Expansion of HIV-1 Infected Cells and the Possible Influence on Tropism Dynamics in Individuals on Antiretroviral Therapy
&
Adapting the geno2pheno[coreceptor] Tool to HIV-1 Subtype CRF01_AE by Phenotypic Validation using Clinical Isolates from South-East Asia

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Nina Marty

2024

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Erstbestreuer/in:	Prof. Dr. rer. nat. Thomas Klimkait
Zweitbetreuer/in:	Prof. Dr. Markus Affolter
Externe/r Experte/in:	PD Dr. Jürg Böni

Basel, den 22.02.2022

Prof. Dr. Marcel Mayor
Dekan der Philosophisch-
Naturwissenschaftlichen
Fakultät

Table of Contents

1 List of Tables

2 List of Figures

3 Abbreviations

AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Therapy
ATL	Adult T-cell Lymphoma/Leukemia
bp	Base Pair
CCR5	CC-Motive Chemokine Receptor 5
CRF	Circulating Recombinant Form
CTL	Cytotoxic T-Lymphocyte
CXCR4	CXC-Motive Chemokine Receptor 4
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleic Triphosphate
Env	Envelope
FPR	False Positive Rate
g2p	Geno2pheno[coreceptor]
gDNA	Genomic DNA
gp120	Glycoprotein 120
HIV-1	Human Immunodeficiency Virus Type 1
HTLV	Human T-cell Lymphotropic Virus Type 1
IC90	90% Inhibitory Concentration
kb	Kilobases
LB	Lysogeny Broth
LTR	Long Terminal Repeats
MSM	Men Who Have Sex With Men
MVC	Maraviroc

nd	Not Determined
nef	Negative Factor Protein
NGS	Next Generation Sequencing
NNRTI	Non-nucleoside Reverse Transcriptase Inhibitor
NRTI	Nucleoside Reverse Transcriptase Inhibitor
PBMC	Peripheral Blood Mononuclear Cell
PBS	Phosphor-buffered Saline
PCR	Polymerase Chain Reaction
R5	Chemokine Receptor CCR5
Rev	Regulator of Expression of Virion Proteins
RNA	Ribonucleic Acid
RRE	Rev Response Element
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SD	Standard Deviation
SHCS	Swiss HIV Cohort Study
T_{CM}	Central Memory T Cell
T_{EM}	Effector Memory T Cell
T_{fh}	T Follicular Helper Cell
T_N	Naïve T Cell
Tat	Trans-Activator of Transcription
V3	Variable Loop 3
VL	Viral Load
qPCR	Quantitative Polymerase Chain Reaction
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

X4	Chemokine Receptor CXCR4
%X4	Frequency of X4-tropic HIV Variants

4 Abstract

For cellular infection HIV-1 interacts with the CD4 receptor and a chemokine receptor, either CCR5 or CXCR4. In the early phase of infection, the virus almost exclusively uses the CCR5-receptor. Only later during disease progression up to 50% of infected individuals without therapy experience a shift to a dominance of CXCR4-tropic viruses, associated with higher CD4 cell depletion rates and an accelerated disease progression. In contrast, under successful therapy a majority of patients show a decrease of CXCR4-tropic variants and only in a few an increase of CXCR4-tropic viruses is observed. As the increase of CXCR4-tropic viruses in patients under therapy was observed to be caused by the outgrowth of a single variant, this study aimed to investigate the possibility of clonal expansion of HIV-1 infected cells by integration site analysis. Although the results confirmed clonal expansion, aberrant cell proliferation due to HIV integration was not observed.

An additional study was performed to validate the popular HIV-1 genotyping tool, Geno2Pheno[coreceptor], for clinical samples of subtype CRF01_AE. The algorithm was shown to have a significant overcalling of X4 variants in this subtype, and by combining phenotypic results with genotypic predictions an adjusted false positive rate (FPR) cut-off was identified to provide more accurate tropism predictions. Included in this work was also the development of a cloning cassette optimized for phenotypic analysis of subtype CRF01_AE samples.

5 Introduction

5.1 HIV-1 Coreceptors

For cellular infection the human immunodeficiency virus 1 (HIV-1) interacts with the CD4 receptor and a chemokine coreceptor, either CCR5 or CXCR4.

HIV variants that infect cells via the CCR5-coreceptor are referred to as R5-tropic, while variants binding to the CXCR4-coreceptor are called X4-tropic.

The identification of chemokine receptors as co-receptors has enhanced the understanding of the cellular entry, viral transmission, and pathogenesis of infection and has played a crucial role in the ongoing development of HIV-1 treatment strategies[1]. A mechanistically new antiretroviral drug, the entry inhibitor Maraviroc (MVC), was developed for the treatment of HIV patients. As MVC specifically blocks the CCR5 coreceptor but not CXCR4, a treatment decision for MVC is linked to the prior determination of coreceptor usage in the respective patient, to ensure that MVC can successfully suppress viral replication[2]. Despite ongoing research, there are currently no entry inhibitors for X4-tropic viruses approved for HIV-1 therapy[3–6].

The co-receptor usage of HIV-1 can either be analyzed functionally *in vitro,* using cell culture assays, termed phenotyping or by sequence analysis of a specific HIV envelope (env) region, termed genotyping. Phenotyping tests often involve a long turnaround time due to the need of sophisticated cell culture formats. Also, most tests use recombinant virus assays that are based on the genetic backbone of HIV subtype B (e.g. NL4-3 or HXB2). Therefore, the analysis of clinical non-B subtype isolates *in vitro* may not be straight-forward.

The genotypic assays are based on the amino acid sequences of the third variable region of gp120 (V3), which is the most critical region in the HIV-1 genome responsible for in coreceptor binding[7]. The V3 loop is generally 35 amino acids long and especially

the positions 11, 24, and 25 in the loop are critical as positively charged amino acids at these positions are strongly associated with an X4-tropism[8].

These specific positions are used for genotypic prediction of tropism (11/25 rule)[9]. Other features that are included in genotypic tools are the net charge and potential N-glycan site in the V3 region[10,11].

The most widely used genotypic tool to predict tropism of HIV is the web tool geno2pheno[coreceptor][12]. Genotypic data pairs and corresponding phenotypic information, mostly based on subtype B, were used to develop the geno2pheno prediction system by machine learning.

The interpretation system of geno2pheno[coreceptor] is designed so that its homology to R5 leads to predicting R5 tropism for a newly sequenced variant and homology to X4 variants leads to prediction of X4 of any newly sequenced variants. In this algorithm, variants with homology to neither R5 nor X4 are classified as X4 with the intention of avoiding possible clinical overuse of MVC. Consequently, sequences with low similarity to the phenotyped variants in the training sets, such as divergent non-B strains of HIV-1, are more frequently predicted as X4. This simplification results in the fact that significant populations of individuals in regions dominated by non-B strains completely miss out on treatment options using MVC.

5.2 HIV-1 Tropism Dynamics

At the early phase of infection, most patients present with about 80-90% R5-tropic virus variants. In fact, the acute phase of a primary HIV infection is characterized by a preferential infection of R5-expressing cells such as macrophages and other monocyte-derived lineages[13,14].

In the absence of effective therapy, X4-tropic variants emerge over time, and late in disease more than 50% of all patients infected with HIV subtype B harbor X4-tropic

viruses in detectable amounts[13,14]. The switch from R5- to X4-tropism is associated with a faster $CD4^+$ T cell depletion and a faster disease progression[13–15]. The occurrence of X4-tropic variants also prohibits the use of MVC, limiting possible therapy and cure strategies[16,17].

The reasons for the selective infection advantage of R5-tropic viruses and the observed viral tropism switch are not entirely clear to date and are rather seen as interplay between different barriers and mechanisms[18]:

One possible explanation is that this HIV-1 coreceptor switch occurs as a result of the evolution of viral populations during the course of infection. The late or absent emergence of X4 virus could be explained with the assumption that intermediate mutants have a fitness disadvantage when compared with the initial R5 and the final X4 virus[19]. However, since relatively few genetic changes are required for a switch from R5-tropic to X4-tropic, it would be expected to see this switch more rapidly and frequently in vivo if genetic changes were solely responsible. It has been shown that minor proportions of X4-tropic viruses exist throughout the infection and X4-tropic variants can emerge independently during the course of infection; in some individuals even during the acute phase of infection[20–22]. Therefore, it is very likely that viral genetic changes alone are not sufficient and that host factors are required to allow X4-tropic variants to become the dominant viral population.

A second possible explanation is that the observed HIV coreceptor switch occurs due to differential susceptibility to immune control of X4 and R5 viruses. X4-tropic viruses may have been transmitted with the R5 variants, but their expansion is prevented by an effective immune response during the primary infection

X4-tropic viruses are often less glycosylated than R5- tropic viruses and it has been suggested this might result in easier recognition and elimination by the host immune

system[11,23,24]. The late emergence of X4-tropic variants might then be associated with a compromised immune system[25,26].

Finally, the HIV-1 coreceptor switch may be explained by the differential target cell ranges of R5 and X4 viruses. CXCR4 is usually expressed on a high fraction of naive $CD4^+$ T cells, compared to a substantial fraction of memory cells that express both CXCR4 and CCR5[27]. In this scenario the changes in the composition of the target cell pool during the course of infection would be responsible for the tropism switch.

So far, it is still unclear if the tropism switch is the cause for or the consequence of a failing immune system and faster disease progression. Recently, a study showed first hints that the host immune activation may be the driving force behind the coreceptor switch[28]. A deeper understanding of the correlates of HIV-1 co-receptor tropism is important to develop future cure strategies.

The influence of co-receptor usage on the course of antiretroviral therapy (ART) or vice versa the effects of ART on tropism are not well established.

Some studies have reported that patients harboring X4 viruses at baseline, display poorer immunological recovery despite comparable viral load suppression than those patients exclusively infected with R5 viruses[29,30]. Other studies have reported no difference in patients harboring an R5 or X4 tropic virus in the recovery of CD4 cell count over several years of effective ART[31,32] and associated X4 tropism with a lower nadir CD4 cell count rather than with current CD4 cell count[33–35].

The effects of ART on HIV co-receptor tropism are currently a popular topic of study, however results have been inconclusive to date. While some studies observe an overall tropism switch from R5 to X4 in patients on ART[36], others show tropism switches in both directions[37], and some studies observe a preferential suppression of X4-tropic strains of HIV-1 by antiviral therapy[38,39]. Bader et al.[39] reported, that while

patients with R5-tropic virus/provirus at therapy initiation remained R5-tropic after 4 years on therapy, almost all patients presenting with X4-tropic virus/provirus prior to therapy initiation showed a decrease of X4-tropic variants after 4 years on treatment. This study also observed that, in some patients (7 out of 35), the relative frequency of X4-tropic proviruses did not decline, but rather increased under ART. During the course of the study, R5-tropic variants were frequently retained as swarms, while in patients with an increased frequency of X4-tropic variants, an outgrowth of a single variant was observed. Subsequent sequence analysis showed that the respective X4-tropic proviruses had already been present as a minority prior to therapy. This emergence of identical HIV-1 variants in individuals on long-term ART has also been observed in other studies[40–43]. The expansion of identical HIV variants under ART is believed to be responsible for the persistence of the HIV reservoir. As the persistence of HIV-1 infected cells is the main barrier to a cure, the source and mechanisms involved in the emergence of these identical variants are of high interest.

5.3 Latent HIV-1 Reservoir

Around 38 million people were living with HIV/AIDS as of 2019[44] and although the advent of ART in 1996 improved the life expectancy and quality of life of these people enormously, there is still no cure for HIV infections. ART blocks new cycles of virus replication; plasma virus levels are reduced to below the clinical assay detection limit (20-50 copies of HIV RNA/mL) and disease progression is stopped. However, if ART is interrupted viremia rebounds rapidly[45]. The principal barrier to a cure is the formation of stable reservoirs of latent HIV, which are established early in infection and persist even in patients under long-term ART with no detectable viremia in the peripheral blood[46,47]. Viral reservoirs have been defined as cell types or anatomical sites in which

replication-competent forms of the virus persist with more stable kinetic properties than in the main pool of actively replicating virus[48,49].

Several factors are known to influence the shape of the viral reservoir, initiation of ART during acute HIV-1 infection substantially accelerates the decay rate[42], whereas viral blips and low-level viremia during ART slow it down[50]. Conversely, treatment intensification does not seem to influence the decay rate[51]. Also, host genetic factors have been found to have little, if any, influence on the size and dynamics of the HIV-1 reservoir[52].

While in the first year after ART initiation HIV DNA in blood decays rapidly, this decrease slows during the following years and subsequently plateaus[53,54]. It has been shown that HIV DNA levels stay stable in individuals receiving long-term ART[42,55,56] and although early ART initiation reduces the reservoir size and genetic complexity[42,57], it does not prevent the generation of latently infected CD4+ cells[58].

The durability of latent reservoirs was first described as the result of the long lifespan of non-dividing resting memory CD4+ T cells endowed with pro-survival capacities[58,59], but more recent studies have demonstrated that the durability of the reservoir is due to massive and sustained clonal expansion of cells harboring both intact and defective proviruses[60–62].

Several mechanisms are thought to have an influence on the formation and dynamics of the clonal expansion of HIV infected cells (Figure 1).

One reason for clonal expansion of HIV infected cells could be HIV integration in or near genes associated with cell growth. Using integration site analysis, cellular clonal expansion of provirus-harboring cells has been investigated[40,41]. Many of the clonally expanded cells had HIV integration sites in cell cycle-regulation genes, such as BACH2, MKL2 and STAT5B[40,41,63,64] and it is therefore hypothesized that HIV

integration in these genes enhances transcription and enables these cells to expand at a higher rate and to persist for long periods of time.

Another explanation for the clonal expansion of HIV infected cells could be homeostatic proliferation. During chronic HIV-1-infection, the proliferation capacity of $CD4^+$ T cells is significantly impaired due to chronic immune activation, decreased IL-7 receptor expression, immune exhaustion and the destruction of lymphoid tissue[65–68]. It has been shown that IL-7 induces proliferation of HIV-1 infected cells without reactivating latent HIV, hence HIV infected $CD4^+$ T cells may undergo homeostatic proliferation without being recognized by immune surveillance[69,70].

A third mechanism responsible for clonal expansion of HIV infected cells could be antigen-driven proliferation. Evidence suggests that chronic or repeated exposure to antigens is able to stimulate the clonal expansion of HIV infected cells[60,71,72]. Antigen-driven proliferation of HIV infected cells could also explain the viral blips and persistent viremia observed in some patients, despite strict adherence to ART[43,73].

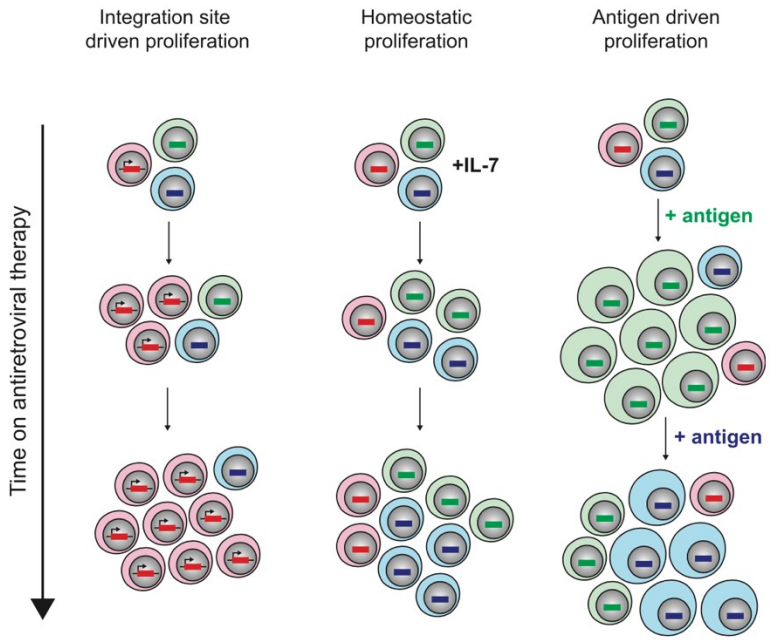

Figure 1 Mechanisms involved in clonal expansion of HIV-1 infected cells. First, the viral integration site may provide a survival advantage allowing preferential proliferation of the infected clone. Second, homeostatic cytokines may signal latently infected cells to divide. Third, latently infected CD4+ T cells with antigen specific T cell receptors may divide in response to recurrent antigen exposure. Adapted from Cohn et al.[74]

The suggested mechanisms responsible for clonal expansion are not mutually exclusive, and it is likely that all mechanisms apply to varying degrees across individuals and perhaps time.

Elucidating the formation and dynamics of the clonal expansion of HIV infected cells will be essential for development of novel cure strategies.

6 Clonal Expansion of HIV-1 Infected Cells and the Influence on Tropism Dynamics

6.1 Aim of the Study

Bader et al.[39,75], recently reported, that while patients with R5-tropic virus/provirus at therapy initiation remained R5-tropic after 4 years of therapy, almost all patients presenting with X4-tropic virus/provirus prior to therapy initiation had exclusively R5-tropic variants after 4 years of treatment. However, a surprise finding was that in some patients (7 of 35), the relative frequency of X4-tropic proviruses did not decline, but rather increased under ART. During the course of the study, R5-tropic variants were often retained as swarms, while a (pseudo-)clonal nature of the emerging X4-tropic proviral sequences was detected. Subsequent sequence analysis showed that the respective X4-tropic proviruses had already been present as a minority prior to therapy. Based on the inclusion criteria of good CD4 recovery (ΔCD4 > 200 cells/mm^3) and absence of a detectable viral load (< 20 cells/mL) or blips in the study population, a contribution from free virus with its error-prone reverse transcriptase and new infection events were excluded as an underlying cause of the observed proviral expansion in peripheral blood mononuclear cells (PBMCs).

This project aims to study the nature and cause of this observed (pseudo-)clonal expansion of X4-tropic proviruses by studying the effects of HIV-1-integration sites on cell proliferation and to further investigate the dynamics of proviral tropism in patients under long-term ART.

6.2 Results

6.2.1 Patient Characteristics

Bader et al.[39] studied the frequency of X4-tropic variants during phases of immune recovery by administration of ART in 35 patients from the Swiss HIV Cohort Study (SHCS) over 4 years. With regards to relative frequencies, the proviral X4-tropic HIV-1 variants decreased or remained stable over time in the majority of patients (28 of the 35 patients, 80%, $p<0.001$). In 7 patients an increase of the frequency of X4-tropic variants was observed.

To further analyze the proviral situation in the 7 patients who showed an increase of frequency of X4-tropic variants (X4-patients), PBMCs from 10 years post (post10) ART initiation were included in this study in addition to the pre-therapy (preT) and post 4 year (post4) samples. PBMCs from 9 patients, which presented in the previous study[39] with a low and stable frequency of X4-tropic variants (R5-patients) were used as control.

The following conditions were stated in the previous study[39] for patients to be included:

i) Patients had to be in the chronic infection phase with low CD4 T-cell count at therapy initiation (median CD4 T-cell count 184.5 cells/µL).

ii) Patients needed to present consistently undetectable virus load under therapy (measured quarterly) with a good CD4 T-cell response (ΔCD4 T-cells above 200 cells/µL in four years of therapy).

The first line regimen consisted of a combination of at least three antiretroviral drugs, either three nucleoside reverse transcriptase inhibitors (NRTIs) or two NRTI combined with one or two protease inhibitors, or with one non-nucleoside transcriptase inhibitor (NNRTI). None of the patients ever received R5-receptor antagonists for therapy.

The majority of the 16 included patients in this study were male (62.5%), of white ethnicity (93.75%) with an average age of 52.1 ± 7.6 years. Most patients were infected with HIV-1 subtype B (75%). Baseline characteristics for all patients and for both, the X4-individuals and the R5-individuals are summarized in Table 1.

Characteristics	All patients (n=16)	X4-patients (n=7)	R5-patients (n=9)
Sex			
Male	10 (62.5%)	5 (71.4%)	5 (55.6%)
Female	6 (37.5%)	2 (28.6%)	4 (44.4%)
Age, mean years ± SD	52.1 ± 7.6	53 ± 8.1	51.4 ± 7.2
Ethnicity			
White	15 (93.8%)	6 (85.7%)	9 (100%)
Black	1 (6.2%)	1 (14.3%)	0 (0.0%)
HIV transmission			
HET	8 (50.0%)	4 (57.1%)	4 (44.4%)
MSM	7 (43.8%)	3 (42.9%)	4 (44.4%)
IDU	1 (6.2%)	0 (0.0%)	1 (11.2%)
Subtype			
B	12 (75.0%)	6 (87.7%)	6 (66.7%)
A/AG	2 (12.5%)	1 (14.3%)	1 (11.1%)
C	1 (6.25%)	0 (0.0%)	1 (11.1%)
D	1 (6.25%)	0 (0.0%)	1 (11.1%)
Baseline HIV RNA load log10 copies/ mL	5.6 (3.8-6.4)	5.5 (5.3-6.4)	4.9 (3.8-6.1)
Baseline proviral load log10 copies/ 10^6 cells	2.4 (2.2-2.9)	2.3 (2.2-2.7)	2.5 (2.2-2.9)
ΔCD4	445 (243-659)	463 (243-659)	350 (252-634)
Baseline CD4	184.5 (7-290)	91 (7-242)	215 (12-290)
Baseline CD8	97 (165-1288)	652 (266-1288)	541 (165-1062)
CCR5-Δ32 genotype			
heterozygous	1 (6.25%)	0 (0.0%)	1 (11.1%)
wild type	15 (93.75%)	7 (100%)	8 (88.9%)

Table 1 Baseline characteristics. Data are presented as no. (%), median (min-max) values unless otherwise indicated. SD: standard deviation; MSM: men who have sex with men; HET: heterosexual; IDU: injecting drug users; PBMC: peripheral blood mononuclear cells.

