Max von Pettenkofer-Institut, Virologie

Institut der Ludwig-Maximilians-Universität München

Vorstand: Professor Dr. med. Oliver T. Keppler

Exploring the role of cell adhesion proteins in retrovirus transmission and establishing CRISPR/Cas9 target enrichment Nanopore sequencing for retrovirus integration analysis

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von
Lisa Falk

aus
Eggenfelden, Deutschland

Jahr
2022

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Betreuerin:	PD Dr. rer. nat. Barbara Adler
Zweitgutachter:	Prof. Dr. rer. nat. Vigo Heissmeyer
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	05. Juni 2023

Table of contents

Summary

Retroviruses like the human-pathogenic human immunodeficiency virus (HIV) and the mouse-pathogenic murine leukemia virus (MLV) can disseminate between leukocytes by cell contact-dependent transmission. Productively infected cells (donor) were shown to transmit virus particles to surrounding lymphocytes (target) across tight cell-cell contacts designated virological synapses (*cis*-infection). These structures resemble the immunological synapse between leukocytes during immune cell priming and support efficient virus spread *in vitro*. While intravital microscopy of MLV infection in the lymph nodes of wild-type (WT) C57BL/6 mice revealed stable cell-cell contacts between infected donor and target lymphocytes, their contribution to retrovirus spread *in vivo* has yet to be established.

In this study, we examine the function of the cell adhesion-mediating proteins LFA1 (CD11a/CD18) and ICAM1 in cell contact-dependent transmission to assess the role of *cis*-infection for retroviral spread. Antibody-mediated blocking of LFA1 and ICAM1 in co-culture assays of MLV-infected (MLV+) primary CD19+ B1 cells with non-infected FoxP3+ T cells indicates a crucial function of LFA1 and ICAM1 in MLV *cis*-infection. To study the individual role of LFA1 and ICAM1 on donor and target cells, we combine primary lymphocytes isolated from ICAM1-knockout (KO) and CD11a-KO (LFA1-deficient) mice with WT-derived cells in the established *in vitro* co-culture assay. Interestingly, efficient MLV cell-to-cell transmission critically depends on expression of LFA1 on non-infected target cells, and expression of LFA1-ligand ICAM1 on MLV-infected donor cells. Adoptive transfer experiments with *in vitro*-transduced MLV+ FoxP3+ T cells are used to characterize the individual contribution of LFA1 and ICAM1 during *cis*-infection *in vivo*. Strikingly, consistent with our *in vitro* findings, LFA1 expression on target cells and the presence of ICAM1 on MLV-infected donor cells determine efficiency of MLV transmission *in vivo*.

During *in vitro* HIV infection, cell-to-cell transmission from productively infected donor cells to target lymphocytes across a cell-cell interface was previously shown to exceed efficiency of cell-free infection by 100- to 10,000-fold. Cell contact-dependent spread protected HIV from some neutralizing antibodies and could overcome the effect of certain cellular restriction factors and anti-retroviral drugs. Importantly, cell-cell contacts supported simultaneous transmission of multiple retroviral genomes, resulting in multicopy integration of HIV genomes. With the objective to characterize the consequences of cell-to-cell transmission at a proviral level, previous studies applied provirus detection techniques, such as fluorescence *in situ* hybridization (FISH) and multifluorescent reporter systems. Unfortunately, all approaches have limitations in their performance to detect the total number of proviruses together with the exact chromosomal integration site.

Here, we establish CRISPR/Cas9 target enrichment Nanopore sequencing as a novel provirus detection technique for localization and quantification of individual proviruses integrated in the genome of an infected host cell. This approach applies CRISPR/Cas9 activity for selective target enrichment without the requirement of PCR amplification and

combines it with the long-read capacity of Nanopore sequencing. We validate our workflow on a library of individual, GFP-expressing (GFP+) cell clones, containing a GFP-encoding retrovirus reporter genome at distinct frequencies and locations. Purified genomic DNA of GFP+ cell clones is subjected to CRISPR/Cas9 target enrichment Nanopore sequencing. Genomic mapping of reads containing GFP sequences identifies retrovirus reporter genome integration sites within the host genome. Strikingly, the number of detected proviruses positively correlates with the GFP expression level (mean fluorescence intensity) and the relative GFP copy number for each analyzed GFP+ cell clone. Breakpoint-spanning PCR and Sanger sequencing validate accuracy of identified GFP-encoding retrovirus reporter integration sites. To implement our novel provirus detection technique for retroviral transmission studies, we combine CRISPR/Cas9 target enrichment Nanopore sequencing with *in vitro* co-culture assays. Interestingly, our first results indicate a contribution of both, provirus frequency and provirus integration site, to efficiency of MLV transmission from an MLV-infected donor cell to a non-infected target cell.

In summary, the findings in our study reveal that cell adhesion-mediating proteins LFA1 and ICAM1 support efficient retroviral transmission *in vitro* and *in vivo*. In addition, with CRISPR/Cas9 target enrichment Nanopore sequencing, we established a long-read sequencing approach that allows for rapid and accurate identification of provirus integration sites. In future experiments, this technique can be applied to quantitatively and qualitatively assess the consequences of cell contact-dependent transmission during retrovirus infection.

Zusammenfassung

Retroviren wie das humanpathogene Humane Immundefizienz-Virus (HIV) und das mauspathogene Murine Leukämievirus (MLV) können Zellkontakt-abhängig zwischen Leukozyten übertragen werden. Produktiv infizierte Zellen (Donoren) transferieren dabei Viruspartikel über Zell-Zell-Kontakte, sogenannte virologische Synapsen (*cis*-Infektion), zu benachbarten Lymphozyten (Akzeptoren). Diese Strukturen ähneln der Immunologischen Synapse, die zwischen Leukozyten während des Primings von T-Zellen geformt wird, und ermöglichen hoch effiziente Virusübertragung. Intravitalmikroskopie der Lymphknoten MLV-infizierter Wildtyp (WT) C57BL/6 Mäuse erlaubte die Visualisierung stabiler Zell-Zell-Kontakte zwischen infizierten Donorzellen und nicht-infizierten Akzeptorzellen. Deren Bedeutung für die Ausbreitung von Retroviren *in vivo* ist allerdings unklar.
In dieser Studie untersuchen wir die Funktion der Zelladhäsionsproteine LFA1 (CD11a/CD18) und ICAM1 während Zellkontakt-abhängiger retroviraler Übertragung, um die Rolle der *cis*-Infektion für die Ausbreitung von Retroviren zu verstehen. Kokulturen von MLV-infizierten (MLV+), primären CD19+ B1-Zellen mit nicht-infizierten FoxP3+ T-Zellen in Gegenwart von blockierenden Antikörpern deuten an, dass LFA1 und ICAM1 eine bedeutsame Rolle bei der *cis*-Infektion zukommt. Um den Beitrag von LFA1 und ICAM1 auf Donor- und Akzeptorzellen individuell zu untersuchen, kombinieren wir primäre Lymphozyten, isoliert von WT C57BL/6, ICAM1-knockout (KO) und CD11a-KO (LFA1-defizienten) Mäusen, in *in vitro cis*-Infektionsexperimenten. Interessanterweise determiniert die Expression von LFA1 auf der Oberfläche nicht-infizierter Akzeptorzellen und die Expression des LFA1-Liganden ICAM1 auf MLV-infizierten Donorzellen die Effizienz der Zell-Zell-Übertragung von MLV. Um den individuellen Beitrag von LFA1 und ICAM1 während der *cis*-Infektion *in vivo* zu charakterisieren, überführen wir mittels adoptivem Zelltransfer *in vitro*-transduzierte MLV+ FoxP3+ T-Zellen in lebende Mäuse. Im Einklang mit unseren Ergebnissen unter *in vitro* Bedingungen, beeinflussen die Präsenz von LFA1 auf Akzeptorzellen und ICAM1 auf Donorzellen die Effizienz der MLV Übertragung *in vivo*.

Frühere *in vitro* Infektionsstudien haben gezeigt, dass die Zell-Zell-Übertragung von produktiv HIV-infizierten Donorzellen zu Akzeptor-Lymphozyten über Zell-Zell-Kontakte 100- bis 10,000-fach effizienter ist als eine zell-freie Infektion. Die Zellkontakt-abhängige Übertragung schützt HIV vor zahlreichen neutralisierenden Antikörpern und setzt die Wirkung bestimmter zellulärer Restriktionsfaktoren und anti-retroviraler Medikamente aus. Vor allem unterstützen Zell-Zell-Kontakte die simultane Übertragung multipler retroviraler Partikel, die in mehrfacher Integration einzelner HIV-Genome resultieren können. Mit dem Ziel die Konsequenzen retroviraler Zell-Zell Kontakte auf proviraler Ebene zu charakterisieren, haben vergangene Studien Provirusdetektionstechniken, wie Fluoreszenz-*in-situ*-Hybridisierung (FISH) und Multifluoreszenz-Reportersysteme, verwendet. Leider liefern diese Ansätze keine Erkenntnisse über die Gesamtzahl integrierter retroviraler Genome zusammen mit deren Integrationsstelle innerhalb des Wirtsgenoms.

In dieser Studie etablieren wir CRISPR/Cas9 Zielanreicherung Nanopore-Sequenzierung als neue Provirusdetektionstechnik, um die Lokalisation und die Frequenz individueller, integrierter Proviren innerhalb des Genoms einer infizierten Wirtszelle zu bestimmen. Dieser Ansatz macht sich die Spezifizität des CRISPR/Cas9 Systems zu Nutze, um selektiv und ohne PCR-basierte Amplifikation Zielgenome anzureichern, und kombiniert dies mit der charakteristischen Eigenschaft von Nanopore-Sequenzierung, lange Reads zu generieren. Wir validieren unseren Ansatz anhand einer Bibliothek GFP-exprimierender (GFP+) Zellklone, die jeweils eine unterschiedliche Anzahl an GFP-kodierenden, retroviralen Reportergenomen an verschiedenen Integrationsstellen in ihrem Genom tragen. Die aufgereinigte, genomische DNA von GFP+ Zellklonen wird mit Hilfe von CRISPR/Cas9 Zielanreicherung Nanopore-Sequenzierung analysiert. Die Zuordnung von Reads, die Teile des GFP-kodierenden, retroviralen Reportergenoms enthalten, zu dem entsprechenden Referenzgenom erlaubt die Identifizierung von Integrationstellen des Reportergenoms innerhalb des Wirtsgenoms. Dabei korrelieren das GFP Expressionslevel (durchschnittliche Fluoreszenzintensität) und die relative GFP Kopienanzahl für jeden analysierten GFP+ Zellklon positiv mit der Anzahl detektierter Proviren. Anhand von PCR-basierter Amplifikation der Übergangsstellen von Wirtszellen- zu proviraler DNA und anschließender Analyse der amplifizierten Genfragmente mittels Sanger-Sequenzierung validieren wir die Präzision mit der einzelne Integrationsstellen des GFP-kodierenden, retroviralen Reportergenoms identifiziert wurden. Um unsere neue Provirusdetektionstechnik für erste Studien zur retroviralen Übertragung einzusetzen, kombinieren wir CRISPR/Cas9 Zielanreicherung Nanopore-Sequenzierung mit *in vitro* Kokultur Assays. Hierbei deuten unsere ersten Ergebnisse einen Einfluss der Provirusanzahl sowie der proviralen Integrationsstelle auf die Effizienz der MLV Übertragung von einer MLV-infizierten Donorzelle zu einer nicht-infizierten Akzeptorzelle an.

Zusammenfassend zeigen die Ergebnisse unserer Studie, dass die Zelladhäsionsproteine LFA1 und ICAM1 zu effizienter, retroviraler Übertragung *in vitro* und *in vivo* beitragen. Zudem haben wir mit CRISPR/Cas9 Zielanreicherung Nanopore-Sequenzierung einen neuen Ansatz etabliert, der die schnelle und akkurate Identifikation von proviralen Integrationsstellen ermöglicht. In zukünftigen Experimenten kann diese Technik angewandt werden, um die quantitativen und qualitativen Konsequenzen von Zellkontakt-abhängiger retroviraler Transmission zu untersuchen.

List of figures

List of tables

Abbreviations

A	Adenine
A. dest.	Distilled water (latin: *Aqua destillata*)
A.U.	Arbitrary unit
Ab	Antibody
APC	Antigen-presenting cell
ART	Antiretroviral therapy
AZT	Azidothymidine
BHQ	Black hole quencher
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
Cas	CRISPR-associated protein
CCR	CC chemokine receptor
CD	Cluster of differentiation
chr.	Chromosome
CIP	Calf intestinal alkaline phosphatase
CpG	5'-cytosine-phosphate-guanine-3'
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
cSMAC	Central supramolecular activation complex
CXCR	CXC chemokine receptor
d	Days
dA	Deoxyadenosine monophosphate
DC	Dendritic cell
ddPCR	Digital droplet polymerase chain reaction
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E	Ecotropic
e.g.	For example (latin: *exempli gratia*)
EDTA	Ethylenediaminetetraacetic acid
Env	Envelope
et al.	And others (latin: *et alii*)
FAM	Fluorescein amidite
FCS	Fetal calf serum
FISH	Fluorescence *in situ* hybridization
FLIPS	Full-length individual proviral sequencing
FOXP3	Forkhead Box P3
Fwd	Forward
g	Gravitational force
G	Guanine
Gag	Group-specific antigen

GFP	Green fluorescent protein
gp	Glycoprotein
GRCm38	Genome Reference Consortium Mouse Build 38
h	Hours
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hg38	Genome Reference Consortium Human Build 38
HIV	Human immunodeficiency virus
Hprt1	Hypoxanthine phosphoribosyltransferase 1 gene
HTLV	Human T-lymphotropic virus
I.U.	Infectious units
ICAM1	Intercellular adhesion molecule 1
IFN	Interferon
Ig	Immunoglobulin
IGV	Integrative Genomics Viewer
IL	Interleukin
IN	Integrase
IRES	Internal ribosome entry site
IS	Integration site
k	Kilo
KO	Knockout
l	Liter
LAD	Leukocyte adhesion deficiency
LAM-PCR	Linear amplification-mediated polymerase chain reaction
LFA1	Lymphocyte function-associated antigen 1
LPS	Lipopolysaccharide
LTR	Long terminal repeat
M	Molar (moles per liter)
MC	Methylcellulose
mCat1	Mouse cationic amino acid transporter 1
MFI	Mean fluorescence intensity
MHC	Major histocompatibility complex
MLV	Murine leukemia virus
MOI	Multiplicity of infection
MTOC	Microtubule-organizing center
N	Nucleotide
NGS	Next-Generation Sequencing
nt	Nucleotide (chain length measure)
ntc	Non-targeting control
O/N	Overnight
OH	Hydroxide
PAM	Protospacer adjacent motif
PBS	Primer binding site
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PI3K	Phosphoinositide 3-kinase

PIC	Pre-integration complex
pLN	Popliteal lymph node
pMHC	Peptide-major histocompatibility complex
Pol	Polymerase
PPT	Polypurine tract
pSMAC	Peripheral supramolecular activation complex
Q4PCR	Quadruplex quantitative polymerase chain reaction
QPCR	Quantitative polymerase chain reaction
Rev	Reverse
RNA	Ribonucleic acid
ROI	Region of interest
rpm	Revolutions per minute
RT	Room temperature
s.c.	Subcutaneous
T	Thymine
TCR	T cell receptor
TGF-β1	Transforming growth factor-β1
TNF	Tumor necrosis factor
tracrRNA	*Trans*-activating crRNA
TRIM5α	Tripartite motif 5 α
TSD	Target site duplication
U	Enzyme unit
vDNA	Viral DNA
VLP	Virus-like particle
WGA	Whole genome amplification
WT	Wild-type
X	Fold
Xpr1	Xenotropic and polytropic retrovirus receptor 1

1. Introduction

1.1 Finding a cure: The worldwide pandemic of HIV infection

With an estimated 38 million infected individuals and more than 690,000 causes of death in the year 2019, human immunodeficiency virus (HIV) infection represents a burden for public health on a global scale (Bracq et al., 2018; Pillay & Johnson, 2021). While the development of antiretroviral therapy (ART) resulted in a significant decline of HIV-associated mortality especially in developed countries, current treatment strategies do not provide a cure to HIV and are constantly challenged by HIV sequence diversity (Bracq et al., 2018; Kiselinova et al., 2015; Leite et al., 2019; Quinn, 2008). Yet, the design of effective treatment and prevention strategies requires a thorough understanding of the mechanistic details underlying HIV transmission early after infection and its consequences for HIV diversification (Bracq et al., 2018; Dixit & Perelson, 2005).
Since its discovery in the 1990s, cell-to-cell transmission of HIV, murine leukemia virus (MLV) and human T-lymphotropic virus (HTLV) has been characterized besides cell-free transmission as an additional mode of retroviral spread (Bracq et al., 2018; Igakura et al., 2003; Jin et al., 2009; Jolly et al., 2004; Phillips, 1994; Sherer et al., 2010). *In vitro*, cell contact-dependent HIV transmission exceeds cell-free transmission efficiency by 100- to 10,000-fold, reduces efficacy of various ART regimens, and represents a potential driving force of intra-host HIV diversification (Bracq et al., 2018; Chen et al., 2007; Del Portillo et al., 2011; Dimitrov et al., 1993; Law et al., 2016; Levy et al., 2004; Sigal et al., 2011). While HIV cell-to-cell transmission is considered as the predominant transmission mode in *in vitro* cell culture settings, the extent and clinical relevance of this mode of spread *in vivo* has yet to be established (Bracq et al., 2018; Chen et al., 2007; Sattentau, 2008). In this study, we apply the mouse-pathogenic MLV as a model retrovirus to study cell contact-dependent retroviral transmission under *in vivo* conditions.

1.2 Murine leukemia virus

1.2.1 Classification, structure and genome composition of MLV

Murine leukemia virus belongs to the genus *Gammaretrovirus*, within the subfamily *Orthoretrovirinae* of the family *Retroviridae*, and can induce myeloid leukemia and lymphoma in mice (Pepersack et al., 1980; Pi et al., 2019; Ru et al., 1993; Troxler & Scolnick, 1978; Young & Bishop, 2021). MLV virions harbor two copies of a non-segmented, single-stranded, positive-sense (+) RNA genome and replicate by integration of a viral DNA (vDNA) intermediate into the genome of a host cell, allocating MLV to the Baltimore Group VI (Young & Bishop, 2021).
Mature MLVs are spherical particles with a diameter of ~ 100 - 120 nm, and composed of a polyhedral core, harboring the RNA genome, which is encircled by a host-cell derived envelope (Fan, 1999; Rein, 2011; Yeager et al., 1998; Young & Bishop, 2021). Within the virion, viral enzymes, including protease, integrase and reverse transcriptase are encapsulated together with cellular transfer RNA, which is required for reverse transcription (Fan, 1999).

As a simple retrovirus, the MLV genome, with an average length of ~ 8.2 kbp, contains three genes, which are successively arranged as 5'-*gag-pol-env*-3' (Figure 1) (Fan, 1999; Young & Bishop, 2021). The *gag* gene encodes all proteins required for virus assembly and release, comprising the Gag polyprotein, and its proteolytic cleavage products matrix, p12, capsid and nucleocapsid (Young & Bishop, 2021). The *pol* gene encodes the catalytic machinery necessary for retroviral replication, including the Pol polyprotein, which upon proteolytic cleavage breaks down into protease, reverse transcriptase and integrase (Young & Bishop, 2021). Structural proteins mediating virus interaction with a permissive target cell are encoded by *env*, and comprise the Env polyprotein, and its products after proteolytic cleavage, the glycosylated surface protein and the transmembrane protein (Fan, 1999; Young & Bishop, 2021).

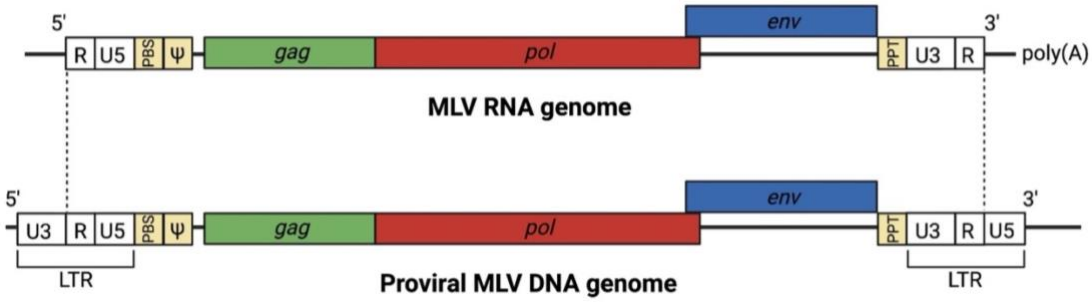

Figure 1: Structure of the MLV genome. The MLV RNA genome comprises coding and non-coding regions (Rein, 2011). Coding regions are successively arranged as 5'-*gag-pol-env*-3' (Fan, 1999; Young & Bishop, 2021). The *gag* gene encodes the Gag polyprotein, which can be proteolytically cleaved by viral protease into matrix, p12, nucleocapsid and capsid (Young & Bishop, 2021). The *pol* gene encodes the Pol polyprotein, which upon proteolytic cleavage yields protease, reverse transcriptase and integrase (Young & Bishop, 2021). The Env glycoprotein and its cleavage products, the surface protein and the transmembrane protein, are encoded by *env* (Fan, 1999; Young & Bishop, 2021). Non-coding regions include the primer binding site (PBS), the packaging signal ψ and the polypurine tract (PPT) (Rein, 2011). After reverse transcription of the MLV RNA genome, the resulting MLV DNA genome harbors characteristic long terminal repeats (LTRs) at both ends, each composed of the regions 5'-U3-R-U5-3'(Fan, 1999). While U3 harbors enhancer and promoter sequences for retroviral expression after integration into the host DNA, motifs within the R sequence mediate genomic cleavage and polyadenylation (Fan, 1999; Rein, 2011). Illustration was created according to (Fan, 1999; Rein, 2011).

Besides its coding regions, the MLV RNA genome harbors several essential, non-coding regions, comprising the primer binding site (PBS), the polypurine tract (PPT), serving as a primer for synthesis of the second vDNA strand, the packaging signal ψ, genomic motifs for integration of vDNA, as well as promoter and enhancer regions, which become part of the long terminal repeats (LTRs) (Rein, 2011). The LTRs are located at both ends of viral DNA as a consequence of reverse transcription of the MLV genome and comprise each three distinct regions in the order 5'-U3-R-U5-3' (Fan, 1999). While the U3 regions harbor enhancer and basal promoter sequences required for retroviral transcription by RNA polymerase II after integration into the host genome, the R regions contain repeat sequences serving as a signal for polyadenylation of the viral RNA genome (Fan, 1999; Rein, 2011).

With respect to the type of envelope protein and its corresponding cell surface receptor, MLV exhibits a broad host range (Fan, 1999; Kozak, 2015; Young & Bishop, 2021). Ecotropic MLV exclusively infects murine cells upon interaction with the mouse cationic amino acid transporter 1 (mCat1), encoded by the *Slc7a1* gene (Albritton et al., 1989; Fujisawa & Masuda, 2007; Wang et al., 1991; Young & Bishop, 2021). The tropism of xenotropic MLV comprises only non-murine cells, expressing the inorganic phosphate exporter Xpr1 (*Xpr1*) as a virus entry receptor (Battini et al., 1999; Fan, 1999; Young & Bishop, 2021). Both polytropic and modified polytropic MLVs share cell entry receptor Xpr1 with xenotropic MLV, yet their cell tropism includes murine and non-murine cells (Battini et al., 1999; Young & Bishop, 2021). Similarly, amphotropic MLV can infect both non-murine and murine cells, by using the phosphate transporter Pit1 (*Slc20a1*) and Pit2 (*Slc20a2*) as cell entry receptors (Miller & Miller, 1994; Rein, 2011; Wilson et al., 1994; Young & Bishop, 2021).

1.2.2 The MLV life cycle

As an enveloped virus, MLV enters a susceptible target cell through interaction of the viral Env glycoprotein with its respective cellular receptor and subsequent fusion of the viral envelope with a cellular membrane (Figure 2A) (Kamiyama et al., 2011; Katen et al., 2001; McClure et al., 1988, 1990; Nussbaum et al., 1993). After virus entry, viral reverse transcriptase catalyzes replication of the viral RNA genome into double-stranded, linear vDNA (Hu & Hughes, 2012; Serrao & Engelman, 2016). In sequential steps, divided in 3'-processing and DNA strand transfer, viral integrase enzymes catalyze the integration of vDNA into the host cell chromosome (Figure 2A and Figure 2B) (Craigie & Bushman, 2012; Serrao & Engelman, 2016). This integration process is initiated with formation of the viral pre-integration complex (PIC), a nucleoprotein complex comprising the vDNA multimerized with cell-derived proteins and virus-derived proteins, such as reverse transcriptase, nucleocapsid and integrase (Chen et al., 1999; Farnet & Haseltine, 1990; Farnet & Bushman, 1997; Hare et al., 2010, 2012; Hindmarsh & Leis, 1999; Lapadat-Tapolsky et al., 1993; Lee & Craigie, 1994; Lee & Coffin, 1991; Li et al., 2006; Serrao & Engelman, 2016). Within the PIC, integrase units interact with both LTR termini of vDNA to form a ternary complex, the so-called intasome (Figure 2B) (Chen et al., 1999; Hare et al., 2010, 2012; Hindmarsh & Leis, 1999; Li et al., 2006; Serrao & Engelman, 2016). During 3'-processing, integrase catalyzes cleavage of phosphodiesterbonds at a conserved motif (5'-CA-3') within the LTRs of the vDNA to release highly reactive hydroxyl groups (CA_{OH}-3') at both ends (Brown et al., 1989; Pauza, 1990; Fujiwara & Mizuuchi, 1988; Hare et al., 2012; Lee & Coffin, 1991; Roth et al., 1989; Serrao & Engelman, 2016). Access of MLV PICs to genomic DNA for proviral integration demands disassembly of the nuclear membrane at mitosis (Lewis et al., 1992; Lewis & Emerman, 1994; Roe et al., 1993; Serrao & Engelman, 2016). Inside the cell nucleus, integrase initiates DNA strand transfer within in a pre-catalytic target capture complex formed between the intasome and its selected target host DNA sequence (Engelman et al., 1991; Hare et al., 2012; Maertens et al., 2010; Serrao & Engelman, 2016). In a nucleophilic attack catalyzed by integrase, the 3'OH groups on CA_{OH}-3'termini of the vDNA react with phosphodiester groups of nucleotides shifted in

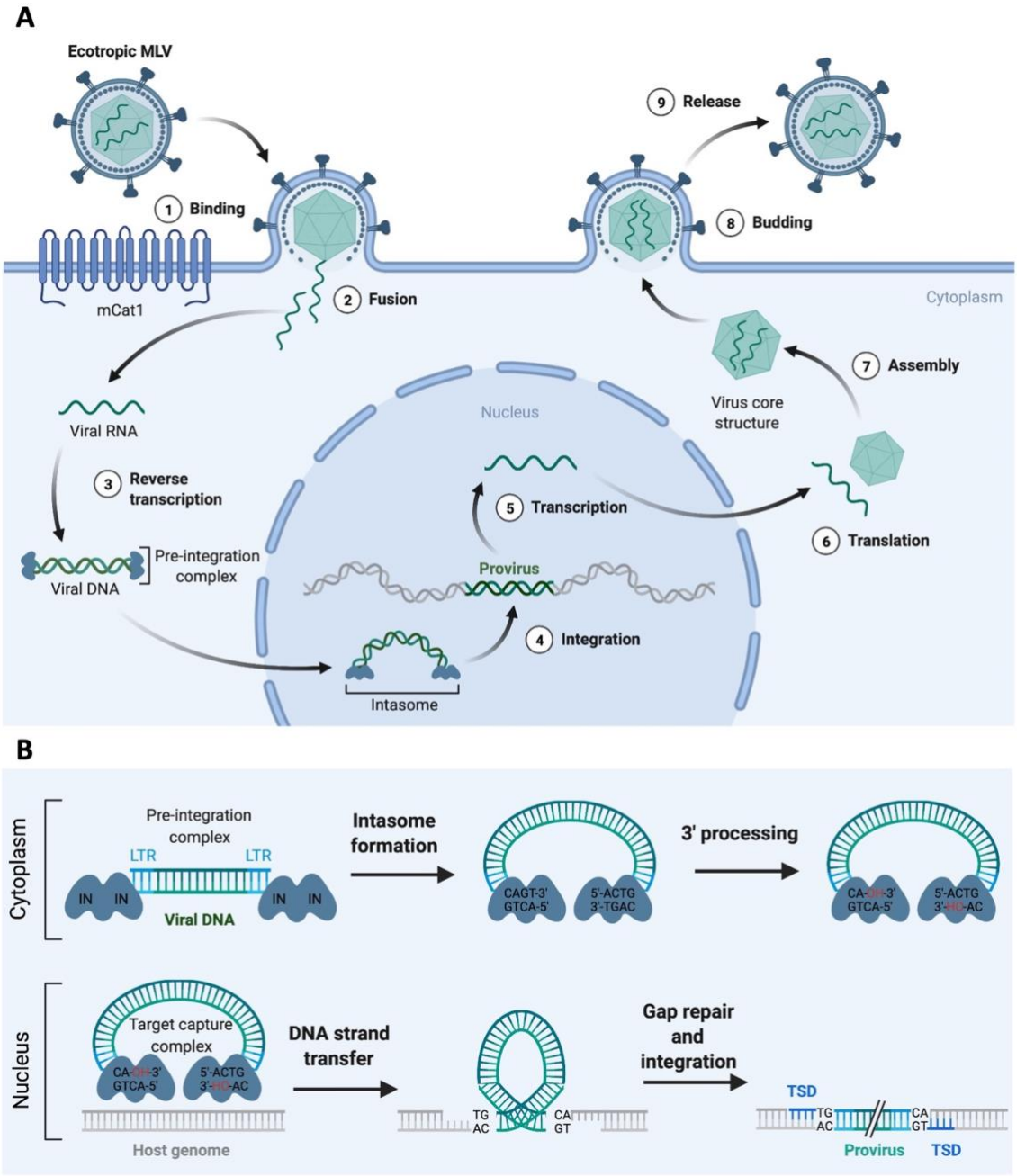

Figure 2: The MLV life cycle. (A) Overview of the MLV life cycle. Entry of ecotropic MLV into a susceptible target cell is initiated through interaction of the viral Env protein with the host cell receptor mouse cationic amino acid transporter 1 (mCat1), followed by fusion of the viral envelope with the cellular membrane (Albritton et al., 1989; Kamiyama et al., 2011; Katen et al., 2001; McClure et al., 1988, 1990; Nussbaum et al., 1993; Wang et al., 1991). In the cytoplasm, viral reverse transcriptase catalyzes replication of viral genomic RNA into linear, double-stranded viral DNA (Hu & Hughes, 2012; Serrao & Engelman, 2016). In complex with viral proteins, such as reverse transcriptase and integrase, viral DNA forms a so-called pre-integration complex (PIC) (Chen et al., 1999; Hare et al., 2010, 2012; Li et al., 2006; Serrao & Engelman, 2016). Nuclear entry of MLV-derived PICs demands the disassembly of the cell nucleus during mitosis (Lewis et al., 1992; Lewis & Emerman, 1994; Roe et al., 1993; Serrao & Engelman, 2016). Inside the nucleus, chromosomal integration of viral DNA into the host cell chromosome is catalyzed by integrase, resulting in a permanently integrated provirus (Serrao & Engelman, 2016). After retroviral protein expression, virus assembly is coordinated synergistically by the matrix, capsid and nucleocapsid domains of the Gag polyprotein at the plasma membrane or plasma membrane invaginations (Balasubramaniam & Freed, 2011; Li et al., 2014; Morita & Sundquist, 2004; Suomalainen et al., 1996). After the immature MLV particle is released from the host cell, maturation is initiated by proteolytic cleavage of the Gag polyproteins to generate infectious virus particles prior binding to another susceptible target cell (Fan, 1999; Rein, 2011;

Young & Bishop, 2021). (B) Individual stages of the retroviral integration process. In the cytoplasm, integrase (IN) molecules associate with the LTR termini of viral DNA to form the so-called intasome (Chen et al., 1999; Hare et al., 2010, 2012; Hindmarsh & Leis, 1999; Li et al., 2006; Serrao & Engelman, 2016). During 3' processing, integrase catalyzes cleavage of phosphodiester bonds to result in reactive CA-$_{OH}$-3' groups at both LTR ends (Brown et al., 1989; Pauza, 1990; Fujiwara & Mizuuchi, 1988; Hare et al., 2012; Lee & Coffin, 1991; Roth et al., 1989; Serrao & Engelman, 2016). Inside the nucleus, the intasome interacts with a target site within the host genome to form a target capture complex (Engelman et al., 1991; Hare et al., 2012; Maertens et al., 2010; Serrao & Engelman, 2016). In a nucleophilic attack, CA-$_{OH}$-3' groups of the viral DNA react with phosphodiester groups between the host DNA nucleotides on opposite strands of the target DNA (Brown, 1997; Craigie & Bushman, 2012; Engelman et al., 1991). As a result of the staggered attack of the target cell DNA, at both ends the integrated viral DNA is flanked by single stranded gaps of target DNA (Brown, 1997; Brown et al., 1989; Fujiwara & Mizuuchi, 1988; Serrao & Engelman, 2016). Once the strand transfer complex is disassembled, these nucleotide gaps are filled by host cell proteins and result in a characteristic target site duplication (TSD) of 4 bp, which flanks the MLV provirus (Kim et al., 2010; Serrao & Engelman, 2016; Vincent et al., 1990; Vink et al., 1990). (A) Illustration modified from Biorender.com template (B) Illustration was created according to (Lusic & Siliciano, 2017).

the 5'direction for four to six bases on opposite strands of the target DNA (Brown, 1997; Craigie & Bushman, 2012; Engelman et al., 1991). Binding energy of the cleaved phosphodiester groups is employed for the generation of a covalent bond between the 3' termini of the vDNA and the target DNA (Brown, 1997). As a result of the staggered attack of the target cell DNA, the integrated viral DNA is flanked at both ends by single stranded gaps of target DNA (Brown, 1997; Brown et al., 1989; Fujiwara & Mizuuchi, 1988; Serrao & Engelman, 2016). Once the strand transfer complex is disassembled, these nucleotide gaps are filled by host cell proteins and result in a target site duplication, which flanks the integrated vDNA (Serrao & Engelman, 2016; Vincent et al., 1990; Vink et al., 1990). The length of target site duplications varies with respect to the retrovirus type, and accounts for 5 bp after HIV integration and 4 bp after MLV integration (Kim et al., 2010; Serrao & Engelman, 2016; Vincent et al., 1990; Vink et al., 1990). Once integrated, viral DNA is denoted as a provirus (Hindmarsh & Leis, 1999). After retroviral protein expression, virus assembly is coordinated synergistically by the matrix, capsid and nucleocapsid domains of the Gag polyprotein at the plasma membrane or plasma membrane invaginations (Figure 2A) (Balasubramaniam & Freed, 2011; Li et al., 2014; Morita & Sundquist, 2004; Suomalainen et al., 1996). MLV virions bud from an infected host cell as immature particles, which harbor a characteristic core composed of Gag polyproteins (Young & Bishop, 2021). Once the MLV particle is released from its host cell, maturation is initiated by proteolytic cleavage of the Gag polyproteins to generate infectious virus particles prior binding to another susceptible target cell (Fan, 1999; Rein, 2011; Young & Bishop, 2021).

1.3 Modes of cell contact-dependent retroviral transmission

While classic models of retroviral infection are based on cell-free retroviral dissemination, early studies in the 1990s indicated that retroviruses can also spread in a cell contact-dependent mode (Bracq et al., 2018; Kilby & Eron, 2003; Mothes et al., 2010; Phillips, 1994; Pierson & Doms, 2003). Initial observations, such as the concentration of HIV particles at sites of cell-cell contact, the efficiency of HIV infection in the presence of dendritic cells, and the rapid spread of HIV in cell culture systems, significantly contributed to the current understanding of retroviral cell-to-cell transmission (Cameron et al., 1992; Dimitrov et al., 1993; Mothes et al., 2010; Phillips, 1994). To date, two distinct modes of cell contact-dependent retroviral spread have been identified, *trans*-infection (infectious synapse) and *cis*-infection (virological synapse) (Figure 3) (Bracq et al., 2018).

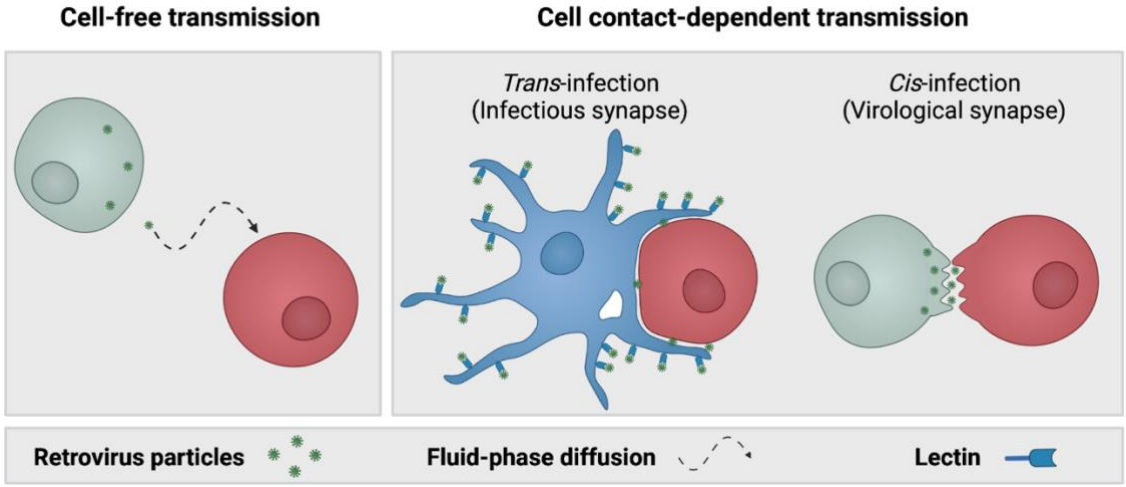

Figure 3: Modes of cell-free and cell contact-dependent retroviral transmission. Illustrations represent the different retroviral transmission modes (Marsh & Helenius, 2006; Mothes et al., 2010; Piguet & Sattentau, 2004). During cell-free transmission, free virions undergo fluid-phase diffusion prior attachment to surface receptors of a permissive target cell (Kilby & Eron, 2003; Pierson & Doms, 2003; Piguet & Sattentau, 2004; Stebbing et al., 2004). There are two distinct mechanisms of cell contact-dependent retroviral transmission: *trans*-infection (infectious synapse) and *cis*-infection (virological synapse) (Bracq et al., 2018). *Trans*-infection describes the cell contact-dependent transmission of retroviruses through binding of virions via cell surface proteins, such as C-type lectins, expressed on the surface of antigen-presenting cells (APCs), without becoming productively infected, and the subsequent presentation of these virions to a permissive target cell (Bobardt et al., 2003; Geijtenbeek et al., 2000; Hu et al., 2004; Nguyen & Hildreth, 2003; Piguet & Sattentau, 2004; Turville et al., 2002). *Cis*-infection is characterized by the directed transfer of retroviruses across a cell-cell interface, formed between a productively infected donor cell and a non-infected, permissive target cell (Agosto et al., 2018; Jolly et al., 2004; Jolly & Sattentau, 2004; Piguet & Sattentau, 2004). Illustration was created according to (Piguet & Sattentau, 2004).

Retroviral *trans*-infection is characterized by the capture of free virions by cell surface proteins, such as C-type lectins, expressed on antigen-presenting cells (APCs) and their subsequent transfer to a target cell (Bobardt et al., 2003; Geijtenbeek et al., 2000; Hu et al., 2004; Nguyen & Hildreth, 2003; Piguet & Sattentau, 2004; Turville et al., 2002). In this cell contact-dependent mode of transmission, productive retroviral infection of the donor cell is not required (Bobardt et al., 2003; Geijtenbeek et al., 2000; Hu et al., 2004; Nguyen & Hildreth, 2003; Piguet & Sattentau, 2004; Turville et al., 2002). The second

mechanism, *cis*-infection, describes the direct cell-to-cell transmission of retroviruses from an infected donor cell to a non-infected target cell across a tight cell-cell contact, designated virological synapse (Jolly et al., 2004; Jolly & Sattentau, 2004; Piguet & Sattentau, 2004).

1.4 Retroviral transmission by *cis*-infection (virological synapse)

1.4.1 Structural similarities of the virological and the immunological synapse

The term "virological synapse" describes the site of cell-cell contacts during retroviral *cis*-infection (Bracq et al., 2018; Piguet & Sattentau, 2004; Vasiliver-Shamis et al., 2010). It refers to its structural and mechanistic homolog, the immunological synapse, which forms between T cells and APCs in the course of antigen presentation (Bracq et al., 2018; Jolly et al., 2004; Vasiliver-Shamis et al., 2010). There, cognate peptide-major histocompatibility complexes (pMHCs) on the surface of APCs bind to T cell receptors (TCRs) to induce T cell activation by intracellular signaling cascades, which mediate cytokine secretion and cell proliferation (Agüera-Gonzalez et al., 2015; Bracq et al., 2018; Huppa & Davis, 2003). The immunological synapse is stabilized by cell adhesion molecules, such as the integrin lymphocyte function-associated antigen 1 (LFA1, CD11a/CD18) and its biologically most relevant ligand intercellular adhesion molecule 1 (ICAM1, CD54), which are arranged as an adhesive circle around the center of the cell-cell interface (Bachmann et al., 1997; de Fougerolles & Springer, 1992; Kukic et al., 2015; Shimaoka et al., 2001; Vasiliver-Shamis et al., 2010; Zimmerman & Blanco, 2008). Both cell adhesion molecules are key mediators of dynamic interactions on a cell-cell and cell-matrix level in a multitude of immunological processes, including T cell migration and T cell-mediated target cell killing (Comrie et al., 2015; Roebuck & Finnegan, 1999; Shimaoka et al., 2001; Springer, 1990; Van De Stolpe & Van Der Saag, 1996; Vasiliver-Shamis et al., 2010; Walling & Kim, 2018). Similar to the immunological synapse, formation of a virological synapse during retroviral *cis*-infection includes the actin-dependent recruitment and clustering of cell adhesion molecules, such as LFA1 and ICAM1, at the cell-cell interface, although their functional contribution to directed retroviral transfer *in vivo* remains unknown (Bracq et al., 2018; Jolly et al., 2004, 2007b; Vasiliver-Shamis et al., 2010).

