A novel hepatitis B virus (HBV) subgenotype D (D8) strain, resulting from recombination between genotypes D and E, is circulating in Niger along with HBV/E strains

Mariama Abdou Chekaraou,1 Ségolène Brichler,1,2† Waël Mansour,3† Frédéric Le Gal,1 Aminata Garba,4 Paul Dény1,5 and Emmanuel Gordien1,2

1Service de Bactériologie, Virologie, Hygiène, Hôpital Avicenne, Assistance Publique – Hôpitaux de Paris, Laboratoire Associé au Centre National de Référence des Hépatites B, C et Delta, UFR Santé Médecine Biologie Humaine, Université Paris 13, Bobigny, France
2INSERM U845, Université Descartes – Paris 5, Paris, France
3Laboratoire de Bactériologie, Virologie, Hygiène Hospitalière, CHU d’Angers, Angers, France
4Laboratoire de Biologie Médicale, Hôpital National de Niamey, Niamey, Niger
5INSERM U871, Lyon, France

Niger is a west African country that is highly endemic for hepatitis B virus (HBV) infection. The seroprevalence for HBV surface antigen (HBsAg) is about 20%; however, there are no reports on the molecular epidemiology of HBV strains spreading in Niger. In the present study, HBV isolates from the sera of 58 consecutive, asymptomatic, HBsAg-positive blood donors were characterized. Genotype affiliation was determined by amplification, sequencing and phylogenetic analysis of the preS1, polymerase/reverse transcriptase (RT/Pol) and precore (preC)/C regions. The first series of results revealed that different genomic fragments clustered with different genotypes on phylogenetic trees, suggesting recombination events. Twenty-four complete genomic sequences were obtained by amplification and sequencing of seven overlapping regions covering the whole genome, and were studied by extensive phylogenetic analysis. Among them, 20 (83.3%) were classified unequivocally as genotype E (HBV/E). The remaining four (16.7%) clustered on a distinct branch within HBV/D with strong bootstrap and posterior probability values. Complete molecular characterization of these four strains was achieved by the Simplot program, bootscanning analysis and cloning experiments, and enabled us to identify an HBV/D–E recombinant that formed a new HBV/D subgenotype spreading in Niger, tentatively named D8. Moreover, 20 new complete HBV/E nucleotide sequences were determined that exhibited higher genetic variability than is generally described in Africa. One was found to be a recombinant containing HBV/D sequences in the preS2 and RT/Pol regions. Taken together, these data suggest that, in Niger, genetic variability of HBV strains is still evolving, probably reflecting ancient endemic HBV infection.

INTRODUCTION

Hepatitis B virus (HBV) infection remains a major public health problem, particularly in Africa, although data for precisely estimating the burden of HBV infection are lacking. It is estimated that 2 billion people worldwide are or have been infected with HBV (WHO: http://www.who.int/csr/disease/hepatitis/HepatitisB_whocdscsrlyo2002_2.pdf), among whom over 360 million are in a chronic carrier state with a high risk of developing cirrhosis and hepatocellular carcinoma (Ganem & Prince, 2004; Lee et al., 1997; Lok, 2004). About 70–140 million of these carriers live in Africa, and about 250 000 of the 1.3 million HBV-related deaths recorded each year throughout the world occur in Africa (Andernach et al., 2009; Hubschen et al., 2008; Kramvis & Kew, 2007; Kramvis et al., 2002; Mulders et al., 2004). HBV belongs to the family Hepadnaviridae and is characterized by a partially double-stranded circular DNA genome of

†These authors contributed equally to this work.

The GenBank/EMBL/DDBJ accession numbers for the sequences determined in this study are FNS94748–FNS94771.

