Infection, Genetics and Evolution 47 (2017) 140–142

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

New pieces on genetic diversity and evolutionary history of hepatitis B virus: Characterization of the novel subgenotype F6

Dear Editor,

Hepatitis B virus (HBV) is a globally distributed human pathogen that causes acute and chronic infections. HBV shows a high genetic heterogeneity and has been classified into at least eight genotypes (A-H), based on an intergroup nucleotide divergence greater than 7.5% in the complete genome sequences. Moreover, "subgenotypes" are defined as groups characterized by a divergence between 4% to 7.5%, and "clades" are subgroups showing less than 4% nucleotide divergence (Kramvis et al., 2008; Shi et al., 2013).

It is widely accepted that genotype F is native from the Americas and it is believed that it might have been endemic in the continent during the pre-Columbian times. At present, it has been found in native populations from Alaska, Central and South America, and it is the most prevalent in admixed populations with Native American ancestry (Arauz-Ruiz et al., 1997; Campos et al., 2005; Devesa et al., 2008; Huy et al., 2006; Livingston et al., 2007; Piñeiro y Leone et al., 2008). This is one of the most divergent HBV lineages, showing a high intra-genotype diversity. Therefore, it is currently classified into five different subgenotypes: F1 to F5 (Kramvis et al., 2008; Martínez et al., 2014; Shi et al., 2013).

In a previous study published in Infection, Genetic and Evolution, we have proposed a division of subgenotype F4 into two clades: F4a and

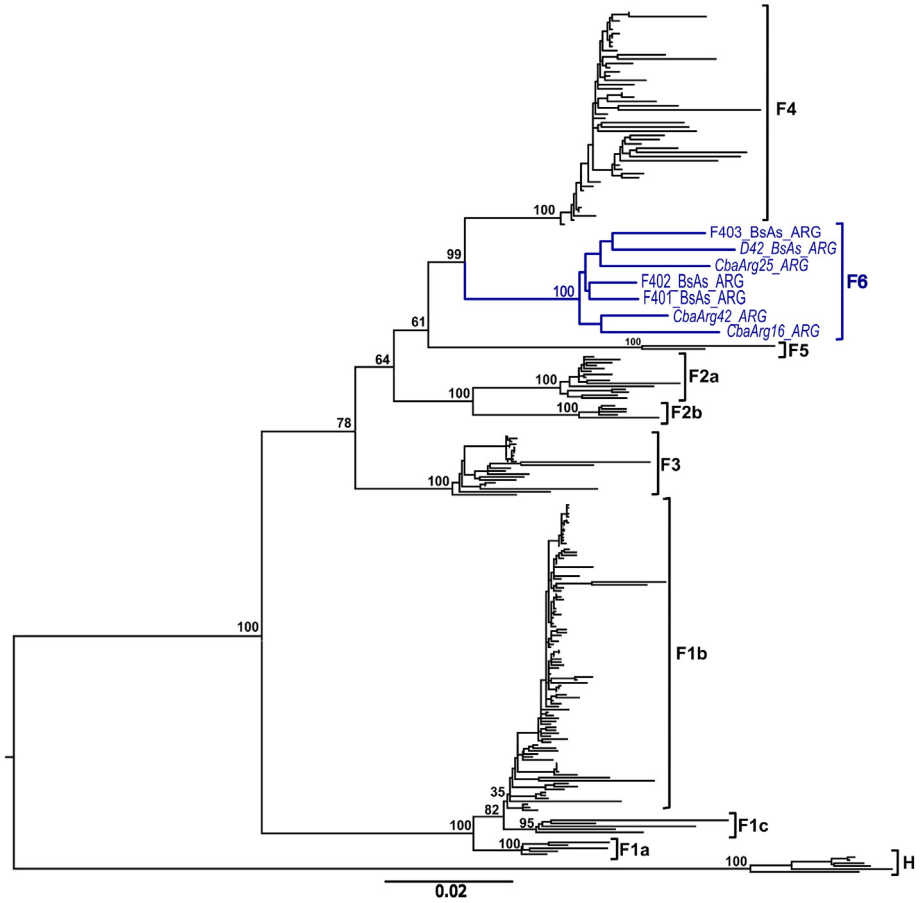

Fig. 1. Phylogenetic analysis of HBV genotype F. Maximum Likelihood tree constructed on complete genome sequences of genotype F ($n = 216$) and genotype H as outgroup ($n = 6$) with RaxML v8.2.8. The new complete genome sequences reported in this work are shown in italics. Statistical support for the main groups is shown in the corresponding nodes (bootstrap values obtained with 1000 pseudoreplicates). The scale bar indicates genetic distances.

