Original Article

Distinctive Drug-resistant Mutation Profiles and Interpretations of HIV-1 Proviral DNA Revealed by Deep Sequencing in Reverse Transcriptase[*]

YIN Qian Qian[1], LI Zhen Peng[1,2], ZHAO Hai[1], PAN Dong[1], WANG Yan[1], XU Wei Si[1],

XING Hui[1], FENG Yi[1], JIANG Shi Bo[3], SHAO Yi Ming[1], and MA Li Ying[1,#]

1. State Key Laboratory of Infectious Disease Prevention and Control, National Center for AIDS/STD Control and Prevention (NCAIDS), Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Chinese Center for Disease Control and Prevention, Beijing 102206, China; 2. State Key Laboratory of Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Chinese Center for Disease Control and Prevention, Beijing 102206, China; 3. Key Laboratory of Medical Molecular Virology (Ministries of Education and Health), Shanghai Medical College and Institute of Medical Microbiology, Fudan University, Shanghai 200433, China

Abstract

Objective   To investigate distinctive features in drug-resistant mutations (DRMs) and interpretations for reverse transcriptase inhibitors (RTIs) between proviral DNA and paired viral RNA in HIV-1-infected patients.

Methods   Forty-three HIV-1-infected individuals receiving first-line antiretroviral therapy were recruited to participate in a multicenter AIDS Cohort Study in Anhui and Henan Provinces in China in 2004. Drug resistance genotyping was performed by bulk sequencing and deep sequencing on the plasma and whole blood of 77 samples, respectively. Drug-resistance interpretation was compared between viral RNA and paired proviral DNA.

Results   Compared with bulk sequencing, deep sequencing could detect more DRMs and samples with DRMs in both viral RNA and proviral DNA. The mutations M184I and M230I were more prevalent in proviral DNA than in viral RNA (Fisher's exact test, $P<0.05$). Considering 'majority resistant variants', 15 samples (19.48%) showed differences in drug resistance interpretation between viral RNA and proviral DNA, and 5 of these samples with different DRMs between proviral DNA and paired viral RNA showed a higher level of drug resistance to the first-line drugs. Considering 'minority resistant variants', 22 samples (28.57%) were associated with a higher level of drug resistance to the tested RTIs for proviral DNA when compared with paired viral RNA.

Conclusion   Compared with viral RNA, the distinctive information of DRMs and drug resistance interpretations for proviral DNA could be obtained by deep sequencing, which could provide more detailed and precise information for drug resistance monitoring and the rational design of optimal antiretroviral therapy regimens.

Key words: HIV-1 drug-resistant mutation; Drug-resistance interpretation; Proviral DNA; Viral RNA; Deep sequencing

INTRODUCTION

Highly active antiretroviral therapy (HAART) has dramatically decreased HIV-related morbidity and mortality, resulting in a longer life expectancy for HIV-infected individuals[1-2]. Currently, HIV is considered a controlled chronic disease. Nevertheless, the presence of HIV drug-resistant mutations (DRMs) is a major barrier to the success of long-term treatment[3-4]. Therefore, HIV drug resistance genotyping, a clinical tool for the detection of viral resistance, is an essential complement to HAART[5]. Drug resistance genotyping is routinely recommended, as it could maximize the benefit of current treatment regimens through identifying drug resistance, facilitating the choice of optimal antiretroviral regimens based on the presence of newly discovered DRMs[6].

The current standard for assessing drug susceptibility relies on the detection of HIV-1 DRMs in plasma viral RNA. Theoretically, DRMs detected in viral RNA are also present in proviral DNA in HIV-infected cells. However, some studies have observed different DRMs between the two sources[7-8]. Some DRMs detected in viral RNA are neither archived nor detectable in proviral DNA. However, drug genotyping based on plasma samples may underestimate the level of drug resistance in the absence of strong drug selective pressure. Several studies have stressed the importance of proviral DNA as an additional source of information for determining the total burden of resistance in an individual[9-10]. Therefore, the use of proviral DNA in combination with viral RNA sequencing could increase the sensitivity of drug resistance testing. Moreover, proviral DNA was suggested to serve as an alternative source to viral RNA present at an undetectable level, or if plasma samples are not available[11].

Therefore, in this study, DRMs detected in viral RNA and paired proviral DNA were analyzed using deep-sequencing technologies to demonstrate differences and distinctive features in the resistance profile and resistance interpretations between viral RNA and paired proviral DNA. Such analysis can provide further evidence of the role that proviral DNA genotyping plays in HIV drug resistance testing.

