Infection, Genetics and Evolution 12 (2012) 291–298

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Reconstruction of the epidemic history of hepatitis B virus genotype D in Albania

Gianguglielmo Zehender [a,*], Renata Shkjezi [a], Erika Ebranati [a], Elena Gabanelli [a], Zhani Abazaj [c], Elisabetta Tanzi [d], Dhimiter Kraja [e], Silva Bino [b], Massimo Ciccozzi [f], Massimo Galli [a]

[a] Department of Clinical Sciences "Luigi Sacco", Chair of Infectious Diseases, University of Milan, Milan, Italy
[b] Control of Infectious Diseases Department, Institute of Public Health, Tirana, Albania
[c] Laboratory of Blood Control for Infectious Agents, Blood Transfusion National Center, Tirana, Albania
[d] Department of Public Health, Microbiology, Virology, University of Milan, Milan, Italy
[e] Infectious Diseases Department, University Hospital Center, "Mother Tereza", Tirana, Albania
[f] Epidemiology Unit, Department of Infectious Parasitic and Immunomediated Diseases, Istituto Superiore di Sanità, Rome, Italy

ARTICLE INFO

Article history:
Received 12 July 2011
Received in revised form 17 November 2011
Accepted 19 November 2011
Available online 28 November 2011

Keywords:
Hepatitis B virus
Coalescent-based analysis
HBV subgenotype D2
Albania
Epidemiology
Public health

ABSTRACT

Despite a recent decrease in the prevalence of HBsAg in the general population, Albania is still highly endemic for HBV infection.

Genotype D is the most prevalent HBV strain in the Mediterranean area. We studied the prevalence and distribution of HBV genotypes and subgenotypes in a total of 73 HBsAg-positive patients living in Albania, and reconstructed the epidemiological history of the most prevalent HBV D subgenotype using a "phylodynamic" framework.

A time-scaled genealogy of the Albanian patients' and reference P gene sequences with known sampling dates was reconstructed using an MCMC Bayesian approach that allows population growth to be estimated on the basis of coalescent theory.

All of the Albanian subjects were infected with the HBV D genotype, and a percentage varying from 44.4% to 100% (depending on the ethnic or risk group) were infected with subgenotype D2, the most prevalent in the study population (72.4%). The other subgenotypes present in a minority of subjects were D1 (13.8%) and D3 (13.8%).

The Bayesian skyline plot population dynamics analysis showed that genotype D2 entered the Albanian population in the late 1960s, and that the effective number of infections grew gradually until the second half of the 1980s and more rapidly until the mid-1990s, when it reached a plateau that still persists today.

Our data suggest that political and socio-economic factors played an important role in determining the rapid spread of HBV infection in Albania.

1. Introduction

HBV is an enveloped DNA virus belonging to the *Hepadnaviridae* family. It has a small circularised genome of about 3.2 kilobases that encompasses four partially overlapping open reading frames (S, C, P and X) coding for seven proteins, which are characterised by considerable variability because of the use of a reverse transcriptase during replication.

HBV is ubiquitous and infects more than 360 million people throughout the world. Between 7000 and 8000 new diagnoses of hepatitis B are made every year in Europe, although there is a general trend towards a decrease (European Centre for Disease Prevention and Control, 2010). The Mediterranean basin includes countries that mainly have intermediate prevalence rates (about 1–2% of carriers in Spain, Greece, and parts of Italy) (European Centre for Disease Prevention and Control, 2010). Although there has been a considerable decrease in HBV infections over the last few years (probably due to the introduction of mass immunisation in 1995), HBV is still highly endemic in Albania, where it is the most important cause of chronic liver disease (Durro et al., 2010; Katsanos et al., 2009; Kondili et al., 2009; Resuli et al., 2009).

So far, eight main viral genotypes (named alphabetically from A to H) and various subgenotypes (indicated by numbers) with different ethno-geographic distributions have been described. Genotypes A and D are the most prevalent in Europe: genotype A mainly in the northern part of the continent, and genotype D in

Abbreviations: HBsAg, hepatitis B surface antigen; MCMC, Markov Chain Monte Carlo; HCWs, health care workers; IVDUs, intravenous drug users; MSMs, males having sex with males; P, HBV polymerase; PreS/S, HBV surface proteins; GTR+G, general time reversible with gamma distribution; BF, Bayes factor; MCC, maximum clade credibility; tMRCA, time of the most recent common ancestor; BSP, Bayesian skyline plot; Ne, effective number; R_0, basic reproductive number; 95% HPD, 95% highest posterior density.

