Short communication

Characterization of drug resistance in antiretroviral-treated patients infected with HIV-1 CRF02_AG and AGK subtypes in Mali and Burkina Faso

Mohamed Sylla[1†], Annie Chamberland[1,2†], Catherine Boileau[3], Hamar A Traoré[4], Souleyman Ag-Aboubacrine[4], Mamadou Cissé[5], Samual Koala[6], Joseph Drabo[7], Ismael Diallo[7], Pascal Niamba[7], Daniel Tremblay-Sher[1], Nimâ Machouf[3,8], Sélim Rashed[9], David C Nickle[10], Vinh-Kim Nguyen[3,8] and Cécile L Tremblay[1,2] on behalf of the ATARAO group[1]*

[1]Département de Microbiologie et Immunologie, Université de Montréal, Montréal, Canada
[2]Centre Hospitalier de l'Université de Montréal, Montréal, Canada
[3]Département de Médecine sociale et préventive, Université de Montréal, Montréal, Canada
[4]Hôpital national du Point-G, Bamako, Mali
[5]CÉSAC, Bamako, Mali
[6]Centre Oasis/AAS, Ouagadougou, Burkina Faso
[7]Centre hospitalier national Yalgado Ouédraogo, Ouagadougou, Burkina Faso
[8]Clinique l'Actuel, Montréal, Canada
[9]Unité de Santé Internationale, Université de Montréal and Centre Hospitalier Maisonneuve-Rosemont, Montréal, Canada
[10]Department of Microbiology, University of Washington School of Medicine, Seattle, Washington, USA

[†]These authors contributed equally to this manuscript
*Corresponding author: E-mail: c.tremblay@umontreal.ca

Background: In a multicentred cohort of patients on antiretroviral therapy (ART) in Burkina Faso and Mali, we analysed the prevalence of HIV drug resistance mutations in patients failing a modified directly observed therapy (mDOT) protocol.

Methods: Patients on ART >6 months and with viral load (VL) >500 copies/ml were enrolled in a mDOT protocol. Genotypic resistance testing was performed on pre- and post-mDOT plasma samples of patients who still had VL >500 copies/ml after mDOT.

Results: Eight hundred and one patients from seven sites participated in the study. One hundred and thirteen patients (14.1%) had VL >500 copies/ml. Most patients were treated with lamivudine along with zidovudine or stavudine and efavirenz or nevirapine. Genotypes were available for 46 patients. The predominant HIV-1 subtypes were CRF02_AG in 26 (56.5%) and AGK/K/AK in 12 (26.1%) patients. The prevalence of drug resistance mutations by class were as follows for nucleoside reverse transcriptase inhibitors: 184I/V (82.6%), 215Y/F (32.6%), 219E/Q (19.6%), 70R (19.6%), 67N (21.7%), 41L (15.2%) and 151M(2.2%). For non-nucleoside reverse transcriptase inhibitors the prevalence was: 103N (50%) and 181C/I (19.6%). Phylogenetic analysis showed that, although the genetic distances were small among isolates, there was no clustering of a particular subtype in a specific region and that the high prevalence of AGK subtype in our drug-resistant population was not due to a circulating resistant strain.

Conclusion: Although CRF02_AG is the dominant clade in the Burkina Faso/Mali region, isolates with subtype K reverse transcriptase were frequent in our cohort. Drug resistance mutation pathways in subtype K reverse transcriptase need to be further evaluated in a larger cohort of non-B HIV-infected individuals.

Introduction

A rapidly growing number of HIV-infected individuals in developing countries now have access to antiretroviral therapy (ART). The World Health Organization estimated that, as of March 2006, 1.3 million people were receiving ART in Africa in middle- and low-income countries, a figure that more than tripled over

the past 2 years [1]. These numbers are likely to expand even more rapidly as large interventions such as the US Presidential Emergency Program for AIDS Relief (PEPFAR), UNAIDS, the Global Fund and other organizations ramp up operations. As antiretroviral drugs are more widely used in countries where non-B HIV-1 subtype infections are prevalent, it is important to assess the drug resistance mutation pathways emerging among various subtypes. Some data suggest that non-B isolates show a different pattern of resistance mutations than the B subtype [2–5]. Reports have shown that mutation V106M confers resistance to non-nucleoside reverse transcriptase inhibitors (NNRTIs) in subtype C HIV [6,7]. Furthermore, the D30N mutation is not preferentially selected by HIV-1 subtype C in the development of resistance to nelfinavir [8]. Differential evolution of resistance patterns among subtypes has important implications for clinical management and drug development. Differences in mutational pathways by different clades could affect our choices of first- and second-line regimens. They could also shed light on the molecular mechanisms by which HIV acquires and maintains resistance to antiretroviral agents. However, little data exist on drug resistance mutation pathways in non-B subtypes. Here, we describe a group of patients in Burkina Faso and Mali in Western Africa. These patients harboured non-B subtypes with resistance to antiretrovirals (ARVs), and were studied in the context of an intervention to support adherence to treatment. We report the specific subtypes and mutations found as well as treatment histories for the patients from whom they were derived.