METHODS

Study Population

A total of 43 patients chronically infected with HIV-1 were enrolled in this study. Starting in 2004, these patients participated in a multicenter AIDS Cohort Study in Anhui and Henan Provinces in China and received a first-line regimen of stavudine (d4T) or zidovudine (AZT), lamivudine (3TC), and nevirapine (NVP) or efavirenz (EFV). Seventy-seven blood samples were collected; some samples were collected at different times from the same patient. Plasma HIV-1 RNA and CD4+ T cell counts were determined after sample collection. This study was approved by the Institutional Research Ethics Community of the Chinese Center for Disease Control and Prevention (IRB00002276), and all subjects provided written informed consent before blood samples were drawn and data collected.

Bulk Sequencing

HIV-1 proviral DNA was extracted from 200 µL of frozen whole blood using the QIAamp DNA blood mini kit (Qiagen, Germany). Viral RNA was extracted from 140 µL samples of plasma using QIAamp viral RNA mini kit (Qiagen, Germany). For bulk PCR, the HIV-1 *pol* gene was amplified using an in-house-designed nested PCR system from plasma-derived viral cDNA and proviral DNA, as previously described[12]. Drug resistance genotyping was analyzed according to the Stanford HIV Drug Resistance Database (HIVdb)[13] (http://hivdb.stan-ford.edu).

Deep Sequencing

Viral RNA and paired proviral DNA samples were used to perform next-generation sequencing. Library construction was achieved according to the instructions of the Illumina DNA sample preparation manual. As mentioned above, the first-round nested PCR was performed using an in-house-designed PCR system. Specific primers (F1: 5'-GCCAGGAATGGATGGCCCAAAAG-3', HXB2: 2588-2610, R1: 5'-TGGTGATCCTTTCCATCCCTGTGG-3', HXB2: 3020-2997; F2: 5'-AGTATACTGCATTTACCATACC-3', HXB2: 2926-2947, R2: 5'-CTTCTGTATGTCATTGACAGTCC-3', HXB2: 3326-3304) were used to amplify two partly overlapping amplicons covering the HIV-1 pol gene from the first-round nested PCR products, including the reverse transcriptase gene codons 14-258. The barcodes were introduced to specific primers to distinguish different samples. Then, the PCR products were purified using Agencourt AMPure XP magnetic beads (Beckman, USA), followed by quantification using the Qubit dsDNA HS Assay Kit (Invitrogen, USA). After mixing and pooling of amplicons, quantification was verified using an Agilent 2100 Bioanalyzer (Agilent, USA). The

Illumina-specific P5 and P7 primer and adaptor were introduced to the pooling product with end-repair by PCR. PhiX DNA was added to the final multiplexed libraries to increase base complexity for improving cluster detection. Sequencing was performed using the MiSeq Reagent Kit v2 (Illumina, USA).

Data Analysis

Deep-sequencing data were analyzed by computer programs, as previously described[14]. Poor quality reads, short reads, and reads with ambiguous bases (Ns) were removed at the first step. The filtered sequences were then partitioned according to the barcode specific to each sample, and the barcodes were subsequently discarded. Consensus sequences were generated from the overlapping forward and reverse sequences. Consensus sequences were aligned to the reference *pol* gene of HXB2 using the bowtie program[15]. After file format conversion, in which the SAM file was converted to a BAM file, sorting, indexing, and variant calling were performed using the Genome Analysis Toolkit[16] (https://www.broadinstitute.org/gatk/index.php). The median coverage per nucleotide position for the two fragments was 1047 and 2272 for proviral DNA and 1451 and 5792 for viral RNA. We selected a 3% threshold for mutation detection to further increase the validity of the results[17].

In this study, variants with DRMs present at a proportion of >20% were considered as majority resistant variants, whereas variants with DRMs present at proportions <20% and >3% were considered as minority resistant variants. Viruses exhibiting either possible resistance or full resistance to a drug were considered to be resistant. DRMs were identified using the most recent mutation list[18], and drug resistance interpretations were classified as high-level resistance, intermediate-level resistance, low-level resistance, and potential low-level resistance for the drugs of the first-line regimen, including AZT or D4T, 3TC, and NVP or EFV, according to the HIVdb.

RESULTS

Characteristics of the Study Population

The subjects' characteristics are given in Supplemental Table S1 (end of this article). This patient group included 16 males and 27 females with a median age of 38 years (range 26-64 years). The median log10 plasma RNA load was 5.10 copies/mL (range 3.21-6.98). The median CD4+ cell count was 217 cells/µL (range 7-515). Twenty-four (55.81%) patients received a treatment regimen of AZT, NVP, and 3TC. Seventeen (39.53%) patients received a treatment regimen of D4T, NVP, and 3TC. Forty-two (97.67%) patients were infected with HIV-1 through plasma donation.