6.2.2 Integration Site Analysis

The increase of relative frequencies of proviral X4-tropic HIV-1 variants in 7 patients after 4 years on therapy was observed to be caused by a single provirus variant

emerging during suppressive therapy, which eventually became solely responsible for the increase of X4 variants. In six out of seven patients the emerging variant had been detected as a minority already prior to therapy[39]. That observation led to the hypothesis that clonal expansion could be the reason for the outgrowth.

To look at this, an integration site analysis for 16 patients of all three sample timepoints (preT, post4, post10) was performed.

In this study, the integration site analysis protocol was adapted from Maldarelli et al.[40]. The method was optimized and validated using gDNA of two cell lines that contain known HIV-1 proviruses: SXR5 cells and Hut4-3 cells. SXR5 cells have been stably transfected with a defective HIV-plasmid (based on pNL4-3) lacking the entire env region (TK, unpublished); Hut4-3 cells are based on a human T4-lymphocyte line (Hut78), which emerged from an acute HIV-1 infection with pNL4-3, and has since been chronically producing infectious virus. For simulating the situation in patients, the infected cells were mixed at defined ratios ($1:10^2$, $1:10^4$, $1:10^6$) with uninfected 293T cells.

As the stable transfection of SXR5 was done with a pNL4-3 based plasmid, the integration sites for this cell line were known to be the 5'cellular flank sequence as well as the 3'cellular flank sequence of this plasmid (5'integration site: chr16:317155, 3'integration site: chr17:82208360). While the 3'integration site in SXR5 cells was confirmed for all three dilutions, the analysis of the 5'integration site failed due to unspecific primer binding. The 5'LTR primers were excluded in further experiments.

For Hut4-3 cells two clonal integration sites were determined, one on chromosome 4 and one on chromosome 7.

After establishing the protocol, integration site analysis was performed by using PBMCs from 16 HIV-infected individuals from three timepoints (preT, post4, popst10).

In total, 1386 different integration sites were mapped; 33 (2.4%) sites of these were associated with reads with different host DNA breakpoints, and therefore determined as clonal. A slight trend of clonal integration site accumulation over the years was seen: In preT samples, a total of 825 different integration sites were discovered, of which 10 (1.2%) sites were clonal sites. In post4 samples 374 unique integration sites and 12 (3.1%) clonal integration sites were mapped. In post10 samples 164 unique integration sites and 11 (6.2%) clonal sites were detected.

In X4 samples a total of 785 integration sites and in R5 samples a total of 601 integration sites were detected. Out of these 16 (2%), respectively 17 (2.8%) integration sites were clonal.

To identify functional associations of the genes that had clonal HIV integration sites in or nearby, gene ontology (GO) analysis was performed using GREAT software with the default settings[76]. An enrichment in genes involved in metanephros morphogenesis was observed, however there was no statistically significant result.

The hypothesized clonal outgrowth of X4 variants due to integration in genes involved in cell cycle regulation could not be confirmed.

6.2.3 Proviral Load

To check if the recovered integration sites were in the expected range, the proviral load for each sample was assessed. The recovery rate of integration sites of 5.6% was just within the expected range of 5 – 15%[77].

As expected in patients under continued ART, proviral loads declined over the years on therapy (Figure 2).

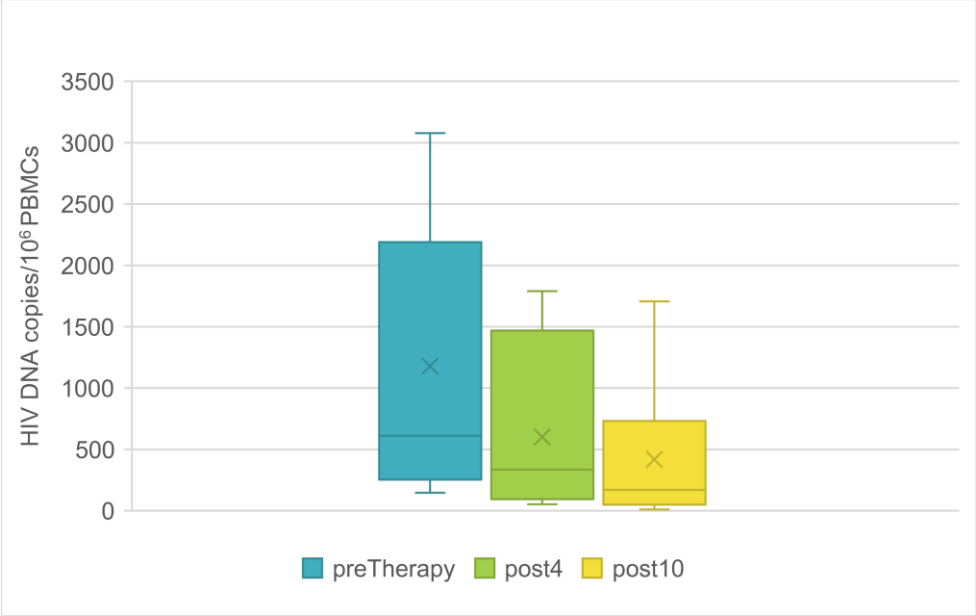

Figure 2 HIV proviral load in copies/10^6 PBMCs for the patient population before therapy (preT), 4 years on ART (post4) and 10 years on ART (post10).

6.2.4 Dynamics of Proviral Tropism under Therapy

To analyze the dynamic of proviral tropism in individuals under long-term suppressive ART, PBMC samples from 10 years after ART initiation were included in this study. The results were subsequently combined with data from Bader et al.[39].

The tropism determining V3 loop was sequenced and analyzed by Geno2Pheno454 to investigate the dynamics of proviral tropism in individuals under long-term suppressive ART.

For the 16 samples the median quality read size was 13861 reads per sample with a median variant count per sample of 119. At the standard false positive rate (FPR) of 3.5%, X4-tropic HIV-1 variants were identified in all samples except two (17637, 18249). These two samples were also classified as solely R5-tropic by applying a less stringent 10% FPR cut-off. Samples, which at FPR 3.5% had a relative abundance below 2% of X4 variants, were assigned as R5-patient.

In 8 out of 9 R5-patients at the standard FPR of 3.5% the relative frequency (%X4) of proviral X4-tropic HIV-1 variants remained stable on a very low level after being on therapy for 10 years (Figure 3).

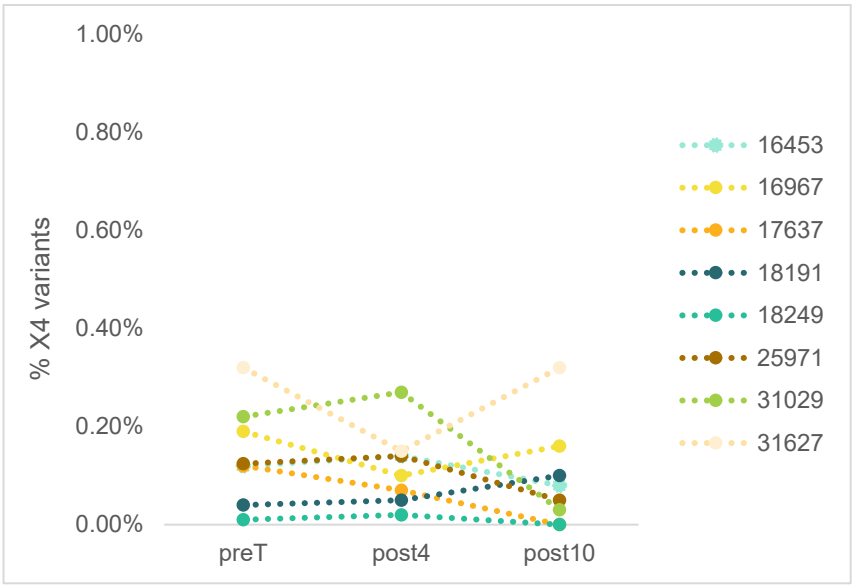

Figure 3 Frequency of X4-tropic variants in R5-patients before therapy start (preT), 4 years on ART (post4) and 10 years on ART (post10).

One R5-patient (25318, not shown) had a frequency of 17.18% X4-tropic variants after 10 years on therapy. At a 3.5% FPR cut-off this patient was an X4-patient at baseline (3.74% of X4 variants), but as the frequency of X4 variants decreased after 4 years on therapy to 0.91% of X4 variants, the patient was defined as R5-patient. Had an 5% FPR cut-off been applied, this patient would have been assigned as an X4-patient for all three timepoints (%X4: 89.87%, 98.07%, 85.51%).

In the 7 X4-patients, in whom an increase in relative frequencies of X4-tropic variants was observed after 4 years on therapy, after 10 years of therapy a fluctuation of the X4-tropic provirus frequency was observed. While 6 out of 7 X4-patients showed a drop in frequency of X4-tropic variants after 10 years on therapy, one (16579) showed an increase (Figure 4).

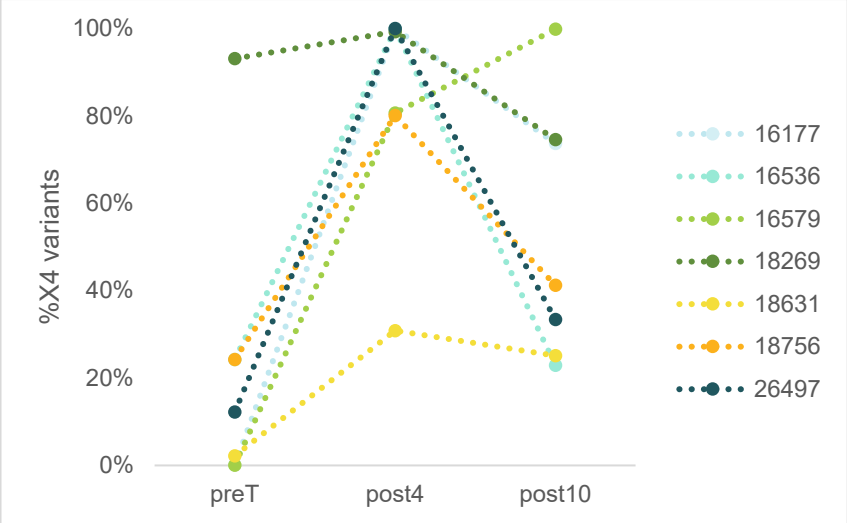

Figure 4 Frequency of X4-tropic variants in X4-patients before therapy start (preT), 4 years on ART (post4) and 10 years on ART (post10).

To exclude variability in sample preparation as the cause of fluctuation, the analysis was repeated for samples from 4 patients (16579, 18191, 25318, 25971). The results were reproducible, except for sample 25318 (Table 2). As already mentioned above, for sample 25318 a 5% FPR cut-off would produce a reproducible outcome.

Sample	variants for this MID	#quality reads	%X4 at 1% FPR	%X4 at 3.5% FPR	%X4 at 5% FPR	%X4 at 10% FPR
16579_preT	283	78055	0	0.06	0.11	0.47
16579_preT_rep	249	50018	0	0.05	0.07	0.49
16579_post4	251	30917	75.97	80.65	81.65	81.7
16579_post10	73	8318	72.45	99.84	99.84	99.84
16579_post10_rep	214	32188	97.60	99.99	99.99	100
18191_preT	285	32703	0	0.04	0.15	0.16
18191_preT_rep	296	39448	0	0.26	0.26	0.27
18191_post4	564	84505	0	0.05	0.15	0.16
18191_post10	214	25203	0	0.1	0.18	0.18
18191_post10_rep	276	41422	0	0.05	0.25	0.25
25318_preT	339	72987	2.68	3.58	89.87	98.93
25318_preT_rep	355	39440	88.71	89.11	89.25	99.93
25318_post4	273	44758	0.13	0.91	98.07	99.58
25318_post4_rep	158	18235	63.84	64.51	99.59	99.99
25318_post10	94	7281	51.15	51.87	85.51	99.89
25318_post10_rep	281	25761	71.91	73.43	73.49	99.77
25971_preT	204	34675	0	0.09	0.13	0.17
25971_preT_rep	325	37290	0	0.14	0.16	0.21
25971_post4	267	48162	0	0.14	0.22	0.4
25971_post4_rep	228	22591	0	0.14	0.15	0.18
25971_post10	151	20060	0	0.05	0.08	0.14

Table 2 Comparison of results to exclude variability in sample preparation. The results labeled in gray were produced by Bader et al.39. The results labeled in white were produced in this study.

6.2.5 Link of Clonal Integration Site to Main V3 Loop

To test if the main V3 variants could be linked to the detected clonal integration sites, for 4 patients with high proviral load, the env-integration site sequence was amplified with integration site-specific primers and subsequently analyzed by Sanger sequencing:

Patient	preT_IS_total	preT_IS_clonal	#integrates	post4_IS_total	post4_IS_clonal	#integrates	post10_IS_total	post10_IS_clonal	#integrates
31627	31	chr11:131126817	4	25	chr22:48631542	5	43	chr1:234143908	3
		chr4:19518716	2		chr11:131126817	2		chr2:102067563	3
25318	136	chr1:186586048	2	36	chr1:186586048	2	18	chr1:202308767	2
					chr11:66690921	2		chr6:22433638	2
					chr6:22433702	3			
16177	163	chr1:28690322	2	37	chr9:131602104	3	54	chr1:226567583	4
								chr19:11560702	12
								chr3:138037542	2
								chr3:140294057	2
18269	56	chr14:71587405	2	43	chr5:38066272	7	37	chr5:38066271	3
		chr17:31928177	6		chr8:142229856	2		chrX:43438671	6
		chr5:38066173	2		chr8:92719217	2			
		chrX:43438703	2						

Table 3 Overview of recovered integration sites in samples from 4 patients. IS_total gives the total number of recovered integration sites, IS_clonal shows the clonal integration site and #integrates shows how often this specific integration site was found.

<u>R5-patient 31627</u>

For patient 31627 five different clonal integration sites over all three timepoints were identified, of which one clonal integration site (chr11:131126817) was recovered in the preT sample, as well as in the post4 sample (Table 3). The PCR amplification of env-integration site(chr11:131126817) sequence was successful for the preT sample and the post4 sample. The amplified sequence was identical to a detected main (>2%) V3 loop, confirming the clonal expansion of the specific variant (Figure 5). Although this V3 loop was also detected as a main variant in the post10 sample, no integration site at chr11:131126817 was recovered and the PCR amplification of env-integration site(chr11:131126817) failed for this sample.

The PCR amplification of env-integration site(chr1:234143908) sequence was successful for the post10 sample, and the amplified sequence was in agreement with the major V3 loop in this sample.

The PCR amplification of env-integration site(chr1:234143908) sequence failed for the preT and post4 sample, although also in these samples the corresponding V3 loop was detected as a main variant.

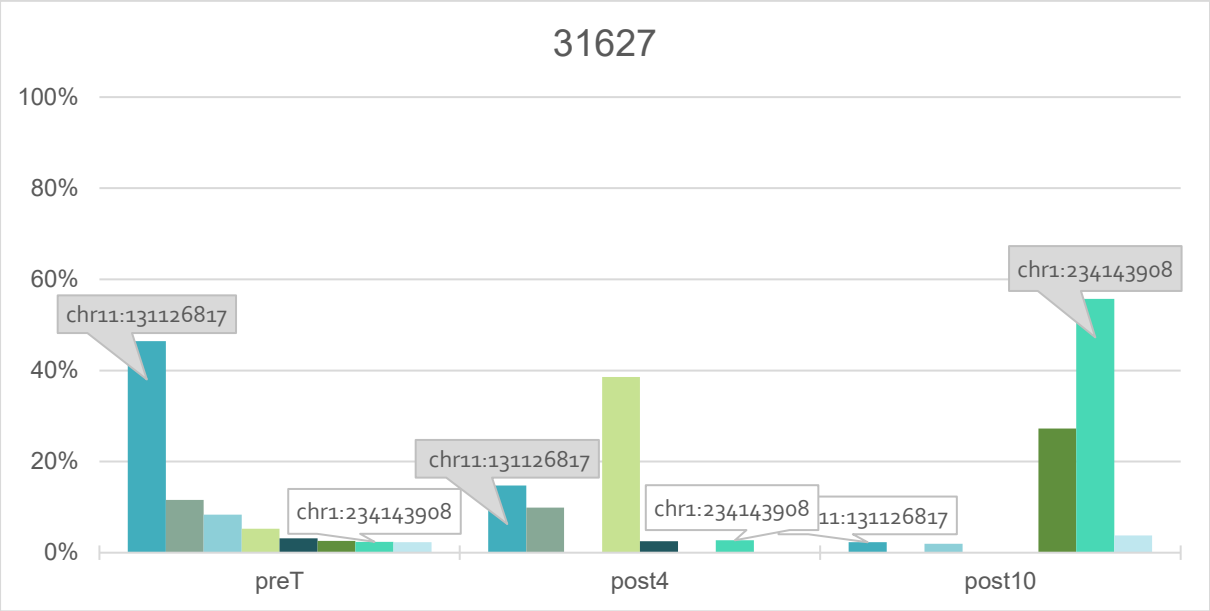

Figure 5 Main (>2%) V3 loops in R5-patient 31627 for all three timepoints. Bars with a blue/green color scheme are R5-tropic variants. It was possible to confirm a link between some V3 loops and specific chromosomal integration site by PCR amplification with integration site specific primers. Confirmed links are labeled in grey; identical V3 loops without confirmed link to the integration site are labeled in white.

R5-patient 25318

For patient 25318 five different clonal integration sites over all three timepoints were identified, of which one clonal integration site (chr1:186586048) was recovered in the preT sample, as well as in the post4 sample (Table 3). The PCR amplification of env-integration site (chr1:186586048) sequence was successful for the preT sample and the post4 sample. The amplified sequence was identical to the major V3 loop in both samples (Figure 6).

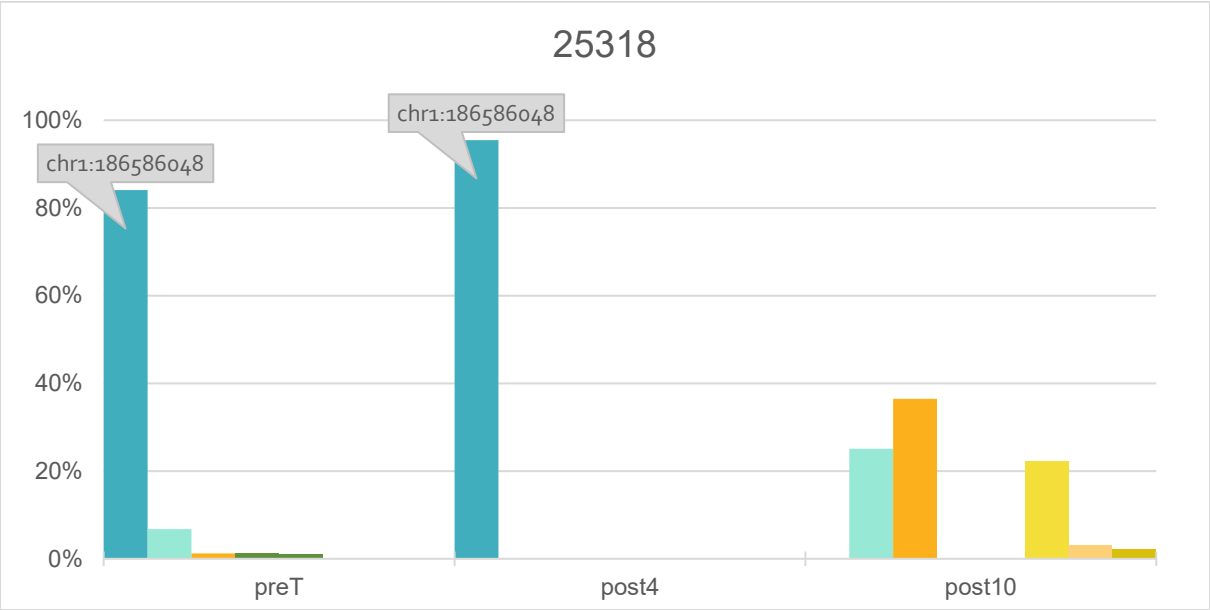

Figure 6 Main (>2%) V3 loops in R5-patient 25318 for all three timepoints. Bars with a blue/green color scheme are R5-tropic variants, bars with a red/yellow color scheme are X4-tropic variants. It was possible to confirm a link between some V3 loops and specific chromosomal integration site by PCR amplification with integration site specific primers. Confirmed links are labeled in grey; identical V3 loops without confirmed link to the integration site are labeled in white.

