

Characterization of HIV Type 1 Genetic Diversity Among South African Participants Enrolled in the AIDS Vaccine Integrated Project (AVIP) Study

Maria A. Papathanasopoulos,¹ Eftyhia Vardas,² Carole Wallis,³ Richard Glashoff,⁴ Stefano Buttó,⁵
Guido Poli,^{6,7} Mauro Malnati,⁶ Mario Clerici,⁸ and Barbara Ensoli⁵

Abstract

The genetic diversity of HIV-1 strains circulating among HIV-1-infected South Africans was investigated in a cohort of 420 individuals enrolled as part of the AIDS Vaccine Integrated Project (AVIP) study. Representative samples (10%) were randomly selected from treatment-naive participants. Viral RNA was extracted for reverse transcriptase-initiated amplification and population-based sequencing of partial *pol* (encompassing *protease* and *reverse transcriptase*) and full-length *integrase*. Overall, HIV-1 sequences confirmed that 97.1% and 96.9% were HIV-1 subtype C in *pol* and *integrase*, respectively. Two participants were infected with unique A1/C and C/A1 recombinants in *pol/integrase*. Further *pol* sequence analysis identified mutation patterns associated with high level resistance to NNRTIs in two participants, whereas no primary mutations conferring resistance to integrase inhibitors were detected. The predominance of HIV-1 subtype C in South African populations is therefore confirmed in the AVIP cohort finalized for testing preventive or therapeutic vaccines against HIV-1 infection.

CONTROL OF THE HIV-1 PANDEMIC depends on the development of a preventive strategy, including an at least partially effective vaccine. One of the greatest obstacles to achieving this goal is the extraordinary genetic diversity of HIV-1, which is reflected by the worldwide distribution of HIV-1 groups, subtypes, circulating recombinant forms (CRFs), and ongoing viral evolution within each infected individual and at the population level.^{1,2} It is widely accepted that a successful vaccine regimen will need to induce strong, cross-subtype HIV-1-specific T cell immunity and broadly cross-reactive, neutralizing antibodies to overcome the challenge of HIV diversity.²

Although the constraints placed by ongoing HIV-1 genetic variation on vaccine efficacy remain unclear, it is highly unlikely that incorporation of proteins or DNA derived from a single natural viral isolate into a vaccine will confer protection against the same or other subtypes. Numerous strategies are being investigated to overcome this central problem, including the use of consensus/ancestral sequences, a combination

of immunogens from different subtypes, and computationally optimized mosaic immunogens.

HIV-1 group M subtype C accounts for over 50% of existing HIV-1 infections and 47% of new annual HIV-1 infections worldwide.³ Furthermore, it is the predominant circulating subtype among the heterosexual population in South Africa,⁴ the country with the highest number of HIV-1-infected individuals (<http://www.unaids.org>). Several South African sites have participated in preparatory studies and site capacity building for human HIV Vaccine Trials (Phases I to III). The AIDS Vaccine Integrated Project (AVIP) consortium is investigating novel HIV-1 vaccine designs combining both regulatory and structural viral proteins as immunogens. A cohort of HIV-1-positive South Africans was established to examine baseline virological and immunological features of infection encompassing both humoral and cellular immune responses. This study focused on the genetic diversity of HIV-1 strains circulating within this cohort, a major obstacle to any vaccine approach.

¹Department of Molecular Medicine and Haematology, University of the Witwatersrand Medical School, Johannesburg, South Africa.

²PHRU, University of the Witwatersrand Medical School, Johannesburg, South Africa.

³University of the Witwatersrand Medical School, Johannesburg, South Africa.

⁴Division of Medical Virology, University of Stellenbosch, Cape Town, South Africa.

⁵Istituto Superiore di Sanita, Rome, Italy.

⁶San Raffaele Scientific Institute, Milan, Italy.

⁷Vita-Salute San Raffaele University, Milan, Italy.

⁸University of Milan, and Don Gnocchi Foundation, IRCCS, Milan, Italy.

