

Analysis of the preS1 gene of hepatitis B virus (HBV) to define epidemiologically linked and un-linked infections in South Africa

D. R. Hardie¹ and C. Williamson^{1,2}

¹Department of Medical Microbiology, University of Cape Town Medical School,
Cape Town, South Africa

²South African Institute for Medical Research, Cape Town, South Africa

Accepted April 9, 1997

Summary. Analysis of the preS1 gene of hepatitis B virus was used to define two nosocomial outbreaks of HBV infection. In an outbreak in an oncology unit we had previously shown by single stranded conformational polymorphism (SSCP) analysis of a 189 bp fragment of the preS1 gene, that 52 children were infected with HBV strains that displayed only 5 different SSCP profiles. Sequencing of a 383 bp fragment encompassing the entire preS1 gene, revealed that isolates with same SSCP profile were identical in sequence across the entire preS1 gene, confirming that those patients with the same SSCP pattern had epidemiologically linked infections. A second outbreak involved 8 liver transplant patients from two different hospitals, 5 of whom were from the same hospital at which the oncology outbreak had occurred. Two of these 5 patients had HBV strains that were identical to strains from the oncology unit and nosocomial transmission probably accounted for the infections in these two, while diversity of both SSCP profiles and sequence data of remaining 6 patients supported the conclusion that they had not been infected from a common source. The donor liver is believed to be the most likely source of infection in these patients. HBV isolates from patients infected in the community were used as a standard for the general degree of preS1 sequence variation of local HBV strains. Phylogenetic analysis and comparison with reference HBV clones revealed that of 27 local HBV strains, genotypes A and D occurred most frequently and were identified in 14 and 12 patients respectively, while genotype C was detected in one patient.

Introduction

Hepatitis B virus (HBV) makes use of overlapping open reading frames to code for its four major proteins. This efficient use of coding sequence puts constraint on the virus's ability to tolerate genomic variation. However, as the virus

replicates via an RNA intermediate, the viral polymerase has reverse transcriptase activity and this process is inherently error prone as the enzyme has no proof reading function [6]. This may account for the genomic diversity of HBV that occurs in different parts of the world.

Historically, HBV variation was defined by means of antigenic analysis of the surface antigen of the virus and on this basis 9 distinct subtypes have been described; namely adw2, adw4, ayw1, ayw2, ayw3, ayw4, ayr, adr^q⁺ and adr^q⁻ [2]. Subsequently, sequence analysis of viral DNA has led to the definition of 6 genotypes (A to F) of HBV, based on the relatedness of whole genomes. However, certain subgenomic regions such as the preS1/preS2 region have also been used in phylogenetic analysis to assign genotypes to isolates with a high degree of confidence [13].

The preS1 gene is the most variable part of the HBV genome and is therefore ideally suited to identify relatedness of HBV isolates that may be epidemiologically linked [18]. Previously we used single stranded conformational polymorphism (SSCP) analysis of a 189 bp fragment from the 3' end of the preS1 gene to analyse an outbreak of HBV in an oncology unit. Based on identity of SSCP profiles, it was shown that 52 oncology patients had been infected with one of 5 different strains of HBV. SSCP is an elegant method for identifying minor differences in the nucleotide sequence of PCR products, but gives no information about the sequence itself, thus the exact degree of relatedness of the SSCP-defined strains was unknown. The present investigation was aimed at assessing the sequence diversity of HBV isolates from those oncology patients infected during the outbreak. In addition, SSCP and sequence analysis was used to investigate the source of HBV infection in eight liver transplant patients who had been infected during the course of their treatment. A 383 bp fragment encompassing the entire preS1 gene was sequenced and subjected to phylogenetic analysis. Sequences of isolates from patients with randomly acquired infections provided a standard of the general degree of divergence of local HBV strains and gave an indication of the prevalent genotypes of HBV in the community.

Materials and methods

Source of HBV DNA

“Community acquired” HBV infections

Blood was obtained from 14 patients who had been infected with HBV in the community. Although all the patients were resident in Cape Town at the time when the blood samples were taken, they were not related and were receiving treatment at different clinics. Their infections were therefore assumed not to be epidemiologically linked. These samples were labelled ZA01 to ZA14. The prefix ZA indicates the country of origin (South Africa).

Oncology patients

A total of 52 children were identified as having been infected nosocomial during an outbreak of HBV infection in an oncology unit. Each had been infected by one of only 5

different strains of HBV, differentiated by SSCP analysis of a 189 bp fragment from the 3' end of the preS1 gene. Three strains predominated and were implicated in infecting 19, 16 and 9 patients respectively [7]. Five examples of each were sequenced. Two additional strains were implicated in infection of 5 and 3 patients, respectively. One of each of these strains was also sequenced.

These samples were labelled ZA23 to ZA38 and the SSCP-defined group to which each sample was assigned, was differentiated by the suffix onc1, onc2, onc3, onc4 or onc5.

