Published in final edited form as:
AIDS. 2019 December 01; 33(15): 2299–2307. doi:10.1097/QAD.0000000000002390.

HIV-1 genetic diversity to estimate time of infection and infer adherence to preexposure prophylaxis

Olivia D. Council[a,*], **Susan Ruone**[a,*], **Philip A. Mock**[b], **George Khalil**[c], **Amy Martin**[a], **Marcel E. Curlin**[d], **Janet M. McNicholl**[a], **Walid Heneine**[a], **Wanna Leelawiwat**[b], **Kachit Choopanya**[e], **Suphak Vanichseni**[e], **Thitima Cherdtrakulkiat**[b], **Rapeepan Anekvorapong**[e], **Michael Martin**[b], **José Gerardo García-Lerma**[a]

[a]Laboratory Branch, Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, US Centers for Disease Control and Prevention, Atlanta, Georgia, USA

[b]Thailand Ministry of Public Health – U.S. Centers for Disease Control and Prevention Collaboration, Nonthaburi, Thailand

[c]Quantitative Sciences and Data Management Branch, Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia

[d]Department of Medicine, Division of Infectious Diseases, Oregon Health and Sciences University, Portland, Oregon, USA

[e]Bangkok Tenofovir Study Group, Bangkok, Thailand.

Abstract

Objective: To estimate time of HIV infection in participants from the Bangkok Tenofovir Study (BTS) with daily oral tenofovir disoproxil fumarate (TDF) for preexposure prophylaxis (PrEP) and relate infection with adherence patterns.

Design: We used the diversity structure of the virus population at the first HIV RNA-positive sample to estimate the date of infection, and mapped these estimates to medication diaries obtained under daily directly observed therapy (DOT).

Methods: HIV genetic diversity was investigated in all 17 PrEP breakthrough infections and in 16 placebo recipients. We generated 10–25 HIV *env* sequences from each participant by single

---

Correspondence to José Gerardo García-Lerma, Laboratory Branch, Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, US Centers for Disease Control and Prevention, MS H18-2, 1600 Clifton Road, Atlanta, GA 30329, USA. jng5@cdc.gov.

[*]Olivia D. Council and Susan Ruone contributed equally to this work.

Author contributions: O.D.C., S.R. and J.G.G.-L. conceived and designed the experiments. O.C., S.R., P.M., A.M., M.E.C., J.M., W.H., W.L., K.C., S.V., T.C., R.A. and M.M acquired data and/or performed the experiments. O.C., S.R., G.K., M.M. and J.G.G.-L. analyzed the data. J.G.G.-L. wrote the article.

Disclaimer: The findings and conclusions of this manuscript are those of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention.

Conflicts of interest

J.G.G.-L. and W.H. are named on US Government patents on ‘Inhibition of HIV infection through chemoprophylaxis’, and on patent applications on ‘HIV post-exposure prophylaxis’ and ‘HIV pre-exposure prophylaxis’.

genome amplification, and calculated time since infection (and 95% confidence interval) using Poisson models of early virus evolution. Study medication diaries obtained under daily DOT were then used to compute the number of missed TDF doses at the approximate date of infection.

Results: Fifteen of the 17 PrEP breakthrough infections were successfully amplified. Of these, 13 were initiated by a single genetic variant and generated reliable estimates of time since infection (median = 47 [IQR = 35] days). Eleven of these 13 were under daily DOT at the estimated time of infection. Analysis of medication diaries in these 11 participants showed 100% adherence in five, 90–95% adherence in two, 55% adherence in one, and nonadherence in three.

Conclusion: We estimated time of infection in participants from BTS and found several infections when high levels of adherence to TDF were reported. Our results suggest that the biological efficacy of daily TDF against parenteral HIV exposure is not 100%.

Keywords

adherence; HIV-1 diversity; preexposure prophylaxis; prevention; tenofovir

Introduction

With approximately 37.9 million people living with HIV-1 at the end of 2018 and 1.7 million people newly infected in 2018, the HIV-1 epidemic continues to be a major public health problem [1]. Daily preexposure prophylaxis (PrEP) with the combination of oral emtricitabine (FTC) and TDF is highly effective in preventing HIV-1 infection [2]. However, for PrEP to be effective, high medication adherence is required. Antiretroviral drug detection in plasma or peripheral blood mononuclear cells (PBMCs) is considered the best indicator of adherence, and in major PrEP trials was a good predictor of PrEP effectiveness [3–8].

Although assessing adherence by measuring drug concentrations in plasma or PBMCs has become essential for understanding PrEP effectiveness, this approach has limitations. Levels of tenofovir or FTC in plasma provide information about recently ingested doses, but not about consistency of drug dosing over time. Concentrations of tenofovir diphosphate (TFV-DP) in PBMCs or dried blood spots provide information about pill-taking behavior and mean number of doses taken per week during the preceding 1–2 or 4–6 weeks, respectively [9–11]. However, these measurements provide average TFV-DP levels and cannot capture cycles of low adherence that may associate with increased HIV risk. A recent modeling study showed that PrEP adherence patterns had a substantial impact on protection even when the same number of pills was taken [12]. Optimally, measures of adherence should rely on pill-taking behavior and drug levels at the precise time of HIV-1 infection. This analysis can be done in trials where PrEP is administered under daily directly observed therapy (DOT) and HIV-1 testing is done frequently enough to identify persons in early stages of HIV acquisition (Fiebig stages I or II of infection).

In this study, we sought to characterize pill-taking behavior at the approximate date of HIV infection in participants from the Bangkok Tenofovir Study (BTS) study with daily oral TDF for PrEP among persons who inject drug (PWID) [6]. We first created adherence maps using medication diaries of participants who came to the clinics daily to receive study drug (daily DOT) in the 3 months prior to seroconversion. We then calculated the approximate date of HIV infection (and 95% confidence interval) for each of these participants using Poisson models of early virus evolution. The Poisson-Fitter and Shifted Poisson Mixture models use the nucleotide diversity of the HIV-1 population to provide estimates of time since infection with 95% confidence intervals [13,14]. The underlying assumption is that during the early weeks of HIV-1 infection and prior to the onset of host immune responses, the virus population grows exponentially with a random accumulation of mutations [15]. The models are not influenced by viremia levels, assume a virus generation time of 2 days, a mutation rate of 2.16×10^{-5} mutations per base pair per cycle, and a basic reproductive ratio of 6 [16–18]. These mathematical tools have been extensively validated using HIV *env* sequences from individuals at Fiebig stages I–V of infection, and provide estimates of time since infection that generally fall within the range of Fiebig-based estimates [13,14,19,20].

