Arch Virol (2006) 151: 589–597
DOI 10.1007/s00705-005-0671-1

Archives of Virology
Printed in Austria

Genomic characterization of HBV genotype F in Bolivia: genotype F subgenotypes correlate with geographic distribution and T^{1858} variant

Brief Report

T. T. T. Huy[1,2], **H. Ushijima**[2], **T. Sata**[1], and **K. Abe**[1]

[1]Department of Pathology, National Institute of Infectious Diseases, Tokyo, Japan
[2]Department of Developmental Medical Sciences, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Received August 15, 2005; accepted September 26, 2005
Published online November 21, 2005

Summary. Hepatitis B virus (HBV) strains were classified into eight genotypes from A to H. Genotype F, an indigenous genotype in Central and South America, has been classified into subgenotypes. An in-depth phylogenetic analysis was performed using two full-length Bolivian HBV sequences and other genotype F strains from the database. A novel nomenclature of subgenotypes of genotype F was proposed, in which Bolivia strains belonged to subgenotype F4. This subgenotype had both Leu^{45} and Ile^{110} in the S gene, and linked to the T^{1858} in the precore. This novel nomenclature demonstrated the relation between variability of the HBV genome and the restricted geographical distribution of the virus in some parts of Central and South America.

*

Hepatitis B virus (HBV) is an important etiologic agent of acute and chronic hepatitis, with over 350 million chronic infected patients around the world [14]. HBV strains are classified into eight genotypes, from A to H [4, 24, 26, 29]. Genotypes A and D are distributed widely in Europe, genotypes B and C are found in East Asia, and genotype E is confined to the sub-Sahara Africa [24]. Genotype G has been sporadically found in North America, Europe and Japan [10, 28–30]. Genotypes F and H, which are genetically divergent from the other

Nucleotide sequence data reported are available in the DDBJ/EMBL/GenBank databases under the accession numbers: AB166850 (HBV-BL592), AB214516 (HBV-BL597).

HBV genotypes, have been detected indigenously in Central and South America [2, 4, 18, 21, 22]. Genotype F strains have been previously classified as four clusters (I–IV) [6, 19], or two clades (clade 1–2) [25], and recently two subgenotypes F1–F2 [12].

Bolivia, a country in the Southern South America, has a low to intermediate prevalence of HBV infection, ranging from 0.4% to 8% [7, 11]. However, there is no report on HBV sequences, as well as the characteristic of HBV genotype F in this area. In this study, we analyzed the pre-S gene and full genome sequences of HBV genotype F from the sera of six HBV infected patients from the Japanese Hospital, Santa Cruz, Bolivia. These sera were obtained from our previous study on the epidemiology of hepatitis virus infections in Bolivia [11]. Six patients were all positive for HBsAg.

HBV DNA was extracted from 100 μl of sera by using the SepaGene RV-R kit (Sanko Junyaku Co., Ltd., Japan). Polymerase chain reaction (PCR) with type-specific primers was used to confirm genotype F of these HBV isolates [20]. The pre-S gene (522 nucleotides, including pre-S1 and pre-S2) and the full-length 3.2 kb of the HBV genome were amplified by using primers and PCR conditions as described [9]. Due to insufficiency of the sera aliquots, only 2 HBV full-length genomes and 4 other pre-S fragments were amplified. PCR products, after purification with the QIAquick gel extraction kit (Qiagen Inc., USA), were subjected to direct sequencing using the ABI PRISM ^{TM} Big Dye Terminator Cycle Sequencing Ready Reaction Kit and automated DNA sequencer ABI 310 (Applied Biosystem, Foster City, USA). Nucleotide sequences and deduced amino acids were multi-aligned [5] with Genetyx for Windows version 6.0 software (Genetyx, Japan). Evolutionary and phylogenetic tree analysis were performed by the MEGA software version 2.1 [13], using the neighbor-joining method, Kimura 2-parameter algorithm with bootstrap 1000 data sets. Genetic distance calculation with standard deviations using bootstrap 1000 data sets were performed by Kimura 2-parameter model (reviewed in Nei and Kumar, 2000 [23]).