Liver transplant patients

A total of 8 patients who had undergone liver transplantation between 1993–95 were found, during post operative screening, to have been infected with HBV. All of the patients were negative for all HBV markers prior to surgery. These patients had been treated at one of two different transplant units in Cape Town. One of these units was at the same hospital as the oncology unit.

These samples were labelled ZA15 to ZA22. The source of these isolates is indicated by the suffix LT.

DNA extraction and PCR

DNA was extracted from 100 µl serum, as described previously [7].

Primers were designed to amplify a fragment encompassing the entire preS1 region of HBV in a semi-nested PCR, incorporating bases 2 829–3 211, numbered as for reference clone pJDW 322 [14].

FP0 5'-TGGGAACAAGAG/TCTAC-3'

RP8 5'-GAACTGGAGCCACCAG-3'

RP7 5'-ACTGCATGGCCTGAGGATGA-3'

Primers FP0 and RP8 were used at a concentration of 10 pmol/µl in the first amplification, and primers FP0 and RP7 were used at a concentration of 40 pmol/µl in the second (nested) amplification. The conditions for PCR were as previously described [7]. PCR products were electrophoresed on 8% polyacrylamide gels, and after staining with ethidium bromide the DNA was visualized by transillumination with ultra-violet light.

Cloning

PCR products were cloned into the *EcoRV* restriction site of Bluescript (SK+) plasmid (Stratagene, USA). DHα competent *E coli* cells were transformed and white colonies were screened for HBV inserts by means of colony blot hybridization, using a digoxigenin-labelled probe (Boehringer-Mannheim, Germany). The probe had the following sequence:

RP4 5'-TCCCTG/AACTGG/CCGATTGGT-3'

Chemiluminescence (CDP star, Boehringer Mannheim) was used to detect specific hybridization. Plasmid DNA was purified by means of nucleobond columns (Macherey Nagel, Germany).

Single strand conformational polymorphism

SSCP was performed on both serum-derived PCR products as well as from cloned material. A 189 bp fragment, starting 117 bases downstream from the ATG at the beginning of the preS1 gene (bases 2965–3153, [14]) was amplified in a semi-nested PCR as previously

described [7]. For SSCP analysis, PCR products were diluted in formamide, heat denatured, snap-cooled and loaded on to 8% polyacrylamide gels, containing 2.5% glycerol. Samples were electrophoresed and DNA was visualised by staining with silver [7].

DNA sequencing and sequence analysis

Plasmid DNA was sequenced by the standard dideoxy chain termination method, either manually (Sequenase version 2.0, United States Biochemical) using ^{35}S dATP, or by means of the automated Alf Express sequencing system. Sequences were initially aligned using Clustal W [17] and final alignments were corrected manually. DNA distance matrices were generated using the Kimura two-parameter model [9] and Clustal W. Neighbour joining phylogenetic trees were constructed using TREECON for Windows 1.0 [20].

Sequences from 39 HBV isolates from Cape Town were compared to 23 reference HBV clones. Sequences were aligned from the ATG at beginning of the preS1 gene to 4 bases past the second ATG (a length of between 331 and 364 bp). Woodchuck hepatitis virus 1 (WHV1) [4] was used as an outgroup. The reliability of branching was assessed by 100 bootstrap analyses.

Results

Correlation between sequence and SSCP

Oncology patients

The five different HBV strains involved in the outbreak in an oncology unit [7] each produced distinctly different SSCP profiles (Fig. 1a), although onc1 and onc2 were similar. Isolates with the same profiles were shown to have complete sequence identity across the 189 bp fragment used for SSCP and in all but two instances were also identical across the entire preS1 region. These two isolates, strains onc2 and onc3, had single base substitutions outside the pre-S1 region. Alignment of the 189 bp sequences from each of the five strains (Fig. 1b) shows that strains onc1, onc2, onc4 and onc5 were similar to one another (differing by 1 to 7 nucleotides), while onc3 was very divergent (differing by up to 32 bases in this region). Single nucleotide changes resulted in different SSCP profiles. The single base difference between strains onc1 and onc2, namely C (onc1) to T (onc2) at position 102, resulted in a slight shift in the mobility of one of the two bands on the gel (Fig. 1a, lanes 1–5 and lanes 6–9). Comparison of the entire preS1 sequence of these two strains revealed no additional base differences. Strains onc4 and onc5 also differed from each other by only one base in the 189 bp region, namely G (onc4) to C (onc5) at position 149, but in this case it resulted in a marked alteration in their SSCP profiles (Fig. 1a, lanes 14 and 15). One further base difference was evident in the preS1 gene, outside the region used for SSCP.