We mapped the approximate date of infection in PrEP breakthrough infections from BTS with medication diaries generated under DOT, and document infections when high levels of adherence to TDF were reported.

Methods

Study samples

The BTS trial was a randomized, double-blind, placebo-controlled trial that evaluated the safety and efficacy of daily TDF in reducing HIV-1 transmission among PWID. In this trial, 2413 participants were randomly assigned to daily TDF or placebo and could choose daily DOT (including weekends) or monthly visits without DOT. Participants were on DOT an average of 86.9% of the time with a median adherence on DOT of 94.8% [6].

We attempted to characterize HIV-1 genetic diversity in all 17 incident infections in the TDF group, and in 16 of the 33 incident infections from the placebo group. Of the 17 participants in the TDF group, amplification was not successful in two (TDF16 and TDF17, Table 1). Genetic diversity was investigated in plasma collected at the first HIV-1 RNA-positive visit. This visit coincided with the seroconversion visit (Genetic Systems HIV-1/HIV-2 plus O enzyme-immunoassay) in 12 of the 15 infections in the TDF group, and in 15 of the 16 infections in the placebo group [6,21]. The concentrations of TFV in plasma collected at this visit were measured as previously described using validated high-performance liquid chromatography–tandem mass spectrometry methods with a lower limit of detection of 5 ng/ml [22].

Single genome amplification of HIV-1 *env* and sequence analysis

Single genome amplification was performed following a standard protocol using primers specific for HIV-1 subtype B or AE [23,24]. Briefly, 30 μl of plasma HIV-1 RNA were added to a cocktail containing reverse primer envB3out (subtype B) or KenvR1 (subtype AE) and Superscript III reverse transcriptase (SSIII reverse transcriptase; Invitrogen). Reverse transcription was done at 50 °C for 60 min and 55 °C for 60 min. The cocktail was heated at 70 °C for 15 min to inactivate SSIII reverse transcriptase, and then treated with RNase H at 37 °C for 20 min to remove the RNA template. Serial dilutions of the reverse transcriptase-generated cDNA were made to identify dilutions of cDNA resulting in less than 30% of positive wells by PCR. On the basis of a Poisson distribution, amplifications at this dilution represent sequences derived from a single cDNA template [25].

The first round of PCR amplification was performed using 1–2 μl of appropriately diluted cDNA and a cocktail containing Platinum *Taq* High-Fidelity (Invitro-gen, Carlsbad, California, USA) and primers envB5out/envB3out (subtype B) or K-R1 or NenvF1/K-R1 or NenvR1 (subtype AE) [23,24]. A 1 μl volume was then used for a nested PCR reaction using primer pair envB5in/envB3in (subtype B) or K-R2 or NenvF2/K-R2 or NenvR2 (subtype AE) [23,24]. PCR products were visualized on a 1% UltraPure Agarose gel (Invitrogen) using a BioRad SubCell GTgel apparatus and purified using the QIAquick PCR Purification Kit (Qiagen, German-town, MD USA). Sequences (~1 kb) representing the V1–V5 and flanking regions of HIV-1 *env* were obtained from 10 to 25 amplicons per sample. Sequencing was done with the BigDye Terminator v1.1 Cycle Sequencing kit in an ABI3130xl automated sequencer. We attempted to generate at least 20 sequences per participant although this was not always possible because of low viremias and limited plasma availability. Sequences with evidence of mixed bases (median of 1 (0–3) per participant) were excluded from the analysis.

Calculation of time since infection and computation of missed tenofovir disoproxil fumarate doses

The degree of virus genetic diversity seen in plasma corresponding to the first HIV-1 RNA positive clinic visit was used to estimate time since infection using the Poisson-Fitter tool or the Shifted Poisson Mixture Model (SPMM). Poisson-Fitter calculates time since infection in homogeneous infections, that is, initiated by a single genetic strain [13]. The tool (available at http://www.hiv.lanl.gov/content/sequence/POISSON_FITTER/poisson_fitter.html) detects if the distribution of mutations in a set of HIV-1 sequences is indicative of a population evolving from a single variant through random accumulation of mutations. When the model is satisfied, Poisson-Fitter provides an estimate of time since infection with 95% confidence intervals. A small goodness-of-fit P value (<0.1) denotes divergence from a Poisson distribution and is indicative of heterogeneous infection or early onset of immune selection pressure [13]. We excluded hypermutated positions and hypermutated sequences found to be significantly enriched with a P value less than 0.1. This correction controls for APOBEC3G/F-mediated mutation enrichment that occurs at a higher rate than the background mutation rate.

A second model called the SPMM was used to calculate time since infection in multivariant infections. This model (available at http://p512.usc.edu/) segregates distinct viral lineages and assumes no preferential selection of a particular lineage within the first few weeks of infection [14]. A P value less than 0.05 implies a significant deviation from the model. Prior to analysis, the alignments were checked for recombination using Simplot and the Recombination Detection Program (RDP v4.94).

We first tested all the specimens by the Poisson-Fitter tool, and retested those that did not fit the model by SPMM. Interval estimates of time since infection (i.e. 95% confidence intervals) generated with Poisson-Fitter or SPMM were then integrated with each participant study medication diaries to calculate the number of TDF pills missed within this period of time. We used Poisson point estimates to define the day of infection, and counted the number of TDF doses missed within the 95% confidence interval for this value (see Fig. 1 for a schematic representation). To minimize potential bias associated with self-reported adherence, we limited our analysis to the 11 participants that were under daily DOT during the 3 months prior to HIV-1 diagnoses.

Analysis of the impact of daily tenofovir disoproxil fumarate on the length of the virus eclipse phase in a macaque model

To understand if drug pressure because of daily TDF might increase the length of the virus eclipse phase and alter Poisson-Fitter estimates, we examined data obtained from macaques infected with simian HIV (SHIV) while receiving placebo or daily PrEP with TDF or FTC/TDF. We monitored time to first RNA detection in plasma after a single exposure to a high (1000 $TCID_{50}$) dose of SHIV162p3. These macaques were part of studies assessing how high virus doses affect PrEP efficacy, and received placebo or daily PrEP with TDF [26] or FTC/TDF.