A supplementary list of additional GenBank sequences used in this study is available with the online version of this paper.
approximately 3.2 kb in length. The HBV genome is generated by reverse transcription from an intermediate 3.5 kb RNA referred to as the pre-genomic RNA. The HBV genome encodes four partially overlapping open reading frames (ORFs): the surface (preS1, preS2, S), core (preC, C), polymerase (Pol) and X genes, respectively. High genetic variability is another feature typical of HBV, related to the absence of proofreading activity of the viral polymerase during the reverse transcription step of genome replication. To date, eight genotypes (A–H) have been established based on intergroup divergence of >7.5% in the complete nucleotide sequence (Arauz-Ruiz et al., 2002; Kramvis et al., 2008; Norder et al., 1994; Okamoto et al., 1988; Stuyver et al., 2000). Recently, a ninth genotype isolated in Laos (Olinger et al., 2008a) and Vietnam (Hannoun et al., 2000; Olinger et al., 2008b; Tran et al., 2008) and tentatively termed ‘I’ was proposed, although it is still subject to debate (Kurbanov, 2008). In addition, a tenth genotype provisionally termed genotype ‘J’ was isolated from a Japanese patient (Tatematsu et al., 2009). Further extensive phylogenetic analyses of HBV genotypes in different studies worldwide have resulted in recognition of several subgenotypes within genotypes, A (A1–A5), B (B1–B8), C (C1–C7), D (D1–D7) and F (F1–F4), defined by >4% intra-genotypic divergence but <7.5% (Cavinta et al., 2009; Hubschen et al., 2009; Kramvis et al., 2008; Lusida et al., 2008; Meldal et al., 2009; Mulyanto et al., 2009; Norder et al., 2004; Nurainy et al., 2008; Schaefer, 2007). Several large-scale studies carried out in a number of African countries have disclosed the emergence of a trend in the distribution of genotypes. HBV/E is the most prevalent genotype in Africa, spreading in a vast crescent, with a span from Senegal to Namibia. In contrast, HBV/A strains are relatively rare, and are found mainly in southern, eastern and central Africa (Hubschen et al., 2009; Kramvis & Kew, 2007; Kramvis et al., 2002; Owiredu et al., 2001), whilst HBV/D is the dominant genotype in northern Africa (Ayed et al., 2007; Bahri et al., 2006; Ezzikouri et al., 2008; Khelifa & Thibault, 2009; Meldal et al., 2009). However, although HBV/E strains are predominant in Africa, they show a low genetic diversity. In contrast, several subgenotypes of genotype A (A1–A5) and a potential new subgenotype identified in Rwanda (Hubschen et al., 2009) have been reported in Africa (reviewed by Andernach et al., 2009; see also references herein). Similarly, three of the seven subgenotypes identified so far for HBV/D (D1, D3 and D7) were isolated in Africa – in Egypt, South Africa and Tunisia, respectively (Kramvis & Kew, 2007; Meldal et al., 2009).

Moreover, in countries in which several genotypes circulate, co-infections have been described and recombination events may occur, leading to the emergence of hybrid strains that can become the dominant subgenotype prevailing in certain geographical regions. For example, a recombinant between HBV/C and HBV/D is the dominant subgenotype circulating in Tibet (Cui et al., 2002). In Africa, such recombination events have also been described, including recombinants A–D (Owiredu et al., 2001), A–E (Garmiri et al., 2009; Kurbanov et al., 2005), D–E–A and the G insertion (Olinger et al., 2006). Very recently, new HBV/D–E recombinants were identified in the Central African Republic, Ireland (in a west African patient), Tunisia and Guinea, with all of them sharing the same recombination profile (Bekondi et al., 2007; Garmiri et al., 2009; Laoli & Crowley, 2008; Meldal et al., 2009). Interestingly, Niger is situated in the Sahara region, between the Maghreb (Algeria, Morocco and Tunisia) in the north, where HBV/D is predominant (Khelifa & Thibault, 2009), and Mali, Burkina Faso, Benin and Nigeria in the west and south, where HBV/E is predominant (Hubschen et al., 2008; Kramvis & Kew, 2007). Two earlier studies showed that Niger is a region highly endemic for HBV infection, with an HBV surface antigen (HBsAg) seroprevalence ranging from 17.6% (Soubiran et al., 1987) to 29.8% (Cenac et al., 1995). In the latter study, this rate reached 73% in patients with hepatic diseases. A more recent survey conducted in the general population gave a rate of 19.2% HBsAg seroprevalence (Mamadou et al., 2006). However, no molecular study on the HBV genome has been performed in Niger. In the present study, we aimed to explore the molecular epidemiology of HBV strains spreading in Niger. We found that, along with HBV/E strains, a new prevalent subgenotype D, provisionally termed D8 and resulting from recombination events between HBV/D and HBV/E, is currently circulating in Niger.

