博士論文

Analysis of HIV-1 co-receptor usage by cell fusion based assay using cloned full-length envelope

（HIV-1 全長 envelope クローンを用いた細胞融合アッセイによる co-receptor usage の解析）

Wei Yazhi

魏 雅稚

Analysis of HIV-1 co-receptor usage by cell fusion based assay using cloned full-length envelope

HIV-1 全長 envelope クローンを用いた 細胞融合アッセイによる co-receptor usage の解析

所属　　東京大学大学院医学系研究科　國際保健学　専攻

指導教員　　岩本　愛吉　教授

Wei Yazhi

魏（ギ）　雅稚（ガチ）

Abbreviations

AIDS	Acquired Immune Deficiency Syndrome
ART	antiretroviral therapy
CA	capsid protein
CCR5	C-C motif chemokine receptor 5
C1~5	constant region 1~5
CXCR4	C-X-C motif chemokine receptor 4
DC-SIGNR	dendritic cell-specific intercellular adhesion-molecule-3-grabbing non-integrin
df	degrees of freedom
DDJB	DNA DATA BANK of Japan
DHHS	Department of Health and Human Services
D/M	dual or mixed
DMEM	Dulbecco's modified Eagle's medium
DSP	dual split protein
DSP_{1-7}	dual split protein $_{1-7}$
DSP_{8-11}	dual split protein $_{8-11}$
Env	Envelope

ER	endoplasmic reticulum
ES Trofile	enhanced sensitivity Trofile HIV tropism assay
FBS	fetal bovine serum
FDA	Food and Drug Administration
FPR	false positive rate
GTAs	genotypic assays
HIV-1	human immunodeficiency virus type 1
HRD	HIV related drugs
IN	integrase
IMSUT	the Institute of Medical Science, the University of Tokyo
UNAID	Joint United Nation Programme on HIV/AIDS
LTR	long terminal repeat
MA	matrix protein
MDM	monocyte-derived macrophages
MEM	modified Eagle's medium
M-tropic	macrophage-tropic
MVC	maraviroc
NC	nucleocapsid

NFAT	nuclear factor of activated T cells
NF-κB	nuclear transcription activator
PBMC	peripheral blood mononuclear cells
Pol	polymerase
PR	protease
PTAs	phenotypic assays
RT	reverse transcriptase
SDF-1	stromal cell-derived factor-1
spGFP	split green fluorescent protein
spRL	split Renilla luciferase
SU	surface protein
TM	transmembrane protein
V1~5	variable region 1~5

Abstract

Human immunodeficiency virus type 1 (HIV-1) uses two receptors upon infection: Cell surface CD4 molecule as a primary receptor and either CXCR4 or CCR5 as a co-receptor. Based on the co-receptor it uses, HIV-1 is categorized as an X4 virus, an R5 virus, or an R5X4 virus.

Co-receptor usage can be estimated by a phenotypic assay based on virus-to-cell infection (virus-cell fusion assay) utilizing pseudoviruses. Trofile™ (Monogram Biosciences Inc., California, USA) is an example. The virus-cell fusion assay mimics the natural infection most accurately, but it has shortcomings such as biosafety considerations and a long turnaround time. Genotypic co-receptor assay utilizing amino acid sequences of the viral envelope has also been developed. While a genotypic assay such as the Geno2pheno assay is safe, rapid and universally accessible, its calculation depends on short amino acid sequences in the V3 region of the envelope. Calculation depends also on historical data with known genotype and phenotype. Therefore, a simple, safe and rapid phenotypic assay is in need. Our laboratory has reported a new rapid and safe co-receptor assay based on cell-cell fusion (DSP-Pheno). In an effort to compare the validity of the three assays, I first applied the envelope glycoprotein gene amplified directly from the plasma of six patients to an in-house virus-cell fusion assay,

Geno2pheno assay, and cell-cell fusion assay. The Geno2pheno assay predicted R5 co-receptor usage in all six patients; however, the results of the phenotypic assays using the whole envelope protein were quite different: The virus-cell fusion assay predicted 3 R5 and 3 R5X4, and the cell-cell fusion assay predicted that all 6 were R5X4 viruses. Patients' plasma viruses are composed of quasispecies, therefore I isolated 25 clones of the whole envelope gene from the 6 patients and tested their co-receptor usage. The Geno2pheno assay (using False Positive Rate <10), predicted that 17 or 8 clones could use CCR5 or CXCR4, respectively. An in-house virus-cell fusion assay, which is supposed to reflect the natural infection, predicted that all 25 clones could use CCR5 and that only 4 out of 8 clones predicted by Geno2Pheno could use CXCR4. These results showed that envelope regions outside V3 should be considered in predicting the co-receptor usage. In patient IMS0718, the amino acid mutation correlating R5 to R5X4 change distributed in C1, C2, and V3, however the frequency of the mutation was not statistically significant. In patient IMS1000, 11 amino acid differences accumulated in C1, V2 and C3. Among them, the clustering of amino acid mutation in V2 was statistically significant.

In the cell-cell fusion assay, 20/25 clones were predicted as R5X4 virus while only 5 clones were predicted as R5 virus. By comparison between the virus-cell fusion

assay and the cell-cell fusion assay, the amino acid mutation correlating the R5 to R5X4 change were scattered in C1, V1, C3, V4, and V5.

These results suggested first that the amino acid change in V2 which covers V3 before CD4 binding could influence the subsequent co-receptor binding. Second, it is likely that the cell-cell fusion assay might over-estimate the CXCR4 usage and it would be difficult to apply the cell-cell fusion assay to the clinical decision on the use of CCR5 antagonist, maraviroc. However, the cell-cell fusion assay may be useful for the research on cell-to-cell infection of HIV-1.

Introduction

Human immunodeficiency virus (HIV) is a pathogen that causes Acquired Immune Deficiency Syndrome (AIDS). It was estimated that 35 million people were living with HIV globally by the end of 2013. About 24.7 million (70%) or 4.8 million people (14%) lived in sub-Saharan Africa or in Asia and the Pacific, respectively. Also, an estimated 2.1 million new infections and 1.5 million AIDS-related deaths were reported by WHO and UNAIDS globally in 2013. The cumulative estimated number of HIV carriers and AIDS patients was 437,000 at the end of 2013 in China.

AIDS was first reported in Los Angeles, United States in 1981 [1]. The hallmark of the laboratory test was the virtual elimination of the Leu-3+ helper/inducer subset, which are now known as $CD4^+$ T cells. Françoise Barré-Sinoussi, et al. isolated a new retrovirus in 1983, which was later named HIV type 1 (HIV-1) [2]. The genomic organization of HIV-1 contains three structural genes typical of a retrovirus, *gag*, *pol*, *env*, and six accessory genes, *vif*, *vpr*, *vpu*, *tat*, *rev* and *nef* [3]. After successful entry into the cytoplasm, the viral core is transported to the cellular nucleus while viral RNA genomes are reverse-transcribed to the DNA by the viral reverse transcriptase (RT) encoded in the *pol* gene (Fig. 1). Viral DNA is then integrated into the host chromosome by the viral integrase (IN) that is also encoded in the *pol* gene. Integrated viral DNA is transcribed into RNA, which is used for the genome of a progeny virus or messenger RNA to be

translated into viral structural proteins (Gag) and enzymes (Pol). After the translation on ribosomes associated with the rough endoplasmic reticulum (ER), the Env precursor protein (gp160) is transported to the Golgi apparatus and cleaved by a host protease into two mature envelope glycoproteins, surface protein (SU) gp120 and transmembrane protein (TM) gp41. Gp120 is expressed on the cell surface via transportation through the intracellular membrane trafficking system. Structural proteins and enzymes form the viral core and are budded from the cell membrane coated with the envelope glycoproteins. On the surface of the cell membrane and virion membrane, mature glycoproteins are present as the trimeric form of gp120 and gp41.

During the HIV-1 replication cycle, the first key step is the binding of the viral gp120 to the cellular receptor on the cell membrane. Since the $CD4^+$ T cell count was significantly deleted when AIDS was first described [1], the CD4 molecule was the first candidate for the cellular receptor for HIV-1 and actually it was; however, CD4 was not sufficient for HIV-1 entry in murine or non-human primate models [4, 5]. A breakthrough experiment using a novel functional complementary DNA (cDNA) cloning strategy identified a G protein-coupled receptor with seven transmembrane domains as a potentially candidate for co-receptor and was named “fusin” [6]. Fusin was later renamed as C-X-C motif chemokine receptor 4 (CXCR4). A CXC chemokine, stromal cell-derived

factor-1 (SDF-1), secreted by a lymphocyte was proven to be a potent inhibitor of HIV-1 entry [7]; however, the identification of CXCR4 as a co-receptor of HIV-1 in addition to the principal receptor, CD4, could not explain the whole picture. Laboratory HIV-1 strains could infect and fuse $CD4^+$ T cell lines that express CXCR4 but primary HIV-1 isolates from the patients could not. Primary HIV-1 isolates showed preferential growth in monocyte-derived macrophages or primary T-cells and did not induce cell fusions. Therefore, laboratory HIV-1 strains were often referred to as T-cell-line-tropic (or fusogenic) HIV-1, while primary isolates were called macrophage-tropic (or non-fusogenic) HIV-1. Another breakthrough in macrophage-tropic HIV-1 was published in 1995. Interestingly, ß-chemokines, such as RANTES, MIP-1α, and MIP-1β, which are the ligands for C-C motif chemokine receptor 5 (CCR5), could suppress the macrophage-tropic HIV-1 [8]. CCR5, a G protein-coupled receptor for a ß-chemokine with seven transmembrane domains, was the co-receptor for the macrophage-tropic HIV-1 [9, 10]. CCR5 is expressed on rather limited cell types, such as monocytes, macrophages, and primary T cells, but CXCR4 is more ubiquitously present on $CD4^+$ T cells, $CD4^+$ T-cell-lines, and other cell types [11].

Berger et al. proposed a new classification for HIV-1 based on co-receptor usage [12]. HIV-1 isolates that use CCR5 but not CXCR4 are termed R5 viruses, isolates using

CXCR4 but not CCR5 are designated as X4 viruses, and isolates able to use both co-receptors with comparable efficiency are called R5X4. (Fig. 2). R5 viruses are the strains most commonly transmitted sexually, which is consistent with the high resistance of infection for individuals lacking CCR5 [13, 14]. A cross-sectional analysis in the early stages of the infection showed that 70-80% of the population with HIV-1 was due to R5 [15]. Individuals harboring X4 viruses had poorer baseline clinical profiles (higher viral load and lower CD4 count) than individuals harboring exclusively R5 viruses [16]. While R5 viruses are detectable over the entire clinical course of HIV-1 infection [17], X4 viruses emerge in 40-50% of infected persons during the clinical course [18]. Why R5 viruses grow preferentially in the early clinical course is still an open question. The appearance of R5X4 or X4 viruses in a patient exacerbates the clinical progression leading to AIDS, since HIV-1 can kill CD4+ T cells more efficiently by taking advantage of the expression of CXCR4 in wider subsets of $CD4^+$ T cells. Therefore, the co-receptor usage of HIV-1 is highly related to the pathogenesis of HIV/AIDS and needs further research.

The crucial role of the V3 region in gp120 was originally considered to be the determinant for T-cell-line-tropic or macrophage-tropic HIV-1 and was designated as the co-receptor binding site [19, 20]. The molecules playing a key role in the binding of

HIV-1 gp120 to cellular receptors are schematically shown in Fig. 3. Upon binding to CD4, gp120 undergoes a major conformational change exposing the V3 region necessary for the co-receptor binding (Fig. 4). The binding of the V3 region of the viral envelope and the cellular co-receptor induce a further conformational change in gp120 as well as in gp41. The fusogenic domain with hydrophobic amino acids in the transmembrane protein gp41 (green bars in Fig.4) plays a crucial role in the membrane fusion between the virus membrane (envelope) and the cell membrane, though the precise mechanism is unknown.

Another important feature of HIV-1 is its tremendous genetic heterogeneity. Highly error-prone RT produces a diverse and changing HIV-1 population and forms quasi-species in patients [21, 22]. It is noteworthy that a highly variable region, such as V3, serves as a co-receptor binding function. In addition, HIV-1 co-receptor usage has implications of treatment and/or prophylaxis. The Food and Drug Administration (FDA) approved an entry inhibitor, maraviroc, for clinical use in 2007. Maraviroc is a specific antagonist against CCR5 and inhibits HIV-1 entry into the cells [23] (Fig. 2). Understandably, maraviroc was not effective for treatment-experienced patients infected with non-R5 HIV-1 [24, 25]. Although the CXCR4 specific entry inhibitor AMD3100 is available for experiments, its clinical development has been discontinued due to hepatotoxicity.

A tropism assay is mandatory before the initiation of treatments if maraviroc is included. Although HIV-1 co-receptor usage can be predicted by a phenotypic assay or a genotypic assay, there are pros and cons of each assay (Fig. 5). From 2007 to 2009, the phenotypic assay used in association with maraviroc treatment was solely Trofile™ (Monogram Biosciences Inc, California, USA), a $CD4^+$ T cell culture assay using replication-defective but infectious pseudoviruses [26]. The method of infecting cells with pseudoviruses (virus-cell fusion assay) was believed to best mimic the natural infection. Actually, according to the present guidelines, a phenotypic assay is preferred rather than a genotypic assay to predict HIV-1 co-receptor usage in patients (http://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-arv-guidelines/8/co-receptor-tropism-assays); however, due to some inconveniences that are shown in Fig. 5, the phenotypic co-receptor assay has not been widely used for clinical practice [27]. The genotypic assay, which is based on amino acid sequences in the V3 region, was developed for wider use. The genotypic co-receptor assay is a rapid and widely applicable test; however, its dependence solely upon the V3 amino acid sequences does not guarantee the co-receptor binding activity of the whole envelope protein that patients' viruses may have. Our laboratory developed and reported a new phenotypic HIV-1 co-receptor assay based on cell-cell fusion, called DSP-Pheno [28]. The assay takes advantage of a newly

developed cell-cell fusion assay system called Dual-Split Protein (DSP) [29]. DSP is a two-way assay system of cell fusion employing Renilla luciferase (RL) and green fluorescent protein (GFP). (Fig. 6). DSP-Pheno is a safe and rapid phenotypic co-receptor assay; however, its application in clinical samples should be validated carefully. Moreover, the results obtained by the DSP-Pheno may overestimate the CXCR4 usage of HIV-1 isolates compared to virus-cell fusion (unpublished observation, Hosoya, N.).

The aim of my experiments was to apply three available co-receptor assays to clinical samples: one genotypic assay (Geno2Pheno) and two phenotypic assays based on virus-cell fusion or cell-cell fusion. Since envelope genes that are amplified by the polymerase chain reaction (PCR) include mixed envelope sequences (quasi-species), I isolated clonal envelope genes from the patients and tested them as well. Through the comparison between an in-house virus-cell fusion assay and a Geno2Pheno assay, I identified a region outside of V3 significantly correlated with the change in co-receptor usage. Then, I compared the two phenotypic co-receptor assays with the clinical samples to identify the role of amino acid mutations which correlated with the difference in co-receptor usage between virus-cell fusion and cell-cell fusion assays.

Materials and Methods

Cell lines

There were three cell lines in the experiments: the 293FT cell line, which served as an HIV-1 Env-expressing cell line, and human glioma derived NP2 cell lines - N4X4 cells and N4R5 cells, which served as the receptor-and-co-receptor-expressing cell line. The N4X4 cells were based on NP2 cells (received from professor Hiroo Hoshino, Gunma University, Gunma, Japan), which stably expressed CD4 and CXCR4 on the surface of the cell. The N4R5 cells were based on NP2 cells, which stably expressed CD4 and CCR5 on the surface of the cell. N4X4 and N4R5 both expressed a Dual split protein$_{1\text{-}7}$ (DSP$_{1\text{-}7}$, 70% split green fluorescent protein and 70% split Renilla luciferase) stably.

The 293FT cells (Invitrogen, California, USA) were cultured in a D10+ medium composed of Dulbecco's modified Eagle's medium (DMEM; Sigma) supplemented with 10% FBS, 100 units/ml of penicillin, and 0.1 mg/ml of streptomycin. The N4X4 and N4R5 cells were maintained in a M10+4 medium composed of modified Eagle's medium (MEM; Sigma, Missouri, USA), 10% heat inactivated fetal bovine serum (FBS), 100 units/ml of penicillin, 0.1 mg/ml of streptomycin, and 4 mg/ml of blasticidin. All cell cultures were maintained at 37℃ in a humidified 5% CO_2 incubator.

Clinical samples

There were six patients enrolled in this study. Treatment-naïve clinical samples from the patients were used. This study has been approved by the institutional review board of the Institute of Medical Science, the University of Tokyo (approval number: 20-31-1120). All patients provided written consent.

Reference HIV-1 Env expression plasmids for positive control

CXCR4-using and CCR5-using reference HIV-1 envelopes expression plasmids were used for the positive control of the cell-cell fusion assay or virus-cell fusion assay. NL4-3 (GenBank accession number U26942) represented the X4 virus, and BaL (GenBank accession number M68893) represented the R5 virus strain. The full-length envelope gene of NL4-3 or BaL was amplified and inserted into pRE11 with the *Xba*I and *Xho*I double restriction enzymes cut.

Preparation of patients' HIV-1 Env expression plasmid

The recombinant DNA experiments used in this research were approved by the Institutional Review Board (approval number 08-30) and by the review board in the

Ministry of Education, Culture, Sports, Science and Technology (MEXT; approval number 23-1927). The full-length HIV-1 envelope (env) was amplified by PCR from patients' plasma as described. Viral RNA was extracted from 140μl patients' plasma using a QIAamp® Viral RNA Mini kit according to the manufacturer recommendations (QIAGEN, Hilden, Germany).