Phylogenetic analysis

Sequences were aligned and edited using Geneious software (version 10.0.9). Nucleotide alignments were prepared using ClustalW; all insertion and deletions were removed from the alignments prior to phylogenetic analysis. Estimates of virus diversity were done by pairwise analysis of individual sequences using the maximum composite likelihood method in MEGA (version 7.0.18), and the results are expressed as number of base substitutions per site. Patterns of *env* diversity in each infection were visualized using the highlighter tool (available at https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html).

Statistical analysis

Medians and interquartile ranges (IQRs) were calculated for virus diversity and plasma HIV-1 RNA concentrations by group (TDF and placebo infections). Statistical differences were evaluated for virus diversity and plasma HIV-1 RNA using the Wilcoxon rank-sum test to compare the TDF and placebo group. The proportion of infections initiated by a single genetic variant in the TDF and placebo groups was compared using the Fisher's exact test. Analyses were completed using SAS version 9.4 (SAS Institute Inc., Cary, North Carolina, USA).

Results

HIV-1 *env* diversity in participants taking daily tenofovir disoproxil fumarate preexposure prophylaxis

We attempted to characterize HIV-1 genetic diversity in the 17 participants from BTS that had incident infection while taking TDF PrEP. HIV diversity was investigated at the visit when HIV-1 RNA was first detected in plasma. At this visit, nine participants had detectable TFV in plasma, seven had TFV levels below the limit of quantification, and one (TDF4) did not have sufficient plasma for testing (Table 1). Median plasma HIV-1 RNA concentration was 4.2 [interquartile range (IQR)=1.6] $\log_{10}$ RNA copies/ml. HIV-1 *env* was successfully amplified and sequenced in 15 of the 17 incident infections. For each participant, we generated 10–25 *env* sequences by single genome amplification. The median *env* diversity was 0.00126 (IQR = 0.00124) base substitutions per site or 1.26 mutations in 1000 bp sequence. Of the 15 infections with available sequences, 13 (86.7%) were initiated by a single genetic variant. Supplementary Fig1, http://links.lww.com/QAD/B546 shows the highlighter plots generated in the 15 infections from the TDF group and illustrates how most of these infections were highly homogeneous. Participants TDF4 and TDF8 had multivariant infection (i.e. P value <0.1 denoting divergence from a Poisson distribution). A close evaluation of trial visits in these two participants showed that they missed their regularly scheduled visit during the 5 and 3 years preceding the specimen collection.

Comparison of virus diversity between infections in the tenofovir disoproxil fumarate and placebo groups

We next compared virus diversity and frequency of single variant infections among PrEP breakthrough infections and infections in the placebo group ($n = 16$). Eleven of the 16 placebo infections (68.7%) were initiated by a single genetic variant ($P = 0.22$ compared with the TDF group) (Table 1 and Supplementary Fig. 2, http://links.lww.com/QAD/B546). Median *env* diversity in the placebo group [0.00151 (IQR = 0.00715) base substitutions per site] was similar to that seen in PrEP breakthrough infections [0.001265 (IQR = 0.00124), $P = 0.71$] (Fig. 2). A sub-analysis of *env* diversity limited to participants with single variant infections also showed no differences of virus diversity between the TDF and placebo groups (medians of 0.00122 (IQR = 0.000959) and 0.000571 (0.00145) base substitutions per site, respectively, $P = 0.56$].

Estimates of date of infection in participants assigned to tenofovir disoproxil fumarate or placebo

We next calculated the approximate date of infection using the Poisson-Fitter and SPMM tools. We generated reliable estimates of infection duration in all the 13 single variant infections from the TDF group, and in 11 of the 16 placebo participants including two (PL3 and PL10) that were infected with multiple variants (Table 1). In the remaining participants, estimates of time since infection were unreliable as the data did not fit the underlying Poisson distribution models as indicated by a low (<0.1 for Poisson-Fitter and <0.05 for SPMM) goodness-of-fit P value (Table 1). The median time since infection in single variant infections from the TDF group was 47 days (IQR= 35) and was not statistically different from that seen in single-variant placebo infections [18 days (IQR= 44)] ($P = 0.40$). Median time since infection in the four participants that were NAT+ but seronegative was 10.5 days (IQR= 6) compared with 46.5 days (IQR= 37) among the 19 participants that were NAT+ and EIA+ ($P < 0.01$).

Analysis of preexposure prophylaxis adherence diaries at the estimated date of infection

We then used PrEP adherence diaries to count the number of TDF doses taken within the 95% confidence interval of the estimated date of infection. Figure 3 shows the adherence diaries for each participant with incident infection as well as the number of TDF doses taken within this period of time. This analysis was limited to the 11 participants who took TDF under daily DOT to minimize potential bias from self-reported adherence [27]. In five of the 11 participants (TDF2, TDF3, TDF6, TDF9, and TDF15), the entire interval estimates of infection fell within periods of 100% adherence according to medication diaries. Participants TDF2, TDF6 and TDF15 were 100% adherent during the 3 months prior to the first RNA-positive visit. Participant TDF3 missed eight TDF doses more than 8 weeks prior to the first RNA-positive visit but was fully adherent during the interval estimate of infection. Participant TDF9 was reported 100% adherent during the estimated date of infection although adherence diaries document several weeks of nonadherence immediately before infection (Fig. 3). Adherence in two other incident infections (TDF11 and TDF12) was also very high according to medication diaries with only a few missed doses recorded immediately before or after the estimated day of infection. Participant TDF11 missed two doses of TDF 4–5 days prior to the estimated day of infection, whereas participant TDF12 missed five TDF doses 4–5 days before or after the estimated day of infection (Fig. 3). In the remaining four participants (TDF1, TDF5, TDF13, and TDF14), infection occurred during periods of little or no adherence to daily TDF. In two participants (TDF1 and TDF12), TFV was undetectable in plasma at the first NAT+ visit despite reported adherence by DOT.

