Infection, Genetics and Evolution 44 (2016) 34–42

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Molecular epidemiology of hepatitis B virus in Misiones, Argentina

Laura Noelia Mojsiejczuk [a,d,*], Carolina Torres [a,d], Ina Sevic [a,e], Inés Badano [b,d], Richard Malan [c], Diego Martin Flichman [a,d], Domingo Javier Liotta [b], Rodolfo Hector Campos [a,d]

[a] Cátedra de Virología, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956 4to Piso, Ciudad Autónoma de Buenos Aires, Argentina
[b] Laboratorio de Biología Molecular Aplicada, Facultad de Ciencias Exactas Químicas y Naturales. Universidad Nacional de Misiones, Av. Mariano Moreno 1375, Posadas, Misiones, Argentina
[c] Banco de Sangre Central de Misiones, Av. Cabred y Av. López Torres, Posadas, Misiones, Argentina
[d] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
[e] Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), Argentina

A R T I C L E I N F O

Article history:
Received 11 December 2015
Received in revised form 10 June 2016
Accepted 15 June 2016
Available online 16 June 2016

Keywords:
Hepatitis B virus
Molecular epidemiology
Argentina
Phylogenetic analysis
Bayesian coalescent analysis

A B S T R A C T

Hepatitis B virus (HBV) infection is a major public health problem worldwide. The aims of this study were to describe the molecular epidemiology of HBV in the Province of Misiones, Argentina and estimate the phylodynamic of the main groups in a Bayesian coalescent framework. To this end, partial or complete genome sequences were obtained from 52 blood donor candidates.
The phylogenetic analysis based on partial sequences of S/P region showed a predominance of genotype D (65.4%), followed by genotype F (30.8%) and genotype A as a minority (3.8%). At subgenotype level, the circulation of subgenotypes D3 (42.3%), D2 (13.5%), F1b (11.5%) and F4 (9.6%) was mainly identified.
The Bayesian coalescent analysis of 29 complete genome sequences for the main groups revealed that the subgenotypes D2 and D3 had several introductions to the region, with ancestors dating back from 1921 to 1969 and diversification events until the late '70s. The genotype F in Misiones has a more recent history; subgenotype F4 isolates were intermixed with sequences from Argentina and neighboring countries and only one significant cluster dated back in 1994 was observed. Subgenotype F1b isolates exhibited low genetic distance and formed a closely related monophyletic cluster, suggesting a very recent introduction.
In conclusion, the phylogenetic and coalescent analyses showed that the European genotype D has a higher circulation, a longer history of diversification and may be responsible for the largest proportion of chronic HBV infections in the Province of Misiones. Genotype F, especially subgenotype F1b, had a more recent introduction and its diversification in the last 20 years might be related to its involvement in new transmission events.
© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Hepatitis B virus (HBV) infection is a major public health problem worldwide. It is estimated that two billion people have been infected and currently more than 240 million are living with chronic HBV infection, increasing their risk to develop liver cirrhosis and hepatocellular carcinoma (WHO, 2014).

HBV belongs to the *Hepadnaviridae* family. Its genome is a partially double-stranded DNA molecule of approximately 3.2 kb in length with a highly compact coding structure consisting of four overlapping open reading frames (ORFs). Despite its DNA nature, HBV employs an error-prone polymerase reverse transcriptase as part of its replication process. As consequence, the nucleotide substitution rate for HBV is higher than those observed in other dsDNA viruses, with the estimated range being mostly between 10^{-4} to 10^{-6} substitutions per site per year (s/s/y) (Hannoun et al., 2000; Wang et al., 2010).

HBV genetic diversity is reflected through the existence of eight well described genotypes (A to H) and two additional (I and J) tentatively proposed (Shi et al., 2013; Tatematsu et al., 2009). These groups show a characteristic geographic distribution and have been associated with a particular ethnic origin and/or anthropological history (Arauz-Ruiz et al., 2002; Norder et al., 1994; Robertson and Margolis, 2002; Stuyver et al., 2000).

The role of genotypes (gts) or subgenotypes (sgts) in the outcome of infection is still controversial, although an increasing number of studies suggest a significant role in the progression and severity of liver disease, seroconvertion rate and antiviral treatment outcome (Cao, 2009; González López Ledesma et al., 2015; Lin and Kao, 2011; Livingston et al., 2007a).

Studies about the distribution of HBV gts in a specific geographical area have shown that it generally reflects the demographic history of the region. Particularly, epidemiological studies in Latin American countries have shown mainly the circulation of gt-F, related to the Native

* Corresponding author at: Cátedra de Virología, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956 4to Piso, C1113AAB Ciudad Autónoma de Buenos Aires, Argentina.
E-mail address: lmojsiejczuk@ffyb.uba.ar (L.N. Mojsiejczuk).

http://dx.doi.org/10.1016/j.meegid.2016.06.032
1567-1348/© 2016 Elsevier B.V. All rights reserved.

American population, and gt-A and gt-D, as a signature of the European colonization that began in the sixteenth century and included slave trade from Africa (Alvarado-Mora and Rebello Pinho, 2013; Campos et al., 2005; Devesa and Pujol, 2007).

The prevalence of HBV in Argentina is unknown, since no population-based studies have been conducted. Nonetheless, in 2011 a national prevalence study in blood donors was carried out showing a prevalence of 0.198% for HBsAg and 2.007% for Anti-HBc; in particular, significantly higher prevalences have been found for the Northwestern and Northeastern regions (Flichman et al., 2014). Previous studies described the HBV molecular epidemiology in the metropolitan area of Buenos Aires, and Central and Northwest regions of the country (Gallego et al., 2014; Pezzano et al., 2011; Piñeiro y Leone et al., 2008). However, there is a paucity of data in regard to the circulating viral gts in the Northeast region, where a higher HBV prevalence and singular demographic characteristics are noticed.

Particularly, the Province of Misiones, which showed a prevalence of 0.656% for HBsAg and 7.039% for Anti-HBc (Flichman et al., 2014), is located in the northeast of Argentina and is sourrounded by over 1000 km of international borders with two neighboring countries: Brazil and Paraguay. This territory has been mostly uninhabited until the late 19th century when a massive wave of migration took place. The immigrants were from different European countries and this process lasted for approximately fifty years (Gallero and Krautstolf, 2009). Additionally, continuous migrations from neighboring regions in the recent years were common. Therefore, the settlement process of this region has been characterized by a recent origin, celerity and heterogeneity, producing a complex scenario in which the current population of the province took shape.

In this context, the aims of this study were to analyze the HBV molecular epidemiology in the Province of Misiones, Argentina, and to estimate the phylodynamic of the main phylogenetic groups in a Bayesian coalescent framework.

2. Materials and methods

2.1. Samples

Serum samples of 76 HBsAg and Anti-HBc positive (AxSYM, Abbott Diagnostics, USA) blood donor candidates from the Central Blood Bank of Misiones Province (*Banco de Sangre Central de la Provincia de Misiones*), randomly collected from 2008 to 2011, were included. The exclusion criteria were: coinfection with human immunodeficiency virus or hepatitis C virus. None of the patients reported having received antiviral treatment. The study protocol was approved by the Bioethical Committee of the *Hospital Escuela de Agudos* "*Dr. Ramón Madariaga*" (Posadas, Misiones).

2.2. HBV DNA extraction, amplification and sequencing

HBV DNA was extracted from 200 μl of serum according to the previously described phenol-chloroform protocol (González López Ledesma et al., 2011). The full genome or partial genomic regions, corresponding to overlapping S/P genes (nt: 227–768) and to basal core promoter/precore/core gene (BCP/pC) (nt: 1661–2358) were amplified by nested PCRs (Supplementary Table S1). For PCR amplification, 5 ul of extracted DNA for the first round, and 2 ul for the second round, were added to a final volume of 25 ul PCR mix.