Liver transplant patients

SSCP of HBV isolates from 6 of the 8 liver transplant patients gave unique profiles, while two, namely ZA15.LT and ZA21.LT, were the same (Fig. 2a, lanes 2 and 8). These two strains were identical in sequence across the 189 bp

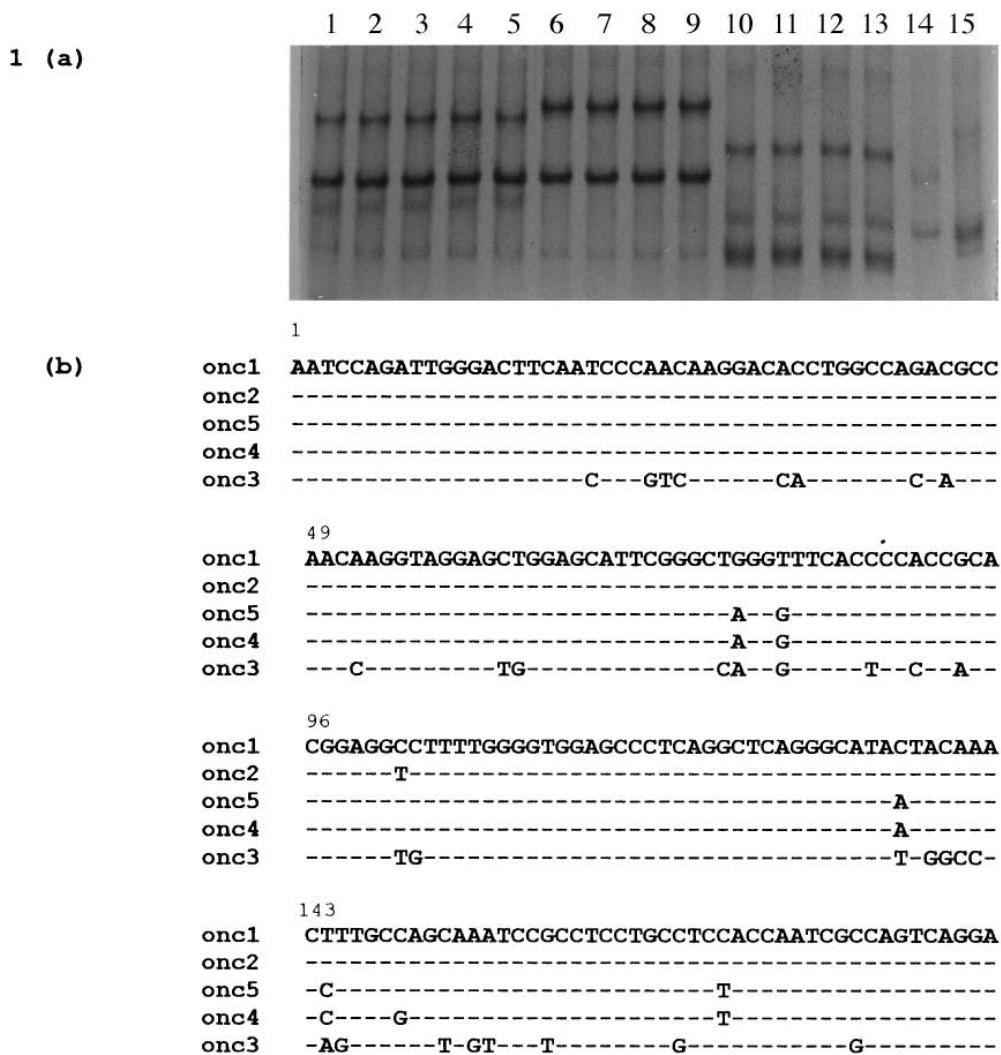


Fig. 1. a Five distinct SSCP profiles were obtained with DNA from patients infected in the oncology unit. 1–5 show profile 1, 6–9 show profile 2, 10–13 show profile 3 and 14 and 15 show profiles 4 and 5 respectively. **b** Sequence alignment of the 189 bp region used for SSCP revealed that strains 1, 2, 4, and 5 were rather similar. Strains 1 and 2 differed by only one base in this region as did 4 and 5. The strain which produced pattern 3 was quite different

region, but in the preS1 gene outside this region, there was one base difference. Alignment of the 189 bp sequence showed that six of the isolates, namely ZA19, ZA20, ZA21, ZA17, ZA15 and ZA22, were similar (differing by a maximum of 13 nucleotides), while the other two sequences, namely ZA16 and ZA18 were more divergent, although similar to each other (differing by 5 nucleotides) (Fig.2b). ZA19.LT and ZA20.LT differed by only one base, namely T (ZA 19.LT) to C (ZA20.LT) at position 152 (as well as 2 further bases outside the region) and ZA17.LT differed from ZA15.LT and ZA21.LT by two bases in the 189 bp region, namely a G at position 18 and a T at position at 140.