1.4.2 Mechanism of retroviral transmission by *cis*-infection

Formation of a virological synapse was initially observed for the directed transfer of HTLV across cell-cell contact sites between HTLV-infected donor and non-infected target cells (Bracq et al., 2018; Igakura et al., 2003). With the first description of cell contact-dependent transmission for HIV, the virological synapse was characterized as a dynamic, actin-dependent, adhesive interaction between an infected donor and a non-infected target cell, which allows for directed transmission of virus particles across the cell-cell interface (Bracq et al., 2018; Jolly et al., 2004; Jolly & Sattentau, 2004; Piguet & Sattentau, 2004). Shortly thereafter, polarized MLV assembly at intercellular contact sites and directed, cell contact-dependent transmission of MLV among fibroblasts were described for the first time (Jin et al., 2009; Sherer et al., 2010).

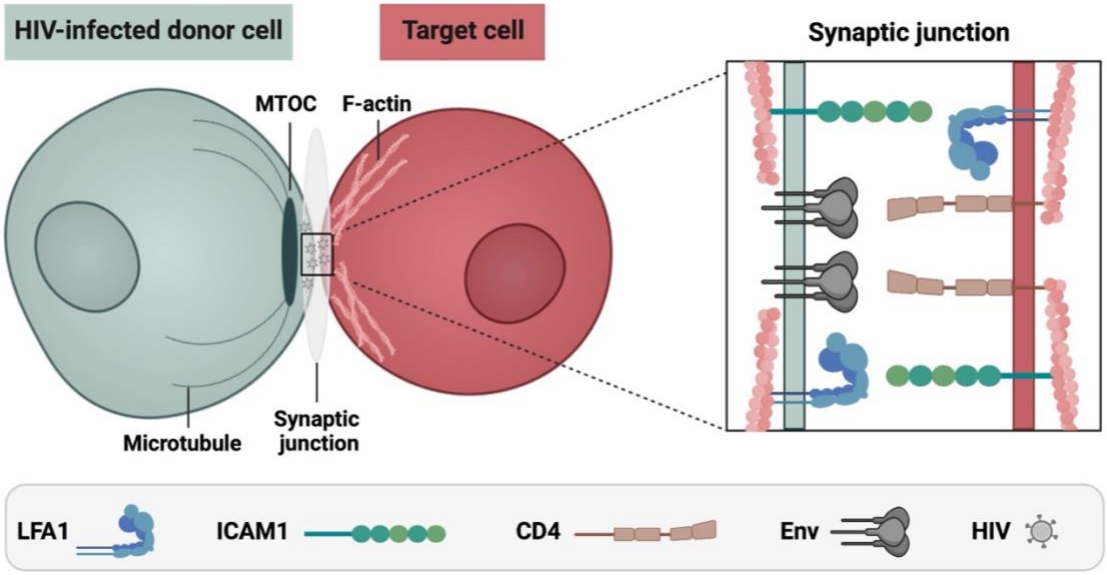

Figure 4: Structural organization of a virological synapse between an HIV-infected donor cell and a non-infected target cell during retroviral *cis*-infection. Formation of a virological synapse (*cis*-infection) between an HIV-infected CD4+ T cell (donor) and a non-infected CD4+ T cell (target) is initiated by interaction of the gp120 subunit of the Env glycoprotein, exposed on the donor cell surface, with the CD4 receptor expressed on the non-infected target T cell (Bracq et al., 2018; Jolly et al., 2004, 2007b). This interaction triggers actin-dependent recruitment of viral Gag proteins, co-receptors, and adhesion-mediating proteins lymphocyte function-associated antigen 1 (LFA1) and intercellular adhesion molecule 1 (ICAM1) toward the intercellular contact site (Bracq et al., 2018; Jolly et al., 2004, 2007a, 2007b; Jolly & Sattentau, 2005, 2007; Starling & Jolly, 2016; Vasiliver-Shamis et al., 2010). At the center of the synaptic junction, gp120 subunits assemble to a microcluster with high similarity to the central supramolecular activation complex (cSMAC) observed during formation of immunological synapses (Campi et al., 2005; Varma et al., 2006; Vasiliver-Shamis et al., 2009, 2010). LFA1/ICAM1 interactions cluster in a ring-like structure, resembling the peripheral supramolecular activation complex (pSMAC) described for immunological synapses, around the central gp120 microcluster (Jolly et al., 2004, 2007b; Vasiliver-Shamis et al., 2008, 2009, 2010). Formation of the virological synapse includes polarization of the microtubule-organizing center (MTOC) toward the synaptic cleft (Jolly et al., 2011; Piguet & Sattentau, 2004; Sol-Foulon et al., 2007; Vasiliver-Shamis et al., 2009). Virus assembly and release are oriented toward the cell-cell interface, followed by virus transmission to the permissive target cell (Fais et al., 1995; Jolly et al., 2007a; Jolly & Sattentau, 2004; Piguet & Sattentau, 2004). Illustration was created according to (Sattentau, 2008).

The concept of the virological synapse is best characterized for interaction between an HIV-infected CD4+ T cell with a non-infected CD4+ T cell, which will serve in the following paragraph as an example to describe its structural and mechanistic characteristics in more detail (Figure 4) (Alvarez et al., 2014; Bracq et al., 2018; Jolly et al., 2004, 2007a). Formation of the virological synapse is initiated through engagement between the HIV Env glycoprotein subunit gp120, exposed on the HIV-infected donor cell, and the CD4 receptor expressed on the target T cell (Bracq et al., 2018; Jolly et al., 2004, 2007b). CD4 receptor engagement induces a reorganization of the actin cytoskeleton and microtubule polarization in the donor cell toward the cell-cell contact site (Bracq et al., 2018; Jolly et al., 2007a). Within both the donor and the target cell, a precursor of the viral Gag protein, co-receptors chemokine receptor 4 (CXCR4) or CCR5, cell adhesion molecules LFA1 and ICAM1 as well as tetraspanins and the lipid raft marker ganglioside GM1 are gathered in an actin-dependent process at the cell-cell contact site (Bracq et al., 2018; Jolly et al., 2004, 2007a, 2007b; Jolly & Sattentau, 2005,

2007; Starling & Jolly, 2016; Vasiliver-Shamis et al., 2010). Similar to the structural organization of the immunological synapse, interacting viral proteins, receptors and cell adhesion-mediating proteins assemble in ringlike, discrete structures around the intercellular contact site of the virological synapse (Piguet & Sattentau, 2004; Vasiliver-Shamis et al., 2010). In the center of the synaptic junction, Env gp120 subunits assemble in a microcluster, which resembles the central supramolecular activation complex (cSMAC) composed of TCR/pMHC complexes found in center of the immunological synapse (Campi et al., 2005; Varma et al., 2006; Vasiliver-Shamis et al., 2009, 2010). Adhesion molecules LFA1 and its ligand ICAM1 cluster around the gp120 subunits, and assemble to a ring of adhesive interactions similar to the peripheral supramolecular activation complex (pSMAC) observed during formation of the immunological synapse (Jolly et al., 2004, 2007b; Vasiliver-Shamis et al., 2008, 2009, 2010).

Assembly of the virological synapse includes polarization of the microtubule-organizing center (MTOC) in the HIV-infected donor cell toward the synaptic cleft (Jolly et al., 2011; Piguet & Sattentau, 2004; Sol-Foulon et al., 2007; Vasiliver-Shamis et al., 2009). The microtubule network is hypothesized to actively support directed transfer of the viral Env glycoprotein to the cell-cell contact site, as Env-containing intracellular compartments are localized in close proximity to the MTOC (Bracq et al., 2018; Jolly et al., 2011; Starling & Jolly, 2016). Frequently, mitochondria and lysosomes participate in cell polarization and re-orient to the site of cell-cell contact, which suggests that HIV might hijack intracellular trafficking machineries to support recruitment and polarized assembly of viral proteins close to the synaptic cleft (Bracq et al., 2018; Jolly et al., 2011).

Recruitment of retroviral proteins toward the intercellular contact site precedes the polarized assembly and release of HIV, which is directed across the synaptic cleft to the non-infected target cell (Fais et al., 1995; Jolly et al., 2007a; Jolly & Sattentau, 2004; Piguet & Sattentau, 2004). Interestingly, controversial reports exist concerning the mechanism of viral entry into a permissive target cell (Bracq et al., 2018; Puigdomènech et al., 2009). Various studies support a model of HIV entry via an endocytic pathway, comprising its internalization and maturation in endocytic compartments prior to HIV fusion with the endosomal membrane (Blanco et al., 2004; Bracq et al., 2018; Dale et al., 2011; Miyauchi et al., 2009; Sloan et al., 2013). Others propose a mechanism of HIV entry by engagement of the CD4 receptor followed by fusion with the cell plasma membrane and the direct transfer of viral components into the cytoplasm (Bracq et al., 2018; Jolly et al., 2004; Jolly & Sattentau, 2004; Martin et al., 2010). Irrespective of the pathway of virus entry, cell contact-dependent retroviral transmission is a highly efficient process with considerable implications for infection status and physiology of the target cell (Chen et al., 2007; Dimitrov et al., 1993; Jolly, 2011; Martin & Sattentau, 2009).

1.5 Consequences of cell contact-dependent retroviral transmission

Several studies indicate that cell contact-dependent transmission offers replicative advantages for retroviruses with significant consequences for viral pathogenesis (Chen et al., 2007; Doitsh et al., 2010; Jolly, 2011; Martin & Sattentau, 2009). While cell-free retroviral spread facilitates host-to-host transmission and is thought to allow for rapid virus dissemination with blood and lymph fluids to establish systemic infection, cell-free

virions experience a variety of potential obstacles, such as impermeable mucus layers, fluid phase diffusion, and the host immune response (Sattentau, 2008).

Most of these environmental barriers can be avoided by retroviruses through cell-to-cell transmission (Sattentau, 2008). During cell contact-dependent transmission, virus assembly and budding are directed toward the synaptic cleft, which generates locally a high multiplicity of infection (MOI) exceeding cell-free infections by 100- to 10,000-fold (Chen et al., 2007; Dimitrov et al., 1993; Jolly, 2011; Martin et al., 2010; Mazurov et al., 2010; Sato et al., 1992; Sourisseau et al., 2007). The likelihood of productive retroviral infection is further increased by clustering of virus entry receptors at the center of the target cell contact site (Jolly, 2011). Remarkably, the efficiency of cell contact-dependent retroviral transmission is high enough to restore infectivity of compromised viruses, which lack fitness for cell-free transmission (Bastarache et al., 2014; Brandenberg et al., 2014). Thereby, direct retroviral transmission across a virological synapse resolves the limiting step of liquid-phase retroviral diffusion (Sattentau, 2008). With regard to the humoral immune response, the structure of the synaptic cleft is hypothesized to shield budding retroviruses from exposure to a variety of neutralizing antibodies, although conclusive experimental evidence for this functional link is missing (Chen et al., 2007; Ganesh et al., 2004; Jolly, 2011; Martin et al., 2010; Massanella et al., 2009). Massive viral entry into the target cell was further shown to reduce efficacy of various antiretroviral therapies and lower efficiency of individual viral restriction factors, including TRIM5α and tetherin, both efficiently restricting cell-free retroviruses (Agosto et al., 2014; Jolly et al., 2010; Richardson et al., 2008; Sigal et al., 2011).

Both physiology and infection status of the target cell are suggested to be affected by the high viral uptake rates during cell contact-dependent transmission (Chen et al., 2007; Dimitrov et al., 1993; Jolly, 2011; Martin & Sattentau, 2009). HIV cell-to-cell transmission coincides with an increased target cell mortality as compared to cell-free transmission, which is suggested to be a result of the high viral uptake rate inducing caspase-1-mediated pyropotosis (Bracq et al., 2018; Galloway et al., 2015). In addition, cell contact-dependent transmission has been shown to correlate with an increased frequency of multicopy integration events in the genome of the target cell, both *in vitro* and *in vivo*, as compared to cell-fee transmission (Del Portillo et al., 2011; Law et al., 2016). Transmission of multiple retroviral copies is suggested to rapidly increase HIV diversification through recombination among co-transmitted viral genotypes and genetic complementation of new sequence variations (Del Portillo et al., 2011; Law et al., 2016; Levy et al., 2004). This genetic diversification potentially enables HIV to evade the host immune response and to acquire resistance to ART (Carvajal-Rodríguez et al., 2007; Levy et al., 2004; Ritchie et al., 2014). Thus, multiploid inheritance as a consequence of cell contact-dependent retroviral transmission would have significant implications for our comprehension of HIV dynamics and the development of therapeutic approaches (Del Portillo et al., 2011; Dixit & Perelson, 2005; Law et al., 2016). Yet, most of our current knowledge of retroviral inheritance and diversification has relied on the technical capabilities of established provirus detection techniques.

1.6 Technical limitations of established provirus detection techniques

There are three main challenges in the detection of integrated proviral DNA in retrovirus-infected cells (Liszewski et al., 2009). First is the distinction between integrated proviral DNA and unintegrated genetic intermediates of the retroviral replication cycle (Chun et al., 1997; Liszewski et al., 2009). A second obstacle is the low abundance of integrated proviruses *in vivo*, and finally, as a consequence of the high mutation rate of retroviruses, the diversity of the proviral population, comprising both defective and intact integrated proviruses (Achaz et al., 2004; Chun et al., 1997; Liszewski et al., 2009; Palmer et al., 2005). The number of provirus detection techniques has increased substantially over the last four decades and they are still rapidly evolving (Serrao & Engelman, 2016). Yet, each available technique addresses above-mentioned challenges to a distinct degree and should therefore be chosen with respect to the specific research question and sample availability (Falcinelli et al., 2019).

To date, quantitative polymerase chain reaction (qPCR) and digital droplet PCR (ddPCR) represent the most widely applied techniques to quantify the frequency of proviral DNA (Falcinelli et al., 2019; Kojabad et al., 2021). Both assays are rapid and highly flexible approaches, which quantify proviruses mostly by targeting conserved retroviral sequences within *gag*, the LTR, or *pol* (Chomont et al., 2009; Falcinelli et al., 2019; Malnati et al., 2008; Rutsaert et al., 2018; Strain et al., 2013). While ddPCR assays exceed sensitivity of provirus qPCR assays, both strategies are prone to overestimate the number of replication-competent proviruses, as quantification is unaffected by mutations external of the targeted proviral sequence (Bruner et al., 2015, 2016; Falcinelli et al., 2019). In addition, as qPCR and ddPCR assays regularly target sequences inside the proviral genomes, they measure the total amount of intracellular proviral DNA, and fail to distinguish integrated proviral DNA from unintegrated forms, such as linear vDNA or circular vDNA products (Bruner et al., 2015; Eriksson et al., 2013; Falcinelli et al., 2019; Martinez-Picado et al., 2018). In contrast, inverse PCR, linker ligation PCR and *Alu* PCR represent provirus detection techniques, which quantify exclusively integrated forms of proviral DNA (Chun et al., 1995, 1997; Liszewski et al., 2009; O'Doherty et al., 2002; Vandegraaff et al., 2001). Common to these strategies is the exponential amplification of a proviral/host junction region by combination of a provirus-specific primer with a host-specific primer in two PCR steps, followed by quantitative analysis using endpoint dilution (Liszewski et al., 2009). Inverse PCR and linker ligation PCR both fragmentize DNA by restriction enzyme digest, which renders these detection strategies susceptible to miss proviral sequences which contain mutations at restriction sites (Liszewski et al., 2009). *Alu* PCR assays are based on amplification of the junction region between *Alu* elements, primate-specific interspersed elements, and a conserved proviral region (Brady et al., 2013). This limits application of *Alu* PCR assays to primate-derived samples and introduces a strong location-bias to proviral detection, as *Alu* repeats are highly abundant in gene-rich regions and exhibit no random chromosomal distribution (Brady et al., 2013; Lander et al., 2001; Venter et al., 2001).

The combination of PCR- or DNA probe-based provirus enrichment techniques with Next-Generation Sequencing (NGS) platforms set new standards in the field of provirus integration site detection (Artesi et al., 2021; Miyazato et al., 2016). Various techniques

including linear amplification-mediated PCR (LAM-PCR) or HIV-1 DNA-capture-seq represent crucial tools to investigate structural diversity and preferred chromosomal target sites of proviruses (Artesi et al., 2021; Iwase et al., 2019; Katsuya et al., 2019; Miyazato et al., 2016; Serrao et al., 2016; Wang et al., 2016). However, PCR-based enrichment approaches are prone to biased amplification, whereas DNA probe-based enrichment techniques are less efficient, especially for target enrichment of long DNA fragments (Aird et al., 2011; Gilpatrick et al., 2020; Kebschull & Zador, 2015; Kozarewa et al., 2015; Miyazato et al., 2016; Van Haasteren et al., 2021). In addition, with an average read length of 200 - 400 bp, NGS-acquired reads cover only subgenomic proviral regions (Besser et al., 2018; Iwase et al., 2019; Paruzynski et al., 2010; Rosewick et al., 2020; Van Haasteren et al., 2021). Consequently, NGS-based provirus detection approaches tend to overestimate the number of replication-competent proviruses, and struggle to quantify proviral frequency in repetitive genomic regions (Asogawa et al., 2020; De Roeck et al., 2019; Van Haasteren et al., 2021; Vondrak et al., 2020).

To capture the full diversity of integrated proviruses, a variety of near full-length sequencing assays emerged, such as the full-length individual proviral sequencing (FLIPS) assay, which resolve more than 90 % of the proviral HIV genome (Bruner et al., 2016; Hiener et al., 2017; Lee et al., 2017; Wang & Palmer, 2018). Although these assays provide unprecedented insights into the diverse HIV reservoir, they reach low-confidence coverage of viral LTRs, tend to overestimate the number of replication-competent proviruses and frequently fail to locate the provirus integration site (Artesi et al., 2021; Einkauf et al., 2019; Falcinelli et al., 2019; Hiener et al., 2017; Lee et al., 2017; Wang & Palmer, 2018).

In conclusion, regardless of numerous available approaches, the simultaneous detection of location, frequency, and genomic integrity of provirus integration sites is difficult due to the technical limitations of various well-established provirus detection techniques and restricts our understanding of retrovirus integration preferences, dynamics, and evolution.

1.7 The CRISPR/Cas9 target enrichment Nanopore sequencing technology

Despite recent advancements in the detection of integrated proviruses, there remains a demand for a rapid, amplification-free provirus detection technology, which provides both, frequency and location of full-length proviruses integrated in the genome of an infected host cell. A recently developed technology for targeted sequencing of a pre-selected genomic region of interest (ROI) combines sequence specificity of the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein (Cas) 9 system with the long-read capacities of Oxford Nanopore sequencing (Gilpatrick et al., 2020; López-Girona et al., 2020).

In this workflow, called CRISPR/Cas9 target enrichment Nanopore sequencing, a selected genomic ROI is enriched through CRISPR/Cas9-mediated on-target restriction followed by selective Nanopore adapter ligation (Figure 5). In detail, genomic DNA is dephosphorylated and a genomic ROI is selectively targeted by CRISPR RNA (crRNA)/Cas9 ribonucleoprotein complexes (Figure 5A) (Gilpatrick et al., 2020; Stangl et al., 2020; Van Haasteren et al., 2021). The crRNA is complementary to a target site within the ROI and is flanked in 3’ direction on the host genome by a “NGG” protospacer

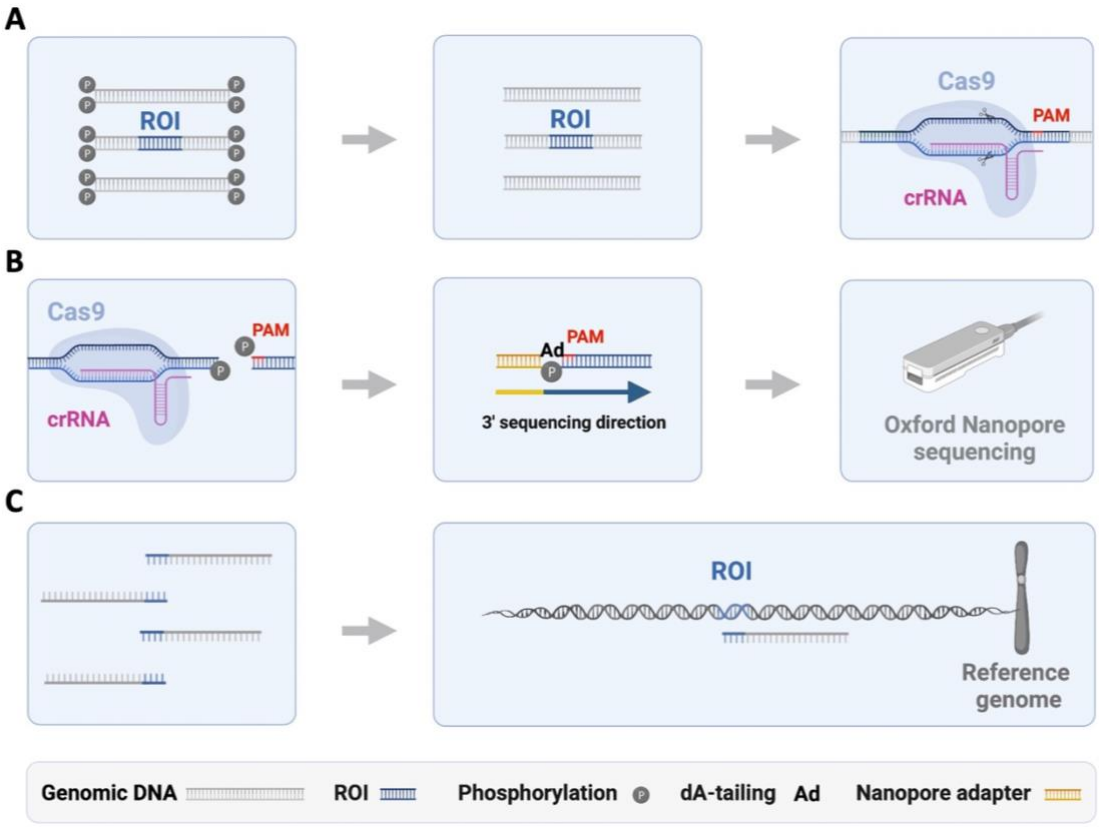

Figure 5: CRISPR/Cas9 target enrichment Nanopore sequencing technology. (A) Genomic DNA is dephosphorylated prior to selective targeting of a defined genomic region of interest (ROI) by crRNA/Cas9 ribonucleoprotein complexes (Gilpatrick et al., 2020; Stangl et al., 2020; Van Haasteren et al., 2021). CrRNAs are flanked on genomic DNA by a protospacer adjacent motif (PAM) (Stangl et al., 2020). (B) After target-specific cleavage of DNA by Cas9 endonuclease activity, phosphorylated PAM-proximal DNA fragments dissociate (Stangl et al., 2020; Van Haasteren et al., 2021). While Cas9 sterically blocks cleavage site of PAM-distal DNA fragments, PAM-proximal DNA fragments are exposed for selective dA-tailing (Gilpatrick et al., 2020; Stangl et al., 2020; Sternberg et al., 2014). Nanopore sequencing adapters are ligated to dA-tailed DNA ends, which mark the direction of Nanopore sequencing on a MinION sequencer (Stangl et al., 2020). (C) Nanopore sequencing reads are analyzed by a bioinformatic pipeline. After filtering of acquired reads for the presence of the enriched ROI using Minimap2, alignment of filtered reads against the respective host reference genome identifies genomic location of the ROI (Stangl et al., 2020; Van Haasteren et al., 2021). Illustration was created according to (Stangl et al., 2020).

adjacent motif (PAM) (Karvelis et al., 2015; Stangl et al., 2020). Hybridization of crRNAs with their complementary target sequence guides Cas9 nuclease to the ROI, where it introduces a sequence-specific double-strand cleavage upstream of the PAM (Stangl et al., 2020). After cleavage, Cas9 nuclease remains stably bound to PAM-distal DNA fragments, allowing for selective dA-tailing of dissociated, PAM-proximal DNA fragments (Figure 5B) (Gilpatrick et al., 2020; Stangl et al., 2020; Sternberg et al., 2014). Design of crRNAs in a strand-specific manner enables targeted ligation of Nanopore sequencing adapters to Cas9-cleaved, PAM-proximal DNA fragments and dictates sequencing into 5’ or 3’ direction (Stangl et al., 2020). The enriched DNA library is sequenced on a Nanopore MinION sequencer. To identify chromosomal location of the ROI by bioinformatic analysis, ROI-containing reads are first filtered from total reads using Minimap2, followed by sequence alignment of filtered reads against the host reference genome (Figure 5C) (Stangl et al., 2020; Van Haasteren et al., 2021).

Since its development, CRISPR/Cas9 target enrichment Nanopore sequencing has been widely applied in diagnostic and clinical research (Gilpatrick et al., 2020; Stangl et al., 2020; Van Haasteren et al., 2021). For instance, the amplification-free technology of CRISPR/Cas9 target enrichment Nanopore sequencing allowed for detection of cancer-related genetic mutations and identification of large heterozygous chromosomal deletions (>70 kb) (Gilpatrick et al., 2020). Most recently, CRISPR/Cas9 target enrichment Nanopore sequencing was employed to map integration sites of lentiviral vectors to provide safety evaluations of preclinical lentiviral gene therapies (Van Haasteren et al., 2021).

Its amplification-free, long-read sequencing capacities of selected genomic regions render CRISPR/Cas9 target enrichment Nanopore sequencing a promising technology for the establishment of a novel provirus detection technique, resolving the technical limitations of previous approaches.

1.8 Aims of this study

Several *in vivo* studies of retroviral infection visualized the formation of stable intercellular contacts, yet a functional contribution of cell-cell contacts to retroviral transmission has remained elusive (Law et al., 2016; Murooka et al., 2012; Sewald et al., 2012). In this study, we aim to define the role of cell adhesion-mediating proteins LFA1 and ICAM1 during retroviral spread to quantify the contribution of cell-cell contacts to retroviral dissemination *in vitro* and *in vivo.*

Our understanding of the consequences of retroviral cell-to-cell transmission for infection and physiology of the target cell is currently limited by the technical capacities of established provirus detection techniques. To simultaneously localize and quantify proviruses integrated in the genome of infected host cells, this study intends to establish CRISPR/Cas9 target enrichment Nanopore sequencing as a novel provirus detection technique. After a detailed workflow evaluation, we aim to demonstrate the capacities of our novel provirus detection technique to elucidate the consequences of retroviral transmission on a proviral level.

Precisely, the aims of this thesis were:

- To establish an MLV *in vitro cis*-infection assay with primary cells of physiological relevance
- To investigate the contribution of LFA1 and ICAM1 during cell contact-dependent MLV transmission *in vitro*
- To assess the role of LFA1 and ICAM1 during cell contact-dependent MLV transmission *in vivo*
- To establish CRISPR/Cas9 target enrichment Nanopore sequencing as a novel provirus detection technique
- To evaluate the quantitative and qualitative provirus detection capacity of CRISPR/Cas9 target enrichment Nanopore sequencing
- To verify the detection accuracy of CRISPR/Cas9 target enrichment Nanopore sequencing
- To demonstrate the flexibility of CRISPR/Cas9 target enrichment Nanopore sequencing
- To introduce CRISPR/Cas9 target enrichment Nanopore sequencing as a novel tool to analyze retroviral transmission from an infected donor to a non-infected target cell

2. Material and Methods

2.1 Material

2.1.1 Instruments

Table 1: Instruments used within this study.

Instrument	Manufacturer
Agarose gel electrophoresis chamber	Bio-Rad
BD FACSAria Fusion	BD Biosciences
BD FACSLyric	BD Biosciences
CO_2 Incubator	Binder
Eclipse Ts2-FL Diascopic and Epi-fluorescence Illumination Microscope	Nikon
Eppendorf ThermoMixer® C	Eppendorf
Eppendorf™ Centrifuge 5425R	Eppendorf
Eppendorf™ Centrifuge 5910R	Eppendorf
Eppendorf™ Mastercycler Nexus Thermal Cyclers	Eppendorf
Gel Documentation System UVP UVSolo touch	AnalytikJena
Ice machine	Manitowoc
Leica S6D stereo dissecting microscope	Leica
Mastercycler nexus X2 Thermal Cycler	Eppendorf
Microwave Inverter	Sharp
MojoSort™ magnets	Biolegend
Multitron shaking incubator	Infors HT
Nanodrop Spectrophotometer ND-1000	ThermoFisher
Pipetting aid, acu2	Integra Biosciences
PowerPacTM Basic Power Supply	Bio-Rad
QuadroMACS™ Separator	Miltenyi Biotec
QuantStudio 3 Real-Time-PCR-System	ThermoFisher
QuantusTM Fluorometer	Promega
Safe 2020 Class II Biological Safety Cabinet	Thermo Scientific™
Tube roller	StarLab
UV transilluminator	Carl Roth
Vortex mixer, Vortex-Genie II	Scientific Industries
Water bath Aqualine AL 12	Lauda
Weighing machine	Satorius

2.1.2 Chemicals and reagents

Table 2: Chemicals and reagents used within this study.

Product name	Manufacturer
Acetic acid, 37 %	Carl Roth
Agarose tablets TopVision	Thermo Scientific
Agencourt AMPure XP beads	Beckman Coulter
Albumin Fraction V, endotoxin-tested	Carl Roth
Ampicillin	Gibco™
BD FC Beads 5-Color Kit	BD Biosciences
CutSmart® Buffer (10X)	New England Biolabs
Cytometer cleaning solution	BD Biosciences
Cytometer quality control beads	BD Biosciences
Cytometer sheath fluid	BD Biosciences
D-Sucrose, ≥99.5 %	Carl Roth
dATP solution (100 mM)	New England Biolabs
Dimethyl sulfoxide (DMSO)	Carl Roth
Dipotassium hydrogen phosphate (K_2HPO_4)	Carl Roth
Dulbecco's phosphate-buffered saline (DPBS)	Gibco™ by Life Technologies
Ethanol, 99 %	Carl Roth
Ethylene diamine tetraacetic acid (EDTA)	AppliChem
Ethylene diamine tetraacetic acid (EDTA), 0.5 M	Invitrogen™
Fetal calf serum	Sigma Aldrich
Fetal calf serum (FCS), One Shot™ format	Gibco™
Fibronectin, human (0.5 mg/ml)	Advanced Biomatrix
Glycerol, ≥99.0 %	Merck
Glycine, pure	Carl Roth
HEPES (1M)	Gibco™ by Life Technologies
Isopropyl alcohol, ≥99.9 %	Carl Roth
1 kbp DNA Ladder	Metabion
MEM Non-Essential Amino Acids (100X)	Gibco™
Methanol	Carl Roth
MojoSort™ Streptavidin Nanobeads	Biolegend
Normal rat serum	Jackson ImmunoResearch
Nuclease-free duplex buffer	Integrated DNA Technologies
Nuclease-free TE (1X), pH 7.5	Integrated DNA Technologies
Nuclease-free water	Integrated DNA Technologies
Opti-MEM™ Reduced serum medium	Gibco™
Orange DNA Loading Dye (6X)	Thermo Scientific™
Paraformaldehyde (PFA)	Electron Microscopy Sciences
Polyethylenimine (PEI)	Polysciences
Potassium dihydrogen phosphate (KH_2PO_4)	Carl Roth

Product name	Manufacturer
Retinoic acid, powder	Sigma Aldrich
ROTIPHORESE® TAE-buffer (50X)	ROTIPHORESE®
RPMI-1640, GlutaMAX™	Gibco™ by Life Technologies
Sodium carboxymethylcellulose	Sigma-Aldrich
Sodium pyruvate (100 mM)	Gibco™
SYBR™ Safe DNA gel stain	Invitrogen™
Terrific Broth medium, powder	CHEMSOLUTE®
Tris(hydroxymethyl)-aminomethane (Tris)	AppliChem
Trypan Blue Solution, 0.4 %	Carl Roth
Tryptone	Carl Roth
Ultrapure lipopolysaccharide (LPS), *E. coli* K12 (EK)	InvivoGen
UltraPure™ DNase/RNase-free distilled water	Thermo Fisher Scientific
Yeast extract	Carl Roth
β-mercaptoethanol	Gibco™

2.1.3 Enzymes

Table 3: Enzymes used within this study.

Enzyme	Manufacturer
Accutase® cell detachment solution	Biolegend
Alt-R® *S. pyogenes* HiFi Cas9 nuclease V3	Integrated DNA Technologies
DNase I recombinant, RNase-free	Roche
Liberase TL Research Grade	Roche
Proteinase K (> 600 mAU/ml)	Qiagen
Quick calf intestinal alkaline phosphatase (CIP)	New England Biolabs
RNase A (100 mg/ml)	Qiagen
T4 DNA ligase	New England Biolabs
Taq polymerase	New England Biolabs
Trypsin-EDTA (1X)	Gibco™

2.1.4 Cytokines

Table 4: Cytokines used within this study.

Cytokine	Manufacturer
Recombinant human IL-2 (carrier-free)	Biolegend
Recombinant human TGF-β1 (carrier-free)	Biolegend
Recombinant mouse IL-15 (carrier-free)	Biolegend
Recombinant mouse IL-4 (carrier-free)	Biolegend
Recombinant mouse IL-5 (carrier-free)	Biolegend

Cytokine	Manufacturer
Recombinant mouse IL-6 (carrier-free)	Biolegend
Recombinant mouse IL-7 (carrier-free)	Biolegend

2.1.5 Antibodies

Table 5: Primary antibodies used within this study. Dilutions applied for immunofluorescence staining prior flow cytometry are enlisted for the respective antibodies.

Antibody name	Clone	Flow cytometry	Manufacturer
Alexa Fluor® 647 anti-mouse FoxP3 antibody	MF-14	1:20	Biolegend
APC anti-mouse CD11a antibody	M17/4	1:500	Biolegend
APC anti-mouse CD19 antibody	6D5	1:1,000	Biolegend
APC anti-mouse CD4 antibody	GK1.5	1:2,000	Biolegend
APC anti-mouse CD54 antibody	YN1/1.7.4	1:500	Biolegend
Biotin anti-mouse CD115 (CSF-1R) antibody	AFS98	-	Biolegend
Biotin anti-mouse CD117 (c-Kit) antibody	2B8	-	Biolegend
Biotin anti-mouse CD11c antibody	N418	-	Biolegend
Biotin anti-mouse CD23 antibody	B3B4	-	Biolegend
Biotin anti-mouse CD4 antibody	RM4-4	-	Biolegend
Biotin anti-mouse CD88 antibody	20/70	-	Biolegend
Biotin anti-mouse CD8a antibody	53-6.7	-	Biolegend
Biotin anti-mouse F4/80 antibody	BM8	-	Biolegend
Biotin anti-mouse Fcε antibody	MAR-1	-	Biolegend
Biotin anti-mouse Ly-6G/Ly-6C (Gr-1) antibody	RB6-8C5	-	Biolegend
Biotin anti-mouse NK-1.1 antibody	PK136	-	Biolegend
Biotin anti-mouse TER-119 antibody	TER-119	-	Biolegend
PE anti-mouse CD3 clone antibody	17A2	1:500	Biolegend
PE anti-mouse F4/80 antibody	BM8	1:1,000	Biolegend
Purified anti-mouse CD16/CD32 antibody	93	-	Biolegend
Ultra-LEAF™ Purified anti-mouse CD11a antibody	M17/4	-	Biolegend
Ultra-LEAF™ Purified anti-mouse CD28 antibody	35.51	-	Biolegend
Ultra-LEAF™ Purified anti-mouse CD3ε antibody	145-2C11	-	Biolegend
Ultra-LEAF™ Purified anti-mouse CD54 antibody	YN1/1.7.4	-	Biolegend

Antibody name	Clone	Flow cytometry	Manufacturer
Ultra-LEAF™ Purified Rat IgG2a,κ Isotype Ctrl antibody	RTK2758	-	Biolegend
Ultra-LEAF™ Purified Rat IgG2b,κ Isotype Ctrl antibody	RTK4530	-	Biolegend

2.1.6 Kits

Table 6: Kits used within this study.

Kit name	Manufacturer
2X Luna Universal Probe qPCR Master Mix	New England Biolabs
Fixation/Permeabilization Kit	BD Biosciences
Flow Cell Priming Kit (EXP-FLP001)	Oxford Nanopore
Ligation Sequencing Kit (SQK-LSK109)	Oxford Nanopore
Monarch HMW DNA Kit	New England Biolabs
Naïve CD4+ T Cell Isolation Kit, mouse	Miltenyi Biotec
NuceloBond Xtra Midi Kit for transfection-grade plasmid DNA	MACHEREY-NAGAL
Phusion® High-Fidelity PCR Master Mix with HF Buffer	New England Biolabs
Qiagen DNeasy Blood & Tissue Kit	Qiagen
QuantiFluor® dsDNA System Kit	Promega
Wizard® SV Gel and PCR Clean-Up System	Promega

2.1.7 Primers and TaqMan probes

Table 7: Primers used within this study.

Primer name	Target	Primer sequence (5'-3')	Manufacturer
GFP_1_Fwd	GFP	CCACATGAAGCAGCACGACTT	Ella Biotech
GFP_1_Rev	GFP	GGTGCGCTCCTGGACGTA	Ella Biotech
GFP_4_Fwd	GFP	GCTGGAGTACAACTACAAC	Ella Biotech
GFP_4_Rev	GFP	TGGCGGATCTTGAAGTTC	Ella Biotech
E10-o2A2_1_Rev	Chr1:155,077,367-155,077,384	TTCTGGCACTGGTGTGAC	Ella Biotech
E10-o2A2_13_Fwd	Chr13:48,997,521-48,997,539	CTTACTACACAGAAGCACG	Ella Biotech

Table 8: TaqMan probes used within this study.

TaqMan probe	Target	TaqMan Probe sequence (5'-3')	Manufacturer
GFP_1_probe	GFP	FAM-TTCAAGTCCGCCATGCCCGAA-BHQ1	Ella Biotech
GFP_4_probe	GFP	FAM-CTTGATGCCGTTCTTCTGCTTGTC-BHQ1	Ella Biotech

2.1.8 CrRNAs and tracrRNA

Table 9: CrRNAs used within this study.

CrRNA Name	CrRNA sequence (5'-3')	Manufacturer
S. pyogenes Cas9 Alt-R® crRNA GFP-D5	GATGCCGTTCTTCTGCTTGT	Integrated DNA Technologies
S. pyogenes Cas9 Alt-R® crRNA GFP-D6	CAAGATCCGCCACAACATCG	Integrated DNA Technologies
S. pyogenes Cas9 Alt-R® crRNA GFP-D7	GCTGAAGCACTGCACGCCGT	Integrated DNA Technologies
Alt-R® CRISPR *Hprt1* Positive Control crRNA	ACCTCTTAGGAGTCTAAAGT	Integrated DNA Technologies

Table 10: TracrRNA used within this study.

TracrRNA Name	Manufacturer
Alt-R® CRISPR-Cas9 tracrRNA	Integrated DNA Technologies

2.1.9 Animal models

C57BL/6 mice were provided by Charles River Laboratory. CD11a-deficient [B6.129S7-Itgaltm1Bll/J], ICAM1-deficient mice [B6.129S4-Icam1tm1Jcgr/J] and RFP+ mice [Tg(CAG-DsRed*MST)1Nagy/J], which express in all nucleated cells cytoplasmic DsRed, were received from Jackson Laboratory. All mouse experiments were conducted according to the legal requirements and approved by the local authorities. Within this study, both male and female mice, between 6- to 12-weeks old, were used.

2.1.10 Strains and plasmids

2.1.10.1 Bacterial strains

Table 11: Bacterial strains used within this study.

Bacterial strain	Plasmid	Resistance
MAX Efficiency *E. coli* Stbl2™	Friend57 MLV WT, full length	Ampicillin
MAX Efficiency *E. coli* Stbl2™	LTR-GFP	Ampicillin
MAX Efficiency *E. coli* Stbl2™	Friend57 MLV Env	Ampicillin
MAX Efficiency *E. coli* Stbl2™	Friend 57 MLV IRES GFP, full length	Ampicillin
MAX Efficiency *E. coli* Stbl2™	MLV Gag-Pol	Ampicillin

2.1.10.2 Cell lines

Table 12: Cell lines used within this study.

Cell line	Species	Description
HEK293 (gold)	Human	Embryonic human kidney cell line
S49.1	Mouse (BALB/c)	Murine T lymphoid cell line
J-Lat, clone 10.6	Human	Subclone of Jurkat cells infected with HIV-1

2.1.10.3 Plasmids

Table 13: Plasmids used within this study.

Plasmid	Plasmid name	Resistance
Friend 57 MLV WT, full length	pLRB303-FrMLV	Ampicillin
LTR-GFP	pMMP-LTR-GFP	Ampicillin
Ecotropic MLV Env	pLZRS-FrMLV Env	Ampicillin
Friend 57 MLV IRES GFP, full length	pLRB303-FrMLV-IRES-GFP	Ampicillin
MLV Gag-Pol	pHIT60	Ampicillin

2.1.11 Bacterial culture medium

Table 14: Ingredients of bacterial culture medium used within this study.

Medium and buffer	Ingredients
Terrific broth (TB) medium	16 g tryptone, 24 g yeast extract, 5 g glycerol, ad 0.9 l A. dest.
Potassium buffer	12.5 g K_2HPO_4, 2.3 g KH_2PO_4, ad 0.1 l A. dest.

Escherichia coli Stbl2™ bacterial strains were cultured in TB medium, supplemented with 1:10 Potassium buffer and 1:1,000 Ampicillin (100 mg/ml).

2.1.12 Culture media for cell lines and primary cells

2.1.12.1 Cultivation medium

All cultures in this study, except for HEK293 cells, were performed in RPMI-1640 GlutaMAX™ medium, supplemented with 10 % heat-inactivated FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, and 55 µM β-mercaptoethanol. HEK293 cells were cultured in RPMI-1640 GlutaMAX™ medium, supplemented with 10 % heat-inactivated FCS.

2.1.12.2 Infection medium

Infection of cell lines and primary cells was carried out in RPMI-1640 GlutaMAX™ medium, supplemented with 10 % heat-inactivated FCS and 10 mM HEPES.

2.1.13 Consumables

Table 15: Consumables used within this study.

