

Molecular and phylogenetic analysis of HIV-1 variants circulating in Italy.

Buonaguro L.¹, Petrizzo A.¹, Tagliamonte M.¹, Vitone F.², Re M.C.^{2,3},
Pilotti E.⁴, Casoli C.⁴, Sbreglia C.⁵, Perrella O.⁵, Tornesello M.L.¹,
Buonaguro F.M.^{1*}

¹Lab of Molecular Biology and Viral Oncogenesis & AIDS Refer. Center, Ist. Naz. Tumori "Fond. G. Pascale", Naples – Italy; ²Section of Microbiology of the Department of Hematology, Oncologic Science, Anatomical Pathology and Microbiology, ³Interuniversity Consortium, National Institute Biostructure and Biosystem (INBB), Rome - Italy; ⁴Dept. Clinical Medicine, Nephrology, and Health Sciences, Univ. of Parma, Parma – Italy; ⁵VII Div. Infect. Diseases, Cotugno Hospital, Naples - Italy.

*Corresponding author:

F.M. Buonaguro, M.D.

Director

Lab. of Molecular Biology and Viral Oncogenesis

& AIDS Ref. Center

Ist. Naz. Tumori "Fond. G. Pascale"

Via Mariano Semmola, 1

80131 NAPLES - ITALY

Tel. +39-081-5903.830

Fax. +39-081-545.1276

E-mail: irccsvir@unina.it

Abstract

Objective: The continuous identification of HIV-1 non-B subtypes and recombinant forms in Italy indicates the need of constant molecular epidemiology survey of genetic forms circulating and transmitted in the resident population.

Methods: The distribution of HIV-1 subtypes has been evaluated in 25 seropositive individuals residing in Italy, most of whom were infected through a sexual route during the 1995-2005 period. Each sample has been characterized by detailed molecular and phylogenetic analyses.

Results: 17 of the 25 samples were positive at HIV-1 PCR amplification. Three samples showed a nucleotide divergence compatible with a non-B subtype classification. The phylogenetic analysis, performed on both HIV-1 *env* and *gag* regions, confirms the molecular sub-typing prediction, given that 1 sample falls into the C subtype and 2 into the G subtype. The B subtype isolates show high levels of *intra*-subtype nucleotide divergence, compatible with a long-lasting epidemic and a progressive HIV-1 molecular diversification.

Conclusion: The Italian HIV-1 epidemic is still mostly attributable to the B subtype, regardless the transmission route, which shows an increasing nucleotide heterogeneity. Heterosexual transmission and the interracial blending, however, are slowly introducing novel HIV-1 subtypes. Therefore, a molecular monitoring is needed to follow the constant evolution of the HIV-1 epidemic.

Introduction

Human immunodeficiency virus type 1 (HIV-1) shows an extensive genetic variability and can be classified into 9 phylogenetic subtypes (A–K), which are approximately equidistant from one another, and several circulating recombinant forms (CRFs), resulting from recombination events occurring between different HIV-1 subtypes co-circulating in a specific geographic region [1].

The first phase of the HIV epidemic in Italy has been mainly confined to the injecting drug users (IDUs) risk group, with an absolute predominance of HIV-1 B subtype, in accordance with other Western Countries. In particular, among the total AIDS cases reported in the adult population during the period between 1982 and 2006, 56.0% were IDUs (including also homosexual IDUs) with similar percentages in men and women groups (57.3% and 51.5%, respectively) [2]. The annual percentages of AIDS cases reported in IDUs have gradually decreased from 65.8% in 1987 to 27.6% in 2006 [2], in part as consequence of prevention programs implemented in Italy to discourage syringe sharing [3,4]. In parallel, the overall AIDS cases reported in heterosexual individuals account for the 19.5% of total epidemic cases, with a significantly higher percentage in the women category compared to men (41.2% vs 13.6%). However, the annual percentage of AIDS cases related to the heterosexual transmission has dramatically increased over the years, becoming in 2006 the most prevalent risk factor for AIDS (40.4%) [2].

Although almost 25% of heterosexual individuals diagnosed with AIDS in Italy are partners of long-term HIV-1 infected individuals, carrying a

“historical” B-subtype virus, more than 10% of them are either immigrants from endemic regions for HIV-1 (6.87%) or their Italian partners (3.03%), while the risk is unidentified for 64% of them [2]. This epidemiological evidence, based only on the AIDS reported cases and not considering all the HIV-1 infections derived also from travelling abroad, suggests that at least 10% of the viruses transmitted through heterosexual contacts could potentially belong to non-B subtypes and CRFs. In fact, HIV-1 isolates genetically related to subtypes novel to the Italian epidemic have been recently increasingly identified and described [5-17].

This has been recently reported in other European Countries, with a higher prevalence of non-B subtypes and CRFs due to an older tradition of immigration waves and much tighter historical as well as economic links with countries endemic for HIV-1 infection [18-30].

In this framework, the introduction and the possible spread of different HIV-1 subtypes and/or recombinant forms, which could require the future development of adequate diagnostic, treatment, and prevention strategies, needs to be constantly monitored.

For the present study, 25 HIV-seropositive individuals residing in Italy were enrolled at sentinel Centers, with a HIV-infection diagnosed in the period 1995-2005. The molecular study has been performed on the hypervariable C2-V3 region of the *env* gene as well as the more conserved 5' region of the *gag* p17 sequence. Three non-B-subtype HIV-1 isolates have been identified and phylogenetically classified as C (1 isolate) and G (2 isolates) subtype.

Methods

Sample Collection

Blood samples were collected from 25 HIV-positive individuals attending Italian sentinel Centers in Bologna, Parma and Naples. For all of them, HIV infection was diagnosed during the 1995 to 2005 period, and most participants were infected through sexual contact (18 of 25, 72%). HIV-1 infection was diagnosed by immunologic methods (ELISA, Western blot), and the viral load was evaluated by viral RNA quantification. The full designation of samples, according to WHO-proposed nomenclature, is CI05.00XE or CI05.00XG, where CI stands for the city of enrolment, 05 stands for the year of study, 00 for the enrolment number and E (or G) stands for *env* (or *gag*). For the sake of simplicity, however, in this paper the samples have been indicated with CI05.01 (i.e. BO05.01).

Polymerase Chain Reaction (PCR)

Peripheral blood mononuclear cells (PBMCs) were purified from fresh HIV-1-positive blood samples by Leucoprep density gradient centrifugation, and cellular lysates (approximately 6×10^6 cells) were prepared by Proteinase K digestion at 56°C.

The quality of target DNA was verified by PCR amplification of the housekeeping p53 cellular gene. The amplification of HIV-1 V3-V5 region of the *env* gene and p17 region of the *gag* gene were performed by nested PCR analysis, using 1.5×10^5 cells (corresponding to approximately 1 µg of genomic DNA) as a template. The V3-V5 region of the HIV-1 *env* gene (666 bp) was amplified, as previously described, using the primer pairs ED5-ED12

and ES7-ES8 for the first and the second round of amplification, respectively [31,32]. The p17 region (474 bp) was amplified, as previously described, using the primer pairs CL1028-CL1033 and CL1029-CL1032 for the first and the second round of amplification, respectively [14,33].

DNA Sequencing

Direct sequencing reactions were performed on PCR products purified with a rapid method developed in our laboratory, following the Sequenase protocol (United States Biochemical, Cleveland OH), modified in the labelling step (3 minutes on ice) [31,34]. The internal annealing oligonucleotides, V3B and GAG B SENSE (annealing to a 19bp fragment of the C2-V3 region and a 17bp of the p17 region, respectively), were used to prime sense sequence reactions. Sequences were then analyzed on 6% polyacrylamide wedge sequencing gel.

