

Full-length genomic sequence of hepatitis B virus genotype C2 isolated from a native Brazilian patient

Mónica Viviana Alvarado-Mora^{1/+}, Rúbia Anita Ferraz Santana², Roberta Sitnik², Paulo Roberto Abrão Ferreira³, Cristovão Luís Pitangueira Mangueira², Flair José Carrilho¹, João Renato Rebello Pinho^{1,2}

¹Laboratório de Gastroenterologia e Hepatologia Tropical, Departamento de Gastroenterologia, Instituto de Medicina Tropical, Faculdade de Medicina, Universidade de São Paulo, Av. Dr. Enéas de Carvalho Aguiar 500 prédio IMT2 2º andar, 05403-000 São Paulo, SP, Brasil ²Departamento de Patologia Clínica, Hospital Israelita Albert Einstein, São Paulo, SP, Brasil ³Universidade Federal de São Paulo, São Paulo, SP, Brasil

The hepatitis B virus (HBV) is among the leading causes of chronic hepatitis, cirrhosis and hepatocellular carcinoma. In Brazil, genotype A is the most frequent, followed by genotypes D and F. Genotypes B and C are found in Brazil exclusively among Asian patients and their descendants. The aim of this study was to sequence the entire HBV genome of a Caucasian patient infected with HBV/C2 and to infer the origin of the virus based on sequencing analysis. The sequence of this Brazilian isolate was grouped with four other sequences described in China. The sequence of this patient is the first complete genome of HBV/C2 reported in Brazil.

Key words: hepatitis B virus - genotype C2 - China - Bayesian analyses - complete genome - Brazilian patient

Hepatitis B virus (HBV) infection is a public health problem; approximately two billion people have been exposed to HBV and more than 300 million are chronically infected with this virus (Grimm et al. 2011). HBV is the prototype member of the Hepadnaviridae, a family of hepatotropic DNA viruses. The virus has a 42 nm diameter viral particle and an extremely compact partially double-stranded circular genome (Dane et al. 1970). HBV is classified into eight genotypes (A-H) based on inter-group divergence of the entire genomic nucleotide sequence (Kramvis et al. 2008). Recently, an additional HBV genotype (I) was proposed (Yu et al. 2010).

HBV genotypes and subgenotypes have distinct geographical distributions and are associated with the severity of liver diseases in different populations. Genotype A is distributed globally and is the main genotype found in Europe, North America, Africa and India. Genotype D is mainly found in the Middle East and the Mediterranean Basin. Genotype E was originally reported in Africa (Kramvis & Kew 2007), while genotype G was initially reported in Europe and North America (Stuyver et al. 2000). HBV genotype F is found in Central and South America, particularly in indigenous populations in South America (Devesa & Pujol 2007, Alvarado-Mora et al. 2011). Genotype H is frequent in Central and North America (Arauz-Ruiz et al. 2002) and is very closely related to genotype F (Alvarado-Mora et al. 2011). Recently, HBV genotype I was described in northwestern China, Vietnam and Laos (Yu et al. 2010).

The most frequent HBV genotypes in Asia are B and C. Genotype C is associated with more aggressive liver disease and poorer response to antiviral therapy compared to genotype B (Chu et al. 2002). Genotype C is associated with hepatocellular carcinoma (HCC) in older patients (Yu et al. 2005), while genotype B is associated with HCC in younger people and with acute hepatitis B in adults (Ni et al. 2004).

Genotype C is classified into 10 subgenotypes that differ in their geographical distribution: C1 is found in Thailand, Vietnam and Myanmar (South Asia), C2 is found in Japan and China (Far East Asia), C3 is found in New Caledonia and Polynesia, C4 is found in Australian Aborigines and C5 is confined to the Philippines and Vietnam (Ni et al. 2004, Chan et al. 2005, Kramvis et al. 2008). The subgenotypes C3-C10 are also found in Indonesia (Mulyanto et al. 2009, 2010).