**RESULTS**

**Nucleotide sequences and phylogenetic analyses**

Several nucleotide sequences of HBV strains from 58 serum samples of HBV-infected blood donors in Niger were characterized: 32 in the preS1 region, 27 in the reverse transcriptase region of the polymerase (RT/Pol) and 33 in the preC/C region. Using this first set of sequences, HBV genotyping was performed by phylogenetic analysis by comparison with HBV sequences of genotypes A–H retrieved from GenBank. Almost all strains were classified unequivocally as HBV/E: 28 (87.5%), 23 (85.2%) and 33 (100%) for the preS1, RT/Pol and preC/C regions, respectively. However, four strains, referred to as bne272, bne281, bne367 and bne442, showed a different affiliation on the phylogenetic trees according to the genomic fragments considered. Indeed, in the preS1 and preC/C regions, these four strains clustered on a distinct branch within HBV/E sequences supported by strong bootstrap and Bayesian posterior probability (BPP) values, whilst in RT/Pol, they belonged to genotype D, suggesting recombination events. Sequencing of the entire genome of all these strains was performed and 24 complete genomic sequences were obtained, which were then subjected to extensive phylogenetic analyses. All previous results were confirmed. Indeed, 20 of the 24 Nigerien sequences (83.3%) were distributed within the HBV/E reference sequences group,
close to two sequences from Ghana and one from Benin, whilst the remaining four (bne272, bne281, bne367 and bne442) formed a cluster within either HBV/D or HBV/E, according to the genomic region analysed. When considering the preS1 region and preC/C ORF, bne272, bne281, bne367 and bne442 formed a branch within HBV/E supported by bootstrap and BPP values of 65/80 and 100/100, respectively (Fig. 1a, e). However, when considering the preS/S, Pol and X ORFs and the complete genomic sequence, these four sequences formed a distinct cluster within HBV/D supported by strong bootstrap and BPP values of 96/100, 100/100, 100/100 and 98/100, respectively (Fig. 1b–d, f). It is noteworthy that, in the X ORF tree, bne272, bne281, bne367 and bne442 strains were close to strains belonging to subgenotypes D4 and D7.

Characterization of new prevalent HBV/D–E recombinant strains in Niger

To further characterize these new isolates (bne272, bne281, bne367 and bne442), we compared their full-length sequence with 202 HBV/D published sequences. As summarized in Table 1, these strains belonged to genotype D, with a mean divergence of 5.1 % (ranging from 3.87 to 6.76 %). These values are within the 7.5 % divergence that defines a subgenotype (Kramvis et al., 2008; Norder et al., 2004). The inter-genotype divergence was >10 % except for HBV/E (7.62 %). As expected, because of recombination, these four recombinant HBV/D–E and bne HBV/E strains were close, with the nucleotide divergence ranging from 6.83 to 7.79 %.

We therefore compared our HBV/D–E isolates with the different HBV/D subgenotypes already described, including the recently proposed HBV/D6 and HBV/D7. The inter-subgenotype divergence ranged from 4.30 % (with HDV/D7) to 5.89 % (with HBV/D5) and the intra-subgenotype divergence within these four bne strains ranged from 1.96 to 2.95 %, with a mean value of 2.55 % (Table 1). Taken together, according to the definition of a subgenotype (Kramvis et al., 2008; Norder et al., 2004), these data clearly indicated a potential novel subgenotype of HBV genotype D (Fig. 1f; Table 1).

On the other hand, clustering in different positions according to phylogenetic analyses strongly indicated hybrid strains resulting from recombination events between HBV/D and HBV/E sequences. Thus, Simplot and bootscanning analyses were carried out using full-length reference sequences of genotypes A–H to determine the recombinant sites in these four isolates. As depicted in Fig. 2, these new HBV/D–E recombinant strains, tentatively named HBV/D8, contained two recombinant fragments: one comprising a large part of the preC/C gene (500–700 bp), including (for bne281 and bne442) or not (for bne272 and 367) preC and the 3’ end of the X region, and the other containing the beginning of the preS1 gene region (~250 bp). Four main breakpoints were identified around nt 1600 (for bne281 and bne442) and 1900 (for bne272 and 367), nt 2400, 2800 and 3000 (Fig. 2a–d).