PCR products were purified using QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA) and submitted to direct nucleotide sequencing reaction in both directions (Macrogen, Inc. Seoul, Korea) with the same primers used for the second PCR round. Nucleotide sequences were deposited in GenBank under accession numbers JN688681, JN688682, JN688684, JN688686, JN688687, JN688692-JN688694, JN688705, KJ647345- KJ647348 for S/P genes; KJ647333-KJ647344 for BCP-pC region and JN688695, JN688699, JN688701, JN688703, JN688708-JN688713, JN688715-JN688717, JN688720, JN688722, KJ647349- KJ647356 for complete genomes.

2.3. Identification of HBV genotypes and subgenotypes

Genotype determination was based on the phylogenetic analysis of S/P region. Sgt assignation was based on phylogenetic analysis of concatenated S/P gene and BCP/pC regions (1216 nt) or complete genome sequences when these were available. Previous to the gt or sgt assignment analyses, recombination in partial and complete genome sequences was evaluated using the RDP4 v4.71 software package, by applying the RDP, GENECONV, MaxChi, BootScan and SisCan detection (Martin et al., 2015). For phylogenetic analysis, sequences obtained in this study and HBV sequences retrieved from GenBank database were aligned with ClustalX v2.1 (Larkin et al., 2007) and edited with BioEdit v7.1.3.0 (Hall, 1999). Evolutionary models were selected according to Akaike Information Criterion (AIC) statistics (Akaike, 1974) obtained with jModeltest v2.1 (Darriba et al., 2012). Despite the influence of overlapping genes on the evolution of HBV, a single substitution model was used for each dataset and gene partitioning models were not evaluated. The phylogenetic signal of datasets was evaluated by likelihood-mapping analysis of 10,000 random quartets using TreePuzzle v5.2 (Schmidt et al., 2002; Strimmer and von Haeseler, 1997). Phylogenetic trees were constructed using Maximum Likelihood (ML) method performed with PhyML v3.0 (Guindon and Gasguel, 2003). The robustness of the reconstructed phylogenies was evaluated by bootstrap analysis (1000 replicates). On the other hand, amino acid and nucleotide signature patterns were assessed in order to approximate the sgt in those samples where only the S/P region could be sequenced. It is worth noting that the subtype assignment by this methodology should not be taken as conclusive. Signature patterns from published works were revised and updated for new sequences (Piñeiro y Leone et al., 2008; Yousif and Kramvis, 2013). Briefly, sgt reference sequences were obtained from GenBank, and dataset of each gt was constructed and analyzed separately with VisSPA v1.6.2 (Korber and Myers, 1992) and inspected manually. A pattern was considered as a signature for a group if it was present in more than 90% of the sequences of the query set (i.e. a particular sgt or cluster) and in less than 10% of the samples of the references background set. Finally, in order to identify mutations with clinical or epidemiological relevance, nucleotide and deduced amino acid sequences belonging to S protein, viral polymerase and the BCP/pC regions of the isolates reported in this study were compared to prototype sequences for each gt/sgt.

2.4. Bayesian coalescent analysis

Bayesian coalescent analyses were performed to estimate the population dynamics and the time to the most recent common ancestor (tMRCA) for the main sgts detected. Datasets of each sgt were constructed and analyzed separately in order to study their individual phylodynamic pattern, since each group might present different evolutionary histories. The datasets included the complete genome sequences obtained in this work (Sgt-D2 $n = 5$, Sgt-D3 $n = 17$, Sgt-F1b $n = 3$, Sgt-F4 $n = 4$) and those obtained from GenBank, with known collection date and country (Sgt-D2 $n = 91$, Sgt-D3 $n = 67$, Sgt-F1b $n = 96$, Sgt-F4 $n = 48$).

The time-scale calibration was based on the isolation date of samples. Analyses were carried out using a single substitution model for the whole dataset, estimated as described above, and the effect of partitioning was not evaluated. The uncorrelated lognormal (UCLN) molecular clock model and Bayesian skyline plot (BSP) demographic model implemented in the BEAST v1.8.1 software package (Drummond and Rambaut, 2007) were used as coalescent prior. Analyses were run for 5×10^7 generations or up to achieve the convergence of parameters, which was assessed on the basis of effective sample sizes (ESS) values higher than 200 after a 10% of the sampling was discarded

as burn-in, using Tracer v1.6. Uncertainty in parameter estimates was evaluated in the 95% highest posterior density (HPD95%) interval. Maximum clade credibility trees (MCCT) were summarized using the "common ancestor heights" algorithm in TreeAnnotator v1.8.2 (Heled and Bouckaert, 2013) and visualized with FigTree v1.4.2 (available at http://tree.bio.ed.ac.uk/soft-ware/figtree/).

3. Results

3.1. Identification of HBV genotypes and subgenotypes

The genetic diversity of HBV in the Province of Misiones was analyzed by characterizing 76 HBsAg and Anti-HBc positive serum samples. Overall, HBV DNA was detected in 52 out of 76 (68%) isolates. The genetic regions that were able to be sequenced in each isolate are detailed in the Supplementary Table S2. No statistical evidence of intergenotype recombination was found in the analyzed sequences. The ML phylogenetic analysis of overlapping S/P region allowed grouping 34 samples as gt-D (65.4%), sixteen samples as gt-F (30.8%) and two samples as gt-A (3.8%) (Fig. 1). The phylogenetic signal of S/P region, concatenated S/P gene and BCP/pC regions, and the complete genome are shown in the Supplementary Fig. S1.

In order to determine the sgt in gt-D samples, full-length genomes, available for 22 out of 34 isolates were analyzed. This approach allowed classifying five samples as sgt-D2 and seventeen as sgt-D3 (Supplementary Fig. S2). Besides, an additional ML phylogenetic analysis using concatenated S/P gene and BCP/pC regions grouped two additional sequences as sgt-D2 and five as sgt-D3 (Supplementary Fig. S3). One sample (HBV234) analyzed by this methodology was unable to be subtyped because of low bootstrap support for the clustering with reference sequences. Finally, based on amino acid signatures, two out of five remaining isolates were classified as sgt-D2 (rt126R) (Yousif and Kramvis, 2013). Subtyping was not achieved in the remaining three samples. It is worth mentioning that, except for sgt-D2, no nucleotide or amino acid signature patterns were identified, precluding a conclusive subgenotyping based on partial genomic sequences.

As regards the gt-F isolates, full-length genomes were obtained in seven out of sixteen samples and the ML phylogenetic analysis classified three of them as sgt—F1b and four as sgt-F4 (Supplementary Fig. S2). In four of the remaining nine samples, the analysis of concatenated S/P gene and BCP/pC regions assigned three to sgt—F1b and one to sgt-F4 (Supplementary Fig. S3). Finally, based on nucleotide or amino acid signatures, four isolates were classified as sgt—F1b (nt 373C and 562T) and one as sgt-F4 (nt 482A and 493T; amino acids s110I, rt116N and rt120Y) (Piñeiro y Leone et al., 2008).

Only partial sequences of S/P region were obtained in the two gt-A samples. The signature pattern suggested that sample HBV025 (amino acids rt127L and rt161I) might be classified as sgt-A1 and the sample HBV182 (nt 514A and 616G; amino acids rt127M and rt161V) as sgt-A2 (Piñeiro y Leone et al., 2008).

Summarizing, the overall distribution of sgt was as follows: D2 (13.5%), D3 (42.3%), gt-D unassigned (9.6%); F1b (11.5%), F4 (9.6%) and gt-F unassigned (9.6%); gt-A unassigned (3.8%). No discrepancies were observed between the subgenotyping methods used.

