HIV-1 Non-B subtypes in Italy: 
a growing trend

Massimo Ciccozzi¹, Maria Mercedes Santoro², Marta Giovanetti¹, 
Laura Andrissi⁴, Ada Bertoli²,³, Marco Ciotti³
¹Department of Infectious, Parasitic and ImmunoMediated Diseases, National Institute of Health, Rome, Italy; 
²Department of Experimental Medicine and Surgery, University of Tor Vergata, Rome, Italy; 
³Tor Vergata Foundation Hospital, Rome, Italy; 
⁴Endocrinology & Diabetes, Biomedical Campus, University of Rome, Italy

INTRODUCTION

In the last 12 years the incidence of new infections by human immunodeficiency virus type 1 (HIV-1) decreased slightly in Italy. The decrease is mainly due to the reduction of new infections among injecting drug users, while the incidence among heterosexual individuals and men who have sex with men (MSM) remains constant. In 2010, 5.5 new cases of HIV-1 infection were reported per 100,000 people. The incidence is higher in Central and North Italy compared to the South and the islands. In addition, new cases are more frequently reported among immigrants living in Italy than Italians (20 new cases among immigrants versus 4 new cases among Italians).

Infection occurs mainly through sexual contact (80.7% of all new cases, heterosexual individuals 49.8%, MSM 30.9%). The mean age of the new infected individuals is 39 years for men, and 35 years for women. It is worth noting that more than one third of the newly diagnosed individuals are in an advanced stage of disease (CD4+ cell counts <200 cells/µl).

In Italy, from the beginning of the HIV-1 epidemic, 64,000 cases of acquired immunodeficiency syndrome (AIDS) have been recorded with about 40,000 deaths. The regions with the highest number of cases are in ranking order: Lombardy, Lazio and Emilia-Romagna. Currently, new cases of AIDS are constantly decreasing thanks to the introduction of the highly

Corresponding author
Massimo Ciccozzi
Department of Infectious Diseases
Istituto Superiore di Sanità
Viale Regina Elena, 299 - Rome, Italy
E-mail: ciccozzi@iss.it

SUMMARY

Sequence analysis plays an important role in the management of patients chronically infected with HIV-1. Knowledge of the viral genotype and drug resistance mutations is crucial for the correct management of these patients. From this point of view, the experience of researchers in the HIV-1 field and the introduction of the HIV-1 genotyping resistance test has been fundamental. Several molecular tools are available to assist the provider in interpreting genotypic test results including phylogenetics. However, it should be remembered that antiretroviral drug designs, resistance studies and interpretation systems have been largely based on HIV-1 subtype B, which has been historically the most prevalent subtype in Western countries. Due to increased migration towards Europe, especially from Africa and South-East Asia, the molecular epidemiology of HIV-1 in Western Europe, including Italy, is changing. HIV-1 non-B subtypes have entered Europe and their prevalence has increased over the last years. In Italy, the estimated percentage of infection with non-B subtypes ranges from 2.4% to 19.4%. However, the true prevalence of HIV-1 non-B subtypes in this country is still not well known and probably underestimated. This may have important clinical and diagnostic implications. A strict molecular epidemiological survey and a reinforced sequencing strategy are required.

KEY WORDS: Phylogeny, HIV-1 non-B subtypes in Italy, Genotypic resistance test.
active antiretroviral therapy (Istituto Superiore di Sanità, 2011).

During the global pandemic, nine lineages of group M HIV-1 (A-D, F-H, J and K) were identified as a result of the high error rate of reverse transcriptase (RT) enzyme during virus replication and selective pressure exerted by the immune system.

Most subtypes and circulating recombinant forms (CRFs) were originally restricted to specific geographic areas and populations (Hemelaar et al., 2006). In western countries, subtype B is the predominant circulating subtype. However, in recent years non-B subtypes entered and are circulating in previously subtype-B restricted areas (Sonnerborg et al., 1997; Balotta et al., 2001; Chaix et al., 2003; Snoeck et al., 2004; Lospitao et al., 2005; Gifford et al., 2007; Palma et al., 2007; Yerly et al., 2007). Migration from or travel to countries at high prevalence of HIV-1 such as sub-Saharan Africa, South America and South-East Asia are responsible for the spread of non-B subtypes in western countries.

