Transmitted drug resistance and phylogenetic analysis of HIV CRF01_AE in Northern Vietnam

Irene Bontell a,*, Do Duy Cuong b,c, Eva Agneskog d, Vinod Diwan c, Mattias Larsson c,e, Anders Sönnerborg a,d

a Department of Medicine, Karolinska Institutet, 141 86 Huddinge, Sweden
b Infectious Diseases Department, Bach Mai Hospital, Hanoi, Vietnam
c Department of Public Health Sciences, Division of Global Health, Karolinska Institutet, 171 77 Solna, Sweden
d Department of Laboratory Medicine, Karolinska Institutet, 141 86 Huddinge, Sweden
e Oxford University Clinical Research Unit, Hanoi, Vietnam

ARTICLE INFO

Article history:
Received 3 February 2011
Received in revised form 19 April 2011
Accepted 22 April 2011
Available online xxx

Keywords:
HIV
Vietnam
CRF01_AE
TDRM
tMRCA

ABSTRACT

The HIV epidemic in Vietnam began in injecting drug users (IDUs), but increasingly affects the general population. It is therefore important to monitor the spread of infection and, since antiretroviral therapy (ART) is now used more frequently, the prevalence of transmitted drug resistance. Sixty-three 1000 bp pol-gene sequences were generated from treatment-naive HIV-1 CRF01_AE infected patients from four clinics in Northern Vietnam. Four drug resistance mutations; Y181C, L210W, L74I and V75M, were found in four different patients, giving a prevalence of 6.3% (4/63). Earlier studies have shown a lower prevalence and the transmission rate should be regularly monitored prospectively in Vietnam. Additional CRF01_AE (N = 190) and outgroup subtype B sequences (N = 4) were retrieved from databases and included for phylogenetic analysis and calculations of the time of the most recent common ancestor (tMRCA). The 63 samples from our study clustered into two distinct groups; one small clade (N = 3) that had a tMRCA in year 1997.5 and a larger group with an estimated tMRCA in 1989.8. The Vietnamese samples in the large group were distinct from CRF01_AE sequences from Thailand, but closely related to previously sequenced isolates from Vietnam, southern China and the Czech Republic, while the samples in the smaller clade appeared to represent a more recent introduction from Southern Vietnam. Our results showed that sequences from IDUs were intermingled with sequences from sexually infected patients, indicating frequent exchange of virus between the transmission risk groups in Northern Vietnam.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The first documented Vietnamese HIV-infection in Ho Chi Minh City in 1990 was a subtype B virus (Lindan et al., 1997), but since then the epidemic has been dominated by the recombinant strain CRF01_AE, which is the predominant genotype in South-East Asia, including Thailand and Southern China (Hemelaar et al., 2006). HIV transmission in Vietnam has so far largely been driven by intravenous drug users (IDUs), and a rapid increase in prevalence from 10.9% to 29.4% was seen in this group between 1996 and 2002 (UNAIDS, 2010). Although the IDU prevalence has since then decreased to 18.4% in 2009, the total HIV epidemic in Vietnam is still on the rise. The estimated number of people living with HIV in Vietnam increased from 160,000 in 2001 to 290,000 in 2007 (UNAIDS, 2008) and a slow but steady increase of the adult HIV prevalence is projected from 0.44% in 2010 to 0.47% in 2012 (UNAIDS, 2010). The HIV epidemic in Vietnam is still limited and increased political recognition of the problem in recent years have resulted in significant efforts to combat the infection both on the prevention and treatment side, including needle exchange programs for IDUs as well as rapid expansion of free antiretroviral therapy (ART) access (UNAIDS, 2010).

In countries where ART has been available for a long time, transmitted drug resistance mutations (TDRMs) are present in a significant proportion of treatment-naive patients; an American study reported a total TDRM prevalence of 14.6% in a study including 11 surveillance areas in the US (Wheeler et al., 2006) and a large European study with data from 20 countries found an

Abbreviations: ART, antiretroviral therapy; IDU, intravenous drug user; TDRM, transmitted drug resistance mutation; tMRCA, time of the most recent common ancestor.

