figure 29 and Figure 30). The strongest difference in the surface area was determined for the most repositioned integration site region 18q22.3 in HeLa cells. Less difference in the surface area was found for integration site region Xq22.1 in HeLa cells corresponding to a less pronounced repositioning.
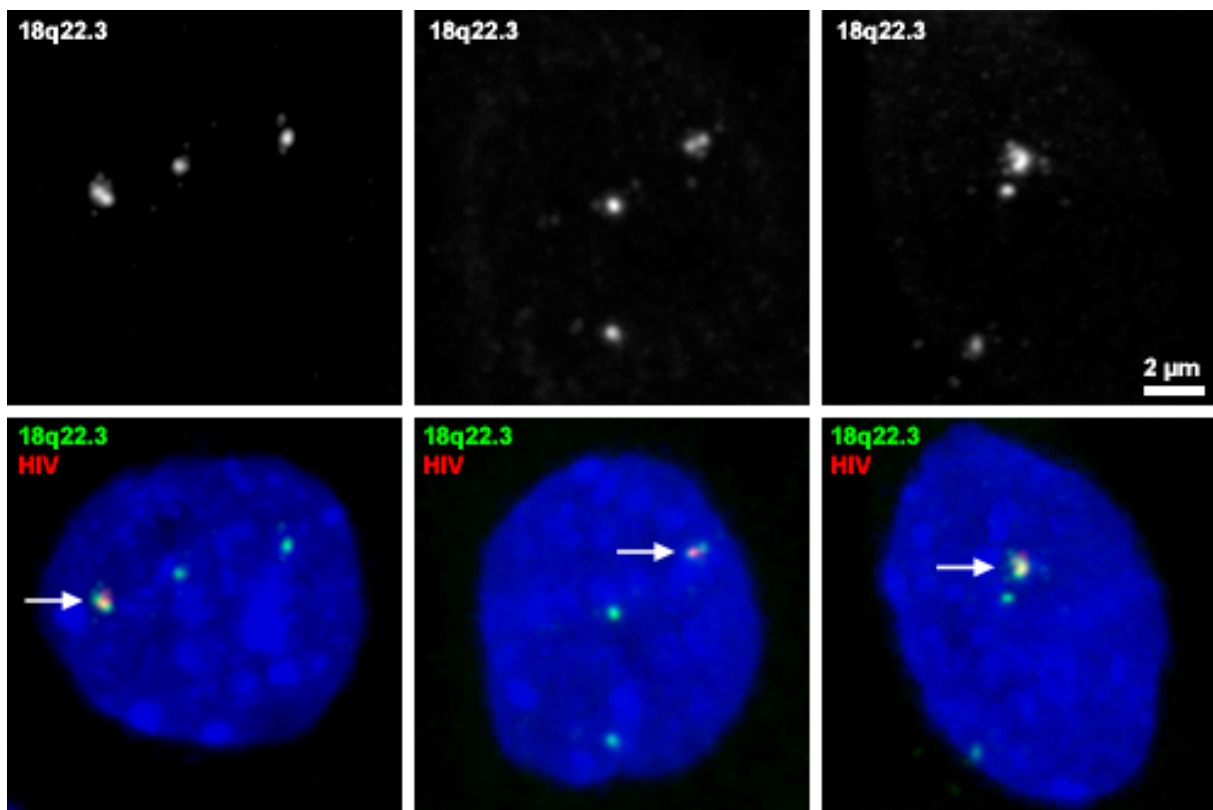
Figure 32: BAC signals in HeLa cell nuclei demonstrating Chromatin decondensation
Z-projections of confocal image stacks; 3D FISH on three different HeLa nuclei with HIV probe and BAC marking integration site region 18q22.3;
upper panel: BAC (18q22.3) signals; lower panel: respective overlay images of BAC (18q22.1, green) and HIV signals (red); counterstain with DAPI (blue); arrows indicate BAC signals co-localizing with HIV signals.
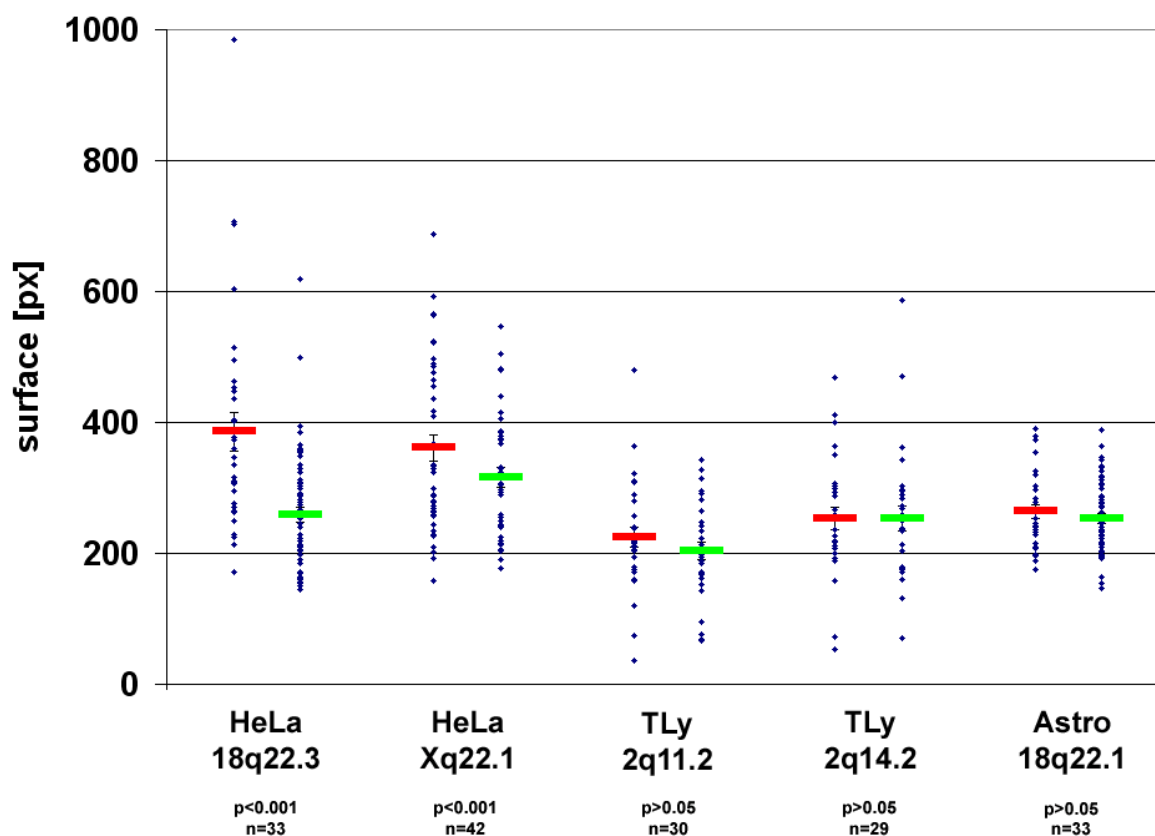
Figure 33: Surface area of BAC signals
Dots – surface area in pixels (px) of each BAC signal; bars – mean values of respective BAC surface pixels co-localizing (red) or not co-localizing (green) with HIV signal, respectively.

4.5 Treatment with sodium butyrate increased HIV transcriptional activity, but did not significantly affect the proviral nuclear position

Sodium butyrate (NaBut) acting as a deacetylase inhibitor has been shown to markedly increase HIV-1 transcription via the transcription factor NF-kB (Quivy et al., 2002). Repositioning towards the nuclear interior upon transcriptional activation was described for transgenes (Tumbar and Belmont, 2001; Dietzel et al., 2004) and several genes or chromosomal subregions (Zink et al., 2004; Chuang et al., 2006; Williams et al., 2006). Whether transcriptional upregulation of the HIV provirus results in an analog repositioning, was tested by treating the cells with NaBut as described in 3.2.3. For HeLa cells, T-lymphocytes and astrocytes both HIV transcription and virus production was measured, in each case with and without NaBut treatment. HIV transcriptional activity was determined via qPCR by comparing to RNA polymerase II RNA levels (described in 3.2.3). Virus production was measured by the amount of

p24 (described in 3.2.3), which is a HIV core protein and therefore correlates with the virus particle count.

In all three cell types HIV transcription was induced upon NaBut treatment (Figure 34a). The highest HIV transcription rate compared to RNA polymerase II RNA levels, with and without NaBut treatment, was determined in T-lymphocytes, whereas in HeLa cells and astrocytes lower amounts of HIV transcripts were found. Except for untreated HeLa cells, the amounts of HIV transcripts exceeded the amounts of RNA polymerase II transcripts, which can be read off the ratio values in Figure 34a.

Among the three investigated cell types, the highest virus production, respectively the virus particle concentration in the cell supernatant was measured in HeLa cells (Figure 34b). In T-lymphocytes and astrocytes virus production was comparatively low. As HIV transcription levels, in HeLa cells, T-lymphocytes and astrocytes virus production was increased upon NaBut treatment. In summary, NaBut treatment increased both HIV transcription and virus production (Figure 34c). HeLa cells showed the highest increase in HIV transcription (~10-fold) compared to T-lymphocytes (~2-fold) and astrocytes (~3-fold). HIV production in astrocytes was increased nearly 7-fold, in HeLa cells and T-lymphocytes virus production was increased only marginally (1.3-fold and 1.1-fold, respectively).

Possible changes in the positioning of the integration loci upon NaBut treatment were determined as described for untreated cells (see 4.3). While transcription of the HIV provirus was increased, the NaBut evaluation revealed very similar distributions for integration loci as well as for control loci (same locus in treated versus untreated cells: $p > 0.05$ for all integration loci). The only noticeable difference is that astrocyte nuclei have become smaller which can be read off the DAPI curve for Astro 18q22.1 (in Figure 29 the curve reaches the x-axes at -3 µm, in Figure 35 at -2 µm indicating a nuclear radius of 3 µm or 2 µm, respectively).

a

HIV transcripts / RNA PolII RNA

b

p24 concentration

c

HIV induction by NaBut

83
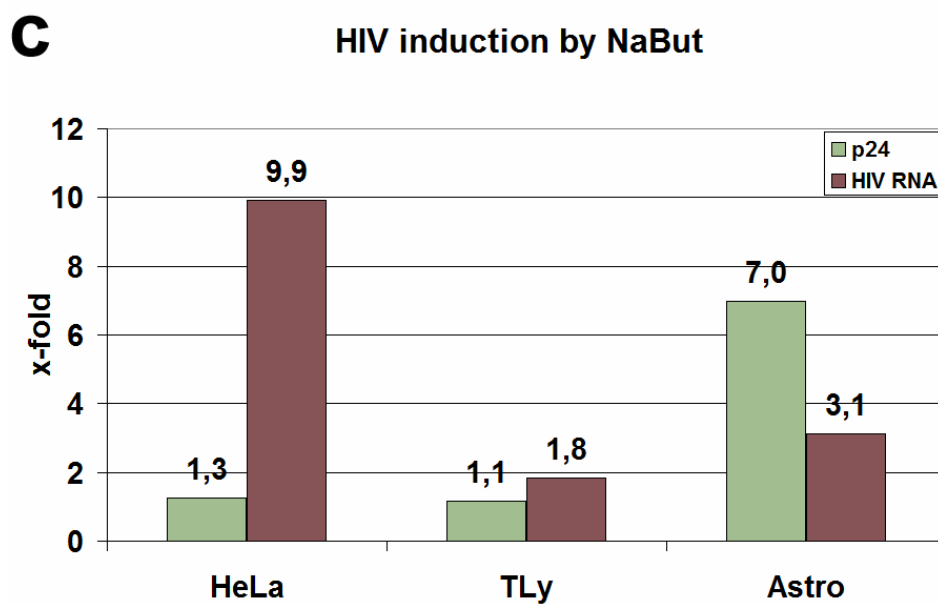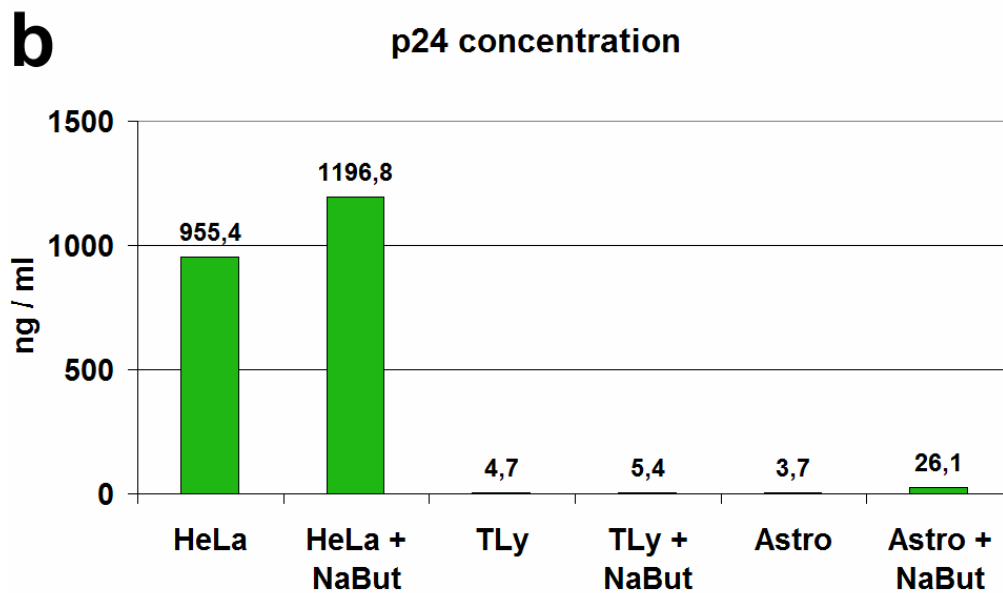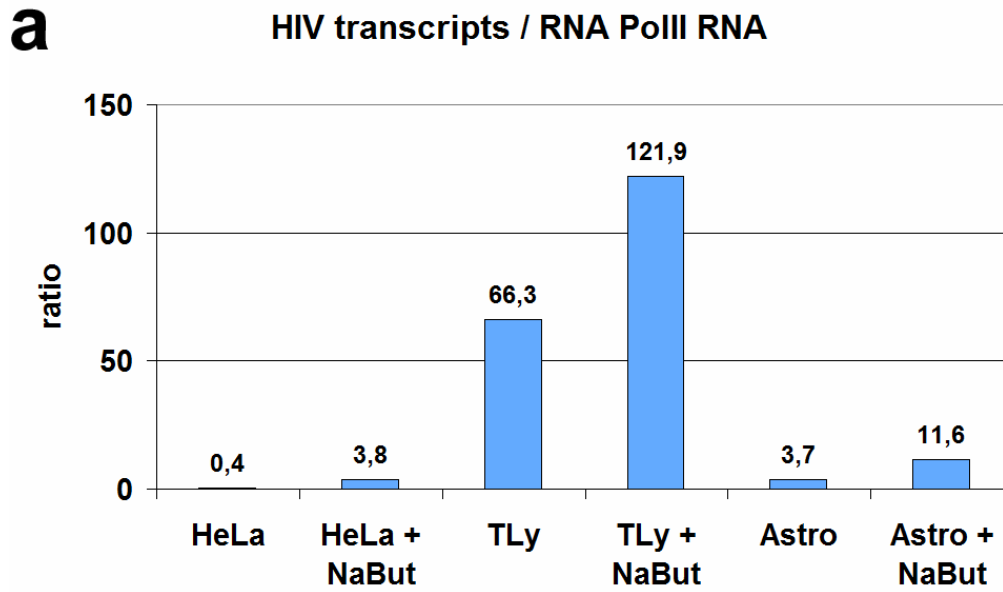
Figure 34 (previous page): Effect of sodium butyrate (NaBut) treatment
A: Ratio of HIV transcripts to RNA polymerase II transcripts in HeLa cells (HeLa), T-lymphocytes (TLy) and astrocytes (Astro) compared to NaBut treated cells (+NaBut); numbers above columns display measured values (data adapted from measurements performed by Manja Meggendorfer, Helmholtz Zentrum München).
B: p24 concentration in supernatant in HeLa cells (HeLa), T-lymphocytes (TLy) and astrocytes (Astro) compared to NaBut treated cells (+NaBut); numbers above columns display measured values (data adapted from measurements performed by Manja Meggendorfer, Helmholtz Zentrum München).
C: x-fold increase of p24 levels an HIV RNA by NaBut treatment in HeLa cells (HeLa), T-lymphocytes (TLy) and astrocytes (Astro) compared to untreated cells.
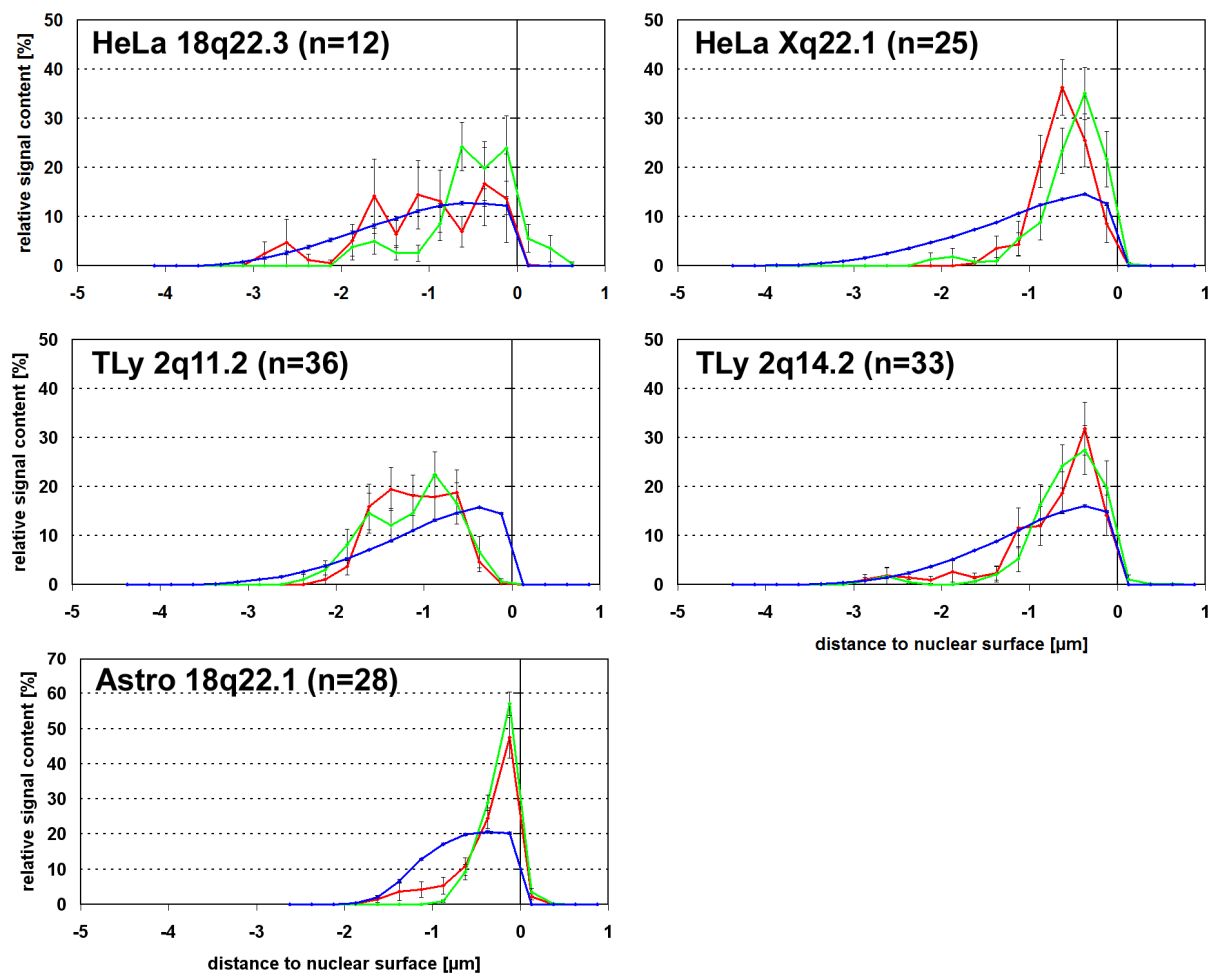
Figure 35: nuclear distribution of HIV integration loci in NaBut treated cells
distance to nuclear surface: negative values – inside nucleus, positive values – outside nucleus; red curves – nuclear distribution of BAC signal co-localizing with HIV signal; green curves – nuclear distribution of BAC signal on homologous chromosome; blue – DNA distribution; HeLa – HeLa cells; TLy – T-lymphocytes; Astro – astrocytes; n – number of evaluated nuclei.

4.6 Integration of HIV had no influence on the positioning of the integration locus relative to SC35 splicing speckles

Since HIV does not have own splicing factors the virus is dependent on the host splicing machinery. To investigate the spatial relation of HIV integration sites towards splicing speckles, the distance of the BAC signals co-localizing with the HIV signal and the respective BAC signals on the homologous chromosomes to SC35 splicing speckles was determined as described in 3.13.5. Therefore 3D FISH in combination with immunocytochemistry was performed (see 3.11). In advance to threshold defining for measurement procedure, the microscopic images were deconvolved as described in 3.13.2. Escpecially for threshold defining, blur hinderes the specification of FISH signals emanating from fine structures such as splicing speckles. The mathematical process of deconvolution reduces blur caused by out of focus light. In (Albiez, 2007) the gain in signal quality and enhancement in contrast were proven by comparing deconvolved confocal images with electron microscope images, the latter providing much higher resolution.

As illustrated in Figure 36 the raw image without deconvolution is much more blurry and noisy compared to the deconvolved image. Hence, deconvolution facilitates proper threshold defining by improved delineating of SC35 signal.

Neither BAC signals co-localizing with HIV signals nor BAC signals of the homologous regions contacted SC35 signals, with the exception of HeLa 18q22.3 where 15 -20% of BAC signals co-localized with or were found adjacent to SC35 Figure 38). Significant differences between the distribution of integration loci and homologous control regions relative to SC35 splicing speckles were not determined (p>0.05; Figure 38). The distribution of the HIV FISH signal towards splicing speckles of the same experiment was determined separately, but revealed similar results as the co-localizing BAC signal (red curves in Figure 38).

