

**Detection of DNA intermediates of the reverse transcription during the infection process of the human endogenous retrovirus-K (HML-2)**

**At the  
Robert Koch Institute  
Berlin**

Under the supervision of PD Dr. Norbert Bannert

Alexander Volkwein

(Student number 4614012)

## Bachelor's thesis statement of originality

I hereby confirm that I have written the accompanying thesis by myself, without contributions from any sources than those cited in the text and acknowledgements. This applies also to all graphics, drawings and images included in the thesis.

.....

Place and date

.....

Signature

## Table of contents

1. Abstract .....	- 5 -
2. Introduction.....	- 6 -
2.1 Viruses .....	- 6 -
2.2 Retroviruses.....	- 7 -
2.3 Endogenous retroviruses .....	- 9 -
2.4 Cell culture systems and viral particles production .....	- 10 -
2.5 Reverse transcription and pre-integration complexes .....	- 12 -
2.6 Objectives .....	- 15 -
3. Materials and methods .....	- 16 -
3.1 Production of pseudotyped HERV-K (HML-2) particles.....	- 16 -
3.1.1 Plasmid DNA.....	- 16 -
3.1.2 Chemical transformation of E. coli.....	- 17 -
3.1.3 Isolation of recombinant DNA (rDNA) and quantification .....	- 18 -
3.1.4 Cell culture.....	- 18 -
3.1.5 Calcium phosphate transfection .....	- 19 -
3.1.6 Harvest, sterile filtration and storage of viral particles .....	- 22 -
3.1.7 Quantification of viral particles - p27 ELISA.....	- 22 -
3.2 Transduction of distinct mammalian cell lines.....	- 23 -
3.2.1 Cell lines, transduction and DNA isolation.....	- 23 -

3.2.2 High-titer viral preparation by ultracentrifugation .....	- 25 -
3.2.3 Complementary DNA PCR amplification and agarose gel analysis .....	- 25 -
4. Results .....	- 29 -
4.1 Quantification of transfected viral particles .....	- 29 -
4.2 Electrophoretic analysis in agarose gel .....	- 29 -
4.2.1 – 1 <sup>st</sup> Transduction .....	- 30 -
4.2.2 – 2 <sup>nd</sup> Transduction .....	- 31 -
4.2.3 – 3 <sup>rd</sup> Transduction .....	- 32 -
5. Discussion .....	- 32 -
6. References .....	- 36 -

## 1. Abstract

Endogenous retroviruses form part of the genome of many vertebrates. In humans they account for approximately 8 % of the entire genetic information. Along evolution the proviral sequences have been mutated until nowadays there is no known HERV able to express infectious particles. However the youngest and probably therefore the less mutated family, HERV-K (HML-2), preserves proviruses able to produce functional proteins. Although it has been shown that not only the glycoproteins responsible for mediating entry have been mutated, but also integration of the viral genes in the host genome will not be properly accomplished.

In order to determine potential inhibited stages of this process, HERV-K pseudotyped viral particles able to mediate penetration of the cell membrane were successfully produced and monitored during reverse transcription on their course towards integration.

For this purpose several set of primers were designed, that bind specific to different stages along the reverse transcription process. Using these primers reverse transcribed DNA could be detected. Consequently, analysis of these characteristics steps could ideally help to detect at some point of significant less cDNA, which may permit to infer the phase where inhibition takes place.

## 2. Introduction

### 2.1 Viruses

Although viruses are present in all three domains of life, namely bacteria, archaea and eukaryote, they may not account for any of these branches [1]. This contemplation is based on the fact that viruses do not fulfil all of the attributes that conventionally characterize life. However this opinion is not unanimous among the scientific community. Furthermore their origins are also a controversial aspect that arouses many question and in spite of the existence of several theories that propose different explanations in this regard a consensus has not yet been found. Nevertheless, since they are present in all three domains of life it could be inferred that they emerged before diversification of these three branches and therefore could be considered as old as life [2].

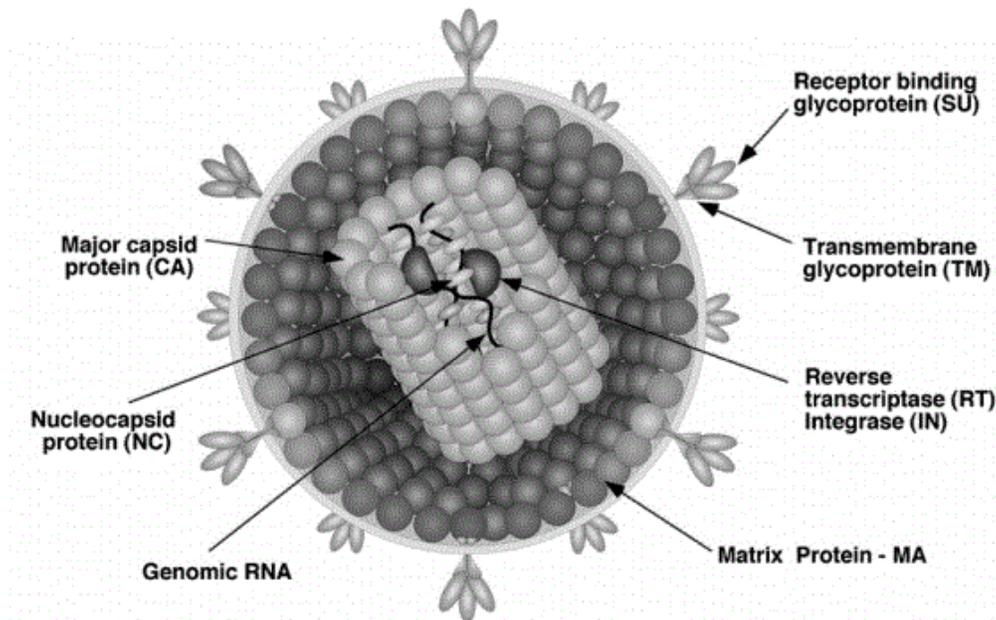
The first scientific evidences about viruses are from the ends of the 19<sup>th</sup> century, when Louis Pasteur, Dimitri Ivanovski, Martinus Beijerinck and Friedrich Loeffler among other researchers, could demonstrate the presence of pathogenic agents, smaller than bacteria and able to proliferate in living tissues [3].

Viruses are obligate intracellular parasites that do not possess own metabolism outside of a living and permissive host cell. The term *permissive* implies that the host cell does not only expresses the receptor that the virus needs to attach itself to, but also it presents intracellular conditions that permit the virus to accomplish effective replication cycle. In other words, viruses rely on infection of living cells in order to persist in time and of course a milestone in the whole viral cycle is the replication of its genetic information. This genetic information can be present in different forms, structures or polymers. Viral particles contain either a DNA or a RNA genome, which is carried by a protein structure termed capsid. Depending on the type of virus this capsid might be enclosed by a lipid membrane or not. [3]. In fact viruses are made of the same building blocks that life is also made of.

## 2.2 Retroviruses

Members of this group were during the first decade of the 20<sup>th</sup> century described for the very first time. In 1908, Vilhelm Ellermann und Oluf Bang shown that filtered fluid of infected tissues, containing no cells, was capable to produce new infections of avian leucosis in healthy cells. A few years after Peyton Rous proved in a similar manner the transmission of the avian sarcoma virus, afterwards named Rous sarcoma virus (RSV) after its discoverer. However they were neither recognized as retroviruses nor gathered into a group until 1970, when their most prominent characteristic was discovered by Temin and Baltimore, namely the property of an enzyme that enables them to reverse the flow of genetic information [4].

Furthermore the *Retroviridae* family present a mutual set of characteristics that led to the formation of their cluster. The virions are enveloped, in a size range between 80 and 100 nm, the genome is a dimer of single stranded positive RNA molecules, and each strand is between 7 and 13 kb long. Their genome also presents a characteristic structure in which the functional and structural genes Gag (group-specific antigen), Pol (polymerase, integrase, protease and reverse transcriptase) and Env (envelope glycoproteins) are flanked by two regulatory sequences termed long terminal repeats (LTRs) [3]



**Figure 1 presents the characteristic morphology of a virion.** [Image source: King, O.E.A., *Virus Taxonomy ninth report of the ICTV. 2011*]

This family can be subdivided in the *Orthoretrovirinae* and *Spumaretrovirinae* subfamilies. The last one differ from the rest, because they bear a double stranded DNA genome, as reverse transcription takes place after the particles leave the host cell and contains only one genus. However it undergoes reverse transcription late during the replication cycle. The *Orthoretrovirinae* comprises the genera *Alpha-*, *Beta-*, *Gamma-*, *Delta-*, *Epsilonretrovirus* and *Lentivirus* [5].