3.2. Characterization of circulating HBsAg and polymerase variants

Based on the comparison with prototypic sequences of each gt/sgt, amino acid modifications within the major hydrophilic region (MHR) of HBsAg (aa 99–169) were observed in ten out of 52 (19.2%) samples. Six of these samples showed changes inside the "a" determinant, all of them previously described (Table 1). Amino acid substitutions on the reverse transcriptase catalytic domains were observed in twelve out of the 52 samples (23.1%), three of which were previously associated to antiviral resistance (rtA200V, rtS202I and rtN238D; Table 1).

3.3. Basal core promoter and preCore mutations

All gt-D samples with sequences of BCP/pC region available ($n = 30$) showed mutations affecting HBeAg expression (Table 1). The mutation G1896A in the pC region, which creates a premature stop codon that prevents production of HBeAg, was the most common, being found in 23 out of 30 (76.6%) isolates. Six samples (20%) showed mutations in the pC initiation codon (nucleotides 1814 to 1816) and one isolate (3.3%) exhibited a stop in the second codon of HBeAg caused by the substitution C1817T. In addition, substitution G1899A was observed in fourteen samples (46.6%). This means that 29 out of 30 gt-D samples (96.7%) presented stop codons that preclude HBeAg expression. On the other hand, the most frequent mutations in BCP region were T1753C in twelve cases (40.0%), G1764A alone in two cases (36.7%), the double mutant A1762T/G1764A in nine cases (30.0%) and C1766T in one case (3.3%).

Regarding gt-F, the six sgtF1b analyzed samples showed a wild type sequence both in BCP and pC regions. Whereas, in sgt-F4 isolates ($n = 5$) no mutations in the BCP were observed, but all of them showed the mutation G1896A and two also presented the substitution G1899A in the pC region (Table 1).

3.4. Bayesian coalescent analyses

The population dynamics and the tMRCA for the main sgts found (D2, D3, F1b and F4) were analyzed in a Bayesian coalescent framework. Complete genome datasets of each subgroup were tested separately under a relaxed molecular clock model and a non-parametric demographic coalescent model (Bayesian Skyline). The global median tMRCAs and substitution rates estimated for each dataset are shown in Table 2.

The MCCT analysis showed four out of five sgt-D2 sequences from Misiones in a monophyletic cluster closely related to Russian and Eastern Europe sequences, dated back in 1921 (HPD95% = 1809–1986) with diversification events until around 1959. Whereas, the sample HBV128 and isolates from Greenland and New Caledonia formed a separated group with a high posterior probability value (Fig. 2 and Supplementary Fig. S4).

For sgt-D3, three monophyletic clades containing Misiones sequences, dated in 1955, 1956 and 1969 (clade D3.a, D3.b and D3.c, respectively, in Fig. 2 and Supplementary Fig. S5), with final diversification events around 1978, were observed. These sequences were related to isolates from Western and Southeastern Europe (Belgium and Serbia, respectively) and Brazilian sequences (Supplementary Fig. S5).

In relation to gt-F, the three sgt—F1b isolates from Misiones were closely related and formed a highly supported monophyletic cluster dated back to 2006 (HPD95% = 2001–2009) (Fig. 2 and Supplementary Fig. S6). This cluster was grouped with sequences from Argentina and Uruguay. Whereas, sgt-F4 isolates from Misiones were intermixed with South American sequences and only one significant clade, including two sequences with a common ancestor in 1994 (HPD95% = 1980–2007), was observed (Fig. 2 and Supplementary Fig. S7).

4. Discussion

The Latin America population is ethnically diverse as a result of a complex genetic admixture between Native American, European and African descendants (Salzano and Sans, 2014). This feature imposes a complex setting for the viral dynamics due to the interaction between viral gts with particular evolutionary histories and hosts with different genetic backgrounds. Particularly, this process of admixture is quite recent in the Province of Misiones, making it an attractive scenario for the study of virus-host interaction. In addition to that, a significant evidence of a differential impact of HBV gts and sgts on clinical outcomes, severity of liver disease and response to antiviral therapy has been found (Kao,

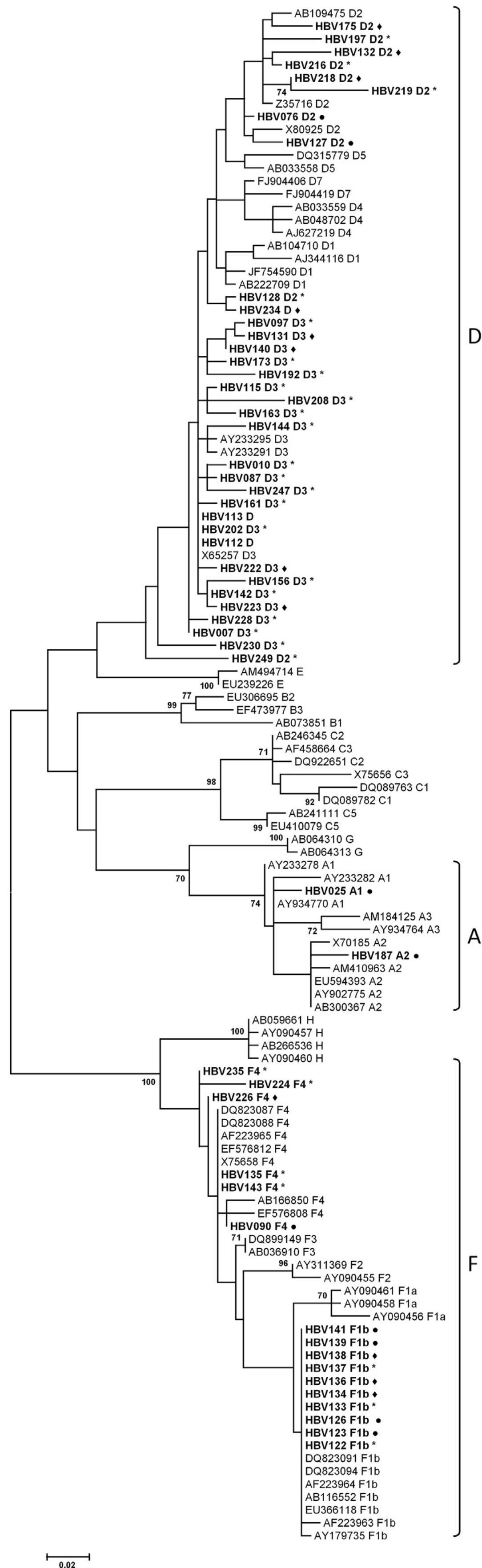


Fig. 1. Maximum-likelihood phylogenetic tree constructed on the small S/P genes partial sequences (525 bp). Analysis includes reference sequences retrieved from GenBank, indicated by their accession numbers, and 52 sequences from Misiones generated in this study (in bold), followed by the respective genotype/subgenotype assigned. Isolates in which the subgenotype assignation was based on phylogenetic analysis of complete genome sequences, phylogenetic analysis of concatenated S/P gene and BCP/pC regions, or nucleotides/amino acid pattern are indicated with *, ♦ and •, respectively. The numbers at each node correspond to bootstrap values (greater than 70%) obtained with 1000 replicates. The scale bar indicates the genetic distances.

2002; Lin and Kao, 2011; Livingston et al., 2007b; Yuan et al., 2007). Therefore, a deeper knowledge of strains that circulate in a population has become a matter of major interest.

In this study, we analyzed the genotype distribution, genetic diversity and population dynamics of HBV in blood donor samples from the Province of Misiones, Argentina.