Differential Drug Resistance Analysis in Viral RNA and Paired Proviral DNA Based on Bulk Sequencing

Using bulk sequencing, approximately 50 DRMs were detected in 26 (33.77%) samples for viral RNA and in 22 (28.57%) samples for proviral DNA. After comparison, we found different distributions of DRMs in viral RNA and proviral DNA (Figure 1A), although there was no statistically significant difference in the number of samples harboring DRMs between viral RNA and proviral DNA (χ^2 test, P=0.49). Overall, analysis of resistance by drug class revealed different drug resistance interpretations between viral RNA and proviral DNA (except for sample 3) among 12 samples (Table 1). In particular, three samples had DRMs (K103N and G190A/E) causing high-level resistance to NVP and EFV, which were detected only in proviral DNA.

DRMs Detection in Viral RNA and Paired Proviral DNA Based on Deep Sequencing

As mentioned above, deep sequencing can generally detect and quantify DRMs at a relatively low frequency. Therefore, to further identify the differences in mutation variation between the viral RNA and paired proviral DNA, we performed deep sequencing. Approximately 80 DRMs were detected in 42 samples for viral RNA and in 40 samples for proviral DNA. As expected, the total number of samples harboring DRMs identified by deep sequencing was significantly higher than that identified by bulk sequencing (χ^2 test, P<0.01). The distributions of DRMs from viral RNA and proviral DNA are shown in Figure 1B. Nineteen mutations at 17 resistance-associated positions and 24 mutations at 18 resistance-associated positions were observed in viral RNA and proviral DNA, respectively. The mutations K65R, L74V, E138A, M184I, G190E, K219R, and M230I were detected only in proviral DNA. The mutations M184I and M230I were more prevalent in proviral DNA than in viral RNA (Fisher's exact test, P<0.05). Among the majority resistant variants (Figure 1Ba), five major mutations, M41L, M184I, E138A, G190E, and M230I, were detected only in proviral DNA. Among the minority resistant variants (Figure 1Bb),
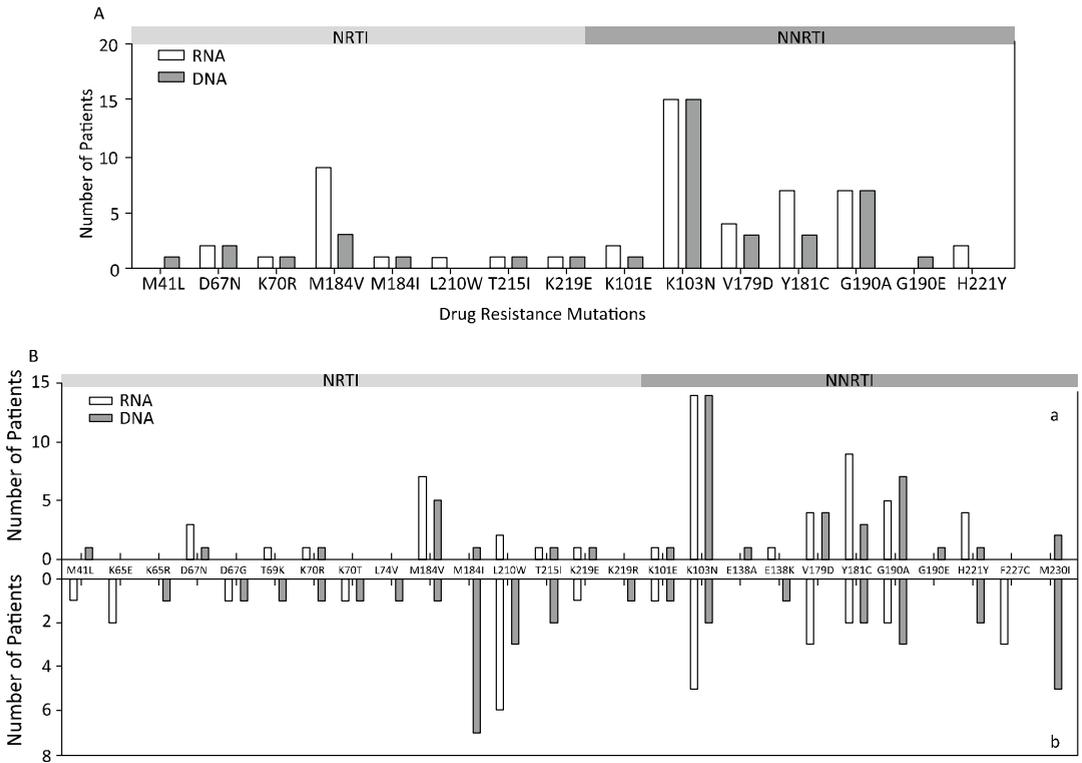
Figure 1. Distributions of DRMs in viral RNA and proviral DNA determined by bulk sequencing (A) and deep sequencing (B) among 77 samples. DRMs were detected from majority resistant variants (>20%) in viral RNA and proviral DNA (Ba). DRMs were detected in minority resistant HIV-1 variants (3%-20%) in viral RNA and proviral DNA (Bb). Bar charts show the number of patients in which specific mutations occurred in viral RNA and proviral DNA.