Impact of daily tenofovir disoproxil fumarate on the length of the virus eclipse phase as defined in a macaque model of infection with simian HIV

We investigated in macaques infected with SHIV if drug pressure because of daily TDF might increase the length of the virus eclipse phase and alter Poisson-Fitter estimates. Figure 4 shows that time to first RNA detection was 7 days in untreated macaques, and that daily TDF increased the length of the virus eclipse phase to 7– 14 days (mean = 11.7; *P* = 0.038). Although significant, the effect was smaller than that observed with daily FTC/ TDF combination (21 days; *P* = 0.015).

Discussion

We characterized early HIV-1 *env* sequences in participants with incident HIV-1 infection in a human PrEP trial with daily TDF in order to estimate the date of infection using well established mathematical models of early virus evolution. We sought to relate HIV infections with PrEP pill-taking behavior at the time when infection was most likely to have occurred. Our analysis found that five of the 11 participants tested had an estimated date of infection that fell within periods of 100% adherence as defined by adherence diaries obtained under daily DOT, though interruptions in PrEP adherence were also noted in two of these participants prior to the estimated time of infection. These cases suggest that daily oral TDF at the doses given may not provide complete protection from HIV infection through intravenous exposure. However, it is clear that oral TDF does provide protection given that in a subanalysis of the BTS data including all participants (i.e. those on and not on DOT), high adherence levels were associated with a 83.5% reduction in the risk of infection [28]. Alternative interpretations of our data can be made. Data in the adherence diaries under DOT may not have accurately reflected drug exposure in some cases, either because of data entry errors or other factors. For instance, we noted two participants (TDF1 and TDF12) with undetectable TFV in plasma at a visit that had reported adherence by DOT. In addition, substantial errors in our estimates of the date of infection may have obscured true infection times during periods of poor regimen adherence. Such errors may be more pronounced among those participants that had a limited number of sequences generated by SGA.

We found that TFV did not have a significant effect on HIV *env* diversity at the first NAT+ sample, as the overall diversity in PrEP breakthrough infections was comparable to that seen in participants receiving placebo. This observation suggests that drug selective pressure in the reverse transcriptase originating from exposure to daily TDF during undiagnosed infection may not significantly affect HIV-1 *env* evolution, though additional studies with larger numbers of infections may be needed. Unfortunately, we did not have longitudinal specimens from macaques infected with SHIV during daily TDF to support this observation. However, estimates of infection duration in BTS participants as defined by *env* diversity correlated well with Fiebig staging as also found in other studies [13,14,19,20]; these estimates were 11 days in participants in Fiebig stages I/II (NAT+, antibody negative) compared with 47 days in those that were both nucleic acid testing (NAT) and antibody positive.

Our molecular approach to estimate time since infection has the potential to be applied to other PrEP modalities including topical products, implants and long-acting injectable formulations, where the availability of infection estimates may be particularly useful for defining drug concentrations associated with breakthrough infection. However, such analysis will require careful validation of Poisson-Fitter models and considerations on how more potent regimens can impact HIV-1 evolution and length of the virus eclipse phase. We noted that daily FTC/TDF in macaques could lengthen the virus eclipse phase from 7 to 21 days. In an earlier study, continued exposure of macaques to FTC/DTF was also associated with a reduction in *env* diversification [29]. Thus, estimations of time since infection with more potent drugs/drug combinations or novel PrEP modalities may require prior validation of Poisson models using for instance animal models of PrEP where exposure is done under controlled conditions and specimens collected frequently enough to precisely define acute infection dynamics. In the context of objective adherence measures, the Poisson tools may potentially help evaluate biological efficacy in real time, and inform decision making on whether a PrEP trial should continue or not.

Our study has several limitations. First, despite regular monitoring, it is possible that errors occurred during DOT study drug administration, data collection, and data entry. This might for instance explain the lack of detectable TFV in plasma at the first NAT-positive sample in participants TDF1 and TDF12 who were reported as adherent in that visit, although testing error or incomplete drug absorption/metabolism might also explain the findings. Second, it is possible that the administration of antiretroviral therapy lengthened the viral eclipse phase, affecting the accuracy of the estimated date of infection. Although not conclusive, our results in rhesus macaques infected rectally with simian HIV demonstrated only modest eclipse phase changes on the order of 4 days, suggesting that this effect may be minimal. However, this modest effect might be sufficient to explain infection in participant TDF9 that was reported as 100% adherent at the estimated date of infection but was nonadherent for several weeks immediately before. If this is the case, this participant would be erroneously reported as 100% adherent in our analysis.

In conclusion, we applied molecular tools to estimate the day of infection among participants from a human PrEP trial with daily TDF, and mapped these estimates with pill-taking behaviors. Using this approach, we found that breakthrough infections may occur among potentially highly adherent PWID taking daily TDF PrEP. This finding is consistent with the overall biological efficacy of daily TDF against parenteral HIV-1 transmission being less than 100%, although we cannot exclude that some of the infections might have occurred through the sexual route. As new methods to measure daily adherence are developed and new PrEP regimens are evaluated in humans, our molecular approach provides a new tool that can help to better define the biological efficacy of PrEP.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Bill Switzer for helping with the analysis of recombination, and Chuong T. Dinh for his assistance in measuring TFV concentrations in some of the plasma specimens.

Funding:

CDC intramural funds.

References

1. HIV/AIDS fact sheet World Health organization [updated 2017]. Available at: http://www.who.int/mediacentre/factsheets/fs360/en/. [Accessed 10 July 2019]
2. Riddell J, Amico KR, Mayer KH. HIV preexposure prophylaxis: a review. JAMA 2018; 319:1261–1268. [PubMed: 29584848]
3. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. , the iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010; 363:2587–2599. [PubMed: 21091279]
4. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. , the Partners PrEP Study Team. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med 2012; 367:399–410. [PubMed: 22784037]
5. Thigpen MC, Kebaabetswe PM, Paxton LA, Smith DK, Rose CE, Segolodi TM, et al. , TDF2 Study Group. Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. N Engl J Med 2012; 367:423–434. [PubMed: 22784038]
6. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, et al. , Bangkok Tenofovir Study Group. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet 2013; 381:2083–2090. [PubMed: 23769234]
7. Marrazzo J, Ramjee G, Nair G, Palanee T, Mkhize B, Nakabiito C, et al. for the VOICE study team. Preexposure prophylaxis for HIV in women: daily oral tenofovir, oral tenofovir/emtricitabine, or vaginal tenofovir gel in the VOICE study (MTN 003). Conference on Retroviruses and Opportunistic Infections. Atlanta; 2013.
8. Van Damme L, Corneli A, Ahmed K, Agot K, Lombaard J, Kapiga S, et al. , FEM-PrEP Study Group. Preexposure prophylaxis for HIV infection among African women. N Engl J Med 2012; 367:411–422. [PubMed: 22784040]
9. Anderson PL, Glidden DV, Liu A, Buchbinder S, Lama JR, Guanira JV, et al. , iPrEx Study Team. Emtricitabine-tenofovir concentrations and preexposure prophylaxis efficacy in men who have sex with men. Sci Transl Med 2012; 4:151ra125.

