AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 22, Number 10, 2006, pp. 985–992

Genotypic Variation in the *pol* Gene of HIV Type 1 in an Antiretroviral Treatment-Naive Population in Rural Southwestern Uganda

CATHERINE V. GALE,[1] DAVID L. YIRRELL,[2–4] EILEEN CAMPBELL,[3] LIEVE VAN DER PAAL,[4] HEINER GROSSKURTH,[4] and PONTIANO KALEEBU[4]

ABSTRACT

The majority of studies of HIV-1 drug resistance have involved subtype B viruses. Here we have characterized subtype distribution and determined the levels of polymorphism at protease (PR) and reverse transcriptase (RT) drug resistance positions, in antiretroviral treatment-naive HIV-positive Ugandan patients. We have also investigated codon usage variability at these positions and assessed intersubtype recombination within the *pol* gene. The study population consisted of 187 patients, from a cohort established by the UK Medical Research Council Programme on AIDS in Uganda in 1990. Results indicate that 28.3% of patients were infected with subtype A ($n = 53$), 64.2% subtype D ($n = 120$), 6.4% A/D recombinant ($n = 12$), and 1.1% subtype C ($n = 2$). Variation in amino acid usage at drug resistance-associated positions was minimal between the two main subtypes (A and D) in RT, but there was appreciable variation in PR. Codon usage, however, was considerably more variable between subtypes A and D in both PR and RT. Thus, while no natural high-level resistance to antiretroviral therapy was detected in this cohort, subtypes A and D may possess different genetic barriers to be overcome in order to achieve resistance. With the increasing introduction of antiretroviral therapy into Africa, such information will be vital in our understanding and evaluation of the development of drug resistance as it occurs, and how to interpret resistance data the type of which has rarely previously been seen. This analysis also significantly increases the number of Ugandan PR and RT sequences characterized to date.

INTRODUCTION

PHYLOGENETIC SEQUENCE ANALYSIS OF HIV-1 has resulted in its classification into three major lineages, M (major), O (outlier), and N (non-M/O or new), each thought to represent a separate transmission of SIVcpz from chimpanzees into the human population.[1] Group M accounts for the vast majority of HIV-1 infections worldwide, and has been subdivided into nine subtypes termed A, B, C, D, F, G, H, J, and K. Furthermore, mosaic genomes of HIV-1 resulting from intersubtype recombination have also been recognized as fully transmissible and pathogenic. Such circulating recombinant forms (CRFs) appear to play an important role in the HIV pandemic and to date 18 have been reported.[2] Unique and complex recombinant forms are also increasingly being reported.[3] Thus the considerable variability between human immunodeficiency virus type-1 (HIV-1) strains has a potential impact on disease epidemiology, diagnosis, therapy, and the prevention of infection.

The HIV-1 epidemic in developed countries has historically been driven almost exclusively by subtype B viruses. It is in these areas that the majority of research and treatment has taken place and thus knowledge regarding antiretroviral (ARV) drug resistance and naturally occurring genotypic variation in HIV-1 protease (PR) and reverse transcriptase (RT) is biased toward subtype B.[4] However, Asia and Africa, where non-B subtypes predominate, bear 90% of the worldwide HIV-1-related disease burden. As ARVs are becoming more readily available throughout the world this imbalance must be addressed such that there

[1]Centre of Virology, University College London, London, UK.
[2]Department of Public Health, University of Glasgow, Glasgow, UK.
[3]Specialist Virology Centre, Gartnavel General Hospital, Glasgow, UK.
[4]MRC/UVRI Uganda Research Unit on AIDS, c/o Uganda Virus Research Institute, Entebbe, Uganda.

is sufficient knowledge accrued to decide whether all subtypes can be considered equal in terms of response to treatment and in development and transmission of drug resistance. Furthermore, in Europe infections with non-B subtypes are being detected with increasing frequency[5] (Gifford *et al.*, submitted), making knowledge of subtype differences pertinent to both the developed and developing world.

Uganda is one of the sub-Saharan countries that has been progressive in treatment scale-up; at the end of December 2004 just over one-third of patients who needed ARV drugs were receiving them (UNAIDS bulletin, 2005). Despite the increasingly widespread use of ARVs in Uganda, however, only a small number of studies have looked at in the two main circulating subtypes (A and D) in terms of levels of polymorphism and development of HIV-1 drug resistance. Such studies have highlighted, in untreated patients, differences in amino acid polymorphisms between subtypes A and D at certain drug resistance-associated positions in RT[6] and PR.[7] One subtype D isolate has also been identified that is naturally highly resistant to NNRTIs, due to the presence of polymorphisms in RT that constitute "nonclassical" resistance mutations.[8] In Ugandan patients receiving ARVs it has been established that levels and patterns of resistance in subtypes A and D are similar to those seen with subtype B.[7] Between subtypes A and D, however, discrepancies have been noted in terms of types of mutations selected after single dose nevirapine treatment[9,10] and the extent of resistance that may develop on treatment.[11] In the latter case, levels of resistance have been found to be lower in subtype A-infected patients, which has also been observed in non-Ugandan studies.[12] Furthermore, a population-based study has shown that intersubtype (A/D) *pol* recombinants are common in Uganda and that these may be perinatally transmitted, illustrating their potential evolutionary stability and thus epidemic significance.[13]

The aim of the present study was to acquire a large dataset of PR and RT sequences from HIV-infected, treatment-naive Ugandans. Several questions could then be asked; first, whether there is any evidence that subtype A or D viruses circulating in the sampled Ugandan population are sufficiently polymorphic within PR and RT to affect the development of classical drug resistance mutations (at IAS-USA positions); and second, whether there are any differences between subtypes A and D in this respect, which may help explain previous observations.[10,12] Third, determination of intersubtype recombination frequency and structure within the study population will be important if we are to understand the potential recombination has to affect sequence evolution within *pol*, and thus the development of drug resistance.

