
The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype adr and adw

Yoshitaka Ono*, Haruo Onda*, Reiko Sasada*, Koichi Igarashi*, Yukio Sugino* and Kusuya Nishioka[†]

*Biotechnology Laboratories, Central Research Division, Takeda Chemical Industries, Ltd., Jusohonmachi 2-17-85, Yodogawa-ku, Osaka 532 and [†]Tokyo Metropolitan Institute of Medical Science, Honkomagome 3-18-22, Bunkyo-ku, Tokyo 113, Japan

Received 12 January 1983; Revised and Accepted 1 March 1983

ABSTRACT

The complete nucleotide sequences of two different subtypes (adr and adw) of hepatitis B virus (HBV) DNA cloned in *E. coli* were determined. The sequence of the viral genome of the adr clone was 3188 nucleotides long, and that of the adw clone was 3200 nucleotides long. The adr and adw clones differed from the reported cloned ayw HBV DNA (3182 nucleotides long) in 11.2% and 10.0% of nucleotides, respectively. Heterogeneity of the HBV genome in the clones with the same subtype was observed.

INTRODUCTION

Viral hepatitis, especially that caused by infection with hepatitis B virus (HBV), is a major worldwide public health problem. HBV is a 42 nm particle (Dane particle) consisting of a core containing the viral genome (3200 base pairs (bp) of partially single stranded DNA) bound to the core protein and its own DNA polymerase (1,2); there are at least four subtypes (adw, adr, ayw, and ayr) of HBV that can be recognized by antigenic differences in the surface antigen (HBsAg)(3). However, since the virus has not been propagated in cultured cells, the biological study and preparation of large quantities of HBsAg, which could be used as a vaccine against HBV infection, have been hampered.

To resolve this problem, it is necessary to analyze the viral genome with a defined subtype. Recently the viral genomes were cloned in *E. coli* (4,5,6), and the nucleotide sequences were determined(7,8,9,10). The complete nucleotide sequence of the subtype ayw HBV DNA has been reported by Galibert *et al.*(8). Although Valenzeula *et al.*(10) also reported the complete nucleotide sequence of HBV DNA, they did not use a defined subtype of HBV, but speculated the subtype by congruence of the restriction endonuclease cleavage pattern with that of an adw2 HBV DNA (4). Information on the nucleotide sequence of the other subtypes of HBV is very limited. We prepared HBV DNA from the Dane particles with defined HBsAg subtypes, adr and adw (these subtypes of HBsAg were mainly found in the Southeast Asia and Japan (11,12,13,14)), and cloned each HBV DNA in *E. coli*. In this paper, we report the nucleotide sequences of the cloned HBV DNA, and discuss the differences in the nucleotide sequences among the different HBV subtypes.

MATERIALS AND METHODS

Enzymes and reagents

All restriction endonucleases were purchased from Takara Shuzo. *E. coli* DNA polymerase I large fragment and bacterial alkaline phosphatase were obtained from Bethesda Research Laboratories. T4 DNA ligase was purchased from New England Biolabs. The enzymes were used as recommended by the suppliers. All reagents for DNA sequencing were supplied by Bethesda Research Laboratories.

Preparation of HBV DNA

Subtype adr HBV Dane particles were isolated from the plasma (20 ml) of a single donor infected persistently with HBV that was defined as subtype adr of HBsAg, and subtype adw HBV Dane particles were isolated from the plasma (20 ml) of a adw HBV-infected chimpanzee as described previously(15). The single stranded region of HBV DNA was repaired by an endogeneous DNA polymerase reaction according to Kaplan *et al.*(16). The repaired double stranded viral DNA was extracted using a proteinase/SDS/phenol procedure(17).

Cloning of HBV DNA in *E. coli*.

HBV DNA (subtype adr), made double stranded and purified from the Dane particles, was digested with BamHI and ligated to BamHI digested, alkaline phosphatase treated pBR322 by T4 DNA ligase. Transformation of *E. coli* λ 1776 was carried out according to Enea *et al.* (18). For cloning of adw HBV DNA, viral DNA was digested with EcoRI, and cloned in the pBR322 EcoRI site. Recombinant plasmid DNA was prepared from each transformant by the alkaline-SDS procedures (19). The plasmids, which contained approximately 3200 bp inserts, were screened by digestion with BamHI or EcoRI. One of these plasmids, which was thought to contain the full length of HBV genome, was named pHBr330 as the subtype adr clone, and another was referred to as pHBV933 for the subtype adw HBV clone. pHBr330 and pHBV933 were used as the DNA sources for the determination of the nucleotide sequences.

