Complete-genome analysis of hepatitis B virus from wild-born chimpanzees in central Africa demonstrates a strain-specific geographical cluster

Maria Makuwa,1 Sandrine Souquière,1 Olivier Bourry,1 Pierre Rouquet,1 Paul Telfer,1 Philippe Mauclevre,2 Mirdad Kazanji,1 Pierre Roques31 and François Simon4

1Centre International de Recherches Médicales de Franceville, Unité de Rétrovirologie et Centre de Primatologie, Franceville, Gabon
2Institut Pasteur, Dakar, Sénégal
3Service de Neurologie, Commissariat Energie Atomique, Fontenay aux Roses, France
4Hôpital St Louis, Service de Microbiologie, Paris, France

In order to determine whether geographical or species clustering accounts for the distribution of hepatitis B virus (HBV) in subspecies of chimpanzees in Africa, four complete chimpanzee HBV (ChHBV) genome sequences were obtained from eight hepatitis B surface antigen-positive wild-born chimpanzees from Cameroon, Republic of Congo and Gabon. The serological profiles of these chimpanzees corresponded to the acute or chronic highly replicative phase of HBV infection, as confirmed by high plasma HBV loads. Analysis of the sequence alignment of 256 aa (S region) from the eight HBV-infected chimpanzees allowed us to determine the HBV amino acid patterns specific to each chimpanzee subspecies and to their geographical origin. Phylogenetic analysis of both the S region and the complete genome confirmed this distinctive clustering of eight novel ChHBV strains within Pan troglodytes. The strong phylogenetic associations of ChHBV sequences with both chimpanzee subspecies and their geographical origin were therefore confirmed.

Hepatitis B virus (HBV), a member of the family Hepadnaviridae, is found in several species of mammal, such as woodchuck (Marmota monax) and ground squirrel (Spermophilus beecheyi, Spermophilus parryii), in birds, such as duck (Anas domesticus) and grey heron (Ardea cinerea) (Summers et al., 1978; Marion et al., 1980; Mason et al., 1980), and it has also been identified in various species of non-human primates, particularly apes. A novel hepadnavirus related distantly to human HBV and chimpanzee HBV (ChHBV) strains was isolated from a woolly monkey (Lagothrix lagotricha), a New World primate, and phylogenetic analysis indicated that it is probably the progenitor of human HBV (Lanford et al., 1998). ChHBV is known to infect chimpanzees in west Africa (Pan troglodytes verus) (MacDonald et al., 2000), west central Africa (P. troglodytes troglodytes and P. troglodytes vellerosus) (Hu et al., 2001; Takahashi et al., 2001) and east Africa (P. troglodytes schweinfurthii) (Vartanian et al., 2002). Only one western lowland gorilla (Gorilla gorilla gorilla) from Cameroon has been reported to be infected, showing a particular strain position within the ChHBV clade (Grethe et al., 2000). The prevalence of HBV-like viruses in wild apes is unknown. Only one case of ChHBV infection has been identified, in a faecal sample from a wild chimpanzee in Gabon (Makuwa et al., 2005). It has been proposed that geographical rather than (sub)species clustering would account for the distribution of ChHBV variants in different subspecies of chimpanzee in Africa (Starkman et al., 2003).

We recently evaluated the occurrence of HBV markers in 568 plasma samples collected from wild-born non-human primates (Cercopithecidae and African great apes) in central Africa. None of the Cercopithecidae tested had HBV, whereas serological markers of HBV infection were present in chimpanzees and gorillas (Makuwa et al., 2003). Data on the genomic structure and genetic diversity of strains in these apes, however, were not available. In the study reported here, we present phylogenetic analyses of complete ChHBV genome sequences and amino acid analysis of partial (S region) sequences from HBsAg-positive wild-born chimpanzees from Cameroon, Republic of Congo and Gabon.
As shown in Table 1, the chimpanzees studied originated from Cameroon (n = 1), Congo (n = 4) and Gabon (n = 2), plus one chimpanzee of unknown origin housed at our Primatologic Center since the age of 10 years. Seven were of the P. troglodytes troglodytes subspecies and one of the P. troglodytes verus subspecies. Six were male and two were female; the mean age of the studied group was 12.3 years (range, 3–27 years). All harboured hepatitis B surface antigen (HBsAg), with the presence of hepatitis B ‘e’ antigen (HBeAg) in six of them. Presence of these HBV serological markers was demonstrated by ELISA (Monolisa AgHBs Plus, Monolisa HBe Plus; BioRad). ChHBV plasma viral load was quantified in seven animals by Monitor HBV (Roche; limit of detection, 600 copies ml\(^{-1}\)). HBV serological testing could not be performed and HBV viral load could not be determined for the chimpanzee from Cameroon, because of a small amount of available plasma sample. As seen in Table 1, the serological profiles of these chimpanzees corresponded to the acute or chronic highly replicative phase of HBV infection, as confirmed by high plasma viral loads, indicating that the infections constituted persistent hepatitis with no clinical signs.