Analysis of Sequences

The *env* and *gag* sequences obtained were aligned. Multiple sequence alignments were performed with the MegAlign application of the Lasergene software (DNASTAR Inc., Madison, WI) using the Clustal method. Phylogenetic trees were generated by using the neighbor-joining method with the PHYLIP software package (version 3.52c; Joseph Felsenstein, University of Washington). Briefly, the SEQBOOT program was carried out to generate 100 data sets that represent randomly re-sampled versions of the input-aligned sequences, to test the reliability of the final tree topology. Evolutionary distances were estimated by the DNADIST program, using either the Kimura 2-parameter method or the maximum likelihood distance method, and the phylogenetic relationships were determined by the NEIGHBOR program. A

consensus tree was constructed using the CONSENSE program with the majority rule criterion and was drawn with the NJPLOT application.

Results

Epidemiologic and clinical parameters.

Overall, 25 HIV-1-seropositive subjects were enrolled in the study. Of them, 32% (8 of 25) were heterosexuals, 36% (9 of 25) were male homosexuals, 4% (1 of 25) was a bisexual and 28% (7 of 25) were IDUs. All but 6 subjects were native Italian, with stable or occasional sexual partners of different nationalities. In particular, the enrolled subjects included an Ethiopian (PR06.03), an Albanian (NA05.04), an Ivorian (PR06.05), a Venezuelan (PR06.06) and two Nigerians (NA05.03, PR06.07). The CD4⁺ counts ranged from 169 to 1387 cells/mm³ (median, 664 cells/mm³) (Table 1).

Amplification of Italian samples by PCR.

The C2-V5 region of the HIV-1 *env* gene and the p17 region of the *gag* gene were amplified by nested PCR using primers and conditions previously described [14]. Overall, 60% of the samples (15 of 25) were positive at the PCR amplification reactions for both *env* and *gag* sub-genomic regions. However, 64% of the samples (16 of 25) were positive at the amplification reaction for *env* and 68% of the samples (17 of 25) were positive to the amplification reaction for *gag*. Therefore, one sample (BO07.14) was positive only in *env* and two samples (NA05.05 e NA05.06) were positive only in *gag*; the remaining 8 samples were negative to both sub-genomic regions. This result might be the consequence of either mutations in the primers' annealing regions, reducing the melting temperature, or proviral DNA quantity below the threshold of sensibility. An example of the results obtained by nested PCR for *env* (666bp) and *gag* (474bp) is shown in Fig 1 A and B.

Molecular analysis of *env* gene.

The nucleotide sequence analysis has been performed on the C2-V3 region of the *env* gene, directly from the PCR products without a sub-cloning step. The Italian HIV-1 nucleotide sequences, aligned with HIV-1 reference standards of different subtypes, Groups and CRF02_AG, have been aligned with Clustal method, to determine the homology values between the analyzed samples. The results, show an average divergence of 19.66% within the Italian group, with values ranging from 7.1 to 33.9% (data not shown). The Italian sequences, showed average divergences of 17.97% versus standards of B clade; 25.87% versus standards of other HIV genotypes of Group M; 43.14% and 48.4% versus the standards of Groups N and O, respectively (Fig. 2A). Nevertheless, specific samples show a >30% divergence versus the B clade, strongly suggesting a non-B clade classification. In particular, the sample PR06.07 shows an average divergence of 15% versus the CRF02_AG and 23.2% versus all other clades (Fig. 2B); the sample PR06.03 shows an average divergence of 18% versus the C clade and 22.7% versus all other clades (Fig. 2C). Therefore, the results suggest a B-clade classification for the vast majority of the analyzed Italian sequences and a C-clade and CRF02_AG classification for the latter samples.

Molecular analysis of *gag* gene.

The nucleotide sequence analysis has been performed on the p17 region of the *gag* gene, directly from the PCR products without a sub-cloning step. As for the *env* gene, the HIV-1 nucleotide sequences have been aligned with Clustal method to HIV-1 reference standards of different subtypes, Groups and

CRF02_AG, in order to determine the homology values between the analyzed samples. The results, show an average divergence of 15.86% within the Italian group, with values ranging from 10.81 to 29.16% (data not shown). The Italian sequences, showed average divergences of 16.88% versus standards of B clade; 25.99% versus standards of other HIV genotypes of Group M; 48.54% and 44.15% versus the standards of Groups N and O, respectively (Fig. 3A). Nevertheless, specific samples show a >22% divergence versus the B clade, strongly suggesting a non-B clade classification. In particular, the sample PR06.07, also in the *gag* region, shows an average divergence of 13% versus the CRF02_AG and 24.34% versus all other clades (Fig. 3B). Similarly, the sample PR06.03, also in the *gag* region, shows an average divergence of 22% versus the C clade and 32.08% versus all other clades (Fig. 3D). Finally, the sample NA05.05, negative for *env*, shows an average divergence of 13.9% versus the CRF02_AG and 25.31% versus all other clades (Fig. 3C). Therefore, also in the *gag* region, the results suggest a B-clade classification for the vast majority of the analyzed Italian sequences and confirm a C-clade or CRF02_AG classification for the same samples identified in *env*.

Phylogenetic analysis of C2-V3 *env* and p17 *gag* sequences.

Nucleotide sequences have been pairwise aligned with HIV-1 reference standards of different subtypes and CRFs, discounting gaps due to nucleotide insertions/deletions. The alignments of HIV-1 *env* and *gag* nucleotide sequences have been used to generate phylogenetic trees by the neighbor-joining method. Confidence values for individual branches have been determined by a bootstrap analysis. The reference standards of different

subtypes as well as CRFs included in this study cluster in the expected distinct phylogenetic branches in both *env* and *gag* sub-genomic regions, indicating that the length of sequence fragment used in this analysis is sufficient for the identification of known subtypes.

All the Italian HIV-1 isolates showing a rate of nucleotide divergence less than 30% in *env* and 20% in *gag* versus the standards of the B clade, phylogenetically cluster with reference standards of this clade in both *env* and *gag* sub-genomic regions, forming several sub-clusters. Similarly, the 3 samples showing a lower nucleotide divergence both in *env* and *gag* versus standards of the C clade or the CRF02-AG, cluster with these non-B subtype standards in both sub-genomic regions, confirming the tentative subtype classification based on nucleotide divergence values (Fig. 4A and B).

Moreover, all isolates show a concordant subtype classification in both *env* and *gag* sub-genomic regions, suggesting the absence of intra-genomic recombination events.

Peptide analysis and comparison.

The V3 and p17 region amino acid sequences have been deduced for each isolate by computer analysis and aligned following their subtype classification. The V3 region of the Italian B-subtype isolates identified in this study shows an amino acid variability mainly localized outside the V3 loop region. The consensus derived from the alignment, in fact, shows that the amino acid residues conserved in 100% of the aligned sequences are all localized in the V3 loop (9 of 35, 25.71%). However, besides the subtype-specific "fingerprint" GPGR sequence at the tip of the V3 loop, less-represented sequences have

been identified at the tip (GPGG, GPGS, GPGQ), suggesting a constant diversification in the B clade (Fig. 5A).

In regards to the non-B clade isolates identified in this study, and distributed in two different clades (C and CRF02-AG), the too limited number of samples hampers the possibility to derive a consensus sequence and to infer conclusions; nevertheless, it is worthwhile to mention that they all show the “fingerprint” GPGQ tetramer at the tip of their V3 loop sequences (data not shown).

Moreover, mutations in amino acid residues conferring resistance to fusion/binding inhibitors, in association with other mutations along the env sequence not analyzed in the present study, have been observed (i.e. Q296K), suggesting the transmission of isolates resistant to this class of antiretrovirals (Fig 5A) [35-37].

The p17 region of the B clade isolates shows a low amino-acidic variability and the consensus derived from the alignment shows an overall rate of amino acid conservation of 44.44% (36 of 81 residues). None of the observed amino acid changes is found in gag residues conferring drug resistance (Fig. 5B). Similar results are observed also for non-B clade isolates.