X4-Patient 18269

For patient 18269 seven different clonal integration sites were identified over all three timepoints, of which one clonal integration site (chr5:38066173) was recovered in the samples for all three timepoints and one clonal integration site (chrX:43438703) was recovered in the preT sample, as well as in the post10 sample (Table 3). Although there were identical major V3-loops detected in different timepoints (Figure 7), neither for integration site (chr5:38066173) nor integration site (chrX:43438703) was it possible to amplify the env-integration site sequence.

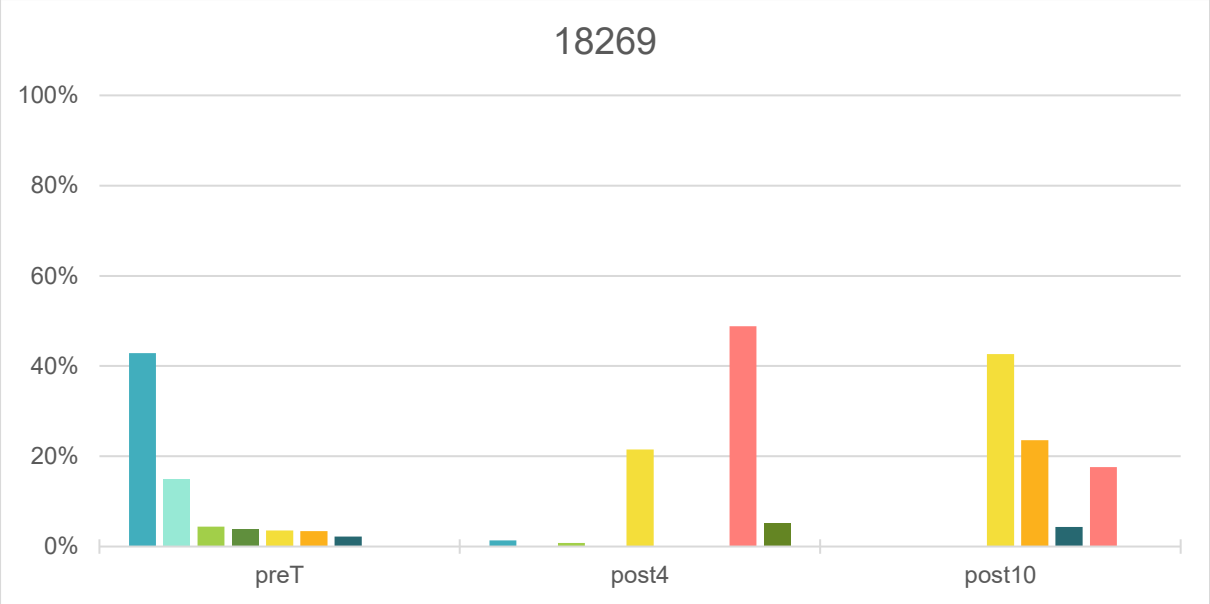

Figure 7 Main (>2%) V3 loops in X4-patient 18269 for all three timepoints. Bars with a blue/green color scheme are R5-tropic variants, bars with a red/yellow color scheme are X4-tropic variants. It was possible to confirm a link between some V3 loops and specific chromosomal integration site by PCR amplification with integration site specific primers. Confirmed links are labeled in grey; identical V3 loops without confirmed link to the integration site are labeled in white.

X4-patient 16177

For patient 16177 six different clonal integration sites over all three timepoints were identified. No clonal integration site was detected in more than one sample. It was possible to link the integration site (chr1:28690322) to the major V3 loop in the preT sample. The PCR amplification of env-integration site (chr19:11560702) was not successful. Although the amplification of the env-integration site (chr9:131602104) was successful for the post4 sample, the corresponding V3 sequence was not identical to any V3 loop. As only one V3 loop of the post4 sample was identical to V3 loops in the preT sample and none to V3 loops in the post10 sample, a sample mix-up in the original data set[39] was suspected. Therefore, the V3 loop of the post4 sample was reanalyzed by NGS (Table 4). Instead of one X4-tropic major variant, three main variants were detected: two R5-tropic variants and one X4-tropic variant. The frequency of X4-tropic of the total detected variants was reduced from 99.99% in the

previous post4 sample analysis to 35.31% in the repeated post4 sample analysis. With this result, patient 16177 is a second patient in which the frequency of X4-tropic variants increased after 10 years of therapy (see chapter 2.4).

Sample	# variants	#quality reads	%X4 at 1% FPR	%X4 at 3.5% FPR	%X4 at 5% FPR	%X4 at 10% FPR	Major V3 loop	FPR of major V3 loop	% of major V3 loop
16177_preT	251	70641	0	0.11	0.13	0.22	CTRPSNNTRKSIHIGPGRAFYATNIIGDIRQAHC	45.78	97.3
16177_post4	121	17931	99.7	99.99	99.99	100	CTRPNNSTRKGIYIGPGRAVYTGEKIIGDIRQAHC	0.58	97.6
16177_post4_rep	119	13909	35.13	35.31	35.31	35.35	CTRPNNNTRKSIHIGPGRAFYATDIIGNIRQAHC	42.59	34.8
16177_post10	125	17456	72.14	73.69	73.73	77.94	CTRPNYSTRRRIHIGPGRAFYATNIIGDIRRAHC	0.29	37.4

Table 4 Reevaluation of V3 analysis of patient 16177. Rows with gray shade indicated the original data set[39], white cells show data generated in this study.

Within the freshly detected V3 loops in the post4 sample, the amplified env-integration site(chr9:131602104) sequence was present (Figure 8).

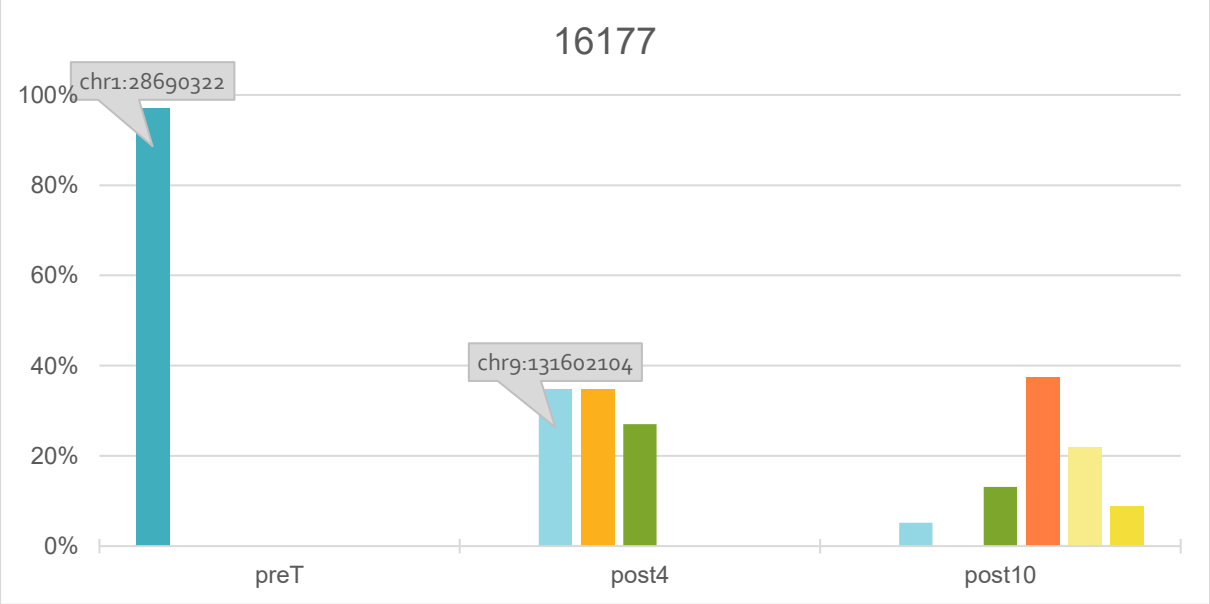

Figure 8 Main (>2%) V3 loops in X4-patient 16177 for all three timepoints. Bars with a blue/green color scheme are R5-tropic variants, bars with a red/yellow color scheme are X4-tropic variants. It was possible to confirm a link between some V3 loops and specific chromosomal integration site by PCR amplification with integration site specific primers. Confirmed links are labeled in grey; identical V3 loops without confirmed link to the integration site are labeled in white.

The successful link of clonal integration sites to the main V3 loops, confirms the clonal feature of these variants.

6.3 Discussion

Recent data of Bader et al.[39] showed in some individuals an outgrowth of a single pretreatment X4-tropic proviral minority after up to 4 years on therapy. The authors discussed that this outgrowth could have caused by clonal expansion. In this study, integration site analysis was used to test this hypothesis. In 10 patients (5 X4-patients, 5 R5-patients) a total of 1386 integration sites were recovered, of which 33 (2.4%) were clonal.

It is important to note that only a very small fraction of infected cells present in the whole body is sampled via peripheral blood, and only a fraction of the integration sites from these samples is recovered. Therefore, it is not possible to identify any integration site as truly “unique”. Due to the small sample size and the additional loss of material during the experimental procedure, it is quite likely, that the actual percentage of clonal integration sites is higher[77].

On the other hand, the combination of linker-mediated PCR and next generation sequencing produces large datasets with a high background noise, which may lead to the false identification of a “clonal integration sites”. Due to scarcity of integration events and the multiple steps of a somewhat error-prone amplification, this cannot be completely prevented.

To limit background noise, several common issues for procedures that involve high levels of PCR amplification have been addressed during the optimization of the experimental integration site analysis protocol over the course of this work. Cross contamination is one of the issues that can influence the outcome of the integration site analysis. Although we had designated rooms for pre-PCR and post-PCR experiments, in some samples we still detected a contamination with the 3’ integration site (chr17:82208360) of plasmid pNL4-3, which is greatly used in our laboratory. By

comparison of recovered integration sites between different patient samples, the possibility of inter-patient cross contamination was excluded.

Another cause of artifacts is PCR mispriming. As proviral DNA is rare and human gDNA is present in vast amounts, even a perfectly designed primer can bind to patient gDNA sequences that are similar to the target sequence. In the integration site analysis, this can lead to incorrectly identifying mispriming sites as integration sites. However, by designing the nested LTR primer to end 7 bp before the end of the HIV sequence, a great extent of misprimed sites could be excluded during bioinformatics analysis. The reads are filtered for those sequences, in which the sequence between the 3' end of the nested LTR primer and the adjacent host sequence exactly matches this 7 HIV bp. Interestingly, a longer end HIV sequence (35 bp instead of 7 bp) with the aim to decrease false amplifications, did not further decrease the background noise.

Further artifacts can also be caused by PCR recombination. If there is incomplete copying of a DNA fragment during the elongation step, the partial DNA product can act as a primer, which leads to recombination if the partial DNA product base pairs to a site elsewhere in the genome. PCR recombination can also be detected and filtered with a proper bioinformatics analysis pipeline.

For a final confirmation of true integration sites, the amplification of both, the 3'LTR-host adjunction and the 5'LTR-host adjunctions would be necessary. Unfortunately, our test experiments with the 5'LTR primer failed. It has previously been shown, that HIV-1 U3 tolerates much more sequence variation than the U5 sequence[78]. As the standard LTR primers were designed based on the reference sequence of HXB2, the possible variation in the U3 sequence could explain why PCRs with the 5'LTR primer

failed. The design and usage of patient-specific 5'LTR primers may improve the identification and confirmation of true integration sites.

While the experimental optimization in this work was recently also described by Wells et al.[77], the inclusion of an optimized bioinformatic analysis could further improve the recovery of true integration sites and therefore help to generate a more detailed insight in HIV-1 latent reservoirs.

By amplifying env – integration site sequences with integration site specific primers, the V3 regions belonging to clonal integration sites were analyzed. For one patient the env – integration site amplification for 2 different clonal integration sites failed. Further PCR optimization was not possible due to a limited amount of available patient DNA and due to the fact that integration site-specificity of primers made the use of an universal positive control impossible. However, for 3 out of 4 patients it was possible to link the major V3 variants to clonal integration sites, confirming the suggested clonal expansion.

As some studies have shown, that integration of HIV into cancer-related genes can drive clonal expansion[41,79], the question arose, if the observed X4 outgrowth could be caused by integration into a gene involved in cell-cycle regulation and therefore could be a cause for cancer development in these patients. So far there are no known cases of T cell malignancies in HIV patients. However, human T lymphotropic virus (HTLV), the first reported human oncogenic retrovirus[80,81], causes adult T cell lymphoma/leukemia (ATL) in around 5% of all HTLV-infected individuals. Several mechanisms are discussed to be involved in the development of ATL, one of the main ones is HTLV interaction with the host chromatin at the integration site[82–84]. It was shown, that over a time of 50-60 years, a clone of HTLV-infected cells grows out of the host immune control and leads to ATL[85,86]. Given the similarity between HIV-1 and

HTLV-1, the observation that the frequency of HIV-1 integration into cancer-related genes is significantly higher than the frequency of cancer-related genes in the human genome[41], and the fact that HIV-infected individuals have a normal life expectancy only since the introduction of ART in 1996 (which is less than 30 years ago), further examination of HIV integration and its effects on gene regulation are needed.

Interestingly, in our study we did not detect any clonal integration sites in or near to specific cancer-related genes, but an enrichment in genes involved in metanephros morphogenesis was observed. However, this enrichment was not statistically significant. Adding to the fact, that first results of integration site analysis did not show the expected trend of a slow and steady increase of clonal HIV-infected cells, as it would be expected with integration site-driven proliferation[87], we investigated the proviral tropism of patient samples 10 years after start of therapy. The idea behind this tropism testing was to check if the outgrown X4-variant, which was observed in patients with increasing frequency of X4-tropic HIV-1 variants, was still present after 10 years of therapy.

In 4 out of 7 X4-patients, in which an increase of the frequency of X4- variants was observed after 4 years on therapy, this frequency dropped again after 10 years of therapy. However, the frequency did not drop below to 2%, which was the assigned cut-off value. So, although there was a decrease of X4-tropic variants, no tropism switch was observed, and the patients were still defined as X4-patients. In two patients (16177, 16579) a further increase of the frequency of X4-tropic variants was observed and in one patient the frequency of X4-tropic variants stayed stable. Interestingly, in both patients with an increase in the frequency of X4-tropic variants, this increase was driven by the appearance of a new X4-tropic main variant. While in patient 16579 the major X4-tropic variant was the same from post4 to post10, in patient 16177 all main

X4-tropic variants from timepoint post4 disappeared at timepoint post10 and new X4-tropic variants were responsible for the increase. In general, there was a higher variation in the composition of the main variants in X4-patients than in R5-patients. This fluctuation of main variants and the lack of a steady increase of any main variant, respectively clonal integration site speaks against integration site-driven clonal expansion.

The fluctuation of main variants over time is in agreement with a study from Wang et al.[88], in which expanded proviral clones were found to wax and wane over time in individuals on ART. They suggest that proliferation of infected cells is balanced by a significant amount of cell loss. The dynamic changes support antigen and cytokines as potential drivers of clonal expansion.

The observed fluctuation of the frequency of X4-tropic variants is also in line with various contradictive studies on proviral tropism development under suppressive ART. Philpott et al.[38] claimed a preferential suppression of X4-tropic strains of HIV-1 by antiviral therapy. They showed that in 15 women with a predominant X4-tropic virus population at baseline, the initiation of ART led to a shift to R5-tropic strains. Multivariate analyses showed that the shift was independent of changes in plasma HIV-1 RNA level and $CD4^+$ cell count. These results agree with the observation of Bader et al.[39], that the majority of X4-patients experienced a decrease of X4-tropic variants.

An opposite observation was made by Delobel et al.[36]. They showed a switch from R5-tropic variants to X4-tropic variants in 11 of the 23 patients who harbored a majority virus population of R5 variants at baseline. X4 variants remained predominant in the 9 patients who harbored mainly X4 variants at baseline. As a possible explanation for this observation, it was suggested that potent antiretroviral therapy produces the

conditions necessary for the gradual emergence of X4 variants in cellular reservoirs. .

While Soulie et al.[37] did not observe any tropism switch in 34 patients being 48 weeks on ART, Saracino et al.[89] observed both, a switch from R5-tropic to X4-tropic, as well as a switch from X4-tropic to R5-tropic in 6, respectively 4 patients out of 36 patients being on ART for 12 months. Castagna et al.[90] also analyzed the dynamics of HIV-1 tropism in 195 patients under ART. Interestingly, although 124 patients presented with persistently detectable viral load, they showed similar rates of R5-X4-switches or X4-R5-switches as the patients with undetectable viral load.

Although some of the discrepancies within these studies could be explained by differences in the experimental setup, different patient cohorts, different study duration, they show that the dynamics of HIV-1 tropism in patients under ART does not yet have a well understood mechanism.

One possible explanation for the proviral tropism fluctuations has been proposed by Raymond et al.[91]. They observed an absence of genetic evolution in individuals infected with R5-tropic variants, while in individuals mainly infected with X4-tropic viruses, diversification in the V3 region was observed. Their proposed explanation for this observation was that R5-tropic proviruses may persist due to proliferation of latently infected cells, while the genetic evolution seen in X4 variants could be caused by residual virus replication despite ART, either by free virions or cell-to-cell spread[92]. However, studies on whether viral replication persists in tissues, such as lymph nodes and gut, to levels that can maintain the HIV reservoir are contradictory[93–96].

Another explanation for the proviral tropism fluctuation could be the different proliferation rates of different cell subpopulations. Circulating integrated HIV proviruses appear to be maintained both by slow turnover of immature CD4

subpopulations, and by clonal expansion as well as cell differentiation into effector cells with faster replacement rates[97].

It has been shown in a study by Zhou et al.[98] that X4-tropic virus variants and R5-tropic virus variants have very limited recombination products, indicating that these viruses might not be found inside the same cell subsets. While X4-tropic variants are mainly found in naïve (T_N) and central memory T cells (T_{CM}), which is consistent with the relatively high expression of CXCR4 on these cells, R5-tropic variants are predominately found in transitional memory T cells (T_{TM}) and effector memory T cells (T_{EM}), also consistent with the relatively high expression of CCR5 on these cells. It has also been indicated that both R5-tropic and X4-tropic variants can infect activated $CD4^+$ T cells, but only X4-tropic variants are able to enter resting $CD4^+$ T cells[99].

$CD4^+$ T cells have many levels of differentiation. The contribution of each differentiated subset to the functional latent reservoir has only recently begun to be studied. Recent findings suggest that cell subpopulations, such as follicular helper T (T_{fh}) cells[100] or naïve T (T_N) cells[101–103] have a greater contribution to the latent reservoir than previously thought, and a profound reshaping of the latent reservoir was suggested to be the cause for proviral tropism switches in patients under ART[104].

The recent development of single-cell approaches will improve our understanding of the composition of the HIV-1 reservoir and the influence of viral tropism on it. The identification of shared cellular markers and metabolic pathways involved in establishing and maintaining these reservoirs will help to find new strategies to cure HIV-1.

6.4 Conclusion & Outlook

This study followed the observation, that while under successful therapy a majority of HIV-1 infected individuals showed a decrease of CXCR4-tropic variants, in a few patients an increase of CXCR4-tropic viruses was observed. As the increase of CXCR4-tropic viruses in patients under therapy was observed to be caused by the outgrowth of a single variant, this study aimed to investigate the possibility of clonal expansion of HIV-1 infected cells by integration site analysis.

By using integration site analysis, this study confirmed clonal expansion of HIV-1 infected cells. However, the following observations speak against the involvement of the HIV-1 integration site in aberrant cell proliferation:

i) the clonal expansion was observed in both, R5-patients and X4-patients

ii) the clones waxed and waned over the observed period of time

Overall, these two main results support the hypothesis that clonal expansion of HIV-1 infected cells could be caused by antigen-driven proliferation rather than by integration site-driven proliferation.

To further investigate this conclusion an ongoing effort should be made to investigate tropism dynamics in the studied patients at even later timepoints.

In this study PBMCs from 7 X4-patients with an increasing frequency of X4-tropic provirus after 4 years on ART and from 9 R5-patients with a very low and stable frequency of X4-tropic viruses were analyzed. In the majority of X4-patients a decreasing frequency of X4-tropic variants was observed after 10 years on ART. Another interesting next step would be the analysis of tropism dynamics in the patients where a decreasing frequency of X4-tropic variants was observed after 4 years on ART[39]. The following two outcomes could be expected:

i) A further decrease of X4-tropic proviruses, or even a switch to being R5-patients. This would support the hypothesis that ongoing ART provides an active selection against X4-tropic provirus. The individual causes of the temporary increase in X4-tropic frequency would then make for an interesting study (e.g. was the patient suffering from an infection and weakened immune system shortly before the sample was taken?).

ii) Fluctuations in the frequency of X4-tropic variants, similar to those observed in this study. This would further support the hypothesis of an antigen-driven proliferation mechanism.

