Using Ultradeep Pyrosequencing to Study HIV-1 Co-receptor Usage in Primary and Dual Infection

Gabriel A. Wagner¹, Mary E. Pacold², Edgar Vigil¹, Gemma Caballero³, Sheldon R. Morris¹, Sergei L. Kosakovsky Pond¹, Susan J. Little¹, Douglas D. Richman¹,³, Sara Gianella*,¹, and Davey M. Smith*,¹,³

¹University of California San Diego, La Jolla, California, USA
²Life Technologies, Foster City, California, USA
³Veterans Affairs San Diego Healthcare System, San Diego, California, USA

Corresponding author: Gabriel A. Wagner, 9500 Gilman Dr. 0679, La Jolla, CA 92093, Phone: 216-789-7059; Fax: 858-552-7445, Email: gawagner@ucsd.edu

*These authors contributed equally to this work
Abstract

HIV-1 dual infection (DI) and CXCR4 (X4) co-receptor usage are associated with accelerated disease progression but frequency and dynamics of co-receptor usage during DI is unknown. Ultradeep sequencing was used to interrogate for DI and infer co-receptor usage in longitudinal blood samples of 102 subjects. At baseline, X4 usage was high (23 subjects harbored X4 variants), and was not associated with infection duration or DI. Co-receptor usage changed over time in 12/47 participants, and X4 usage emerged in 4/41 monoinfections vs. 2/5 superinfections ($p=0.12$), suggesting a weak statistical trend towards occurrence of superinfection and acquiring X4 usage.
Introduction

HIV-1 dual infection (DI) is manifested as two or more distinct viral subpopulations within the same host. DI cases can be divided into co-infection, i.e. acquisition of at least two strains from different source partners simultaneously or within a brief period of time, and superinfection, i.e. acquisition of a second strain after an immune response to the first infection has been established [1]. DI has been associated with increased viral load, faster decline in CD4+ T-cells, and shorter time to AIDS diagnosis [2]. Similarly, infection with CXCR4 (X4) co-receptor-using virus has been linked to decreased CD4+ T-cell counts [3] and accelerated disease progression [4]. Although a strong preference for viral variants using the CCR5 (R5) co-receptor is displayed during HIV-1 transmission, the use of the alternative X4 co-receptor emerges in approximately half of individuals over the course of disease [5].

Ultradeep pyrosequencing (UDS) is a useful technique for increasing the sensitivity of HIV-1 co-receptor usage prediction [6], and for uncovering cases of DI [2]. This study exploits UDS and computational analyses to quantify the impact of DI on co-receptor usage in a large primary infection cohort of men who have sex with men (MSM).

Methods

Study Participants and Clinical Data

This study included MSM enrolled in the San Diego Primary Infection Cohort between January 1998 and January 2007, who deferred antiretroviral therapy (ART) for at least 6 months after their estimated date of infection (EDI), and had at least two blood plasma
samples available for UDS, one of which was collected within 3 years of EDI. At all time points, CD4 cell counts (LabCorp), and blood plasma HIV-1 RNA levels (Amplicor HIV-1 Monitor Test, Roche Molecular Systems Inc.) were quantified. EDI was calculated using baseline measurements following an established protocol [7].

**Sequencing and Dual Infection Screening**

HIV-1 RNA was isolated from blood plasma (QIAamp Viral RNA Mini Kit, Qiagen, Hilden, Germany), cDNA generated (RETROscript® Kit, Applied Biosystems/Ambion, Austin, TX, USA), and UDS (454 GS FLX Titanium Roche, Branford, CT) of env, gag, and pol was performed as previously described [2]. For each sample, the cDNA template input was calculated assuming 43% reverse transcription efficiency and was expressed as the number of templates (log10) for the first round of nested PCR. Average efficiency was validated by quantifying the input cDNA for 19 random samples from the total sample pool with real-time quantitative PCR (qPCR) and comparing these results with the HIV RNA viral load measurements (data not shown) [8]. UDS data were interrogated for DI with a previously published bioinformatics pipeline [2]. Notably, UDS can detect circulating minority variants as low as 0.25% of the population, and recent studies have found even lower limits of detection [9].

**Genotypic Co-receptor Usage Prediction**

Co-receptor usage was predicted using an online prediction tool, *geno2pheno 454* [6], for all samples containing the V3 env region. Based on the sensitivity of UDS, samples were conservatively classified as bearing X4-capable virus when *geno2pheno* predicted >1% of the viral population as X4 variants. Based on optimization studies by other groups [10, 11],
a geno2pheno false positive rate (FPR) level of 5% was selected. Prediction agreement with phenotypic co-receptor usage was in line with previous studies [12].

Statistical analyses

Fisher’s exact test was applied to detect associations between categorical data. Spearman’s rank correlation analysis was used to test for association between continuous variables. The differences in continuous variables (age, blood HIV RNA level, and CD4 cell count) between categories were evaluated with the Mann-Whitney test. For all statistical tests, a 2-tailed p-value ≤ 0.05 was considered statistically significant.