The peculiar ability of these viruses to convert RNA into DNA contradicted what was considered to be the conventional direction of the genetic information flow. Therefore this distinctive process was named reverse transcription, term that also shaped the name of the family. The enzyme that catalyses this reaction is called reverse transcriptase (RT). This enzyme is encoded by the virus and packed in every virion during assembly, so that after infection of a new cell, the RT molecules can transcribe the viral RNA genome in linear complementary DNA (cDNA). Moreover the RT also possess RNase H activity, feature that allows it to degrade RNA, in this case, the viral genome after it has been copied into DNA. Subsequently the double stranded viral DNA molecules can bond to cellular proteins and form-a pre-integration complex (PIC). The integrase (IN) is another indispensable viral enzyme that will be attached to the PIC and participate in several steps. For example in the reverse transcription itself, the nuclear import of the PIC, chromatin targeting and integration [6].

Retroviruses also differ among them within the subfamily *Orthoretrovirinae* in different aspects, for example regarding the selection of the site for chromosomal integration. However their replication cycle invariably requires integration of the viral DNA sequence in the host genome in order to have eventually the chance to be expressed by cellular factors [7]. The step of successful integration constitutes also the formation of a provirus [8].

## 2.3 Endogenous retroviruses

Even closer related to the main subject of this work is another betaretrovirus denominated mouse mammary tumour virus (MMTV), which was proved in 1936 by John Bittner to also possess tumorigenic properties [9]. Besides the more obvious way of proliferation by infection of further permissive cells, that the virions might be able to reach after being released from the cell where it was replicated, the MMTV exhibited also vertical gene transmission, which means that the viral genes will be passed on from the parents to the next generation in the own genome of the host organism, after a retrovirus has become a provirus within a germ line cell [10]. Unlike proviral sequences in the chromosomes of a somatic cell that will perish together with their host, endogenous retroviruses will be inherited to further generations and yet be present in every single cell of their host organisms. Although this feature per se does not necessarily entail expression of the provirus.

The access to numerous whole genome sequences in the post-genomic era has revealed to which extends ERVs have conquer the genomes of diverse organisms, including the human. It might appear surprising that the human genome comprise approximately 8 % of ERVs sequences and less than 2 % of protein coding regions. In others words 4-fold more ERVs sequences than protein coding sequences [11]. However all the known human proviruses are defective, which means that are not able to express infectious particles anymore. The reason for this are mainly lost-of-function mutations that have at least partially deteriorate the products of these genes [12].

This exclusion of proviral expression in human cells seems to be the result of adaptation after long periods of coexistence, in which the host has been able to turn the virus non-infectious. The integration of foreign sequences in the chromosomal genome of an organism can provide favourable conditions for molecular evolution, especially regarding two aspects. The first one can be the proviral sequence itself, which may constitute the substrate for the origin of novel sequences. The second issue can be the modification of the existing host sequence after the integration of the viral DNA. Both factors might lead to outcomes of three possible nature for the host, viz. neutral, detrimental or beneficial. Clearly in this regard time will be the factor with a

determinant influence, then the fate of the novel and/or the alteration of the existent sequences will be exposed to selective pressures.

Consequently the most active ERV in the human lineage correspond to the remnants of the youngest integrated group of proviral elements, clustered as human endogenous retrovirus K human mouse mammary tumour virus like -2 (HML-2). Several proviral genes of this group are still capable of expressing functional proteins and even complete viral particles. However and despite the fact that it has been shown that consensus sequences could result in infectious particles [13], these will still not be able to undergo efficient replication, due to the innate immune response of the host. These intracellular factors like APOBEC and tetherin proteins have probably coevolved with the virus within the last 30 to 40 million of years and are now capable of preventing further infection [14].

Remarkably the mutations that have inactivated the expression of infectious particles have been successfully reverted or back-mutated by several groups, what lead to original expression of all reading frames and consequently the reconstitution of infectious particles [15].

Moreover evidence shows that some attributions of the viral genes might have been co-opted, modified and incorporated by the host, to its own functions and for its own benefit. An example could be the evolution of the *FV1* gene in mice, which confers resistance against a related gammaretrovirus, the murine leukaemia virus (MLV) [16].

## 2.4 Cell culture systems and viral particles production

In order to study different characteristics of viruses, it is indispensable to establish adequate and efficient cell cultures as replication systems. The produced viral particles, do not correspond to the naturally occurring virus. They are manipulated in order to confer them previously specified features and so make possible specific posterior analysis.

The production of viral particles was performed by means of transient transfection [17]. In this procedure different DNA sequences are putted together, in a specific concentration, in presence

of eukaryotic cells. These DNA sequences in form of circular bacterial plasmids, encode for different viral proteins, in other words for different elements of the subsequent mature virions. This particular step of the process allows to some extents the possibility of exchange some of the constituents of the viruses. Therefore also the corresponding characteristics of the resulting particles are altered. This type of modification it is known as pseudotyping [18].

Since this manipulation concerns the exchange of envelop proteins, which characterize and mediate the entry into the host cell, it will be obviously performed on enveloped viruses. Interestingly, the foreign envelop proteins will only become part of the virions produced by the transfected cells, because only during the transfection is the envelop vector present. This vector carries the sequence of the foreign proteins. However the produced virions do not possess the necessary genes for expressing the foreign proteins and therefore consecutive replicas of the viruses assembled by the transfected cells and therefore eventual copies of these virions, i.e. after replication cycle, will correspond to the genotype that they contain and not to their original phenotype [19].

One of the transfection procedures consists in mixing plasmid DNA with calcium and phosphate buffer, in order to create a precipitation, which also binds the plasmid DNA and settles down. The adherent cells at the bottom incorporate these complex and eventually perform translation with it. Although the reason for the uptake of external DNA by the producing cells still remains unknown, the efficiency of this method is well established [19].

Thus, two different kinds of the reconstituted HERV-K pseudotyped virions were produced. The first one carries a, due to a deletion in its sequence, defective envelop glycoprotein of HIV-1. This mutant is called  $\Delta$ KS and is not able to provide transport for the viral genes through the cell membrane. The second pseudotype was produced with the envelope protein of the vesicular stomatitis virus (VSV-G), which permits efficient entry in a high number of different cells [20]. This ensures competent transfer of the viral capsid across the cellular membrane in the cytoplasm. As a result of the distinctive capacities of the two pseudotypes a meaningful and reliable effect can be tested and compared.

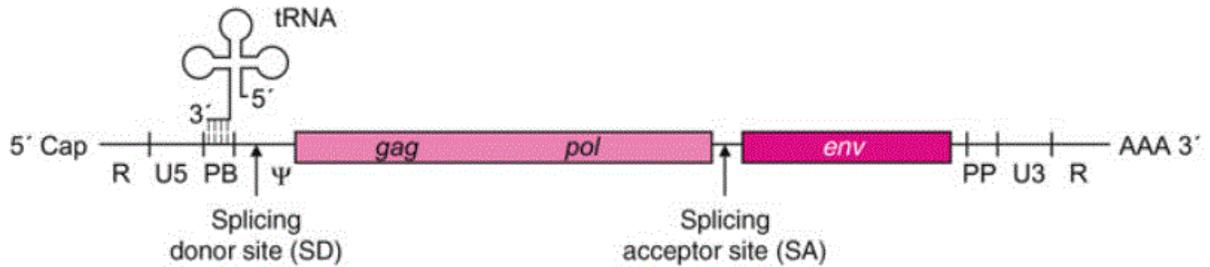
The productivity of the transfected cells that express the viral genes it is subjected to diverse factors [17]. For example the period of time between transfection and harvest of particles. Production can be also specifically incremented, for instance through addition of sodium butyrate  $-\text{Na}(\text{C}_3\text{H}_7\text{COO})-$ . Regarding the time it is relevant to consider that in transient transfection, as the name suggests, the foreign genes do not integrate in the genome of the host and therefore they are present in the nucleus extra chromosomal, in other words as an episome. These episomes can be transcribed, replicated and even passed on to the next generation of cells after cell division. However, all these events are conducted by cellular factors, hence some of which might also act on the foreign DNA, degrading it [3]. As a consequence the amount of transfected DNA and its corresponding expression will decrease as mitosis takes place.

The quantity of the virions produced may also be influenced by sodium butyrate [21]. This compound has several effects on mammalian cells but the one observed here is the increase of protein expression in order to reach a higher expression of the transfected genes and consequently a larger viral yield. This is accomplished by manipulating the conformation in which the DNA is present, for example by histone hyper acetylation. In this case sodium butyrate provokes the inhibition of the enzymes that are responsible for histone deacetylation, named histone deacetylases (HDACs), inducing so the mentioned histone hyper acetylation [22]. This change in the arrangement of the DNA structure allows higher transcriptional activity by unwinding the chromatin around the nucleosomes and therefore exposing larger segments of DNA to the DNA-dependent RNA polymerase.

