

The HIV Type 1 Epidemic in Bulgaria Involves Multiple Subtypes and Is Sustained by Continuous Viral Inflow from West and East European Countries

Marco Salemi,¹ Maureen M. Goodenow,¹ Stefania Montieri,² Tulio de Oliveira,³ Maria Mercedes Santoro,⁴ Danail Beshkov,⁵ Ivailo Alexiev,⁵ Ivailo Elenkov,⁶ Ivan Elenkov,⁷ Tsvetana Yakimova,⁸ Tonka Varleva,⁸ Giovanni Rezza,² and Massimo Ciccozzi²

Abstract

Little is known about the HIV-1 epidemic in Balkan countries. To fill the gap, we investigated the viral genetic diversity in Bulgaria, by sequencing and phylogenetic characterization of 86 plasma samples collected between 2002 and 2006 from seropositive individuals diagnosed within 1986–2006. Analysis of *pol* gene sequences assigned 51% of the samples to HIV-1 subtype B and 27% to subtype A1. HIV-1 subtype C, F, G, H, and a few putative recombinant forms were also found. Phylogenetic and molecular clock analysis showed a continuous exchange of subtype A and B between Bulgaria and Western as well as other Eastern European countries. At least three separate introductions of HIV-1 subtype A and four of HIV-1 subtype B have occurred within the past 25 years in Bulgaria. The central geographic location of Bulgaria, the substantial genetic heterogeneity of the epidemic with multiple subtypes, and the significant viral flow observed to and from the Balkan countries have the potential to modify the current HIV-1 epidemiological structure in Europe and highlight the importance of more extensive and continuous monitoring of the epidemic in the Balkans.

Introduction

CURRENT EPIDEMIOLOGICAL DATA indicate that the impact of the HIV-1 epidemic on Balkan countries, which are part of the central European region, is still limited.¹ In Bulgaria, up to the end of 2005, 689 cumulative cases of HIV infection, including 164 who had developed AIDS, have been reported. The annual number of HIV-1 diagnoses increased from 15–20 in the early 1990s to 40–60 within the past 5 years, with a peak of 91 diagnoses in 2005, following implementation of second generation surveillance systems. Most of the infections (87%) are attributable to heterosexual contact, whereas only 5.1% have been attributed to injecting drug use and 3.6% to blood transfusion. Moreover, most cases (about 75%) have been identified in the capital city of Sofia, followed by the towns of Buras, Plovdiv, and Varna.^{1,2}

According to recent serological surveys conducted in Bulgaria's largest cities and towns, the prevalence of HIV-1 infection is only 0.73% among commercial sex workers and 0.59% among injecting drug users, whereas no infection has been found among the Roma population.² However, the high incidence of HCV infection and syphilis and the high level of risk behavior in these population groups suggest that they are quite vulnerable to HIV infection and thus represent a potential source of infection.^{3–5}

Following the end of the cold war a decade ago, most Balkan countries have been affected by dramatic political and socioeconomic changes, which could contribute to the spread of HIV infection. Given the central geographic location of Bulgaria at the crossing point between Western Europe, Eastern Europe, and the Middle East, defining the origin and evolution of HIV in Bulgaria has obvious epi-

¹Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, Florida 32610.

²Epidemiology Unit, Department of Infectious Parasitic and Immunomediate Diseases, Istituto Superiore di Sanità, Rome, Italy.

³MRC Bioinformatics Unit, South African National Bioinformatics Institute, University of Western Cape, Cape Town, South Africa.

⁴Department of Experimental Medicine, University Roma Tor Vergata, Rome, Italy.

⁵National HIV Confirmatory Laboratory—National Center of Infectious and Parasitic Diseases, Sofia, Bulgaria.

⁶Hospital of Infectious Diseases, Sofia, Bulgaria.

⁷Faculty of Biology, University of Sofia, Sofia, Bulgaria.

⁸Ministry of Health, Program of "Prevention and Control of HIV/AIDS," Sofia, Bulgaria.

demological relevance. However, although basic information on the characteristics of HIV-1-infected individuals in Bulgaria is available, little is known about the HIV origin and distribution of different circulating subtypes. The small amount of information from other Balkan countries indicates that there is high genetic diversity, with subtype B being predominant in Yugoslavia and subtype A in Albania.^{6,7}

The objective of the present study was to investigate the molecular diversity and epidemiology of HIV-1 subtypes circulating in Bulgaria. To this end, we sequenced a fragment of the HIV-1 *pol* gene amplified from serum samples collected between 2002 and 2006 within a cohort of infected subjects under monitoring in different Bulgarian hospitals. By employing high-resolution phylogenetic and phylogeographic analysis, as well as molecular clock estimates, we obtained, for the first time, detailed information about the origin and the epidemiological history of the two most prevalent viral subtypes (HIV-1 A and B) in this Balkan country.

Materials and Methods

Patient samples

Eighty-six serum samples were obtained from persons who were diagnosed with HIV infection [58 treated with highly active antiretroviral therapy (HAART), 23 naive, and 5 no treatment data available] between 1986 and 2006 in Bulgaria. The serum samples were obtained from EDTA-whole blood (centrifugation at 3000 rpm at 18°C for 15 min), and stored at -80°C at the National HIV Confirmatory Laboratory in Sofia prior to genotyping analysis. The samples were linked to demographic and clinical data through an anonymous numerical code, in accordance with the ethical standards of the National HIV Confirmatory Laboratory of Sofia. We successfully performed HIV-1 genotyping on 83 (97%) of these samples (56 males, 27 females).

Pol gene sequencing

Sequencing and genotyping of the protease (PR) and reverse transcriptase (RT) of HIV-1 *pol* gene were performed in the National HIV laboratory in Sofia, Bulgaria by a commercially available diagnostic test. PR and RT *pol* sequences were generated after RNA extraction, reverse transcriptase polymerase chain reaction (RT-PCR) amplification, and automatic DNA sequencing, performed with the Applied Biosystems Viroseq HIV-1 Genotyping System (Abbott Wiesbaden Germany) following the manufacturer's instructions. Viral RNA extraction was performed using a Roche Amplicor HIV-1 Monitor with an Ultrasensitive Specimen Preparation kit (alcohol extraction). The PCR products (1.8-kb amplicons) were sequenced using seven different overlapping sequence-specific primers, capillary electrophoresis, and laser detection; the sequencing samples were analyzed with the ABI Prism—310 Genetic Analyzer using Viroseq analysis software (Applied Biosystems, Foster City, CA).