In Italy, the estimated percentage of infection with non-B subtypes has been reported to range from 2.4% to 19.4%, confirming a significant increase in non-B subtypes (Balotta et al., 2001; Tramuto et al., 2004; Monno et al., 2005; Ciccozzi et al., 2007; Baldanti et al., 2008; Longo et al., 2008; Buonaguro et al., 2008; Giuliani et al., 2009; Lai et al., 2010; De Paschale et al., 2011; Rossotti et al., 2011).

A recent report (Lai et al., 2010) found that the F1 subtype and the B/F recombinant form predominate among the non-B subtypes. The F1 subtype was acquired mainly through sexual contact, confirming the importance of this route in the acquisition of non-B subtypes (Tramuto et al., 2004; Buonaguro et al., 2007; Baldanti et al., 2008; Buonaguro et al., 2008; De Paschale et al., 2011; Franzetti et al., 2012). Most of the individuals infected by non-B subtypes are of African origin, although a relatively high number of individuals are Caucasians indicating a spread of non-B subtypes among Europeans (Lai et al., 2010; Saladini et al., 2012) with important clinical and diagnostic implications (Riva et al., 2010; Franzetti et al., 2012).

This review provides an overview of sequence analysis studies performed on HIV-1 non-B subtypes circulating in Italy. Particular attention has been given to studies focused on phylogeny and drug resistance. Moreover, the different propensity to develop mutations in the different subtypes was also evaluated.

**Phylogeny of HIV-1 in Italy**

Phylogenetics is a branch of molecular biology that infers knowledge about taxonomy and evolution of species (Lemey et al., 2009). It is a powerful tool widely used in the study of rapidly evolving RNA viruses that establish chronic infections such as HIV-1. In recent years a number of methods that infer phylogenetic trees have been introduced. These methods are based on genetic distances, evolutionary parsimony, maximum-likelihood and Bayesian theory (Felsenstein, 2004; Fitch et al., 1967; Saitou et al., 1987; Yang et al., 1997). Genetic distances and phylogenetic trees (coupled with a correct epidemiological design i.e., cross sectional studies), inferred via different sequence evolutionary models and model selection criteria, are normally used to assign the genotype (Pol, 2004). In addition, phylogenetic and evolutionary methods have been widely used to define circulating recombinant forms (CRFs), discovering mosaics and complex form of the virus (Zehender et al., 2010; Lai et al., 2010; Callegaro et al., 2011; Ciccozzi et al., 2011; De Paschale et al., 2011; Ciccozzi et al., 2012).

Coalescent theory and the molecular clock hypothesis are instead used to study the ancestral relationships of individuals sampled from a population (i.e. longitudinal studies) which can be inferred from a gene genealogy (phylogenetic tree) (De Oliveira et al., 2006; Salemi et al., 2008; Bon et al., 2010; Zehender et al., 2010; Callegaro et al., 2011; Ciccozzi et al., 2012).

Few cross-sectional phylogenetic studies on HIV genotypes have been carried out in Italy (Balotta et al., 2001; Visco-comandini et al., 2003; Tramuto et al., 2004; Monno et al., 2005; Tagliamonte et al., 2006; Buonaguro et al., 2007; Ciccozzi et al., 2007; Baldanti et al., 2008; Longo et al., 2008; Giuliani et al., 2009; Ciccozzi et al., 2010; Lai et al., 2010; Torti et al., 2010; De Paschale et al., 2011). The remaining studies focused mainly on HIV-1 drug resistance in relation to the viral subtype (Bracciale et al., 2009; Monno et al., 2009; Riva et al., 2010; Piralla et al., 2011; Franzetti et al., 2012; Santoro et al., 2012). Moreover, pol gene has nor-
mally been used because of the huge number of sequences available from the genotypic resistance test (GRT, performed mainly for clinical purposes). While this approach is acceptable when studying pure subtypes, it can be misleading when dealing with CRFs. The paradigm of phylogeny is shown in Figure 1. The link between epidemiological design and phylogeny inference can be seen as a hypothetical cycle starting with the specific evolutionary hypothesis and going on through data generation, data analysis and test hypothesis. The final step can also lead to a revision of the initial hypothesis or to the generation of a new hypothesis. The study design is the first important step of the cycle and it is connected with the construction of the data set. Depending on the hypothesis under study, the epidemiological design may include the whole population or local and restricted geographic areas and populations. Sequence data set may include inter-host strains sampled in a cross-sectional study or strains from the same host collected in a longitudinal study.