Viral RNA was reverse transcripted with the one-step RT-PCR using SuperScript Ⅲ (Invitrogen, California, USA) and Platinum® Taq DNA polymerase high Fidelity (Invitrogen). One sample was amplified separately in three reactions containing the same15μL mixture for the latter combination to minimize experiment bias. The15μL reaction mixture contained 2μL of RNA template, 7.5μL of 2×reaction buffer, 0.3μL of 5mM $MgSO_4$, 0.3μL of each 10μM of forward primer Env-1F, reverse primer Env-3Rmix (equimolar mixture of two 30-mer sequences) (Table 1), 0.6μL of SuperScript Ⅲ and Platinum Taq DNA polymerase high Fidelity, 0.25μL RNAse OUT, and 3.75μL of SQ. The one-step RT-PCR was carried out in the following conditions. The first step was reverse transcription at 55℃ for 30min, followed by inactivation at 94℃ for 2 min. The second step included thirty cycles; each cycle composed of three sequential steps: denaturation at 94℃ for 20 sec, annealing at 55℃ for 30 sec, extension at 68℃ for 4 min. After the 30 cycles, final extension at 68℃ for 5 min was added. The amplification

product spanned the positions of the envelope gene from position 5,853 to position 8,936 referred to NL4-3. From a combination of the three reactions of the first-round PCR product, 1.2μL was used as a template in a 15μL system for the second-round PCR with the forward primer EnvB-2F-XbaMlu and the reverse primer EnvB-4R-AgeXho. The second-round PCR was carried out for five separate reactions. The second-round PCR product spanned the positions of the envelope gene from position 5,957 to 8,817 referred to NL4-3. The five reactions were mixed and purified using a QIAquick PCR Purification Kit (QIAGEN) by first adding 5 volumes of Buffer PB to 1 volume of the PCR sample and mixing and then adding 10 μL of 3 M sodium acetate if the color of the mixture was orange or violet and then mixing. The sample was applied to the QIAquick column and centrifuged for 30–60s, washed with 0.75 ml Buffer PE, and eluted with 30μL EB. The PCR purification product around 3000bp with a variation in each product was purified by 1% agarose gel and the QIAquick gel extraction kit (QIAGEN). The purified product was double digested by the *Xba*I (5'C|TCGAG) and *Xho*I (5' T|CTAGA) restriction enzymes at 10μg plasmid with a 30 unit *Xba*I and a 30 unit *Xho*I in a 50μL system at 37℃ overnight. Therefore, the double restriction enzyme cut mixture was ligated to a double restriction enzyme cut pRE11 by ligation at a 3:1 molar ratio at a total volume of 5μL in a system containing the 5μL solution I of the DNA Ligation Kit Ver. 2.1 (TAKARA, Shiga,

JAPAN) at 16℃ for 3-4 hours. The 10μL ligation mixture was transformed into 100μL JM109 competent cells with incubation on ice for 10 minutes and heat shock at 37℃ for 90 seconds followed by a 2 minute incubation before an additional culture for 30 minutes in a 450μL Super optimal broth with catabolite repression (S.O.C.), which was composed of 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose. The culture was pelleted down by a 1 minute centrifuge at 13,000 rpm, resuspended in 50μL S.O.C, and plated on an Amp+ LB-agar plate. The plate was cultured at 37℃ overnight, and all of the colonies on the plate were collected in a 1mL LB medium and prepared in bulk (pRE11-envbulk) with the PureYield™ Plasmid Miniprep System (Promega, Wisconsin, USA). The purified plasmid DNA was applied to a 0.8% gel electrophoresis, and the DNA band according to the positive control pRE11-NL4-3 was cut and purified using the QIAquick Gel Extraction Kit (QIAGEN). The purified DNA was transformed again into the JM109 competent cells, and single colonies were picked up and amplified by an overnight culture in 2mL LB-amp in 37℃, 5% CO_2. The plasmid DNA were purified by the QIAquick Gel Extraction Kit and confirmed by 0.8% gel electrophoresis, and the envelope DNA were applied to sequencing PCR using the BigDye® Terminator v3.1 Cycle Sequencing Kit with two primers of sense strands and antisense strands covering the V3 region of the envelope

gene in condition of pre-denaturation at 96℃ for 1 minute and denaturation at 96℃ for 10 seconds, annealing at 50℃ for 5 seconds and extension at 60℃ for 4 minutes for 25 cycles. The PCR product was purified in Sephadex and applied to ABI 3730xl DNA Analyzers (Applied Biosystems, California, USA). The sequencing product was read by sequencher 4.8 (Gene Codes Corporation, Michigan, USA) and analyzed by Genetyx 12.0.0 (Genetyx Corporation, Tokyo, Japan).

In-house virus-cell fusion assay

The virus-cell fusion assay was also called the single round assay, which used the Env-expressing vector pRE11-env to pseudotype 293FT cells with a HIV-1 genomic vector pNL4-3.Luc.R-E-. The co-receptor-expressing cells, also called indicator cells, were N4X4 and N4R5 cells. The virus-cell fusion assays were carried out to examine the co-receptor usage of the clones phenotypically. The HIV-1 genomic vector pNL4-3.Luc.R-E- carried a Luciferase gene, which can easily embody the infectivity of a pseudovirus. Also, pNL4-3.Luc.R-E- had envelope deletion (Δenv), which should be compensated with an Env-expressing vector to produce the full-length HIV-1 without Vpr, that indicates an incompetent replication or single round infection (Fig. 7).

On day 1, 1.4×10^6 cells/well of 293FT cell in 2mL D10- were seeded in a 6-well

tissue culture plate in order to achieve 100% confluency on the next day.

One day 2, pRE11-env and pNL4-3.Luc.R-E- were co-transfected into 293FT cells. 123.7μL OPTI-MEM I Reduced Serum Medium (GIBCO) were aliquoted into each 5mL polystyrene round-bottom tube (BD Falcon), and 1250ng plasmid DNA of pRE11-env and 1250ng pNL4-3.Luc.R-E- were added into the medium followed by a brief vortex and incubated at room temperature for 10 minutes. Then, a 8.7μL Fugene6 transfection reagent was added into each tube for another 20 minutes of incubation. Finally, the mixture was added into the medium of 293FT cells in a 6-well tissue culture plate in a manner of three dot scattering.

On day 3, N4X4 cells and N4R5 cells were plated on the 96 Well Optical Btm Plt PolymerBase White w/Lid Cell Culture Sterile PS (Thermo Fish Scientific, Maltham, USA) at 4.0×10^4 cells/well in a total volume of 100μL/well of M10+4.

On day 4, 8.9μM of AMD3100 or MVC (Sigma) in 80μL M10+4 was added to each well of a 96 well optical bottom plate according to N4X4 cells or N4R5 cells. Ninety minutes later, the supernatant of transfected 293FT cells in a 6-well tissue culture plate was recovered and filtered with a 0.45μm syringe filter (Sartorius stedim biotech, Goettingen, Germany), and 1.5mL DEAE was added to the filtered virus per mL and fully mixed. Each 100μL recombinant virus was added into each well of N4X4 cells and N4R5

cells in the 96-well optical bottom plate with AMD3100 or MVC pretreatment or without AMD3100 or MVC treatment and duplicated.

On day 7, a 185μL supernatant was removed, and a 75μL reagent of ONE-Glo™ Luciferase Assay System (Promega) was added into the medium of each well of the 96-well optical bottom plate. After incubation for 3 minutes at room temperature, the plate was read with a GloMax® 96 Microplate Luminometer. The luciferase activity was measured three times consecutively and was recorded as the relative light unit (RLU). The average of three measurements represented the actual mean luciferase activity of each well. HIV-1 envelope references, NL4-3 and BaL, in pRE11 backbone were transfected into 293FT cells, and 293FT cells were co-cultured with N4X4 cells and N4R5 cells, respectively, serving as positive controls. pRE11 served as negative control and was transfected into 293FT cells and co-cultured with both indicator cells.

Cell-cell fusion assay

A new phenotypic co-receptor assay based on cell-cell fusion was used [28]. To make the virus-cell fusion assay comparable to the cell-cell fusion assay (Fig. 8), the Env-expressing cells and the co-receptor-expressing cells were the same as were used in the virus-cell fusion assay. Thus, the difference between the two phenotypic assays

existed only in cell-cell fusion or virus-cell fusion, and the virus-cell fusion assay indicates the actual infection.

293FT cells were used as envelope expression cells and were seeded at 2.8×10^5 cells/well in a 24-well tissue culture plate with a 500μL D10- medium.

On the next day, pRE11-env clones were transfected into the 293FT cells of a 60-80% confluency. The transfections were carried out as follows: a 25μL OPTI-MEM® I Reduced Serum Medium (GIBCO) was aliquoted into each 5mL polystyrene round-bottom tube (BD Falcon), and each 500ng plasmid DNA of pRE11-env was added followed by a brief vortex and incubated at room temperature for 10 minutes. Next, a 1.75μL Fugene6 transfection reagent (Promega) was added into each tube for another 20 minutes of incubation. Finally, the transfection mixture was added into corresponding wells of 293FT cell cultures in the 24-well tissue culture plate. On the same day, N4X4 cells and N4R5 cells were plated on a 96 Well Optical Btm Plt PolymerBase White w/Lid Cell Culture Sterile PS (Thermo Fish Scientific) at 2.0×10^4 cells/well in a 100μL M10+4 medium.

48 hours post-transfection, transfection efficiency was examined by red fluorescence observed with an IN Cell Analyzer 1000 (GE Healthcare Limited, Buckinghamshire, UK). The co-culturing of Env-expressing 293FT cells and indicator

N4X4 or N4R5 cells was carried out by replacing the medium of 293FT cells with 1mL phosphate buffer saline (PBS) (Sigma), resuspended by gently pipetting, and each 100μL was aliquoted into N4X4 cells and N4R5 cells in a 96-well optical bottom plate in a duplicate manner, respectively. If needed, 8.9μM of AMD3100 or MVC (Sigma) in 80μL M10+4 was added to each well. The co-culture cells were incubated at 37℃ in a humidified 5% CO_2 incubator for 6 hours before the 90μL supernatant was discarded and the 5μL EnduRen™ Live Cell Substrate (Promega) was added to each well. The plate was incubated at 37℃ for an additional 90 minutes. The luciferase activity was measured three times by a GloMax® 96 Microplate Luminometer (Promega) and recorded as a relative light unit (RLU). The average of three measurements represented the actual mean of luciferase activity in each well. HIV-1 envelope references, pRE11-NL4-3 and pRE11-BaL, were transfected into 293FT cells and co-cultured with N4X4 cells and N4R5 cells, respectively, serving as positive controls; pRE11 served as a negative control and was transfected into 293FT cells and co-cultured with both indicator cells.

Geno2pheno assay

The V3 region of clones were sequenced by an E110 forward primer and an Er115 reverse primer. The subtype and co-receptor usage were predicted by a

Geno2pheno algorithm with a false positive rate of 10%. The A or A/G clones were aligned by clustalW 2.1 on the DNA DATA BANK of Japan (DDJB) with the HIV-1 reference genes representing subtype A1, A2, B, C, D, F1, F2, G, H, J, K, N, O, and CRF01_ AE and confirmed again with a phylogenetic tree draw by TreeView. An FPR above 10% was regarded as R5, and an FPR below 10% was regarded as X4 clones (Fig. 9).

Full-length env sequencing

Twelve primers, including forward primers PGK-F2, E70, E110, E130, E170, and E250 and reverse primers Er35, Er65, Er115, Er145, and Er155, and ori-seq-R2 were used to cover the full-length env, which was 2551bp~2640bp in length with a variation in each clone (Fig. 10). The BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, California, USA) was used to amplify the corresponding fragment. The 10μL reaction mixture contained 1μL of DNA template, 2μL of 5×reaction buffer, 1μL of BigDye® Terminator v3.1, 0.25μL of each 10μM of forward primer and reverse primer, and 6.25μL of SQ. PCR was carried out under the following condition. Predenaturation at 96℃ for 1 minute, followed by 25 cycles of denaturation at 96℃ for 10 seconds, annealing at 50℃ for 5 seconds and extension at 60℃ for 4 minutes. For IMS0328, the

additional primer E2350 was used, and for IMS0718 and IMS1000, the additional primer E510 was used. The sequences of the total 14 primers are shown in Table 1. Sequence purification included filling illustra™ Sephadex™ G-50 Fine DNA Grade (GE Healthcare, Buckinghamshire, UK) into a Multiscreen plate (Merck Millipore, Massachusetts, USA) by a sephadex loading plate and filling 300μL SQ in a Sephadex-contained well leaving the sephadex to swell at room temperature for at least 1 hour before using it in the purifying step. The purifying step included spinning down the plate at 900×g for 5 minutes and loading the 10μL labelling PCR product on the middle top of the swelled sephadex and then filling 10μL SQ in a new skirted 96-well plate. Finally, the labelling product was spinned down to a new plate at 900×g for 5 minutes before being applied to the ABI 3100 Genetic Analyzer (Applied Biosystems, California, USA). The instrument protocol was RapidSeq36_POP7_BDv3, and the analysis protocol was KB_POP7_BDv3.

Fisher's exact test

The numbers of conserved amino acids in each appropriate region (C1, C2, C3, C4, C5, V1, V2, V3, V4, and V5) was subjected to a contingency table analysis. The null hypothesis determined that the mutations in each region had no difference regardless of how many amino acids the regions contained. The degrees of freedom (df) was 9. To

calculate whether each clustering of mutations in a certain region was different from the others or not, the 2×2 contingency table was used. The df was 2. Fisher's exact tests were applied.

Results

Estimation and determination of cutoff level of each assay

Ten percent was chosen as the false positive rate (FPR) in Geno2pheno as recommended by the European consensus groups on the clinical management of HIV-1 tropism testing, which determined that the significance level of the conservation of detection of CXCR4 usage should be 10% (http://coreceptor.geno2pheno.org/index.php).

The laboratory strains NL4-3 and BaL were used as positive controls in N4X4 and N4R5 cells, respectively, and as negative controls in N4R5 and N4X4 cells, respectively. Each virus-cell fusion assay contained the positive control and the negative control and were duplicated in the cell-culture plates. The values of NL4-3 in N4R5 cells and BaL in N4X4 cells in each assay were used in the equation of the mean plus two standard derivations (SD) as a threshold calculation, which was 465 for the luciferase activity in N4X4 cells and 4686 for the luciferase activity in N4R5 cells.

Each cell-cell fusion assay contained the positive control and the negative control and were duplicated in the cell-culture plates. The cell-cell fusion assays used the threshold of 911 relative light unit (RLU) in N4X4 cells and 435 relative light unit (RLU) in N4R5 cells, respectively (Fig. 11).

Patient characterization and the results of the three co-receptor assays on patients' plasma viruses

The patients enrolled in this study had not received ART therapy and were selected randomly. The numbers IMS0025, IMS0155, IMS0328, IMS0718, IMS0782, and IMS1000 assigned to the patients were used in this research. Among them, IMS0025 and IMS0155 had a low CD4 count ranging from 19 to 33, and IMS0328, IMS0718, IMS0782, and IMS1000 had a relatively higher CD4 count that ranged from 233 to 399 (Table 2).

The envelope DNAs reversely transcripted from the patient serums were ligated to the pRE11 vector, and each contained the whole quasispecies infecting one individual. The DNA plasmids were called pRE11-env_{bulk} or $bulk_{envs}$. The six $bulk_{envs}$ were applied to the Geno2pheno assay, the virus-cell fusion assay, and the cell-cell fusion assay, respectively. The false positive rates of the six $bulk_{envs}$ of IMS0025, IMS0155, IMS0328, IMS0718, IMS0782, and IMS1000 were 54.8%, 55.3%, 37.8%, 85.5%, 35.6%, and 10.5%, respectively, as they were all above 10%. The six $bulk_{envs}$ were all predicted to be the R5 genotype by the Geno2pheno assay. The virus-cell fusion assay showed that the luciferase activities of the six $bulk_{envs}$ IMS0025, IMS0155, IMS0328, IMS0718, IMS0782, and IMS1000 in the N4X4 cells were 1173, 574, 558, 235, 204, and 254

relative light unit (RLU), respectively, and the luciferase activities of the six $bulk_{envs}$ in the N4R5 cells were 1413289, 1622193, 215228, 10938, 74071, and 249309 relative light unit (RLU), respectively. Based on this, the six $bulk_{envs}$ IMS0025, IMS0155, IMS0328, IMS0718, IMS0782, and IMS1000 were divided into two phenotypes: the R5 virus (CCR5-using) and the R5X4 virus (CXCR4-using). The $bulk_{envs}$ of IMS0025, IMS0155, and IMS0328 were predicted to be R5X4, and IMS0718, IMS0782, and IMS1000 were predicted to be R5. The cell-cell fusion assay had a result of all R5X4 phenotypes for the six $bulk_{envs}$. The luciferase activities in the N4X4 cells were 2725, 2379, 1857, 27294, 1778, and 18432 relative light unit (RLU), respectively, and the luciferase activities in the N4R5 cells were 14070, 30799, 75923, 25342, 29176, and 66013 relative light unit (RLU), respectively. This caused the $bulk_{envs}$ of IMS0025, IMS0155, and IMS0328 to have concordant results with the virus-cell fusion assay, while IMS0718, IMS0782, and IMS1000 had discordant results with the virus-cell fusion assay, by cell-cell fusion assay (Fig. 12). To clarify the mechanism within the discordant results, single colonies were picked up after transforming the $bulk_{envs}$ into JM109 competent cells, and the DNA clones were purified for the Geno2pheno assay, virus-cell fusion assay, and cell-cell fusion assay.

Clonal analysis of viruses originating from patients

From IMS0025, 3 clones were separated. From IMS0155, 1 clone was separated. From IMS0328, 5 clones were separated. From IMS0718, 7 clones were separated. From IMS0782, 1 clone was separated. From IMS1000, 9 clones were separated. All of the 25 clones were applied to a full-length sequencing. According to the references from subtype A, subtype B which included laboratory strains BaL, NL4-3, HXB2, and SF2, and subtype C, the phylogenetic tree was created (Fig. 13). Because the primary prevalent HIV-1 strains were subtype B, it was not surprising that all of the clones from the six patients were all subtype B. The clones from each individual were clustered in one branch and were separate from the others.

Comparison between Geno2pheno assay and virus-cell fusion assay

From $bulk_{env}$ IMS0025, clones IMS0025-9, IMS0025-13, and IMS0025-15 all had an FPR of 54.8%. From $bulk_{env}$ IMS0155, clones IMS0155-22 had an FPR of 41.9%. From $bulk_{env}$ IMS0328, clones IMS0328-6, IMS0328-7, IMS0328-27, and IMS0328-29 had an FPR of 37.8%. From $bulk_{env}$ IMS0718, clones IMS0718-6, IMS0718-11, IMS0718-15, IMS0718-22, IMS0718-23, IMS0718-27 and IMS0328-63 had an FPR of 52.8, 96.5, 52.8, 1.1, 52.8, 52.8, and 1.1%. From $bulk_{env}$ IMS0782, clones IMS0782-35

had an FPR of 49.9%. From $bulk_{env}$ IMS1000, clones IMS1000-27, IMS1000-32, IMS1000-35, IMS1000-42, IMS1000-63, IMS1000-64, IMS1000-81, IMS1000-109, and IMS1000-111 had an FPR of 1.7, 9.6, 13.2, 13.2, 9.6, 9.6, 5.0, 13.2, and 9.6%. Out of the 25 clones, 17 were predicted to be R5 viruses, and 8 were predicted to be R5X4 or X4 viruses (Fig. 14). These 8 CXCR4-using clones were derived from two individuals, IMS0718 and IMS1000. The V3 sequences were aligned. Geno2pheno used the following rules in the algorithm, interpreted as the positive charge should be calculated as the total number of positively charged (R/K/H) amino acid residues. The positive charges of the R5 clones were from 7 to 8 and from 8 to 10 for the CXCR4-using clones. A higher positive charge was correlated with the likelihood of CXCR4 use. The net charge was calculated as the number of positively charged (R/K/H) amino acid residues minus the number of negatively charged (D/E) residues. The 11/25 rule is an addition to the positive charge. Basic residues (R or K) at either or both of these sites is predictive of CXCR4 use. For the 11/25 rule, two basic R in the 11^{th} position and one basic K in the 25^{th} position from three CXCR4-using viruses, IMS0718-22, IMS0718-63, and IMS1000-27, obeyed the rule; however, the other 5 CXCR4-using clones did not obey the rule. Three R5 clones with an FPR of 13.2% had the same relatively high positive charge and net charge as the CXCR4-using viruses. The net charge of the CXCR4-using clones ranged from 7 to 10.