Material	Manufacturer
0.2 ml thin-walled PCR tubes	Applied Biosystems
3-piece Omnifix® Luer Solo syringe	B. Braun
Aspiration pipette, 2 ml	Greiner Bio
Autoclave tape	Hartenstein
Cell strainers; 70 µm	Croning
Cellstar® serological pipettes; 5 ml, 10 ml, 25 ml	Greiner
Corning® HTS Transwell® 96 well permeable supports; 3 µm	Corning
CryoPure tubes; 2 ml	Sarstedt
DNAase/RNase-free Eppendorf tubes; 1.5 ml	Eppendorf
Eppendorf DNA LoBind tubes; 1.5 ml	Eppendorf
Eppendorf safe-lock tubes; 1.5 ml, 2 ml, 5 ml	Eppendorf
Falcon tubes; 15 ml, 50 ml	Corning
Falcon™ Round-bottom Polystyrene Tubes with Cell Strainer Snap Cap; 5 ml	Corning
Falcon™ standard tissue culture dishes; 3.5 cm, 10 cm	Falcon™
FLO-MIN106 flow cell	Oxford Nanopore
Flongle	Oxford Nanopore
LS columns	Miltenyi Biotech
Millex™ sterile filter units; 0.22 µm, 0.45 µm	Merck
Neubauer counting chambers	Carl Roth
Parafilm® M	Sigma Aldrich
Pipette filter tips; 10 µl, 20 µl, 200 µl, 1000 µl	Starlab
Polystyrene tubes	Applied Biosystems
Sterican needles; 21 gauge, 26 gauge, 27 gauge	B. Braun
Sterile syringes; 10 ml	Fisher Scientific
Tissue culture plates; 48-well, 24-well, 12-well, 6-well	Sarstedt
Tissue culture plates; 96-well flat bottom, round bottom	Sarstedt

2.1.14 Computational data analysis

2.1.14.1 Oligonucleotide design and evaluation

GFP-targeting crRNAs applied in this study were designed and evaluated using Custom Alt-R® CRISPR-Cas9 guide RNA software tool from Integrated DNA Technologies.

2.1.14.2 Flow cytometry data analysis

Flow cytometry data, collected on a BD FACSLyric or BD FACSAria Fusion flow cytometer, were analyzed using FlowJo software (Version 10.7.1, Treestar).

2.1.14.3 CRISPR/Cas9 target enrichment Nanopore sequencing data analysis

For analysis of sequencing results after CRISPR/Cas9 target enrichment Nanopore sequencing, reads were aligned against the corresponding proviral genome using Minimap2. Filtered reads were mapped to the mouse (GRCm38) or human (hg38) reference genome, and visualized using Integrative Genomics Viewer (IGV), Version 2.9.0 provided by Broad Institute and the University of California.

2.1.14.4 Quantitative PCR data analysis

Quantitative PCR data, acquired using QuantStudio 3 Real-Time-PCR-System, were analyzed using QuantStudio Design and Analysis software, Version 2.5.0 by ThermoFisher. For generation of a standard curve, a range of nine serial dilutions (10-fold each) of plasmid DNA containing the target DNA sequence, were amplified.

2.1.14.5 Figures

Figures were edited using Adobe Photoshop Version 6.0 and Illustrator CC. Schematic visualizations were designed using BioRender.com.

2.1.14.6 Statistical analysis

Under the presumption that the samples did not follow a Gaussian distribution, the corresponding statistical analyses were conducted using GraphPad Prism 9.2 software. For comparison of two groups, a non-parametric Mann-Whitney test (two-tailed) was applied. In each figure the exact P values are included. If $P < 0.05$ (two-tailed), values were interpreted as statistically significant. The figure legend of each experiment specifies the number of independent replicates.

2.2 Methods

2.2.1 Molecular biological methods

2.2.1.1 Polymerase chain reaction

Polymerase chain reaction is a sensitive technique for the amplification of a selected DNA region. Amplification reactions were performed with Phusion DNA polymerase according to the manufacturer's protocol (see Table 7 and Table 16). PCR reaction was conducted in the Mastercycler Nexus X2 Thermal Cycler at the following temperature steps: after an initial DNA denaturation for 30 s at 98 °C, 35 cycles, comprising denaturation for 30 s at 98 °C, primer annealing for 30 s at 56 °C and elongation for 90 s at 72 °C, were performed. Following a final extension step for 10 min at 72 °C, PCR products were loaded on an agarose gel and DNA fragments were analyzed after agarose gel electrophoresis (described in 2.2.1.2).

Table 16: Reaction components for DNA amplification using Phusion DNA Polymerase.

Component	For 25 µl PCR reaction	Final concentration
2X Phusion Master Mix	12.5 µl	1X
Forward primer (10 µM)	1.25 µl	0.5 µM
Reverse primer (10 µM)	1.25 µl	0.5 µM
Template DNA	variable	10 ng
Nuclease-free water	to 25 µl	-

2.2.1.2 Agarose gel electrophoresis

50X TAE buffer: 2 M Tris, 1 M acetic acid, 50 mM EDTA in A. dest., pH 8.5;

Agarose gel electrophoresis and purification of DNA fragments from agarose gels were performed according to standard protocols. Shortly, agarose gels were prepared from agarose tablets in 1X TAE buffer supplemented with 1X SYBR™ Safe DNA Gel Stain at a concentration range between 0.7 % to 1.0 %. Approximately 20 µl of DNA sample, mixed with 1X Orange DNA Loading Dye, were loaded onto agarose gels, and covered with 1X TAE running buffer. Along with the DNA samples, 5 µl of 1 kbp DNA Ladder were loaded onto the agarose gel. DNA fragments were separated by applying 90 V for 1 h using a PowerPacTM Basic Power Supply. DNA fragments were analyzed by UV exposure for 2 s using the Gel Documentation System UVP UVSolo touch. Selected DNA fragments were extracted from agarose gels using the Promega Kit Wizard® SV Gel and PCR Clean-Up System according to the manufacturer's protocol.

2.2.1.3 Quantitative PCR

Quantitative PCR is a technique which applies PCR-based amplification of a specific DNA target sequence to measure its absolute or relative amount in the reaction. Therefore, 10 ng of template DNA were mixed with a specific primer pair, TaqMan probe, 2X Luna Universal Probe qPCR Master Mix and nuclease-free water, reaching a total volume of 20 µl (see Table 7, Table 8 and Table 17). QPCR assays were performed

on the QuantStudio 3 Real-Time-PCR-System using the Comparative C_t value program. In detail, after initial denaturation of template DNA for 1 min at 95 °C, within 45 cycles DNA was sequentially denatured for 15 s at 95 °C, and primers were extended for 30 s at 60 °C. For absolute quantification, a standard curve was prepared using 10-fold serial dilutions of plasmid DNA harboring the target DNA sequence (see Table 13). Data analysis was performed using the QuantStudio Design and Analysis software, Version 2.5.0 by ThermoFisher.

Table 17: Reaction components for quantitative PCR using Luna Universal Probe Master Mix.

Component	For 20 µl qPCR reaction	Final concentration
2X Luna Universal Probe qPCR Master Mix	10 µl	1X
Forward primer (10 µM)	0.8 µl	0.4 µM
Reverse primer (10 µM)	0.8 µl	0.4 µM
TaqMan probe (10 µM)	0.4 µl	0.2 µM
Template DNA	variable	10 ng
Nuclease-free water	to 20 µl	-

2.2.1.4 CRISPR/Cas9 target enrichment Nanopore sequencing

CRISPR/Cas9 target enrichment was performed according to Oxford Nanopore protocol (Version ENR_9084_v109_revA_04Dec2018). In detail, *S. pyogenes* Cas9 Alt-R® crRNAs and Cas9 Alt-R® tracrRNA were resuspended in nuclease-free TE (pH 7.5) to a final concentration of 100 µM (see Table 9 and Table 10). CrRNAs were equimolarly pooled and 1 µl of crRNA pool was mixed with 1 µl tracrRNA in 8 µl nuclease-free duplex buffer, followed by incubation at 95 °C for 5 min in a thermocycler. For duplex formation, reaction mix was left to cool to room temperature (RT) for 30 min. Subsequently, 10 µl of crRNA-tracrRNA duplex mix were combined with 0.8 µl Alt-R® *S. pyogenes* HiFi Cas9 nuclease V3, 79.2 µl H_2O and 10 µl CutSmart Buffer (10X). CrRNA/Cas9 complex formation was carried out at RT for 30 min. For dephosphorylation of genomic DNA, 3 µg of genomic DNA was mixed with 10 µl CutSmart buffer (10X) and 1 µl CIP Enzyme, followed by incubation at 37 °C for 10 min and subsequent enzyme inactivation for 5 min at 72 °C in a thermocycler. For Cas9-mediated *in vitro* restriction and dA-tailing, dephosphorylated DNA was mixed with 10 µl crRNA/Cas9 complexes, 1 µl ATPs (10 µM) and Taq polymerase, followed by incubation for 30 min at 30 °C and 80 °C for 5 min. Adapter ligation and Nanopore sequencing was performed by UMC Utrecht Sequencing Facility according to Oxford Nanopore protocol (Version ENR_9084_v109_revA_04Dec2018). Shortly, the DNA sample was mixed with 20 µl of adapter ligation mix from the Ligation Sequencing Kit (SQK-LSK109, Oxford Nanopore), and subsequently incubated for 10 min at RT. For sample purification, 0.3X volume of AMPure XP Beads were added to the ligation sample, followed by incubation for 10 min at RT. DNA sample was placed in a magnetic rack, to allow 2X washing with long fragment buffer (SQK-LSK109, Oxford Nanopore). DNA fragments were eluted in 12 µl elution buffer. Samples were sequenced on a Flongle

or a MinION device equipped with a FLO-MIN106 flow cell (Oxford Nanopore) for a duration of 24 h.

2.2.1.5 Plasmid DNA extraction

TB medium: 1.6 % tryptone, 2.4 % yeast extract, 0.5 % glycerol, 0.9 ml A. dest; Potassium buffer: 0.72 M K_2HPO_4, 0.17 M KH_2KO_4, 0.1 ml A. dest.;

For extraction of plasmid DNA, transformed *E. coli* stable II (see Table 11) were grown overnight (O/N) at 32 °C in a shaking culture system at 180 rpm in 12 ml TB medium supplemented with 1:10 Potassium buffer and Ampicillin (100 µg/ml). First O/N culture was transferred into 225 ml TB medium supplemented with 1:10 Potassium buffer and Ampicillin (100 µg/ml) and grown O/N at 32 °C and 180 rpm. Plasmid DNA was extracted using the NuceloBond Xtra Midi Kit for transfection-grade plasmid DNA by MACHEREY-NAGEL according to the manufacturer's instructions. Concentration of eluted DNA was measured on a Nanodrop Spectrophotometer ND-1000.

2.2.1.6 Cryopreservation of transformed bacteria

TB medium: 1.6 % tryptone, 2.4 % yeast extract, 0.5 % glycerol, 0.9 ml A. dest; Potassium buffer: 0.72 M K_2HPO_4, 0.17 M KH_2KO_4, 0.1 ml A. dest.;

For generation of bacterial glycerol stocks, transformed *E.coli* stable II (see Table 11) were grown O/N at 32 °C in a shaking culture system at 180 rpm in 12 ml TB medium supplemented with 1:10 Potassium buffer and Ampicillin (100 µg/ml). Afterwards, 750 µl of transformed *E. coli* O/N culture were added to 250 µl 50 % glycerol and transferred into a 2 ml cryotube for storage at - 80 °C.

2.2.1.7 Sequencing

Plasmids applied within this study (see Table 13) were analyzed by Sanger sequencing (Eurofins) and Oxford Nanopore sequencing (AG Blum, LMU München). Amplified DNA products extracted after DNA gel electrophoresis were sequenced by Sanger Sequencing (Eurofins). For analysis of sequencing data, CodonCode Aligner, Version 9.0.2 and NCBI BLAST were used.

2.2.2 Cell biological methods

2.2.2.1 Cell cultivation

Primary cells and cell lines were cultured in cell culture medium, if not stated otherwise, at 37 °C and 5 % CO_2. Cell lines were splitted prior to reaching 100 % confluence. Therefore, old culture medium was removed, and cells were centrifuged for 5 min at 400 x g and RT. After resuspension in fresh culture medium, cells were seeded at a ratio of 1:2 and incubated at 37 °C.

2.2.2.2 Freezing and thawing of cells

Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

For freezing of cell lines, cells were harvested and centrifuged for 5 min at 400 x g and RT. After resuspension in fresh culture medium supplemented with 10 % DMSO, cells were aliquoted into freezing vials and stored for 48 h in an isopropanol container at -80°C, prior to storage in liquid nitrogen.
Cells were thawed by incubation for 2 min in a 37 °C water bath, followed by transfer into 2 ml culture medium. DMSO-containing medium was removed by centrifugation for 5 min at 400 x g and RT, and cells were resuspended in fresh culture medium. After 24 h of cultivation, dead cells were removed by centrifugation for 8 min at 60 x g at RT and seeded in fresh culture medium for cultivation at 37 °C.

2.2.2.3 Cell counting

Cell counting was performed using a Neubauer counting chamber according to manufacturer's instructions. Shortly, cells were mixed with Trypan Blue at a ratio corresponding to the desired end dilution. To determine the number of cells in suspension, 10 µl of the cell suspension was introduced into the Neubauer chamber. The number of cells was assessed by counting of the four 4x4 corner squares under a bright field microscope. Cell concentration was determined according to the following equation:

$$\text{concentration (cells/ml)} = \frac{Number\ of\ cells\ counted}{Number\ of\ corner\ squares}\ \text{x}\ 10^4\ \text{x dilution factor x ml}^{-1}$$

2.2.2.4 Isolation of primary peritoneal B1 cells

MACS buffer: 0.5 % BSA, 2 mM EDTA in PBS; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

Naïve mouse B1 cells were enriched from peritoneal washes of C57BL/6 mice, CD11a-/-, ICAM1-/- and RFP+ mice. Therefore, 5 ml MACS buffer were injected into the peritoneal cavity of euthanized mice using a 26 gauge needle. Peritoneal wash was extracted using a 21 gauge needle, and centrifuged for 8 min at 400 x g and 12 °C. After resuspension in 400 µl MACS buffer, cell suspension was transferred into a polystrene tube and 20 µl of a mix of biotinylated monoclonal antibodies, targeting CD4, CD8, CD23, F4/80, CD115, CD117, TER-119, Gr-1, NK1.1, CD11c, CD23, FcεRIα and CD88 (see Table 18), were added to the cellular suspension. After incubation for 12 min on ice, 25 µl of magnetic streptavidin Nanobeads were added to the suspension followed by 5 min incubation on ice. Cell suspension was resuspended in 2 ml MACS buffer and transferred into a MojoSort™ magnet for 3 min. Enriched cells were poured into 6 ml cold cell culture medium and centrifuged for 10 min at 400 x g and 12 °C. Cells were resuspended in 1 ml culture medium, and cell number was determined (described in 2.2.2.3). Cell activation

(described in 2.2.2.6) was initiated in a 96-well flat bottom plate at a seeding concentration of 4 x 10^5 naïve B1 cells per well by incubation for 24 h at 37 °C.

Table 18: Volumes of biotin-labeled antibodies combined for negative isolation of B1 cells from the peritoneal wash of a single mouse.

Antibody	Volume [µl]
Biotin anti-mouse F4/80 antibody	7.0
Biotin anti-mouse CD8a antibody	1.0
Biotin anti-mouse CD4 antibody	0.75
Biotin anti-mouse Ly-6G/Ly-6C (Gr-1) antibody	0.75
Biotin anti-mouse TER-119 antibody	1.0
Biotin anti-mouse CD11c antibody	2.0
Biotin anti-mouse NK-1.1 antibody	2.0
Biotin anti-mouse CD23 antibody	0.5
Biotin anti-mouse Fcε antibody	1.0
Biotin anti-mouse CD88 antibody	1.0
Biotin anti-mouse CD117 (c-Kit) antibody	1.0
Biotin anti-mouse CD115 (CSF-1R) antibody	2.0

2.2.2.5 Isolation of primary naïve CD4+ T cells

MACS buffer: 0.5 % BSA, 2 mM EDTA in PBS; Washing buffer: 2 mM EDTA in PBS; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

Splenocytes derived from C57BL/6, CD11a-/- and ICAM1-/- mice were used for isolation of naïve CD4+ T cells. Therefore, excised spleen was dissected into small fragments and grinded in RPMI-1640 GlutaMAX™ supplemented with 10 % heat-inactivated FCS through a 70 µm strainer. Cell suspension was centrifuged for 12 min at 400 x g and 12 °C and washed once with PBS/2 mM EDTA. Cells were resuspended in 10 ml MACS buffer and counted at a 1:50 dilution in Trypan Blue (described in 2.2.2.3). A total of 1 x 10^8 cells were resuspended in 400 µl MACS buffer as preparation for negative isolation using the Miltenyi Naïve CD4+ T cell isolation Kit, mouse. In detail, 100 µl of biotin antibody cocktail was added to the cell suspension, followed by incubation for 5 min on ice. Sequentially, 200 µl MACS buffer, 200 µl anti-biotin microbeads and 100 µl CD44 microbeads were added to the suspension. After incubation for 10 min on ice, 1 ml MACS buffer was added, and cells were centrifuged for 10 min at 300 x g and 12 °C. After resuspending the cellular pellet in 600 µl MACS buffer, cell suspension was introduced into an equilibrated LS magnetic column, and the collected flow through was centrifuged for 10 min at 400 x g and 12 °C. Isolated naïve CD4+ T cells were resuspended in fresh culture medium, followed by quantification at 1:10 dilution in Trypan Blue (described in 2.2.2.3). Cell activation and differentiation into FoxP3+ T cells (described in 2.2.2.6) was initiated at a seeding concentration of 1.5 x 10^5 cells/well in a 96-well flat bottom plate over 48 h at 37 °C.

2.2.2.6 Primary cell activation and differentiation

Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

For cell activation, enriched B1 cells were cultured in the presence of LPS (2.5 µg/ml), mouse IL-4 (100 ng/ml), mouse IL-5 (100 ng/ml) and mouse IL-6 (100 ng/ml) for 24 h at 37 °C (see Table 2 and Table 4). Isolated naïve CD4+ T cells were activated by co-stimulation with anti-CD3ε/CD28 surface-coating (1 µg/100 µl) and differentiation to FoxP3+ T cells was initiated by cultivation in culture medium supplemented with human IL-2 (20 ng/ml), mouse IL-7 (100 ng/ml), mouse IL-15 (100 ng/ml), retinoic acid (10 nM) and human TGF-β1 (5 ng/ml) for 48 h at 37 °C (see Table 2 and Table 4).

2.2.2.7 DNA extraction from cultured cells

Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

Genomic DNA was isolated from cultured cells using the Qiagen DNeasy Blood & Tissue Kit according to the manufacturer's instructions. In detail, 5 x 10^6 cells were resuspended in 200 µl PBS and incubated with 20 µl Proteinase K and 4 µl RNase A for 2 min at RT. Cells were lysed by adding 200 µl AL buffer and incubation for 10 min at 56 °C. After addition of 200 µl 99 % ethanol, cellular lysate was transferred into a DNeasy Mini spin column and centrifuged for 1 min at 6,000 x g. After two washing steps with 500 µl wash buffers AW1 and AW2, respectively, DNA was eluted from the membrane by applying 100 µl elution buffer followed by centrifugation for 1 min at 6,000 x g. DNA concentration of the collected eluate was determined using the QuantiFluor® dsDNA System Kit according to the manufacturer's instructions and measured on a Promega Quantus Fluorometer.

2.2.3 Microbiological methods

2.2.3.1 Virus production

Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES;

Replication-competent MLV LTR-GFP, replication-competent MLV IRES GFP, and single-round virus like particles (VLPs) harboring the LTR-GFP genome were generated by co-transfecting HEK293 cells. Therefore, HEK293 cells were cultured to reach 70 - 80 % confluence in a 10 cm cell culture dish. Old culture medium was removed, and fresh culture medium was carefully transferred onto the cellular layer, followed by incubation at 37 °C. A total of 12 µg plasmid DNA was mixed with 550 µl OptiMEM and 36 µl PEI, followed by incubation for 30 min at RT (see Table 19, Table 20, Table 21). Reaction mix was added to HEK293 cells and incubated for a total of 48 h at 37 °C. After 24 h and 48 h, cell culture supernatant was harvested by transfer through a 0.45 µm nylon membrane filter, aliquoted in 2 ml Eppendorf tubes and stored at - 80 °C.

Table 19: Co-transfected plasmids for generation of MLV LTR-GFP particles.

Plasmid name	Amount
Friend 57 MLV WT, full length	10 µg
LTR-GFP	1 µg
Ecotropic MLV Env	1 µg

Table 20: Co-transfected plasmids for generation of MLV IRES GFP particles.

Plasmid name	Amount
Friend 57 MLV IRES GFP, full length	11 µg
Ecotropic MLV Env	1 µg

Table 21: Co-transfected plasmids for generation of virus-like particles (VLPs) harboring LTR-GFP genome.

Plasmid name	Amount
MLV Gag-Pol	10 µg
LTR-GFP	1 µg
Ecotropic MLV Env	1 µg

2.2.3.2 Virus concentration

Sucrose gradient: 15 % sucrose in PBS; Infection medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES;

Viruses were concentrated from pooled (24 h and 48 h) cell culture supernatants post-transfection (described in 2.2.3.1). A cushion of 130 µl 15 % sucrose in PBS was used for sedimentation of pooled virus supernatants at 20,000 x g and 4 °C for 2 h. Supernatant was discarded, and virus pellets were resuspended in a total of 50 µl infection medium for 45 min at 4 °C. After removal of debris by centrifugation for 5 min at 5,200 x g and 4 °C, concentrated virus was collected in a single 1.5 ml Eppendorf tube and stored at 4 °C until downstream applications.

2.2.3.3 Virus titration on S49.1 cells

Infection medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

To determine viral titers, concentrated virus stock was titrated in serial dilutions on murine S49.1 cells. In detail, 2 x 10^5 S49.1 cells were seeded in a flat bottom 96-well plate in a total volume of 50 µl cell culture medium per well. After virus concentration (described in 2.2.3.2), S49.1 cells were inoculated with 0.1 µl, 0.3 µl, 1 µl, 3 µl and 10 µl virus suspension, followed by incubation for 24 h at 37 °C. After 18 h, 100 µl fresh culture

medium was added to each well. Corresponding virus titer in infectious units (I.U.) per µl was assessed by measuring GFP expression of infected S49.1 cells using flow cytometry.

2.2.3.4 Generation of GFP-expressing cell clones

Culture medium with 1 % methylcellulose: RPMI-1640 GlutaMAX™, 1 % sodium carboxymethylcellulose, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol; Culture medium without methylcellulose: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

To yield discrete differences in the number of integration sites, S49.1 cells were transduced in the presence or absence of methylcellulose with varying titers of GFP-reporter virus particles. In detail, 2 x 10^5 S49.1 cells were seeded in a 96-well flat bottom plate and cultivated in cell culture medium with or without 1 % methylcellulose (see Table 2). Cells were inoculated with single-round VLPs or MLV IRES GFP suspension (described in 2.2.3.1 and 2.2.3.2). After 24 h, cells were seeded by limiting dilution at 0.5 cells/100 µl in a 96-well round bottom well plate in cell culture medium without methylcellulose, and single, GFP-expressing cells were identified by visual screening using the Eclipse Ts2-FL Diascopic and Epi-fluorescence Illumination Microscope. Exclusively single, GFP-expressing cells were expanded over a duration of 6 weeks at 37 °C to a 6-well plate format and characterized by flow cytometry using a BD FACSLyric for differences in GFP expression. For long-term storage, expanded individual cell clones were frozen in liquid nitrogen (described in 2.2.2.2).

2.2.3.5 *In vitro* transduction of FoxP3+ T cells and B1 cells

Infection medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

LPS-activated B1 cells and *in vitro* differentiated FoxP3+ T cells were infected with MLV LTR-GFP using spin infection. Therefore, cells were seeded at 4 x 10^5 cells/well in a 96-well flat bottom plate and inoculated with concentrated MLV LTR-GFP suspension (described in 2.2.3.1 and 2.2.3.2) at 2 x 10^5 I.U./well. Cells were spin infected by centrifugation for 1:30 h at 1,100 x g at 37 °C. Prior to cultivation for 24 h at 37 °C, 100 µl cell culture medium, supplemented with LPS (2.5 µg/ml), mouse IL-4 (100 ng/ml), mouse IL-5 (100 ng/ml) and mouse IL-6 (100 ng/ml) for B1 cell cultivation or human IL-2 (20 ng/ml), mouse IL-7 (100 ng/ml), mouse IL-15 (100 ng/ml), retinoic acid (10 nM) and human TGF-β1 (5 ng/ml) for FoxP3+ T cell cultivation, were added to each well.
Without spin infection, *in vitro* transduction was performed by seeding of *in vitro* differentiated FoxP3+ T cells at 2 x 10^5 cells/well in a 96-well flat bottom plate in cell culture medium supplemented with human IL-2 (20 ng/ml), mouse IL-7 (100 ng/ml), mouse IL-15 (100 ng/ml), retinoic acid (10 nM) and human TGF-β1 (5 ng/ml). Cells were inoculated with concentrated MLV LTR-GFP (described in 2.2.3.1 and 2.2.3.2) at

2 x 10^5 I.U./well. Following an incubation period of 24 h at 37 °C, the number of infected FoxP3+ T cells was assessed by GFP expression using flow cytometry (BD FACSLyric).

2.2.3.6 Trypsinization of MLV-infected FoxP3+ T cells and B1 cells

Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

To remove extracellular virus particles, MLV-infected cells were treated with trypsin. In detail, 24 h after *in vitro* transduction using spin infection (described in 2.2.3.5) MLV-infected cells were harvested and transferred into 1 ml PBS/2 mM EDTA, followed by centrifugation for 8 min at 200 x g. For removal of extracellular virus particles, cells were resuspended in 250 µl 0.05 % Trypsin/EDTA and incubated for 2 min at 37 °C. Trypsin activity was inhibited by adding 1 ml RPMI-1640 GlutaMAX™ supplemented with 30 % FCS, followed by sequential incubation periods at 4 °C for 5 min and 37 °C for 8 min, respectively. After centrifugation for 8 min at 200 x g, cells were resuspended in fresh cell culture medium and stored at 37 °C.

2.2.3.7 *In vitro cis*-infection assay

Infection medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

Co-culture of MLV-infected donor cells and target lymphocytes was established to investigate MLV *cis*-infection *in vitro*. Therefore, using spin infection, LPS-activated B1 cells and *in vitro* differentiated FoxP3+ T cells, respectively, were *in vitro* transduced with MLV LTR-GFP (described in 2.2.3.5). After 24 h, MLV-infected cells were trypsinized to remove extracellular MLV LTR-GFP particles (described in 2.2.3.6). In case of MLV-infected FoxP3+ T cells as donor cells, cells were stained with CellTrace Cell Proliferation Kit FarRed according to the manufacturer's protocol (described in 2.2.4.3). Trypsinized MLV-infected cells were seeded at 1 x 10^5 cells/well in a flat bottom 96-well plate to serve as donor cell population. *In vitro* differentiated FoxP3+ T cells were added at a density of 2 x 10^5 cells/well to reach a final ratio of 1:2 (donor cells:target cells). Co-cultures were incubated for 24 h at 37 °C in 200 µl culture medium per well supplemented with mouse IL-5 (100 ng/ml), mouse IL-6 (100 ng/ml), mouse IL-7 (100 ng/ml) and mouse IL-15 (100 ng/ml) for B1 / FoxP3+ T cell co-cultures, and mouse IL-7 (100 ng/ml) and mouse IL-15 (100 ng/ml) for FoxP3+ T / FoxP3+ T cell co-cultures. MLV transmission was calculated as the number of GFP-expressing target cells relative to the number of GFP-expressing donor cells using a BD FACSLyric flow cytometer. In case of B1/FoxP3+ T co-cultures, B1 cells were discriminated from FoxP3+ T cells using anti-CD19 immunostaining (described in 2.2.4.1). For FoxP3+ T/FoxP3+ T cell co-cultures, populations were differentiated according to FarRed-staining in FarRed-positive (donor cells) and FarRed-negative (target cells) populations by flow cytometry.

2.2.3.8 Transwell co-culture assay

Infection medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

Co-cultures were physically separated to determine the contribution of cell-cell contacts to MLV transmission. Therefore, a 96-well transwell plate, containing inserts with a pore size of 3 µm, was coated in its bottom compartment with 100 µl human fibronectin in PBS to reach a final concentration of 1 µg per well. After incubation for 30 min at RT, coating was washed twice with PBS and once with cell culture medium. Donor and target cell populations were prepared according to 2.2.3.7. FoxP3+ T target cells were seeded into the bottom compartment at 2 x 10^5 cells/well. Separated by the transwell insert, trypsinized MLV-infected B1 donor cells were transferred at 1 x 10^5 cells/well into the upper compartment. Cells were co-cultured for 24 h at 37 °C in 300 µl culture medium per well, supplemented with mouse IL-5 (100 ng/ml), mouse IL-6 (100 ng/ml), mouse IL-7 (100 ng/ml) and mouse IL-15 (100 ng/ml). Anti-CD19 immunostaining allowed discrimination of FoxP3+ T cells from B1 cells (described in 2.2.4.1). MLV transmission was assessed based on the number of GFP-expressing target cells using a BD FACSLyric flow cytometer.

2.2.3.9 Antibody-mediated blocking *cis*-infection assay

Infection medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

Antibody-mediated blocking of cell adhesion-mediating proteins LFA1 and ICAM1 was used to determine the contribution of cell-cell contacts to MLV spread. Using spin infection, LPS-activated B1 cells were *in vitro* transduced with MLV LTR-GFP (described in 2.2.3.5). After 24 h, MLV-infected cells were trypsinized to remove extracellular MLV LTR-GFP particles (described in 2.2.3.6). Trypsinized MLV-infected cells were seeded at 1 x 10^5 cells/well in a flat bottom 96-well plate to serve as donor cell population. FoxP3+ T cells were treated with soluble anti-CD11a, anti-CD54, Isotype Ctrl IgG2a, Isotype Ctrl IgG2b antibodies (see Table 5) at 10 µg/ml for 30 min at 37 °C. FoxP3+ T cells were added to MLV-transduced donor cells at a density of 2 x 10^5 cells/well to reach a final ratio of 1:2 (donor cells:target cells). Co-cultures were incubated for 24 h at 37 °C in 200 µl culture medium per well supplemented with mouse IL-5 (100 ng/ml), mouse IL-6 (100 ng/ml), mouse IL-7 (100 ng/ml) and mouse IL-15 (100 ng/ml). B1 cells were discriminated from FoxP3+ T cells using anti-CD19 immunostaining (described in 2.2.4.1). MLV transmission was calculated as the number of GFP-expressing target cells relative to the number of GFP-expressing donor cells by flow cytometry using a BD FACSLyric.

2.2.3.10 *In vivo cis*-infection - adoptive cell transfer

Infection medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES; Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

MLV-infected FoxP3+ T cells were adoptively transferred into the footpads of mice to investigate retroviral *cis*-infection *in vivo*. Therefore, using spin infection, *in vitro* differentiated FoxP3+ T cells were *in vitro* transduced with MLV LTR-GFP, followed by incubation for 24 h at 37 °C (described in 2.2.3.5). To remove extracellular virus particles, FoxP3+ T cells were treated for 2 min at 37 °C with trypsin (described in 2.2.3.6). MLV-infected FoxP3+ T cells were labeled using the CellTrace Cell Proliferation Kit FarRed (described in 2.2.4.3). For adoptive cell transfer, C57BL/6, CD11a-/- and ICAM1-/- mice were anesthetized and MLV-infected, FarRed-positive FoxP3+ T cells were subcutaneously injected into the hind hock. Approximately 60 h after cell transfer, draining popliteal lymph nodes of euthanized mice were extracted and transferred into 115 µl serum-free RPMI-1640 GlutaMAX™ supplemented with 10 mM HEPES. After adding 25 µl Liberase TE (0.2 µg/ml) and 10 µl DNase I Mix (10 U/sample), lymph nodes were incubated for 20 min at 37 °C and 160 x g. Enzymatic process was inhibited by addition of 1 ml RPMI-1640 GlutaMAX™ supplemented with 10 % FCS. Lymph nodes were minced through a 70 µm cell strainer into a 50 ml falcon tube. Cells were centrifuged for 12 min at 160 x g and 12 °C, resuspended in 500 µl MACS buffer and transferred into the cell strainer lid of a polystyrene tube. Cells were filtered into the tube by centrifugation for 5 min at 160 x g and 12 °C. The amount of newly infected target cells was determined using flow cytometry as the number of GFP-expressing, FarRed-negative cells relative to the number of GFP-expressing, FarRed-positive donor cells.

2.2.4 Biochemical methods

2.2.4.1 Immunofluorescence staining

Blocking buffer: PBS, 1 % BSA, 10 % rat serum, 1 µl/(1 x 10^6 cells) Fc-blocking CD16/CD32 antibody; Washing buffer: PBS, 1 % BSA, BD Fixation/Permeabilization solution, BD Perm/Wash™ buffer;

Immunofluorescence staining was performed to examine the presence and relative expression levels of specific surface and intracellular proteins. Therefore, 1.0 x 10^6 cells were harvested and washed once with PBS supplemented with 1 % BSA, followed by centrifugation for 10 min at 400 x g and 12 °C. As the samples were not fixed, all steps were performed on ice. Prior to immunofluorescence staining, cell samples were blocked for 30 min in 50 µl PBS supplemented with 1 % BSA, 10 % rat serum and 1 µl Fc-blocking antibody against CD16/CD32 at 4 °C. 50 µl of primary antibodies diluted in PBS supplemented with 1 % BSA to reach final concentrations according to Table 5 were added to the cellular sample followed by incubation for 30 min at 4 °C.

For intracellular immunostaining, cell fixation and permeabilisation were performed using the Fixation/Permeabilization Kit by BD Biosciences according to the manufacturer's protocol. Thereafter, cells were incubated with 100 µl primary antibody mix (see Table 5) in 1X BD Perm/Wash™ buffer for 30 min at 4 °C. After two washing steps with 1X BD Perm/Wash™ buffer, cells were resuspended in 400 µl PBS supplemented with 1 % BSA and analyzed by flow cytometry on a BD FACSLyric.

2.2.4.2 Flow cytometry

MACS buffer: 0.5 % BSA, 2 mM EDTA in PBS;

Flow cytometric immunophenotyping allows for a precise differentiation and quantification of distinct cell populations and subpopulations by analysis of cell characteristic parameters, including cell size, granularity and relative levels of antigen expression. Therefore, cells were harvested, washed once in PBS/2mM EDTA and cell number was determined (described in 2.2.2.3). For cell type identification and distinction, immunofluorescence staining was performed according to 2.2.4.1. Prior to analysis, 100 µl cell suspension was transferred into 200 µl MACS buffer and flow cytometry data were collected using a BD FACSLyric.

2.2.4.3 CellTrace FarRed staining

Culture medium: RPMI-1640 GlutaMAX™, 10 % FCS, 10 mM HEPES, 10 mM sodium pyruvate, 1 % MEM non-essential amino acids, 55 µM β-mercaptoethanol;

For discrimination of a selected cell population under *in vitro* and *in vivo* conditions, cells were labeled using the CellTrace Cell Proliferation Kit FarRed according to the manufacturer's protocol. In detail, CellTrace FarRed dye was diluted to a working concentration of 1 µM in OPTiMEM. For FarRed staining, cell concentration was adjusted to 1 x 10^6 cells/ml prior to incubation in FarRed working solution for 20 min at 37 °C. To adsorb unbound dye, 5X the current volume of fresh culture medium was added to the cell suspension followed by incubation for 10 min at 37 °C, and resuspension in fresh culture medium.

3. Results

3.1 The role of LFA1 and ICAM1 in MLV cell-to-cell transmission

After subcutaneous (s.c.) virus delivery, retroviruses spread *in vivo* among lymphocyte populations of popliteal lymph nodes (pLNs) (Pi et al., 2019; Sewald et al., 2015; Uchil et al., 2019). Inside secondary lymphoid tissues, intravital microscopy revealed the formation of stable, glycoprotein Env-dependent cell-cell contacts between MLV-infected donor cells and non-infected target cells (Law et al., 2016; Murooka et al., 2012; Sewald et al., 2012). Infections with retroviruses deficient in Env or harboring non-functional Env result in a significant reduction in the frequency of cell-cell contacts (Law et al., 2016; Sewald et al., 2012). While these findings indicate that the formation of cell-cell contacts is initiated and/or stabilized by retroviral protein expression, a causal link between the occurrence of intercellular contacts and MLV transmission between cells *in vivo* has yet to be established.

3.1.1 Establishment of an *in vitro cis*-infection assay

To study the contribution of cell-cell contacts between MLV-infected donor cells and non-infected target cells to retroviral dissemination, we first established an *in vitro cis*-infection assay with primary cells of physiological relevance. Previous *in vivo* studies characterized B1 cells as target cells among MLV-infected lymphocyte populations in pLNs of living mice (Pi et al., 2019; Sewald et al., 2015). Within the CD4+ T cell population, our group recently characterized CD4+ memory cells and FoxP3-expressing (FoxP3+) CD4+ T cells as specific target cell populations permissive to MLV infection in pLNs *in vivo* (Engels, Falk et al., 2022).

As a first step, we isolated primary B1 cells and CD4+ T cells from mice for *in vitro* co-culture assays. B1 cells were enriched from the peritoneal cavity of wild-type (WT) C57BL/6 mice by negative selection and cell purity was evaluated through immunophenotyping using flow cytometry (described in 2.2.2.4 and 2.2.4.1). Directly after cell separation, CD19-expressing B1 cells constituted the predominant cell type with ~ 88 % of total cells, followed by a CD19-negative, F4/80-negative population with a relative frequency of ~ 9.5 % (Figure 6A). Post B1 cell isolation, CD4+ and CD8+ T cells accounted for less than 1 % of total isolated cells.

FoxP3+ T cells were differentiated *in vitro* from naïve CD4+ T cells to generate sufficient cells for our co-culture experiments. For this purpose, after isolation of naïve CD4+ T cells from splenocytes of WT C57BL/6 mice by negative selection, cells were differentiated for 48 h and FoxP3-expression was assessed by immunophenotyping using flow cytometry (described in 2.2.2.5, 2.2.2.6 and 2.2.4.1). As depicted in Figure 6B, the number of successfully differentiated FoxP3+ T cells reached ~ 96 % in response to CD3/CD28-mediated activation in combination with interleukin 7 (IL-7), IL-15, trans-forming growth factor-β1 (TGF-β1), IL-2 and retinoic acid treatment.

Finally, we examined the capacity of both primary cell populations to support retroviral cell-to-cell transmission *in vitro.* Therefore, we transduced B1 cells with MLV LTR-GFP using spin infection and removed extracellular viral particles by trypsin digestion prior to

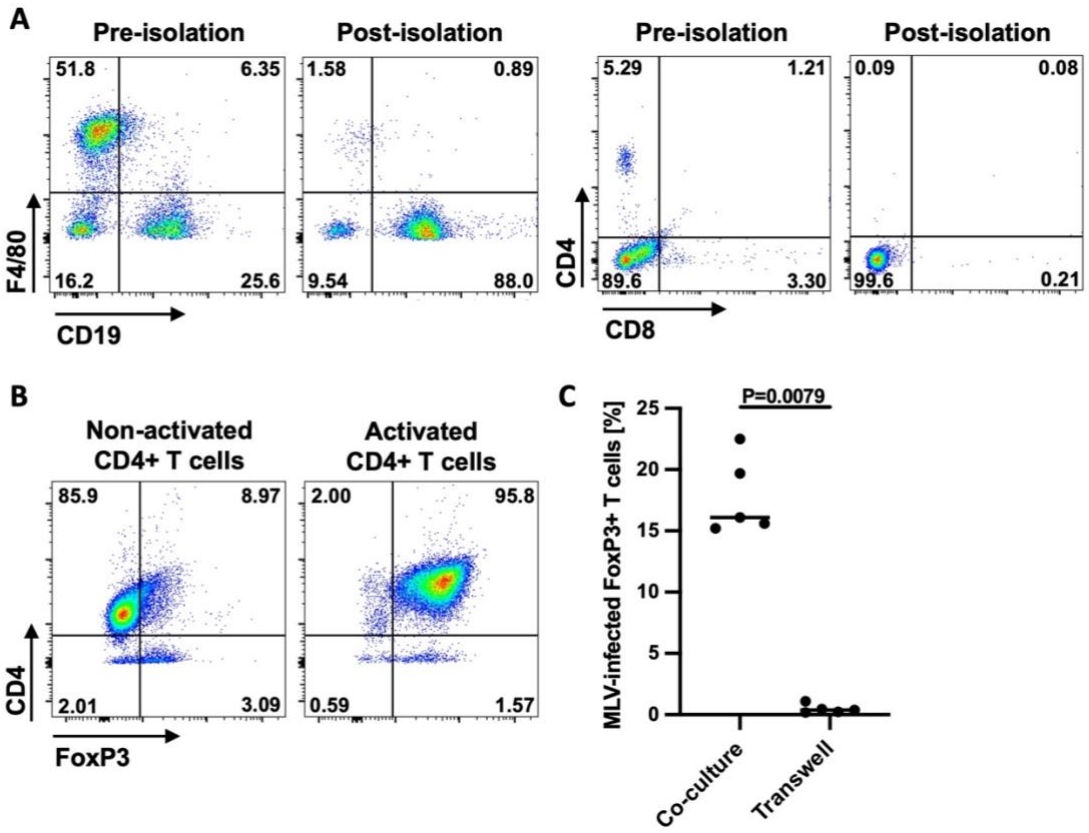

Figure 6: B1 and FoxP3+ T cells support MLV cell-to-cell transmission *in vitro*. (A) Flow cytometric analysis pre- and post-enrichment of primary B1 cells by negative selection of peritoneal washes from wild-type (WT) C57BL/6 mice. Immunophenotyping includes B1 cells (CD19-positive), macrophages (F4/80-positive), CD4+ T cells (CD4-positive), and CD8+ T cells (CD8-positive). Analysis is representative for three independent experiments. (B) Comparative flow cytometric analysis to assess *in vitro* differentiation of naïve CD4+ T cells into FoxP3-expressing (FoxP3+) CD4+ T cells. For activation and differentiation, CD4+ T cells were cultured on anti-CD3/CD28-coated surfaces in the presence of interleukin 7 (IL-7), IL-15, transforming growth factor-β1 (TGF-β1), IL-2 and retinoic acid. Non-activated CD4+ T cells cultured in the presence of above-mentioned cytokines were used as control. Analysis is representative for three independent experiments. (C) Quantification of MLV-infected FoxP3+ T cells post direct co-cultivation with MLV-infected B1 cells as compared to co-cultivation in transwells (n=5).

co-culture with non-infected FoxP3+ T cells at a 1:2 ratio (described in 2.2.3.5, 2.2.3.6 and 2.2.3.7). To assess the contribution of direct cell-cell contacts to MLV transmission, we cultured *in vitro* transduced, MLV-infected B1 donor cells with non-infected FoxP3+ T target cells under two distinct culture conditions: a direct co-culture in comparison to a co-culture using transwell inserts, which allow for physical separation of the donor and target cell population and thereby limit MLV spread to cell-free transmission (described in 2.2.3.7 and 2.2.3.8). After 24 h, we quantified MLV transmission as the number of GFP-expressing (GFP+) FoxP3+ T target cells using flow cytometry (described in 2.2.4.2). Physical separation of MLV-infected donor cells from non-infected target cells resulted in a significant reduction in the number of MLV-infected target cells, as compared to direct co-culture of both cell populations (Figure 6C).