In general, the differences of X4-tropic and R5-tropic proviral variants and their cellular origin should be further investigated to gain a better understanding of the HIV-1 reservoir and also a better understanding of the variation of therapy outcome between individual patients.

7 Adapting the geno2pheno[coreceptor] tool to HIV-1 subtype CRF01_AE by phenotypic validation using clinical isolates from South-East Asia

7.1 Aim of the study

The determination of coreceptor usage became clinically important in diagnostic settings when a mechanistically new antiretroviral drug, the entry inhibitor MVC was licensed for treatment of HIV patients. As MVC specifically blocks the CCR5 coreceptor but not CXCR4, a treatment decision for MVC requires prior tropism determination in the blood of the respective patient[2].

Today, genotyping tools such as geno2pheno[coreceptor][12] with their specific algorithms are available as standard methods for viral tropism determination.

The geno2pheno[coreceptor] tool has been developed mostly based on subtype B viruses, and was until recently mainly been used for this subtype[105]. Herein, the algorithm demonstrated an excellent agreement between genotypic and phenotypic methods.

However, in a recent study[106], it was shown that there are significant discrepancies between geno- and phenotyping in other subtypes, i.e. CRF01_AE, where geno2pheno[coreceptor] predicted an excessive number of X4-tropic envelopes. The HIV-1 subtype CRF01_AE, predominantly circulating in South-East Asia, is among those subtypes diverging the most from European subtype B viruses. It has been suggested that patients infected with subtype CRF01_AE may have a more rapid decline of $CD4^+$ T cell count compared with patients infected with subtype B virus, as well as a shorter time to needing antiretroviral therapy and a higher virulence during the course of infection [107,108].

As geno2pheno[coreceptor] has been used in several recent studies performed in South-East Asia[107,109], a thorough examination and, if needed, a correction of the geno2pheno tool for the genotypic prediction of CRF01_AE coreceptor-usage is urgently needed.

The aim of this study was to provide the necessary verification and to provide a basis for adjustments of the geno2pheno tool for CRF01_AE in diagnostic settings.

7.2 Results

7.2.1 Sample Characteristics

Twenty patient-derived env (gp120) HIV-1 CRF01_AE samples from a cohort in Thailand were used for simultaneous phenotyping and genotyping. The samples were randomly chosen from 144 CRF01_AE plasma samples available through the Thailand's National HIV Drug Resistance Surveillance Program from a study among female sex workers[110]; informed consent and ethical approval from the responsible IHRP have been obtained (approval 3/2557).

7.2.2 Construction of CRF01_AE Cassette

The viral gp120 region in patient-derived samples was amplified by RT-PCR and cloned into a pNL4-3 cassette (pNL-K7)[111], where it reconstituted fully functional HIV-1 genomes. Only a very low viral infection rate was obtained in cell culture by the HIV-1 genome reconstitution inserting exclusively the gp120 region from CRF01_AE samples into the NL4-3 background. As strategy for improving viral competence, the replication properties of a whole array of recombinant HIV-1 clones, carrying various genomic segments of CRF01_AE-origin were compared side-by-side in the backbone of a prototypic subtype B virus (NL4-3). After each cloning step (initially only the entire env gene, then env plus vpu, then vpu plus nef and eventually the entire region from

gag to env), replication of the resulting viral subtype B/AE recombinants were analyzed. As final result, the HIV-1 genome from the BssHII site at nt 712 to NgoMIV at nt 8338 (pNL-AE-K7_short) or to XmaI at nt 8888 (pNL-AE-K7) was substituted in frame by patient-derived CRF01_AE sequences, retaining only the LTRs and the 3' end of *nef* of NL4-3. The insertion of the respective gp120 sequences from clinical specimens into pNL-AE-K7_short allowed to phenotypically re-assess the tropism of the respective patient-derived viral envelopes.

7.2.3 Phenotyping and Genotyping

Using the new pNL-AE-K7_short cassette, we were able to determine the phenotype in 19 clinical samples (Table 5, column "Phenotype") by judging drug-based inhibition of viral replication and potential syncytia formation in the presence of either the R5-antagonist MVC or the X4-antagonist AMD3100. In this assessment using a virus-replication system, only one sample (Th026) was found to be X4-tropic, while 18 samples were determined to contain R5-tropic virus. For one sample (Th049), no clear tropism determination was possible, since small fusion events of 2-3 HIV-infected cells had formed in the cultures in the presence of either inhibitor. For this case, the presence of a dual-tropic virus could not be excluded.

Noteworthy, for all tested B/AE-recombinants the average syncytium sizes and the overall number of viral infection events in the culture dish remained low (approximately 10% of the control) when compared to the control plasmid pNL-NF.

Sample	V3-LOOP	Phenotype	g2p: 10% FPR	g2p: 5% FPR	g2p: 2.5% FPR	g2p: 1% FPR
TH026	CTRPSKTVR-SMRIGPGKVFYRIEGIGGDIRKAYC	X4	0.3	0.3	0.3	0.3
TH012	CTRPSNNIRTSMTVGPGQVIYKTGDITGNIRKARC	R5	1.7	1.7	1.7	1.7
TH016	CTRPSNYTRTSTRIGPGQVWYRTGDIIGNPRKAYC	R5	1.7	1.7	1.7	1.7
TH034	CTRPSNNTRTSIGIGPGQVFYRTGDIIGDIRRAYC	R5	2.6	2.6	2.6	2.6
TH010	CTRPSNNIRTSVHIGPGQVFYKTGDIIGDIRQAHC	R5	4.7	4.7	4.7	4.7
TH041	CTRPSTNRRTGVHIGPGQVFYRTGEIIGDIRKAYC	R5	7.2	7.2	7.2	7.2
TH023	CTRPSNNTRTSSRIGPGAVFYRTGDITGDIRQAHC	R5	8.7	8.7	8.7	8.7
TH020	CTRPSNNTRTSVTMGPGHVFYRTGDIIGDIRKAHC	R5	15.5	15.5	15.5	15.5
TH043	CTRPSNNTRTSMTIGPGQVFYRTGDIIGDIRKAYC	R5	16.6	16.6	16.6	16.6
TH037	CTRPSNNTRTS-HIGPGQVFYRTGDIIGDIRKAHC	R5	17.8	17.8	17.8	17.8
TH040	CIRPSNNTRTSIPIGPGQVFYKTGEIIGDIRKAYC	R5	20.7	20.7	20.7	20.7
TH028	CTRPFNNTRTSITIGPGQMFYRTGEIIGDIRKAYC	R5	24.9	24.9	24.9	24.9
TH038	CTRPSNNTRTSITIGPGQVFYRTGDITGDIRKAFC	R5	25.4	25.4	25.4	25.4
TH005	CTRPSNNTRTSITIGPGQVFYRTGDIIGDIRKAYC	R5	33.1	33.1	33.1	33.1
TH046	CTRPSNNTRKGIHLGPGQVFYRTGDIIGDIRQAYC	R5	35.7	35.7	35.7	35.7
TH022	CTRPSNNTRQSINIGPGRVFYRPGDIIGDIRKAYC	R5	44.9	44.9	44.9	44.9
TH039	CTRPSNNTRTSVHIGPGQVFYRTGDIIGDIRKAYC	R5	44.9	44.9	44.9	44.9
TH049	CTRPSNNTRTSIHMGPGQVFYRTGEIIGDIRQAHC	ND	50.9	50.9	50.9	50.9
TH032	CTRPSNNTRTSITMGPGQVLYRTGDITGDIRKAYC	R5	63.6	63.6	63.6	63.6
TH035	CTRPSNNTRKSVPIGPGQVFYRTGDIIGDIRQAHC	R5	89.6	89.6	89.6	89.6

X4
R5

Table 5 The phenotyped CRF01_AE samples with their confirmed sequence (V3-loop) and the genotypically predicted respective tropism. Blue = X4-tropic, green = R5-tropic, ND = not determined.

In parallel, the most prevalent genotype present in these 20 clinical samples was predicted using the standard version of geno2pheno[coreceptor] (Table 5, columns "g2p"). When linking these results to the phenotypic findings, the suspected systematic overcalling of X4-tropism in subtype CRF01_AE by the current version of Geno2Pheno[coreceptor] became apparent, reaching only a low assay specificity of 66% when the standard FPR cut-off of 10%, was used. By lowering the FPR cut-off to 2.5% the specificity increased to 89%.

7.2.4 HIV-GRADE

For confirmation beyond the small initial data set from Thailand, the newly suggested CRF01_AE-specific FPR cut-off of 2.5% was re-applied to a large data set from a German HIV-GRADE cohort on CRF01_AE samples (Table 6). When applying this new rule to all available CRF01_AE isolates, the significant discrepancy in the X4/R5 tropism ratio for CRF01_AE isolates, as depicted in Figure 9, completely disappeared

and rendered this subtype similar to the general, subtype independent distribution of clinical samples.

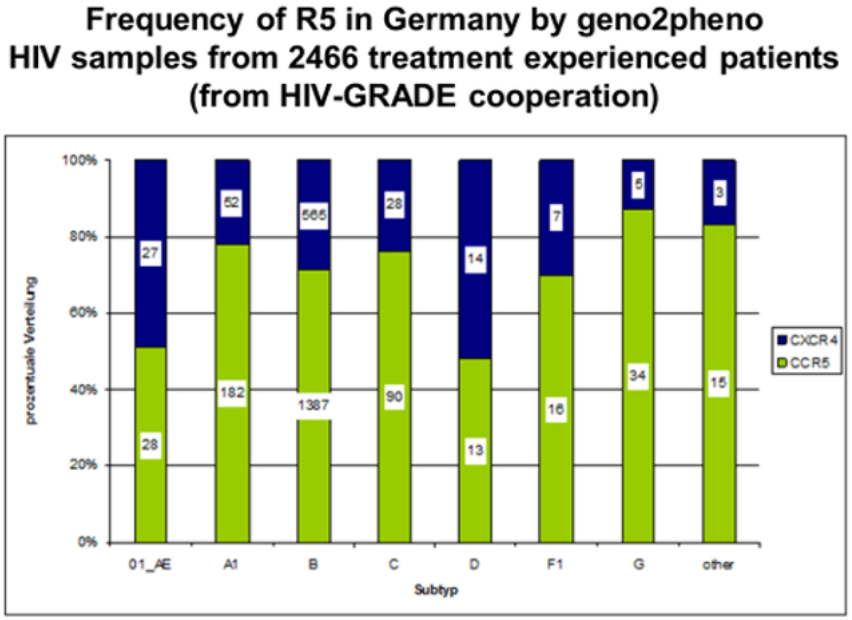

Figure 9 Frequency of R5- and X4-tropism by geno2pheno using the standard FPR cut-off of 10%[112] in 2466 treatment-experienced patients, for whom tropism testing was performed at baseline prior to potential maraviroc administration. Green = R5, Blue = X4

When disregarding differences between subtypes, the overall tropism distribution across all isolates (including all subtypes) would be 72% R5 and 28% X4. When the 10% FPR cut-off was applied specifically to subtype CRF01_AE isolates, this ratio shifts to 51% R5 and 49% X4, indicating a dramatic deviation with a Chi^2 of <0.001. When we now apply the phenotype-supported new CRF01_AE-FPR cut-off of 2.5%, a 76% R5 and 24% X4 distribution is seen for the CRF01_AE isolates with a Chi^2 value of 0.43, which is no longer significantly different from the calculated global average of isolates irrespective of their subtype.

For verification and potential fine-tuning, our phenotype-matched geno2pheno values were also subjected to the FPR cut-offs of 1% and 5% as well as to the FPR cut-off of 3.75%, which is used for NGS data. No significant difference from the expected distribution was seen for cut-offs at and above 2.5%.

FPR-cutoff	CCR5	CXCR4	Chi²
Expected normalized distribution	*72%*	*28%*	
10%	51%	49%	<0.001
5%	64%	36%	0,22
3.75%	69%	30%	0,68
2.50%	76%	24%	0,43
1%	89%	11%	0,002

Table 6 The frequency of X4 in the patients with subtype CRF01_AE in HIV-GRADE was likely to be overstated with an FPR cut-off of 10% in comparison to the frequencies of the other subtypes. Lowering the FPR to the cut-off of 2.5% as phenotypically determined in this analysis, the subtype AE-specific polymorphisms were correctly accommodated. The relevant FPR range between 2.5% and 5% (Chi^2 >0.2) has been shaded.

7.3 Discussion

In this study, phenotypically determined co-receptor usage was compared to and combined with genotypic data to improve the prediction of geno2pheno[coreceptor] for subtype CRF01_AE isolates of HIV-1.

For the phenotypic determination of the co-receptor usage major challenges had to be overcome. A recombinant plasmid-based system (pNL-K7), previously developed by our group[111], was used to reconstitute HIV-1 variants. This cassette permits the exchange of env segments by cleavage with unique restriction endonucleases and placing PCR-amplified HIV-1 env derived from patient plasma directly into a complete viral genome. After transfection into a human indicator cell line, viral replication of the recombinant HIV-1 variant in the presence of inhibitors can be quantitatively analyzed[113]. One hurdle in this process was a poor PCR amplification rate of the env fragments. Standard PCR-primers were derived from a reliably working subtype B consensus sequence. The observed low amplification rates were a strong indicator for the vast sequence heterogeneity of our HIV-1 isolates in the viral env region, suggesting that the validated recombination protocol at predefined sites in Env may not be optimal for generating replicating viral subtype CRF01_AE genomes. Another technical hurdle was the low replicative fitness of recombinant HIV-1 genomes encountered when using the in-house subtype B-based HIV-1 cassette (pNL-K7). We attributed the poor replicative capacity to previously reported observations that Env may critically depend on interactions with subtype-matched corresponding regions in Gag-Pol[114,115]. It is further possible that Env functions best in a subtype-unique context including its own co-evolved Vpu[116] or other viral proteins[113,117]. To improve the replicative fitness, we therefore designed a new cassette, carrying a near full-length subtype CRF01_AE backbone. With this construct we were able to obtain sufficiently

replicative virus to phenotypically determine the tropism of CRF01_AE patient samples.

Based on our comparison between phenotypically and genotypically determined tropism, our study supports implementing a significantly lower FPR cut-off of 2.5% (compared to the standard of 10%) as a critical adjustment for appropriate tropism prediction for CRF01_AE samples by geno2pheno[coreceptor]. The suggestion to lower the FPR cut-off for CRF01_AE virus variants is supported by others: One study had shown for CRF01_AE samples a specificity of only 50% at a 10% FPR cut-off, whereas the specificity increased to 77% by lowering the FPR cut-off to 5%[118]. Another study concluded in a comparison of different genotypic tools that for clinical practice, a geno2pheno[coreceptor] FPR cutoff of 5% could be used to predict CRF01_AE tropism[119].

Also, for genotyping of clinical samples using deep V3 sequencing (NGS), the interpretation of the analysis combines the information on FPR on each of the sequences and the corresponding frequency of these different variants in the sample. Currently, this two-dimensional cutoff predicts a sample with >2% of the variants with an FPR <3.5% as not suitable for maraviroc treatment. This recommendation is so far independent of the HIV-1 subtype[120]. We re-adjusted the FPR-value for subtype CRF01_AE in order to improve the clinical application of geno2pheno[coreceptor] also for the use for consensus, Sanger-sequencing.

As the R5-antagonist MVC proved to be a well-tolerated drug, lowering the FPR cut-off would potentially allow for more patients benefitting from MVC administration, especially in South-East Asia. Therefore, taking our results into consideration, we suggest setting an FPR cut-off of 2.5% for the tropism prediction of clinical subtype

CRF01_AE samples by geno2pheno[coreceptor]. However, further studies on a larger cohort are needed to verify this suggestion.

One proposed reason for the observed X4-overcall using the standard FPR cut-off of 10% is a difference in common sequence motifs. The typical CRF01_AE envelope contains several otherwise uncommon amino acids in the V3 region as unique and inherent feature. The motif GPGQVF at the tip of the V3 loop occurred very frequently in the HIV-1 CRF01_AE samples of this study (Figure 10).

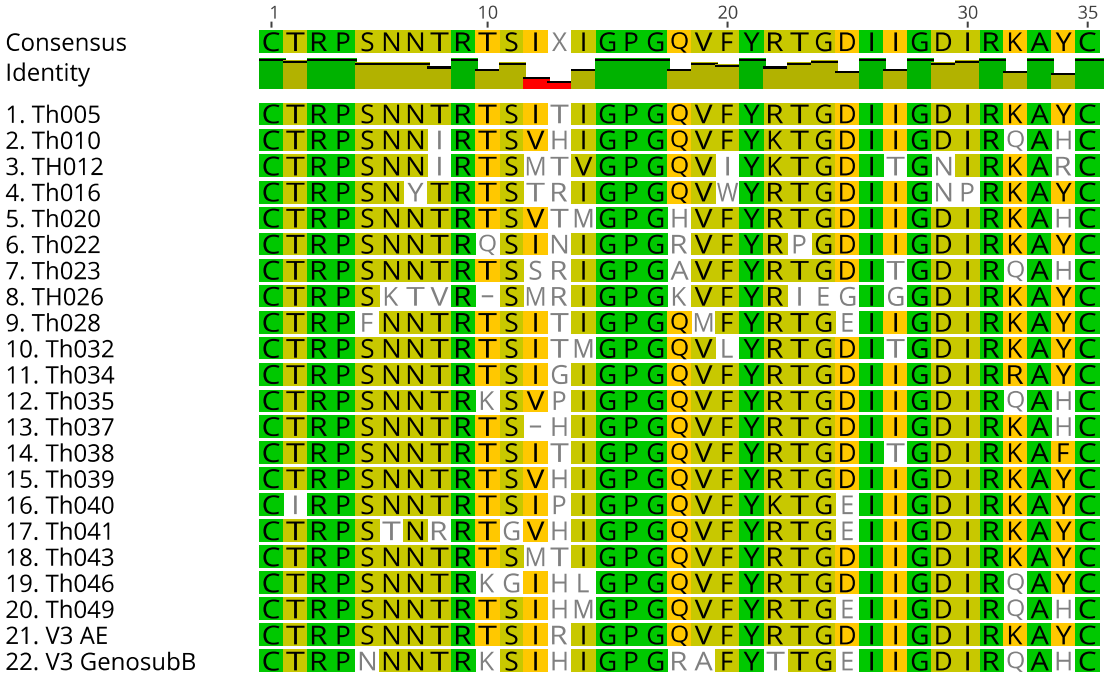

Figure 10 Comparison of CRF01_AE patient V3 sequence (Th0xx samples), the reference sequence of CRF01_AE (V3 AE) and the reference sequence of subtype B used for the training of geno2pheno[coreceptor][121] (V3 GenosubB). Green = 100% similar, olive = 80%-100% similar, yellow = 60%-80% similar, white = less than 60% similar.

Using the standard version of geno2pheno[coreceptor], these samples were often predicted as X4-tropic HIV-1 isolates in contrast to samples with the GPGRAF motif being frequent in R5-variants of subtype B. This significant deviation raised the speculation that the GPGQVF motif alone might result in incorrect X4 predictions. However, when taking the raw data of the prediction system into account, also

additional minor changes outside this very tip region contribute to the X4-overcalling of subtype CRF01_AE variants.

For further improvement of geno2pheno[coreceptor], more information about specific V3-loop characteristics of subtype CRF01_AE variants will be incorporated in future versions of the algorithm.

Although the results are clear within experimental setting, the low number of X4-tropic samples, identified within the study, poses a limitation. Even by including the findings reported in Matsuda et al.[106] to our data, only two samples were phenotypically identified as X4-tropic, which reflects an overall percentage of 4.7% (2 out of 42 samples). Although a relatively high proportion of R5-tropic isolates is common for HIV-1 in ART treated individuals, consistent with results reported by Cui et al.[122], a validation with more X4-tropic samples is recommended to corroborate our results.

It should further be noted, although gender is not currently known to play a role in the tropism distribution, that all specimens of this study came from female sex workers in a cohort in Thailand.

7.4 Conclusion

By combining genotypic data with phenotypically determined results, this study demonstrates the previously suspected systematic overcalling of X4-tropism in subtype CRF01_AE by the current version of the geno2pheno[coreceptor] tool. A suitable solution was obtained by adjusting the FPR cut-off for CRF01_AE samples to 2.5%, This results in a correct tropism prediction for clinical CRF01_AE samples and may guide the safe beneficial use of MVC in a broader group of HIV-1 infected individuals, who carry virus of this subtype.