The false positive rate was the probability of classifying a CXCR4-not-using virus falsely as a CXCR4-using virus. The FPR of the R5 clones were from 13.2% to 96.5% and from 1.1% to 9.6% for the CXCR4-using clones.

In the virus-cell fusion assay, the recovered DSP activity was detected by the luciferase activity (Luc), and each clone had values in each N4X4 and N4R5 cell, respectively. The values of the luciferase activity of each clone in N4X4 and N4R5 were estimated by comparing them with the cutoff in N4X4 or N4R5, and the 25 clones were divided into two groups: one had luciferase activities above the cutoff for N4R5 but lower than the cutoff for N4X4, and the other group had both values positive for N4R5 and N4X4. The former group represented R5 clones, and the latter represented R5X4 clones. The number of R5 clones was 21, and the number of R5X4 clones was 4. These 4 R5X4 were derived from two individuals, IMS0718 and IMS1000 (Fig. 15). The clones from IMS0718 and IMS1000 by the Geno2pheno assay and the virus-cell fusion assay were shown in the upper and lower panels. Panel (a) includes the results of the Geno2pheno assay, panel (b) includes the results of the virus-cell assay, and panel (c) shows the alignment of the V3 sequences. For IMS0718, the Geno2pheno and virus-cell assays predicted the same results: 5 R5 and two with CXCR4 usage. The alignment of the V3 sequences showed that four consistent amino acid changes were related to the co-receptor

usage. For IMS1000, the results from the Geno2pheno and virus-cell assays did not match. In the virus-cell assay, 7 were predicted to be R5, and 2 were predicted to be R5X4. In the Geno2pheno assay, four of the 7 R5 were predicted to be CXCR4 usage. The V3 alignment showed that E/K in the 25^{th} position was related to the co-receptor usage; however, it did not appear to obey the rules. The alignment also showed that there were 4 clones with identical V3 sequences with different phenotypes. Three were R5 clones, and one was an R5X4 clone, which suggests that the co-receptor usage could be determined outside of the V3.

Figure 16 shows the full-length Env alignment of the clones from IMS0718. The clone numbers are shown on the left. The R5 clones are shown in the upper panel of the aligned sequences separated by a green line from R5X4 clones in the lower panel. The sequences were divided into C1 to V5 regions, respectively. The consistent mutations $T_{303}I$, $R_{304}K$, $S_{306}R$, and $D_{321}G$ in V3 are shown in the yellow bar. The $S_{306}R$ and $D_{321}G$ positioned in the 11^{th} and 25^{th} amino acid of the V3 sequence of Env, respectively, obeyed the rules of positive charge. There are two consistent mutations, E_5G in C1 and $E_{268}K$ in C2, shown in the pink bar. The 3D structure of gp120 with a PDB file name, 4TVP, summarizes the consistent amino acid usages that were different in the R5 and R5X4 clones. After CD4 binding, V1V2 was repositioned relative to V3, and the conserved

mutations were predicted (Fig. 17). The mutations of the amino acid in V3 are shown in red in the left panel, and the mutations of the amino acid outside of V3 are shown in the right panel. The mutations in V3 were in the co-receptor binding site; however, the mutations in C1 and C2 appear to be far from the co-receptor binding site. Figure 18 shows the full-length Env alignment of another patient, IMS1000. The sequences in V1 were quite heterogeneous. The 9 clones can be differentiated into two groups, as shown in red and pink. Sequence IMS1000-27 does not appear to belong to either of the two groups because of some mutations that are scattered throughout the full-length Env. From the analysis of the phylogenetic tree, we can see that the 9 clones that originated from the same branch evolved into one single clone and two groups, which are colored in red and pink, respectively. Group 1 had one R5X4 clone, IMS1000-32, and four R5 clones. Group 2 had 3 R5 clones (Fig. 19). To keep the homogeneity of the sequences, the clones from group 1 only were aligned. Figure 20 shows the alignment of the R5 and R5X4 clones in group 1. It can be observed that the consistent mutations K_2R, A_3V, R_9K in C1, $T_{161}A$, $T_{162}P$, $Y_{173}H$, $Y_{177}S$, $Q_{183}R$, $S_{190}R$, and $N_{195}H$ in V2 and $M_{347}I$ in C3 were responsible for the difference between the R5 and R5X4 clones. Figure 21 shows the crystal structure of the consistent mutations in IMS1000. In the left structure, seven amino acid changes in V2 were found, and they were in the co-receptor binding site. In the right structure, 4

amino acid changes in C1 and C3 were found, which appear far from the co-receptor binding site. For fisher's exact test, the number of conserved amino acids compared to the full-length of each appropriate region (C1, C2, C3, C4, C5, V1, V2, V3, V4, and V5) were applied to a null hypothesis, and the groups did not have significant differences; however, the hypothesis was rejected by $p<0.05$ (Fig. 22). The region with significantly different mutations needed to be identified. Thus, C1, V2, or C3 was extracted from the full-length Env. The frequencies of amino acid mutations in each of the three regions and amino acid mutations in the remaining total of envelope were applied to a 2×2 contingency table analysis. When C1 was extracted to be compared with the remaining regions, it showed $p>0.05$. When V2 was extracted to be compared with the remaining regions, it showed $p<0.05$. When C3 was extracted to be compared with the remaining regions, it showed $p>0.05$. To eliminate the effect of the significant influence of amino acids in V2, the frequencies of amino acid mutations in C1 and C3 and frequencies in the remaining total of envelope excluding V2 were applied to another two 2×2 contingency table analysis. They still showed no significant difference of C1 or C3 from the contrast, $p>0.05$. This means that the number of consistent amino acid changes in V2 had significant differences in comparison with the full-length Env.

The comparison between the Geno2pheno assay and the virus-cell fusion assay

can be summarized as: an amino acid outside of V3 could influence the co-receptor usage by the virus-cell fusion assay. By the summary of the mutations from IMS0718 and IMS1000, there were a total of 4 mutations in C1, 7 mutations in V2, 1 mutation in C2, and 4 mutations in V3. Generally speaking, the amino acid differences that were correlated with CXCR4 usage clustered in V2 and C1, especially in V2. The amino acid changes in C2 and C3 may be related to CXCR4 usage and may also be a consequence of the small number of samples.

Comparison between the virus-cell fusion assay and cell-cell fusion assay

Figure 23 shows the results of the 25 clones in the cell-cell fusion assay. Panel (a) is the result of the Geno2pheno assay, (b) is the result of the virus-cell assay, and (c) is the result of cell-cell fusion assay. Co-receptor usage predicted by cell-cell fusion assay of the 25 clones is described as the following. 5 clones were predicted as R5, and 21 clones were predicted as R5X4. Compared with the virus-cell assay, there were 16 R5 clones in the virus-cell fusion assay turning out to be R5X4 clones in the cell-cell fusion assay. The clones were from four patients: IMS0025, IMS0718, IMS0328, and IMS1000. Panel (d) shows the location of the consistent mutations $S_{49}T$, $I_{100}M$ located in C1, $S_{134}V$ located in V1, $E_{336}R$, $R_{344}K$ located in C3, $K_{410}E$ located in V4, and $I_{465}T$ located in V5.

The crystal structure shows the sites of the consistent mutations by the cell-cell fusion assay. It can be observed that the amino acids that differentiate in the discordant clones between the virus-cell and cell-cell assays were scattered in different regions of the envelope.

Discussion

The most important parameters in an HIV-1 infected patient are the CD4+ T cell counts and the HIV RNA loads in the plasma. It has been discovered that the HIV-1 infected patients with CD4+ T cell counts that ranged from 200 to 350/μL had a high mortality related to complications such as liver diseases and coronary artery diseases compared to those with CD4+ T cell counts above 350/μL. An HIV-1 infected individual with CD4+ T cell counts less than 200/μL has a higher possibility of complications of opportunistic infections. The relationship between the CD4+ T cell counts and the HIV RNA loads appear to be negative. Research on the proportion of variability in the rate of CD4+ T cell loss in untreated HIV-infected persons shows that an average decreasing percentage of CD4+ T cells by year has a significantly higher value in the group with higher HIV RNA loads than in the group with lower HIV RNA loads [30]. An analysis of NA-ACCORD in 2009 shows that groups starting antiretroviral therapy (ART) when the CD4+ T counts ranged from 351~500/μL have a higher mortality than groups starting ART when the CD4+ T counts dropped to below 350/μL [31]. The data from HIV related drugs (HRD) shows that the mortality between the patients with CD4+ T cell counts ranging from 200 to 350/μL and those with CD4+ T cell counts over 350/μL had a significant difference [32].

In this study, two patients had CD4+ T cell counts lower than 200/μL, and four patients had CD4+ T cell counts ranging from 200 to 350/μL. Three of the four patients with CD4+ T cell counts ranging from 200 to 350/μL were predicted to be R5 by the $bulk_{env}$ virus-cell fusion assay, and the remaining one patient with CD4+ T cell counts ranging 200 to 350/μL and two patients with CD4+ T cell counts lower than 200/μL were predicted to be R5X4 by the $bulk_{env}$ virus-cell fusion assay. This suggests that the CXCR4-using viruses were comparably increasing for the patients with lower CD4+ T cells. It has been reported that faster disease progression and rapid CD4+ T cell loss is associated with the emergence of CXCR4-using HIV-1 during the disease's course in some AIDS patients [33]. The prognosis of the patients was unknown; however, the co-receptor usages reflected the HIV-1 infection stages of each individual properly.

The clones showed a discrepancy with the $bulk_{envs}$ in co-receptor usages by the Geno2pheno assay or virus-cell fusion assay. As for the Geno2pheno assay, the discrepancy for the $bulk_{envs}$ and clones existed in IMS0718 and IMS1000, which had CXCR4-usage clones, but the $bulk_{envs}$ were both R5. The other four $bulk_{envs}$ and clones were R5. This can be explained because the Geno2pheno assay has a low sensitivity in the detection of X4 minority variants in a $bulk_{env.}$ For the maraviroc treatment, the presence of CXCR4-using viruses could not be included; however, because the limited sensitivity

of the Geno2pheno for the detection of a CXCR4-using virus is 10%, it had been implied that the genotypic assay is inadequate for the prediction of the X4 variant in a clinical setting [34]. Many methods have been suggested to improve the genotypic assay, including raising the repeating times on samples from the same patient, combining several algorithms together, changing the assay threshold parameters, and incorporating the results based on regions other than the V3 [35-38]. V3 loop mutagenesis studies have demonstrated that co-receptor binding is strongly related to basic residues, including R_{298}, R_{308}, R_{315}, R_{327}, and two hydrophobic residues, F_{317} and I_{323}, both in the R5 and X4 virus. For R5 co-receptor usage, residues N_{302}, D_{325}, and I_{326} were proven to be important for determining co-receptor preference [39]. For X4 co-receptor usage, two basic residues, R_{306} and K_{322}, were proven to strongly correlate to CXCR4 use [40]. Because CCR5 structures have a more open ligand binding pocket, the V3 loop backbone residues 318-320 are engaged in hydrogen bond interactions with the ECL2 of CCR5, and the crown GPGR motif has residue P_{313} engaged in hydrophobic contact with helices I and helices II of CCR5. The V3 mutation $S_{306}R$ from the CCR5/R5-V3 model to the CXCR4/X4-V3 model showed a decreased binding ability to CCR5 and a suggested determinant of co-receptor preference. It was proven that Q_{310}-R_{311} in V3 were commonly absent in most HIV-1 isolates, except particular X4-tropic viruses. When the prototypical

X4 isolate HXBc2 is removed from the two residues, it becomes an R5X4 strain, which suggests that the R5 tropic should not have the residues 310-311. $I_{323}V$ was reported to be related to the resistance to the CCR5 antagonists, including the small molecule maraviroc and the polypeptide 5P12-RANTE [41-44]. It is widely accepted that reductions in CCR5 affinity can reduce fusion levels and viral infectivity [45, 46].

By the virus-cell fusion assay, three $bulk_{envs}$ showed R5X4 for IMS0025, IMS0155, and IMS0328, but the clones from the three patients all showed R5 viruses. Two $bulk_{envs}$ showed R5 for IMS0718 and IMS1000, but the clones showed R5 as well as R5X4. Only for one $bulk_{env}$ IMS0782, virus-cell fusion assay predicted R5 in concordance with clones. The discordance of $bulk_{envs}$ and clones may be explained by the high sensitivity on X4 minorities of $bulk_{envs}$ and the small number of clones separated from each $bulk_{env.}$ Dr. Noriaki Hosoya tested the sensitivity for the minor variant detection of the virus-cell fusion assay using the reference strains pRE11-NL4-3 or pRE11-BaL. Based on the luciferase activity, up to 0.3% in the pRE11-NL4-3 and 5% in the pRE11-BaL detection limit showed a significant level of the statistic test when mixing the two strains by different mixing ratios [28]. Because the sensitivity of detection on CXCR4-using minor variants was higher than the Geno2pheno assay, the virus-cell fusion assay could detect three R5X4 from the $bulk_{envs,}$ but the Geno2pheno assay could

not. The virus-cell fusion assay appeared to have some limitations on viral load. The viral load for the Trofile assay should be >1,000 copies/mL; however, the recommended sample for the genotypic assay is plasma with a viral load >500 copies/mL. By comparison of the Geno2pheno and virus-cell fusion assay, it showed four discordant clones which were predicted CXCR4-using by Geno2pheno assay, but R5 by virus-cell fusion assay. The FPR of the four discordant clones IMS1000-81, IMS1000-63, IMS1000-64, and IMS1000-111, predicted by the Geno2pheno assay and the virus-cell fusion assay, were 5, 9.6, 9.6, and 9.6%, respectively. And remaining four CXCR4-using clones predicted by Geno2pheno assay were R5X4 by virus-cell fusion assay. The FPR of the four concordant CXCR4-using clones IMS0718-22, IMS0718-63, IMS1000-27, and IMS1000-32, predicted by the Geno2pheno assay and the virus-cell fusion assay, were 1.1, 1.1, 1.7, and 9.6%, respectively. The two groups showed a statistical significance by the one-tailed unpaired student *t*-test, $p<0.05$. The luciferase activities in N4X4 by the virus-cell fusion assay of the two groups also showed a statistical significance, $p<0.05$, when the logarithmic values of the two groups in N4X4 were applied to the one-tailed unpaired student *t*-test; however, the luciferase activities in N4R5 did not show a statistical significance. The significant level of a false positive rate had been recommended to be 10%. If the cutoff is 5%, Geno2pheno assay would predict the 25

clones concordantly with virus-cell fusion assay. However, research on assessing the performance of the Geno2pheno assay in predicting CXCR4-using viruses compared with the enhanced version of the Trofile HIV tropism assay (ES Trofile) showed that when the significance level of the false positive rate was 20%, the Geno2pheno assay showed the highest sensitivity (76.7%) [47]. Although the Geno2pheno assay had been approved for clinical diagnostics to predict the co-receptor usage of an HIV-1 infection, the threshold for the CXCR4-usage level had not yet been conclusively established. In general, V3 genotyping samples from plasma HIV RNA or proviral DNA were recommended to be interpreted with a false-positive rate of 10%; however, if only one sequence can be generated from a DNA sample, it was recommended to interpret it with a false-positive rate of 20%. The ES Trofile was developed by Monogram Biosciences and improved by a 30-fold increase in analytical sensitivity on the detection of the CXCR4-using virus. Samples from naïve patients and CD4+ T cell counts that ranged from 200 to 500 cells/mm^3 showed the best predictive performance. Another cell-cell fusion assay called VERITROP uses a homologous recombination by an innovative yeast-based cloning strategy.

The discrepancy of the two R5 $bulk_{envs}$ with the R5X4 clones from IMS0718 and IMS1000 showed that the sensitivity of the X4 minor variant was not high enough. In this

research, the E_{322} in V3 of the 4 discordant clones determined by the Geno2pheno assay and the virus-cell fusion assay were from one $bulk_{env}$ IMS1000, and another two X4 clones determined by the Geno2pheno assay showed K_{322} and E_{322}, which may suggest that E_{322} may be the characteristic for the CXCR4 co-receptor usage at least for the one patient, IMS1000. It was reported that the combination of the mutations $N_{197}D$, $P_{311}R$, and $E_{320}K$ has the ability of a co-receptor switch in CCR5:CXCR4 chimeric co-receptors [48]. The result in this study of E_{322} is a new finding, which clearly suggests that the amino acid outside of V3 should be the determinant for co-receptor usage. The variable regions V1, V2, V3, V4, and V5 form an exposed loop of gp120, while the conserved regions C1, C2, C3, C4, and C5 form the gp41 binding surface and the CD4 binding surface [49]. It was generally considered that the coding of the variable region 3 (V3 loop) of gp120 of the HIV envelope is the major genetic determinant of CCR5 or CXCR4 associated with R5 or X4 virus binding. In the V3 regions, it has been reported that X4 tropic viruses have more positively charged amino acid residues and less N-linked glycosylation sites than R5 viruses [50]. To differentiate the CCR5-using viruses from the CXCR4-using viruses, the amino acid variability, net charge, amino acid length, frequency of insertions and substitutions within the GPGQ crown motif, and frequency of glycosylation were investigated by the CXCR4-using V3 sequences [51]. Structural studies showed the

structure rearrangement of gp120 is primarily in the core structure. The gp120 3D structure analysis of the V3 loop from a R5 virus showed that the V3 loop protrudes from the gp120 core by 30 Å and likely interacts with the co-receptors following CD4 binding with this length, and 3D structure from both R5 virus and X4 virus showed the bridging sheet formed between the C1, C2, and C4 domains had a corresponding entropy loss in CD4-gp120 complexation [49, 52]. The V3 region and the bridging sheet formed after CD4 binding were reported to contact the N-terminal domain of CCR5 by using a ΔV1/V2 HIV-1 clone TA1, which is derived from the R3A clone with a one-half truncation of the V3 loop [53]. The V1/V2 region was also reported to have an implication on co-receptor usage [50, 54] [48, 55-59], but it was controversial regarding whether or not V1/V2 plays a role in co-receptor selection [60, 61]. The co-receptor usage-related amino acid substitution is also mapped to the gp41 TM subunit and the fusion peptide of TM [62]. In this research, 7 conserved mutations $T_{161}A$, $T_{162}P$, $Y_{173}H$, $Y_{177}S$, $Q_{183}R$, $S_{190}R$, and $N_{195}H$ in V2 predicted by the virus-cell fusion assay were related to R5 to R5X4 changes and have statistical significance.