Since splicing of the HIV transcripts is mandatory, the question arises whether HIV RNA signals contact SC35 speckles. Therefore, RNA FISH in combination with immunocytochemistry was performed (described in 3.11). In HeLa cells and T-lymphocytes, ~30% of the HIV RNA signals co-localized with or contacted splicing speckles (Figure 39b). BAC probes were not used since RNA signals were not expected, which was proven in preliminary RNA FISH test experiments. Hence, it could not be identified from which integration site the RNA signals derived.
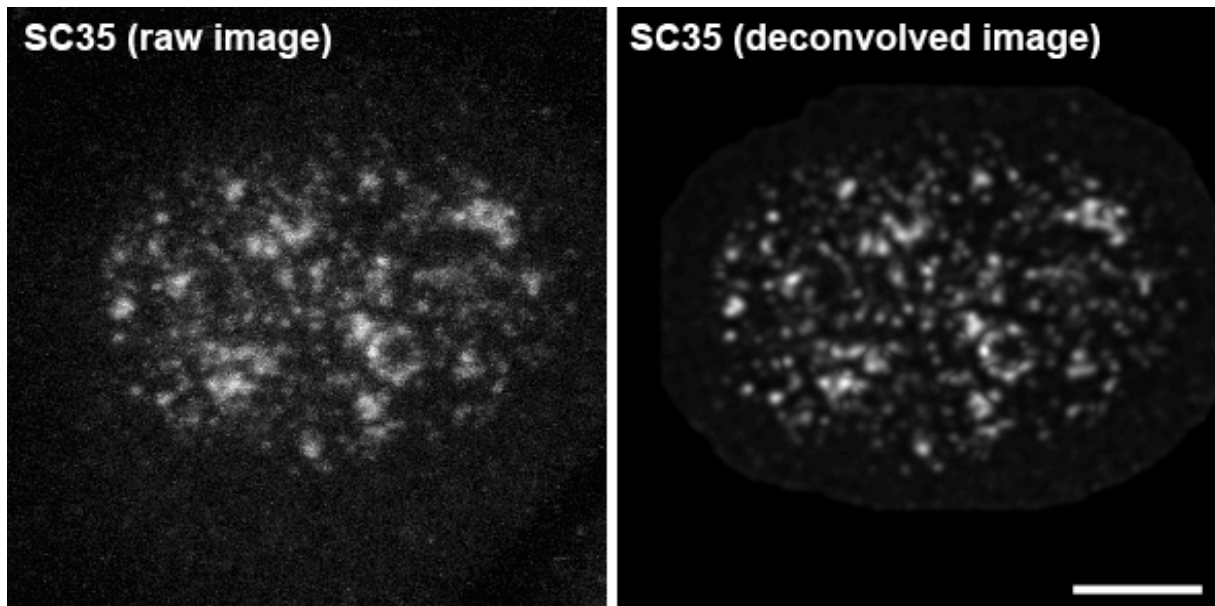
Figure 36: Effect of deconvolution on SC35 signal in an astrocyte nucleus
Z-projection of three mid confocal sections of the same SC35 splicing speckles image; before (raw image) and after deconvolution (deconvolved image); scale bar 5 µm.
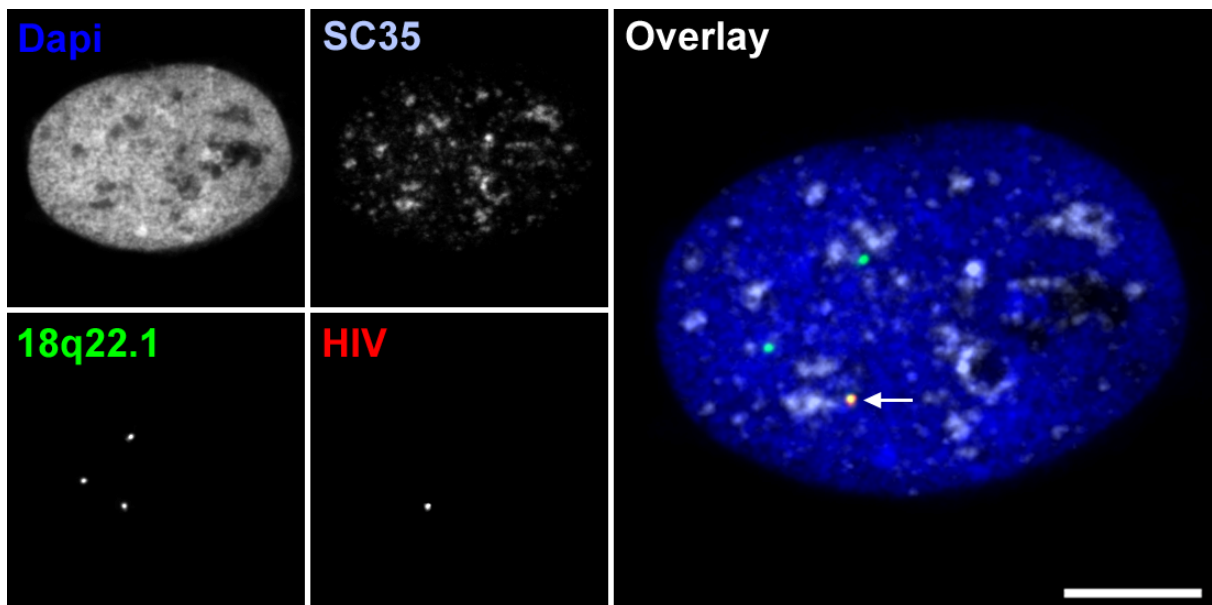
Figure 37: SC35 antibody staining in combination with 3D FISH
Z-projection of three mid confocal sections of the respective astrocyte nucleus from Figure 36; FISH probes for BAC covering integratiuon site region 18q22.1 (green) and HIV (red); counterstain with DAPI (blue); antibody against SC35 splicing speckles (light blue); Images are deconvolved; arrow in overlay indicates BAC signal co-localizing with HIV; to benefit HIV FISH signal quality and to reduce background signal, the brightness/contrast tool in ImageJ was applied for the HIV channel; scale bar 5 µm.
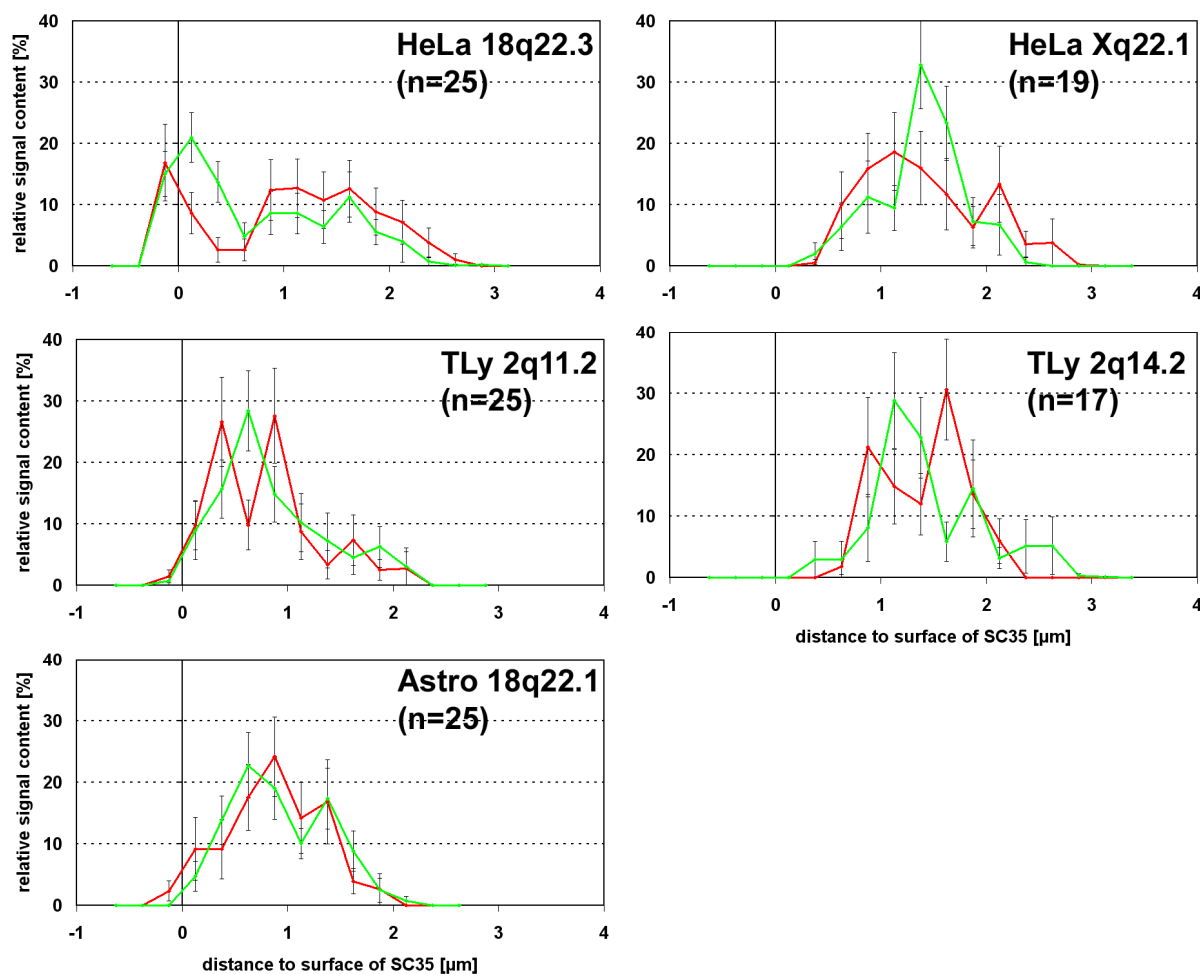
Figure 38: Distribution of BAC signals with regard to SC35 splicing speckles
Distribution of BAC signal co-localizing with HIV signal (red) and on homologous chromosome (green) towards surface of SC35 splicing speckles; x-axis: negative values – inside SC35 signals, positive values – outside SC35 signals; n – number of evaluated nuclei.
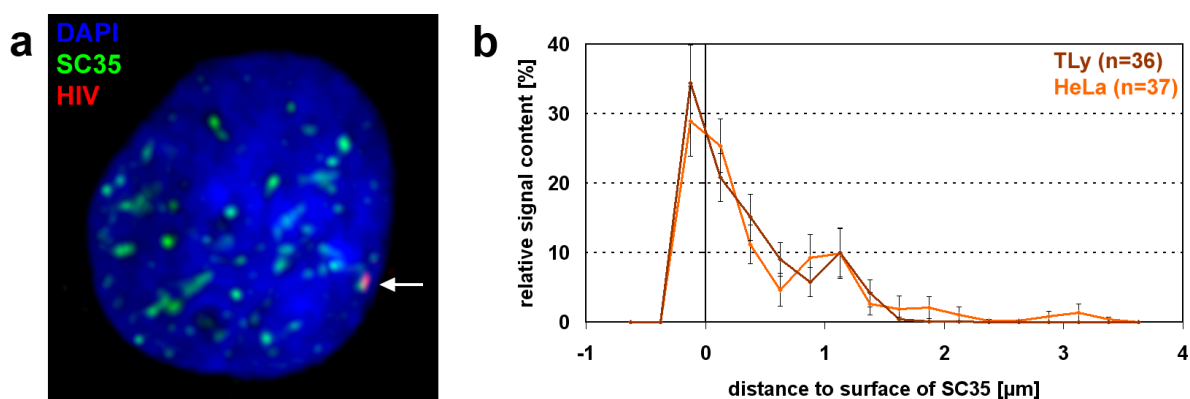
Figure 39: RNA FISH in combination with SC35 antibody staining and according eADS evaluation
a: z-projection of mid confocal sections, HeLa cell nucleus; green - SC35 splicing speckles, red – HIV RNA (indicated by arrow), blue - counterstain with DAPI; image is deconvolved; scale bar 5 µm.
b: distribution of HIV RNA with regard to the surface of SC35 splicing speckles in T-lymphocytes (TLy, (brown) and HeLa cells (HeLa, orange); negative values on x-axis – inside SC35 signal; positive values – outside SC35 signal; n – number of evaluated nuclei.

87

4.7 Karyotype analysis and genomic properties of the integration sites in the mouse cell line cloneB

The mouse hematopoietic precursor cell line cloneB was transduced with the LTR-driven gamma-retroviral gene transfer vector pSF91-GFP (Modlich et al., 2006). In this publication, five integration sites are described. Four additional integration sites were identified in the lab of Manuel Grez (Georg-Speyer-Haus, Frankfurt), who kindly provided the cell line (for overview of identified integration sites see Table 4).

FISH on metaphases and karyotype determination of the cell line, test for correct chromosomal position of the BAC clones and gene density, GC content and chromatin condensation analysis of the integration loci were part of this thesis, while 3D FISH experiments and data evaluation (see 4.8.2) were part of the supervised diploma thesis of Birgit Groß (Groß, 2009).

Integration site	MMU 3A3[+]	MMU 5G1[+]	MMU 9A3	MMU 11B5	MMU 11E1[+]	MMU 12B3[+]	MMU 12D2	MMU 16B2	MMU XA1.1[+]
Nearest gene	Evi1	Mad1like1	Dnm2	Lig3	Gm885	Akap6	Batf	AI480653	Wdr45
Distance to nearest gene	121 kb	0	0	0.5 kb	0	0.4 kb	13.5 kb	0	0.4 kb
Gene ID	14013	17120	13430	16882	380732	238161	53314	268880	54636

Table 4: gene transfer vector integration sites in cloneB
Overview of gene vector integration sites in mouse hematopoietic cell line cloneB. Integration site description defines the chromosomal region; Integration site mapping was performed in the lab of Manuel Grez (Georg-Speyer-Haus, Frankfurt); integration sites marked by + have been described in (Modlich et al., 2006); gene ID from NCBI mouse genome database.

4.7.1 The cloneB cell line has a normal karyotype

To verify the chromosomal position of BAC clones, each clone was co-hybridzed with the integration site harboring chromosome paint probe. Karyotype analysis of cloneB revealed a normal female mouse karyotype (40,XX). As exemplified for integration sites MMU11B5, MMU11E1, MMU3A3 and MMU5G2 in Figure 40a,c, no chromosomal aberrations of the integration site harboring chromosomes were determined. Hence, all nine integration loci could be evaluated in terms of their 3D nuclear position. The used BAC clones are summarized in Table 5.
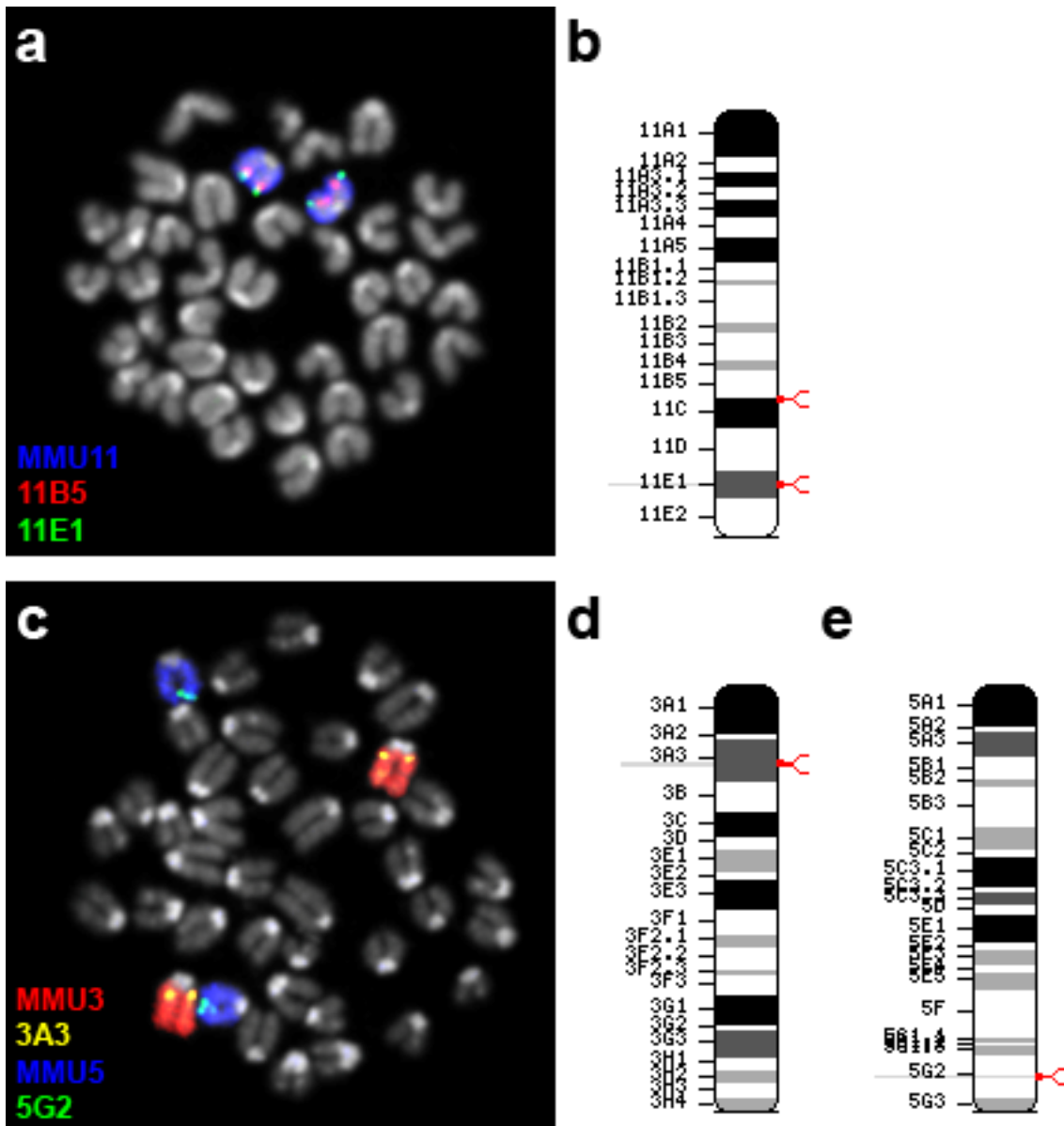
Figure 40: BAC test and karyotype anaysis of cloneB
a, c: FISH on cloneB metaphase spreads; paint probes for chromosome 11 (a, MMU11, blue), chromosome 3 (c, MMU3, red) and chromosome 5 (c, MMU5, blue); BAC probes covering integration site regions 11B5 (a, red), 11E1 (a, green), 3A3 (c, yellow) and 5G2 (c, green); counterstain with DAPI (grey).
b, d, e: ideograms for respective chromosomes stained in metaphase spreads on the left; red marks indicate respective integration loci.

Integration site	MMU 3A3	MMU 5G1	MMU 9A3	MMU 11B5	MMU 11E1	MMU 12B3	MMU 12D2	MMU 16B2	MMU XA1.1
BAC name	RP23-439N22	RP23-217N11	RP23-317E10	RP23-316C11	RP23-247J12	RP24-267H3	RP23-369N11	RP24-206H1	RP23-54C14

Table 5: BAC clones for integration sites in cloneB

4.7.2 Gene density and GC content of integration sites

Gene density
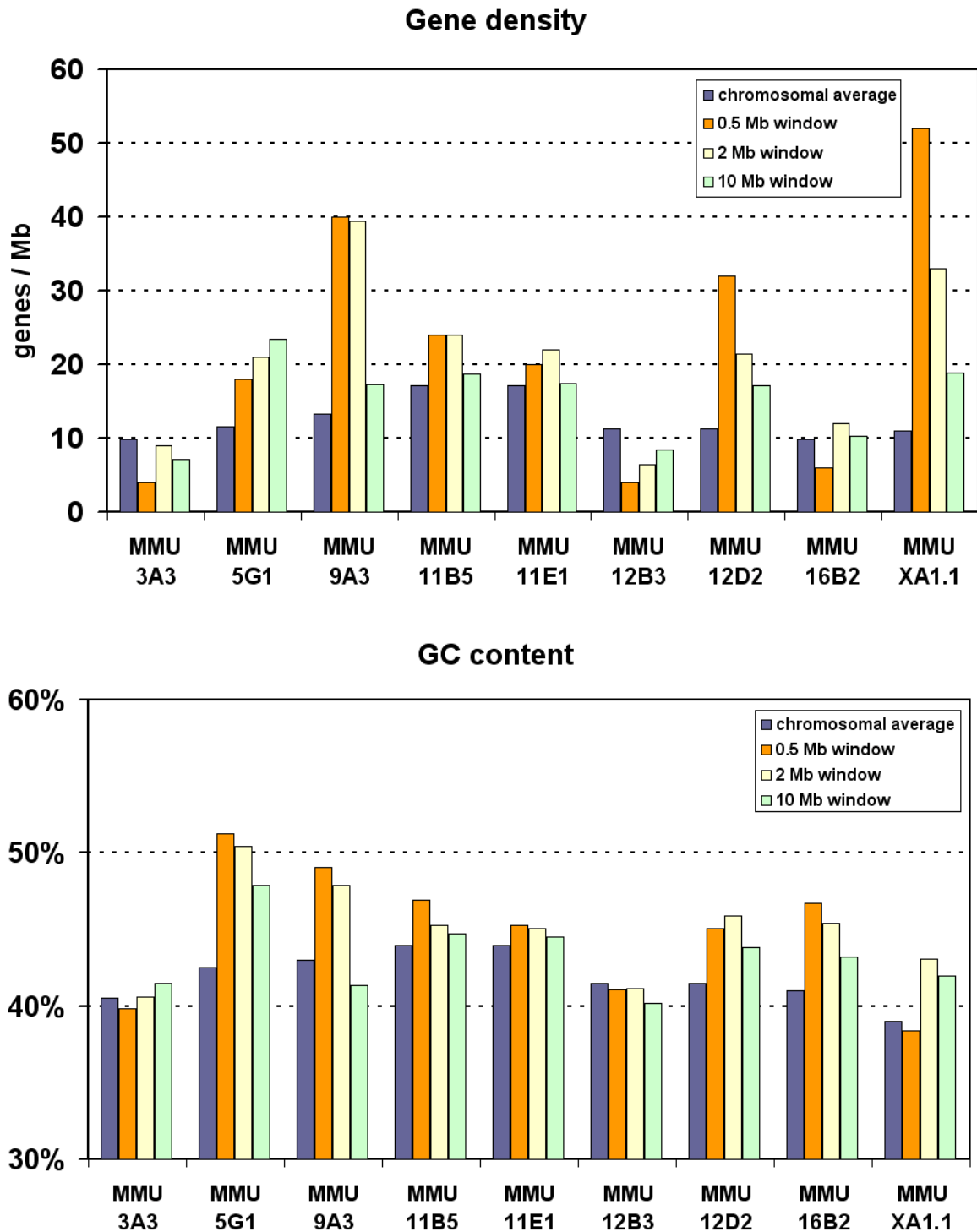
GC content

Figure 41: CloneB gene density and GC content of gene vector integration site regions
Gene density was read off the NCBI homepage; the sequences were submitted to the Repeatmasker homepage to obtain the GC content. Chromosomal averages were taken from NCBI homepage (gene density) and from (Waterston et al., 2002) (GC content).

