SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Comparative transcriptomic and proteomic analysis of HIV-1 and HIV-2 pseudovirion transduction in the early phase of viral infection

Tamás Richárd Linkner

Supervisor: Dr. Mohamed Mahdi

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR CELLULAR AND IMMUNE BIOLOGY

Debrecen, 2025

Comparative transcriptomic and proteomic analysis of HIV-1 and HIV-2 pseudovirion transduction in the early phase of viral infection

By Tamás Richárd Linkner, MSc, molecular biologist

Supervisor: Dr. Mohamed Mahdi

Doctoral School of Doctoral School of Molecular Cellular and Immune Biology, University of Debrecen

Head of the Defense Committee:	Prof. Dr. Attila Bácsi, PhD, DSc
Reviewers:	Dr. Eszter Csoma, PhD
	Dr. Aladár Pettkó-Szandtner, PhD
Members of the Defense Committee:	Dr. Krisztina Tar, PhD
	Dr. Árpád Patai, PhD

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen at 13:00 p.m., 13th of March, 2026.

Introduction

Epidemiology of HIV

According to data from the World Health Organization (WHO), more than 39 million people worldwide are infected with HIV (HIV.gov 2024). HIV-1 is responsible for the majority of those infections, however, infections caused by HIV-2, and dual HIV-1/2 infections are also thought to be included in the pooled data. To date, no accurate statistics of the prevalence of HIV-2 infection exists, with very outdated estimates of around 2 million infections majorly in the West African region, including cases that are infected with both HIV-1 and 2 (Gottlieb et al., 2008). More recently, prevalence of HIV-2 infection was shown to extend to countries with significant colonial or social and economic connections with the West-African region, such as France, Spain and Portugal (HIV.uw.edu 2024), (Reeves & Doms, 2002), Campbell-Yesufu & Gandhi, 2011).

Genetic structure of HIV

The genome of HIV consists of two identical copies of single stranded RNA which is transcribed into DNA via the reverse transcriptase during infection. There is a long terminal repeat (LTR) in each end of the HIV genome, and their 5' end region code for a transcriptional promoter. The *gag* gene encodes for the proteins of the outer core membrane, capsid protein, nucleocapsid and smaller, nucleic acid stabilizing protein. This is followed by the *pol* gene encoding for the enzymes protease (PR), reverse transcriptase (RT) and integrase (IN). Adjacent to the *pol* gene is the *env* reading frame, which codes for the envelope glycoproteins. HIV genome encodes for regulatory proteins Tat and Rev which are crucial for the initiation of HIV replication. Additionally, multiple auxiliary proteins such as Nef, Vif, Vpr and Vpu are also coded by its genome. HIV-2 encodes for the Vpx instead of Vpu (German Advisory Committee Blood, 2016) (Li et al., 2005). Despite their similar ancestral origin, HIV-1 and HIV-2 contain only 50-55% identity in their nucleotide sequence of their genome. There is roughly 54% amino acid sequence similarity in their Gag, 55% in Pol and 35% in Env proteins, depending on the viral groups (Li et al., 2015) (Motomura et al., 2008).

Life cycle of HIV

The life cycle of HIVs can be divided into two parts: and early, and a late phase. In the beginning of the early phase, the viral surface glycoprotein binds to the target receptor which triggers the fusion of the viral envelope and the cell membrane (Freed, 2015) (Chen, 2019**) (**Melikyan, 2014). Following entry, the viral core travels through the cytoskeleton to the nucleus. The viral

core encompasses the capsid, nucleocapsid and viral proteins such as RT, PR and IN. During cytoskeletal travel, the viral capsid disassembles in a process called uncoating, which is required for the completion of the reverse transcription (Ambrose & Aiken, 2014). The reverse transcription starts inside the viral core during its transport through the cytoskeleton. The process is mediated by the RT enzyme during which the viral RNA genome is transcribed into double stranded DNA (Goff, 2001) **(**Xavier Ruiz & Arnold, 2020). In close proximity to, or within the nucleus, the pre-integration complex (PIC) assembles which contains the transcribed viral DNA as well as viral and host proteins (Goff, 2001) **(**Arhel, 2010**)** (Li, 2015). Moreover, recent data indicates that the uncoating and reverse transcription continues following the nuclear entry (Dharan et al., 2020) (Müller et al., 2022). The PIC enters into the nucleus via the nucleopore complex in an importin/karyophilin dependent manner. Host proteins which are involved in the nuclear entry are recruited through the nuclear localization signal of the Vpr and integrase proteins. Integrase interacts with members of the nuclear transport machinery and induces the translocation of the PIC into the nucleus (Jayappa et al., 2012) **(**Popov et al., 1998). After nuclear entry, the viral genome is integrated into the host DNA which is mediated by the viral IN (Hokello et al., 2024). The integration marks the transition from early to late phase of the HIV life cycle.

In late phase, the viral transcription began during which new viral proteins are synthetized and new virons are assembled (Freed, 2015). In infected cells the integrated viral DNA acts as a template for viral mRNA and genomic RNA synthesis. During the beginning of the late phase the viral transcripion is dependent on cellular factors like nuclear factor kappa B (NF-κB) until the expression of tat. Tat binds to the viral TAR region after which its recruit's cellular proteins like positive transcription elongation factor b (P-TEFb), comprising cyclin-dependent kinase 9 (CDK9) and cyclin T1 (CCNT1). This allows the virus to synthetize full length viral RNA and different mRNA products (Liu et al., 2014). The transport of the viral mRNA from the nucleus to the cytoplasm is managed by the Rev protein (Hokello et al., 2024). HIV-1 virion assembly happens at the cellular plasma membrane and the process is mediated by the Gag polyprotein. The Gag recruits the Env proteins and viral genomic RNA to the site of virus assembly. The virion acquires its lipid envelope and Env protein spikes during the budding from the plasma membrane. The Gag polyprotein first assembles into immature particles which following budding is processed by the activated viral PR forming new infectious mature viral particles (Sundquist & Kräusslich, 2012) (Ganser et al., 2012).

Lentiviral Vectors

Replication incompetent lentiviral vectors serve as convenient and effective tools for stable and efficient transfer of genes into cells of human origin, making them attractive candidates for human gene therapy. Their main appeal is that they stably integrate into the DNA of their target cells ensuring long-term gene expression. Lentivirus based vectors offer a notable advantage over oncoretroviral delivery systems which is that they are able to introduce genes into non-dividing cells. Successful treatment was reported from patients with severe monogenic disorders; such as skin adhesion disorders, immunodeficiency's, neurometabolic and hemoglobinopathies diseases, based on somatic stem cells transduced with viral vectors (Zufferey et al., 1998) (Tucci et al., 2021).

The first lentivirus-based vectors contained large parts of the HIV genome. As for the envelope, G protein of the vesicular stotamitis virus (VSV-G) was utilized. The target receptor for VSV-G is a low density lipoprotein (LDL) receptor which allowed the lentivirus-based vectors to target a large variety of cells. In the next iteration of lentivectors, the accessory proteins were removed, further improving the design and safety. For further upgrading vector safety, in third generation systems, the genome of the virus was divided into two plasmids which makes it less unlikely to generate recombinant viruses. Lentivirus-based vectors became attractive from a clinical perspective due to their abilities of transducing slowly or non-proliferating cells including CD34+ stem cell. Genetic disease such as metachromatic leukodystrophy, Wiskot-Aldrich syndrome, β-thalassemia and X-linked adrenoleukodystrophy were successfully treated with lentivirus-based vectors (Milone & O'Doherty, 2018) (Dull el al., 1998) (Miyoshi et al., 1998) (Zufferey et al., 1998). There are also fourth generation lentiviral vectors, wherein they modified the packaging sequences in order to avoid the formation of replication competent particles. In these vectors, the HIV-1 ψ and RRE packaging sequences are located downstream of the self-inactivating LTR. Some viral components required for virus production, such as structural and regulatory genes, are provided during vector manufacturing, but are excluded from the final viral genome. This overall further improves the safety of the vectors (Vink et al., 2017) (Berkhout, 2017).

HIV-induced transcriptomic and proteomic changes

Researchers discovered multiple genes that are regulated differentially in response to HIV infection. Coelho et al showed that 208 genes were upregulated in infected CD4+ T cells, with

toll like receptor 7 (*TLR7*), interferon beta 1 (*IFNB1*) and TNF superfamily member 4 (*TNFSF4*) amongst them (Coelho et al., 2021). In their study, Wu et al focused on genes that are expressed differentially in CD4+, CD8+ T cells, and macrophages from patients infected with HIV. They found elevated levels of genes that regulate proteasome, proton transporting ATPase complex, actin filaments, and complement activation. Pathways such as mitochondrial signatures of disease progression and pathways linked to energy production, apoptosis, cell-cycle dysregulation and metabolism showed enrichment during the analysis (Wu et al., 2011).

Pollara et al. employed single-cell transcriptome sequencing to elucidate the alterations and cellular dynamics linked to HIV-1 infection, changes which may remain despite therapy. They found that 45 genes were downregulated and 96 were upregulated in cells from un-treated HIV-1 infected patients. Notably, HIV-1 infection induced a proinflammatory state across all observed immune cell types, characterized by the upregulation of interferon response genes. Additionally, compared to seronegative individuals, they found 83 upregulated and 31 downregulated genes in HIV-1 infected individuals under antiretroviral therapy (Pollara et al., 2022).

Many studies focus on the proteomic changes caused by HIV infection. In their work Chan et al detected 3255 cellular proteins from HIV-1 infected CD4+ T cell line. Out of those detected proteins, 344 were upregulated and 343 were downregulated at 36 hours post-infection. This time-point denotes the peak of viral production. The differentially expressed proteins were concentrated to select biological pathways according to their pathway analysis, including cell cycle progression, citrate cycle pathways, ubiquitination and nucleocytoplasmic transport. They also observed alterations in the level of proteins with known interaction with proteins of HIV-1 (Chan et al., 2007).