Methods

Study population and design

Eight hundred and one patients from Mali (398) and Burkina Faso (403) were enrolled in a study to assess adherence to ART. Viral loads were measured in all 801 patients, who had received ARVs for >6 months. In 113 patients whose viral load was >500 c/ml, a 1-month adherence intervention (modified directly observed therapy; mDOT) was proposed; 77/113 patients agreed and completed the intervention. Viral load was measured 1 month after the intervention. Genotypic analysis of drug resistance mutations was performed before and within a month following the completion of the adherence intervention.

Viral load

Viral load was measured with COBAS AMPLICOR HIV-1 MONITOR Test, version 1.5 (Roche Diagnostics, Branchburg, NJ, USA), according to the manufacturer's instructions.

Genotypic analysis

Plasma viral RNA was extracted using QIAamp Viral Mini Spin Kit (Qiagen, Mississauga, Ontario, Canada) and concentrated three times, according to the manufacturer's instructions. cDNA synthesis was performed using SuperScript III One-Step reverse transcriptase (RT)-PCR System with Platinum Taq Polymerase (Invitrogen, Carlsbad, CA, USA) with primers covering Gag–Pol domains (Gag-1824F, T_m 65°C and Pol-4295R, T_m 63.5°C) using a 30 min cycle at 50°C. Amplification conditions were as follows: an initial 2 min cycle at 94°C for denaturation, 40 cycles with three steps (15 s at 94°C, 30 s at 58°C and 5 min at 68°C) and a final extension cycle for 5 min at 68°C. Samples with low viral load or that were difficult to amplify were submitted to nested PCR with primers (Gag-1824F and Pol-4002R or Pol-2137F and Pol-4295R) in order to obtain the expected band. PCR products were separated on a 1% agarose gel to confirm the presence of a 2020 bp band. They were purified using QIAprep Spin Miniprep Kit 50 (Qiagen). Sequencing was performed at Génome Québec (McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada) using six primers covering *pol*. All nucleotide sequences were submitted to GenBank and provided with accession numbers (DQ001615–DQ991673, EF010976, EF010977, EF525643–EF525650, EF525652–EF525667 and EF588323–EF588326).

Data analysis

Sequences were analysed using Sequencer 4.5 from Gene Codes Corporation software (Ann Arbor, MI, USA). Analysis of drug resistance mutations and subtype analysis were performed using the Stanford University HIV Resistance Database. Results were confirmed using the Los Alamos HIV blast tool (http://www.hiv.lanl.gov/content/hiv-db/BASIC_BLAST/basic_blast.html; Los Alamos National Laboratory [LANL]). The comparison of drug mutations according to subtype was analysed using a non-parametric statistical analysis, using Mann–Whitney test on SPSS 12.0 for Windows. (Chicago, IL, USA).

Phylogenetic analysis

Sequences were aligned using Clustal W version 1.83 [9] followed by manual alignment editing using BioEdit version 7.0.4.1 (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Pairwise alignment parameters were set to the dynamic 'slow-accurate' programming, using 10 as the gap-opening penalty and 0.1 as the gap extension penalty. Multiple alignment parameters included a gap extension penalty equal to 0.2 [10]. Subtype reference sequences were obtained from LANL HIV sequence database (http://www.hiv.lanl.gov/). We estimated a maximum likelihood (ML) tree with our

data (AG and AGK) and each available sequence from the LANL HIV database from the same subtypes and gene region (97 sequences). Phylogenetic analysis was done by first using PAUP* to estimate a starting NJ tree and then used the swapping algorithm SPR under an HKY+G+I model of evolution.

Population substructure

We used two methods to determine whether our study population was distinct from previously described AG and AGK clades. First, we used a cladistic-based method [11,12] to determine if HIV sequences from within our study population had a history of migration between the local population and AG and AGK strains found elsewhere. We therefore estimated a maximum likelihood (ML) tree with our data (AG and AGK) and as many database sequences from the LANL HIV database of the same subtypes and gene region (97 sequences) using PAUP*(Sinaver Associates Inc. Sunderland, MA, USA). [13]. We then mapped the number of potential migration events between populations onto the phylogeny using the software program MacClade [14].