Table 1. Different Distribution of DRMs in Viral RNA and Proviral DNA Detected by Bulk Sequencing and Related Drug-resistant Interpretations

PID	CD4+ Cells/μL	VL Copies/mL	Treatment	Viral RNA		Proviral DNA	
				Mutations	Drug-resistance Interpretation[*]	Mutations	Drug-resistance Interpretation[*]
2	154	173,700	naive			K103N	H: NVP, EFV
3	219	16,750	naive	K101E, G190A	H: NVP, EFV	K103N, G190A	H: NVP, EFV
32	112	361,300	naive	G190A	H: NVP I: EFV		
35	382	61,100	naive			K103N	H: NVP, EFV
44	623	16,200	AZT+**NVP**+3TC	M184V,Y181C	H: **NVP**, **3TC** I: EFV		
47	287	626,000	AZT+**EFV**+3TC			G190E	H: NVP, **EFV**
59	195	5,470	**AZT**+**NVP**+**3TC**	M184V, K103N	H: **3TC**, **NVP**, EFV L: **AZT**, D4T	K103N	H: **NVP**, EFV
63	337	19,400	D4T+**EFV**+3TC	K103N	H: NVP, **EFV**		
65	213	1,060	AZT+**NVP**+3TC	G190A	H: **NVP**, EFV		
66	286	11,000	NVP+D4T+**3TC**	M184V	H: **3TC**		
67	250	1,900	AZT+**NVP**+3TC	M184V, K103N	H: **3TC**, **NVP**, EFV		
73	416	1,370	D4T+**NVP**+3TC	M184V, K103N Y181C, H221Y	H: **NVP**, EFV, **3TC**	K103N, G190A	H: **NVP**, EFV

Note. [*]Drug resistance interpretations were evaluated for first-line drugs using the HIV Drug Resistance Database (HIVdb). Drugs in bold are the self-reported drugs taken by study subjects. H=high level drug-resistance, I=intermediate drug-resistance, L=low level drug-resistance, PL: potential low level drug-resistance.

K65R, D67G, T69N, K70R, L74V, M184V/I, T215I, K219R, E138K, H221Y, and M230I were detected only in proviral DNA.

Comparison of Drug-resistance Interpretations in Viral RNA and Paired Proviral DNA by Deep Sequencing

As described above, high-frequency resistance-conferring mutations were classified as 'majority resistant variants', whereas low-frequency resistance-conferring mutations were termed 'minority resistant variants'. Therefore, based on the different distribution of DRMs, we compared the genotypic susceptibility of the identified DRMs to the first-line reverse transcriptase inhibitors (RTIs) tested in this study.

Considering only the majority resistant variants, 16 (20.78%) samples showed identical drug resistance interpretation according to drug class between viral RNA and paired proviral DNA. Fifteen samples (19.48%) showed differences in drug resistance interpretation between viral RNA and proviral DNA (Table 2), noting that the DRMs identified in all samples correlated well with the treatment regimens reported. Except for seven samples (9.09%) with DRMs only detected in viral RNA, 5 samples (6.49%) in which the DRMs detected in proviral DNA differed from those detected in viral RNA showed a higher level of drug resistance to one or more of the first-line drugs tested, including 3 samples (3.90%) with DRMs only detected in proviral DNA.

Considering the minority resistant variants, only 3 samples (3.9%) showed identical drug resistance according to drug class between viral RNA and paired proviral DNA. Further analysis of resistance by drug class revealed different interpretations between viral RNA and proviral DNA among 35 samples (45.45%), as shown in Table 3. There were 12 samples (15.58%) with DRMs only detected in viral RNA, whereas there were 22 samples (28.57%) with DRMs detected only in proviral DNA. These samples were associated with a higher level of drug resistance to one or more of the tested RTIs when compared to viral RNA, including 16 samples that had different DRMs present only in proviral DNA.

Altogether, DRM-associated drug-resistance interpretations differed when considering viral RNA and proviral DNA.

DISSCUSSION

Antiretroviral therapy can effectively reduce viral replication and thus limits the transmission of HIV-1. Moreover, antiretroviral therapy is not only a treatment method but could also serve as a new approach to preventing HIV infection[19-21]. In China, a government initiative was undertaken to control HIV transmission between spouses through full-scale antiretroviral therapy[22]. Nevertheless, drug mutations remain a major obstacle to both HIV prevention and treatment, and can influence the efficacy of the continued expansion of this nationwide antiretroviral therapy strategy.

HIV drug resistance genotyping is an important clinical tool and an essential complement to HAART. At present, HIV-1 drug resistance genotyping is still routinely performed from plasma samples. However, increasing numbers of studies are demonstrating the value of proviral DNA as an additional source of drug resistance data. DRMs archived in proviral DNA may be expressed upon virological failure, subsequently transmitting the drug-resistant virus and, hence, affecting long-term antiretroviral therapy.