8 Material & Methods

8.1 Material

8.1.1 Reagents

NAME	SUPPLIER
gDNA ISOLATION	
Maxwell® RSC Cultured Cells DNA Kit	Promega
LIBRARY PREPARATION	
NEBNext Ultra II FS DNA Library Prep Kit for Illumina	New England Biolabs
NEBNext Ultra II End Repair/ dA-Tailing Module	New England Biolabs
NEBNext Ultra II Ligation Module	New England Biolabs
SPRIselect	Beckman Coulter
STANDARD PCR	
Herculase II Fusion DNA Polymerase	Agilent Technologies
DNA PURIFICATION	
NucleoSpin® Gel and PCR clean-up	Macherey-Nagel
ExoProStar 1-step	GE Healthcare
AGAROSE GEL ELECTROPHORESIS	
Agarose	BioConcept
TAE Buffer 10x	Biosolve
Gel Loading Dye Purple (6x)	New England Biolabs
100 bp DNA Ladder (500 µg/mL)	New England Biolabs
1 kb DNA Ladder (500 µg/mL)	New England Biolabs
SYBR™ Safe DNA Gel Stain	Thermo Fisher
qPCR	
LUNA® Universal Probe qPCR Master mix	New England Biolabs
GENERAL	
All primers used	Microsynth
All enzymes used	New England Biolabs

Table 7 Reagents used with respective supplier.

8.1.2 Instruments

NAME	SUPPLIER
THERMO-CYCLER	
Biometra Trio 48 PCR Cycler	Biometra
7500 Fast Real Time PCR	Applied Biosystems
NUCLEIC ACID EXTRACTION	

Maxwell® RSC Instrument	Promega
NUCLEIC ACID QUANTIFICATION	
NanoDrop 1000	Thermo Scientific
SONICATOR	
Q700 Sonicator	Qsonica
CENTRIFUGE	
Centrifuge 5417C	Eppendorf

Table 8 Instruments/machinery used with their respective supplier.

8.1.3 Primer and Probes

NAME	SEQUENCE 5' – 3'
INTEGRATION SITE ANALYSIS	
Linker (+)	GTA ATA CGA CTC ACT ATA GGG CTC CGC TTA AGG GAC T
Linker (-)	PO4-GTC CCT TAA GCG GAG-C6
3LTR	TGT GAC TCT GGT AAC TAG AGA TCC CTC
3LTRmore	CTT AAG CCT CAA TAA AGC TTG CCT TGA G
5LTR	TCA GGG AAG TAG CCT TGT GTG TGG T
Linker	GTA ATA CGA CTC ACT ATA GGG CTC C
3'LTRnest	CCC TTT TAG TCA GTG TGG AAA ATC
5'LTRnest	CAC TGT TGT CTT TTC TGG GAG TGA ACT AGC C
Linkernest	AGG GCT CCG CTT AAG GGA C
V3 AMPLIFICATION	
F-6553 (1st PCR)	ATG GGA TCA AAG CCT AAA GCC ATG TG
R-7801 (1st PCR)	AGT GCT TCC TGC TGC TCC CAA GAA CCC AAG
F-6848 (2nd PCR)	CCA ATT CCC ATA CAT TAT TGT GCC CCG GCT GG
R-7371 (2nd PCR)	AGT TAC AGT AGA AAA ATT CCC CTC CAC AAT TAA A
D-6846 (Sequencing)	TGT TAA ATG GCA GTC TAG CA
R-7365 (Sequencing)	AGT AGA AAA ATT CYC CTC YAC AAT TAA A
D-6991 (Sequencing)	AGG CCT GTC CAA AGG TAT CCT TTG A
INTEGRATION SITE – V3 LINK	
F-LAM	CCC CRG CTG GTT TTG CGA TTC TAA AGT GTA
F-ISAenv1	ACA GTA CAA TGT ACA CAT GGA ATT A
F-ISAenv2	TGG AAT TAR GCC AGT AGT ATC AAC TCA
16177-3'LTR_chr1_out	GAC TTT AGC TTC CTT GGT TGA GT
16177-3'LTR_chr1	TGT GCC TTA TGT TTT TCT CCC AT
16177-3'LTR_chr9_out	GAA GGA AGA GTT TCA ACT TGA AAT A
16177-3'LTR_chr9	AGG GAC GCC GCT GCT AGA
16177-3'LTR_chr19_out	TTG CAT TAG AGT GTT CAG GAA GAT A
16177-3'LTR_chr19	TAA TCA AGT CAG GGG CTG TGC TAG A
18269-3'LTR_chr17_out	TCA GAA ACT CAC TAG ACA GTG A

18269-3'LTR_chr17	AGC TTT CTT TTG CAG TTA TGC TAG A
18269-3'LTR_chr5_out	TTT CAC CCT CAA AGC CAC CCT A
18269-3'LTR_chr5	ACA TGA AAA AGA GGC CTA ACT TTC AA
25318-3'LTR_chr1_out	TGT AAG CAC CAT CCG ATA CAG T
25318-3'LTR_chr1	**ACC TCC ACC TCC GGG TGC TAG A**
31627-3'LTR_chr1_out	TTG CAG ACA TTG CTA TGA GTT GCT
31627-3'LTR_chr1	GAC TTC ATT TCC TTT AGG ATA AAT A
31627-3'LTR_chr11_out	TCT TTT TAT GCC TTT TGA CAT GAA GT
31627-3'LTR_chr11	TAA TGT AAA CGT AGC CAC TGC TAG A
CRF01_AE CLONING	
KVL008	GGT CAK GGR GTC TCC ATA GAA TGG A
KVL009	GCC AAT CAG GGA AGW AGC CTT GTG T
F-6435alt	CYA CCA ACG CGT GTG TAC CCA C
R-8319NaeI	TGA RTA TCC CTG CCG GCC TCT ATT YAY TAT AGA AA
F-707	TGA AGC GCG CAC RGC AAG A
F-3475-AgeI_AE	AAA AAC CAC CGG TGC ATG GRG TAT A
R-3499-AgeI_AE	ACY CCA TGC ACC GGT GGT TTT TAG AA
F-6963-XmaI_AE	GAA TTA AGC CCG GGG TAT
R-8169-Nhe_AE	TAA TTT GCT AGC TAC CTG TTT TAA ARY TTT A
qPCR	
F-522	GCC TCA ATA AAG CTT GCC TTG A
R-643	GGG CGC CAC TGC TAG AGA
F-CCR5	ATG ATT CCT GGG AGA GAC GC
R-CCR5	AGC CAG GAC GGT CAC CTT
HIV LTR Probe	FAM-CCA GAG TCA CAC AAC AGA CGG GCA CA-BHQ
CCR5 Probe	VIC-AAC ACA GCC ACC ACC CAA GTG ATC A-BHQ

Table 9 Name and Sequence of Primers and Probes. Nomenclature: F = forward primer, R = reverse primer, followed by the number referring to the position in the HIV-1 genome (pNL4-3). All primers and probes were supplied by Microsynth AG.

8.2 Methods

8.2.1 Integration Site Analysis

Descendants of an HIV-infected cell will inherit a copy of a provirus integrated at exactly the same specific site in the host genome as in the parental cell.

8.2.1.1 Isolation of genomic DNA

Genomic DNA (gDNA) was isolated from 2.5 – 5x10^6 PBMCs using the Maxwell® RSC Cultured Cells DNA Kit (Promega) according to the manufacturers protocol.

8.2.1.2 DNA Fragmentation and End Repair

2 µg of gDNA were sheared into into 150 bp-500 bp random fragments. By randomly shearing gDNA, the influence of PCR bias on the integration site results is reduced, as only results with the same integration site but different shearing site (at least 3bp apart) were counted as clonal. 2 µg of gDNA is equivalent to the amount of DNA in approximately 330'000 cells. Because the number of provirus in patient samples is low, this may yield a low number of detected integration site. If enough material was available, the experiment was performed in triplicates.

For the random fragmentation of gDNA two different methods were used:

A) *Sonication* was performed with Q700 Sonicator (Qsonica). In a 0.5 mL Eppendorf tube 2 µg of gDNA were filled up to an end volume of 100 µL with H_2O. The samples were sonicated at 40% amplitude for 7 min with a 15 s pulse on/ off setting. The sheared DNA fragments were end-repaired and a single dA was added to the 3' ends using the EndRepair Module (New England Biolabs).

B) *Enzymatic fragmentation* was performed using the NEB Next Ultra II FS DNA Library Prep Kit for Illumina (New England Biolabs). The dsDNA fragmentase is designed specifically to produce random fragments that are appropriate for next generation sequencing. By combining random DNA fragmentation, end repair and dA-tailing in one reaction, the chances of contamination are reduced. The samples were prepared in duplicates. The reaction mix with 1 µg of gDNA

was prepared according to the manufacturers protocol with an incubation step of 37°C for 20 min, with expected fragment sizes of ~300bp.

8.2.1.3 Linker Preparation, Linker Ligation

A partially double stranded linker with a one nucleotide 3' T overhang was prepared by annealing two linker strands. The 5'end of the shorter strand (Linker(-)) was phosphorylated to enable an efficient linker ligation and the 3'end had a six-carbon glycol modification to improve PCR specificity by blocking the extension of the short strand. To anneal Linker(+) and Linker(-), 20 µL of each linker were mixed and incubated at 95°C for 5 min in a thermo cycler. The mixture was cooled down gradually by decreasing the temperature by 1°C every 2 min until the temperature reached 21°C. To ligate the linker to gDNA, the Ultra II Ligation kit (New England Biolabs) was used, following the manufacturer's instructions.

The 3′dA overhang of gDNA fragments and the 3′T overhang of the linker increase the efficiency of gDNA-linker ligation and reduce gDNA-gDNA and linker-linker ligations.

8.2.1.4 Purification / Size Selection

After the ligation step a size selection was performed using size selection SPRIbeads (Beckman Coulter) to remove excess linkers, small (<100bp) and big (>500 bp) DNA fragments.

After the purification step the triplicates of each sample were pooled and in the following steps the samples were prepared in triplicates again unless otherwise stated.

8.2.1.5 1st PCR

As cross contamination is a common issue for procedures that involve high levels of PCR amplification and can falsify the results of integration site analysis, good

laboratory practice is important. To avoid cross contamination rooms for pre-PCR and post-PCR experiments were designated and samples from the same patient but different timepoint were prepared separately.

An initial PCR step was performed to selectively amplify the host-virus integration junctions using one primer that matches the LTR sequence of HIV-1 (3LTR or 5LTR) and a second primer that matches the single stranded portion of the linker (Linker). The reaction was performed in triplicates with Herculase II Fusion DNA Polymerase (Agilent) in a volume of 50 µL. The master mix contained 31.5 µL of MilliQ H_2O, 10 µL of Herculase II Fusion DNA Reaction Buffer, 1.25 µL of 3LTR or 5LTR, 1.25 µL of Linker, 0.5 µL dNTPs and 0.5 µL of Herculase II Fusion DNA Enzyme. To each 45 µL of master mix, 5 µL of gDNA were added. First PCR reaction started with and initial step of 2 min at 95°C, followed by 35 cycles of denaturing at 95°C for 20 s, annealing at 60°C for 20 s and elongation at 72°C for 30 s. The final elongation step was at 72°C for 3 min.

8.2.1.6 EXOProSTAR™ Treatment

After the first PCR the triplicates were pooled and treated with ExoProStar™ (GE Healthcare) to remove leftover primers and other single stranded DNA. 2 µL of the ExoProStar™ solution were added to 15 µL of PCR product and mixed briefly. The reaction was incubated for 15 min at 37°C, followed by an enzyme inactivation at 80°C for 15 min.

8.2.1.7 2nd PCR

After purification, a nested PCR was used to increase the yield of LTR-host junction products. Nested PCR primers (3LTRnest or 5LTRnest and Linkernest) matched

sequences inside the first round PCR primers and were used to increase both the yield and the specificity of the amplification. The reaction was performed in triplicates with Herculase II Fusion DNA Polymerase (Agilent Technologies) in a volume of 50 µL. The master mix contained 31.5 µL of MilliQ H_2O, 10 µL of Herculase II Fusion DNA Reaction Buffer, 1.25 µL of 3LTRnest or 5LTRnest, 1.25 µL of Linker, 0.5 µL dNTPs and 0.5 µL of Herculase II Fusion DNA Enzyme. To each 45 µL of master mix, 5 µL of first PCR product were added. The second PCR reaction started with and initial step of 2 min at 95°C, followed by 30 cycles of denaturing at 95°C for 20 s, annealing at 55°C for 20 s and elongation at 72°C for 30 s. The final elongation step was at 72°C for 3 min.

Generally, a 5 µL aliquot of the PCR product was ran for 30 min at 100 V on a 1% agarose gel to confirm the presence of the expected smear in the range of 150-500bp.

8.2.1.8 Next Generation Sequencing (NGS)

Both ends of the amplified junction fragments were sequenced on the Illumina platform (San Diego, CA) to determine the viral/host junctions and the breakpoints in the host DNA. The sequence of the viral/host junction identifies the exact position and orientation in which the HIV-1 DNA was integrated. The breakpoints in the host DNA can be used to identify the integration sites in clonally expanded cells. If several cells with the same integration site are present, shearing their DNA will give rise to multiple fragments in which the integration site is the same, but the host DNA breakpoints differ.

The nested PCR products were sent to Kaiserslautern for NGS

NGS was carried out using a MiSeq 2 x 150bp paired end kit (Illumina). Read1 sequences include the LTR-host DNA junction and read2 sequences include the broken end of the host DNA.

Read 1 and read 2 were paired and primer sequences were trimmed. The trimmed reads were mapped against hg38 reference genome using bwa mem. Integration site was determined: “iSite” = mapping position + read_length (= sonic length) and for each integration site the number of different sonic lengths was determined.

True integration site had to fullfill the following criteria:

i) read1 contained the LTR primer and the last 7 bp of the LTR sequence

ii) followed by >20 bp DNA sequence with an average quality score Q = 60 with a >95% match to genomic DNA starting within 3bp of the LTR junction

8.2.2 Analysis of HIV-1 V3 Linked to Integration Sites

8.2.2.1 Primer Design

Patient specific primers for the amplification of the HIV-1 env sequences linked to integration sites were designed in the human sequence 3’ to the main clonal integration sites for 4 individuals. Primer features were analyzed using the Integrated DNA Technologies’ OligoAnalyzer tool (https://www.idtdna.com). To identify whether primers are likely to bind unspecific and amplify unintended regions of the human genome, primer sequences were analyzed using Primer-BLAST[123].

8.2.2.2 1st PCR

The reaction was performed in triplicates with Herculase II Fusion DNA Polymerase (Agilent) in a volume of 25 µL. The master mix contained 17.1 µL of H_2O, 5 µL of Herculase II Fusion DNA Reaction Buffer, 0.6 µL of F-ISAenv1 (1st PCR) or F-ISAenv2 (2nd PCR), 0.6 µL of integration site specific primer Primer_out (1st PCR) or Primer (2nd

PCR), 0.25 µL dNTPs and 0.25 µL Herculase II Fusion DNA Enzyme. To each 24 µL of master mix, 1 µL of gDNA was added.

First PCR reaction started with and initial step of 2 min at 95°C, followed by 30 cycles of denaturing at 95°C for 20 s, annealing at T_a-5°C for 20 s and elongation at 72°C for 3 min. The final elongation step was at 72°C for 3 min.

Because each *env-integration site* amplification used a patient specific primer, there were no positive control templates to optimize the PCR conditions.

8.2.2.3 2nd PCR

Conditions for 2nd PCR were the same as for the 1st PCR, only the elongation time was decreased to 2 min.

Generally, a 5 µL aliquot of the PCR product was ran for 30 min at 100 V on a 1% agarose gel to confirm the presence of the expected amplified product (~3 kb). The 2nd PCR product was purified by gel purification according to the protocol of NucleoSpin® Gel and PCR Clean-Up kit (Macherey Nagel). DNA was eluted in 25 µL elution buffer.

8.2.2.4 Sanger Sequencing

The V3 region of HIV-1 env was bi-directionally sequenced with the primers D-6991 and R-7365 by Microsynth.

8.2.3 Proviral Load Determination

To measure the proviral load, a multiplex qPCR reaction was performed with a VIC labeled CCR5 probe and a FAM labeled HIV-1 LTR probe.

15 µL of master mix were mixed with 5 µL of extracted patient gDNA. For qPCR standards pre-determined gDNA amounts of HUT4-3 cells were used. For one reaction the master mix was composed of 0.4 µL of 10 µM forward primers (F-522 and F-CCR5), 0.4 µL of 10 µM reverse primers (R-643 and R-CCR5), 0.25 µL of 10 µM CCR5- and LTR probe and 10 µL of 2x Luna® Universal qPCR master mix) in a final volume of 20 µL. Cycling conditions were as follows: one minute at 95°C, 45 cycles of alternating 15 seconds at 95°C and one minute at 60°C.

All measurements were performed in triplicates, and a non-template control was always included to detect the presence of HIV-1 contaminants.

8.2.4 V3 Tropism Analysis

To investigate viral tropism and the stability of observed outgrown X4 variants under therapy, gDNA from PBMCs from selected HIV-infected patients after 10 years was analyzed (see 5.2.1.1 Isolation of genomic DNA).

8.2.4.1 1st PCR

The variable loop 3 (V3) region of HIV-1 was amplified by a nested PCR, both reactions were performed with Herculase II Fusion DNA Polymerase (Agilent Technologies) in a volume of 25 µL. To reduce the impact of PCR bias, the nested PCR for each patient sample was performed in triplicates. The master mix contained 17.1 µL of H_2O, 5 µL of Herculase II Fusion DNA Reaction Buffer, 0.6 µL of F-6553 (1st PCR) or F-6848 (2nd PCR), 0.6 µL of R-7801 (1st PCR) or R-7371 (2nd PCR), 0.25 µL dNTPs and 0.25 µL Herculase II Fusion DNA Enzyme. To each 24 µL of master mix, 1 µL of gDNA (1st PCR) or 1st PCR product (2nd PCR) was added.

First PCR reaction started with and initial step of 2 min at 95°C, followed by 30 cycles of denaturing at 95°C for 20 s, annealing at 60°C for 20 s and elongation at 72°C for 1 min. The final elongation step was at 72°C for 3 min.

Before continuing with the 2nd PCR, the triplicates of the 1st PCR were pooled.

8.2.4.2 2nd PCR

Conditions for 2nd PCR were the same as for the 1st PCR, only the annealing temperature was changed to 56°C and the elongation time was decreased to 30 s.

Generally, a 5 µL aliquot of the PCR product was ran for 30 min at 100V on a 1% agarose gel to confirm the presence of the expected amplified product (~500 bp). Each replicate of the 2nd PCR was pooled and purified according to the protocol of NucleoSpin® Gel and PCR Clean-Up kit (Macherey Nagel). DNA was eluted in 25 µL elution buffer.

8.2.4.3 Next Generation Sequencing (NGS)

After PCR the DNA product was cleaned with Agencourt AMPour XP beads (Beckmann Coulter) according to the manufacturer's protocol.

8.2.4.4 DNA Quantification

For quantification the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen) was used according to the manufacturer's protocol.

8.2.4.5 Library Preparation

DNA concentration was adjusted to 0.2 ng/µL and the Nextera XT DNA Library Preparation Kit (Illumina) was used to prepare the library according to kit instructions.

8.2.4.6 Sequencing

Sequencing was performed with a Illumina MiSeq Benchtop sequencer with 2x250bp reads.

8.2.4.7 Data Analysis

Tropism determination was done with the Geno2Pheno454 tool with a FPR cut-off value of 3.5%. Patient samples were defined as R5-tropic if the relative frequency of X4-variants in the patient pool was below 2%. If the frequency of X4-variants in the same patient changed by less than 5% between timepoints it was designated as a stable tropism.

8.2.4.8 Sequence Analysis

The sequences were analyzed with Geneious Prime® (version 2020.2.4).

8.2.5 Phenotyping and Genotyping of CRF01_AE clinical samples

In order to improve the tropism prediction for CRF01_AE samples by geno2pheno[coreceptor], phenotypic results were combined with genotypic predictions.

8.2.5.1 Sample preparation in Thailand

RNA was extracted from 150 µL plasma using the NucleoSpin viral RNA extraction kit (Macherey Nagel AG) following manufacturer's instructions. The cDNA and PCR product were then obtained using the SuperScript III One Step RT-PCR kit (Invitrogen). Outer PCR was performed with the primer pair KVL008 and KVL009; a nested PCR reaction was performed with the primer pair F-6435alt and R-8319NaeI.

The temperature profile for the outer PCR was 40 min at 55°C, 2 min at 95°C followed by 40 cycles of 15 s at 95°C, 15 s at 55°C and 90 s at 72°C, and 7 min at 72°C. The reaction mixture for the nested PCR contained 3.5 µL of the product from the first PCR. The amplification profile in the second PCR was 2 min at 92°C followed by 30 cycles of 10 s at 94°C, 15 s at 60°C and 60 s at 72°C, and 7 min at 72°C.

