Publications
Distinct Geographical Clustering of HIV-1 and a Signature Amino Acid at Position 41 of the p24 Unveiled by gag Variability in India

Arpan Acharya1, Salil Vaniawala2, Harsh Parekh3, Anju Nagee3, Rabindra N. Misra4, Minal Wani1 and Pratap N. Mukhopadhyaya*5

1Dr. D.Y. Patil Biotechnology & Bioinformatics Institute, Tathawade, Pune, India; 2SN Genelab, 2nd Floor, President Plaza, Tower A, Nanpura, Ring Road, Surat, Gujarat, India; 3Ashok & Rita Patel Inst. of Integrated Study & Res. in Biotech & Allied Sciences, New V.V Nagar, Gujarat, India; 4Department of Microbiology, Padmashree Dr. DY. Patil Medical College & Research Centre, Sant Tukaram Nagar, Pimpri, Pune - 411 018, Maharashtra, India; 5Interdisciplinary Science, Technology and Research Academy, AISC, Hidayatullah Road, CAMP, Pune, India

Abstract: A portion of the gag gene cDNA for p24 protein from 30 Indian HIV-1 proviral DNA was amplified by PCR and sequenced. Phylogenetic analysis with reference samples of A1, A2, B, C, D, F1, F2, G, H, J, K, N and O subtypes revealed that 29 test samples aligned with subtype C reference strain while 1 matched with HIV-1 subtype A. Multiple alignment of predicted amino acid sequence of the Indian test samples and reference C subtype of HIV-1 samples from other countries indicated a molecular signature by way of rigid conservation of the amino acid ‘S’ at position 41 of the gag p24 protein in all Indian HIV-1 samples analyzed in this study as opposed to ‘T’in the same position in C subtype sequences from other parts of the world. A phylogenetic analysis and visualization of the resulting tree in radial position showed distinct clubbing of all Indian C subtypes and formation of a cluster when compared to C subtype sequences from other countries with a single Chinese sample as an exception which was found in the Indian cluster. The use of a portion of p24 gene sequence as tool for subtyping as well as phylogenetic grouping with special reference to its geographical location is discussed.

Keywords: HIV-1, gag, p24, phylogeny, subtype.

INTRODUCTION

In the year 1986, India reported its first case of HIV-1 infection from the state of Tamil Nadu in southern part of the country [1]. Since then, the spread of this virus has been phenomenal and as per UNAIDS report (2010), with a population of around 1.10 billion, India is staring at almost 2.4 million HIV-1 infected people to care for. This makes India one of the largest HIV-1 infected population bearing countries in the world with most infection routes being acquired through unprotected sex although certain parts of the country, especially those in the north-east record transmission through use of contaminated needles also (UNGASS country progress report, 2007).

Division of HIV-1 into subgroups is primarily based on phylogenetic analysis of the envelope (env) and group-specific antigen (gag) gene sequences. The three main groups identified are M (main), O (outlier) and N (non M, non O). The M group is further classified into subtypes which range from A-H, J and K, respectively [1-3]. The region of central Africa is dominated by subtype A & D while B is mainly found in North America, Western Europe, Australia and Japan. The subtype C is found in South Africa and India while E is present in Thailand and F in Brazil and Romania, respectively [4].

It has been observed that the subtype C sequences generated from samples from different parts of India are more closely related to each other than with those from other parts of the world [5]. International travel has increased in recent times and there is a continuous follow of HIV-1 samples in one country from another through male and female carriers [6] with India being no exception. However, despite this phenomenon the close similarity observed within HIV-1 C subtype samples circulating within the country is intriguing. This is further so since in a study it had been found that HIV-1 C subtype samples in India are closely related to those found in China, Myanmar and Taiwan [7].

Cellular toxicity [8] as a form of immune reaction has been detected against conserved p24 antigen of HIV-1 although primarily they are directed towards the hypervariable V3 domain of the env gene product gp120 [9, 10]. Further, there is distinct relation between change from asymptomatic to symptomatic stage of HIV-1 infection in a patient and parallel reduction of antibodies against p24 and p17 in the blood [11]. It was subsequently known that gag proteins play a significant role in the onset of AIDS raising the importance of the antigenic structure of p24 and p17.

In India while initial reports indicated prevalence of subtype C [5, 12, 13], a small percentage of subtype A and Thai B [14, 15], later stage studies indicated presence of multiple subtypes and recombinant samples possibly because of improved detection technologies and a wider screening spectrum [16].
The env gene had traditionally been the prime target for HIV-1 subtyping in India [5, 17, 18]. However, apart from the popular C2-V3 domain based genotyping [19-21] the critical gag p24-p7 domain has also been emerging now as an important target for diversity study of HIV-1 [21-23] particularly with the development of heteroduplex mobility assay (HMA) exploiting genetic variation within the gag region of HIV-1 genome [21, 24]. Mandal and coworkers (2002) had earlier analyzed HIV-1 C subtypes based on the C2-V3 and C2-V5 genotype in a set of female population in eastern India. However, such analysis based on the gag gene is very limited with one study by Kurle and co workers [15] on a small population from western India (Pune) and a year later by Sengupta et al. (2005) [25] on another population from eastern India (Kolkata).

With rapidly evolving genome complexity of HIV-1, we felt the need to re-look at the phylogeny of HIV-1 proviral DNA from western India based on the less focused gag gene polymorphism. In the past several groups focused on the full length gag gene or a part of it covering p24-p7 region for HIV-1 subtyping and genetic diversity analysis from India and other parts of the world [25, 26]. In contrast, in this study, we analyzed a portion of the N-terminal region of p24 gene to ascertain if this conserved region may be employed for HIV-1 subtyping and further, to screen for the presence of some cryptic as well as known p24 escape mutations located within this part of the gag gene of the HIV-1 genome.

MATERIALS AND METHODS

A prospective pilot study was carried out between year 2007 to 2008 for screening HIV status of high risk population in India. A total of 285 persons were included in the study after obtaining approval from bio-safety ethics committee (SNGL/2007/87H). This study group included Commercial Sex Worker (CSW), Men having Sex with Men (MSM), Intra-Venus Drug User (IDA), blood transfusion recipients and persons having sex with CWS. Blood samples were collected in K_2-EDTA vacutainer tubes (Becton Dickinson, san Diego, Calif) and HIV-1 infection was confirmed by (a) Tridot rapid assay (J. Mitr & Co. Ltd., New Delhi, India) (b) multiple enzyme-linked immunosorbent assays (Biochem Immunosystems, Montreal, Canada) and (c) western blots analysis (LAV Blot HIV-1, Bio- Rad, France). Thirty out of 285 samples (10.5%) were HIV-1 positive. The study population never received anti retroviral therapy. The detailed background of the HIV-1 seropositive persons is described in Table 1. HIV-1 proviral DNA of these 30 HIV positive subjects were analyzed for studying the genetic diversity within p24 gag gene. Genomic DNA was extracted by using the QIAamp DNA blood mini kit 250 (QIAGEN, Germany) according to the manufacturer’s instructions. Concentration of DNA was measured using a qubit® 2.0 Fluorometer (Invitrogen, USA) and qubit™ dsDNA BR Assay Kit (Invitrogen, USA).

The HIV-1 DNA comprising of a 356 bp gag gene fragment corresponding to the region between amino acid positions 41 to 159 of p24 core nucleocapsid protein was amplified by nested PCR using a Veriti 96 well thermal cycler (Applied Biosystems; USA). The 1st round of PCR was performed using 5' – TCACCTAGACTTTGAATGCATGGG – 3' (outer forward) and 5' – CTAATACTGTATCCTGCTCCTGT – 3' (outer reverse) oligonucleotide primers as described by Sengupta et al. (2005). Briefly, 1 microgram of human genomic DNA was used as template in a 50 microliter reaction volume containing 1.5 mM MgCl₂, 200 mM (each) deoxynucleotide triphosphates, 10 pmoles of each primer and 2.0 unit of Taq DNA polymerase (Life Technologies, USA). The first-round thermal cycling conditions were as follows: 94°C for 2 min followed by 94°C for 30 sec, 50°C for 30 sec and 72°C for 90 sec for a

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Age</th>
<th>Gender</th>
<th>Risk Factor</th>
<th>Antiretroviral Treatment History</th>
</tr>
</thead>
<tbody>
<tr>
<td>APS001</td>
<td>28</td>
<td>Male</td>
<td>IDA</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS002</td>
<td>39</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS003</td>
<td>49</td>
<td>Male</td>
<td>MSM</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS004</td>
<td>21</td>
<td>Male</td>
<td>Transfusion</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS005</td>
<td>51</td>
<td>Male</td>
<td>Sex with CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS006</td>
<td>24</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS007</td>
<td>27</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS008</td>
<td>35</td>
<td>Male</td>
<td>Sex with CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS009</td>
<td>45</td>
<td>Male</td>
<td>Sex with CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS010</td>
<td>54</td>
<td>Male</td>
<td>MSM</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS011</td>
<td>19</td>
<td>Male</td>
<td>Transfusion</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS012</td>
<td>28</td>
<td>Male</td>
<td>MSM</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS013</td>
<td>34</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS014</td>
<td>31</td>
<td>Male</td>
<td>MSM</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS015</td>
<td>37</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS016</td>
<td>42</td>
<td>Male</td>
<td>Sex with CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS017</td>
<td>47</td>
<td>Male</td>
<td>Sex with CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS018</td>
<td>29</td>
<td>Male</td>
<td>IDA</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS019</td>
<td>24</td>
<td>Male</td>
<td>IDA</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS020</td>
<td>27</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS021</td>
<td>35</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS022</td>
<td>39</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS023</td>
<td>34</td>
<td>Female</td>
<td>Transfusion</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS024</td>
<td>42</td>
<td>Male</td>
<td>Sex with CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS025</td>
<td>47</td>
<td>Male</td>
<td>Transfusion</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS026</td>
<td>38</td>
<td>Female</td>
<td>CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS027</td>
<td>27</td>
<td>Female</td>
<td>IDA</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS028</td>
<td>23</td>
<td>Female</td>
<td>IDA</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS029</td>
<td>46</td>
<td>Male</td>
<td>Sex with CSW</td>
<td>Drug Naïve</td>
</tr>
<tr>
<td>APS030</td>
<td>32</td>
<td>Female</td>
<td>IDA</td>
<td>Drug Naïve</td>
</tr>
</tbody>
</table>

ID: Intra-Venus Drug User; CSW: Commercial Sex Worker; MSM: Men Sex with Men.