Genetic subtyping and phylogenetic analysis

All 83 HIV-1 *pol* sequences were first analyzed using the REGA HIV-1 Subtyping Tool.⁸ The sequences were aligned using ClustalX software⁹ followed by manual editing using the Bioedit software.¹⁰ The final data sets included HIV-1 *pol* Bulgarian strains, as well as subtype-specific and circulating

recombinant form (CRF) reference sequences downloaded from the HIV Los Alamos database (<http://hiv-web.lanl.gov/>). A BLAST search for B and non-B subtype Bulgarian sequences was performed, and separate alignments were developed with the most similar GenBank sequences for each subtype and CRF. For each aligned data set, the best fitting nucleotide substitution model was tested with a hierarchical likelihood ratio test, using a neighbor-joining (NJ) tree with LogDet corrected distances as base tree.¹¹ A maximum likelihood (ML) tree was then inferred with the selected model and ML-estimated substitution parameters. The heuristic search for the best tree was performed using an NJ tree as the starting tree and the TBR branch-swapping algorithm. NJ trees were also obtained with ML estimated pairwise distances using the best fitting nucleotide substitution model. Calculations were performed with PAUP* software (version 4.0).¹² Bayesian phylogenetic trees were also obtained with the program MrBayes version 3 [<http://mrbayes.csit.fsu.edu/>] running for each alignment two parallel Markov Chain Monte Carlo (MCMC) for 10,000,000 generations, with a sampling frequency of 1/1000. Bayesian consensus trees were obtained using the SUMT command in MrBayes. Convergence of the MCMC procedure was assessed by calculating the effective sampling size (ESS) with the program Tracer version 1.3 [<http://evolve.zoo.ox.ac.uk/beast/>]. An ESS >500 was considered significant.

Statistical support for specific clades in each phylogeny was obtained with the ML-based zero branch length test for the ML tree,¹² by bootstrapping (1000 replicates) for the NJ trees, and by calculating the posterior probability of each monophyletic clade for the Bayesian tree. Trees were rooted by ML rooting by selecting the rooted tree with the best likelihood under the molecular clock constraint or by outgroup rooting. The nucleotide sequences obtained in this study have been submitted to GenBank under accession numbers EF517409–EF517489. Multiple sequence alignments of Bulgarian and reference HIV strains are available from the authors upon request.

Molecular clock analysis

Since both subtype A and subtype B strains (the two main subtypes responsible for about 75% of the infections in our sample, see Results) were collected at different times within a 4-year period, it was possible to use their different sampling dates to calibrate a molecular clock. We assembled two separate data sets for HIV-1A and HIV-1B sequences, each one including Bulgarian strains, as well as a subset of those reference sequences for which sampling dates were available in the Los Alamos HIV database. Ultrametric trees were obtained by enforcing a molecular clock on the inferred genealogy and reestimating the branch lengths and substitution parameters with ML using the previously selected evolutionary model. The clock hypothesis was tested with the likelihood ratio test. Evolutionary rates and divergence times for the major Bulgarian clades in the HIV-1 subtype A and B phylogenies were also estimated with the Bayesian framework implemented in the BEAST program using the HKY+, and a relaxed molecular clock model that employed an exponential distribution of evolutionary rates as before.¹³ For each alignment, three parallel MCMC were run and combined with the program LogCombiner v1.4 [<http://evolve.zoo.ox.ac.uk/beast/>]. Each chain was run for 20,000,000 gen-

erations, with sampling every 1000 generations. Convergence of the MCMC procedure was assessed by calculating the ESS with the program Tracer version 1.3 [http://evolve.zoo.ox.ac.uk/beat/]. An ESS >500 was considered significant.

Phylogeographic analysis

The hypothesis of restricted gene flow among distinct HIV-1 populations in Bulgaria and other Eastern and Western European countries was tested with a modified version of the Slatkin and Maddison test¹⁴ that counts migration to/from different viral subpopulations,¹⁵ using the MacClade version 4 program (Sinauer Associates, Sunderland, MA). A one-character data matrix is obtained from the original data set by assigning to each taxon in the tree a one-letter code indicating its country (or geographic region) of origin. Then, the putative origin of each ancestral sequence (i.e., internal node) in the tree is inferred by finding the most parsimonious reconstruction (MPR) of the ancestral character. The final tree length, i.e., the number of observed migrations in the genealogy, can easily be computed and compared to the tree-length distribution of 10,000 trees obtained by random joining-splitting. Observed genealogies significantly shorter than random trees indicate the presence of subdivided populations with restricted gene flow. Specific migrations among different countries (character states) were traced with the *State changes and stasis* tool (MacClade software), which counts the number of changes in a tree for each pairwise character state. When multiple MPRs were present (as in our data sets), the algorithm calculated the average migration count over all possible MPRs for each pair. The resulting pairwise migration matrix was then normalized, and a randomization test with 10,000 matrices obtained from 10,000 random trees (by random joining-splitting of the original tree) was performed to assess the statistical significance of the observed migration counts.

Results

Diversity of HIV-1 subtypes circulating in Bulgaria

Of the 83 *pol* gene sequences, 41 (51%) were classified as subtype B, 18 (22%) as subtype A1, 6 (7%) as subtype C, 3 (3.7%) as subtype F, and 2 (2.5% each) as subtype H or G, respectively. The subtype analysis also classified eight sequences (10%) as putative CRFs (five as 01_AE, one as 02_AG, and two as 05_DF), whereas the remaining three (3.5%) sequences could not be classified into known groups (data not shown). The distribution of the different subtypes by age, gender, HIV exposure category, presumed place of infection, and year of HIV/AIDS diagnosis is reported in Table 1. Subtype A represented 27% of all subtypes identified after 1995, but was detected in only one individual diagnosed up to 1995, suggesting a later introduction of HIV-1A in Bulgaria. After 1995, the detection frequency of subtype A versus subtype B did not significantly change over time (χ^2 for trend >0.05).