DISCUSSION

A molecular and phylogenetic characterization was performed on HIV-1 variants identified in individuals residing in Italy and infected in the 1995-2005 period. The groups at high risk of HIV-1 infection (heterosexuals, homosexuals and IDUs) were equally represented in the present study. The C2-V5 and p17 regions of the *env* and *gag* gene have been amplified from uncultured PBMCs of Italian samples by the standardized nested PCR conditions.

Informative regions of the 2 structural *gag* and *env* genes have been studied in parallel, to have an immediate picture of the evolution pattern in viral regions under different immune pressure and to identify possible intra-genomic recombinants. In fact, considering the worldwide represented CRFs described until now, the p17 *gag* and the C2-V3 *env* regions show a different subtype designation in most of cases and are highly diagnostic for the identification of novel recombinants.

The average nucleotide divergences versus standard sequence of different subtypes suggest the presence of B and non-B subtypes among the Italian sequences identified in the present study. Moreover, within the B subtype sequences, the average divergence in the *env* C2-V3 region is 18.08%, with values ranging from 7.1 to 29.9%. On the contrary, the average divergence in the *gag* p17 region is 11.13%, with values ranging from 7.79 to 16.86%. These observed divergence values are compatible with samples identified in a geographical area characterized by a long-lasting HIV epidemic and confirm the more pronounced genetic evolution of *env* gene compared to *gag*.

The phylogenetic analysis confirmed the nucleotide divergence subtype prediction for both B and non-B-subtype classification; in particular, of the 3 non-B isolates, 2 clusters with CRF02-AG (NA05.05 and PR06.07), 1 with the C subtype (PR06.03). Moreover, for all the B- and non-B-subtype sequences, the phylogenetic classification matches in the *gag* and *env* sub-genomic regions, suggesting the absence of intra-genomic recombination events.

The phenetic analysis of the V3 region shows a significant amino acid stability in 9 residues of the V3 loop, for B-subtype isolates, confirming the strong selection for specific sequences involved in strategic functions regarding the immune response as well as cellular tropism and transmission. The different HIV-1 isolates identified in the current study show the subtype-specific tetrameric sequence at the apex of the V3 loop (GPGR for B clade, GPGQ for non-B clades), which is considered the target for the anti-V3 neutralizing antibodies. However, less-represented sequences have been identified at the tip of the B clade isolates (GPGG, GPGS, GPGQ), suggesting a constant diversification in this clade.

The overall results suggest that the B subtype is still largely predominant in the HIV-1 epidemic in Italy and is circulating among all risk groups. On the contrary, HIV-1 non-B subtypes in Italy are strictly associated with the heterosexual transmission and are identified in infections acquired in the last period (2001–2005). The three non-B-subtype HIV-1 isolates are strictly associated with heterosexual transmission. The samples PR06.03 and PR06.07, corresponding to an Ethiopian and Nigerian subject respectively, belong to the subtype/CRF predominant in their respective Country of origin, suggesting an

infection either prior to the immigration to Italy or subsequent but through heterosexual contacts with HIV-infected subjects from those Countries. The sample NA05.05, on the contrary, corresponds to an Italian-born subject with a possible partner/s from Regions with high endemicity for HIV infections.

These results confirm that in Italy, as in other Western European countries, non-B subtypes or recombinant forms are introduced by immigrants/migrants and transmitted at a low rate to the indigenous population. This would also explain the lower prevalence of non-B subtypes in Italy compared with other European countries with an older tradition of immigration waves and much tighter historic and economic links with African countries.

The presented data are representative of a nationwide molecular survey and, regardless the small sample size, are part of recurrent studies giving a constant updated picture of the genetic evolution of the HIV-1 epidemics in different risk groups in Italy.

ACKNOWLEDGEMENTS

This study was supported by grant from the Ministero Italiano della Sanità (Ricerca Corrente; Progetto Finalizzato AIDS 2006, #20G.19 and the SIVIM study group for DNA test standardization).

Reference List

1. Buonaguro L, Tornesello ML, Buonaguro FM: Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol* 2007, 81: 10209-10219.
2. Italian AIDS Operative Center: Aggiornamento dei casi di AIDS notificati Italia - Dec 2006. *Not Ist Sup Sanità* 2007, 20 S1.
3. Nicolosi A, Molinari S, Musicco M, Saracco A, Ziliani N, Lazzarin A: Positive modification of injecting behavior among intravenous heroin users from Milan and Northern Italy: 1987-1989. *Br J Addiction* 1991, 86: 91-102.
4. Rezza G, De Rose A, Dorrucchi M, Arpino C, Serafin I: Declining prevalence of HIV infection among injecting drug users entering drug treatment in Italy: 1990-1991. *Eur J Epidemiol* 1993, 9: 663-666.
5. Balotta C, Violin M, Monno L, Bagnarelli P, Riva C, Facchi G, Berlusconi A, Lippi M, Rusconi S, Clementi M, Galli M, Angarano G: Prevalence of multiple dideoxynucleoside analogue resistance (MddNR) in a multicenter cohort of HIV-1-infected Italian patients with virologic failure. *J AIDS* 2000, 24: 232-240.
6. Balotta C, Facchi G, Violin M, Van Dooren S, Cozzi-Lepri A, Forbici F, Bertoli A, Riva C, Senese D, Caramello P, Carnevale G, Rizzardini G, Cremonini L, Monno L, Rezza G, Perno CF, Ippolito G, d'Arminio-Monforte A, Vandamme AM, Moroni M: Increasing prevalence of non-clade B HIV-1 strains in heterosexual men and women, as monitored by analysis of reverse transcriptase and protease sequences. *J Acquir Immune Defic Syndr* 2001, 27: 499-505.
7. Ceccherini-Silberstein F, Erba F, Gago F, Bertoli A, Forbici F, Bellocchi MC, Gori C, d'Arrigo R, Marcon L, Balotta C, Antinori A, Monforte AD, Perno CF: Identification of the minimal conserved structure of HIV-1 protease in the presence and absence of drug pressure. *AIDS* 2004, 18: F11-F19.
8. Monno L, Brindicci G, Lo CS, Punzi G, Scarabaggio T, Riva C, Di BC, Pierotti P, Saracino A, Lagioia A, Mazzotta F, Balotta C, Angarano G: HIV-1 subtypes and circulating recombinant forms (CRFs) from HIV-1 infected patients residing in two regions of central and southern Italy. *J Med Virol* 2005, 75: 483-490.
9. Venturi G, Romano L, Carli T, Corsi P, Pippi L, Valensin PE, Zazzi M: Divergent distribution of HIV-1 drug-resistant variants on and off antiretroviral therapy. *Antivir Ther* 2002, 7: 245-250.