In another study, Al-Mozaini et al identified over 314 unique peripheral blood plasma proteins, out of which 100 were significantly differentially regulated between HIV-1, HIV-2 and HIV-1 elite controller samples. Between the two HIV strains, the protein profiles were remarkably similar. However, despite these similarities, 6 proteins showed significant differences between HIV-1 and HIV-2 infected samples including prohibitin-2 (PHB2), calcineurin subunit B type 1 (PPP3R1), electron transfer flavoprotein subunit beta (ETFB), rhombotin-2 (LMO2), protein S100-A94s (S100A9) and also viral Vif. In the study they concluded that these different proteins could be utilized as diagnosis and prognostic molecules of HIV infected individuals. As using single makers might not provide accurate prognostic data it is better to use disease-specific protein panel markers which provide more prominent information (Al-Mozaini et al., 2021**)**.

Aims

Despite the almost 40 years of HIV research so far, very few studies aimed to characterize the early phase of the HIV lifecycle in different cell types. Moreover, few articles aimed to characterize HIV-2 generated cellular changes in comparison to HIV-1. Additionally, the use of lentiviral-derived vectors for gene therapy is continuously being adopted, and great efforts to improve their safety and stability are being spent in their research and development, yet their effects on the host cell transcriptome and proteome is understudied.

In order to enrich our knowledge of HIV-1 and HIV-2 generated proteomic and transcriptomic changes, we set the following aims for our study:

1, Analysis of the Proteo-transcriptomic changes in HEK-293T cells induced by HIV-1 and 2-based lentiviral vectors in the early-phase of lentiviral transduction

2, Comparative analysis of the proteo-transcriptomic changes between HIV-1 and HIV-2 highlighting the similarities and differences in the pathways involved in HEK-293T and Jurkat cells

The study was carried out between 2018 and 2022, conducted at the department of Biochemistry and Molecular Biology, utilizing services from the Proteomics Core Facility and the Center for Clinical Genomics and Personalized Medicine at the University of Debrecen.

Materials and methods

Utilized plasmids

We used a 2nd generation lentiviral vector system to generate HIV-1, HIV-2, and control pseudovirions. For the generation of HIV-1 virions we utilized the following plasmids: the transfer vector pWOX-CMV-GFP, which was adapted to express mCherry instead of green fluorescent protein (GFP), the packaging plasmid psPAX2 (a kindly gift from Dr. D. Trono at the University of Geneva Medical School), and pMD.G coding for the vesicular stomatitis virus G protein. For HIV-2 pseudovirion production, the subsequent vectors were utilized: a ROD-based HIV-2 protein expression vector named CGP, CRU5SINCGW transfer vector featuring a GFP expression cassette under the CMV promoter (both are a kindly gift from Joseph P. Dougherty at the Robert Wood Johnson Medical School), and the pMD.G plasmid. In the case of control virion generation, the vector pTY-EFeGFP, a lentiviral transducing vector containing a GFP expression cassette under an EF1α promoter, and the pMD.G vectors were utilized.

Cell lines and maintenance:

Human embryonic kidney (HEK-293T) cell line were obtained from the American Type Culture Collection (ATCC) (Manassas, VA, USA). HEK-293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% FBS, 1% L-glutamine, and 1% penicillin/streptomycin. For the passage of HEK-293T cells, first the culture medium was discarded, and cells were rinsed with 1x Phosphate Buffer Saline (PBS). This was followed by the detachment of cells from the flask surface using 0.25% (w/v) Trypsin-EDTA solution. This was followed by centrifugation at 100 g, 24°C, for 5 minutes. Next, the cells were resuspended in fresh DMEM, and seeded back into the flask with a seeding density ranging from $0.5x10^6$ to $1.5x10^6$ cells. In the case of Jurkat cell line, the cells first were centrifuged at 100 g, 24°C, for 5 minutes. This was followed by resuspending the cell pellet into fresh RPMI (Sigma-Aldrich, St. Louis, MO, USA) medium and $1x10^6$ cells were seeded back into the flask.

HIV-1, HIV-2 and control pseudovirion production

For HIV-1 pseudovirion generation pWOX-CMV-mCherry, psPAX2, and pMD.G plasmids in a 3:2:1 ratio were utilized. CGP, CRU5SINCGW and pMD.G plasmids were employed in a 1:1:1 ratio for the generation of HIV-2 virions. pTY-EFeGFP and p.MDG in a 1:1 ratio were utilized for control production. The day before transfection HEK-293T cells were passaged in order to get an approximate confluence of 70% (~$3x10^6$ cells/flask (growth surface was 75cm^2/flask)) on the following day. HEK-293T cells were transfected with the above mentioned plasmid ratios, for the procedure we utilized the polyethylenimine (PEI) method. After the process the cells were incubated for five hours at 37°C with 5% CO2 in antibiotic-free, 1% FBS containing DMEM. Five hours later, the medium was replaced with DMEM supplemented with 10% FBS, 1% penicillin-streptomycin, and 1% glutamine. At 24, 48, 72 hours post transfection the medium containing the virions was collected and filtered through a 0.45-µm polyvinylidene fluoride filter (Merck Millipore, Darmstadt, Germany). The supernatants were pooled together, and concentrated via ultracentrifugation (100,000 × g for 2 hours and 10 minutes at 4°C). Next the pellet containing the virions were resuspended into PBS and stored at -70°C. A colorimetric assay based on enzyme-linked immunosorbent assay (ELISA) was utilized for the measurement of the reverse transcriptase activity (Roche Applied Science, Mannheim, Germany). To

determine the transduction units/ml (TU/ml) of the produced virions transduction experiment were carried out on HEK293T cells. The quantity of the control pseudovirions was subsequently calibrated to the reverse transcriptase (RT) equivalence of HIV-1 pseudovirions, ensuring equal virion quantities for the transduction experiments.

Transduction of- and RNA isolation from HEK-293T cells for transcriptomic analysis

On the day before the transduction, HEK-293T cells were seeded into 6-well plates ($5x10^5$ cells/well (growth surface was 9,6 cm^2/well)) in DMEM supplemented with 10% FBS, 1% L-glutamine, and 1% penicillin-streptomycin. On the day of transduction, the medium was removed, and the cells were transduced with 5 ng RT-equivalent of HIV-1/HIV-2 or control pseudovirions in serum and antibiotic-free media, supplemented with 8 µg/ml polybrene. At zero, two, eight, 12 and 26 hours post transduction the medium was discarded, and cells were washed with PBS before suspension into TRIzol reagent (Thermo Fisher Scientific, MA, USA). RNA isolation was carried out following the manufacturer's instructions. The quality of the RNA was assessed using the Agilent RNA 6000 Nano kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). High-throughput sequencing was carried out on the MGI DNBSEQ G400 (MGI tech, Shenzhen, China) sequencer using the MGIEasy RNA Library Prep Set at the Genomic Medicine and Bioinformatics Core Facility of the University of Debrecen.

Transduction of HEK-293T cells for proteomic analysis

The day before transduction, HEK293T cells were seeded into T-25 flasks at a density of 0.7-$1x10^6$ cells per flask (growth surface was 25cm^2/flask), in 5 ml of DMEM supplemented with 10% FBS, 1% L-glutamine, and 1% penicillin-streptomycin. At the next day, the medium was aspired, and the cells were transduced with 15 ng RT-equivalent of HIV-1/HIV-2 or control pseudovirions in serum and antibiotic-free media supplemented with 8 µg/ml polybrene. This was followed by an incubation at 37 °C with 5% CO2 for zero, two, eight, 12, and 26 hours. Next, the medium was discarded, and the cells were mechanically detached and suspended into PBS. After a brief centrifugation, the pellet was stored at -20°C for subsequent analysis.

Transduction of - and RNA isolation from Jurkat cells for transcriptomic analysis

On the day of the transduction Jurkat cells were seeded in 24-well plates at a density of 1.25 × 10^5 cells/well (working volume was 0.5-1.0 ml/well) in RPMI supplemented with 10% FBS and 1% L-glutamine, along with 8 µg/ml polybrene. This was followed by transduction with HIV-1, HIV-2 or control pseudovirions corresponding to 250,000 transduction units/ml (TU/ml). Jurkat cells were collected at zero, two, eight, 12, and 26 hours post-transduction, washed with PBS, and then lysed using TRIzol reagent (Thermo Fisher Scientific, MA, USA). RNA isolation was carried out following the manufacturer's instructions. The RNA quality was assessed using the Agilent RNA 6000 Nano kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). Subsequently, high-throughput sequencing was conducted on the NextSeq 500 sequencer (Illumina, CA, USA) at the Genomic Medicine and Bioinformatics Core Facility of the University of Debrecen.

Transcriptomic data analysis of HEK293T cells from the immediate and early phase of lentiviral transduction

RNA-seq raw fastq data underwent adaptor removal and quality trimming (phred score 30) utilizing Trimmomatic v0.36 (Bolger et al., 2014). The minimal trimmed read length was set to 36 bp. Quality check was conducted using FastQC v0.11.9 (Andrews S. Fastqc: A Quality Control Tool for High Throughput Sequence Data). The reads were mapped to the GRCh38 Human Genome Assembly reference genome using HISAT2 v2.1.0 (Langmead et al., 2009, Kim et al., 2015). The reference genome index files were created with Bowtie v.1.2.2 (Langmead et al., 2009). Aligned reads were quantified using FeatureCounts v2.0.1 (Liao et al., 2014, Yates et al., 2016). The obtained count matrix was used for analysis with R v4.2.3 (The R Development Core Team. R: A Language and Environment for Statistical Computing). Normalized counts, obtained through the median-of-ratios method were used for estimating differential expression (DE) with DESeq2 v1.38.3 (Love et al., 2014). To improve the accuracy of fold change estimates, we applied the "Adaptive Shrinkage" package version 2.2-63 (Stephens, 2017). Transcripts from the immediate early phase time points (0 and 2 hours) with an adjusted p-value lower than 0.05 and absolute log2 fold changes (abs(LFC)) >0.58 were considered as differentially expressed. In case of transcripts from the later time points (8, 12 and 26 hours) the adjusted p-value were the same but the (abs(LFC)) were set to >1. Gene ontology (GO) enrichment analysis of the DE genes was conducted using the clusterProfiler package (Yu et al., 2012).