Next, we used a genetic distance-based method called nearest-neighbour statistic (S_{nn}), which is reported to be one of the most statistically powerful methods to detect a restriction of gene flow among populations using genetic distance data [15]. This method enumerates the number of individuals that are the closest in terms of genetic distance (that is, nearest neighbour [nn]) to each individual sampled from both putative populations and then simply sums the number of nn individuals within each populations divided by the total number of nn across all putative groups arriving at the S_{nn} test statistic. *P*-values are determined non-parametrically by randomizing the genetic distance between individuals.

Results

To be eligible for ART, most patients had to have $CD4^+$ T-cell counts ≤200 cells/mm^3. Most patients were receiving a NNRTI plus two nucleoside reverse transcriptase inhibitors (NRTI), with stavudine, lamivudine and nevirapine being the most commonly used drugs. Of the 801 patients enrolled on therapy, viral load was measured in 798 patients. After 6 months of therapy, 598 patients (74.9%) had plasma HIV-1 RNA <50 copies/ml. Adherence intervention, including counselling and home visits, was performed in 77/113 patients with viral load >500 copies/ml at that time. Genotyping was attempted in these 77 patients before and after this intervention. Thirteen more patients achieved an undetectable viral load after this intervention; consequentally, a total of 611 patients had plasma HIV-1 RNA levels <50 copies/ml (76.5%).

Resistance testing was successfully performed in 46/113 patients with viral load >500 copies/ml. Some viral sequences could not be amplified because of the poor quality of the samples. Characteristics of these patients are summarized in Table 1. In both countries, the predominant form of virus was the circulating recombinant form CRF02_AG, accounting for 56.5% of the patients tested, followed by CRF_AGK in 26.1% of patients. These viruses were composed of AG subtype sequences in protease and K sequence in RT. One patient's virus was K subtype over both protease and RT and another patient's was AK. Other subtypes represented in single patients were: A, C, AGG (which differs

Table 1. Patient characteristics

	CRF02_AG	CRF_AGK/K	Others	Total
Female, *n* (%)	11 (44)	10 (40)	4 (16)	25 (100)
Male, *n* (%)	12 (57.1)	4 (19.04)	5(23.8)	21(100)
Subtype distribution, *n* (%)	26 (56.5)	12 (26.1)	8 (17.4)	46 (100)
Mean $CD4^+$ T-cell count, cells/mm^3	161	210	197	189 (100)
Drug regimen, *n* (%)*				
ZDV + 3TC + NVP	5 (19.2)	5 (41.6)	2 (25)	12 (100)
ZDV + 3TC + EFV	13 (50)	1 (8.3)	0	14 (100)
D4T + 3TC + NVP	0	3 (25)	1 (12.5)	4 (100)
D4T + 3TC + EFV	2 (7.7)	0	0	2 (100)
DDI + ZDV + 3TC	1 (3.8)	1 (8.3)	1 (12.5)	3 (100)
DDI + D4T + IDV	2 (7.7)	1 (8.3)	0	3 (100)
DDI + ZDV + IDV	0	0	1 (12.5)	1 (100)
D4T + 3TC + IDV	1 (3.8)	1 (8.3)	1 (12.5)	3 (100)
ZDV + 3TC + IDV	2 (7.7)	0	1 (12.5)	3 (100)
DDI + 3TC + IDV	0	0	1 (12.5)	1 (100)

*DDI, didanosine; D4T, stavudine; EFV, efavirenz; IDV, indinavir; NVP, nevirapine; 3TC, lamivudine; ZDV, zidovudine.

from our other AGs only in the protease domain with clade G), GH, BD and B; CRF01_AE was found in two patients. Some previous reports indicated that CRF02_AG is common in Western Africa [16–18], whereas subtype K is rare [18]. Most of our AGK isolates were from Burkina Faso (8/10), which is consistent with recent reports describing CRF 06_cpx (a mosaic of A, G, J and K subtypes) and CRF 09_cpx in this country [19,20]. There was no known epidemiological link among the patients as assessed by our demographic questionnaire and our phylogenetic analysis. However, the genetic distances between all AGK isolates were small. Our subtype K isolate, originating from Mali, clustered with the only two other subtype K isolates described in the literature (from Cameroon and Congo) as well as some CRF 09_cpx strains. Isolates from Mali and from Burkina Faso did not cluster preferentially within a country (Figure 1).

Resistance mutations found in the 46 patients are shown in Figure 2. The most common NRTI resistance mutation encountered was the 184I/V (82.6% of patients). Thymidine-associated mutations (TAMs) were detected in 52.1% of patients: 21.7% had mutations characteristic of the TAM1 pathway (M41L, L210W, T215Y) and 30.4% of the TAM2 pathway (D67N, K70R, T215F, K219Q/E). Of interest, a greater proportion of mutations associated with the TAM2 pathway was observed in patients harbouring viruses of the AGK or K subtypes compared with CRF02_AG subtypes (41.7% versus 19.2%, P=0.150), although this trend was not statistically significant perhaps due to the small sample size. One patient had the Q151M MDR mutation complex. Very little viral evolution occurred between the time points. Only three patients acquired additional NRTI mutations between the two time points (Figure 3a).