Whole blood samples were collected between 2004 and 2007 from 420 HIV-1-positive antiretroviral (ARV) drug-naive individuals attending the Perinatal HIV Research Unit (PHRU) at Chris Hani Baragwaneth Hospital in Johannesburg, South Africa. Plasma and peripheral blood mononuclear cells (PBMCs) were isolated and stored at -80°C and liquid nitrogen, respectively, for each participant, using standard methodologies. Ethical clearance for the study was obtained for Research on Human Subjects (Medical) at the University of the Witwatersrand.

Ten percent of all plasma samples were randomly selected for viremia determination from treatment-naive participants with $\text{CD4} > 400$ cells/ μl ($n = 21$) and $\text{CD4} < 400$ cells/ μl ($n = 21$). Viral RNA was extracted from clarified plasma samples using the automated Roche MagNA Pure LC analyzer and the MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche, Germany), according to the manufacturer's instructions. Extracted viral RNA was used to reverse transcriptase polymerase chain reaction (RT-PCR) amplify the partial *pol* (encompassing approximately 1.4 kb of the *protease* and *reverse transcriptase*) and full-length *integrase*, using in-house assays.^{5,6} PCR amplicons were purified using the High Pure PCR Product Purification kit (Roche), as per the manufacturer's instructions. Primers spanning the partial *pol* and *integrase* were used for bidirectional sequencing of the amplicons on the ABI Prism 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, CA) at a population-based level. Sequencing results were edited using the Sequencing Analysis V3.3 program (Applied Biosystems), and the complete sequences were assembled and manually edited using Sequencher V4.7 (Genecodes, Ann Arbor, MI).

A multiple alignment of the regions of interest with references from HIV-1 subtypes A to K, CRF01_AE, and CRF02_AG (<http://hiv-web.lanl.gov>) was generated in Clustal X. Aligned sequences were converted to MEGA V3.0 format and used in phylogenetic and molecular evolutionary analyses. A phylogeny reconstruction of each gene was performed by neighbor-joining using the Kimura two-parameter distance matrix. The stability of the nodes was assessed by bootstrap analysis (1000 replicates) and bootstrap values greater than 70% were considered significant. The partial *pol* and *integrase* phylogenetic subtyping results were also compared to those obtained from the REGA HIV-1 subtyping tool (version 2.0) from the HIV drug resistance database maintained by Stanford University (<http://hivdb.stanford.edu>). The partial *pol* sequences were submitted directly to the HIV-1 drug resistance database to infer the presence/absence of different antiretroviral drug resistance mutations and levels of resistance to the commonly used reverse transcriptase and protease inhibitors. The *integrase* nucleotide sequences for each patient were converted to the respective amino acid sequences, aligned, and extensively analyzed for the presence of mutations affecting viral susceptibility/resistance to raltegravir, elvitegravir, and other investigational integrase inhibitors (<http://hivdb.stanford.edu>).

Overall, HIV-1 sequences were obtained from 35 of the 42 participants (83% success rate). However, matched partial *pol* and *integrase* HIV-1 sequences were obtained from 32 participants, including 17 individuals with CD4 counts > 400 cells/ μl and 15 subjects with CD counts < 400 cells/ μl . Epidemiological, clinical, and subtyping data analysis of the 35 participants is summarized in Table 1. Seventy-five percent of

the participants were female, with a median age of 33 years, and median viral loads of 21,700 RNA copies/ml (ranges 645 to $> 1,000,000$ RNA copies/ml). Phylogenetic tree analysis of the partial *pol* (Fig. 1A) and *integrase* (Fig. 1B) sequences with viruses from the major subtypes showed that 30 of the 32 were HIV-1 subtype C.