10. Castillo-Mancilla JR, Zheng JH, Rower JE, Meditz A, Gardner EM, Predhomme J, et al. Tenofovir, emtricitabine, and tenofovir diphosphate in dried blood spots for determining recent and cumulative drug exposure. AIDS Res Hum Retroviruses 2013; 29:384–390. [PubMed: 22935078]

11. Anderson PL, Liu AY, Castillo-Mancilla JR, Gardner EM, Seifert SM, McHugh C, et al. Intracellular tenofovir-diphosphate and emtricitabine-triphosphate in dried blood spots following directly observed therapy. Antimicrob Agents Chemother 2017; 62:e01710–e01717.

12. Dimitrov DT, Maˆsse BR, Donnell D. PrEP adherence patterns strongly affect individual HIV risk and observed efficacy in randomized clinical trials. J Acquir Immune Defic Syndr 2016; 72:444–451. [PubMed: 26990823]

13. Giorgi EE, Funkhouser B, Athreya G, Perelson AS, Korber BT, Bhattacharya T. Estimating time since infection in early homogeneous HIV samples using a poisson model. BMC Bioinformatics 2010; 11:532. [PubMed: 20973976]

14. Love TM, Park SY, Giorgi EE, Mack WJ, Perelson AS, Lee HY. SPMM: estimating infection duration of multivariant HIV-1 infections. Bioinformatics 2016; 32:1308–1315. [PubMed: 26722117]

15. Ribeiro RM, Qin L, Chavez LL, Li D, Self SG, Perelson AS. Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. J Virol 2010; 84:6096–6102. [PubMed: 20357090]

16. Markowitz M, Louie M, Hurley A, Sun E, Di Mascio M, Perelson AS, Ho DD. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. J Virol 2003; 77:5037–5038. [PubMed: 12663814]

17. Mansky LM, Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J Virol 1995; 69:5087–5094. [PubMed: 7541846]

18. Stafford MA, Corey L, Cao Y, Daar ES, Ho DD, Perelson AS. Modeling plasma virus concentration during primary HIV infection. J Theor Biol 2000; 203:285–301. [PubMed: 10716909]

19. Lee HY, Giorgi EE, Keele BF, Gaschen B, Athreya GS, Salazar-Gonzalez JF, et al. Modeling sequence evolution in acute HIV-1 infection. J Theor Biol 2009; 261:341–360. [PubMed: 19660475]

20. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A 2008; 105:7552–7557. [PubMed: 18490657]

21. Suntharasamai P, Martin M, Choopanya K, Vanichseni S, Sangkum U, Tararut P, et al. Assessment of oral fluid HIV test performance in an HIV pre-exposure prophylaxis trial in Bangkok, Thailand. PLoS One 2015; 10:e0145859.

22. Kuklenyik Z, Martin A, Pau CP, Garcia-Lerma JG, Heneine W, Pirkle JL, Barr JR. Effect of mobile phase pH and organic content on LC-MS analysis of nucleoside and nucleotide HIV reverse transcriptase inhibitors. J Chromatogr Sci 2009; 47:365–372. [PubMed: 19476704]

23. Boonchawalit S, Jullaksorn D, Uttiyoung J, Yowang A, Krathong N, Chautrakul S, et al. Molecular evolution of HIV-1 CRF01_AE Env in Thai patients. PLoS One 2011; 6:e27098.

24. Jesus Salazar-Gonzalez Kimmy Pham, Keele Brandon, David McPherson. Standard operating procedure for: single genome amplification of HIV-1 envelope. In. Edited by (CHAVI) CHAVI: Molecular Biology and Sequencin Core: University of Alabama at Birmingham; 2007.

25. Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, Keele BF, et al. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. J Virol 2008; 82:3952–3970. [PubMed: 18256145]

26. Subbarao S, Ramos A, Kim C, Adams D, Monsour M, Butera S, et al. Direct stringency comparison of two macaque models (single-high vs. repeat-low) for mucosal HIV transmission using an identical anti-HIV chemoprophylaxis intervention. J Med Primatol 2007; 36:238–243. [PubMed: 17669212]

27. Amico KR, Marcus JL, McMahan V, Liu A, Koester KA, Goicochea P, et al. Study product adherence measurement in the iPrEx placebo-controlled trial: concordance with drug detection. J Acquir Immune Defic Syndr 2014; 66:530–537. [PubMed: 24853306]

28. Martin M, Vanichseni S, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, et al. , Bangkok Tenofovir Study Group. The impact of adherence to preexposure prophylaxis on the risk of HIV infection among people who inject drugs. AIDS 2015; 29:819–824. [PubMed: 25985403]
29. Zheng Q, Ruone S, Switzer WM, Heneine W, Garcia-Lerma JG. Limited SHIV env diversification in macaques failing oral antiretroviral preexposure prophylaxis. Retrovirology 2012; 9:40. [PubMed: 22571771]