Table 1. Characteristics of Study Subjects
total of 35 cycles and a 7 min final extension at 72°C. The nested PCR amplicon was generated using 5'- TCA GCA TTA TCA GAA GGA GCC AC - 3' (inner forward) and 5'- TCC TTT GGC CCT TGT TTAT GC - 3' (inner reverse) primers. Five microliters of first round PCR product was used as a template for the 2nd round PCR which comprised of 1.5 mM MgCl2, 200 mM (each) deoxyribonucleotide triphosphates, 10 pmoles of each primer and 1.0 unit of Taq DNA polymerase (Life Technologies, USA) in a total volume of 25 microliters. The thermal cycling conditions were as follows: 95°C for 5 min followed by 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec for a total of 35 cycles and a 5 min final extension at 72°C. PCR products were resolved on a 2% agarose gel (Promega Corporation, Madison, USA); stained with ethidium bromide (0.5 µg/ml) and image captured using a Bio Rad gel documentation system (BIORAD, USA).

PCR amplicons were purified using QIAGEN PCR purification kit (QIAGEN, Germany), and the purified products were subjected double strand fluorescent nucleotide sequencing using the BigDye Terminator v3.1 Cycle Sequencing Kit and ABI PRISM® 3500 Dx Genetic Analyzer (Applied Biosystems, Calif) as per manufacturer’s instructions. Each PCR product was sequenced once and the nucleotide sequences thus generated were submitted to GenBank, and accession numbers assigned for further reference (see below). Throughout the text single letter amino acid codes were used as per IUPAC nomenclature.

All 30 nucleotide sequences were aligned with a reference panel of reported sequences from different geographic regions available in the HIV sequence database (http://www.hiv.lanl.gov/content/index) provided by the Los Alamos National Laboratory operated by the University of California, USA to study the nucleotide substitution pattern among them. The reference panel included 27 sequences of different global samples with the same p24 region of HIV-1 and consisting of all subtypes (A to K) including group N and O samples (used as outlier sample). Two sequences of each reference subtypes were taken for comparison. The multiple alignments were done using the ClustalW program. Phylogenetic and molecular evolutionary analyses were conducted using MEGA program, version 5.1 [27]. The phylogenetic tree was constructed using the Neighbor-Joining method [28]. The optimal tree with the sum of branch length = 1.03543829 is shown. The tree is drawn to scale, with branch lengths in the same unit as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and is in the unit of the number of base substitutions per site. The rate of variation among sites was modeled with a gamma distribution (shape parameter = 1).

**RESULTS**

Thirty different HIV-1 proviral DNA samples were analyzed in this study. Out of them, 29 were of subtype C while 1 was of subtype A. The subtyping results were consistent with the findings from our phylogenetic analysis and that generated using REGA HIV-1 Automated Subtyping Tool (31, 32). A ‘t’ test by way of interior branching test of phylogeny and using the bootstrapping procedure was used to construct phylogenetic tree based on the neighbor joining method. One thousand replicates were included in the boot strapping analysis and the constructed tree was viewed using Tree Explorer in MEGA 5.1. This phylogenetic tree generated from all 57 samples (30 ‘test’ and 27 ‘reference’) is shown in Fig. (1).

Multiple alignment of the computed amino acid sequences obtained from experimentally generated corresponding nucleotide sequence from test samples as well as retrieved reference ones indicated that within the TW10 epitope spanning 240 to 249 amino acid position of gag poly protein (i.e. 108 to 117 position of p24) [30] T110N mutation was present in samples APS023, APS025 & APS029 and T110S in sample APS021. Twenty one out of the 30 test samples harbored a G to A mutation at position 116 (G116A) of the gag poly protein. The escape mutation T54A was detected in APS020 and T107A in APS008 and APS020 respectively.

Similar analysis of the amino acid sequences from position 41 to 120 of the p24 gene of all 30 test sequences (APS001-APS030), a B subtype (B.FR.HXB2_K03455) as well as two C subtype reference sequences [C.BR.BR0 25_D_(U52953) and C.IN.95IN21068_(AP067155)] were performed. Result indicated that at amino acid position 120, all ‘APS’ sequences had either S, A or occasionally G amino acids barring two (APS006_HM569715 & APS026_HM921071) where the amino acid was ‘N’ as seen in the two reference C subtype sequences. Similarly, at amino acid position 41, the reference C subtype sequences had ‘T’ while
Fig. (1). Phylogenetic analysis of HIV-1 proviral DNA N terminal region of p24 of gag gene from 30 HIV-1 positive patients in Western India. The construction of the tree is described in the text. In the tree the test sequence names appears as sample number followed by accession number, as for example APS030- HM921075, APS002-HM560978. All the reference sequences used to construct the tree were obtained from Los Alamos National Laboratory HIV sequence database (http://www.hiv.lanl.gov/content/index). The reference sequence IDs shown in the tree are in the following sequence: subtype, country of origin, sample number and accession number in parenthesis. Example: A1.AU.PS1044 (DQ676872): Subtypes A1, reference sequence from Australia, sample number PS1044 and accession number DQ676872. Subtype N and O reference sequence were used as out groups during construction of the tree. Test samples included in this study were designated as APS. Reference samples of different subtypes used were as follows: subtype A1: PS1044, 92RW008; subtype A2: 97CDKTB48, 1445MV; subtype B: HXB2, 1058_11; subtype C: BR025_D, 95IN21068; subtype D: 01CM_4412HAL, TZ.A280; subtype F1: VI850, 93BR020_1; subtype F2: 02CM_0016BBY, CM53657; subtype G: DRCBL, HH8793_12_1; subtype H: 00GBAC4001, VI991; subtype J: 04CMU11421, SE9280_7887; subtype K: 97ZR_EQTB11, 96CM_MP535; subtype N: YBF_30, YBF_106; subtype O: ANT70, MPV-5180.
all 'APS' sequences had S as seen in the reference B subtype sequence (Fig. 2, black box, left) while at amino acid position 86/87, there was a tendency to incorporate a A/Q in the 'APS' sequences. Amino acids that correspond to the numbers shown in the figure are underlined in reference B subtype sequence (B.FR.HXB2_(K03455)).
A part of N-terminal region of gag p24 protein was analyzed to gather information about the genetic variability of HIV C subtype - APS samples (Black box). At amino acid position 86/87, the amino acid pair 'VH' (in reference sequence HXB2) is conserved throughout all the reference C subtype samples barring one (F286228_Brazil) which had the 'Q' amino acid at position 87. The sequences coded HXB2 (K03455) (B subtype) was taken as the reference sequence. Sequences AF286228_CHINA, AF286226_CHINA, AF286230_CHINA, U52953_Brazil, F286228_Brazil, AF443074_Botswana, AF286227_South Africa, AF286224_Zambia, AF286225_Zambia, AF443075_Botswana, U46016_Ethiopia, AF286235_Tanzania and AF286234_Tanzania were used as non Indian, reference C subtype sequences. At position 41, there is conservation of amino acid S in all the HIV C subtype - APS samples (Black box). At amino acid position 86/87, the amino acid pair ‘VH’ (in reference sequence HXB2) is conserved throughout all the reference C subtype samples barring one (F286228_Brazil) which had the 'Q' amino acid at position 87. Several of the APS samples deviate from this conservation and show 'A' or 'Q' or both in position 86/87 except APS022_HM921063 which shows 'I' at position 86 and APS018_HM921059 with ‘PP’ at positions 86/87.

**DISCUSSION**

The subtype of all thirty samples used in this study was identified in prior using the REGA HIV-1 Automated Subtyping Tool (32, 33). Twenty nine of them were found to be of subtype C while 1 belonged to subtype A category. The phylogenetic analysis of all the 30 test samples (29 C and 1 A subtype) were done wherein the sequences were challenged with a range of HIV subtype sequences retrieved from the public domain (Subtypes A1, A2, B, C, D, F1, F2, G, H, J, K, N and O, respectively). The result from both the methods of analysis was consistent with regard to the subtype-categorization of all 30 test HIV-1 samples.

A part of N-terminal region of gag p24 protein was analyzed to gather information about the genetic variability within one of the apparently most conserved regions of this viral genome. This specific fragment under focus included the TW10 epitope of p24 capsid responsible for HLA mediated cytotoxic T-lymphocyte responses which play an active role in the control of HIV infection. Escape-mutations within this epitope lead to reduced immune response and rapid progression towards AIDS. The two hotspot escape mutations within this epitope are T110N and G116V. In 3 of our test samples, T110N while in 1, T110S was detected. These mutations may lead to rapid progression towards AIDS in the patient. The G116A mutation acts as an escape mutation in subtype B samples whereas the same is a conserved substitution for subtype C samples. In our study, we observed this consensus ‘A’ at amino acid position 116 of the gag p24 protein in 21 out of 30 of our study subjects.
Thus 8 subtype C test samples did not carry the conserved G116A mutation. Few more escape mutations encountered in subtype C samples were detected in our test samples also. These were T54A in APS020 and T107A in APS008 & APS020 respectively. These observations broadly indicate the significant genetic variance presented by this particular portion of the gag p24 region of the HIV-1 genome that qualifies it for focused study and analysis.

Phylogenetic tree construction despite the recent advances in computing algorithm that creates them does not generate sufficient location-specific information of amino acid substitutions within a gene. In order to address this lacuna, we undertook a multiple alignment analysis of the amino acid sequences from position 41 to 120 of the gag p24 gene of our test sequences. This was done in two phases. First, the sequences were aligned with a limited number of reference samples comprising of one standard B and two C subtypes. In the second phase, the reference samples were expanded to include 14 HIV-1 C subtype samples from different countries and 1 standard HIV-1 B subtype as master reference sequence. One of the objectives of second phase analysis was to observe the clustering pattern of our test HIV-1 sequences when analyzed along with other sequences.

**Fig. (4)**. Phylogenetic study of p24 gene sequences of 30 HIV-1 seropositive samples from India (29 C subtype and 1 A subtype) with 12 and 1 other C and B samples from different countries respectively. For Indian samples, the name is indicated by ‘APS’ followed by a 3 digit number and genebank accession number. For samples from other countries, the Genebank accession number is shown followed by the name of the country. ‘South’ refers to South Africa. The different samples used for the study are as follows: APS 001-030 with Genebank accession numbers GU647157, HM569757, HM569712, HM569713, HM569714, HM569715, HM569716, HM589196, HM589197, HM589198, HM921053, HM921054, HM921055, HM921056, HM921057, HM921058, HM921059, HM921060, HM921061, HM921062, HM921063, HM921064, HM921065, HM921066, HM921070, HM921071, HM921072, HM921073, HM921074, HM921075, JN173782 respectively; Non Indian C subtype samples: AF286229_CHINA, AF286226_CHINA, AF286230_CHINA, U52953_Brazil, F286228_Brazil, U46016_Ethiopia, AF286234_Tanzania and AF286235_Tanzania; Non Indian B subtype: K03455_HXB2.
similar ones of identical subtype category but reported from different parts of the world. The results obtained from both the analyses were similar.