Groups of 16 patients each were on NNRTI-containing regimens, either nevirapine or efavirenz. However, 36 patients had NNRTI resistance-associated mutations: 31/32 patients on an NNRTI-containing regimen had NNRTI-associated mutations, and three patients on NRTI-only regimens and two patients on a PI-containing regimen had NNRTI mutations, possibly due to previous exposure to an NNRTI.

K103N was the most common (50%), followed by Y181C (19.6%), P225H (19.6%), G190A/S (17.4%) and K101E (13%). The K103N mutation was more common in CRF02_AG infections as opposed to CRF AGK/K infections (69.2 versus 33.3%, P=0.042). However, more patients with CRF02_AG were on efavirenz-containing regimens (56% compared with 17% of CRF AGK/K subtypes). K103N mutations were present in 13/16 (81%) patients on efavirenz-containing regimens (all of them were subtype AG), in contrast to 6/17 (35%) patients on nevirapine-containing regimens (4/5 AG and 2/8 AGK on nevirapine-containing regimens). K101E was predominantly identified in CRF AGK/K subtypes (25% versus 0% in CRF02AG subtype; Figure 3b).

Eleven patients received a PI-containing regimen (indinavir). Four of them harboured PI resistance-associated mutations (G48W, I54V, V82A/T). Several polymorphisms, such as K20I and M36I and L10V/M/I, were observed.

Drug resistance evolution

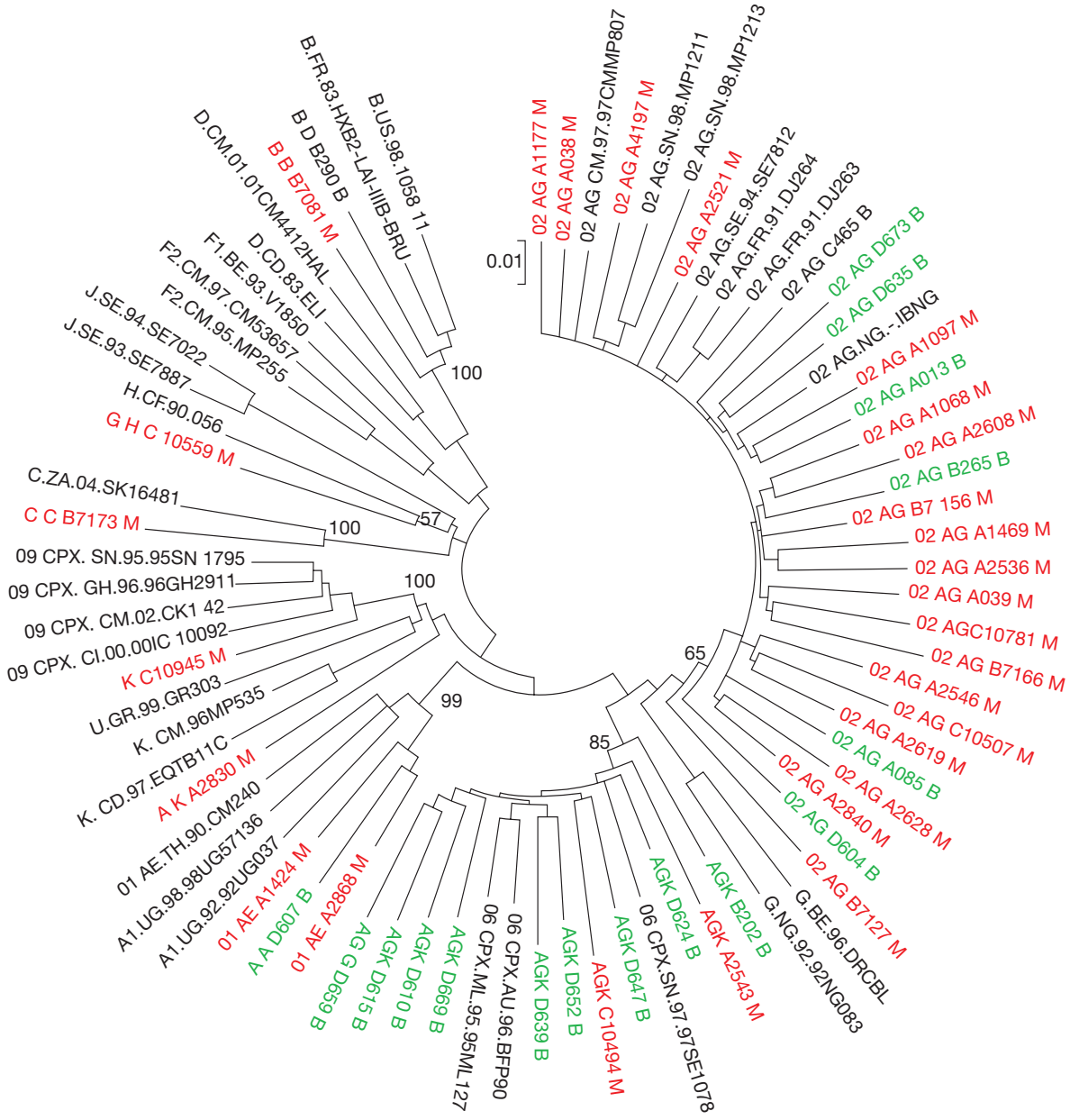
We used two methods to ascertain the probability that the observed drug resistance patterns in the population we studied was driven by transmitted drug resistance virus or by independent evolution of the drug-resistant genotypes. To test this hypothesis, we compared our observed number of drug-sensitive genotypes evolving into drug-resistant genotypes on the ML tree with the null hypothesis of random independent evolution of drug resistance. If drug resistance mutations were being transmitted in the population, we would expect our observation to be significantly outside the null distribution. To obtain the P-value, we compared the observed number of drug resistance mutations evolving on the ML tree to our null hypothesis, which is the distribution of same evolutionary measure across 1,000 independent random trees given the drug resistance mutations we observed.