High-resolution phylogenetic analysis of HIV-1 subtype B in Bulgaria

The ML *pol* gene tree of Bulgarian and other available HIV-1B reference sequences is shown in Fig. 1. NJ and Bayesian trees gave similar topologies (data not shown). The Bulgarian

subtype B strains were highly interspersed within the tree. Four monophyletic clades of Bulgarian sequences, highlighted by the boxes in Fig. 1, were statistically supported in the NJ (bootstrap values >75%), ML (zero branch length test p -value <0.001), and Bayesian tree (posterior probability >90%). All four clades clustered subtype B sequences from individuals infected through sexual transmission. Sequences from individuals infected through other transmission routes were intermixed in the tree and failed to cluster within any supported monophyletic clade. Bulgarian subtype B clades I and III appeared to be more closely related to HIV subtype B sequences from Western Europe, whereas II and IV clades from Bulgaria clustered with HIV-1 sequences from the Former Soviet Union (FSU). Clade IV in particular was closely related to strains from Ukraine. The tree also showed a clustering between two Bulgarian strains with two strains from the FSU (Russia and Bielorrussia), that was not statistically supported (bootstrap <50%). The high degree of intermixing between Bulgarian and foreign sequences and the phylogenetic position of the four statistically supported clades are evidence of several independent introductions of HIV-1 subtype B into Bulgaria from both East and West Europe. Moreover, the existence of independent monophyletic clades implies the presence of separate transmission networks for subtype B within the Balkan countries.

High-resolution phylogenetic analysis of HIV-1 subtype A1 in Bulgaria

Figure 2 shows the maximum ML *pol* gene tree of Bulgarian and other available HIV-1A1 reference sequences. NJ and Bayesian trees gave similar topologies (data not shown). The Bulgarian subtype A strains are distributed within three statistically supported monophyletic clades (bootstrap values >75%, zero branch length test p -value <0.001, Bayesian posterior probability >90%). Strains 21A and 172A, which clustered together at the top of the tree, were isolated from two subjects infected through blood transfusion, whereas the strains in the remaining two clades (indicated in the figure as Bulgaria West Europe clade and FSU clade, respectively) were isolated from persons infected through sexual contact. The Bulgarian blood transfusion subtype A1 clade appears to be closely related to a monophyletic clade that includes both West European and African sequences. A second statistically supported monophyletic clade included most of the Bulgarian subtype A1 strains (77.8%) and HIV-1 subtype A1 sequences from Western Europe, including France and Spain. Although a monophyletic group joining a strain from the former Yugoslavia and one from Greece appeared at the origin of the clade (see Fig. 2), the clustering was not significantly supported in the ML tree and was not confirmed by the NJ or Bayesian trees. The presence of HIV-1A sequences from Africa (Cameroon), close to the root of the clade in the ML, NJ, and Bayesian trees (data not shown), with Western European strains branching off suggests an early introduction of subtype A1 from Africa to West Europe. The third statistically supported clade included two Bulgarian subtype A strains, in addition to sequences originating from different Eastern European countries. Overall, each cluster is likely to represent a separate introduction into Bulgaria of HIV-1A1 from at least two different geographic areas (West and East Europe) by different transmission routes (sexual transmission and blood transfusion).

