Phylogenetic analysis and subgenotypic distribution of the hepatitis B virus in Recife, Brazil

Izolda Fernandes Moura a, Edmundo Pessoa Lopes b,*, Mónica Viviana Alvarado-Mora c, João Renato Pinho c, Flair José Carrilho c

a Central Laboratory of the Hospital das Clínicas, Federal University of Pernambuco, Brazil
b Division of Gastroenterology, Department of Internal Medicine, Federal University of Pernambuco, Brazil
c Laboratory of Gastroenterology and Hepatology, São Paulo Institute of Tropical Medicine and Department of Gastroenterology, School of Medicine, University of São Paulo, Brazil

Abstract

The analysis of the genomes of the hepatitis B virus in human hosts identifies phylogenetic variants called viral genotypes. Indeed, clinical and epidemiological observations suggest that differences in viral genotypes lead to distinct biological and clinical behaviors. The aim of this study was to evaluate the subgenotypic distribution and to conduct a phylogenetic analysis by Bayesian method of the hepatitis B virus (HBV) in patients from Recife, Brazil. From July 2009 to December 2010, 60 HBV infected patients were examined, 39 (65%) were males, whose mean age was 50 years old. 33 (55%) were genotyped by obtaining and amplifying a 1306 bp fragment comprising part of the DNA polymerase and the surface antigen of the HBV. The sequencing was performed on an ABI 3500 Automatic Sequencer and the consensus sequences were obtained by aligning both the sequenced strands (clockwise and anti-clockwise) using SEQUENCHER software. Phylogenetic analysis was conducted using the Markov Chain Monte Carlo simulation implemented by Bayesian evolutionary method by sampling trees. The following subgenotypic distribution was observed: A1 (79%), F2a (12%), A2 (6%) and F4 (3%) as was that those identified as subgenotype A1 were in the same cluster in the phylogenetic tree. In this study, the majority of patients presented the A1 subgenotype from the same viral strain. As per the distribution in the phylogenetic tree by Bayesian method, possibly this subgenotype was in the genetic make-up of Africans brought in centuries past to Brazil as slaves.

Keywords: HBV genotype, HBV subgenotype, Phylogenetic analysis, Bayesian analysis, Hepatitis B virus

1. Introduction

The analysis of the genomes of hepatitis viruses in the human population identifies phylogenetic variants called viral genotypes. The genotype is characterized by nucleotide differences of 8% or more over the entire length of the genome of the hepatitis B virus (HBV) or 4% or more within the S gene of the viral surface (HBsAg) antigen (Kim et al., 2011; Tanwar and Dusheiko, 2012). In fact, biological and epidemiological observations suggest that differences in HBV genotypes induce distinct clinical and therapeutic behaviors (Raimondi et al., 2010). For example, patients infected with genotype A seem to respond better to interferon therapy than those infected with genotype D, while genotype B infected patients respond better than those infected with genotype C (Cao, 2009; Mello et al., 2012).

HBV is classified into nine genotypes from A through I, which are subdivided into subgenotypes that can be characterized by a few amino acid substitutions in the determining “a” of gene S and named with digits A1 ... A5, F1 ... F4. These genotypes have distinct geographical distributions (Scheiblauer et al., 2010; Yu et al., 2010).

The genetic diversity of HBV and the geographical distribution of its subgenotypes are tools for reconstructing its evolutionary history, but this also can assist in coming to understand the evolution and migrations of humans in the past (Cao, 2009; Scheiblauer et al., 2010). It is worth mentioning that the flow of global migration can determine changes in the geographical distribution of HBV genotypes or subgenotypes, with the presence of some genotypes in areas where they had not previously circulated. Indeed, countries in South America have shown emigration flows to European countries such as Italy, which have led to the distribution of HBV genotypes, such as A and D (Palumbo et al., 2007).

In Brazil, studies on identifying HBV genotypes and subgenotypes show that variability occurs in the geographical distribution of the virus. As result of individuals coming from Europe and Africa to our country over many decades, the population became of mixed races and started to show a different genotypic circulation pattern to that found in other Latin American countries.
The presence of genotypes A and D suggest influences from African descendants as a result of the period of slavery and European colonization, respectively (Devesa et al., 2008; Alcalde et al., 2009; Tonetto et al., 2009). Nonetheless, in the native populations (indigenous) a higher prevalence of genotype F is observed (Mello et al., 2007; Dias et al., 2012).

Fig. 1 shows subgenotype distribution of HBV in some Brazilian states. It is noted that it has only been presented the articles that the sequencing of the subgenotype has been made (Andermach et al., 2009; Matos et al., 2009; Alvarado-Mora et al., 2011a; Santos et al., 2011; Bertolini et al., 2012).

Due to the scarcity of data on the distribution of subtypes of HBV in our region, this study set out to evaluate the subgenotypic distribution and to construct the phylogenetic tree and to conduct the Bayesian analysis of this virus.

