**Supplementary information**

**Methods**

Viral RNA was extracted from plasma samples and single genome sequencing was carried out for both *integrase* and *envelope* region.The entire integrase region (HXB2 nt 4168 to 5190) was amplified using protocol by Lataillade et al., 2007 [4] and the envelope region (HXB2 positions: 5957 to 8904) described by Rothenberger et al., 2015 ([5] Supplementary Table 1). An Illumina based next generation sequencing procedure with primer-ID was utilized to obtain large numbers of sequences from HIV RT region (HXB2 nt positions: 2704-2943 and 3046-3253) with modified primer for reverse-transcriptase reaction (Supplementary Table 1, [6]).

To amplify the integrase region, reverse transcription was carried out in 100 µl reaction using Superscript III (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), 0.1 µM reverse outer primer (Supplementary Table 1) and 40 µl (10,000 copies) resuspended viral RNA following manufacturer’s guidelines. Reverse transcription was carried out incubating the reaction mix for 1 hour at 45 ֯C. Nested-PCR was performed in 10 µl using 0.2 mM dNTPs, 2.5 mM MgCl2, 0.4 µM primers (Supplementary Table 1), Platinum 1x PCR buffer and 0.8 unit Platinum Taq (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) as final concentrations.

For envelope region, 20,000 and 3,000 HIV-1 copies from prior to DTG initiation (19th of August 2014) and after DTG failure (4th of June 2015), respectively, was used as a template for reverse transcription in 60 ul of reaction using Superscript III (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), 0.5 mM dNTPs and 0.25 µM reverse outer primer (Supplementary Table 1) following manufacturer’s guidelines. Reverse transcription was carried out incubating the reaction mix for 1 hour at 50 ֯C, 1 hour at 55 ֯C and 15 min at 70 ֯C, followed by RNaseH treatment using 1 µl of RNaseH per sample and incubation for 20 min at 37 ֯C. Nested-PCR was performed in 10 µl using 0.2 mM dNTPs, 2 mM MgCl2, 0.2 µM primers (Supplementary Table 1), Platinum 1x PCR buffer and 0.5 unit Platinum Taq as final concentrations. For both regions, the second round PCR used 2 µl of diluted first round product. PCR programs are outlined in Supplementary Table 2.

**Data analysis**

Single genome sequences of 1024 bp and 2928 bp in HIV-1 integrase and *env* region, respectively, were aligned and subjected to neighbor-joining phylogenetic analysis, and topologies were investigated with bootstrap analysis (MEGA), HIV-1 INSTI drug resistance mutations determination was carried out by Stanford University HIV Drug Resistance Database (<https://hivdb.stanford.edu/>). To investigate the presence of population shift during the course of the evaluation period, a geographic subdivision test for evidence of panmixia was used as previously described [4] with probability of panmixia of < 1x10-9 considered to be evidence of population shift. Replicating population sizes were estimated from temporal changes in allele frequencies [5] and linkage disequilibrium analysis was performed using DNA Sequence Polymorphism software version 6 [6]. Recombination analyses were performed using the SIMPLOT Package (<https://sray.med.som.jhmi.edu/scroftware/simplot/>). All sequences will be submitted to GENBANK on publication

**Supplementary Table 1.** The list of primers used to amplify and sequence HIV-1 integrase, envelope and reverse transcriptase regions.