Bulk sequencing is the current standard methodology with which to analyze HIV drug resistance, but is limited to the detection of variants with DRMs present at a frequency of over 20%. In contrast, deep-sequencing techniques can improve the sensitivity of drug resistance genotyping through detecting drug-resistant variants present at a frequency between 1% and 20%[8,23-24]. Moreover, the use of deep sequencing can reduce the cost when multiplexing a large number of samples. Therefore, we compared the distribution of DRMs across the HIV-1 reverse transcriptase coding region from viral RNA and corresponding proviral DNA to study differences in drug resistance profiles to the common first-line RTIs.

Using bulk sequencing, differences in the distribution of DRMs in viral RNA and proviral DNA were found in 12 out of 77 samples, resulting in very different resistance profiles. Interestingly, the majority DRMs, as defined above, were detected only in proviral DNA in association with high-level resistance to one or more of the tested RTIs without detection of DRMs in viral RNA. Other studies have revealed corroborative results for mutations detected in proviral DNA compared to viral RNA[7,25-26], indicating the need to further investigate the deep sequencing of DRMs associated with proviral DNA.

Indeed, the use of deep-sequencing technologies further demonstrated that DRMs from viral RNA and paired proviral DNA varied in frequency as well as position. More mutations at

Table 2. Different Distribution of DRMs in Viral RNA and Proviral DNA by Deep Sequencing (mutation frequency: >20%) and Related Drug-resistant interpretationsl

PID	CD4+ Cells/µL	VL Copies/mL	Treatment	Viral RNA Mutation (freq. %) NRTI	Viral RNA Mutation (freq. %) NNRTI	Viral RNA Drug-resistance Interpretation*	Proviral DNA Mutation (freq. %) NRTI*	Proviral DNA Mutation (freq. %) NRTI	Proviral DNA Drug Resistance Interpretation*
16	152	620,100	naive		E138K (99.2)	PL: NVP, EFV			
35	382	61,100	naive						H: NVP, EFV
44	624	16,200	AZT+**NVP**+**3TC**	T69N (40.8), M184V (99.8)	Y181C (99.7)	H: **3TC, NVP,** EFV	M184V (24.7)	K103N (97.2)	H: **3TC**
46	132	16,100	NVP+**D4T**+3TC	D67N (74.6)	Y181C (97.7), H221Y (93.8)	H: **NVP** I: EFV L: AZT, **D4T**			
47	287	626,000	AZT+**EFV**+3TC					G190E (93.6)	H: NVP, **EFV**
50	151	196,000	AZT+**NVP**+**3TC**	D67N (94.2), K70R (98.3), M184V (98.3), T215I (32.8), K219E (97.2)	G190A (98.3)	H: **NVP, 3TC,** **AZT,** D4T I: EFV	D67N (93), K70R (85.8), M184V (98.2), T215I (40.7), K219E (86.9)	K103N (57.7), G190A (98.2)	H: **NVP, 3TC** **AZT,** D4T, EFV
51	323	172,000	D4T+**EFV**+3TC	L210W (97.5)	K103N (97.5)	H: NVP, **EFV** L:AZT, **D4T**		K103N (87.5)	H: NVP, **EFV**
53	144	533,000	AZT+**NVP**+**3TC**				M184V (28.1)	M230I (28)	H: **3TC** I: **NVP** L: EFV
57	317	53,596	NVP+**D4T**+**3TC**	D67N (21.5), M184V (98.6), L210W (33.1)	K103N (98.8), Y181C (45.3)	H: **3TC,** EFV, **NVP** L:AZT, **D4T**	M41L (94.1), M184V (94)	K101E (64.7), Y181C (36.2), G190A (58)	H:**3TC,** EFV, **NVP**
59	195	5,470	AZT+**NVP**+3TC		K103N (96.2)	H: **NVP,** EFV	M184I (28.1)	M230I (29.9)	H: **3TC** I: **NVP** L: EFV
63	337	19,400	D4T+**EFV**+3TC		K103N (24.2)	H: NVP, **EFV**			
64	535	4,280	AZT+**NVP**+3TC		Y181C (99.2)	H: **NVP** I: EFV			
65	213	1,060	AZT+**NVP**+3TC		K103N (33.9), Y181C (52.4), G190A (33.6)	H:**NVP,** EFV			
66	286	11,000	NVP+D4T+**3TC**	M184V (98.7)	Y181C (98.9), H221Y (94.3)	H: **3TC, NVP** I: EFV			
67	250	1,900	AZT+**NVP**+**3TC**	M184V (98)	K103N (97.6)	H: **3TC,** EFV, **NVP**			

Note. *Drug resistance interpretations were evaluated for first-line drugs using the HIV Drug Resistance Database (HIVdb). Drugs in bold are the self-reported drugs taken by study subjects. H=high level drug-resistance, I=intermediate drug-resistance, L=low level drug-resistance, PL: potential low level drug-resistance.