The remaining two participants were infected with unique intersubtype recombinant strains: virus 07ZA17579 was subsubtype A1 and C in the *pol* and *integrase* regions, respectively, whereas virus 07ZA17657 was subtype C and subsubtype A1 in the partial *pol* and *integrase* regions, respectively (Fig. 1A and B). These recombinant strains combine segments of subsubtype A1 that are typically found in East and Central Africa with subtype C, which is more common in Southern Africa.^{3,6} Thus, they could have originated either in East or Central Africa, or they could represent recombination events that occurred between East or Central African and South African strains. The remainder three samples (07ZA17751, 07ZA17865, and 06ZA15274) could not be amplified in the *integrase* region and were designated HIV-1 subtype C by analysis of their partial *pol* regions. Since participants 17751, 17865, and 15274 had viral loads of 1140, 1300, and 1190 RNA copies/ml, respectively, it is conceivable that the in-house *integrase* assay has a different lower limit of detection than the in-house partial *pol* assay. Alternatively, primer mismatches could account for the lack of amplification. Furthermore, six of the seven samples that could not be amplified in *pol* and *integrase* were from participants with low/undetectable viral loads (viral load ranges < 400 to 1450 RNA copies/ml).

Analysis of antiretroviral drug resistant mutations among the therapy-naive study participants showed that most of them had viruses susceptible to all known reverse transcriptase and protease inhibitors ($n = 33/35$, Table 1). However, mutation patterns associated with high level resistance to nonnucleoside reverse transcriptase inhibitors (NNRTIs) were identified in two (5.7%) participants (17657 and 16923). It is possible that participants 17657 and 16923 may have taken antiretroviral drugs previously, and not disclosed this to the study coordinator. Alternatively, they could have been infected with antiretroviral drug resistant virus, or participant 16923, who is a female of childbearing age, may have taken nevirapine as part of the national PMTCT (prevention of mother-to-child transmission) program in South Africa. Participant 17193 contained K103R, a mutation known to occur in 1–2% of untreated individuals. Because K103R was not found in combination with V179D (which reduces nevirapine, delavirdine, and efavirenz susceptibility by about 15-fold), this mutation has no effect on NNRTI susceptibility.⁷ In addition, the presence of V118I, a subtype C-specific polymorphism associated with NRTIs, was detected in participants 17503 and 17898. V118I is found in approximately 2% of therapy-naive individuals. It causes low-level resistance to lamivudine (3TC) and possibly to other NRTIs when present with E44A/D and/or one or more thymidine analogue mutations.⁸ However, these associated mutations were not found in these participants. Furthermore, participant 17503 had the E138A mutation, a polymorphism recently associated with decreased responses to the NNRTI inhibitor etravirine.⁹

The presence of subtype C-specific polymorphisms and minor resistance mutations to protease inhibitors was detected among 11 other participants. Seven participants (17332,