To confirm the recombination event, cloning and sequencing of the preS1 and X–preC regions that encompassed the recombination points were performed. Ten clones for each recombinant region for the four strains were sequenced and phylogenetically analysed. No mixed infection was observed and all sequences of the corresponding region were identical to the initial sequence and clustered together on phylogenetic trees (data not shown).

As for the HBV/E strains, the overall genome size of these recombinant D–E strains was 3212 nt, except for bne272, which was 3253 nt long. Whilst classified in genotype D, they did not have the common 33 nt deletion in the pre-S1 region specific for genotype D. The bne272 strain possessed a 41 bp insertion, just after nt 1826 (Fig. 2c), with the sequence GAAGAGCTCAAGCTTTCGGAAGCTTTGAGACCTCTTTTCTT.

We therefore compared our HBV/D–E isolates with the different HBV/D subgenotypes already described, including the recently proposed HBV/D6 and HBV/D7. The inter-subgenotype divergence ranged from 4.30 % (with HDV/D7) to 5.89 % (with HBV/D5) and the intra-subgenotype divergence within these four bne strains ranged from 1.96 to 2.95 %, with a mean value of 2.55 % (Table 1). Taken together, according to the definition of a subgenotype (Kramvis et al., 2008; Norder et al., 2004), these data clearly indicated a potential novel subgenotype of HBV genotype D (Fig. 1f; Table 1).

On the other hand, clustering in different positions according to phylogenetic analyses strongly indicated hybrid strains resulting from recombination events between HBV/D and HBV/E sequences. Thus, Simplot and bootscanning analyses were carried out using full-length reference sequences of genotypes A–H to determine the recombinant sites in these four isolates. As depicted in Fig. 2, these new HBV/D–E recombinant strains, tentatively named HBV/D8, contained two recombinant fragments: one comprising a large part of the preC/C gene (500–700 bp), including (for bne281 and bne442) or not (for bne272 and 367) preC and the 3’ end of the X region, and the other containing the beginning of the preS1 gene region (~250 bp). Four main breakpoints were identified around nt 1600 (for bne281 and bne442) and 1900 (for bne272 and 367), nt 2400, 2800 and 3000 (Fig. 2a–d).

To confirm the recombination event, cloning and sequencing of the preS1 and X–preC regions that encompassed the recombination points were performed. Ten clones for each recombinant region for the four strains were sequenced and phylogenetically analysed. No mixed infection was observed and all sequences of the corresponding region were identical to the initial sequence and clustered together on phylogenetic trees (data not shown).

As for the HBV/E strains, the overall genome size of these recombinant D–E strains was 3212 nt, except for bne272, which was 3253 nt long. Whilst classified in genotype D, they did not have the common 33 nt deletion in the pre-S1 region specific for genotype D. The bne272 strain possessed a 41 bp insertion, just after nt 1826 (Fig. 2c), with the sequence GAAGAGCTCAAGCTTTCGGAAGCTTTGAGACCTCTTTTCTT. This sequence was composed of an 18 nt motif (underlined) followed by the exactly inverted complementary sequence (bold) of this motif, followed by five Ts. Interestingly, this sequence was inserted between the start codon of the preC gene, just 9 nt after the ATG and the direct repeat 1 (DR1) region corresponding to the 5’ end of the negative strand of the HBV genome. This insertion creates a stop codon at position 1836, in the preC gene, leading to the discontinuation of HBV e antigen (HBeAg) synthesis. It should be note that bne272 also had the specific mutations/variants in the basal core promoter (A1762T and G1764A) and preC region (G1896A). This latter mutation was also found in bne281 and bne367 but not in bne442. In addition, this 41 bp insertion led to disruption of the stop codon of the X gene, leading to an aberrant X ORF with 126 additional nucleotides.

In summary, this study characterized a prevalent new subgenotype D, which we propose to name HBV/D8, found in nearly 20 % of our samples, resulting from well-defined recombination events between HBV/D and HBV/E sequences involving the X–preC and preS1 regions.