Patterns at 3 different amino acid positions, viz., 120, 86/87 and 41 within the gag p24 region were found to be of interest. At position 120 occurrence of S, A or occasionally G amino acid appeared to be a uniform phenomenon in all the HIV-1 C subtype sequences irrespective of their geographical origin. At position 86/87 we observed a trend. Here the amino acid pair ‘VH’ seen in the master reference strain HXB2 (Fig. 3) was incidentally found to be conserved throughout all the reference C subtype samples barring one reported from Brazil. However, APS samples included in our study deviated from this conservatism and showed ‘A’ or ‘Q’ or both at position 86/87, APS022_HM921063 and APS018_HM921059 being the 2 exceptions with ‘I’ and a ‘PP’ at positions 86 and 86/87 respectively (Fig. 3).

However, the pattern at position 41 is most intriguing. Here all the reference C subtype sequences had the amino acid ‘T’ while all ‘APS’ sequences had the amino acid S as was observed in the master reference B subtype sequence. This therefore appeared to be a molecular signature associated with all the APS test sequences analyzed in this study. Substitutions at position 86/87 on the other hand appeared to be rather like a trend unlike that at amino acid position 41 where the occurrence of amino acid S was rigid and a characteristic which appeared to be associated only with the Indian “APS” HIV-1 subtype C sequences.

Due to occurrence of such molecular signatures within the p24 protein in the ‘APS’ Indian HIV-1 proviral DNA (29 C and 1 A subtypes) we wanted to see their overall clustering pattern when grouped along with all the 14 reference HIV-1 C subtype P24 amino acid sequences used for analysis in Fig. (3). For this, the clustalW multiple alignments were used to construct a phylogenetic tree using the neighbor-joining method by MEGA version 5.1 program. A radial pattern of the tree was generated for analysis.

Despite all the sequences taken up for this analysis except APS030-HM921075 and K03455_HXB2 were of C subtype, all Indian C subtype APS samples formed a distinct cluster in the ‘left’ except APS030-HM921075. On the other hand, the cluster in the ‘right’ comprised of all other non-Indian HIV-1 C subtypes (AF286229_CHINA, AF286230_CHINA, U52953_Brazil, F286228_Brazil, AY043173_South Africa, AF443074_Botswana, AF286227_South Africa, AF286224_Zambia, AF286225_Zambia, AF443075_Botswana, U46016_Ethiopia, AF286234_Tanzania and AF286235_Tanzania) except AF286226_CHINA which clubbed with the ‘left’ APS group. As expected, APS030-HM921075 (the only Indian HIV-1 A subtype) despite its link with the ‘right’ cluster, occupied a distal position with the reference K03455_HXB2 and AF286229_CHINA as its closest neighbors (Fig. 4). This is understandable since the genetic sequence pattern of A subtype of APS030-HM921075 prevented it from closely merging with the C subtype group and forced it to position itself to an isolated, distal location within the tree.

It appears that possibly there is a geography-specific retention of specific amino acids of choice at precise location(s) within the p24 gene of HIV-1 that confer some yet to be defined advantage to the pathogen. This study clearly indicates p24 gene-specific mutations in Indian samples of HIV-1 C subtypes to follow a pattern which despite its apparent randomness translated p24 proteins the amino acid sequences of which formed a distinct India-specific cluster when compared to C subtypes from other parts of the world. A single Chinese strain (AF286226_CHINA) clubbing with the ‘right’ Indian subtype C cluster is intriguing and might reflect merging genotype trend thus warranting further investigation.

CONCLUSION

In summary, from this study it was observed that as concluded earlier, HIV-1 p24 gene sequence could be suitably used for genotype identification as well as construction of informative phylogenetic tree. Retention of some specific amino acids at certain locations of the p24 gene within the test HIV-1 proviral DNA samples as was detected from multiple alignment analysis indicated possible geographic molecular signature which might confer adaptive advantage to the pathogen for growth and survival.

To our knowledge, this is the first study in recent times to indicate a distinct clustering pattern of HIV-1 C subtypes from India when compared to those samples from other countries on the basis on p24 amino acid sequence. There is a need for continuous monitoring of the emergence or conservation of sequence variations within the endemic subtypes in the country and since year 2005, this is also the first report of p24 N-terminal sequence-based HIV-1 subtyping and genetic diversity study from a selected HIV-1 infected population from India.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

PATIENT CONSENT

Declared none.

HUMAN/ANIMAL RIGHTS

Declared none.

REFERENCES


A robust HIV-1 viral load detection assay optimized for Indian sub type C specific strains and resource limiting setting

Arpan Acharya1, Salil Vaniawala2, Parth Shah3, Harsh Parekh4, Rabindra Nath Misra5, Minal Wani1 and Pratap N Mukhopadhyaya4*

Abstract

Background: Human Immunodeficiency Virus Type 1 (HIV-1) viral load testing at regular intervals is an integral component of disease management in Acquired Immunodeficiency Syndrome (AIDS) patients. The need in countries like India is therefore an assay that is not only economical but efficient and highly specific for HIV-1 sub type C virus. This study reports a SYBR Green-based HIV-1 real time PCR assay for viral load testing and is designed for enhanced specificity towards HIV-1 sub type C viruses prevalent in India.

Results: Linear regression of the observed and reference concentration of standards used in this study generated a correlation coefficient of 0.998 (p < 0.001). Lower limit of detection of the test protocol was 50 copies/ml of plasma. The assay demonstrated 100% specificity when tested with negative control sera. The Spearman coefficient of the reported assay with an US-FDA approved, Taqman probe-based commercial kit was found to be 0.997. No significant difference in viral load was detected when the SYBR Green based assay was used to test infected plasma stored at −20°C and room temperature for 7 days respectively (Wilcoxon signed rank test, p = 0.105). In a comparative study on 90 pretested HIV-1 positive samples with viral loads ranging from 5,000–25,000 HIV-1 RNA copies/ml and between two commercial assays it was found that the later failed to amplify in 13.33% and 10% samples respectively while in 7.77% and 4.44% samples the copy number values were reduced by >0.5 log value, a figure that is considered clinically significant by physicians.

Conclusion: The HIV-1 viral load assay reported in this study was found to be robust, reliable, economical and effective in resource limited settings such as those existing in India. PCR probes specially designed from HIV-1 Subtype C-specific nucleotide sequences originating from India imparted specificity towards such isolates and demonstrated superior results when compared to two similar commercial assays widely used in India.

Keywords: HIV-1, Gag, SYBR Green, RT PCR

Background

At present estimating HIV-1 RNA plasma concentration is the standard of care in monitoring the progression of HIV/AIDS and is believed to have clinically useful prognostic value [1,2]. Three different commercial assays [NucliSens HIV-1 QT RNA Assay (bioMérieux, France), Quantiplex HIV-1 RNA 3.0 Assay (Bayer Diagnostics Division, USA) and COBAS TaqMan HIV-1 48 test (Roche Molecular Systems, USA)] presently dominate the HIV-1 viral load testing market but are prohibitively expensive [3].

Real time PCR technology is currently considered the gold standard for HIV-1 viral load detection because of its high assay specificity, sensitivity and wide linear range of detection [4]. TaqMan chemistry is widely popular and accepted method but in order to develop this on real time PCR platform for quantification of HIV-1 viral RNA, three conserved regions within the viral genome are required all in close proximity which is often difficult due to the heterogenic genomic structure of HIV-1. The
problem of sub-optimal performance by TaqMan chemistry based commercial assay is mostly encountered in countries where non B subtype of HIV-1 is prevalent because these assays are widely validated in USA and Europe where HIV-1 subtype B dominates [3,5]. In addition to the TaqMan chemistry, SYBR Green chemistry based real time PCR is also widely used for HIV-1 viral RNA quantification assays.

In this study, we report a real time RT-PCR assay specifically optimized for HIV-1 C subtype that uses SYBR Green chemistry for quantitative detection of HIV-1 viral RNA in human plasma. The assay is designed to target a conserved region within the gag gene of the viral genome and is less expensive compared to those which employ dual-labeled TaqMan probes. The protocol gave consistent results across a wide range of clinical samples.

Results

Linear range, detection limit and inter-assay precision

The linearity of the assay was determined using reference standards constructed specifically for this assay by using a recombinant plasmid harboring the amplified gag gene fragment as insert. Linear regression of the observed and reference concentrations yielded a correlation coefficient of 0.998 (p < 0.001). A standard curve was generated by plotting the threshold cycles of reference standards versus their log concentrations (Figure 1). The dissociation curve of all reference standards showed a melting temperature of 85.5°C (Figure 2).

In order to compute the detection limit samples with known viral load in the range of 10 to 500 HIV-1 RNA copies/ml were analyzed. The results showed detectable amplification in 5 (25%) of the replicates for standard having 10 HIV-1 RNA copies/ml, in 10 (50%) of the replicates for those having 30 HIV-1 RNA copies/ml, and 20 (100%) having 50, 100 and 500 copies/ml of HIV-1 RNA respectively. The data is represented as a bar diagram where the log concentrations are compared with their corresponding threshold cycles (Figure 3). From this analysis, the lower limit of the assay was fixed at 50 copies/ml of HIV-1 RNA.

In order to evaluate the inter-assay precision of the test protocol the reference panel comprising of 25 samples with known HIV-1 viral load were tested in triplicates through 10 independent experiments. Statistical analysis of the data for standard deviation (SD) and coefficient of variation are shown in Table 1. The coefficient of variation was found to increase as the reference viral load approached detection limit of the assay.

Sensitivity and specificity of the in-house HIV-I RT PCR assay

Clinical sensitivity of the assay was determined using samples from the reference panel. All 25 samples were found to be positive using the in-house SYBR Green RT PCR assay indicating 100% sensitivity. The specificity of the assay was also found to be 100% since none of the samples showed any detectable amplification from our reference panel of HIV-1-negative samples (Figure 4). This was further confirmed by analyzing the melting temperature profile of the amplicons using dissociation curve analysis. No significant fluorescent signals were detected from any non-specific amplification including primer dimers.

Comparison of viral load measurements using the in-house and the US-FDA approved artus HIV-1 RG RT-PCR kit

Viral load of 100 HIV-1 sero-positive samples was determined both by a commercially available US FDA-approved artus HIV-1 RG RT-PCR kit (Qiagen, Germany) as well as the SYBR Green RT PCR assay described in this study. The log concentration of HIV-1 RNA copies/ml of samples were compared by plotting results obtained using the present study protocol with those obtained using artus HIV-1 RG RT-PCR kit (Figure 5). The spearman coefficient of the in house assay with that of artus HIV-1 RG RT-PCR kit was found to be 0.997.

Comparison of sensitivity with two commercial HIV-1 RNA detection assays

The detection limit of HIV RNA Real-Quant (Sacace Biotechnologies S.r.l, Italy) and HIV-1 Real Time RT PCR kit (Life River, Shanghai ZJ Biotech, China) are 500 and 1,000 HIV-1 RNA copies/ml respectively. Out of the 90 samples whose viral load varied between 5000 - 25,000 HIV-1 RNA copies/ml as determined by artus HIV-1 RG RT-PCR kit (Qiagen, Germany), 12 (13.33%) and 9 (10.0%) samples fail to show any detectable amplification when tested using HIV RNA Real-TM Quant (Sacace Biotechnologies S.r.l, Italy) and HIV-1 Real Time RT PCR kit (Life River, Shanghai ZJ Biotech, China) respectively.
While 7 (7.77%) and 4 (4.44%) samples showed underestimation of viral load at a value that is greater than 0.5 log when compared with reference panel results. All the above mentioned samples had comparable results when tested using the protocol described in this study as shown in Table 1.