Table 3. Different Distribution of DRMs in Viral RNA and Proviral DNA by Deep Sequencing (mutation frequency: 3%-20%) and Related Drug-resistant Interpretations

PID	CD4+ Cells/μL	VL Copies/mL	Treatment	Viral RNA		Proviral DNA	
				Mutation (Freq .%)	Drug-resistance Interpretation[*]	Mutation (Freq .%)	Drug-resistance Interpretation[*]
3	219	16,750	naive			K101E (19.9)	I: NVP L: EFV
4	459	127,600	naive			K70R (10)	I: AZT L: D4T
5	271.5	175,100	naive	L210W (4.4)	L: D4T, AZT		
8	83	552,400	naive	K103N (13.6) Y181C (9.4)	H: NVP, EFV		
10	423	116,700	naive			M184I (16.8) M230I (15.6)	H: 3TC I: NVP L: EFV
14	200	200,000	naive			L210W (11.8)	L: D4T, AZT
15	298	170,000	naive	V179D (12.6) K70T (6.1)	P: 3TC, EFV NVP, D4T		
18	239.55	48,000	naive	V179D (14.7)	P: NVP, EFV		
25	7	2808,000	naive	V179D (6.9)	P: NVP, EFV		
26	125	1611,000	naive			T215I (9.8)	L: AZT, D4T
27	74	84,650	naive	K65E (4)	PL: D4T		
28	15	4429,000	naive			D67G (4.7)	PL: AZT, D4T
30	136	371,000	naive	L210W (3) F227C (3.1)	I: EFV, NVP L: D4T, AZT	M184I (4.2) M230I (3.6)	H: 3TC, NVP I: EFV
32	112	361,300	naive	K103N (3) Y181C (7.8)	H: NVP, EFV	Y181C (18.1)	H: NVP I: EFV
35	382	61,100	naive	K101E (10.2) K103N (12.3)	H: NVP, EFV		
37	164	236,500	naive	K65E (3.4)	PL: D4T		
42	515	4,500	naive			L210W (6.3)	L: D4T, AZT
44	624	16,200	AZT+**NVP**+3TC			T69N (10.8) H221Y (6.6)	PL: **NVP**, EFV
48	82	2430,000	NVP+**D4T**+3TC			T215I (4.6)	L: AZT, **D4T**
49	202	935,000	AZT+**NVP**+3TC			G190A (14.3)	H: **NVP,** I: EFV
52	309	845,000	**AZT**+NVP+**3TC**	L210W (3.3)	L: D4T, **AZT**	K70T (4.3)	PL: **3TC**, D4T
57	317	53,596	NVP+**D4T**+**3TC**	M41L (8.2)	L: AZT, **D4T**	L74V (6.5) M184I (4.9) L210W (8.5)	H: **3TC**
59	195	5,470	**AZT**+**NVP**+3TC	L210W (3.6)	L: D4T, **AZT**	K103N (6.8)	H: **NVP**, EFV
60	242	30,000	AZT+**EFV**+3TC	K103N (16.5)	H: NVP, **EFV**		
61	63	15,200	AZT+**NVP**+3TC			M230I (4.8)	I: **NVP,** L: EFV
63	337	19,400	**D4T**+EFV+3TC			K219R (3.2)	PL: AZT, **D4T**
64	535	4,280	**AZT**+NVP+3TC	L210W (3.1)	L: D4T, **AZT**		
66	286	11,000	NVP+**D4T**+**3TC**			K65R (6.1)	I: **3TC, D4T**
69	193	622	**D4T**+EFV+3TC	L210W (7.6)	L: **D4T**, AZT		
70	257	290,000	AZT+**NVP**+**3TC**			M184I (12.1) M230I (11.2)	H: **3TC** I: **NVP** L: EFV
72	20	138,000	**NVP**+D4T+**3TC**			E138K (3.4) M184I (3.6)	H: **3TC** PL: EFV, **NVP**
73	416	1,370	AZT+**NVP**+**3TC**	F227C (3)	I: EFV, **NVP**	Y181C (8.1) M184V (7.9) H221Y (6.8)	H: **3TC, NVP** I: EFV
74	255	406	AZT+**NVP**+**3TC**			M184I (4) M230I (3.9)	H: **3TC** I: **NVP** L: EFV
75	604	132,000	AZT+**NVP**+3TC	F227C (3)	I: EFV, **NVP**		
77	245	11,800	**AZT**+NVP+**3TC**	D67G (5.7)	PL: **AZT**, D4T	M184I (6.9)	H: **3TC**