TABLE 1. DATA FOR 35 PARTICIPANTS ENROLLED IN THE AVIP STUDY FROM WHOM SEQUENCES WERE OBTAINED

AVIP patient number	Sex ^a	Age (years)	Year of collection	Viral load (RNA copies/ml)	CD4 count (cells/ μ l)	Subtype designation ^b	Drug resistance genotype reverse transcriptase and protease ^c	Drug resistance genotype integrase ^d
CD4 counts >400 cells/μl								
17332	M	28	2007	12,800	709	C/C	Susceptible	Susceptible
17409	M	41	2007	8,170	421	C/C	Susceptible	Susceptible
17462	F	42	2007	4,960	438	C/C	Susceptible	Susceptible
17503	F	48	2007	8,680	440	C/C	Susceptible, E138A (NNRTI)	Susceptible
17491	F	30	2007	2,070	425	C/C	Susceptible, PI minor resistance mutation Q58E	Susceptible
17556	F	33	2007	22,700	557	C/C	Susceptible	Susceptible
17579	M	57	2007	415,000	403	A1/C	Susceptible	Susceptible, G163R ^e
17568	F	34	2007	645	431	C/C	Susceptible	Susceptible
17604	M	32	2007	116,000	863	C/C	Susceptible	Susceptible
17657	M	36	2007	11,200	424	C/A1	High level NNRTI resistance (K103N)	Susceptible
17699	F	28	2007	121,000	708	C/C	Susceptible	Susceptible
17716	M	35	2007	20,300	452	C/C	Susceptible	Susceptible
17751	M	38	2007	1,140	403	C	Susceptible	NA ^f
17865	F	24	2007	1,300	445	C	Susceptible	NA
17898	F	41	2007	70,200	539	C/C	Susceptible	Susceptible
17854	M	30	2007	3,140	504	C/C	Susceptible	Susceptible
17872	F	NA	2007	9,180	426	C/C	Susceptible	Susceptible
16853	F	37	2006	201,000	558	C/C	Susceptible	Susceptible
16985	F	21	2006	28,800	425	C/C	Susceptible	Susceptible
CD4 counts <400 cells/μl								
17188	F	31	2007	21,700	303	C/C	Susceptible	Susceptible
17200	F	36	2007	42,100	224	C/C	Susceptible	Susceptible
17193	F	51	2007	32,400	284	C/C	Susceptible, K103R	Susceptible
17241	F	25	2007	17,241	336	C/C	Susceptible	Susceptible
17058	F	33	2006	367,000	163	C/C	Susceptible	Susceptible
17067	F	31	2006	47,100	332	C/C	Susceptible	Susceptible
15274	F	33	2006	1,190	273	C	Susceptible	NA
14998	F	37	2006	14,900	270	C/C	Susceptible	Susceptible
15141	F	28	2006	81,200	219	C/C	Susceptible	Susceptible
16976	F	24	2006	227,000	111	C/C	Susceptible	Susceptible
17071	F	46	2006	4,900	324	C/C	Susceptible, PI minor resistance mutations L10V, V11IV	Susceptible
16938	F	30	2006	343,000	260	C/C	Susceptible	Susceptible, V151I ^e
17212	F	30	2007	>1,000,000	87	C/C	Susceptible, PI minor resistance mutation N88K	Susceptible
16997	F	31	2006	102,000	394	C/C	Susceptible	Susceptible
16923	F	26	2006	48,800	6	C/C	High Level NNRTI resistance (K103N); PI minor resistance mutation A71T	Susceptible
16642	M	43	2006	8,170	324	C/C	Susceptible	Susceptible

^aM, male; F, female.

^bPartial *pol/integrase*, based on Rega HIV-1 subtyping tool (<http://hivdb.stanford.edu/>) and phylogenetic tree analysis.

^cFor nucleoside reverse transcriptase inhibitors (NRTI), nonnucleoside reverse transcriptase inhibitors (NNRTI), and protease inhibitors (PI).

^dIntegrase inhibitors (<http://hivdb.stanford.edu/>, accessed September 2008).

^ePolymorphic mutations associated with raltegravir resistance.

^fNot available.