The cell-cell fusion assay showed a high sensitivity on CXCR4-using viruses. The 25 clones measured by the cell-cell fusion assay showed no discrepancy with the $bulk_{envs}$. The $bulk_{envs}$ were all predicted to be R5X4, and the clones from the $bulk_{envs}$ all

contained R5X4. The clones also confirmed that the cell-cell fusion assay has a better sensitivity in the detection of the minority of CXCR4-using viruses than the virus-cell fusion assay. It would be too sensitive to apply for a clinical usage. So, I compared the virus-cell fusion assay and cell-cell fusion assay in order to find the amino acid differences in R5X4 prediction between these two assays. In Fig. 23, there were 16 discordant clones out of the 25 clones that were predicted to be R5 by the virus-cell fusion assay but R5X4 by the cell-cell fusion assay. The luciferase activity in the N4X4 cells by the cell-cell fusion assay of the 16 discordant clones predicted by the virus-cell fusion assay and the cell-cell fusion assay, had a significant difference from the five concordant R5 clones predicted by the virus-cell fusion assay and the cell-cell fusion assay, IMS0025-9, IMS0025-15, IMS0328-27, IMS1000-109, and IMS1000-111, in the log transfer, $P<0.05$; however, the luciferase activity in the N4X4 cells by the virus-cell fusion assay of the discordant group and the concordant R5 group did not show a significant difference. In N4R5 cells, the luciferase activity of the 16 discordant clones and the five concordant R5 clones by both assays had a significant difference, $P<0.05$.

The conserved mutations $S_{49}T$, $I_{100}M$ located in C1, $S_{134}V$ located in V1, $E_{336}R$, $R_{344}K$ located in C3, $K_{410}E$ located in V4, and $I_{465}T$ located in V5 had the tendency to affect the results of the cell-cell fusion assay. The mutations were scattered through the

envelope protein, and they may have affected the cell-cell fusion in other ways.

Conclusion

1. In addition to V3, the V2 domain or C1 domain may be involved in co-receptor usage.

2. Two phenotype assays, virus-cell and cell-cell fusion assays, provided many discordant results on CXCR4 indicator cells.

3. The cell-cell fusion assay may be applied for CCR5 but not for CXCR4 usage.

Acknowledgement

During my stay at the University of Tokyo for three and half a year, I have proceeded in the path of a doctor. A foreigner student struggling in the academic study has so many difficulties. Without the following people, it is impossible to achieve.

First, I would give my deep acknowledgement to my professor Aikichi Iwamoto in division of infectious diseases, advanced clinical research center, the institute of medical science, the University of Tokyo leading me into the research of HIV and gives me warm care, patient and splendid supervision. And I wish to appreciate for Dr. Noriaki Hosoya in division of infectious diseases, advanced clinical research center, the institute of medical science, the University of Tokyo for enthusiastic supervising me in designing, operating, and summery of experiments and in performing presentations. I would give my respect for Dr. Ai Tachikawa in division of infectious diseases, advanced clinical research center, the institute of medical science, the University of Tokyo for discussion of experiment results. Furthermore, I am grateful to technician Stuchi Heinz and Yoshimi Kashi and doctor course student Michio Okame for teaching me RNA extraction, PCR technique from plasma, transfection, and virus-cell fusion assay, respectively. I would thank for all the members in division of infectious diseases, advanced clinical research center, the institute of medical science, the University of Tokyo for endless help and

communication in the studying life. Thanks for professor Zene Mazuda from research center for Asian infectious diseases, advanced clinical research center, the institute of medical science, the University of Tokyo for cooperation and provision of DSP_{1-7} and DSP_{8-11} vectors. I would like to thank Dr. Tadashi Kikuchi and Dr. Jiro Kato teaching me mathematics and statistics.

Thanks for the staff in the graduate school of medicine; they have given me affordable help.

In the last, I would thank for my parents for giving me birth and my friends in Japan for the support and friendship.

Reference

[1] G. J. Gottlieb, A. Ragaz, J. V. Vogel, A. Friedman-Kien, A. M. Rywlin, E. A. Weiner, and A. B. Ackerman, "A preliminary communication on extensively disseminated Kaposi's sarcoma in young homosexual men," *Am J Dermatopathol,* vol. 3, no. 2, pp. 111-4, Summer, 1981.

[2] F. Barre-Sinoussi, J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier, "Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)," *Science,* vol. 220, no. 4599, pp. 868-71, May 20, 1983.

[3] D. M. Knipe, and P. M. Howley, *FIELDS VIROLOGY*, 6 ed.: Wolters Kluwer Health | Lippincott Williams & Wilkins, 2013.

[4] P. A. Ashorn, E. A. Berger, and B. Moss, "Human immunodeficiency virus envelope glycoprotein/CD4-mediated fusion of nonprimate cells with human cells," *J Virol,* vol. 64, no. 5, pp. 2149-56, May, 1990.

[5] D. B. Weiner, K. Huebner, W. V. Williams, and M. I. Greene, "Human genes other than CD4 facilitate HIV-1 infection of murine cells," *Pathobiology,* vol. 59, no. 6, pp. 361-71, 1991.

[6] Y. Feng, C. C. Broder, P. E. Kennedy, and E. A. Berger, “HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor,” *Science,* vol. 272, no. 5263, pp. 872-7, May 10, 1996.

[7] E. Oberlin, A. Amara, F. Bachelerie, C. Bessia, J. L. Virelizier, F. Arenzana-Seisdedos, O. Schwartz, J. M. Heard, I. Clark-Lewis, D. F. Legler, M. Loetscher, M. Baggiolini, and B. Moser, “The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1,” *Nature,* vol. 382, no. 6594, pp. 833-5, Aug 29, 1996.

[8] F. Cocchi, A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, and P. Lusso, “Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells,” *Science,* vol. 270, no. 5243, pp. 1811-5, Dec 15, 1995.

[9] H. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, and N. R. Landau, “Identification of a major co-receptor for primary isolates of HIV-1,” *Nature,* vol. 381, no. 6584, pp. 661-6, Jun 20, 1996.

[10] G. Alkhatib, C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and E. A. Berger, “CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a

fusion cofactor for macrophage-tropic HIV-1," *Science,* vol. 272, no. 5270, pp. 1955-8, Jun 28, 1996.

[11] P. R. Clapham, J. D. Reeves, G. Simmons, N. Dejucq, S. Hibbitts, and A. McKnight, "HIV coreceptors, cell tropism and inhibition by chemokine receptor ligands," *Mol Membr Biol,* vol. 16, no. 1, pp. 49-55, Jan-Mar, 1999.

[12] E. A. Berger, R. W. Doms, E. M. Fenyo, B. T. Korber, D. R. Littman, J. P. Moore, Q. J. Sattentau, H. Schuitemaker, J. Sodroski, and R. A. Weiss, "A new classification for HIV-1," *Nature,* vol. 391, no. 6664, pp. 240, Jan 15, 1998.

[13] R. Liu, W. A. Paxton, S. Choe, D. Ceradini, S. R. Martin, R. Horuk, M. E. MacDonald, H. Stuhlmann, R. A. Koup, and N. R. Landau, "Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection," *Cell,* vol. 86, no. 3, pp. 367-77, Aug 9, 1996.

[14] M. Samson, F. Libert, B. J. Doranz, J. Rucker, C. Liesnard, C. M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cognaux, C. Forceille, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y. Yi, R. J. Smyth, R. G. Collman, R. W. Doms, G. Vassart, and M. Parmentier, "Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene," *Nature,* vol. 382, no. 6593, pp. 722-5, Aug 22, 1996.

[15] G. Pantaleo, and A. S. Fauci, “Immunopathogenesis of HIV infection,” *Annu Rev Microbiol,* vol. 50, pp. 825-54, 1996.

[16] Z. L. Brumme, J. Goodrich, H. B. Mayer, C. J. Brumme, B. M. Henrick, B. Wynhoven, J. J. Asselin, P. K. Cheung, R. S. Hogg, J. S. Montaner, and P. R. Harrigan, “Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals,” *J Infect Dis,* vol. 192, no. 3, pp. 466-74, Aug 1, 2005.

[17] H. Schuitemaker, N. A. Kootstra, R. E. de Goede, F. de Wolf, F. Miedema, and M. Tersmette, “Monocytotropic human immunodeficiency virus type 1 (HIV-1) variants detectable in all stages of HIV-1 infection lack T-cell line tropism and syncytium-inducing ability in primary T-cell culture,” *J Virol,* vol. 65, no. 1, pp. 356-63, Jan, 1991.

[18] E. A. Berger, P. M. Murphy, and J. M. Farber, “Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease,” *Annu Rev Immunol,* vol. 17, pp. 657-700, 1999.

[19] T. Shioda, J. A. Levy, and C. Cheng-Mayer, “Macrophage and T cell-line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene,” *Nature,* vol. 349, no. 6305, pp. 167-9, Jan 10, 1991.

[20] S. S. Hwang, T. J. Boyle, H. K. Lyerly, and B. R. Cullen, “Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1,” *Science,* vol. 253, no. 5015, pp. 71-4, Jul 5, 1991.

[21] C. Briones, A. de Vicente, C. Molina-Paris, and E. Domingo, “Minority memory genomes can influence the evolution of HIV-1 quasispecies in vivo,” *Gene,* vol. 384, pp. 129-38, Dec 15, 2006.

[22] E. Domingo, V. Martin, C. Perales, A. Grande-Perez, J. Garcia-Arriaza, and A. Arias, “Viruses as quasispecies: biological implications,” *Curr Top Microbiol Immunol,* vol. 299, pp. 51-82, 2006.

[23] P. Dorr, M. Westby, S. Dobbs, P. Griffin, B. Irvine, M. Macartney, J. Mori, G. Rickett, C. Smith-Burchnell, C. Napier, R. Webster, D. Armour, D. Price, B. Stammen, A. Wood, and M. Perros, “Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity,” *Antimicrobial agents and chemotherapy,* vol. 49, no. 11, pp. 4721-32, Nov, 2005.

[24] G. Fatkenheuer, M. Nelson, A. Lazzarin, I. Konourina, A. I. Hoepelman, H. Lampiris, B. Hirschel, P. Tebas, F. Raffi, B. Trottier, N. Bellos, M. Saag, D. A. Cooper, M. Westby, M. Tawadrous, J. F. Sullivan, C. Ridgway, M. W. Dunne, S.

Felstead, H. Mayer, and E. van der Ryst, “Subgroup analyses of maraviroc in previously treated R5 HIV-1 infection,” *The New England journal of medicine,* vol. 359, no. 14, pp. 1442-55, Oct 2, 2008.

[25] M. Saag, J. Goodrich, G. Fatkenheuer, B. Clotet, N. Clumeck, J. Sullivan, M. Westby, E. van der Ryst, and H. Mayer, “A double-blind, placebo-controlled trial of maraviroc in treatment-experienced patients infected with non-R5 HIV-1,” *The Journal of infectious diseases,* vol. 199, no. 11, pp. 1638-47, Jun 1, 2009.

[26] J. M. Whitcomb, W. Huang, S. Fransen, K. Limoli, J. Toma, T. Wrin, C. Chappey, L. D. Kiss, E. E. Paxinos, and C. J. Petropoulos, “Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism,” *Antimicrob Agents Chemother,* vol. 51, no. 2, pp. 566-75, Feb, 2007.

[27] T. Lengauer, O. Sander, S. Sierra, A. Thielen, and R. Kaiser, “Bioinformatics prediction of HIV coreceptor usage,” *Nat Biotechnol,* vol. 25, no. 12, pp. 1407-10, Dec, 2007.

[28] P. Teeranaipong, N. Hosoya, A. Kawana-Tachikawa, T. Fujii, T. Koibuchi, H. Nakamura, M. Koga, N. Kondo, G. F. Gao, H. Hoshino, Z. Matsuda, and A. Iwamoto, “Development of a rapid cell-fusion-based phenotypic HIV-1 tropism

assay," *J Int AIDS Soc,* vol. 16, no. 1, pp. 18723, 2013.

[29] N. Kondo, K. Miyauchi, F. Meng, A. Iwamoto, and Z. Matsuda, "Conformational changes of the HIV-1 envelope protein during membrane fusion are inhibited by the replacement of its membrane-spanning domain," *The Journal of biological chemistry,* vol. 285, no. 19, pp. 14681-8, May 7, 2010.

[30] B. Rodriguez, A. K. Sethi, V. K. Cheruvu, W. Mackay, R. J. Bosch, M. Kitahata, S. L. Boswell, W. C. Mathews, D. R. Bangsberg, J. Martin, C. C. Whalen, S. Sieg, S. Yadavalli, S. G. Deeks, and M. M. Lederman, "Predictive value of plasma HIV RNA level on rate of CD4 T-cell decline in untreated HIV infection," *JAMA,* vol. 296, no. 12, pp. 1498-506, Sep 27, 2006.

[31] M. M. Kitahata, S. J. Gange, A. G. Abraham, B. Merriman, M. S. Saag, A. C. Justice, R. S. Hogg, S. G. Deeks, J. J. Eron, J. T. Brooks, S. B. Rourke, M. J. Gill, R. J. Bosch, J. N. Martin, M. B. Klein, L. P. Jacobson, B. Rodriguez, T. R. Sterling, G. D. Kirk, S. Napravnik, A. R. Rachlis, L. M. Calzavara, M. A. Horberg, M. J. Silverberg, K. A. Gebo, J. J. Goedert, C. A. Benson, A. C. Collier, S. E. Van Rompaey, H. M. Crane, R. G. McKaig, B. Lau, A. M. Freeman, R. D. Moore, and N.-A. Investigators, "Effect of early versus deferred antiretroviral therapy for HIV on survival," *N Engl J Med,* vol. 360, no. 18, pp. 1815-26, Apr 30, 2009.

[32] “ HIV-related drug joint survey, 16 annual report,” *CMIC-PMS Co. Ltd.*

[33] P. R. Clapham, and A. McKnight, “HIV-1 receptors and cell tropism,” *Br Med Bull,* vol. 58, pp. 43-59, 2001.

[34] A. J. Low, W. Dong, D. Chan, T. Sing, R. Swanstrom, M. Jensen, S. Pillai, B. Good, and P. R. Harrigan, “Current V3 genotyping algorithms are inadequate for predicting X4 co-receptor usage in clinical isolates,” *Aids,* vol. 21, no. 14, pp. F17-24, Sep 12, 2007.

[35] A. Moores, A. Thielen, W. Dong, A. Low, C. Woods, M. Jensen, B. Wynhoven, D. Chan, C. Glasock, and P. R. Harrigan, “Improved detection of X4 virus by V3 genotyping: application to plasma RNA and proviral DNA,” *Antiviral Therapy,* vol. 13, no. 4, pp. A99-A99, 2008.

[36] N. Chueca, L. Martin, M. Alvarez, A. Pena, V. Guillot, V. Garcia-Casas, R. Sanchez, J. Hernandez-Quero, M. C. Maroto, and F. Garcia, “A combination of bioinformatic tools can be accurately used for the screening of coreceptor usage in clinical samples,” *Antiviral Therapy,* vol. 13, no. 4, pp. A106-A106, 2008.

[37] E. Poveda, E. Seclen, M. M. Gonzalez, C. Garrido, C. de Mendoza, R. Sanchez, P. Mallaina, and V. Soriano, “Improvement of genotypic algorithms to predict HIV-1 tropism in treatment-experienced patients - correlation with Trofile results,”

Antiviral Therapy, vol. 13, no. 4, pp. A109-A109, 2008.

[38] A. Thielen, P. R. Harrigan, A. J. Low, A. Moores, A. Altmann, K. Bozek, E. Heger, N. Sichtig, R. Kaiser, and T. Lengauer, "Improved genotypic prediction of HIV-1 coreceptor usage by incorporating V2 loop sequence variation," *Antiviral Therapy,* vol. 13, no. 4, pp. A100-A100, 2008.

[39] S. H. Xiang, B. Pacheco, D. Bowder, W. Yuan, and J. Sodroski, "Characterization of a dual-tropic human immunodeficiency virus (HIV-1) strain derived from the prototypical X4 isolate HXBc2," *Virology,* vol. 438, no. 1, pp. 5-13, Mar 30, 2013.

[40] N. G. Hoffman, F. Seillier-Moiseiwitsch, J. Ahn, J. M. Walker, and R. Swanstrom, "Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop," *J Virol,* vol. 76, no. 8, pp. 3852-64, Apr, 2002.

[41] "Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry," *J Virol.,* vol. 81, no. 5, pp. 2359-71. Epub 2006 Dec 20., Mar, 2007.

[42] J. C. Tilton, C. B. Wilen, C. A. Didigu, R. Sinha, J. E. Harrison, C. Agrawal-Gamse, E. A. Henning, F. D. Bushman, J. N. Martin, S. G. Deeks, and R.

W. Doms, “A maraviroc-resistant HIV-1 with narrow cross-resistance to other CCR5 antagonists depends on both N-terminal and extracellular loop domains of drug-bound CCR5,” *J Virol,* vol. 84, no. 20, pp. 10863-76, Oct, 2010.

[43] R. Nedellec, M. Coetzer, M. M. Lederman, R. E. Offord, O. Hartley, and D. E. Mosier, “Resistance to the CCR5 inhibitor 5P12-RANTES requires a difficult evolution from CCR5 to CXCR4 coreceptor use,” *PLoS One,* vol. 6, no. 7, pp. e22020, 2011.

[44] M. Armand-Ugon, G. Moncunill, E. Gonzalez, M. Mena, E. Ballana, B. Clotet, and J. A. Este, “Different selection patterns of resistance and cross-resistance to HIV-1 agents targeting CCR5,” *J Antimicrob Chemother,* vol. 65, no. 3, pp. 417-24, Mar, 2010.

[45] J. D. Reeves, J. L. Miamidian, M. J. Biscone, F. H. Lee, N. Ahmad, T. C. Pierson, and R. W. Doms, “Impact of mutations in the coreceptor binding site on human immunodeficiency virus type 1 fusion, infection, and entry inhibitor sensitivity,” *J Virol,* vol. 78, no. 10, pp. 5476-85, May, 2004.

[46] M. Roche, H. Salimi, R. Duncan, B. L. Wilkinson, K. Chikere, M. S. Moore, N. E. Webb, H. Zappi, J. Sterjovski, J. K. Flynn, A. Ellett, L. R. Gray, B. Lee, B. Jubb, M. Westby, P. A. Ramsland, S. R. Lewin, R. J. Payne, M. J. Churchill, and P. R.

Gorry, “A common mechanism of clinical HIV-1 resistance to the CCR5 antagonist maraviroc despite divergent resistance levels and lack of common gp120 resistance mutations,” *Retrovirology,* vol. 10, pp. 43, 2013.

[47] V. Sanchez, M. Masia, C. Robledano, S. Padilla, J. M. Ramos, and F. Gutierrez, “Performance of genotypic algorithms for predicting HIV-1 tropism measured against the enhanced-sensitivity Trofile coreceptor tropism assay,” *J Clin Microbiol,* vol. 48, no. 11, pp. 4135-9, Nov, 2010.

[48] C. Pastore, R. Nedellec, A. Ramos, S. Pontow, L. Ratner, and D. E. Mosier, “Human immunodeficiency virus type 1 coreceptor switching: V1/V2 gain-of-fitness mutations compensate for V3 loss-of-fitness mutations,” *J Virol,* vol. 80, no. 2, pp. 750-8, Jan, 2006.