The study population consisted of HIV-1-infected drug-naive individuals who were part of a clinical cohort living in rural southwestern Uganda, established by the UK Medical Research Council Programme on AIDS in Uganda in 1990.[14] This population has been extensively studied in the past[15–20] but with the introduction of ARVs into this cohort as of October 2004, this is the first consideration of the importance of viral genetic variation within the *pol* gene. This study also provides a basis for evaluation of the transmission of resistance,[21,22] choice of initial treatment,[23] and identification of the genetic changes resulting from treatment failure[24] in Uganda.

MATERIALS AND METHODS

Study population

The Medical Research Council (MRC) Programme on AIDS in Uganda established a Rural Clinical Cohort (RCC) of HIV-1-infected individuals and HIV-1-uninfected controls living in rural southwestern Uganda in 1990.[14] Participants were, and continue to be, recruited from a larger population study where the dynamics of HIV-1 infection were examined by annual census, questionnaire, and serological surveys.[25] Samples for this study were collected between 1996 and 2003.

RNA extraction, RT-PCR, and sequencing

HIV-1 RNA was extracted from 200 μl of blood plasma, using RNA extraction columns (Qiagen, Chatsworth, CA). Polymerase gene-specific primers were used for reverse transcription (RT), followed by nested polymerase chain reactions (PCRs) to amplify the *pol* gene encompassing codons 1–242 of RT and codons 1–99 of protease (PR). DNA sequencing of PR and RT-PCR products was achieved using a Beckman CEQ 2000 automated capillary DNA sequencer (Beckman Coulter, UK).

Phylogenetic analysis

Sequence data were analyzed using BioEdit v5.0.9 (Hall, 1999), ClustalW,[26] and Sequencher (GeneCodes, USA) programs. Neighbor-joining (NJ) trees were constructed using the NJ algorithm implemented in ClustalW and edited using Tree Explorer v2.12 from the MEGA2 software package.[27] Subtype and recombination frequency were determined using STAR[28] and STARrec,[29] and recombination was also assessed using Simplot v3.5.1[30] (100 bootstrap replicates, 200-bp sliding window, 10-bp step—any recombinants identified confirmed with 400-bp sliding window).

Amino acid variation analysis

Amino acid variation at each drug resistance-associated mutation position listed by the International AIDS Society—USA (IAS-USA)[31] was calculated. Specifically, the percentage of wild-type (WT), mutant, and any other amino acid was calculated, per position, and expressed as a percentage of amino acid frequency. This analysis was performed for subtypes A and D sequences separately. To determine differences between subtypes A and D, the normalized frequency of WT, drug resistance-associated mutant, and any other amino acid at each IAS-USA mutation position for subtype A were subtracted from those for subtype D and the results were plotted graphically.

Codon variation analysis

Codon variation at each drug-resistance mutation position listed by the IAS-USA was assessed. Where codon usage differed between subtypes A and D this was recorded as the percentage of sequences in each subtype grouping that possessed one of the two main variant codons.

RESULTS

Sequence processing and subtype classification

Subtype classification of all sequences was achieved using the subtyping tool STAR.[28,29] Of the 187 sequences, 28.3% were subtype A ($n = 53$), 64.2% subtype D ($n = 120$), 6.4% A/D recombinant ($n = 12$), and 1.1% subtype C ($n = 2$). For phylogenetic analysis, alignments of all *pol* sequences from the study cohort were edited to remove sequences that were less than full length ($n = 22$), or had been found to be recombinant ($n = 12$). These sequences were of good quality with a mean number of nucleotide mixtures called per complete sequence of 6.7 (0.7%), ranging from 0 to 42 (0–4.1%). The recombinant sequences identified were analyzed separately using STARrec and Simplot to confirm recombination. For amino acid and nucleotide mutation analysis, all full-length sequences were separated into subtype-specific alignments for subtypes A and D.

Phylogenetic analysis

Neighbor-joining phylogenetic analysis revealed that all *pol* sequences from the study cohort classified as subtypes A, D, or C by STAR clustered with reference sequences of the relevant subtype (Fig. 1). The subtype A, D, and C radiations were supported by bootstrap values of 98.4%, 72.6%, and 100%, respectively. The lower value for the subtype D radiation is due to the presence of two monophyletic subtype D sequences (151634 and 351180). Closer examination of these sequences revealed that they had scored poorly using the STAR subtyping algorithm, suggesting a weak subtype assignation but not sufficient to justify STARrec recombination analysis.[29] Simplot analysis, however, revealed that these two sequences were D/A recombinants of complex architecture, consisting mainly of subtype D, explaining their monophyletic distribution within subtype D radiation. Within the subtype A radiation no monophyletic sequences were present, suggesting that all sequences included were correctly classified as pure subtype A in the *pol* region being studied.