## 2.5 Reverse transcription and pre-integration complexes

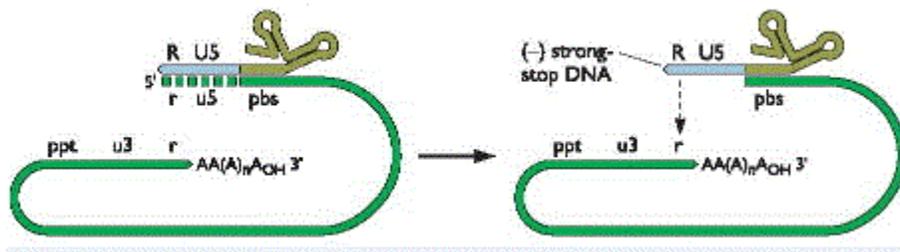
The genome of all retroviruses presents a similar structure. Further the 5' terminus is provided with a 5' cap structure and the 3' end polyadenylated, in others words, these aspects resemble eukaryotic mRNAs. The two regulatory LTRs in the viral RNA genome differ from the LTRs in the proviral sequences. Indeed some of the most peculiar stages of the reverse transcription may aim the modification of these LTRs. Furthermore this process involves viral and cellular factors. In fact the designation K in the denomination HERV-K (HML-2), stands for the amino acid lysine

(abbreviated Lys or K) and refers to the tRNA<sub>Lys</sub> molecule from cellular origin, that it is packaged in virions during assembly and attached to the 5'-end of each linear RNA molecule in the viral genome [23].



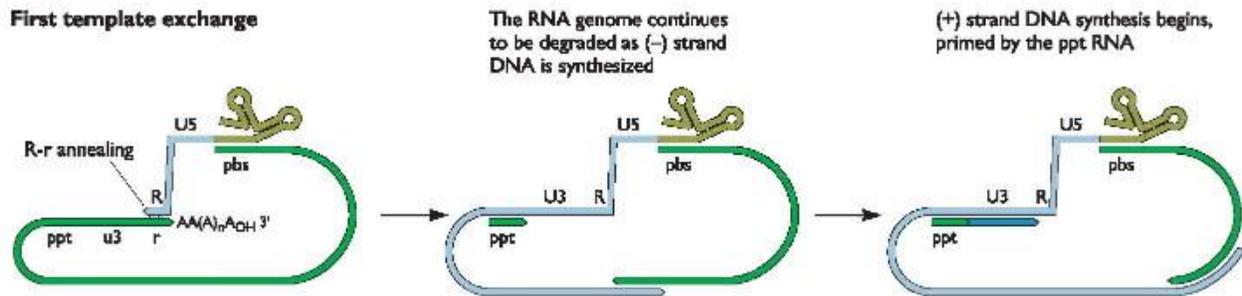
**Figure 2 shows a retroviral genome prior to reverse transcription, with a cellular tRNA annealed to the primer binding site (PB), which is situated between the U5 and Ψ-site, close to the 5' region and upstream of all genes.** [Image source: Susanne Modrow, D.F., Uwe Truyen and Hermann Schätzl, *Molecular Virology 3<sup>rd</sup> Edition 2013*]

This tRNA<sub>Lys</sub> will be used as a primer for the RT, which will start to synthesize nucleotides at the 3'-end of this molecule that is to say towards the U5 and redundant (R) regions. This first segment of cDNA constitutes the minus strong stop DNA and will be the first segment targeted for amplification. Consequently the newly synthesized R region will anneal to complementary R region of the 3' terminus, while the RNase H activity of the RT degrades the RNA template after being copied.



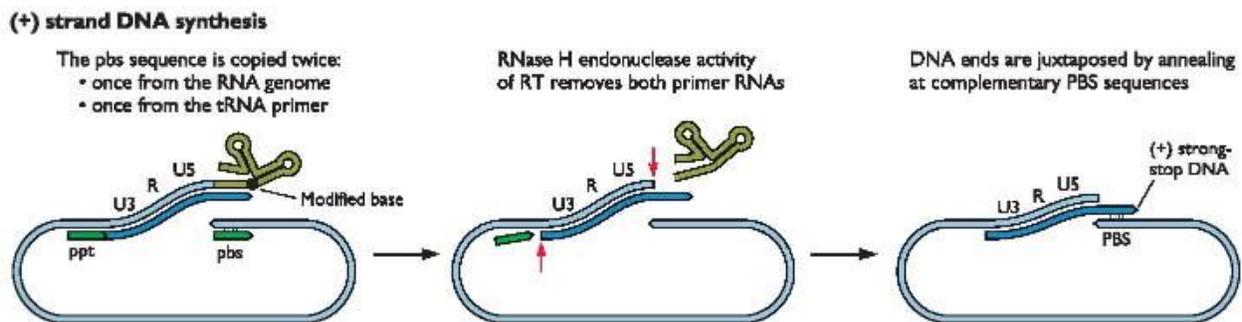
**Figure 3 illustrates the transfer of nascent cDNA minus strong stop (light blue) from the 5' to the 3' end of the viral genome and the simultaneously digested RNA template (segmented green line).** [Image source: S. Jane Flint, *American Society for Microbiology, Principles of virology, Washington, DC ASM Press, ©2009*]

Then, after the first template exchange (R-r) the RT will continue extending at the 3' end of the minus strong stop DNA and digesting the RNA template with one exception. Only the region termed polypurine tract (PPT) of the RNA template will be kept and used as a primer for the synthesis of the complementary +DNA.



**Figure 4 presents the DNA synthesis of functional and structural genes, i.e. elongation of minus strong stop DNA, further RNA template decomposition and polymerisation beginning of the plus strand DNA (darker blue) at the 3' terminus of the PPT.** [Image source: S. Jane Flint, American Society for Microbiology, Principles of virology, Washington, DC ASM Press, ©2009]

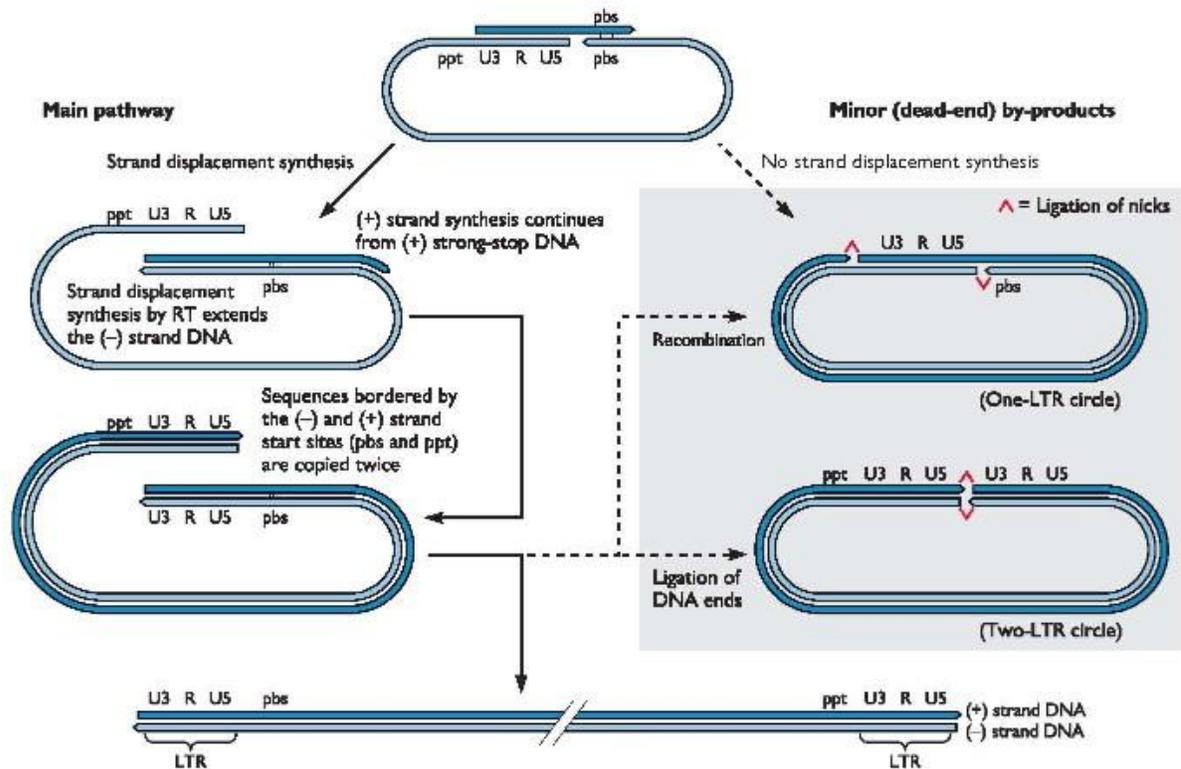
The plus strand DNA will be extended over the PBS and discard the tRNA primer by degrading the template. Now this overhang can anneal to its homologous sequence at the 3' end of the minus strong stop DNA and so continue polymerisation.