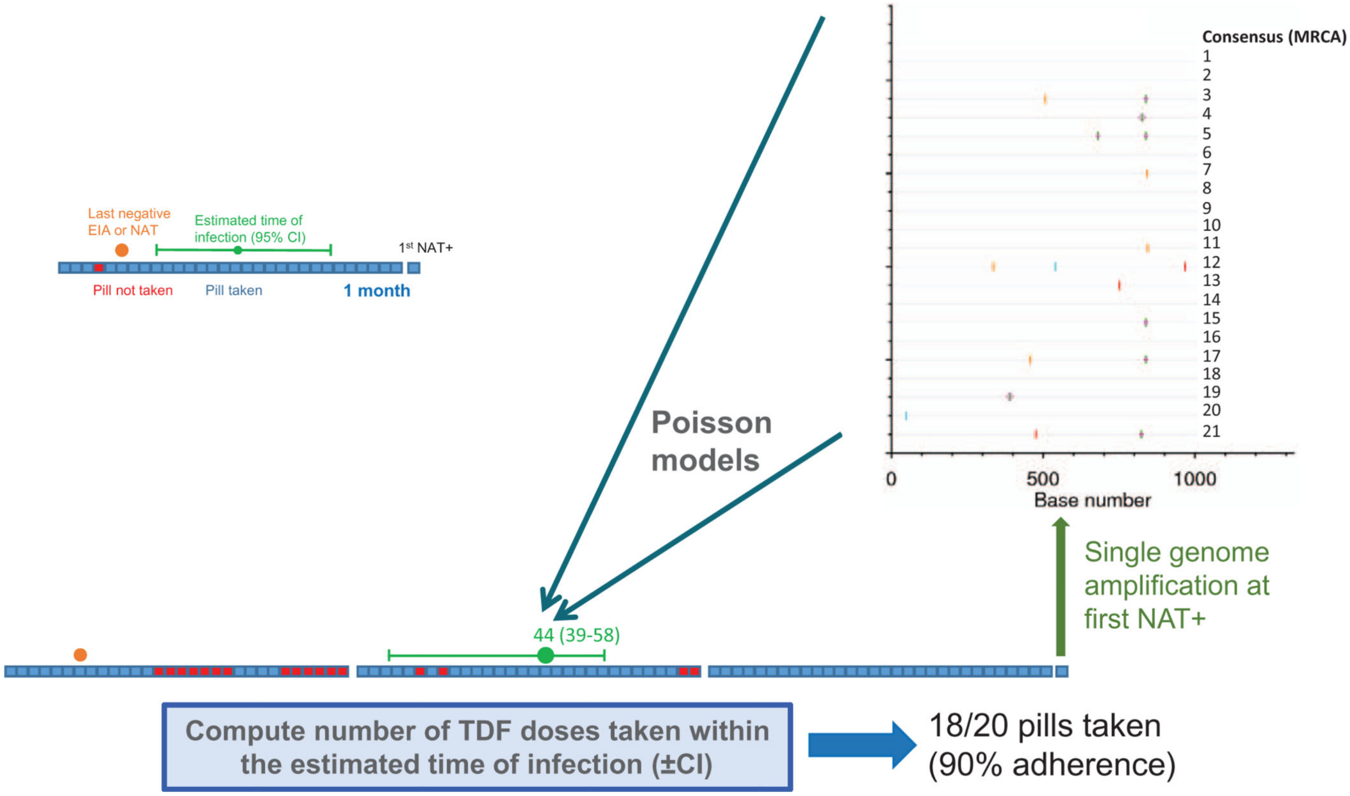

Fig. 1. Computation of number of tenofovir disoproxil fumarate doses taken within the estimated date of infection.

The diversity structure of the virus population (highlighter plot) defined at the first nucleic acid testing positive (NAT+) sample using single genome amplification and sequence analysis is used to estimate date of infection using the Poisson Fitter or Shifted Poisson Mixture Model (SPMM). Estimates of date of infection are then integrated with adherence maps generated under daily directly observed therapy to compute the number of TDF doses taken within this period of time. One square represents 1 day (red = missed TDF dose; blue = TDF dose taken). Orange circle denotes the last negative EIA or NAT specimen. The estimated time of infection (and 95% CI) is indicated with a green line. The highlighter plot illustrates the changes observed in each individual sequence relative to the consensus (A: green, T: red, G: orange, C: light blue, IUPAC: dark blue, gaps: gray, circle: APOBEC signatures, diamond: G to A conversions). EIA, enzyme immunoassay; NAT, nucleic acid testing; TDF, tenofovir disoproxil fumarate.

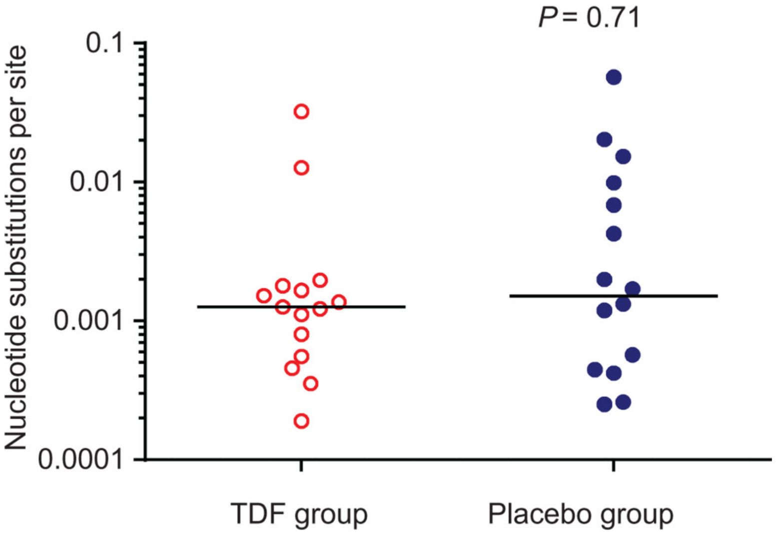

Fig. 2. HIV *env* diversity in incident infections from the Bangkok Tenofovir Study (BTS) with daily tenofovir disoproxil fumarate.

Diversity was measured when HIV RNA was detected in plasma for the first time. This visit coincided with the seroconversion visit in 14 of the 17 infections in the TDF group, and in 15 of the 16 infections in the placebo group. Estimates of diversity were determined by measuring mean pairwise distances between individual sequences from each participant. Results are expressed as number of nucleotide substitutions per site. Horizontal lines denote median values for each group.