Phylogenetic analysis of the pre-S gene of HBV was performed with a total of 51 strains, including 6 Bolivian strains, 25 genotypes F and 3 genotypes H; together with 17 strains of other reference genotypes from GenBank (Table 1). The woolly monkey HBV strain (Accession no. NC_001896) was used as an outgroup. The division of genotype F into clusters demonstrated previously [6, 19] were supported by this pre-S gene tree (data not shown). Clusters I and II possessed strains reported by Mbayed et al. (2001) [19]. The original cluster III of Mbayed et al. [19] was later renamed genotype H by Arauz-Ruiz et al. [4] due to its high divergence. More recently, the designation of cluster III was adopted for another subclade of F strains from Venezuela [27]. Also Devesa et al. [6] adopted this new cluster III which comprised all Venezuela strains in this tree. In the present study, all the pre-S genes of Bolivian HBV strains were located in a cluster comprised of reported strains from Mbayed cluster IV [19].

Full-length genome phylogenetic analysis on these two Bolivian HBV strains, together with 25 genotype F strains from the database (Fig. 1A) confirmed the existence of the previously observed four clusters. Recently, Kramvis et al. [12]

Table 1. HBV genotype F strains from GenBank used for analysis

Isolate names	Accession numbers	Countries	References	Proposed subgenotypes
1024sal	U91831	El Salvador	Arauz-Ruiz et al. (1997) [2]	F1
1073sal	U91830	-	-	-
1116sal	**AY090461**	-	-	-
567teg	U91811	Honduras	-	-
529teg	U91814	-	-	-
553teg	U91812	-	-	-
543teg	U91813	-	-	-
496teg	U91815	-	-	-
8884h	U91805		-	-
3504g	U91818	Guatemala	-	-
70h	**AY090458**	Costa Rica	-	-
7242bsj	U91807	-	-	-
7768h	**AY090459**	-	-	-
1980nic	**AY090456**	Nicaragua	-	-
2091nic	U91822	-	-	-
2187nic	U91821	-	-	-
Sa11	**AF223963**	Argentine	Alestig et al. (2000)*	-
Sa14	**AF223964**	-	-	-
BA10	**AY179735**	-	Pineiro Y Leone et al. (2003) [27]	-
BA40	AF288627	-	Mbayed et al. (2001) [19]	-
BA5	AF043561	-	Mbayed et al. (1998) [18]	-
SC1214	U91803	Costa Rica	Arauz-Ruiz et al. (1997) [2]	F2
1889nic	**AY090455**	Nicaragua	-	-
BD513	AY311347	Venezuela	Devesa et al. (2004) [6]	-
HBVADW4A	**X69798**	Brazil	Naumann et al. (1993) [22]	-
VNZ8248	**AB036905**	Venezuela	Nakano et al. (2001) [21]	F3
VNZ8337	**AB036908**	-	-	-
VNZ8339-1	**AB036909**	-	-	-
VNZ8251	**AB036910**	-	-	-
VNZ8255	**AB036911**	-	-	-
VNZ8323	**AB036912**	-	-	-
VNZ8346	**AB036913**	-	-	-
VNZ8349	**AB036914**	-	-	-
VNZ8351	**AB036915**	-	-	-
VNZ8371	**AB036916**	-	-	-
VNZ8375-1	**AB036917**	-	-	-
VNZ8381	**AB036919**	-	-	-
VNZ8624	**AB036920**	-	-	-
HHVBF	X75663	Colombia	Norder et al. (1994)	-
BA30	AF288628	Argentine	Mbayed et al. (2001) [19]	F4
BA33	AF288624	-	-	-
BA36	AF288626	-	-	-
BA38	AF288625	-	-	-
BA43	AF288623	-	-	-
BA8	AF043573	-	Mbayed et al. (1998) [18]	-
BA9	**AY179734**	-	Pineiro Y Leone et al. (2003) [27]	-
Sa16	**AF223965**	-	Alestig et al. (2000)*	-
Sa4	**AF223962**	-	-	-
HHVBFFou	**X75658**	France	Norder et al. (1994) [24]	-