* Corresponding author at: Division of Clinical Microbiology, F68, Karolinska Institutet, Department of Medicine, Huddinge 141 86 Stockholm, Sweden.
Tel.: +46 858 587 935/739 799 160.
E-mail addresses: irene.bontell@ki.se, irenebontell@gmail.com (I. Bontell).

1567-1348/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
overall TDRM prevalence of 8.4% (Vercauteren et al., 2009). These high levels are largely explained by the long history of ART including the early use of suboptimal therapies in these countries. Recent ART roll-outs in resource-limited settings utilize more potent regimens with higher resistance thresholds, but the frequent absence of viral load testing and limited availability of second-line ART may result in delayed treatment switches, promoting TDRM development. WHO therefore recommends surveillance of transmitted drug resistant HIV in countries scaling up ART access (Bennett et al., 2008).

The spread of HIV in Vietnam increasingly appears to occur through sexual transmission (Thanh et al., 2009), which suggests that the epidemic may become more and more difficult to control. It is thus important to monitor infection patterns and the prevalence of TDRMs in order to direct diagnostic and treatment efforts in an efficient manner to minimize the number of new infections. The aims of this study were to assess the prevalence of TDRMs in a North Vietnamese cohort consisting of both sexually infected and drug using patients, and to perform phylogenetic analyses including molecular clock calculations to investigate the HIV transmission patterns in this area.

2. Materials and methods

2.1. Study population

Baseline samples were collected at the time of ART initiation from sixty-three patients in four districts in North-East Vietnam: Dong Trieu (N = 7), Uong Bi (N = 8), yen Hung (N = 4) and Ha Long (N = 44). All patients were ethnically Vietnamese (Kinh), 29 were IDUs, 27 were sexually infected and 7 had an unknown mode of transmission. Thirty patients (47.6%) had a viral load above 100,000 copies/ml at baseline, 25 (39.7%) had between 10,000–100,000 and 8 (12.7%) had less than 10,000 copies/ml. Most patients had only recently been diagnosed with HIV (median 3.05 months before ART initiation, IQR: 1.65–14.2 months), but the generally low level of CD4 cells (median 56 CD4-cells/μl, IQR: CD4 26–163) suggested that most of the patients had been infected several years before diagnosis.

All participating patients belong to a study cohort for directly observed treatment with antiretroviral drugs (DOTARV), where samples collected between December 2008 and January 2009 were included in the current study (N = 66). A small number of patients (N = 3) were excluded after admitting previous exposure to ART. Approval has been obtained from ethical committees in Vietnam as well as Sweden.

2.2. Amplification and sequencing

Viral RNA was isolated from 1 ml plasma, which was concentrated through high-speed centrifugation (20,000 × g for 80 min at 4 °C), and 140 μl was used for RNA extraction using QiAamp ViralRNA kit (QiAgen GmbH, Hilden, Germany) according to the manufacturer’s instructions. cDNA was synthesized using SuperScript III First-Strand Synthesis Supermix (Invitrogen, Carlsbad, CA, USA) with random hexamer primers, and a product spanning protease and the first two-thirds of reverse transcriptase gene of the HIV-1 pol-gene (ref HKB2: 2135–3338) was amplified using the primers JA204F-AR (5′-CTCAGCAAGCAGGAAACACAGC-3′) and JA205R-AR (5′-TTTCTCCATATACTGGAATAT-3′). PCR-products were purified using the QIAquick PCR-purification kit (QiAgen GmbH, Hilden, Germany) and sent to Eurofins MWG Operon, Ebersberg, Germany for sequencing with the PCR-primers JA204F-AR and JA205R-AR and plus an additional primer, SeqR-AR (5′-TACATA-CAAGCTCATCCTGATTG-3′).