GeLC-MS/MS analysis

Lysis of the HIV-1, HIV-2 or control pseudovirion transduced cells was done in 100 µL lysis buffer (50 mM Tris pH 8.3, 1 mM EDTA, 17 mM β-mercaptoethanol, 0.5% (v/v) Triton-X100) using three freeze–thaw cycles. Bradford method was utilized to determine the protein concentration, and 100 µg of protein for each case underwent in-gel digestion followed by liquid chromatography–tandem mass spectrometry (GeLC–MS/MS) analysis (Dzieciatkowska et al., 2014). Briefly, samples were run into a 5% SDS-polyacrylamide gel using a 100 V current for 20 min. The proteins were stained with PageBlue Protein Staining solution (Thermo Scientific, Waltham, MA, USA), and the stained gel slice was excised, separated into three equal portions and submitted to in-gel trypsin digestion. Reduction was performed with 20 mM dithiothreitol (Bio-Rad, Hercules, CA, USA) for 1 h at 56 °C, followed by alkylation with 55 mM iodoacetamide (Bio-Rad, Hercules, CA, USA) for 45 min at room temperature in the dark. Overnight trypsin digestion was performed at 37 °C utilizing stabilized MS-grade TPCK-treated bovine trypsin (ABSciex, Framingham, MA, USA). The digested peptides were extracted and dried in a speed-vac (Thermo Scientific, Waltham, MA, USA). The peptides were re-dissolved in 33 µL 1% formic acid (VWR Ltd., Radnor, PA, USA) before LC–MS/MS analysis. BCA method was used to determine the peptide concentration of the samples. Before mass spectrometry analyses, the samples were spiked with equal amounts of indexed retention time (iRT) peptide mixtures (Biognosys, Schlieren, Switzerland), and the samples were analysed in duplicate.

Prior to the mass spectrometric analysis, peptides were separated in a 180 min water/acetonitrile gradient using an Easy nLC 1200 nano UPLC (Thermo Scientific, Waltham, MA, USA). The peptide mixtures were desalted in an ACQUITY UPLC Symmetry C18 trap column (20 mm × 180 µm, 5 µm particle size, 100 Å pore size, Waters, Milford, MA, USA), followed by separation in Acclaim PepMap RSLC C18 analytical columns (150 mm × 50 µm × 2 µm particle size, 100 Å pore size, Thermo Scientific, Waltham, MA, USA). Chromatographic separation was performed utilizing a gradient of 5–7% solvent B over 5 min, followed by a rise to 15% of solvent B over 50 min and then to 35% solvent B over 60 min. Thereafter, solvent B was increased to 40% over 28 min and then to 85% over 5 min, followed by a 10 min rise to 85% of solvent B, after which the system returned to 5% solvent B in 1 min for a 16 min hold-on. Solvent A was 0.1% formic acid in LC water (Sigma, St. Louis, MO, USA); solvent B was 95% acetonitrile (Sigma, St. Louis, MO, USA) containing 0.1% formic acid. The flow rate was set to 300 nL/min.

Data-dependent analyses were conducted using an Orbitrap Fusion mass spectrometer (Thermo Scientific, Waltham, MA, USA). The 14 most abundant multiply charged positive ions were selected from each survey MS scan using a scan range of 350–1600 m/z for MS/MS analyses (Orbitrap analyzer resolution: 60,000, AGC target: 4.0 × 105, acquired in profile mode). Collision-induced dissociation (CID) fragmentation was performed in the linear ion trap with 35% normalized collision energy (AGC target: 2.0 × 103, acquired in centroid mode). Dynamic exclusion was enabled during the cycles (exclusion time: 45 s).

Data analysis of mass spectrometry

The acquired LC-MS/MS data were used for protein identification with the utilization of MaxQuant 1.6.2.10 software (Cox & Mann, 2008) searching against the Human SwissProt database (release: 2020.02, 20394 sequence entries), the HIV-1 and HIV-2 SwissProt databases (release: 2020.02, 381 sequence entries for HIV-1 and 109 sequence entries for HIV-2), and against the contaminants database provided by the MaxQuant software. Cys carbamidomethylation, Met oxidation, and N-terminal acetylation were set as variable modifications with a maximum of 2 allowed missed cleavage sites. Results were imported into Scaffold 4.8.9 software (ProteomeSoftware Inc., Portland, OR, USA). Proteins were considered acceptable with at least 3 identified peptides using 1% protein false discovery rate (FDR) and 0.1% peptide FDR. For label-free quantification, the normalized total precursor intensities were utilized, and quantitative values of the identified proteins were normalized to the quantitative values of the iRT mixture. Data were further normalized to the concentration of the peptides in the samples determined by BCA method after the digestion.

A mass of mixed effects of ANOVA models were applied, one for each protein, to select the significantly different protein quantities between the investigated groups. Sample and measurement repetitions were modelled as random effects, while the transduced cell groups were modelled as fixed effects (Oberg & Vitek, 2009). Following linear model fitting, post-hoc tests were applied to determine the p-values of group differences, and significant results with an FDR < 0.05 criteria were retained.

Proteomic data analysis of HEK293T cells from the early phase of lentiviral transduction

The compilation of identified proteins with intensity values was exported from Scaffold and further processed in the R environment (v4.3.1) (The R Development Core Team. R: A

Language and Environment for Statistical Computing). Normalized values were calculated accounting for variations in overall sample protein concentrations (determined post-digestion using the BCA method), injection volumes and the sum of detected protein intensities per sample. A series of mixed-effects ANOVA models were employed, one for each protein, to discern statistically significant differences in protein abundances between the investigated groups. Sample and measurement replicates were treated as random effects, meanwhile the different transduced cell groups were modelled as fixed effects (Oberg & Vitek, 2009). We ran the computations using the emmeans (v1.8.8) and lme4 (1.1.34) packages (Bates et al., 2015). Following linear model fitting, post-hoc tests were applied to determine the p-values of group differences and significant results with an FDR < 0.05 criteria were retained. The protein coding transcript annotation was obtained from the daily updated 'gene_info', 'gene2go' files available on the NCBI FTP site (https://ftp.ncbi.nlm.nih.gov/gene/DATA/, accessed on 15 November 2023) and reorganized using in-house shell scripts. GO enrichment analysis was carried out with the topGO R package (v2.52.0) using a custom-generated gene-to-GOs mapping file (Alexa et al., 2006). The protein universe consisted of the entire set of proteins (807) detected in all 26 h post-transduction samples; significant proteins were filtered according to the FDR < 0.05 threshold. In the topGO algorithm, the default 'weight01' method was chosen with the Kolmogorov–Smirnov statistical test and selected the enriched GO terms having FDR-adjusted p-values < 0.05. Data visualization was performed with ggplot2 (v3.4.3), ggrepel (v0.9.3), ggpubr (v0.6.0), GOplot (v1.0.2) (Elio Campitelli Ggnewscale: Multiple Fill and Colour Scales in 'ggplot2'. R Package Version 0.4.9.), ggnewscale (v0.4.9), VennDiagram (v1.7.3) R packages.

Transcriptomic data analysis of Jurkat cells from the early phase of lentiviral transduction

The reads were mapped to the GRCh38 Human Genome Assembly reference genome using Hisat2 v2.1.0 (Langmead et al., 2009). The BAM files generated during alignment were imported into StrandNGS software for analysis. The quantification of the reads were done with the integrated DESeq algorithm. Moderated T-test were utilized for the determination of the DETs. The downstream analysis consisted of genes passing significance filters with adjusted p-value (padj) < 0.05 and absolute log2 fold-change (abs(LFC)) > 1. For the correction of errors originating from the multiple testing Benjamini-Hochberg FDR analysis were utilized.

GO analysis were carried out in CytoScape 3.4.0 with the utilization of the ClueGO v2.3.5 application. The GO biological process categories were determined with the use of two sided hypergeometric test and Bonferroni step down correction.

Results

Analysis of transcriptomic changes in the early phase of HIV-1 transduction in HEK-293T cells

Compared to cells transduced with the control vector, at the 0-hour time point, HIV-1 influenced the expression of a total of 2676 transcripts from which 1533 were up-, and 1143 were downregulated. From the upregulated ones, 1256 encoded proteins, 226 non-coding RNAs, 1 mitochondrial RNA, 42 pseudogenes and 8 were the products of unknown genes. From the downregulated transcripts, 710 were protein-coding, 294 non-coding RNAs, 11 were mitochondrial RNAs, 94 were pseudogenes, 2 were ribozymes, and 32 were unknown transcripts.

Two hours after transduction, HIV-1 induced significant changes on 4551 transcripts. Among these, the expression of 2366 transcripts were upregulated, from which 1916 were encoding for proteins, 98 coding for pseudogenes, and 334 were non-coding RNAs. Additionally, 18 gene products were categorized as unknown transcripts. Moreover, 2185 genes were downregulated by HIV-1, among these were 1634 transcripts coding for proteins, 153 for pseudogenes, 363 for non-coding RNA transcripts, 18 for mitochondrial RNA, one for ribozyme, and 16 were products of unknown genes.

Compared to the control, 8 hours post-transduction, HIV-1 altered the cellular level of 158 transcripts, out of which 115 were increased and 43 were decreased. Amongst the increased ones; 71 were protein-coding, 23 were non-coding RNAs, 5 were mitochondrial RNAs, 12 pseudogene, and 4 derived from uncategorized genes. Amongst the decreased ones, 32 encoded proteins, 7 non-coding RNAs and 4 pseudogenes.

In comparison to cells transduced with the control l, at the 12-hour mark, HIV-1 altered the expression of a total of 114 transcripts. Out of these, 18 were positively-, and 96 negatively regulated. Of the positively regulated ones, 11 coded for proteins, 6 non-coding RNAs and 1 was the product of pseudogene. Of the downregulated transcripts, 59 were products of protein -coding transcript, 27 were non-coding RNAs, 5 were mitochondrial RNAs, 1 was a pseudogene and 4 were unknown gene products.