**Figure 5 represents the transfer of the plus strand DNA from one PBS to the opposite to enable further elongation.** [Image source: S. Jane Flint, American Society for Microbiology, Principles of virology, Washington, DC ASM Press, ©2009]

Finally the plus strand DNA will be completed and the RNA template entirely dissolved, as a result the now fully retro-transcribed double stranded DNA is prepared to attempt nucleus import and subsequent integration [24].

**Second template exchange is facilitated by annealing of PBS sequences**



*Figure 6 displays two possible outcomes beside each other. On the left the proper completed double stranded cDNA, which could attempt integration. On the right two results and two different paths through which the sequence will turn incapable of performing integration. At the bottom is the linear proviral form of the viral genome depicted with LTRs capable of eventually accomplish replication cycle.*

*[Image source: S. Jane Flint, American Society for Microbiology, Principles of virology, Washington, DC ASM Press, ©2009]*

## 2.6 Objectives

The objectives of this work are to produce infectious pseudotyped viral particles of the reconstituted betaretrovirus HERV-K (HML-2) and use them to transduce different cell lines in

order to detect and quantify selected steps of the pre-integration complexes along the reverse transcription process. Ideally this would provide necessary data to identify a state during reverse transcription where presumed inhibition of this process occur, which in turn could make possible to infer the cause of the blockade and therefore contribute to further elucidation and understanding of the innate response of the immune defence to retroviruses.

### 3. Materials and methods

#### 3.1 Production of pseudotyped HERV-K (HML-2) particles

##### 3.1.1 Plasmid DNA

The following constructs were applied in the production of the two different pseudotyped viral particles (VSV-G and  $\Delta$ KS). All of them provided by the FG 18 division of the RKI.

*Table 1 – Expression vectors for viral particles production.*

<b>Construct</b>	<b>Characteristics</b>
CMVoriEGFP	CMV promoter in the $\Delta$ U3 of the 5' LTR, $\Delta$ splice donor and acceptor sites, internal promoter EF1 $\alpha$ for the expression of the subsequent eGFP sequence in the $\Delta$ env position
<b>oricoGPP</b>	Co-transfection of gag, pro and pol reading frames
<b>oricoRec</b>	Co-transfection of Rec protein sequence
<b>VSV-G</b>	Functional envelop plasmid, containing promoter and env gene of the vesicular stomatitis virus glycoprotein

**ΔKS**

Defective envelop plasmid, containing env gene of HIV-1, impaired through a deletion, for use as a negative control

**Staufen 1**

Increase particle production by enhancing export of transcripts in the cytoplasm and/or translation

---

### 3.1.2 Chemical transformation of E. coli

Through the transformation of One Shot TOP10 Chemically Competent E.coli cells were the recombinant plasmids augmented. The competent E. coli were taken out of the -80°C lager and thawed slowly on ice. Parallel Super optimal broth with catabolite repression (SOC) medium was warmed up at 42°C as well as agar plates were brought at 37°C. After approx. 30 min. were added to each transformation reaction 5 µl of the 1 [ng/µl].

Consequently the mixture was stirred and incubated for 20 min. Then a heat shock of 42°C for exactly 45 sec. was performed before transferring the suspension on ice for two further min. of incubation. Afterwards 500 µl of SOC medium were added and another incubation step on the shaker at 37°C for 60 min. was conducted.

Thereupon 3 different volumes (50, 100 and 350 µl) were spread on 3 different agar plates. Since the recombinant plasmids carry the ampicillin resistance gene, the agar plates contained obviously ampicillin in order to provide an advantage to transformed cells and select for them. The plates were bred over night at 37°C in an incubator.

The inoculation was performed one day after by pricking one single colony of each sample and transferring it into a conical flask with 200 ml of lysogenic broth and 200 µl ampicillin, which correspond to a dilution of approx. 1:1000. The suspension was incubated overnight again in a shaker at 37°C.

### 3.1.3 Isolation of recombinant DNA (rDNA) and quantification

For isolation of plasmid DNA from the transformed bacterial culture the Qiagen Endo Free Plasmid Maxi Kit and Plasmid Maxi Kit were used accordingly to the protocol provided by the manufacturer. Exceptionally and only after the 3<sup>rd</sup> transduction the removal of bound DNA of the spin column, which constitutes the last step of the utilized Qiagen kit, was modified. In order to achieve a higher yield of template DNA for PCR amplification, the columns were eluted with only 100 µl instead of the 200 µl recommended, modifying so, slightly the recommended procedure.

Following, the concentrations of the dissolved DNA were determined by measuring the solutions with a NanoDrop ND-1000 spectrophotometer (ratio of purity assessed, quotients of absorbance at 260nm and 280nm, between 1,8 and 1,9).

### 3.1.4 Cell culture

For the production of viral particles was used the cell line human embryonic kidney 293T (HEK 293T), which were also provided by the FG 18 division of RKI. An aliquot of the cells was thawed out of liquid nitrogen by warming up the suspension in a 37°C water bath for approx. 60 sec. Consequently they were transferred to prewarmed medium and centrifuged at 350 rcf (relative centrifugal force) for 8 min in order to pellet the cells and so separate them of toxic components of the cryomedium (DMSO). Finally the cells were seeded at a density of  $2 \times 10^6$  in a T-75 flask and grown in Dulbecco's Modified Eagle's Medium (DMEM) high glucose, to which was added 10 % of foetal calf serum and 0,5 % of penicillin/streptomycin. The cells were passaged every two days, detaching them with Trypsin/EDTA 0,05 %/0,02% solution and in this manner expanded for 10 days before transfection in order to allow them to stabilize their metabolism after cryopreservation period.

### 3.1.5 Calcium phosphate transfection

Transfection was performed in two different formats. First in 10 cm Petri dishes and later in T-150 tissue culture flasks for higher yield. The cells were seeded in the respective containers approx. 6 hours before transfection at a density corresponding to approximately 55 % of confluence. Cell number determined by haemocytometer during all steps of this work.

*Table 2 – Seeding density for transfection*

<b>Format</b>	<b>Surface (mm<sup>2</sup>)</b>	<b>Seeding density (e<sup>6</sup>)</b>	<b>Medium (ml)</b>
<b>100 mm Petri dish</b>	7,845	4,9	10
<b>T-150 flask</b>	15	9,5	20

*Table 3 – Constitution of the transfection mixture for one 10 cm Petri dish and 33 µg of total DNA.*

<b>Construct</b>	<b>Proportion (%)</b>	<b>DNA quantity [µg]</b>	<b>Plasmid DNA [µg/µl]</b>	<b>Volumes [µl]</b>
<b>CMVoriEGFP</b>	54,6	18	0,14	128,57
<b>oricoGPP</b>	15,2	5	0,11	47,17
<b>oricoREC</b>	15,2	5	0,2	25

VSV-G/ $\Delta$ KS	9,1	3	0,21/0,12	14,08/25
Staufen 1	6,1	2	0,23	34,33
HEPES 1 M				500
CaCl <sub>2</sub> 2,5 M				30
<b>Total</b>		<b>33</b>		<b>753,41</b>

Table 4 – Constitution of the transfection mixture for one T-150 tissue culture flask and 77,1  $\mu$ g of total DNA.

Construct	Proportion (%)	DNA quantity [ $\mu$ g]	Plasmid DNA [ $\mu$ g/ $\mu$ l]	Volumes [ $\mu$ l]
CMVoriEGFP	54,5	42	0,14	300
oricoGPP	15,2	11,7	0,11	47,17
oricoREC	15,2	11,7	0,2	58,5
VSV-G/ $\Delta$ KS	9,1	7	0,21/0,12	32,86/58,33
Staufen 1	6,1	4,7	0,23	20,17

HEPES 1 M		1300
CaCl <sub>2</sub> 2,5 M		77
<b>Total</b>	<b>77,1</b>	<b>2725,11</b>

Respectively 500 or 1300 µl of 1M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) and 30 or 77 µl of 2,5M CaCl<sub>2</sub> (depending on the format) were added to each reaction and mixed by vigorous vortexing. Next the transfection mixtures were incubated for 20 min. at room temperature and in the meantime, the culture medium of the cells was replaced by 8 or 20,5 ml of fresh medium (transfection medium). Consequently the transfection mixtures were carefully transferred to each recipient and these were rocked softly to distribute transfection mixture evenly in medium.