http://dx.doi.org/10.1016/j.meegid.2016.11.023

Table 1
Estimates of mean divergence (p-distance, % ± SE) between groups of HBV genotype F. The analysis was conducted in MEGA6 and involved 216 genotype F complete genome sequences (F1 = 109, F2 = 21, F3 = 26, F4 = 51, F5 = 2, F6 = 7). Divergence between F6 and the closest group is shown in bold.

	F1	F2	F3	F4	F5
F2	5.79 ± 0.38				
F3	5.37 ± 0.36	4.39 ± 0.31			
F4	5.69 ± 0.41	4.54 ± 0.28	4.23 ± 0.34		
F5	6.61 ± 0.40	5.39 ± 0.32	5.34 ± 0.37	5.19 ± 0.36	
F6	6.24 ± 0.41	5.14 ± 0.27	4.94 ± 0.35	**4.11 ± 0.27**	5.66 ± 0.37

F4b. The latter was defined based on the report and analysis of three complete genome sequences isolated from patients from the city of Buenos Aires that confo\rmed a separate monophyletic cluster and showed 3.8% divergence in genomic sequence with the previously known F4 strains (Mojsiejczuk et al., 2015). While these features allowed establishing a new clade within the genotype F, the genetic distance value is close to the lower limit conventionally accepted to consider a group as a new subgenotype (4.0%). Recently, four new strains from Argentina that grouped with the previously reported F4b sequences were detected. One of them was obtained from a patient with chronic HBV infection resident in the city of Buenos Aires. The other three additional sequences were identified in patients from the province of Córdoba, in the central region of the country, and corresponded to partial genomic sequences previously reported by Gallego et al. (2014).

Hence, we characterized these four strains and analyzed them in the context of the genotype F evolutionary history. Complete genome sequences were obtained by PCR and direct sequencing, and recombination signals were discarded with RDP4 v4.71 software package (Martin et al., 2015). Then, a Maximum Likelihood phylogenetic analysis and genetic distances calculation were conducted as previously described (Mojsiejczuk et al., 2015). Datasets were updated for genotype F complete genomes newly reported in GenBank, including Panamanian sequences from the recently described groups: F1c and F5 (Martínez et al., 2014).

The phylogenetic tree showed that the four new complete genomes (D42_BsAs, CbaArg16, CbaArg25, CbaArg42) clustered together with the previous named F4b sequences with maximal bootstrap support (Fig. 1). Subsequently, the mean nucleotide divergence (mean ± standard error) between groups was calculated. After the addition of the four novel strains, the nucleotide divergence over the complete genome sequence between "F4a" and the previously named "F4b" cluster was 4.11 ± 0.27 (Table 1). The highly supported monophyletic group and the genetic distance over 4% suggest that these two lineages should be classified as separate subgenotypes (hereafter subgenotypes F4 and F6). Moreover, the novel subgenotype F6, formed by the four novel strains and the former "F4b" sequences, had distinctive nucleotide and amino acid signature patterns (Supplementary Fig. Fig. S1).

To date, several recommendations were introduced for an accurate designation of HBV genotypes and subgenotypes, including monophyletic origin, adequate bootstrap support and minimum sequence divergences among the most relevant (Kramvis et al., 2008; Pourkarim et al., 2010). Even the existence of specific nucleotide and amino acid patterns and the minimum sequence number have been discussed to suggest the existence of a new group (Pourkarim et al., 2010). Notwithstanding, the