26 hours after transduction, HIV-1 significantly altered the expression of 380 transcripts, from which 48 were upregulated and 332 were downregulated. From the upregulated transcripts 26 encoded for proteins, 14 for non-coding RNAs, 7 products of pseudogenes and 1 coded for an unknown gene product. From the 332 downregulated transcripts, 264 were protein-coding, 56 were non-coding RNAs, 8 were pseudogenes and 4 were transcripts derived from unknown genes.

Differentially induced transcripts by HIV-1 in the immediate early phase of transduction in HEK-293T cells

At zero-hour cells treated with HIV-1 showed upregulation of insulin receptor substrate 4 (*IRS4*), host cell factor C1 (*HCFC1*), heparan sulfate proteoglycan 2 (*HSPG2*), CD109 molecule, 2'-5'-oligoadenylate synthetase 3 (*OAS3*), nidogen 1 (*NID1*), lysine methyltransferase 2D (*KMT2D*), SRY-box transcription factor 5 (*SOX5*), RNA polymerase II subunit A (*POLR2A*) and FRY microtubule binding protein (*FRY*).

H4 clustered histone 3 (*H4C3*), argininosuccinate synthase 1 (*ASS1*) immunoglobulin superfamily member 6 (*IGSF6*), pleckstrin homology like domain family A member 3 (*PHLDA3*), interleukin 3 receptor subunit alpha (*IL3RA*), apolipoprotein E (*APOE*), ankyrin repeat domain 18A (*ANKRD18A*), cholinergic receptor nicotinic gamma subunit (*CHRNG*), major facilitator superfamily domain containing 4B (*MFSD4B*) and mannosidase alpha class 1A member 2 (*MAN1A2*) were downregulated.

At the 2-hour time-point, HIV-1 resulted in the up-regulation of solute carrier family 7 member 11 (*SLC7A11*), assembly factor for spindle microtubules (*ASPM*), sacsin molecular chaperone (*SACS*), CD109 molecule, zinc finger and BTB domain containing 41 (*ZBTB41*), BRCA2 DNA repair associated (*BRCA2*), testis expressed 15, meiosis and synapsis associated (*TEX15*), zinc finger with KRAB and SCAN domains 8 (*ZKSCAN8*), chaC glutathione specific gamma-glutamylcyclotransferase 1 (*CHAC1*) and protein prenyltransferase alpha subunit repeat containing 1 (*PTAR1*), respectively. On the other hand, H4 clustered histone 3 (*H4C3*) epithelial membrane protein 3 (*EMP3*), pleckstrin homology like domain family A member 3 (*PHLDA3*), argininosuccinate synthase 1 (*ASS1*), inhibitor of DNA binding 1, and 3 HLH protein (*ID1/3*), ribosomal protein L39 like (*RPL39L*), SHC adaptor protein 2 (*SHC2*), syntaxin 8 (*STX8*), and regulator of G protein signaling 16 (*RGS16*) were downregulated.

Differentially induced transcripts by HIV-1 in the early phase of transduction in HEK-293T cells

At the eight hour time-point, HIV-1 upregulated the expression of the leucine rich repeats and IQ motif containing 1 (*LRRIQ1*), centromere protein E (CENPE), A-kinase anchoring protein 9 (*AKAP9*), ankyrin repeat domain 12 (*ANKRD12*), coiled-coil domain containing 88A (*CCDC88A*), biorientation of chromosomes in cell division 1 like 1 (*BOD1L1*), GRIP and coiled domain containing 2 (*GCC2*), ankyrin repeat domain 26 (*ANKRD26*), structural maintenance of chromosomes 4 (*SMC4*) and dopamine receptor D4 (*DRD4*). In the same time-point, WD repeat domain 38 (*WDR38*), fos proto-oncogene AP-1 transcription factor subunit (*FOS*), bolA family member 2B (*BOLA2B*), cyclin dependent kinase inhibitor 1A (*CDKN1A*), inhibitor of DNA binding 3, HLH protein (*ID3*), H4 clustered histone 3 (*H4C3*), collagen type XI alpha 1 chain (*COL11A1*), epithelial membrane protein 3 (*EMP3*), RELB proto oncogene NF-kB subunit (*RELB*) and pleckstrin homology like domain family A member 3 (*PHLDA3*) were downregulated.

In decreasing order of magnitude, the top 10 upregulated protein-coding transcripts induced by HIV-1 at the 12-hour time-point were the secretogranin III (*SCG3*), forkhead box D4 (*FOXD4*), neuregulin 4 (*NRG4*), leucine rich repeat and coiled-coil and coiled centrosomal protein 1 (*LRRCC1*), nucleosome assembly protein 1 like 2 (*NAP1L2*), BRCA2 DNA repair associated (*BRCA2*), nibrin (*NBN*), taste 2 receptor member 20 (*TAS2R20*), cytochrome c oxidase assembly factor COX20 (*COX20*) and transmembrane protein 145 (*TMEM145*). Meanwhile, at the same time-point, H4 clustered histone 3 (*H4C3*), collagen type XI alpha 1 chain (*COL11A1*), transmembrane protein 132E (*TMEM132E*), calcium/calmodulin dependent protein kinase ID (*CAMK1D*), SH2 domain containing 3C (*SH2D3C*), tetraspanin 11 (*TSPAN11*), programed cell death 11 (*PDCD11*), NOP9 nucleolar protein (*NOP9*), early growth response 1 (*EGR1*) and kinesin family member 1A (*KIF1A*) were downregulated.

At the 26 hours post-transduction, the expression of heme oxygenase 1 (*HMOX1*), oxidative stress induced growth inhibitor 1 (*OSGIN1*), VGF nerve growth factor inducible (*VGF*), NAD(P)H quinone dehydrogenase 1 (*NQO1*), nucleosome assembly protein 1 like 2 (*NAP1L2*), heat shock protein family A (Hsp70) member 1A (*HSPA1A*), dehydrogenase/reductase 2 (*DHRS2*), ETS variant transcription factor 4 (*ETV4*), metallotioenin 2A (*MT2A*) and glutamate-cysteine ligase modifier subunit (*GCLM*) were upregulated; while H4 clustered histone 3 (*H4C3*), kelch domain containing 7B (*KLHDC7B*), inositol polyphosphate-5-phosphatase D (*INPP5D*), glutamate ionotropic receptor kainate type subunit 3 (*GRIK3*), transmembrane protein 132E (*TMEM132E*), ATP binding cassette subfamily G member 1 (*ABCG1*), AT-hook transcription factor (*AKNA*), dehydrogenase/reductase 3 (*DHRS3*), hes family bHLH

transcription factor 5 (*HES5*) and ABI family member 3 binding protein (*ABI3BP*) were found to be downregulated by HIV-1.

Analysis of transcriptomic changes in the early phase of HIV-2 transduction in HEK-293T cells

At 0 hours post-transduction, HIV-2 significantly altered the expression of a total of 4075 transcripts. Out of these the level of 2015 were increased while 2060 were decreased. Amongst the upregulated transcripts, 1672 were found to be coding for proteins, 240 were non-coding RNAs, 92 products of pseudogenes, 1 product of an Ig-V gene and 10 were products of unknown genes. Out of the downregulated transcripts, 1646 encoded for proteins, 273 non-coding RNAs, 12 mitochondrial RNAs, 106 pseudogenes, 2 ribozymes and 21 were products of uncategorized genes.

Compared to the control, the expression of 3422 genes were differentially altered in cells transduced with HIV-2 at two hours after transduction. 2114 transcripts were upregulated, amongst these were 1716 transcripts which encoded for proteins, 128 for pseudogenes, 261 for non-coding RNAs and nine were products of uncharacterized genes. On the other hand, HIV-2 downregulated the expression of 1308 genes, consisting of 875 protein-coding transcripts, 306 non-coding RNAs, and 16 genes encoding mitochondrial RNAs. Additionally, 83 were products of pseudogenes, one ribozyme transcript, and 27 transcripts were categorized as unknown genes.

At 8 hours, HIV-2 managed to alter the expression of 283 transcripts from which 211 were up-, and 72 were downregulated. Of the upregulated ones 99 encoded for proteins, 31 non-coding RNAs, 3 mitochondrial RNAs, 75 pseudogenes and 3 were products of unknown genes. Amongst the downregulated transcripts, 43 encoded for proteins, 27 were non-coding RNAs and 2 pseudogenes.

12 hours post transduction, HIV-2 differently regulated the expression of 299 transcripts from which 225 were up-, and 74 were downregulated. Of them 107 were protein-coding, 26 were non-coding RNAs, 91 were pseudogenes, and 1 coded for uncategorized gene product. From the downregulated transcripts, 46 were protein coding, 24 were non-coding RNAs, 1 was mitochondrial RNA, 2 were products of pseudogene, and 1 was an unknown gene product.

At the 26 hours time-point, HIV-2 altered the expression of a total of 182 transcripts from which 121 were up- and 61 were downregulated. Out of the upregulated transcripts 35 coded for

proteins, 8 non-coding RNAs and 78 were products of pseudogenes. Of the downregulated transcripts 35 encoded proteins, 21 non-coding RNAs, 4 mitochondrial RNAs and 1 was a product of unknown genes.