* Submitted only in database

suggested these clusters or clades of genotype F should fall into two subgenotypes F1, F2. However, when the two complete genomes of Bolivia HBV have become available, the existence of these four clusters was evident with significant bootstrap value (98–100%) in the full-length genome tree. We designated these clusters as subgenotypes of genotype F (F1, F2, F3, and F4). Subgenotype F1 corresponded to Mbayed I; F2 to Mbayed II [19], F3 to Devesa III [6] and F4 to Mbayed IV [19]. Inter-subgenotype nucleotide distances of these proposed groups ranged from 4.3% to 6.1% (Table 2), in which subgenotype F1 was the most distant group (5.5–6.1%). In addition, intra-subgenotype nucleotide distances were much lower (0.7% to 2%).

Due to the limited number in the full-length genotype F strains, phylogenetic analysis over the small S gene was further performed to support the existence of these four subgenotypes. A total of 40 genotype F strains from the database, two

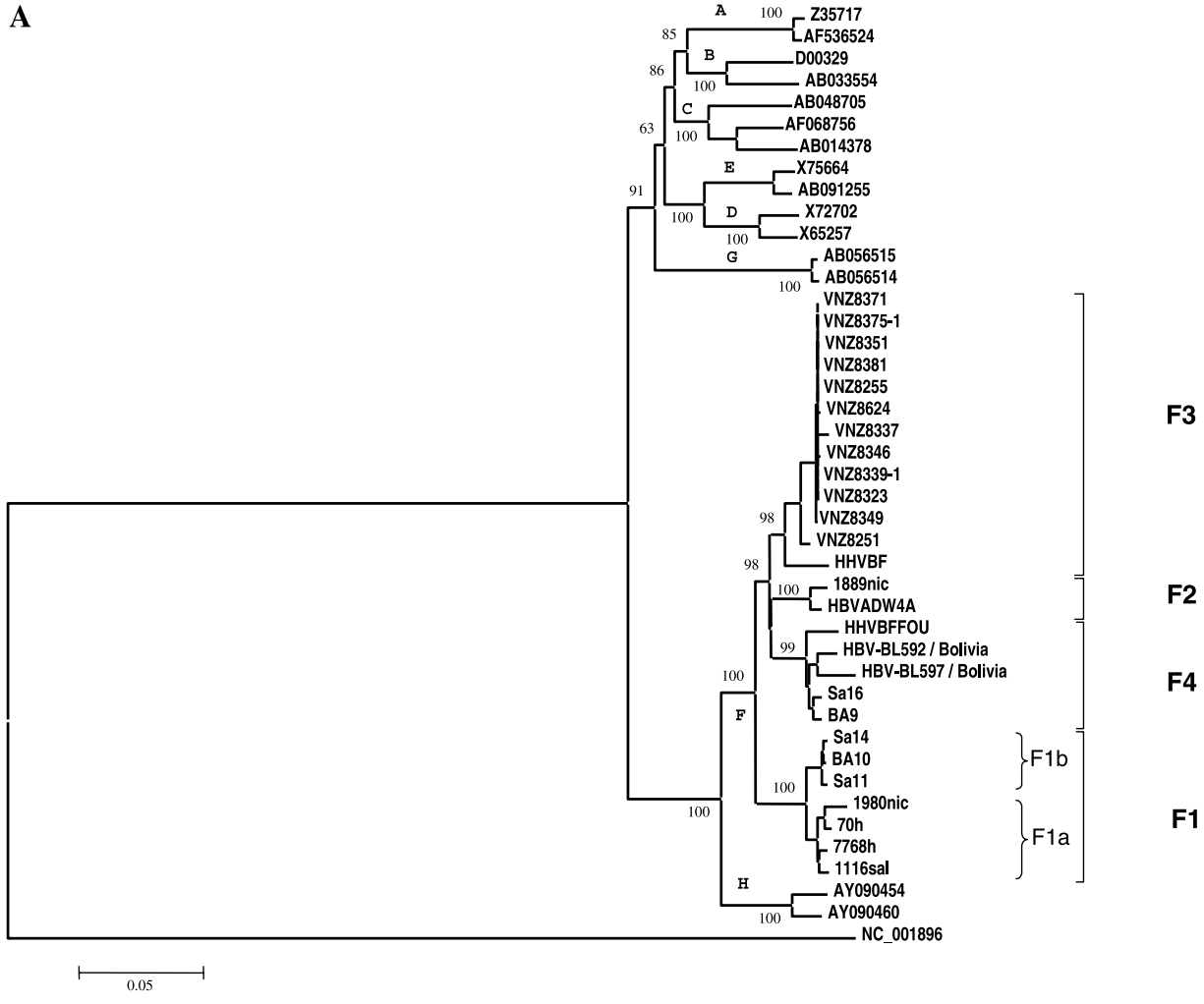


Fig. 1 (*continued*)

Bolivia strains and other reference strains from genotypes A–E, G–H were investigated (Fig. 1B). The phylogenetic tree based on the small S gene (681 nucleotides) also identified the four subgenotypes with high bootstrap values.