We isolated viral DNA from clarified plasma (centrifugation at 700 \(g\) for 10 min) and undertook PCR amplification of partial and complete genomes with sequence analysis. Primers were selected from conserved HBV genome regions such that the resulting fragments overlapped adjacent amplicons by at least two-thirds of their length, as described previously (Hu et al., 2000; Makuwa et al., 2003). The PCR-positive products were sequenced directly (Macrogen). A 769 bp fragment of the HBV S gene was amplified for HBV from eight chimpanzees. The four full-length HBV genomes from HBsAg-positive chimpanzees were all 3182 nt in length.

Alignments were carried out for the partial (769 bp) HBV S gene, encompassing the pre-S1, S2 and S domains, and for the complete HBV genome sequences, with CLUSTAL_W (v. 1.7) (Thompson et al., 1994), which compiles neighbour-joining trees under a Kimura two-parameter model (transition/transversion ratio = 2). The alignments obtained by using CLUSTAL_W (v. 1.7) were imported into MEGA (v. 3.1) software (Kumar et al., 2004) to perform a complete phylogenetic analysis with bootstrapping (500 replicates) and amino acid analysis.

Each group of non-human primates appears to have a distinct strain of HBV, as described for humans. Other authors have also found grouping of non-human primate HBV sequences into subspecies or geographically associated clades when partial S gene sequences were analysed (Hu et al., 2001; Robertson & Margolis, 2002; Starkman et al., 2003).

We analysed a sequence alignment of 256 aa (Fig. 1), corresponding to a fragment of hepatitis B surface protein (pre-S1, pre-S2 and S domains) obtained from HBV-positive chimpanzees and gorilla strains, taking Ch-Louise-Cam (GenBank accession no. AY330911) as the reference strain. This analysis allowed us to determine the HBV

### Table 1. Samples collected from wild-born chimpanzees with reference to taxonomic classification, geographical origin, serological and viral HBV statuses and GenBank accession numbers

<table>
<thead>
<tr>
<th>Country</th>
<th>Origin within Africa</th>
<th>Name</th>
<th>Designation</th>
<th>Species</th>
<th>Estimated age (years) at date of sampling</th>
<th>Sex</th>
<th>HBsAg</th>
<th>HBeAg</th>
<th>HBeAb</th>
<th>Viral load (copies ml(^{-1}))</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gabon</td>
<td>East</td>
<td>Makata</td>
<td>Ch-Mak</td>
<td>P. troglodytes troglodytes</td>
<td>19</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>9.0 × 10^6</td>
<td>AM117395*</td>
</tr>
<tr>
<td>Gabon</td>
<td>East</td>
<td>Mangoustan</td>
<td>Ch-Mang</td>
<td>P. troglodytes troglodytes</td>
<td>27</td>
<td>M</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>9.0 × 10^6</td>
<td>AM396403</td>
</tr>
<tr>
<td>Gabon</td>
<td>Unknown</td>
<td>Edgar</td>
<td>Ch-Ed</td>
<td>P. troglodytes verus</td>
<td>27</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>2.2 × 10^4</td>
<td>AM117397*</td>
</tr>
<tr>
<td>Congo</td>
<td>South-west</td>
<td>Jeannette</td>
<td>Ch-Jea</td>
<td>P. troglodytes troglodytes</td>
<td>5</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>28.9 × 10^4</td>
<td>AM396406</td>
</tr>
<tr>
<td>Congo</td>
<td>South-east</td>
<td>Mickey</td>
<td>Ch-Mic</td>
<td>P. troglodytes troglodytes</td>
<td>5</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>73.7 × 10^4</td>
<td>AM396402</td>
</tr>
<tr>
<td>Congo</td>
<td>South-west</td>
<td>Bateko</td>
<td>Ch-Ba</td>
<td>P. troglodytes troglodytes</td>
<td>7</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>5.9 × 10^6</td>
<td>AM117396*</td>
</tr>
<tr>
<td>Congo</td>
<td>South-west</td>
<td>Lucie</td>
<td>Ch-Lu</td>
<td>P. troglodytes troglodytes</td>
<td>3</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>8.5 × 10^5</td>
<td>AM396407</td>
</tr>
<tr>
<td>Cameroon</td>
<td>South-east</td>
<td>Dja</td>
<td>Ch-Dja</td>
<td>P. troglodytes troglodytes</td>
<td>5</td>
<td>M</td>
<td>+</td>
<td>ND</td>
<td>ND</td>
<td></td>
<td>AM117394*</td>
</tr>
</tbody>
</table>