The second method we used to determine if drug resistance was being transmitted among the individuals within our local study population was to derive ML estimates of the ancestral state of each independent cluster of viruses found in our study population.

We found that viruses in our study population are genetically distinct from previously described strains (Figure 1). First, we found that the ML tree topology supports the notion that our population is more isolated then we would expect by chance alone (P<0.001). We observed 18 migration events between our AGK and AG variants and similar CRFs found in the LANL database. We also found that the genetic distance data supported the same conclusion. The S_{nn} value was 0.64 and the probability of observing a value this small or smaller under the null of free gene flow is P=0.013.

Our cladistic-based method revealed that the number of times drug resistance evolved in our population appears to be no different than random expectation. If the population of infected HIV individuals were transmitting drug-resistant viruses to each other at appreciable frequencies, we would see many fewer evolutionary events on the phylogeny. The ancestral state would harbour drug resistance mutations and all of the descendants from the ancestor would also be drug resistant, thus requiring few changes on the phylogeny

Figure 1. Phylogenetic analysis of 46 viral sequences



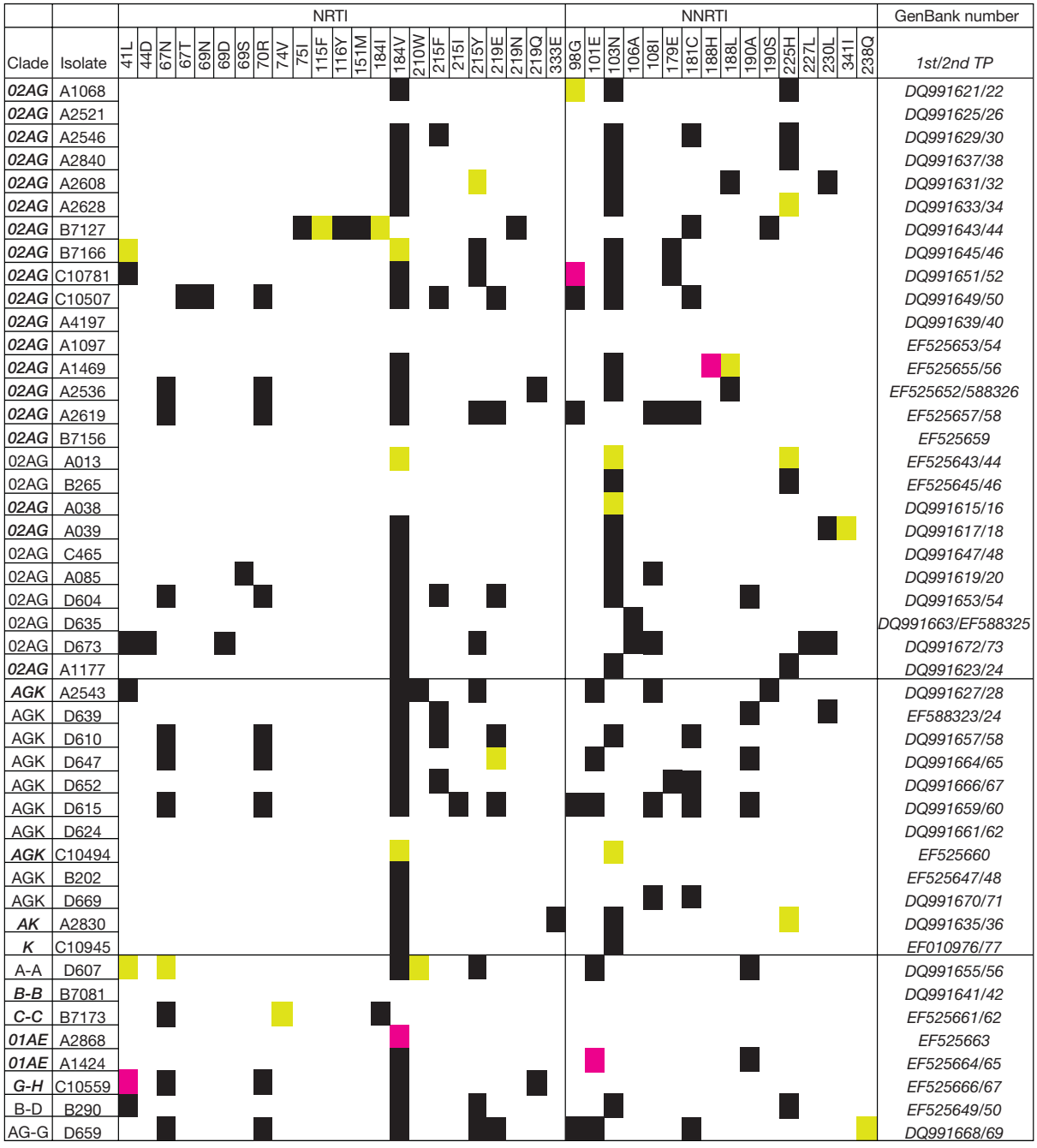
Sequences from Mali are shown in red and from Burkina Faso in green. Reference sequences are shown in black font. Bootstrap values >50 are indicated.

to adapt to this selective force. We observed 57 changes from drug-sensitive to drug-resistant genotypes in the AG dataset, which was no different than random expectation (P=0.7). In the AGK dataset, we observed 17 changes from drug-sensitive to drug-resistant genotypes (P=0.058). The marginal significance for the AGK set was perhaps due to the smaller sample size decreasing the statistical power. Ancestral states estimated on the ML tree were drug sensitive, further supporting the notion that antiretroviral drug resistance arrived in this population due to *de novo* mutations.

Discussion

This study is the largest to date on viruses encoding subtype K RT and reports some intriguing findings. First, in our group experiencing virological failure, patients with HIV AGK subtype were common. The

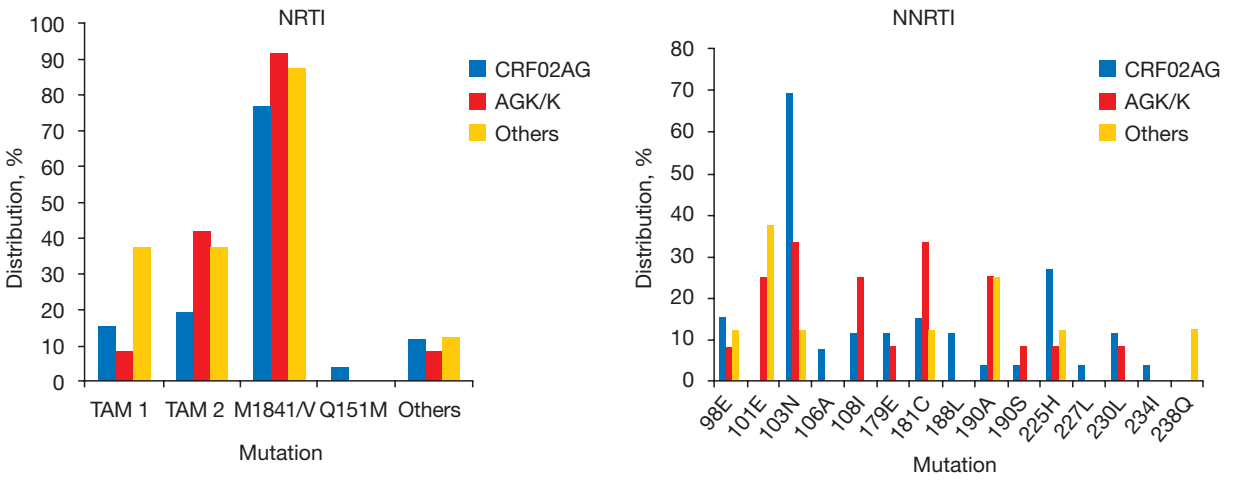
Figure 2. NRTI and NNRTI resistance mutations identified in 46 isolates