**Effect of ambient temperature on stability of HIV-1 RNA**

Effect of storage of plasma at room temperature was studied by analyzing 10 plasma samples each stored in (1) -20°C freezer and (2) room temperature (25°C) for 7 days. These samples were subjected to viral load testing and each reaction was performed in triplicate. Results indicate a w- value of 11.5, mean difference of 1.01, sum of positive ranks equaling to 43.5 and sum of negative ranks equaling to 11.5 respectively. As the critical value of w for n = 10 (p ≤ 0.05) was 8, the result was not significant meaning that there was no significant difference in viral load detected in both the groups (Wilcoxon

![Figure 2: Dissociation curve of the HIV-1 reference standards.](image)

*Figure 2* Dissociation curve of the HIV-1 reference standards. The negative first derivative of fluorescent signal (Y-axis) was plotted against Temperature (X-axis).

![Figure 3: Bar diagram used for finding the detection limit of the assay.](image)

*Figure 3* Bar diagram used for finding the detection limit of the assay. Log concentration of the dilution series plotted with their corresponding Threshold Cycles. Series 1: Threshold Cycles; Series 2: Log concentration of viral load; CTN10: 10 HIV-1 RNA copies/ml; CTN30: 30 HIV-1 RNA copies/ml; CTN50: 50 HIV-1 RNA copies/ml; CTN100: 100 HIV-1 RNA copies/ml; CTN500: 500 HIV-1 RNA copies/ml.
signed rank test, p = 0.105) (Figure 6). The box and whisker plot in Figure 6 show differences between medians (DBM), over all visible spread (OVS) and DBM as a percentage of OVS within the two groups equaling to 0.013, 1.114 and 1.17% respectively indicating no significant difference between the two data sets.

Discussion

At present SYBR Green chemistry based real time PCR is widely used for development of sensitive and economical quantitative PCR assays due to its competitive performance when compared to TaqMan chemistry based counterpart tests. In this study we described a SYBR Green chemistry based sensitive real time PCR assay for quantitative detection of HIV-1 viral load in human plasma samples.

Two assay panels were created for the study. The reference study panel comprised of 25 HIV-1 sero-positive samples with reference viral load assigned to them using artus HIV-1 RG RT-PCR kit (Qiagen, Germany) and an equal number of samples from healthy individuals. This panel was primarily used for development of the assay. The clinical study panel on the other hand was a secondary one comprising of 100 samples and was used to evaluate the performance of the assay. Reference viral load values were also assigned to this panel using artus HIV-1 RG RT-PCR kit (Qiagen, Germany).

Most commercial HIV-1 viral load assays are designed and optimized for the detection of viruses found in North America and Europe. As a result, these assays either under-quantify or fail to amplify some of the HIV-1 non-B subtypes [6-8]. This observation motivated us to identify the gag gene from Indian subtype C HIV-1 viruses as the target to develop this SYBR Green chemistry based HIV-1 viral load estimation assay. The linearity of our assay ranged from 50 HIV-1 RNA copies to 10^7 HIV-1 RNA copies/ml of plasma sample and had a lower limit of detection of 50 HIV-1 RNA copies/ml.

The normalized measure of dispersion of the probability distribution of data generated from inter assay precision testing and expressed as coefficient of variation was found to increase as the values came closer to the detection limit. This was found to be acceptable since the linear range of detection was wide and at par with standard quantitative real time PCR assays [9-11].

The sensitivity of the assay was found to be 100%. No amplification in 25 clinical samples from HIV-1 sero negative reference panel indicated 100% specificity. This finding was further consolidated by dissociation curve analysis of all the sero positive reference panel samples. Sharp peaks at 85.5°C were detected in all sero positive reference panel samples and was absent in samples from healthy individuals. The results obtained from clinical study panel indicated a high level of concordance between the viral loads generated using the commercial kit and the assay described in this study.

The poor sensitivity of two commercial assays [HIV RNA Real-TM Quant (Sacace Biotechnologies S.r.l, Italy) and HIV-1 Real Time RT PCR kit (Life River, Shanghai ZJ Biotech, China) respectively] shown in this study may be due to presence of nucleotide mismatches within the primer and probe binding regions of HIV-1 genome. We believe that suboptimal attention towards HIV-1 subtype C specific sequences originating from India was the primary reason for such results. This finding is in line with data presented by Drexler et al. which demonstrated that HIV-1 positive samples which were reported negative by three commercial assays were shown to have a HIV-1 viral load of >5,000 copies/ml when tested using a home brew real time PCR protocol [7].

To evaluate the performance of the assay in a resource limited setting we used the Wilcoxon signed-rank test

<table>
<thead>
<tr>
<th>Reference concentration (copies/ml)</th>
<th>Measured concentration (copies/ml)</th>
<th>Standard deviation (SD)</th>
<th>Coefficient of variation (CV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>50</td>
<td>1.65</td>
<td>0.0325</td>
</tr>
<tr>
<td>124</td>
<td>120</td>
<td>3.06</td>
<td>0.0251</td>
</tr>
<tr>
<td>344</td>
<td>345</td>
<td>0.47</td>
<td>0.0014</td>
</tr>
<tr>
<td>871</td>
<td>863</td>
<td>5.42</td>
<td>0.0063</td>
</tr>
<tr>
<td>1,235</td>
<td>1,250</td>
<td>10.14</td>
<td>0.0082</td>
</tr>
<tr>
<td>16,383</td>
<td>16,358</td>
<td>17.68</td>
<td>0.0011</td>
</tr>
<tr>
<td>267,816</td>
<td>267,098</td>
<td>507.70</td>
<td>0.0019</td>
</tr>
<tr>
<td>31,293</td>
<td>31,299</td>
<td>4.48</td>
<td>0.0001</td>
</tr>
<tr>
<td>541</td>
<td>534</td>
<td>4.71</td>
<td>0.0088</td>
</tr>
<tr>
<td>1,476</td>
<td>1,466</td>
<td>6.00</td>
<td>0.0045</td>
</tr>
<tr>
<td>4,572</td>
<td>4,555</td>
<td>12.26</td>
<td>0.0027</td>
</tr>
<tr>
<td>2571,210</td>
<td>2569,050</td>
<td>1,526.88</td>
<td>0.0006</td>
</tr>
<tr>
<td>76</td>
<td>89</td>
<td>9.43</td>
<td>0.1145</td>
</tr>
<tr>
<td>43,110</td>
<td>43,135</td>
<td>18.15</td>
<td>0.0004</td>
</tr>
<tr>
<td>5,447</td>
<td>5,414</td>
<td>23.10</td>
<td>0.0043</td>
</tr>
<tr>
<td>21,312</td>
<td>21,393</td>
<td>57.28</td>
<td>0.0027</td>
</tr>
<tr>
<td>96</td>
<td>94</td>
<td>1.65</td>
<td>0.0174</td>
</tr>
<tr>
<td>635,469</td>
<td>635,592</td>
<td>86.97</td>
<td>0.0001</td>
</tr>
<tr>
<td>626</td>
<td>620</td>
<td>4.48</td>
<td>0.0072</td>
</tr>
<tr>
<td>54,886</td>
<td>54,853</td>
<td>23.81</td>
<td>0.0004</td>
</tr>
<tr>
<td>8,742</td>
<td>8,781</td>
<td>27.58</td>
<td>0.0031</td>
</tr>
<tr>
<td>917,838</td>
<td>917,648</td>
<td>133.88</td>
<td>0.0001</td>
</tr>
<tr>
<td>74,511</td>
<td>74,552</td>
<td>29.23</td>
<td>0.0004</td>
</tr>
<tr>
<td>944</td>
<td>965</td>
<td>15.32</td>
<td>0.0161</td>
</tr>
<tr>
<td>9,648</td>
<td>9,984</td>
<td>237.82</td>
<td>0.0242</td>
</tr>
</tbody>
</table>

The standard deviation and the coefficient of variation were determined using data from 10 independent experiments.
(similar to the paired student's t-test) as a non-parametric statistical hypothesis test to compare the two related samples (plasma samples stored in −20°C and at room temperature respectively) and ascertain whether their population mean ranks differed. Results indicated that there were no significant differences in viral loads detected in both the groups (Wilcoxon signed rank test, \( p = 0.105 \)) thus confirming the conclusions drawn from similar studies undertaken elsewhere [12]. In resource limiting setting like India cold chain facility is not always available for transportation of plasma samples from collection centers to the reference laboratories. Stability of the samples at ambient temperature is therefore important for proper clinical management of the patients.

**Conclusions**

There has been a significant increase in sensitivity of HIV-1 RNA quantification due to systematic involvement of advanced molecular techniques over the years. However, enhanced cost and demand for greater reagent and instrument sophistication is a necessary evil that needs to be traded off to achieve high level of accuracy and consistency. This is therefore a compelling reason for exploring areas where economy and ease of work can be linked to accuracy, sensitivity and specificity in HIV-1 RNA quantification. To address these issues we developed and evaluated the performance of SYBR Green chemistry based HIV-1 viral load detection assay. The significance of this study is high for a developing country like India which has a high rate of HIV-1 infection. Some commercial assays originating from Europe and China with poor attention to India-specific HIV-1 Subtype C genome sequences generated sub optimal data in our study and therefore carry a risk of generating error prone data with potential to mislead therapeutic decisions. In these contexts the assay described in this study specifically designed for Indian Subtype C viruses is useful, economical and reliable for routine monitoring of HIV-1 viral load in resource limited settings.

**Methods**

**Clinical sample collection and processing**

The reference panel comprised of blood samples collected from 25 each of HIV-1 sero-positive (drug naïve) and sero-negative (with no AIDS related risk factors) individuals from India in K₂-EDTA vacutainer tubes (Becton Dickinson, San Diego, California, USA) after obtaining written consent from the donors. The status
of HIV-1 infection was confirmed by (a) Tridot rapid assay (J. Mitra & Co. Ltd., New Delhi, India) (b) multiple enzyme-linked immunosorbent assays (Biochem Immunosystems, Montreal, Canada) and (c) western blot analysis (LAV Blot HIV-1, Bio- Rad, France). The viral load for each of these samples was determined using a US FDA-approved kit (artus HIV-1 RG RT-PCR kit; Qiagen, Germany) and used as reference concentration.

Plasma were separated from the blood samples within 4 hours of collection and stored in a −20°C freezer. Details of the study participants are described in Table 2.

The clinical study panel comprised of blood samples collected from 100 HIV-1 sero-positive individuals. The viral load for each of these samples was determined using a US FDA-approved kit (artus HIV-1 RG RT-PCR kit; Qiagen, Germany) and used for analysis and comparison.