*GenBank accession numbers of complete HBV genomes.
amino acid patterns specific to each chimpanzee subspecies and to confirm the presence of specific amino acid patterns corresponding to their geographical origin.

Within the *P. troglodytes troglodytes* ChHBV group, the ChHBV-Cam strain (Ch-Dja) was related closely to the reference ChHBV strain, and no specific non-synonymous substitutions were observed within the fragment of hepatitis B surface protein analysed. Interestingly, two specific amino acid patterns were found in the pre-S1 domain (E/16/D, A/78/T) for the two ChHBV-Gab strains (Ch-Mang-Gab and Ch-Mak-Gab) and one ChHBV-Congo strain (Ch-Mic-Congo), whilst a unique amino acid substitution (A/78/I) characterized the remaining three ChHBV-Congo strains. No characteristic amino acid substitution patterns were observed in the pre-S2 domain.

Fig. 1. Sequence alignment of 256 aa corresponding to a fragment of the hepatitis B surface protein (pre-S1, pre-S2 and S domains) obtained from HBV-positive chimpanzee and gorilla strains, with Ch-Louisa (GenBank accession no. AY330911) as the reference strain. Novel Ch-HBV strains and corresponding amino acid patterns are shown in bold and complete-genome strains are denoted by an asterisk.
of any of the ChHBV strains analysed. Conversely, a unique specific amino acid substitution (G/181/V) was present in three of four ChHBV strains from Republic of Congo.

With reference to Table 1, the chimpanzee Edgar (Ch-Ed) of unknown origin, housed at the Primate Center in Gabon, belonged to the *P. troglodytes verus* HBV group. Within this group, two characteristic amino acid substitutions were present in the pre-S1 domain (N/40/Y/H and L/56/F) and four in the pre-S2 domain (H/109/Q, V/140/L, I/153/V and R/156/T). The presence of *P. troglodytes verus* in central Africa is uncommon and, according to phylogenetic analysis, this chimpanzee very probably originated from west Africa.

These findings strongly suggest a correlation with geographical origin and a founding effect of these virus variants in the subpopulation of the host.
These findings were also confirmed by phylogenetic analysis of a partial HBV S gene (769 bp) (Fig. 2a). All human HBV sequences cluster within their genomic groups (A–H), as do those of the gibbon and orang-utan HBV strains. Likewise, all HBV strains from <i>P. troglodytes</i> species fell into a single, separate cluster with 99 % bootstrap support, including three chimpanzee subspecies clusters: <i>P. troglodytes troglodytes</i> (57 % bootstrap value), <i>P. troglodytes verus</i> (100 % bootstrap value) and <i>P. troglodytes vellerosus</i> (78 % bootstrap value). A recent phylogenetic reanalysis of the <i>P. troglodytes schweinfurthii</i> HBV strain showed evidence for interspecies recombination between HBV infecting chimpanzees and the human HBV-C genotype strain (Magiorkinis <i>et al.</i>, 2005). This finding of species-specific HBV sequences in Old World great apes suggests species-specific evolution among primates, and the correlation between the mitochondrial DNA sequences of different geographical chimpanzees’ subspecies and chimpanzee HBV sequences supports the species specificity of HBV (Hu <i>et al.</i>, 2001; Robertson & Margolis, 2002).

