# *Sequence Note*

# Molecular Epidemiology of HIV Type 1 in Newly Diagnosed Patients in Southern Spain

Michael J. McConnell,<sup>1</sup> Fernando Docobo-Pérez,<sup>1</sup> Rosario C. Mata,<sup>1</sup> Felipe Fernandez-Cuenca,<sup>2</sup> Pompeyo Viciana,<sup>1</sup> Luis F. López-Cortés,<sup>1</sup> Mónica Trastoy,<sup>1</sup> Jerónimo Pachón,<sup>1</sup> and Pilar Pérez-Romero<sup>1</sup>

## **Abstract**

The prevalence of different HIV-1 subtypes in Spain varies by geographic region. In the present study isolates were collected from 72 newly diagnosed individuals in western Andalucia from 2004 to 2006. Viral sequences were amplified and the subtype diversity and prevalence of resistance mutations in the reverse transcriptase and protease genes were determined. The results presented here demonstrate that subtype B virus predominates in this region (88.9%), with the non-B subtypes CRF02\_AG (9.7%) and B/G (1.4%) also present. Only two isolates (2.9%) carried resistance mutations in the reverse transcriptase gene and none of the isolates had major resistance mutations in the protease gene. Minor mutations in the protease gene were more prevalent with 86.1% of isolates containing at least one minor mutation. These results elucidate the subtype diversity present in this region and suggest that the transmission of highly resistant virus variants does not occur at a high frequency in this population.

**V**ARIANTS OF HIV-1 HAVE BEEN CLASSIFIED INTO M, N, AND O GROUPS based on sequence homology. Group M viruses have been further subdivided into nine subtypes, A, B, C, D, F, G, H, J, and K, in addition to many circulating recombinant forms (CRFs) and unique recombinant forms (URFs) that have arisen via intersubtype recombination. The prevalence of HIV-1 subtypes varies greatly depending on the geographic region, with subtype B predominating in western Europe and North America, whereas multiple non-B subtypes account for the majority of infections in sub-Saharan Africa and India. Characterization of HIV variants endemic to distinct geographic regions is necessary since variation between the different subtypes of HIV-1 may affect antiretroviral treatment and future vaccination.

The use of antiretroviral therapy (ART) in developed countries has resulted in populations harboring resistant viruses. The transmission of resistant viruses has been previously reported, and in some populations the rates of transmitted resistance have increased as the use of ART becomes more widespread.<sup>1,2</sup> This is of particular concern since it has been demonstrated that more time is required to achieve virologic suppression with ART in patients infected with resistant viruses than in patients infected with viruses free of resistance mutations.<sup>1,2</sup> In addition, the duration of virologic suppression before virologic failure is shorter in patients infected with resistant viruses.<sup>2</sup> Resistance mutations present in the protease gene (PR) have been classified as either major mutations or minor mutations (also called polymorphisms). Major mutations are those that occur within the active site of PR and directly confer phenotypic resistance to protease inhibitors, and it has been well established that the presence of major mutations in PR is predictive of virologic failure in patients receiving ART.<sup>3</sup> However, the role played by minor mutations in PR in conferring resistance to PIs is less clear. It has been suggested that minor mutations in PR are compensatory mutations that increase the fitness of viruses harboring major mutations, $4.5$  and some studies have indicated that these mutations can affect treatment outcome.3,6

HIV surveillance studies have shown that the geographic regions of Spain are subject to unique migration patterns, resulting in populations that harbor different HIV subtype diversities.<sup>7–9</sup> The goal of the present study is to characterize the subtype diversity and determine the prevalence of re-

<sup>1</sup>Servicio de Enfermedades Infecciosas, Hospitales Universitarios Virgen del Rocío, Seville, Spain.

<sup>2</sup>Servicio de Microbiologia, Hospitales Universitarios Virgen Macarena, Seville, Spain.





Accession number	Patient ID	Subtype	Type of infection	Risk factor	Age	Sex	Country	Year of diagnosis
EU288149	NV63	B	Chr	<b>HTSX</b>	63	F	Spain	2006
EU288150	NV64	B	Chr	ND	38	Μ	Spain	2006
EU288151	NV65	B	Chr	<b>HTSX</b>	51	Μ	Spain	2006
EU288152	NV66	B	Chr	<b>IVDU</b>	38	М	Spain	2006
EU288153	NV67	B	Chr	<b>HTSX</b>	58	Μ	Spain	2006
EU288154	NV68	B	Chr	<b>HTSX</b>	50	F	Spain	2006
EU288155	NV69	B	Chr	HO/BI	24	Μ	Spain	2006
EU288156	NV70	B	Recent	HO/BI	24	Μ	Spain	2006
EU288157	<b>NV71</b>	B	Chr	<b>HTSX</b>	37	F	Spain	2006
EU288158	<b>NV72</b>	CRF02_AG	Chr	HO/BI	40	Μ	Spain	2006