10. Bagnarelli P, Vecchi M, Burighel N, Bellanova D, Menzo S, Clementi M, De RA: Genotypic and phenotypic correlates of the HIV Type 1 env gene evolution in infected children with discordant response to antiretroviral therapy. *AIDS Res Hum Retroviruses* 2004, 20: 1306-1313.
11. Binley JM, Wrin T, Korber B, Zwick MB, Wang M, Chappey C, Stiegler G, Kunert R, Zolla-Pazner S, Katinger H, Petropoulos CJ, Burton DR: Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 2004, 78: 13232-13252.
12. Tramuto F, Vitale F, Bonura F, Romano N: Detection of HIV type 1 non-B subtypes in Sicily, Italy. *AIDS Res Hum Retroviruses* 2004, 20: 251-254.
13. Buonaguro L, Tagliamonte M, Tornesello ML, Ciuffreda D, Capiluppi B, Lopalco L, Lazzarin A, Tambussi G, Buonaguro FM: Molecular and phylogenetic characterization of HIV variants in Italian primary HIV infections (PHI): identification of non-B subtype variants. *Journal of Biological Regulators and Homeostatic Agents* 2002, 16: 44-48.
14. Buonaguro L, Tagliamonte M, Tornesello ML, Pilotti E, Casoli C, Lazzarin A, Tambussi G, Ciccozzi M, Rezza G, Buonaguro FM: Screening of HIV-1 isolates by *reverse* Heteroduplex Mobility Assay and identification of non-B subtypes in Italy. *J AIDS* 2004, 37: 1295-1306.
15. Buonaguro L, Tagliamonte M, Tornesello ML, Buonaguro FM: Evolution of the HIV-1 V3 region in the Italian epidemic. *New Microbiologica* 2007, 30: 1-11.
16. Buonaguro L, Tagliamonte M, Tornesello ML, Buonaguro FM: Genetic and phylogenetic evolution of HIV-1 in a low subtype heterogeneity epidemic: the Italian example. *Retrovirology* 2007, 4: 34.
17. Tagliamonte M, Vidal N, Tornesello ML, Peeters M, Buonaguro FM, Buonaguro L: Genetic and phylogenetic characterization of structural genes from non-B HIV-1 subtypes in Italy. *AIDS Res Hum Retroviruses* 2006, 22: 1045-1051.
18. Snoeck J, Van Dooren S, Van Laethem K, Derdelinckx I, Van Wijngaerden E, De Clercq E, Vandamme AM: Prevalence and origin of HIV-1 group M subtypes among patients attending a Belgian hospital in 1999. *Virus Research* 2002, 85: 95-107.
19. Fransen K, Buvé A, Nkengasong JN, Laga M, van der Groen G: Longstanding presence in Belgians of multiple non-B HIV-1 subtypes. *Lancet* 1996, 347: 1403.

20. Iversen AKN, Learn GH, Fugger L, Gerstoft J, Mullins JI, Skinhoj P: Presence of multiple HIV subtypes and a high frequency of subtype chimeric viruses in heterosexually infected women. *J Acquir Immune Defic Syndr* 1999, 22: 325-332.
21. Leitner T, Escanilla D, Marquina S, Wahlberg J, Brostrom C, Hansson HB, Uhlen M, Albert J: Biological and molecular characterization of subtype D, G, and A/ D recombinant HIV-1 transmission in Sweden. *Virology* 1995, 209: 136-146.
22. Boni J, Pyra H, Gebhardt M, Perrin L, Burgisser P, Matter L, Fierz W, Erb P, Piffaretti JC, Minder E, Grob P, Burckhardt JJ, Zwahlen M, Schupbach J: High frequency of non-B subtypes in newly diagnosed HIV-1 infections in Switzerland. *J Acquir Immune Defic Syndr* 1999, 22: 174-179.
23. Couturier E, Damond F, Roques P, Fleury HJ, Barin F, Brunet JB, Brun-Vezinet F, Simon F: HIV-1 diversity in France, 1996-1998. The AC 11 laboratory network. *AIDS* 2000, 14: 289-296.
24. Paraskevis D, Magiorkinis E, Theodoridou M, Mostrou G, Papaevangelou V, Kiosses VG, Hatzakis A, Matsianotis N: Molecular epidemiology of vertical human immunodeficiency virus type 1 transmission in Greece: evidence of non-B subtypes. *J Hum Virol* 1999, 2: 239-243.
25. Devereux H, Loveday C, Burke A, Dann L, Johnson M, Phillips A: The prevalence of non-B subtype HIV-1 in a London HIV/ AIDS outpatient clinic. *AIDS* 1999, 13: 142.
26. Holguin A, Alvarez A, Soriano V: HIV-1 subtype J recombinant viruses in Spain. *AIDS Res Hum Retroviruses* 2002, 18: 523-9.
27. Holguin A, Alvarez A, Soriano V: High prevalence of HIV-1 subtype G and natural polymorphisms at the protease gene among HIV-infected immigrants in Madrid. *AIDS* 2002, 16: 1163-70.
28. Falkensammer B, Doerler M, Kessler HH, Puchhammer-Stoeckl E, Parson W, Duftner C, Dierich MP, Stoiber H: Subtype and genotypic resistance analysis of HIV-1 infected patients in Austria. *Wien Klin Wochenschr* 2007, 119: 181-185.
29. Vidal N, Frange P, Chaix ML, Mulanga C, Lepira F, Bazepeo SE, Goujard C, Meyer L, Rouzioux C, Delaporte E, Peeters M: Characterization of an Old Complex Circulating Recombinant Form, CRF27_cpx, Originating from the Democratic Republic of Congo (DRC) and Circulating in France. *AIDS Res Hum Retroviruses* 2008, 24: 315-321.
30. Holguin A, Lospitao E, Lopez M, de Arellano ER, Pena MJ, del RJ, Martin C, Soriano V: Genetic characterization of complex inter-recombinant

HIV-1 strains circulating in Spain and reliability of distinct rapid subtyping tools. *J Med Virol* 2008, 80: 383-391.

31. Buonaguro L, Del Gaudio E, Monaco M, Greco D, Corti P, Beth-Giraldo E, Buonaguro FM, Giraldo G: Heteroduplex mobility assay and phylogenetic analysis of V3 region sequences of HIV 1 isolates from Gulu - Northern Uganda. *J Virol* 1995, 69: 7971-7981.
32. Naderi HR, Tagliamonte M, Tornesello ML, Ciccozzi M, Rezza G, Farid R, Buonaguro FM, Buonaguro L: Molecular and phylogenetic analysis of HIV-1 variants circulating among injecting drug users in Mashhad-Iran. *Infectious Agents and Cancer* 2006, 1: 4.
33. Buonaguro L, Tagliamonte M, Tornesello ML, Buonaguro FM: Evaluation of a modified version of Heteroduplex Mobility Assay for rapid screening of HIV-1 isolates in epidemics characterized by mono/ dual clade predominance. *J Virol Methods* 2005, 124: 123-134.
34. Buonaguro L, Greco D, Tornesello ML, Rago M, Zaccarelli M, Curcio F, Sguazzo N, Pipolo G, Beth-Giraldo E, Buonaguro FM, Giraldo G: Analysis of HIV-1 *env* gene V3 loop sequence in a southern italian cohort of intravenous drug users. *AIDS* 1994, 8: 268-269.
35. Kanbara K, Sato S, Tanuma J, Tamamura H, Gotoh K, Yoshimori M, Kanamoto T, Kitano M, Fujii N, Nakashima H: Biological and genetic characterization of a human immunodeficiency virus strain resistant to CXCR4 antagonist T134
11. *AIDS Res Hum Retroviruses* 2001, 17: 615-622.
36. De VK, Reymen D, Griffin P, Steinkasserer A, Werner G, Bridger GJ, Este J, James W, Henson GW, Desmyter J, Anne J, De C, I: The bicyclams, a new class of potent human immunodeficiency virus inhibitors, block viral entry after binding
4. *Antiviral Res* 1996, 29: 209-219.
37. Schols D, Este JA, Cabrera C, De CE: T-cell-line-tropic human immunodeficiency virus type 1 that is made resistant to stromal cell-derived factor 1alpha contains mutations in the envelope gp120 but does not show a switch in coreceptor use
32. *J Virol* 1998, 72: 4032-4037.