Results

Study cohort

HIV-infected participants (n=102) were predominantly white (82%) men with a median age of 31 years (IQR: 25.3–37 years) who reported sex with other men as their main risk factor for HIV transmission. At the first sampled timepoint, the median CD4+ T-cell count was 509 cells/ml (IQR: 412–708 cells/ml), the median time of baseline sample from EDI was 85 days (IQR: 45–85 days), and the median blood plasma level of HIV-1 was 4.76 HIV-1 RNA log_{10} copies/ml (IQR: 4.27–5.24 log_{10} copies/ml). The median calculated cDNA input in the first-round nested PCR was 3.42 log_{10} copies/10 ul (IQR: 3.13–3.90 log_{10} copies), and the median number of UDS reads passing quality filters was 4450 (IQR: 1911–8722). Based on analysis of UDS data, 16 cases of DI (11 co-infections and 5 superinfections) were
identified and investigated further.

**Co-receptor Usage Prediction and DI**

At the first sampled timepoint, 23 out of 102 (22.5%) participants harbored detectable X4-using variants in their viral population, and within the first three months from EDI (N=48), the prevalence of X4 usage was 18.8% (Figure 1). The median intrasample proportion of predicted X4 variants in participants harboring X4 usage was 11.9% (IQR: 1.8%−89.0%) observed at a median time since EDI of 3.1 months (IQR: 2.8−9.3 months). No association between detection of X4 variants and co-infection was found at baseline (Fisher’s exact test, *p*>0.29). The intrasample proportion of predicted X4-using variants was not significantly associated with time since EDI (Spearman’s rho, *p*>0.29). No difference was found in age, viral load, or CD4+ T-cell count between co-infection and monoinfection groups (Mann-Whitney Test, *p*>0.29). No differences were found in age, viral load, or CD4+ T-cell count between samples with predicted R5 and X4 viral populations (Mann-Whitney Test, *p*>0.29).

Forty-one monoinfected, one co-infected, and five superinfected individuals were followed longitudinally. Superinfection occurred at a median of 20.1 months from EDI (IQR: 7.0−27.7 months), with a median follow up of 24.3 months (IQR: 7.0−32.5 months). Median follow up for the monoinfection group was 12.4 months (IQR: 4.6−29.3 months). During follow up, the viral populations in 12 out of 47 subjects changed co-receptor usage (6 towards X4, and 6 towards R5). Viral co-receptor usage in 5 superinfected (and 1 co-infected) subjects did not have statistically significantly different rates of change (towards R5 or X4) compared to monoinfected individuals. For the 41 monoinfections, inferred co-receptor usage changed from R5 to X4 in 4 participants. For the 5 superinfections, predictions changed from R5 to
X4 in 2 participants (Fisher’s exact test, $p=0.12$). Viral co-receptor usage did not change in the single co-infected participant.

Interrogation of the first superinfection case where co-receptor usage changed (G59) revealed an R5-using viral population at 1.5 months after the EDI; however, X4-using variants were detected 18 months later (Figure 2A). Simultaneously, DI was confirmed. Geno2pheno predicted X4-using variants in the second, but not the initial strain.

Phylogenetic analysis demonstrated both strains present from 20 to 29 months after the EDI; however, the second strain disappeared at 31 months after the EDI, and concurrently the viral population returned to R5. In the second case (L78), the viral population was R5 at baseline, then low-level X4 variants were detected at 6.6 months after the EDI, but the superinfecting strain was not identified until three months later. At this timepoint, X4 variants decreased below 1% (Figure 2B).

**Discussion**

The use of next-generation sequencing in this study provided the opportunity to identify with more sensitivity both HIV-1 DI and minority variants with different co-receptor usage in a cohort of patients followed longitudinally after primary infection. The cohort had a high prevalence of inferred X4 usage at the first sampled timepoint (22.5%), similar to previous reports using UDS [12]. Baseline X4 usage in the cohort was not associated with DI or time after EDI. The analysis of subjects followed over time found that viral populations in 26% of subjects changed co-receptor usage during follow-up (13% towards X4, and 13% towards R5). When compared to a similar group of 41 monoinfected individuals followed longitudinally, the viral co-receptor usage in 5 superinfected (and 1
co-infected) subjects did not have statistically significantly different rates of change towards either R5 or X4. However, co-receptor usage did change towards X4 (>1% intrasample X4 variants) in 4 out of 41 (9.8%) monoinfections vs. 2 out of 5 (40%) superinfections \( p=0.12 \), suggesting a possible association between X4 usage gain and superinfection.

In the superinfection cases where inferred co-receptor usage changed, the second strain harbored X4 variants, and in the case of transient superinfection, X4 variants became undetectable with the disappearance of the second strain. Since superinfection can occur at any time during primary infection [2], high-risk (i.e. having multiple sex partners) R5-using infected individuals could develop X4-capable superinfection, even in the earliest stages of infection. If superinfection is transient, co-receptor usage can also revert to the prior phenotype as observed here. This rapid change in viral genetic diversity associated with DI has implications for co-receptor tropism testing in the clinical setting, which assumes that the viral population at the time of testing is an accurate representation of the moment when CCR5 antagonists are actually introduced into an individual’s antiretroviral regimen.