TABLE 1. CHARACTERISTICS OF 72 HIV-INFECTED, TREATMENT-NAIVE PATIENTS RECEIVING CARE AT VIRGEN DEL ROCÍO UNIVERSITY HOSPITAL<sup>a</sup> (CONT'D)

aRecent, recently infected patients; Chr, chronically infected patients; HO/BI, homosexual/bisexual; HTSX, heterosexual; IVDU, intravenous drug use; ND, not determined; F, female; M, male.

sistance mutations in the reverse transcriptase (RT) and PR regions of *pol* in treatment-naive individuals at a tertiary medical center in western Andalucia from 2004 to 2006. For this purpose, 72 HIV-infected, treatment-naive patients receiving care at Virgen del Rocío University Hospital between 2004 and 2006 were included in the study. Table 1 shows the characteristics of individuals included in the study. Of the 72 patients, 67 were Spanish, two were from Nigeria (NV25 and NV28), one was from Argentina (NV12), one was from India (NV35), and one was from Morocco (NV42). Patient age, sex, risk factor, coinfection status (hepatitis B and C), viral load, and CD4<sup>+</sup> T cell counts were determined at the time of diagnosis. Plasma HIV RNA was quantified using the Amplicor HIV-1 Monitor Test 1.5 (Roche Diagnostics, Alamida, CA) with a limit of detection of 50 copies/ml. Patients were categorized as either recently or chronically infected using criteria similar to what has been used previously.12 Recently infected patients were defined as patients with documented seroconversion in the preceding 12 months or clinical evidence of acute HIV infection within 12 months of the first documented seroconversion. Chronically infected patients were defined as patients for whom there was no documented seroconversion or clinical evidence of acute infection in the preceding 12 months. For quantitative measurements, data sets were compared parametrically using the Student's *t* test or nonparametrically using the Mann–Whitney *U* test. Categorical data were analyzed using the  $\chi^2$  test. For all statistical tests a  $p$  value  $\leq 0.05$  was considered significant. All statistical analyses were performed using SPSS version 14.0 (SPSS Inc., Chicago, IL).

For genotypic analysis, viral RNA was isolated from 1 ml of plasma using the MagNA Pure LC instrument (Roche Applied Science, Indianapolis, IN) and sequenced using the TruGene Genotyping System (Bayer Healthcare), which sequences codons 10–99 of PR and 38–247 of RT. Virus subtypes were determined by analyzing sequence data using the Stanford University Drug Resistance Database (http:// hivdb.stanford.edu). Polymorphisms were determined by comparison with the HIV-1 reference strain LAV- $1_{BRU}$  (accession number K02013). Resistance conferring mutations in RT and PR were defined according to the recommendations of the International AIDS Society-USA.10

Phylogenetic analysis of the *pol* region sequences was performed using the Vector NTI software package (version 13.3.0, Invitrogen Corporation). Multiple nucleotide sequence alignment was performed using the AlignX program, an algorithm based on a modified ClustalW algorithm in which pairwise alignments are performed between selected sequences and HIV-1 group M reference sequences. A phylogenetic tree was constructed using CLC Free workbench (version 4.0.2, CLC bio A/S) through the neighbor-joining method in which a matrix of distances between all pairs of sequence was calculated after the sequences were aligned. Branch reproducibility was assessed using bootstrap analysis (500 replicates) and the results are presented as percentages. The tree was rooted with the simian immunodeficiency virus strain SIVcpz (accession number AF115393). Reference sequences for subtype B (accession numbers M19921 and NC001802), subtype B/G (accession number AY900576), subtype F (accession number DQ358812), subtype A (accession number EF545108), subtype C (accession number AY968312), subtype E (accession number AB052995), and CRF02\_AG (accession numbers AF447832, DQ845388) viruses were included in the phylogenetic analysis.

Baseline characteristics at the time of diagnosis for the 72 individuals included in the study are presented in Table 2. Seventeen individuals were identified as recently infected, whereas 55 were characterized as chronically infected due to evidence of infection for greater than 1 year or no evidence of recent infection. Significant differences between recently infected and chronically infected individuals included the following: a higher proportion of the recently infected individuals identified homosexual contact as their primary risk factor, recently infected individuals were significantly younger than chronically infected patients, and CD4<sup>+</sup> T cell levels were lower in chronically infected individuals than in those who were recently infected.