Taken together, we successfully established a co-culture assay with primary cell populations relevant for MLV infection and demonstrated cell contact-dependent transmission of MLV *in vitro*.

3.1.2 LFA1 and ICAM1 are crucial for efficient MLV transmission *in vitro*

In their function as mediators of cellular adhesion and migration, LFA1 and its ligand ICAM1 are essential for various steps of the host immune response, including lymphocyte homing, antigen presentation and T cell-mediated killing (Núñez et al., 2017; Springer & Dustin, 2012). To investigate the relevance of cell-cell interactions for retroviral transmission, we studied the role of LFA1 and ICAM1 during *cis*-infection of MLV using our established *in vitro* co-culture assay.

For this purpose, we first characterized surface expression of LFA1 and its ligand ICAM1 on primary cells applied for *in vitro* co-cultures using flow cytometry (described in 2.2.4.1 and 2.2.4.2). Both cell surface proteins were expressed at high levels on B1 and FoxP3+ T cells irrespective of the infection status of the cell (Figure 7A).

Next, we used antibody-mediated blocking to investigate a potential contribution of both surface proteins to cell contact-dependent MLV transmission during *cis*-infection (described in 2.2.3.9). Therefore, MLV-infected B1 donor cells were co-cultured with non-infected FoxP3+ T target cells in the presence of anti-CD11a or anti-ICAM1 blocking antibodies. After 24 h, MLV transmission was quantified as the number of GFP+ FoxP3+ T cells relative to the number of GFP+ B1 cells using flow cytometry (described in 2.2.4.2). Blocking of CD11a resulted in a significant decline in the number of MLV-infected target cells compared to isotype IgG-treated control cells, suggesting a contribution of LFA1 to MLV spread (Figure 7B). Treatment with ICAM1-blocking antibodies had only a minor effect on MLV transmission as compared to isotype control. These findings indicate a contributive effect of LFA1/ICAM1 interaction to efficient, cell contact-dependent MLV spread *in vitro*.

However, antibody-blocking assays cannot differentiate the individual contribution of cell adhesion proteins to MLV transmission with respect to the donor and target cell population. To study the role of LFA1 and its ligand ICAM1 in cell contact-dependent MLV transmission on donor and target cells separately, we combined primary cells isolated from WT C57BL/6 mice with ICAM1-knockout (KO) and CD11a-KO (LFA1-deficient) mouse lines (described in 2.1.9). To function as donor cells, B1 cells and *in vitro* differentiated FoxP3+ T cells from WT C57BL/6 and KO mice, were *in vitro* transduced with MLV LTR-GFP using spin infection (described in 2.2.3.5). After 24 h to allow for MLV protein expression, MLV-infected donor cells were co-cultured with non-infected FoxP3+ T cells from both WT C57BL/6 and KO mice as target cells (described in 2.2.3.7). MLV transmission was determined by flow cytometry as the number of GFP+ FoxP3+ T target cells relative to the number of GFP+ B1 and GFP+FoxP3+ T donor cells, respectively, 24 h after start of the co-culture (described in 2.2.4.2).

Co-culture of MLV-infected donor cells resulted in a significant reduction in the infection rate of CD11a-KO (LFA1-deficient) FoxP3+ T target cells (Figure 7C and Figure 7E). By contrast, infection rate of FoxP3+ T target cells reached similar levels after co-culture with LFA1-deficient and WT C57BL/6-derived donor cells. The presence of ICAM1 on donor cells supported MLV transmission to the target cell population (Figure 7D and Figure 7F). Thus, co-culture of ICAM1-deficient donor cells significantly decreased MLV transmission to FoxP3+ T target cells, whereas infection levels of WT C57BL/6-derived and ICAM1-deficient target cells were comparable after co-culture with WT

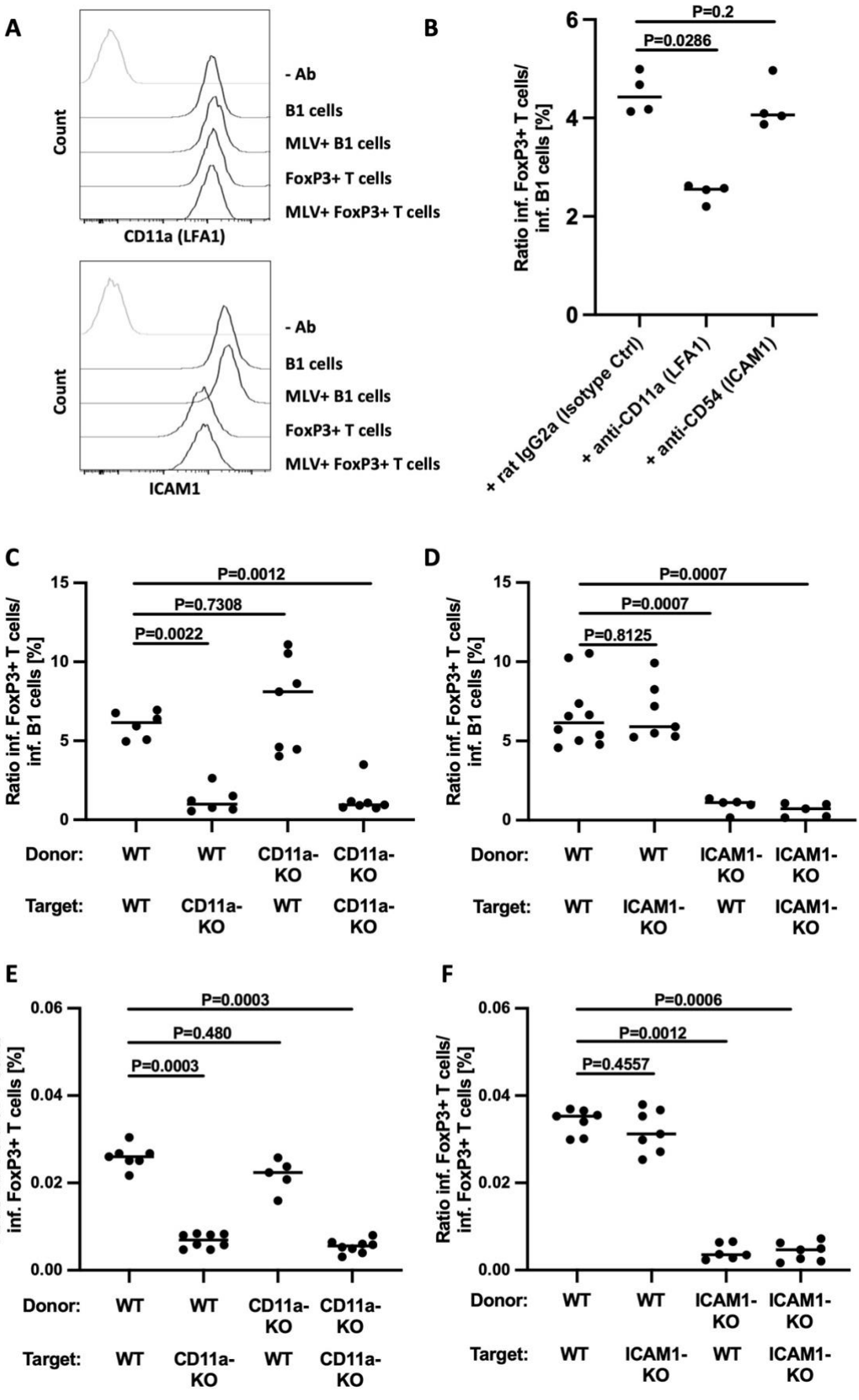
A
Count
- Ab
B1 cells
MLV+ B1 cells
FoxP3+ T cells
MLV+ FoxP3+ T cells
CD11a (LFA1)
Count
- Ab
B1 cells
MLV+ B1 cells
FoxP3+ T cells
MLV+ FoxP3+ T cells
ICAM1
B
Ratio inf. FoxP3+ T cells/ inf. B1 cells [%]
P=0.2
P=0.0286
+ rat IgG2a (Isotype Ctrl)
+ anti-CD11a (LFA1)
+ anti-CD54 (ICAM1)
C
Ratio inf. FoxP3+ T cells/ inf. B1 cells [%]
P=0.0012
P=0.7308
P=0.0022
Donor: WT WT CD11a-KO CD11a-KO
Target: WT CD11a-KO WT CD11a-KO
D
Ratio inf. FoxP3+ T cells/ inf. B1 cells [%]
P=0.0007
P=0.0007
P=0.8125
Donor: WT WT ICAM1-KO ICAM1-KO
Target: WT ICAM1-KO WT ICAM1-KO
E
Ratio inf. FoxP3+ T cells/ inf. FoxP3+ T cells [%]
P=0.0003
P=0.480
P=0.0003
Donor: WT WT CD11a-KO CD11a-KO
Target: WT CD11a-KO WT CD11a-KO
F
Ratio inf. FoxP3+ T cells/ inf. FoxP3+ T cells [%]
P=0.0006
P=0.0012
P=0.4557
Donor: WT WT ICAM1-KO ICAM1-KO
Target: WT ICAM1-KO WT ICAM1-KO

Figure 7: LFA1 and ICAM1 are crucial for cell contact-dependent transmission of MLV *in vitro*. (A) Surface expression levels of CD11a (LFA1) and ICAM1 on non-infected and MLV-infected (MLV+) B1 cells and *in vitro* differentiated FoxP3+ T cells, in comparison to a representative, unstained control (-Ab), assessed by flow cytometry. Expression analysis is representative for three independent experiments. (B) Quantification of MLV-infected FoxP3+ T target cells relative to MLV-infected B1 donor cells after 24 h co-culture in the presence of anti-CD11a or anti-ICAM1 blocking antibodies. Treatment with isotype IgG antibodies served as Fc control (n=4). (C, D, E, F) Quantification of MLV-infected target FoxP3+ T cells relative to (C, D) MLV-infected B1 donor cells or (E, F) MLV-infected FoxP3+ T donor cells after 24 h of co-culture. Donor and target cell populations were isolated from WT C57BL/6, CD11a-knockout (KO) and ICAM1-KO mice, respectively. Combination of WT C57BL/6- with CD11a-KO-derived primary cells is shown in (C, E), and combination of WT C57BL/6- with ICAM1-KO-derived primary cells in (D, F), (n=6-7; C), (n=5-10; D), (n=5-8; E), (n=6-7; F).

C57BL/6-derived donor cells. At this point, it is important to note, that our findings were consistent for both donor cell types, B1 and FoxP3+ T cells, although MLV transmission efficiency in B1/FoxP3+ T co-cultures exceeded FoxP3+ T/FoxP3+ T cell co-cultures.
In conclusion, our results suggest a crucial role of LFA1 on the target cell population and its ligand ICAM1 on the donor cell population for efficient MLV transmission during *cis*-infection *in vitro.*

3.1.3 LFA1 and ICAM1 are crucial for efficient MLV transmission *in vivo*

Previous *in vivo* studies described the formation of stable cell-cell interactions between retrovirus-infected donor cells and non-infected target cells in lymphoid tissues (Law et al., 2016; Sewald et al., 2012). Yet, a direct correlation between cell-cell interactions and retroviral spread *in vivo* has yet to be established. To address this, we investigated the individual function of LFA1 and ICAM1 expression on donor and target cells for retroviral transmission *in vivo*.
Therefore, *in vitro* differentiated FoxP3+ T cells, isolated from WT C57BL/6, CD11a-KO and ICAM1-KO mice, were *in vitro* transduced with MLV LTR-GFP and labeled by intracellular FarRed staining (FarRed+) prior to s.c. injection into the foot pads of WT C57BL/6, CD11a-KO and ICAM1-KO acceptor mice (described in 2.2.3.10, Figure 8A). After 72 h, draining popliteal lymph nodes were removed to generate a single cell suspension, and GFP-expressing leukocytes were quantified by flow cytometry (described in 2.2.4.2). MLV transmission was determined as the number of newly MLV-infected (GFP+FarRed-) leukocytes relative to the number of MLV-infected, adoptively transferred (GFP+FarRed+) FoxP3+ T donor cells.
MLV transmission from WT C57BL/6 donor cells was significantly impaired in pLNs of CD11a-KO mice, suggesting a critical function of LFA1 on target cells for efficient retroviral *cis*-infection *in vivo* (Figure 8B). LFA1-deficient donor cells transmitted MLV in pLNs of WT C57BL/6 mice at rates similar to WT C57BL/6-derived donor cells. The LFA1 ligand ICAM1 supported efficient retroviral spread when expressed on the surface of donor cells, whereas its contribution to MLV transmission on the target cell surface was dispensable. Thus, ICAM1-deficient donor cells transmitted MLV in pLNs of WT C57BL/6 mice at significantly lower rates as compared to WT C57BL/6 donor cells. MLV transmission from WT C57BL/6-derived donor cells in pLNs of ICAM1-KO mice, however, was comparable to spread in pLNs of WT C57BL/6 acceptor mice.

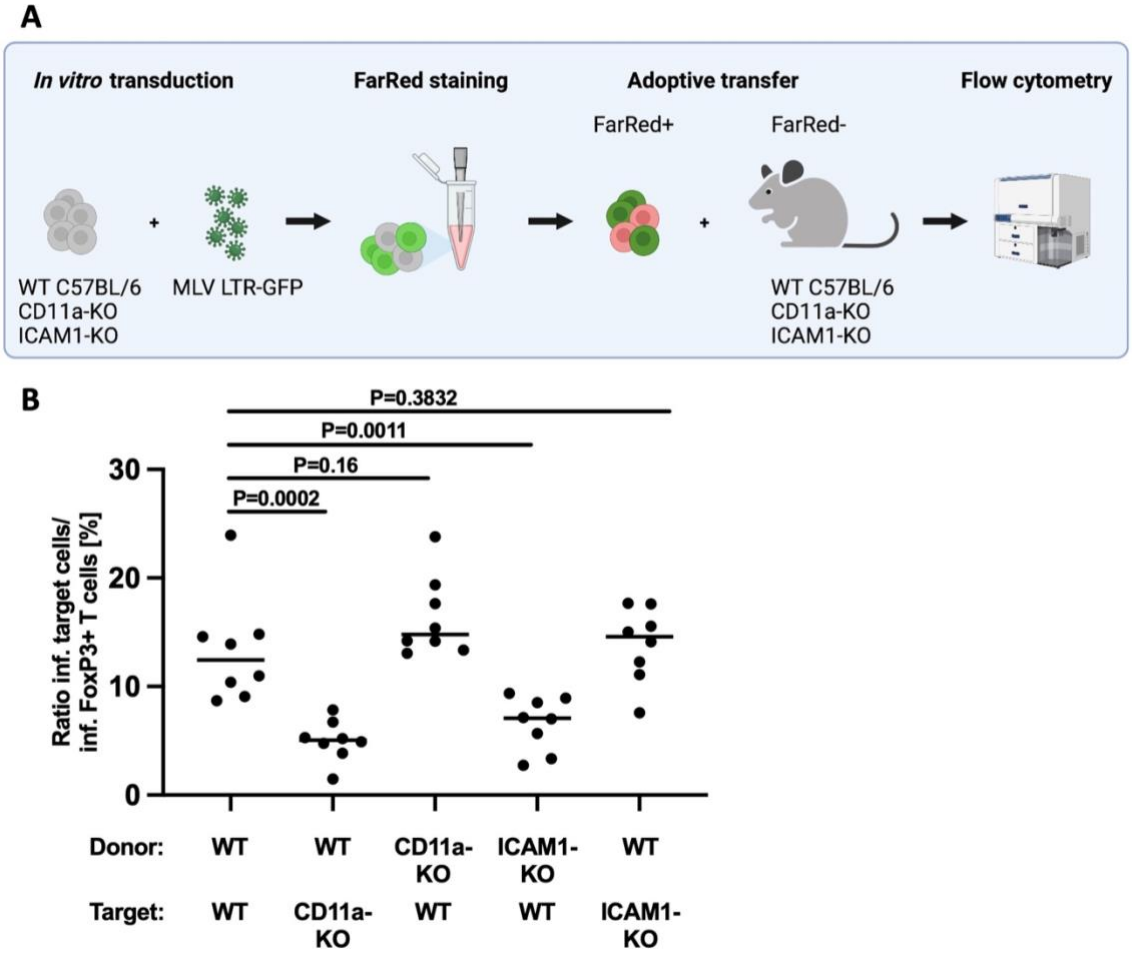

Figure 8: LFA1 and ICAM1 are crucial for cell contact-dependent transmission of MLV *in vivo*. (A) Schematic representation of the adoptive cell transfer workflow. Naïve CD4+ T cells were enriched from splenocytes of WT C57BL/6, CD11a-knockout (KO) and ICAM1-KO mice, *in vitro* differentiated into FoxP3-expressing (FoxP3+) T cells and transduced with MLV LTR-GFP reporter suspension (2 x 10^5 I.U.). After 24 h, transduced FoxP3+ T cells were trypsin treated and intracellularly stained with FarRed (FarRed+) prior to subcutaneous injection into foot pads of WT C57BL/6, CD11a-KO and ICAM1-KO acceptor mice (FarRed-). After 72 h, popliteal lymph nodes were removed and MLV transmission was quantified as the number of newly infected leukocytes (GFP+FarRed-) relative to the number of MLV-infected, adoptively transferred FoxP3+ T cells (GFP+FarRed+). (B) Quantification of MLV transmission during *in vivo cis*-infection after adoptive cell transfer, as described in (A), as the number of newly MLV-infected, lymph node-derived cells (GFP+FarRed-) relative to the number of MLV-infected, adoptively transferred FoxP3+ T cells (GFP+FarRed+), (n=8).

Taken together, our observations indicate that efficient MLV transmission *in vivo* requires expression of LFA1 on target cells and ICAM1 on donor cells. Thereby, our adoptive transfer experiments suggest a critical role of cell-cell contacts for retroviral dissemination *in vivo.*

3.2 Establishment of CRISPR/Cas9 target enrichment Nanopore sequencing as a novel provirus detection technique

The efficiency of HIV transmission across cell-cell contacts during *cis*-infection exceeds cell-free transmission by 100- to 10,000-fold, partly explained by a high local MOI with elevated virus transfer rates to target cells (Chen et al., 2007; Del Portillo et al., 2011; Dimitrov et al., 1993; Law et al., 2016; Russell et al., 2013). Yet, the consequences of cell-to-cell transmission for infection at the cellular level and retroviral pathogenesis remain incompletely understood. One major limitation is that previously applied experimental approaches, such as multifluorescent reporter systems, fluorescence *in situ* hybridization (FISH) and qPCR, provide only limited insights into provirus frequency and the pattern of retroviral integration. Given the technical limitations of established provirus detection techniques, we aimed to establish CRISPR/Cas9 target enrichment Nanopore sequencing as an analytical approach to localize and quantify retroviral DNA integrated in the genome of a host cell.

3.2.1 Generation of a monoclonal cell model to study retroviral integration

To study retroviral integration on a representative cellular model, we first generated a library of monoclonal cell lines displaying high heterogeneity in the distribution and frequency of their proviral integration events.

For the generation of cell clones with a broad range of integrated proviruses, murine S49.1 cells were transduced with ecotropic (E), replication-incompetent virus-like particles (VLPs), containing a 3.3 kb LTR-GFP reporter genome, at low (2×10^4 I.U., E1) and high (2×10^5 I.U., E10) infectious titers (described in 2.2.3.4). Transduction was performed in the absence (-) or presence (+) of 1 % methylcellulose (MC), enhancing transduction efficiency of S49.1 cells with MLV LTR-GFP by 30-fold (Figure 9A). Increasing transduction titers in the presence or absence of methylcellulose were chosen to gain different cell clones with a broad quantitative range of integrated LTR-GFP genomes. After 24 h, single cell isolation was performed by limiting dilution at a seeding concentration of 0.5 cells per well on a 96-well plate and validated by visual screening using bright field microscopy (Figure 9B). Exclusively GFP-expressing (GFP+) single cells were expanded for six weeks at 37°C and subsequently characterized by GFP mean fluorescence intensity (MFI) using flow cytometry and relative GFP copy number by Taqman qPCR (described in 2.2.4.2 and 2.2.1.3).

As illustrated in Figure 9C, GFP expression levels varied considerably between GFP+ cell clones with respect to the applied transduction parameters. Clones E1-, generated by transduction with low infectious titers in the absence of MC, yielded lowest average GFP MFIs (4,801 - 12,538 A.U.), compared to parental S49.1 cells (GFP-negative, 55.8 A.U.). Clones E10+, transduced with high infectious titers in the presence of MC, displayed highest GFP MFIs (132,499 A.U. and 178,313 A.U.) among all monoclonal cell populations. GFP+ clone subsets E1+ and E10- showed comparable GFP expression levels with MFIs values ranging between 4,812 A.U. - 68,923 A.U. To cover a maximal range of GFP expression levels, we selected a total of 25 GFP+ cell clones, all exhibiting highest GFP MFIs within their transduction group. Effects of distinct transduction parameters were likewise reflected in relative GFP copy numbers, determined by qPCR,

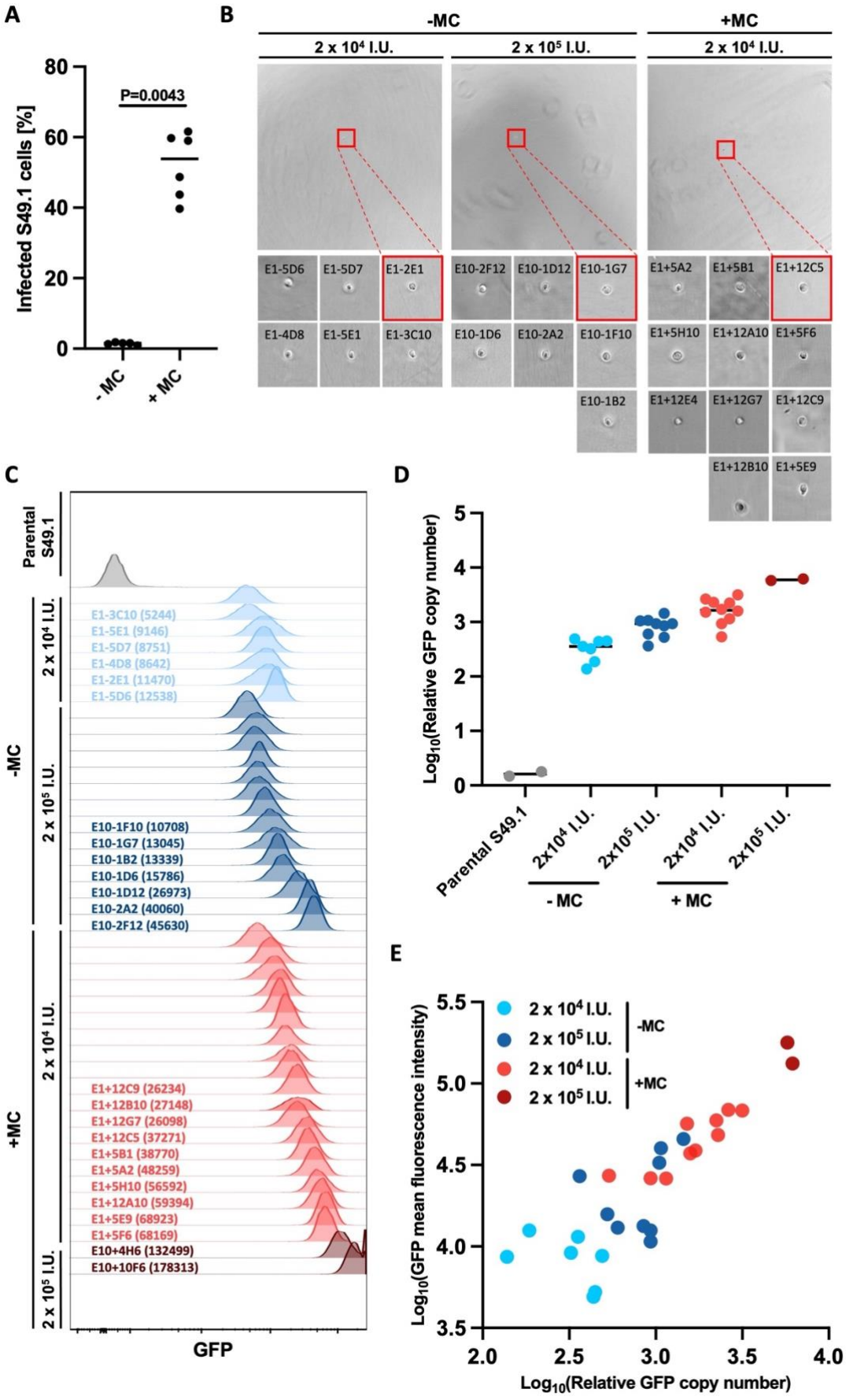
A
Infected S49.1 cells [%]
80
60
40
20
0
P=0.0043
- MC
+ MC
B
-MC
+MC
2 x 10⁴ I.U.
2 x 10⁵ I.U.
2 x 10⁴ I.U.
E1-5D6
E1-5D7
E1-2E1
E10-2F12
E10-1D12
E10-1G7
E1+5A2
E1+5B1
E1+12C5
E1-4D8
E1-5E1
E1-3C10
E10-1D6
E10-2A2
E10-1F10
E1+5H10
E1+12A10
E1+5F6
E10-1B2
E1+12E4
E1+12G7
E1+12C9
E1+12B10
E1+5E9
C
Parental S49.1
-MC
2 x 10⁴ I.U.
2 x 10⁵ I.U.
+MC
2 x 10⁴ I.U.
2 x 10⁵ I.U.
E1-3C10 (5244)
E1-5E1 (9146)
E1-5D7 (8751)
E1-4D8 (8642)
E1-2E1 (11470)
E1-5D6 (12538)
E10-1F10 (10708)
E10-1G7 (13045)
E10-1B2 (13339)
E10-1D6 (15786)
E10-1D12 (26973)
E10-2A2 (40060)
E10-2F12 (45630)
E1+12C9 (26234)
E1+12B10 (27148)
E1+12G7 (26098)
E1+12C5 (37271)
E1+5B1 (38770)
E1+5A2 (48259)
E1+5H10 (56592)
E1+12A10 (59394)
E1+5E9 (68923)
E1+5F6 (68169)
E10+4H6 (132499)
E10+10F6 (178313)
GFP
D
Log10(Relative GFP copy number)
5
4
3
2
1
0
Parental S49.1
2x10⁴ I.U.
2x10⁵ I.U.
2x10⁴ I.U.
2x10⁵ I.U.
- MC
+ MC
E
Log10(GFP mean fluorescence intensity)
5.5
5.0
4.5
4.0
3.5
2.0
2.5
3.0
3.5
4.0
Log10(Relative GFP copy number)
2 x 10⁴ I.U.
2 x 10⁵ I.U.
-MC
2 x 10⁴ I.U.
2 x 10⁵ I.U.
+MC

Figure 9: Generation and characterization of GFP-expressing (GFP+) cell clones. Murine T cell line S49.1 was transduced with ecotropic (E), single-round virus-like particles (VLPs) containing an LTR-GFP reporter genome at low (2 x 10^4 I.U., E1) or high (2 x 10^5 I.U., E10) infectious titers in the presence (+) or absence (-) of 1 % methylcellulose (MC). (A) Quantification of infected S49.1 cells after infection with MLV LTR-GFP in the presence and absence of 1 % methylcellulose. (B) Single cell isolation was performed by limiting dilution (0.5 cells/100 µl) in 96-well round bottom plates and validated by visual screening using bright field microscopy. (C) Flow cytometric analysis of GFP-expressing (GFP+) cell clones after single cell expansion. Cell clones selected for downstream analysis are labeled with their respective GFP+ cell clone name and GFP mean fluorescence intensity in brackets. (D) Relative GFP copy number of selected GFP+ cell clones in 10 ng genomic DNA as assessed by TaqMan quantitative PCR in comparison to parental S49.1 cells serving as negative control. (E) Correlation graph of GFP mean fluorescence intensity and relative GFP copy number of selected GFP+ cell clones. GFP mean fluorescence intensity analysis of selected clones is shown in (C) and relative GFP copy number in (D).

with an increase of relative GFP copy numbers from clones E1- to clones E10+ by ~ 20-fold (Figure 9D). Consistently, relative GFP copy numbers positively correlated with GFP expression levels for the individual transduction groups (Figure 9E).
In conclusion, we generated and characterized a set of GFP+ cell clones, stably expressing GFP at distinct intensities, which correspond to differences in their relative GFP copy numbers. The observed variations in GFP expression levels indicate that selected GFP+ cell clones harbor distinct numbers and/or genome-wide distributions of integrated LTR-GFP reporter genomes, which can serve as a target to validate performance of a novel provirus detection technique during its establishment.

3.2.2 Adaption of CRISPR/Cas9 enrichment strategy for targeting integrated GFP sequences

CRISPR/Cas9-mediated enrichment in combination with long-read Nanopore sequencing has proven as a valuable technology for targeted sequencing of a specific ROI within genomic DNA (Gilpatrick et al., 2020; Stangl et al., 2020; Van Haasteren et al., 2021). To establish CRISPR/Cas9 target enrichment Nanopore sequencing as a novel provirus detection technique, we selected the GFP sequence as ROI to allow for targeted enrichment of GFP-containing proviruses in genomic DNA of retrovirus-infected cells. Therefore, we designed a panel of three different crRNAs (D5-GFP, D6-GFP, D7-GFP), which target multiple sites within the GFP sequence (Figure 10A). Each of the three 20 nt crRNAs is complementary to different target sequences within GFP and is flanked in 3’ direction by a “NGG” PAM on the sense-strand (D6-GFP) or anti-sense-strand (D5-GFP, D7-GFP) of the reporter virus genome (Figure 10B). Hybridization of crRNAs with their complementary GFP target sequence guides Cas9 nuclease to target DNA sites, where it introduces a sequence-specific double-strand cleavage upstream of a PAM (Stangl et al., 2020). All GFP-targeting crRNAs were evaluated by IDT algorithms for optimal on-target efficiency and minimal potential off-target restriction in both, the murine and human genome. To further assess target specificity of crRNAs, we subjected LTR-GFP plasmid (~ 9.3 kbp) to crRNA/Cas9 complexes. Non-targeting control (ntc) crRNA in complex with Cas9 served as control. Reaction products were analyzed after 30 min incubation by agarose gel electrophoresis, validating Cas9-mediated linearization

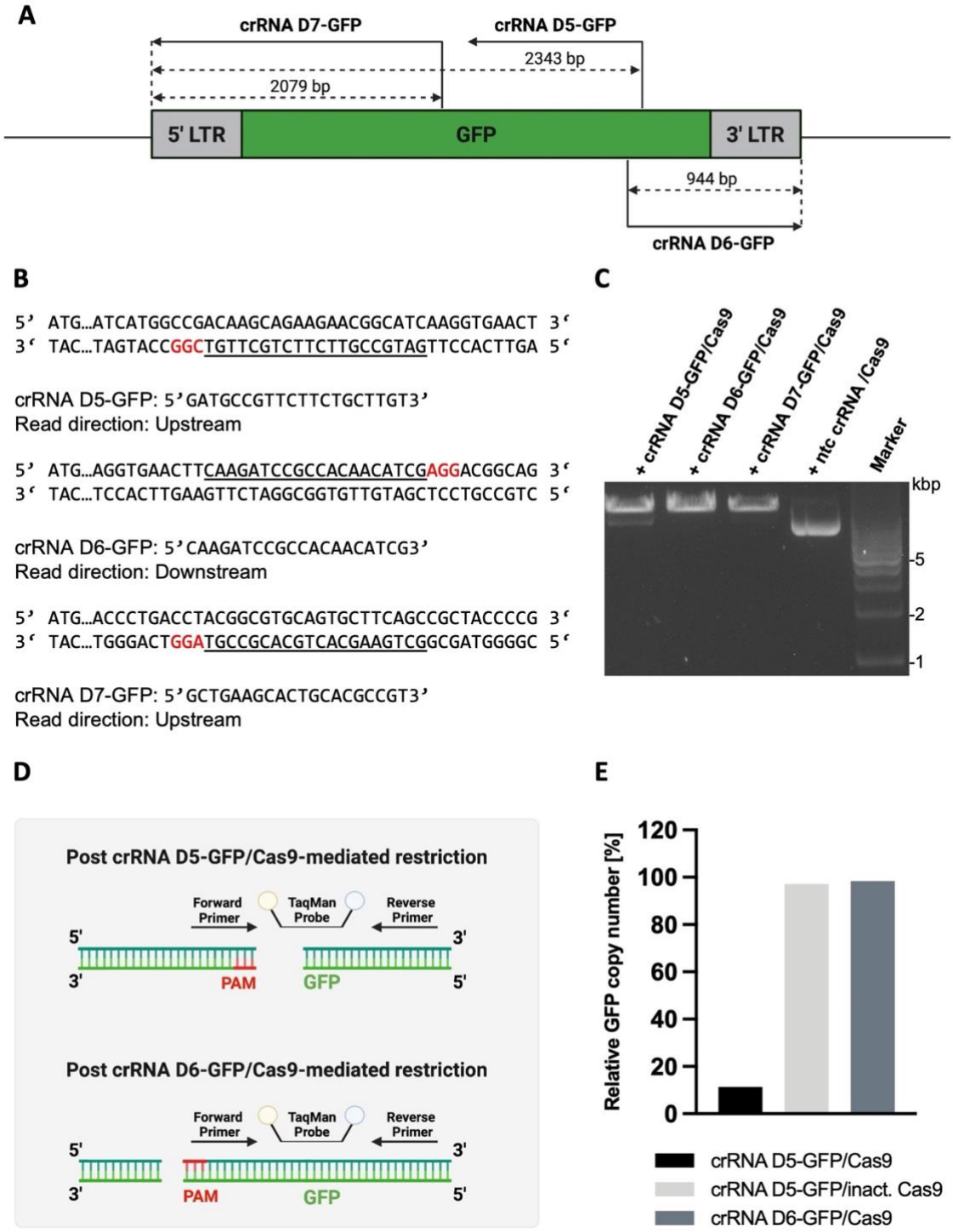

Figure 10: Design and evaluation of GFP-targeting crRNAs. (A) A total of three crRNAs (D5-GFP, D6-GFP, D7-GFP) was designed to allow targeting of GFP at multiple sites. Arrows highlight location of Cas9-introduced sequence cuts and point in the sequencing direction of enriched reads after Cas9-mediated GFP-enrichment for each crRNA. (B) Sequencing directionality is achieved through strand-specificity of crRNAs. CrRNAs D5-GFP and D7-GFP were designed to harbor a 20 bp GFP sequence (underlined) of the reverse strand, yielding CRISPR/Cas9-mediated enrichment of DNA fragments upstream of the cleavage site. In contrast, crRNA D6-GFP contains a 20 bp GFP sequence (underlined) of the forward strand, consequently enabling Cas9-mediated enrichment of DNA fragments downstream of the target sequence. All crRNAs are flanked in 3' direction by a "NGG" protospacer adjacent motif (PAM; highlighted in red). (C) Agarose gel electrophoresis of reaction products after applying CRISPR/Cas9 enrichment workflow to LTR-GFP plasmid. Restriction activity was assessed for all GFP-targeting crRNAs (D5-GFP, D6-GFP, D7-GFP)/Cas9 individually, non-targeting control crRNA (ntc crRNA)/Cas9 served as control. (D) Design of breakpoint-spanning TaqMan quantitative PCR (qPCR) to quantify *in vitro*

restriction efficiency of crRNA D5-GFP/Cas9. Arrows represent orientation of GFP-specific primers, flanking the predicted breakpoint after crRNA D5-GFP/Cas9- and crRNA D6-GFP/Cas9-mediated cleavage. Horizontal line marks position of TaqMan probe, spanning the predicted breakpoint. Predicted breakpoints are each adjacent to a "NGG" protospacer adjacent motif (PAM; highlighted in red). Restriction by crRNA D6-GFP/Cas9 does not interfere with performance of TaqMan qPCR. (E) *In vitro* restriction efficiency of crRNA D5-GFP/Cas9 quantified by breakpoint-spanning TaqMan qPCR. LTR-GFP plasmid was subjected to CRISPR/Cas9 enrichment workflow using crRNA D5-GFP/Cas9. As negative controls, crRNA D5-GFP/heat-inactivated (inact.) Cas9 and crRNA D6-GFP/Cas9, respectively, were used. After CRISPR/Cas9 enrichment workflow, percentage of on-target cleaved plasmid was assessed as the decline in TaqMan fluorescence signal during qPCR in comparison to negative controls. Analysis is representative for five independent experiments.

of LTR-GFP plasmid for all GFP-targeting crRNAs, in contrast to non-restricted plasmid applying ntc crRNA (described in 2.2.1.2 and 2.2.1.4, Figure 10C). Next, we quantified *in vitro* restriction efficiency through the example of crRNA D5-GFP/Cas9-mediated cleavage of LTR-GFP plasmid by breakpoint-spanning TaqMan qPCR. Therefore, we designed a flanking primer set and a TaqMan probe spanning the predicted cleavage site (described in 2.2.1.3, Figure 10D). Consequently, after crRNA D5-GFP/Cas9-mediated restriction, the number of on-target cleaved DNA fragments corresponds to the decline of fluorescence signal detected by qPCR, as compared to controls. To exclude false-positive results by hybridization of crRNA D5-GFP with qPCR primers, we included target enrichment with heat-inactivated Cas9 as a first negative control. In addition, restriction specificity was assessed by crRNA D6/Cas9-mediated restriction at a different target site, serving as second negative control. After subjecting LTR-GFP plasmid to crRNA D5-GFP-guided Cas9 cleavage, breakpoint-spanning TaqMan qPCR quantified on-target restriction of LTR-GFP plasmid with ∼90 % as compared to negative controls (Figure 10E).
Taken together, we validated individual restriction activity of GFP-targeting crRNAs and quantified restriction efficiency of crRNA D5-GFP by applying crRNA/Cas9 complexes to LTR-GFP plasmid.

3.2.3 CRISPR/Cas9 target enrichment Nanopore sequencing detects provirus location and frequency in GFP+ cell clones

Using the characterized monoclonal cell lines, we aimed to evaluate the capacity of CRISPR/Cas9 target enrichment Nanopore sequencing to simultaneously determine frequency and genomic location of GFP-containing proviruses integrated in genomic DNA of GFP+ cell clones (see section 3.2.1). For this purpose, we subjected genomic DNA of nine selected GFP+ cell clones to our CRISPR/Cas9 target enrichment strategy and analyzed the enriched DNA library by long-read Nanopore sequencing (Figure 11). In detail, genomic DNA from GFP reporter virus-infected cells was dephosphorylated and GFP was simultaneously targeted by crRNA/Cas9 ribonucleoprotein complexes D6-GFP and D7-GFP to achieve bi-directional sequencing (described in 2.2.1.4 and 2.2.2.7, Figure 11A). As internal positive control, we included crRNA targeting the location of the mouse *Hprt1* gene. Hybridization of crRNAs with their complementary GFP target sequence guided Cas9 nuclease to target DNA sites, where it introduced a sequence-specific double-strand cleavage upstream of a PAM. After cleavage, Cas9 nuclease

remained stably bound to PAM-distal DNA fragments, allowing selective dA-tailing of dissociated, PAM-proximal DNA fragments (Figure 11B) (Gilpatrick et al., 2020; Stangl et al., 2020; Sternberg et al., 2014). Design of crRNAs in a strand-specific manner targeted ligation of Nanopore sequencing adapters selectively to Cas9-cleaved PAM-proximal DNA fragments and dictated sequencing into 5' (D7-GFP) or 3'direction (D6-GFP). Enriched DNA library was sequenced in a single Flongle flow cell on a MinION sequencer. In collaboration with the Utrecht Sequencing Facility and the Department of Genetics at the University Medical Center Utrecht, we established a bioinformatic approach to identify the location, orientation, and structural integrity of GFP-containing proviruses from Nanopore sequencing data (Figure 11C). In sequential steps, Nanopore-acquired sequencing reads were first mapped to the LTR-GFP genome using Minimap2, followed by sequence alignment of filtered reads against the host reference genome, here the mouse reference genome GRCm38. The quality of sample processing and Nanopore sequencing performance were judged by enrichment of *Hprt1* reads. MinION runs which did not result in *Hprt1* enrichment were excluded from downstream processing.

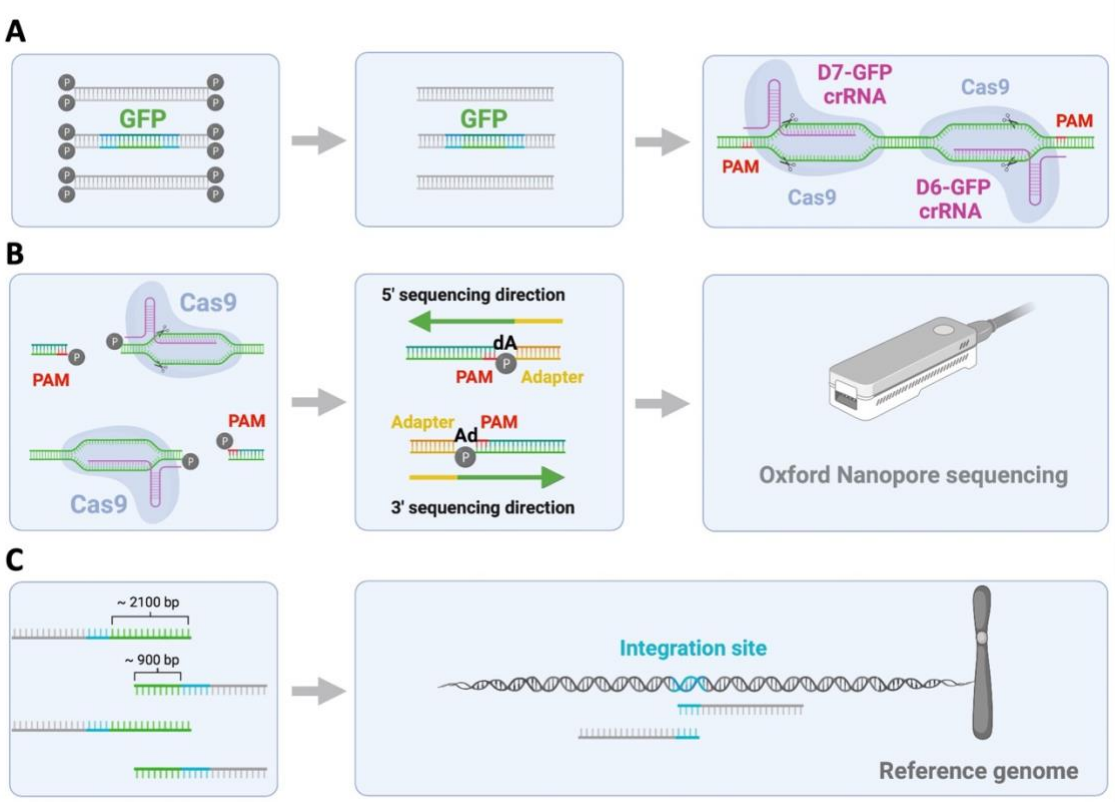

Figure 11: CRISPR/Cas9 target enrichment Nanopore sequencing workflow for detection of GFP-containing proviruses in genomic DNA. (A) Genomic DNA, extracted from GFP reporter virus-infected cells, was dephosphorylated and GFP insert was targeted by crRNA/Cas9 ribonucleoprotein complexes. CrRNAs, D6-GFP and D7-GFP, were each flanked by a protospacer adjacent motif (PAM). (B) Cas9 nuclease introduced a target-specific double-strand DNA cleavage, leading to dissociation of phosphorylated PAM-proximal DNA fragments. Cas9 remained bound to PAM-distal DNA fragments, restricting dA-tailing to PAM-proximal DNA fragments (Gilpatrick et al., 2020; Stangl et al., 2020; Sternberg et al., 2014). Ligation of Nanopore sequencing adapters to dA-tailed DNA ends enabled sequencing in 5' (D7-GFP) or 3' (D6-GFP) direction. Enriched DNA library was sequenced on a MinION sequencer. (C) Nanopore sequencing reads were analyzed by a bioinformatic pipeline, which first mapped raw reads to the GFP-containing proviral genome using Minimap2, and subsequently aligned filtered reads against the host reference genome, which identified the genomic location of the integrated provirus.