* Corresponding author. Address: Dipartimento di Scienze Cliniche "L. Sacco", Sezione Malattie Infettive e Tropicali, Università di Milano, c/o Ospedale L. Sacco, Via G.B. Grassi 74, 20157 Milano, Italy. Tel.: +39 02 50319770; fax: +39 02 50319768.
E-mail address: gianguglielmo.zehender@unimi.it (G. Zehender).

1567-1348/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2011.11.009

eastern Europe and Mediterranean countries, including Albania (Kondili et al., 2005; Schaefer, 2007). The distribution of D subgenotypes in Europe and the Mediterranean basin is still unclear because only a few studies have been carried out (Schaefer, 2007; Zehender et al., 2008), and there are no data concerning their distribution in Albania.

The aims of this study were to describe the molecular epidemiology of HBV in Albania by characterising the viral genotypes and subgenotypes infecting HBsAg-positive inhabitants and their distribution by population groups, and to investigate its phylodynamics on the basis of a coalescent framework that has recently been used to reconstruct the epidemiological history of various infections including HBV (Drummond et al., 2005; Zehender et al., 2008; Zhou and Holmes, 2007).

2. Patients and methods

2.1. Patients and datasets

A total of 73 HBsAg-positive serum samples collected between 2005 and 2007 from patients with chronic HBV infection living in Albania (57 males [78%] and 16 females [22%] with a median age of 34.5 years, range 16–58) were analysed for HBV DNA. The samples were collected at the Institute of Public Health, Tirana during a serosurveillance of HBV markers in a group of subjects at high and low risk of infection, and at the National Blood Transfusion Centre, Tirana. All of the patients were anti-HCV negative. The possible route of exposure to HBV infection was known in 32 cases: nine healthcare workers (HCWs), 13 intravenous drug users (IVDUs) and 10 males having sex with males (MSMs). In all of the other cases, the route of transmission was unknown. Nine subjects belonged to the Roma ethnic group. The group of MSMs included two patients with dual HIV-1/HBV infection.

2.2. Methods

2.2.1. HBV DNA extraction and amplification by means of nested PCR

Viral DNA was extracted from the patients' serum (300 μl stored at −40 °C) using the High Pure PCR Template Preparation Kit (Roche Diagnostics GmbH, Mannheim, Germany) in accordance with the manufacturer's instructions. The DNA was eluted in 50 μl of nuclease-free distilled water, and a sequence of about 1348 nucleotides was amplified using a previously described multiple nested PCR protocol in order to obtain the HBV polymerase (P) and surface antigen (PreS/S) sequences (De Maddalena et al., 2007). The outer degenerated primers used for the first fragment (911 pb:nct 2960–655) were HBV-7 [5′-A(AG)T CC(AC) GAT TGG GAC (CT)(AT)C AA-3′] and ERS-IN3 [5′-TGA GGC CCA CTC CCA TAG G-3′], and those used for the second fragment (907 pb:nct 202–1108) were HBV-9 [5′-GGC GG(GT) GT(GT) TTT CTT GTT G-3′] and HBV-6 [5′-AAG TTG GCG A(AG)A A(AG)(AG) (CT)(AG)A AAG CCT-3′]. Subsequently, in order to amplify three inner overlapping fragments, we used the internal primers HBV-7 [5′-A(AG)T CC(AC) GAT TGG GAC (CT)(AT)C AA-3′] and HBV-4 [5′-TAG AAA ATT GAG AGA AGT CCA CCA-3′] for the first fragment (536 pb:nct 2960–280), HBV-5 [5′-TAG GAC CCC T(GT)C TCG TGT TAC AGG-3′] and ERS-IN3 [5′-TGA GGC CCA CTC CCA TAG G-3′] for the second fragment (476 pb:nct 180–655), and ERS-IN2 [5′-CAT CCT GCT GCT ATG CCT CA-3′] and HBV-6 [5′-AAG TTG GCG A(AG)A A(AG)(AG) (CT)(AG)A AAG CCT-3′] for the third fragment (700 pb:nct 409–1108). The nucleotide numeration refers to isolate NDR260 (Okamoto et al., 1988). About 10 μl of the extracted DNA were amplified in a 50 μl mixture containing 10 μl of 5× Colorless Flexi Buffer, 2 mmol/L $MgCl_2$, 0.2 mmol/L dNTPs, 0.5 μmol/L of the outer primers, and 1 U of GoTaq Polymerase (GoTaq® Hot Start Polymerase, Promega, Madison, USA). The PCR cycling conditions consisted of 9 min at 95 °C, 45 cycles of 95 °C for 45″, 57.5 °C for 45″ and 72 °C for 1′30″, a prolonged extension step at 72 °C for 10 min, and finally holding at 4 °C.