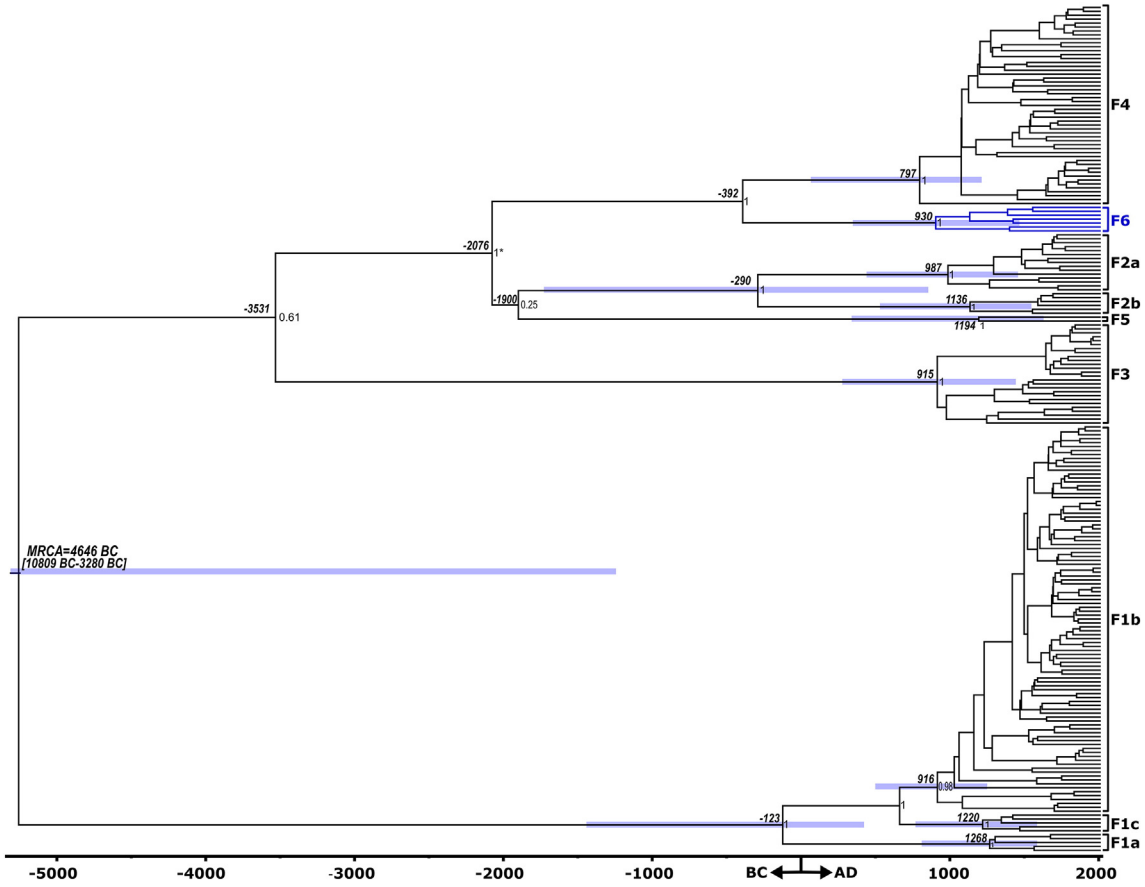

Fig. 2. Time-scaled phylogeny of HBV genotype F. Reconstructed on complete genome sequences of genotype F ($n = 216$) in BEAST v1.8.2. The analysis was performed under the Bayesian skyline demographic model and a relaxed clock model (UCLN) by calibration with a mean fixed on 1×10^{-5} s/s/y. Maximum clade credibility tree was summarized using the "common ancestor heights" algorithm. Posterior clade probabilities of main groups are shown on the nodes (*constrained group). For selected nodes, the ancestor date (based on tMRCA median value) are displayed and high posterior density interval of the nodal dates are represented by the blue bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

suitability of an "appropriate sampling" cannot be judged *a priori* and there is always the possibility that the incorporation of new sequences may change the hierarchical status of the group. Thereby, increasing the amount of data from different geographical locations might be particularly relevant in the case of genotype F, considering that it shows a high intragroup divergence and probably it evolved in isolated native populations scattered throughout the continent. In this context, a permanent review of classifications will keep being of utmost importance in the future.