Enriched DNA libraries, prepared from 3 µg genomic DNA per GFP+ cell clone, yielded highly variable total read counts, ranging between 4,208 to 62,160 (Figure 12, Table 22). GFP-targeted regions were enriched between 1.3- to 4.5-fold, with an on-target coverage ranging from 1- to 39-fold, with one clone yielding an on-target coverage of 201-fold (E10+o10F6) (Figure 12A, Figure 12B). Interestingly, total read counts did not correlate with on-target coverage nor GFP-enrichment for all analyzed clones. The performance of

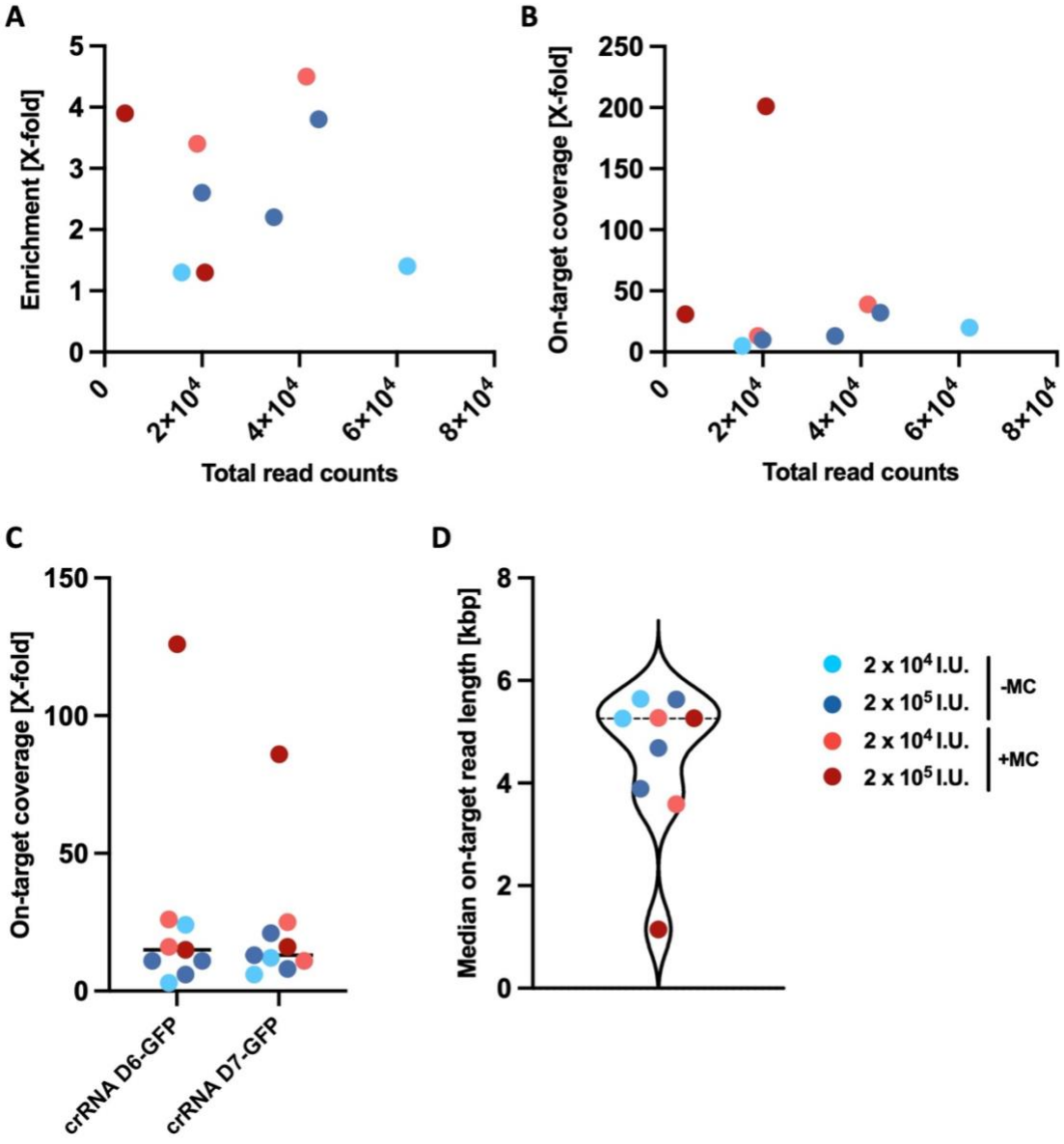

Figure 12: Enrichment, on-target coverage and read-length after CRISPR/Cas9 target enrichment Nanopore sequencing of GFP-expressing (GFP+) cell clones. Murine T cell line S49.1 was transduced with ecotropic, single-round virus-like particles (VLPs), containing an LTR-GFP reporter genome, at low (2 x 10^4 I.U.) or high (2 x 10^5 I.U.) infectious titers in the presence (+) or absence (-) of 1 % methylcellulose (MC). Genomic DNA of nine GFP-expressing (GFP+) cell clones was subjected to CRISPR/Cas9 target enrichment Nanopore sequencing workflow using GFP-targeting crRNAs D6-GFP and D7-GFP. (A) Target enrichment across the mouse genome GRCm38 relative to total read counts for all analyzed GFP+ cell clones. (B) On-target coverage relative to total read counts for all analyzed GFP+ cell clones. (C) On-target read coverage with respect to crRNA D6-GFP and crRNA D7-GFP for all analyzed GFP+ cell clones. (D) Median on-target read length for all analyzed GFP+ cell clones.

crRNAs D6-GFP and D7-GFP was highly comparable as assessed by their relative contribution to on-target reads (Figure 12C). Average read length distribution of on-target reads ranged between 3.6 - 5.6 kbp, whereas one GFP+ cell clone (E10+o10F6) yielded with 1.1 kbp a smaller average fragment length (Figure 12D). Thus, for the majority of GFP+ cell clones, a minimum of ~ 1.5 kbp of each GFP-containing read could be aligned against the mouse reference genome to identify unique LTR-GFP genome integration sites.

Alignment of on-target reads against a host reference genome can be visualized and manually checked using Integrative Genomics Viewer (IGV). Figure 13 displays a representative analysis report for GFP+ cell clone E1-o4D8. After filtering for GFP-containing reads, Nanopore sequencing reads aligned to a genomic region downstream of the *Pds5a* gene on chromosome (chr.) 5 of the mouse reference genome GRCm38. Bi-directional sequencing resulted in reads downstream (crRNA D6-GFP) and upstream (crRNA D7-GFP) from a common breakpoint and identified chr.5:65,699,622 as potential LTR-GFP genome integration site (Figure 13A). Importantly, up- and downstream reads shared a 4 bp overlap at the breakpoint site, which is likely to represent the characteristic target site duplication, flanking an integrated MLV provirus after integration (Figure 13B) (Kim et al., 2010; Serrao & Engelman, 2016).

After filtering raw data for GFP-containing reads, genome mapping identified LTR-GFP integration sites in all analyzed cell clones. Table 22 lists frequency, location, and orientation of LTR-GFP integration sites in GFP+ cell clones based on Cas9-mediated integration site enrichment. Overall, we detected between 1 and 20 unique integration sites per GFP+ cell clone, with seven cell clones harboring multiple LTR-GFP insertions. LTR-GFP genome integration sites were found in different chromosomes. Interestingly, the majority of clones with more than 2 integration events harbored several integration events in the same chromosome (E1+o12B10, E1+5B1, E10-o2A2, E10+4H6, E10+o10F6). Provirus orientations were equally distributed between sense and anti-sense orientations relative to host gene transcription.

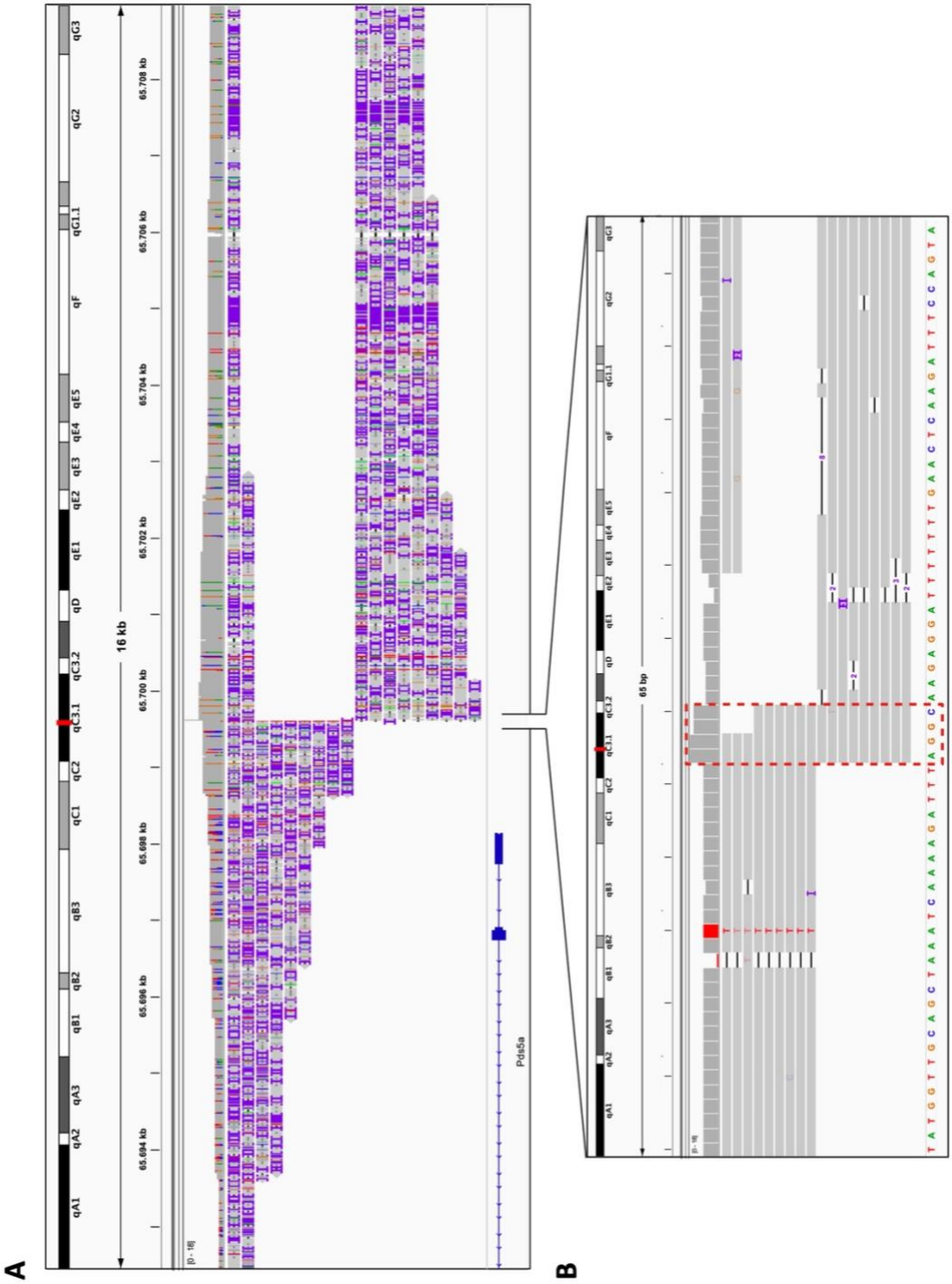

Figure 13: Identification of LTR-GFP genome integration site location by read alignment against the mouse reference genome. After CRISPR/Cas9 target enrichment, applying crRNAs D6-GFP and D7-GFP, and Nanopore sequencing of clone E1-o4D8 (2 x 10^4 I.U., -MC), raw reads were first filtered for LTR-GFP genome mapping and subsequently aligned to the mouse reference genome GRCm38 for identification of LTR-GFP integration site locations. (A) Integrative Genomics Viewer (IGV) allows visualization of mapped reads aligning in up- and downstream direction against chromosome (chr.) 5 and identifies chr.5:65,699,622, down-stream of the *Pds5a* gene, as LTR-GFP integration site location. (B) Magnified illustration of predicted LTR-GFP integration site location at single-nucleotide resolution. At the predicted location of the LTR-GFP integration site, bi-directional reads share a 4 bp overlap (red box), which might represent the characteristic target site duplication of host nucleotides, flanking MLV integration sites (Kim et al., 2010; Serrao & Engelman, 2016).

Table 22: LTR-GFP genome integration sites identified in GFP-expressing (GFP+) cell clones. Genomic DNA of nine GFP-expressing (GFP+) cell clones was subjected to CRISPR/Cas9 target enrichment Nanopore sequencing using GFP-targeting crRNAs D6-GFP and D7-GFP. Nanopore sequencing reads were filtered by mapping to the LTR-GFP genome using Minimap2 and genome-wide LTR-GFP locations were identified by alignment of filtered reads against the mouse reference genome GRCm38.

GFP+ cell clone	Total read counts	Fold enrichment	Mean read length [bp]	LTR-GFP genome integration frequency	On-target reads	Chromosome base location	LTR-GFP genome orientation
E1-o4D8	62,160	1.4x	5,262	1	20	Chr.5: 65,699,622	+
E1-o3F11	15,836	1.3x	5,640	1	5	Chr.12: 118,374,865	+
E1+o12B10	19,012	3.4x	5,272	3	4	Chr.17: 15,830,360	-
					7	Chr.18: 34,859,003	-
					2	Chr.18: 38,604,511	-
E1+5B1	41,428	4.5x	3,589	7	9	Chr.2: 164,696,168	+
					8	Chr.3: 95,282,218	+
					6	Chr.10: 86,391,131	+
					6	Chr.15: 8,563,605	+
					3	Chr.15: 101,233,780	+
					3	Chr.15: 102,206,085	-
					4	Chr.19: 6,549,571	-
E10-1D6	43,968	3.8x	4,685	2	16	Chr.2: 103,680,201	+
					16	Chr.11: 115,901,212	-
E10-o2D1	34,768	2.2x	5,633	3	1	Chr.4: 46,526,790	-
					7	Chr.5: 149,321,386	-
					5	Chr.11: 83,262,278	+
E10-o2A2	19,988	2.6x	3,890	3	4	Chr.1: 155,077,150	-
					2	Chr.1: 180,701,650	-
					4	Chr.13: 48,997,321	+
E10+4H6	4,208	3.9x	5,258	12	2	Chr.3: 89,664,477	+
					3	Chr.7: 124,909,399	-
					3	Chr.10: 21,446,165	-
					2	Chr.10: 26,567,023	-
					3	Chr.11:80,077,594	+
					3	Chr.12: 92,884,413	+
					2	Chr.12: 106,310,364	-
					2	Chr.14: 64,322,595	-
					2	Chr.14: 52,450,466	-
					3	Chr.17: 45,506,550	+
					4	Chr.18: 61,053,342	-
					2	Chr.X: 53,015,969	+
E10+o10F6	20,627	1.3x	1,148	20	20	Chr.1: 64,871,864	+
					8	Chr.1: 133,543,169	-
					17	Chr.2: 92,929,527	+
					13	Chr.2: 128,153,587	-
					10	Chr.2: 11,613,236	-
					6	Chr.2: 31,114,830	-
					9	Chr.4: 138,722,188	-
					12	Chr.5: 137,705,089	+
					9	Chr.6: 142,957,575	-
					11	Chr.10: 116,558,200	+
					10	Chr.11: 115,911,990	-
					3	Chr.11: 115,912,028	-
					3	Chr.11: 70,451,774	+
					15	Chr.12: 102,697,821	+
					9	Chr.12: 85,840,402	-
					13	Chr.18: 80,612,346	+
					12	Chr.18: 15,084,223	-
					9	Chr.19: 41,035,120	+
					7	Chr.19: 44,279,933	-
					15	Chr.X: 53,015,970	+

To evaluate quantitative performance of our approach, we directly compared the frequency of LTR-GFP integration events with GFP+ cell clone characterizations based on GFP MFI and relative GFP copy number (Figure 14). Clones E1-, transduced with low infectious titers in the absence of MC and showing lowest GFP MFIs and relative GFP copy numbers, contained a single copy of the LTR-GFP genome. By contrast, 2 - 7 integration events were identified in GFP+ cell clone subsets E1+ and E10-, which exhibited comparable GFP expression pattern and relative GFP copy numbers. Highest frequencies of LTR-GFP integration events were detected in clones E10+, harboring 12 and 20 LTR-GFP inserts, respectively, which had been transduced with high infectious titers in the presence of 1 % methylcellulose and expressed very high GFP levels and relative GFP copy numbers.

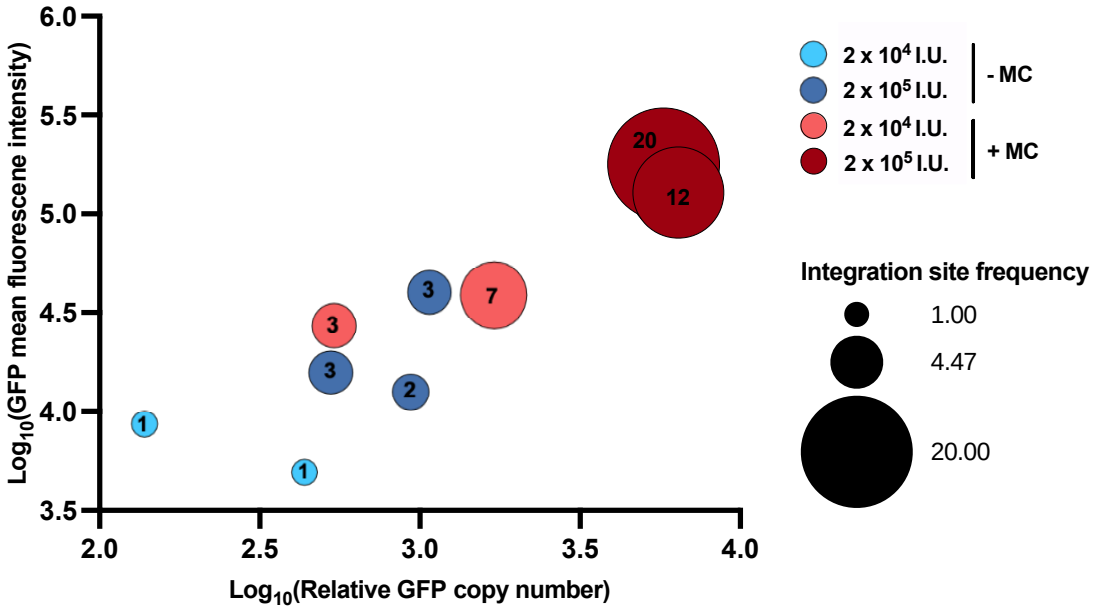

Figure 14: Multivariate analysis of nine GFP-expressing (GFP+) cell clones based on GFP mean fluorescence intensity, relative GFP copy number and frequency of LTR-GFP integration sites. Individual GFP-expressing (GFP+) cell clones were generated by transduction of murine T cell line S49.1 with ecotropic, single-round virus-like particles (VLPs), containing an LTR-GFP reporter genome, at low (2 x 10^4 I.U.) or high (2 x 10^5 I.U.) infectious titers in the presence (+) or absence (-) of 1 % methylcellulose (MC). Expanded GFP+ cell clones were characterized by GFP mean fluorescence intensity using flow cytometry and relative GFP copy number by TaqMan quantitative PCR of extracted genomic DNA. Genomic DNA of nine GFP+ cell clones was subjected to CRISPR/Cas9 target enrichment Nanopore sequencing applying GFP-targeting crRNAs D6-GFP and D7-GFP. Bioinformatic analysis revealed frequency of unique LTR-GFP integration sites in genomic DNA of each analyzed GFP+ cell clone.

In conclusion, our findings demonstrate the capacity of CRISPR/Cas9 target enrichment Nanopore sequencing to quantify and localize GFP-containing target regions in retrovirus-infected cells. Both, the detection of bi-directional reads originating from a common genomic location and the characteristic duplication of 4 bp flanking these locations, support identification of LTR-GFP genome integration sites with high confidence. In addition, the preliminary quantitative evaluation of our workflow demonstrates a correlation of detected LTR-GFP genome integration frequencies with GFP expression levels and relative GFP copy numbers for the analyzed GFP+ cell clones.

3.2.4 Validation of LTR-GFP integration events in GFP+ cell clones by breakpoint-spanning PCR and Sanger sequencing

To validate integration site locations predicted by CRISPR/Cas9 target enrichment Nanopore sequencing, we applied breakpoint-spanning PCR to two potential LTR-GFP integration site locations (chr.1:155,077,150, chr.13:48,997,321) of clone E10-o2A2. To this end, integration sites were amplified by PCR using a combination of primers specific to integration site flanking regions of the host DNA with GFP-specific primers (described in 2.2.1.1, Figure 15A, Figure 15D). Genomic DNA extracted from parental S49.1 cells served as control.

PCR amplification yielded two reaction products of ~ 2,400 bp (PCR product 1) and ~ 1,500 bp (PCR product 2), each specific for clone E10-o2A2 in comparison to parental cell line control (Figure 15B, Figure 15E). Precise sequences of amplified fragments were

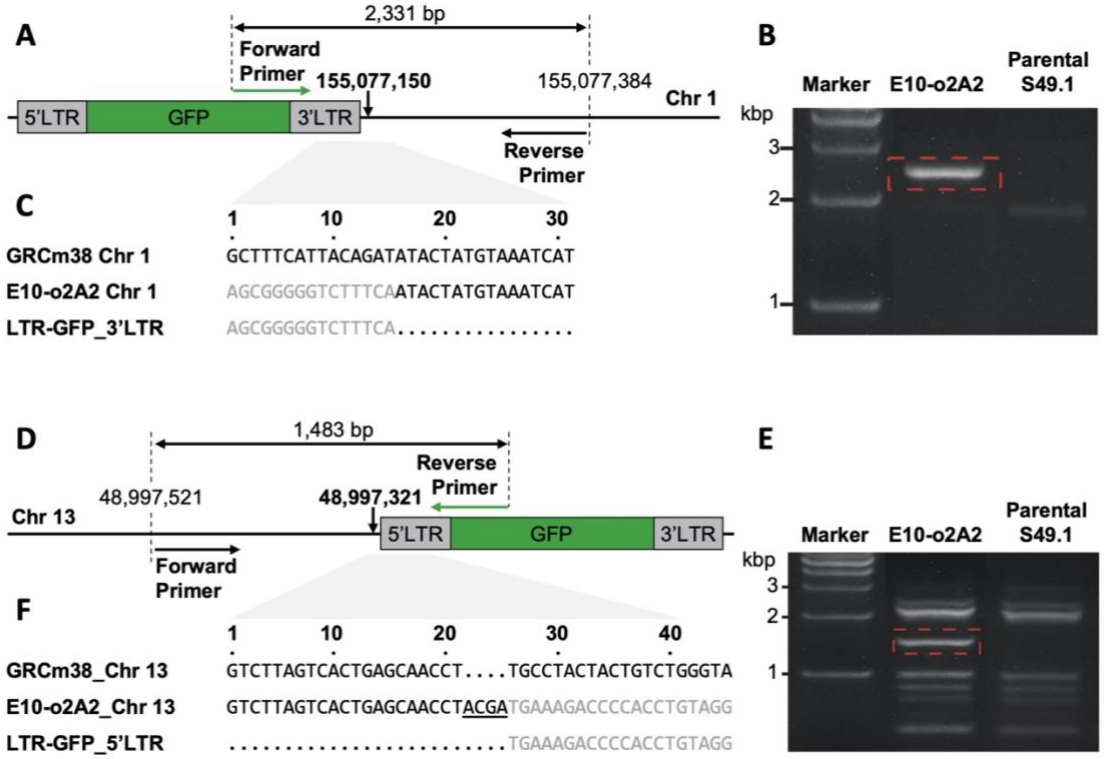

Figure 15: Validation of two predicted LTR-GFP integration site locations by breakpoint-spanning PCR and Sanger sequencing. (A, D) Schematic representation of predicted integration sites in clone E10-o2A2 at chromosome (chr.) base locations chr.1:155,077,150 and chr.13:48,997,321, respectively. Vertical black arrows specify location of predicted integration sites. Black horizontal arrows represent position and orientation of integration site-specific primers, green horizontal arrows symbolize orientation and position of GFP-specific primers designed for breakpoint-spanning PCR. (B, E) Agarose gel electrophoresis after genomic DNA extracted from clone E10-o2A2 was subjected to breakpoint-spanning PCR applying primer combinations assigned in (A, D). PCR amplification revealed a >2 kbp fragment (B, red box) and a >1 kbp fragment (E, red box), each unique for clone E10-o2A2 in comparison to parental S49.1 cells. (C, F) Both DNA fragments were analyzed by Sanger sequencing. Alignment of sequencing data against LTR-GFP genome (grey) and mouse reference genome GRCm38 (black) identified amplified PCR products as LTR-GFP/host junction fragments. Underlined bases indicate 4 bp insertion not aligning to LTR-GFP genome nor GRCm38 reference genome.

determined by Sanger sequencing and subsequently aligned against the LTR-GFP genome and the mouse reference genome GRCm38 (Figure 15C, Figure 15F). We successfully identified PCR product 1 as LTR-GFP/host junction fragment, containing the junction between the 3'LTR of the LTR-GFP genome and flanking host DNA with a breakpoint at chr.1:155,077,150 (Figure 15C). PCR product 2 covers the junction between flanking host DNA and the 5'LTR of the LTR-GFP genome, including predicted integration site chr.13:48,997,321, followed downstream by a 4 bp insertion not aligning to the LTR-GFP genome nor the mouse reference genome (Figure 15F).
Taken together, these findings confirmed LTR-GFP integration site locations reported by CRISPR/Cas9 target enrichment Nanopore sequencing for a selected GFP+ cell clone. In addition, we characterized breakpoint-spanning PCR as an accurate and rapid tool to validate findings by CRISPR/Cas9 target enrichment Nanopore sequencing. Further analysis of the remaining, detected LTR-GFP integration sites is in progress.

3.2.5 Versatile applicability of CRISPR/Cas9 target enrichment Nanopore sequencing

We specifically designed our CRISPR/Cas9 target enrichment Nanopore sequencing workflow to provide high flexibility with respect to the applied retrovirus by targeting regions within the GFP sequence. To demonstrate the broad applicability and detection accuracy of our provirus detection approach, we exemplarily subjected the HIV-transduced cell line J-Lat clone 10.6 to CRISPR/Cas9 target enrichment Nanopore sequencing. J-Lat clone 10.6 is a Jurkat-based, clonal cell line carrying a single proviral copy at chr.9:136,468,579 of replication-incompetent, full-length HIV with a frameshift mutation in *env* and GFP as a substitute for the *nef* gene (Bieniasz & Cullen, 2000; Jordan et al., 2003; Symons et al., 2017).
GFP-target enrichment was performed by applying CRISPR/Cas9-enrichment strategy using crRNAs D6-GFP and D7-GFP to genomic DNA isolated from J-Lat clone 10.6 (described in 2.2.1.4 and 2.2.2.7). After sequencing of enriched DNA library on a MinION sequencer, Nanopore reads were mapped against the full proviral HIV sequence and subsequently aligned against the human reference genome hg38. Bioinformatic analysis detected a potential proviral integration site in gene *SEC16A* at chr.9:136,468,584 by three reads with an average read length of 1,345 bp (Table 23). The

Table 23: Identification of GFP-containing provirus integration sites in J-Lat clone 10.6. Genomic DNA of J-Lat clone 10.6 was subjected to CRISPR/Cas9 target enrichment Nanopore sequencing, using GFP-targeting crRNAs D6-GFP and D7-GFP. Nanopore sequencing reads were filtered by mapping against the proviral genome using Minimap2, and genome-wide provirus locations were identified by alignment of filtered reads against the human reference genome hg38.

Sample name	Total read counts	Mean read length [bp]	On-target reads	Chromosome base location	Provirus orientation
J-Lat 10.6	52,017	1,345	3	Chr.9:136,468,584	-
			1	Chr.9:41,013,247	-
			1	Chr.X:15,327,904	+

identified proviral integration site deviates by only 5 bp from the reported integration site at chr.9:136,468,579 (Symons et al., 2017). In addition, we detected two potential provirus locations at chr.9:41,013,247 and chr.X:15,327,904, each confirmed by only one read.

In conclusion, we successfully confirmed the described HIV provirus location in J-Lat clone 10.6, which emphasizes the accuracy and flexibility of our CRISPR/Cas9 target enrichment Nanopore sequencing workflow. However, our strategy could not resolve the full proviral HIV sequence, due to short average read length and low on-target sequencing coverage. To confirm the existence of two additional integration sites at chr.9:41,013,247 and chr.X:15,327,904 combined application of breakpoint-spanning PCR and Sanger sequencing is required.

3.3 Studying *in vitro* retrovirus transmission using CRISPR/Cas9 target enrichment Nanopore sequencing

3.3.1 Generation of MLV IRES GFP-transduced (MLV IRES GFP+) cell clones

Previous studies documented the co-transmission of multiple retroviral copies from an HIV-infected donor cell to a non-infected target cell during *cis*-infection (Del Portillo et al., 2011; Law et al., 2016; Russell et al., 2013). Yet, both the exact proviral frequency as well as the proviral integration pattern after MLV *cis*-infection remain elusive due to quantitative and qualitative limitations of previously applied provirus detection techniques. Given the quantitative capacity and detection sensitivity of CRISPR/Cas9 target enrichment Nanopore sequencing, we aimed to study efficiency and characteristics of MLV transmission at a proviral level.

To approach this, we first generated MLV-infected cell clones to serve as genetically characterized donor cells in *in vitro* co-culture assays. Therefore, S49.1 cells were transduced with replication-competent MLV IRES GFP in the presence of 1 % methylcellulose (described in 2.2.3.4). After 48 h, allowing for retroviral integration and GFP expression, single cells were isolated by limiting dilution at 0.5 cells/well and treated with reverse transcriptase inhibitor Azidothymidine (AZT) at 1 µM to prevent retroviral integration and therefore spread within the monoclonal population during cellular expansion (Figure 16A). Single MLV IRES GFP-infected (MLV IRES GFP+) cell clones were identified by visual screening for GFP expression and expanded over 6 weeks at 37 °C in the presence of 5 µM AZT. MLV IRES GFP+ cell clones were characterized by GFP MFI using flow cytometry and relative GFP copy number in 10 ng genomic DNA of each individual cell clone was determined using Taqman qPCR (described in 2.2.1.3 and 2.2.4.2).

As depicted in Figure 16B, GFP expression levels were highly variable ranging from 3,873 A.U. (clone 9+C3) to 68,249 A.U. (clone 9+A9), within a selection of eleven MLV IRES GFP+ cell clones. Interestingly, the observed differences in GFP expression levels between individual cell clones were not consistent with relative GFP copy numbers. Except for clone 9+A9 with 1.3×10^3 GFP copies, MLV IRES GFP+ cell clones exhibited highly homogenous relative GFP copy numbers (2.2×10^2 - 3.6×10^2) (Figure 16C). This

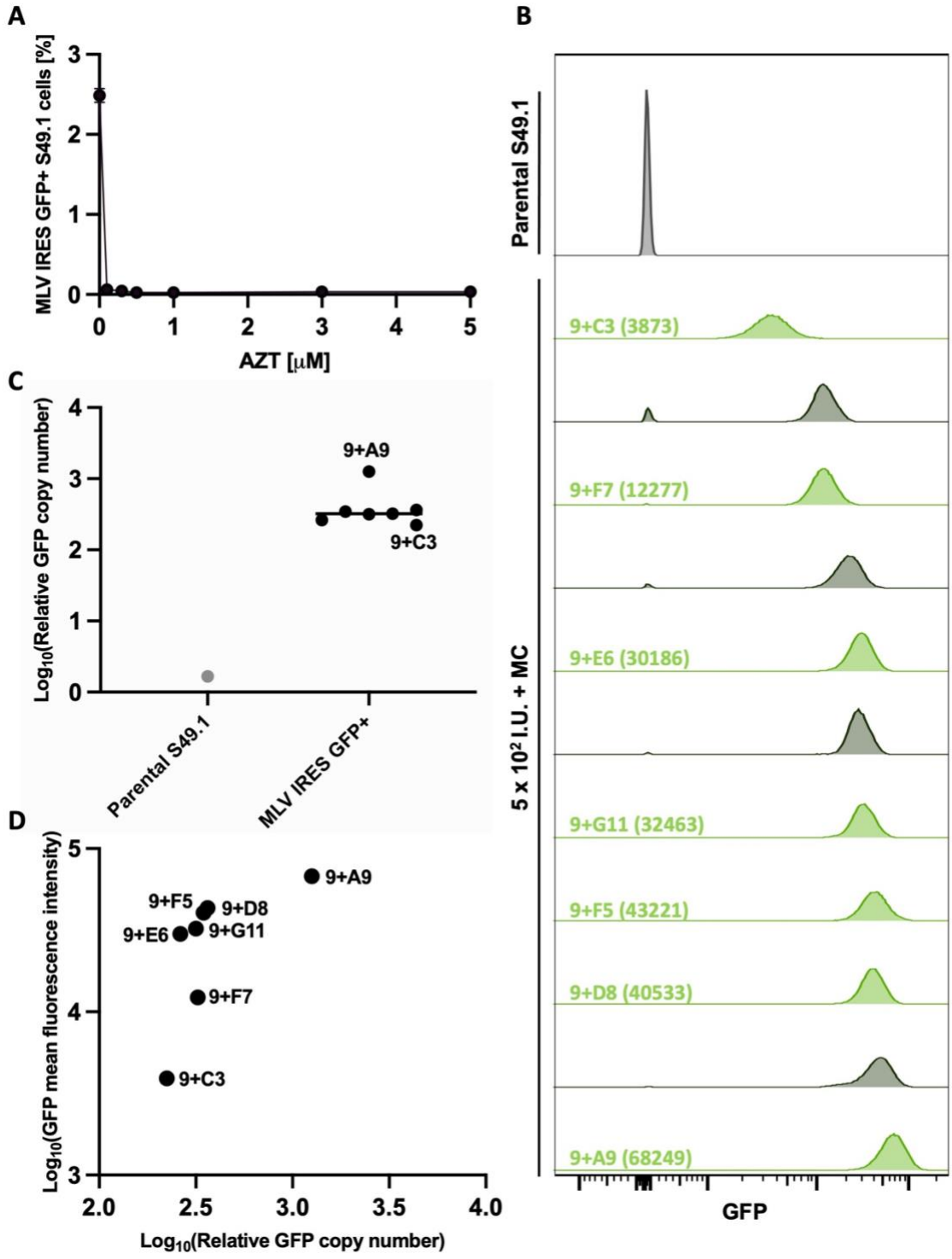

Figure 16: Characterization of MLV IRES GFP-infected (MLV IRES GFP+) cell clones after transduction with replication-competent MLV IRES GFP. Murine T cell line S49.1 was transduced with replication-competent MLV IRES GFP at 5 x 10^2 I.U. in the presence (+) of 1 % methylcellulose (MC). Single cell isolation was performed by limiting dilution (0.5 cells/100 µl) into culture medium supplemented with reverse transcriptase inhibitor Azidothymidine (AZT) at 1 µM. Exclusively GFP-expressing, single cells were expanded over 6 weeks in 5 µM AZT culture medium. (A) Quantification of MLV IRES GFP-infected (MLV IRES GFP+) S49.1 cells in the presence of varying concentrations of AZT. Murine T cell line S49.1 was transduced with 5 x 10^2 I.U. of MLV IRES GFP in culture medium containing 0.1, 0.3, 0.5, 1.0, 3.0 or 5.0 µM AZT. After 48 h, percentage of MLV IRES GFP+ S49.1 cells was assessed by flow cytometry. (B) Flow cytometry analysis of MLV IRES GFP+ cell clones after cellular expansion. Individual cell clones selected for downstream analysis are labeled with MLV IRES GFP+ cell clone name and GFP mean fluorescence intensity in brackets. (C) Relative GFP copy number of each

selected MLV IRES GFP+ cell clone after quantitative PCR of extracted genomic DNA. Clones exhibiting highest (9+A9) and lowest (9+C3) relative GFP copy number are highlighted. (D) Correlation analysis of GFP mean fluorescence intensity and relative GFP copy number of selected MLV IRES GFP+ cell clones. GFP mean fluorescence intensity analysis and cell clone selection is shown in (B) and relative GFP copy number in (C).

is equally reflected in direct comparison of GFP expression levels and relative GFP copy numbers for each MLV IRES GFP+ cell clone (Figure 16D).
Taken together, we successfully generated MLV IRES GFP+ cell clones, characterized by high variances in their GFP expression levels, which can be indicative for differences in provirus frequency and/or location.

3.3.2 Provirus location and frequency influence GFP expression levels in MLV IRES GFP+ cell clones

To identify whether the observed differences in GFP expression levels among MLV IRES GFP+ cell clones can be attributed to variations in provirus frequency and/or a specific chromosomal integration site, we subjected seven MLV IRES GFP+ cell clones to CRISPR/Cas9 target enrichment Nanopore sequencing. For bi-directional targeting, we applied GFP-targeting crRNAs D6-GFP and D7-GFP and included crRNA targeting mouse *Hprt1* as internal control (described in 2.2.1.4). Nanopore sequencing reads were filtered by mapping against the MLV IRES GFP genome and potential integration sites were identified by subsequent read alignment against the mouse reference genome GRCm38.
As bioinformatic analysis was still ongoing at the time of writing, Table 24 displays only integration site analysis for a subset of four MLV IRES GFP+ cell clones. Sequencing of enriched genomic DNA libraries yielded lower average lead lengths, ranging between 773 - 1,333 bp, in comparison to our workflow evaluation (see section 3.2.3). On-target

Table 24: MLV IRES GFP genome integration sites identified in MLV IRES GFP-infected (MLV IRES GFP+) cell clones. Genomic DNA of seven MLV IRES GFP-infected (MLV IRES GFP+) cell clones was subjected to CRISPR/Cas9 target enrichment Nanopore sequencing, using GFP-targeting crRNAs D6-GFP and D7-GFP. Nanopore sequencing reads were filtered by mapping to MLV IRES GFP genome using Minimap2 and genome-wide provirus locations were identified by alignment of filtered reads against mouse reference genome GRCm38.

MLV IRES GFP+ cell clone	On-target reads	Median read length [bp]	Provirus frequency	Chromosome base location	Gene	Provirus orientation
9+A9	1	773	2	Chr.5:34,247,892	*Poln*	+
	1			Chr.17:73,259,584	Non-coding	+
9+D8	6	1,110	1	Chr.5:147,426,187	*Pan3*	+
9+F7	2	1,092	1	Chr.2:128,688,205	*Tmem87b*	+
9+G11	1	1,333	1	Chr.11:59,193,197	Non-coding	-
9+C3	Bioinformatic analysis in progress					
9+E6	Bioinformatic analysis in progress					
9+F5	Bioinformatic analysis in progress					

coverage varied between 1 - 6 reads, whereas majority of reads were directed downstream of the Cas9-cleaved DNA breakpoints, indicating an out-performance of crRNA D6-GFP in comparison to crRNA D7-GFP. With exception of clone 9+A9 harboring two proviruses, all other analyzed MLV IRES GFP+ cell clones contained only one MLV IRES GFP genome each. A detailed positional analysis of detected integration sites revealed that majority of MLV IRES GFP inserts were located in transcriptional units. With one exception, orientation of provirus sequences was in sense-direction relative to host gene transcription. Chromosome-wide distribution of proviral integrations appeared random, yet low proviral frequency per analyzed MLV IRES GFP+ cell clone allowed no detailed analysis of overall preferred chromosomal target locations.

Multivariate analysis provided direct comparison of detected proviral integration site frequencies with GFP MFI and relative GFP copy number for each analyzed MLV IRES GFP+ cell clone (Figure 17). Cell clones 9+D8, 9+F7, 9+G11, which covered a broad range of GFP expression levels (12,277 A.U. to 40,533 A.U), each contained a single copy of MLV IRES GFP at distinct genomic locations. Clone 9+A9, which exhibited by far highest GFP mean fluorescence intensity (68,249 A.U.), was the only clone harboring two provirus copies, based on the current status of our bioinformatic analysis.