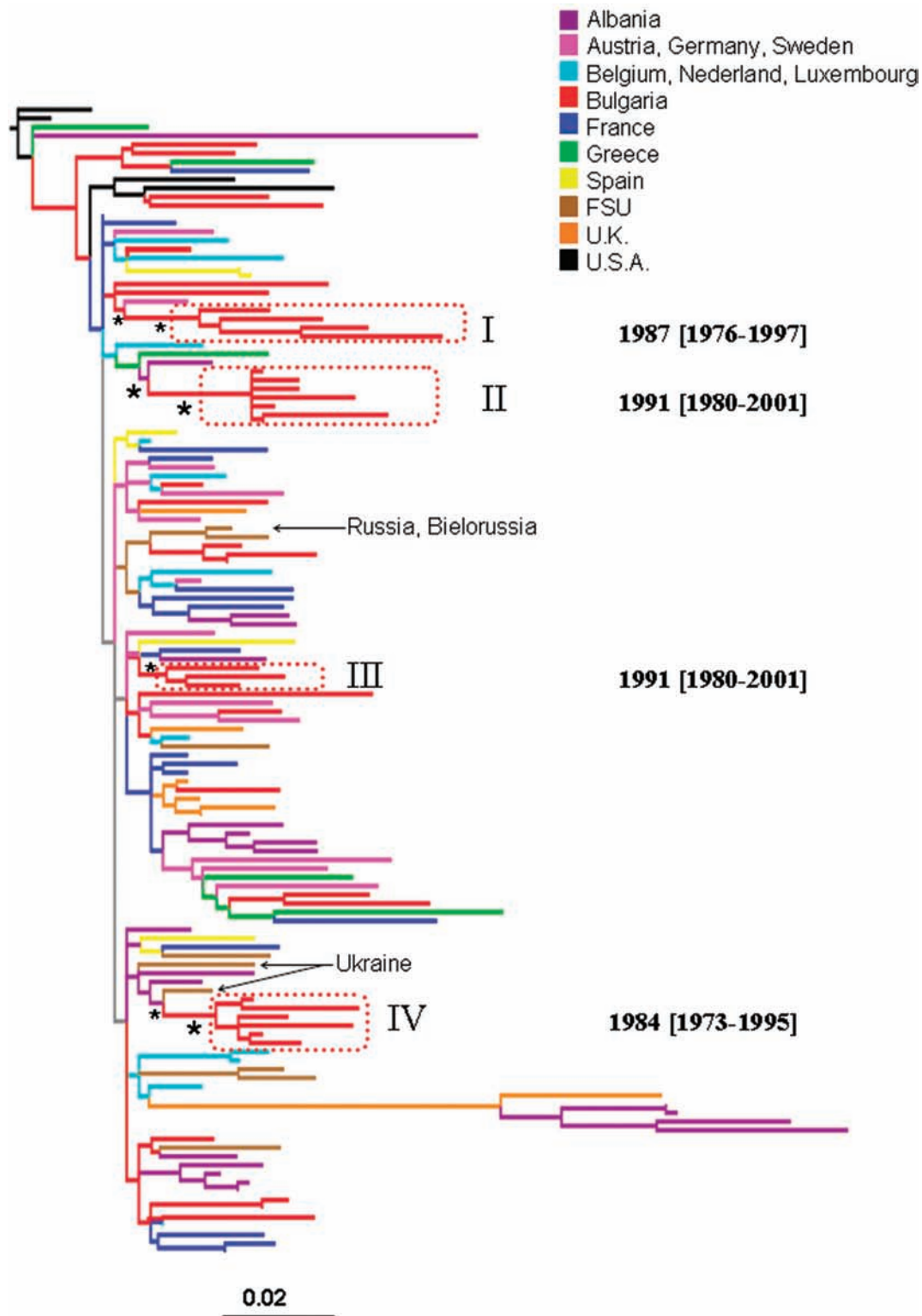


FIG. 1. Maximum likelihood phylogenetic analysis of HIV-1B *pol* sequences. The data set included 41 HIV-1B strains from Bulgaria and 91 subtype B reference sequences downloaded from the Los Alamos HIV database. The tree was rooted by using two HIV-1A strains as the outgroup. Branch lengths were estimated with the best fitting nucleotide substitution model according to a hierarchical likelihood ratio test,¹¹ and were drawn in scale with the bar at the bottom indicating 0.02 nucleotide substitutions per site. One asterisk (*) along a branch represents significant statistical support for the clade subtending that branch ($p < 0.001$ in the zero-branch-length test, Bayesian posterior probability $>90\%$, and bootstrap support $>75\%$). The color of each external branch represents the country of origin of the sequence corresponding to that branch, according to the legend in the figure. FSU (Former Soviet Union) strains include sequences from Russia, Bielorrussia, Ukraine, and Slovenia (the specific country of origin is indicated by the arrows). Broken boxes highlight statistically supported monophyletic clades including Bulgarian strains. The inferred time of the most recent common ancestor of each supported Bulgarian clade is also indicated on the right of the highlighted clade. Dates and 95% high posterior density interval (within square brackets) were estimated using a Bayesian relaxed molecular clock (see Materials and Methods).