Recombination analysis

Using STAR, 12 sequences (6.4%) were identified as being recombinant. STARrec analysis was then used to give an estimate of recombinant architecture. These data indicate that in the case of the 12 sequences here analyzed, there is some conservation among recombination breakpoints within *pol* (Table 1). This feature of STARrec, however, has never been formally tested. Accordingly, in order to further analyze the recombinant architecture of these sequences they were also analyzed using the well-described recombination detection program, Simplot.[30] Bootscanning plots are shown in Fig. 2 for two of the recombinants detected using STAR (412+530787 and 407) together with the STARrec output, illustrating the agreement between the two methods. Table 1 summarizes these data and illustrates that recombinants and recombination breakpoints identified using STAR are in good agreement with those identified with Simplot, although Simplot does tend to predict a greater number of breakpoints for each sequence. In the case of these, however, the mean length of "extra" recombinant sequences pre-

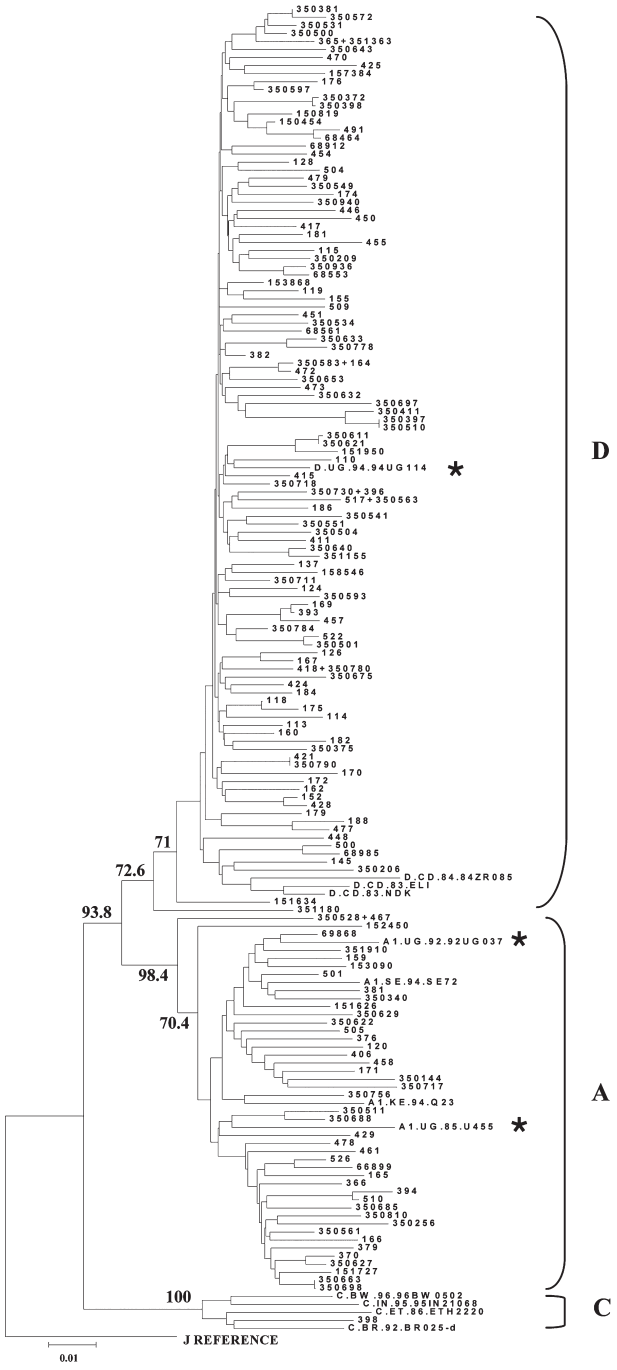

FIG. 1. Phylogenetic tree of *pol* gene sequences from 154 HIV-1 infected Ugandan patients. Subtype A (n = 4), C (n = 4) and D (n = 4) reference sequences are included (those of Ugandan origin marked as *). The tree is rooted with a subtype J reference sequence. 33 sequences from the original dataset were removed from the alignment used to build this tree as they were either partial in length, or found to be recombinant. Bootstrap values for the major internal tree nodes are shown (1000 replicates) and subtype A, D and C radiations are indicated. The scale bar represents a 1% nucleotide difference.

TABLE 1. DETAILS OF RECOMBINANT STRUCTURE OF SEQUENCES IDENTIFIED AS BEING A/D RECOMBINANT BY STAR[a]

		Recombinant regions[b]		
Isolate	*Main subtype*	*First*	*Second*	*Third*
532	A	400–450 (C)[c]	450–550 (D)	***800+ (D)***[d]
132	A	550–570 (D)	600–610 (D)	***750+ (D)***
435	A	400–450 (C)	***450–720 (D)***	—
15312	A	***700+ (D)***	—	—
412+350787[e]	A	***450–710 (D)***	—	—
149	A	***350–650 (D)***	750–800 (D)	***850+ (D)***
407[e]	D	***0–400 (A)***	—	—
419	D	***0–375 (A)***	650–750 (A)	—
527	D	***500–800 (A)***	—	—
138+187	D	***0–400 (A)***	—	—
350566	D	***0–300 (A)***	300–419 (C)	***425–600 (A)***
380	D	***750+ (A)***	—	—

[a]Analysis performed using STARrec and Simplot. Agreement between methods is indicated.
[b]Recombinant regions are segments of the sequence analyzed that are not the main subtype. First to third refers to the order the recombinant segments appear in a 5′ to 3′ direction, from the beginning of protease to the end (position 242) of RT.
[c]Standard font indicates recombinant regions that Simplot alone identified.
[d]Bold italic font indicates recombinant regions that both STARrec and Simplot identify.
[e]Comparison between STARrec and Simplot outputs shown in Fig. 2.

dicted by Simplot was low (63 nucleotides) and is likely to represent regions where the sequence being analyzed was less like both A and D sets of reference sequence and more like the outlier reference group (subtype C), without necessarily being subtype C itself. No temporal differences in the detection of subtypes A, D, or A/D recombinants were found within the sampling period.