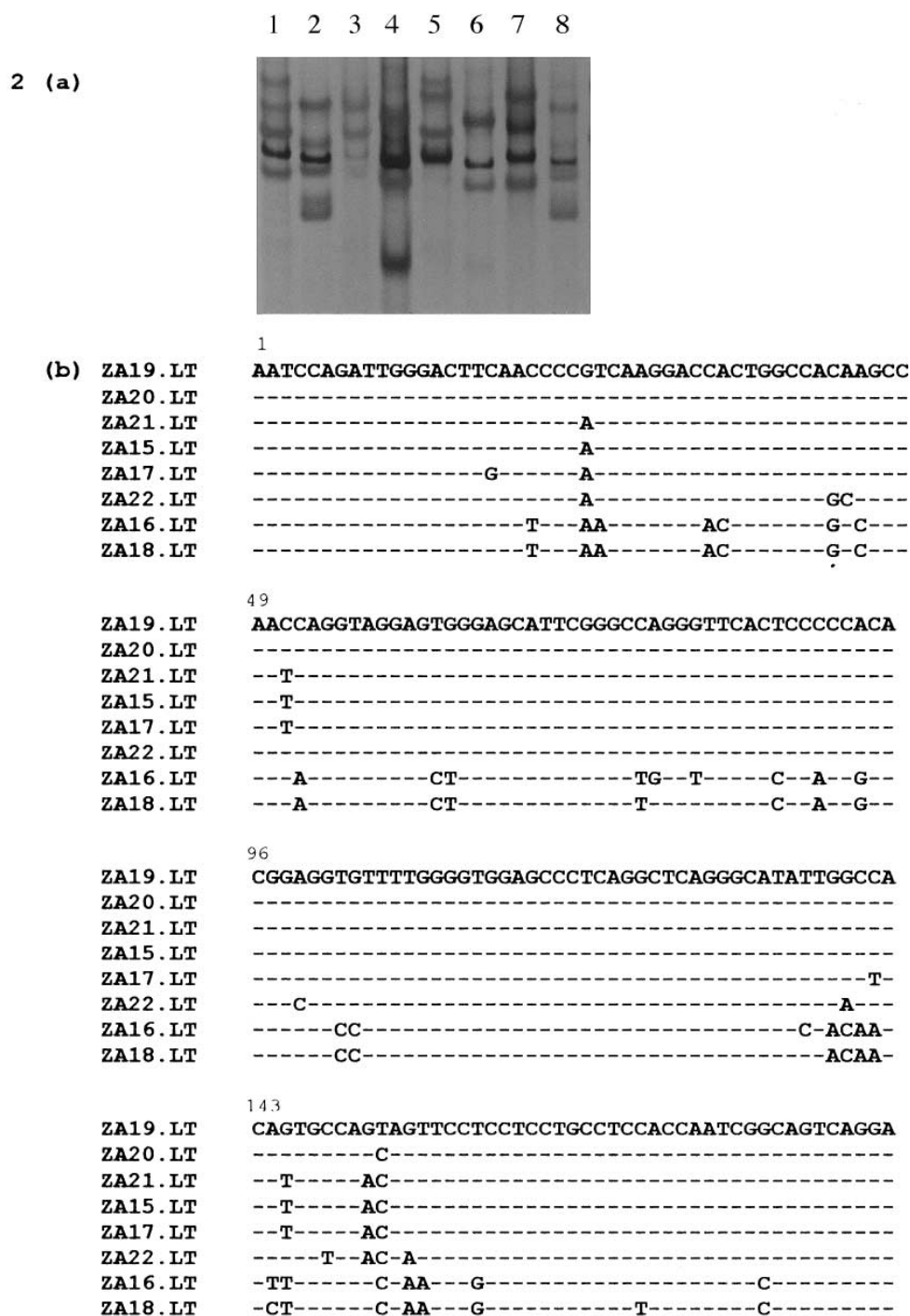


Fig. 2. a SSCP profiles of 8 liver transplant patients. Six were unique, but 2 (2 and 8) shared the same profile. **b** Sequence analysis revealed that 6 of the sequences were similar and the remaining two were similar to each other. The sequence of two patients who exhibited the same SSCP profile, namely ZA15.LT and ZA21.LT were identical in this region. ZA19.LT and ZA20.LT differed by one base in this region and ZA17.LT differed from ZA15.LT and ZA21.LT by two bases. In all cases a difference of even a single base was reflected by a change in the SSCP profile

The SSCP profiles of HBV isolates from the liver transplant patients all contained multiple single stranded bands (most contained between 4–5 bands). However, profiles of cloned and serum derived PCR products from the liver transplant were all identical (Fig. 3). This suggests that the multiplicity of bands is due to alternative conformations of a single sequence rather than to the presence of different HBV species in the same patient.

The entire preS1 gene of two HBV isolates from liver transplant patients were identical in sequence to strains found in the oncology children. ZA19.LT was identical to the onc3 strain and ZA16.LT was identical to the onc1 strain. The liver transplant patients from whom these strains were isolated were from the same hospital as the oncology patients.