[49] C. Balasubramanian, G. Chillemi, I. Abbate, M. R. Capobianchi, G. Rozera, and A. Desideri, “Importance of V3 Loop Flexibility and Net Charge in the Context of Co-Receptor Recognition. A Molecular Dynamics Study on HIV gp120,” *J Biomol Struct Dyn,* vol. 29, no. 5, pp. 1-13, Apr, 2012.

[50] J. J. De Jong, A. De Ronde, W. Keulen, M. Tersmette, and J. Goudsmit, “Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution,” *J*

Virol, vol. 66, no. 11, pp. 6777-80, Nov, 1992.

[51] K. Cashin, L. R. Gray, M. R. Jakobsen, J. Sterjovski, M. J. Churchill, and P. R. Gorry, "CoRSeqV3-C: a novel HIV-1 subtype C specific V3 sequence based coreceptor usage prediction algorithm," *Retrovirology,* vol. 10, pp. 24, 2013.

[52] D. G. Myszka, R. W. Sweet, P. Hensley, M. Brigham-Burke, P. D. Kwong, W. A. Hendrickson, R. Wyatt, J. Sodroski, and M. L. Doyle, "Energetics of the HIV gp120-CD4 binding reaction," *Proc Natl Acad Sci U S A,* vol. 97, no. 16, pp. 9026-31, Aug 1, 2000.

[53] M. M. Laakso, F. H. Lee, B. Haggarty, C. Agrawal, K. M. Nolan, M. Biscone, J. Romano, A. P. Jordan, G. J. Leslie, E. G. Meissner, L. Su, J. A. Hoxie, and R. W. Doms, "V3 loop truncations in HIV-1 envelope impart resistance to coreceptor inhibitors and enhanced sensitivity to neutralizing antibodies," *PLoS Pathog,* vol. 3, no. 8, pp. e117, Aug 24, 2007.

[54] R. A. Fouchier, M. Groenink, N. A. Kootstra, M. Tersmette, H. G. Huisman, F. Miedema, and H. Schuitemaker, "Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule," *J Virol,* vol. 66, no. 5, pp. 3183-7, May, 1992.

[55] S. Modrow, B. H. Hahn, G. M. Shaw, R. C. Gallo, F. Wong-Staal, and H. Wolf,

"Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions," *J Virol,* vol. 61, no. 2, pp. 570-8, Feb, 1987.

[56] C. Pastore, R. Nedellec, A. Ramos, O. Hartley, J. L. Miamidian, J. D. Reeves, and D. E. Mosier, "Conserved changes in envelope function during human immunodeficiency virus type 1 coreceptor switching," *J Virol,* vol. 81, no. 15, pp. 8165-79, Aug, 2007.

[57] A. A. Nabatov, G. Pollakis, T. Linnemann, A. Kliphius, M. I. Chalaby, and W. A. Paxton, "Intrapatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies," *J Virol,* vol. 78, no. 1, pp. 524-30, Jan, 2004.

[58] B. Labrosse, C. Treboute, A. Brelot, and M. Alizon, "Cooperation of the V1/V2 and V3 domains of human immunodeficiency virus type 1 gp120 for interaction with the CXCR4 receptor," *J Virol,* vol. 75, no. 12, pp. 5457-64, Jun, 2001.

[59] A. Carrillo, and L. Ratner, "Human immunodeficiency virus type 1 tropism for T-lymphoid cell lines: role of the V3 loop and C4 envelope determinants," *J Virol,* vol. 70, no. 2, pp. 1301-9, Feb, 1996.

[60] M. Groenink, R. A. Fouchier, S. Broersen, C. H. Baker, M. Koot, A. B. van't Wout, H. G. Huisman, F. Miedema, M. Tersmette, and H. Schuitemaker, “Relation of phenotype evolution of HIV-1 to envelope V2 configuration,” *Science,* vol. 260, no. 5113, pp. 1513-6, Jun 4, 1993.

[61] N. Wang, T. Zhu, and D. D. Ho, “Sequence diversity of V1 and V2 domains of gp120 from human immunodeficiency virus type 1: lack of correlation with viral phenotype,” *J Virol,* vol. 69, no. 4, pp. 2708-15, Apr, 1995.

[62] W. Huang, J. Toma, S. Fransen, E. Stawiski, J. D. Reeves, J. M. Whitcomb, N. Parkin, and C. J. Petropoulos, “Coreceptor tropism can be influenced by amino acid substitutions in the gp41 transmembrane subunit of human immunodeficiency virus type 1 envelope protein,” *J Virol,* vol. 82, no. 11, pp. 5584-93, Jun, 2008.

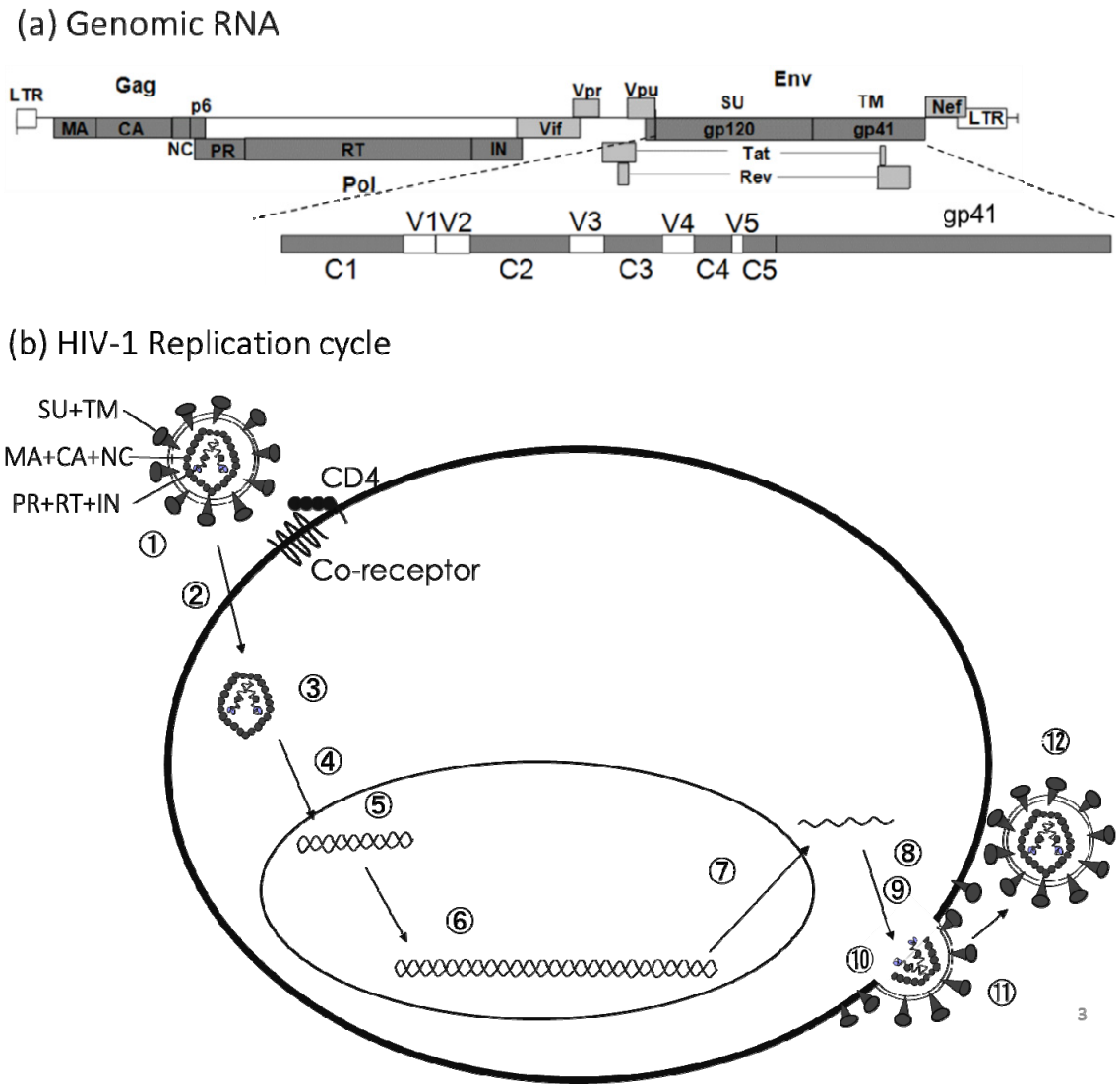
Table 1. Sequences of primers Env-1F and Env-3Rmix for Env amplification and PGK-F2, E70, E110, E130, E170, E250, Er35, Er65, Er115, Er145, Er155, and ori-seq-R2 for sequencing.

PCR primers	sequence
Env-1F	5′-TAGAGCCCTGGAAGCATCCAGGAAG-3′
Env-3Rmix	5′-TGCTGTATTGCTACTTGTGATTGCTCCATR-3′
EnvB-2F-XbaMlu	5′-TAGCTCTAGAACGCGTCTTAGGCATCTCCTATGGCAGGAAG-3′
EnvB-4R-AgeXho	5′-TAGCCTCGAGACCGGTTACTTTTTGACCACTTGCCACCCAT-3′
Sequencing primers	**sequences**
PGK-F2	5′- CTTCAAAAGCGCACGTCTGC-3′
E70	5′- GGGATCAAAGCCTAAAGCCATGTGTAA-3′
E110	5′- CTGTTAAATGGCAGTCTAGCAGAA-3′
E130	5′- ACAAATTATAAACATGTGGCAGG-3′
E170	5′- AGCAGGAAGCACTATGGG-3′
E250	5′- GGAGGCTTGATAGGTTTAAGAATA-3′
Er35	5′- GGTGAGTATCCCTGCCTAACTCTATT-3′
Er65	5′- AGTGCTTCCTGCTGCTCC-3′
Er115	5′- AGAAAAATTCCCCTCCACAATTAA-3′
Er145	5′- CAGCAGTTGAGTTGATACTACTGG-3′
Er155	5′- CTGTTCTACCATGTTATTTTTCCACATGT-3′
ori-seq-R2	5′- AAGCAGCAGATTACGCGCAG-3′
E510	5′- TACATAATGTTTGGGCCACAC-3′
E2350	5′- AGAGAAATTGACAATTACAC-3′

Table 2. The CD4+ cell count and viral load characteristics of enrolled patients.

IMS number	CD4+ (cells/µL)	Viral load (log_{10} copies/mL)
IMS0025	19	5.88
IMS0155	33	6.23
IMS0328	248	4.18
IMS0718	310	3.76
IMS0782	339	4.08
IMS1000	233	6.62

Fig. 1 The genomic RNA structure (a) with protein products and the replication cycle (b) of HIV-1.



HIV-1 genomic RNA includes three structural genes, gag, pol, and env, and six accessory genes. The structural genes encode the polypeptide of the Gag, Pol, and Env protein. By the cleavage of proteasome, precursor proteins are processed to MA, CA, and NA, PR, RT, and IN, SU, and TM, respectively. SU is also called gp120, and TM is also called gp41. The accessory genes encode for enzymes necessary for the infectivity and replication of HIV-1. The replication cycle of HIV-1 includes these steps with protein involved: ① Absorption (gp120), ② Membrane fusion・entry (gp41), ③ Uncoating, ④ Reverse transcript (RT), ⑤Nuclear import, ⑥Integration (IN), ⑦ Transcription (Tat), ⑧ Splicing and Nuclear export (Rev), ⑨ Translation, ⑩ Assembly, ⑪ budding (Vpu), ⑫ maturation (PR).

Fig. 2 Definition of the co-receptor usage of HIV-1.

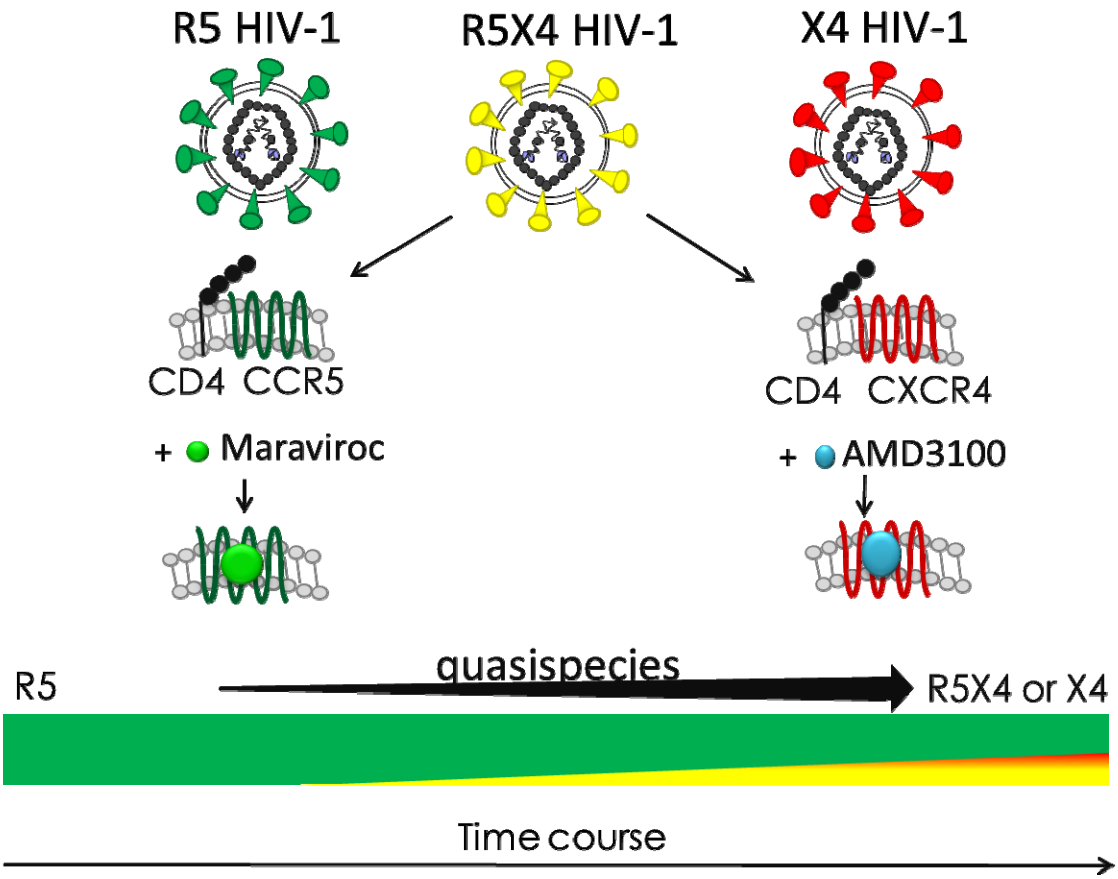
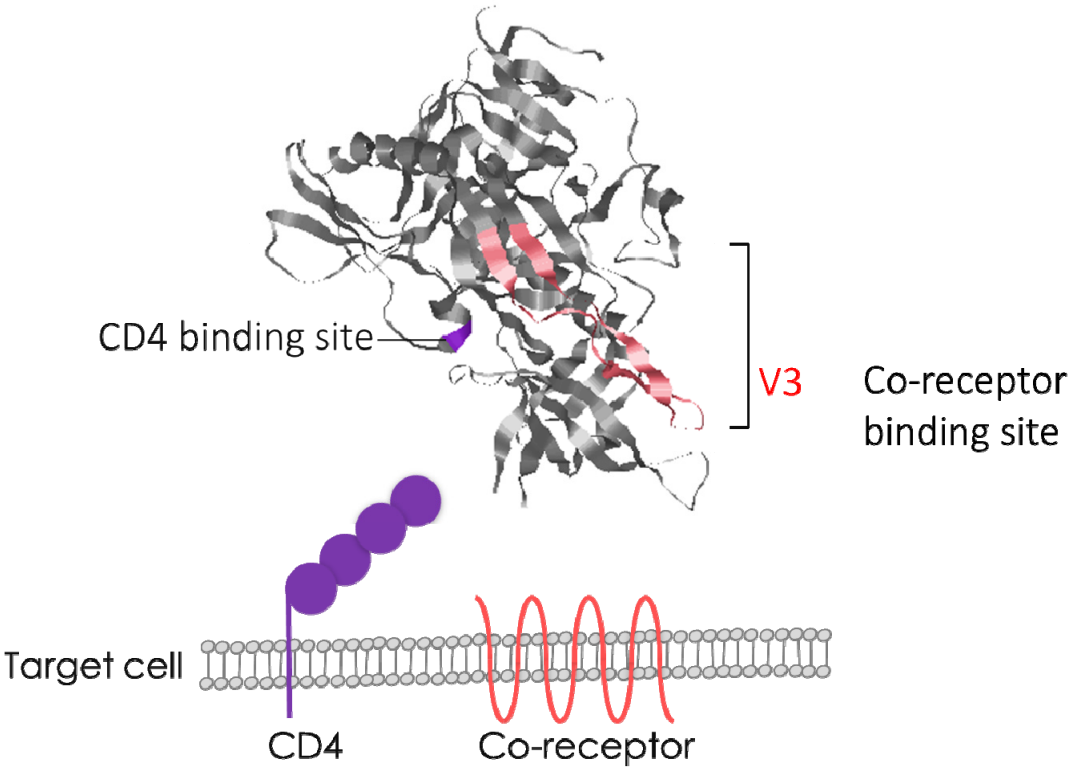


Fig. 3 The crystal structure of the envelope protein.



Pancera, M. et al. Nature, Vol.514, p.455,2014

Fig. 4 HIV-1 entry process involving the transformation of the envelope protein after binding to CD4 and co-receptor CCR5 or CXCR4.