2.3. Baseline resistance

Sixty-three pol-gene sequences obtained from ART-naive Vietnamese HIV-patients were aligned and edited using the BioEdit and ReCall software (Hall, 1999; Harrigan et al., 2002) and a consensus sequence spanning 1000 bp was created for each sample, covering codons 1–99 for the protease gene and 1–234 for the reverse transcriptase gene. Secondary peaks were called automatically in ReCall if they reached ≥20% of the primary peak, but visual inspection of chromatograms was also done and minor manual adjustments were made. All sequences are available in GenBank (accession no HQ852853–HQ852915). Genotypic resistance analyses of all sequences were performed using the Stanford HIVdb Sequence Analysis (http://sierra2.stanford.edu/sierra/servlet/JSierra?action=sequencelinput, Liu and Shafer, 2006), and detected resistance mutations were compared against the TDRM surveillance list (Bennett et al., 2009) as well as the IAS-USA 2010 update (Johnson et al., 2010). Subtype classification was done using the REGA HIV Subtyping tool (de Oliveira et al., 2005).

2.4. Phylogenetic analysis

In addition to the 62 Vietnamese sequences obtained in the current study, a total of 194 reference sequences were included in the phylogenetic analysis. All full-length CRF01_AE strains available in the Los Alamos database were used (N = 71). Sixty-nine CRF01_AE sequences were retrieved from patients included in the national Swedish database InCare HIV, where the first available sequence from each patient was used and no more than two sequences from the same country and sampling year were included. In addition, 50 sequences were retrieved from GenBank on the basis of high BLAST similarity to the Vietnamese samples. Finally, four subtype B reference strains from Los Alamos were included as outgroup.

Alignments were made using ClustalX2 (Larkin et al., 2007) and phylogenetic analyses were performed in BEAST v1.6.1 (Drummond and Rambaut, 2007). The GTR substitution model with inverse gamma distribution (4 categories), empirical base frequencies and three codon partitions were used in all BEAST runs. Three molecular clock models (‘Strict, ‘Relaxed: exponential’ and ‘Relaxed: log-normal’) were tested in combination with five different coalescent tree priors (‘Constant Size’, ‘Exponential Growth’, ‘Logistic Growth’, ‘Bayesian Skyline’ and ‘GMRF Bayesian Skyride’), resulting in a total of 15 parallel analyses. Each analysis was run for 30 million generations and sampled every 3000th generation. Log-files were analyzed in Tracer v1.6.1 (Drummond and Rambaut, 2007), where Bayes Factor calculations were performed to determine which model was most appropriate for the data. The best model, using the Relaxed: log-normal clock with Logistic growth tree prior (‘ln_log’), was significantly better compared to most other models (Bayes factor range 17.5–300.2). However, the difference to the model using Relaxed: log-normal clock with Exponential growth tree prior (‘ln_exp’) was less pronounced at 8.2. These two models were therefore used for further analyses where each model was run in triplicate, using one UPGMA generated and two different random starting trees, for 100 million generations each, sampled every 10,000 generations. These six runs showed comparable performances (Bayes factor range 0.985–4.184), with the highest likelihood for the ‘ln_exp’ run with random starting Tree 2. The 10,000 sampled trees from this run were annotated using TreeAnnotator v1.6.1 and visualized in FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). Sampling dates for all included samples were used to calibrate the molecular clock and a previous estimate of tMRCA in the year 1975.5 for CRF01_AE (Abecasis et al., 2009) was used as a prior for the CRF01_AE taxon, which contained all but the four subtype B
sequences (the prior was set to Normal distribution, 35.5 ± 2 years since the last year of sampling, 2009).

3. Results and discussion

3.1. Drug resistance mutations in ART-naive patients

All study participants were found to be infected with HIV-1 subtype CRF01_AE with ≥95% bootstrap support. The 63 included samples all originated from ART-naive individuals and most viruses were fully susceptible to all protease and reverse transcriptase inhibitors; 39 (61.9%) had no resistance associated mutations at all, while 20 sequences (31.7%) had one or two polymorphic mutations that frequently occur in untreated patients. Four patients were, however, infected with viruses carrying transmitted resistance mutations, giving a TDM prevalence of 6.3%. An overview of all detected resistance associated mutations is shown in Table 1.