90

The chromosomal neighborhoods in 0.5 Mb, 2 Mb and 10 Mb windows around the integration sites vary both in gene density levels and GC content. MMU3A3, MMU12B3 and MMU16B2 for example locate in gene poor, whereas MMU9A3, MMU12D2 and MMUXA1.1 are found in rather gene dense regions (Figure 41). Concerning the GC content similar observations can be made, whereby again no preferences are obvious.

4.8 FISH on intact mouse hematopoietic precursor nuclei and data evaluation

All nine integration sites were investigated in terms of their radial nuclear distribution. To investigate possible repositioning upon gene transfer vector integration, the position in 3D of BACs covering the integration sites were compared to the position of the homologous BACs. This chapter summarizes data from the diploma thesis of Birgit Groß (Groß, 2009). However, these data are essential to understand the following chapter and parts of discussion and were therefore included here.

4.8.1 No more than four integration sites were visualized in one nucleus

In 3D FISH experiments with the gene transfer vector probe and different BAC probes marking the integration loci (Figure 42), a varying number of vector signals per nucleus were found. The maximum count was never more than four. The target sequence of the vector probe is 3.2 kb and therefore even smaller than the HIV provirus. It is most probably that, just like in case of HIV, RNA contributed to the enlargement of the FISH signal and thus to its visualization. This suggestion was supported by 2D FISH experiments on metaphases, where only DNA for hybridization is available. Here, it was not possible to visualize the gene transfer vector signal.

In summary, a maximum of four integrations per nucleus were transcriptionally active simultanously or at least active at detectable levels. Combinations of certain actively transcribed integration sites were not found, but all integration sites were transcriptionally active at any time.
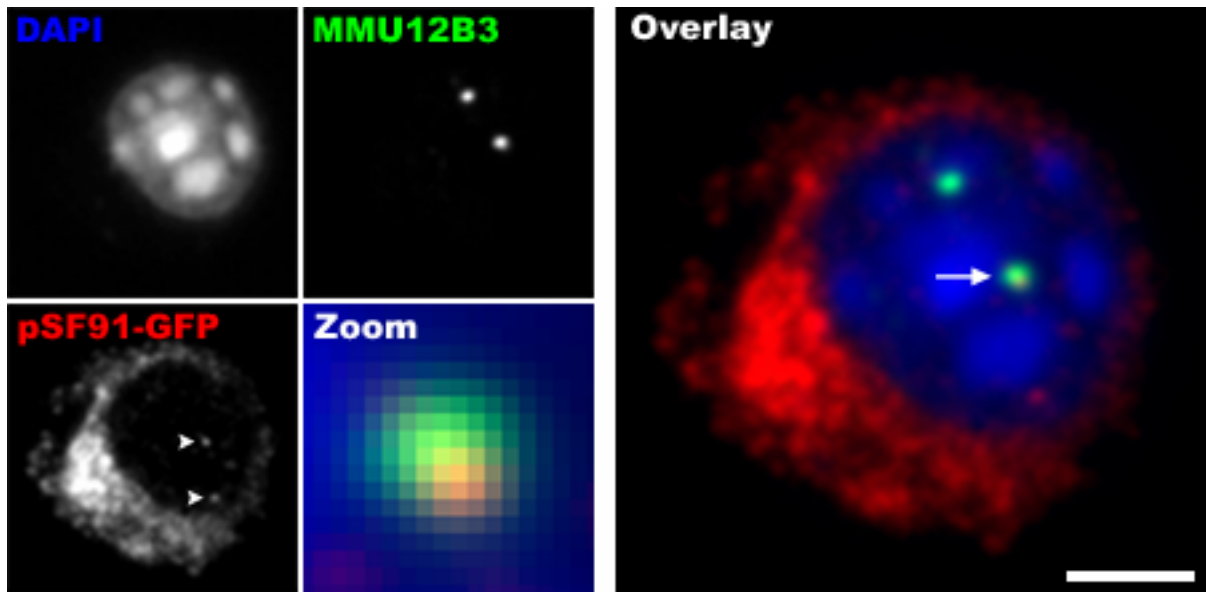
Figure 42: 3D FISH on mouse hematopoietic precursor cell line cloneB nucleus
Z-projections of confocal image stacks; Overlay image: green – BAC signal marking integration site region (MMU12B3), red – gene transfer vector signal (pSF91-GFP), blue – DAPI counterstain, co-localization of gene transfer vector and BAC signal indicated by arrow, scale bar 3 µm; Zoom image: magnification of co-loacalization of BAC and gene transfer vector signal from overlay image; arrowheads in pSF91-GFP image indicate gene transfer vector signals; image adapted and modified from (Groß, 2009).

4.8.2 All nine integration loci in the mouse cell line cloneB tend to locate more internal in the nucleus than their homologous regions

Integration site region MMU3A3 showed a highly significant shift towards the nuclear interior when compared to its corresponding site on the homologous chromosome (p<0.001, Figure 43a,j). Notably, MMU3A3 is in the immediate vicinity of the protooncogene Evi1 (ecotropic virus integration site 1), a region with frequently observed viral and gene transfer vector integrations (Modlich et al., 2006; Ott et al., 2006; Wieser, 2007; Metais and Dunbar, 2008). MMU5G1 showed a less pronounced but still significant difference (p=0.033, Figure 43b).

Although no significant difference in the radial nuclear distribution in case of the other seven integration sites was determined (p>0.05, Figure 43c-i), interestingly, they were all located more internal than their homologous regions (Figure 43j). However, for a random radial distribution of nine integration site regions, the probability that all nine are localized more interior than their homologues is only $(0.5)^9$, corresponding to p<0.002. For all nine sites combined, the deviation from a random distribution towards more internal positions of integration sites is thus highly significant. Even if the integration site regions with the smallest differences,

92

MMU11E2 and MMU12D2, would be excluded from the calculation, the p value still would be highly significant with $(0.5)^7 = 0.008$.

93

4.8.3 Chromatin condensation at integration sites was altered

Since repositioning in HIV infected cells was accompanied by chromatin decondensation (Figure 33), the gene transfer vector integration sites were also investigated in this regard (described in 3.13.6). A significant difference in chromatin condensation between integration loci and homologous regions was determined for MMU11B5 and the repositioned integration site region MMU3A3. In both cases chromatin of the homologous regions was less condensed, in principle the opposite effect as observed in HIV infected cells. For all other integration sites no significant difference was detectable ($p > 0.05$).
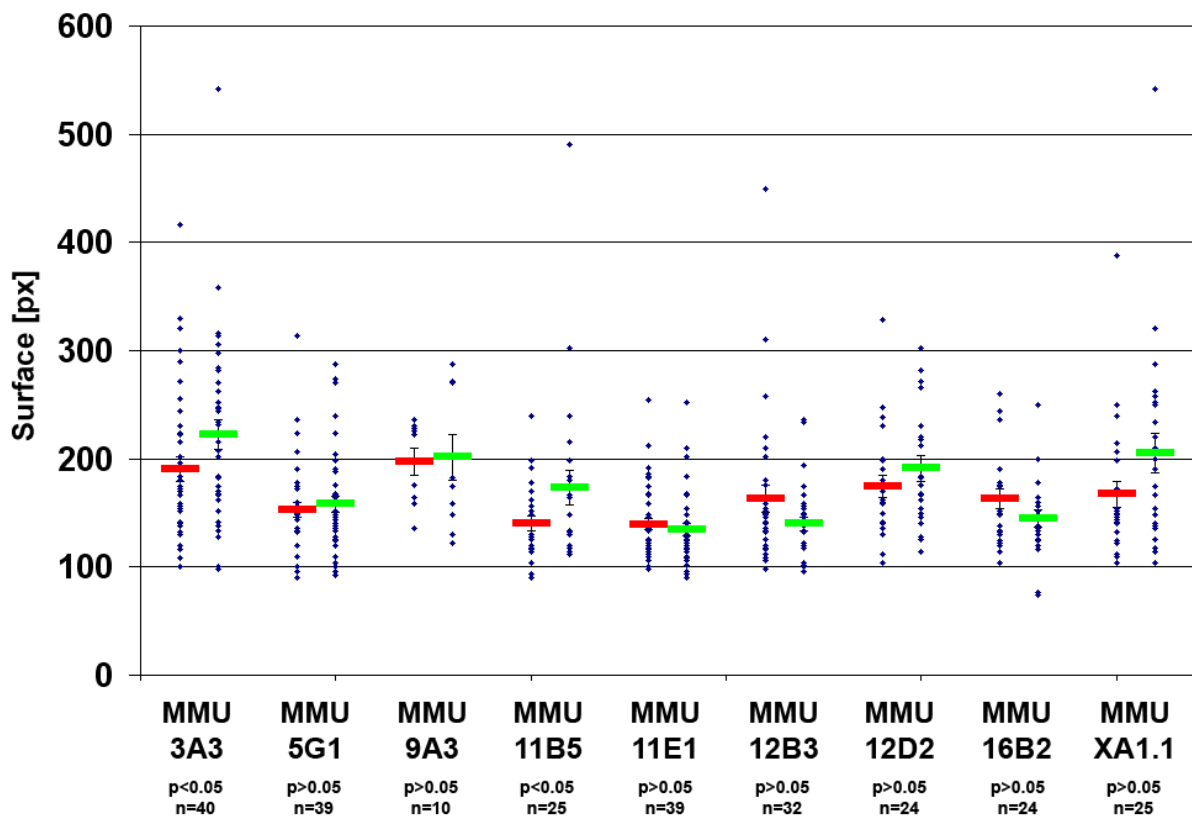
Figure 44: Surface area of BAC signals in the cloneB cell line
Blue dots indicate the surface area of each single BAC signal; red (BAC signal co-localizing with gene vector signal) and green (respective BAC signals on homologous chromosomes) bars indicate the mean values in the respective columns; for method of determination of BAC surface area see 3.13.6.

5. DISCUSSION

5.1 Development of a protocol to visualize small FISH targets

The first experimental challenge in this thesis was the visualization of a fluorescence signal for a <10 kb target sequence with FISH methods. The probe is conventionally labeled with fluorochromes or haptens and in the latter case subsequently detected by a fluorochrome labeled antibody. Hybridization ought to be carried out on structurally preserved nuclei. The HIV provirus, i.e. the integrated HIV genome within the host genome, is about 9.7 kb in size (Ratner et al., 1985). The integrated gene transfer vector in the mouse hematopoietic precursor cell line cloneB with a size of 3.2 kb was even three times smaller.

Alternative FISH techniques for critical target signal sizes have also been described. A dramatic increase of sensitivity can be achieved by enzymatic-mediated signal amplification such as catalyzed reporter deposition (CARD)-FISH (Schönhuber et al., 1997; Zwirglmaier, 2005). Following this method, either the probe nucleotides or the antibody detecting the probe are conjugated to horseradish peroxidase (HRP). HRP activates its substrate tyramide by oxidization which than covalently binds to tyrosine residues in close vicinity to HRP. The accumulation of tyramides conjugated to fluorochromes results in signal intensities up to 12 times stronger than with conventional labeled FISH probes (Schönhuber et al., 1997; Zwirglmaier, 2005). This techinique has already been used to visualize the integrated HIV provirus (Dieudonne et al., 2009). The drawback of this technique is the HRP molecule size of 44 KDa. To assure cell access lysozyme or proteinase treatment is required, which concomitantly affects cell morphology. Hence, this might be a useful tool for studies on metaphases, which was already shown (Deichmann et al., 1997), but possibly limited for 3D analysis on structurally preserved interphase nuclei. Additionally, the precise position of the target sequence gets delocated by the covalently binding of tyramides in the vicinity.

Integrated HIV DNA was already visualized with FISH techniques in (Bell et al., 2001). There, cells were denatured together with the probe by heating at 94°C for 4 min. This robust denaturing process normally should severely irritate if not destroy the 3D nuclear structure. The difference in (Bell et al., 2001) was that the cells were dehydrated in an ethanol series and air dried in advance to hybridization. This process flattens the nucleus and impedes reliable 3D analysis.

So, for the purposes in this thesis, FISH technique for structurally preserved nuclei as described in (Hepperger et al., 2007; Müller et al., 2007; Cremer et al., 2008) had to be used, whereas standard protocols had to be costumized. As FISH probe template for HIV, the whole pNL4-3 plasmid was used, which contains the HIV proviral genome framed by the LTR sequences (Adachi et al., 1986). For the gene transfer vector in the mouse cell line, the whole pSF91-eGFP plasmid was used.

The first question was how to label the FISH probe. The variety of hapten and fluorophore labeled nucleotides offers plenty possibilities. Using hapten labeled probes depends on antibody detection, which may account for worsening background signal due to unspecific antibody binding. Therefore, directly labeled probes were tested first. During recording nuclei at the confocal microscope faster bleaching compared to hapten labeled probes was observed. The best results were achieved by digoxigenin and directly Texas Red labeled probes in repeated experiments. Digoxigenin was detected with the antibody combination mouse α dig-Cy5 and goat α mouse-Cy5 leading to less background signal and marginal bleaching. Directly Texas red labeled HIV probes resulted in less background signal, but bleaching was worse, even though acceptable, compared to hapten labeled probes. Taken together, the digoxigenin labeled HIV probe detected with mouse α dig-Cy5 and goat α mouse-Cy5 was the first choice in this thesis and is most advisable. The gene transfer vector probe

The next modification was made with the composition of the hybridization mixture. For non-repetitive probes it is recommended to use 20-100 ng DNA/µl hybridization mixture and additionally unlabeled Cot-1 competitor DNA, representing 10-50-fold the concentration of probe DNA (Müller et al., 2007). Although the HIV genome does not harbor larger repetitive sequences, adding the 50-fold concentration of Cot-1 DNA increased markedly the FISH signal quality, so did high amounts of probe (100-150 ng DNA/µl hybridization mixture). FISH control experiments revealed that the HIV probe detected HIV DNA and RNA which extended the fluorescence signal (see Figure 18 and Figure 19). Since RNA is prone to degradation it was important not to store the fixed cells too long until use. Finally, denaturation time and temperature were reduced. The cells were placed for 100 sec. at 72°C instead of the recommended 3 min. at 76°C on the hotblock.

In summary, higher amounts of probe DNA in the hybridization mixture, freshly fixed cells and gentle denaturation were pivotal criteria for visualizing a FISH signal derived from a <10 kb target sequence.

5.2 Gene transfer vector integrations show preferences for gene dense and GC rich regions

The mouse hematopoietic precursor cell line cloneB was transduced with the gamma-retroviral gene transfer vector pSF91-eGFP (Schambach et al., 2000; Modlich et al., 2006). Four of the nine identified integration sites were located within genes (see Table 7 page 114). Since in gene therapy insertional mutagenesis caused by gene vector integration within host genes is a major concern, integration between genes may be an advantage. Depending on the chromosomal location of the integration site gene function might be irritated by activation or inactivation potentially provoking leukemia (Bokhoven et al., 2009; Kustikova et al., 2009). The frequency of severe phenotypic alterations evoked from gene vector integration, however, is extremely low. This was shown both in animal models and clinical studies (Baum et al., 2003; Kohn et al., 2003). This may be underlined by the fact that the vast majority of the human genome is occupied by non-coding sequences although these DNA areas provide regulatory elements with essential functions in regulating gene expression. Hence, variations of these sequences by inserting genetic material may also have serious consequences.

In the mouse cell line cloneB gene dense areas were slightly favored for integration site selection compared to gene poor areas. Six (MMU 5G1, MMU 9A3, MMU 11B5, MMU 11E1, MMU 12D2, MMU XA1.1) of the nine integration sites were in more gene dense regions compared to the respective chromosomal average, three of them (MMU 9A3, MMU 12D2, MMU XA1.1) in very gene dense regions (>30 genes/Mb in a 0.5 Mb window) (see Figure 41 and Table 9). The highest gene density with more than 50 genes/Mb in a 0.5 Mb window was found for integration site on the X chromosome. A more obvious preference seems to exist for the GC content. Five of the nine integration sites are in regions with GC content higher than 45%. The lowest GC content in a 0.5 Mb window was determined for integration sites MMU 3A3, MMU 12B3 and MMU XA1.1, but in these cases it was still in the range of

the respective chromosomal average. The highest percentage was found for the integration on chromosome five with more than 50% GC content.

In contrast to HIV integrations, a preference for gene transfer vector integrations within genes was not found here, as only five of nine were within genes (Table 4 page 88). On the other hand, a tendency of gene transfer vector integrations towards rather gene dense and GC rich regions was determined. These results show that the risk of insertional mutagenesis is low, although cannot be excluded, since the probability to hit a gene for integration in gene dense regions is given.

5.3 All gene transfer vector integration loci tend to locate more internal of the nucleus than their homologous regions

In the mouse hematopoietic precursor cell line cloneB transduced with a gamma-retroviral gene transfer vector all nine mapped and investigated integration sites located more internal than their respective loci on the homologous chromosomes (see Figure 43). Notwithstanding only two integration sites were repositioned significantly (MMU3A3 and MMU5G1, whereas MMU3A3 revealed higher significance; Figure 43), taking all nine integration sites together, the result is highly significant. For a random radial nuclear distribution of nine integration sites the probability that all nine are localized more interior than their homologues is only $(0.5)^9$, corresponding to p<0.002. Even if integration site regions MMU11E1 and MMU12D2, which showed the least change, are excluded from the calculation, the p value still would be highly significant with $(0.5)^7$=0.008.

In (Williams et al., 2006) a scenario is described that the upregulation of the Mash1 gene was accompanied by repositioning of the respective gene locus towards the center. Several neighboring genes were activated upon this process. If oncogenes would be affected by this activation, repositioning of gene transfer vector integration sites may have severe consequences in the context of gene therapy.

Interestingly, the higher significantly repositioned integration locus MMU3A3 was in the immediate vicinity of Evi1 (ecotropic virus integration site 1), a region with frequently observed viral and gene vector integrations, whereas in most cases the integration leads to overexpression of Evi1 (Buonamici et al., 2004; Modlich et al., 2006; Ott et al., 2006; Wieser, 2007; Metais and Dunbar, 2008). In murine hematopoietic precursor (lineage negative) cells, Evi11 overexpression was shown to

enhance the self-renewal capacity and clonal dominance resulting from Evi1 upregulation was reported in several animal studies and human clinical trials (Li et al., 2002; Du et al., 2005; Kustikova et al., 2005; Ott et al., 2006; Laricchia-Robbio and Nucifora, 2008; Bosticardo et al., 2009; Komeno et al., 2009; Stein et al., 2010). It was shown that in the cell line cloneB the Evi1 gene was up-regulated upon the upstream gene transfer vector integration, which was essential for immortalization and cell survival (Modlich et al., 2006). Increased transcription of Evi 1 and potentially of the gene transfer vector may have taken influence on the nuclear position of the affected chromosomal region. Taken together, the nuclear repositioning of the MMU3A3 integration locus is linked to activation of Evi1 and clonal viability.

In 3D FISH experiments never more than four gene transfer vector signals per nucleus were found. The small size of 3.2 kbp and results from HIV experiments suggest, that the gene vector is only visualized when it is sufficiently transcriptionally active. On metaphase spreads, where only DNA is detected, a distinctive signal for the 3.2 kbp gene transfer vector was never visible. Apparently, there is a transcriptional limitation for four integration sites per nucleus. Whether the other integration sites are transcriptionally silent or transcribed at undetectable levels remains elusive. It seems to be random which integration sites are transcribed simultaneously, since no preferences for certain combinations were identified.

5.4 In the mouse cell line cloneB alterations in chromatin condensation upon gene vector integrations were determined

Of the nine investigated gene transfer vector integration sites in the mouse cell line cloneB two differed significantly (MMU3A3, MMU11B5; $p < 0.05$) in chromatin condensation compared to their respective regions on the homologous chromosomes. The next largest change was determined for integration locus MMUXA1.1 and reached a p-value of 0.07. Contrary to the observations made in HIV infected cells, in all three cases of the mouse cells the surface area of the integration site regions was less than on the homologous chromosomes, suggesting more condensed chromatin in the integration site region. This was also the case for the repositioned integration site (MMU3A3). That means if any changes in chromatin condensation occurred, gene vector integrations induced the opposite effect than observed for HIV integrations in this thesis. The gene transfer vector pSF91-eGFP

contains the spleen focus-forming virus (SFFV) LTR as promoter element (Modlich et al., 2006). Comparisons concerning the strength of the SFFV and HIV LTR promoter are not described in the literature. It is assumed here that transcription of the HIV provirus exceeded the transcription of the pSF91-eGFP vector since the HIV Tat protein boosts transcription of the integrated provirus (Easley et al., 2010).