The most common genotype in Brazil is genotype A, followed by genotypes D and F. The North, Northeast and Southeast Regions have a higher frequency of genotype A, while genotype D is the most frequent in the South Region (Mello et al. 2007). In some regions, genotypes B and C are also detected at low prevalences, reflecting the presence of Asian descendants within the populations of these regions (Sitnik et al. 2004). The aim of this study was to determine the complete genome sequence of an HBV subgenotype C2 isolate from a native Brazilian patient and to infer the origin of this virus.

The patient was a 54-year-old Caucasian female who was born in state of São Paulo (SP), Brazil. She was unemployed at the time of the study (but worked in Ota, Gunma, Japan, from 2000-2009) and had been diagnosed with chronic hepatitis B in Japan, at which time antiviral treatment was started. No other comorbidities or relevant data on her clinical history were reported and her physical examination was normal. At that time, she had the following clinical assay results: aspartate

Financial support: FAPESP (2007/53457-7, 2008/50461-6), CNPq, IIRS/HIAE

+ Corresponding author: monica.viviana@usp.br

Received 4 February 2011

Accepted 9 May 2011

transaminase (AST) 73 IU/mL, alanine transaminase (ALT) 93 IU/mL, HBsAg (+), HBeAg (+) and anti-HBc (+). She received lamivudine 150 mg/day from August 2005-October 2008. As HBV DNA was always detected, the treatment was changed to adefovir 10 mg/day. At that point, her clinical assay results were as follows: anti-HCV (-), anti-HIV (-), HBeAg (+), AST 37 IU/mL, ALT 32 IU/mL, HBV DNA 13,900,000 IU/mL (log 6.14) and genotype C; no resistance-related mutations were found in the DNA polymerase coding region. We started lamivudine plus tenofovir in June 2010 and HBV DNA became undetectable in September 2010. All procedures followed were in accordance with the ethical standards.

The sample collected when the patient arrived in Brazil was processed to amplify the complete HBV genome. HBV DNA extraction was conducted from 100 μ L of serum using the acid guanidinium thiocyanate/phenol/chloroform method (Chomczynski & Sacchi 1987). Amplification of the whole HBV genome was performed with the P1 and P2 primers described previously, with slight modifications (Gunther et al. 1995). The sample was submitted to DNA amplification by polymerase chain reaction (PCR) (Gomes-Gouvea 2005).

Sequencing was performed in an ABI Prism[®] 3100 Automatic Sequencer (Applied Biosystems, Foster City, CA, USA) (Sanger et al. 1977) using dideoxy nucleoside triphosphates containing fluorescent markers (Big Dye[®] Terminator v3.1 Cycle Sequencing Ready Reaction kit, Applied Biosystems, Foster City, CA, USA). For sequencing, we used 18 primers internal to the PCR fragment, generating sequences around 300-500 bp in length. The primers were as follows: PS3076F (Stuyver et al. 2000), RHBS2 (Sitnik et al. 2004), P1.2, P2.2 (Gunther et al. 1995), 582R, LAM1F, 2817R, P184, P192, P194, P197, P1193R, P970F, 2029R, X1577F (Gomes-Gouvea 2005), EP2.2. (Takahashi et al. 1995), PS3216R (Norder et al. 1994) and 5LAM5 (Da Silva et al. 2000). The quality of each electropherogram was evaluated using Phred-Phrap software (Ewing et al. 1998) and consensus sequences were obtained by alignment of both sequenced strands using CAP3 software available from the Electropherogram quality analysis web page (asparagin.cenargen.embrapa.br/phph/).