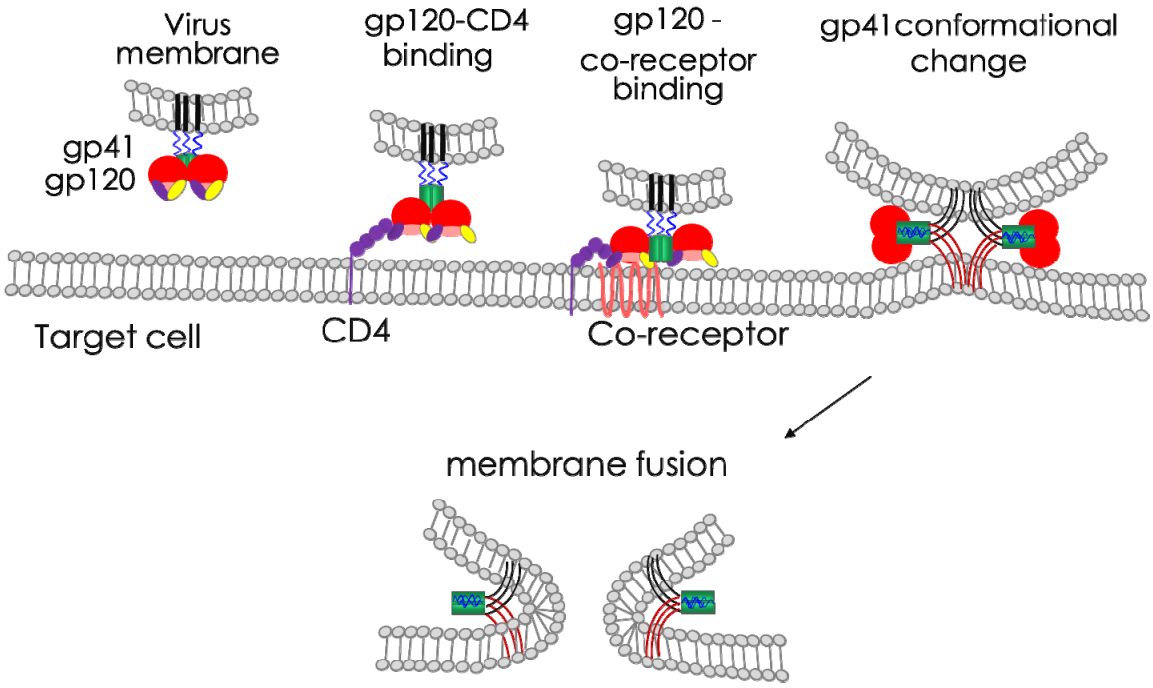


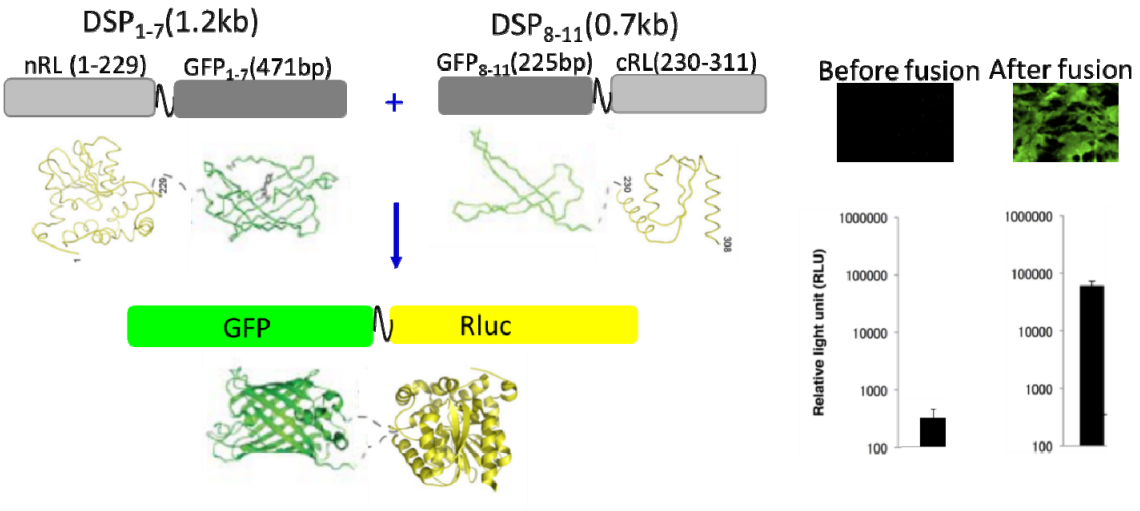
Fig. 5 Description of the three co-receptor assays.

- Genotypic assays – Geno2pheno assay ★ Approved by FDA
  (Lengauer, T., et al. Nat Biotechnol Vol. 25, p.1407, 2007)
- Phenotypic assays – (1) Virus-cell fusion assay ★ Approved by FDA
  (Trofile assay, Monogram, 2007)

  (2) Cell-cell fusion assay
  (Teeranaipong. P. et al. JIAS, vol. 16: p. 18723, 2013)

	Pros	Cons
Geno2pheno assay	Rapid (1~2 days), cheap	Information is only for V3 region
Virus-cell fusion assay	Mimic the natural infection	Biosafety requirement, cost, long turnaround time (1 month), only available in US, monopolized by Monogram
Cell-cell fusion assay	No virus shorter turnaround time (5 days)	To be determined

Fig. 6 Genetic composition, structure, and availability of DSP.



Kondo, N., et al. J Biol Chem Vol. 285: p.14681. 2010

Fig. 7 The schema of the virus-cell fusion assay.

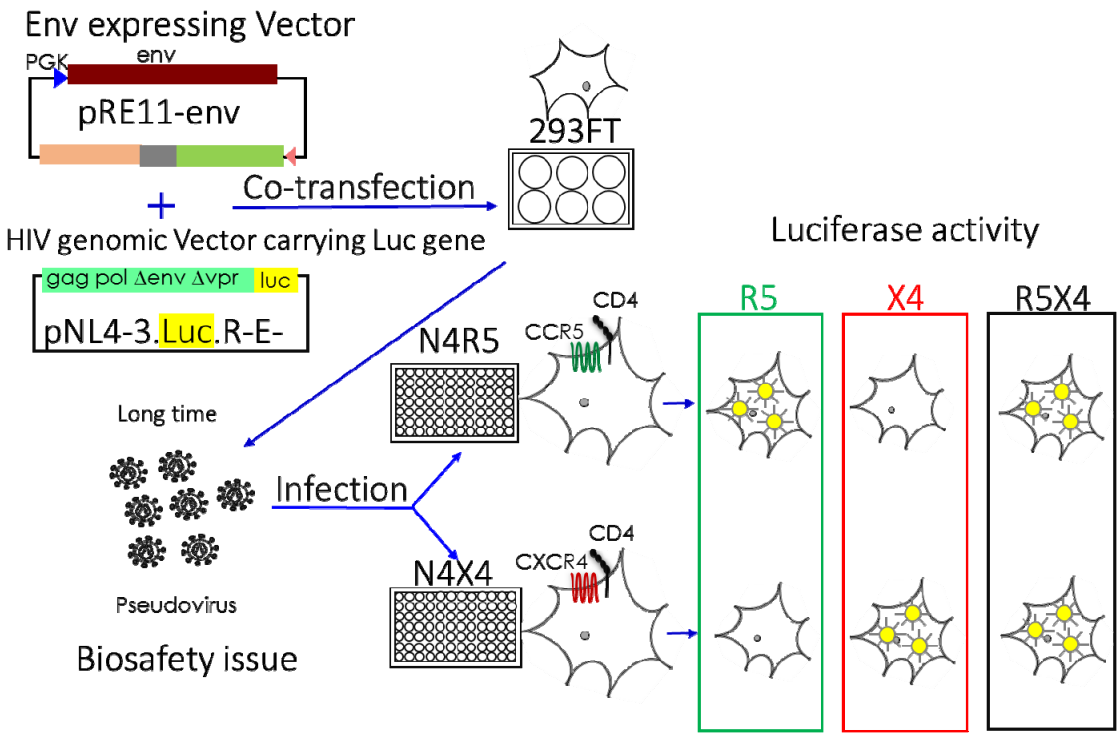
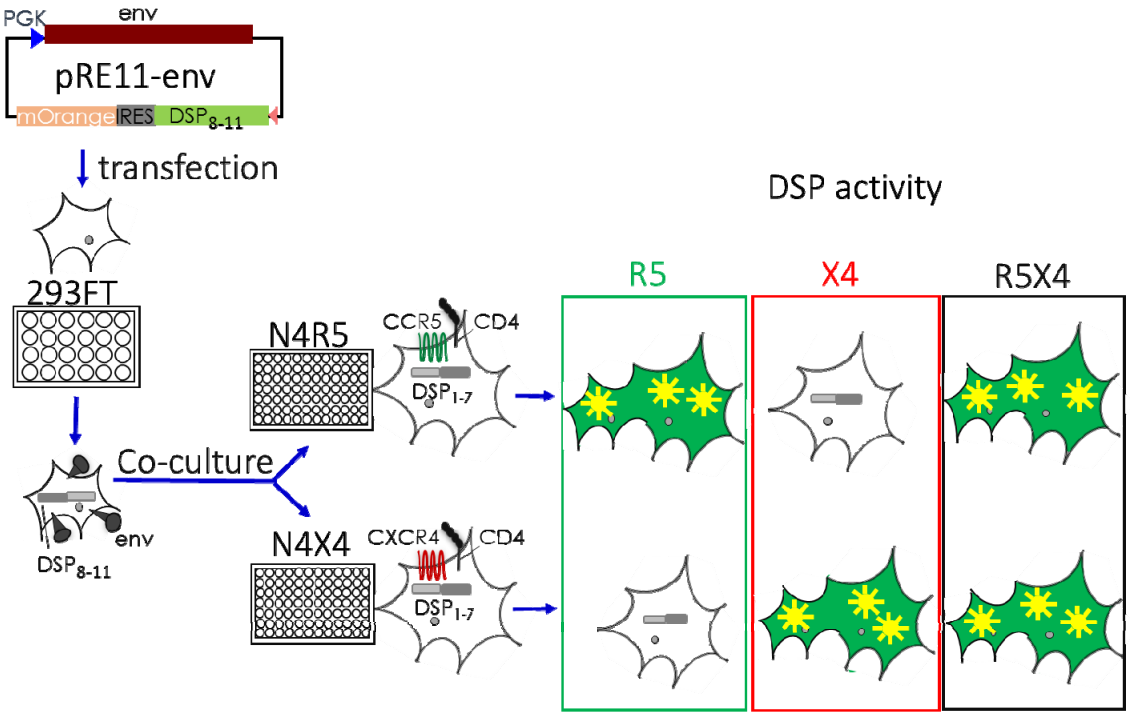
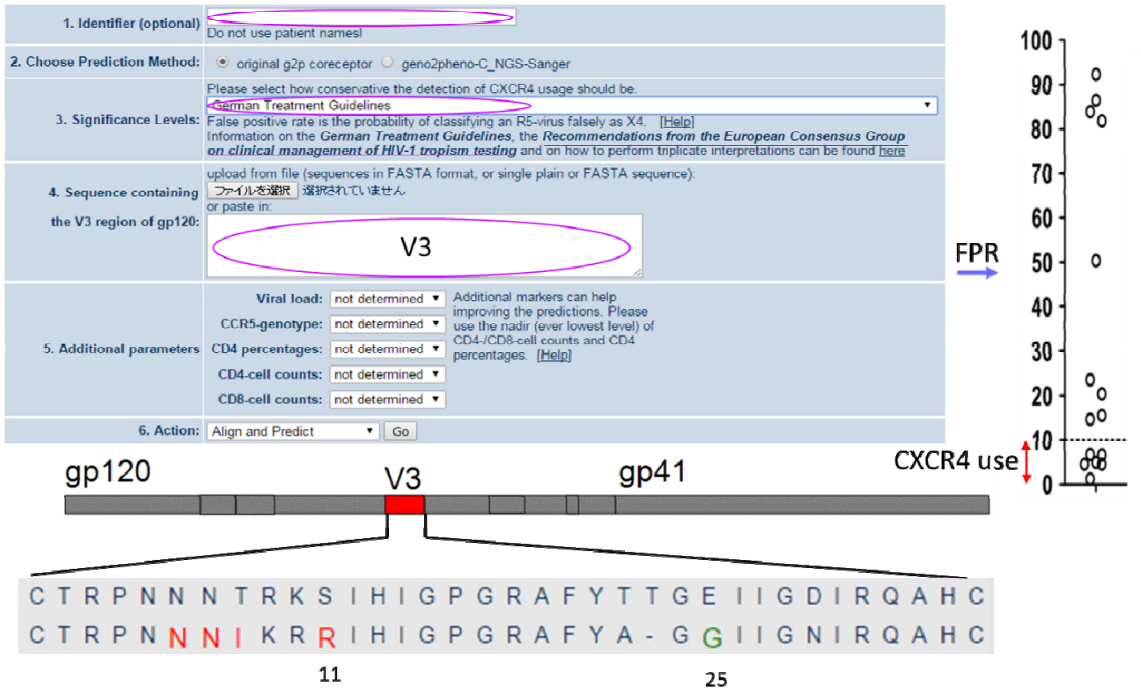


Fig. 8 The schema of the cell-cell fusion assay.



Teeranaipong, P., et al. J Int AIDS Soc Vol. 16, p. 18723, 2013

Fig. 9 Website interface of the Geno2pheno assay and the interpretation.



Co-receptor is predicted on the amino acid sequence of V3 in Env protein

Fig. 10 pRE11-based plasmid preparation, co-receptor assays, clonal analysis, and sequencing.

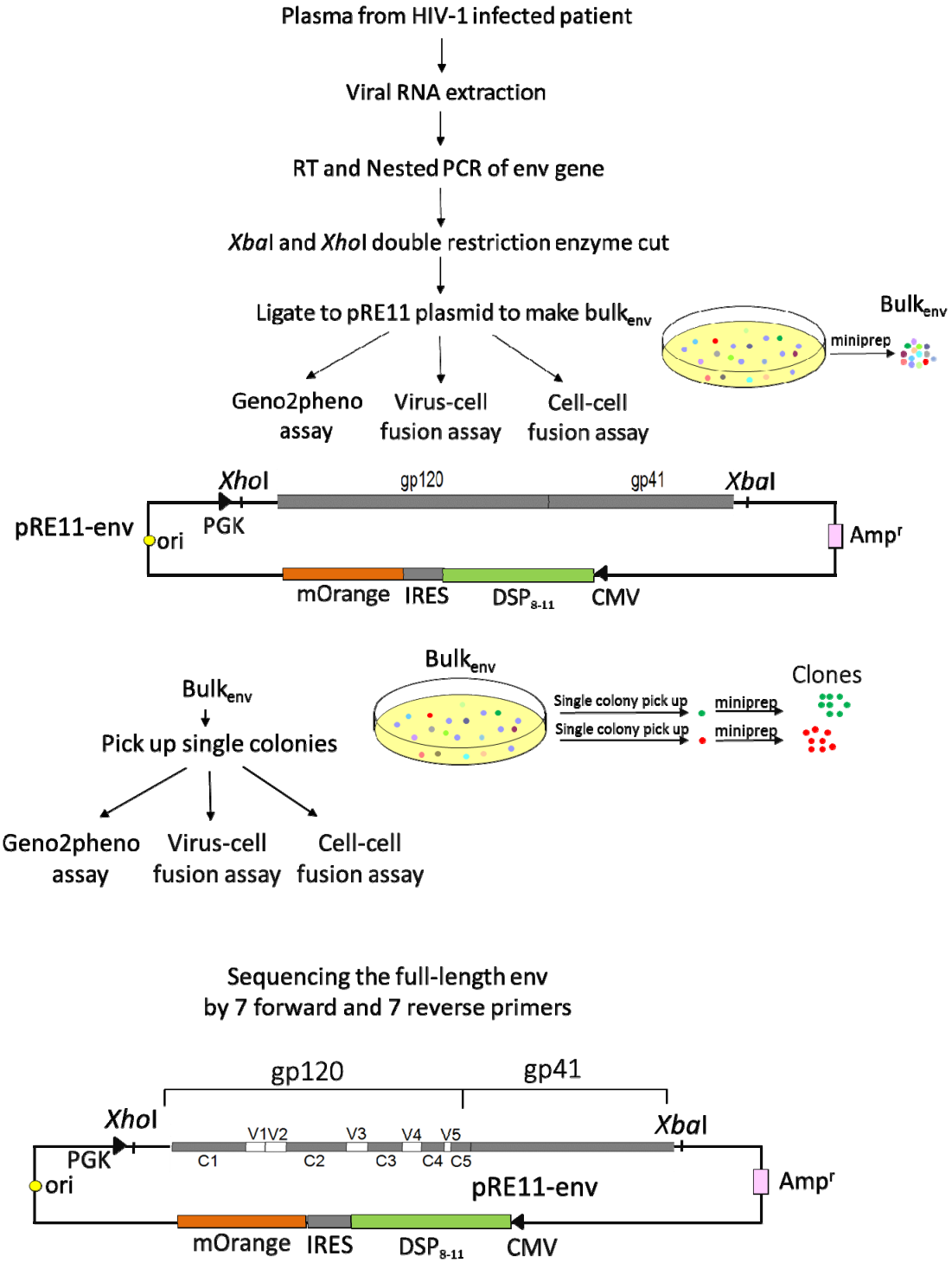
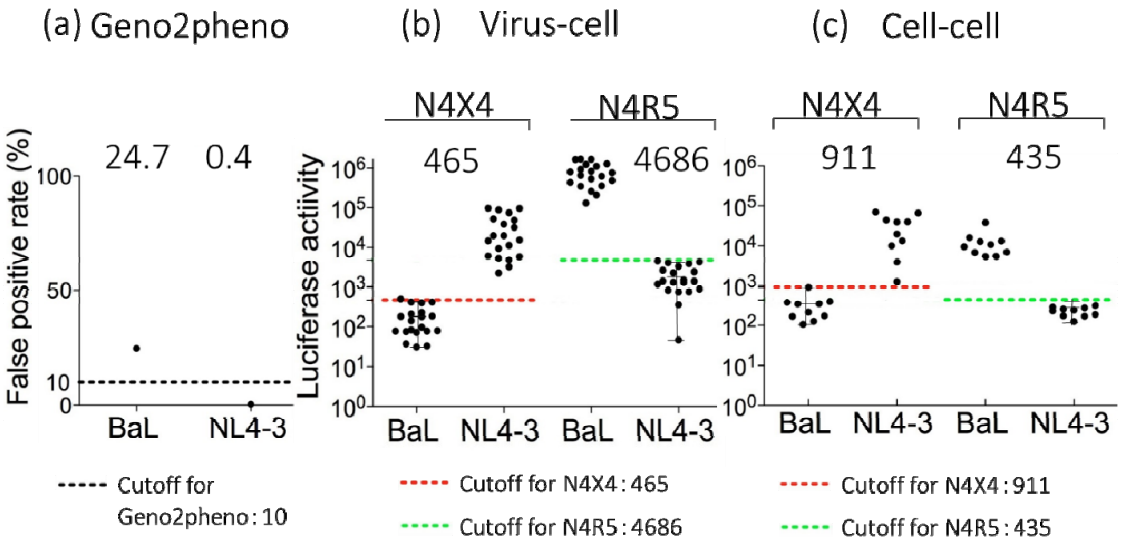
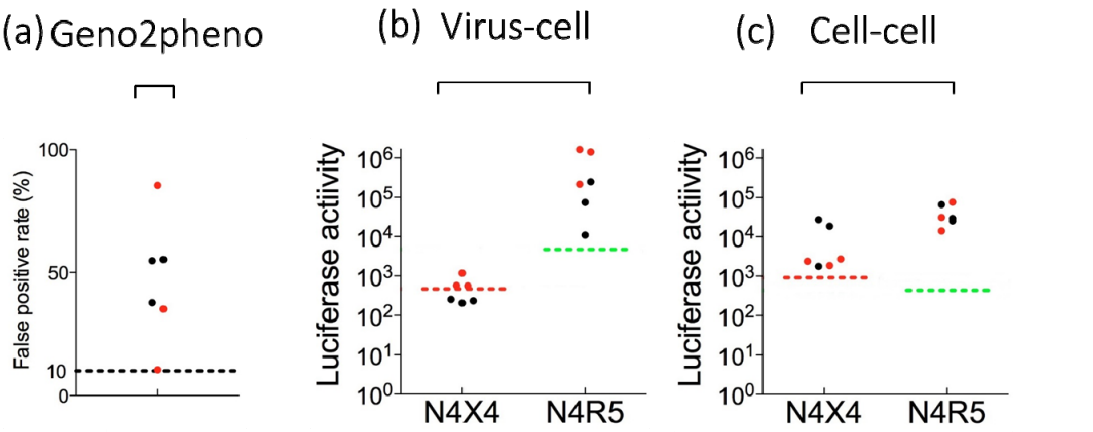


Fig. 11 Validation of (a) the Geno2pheno assay, (b) the virus-cell fusion assay, and (c) the cell-cell fusion assay with laboratory strains BaL and NL4-3, respectively.