A phylogenetic tree constructed using sequences from the *pol* region of isolates included in the study is shown in Fig. 1. Our analysis resulted in one major cluster of 21 sequences (subtype B, NV02–NV31) and two minor clusters, each with three sequences (subtype B, NV01–NV27, and CRF02\_AG, NV18–NV37). The 21 sequences comprising the major cluster were isolated from Spanish males between 23 and 46

	Recently	Chronically	
	infected	infected	Total
Participants, No. (%) Sex, No. $(\% )$	17(23.6)	55 (76.4)	72 (100.0)
Male	16(94.1)	43 (78.2)	59 (81.9)
Female	1(5.9)	12(21.8)	13(18.1)
Age, mean (SD)	32.0(5.8)	40.0 $(12.2)^{b}$	38.1(11.5)
Risk factor, No. (%)			
Homosexual contact	$15(88.2)^c$	28 (50.9)	43 (59.7)
Heterosexual contact	$1(5.9)^c$	19 (34.5)	20(27.8)
Intravenous drug user	1(5.9)	3(5.5)	4(5.6)
Other/unknown	0(0.0)	5(9.1)	5(6.9)
Coinfection, No. (%)			
<b>HBV</b>	1(5.6)	14 (20.9)	15(17.6)
<b>HCV</b>	1(5.6)	9(13.4)	10(11.8)
$CD4^+/\mu l$ , median (IQR)	510 (354–719)	333 $(168 - 574)$ <sup>d</sup>	395 (226–617)
$HIV-RNA log10 copies/ml,$ median (IQR)	$4.6(4.3-5.3)$	$4.7(4.1-5.1)$	$4.6(4.1-5.1)$

TABLE 2. BASELINE CHARACTERISTICS OF PATIENTS INCLUDED IN THE STUDY<sup>a</sup>

<sup>a</sup>No, number; SD, standard deviation; IQR, interquartile range.

 $b$ *p*  $<$  0.05 compared to recently infected; Student's *t* test.  $c$ *p*  $<$  0.05; chi-squared test.

 $p^d p < 0.05$  compared to recently infected; Mann–Whitney *U* test.

years of age, with 20 of the 21 identifying homosexual/bisexual contact as their primary risk factor. This cluster may therefore represent an epidemiologically linked group of Spanish homosexual/bisexual males in this region of Spain. The minor cluster including the CRF02\_AG sequences from NV18, NV21, and NV37 were isolated from one Spanish male and two Spanish females between the ages of 27 and 34 years, all of whom were diagnosed within 1 month of each other and identified heterosexual contact as their primary risk factor. The minor cluster including the subtype B sequences from NV01, NV26, and NV27 consist of three Spanish males two of whom identified homosexual/bisexual contact as their primary risk factor. Interestingly, sequences from NV26 and NV27 showed almost 100% sequence identity, although no social connection could be established between the two individuals.

Subtype analysis revealed that the majority of the 72 individuals included in the study were infected with subtype B virus (64 of 72; 88.9%). Of those infected with non-B subtypes, seven (9.7%) were infected with the circulating recombinant form CRF\_02AG and one (1.4%) was infected with the subtype recombinant B/G. Of 17 newly infected patients 15 (88.2%) were infected with subtype B, one (5.9%) was infected with CRF02\_AG, and one (5.9%) with B/G. Of 55 chronically infected patients, 49 (89.1%) were infected with subtype B and six (10.9%) were infected with CRF\_02AG. Interestingly, six of the 20 individuals (30%) identifying heterosexual contact as their primary risk factor were infected with CRF02\_AG, whereas only one of 43 individuals (2.3%) identifying homosexual contact as their primary risk factor were infected with CRF02\_AG ( $p < 0.01$ , chi-squared).

In order to characterize resistance mutations present in this population all mutations in the RT gene and both major and minor resistance mutations in the PR gene were considered.<sup>10</sup> Table 3 shows the prevalence of mutations identified in each viral subtype at the time of diagnosis. Of the 72 isolates analyzed, 62 (86.1%) had at least one resistanceassociated mutations in either the RT or PR genes. Sixty of the 72 isolates (83.3%) contained resistance-associated mutations to one class of antiretroviral drug, whereas 2 (2.9%) contained resistance-associated mutations to two classes of antiretroviral drug. No isolates had mutations associated with resistance to all three drug classes.