Three of the TDRMs present in the analyzed samples confer reduced susceptibility to NRTIs; L74I (N = 1) and V75M (N = 1) confer low-level resistance to ddl (both), d4T (V75M) and ABC (L74I), while L210W (N = 1) causes a low-level of resistance to all NRTIs except 3TC and FTC. The fourth TDRM was Y181C (N = 1), which provides intermediate to high level of resistance to all NNRTIs. Minor mutations found for reverse transcriptase were: A98G (N = 1), V179D (N = 2), V106I (N = 9), while L10I/V was found in the protease region of 18 sequences. No clinically significant resistance mutation for protease inhibitors was found in this study.

The resistance mutations detected in this study, Y181C (in a sexually infected patient), L210W (in an IDU), L74I (in an IDU) and V75M (in a patient with unknown mode of transmission) are in line with what could be expected, as the first-line treatment in Vietnam since the nationwide PEPFAR-funded ART roll-out in 2005 has been 3TC + d4T/3TC + NVP/Efavirenz. Protease inhibitors have not been widely used, which is mirrored in the absence of PI-associated mutations. The total TDM prevalence observed in this study (6.3%) is slightly higher compared with other recent studies performed in the same geographic region: China 3.8%, Vietnam 2.9%, Thailand 2%, Cambodia 1.5% (Liao et al., 2009; Ishizaki et al., 2009; Apisarnthanarak et al., 2008; Nouhin et al., 2009). However, none of the study participants had virus harboring more than one TDM and the total prevalence for the three relevant drug classes were thus 4.7% (NRTI), 1.6% (NNRTI) and 0% (PI), all falling below the 5% threshold level defined by the World Health Organization (Bennett et al., 2008). Apart from the patient who had virus with the Y181C mutation, the TDRMs detected in our cohort of ART-naive patients in North Vietnam are of limited clinical importance and do not rule out the use of the standard first-line treatment regimen. However, in view of the increasing use of different antiretroviral drugs in Vietnam it is important to monitor the rate of TDRMs on a regular basis.

3.2. Phylogenetic relationships and tMRCA calculations

Sixty-three pol-sequences from ART-naive Vietnamese HIV-patients were aligned with 190 CRF01_AE and four subtype B sequences retrieved from public and local databases. The initial analysis in BEAST revealed three clearly demarcated clades which all had a posterior probability support = 1. These defined three taxa that were used for the subsequent TMRCA calculations; ‘CRF01_AE’ (which included all the Vietnamese samples plus the 190 CRF01_AE reference sequences), ‘Vietnam large clade’ (which contained 60/63 Vietnamese strains in this study), and ‘Vietnam small clade’ (which contained the three Vietnamese samples that clustered separately from the others), see Fig. 1.

‘Vietnam large clade’ clustered together with other samples originating from Vietnam, plus a small number of samples from China and the Czech Republic, but not with CRF01_AE strains from Thailand. A sequence identity analysis revealed that the nucleotide consensus sequence of these 60 Vietnamese strains compared with the 84 included Thai strains were only 90% identical. In contrast, the remaining three Vietnamese strains formed a separate clade together with one other sample from Hai Phong in North Vietnam and one strain from Japan. These strains were more closely related to CRF01_AE strains from Thailand, with only three divergent amino acids in the consensus sequences: PR16 (G — E), PR63 (I — C) and RT43 (E — K).

The TMRCA of ‘CRF01_AE’, ‘Vietnam large clade’ and ‘Vietnam small clade’ were calculated using the log-file analyses of three independent runs for the two selected evolutionary models ‘in_log’ and ‘ln_exp’ (Fig. 2). The six TMRCA estimates of the CRF01_AE clade based on 190 samples were calculated to 1975.7–1977.2, median 1976.4. This was very close to the previous estimate in 1975.5 (Abecasis et al., 2009), which was also used as a prior for this study. The median estimates of the tMRCA for the Vietnamese large clade ranged from 1988.5 to 1990.7 with an overall median in the year 1989.8, while the tMRCA estimates of the small clade was in the range 1996.3–1998.1, median 1997.5.