Black boxes correspond to mutations identified at both time points (before and 30 days after a 1 month-adherence intervention; TP), yellow boxes correspond to mutations identified at the second time point only, and pink boxes correspond to mutations identified at the first time points only. For three patients, only one time-point sequence is available. Clades in bold and italic = Mali; normal font = Burkina Faso. NNRTI, non-nucleoside reverse transcriptase inhibitor; NRTI, nucleoside reverse transcriptase inhibitor.

epidemiological significance of this finding is difficult to ascertain given that the majority of the 801 patients sampled had undetectable viral loads and, therefore, their subtype could not be determined. The prevalence of CRF_AGK in the Western African countries of Burkina Faso and Mali has not been extensively studied [16,18,21]. According to our phylogenetic analysis, it is unlikely that the high prevalence of AGK in our drug-resistant patients was due to primary infection with a circulating resistant strain. Furthermore, the patterns of

Figure 3. Distribution of NRTI and NNRTI drugs resistance mutations according to subtype



NNRTI, non-nucleoside reverse transcriptase inhibitor; NRTI, nucleoside reverse transcriptase inhibitor.

drug resistance mutations differed among patients. Although eight of the nine AGK-resistant patients were cared for at the same site, they came from different neighbourhoods. Furthermore, our phylogenetic analysis illustrates that although the genetic distances among the isolates are small, there was no clustering of a particular subtype in a specific region.

In our population, subtype K RT isolates tended to evolve NRTI resistance predominantly through the TAM2 pathway unlike our CRF02_AG strains and previous observations on subtype B strains [22,23]. This suggests that there may be differences in the molecular evolution of resistance in this clade. In subtype B, TAM1s are more common than TAM2s, possibly because of long-term trends in sequential use of ARVs, which has resulted in TAM1s evolving over time or possibly due to a fitness advantage conferred by certain mutations, such as T215Y or L210W, in a TAM1 background [24]. In our cohort of patients, treatment regimens were similar among subtypes. Furthermore, none of the patients who were on a stavudine regimen developed any TAMs. Therefore, the differences observed in the predominance of TAM2s could not be attributed to stavudine. Further evidence supporting the hypothesis that AGK RT may differ in its molecular evolution toward resistance stems from the finding that NNRTI-resistant AGK mutants were more likely to have Y181C than K103N, whereas the converse was true for CRF_AG02 mutants. However, patients with CRF02_AG had been more exposed to efavirenz-containing regimens.

Six patients did not have any resistance mutations, although only one of them was reported to be non-adherent. At the second time point, 20/46 patients harbouring resistance mutations had NNRTI mutation as well as TAMs and the 184V mutation, which compromised their future therapeutic options. The other 20 patients had various patterns of mutations: 14 patients had NNRTI and the 184V mutations, two only had the 184V mutation, one had NNRTI plus TAM, one had the Q151M complex and NNRTI mutations, one had 69S plus NNRTI mutations and one only had NNRTI resistance mutations.

This is the first analysis of the molecular evolution of a group of subtype K RT under ARV selective pressure. Our data suggest that subtype K viruses may evolve differently from other subtypes after exposure to RT inhibitors. This may have important consequences on treatment options and second-line regimens as more drugs become available to individuals in resource-poor settings. These findings indicate that further molecular studies in AGK viruses are warranted.

Acknowledgements

We wish to acknowledge G Bélanger-Jasmin for her laboratory support, the entire ATARAO team and the patients from Mali and Burkina-Faso for participating in this study.

Support

This work was supported by CIHR grant no 24141 and US Public Health Service Center grant P30 27757 to Dr James I Mullins for the support of DCN. This work was also supported by an unrestricted grant from Boehringer Ingelheim and le Réseau FRSQ-SIDA. CLT and VKN are CIHR New Investigators.

Disclosure statement

The authors declare no conflict of interests.