Afterwards the transfections recipients were incubated at 37°C and 5% CO<sub>2</sub> overnight. On the following day medium containing transfection mixtures was replaced by 10 and 24 ml of fresh culture medium. In addition 20 or 48 µl of sodium butyrate were added, to reach a final concentration of 2 mM in every reaction container. Since a GFP gene was also transfected inside the transfer plasmid, transfection progress could be controlled under the fluorescence microscope.

A total of eight 10 cm dishes for the two pseudotypes were transfected, four containing VSV-G plasmid and four containing ΔKS plasmid instead. For the second transfection ten T-150 flasks were used for higher yield of particles. As in the 1<sup>st</sup> transfection the two pseudotypes were distributed evenly among the flasks.

### 3.1.6 Harvest, sterile filtration and storage of viral particles

The supernatants of the transfected cells were collected three and six days after transfection, i.e. two harvest pro transfection. It is known from previous experiments that further harvest only deliver a poor yield of particles, which are not worth to cultivate for. The viral containing supernatants were centrifuged at 1800 rcf for 7 min. at 4 °C in order to pelletize cells and cellular debris prior to sterile filtration with a 0,45 µm pore diameter syringe filter. The viral suspension was kept in ice at all times to prevent inactivation and after sterile filtration it was aliquoted to make possible later partial use of it. Finally the aliquots were moved to the -84°C fridge.

### 3.1.7 Quantification of viral particles - p27 ELISA

The amount of the capsid protein p27 of HERV-K was quantified by ELISA assay, as described next. 96 well plate were coated with antibody HERMA 7 diluted 1:100 in bicarbonate buffer. Then 50 µl of the solution were given into each well. The plates were incubated overnight at 4°C and washed with PBST (PBS plus 0,05 % tween) prior to use. The epitopes of the HERMA 7 were blocked by administrating 200 µl of PBS and 2% milk powder (PM) to each well. Followed by incubation for 1 hr. at 37°C and subsequent wash with PBST.

The 200 µl samples were inactivated with 4 µl of 10 % tween and incubated for 5 min. at room temperature. 50 µl PMT (PM + 0,05 % tween) were applied pro well between rows B and H. Then 75 µl of the pure samples were added to the first row (A). Following, in order to reach a 3-fold dilution, 25 µl of the A-row were transferred in the second row (B, containing 50 µl of PMT).and mixed before continuing repeating the procedure till completing the plate. The standards with an initial concentration of 60 ng/ml were also applied in this form. Subsequent incubation for 1 hr. at 37°C, followed by wash with PBST.

The first antibody, HERV anti-CARatte 5/2 was added in a dilution of 1:5000 (1 µl antibody in 5 ml PMT) and then pipetted to the plate in a volume of 50 µl pro well. Further incubation for 1 hr. at 37°C and subsequent wash with PBST. Then the second antibody antiRat Biologend diluted 1:2500

(2 µl in 5 ml) in PMT was also applied in 50 µl pro well volume, incubated for 1 hr. at 37°C and finally washed again with PBST.

The reaction was then developed with 50 µl pro well of the previously prepared solution of citrate-phosphate buffer, OPD (o-Phenylenediamine) 5 mg and hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) solution 30 % for approx. 15 min. Consecutive the reaction was stopped with 20 µl pro well of 25% sulphuric acid (H<sub>2</sub>SO<sub>4</sub>). Finally the products of the reaction were measured by 492nm in the ELISA reader.

All conducted p27-ELISA assays were performed as above.

## 3.2 Transduction of distinct mammalian cell lines

### 3.2.1 Cell lines, transduction and DNA isolation

Three cell lines of different origin were used for transduction with the pseudotyped HERV-K particles as shown in table number five. However all three have different origin, they present some mutual features, for example all three are epithelial and adherent.

*Table 5 – Transduced cell lines*

<b>Name</b>	<b>Organism</b>	<b>Tissue</b>
<b>CrFK</b>	Felis catus	Kidney, cortex
<b>Cf2TH</b>	Canis familiaris	Thymus
<b>HEK 293T</b>	Homo sapiens	Kidney

Transduction was performed in 24 wells plates, where the cells were seeded 24 hrs. prior exposition to the virus. The surface of each well is 200 mm<sup>2</sup> and at a confluence of 100 % approx. 0,2 x 10<sup>6</sup> cells are present. The chosen seeding density of 0,06 x 10<sup>6</sup> cells correspond to a confluence of 30 %.

0h	1h	2h	4h	8h	GFP control
0h	1h	2h	4h	8h	Mock
0h	1h	2h	4h	8h	Mock
0h	1h	2h	4h	8h	GFP control

**Figure 7 pictures the arrangement a plate set for transduction. Two rows for each viral pseudotype (VSV-G pseudotyped particles in green and  $\Delta$ KS in red) in repeat determination, five intervals and controls.**

For each cell line was prepared a separate 24 well plate and they were incubated overnight at 37°C and 5 % CO<sub>2</sub> in 1 ml culture medium. Also, the night before transduction, the viral suspension were transferred from -80° C to the 4° C fridge, so they can thaw out slowly overnight. Before exposing the cells to the viral particles, the viral suspension was treated with RNase-Free DNase from Qiagen at a ratio of 10 Units enzyme pro ml viral suspension at 37° C for 60 min. and rattled softly a couple of times. This step attempts to digest residual plasmid DNA from transfection and so avoid a posterior amplification of it.

Following, the culture media was removed and the cells washed with phosphate buffer saline (PBS). Than the viral suspension was applied in normalized concentrations, aiming to maintain equals amount of viruses in all the samples. The ratio of viral containing medium and fresh medium was set to 2:1 in a total volume of 330 µl (approx. 220 µl viral suspension and 110 µl fresh medium). To improve infectivity also 2,64 µl of hexadimethrine bromide (polybrene) were added pro well in order to reach a final concentration of 8 µg/ml in each well.

After allowing the viral particles to attach and penetrate the cells during one hour at 37° C the supernatant was exchanged by fresh culture medium. The cells of the first interval (0 hrs. post-

infection) were washed again with PBS and then harvested accordingly to the QIAamp DNA Mini and Blood Mini kit from Qiagen, which was also used for posterior DNA isolation (Appendix B: Protocol for Cultured Cells, B1b cells grown in a monolayer and consecutive protocol: DNA purification from blood or body fluids (spin protocol)). Correspondingly to the previously established time periods the further cells were, in the same manner, then harvested 1,2,4 and 8 hrs. post-infection (p.i.). Thereupon the amounts of DNA were quantified with a NanoDrop ND-1000 spectrophotometer (ratio of purity assessed, quotients of absorbance at 260nm and 280nm, between 1,8 and 1,9).

All the procedures just described above were conducted three times and some variations among their execution will be remarked later.

### 3.2.2 High-titer viral preparation by ultracentrifugation

Attempting to achieve more efficient infection, ultracentrifugation was also performed and tested. First of all for this propose a 20 % sucrose sterile solution was prepared and chilled at 4° C before its use. To ensure sterile conditions the centrifuge tubes 1 x 3 ½ inches (25 x 89 mm) were submerged in a 70 % ethanol solution one day before use and left in there overnight. The viral solution was also moved from -80° C to 4° C the night before its use.

The rotor sw32 was used, which offers place for six tubes. Each of ones was filled with 9 ml of the chilled 20 % sucrose sterile solution, then topped up with approx. 30 ml of viral medium and tared. The viruses were than in the precooled centrifuge palletised at 175.000 rcf for four hrs. After the supernatant was discarded the viral pellet was resuspended in 250 µl cooled PBS, aliquoted and storage at - 80° C.

### 3.2.3 Complementary DNA PCR amplification and agarose gel analysis

For specific amplification of the sequences of interest the following primers set were supplied and designed by the FG 18 division using the primer design software of DNASTAR. The volumes

for the PCR reactions was conceived to maximize the amount of isolated DNA template. The positive control used for amplification was solution containing  $2 \times 10^{-2}$  ng/ $\mu$ l of plasmid DNA of the construct CMVoriEGFP. The same mixture with nuclease free water instead of the plasmid DNA template was employed as negative control.