**Fig. 1.** Phylogenetic analysis of HBV strains isolated from Niger compared with reference sequences of each genotype (A–H). Genetic distances were estimated by the Kimura two-parameter matrix, and phylogenetic trees were constructed by the neighbour-joining method. GenBank accession numbers of all published sequences used, followed by the letter of the corresponding genotype or subgenotype, are indicated. Nigerian sequences are referred to as ‘bne’ followed by the number of the strain. Black circles indicate HBV/E bne strains; the black square corresponds to the recombinant HBV/E–D described in Results. Black stars indicate newly characterized subgenotype D8 strains (shaded). Bootstrap (numerator) and BPP (denominator) values are shown along each main node. The regions of the genome analysed were: (a) pre-S fragment; (b) Pol ORF; (c) pre/S/S ORF; (d) preC/C ORF; (e) X ORF; (f) entire genome. Bars indicate nucleotide substitutions per site.
HBV/E strains are predominant in Niger and show greater genetic variability than generally described in Africa

More than 80% of HBV strains spreading in Niger belonged to genotype E, as reported in western and sub-Saharan Africa. Despite its hyperendemicity, HBV/E is characterized by weak genetic variability throughout Africa, i.e. about 1.75% (Andernach et al., 2009; n=69). Here, we have provided 20 new complete HBV/E sequences. The mean intra-genotype divergence, excluding bne127 (see below), with 99 published HBV/E sequences was 2.55%, ranging from 0.009 to 4.42%, and among HBV/E Nigerien sequences themselves, divergence ranged from 1.34 to 3.80% with a mean value of 2.85%. Compared with the study of Andernach et al. (2009), the genetic variability of Nigerien sequences was significantly higher than other available HBV/E sequences (P<0.004).

Moreover, in this study, another recombinant E–D strain was characterized. The bne127 strain formed a distinct branch within the genotype E group on phylogenetic trees in the preS1 and Pol ORFs, and in complete genome sequences with good bootstrap and BPP values (Fig. 1b, c, f). This recombinant sequence comprises the first 340 nt of the preS2 (nt 98–438) and 495 nt (nt 778–1273) within the 3’ end of the Pol region of the HBV/D genotype (Fig. 2e).

In summary, our data show that HBV/E is the dominant genotype circulating in Niger. HBV/E sequences exhibited higher genetic variability than that generally described in Africa and are still evolving, probably reflecting ancient endemic HBV infection.

**DISCUSSION**

Niger is highly endemic for HBV infection, with an HBsAg seroprevalence of about 20%. We provide here the first molecular study on HBV strains isolated in a cohort of infected Nigerien blood donors. Near the borders (Mali, Burkina Faso, Benin, Guinea and Nigeria) where the majority of the population lives, over 80% of circulating strains in Niger belong, as expected, to the HBV/E genotype. However, sequences from Nigerien HBV strains showed higher variability over the complete genome (3% mean genetic diversity) than that generally found within the HBV/E group (1.73%) (Hubschen et al., 2008). Such diversity has also been described in Benin and is probably related to the large number of infected individuals and to more ancient infection in those countries.

Another important finding of this study was that nearly 20% of the characterized HBV strains were recombinants between HBV/D and HBV/E, with a well-defined pattern. These strains clearly form a new subgenotype within genotype D that we propose to name HBV/D8. Firstly, these strains were isolated in healthy unrelated blood donors. Secondly, the four complete genomic sequences characterized here formed a strongly supported phylogenetic group within HBV/D sequences. Lastly, the mean inter-subgenotype nucleotide divergence compared with 202 HBV/D reference sequences of all described HBV/D subgenotypes was 5.1±0.58%.

Recombination events require co-infection with more than one genotype in the same patient. Interestingly, Niger is located between the Maghreb (Algeria in the north) where HBV/D strains are dominant (Ayed et al., 2007; Bahri et al., 2006; Ezzikouri et al., 2008; Khelifa & Thibault, 2009; Meldal et al., 2009) and tropical sub-Saharan west Africa where HBV/E is prevalent (Andernach et al., 2009; Garmiri et al., 2008; Hubschen et al., 2008). Moreover, Niger is also historically a pastoral nomad society, with records of several population migrations, which may well account for the spread of HBV strains of different genotypes in the Nigerian population, leading to the emergence of recombinant strains. Indeed, further analysis of the molecular organization of these recombinant strains disclosed common identical points of recombination: (i) in X–preC (nt 1600–2000); (ii) at the 3’ end of the core gene, around nt 2400; (iii) in the Pol–preS1 region, around nt 2800; and (iv) at the end of the preS1 region, around nt 3000. Interestingly, these breakpoints were located in previously described hot-spot recombination and integration regions of the HBV genome (Dejean et al., 1984; Nagaya et al., 1987; Pineau et al., 1998). The so-called cohesive region of the genome, comprising the direct repeat regions (DR1 and DR2), was often involved in this process. Similarly, Hino et al. (1991) showed that this same HBV DNA region