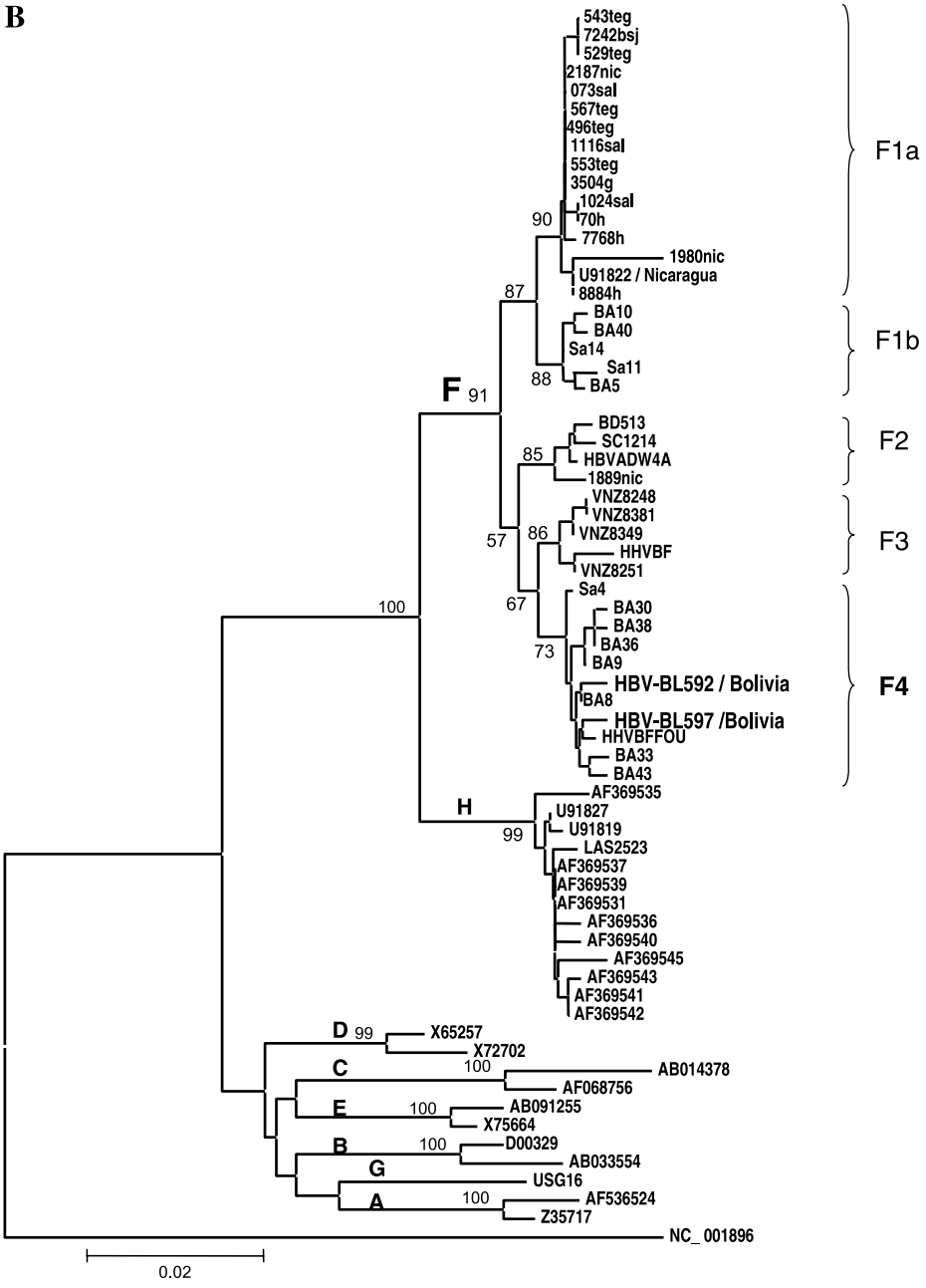


Fig. 1. Phylogenetic tree constructed on the HBV full-length genome (**A**) and the small S gene (**B**). Bootstrap values are shown at the beginning of each main node. The length of the horizontal bar indicates the number of nucleotide substitutions per site. Subgenotype F4 possessed strains from Bolivia, Argentine and France