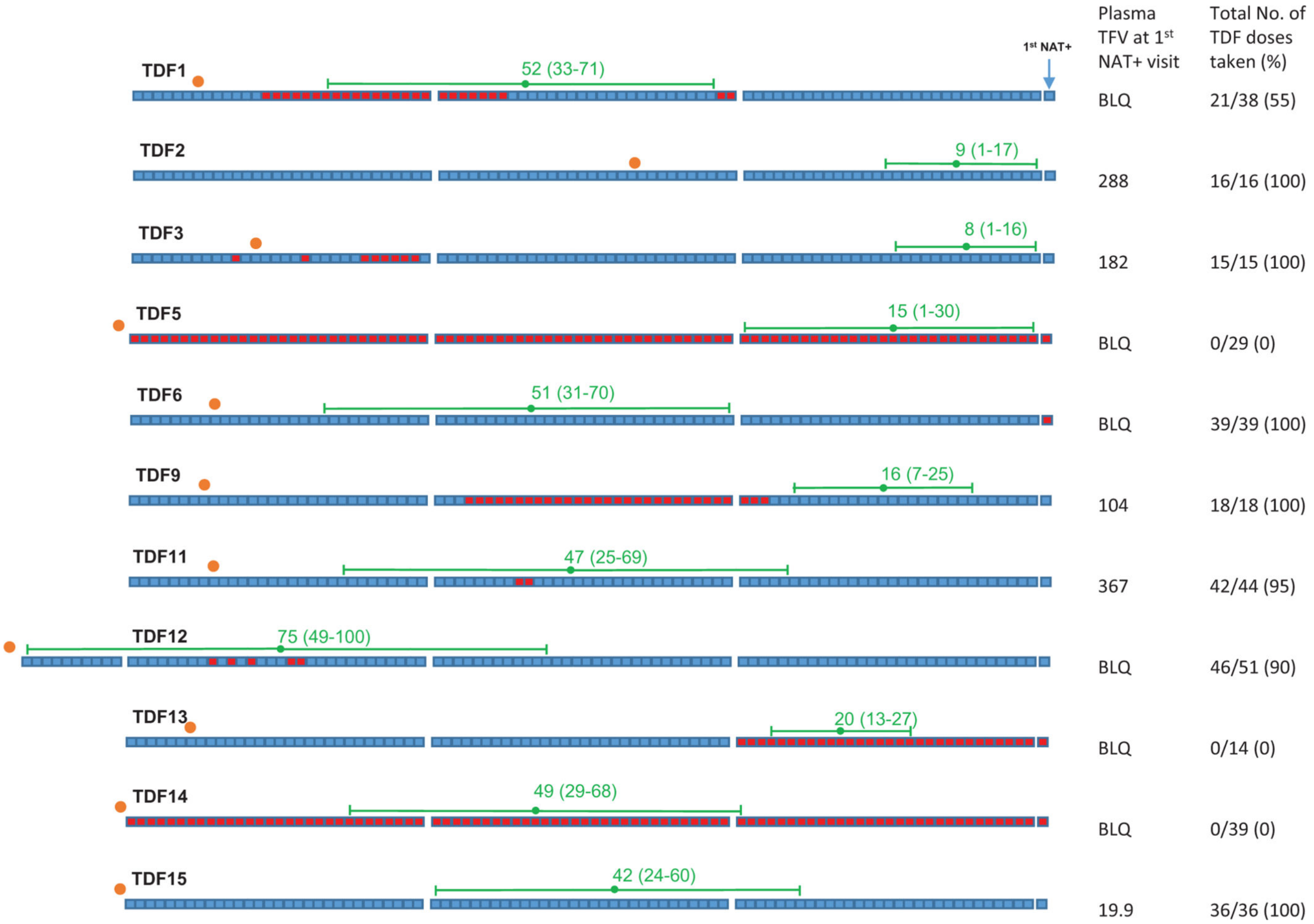

Fig. 3. Analysis of pill taking behavior at the estimated time of infection.

The estimated infection time (and 95% CI) is indicated with a green line. Adherence diaries obtained under daily directly observed therapy (horizontal lines) were used to manually count the number of missed TDF doses within this period of time. One square represents 1 day (red = missed TDF dose; blue = TDF dose taken). Each block represents 1 month. Orange circle denotes the last negative EIA or NAT specimen. CI, confidence interval; TDF, tenofovir disoproxil fumarate.

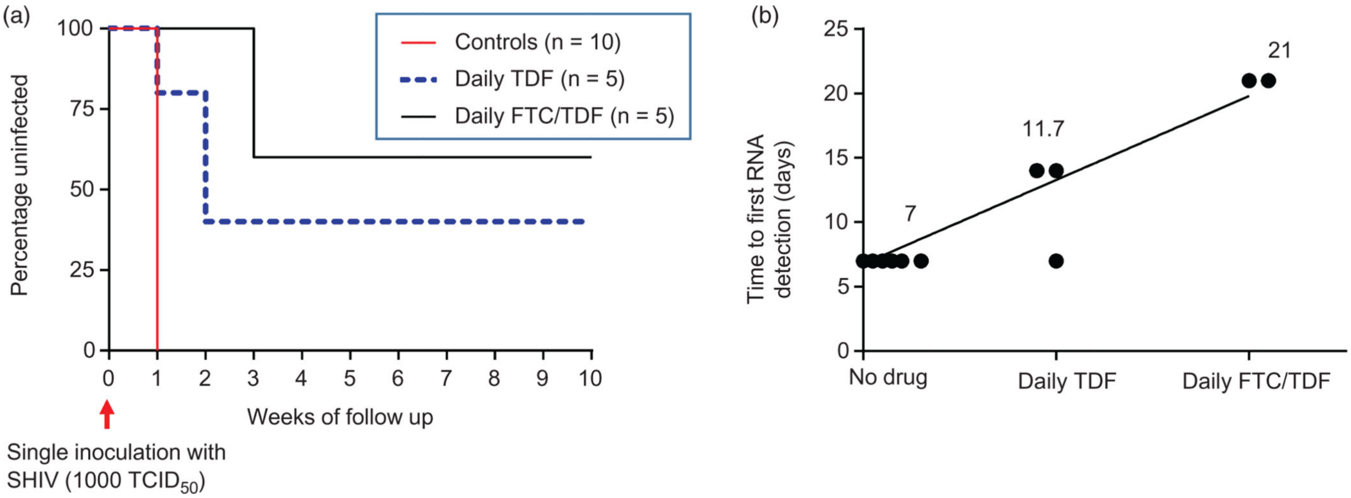

Fig. 4. Effect of preexposure prophylaxis with tenofovir disoproxil fumarate or emtricitabine/tenofovir disoproxil fumarate on the length of the virus eclipse phase in macaques infected with simian HIV.