The study program was duly approved by a bio-safety and bio-ethics committee (Approval number SNGL/2012/GY06) overseeing ethical aspect of such research programs within the organization.

![Comparison of viral load from 10 plasma samples after storage at (X) -20°C degree and (Y) room temperature (25°C) for 7 days prior to SYBR Green Real Time PCR assay presented in a box and whisker plot.](image)

**Table 2 Characteristics of reference study panel**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Age</th>
<th>Gender</th>
<th>Sample ID</th>
<th>Age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBL001</td>
<td>28</td>
<td>Male</td>
<td>DBN001</td>
<td>29</td>
<td>Male</td>
</tr>
<tr>
<td>DBL002</td>
<td>39</td>
<td>Female</td>
<td>DBN002</td>
<td>24</td>
<td>Female</td>
</tr>
<tr>
<td>DBL003</td>
<td>44</td>
<td>Male</td>
<td>DBN003</td>
<td>21</td>
<td>Male</td>
</tr>
<tr>
<td>DBL004</td>
<td>21</td>
<td>Male</td>
<td>DBN004</td>
<td>46</td>
<td>Male</td>
</tr>
<tr>
<td>DBL005</td>
<td>47</td>
<td>Male</td>
<td>DBN005</td>
<td>35</td>
<td>Male</td>
</tr>
<tr>
<td>DBL006</td>
<td>24</td>
<td>Female</td>
<td>DBN006</td>
<td>39</td>
<td>Female</td>
</tr>
<tr>
<td>DBL007</td>
<td>27</td>
<td>Female</td>
<td>DBN007</td>
<td>42</td>
<td>Female</td>
</tr>
<tr>
<td>DBL008</td>
<td>35</td>
<td>Male</td>
<td>DBN008</td>
<td>31</td>
<td>Male</td>
</tr>
<tr>
<td>DBL009</td>
<td>45</td>
<td>Male</td>
<td>DBN009</td>
<td>36</td>
<td>Male</td>
</tr>
<tr>
<td>DBL010</td>
<td>54</td>
<td>Male</td>
<td>DBN010</td>
<td>45</td>
<td>Female</td>
</tr>
<tr>
<td>DBL011</td>
<td>19</td>
<td>Female</td>
<td>DBN011</td>
<td>26</td>
<td>Male</td>
</tr>
<tr>
<td>DBL012</td>
<td>28</td>
<td>Male</td>
<td>DBN012</td>
<td>21</td>
<td>Female</td>
</tr>
<tr>
<td>DBL013</td>
<td>34</td>
<td>Female</td>
<td>DBN013</td>
<td>37</td>
<td>Male</td>
</tr>
<tr>
<td>DBL014</td>
<td>31</td>
<td>Male</td>
<td>DBN014</td>
<td>41</td>
<td>Male</td>
</tr>
<tr>
<td>DBL015</td>
<td>25</td>
<td>Female</td>
<td>DBN015</td>
<td>28</td>
<td>Male</td>
</tr>
<tr>
<td>DBL016</td>
<td>42</td>
<td>Male</td>
<td>DBN016</td>
<td>37</td>
<td>Female</td>
</tr>
<tr>
<td>DBL017</td>
<td>47</td>
<td>Male</td>
<td>DBN017</td>
<td>35</td>
<td>Male</td>
</tr>
<tr>
<td>DBL018</td>
<td>29</td>
<td>Female</td>
<td>DBN018</td>
<td>32</td>
<td>Female</td>
</tr>
<tr>
<td>DBL019</td>
<td>24</td>
<td>Male</td>
<td>DBN019</td>
<td>27</td>
<td>Male</td>
</tr>
<tr>
<td>DBL020</td>
<td>27</td>
<td>Female</td>
<td>DBN020</td>
<td>26</td>
<td>Female</td>
</tr>
<tr>
<td>DBL021</td>
<td>29</td>
<td>Female</td>
<td>DBN021</td>
<td>29</td>
<td>Female</td>
</tr>
<tr>
<td>DBL022</td>
<td>36</td>
<td>Female</td>
<td>DBN022</td>
<td>38</td>
<td>Male</td>
</tr>
<tr>
<td>DBL023</td>
<td>34</td>
<td>Female</td>
<td>DBN023</td>
<td>39</td>
<td>Female</td>
</tr>
<tr>
<td>DBL024</td>
<td>31</td>
<td>Male</td>
<td>DBN024</td>
<td>34</td>
<td>Male</td>
</tr>
<tr>
<td>DBL025</td>
<td>50</td>
<td>Male</td>
<td>DBN025</td>
<td>38</td>
<td>Male</td>
</tr>
</tbody>
</table>
Oligonucleotide primer design
In order to avoid primer mismatch the most conserved region of HIV-1 gag gene was chosen for designing of the oligonucleotide primers. This region was identified by analyzing several full length HIV-1 gag gene nucleotide sequences submitted from India and obtained from NCBI GenBank database (www.ncbi.nlm.nih.gov/Genbank) [13]. The sequences were aligned with the help of multiple sequence alignment program ClustalW v.2 [14] and conserved regions identified were used to manually design PCR primers. Five different sets of forward and reverse primers were designed, their reaction kinetics as well as self-complimentarity checked using Primer Express 3.0 program (Life Technologies, USA) and experimentally evaluated for their efficiency in amplifying target DNA.

Construction of reference standards
For construction of reference standards, the HIV-1 gag gene PCR fragment generated using the newly designed PCR primers was cloned in the pCR4-TOPO vector (Life Technologies, USA) following manufacturer's instructions. The recombinant plasmid was purified and identity of the cloned insert confirmed by double strand DNA sequencing using the BigDye Terminator v3.1 Cycle Sequencing Kit and ABI PRISM® 3500 Dx Genetic Analyzer (Applied Biosystems, USA).

In order to generate quantization standard, the recombinant plasmid was serial diluted and calibrated against reference standards of the US FDA approved artus HIV-1 RG RT-PCR kit (Qiagen, Germany). A seven dilution series of the reference standard was prepared with copy numbers as follows: 10^1, 10^2, 10^3, 10^4, 10^5, 10^6 and 10^7 HIV-1 viral RNA copies/ml.

RNA extraction and reverse transcription–real time PCR
HIV-1 RNA was extracted from plasma samples using a QIAamp Viral RNA mini kit according to manufacturer’s instructions (Qiagen, Germany). SYBR Green 1 - based one-step real time quantitative RT-PCR amplification was performed using a StepOnePlus Real Time PCR System (Applied Biosystems, USA). Fifty μL of a reaction mixture comprised of 25 μL of 2X QuantiFast SYBR Green RT-PCR master mix (Qiagen, Germany) which contained HotStarTaq DNA Polymerase, 0.5 μL 100X QuantiFast RT Mix, 20 μL of RNA and each primers (forward and reverse) at a final concentration of 5 pmol each reaction.

The nucleotide sequence of the primers used in this study is as follows: 5’- ACATCAAGCAGGCCATGCAAAT - 3’ (forward) and 5’- TACTAGTAGTCTCTGCTATGTC - 3’ (reverse). ROX dye was used as a passive reference to normalize fluorescent signals. The thermal cycling profile comprised of 30 minutes of reverse transcription at 45°C, 5 minutes of heating at 95°C to activate HotStart Taq DNA Polymerase followed by 45 cycles of PCR amplification each comprising of 95°C for 15 seconds, 55°C for 15 seconds, and 72°C for 30 seconds. Data collection was done at 72°C of each cycle. Following amplification, a melting curve (dissociation curve) analysis was performed to verify the authenticity of the amplified products by checking their specific melting temperatures (Tm). In each batch all reference standards as well as samples were tested in triplicate and appropriate negative as well as reagent controls were included to exclude the possibility of cross contamination during the setting up of reactions.

Intra- and inter-assay validation
For intra- and inter-assay validation 25 HIV-1 sero-positive samples were tested 10 times each in different batches independently and in triplicate including the reference standards and controls and their coefficient of variation was calculated.

Sensitivity and specificity
For finding limit of detection of the assay 10, 30, 50, 100 and 500 copies/ml of serially diluted standards were tested in triplicate in 20 independent experiments. In order to measure the specificity of the assay all known HIV-1 sero-negative samples (Table 2) were subjected to real time PCR amplification followed by comparison of the melting-temperature data with known positives as well as no-template control (NTC) samples.

Comparing sensitivity of detection against two commercial HIV-1 copy number detection assays
A panel of 90 plasma samples which was a part of the study panel and with HIV-1 viral copy number ranging from 5,000–25,000 HIV-1 RNA copies/mL as estimated using artus HIV-1 RG RT-PCR kit (Qiagen, Germany) and confirmed as sub type C using method described by Siddappa et al. [15] were simultaneously subjected to quantitative detection of HIV-1 RNA using (a) HIV RNA Real-TM Quant (Sacace Biotechnologies S.r.l, Italy) (b) HIV-1 Real Time RT PCR kit (Life River, Shanghai Z) Biotech, China) and (c) assay method developed in this study. All RT-PCR reactions were set in duplicate in two separate thermal cyclers of identical make (StepOne Plus Real Time PCR System; Applied Biosystems, USA) in order to minimize the influence of experimental errors.

Effect of storage time and temperature on stability of plasma samples
Plasma samples of 10 patients with known viral load were collected and two aliquots were made from each of them. One set of plasma were stored at −20°C and the other at 25°C. RNA was extracted from both set of samples at day seven and their viral load estimated in triplicate using the method described in this study.
Statistical analysis
The linearity of the assay was determined by linear regression method using the MedCalc v.12.7.2 (Acacilaan 22, B-8400 Ostend, Belgium) program. Sensitivity of the assay was expressed as the percentage of samples that scored positive using the protocol described in this study divided by the total number of samples that scored positive using an US -FDA approved commercial assay. Specificity of the assay was calculated as the number of negative samples scored using the present study divided by the total number of known negative samples obtained from healthy donors and was expressed in percentage. Performance of the assay was measured by doing a comparative analysis with artus HIV-1 RG RT-PCR kit (Qiagen, Germany) and applying the Spearman correlation coefficient formula (http://www.wessa.net/rankcorr.wasp). The effect of ambient temperature on the stability of HIV-1 RNA was analyzed using the Wilcoxon signed rank test [16] and the box and a whisker plot was created using BioStat program (AnalystSoft Inc, USA).

Competing interests
The authors declare no competing interests.

Authors’ contributions
AA and PS carried out the viral RNA extraction and real time PCR reactions. HP generated the recombinant plasmids used in this study. AA and SV analyzed the amplification data and advised on the resource population. RNM and MW assisted in drafting the manuscript. PNM and SV conceptualized the work and drafted the complete manuscript. All authors read and approved the final manuscript.

Acknowledgments
The authors are grateful to all the study participants for agreeing to provide their valuable clinical samples for this study.