*Table 6 – Complete list of available primer sets for targeted amplification*

Number	Primer	Sequence	Length
1	qHervU5-For	TATGCTGAACGCTGGTTCC	19
2	qHervU5-Rev	CTACACACCTGTGGGTGTTTCT	22
3	HervRU5_1649-1668_For	TACTAAGGGAACTCAGAGGC	20
4	HervU5_1802-1818_Rev	TGTAGGGGTGGGTTGCC	17
5	HervPBS_1829-1849_Rev	CTAGAGAAAAGCCTCCACGTT	21
6	HervPBS_1858-1875_Rev	TGACCACGCTCGAGCGTA	18
7	HervU3_1032-1049_Rev	TGCACCGCCCTTAATCCA	18

Table 7 – Exemplary PCR reaction set up for a 50 µl final volume

Components	Vol in µl
10x Maxima Hot Start Taq buffer	5
25 mM MgCl <sub>2</sub>	5
Forward primer (3)	2,5
Reverse primer (4)	2,5
Maxima Hot Start Taq DNA Polymerase	2
2,5 mM dNTP Mix	2,5
Water, nuclease free	0
Total	19,5
Template DNA	30,5

Table 8 – Exemplary PCR cycling

Step	Temp (°C)	Time (min)	Cycles
Initial denaturation	95	4	1
Denaturation	95	1	22 – 32
Annealing	63	1	22 – 32
Extension	72	0,75	22 - 32
Final extension	72	15	1
Preserve	4	∞	1

Samples for electrophoresis were always prepared in a 12 µl final volume, containing 10 µl of the sample itself and 2 µl of a 6x loading buffer. Due to the size of the expected fragments, 2 % agarose concentration was used in the blend of gels. For the same reason a 100 bp DNA ladder was utilized for sizing and for staining 4 µl of ethidium bromide were added to each 100 ml gel. The gels were run for approx. 35 - 40 min. at 110 V in TAE buffer.

## 4. Results

### 4.1 Quantification of transfected viral particles

The viral yields were collected 3 and 5 days after transfection and quantified by p27-ELISA as described above. This procedure was conducted in total three times, once after each transfection and one more time after ultracentrifugation. This data reflects strictly the concentration of the capsid protein p27.

*Table 9 – Concentrations of p27 in pg/ml of the two pseudotypes in two different formats*

	VSV-G		ΔKS	
	1 <sup>st</sup> harvest	2 <sup>nd</sup> harvest	1 <sup>st</sup> harvest	2 <sup>nd</sup> harvest
<b>10 cm Petri</b>	790,64	573,32	860,78	625,68
<b>T-150 flask</b>	1325,42	250,67	1012,07	758,42

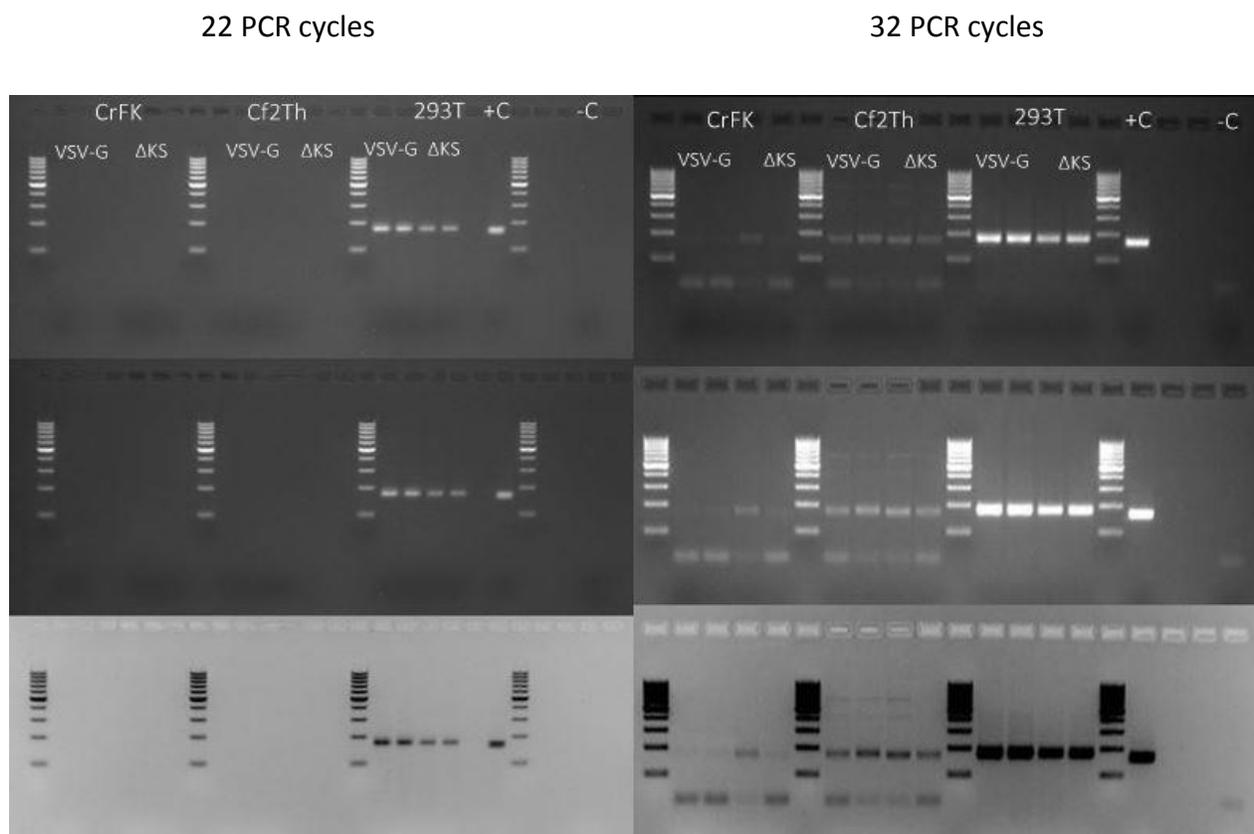
### 4.2 Electrophoretic analysis in agarose gel

Following will be presented the results for detection of early products of reverse transcription, which were amplified with primers no. 3 (HervRU5\_1649-1668\_For) forward and no. 4 (HervU5\_1802-1818\_Rev) reverse. This set of primers binds specific to the first retro-transcribed DNA sequence, the minus strong stop strand, before the first template exchange. The forward primer binds across the intersection of R and U5 and the reverse primer hybridizes to the last 17 bp of U5 until immediately before the PBS, at the 5' end LTR.

The transductions procedures lead partially to slightly differing data. Therefore they will be presented associated to the transduction process through which they were obtained and in dependence of the number of PCR cycles.

All results are presented in homologous manner. 22 and 32 amplification cycles on left and right correspondingly and three shots with different camera settings vertically disposed. All the samples of first and second transductions in replicate. Third transduction conducted only with single samples. Each cell line is separated through the GeneRuler 100 bp DNA ladder, where the second band, bottom to top, 200 bp represents. The amplicon as well as the positive control (+C) are 169 bp long.

#### 4.2.1 – 1<sup>st</sup> Transduction



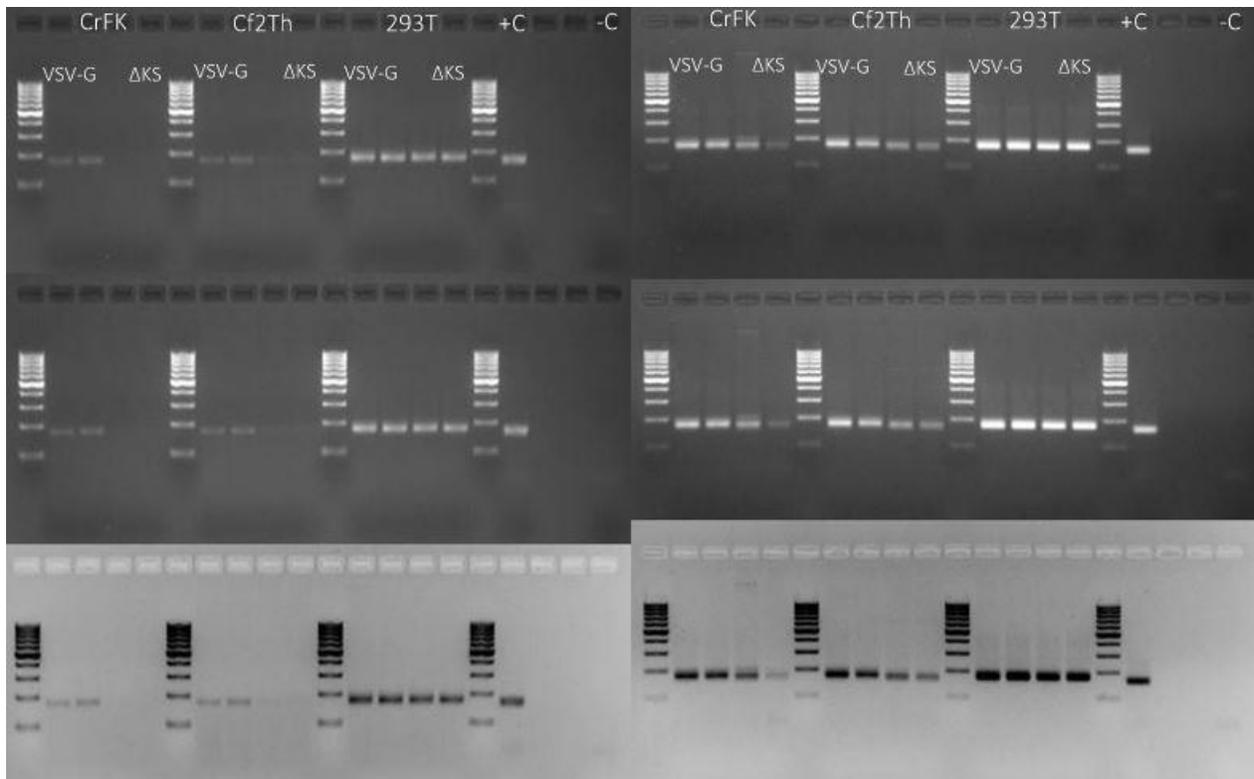
**Figure 8 shows already after 22 cycles amplified sequences of the human cell line, where the targeted sequences also belong to the isolated chromosomal material. However it is clear that the two 293T**