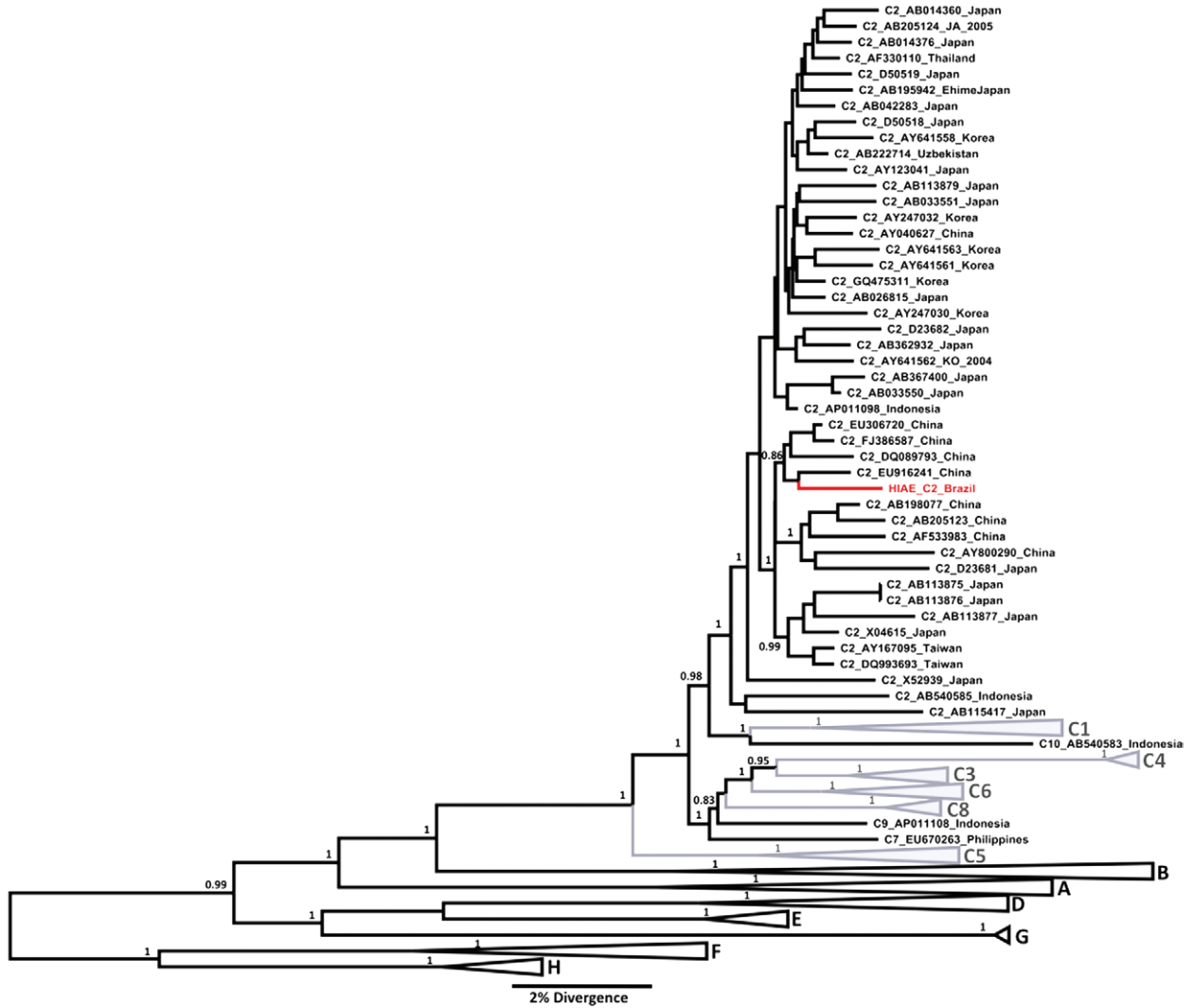
The isolated complete genome was aligned with other previously reported complete genome sequences ($n = 192$) using ClustalX software (Thompson et al. 1997) and was edited with the SE-AL software (available from: tree.bio.ed.ac.uk/software/seal/). The Bayesian Markov chain Monte Carlo simulation implemented in BEAST v.1.4.8 (Drummond & Rambaut 2007) was applied to obtain the best possible estimates using both the relaxed uncorrelated log_{normal} and exponential molecular clock models and using the model of nucleotide substitution (general time-reversible + gamma + proportion invariant). The molecular clock that best fits the data was chosen by Bayes factor comparison. After 10 million generations, the maximum clade credibility (MCC) tree was obtained by summarising 10,000 substitution trees and was then modified using a 10% burn-in using Tree Annotator v.1.5.3 (Drummond & Rambaut 2007). Phylogenetic trees were visualised and midpoint rooted in FigTree v1.2.2 (tree.bio.ed.ac.uk/software/figtree/).