Table 2. Mean of nucleotide distances among and within subgenotypes of genotype F

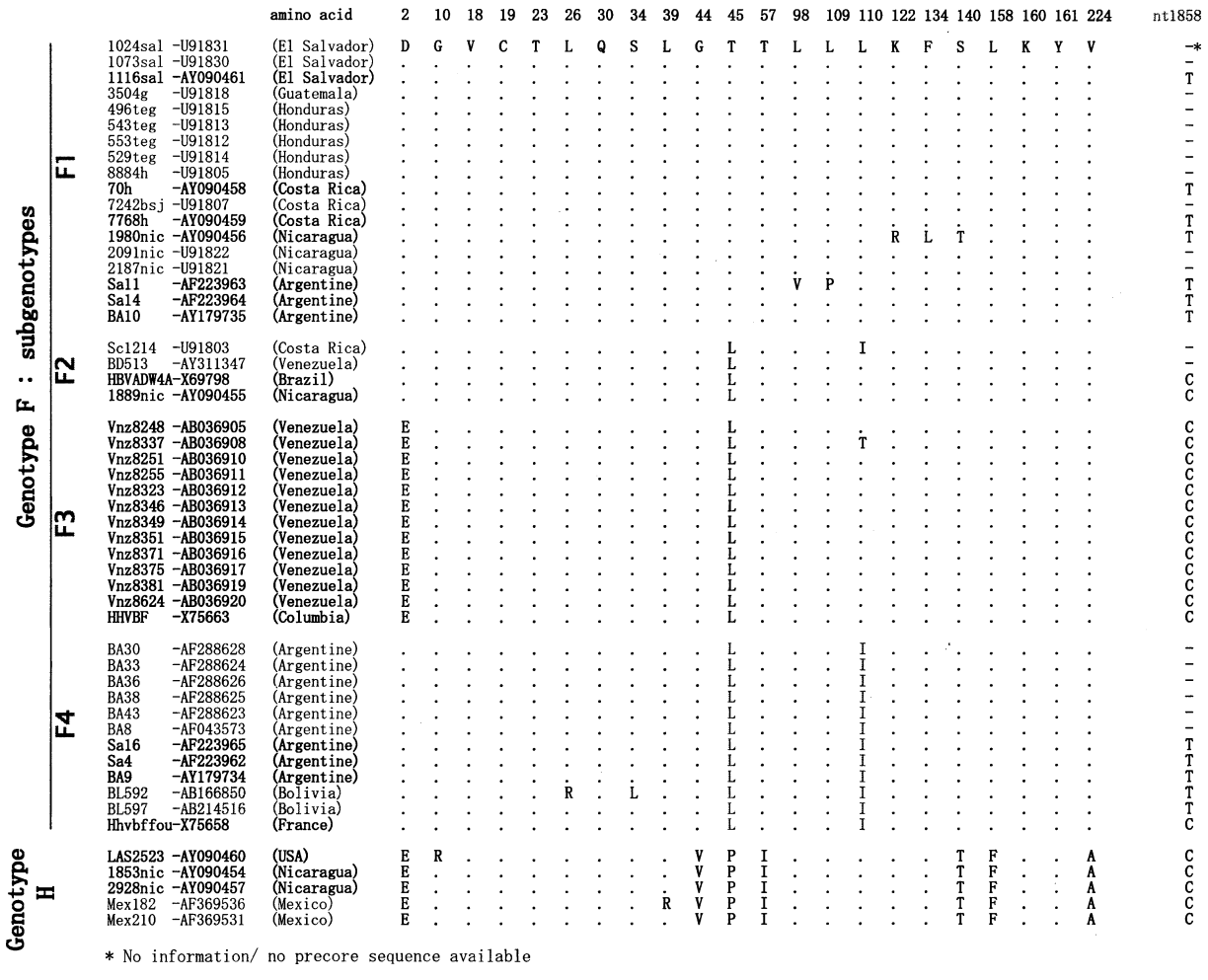
Genotype F*	F1	F2	F3	F4	Genotype H
F1	**1.5 ± 0.1**				
F2	5.5 ± 0.4	**1.1 ± 0.2**			
F3	5.7 ± 0.4	4.3 ± 0.3	**0.7 ± 0.1**		
F4	6.1 ± 0.4	4.6 ± 0.3	4.7 ± 0.3	**2.0 ± 0.2**	
Genotype H	8.6 ± 0.5	7.8 ± 0.4	8.4 ± 0.5	8.8 ± 0.4	**1.9 ± 0.2**