Legends to Figures

- Fig. 1 Analysis of DNA fragments obtained by nested PCR on the C2-V5 *env* (A) and the p17 *gag* (B) genes of HIV-1. Examples of specific *env* fragments of 666bp (panel A) and *gag* fragments of 474 (panel B), obtained by nested PCR have been separated on a 1% agarose. M = λ /Hind III, C1 = I round negative control; C2 = II round negative control; + = positive control.
- Fig. 2 Average nucleotide divergence of C2-V5 *env* gene versus standard sequences of different clades. The shown average divergence values have been obtained aligning the Italian sequences with standard sequences of individual clades, Groups and CRF02_AG. (A) The lowest average divergence value of the whole group of sequences is versus B standard sequences (arrow); (B) the sample PR06.07 shows the lowest average divergence value versus CRF02_AG standards (arrow); (C) the sample PR06.03 shows the lowest average divergence versus clade C standard sequences (arrow) ($p < 0.01$).
- Fig. 3 Average nucleotide divergence of p17 *gag* gene versus standard sequences of different clades. The shown average divergence values have been obtained aligning the Italian sequences with standard sequences of individual clades. (A) The lowest average divergence value of the whole group of sequences is versus B standard sequences (arrow); (B) the sample PR06.07 and (C) the sample NA05.05 show the lowest average divergence value versus CRF02_AG

standards (arrow); (D) the sample PR06.03 shows the lowest average divergence versus clade C standard sequences (arrow) ($p < 0.01$).

Fig. 4 Phylogenetic classification of the Italian samples. The Italian samples have been aligned with standard sequences of HIV-1 group M including some known CRF. (A) C2-V3 region of *env* gene and (B) p17 region of *gag* gene. Sequences from groups N and O have been used as outgroup. Italian sequences are indicated as underlined. Reliability has been estimated by boot-strap analysis. The bar shows a 10% divergence.

Fig. 5 Alignment of the amino acid sequences. (A) V3 env region and (B) p17 gag region have been aligned, and a consensus sequence (top) has been generated. Question marks (?) have been introduced in the consensus when, in the specific position, no residue is found in more than 50% of sequences. Dots (.) indicate agreement with the consensus and dashes (-) indicate a gap inserted to maintain the alignment. Gray areas indicate perfectly conserved residues. Boxed residues indicate mutations possibly conferring resistance to class of drugs.

Table 1. Epidemiologic and clinical parameters of enrolled subjects.

Sample Code	Nationality	Risk Factor	Presumed infection date	CD4+/ μ l
BO07.01	Italian	HOMOSEX	2002	886
BO07.06	Italian	IDU	1998	1110
BO07.09	Italian	HETEROSEX	1997	533
BO07.12	Italian	HOMOSEX	1999	515
BO07.13	Italian	IDU	1995	718
BO07.14	Italian	IDU	1998	659
BO07.18	Italian	HOMOSEX	2001	1387
BO07.20	Italian	IDU	2000	515
BO07.23	Italian	HETEROSEX	2002	765
BO07.24	Italian	IDU	2003	917
BO07.25	Italian	HOMOSEX	2001	462
BO07.26	Italian	IDU	2000	279
NA05.01	Italian	IDU	1999	664
NA05.02	Italian	HOMOSEX	2000	427
NA05.03	Nigerian	HETEROSEX	2002	1203
NA05.04	Albanian	HETEROSEX	2003	391
NA05.05	Italian	HETEROSEX	2005	898
NA05.06	Italian	HOMOSEX	2005	459
PR06.01	Italian	HOMOSEX	2001	764
PR06.02	Italian	BISEX	1999	169
PR06.03	Ethiopic	HETEROSEX	2002	481
PR06.04	Italian	HOMOSEX	2000	852
PR06.05	Ivorian	HETEROSEX	1998	1100
PR06.06	Venezuelan	HOMOSEX	2001	1244
PR06.07	Nigerian	HETEROSEX	2001	389

Fig. 1

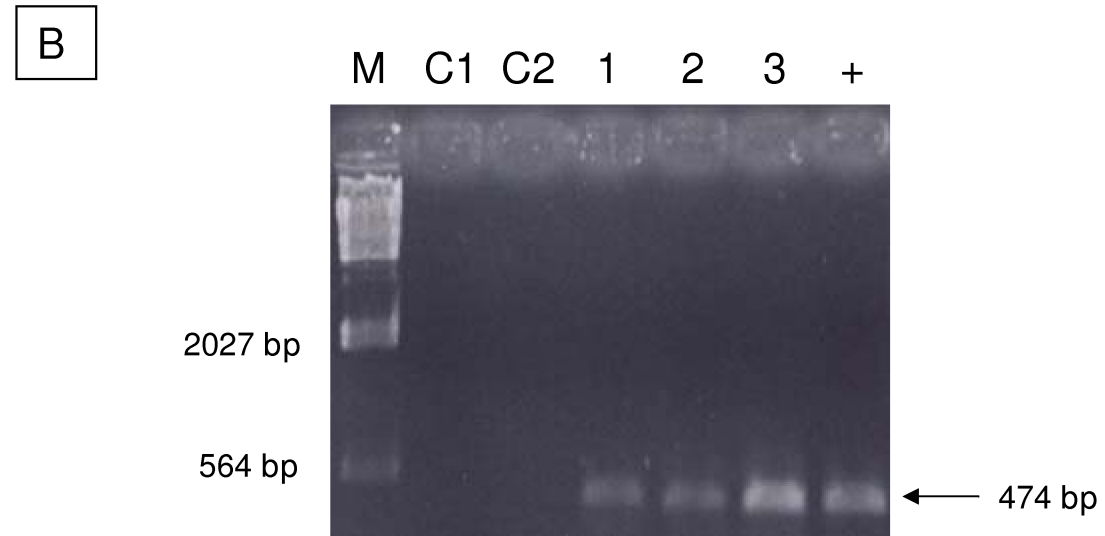
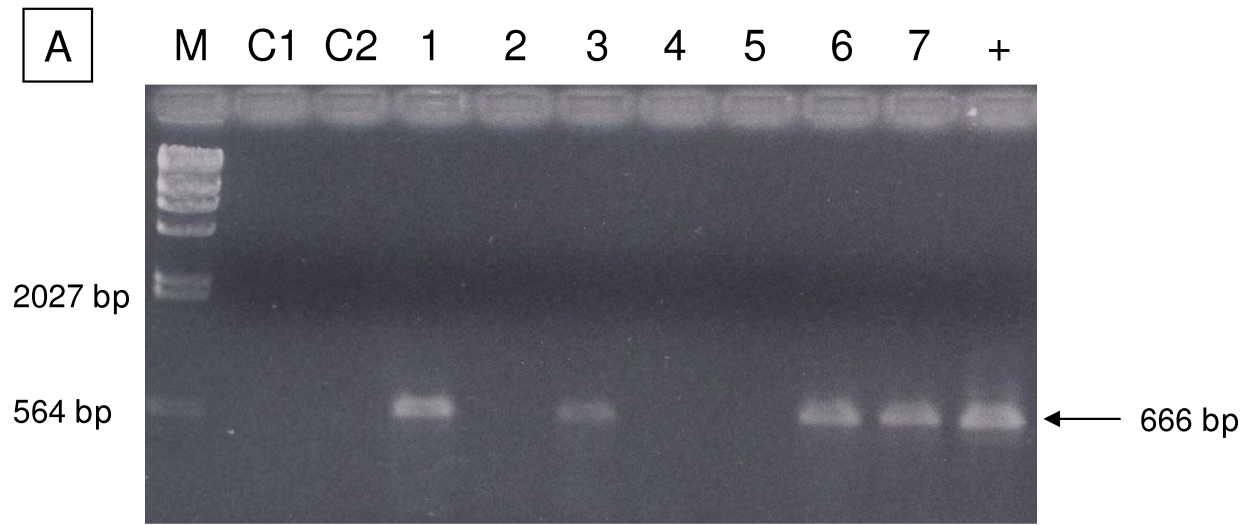