Depending on the scientific question raised and on the availability of additional data, the former data set is normally analyzed using maximum-likelihood criteria (Ciccozzi et al., 2011), whereas the latter is analyzed using phylodynamic criteria applying the molecular clock hypothesis and Bayesian techniques (Ciccozzi et al., 2012).

After the generation of the data set, the alignment with reference sequences, manual editing to delete “indels” (insertions/deletions), absence of contamination, and the determination of the phylogenetic signal are required. All these steps are essential preconditions for the generation of “clean” data sets. Computational method and phylogenetic inference are sensitive to the para-

**FIGURE 1** - Flow-chart representing the paradigm of phylogeny. The major steps in phylogeny inference linking experimental design and data analysis are represented.
digm: bad data will produce unreliable or even nonsensical output (Butler et al., 2010). Phylogenetic and/or phylodynamic analyses represent the “core” of the data analysis and hypothesis testing. To test for the best substitution model, to infer phylogeny using different algorithms (e.g., genetic distance, maximum likelihood, Bayesian methods), and to test the trees’ reliability (e.g. by bootstrapping), are essential steps for evolutionary analyses. This hypothetical-deductive cycle shown in Figure 1 is the stronghold of the phylogenetic approach.

**Evaluation of drug-resistance in non-B HIV-1 subtypes circulating in Italy**

GRT is a molecular tool widely used in clinical practice to guide the antiviral therapy in both drug naïve and experienced patients as recommended by the international guidelines (Hirsch et al., 2008; Vandamme et al., 2011; European AIDS Clinical Society Guidelines [EACS], 2011; The US Department of Health and Human Services [DHHS], 2012). The GRT assay is routinely performed on the *pol* region (including the all protease and the first part of RT), the main target of antiretroviral therapy. Transmitted HIV-1 drug resistance (TDR) can have a negative influence on the virological response to first-line antiretroviral therapy. This supports the implementation of GRT as a standard of care in all treatment-naïve patients to increase the success of first-line regimens. Studies of the last decade showed a limited prevalence (3.2%-11%) of TDR among individuals carrying a non-B clade in Europe, depending on the area and time of survey (Chilton et al. 2003; Descamps et al., 2010).

The detection rate of TDR in individuals living in Italy and harboring non-B subtype viruses ranges from 5% to 9% (Bracciale et al., 2009; Riva et al., 2010; Colafigli et al., 2012) which is lower than that detected in patients infected with subtype B virus (Bracciale et al., 2009; Riva et al., 2010; Salpini et al., 2011; Colafigli et al., 2012). This observation may be explained by the low exposure to treatment regimens of individuals harboring non-B subtypes, since they migrated in recent years from countries where antiretroviral drugs were not widely available. It is worth noting that these findings come from studies where non-B clades were grouped together. Indeed, TDR may vary in different subtypes as a consequence of their specific spread in various populations of different countries, as observed by a recent study showing a remarkable frequency of TDR (15.4%) in the F1 subtype-infected population, who was mainly infected by sexual transmission route (Franzetti et al., 2012). Regarding treated patients, Santoro et al. (2012) performed a comparative study of drug resistance to nucleoside/nucleotide RT inhibitors (NRTIs), non-NRTI (NNRTIs), and protease inhibitors (PIs) among B and the most prevalent non-B HIV-1 subtypes (F, C and CRF02_AG) in Central Italy from 2001-2008. Among patients treated with NRTI and with experience to thymidine analogues (TA), TA mutations of pathway 1 (TAMs1) M41L and L210W were less prevalent in CRF02_AG, while TAMs2 T215F and K219E were more prevalent in the F subtype. In NRTI-treated patients, with experience with abacavir, didanosine, tenofovir, or stavudine, the K65R mutation was mostly prevalent in the C subtype. In NNRTI-treated patients infected by the C subtype the prevalence of K103N was lower than in patients infected with other subtypes, while the prevalence of Y181C and Y188L was higher compared to subtype B. The prevalence of Y181C was higher also in subtype F as compared to subtype B. In patients treated with PIs, L89V was predominantly found in CRF02_AG, while the TPV resistance mutation T74P was predominantly found in the C subtype. The authors believe that the different prevalence of mutations among B and non-B HIV-1 subtypes analyzed is due to the genetic diversity within HIV-1 group M (as largely discussed in the following paragraph), and this viral characteristic could play a role in the development of resistance to antiviral drugs.