The validation of the three co-receptor assays by laboratory strains BaL and NL4-3 are practiced ten to twenty times in the two phenotypic assays, respectively. Panel (a) shows the false positive rate of the two strains by Geno2pheno, and the values are 24.7 and 0.4, respectively, predicted to be R5 and X4. Panel (b) shows the 20 time repetition of the two strains in the virus-cell fusion assay. The horizontal axis shows BaL and NL4-3 infecting the N4X4 and N4R5 cells, respectively, and the vertical axis shows the luciferase activity. By calculation of the mean plus 2SD of the R5 strain in the N4X4 cell and the X4 strain in the N4R5 cells, the cutoff for the N4X4 cell and the N4R5 cells are calculated to be 465 and 4686 for luciferase activity. Panel (c) shows the 10 time repetition of the two strains in the cell-cell fusion assay. The cutoff is calculated to be 911 and 435, respectively.

Fig. 12 Results of bulk$_{envs}$ by Geno2pheno assay, virus-cell fusion assay, and cell-cell fusion assay.

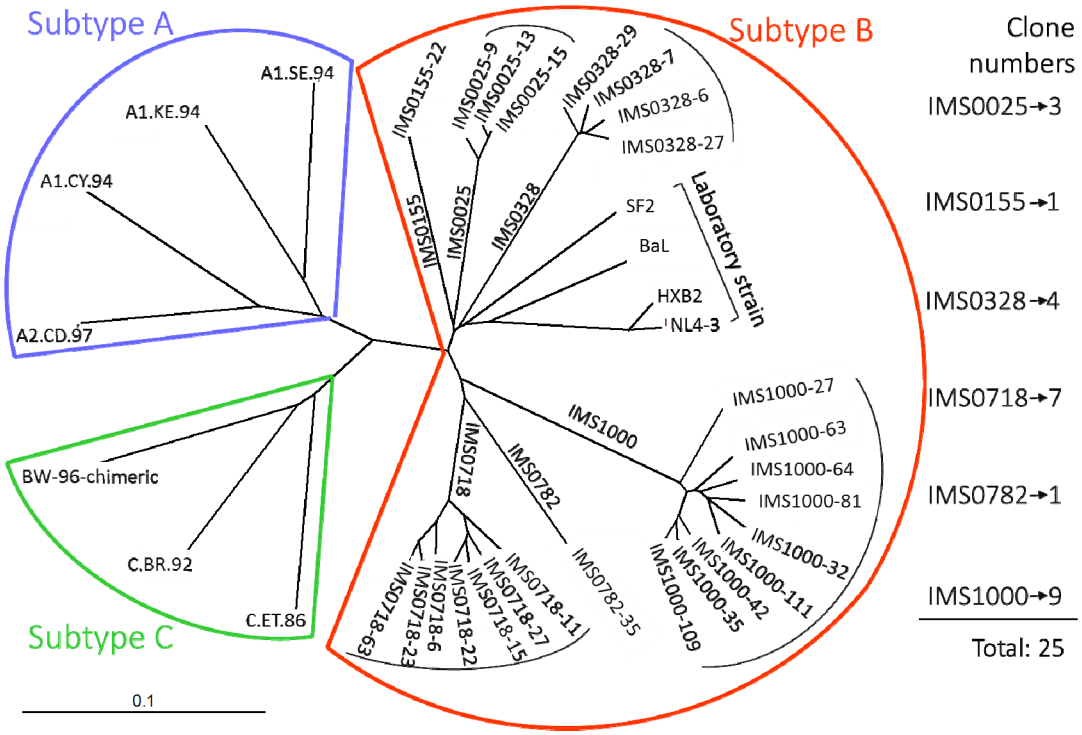


(d) Summery

	Geno2pheno	Virus-cell	Cell-cell
CCR5 only	6	3	0
CCR5 + CXCR4	0	3	6

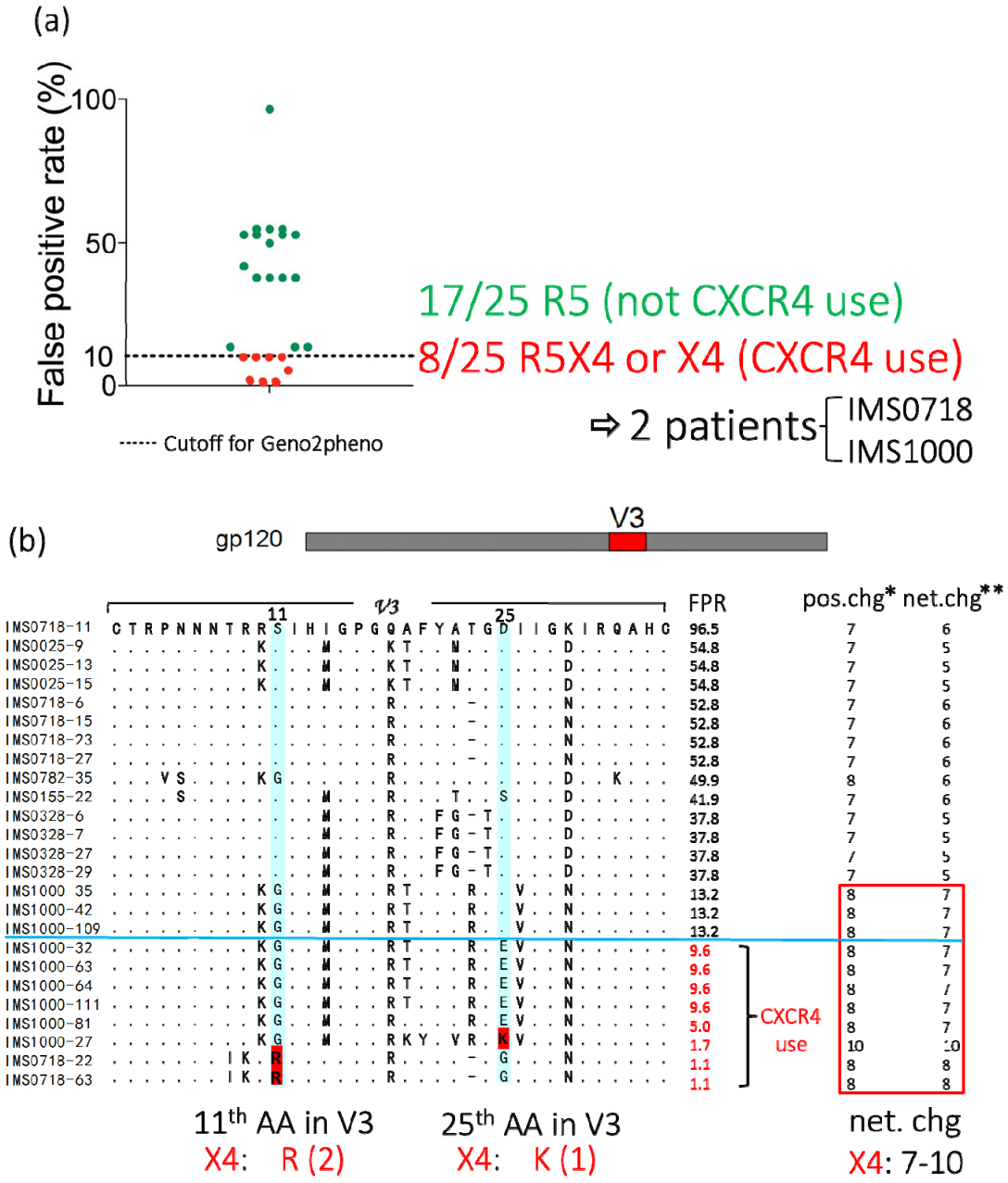
Envelope genes in 6 patients' plasma were amplified and subjected to three co-receptor assays.

Fig. 13 Phylogenetic tree of the full-length Env sequences.



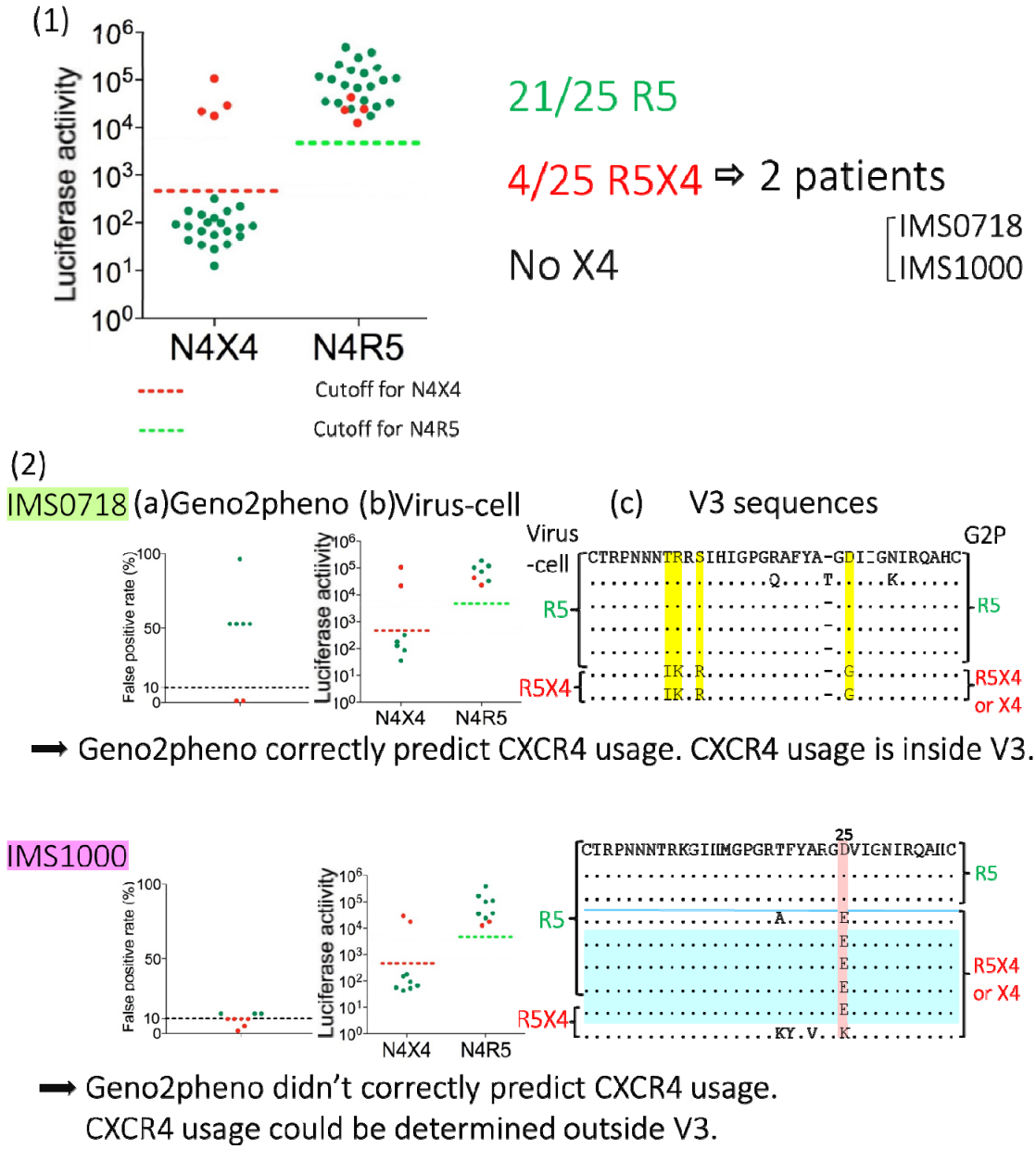
Plylogenetic tree of amino acid sequences of the whole envelope that was analyzed by clustalW 2.1 on the DNA DATA BANK of Japan (DDJB) with the HIV-1 reference genes representing subtypes A1, A2, B, and C drawn by TreeView.

Fig. 14 False positive rate and alignment of V3 amino acid sequences of the 25 clones.



In panel (a), false positive rates were shown for the 25 clones. Panel (b) showed the alignment of the V3 sequences. The horizontal line shows the boundary of 10% FPR. FPR under 10% are shown in red. The 11th and 25th amino acids of the V3 region are highlighted. The 11/25 rule means that the basic residues (R or K) at either or both of these sites is predictive of CXCR4 use.* pos.chg: total number of positively charged (R/K/H) amino acid residues. A higher positive charge has been correlated with the likelihood of CXCR4 use. **net.chg: number of positively charged (R/K/H) amino acid residues, minus the number of negatively charged (D/E) residues. A false positive rate indicates the probability of classifying an R5 virus falsely as an R5X4 or X4 virus.

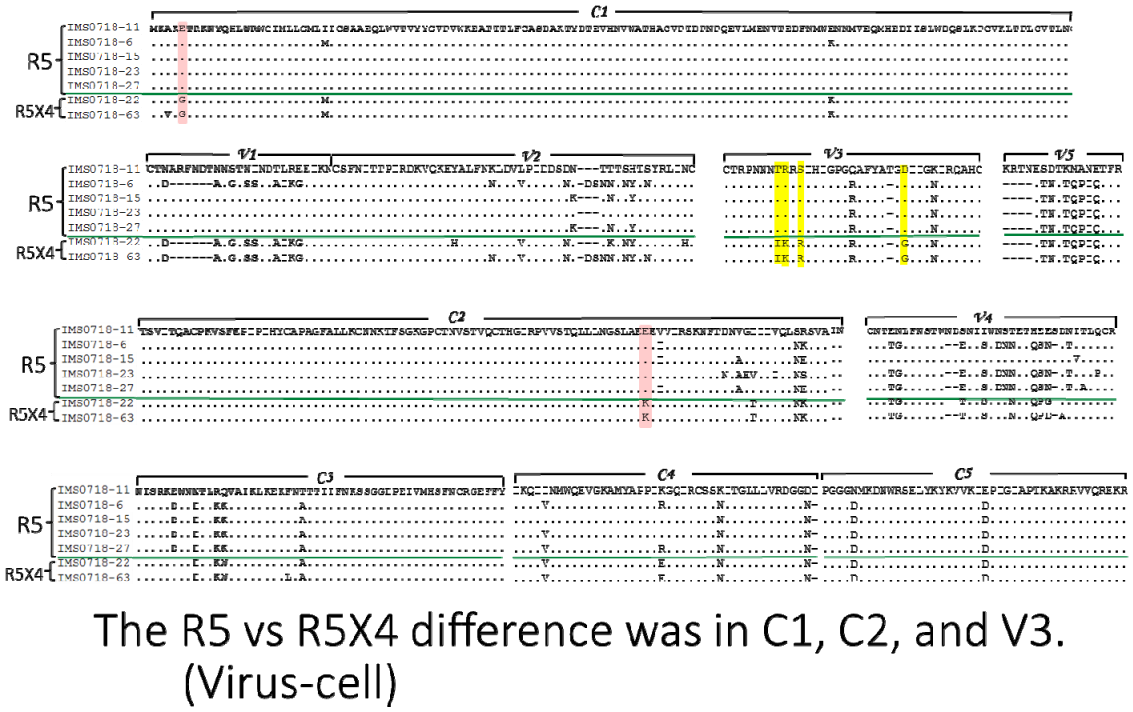
Fig. 15 Results of the virus-cell assay and the details of the R5X4 results for each individual.



Panel (1) shows the results of the 25 clones by the virus-cell fusion assay. The N4X4 cells were NP2 cell lines that were stably transfected CD4 and CXCR4. N4R5 cells were NP2 cell lines stably transfected CD4 and CCR5. Panel (2) shows clones from IMS0718 and IMS1000 by the Geno2pheno assay and the virus-cell fusion assay in the upper and lower panel, respectively. Panel (a) shows the result of the Geno2pheno assay, panel (b) shows the result of the virus-cell, and panel (c) shows the alignment of the V3 sequences. In IMS0718, the amino acid usage that correlated with the CXCR4 usage by two assays is highlighted in yellow. In IMS1000, the blue horizontal line is the boundary of 10% FPR. The amino acid usage that correlated with CXCR4 usage by the virus-cell

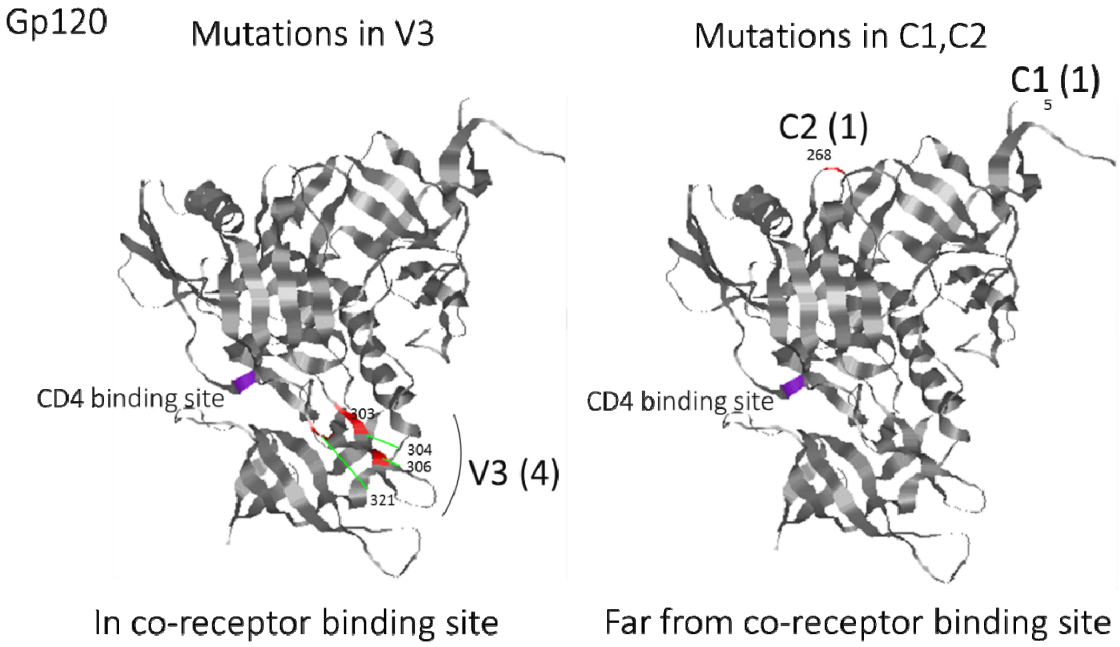
fusion assays is highlighted in pink.