Fig. 2

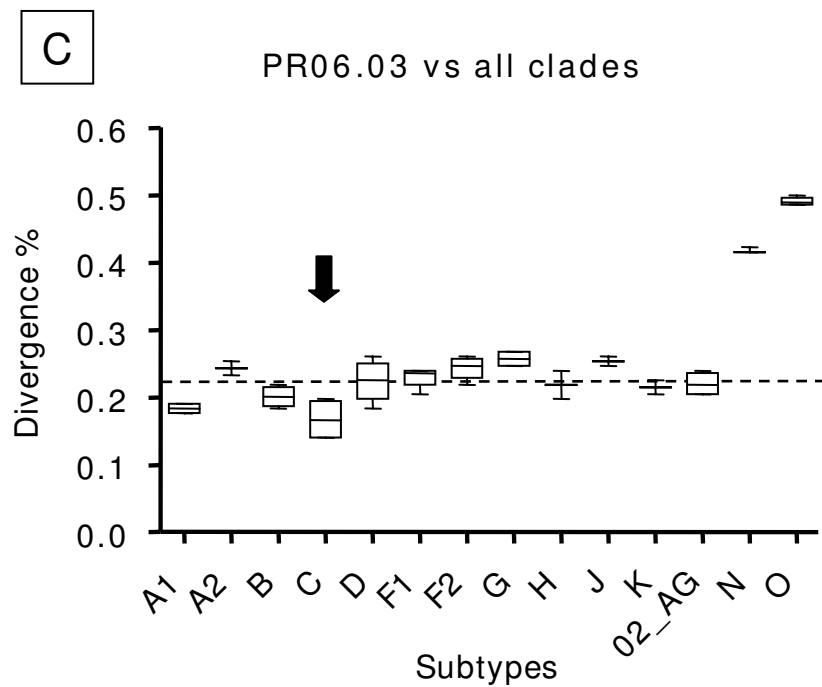
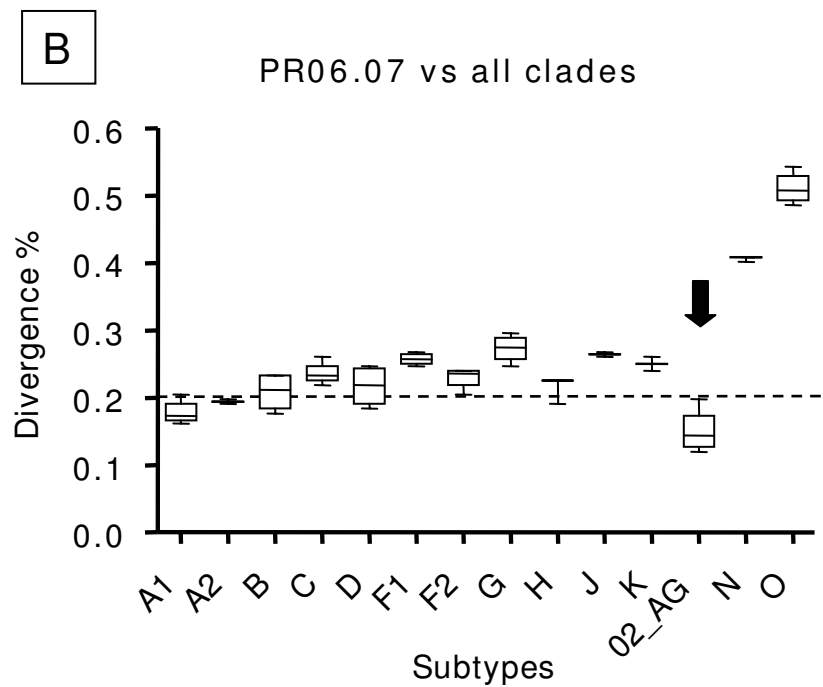
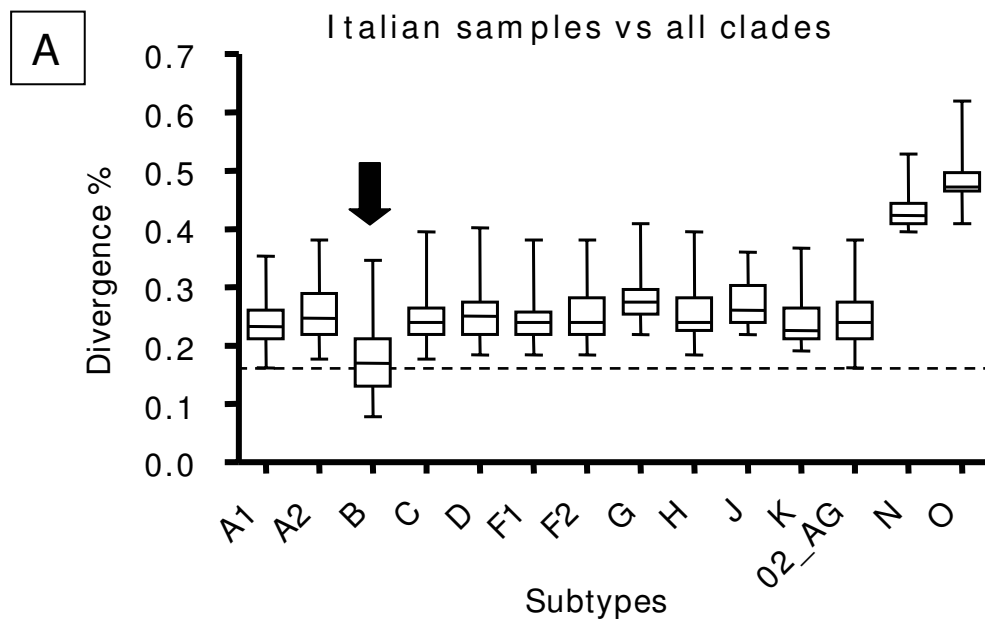