Previous studies of the CRF01_AE epidemiology in Vietnam have shown that HIV was first introduced in the southern part of the country and by 1993 over 950 infections had been diagnosed in Vietnam, of which only three cases were found in the north (Hien and Wolffers, 1994). The introduction of HIV-1 CRF01_AE in Vietnam has been estimated to have occurred at least a decade prior to the first detections of clinical cases and by the late 1980s the disease is believed to have been spreading among IDUs in South Vietnam and thereafter to IDUs in the northern part of the country around 1993–1994 (Liao et al., 2009). Our results date the tMRCA of the clade currently spreading through sexual and intravenous transmission in North Vietnam a few years prior to this, around 1990. This clade, ‘Vietnam large clade’ includes samples from Ha Long, Uong Bi, Dong Trieu and Yen Hung from the current study (N = 60), as well as sequences from Hai Phong (Ishizaki et al., 2009), Bac Giang and Hai Duong (Liao et al., 2009), also located in the coastal North-Eastern part of Vietnam (N = 22), plus a number of intermixed strains from China and the Czech Republic (N = 13). The tMRCA for the North Vietnam cluster calculated by Liao et al. (2010) was based on a smaller number of samples (8 Vietnamese + 2 Chinese samples), which explains the discrepancy between these studies. Indeed, six of these strains were included in the current study and the tMRCA of these strains fell around 1993–1994 (Fig. 1. Vietnamese strains sampled 1998). It is therefore likely that larger sampling rather than methodological differences accounts for the different time estimates, and that HIV first spread to Northern Vietnam around 1990 or earlier.

The ‘Vietnam small clade’ has an estimated tMRCA around 1997, but since the number of strains is small it is difficult to say if they represent an emerging cluster in the north or if the three
Fig. 1. Phylogenetic trees showing the nodes used for tMRCA calculations and the intermixture of strains from intravenous drug users and sexually infected patients in Northern Vietnam. The small inset tree shows all 257 strains with the Vietnam large and small clades encircled. In the larger tree some clades have been collapsed for clarity. The branch length corresponds to the year of sampling. Node markings: red circles, posterior probability > 0.99; blue circles, posterior probability > 0.90. Tip markings: filled circles, intravenous drug users; open circles, sexually infected patients. No tip marking, unknown mode of transmission.

infections were unrelated. BLAST searches confirmed that these strains were more similar to samples from southern Vietnam (Ho Chi Minh City, An Giang) and Thailand than to North Vietnamese and Chinese CRF01_AE strains. One of these samples originated from a truck driver, who had travelled widely throughout Vietnam in his job, and the other two samples came from women who were/had been married to drivers. It is therefore possible that these strains were independently introduced from the southern part of the country. None of these genetically divergent strains carried TDRMs.

The Vietnamese samples analyzed in this study originated from four clinics in the Quang Ninh province in Northeastern Vietnam, near the border to China. These clinics are all located within a radius of approximately 35 km, and no local clustering was found for the respective sites. Twenty-nine samples originated from patients with a history of intravenous drug use, 27 individuals were infected through sexual transmission and the mode of transmission for the remaining seven patients was unknown. Samples from patients with different modes of infection were completely intermixed in the phylogeny (Fig. 1), indicating that HIV-transmission frequently occurs between intravenous drug users and non-drug users in northern Vietnam.

One important limitation of this study is that it may not reflect the most recent trends in transmission as the majority of the patients (84%) had advanced immunodeficiency (CD4 < 200) and had most likely been infected for several years at the time of sampling. Consistent with this, the phylogenetic tree (Fig. 1) revealed that the tMRCA of the most closely related sequences often occurred around 7–12 years ago. Thus, late testing appears to be a major problem in Vietnam and it is likely that important transmission networks still remain undetected.

Acknowledgements

We would like to thank the following contributors for financial support: EU FP7 Marie Curie ERG (GlobalART #246599), EU FP7 CHAIN (#223131), the Swedish Research Council (K2010-56X-13414-11-3), the Swedish International Development Cooperation Agency (SWE-2009-151), the Swedish Civil Contingencies Agency.

References