Second round PCR was performed using 6 μl of the first round PCR product in 50 μl of the same mixture, but with 10 μl of 5× Green Flexi Buffer and 0.5 μmol/L of the inner primers. In this round, two different amplification conditions were used: 95 °C for 9′, 45 cycles of 95 °C for 40″, 55 °C for 40″, 72 °C for 1′20″, followed by 72 °C for 10′ and holding at 4 °C for the first fragment; and 95 °C for 9′, 45 cycles of 95 °C for 45″, 59 °C for 45″, 72 °C for 1′30″, followed by 72 °C for 10′ and holding at 4 °C for the second and third fragments.

All of the PCRs were performed using an automated thermocycler (Applied Biosystems, Foster City, CA, USA), and each amplified DNA sample (10 μl) was separated by means of electrophoresis in a 1.5% agarose gel stained with ethidium bromide, and visualised under UV light. Optimal positive and negative controls were included in the extraction and amplification processes.

2.2.2. HBV DNA sequencing

The three DNA fragments obtained by means of PCR were purified using a commercial purification kit (microCLEAN, Labogen, Catania, Italy), and then sequenced bidirectionally using a BigDye Terminator Kit version 3.1 (Applied Biosystems).

An optimal quantity of purified PCR product was mixed with 3.2 pmol of the sequencing primer and 1.1 μl of BigDye, in a final volume of 10 μl. The cycle sequencing reactions were performed for 25 cycles, each of 20 s at 95 °C, 20 s at 55 °C, and 3 min at 60 °C.

The sequencing products were purified by precipitating 10 μl in an ethanol/sodium acetate mixture. Finally, after suspending the heat-dried samples in 10 μl of nuclease-free distilled water, 2.2 μl were loaded into an automated DNA sequencer ABI PRISM® 3100 Genetic Analyser (Applied Biosystems). The sequences were submitted to GenBank (Accession Nos. JQ244777–JQ244834).

2.2.3. Phylogenetic analysis

2.2.3.1. Genotype characterisation. All of the Albanian P gene sequences were aligned with 61 genotype/subgenotype-specific reference sequences downloaded from GenBank, and cropped to a length of 1182 nucleotides (http://www.ncbi.nlm.nih.gov/Genbank) using ClustalW (Thompson et al., 1994) included in Bioedit (created by Tom Hall, 1999), followed by manual editing. The hierarchical likelihood ratio test (LRT) method implemented in ModelTest version 3.07 (Posada and Crandall, 1998) was used to select the evolutionary model that best fitted the sequence data. A general time reversible model with gamma distributed rate heterogeneity (GTR+G) (Lanave et al., 1984) was selected. The phylogeny of the P gene sequences was reconstructed using a distance-based neighbour-joining method (NJ) implemented in version 4 of the MEGA program (Kumar et al., 2004).

The reliability of the observed clades was established on the basis of an internal node bootstrap support value of more than 0.60 (after 1000 replicates).

2.2.3.2. HBV genotype D dataset. After removing the isolates containing long deletions/insertions and recombinant strains, the remaining Albanian genotype D sequences were aligned with 94 Italian P gene sequences of the same genotype sampled in known years between 1980 and 2007, and included in a previously published study (Zehender et al., 2008) (available at GenBank under Accession No.: EF514251–EF514344). All of these sequences matched the following criteria: (1) they had been previously published in peer-reviewed journals (except for those characterised in this study); (2) there was no uncertainty about the subtype assignment of each sequence, and they were classified as non-recombinant; (3) the sequences