The liver transplant isolates and their possible linkage to nosocomial sources of HBV are listed in Table 1.

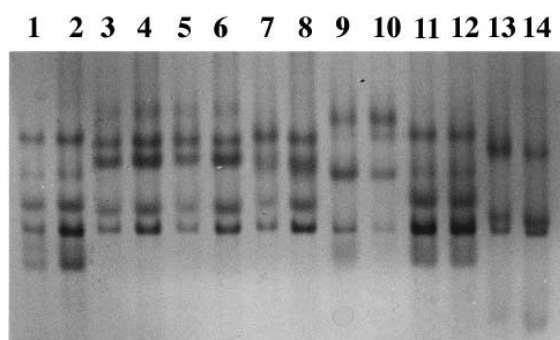


Fig. 3. The SSCP profiles of cloned and serum derived PCR products were identical in all the transplant patients. For each patient, the first lane shows the SSCP profile of PCR products derived from cloned DNA and the second lane that from serum derived DNA. The samples loaded are ZA20.LT (1, 2) ZA21.LT (3, 4), ZA15.LT (5, 6), ZA17.LT (7, 8), ZA16.LT (9, 10), ZA19.LT (11, 12) and ZA18.LT (13, 14)

Table 1. Epidemiological linkage of HBV isolates from liver transplant patients

Hospital	Isolate	Putative linkage
1 (total=3)	ZA20.LT	donor
	ZA21.LT	donor
	ZA22.LT	donor
2 ^a (Total=5)	ZA19.LT	onc3
	ZA16.LT	onc1
	ZA15.LT ^b	donor or? ZA17.LT
	ZA17.LT ^b	donor or? ZA15.LT
	ZA18.LT	donor

^a At the same hospital as the oncology unit

^b Differed from each other by 2 bases

Table 2. Genotypes of local HBV isolates

Source of HBV	Genotype		
	A	D	C
Community acquired infections	7	6	1
Oncology outbreak	1	4	
Liver transplant patients	6	2	
Total	14	12	1

Phylogenetic analysis

The phylogenetic tree shown in Fig. 4 compares preS1 sequences of HBV isolates from Cape Town with reference clones. All the reference clones grouped according to the genotypes previously assigned to them except for reference clone adw/LSH. Although this clone was previously defined as belonging to genotype D based on analysis of the full length genome sequence, it fell as an outgroup in the phylogenetic tree based on the preS1 gene. The bootstrap values for the branches defining different genotypes were all greater than 80%, indicating that this region is adequate for assigning genotypes to HBV isolates.

Of the 27 different local strains sequenced (14 from the community, 8 from liver transplant patients and 5 from oncology patients), 14 were of genotype A, 12 of genotype D and 1 genotype C (Table 2). The percentage DNA difference between the different genotypes ranged from 6–27%, while less sequence variation was noted within each genotype. The average DNA difference between local isolates of genotype A was 2.3%, and between isolates of genotype D was 2.8%.

Local genotype A sequences clustered in two subgroups. One subgroup included the onc3 strains, 7 strains from the community and 5 strains from liver transplant patients, as well as the reference sequence pFDW294 [3]. Within this subgroup, three of the liver transplant strains (ZA15, ZA17 and ZA21) clustered closely together (bootstrap value was 98%). The other subgroup contained only one South African isolate and included the reference clones pHBV933 [15] and pHBV3200 [19].

Local genotype D sequences also clustered in two subgroups. One subgroup consisted exclusively of Cape Town sequences (2 of the oncology isolates, 1 from a liver transplant patient and 3 from the community). Most of the local sequences in the other subgroups were closely related to the reference sequence ecoHBVDNA [5]. Two local HBV strains, namely onc1 and ZA16.LT were identical in the preS1 region to the reference clone ecoHBVDNA.

The single local isolate of genotype C, namely ZA01 came from a patient infected in the community.