In the previous chapter was suggested that the increased transcription of the Evi1 gene in integration site region MMU3A3 may have taken influence on the radial position of this locus. The surrounding chromatin, in contrast, was more condensed than at the respective region on the homologous chromosome, which was also the case for integration sites MMU11B5 and MMUXA1.1. Several studies have described transcription in condensed chromatin and offer possible explanations. In (Hu et al., 2009) transcription of a condensed transgene was linked to a local chromatin decondensation, which was not visible with the applied imaging approaches. Reorganization within the histone octamere resulting in a locally small open transcriptionally permissive configuration is also an option as described in (Becavin et al., 2010).

5.5 HIV integration sites were preferentially found within genes, but independent of splicing speckles location

Generally, HIV proviruses in the human genome are predominantly found within active transcription units (Wang et al., 2007; Brady et al., 2009). In this thesis seven of nine identified HIV integration sites were located within genes, in line with these reports (see Table 3 page 62 and Table 6 page 113). Due to their predicted functions it is most likely that they are transcriptionally active (see Table 6). In (Mack et al., 2003) genes adjacent to HIV integration sites frequently encode receptor-, signal transduction- and transcription-associated proteins. Such preferences were also observed in this thesis. Proteins with known functions are amongst others a cadherin (Astro 18q22.1), a zinc finger protein (HeLa 15q21.3), a protein tyrosine phosphatase family member (TLy 2q14.2) and a RNA processing protein (HeLa Xq22.1).

Consistent with (Schröder et al., 2002) all seven integration sites within genes were located in intron sequences (Table 1 page 57). Concerning the gene density of the surrounding environment, integration sites were found in gene poor (HeLa 11q22.3, HeLa 18q22.3, Astro 18q22.1) as well as in gene dense regions (HeLa

16p13.3, HeLa Xq22.1, TLy 2q11.2) (see Figure 15, Table 8). The GC content varies too. Taken together, the only preference of HIV integration site selection noticed here was integration within genes. Both the gene density and GC content of the surrounding environment did not seem to be pivotal or cannot be brought in a considerable relationship. Neither did other genomic properties such as CpG islands or LTR content seem to play an important role (see Table 8 page 115).

Since HIV has to rely on the host cell transcription machinery and therefore makes use of host splicing components, integration sites were investigated in terms of their position towards SC35 splicing speckles. One might expect that integration and strong transcription in a gene poor area such as 18q22.3 in HeLa cells could lead to accumulation of splicing factors to this site. A significant difference between integration sites and respective regions on the homologous chromosomes with respect to splicing speckles, however, was not detected here (see Figure 38). Moreover, none of the BAC signals covering the integration sites were located adjacent to splicing speckles, except for HeLa 18q22.3 where 15 - 20% contacted or co-localized with SC35 signals (see Figure 38). This is consistent with (Bell et al., 2001), where HIV integration site location towards splicing speckles was observed to be random. By contrast 30% of HIV RNA signals in HeLa cells and T-lymphocytes contacted SC35 signals (see Figure 39). However, this can be explained with the larger volume of RNA signals compared to BAC signals.

On the molecular level of the linear genome HIV preferentially integrates within active transcription units (Wang et al., 2007). Whether preferences with regard to the 3D nuclear space exist, however, has not been reported yet. The distribution curves of the BAC signals not co-localizing with the HIV signals (Figure 29, green curves) should in principle reflect the nuclear radial location for integration site selection. Even though comparing results determined in nuclei of different shapes is critical and should not be taken exclusively, in the five scenarios in Figure 29 no obvious preferences or tendencies were apparent. The peaks of the distribution curves in Figure 29 show both peripheral (HeLa 18q22.3, Astro 18q22.1) and more internal (TLy 2q11.2) positions. Furthermore the mean values of the medians (Figure 30, green circles) range from -1.42 µm (TLy 2q11.2) to -0.75 µm (Astro 18q22.1).

5.6 HeLa cells and T-lymphocytes harbored multiple HIV integration sites

Target cells of HIV were generally considered to bear only one single infection per cell. During the last decade this prevalent assumption was questioned by several counterexamples based on experiments both in in vitro infected and in patient cells. Spleen cells of HIV infected individuals, for example, were averaging three or four proviruses, up to eight (Jung et al., 2002). In a T cell line infected with a HIV-1-derived virus double infections were found significantly more often than expected from random events (Dang et al., 2004). The probability of multiple infections was based on calculations from the multiplicity of infection value. The findings in this thesis join these counterexamples, since in two of three cell types in vitro infected with HIV-1 more than one provirus per nucleus was detected (see 4.2.3). Only astrocytes had constantly one integration site per nucleus. In the T-lymphocyte cell type KE37/1-IIIB up to four integrations per nucleus were microscopically visualized, with one or two integrations on average. In the HIV infected HeLa cells rarely two integration sites per nucleus were visualized, mostly only one.

In advance, a quantification via LightCycler qPCR revealed statistically for the astrocytes one provirus copy per nucleus, for the T-lymphocytes six and for the HeLa cells 0.6 (Manja Meggendorfer, Helmholtz Zentrum München, personal communication). Thus, among the HeLa cell population not all cells were infected. Determining the number of integration sites via FISH methods is based on single cell observations. Since the provirus was detected with a probe that subsequently was visualized by antibody detection, this method of analysis is indirect and might be affected by various factors. As described in the previous chapter detecting sequences of critical sizes less than 10 kb with conventionally labeled FISH probes is challenging. Hence, the number of visualized integration sites can vary. Never the less, results from both 3D FISH (Figure 23) and RNA FISH (Figure 27 and Figure 28) experiments support the assumption that HIV is capable of multiple infections.

As mentioned, in most HeLa cells one integration site was visualized, rarely two (Figure 24). Integration sites 18q22.3 and Xq22.1 were never detected within the same nucleus. Possible explanations would be that in HeLa cells either multiple HIV integrations are a rare event or selectively disadvantageous. In the KE37/1-IIIB T-lymphocytes both mapped integration sites were located on chromosome two. Microscopic Images revealed that they were located on the same chromosome, not

on different homologs. Frequently two additional integration sites were visible simultaneously.

Besides direct infection, HIV is able to choose a cell mediated entry pathway. Dendritic cells are known to capture several virions of HIV and transport them to their target T-cells (Grouard and Clark, 1997; Geijtenbeek et al., 2000; McDonald et al., 2003). In case of multiple integrations a certain importance is assigned to this cell mediated entry pathway (McDonald et al., 2003; Chen et al., 2005). The observations made in this thesis, however, offer an example of multiple HIV integrations in in vitro infected cells.

5.7 The integration locus Xq22.1 harboring clone was the most dominant among the LC5-HIV HeLa cell population

In 4.2.4 it was determined that the LC5-HIV HeLa cells are a pool of clones with different integration sites. The cell line was obtained by infection of LC5 cells with the HIV-1 isolate HTLV-III$_B$ and subsequent end-point dilution technique (Mellert et al., 1990). qPCR experiments revealed 0.6 HIV integrations per nucleus. In this thesis six integration sites were mapped but never more than two were microscopically visible, in most cases one (Figure 24). As described in 4.2.2 the HIV probe detected DNA and RNA, which increased the signal size and facilitated the visualization of the HIV provirus in FISH experiments. If an integration site is insufficiently transcribed or inactive, it is possible that due to the lack of RNA it is not visible. In FISH experiments on metaphases with the HIV probe and BACs marking integration site regions 11q22.3, 18q22.3 and Xq22.1, only one provirus per metaphase spread was found (see 4.2.4). Hence, these three integration sites are found in different clones.

Interestingly, clones with different integration sites were resented among the HeLa cell population with variable frequency (Figure 25). The most prominent member was the clone harboring integration site Xq22.1 with 40%, whereas integration site clone 18q22.3 accounted only for 7% of the population. Obviously, some clones have selective advantages over others, which might be traced back to different genomic positions of integration sites. An in vivo clonal selection upon retroviral integrations has been described earlier for blood cells (Deichmann et al., 2007). There, in an X-linked SCID (severe combined immunodeficiency) gene therapy trial CD34[+] cells of patients were transduced and retransplanted. In post-

transplanted cells (4 – 41 months after retransplantation) integration site distribution among the population was different indicating selective or growth advantages for clones harboring distinct integration sites. Gene activation upon integration in gene loci that govern growth or survival of the affected cells was suggested to account for this observation. In another gene therapy trial, where X-linked chronic granulomatous disease (CGD) was treated, similar in vivo expansion of distinct clones was observed (Ott et al., 2006). There, clones showing gene vector integration sites in or near the genetic loci MDS1-EVI1, PRDM16 or SETBP1 became predominant approximately five months after reinfusion of CD34$^+$ cells. Transcriptional levels of the affected genes increased upon integration. A relation between gene vector integrations in the MDS1-EVI1 region and an in vivo clonal dominance has also been described before for mouse bone marrow cells (Kustikova et al., 2005). It was frequently observed that expression of genes in the vicinity of retroviral integration sites can be altered upon integration (see 5.9).

The HeLa cell clone with integration in region Xq22.1 accounted for 40% and therefore dominated the cell population. The provirus in Xq22.1 was located within the CSTF2 gene (Table 1 page 57) which encodes a protein involved in RNA processing. Genes at integration sites in less represented clones encode for a transcription initiation protein (integration locus 6p12.3), a zinc finger protein (integration locus 15q21.3) and a protein of the Wnt signalling pathway (integration locus 16p13.3). These clones together made up 10% of the HeLa cell population. Whether transcription of these genes was upregulated or affected somehow by HIV integration, which might influence growth behaviour of the cell, was not investigated, since it was not in the focus of this thesis. Only the gene AXIN1 in close vicinity to integration locus 16p13.3 with its ability to induce apoptosis allows drawing an obvious connection to cell survival. How exactly the HeLa cell sub-clone population arose, is not clear.

For the KE37/1-IIIB T-lymphocytes results from qPCR experiments suggested up to six integrations per nucleus, with the LM-PCR technique two of them were mapped in this thesis. In the subsequent FISH experiments between one and four HIV signals per nucleus were visible (see Figure 24), the count of HIV RNA signals varied too (see Figure 26). It is not clear whether the T-lymphocytes were also a population of subclones, which would explain the non-constant number of DNA FISH signals. The qPCR results and the varying number of HIV RNA signals rather

suggest that the T-lymphocytes harbor four or more integration sites, which might have different transcriptional activities. Therefore, the T-lymphocytes might be an example for latent HIV infections within a population. It remains elusive on which mechanism this putative latency was based. High levels of CpG methylation of the viral promoter have been shown to result in viral latency (Blazkova et al., 2009) as well as distinct epigenetic modifications like hypoacetylation (Coull et al., 2000).

5.8 Radial nuclear repositioning was determined for two HIV integration loci

Five of the nine mapped HIV integration sites could be investigated in terms of their 3D nuclear position (4.1.4). Two of them (HeLa 18q22.3 and HeLa Xq22.1) were significantly located more internal than their respective regions on the homologous chromosomes (see Figure 29). The change was more pronounced for 18q22.3 than for Xq22.1 (0.24 µm compared to 0.12 µm mean values of median in Figure 30).

Peripheral chromatin is in molecular contact with the nuclear lamina (Kind and van Steensel, 2010). These chromatin properties might constrain repositioning of distinct genetic loci. Repositioning in a flat astrocyte nucleus, where more chromatin is in closer proximity of the lamina, might be more limited than in round nuclei like T-lymphocytes. Here, in neither of them repositioning of the investigated integration loci was observed. The repositioned integration locus 18q22.3 was in hemispherical HeLa nuclei and in close vicinity to the nuclear surface (HeLa 18q22.3, peak of green curve at -0.125 µm, Figure 29). Within that distance away from the nuclear lamina, attachment did not seem to prevent repositioning. Independent from lamina attachment, due to its absence, would be repositioning during mitosis. At which particular time after integration repositioning took place and at which cell cycle stage could not be determined.

The repositioned integration loci HeLa 18q22.3 and HeLa Xq22.1 differed both in gene density and GC content. Integration site in region 18q22.3 was between genes and in a very gene poor region (no genes within a 0.5 Mb window). On the other hand, integration site in region HeLa Xq22.1 was located within a gene and in a gene dense region. The grade of gene density seemed not to elicit repositioning. Taking all five integration loci into account, a high variation in gene density, GC content and LTR content was found (Figure 15 and Table 8).

Repositioning is often implicated in transcriptional alterations. Several publications report gene activation upon movements towards the nuclear interior (see introduction 2.3.2). In (Gierman et al., 2007) a scenario of an integrated transgene acquiring the expression level of the surrounding chromosomal area is described. In particular, epigenetic properties at HIV integration sites and the surrounding genomic regions can influence the degree of provirus expression (Jordan et al., 2001; Wang et al., 2007). Notwithstanding its preference for active transcription units, HIV integration sites might be surrounded by transcriptional environment not satisfying for viral purposes. Since gene expression is proposed ot be affected by the nuclear location (Zhao et al., 2009), repositioning towards more suitable regions offers a solution to improve transcriptional competence. In (Zink et al., 2004) transcriptional activation is proposed to be the driving force for repositioning. Transcriptional activity affects its nuclear position rather than vice versa. This theory would provoke repositioning of the HIV integration sites as a result of increased transcription.

In (Lilley et al., 2009) viral interactions in advance to integration with host chromatin modifications such as histone methylations, phosphorylations or acetylations and repair mechanisms induced by the virus are described. These events might lead to alterations in nuclear architecture which provoke repositioning. Actually, retrovirally caused changes in chromatin condensation were observed (see 4.4 and following chapter).

Considering repositioning under the light of increasing transcription, it is not a must for the integration site to move to another nuclear location if transcriptional competence is already given or sufficient. Another suggestion to guarantee transcription could be a reorganization in the chromatin environment around the integration site as described in (Dieudonne et al., 2009). In the silent state the HIV-1 provirus was observed to be associated with pericentromeric chromatin regions, being transcriptional activated this association got lost, without taking influence on its nuclear location. In this thesis all investigated integration sites were transcriptionally active, because the HIV probe always detected certain amounts of RNA (see 4.2.2). Having that in mind one can guess that the integration sites that did not reposition either were already convenient for transcription or became so by modification of the neighboring chromatin.

A hint for that is given by the results of the sodium butyrate (NaBut) experiments (see Figure 35). NaBut treatment verifiably increased HIV transcription

(see Figure 34c), whereas the strongest effect was measured for the HeLa cells with a ten time higher amount of HIV transcripts than in untreated cells. An explanation for the quite minor increase in T-lymphocytes could be the already high transcriptional activity in untreated cells (Figure 34a), so a tremendous increase upon NaBut treatment might not be compulsorily expectable. Nuclear positions of HIV integration sites and their respective loci on the homologous chromosomes in NaBut treated cells did not differ significantly compared to untreated cells (compare Figure 29 and Figure 35). That could mean, although transcription was boosted, repositioning was not necessary as the current location seems probably transcriptionally satisfying. It has to be taken into account that the measured effect on transcription of the NaBut treatment is representative for the whole cell population of each cell type (astrocytes, HeLa cells and T-lymphocytes). This means, in case of the latter two cell types, increase in transcription was determined for all integration sites together and not for each single one.

The observed repositioning was in both cases (HeLa 18q22.3 and HeLa Xq22.1) with regard to the nuclear surface. The integration site position did not change with regard to the surface of the respective chromosome territory (see Figure 31c,d). Hence, it is no looping out scenario with an emanating chromatin fiber from its harboring chromosome territory as described in several studies for gene loci upon transcriptional activation (see introduction 2.3.2). There are several mechanistic explanations. The integration site changed its location within the chromosome territory but not the distance to its surface. The other possibility might be that a larger area or component within the chromosome territory around the integration site changed its position or the whole chromosome territory rotated. Both repositioned integration sites harboring chromosome territories X and 18 did not change their nuclear distribution (see Figure 31c,d). A contrary event is reported in (Szczerbal et al., 2009). There, adipogenesis genes were found to move towards the nuclear center upon transcriptional activation and the respective chromosome territories also altered their nuclear distribution.

The LC5-HIV HeLa cells were a population of sub-clones with several integration site clones. Transcription of genes in the integration site regions can therefore not be determined by qPCR since the results would display the properties of the whole population. It would be interesting to see whether repositioning correlates with an up-regulation of genes at the integration sites. Gene function and

expression levels can be tremendously affected by insertion of a provirus directly within or in the vicinity of a gene (Thomas et al., 2003; Bushman et al., 2005).

Among the HeLa cell population, the clone harboring the most repositioned integration site region 18q22.3 only accounted for 7%. The cell clone harboring the other repositioned integration site Xq22.1 dominated the population with 40%. Whether less repositioning contributes to this selective advantage or the genomic position of the provirus takes the main influence, remains elusive.

5.9 Repositioning of HIV integration loci was accompanied by chromatin decondensation

Depending on the grade of compaction and transcriptional competence, chromatin is classified in euchromatin and heterochromatin. Compared to euchromatin, heterochromatin is more condensed and transcriptionally less active (see introduction 2.3.1). In this thesis both upon HIV integration repositioned loci (HeLa 18q22.3 and HeLa Xq22.1) were significantly decondensed compared to their respective loci on the homologous chromosomes. This was determined by the different surface areas of BACs which marked the integration sites (Figure 33). Since the BAC probes did not detect the provirus but only the surrounding host genome, the measured increase in signal size clearly evolved from chromatin enlargement and not from the accretion of nucleotides. Obviously, neighboring host cell chromatin structure was influenced by the integrated HIV provirus. Chromatin decondensation accompanying transcriptional activation of a genetic locus has been observed earlier, e.g. for the Hoxb locus (Chambeyron and Bickmore, 2004) or an activated transgene (Dietzel et al., 2004). In the last case, repositioning was observed too.

HIV integrations as well as integrations of HIV based vectors are known to modulate expression levels of certain host genes, whereas both upregulation and downregulation was found (Mitchell et al., 2003; van 't Wout et al., 2003). Expressional alteration was described for several transcription factors (van 't Wout et al., 2003) or genes involved in the interferon response, DNA-damage response and apoptosis (Corbeil et al., 2001). Transcription of genes proximal to integration sites can be affected especially, when located downstream of the provirus. A read through the 3´LTR polyadenylation signal, by which subsequent host genes might be activated, was suggested in (Ashe et al., 1995).

Whether the decondensation led to transcriptional up-regulation of adjacent genes remains elusive. At least for integration site HeLa 18q22.3, which was 700 kb away from the next gene, the BAC probe, due to its size of 146 kb (see Table 10 page 117), was not capable of detecting RNA of possible transcribed genes in this region. The genomic region covered by the BAC probe at integration site region Xq22.3 in HeLa cells contains two genes (CSTF2 and NOX1) which possibly would allow by transcription certain amounts of RNA to be detected. This might result in a measurable increase in signal size. However, this assumption tends to be unlikely since in RNA FISH experiments using the BAC probes a distinct RNA signal was not visible (data not shown). For both repositioned integration site regions, HeLa 18q22.3 and HeLa Xq22.1, the contribution of RNA to the increase in BAC signal size can thus be disregarded.

Integration locus 18q22.3 in HeLa cells was repositioned and decondensed strongest. This locus is gene poor and 700 kb distant from the next gene. The less repositioned and decondensed integration locus Xq22.1 was located within a gene and in a gene dense region. A possible explanation for the different alterations in chromatin compaction might be that the gene blank chromatin was more compacted and therefore could be decondensed more than already less condensed, gene richer chromatin. This assumption is underlined by the signal sizes of the BACs that did not co-localize with the HIV signal (green bars in Figure 33). In principle, these BACs display the properties of the integration site before integration. The BACs marking integration loci 18q22.3 and Xq22.1 in HeLa cell feature approximately the same sizes (146 and 155 kb). At same chromatin compaction they should posses the same surface area. Apparently, the BAC signal for region 18q22.3 was smaller and therefore the chromatin was more condensed than the BAC for integration locus Xq22.1 (259 px to 316 px).

The influence of surrounding chromosomal regions on the transcription level of a given inserted gene is exemplified in (Gierman et al., 2007). There, embryonic kidney cells were transduced with a lentiviral construct containing a GFP reporter gene. GFP expression gave information about the transcriptional level of the inserted construct. GFP expression of integration sites located in highly expressed domains called ridges was 4-fold higher than the identical gene integrated in anti-ridges. Furthermore FISH with BAC clones for the integration sites revealed a significantly larger diameter of the clones covering the highly expressed reporter gene region than

the respective region with low expression in anti-ridges. Transferred to HIV integrations, this would suggest that, depending on the genomic environment of integration sites, HIV may be transcribed at varying intensities and therefore provide different chromatin condensation states.