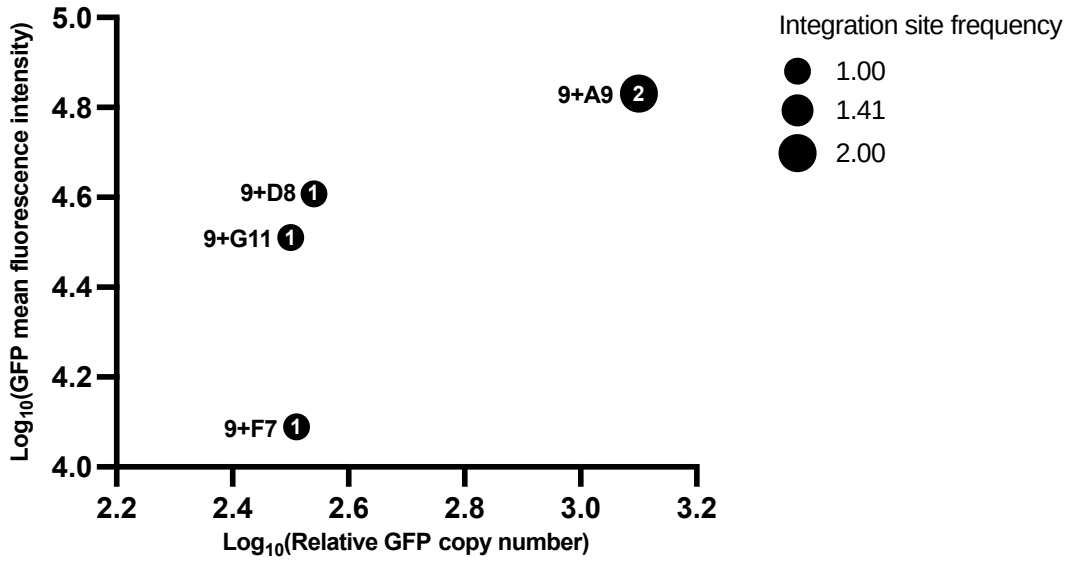

Figure 17: Multivariate analysis visualizing correlation between GFP mean fluorescence intensity, relative GFP copy number and provirus frequency of selected MLV IRES GFP-infected (MLV IRES GFP+) cell clones. MLV IRES GFP-infected (MLV IRES GFP+) cell clones were generated by transduction of S49.1 cells with 5 x 10^2 I.U. of MLV IRES GFP in the presence of 1 % methylcellulose. After single cell isolation in culture medium supplemented with 1 µM AZT to prevent viral spread, single cells were expanded in culture medium supplemented with 5 µM AZT. Expanded MLV IRES GFP+ cell clones were characterized by flow cytometry for GFP expression levels and for relative GFP copy number by TaqMan quantitative PCR of extracted genomic DNA. CRISPR/Cas9 target enrichment Nanopore sequencing was applied to genomic DNA of a selection of MLV IRES GFP+ cell clones, using GFP-targeting crRNAs D6-GFP and D7-GFP. Number of unique integration sites for each analyzed clone was determined by mapping of Nanopore reads to the MLV IRES GFP genome and subsequent alignment of filtered reads to the mouse reference genome GRCm38.

Taken together, our results suggest that provirus expression is associated with provirus frequency and integration site location. However, the effect of other considerable factors, such as promoter methylation status and structural integrity of the provirus, could not be addressed due to short on-target reads, and should be considered in future applications of this workflow.

3.3.3 Efficiency of MLV transmission correlates with GFP expression levels of MLV IRES GFP+ donor cell clones

To quantify retroviral transmission to a target cell population with respect to the proviral characteristics of the MLV-transduced donor cell population, we employed MLV IRES GFP+ cell clones as donor cells in *in vitro* co-culture assays.

Therefore, MLV IRES GFP+ cell clones were each fluorescently labeled with cytoplasmic FarRed dye (described in 2.2.4.3). Subsequently, we co-cultured 1 x 10^5 FarRed-positive MLV IRES GFP-infected cell clones with non-infected parental S49.1 cells at a 1:2 ratio. After cultivation over 48 h in AZT-free culture medium, the number of MLV IRES GFP-infected S49.1 target cells was determined as the percentage of GFP-expressing cells among the FarRed-negative cell population using flow cytometry (described in 2.2.4.2).

With exception of one MLV IRES GFP+ cell clone, co-cultivation resulted in MLV IRES GFP transmission from all donor cell populations to parental S49.1 target cells (Figure 18). Interestingly, the number of infected target cells increased with rising GFP expression levels of the respective monoclonal donor cell population. Starting from clone 9+C3 (3,873 A.U.), which yielded no detectable infection of S49.1 cells, co-culture with clone 9+A9 (68,249 A.U.) resulted in highest MLV IRES GFP target cell infection rate, with an average of 1.53 %. In addition, cell clone 9+A9 was the only clone harboring two

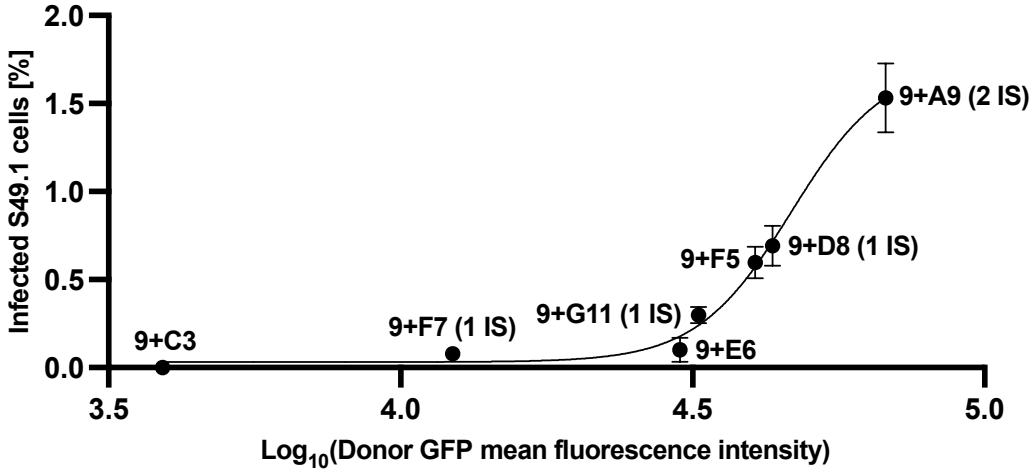

Figure 18: MLV transmission positively correlates with GFP expression levels of MLV IRES GFP-infected (MLV IRES GFP+) donor cells. MLV IRES GFP-infected (MLV IRES GFP+) cell clones (names highlighted), each characterized for GFP mean fluorescence intensity using flow cytometry and selectively characterized for their provirus integration site (IS) frequency (number in brackets), were co-cultured with S49.1 cells for 48 h. MLV transmission rate was determined by quantification of GFP-expressing S49.1 cells using flow cytometry. Correlation analysis illustrates the number of MLV-infected S49.1 target cells relative to GFP mean fluorescence intensity of MLV IRES GFP+ donor cell clones (n=3).

integrated copies of the MLV IRES GFP genome, whereas all other analyzed MLV IRES GFP+ cell clones contained a single integrated provirus, according to the current status of our bioinformatic analysis.

Taken together, we successfully validated retroviral transmission from MLV IRES GFP+ donor cells, genetically characterized for their provirus frequencies and locations, to non-infected target cells. Tendencies of a potential correlation between donor cell GFP expression levels and retroviral transmission rate need to be validated in future experiments with a higher number of donor cells, covering a broad range of proviral frequencies and locations.

4. Discussion

After entering the host organism, retroviruses like HIV and MLV disseminate to lymphoid tissues to establish retroviral infection (Haase, 2011; Murooka et al., 2012; Sewald et al., 2012, 2015). Despite their clinical relevance, the mechanisms involved in local retrovirus transmission during early time points after infection are incompletely understood (Haase, 2011; Haugh et al., 2021; Murooka et al., 2012; Sewald et al., 2016). *In vivo* studies of retroviral infection describe the formation of stable cell-cell contacts in secondary lymphoid tissues, which are indicative for local retroviral spread across virological synapses (Law et al., 2016; Murooka et al., 2012; Sewald et al., 2012). Further findings indicate that retroviruses initiate and/or stabilize intercellular contacts, as the frequency of cell-cell contacts is reduced during infection with mutant HIV and MLV deficient in the Env glycoprotein or harboring non-functional Env (Law et al., 2016; Sewald et al., 2012). However, due to the application of non-infectious virus mutants, the role of intercellular contacts in retrovirus transmission *in vivo* has yet to be established. In this study, we focus on the cell adhesion-mediating proteins LFA1 and ICAM1 to examine the contribution of stable cell-cell contacts to local retroviral transmission *in vitro* and *in vivo*.

4.1 LFA1 and ICAM1 are crucial for MLV cell-to-cell transmission *in vitro* and *in vivo*

Adhesive interactions between LFA1 and its ligand ICAM1 mediate a variety of immunological functions, such as T cell migration, T cell activation and target cell killing (Walling & Kim, 2018). The formation of immunological synapses *in vitro* and *in vivo* requires engagement of ICAM1 with LFA1 at the cell-cell contact site (Donnadieu et al., 1994; Dustin et al., 1997; Dustin, 2009; Monks et al., 1998; Negulescu et al., 1996; Scholer et al., 2008; Sims et al., 2007; Vasiliver-Shamis et al., 2010). There, assembly of interacting LFA1/ICAM1 complexes in ringlike structures around central microclusters of TCR/pMHC complexes provides structural stability, which enhances antigen sensitivity of T cells during TCR-engagement (Bachmann et al., 1997; Vasiliver-Shamis et al., 2010; Walling & Kim, 2018). Interestingly, the cell-cell interface during formation of virological synapses between retrovirus-infected donor cells and non-infected target cells resembles the morphology and structural organization of immunological synapses (Jolly et al., 2004, 2007b; Len et al., 2017; Starling & Jolly, 2016; Vasiliver-Shamis et al., 2008, 2010). Yet, the role of LFA1 and ICAM1 in retrovirus transmission during *cis*-infection *in vivo* has yet to be established.

To study the function of LFA1 and ICAM1 in MLV *cis*-infection, we first established an *in vitro* co-culture assay with primary cells of physiological relevance, which support cell contact-dependent MLV transmission (see section 3.1.1, Figure 6). Both cellular subsets, murine B1 and FoxP3+ T cells, expressed LFA1 and ICAM1 at high levels irrespective of the infection status of the cell (see section 3.1.2, Figure 7A). Using CD11a- and ICAM1-blocking antibodies, we aimed to identify the functional role of LFA1 and ICAM1 in cell contact-dependent MLV transmission. Although our findings indicated a contribution of LFA1 to cell contact-dependent retroviral transfer *in vitro*, this

experimental approach delivered no conclusive results (see section 3.1.2, Figure 7B). Whereas MLV transmission from infected B1 donor cells to non-infected FoxP3+ T target cells was significantly reduced in the presence of CD11a-blocking antibodies, antibody-mediated blocking of ICAM1 reduced MLV spread only marginally. This is in line with previous studies, which report conflicting findings after employing antibody blocking of LFA1 and ICAM1 to assess their role in retroviral cell-to-cell transmission. Whereas some *in vitro* studies document absence of cell conjugate formation and reduced cell contact-dependent transmission of HIV in the presence of LFA1- and ICAM1-blocking antibodies, others describe enhanced contact-dependent HIV transmission upon LFA1 blocking (Jolly et al., 2007b; Puigdomènech et al., 2008). Potential causes of these contradicting results include usage of different cellular models, which coincide with varying expression levels of LFA1 and ICAM1, and application of antibodies targeting distinct epitopes (Bracq et al., 2018). Within our antibody-blocking assay, we specifically considered cellular proliferation and simultaneous blocking of donor and target cell proteins as critical factors potentially impairing significance of our findings.

Given the above-mentioned limitations of antibody-blocking experiments, we applied a different approach to distinguish the contribution of LFA1 and ICAM1 to MLV cell-to-cell transmission with respect to the donor and target cell population. By application of primary cells derived from WT C57BL/6, CD11a-KO (LFA1-deficient) and ICAM1-KO mice, we could demonstrate a functional role of polarized LFA1/ICAM1 interaction for efficient MLV transmission during *in vitro cis*-infection. Whereas LFA1 expression was critical on target cells, the presence of ICAM1 was essential on MLV-infected donor cells for efficient cell contact-dependent MLV transmission (see section 3.1.2, Figure 7C - Figure 7F). Importantly, we confirmed these findings for two distinct donor cell types, B1 and FoxP3+ T cells. Our findings are supported by previous studies applying LFA1-deficient or -mutant cell lines to demonstrate requirement of LFA1 expression for efficient HIV cell-to-cell transfer (Hioe et al., 2001; Jolly et al., 2007b). Consistently, silencing of ICAM1 significantly impaired efficiency of HIV transmission during *trans*-infection (Wang et al., 2009). The importance of LFA1 for HIV transmission by dendritic cells during *trans*-infection could further be demonstrated using *in vitro* co-cultures with T cells isolated from Leukocyte Adhesion Deficiency type 1 (LAD-1) patients, expressing non-functional LFA1 (Groot et al., 2006).

There has been no experimental demonstration of the role of LFA1 and ICAM1 in cell contact-dependent retroviral transmission *in vivo.* In this study, we validate for the first time a contributive effect of polarized LFA1/ICAM1 interaction to efficient retroviral spread in the lymph nodes of living mice (see section 3.1.3, Figure 8B and Figure 19). Our *in vivo* model identified LFA1 expression on target cells as a critical factor for efficient MLV *cis*-infection of cellular subsets by MLV-infected FoxP3+ T cells. On MLV-infected donor cells, expression of ICAM1 was crucial for efficient retroviral transmission by *cis*-infection *in vivo*. Unfortunately, due to the limited number of B1 cells derived from the murine peritoneal cavity, we had to restrict our *in vivo* experiments to adoptive transfer of MLV-transduced FoxP3+ T cells. In the future, we should particularly focus on adoptive transfer of MLV-infected B1 cells to validate our initial findings for a broad spectrum of cell populations relevant for MLV infection *in vivo*.

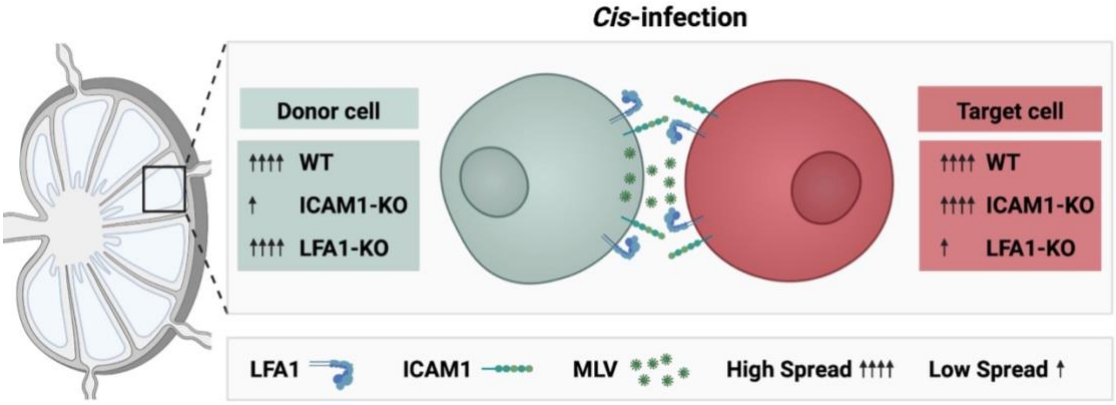

Figure 19: Polarized LFA1/ICAM1 interaction is critical for efficient retroviral *cis*-infection *in vivo*. Schematic representation of efficient cell contact-dependent retroviral transmission in popliteal lymph nodes of MLV-infected mice according to the findings within our study. Adoptive transfer experiments allowed combination of WT C57BL/6, ICAM1-knockout (KO) and CD11a-KO (LFA1-deficient) mice to assess the function of LFA1 and ICAM1 on donor and target cells separately. Efficient MLV transmission during *cis*-infection required expression of LFA1 on non-infected target cells and ICAM1 expression on MLV-infected donor cells, suggesting that polarized interaction between LFA1 and ICAM1 supports retroviral cell-to-cell transmission *in vivo*.

The biological reason for the polarized nature of LFA1/ICAM1 interaction during efficient cell contact-dependent retroviral transmission is still speculative. One may hypothesize that besides their structural function in the stabilization of cell-cell interactions, adhesion protein-mediated signaling cascades might actively support formation of the virological synapse and retroviral integration (Starling & Jolly, 2016). For instance, in HTLV-infected cells, the retroviral protein Tax and ICAM1-induced signaling cascades synergistically initiate MTOC polarization during formation of the virological synapse (Barnard et al., 2005; Nejmeddine et al., 2009). Previous studies further indicate that signaling cascades triggered by binding of LFA1 to ICAM1 mediate T cell activation and thereby might increase target cell susceptibility (Lebedeva et al., 2005; Van Seventer et al., 1990). Since MLV requires disassembly of the nucleus during mitosis for successful integration, activation-induced proliferation defines susceptibility of a target cell for MLV infection (Roe et al., 1993). Functional studies identified that LFA1 engagement by ICAM1 releases a co-stimulating signal supporting activation of resting T cells (Van Seventer et al., 1990). In addition, TCR/pMHC engagement during formation of immunological synapses was shown to activate a PI3Kδ-dependent signaling cascade, which enhances LFA1-mediated adhesion and thereby determines efficiency of T cell activation (Deng & Huttenlocher, 2012; Garçon & Okkenhaug, 2016; Scholer et al., 2008; Stadtmann et al., 2011; Vasiliver-Shamis et al., 2010; Walling & Kim, 2018). Since the immunological and the virological synapse share a variety of structural characteristics, one may speculate that LFA1/ICAM1-interaction also influences target cell activation during cell contact-dependent retroviral transmission. In the future, analysis of signaling cascades downstream of the cell adhesion-mediating proteins LFA1 and ICAM1 might dissect the functional role of polarized LFA1/ICAM1 interaction during retroviral *cis*-infection on a molecular level.

4.2 Consequences of cell contact-dependent retroviral transmission

Cell contact-dependent transmission of HIV is suggested to have critical implications for retroviral pathogenesis (Jolly, 2011). The efficient transmission of HIV at the cell-cell interaction site generates a high local MOI, exceeding efficiency of cell-free transmission by 100- to 10,000-fold (Bracq et al., 2018; Carr et al., 1999; Chen et al., 2007; Del Portillo et al., 2011; Dimitrov et al., 1993; Martin et al., 2010; Sourisseau et al., 2007). Transfer of multiple retroviral copies during cell contact-dependent HIV transmission has been documented to reduce effectiveness of various antiretroviral drugs and cellular restriction factors (Agosto et al., 2014; Jolly et al., 2010; Richardson et al., 2008; Sigal et al., 2011). Previous studies provide evidence for HIV multicopy integration in the genome of target cells after cell contact-dependent retroviral transmission, potentially promoting HIV genetic diversity through enhanced frequency of recombination events (Del Portillo et al., 2011; Jung et al., 2002; Law et al., 2016; Russell et al., 2013). Despite its clinical relevance, previous transmission studies have only provided limited insights into the concept of multicopy infection as a result of cell contact-dependent retroviral transmission (Del Portillo et al., 2011; Law et al., 2016; Russell et al., 2013). General limitations of previously applied experimental approaches include a restricted quantitative detection range and the absence of provirus integration site analysis (Del Portillo et al., 2011; Law et al., 2016; Liszewski et al., 2009; Russell et al., 2013). Given the limitations of established provirus detection techniques, we aimed to establish a novel approach to simultaneously localize and quantify proviruses in the genome of a host cell.

4.3 CRISPR/Cas9 target enrichment Nanopore sequencing - a novel approach for provirus detection in retrovirus-infected cells

In this study, we established CRISPR/Cas9 target enrichment Nanopore sequencing as a novel provirus detection technique resolving proviral genomes and their adjacent host DNA sequences (see section 3.2.3, Figure 11). Our approach combines the sequence-specificity of the CRISPR/Cas9 system for selective enrichment of a defined genomic target region with the long-read capacity of Nanopore sequencing. Subsequent bioinformatic processing of Nanopore-acquired data identifies location, frequency, and orientation of proviral target regions within the host genome. Along with the evaluation of CRISPR/Cas9 target enrichment Nanopore sequencing on a library of GFP+ cell clones (see section 3.2.1, Figure 9), we identified various advantages emphasizing the potential of our novel approach as compared to established provirus detection techniques.

4.3.1 Advantages of CRISPR/Cas9 target enrichment Nanopore sequencing

4.3.1.1 CRISPR/Cas9 system allows for rapid and amplification-free target enrichment

One of the main challenges in measuring retroviral integration is the detection of the small retroviral genome (~9 x 10^3 bp) against a predominant background of human genomic DNA (~3 x 10^9 bp) (Miyazato et al., 2016). To reduce sequencing capacity occupied by non-informative reads, established techniques apply PCR-based amplification or selective capture by DNA probe-mediated hybridization to specifically enrich for proviral

genomes prior sequencing (Burgess, 2021; Iwase et al., 2019; Schmidt et al., 2001, 2007; Van Haasteren et al., 2021). However, both strategies harbor various limitations impairing provirus detection accuracy. PCR-based enrichment methods are prone to stochastic amplification, sensitive to structural variations in primer-binding sites and represent the main source for GC-bias in enriched DNA libraries, whereas DNA probe-based enrichment techniques struggle with low yield and capture of long DNA fragments (Aird et al., 2011; Gilpatrick et al., 2020; Kebschull & Zador, 2015; Kozarewa et al., 2015; Miyazato et al., 2016; Van Haasteren et al., 2021). In addition, both provirus enrichment techniques are frequently combined with NGS platforms, which require either restriction enzyme digest or sonication for DNA fragmentation, further introducing a detection bias (Harkey et al., 2007; Van Haasteren et al., 2021). Besides, these techniques are laborious and require up to seven days (e.g. LAM-PCR) for provirus detection (Van Haasteren et al., 2021).

In this study, we established CRISPR/Cas9 target enrichment Nanopore sequencing as an amplification-free provirus detection technique. Our approach overcomes the need for PCR-based amplification by enrichment of a selected proviral target region in genomic DNA through incubation with crRNA/Cas9 ribonucleoprotein complexes (see section 3.2.3, Figure 11). Upon target sequence recognition, Cas9 introduces a double strand cleavage, allowing for selective adapter ligation and targeted Nanopore sequencing of long, native DNA fragments. Thereby, CRISPR/Cas9 target enrichment Nanopore sequencing is not prone to bias introduced by target site amplification nor DNA library preparation as compared to various established provirus detection techniques (Van Haasteren et al., 2021). Remarkably, despite the lack of PCR-based amplification, our approach reaches high target enrichment efficiency as demonstrated by target site-specific enrichment and the number of on-target reads after analysis of genomic DNA purified from GFP+ cell clones (see section 3.2.3, Figure 12). In our study, mapped target enrichment accounted for 1.3- to 4.5-fold with an on-target coverage between 1- to 39-fold, while one clone reached 201-fold.

This is in line with recent studies applying CRISPR/Cas9 target enrichment Nanopore sequencing, which reached comparable target enrichment, while their documented on-target coverage was equally characterized by a high variance. Thus, Gilpatrick et al. documented mapping of 1.8 % of reads against a specific target region with an on-target read depth ranging from 18-fold to 846-fold, whereas Stangl et al. reached an average on-target coverage of 68-fold (Gilpatrick et al., 2020; Stangl et al., 2020). In addition to its enrichment efficiency, CRISPR/Cas9 target enrichment Nanopore sequencing is an unprecedentedly rapid provirus detection technique. From purification of genomic DNA to bioinformatic analysis, our approach can complete provirus detection within less than 48 h, and is thereby outperforming various established provirus detection techniques (Falcinelli et al., 2019; Gaebler et al., 2019; Liszewski et al., 2009).

4.3.1.2 CRISPR/Cas9 enrichment strategy enables targeting of multiple regions

A particular challenge for established provirus detection techniques is the high mutation rate of retroviruses (Achaz et al., 2004; Coffin, 1995; Liszewski et al., 2009; Palmer et al., 2005). For instance, in HIV-1 infected adults, proviral variants constitute between

60 – 71 % of the proviral landscape, with an average of 14 % defective sequences (Abrahams et al., 2019; Brodin et al., 2016; Brooks et al., 2020). Provirus detection assays based on restriction-enzyme digest, such as linker ligation PCR or inverse PCR, are particularly impaired by mutations within the proviral sequence, which result in diminished workflow performance, likely leaving variant proviral sequences undetected (Liszewski et al., 2009; Wells et al., 2020). Other provirus detection assays, such as quadruplex qPCR (Q4PCR), address this challenge by targeting of up to four highly conserved regions, such as ψ, *env*, *pol* and *gag*, however these techniques are prone to overestimate the replication-competent reservoir due to inter-patient provirus variance (Bruner et al., 2019; Falcinelli et al., 2019; Gaebler et al., 2019).

In contrast to above-mentioned approaches, CRISPR/Cas9-mediated enrichment is highly flexible with respect to the number of target regions. Protocols by Oxford Nanopore enable simultaneous targeting of multiple regions of interest without significant impairment of individual on-target restriction efficiencies (Bruijnesteijn et al., 2021). This strongly suggests that simultaneous targeting of a wide range of proviral sequences using a multi-target CRISPR/Cas9 enrichment strategy could compensate for deficient enrichment of individual variant target sites and capture the full diversity of the proviral landscape, including intact and defective proviruses.

For our approach, we specifically designed CRISPR/Cas9 strategy to enrich for GFP-containing retroviruses integrated in genomic host DNA (see section 3.2.2, Figure 10). By targeting a single region within the proviral genome, our approach provided high flexibility regarding the applied retrovirus. In addition, GFP fluorescence intensity could be utilized as quantitative readout to assess the relative number of integrated and intact GFP copies in a monoclonal population, thus allow for workflow evaluation during its establishment. However, proviruses harboring mutations in the GFP regions targeted for CRISPR/Cas9 enrichment could potentially remain undetected within our current approach. In the future, an extension of our workflow to multi-target CRISPR/Cas9-mediated enrichment could particularly address heterogeneity of provirus variants, emphasizing the detection accuracy and flexibility of our approach.

4.3.1.3 CRISPR/Cas9 target enrichment Nanopore sequencing provides quantitative provirus detection

Various established provirus detection techniques approach provirus quantification through endpoint dilution analysis combined with near-full length sequencing (Hiener et al., 2017; Patro et al., 2019). Although these methods provide reliable estimates of the proviral reservoir, their application is laborious and too extensive for clinical applications (Levy et al., 2021). To test applicability of CRISPR/Cas9 target enrichment Nanopore sequencing for quantitative provirus detection, Van Haasteren et al. applied this technique to a polyclonal population of HEK 293T cells transduced with replication-competent HIV (Van Haasteren et al., 2021). There, frequencies of detected proviruses varied by 92 % between samples due to viral transmission during cell cultivation, complicating assessment of this workflow for its quantitative robustness and data reproducibility.

To overcome these limitations in our study, we evaluated quantitative performance of CRISPR/Cas9 target enrichment Nanopore sequencing on a monoclonal set of transduced

S49.1 cell clones, each harboring stable insertions of replication-incompetent LTR-GFP genomes at variable frequencies (see section 3.2.1, Figure 9). To assess differences in the frequency of provirus integrations, we characterized each GFP+ cell clone using two independent GFP reporter read-outs, GFP expression and relative GFP copy number. With these reliable estimates, our monoclonal cell model provided a robust tool to assess quantitative performance of CRISPR/Cas9 target enrichment Nanopore sequencing in detection of integrated LTR-GFP genomes in genomic DNA of GFP+ cell clones.

For quantification of LTR-GFP integration sites, we applied CRISPR/Cas9 target enrichment to genomic DNA of nine GFP+ cell clones and sequenced each enriched DNA library on a single Nanopore flow cell (see section 3.2.3, Figure 11). Using IGV for on-target read alignment against the mouse reference genome, we identified a widespread genomic distribution between 1 and 20 unique LTR-GFP integration sites per individual GFP+ cell clone (see section 3.2.3, Table 22). Most importantly, we were able to detect discrete differences in the frequency of LTR-GFP genome integrations, which were positively correlated with previous characterizations of GFP+ cell clones based on GFP expression levels and relative GFP copy number (see section 3.2.3, Figure 14). Taken together, these findings strongly indicate that CRISPR/Cas9 target enrichment Nanopore sequencing provides a quantitative evaluation of LTR-GFP genomes integrated in the genome of transduced GFP+ cell clones.

In the future, we should additionally determine the quantitative detection limit of CRISPR/Cas9 target enrichment Nanopore sequencing. To quantify the maximal provirus detection capacity of CRISPR/Cas9 target enrichment and Nanopore sequencing individually in upcoming experiments, we could sequentially increase the number of LTR-GFP inserts by pooling genomic DNA from different, characterized GFP+ cell clones. Analysis of samples pooled prior CRISPR/Cas9-mediated enrichment or prior Nanopore sequencing could provide insights into the maximal quantitative detection capacity of each technique and identify potential limiting factors.

4.3.1.4 Bi-directional long-read Nanopore sequencing resolves the chromosomal integration site and the provirus sequence

Recent advancements in the detection of integrated proviruses aim to identify both the integration site and the corresponding proviral sequence (Hiener et al., 2017; Patro et al., 2019). A critical determinant of confident provirus sequence detection and localization by genomic mapping is the on-target read length (Asogawa et al., 2020; De Roeck et al., 2019; Van Haasteren et al., 2021; Vondrak et al., 2020). Established provirus detection techniques based on Next Generation Sequencing platforms reach an average total read-length of 200 - 400 bp, which includes parts of the proviral sequence (Besser et al., 2018; Paruzynski et al., 2010; Van Haasteren et al., 2021). While these approaches can assess replication-competence of a provirus, they frequently fail to identify the integration site location (Einkauf et al., 2019; Falcinelli et al., 2019; Hiener et al., 2017; Lee et al., 2017; Wang & Palmer, 2018). In addition, short reads negatively affect confidence of sequence alignment against the host reference genome and impair sensitive provirus detection, particularly in repetitive regions (Asogawa et al., 2020; De Roeck et al., 2019; Van Haasteren et al., 2021; Vondrak et al., 2020).

In our study, we strongly benefit from the long-read sequencing capacities of Oxford Nanopore devices. For LTR-GFP integration site identification, we applied CRISPR/Cas9 target enrichment to genomic DNA of nine GFP+ cell clones and sequenced each enriched DNA library on a single Nanopore flow cell (see section 3.2.3, Figure 11). By targeting integrated LTR-GFP genomes at opposite ends, CRISPR/Cas9 enrichment generated DNA fragments containing parts of the GFP sequence and either the proviral 5'LTR or 3'LTR and its flanking host DNA sequence. Following analysis by long-read Nanopore sequencing, average on-target read length ranged between 3.6 - 5.6 kbp (see section 3.2.3, Figure 12). Consequently, after subtraction of regions taken up by the LTR-GFP sequence (944 bp (D6-GFP), 2,079 bp (D7-GFP)) a minimum of ~ 1.5 kbp could be mapped against the mouse reference genome. In combination with bi-directional sequencing in 5' and 3' direction, our approach yielded high-confidence genomic alignment for identification of LTR-GFP integration sites in the host cell genome (see section 3.2.3, Figure 13, Table 22).
Given that retroviral genome sizes range between 7 - 10 kbp, future optimizations of our approach should focus on increasing on-target read length to resolve the full proviral genome and its flanking host DNA to identify potential proviral mutations or structural variations with high confidence (Ryu, 2017). Here, processing parameters, such as DNA extraction, library preparation and delivery into the Nanopore device, should be assessed for their potential to increase the length of Nanopore-acquired reads in future experiments (Amarasinghe et al., 2020; Gilpatrick et al., 2020; Jain et al., 2018).

4.3.1.5 CRISPR/Cas9 target enrichment Nanopore sequencing provides highly accurate provirus detection

In this study, we determined provirus detection accuracy of CRISPR/Cas9 target enrichment Nanopore sequencing through several approaches. A first indication for detection confidence of CRISPR/Cas9 target enrichment Nanopore sequencing was provided by IGV analysis of identified LTR-GFP genome integration sites in GFP+ cell clones (see section 3.2.3, Figure 13). Strikingly, a large majority of bi-directional on-target reads overlapped for 4 bp at detected LTR-GFP integration sites. This phenomenon is likely to represent the characteristic host target site duplication of 4 bp during MLV integration into the host genome and strongly supported provirus detection confidence of our workflow (Kim et al., 2010). Second, we directly validated selected LTR-GFP integration site locations by breakpoint-spanning PCR in combination with Sanger sequencing (see section 3.2.4, Figure 15). Each integration site required design of a location site-specific primer pair, which in combination with GFP-specific primers exclusively amplified on-site integrated LTR-GFP genomes. Remarkably, Sanger sequencing of host/LTR-GFP junction regions from two randomly selected integration sites corresponded precisely to location predicted by CRISPR/Cas9 target enrichment Nanopore sequencing, demonstrating provirus detection accuracy of our approach. Finally, we extended validation of CRISPR/Cas9 target enrichment Nanopore sequencing provirus detection accuracy from our monoclonal cell model to a described cellular model. For this purpose, we applied our provirus detection workflow to J-Lat cells, which represent one of the best-characterized *in vitro* HIV latency models (Hakre et al., 2012;

Jordan et al., 2003). By subjecting genomic DNA of J-Lat clone 10.6 to CRISPR/Cas9 target enrichment Nanopore sequencing, we successfully confirmed the described provirus integration site at chr.9:136,468,579 by three on-target reads (see section 3.2.5, Table 23) (Jordan et al., 2003; Symons et al., 2017). In conclusion, all three evaluation approaches indicated high provirus detection confidence of CRISPR/Cas9 target enrichment Nanopore sequencing.

Nevertheless, the techniques utilized within this study to evaluate accuracy of CRISPR/Cas9 target enrichment Nanopore sequencing are rather laborious and only applicable to confirm previously detected integration site locations. Future assessments of CRISPR/Cas9 target enrichment Nanopore sequencing should additionally characterize its general qualitative accuracy as compared to NGS-based provirus detection techniques. A commonly raised concern associated with long-read Nanopore sequencing is its comparatively low read accuracy. Although read identities of Nanopore-acquired data already improved from formerly 60 % to more than 90 % due to generation of optimized pores and base calling algorithms, their accuracy is still inferior in comparison to NGS platforms with an average error rate of 0.24 ± 0.06 % per base (Artesi et al., 2021; Goodwin et al., 2015; Pfeiffer et al., 2018; Wick et al., 2019). To specifically define read accuracy of our workflow, we could subject selected GFP+ cell clones to an established NGS-based provirus detection technique, and directly compare NGS- with Nanopore-acquired reads for each predicted LTR-GFP integration site.

4.3.2 Limitations of CRISPR/Cas9 target enrichment Nanopore sequencing

During evaluation of CRISPR/Cas9 target enrichment Nanopore sequencing as a novel provirus detection technique, we further identified a few technical limitations, which should be specifically considered for future optimizations of our workflow.

4.3.2.1 CRISPR/Cas9 off-target effects potentially impair optimal provirus detection

First limitations of CRISPR/Cas9 target enrichment Nanopore sequencing arise in the number of suitable and distinctive CRISPR/Cas9-restriction sites within a specific target sequence. Critical considerations in the design of optimal crRNAs comprise the presence of PAM regions, predicted on-target specificity and off-target activity with respect to the host genome (Gilpatrick et al., 2020; Stangl et al., 2020; Van Haasteren et al., 2021).

In our study, we designed three crRNAs for optimal targeting of GFP sequences integrated in both the murine and human host genome (see section 3.2.2, Figure 10). As computational prediction models can vary from experimental restriction efficiency and specificity, we determined individual crRNA/Cas9 activity through LTR-GFP plasmid digestion and established breakpoint-spanning TaqMan qPCR to quantify Cas9-mediated restriction. The latter quantified crRNA D5-GFP/Cas9 restriction efficiency with ~ 90 %, indicating high on-target restriction efficiency (see section 3.2.2, Figure 10E). However, our approach did not evaluate restriction efficiency with respect to distinct host species. This impact was particularly evident during CRISPR/Cas9-mediated target enrichment of HIV provirus sequences integrated in the human host genome. Thus, CRISPR/Cas9 target enrichment Nanopore sequencing of the human-derived J-Lat clone 10.6 yielded

low on-target coverage and detected two additional potential integration sites (chr.9:41,013,247, chr.X:15,327,904), each confirmed by one on-target read, which will require further investigation (see section 3.2.5, Table 23). A potential reason, besides sample processing, could be increased off-target effects within the human genome as compared to the mouse genome, impairing efficient CRISPR/Cas9-mediated targeting of integrated proviruses. This is in line with previous studies, indicating that incomplete on-target exonuclease digestion and off-target binding are among the main causes for reduced yield and coverage of target loci (Gilpatrick et al., 2020; Van Haasteren et al., 2021; Wallace et al., 2021).

In conclusion, these findings highlight the significance of assessing crRNA/Cas9 restriction specificity prior sequencing by both computational and experimental approaches with respect to the host species. While computational screening for species-specific off-target sites represents an essential tool during design of optimal crRNAs, TaqMan qPCR can provide a fast and reliable practical assessment of crRNA/Cas9 on-target restriction performance in future experiments.

4.3.2.2 High DNA input requirements limit applicability of CRISPR/Cas9 target enrichment Nanopore sequencing

In vivo, latently HIV-infected cells harboring replication-competent, integrated proviral DNA constitute less than one in 10^7 cells (Chun et al., 1997). This low frequency of integrated proviral DNA within a cellular population represents one of the general challenges for accurate quantification of proviral integration (Chun et al., 1997; Liszewski et al., 2009).

Unfortunately, the amount of input genomic DNA is one of the major limitations of CRISPR/Cas9 target enrichment Nanopore sequencing (Van Haasteren et al., 2021). Due to the lack of PCR-based amplification, CRISPR/Cas9 target enrichment Nanopore sequencing requires an input between 1 - 10 µg genomic DNA (Stangl et al., 2020; Van Haasteren et al., 2021). By comparison, other provirus detection techniques, such as LAM-PCR, reach a sensitivity comparable to CRISPR/Cas9-based approaches with only 0.5 - 1 µg input genomic DNA (Paruzynski et al., 2010; Van Haasteren et al., 2021). Although genomic DNA input requirements were met by our monoclonal cell model, this limiting factor must be considered for future applications based on CRISPR/Cas9 target enrichment, such as analysis of single cells in HIV latency studies.

Most recently, Stangl et al. expanded applicability of CRISPR/Cas9 target enrichment Nanopore sequencing to low input DNA by subjecting small amounts of genomic material to whole genome amplification (WGA) prior target enrichment (Stangl et al., 2020). However, WGA of low input DNA is prone to introduce an amplification bias, which could significantly impair sensitive provirus detection within our study (Deleye et al., 2017). In the future, our monoclonal cell model could serve as a valuable tool to assess the extent of amplification bias generated by WGA of single cell genomes (see section 3.2.1, Figure 9). For instance, we could subject DNA of a GFP+ cell clone after WGA and genomic DNA purified from the same GFP+ cell clone, expanded to a monoclonal population, to CRISPR/Cas9 target enrichment Nanopore sequencing. By comparing workflow performance with respect to the distinct genomic input material, we could gain

insights into DNA quality, genome coverage, provirus detection reliability and reproducibility after WGA and potentially expand our workflow to accurate provirus detection in low input genomic material.

In conclusion, we successfully established CRISPR/Cas9 target enrichment Nanopore sequencing as a novel, amplification-free provirus detection technique. Within this study, we demonstrate for the first time the capacity of CRISPR/Cas9 target enrichment Nanopore sequencing to simultaneously quantify and localize integrated proviral genomes within less than 48 h, which exceeds the performance of various established provirus detection techniques. With its target sequence flexibility, on-target specificity, and provirus detection accuracy, CRISPR/Cas9 target enrichment Nanopore sequencing has the potential to serve as a crucial technique for retrovirus research in the future.

4.4 Studying retroviral transmission using CRISPR/Cas9 target enrichment Nanopore sequencing

4.4.1 Provirus integration site location and/or frequency influence retroviral transmission efficiency

Given the capacities of CRISPR/Cas9 target enrichment Nanopore sequencing, we aimed to study the effects of MLV integration site frequency and distribution within the genome of a donor cell on retroviral transmission. Previous studies indicate that the expression of viral proteins alters the phenotype of infected host cells and thereby represents a critical determinant of retroviral transmission and pathogenesis (Law et al., 2016; Murooka et al., 2012; Sewald et al., 2012). For instance, during HIV infection, expression of viral proteins is suggested to influence the motility of infected host cells (Law et al., 2016; Murooka et al., 2012; Stolp et al., 2012). Thus, the average migratory velocity of HIV-infected cells in splenocytes and lymph nodes of humanized mice is significantly reduced in comparison to non-infected cells (Law et al., 2016; Murooka et al., 2012). Importantly, the phenotype of an HIV-infected cell can be a vital determinant of HIV dissemination from local infection to a systemic level. While motile HIV-infected cells migrate to other organs to promote systemic dissemination, within lymphoid tissues, reduced motility of a subpopulation of HIV-infected cells might support local spread across long-lasting, stable cell-cell contacts (Law et al., 2016; Murooka et al., 2012). However, previous HIV transmission studies have not experimentally addressed the contribution of the full proviral landscape of an infected host cell to the efficiency of retroviral spread.
For this purpose, we generated MLV IRES GFP+ cell clones and determined their proviral characteristics by CRISPR/Cas9 target enrichment Nanopore sequencing (see sections 3.3.1, 3.3.2, Figure 16 and Table 24). Each clone harbored between 1 - 2 integrated proviral copies, which were predominantly located in transcriptional units, consistent with previous studies characterizing MLV target site preferences (Bushman et al., 2005; Rohdewohld et al., 1987; Scherdin et al., 1990). Using monoclonal MLV IRES GFP+ cell clones as donor cells in *in vitro* co-culture assays, we revealed that the frequency and/or location of proviruses integrated in the genome of the donor cell critically influence MLV transmission efficiency to the target cell population (see section

3.3.3, Figure 18). Thus, the number of MLV-infected target cells increased with rising GFP expression levels of the donor cell population. Further, co-culture using a MLV IRES GFP+ cell clone harboring two proviral copies as donor yielded higher cell MLV transmission rates than donor cells containing a single proviral copy. Thus, our findings suggest that location and/or frequency of proviruses integrated in the genome of a host cell influence retrovirus production and/or retroviral transmission efficiency.
At this point, it is important to note, that due to ongoing bioinformatic analysis and the small number of MLV IRES GFP+ cell clones, each harboring only 1 - 2 proviruses with no information about their structural integrity, our results represent preliminary data and only allow for contingent conclusions. To support our findings, further analysis with MLV-transduced donor cells, containing a broad quantitative range of integrated proviral copies, and transcriptomic analysis of retroviral protein expression are indispensable. Nonetheless, our findings indicate that defining an MLV-infected donor cell exclusively based on its infection status might fail to capture its full complexity. Other potential factors include provirus location, frequency, orientation, replication-competence, and silencing, and should equally be characterized for their impact on retrovirus protein expression and the phenotype of a host cell. In the future, the combination of CRISPR/Cas9 target enrichment Nanopore sequencing with transcriptomic analysis and live cell imaging could provide unprecedented insights into the effects of the proviral landscape of a host cell on its dynamic behavior, morphological characteristics, virus production and its likelihood to establish stable cell-cell contacts to allow for local retroviral transmission.