<table>
<thead>
<tr>
<th>Genotype/subgenotype</th>
<th>n</th>
<th>Divergence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Genotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>10</td>
<td>11.16</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>11.94</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>11.48</td>
</tr>
<tr>
<td>D</td>
<td>202</td>
<td>5.1</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>7.62</td>
</tr>
<tr>
<td>F</td>
<td>10</td>
<td>14.75</td>
</tr>
<tr>
<td>G</td>
<td>4</td>
<td>12.06</td>
</tr>
<tr>
<td>H</td>
<td>4</td>
<td>15.0</td>
</tr>
<tr>
<td>Subgenotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>79</td>
<td>5.06</td>
</tr>
<tr>
<td>D2</td>
<td>47</td>
<td>5.14</td>
</tr>
<tr>
<td>D3</td>
<td>62</td>
<td>5.54</td>
</tr>
<tr>
<td>D4</td>
<td>5</td>
<td>4.40</td>
</tr>
<tr>
<td>D5</td>
<td>2</td>
<td>5.89</td>
</tr>
<tr>
<td>D6</td>
<td>3</td>
<td>5.37</td>
</tr>
<tr>
<td>D7</td>
<td>4</td>
<td>4.30</td>
</tr>
<tr>
<td>D8</td>
<td>4</td>
<td>2.55</td>
</tr>
</tbody>
</table>
Fig. 2. Nucleotide similarity comparisons with consensus sequences representing each genotype (A–H) of the four recombinant HBV/D–E isolates (HBV/D8), bne281 (a), bne442 (b), bne272 (c) and bne367 (d), and of the recombinant HBV/E–D, referred to as bne127 (e), using Simplot and bootscanning analysis. Each genotype is represented by a different colour as indicated. Each isolate was compared over the full-length genome using a 200 bp window size, 20 bp step size, 500 bootstrap replicates, gapstrip on and neighbour-joining analysis. Vertical dotted lines show recombination breakpoints. Schematic diagrams of the ORF of the HBV genome and of the recombinant strains are indicated above and below each figure, respectively. bne281 and bne442, and bne272 and bne367 showed a similar recombination profile (HBV/D in pink and HBV/E in blue). The hatched region within bne272 corresponds to the duplicated inverted 41 nt insert (see Results).
covering nt 1855–1915, comprising DR1, was indispensable for enhancement of a recombination assay (Hino et al., 1991). After analysing 99 complete HBV sequences, Morozov et al. (2000) described homologous recombination between different genotypes and confirmed similar regularity in the distribution of recombinant sites in the vicinity of DR1 and at the 3’ end of the core gene.

Elsewhere in Africa, several hybrid strains have been described as involving HBV/E and genotypes A, D and G (Bekondi et al., 2007; Garmiri et al., 2009; Kurbanov et al., 2005; Laoi & Crowley, 2008; Mulders et al., 2004; Olinger et al., 2006). Molecular analyses often found these same recombinant sites. It is noteworthy that these hot-spot sites are also involved in virus–virus and virus–cell DNA integration events in hepatocellular carcinoma (HCC), linked to HBV. In this setting, in the bne272 strain, we found a 41 nt insert composed in part of an 18 nt motif followed by the exact inverted complementary sequence followed by five Ts. Such an inverted duplication has already been described and was associated with integrated HBV in HCC (Tokino et al., 1987). Whether or not such a strain might have a higher capacity for integration into the genome of the host cell remains to be explored.

Taken together, these data and our findings in HBV sequences in Niger strongly suggest that these new recombinant HBV/D–E strains circulate and are transmitted within the population. However, larger-scale surveys are necessary in Niger and in the Saharan area to confirm the spread and high (17 %) prevalence of this new HBV subgenotype, HBV/D8. Similarly, studies in chronically infected patients will need to be conducted to explore a potential association with HCC occurrence in these patients.