DNA sequence determination

Plasmid DNA was prepared from each clone by alkaline-SDS procedures followed by equilibrium density gradient centrifugation. In order to determine the nucleotide sequences by the dideoxy chain termination methods (20,21,22), approximately 3200bp inserts of pHBr330 and pHBV933 were digested with several restriction enzymes. The digested fragments were subcloned in M13mp7 phage. Using this M13mp7 phage subclone bank, the cloned HBV DNAs were sequenced. Preparation of the recombinant single stranded phage DNA as a template, DNA polymerase reaction, and polyacrylamide-urea gel electrophoresis were performed according to the manuals of the sequencing kit supplied by Bethesda Research Laboratories. The strategies for sequencing of the pHBr330 and pHBV933 inserts are shown in Fig. 1.

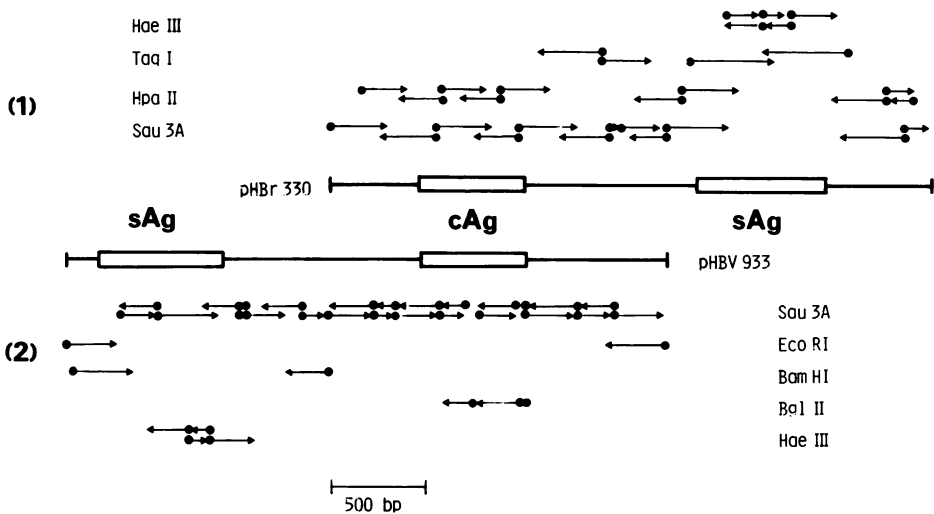


Fig. 1. The sequence strategies for the cloned HBV DNA (1) adr clone, pHBr330 insert, (2) adw clone, pHBV933 insert. The open boxes represent the coding regions of HBsAg and HBcAg. The dots indicate the position of each restriction endonuclease cleavage site, and the arrows indicate the directions and ranges of sequence determination. Only restriction endonuclease cleavage sites used for sequence are shown.

RESULTS

Cloning of subtype adr and subtype adw HBV DNA

In order to clone the full length of HBV DNA, we first of all searched for the restriction enzymes that digested only one position on the viral genome. As reported for the subtype of ayw or adw2 HBV DNA, adw HBV DNA was found to have an unique EcoRI site (data not shown). So we chose the EcoRI site of pBR322 for the cloning of adw HBV DNA. On the other hand, the greatest portion of our adr HBV DNA was not digested with EcoRI, but was digested with BamHI at one position. We, therefore, cloned the subtype adr HBV DNA at the BamHI site of pBR322 (data not shown). The recombinant plasmids which contained about 3200bp inserts were chosen by the size selection of inserts. From these clones, pHBr330 and pHBV933 were chosen as the DNA source for the DNA sequencing of the subtype adr and adw HBV DNA, respectively.