Although the analysis of blood donors has limitations to describe the epidemiological scenario in the general population, in this work, the infecting viral gt was determined in 52 out of 76 (68%) analyzed samples. Moreover, twenty-nine complete genome sequences were obtained. This allowed us, on the one hand, to perform a deeper molecular characterization of HBV in a group of individuals that are not frequently studied since they are non-symptomatic and, on the other hand, to study and make hypothesis about the events that shaped the current epidemiology.

In the cohort analyzed in this study gt-D was the most prevalent, followed by gt-F and gt-A in a lesser extent. Gt-D is one of the most widespread gts worldwide, predominating in the Mediterranean area, Middle East and Indian subcontinent (Yousif and Kramvis, 2013). It has also been reported in Latin American countries, and it is related to an European origin in these populations (Alvarado-Mora and Rebello Pinho, 2013; Campos et al., 2005; Devesa and Pujol, 2007). Previous studies have described the circulation of this gt in Argentina, particularly in the metropolitan area of Buenos Aires and the Central region. However the prevalence reported in this areas were lower than that found in Misiones (Barbini et al., 2013; Gallego et al., 2014; González López Ledesma et al., 2015; Piñeiro y Leone et al., 2008). On the contrary, a higher prevalence has been reported in the Southern region of Brazil (Bertolini et al., 2012; Carrilho et al., 2004; Mello et al., 2007; Reis et al., 2011), suggesting that the HBV epidemiology of Misiones resembles that found in this neighboring area more than that described for other regions of Argentina. Particularly, Bertolini et al. (2012) found a significant association between the infection with gt-D and the European ancestry in blood donors from Southern Brazil. Unfortunately, there is no available data in regard to the ethnic origin of patients included in the present study, although different genetic molecular studies performed in the region confirmed a high proportion of European genomic ancestry in several populations of Misiones (Badano et al., 2012; Corach et al., 2010).

These facts are supported by the human settlement process of Misiones which has involved both a spontaneous migration, mostly from Italian and German colonies previously settled in Southern Brazil, and official campaigns promoted by the Argentinean government. The last included several migratory waves from Europe and Middle East countries that, during the first half of the 20th century, increased sevenfold the population of the province (Gallero and Krautstolf, 2009).

In agreement with the demographic records, the coalescent analysis showed that HBV sgt-D2 isolates from Misiones could have a common ancestor in the early 20th century and might be closely related to sequences from Russia, Estonia, Latvia, Poland, Belarus and Kazakhstan. These results suggest that this sgt would have been introduced by immigrants from these regions and subsequently underwent a diversification process in the province. On the other hand, the analysis of the sgt-D3 sequences from Argentina suggests at least three introductions to the region, with isolates mainly related to sequences from Belgium, Serbia and Brazil, and sequences from Misiones reported previously (Delfino

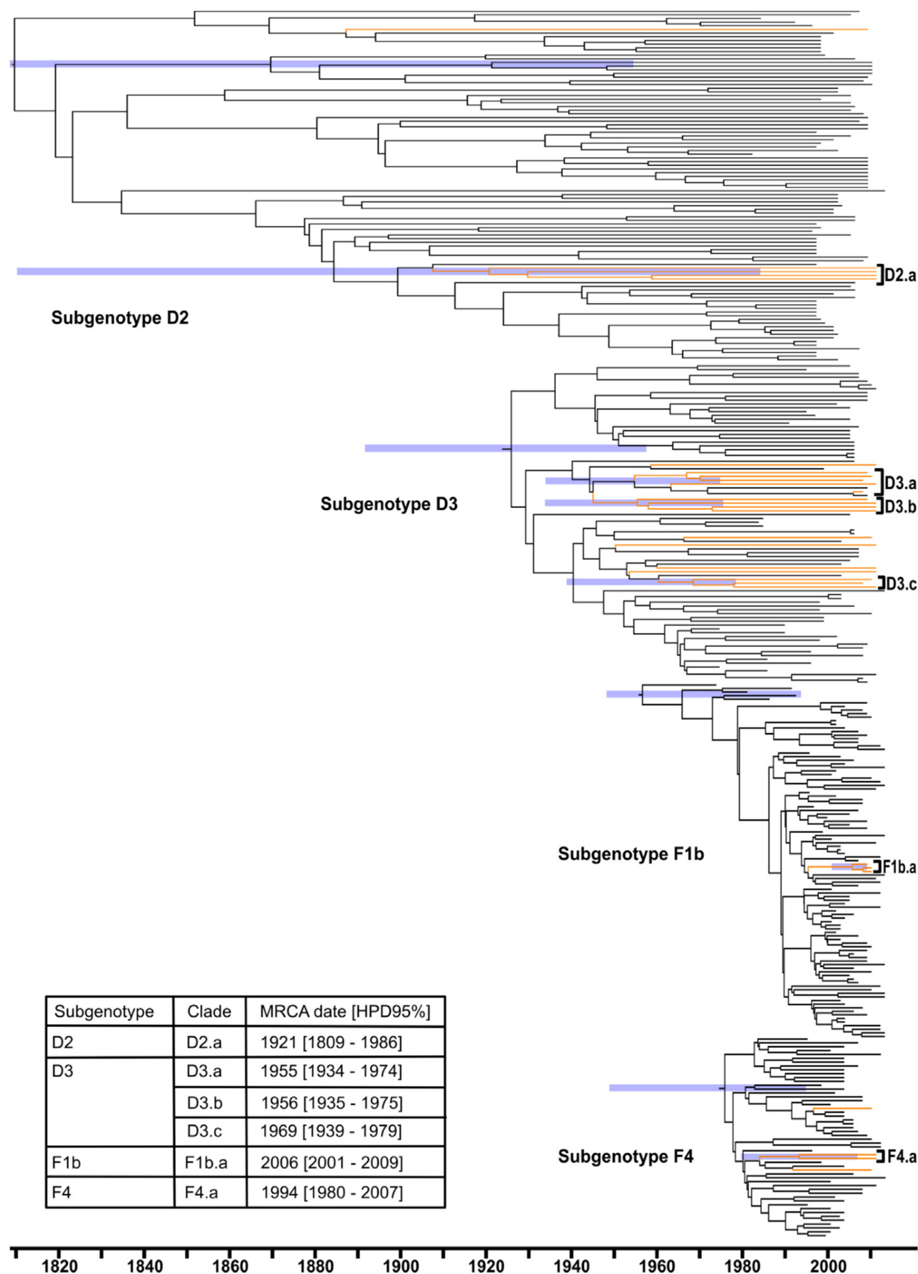
Table 1
Nucleotide or amino acid substitutions on the MHR of HBsAg, rtPol catalytic domain, basal core promoter/precore (BCP-pC) and core genes open reading frames (ORFs).