*replicates transduced with the VSV-G pseudotyped particles exhibit more pronounced bands than the ones transduced with  $\Delta$ KS. This effect is also lightly exhibited after 32 cycles. Although the bands of the other two cell lines represent a discrepancy to the rest of the results and are considered not representative. Moreover they can be attributed to a faulty dislodge method.*

#### 4.2.2 – 2<sup>nd</sup> Transduction

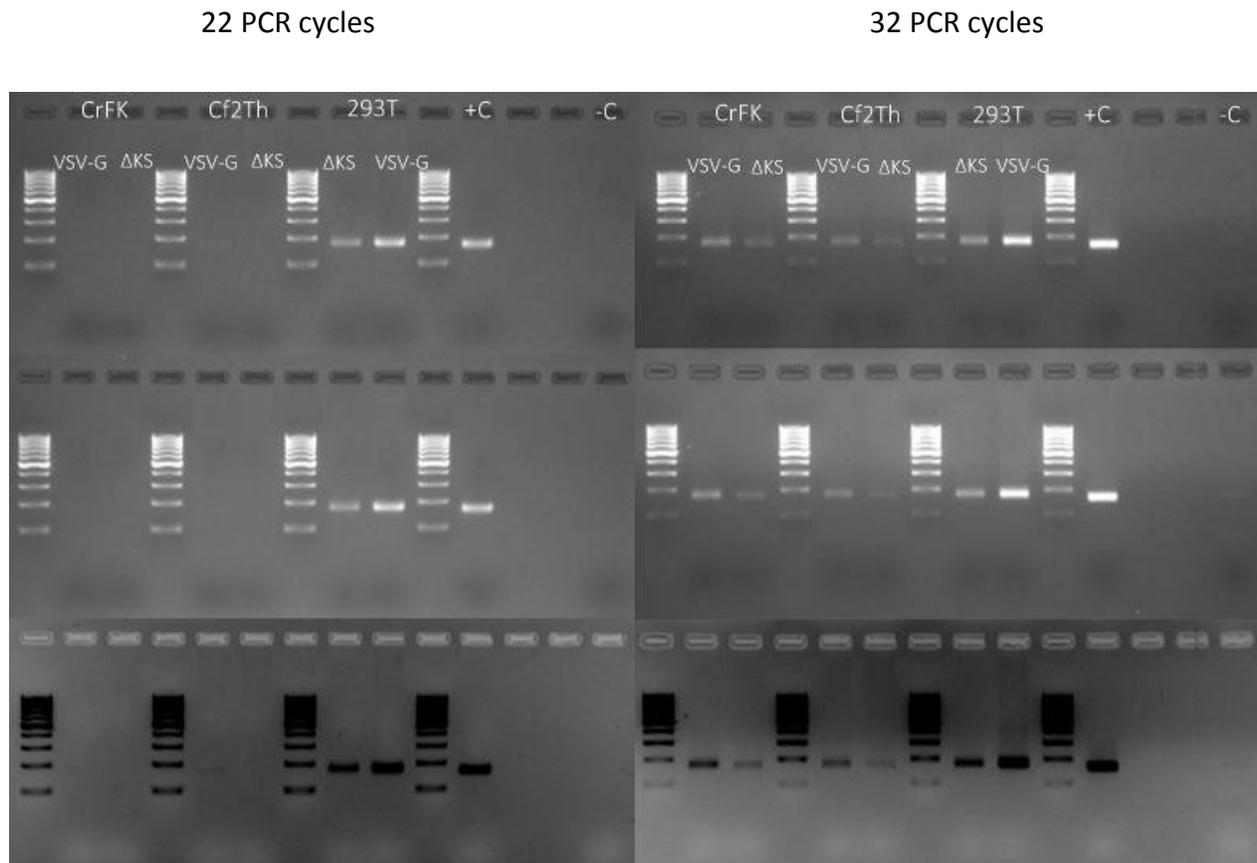
22 PCR cycles

32 PCR cycles



*Figure 9 displays distinct bands after 22 cycles in the VSV-G transduced samples and also accordingly difference of intensity in the bands of the 293T cells samples. The 32cycles picture exhibits analogous results.*

### 4.2.3 – 3<sup>rd</sup> Transduction



*Figure 10 depicts results that correspond to the ones from the 2<sup>nd</sup> transduction and partially to the 1<sup>st</sup> one also. However after 22cycles the increased DNA presence absent in the CrFK cells and weak in the Cf2Th samples. While after ten more cycles the results corroborate the previous ones.*

## 5. Discussion

The analysis of interactions between ERVs and their hosts provides remarkable conditions for further understanding of evolutionary events and processes that have developed along great periods of time. Moreover comprehension of the current situation and the facts that lead to it can supply valuable information, reflecting strategies of invasion and defence. That is why elucidation of the outcomes in this relationship could result extremely beneficial. Especially if

these facts reveal knowledge that can be applied to gain advantage in this evolutionary “arms race”.

In humans even the endogenous retroviral family with the most effective members, HERV-K (HML-2), is not able of producing infectious particles any more [25]. Further, it has been shown that functionally reconstituted virions of this family are still incapable of accomplish replication, although several different types of host cells are susceptible but no permissive. In other words the viral particles are capable to move across the cell membrane into the cytoplasm, however they fail to produce infectious particles. This issue presumptively relies on inhibition through cellular factors [26].

Nevertheless the obstruction of retroviral replication cycle may take place at different levels. For example after integration at pre-transcriptional level through epigenetic mechanisms, i.e. methylation of DNA containing proviral sequences. Also by modification of host tropism through mutations of receptors that alter cellular susceptibility to the incoming particles [14]. Maybe it can be presumed that loss of infectivity was first achieved through restriction factors and afterwards at the level of gene expression like mutations or gene silencing. This chronological order could maintain the selective pressure on restriction factors until they are able to turn the virus harmless. On the contrary if infectiousness was first disable at the level of gene expression, the probability of dysfunctional gene products applying selective pressure on the evolution of restriction can be lower. This could be supported by the fact of reconstituted virions still being blocked. However it has also been shown that restriction factors efficiently inhibit a variety of different retroviruses [16] and even other non-retroviral viruses [27], which could still left open the possibility of blocking more viruses by the presence of similar structures or even by mere coincidence.

Nonetheless in the context of this work it will be observed the stage of reverse transcription. Here the proteins most likely to produce a blockade are the ones targeting the steps of retroviral replication after entry and prior integration. These proteins, termed restriction factors, form part of the innate immune response and have coevolved together with retroviruses to become a cellular response to viral invasion [28]. Some of the best known are tripartite motif-containing 5 $\alpha$

(TRIM5 $\alpha$ ) which inhibits retroviral uncoating and RT, SAM domain- and HD domain-containing 1 (SAMHD1) that hinders RT by reducing dNTPs available and apolipoprotein B mRNA-editing complex 3 (APOBEC3) which also impedes RT [29].