In this study, we also described one HBV/E sequence (bne127) that formed a strong distinct branch with 100 % bootstrap and BPP within genotype E, over the preS/S and Pol regions and the complete genome sequence (Fig. 1b, c, f). The breakpoint sites were located within the preS2–S and 3’-end Pol regions (Fig. 2e). Depending on the variability of Nigerien HBV/E strains, a study of the overall prevalence of this HBV/E–D recombinant might be of interest. Surprisingly, whilst recombinant HBV/D–E strains were found at a rather high prevalence in our sampling, no HBV/D strain was identified in Niger. This could be due to the small size of our cohort and to the greater severity of the disease linked to HBV/D (Ganne-Carrié et al., 2006; Schaefer, 2007). Indeed, our cohort was composed of healthy blood donors. Further studies in larger cohorts are needed to fully investigate the genetic variability of HBV in Africa.

Finally, this study highlights the question of the genetic variability and classification of HBV genetic groups (Bollyky et al., 1996; Simmonds & Midgley, 2005). Homologous or heterologous recombination, as well as the replication process of the HBV genome involving an RNA polymerase lacking a proofreading function, along with immune pressure on the scale of individuals or in ethnic groups, may account for this diversity. The molecular epidemiology of HBV strains in Niger, with the emergence of new prevalent HBV subgenotype D8, strongly indicates that HBV is still evolving genetically.

Table 2. Primers used for DNA amplification and sequencing

<table>
<thead>
<tr>
<th>Fragment/primers</th>
<th>Nucleotide sequence (5’→3’)</th>
<th>Position (nt)</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>PreS1</td>
<td></td>
<td>2817–2839</td>
<td>Sense</td>
</tr>
<tr>
<td>P1S</td>
<td>TCACCATATTCTTGGGAAACAAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2AS</td>
<td>TTCCGAACTGGAGCCACCA</td>
<td>80–61</td>
<td>Antisense</td>
</tr>
<tr>
<td>PreS–Pol</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P8S</td>
<td>TGAGAGACACTCATCCTCA</td>
<td>3179–3197</td>
<td>Sense</td>
</tr>
<tr>
<td>P6AS</td>
<td>CTTGAACARGAGTGTGCA</td>
<td>543–524</td>
<td>Antisense</td>
</tr>
<tr>
<td>RT/Pol</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pol1b</td>
<td>GGACTATCAAGGTATGTT</td>
<td>448–465</td>
<td>Sense</td>
</tr>
<tr>
<td>Pol2</td>
<td>TAACCACATCTTTCTTGTGTT</td>
<td>863–844</td>
<td>Antisense</td>
</tr>
<tr>
<td>Pol–X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P7S</td>
<td>TATATGATGATGTGGTATTGGG</td>
<td>736–758</td>
<td>Sense</td>
</tr>
<tr>
<td>P9AS</td>
<td>GCAGATCCAGTTGGCAG</td>
<td>1408–1391</td>
<td>Antisense</td>
</tr>
<tr>
<td>X–preC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P10S</td>
<td>GATCCATAGCGGGGAACTCC</td>
<td>1263–1282</td>
<td>Sense</td>
</tr>
<tr>
<td>P5AS</td>
<td>CCTCAAGTGCGTGCTGAGC</td>
<td>1701–1682</td>
<td>Antisense</td>
</tr>
<tr>
<td>PreC/C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HB7ES</td>
<td>TGAGAGACACTCGGGAAAGC</td>
<td>1609–1629</td>
<td>Sense</td>
</tr>
<tr>
<td>HB7R2d</td>
<td>CCTGAGTGCGGATGCTGAGG</td>
<td>2068–2084</td>
<td>Antisense</td>
</tr>
<tr>
<td>PreC–preS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4S</td>
<td>TAGATCCGCGCTCACGCT</td>
<td>1992–2009</td>
<td>Sense</td>
</tr>
<tr>
<td>P3AS</td>
<td>TTGGTGAATGATTCTTCCC</td>
<td>2899–2881</td>
<td>Antisense</td>
</tr>
</tbody>
</table>
METHODS

Samples. Sera were collected from 58 consecutive, asymptomatic, unrelated, HBsAg-positive blood donors from Niamey in Niger who gave informed written consent. Samples were stored at −20 °C until assayed.