(Sub) Genotype	Samples	Genomic Region										
		Bcp- pC (1690–1900 nt)										
		MHR (99–169 aa)[1]	rtPol catalytic domains[2]	T1753	A1762	G1764	A1814	T1815	G1816	G1896	G1899	Others
Gt-A unassigned	HBV025	–	–				NA					
	HBV187	Q101P	–				NA					
Gt-D unassigned	HBV076	–	–				NA					
	HBV112	–	–				NA					
	HBV113	–	–				NA					
	HBV127	–	–				NA					
Sgt-D2	HBV234	–	L164 M	C	–	–	–	C	–	A	A	
	HBV128	–	L164 M	–	–	–	–	–	–	A	A	
	HBV175	–	–	–	–	–	–	–	–	A	–	
	HBV197	–	C256S	C	T	A	–	–	–	A	–	
	HBV216	–	–	–	–	A	–	–	–	A	–	
	HBV218	Y100F	–	–	–	–	–	C	–	–	–	
	HBV219	Y100F, **R122Q**, G159E	G172E/G, **S202I**	–	–	–	–	–	–	A	–	
	HBV249	**Y134F**, F161I	I169Y	–	–	–	–	–	–	A	–	
Sgt-D3	HBV007	–	R242K	–	–	–	C	–	–	–	–	
	HBV010	–	–	–	–	A	–	–	–	A	–	
	HBV087	–	–	C	–	–	–	C	–	A	A	
	HBV097	–	–	C	T	A	–	–	–	–	A	
	HBV115	–	–	–	–	–	–	–	–	A	A	C1766T
	HBV131	–	–	–	–	–	–	–	–	A	–	
	HBV132	T127 A, T131P	–	Y	–	–	–	–	–	A	A	
	HBV140	–	–	–	–	–	–	–	–	A	A	
	HBV142	–	–	–	T	A	–	–	–	A	A	
	HBV144	–	N76D, **N238D**	C	T	A	–	–	–	A	A	
	HBV156	–	N76D	Y	T	A	–	–	–	A	–	
	HBV161	–	–	C	–	–	–	–	–	A	–	
	HBV163	E164G	–	C	T	A	–	–	–	A	A	
	HBV173	–	–	G	T	A	–	–	T	–	–	
	HBV192	–	I233V	–	–	–	–	–	–	A	A	
	HBV202	–	–	C	–	–	–	–	T	–	A	
	HBV208	–	N76D	C	T	A	–	–	–	A	A	
	HBV222	M103I, **Q129H**	–	–	T	A	–	–	–	A	–	
	HBV223	–	–	–	–	–	–	–	–	A	–	
	HBV228	–	–	–	–	–	–	–	–	–	–	C1817T
	HBV230	**R122Q**, **Y134F**	–	–	–	–	–	–	–	A	A	
	HBV247	**R122Q**, D144D/N, E164E/K	G172E/I	–	–	–	–	–	–	A	–	
Gt-F unassigned	HBV090	–	–				NA					
	HBV123	–	–				NA					
	HBV126	–	–				NA					
	HBV139	–	–				NA					
	HBV141	–	–				NA					
Sgt-F1b	HBV122	–	–	–	–	–	–	–	–	–	–	
	HBV133	–	–	–	–	–	–	–	–	–	–	
	HBV134	–	–	–	–	–	–	–	–	–	–	
	HBV136	–	–	–	–	–	–	–	–	–	–	
	HBV137	–	–	–	–	–	–	–	–	–	–	
	HBV138	–	–	–	–	–	–	–	–	–	–	
Sgt-F4	HBV135	–	–	–	–	–	–	–	–	A	–	
	HBV143	–	–	–	–	–	–	–	–	A	–	T1845Y
	HBV224	–	–	–	–	–	–	–	–	A	–	T1845G
	HBV226	–	–	–	–	–	–	–	–	A	A	T1845C
	HBV235	D144D/N, A157 A/V, G159E/G, E164K/E	W79G/W, G172 E/G, A200 A/V	–	–	–	–	–	–	A	A	T1845C C2409A

Mutations associated with: vaccine escape, failure in diagnostic assays, immunotherapy escape or antiviral resistance are shown in bold.
[1] Amino acidic position on small S gene.
[2] Amino acidic position on the retrotranscriptase catalytic domains A (75–90 aa), B (159–182 aa), C (200–210 aa), D (230–241 aa), E (247–257 aa).

Table 2
Global median tMRCAs and substitution rates estimates for each dataset. The tMRCA and MRCA values are informed in years; substitution rates as s/s/y (substitution sites per year).

Subgenotype	tMRCA [HPD95%]	MRCA date [HPD95%]	Substitution rate [HPD95%]
D2	203 [456–57]	1810 [1557–1956]	1.4×10^{-4} [2.4×10^{-5} – 2.6×10^{-4}]
D3	86 [121–56]	1927 [1892–1957]	2.9×10^{-4} [2.0×10^{-4} – 4.0×10^{-4}]
F1b	55 [79–38]	1958 [1934–1975]	1.9×10^{-4} [1.2×10^{-4} – 2.5×10^{-4}]
F4	35 [64–18]	1978 [1949–1995]	4.7×10^{-4} [1.3×10^{-4} – 7.0×10^{-4}]



Subgenotype	Clade	MRCA date [HPD95%]
D2	D2.a	1921 [1809 - 1986]
D3	D3.a	1955 [1934 - 1974]
	D3.b	1956 [1935 - 1975]
	D3.c	1969 [1939 - 1979]
F1b	F1b.a	2006 [2001 - 2009]
F4	F4.a	1994 [1980 - 2007]

Fig 2. Time-scaled Maximum clade credibility trees (MCCT) for HBV complete genome datasets of sgt-D2, D3, F1b and F4. The branches corresponding to lineages reported in this study are colored. The monophyletic clades from Misiones are indicated with brackets and the respective MRCA dates are detailed in the lower left inset. The HPD95% values for MRCA are shown in horizontal bars; upper limit of the HPD95% interval for Subgenotype D2 tMRCA was trimmed for clarity purpose.

et al., 2014a). Although the coalescent analysis showed that sgt-D2 isolates could have the oldest ancestors among Misiones samples, the diversification process for both European sgts detected (D2 and D3) would have taken place simultaneously from the early 20th century to around 1980. Therefore, it can be speculated that the gt-D currently circulating in the Province of Misiones probably derives from the strains introduced by immigrants from Eastern Europe, Mediterranean and Middle East countries in the process of settlement of the region in early 20th century.

The phylogenetic analysis also revealed the circulation of gt-F strains in Misiones. This gt has been found in Native American populations of Alaska, Central and South America (Arauz-Ruiz et al., 1997; Devesa et al., 2008; Livingston et al., 2007b). In Argentina, the Northern region presents the highest prevalence of the country, mainly by the presence of sgt—F1b and F4, related to a major Native American ethnical background in this region (Piñeiro y Leone et al., 2008; Torres et al., 2011). Their circulation has been previously reported in the Province of Misiones, both in blood donor population and Native American communities, although the low number of sequences obtained makes the prevalence analysis difficult (Delfino et al., 2014b, 2012). There are about 80 small Native American communities in the rainforest of Misiones, but their inhabitants rarely contribute to the blood donor population. Therefore, the presence of gt-F in our cohort is unlikely to be directly related to Native American communities only. In the present study, a 30.8% of isolates corresponded to gt-F. This proportion is lower than that found in other provinces of Northern Argentina (91.7%) or in the metropolitan area of Buenos Aires (54.3%) (Barbini et al., 2013; Gallego et al., 2014; González López Ledesma et al., 2015; Piñeiro y Leone et al., 2008). This finding highlights the relevance of gt-D introduction to the region associated to European migrations, as described above.

The analysis of sgt—F1b sequences from Misiones showed a very low genetic divergence between them. The time-scaled MCCT showed

that the three sequences would have a common ancestor dated only four years before the sampling date, suggesting that, unlike the gt-D, the F1b has been recently introduced in Misiones. This is consistent with the current epidemiological data from other regions of Argentina, pointing to this sgt as responsible for most of the acute and new chronic HBV infections (Barbini et al., 2013; Gallego et al., 2014; González López Ledesma et al., 2015; Pezzano et al., 2011).

The sgt-F4 isolates from Misiones exhibited a higher genetic distance and an older ancestor than F1b, and a disperse distribution in the MCCT, with the exception of one cluster having two sequences. This suggests several introductions of F4 strains that may be related to the geographical location of Misiones, i.e. the proximity to regions with high F4 prevalence such as Bolivia and Northern Argentina, and the frequent contact with these neighboring populations. Alternatively, they could represent autochthonous strains that circulated in the region before the massive introduction of European HBV gts in the early 20th century. The results of coalescent analysis must be interpreted with caution given the low number of available sequences and their short sampling time dispersion, which could lead to an underestimation of the tMRCA for the global dataset.