8.2.5.2 Sample amplification by PCR

The DNA products were re-amplified by PCR, using the enzyme Herculase II Fusion DNA polymerase (Agilent Technologies). The total reaction volume was 50 µL containing 45 µL of master mix and 5µL of purified PCR product. The master mix included 34 µL of MilliQ H_2O, 10 µL of Herculase II Reaction Buffer, 1.25 µL of 10 µM F-6435alt primer, 1.25 µL of 10 µM R-8319NaeI primer, 0.5 µL of (10 µM) dNTPs and 1µL Herculase II Fusion DNA Polymerase. After an initial step of two minutes at 95°C for denaturation of DNA, 30 cycles of denaturing at 95°C for 15 seconds, annealing at 60°C for 20 seconds and elongation for 60 seconds at 72°C followed. The final elongation step was for three minutes at 68°C.

8.2.5.3 PCR purification

For verification, the amplified DNA was loaded onto a 1% agarose gel and run at 10 V/cm for 30 minutes. The band with the expected size (1.9 kb) was cut and gel extraction was performed according to the protocol of the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel AG). DNA was eluted in 20 µL of elution buffer.

8.2.5.4 Cloning

1 µg of inserts and 2 µg of the corresponding plasmids (pNL-K7, pNL-AE-K7) were digested with 1 unit of MluI-HF and NgoMIV (New England Biolabs). The digested plasmid underwent a CIP-treatment to impede vector re-ligation and purification by agarose gel electrophoresis. For ligating the inserts into the respective plasmids, one unit of T4 DNA ligase (New England Biolabs) was added to 60 ng of vector and insert (at a molar ratio of 1:3) and incubated for 10 min at room temperature. Chemically competent stbl3 *E. coli* (ThermoFisher) was transformed with the ligated plasmid DNA according to the manufacturer's instructions. 4 mL of LB + Amp (100 µg/mL) was directly inoculated without plating in order to retain the viral diversity present in the amplified HIV-1 DNA, reflecting non-clonal HIV-1 in the clinical specimens. Transformed bacteria were grown overnight at 37°C in a shaker incubator.

Plasmid isolation was then performed using the NucleoSpin Plasmid Transfection-grade MiniPrep kit (Macherey-Nagel AG) according to the manufacturer's manual. To verify the integrity of the yielded plasmid, a restriction digest with HindIII-HF (New England Biolabs) was performed. The digestion pattern was checked after electrophoresis in a 1% agarose gel and compared to the known pattern of the plasmids.

8.2.5.5 Transfection

Plasmid DNA transfection of mammalian cells was performed using the jetPRIME Transfection Kit (Polyplus-transfection) according to the manufacturer's protocol. pNL-NF, a prototypic CXCR4-tropic subtype B plasmid of HIV-1[124], was used as positive control in all transfection experiments. For all experiments using the plasmid pNL-K7, 200'000 SXR5 cells per well were seeded in a 12-well plate and directly transfected with 1 µg of plasmid DNA. In SXR5 cells, viral replication simultaneously induces an

endogenous, LTR-driven lacZ gene via the expression of HIV-1 Tat in the transfected cell; lacZ activity is analyzed by detecting ß-galactosidase activity using X-Gal staining. After adding the transfection mix, cells were incubated for 4 h at 37°C. Then the transfection mix was removed, cells were washed with 1 mL of medium before 1 mL of cDMEM was added. One well remained without drug addition, to one well AMD3100 was added as bona-fide CXCR4-inhibitor[125] at concentrations exceeding the IC90 concentration, and the third well was incubated in the presence of fully inhibitory concentrations of maraviroc (MVC), a CCR5-antagonists[126].

Experiments with the plasmid pMN-AE-env-K7 were done as co-cultures: 293T cells were transfected with 1 µg of plasmid DNA, and the reporter cells SXR5 were added after a wash step, at the same time as adding the drugs. Cultures were incubated for 48 hours at 37°C in a BSL-3 environment.

8.2.5.6 Phenotyping

After 48 hours of incubation, culture media was removed, and cells were fixed with a PBS + 2% PFA solution for 15 minutes. Fixation buffer was aspirated, and X-Gal staining solution added to the cells. After 1h incubation at 37°C, the cells were analyzed by optical microscopy. Viral replication was judged as replication in the presence of either inhibitor. This allowed to obtain direct information about the viral tropism: CCR5-tropic HIV-1 is only inhibited by MVC but not by AMD3100 while CXCR4-tropic HIV-1 is only inhibited by AMD3100. Accordingly, the phenotypes of the isolates in this study were identified by determining the inhibition of viral replication in presence of either AMD3100 or MVC and by the formation of syncytia.

8.2.5.7 Genotyping

DNA of the recombinant plasmids, used for phenotyping, was subsequently sequenced to correlate genotypic information with the phenotypic values. For this, 20 µL of each patient-plasmid underwent V3 sequencing and coreceptor prediction using the geno2pheno [coreceptor] tool. Plasmid population sequencing was performed with the ABI 3130xl Genetic Analyzer following the sequencing protocol of Sierra et al.[105], using the following sequencing primers: ENV-2, ENV-11 and subtype CRF01_AE adapted primers ENV-6_AE (*AGCCAGTGGTATCAACTCAAT*) and ENV-7_AE (*TTTCCACTGATGGGAGGAGC*).

9 Literature

1. Fauci, A. S. Host factors and the pathogenesis of HIV-induced disease. *Nature* **384**, 529–534 (1996).
2. Pérez-Olmeda, M. & Alcami, J. Determination of HIV tropism and its use in the clinical practice. *Expert Rev. Anti. Infect. Ther.* **11**, 1291–1302 (2013).
3. Zhao, H., Prosser, A. R., Liotta, D. C. & Wilson, L. J. Discovery of novel N-aryl piperazine CXCR4 antagonists. *Bioorg. Med. Chem. Lett.* **25**, 4950–4955 (2015).
4. Tan, S. *et al.* Penicillixanthone A, a marine-derived dual-coreceptor antagonist as anti-HIV-1 agent. *Nat. Prod. Res.* **33**, 1467–1471 (2019).
5. Murakami, T. *et al.* The novel CXCR4 antagonist KRH-3955 is an orally bioavailable and extremely potent inhibitor of human immunodeficiency virus type 1 infection: comparative studies with AMD3100. *Antimicrob. Agents Chemother.* **53**, 2940–2948 (2009).
6. Doranz, B. J. *et al.* A small-molecule inhibitor directed against the chemokine receptor CXCR4 prevents its use as an HIV-1 coreceptor. *J. Exp. Med.* **186**, 1395–1400 (1997).
7. Cocchi, F. *et al.* The V3 domain of the HIV–1 gp120 envelope glycoprotein is critical for chemokine–mediated blockade of infection. *Nat. Med.* **2**, 1244–1247 (1996).
8. Fouchier, R. A. *et al.* Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J. Virol.* **66**, 3183–3187 (1992).
9. Resch, W., Hoffman, N. & Swanstrom, R. Improved Success of Phenotype Prediction of the Human Immunodeficiency Virus Type 1 from Envelope Variable

Loop 3 Sequence Using Neural Networks. *Virology* **288**, 51–62 (2001).

10. Carrillo, A. & Ratner, L. Cooperative effects of the human immunodeficiency virus type 1 envelope variable loops V1 and V3 in mediating infectivity for T cells. *J. Virol.* **70**, 1310–1316 (1996).
11. Pollakis, G. *et al.* N-Linked Glycosylation of the HIV Type-1 gp120 Envelope Glycoprotein as a Major Determinant of CCR5 and CXCR4 Coreceptor Utilization*. *J. Biol. Chem.* **276**, 13433–13441 (2001).
12. Lengauer, T., Sander, O., Sierra, S., Thielen, A. & Kaiser, R. Bioinformatics prediction of HIV coreceptor usage. *Nat. Biotechnol.* **25**, 1407–1410 (2007).
13. Schuitemaker, H., van 't Wout, A. B. & Lusso, P. Clinical significance of HIV-1 coreceptor usage. *J. Transl. Med.* **9**, S5–S5 (2011).
14. Connor, R. I., Sheridan, K. E., Ceradini, D., Choe, S. & Landau, N. R. Change in Coreceptor Use Correlates with Disease Progression in HIV-1–Infected Individuals. *J. Exp. Med.* **185**, 621 (1997).
15. Moore, J. P., Kitchen, S. G., Pugach, P. & Zack, J. A. The CCR5 and CXCR4 Coreceptors—Central to Understanding the Transmission and Pathogenesis of Human Immunodeficiency Virus Type 1 Infection. *AIDS Res. Hum. Retroviruses* (2004) doi:10.1089/088922204322749567.
16. Tsibris, A. M. N. *et al.* Quantitative Deep Sequencing Reveals Dynamic HIV-1 Escape and Large Population Shifts during CCR5 Antagonist Therapy In Vivo. *PLoS One* **4**, e5683 (2009).
17. Symons, J. *et al.* Maraviroc is able to inhibit dual-R5 viruses in a dual/mixed HIV-1-infected patient. *J. Antimicrob. Chemother.* **66**, 890–895 (2011).
18. Grivel, J.-C., Shattock, R. J. & Margolis, L. B. Selective transmission of R5 HIV-1 variants: where is the gatekeeper? *J. Transl. Med.* **9**, S6 (2011).

19. Bunnik, E. M. *et al.* Detection of Inferred CCR5- and CXCR4-Using HIV-1 Variants and Evolutionary Intermediates Using Ultra-Deep Pyrosequencing. *PLOS Pathog.* **7**, e1002106 (2011).

20. Ceresola, E. R. *et al.* Performance of commonly used genotypic assays and comparison with phenotypic assays of HIV-1 coreceptor tropism in acutely HIV-1-infected patients. *J. Antimicrob. Chemother.* **70**, 1391–1395 (2015).

21. Wagner, G. A. *et al.* Using Ultradeep Pyrosequencing to Study HIV-1 Coreceptor Usage in Primary and Dual Infection. *J. Infect. Dis.* **208**, 271–274 (2013).

22. de Mendoza, C. *et al.* Prevalence of X4 tropic viruses in patients recently infected with HIV-1 and lack of association with transmission of drug resistance. *J. Antimicrob. Chemother.* **59**, 698–704 (2007).

23. Polzer, S. *et al.* Loss of N-linked glycans in the V3-loop region of gp120 is correlated to an enhanced infectivity of HIV-1. *Glycobiology* **11**, 11–19 (2001).

24. Tsuchiya, K. *et al.* Arginine insertion and loss of N-linked glycosylation site in HIV-1 envelope V3 region confer CXCR4-tropism. *Sci. Rep.* **3**, 2389 (2013).

25. Miedema, F., Tersmette, M. & van Lier, R. . AIDS pathogenesis: a dynamic interaction between HIV and the immune system. *Immunol. Today* **11**, 293–297 (1990).

26. Wodarz, D., Lloyd, A. L., Jansen, V. A. A. & Nowak, M. A. Dynamics of Macrophage and T Cell Infection by HIV. *J. Theor. Biol.* **196**, 101–113 (1999).

27. Bleul, C. C., Wu, L., Hoxie, J. A., Springer, T. A. & Mackay, C. R. The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proc. Natl. Acad. Sci.* **94**, 1925 LP – 1930 (1997).

28. Connell, B. J. *et al.* Immune activation correlates with and predicts CXCR4 co-receptor tropism switch in HIV-1 infection. *Sci. Rep.* **10**, 15866 (2020).

29. Weiser, B. *et al.* HIV-1 coreceptor usage and CXCR4-specific viral load predict clinical disease progression during combination antiretroviral therapy. *AIDS* **22**, (2008).

30. Brumme, Z. L. *et al.* Clinical and immunological impact of HIV envelope V3 sequence variation after starting initial triple antiretroviral therapy. *AIDS* **18**, (2004).

31. Parisi, S. G. *et al.* A stable CC-chemokine receptor (CCR)-5 tropic virus is correlated with the persistence of HIV RNA at less than 2.5 copies in successfully treated naïve subjects. *BMC Infect. Dis.* **13**, 314 (2013).

32. Waters, L. *et al.* The impact of HIV tropism on decreases in CD4 cell count, clinical progression, and subsequent response to a first antiretroviral therapy regimen. *Clin. Infect. Dis.* **46**, 1617–23 (2008).

33. Soulié, C. *et al.* Factors associated with proviral DNA HIV-1 tropism in antiretroviral therapy-treated patients with fully suppressed plasma HIV viral load: implications for the clinical use of CCR5 antagonists. *J. Antimicrob. Chemother.* **65**, 749–751 (2010).

34. Seclén, E. *et al.* Impact of Baseline HIV-1 Tropism on Viral Response and CD4 Cell Count Gains in HIV-Infected Patients Receiving First-line Antiretroviral Therapy. *J. Infect. Dis.* **204**, 139–144 (2011).

35. Lombardi, F. *et al.* HIV-1 non-R5 tropism correlates with a larger size of the cellular viral reservoir and a detectable residual viremia in patients under suppressive ART. *J. Clin. Virol.* **103**, 57–62 (2018).

36. Delobel, P. *et al.* R5 to X4 Switch of the Predominant HIV-1 Population in Cellular Reservoirs During Effective Highly Active Antiretroviral Therapy. *JAIDS J. Acquir. Immune Defic. Syndr.* **38**, (2005).

37. Soulié, C. *et al.* HIV-1 X4/R5 co-receptor in viral reservoir during suppressive HAART. *AIDS* **21**, (2007).

38. Philpott, S. *et al.* Preferential suppression of CXCR4-specific strains of HIV-1 by antiviral therapy. *J. Clin. Invest.* **107**, 431–438 (2001).

39. Bader, J. *et al.* Therapeutic Immune Recovery and Reduction of CXCR4-Tropic HIV-1. *Clin. Infect. Dis.* **64**, 295–300 (2017).

40. Maldarelli, F. *et al.* Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science (80-.).* (2014) doi:10.1126/science.1254194.

41. Wagner, T. A. *et al.* Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science (80-.).* **345**, 570 (2014).

42. Josefsson, L. *et al.* The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc. Natl. Acad. Sci.* **110**, E4987–E4996 (2013).

43. Kearney, M. F. *et al.* Lack of Detectable HIV-1 Molecular Evolution during Suppressive Antiretroviral Therapy. *PLOS Pathog.* **10**, e1004010 (2014).

44. UNAIDS. Global Hiv Statistics 2020. *End. AIDS epidemic* 1–3 (2020).

45. Davey Jr, R. T. *et al.* HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 15109–15114 (1999).

46. Archin, N. M. *et al.* Immediate antiviral therapy appears to restrict resting $CD4^+$ cell HIV-1 infection without accelerating the decay of latent infection. *Proc. Natl. Acad. Sci.* **109**, 9523 LP – 9528 (2012).

47. Coffin, J. M. *et al.* Clones of infected cells arise early in HIV-infected individuals. *JCI Insight* **4**, (2019).

48. Blankson, J. N., Persaud, D. & Siliciano, R. F. The Challenge of Viral Reservoirs in HIV-1 Infection. *Annu. Rev. Med.* **53**, 557–593 (2002).

49. Eisele, E. & Siliciano, R. F. Redefining the Viral Reservoirs that Prevent HIV-1 Eradication. *Immunity* **37**, 377–388 (2012).

50. Bachmann, N. *et al.* Determinants of HIV-1 reservoir size and long-term dynamics during suppressive ART. *Nat. Commun.* **10**, 3193 (2019).

51. Deeks, S. G. *et al.* Towards an HIV cure: a global scientific strategy. *Nat. Rev. Immunol.* **12**, 607–614 (2012).

52. Thorball, C. W. *et al.* Host genomics of the HIV-1 reservoir size and its decay rate during suppressive antiretroviral treatment. *medRxiv* 19013763 (2019) doi:10.1101/19013763.

53. Besson, G. J. *et al.* HIV-1 DNA Decay Dynamics in Blood During More Than a Decade of Suppressive Antiretroviral Therapy. *Clin. Infect. Dis.* **59**, 1312–1321 (2014).

54. Siliciano, J. D. *et al.* Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat. Med.* **9**, 727–728 (2003).

55. Kearney, M. F. *et al.* Lack of Detectable HIV-1 Molecular Evolution during Suppressive Antiretroviral Therapy. *PLOS Pathog.* **10**, e1004010 (2014).

56. von Stockenstrom, S. *et al.* Longitudinal Genetic Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool During Effective HIV Therapy. *J. Infect. Dis.* **212**, 596–607 (2015).

57. Chomont, N. *et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat. Med.* (2009) doi:10.1038/nm.1972.

58. Chun, T.-W. *et al.* Early establishment of a pool of latently infected, resting CD4 T cells during primary HIV-1 infection (latency¯primary infection¯HAART therapy).

Med. Sci. **95**, 8869–8873 (1998).

59. Finzi, D. *et al.* Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat. Med.* **5**, 512–517 (1999).

60. Simonetti, F. R. *et al.* Clonally expanded CD4 $^{+}$ T cells can produce infectious HIV-1 in vivo. *Proc. Natl. Acad. Sci.* (2016) doi:10.1073/pnas.1522675113.

61. Bruner, K. M. *et al.* Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat. Med.* (2016) doi:10.1038/nm.4156.

62. Ho, Y.-C. *et al.* Replication-Competent Noninduced Proviruses in the Latent Reservoir Increase Barrier to HIV-1 Cure. *Cell* **155**, 540–551 (2013).

63. Ikeda, T., Shibata, J., Yoshimura, K., Koito, A. & Matsushita, S. Recurrent HIV-1 Integration at the BACH2 Locus in Resting CD4+ T Cell Populations during Effective Highly Active Antiretroviral Therapy. *J. Infect. Dis.* **195**, 716–725 (2007).

64. Cesana, D. *et al.* HIV-1-mediated insertional activation of STAT5B and BACH2 trigger viral reservoir in T regulatory cells. *Nat. Commun.* **8**, 498 (2017).

65. Rethi, B. *et al.* Loss of IL-7Rα is associated with CD4 T-cell depletion, high interleukin-7 levels and CD28 down-regulation in HIV infected patients. *AIDS* **19**, (2005).

66. Day, C. L. *et al.* PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. *Nature* **443**, 350–354 (2006).

67. D'Souza, M. *et al.* Programmed Death 1 Expression on HIV-Specific CD4⁺ T Cells Is Driven by Viral Replication and Associated with T Cell Dysfunction. *J. Immunol.* **179**, 1979 LP – 1987 (2007).

68. Zeng, M. *et al.* Lymphoid Tissue Damage in HIV-1 Infection Depletes Naïve T

Cells and Limits T Cell Reconstitution after Antiretroviral Therapy. *PLOS Pathog.* **8**, e1002437 (2012).

69. Napolitano, L. A. *et al.* Increased production of IL-7 accompanies HIV-1–mediated T-cell depletion: implications for T-cell homeostasis. *Nat. Med.* **7**, 73–79 (2001).

70. Vandergeeten, C. *et al.* Interleukin-7 promotes HIV persistence during antiretroviral therapy. *Blood* **121**, 4321–4329 (2013).

71. Mendoza, P. *et al.* Antigen-responsive CD4+ T cell clones contribute to the HIV-1 latent reservoir. *J. Exp. Med.* **217**, (2020).

72. Simonetti, F. R. *et al.* Antigen-driven clonal selection shapes the persistence of HIV-1–infected CD4+ T cells in vivo. *J. Clin. Invest.* **131**, (2021).

73. Bailey, J. R. *et al.* Residual Human Immunodeficiency Virus Type 1 Viremia in Some Patients on Antiretroviral Therapy Is Dominated by a Small Number of Invariant Clones Rarely Found in Circulating CD4⁺ T Cells. *J. Virol.* **80**, 6441 LP – 6457 (2006).

74. Cohn, L. B., Chomont, N. & Deeks, S. G. The Biology of the HIV-1 Latent Reservoir and Implications for Cure Strategies. *Cell Host Microbe* **27**, 519–530 (2020).

75. Bader, J. *et al.* Correlating HIV tropism with immunological response under combination antiretroviral therapy. *HIV Med.* (2016) doi:10.1111/hiv.12365.

76. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

77. Wells, D. W. *et al.* An analytical pipeline for identifying and mapping the integration sites of HIV and other retroviruses. *BMC Genomics* **21**, 216 (2020).

78. Brin, E. & Leis, J. HIV-1 Integrase Interaction with U3 and U5 Terminal

Sequences in Vitro Defined Using Substrates with Random Sequences *. *J. Biol. Chem.* **277**, 18357–18364 (2002).

79. Maldarelli, F. *et al.* Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science (80-.).* **345**, 179 (2014).

80. Poiesz, B. J. *et al.* Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 7415–7419 (1980).

81. Yoshida, M., Miyoshi, I. & Hinuma, Y. Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 2031–2035 (1982).

82. Satou, Y. *et al.* The retrovirus HTLV-1 inserts an ectopic CTCF-binding site into the human genome. *PNAS* **113**, 3054–3059 (2016).