6. APPENDIX

6.1 Abbreviations

3C	chromosome conformation capture
4C	circular chromosome conformation capture
2D, 3D	2-dimensional, 3-dimensional
Astro	astrocytes
BAC	bacterial artificial chromosome
bp	base pair
CT	chromosome territory
DAPI	4',6-diamidino-2-phenylindole
dH$_2$O	distilled water
ddH$_2$O	double-distilled water
eGFP	enhanced green fluorescent protein
EtOH	ethanol
Evi-1	ecotropic virus integration site 1
FISH	fluorescence in situ hybridization
gDNA	genomic DNA
h	hour
HM	hybridization mixture
HIV	humane immunodeficiency virus 1
HSA	homo sapiens
kb	kilobase
LTR	long terminal repeat
LM-PCR	ligation (linker) mediated polymerase chain reaciton
Mb	megabase
min	minutes
MMU	mus musculus
NaBut	sodium butyrate
ng, µg, mg	nanogram, microgram, milligram
µl, ml	microliter, milliliter
nm, µm	nanometer, micrometer
o/n	over night
pmol	pico mol

PSF	point spread function
qPCR	quantitative polymerase chain reaction
RT	room temperature
sec	seconds
siRNA	small interfering RNA
SFFV	spleen focus-forming virus
TLy	T-lymphocytes
VRC	*V*anadyl ribonucleoside complex

6.2 Location of integration sites

6.2.1 HIV integration sites

Integration locus	HeLa 6p12.3	HeLa 11q22.3	HeLa 15q21.3	HeLa 16p13.3	HeLa 18q22.3	HeLa Xq22.1	TLy 2q11.2	TLy 2q14.2	Astro 18q22.1
Nearest gene	SUPT3H	PDFGD	ZNF280D	AXIN1	LOC100505817	CSTF2	KIAA1310	PTPN4	CDH19
Gene ID	8464	80310	54816	8312	100505817	1478	55683	5775	28513
Distance to nearest gene	0	250 kb	0	0	700kb	0	0	0	0
Gene description	Encodes Transcription initiation protein SPT3 homolog	platelet derived growth factor D	zinc finger protein 280D	Cytoplasmic protein; Wnt signaling pathway; can induce apoptosis	misc. RNA	cleavage stimulation factor, 3' pre-RNA, subunit 2	hypothetical protein	protein tyrosine phosphatase, non-receptor type 4 (megakaryocyte)	cadherin 19, type 2
Chromosome integrity	-	-	-	-	+	+	+	+	+
sequence	GCAAGTTTTCAGTTACTGTGCTTGATAGGC	AAATCATTAAGTTGTTAATGGAATTTTAAG	CAGAAAAATGCAAATAGGGTCATTTTAGTC	GGGCTTTCCCTAGGTTCTGTTCATCAGGTG	CATAGAGCAAGGGGCCAACAGACCCCCTTA	ACAGCACAACTGGATACCACTGAATTAATC	GGTGGGTCTGGACCACTCCTGGATTCTGGG	TGCCTTCCTCAGACTTTGTTACAGCATAGG	ATAACAATATTGTATAATATATGAAGAAAT

Table 6: Position of HIV integration sites
Loci are identified by the chromosomal band in which they sit. A distance of 0 to the nearest gene means the integration is inside the gene. Gene ID and description are from the NCBI human genome database. Chromosome integrity was essential to compare the nuclear distribution to the homologous locus (see main text). 'Sequence' identifies the start of the sequenced host genome from LM-PCR directly adjacent to the 5´LTR of the integrated provirus.

6.2.2 Gene transfer vector integration sites in mouse cells

Integration locus	MMU 3A3[+]	MMU 5G1[+]	MMU 9A3	MMU 11B5	MMU 11E1[+]	MMU 12B3[+]	MMU 12D2	MMU 16B2	MMU XA1.1[+]
Nearest gene	Evi1	Mad1like1	Dnm2	Lig3	Gm885	Akap6	Batf	AI480653	Wdr45
Gene ID	14013	17120	13430	16882	380732	238161	53314	268880	54636
Distance to nearest gene	121 kb	0	0	2.6 kb	0	0.4 kb	13.6 kb	0	0.4 kb
Sequence	AATTAGGA	GGAGGGCCTGAGGGGAAACTCCAATAAATT	TCACAGACACTCTGTATCTATTTGCAAATT	TGAGGGGTCACTATAGTTGAAGCTGGTCAG	GAAGTCAGAAGACACTCTATGGGAGTTGAT		GAGAACGTGGTCTTGCGTGTCCTGAGCCTA	GAGATCCAACTATAAGGCATTTTTTCAATT	GCTGGAATAAAAGATCTTGAAAGTTGAGAT
Start of sequenced gDNA (Tsp5901 cut)		140404209	21206705	82596948	106575437	53620480	86562465	30913115	6879245
End of sequenced gDNA (Begin of 5'LTR)		140404081	21206612	82596769	106575524	53620440	86562199	30912770	6878793

Table 7 Position of gene transfer vector integration sites
Integration locus description defines the chromosomal band; Loci marked by + have been described before (Modlich et al., 2006); gene ID from NCBI mouse genome database; Sequences from LM-PCR performed in the lab of Manuel Grez: end of sequenced host genome adjacent to 5´LTR of integrated gene vector except for MMU11E1 where the sequence displays the start of the sequenced host genome adjacent to the 5´LTR; Start and end positions of gDNA state 02/07.

6.3 Genomic properties of integration site environment

6.3.1 HIV integration sites

	HeLa 6p12.3	HeLa 11q22.3	HeLa 15q21.3	HeLa 16p13.3	HeLa 18q22.3	HeLa Xq22.1	Tly 2q11.2	Tly 2q14.2	Astro 18q22.1
Total Genes on chromosome	1959	2051	1202	1318	517	1606	2204	2204	517
Total Mbp of chromosome	171	134	100	88,8	76,7	153	243	243	76,7
Genes on chromosome per Mb	11,46	15,31	12,02	14,84	6,74	10,50	9,07	9,07	6,74
Total CpG islands on chromosome	16263	14376	10089	15182	7031	13486	22388	22388	7031
Total Hs RNA on chromosome	410021	473462	232976	314942	112196	239916	466689	466689	112196
Total Repeats on chromosome	236957	199288	126356	143735	104675	211180	342695	342695	104675
Total GC content on chromosome	40%	42%	42%	44%	40%	39%	40%	40%	40%
0.5 Mb window									
Genes On Sequence	3	1	5	27	0	17	12	6	1
Genes On Sequence per Mb	6	2	10	54	0	34	24	12	2
Hs RNA on Sequence	158	95	465	12134	51	2014	2313	618	95
Hs RNA per Mb	316	190	930	24268	102	4028	4626	1236	190
CpG islands on Sequence	23	31	26	331	30	83	113	65	28
CpG islands per Mb	46	62	52	662	60	166	226	130	56
Repeats on Sequence	657	735	747	818	688	805	869	806	646
Repeats per Mb	1314	1470	1494	1636	1376	1610	1738	1612	1292
GC content	35.93%	35.76%	37.46%	55.26%	35.88%	41.68%	47.71%	39.53%	34.38%
LTR content	5.11%	9.27%	9.55%	1.79%	10.22%	8.22%	9.33%	6.86%	11.94%
2 Mb window									
Genes On Sequence	17	15	17	83	8	43	40	22	4
Genes On Sequence per Mb	8,5	7,5	8,5	41,5	4	21,5	20	11	2
Hs RNA on Sequence	12684	1185	3883	19645	1015	6961	9007	3185	253
Hs RNA per Mb	6342	592,5	1941,5	9822,5	507,5	3480,5	4503,5	1592,5	126,5
CpG islands on Sequence	168	82	160	937	133	228	292	272	115
CpG islands per Mb	84	41	80	468,5	66,5	114	146	136	57,5
Repeats on Sequence	2965	2790	2990	1686	2612	2856	2866	3179	2746
Repeats per Mb	1482,5	1395	1495	843	1306	1428	1433	1589,5	1373
GC content	41.72%	44.75%	39.25%	58.6%	38.36%	40.17%	44.88%	45.55%	34.75%
LTR content	6.66%	7.06%	6.32%	2.62%	12.34%	7.47%	6.88%	7.18%	11.1%
10 Mb window				+/-2.7 Mb			+/-3.5 Mb		
Genes On Sequence	147	78	85	293	38	116	118	47	45
Genes On Sequence per Mb	14,7	7,8	8,5	29,3	3,8	11,6	11,8	4,7	4,5
Mm RNA on Sequence	35579	14663	23872	69186	27497	22415	21064	7114	5571
Mm RNA per Mb	3557,9	1466,3	2387,2	6918,6	2749,7	2241,5	2106,4	711,4	557,1
CpG islands on Sequence	1057	587	917	3368	1075	757	951	757	628
CpG islands per Mb	105,7	58,7	91,7	336,8	107,5	75,7	95,1	75,7	62,8
Repeats on Sequence	15632	14414	15358	18546	12669	13095	10088	14346	13953
Repeats per Mb	1563,2	1441,4	1535,8	1854,6	1266,9	1309,5	1008,8	1434,6	1395,3
GC content	41.67%	37.19%	39.61%	53.61%	39.22%	38.45%	43.29%	39.99%	37.02%
LTR content	7.38%	9.57%	6.05%	3.81%	9.6%	10.75%	8.33%	9.96%	10.08%

Table 8: Genomic properties of HIV integration sites (state 08/09)
HeLa – HeLa cells; Tly – T-lymphocytes; Astro – astrocytes; data taken from NCBI homepage by blasting respective sequences except for GC content and LTR content which were taken from Repeatmasker homepage; total GC content on chromosome from (Venter et al., 2001). Gene density is given as genes per megabase, calculated for 0.5, 2 and 10 megabase windows around the integration site, and for the whole chromosome. The GC content in percent is given for the same windows. 10 Mb window incomplete due to genomic position in case of HeLa 16p13.3 (telomeric) and Tly 2q11.2 (gap in sequenced contig).

6.3.2 Gene transfer vector integration sites in mouse cells

	MMU 3A3	MMU 5G1	MMU 9A3	MMU 11B5	MMU 11E1	MMU 12B3	MMU 12D2	MMU 16B2	MMU XA1.1
Total Genes on chromosome	1579	1773	1640	2095	2095	1371	1371	973	1847
Total Mbp of chromosome	160,0	153,0	124,0	122,0	122,0	121,0	121,0	98,0	167,0
Genes on chromosome per Mb	9,9	11,6	13,2	17,2	17,2	11,3	11,3	9,9	11,1
Total CpG islands on chromosome	9164	12975	8973	10675	10675	7671	7671	5849	6355
Total Mm RNA on chromosome	218633	286526	272565	380713	380713	157540	157540	126834	133215
Total Repeats on chromosome	239326	261668	211925	225939	225939	191084	191084	152863	254111
Total GC content on chromosome	40.5%	42.5%	43.0%	44.0%	44.0%	41.5%	41.5%	41.0%	39.0%
0.5 Mb window									
Genes On Sequence	2	9	20	12	10	2	16	3	26
Genes On Sequence per Mb	4	18	40	24	20	4	32	6	52
Mm RNA on Sequence	122	897	4781	1417	3549	474	3241	486	4135
Mm RNA per Mb	244	1794	9562	2834	7098	948	6482	972	8270
CpG islands on Sequence	22	75	93	40	52	47	80	48	66
CpG islands per Mb	44	150	186	80	104	94	160	96	132
Repeats on Sequence	617	1017	1649	1017	1549	817	1264	834	1177
Repeats per Mb	1234	2034	3298	2034	3098	1634	2528	1668	2354
GC content	39.85%	51.25%	49.05%	46.92%	45.25%	41.06%	45.05%	46.69%	38.38%
LTR content	3.93%	7.26%	4.8%	9.9%	6.35%	7.74%	7.18%	15.35%	13.53%
2 Mb window									
Genes On Sequence	18	42	79	48	44	13	43	24	66
Genes On Sequence per Mb	9	21	39,5	24	22	6,5	21,5	12	33
Mm RNA on Sequence	1414	6507	15203	4986	12512	1565	8748	3480	7651
Mm RNA per Mb	707	3253,5	7601,5	2493	6256	782,5	4374	1740	3825,5
CpG islands on Sequence	102	311	352	133	247	124	261	206	163
CpG islands per Mb	51	155,5	176	66,5	123,5	62	130,5	103	81,5
Repeats on Sequence	2902	4770	6189	3517	5285	3285	4719	4010	4147
Repeats per Mb	1451	2385	3094,5	1758,5	2642,5	1642,5	2359,5	2005	2073,5
GC content	40.58%	50.46%	47.91%	45.29%	45.06%	41.12%	45.87%	45.4%	43.08%
LTR content	7.12%	7,02%	10.51%	13.86%	5.72%	6.94%	7.55%	13.57%	1.82%
10 Mb window									
Genes On Sequence	72	234	173	187	174	85	171	103	188
Genes On Sequence per Mb	7,2	23,4	17,3	18,7	17,4	8,5	17,1	10,3	18,8
Mm RNA on Sequence	9661	41296	18702	34635	34412	10449	23501	16617	12848
Mm RNA per Mb	966,1	4129,6	1870,2	3463,5	3441,2	1044,9	2350,1	1661,7	1284,8
CpG islands on Sequence	633	1436	709	923	1157	590	853	819	576
CpG islands per Mb	63,3	143,6	70,9	92,3	115,7	59	85,3	81,9	57,6
Repeats on Sequence	16103	25388	16869	23103	21948	16677	19371	18440	14943
Repeats per Mb	1610,3	2538,8	1686,9	2310,3	2194,8	1667,7	1937,1	1844	1494,3
GC content	41.47%	47.89%	41.34%	44.75%	44.51%	40.21%	43.86%	43.23%	41.94%
LTR content	10.91%	9.74%	12.83%	7.99%	8.16%	9.65%	10.26%	10.88%	15.53%

Table 9: Genomic properties of gene transfer vector integration sites (state 11/10)
Data taken from NCBI homepage by blasting respective sequences except for GC content and LTR content that were taken from Repeatmasker homepage; total GC content on chromosome from (Waterston et al., 2002). Gene density is given as genes per megabase, calculated for 0.5, 2 and 10 megabase windows around the integration site, and for the whole chromosome. The GC content in percent is given for the same windows.

6.4 BAC properties

6.4.1 BACs covering HIV integration sites

Integration Site	HeLa 6p12.3	HeLa 11q22.3	HeLa 15q21.3	HeLa 16p13.3	HeLa 18q22.3	HeLa Xq22.1	TLy 2q11.2	TLy 2q14.2	Astro 18q22.1
BAC name	RP11-818O21	RP11-63H12	RP11-566D24	RP11-517F15	RP11-693I21	RP11-255J06	RP11-67L23	RP11-132N24	RP11-831H17
Start [Mb]	45.10	103.66	54.60	0.23	96.38	99.88	97.23	120.34	-
Stop [Mb]	45.28	103.85	54.79	0.38	96.53	100.03	97.33	120.48	-
Length [kb]	182	194	189	150	146	155	100	139	-
State	06/09	01/08	06/09	06/09	01/08	01/08	07/09	07/09	07/07
Database version	54	49	54	54	49	49	55	36.3	-

Table 10: HIV integration sites – BAC properties
BACs were chosen from Ensembl database except RP11-132N24, which was chosen from NCBI database and RP11-831H17, which was kindly provided by Mariano Rocchi, University of Bari, Italy.

6.4.2 BACs covering gene transfer vector integration sites

Integration Site	MMU 3A3	MMU 5G1	MMU 9A3	MMU 11B5	MMU 11E1	MMU 12B3	MMU 12D2	MMU 16B2	MMU XA1.1
BAC name	RP23-439N22	RP23-217N11	RP23-317E10	RP23-316C11	RP23-247J12	RP24-267H3	RP23-369N11	RP24-206H1	RP23-54C14
Start [Mb]	30.15	140.32	21.13	82.50	106.46	53.56	86.49	30.81	6.69
Stop [Mb]	30.37	140.52	21.31	82.68	106.67	53.67	86.64	31.02	6.90
Length [kb]	220	200	175	180	210	113	144	212	215
State	10/06	12/07	07/08	02/07	02/07	02/07	02/07	02/07	02/07
Database version	41	48	50	43	43	43	43	43	43

Table 11: Gene transfer vector integration sites – BAC properties
All BACs were chosen from Ensembl database.

6.5 Linker and Primer sequences

6.5.1 LM-PCR described in 3.4.2:

Linker +: 5'-GACCCGGGAGATCTGAATTCAGTGGCACAGCAGTTAGG**T** -3`

Linker -: 5'-Pho-CCTAACTGCTGTGCCACTGAATTCAGATCTCCCG - 3'

<u>Linker specific primers for nested PCRs:</u>

LSP1 (outer): 5' - GACCCGGGAGATCTGAATTC - 3'

LSP2 (inner): 5' - AGTGGCACAGCAGTTAGG - 3'

<u>HIV provirus specific primers:</u>

Bio-LMPCR-1: 5´- TAACTTCTCTGGCTCAACTGGTAC - 3´

LMPCR-2 (outer): 5´ - TGGCCCTGGTGTGTAGTTCTG - 3´

LMPCR-3 (inner): 5´- AGGATATCTTGTCTTCGTTGGGAG - 3´

6.5.2 Control PCRs described in 4.1.2:

Primer within 3´LTR (25 bp before end of provirus) towards host genome:

GSP3LTR: 5´- AACTAGAGATCCCTCAGACCCTTTTAGTCAG -3´

Primers within host genome towards provirus 3´LTR (approx. 1 kb away); specific for each integration site (indicated in brackets):

LC5_#11 (HeLa 11q22.3): 5´-AGACGTCTAGCTAAAAGAAAATTCTGATAT-3´

LC5_#18 (HeLa 18q22.3): 5´-GTATGGCAAAGATGGAAAATCTCCACAAATC-3´

LC5_#X (HeLa Xq22.1): 5´-AGCTACTGTAGTGGGATATGAAGCAAAATGGA-3´

Primer within 3´LTR (100 bp away from provirus end) towards host genome:

GSP3LTR_100: 5´-AACTAGGGAACCCACTGCTTAAGCCTCAAT-3´

6.5.3 Universal Genome Walker Kit described in 3.4.3:

<u>HIV provirus specific primers:</u>

GSP1 (outer): 5´- AAAGGTCAGTGGATATCTGATCCCTGGCCC-3´

GSP2 (inner): 5´ - CAGGGAAGTAGCCTTGTGTGTGGTAGATCC-3´

118

Linker (adaptor) sequence and adaptor primers:
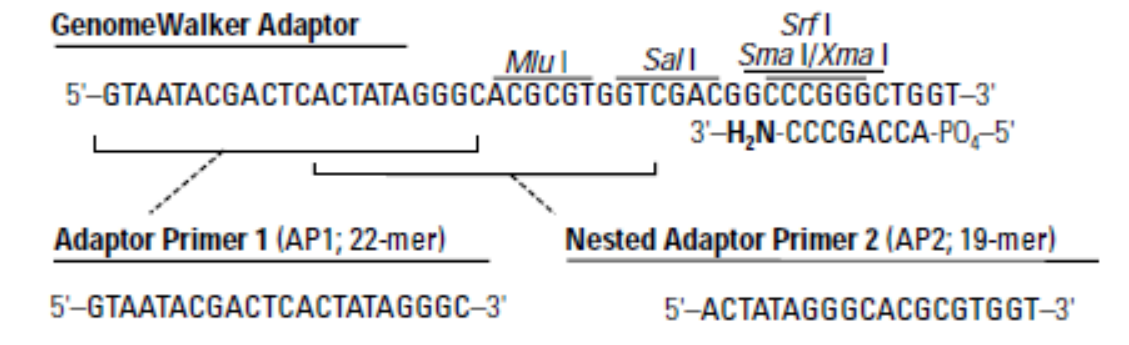
Image taken from Genome Walker Universal Kit User Manuel (Clontech Laboratories)

6.5.4 qPCR described in 3.2.3:

RNA Pol II (s): 5´-GCACCACGTCCAATGACAT-3´

RNA Pol II (as): 5´-GTGCGGCTGCTTCCATAA-3´

HIV-transcripts F: 5´-CCAGTCACACCTCAGGTACCTTTAAGACC-3´

HIV-transcripts R: 5´-GTGTGTGGTAGATCCACAGATCAAGG-3´

6.6 pNL4-3 (HIV) plasmid map
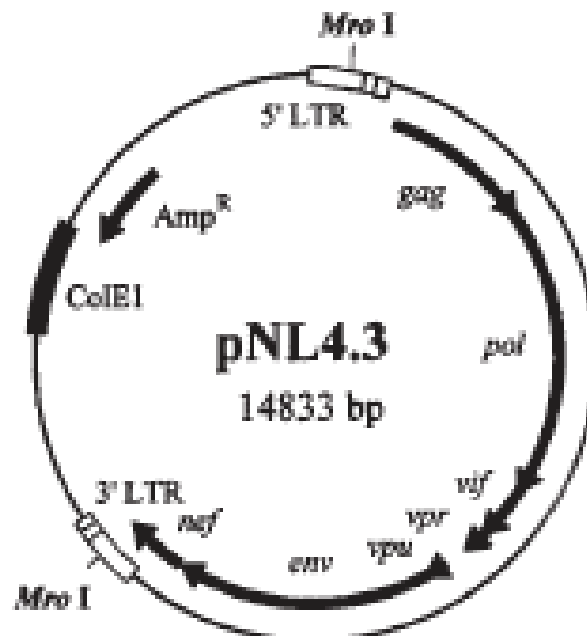
6.7 pSF91-eGFP (gene transfer vector) plasmid map
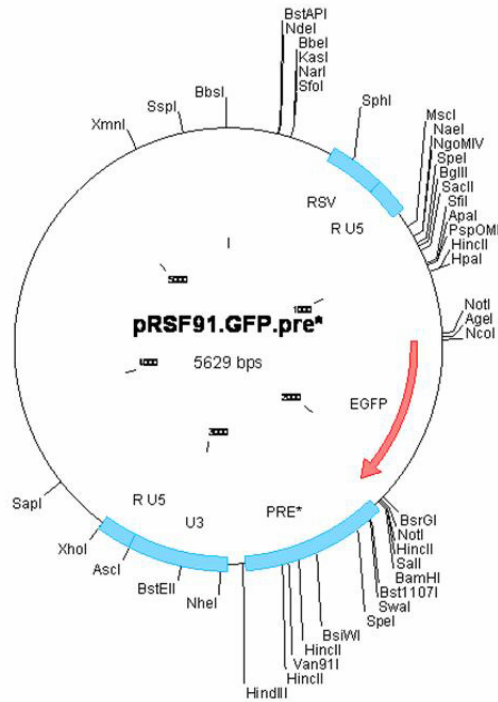
Linear structure of the pSF91-eGFP gene transfer vector (3.188 kb in size):

LTRSF – long terminal repeat from friend spleen focus-forming virus; SD – splice donor site; SA – splice acceptor site; θ – primer binding site; ψ – packaging signal; eGFP – reportergene; PRE – posttranscriptional regulatory element from Woodchuck Hepatitis virus; image taken from (Modlich et al., 2006).