2. Patients and methods

2.1. Patients

This was a cross sectional study to assess the genotypic and subgenotypic distribution of HBV in the outpatient’s clinic for Hepatology at the Hospital das Clínicas of Federal University of Pernambuco between July 2009 and December 2010.

Patients were included consecutively provided they were above 18 years old, and had been HBsAg and anti-HBc positive for more than 6 months. The exclusion criteria were patients with positive serology for the hepatitis C virus (HCV) and immune-compromised patients.

The project was submitted to and approved by the Committee for Ethics in Research, Center for Health Sciences, UFPE and all patients signed an informed term of consent.

2.2. Methods

Blood samples were collected and taken to the Central Laboratory of the Hospital das Clínicas, where they were centrifuged and frozen at −20 °C.

The frozen samples were then taken to the Laboratory of Gastroenterology and Hepatology of Institute of Tropical Medicine, University of São Paulo, where HBV genotyping and subgenotyping were conducted as was Bayesian and phylogenetic analysis.

HBV-DNA was extracted by withdrawing 100 μL of serum from each sample using the acid guanidinium thiocyanate–phenol–chloroform method (Chomczynski and Sacchi, 1987; Kwok and Higuchi, 1989). A 1306 bp fragment comprising partially HBsAg and the DNA polymerase (S/Pol) coding regions was amplified by the nested PCR method using the primers PS3132F/2920R and PS3201F/P1285R (Alvarado-Mora et al., 2011b).

The amplified DNA was purified using the ChargeSwitch PCR Clean-Up Kit (Invitrogen). Sequencing was performed on an ABI 3500 Automatic Sequencer (Applied Biosystems, Foster City, CA) using dideoxynucleoside triphosphates (ddNTPs) containing fluorescent markers (Big Dye1 Terminator v3.1 Cycle Sequencing Ready Reaction Kit, Applied Biosystems).

The consensus sequences were obtained by aligning both sequenced strands (clockwise and anti-clockwise) using SEQUENCHER software (Gene Codes Corporation Ann Arbor, Michigan, United States of America).

The sequences were genotyped by means of phylogenetic reconstruction using the reference sequences of each genotype obtained from the GenBank (n = 267). The sequences were aligned using Muscle software and edited with SE-AL software. Phylogenetic analysis was conducted using the Markov Chain Monte Carlo (MCMC) simulation implemented in BEAST (Bayesian evolutionary analysis by sampling trees) v.1.5.3. Chains were run for 10 million
generations which were sufficient to achieve the convergence of parameters. The maximum clade credibility (MCC) tree was obtained which summarized the 10,000 substitution trees and then 10% of the burn-in was removed using Tree Annotator v.1.5.3 (Drummond and Rambaut, 2007).

The sequences obtained were analyzed and grouped in the same “cluster”. Then phylogenetics analysis was performed using BEAST v.1.5.4. A Bayesian skyline plot was drawn using the best model of nucleotide substitution (GTR + I + G), previously obtained in the Modeltest (Posada and Crandall, 1998). Ten million generations were sufficient to converge all parameters (effective sampling size > 200). The molecular clock with the best results was chosen by Bayes Factor comparison.

3. Results

We investigated 60 HBV infected patients, whose mean age was 50 years old. 39 (65%) were male and 21 (35%) were female. In 27 patients (45%) it was not possible to identify subgenotypes, and therefore, 33 HBV infected patients were analyzed. The samples from these patients revealed the following distribution: 28 A genotype patients (85%), there being 26 A1 subgenotype (79%) and 2 A2 (6%), and 5 F genotype patients (15%) of whom 4 were F2a subgenotype (12%) and 1 F4 (3%).

According to the Bayesian analysis, of the total number of sequences obtained, 24 of the A1 subgenotype were grouped in the same “cluster” with great support. The subgenotypes identified in this study are placed in the phylogenetic tree (Fig. 2).

4. Discussion

How the colonization of Brazil took place greatly influenced the distribution of HBV genotypes in each region of the country. In fact, a study involving 49 patients on hemodialysis in the state of Santa Catarina revealed that genotype D (57%) was the one that occurred most frequently, followed by the A (31%) and F (12%) genotypes (Carrilho et al., 2004). Another study of 67 patients from the state of Rio Grande do Sul also found that genotype D (60%) was the most frequent, followed by the A (34%) and F (5%) genotypes (Becker et al., 2010). Furthermore, a recently published study revealed that the most common genotype in state of Paraná was D (83%) followed by genotype A (14%) (Bertolini et al., 2012).

This finding should be related to the population of the South of Brazil being of predominantly European ancestry, mainly Italians and Germans, which would explain why genotype D is the one that has the highest occurrence.

On the other hand, in the Southeast region of Brazil was observed a predominance of HBV genotype A followed by D. In addition, the genotypes C and F were also found, suggesting that different HBV strain were introduced in this region (Alcalde et al., 2009; Tonetto et al., 2009).