None of the patient isolates contained major resistance mutations within the PR gene. However, 62 of the 72 isolates (86.1%) contained at least one minor mutation within the PR gene that is associated with resistance to protease inhibitors.10 Of the mutations identified, M36I and L63P were the most common, present at frequencies of 41.7% and 48.6%, respectively. Mutations in the RT region were less prevalent than those in the PR region. Of the 72 isolates analyzed, only two had mutations in RT that are associated with resistance to nucleoside or nonnucleoside reverse transcriptase inhibitors (NRTIs or NNRTIs). The K103N mutation, which confers resistance to all NNRTIs, was present in two isolates (2.8%), while Y188C, which has been associated with resistance to nevirapine, was present in one isolate (1.4%). When data from newly infected patients were analyzed alone in order to determine the level of transmitted resistance, none of the newly infected individuals carried viruses with major mutations, whereas 13 of 17 carried minor mutations in PR.

In the present study we have evaluated the prevalence of HIV-1 subtypes and resistance mutations in RT and PR in ART-naive patients at a tertiary medical center in Seville, Spain. Our data show that subtype B virus predominates in this region with a frequency similar to what has been reported for other regions of mainland  $Spin<sup>7,8</sup>$ and France.<sup>11</sup> However, the prevalence of subtype  $\overline{B}$  virus is higher in western Andalucia than has been reported in neighboring Portugal<sup>12</sup> and the Spanish Canary Islands,<sup>9</sup> likely due to the differing immigration patterns in these regions. The data presented here also demonstrate a rela-



**FIG. 1.** Phylogenetic tree derived using *pol* sequences from the 72 HIV-1 isolates from western Andalucia. The tree was rooted using the simian immunodeficiency virus strain SIVcpz as an outgroup. Bootstrap values for selected branches are shown as percentages of 500 replicates. Reference sequences for subtype B, CRF02\_AG, subtype B/G, subtype A, subtype C, subtype E, and subtype F viruses are included.

Drug class	<b>Mutations</b>	Subtype B $(n = 64)$	$CRF_02AG$ $(n = 7)$	B/G $(n = 1)$	Total $(n = 72)$
Protease inhibitors	L10I	2(3.1)			2(2.8)
	L <sub>10</sub> V	1(1.6)			1(1.4)
	K <sub>20</sub> R	4(6.3)			4(5.6)
	K20I		7(100.0)		7(9.7)
	M36I	24 (37.5)	5(71.4)	1(100.0)	30(41.7)
	M36P	2(3.1)	1(14.3)		3(4.2)
	M36L	1(1.6)			1(1.4)
	L63P	32(50.0)	2(28.6)	1(100.0)	35 (48.6)
	A71V	3(4.7)			3(4.2)
	A71T	3(4.7)			3(4.2)
	V77I	1(1.6)			1(1.4)
Reverse transcriptase inhibitors	<b>K103N</b>	1(1.6)	1(14.3)		2(2.8)
	Y188C	1(1.6)			1(1.4)
No resistance mutations		10(15.6)	$\Omega$	$\Omega$	10(13.9)
Mutations to one or more classes		54 (84.4)	7(100.0)	1(100.0)	62(86.1)
Mutations to $>1$ class		1(1.6)	1(14.3)	$\theta$	2(2.8)
Mutations to $>2$ classes		0	$\theta$	$\theta$	0

TABLE 3. PREVALENCE OF RESISTANCE MUTATIONS IN RT AND PR GENES<sup>a</sup>

aNumbers in parentheses indicate percentages.

tively high prevalence of the CRF02\_AG recombinant form in infected individuals identifying heterosexual contact as their primary risk factor compared to those identifying homosexual contact as their primary risk factor. These results may indicate that distinct virus subtype diversities exist within the different subpopulations of this geographic region, similar to what has been reported in other regions of Europe.<sup>13</sup>

Our analysis also indicates that there are few major resistance mutations present in either RT or PR in this population. Of the 72 isolates analyzed, only two had resistance mutations in RT and none had major resistance mutations in PR. While the prevalence of primary mutations in this population was low, the frequency of isolates with at least one minor mutation in the PR gene was 86.1%. The effect that these mutations have on ART in the absence of major resistance mutations is unclear. These findings are in agreement with previous reports evaluating the presence of resistance mutations in drug-naive individuals in which the prevalence of major resistance mutations is low.14,15 The low prevalence of major resistance mutations in the chronically infected group may simply reflect the persistence of virus in the absence of selective pressure from antiretrovirals. However, the low prevalence of major resistance mutations in the recently infected group likely reflects a low level of transmission of resistant strains in this population. There are reports describing higher levels of major resistance mutations in some ART-naive cohorts due to the transmission of resistant strains.<sup>1,2</sup> These differences may be due to either the virus variants circulating in that population, or the availability of antiretroviral drugs in that region.