Fig. 16 The alignment of the full-length Env by the virus-cell fusion assay. (Patient IMS0718)



Full-length Env alignment of clones from IMS0718. The clone numbers are shown in the left. The green line is the boundary of R5 sequences and R5X4 sequences. The R5 clones are shown in the upper, and the R5X4 clones are shown in the lower. The sequences were divided into C1 to V5 regions, respectively. Consistent mutations in V3 are shown in the yellow bar. Consistent mutations in C1 and C2 are shown in the pink bar.

Fig. 17 Positions of the consistent mutations of the R5X4 clones by the virus-cell fusion assay in the crystal structure of envelope. (Patient IMS0718)



The crystal structure of the envelope is shown. Purple shows the binding site for the CD4 molecule. Red shows the positions of the consistent mutations.

Fig. 18 Alignment of clones by the virus-cell fusion assay. (Patient IMS1000)

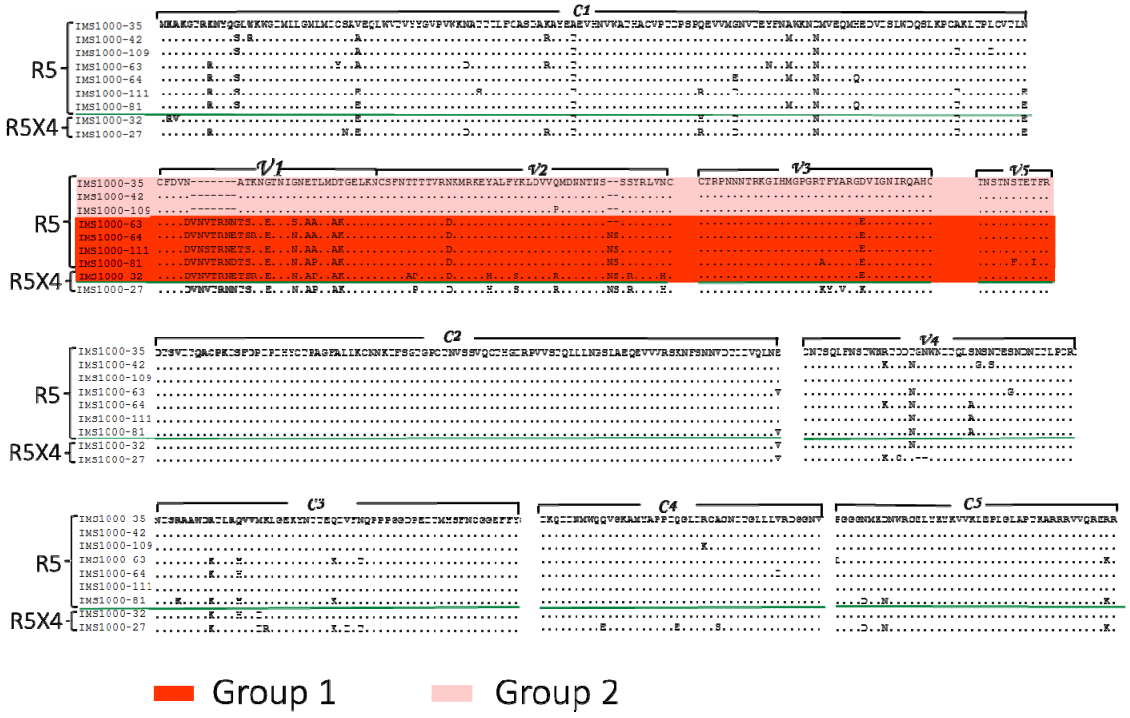
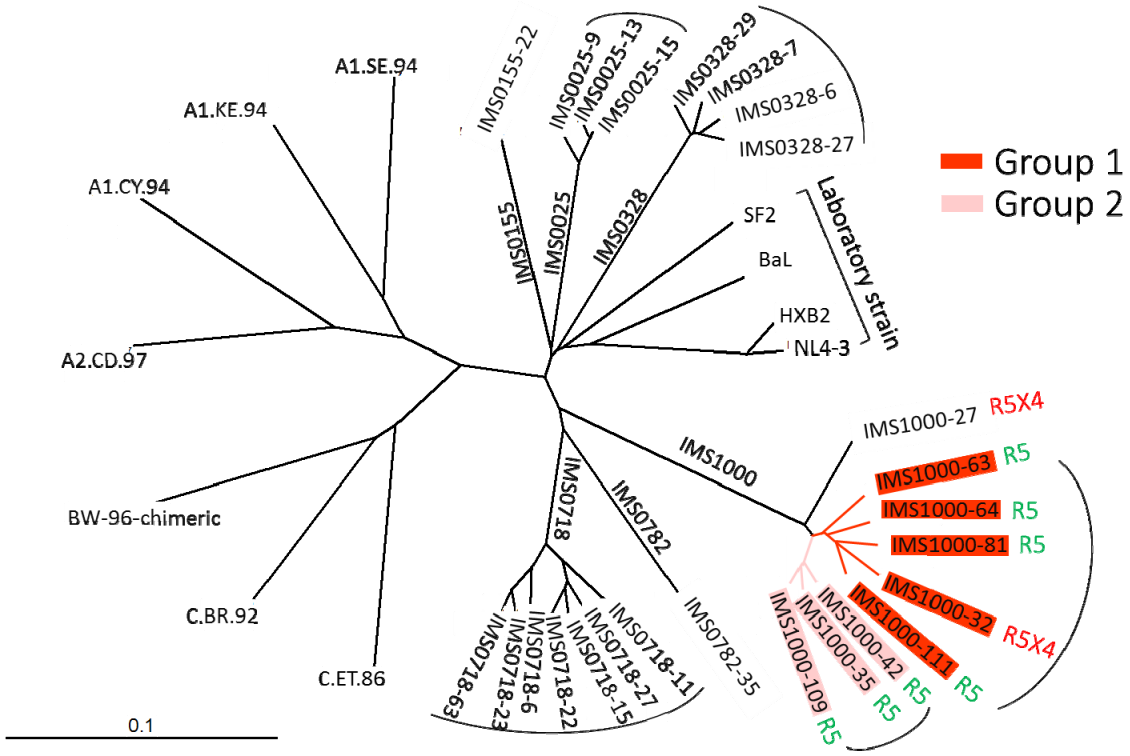
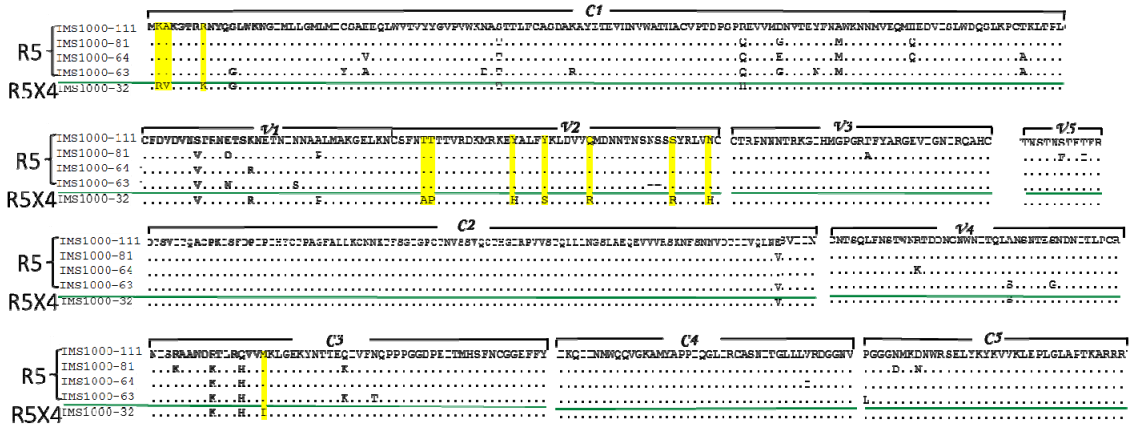


Fig. 19 Phylogenetic relationships of group1, group2, and IM1000-27 from IMS1000.



By the analysis of the phylogenetic tree, IMS1000 originated 9 clones that evolved into two groups and an isolated clone IMS1000-27. The clones in red with red branches are group1, and the clones in pink with pink branches are group2. Group 1 has one R5X4 clone, IMS1000-32, and four R5 clones. Group 2 has 3 R5 clones.

Fig. 20 Alignment of the clones from group 1 predicted by the virus-cell fusion assay. (Patient IMS1000)



The IMS1000-32 vs R5 difference was in C1, V2, and C3. (Virus-cell)

Full-length Env alignment of the clones from group1 of IMS1000. The clone numbers are shown on the left. The green line is the boundary of R5 sequences and R5X4 sequences. The R5 clones are shown in the upper of the green line, and the R5X4 clones are in the lower of the green line. The sequences were divided into C1 to V5 regions, respectively. The consistent mutations are shown in the yellow bar.

Fig. 21 Positions of the consistent mutations of the R5X4 clones of group1 in IMS1000 in the crystal structure of the envelope by the virus-cell fusion assay.

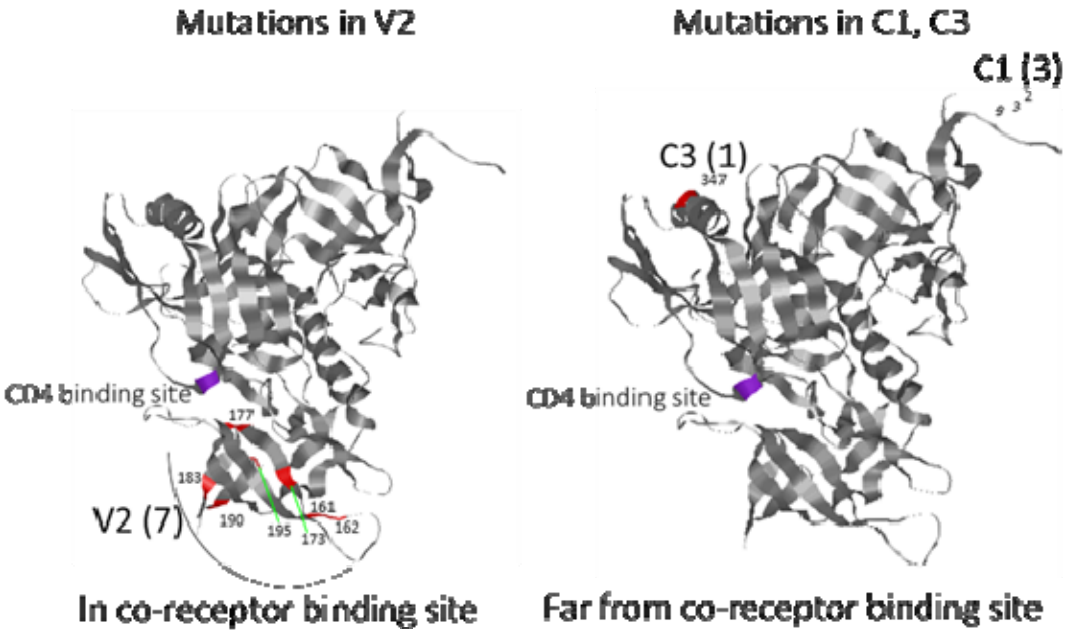
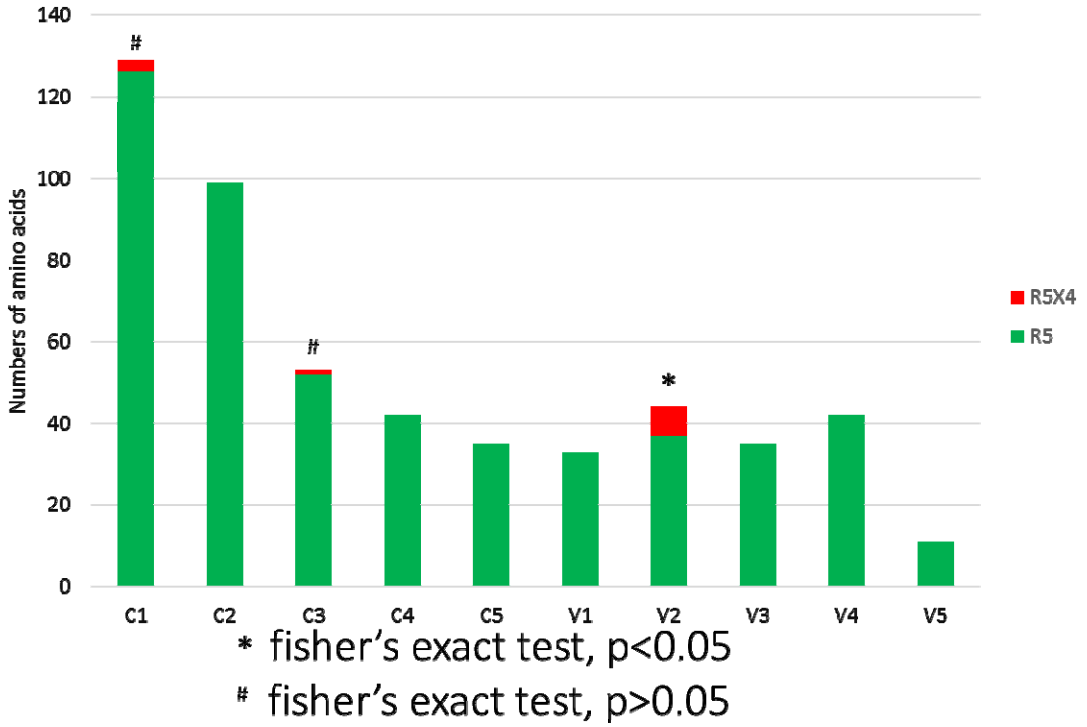
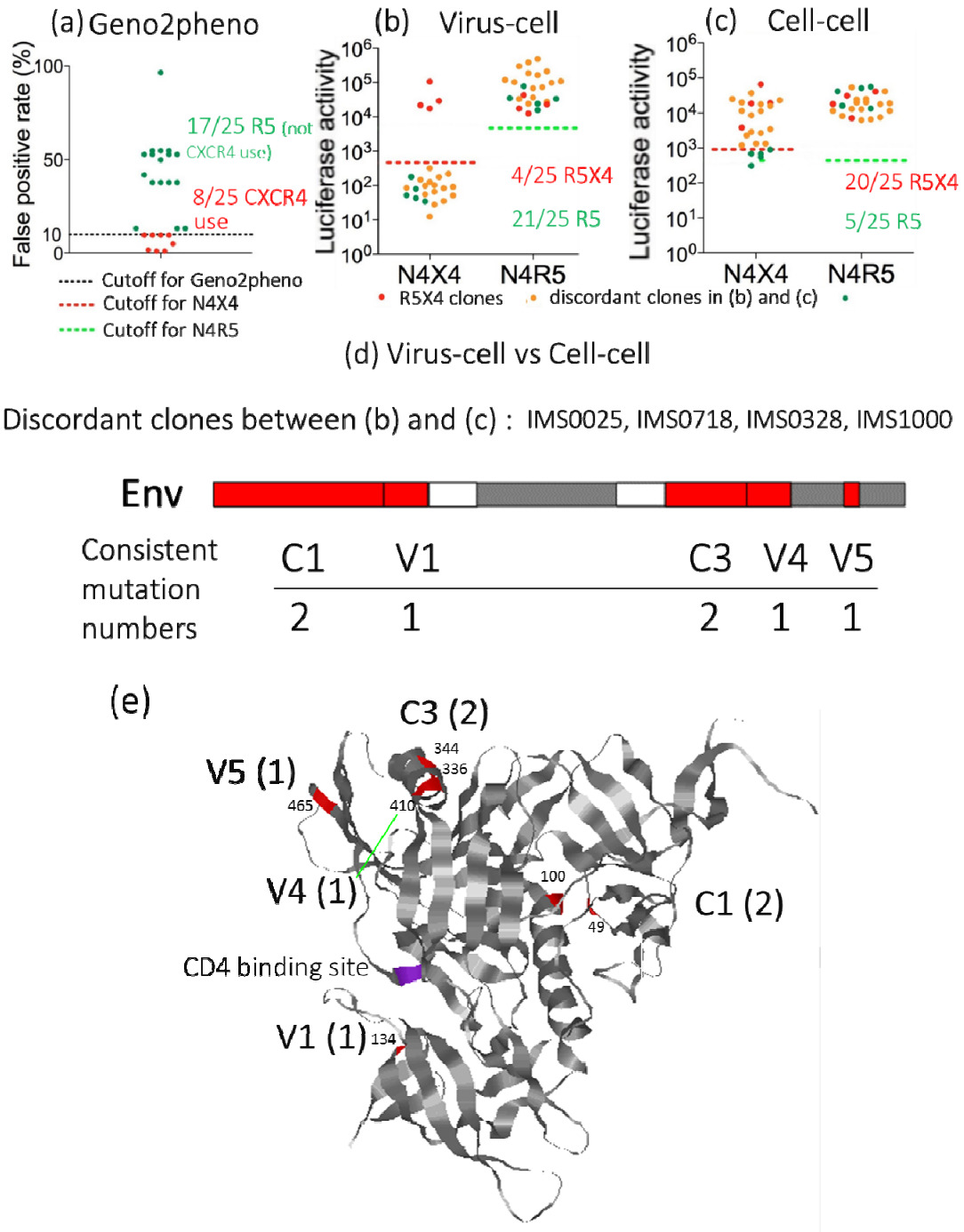


Fig. 22 Summary of the consistent mutations in each region and the statistical difference of the number of amino acid changes.



The number of the conserved amino acids by the virus-cell fusion assay in each appropriate region was analyzed by the contingency table. The mutation numbers of the R5 to the R5X4 change was compared between each group (C1, C2, C3, C4, C5, V1, V2, V3, V4, and V5). The total amino acid numbers were calculated in the contingency table. The dF was 9. Fisher's exact tests were applied to a null hypothesis that the 10 groups do not have differences. The ten groups showed that they had significant differences among them, $p<0.05$. Further analysis was conducted. The clustering of mutations in a certain region C1, V2, or C3 were compared with mutations/nonmutations in the full-length Env except C1, V2, and C3. The dF was 2. Fisher's exact tests were applied. When V2 were extracted to be compared with the remaining regions, it showed $p<0.05$. When C1 were extracted to be compared with the remaining regions except V2, it showed $p>0.05$. When C3 were extracted to be compared with the remaining regions except V2, it showed $p<0.05$. This means that the number of the consistent amino acid changes in V2 have a significant difference.

Fig. 23 Comparison of the cell-cell fusion assay with the virus-cell fusion assay.



Panel (a) is the result of the Geno2pheno assay, (b) is the result of the virus-cell assay, and (c) is the result of the cell-cell fusion assay of the 25 clones. The red dots in panel (b) and (c) show the concordant R5X4 clones by the cell-cell and virus-cell fusion assay. The orange dots show the discordant clones by the two assays. The green dots show the concordant R5 clones. Panel (d) shows the location of the consistent mutations. Panel (e) shows the sites of the consistent mutations in the crystal structure of gp120.