Lastly, we identified the presence of gt-A as a minority, with two isolates that might belong to sgt-A1 and sgt-A2 each. Sgt-A1 is mainly found in the Southern and Eastern regions of Africa and Asia (Kramvis and Kew, 2007), while A2 is present in Northwestern Europe (Norder et al., 2004). Sgt-A2 might have been continuously introduced in Argentina by the successive European migrations since the 16th century (Arauz-Ruiz et al., 1997; Mbayed et al., 2009; Mello et al., 2007). Likewise, the introduction of sgt-A1 may be associated to the slave trade from Africa in the past (Motta-Castro et al., 2008). Unfortunately, only partial sequences were obtained from samples infected with this gt, making it difficult a deeper characterization.

Additionally, the presence of mutations with clinical or epidemiological relevance was evaluated. The analysis of HBsAg variants showed that twelve out of 52 analyzed sequences presented mutations within the MHR. These included several amino acid substitutions inside the “a” determinant or in sites related to immunological evasion and/or to diagnostic failure (Ma and Wang, 2012), which in contrast did not hamper the performance of the HBsAg screening assay used. Three gt-D isolates showed the R122Q substitution in a residue involved in the HBV attachment to hepatocytes, which could impair the viral infectivity (Sureau and Salisse, 2013).

The deduced amino acid sequence of the reverse transcriptase domain of the polymerase showed that three isolates had substitutions in residues associated with antiviral resistance. While the presence of resistant variants in a group of individuals not undergoing antiviral therapy (“healthy” blood donors) is noteworthy, the emergence of these mutations without antiviral selective pressure and their circulation in naïve patients has been previously described (Tan et al., 2012; Zheng et al., 2012). This points out their existence as a component of viral population diversity.

Several studies highlight the implication of HBV gt in the progression of infection related to the extension of the HBeAg positive stage, proposing that certain gts or sgts could exhibit a delayed HBeAg seroconversion and consequently would be associated with a more severe clinical course of infection (Chu et al., 2004; Yang et al., 2002). Different mutations have been associated with HBeAg seroconversion and their distribution in Misiones population was found to be biased according to the gt, as previously described (Kramvis et al., 2008; Lindh et al., 1997). It is worth mentioning that almost all gt-D samples presented stop codons that preclude the HBeAg expression and one-third of them exhibit more than one mutation related to HBeAg negative phenotype. Particularly, the substitution G1896A was frequently observed in both: gt-D and sgt-F4 isolates (76.6% and 100% respectively), consistent with previous reports (González López Ledesma et al., 2015). In addition, all the F1b strains analyzed in this study displayed a wild type BCP sequence, possibly related to more recent infections, as discussed above.

This study presents some limitation. Firstly, some partial sequences for which the phylogenetic analysis did not result in a conclusive assignment of subtype were classified by alternative methods. Despite these approaches allow a detailed epidemiological description, they might not be reliable on subtyping; particularly, a consensus could be necessary for an unambiguous classification of HBV variants, especially for gt-D. Secondly, no evidence of recombination was found in the sequences analyzed, although the lack of whole genome amplification and the methodology employed, i.e. direct sequencing of PCR fragments, prevents to reject the circulation of recombinant strains or gt mixtures and assess their impact on local epidemiology. Finally, unfortunately, the clinical characterization of patients and data about HBV DNA levels, HBeAg and Anti-HBe reactivity were not available. Therefore, the association between the clinical phase of infection and the gt or sgt found could not be evaluated.

In summary, the distribution of HBV gts reflects the demographic history of the Province of Misiones, Argentina. The phylogenetic and coalescent analyses suggest that several strains of the European gt-D would have been introduced during the first half of 20th century, showing a long history of diversification and resulting in the largest proportion of current chronic HBV infections. In contrast, the introduction of gt-F would be related to internal migrations, and particularly the sgt—F1b might have been more recently introduced in the region and might have diversified over the last 20 years, mainly associated with new transmission events.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.meegid.2016.06.032.

Acknowledgments

This work was supported by grants from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET; PIP2012-11220110100215), Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT; PICT2011-0705), and Comisión Nacional Salud Investiga (Ministerio de Salud de la Nación). LNM receives a doctoral fellowship from CONICET, Argentina. IS receives a doctoral fellowship from ANPCyT, Argentina. CT, IB, DMF and RHC are members of the researcher career of CONICET, Argentina. The authors gratefully acknowledge for technical assistance provided by Gabriela Decombard, Williams Pedrozo and Alberto Krupp.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19, 716–723. http://dx.doi.org/10.1109/TAC.1974.1100705.

Alvarado-Mora, M.V., Rebello Pinho, J.R., 2013. Distribution of HBV genotypes in Latin America. Antivir. Ther. 18, 459–465. http://dx.doi.org/10.3851/IMP2599.

Arauz-Ruiz, P., Norder, H., Robertson, B.H., Magnius, L.O., 2002. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J. Gen. Virol. 83, 2059–2073.

Arauz-Ruiz, P., Norder, H., Visoná, K.A., Magnius, L.O., 1997. Molecular epidemiology of hepatitis B virus in Central America reflected in the genetic variability of the small S gene. J. Infect. Dis. 176, 851–858. http://dx.doi.org/10.1086/516507.

Badano, I., Stietz, S.M., Schurr, T.G., Picconi, A.M., Fekete, D., Quintero, I.M., Cabrera, M.D.E., Campos, R.H., Liotta, J.D., 2012. Analysis of TNFα promoter SNPs and the risk of cervical cancer in urban populations of Posadas (Misiones, Argentina). J. Clin. Virol. 53, 54–59. http://dx.doi.org/10.1016/j.jcv.2011.09.030.

Barbini, L., Elizalde, M., Torres, C., Campos, R.H., 2013. Molecular epidemiology and genetic diversity of hepatitis B virus in Mar del Plata City. Argentina. Infect. Genet. Evol. 19, 152–163. http://dx.doi.org/10.1016/j.meegid.2013.07.007.

Bertolini, D.A., Gomes-Gouvêa, M.S., Vicente Guedes de Carvalho-Mello, I.M., Saraceni, C.P., Sitnik, R., Grazziotin, F.G., Laurindo, J.P., Rosa Fagundes, N.J., Carrilho, F.J., Rebello Pinho, J.R., 2012. Hepatitis B virus genotypes from European origin explains the high endemicity found in some areas from southern Brazil. Infect. Genet. Evol. 12, 1295–1304. http://dx.doi.org/10.1016/j.meegid.2012.04.009.

Campos, R.H., Mbayed, V.A., Piñeiro y Leone, F.G., 2005. Molecular epidemiology of hepatitis B virus in Latin America. J. Clin. Virol. 34, S8–S13. http://dx.doi.org/10.1016/S1386-6532(05)80028-9.

Cao, G.-W., 2009. Clinical relevance and public health significance of hepatitis B virus genomic variations. World J. Gastroenterol. 15, 5761–5769.