These macaques were part of studies assessing how high virus doses affect PrEP efficacy, and received placebo or daily PrEP with TDF [26] or FTC/TDF. Animals were exposed to a single high (1000 TCID50) dose of a SHIV162p3 isolate containing the HIV-1 envelope, and received clinically equivalent doses of TDF (22 mg/kg) or a combination of TDF (22 mg/kg) and FTC (20 mg/kg). Treatment was initiated 7 days prior to SHIV exposure and stopped 28 days after virus inoculation. (a) Time for SHIV RNA detection in plasma. (b) Impact of TDF or FTC/TDF on the length of the virus eclipse phase as defined by RNA detection in plasma. Time of RNA detection increased from 7 days in untreated animals (*n* = 10), to 11.7 in animals receiving daily TDF (*n* = 3) and 21 days in animals receiving daily FTC/TDF ([26] and unpublished observation). FTC, emtricitabine; PrEP, preexposure prophylaxis; SHIV, simian HIV; TDF, tenofovir disoproxil fumarate.

Table 1.

Characteristics of the study population and estimates of time since infection in 33 participants from the Bangkok Tenofovir Study trial with daily tenofovir disoproxil fumarate who had incident HIV-1 infection.

Group	Sex	Visit	Plasma HIV RNA (copies/ml)	Plasma TFV (ng/ml)	Number of *env* sequences	Number of base substitutions per site	Number of founder viruses	Time since infection (95% CI)	Goodness of fit *P* value[a,b]
TDF group									
TDF1	M	NAT FP/EIA+	434	BLQ	21	0.0016570	1	52 (33–71)	0.89
TDF2	M	NAT FP	675 000	288	22	0.0001905	1	9 (1–17)	0.70
TDF3	M	NAT FP	36 100	182	22	0.0004545	1	8 (1–16)	0.70
TDF4	M	NAT FP/EIA+	9671	–	20	0.0323300	18	31 (15–49)	0.00
TDF5	M	NAT FP	536 799	BLQ	18	0.0005556	1	15 (1–30)	0.42
TDF6	M	NAT FP/EIA+	61 397	BLQ	22	0.0012170	1	51 (31–70)	0.77
TDF7	M	NAT FP/EIA+	1930	60	18	0.0013660	1	50 (32–68)	0.82
TDF8	M	NAT FP/EIA+	131 689	BLQ	18	0.0126900	2	382 (323–441)	0.00
TDF9	M	NAT FP/EIA+	5530	104	17	0.0003529	1	16 (7–25)	0.65
TDF10	M	NAT FP/EIA+	520 042	136	25	0.0019670	1	85 (61–109)	0.26
TDF11	F	NAT FP/EIA+	2340	367	17	0.0015150	1	47 (25–69)	0.33
TDF12	M	NAT FP/EIA+	21 400	BLQ	23	0.0017940	1	75 (49–100)	0.48
TDF13	M	NAT FP/EIA+	15 612	BLQ	22	0.0008009	1	20 (13–27)	0.53
TDF14	M	NAT FP/EIA+	20 194	BLQ	17	0.0012650	1	49 (29–68)	0.68
TDF15	M	NAT FP/EIA+	1276	19.9	10	0.0011110	1	42 (24–60)	0.32
TDF16	M	NAT FP/EIA+	771	86	–	–	–	–	–
TDF17	F	NAT FP/EIA+	315	56	–	–	–	–	–
Placebo group									
PL1	F	NAT FP/EIA+	3635	–	17	0.0019850	1	84 (58–109)	0.43
PL2	F	NAT FP/EIA+	425 225	–	21	0.0202400	6	112 (82–142)	0.02
PL3	F	NAT FP/EIA+	15 010	–	18	0.0152300	5	74 (48–101)	0.05
PL4	F	NAT FP/EIA+	95 892	–	20	0.0013210	1	46 (34–59)	0.55
PL5	M	NAT FP/EIA+	374 746	–	18	0.0068250	1	295 (243–348)	<0.01
PL6	M	NAT FP/EIA+	294 000	–	20	0.0011890	1	18 (9–27)	0.60
PL7	M	NAT FP	1 950 000	–	23	0.0002609	1	12 (4–20)	0.67

Group	Sex	Visit	Plasma HIV RNA (copies/ml)	Plasma TFV (ng/ml)	Number of *env* sequences	Number of base substitutions per site	Number of founder viruses	Time since infection (95% CI)	Goodness of fit *P* value[a,b]
PL8	M	NAT FP/EIA+	35 752	BLQ	16	0.0568900	7	197 (150–244)	0.00
PL9	M	NAT FP/EIA+	9730	–	13	0	1	nc	–
PL10	M	NAT FP/EIA+	23 891	BLQ	14	0.0042530	3	54 (29–80)	0.82
PL11	M	NAT FP/EIA+	324 797	BLQ	17	0.0017060	1	65 (48–82)	0.18
PL12	M	NAT FP/EIA+	3 030 000	–	16	0.0002500	1	11 (2–20)	0.68
PL13	M	NAT FP/EIA+	180 000	–	19	0.0004211	1	5 (−2–12)	0.75
PL14	M	NAT FP/EIA+	284 920	BLQ	18	0.0004444	1	16 (6–25)	0.65
PL15	M	NAT FP/EIA+	687 126	BLQ	22	0.0005714	1	22 (7–37)	0.13
PL16	M	NAT FP/EIA+	563 000	–	17	0.0098670	4	45 (23–67)	0.00

BLQ, below the limit of quantification; EIA+, enzyme immunoassay positive; NAT FP, nucleic acid testing first positive; nc, not computed; SPMM, Shifted Poisson Mixture Model; TDF, tenofovir disoproxil fumarate.

[a] A small (<0.1 for Poisson-Fitter and <0.05 for SPMM) *P* value denotes divergence from a Poisson distribution and a bad fit to the model.

[b] Reported values for samples TDF4, TDF8, TDF10, PL2, PL3, PL5, PL8, and PL16 were generated by SPMM.