4.4.2 Challenging the concept of multicopy provirus integration using CRISPR/Cas9 target enrichment Nanopore sequencing

In vivo, multiply infected cells are consistently detectable in spleen, lymphoid tissue and blood of HIV-infected patients (Gratton et al., 2000; Josefsson et al., 2011; Josefsson et al., 2013; Jung et al., 2002; Law et al., 2016). While the majority of HIV-infected cells *in vivo* contains a single proviral copy, between 5 - 7 % of CD4+ T cells harbor multiple proviral copies (Josefsson et al., 2011; Josefsson et al., 2013; Law et al., 2016). In human splenic tissue of HIV-infected patients, multiply HIV-infected cells carry an average of 3 - 4 genetically variant proviruses (Gratton et al., 2000; Jung et al., 2002; Law et al., 2016). Yet, the reason for multicopy infection of single cells remains elusive. Previous studies challenging the concept of multicopy infection as a consequence of retroviral cell-to-cell transmission applied multicolor infection strategies, FISH and quantitative PCR assays (Del Portillo et al., 2011; Law et al., 2016; Russell et al., 2013). However, these experimental approaches share various limitations in their quantitative and qualitative provirus detection. While application of distinct fluorescent retroviral constructs demonstrated that cell contact-dependent transmission increased the frequency of HIV multicopy infection within the target cell population, this experimental approach limits confident provirus quantification to two intact, expressed proviruses (Del Portillo et al., 2011; Law et al., 2016). Provirus detection by FISH, on the other hand, covers a broad quantitative range of integrated and non-integrated proviruses, yet the precise provirus integration site remains unspecified (Del Portillo et al., 2011; Russell et al., 2013).

Given the capacities of CRISPR/Cas9 target enrichment Nanopore sequencing, we aimed to determine whether MLV transmission across virological synapses during *cis*-infection results in multicopy MLV infection. Unfortunately, due to time limitations within our study, we could not proceed to analyze the proviral profile of a target cell population after cell contact-dependent MLV transmission. Irrespective of the above, our preliminary findings demonstrate the potential provided by CRISPR/Cas9 target enrichment Nanopore sequencing in combination with *in vitro* co-culture assays for future experiments (see section 3.3.3, Figure 18). In consideration of the limitations of previous transmission studies, an extension of our initial workflow could provide unprecedented insights into the consequences of cell contact-dependent retroviral transmission for infection of the target cell. For instance, using a DNA barcoded MLV library would allow for unique identification of each provirus within an infected host cell (Chen et al., 2018). In combination with our *in vitro cis*-infection assay, we could trace transmission of each individual provirus from an infected donor cell population to the target cell population. Using CRISPR/Cas9 target enrichment Nanopore sequencing, we could determine efficiency of cell contact-dependent transmission in comparison to cell-free transmission over a broad quantitative range. Besides its quantitative capacities, the application of a barcoded MLV library could provide insights into the dynamics of MLV in multiply infected cells. During HIV infection, multiploid inheritance of genetically distinct proviral genomes is suggested to increase the genetic diversity of virus progeny through genetic recombination (Dang et al., 2004; Del Portillo et al., 2011; Dixit & Perelson, 2005; Gratton et al., 2000; Jung et al., 2002; Levy et al., 2004). Recombinant HIV variants are prone to develop resistance to multidrug ART or to evade both the innate and adaptive immune response (Blackard et al., 2002; Dixit & Perelson, 2005; Rambaut et al., 2004). Consequently, an understanding of the molecular mechanisms driving multiploid inheritance is crucial to study evolution of HIV and improve the effectiveness of ART (Dixit & Perelson, 2005). Here, the combined application of our *in vitro cis*-infection assay with CRISPR/Cas9 target enrichment Nanopore sequencing could provide a prospective tool to challenge the concept of multicopy infection after cell contact-dependent retroviral transmission and identify potential therapeutic targets for antiretroviral drug development.

4.5 Future perspectives

Within this study, we emphasized the potential of CRISPR/Cas9 target enrichment Nanopore sequencing with a particular focus on cell contact-dependent transmission of retroviruses. Yet, our novel provirus detection approach could be extended to other areas in retrovirus research. For instance, CRISPR/Cas9 target enrichment Nanopore sequencing could be applied for HIV latency studies. Methylation of integrated proviral DNA, in particular LTR promoter methylation, is one of the primary mediators of retroviral silencing (Hakre et al., 2012; Shalginskikh et al., 2013). However, the exact mechanisms underlying initiation and maintenance of retroviral silencing are poorly understood. As CRISPR/Cas9 target enrichment leaves genomic DNA in its native state, it conserves DNA modifications such as cytosine methylation pattern, which can be detected in Nanopore sequencing devices based on electric signals (Gigante et al., 2019;

Gilpatrick et al., 2020; Ni et al., 2019; Simpson et al., 2017). Strikingly, differential CpG promoter methylation analysis by CRISPR/Cas9 target enrichment Nanopore sequencing reaches a detection sensitivity comparable to established DNA methylation status approaches such as whole genome bisulfite sequencing (Gilpatrick et al., 2020). Thus, CRISPR/Cas9 target enrichment Nanopore sequencing could provide insights into the promoter methylation status of individual proviruses within a selected host in a rapid and accurate way. Furthermore, our approach could be expanded for analysis of host factors involved in control of HIV infection. Recent studies identified a positive correlation between hypermethylation of antiviral host factors and interferon-stimulated genes with viral load in HIV-infected individuals (Oriol-Tordera et al., 2020). As CRISPR/Cas9 target enrichment Nanopore sequencing allows for simultaneous enrichment of distinct ROIs, our technique could be expanded for multi-targeting of proviral sequences and host factors involved in retroviral infection control. Hence, with its flexible target selection, high target enrichment specificity and capacity to reveal epigenetic pattern, CRISPR/Cas9 target enrichment Nanopore sequencing has the potential to serve as a crucial technique in diagnostic and clinical HIV research in the future.

References

Abrahams, M. R., Joseph, S. B., Garrett, N., Tyers, L., Moeser, M., Archin, N., Council, O. D., Matten, D., Zhou, S., Doolabh, D., Anthony, C., Goonetilleke, N., Karim, S. A., Margolis, D. M., Pond, S. K., Williamson, C., & Swanstrom, R. (2019). The replication-competent HIV-1 latent reservoir is primarily established near the time of therapy initiation. *Science Translational Medicine*, *11*(513), eaaw5589. https://doi.org/10.1126/scitranslmed.aaw5589

Achaz, G., Palmer, S., Kearney, M., Maldarelli, F., Mellors, J. W., Coffin, J. M., & Wakeley, J. (2004). A robust measure of HIV-1 population turnover within chronically infected individuals. *Molecular Biology and Evolution*, *21*(10), 1902–1912. https://doi.org/10.1093/molbev/msh196

Agosto, L. M., Herring, M. B., Mothes, W., & Henderson, A. J. (2018). HIV-1-Infected CD4+ T Cells Facilitate Latent Infection of Resting CD4+ T Cells through Cell-Cell Contact. *Cell Reports*, *24*(8), 2088–2100. https://doi.org/10.1016/j.celrep.2018.07.079

Agosto, L. M., Zhong, P., Munro, J., & Mothes, W. (2014). Highly Active Antiretroviral Therapies Are Effective against HIV-1 Cell-to-Cell Transmission. *PLoS Pathogens*, *10*(2), e1003982. https://doi.org/10.1371/journal.ppat.1003982

Agüera-Gonzalez, S., Bouchet, J., & Alcover, A. (2015). Immunological Synapse. In *eLS* (pp. 1–9). John Wiley & Sons. https://doi.org/10.1002/9780470015902.a0004027.pub2

Aird, D., Ross, M. G., Chen, W. S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D. B., Nusbaum, C., & Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biology*, *12*(2), R18. https://doi.org/10.1186/gb-2011-12-2-r18

Albritton, L. M., Tseng, L., Scadden, D., & Cunningham, J. M. (1989). A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell*, *57*(4), 659–666. https://doi.org/10.1016/0092-8674(89)90134-7

Alvarez, R. A., Barría, M. I., & Chen, B. K. (2014). Unique Features of HIV-1 Spread through T Cell Virological Synapses. *PLoS Pathogens*, *10*(12), e1004513. https://doi.org/10.1371/journal.ppat.1004513

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, *21*(1), 30. https://doi.org/10.1186/s13059-020-1935-5

Artesi, M., Hahaut, V., Cole, B., Lambrechts, L., Ashrafi, F., Marçais, A., Hermine, O., Griebel, P., Arsic, N., van der Meer, F., Burny, A., Bron, D., Bianchi, E., Delvenne, P., Bours, V., Charlier, C., Georges, M., Vandekerckhove, L., Van den Broeke, A., & Durkin, K. (2021). PCIP-seq: simultaneous sequencing of integrated viral genomes and their insertion sites with long reads. *Genome Biology*, *22*(1), 97. https://doi.org/10.1186/s13059-021-02307-0

Asogawa, M., Ohno, A., Nakagawa, S., Ochiai, E., Katahira, Y., Sudo, M., Osawa, M., Sugisawa, M., & Imanishi, T. (2020). Human short tandem repeat identification using a nanopore-based DNA sequencer: a pilot study. *Journal of Human Genetics*, *65*(1), 21–24. https://doi.org/10.1038/s10038-019-0688-z

Bachmann, M. F., McKall-Faienza, K., Schmits, R., Bouchard, D., Beach, J., Speiser, D. E., Mak, T. W., & Ohashi, P. S. (1997). Distinct Roles for LFA-1 and CD28 during Activation of Naive T Cells: Adhesion versus Costimulation. *Immunity*, *7*(4), 549–557. https://doi.org/10.1016/S1074-7613(00)80376-3

Balasubramaniam, M., & Freed, E. O. (2011). New insights into HIV assembly and trafficking. *Physiology*, *26*(4), 236–251. https://doi.org/10.1152/physiol.00051.2010

Barnard, A. L., Igakura, T., Tanaka, Y., Taylor, G. P., & Bangham, C. R. M. (2005). Engagement of specific T-cell surface molecules regulates cytoskeletal polarization in HTLV-1-infected lymphocytes. *Blood*, *106*(3), 988–995. https://doi.org/10.1182/blood-2004-07-2850

Bastarache, S. M., Mesplède, T., Donahue, D. A., Sloan, R. D., & Wainberg, M. A. (2014). Fitness impaired drug resistant HIV-1 is not compromised in cell-to-cell transmission or establishment of and reactivation from latency. *Viruses*, *6*(9), 3487–3499. https://doi.org/10.3390/v6093487

Battini, J. L., Rasko, J. E. J., & Miller, A. D. (1999). A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: Possible role in G protein-coupled signal transduction. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(4), 1385–1390. https://doi.org/10.1073/pnas.96.4.1385

Besser, J., Carleton, H. A., Gerner-Smidt, P., Lindsey, R. L., & Trees, E. (2018). Next-generation sequencing technologies and their application to the study and control of bacterial infections. In *Clinical Microbiology and Infection* (Vol. 24, Issue 4, pp. 335–341). https://doi.org/10.1016/j.cmi.2017.10.013

Bieniasz, P. D., & Cullen, B. R. (2000). Multiple Blocks to Human Immunodeficiency Virus Type 1 Replication in Rodent Cells. *Journal of Virology*, *74*(21), 9868–9877.

https://doi.org/10.1128/JVI.74.21.9868-9877.2000

Blackard, J. T., Cohen, D. E., & Mayer, K. H. (2002). Human immunodeficiency virus superinfection and recombination: Current state of knowledge and potential clinical consequences. *Clinical Infectious Diseases*, *34*(8), 1108–1114. https://doi.org/10.1086/339547

Blanco, J., Bosch, B., Fernández-Figueras, M. T., Barretina, J., Clotet, B., & Esté, J. A. (2004). High level of coreceptor-independent HIV transfer induced by contacts between primary CD4 T cells. *Journal of Biological Chemistry*, *279*(49), 51305–51314. https://doi.org/10.1074/jbc.M408547200

Bobardt, M. D., Saphire, A. C. ., Hung, H. C., Yu, X., Van Der Schueren, B., Zhang, Z., David, G., & Gallay, P. A. (2003). Syndecan captures, protects, and transmits HIV to T lymphocytes. *Immunity*, *18*(1), 27–39. https://doi.org/10.1016/S1074-7613(02)00504-6

Bracq, L., Xie, M., Benichou, S., & Bouchet, J. (2018). Mechanisms for cell-to-cell transmission of HIV-1. *Frontiers in Immunology*, *9*(FEB), 260. https://doi.org/10.3389/fimmu.2018.00260

Brady, T., Kelly, B. J., Male, F., Roth, S., Bailey, A., Malani, N., Gijsbers, R., O'Doherty, U., & Bushman, F. D. (2013). Quantitation of HIV DNA integration: Effects of differential integration site distributions on Alu-PCR assays. *Journal of Virological Methods*, *189*(1), 53–57. https://doi.org/10.1016/j.jviromet.2013.01.004

Brandenberg, O. F., Rusert, P., Magnus, C., Weber, J., Böni, J., Günthard, H. F., Regoes, R. R., & Trkola, A. (2014). Partial rescue of V1V2 mutant infectivity by HIV-1 cell-cell transmission supports the domain's exceptional capacity for sequence variation. *Retrovirology*, *11*(1), 75. https://doi.org/10.1186/s12977-014-0075-y

Brodin, J., Zanini, F., Thebo, L., Lanz, C., Bratt, G., Neher, R. A., & Albert, J. (2016). Establishment and stability of the latent HIV-1 DNA reservoir. *ELife*, *5*(November2016), e18889. https://doi.org/10.7554/eLife.18889

Brooks, K., Jones, B. R., Dilernia, D. A., Wilkins, D. J., Claiborne, D. T., McInally, S., Gilmour, J., Kilembe, W., Joy, J. B., Allen, S. A., Brumme, Z. L., & Hunter, E. (2020). HIV-1 variants are archived throughout infection and persist in the reservoir. *PLoS Pathogens*, *16*(6), e1008378. https://doi.org/10.1371/journal.ppat.1008378

Brown, P. O. (1997). Integration. In J. M. Coffin, S. H. Hughes, & H. E. Varmus (Eds.), *Retroviruses*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. https://www.ncbi.nlm.nih.gov/books/NBK19392/

Brown, P. O., Bowerman, B., Varmus, H. E., & Bishop, J. M. (1989). Retroviral integration: Structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(8), 2525–2529. https://doi.org/10.1073/pnas.86.8.2525

Bruijnesteijn, J., van der Wiel, M., de Groot, N. G., & Bontrop, R. E. (2021). Rapid Characterization of Complex Killer Cell Immunoglobulin-Like Receptor (KIR) Regions Using Cas9 Enrichment and Nanopore Sequencing. *Frontiers in Immunology*, *12*(September), 722181. https://doi.org/10.3389/fimmu.2021.722181

Bruner, K. M., Hosmane, N. N., & Siliciano, R. F. (2015). Towards an HIV-1 cure: Measuring the latent reservoir. *Trends in Microbiology*, *23*(4), 192–203. https://doi.org/10.1016/j.tim.2015.01.013

Bruner, K. M., Murray, A. J., Pollack, R. A., Soliman, M. G., Laskey, S. B., Capoferri, A. A., Lai, J., Strain, M. C., Lada, S. M., Hoh, R., Ho, Y. C., Richman, D. D., Deeks, S. G., Siliciano, J. D., & Siliciano, R. F. (2016). Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nature Medicine*, *22*(9), 1043–1049. https://doi.org/10.1038/nm.4156

Bruner, K. M., Wang, Z., Simonetti, F. R., Bender, A. M., Kwon, K. J., Sengupta, S., Fray, E. J., Beg, S. A., Antar, A. A. R., Jenike, K. M., Bertagnolli, L. N., Capoferri, A. A., Kufera, J. T., Timmons, A., Nobles, C., Gregg, J., Wada, N., Ho, Y. C., Zhang, H., … Siliciano, R. F. (2019). A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature*, *566*(7742), 120–125. https://doi.org/10.1038/s41586-019-0898-8

Burgess, D. J. (2021). Complex targeted sequencing in real time. *Nature Reviews Genetics*, *22*(2), 67. https://doi.org/10.1038/s41576-020-00324-6

Bushman, F., Lewinski, M., Ciuffi, A., Barr, S., Leipzig, J., Hannenhalli, S., & Hoffmann, C. (2005). Genome-wide analysis of retroviral DNA integration. *Nature Reviews Microbiology*, *3*(11), 848–858. https://doi.org/10.1038/nrmicro1263

Cameron, P. U., Freudenthal, P. S., Barker, J. M., Gezelter, S., Inaba, K., & Steinman, R. M. (1992). Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells. *Science*, *257*(5068), 383–387. https://doi.org/10.1126/science.1352913

Campi, G., Varma, R., & Dustin, M. L. (2005). Actin and agonist MHC-peptide complex-dependent T cell receptor microclusters as scaffolds for signaling. *Journal of Experimental Medicine*, *202*(8), 1031–1036. https://doi.org/10.1084/jem.20051182

Carr, J. M., Hocking, H., Li, P., & Burrell, C. J. (1999). Rapid and efficient cell-to-cell transmission of human immunodeficiency virus infection from monocyte-derived macrophages to peripheral blood

lymphocytes. *Virology*, *265*(2), 319–329. https://doi.org/10.1006/viro.1999.0047

Carvajal-Rodríguez, A., Crandall, K. A., & Posada, D. (2007). Recombination favors the evolution of drug resistance in HIV-1 during antiretroviral therapy. *Infection, Genetics and Evolution*, *7*(4), 476–483. https://doi.org/10.1016/j.meegid.2007.02.001

Chen, H. C., Zorita, E., & Filion, G. J. (2018). Using Barcoded HIV Ensembles (B-HIVE) for Single Provirus Transcriptomics. *Current Protocols in Molecular Biology*, *122*(1), e56. https://doi.org/10.1002/cpmb.56

Chen, H., Wei, S. Q., & Engelman, A. (1999). Multiple integrase functions are required to form the native structure of the human immunodeficiency virus type I intasome. *Journal of Biological Chemistry*, *274*(24), 17358–17364. https://doi.org/10.1074/jbc.274.24.17358

Chen, P., Hübner, W., Spinelli, M. A., & Chen, B. K. (2007). Predominant Mode of Human Immunodeficiency Virus Transfer between T Cells Is Mediated by Sustained Env-Dependent Neutralization-Resistant Virological Synapses. *Journal of Virology*, *81*(22), 12582–12595. https://doi.org/10.1128/jvi.00381-07

Chomont, N., El-Far, M., Ancuta, P., Trautmann, L., Procopio, F. A., Yassine-Diab, B., Boucher, G., Boulassel, M. R., Ghattas, G., Brenchley, J. M., Schacker, T. W., Hill, B. J., Douek, D. C., Routy, J. P., Haddad, E. K., & Sékaly, R. P. (2009). HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature Medicine*, *15*(8), 893–900. https://doi.org/10.1038/nm.1972

Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., Kuo, Y. H., Brookmeyer, R., Zeiger, M. A., Barditch-Crovo, P., & Siliciano, R. F. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, *387*(6629), 183–188. https://doi.org/10.1038/387183a0

Chun, T. W., Finzi, D., Margolick, J., Chadwick, K., Schwartz, D., & Siliciano, R. F. (1995). In vivo fate of HIV-1-infected T cells: Quantitative analysis of the transition to stable latency. *Nature Medicine*, *1*(12), 1284–1290. https://doi.org/10.1038/nm1295-1284

Coffin, J. M. (1995). HIV population dynamics in vivo: Implications for genetic variation, pathogenesis, and therapy. *Science*, *267*(5197), 483–489. https://doi.org/10.1126/science.7824947

Comrie, W. A., Babich, A., & Burkhardt, J. K. (2015). F-actin flow drives affinity maturation and spatial organization of LFA-1 at the immunological synapse. *Journal of Cell Biology*, *208*(4), 475–491. https://doi.org/10.1083/jcb.201406121

Craigie, R., & Bushman, F. D. (2012). HIV DNA Integration. *Cold Spring Harbor Perspectives in Medicine*, *2*(7), a006890–a006890. https://doi.org/10.1101/cshperspect.a006890

Dale, B. M., McNerney, G. P., Thompson, D. L., Hubner, W., De Los Reyes, K., Chuang, F. Y. S., Huser, T., & Chen, B. K. (2011). Cell-to-cell transfer of HIV-1 via virological synapses leads to endosomal virion maturation that activates viral membrane fusion. *Cell Host and Microbe*, *10*(6), 551–562. https://doi.org/10.1016/j.chom.2011.10.015

Dang, Q., Chen, J., Unutmaz, D., Coffin, J. M., Pathak, V. K., Powell, D., KewalRamani, V. N., Maldarelli, F., & Hu, W. S. (2004). Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(2), 632–637. https://doi.org/10.1073/pnas.0307636100

de Fougerolles, A. R., & Springer, T. A. (1992). Intercellular adhesion molecule 3, a third adhesion counter-receptor for lymphocyte function-associated molecule 1 on resting lymphocytes. *Journal of Experimental Medicine*, *175*(1), 185–190. https://doi.org/10.1084/jem.175.1.185

De Roeck, A., De Coster, W., Bossaerts, L., Cacace, R., De Pooter, T., Van Dongen, J., D'Hert, S., De Rijk, P., Strazisar, M., Van Broeckhoven, C., & Sleegers, K. (2019). NanoSatellite: Accurate characterization of expanded tandem repeat length and sequence through whole genome long-read sequencing on PromethION. *Genome Biology*, *20*(1), 239. https://doi.org/10.1186/s13059-019-1856-3

Del Portillo, A., Tripodi, J., Najfeld, V., Wodarz, D., Levy, D. N., & Chen, B. K. (2011). Multiploid Inheritance of HIV-1 during Cell-to-Cell Infection. *Journal of Virology*, *85*(14), 7169–7176. https://doi.org/10.1128/JVI.00231-11

Deleye, L., Tilleman, L., Van Der Plaetsen, A. S., Cornelis, S., Deforce, D., & Van Nieuwerburgh, F. (2017). Performance of four modern whole genome amplification methods for copy number variant detection in single cells. *Scientific Reports*, *7*(1), 3422. https://doi.org/10.1038/s41598-017-03711-y

Deng, Q., & Huttenlocher, A. (2012). Leukocyte migration from a fish eye's view. *Journal of Cell Science*, *125*(17), 3949–3956. https://doi.org/10.1242/jcs.093633

Dimitrov, D. S., Willey, R. L., Sato, H., Chang, L. J., Blumenthal, R., & Martin, M. A. (1993). Quantitation of human immunodeficiency virus type 1 infection kinetics. *Journal of Virology*, *67*(4), 2182–2190. https://doi.org/10.1128/jvi.67.4.2182-2190.1993

Dixit, N. M., & Perelson, A. S. (2005). HIV dynamics with multiple infections of target cells. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(23), 8198–

8203. https://doi.org/10.1073/pnas.0407498102

Doitsh, G., Cavrois, M., Lassen, K. G., Zepeda, O., Yang, Z., Santiago, M. L., Hebbeler, A. M., & Greene, W. C. (2010). Abortive HIV infection mediates CD4 T cell depletion and inflammation in human lymphoid tissue. *Cell*, *143*(5), 789–801. https://doi.org/10.1016/j.cell.2010.11.001

Donnadieu, E., Bismuth, G., & Trautmann, A. (1994). Antigen recognition by helper T cells elicits a sequence of distinct changes of their shape and intracellular calcium. *Current Biology*, *4*(7), 584–595. https://doi.org/10.1016/S0960-9822(00)00130-5

Dustin, M. L. (2009). Multiscale analysis of T cell activation: Correlating in vitro and in vivo analysis of the immunological synapse. *Current Topics in Microbiology and Immunology*, *334*(1), 47–70. https://doi.org/10.1007/978-3-540-93864-4_3

Dustin, M. L., Bromley, S. K., Kan, Z., Peterson, D. A., & Unanue, E. R. (1997). Antigen receptor engagement delivers a stop signal to migrating T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(8), 3909–3913. https://doi.org/10.1073/pnas.94.8.3909

Einkauf, K. B., Lee, G. Q., Gao, C., Sharaf, R., Sun, X., Hua, S., Chen, S. M. Y., Jiang, C., Lian, X., Chowdhury, F. Z., Rosenberg, E. S., Chun, T. W., Li, J. Z., Yu, X. G., & Lichterfeld, M. (2019). Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy. *Journal of Clinical Investigation*, *129*(3), 988–998. https://doi.org/10.1172/JCI124291

Engelman, A., Mizuuchi, K., & Craigie, R. (1991). HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. *Cell*, *67*(6), 1211–1221. https://doi.org/10.1016/0092-8674(91)90297-C

Engels, R., Falk, L., Albanese, M., Keppler, O. T., & Sewald, X. (2022). LFA1 and ICAM1 are critical for fusion and spread of murine leukemia virus in vivo. *Cell Reports*, *38*(3), 110279. https://doi.org/10.1016/j.celrep.2021.110279

Eriksson, S., Graf, E. H., Dahl, V., Strain, M. C., Yukl, S. A., Lysenko, E. S., Bosch, R. J., Lai, J., Chioma, S., Emad, F., Abdel-Mohsen, M., Hoh, R., Hecht, F., Hunt, P., Somsouk, M., Wong, J., Johnston, R., Siliciano, R. F., Richman, D. D., … Siliciano, J. D. (2013). Comparative Analysis of Measures of Viral Reservoirs in HIV-1 Eradication Studies. *PLoS Pathogens*, *9*(2), e1003174. https://doi.org/10.1371/journal.ppat.1003174

Fais, S., Capobianchi, M. R., Abbate, I., Castilletti, C., Gentile, M., Fei, P. C., Ameglio, F., & Dianzani, F. (1995). Unidirectional budding of HIV-1 at the site of cell-to-cell contact is associated with co-polarization of intercellular adhesion molecules and HIV-1 viral matrix protein. *AIDS*, *9*(4), 329–335. https://doi.org/10.1097/00002030-199509040-00003

Falcinelli, S. D., Ceriani, C., Margolis, D. M., & Archin, N. M. (2019). New Frontiers in Measuring and Characterizing the HIV Reservoir. *Frontiers in Microbiology*, *10*, 2878. https://doi.org/10.3389/fmicb.2019.02878

Fan, H. (1999). MURINE LEUKEMIA VIRUSES (RETROVIRIDAE). In A. Granoff & R. G. Webster (Eds.), *Encyclopedia of Virology* (Second Edi, pp. 995–1001). Elsevier. https://doi.org/10.1006/rwvi.1999.0186

Farnet, C. M., & Bushman, F. D. (1997). HIV-1 cDNA integration: Requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell*, *88*(4), 483–492. https://doi.org/10.1016/S0092-8674(00)81888-7

Farnet, C. M., & Haseltine, W. A. (1990). Integration of human immunodeficiency virus type 1 DNA in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(11), 4164–4168. https://doi.org/10.1073/pnas.87.11.4164

Fujisawa, R., & Masuda, M. (2007). Ecotropic murine leukemia virus envelope protein affects interaction of cationic amino acid transporter 1 with clathrin adaptor protein complexes, leading to receptor downregulation. *Virology*, *368*(2), 342–350. https://doi.org/10.1016/j.virol.2007.06.036

Fujiwara, T., & Mizuuchi, K. (1988). Retroviral DNA integration: Structure of an integration intermediate. *Cell*, *54*(4), 497–504. https://doi.org/10.1016/0092-8674(88)90071-2

Gaebler, C., Lorenzi, J. C. C., Oliveira, T. Y., Nogueira, L., Ramos, V., Lu, C. L., Pai, J. A., Mendoza, P., Jankovic, M., Caskey, M., & Nussenzweig, M. C. (2019). Combination of quadruplex qPCR and next-generation sequencing for qualitative and quantitative analysis of the HIV-1 latent reservoir. *Journal of Experimental Medicine*, *216*(10), 2253–2264. https://doi.org/10.1084/jem.20190896

Galloway, N. L. K., Doitsh, G., Monroe, K. M., Yang, Z., Muñoz-Arias, I., Levy, D. N., & Greene, W. C. (2015). Cell-to-Cell Transmission of HIV-1 Is Required to Trigger Pyroptotic Death of Lymphoid-Tissue-Derived CD4 T Cells. *Cell Reports*, *12*(10), 1555–1563. https://doi.org/10.1016/j.celrep.2015.08.011

Ganesh, L., Leung, K., Loré, K., Levin, R., Panet, A., Schwartz, O., Koup, R. A., & Nabel, G. J. (2004). Infection of Specific Dendritic Cells by CCR5-Tropic Human Immunodeficiency Virus Type 1 Promotes Cell-Mediated Transmission of Virus Resistant to Broadly Neutralizing Antibodies.

Journal of Virology, *78*(21), 11980–11987. https://doi.org/10.1128/jvi.78.21.11980-11987.2004

Garçon, F., & Okkenhaug, K. (2016). PI3Kδ promotes CD4 + T-cell interactions with antigen-presenting cells by increasing LFA-1 binding to ICAM-1. *Immunology & Cell Biology*, *94*(5), 486–495. https://doi.org/10.1038/icb.2016.1

Geijtenbeek, T. B. ., Kwon, D. S., Torensma, R., Van Vliet, S. J., Van Duijnhoven, G. C. F., Middel, J., Cornelissen, I. L. M. H. ., Nottet, H. S. L. ., KewalRamani, V. N., Littman, D. R., Figdor, C. G., & Van Kooyk, Y. (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, *100*(5), 587–597. https://doi.org/10.1016/S0092-8674(00)80694-7

Gigante, S., Gouil, Q., Lucattini, A., Keniry, A., Beck, T., Tinning, M., Gordon, L., Woodruff, C., Speed, T. P., Blewitt, M. E., & Ritchie, M. E. (2019). Using long-read sequencing to detect imprinted DNA methylation. *Nucleic Acids Research*, *47*(8), e46. https://doi.org/10.1093/nar/gkz107

Gilpatrick, T., Lee, I., Graham, J. E., Raimondeau, E., Bowen, R., Heron, A., Downs, B., Sukumar, S., Sedlazeck, F. J., & Timp, W. (2020). Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nature Biotechnology*, *38*(4), 433–438. https://doi.org/10.1038/s41587-020-0407-5

Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M. C., & McCombie, W. R. (2015). Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome Research*, *25*(11), 1750–1756. https://doi.org/10.1101/gr.191395.115

Gratton, S., Cheynier, R., Dumaurier, M. J., Oksenhendler, E., & Wain-Hobson, S. (2000). Highly restricted spread of HIV-1 and multiply infected cells within splenic germinal centers. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(26), 14566–14571. https://doi.org/10.1073/pnas.97.26.14566

Groot, F., Kuijpers, T. W., Berkhout, B., & de Jong, E. C. (2006). Dendritic cell-mediated HIV-1 transmission to T cells of LAD-1 patients is impaired due to the defect in LFA-1. *Retrovirology*, *3*(1), 75. https://doi.org/10.1186/1742-4690-3-75

Haase, A. T. (2011). Early events in sexual transmission of HIV and SIV and opportunities for interventions. *Annual Review of Medicine*, *62*(1), 127–139. https://doi.org/10.1146/annurev-med-080709-124959

Hakre, S., Chavez, L., Shirakawa, K., & Verdin, E. (2012). HIV latency: Experimental systems and molecular models. *FEMS Microbiology Reviews*, *36*(3), 706–716. https://doi.org/10.1111/j.1574-6976.2012.00335.x

Hare, S., Gupta, S. S., Valkov, E., Engelman, A., & Cherepanov, P. (2010). Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature*, *464*(7286), 232–236. https://doi.org/10.1038/nature08784

Hare, S., Maertens, G. N., & Cherepanov, P. (2012). 3′-Processing and strand transfer catalysed by retroviral integrase in crystallo. *EMBO Journal*, *31*(13), 3020–3028. https://doi.org/10.1038/emboj.2012.118

Harkey, M. A., Kaul, R., Jacobs, M. A., Kurre, P., Bovee, D., Levy, R., & Blau, C. A. (2007). Multiarm high-throughput integration site detection: Limitations of LAM-PCR technology and optimization for clonal analysis. *Stem Cells and Development*, *16*(3), 381–392. https://doi.org/10.1089/scd.2007.0015

Haugh, K. A., Ladinsky, M. S., Ullah, I., Stone, H. M., Pi, R., Gilardet, A., Grunst, M. W., Kumar, P., Bjorkman, P. J., Mothes, W., & Uchil, P. D. (2021). In vivo imaging of retrovirus infection reveals a role for siglec-1/cd169 in multiple routes of transmission. *ELife*, *10*, e64179. https://doi.org/10.7554/eLife.64179

Hiener, B., Horsburgh, B. A., Eden, J. S., Barton, K., Schlub, T. E., Lee, E., von Stockenstrom, S., Odevall, L., Milush, J. M., Liegler, T., Sinclair, E., Hoh, R., Boritz, E. A., Douek, D., Fromentin, R., Chomont, N., Deeks, S. G., Hecht, F. M., & Palmer, S. (2017). Identification of Genetically Intact HIV-1 Proviruses in Specific CD4+ T Cells from Effectively Treated Participants. *Cell Reports*, *21*(3), 813–822. https://doi.org/10.1016/j.celrep.2017.09.081

Hindmarsh, P., & Leis, J. (1999). Retroviral DNA integration. *Microbiology and Molecular Biology Reviews : MMBR*, *63*(4), 836–843. https://doi.org/10.1128/mmbr.63.4.836-843.1999

Hioe, C. E., Chien, P. C., Lu, C., Springer, T. A., Wang, X.-H., Bandres, J., & Tuen, M. (2001). LFA-1 Expression on Target Cells Promotes Human Immunodeficiency Virus Type 1 Infection and Transmission. *Journal of Virology*, *75*(2), 1077–1082. https://doi.org/10.1128/jvi.75.2.1077-1082.2001

Hu, Q., Frank, I., Williams, V., Santos, J. J., Watts, P., Griffin, G. E., Moore, J. P., Pope, M., & Shattock, R. J. (2004). Blockade of Attachment and Fusion Receptors Inhibits HIV-1 Infection of Human Cervical Tissue. *Journal of Experimental Medicine*, *199*(8), 1065–1075. https://doi.org/10.1084/jem.20022212

Hu, W. S., & Hughes, S. H. (2012). HIV-1 reverse transcription. *Cold Spring Harbor Perspectives in Medicine*, *2*(10), a006882–a006882. https://doi.org/10.1101/cshperspect.a006882

Huppa, J. B., & Davis, M. M. (2003). T-cell-antigen recognition and the immunological synapse. *Nature*

Reviews Immunology, *3*(12), 973–983. https://doi.org/10.1038/nri1245

Igakura, T., Stinchcombe, J. C., Goon, P. K. C., Taylor, G. P., Weber, J. N., Griffiths, G. M., Tanaka, Y., Osame, M., & Bangham, C. R. M. (2003). Spread of HTLV-I between lymphocytes by virus-induced polarization of the cytoskeleton. *Science*, *299*(5613), 1713–1716. https://doi.org/10.1126/science.1080115

Iwase, S. C., Miyazato, P., Katsuya, H., Islam, S., Yang, B. T. J., Ito, J., Matsuo, M., Takeuchi, H., Ishida, T., Matsuda, K., Maeda, K., & Satou, Y. (2019). HIV-1 DNA-capture-seq is a useful tool for the comprehensive characterization of HIV-1 provirus. *Scientific Reports*, *9*(1), 12326. https://doi.org/10.1038/s41598-019-48681-5

Jain, M., Koren, S., Miga, K. H., Quick, J., Rand, A. C., Sasani, T. A., Tyson, J. R., Beggs, A. D., Dilthey, A. T., Fiddes, I. T., Malla, S., Marriott, H., Nieto, T., O'Grady, J., Olsen, H. E., Pedersen, B. S., Rhie, A., Richardson, H., Quinlan, A. R., … Loose, M. (2018). Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology*, *36*(4), 338–345. https://doi.org/10.1038/nbt.4060

Jin, J., Sherer, N. M., Heidecker, G., Derse, D., & Mothes, W. (2009). Assembly of the murine leukemia virus is directed towards sites of cell-cell contact. *PLoS Biology*, *7*(7), e1000163. https://doi.org/10.1371/journal.pbio.1000163

Jolly, C. (2011). Cell-to-cell transmission of retroviruses: Innate immunity and interferon-induced restriction factors. *Virology*, *411*(2), 251–259. https://doi.org/10.1016/j.virol.2010.12.031

Jolly, C., Booth, N. J., & Neil, S. J. D. (2010). Cell-Cell Spread of Human Immunodeficiency Virus Type 1 Overcomes Tetherin/BST-2-Mediated Restriction in T cells. *Journal of Virology*, *84*(23), 12185–12199. https://doi.org/10.1128/jvi.01447-10

Jolly, C., Kashefi, K., Hollinshead, M., & Sattentau, Q. J. (2004). HIV-1 Cell to Cell Transfer across an Env-induced, Actin-dependent Synapse. *Journal of Experimental Medicine*, *199*(2), 283–293. https://doi.org/10.1084/jem.20030648

Jolly, C., Mitar, I., & Sattentau, Q. J. (2007a). Requirement for an Intact T-Cell Actin and Tubulin Cytoskeleton for Efficient Assembly and Spread of Human Immunodeficiency Virus Type 1. *Journal of Virology*, *81*(11), 5547–5560. https://doi.org/10.1128/jvi.01469-06

Jolly, C., Mitar, I., & Sattentau, Q. J. (2007b). Adhesion Molecule Interactions Facilitate Human Immunodeficiency Virus Type 1-Induced Virological Synapse Formation between T Cells. *Journal of Virology*, *81*(24), 13916–13921. https://doi.org/10.1128/jvi.01585-07

Jolly, C., & Sattentau, Q. J. (2004). Retroviral spread by induction of virological synapses. *Traffic*, *5*(9), 643–650. https://doi.org/10.1111/j.1600-0854.2004.00209.x

Jolly, C., & Sattentau, Q. J. (2005). Human Immunodeficiency Virus Type 1 Virological Synapse Formation in T Cells Requires Lipid Raft Integrity. *Journal of Virology*, *79*(18), 12088–12094. https://doi.org/10.1128/jvi.79.18.12088-12094.2005

Jolly, C., & Sattentau, Q. J. (2007). Human Immunodeficiency Virus Type 1 Assembly, Budding, and Cell-Cell Spread in T Cells Take Place in Tetraspanin-Enriched Plasma Membrane Domains. *Journal of Virology*, *81*(15), 7873–7884. https://doi.org/10.1128/jvi.01845-06

Jolly, C., Welsch, S., Michor, S., & Sattentau, Q. J. (2011). The regulated secretory pathway in cd4 + t cells contributes to human immunodeficiency virus type-1 cell-to-cell spread at the virological synapse. *PLoS Pathogens*, *7*(9), e1002226. https://doi.org/10.1371/journal.ppat.1002226

Jordan, A., Bisgrove, D., & Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO Journal*, *22*(8), 1868–1877. https://doi.org/10.1093/emboj/cdg188

Josefsson, L., King, M. S., Makitalo, B., Brännström, J., Shao, W., Maldarelli, F., Kearney, M. F., Hu, W. S., Chen, J., Gaines, H., Mellors, J. W., Albert, J., Coffin, J. M., & Palmer, S. E. (2011). Majority of CD4 + T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(27), 11199–11204. https://doi.org/10.1073/pnas.1107729108

Josefsson, L., Palmer, S., Faria, N. R., Lemey, P., Casazza, J., Ambrozak, D., Kearney, M., Shao, W., Kottilil, S., Sneller, M., Mellors, J., Coffin, J. M., & Maldarelli, F. (2013). Single Cell Analysis of Lymph Node Tissue from HIV-1 Infected Patients Reveals that the Majority of CD4+ T-cells Contain One HIV-1 DNA Molecule. *PLoS Pathogens*, *9*(6), e1003432. https://doi.org/10.1371/journal.ppat.1003432

Jung, A., Maier, R., Vartanian, J. P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S., & Meyerhans, A. (2002). Recombination: Multiply infected spleen cells in HIV patients. *Nature*, *418*(6894), 144. https://doi.org/10.1038/418144a

Kamiyama, H., Kakoki, K., Yoshii, H., Iwao, M., Igawa, T., Sakai, H., Hayashi, H., Matsuyama, T., Yamamoto, N., & Kubo, Y. (2011). Infection of XC cells by MLVs and Ebola virus is endosome-dependent but acidification-independent. *PLoS ONE*, *6*(10), e26180. https://doi.org/10.1371/journal.pone.0026180

Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., & Siksnys, V. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biology*, *16*(1), 253. https://doi.org/10.1186/s13059-015-0818-7

Katen, L. J., Januszeski, M. M., Anderson, W. F., Hasenkrug, K. J., & Evans, L. H. (2001). Infectious Entry by Amphotropic as well as Ecotropic Murine Leukemia Viruses Occurs through an Endocytic Pathway. *Journal of Virology*, *75*(11), 5018–5026. https://doi.org/10.1128/jvi.75.11.5018-5026.2001

Katsuya, H., Islam, S., Tan, B. J. Y., Ito, J., Miyazato, P., Matsuo, M., Inada, Y., Iwase, S. C., Uchiyama, Y., Hata, H., Sato, T., Yagishita, N., Araya, N., Ueno, T., Nosaka, K., Tokunaga, M., Yamagishi, M., Watanabe, T., Uchimaru, K., … Satou, Y. (2019). The Nature of the HTLV-1 Provirus in Naturally Infected Individuals Analyzed by the Viral DNA-Capture-Seq Approach. *Cell Reports*, *29*(3), 724-735.e4. https://doi.org/10.1016/j.celrep.2019.09.016

Kebschull, J. M., & Zador, A. M. (2015). Sources of PCR-induced distortions in high-throughput sequencing data sets. *Nucleic Acids Research*, *43*(21), e143. https://doi.org/10.1093/nar/gkv717

Kilby, J. M., & Eron, J. J. (2003). Novel Therapies Based on Mechanisms of HIV-1 Cell Entry. *New England Journal of Medicine*, *348*(22), 2228–2238. https://doi.org/10.1056/nejmra022812

Kim, S., Rusmevichientong, A., Dong, B., Remenyi, R., Silverman, R. H., & Chow, S. A. (2010). Fidelity of Target Site Duplication and Sequence Preference during Integration of Xenotropic Murine Leukemia Virus-Related Virus. *PLoS ONE*, *5*(4), e10255. https://doi.org/10.1371/journal.pone.0010255

Kiselinova, M., Geretti, A. M., Malatinkova, E., Vervisch, K., Beloukas, A., Messiaen, P., Bonczkowski, P., Trypsteen, W., Callens, S., Verhofstede, C., De Spiegelaere, W., & Vandekerckhove, L. (2015). HIV-1 RNA and HIV-1 DNA persistence during suppressive ART with PI-based or nevirapine-based regimens. *Journal of Antimicrobial Chemotherapy*, *70*(12), 3311–3316. https://doi.org/10.1093/jac/dkv250