A limitation of this study is the lack of standardized cutoffs, both for the FPR level and for the X4 variant proportion above which an individual sample can be called X4 with confidence. Previous work in the MOTIVATE studies determined that for their UDS informatics pipelines, the FPR was 3.5% when a viral population had >2% predicted X4 variants [12]. These results set a good benchmark that should be independently replicated as the field matures. Another limitation is that while UDS is very sensitive at detecting DI, it is subject to experimental biases when preceded by PCR amplification [13]. A separate analysis of 29 cohort participants showed a correlation between env maximum diversity
and infection duration, while no association was observed between viral diversity and HIV RNA viral load or template input (data not shown). These observations suggest that the generated UDS reads are representative of the actual plasma virus populations in vivo. Other limitations include the limited number of participants with identifiable DI and the follow-up difference compared to monoinfected subjects, both of which can confound the detection of statistically significant differences in viral co-receptor usage; as well as the inevitable sampling bias that occurs in clinically collected specimens at discrete timepoints.

In summary, this study generated one of the largest sets of UDS data on a well-characterized cohort of HIV-1 infected MSM and analyzed the frequency and associations between DI and inferred X4-using virus. These observations raise the possibility of superinfection as a risk factor for more rapid development of X4-tropic infection. The study's findings can guide future lines of research, such as the actual frequency with which superinfection can change viral co-receptor usage, and the identification of genetic determinants of co-receptor usage outside the V3 loop.

Funding

This work was supported by the Department of Veterans Affairs and grants from: the National Institutes of Health [AI090970, AI100665, AI080353, MH097520, DA034978, MH83552, AI36214, MH62512, AI43638, AI74621, TW008908, AI69432, AI096113, AI47745, GM093939], the International AIDS Vaccine Initiative (IAVI); and the James B. Pendleton Charitable Trust. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Acknowledgements

We are grateful to Mehdi Bouhaddou for his assistance with data organization and analysis, Alexander Thielen for his technical support and helpful discussion, Caroline Ignacio for her wonderful technical support, and Demetrius Dela Cruz for his administrative assistance, as well as all the participants in the San Diego Primary Infection Cohort. Phenotypic HIV co-receptor tropism assays in this study were performed at Monogram Biosciences Clinical Reference Laboratory.
Potential Conflicts of Interest

GW does not have any commercial or other associations that might pose a conflict of interest. DDR has served as a consultant for Biota, Bristol-Myers Squibb, Chimerix, Gen-Probe, Gilead Sciences, Merck & Co, Monogram Biosciences, and Tobira Therapeutics. DMS has received research support from ViiV Pharmaceuticals and has served as a consultant to Gen-Probe. SKP has served as a consultant to Monogram Biosciences and Gen-Probe. The remaining authors do not report any conflicts of interest.

Financial Support

This work was supported by the Department of Veterans Affairs and grants from: the National Institutes of Health [AI090970, AI100665, AI080353, MH097520, DA034978, MH83552, AI36214, MH62512, AI43638, AI74621, TW008908, Al69432, AI096113, AI47745, GM093939], the International AIDS Vaccine Initiative (IAVI); and the James B. Pendleton Charitable Trust.

Presented in Part

The 19th Conference on Retroviruses and Opportunistic Infections, Seattle, Washington, March 5-8, 2012 (Abstract Poster 573).

Reprints or Correspondence

Gabriel A. Wagner, University of California San Diego, 9500 Gilman Dr. 0679, La Jolla, CA 92037; Phone: 216-789-7059, Fax: 858-552-7445, Email: gawagner@ucsd.edu
References

Figure Legends

Figure 1. Prevalence of predicted R5 and X4 co-receptor usage.

Participants first-sampled within three years after EDI harbored mostly R5 usage (light gray bars), but 23 out of 102 (22.5%) had detectable X4 variants (dark gray bars). A significant prevalence of X4 usage was detected at all time intervals, even in the first three months after infection (18.8%).

Abbreviations: R5 usage: viral population demonstrated R5 co-receptor usage; X4 usage: viral population with >1% X4 variants detected at a false positive rate of 5%; EDI: estimated date of infection.

Figure 2. HIV-1 superinfection and inferred X4 co-receptor usage over time.

In case G59 (A), an increase in the proportion of X4 usage (black solid diamonds) coincided with the detection of a superinfecting strain (pie charts), and X4 variants disappeared when this strain was no longer detected. In case L78 (B), low-level detection of X4 usage occurred right before detection of the superinfecting strain, and decreased below 1% when the second strain was detected. HIV-1 RNA viral load (gray squares) over time for each participant is superimposed.

Abbreviations: X4 usage: viral population with >1% X4 variants detected at a false positive rate of 5%; EDI: estimated date of infection.
Figure 1. Relative frequencies of predicted co-receptor usage

- R5 usage
- X4 usage

Percent of participants with R5- or X4-using virus

Month Intervals from EDI

0 - 3
3 - 6
6 - 12
12 - 24
24 - 36