Monno et al. (2009) used Virtual Phenotype with the purpose of improving drug resistance interpretation algorithms for non-B subtypes. In particular, they analyzed the protease and RT sequences of naïve and drug-experienced subjects with non-B HIV infection to identify some unknown resistance mutations in non-B strains which might be relevant for resistance to current antiretroviral drugs and which are not included in the list of known resistance mutations for B subtype. There were no substantial differences regarding known resistance-associated mutations and the newly emerging substitutions between non-B and B subtype strains. This evidence has
been recently supported by a study on the impact of specific patterns of mutations on virological response at 12 weeks of treatment in patients harboring different HIV-1 subtypes (Franzetti et al., 2012). Indeed, no difference in response was detected in patients harboring B and non-B subtypes with the same drug resistance pattern.

**Subtype propensities to develop mutations**

Antiretroviral drug designs, resistance mutation studies, and interpretation systems have been largely based on subtype B HIV-1 group M viruses, which represent the predominant variant in developed countries (Kantor et al., 2005; Peeters et al., 2010; Hemelaar et al., 2011). However, improving the knowledge of resistance patterns of non-B subtypes and their interpretation has become mandatory not only because antiretroviral therapy is being introduced in countries where non-B subtypes are driving the epidemic, but also because the number of infections by these variants is increasing sharply in several European countries including Italy (Balotta et al., 2001; Monno et al., 2005; Lai et al., 2010; Torti et al., 2010; De Pascale et al., 2011; Véras et al., 2011; Santoro et al., 2012). In this regard, it should be noted that mutations that cause resistance in subtype B viruses also cause resistance in each of the other non-B subtypes. However, in different subtypes the development of specific mutations varies, and this can be explained mainly by the intrinsic properties of the virus and not only by different pressure of antiretroviral drugs as suggested by Santoro et al. (2012). In particular, three factors are involved in this phenomenon: a) **Inter-subtype differences in codon usage.** Subtype differences in nucleotide and mutational motifs, which are defined as the number of transitions or transversions needed to develop resistance to different classes of antiretroviral drugs, may affect the genetic barrier for resistance, as shown in the development of mutations associated with resistance to NNRTIs at the codon 106 of the RT (Brenner et al., 2006). Indeed, the propensity for subtype C viruses to develop V106M during NNRTI treatment - rather than V106A, which is more commonly observed in subtype B viruses - results from the fact that V106 is encoded by GTA in subtype B viruses and GTG in subtype C viruses. A single G-to-A transition at the first position of codon 106 in subtype C viruses results in V106M, which confers high-level resistance to efavirenz and nevirapine. In contrast, in subtype B viruses, V106M requires two nucleotide substitutions (GTA-ATG) and therefore occurs infrequently (Brenner et al., 2003; Grossman et al. 2004). A similar phenomenon has been observed at protease codons 74 in C subtypes, 82 in G subtypes, and 89 in several non-B subtypes such as C and CRF02_AG (van De Vijer et al., 2006; Palma et al., 2012; Santoro et al., 2012). It should be remembered that some natural polymorphisms (such as T74S and L89M) could favor the emergence of resistance mutations. For example, the presence of methionine at protease position 89 in non-B subtype viruses has been proposed to favor the selection of mutation I/V at position 89 and the major mutation L90M associated with resistance to PIs, rather than the lysine in subtype B (Abecasis et al., 2005). Interestingly, the two protease positions 74 and 89 seem to be closely related to each other. As shown by Gonzales et al. (2008), the protease mutation L89I/V is stabilized by the acquisition of T74S in subtype G. It is conceivable that in subtype C, the mutation T74S, a frequent protease polymorphism in drug-naïve patients, could favor the 89I/V selection under treatment. Inter-subtype differences are also found in the integrase gene at codons 140, 148, 151, 157, and 160 (Maïga et al., 2009). These differences may determine a higher genetic barrier for the development of some major mutations associated with resistance to integrase inhibitors (such as G140S and Q148R/H/K) in subtypes A, CRF02_AG, and C in comparison to B subtypes (Maïga et al., 2009; Brenner et al., 2011).