Differentially induced transcripts by HIV-2 in the immediate early phase of transduction in HEK-293T cells

HIV-2, at the zero-hour time-point upregulated the expression of collagen type I alpha 2 chain (*COL1A2*), keratin 5 (*KRT5*), collagen type VI alpha 3 chain (*COL6A3*), keratin 7 (*KRT7*), decorin (*DCN*), S100 calcium binding protein A2 (*S100A2*), collagen type III alpha 1 chain (*COL3A1*), thrombospondin 1 (*THBS1*), keratin 14 (*KRT14*) and S100 calcium binding protein A6 (*S100A6*). Meanwhile HIV-2 downregulated the expression of H4 clustered histone 3 (*H4C3*), argininosuccinate synthase 1 (*ASS1*), pleckstrin homology like domain family A member 3 (*PHLDA3*), ankyrin repeat domain 36B (*ANKRD36B*), ankyrin repeat domain 36C (*ANKRD36C*), src kinase associated phosphoprotein 1 (*SKAP1*), serine dehydratase like (*SDSL*), epithelial membrane protein 3 (*EMP3*), mannosidase alpha class 1A member 2 (*MAN1A2*) and family with sequence similarity 133 member B (*FAM133B*). At the 2 hour time-point, HIV-2 increased the expression of collagen type I alpha 2 chain (*COL1A2*), collagen type VI alpha 3 chain (*COL6A3*), thrombospondin 1 (*THBS1*), keratin 7 (*KRT7*), keratin 5 (*KRT5*), collagen type III alpha 1 chain (*COL3A1*), keratin 14 (*KRT14*), chaC glutathione specific gamma-glutamylcyclotransferase 1 (*CHAC1*), S100 calcium binding protein A2 (*S100A2*), and insulin receptor substrate 4 (*IRS4*). At the same time the expression of H4 clustered histone 3 (*H4C3*), ankyrin repeat domain 36C (*ANKRD36C*), cholinergic receptor nicotinic gamma subunit (*CHRNG*), ankyrin repeat domain 36B (*ANKRD36B*), ankyrin repeat domain 18A (*ANKRD18A*), ankyrin repeat domain 36 (*ANKRD36*), immunoglobulin superfamily member 6 (*IGSF6*), family with sequence similarity 133 member B (*FAM133B*), WD repeat domain 38 (*WDR38*), and lysine rich nucleolar protein 1 (*KNOP1*) were downregulated.

Differentially induced transcripts by HIV-2 in the early phase of transduction in HEK-293T cells

At the eight-hour time-point, HIV-2 upregulated the expression of collagen type I alpha 2 chain (*COL1A2*), keratin 5 (*KRT5*), serpin family E member 1 (*SERPINE1*), keratin 14 (*KRT14*), S100 calcium binding protein A6 (*S100A6*), keratin 7 (*KRT7*), decorin (*DCN*), thrombospondin 1 (*THBS1*), ETS variant transcription factor 5 (*ETV5*) and collagen type VI alpha 3 chain

(*COL6A3*). Meanwhile H4 clustered histone 3 (*H4C3*), inhibitor of DNA binding 3, HLH protein (*ID3*), ectodysplasin A2 receptor (*EDA2R*), AHNAK nucleoprotein (*AHNAK*), transmembrane protein 132E (*TMEM132E*), inhibitor of DNA binding 1, HLH protein (*ID1*), AHNAK nucleoprotein 2 (*AHNAK2*), paralemmin 3 (*PALM3*), hes family bHLH transcription factor 5 (*HES5*) and ArfGAP with dual PH domains 1 (*ADAP1*) were downregulated in the same time-point.

12 hours post-transduction, the most upregulated protein-coding transcripts were the collagen type I alpha 2 chain (*COL1A2*), keratin 7 (*KRT7*), decorin (*DCN*), alanyl aminopeptidase, membrane (*ANPEP*), dermatopontin (*DPT*), integrin subunit beta like 1 (*ITGBL1*), keratin 5 (*KRT5*), transforming growth factor beta induced (*TGFBI*), S100 calcium binding protein A6 (*S100A6*) and fatty acid binding protein 4 (*FABP4*). Additionally, the expression level of H4 clustered histone 3 (*H4C3*), ATP binding cassette subfamily G member 1 (*ABCG1*), inhibitor of DNA binding 3, HLH protein (*ID3*), ankyrin repeat domain 36C (*ANKRD36C*), transmembrane protein 132E (*TMEM132E*), tetraspanin 11 (*TSPAN11*), ankyrin repeat domain 18A (*ANKRD18A*), regulating synaptic membrane exocytosis 3 (*RIMS3*), pleckstrin homology like domain family A member 3 (*PHLDA3*) and calcium/calmodulin dependent protein kinase ID (*CAMK1D*) were downregulated by HIV-2.

At the 26-hour time-point, HIV-2 increased the cellular transcription of collagen type I alpha 2 chain (*COL1A2*), keratin 7 (*KRT7*), keratin 5 (*KRT5*), keratin 14 (*KRT14*), decorin (*DCN*), S100 calcium binding protein A6 (*S100A6*), serpin family E member 1 (*SERPINE1*), S100 calcium binding protein A2 (*S100A2*), SLX1 homolog B, structure-specific endonuclease subunit (*SLX1B*), and thrombospondin 1 (*THBS1*) were upregulated by HIV-2, while H4 clustered histone 3 (*H4C3*), early growth response 1 (*EGR1*), kelch domain containing 7B (*KLHDC7B*), RELB proto-oncogene NF-κB subunit (*RELB*), ATP binding cassette subfamily G member 1 (*ABCG1*), transmembrane protein 132E (*TMEM132E*), SHC adaptor protein 2 (*SHC2*), dehydrogenase/reductase 3 (*DHRS3*), tetraspanin 11 (*TSPAN11*) and AT-hook transcription factor (*AKNA*) were found to be downregulated.

Gene ontology analysis of the significantly altered transcripts in the immediate early-phase of HIV transduction in HEK-293T cells

To further categorize the biological functions of the detected transcripts Gene ontology (GO) enrichment analysis was carried out. We utilized the protein-coding transcripts that showed statistically significant differential expression ($p > 0.05$, $\log2FC > 0.58$) at the two-hour time-

point from HIV-1 and HIV-2 transduced cells in comparison to control treated ones. The results revealed that both HIVs similarly affected GO terms such as protein serine/threonine kinase activity, structural constituent of ribosome, GTPase activator activity, nucleoside triphosphatase regulator and ubiquitin-like protein transferase activity. Amongst the GO terms influenced differentially by HIV-1 were rRNA and tau protein binding, tau-protein kinase activity, and catalytic activity, acting on DNA. Conversely, HIV-2 affected protein coding genes were involved in Rho GTPase binding, DNA-binding transcription factor binding, transcription coactivator activity, and RNA polymerase II-specific DNA-binding transcription factor binding.

Gene ontology analysis of the significantly altered transcripts in the early phase of HIV transduction in HEK-293T cells

Similarly, in immediate early-phase of transduction, GO analysis was carried out on identified transcripts using the most significant differentially expressed ($p < 0.05$, $log2FC > 1$) protein-coding transcripts at the 8, 12, and 26-hour time points from both HIV-1 and HIV-2 transduced cells, compared to control-transduced ones. This analysis revealed further differences between HIV-1 and HIV-2 transduced cells in comparison to the control.

At the 8-hour mark, HIV-1 affected transcripts which are associated with the cell cycle, positive regulation of RNA metabolic processes and positive regulation of metabolic processes. Meanwhile, HIV-2 influenced the expression of genes involved in tissue development, extracellular matrix organization and cell adhesion.

At the 12-hour mark, no significantly altered GO terms were detected in HIV-1 transduced HEK-293T cells. In contrast, HIV-2 affected the expression of transcripts involved in cell differentiation, cell adhesion, and cell migration.

At 26 hours post-transduction, HIV-1 altered transcripts involved in homeostatic processes, response to toxic substances and response to oxidative stress. In comparison, HIV-2 influenced transcripts associated with peptide cross-linking, negative regulation of cell migration and negative regulation of cell motility.

Proteomic changes in the immediate early-phase post viral transduction in HEK-293T cells

Over 1000 protein were identified by analysis of transduced host cell proteome. Following false discovery rate analysis, seven proteins exhibited statistically significant downregulation in comparison to the control. Notably, five proteins were found to be downregulated in HIV-1

transduced cells, and an additional two proteins showed decreased expression in cells transduced with HIV-2 compared to the control.

Both HIV-1 and HIV-2 downregulated the expression of mitochondrial 60kDa heat shock protein (HSPD1), non-POU domain-containing octamer binding protein (NONO), heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1), serine/arginine-rich splicing factor 6 (SRSF6) and histone H1.4 (H1-4). Furthermore, complex protein 1 subunit theta (CCT8) and FK506-binding protein (FKBP4) showed significant decrease only in HIV-2 transduced cells compared to HIV-1.

Following the identification of the seven significantly downregulated proteins, interaction network and biological process enrichment analyses were carried out using Cytoscape 3.8.1 **(**Shannon et al., 2003). The integrated STRING database was used to generate the physical interaction network, with a confidence level of 0.7–100 first shell interactors was queried, and for the gene ontology enrichment Cytoscape's ClueGO v2.5.7. plugin was used **(**Bindea et al., 2009). In the gene ontology analysis of the enriched network compromising 107 proteins altogether, functional clusters of GO terms were first generated. ClueGO utilizes term similarity to define functional groups of multiple terms. In our analysis, initial group size was set to three terms (two being the default value), and the percentage for group merge was left at the default value of 50%. Further, 33 of the 72 resulting GO terms were grouped into six functional clusters, the rest of them, 39 in total, did not reach the group merge threshold. However, among these was the GO term mRNA splicing (GO:00000398) including 97 proteins in the enriched network. Due to the redundancy of GO terms in the clusters, we selected seven biological processes that were the most representative and significant according to their *p*-values corrected with the Bonferroni step-down method ($p < 0.05$). Altogether, 105 proteins were covered by these biological processes in the interaction network. Moreover, 77 of these 105 proteins were also detected by MS/MS, indicating a wider interaction network centred around the significantly downregulated proteins. However, only 67 of these 77 proteins had corresponding quantitative data of sufficient quality to enable statistical analysis, as they could not be quantified in some of the replicate measurements.

The terms we employed for matching, encompassing processes such as ribonucleoprotein complex assembly (GO:0022618), RNA transport (GO:0050658), regulation of mRNA stability (GO:0043488), regulation of DNA metabolic processes (GO:0051052), protein folding (GO:0006457), mRNA splicing via spliceosome (GO:0000398), and cellular response to stress (GO:0033554).