*Complete genome sequences included in this analysis are described in Fig. 1A. Mean ± standard deviations of nucleotide distances were estimated by bootstrapping with 1000 replicates. Intra-subgenotype values are displayed in bold. Nucleotide distances of genotype H are added as a reference

Subgenotype F1 included 2 subclades described as F1a (Mbayed Ia) and F1b (Mbayed Ib) [19]. F1a contained strains from Central America (Honduras, El Salvador, Costa Rica, Nicaragua); F1b comprised strains from Argentine. Subgenotype F2 included strains from Brazil, Venezuela, Nicaragua and Costa Rica. Subgenotype F3 possessed strains from Venezuela and Colombia. Finally, subgenotype F4 contained strains from Bolivia, France and most of reported Argentine isolates. As previously demonstrated [6, 19, 27], subgenotype F1b and F3 exhibited a restricted geographic distribution (Southern and Northern South America, respectively). Subgenotype F1a and F2 were scattered throughout Central and Northern South America; and subgenotype F4 was found in Southern South America (i.e. Bolivia and Argentine) and in one infected person resident in France.

In addition, it became evident that genotype F strains in Central America have T^{1858} and Thr^{45} in the S gene, whereas those derived from South America and Polynesia have C^{1858} and Leu^{45} [12, 25]. Taking into account of these characteristics, multi-alignment of the nucleotides and deduced amino acid were performed with strains from subgenotypes F1–F4, and from genotype H (Fig. 2). Subgenotype F1 possessed Thr^{45} and Leu^{110} in the S gene and T^{1858} in the core region [12, 25], the remaining subgenotypes F2–F4 had Leu^{45} instead of Thr^{45} in the S gene. However, while strains from F2, F3 carried C^{1858} and had the common characteristic of Norder's clade 2; strains from F4 possessed T^{1858} in 5/6 isolates. In fact, subgenotype F4, not only had Leu^{45}, but also possessed Ile^{110} instead of Leu^{110}. This Ile^{110} substitution is conserved in all strains of this subgenotype, and is frequently found in other genotypes A, B, D, E and G [4, 17]. Therefore, the Ile^{110} substitution and the correlation to T^{1858} of the subgenotype F4 keep itself apart from clade 2 of Norder's classification. This further supports for our novel nomenclature of four subgenotypes of genotype F (i.e. F1–F4).