Figure 3
Fig. 3

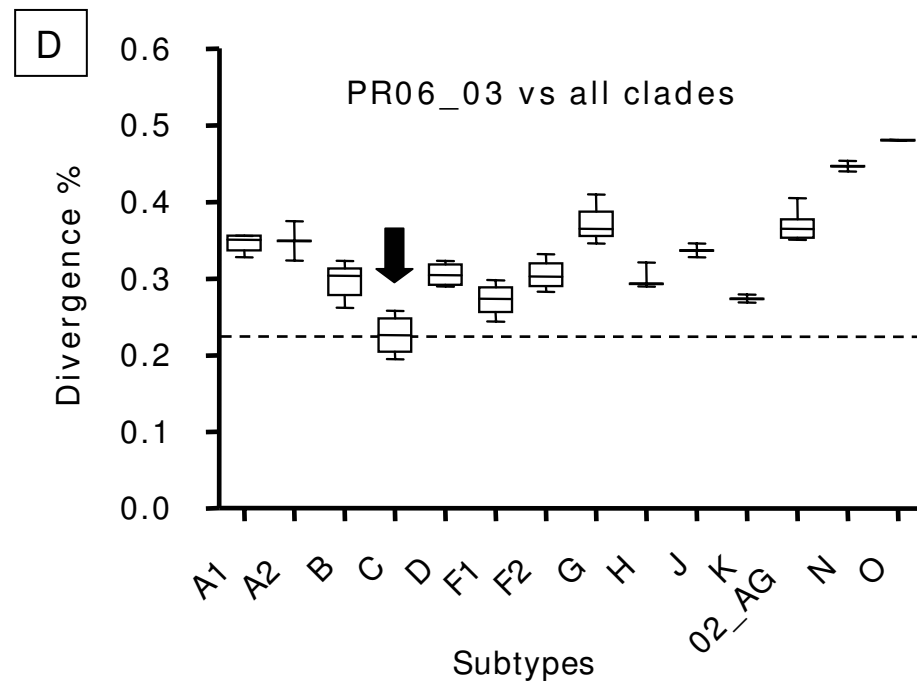
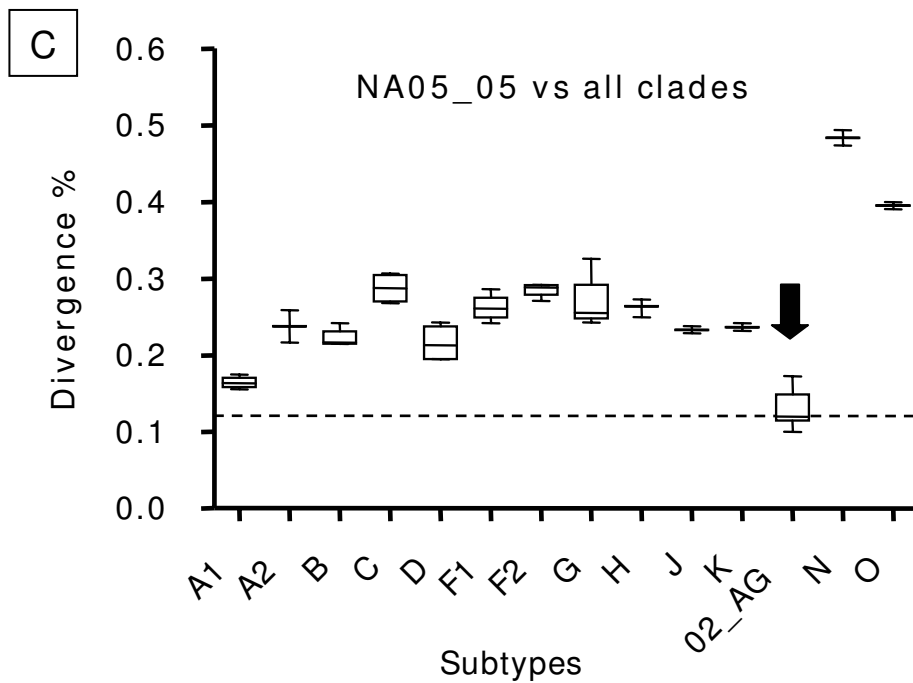
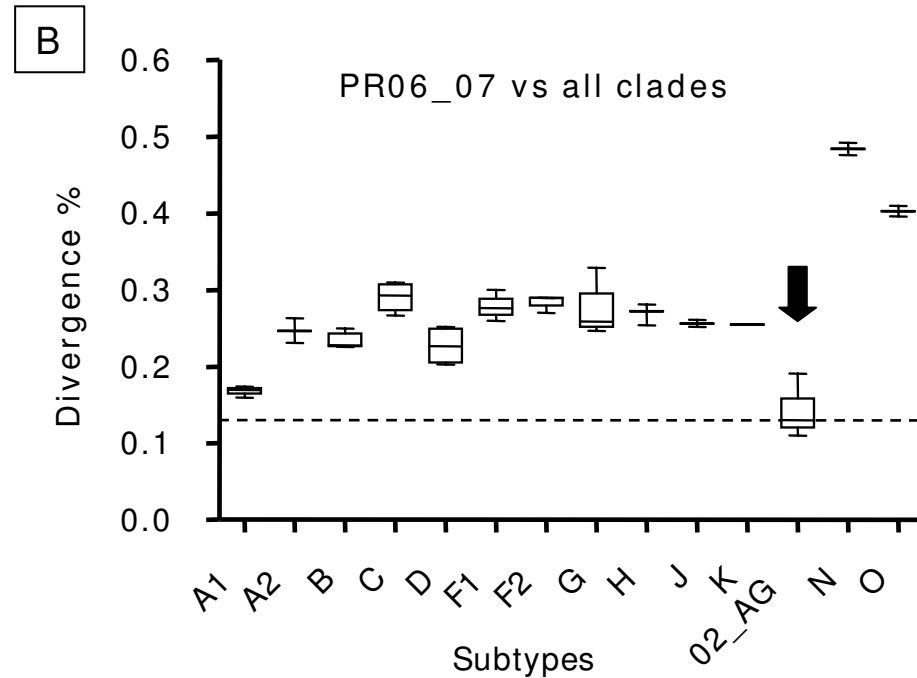
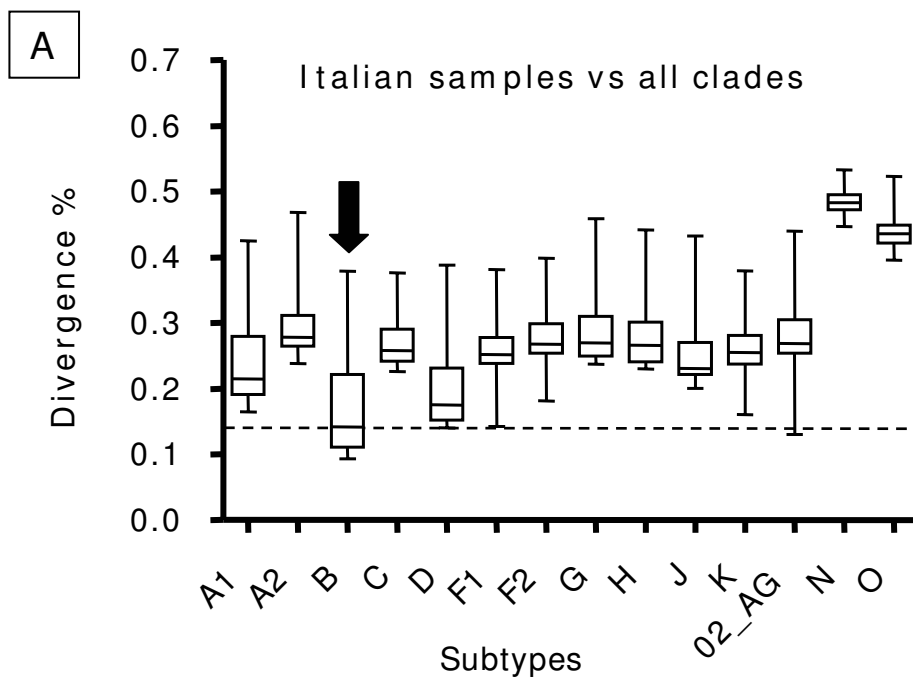