83. Cook, L. B. *et al.* The role of HTLV-1 clonality, proviral structure, and genomic integration site in adult T-cell leukemia/lymphoma. *Blood* (2014)

84. Gillet, N. A. *et al.* The host genomic environment of the provirus determines the abundance of HTLV-1–infected T-cell clones. *Blood* **117**, 3113–3122 (2011).

85. Etoh, K. *et al.* Persistent Clonal Proliferation of Human T-lymphotropic Virus Type I-infected Cells in Vivo *Cancer Res.* **57**, 4862 LP – 4867 (1997).

86. Cavrois, M., Wain-Hobson, S., Gessain, A., Plumelle, Y. & Wattel, E. Adult T-Cell Leukemia/Lymphoma on a Background of Clonally Expanding Human T-Cell Leukemia Virus Type-1–Positive Cells. *Blood* **88**, 4646–4650 (1996).

87. Liu, R., Simonetti, F. R. & Ho, Y. C. The forces driving clonal expansion of the HIV-1 latent reservoir. *Virology Journal* vol. 17 (2020).

88. Wang, Z. *et al.* Expanded cellular clones carrying replication-competent HIV-1

persist, wax, and wane. *Proc. Natl. Acad. Sci.* (2018) doi:10.1073/pnas.1720665115.

89. Saracino, A. *et al.* Co-receptor switch during HAART is independent of virological success. *J. Med. Virol.* **81**, 2036–2044 (2009).

90. Castagna, A. *et al.* Switch of predicted HIV-1 tropism in treated subjects and its association with disease progression. *Medicine (Baltimore).* **95**, e5222–e5222 (2016).

91. Raymond, S. *et al.* Evolution of HIV-1 quasispecies and coreceptor use in cell reservoirs of patients on suppressive antiretroviral therapy. *J. Antimicrob. Chemother.* **69**, 2527–2530 (2014).

92. Sigal, A. *et al.* Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. *Nature* **477**, 95–98 (2011).

93. Lorenzo-Redondo, R. *et al.* Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature* **530**, 51–56 (2016).

94. Fletcher, C. V *et al.* Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues. *Proc. Natl. Acad. Sci.* **111**, 2307 LP – 2312 (2014).

95. Imamichi, H. *et al.* Lack of Compartmentalization of HIV-1 Quasispecies Between the Gut and Peripheral Blood Compartments. *J. Infect. Dis.* **204**, 309–314 (2011).

96. van Marle, G. *et al.* Compartmentalization of the gut viral reservoir in HIV-1 infected patients. *Retrovirology* **4**, 87 (2007).

97. Bacchus-Souffan, C. *et al.* Relationship between CD4 T cell turnover, cellular differentiation and HIV persistence during ART. *PLOS Pathog.* **17**, e1009214 (2021).

98. Zhou, S., Bednar, M. M., Sturdevant, C. B., Hauser, B. M. & Swanstrom, R. Deep Sequencing of the HIV-1 env Gene Reveals Discrete X4 Lineages and Linkage Disequilibrium between X4 and R5 Viruses in the V1/V2 and V3 Variable Regions. *J. Virol.* **90**, 7142–7158 (2016).

99. Roche, M. *et al.* CXCR4-Using HIV Strains Predominate in Naive and Central Memory CD4(+) T Cells in People Living with HIV on Antiretroviral Therapy: Implications for How Latency Is Established and Maintained. *J. Virol.* **94**, e01736-19 (2020).

100. Yu, F. *et al.* X4-Tropic Latent HIV-1 Is Enriched in Peripheral Follicular Helper T Cells and Is Correlated with Disease Progression. *J. Virol.* **94**, e01219-19 (2020).

101. Virgilio, M. C. & Collins, K. L. The Impact of Cellular Proliferation on the HIV-1 Reservoir. *Viruses* vol. 12 (2020).

102. Venanzi Rullo, E. *et al.* Persistence of an intact HIV reservoir in phenotypically naive T cells. *JCI insight* **5**, e133157 (2020).

103. Zerbato, J. M., McMahon, D. K., Sobolewski, M. D., Mellors, J. W. & Sluis-Cremer, N. Naive CD4+ T Cells Harbor a Large Inducible Reservoir of Latent, Replication-competent Human Immunodeficiency Virus Type 1. *Clin. Infect. Dis.* **69**, 1919–1925 (2019).

104. Rozera, G. *et al.* Evolution of HIV-1 tropism at quasispecies level after 5 years of combination antiretroviral therapy in patients always suppressed or experiencing episodes of virological failure*. *J. Antimicrob. Chemother.* **69**, 3085–3094 (2014).

105. AU - Sierra, S. *et al.* Prediction of HIV-1 Coreceptor Usage (Tropism) by Sequence Analysis using a Genotypic Approach. *JoVE* e3264 (2011)

doi:doi:10.3791/3264.

106. Matsuda, M. *et al.* Performance Evaluation of a Genotypic Tropism Test Using HIV-1 CRF01_AE Isolates in Japan. *Jpn. J. Infect. Dis.* **71**, 264–266 (2018).

107. Chu, M. *et al.* HIV-1 CRF01-AE strain is associated with faster HIV/AIDS progression in Jiangsu Province, China. *Sci. Rep.* **7**, (2017).

108. Ng, O. T. *et al.* Increased rate of CD4+ T-cell decline and faster time to antiretroviral therapy in HIV-1 subtype CRF01_AE infected seroconverters in Singapore. *PLoS One* **6**, e15738 (2011).

109. Li, Y. *et al.* CRF01-AE subtype is associated with X4 tropism and fast HIV progression in Chinese patients infected through sexual transmission. *AIDS* **28**, 521–530 (2014).

110. Saeng-aroon, S. *et al.* Circulation of HIV-1 Multiple Complexity Recombinant Forms Among Female Sex Workers Recently Infected with HIV-1 in Thailand. *AIDS Res. Hum. Retroviruses* **32**, 694–701 (2016).

111. Fehr, J. *et al.* Replicative phenotyping adds value to genotypic resistance testing in heavily pre-treated HIV-infected individuals - the Swiss HIV Cohort Study. *J. Transl. Med.* **9**, 14 (2011).

112. Obermeier, M., Symons, J. & Wensing, A. M. J. HIV population genotypic tropism testing and its clinical significance. *Curr. Opin. HIV AIDS* **7**, 470–477 (2012).

113. Klimkait, T., Stauffer, F., Lupo, E. & Sonderegger-Rubli, C. Dissecting the mode of action of various HIV-inhibitor classes in a stable cellular system. *Arch. Virol.* **143**, 2109–2131 (1998).

114. Gardiner, J. C., Mauer, E. J. & Sherer, N. M. HIV-1 Gag, Envelope, and Extracellular Determinants Cooperate To Regulate the Stability and Turnover of

Virological Synapses. *J. Virol.* **90**, 6583–6597 (2016).

115. Chojnacki, J. *et al.* Envelope glycoprotein mobility on HIV-1 particles depends on the virus maturation state. *Nat. Commun.* **8**, 545 (2017).
116. Schwartz, S., Felber, B. K., Fenyö, E. M. & Pavlakis, G. N. Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *J. Virol.* **64**, 5448 LP – 5456 (1990).
117. Wildum, S., Schindler, M., Münch, J. & Kirchhoff, F. Contribution of Vpu, Env, and Nef to CD4 down-modulation and resistance of human immunodeficiency virus type 1-infected T cells to superinfection. *J. Virol.* **80**, 8047–8059 (2006).
118. Hongjaisee, S., Nantasenamat, C., Carraway, T. S. & Shoombuatong, W. HIVCoR: A sequence-based tool for predicting HIV-1 CRF01_AE coreceptor usage. *Comput. Biol. Chem.* **80**, 419–432 (2019).
119. Soulie, C. *et al.* Performance of genotypic algorithms for predicting tropism for HIV-1 CRF01_AE recombinant. *J. Clin. Virol.* **99**–**100**, 57–60 (2018).
120. Swenson, L. C. *et al.* Deep Sequencing to Infer HIV-1 Co-Receptor Usage: Application to Three Clinical Trials of Maraviroc in Treatment-Experienced Patients. *J. Infect. Dis.* **203**, 237–245 (2011).
121. Sing, T. *et al.* Predicting HIV coreceptor usage on the basis of genetic and clinical covariates. *Antivir. Ther.* **12**, 1097–1106 (2007).
122. Cui, H. *et al.* Rapid CD4+ T-cell decline is associated with coreceptor switch among MSM primarily infected with HIV-1 CRF01_AE in Northeast China. *AIDS* **33**, 13–22 (2019).
123. Ye, J. *et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**, 134 (2012).
124. Bleiber, G., Munoz, M., Ciuffi, A., Meylan, P. & Telenti, A. Individual

Contributions of Mutant Protease and Reverse Transcriptase to Viral Infectivity, Replication, and Protein Maturation of Antiretroviral Drug-Resistant Human Immunodeficiency Virus Type 1. *J. Virol.* **75**, 3291 LP – 3300 (2001).

125. Stalmeijer, E. H. B. *et al.* In Vivo Evolution of X4 Human Immunodeficiency Virus Type 1 Variants in the Natural Course of Infection Coincides with Decreasing Sensitivity to CXCR4 Antagonists. *J. Virol.* **78**, 2722 LP – 2728 (2004).
126. Baba, M. *et al.* A small-molecule, nonpeptide CCR5 antagonist with highly potent and selective anti-HIV-1 activity. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 5698–5703 (1999).

10 Acknowledgments

I would like to express my gratitude to Prof. Dr. Thomas Klimkait for the opportunity to carry out my PhD work in his lab group. I am grateful for the time and many learning experiences he gave me.

I would also like to thank Prof. Dr. Markus Affolter and Dr. Jürg Böni for for agreeing to be on my PhD committee and sharing their time and experience.

My heartfelt thanks go to all Molecular Virology group members, Ulrike Seeburg, Yuepeng Zhang, Lorena Urda, Fabian Otte and Jennifer Brown. I am also grateful to have been able to share some enjoyable times with the Master students in the Molecular Virology group: Silvia Caimi and Benjamin Schwob. The supervision of the interns Sina Berlin and Siro Ellenberger was great fun and a learning experience – thanks for being curious and asking many questions.

I am also grateful for the support of my family; a special thanks goes to my sister Anna, who sent me a box of chocolate just at the right time.

Last but not least: Ta Robert, for still sticking to the couch with me.

11 Publications

Marty N, Saeng-Aroon S, Heger E el al. (2021) Adapting the geno2pheno[coreceptor] tool to HIV-1 subtype CRF01_AE by phenotypic validation using clinical isolates from South-East Asia. Journal of Clinical Virology, 136

Journal of Clinical Virology 136 (2021) 104755

Contents lists available at ScienceDirect

Journal of Clinical Virology

journal homepage: www.elsevier.com/locate/jcv

Adapting the geno2pheno[coreceptor] tool to HIV-1 subtype CRF01_AE by phenotypic validation using clinical isolates from South-East Asia

Nina Marty [a,*], Siriphan Saeng-Aroon [b], Eva Heger [c], Alexander Thielen [d], Martin Obermeier [e], Nico Pfeifer [f], Rolf Kaiser [c,1], Thomas Klimkait [a,1]

[a] Molecular Virology, Department Biomedicine-Petersplatz, University of Basel, Petersplatz 10, 4055 Basel, Switzerland
[b] Hazardous Pathogen Laboratory, National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand
[c] Institute of Virology, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany
[d] Seq-It GmbH & Co. KG, Kaiserslautern, Germany
[e] Medizinisches Infektiologiezentrum Berlin, Berlin, Germany
[f] Max Planck Institute for Informatics, Saarland Informatics Campus E1 4, Saarbruecken, Germany

ARTICLE INFO

Keywords:
HIV-1
Tropism
Subtype CRF01_AE
Coreceptor
geno2pheno
Phenotyping

ABSTRACT

Objectives: Geno2pheno[coreceptor] is a widely used tool for the prediction of coreceptor usage (viral tropism) of HIV-1 samples. For HIV-1 CRF01_AE, a significant overcalling of X4-tropism is observed when using the standard settings of Geno2pheno[coreceptor]. The aim of this study was to provide the experimental backing for adaptations to the geno2pheno[coreceptor] algorithm in order to improve coreceptor usage predictions of clinical HIV-1 CRF01_AE isolates

Study design: V3-sequences of 20 clinical HIV-1 subtype CRF01_AE samples were sequenced and analyzed by geno2pheno[coreceptor]. In parallel, coreceptor usage was determined for these samples by replicative phenotyping in human cells in the presence of specific X4- or R5-inhibitors.

Results: The sole introduction of the CRF01_AE V3 region into a full-length otherwise subtype B provirus failed to produce replication-competent viral progeny. A successive genome-replacement strategy revealed that also CRF01_AE derived gag and pol sequences are necessary to generate HIV genomes with sufficient replication competence. Subsequent phenotypic analysis confirmed overcalling of X4-tropism for CRF01_AE viruses using the current version and the standard cut-off at 10% false positive rate (FPR) of geno2pheno[coreceptor]. Lowering the FPR cut-off to 2.5% reduced the X4-overcalling in our sample collection, while still allowing a safe administration of Maraviroc (MCV).

Conclusion: This study demonstrates the successful adjustment of geno2pheno[coreceptor] rules for subtype CRF01_AE. It also supports the unique strength of combining complementing methods, namely phenotyping and genotyping, for validating new bioinformatics tools prior to application in diagnostics.

1. Introduction

HIV-1 uses either CCR5 (R5-tropic virus) or CXCR4 (X4-tropic virus) as coreceptor for cell entry. Virus variants using the CCRR5-coreceptor generally predominate during early stages of infection. This has been assigned to the mechanism of entry, often involving monocytoid and other cells in mucosal tissue. As the HIV infection progresses in the absence of therapy, viral strains experience an increasing variability within the infected host, also with respect to the cellular tropism. In the late stages of infection, X4-tropic strains then become dominant in more than 50% of patients [1,2]. As far as studied so far, these coreceptor dynamics seem to be similar for the various HIV-1 subtypes.

The determination of coreceptor usage became clinically important in diagnostic settings when a mechanistically new antiretroviral drug, the entry inhibitor Maraviroc (MVC), was licensed for treatment of HIV patients. As MVC specifically blocks the CCR5 coreceptor but not CXCR4, a treatment decision for MVC requires prior determination of coreceptor usage in the blood of the respective patient [3].

Principally, co-receptor usage can either be analyzed functionally *in vitro*, using cell culture assays, termed phenotyping [4], or by sequence analysis of a specific envelope region (env-V3), termed genotyping [5,6]. Phenotyping tests often involve long turnaround times due to the need of sophisticated cell culture formats and can only be performed in a biosafety level 3 laboratory. Also, most tests use DNA recombination into an existing proviral backbone, mostly based on the genetic backbone of HIV-1 subtype B (e.g. NL4-3 or HXB2). Therefore, the analysis of clinical non-B subtype isolates *in vitro* may not be straight-forward.

One of the most widely used tools to genotypically predict tropism of HIV is the geno2pheno[coreceptor] web service [7]. Pairs of genotypic data and corresponding phenotypic information were used to develop and train the geno2pheno prediction system with machine learning methods. The resulting web tool geno2pheno[coreceptor] [8] has been validated in large subtype B studies, including MOTIVATE [9] and MERIT [10]. It allows for predicting the coreceptor usage based upon the V3 sequence of a given viral genome. The system uses a support vector machine to classify viruses as R5- or X4-capable based on informative patterns in the V3 sequence. HIV-1 isolates that do not exhibit sequence patterns indicative of R5 viruses are typically classified as X4-capable. Many viruses from divergent non-B strains of HIV have V3 sequences, which do not display strong sequence patterns of being R5-tropic and are therefore predicted as X4-capable. This is particularly true for subtype CRF01_AE viruses.

The HIV-1 subtype CRF01_AE, predominantly circulating in South-East Asia, is among those subtypes diverging the most from European subtype B viruses. Correlating the different genetic sequence, differences in the clinical properties have also been reported. It has been suggested that patients infected with subtype CRF01_AE may have a more rapid decline of CD4 + T cell count compared with patients infected with subtype B virus. Further, a shorter time to needing antiretroviral therapy and a higher virulence during the course of infection, have also been documented [11,12].

In a recent study by HIV-GRADE, analyzing the R5/X4-frequency in 2466 clinical HIV-1 isolates in Germany, the overall proportion of X4-tropic virus variants was found to be 15–30% overall, applying the 10% false positive rate (FPR) cut-off. However, while the X4/R5 ratio was observed in this range for most subtypes, this ratio was markedly different for samples of subtypes D and CRF01_AE. Here the study predicted an X4 frequency of 50% [13] (Fig. 1).

Potential reasons for the unexpectedly high frequency of X4-capabe virus include i) a suggested true higher prevalence of X4-capable viruses in CRF01_AE infected patients, or, alternatively, ii) a principal false overcalling in CRF01_AE isolates by the current geno2pheno[coreceptor] algorithm [14–16].

Matsuda et al. [17]. recently showed by phenotyping that for HIV-1 CRF01_AE there is indeed a significant X4-overcalling when using the 10% FPR cut-off of the classical version of geno2pheno[coreceptor]. As this algorithm has been used in several recent studies performed in South-East Asia [11,18,19], a thorough examination and, if needed, a correction of the geno2pheno tool for the genotypic prediction of CRF01_AE coreceptor-usage is urgently indicated.

The aim of this study was to provide the necessary verification and to provide a basis for adjustments of the geno2pheno tool for CRF01_AE in diagnostic settings.

2. Study design

Twenty patient-derived env (gp120) HIV-1 CRF01_AE samples from a cohort in Thailand were used for simultaneous phenotyping and genotyping (sequence analysis of the V3 region of the env gene of HIV-1) of these specimens. The samples were randomly chosen from 144 CRF01_AE plasma samples available through the Thailand's National HIV Drug Resistance Surveillance Program from a study among female sex workers [20]; informed consent and ethical approval from the responsible IHRP have been obtained (approval 3/2557).

The viral *env* region in patient-derived samples was amplified by RT-PCR and cloned into a pNL4-3 cassette or a newly designed CRF01_AE plasmid cassette where it reconstituted fully functional HIV genomes. The new CRF01_AE cassette has been made available through the portal of the European Horizon2020 project of EVAg.

After DNA transfection of 293 T cells in co-culture with the HIV-competent SXR5-reporter cell line, viral replication and syncytium formation were phenotypically determined in the presence of either the R5-

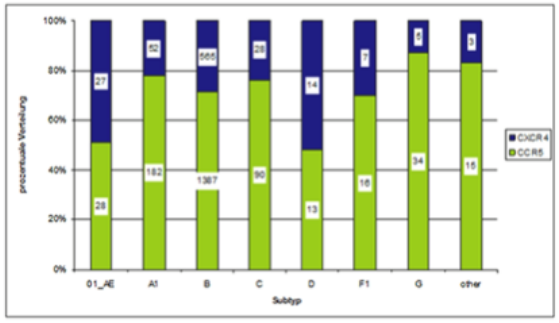

Fig. 1. Frequency of R5- (green) and X4-tropism (blue) by geno2pheno using the standard FPR cut-off of 10% [13] in 2466 treatment-experienced patients, for whom tropism testing was performed at baseline prior to potential Maraviroc administration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

* Corresponding author.
E-mail address: nina.marty@unibas.ch (N. Marty).
[1] Contributed equally.

https://doi.org/10.1016/j.jcv.2021.104755
Received 25 August 2020; Received in revised form 24 December 2020; Accepted 1 February 2021
Available online 9 February 2021

antagonist MVC or the X4-antagonist AMD3100.

Phenotypic results were compared and combined with genotypic predictions and used for adapting geno2pheno[coreceptor] to the available CRF01_AE samples in order to reduce the prior X4-overcalling.

3. Results

We had noticed early on that the HIV-1 genome reconstitution inserting exclusively the gp120 region from the samples from Thailand into the NL4-3 background, only a very low viral infection rate was obtained in cell culture. As strategy for improving viral competence, the replication properties of a whole array of recombinant HIV-1 clones, carrying various genomic segments of CRF01_AE-origin were compared side-by-side in the backbone of a prototypic subtype B virus (NL4-3). After each cloning step (initially only the entire env gene, then env plus vpu, then vpu plus nef and eventually the entire region from gag to env), replication of the resulting viral subtype B/AE recombinants were analyzed. As final result, the HIV-1 genome from the BssHII site at nt 712 to NgoMIV at nt 8338 (pNL-AE-K7_short) or to XmaI at nt 8888 (pNL-AE-K7) was substituted in frame by patient-derived CRF01_AE sequences, retaining only the LTRs and the 3′ end of *nef* of NL4-3. Into pNL-AE-K7 we inserted the respective *env* sequences from clinical specimens in this study. This allowed to phenotypically re-assess the tropism of the respective patient-derived viral envelopes. Passage of cell-free recombinant virus onto human lymphocytes verified the infectivity of progeny CRF01_AE viruses, carrying patient-derived HIV-1 *env* sequences. By comparing viral V3 sequences after 30 days of passaging to the corresponding patient V3 sequences, possible mutations and contamination could be excluded.