Amino acid variation analysis

Amino acid variation at IAS-USA resistance positions within the region of *pol* being studied is presented in Fig. 3. It is immediately clear that the majority of variations between the two subtypes, within this dataset, are in the protease gene at positions 10, 33, 36, 63, and 77. Chi-square testing revealed a significant difference in frequency between the wild-type and mutant amino acids, between subtypes A and D, at positions 10, 36, and 63 ($p < 0.05$).

In more detail, of 53 subtype A and 120 subtype D sequences, at position 10, 2% of the subtype D isolates possess the mutant amino acid (I) compared to 12% of subtype A isolates. At position 33, 100% of subtype A isolates possess the wild-type amino acid (L), whereas 11% of subtype D isolates have an amino acid that is neither WT nor mutant at this position. At position 36, while 32% of subtype D viruses encode the WT amino acid (M) and 66% the mutant residue (I), the majority of subtype A isolates possess either the resistance-associated mutant (I–94%) or an alternative residue (6%). At position 63, subtype A has a greater frequency of the WT amino acid (L—71%) compared to isolates of subtype D, which largely possess either the mutant (P—24%) or a different amino acid (31%). Finally at position 77 of protease, all subtype A isolates have the WT residue (V), but 8% of subtype D isolates have the mutant residue (I). Position 36 of protease therefore represents the greatest discrepancy between subtypes A and D in terms of amino acid usage at resistance positions.

In this sequence dataset, within RT, there is very little difference between subtypes A and D in terms of amino acid variation at resistance positions.

Codon variation analysis

While none of the 29 mutation positions in RT was variable at the amino acid level (Fig. 3), 7 of these 29 positions (24%) were found to have variable codon usage between subtypes A and D (Table 2). These include positions 62, 67, 70, 77, 151, 215 [nucleoside reverse transcriptase inhibitor (NRTI) resistance associated], and 100 [non-NRTI (NNRTI) resistance associated]. While drug resistance positions are, by definition, more plastic, i.e., a mutation can be accommodated in order to facilitate reduced drug sensitivity without overt loss of functionality of the gene product in question, it is interesting that where the codon score differed at five of the seven positions, this was by the substitution of a thymine (in subtype A) with a cytosine (in subtype D). This pattern is also borne out in protease (positions 10 and 63). Of the synonymously variable positions identified, only the variation at position 151 (Q) has the potential to influence the number of nucleotide changes required to achieve the resistance-associated codon (M).

DISCUSSION

The majority of research addressing HIV-1 drug resistance has involved the study of subtype B viruses, the outcome being the classification of both minor and major mutations associated with drug resistance. Minor mutations in subtype B have little effect on viral susceptibility to a drug *in vitro*: they mainly

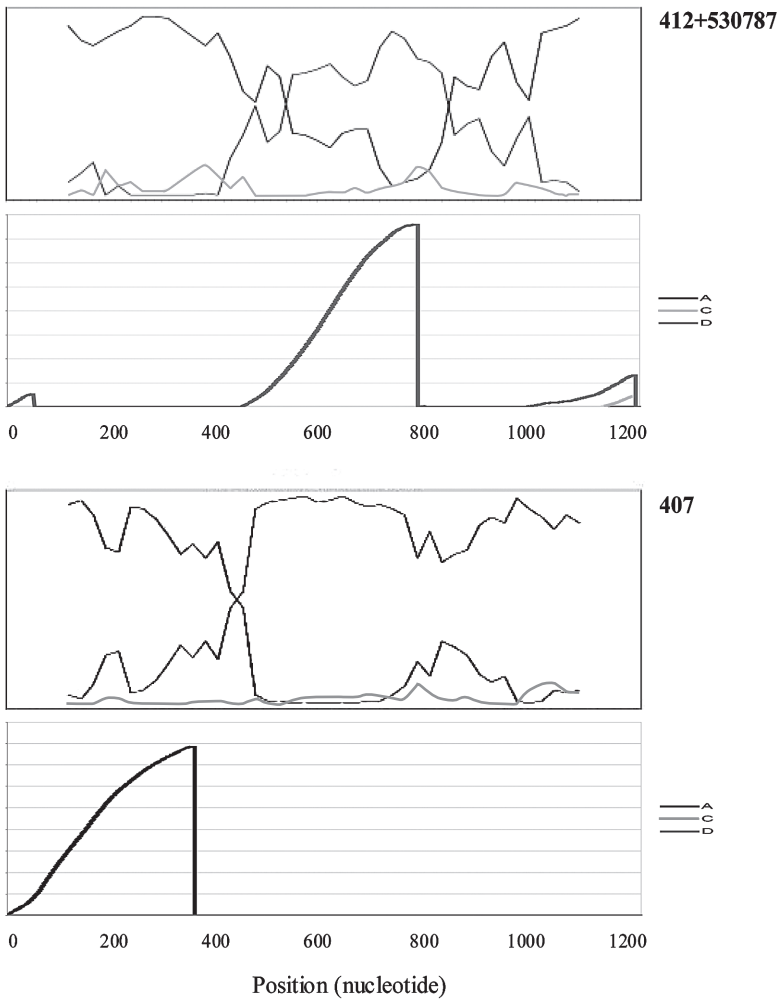

FIG. 2. Graphical outputs from bootscanning (Simplot) and STARrec analysis, for two of the recombinant *pol* sequences, identified using STAR. Bootscanning charts show a derived score representing recombinant architecture of the query sequence between the two 'expected' subtypes (A and D) and an outlier (C), whereas STARrec charts show a cumulative score of any subtype for which the score is greater than that for the predicted query subtype (giving the distinctive "cliff edge' shape to the graph at the breakpoint where the recombinant returns to the major subtype). For both sequences, using both methods recombination breakpoints are in close agreement. For the remaining 10 sequences recombination characteristics, as defined by both methods, are summarised in Table 1.