Proteomic changes in the early phase of viral transduction in HEK-293T cells

Through mass spectrometry analysis a total of 871 proteins were detected at eight, 817 at 12, and 810 at 26 hours post-transduction. At eight hours, proteomic analysis revealed no significant changes in the cellular proteome. However, at 12 hours, we detected 17 proteins which showed difference in their expression, with 5 altered by HIV-1, 4 by HIV-2, and 8 by both viruses compared to the control. Among the proteins influenced by HIV-1 were the ATP synthase F1 subunit beta (ATP5F1B), valosin-containing protein (VCP) and synaptotagmin-binding cytoplasmic RNA interacting protein (SYNCRIP). HIV-2 significantly altered the expression of non-SMC condensin I complex subunit H (NCAPH), phosphoribosylformylglycinamidine synthase (PFAS), valyl-tRNA synthetase 1 (VARS1) and proline-rich coiled-coil 2A (PRRC2A). Altered proteins by both viruses included Zyxin (ZYX), arginyl-tRNA synthetase 1 (RARS1), heat shock protein family A member 1A (HSPA1A), nucleophosmin 1 (NPM1), and deoxyuridine triphosphatase (DUT). At 26 hours, a total of 117 proteins were differentially regulated following transduction with HIVs. HIV-1 significantly altered the expression of 25 proteins, meanwhile HIV-2 influenced 39 proteins. Moreover, both HIVs altered the expression of 48 proteins respectively, compared to the control. Additionally, 5 proteins were exclusively regulated by HIV-2 in comparison to HIV-1. HIV-1 altered proteins included insulin-like growth factor 2 receptor (IGF2R), DEAD-box helicase 3 X-linked (DDX3X), protein disulfide isomerase family A member 3 (PDIA3), HIV-1 Tat Specific Factor 1 (HTASF1), lysyl-tRNA synthetase 1 (KARS1), proteasome 20S subunit alpha 2 (PSMA2), and ATPase Family AAA Domain Containing 3A (ATAD3A). HIV-2 managed to influence the cellular level of nucleolin (NCL), eukaryotic translation elongation factor 1 alpha 1 (EEF1A1) and serine and arginine-rich splicing factor 1 (SRSF1). Both viruses affected the expression of heterogeneous nuclear ribonucleoprotein A2/B1 (HNRNPA2B1), heterogeneous nuclear ribonucleoprotein K (HNRNPK), serine and arginine-rich splicing factor 2 (SRSF2), dynein cytoplasmic 1 light intermediate chain 1 (DYNC1LI1) and eukaryotic translation initiation factor 2B subunit delta (EIF2B4). When HIV-2 was compared to HIV-1, 17 proteins had showed altered expression, including transferrin receptor (TFRC), heat shock protein family A members, ribosomal protein L23a (RPL23A) and RAN binding protein 1 (RANBP1).

Gene ontology analysis of proteins in the 12 and 26 hours time-points in HEK-293T cells

To better understand the proteomic changes in transduced cells, GO enrichment analysis was carried out on the HIV-1 and HIV-2 transduced samples at 12 and 26 hours post-transduction.

In comparison to the control, at 12 hours, HIV-1 altered the expression of proteins involved in ubiquitin protein ligase binding, ATP metabolic processes and mRNA splicing via spliceosome. Meanwhile, HIV-2 affected the expression of proteins associated with regulation of localization, regulation of cellular response to stress and negative regulation of macromolecule biosynthetic processes. At the 26-hour time-point, HIV-1 influenced the cellular level of proteins related to the RNA polymerase II transcription regulator complex, viral genome replication and response to virus. Moreover, HIV-2 altered proteins associated with viral genome replication, regulation of viral genome replication, and the spliceosomal small nuclear ribonucleoprotein (snRNP) complex.

Differentially induced transcripts in the early-phase of transduction in Jurkat cells

We detected no significantly altered transcript at the zero-, and two-hour time-points. Eight hours post-transduction, HIV-1 altered the expression of 39 transcripts. Out of the 39, 38 were downregulated while only 1, a non-coding RNA transcript was upregulated. Of the downregulated transcripts, 29 coded for proteins, 3 for pseudogenes and 6 for various RNA transcripts. Among the differentially downregulated transcripts were the argininosuccinate synthase 1 (*ASS1*), cyclin G2 (*CCNG2*) and protein phosphatase, Mg2+/Mn2+ dependent 1K (*PPM1K*). At twelve hours, HIV-1 differentially regulated 45 transcripts compared to the control transduced Jurkat cells. From the 45, 43 were down- and 2 were upregulated. Amongst the downregulated transcripts, 35 coded for proteins, 2 for pseudogenes and 6 were coding for various RNA products. Meanwhile, two non-coding RNA transcripts were upregulated. The differentially downregulated transcripts include methylenetetrahydrofolate reductase (*MTHFR*) and spondin 2 (*SPON2*). At the 26-hour time- point, the expression of 42 transcripts were altered by HIV-1 with 40 down- and 2 upregulated gene products. Amongst the differentially downregulated transcripts, 36 coded for proteins, 1 for pseudogenes and 3 for RNA products. Additionally, 1 protein and 1 pseudogene encoding transcript were upregulated. The protein-coding downregulated transcripts included cluster of differentiation 300a (*CD300a*), laminin subunit gamma 2 (*LAMC2*) and serine peptidase inhibitor, Kazal type 2 (*SPINK2*).

Eight hours post-transduction, HIV-2 altered the expression of 15 transcripts with 12 being down-, and 3 upregulated. Amongst the differentially downregulated transcripts 10 coded for proteins, 1 for a pseudogene and 1 for an RNA product. Moreover, 1 pseudogene and 2 assorted RNA products were upregulated. There were no unique genes regulated only by HIV-2 at the 8-hour time-point other than a pseudogene and an RNA product compared to control transduced

cells. 12 hours post-transduction, HIV-2 affected the expression of 23 genes out of which 21 were down-, and 2 were upregulated. Of the downregulated ones, 19 coded for proteins and 2 for different RNA transcripts. Additionally, 1 protein coding and 1 non-coding RNA transcript were upregulated. The differentially regulated genes by HIV-2 at the 12-hour time point included cysteine rich angiogenic inducer 61 (*CYR61*) and ubiquinol-cytochrome c reductase complex assembly factor 3 (*UQCC3*). At the 26-hour time point, HIV-2 altered the expression of 36 transcripts, with 33 being down-, and 3 being upregulated. Of the downregulated ones 31 coded for proteins and 2 for different RNA products. Of the upregulated ones, 1 coded for protein and 2 for pseudogenes. The downregulated transcripts included the early growth response 1 (*EGR1*) and TOX high mobility group box family member 2 (*TOX2*).

Gene ontology analysis of the significantly altered transcripts at the early phase of HIV transduction in Jurkat cells

GO analysis was carried out on the significantly altered transcripts at 26 hours post-transduction. Our analysis revealed that HIV-1 altered the expression of genes related to steroid biosynthetic process, cholesterol biosynthetic process and isoprenoid biosynthetic process. Meanwhile, HIV-2 only affected genes which are involved in the cholesterol biosynthetic process. This indicates that HIV-1 has a greater effect on the host cell at 26 hours post-transduction compared to HIV-2 at the same time-point.

Discussion

HIV has infected more than 40 million people at the time of writing this dissertation. The majority of infections are caused by HIV-1, however, a significant number of cases are caused by HIV-2. There are multiple similarities between the two HIV viruses such as their method of transmission and pathogenesis. However, there are also numerous differences in their replication dynamics and course of infection. There is an initial increased viral production phase in the early stage in case of HIV-2 infection, followed by an extended latency period. Additionally, compared to HIV-1, HIV-2 exhibits reduced viral replication and a slowed progression of infection towards AIDS (Campbell-Yesufu & Gandhi, 2011**)** (MacNeil et al., 2007**)** (Bock & Markovitz, 2001**)**. The variables contributing to these unique replication dynamics and the mechanism influencing the preference for an extended latency period of HIV-2 infection are not fully understood.

There are multiple studies focusing on the cellular effects of HIV-1, however there are very few studies in regards to HIV-2. Additionally, lentivirus based vectors are valuable clinical and research and thereby, understanding the intracellular effects of the vectors on the host cell is important for their safe and targeted application. We set the goal to analyse the effect of HIV-1 and 2-based lentiviral vectors on the cellular transcriptome and proteome at the early phase of transduction, focusing on 5 distinct time-points.

Firstly, in order to define the changes in the gene expression of the HIV-1 and HIV-2 transduced cells, RNA-seq analysis was carried out. In the immediate early-phase of HIV transduction, both viruses were influenced the expression profile thousands of transcripts. 4551 transcripts were regulated differentially by HIV-1, and 3422 by HIV-2. We detected significant differences between the top 10 most up and downregulated transcripts between HIV-1 and HIV-2 transduced cells. Compared to HIV-2, HIV-1 regulated transcripts were much more diverse.

2 hours post transduction HIV-1 managed to alter the expression of transcripts involved in the cellular transport (*SCL711A*), DNA repair (*TEX15*, *BRCA2*) and chaperones (*SACS*). Interestingly most of the transcripts affected by HIV-2 transduction were related to the intra and extracellular matrix. Protein-coding transcripts such as *COL1A2*, *COL6A3*, *COL3A1*, *KRT14*, *KRT7* and *KRT5* were upregulated by HIV-2 at 2 hours post-transdction.

In order to gather more information about the affected transcriptome 2 hours post transduction, GO analysis was carried out. Both viruses induced changes in the expression of genes related to GTPase regulator activity, GTPase activator activity, protein serine/threonine kinase activity, protein processing, structural constituent of ribosome and nucleoside-triphosphatase regulator activity. We observed changes in transcripts coding for proteins with iron-sulfur cluster binding, metal cluster binding, tau protein binding, rRNA binding ability and tau-protein kinase activity in case of HIV-1 transduced cells.

Moreover, protein-coding transcripts from HIV-2 transduced cells were associated with RNA polymerase II-specific DNA binding transcription factor binding, DNA-binding transcription factor binding, phosphatidylinositol binding, Rho GTPase binding ability and guanyl-nucleotide exchange factor activity.