DNA extraction, amplification and sequencing. DNA extraction was performed using a QIAmp MinElute Virus Vacuum kit (Qiagen) according to the manufacturer’s instructions. In a first series of experiments, three regions of the HBV genome were amplified and sequenced: preS1 (nt 2817–80), the reverse transcriptase region of the polymerase (RT/Pol) (nt 448–863) and preC/C (nt 1609–2068). Sequencing of the whole genome was performed by amplification and sequencing of four additional fragments to cover the entire genome: preS–Pol (nt 3179–543), Pol–X (nt 736–1408), X–preC (nt 1263–1701) and preC–preS1 (1992–2899). The different sets of primers used are listed in Table 2. To reduce the risk of failure, the entire HBV genome was amplified first, using primers P1 and P2 as described by Gunther et al. (1995), followed by nested PCR using the above primers. Purified PCR products were sequenced bidirectionally using a BigDye Terminator v. 3.1 kit (Applied Biosystems). After purification, sequences of amplified nucleic acids were determined using an automated DNA sequencer ABI PRISM 3100 Analyzer (Applied Biosystems).

Phylogenetic analyses. The sequences obtained were compared with sequences of the eight HBV genotypes (A–H) retrieved from GenBank. Alignments were carried out using CLUSTAL X software. The Kimura two-parameter model integrated into PAUP* v. 4.0 b 6 software was used to calculate genetic distance and pairwise distance comparisons. Phylogenetic trees were constructed by the neighbour-joining method. GenBank accession numbers for the reference sequences used in the phylogenetic analyses are specified in Fig. 1. For easier reading of phylogenetic trees, each accession number is followed by the letter and number of the corresponding genotype and subgenotype. To confirm the reliability of phylogenetic tree topologies, bootstrap reconstruction was carried out 10 000 times. Bayesian inference (MrBayes software v. 3.1.2) was also used to reconstruct the phylogenetic trees and to attribute a clade credibility of the branching patterns. For each set, two parallel Markov chain Monte Carlo searches of four classes, including one cold chain with 106 generations and a heated chain of 106 generations were run for 2 × 106 generations using a generalized time reversible (GTR) model of evolution with a gamma distribution of variable sites of four classes (alpha parameter=0.5). To focus on the results reaching apparent stationarity phase, the first 25 % of the runs were discarded and the remaining tree dataset was computed using strict consensus and majority rule algorithms to evaluate the posterior probabilities of the branching pattern (Altekar et al., 2004; Huelsenbeck & Ronquist, 2001).

Recombination investigation. Recombination events for sequences with conflicting phylogenetic positions were searched for using the Simplot program and bootscanning analysis (Lole et al., 1999). Briefly, each sequence was compared with a consensus sequence of each HBV genotype (A–H) in order to identify the breakpoints. The analysis was carried out using a window size of 200 bp, a step size of 20 bp, 500 bootstrap replicates, gapstrip on and neighbour-joining analysis.

Furthermore, cloning experiments were performed to confirm breakpoint regions in the intergenotypic recombinants. The PCR products of the preS1 and X–preC regions were cloned into the TOPO vector (Invitrogen) according to the manufacturer’s instructions. White colonies were picked and grown in Luria–Bertani medium with ampicillin (100 μg ml−1). For each recombinant strain, ten clones were sequenced using M13 universal primers.

ACKNOWLEDGEMENTS

We wish to thank the staff of the National Blood Transfusion Center in Niamey, Niger, for collecting and testing donor samples. We thank Dr Olivier Lada for technical help in sequencing and cloning experiments. We thank Professor Françoise Lunel-Fabiani for critical reading of this paper. We thank Andréa Gordion for help in redaction of the manuscript. M. A. C. received financial support from the ANRS (Agence Nationale de la Recherche contre le SIDA et les Hépatites Virales) and the INVS (Institut National de Veille Sanitaire) via the University Laboratory of Virology of Avicenne Hospital, associated with the ‘Centre National de Référence des Hépatites B, C et Delta’. P. D. is a recipient of a grant from INSERM (Institut National de la Santé et de la Recherche Médicale) attributed to the University Laboratory of Virology of Avicenne Hospital.

REFERENCES