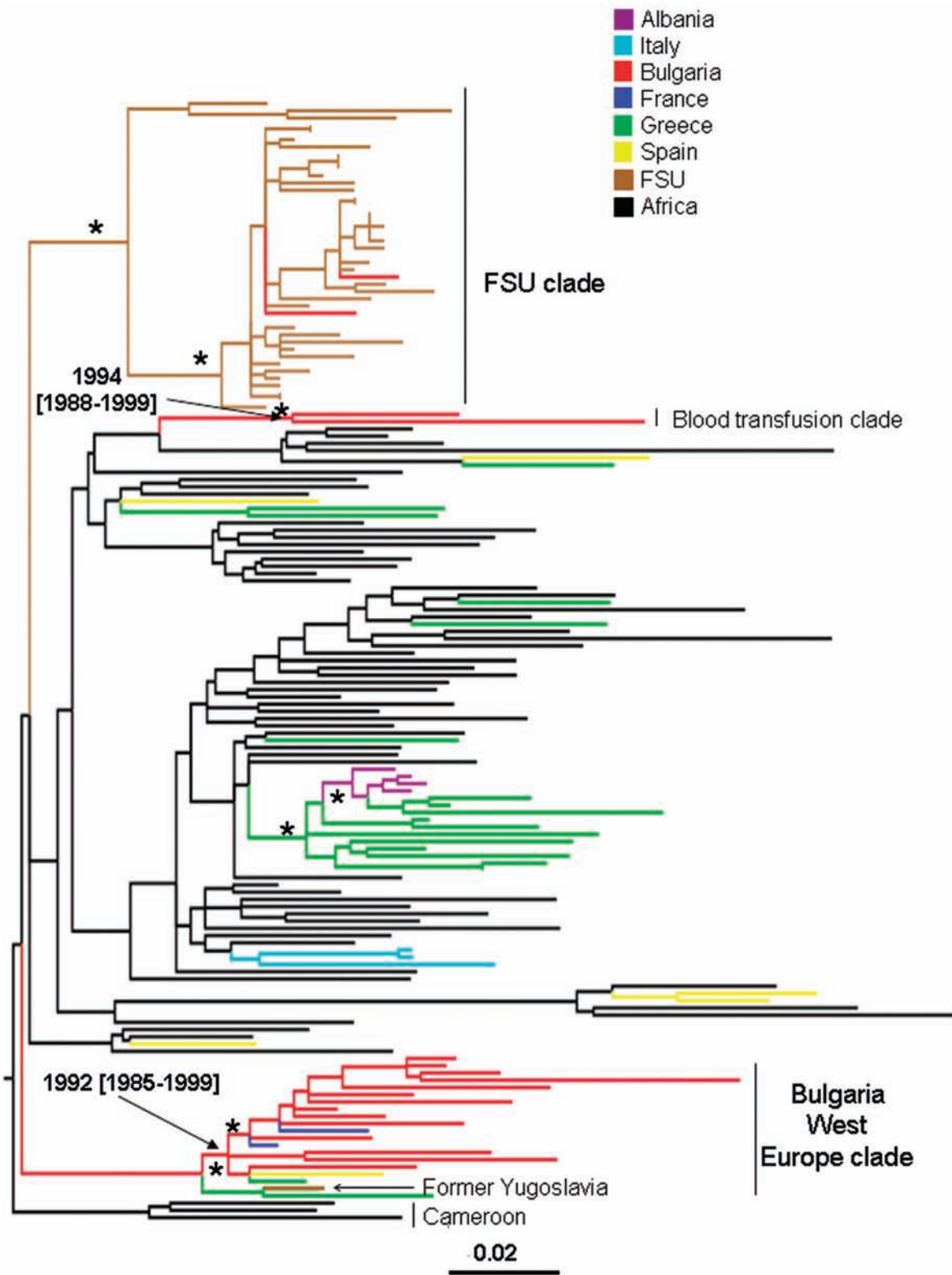


FIG. 2. Maximum likelihood phylogenetic analysis of HIV-1A1 *pol* sequences. The data set included 18 HIV-1A strains from Bulgaria and 137 subtype A1 reference sequences downloaded from the Los Alamos HIV database. The tree was rooted by using two HIV-1B strains as the outgroup. Branch lengths were estimated with the best fitting nucleotide substitution model according to a hierarchical likelihood ratio test,¹¹ and were drawn to scale with the bar at the bottom indicating 0.01 nucleotide substitutions per site. One asterisk (*) along a branch represents significant statistical support for the clade subtending that branch ($p < 0.001$ in the zero-branch-length test, Bayesian posterior probability $>90\%$, and bootstrap support $>75\%$). The color of each external branch represents the country of origin of the sequence corresponding to that branch, according to the legend in the figure. The Blood Transfusion clade and Bulgaria West Europe clade are the two major monophyletic clades containing the Bulgarian isolates. The FSU (Former Soviet Union) clade includes strains from Russia, Bielorusia, Ukraine, Slovenia, Czech Republic, and two strains from Bulgaria. Dates and 95% high posterior density intervals (within square brackets) were estimated using a Bayesian relaxed molecular clock (see Materials and Methods).

Molecular clock analysis of the major HIV-1 Bulgarian clades

To estimate the time of the most recent common ancestor (tMRCA) of the supported Bulgarian clades within subtype A and B, we implemented a strict molecular clock model, assuming a constant evolutionary rate along each subtype tree. Since the molecular clock hypothesis was rejected for both subtype A and subtype B isolates ($p < 0.0001$), we calibrated the rate and the origin of the different HIV-1 Bulgarian clades by relaxing the strict clock model and allowing the evolutionary rate to vary along the tree by sampling branch-specific rates from an exponential distribution (see Materials and Methods). The tMRCA for the different clades within subtype B and A are shown in the trees in Figs. 1 and 2, respectively. It was not possible to date the introduction of HIV-1A in Bulgaria from Eastern Europe (FSU clade in Fig. 2) because only two Bulgarian strains intermixed in the FSU clade were available. The other two subtype A Bulgarian clades (Fig. 2), as well as clade II and III within subtype B (Fig. 1), originated in the early 1990s, while the origin of subtype B clade I and III was estimated to be in the mid-1980s (Fig. 1).

Migration patterns of HIV-1 A1 and B subtypes to/from Bulgaria

The trees given in Figs. 1 and 2 were also analyzed with a modified version of the Slatkin and Maddison method^{14,15} to infer the HIV-1 gene flow (migration) between Bulgaria and other geographic areas. To perform the analysis, the HIV-1 non-Bulgarian sequences were assigned to three distinct groups: Africa, West Europe, and East Europe for subtype A1, or the United States, West Europe, and East Europe for subtype B. The null hypothesis of panmixia (i.e., no population subdivision or complete intermixing of sequences from different geographic areas) was rejected by the randomization test ($p < 0.0001$) for both subtypes.¹⁴ Almost 19% of the observed HIV-1 subtype A1 gene flow to Bulgaria was from Africa and East and West Europe (Fig. 3A).^{16,17} In contrast, the HIV-1A1 gene flow from Bulgaria to West and East European countries accounted for only 5.4% of the observed migrations, which was the same order of magnitude as the migrations observed between East European and West European groups (Fig. 3A). A substantial gene outflow of HIV-1A1 strains from Africa to the other regions, including Bulgaria, was also evident (71.9% of the observed migrations, Fig. 3A). Within the HIV-1 subtype B tree, the gene flow to Bulgaria from the United States, as well as West and East European countries, accounted for 45.2% of the observed migrations (Fig. 3B). The gene flow from Bulgaria to West Europe was 5.4%, slightly lower than the gene flow from other East European countries to West Europe (8%). On the other hand, the HIV-1B gene flow from West European countries to Bulgaria and to other Eastern European countries accounted for 34.2% and 33.8% of the observed migrations, respectively. The proportion of gene flow to/from the United States displayed in Fig. 3B was probably underestimated because HIV-1B migration analysis was focused primarily on the relationship between the epidemic in Bulgaria and other European countries and included only a few U.S. isolates for comparison.¹⁸⁻²⁶

Discussion

The present study is the first detailed investigation of the HIV-1 molecular epidemiology in Bulgaria. Although AIDS and HIV surveillance data suggest that Bulgaria, like other countries in Central Europe, still has a low AIDS/HIV incidence, data on other sexually transmitted infections indicate a potential for spread of HIV infection. For example, 21.5% of commercial sex workers, 6.7% of the Roma community,