The complete HBV genome was successfully amplified. The sequence was analysed for all reported mutations of antiviral resistance and it did not show any mutation-related response. In addition, the patient harboured a virus with a precore mutation at position 1896 (G-A), but no mutation at position 1899; this patient's virus also lacked basal core promoter mutations. Once the MCC tree was obtained, the Brazilian isolate sequence grouped within genotype C, subgenotype HBV/C2, along with four other sequences from China (Figure). This cluster was supported by an *a posteriori* probability of approximately 0.83. Sequence was deposited in GenBank under the accession HQ622095.

In Brazil, there is a highly miscegenated population and the HBV genotype distribution pattern in Brazil is different from that of other Latin American countries, with genotypes A, D and F being the most prevalent among HBV carriers in Brazil (Devesa & Pujol 2007, Mello et al. 2007, Santos et al. 2010). In the Southeast and South Regions, other HBV genotypes have been found that reflect the migrant origin of the population. In a previous study performed with chronic HBV carriers from different Brazilian regions, it was reported that genotype C was found in 13.2% of the patients, all of whom had an Asian background (Sitnik et al. 2004). In another study in Campinas, SP, the presence of the genotype C (3%) was reported in two Asian descendants and also in two Caucasians (Tonetto et al. 2009). In the Brazilian Amazon Region, one case of genotype C was reported among 44 patients (de Oliveira et al. 2008). Finally, in another study carried out in the states of Rio de Janeiro and Mato Grosso, Brazil, one case of genotype C was found (Bottecchia et al. 2008). In conclusion, the presence of genotype C in the Brazilian population is very low and most cases are restricted to the Asian population living in the country or to those individuals with some contact with Asian people. In addition, because only a short HBV genome region was amplified in each of these previous studies, it is not possible to infer the specific origin of each sequence.

HBV infection is highly prevalent in China. A previous study verified that genotypes B and C are the predominant genotypes in China (Zhu & Dong 2009). The two major subgenotypes of genotype C HBV are subgenotype C1 (predominant in Southeast Asia) and C2 (predominant in East Asia). In an interesting study, it was reported that in China the frequency of HBV genotype B tends to decrease gradually and that the prevalence of genotype C increases with increasing latitude (Zhu & Dong 2009). Furthermore, genotype C prevails in the northern provinces of China, while genotype B is more prevalent in southern China (Zeng et al. 2005). Likewise, there are significant differences in geographical distribution among the patients with genotypes B and C in Japan (Orito et al. 2001).

The complete HBV genome reported in this study grouped with other sequences from China. Based on the demographic and clinical information for this patient, we inferred that she may have acquired the infection during the period that she lived in Japan (9 years), although the origin of this sequence does not appear to be Japan. The



The maximum clade credibility tree was estimated by Bayesian analysis of 192 complete genomes of hepatitis B virus (HBV) strains. The posterior probabilities (> 0.80) of the key nodes are depicted above the respective nodes. The HBV/C2 isolate from the Brazilian patient is represented in red and was analyzed together other worldwide strains. The collapsed clades correspond to the other genotypes of HBV.

patient did not report any other information about the possible mode of transmission that could have aided us in making more inferences about this case. Phylogenetic analysis showed that within the group of genotype C2, the sequences previously reported from different countries (China, Japan, Korea, Taiwan, Thailand, Uzbekistan) are interspersed. The sequence from this Brazilian case grouped with four other sequences described in China, raising the possibility that someone who had direct or indirect contact with people living in China might have infected our patient. It was not possible to determine the specific region of China from which this sequence could have originated because there is no specific publication about these sequences and we could not determine from which county these sequences were collected. The sequence of our patient is the first complete genome sequence of HBV subgenotype C2 reported in Brazil and it is noteworthy that this patient was not of Asian descent.