impart increased fitness.[32] Major mutations, however, are associated with virologic and thus therapeutic failure, and have a marked impact on *in vitro* susceptibility of HIV to the drug/drug class in question. In non-B subtypes, while most evidence points to the fact that major mutations developing at therapeutic failure are the same as those characterized in subtype B viruses,[33] minor mutations can exist as natural polymorphisms in viruses from drug-naive patients.[34] Since HIV-1 subtypes differ from one another by 10–12% of their nucleotides and 5–6% of their amino acids in PR and RT,[35] these differences may influence the spectrum of mutations that develops during selective drug pressure.

Since ARVs have recently been introduced into a well-characterized cohort in Uganda, PR and RT gene sequencing of 187 cryopreserved serum samples from these individuals (collected between 1996 and 2003, prior to introduction of ARVs) provided the opportunity to examine genetic variation in the *pol* gene in drug-naive individuals. These sequences provide a baseline for the inevitable future development of drug resistance in this population. It will be important to understand the baseline level of drug-resistance-associated polymorphism and genetic recombination both within and between the major genetic variants affecting this population in order to determine the choice of initial treatment, to evaluate future transmission of resistance, and to identify the genetic changes resulting from treatment failure.[4]

While the subtype distribution characteristics described for this dataset are consistent with those previously published (dominated by subtype A and D, with a small number of subtype C infections), the frequency of recombinants within *pol* (6.4%) is much lower than the figure of 30% previously determined for the same population using matched *gag* and *env* sequences[20] and in other Ugandan studies using complete genome sequences.[36] This suggests that recombination, while frequent

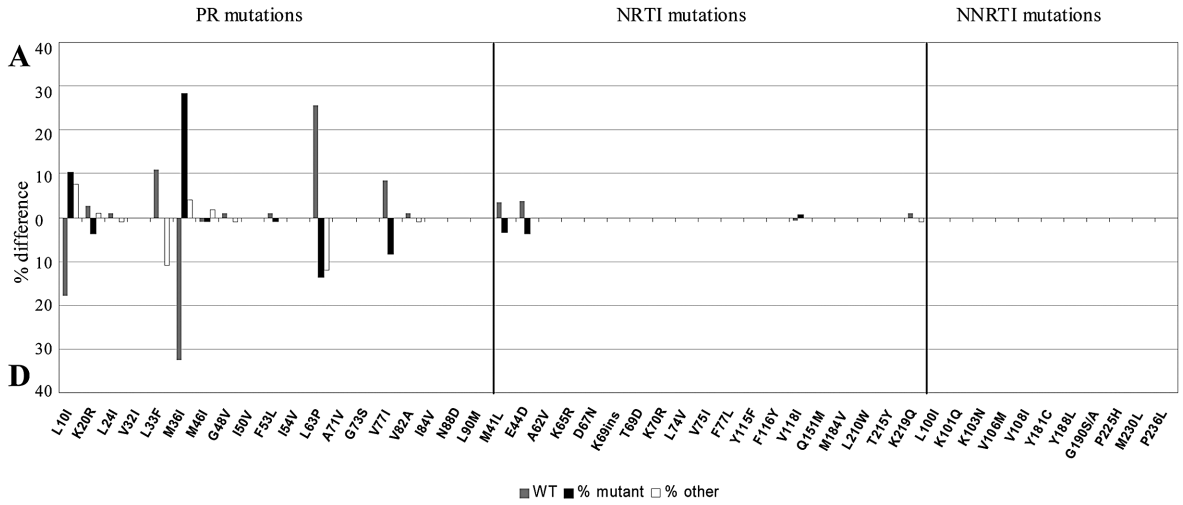

FIG. 3. Amino acid variation in Pol. Differences between subtypes A and D were expressed as the normalised frequency of WT, mutant and any other amino acid at each IAS-USA mutation position for subtype A, subtracted from those for subtype D. Columns above the x-axis represent a greater frequency in subtype A, whereas those below the x-axis represent a greater frequency in subtype D. Where amino acid frequency at each IAS-USA position are equivalent, between subtypes A and D, the frequency scores cancel out.

between *gag* and *env* within this cohort and indeed other African study cohorts where subtypes A and D predominate,[37] is relatively less frequent within the *pol* gene. While double the level of A/D intersubtype recombination has previously been reported in a cohort of Ugandan women enrolled in the HIVNET 012 Trial (12%),[13] this remains significantly lower than the rates estimated for the *env* and *gag* genes. This may be, in part, because the viral structural products of the *env* and *gag* genes are under greater selective pressure from the immune system, compared to the enzymatic gene products of *pol*. If it is this lack of host-driven selective pressure that has historically limited *pol* recombination in this cohort; we may anticipate a rise following widespread ARV usage. It is also the case, however, that low levels of dual infection have been observed in this cohort (P. Kaleebu, personal communication), which may reduce the potential for generation of complex *pol* recombinants.

Recombination is potentially a powerful means of overcoming fitness deficits imparted by the presence of resistance mutations in this region of the viral genome. As currently there is a low level and little conservation between *pol* recombinants within this population, it will be interesting and important to monitor *pol* recombination structure and frequency (both inter- and intrasubtype) over time, with the increasing introduction of ARV therapy. This may be especially important in countries such as Uganda, since the absence of close patient monitoring may result in prolonged periods of virological failure, prior to observed clinical failure. An ongoing database analysis of treated subtype A- and D-infected patients in the UK should also shed light on this.