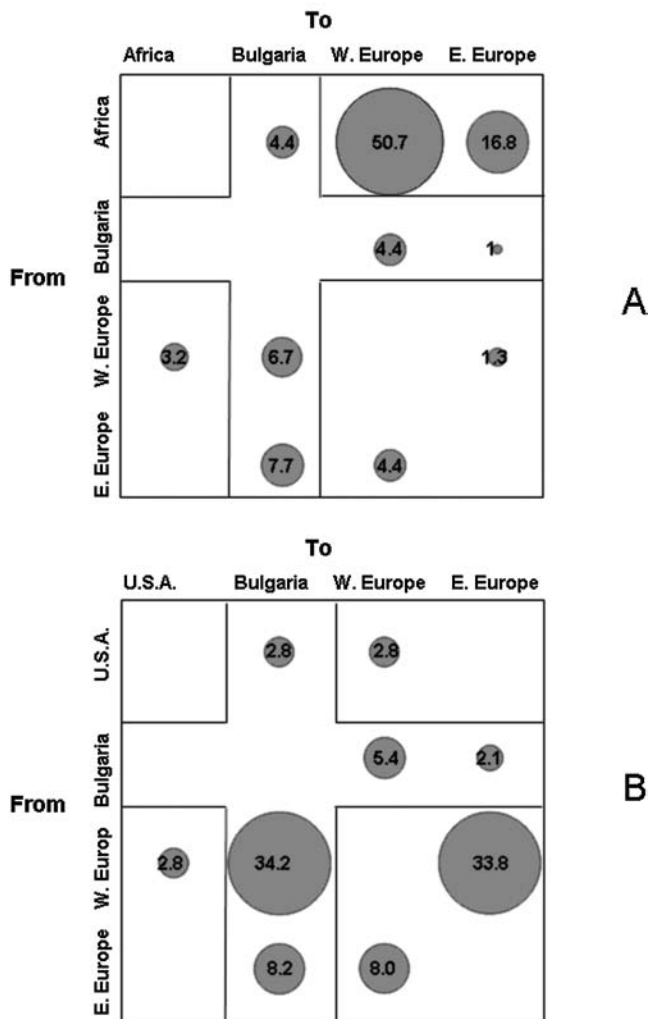


FIG. 3. Maximum parsimony migration patterns of Bulgarian sequences to/from different geographic areas. The bubblegram shows the frequency of gene flow (migrations) to/from different geographic areas, as the percentage of the total observed migrations estimated from the maximum likelihood trees for different subtypes with a modified version of the Slatkin and Maddison test (see Materials and Methods). The surface of each circle is proportional to the percentage of observed migrations given within the circle. The highlighted row indicates inferred HIV-1 outflow from Bulgaria to different geographic areas. The highlighted column indicates inferred HIV-1 inflow to Bulgaria from different geographic areas. (A) HIV-1A1 subtype migration from/to Africa, Bulgaria, and West Europe estimated from the tree given in Fig. 1. (B) HIV-1B subtype migration from/to the United States, Bulgaria, East Europe, and West Europe estimated from the tree given in Fig. 2.

TABLE 1. DISTRIBUTION OF HIV-1 SUBTYPES BY AGE, GENDER, ROUTE OF INFECTION, PLACE OF PRESUMED INFECTION, AND YEAR OF DIAGNOSIS IN BULGARIA

	Subtype A		Subtype B		Subtype other		Subtype unclassified		Total
	(N = 18)	(%)	(N = 41)	(%)	(N = 21)	(%)	(N = 3)	(%)	
Age (in years)									
≤38	10	25	21	52.5	8	20	1	2.5	40
>38	8	18.6	20	46.5	13	30	2	4.9	43
Gender									
Male	8	14.5	34	61.8	12	21.8	1	1.9	55
Female	10	35.7	7	25	9	32.1	2	7.2	28
Route of infection									
Sexual	15	22.1	31	45.6	19	27.9	3	4.4	68
Vertical	1	50.0	None		1	50.0	None		2
MSM	None		5	100	None		None		5
Blood	2	30.0	3	50	1	20	None		6
IDUs	None		2	100	None		None		2
Presumed place of infection									
Bulgaria	15	22.1	35	51.5	16	23.5	2	2.9	68
Other	3	20.0	6	40.0	5	33.3	1	6.7	15
Distribution by year of presumed infection									
1986–1995	1	5.3	8	42.1	8	42.1	2	10.5	19
1996–2000	7	28	11	44	7	28	None		25
>2001	10	25.6	22	56.4	6	15.4	1	2.6	39

and 2.4% of intravenous drug users were found positive for syphilis.^{2,3} Our findings suggest that HIV-1 infection has been present in Bulgaria for at least the past two decades and established three main characteristics of the epidemic: first, a high genetic heterogeneity in terms of circulating subtypes; second, several independent transmission clusters within subtype A1 and B (the two most represented subtypes in the screened sample), indicating multiple introductions over time but only limited introduction from countries of the FSU; and third, a dynamic HIV-1 gene inflow and outflow from/to West Europe and other East European countries.

Results of the screened population sample, which included about 15% of all known diagnosed HIV subjects in this Balkan country, clearly showed a wide genetic diversity of the circulating virus. Although subtype B was the most prevalent subtype and was identified in about 50% of the screened samples, subtype A1 was identified in approximately 25% of samples, while non-A/non-B subtypes or several putative CRFs comprised another 25%. The molecular clock analysis estimated that HIV-1B was present in Bulgaria since the early 1980s, with a lower limit estimate for the first introduction in 1973 that is within the same time frame as the introduction of HIV-1 subtype B in North America.²⁷ In contrast to subtype B, the introduction of HIV-1 subtype A1 appeared to be more recent (early 1990s). Both molecular clock and viral gene flow analyses indicated a dynamic epidemic over the past 25 years, with multiple introductions from foreign countries and ongoing gene outflow from Bulgaria to West and East Europe. Although fewer than 600 cumulative AIDS cases have been identified in Bulgaria to date, our data suggest that the actual prevalence of HIV-1 in Bulgaria may be underestimated. The dynamic patterns of the HIV-1 epidemic within the country point out an enormous

potential for spreading within the next few years, both within the country and to other neighboring countries in Europe. In this regard, three factors seem to be highly relevant: (1) the central geographic location of Bulgaria at the crossing between Western Europe, Eastern Europe, and the Middle East, (2) the high genetic heterogeneity of the HIV-1 epidemic in the Balkan countries revealed by the molecular characterization of our samples, and (3) the ongoing sociopolitical changes that have been affecting Bulgaria, as well as other countries in Eastern and Central Europe, since the end of the cold war.