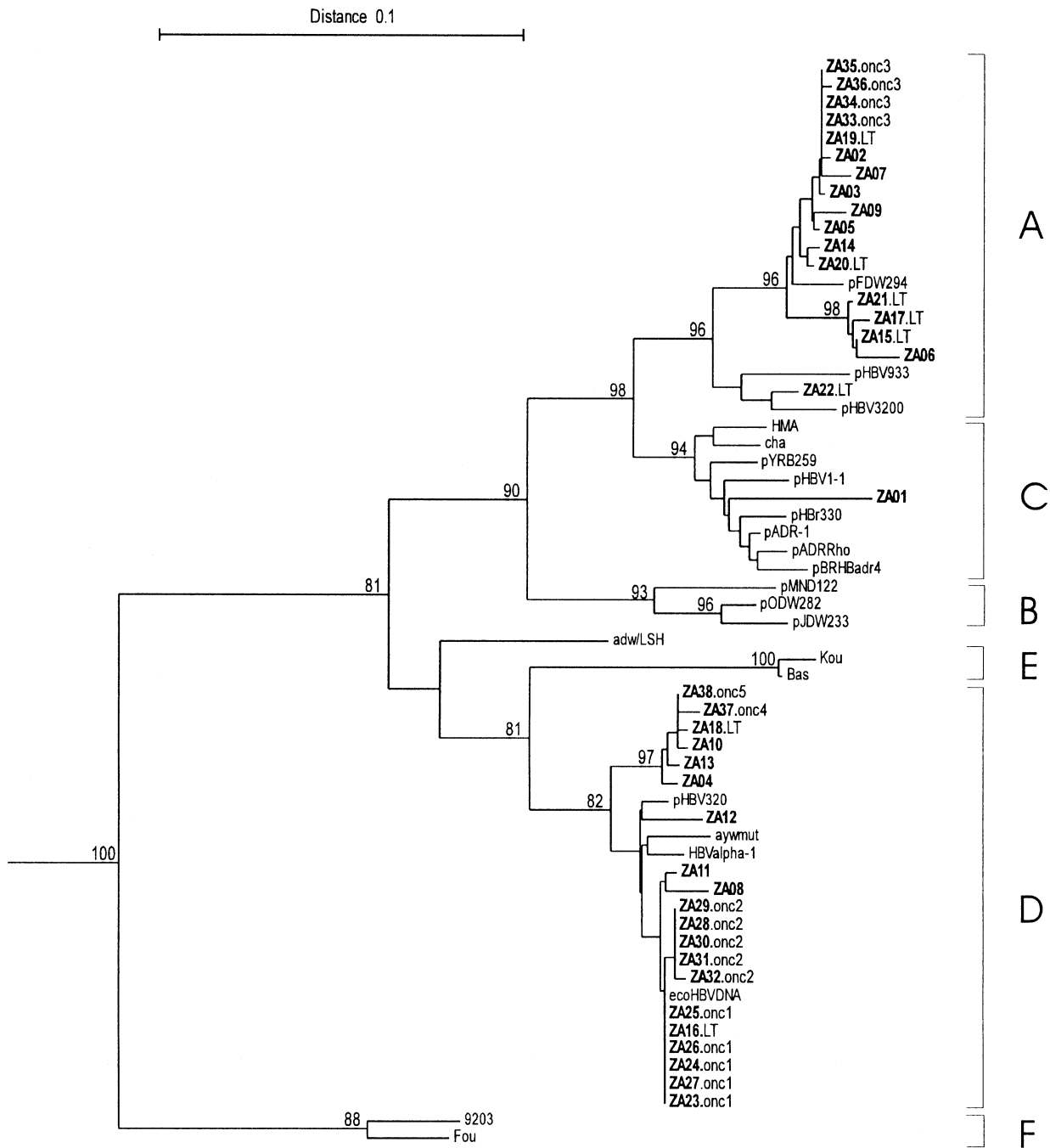


Fig. 4. Phylogenetic tree based on the preS1 gene of HBV. WHV1 was used as an outgroup. Bootstrap values of greater than 80% are shown. Reference clones grouped according to the genotypes previously assigned to them. Of the 27 unique sequences from Cape Town, 14 were of genotype A, 12 of genotype D and 1 of genotype C. The accession numbers of the control sequences were as follows: pFDW294, M57663; pHBV933, V00866; pHBV3200, X02763; HMA, X75665; cha X75656; pYRB259, X04615; pHBV1-1, M12906; pHB330, D00630; pADR-1, M38454; pADRRho, X14193; pBRHBadr4, X01587; pMND122, M54923; pODW282, D00330; pJDW233, D00329; adw/LSH, D00220; Kou, X75664; Bas, X75657; aywmut, X59795; HB Valpha-1, M32138; pHBV320, X02496; ecoHBVDNA, V01460