Kojabad, A. A., Farzanehpour, M., Galeh, H. E. G., Dorostkar, R., Jafarpour, A., Bolandian, M., & Nodooshan, M. M. (2021). Droplet digital PCR of viral DNA/RNA, current progress, challenges, and future perspectives. *Journal of Medical Virology*, *93*(7), 4182–4197. https://doi.org/10.1002/jmv.26846

Kozak, C. A. (2015). Origins of the endogenous and infectious laboratory mouse gammaretroviruses. *Viruses*, *7*(1), 1–26. https://doi.org/10.3390/v7010001

Kozarewa, I., Armisen, J., Gardner, A. F., Slatko, B. E., & Hendrickson, C. L. (2015). Overview of target enrichment strategies. *Current Protocols in Molecular Biology*, *112*, 7.21.1.-7.21.23. https://doi.org/10.1002/0471142727.mb0721s112

Kukic, P., Alvin Leung, H. T., Bemporad, F., Aprile, F. A., Kumita, J. R., De Simone, A., Camilloni, C., & Vendruscolo, M. (2015). Structure and dynamics of the integrin LFA-1 I-domain in the inactive state underlie its inside-out/outside-in signaling and allosteric mechanisms. *Structure*, *23*(4), 745–753. https://doi.org/10.1016/j.str.2014.12.020

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., Levine, R., McEwan, P., … Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

Lapadat-Tapolsky, M., De Rocquigny, H., Van Gent, D., Roques, B., Plasterk, R., & Darlix, J. L. (1993). Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. *Nucleic Acids Research*, *21*(4), 831–839. https://doi.org/10.1093/nar/21.4.831

Law, K. M., Komarova, N. L., Yewdall, A. W., Lee, R. K., Herrera, O. L., Wodarz, D., & Chen, B. K. (2016). In Vivo HIV-1 Cell-to-Cell Transmission Promotes Multicopy Micro-compartmentalized Infection. *Cell Reports*, *15*(12), 2771–2783. https://doi.org/10.1016/j.celrep.2016.05.059

Lebedeva, T., Dustin, M. L., & Sykulev, Y. (2005). ICAM-1 co-stimulates target cells to facilitate antigen presentation. *Current Opinion in Immunology*, *17*(3), 251–258. https://doi.org/10.1016/j.coi.2005.04.008

Lee, G. Q., Orlova-Fink, N., Einkauf, K., Chowdhury, F. Z., Sun, X., Harrington, S., Kuo, H. H., Hua, S., Chen, H. R., Ouyang, Z., Reddy, K., Dong, K., Ndung'U, T., Walker, B. D., Rosenberg, E. S., Yu, X. G., & Lichterfeld, M. (2017). Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. *Journal of Clinical Investigation*, *127*(7), 2689–2696. https://doi.org/10.1172/JCI93289

Lee, M. S., & Craigie, R. (1994). Protection of retroviral DNA from autointegration: Involvement of a cellular factor. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(21), 9823–9827. https://doi.org/10.1073/pnas.91.21.9823

Lee, Y. M., & Coffin, J. M. (1991). Relationship of avian retrovirus DNA synthesis to integration in vitro. *Molecular and Cellular Biology*, *11*(3), 1419–1430. https://doi.org/10.1128/mcb.11.3.1419-1430.1991

Leite, T. F., Delatorre, E., Côrtes, F. H., Ferreira, A. C. G., Cardoso, S. W., Grinsztejn, B., De Andrade, M. M., Veloso, V. G., Morgado, M. G., & Guimarães, M. L. (2019). Reduction of HIV-1 reservoir size and diversity after 1 year of cART among Brazilian individuals starting treatment during early stages of acute infection. *Frontiers in Microbiology*, *10*(FEB), 145. https://doi.org/10.3389/fmicb.2019.00145

Len, A. C. L., Starling, S., Shivkumar, M., & Jolly, C. (2017). HIV-1 Activates T Cell Signaling Independently of Antigen to Drive Viral Spread. *Cell Reports*, *18*(4), 1062–1074. https://doi.org/10.1016/j.celrep.2016.12.057

Levy, C. N., Hughes, S. M., Roychoudhury, P., Reeves, D. B., Amstuz, C., Zhu, H., Huang, M. L., Wei, Y., Bull, M. E., Cassidy, N. A. J., McClure, J., Frenkel, L. M., Stone, M., Bakkour, S., Wonderlich, E. R., Busch, M. P., Deeks, S. G., Schiffer, J. T., Coombs, R. W., … Hladik, F. (2021). A highly multiplexed droplet digital PCR assay to measure the intact HIV-1 proviral reservoir. *Cell Reports Medicine*, *2*(4), 100243. https://doi.org/10.1016/j.xcrm.2021.100243

Levy, D. N., Aldrovandi, G. M., Kutsch, O., & Shaw, G. M. (2004). Dynamics of HIV-1 recombination in its natural target cells. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(12), 4204–4209. https://doi.org/10.1073/pnas.0306764101

Lewis, P. F., & Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *Journal of Virology*, *68*(1), 510–516. https://doi.org/10.1128/jvi.68.1.510-516.1994

Lewis, P., Hensel, M., & Emerman, M. (1992). Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO Journal*, *11*(8), 3053–3058. https://doi.org/10.1002/j.1460-2075.1992.tb05376.x

Li, F., Sewald, X., Jin, J., Sherer, N. M., & Mothes, W. (2014). Murine Leukemia Virus Gag Localizes to the Uropod of Migrating Primary Lymphocytes. *Journal of Virology*, *88*(18), 10541–10555. https://doi.org/10.1128/jvi.01104-14

Li, M., Mizuuchi, M., Burke, T. R., & Craigie, R. (2006). Retroviral DNA integration: Reaction pathway and critical intermediates. *EMBO Journal*, *25*(6), 1295–1304. https://doi.org/10.1038/sj.emboj.7601005

Liszewski, M. K., Yu, J. J., & O'Doherty, U. (2009). Detecting HIV-1 integration by repetitive-sampling Alu-gag PCR. *Methods*, *47*(4), 254–260. https://doi.org/10.1016/j.ymeth.2009.01.002

López-Girona, E., Davy, M. W., Albert, N. W., Hilario, E., Smart, M. E. M., Kirk, C., Thomson, S. J., & Chagné, D. (2020). CRISPR-Cas9 enrichment and long read sequencing for fine mapping in plants. *Plant Methods*, *16*(1), 121. https://doi.org/10.1186/s13007-020-00661-x

Lusic, M., & Siliciano, R. F. (2017). Nuclear landscape of HIV-1 infection and integration. *Nature Reviews Microbiology*, *15*(2), 69–82. https://doi.org/10.1038/nrmicro.2016.162

Maertens, G. N., Hare, S., & Cherepanov, P. (2010). The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature*, *468*(7321), 326–329. https://doi.org/10.1038/nature09517

Malnati, M. S., Scarlatti, G., Gatto, F., Salvatori, F., Cassina, G., Rutigliano, T., Volpi, R., & Lusso, P. (2008). A universal real-time PCR assay for the quantification of group-M HIV-1 proviral load. *Nature Protocols*, *3*(7), 1240–1248. https://doi.org/10.1038/nprot.2008.108

Marsh, M., & Helenius, A. (2006). Virus entry: Open sesame. *Cell*, *124*(4), 729–740. https://doi.org/10.1016/j.cell.2006.02.007

Martin, N., & Sattentau, Q. (2009). Cell-to-cell HIV-1 spread and its implications for immune evasion. *Current Opinion in HIV and AIDS*, *4*(2), 143–149. https://doi.org/10.1097/COH.0b013e328322f94a

Martin, N., Welsch, S., Jolly, C., Briggs, J. A. G., Vaux, D., & Sattentau, Q. J. (2010). Virological Synapse-Mediated Spread of Human Immunodeficiency Virus Type 1 between T Cells Is Sensitive to Entry Inhibition. *Journal of Virology*, *84*(7), 3516–3527. https://doi.org/10.1128/jvi.02651-09

Martinez-Picado, J., Zurakowski, R., Buzón, M. J., & Stevenson, M. (2018). Episomal HIV-1 DNA and its relationship to other markers of HIV-1 persistence. *Retrovirology*, *15*(1), 15. https://doi.org/10.1186/s12977-018-0398-1

Massanella, M., Puigdoménech, I., Cabrera, C., Fernandez-Figueras, M. T., Aucher, A., Gaibelet, G., Hudrisier, D., García, E., Bofill, M., Clotet, B., & Blanco, J. (2009). Antigp41 antibodies fail to block early events of virological synapses but inhibit HIV spread between T cells. *AIDS*, *23*(2), 183–188. https://doi.org/10.1097/QAD.0b013e32831ef1a3

Mazurov, D., Ilinskaya, A., Heidecker, G., Lloyd, P., & Derse, D. (2010). Quantitative comparison of HTLV-1 and HIV-1 cell-to-cell infection with new replication dependent vectors. *PLoS Pathogens*, *6*(2), e1000788. https://doi.org/10.1371/journal.ppat.1000788

McClure, M. O., Marsh, M., & Weiss, R. A. (1988). Human immunodeficiency virus infection of CD4-bearing cells occurs by a pH-independent mechanism. *The EMBO Journal*, *7*(2), 513–518. https://doi.org/10.1002/j.1460-2075.1988.tb02839.x

McClure, M. O., Sommerfelt, M. A., Marsh, M., & Weiss, R. A. (1990). The pH independence of

mammalian retrovirus infection. *Journal of General Virology*, *71*(4), 767–773. https://doi.org/10.1099/0022-1317-71-4-767

Miller, D. G., & Miller, A. D. (1994). A family of retroviruses that utilize related phosphate transporters for cell entry. *Journal of Virology*, *68*(12), 8270–8276. https://doi.org/10.1128/jvi.68.12.8270-8276.1994

Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., & Melikyan, G. B. (2009). HIV Enters Cells via Endocytosis and Dynamin-Dependent Fusion with Endosomes. *Cell*, *137*(3), 433–444. https://doi.org/10.1016/j.cell.2009.02.046

Miyazato, P., Katsuya, H., Fukuda, A., Uchiyama, Y., Matsuo, M., Tokunaga, M., Hino, S., Nakao, M., & Satou, Y. (2016). Application of targeted enrichment to next-generation sequencing of retroviruses integrated into the host human genome. *Scientific Reports*, *6*(1), 28324. https://doi.org/10.1038/srep28324

Monks, C. R. F., Freiberg, B. A., Kupfer, H., Sciaky, N., & Kupfer, A. (1998). Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature*, *395*(6697), 82–86. https://doi.org/10.1038/25764

Morita, E., & Sundquist, W. I. (2004). Retrovirus budding. *Annual Review of Cell and Developmental Biology*, *20*(1), 395–425. https://doi.org/10.1146/annurev.cellbio.20.010403.102350

Mothes, W., Sherer, N. M., Jin, J., & Zhong, P. (2010). Virus Cell-to-Cell Transmission. *Journal of Virology*, *84*(17), 8360–8368. https://doi.org/10.1128/jvi.00443-10

Murooka, T. T., Deruaz, M., Marangoni, F., Vrbanac, V. D., Seung, E., Von Andrian, U. H., Tager, A. M., Luster, A. D., & Mempel, T. R. (2012). HIV-infected T cells are migratory vehicles for viral dissemination. *Nature*, *490*(7419), 283–289. https://doi.org/10.1038/nature11398

Negulescu, P. A., Krasieva, T. B., Khan, A., Kerschbaum, H. H., & Cahalan, M. D. (1996). Polarity of T cell shape, motility, and sensitivity to antigen. *Immunity*, *4*(5), 421–430. https://doi.org/10.1016/S1074-7613(00)80409-4

Nejmeddine, M., Negi, V. S., Mukherjee, S., Tanaka, Y., Orth, K., Taylor, G. P., & Bangham, C. R. M. (2009). HTLV-1-Tax and ICAM-1 act on T-cell signal pathways to polarize the microtubule-organizing center at the virological synapse. *Blood*, *114*(5), 1016–1025. https://doi.org/10.1182/blood-2008-03-136770

Nguyen, D. G., & Hildreth, J. E. K. (2003). Involvement of macrophage mannose receptor in the binding and transmission of HIV by macrophage. *European Journal of Immunology*, *33*(2), 483–493. https://doi.org/10.1002/immu.200310024

Ni, P., Huang, N., Zhang, Z., Wang, D. P., Liang, F., Miao, Y., Xiao, C. Le, Luo, F., & Wang, J. (2019). DeepSignal: Detecting DNA methylation state from Nanopore sequencing reads using deep-learning. *Bioinformatics*, *35*(22), 4586–4595. https://doi.org/10.1093/bioinformatics/btz276

Núñez, D., Comas, L., Lanuza, P. M., Sánchez-Martinez, D., Pérez-Hernández, M., Catalán, E., Domingo, M. P., Velázquez-Campoy, A., Pardo, J., & Gálvez, E. M. (2017). A functional analysis on the interspecies interaction between mouse LFA-1 and human intercellular adhesion molecule-1 at the cell level. *Frontiers in Immunology*, *8*(DEC), 1817. https://doi.org/10.3389/fimmu.2017.01817

Nussbaum, O., Roop, A., & Anderson, W. F. (1993). Sequences determining the pH dependence of viral entry are distinct from the host range-determining region of the murine ecotropic and amphotropic retrovirus envelope proteins. *Journal of Virology*, *67*(12), 7402–7405. https://doi.org/10.1128/jvi.67.12.7402-7405.1993

O'Doherty, U., Swiggard, W. J., Jeyakumar, D., McGain, D., & Malim, M. H. (2002). A Sensitive, Quantitative Assay for Human Immunodeficiency Virus Type 1 Integration. *Journal of Virology*, *76*(21), 10942–10950. https://doi.org/10.1128/jvi.76.21.10942-10950.2002

Oriol-Tordera, B., Berdasco, M., Llano, A., Mothe, B., Gálvez, C., Martinez-Picado, J., Carrillo, J., Blanco, J., Duran-Castells, C., Ganoza, C., Sanchez, J., Clotet, B., Calle, M. L., Sánchez-Pla, A., Esteller, M., Brander, C., & Ruiz-Riol, M. (2020). Methylation regulation of Antiviral host factors, Interferon Stimulated Genes (ISGs) and T-cell responses associated with natural HIV control. *PLoS Pathogens*, *16*(8), e1008678. https://doi.org/10.1371/JOURNAL.PPAT.1008678

Palmer, S., Kearney, M., Maldarelli, F., Halvas, E. K., Bixby, C. J., Bazmi, H., Rock, D., Falloon, J., Davey, R. T., Dewar, R. L., Metcalf, J. A., Hammer, S., Mellors, J. W., & Coffin, J. M. (2005). Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *Journal of Clinical Microbiology*, *43*(1), 406–413. https://doi.org/10.1128/JCM.43.1.406-413.2005

Paruzynski, A., Arens, A., Gabriel, R., Bartholomae, C. C., Scholz, S., Wang, W., Wolf, S., Glimm, H., Schmidt, M., & von Kalle, C. (2010). Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nature Protocols*, *5*(8), 1379–1395. https://doi.org/10.1038/nprot.2010.87

Patro, S. C., Brandt, L. D., Bale, M. J., Halvas, E. K., Joseph, K. W., Shao, W., Wu, X., Guo, S., Murrell,

B., Wiegand, A., Spindler, J., Raley, C., Hautman, C., Sobolewski, M., Fennessey, C. M., Hu, W. S., Luke, B., Hasson, J. M., Niyongabo, A., … Kearney, M. F. (2019). Combined HIV-1 sequence and integration site analysis informs viral dynamics and allows reconstruction of replicating viral ancestors. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(51), 25891–25899. https://doi.org/10.1073/pnas.1910334116

Pauza, C. D. (1990). Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. *Virology*, *179*(2), 886–889. https://doi.org/10.1016/0042-6822(90)90161-J

Pepersack, L., Lee, J. C., McEwan, R., & Ihle, J. N. (1980). Phenotypic heterogeneity of Moloney leukemia virus-induced T cel lymphomas. *Journal of Immunology (Baltimore, Md. : 1950)*, *124*(1), 279–285.

Pfeiffer, F., Gröber, C., Blank, M., Händler, K., Beyer, M., Schultze, J. L., & Mayer, G. (2018). Systematic evaluation of error rates and causes in short samples in next-generation sequencing. *Scientific Reports*, *8*(1), 10950. https://doi.org/10.1038/s41598-018-29325-6

Phillips, D. M. (1994). The role of cell-to-cell transmission in HIV infection. *AIDS*, *8*(6), 719–731. https://doi.org/10.1097/00002030-199406000-00001

Pi, R., Iwasaki, A., Sewald, X., Mothes, W., & Uchil, P. D. (2019). Murine Leukemia Virus Exploits Innate Sensing by Toll-Like Receptor 7 in B-1 Cells To Establish Infection and Locally Spread in Mice. *Journal of Virology*, *93*(21), e00930-19. https://doi.org/10.1128/jvi.00930-19

Pierson, T. C., & Doms, R. W. (2003). HIV-1 entry and its inhibition. In J. A. T. Young (Ed.), *Cellular Factors Involved in Early Steps of Retroviral Replication. Current Topics in Microbiology and Immunology* (Vol. 281, pp. 1–27). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-19012-4_1

Piguet, V., & Sattentau, Q. (2004). Dangerous liaisons at the virological synapse. *Journal of Clinical Investigation*, *114*(5), 605–610. https://doi.org/10.1172/JCI22812

Pillay, Y., & Johnson, L. (2021). World AIDS day 2020: Reflections on global and South African progress and continuing challenges. *Southern African Journal of HIV Medicine*, *22*(1), 1205. https://doi.org/10.4102/SAJHIVMED.V22I1.1205

Puigdomènech, I., Massanella, M., Cabrera, C., Clotet, B., & Blanco, J. (2009). On the steps of cell-to-cell HIV transmission between CD4 T cells. *Retrovirology*, *6*(1), 89. https://doi.org/10.1186/1742-4690-6-89

Puigdomènech, I., Massanella, M., Izquierdo-Useros, N., Ruiz-Hernandez, R., Curriu, M., Bofill, M., Martinez-Picado, J., Juan, M., Clotet, B., & Blanco, J. (2008). HIV transfer between CD4 T cells does not require LFA-1 binding to ICAM-1 and is governed by the interaction of HIV envelope glycoprotein with CD4. *Retrovirology*, *5*(1), 32. https://doi.org/10.1186/1742-4690-5-32

Quinn, T. C. (2008). HIV epidemiology and the effects of antiviral therapy on long-term consequences. *AIDS*, *22*(SUPPL. 3), S7–S12. https://doi.org/10.1097/01.aids.0000327510.68503.e8

Rambaut, A., Posada, D., Crandall, K. A., & Holmes, E. C. (2004). The causes and consequences of HIV evolution. *Nature Reviews Genetics*, *5*(1), 52–61. https://doi.org/10.1038/nrg1246

Rein, A. (2011). Murine leukemia viruses: Objects and organisms. *Advances in Virology*, *2011*, 403419. https://doi.org/10.1155/2011/403419

Richardson, M. W., Carroll, R. G., Stremlau, M., Korokhov, N., Humeau, L. M., Silvestri, G., Sodroski, J., & Riley, J. L. (2008). Mode of Transmission Affects the Sensitivity of Human Immunodeficiency Virus Type 1 to Restriction by Rhesus TRIM5α. *Journal of Virology*, *82*(22), 11117–11128. https://doi.org/10.1128/jvi.01046-08

Ritchie, A. J., Cai, F., Smith, N. M., Chen, S., Song, H., Brackenridge, S., Abdool Karim, S. S., Korber, B. T., McMichael, A. J., Gao, F., & Goonetilleke, N. (2014). Recombination-mediated escape from primary CD8+ T cells in acute HIV-1 infection. *Retrovirology*, *11*(1), 1–10. https://doi.org/10.1186/s12977-014-0069-9

Roe, T., Reynolds, T. C., Yu, G., & Brown, P. O. (1993). Integration of murine leukemia virus DNA depends on mitosis. *EMBO Journal*, *12*(5), 2099–2108. https://doi.org/10.1002/j.1460-2075.1993.tb05858.x

Roebuck, K. A., & Finnegan, A. (1999). Regulation of intercellular adhesion molecule-1 (CD54) gene expression. *Journal of Leukocyte Biology*, *66*(6), 876–888. https://doi.org/10.1002/jlb.66.6.876

Rohdewohld, H., Weiher, H., Reik, W., Jaenisch, R., & Breindl, M. (1987). Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase I-hypersensitive sites. *Journal of Virology*, *61*(2), 336–343. https://doi.org/10.1128/jvi.61.2.336-343.1987

Rosewick, N., Hahaut, V., Durkin, K., Artesi, M., Karpe, S., Wayet, J., Griebel, P., Arsic, N., Marçais, A., Hermine, O., Burny, A., Georges, M., & Van den Broeke, A. (2020). An Improved Sequencing-Based Bioinformatics Pipeline to Track the Distribution and Clonal Architecture of Proviral Integration Sites. *Frontiers in Microbiology*, *11*, 587306. https://doi.org/10.3389/fmicb.2020.587306

Roth, M. J., Schwartzberg, P. L., & Goff, S. P. (1989). Structure of the termini of DNA intermediates in the integration of retroviral DNA: Dependence on IN function and terminal DNA sequence. *Cell*, *58*(1), 47–54. https://doi.org/10.1016/0092-8674(89)90401-7

Ru, M., Shustik, C., & Rassart, E. (1993). Graffi murine leukemia virus: molecular cloning and characterization of the myeloid leukemia-inducing agent. *Journal of Virology*, *67*(8), 4722–4731. https://doi.org/10.1128/jvi.67.8.4722-4731.1993

Russell, R. A., Martin, N., Mitar, I., Jones, E., & Sattentau, Q. J. (2013). Multiple proviral integration events after virological synapse-mediated HIV-1 spread. *Virology*, *443*(1), 143–149. https://doi.org/10.1016/j.virol.2013.05.005

Rutsaert, S., Bosman, K., Trypsteen, W., Nijhuis, M., & Vandekerckhove, L. (2018). Digital PCR as a tool to measure HIV persistence. *Retrovirology*, *15*(1), 16. https://doi.org/10.1186/s12977-018-0399-0

Ryu, W.-S. (2017). Retroviruses. In W.-S. Ryu (Ed.), *Molecular Virology of Human Pathogenic Viruses* (pp. 227–246). Academic Press. Elsevier. https://doi.org/10.1016/b978-0-12-800838-6.00017-5

Sato, H., Orensteint, J., Dimitrov, D., & Martin, M. (1992). Cell-to-cell spread of HIV-1 occurs within minutes and may not involve the participation of virus particles. *Virology*, *186*(2), 712–724. https://doi.org/10.1016/0042-6822(92)90038-Q

Sattentau, Q. (2008). Avoiding the void: Cell-to-cell spread of human viruses. *Nature Reviews Microbiology*, *6*(11), 815–826. https://doi.org/10.1038/nrmicro1972

Scherdin, U., Rhodes, K., & Breindl, M. (1990). Transcriptionally active genome regions are preferred targets for retrovirus integration. *Journal of Virology*, *64*(2), 907–912. https://doi.org/10.1128/jvi.64.2.907-912.1990

Schmidt, M., Hoffmann, G., Wissler, M., Lemke, N., Müßig, A., Glimm, H., Williams, D. A., Ragg, S., Hesemann, C. U., & von Kalle, C. (2001). Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. *Human Gene Therapy*, *12*(7), 743–749. https://doi.org/10.1089/104303401750148649

Schmidt, M., Schwarzwaelder, K., Bartholomae, C., Zaoui, K., Ball, C., Pilz, I., Braun, S., Glimm, H., & von Kalle, C. (2007). High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nature Methods*, *4*(12), 1051–1057. https://doi.org/10.1038/nmeth1103

Scholer, A., Hugues, S., Boissonnas, A., Fetler, L., & Amigorena, S. (2008). Intercellular Adhesion Molecule-1-Dependent Stable Interactions between T Cells and Dendritic Cells Determine CD8+ T Cell Memory. *Immunity*, *28*(2), 258–270. https://doi.org/10.1016/j.immuni.2007.12.016

Serrao, E., Cherepanov, P., & Engelman, A. N. (2016). Amplification, next-generation sequencing, and genomic DNA mapping of retroviral integration sites. *Journal of Visualized Experiments*, *2016*(109), 53840. https://doi.org/10.3791/53840

Serrao, E., & Engelman, A. N. (2016). Sites of retroviral DNA integration: From basic research to clinical applications. *Critical Reviews in Biochemistry and Molecular Biology*, *51*(1), 26–42. https://doi.org/10.3109/10409238.2015.1102859

Sewald, X., Gonzalez, D. G., Haberman, A. M., & Mothes, W. (2012). In vivo imaging of virological synapses. *Nature Communications*, *3*(1), 1320–1329. https://doi.org/10.1038/ncomms2338

Sewald, X., Ladinsky, M. S., Uchil, P. D., Beloor, J., Pi, R., Herrmann, C., Motamedi, N., Murooka, T. T., Brehm, M. A., Greiner, D. L., Shultz, L. D., Mempel, T. R., Bjorkman, P. J., Kumar, P., & Mothes, W. (2015). Retroviruses use CD169-mediated trans-infection of permissive lymphocytes to establish infection. *Science*, *350*(6260), 563–567. https://doi.org/10.1126/science.aab2749

Sewald, X., Motamedi, N., & Mothes, W. (2016). Viruses exploit the tissue physiology of the host to spread in vivo. *Current Opinion in Cell Biology*, *41*, 81–90. https://doi.org/10.1016/j.ceb.2016.04.008

Shalginskikh, N., Poleshko, A., Skalka, A. M., & Katz, R. A. (2013). Retroviral DNA Methylation and Epigenetic Repression Are Mediated by the Antiviral Host Protein Daxx. *Journal of Virology*, *87*(4), 2137–2150. https://doi.org/10.1128/jvi.02026-12

Sherer, N. M., Jin, J., & Mothes, W. (2010). Directional Spread of Surface-Associated Retroviruses Regulated by Differential Virus-Cell Interactions. *Journal of Virology*, *84*(7), 3248–3258. https://doi.org/10.1128/jvi.02155-09

Shimaoka, M., Lu, C., Palframan, R. T., Von Andrian, U. H., McCormack, A., Takagi, J., & Springer, T. A. (2001). Reversibly locking a protein fold in an active conformation with a disulfide bond: Integrin αL I domains with high affinity and antagonist activity in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(11), 6009–6014. https://doi.org/10.1073/pnas.101130498

Sigal, A., Kim, J. T., Balazs, A. B., Dekel, E., Mayo, A., Milo, R., & Baltimore, D. (2011). Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. *Nature*, *477*(7362), 95–98. https://doi.org/10.1038/nature10347

Simpson, J. T., Workman, R. E., Zuzarte, P. C., David, M., Dursi, L. J., & Timp, W. (2017). Detecting

DNA cytosine methylation using nanopore sequencing. *Nature Methods*, *14*(4), 407–410. https://doi.org/10.1038/nmeth.4184

Sims, T. N., Soos, T. J., Xenias, H. S., Dubin-Thaler, B., Hofman, J. M., Waite, J. C., Cameron, T. O., Thomas, V. K., Varma, R., Wiggins, C. H., Sheetz, M. P., Littman, D. R., & Dustin, M. L. (2007). Opposing Effects of PKCθ and WASp on Symmetry Breaking and Relocation of the Immunological Synapse. *Cell*, *129*(4), 773–785. https://doi.org/10.1016/j.cell.2007.03.037

Sloan, R. D., Kuhl, B. D., Mesplède, T., Münch, J., Donahue, D. A., & Wainberg, M. A. (2013). Productive Entry of HIV-1 during Cell-to-Cell Transmission via Dynamin-Dependent Endocytosis. *Journal of Virology*, *87*(14), 8110–8123. https://doi.org/10.1128/jvi.00815-13

Sol-Foulon, N., Sourisseau, M., Porrot, F., Thoulouze, M. I., Trouillet, C., Nobile, C., Blanchet, F., Di Bartolo, V., Noraz, N., Taylor, N., Alcover, A., Hivroz, C., & Schwartz, O. (2007). ZAP-70 kinase regulates HIV cell-to-cell spread and virological synapse formation. *EMBO Journal*, *26*(2), 516–526. https://doi.org/10.1038/sj.emboj.7601509

Sourisseau, M., Sol-Foulon, N., Porrot, F., Blanchet, F., & Schwartz, O. (2007). Inefficient Human Immunodeficiency Virus Replication in Mobile Lymphocytes. *Journal of Virology*, *81*(2), 1000–1012. https://doi.org/10.1128/jvi.01629-06

Springer, T. A. (1990). Adhesion receptors of the immune system. *Nature*, *346*(6283), 425–434. https://doi.org/10.1038/346425a0

Springer, T. A., & Dustin, M. L. (2012). Integrin inside-out signaling and the immunological synapse. *Current Opinion in Cell Biology*, *24*(1), 107–115. https://doi.org/10.1016/j.ceb.2011.10.004

Stadtmann, A., Brinkhaus, L., Mueller, H., Rossaint, J., Bolomini-Vittori, M., Bergmeier, W., Van Aken, H., Wagner, D. D., Laudanna, C., Ley, K., & Zarbock, A. (2011). Rap1a activation by CalDAG-GEFI and p38 MAPK is involved in E-selectin-dependent slow leukocyte rolling. *European Journal of Immunology*, *41*(7), 2074–2085. https://doi.org/10.1002/eji.201041196

Stangl, C., de Blank, S., Renkens, I., Westera, L., Verbeek, T., Valle-Inclan, J. E., González, R. C., Henssen, A. G., van Roosmalen, M. J., Stam, R. W., Voest, E. E., Kloosterman, W. P., van Haaften, G., & Monroe, G. R. (2020). Partner independent fusion gene detection by multiplexed CRISPR-Cas9 enrichment and long read nanopore sequencing. *Nature Communications*, *11*(1), 2861. https://doi.org/10.1038/s41467-020-16641-7

Starling, S., & Jolly, C. (2016). LFA-1 Engagement Triggers T Cell Polarization at the HIV-1 Virological Synapse. *Journal of Virology*, *90*(21), 9841–9854. https://doi.org/10.1128/jvi.01152-16

Stebbing, J., Gazzard, B., & Douek, D. C. (2004). Where Does HIV Live? *New England Journal of Medicine*, *350*(18), 1872–1880. https://doi.org/10.1056/nejmra032395

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, *507*(7490), 62–67. https://doi.org/10.1038/nature13011

Stolp, B., Imle, A., Coelho, F. M., Hons, M., Gorina, R., Lyck, R., Stein, J. V., & Fackler, O. T. (2012). HIV-1 Nef interferes with T-lymphocyte circulation through confined environments in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(45), 18541–18546. https://doi.org/10.1073/pnas.1204322109

Strain, M. C., Lada, S. M., Luong, T., Rought, S. E., Gianella, S., Terry, V. H., Spina, C. A., Woelk, C. H., & Richman, D. D. (2013). Highly Precise Measurement of HIV DNA by Droplet Digital PCR. *PLoS ONE*, *8*(4), e55943. https://doi.org/10.1371/journal.pone.0055943

Suomalainen, M., Hultenby, K., & Garoff, H. (1996). Targeting of moloney murine leukemia virus gag precursor to the site of virus budding. *Journal of Cell Biology*, *135*(6 II), 1841–1852. https://doi.org/10.1083/jcb.135.6.1841

Symons, J., Chopra, A., Malantinkova, E., Spiegelaere, W., Leary, S., Cooper, D., Abana, C. O., Rhodes, A., Rezaei, S. D., Vandekerckhove, L., Mallal, S., Lewin, S. R., & Cameron, P. U. (2017). HIV integration sites in latently infected cell lines: Evidence of ongoing replication. *Retrovirology*, *14*(1), 2. https://doi.org/10.1186/s12977-016-0325-2

Troxler, D. H., & Scolnick, E. M. (1978). Rapid leukemia induced by cloned Friend strain of replicating murine type-C virus Association with induction of xenotropic-related RNA sequences contained in spleen focus-forming virus. *Virology*, *85*(1), 17–27. https://doi.org/10.1016/0042-6822(78)90408-7

Turville, S. G., Cameron, P. U., Handley, A., Lin, G., Pöhlmann, S., Doms, R. W., & Cunningham, A. L. (2002). Diversity of receptors binding HIV on dendritic cell subsets. *Nature Immunology*, *3*(10), 975–983. https://doi.org/10.1038/ni841

Uchil, P. D., Pi, R., Haugh, K. A., Ladinsky, M. S., Ventura, J. D., Barrett, B. S., Santiago, M. L., Bjorkman, P. J., Kassiotis, G., Sewald, X., & Mothes, W. (2019). A Protective Role for the Lectin CD169/Siglec-1 against a Pathogenic Murine Retrovirus. *Cell Host and Microbe*, *25*(1), 87-100.e10. https://doi.org/10.1016/j.chom.2018.11.011

Van De Stolpe, A., & Van Der Saag, P. T. (1996). Intercellular adhesion molecule-1. *Journal of Molecular Medicine (Berlin, Germany)*, *74*(1), 13–33. https://doi.org/10.1007/BF00202069

Van Haasteren, J., Munis, A. M., Gill, D. R., & Hyde, S. C. (2021). Genome-wide integration site detection using Cas9 enriched am plification-free long-range sequencing. *Nucleic Acids Research*, *49*(3), e16. https://doi.org/10.1093/nar/gkaa1152

Van Seventer, G. A., Shimizu, Y., Horgan, K. J., & Shaw, S. (1990). The LFA-1 ligand ICAM-1 provides an important costimulatory signal for T cell receptor-mediated activation of resting T cells. *Journal of Immunology (Baltimore, Md. : 1950)*, *144*(12), 4579–4586. http://www.ncbi.nlm.nih.gov/pubmed/1972160

Vandegraaff, N., Kumar, R., Burrell, C. J., & Li, P. (2001). Kinetics of Human Immunodeficiency Virus Type 1 (HIV) DNA Integration in Acutely Infected Cells as Determined Using a Novel Assay for Detection of Integrated HIV DNA. *Journal of Virology*, *75*(22), 11253–11260. https://doi.org/10.1128/jvi.75.22.11253-11260.2001

Varma, R., Campi, G., Yokosuka, T., Saito, T., & Dustin, M. L. (2006). T Cell Receptor-Proximal Signals Are Sustained in Peripheral Microclusters and Terminated in the Central Supramolecular Activation Cluster. *Immunity*, *25*(1), 117–127. https://doi.org/10.1016/j.immuni.2006.04.010

Vasiliver-Shamis, G., Cho, M. W., Hioe, C. E., & Dustin, M. L. (2009). Human Immunodeficiency Virus Type 1 Envelope gp120-Induced Partial T-Cell Receptor Signaling Creates an F-Actin-Depleted Zone in the Virological Synapse. *Journal of Virology*, *83*(21), 11341–11355. https://doi.org/10.1128/jvi.01440-09

Vasiliver-Shamis, G., Dustin, M. L., & Hioe, C. E. (2010). HIV-1 virological synapse is not simply a copycat of the immunological synapse. *Viruses*, *2*(5), 1239–1260. https://doi.org/10.3390/v2051239

Vasiliver-Shamis, G., Tuen, M., Wu, T. W., Starr, T., Cameron, T. O., Thomson, R., Kaur, G., Liu, J., Visciano, M. L., Li, H., Kumar, R., Ansari, R., Han, D. P., Cho, M. W., Dustin, M. L., & Hioe, C. E. (2008). Human Immunodeficiency Virus Type 1 Envelope gp120 Induces a Stop Signal and Virological Synapse Formation in Noninfected CD4 + T Cells. *Journal of Virology*, *82*(19), 9445–9457. https://doi.org/10.1128/jvi.00835-08

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., … Zhu, X. (2001). The sequence of the human genome. *Science*, *291*(5507), 1304–1351. https://doi.org/10.1126/science.1058040

Vincent, K. A., York-Higgins, D., Quiroga, M., & Brown, P. O. (1990). Host sequences flanking the HIV provirus. *Nucleic Acids Research*, *18*(20), 6045–6047. https://doi.org/10.1093/nar/18.20.6045

Vink, C., Groenink, M., Elgersma, Y., Fouchier, R. A., Tersmette, M., & Plasterk, R. H. (1990). Analysis of the junctions between human immunodeficiency virus type 1 proviral DNA and human DNA. *Journal of Virology*, *64*(11), 5626–5627. https://doi.org/10.1128/jvi.64.11.5626-5627.1990

Vondrak, T., Ávila Robledillo, L., Novák, P., Koblížková, A., Neumann, P., & Macas, J. (2020). Characterization of repeat arrays in ultra-long nanopore reads reveals frequent origin of satellite DNA from retrotransposon-derived tandem repeats. *Plant Journal*, *101*(2), 484–500. https://doi.org/10.1111/tpj.14546

Wallace, A. D., Sasani, T. A., Swanier, J., Gates, B. L., Greenland, J., Pedersen, B. S., Varley, K. E., & Quinlan, A. R. (2021). CaBagE: A Cas9-based Background Elimination strategy for targeted, long-read DNA sequencing. *PLOS ONE*, *16*(4), e0241253. https://doi.org/10.1371/journal.pone.0241253

Walling, B. L., & Kim, M. (2018). LFA-1 in T cell migration and differentiation. *Frontiers in Immunology*, *9*(MAY), 952. https://doi.org/10.3389/fimmu.2018.00952

Wang, H., Kavanaugh, M. P., North, R. A., & Kabat, D. (1991). Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter. *Nature*, *352*(6337), 729–731. https://doi.org/10.1038/352729a0

Wang, J.-H., Kwas, C., & Wu, L. (2009). Intercellular Adhesion Molecule 1 (ICAM-1), but Not ICAM-2 and -3, Is Important for Dendritic Cell-Mediated Human Immunodeficiency Virus Type 1 Transmission. *Journal of Virology*, *83*(9), 4195–4204. https://doi.org/10.1128/jvi.00006-09

Wang, W., Bartholomae, C. C., Gabriel, R., Deichmann, A., & Schmidt, M. (2016). The LAM-PCR method to sequence LV integration sites. In M. Federico (Ed.), *Lentiviral Vectors and Exosomes as Gene and Protein Delivery Tools. Methods in Molecular Biology* (Vol. 1448, pp. 107–120). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-3753-0_9

Wang, X. Q., & Palmer, S. (2018). Single-molecule techniques to quantify and genetically characterise persistent HIV. *Retrovirology*, *15*(1), 3. https://doi.org/10.1186/s12977-017-0386-x

Wells, D. W., Guo, S., Shao, W., Bale, M. J., Coffin, J. M., Hughes, S. H., & Wu, X. (2020). An analytical pipeline for identifying and mapping the integration sites of HIV and other retroviruses. *BMC Genomics*, *21*(1), 216. https://doi.org/10.1186/s12864-020-6647-4

Wick, R. R., Judd, L. M., & Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biology*, *20*(1), 129. https://doi.org/10.1186/s13059-019-1727-y

Wilson, C. A., Farrell, K. B., & Eiden, M. V. (1994). Properties of a unique form of the murine

amphotropic leukemia virus receptor expressed on hamster cells. *Journal of Virology*, *68*(12), 7697–7703. https://doi.org/10.1128/jvi.68.12.7697-7703.1994

Yeager, M., Wilson-Kubalek, E. M., Weiner, S. G., Brown, P. O., & Rein, A. (1998). Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: Implications for retroviral assembly mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7299–7304. https://doi.org/10.1073/pnas.95.13.7299

Young, G. R., & Bishop, K. N. (2021). Murine Leukemia and Sarcoma Viruses (Retroviridae). In D. H. Bamford & M. Zuckerman (Eds.), *Encyclopedia of Virology* (Fourth Edi, pp. 643–647). Academic Press, Elsevier. https://doi.org/10.1016/b978-0-12-814515-9.00038-2

Zimmerman, T., & Blanco, F. (2008). Inhibitors Targeting the LFA-1/ICAM-1 Cell-Adhesion Interaction: Design and Mechanism of Action. *Current Pharmaceutical Design*, *14*(22), 2128–2139. https://doi.org/10.2174/138161208785740225

Acknowledgments

First of all, I would like to thank PD Dr. Barbara Adler for her outstanding supervision, guidance, and support during my work on this exciting and interesting project. My special thanks to Dr. Xaver Sewald for his advice during the course of my thesis. Furthermore, I would like to thank Prof. Dr. Oliver Keppler for giving me the opportunity to work on this project in the Department of Virology at the Max von Pettenkofer-Institute.

Further on, I would like to thank Dr. Glen R. Monroe, Dr. Roy Straver and Ivo Renkens from the Department of Genetics at the University Medical Center Utrecht and all members of the AG Blum at the Gene Center Munich, for their kind assistance, support and helpful advice during the establishment of the CRISPR/Cas9 target enrichment Nanopore sequencing technique.

My special thanks to all members of the AG Sewald, AG Schölz, AG Baldauf, AG Adler and the AG Keppler for their kind support and helpful suggestions during the work on this project. In particular, I would like to thank Rebecca Engels, Dr. Manuel Albanese, Dr. Ernesto Mejías-Pérez, Linda Jocham, Laura Amann, Marie Bischof, Thimo Fuchs and Ramya Nair for their assistance, cooperativeness, and the pleasant working atmosphere. Furthermore, my thanks go to my friends, who have accompanied me through my studies and supported me in many ways.

Celia and Walerija, my dear friends, thank you for enriching my life with your uniqueness. Special thanks are due to my parents and entire family for their unlimited support, motivation, patience and for always believing in me.

My very special thanks are directed to Max, Leo, Asti, Fritz, Benny, Saruska, Rico and Blitz for being the most marvellous individuals of the Laniakea Supercluster and all groups of galaxy clusters that are about to be discovered.

Affidavit

Eidesstattliche Versicherung

Falk, Lisa

……..……..

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Exploring the role of cell adhesion proteins in retrovirus transmission and establishing

CRISPR/Cas9 target enrichment Nanopore sequencing for retrovirus integration analysis

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Neumarkt-St. Veit, 11.6.23
Ort, Datum

Lisa Falk
Unterschrift Doktorandin bzw. Doktorand

List of publications

Figures 6, 7, 8 and 19 of this thesis incorporate graphs that were published in:
Engels, R.*, **Falk, L.***, Albanese, M., Keppler, O. T. and Sewald, X. (2022). LFA1 and ICAM1 are critical for fusion and spread of murine leukemia virus *in vivo*. *Cell Reports*, *38*(3), 110279. *Contributed equally

Hanschmann, E. M., Wilms, C., **Falk, L.**, Holubiec, M. I., Mennel, S., Lillig, C. H. and Godoy, J. R. (2022). Cytosolic glutaredoxin 1 is upregulated in AMD and controls retinal pigment epithelial cells proliferation via β-catenin. *Biochemical and Biophysical Research Communications*, *618*, 24-29.