To further outline the changes in the cellular transcriptome caused by HIV-based vectors, we conducted examination of later time-points, at 8, 12 and 26 hours to be exact. Following transduction with HIV-1, HEK-293T cells showed a dynamic fluctuation in the count of DETs, transitioning from 158 at 8 hours, to 114 at 12 hours, and eventually reaching 380 at 26 hours post-transduction. In contrast, at the 8-hour time point, HIV-2 exhibited a differential regulation

of 283 transcripts, a significantly higher number compared to HIV-1 at the same time point. By 12 hours, HIV-2 induced alterations in the expression of 299 transcripts, which decreased to 182 at 26 hours. The increased number of transcripts influenced by HIV-2 at the initial time-points suggests a more varied impact on the transduced cell, followed by a plateau at 26 hours, in contrast to HIV-1, which consistently altered the transcriptome of the transduced cell.

HIV-1 influenced transcripts at 8 hours post transduction included the AKAP9 a protein involved in the cAMP mediated signaling (Tröger et al., 2012). Another protein coding transcritps which was altered by HIV-1 were the GCC2 which is plays a part in the Nef mediated downregulation of MHCI (Kumari et al., 2019). HIV-1, at 12 hours post transduction also increased the cellular level of COX20, a cytochrome C oxidase member and an integral contributor to the mitochondrial respiratory chain complex IV (Li et al., 2022). During HIV-1 uncoating the p2 peptide is released which activates the cytochrome C oxidase thus increasing ATP production which is required for efficient reverse transcription and nuclear import of the PIC (Ogawa et al., 2015). 26 hours after transduction, HIV-1 induced the upregulation of protein coding transcripts that are associated with the oxidative stress response. HIV-1 is able to increase the production of reactive oxygen species which leads to the disruption of the oxidative stress pathways and the induction of mitochondrial dysfunction (Ivanov et al., 2016). The cellular level of transcripts encoding for the NQO1, HMOX1 and OSGIN1 were increased by HIV-1, which indicates an active defence against the generated oxidative stress.

Transduction with HIV-2 significantly increased the cellular level of a transcripts encoding for the DCN, a protein involved in cell proliferation and intracellular communication in all of our observed time-points. Increased expression of DCN can initiate apoptosis via caspase-3 and initiate cell cycle arrest at the G0-G1 phase through the upregulation of p21 (Wu et al., 2008) (Ständer et al., 1999).

Just as observed at earlier time-points, transduction with HIV-2 increased the expression of genes coding for components of the intra- and extracellular matrix, including KRT4, 5, 8, COL1A2 and COL6A3. At the same time, HIV-1 only managed to affect the cellular level of collagens; such as COL11A1.

Both HIVs managed to influence the cellular level of *EGR1*, a transcription factor involved in the Tat-dependent HIV gene expression. EGR1 plays a prominent part in the therapeutic strategy called "kick and kill" wherein cells which are latently infected are subjected to latency reversal agents to reactivate dormant HIV reservoirs. Following treatment with various agents, EGR1 was shown to be upregulated, and it has been established that EGR1 directly interacts

with the HIV-1 promoter, thereby inducing proviral transcription (Woodson & Kehn-Hall, 2022) (Wong et al., 2022).

Our data showed the differential regulation of pseudogenes in all of the observed time-points by both HIVs which is indeed a notable find. Emerging evidence suggests that self-derived mRNAs, including pseudogenes, have a crucial part in regulating immune responses against viruses and tumours (Han et al., 2021)

Cells transduced with HIV-1 showed miniscule changes in the number of pseudogenes compared to transduced controls, while HIV-2 transduced cells displayed a significantly higher number of differentially regulated pseudogenes compared to HIV-1 and control. At 8 hours post-transduction in HIV-2 transduced cells, 77 pseudogenes showed differential regulation, this number increased to 93 at 12 hours and decreased to 78 pseudogenes at 26 hours. As pseudogenes traditionally known as negative regulators of gene expression, the differential regulation we observed in the expression of several pseudogenes might contribute to the unique effect HIV-2 has on the host cell.

Similar to earlier time points, we employed GO analysis to elucidate differences between cells transduced with HIV-1 and HIV-2. At 8 hours, HIV-1 had a pronounced influence on cellular metabolism, as shown by the significant GO terms such as positive regulation of RNA metabolic processes, regulation of primary metabolic processes and positive regulation of metabolic processes. In contrast, HIV-2 primarily influenced transcripts associated with keratinocyte differentiation, extracellular matrix organization or cell adhesion. Increased metabolic activity just as the one induced by HIV-1 increased viral replication and infectivity, as it is known that the inhibition of glycolysis leads to a concurrent inhibition of reverse transcription (Taylor & Palmer, 2020) (Kang & Tang, 2020).

Surprisingly, no significant GO terms were detected at the 12-hour time point from cells transduced with HIV-1. We detected the lowest number of altered transcripts at 12 hours compared to other time points from HIV-1-infected HEK-293T cells. Meanwhile, at the 12-hour mark, HIV-2 affected animal organ development, system development and cell differentiation related genes. At the 26 hour time point, HIV-1 differentially regulated transcripts which are associated with the response to detoxification, cellular response to toxic substances, and oxidative stress. In contrast, HIV-2 showed less impactful influence on the cellular environment, affecting genes primarily related to the skin development, negative regulation of cell migration, and negative regulation of cell motility.

To further elucidate the changes detected from the transcriptomic analysis, proteomic investigation was also carried out. It was revealed by our analysis that in the first two hours of

transduction, both virus managed to alter the expression of hnRNPA1, NONO, H1-4, HSPD1 and SRSF6. Moreover, HIV-2 also downregulated the expression of FKBP4 and CCT8.

hnRNPA1 belongs to a complex family of ribonucleoproteins and had an important role in the transcription, stability, and transport of newly synthesized cellular mRNAs. It has been shown that through binding to specific sites on the newly synthesized transcript, it can function as a splice inhibitor of HIV-1 *tat* mRNA upon expression from the integrated proviral genome (Tange et al., 2001). Tat is known for its highly apoptotic nature, so this inhibition proves advantageous for the virus through the prevention of overexpression of Tat. Furthermore, through Rev stimulation, hnRNPA1 facilitates the transport of viral mRNA to the cytoplasm from the nucleus (Hallay et al., 2006).

Nuclear protein NONO has a role in transcriptional regulation and RNA splicing. It was shown to associate with reverse transcription and PIC during HIV infection (Schweitzer et al., 2013) (St Gelais et al., 2015). Overexpression of NONO leads to a decrease in the infectivity of HIV-1, which negatively impacts reverse transcription and viral gene expression in Jurkat cells. It has been postulated that the downregulation of reverse transcription is achieved either through direct contact with the reverse transcriptase complex or through indirect interaction with other components of the complex. However, the exact mechanism remains unclear (St Gelais et al., 2015).

Using STRING and GO terms, an intricate network of cellular processes in which the differentially regulated proteins are involved was visualized. The GO terms included cellular response to stress, protein folding, mRNA stability and mRNA splicing. It is noteworthy to mention that the initial low level of proteins may be overridden at later time-points later as the infection progresses.

Similarly, to the transcriptomic analysis, we conducted analysis of proteomic data from later time-points as well. It was revealed from our data that a significant portion of the altered proteins are connected to mRNA processing, proteasome and chaperon function. We did not detect any significantly regulated protein at 8 hours post-transduction, however, our analysis showed that both HIVs influenced the expression of 17 and 117 different proteins at 12 and 26 post-transduction, respectively.

HSPA1A, hnRNPK, PA2G4 and Zyx were altered at 12 hours post-transduction by both HIVs. In addition to actin polymerization and focal adhesion, Zyx also takes part in the intracellular signalling (Rauskolb et al., 2011). Besides its involvement in cell growth, apoptosis and differentiation, PA2G4 is also downregulated by HIV-1 Vpr, resulting in G2 arrest and apoptosis in U87MG cells (Zhang et al., 2014).

Only HIV-1 downregulated the expression of hnRNPQ, a member of a protein family involved in mRNA transcription, splicing, export, stability and translation. Through interaction with Rev protein, hnRNPQ also takes part in the proviral transcription of HIV-1. It is involved in the mRNA stability, splicing and RNA replication of Hepatitis C virus (Vincendeau et al., 2013).
Moreover, HIV-2 was able to decrease the cellular level of RANBP1 compared to HIV-1. RANBP1 possesses a nuclear export signal (NES) domain which is analogous in both function and structure to the one present in HIV Rev, this may indicate that RANBP1 might share a step with Rev in the post-transcriptional pathway (Audia et al., 2023). The downregulation of RANBP1 by HIV-2 prompts inquiries about the proteins role in hindering the transport of proviral transcripts.
Similarly, to the very early time point, GO term analysis was carried out from the differentially regulated proteins to further outline the differences between the two pseudovirions. In case of both HIVs, only a low number of GO terms were affected at 12 hours post-transduction. However, at the 26-hour time-point, both virus altered the expression of proteins related to post-transcriptional regulation of gene expression, regulation of DNA-templated transcription elongation and regulation of alternative mRNA splicing via the spliceosome. Moreover, differences were observed between the two viruses at the 26-hour time point. In case of HIV-1, GO terms like positive regulation of apoptotic processes, regulation of gene expression and cellular response to unfolded protein were present. Meanwhile, HIV-2 altered the expression of proteins related to positive regulation of translation, negative regulation of protein ubiquitination and negative regulation of DNA metabolic processes.
In order to gather more information about the effect of HIV-based pseudovirions on the host cells, we also carried out transcriptomic analysis on Jurkat T cells
We noted that similarly to HEK-293T cells, HIV-1 exerted a larger effect on the cells, represented by the higher number of DETs compared to HIV-2. We detected 39 at eight, 45 at 12, and 42 transcripts at the 26 hours time-point from HIV-1 transduced Jurkat. In contrast, HIV-2 only altered the expression of 15 at eight, 23 at 12 and 36 transcripts at 26 hours post transduction.

HIV-1 differentially regulated the expression of the insulin induced gene 1 (*INSIG1*). Research showed that INSIG1 is upregulated during HIV-1 infection and it acts as a sentinel responding to HIV replication and inhibits via the acceleration of HIV-1 gag protein degradation (Zhang et al., 2019). As we detected the downregulation of *INSIG1* at 8 hours post-transduction, it is

possible that during the early phase of infection, it has a yet undiscovered role in the viral life cycle.