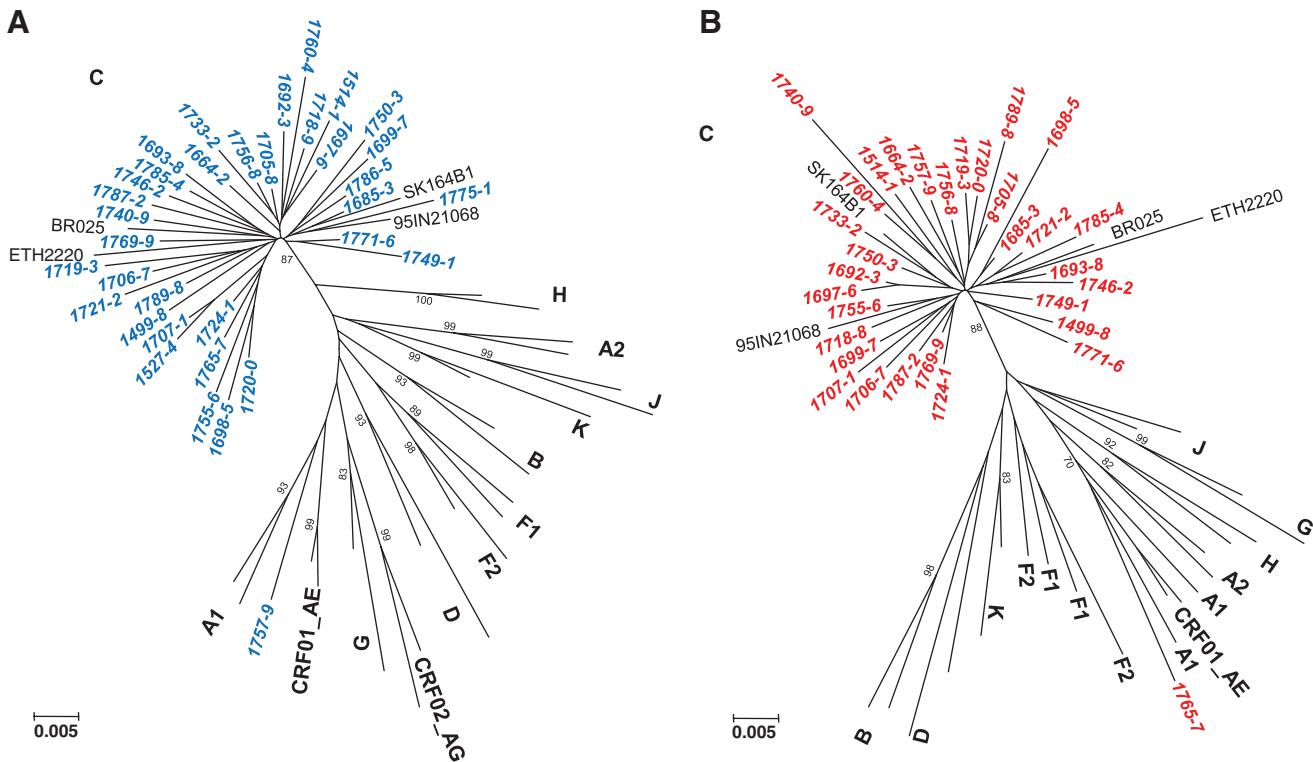


FIG. 1. Phylogenetic relationships of the newly characterized partial *pol* (protease and reverse transcriptase) sequences (**A**) and full length *integrase* sequences (**B**) with HIV-1 subtype reference sequences from the Los Alamos database (<http://hiv-web.lanl.gov>). Phylogenetic trees were constructed from nucleotide alignments, using neighbor joining and the Kimura two-parameter distance matrix. Only bootstrap values of 70% or higher are shown. The new sequences are shown italicized in bold.

17462, 17699, 17716, 17872, 16938, and 16997) had the T74S mutation that has been described to confer potential low level resistance to nelfinavir.¹⁰ However, this is a naturally occurring subtype C-specific polymorphism, and was thus not listed in Table 1 (as per WHO recommendations). Of the remaining four participants, 17491, 17071, 17212, and 16923 had the Q58E, L10V and V11I, N88K, and A71T, respectively (Table 1). Q58E is a major tipranavir resistance mutation associated with decreased susceptibility to tipranavir/ritonavir and possibly other protease inhibitors.¹¹ L10V is a minor tipranavir/ritonavir resistance mutation, resulting in resistance to most protease inhibitors when present with other mutations.¹¹ V11I is a protease inhibitor-selected mutation associated with decreased response to darunavir/ritonavir.¹² The presence of N88D/S/T/G has diverse effects on several protease inhibitors.¹³ Finally, the A71T/V polymorphisms are associated with protease inhibitor-based treatment.¹⁴

As expected, none of the 32 participants had primary mutations associated with resistance to the *integrase* inhibitors raltegravir and elvitegravir (Table 1). Samples 17579 and 16938 contained the G163R and V151I polymorphic mutations, respectively, that have been described with raltegravir use.¹⁵ Sequencing results from *integrase* obtained in this study confirms previous findings from our laboratory that no baseline polymorphisms affecting the susceptibility to raltegravir could be detected in therapy-naïve individuals, and therefore, it can be considered for use in the HIV-1-infected South African population.⁵

The overall predominance of HIV-1 subtype C in the infected South African population is confirmed in our AIVP cohort. However, the presence of unique recombinant forms, together with a small, but significant proportion (5.7%) of participants harboring antiretroviral drug-resistant virus with primary mutations, highlights the need for ongoing surveillance of circulating HIV-1 isolates within South Africa. This is critical to ensure that any vaccine efficacy is not adversely affected by the continuous dynamic evolution of HIV-1.

