Infection, Genetics and Evolution 12 (2012) 448–452

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

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 19 May 2011

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.

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-naïve 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 overall

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

☆ This work was the result of a collaborative effort fueled by the International Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology (http://www.rega.kuleuven.be/cev/workshop/).

* 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.
doi:10.1016/j.meegid.2011.04.034

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), and140 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 HXB2: 2135–3338) was amplified using the primers JA204F-AE (5′-CTCAGAGCAGACCAGAGCCAACAG-3′) and JA205R-AE (5′-TTTTCCCACTAATTTCTGTATATC-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-AE and JA205R-AE and plus an additional primer, SeqR-AE (5′-TACATACAAGTCATCCATGTATTG-3′).

2.3. Baseline resistance

Sixty-three *pol*-gene sequences obtained from ART-naïve 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 $\geq$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 = sequenceInput, 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 63 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 InfCare 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 v.1.6.1 (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 v.1.3.1 (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