Author details
Dr. D. Y. Patil Biotechnology & Bioinformatics Institute, Tathawade, Pune, INDIA. 5SN Genelab, 2nd Floor, President Plaza, Tower A, Nanpura, Ring Road, Surat, Gujarat; India. 6Supratex Micropath Laboratory & Research Institute, Kedar Building, Elphinstone, Opposite Krupa Petrol Pump, Parimal Garden, Ahmedabad, Gujarat, India. 7Interdisciplinary Science, Technology and Research Academy, AISC, Hidayatullah Road, Camp, Pune 411 001, Maharashtra, India. 8Department of Microbiology, Padmashree Dr D Y Patil Medical College & Research Centre, Sant Tukaram Nagar, Pimpri, Pune 411 018, Maharashtra, India.

Received: 21 January 2014 Accepted: 24 May 2014
Published: 30 May 2014

References

Cite this article as: Acharya et al.: A robust HIV-1 viral load detection assay optimized for Indian sub type C specific strains and resource limiting setting. Biological Research 2014, 47:22.

Submit your manuscript to BioMed Central and take full advantage of:

• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit
Development, Validation and Clinical Evaluation of a Low Cost In-House HIV-1 Drug Resistance Genotyping Assay for Indian Patients

Arpan Acharya¹, Salil Vaniawala², Parth Shah³, Rabindra Nath Misra⁴, Minal Wani¹, Pratap N. Mukhopadhyaya⁵*

¹Department of Molecular Biology, Dr. D. Y. Patil Biotechnology & Bioinformatics Institute, Dr. D. Y. Patil Vidyapeeth, Pune, Maharashtra, India, ²Department of Molecular Diagnostics, SN Genelab, Surat, Gujarat, India, ³Department of Molecular Diagnostics, Supratech Micropath Laboratory & Research Institute, Ahmedabad, Gujarat, India, ⁴Department of Microbiology, Padmashree Dr. D. Y. Patil Medical College & Research Centre, Dr. D. Y. Patil Vidyapeeth, Pune, Maharashtra, India, ⁵Department of Microbiology, Interdisciplinary Science, Technology and Research Academy, AISC, Pune, Maharashtra, India

Abstract

Human Immunodeficiency Virus-1 (HIV-1) drug resistance genotyping assay is a part of clinical management of HIV-1 positive individuals under treatment with highly active antiretroviral therapy (HAART). Routine monitoring of drug resistance mutations in resource limited settings like India is not possible due to high cost of commercial drug resistance assays. In this study we developed an in-house, cost effective HIV-1 drug resistance genotyping assay for Indian patients and validated it against the US-FDA-approved ViroSeq HIV-1 drug resistance testing system. A reference panel of 20 clinical samples was used to develop and validate the assay against ViroSeq HIV-1 drug resistance testing system which was subsequently used to genotype a clinical panel of 225 samples. The Stanford HIV database was used to identify drug resistant mutations. The analytical sensitivity of the assay was 1000 HIV-1 RNA copies/ml of plasma sample while precision and reproducibility was 99.68±0.16% and 99.76±0.18% respectively. One hundred and one drug resistant mutations were detected by the in-house assay compared to 104 by ViroSeq system in the reference panel. The assay had 91.55% success rate in genotyping the clinical panel samples and was able to detect drug resistant mutations related to nucleoside reverse transcriptase inhibitor (NRTI), non-nucleoside reverse-transcriptase inhibitor (NNRTI) as well as protease inhibitor (PI) classes of antiretroviral drugs. It was found to be around 71.9% more cost effective compared to ViroSeq genotyping system. This evaluation of the assay on the clinical panel demonstrates its potential for monitoring clinical HIV-1 drug resistance mutations and population-based surveillance in resource limited settings like India.

Introduction

Introduction of antiretroviral therapy (ART) all over the world has dramatically improved the health condition of HIV-1 infected individuals [1]. This is reflected by way of sharp fall in mortality and morbidity associated with HIV-1/AIDS [2]. Developing countries like India with heavy burden of HIV-1 population has also seen a gradual increase in use of ART and at present 0.6 million individuals are taking 1st line ART [3]. In order to sustain the success of this ART program it is important to monitor the patients in regular interval and keep switching the treatment regimen [4] as per need of the patient. As HIV-1 drug resistance is one of the formidable causes for treatment failure and choice for drugs are limited, there is need for a cost effective HIV-1 drug resistance monitoring system in resource limited settings like India [5].

In order to monitor progression of disease in patients undergoing ART, a number of tools such as point-of-care CD4 testing, HIV-1 viral load testing and HIV-1 drug resistance analysis using plasma/dried blood spots are being evaluated for their suitability in resource limited setting in terms of cost and efficiency [6–9]. Among these, HIV-1 drug resistance genotyping assay is of special relevance because of its use in monitoring transmitted drug resistance mutations in drug naïve individuals [10]. Information regarding emergence and prevalence of HIV-1 drug resistance mutations is important in guiding the national strategy for implementing ART program in resource limited settings with limited treatment options [11].
The commercial kits available for HIV-1 drug resistance genotyping are expensive and the prohibitive running costs hinder their routine use in resource limited countries like India [12]. These assays are designed specifically for HIV-1 subtype B strains prevalent in North America and Europe. Hence, genetic diversity of non B subtype HIV-1 strains prevalent in Africa and Asia pacific region poses challenge for use of these kits in such geographical regions [13]. This has led to continuous efforts to develop homebrew HIV-1 drug resistance genotyping assays across the world with varying success [14–16]. India is dominated by HIV-1 subtype C virus [17]. Recently we have demonstrated a definite geographical bias of nucleotide sequence motifs within specific regions of HIV-1 genome [18] that indicate the need for not only cost effective but also a region-specific, high quality HIV-1 drug resistance assay for periodic monitoring of patients undergoing ART. 

In this study we report a cost effective HIV-1 drug resistance genotyping assay for Indian population that is validated against US-FDA approved ViroSeq Genotyping System 2.0 (Celera Diagnostics, USA) and demonstrate its efficiency in identifying all drug resistant mutations related to NRTI, NNRTI and PI classes of antiretroviral drugs.

### Materials and Methods

#### Study Population

Plasma samples collected from 20 treatment experienced HIV-1 positive patients from India and those experiencing virologic failure were used as reference panel in this study. The clinical characteristics of this resource population are described in Table 1. The clinical panel used in this study comprised of 225 samples. The patients were enrolled during the period of May 2012 and September 2013 from India and all were under first-line of antiretroviral therapy for more than 6 months and failing the highly active anti retroviral therapy [HAART] as per current HIV treatment guidelines followed in India [19]. After developing the in-house genotyping assay described here, this clinical panel was used to study the prevalence of HIV-1 drug resistance mutations in India. The details of demographic and clinical characteristics of these 225 samples are summarized in Table 2. The study was duly approved by SN Gene laboratory clinical research, institutional bio-safety and bio-ethics committee (Approval number DGL/2012/WY07). Written informed consent was obtained from all participants enrolled in this program.

#### Specimen Collection and storage

Ten ml of blood was collected from each panel members in K2 EDTA vacutainer tubes (Becton Dickinson, San Diego, California, USA). Out of this, 2–3 ml was used for CD4+ T cell counting and the remaining for plasma separation. Plasma samples were stored in 1 ml aliquots at −20°C till further use.

#### HIV-1 Viral Load and CD4+ T cell count estimation

The viral load of reference and clinical panel samples were determined using artus HIV-1 RG RT-PCR kit (Qiagen, Germany) while CD4/CD8+ T cell counts were estimated using a FACS CALIBUR flow cytometer (BD Biosciences, California, USA), both according to respective manufacturer’s instructions.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Plasma Viral Load (HIV-1 RNA copies/ml)</th>
<th>ART Regimen</th>
<th>HIV-1 subtype</th>
<th>Plasma genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-IN-638</td>
<td>1680600</td>
<td>AZT+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-642</td>
<td>2576440</td>
<td>AZT+3TC+NVP</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-648</td>
<td>6598655</td>
<td>AZT+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-652</td>
<td>1154880</td>
<td>TDF+3TC+NVP</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-655</td>
<td>2237655</td>
<td>ATV+r/TDF+3TC</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-661</td>
<td>875600</td>
<td>TDF+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-662</td>
<td>165895</td>
<td>AZT+3TC+NVP</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-665</td>
<td>436278</td>
<td>TDF+3TC+NVP</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-669</td>
<td>278354</td>
<td>AZT+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-670</td>
<td>122400</td>
<td>AZT+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-672</td>
<td>24388</td>
<td>TDF+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-675</td>
<td>46544</td>
<td>ATV+r/TDF+3TC</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-678</td>
<td>75390</td>
<td>TDF+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-679</td>
<td>33175</td>
<td>TDF+3TC+NVP</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-683</td>
<td>12540</td>
<td>AZT+3TC+NVP</td>
<td>A</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-685</td>
<td>6755</td>
<td>AZT+3TC+NVP</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-688</td>
<td>4655</td>
<td>ATV+r/TDF+3TC</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-690</td>
<td>2690</td>
<td>AZT+3TC+EFV</td>
<td>C</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-694</td>
<td>2100</td>
<td>AZT+3TC+EFV</td>
<td>A</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-695</td>
<td>1350</td>
<td>AZT+3TC+NVP</td>
<td>C</td>
<td>+</td>
</tr>
</tbody>
</table>

*’+’: Positive amplification; ART: Anti retroviral therapy; AZT: Zidovudine; 3TC: Lamivudine; EFV: Efavirenz; NVP: Nevirapine; TDF: Tenofovir; ATV/r: Atazanavir/r.*

doi:10.1371/journal.pone.0105790.t001
ViroSeq genotyping system

The HIV-1 drug resistance genotyping of reference panel samples were carried out using US-FDA approved ViroSeq genotyping system according to manufacturer’s instructions. ViroSeq HIV-1 genotyping system software v2.6 and Stanford HIVDB [20] were used for drug resistance interpretation.

RNA Extraction and In-house HIV-1 Drug Resistance Genotyping

The drug resistance genotyping analysis of reference as well as clinical panel samples were carried out according to the method described as follows:

HIV-1 RNA was extracted from plasma samples stored at −20°C within 7 days of collection using a QIAamp Viral RNA mini kit (Qiagen, Germany) and subjected to one step RT PCR. This was followed by nested PCR to generate a 1614 bp amplicon that covered the entire protease gene and more than 300 initial amino acids of reverse transcriptase gene. All primers described in this study were designed using HIV-1 pol gene sequences reported from India and available from NCBI GenBank.