The relatively few studies on the HIV-1 epidemic in other Eastern European countries have shown that non-B subtypes are quite common, with HIV-1A being highly prevalent in Russia, Belarus, and Ukraine,^{28–30} as well as the former Yugoslavia and Albania.^{6,7} It has been demonstrated that shortly after the beginning of the subtype A epidemic in the FSU the virus was introduced into Ukraine and then spread, as shown in our analysis by the clustering of Bulgarian strains within the Eastern European clade, to other countries in Eastern Europe,³¹ including Bulgaria. Surprisingly, only 1% of the subtype A1 Bulgarian strains actually clustered within the Eastern European clade. In fact, phylogenetic analysis indicated that at least three independent introductions of the HIV-1A1 subtype have occurred in Bulgaria, and that the majority of the HIV-1A Bulgarian strains in our sample were more closely related to West European subtype A sequences. Phylogeographic (migration) analysis showed that the largest HIV-1A1 outflow occurred from the African continent, a finding consistent with the African origin and the high seroprevalence of subtype A1 in Africa.¹ In addition, a significant HIV-1A1 gene flow from West Europe as well as other East European countries to Bulgaria was ob-

served. The analysis revealed a complex web of introductions/transmissions of the virus between Bulgaria and other European countries, with at least four separate introductions of HIV-1B supported by both phylogeny and molecular clock techniques. The findings are consistent with the central geographic location of Bulgaria and point out that continuous surveillance of the epidemic in this Balkan country may be of great importance for the monitoring and prediction of future HIV-1 epidemiological trends in the rest of Europe.

A recent study explored the history of the HIV epidemic in the United Kingdom within defined risk groups.³² The study revealed the complex epidemiological structure of different transmission clusters characterized by an initial period of exponential growth, followed by a slower spread as a consequence of behavior change within the risk group and the introduction of HAART. Most of the Bulgarian clades included viral strains from individuals infected exclusively through sexual contact. However, the presence of distinct and well-supported monophyletic clades within both HIV-1A1 and HIV-1B phylogeny suggests that after each initial introduction of the virus in a specific population, separate transmission clusters have been evolving along independent phylogenetic lineages. Better characterization and continuous monitoring of such groups are going to be crucial to understand in detail the epidemiology of HIV-1 in Bulgaria and to assess the efficacy of prevention and therapy in controlling the epidemic.

Some possible limits and biases of the present study should be mentioned. First, we cannot exclude the possibility that the study population is not truly representative of all HIV cases in Bulgaria. In particular, most persons included in our analysis had treatment failure; thus it is possible that persons in an advanced stage of disease (i.e., AIDS cases) were overrepresented. The possible consequences could be (1) underestimation of the increase in the circulation of different clades in recent infections or (2) overrepresentation of transmission clusters due to the elevated virulence of the circulating virus. Underestimation might occur only if changes in the circulation of different clades took place over recent years, which seems unlikely in the light of phylogeny and molecular clock analysis that showed a consistent evolution of independent transmission clusters after the initial HIV-1 introduction between 1980 and 1990. Overestimation is also unlikely to have affected our findings since there is no definitive evidence that some clades within a specific subtype (i.e., A vs. B) are more virulent than others, and no clear-cut association between a specific subtype and virulence has been reported, although increased virulence of subtype D over A has been noted.³³ Another possible limit of our study is use of subgenomic, rather than full genome, sequences, which could underestimate recombinant strains or overestimate subtype representation in the population. Even so, given the high statistical association of Bulgarian strains with specific Eastern and Western European sequences within both subtype A1 and B data sets, neither the conclusion of multiple HIV-1 introductions in Bulgaria nor the migration analysis was likely to be affected. In conclusion, we found a wide genetic diversity of HIV-1 in Bulgaria. Both subtype A1 and B appeared to have been introduced multiple times within the past two decades. Our data also suggest the extreme dynamic nature of the Bulgarian epidemic characterized by considerable viral inflow and out-

flow to/from other European countries, and emphasize the need for improved surveillance and control of HIV-1 infection in the Balkans.

Acknowledgments

The research was supported by PHS R01 awards HD032259, AI047723, and AI065265; General Clinical Research Center of the University of Florida MO1-RR00082; University of Florida Children's Miracle Network; Center for Research in Pediatric Immune Deficiency; Laura McClamma Research Fellowship; Pediatric Clinical Research Center of All Children's Hospital, University of South Florida, and the Maternal and Child Health Bureau, R60 MC 00003-01, Department of Health and Human Services, Health Resources and Services Administration; and Robert A. Good Chair for Immunology (J.W.S.) and Stephany W. Holloway University Chair for AIDS Research (M.M.G.).

References

1. EuroHIV (European Centre for the Epidemiological Monitoring of AIDS): HIV/AIDS Surveillance in Europe: Mid-year report, 2005.
2. Matic S, Lazarus JV, and Donoghoe MC, eds.: *HIV/AIDS in Europe: Moving from Death Sentence to Chronic Disease Management*. WHO Regional Office for Europe, Copenhagen, Denmark, 2006.
3. Kabakchieva E, Vassileva S, Kelly JA, et al.: HIV risk behaviour patterns, predictors, and sexually transmitted disease prevalence in the social networks of young Roma (Gypsy) men in Sofia, Bulgaria. *Sex Transm Dis* 2006;33(8):485-490.
4. Kabakchieva E, Amirkhanian YA, Kelly JA, McAuliffe TL, and Vassileva S: High levels of sexual HIV/STD risk behaviour among Roma (Gypsy) men in Bulgaria: Pattern and predictors of risk in a representative community sample. *Int J STD AIDS* 2002;13(3):184-191.
5. Kelly JA, Amirkhanian YA, Kabakchieva E, et al.: Gender roles and HIV sexual risk vulnerability of Roma (Gypsies) men and women in Bulgaria and Hungary: An ethnographic study. *AIDS Care* 2004;16(2):231-245.
6. Stanojevic M, Papa A, Papadimitriou E, et al.: HIV-1 subtypes in Yugoslavia. *AIDS Res Hum Retroviruses* 2002;18:519-522.
7. Ciccozzi M, Gori C, Boros S, et al.: Molecular diversity of HIV in Albania. *J Infect Dis* 2005;192:475-479.
8. De Oliveira T, Deforche K, Cassol S, et al.: An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* 2005;21(19):3797-3800.
9. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG: CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1994;25:4876-4882.
10. Hall TA: Bioedit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98 NT. *Nucleic Acids Symp Ser* 1999;41:95-98.
11. Swofford D and Sullivan J: Phylogeny inference based on parsimony and other methods with PAUP*. In: *The Phylogenetic Handbook—A Practical Approach to DNA and Protein Phylogeny* (Salemi M and Vandamme AM, eds.). Cambridge University Press, New York, 2003, pp. 160-206.
12. Swofford DL: PAUP: Phylogenetic analysis using parsimony (*and other methods). Version 4.0, Sinauer Associates, Sunderland, MA, 2002.