6.8 Antibodies

Antibody	Dilution	Company, Distributor
Mouse-α-Ki67	1:100	Roche, Mannheim, Germany
Mouse-α-DIG-Cy3	1:100	Dianova, Hamburg, Germany
Mouse-α-DIG-Cy5	1:75	Dianova
Mouse-α-SC35	1:100	Sigma-Aldrich, Deisenhofen, Germany
Mouse-α-DIG	1:500	Sigma-Aldrich
Avidin Alexa488	1:200	Molecular Probes (Invitrogen), Karlsruhe, Germany
Streptavidin-Cy5	1:200	Rockland, Gilbertsville, PA, USA
Rabbit-α-DNP	1:200	Sigma-Aldrich
Rabbit-α-triH3K4*	1:500	Abcam, Cambridge, UK
Rabbit-α-triH3K9*	1:400	T. Jenuwein, IMP, Wien, Austria
Rabbit-α-triH3K27*	1:400	Upstate, Lake Placid, NY, USA
Mouse-α-macroH2A*	1:5	Kindly provided by A. Ladurner, EMBL, Heidelberg, Germany
Sheep-α-DIG-FITC	1:100	Roche
Goat-α-Avidin-FITC	1:200	Vector, Burlingame, CA, USA
Goat-α-Mouse-Cy5	1:100	Dianova
Goat-α-Mouse-Alexa488 (hca)	1:400	Molecular Probes (Invitrogen)
Goat-α-Mouse-Alexa350	1:100	Molecular Probes (Invitrogen)
Goat-α-Rabbit-Alexa488 (hca)	1:200	Molecular Probes (Invitrogen)
Goat-α-Rabbit TexasRed	1:50	Dianova
Goat-α-Rabbit-Cy3 (Fab)	1:200	Dianova

Antibodies marked by "*" diluted in PBS buffer; others work both in PBS and SSC buffered solutions.

6.9 Buffers, media and solutions

Buffer/solution/medium	Ingredients	Constitution
Agarose gel	Agarose, TAE-buffer	Appropriate amount of agarose dissolved in respective amount of TAE-buffer by heating-up in microwave
4% BSA in PBS/Tween blocking solution	BSA, 1x PBS/Tween	4 g BSA dissolved in 100 ml 1x PBS/Tween
4% BSA in 4x SSC/Tween blocking solution	BSA, 4x SSC/Tween	4 g BSA dissolved in 100 ml 4x SSC/Tween
CSK buffer	NaCl, Sucrose, Pipes, $MgCl_2$, PMSF	Diluted in DEPC-treated water; final conc.: 100 mM NaCl 300 mM Sucrose 10 mM Pipes 3 mM $MgCl_2$ 1.2 mM PMSF
DEPC-treated water	ddH_2O, DEPC	0.1 ml DEPC are added to 100 ml ddH_2O, incubated 12 h at 37°C and autoclaved for 15 min.;

		stored at 4°C for months until use
Dulbeccos MEM-medium (10% FCS)	Dulbeccos MEM-Medium FCS Penicillin/Streptomycin	500 ml Dulbeccos MEM 50 ml FCS 5 ml Penicillin/Streptomycin
Fixative (fixing metaphases)	Methanol, Acetic Acid	Methanol:Acetic Acid = 3:1
50 % formamide solution	Formamide, 4x SSC	500 ml formamide, 500 ml 4X SSC; pH adjusted on 7.0
20% glycerol	Glycerol, ddH$_2$O	20 ml glycerol diluted in 80 ml H2O bidest.
0.1 N HCl solution	1N HCl, H2O bidest	10 ml 1N HCl diluted in 90 ml H2O bidest
Hybridization mastermix	20% dextran sulfate, 2x SSC	8g dextran sulfate dissolved in 40 ml 2xSSC (vortex) and filtrated (0.45 µm filter)
50 mM MgCl2 solution	1 M MgCl2, 1x PBS	10 ml MgCl2 diluted in 190 ml 1x PBS
10x NT buffer	0.5 M Tris-HCl (pH 7,5) 50 mM MgCl2 0.05% BSA	50 ml 1M Tris-HCl (pH7.5) 5 ml 1M MgCl2 50 mg BSA 45 ml H2O
4 % paraformaldehyde fixation solution	Paraformaldehyd, 1x CMF-PBS	4 g PFA dissolved in 100 ml 1x CMF-PBS (or 0.3x CMF-PBS for fixing in suspension growing cells) by heating-up to approx. 70°C (transparent solution); cooling down to RT before use
20x PBS	140 mM NaCl 2.7 mM KCl 6.5 mM Na2HPO4 1.5 mM KH2PO4	320 g NaCl 8g KCl 57.6g Na2HPO4 9.6g KH2PO4 ad 2l H2O bidest.; adjust pH to 7.4 by adding HCl; autoclaved if used for cell culture
1x PBS	20x PBS H$_2$O bidest.	20x PBS diluted 1:20 in H$_2$O bidest.
1x PBST	Tween 20 1xPBS	100µl Tween 20 ad 1l 1xPBS
Pepsin solution (50µg/ml)	Pepsin stocksolution (100mg/ml) 10mM HCl solution	50µl Pepsin stock solution diluted in 100ml 10mM HCl prewarmed to 37°C
Poly L lysin solution	Poly L lysine, ddH$_2$O	Poly L lysine diluted 1:7 in ddH$_2$O
RNase solution (20mg/ml)	RNase, 1M Tris-HCl	200mg RNase dissolved in 100µl Tris-HCl, 30µl NaCl,

	5 M NaCl ddH$_2$O	ad 10ml ddH$_2$O; heat for 15 min. at 90°C
RPMI 1640 medium (10% FCS)	RPMI 1640 Medium 10% FCS Penicillin/Streptomycin	500ml RPMI 1640 50ml FCS 5ml Penicillin/Streptomycin
20x SSC (pH 7,0)	150mM NaCl 15mM Na-Citrat	175.3g NaCl 88.2g Na-Citrat ad 1l H2O bidest.; adjust pH to 7.0 by adding NaOH
4x/2x/0.1x SSC	20x SSC H2O bidest.	20x SSC diluted 1:5/1:10/1:200 with H$_2$O bidest.
4x SSCT	Tween 20 4x SSC	2ml Tween 20 in 1l 4x SSC
Nicktranslation stopping mix	0.1% bromophenol blue 0.5% dextran blue 0.1 M NaCl 20 mM EDTA 20 mM Tris-HCl pH7.5	40 mg bromophenol blue 200 mg dextran blue 800 µl 5 M NaCl 1600 µl 0.5 M EDTA 800 µl 1 M Tris-HCl pH7.5 ad 40 ml H2O bidest.
50x TAE-Puffer (pH 8.0)	40 mM Tris acetic acid 1 mM EDTA	242g Tris 18,6 g EDTA 57,1 ml acetic acid ad 1l H2o bidest.
0.5% Triton solution	Triton X-100 1x PBS	5ml Triton X-100 ad 1l 1xPBS
Trypsin / EDTA solution	0.5% trypsin CMF-PBS	0.5% trypsin diluted 1:10 in CMF-PBS; warming-up to 37°C before use

6.10 Chemicals and reagents

Chemical / reagent	Company, Distributor
Acetic acid	Sigma-Aldrich, Deisenhofen, Germany
Agarose Seakem ME	Cambrex, Rockland, Maine,USA
Biotin succinimidyl ester (bio)	Molecular Probes (Invitrogen), Karlsruhe, Germany
Bovine serumalbumin (BSA) dissolved in SSC	ICN Biomedicals, Eschwege, Germany
Bovine serumalbumin (BSA) dissolved in PBS	Sigma-Aldrich
Bromophenol blue	Sigma-Aldrich
Calcium chloride	Merck, Darmstadt, Germany
Cetus II buffer	Applied Biosystems
Chloroform	Roth, Karlsruhe, Germany
Colcemide	Biochrom, Berlin; Germany
Cy3 mono NHS ester (Cy3)	Amersham Biosciences, Freiburg, Germany
Cy5 mono NHS ester (Cy5)	Amersham Biosciences, Freiburg, Germany

DAPI (4,6 diamino-2-phenylindole)	Sigma-Aldrich
Dextran blue	Sigma-Aldrich
Dextran sulfate	Amersham Biosciences, Freiburg, Germany
Digoxigenin succinimidyl ester (dig)	Molecular Probes
Dinitrophenyl aminohexanoid acid succinimidyl ester (dnp)	Molecular Probes
DMEM medium	Biochrom, Berlin, Germany
DMSO (Dimethyl sulphoxide)	Sigma-Aldrich
EDTA	Merck
Ethanol absolute p.a.	Merck
Ethidium bromide	Sigma-Aldrich
Fetal calf serum (FCS)	Biochrom
FITC succinimidyl ester (FITC)	Molecular Probes
Formamide	Merck
Gel loading buffer	Sigma-Aldrich
Glycerol	Merck
Glycine	Amersham Biosciences
Iscove medium (incl stable glutamine)	Biochrom
Isopropanol	Merck
Potassium chloride (KCl)	Merck
Potassium phosphate (KH_2PO_4)	Calbiochem, Darmstadt, Germany
Magnesium chloride	Merck
$MgCl_2$ solution (25 mM) PCR	Perkin Elmer, Wellesley, MA, USA
β-mercaptoethanol	Merck
m-IL3 (murine Interleukin-3)	Biochrom
m-SCF (murine stem cell factor)	Biochrom
NaHCO3	Sigma-Aldrich
Nick translation buffer (10×)	Roche, Mannheim, Germany
Sodium acetate	Merck
Sodium azide	Merck
Sodium chloride	Merck
Sodium citrate	Calbiochem
Sodium hydrogen phosphate (Na_2HPO_4)	Merck
Paraformaldehyde	Merck
PCR buffer D (5×)	Invitrogen
PCR buffer I (10×) without MgCl2	Perkin Elmer, Jugesheim, Germany
Penicillin/Streptomycin	Biochrom
Phenol	Roth
Poly-L-lysine	Sigma-Aldrich
Buffer D (5x)	Invitrogen
RPMI 1640 Medium	Biochrom
Hydrochloric acid	Merck
Nitrogen (liquid)	Provided by LMU
TAMRA succinimidyl ester (TAMRA)	Molecular Probes
Texas Red succinimidyl ester (Texas Red)	Molecular Probes
Tris-HCl	Sigma-Aldrich
Triton X-100	Sigma-Aldrich
Tween 20	Calbiochem

Vectashield	Vector Laboratories Inc., Burlingame, CA, USA
VRC (Ribonucleoside-Vanadyl complex)	New England Biolabs, Frankfurt, Germany
W1 (Polyoxyethylene ether W1)	Sigma-Aldrich

6.11 Enzymes and nucleotides

Enzyme	Company, Distributor
DNase I	Roche, Mannheim, Germany
Pepsin	Sigma-Aldrich, Deisenhofen, Germany
DNA polymerase I (Kornberg Pol)	Roche
Rnase A	Roche
Taq polymerase	GE Healthcare, München, Germany
All restriction enzymes, according buffers and BSA used in 3.4.2	New England Biolabs, Frankfurt, Germany

Nucleotide(s)	Company, Distributor
6MW-Primer	MWG-Biotech, Ebersberg, Germany
Aminoallyl-dUTP	Sigma-Aldrich, Deisenhofen, Germany
Biotin-16-dUTP	home labeled
Cot-1 DNA	Invitrogen, Karlsruhe, Germany
Cy3-dUTP	home labeled
Digoxigenin-11-dUTP	home labeled
DNP-11-dUTP	home labeled
dNTPs (each 2,5 mM)	Invitrogen
dTTP (1mM)	Invitrogen
Lambda/Hind III marker	Fermentas, St. Leon-Rot, Germany
1kb ladder	Fermentas
Salmonsperm-DNA	Invitrogen
humane Cot-1 DNA	Invitrogen
Tamra-6-dUTP	home labeled
TexasRed-12-dUTP	home labeled
All primers, Linker+ and Linker- used in 3.4.2	Metabion, Martinsried, Germany

→ Further information about solutions, chemicals and protocols see (Cremer et al., 2008).

6.12 Kits

Kit	Company, Distributor
Advantage 2 PCR Kit	Clontech, CA, USA
DNeasy Blood & Tissue Kit	Qiagen, Hilden, Germany
Dynabeads kilobase BINDER Kit	Invitrogen, Karlsruhe, Germany

First Strand Synthesis System for RT-PCR	Invitrogen
GenomiPhiTM V2 DNA Amplification Kit	GE Healthcare, München, Germany
HIV-1 p24-Antigen Capture ELISA	Advanced BioScience Lab Inc., CA, USA
Light Cycler 480 SYBR Green I Master	Roche, Mannheim, Germany
NucleoBond Xtra Midi/Maxi Kit	Macherey-Nagel, Düren, Germany
peqGOLD Plasmid Miniprep Kit II	Peqlab, Erlangen, Germany
Qiaquick Gel Extraction Kit	Qiagen
Qiaquick PCR Purification Kit	Qiagen
Rapid DNA Ligation Kit	Roche, Mannheim, Germany
RNeasy RNA Mini Kit	Qiagen
Universal Genome Walker Kit	Clontech, CA, USA

6.13 Consumables, technical equipment and microscopes

Consumables	Company, Distributor
Cryotubes (1.5 ml)	Greiner Bio One, Frickenhausen
Coverslip (15x15/76x26mm)	Superior
Glass pipette (1/5/10/25ml)	Costar, Fernwald
Tubes (1.5/2ml)	Schubert & Weiss, München
Cell culture flasks (25/75cm2)	Greiner Bio One, Frickenhausen
Tubes (15/50ml)	Greiner Bio One, Frickenhausen
Fixogum	Marabuwerke, Tamm
Glass flasks (50/100/250/500/1000ml)	Schott
Cuvettes (50/100ml)	Schubert & Weiss, München
Microscopic slides (76x26mm)	R.Langenbrink
PCR tubes (0.5ml)	Molecular Bio Products TM, San Diego, USA
Petri dishes (35x10mm)	Greiner Bio One, Frickenhausen
Pipette tips (white)	Molekular Bio Produkts
Pipette tips (yellow, blue)	Greiner Labortechnik
Pipette aid (Accu-jet® BRAND)	Schubert & Weiss, München
Quadri PERM® plus (4x 76x26mm)	Heraeus, Hanau
Sterile filter (pore size 20/40 µm)	Satorius, Göttingen
6 well plates (6x 35x10mm)	Greiner Bio One, Frickenhausen

Technical equipment	Company, Distributor
Incubator (B 5060)	Haraeus, Hanau
Autoclav	H+P Labortechnik GmbH
Gel electrophoresis apparatus	Owl, Portsmouth, USA
Centrifuge (1,5 ml, Biofuge pico)	Heraeus, Hanau
Centrifuge (15/50ml, Rotanta/S)	Hettich, Tuttlingen
Hot block (Dri Block DB-2D)	Techne Progene, Cambridge, UK
Laminar flow (Airflow)	Biohit
Pipette aid (Accu-jet® BRAND)	Schubert & Weiss, München
Thermal cycler (TC312, TC3000)	Techne Progene, Cambridge, UK
Thermal cycler (Thermomixer 5436)	Eppendorf
Transilluminator (Gel Imager)	MWG Biotech, Ebersberg

Vacuum centrifuge (Vacuum Concentrator)	Bachhofer, Reutlingen
pH-meter	WTW, Weilheim
Vortex (VF2)	Janke & Kunkel
Liquid nitrogen tank (Apollo)	Cryotherm
Icemachine (Scotsman AF-10)	Masare, München
Freezer -20°C	Bosch
Freezer -80°C	Sanyo
Water baths	Jukabo (small), GFL (large)
Parafilm	Pechiney, Menasha, USA
Paper wipes	Kimberly-Clark / Roth
forceps	Dumont, Montiguez
Pipette (10/20/200/1000µl)	Gilson

Phase contrast microscope: Axiovert 25C (Zeiss, Jena)
Objectives: CP Achromat 5x / 0.12 CP Achromat 10x / 0.25 Ph1 LD Achrostigmat 20x /0.3 Ph1 LD Achrostigmat 40x / 0.55 Ph2

Epifluorescence microscope: Axiophot 2 (Zeiss, Jena)
Objectives: Plan Neofluar 16x / 0.5 Plan Neofluar 40x / 1.3 oil Plan Neofluar 63x / 1.4 oil Plan Neofluar 100x / 1.3 oil Filter: DAPI (BP 365, FT 395, LP 450-490 nm) FITC (BP 450-490, FT 510, LP 515-565 nm) Cy3 (BP 546, FT 580, LP 590 nm) Cy5 (BP 575-625, FT 645, LP 660-710 nm) Camera: Coolview CCD (Applied Imaging) Software: Metavue

Confocal laser scanning microscope: Leica TCS-SP5 (Leica , Heidelberg)
Objectives: Objektiv: Plan-Achromat 63x oil / 1.4 Lasers: Ar 100 mW 485, 476, 488, 496, 512 nm HeNe 2 mW 594 nm HeNe 10 mW 633 nm DPSS 10 mW 561 nm

5 photomultiplier with continuously adjustable bandwidth Software: Leica Application Suite AF Version: 2.0.0 build 1934	

6.14 Software

Software	Company, Distributor
Amira	TGS Europe, Düsseldorf, Germany
Endnote X3	Thomson Reuters, NY, USA
Enhanced Distance to Surface (eADS) v3.00a	T. Thormeyer (Thormeyer, 2005; Albiez et al., 2006)
Gene Construction Kit 2.5.13	Textco Inc., NH, USA
ImageJ 1.37c	Wayne Rasband, NIH, USA
Leica confocal software	Leica Microsystems, Wetzlar, Germany
Metavue 6.2r2	Universial Imaging Corp.
Microsoft Office 2003	Microsoft Corporation, WA, USA
Photoshop CS2	Adobe Corporation, CA, USA
Sigma Stat 3.5	SPSS Inc. (IBM), IL, USA

6.15 Homepages / Databases

Ncbi: http://www.ncbi.nlm.nih.gov/

Ensemble: http://www.ensembl.org/index.html

Repeatmasker: http://www.repeatmasker.org/

BAC Pac Resources: http://bacpac.chori.org/

6.16 List of publications

Nagel J., Groß B., Meggendorfer M., Preiss C., Grez M., Brack-Werner R., Dietzel S.
"Stably integrated and expressed retroviral sequences can influence nuclear location and chromatin condensation of the integraton locus "
Chromosoma, 2012 Mar; DOI: 10.1007/s00412-012-0366-9

Rienzo M., **Nagel J.**, Casamassimi A., Giovane A., Dietzel S., Napoli C.
"Mediator subunits: gene expression pattern, a novel transcript identification and nuclear localization in human endothelial progenitor cells"
Biochim Biophys Acta. 2010 Jul;1799(7):487-95

Kim IH., **Nagel J.**, Otten S., Knerr B., Eils R., Rohr K., Dietzel S.
"Quantitative comparison of DNA detection by GFP-lac repressor tagging, fluorescence in situ hybridization and immunostaining"
BMC Biotechnol. 2007 Dec 20;7:92

6.17 Poster presentations and talks

07/06: Düsseldorf, Germany (*short talk*): Kick-off Meeting of the DFG Schwerpunkt-Programm (SPP)1230 („Mechanisms of gene vector entry and persistence") in line with the 13th Annual Meeting of the DG-GT (German Society of Gene Therapy)

09/06: Reimlingen, Germany (*poster*): 1st Chromatin Summer School, SFB/Transregio 5

03/07: München, Germany (*talk*): SPP1230 workshop

11/07: Frankfurt, Germany (*talk*): SPP1230 Annual Meeting

03/08: München, Germany (*talk*): SPP1230 workshop

05/08: Madrid, Spain (*poster*): 2nd MC-GARD (Marie Curie: genome architecture in relation to disease) conference: Interplay among genetics, epigenetics and non-coding RNA's.

10/08: Berlin, Germany (*talk*): SPP1230 Annual Meeting

04/09: Edinburgh, Scotland (*poster*): 3rd MC-GARD conference: Higher order genome architecture

10/10: München, Germany (*poster*): 3rd Transregio 5 Symposium, Chromatin: Assembly and Inheritance of Functional States

6.18 Curriculum vitae

Name:	Jens Nagel
Geburtsdatum:	19.12.1978
Geburtsort:	Heidelberg
Familienstand:	ledig
Nationalität:	deutsch

Schulbildung

1998	Allgemeine Hochschulreife, St.Raphael-Gymnasium Heidelberg

Zivildienst

08/1998 – 09/1999	St.Elisabeth Klinik, Heidelberg

Universitätslaufbahn

10/1999 – 09/2000	Studium der Rechtswissenschaften, Ruprecht-Karls-Universität, Heidelberg
10/2000 – 10/2005	Studium der Biologie, Ruprecht-Karls-Universität, Heidelberg
10/2005 – 06/2006	Diplomarbeit am Institut für Anthropologie und Humangenetik, Ludwig-Maximilians-Universität, München
20.06.2006	Diplom
seit 07/2006	Doktorarbeit am Institut für Anthropologie und Humangenetik, Ludwig-Maximilians-Universität, München

7. Bibliography

Abney, J.R., B. Cutler, M.L. Fillbach, D. Axelrod, and B.A. Scalettar. 1997. Chromatin dynamics in interphase nuclei and its implications for nuclear structure. *J. Cell Biol.* 137:1459-68.

Adachi, A., H.E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M.A. Martin. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol*. 59:284-91.

Adkins, N.L., M. Watts, and P.T. Georgel. 2004. To the 30-nm chromatin fiber and beyond. *Biochim Biophys Acta*. 1677:12-23.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. 2011. Molekularbiologie der Zelle. Wiley-VCH Verlag GmbH, 5. Auflage.

Albiez, H. 2007. Manipulation of global chromatin architecture in the human cell nucleus and critical assessment of current model views. *In* Biozentrum. LMU, München.

Albiez, H., M. Cremer, C. Tiberi, L. Vecchio, L. Schermelleh, S. Dittrich, K. Kupper, B. Joffe, T. Thormeyer, J. von Hase, S. Yang, K. Rohr, H. Leonhardt, I. Solovei, C. Cremer, S. Fakan, and T. Cremer. 2006. Chromatin domains and the interchromatin compartment form structurally defined and functionally interacting nuclear networks. *Chromosome Res*. 14:707-733.