Sequence Data

All sequences were submitted to GenBank using Sequin V9.2 (<http://www.ncbi.nlm.nih.gov>), and are available under the accession numbers GU253395 to GU253429 for *protease/reverse transcriptase* and GU253430 to GU253461 for *integrase*.

Acknowledgments

The AIDS Vaccine Integrated Project is funded by the FP6 project of the European Commission.

Author Disclosure Statement

No competing financial interests exist.

References

1. Gaschen B, Taylor J, Yusim K, *et al.*: Diversity considerations in HIV-1 vaccine selection. *Science* 2002;296:2354–2360.

2. Barouch DH: Challenges in the development of an HIV-1 vaccine. *Nature* 2008;455:613–619.
3. Hemelaar J, Gouws E, Ghys PD, and Osmanov S: Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006;20:W13–23.
4. Wallis CL, Papathanasopoulos MA, Lakhi S, *et al.*: Affordable in-house antiretroviral drug resistance assay with good performance in non-subtype B HIV-1. *J Virol Methods* 2010;163(2):505–508.
5. Fish M, Hewer R, Wallis CL, *et al.*: Natural polymorphisms of integrase amongst HIV-1 infected South African patients. *AIDS Research and Human Retroviruses* (in press).
6. Papathanasopoulos MA, Hunt GM, and Tiemessen C: Review: Evolution and diversity in HIV-1 in Africa. *Virus Genes* 2003;6:151–163.
7. Parkin NT, Gupta S, Chappay C, and Petropoulos CJ: The K101P and K103R/V179D mutations in human immunodeficiency virus type 1 reverse transcriptase confer resistance to nonnucleoside reverse transcriptase inhibitors. *Antimicrob Agents Chemother* 2006;50:351–354.
8. Stoeckli TC, MaWhinney S, Uy J, *et al.*: Phenotypic and genotypic analysis of biologically cloned human immunodeficiency virus type 1 isolates from patients treated with zidovudine and lamivudine. *Antimicrob Agents Chemother* 2002;46:4000–4003.
9. Kagan RM, Sista P, Pattery T, Bachelier L, and Schwab DA: Additional HIV-1 mutation patterns associated with reduced phenotypic susceptibility to etravirine in clinical samples. *AIDS* 2009;23:1602–1605.
10. Deforche K, Silander T, Camacho R, *et al.*: Analysis of HIV-1 pol sequences using Bayesian networks: Implications for drug resistance. *Bioinformatics* 2006;22:2975–2979.
11. de Mendoza C, Morelló J, Garcia-Gascó P, Rodríguez-Novoa S, and Soriano V: Tipranavir: a new protease inhibitor for the treatment of antiretroviral-experienced HIV-infected patients. *Expert Opin Pharmacother* 2007;8:839–850.
12. De Meyer S, Vangeneugden T, van Baelen B, *et al.*: Resistance profile of darunavir: Combined 24-week results from the POWER trials. *AIDS Res Hum Retroviruses* 2008;24:379–388.
13. Mitsuya Y, Winters MA, Fessel WJ, *et al.*: N88D facilitates the co-occurrence of D30N and L90M and the development of multidrug resistance in HIV type 1 protease following nelfinavir treatment failure. *AIDS Res Hum Retroviruses* 2006;22:1300–1305.
14. Masquelier B, Breilh D, Neau, Lawson-Ayayi S, *et al.*: Human immunodeficiency virus type 1 genotypic and pharmacokinetic determinants of the virological response to lopinavir-ritonavir-containing therapy in protease inhibitor-experienced patients. *Antimicrob Agents Chemother* 2002;46:2926–2932.
15. Lataillade M, Chiarella J, and Kozal MJ: Natural polymorphism of the HIV-1 integrase gene and mutations associated with integrase inhibitor resistance. *Antivir Ther* 2007;12:563–570.

Address correspondence to:

Maria A. Papathanasopoulos

HIV Pathogenesis Research Laboratory

Department of Molecular Medicine and Haematology

University of the Witwatersand Medical School

7 York Road

Parktown, 2193 Johannesburg, South Africa

E-mail: Maria.Papathanasopoulos@wits.ac.za