Additionally, a Bayesian coalescent analysis was performed in order to estimate the time to the most recent common ancestor (tMRCA) for F6 and to reevaluate the genotype F origin and divergence times. The dataset included all non-recombinant American genotype F complete genome sequences available at GenBank ($n = 216$). We previously reported that Maximum Likelihood and Bayesian coalescent analyses of genotype F suggest dissimilar hypotheses about subgenotypes diversification, being the first better supported (Mojsiejczuk et al., 2015). Accordingly, topological constraints were applied to the Bayesian analysis, to force the best supported hypothesis, grouping subgenotypes F2, F4, F5 and F6. A more detailed description of analytical conditions is provided in Fig. 2 legend. The analysis showed that the starting point of F6 diversification dates from approximately 930 CE [HPD95% = 251–1466 CE] and lasted until around 1600 CE. The MRCAs for each subgenotype, including F6, dates from 1250 to 700 years ago, suggesting that the expansion of genotype F strains that are currently circulating mainly took place during the post-Columbian era (Fig. 2). The genotype F tMRCA was similar to those previously calculated using the same median evolutionary rate, estimating an age of around 6600 years [HPD95% = 3280–12,856] (Mello et al., 2013; Mojsiejczuk et al., 2015; Torres et al., 2011).

In summary, new complete genome sequences of HBV genotype F from Argentina were described and, based on the genetic divergence and phylogenetic evidences, it is proposed to classify these new isolates along with the sequences previously reported as a novel "subgenotype F6". Accordingly, it is considered that the accurate classification of HBV subgenotypes and a more exhaustive sampling might provide new insights for understanding the evolutionary history of HBV genotype F and proposing new hypotheses addressing its diversification process.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.meegid.2016.11.023.

Acknowledgments

This work was supported by grants from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET; PIP2012-11220110100215), and the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT; PICT2011-0705).

The following is the supplementary data related to this article.

References

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.

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.

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.

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.

Huy, T.T.T., Ushijima, H., Sata, T., Abe, K., 2006. Genomic characterization of HBV genotype F in Bolivia: genotype F subgenotypes correlate with geographic distribution and T(1858) variant. Arch. Virol. 151:589–597. http://dx.doi.org/10.1007/s00705-005-0671-1.

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.

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., 2007. 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.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 1:1–5. http://dx.doi.org/10.1093/ve/vev003.

Martínez, A.a., Zaldivar, Y.Y., De Castillo, Z., Ortiz, A.Y., Mendoza, Y., Cristina, J., Pascale, J.M., Jimenez, K., Aguilar, J., Gomez, C., Real, F., 2014. High diversity of hepatitis B virus genotypes in Panamanian blood donors: a molecular analysis of new variants. PLoS One 9, e103545. http://dx.doi.org/10.1371/journal.pone.0103545.

Mello, F.C.a., Araujo, O.C., Lago, B.V., Motta-Castro, A.R.C., Moraes, M.T.B., Gomes, S.A., Bello, G., Araujo, N.M., 2013. Phylogeography and evolutionary history of hepatitis B virus genotype F in Brazil. Virol. J. 10:236. http://dx.doi.org/10.1186/1743-422X-10-236.

Mojsiejczuk, L.N., Torres, C., Fainboin, H.A., Galdame, O.A., Campos, R.H., Flichman, D.M., 2015. Identification of a new clade of hepatitis B virus genotype F. Infect. Genet. Evol. 34:122–125. http://dx.doi.org/10.1016/j.meegid.2015.06.007.

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.

Pourkarim, M.R., Amini-Bavil-Olyaee, S., Lemey, P., Maes, P., Van Ranst, M., 2010. Are hepatitis B virus "subgenotypes" defined accurately? J. Clin. Virol. 47:356–360. http://dx.doi.org/10.1016/j.jcv.2010.01.015.

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.

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.

Laura Noelia Mojsiejczuk
Carolina Torres
Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Departamento de Microbiología, Inmunología y Biotecnología, Cátedra de Virología, Junín 956 4to piso, Ciudad Autónoma de Buenos Aires, Argentina.
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

María Belén Pisano
Viviana Re
Universidad Nacional de Córdoba, Facultad de Ciencias Médicas, Instituto de Virología "Dr. Vanella", Enfermera Gordillo Gómez s/n, Ciudad Universitaria, Córdoba, Argentina.
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Rodolfo Héctor Campos
Diego Martín Flichman*
Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Departamento de Microbiología, Inmunología y Biotecnología, Cátedra de Virología, Junín 956 4to piso, Ciudad Autónoma de Buenos Aires, Argentina.
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
*Corresponding author at: Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Departamento de Microbiología, Inmunología y Biotecnología, Cátedra de Virología, Junín 956 4to piso, (C1113AAB), Ciudad Autónoma de Buenos Aires, Argentina.
E-mail address: dflichman@ffyb.uba.ar.

8 September 2016
Available online 26 November 2016