The finding of T^{1858} entity in subgenotype F1 [25] and F4 demonstrated the relation between the variability of the HBV genome, in order to overcome the error threshold for survival, with the geographical distribution of the virus. Most of the genotype F strains from Southern South America (F1b and F4) and Central



* No information/ no precore sequence available

Fig. 2. Multi-alignment of the deduced amino acids from all subgenotypes of genotype F and genotype H strains was performed on the S genes. The characteristic of the nucleotide T^{1858} or C^{1858} in the precore gene of each strain was shown at the end of each alignment. Strains with bold characters had the full-length genome sequences

America (F1a) had T^{1858}. In contrast, strains from Northern South America (F3) carried C^{1858}. Generally, C^{1858} strains found in genotype A, C, F and H, constrains the G to A mutation at position 1896 in the precore gene [1, 15]. In other words, the T^{1858} variant found in F1 and F4 strains might frequently cause the precore stop codon mutation, as previously observed in Central America strains [3]. This observation was supported by recent reports on genotype F strains in Argentine, which showed the high association (50–66.7%) with precore stop codon mutation [8, 16].

In conclusion, the finding of Bolivian HBV genotype F strains, grouped into the subgenotype F4 and linked with T^{1858}, corroborated for our novel nomenclature of subgenotypes of genotype F. The correlation between their geographic distribution and the ability of predisposing to precore stop codon mutation contribute to the

knowledge about the HBV evolution in the New World. It is noteworthy to further investigate whether the variability of genotype F has an affect on the differences in HBV prevalence in this region [7].

Acknowledgments

The authors thank to Dr. Akira Muraoka, Chiaki Myoshi (International Medical Center of Japan, Tokyo, Japan) and Dr. Carlos La Fuente Zerain (Japanese Hospital, Santa Cruz, Bolivia). Dr. T. T. T. Huy was supported by scholarship from Sato International Scholarship Foundation (Japan).

References

1. Alestig E, Hannoun C, Horal P, Lindh M (2001) Phylogenetic origin of hepatitis B virus strains with precore C-1858 variant. J Clin Microbiol 39: 3200–3203
2. Arauz-Ruiz P, Norder H, Visona KA, Magnius LO (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
3. Arauz-Ruiz P, Norder H, Visona KA, Magnius LO (1997) Genotype F prevails in HBV infected patients of hispanic origin in Central America and may carry the precore stop mutant. J Med Virol 51: 305–312
4. Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO (2002) Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol 83: 2059–2073
5. Barton GJ, Sternberg MJ (1987) A strategy for the rapid multiple alignment of protein sequences. Confidence levels from tertiary structure comparisons. J Mol Biol 198: 327–337
6. Devesa M, Rodriguez C, Leon G, Liprandi F, Pujol FH (2004) Clade analysis and surface antigen polymorphism of hepatitis B virus American genotypes. J Med Virol 72: 377–384
7. Echevarria JM, Leon P (2003) Epidemiology of viruses causing chronic hepatitis among populations from the Amazon Basin and related ecosystems. Cad Saude Publica 19: 1583–1591
8. Franca PH, Gonzalez JE, Munne MS, Brandao LH, Gouvea VS, Sablon E, Vanderborght BO (2004) Strong association between genotype F and hepatitis B virus (HBV) e antigen-negative variants among HBV-infected argentinean blood donors. J Clin Microbiol 42: 5015–5021
9. Huy TT, Ushijima H, Quang VX, Win KM, Luengrojanakul P, Kikuchi K, Sata T, Abe K (2004) Genotype C of hepatitis B virus can be classified into at least two subgroups. J Gen Virol 85: 283–292
10. Kato H, Gish RG, Bzowej N, Newsom M, Sugauchi F, Tanaka Y, Kato T, Orito E, Usuda S, Ueda R, Miyakawa Y, Mizokami M (2004) Eight genotypes (A–H) of hepatitis B virus infecting patients from San Francisco and their demographic, clinical, and virological characteristics. J Med Virol 73: 516–521
11. Konomi N, Miyoshi C, La Fuente Zerain C, Li TC, Arakawa Y, Abe K (1999) Epidemiology of hepatitis B, C, E, and G virus infections and molecular analysis of hepatitis G virus isolates in Bolivia. J Clin Microbiol 37: 3291–3295
12. Kramvis A, Kew M, Francois G (2005) Hepatitis B virus genotypes. Vaccine 23: 2409–2423
13. Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17: 1244–1245
14. Lee WM (1997) Hepatitis B virus infection. N Engl J Med 337: 1733–1745