Carrilho, F.J., Moraes, C.R., Pinho, J.R.R., Mello, I.M., Bertolini, D.A., Lemos, M.F., Moreira, R.C., Bassit, L.C., Cardoso, R.A., Ribeiro-dos-santos, G., Silva, L.C.D., 2004. Hepatitis B virus

infection in Haemodialysis Centres from Santa Catarina State , Southern Brazil . Predictive risk factors for infection and molecular epidemiology. BMC Public Health 11, 1–11.
Chu, C.-M., Hung, S.-J., Lin, J., Tai, D.-I., Liaw, Y.-F., 2004. Natural history of hepatitis b e antigen to antibody seroconversion in patients with normal serum aminotransferase levels. Am. J. Med. 116, 829–834. http://dx.doi.org/10.1016/j.amjmed.2003.12.040.
Corach, D., Lao, O., Bobillo, C., Gaag, K.V.D., Zuniga, S., Vermeulen, M., Duijn, K.V., Goedbloed, M., Vallone, P.M., Parson, W., Knijff, P.D., Kayser, M., 2010. Inferring continental ancestry of Argentineans from autosomal , Y-chromosomal and mitochondrial DNA. Ann. Hum. Genet. 65–76. http://dx.doi.org/10.1111/j.1469-1809.2009.00556.x.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. http://dx.doi.org/10.1038/nmeth.2109.
Delfino, C.M., Berini, C., Eirin, M.E., Malan, R., Pedrozo, W., Krupp, R., Blejer, J., Espejo, R., Fierro, L., Puca, A., Oubiña, J.R., Mathet, V.L., Biglione, M.M., 2012. New natural variants of hepatitis B virus among Amerindians from Argentina with mainly occult infections. J. Clin. Virol. 54, 174–179. http://dx.doi.org/10.1016/j.jcv.2012.02.023.
Delfino, C.M., Berini, C., Pedrozo, W., Malan, R., Blejer, J., Oubiña, J.R., Mathet, V.L., Biglione, M.M., 2014a. First hepatitis B virus full-length genomic study among blood donors from Argentina: unexpected mutations in the circulating subgenotypes' proteins. http://dx.doi.org/10.1007/s11262–014–1159-4.
Delfino, C.M., Gentile, E.A., Castillo, A.I., Cuestas, M.L., Pataccini, G., Cánepa, C., Malan, R., Blejer, J., Berini, C., Eirin, M.E., Pedrozo, W., Oubiña, J.R., Biglione, M.M., Mathet, V.L., 2014b. Hepatitis B virus and hepatitis D virus in blood donors from Argentina: circulation of HBsAg and reverse transcriptase mutants. Arch. Virol. 159, 1109–1117. http://dx.doi.org/10.1007/s00705-013-1917-y.
Devesa, M., Loureiro, C.L., Rivas, Y., Monsalve, F., Cardona, N., Duarte, M.C., Poblete, F., Gutierrez, M.F., Botto, C., Pujol, F.H., 2008. Subgenotype diversity of hepatitis B virus American genotype F in Amerindians from Venezuela and the general population of Colombia. J. Med. Virol. 80, 20–26. http://dx.doi.org/10.1002/jmv.
Devesa, M., Pujol, F.H., 2007. Hepatitis B virus genetic diversity in Latin America. Virus Res. 127, 177–184. http://dx.doi.org/10.1016/j.virusres.2007.01.004.
Drummond, A.J., Rambaut, A., 2007. BEAST : Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214. http://dx.doi.org/10.1186/1471-2148-7-214.
Flichman, D.M., Blejer, J.L., Livellara, B.I., Re, V.E., Bartoli, S., Bustos, Ansola, C.P., Hidalgo, S., Cerda, M.E., Levin, A.E., Huenul, A., Riboldi, V., Treviño, E.M.C., Salamone, H.J., Nuñez, Fernández, R.J., Reybaud, J.F., Campos, R.H., 2014. Prevalence and trends of markers of hepatitis B virus, hepatitis C virus and human immunodeficiency virus in Argentine blood donors. BMC Infect. Dis. 14, 218. http://dx.doi.org/10.1186/1471-2334-14-218.
Gallego, F., Pisano, M.B., Torres, C., Caeiro, L., Martínez Wassaf, M., Balangero, M., Campos, R.H., Ré, V., 2014. Molecular epidemiology of hepatitis B virus in Córdoba. Argentina. J. Clin. Virol. http://dx.doi.org/10.1016/j.jcv.2014.06.030.
Gallero, C.M., Krautstolf, E.M., 2009. Proceso de poblamiento y migraciones en la provincia de Misiones, Argentina (1881–1970). Rev. Antropol. Avá 16, 245–264.
González López Ledesma, M.M., Galdame, O., Bouzas, B., Tadey, L., Livellara, B., Giuliano, S., Viaut, M., Paz, S., Fainboim, H., Gadano, A., Campos, R.H., Flichman, D.M., 2011. Characterization of the basal core promoter and precore regions in anti-HBe-positive inactive carriers of hepatitis B virus. Int. J. Infect. Dis. 15, e314–e320. http://dx.doi.org/10.1016/j.ijid.2010.12.009.
González López Ledesma, M.M., Mojsiejczuk, L.N., Rodrigo, B., Sevic, I., Mammana, L., Galdame, O., Gadano, A., Fainboim, H., Campos, R.H., Flichman, D.M., 2015. Hepatitis B virus genotype distribution and genotype-specific BCP/preCore substitutions in acute and chronic infections in Argentina. PLoS One 10, e0121436. http://dx.doi.org/10.1371/journal.pone.0121436.
Guindon, S., Gasguel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704. http://dx.doi.org/10.1080/10635150390235520.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Hannoun, C., Horal, P., Lindh, M., 2000. Long-term mutation rates in the hepatitis B virus genome. J. Gen. Virol. 81, 75–83.
Heled, J., Bouckaert, R.R., 2013. Looking for trees in the forest: summary tree from posterior samples. BMC Evol. Biol. 13, 1–11. http://dx.doi.org/10.1186/1471-2148-13-221.
Kao, J., 2002. Hepatitis B viral genotypes : clinical relevance and molecular characteristics. J. Gastroenterol. Hepatol. 17, 643–650. http://dx.doi.org/10.1046/j.1440-1746.2002.02737.x.
Korber, B., Myers, G., 1992. Signature pattern analysis: a method for assessing viral sequence relatedness. AIDS Res. Hum. Retrovir. 8, 1549–1560.
Kramvis, A., Arakawa, K., Yu, M.C., Nogueira, R., Stram, D.O., Kew, M.C., 2008. Relationship of serological subtype , basic Core promoter and precore mutations to genotypes/subgenotypes of hepatitis B virus. J. Med. Virol. 80, 27–46. http://dx.doi.org/10.1002/jmv.
Kramvis, A., Kew, M.C., 2007. Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes. Hepatol. Res. 37, S9–S19. http://dx.doi.org/10.1111/j.1872-034X.2007.00098.x.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948. http://dx.doi.org/10.1093/bioinformatics/btm404.
Lin, C., Kao, J.-H., 2011. The clinical implications of hepatitis B virus genotype: recent advances. J. Gastroenterol. Hepatol. 26 (Suppl. 1), 123–130. http://dx.doi.org/10.1111/j.1440-1746.2010.06541.x.
Lindh, M., Andersson, A., Gusdal, A., 1997. Genotypes, nt 1858 variants, and geographic origin of hepatitis B virus — large-scale analysis using a new genotyping method. J. Infect. Dis. 175, 1285–1293.
Livingston, S.E., Simonetti, J.P., Bulkow, L.R., Homan, C.E., Snowball, M.M., Cagle, H.H., Negus, S.E., McMahon, B.J., 2007a. Clearance of hepatitis B e antigen in patients with chronic hepatitis B and genotypes A, B, C, D, and F. Gastroenterology 133, 1452–1457. http://dx.doi.org/10.1053/j.gastro.2007.08.010.