In the North region of Brazil, in Western Brazilian Amazon at Lábrea District, a study evaluating 86 serum samples HBsAg reactive which whom 39 were found to be HBV-DNA, obtaining the following genotypic distribution A (60%), D (35%) and F (5%) (Dias et al., 2012). Nevertheless, an evaluation of 40 HBV infected patients in the Rondônia state found genotype D in 43%, A in 37% and F in 20% of the patients. The subgenotypic analysis showed the following distribution: A1 (37%), D3 (23%), F2a (20%), D4 (17%) and D2 (3%) (Santos et al., 2010). It is noteworthy that the HBV found in this study, according to phylogenetic analysis, did not come from a common strain, which suggests that different subgenotypic inputs occurred in this region of the Amazon. Probably, the genotype D was introduced there by people from South region of Brazil that immigrates during the 1970s to 1980s decades. The F2a subgenotype was identified, even if its prevalence was low, as in our study, which is probably related to the influence of native South Americans. This genotype was found primarily in the indigenous populations of South America, especially Venezuela, Colombia and Brazil (Mello et al., 2007; Devesa et al., 2008; Alvarado-Mora et al., 2011b).

In the Midwest region, a study involving 26 hemodialysis patients from the state of Goiás verified the occurrence of genotype A in half the cases, followed by D (46%) and F (4%) (Teles et al., 1999). Two other studies conducted in this region on individuals from communities with African descendants (quilombos), as expected, showed a great predominance of the A genotype and a low percentage of genotype D (Motta-Castro et al., 2008; Matos et al., 2009).

In the Northeast, an evaluation conducted in a rural community of the semi-arid region in Bahia, involving 1476 individuals, found that 38 (2.6%) had chronic HBV infection and of these, 9 subjects (23.6%) presented positive HBV-DNA and were identified as A genotype (Almeida et al., 2006). Also in Bahia state, at Salvador city, were evaluated 61 HBV infected patients and revealed genotype distribution similar to our results. They found the following: genotype A (87%), D (11%) and F (2%), but the virus subgenotypes were not evaluated (Ribeiro et al., 2006). Another study analyzing subgenotyping and phylogenetic analysis in this region evaluated 72 individuals of African descent, residents of a quilombo in the state of Maranhão, and revealed that 4 (5.6%) were positive for HBsAg and all of them were identified as A1 subgenotype (Alvarado-Mora et al., 2011).

The present study which evaluated patients from a large city in the Northeast region of Brazil revealed that the A1 subgenotype was the most prevalent (79%). The subgenotypic placement in the phylogenetic tree indicated that most of them came from a common single strain from Africa (Fig. 2). It is noteworthy that no patients were found with genotype D, despite the influence of Portuguese and Dutch colonization in our region. The absence of genotype D could be justified because the patients included were from a public hospital and their socio-economic level was low. In other words, it may well be that the genetic make-up of these patients have minimal or no European elements in them. As the population evaluated had low incomes, the transmission could have occurred in the past by sexual or vertical pathway. Differently, a previous study in our city involving 24 blood donors revealed 4 individuals (17%) with genotype D, 54% and 29% with genotypes A and F, respectively (Mello et al., 2007). In our region, usually blood donors are of urban origin and have higher socio-economic level. Nor did our study find patients with genotypes B and C which have been found in Brazil (Tonetto et al., 2009; Bertolini et al., 2012), probably due to the immigration of people from the Far East to the Northeast being small.

All these data indicate that the HBV A genotype is widely prevalent in Brazil, except in the South region. As this genotype had an African origin, its introduction into Brazil could have occurred during the slave trade between the 16th and 19th centuries, mainly in the Northeast region where slaves were sent to work in the sugar cane fields. After the arrival in this region, some slaves ran away to other regions of the country especially North and Midwest where they raised the quilombos.

Similarly, in Haiti, where the population is predominantly of African origin, the A genotype was found to be the most prevalent, above all A1 and A5, followed by genotypes D and E (Kramvis and Kew, 2007; Andermach et al., 2009).

Current data unexpectedly show that the most common genotype in Angola is E and that the region which that country then covered was the main source of the slaves brought by the Portuguese to the Northeast of Brazil (Valente et al., 2010).
Fig. 2. Phylogenetic tree with the hepatitis B virus subgenotypes distribution of patients in Recife (2011).
As the A genotype has five subgenotypic differences (A1–A5), studies on evolution suggest that it has a long natural history in the African continent. As to the E genotype, it has less genetic diversity and must have arisen more recently in Africa (Kimbi et al., 2004; Kramvis and Kew, 2007). Thus, the high occurrence of the A genotype in the African–American descendants would be justified because of its having spread out in past centuries, when it was still endemic in Africa. As the emergence of the E genotype was more recent, it probably spread after the slave trade ended, namely from around the mid-to late 19th century.

5. Conclusion

In conclusion, this study demonstrated that most patients had the A genotype of HBV, the A1 subgenotype being the most common. The phylogenetic tree by Bayesian method indicates that its spread occurred from the same viral strain, suggesting that the introduction of this subgenotype in our region began when African slaves were brought here between the 16th and 19th centuries. Further studies in other cities of the Northeast and other regions of Brazil, however, are necessary to confirm these findings.

References