Briefly, the extracted RNA samples were reverse transcribed and then amplified using SuperScript III One-Step RT-PCR System (Life Technologies, Foster City, USA). Fifty µl of a reaction mixture comprised of 25 µl of 2X reaction mix, 2 µl of enzyme mix (SuperScript III RT and Platinum Taq), 20 µl of RNA and each primers (forward and reverse) at a final concentration of 10 pmoles per reaction respectively. The primer sequences used in the reaction were: 5′- GCTGTTGGAAATGTGGAA–3′ (forward) and 5′- TGGCTTGCCAATAGTCTGT–3′ (reverse). The thermal cycling profile comprised of 60 minutes of reverse transcription at 45°C followed by 5 minutes of heating at 95°C and 35 subsequent cycles of PCR amplification, each comprising of 95°C for 30 seconds, 59°C for 45 seconds, and 72°C for 180 seconds followed by a 10 minutes final extension at 72°C. For nested PCR, 5 µl of the 1st round PCR product was used as template. The reaction mixture comprised of 25 µl 2X PCR Mix v.2.0 (TaKaRa-bio, CA, USA), 10 pmoles of each primer (forward and reverse) and nuclease free water to make the volume to 50 µl. The thermal cycling profile comprised of an initial denaturation step of 5 minutes at 95°C followed by 35 cycles of PCR amplification, each comprising of 95°C for 30 seconds, 56°C for 45 seconds, and 72°C for 180 seconds followed by a 10 minutes final extension at 72°C.

For nested PCR, 5 µl of the 1st round PCR product was used as template. The reaction mixture comprised of 25 µl 2X PCR Mix v.2.0 (TaKaRa-bio, CA, USA), 10 pmoles of each primer (forward and reverse) and nuclease free water to make the volume to 50 µl. The thermal cycling profile comprised of an initial denaturation step of 5 minutes at 95°C followed by 35 cycles of PCR amplification, each comprising of 95°C for 30 seconds, 59°C for 45 seconds, and 72°C for 180 seconds followed by a 10 minutes final extension at 72°C. The primer sequences used for the nested PCR reaction were 5′- GTGGAAAGGAAGGACACCA–3′ (forward) and 5′- TGGTTTTACATCATTAGTGT–3′ (reverse). PCR products generated from nested PCR were run on a 1% agarose gel (Promega Corporation, Madison, USA), stained with ethidium bromide (0.5 µg/ml), visualized under a UV source (260 nm) and documented using an automated gel documentation system.

### Table 2. Demographic characteristics and laboratory results of the clinical panel.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Summary, n = 225</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yrs), median (IQR)</td>
<td>32 (26–41)</td>
</tr>
<tr>
<td>Gender, n (%):</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>121 (53.7%)</td>
</tr>
<tr>
<td>Female</td>
<td>90 (40.0%)</td>
</tr>
<tr>
<td>Child</td>
<td>14 (6.3%)</td>
</tr>
<tr>
<td>Median CD4T cell count, cells/µl (IQR)</td>
<td>137 (100–180)</td>
</tr>
<tr>
<td>Median Viral load, log10 copies/ml (IQR)</td>
<td>5.055 (4.39–5.47)</td>
</tr>
<tr>
<td>Risk exposure, n (%):</td>
<td></td>
</tr>
<tr>
<td>Heterosexual (%)</td>
<td>115 (51.2%)</td>
</tr>
<tr>
<td>Bisexual (%)</td>
<td>40 (17.7%)</td>
</tr>
<tr>
<td>MSM (%)</td>
<td>50 (22.2%)</td>
</tr>
<tr>
<td>MTC (%)</td>
<td>20 (8.9%)</td>
</tr>
<tr>
<td>Other co-infections, n (%):</td>
<td>48 (21.3%)</td>
</tr>
<tr>
<td>HIV-1 subtypes, n (%):</td>
<td></td>
</tr>
<tr>
<td>Subtype C</td>
<td>176 (85.5%)</td>
</tr>
<tr>
<td>Subtype A</td>
<td>29 (14.0%)</td>
</tr>
<tr>
<td>Subtype B</td>
<td>1 (0.5%)</td>
</tr>
<tr>
<td>Treatment regimen:</td>
<td></td>
</tr>
<tr>
<td>AZT, 3TC, EFV</td>
<td>35 (15.56%)</td>
</tr>
<tr>
<td>AZT, 3TC, NVP</td>
<td>50 (22.22%)</td>
</tr>
<tr>
<td>TDF, 3TC, NVP</td>
<td>45 (20.00%)</td>
</tr>
<tr>
<td>TDF, 3TC, EFV</td>
<td>35 (15.56%)</td>
</tr>
<tr>
<td>ATV/r, TDF, 3TC</td>
<td>35 (15.56%)</td>
</tr>
<tr>
<td>LPV/r, AZT, 3TC</td>
<td>25 (11.10%)</td>
</tr>
</tbody>
</table>

IQR: Interquartile range; MSM: Men who have Sex with Men; MTC: Mother to Child Transmission; AZT: Zidovudine; 3TC: Lamivudine; EFV: Efavirenz; NVP: Nevirapine; ATV/r: Atazanavir/r; LPV/r: Lopinavir/r.
doi:10.1371/journal.pone.0105790.t002
system (BIORAD, USA). For quality control, negative, low positive and high positive control samples were run with every batch of reactions.

PCR products generated from nested PCR were purified using a PureLink Quick PCR Purification Kit (Invitrogen, USA) and subjected to double strand DNA sequencing using 4 pairs of sequencing primers. The sequencing reactions were carried out using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) as per manufacturer’s instructions followed by capillary electrophoresis performed on an ABI PRISM 3500 Dx Genetic Analyzer (Applied Biosystems, USA). The nucleotide sequences of all oligonucleotide primers used to generate bidirectional sequence data, apart from those used for nested PCR amplification, were as follows: 5'- GTACAGTATTAG-3', 5'- ATATCAATATAATGTGC-3', 5'- ATGATATACAGAAGTGTCG-3', 5'- TACTGGTACAGTTT-

Table 3. Comparison of drug resistance mutations identified by ViroSeq genotyping system and the in-house assay.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Protease gene</td>
</tr>
<tr>
<td>HIV-IN-638</td>
<td>None</td>
</tr>
<tr>
<td>HIV-IN-642</td>
<td>None</td>
</tr>
<tr>
<td>HIV-IN-648</td>
<td>M46V, I54V, V82A</td>
</tr>
<tr>
<td>HIV-IN-652</td>
<td>L10F, I54V, V82F</td>
</tr>
<tr>
<td>HIV-IN-655</td>
<td><strong>L10I</strong></td>
</tr>
<tr>
<td>HIV-IN-661</td>
<td>A71T</td>
</tr>
<tr>
<td>HIV-IN-669</td>
<td>None</td>
</tr>
<tr>
<td>HIV-IN-675</td>
<td>I54V, A71V, N88D, L90M</td>
</tr>
<tr>
<td>HIV-IN-678</td>
<td>None</td>
</tr>
<tr>
<td>HIV-IN-688</td>
<td>None</td>
</tr>
<tr>
<td>HIV-IN-694</td>
<td>None</td>
</tr>
<tr>
<td>HIV-IN-695</td>
<td>I54V, V82A</td>
</tr>
</tbody>
</table>

The discordant mutations are shown in bold and underlined letters.

doi:10.1371/journal.pone.0105790.t003
CAATA–3’, 5’- TGTTTATCTAGGTATGGT–3’ and 5’-CTGGCAGGCTGATAGGCTGTA–3’.

The raw nucleotide sequence data generated were manually edited and assembled into a single contiguous sequence, archived and compared with standard HIV-1 reference strain sequence (HXB2) to obtain the nucleotide variation data. For determining HIV-1 subtype and obtaining HIV-1 drug resistance mutation profile, the edited nucleotide sequences were analyzed using Stanford HIVDB [20]. Single letter amino acid codes were used throughout the manuscript as per the standard IUPAC nomenclature.

Validation criteria

Validation of the assay methodology described in this study was performed according to WHO guidelines [21].

Accuracy

Accuracy of the assay was evaluated by analyzing the degree of concordance between drug resistance mutations identified by ViroSeq genotyping system and in-house assay using the reference panel as per IAS mutation list [22].

Sensitivity

The sensitivity of the assay was evaluated using 5 clinical samples taken from the reference panel. A dilution series was prepared for each sample with viral load of 100000, 10000, 5000, 1000 and 500 HIV-1 RNA copies/ml and tested in triplicate using the in-house assay protocol.

Precision and Reproducibility

Precision and reproducibility of the assay were evaluated using five clinical samples taken from the reference panel and tested in five replicates each. The degree of concordance of drug resistance associated mutations and nucleotide sequence identity was used to estimate the precision and reproducibility.

Phylogenetic analysis

Phylogenetic tree was constructed using Neighbor-Joining method [23] in MEGA 5.1 software [24] where the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown above the branches [25]. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in units of number of base substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1).

The reference panel tree had 129 sequences, which included 20 reference panel samples each tested with in-house genotyping assay (n = 20) and ViroSeq genotyping system (n = 20) respectively, 50 sequence data generated from 5 clinical samples, each tested in 5 replicates for precision data (n = 25) and reproducibility data (n = 25) respectively and 39 HIV-1 group M reference sequences obtained from HIV sequence database (http://www.hiv.lanl.gov/content/index) maintained by the Los Alamos National Laboratory, University of California, USA. The clinical panel tree comprised of 245 sequences which included 39 HIV-1 subtype reference sequences as described above and 206 clinical panel sequences.

Statistical Analysis

The clinical and biological parameter of study subjects in reference and clinical panels are presented in frequency (%) for categorical variables. For quantitative variables, data are presented in mean ± standard deviation (SD) or median [Interquartile range (IQR)]. Prevalence of drug resistance mutations were computed with 95% confidence interval (CI).

Nucleotide sequence accession numbers. The GenBank accession number of the sequences generated in this study is KJ185171–KJ185376.

Results

Assay design

The present assay is optimized on a nested RT-PCR based protocol to achieve maximum possible sensitivity. The PCR amplicon covered entire protease gene and 1st 300 amino acids of RT gene so as to include all major drug resistance mutations as per the IAS mutation list. The primers used in the study were designed using a database of HIV-1 pol gene sequences reported from India and archived at NCBI GenBank, USA.

Accuracy

The 20 reference panel samples were genotyped using the ViroSeq genotyping system as well as the in-house genotyping assay. The mean nucleotide and amino acid identity between the two tests were 99.21±0.58% and 99.65±0.43% respectively. A total 101 drug resistance mutations were detected by the in-house assay compared to 104 using the ViroSeq genotyping system. A comparative analysis of drug resistance mutations detected by both the methods is described in Table 3.

Sensitivity

The assay was optimized for amplification of plasma samples having 1000 HIV-1 RNA copies/ml and above. This limit of detection was established by testing a dilution series of 5 samples in triplicates. The assay result is described in Table 4. It was found

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Sensitivity at a dilution (HIV-1 RNA copies/ml) of:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100000</td>
</tr>
<tr>
<td>HIV-IN-652</td>
<td>+++</td>
</tr>
<tr>
<td>HIV-IN-655</td>
<td>+++</td>
</tr>
<tr>
<td>HIV-IN-665</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-669</td>
<td>+</td>
</tr>
<tr>
<td>HIV-IN-670</td>
<td>+</td>
</tr>
</tbody>
</table>

**Table 4.** Assay sensitivity results using a dilution series from 5 reference panel samples, each tested in triplicate.