Livingston, S.E., Simonetti, J.P., McMahon, B.J., Bulkow, L.R., Hurlburt, K.J., Homan, C.E., Snowball, M.M., Cagle, H.H., Williams, J.L., Chulanov, V.P., 2007b. Hepatitis B virus genotypes in Alaska native people with hepatocellular carcinoma : preponderance of genotype F. J. Infect. Dis. 195, 5–11. http://dx.doi.org/10.1086/509894.
Ma, Q., Wang, Y., 2012. Comprehensive analysis of the prevalence of hepatitis B virus escape mutations in the major hydrophilic region of surface antigen. J. Gen. Virol. 206, 198–206. http://dx.doi.org/10.1002/jmv.
Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4 : Detection and Analysis of Recombination Patterns in Virus Genomes 1–5. http://dx.doi.org/10.1093/ve/vev003.
Mbayed, V.A., Piñeiro y Leone, F.G., Pezzano, S.C., Campos, R.H., 2009. Molecular characterization of hepatitis B virus genotype A from Argentina and Brazil. Arch. Virol. 154, 525–529. http://dx.doi.org/10.1007/s00705-009-0328-6.
Mello, F.C., Souto, F.J.D., Nabuco, L.C., Villela-Nogueira, C.A., Coelho, H.S.M., Franz, H.C.F., Saraiva, J.C.P., Virgolino, H.A., Motta-Castro, A.R., Melo, M.M.M., Martins, R.M.B., Gomes, S.A., 2007. Hepatitis B virus genotypes circulating in Brazil: molecular characterization of genotype F isolates. BMC Microbiol. 7, 103. http://dx.doi.org/10.1186/1471-2180-7-103.
Motta-Castro, A.R., Martins, R.M.B., Araujo, N.M., Niel, C., Facholi, G.B., Lago, B.V., Mello, F.C.A., Gomes, S.A., 2008. Molecular epidemiology of hepatitis B virus in an isolated Afro-Brazilian community. Arch. Virol. 153, 2197–2205. http://dx.doi.org/10.1007/s00705-008-0237-0.
Norder, H., Couroucé, A.M., Coursaget, P., Echevarria, J.M., Lee, S.-D.D., Mushahwar, I.K., Robertson, B.H., Locarnini, S.A., Magnius, L.O., 2004. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. Intervirology 47, 289–309. http://dx.doi.org/10.1159/000080872.
Norder, H., Couroucé, A.-M., Magnius, L.O., 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. http://dx.doi.org/10.1006/viro.1994.1060.
Pezzano, S.C., Torres, C., Fainboim, H.A., Bouzas, M.B., Schroder, T., Giuliano, S.F., Paz, S., Alvarez, E., Campos, R.H., Mbayed, V.A., 2011. Hepatitis B virus in Buenos Aires, Argentina: genotypes, virological characteristics and clinical outcomes. Clin. Microbiol. Infect. 17, 223–231. http://dx.doi.org/10.1111/j.1469-0691.2010.03283.x.
Piñeiro y Leone, F.G., Pezzano, S.C., Torres, C., Rodríguez, C.E., Garay, M.E., Fainboim, H.A., Remondegui, C., Sorrentino, A.P., Mbayed, V.A., Campos, R.H., 2008. Hepatitis B virus genetic diversity in Argentina: dissimilar genotype distribution in two different geographical regions; description of hepatitis B surface antigen variants. J. Clin. Virol. 42, 381–388. http://dx.doi.org/10.1016/j.jcv.2008.01.018.
Reis, L.M.W., Soares, M.A., Franca, P.H., Soares, E.A., Bonvicino, C.R., 2011. Clonal analysis of hepatitis B viruses among blood donors from Joinville , Brazil : evidence of dual infections, Intragenotype recombination and markers of risk for hepatocellular carcinoma. J. Med. Virol. 83, 2103–2112. http://dx.doi.org/10.1002/jmv.
Robertson, B.H., Margolis, H.S., 2002. Primate hepatitis B viruses – genetic diversity, geography and evolution. Rev. Med. Virol. 12, 133–141. http://dx.doi.org/10.1002/rmv.348.
Salzano, F.M., Sans, M., 2014. Interethnic admixture and the evolution of Latin American populations. Genet. Mol. Biol. 37, 151–170. http://dx.doi.org/10.1590/S1415-47572014000200003.
Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18, 502–504. http://dx.doi.org/10.1093/bioinformatics/18.3.502.
Shi, W., Zhang, Z., Ling, C., Zheng, W., Zhu, C., Carr, M.J., Higgins, D.G., 2013. Hepatitis B virus subgenotyping: history, effects of recombination, misclassifications, and corrections. Infect. Genet. Evol. 16, 355–361. http://dx.doi.org/10.1016/j.meegid.2013.03.021.
Strimmer, K., von Haeseler, A., 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc. Natl. Acad. Sci. U. S. A. 94, 6815–6819. http://dx.doi.org/10.1073/pnas.94.13.6815.
Stuyver, L., De Gendt, S., Van Geyt, C., Zoulim, F., Fried, M., Schinazi, R.F., Rossau, R., 2000. A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. J. Gen. Virol. 81, 67–74.
Sureau, C., Salisse, J., 2013. A conformational heparan sulfate binding site essential to infectivity overlaps with the conserved hepatitis B virus a-determinant. Hepatology 57, 985–994.
Tan, Y., Ding, K., Su, J., Trinh, X., Peng, Z., Gong, Y., Chen, L., Cui, Q., Lei, N., Chen, X., Yu, R., 2012. The naturally occurring YMDD mutation among patients chronically infected HBV and untreated with lamivudine: a systematic review and meta-analysis. PLoS One 7, e32789. http://dx.doi.org/10.1371/journal.pone.0032789.
Tatematsu, K., Tanaka, Y., Kurbanov, F., Sugauchi, F., Mano, S., Maeshiro, T., Nakayoshi, T., Wakuta, M., Miyakawa, Y., Mizokami, M., 2009. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. J. Virol. 83, 10538–10547. http://dx.doi.org/10.1128/JVI.00462-09.
Torres, C., Piñeiro y Leone, F.G., Pezzano, S.C., Mbayed, V.A., Campos, R.H., 2011. New perspectives on the evolutionary history of hepatitis B virus genotype F. Mol. Phylogenet. Evol. 59, 114–122. http://dx.doi.org/10.1016/j.ympev.2011.01.010.
Wang, H., Chien, M., Huang, H., Chang, H., Wu, C., Chen, P.-J., Chang, M., Chen, D.-S., 2010. Distinct hepatitis B virus dynamics in the Immunotolerant and early immunoclearance phases. J. Virol. 84, 3454–3463. http://dx.doi.org/10.1128/JVI.02164-09.
WHO, 2014. Hepatitis B. fact sheet no. 204. [WWW document]. URL. http://www.who.int/mediacentre/factsheets/fs204/en/ (accessed 7.1.15).
Yang, H., Lu, S., Liaw, Y., You, S., Sun, C., Wang, L., Hsiao, C., Chen, P., Chen, D., Chen, C., Group, T.C.-B.C.S.P., 2002. Hepatitis B e antigen and the risk of hepatocellular carcinoma. N. Engl. J. Med. 347, 168–174.

Yousif, M., Kramvis, A., 2013. Genotype D of hepatitis B virus and its subgenotypes: an update. Hepatol. Res. 43, 355–364. http://dx.doi.org/10.1111/j.1872-034X.2012.01090.x.

Yuan, J., Zhou, B., Tanaka, Y., Kurbanov, F., Orito, E., Gong, Z., Xu, L., Lu, J., Jiang, X., Lai, W., Mizokami, M., 2007. Hepatitis B virus (HBV) genotypes/subgenotypes in China : mutations in core promoter and precore/core and their clinical implications. J. Clin. Virol. 39, 87–93. http://dx.doi.org/10.1016/j.jcv.2007.03.005.

Zheng, J., Zeng, Z., Zhang, D., Yu, Y., Wang, F., Pan, C.Q., 2012. Prevalence and significance of hepatitis B reverse transcriptase mutants in different disease stages of untreated patients. Liver Int. 32, 1535–1542. http://dx.doi.org/10.1111/j.1478-3231.2012.02859.x.