Ashe, M.P., P. Griffin, W. James, and N.J. Proudfoot. 1995. Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev*. 9:3008-25.

Bannwarth, S., and A. Gatignol. 2005. HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. *Curr HIV Res*. 3:61-71.

Barr, M.L., and E.G. Bertram. 1949. A morphological distinction between neurons of the male and female, and the behavior of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*. 163:676-677.

Batsche, E., M. Yaniv, and C. Muchardt. 2006. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol*. 13:22-9.

Baum, C., J. Dullmann, Z. Li, B. Fehse, J. Meyer, D.A. Williams, and C. von Kalle. 2003. Side effects of retroviral gene transfer into hematopoietic stem cells. *Blood*. 101:2099-114.

Baum, C., O. Kustikova, U. Modlich, Z. Li, and B. Fehse. 2006. Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. *Hum Gene Ther*. 17:253-63.

Becavin, C., M. Barbi, J.M. Victor, and A. Lesne. 2010. Transcription within condensed chromatin: Steric hindrance facilitates elongation. *Biophys J*. 98:824-33.

Bell, P., L.J. Montaner, and G.G. Maul. 2001. Accumulation and intranuclear distribution of unintegrated human immunodeficiency virus type 1 DNA. *J Virol*. 75:7683-91.

Belmont, A.S. 2006. Mitotic chromosome structure and condensation. *Curr Opin Cell Biol*. 18:632-8.

Berthold, E., and F. Maldarelli. 1996. cis-acting elements in human immunodeficiency virus type 1 RNAs direct viral transcripts to distinct intranuclear locations. *J Virol*. 70:4667-82.

Birnboim, H.C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol*. 100:243-55.

Bisgrove, D., M. Lewinski, F. Bushman, and E. Verdin. 2005. Molecular mechanisms of HIV-1 proviral latency. *Expert Rev Anti Infect Ther*. 3:805-14.

Blazkova, J., K. Trejbalova, F. Gondois-Rey, P. Halfon, P. Philibert, A. Guiguen, E. Verdin, D. Olive, C. Van Lint, J. Hejnar, and I. Hirsch. 2009. CpG methylation controls reactivation of HIV from latency. *PLoS Pathog*. 5:e1000554.

Boe, S.O., B. Bjorndal, B. Rosok, A.M. Szilvay, and K.H. Kalland. 1998. Subcellular localization of human immunodeficiency virus type 1 RNAs, Rev, and the splicing factor SC-35. *Virology*. 244:473-82.

Bokhoven, M., S.L. Stephen, S. Knight, E.F. Gevers, I.C. Robinson, Y. Takeuchi, and M.K. Collins. 2009. Insertional gene activation by lentiviral and gammaretroviral vectors. *J Virol*. 83:283-94.

Bolte, S., and F.P. Cordelieres. 2006. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc*. 224:213-32.

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

Bornfleth, H., P. Edelmann, D. Zink, T. Cremer, and C. Cremer. 1999. Quantitative motion analysis of subchromosomal foci in living cells using four-dimensional microscopy. *Biophys J*. 77:2871-86.

Bosticardo, M., A. Ghosh, Y. Du, N.A. Jenkins, N.G. Copeland, and F. Candotti. 2009. Self-inactivating retroviral vector-mediated gene transfer induces oncogene activation and immortalization of primary murine bone marrow cells. *Mol Ther*. 17:1910-8.

Boyle, S., S. Gilchrist, J.M. Bridger, N.L. Mahy, J.A. Ellis, and W.A. Bickmore. 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet*. 10:211-219.

Brack-Werner, R., A. Kleinschmidt, A. Ludvigsen, W. Mellert, M. Neumann, R. Herrmann, M.C. Khim, A. Burny, N. Muller-Lantzsch, D. Stavrou, and et al. 1992. Infection of human brain cells by HIV-1: restricted virus production in chronically infected human glial cell lines. *AIDS*. 6:273-85.

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

Brady, T., L.M. Agosto, N. Malani, C.C. Berry, U. O'Doherty, and F. Bushman. 2009. HIV integration site distributions in resting and activated CD4+ T cells infected in culture. *AIDS*. 23:1461-71.

Brown, J.M., J. Green, R.P. das Neves, H.A. Wallace, A.J. Smith, J. Hughes, N. Gray, S. Taylor, W.G. Wood, D.R. Higgs, F.J. Iborra, and V.J. Buckle. 2008. Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J Cell Biol*. 182:1083-97.

Brown, J.M., J. Leach, J.E. Reittie, A. Atzberger, J. Lee-Prudhoe, W.G. Wood, D.R. Higgs, F.J. Iborra, and V.J. Buckle. 2006. Coregulated human globin genes are frequently in spatial proximity when active. *J Cell Biol*. 172:177-87.

Buonamici, S., D. Li, Y. Chi, R. Zhao, X. Wang, L. Brace, H. Ni, Y. Saunthararajah, and G. Nucifora. 2004. EVI1 induces myelodysplastic syndrome in mice. *J Clin Invest*. 114:713-9.

Bushman, F., M. Lewinski, A. Ciuffi, S. Barr, J. Leipzig, S. Hannenhalli, and C. Hoffmann. 2005. Genome-wide analysis of retroviral DNA integration. *Nat Rev Microbiol*. 3:848-58.

Campbell, E.M., and T.J. Hope. 2008. Live cell imaging of the HIV-1 life cycle. *Trends Microbiol*. 16:580-7.

Campos, E.I., and D. Reinberg. 2009. Histones: annotating chromatin. *Annu Rev Genet*. 43:559-99.

Chambeyron, S., and W.A. Bickmore. 2004. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev*. 18:1119-30.

Chen, J., Q. Dang, D. Unutmaz, V.K. Pathak, F. Maldarelli, D. Powell, and W.S. Hu. 2005. Mechanisms of nonrandom human immunodeficiency virus type 1 infection and double infection: preference in virus entry is important but is not the sole factor. *J Virol*. 79:4140-9.

Chow, J., and E. Heard. 2009. X inactivation and the complexities of silencing a sex chromosome. *Curr Opin Cell Biol*. 21:359-66.

Chuang, C.H., A.E. Carpenter, B. Fuchsova, T. Johnson, P. de Lanerolle, and A.S. Belmont. 2006. Long-range directional movement of an interphase chromosome site. *Curr Biol*. 16:825-31.

Ciuffi, A., and F.D. Bushman. 2006. Retroviral DNA integration: HIV and the role of LEDGF/p75. *Trends Genet*. 22:388-95.

Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J.R. Ecker, and F. Bushman. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med*. 11:1287-9.

Coiras, M., M.R. Lopez-Huertas, M. Perez-Olmeda, and J. Alcami. 2009. Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat Rev Microbiol*. 7:798-812.

Corbeau, P. 2008. Interfering RNA and HIV: reciprocal interferences. *PLoS Pathog*. 4:e1000162.

Corbeil, J., D. Sheeter, D. Genini, S. Rought, L. Leoni, P. Du, M. Ferguson, D.R. Masys, J.B. Welsh, J.L. Fink, R. Sasik, D. Huang, J. Drenkow, D.D. Richman, and T. Gingeras. 2001. Temporal gene regulation during HIV-1 infection of human CD4+ T cells. *Genome Res*. 11:1198-204.

Cotrim, A.P., and B.J. Baum. 2008. Gene therapy: some history, applications, problems, and prospects. *Toxicol Pathol*. 36:97-103.

Coull, J.J., G. He, C. Melander, V.C. Rucker, P.B. Dervan, and D.M. Margolis. 2002. Targeted derepression of the human immunodeficiency virus type 1 long terminal repeat by pyrrole-imidazole polyamides. *J Virol*. 76:12349-54.

Coull, J.J., F. Romerio, J.M. Sun, J.L. Volker, K.M. Galvin, J.R. Davie, Y. Shi, U. Hansen, and D.M. Margolis. 2000. The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. *J Virol*. 74:6790-9.

Cremer, M., F. Grasser, C. Lanctot, S. Muller, M. Neusser, R. Zinner, I. Solovei, and T. Cremer. 2008. Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. *Methods Mol Biol*. 463:205-39.

Cremer, M., J. von Hase, T. Volm, A. Brero, G. Kreth, J. Walter, C. Fischer, I. Solovei, C. Cremer, and T. Cremer. 2001. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res*. 9:541-67.

Cremer, T., and C. Cremer. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet*. 2:292-301.

Cremer, T., and M. Cremer. 2010. Chromosome territories. *Cold Spring Harb Perspect Biol*. 2:a003889.

Cremer, T., M. Cremer, S. Dietzel, S. Muller, I. Solovei, and S. Fakan. 2006. Chromosome territories--a functional nuclear landscape. *Curr Opin Cell Biol*. 18:307-16.

Dahl, V., L. Josefsson, and S. Palmer. 2010. HIV reservoirs, latency, and reactivation: prospects for eradication. *Antiviral Res*. 85:286-94.

Dang, Q., J. Chen, D. Unutmaz, J.M. Coffin, V.K. Pathak, D. Powell, V.N. KewalRamani, F. Maldarelli, and W.S. Hu. 2004. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. *Proc Natl Acad Sci U S A*. 101:632-7.

Daniel, R., J. Ramcharan, E. Rogakou, K.D. Taganov, J.G. Greger, W. Bonner, A. Nussenzweig, R.A. Katz, and A.M. Skalka. 2004. Histone H2AX is phosphorylated at sites of retroviral DNA integration but is dispensable for postintegration repair. *J Biol Chem*. 279:45810-4.

Daniel, R., and J.A. Smith. 2008. Integration site selection by retroviral vectors: molecular mechanism and clinical consequences. *Hum Gene Ther*. 19:557-68.

de Laat, W. 2007. Long-range DNA contacts: romance in the nucleus? *Curr Opin Cell Biol*. 19:317-20.

Dean, F.B., J.R. Nelson, T.L. Giesler, and R.S. Lasken. 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res*. 11:1095-9.

Deichmann, A., S. Hacein-Bey-Abina, M. Schmidt, A. Garrigue, M.H. Brugman, J. Hu, H. Glimm, G. Gyapay, B. Prum, C.C. Fraser, N. Fischer, K. Schwarzwaelder, M.L. Siegler, D. de Ridder, K. Pike-Overzet, S.J. Howe, A.J. Thrasher, G. Wagemaker, U. Abel, F.J. Staal, E. Delabesse, J.L. Villeval, B. Aronow, C. Hue, C. Prinz, M. Wissler, C. Klanke, J. Weissenbach, I. Alexander, A. Fischer, C. von Kalle, and M. Cavazzana-Calvo. 2007. Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. *J Clin Invest*. 117:2225-32.

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

Delelis, O., A. Zamborlini, S. Thierry, and A. Saib. 2010. Chromosomal tethering and proviral integration. *Biochim Biophys Acta*. 1799:207-16.

Dietzel, S., A. Jauch, D. Kienle, G. Qu, H. Holtgreve-Grez, R. Eils, C. Munkel, M. Bittner, P.S. Meltzer, J.M. Trent, and T. Cremer. 1998. Separate and variably shaped chromosome arm domains are disclosed by chromosome arm painting in human cell nuclei. *Chromosome Res.* 6:25-33.

Dietzel, S., K. Schiebel, G. Little, P. Edelmann, G.A. Rappold, R. Eils, C. Cremer, and T. Cremer. 1999. The 3D positioning of ANT2 and ANT3 genes within female X chromosome territories correlates with gene activity. *Exp. Cell Res.* 252:363-75.

Dietzel, S., K. Zolghadr, C. Hepperger, and A.S. Belmont. 2004. Differential large-scale chromatin compaction and intranuclear positioning of transcribed versus non-transcribed transgene arrays containing beta-globin regulatory sequences. *J Cell Sci*. 117:4603-14.

Dieudonne, M., P. Maiuri, C. Biancotto, A. Knezevich, A. Kula, M. Lusic, and A. Marcello. 2009. Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery. *Embo J*. 28:2231-43.

Du, Y., N.A. Jenkins, and N.G. Copeland. 2005. Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. *Blood*. 106:3932-9.

Easley, R., R. Van Duyne, W. Coley, I. Guendel, S. Dadgar, K. Kehn-Hall, and F. Kashanchi. 2010. Chromatin dynamics associated with HIV-1 Tat-activated transcription. *Biochim Biophys Acta*. 1799:275-85.

Edelstein, M.L., M.R. Abedi, and J. Wixon. 2007. Gene therapy clinical trials worldwide to 2007--an update. *J Gene Med*. 9:833-42.

Eils, R., S. Dietzel, E. Bertin, E. Schrock, M.R. Speicher, T. Ried, M. Robert-Nicoud, C. Cremer, and T. Cremer. 1996. Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome territories have similar volumes but differ in shape and surface structure. *J. Cell Biol.* 135:1427-40.

Engelman, A., and P. Cherepanov. 2008. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog*. 4:e1000046.

Favaro, J.P., K.T. Borg, S.J. Arrigo, and M.G. Schmidt. 1998. Effect of Rev on the intranuclear localization of HIV-1 unspliced RNA. *Virology*. 249:286-96.

Felsenfeld, G., and M. Groudine. 2003. Controlling the double helix. *Nature*. 421:448-53.

Frankel, A.D., and J.A. Young. 1998. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem*. 67:1-25.

Fraser, P., and W. Bickmore. 2007. Nuclear organization of the genome and the potential for gene regulation. *Nature*. 447:413-7.

Geijtenbeek, T.B., D.S. Kwon, R. Torensma, S.J. van Vliet, G.C. van Duijnhoven, J. Middel, I.L. Cornelissen, H.S. Nottet, V.N. KewalRamani, D.R. Littman, C.G. Figdor, and Y. van Kooyk. 2000. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*. 100:587-97.

GenTG. 2010. Gesetz zur Regelung der Gentechnik. B. Deutschland, editor.

Gerlich, D., J. Beaudouin, B. Kalbfuss, N. Daigle, R. Eils, and J. Ellenberg. 2003. Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell*. 112:751-64.

Gierman, H.J., M.H. Indemans, J. Koster, S. Goetze, J. Seppen, D. Geerts, R. van Driel, and R. Versteeg. 2007. Domain-wide regulation of gene expression in the human genome. *Genome Res*. 17:1286-95.

Gondor, A., and R. Ohlsson. 2009. Chromosome crosstalk in three dimensions. *Nature*. 461:212-7.

Graveley, B.R. 2000. Sorting out the complexity of SR protein functions. *RNA*. 6:1197-211.

Grewal, S.I., and S. Jia. 2007. Heterochromatin revisited. *Nat Rev Genet*. 8:35-46.

Groß, B. 2009. Positionierung von Integrationsstellen eines retroviralen Genvektors in Zellkernen einer Blutstammzelllinie der Maus. *In* Biozentrum. LMU, München.

Grouard, G., and E.A. Clark. 1997. Role of dendritic and follicular dendritic cells in HIV infection and pathogenesis. *Curr Opin Immunol*. 9:563-7.

Habermann, F.A., M. Cremer, J. Walter, G. Kreth, J. von Hase, K. Bauer, J. Wienberg, C. Cremer, T. Cremer, and I. Solovei. 2001. Arrangements of macro- and microchromosomes in chicken cells. *Chromosome Res*. 9:569-84.

Hacein-Bey-Abina, S., A. Garrigue, G.P. Wang, J. Soulier, A. Lim, E. Morillon, E. Clappier, L. Caccavelli, E. Delabesse, K. Beldjord, V. Asnafi, E. MacIntyre, L. Dal Cortivo, I. Radford, N. Brousse, F. Sigaux, D. Moshous, J. Hauer, A. Borkhardt, B.H. Belohradsky, U. Wintergerst, M.C. Velez, L. Leiva, R. Sorensen, N. Wulffraat, S. Blanche, F.D. Bushman, A. Fischer, and M. Cavazzana-Calvo. 2008. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest*. 118:3132-42.

Harouse, J.M., C. Kunsch, H.T. Hartle, M.A. Laughlin, J.A. Hoxie, B. Wigdahl, and F. Gonzalez-Scarano. 1989. CD4-independent infection of human neural cells by human immunodeficiency virus type 1. *J Virol*. 63:2527-33.

BIBLIOGRAPHY

He, G., and D.M. Margolis. 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*. 22:2965-73.

Hepperger, C., S. Otten, J. von Hase, and S. Dietzel. 2007. Preservation of large-scale chromatin structure in FISH experiments. *Chromosoma*. 116:117-33.

Herschhorn, A., and A. Hizi. 2010. Retroviral reverse transcriptases. *Cell Mol Life Sci*. 67:2717-47.

Hewitt, S.L., F.A. High, S.L. Reiner, A.G. Fisher, and M. Merkenschlager. 2004. Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *Eur J Immunol*. 34:3604-13.

Holman, A.G., and J.M. Coffin. 2005. Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. *Proc Natl Acad Sci U S A*. 102:6103-7.

Holmes-Son, M.L., R.S. Appa, and S.A. Chow. 2001. Molecular genetics and target site specificity of retroviral integration. *Adv Genet*. 43:33-69.

Howe, S.J., M.R. Mansour, K. Schwarzwaelder, C. Bartholomae, M. Hubank, H. Kempski, M.H. Brugman, K. Pike-Overzet, S.J. Chatters, D. de Ridder, K.C. Gilmour, S. Adams, S.I. Thornhill, K.L. Parsley, F.J. Staal, R.E. Gale, D.C. Linch, J. Bayford, L. Brown, M. Quaye, C. Kinnon, P. Ancliff, D.K. Webb, M. Schmidt, C. von Kalle, H.B. Gaspar, and A.J. Thrasher. 2008. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest*. 118:3143-50.

Hu, Y., I. Kireev, M. Plutz, N. Ashourian, and A.S. Belmont. 2009. Large-scale chromatin structure of inducible genes: transcription on a condensed, linear template. *J Cell Biol*. 185:87-100.

Jacob, D.T., and J.J. DeStefano. 2008. A new role for HIV nucleocapsid protein in modulating the specificity of plus strand priming. *Virology*. 378:385-96.

Jordan, A., D. Bisgrove, and E. Verdin. 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *Embo J*. 22:1868-77.

Jordan, A., P. Defechereux, and E. Verdin. 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *Embo J*. 20:1726-38.

Jung, A., R. Maier, J.P. Vartanian, G. Bocharov, V. Jung, U. Fischer, E. Meese, S. Wain-Hobson, and A. Meyerhans. 2002. Recombination: Multiply infected spleen cells in HIV patients. *Nature*. 418:144.

Kao, S.Y., A.F. Calman, P.A. Luciw, and B.M. Peterlin. 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature*. 330:489-93.

Kay, M.A., J.C. Glorioso, and L. Naldini. 2001. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat Med*. 7:33-40.

Kind, J., and B. van Steensel. 2010. Genome-nuclear lamina interactions and gene regulation. *Curr Opin Cell Biol*. 22:320-5.

Klimas, N., A.O. Koneru, and M.A. Fletcher. 2008. Overview of HIV. *Psychosom Med*. 70:523-30.

Köhler, D., V. Zakhartchenko, L. Froenicke, G. Stone, R. Stanyon, E. Wolf, T. Cremer, and A. Brero. 2009. Changes of higher order chromatin arrangements during major genome activation in bovine preimplantation embryos. *Exp Cell Res*. 315:2053-63.

136

Kohn, D.B., M. Sadelain, C. Dunbar, D. Bodine, H.P. Kiem, F. Candotti, J. Tisdale, I. Riviere, C.A. Blau, R.E. Richard, B. Sorrentino, J. Nolta, H. Malech, M. Brenner, K. Cornetta, J. Cavagnaro, K. High, and J. Glorioso. 2003. American Society of Gene Therapy (ASGT) ad hoc subcommittee on retroviral-mediated gene transfer to hematopoietic stem cells. *Mol Ther*. 8:180-7.

Komeno, Y., J. Kitaura, and T. Kitamura. 2009. Molecular bases of myelodysplastic syndromes: lessons from animal models. *J Cell Physiol*. 219:529-34.

Kumar, A., and K.T. Jeang. 2008. Insights into cellular microRNAs and human immunodeficiency virus type 1 (HIV-1). *J Cell Physiol*. 216:327-31.

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

Kumaran, R.I., R. Thakar, and D.L. Spector. 2008. Chromatin dynamics and gene positioning. *Cell*. 132:929-34.

Kupper, K. 2007. Untersuchungen zur Zellkernarchitektur: Der einfluss verschiedener Chromatineigenschaften auf die räumliche Anordnung subchromosomaler Segmente bezüglich der Zellkerne sowie der Chromosomenterritorien anhand der humanen Chromosomen 11, 12, 18 und19. *In* Biozentrum. LMU, München.

Kupper, K., A. Kolbl, D. Biener, S. Dittrich, J. von Hase, T. Thormeyer, H. Fiegler, N.P. Carter, M.R. Speicher, T. Cremer, and M. Cremer. 2007. Radial chromatin positioning is shaped by local gene density, not by gene expression. *Chromosoma*. 116:285-306.

Kustikova, O., B. Fehse, U. Modlich, M. Yang, J. Dullmann, K. Kamino, N. von Neuhoff, B. Schlegelberger, Z. Li, and C. Baum. 2005. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science*. 308:1171-4.

Kustikova, O.S., C. Baum, and B. Fehse. 2008. Retroviral integration site analysis in hematopoietic stem cells. *Methods Mol Biol*. 430:255-67.

Kustikova, O.S., B. Schiedlmeier, M.H. Brugman, M. Stahlhut, S. Bartels, Z. Li, and C. Baum. 2009. Cell-intrinsic and Vector-related Properties Cooperate to Determine the Incidence and Consequences of Insertional Mutagenesis. *Mol Ther*.

Lanctot, C., T. Cheutin, M. Cremer, G. Cavalli, and T. Cremer. 2007. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet*. 8:104-15.

Laricchia-Robbio, L., and G. Nucifora. 2008. Significant increase of self-renewal in hematopoietic cells after forced expression of EVI1. *Blood Cells Mol Dis*. 40:141-7.