Analysis of amino acid variation in this dataset at resistance positions revealed that while variation was minimal between subtypes A and D in RT, there is appreciable variation in PR

TABLE 2. DETAILS OF IAS-USA MUTATION POSITIONS SYNONYMOUSLY VARIABLE BETWEEN SUBTYPES A AND D

			Codon (%)[a]		
Gene	*Position*	*Amino acid*	*Subtype A*	*Subtype D*	*Difference*
PR	10	Leucine (L)	CTT (89)	CTC (53)	T → C
PR	53	Phenylalanine (F)	TTC (16)	TTT (97)	C → T
PR	63	Leucine (L)	CTT (89)	CTC (42)	T → C
RT	62	Alanine (A)	GCT (98)	GCC (87)	T → C
RT	67	Aspartic acid (D)	GAT (49)	GAC (97)	T → C
RT	70	Lysine (K)	AAA (100)	AAG (90)	A → G
RT	77	Phenylalanine (F)	TTT (41)	TTC (92)	T → C
RT	100	Leucine (L)	TTA (80)	CTA (80)	T → C
RT	151	Glutamine (Q)	CAG (86)	CAA (85)	G → A
RT	215	Threonine (T)	ACT (95)	ACC (97)	T → C

[a]Denotes the percentage of sequences possessing the shown codon at the position in question.

at minor mutation positions including 10, 33, 36, 63, and 77. While in isolation these changes do not have a direct impact on protease inhibitor (PI) resistance per se, there is some evidence that these minor mutations are similarly compensatory for many PIs.[38] As a result, the preexistence of minor compensatory mutations in subtypes A and D might result in the faster emergence of viruses resistant to PIs. Until the spectrum of accessory mutations for subtype A and D viruses is more fully characterized, however, this cannot be confirmed.

It is analysis of codon usage that shows much greater variability between subtypes A and D. Strikingly, while none of the IAS mutation positions in RT was variable at the amino acid level (Fig. 3), 24% had differential codon usage between subtypes A and D, including positions 62, 67, 70, 77, 151, 215 (NRTI resistance associated), and 100 (NNRTI resistance associated). Positions 10, 53, and 63 in PR were also variable.

Of greatest interest is the variability at position 151 of RT. The mutation Q151M, which emerges in approximately 16% of patients treated with AZT in combination with either zalcitabine (ddC) or didanosine (ddI), confers cross-resistance to all NRTIs.[39] The mutational pathway, described by Kijak *et al.*,[40] involves leucine (L) as an intermediate: Q to L to M. As described in this study, position 151 of RT is encoded by CAG in 86% of subtype A viruses, which gives the mutational pathway Q (CAG) to L (CTG) to M (ATG).[41] In the subtype D viruses studied only 15% have CAG at this position; 85% have CAA for which the path to M, through L, is longer: Q (CAA) to L (CTA) to L (CTG) to M (ATG). It may be hypothesized, therefore, that such a difference may cause resistance to develop more slowly or take a different path in subtype D. This example, however, is "ideal" as the resistance-associated amino acid has only one potential codon (ATG). To truly hypothesize about the importance of synonymous variability at other resistance mutation positions between subtypes, the variability at the codon when encoding a resistance-associated amino acid must also be known. A large database analysis of treated subtype A- and D-infected patients is underway in order to characterize codon variability at these positions in resistant viruses.

In conclusion, we have shown that while in this cohort there is no natural high-level resistance to antiretroviral therapy, by virtue of intersubtype genetic divergence subtypes A and D may possess different genetic barriers to be overcome in order to achieve resistance. It will only be by assessing codon usage at resistance positions in resistant viruses, however, that such an analysis can be complete. It will also be important to monitor recombination, given the power of such a process to drive evolution and presumably increase viral fitness and resistance in treated populations in countries like Uganda. It is likely to be in resource-poor settings where access to treatment and monitoring and adherence to treatment may be suboptimal, that the importance of genotypic variation may be more apparent.

ACKNOWLEDGMENTS

The authors would like to thank the members of the Rural Clinical Cohort for their participation in this and other studies. We also thank Deenan Pillay for critical manuscript review. This work was supported by the MRC/UVRI Uganda Research Unit on AIDS, which is funded by the MRC (UK) and the Department for International Development (DFID) UK. C.V.G. is funded by the MRC, UK.