<sup>+: Positive amplification; −: Negative amplification.</sup>

<sup>doi:10.1371/journal.pone.0105790.t004</sup>
that up to 1000 HIV-1 RNA copies/ml, all replicates of 5 clinical samples amplified successfully but at 500 HIV-1 RNA copies/ml the results were inconsistent.

Precision and Reproducibility
All five replicates of 5 clinical samples could be amplified and sequenced successfully. The mean nucleotide sequence identity for precision varied between 99.68±0.16% and 100% whereas the mean nucleotide sequence identity for reproducibility varied between 99.76±0.18% and 100%. The results are summarized in Table 5. No discordant drug resistance mutations were detected in the replicate data generated for precision and reproducibility.

The maximum likelihood tree constructed using Mega 5.1 confirmed absence of any sample mix-up or cross contamination and sequences generated from the same sample clustered together (Figure 1).

Clinical Panel result
After development of the in-house assay it was used for testing a clinical panel specially created for this purpose. The viral load and CD4+ T cell count of clinical panel samples are described in Table 2. Out of the 225 samples, 210 responded successfully to PCR amplification out of which 206 could be successfully sequenced and analyzed for HIV-1 drug resistance mutations. Among 110 samples with viral load of 10^4 HIV-1 RNA copies/ml of plasma and above, 107 could be successfully amplified and both the DNA strands sequenced. On the other hand, out of 90 samples with viral load between 10^3 and 10^4 HIV-1 RNA copies/ml of plasma, 82 could be successfully amplified but only 80 among them could be sequenced for both the strands. Among 13 samples with viral load in between 5×10^3 and 10^3 HIV-1 RNA copies/ml of plasma, 12 could be successfully amplified but double strand DNA sequence could be generated for 11 of them. Among 12 samples with viral load between 10^3 and 5×10^3 HIV-1 RNA copies/ml of plasma, 9 could be successfully amplified and double strand DNA sequence could be generated from 8 of them.

Out of 206 samples genotyped, 176 (85.5%), 29 (14.0%) and 1 (0.5%) were from patients infected with HIV-1 subtype C, subtype A and subtype B respectively (Figure 2). Samples from 28 (13.59%) patients did not show any mutations related to HIV-1 drug resistance and 178 (86.41%) of them had at least one HIV-1 drug resistance mutation(s). One hundred and fifty nine (77.18%) samples had at least one NRTI resistance mutation while 161 (78.16%) harbored at least one NNRTI mutation. Samples from 41 (19.90%) patients had at least one PI mutation(s).

All three classes of mutations were detected in samples from 29 (14.08%) patients. Among 147 patients on NRTI and NRTI-based 1st line ART regimen, 116 had mutations belonging to these 2 categories of drugs. M184V was the most common NRTI resistance mutation detected in 132 (64.08%) patients. 76 (36.89%) of the 213 patients harbored both PI major and minor mutations while 18 and 3 had only PI minor and only PI major mutations respectively. Eighteen samples did not have any PI-related mutations. The most common major PI mutations detected were M461L/I-14 (6.80%), I54A/T/V-16 (7.77%), V82A/C/G/F-16 (7.77%) and L90M-8 (3.88%) respectively. L10F/I/V-29 (14.08%) and A71T/V-14 (6.80%) were the two common PI minor mutations observed in the clinical panel. The details of other PI-related mutations are described in Figure 3C.

Comparative analysis of cost and hands-on time of in-house assay with ViroSeq genotyping system
Table 6 describes the hands-on time and cost of different stages of the analysis starting from collection of clinical sample to interpretation of drug resistance mutations. All costs are presented in US dollars. Cost of establishing the reference laboratory capable of performing the assay is however not included in the analysis. Further, some major costs common to both the assays such as logistics and manpower were also excluded. The hands-on times for the ViroSeq genotyping system and the in-house method were 18 hour 45 min and 17 hour 15 min respectively while the running cost of the in-house assay was computed at $85 compared to $303 for the ViroSeq genotyping system.

Discussion
The HIV-1 drug resistance genotyping assay is not feasible for routine monitoring of patients taking 1st line antiretroviral drugs in resource limited settings like India mainly due to high cost of commercial HIV-1 genotyping assays presently available in the market [26]. But increased access to antiretroviral drugs without proper monitoring results in transmission of drug resistant HIV-1 strains in newly infected individuals [27]. Laboratory methods to monitor the treatment outcome and proper guidelines regarding course of action in case of therapeutic failure is critical in management of HIV-1/AIDS. HIV-1 drug resistance genotyping assay for patients with virologic failure acts as a guiding tool during switching to next line of treatment [12]. We performed a cost analysis of the drug resistance genotyping assay described in this study and compared it with the running cost of ViroSeq genotyping system which indicated that our assay is around 71.9% cost effective compared to the later. This attribute make
<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Plasma Viral Load (HIV-1 RNA copies/ml)</th>
<th>HIV-1 Subtype</th>
<th>% Nucleotide sequence identity</th>
<th>Replicate Tests</th>
<th>Number of discordant mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-IN-638</td>
<td>1680600</td>
<td>C</td>
<td>100.00±0.00%</td>
<td>3 3 3 3 3 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>HIV-IN-661</td>
<td>875600</td>
<td>C</td>
<td>99.81±0.15%</td>
<td>4 4 4 4 4 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>HIV-IN-662</td>
<td>165895</td>
<td>C</td>
<td>99.96±0.07%</td>
<td>11 11 11 11 11 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>HIV-IN-694</td>
<td>2100</td>
<td>A</td>
<td>99.76±0.18%</td>
<td>0 0 0 0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>HIV-IN-695</td>
<td>1350</td>
<td>C</td>
<td>99.84±0.09%</td>
<td>4 4 4 4 4 0</td>
<td>0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

**Reproducibility:**

**Precision:**
this assay more suitable for routine monitoring of transmitted HIV-1 drug resistance strains as well as for detection of drug resistance mutations in patients with virologic failure.

The drug resistance mutations detected by our in-house assay exhibited excellent concordance when compared with corresponding results from the ViroSeq genotyping system. The assay was able to detect all clinically relevant mutations according to the IAS 2013 mutation list [22]. These findings demonstrate both utility and feasibility of this home brew assay in HIV-1 drug resistance surveillance and monitoring in resource limited settings like India.

None of the HIV-1 drug resistance genotyping assays including the US-FDA approved commercial assays as well as various home brew assays can successfully amplify 100% clinical samples mainly due to high genetic variability of HIV-1 [28] and occurrence of spontaneous mutation within primer binding regions of the viral genome [29]. In this backdrop, the home brew assay described in this study could successfully genotype 91% of samples from the clinical panel which was found to be satisfactory. This high rate of success is possibly due to the geographical region-specific primers designed for this assay coupled with incorporation of a nested PCR protocol.

The in-house assay described in this study was validated as per WHO guidelines for HIV-1 drug resistance genotyping and demonstrated a high degree of precision and reproducibility. The limit of detection of this assay was 1000 HIV-1 RNA copies/ml of plasma sample. This is in line with similar studies from India and other parts of the world [30–31]. The assay has ability to detect all major HIV-1 subtypes (HIV-1 subtype A, B and C) predominant in India [32] as revealed from the clinical panel genotyping results.

The main limitation of this study is the lack of subtype diversity in the reference panel. An ideal panel should comprise of all HIV-1 group M subtypes including circular recombinant forms which were not included here due to scarcity of such samples among HIV-1 sero-positive individuals in India. In spite of this limitation, due to the rigorous validation of assay parameters as per WHO guidelines there is enhanced confidence and reliability seen to be associated with our assay. Genotyping of the clinical panel in this study simulated real time field conditions and demonstrated good
performance in detecting all clinically relevant HIV-1 drug resistance mutations in the protease and reverse transcriptase genes. This result is also in line with observations made from other similar studies reporting patterns of HIV-1 drug resistance mutations in patients failing 1st line ART from India [33–37].

In conclusion, we report development and validation of a low cost HIV-1 drug resistance genotyping assay for resource limited settings like India with potential to serve the increasing demand of HIV-1 drug resistance genotyping assay for resource limited areas. J Chin Clin Med 2: 469–480.

Acknowledgments

The authors are grateful to all the study participants for agreeing to provide their valuable clinical samples for this study.

Author Contributions

Conceived and designed the experiments: AA PNM RNM MW. Performed the experiments: AA PS. Analyzed the data: AA PNM SV. Contributed reagents/materials/analysis tools: PNM SV. Contributed to the writing of the manuscript: AA PNM RNM MW.

References


Table 6. Hands - on time and cost comparison between ViroSeq and the in-house HIV-1 genotyping assay.

<table>
<thead>
<tr>
<th>Process</th>
<th>Step(s)</th>
<th>Assay type</th>
<th>ViroSeq</th>
<th>In-house</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Hands-on time</td>
<td>cost/test ($)</td>
<td>Hands-on time</td>
</tr>
<tr>
<td>Sample preparation</td>
<td>Sample collection and plasma separation</td>
<td>45 min</td>
<td>1.0</td>
<td>45 min</td>
</tr>
<tr>
<td>Nucleic acid extraction</td>
<td>RNA Extraction</td>
<td>3 h</td>
<td>250</td>
<td>1 h</td>
</tr>
<tr>
<td>Amplification</td>
<td>One step RT-PCR</td>
<td>5 h 30 min</td>
<td>15.0</td>
<td>3 h 30 min</td>
</tr>
<tr>
<td></td>
<td>Nested PCR</td>
<td></td>
<td></td>
<td>2 h 30 min</td>
</tr>
<tr>
<td>Gel Documentation</td>
<td>Agarose gel electrophoresis</td>
<td>45 min</td>
<td>2.0</td>
<td>45 min</td>
</tr>
<tr>
<td>Genotyping</td>
<td>Amplicon purification</td>
<td>1 h</td>
<td>1 h</td>
<td>1 h</td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>2 h 30 min</td>
<td>30.0</td>
<td>2 h 30 min</td>
</tr>
<tr>
<td></td>
<td>Sequence amplicon purification</td>
<td>1 h 30 min</td>
<td>10.0</td>
<td>1 h 30 min</td>
</tr>
<tr>
<td></td>
<td>Sequencing sample run</td>
<td>2 h 30 min</td>
<td>10.0</td>
<td>2 h 30 min</td>
</tr>
<tr>
<td>Sequence analysis</td>
<td>Sequence data validation</td>
<td>30 min</td>
<td>30 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequence assembly</td>
<td>15 min</td>
<td>30 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interpretation and quality analysis</td>
<td>30 min</td>
<td>30 min</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>18 h 45 min</td>
<td>303.0</td>
<td>17 h 15 min</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0105790.t006