Note. [*]Drug resistance interpretations were evaluated for first-line drugs using the HIV Drug Resistance Database (HIVdb). Drugs in bold are the self-reported drugs taken by study subjects. H=high level drug-resistance, I=intermediate drug-resistance, L=low level drug-resistance, PL: potential low level drug-resistance.

resistance-associated positions were observed in proviral DNA than in viral RNA, especially in the minority resistant variants. Specifically, analysis by drug class revealed different drug resistance interpretations between viral RNA and paired proviral DNA. Of the samples with majority DRMs detected in proviral DNA that differed from those detected in paired viral RNA, 5 samples showed a higher level of drug resistance to the tested drugs. When using the 3% threshold, the prevalence of mutations in proviral DNA associated with resistance to first-line drugs was increased compared to that identified in viral RNA. Among samples with DRMs detected in proviral DNA, 22 samples were associated with a higher level of drug resistance to the tested RTIs when compared to the paired RNA. The valuable information provided by proviral DNA should not be ignored, given that plasma-based drug genotyping can underestimate the level of drug resistance. These data suggest that the deep sequencing of proviral DNA might provide an additional source of information with respect to the total burden of resistance, and should be used in the rational design of optimal therapy and prevention regimens. Moreover, a drug resistance genotyping assay based on patients' self-reported treatment regimens should be considered as a routine clinical protocol[27].

Our study showed that minority DRMs (frequency of 3%-20%) were more prevalent in proviral DNA than in viral RNA. Specifically, for 22 patients, including 8 treatment-naïve patients, minority DRMs detected in proviral DNA showed a higher level of resistance to first-line drugs. The DRMs that emerge before treatment with resistance to the initial treatment drugs may thus influence the treatment effect. Moreover, if these minority mutations turn into majority mutations during therapy, they may lead to treatment failure. The contribution of the minority mutations that emerge in proviral DNA during therapy to the high viral load observed in cases with no DRMs detected in viral DNA is not clear. Although previous reports indicated that low-frequency DRMs detected at baseline are associated with virological failure[28-30], whether minority mutations detected at baseline could influence the treatment outcome and how they change during long-term therapy should be further studied.

In conclusion, although limited by sample size, our study clearly demonstrates that the detailed and distinctive information of DRMs and drug resistance

interpretations for proviral DNA could be provided by deep sequencing comparing with paired viral RNA. The complementary information provided by deep sequencing of proviral DNA could provide more detailed and precise drug resistance information that is useful for drug resistance monitoring and in the rational design of optimal antiretroviral therapy regimens. As such, the combined analysis of viral RNA and proviral DNA could serve as a meaningful method to increase the sensitivity of HIV drug resistance testing. In addition, proviral DNA resistance genotyping may be a particularly important tool when the plasma viral load is undetectable or when plasma samples are not available.

ACKNOWLEDGEMENTS

CONFLICT OF INTREST

The authors have no commercial or other association that might pose a conflict of interest.

REFERENCES

1. Ray M, Logan R, Sterne JA, et al. The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals. Aids, 2010; 24, 123-37.
2. Palella FJ Jr, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med, 1998; 338, 853-60.
3. Gupta RK, Hill A, Sawyer AW, et al. Virological monitoring and resistance to first-line highly active antiretroviral therapy in adults infected with HIV-1 treated under WHO guidelines: a systematic review and meta-analysis. Lancet Infect Dis, 2009; 9, 409-17.
4. Clavel F, Hance AJ. HIV drug resistance. N Engl J Med, 2004; 350, 1023-35.
5. Rojas Sanchez P, Holguin A. Drug resistance in the HIV-1-infected paediatric population worldwide: a systematic review. J Antimicrob Chemother, 2014; 69, 2032-42.
6. De Luca A, Hamers RL, Schapiro JM. Antiretroviral treatment sequencing strategies to overcome HIV type 1 drug resistance in adolescents and adults in low-middle-income countries. J Infect Dis, 2013; 207, S63-69.

7. Lubke N, Di Cristanziano V, Sierra S, et al. Proviral DNA as a Target for HIV-1 Resistance Analysis. Intervirology, 2015; 58, 184-9.

8. Glenn TC. Field guide to next-generation DNA sequencers. Mol Ecol Resour, 2011; 11, 759-69.

9. Derache A, Shin HS, Balamane M, et al. HIV drug resistance mutations in proviral DNA from a community treatment program. PLoS One, 2015; 10, e0117430.

10. Palmisano L, Galluzzo CM, Giuliano M. The importance of testing genotypic resistance in proviral DNA of patients fully responding to highly active antiretroviral therapy. J Acquir Immune Defic Syndr, 2009; 51, 233-4.