13. Drummond AJ, Ho SYW, Phillips MJ, and Rambaut A: Relaxed phylogenetics and dating with confidence. *PLoS Biol* 2006;4:1–12.
14. Slatkin M and Maddison WP: A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 1989; 123:401–416.
15. Salemi M, Lamers SL, Yu S, de Oliveira T, Fitch WM, *et al.*: HIV-1 phylodynamic analysis in distinct brain compartments provides a model for the neuropathogenesis of AIDS. *J Virol* 2005;79:11343–11352.
16. McCutchan FE: Global epidemiology of HIV. *J Med Virol* 2006;78:S7–S12.
17. Strimmer K and Pybus OG: Exploring the demographic history of DNA sequences using the generalized skyline plot. *Mol Biol Evol* 2001;18:2298–2305.
18. Smolskaya T, Liitsola K, Zetterberg V, Golovanova E, *et al.*: HIV epidemiology in the Northwestern Federal District of Russia: Dominance of HIV type 1 subtype A. *AIDS Res Hum Retroviruses* 2006;22(11):1074–1080.
19. Saad MD, Aliev Q, Botros BA, *et al.*: Genetic forms of HIV type 1 in the former Soviet Union dominate the epidemic in Azerbaijan 2006. *AIDS Res Hum Retroviruses* 22(8):796–800.
20. Zarandia M, Tsertsvadze T, Carr JK, Nadai Y, Sanchez JL, and Nelson AK: HIV-1 genetic diversity and genotyping drug susceptibility in the Republic of Georgia. *AIDS Res Hum Retroviruses* 2006;22(5):470–476.
21. 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–W23.
22. Kato S, Saito R, and Hiraishi Y: Differential prevalence of HIV type 1 subtype B and CRF01_AE among different sexual transmission groups in Tokyo, Japan, as revealed by subtype-specific PCR. *AIDS Res Hum Retroviruses* 2003;19: 1057–1063.
23. Lan NT, Recordon-Pinson P, Hung PV, *et al.*: HIV type 1 isolates from 200 untreated individuals in Ho Chi Minh City (Vietnam): ANRS 1257 Study. Large predominance of CRF01_AE and presence of major resistance mutation to antiretroviral drugs. *AIDS Res Hum Retroviruses* 2003;19:925–928.
24. Wichukchinda N, Shiino T, Srisawat J, *et al.*: Heterosexual transmission of novel CRF01 AE and subtype B recombinant forms of HIV type 1 in Northern Thailand. *AIDS Res Hum Retroviruses* 2005;21:734–738.
25. Papa A, Papadimitriou E, Papoutsis A, Kiosses V, and Antonidas A: HIV-1 subtypes and circulating recombinant forms (CRFs) in Northern Greece. *Virus Res* 2002;85:85–93.
26. Paraskevis D, Magiorkinis E, Theodoridou M, *et al.*: Molecular epidemiology of vertical human immunodeficiency virus type 1 transmission in Greece: Evidence of non-B subtype. *J Hum Virol* 1999;2(6):339–343.
27. Robbins KE, Lemey P, Pybus OG, Jaffe HW, Saekhou AM, Brown TM, Salemi M, Vandamme A-M, and Kalish ML: Characterization of early U.S. HIV-1 strains, dating of the introduction of HIV-1 into the U.S. and estimation of population history of the U.S. epidemic. *J Virol* 2003;77:6359–6366.
28. Lukashov VV, Huisman R, Rakhmanova AG, Lisistina ZN, Akhtyrskaya NA, Vlasov NN, Melnick OB, and Goudsmit J: Circulation of subtype A and gagA/envB recombinant HIV type 1 strains among injecting drug users in St. Petersburg, Russia, correlates with geographical origin of infections. *AIDS Res Hum Retroviruses* 1999;15:1577–1583.
29. Lazouskaya NV, Eremin VF, Adema KW, Gasich EL, Baan E, and Lukashov VV: The HIV type 1 epidemic in Belarus: Predominance of Eastern European subtype A strains and circulation of subtype B strains. *AIDS Res Hum Retroviruses* 2005;21(9):830–833.
30. Saad MD, Shcherbinskaya AM, Nadai Y, Kruglov YV, Antonenko SV, *et al.*: Molecular epidemiology of HIV type 1 in Ukraine: Birthplace of an epidemic. *AIDS Res Hum Retroviruses* 2006;22(8):709–714.
31. Nabatov AA, Kravchenko ON, Lyulchuk MG, Shcherbinskaya AM, and Lukashov VV: Simultaneous introduction of HIV type 1 subtype A and B viruses into injecting drug users in Southern Ukraine at the beginning of the epidemic in former Soviet Union. *AIDS Res Hum Retroviruses* 2002;28: 891–895.
32. Hue S, Pillay D, Clewly JP, and Pybus OG: Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc Natl Acad Sci USA* 2005;102 (12): 4425–4429.
33. Kaleebu P, French N, Mahe C, Yrrell D, Watera C, *et al.*: Effect of human immunodeficiency virus type 1 (HIV-1) envelope subtypes A and D on disease progression in a large cohort of HIV-1 positive persons in Uganda. *J Infect Dis* 2002;185(9):1244–1250.

Address reprint requests to:
Massimo Ciccozzi
 Department of Infectious Diseases
 Istituto Superiore Di Sanita
 Viale Regina Elena 299
 00161 Rome, Italy

E-mail: Ciccozzi@iss.it