The results for all transductions show clearly that already after 22 PCR cycles the HERV-K sequences integrated in the genome of the human cell line become detectable on gel. This fact certainly relies on the amount of template present at the beginning of the amplification [30]. Moreover the samples transduced with the VSV-G pseudotyped virions deliver stronger bands than the ones transduced with  $\Delta$ KS, which also argues the size of the bands in dependence of the initial amount of template. Still longer cycling amplifies also segments of DNA in the feline and canine samples transduced with the virions carrying the incompetent envelop protein, what probably can be ascribed to spontaneous plasmid DNA uptake during transfection

Obviously this demonstrate that the particles are able to reach the cytoplasm, initialize RT and complete the retrotranscription of the minus strong stop DNA before the first template exchange from the 5'- to the 3' LTR occurs. In this case the cytidine deaminase action of APOBEC-3G and 3F might not be detected, because this restriction factor cause point mutations in the synthesis of the minus strong stop DNA, exchanging C to U and therefore provoking G to A mutations in the complementary plus strand of DNA [31], which could be later on deleterious for viral replication, without necessarily interrupting elongation. However there is also evidence of RT inhibition by these proteins, which in the case of HIV-1 can be antagonized by the viral infectivity factor (Vif) [32]. Instead HERV-K does only code for one more protein besides the four minimal genes required for replication. This protein is either Rec or Np9, which reassemble the function of REC in HIV. In this regard it can be also interesting to infer to what extent the expression levels of different restriction factors differ in diverse cell types.

In a similar way the bands of amplified DNA after 32 PCR cycles corroborate the previous mentioned results with exception of the 1<sup>st</sup> transfection, where only the humans HEK 293T cells show more DNA in the VSV-G transduced samples. In contrast the sequences amplified from feline and canine cells do not coincide with the rest of the results. The reason for this it is probably a different method during the harvest of the cells. Samples of the 1<sup>st</sup> transduction were partially

detached of the culture plate by using a cell scrapper (unlike the following two transductions, where dislodge of samples was performed with trypsin), which probably lead to adulteration of the results.

Another factor that can be considered is the degradation of plasmid DNA remaining in the viral suspension after transfection. Although the clearance of the mentioned “carry-over plasmid” was only performed in two of the three transduction experiments no variation in the results was detected. However this could turn relevant when trying to quantify the transcribed DNA by means of qRT-PCR, method that possess a higher sensitivity. Jean L. Mbisa et al. published in 2009 such analysis of post-entry events of HIV-1 [33] and their work stimulated the present study for the particular case of HERV-K. Although in the case of HIV besides the immediate therapeutic aspects, i.e. resolving the role of unintegrated DNA [34] can be relevant to point out that within retroviruses only lentiviruses possess the capability to translocate the PIC from the cytoplasm, through nuclear pores, into the nucleus [3]. Then betaretroviruses and all other retroviruses in general are only able to reach the nucleus of dividing cells. This characteristic has a major influence in their tropism, determining permissiveness by quiescent cells.

Certainly there are a copious amounts of aspects that can be improved regarding the procedures conducted in this work. For example it could be helpful to establish a stable transfected cell line in order to enhance performance of virus production, which will make easier to transduce larger amount of cells and so isolate higher concentrations of template DNA for amplification. It is also worth to mention that when testing in an isolated system, such as a cell culture, potential synergetic aspects might not be present and also increasing number of controlled variables very often positively correlate with a raising loss of realism.

## 6. References

1. Prangishvili, D., *The wonderful world of archaeal viruses*. Annu Rev Microbiol, 2013. **67**: p. 565-85.
2. Forterre, P., *Defining life: the virus viewpoint*. Orig Life Evol Biosph, 2010. **40**(2): p. 151-60.
3. Susanne Modrow, D.F., Uwe Truyen and Hermann Schätzl, *Molekulare Virologie 3. Auflage*. 2010.
4. Temin, H.M. and S. Mizutani, *RNA-dependent DNA polymerase in virions of Rous sarcoma virus*. Nature, 1970. **226**(5252): p. 1211-3.
5. King, O.E.A., *Virus Taxonomy ninth report of the ICTV*. 2011: p. 477-495.
6. Levy, *Structural and functional studies of the HIV-1 pre-integration complex*. Retrovirology, 2013. **10**: p. 76.
7. Kvaratskhelia, M., et al., *Molecular mechanisms of retroviral integration site selection*. Nucleic Acids Res, 2014. **42**(16): p. 10209-25.
8. Campbell, N.A. and J.B. Reece, *Biology*. 8. ed. 2008, San Francisco, CA u.a.: Pearson Education. XLVI, 1267, 152 S.
9. Bittner, J., *Some possible effects of nursing on the mammary gland tumor incidence in mice*. Science, 1936. **84**(2172): p. 162.
10. Oliveira, N.M., et al., *Changes in viral protein function that accompany retroviral endogenization*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(44): p. 17506-17511.
11. Munoz, J. and A.J.R. Heck, *From the Human Genome to the Human Proteome*. Angewandte Chemie-International Edition, 2014. **53**(41): p. 10864-10866.
12. Agoni, L., C. Guha, and J. Lenz, *Detection of Human Endogenous Retrovirus K (HERV-K) Transcripts in Human Prostate Cancer Cell Lines*. Front Oncol, 2013. **3**: p. 180.
13. Dewannieux, M., et al., *Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements*. Genome Res, 2006. **16**(12): p. 1548-56.
14. Stoye, J.P., *Studies of endogenous retroviruses reveal a continuing evolutionary saga*. Nat Rev Microbiol, 2012. **10**(6): p. 395-406.
15. Lee, Y.N. and P.D. Bieniasz, *Reconstitution of an infectious human endogenous retrovirus*. PLoS Pathog, 2007. **3**(1): p. e10.

16. Yap, M.W., et al., *Evolution of the Retroviral Restriction Gene Fv1: Inhibition of Non-MLV Retroviruses*. Plos Pathogens, 2014. **10**(3).
17. Mülhardt, C., *Der Experimentator Molekularbiologie/Genomics*. 7., aktualisierte Auflage ed. Der Experimentator. 2013, Berlin, Heidelberg: Springer.
18. James Cronin, X.-Y.Z.a.J.R., *Altering the tropism of lentiviral vectors through pseudotyping*. 2005.
19. Jordan, M. and F. Wurm, *Transfection of adherent and suspended cells by calcium phosphate*. Methods, 2004. **33**(2): p. 136-43.
20. Roche, S., et al., *Structures of vesicular stomatitis virus glycoprotein: membrane fusion revisited*. Cell Mol Life Sci, 2008. **65**(11): p. 1716-28.
21. Howard, C.M.G.a.B.H., *Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate*. Nucleic Acid Research, 1983. **11**(21).
22. Davie, J.R., *Inhibition of Histone Deacetylase Activity by Butyrate*. The Journal of Nutrition, 2003. **133**: p. 2485-2493.
23. Subramanian, R.P., et al., *Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses*. Retrovirology, 2011. **8**: p. 90.
24. Wilhelm, M. and F.X. Wilhelm, *Reverse transcription of retroviruses and LTR retrotransposons*. Cellular and Molecular Life Sciences, 2001. **58**(9): p. 1246-1262.
25. Hohn, O., K. Hanke, and N. Bannert, *HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease*. Front Oncol, 2013. **3**: p. 246.
26. Lausch, V., *Etablierung eines Verfahrens zur Kombination der Elektronenmikroskopie (EM) mit der Spitzenverstärkten Raman-Spektroskopie (TERS) sowie Untersuchungen zum Infektionsprozess und der Ultrastruktur des Humanen Endogenen Retrovirus K (HERV-K)*. 2013. p. VI, 175 S.
27. Wolf, D. and S.P. Goff, *Host restriction factors blocking retroviral replication*. Annu Rev Genet, 2008. **42**: p. 143-63.
28. Duggal, N.K. and M. Emerman, *Evolutionary conflicts between viruses and restriction factors shape immunity*. Nat Rev Immunol, 2012. **12**(10): p. 687-95.
29. Armitage, A.E., et al., *Conserved footprints of APOBEC3G on Hypermutated human immunodeficiency virus type 1 and human endogenous retrovirus HERV-K(HML2) sequences*. J Virol, 2008. **82**(17): p. 8743-61.

30. Hermey, G., *Der Experimentator: Neurowissenschaften*. <<Der>> Experimentator. 2010, Heidelberg: Spektrum. X, 273 S.
31. Isel, C., C. Ehresmann, and R. Marquet, *Initiation of HIV Reverse Transcription*. *Viruses*, 2010. **2**(1): p. 213-43.
32. Refsland, E.W. and R.S. Harris, *The APOBEC3 family of retroelement restriction factors*. *Curr Top Microbiol Immunol*, 2013. **371**: p. 1-27.
33. Vinayaka R. Prasad, G.V.K.e., *Real-Time PCR Analysis of HIV-1 Replication Post-entry Events*. *HIV Protocols: Second Edition*, 2009. **485**: p. 55-72.
34. Wainberg, R.D.S.a.M.A., *The role of unintegrated DNA in HIV infection*. *Retrovirology*, 2011.