Using the new pNL-AE-K7 we were able to determine the phenotype initially in 20 clinical samples (Table 1, column "Phenotype") by judging drug-based inhibition of viral replication and potential syncytia formation in the presence of either the R5-antagonist MVC or the X4-antagonist AMD3100. In this assessment using a virus-replication system, only one sample (Th026) was found to be X4-tropic, while 18 samples were determined to contain R5-tropic virus. For one sample (Th049), no clear tropism determination was possible, since small fusion events of 2–3 HIV-infected cells had formed in the cultures in the presence of either inhibitor. For this case, the presence of a dual-tropic virus could not be excluded.

Noteworthy, for all tested B/AE-recombinants the average syncytium sizes and the overall number of viral infection events in the culture dish remained low (approximately 10% of the control) when compared to the control plasmid pNL-NF.

In parallel, the most prevalent genotype present in these 20 clinical samples was predicted using the standard version of geno2pheno[coreceptor] (Table 1, columns "g2p"). When linking these results to the phenotypic findings, the suspected systematic overcalling of X4-tropism in subtype CRF01_AE by the current version of the geno2pheno[coreceptor] tool became apparent, reaching only a low assay specificity of 66% when the standard FPR cut-off of 10%, was used. By lowering the FPR cut-off to 2.5% the specificity increased to 89%.

For confirmation beyond the small initial data set from Thailand, the newly suggested CRF01_AE-specific FPR cut-off of 2.5% was re-applied to a large data set from a German HIV-GRADE cohort on CRF01_AE samples (Table 2). When applying this new rule to all available CRF01_AE isolates, the significant discrepancy in the X4/R5 tropism ratio for CRF01_AE isolates, as depicted in Fig. 1, completely disappeared and rendered this subtype similar to the general, subtype independent distribution of clinical samples.

When disregarding differences between subtypes, the overall tropism distribution across all isolates (including all subtypes) would be 72% R5 and 28% X4. When the 10% FPR cut-off was applied specifically to subtype CRF01_AE isolates, this ratio shifts to 51% R5 and 49% X4, indicating a dramatic deviation with a Chi^2 of <0.001. When we now apply the phenotype-supported new "CRF01_AE-FPR" cut-off of 2.5%, a 76% R5 and 24% X4 distribution is seen for the CRF01_AE isolates with a Chi^2 value of 0.43, which is no longer significantly different from the calculated global average of isolates irrespective of their subtype.

For verification and potential fine-tuning, our phenotype-matched geno2pheno values were also subjected to the FPR cut-offs of 1% and 5% as well as to the FPR cut-off of 3.75%, which is used for NGS data. No significant difference from the expected distribution was seen for cut-offs at and above 2.5%.

4. Discussion

In this study, phenotypically determined co-receptor usage was compared to and combined with genotypic data to improve the prediction of geno2pheno[coreceptor] for subtype CRF01_AE isolates of HIV-1.

For the phenotypic determination of the co-receptor usage major challenges had to be overcome. A recombinant plasmid-based system (pNL-K7), previously developed by our group [21], was used to reconstitute HIV-1 variants. This cassette permits the exchange of env segments by cleavage with unique restriction endonucleases and placing PCR-amplified HIV-1 env derived from patient plasma directly into a complete viral genome. After transfection into a human indicator cell line, viral replication of the recombinant HIV-1 variant in the presence of inhibitors can be quantitatively analyzed [22]. One hurdle in this process was a poor PCR amplification rate of the env fragments. Standard PCR-primers were derived from a reliably working subtype B consensus sequence. The observed low amplification rates were a strong indicator for the vast sequence heterogeneity of our HIV-1 isolates in the viral env region, suggesting that the validated recombination protocol at predefined sites in Env may not be optimal for generating replicating viral subtype CRF01_AE genomes. Another technical hurdle appeared to be the low replicative fitness of recombinant HIV genomes encountered when using the in-house subtype B-based HIV-1 cassette (pNL-K7). We attributed the poor replicative capacity with previously reported observations that Env may critically depend on interactions with subtype-matched corresponding regions in Gag-Pol [23,24]. It is further possible that Env functions best in a subtype-unique context including its own co-evolved Vpu [25] or other viral proteins [22,26]. To improve the replicative fitness, we therefore designed a new cassette, carrying a near full-length subtype CRF01_AE backbone. With this construct we were able to obtain sufficiently replicative virus to phenotypically determine the tropism of CRF01_AE patient samples.

Based on our comparison between phenotypically and genotypically determined tropism, our study supports implementing a significantly lower FPR cut-off of 2.5% (compared to the standard of 10%) as a critical adjustment for appropriate tropism prediction for CRF01_AE samples by geno2pheno[coreceptor]. The suggestion to lower the FPR cut-off for CRF01_AE virus variants is supported by others: One study had shown for CRF01_AE samples a specificity of only 50% at a 10% FPR cut-off, whereas the specificity increased to 77% by lowering the FPR cut-off to 5% [27]. Another study concluded in a comparison of different genotypic tools that for clinical practice, a geno2pheno[coreceptor] FPR cut-off of 5% could be used to predict CRF01_AE tropism [28].

Also, for genotyping of clinical samples using deep V3 sequencing (NGS), the interpretation of the analysis combines the information on FPR on each of the sequences and the corresponding frequency of these different variants in the sample. Currently, this two-dimensional cut-off predicts a sample with >2% of the variants with an FPR < 3.5% as not suitable for maraviroc treatment. This recommendation is so far independent of the HIV subtype [29]. We re-adjusted the FPR-value for a specific variant, namely for subtype CRF01_AE, in order to improve the clinical application of geno2pheno[coreceptor] also for the use for consensus, Sanger-sequencing.

As the R5-antagonist MVC proved to be a well-tolerated drug, lowering the FPR cut-off would potentially allow for more patients benefitting from MVC administration, especially in South-East Asia. Therefore, taking our results into consideration, we suggest setting an FPR cut-off of 2.5% for the tropism prediction of clinical subtype CRF01_AE samples by geno2pheno[coreceptor]. However, further studies on a larger cohort are needed to verify this suggestion.

One proposed reason for the observed X4-overcall using the standard FPR cut-off of 10% is a difference in common sequence motifs. The typical CRF01_AE envelope contains several otherwise uncommon amino acids in the V3 region as unique and inherent feature. The motif GPGQVF at the tip of the V3 loop occurred very frequently in the HIV-1 CRF01_AE samples of this study (Fig. 2).

Using the standard version of geno2pheno[coreceptor], these samples were often predicted to belong to X4-tropic HIV-1 isolates in contrast to samples with the GPGRAF motif being frequent in R5-variants of subtype B. This significant deviation raised the speculation that the GPGQVF motif alone might result in incorrect X4 predictions. However, when taking the raw data of the prediction system into account, also additional minor changes outside this very tip region contribute to the X4-overcalling of subtype CRF01_AE variants.

For further improvement of geno2pheno[coreceptor], more information about specific V3-loop characteristics of subtype CRF01_AE variants will be incorporated in future versions of the algorithm.

4.1. Study limitation

Although the results are clear within experimental setting, the low number of X4-tropic samples, identified within the study, poses a limitation. One limitation is that most samples appeared to be R5-tropic by phenotypic determination although the genotypic prediction was X4-tropic for several samples in the investigated cohort. However, a relatively high proportion of R5-tropic isolates is common for HIV-1 in ART treated individuals, consistent with results reported by Cui et al. [31]. When including our findings reported in Matsuda et al. [17], only two samples were phenotypically identified as X4-tropic, which reflects an overall percentage of 4.8% (2 out of 42 samples). A validation with more X4-tropic samples is recommended to corroborate our results.

It should be noted, although gender is currently not known to play a

Table 1
Overview of results. The phenotyped CRF01_AE-samples with their confirmed sequence (V3-loop) and the genotypically predicted respective tropism. Blue = X4-tropic, green = R5-tropic, ND = not determined.

Sample	V3-LOOP	Phenotype	g2p: 10% FPR	g2p: 5% FPR	g2p: 2.5% FPR	g2p: 1% FPR
TH026	CTRPSKTVR-SMRIGPGKVFYRIEGIGGDIRKAYC	X4	0.3	0.3	0.3	0.3
TH012	CTRPSNNIRTSMTVGPGQVIYKTGDITGNIRKARC	R5	1.7	1.7	1.7	1.7
TH016	CTRPSNYTRTSTRIGPGQVWYRTGDIIGNPRKAYC	R5	1.7	1.7	1.7	1.7
TH034	CTRPSNNTRTSIGIGPGQVFYRTGDIIGDIRRAYC	R5	2.6	2.6	2.6	2.6
TH010	CTRPSNNIRTSVHIGPGQVFYKTGDIIGDIRQAHC	R5	4.7	4.7	4.7	4.7
TH041	CTRPSTNRRTGVHIGPGQVFYRTGEIIGDIRKAYC	R5	7.2	7.2	7.2	7.2
TH023	CTRPSNNTRTSSRIGPGAVFYRTGDITGDIRQAHC	R5	8.7	8.7	8.7	8.7
TH020	CTRPSNNTRTSVTMGPGHVFYRTGDIIGDIRKAHC	R5	15.5	15.5	15.5	15.5
TH043	CTRPSNNTRTSMTIGPGQVFYRTGDIIGDIRKAYC	R5	16.6	16.6	16.6	16.6
TH037	CTRPSNNTRTS-HIGPGQVFYRTGDIIGDIRKAHC	R5	17.8	17.8	17.8	17.8
TH040	CTRPSNNTRTSIPIGPGQVFYKTGEIIGDIRKAYC	R5	20.7	20.7	20.7	20.7
TH028	CTRPFNNTRTSITIGPGQMFYRTGEIIGDIRKAYC	R5	24.9	24.9	24.9	24.9
TH038	CTRPSNNTRTSITIGPGQVFYRTGDITGDIRKAFC	R5	25.4	25.4	25.4	25.4
TH005	CTRPSNNTRTSITIGPGQVFYRTGDIIGDIRKAYC	R5	33.1	33.1	33.1	33.1
TH046	CTRPSNNTRKGIHLGPGQVFYRTGDIIGDIRQAYC	R5	35.7	35.7	35.7	35.7
TH022	CTRPSNNTRQSINIGPGRVFYRPGDIIGDIRKAYC	R5	44.9	44.9	44.9	44.9
TH039	CTRPSNNTRTSVHIGPGQVFYRTGDIIGDIRKAYC	R5	44.9	44.9	44.9	44.9
TH049	CTRPSNNTRTSIHMGPGQVFYRTGEIIGDIRQAHC	ND	50.9	50.9	50.9	50.9
TH032	CTRPSNNTRTSITMGPGQVLYRTGDITGDIRKAYC	R5	63.6	63.6	63.6	63.6
TH035	CTRPSNNTRKSVPIGPGQVFYRTGDIIGDIRQAHC	R5	89.6	89.6	89.6	89.6

X4
R5

Table 2
The frequency of X4 in the patients with subtype CRF01_AE in the HIV-GRADE study was likely to be overstated with an FPR cut-off of 10% in comparison to the frequencies of the other subtypes. Lowering the FPR to the cut-off of 2.5% as phenotypically determined in this analysis, the subtype CRF01_AE-specific polymorphisms were correctly accommodated. The relevant FPR range between 2.5% and 5% (Chi [2] >0.2) has been shaded.

FPR cut-off	CCR5	CXCR4	Chi^2
Expected normalized distribution	*72%*	*28%*	
10%	51%	49%	<0.001
5%	64%	36%	0,22
3.75%	69%	30%	0,68
2.50%	76%	24%	0,43
1%	89%	11%	0,002

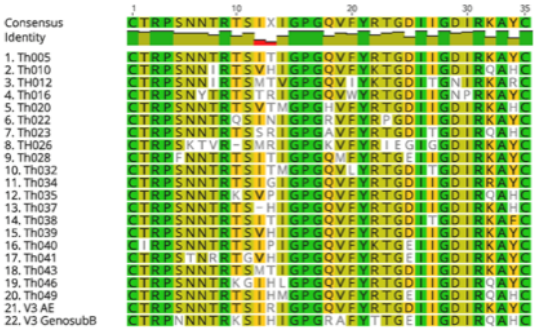

Fig. 2. Comparison of CRF01_AE patient V3 sequence (Th0xx samples), the reference sequence of CRF01_AE (V3 AE) and the reference sequence of subtype B used for the training of geno2pheno[coreceptor] [30] (V3 GenosubB). Green = 100% similar, olive = 80%-100% similar, yellow = 60%-80% similar, white = less than 60% similar.

role in the tropism distribution, that all specimens of this study came from female sex workers in a cohort in Thailand.

5. Conclusion

By combining genotypic data with phenotypically determined results, this study demonstrates the previously suspected systematic overcalling of X4-tropism in subtype CRF01_AE by the current version of the geno2pheno[coreceptor] tool. A suitable solution was obtained by adjusting the FPR cut-off for CRF01_AE samples to 2.5%, This results in a correct tropism prediction for clinical CRF01_AE samples and may guide the safe beneficial use of MVC in a broader group of HIV infected individuals, who carry virus of this subtype.

Declaration of Competing Interest

None.

Acknowledgments

This study received financial support from ViiV Healthcare. ViiV Healthcare did not have any role in the design or conduct of the work and did not take any influence on conclusions or manuscript writing.

References

[1] R.R. Regoes, S. Bonhoeffer, The HIV coreceptor switch: a population dynamical perspective, Trends Microbiol. 13 (2005) 269–277, https://doi.org/10.1016/j.tim.2005.04.005.
[2] D.E. Mosier, How HIV changes its tropism: evolution and adaptation? Curr. Opin. HIV AIDS 4 (2009) 125–130, https://doi.org/10.1097/COH.0b013e3283223d61.
[3] M. Perez-Olmeda, J. Alcami, Determination of HIV tropism and its use in the clinical practice, Expert Rev. Anti. Ther. 11 (2013) 1291–1302, https://doi.org/10.1586/14787210.2013.852469.
[4] A.B. Siddik, et al., Phenotypic co-receptor tropism and Maraviroc sensitivity in HIV-1 subtype C from East Africa, Sci. Rep. 8 (2018) 2363, https://doi.org/10.1038/s41598-018-20814-2.
[5] L.P. Vandekerckhove, et al., European guidelines on the clinical management of HIV-1 tropism testing, Lancet Infect. Dis. 11 (2011) 394–407, https://doi.org/10.1016/S1473-3099(10)70319-4.
[6] S. Edwards, et al., A diagnostic HIV-1 tropism system based on sequence relatedness, J. Clin. Microbiol. 53 (2015) 597–610, https://doi.org/10.1128/JCM.02762-14.
[7] N. Beerenwinkel, et al., Geno2pheno: estimating phenotypic drug resistance from HIV-1 genotypes, Nucleic Acids Res. 31 (2003) 3850–3855, https://doi.org/10.1093/nar/gkg575.
[8] T. Lengauer, O. Sander, S. Sierra, A. Thielen, R. Kaiser, Bioinformatics prediction of HIV coreceptor usage, Nat. Biotechnol. 25 (2007) 1407–1410, https://doi.org/10.1038/nbt1371.
[9] R.A. McGovern, et al., Population-based V3 genotypic tropism assay: a retrospective analysis using screening samples from the A4001029 and MOTIVATE studies, AIDS 24 (2010).
[10] L.C. Swenson, et al., Deep V3 sequencing for HIV type 1 tropism in treatment-naive patients: a reanalysis of the MERIT trial of maraviroc, Clin. Infect. Dis. 53 (2011) 732–742, https://doi.org/10.1093/cid/cir493.
[11] M. Chu, et al., HIV-1 CRF01_AE strain is associated with faster HIV/AIDS progression in Jiangsu Province, China, Sci. Rep. 7 (2017) 1570, https://doi.org/10.1038/s41598-017-01858-2.
[12] O.T. Ng, et al., Increased rate of CD4+ T-Cell decline and faster time to antiretroviral therapy in HIV-1 subtype CRF01_AE infected seroconverters in Singapore, PLoS One 6 (2011) e15738, https://doi.org/10.1371/journal.pone.0015738.
[13] M. Obermeier, J. Symons, A.M. Wensing, HIV population genotypic tropism testing and its clinical significance, Curr. Opin. HIV AIDS 7 (2012) 470–477, https://doi.org/10.1097/COH.0b013e328356eaa7.
[14] S. Raymond, et al., Genotypic prediction of HIV-1 CRF01-AE tropism, J. Clin. Microbiol. 51 (2013) 564–570, https://doi.org/10.1128/JCM.02328-12.
[15] A. Phuphuakrat, et al., Coreceptor tropism determined by genotypic assay in HIV-1 circulating in Thailand, where CRF01_AE predominates, HIV Med. 15 (2014) 269–275, https://doi.org/10.1111/hiv.12108.
[16] M. Mulinge, et al., HIV-1 tropism determination using a phenotypic Env recombinant viral assay highlights overestimation of CXCR4-usage by genotypic prediction algorithms for CRF01_AE and CRF02_AG [corrected], PLoS One 8 (2013) e60566, https://doi.org/10.1371/journal.pone.0060566.
[17] M. Matsuda, et al., Performance evaluation of a genotypic tropism test using HIV-1 CRF01_AE isolates in Japan, Jpn. J. Infect. Dis. 71 (2018) 264–266, https://doi.org/10.7883/yoken.JJID.2017.482.
[18] X. Li, et al., Evidence that HIV-1 CRF01_AE is associated with low CD4+T cell count and CXCR4 co-receptor usage in recently infected young men who have sex with men (MSM) in Shanghai, China, PLoS One 9 (2014) e89462, https://doi.org/10.1371/journal.pone.0089462.
[19] Y. Li, et al., CRF01_AE subtype is associated with X4 tropism and fast HIV progression in Chinese patients infected through sexual transmission, AIDS 28 (2014) 521–530, https://doi.org/10.1097/qad.0000000000000125.
[20] S. Saeng-aroon, et al., Circulation of HIV-1 multiple complexity recombinant forms among female sex workers recently infected with HIV-1 in Thailand, AIDS Res. Hum. Retroviruses 32 (2016) 694–701, https://doi.org/10.1089/aid.2015.0371.
[21] J. Fehr, et al., Replicative phenotyping adds value to genotypic resistance testing in heavily pre-treated HIV-infected individuals - the Swiss HIV Cohort Study, J. Transl. Med. 9 (2011) 14, https://doi.org/10.1186/1479-5876-9-14.
[22] T. Klimkait, F. Stauffer, E. Lupo, C. Sonderegger-Rubli, Dissecting the mode of action of various HIV-inhibitor classes in a stable cellular system, Arch. Virol. 143 (1998) 2109–2131, https://doi.org/10.1007/s007050050447.

[23] J.C. Gardiner, E.J. Mauer, N.M. Sherer, HIV-1 Gag, envelope, and extracellular determinants cooperate to regulate the stability and turnover of virological synapses, J. Virol. 90 (2016) 6583–6597, https://doi.org/10.1128/JVI.00600-16.
[24] J. Chojnacki, et al., Envelope glycoprotein mobility on HIV-1 particles depends on the virus maturation state, Nat. Commun. 8 (2017) 545, https://doi.org/10.1038/s41467-017-00515-6.
[25] S. Schwartz, B.K. Felber, E.M. Fenyo, G.N. Pavlakis, Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs, J. Virol. 64 (1990) 5448–5456.
[26] S. Wildum, M. Schindler, J. Munch, F. Kirchhoff, Contribution of Vpu, Env, and Nef to CD4 down-modulation and resistance of human immunodeficiency virus type 1-infected T cells to superinfection, J. Virol. 80 (2006) 8047–8059, https://doi.org/10.1128/JVI.00252-06.
[27] S. Hongjaisee, C. Nantasenamat, T.S. Carraway, W. Sboombustong, HIVCoR: a sequence-based tool for predicting HIV-1 CRF01_AE coreceptor usage, Comput. Biol. Chem. 80 (2019) 419–432, https://doi.org/10.1016/j.compbiolchem.2019.05.006.
[28] C. Soulie, et al., Performance of genotypic algorithms for predicting tropism for HIV-1 CRF01_AE recombinant, J. Clin. Virol. 99-100 (2018) 57–60, https://doi.org/10.1016/j.jcv.2017.12.014.
[29] L.C. Swenson, et al., Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients, J. Infect. Dis. 203 (2011) 237–245, https://doi.org/10.1093/infdis/jiq030.
[30] T. Sing, et al., Predicting HIV coreceptor usage on the basis of genetic and clinical covariates, Antivir. Ther. 12 (2007) 1097–1106.
[31] H. Coi, et al., Rapid CD4+ T-cell decline is associated with coreceptor switch among MSM primarily infected with HIV-1 CRF01_AE in Northeast China, AIDS 33 (2019) 13–22, https://doi.org/10.1097/qad.0000000000001981.