REFERENCES

1. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S, and Korber B: HIV-1 nomenclature proposal. Science 2000;288(5463):55–56.
2. Gomez-Carrillo M, Quarleri JF, Rubio AE, Carobene MG, Dilernia D, Carr JK, and Salomon H: Drug resistance testing provides evidence of the globalization of HIV type 1: A new circulating recombinant form. AIDS Res Hum Retroviruses 2004;20(8):885–888.
3. Kijak GH and McCutchan FE: HIV diversity, molecular epidemiology, and the role of recombination. Curr Infect Dis Rep 2005;7(6):480–488.
4. Kantor R and Katzenstein D: Polymorphism in HIV-1 non-subtype B protease and reverse transcriptase and its potential impact on drug susceptibility and drug resistance evolution. AIDS Rev 2003;5(1):25–35.
5. Yirrell DL, Shaw L, Burns SM, Cameron SO, Quigg M, Campbell E, and Goldberg D: HIV-1 subtype in Scotland: The establishment of a national surveillance system. Epidemiol Infect 2004;132(4): 693–698.
6. Becker-Pergola G, Kataaha P, Johnston-Dow L, Fung S, Jackson JB, and Eshleman SH: Analysis of HIV type 1 protease and reverse transcriptase in antiretroviral drug-naive Ugandan adults. AIDS Res Hum Retroviruses 2000;16(8):807–813.
7. Weidle PJ, Kityo CM, Mugyenyi P, Downing R, Kebba A, Pieniazek D, Respess R, Hertogs K, De Vroey V, Dehertogh P, Bloor S, Larder B, and Lackritz E: Resistance to antiretroviral therapy among patients in Uganda. J Acquir Immune Defic Syndr 2001;26(5):495–500.
8. Gao Y, Paxinos E, Galovich J, Troyer R, Baird H, Abreha M, Kityo C, Mugyenyi P, Petropoulos C, and Arts EJ: Characterization of a subtype D human immunodeficiency virus type 1 isolate that was obtained from an untreated individual and that is highly resistant to nonnucleoside reverse transcriptase inhibitors. J Virol 2004;78(10):5390–5401.
9. Jackson JB, Becker-Pergola G, Guay LA, Musoke P, Mracna M, Fowler MG, Mofenson LM, Mirochnick M, Mmiro F, and Eshleman SH: Identification of the K103N resistance mutation in Ugandan women receiving nevirapine to prevent HIV-1 vertical transmission. AIDS 2000;14(11):F111–115.
10. Eshleman SH, Guay LA, Wang J, Mwatha A, Brown ER, Musoke P, Mmiro F, and Jackson JB: Distinct patterns of emergence and fading of K103N and Y181C in women with subtype A vs. D after single-dose nevirapine: HIVNET 012. J Acquir Immune Defic Syndr 2005;40(1):24–29.
11. Richard N, Juntilla M, Abraha A, Demers K, Paxinos E, Galovich J, Petropoulos C, Whalen CC, Kyeyune F, Atwine D, Kityo C, Mugyenyi P, and Arts EJ: High prevalence of antiretroviral resistance in treated Ugandans infected with non-subtype B human immunodeficiency virus type 1. AIDS Res Hum Retroviruses 2004;20(4): 355–364.
12. Gupta RK, Chrystie IL, O'Shea S, Mullen JE, Kulasegaram R, and Tong CY: K65R and Y181C are less prevalent in HAART-experienced HIV-1 subtype A patients. AIDS 2005;19(16):1916–1919.
13. Eshleman SH, Gonzales MJ, Becker-Pergola G, Cunningham SC, Guay LA, Jackson JB, and Shafer RW: Identification of Ugandan HIV type 1 variants with unique patterns of recombination in pol involving subtypes A and D. AIDS Res Hum Retroviruses 2002;18(7):507–511.