Our data showed that both HIV-1 and HIV-2 are able to downregulate the expression of the aldolase, fructose bisphosphate C (*ALDOC*) 26 hours post transduction compared to control transduced cells. Downregulation of the metabolic pathways by HIV is a sign of HIV latency, and during viral reactivation, the negative effect on metabolism is reverted as elevated glycolysis is required for efficient viral replication (Shytaj et al., 2021) (Barrero et al., 2013).

An interesting finding was the differential regulation of transcripts, products of which are related to the mevalonate pathway. Transcripts such as, 3-hydroxy-3-methylglutaryl-CoA reductase (*HMGCR1*), *HMGCS1* and mevalonate diphosphate decarboxylase (*MVD*) were altered in almost all of the observed time point by both viruses. HIV-1 further altered the expression of mevalonate kinase (*MVK*) at 12 hours post-transduction. The mevalonate pathway contributes greatly to immunity, so we hypothesize that the observed reduction in the expression of mevalonate synthesis genes contributes to the HIV generated immune dysfunction (Mu et al., 2024). Inhibitors against the HMG-CoA reductase enzyme called statins are known to be decrease HIV-related co-morbidities such as HIV-associated cardiovascular disease (CVD) and non-Hodgkin lymphoma (Grinspoon et al., 2019) (Grinspoon et al., 2023) (Chao et al., 2011).

Similarly to the transduced HEK-293T cells, we carried out GO analysis to further reveal the affected biological processes. Both HIVs significantly altered genes that are part of the cholesterol biosynthetic process.

There were very few transcripts that were regulated similarly in both HEK293T and Jurkat cells by HIV-1 and HIV-2. Most of the detected transcripts were unique to the given cell line. One exception is the arginosuccinate synthase 1 (*ASS1*) which was downregulated by both HIV-1 and HIV-2 at 8 hours in HEK293T cells and also by HIV-1 at 8 hours in Jurkat cells. The product of this transcript is involved in the de novo synthesis of arginine, and also described in the context of Herpes simplex virus 1 (HSV1) infection as the knockdown of ASS1 enhances viral replication (Grady et al., 2013). However, there is no information about the involvement of ASS1 in the life cycle of HIV. Additionally, the previously described CDKN1A was also differentially regulated at 26 hours post transduction by HIV-1 in both HEK293T and Jurkat cells. Moreover, HIV-2 at 26 hours post transduction significantly downregulated the expression of the earlier discussed *EGR1* and *HMGCS1* in both HEK239T and Jurkat cells.

In conclusion, the analysis of very early and early time points of HIV-1 and HIV-2 transduction revealed significant differences between the two HIV based pseudovirions. These observations may contribute to a better understanding of the pathomechanisms associated with lentiviral vector transduction. Additionally, our results offer potential insights into the distinctive features of HIV-2's replication cycle, an area that needs future investigations. For future research involving other cell lines such as monocyte based ones might be a good addition to the knowledge pool about HIV generated proteo-transcriptomic changes. Moreover, a study involving primary cells such as $CD4^+$ T cells might provide a further information about the early phase of lentiviral transduction.

Summary

Our goal was to analyse the effect of HIV-1 and HIV-2 pseudovirions on the cellular transcriptome and proteome focusing on the early phase of transduction. Moreover, we wanted to shed light on the unique characteristics of HIV-2 as studies focusing on this virus are lacking. In the very early phase of HIV transduction, significant changes are detected between HIV-1 and HIV-2 transduced HEK-293T cells. HIV-1 managed to alter the expression of a diverse array of transcripts including DNA repair proteins, transporters and chaperons. Meanwhile at the same time point changes generated by HIV-2 were less diverse and mainly affected transcripts coding for members of the extra and intracellular matrix. At later time points this pattern was also observed as HIV-1 altered the cellular transcriptome steadily compared to HIV-2 which at the early time points had a diverse effect on the transcripts but reached a plateau at 26 hours post transduction. Product of transcripts altered by HIV-1 function as signalling, transport and oxidative stress involved molecules. While HIV-2 influenced the expression of transcripts with role in the lipid homeostasis. Just at with the very early time points, HIV-2 had a significant effect on the transcripts coding for proteins involved in the extra and intracellular matrix including keratins and collagens. Utilizing the altered protein coding transcripts GO analysis was carried out which revealed multiple common pathways influenced by both HIV strains. Moreover, significant differences were also observed between the two virus such as the different regulation of oxidative stress related genes by HIV-1 at 26 hours post transduction.

A notable find was the different alteration of the cellular level of pseudogenes by both HIV strains. Compared to HIV-1 HIV-2 showed higher capabilities to alter the cellular level of pseudogenic RNA. As pseudogenes are described as negative regulators of gene expression,

the effect HIV-2 had on the expression of pseudogenes might give us an answer to its unique pathophysiology.

To further outline the differences caused by the HIV transduction proteomic analysis was carried out alongside the transcriptomic investigation. Analysis of the very early phase of HIV transduction revealed five proteins that were differentially regulated by both HIV strains and two more which are influenced only by HIV-2. Amongst the altered proteins were molecules with role in mRNA splicing and protein folding. We did not detected any significantly altered protein at eight hours post transduction. However, at the 12 hour time point 17 different proteins were altered by HIV with roles in intracellular signalling, apoptosis and mRNA splicing. At 26 hours post transduction a total of 117 protein were differentially regulated by HIV-1 and HIV-2. The functions of the altered proteins were diverse with roles in mRNA splicing, cell proliferation and protein folding.

GO analysis of the differentially regulated proteins revealed further similarities and significant differences between the two HIV strains. Both virus managed to alter proteins with roles in post translational regulation of gene expression and regulation of alternative mRNA splicing via the spliceosome. Meanwhile only HIV-1 influenced the expression of proteins taking part in cellular response to unfolded protein while HIV-2 influenced the expression of negative regulation of protein ubiquitination involved proteins.

Analysis of transduced Jurkat cells revealed further differences between the two HIVs. We detected that HIV-1 altered a larger and more diverse array a of transcripts compared to HIV-2. Both virus altered the expression of transcripts involved in the cholesterol biosynthetic processes with affected genes such as HMGCR and MVD.

In conclusion, we hope that our study help researcher better understand the complex changes generated by the infection with HIV-1 and HIV-2 as well as transduction with lentivirus-based vectors. We also aimed to shed light on the unique characteristics of HIV-2.

List of publications prepared by the Kenézy Life Science Library

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/10/2025.PL
Subject: PhD Publication List

Candidate: Tamás Richárd Linkner
Doctoral School: Doctoral School of Molecular Cellular and Immune Biology

List of publications related to the dissertation

1. **Linkner, T. R.**, Ambrus, V. A., Kunkli, B., Szojka, Z., Kalló, G., Csősz, É., Kumar, A., Emri, M., Tőzsér, J., Mahdi, M.: Comparative Analysis of Differential Cellular Transcriptome and Proteome Regulation by HIV-1 and HIV-2 Pseudovirions in the Early Phase of Infection.
   Int. J. Mol. Sci. 25 (1), 1-26, 2024.
   DOI: http://dx.doi.org/10.3390/ijms25010380
   IF: 4.9 (2023)
2. **Linkner, T. R.**, Ambrus, V. A., Kunkli, B., Szojka, Z., Kalló, G., Csősz, É., Kumar, A., Emri, M., Tőzsér, J., Mahdi, M.: Cellular Proteo-Transcriptomic Changes in the Immediate Early-Phase of Lentiviral Transduction.
   Microorganisms. 9 (11), 1-21, 2021.
   DOI: http://dx.doi.org/10.3390/microorganisms9112207
   IF: 4.926

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

List of other publications

3. Miltner, N., **Linkner, T. R.**, Ambrus, V. A., Al-Muffti, A. S., Ahmad, H., Mótyán, J. A., Benkő, S., Tőzsér, J., Mahdi, M.: Early suppression of antiviral host response and protocadherins by SARS-CoV-2 Spike protein in THP-1-derived macrophage-like cells.
Front. Immunol. 13, 999233, 2022.
DOI: http://dx.doi.org/10.3389/fimmu.2022.999233
IF: 7.3

Total IF of journals (all publications): 17,126
Total IF of journals (publications related to the dissertation): 9,826

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

16 January, 2025

Acknowledgments

First and foremost, I would like to thank to my supervisor Dr. Mohamed Mahdi for his guidance, immense patience, help and support during my years as a Phd student.

Also, I would like thank to Prof. József Tőzsér for accepting me into his workgroup and providing me the opportunity to work in the field of science.

I'm grateful to Dr. Zsófia Szojka for teaching me how to work in a molecular biology laboratory and how to handle viruses.

I would like to thank the workers of the Retroviral Biochemistry laboratory; Dr. Mária Golda, Dr. András Szabó, Dr. János András Mótyán, Noémi Miltner, Almuffti Aya Shamal Abdullah, Kiarie Irene Wanjiru, Dr. Vanda Toldi and Dr. Márió Miczi for their help and fun conversations that we had in the lab.

I wanna give a special thank you to Kiarie Irene Wanjiru for her help in the production of the transduced Jurkat samples.

I would like to thank Viktor Attila Ambrus and Balázs Kunkli for their immense knowledge in the field of bioinformatics and their work in the analysis of the transcriptomic and proteomic data respectively.

I thank our great laboratory assistant Szilvia Janics-Pető for her help and assistance in laboratory work.

I'm greatfull to Dr. Gergő Kalló, Dr. Ajneesh Kumar, Dr. Miklós Emri and the workers and collegues of the Proteomics Core facility for their help in the mass specometry measurments and analysis of the proteomic data.

I'm thankfull to Dr. Szilárd Póliska and collegues of the Genomic Medicine and Bioinformatics Core Facility for their help in the sequencing of the transcriptomic data and data analysis.

Lastly I want to thank my parents and sister for their continuous support during my Phd years and uplifting me during harder months.