DNA sequencing analysis

The nucleotide sequences of pHBr330 insert (adr clone) and pHBV933 insert (adw clone) were determined by the dideoxy chain termination procedure employing templates derived from single stranded M13mp7 phage derivatives that were a subclone bank prepared from each insert DNA. As shown in Fig. 2, the sequence of the cloned adr HBV DNA was 3188 nucleotides long, and that of the adw clone was 3200 nucleotides long.

```

10      20      30      40      50      60      70      80      90      100
(1) TTGCACAGCATTCACCAAGCTCTGCTAGATTCACAGTACAGGGCTTATAATTTCTGTTGGTGGCTTCTAGTTCGGAAATAGTAAGCTCTGTTCCGAGT
(2) -----TG-C-G-----AG-----C-----T-G-----A-----A-----T-----C-----T-----A-----
(3) -----C-----A-----A-----C-----A-----C-----C-----C-----T-----A-----T-----T-----

(1) ACTGCCACCCATATCGTCAATETTTCCAGACTGGGGACCTTCACCCAAATTCAGAAACAAACATACAGGATTCCTAGGACCCCTGGCTGCTGTAC
(2) -T-----T-A-----C-----C-G-----TGA-----T-----TC-----TC-----T-----T-----
(3) -----T-----T-----G-T-----T-----TC-----T-----T-----T-----

(1) AGGGGGGTTTTCTGTGACAGAACTCTACAAATCCACAGAGCTTAGAATCGTGGTGGACTCTCTCAATTTCTAGGGGACCCACCCAGCTGCTCC
(2) -----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----
(3) -----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----

(1) TGGCCAAAATTCGGAGTCCCAACCTCCAATCACTACCAAACTCTTGCTCCCAATTTGCTCCGCTATCGGTGATGCTGCTGGCCGTTTTATCATA
(2) -----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
(3) -----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----

(1) TTGCTCTCATCTGCTGCTATGCTCTCTCTCTGCTTGGTCTCTGGACTACCAAGGATGTTGGCCGTTTGTCTCTACTCCAGGAACATCAACCA
(2) -----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
(3) -----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----

(1) CCAGCAGCGGCCATGCAAGACTCCAGCATTCCTGCTCAAGCAACCTCTATGTTCCCTCTGTTGCTGACAAAACCTTCGGAGGAAATGCCACTG
(2) -----T-----A-----A-----A-----C-----C-----A-----A-----A-----A-----A-----A-----T-----C-----
(3) -----C-----A-----CG-----T-----C-----A-----A-----A-----A-----A-----C-----C-----A-----T-----C-----

(1) TATTCGCATCCATCACTCCGGCTTCGCAAGATTCCATGCGGAGTGGGCCAGTCCGTTTCTCCGGCTCAGTTACTAGTGCATTGTTCAAGTGG
(2) -----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----T-----
(3) -----G-----A-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----

(1) TTGCTAGGGCTTTCCCCACTGTTTGGCTTCAGTTATATGGATGATGGATTTGGGGGCCAAGTCTACAAACATCTCAGTCCCTTTTACCCTAT
(2) -----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
(3) -----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----

(1) TAACTAATTTCTTTTCTTTCGGTATACATTTAAACCTCAATAAAACCAAACTTGGGGCTACTCCCTTAACTTCATGGGATATCTAATTGGAACTTG
(2) -----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----T-----T-----
(3) -----C-----A-----GA-----A-----T-----T-----A-----T-----T-----T-----C-----T-----AT-----

(1) CGTACTTTACCCAGGAACTATTTGCTACTAAACTCAAGCAATGTTTTCGAAAATTCCTGTAATAGCCCTATTGATTCGAAAGTATGCAAGAAATG
(2) -----A-----G-----A-----T-----A-----GA-----A-----C-----A-----C-----T-----T-----C-----G-----
(3) -----GT-----C-----G-----A-----A-----C-----CA-----A-----A-----AG-----A-----A-----C-----T-----A-----T-----C-----G-----

(1) TGGGCTTTTGGGCTTTCGCTCCCTTTTACAAATTCGGGCTACTCTGCTCTGATGCTTTTATATGCAATGATACAATCAAGCAGGCTTTCACCTTCTC
(2) -----T-----A-----T-----A-----T-----A-----G-----G-----G-----G-----G-----G-----A-----A-----C-----CA-----
(3) -----T-----T-----T-----T-----T-----T-----C-----G-----C-----C-----C-----C-----A-----A-----C-----CA-----