Discussion

The pre-S1 gene is one of the most variable regions of the HBV genome. A comparison of this region from 27 HBV genomes comprising multiple genotypes [13], revealed that only 49.6% of the nucleotides were conserved. This variation is particularly extensive between the different genotypes, ranging between 8.4 to 25.8%, but within particular genotypes it is more limited, ranging between 1.4 and 4.9%. An indication of the degree of variation in the preS1 gene of local isolates of HBV was gained by SSCP and sequence analysis of isolates from random infected individuals in Cape Town. PreS1 DNA from local HBV isolates from patients with community acquired infections displayed a wide variety of SSCP profiles [7], implying considerable sequence variation. However, while differences in the mobility of ssDNA fragments reflect differences in their nucleotide sequence, they give no indication of the degree of variation. The present investigation has revealed that the percentage DNA difference in the preS1 gene of local isolates was only moderate, varying from 2.3% for genotype A to 2.8% for genotype D. Within a particular genotype, some strains differed by as little as a single base, but they could nonetheless be differentiated by SSCP, substantiating the sensitivity of the technique. Since only five different SSCP profiles were obtained with preS1 DNA from 52 oncology patients screened during an outbreak of HBV, it was concluded that no more than 5 different strains were implicated, and that the infections of those with the same SSCP profile were linked epidemiologically. Sequence analysis of examples of each of the 5 strains confirmed that those with the same SSCP profiles had complete sequence identity across the 189 bp region used for SSCP (and with two exceptions, these strains were identical over the whole preS1 gene). Genetic variation between the five strains ranged between 1 and 32 bases. Two pairs of strains (onc1 and onc2, and onc4 and onc5) differed from each other by only one and two bases respectively. In view of this conformity, it is possible that each of the closely related pairs may have arisen from a common strain. Although all patients were tested for hepatitis B surface antigen (HBsAg) prior to the initiation of therapy, many were not tested for as long as 5 years thereafter. It is therefore not known how much time had elapsed before their infections were identified and mutations might have arisen during this period. However, as unlinked HBV isolates from the community have also been shown to differ by as little as one or two bases, we cannot be sure how many strains were originally implicated in the extensive outbreak.

It was shown that the sensitivity of SSCP is such that it can be used to establish a common source of infection [7]. Therefore, when HBV infection was identified in eight liver transplant patients (albeit from two different transplant units), SSCP was once more utilised to investigate the possibility of cross-infection. However, since HBV isolates from only two of the eight patients had the same SSCP pattern, and the patients harbouring these two strains were from different hospitals, a common source for their infection was unlikely. Sequence analysis of the preS1 region confirmed that 5 of the strains

were clearly different, while three of the isolates (ZA21.LT, ZA17.LT and ZA15.LT) clustered together on the phylogenetic tree. The patients harbouring isolates ZA17.LT and ZA15.LT were from the same hospital and although the sequences differed from each other by two bases, the possibility that the patients were infected from a common source cannot be excluded. It was also noted that two HBV isolates from liver transplant patients had preS1 sequences that were identical to two of the HBV strains implicated in the oncology unit outbreak. Since both the transplant unit and the oncology unit were at the same hospital and some facilities are shared by both units, cross infection remains a possibility.

The origin of HBV infections following liver transplantation was a matter of concern. As a common source of infection had been excluded for most of the patients, the most probable source of infection was the donor organ. In the past, liver donors were screened for HBsAg only, and not for other markers of HBV infection, but this practice has now been re-assessed. The prevalence of HBV infection in South Africa is relatively high, with up to 55% of individuals showing evidence of past infection (anti-HBs positive or anti-HBc positive) [8]. Although prevalence rates in the Western Cape urban areas are not quite as high as in the more outlying districts, a significant percentage of donors are nonetheless likely to have been exposed to HBV. It has been reported [1, 11] that recipients may acquire HBV from the engrafted liver of donors who have had HBV infection in the past (even though HBsAg may have been cleared from their blood). Presumably this occurs as a result of reactivation of virus latent in the graft, following introduction into a patient who is severely immunosuppressed. Relevant to this hypothesis is a report by Rehmann et al. [16] which provides evidence that HBV may persist for many years in patients who have apparently cleared the infection. Consequently, our current policy is to include only those liver donors who are negative for both HBsAg and anti-HBc.

Most of the information regarding HBV variation in Africa is based on sero-epidemiological surveys performed in the 1980s. Courouce-Pauty et al. [2] showed that the subtypes of HBV found in Africa are mainly ayw4, ayw2, and adw2, but the prevalence of these subtypes varies in different parts of the continent: subtype ayw2 (genotype D) predominates in North and Central Africa and ayw4 (genotype D or E) in West Africa, while in sub-saharan Africa most infections are caused by subtype adw2 (genotype A or B). In South Africa 95% of HBV isolates have been shown to belong to subtype adw2. However, subtypes determined serologically do not segregate perfectly into the 6 genomic groups, and information regarding the genotypes of HBV that are prevalent in this part of the world is limited. Therefore, sequence data gained during the current investigation was subjected to phylogenetic analysis to give an indication of the genotypes of the local HBV strains. The group of 27 different HBV isolates from Cape Town were almost exclusively genotypes A and D, but included a single isolate of genotype C. While genotypes A and D have a widespread distribution around the world, including Western Europe, USA,

Sub-Saharan Africa, the Middle East and Asia as far east as the Phillipines [12], it appears that genotype C is mainly confined to the Far East. The relative preponderance of genotypes A and D in our local population is consistent with the prevalence of these two genotypes on the rest of the African continent, while the isolate of genotype C (from a patient of Asian extraction) might have been introduced from the East.