In conclusion, our results increase the understanding of the dynamics of HBV and demonstrate the application of the Bayesian inference to determine the origin of an infection. We hypothesise that the native Brazilian patient was infected with an HBV/C2 strain that is prevalent in China when she lived in Japan.

REFERENCES

Alvarado-Mora MV, Romano CM, Gomes-Gouvea MS, Gutierrez MF, Botelho L, Carrilho FJ, Pinho JR 2011. Molecular characterization of the hepatitis B virus genotypes in Colombia: a Bayesian inference on the genotype F. *Infect Genet Evol* 11: 103-108.

Arauz-Ruiz P, Norder H, Robertson BH, Magnus LO 2002. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *J Gen Virol* 83: 2059-2073.

Bottechia M, Souto FJ, O KM, Amendola M, Brandão CE, Niel C, Gomes SA 2008. Hepatitis B virus genotypes and resistance mutations in patients under long term lamivudine therapy: characterization of genotype G in Brazil. *J Gen Virol* 8: 11.

- Chan HL, Tsui SK, Tse CH, Ng EY, Au TC, Yuen L, Bartholomeusz A, Leung KS, Lee KH, Locarnini S, Sung JJ 2005. Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. *J Infect Dis* 191: 2022-2032.
- Chomczynski P, Sacchi N 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162: 156-159.
- Chu CJ, Hussain M, Lok AS 2002. Hepatitis B virus genotype B is associated with earlier HBeAg seroconversion compared with hepatitis B virus genotype C. *Gastroenterology* 122: 1756-1762.
- Dane DS, Cameron CH, Briggs M 1970. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet* 1: 695-698.
- Da Silva LC, da Fonseca LE, Carrilho FJ, Alves VA, Sitnik R, Pinho JR 2000. Predictive factors for response to lamivudine in chronic hepatitis B. *Rev Inst Med Trop Sao Paulo* 42: 189-196.
- de Oliveira CM, Farias IP, Ferraz da Fonseca JC, Brasil LM, de Souza R, Astolfi-Filho S 2008. Phylogeny and molecular genetic parameters of different stages of hepatitis B virus infection in patients from the Brazilian Amazon. *Arch Virol* 153: 823-830.
- Devesa M, Pujol FH 2007. Hepatitis B virus genetic diversity in Latin America. *Virus Res* 127: 177-184.
- Drummond AJ, Rambaut A 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214.
- Ewing B, Hillier L, Wendl MC, Green P 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8: 175-185.
- Gomes-Gouveia MS 2005. *Characterization of complete genomes of the hepatitis B virus of different genotypes isolated in Brazil*, Secretaria da Saúde/Coordenadoria de Controle de Doenças/Programa de Pós-Graduação em Ciências para obtenção do grau de Mestre, São Paulo, 130 pp.
- Grimm D, Thimme R, Blum HE 2011. HBV life cycle and novel drug targets. *Hepatology* 53: 644-653.
- Gunther S, Li BC, Miska S, Kruger DH, Meisel H, Will H 1995. A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients. *J Virol* 69: 5437-5444.
- Kramvis A, Arakawa K, Yu MC, Nogueira R, Stram DO, Kew MC 2008. Relationship of serological subtype, basic core promoter and precore mutations to genotypes/subgenotypes of hepatitis B virus. *J Med Virol* 80: 27-46.
- Kramvis A, Kew MC 2007. Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes. *Hepatology* 46 (Suppl.): S9-S19.