References

1. Kuehn BM. UNAIDS report: AIDS epidemic slowing, but huge challenges remain. *JAMA* 2006; **296:**29–30.
2. Ariyoshi K, Matsuda M, Miura H, Tateishi S, Yamada K, Sugiura W. Patterns of point mutations associated with antiretroviral drug treatment failure in CRF01_AE (subtype E) infection differ from subtype B infection. *J Acquir Immune Defic Syndr* 2003; **33:**336–342.
3. Bussmann H, Wester CW, Masupu KV, *et al*. Low CD4$^+$ T-lymphocyte values in human immunodeficiency virus-negative adults in Botswana. *Clin Diagn Lab Immunol* 2004; **11:**930–935.
4. Calazans A, Brindeiro R, Brindeiro P, *et al*. Low accumulation of L90M in protease from subtype F HIV-1 with resistance to protease inhibitors is caused by the L89M polymorphism. *J Infect Dis* 2005; **191:**1961–1970.
5. Kantor R, Katzenstein DA, Efron B, *et al*. Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. *PLoS Med* 2005; **2:**e112.
6. Brenner B, Turner D, Oliveira M, *et al*. A V106M mutation in HIV-1 clade C viruses exposed to efavirenz confers cross-resistance to non-nucleoside reverse transcriptase inhibitors. *Aids* 2003; **17:**F1–F5.
7. Quan Y, Brenner BG, Marlink RG, Essex M, Kurimura T, Wainberg MA. Drug resistance profiles of recombinant reverse transcriptases from human immunodeficiency virus type 1 subtypes A/E, B, and C. *AIDS Res Hum Retroviruses* 2003; **19:**743–753.
8. Grossman Z, Paxinos EE, Averbuch D, *et al*. Mutation D30N is not preferentially selected by human immunodeficiency virus type 1 subtype C in the development of resistance to nelfinavir. *Antimicrob Agents Chemother* 2004; **48:**2159–2165.
9. Thompson JD, Higgins DG, Gibson TJ. Clustal W improving the sensitivity of progressive multiple sequence alignment through sequence weighting poition specific gap penalties. *Nucleic Acids Res* 1994. **22:**4693–4680.
10. Kumar M, Jain SK, Pasha ST, Chattopadhaya D, Lal S, Rai A. Genomic diversity in the regulatory nef gene sequences in Indian isolates of HIV type 1: emergence of a distinct subclade and predicted implications. *AIDS Res Hum Retroviruses* 2006; **22:**1206–1219.
11. Slatkin M, Maddison WP. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 1989; **123:**603–613.
12. Slatkin M, Maddison WP. Detecting isolation by distance using phylogenies of genes. *Genetics* 1990; **126:**249–260.
13. Swofford DL. *PAUP* 4.0: Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sunderland, MA: Sinauer Associates, Inc. 2002.
14. Maddison WP, Maddison DR. *MacClade - Analysis of Phylogeny and Character Evolution - Version 4*. Sunderland, MA: Sinauer Associates, Inc. 2001.
15. Hudson RR. A new statistic for detecting genetic differentiation. *Genetics* 2000; **155:**2011–2014.
16. Laurent C, Bourgeois A, Faye MA, *et al*. No difference in clinical progression between patients infected with the predominant human immunodeficiency virus type 1 circulating recombinant form (CRF) 02_AG strain and patients not infected with CRF02_AG, in Western and West-Central Africa: a four-year prospective multicenter study. *J Infect Dis* 2002; **186:**486–492.
17. Osmanov S, Pattou C, Walker N, Schwardlander B, Esparza J. Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J Acquir Immune Defic Syndr* 2002; **29:**184–190.
18. Triques K, Bourgeois A, Vidal N, *et al*. Near-full-length genome sequencing of divergent African HIV type 1 subtype F viruses leads to the identification of a new HIV type 1 subtype designated K. *AIDS Res Hum Retroviruses* 2000; **16:**139–151.
19. Nadembega WM, Giannella S, Simpore J, *et al*. Characterization of drug-resistance mutations in HIV-1 isolates from non-HAART and HAART treated patients in Burkina Faso. *J Med Virol* 2006; **78:**1385–1391.
20. Tebit DM, Ganame J, Sathiandee K, Nagabila Y, Coulibaly B, Krausslich HG. Diversity of HIV in rural Burkina Faso. *J Acquir Immune Defic Syndr* 2006; **43:**144–152.
21. McCutchan FE, Carr JK, Bajani M, *et al*. Subtype G and multiple forms of A/G intersubtype recombinant human immunodeficiency virus type 1 in Nigeria. *Virology* 1999; **254:**226–234.
22. Cozzi-Lepri A, Ruiz L, Loveday C, *et al*. Thymidine analogue mutation profiles: factors associated with acquiring specific profiles and their impact on the virological response to therapy. *Antivir Ther* 2005; **10:**791–802.
23. Kuritzkes DR, Bassett RL, Hazelwood JD, *et al*. Rate of thymidine analogue resistance mutation accumulation with zidovudine- or stavudine-based regimens. *J Acquir Immune Defic Syndr* 2004; **36:**600–603.
24. Hu Z, Giguel F, Hatano H, Reid P, Lu J, Kuritzkes DR. Fitness comparison of thymidine analog resistance pathways in human immunodeficiency virus type 1. *J Virol* 2006; **80:**7020–7027.

Accepted for publication 2 October 2007