Fig. 4

A

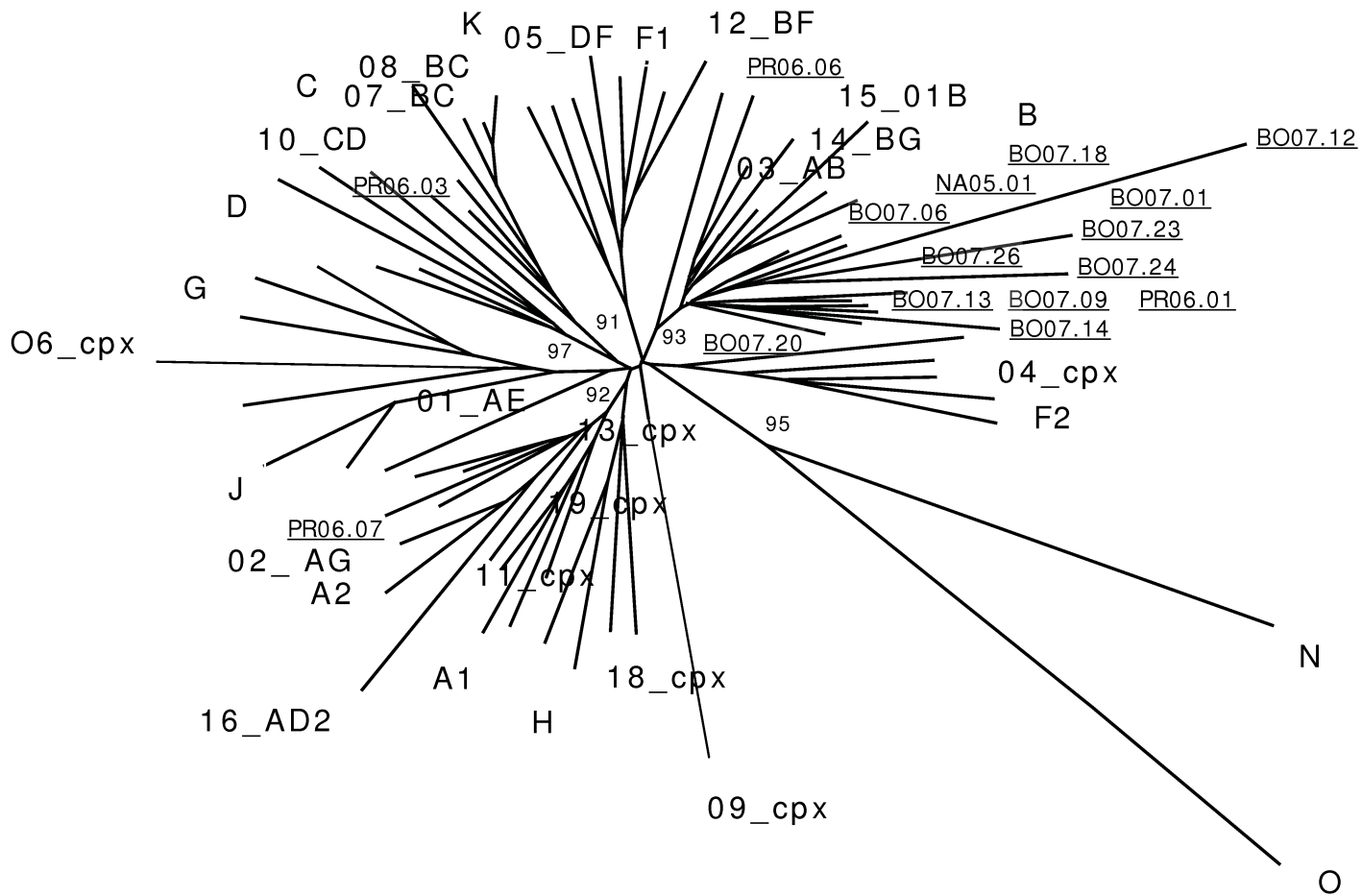


Fig. 4

Figure 5

B

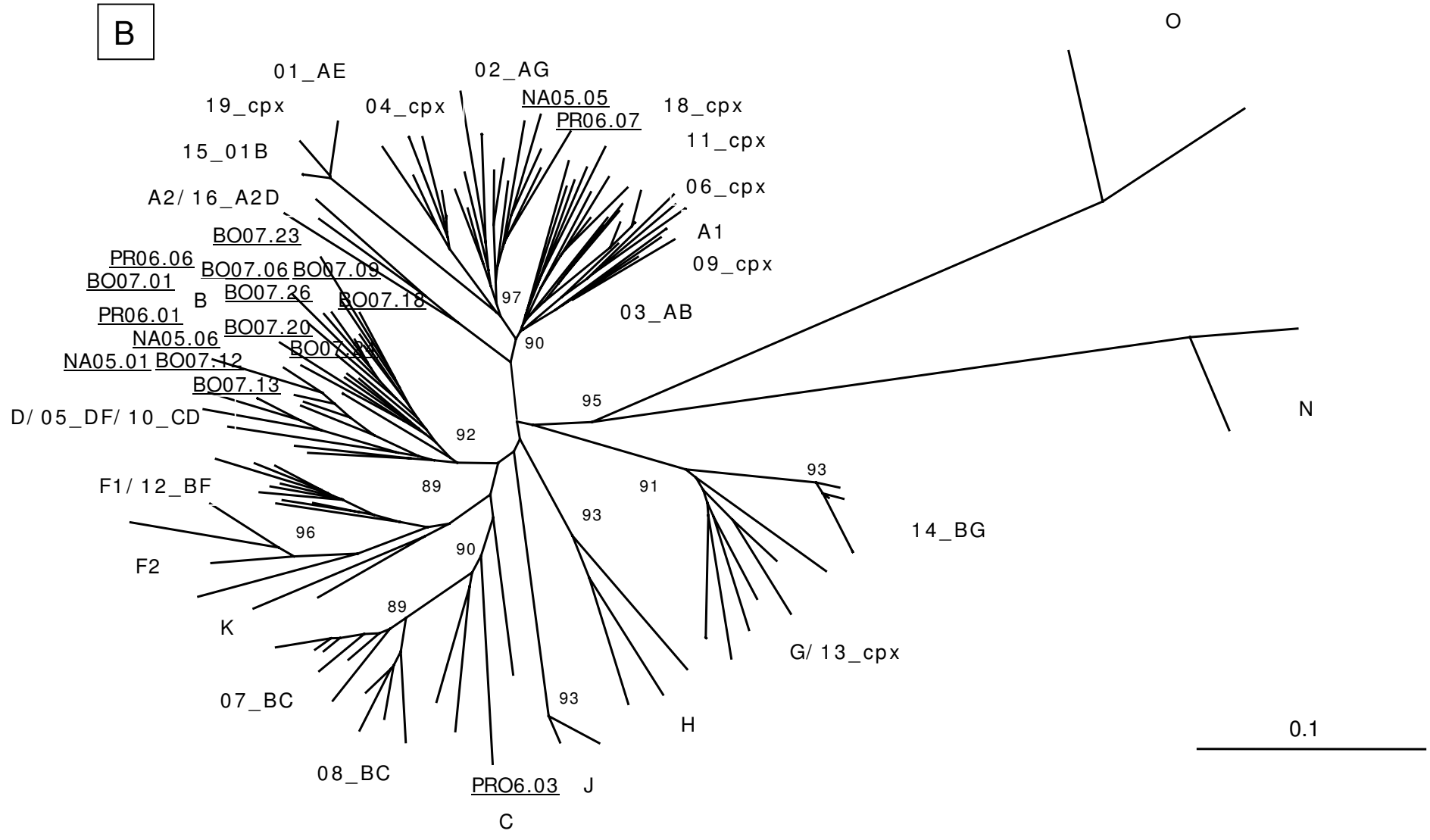


Fig. 5

	Consensus	CTRPNNNTRK?I??GPGRAF ^Y ATG?IIGDIRQAH ^C
BO07.01S.HI.....E.....	
BO07.06S.HM.....D.....SL	
BO07.09G.HM...G~V.QQ.Q...N.....	
BO07.12G.YI...QV...EK.....KHTR	
BO07.13S.QI...V...EK.....K...	
BO07.14RC~TM...GE...TR.E.TRRYKER.I	
BO07.18S.PI.....D.....Y.	
BO07.20	...G...G.RM...T...DR.V.N.....	
BO07.23	...D...A.RSLLL...S...V.AD.....FL	
BO07.24G.HM...GSS...D.....PLV	
BO07.26I.G.HM.....QV.....	
NA05.01S.PI.....T..D.....RII	
PR06.01S.HM.....TA.N.....K...	
PR06.06	..G.....G.HI...Q.V.V..Q.....	

B

	Consensus	TWATLYCVHQKIDVKDTKEAL?KIEEEQNKSKKKAQQA--?----?AA?TGNSSQVSNYP ^I VQN?QGQMVHQ ^A IS
BO07.01	..L.....E.....D.....--A-~-~-G..NN.....V	
BO07.06SL.....E.....G.....F.....I	
BO07.09	..N.....NQ...E.....C.R...QQA---AA.TGR.....L	
BO07.12	..N...R.E.....D.....L.....D.....L	
BO07.13	..N...E.R.....D.....D.....L	
BO07.14	..L.....E.....M.....A.D.R.....M	
BO07.18	..L.....E.....Q.....A.D.....I	
BO07.20	..N...EY...I...E.....~~-~-AA.G.....I	
BO07.23	..L.....E.ER...D.....-A-~-~-D..N.....L	
BO07.24R.....D.....TR...-A-~-~-D..N.....V	
BO07.26	..N...ER.....D.....-H.AAD-----G..N.....I	
NA05.01E.R...D.....T.....D.....M	
PR06.01R.R...D.....N...VLA--KETTLV..NN.....L	
PR06.06	..L.....N...A...E...K...-P...-~-~-N.....L	