b) **Inter-subtype amino acid differences that can create subtle structural differences in the targets of therapy.** In this situation, different mutations emerge under the same drug pressure. For example, subtype B-infected patients receiving nelfinavir are more likely to develop D30N than those with viruses belonging to subtypes C, F, G and CRF01_AE, which are more likely to develop L90M or N88S (Cane et al., 2001; Sugiura et al., 2002; Calazans et al., 2005). c) **Inter-subtype differences in the sequence context surrounding a nucleotide whose substitution results in a drug resistance mutation.** Several studies highlighted a higher prevalence of K65R mutation in the RT of patients infected with subtype
C viruses and that experienced stavudine, abacavir, didanosine or tenoforv than in those infected by other subtypes such as F, CRF02_AG, and B (Garcia-Lerma et al., 2003; Doualla-Bell et al., 2006; Miller et al., 2007). Biochemical findings suggest that a nucleotide template-based mechanism facilitates the acquisition of the K65R mutation in subtype C (Invernizzi et al., 2005; Coutsinos et al., 2009). All these findings suggest that some natural variations in non-B subtypes may occur with possible clinical implications in the management of patients infected by these viral variants.

CONCLUSIONS

Sequence analysis plays a crucial role in the management of HIV-1 infected patients. HIV-1 genotyping permits not only the detection of mutations responsible for drug resistance but it also allows surveillance of epidemiological trends in various geographic areas and populations. Information regarding resistance to antiretroviral drugs is mainly derived from patients infected with HIV-1 subtype B. However, the spread of non-B viruses in Western countries and the introduction of antiretroviral drugs in developing countries require a deeper knowledge of anti-HIV drug-resistance patterns in non-B strains. Sequences obtained by genotypic assays can be also used to evaluate the HIV-1 genetic variability. A long time has passed since the first publication on the prevalence of HIV-1 non-B subtypes in Italy (Balotta et al., 2001). Despite that, the true prevalence of HIV-1 non-B subtypes in Italy is not well known (Ciccozzi et al. (2010) and probably underestimated. Indeed, most of the studies performed in Italy included limited geographic areas (Tramuto et al., 2004; Monno et al., 2005; De Paschale et al., 2006; Baldanti et al., 2008; Longo et al., 2008; Giuliani et al., 2009; Zehender et al., 2010; Torti et al., 2010; Callegaro et al., 2011; Piralla et al., 2011; Santoro et al., 2012). The analytical power of the phylogenetic methods available today should prompt researchers to use datasets as large as possible to monitor the epidemiological changes of HIV-1 over time. Molecular characterization of several viral sub-genomic regions is advisable. Performing phylogenetic analyses only on the pol genomic region could result in the lack of identification of novel CRFs in the Italian epidemic. Therefore, where possible, the strategy of sequencing should be changed, and major efforts should be directed towards other sub-genomic regions such as Env and Gag regions or when possible consider the whole genome to better identify the CRFs. Furthermore, it should be kept in mind that hidden modifications of HIV-1 genome in non-B clades under antiviral pressure, favour the appearance of drugs resistance. The estimate of putative origin and time frame introduction of new HIV-1 clades and CRFs in Italy is another important aspect to tackle. Monitoring the genetic evolution of HIV-1 using large dataset represents an essential strategy to control the local as well as global HIV-1 epidemic and to develop efficient preventive and therapeutic strategies with a major impact in clinical practice.

REFERENCES


Descamps D., Chaix M.L., Montes B., Pakianather S.,


line at: www.epicentro.iss.it/problemi/aids/pdf/Coa2011_commento.pdf


CRF02_AG) in Italy. AIDS Res. Hum. Retrovir. [Epub ahead of print].