11. Gallien S, Charreau I, Nere ML, et al. Archived HIV-1 DNA resistance mutations to rilpivirine and etravirine in successfully treated HIV-1-infected individuals pre-exposed to efavirenz or nevirapine. J Antimicrob Chemother, 2015; 70, 562-5.

12. Ma L, Sun J, Xing H, et al. Genotype and phenotype patterns of drug-resistant HIV-1 subtype B' (Thai B) isolated from patients failing antiretroviral therapy in China. J Acquir Immune Defic Syndr, 2007; 44, 14-9.

13. HIV DRUG RESISTANCE DATABASE [database online]. Stanford, California: Stanford University. Updated February 27, 2014.

14. Ekici H, Rao SD, Sonnerborg A, et al. Cost-efficient HIV-1 drug resistance surveillance using multiplexed high-throughput amplicon sequencing: implications for use in low- and middle-income countries. J Antimicrob Chemother, 2014; 69, 3349-55.

15. Bowtie 2 [computer program]. version 2.0.0. Baltimore:Johns Hopkins Unversity.

16. GATK [computer program]. Cambridge: Broad Institute.

17. Ram D, Leshkowitz D, Gonzalez D, et al. Evaluation of GS Junior and MiSeq next-generation sequencing technologies as an alternative to Trugene population sequencing in the clinical HIV laboratory. J Virol Methods, 2015; 212, 12-6.

18. Wensing AM, Calvez V, Gunthard HF, et al. 2014 Update of the drug resistance mutations in HIV-1. Top Antivir Med, 2014; 22, 642-50.

19. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med, 2011; 365, 493-505.

20. Brown AE, Gill ON, Delpech VC. HIV treatment as prevention among men who have sex with men in the UK: is transmission controlled by universal access to HIV treatment and care? HIV Med, 2013; 14, 563-70.

21. Choopanya K, Martin M, Vanichseni S, Mock PA, Suntharasamai P, Sangkum U. HIV antiretroviral prophylaxis for injecting drug users - Authors' reply. Lancet, 2013; 382, 855.

22. National Health and Family Planning Commission of the PRC.Further promoting prevention and treatment of HIV/AIDS. Available at http://www.nhfpc.gov.cn/jkj/s3585/201312/60987b42c19e4d649030f7c6fd399cf2.shtml. [2013-12-13]

23. Gianella S, Delport W, Pacold ME, et al. Detection of minority resistance during early HIV-1 infection: natural variation and spurious detection rather than transmission and evolution of multiple viral variants. J Virol, 2011; 85, 8359-67.

24. Todesco E, Rodriguez C, Morand-Joubert L, et al. Improved detection of resistance at failure to a tenofovir, emtricitabine and efavirenz regimen by ultradeep sequencing. J Antimicrob Chemother, 2015; 70, 1503-6.

25. Laur JJ, Weinberg GL. Comparing safety in surface landmarks versus ultrasound-guided peripheral nerve blocks: an observational study of a practice in transition. Reg Anesth Pain Med, 2012; 37, 569-70.

26. Kabamba-Mukadi B, Duquenne A, Henrivaux P, et al. HIV-1 proviral resistance mutations: usefulness in clinical practice. HIV Med, 2010; 11, 483-92.

27. Dudley DM, Bailey AL, Mehta SH, et al. Cross-clade simultaneous HIV drug resistance genotyping for reverse transcriptase, protease, and integrase inhibitor mutations by Illumina MiSeq. Retrovirology, 2014; 11, 122.

28. Li JZ, Paredes R, Ribaudo HJ, et al. Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. Jama, 2011; 305, 1327-35.

29. Cozzi-Lepri A, Noguera-Julian M, Di Giallonardo F, et al. Low-frequency drug-resistant HIV-1 and risk of virological failure to first-line NNRTI-based ART: a multicohort European case-control study using centralized ultrasensitive 454 pyrosequencing. J Antimicrob Chemother, 2015; 70, 930-40.

30. Vandenhende MA, Bellecave P, Recordon-Pinson P, et al. Prevalence and evolution of low frequency HIV drug resistance mutations detected by ultra deep sequencing in patients experiencing first line antiretroviral therapy failure. PLoS One, 2014; 9, e86771.

Supplemental Table S1. Characteristics of HIV-1 Infected Patients in China

Characteristics	Patients (*N*=43)
Age, median years (Range)	38 (26-64)
Male, no. (%)	16 (37.21)
Viral Load, median log10 copies/mL (Range)	5.10 (3.21-6.98)
CD4+ count, median cells/μL (Range)	217 (7-515)
Treatment, no. (%)	
AZT+NVP+3TC	24 (55.81)
D4T+NVP+3TC	17 (39.53)
D4T+EFV+3TC	1 (2.33)
AZT+EFV+3TC	1 (2.33)
High-risk behavior, no. (%)	
Former plasma donation	42 (97.67)
Sex	1 (2.33)