#### **Sequence Data**

GenBank accession numbers for the sequences reported here are from EU288087 to EU288158.

#### **Acknowledgments**

Financial support was provided through Plan Nacional de Investigacion Cientifica, Desarrollo e Innovacion Tecnologica (I+D+I) PI05/0226, Instituto de Salud "Carlos III"-Fondo de Investigacion Sanitaria; Consejeria de Salud, Junta de Andalucia 0183/2006; Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía P06-CTS-01915; the Spanish Ministry of Health; and Red de Investigacion en SIDA, RD06/ 0006 (RIS network). M.J.M was supported by T32 GM07863 and T32 GM08353 from the NIH.

# **References**

- 1. Grant RM, Hecht FM, Warmerdam M, *et al.*: Time trends in primary HIV-1 drug resistance among recently infected persons. JAMA 2002;288:181–188.
- 2. Little SJ, Holte S, Routy JP, *et al.*: Antiretroviral-drug resistance among patients recently infected with HIV. N Engl J Med 2002;347:385–394.
- 3. Zolopa AR, Shafer RW, Warford A, *et al.*: HIV-1 genotypic resistance patterns predict response to saquinavir-ritonavir therapy in patients in whom previous protease inhibitor therapy had failed. Ann Intern Med 1999;131:813–821.
- 4. Maeda Y, Venzon DJ, and Mitsuya H: Altered drug sensitivity, fitness, and evolution of human immunodeficiency virus type 1 with pol gene mutations conferring multidideoxynucleoside resistance. J Infect Dis 1998;177:1207– 1213.
- 5. Martinez-Picado J, Savara AV, Sutton L, and D'Aquila RT: Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. J Virol 1999;73: 3744–3752.
- 6. Perno CF, Cozzi-Lepri A, Balotta C, *et al.*: Secondary mutations in the protease region of human immunodeficiency virus and virologic failure in drug-naive patients treated with protease inhibitor-based therapy. J Infect Dis 2001; 184:983–991.

### **HIV-1 DRUG RESISTANCE IN SOUTHERN SPAIN 887**

- 7. Casado C, Urtasun I, Martin-Walther MV, *et al.*: Genetic analysis of HIV-1 samples from Spain. J Acquir Immune Defic Syndr 2000;23:68–74.
- 8. Alvarez M, Garcia F, Martinez NM, *et al.*: Introduction of HIV type 1 non-B subtypes into Eastern Andalusia through immigration. J Med Virol 2003;70:10–13.
- 9. Holguin A, Pena MJ, Troncoso F, and Soriano V: Introduction of non-B subtypes among Spaniards newly diagnosed with HIV type 1 in the Canary Islands. AIDS Res Hum Retroviruses 2007;23:498–502.
- 10. Johnson VA, Brun-Vezinet F, Clotet B, *et al.*: Update of the drug resistance mutations in HIV-1: Fall 2006. Top HIV Med 2006;14:125–130.
- 11. Couturier E, Damond F, Roques P, *et al.*: HIV-1 diversity in France, 1996–1998. The AC 11 laboratory network. AIDS 2000;14:289–296.
- 12. Esteves A, Parreira R, Venenno T, *et al.*: Molecular epidemiology of HIV type 1 infection in Portugal: High prevalence of non-B subtypes. AIDS Res Hum Retroviruses 2002; 18:313–325.
- 13. Casado C, Urtasun I, Saragosti S, *et al.*: Different distribution of HIV type 1 genetic variants in European patients with

distinct risk practices. AIDS Res Hum Retroviruses 2000; 16:299–304.

- 14. Perno CF, Cozzi-Lepri A, Balotta C, *et al.*: Low prevalence of primary mutations associated with drug resistance in antiviral-naive patients at therapy initiation. AIDS 2002;16:619– 624.
- 15. Rodrigues R, Scherer LC, Oliveira CM, *et al.*: Low prevalence of primary antiretroviral resistance mutations and predominance of HIV-1 clade C at polymerase gene in newly diagnosed individuals from south Brazil. Virus Res 2006;116: 201–207.

Address reprint requests to: *Pilar Pérez-Romero Hospitales Universitarios Virgen del Rocío Avenida Manuel Siurot s/n 48013 Seville Spain*

*E-mail:* mariap.perez.exts@juntadeandalucia.es