Both episodes of nosocomial infection involved HBV strains of genotypes A and D. Despite the reported high degree of variability of the preS1 gene [13], we observed little variation in local strains within the two genotypes. Nevertheless sequence data confirmed that the SSCP technique was able to distinguish between closely related isolates, even where they differed by as little as a single base in the region examined. Epidemiological linkage was confirmed in 52 patients who were infected nosocomially during an outbreak in an oncology unit and, in addition a further two liver transplant patients were probably also infected from this source. However, we were unable to link the remaining 6 post operative HBV infections and the donor liver appears to be the most likely source of infection. Further work needs to be done to test this hypothesis.

Acknowledgements

This work was funded in part by the Poliomyelitis Research Foundation. We gratefully acknowledge the help of Dr. Linda Stannard who assisted with the preparation of the manuscript.

References

1. Chazouilleres O, Mamish D, Kim M, Carey K, Ferrell L, Roberts JP, Ascher N, Wright TL (1994) "Occult" hepatitis B virus as a source of infection in liver transplant recipients. *Lancet* 343: 142–145
2. Courouce-Pauty A-M, Plancon A, Soulier JP (1983) Distribution of HBsAg subtypes in the world. *Vox Sang* 44: 197–211
3. Estacio RC, Chavez CC, Okamoto H, Lingao AL, Reyes MT, Domingo E, Mayumi M (1988) Nucleotide sequence of a hepatitis B virus genome of subtype adw isolated from a Phillipino: comparison with the reported three genomes of the same subtype. *J Gastroenterol Hepatol* 3: 215–222
4. Galibert F, Nan Chen T, Mandart E (1982) Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence. *J Virol* 41: 51–65
5. Galibert F, Mandart E, Fitoussi F, Tiollais P, Charnay P (1979) Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli*. *Nature* 281: 646–650
6. Genem D (1995) Hepadnaviridae and their replication. In: Fields BN, Knipe DM, Howley PM (eds) *Fields virology*, 3rd ed. Lippencott-Raven, Philadelphia, pp 2703–2737
7. Hardie DR, Kannemeyer, J, Stannard LM (1996) DNA single stranded conformational polymorphism identifies five defined strains of hepatitis B virus (HBV) during an outbreak of HBV infection in an oncology unit. *J Med Virol* 49: 49–54
8. Kew MC, Kassianides C, Hodgkinson J, Coppin A, Paterson AC (1986) Hepatocellular carcinoma in urban born blacks: frequency and relation to hepatitis B virus infection. *Br Med J* 293: 1339–1341

9. Kimura MA (1980) Simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences *J Mol Evol* 16: 111–120
10. Kobayashi M, Koike K (1984) Complete nucleotide sequence of hepatitis B virus DNA of subtype adr and its conserved gene organization. *Gene* 30: 227–232
11. Lowell JA, Howard TK, White HM, Shenoy S, Huettner PC, Brennan DC, Peters MG (1995) Serological evidence of past hepatitis B infection in liver donor and hepatitis B infection in liver allograft. *Lancet* 345: 1 084–1 085
12. Norder H, Hammas B, Lee-S-D, Bile K, Courouce A-M, Mushahwar IK, Magnius LO(1993) Genetic relatedness of hepatitis B viral strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. *J Gen Virol* 74: 1 341–1 348
13. Norder H, Courouce A-M, Magnius LO(1994) Complete genomes, phylogenetic relatedness, and structural proteins of the hepatitis B virus, four of which represents two new genotypes. *Virology* 198: 489–503
14. Okamoto H, Tsuda F, Sakugawa H, Sastrosoewigijo 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: 2 575–2 583
15. Ono Y, Onda H, Sasada H, Igarashi K, Sugino Y, Nishioka K (1983) The complete nucleotide sequence of the cloned hepatitis B virus DNA: subtype adr and adw. *Nucleic Acids Res* 11: 1 747–1 757
16. Rehmann B, Ferrari C, Pasquinelli C, Chisari F (1996) The hepatitis B virus persists for decades after patients recovery from acute viral hepatitis despite active maintenance of a cytotoxic T lymphocyte response. *Nature Med* 2: 1 104–1 108
17. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4 673–4 680
18. Uy A, Wunderlich G, Olsen DB, Heerman K, Gerlich WH, Thomssen R (1992) Genomic variability in the preS1 region and determination of routes of transmission of hepatitis B virus. *J Gen Virol* 73: 3 005–3 009
19. Valenzuela P, Quiroga M, Zaldivar J, Gray P, Rutter WJ (1980) The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. In: Fields B, Jahnisch R, Fox C (eds) *ICN-UCLA symposia on animal virus genetics*. Academic Press, New York, pp 57–70
20. Van de Peer Y (1994) TREECON for Windows, Version 1.0. Computer programme distributed by Y van de Peer, Dept Biochemistry, Univ Antwerp, Belgium

Authors' address: Dr. D. R. Hardie, Department of Medical Microbiology, University of Cape Town Medical School, Anzio Road, Observatory, 7925 Cape Town, South Africa.

Received January 27, 1997