|  |  |  |  |
| --- | --- | --- | --- |
| Primer name | Primer’s designation | Primer sequence 5’ → 3’ | Reference |
| **Integrase region** | | | |
| INREV-I | Reverse transcription and I round PCR | TCTCCTGTATGCAGACCCCAATAT | 1 |
| INFORI | I round PCR | GGAATCATTCAAGCACAACCAGA | 1 |
| HIV+4141 | II round PCR and sequencing | TCTACCTGGCATGGGTACCA |  |
| INREVII | II round PCR and sequencing | CCTAGTGGGATGTGTACTTCTGA |  |
| POLK- | Sequencing | CCT TGA CTT TGG GGA TTG TAG GGA A |  |
| POLZ | Sequencing | AAT TTT CGG GTT TAT TAC AG |  |
| POLJ | Sequencing | GAAGCCATGCATGGACAAGTAGA |  |
| **Envelope region** | | | |
| envB5out | I round PCR | TAGAGCCCTGGAAGCATCCAGGAAG | 2 |
| envB3out | I round PCR | TTGCTACTTGTGATTGCTCCATGT | 2 |
| envB5in | II round PCR and sequencing | TTAGGCATCTCCTATGGCAGGAAGAAG | 2 |
| envB3in | II round PCR and sequencing | GTCTCGAGATACTGCTCCCACCC | 2 |
| For14 | Sequencing | TATGGGACCAAAGCCTAAAGCCATGTG |  |
| For16 | Sequencing | TTTAATTGTGGAGGAGAATTTTTCTA |  |
| For18 | Sequencing | CATATCAAATTGGCTGTGGTATAT |  |
| Rev15 | Sequencing | CTGCCATTTAACAGCAGTTGAGTTGA |  |
| Rev16 | Sequencing | ATGGGAGGGGCATACATTGCT |  |
| Rev18 | Sequencing | GGTGAGTATCCCTGCCTAACTCTAT |  |
| **uSGS assay of reverse transcriptase region** | | | |
| DRPTEAMG\_2834\_Rmoved | Reverse  Transcription USGS assay | GGTATCGAAGTCATCCTGCTAGNNNNNNNNNNTAACCCATCCAAAGGAATGGAG | This paper |
| PCR1 28-2195F-dU | I round PCR primer | AAACAAUGGCCAUTGACAGAAGA | 3 |
| DRPTEAMG\_R-5Us | I and II round PCR | GGUAUCGAAGUCAUCCUGCTAG | 3 |
| PCR2 2286-F-dUs | II round PCR | CUGAAAAUCCAUACAAUACTCCAGTATTTGC | 3 |

**Supplementary Table 2.** The PCR programs used to amplify HIV-1 reverse transcriptase and envelope regions.

|  |  |  |
| --- | --- | --- |
| Temperature | Time | Cycles |
| Integrase region | | |
| 94 ⁰C | 2 min (5 min for II round PCR) | 1 |
| 94 ⁰C | 30 sec | 35 |
| 60 ⁰C | 1 min |
| 72 ⁰C | 2 min |
| 72 ⁰C | 7 min | 1 |
| 4 ⁰C | ∞ |  |
| Envelope region | | |
| 94 ⁰C | 2 min (5 min for II round PCR) | 1 |
| 94 ⁰C | 15 sec | 35 (45 for II round PCR) |
| 57 ⁰C | 30 min |
| 68 ⁰C | 6 min |
| 68 ⁰C | 15 min | 1 |
| 4 ⁰C | ∞ |  |

**Supplementary Figure Legend**

Detection of recombination in integrase sequence during DTG therapy.

A. Positions of 17 polymorphisms present in HIV in sequence in the rebound HIV and subsequent time points during DTG that contained T97A were identified. Based on polymorphisms, 4 distinct variants emerged at rebound viremia, and are color coded. In subsequent time points, the individual variants persisted, and evidence of recombination was present by identifying new variants containing regions with polymorphisms characteristic of two or more of the initial four rebound variants. Additional new singleton changes in the on DTG SGS are indicated in red.

B. Bootscanning was performed using SimPlot (4) identify the presence of recombination, using query sequences against parental sequences

Recombination threshold of 85 window size of 60, step size of 60, 1000 bootstrap replication. In the example , the sequence indicated in the unfilled black arrow in panel A was queried against sequences indicated with filled blac arrows in panel A. Position of two recombination breakpoints is indicated by color-coded arrows.

**References**

1. Lataillade M, Chiarella J, Kozal MJ. Natural polymorphism of the HIV-1 integrase gene and mutations associated with integrase inhibitor resistance. Antivir Ther.2007;12(4):563-70.
2. Rothenberger MK, Keele BF, Wietgrefe SW, Fletcher CV, Beilman GJ, Chipman JG, Khoruts A, Estes JD, Anderson J, Callisto SP, Schmidt TE, Thorkelson A, Reilly C, Perkey K, Reimann TG, Utay NS, Nganou Makamdop K, Stevenson M, Douek DC, Haase AT, Schacker TW. Large number of rebounding/founder HIV variants emerge from multifocal infection in lymphatic tissues after treatment interruption. Proc Natl Acad Sci U S A. 2015 Mar 10;112(10):E1126-34. doi: 10.1073/pnas.1414926112. Epub 2015 Feb 23
3. Boltz VF, Rausch J, Shao W, Hattori J, Luke B, Maldarelli F, Mellors JW, Kearney MF, Coffin JM. Ultrasensitive single-genome sequencing: accurate, targeted, next generation sequencing of HIV-1 RNA. Retrovirology. 2016 Dec 20;13(1):87. doi: 10.1186/s12977-016-0321-6.
4. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol. 1999 Jan;73(1):152-60. doi: 10.1128/JVI.73.1.152-160.1999.