14. Morgan D, Malamba SS, Maude GH, Okongo MJ, Wagner HU, Mulder DW, and Whitworth JA: An HIV-1 natural history cohort and survival times in rural Uganda. AIDS 1997;11(5):633–640.
15. Morgan D, Rutebemberwa A, Malamba S, Ross A, Whitworth J, Kaleebu P, and Gotch F: HIV-1 RNA levels in an African population-based cohort and their relation to CD4 lymphocyte counts and World Health Organization clinical staging. J Acquir Immune Defic Syndr 1999;22(2):167–173.
16. Kaleebu P, Whitworth J, Hamilton L, Rutebemberwa A, Lyagoba F, Morgan D, Duffield M, Biryahwaho B, Magambo B, and Oram J: Molecular epidemiology of HIV type 1 in a rural community in southwest Uganda. AIDS Res Hum Retroviruses 2000;16(5):l393–401.
17. Kaleebu P, Ross A, Morgan D, Yirrell D, Oram J, Rutebemberwa A, Lyagoba F, Hamilton L, Biryahwaho B, and Whitworth J: Relationship between HIV-1 Env subtypes A and D and disease progression in a rural Ugandan cohort. AIDS 2001;15(3):293–299.
18. Yirrell DL, Kaleebu P, Morgan D, Watera C, Magambo B, Lyagoba F, and Whitworth J: Inter- and intra-genic intersubtype HIV-1 recombination in rural and semi-urban Uganda. AIDS 2002;16(2): 279–286.
19. Senkaali D, Muwonge R, Morgan D, Yirrell D, Whitworth J, and Kaleebu P: The relationship between HIV type 1 disease progression and V3 serotype in a rural Ugandan cohort. AIDS Res Hum Retroviruses 2004;20(9):932–937.
20. Yirrell DL, Kaleebu P, Morgan D, Hutchinson S, and Whitworth JA: HIV-1 subtype dynamics over 10 years in a rural Ugandan cohort. Int J STD AIDS 2004;15(2):103–106.
21. Boden D, Hurley A, Zhang L, Cao Y, Guo Y, Jones E, Tsay J, Ip J, Farthing C, Limoli K, Parkin N, and Markowitz M: HIV-1 drug resistance in newly infected individuals. JAMA 1999;282(12): 1135–1141.
22. Little SJ, Holte S, Routy JP, Daar ES, Markowitz M, Collier AC, Koup RA, Mellors JW, Connick E, Conway B, Kilby M, Wang L, Whitcomb JM, Hellmann NS, and Richman DD: Antiretroviral-drug resistance among patients recently infected with HIV. N Engl J Med 2002;347(6):385–394.
23. Weidle PJ, Downing R, Sozi C, Mwebaze R, Rukundo G, Malamba S, Respess R, Hertogs K, Larder B, Ochola D, Mermin J, Samb B, and Lackritz E: Development of phenotypic and genotypic resistance to antiretroviral therapy in the UNAIDS HIV Drug Access Initiative—Uganda. AIDS 2003;17(Suppl 3):S39–48.
24. Kantor R, Zijenah LS, Shafer RW, Mutetwa S, Johnston E, Lloyd R, von Lieven A, Israelski D, and Katzenstein DA: HIV-1 subtype C reverse transcriptase and protease genotypes in Zimbabwean patients failing antiretroviral therapy. AIDS Res Hum Retroviruses 2002;18(18):1407–1413.
25. Mulder DW, Nunn AJ, Kamali A, Nakiyingi J, Wagner HU, and Kengeya-Kayondo JF: Two-year HIV-1-associated mortality in a Ugandan rural population. Lancet 1994;343(8904):1021–1023.
26. Higgins DG: CLUSTAL V: Multiple alignment of DNA and protein sequences. Methods Mol Biol 1994;25:307–318.
27. Kumar S, Tamura K, Jakobsen IB, and Nei M: MEGA2: Molecular evolutionary genetics analysis software. Bioinformatics 2001; 17(12):1244–1245.
28. Gale CV, Myers R, Tedder RS, Williams IG, and Kellam P: Development of a novel human immunodeficiency virus type 1 subtyping tool, Subtype Analyzer (STAR): Analysis of subtype distribution in London. AIDS Res Hum Retroviruses 2004;20(5):457–464.
29. Myers RE, Gale CV, Harrison A, Takeuchi Y, and Kellam P: A statistical model for HIV-1 sequence classification using the subtype analyser (STAR). Bioinformatics 2005;21(17):3535–3540.
30. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, and Ray SC: Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 1999;73(1):152–160.
31. Johnson VA, Brun-Vezinet F, Clotet B, Conway B, Kuritzkes DR, Pillay D, Schapiro J, Telenti A, and Richman D: Update of the Drug Resistance Mutations in HIV-1: 2005. Top HIV Med 2005;13(1):51–57.
32. Spira S, Wainberg MA, Loemba H, Turner D, and Brenner BG: Impact of clade diversity on HIV-1 virulence, antiretroviral drug sensitivity and drug resistance. J Antimicrob Chemother 2003; 51(2):229–240.
33. Kantor R, Katzenstein DA, Efron B, Carvalho AP, Wynhoven B, Cane P, Clarke J, Sirivichayakul S, Soares MA, Snoeck J, Pillay C, Rudich H, Rodrigues R, Holguin A, Ariyoshi K, Bouzas MB, Cahn P, Sugiura W, Soriano V, Brigido LF, Grossman Z, Morris L, Vandamme AM, Tanuri A, Phanuphak P, Weber JN, Pillay D, Harrigan PR, Camacho R, Schapiro JM, and Shafer RW: Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: Results of a global collaboration. PLoS Med 2005;2(4):e112.
34. Vergne L, Peeters M, Mpoudi-Ngole E, Bourgeois A, Liegeois F, Toure-Kane C, Mboup S, Mulanga-Kabeya C, Saman E, Jourdan J, Reynes J, and Delaporte E: Genetic diversity of protease and reverse transcriptase sequences in non-subtype-B human immunodeficiency virus type 1 strains: Evidence of many minor drug resistance mutations in treatment-naive patients. J Clin Microbiol 2000;38(11):3919–3925.
35. Gonzales MJ, Machekano RN, and Shafer RW: Human immunodeficiency virus type 1 reverse-transcriptase and protease subtypes: Classification, amino acid mutation patterns, and prevalence in a northern California clinic-based population. J Infect Dis 2001;184(8):998–1006.
36. Harris ME, Serwadda D, Sewankambo N, Kim B, Kigozi G, Kiwanuka N, Phillips JB, Wabwire F, Meehen M, Lutalo T, Lane JR, Merling R, Gray R, Wawer M, Birx DL, Robb ML, and McCutchan FE: Among 46 near full length HIV type 1 genome sequences from Rakai District, Uganda, subtype D and AD recombinants predominate. AIDS Res Hum Retroviruses 2002;18(17):1281–1290.
37. Renjifo B, Gilbert P, Chaplin B, Vannberg F, Mwakagile D, Msamanga G, Hunter D, Fawzi W, and Essex M: Emerging recombinant human immunodeficiency viruses: Uneven representation of the envelope V3 region. AIDS 1999;13(13):1613–1621.
38. Erickson JW, Gulnik SV, and Markowitz M: Protease inhibitors: Resistance, cross-resistance, fitness and the choice of initial and salvage therapies. AIDS 1999;13(Suppl A):S189–204.
39. Shirasaka T, Kavlick MF, Ueno T, Gao WY, Kojima E, Alcaide ML, Chokekijchai S, Roy BM, Arnold E, Yarchoan R, *et al*.: Emergence of human immunodeficiency virus type 1 variants with resistance to multiple dideoxynucleosides in patients receiving therapy with dideoxynucleosides. Proc Natl Acad Sci USA 1995;92(6):2398–2402.
40. Kijak GH, Currier JR, Tovanabutra S, Cox JH, Michael NL, Wegner SA, Birx DL, and McCutchan FE: Lost in translation: Implications of HIV-1 codon usage for immune escape and drug resistance. AIDS Rev 2004;6(1):54–60.
41. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, and Shafer RW: Human immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res 2003;31(1):298–303.

Address reprint requests to:
Catherine V. Gale
Centre of Virology
University College London
Windeyer Institute
46 Cleveland Street
London W1T 4JF, UK

E-mail: c.gale@ucl.ac.uk