The results of previous sequence analyses and our phylogenetic analysis confirm the splitting of the <i>P. troglodytes troglodytes</i> cluster into two subclusters with a 60 % bootstrap value. The first subcluster is subdivided into two groups, the ChHBV strains from Gabon (group 1) and Cameroon (group 2), and the second subcluster consists of the three ChHBV strains from Republic of Congo (group 3). The fourth chimpanzee, Mickey (Ch-Mic-Congo), which originated from the south-east of Congo near the Gabonese border (Brazzaville area), harboured an HBV strain that was related more closely to those of chimpanzees from Gabon than to those found in south-west Congo. The estimated inter-group nucleotide divergence in the partial ChHBV S gene (769 bp, groups 1–3) was 1.4 % for groups 1 and 2 and 4.2 % for groups 1 and 3. The estimated intra-group nucleotide divergences were 0.9 % for group 1, 0.4 % for group 2 and 1.1 % for group 3.

Phylogenetic analysis of the complete genome confirmed this distinctive clustering of ChHBV strains within <i>P. troglodytes</i> species (100 % bootstrap value). This cluster was also confirmed by a separate phylogenetic analysis of the partly overlapping open reading frames coding for each gene, with the HBV-A genotype (GenBank accession no. M57663) as the reference strain: C (core protein; bootstrap value 99 %), P (polymerase reverse transcriptase protein Pol; bootstrap value 100 %) and S (envelope proteins S, M and L; bootstrap value 100 %). Even though the bootstrap
value was not particularly strong for the X gene (transcriptional trans-activator protein; bootstrap value <50 %), the novel ChHBV sequences clustered more closely with other viral sequences obtained from the P. troglodytes subspecies (data not shown).

As described previously (Takahashi et al., 2000; Robertson & Margolis, 2002), the amino acid replacements of non-human primate HBV sequences differ somewhat from representatives of the human genotypes. Thus, the human HBV E and F genotypes (African and South American, respectively) are related more closely to non-human HBV isolates than the remaining HBV-A/B/C/D genotypes when compared within the core region sequences. The nucleotide and deduced amino acid sequences in this region are highly conserved in the different genotypes of HBV and other primate and non-primate hepadnaviruses. The human HBV-E/F genotype strains in the core–tail region were thus suggested to be intermediary between HBV-A/B/C/D and the non-primate mammalian hepadnaviruses (Takahashi et al., 2000).

As reported previously (Takahashi et al., 2000), we found an arginine-rich nucelophilic peptide (repeats of an SPRRR motif), common to human HBV-E/F/F genotypes, great apes and woolly monkey HBV strains, that is not exposed to humoral immune pressure. Moreover, we confirmed the presence of the highly conserved nucleotide and deduced amino acid sequences in this region for the different genotypes of HBV and other primate and non-primate hepadnaviruses, when analysing the 3′ end of the core region in our new ChHBV isolates (Fig. 2b).

During HBV screening of chimpanzees from central Africa, including the present phylogenetic analysis, all HBsAg-positive chimpanzees were positive for anti-HBc (hepatitis B core antigen). None of them showed any significant difference in the full-genome nucleotide sequence from those reported so far from primates, including humans, chimpanzees and gorillas, as was described for HBV strain Ch-Bassi (difference of 9–26 %), suggesting a novel strain presumably indigenous to P. troglodytes troglodytes in central Africa (Takahashi et al., 2001). Moreover, all non-human primate hepadnaviruses except for the ‘Bassi’ HBV strain genotype, including those in the present study, have a 36 nt deletion in the core region. As suggested previously, this virus is likely to be a recombinant between a troglodyte-associated HBV variant and another highly divergent HBV variant of unknown origin (Takahashi et al., 2001; Starkman et al., 2003).

In conclusion, we have confirmed the restriction of HBV infection to apes by using the sensitivity/specificity levels of methods presently available to us, and demonstrated a strong phylogenetic association of ChHBV sequences with chimpanzee subspecies and/or their geographical origin.

Acknowledgements
The Centre International de Recherches Médicales Franceville (CIRMF), Franceville, Gabon, is supported by the Gabonese Government, Total-Fina-Elf Gabon and the Ministère des Affaires Etrangères, France. This work was supported by funds from CIRMF and NIH grant R01 A144596 (to Professor P. A. Marx, Tulane University, New Orleans, LA, USA). The authors thank the Habitat Ecologique et Liberté des Primates (HELP) Sanctuary in Conkouati-Douli National Park, Congo, and Dr Marc Ancrenaz for samples from chimpanzees collected in 1992 and 1996, and Mélanie Caron for critical review of the manuscript.

References