15. Lok AS, Akarca U, Greene S (1994) Mutations in the pre-core region of hepatitis B virus serve to enhance the stability of the secondary structure of the pre-genome encapsidation signal. Proc Natl Acad Sci USA 91: 4077–4081
16. Lopez JL, Mbayed VA, Telenta PF, Gonzalez JE, Campos RH (2002) 'Hbe minus' mutants of hepatitis B virus. Molecular characterization and its relation to viral genotypes. Virus Res 87: 41–49
17. Magnius LO, Norder H (1995) Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene. Intervirology 38: 24–34
18. Mbayed VA, Lopez JL, Telenta PF, Palacios G, Badia I, Ferro A, Galoppo C, Campos R (1998) Distribution of hepatitis B virus genotypes in two different pediatric populations from Argentina. J Clin Microbiol 36: 3362–3365
19. Mbayed VA, Barbini L, Lopez JL, Campos RH (2001) Phylogenetic analysis of the hepatitis B virus (HBV) genotype F including Argentine isolates. Arch Virol 146: 1803–1810
20. Naito H, Hayashi S, Abe K (2001) Rapid and specific genotyping system for hepatitis B virus corresponding to six major genotypes by PCR using type-specific primers. J Clin Microbiol 39: 362–364
21. Nakano T, Lu L, Hu X, Mizokami M, Orito E, Shapiro C, Hadler S, Robertson B (2001) Characterization of hepatitis B virus genotypes among Yucpa Indians in Venezuela. J Gen Virol 82: 359–365
22. Naumann H, Schaefer S, Yoshida CF, Gaspar AM, Repp R, Gerlich WH (1993) Identification of a new hepatitis B virus (HBV) genotype from Brazil that expresses HBV surface antigen subtype adw4. J Gen Virol 74 (Pt 8): 1627–1632
23. Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxford University Press, New York
24. 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
25. Norder H, Arauz-Ruiz P, Blitz L, Pujol FH, Echevarria JM, Magnius LO (2003) The T1858 variant predisposing to the precore stop mutation correlates with one of two major genotype F hepatitis B virus clades. J Gen Virol 84: 2083–2087
26. Okamoto H, Tsuda F, Sakugawa H, Sastrosoewignjo RI, Imai M, Miyakawa Y, Mayumi M (1988) Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. J Gen Virol 69 (Pt 10): 2575–2583
27. Pineiro Y Leone FG, Mbayed VA, Campos RH (2003) Evolutionary history of hepatitis B virus genotype F: an in-depth analysis of Argentine isolates. Virus Genes 27: 103–110
28. Shibayama T, Masuda G, Ajisawa A, Hiruma K, Tsuda F, Nishizawa T, Takahashi M, Okamoto H (2005) Characterization of seven genotypes (A to E, G and H) of hepatitis B virus recovered from Japanese patients infected with human immunodeficiency virus type 1. J Med Virol 76: 24–32
29. 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
30. Vieth S, Manegold C, Drosten C, Nippraschk T, Gunther S (2002) Sequence and phylogenetic analysis of hepatitis B virus genotype G isolated in Germany. Virus Genes 24: 153–156

Author's address: Kenji Abe, D.V.M., Ph.D., Department of Pathology, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan; e-mail: kenjiabe@nih.go.jp