Lau, A., K.M. Swinbank, P.S. Ahmed, D.L. Taylor, S.P. Jackson, G.C. Smith, and M.J. O'Connor. 2005. Suppression of HIV-1 infection by a small molecule inhibitor of the ATM kinase. *Nat Cell Biol*. 7:493-500.

Li, B., M. Carey, and J.L. Workman. 2007. The role of chromatin during transcription. *Cell*. 128:707-19.

Li, Z., J. Dullmann, B. Schiedlmeier, M. Schmidt, C. von Kalle, J. Meyer, M. Forster, C. Stocking, A. Wahlers, O. Frank, W. Ostertag, K. Kuhlcke, H.G. Eckert, B. Fehse, and C. Baum. 2002. Murine leukemia induced by retroviral gene marking. *Science*. 296:497.

Lilley, C.E., M.S. Chaurushiya, and M.D. Weitzman. 2009. Chromatin at the intersection of viral infection and DNA damage. *Biochim Biophys Acta*.

Lilley, C.E., R.A. Schwartz, and M.D. Weitzman. 2007. Using or abusing: viruses and the cellular DNA damage response. *Trends Microbiol*. 15:119-26.

Lin, S., and X.D. Fu. 2007. SR proteins and related factors in alternative splicing. *Adv Exp Med Biol*. 623:107-22.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature*. 362:709-15.

Luger, K., A.W. Mader, R.K. Richmond, D.F. Sargent, and T.J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*. 389:251-60.

Mack, K.D., X. Jin, S. Yu, R. Wei, L. Kapp, C. Green, B. Herndier, N.W. Abbey, A. Elbaggari, Y. Liu, and M.S. McGrath. 2003. HIV insertions within and proximal to host cell genes are a common finding in tissues containing high levels of HIV DNA and macrophage-associated p24 antigen expression. *J Acquir Immune Defic Syndr*. 33:308-20.

Mahy, N.L., P.E. Perry, and W.A. Bickmore. 2002. Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol*. 159:753-63.

Mancheno-Corvo, P., and P. Martin-Duque. 2006. Viral gene therapy. *Clin Transl Oncol*. 8:858-67.

Mandell, G., J. Bennett, and R. Dolin. 2004. Principles and Practice of Infectious Diseases. Sixth Edition Churchill Livingstone.

Mayer, R., A. Brero, J. von Hase, T. Schroeder, T. Cremer, and S. Dietzel. 2005. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol*. 6:44.

McDonald, D., L. Wu, S.M. Bohks, V.N. KewalRamani, D. Unutmaz, and T.J. Hope. 2003. Recruitment of HIV and its receptors to dendritic cell-T cell junctions. *Science*. 300:1295-7.

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

Mellert, W., A. Kleinschmidt, J. Schmidt, H. Festl, S. Emler, W.K. Roth, and V. Erfle. 1990. Infection of human fibroblasts and osteoblast-like cells with HIV-1. *AIDS*. 4:527-35.

Metais, J.Y., and C.E. Dunbar. 2008. The MDS1-EVI1 gene complex as a retrovirus integration site: impact on behavior of hematopoietic cells and implications for gene therapy. *Mol Ther*. 16:439-49.

Miller, M.D., C.M. Farnet, and F.D. Bushman. 1997. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J Virol*. 71:5382-90.

Mitchell, R., C.Y. Chiang, C. Berry, and F. Bushman. 2003. Global analysis of cellular transcription following infection with an HIV-based vector. *Mol Ther*. 8:674-87.

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

Miyauchi, K., Y. Kim, O. Latinovic, V. Morozov, and G.B. Melikyan. 2009. HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. *Cell*. 137:433-44.

Modlich, U., J. Bohne, M. Schmidt, C. von Kalle, S. Knoss, A. Schambach, and C. Baum. 2006. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood*. 108:2545-53.

Molenaar, C., K. Wiesmeijer, N.P. Verwoerd, S. Khazen, R. Eils, H.J. Tanke, and R.W. Dirks. 2003. Visualizing telomere dynamics in living mammalian cells using PNA probes. *Embo J*. 22:6631-41.

Morey, C., N.R. Da Silva, P. Perry, and W.A. Bickmore. 2007. Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development*. 134:909-19.

Müller, S., M. Neusser, D. Köhler, and M. Cremer. 2007. Preparation of Complex DNA Probe Sets for 3D FISH with up to Six Different Fluorochromes. *CSH Protocols*. pdb.prot4730.

Neusser, M., V. Schubel, A. Koch, T. Cremer, and S. Muller. 2007. Evolutionarily conserved, cell type and species-specific higher order chromatin arrangements in interphase nuclei of primates. *Chromosoma*. 116:307-20.

Nolan, G. 2009. Retroviral Life cycle. *In* Tutorials Nolan Lab. Stanford University.

Nunnari, G., E. Argyris, J. Fang, K.E. Mehlman, R.J. Pomerantz, and R. Daniel. 2005. Inhibition of HIV-1 replication by caffeine and caffeine-related methylxanthines. *Virology*. 335:177-84.

Osborne, C.S., L. Chakalova, K.E. Brown, D. Carter, A. Horton, E. Debrand, B. Goyenechea, J.A. Mitchell, S. Lopes, W. Reik, and P. Fraser. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet*. 36:1065-71.

Osborne, C.S., L. Chakalova, J.A. Mitchell, A. Horton, A.L. Wood, D.J. Bolland, A.E. Corcoran, and P. Fraser. 2007. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol*. 5:e192.

Ott, M.G., M. Schmidt, K. Schwarzwaelder, S. Stein, U. Siler, U. Koehl, H. Glimm, K. Kuhlcke, A. Schilz, H. Kunkel, S. Naundorf, A. Brinkmann, A. Deichmann, M. Fischer, C. Ball, I. Pilz, C. Dunbar, Y. Du, N.A. Jenkins, N.G. Copeland, U. Luthi, M. Hassan, A.J. Thrasher, D. Hoelzer, C. von Kalle, R. Seger, and M. Grez. 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med*. 12:401-9.

Pace, M.J., L. Agosto, E.H. Graf, and U. O'Doherty. 2011. HIV reservoirs and latency models. *Virology*.

Palella, F.J., Jr., K.M. Delaney, A.C. Moorman, M.O. Loveless, J. Fuhrer, G.A. Satten, D.J. Aschman, and S.D. Holmberg. 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med*. 338:853-60.

Permanyer, M., E. Ballana, and J.A. Este. 2010. Endocytosis of HIV: anything goes. *Trends Microbiol*. 18:543-51.

Pfeifer, G.P., S.D. Steigerwald, P.R. Mueller, B. Wold, and A.D. Riggs. 1989. Genomic sequencing and methylation analysis by ligation mediated PCR. *Science*. 246:810-3.

Popovic, M., E. Read-Connole, and R.C. Gallo. 1984. T4 positive human neoplastic cell lines susceptible to and permissive for HTLV-III. *Lancet*. 2:1472-3.

Quivy, V., E. Adam, Y. Collette, D. Demonte, A. Chariot, C. Vanhulle, B. Berkhout, R. Castellano, Y. de Launoit, A. Burny, J. Piette, V. Bours, and C. Van Lint. 2002. Synergistic activation of human immunodeficiency virus type 1 promoter activity by NF-kappaB and inhibitors of deacetylases: potential perspectives for the development of therapeutic strategies. *J Virol*. 76:11091-103.

Ragoczy, T., M.A. Bender, A. Telling, R. Byron, and M. Groudine. 2006. The locus control region is required for association of the murine beta-globin locus with

engaged transcription factories during erythroid maturation. *Genes Dev*. 20:1447-57.

Ratner, L., W. Haseltine, R. Patarca, K.J. Livak, B. Starcich, S.F. Josephs, E.R. Doran, J.A. Rafalski, E.A. Whitehorn, K. Baumeister, and et al. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature*. 313:277-84.

Reddy, K.L., J.M. Zullo, E. Bertolino, and H. Singh. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*. 452:243-7.

Richmond, T.J., and C.A. Davey. 2003. The structure of DNA in the nucleosome core. *Nature*. 423:145-50.

Rosenberg, S.A., P. Aebersold, K. Cornetta, A. Kasid, R.A. Morgan, R. Moen, E.M. Karson, M.T. Lotze, J.C. Yang, S.L. Topalian, and et al. 1990. Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *N Engl J Med*. 323:570-8.

Rynditch, A.V., S. Zoubak, L. Tsyba, N. Tryapitsina-Guley, and G. Bernardi. 1998. The regional integration of retroviral sequences into the mosaic genomes of mammals. *Gene*. 222:1-16.

Schambach, A., H. Wodrich, M. Hildinger, J. Bohne, H.G. Krausslich, and C. Baum. 2000. Context dependence of different modules for posttranscriptional enhancement of gene expression from retroviral vectors. *Mol Ther*. 2:435-45.

Schmidt, M., G. Hoffmann, M. Wissler, N. Lemke, A. Mussig, H. Glimm, D.A. Williams, S. Ragg, C.U. Hesemann, and C. von Kalle. 2001. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. *Hum Gene Ther*. 12:743-9.

Schönhuber, W., B. Fuchs, S. Juretschko, and R. Amann. 1997. Improved sensitivity of whole-cell hybridization by the combination of horseradish peroxidase-labeled oligonucleotides and tyramide signal amplification. *Appl Environ Microbiol*. 63:3268-73.

Schröder, A.R., P. Shinn, H. Chen, C. Berry, J.R. Ecker, and F. Bushman. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*. 110:521-9.

Shun, M.C., N.K. Raghavendra, N. Vandegraaff, J.E. Daigle, S. Hughes, P. Kellam, P. Cherepanov, and A. Engelman. 2007. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev*. 21:1767-78.

Simonis, M., P. Klous, E. Splinter, Y. Moshkin, R. Willemsen, E. de Wit, B. van Steensel, and W. de Laat. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet*. 38:1348-54.

Singh, D.P., E. Kubo, Y. Takamura, T. Shinohara, A. Kumar, L.T. Chylack, Jr., and N. Fatma. 2006. DNA binding domains and nuclear localization signal of LEDGF: contribution of two helix-turn-helix (HTH)-like domains and a stretch of 58 amino acids of the N-terminal to the trans-activation potential of LEDGF. *J Mol Biol*. 355:379-94.

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

Spector, D.L., and A.I. Lamond. 2010. Nuclear Speckles. *Cold Spring Harb Perspect Biol*.

Stein, S., M.G. Ott, S. Schultze-Strasser, A. Jauch, B. Burwinkel, A. Kinner, M. Schmidt, A. Kramer, J. Schwable, H. Glimm, U. Koehl, C. Preiss, C. Ball, H. Martin, G. Gohring, K. Schwarzwaelder, W.K. Hofmann, K. Karakaya, S. Tchatchou, R. Yang, P. Reinecke, K. Kuhlcke, B. Schlegelberger, A.J. Thrasher, D. Hoelzer, R. Seger, C. von Kalle, and M. Grez. 2010. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med*. 16:198-204.

Sun, H.B., J. Shen, and H. Yokota. 2000. Size-dependent positioning of human chromosomes in interphase nuclei. *Biophys. J.* 79:184-90.

Sutherland, H., and W.A. Bickmore. 2009. Transcription factories: gene expression in unions? *Nat Rev Genet*. 10:457-66.

Suzuki, Y., and R. Craigie. 2007. The road to chromatin - nuclear entry of retroviruses. *Nat Rev Microbiol*. 5:187-96.

Szczerbal, I., H.A. Foster, and J.M. Bridger. 2009. The spatial repositioning of adipogenesis genes is correlated with their expression status in a porcine mesenchymal stem cell adipogenesis model system. *Chromosoma*. 118:647-63.

Takizawa, T., P.R. Gudla, L. Guo, S. Lockett, and T. Misteli. 2008a. Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. *Genes Dev*. 22:489-98.

Takizawa, T., K.J. Meaburn, and T. Misteli. 2008b. The meaning of gene positioning. *Cell*. 135:9-13.

Tazi, J., N. Bakkour, V. Marchand, L. Ayadi, A. Aboufirassi, and C. Branlant. 2010. Alternative splicing: regulation of HIV-1 multiplication as a target for therapeutic action. *FEBS J*. 277:867-76.

Telenius, H., N.P. Carter, C.E. Bebb, M. Nordenskjold, P.A.I. Ponder, and A. Tunnacliffe. 1992. Degenerate oligonucleotide PCR: general amplification of target DNA by a single degenerate primer. *Genomics*. 13:718-725.

Thiel, G., Rössler, O. 2007. Viren als Werkzeuge der Gentherapie. *In* Biol. Unserer Zeit. Vol. 4. 241-248.

Thomas, C.E., A. Ehrhardt, and M.A. Kay. 2003. Progress and problems with the use of viral vectors for gene therapy. *Nat Rev Genet*. 4:346-58.

Thompson, M., R.A. Haeusler, P.D. Good, and D.R. Engelke. 2003. Nucleolar clustering of dispersed tRNA genes. *Science*. 302:1399-401.

Thormeyer, T. 2005. Untersuchungen zur Lokalisation der nukleären Matrix in Zellkernen nach induzierter Chromatinkondensation. *In* Biozentrum. LMU, München.

Tremethick, D.J. 2007. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell*. 128:651-4.

Trojer, P., and D. Reinberg. 2007. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell*. 28:1-13.

Tsyba, L., A.V. Rynditch, E. Boeri, K. Jabbari, and G. Bernardi. 2004. Distribution of HIV-1 in the genomes of AIDS patients. *Cell Mol Life Sci*. 61:721-6.

Tumbar, T., and A.S. Belmont. 2001. Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat Cell Biol*. 3:134-9.

van 't Wout, A.B., G.K. Lehrman, S.A. Mikheeva, G.C. O'Keeffe, M.G. Katze, R.E. Bumgarner, G.K. Geiss, and J.I. Mullins. 2003. Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4(+)-T-cell lines. *J Virol*. 77:1392-402.

Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew,

D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, et al. 2001. The sequence of the human genome. *Science*. 291:1304-51.

Volpi, E.V., E. Chevret, T. Jones, R. Vatcheva, J. Williamson, S. Beck, R.D. Campbell, M. Goldsworthy, S.H. Powis, J. Ragoussis, J. Trowsdale, and D. Sheer. 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.* 113:1565-76.

Wang, G.P., A. Ciuffi, J. Leipzig, C.C. Berry, and F.D. Bushman. 2007. HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res*. 17:1186-94.

Waterston, R.H., K. Lindblad-Toh, E. Birney, J. Rogers, J.F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S.E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M.R. Brent, D.G. Brown, S.D. Brown, C. Bult, J. Burton, J. Butler, R.D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A.T. Chinwalla, D.M. Church, M. Clamp, C. Clee, F.S. Collins, L.L. Cook, R.R. Copley, A. Coulson, O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K.D. Delehaunty, J. Deri, E.T. Dermitzakis, C. Dewey, N.J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D.M. Dunn, S.R. Eddy, L. Elnitski, R.D. Emes, P. Eswara, E. Eyras, A. Felsenfeld, G.A. Fewell, P. Flicek, K. Foley, W.N. Frankel, L.A. Fulton, R.S. Fulton, T.S. Furey, D. Gage, R.A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Grafham, T.A. Graves, E.D. Green, S. Gregory, R. Guigo, M. Guyer, R.C. Hardison, D. Haussler, Y. Hayashizaki, L.W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D.B. Jaffe, L.S. Johnson, M. Jones, T.A. Jones, A. Joy, M. Kamal, E.K. Karlsson, et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 420:520-62.

Watts, J.M., K.K. Dang, R.J. Gorelick, C.W. Leonard, J.W. Bess, Jr., R. Swanstrom, C.L. Burch, and K.M. Weeks. 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*. 460:711-6.

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

Wellensiek, B.P., R. Ramakrishnan, V. Sundaravaradan, R. Mehta, D.T. Harris, and N. Ahmad. 2009. Differential HIV-1 integration targets more actively transcribed host genes in neonatal than adult blood mononuclear cells. *Virology*. 385:28-38.

Wieser, R. 2007. The oncogene and developmental regulator EVI1: expression, biochemical properties, and biological functions. *Gene*. 396:346-57.

Williams, R.R., V. Azuara, P. Perry, S. Sauer, M. Dvorkina, H. Jorgensen, J. Roix, P. McQueen, T. Misteli, M. Merkenschlager, and A.G. Fisher. 2006. Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. *J Cell Sci*. 119:132-40.

Williams, R.R., S. Broad, D. Sheer, and J. Ragoussis. 2002. Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. *Exp Cell Res*. 272:163-75.

Woodcock, C.L. 2006. Chromatin architecture. *Curr Opin Struct Biol*. 16:213-20.

Woodcock, C.L., and S. Dimitrov. 2001. Higher-order structure of chromatin and chromosomes. *Curr Opin Genet Dev*. 11:130-5.

Woodcock, C.L., and R.P. Ghosh. 2010. Chromatin higher-order structure and dynamics. *Cold Spring Harb Perspect Biol*. 2:a000596.

Worobey, M., M. Gemmel, D.E. Teuwen, T. Haselkorn, K. Kunstman, M. Bunce, J.J. Muyembe, J.M. Kabongo, R.M. Kalengayi, E. Van Marck, M.T. Gilbert, and S.M. Wolinsky. 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature*. 455:661-4.

Worobey, M., P. Telfer, S. Souquiere, M. Hunter, C.A. Coleman, M.J. Metzger, P. Reed, M. Makuwa, G. Hearn, S. Honarvar, P. Roques, C. Apetrei, M. Kazanji, and P.A. Marx. 2010. Island biogeography reveals the deep history of SIV. *Science*. 329:1487.

Wu, X., Y. Li, B. Crise, and S.M. Burgess. 2003. Transcription start regions in the human genome are favored targets for MLV integration. *Science*. 300:1749-51.

Wu, X., Y. Li, B. Crise, S.M. Burgess, and D.J. Munroe. 2005. Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. *J Virol*. 79:5211-4.

Zhang, G., M.L. Zapp, G. Yan, and M.R. Green. 1996. Localization of HIV-1 RNA in mammalian nuclei. *J Cell Biol*. 135:9-18.

Zhao, R., M.S. Bodnar, and D.L. Spector. 2009. Nuclear neighborhoods and gene expression. *Curr Opin Genet Dev*. 19:172-9.

Zink, D., M.D. Amaral, A. Englmann, S. Lang, L.A. Clarke, C. Rudolph, F. Alt, K. Luther, C. Braz, N. Sadoni, J. Rosenecker, and D. Schindelhauer. 2004. Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. *J Cell Biol*. 166:815-25.

Zwirglmaier, K. 2005. Fluorescence in situ hybridisation (FISH)--the next generation. *FEMS Microbiol Lett*. 246:151-8.

8. Danksagung

An erster Stelle möchte ich mich ganz besonders bei meinem Betreuer und Doktorvater PD Dr. Steffen Dietzel für seine Unterstützung und Bemühungen bedanken. Er hat stets ein offenes Ohr für meine Anliegen und trägt maßgeblichen Anteil am Gelingen dieser Arbeit.

Prof. Dr. Thomas Cremer danke ich herzlichst, dass ich meine Doktorarbeit in seiner Arbeitsgruppe mit einem äußerst angenehmen Arbeitsklima anfertigen durfte, so dass ich – auch wenn die Experimente manchmal etwas nervenbeanspruchend waren – immer gerne auf die Arbeit kam.

Bei Prof. Dr. Ruth Brack-Werner möchte ich mich für die langjährige Zusammenarbeit und ergiebige Diskussionen und Anregungen, die zur Entwicklung und Gestaltung dieser Arbeit beigetragen haben, bedanken.

Marion, danke für deine beruhigenden Worte bei ungeplanten und völlig unerwarteten Entsorgungsaktionen meinerseits und für die herzliche Aufnahme an meinem neuen Büroplatz.

Pfiffige Katrin der guten Laune: Hey Hey uuhhuuhuuhuuhuuhuuuuu…Yeah Yeah uuhhuuhuuhuuhuuhuuuuu (gesendet von meinem Palm-Pre). Beim Wipe-Out teilnehmen zusammen unbedingt mal wir müssen oder im Spaßlabor Fürstenried-Wurst uns wir sehen. Jetzt bin ich erstmal „wie weg". Putzig!

Yola, thank you for emotional cooking and baking sessions and your preference for birds – in general. Brrravo!

Hilli, danke nochmals für die Übernachtungsmöglichkeit in dieser eisig kalten Dezembernacht. Da werde ich mich auf jeden Fall einmal revanchieren. Versprochen.

Zuni, mein langjähriger Begleiter auf beruflichen und privaten Wegen, danke für die gemeinsame Zeit an der Uni, wodurch der Forschungsalltag immer mit diversen Perlen und einer gewissen trockenen, direkten und auch subtilen Note an Spass verfeinert wurde. Und nochmals danke für dein Verständnis, dass die Technik und Tankanzeige neuerer Autos manchmal etwas unberechenbar ist.

Püffi, danke fürs Zeigen, wie man mit Niederlagen am Kickertisch umgeht - und natürlich für die gemeinsame Leidenschaft für bayuvarischer Traditionsfeste.

Kathi, danke für die freundlichen und äußerst unterhaltsamen Fahrdienste und für deine kulinarische Vorliebe für Apfelstrudel und flüssige Raufaser.

Doris, danke für die leckeren Kekse und die ein oder andere Überraschung in Form von Süßigkeiten an meinem Büroplatz.

Barbara, ebenfalls danke für die herzliche Aufnahme an meinem neuen Büroplatz und für immer nette Gespräche.

Ehrenwörtliche Versicherung

Hiermit bestätige ich, dass ich die vorliegende Dissertation selbstständig und nur mit den angegebenen Materialien und Quellen durchgeführt habe.

München, den 10.08.2011

Jens Nagel

Erklärung über frühere Promotionsversuche

Hiermit bestätige ich, dass ich vor dieser Arbeit keine anderen Promotionsversuche unternommen habe.

München, den 10.08.2011

Jens Nagel