- Mello FC, Souto FJ, Nabuco LC, Villela-Nogueira CA, Coelho HS, Franz HC, Saraiva JC, Virgolino HA, Motta-Castro AR, Melo MM, Martins RM, Gomes SA 2007. Hepatitis B virus genotypes circulating in Brazil: molecular characterization of genotype F isolates. *BMC Microbiol* 7: 103.
- Mulyanto, Depamede SN, Surayah K, Tjahyono AA, Jirintai, Nagashima S, Takahashi M, Okamoto H 2010. Identification and characterization of novel hepatitis B virus subgenotype C10 in Nusa Tenggara, Indonesia. *Arch Virol* 155: 705-715.
- Mulyanto, Depamede SN, Surayah K, Tsuda F, Ichiyama K, Takahashi M, Okamoto H 2009. A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. *Arch Virol* 154: 1047-1059.
- Ni YH, Chang MH, Wang KJ, Hsu HY, Chen HL, Kao JH, Yeh SH, Jeng YM, Tsai KS, Chen DS 2004. Clinical relevance of hepatitis B virus genotype in children with chronic infection and hepatocellular carcinoma. *Gastroenterology* 127: 1733-1738.
- Norder H, Courouce AM, Magnius LO 1994. Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* 198: 489-503.
- Orito E, Ichida T, Sakugawa H, Sata M, Horiike N, Hino K, Okita K, Okanoue T, Iino S, Tanaka E, Suzuki K, Watanabe H, Hige S, Mizokami M 2001. Geographic distribution of hepatitis B virus (HBV) genotype in patients with chronic HBV infection in Japan. *Hepatology* 34: 590-594.
- Sanger F, Nicklen S, Coulson AR 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463-5467.
- Santos AO, Alvarado-Mora MV, Botelho L, Vieira DS, Pinho JR, Carrilho FJ, Honda ER, Salcedo JM 2010. Characterization of hepatitis B virus (HBV) genotypes in patients from Rondonia, Brazil. *Virol J* 7: 315.
- Sitnik R, Pinho JR, Bertolini DA, Bernardini AP, Da Silva LC, Carrilho FJ 2004. Hepatitis B virus genotypes and precore and core mutants in Brazilian patients. *J Clin Microbiol* 42: 2455-2460.
- Stuyver L, De Gendt S, Van Geyt C, Zoulim F, Fried M, Schinazi RF, Rossau R 2000. A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. *J Gen Virol* 81: 67-74.
- Takahashi K, Aoyama K, Ohno N, Iwata K, Akahane Y, Baba K, Yoshizawa H, Mishiro S 1995. The precore/core promoter mutant (T1762A1764) of hepatitis B virus: clinical significance and an easy method for detection. *J Gen Virol* 76: 3159-3164.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876-4882.
- Tonetto PA, Goncalves NS, Fais VC, Vigani AG, Goncalves ES, Feltrin A, Goncalves FL Jr 2009. Hepatitis B virus: molecular genotypes and HBeAg serological status among HBV-infected patients in the Southeast of Brazil. *BMC Infect Dis* 9: 149.
- Yu H, Yuan Q, Ge SX, Wang HY, Zhang YL, Chen QR, Zhang J, Chen PJ, Xia NS 2010. Molecular and phylogenetic analyses suggest an additional hepatitis B virus genotype "I". *PLoS ONE* 5: e9297.
- Yu MW, Yeh SH, Chen PJ, Liaw YF, Lin CL, Liu CJ, Shih WL, Kao JH, Chen DS, Chen CJ 2005. Hepatitis B virus genotype and DNA level and hepatocellular carcinoma: a prospective study in men. *J Natl Cancer Inst* 97: 265-272.
- Zeng G, Wang Z, Wen S, Jiang J, Wang L, Cheng J, Tan D, Xiao F, Ma S, Li W, Luo K, Naoumov NV, Hou J 2005. Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotypes in China. *J Viral Hepat* 12: 609-617.
- Zhu CT, Dong CL 2009. Characteristics of general distribution of hepatitis B virus genotypes in China. *Hepatobiliary Pancreat Dis Int* 8: 397-401.