(1) GCGAAGTTATAAGGCTTTCTGTGTAACAATATCTGAAGCTTTACCCGTTTGGCCGGAACGGTACAGTCTCTGGCAAGTCTTTGCTGAGCAAGCTCC
(2) -----C-----AA-----G-----CA-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----
(3) -----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----G-----G-----

(1) ACTGGATGGGCTTGGCCATAGCCATTCGGGCTATCGCTGGAAACCTTTGGCTCCTCTGCCATCCCACTTCGGAAATCTTAGGAGCTGTTTCTGCTC
(2) -----C-----A-----T-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
(3) -----C-----T-----G-----A-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----

(1) GAGCGGCTCTGGAGGAAACTTATCGGAGCGGAGAACTCTGTTCTCTCTCTGCAAAATACAGCTCTTTTCCATGGCTGGTAGGCTGTGTGGCAAGTG
(2) -----A-----G-----C-----T-----A-----C-----T-----T-----A-----A-----C-----C-----C-----C-----C-----L-----A-----
(3) -----A-----A-----CA-----G-----T-----T-----T-----C-----A-----C-----C-----T-----A-----G-----C-----A-----C-----C-----

(1) GATCGTGGCGGAGCTCTTTGCTATCTGCTGGGCTGAAATCCCGGGACACCGCTGCGGGCCGCTTGGGGCTTACCGCTCTCTCTCTCTCTCT
(2) -----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----
(3) -----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----

(1) CTGCGCTTTCGGGACACGGGGGCACTCTCTTAGCGGGTCTCCCGCTGTGGCTTCTGATCTGGCGAGCCGTGGCAATTCGGTTACCTCTGCTG
(2) -----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----T-----
(3) -----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----T-----

(1) AGCTGGGATCGAGAGCAGCCGTAACGCGCAGCACTCTTCCCAAGCTCTTACATAAGAGGACTCTGGACTCTCAGCCATGCTCAACGACCGACCTTAG
(2) -----T-----C-----C-----C-----T-----A-----C-----G-----G-----G-----G-----G-----G-----C-----A-----
(3) -----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----AA-----AA-----

(1) GCATATTCGAAGACTGCTGCTTTAAGACTGGGAGGAGTTGGGGAGGAGATAGGTTAAGGCTCTTTGCTACTAGGAGCTGTAGGCAAT////////
(2) -----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----T-----A-----
(3) -----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----AAATGGCTCT
AAATGGCTCT

(1) ////////////AACTTTTCACTCTGCTCAATCATCTCATGTTCAATGCTCTACTGTTCAAGCCCTCAAGCTGTGCCCTTGGGTTGGCTTTGGGG
(2) GCGACAGGACCCATGC-----T-----A-----C-----
(3) GCGACAGGACCCATGC-----T-----A-----C-----

(1) ATTACATTGACCCGTATAAGAAATTTGGAGCACTGTGGAGTACTCTCTTTTTCGCTCTGCACTCTTTCCGCTATTCGAGATCTCTTGCACCCG
(2) -----T-----T-----IA-----T-----T-----A-----C-----C-----C-----G-----A-----G-----C-----
(3) -----C-----T-----T-----TA-----C-----G-----T-----A-----G-----A-----T-----A-----T-----

(1) CCTCTGCTCTGATCGGAGGCTTAGAGTTCCTGGCAATTTGCTACCTCAACCATACAGCACTCAGGCAAGCTATTCTGTTGGGGTGAGTTAATGAA
(2) -----A-----A-----G-----A-----T-----T-----G-----C-----A-----T-----T-----C-----C-----G-----A-----G-----C-----
(3) -----A-----A-----T-----G-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----G-----AC-----C-----

(1) TCTGGCCACTGGTGGGAAATAATTTGGAAGCCACAGATCCAGCAATTAAGTACTCAGCTATGCTCAATTTAATATGGCCATAAAATCAAGAACATTA
(2) -----A-----T-----T-----A-----C-----T-----A-----TC-----AT-----T-----AC-----C-----TT-----G-----G-----
(3) -----A-----T-----T-----T-----G-----T-----G-----T-----A-----CC-----T-----CAC-----T-----GT-----G-----C-----

(1) TTCTGGTTTACATTTCTCGCTTACTTTTGGAGAAACTGTTTGGACTATTGGTATCTTTTGGAGTGGAAATCGCACTCCCTCCGCTTACAGAC
(2) -----T-----A-----T-----T-----G-----AC-----T-----A-----C-----C-----C-----C-----A-----C-----T-----
(3) -----T-----T-----T-----C-----A-----A-----A-----A-----G-----C-----C-----C-----C-----A-----T-----

```

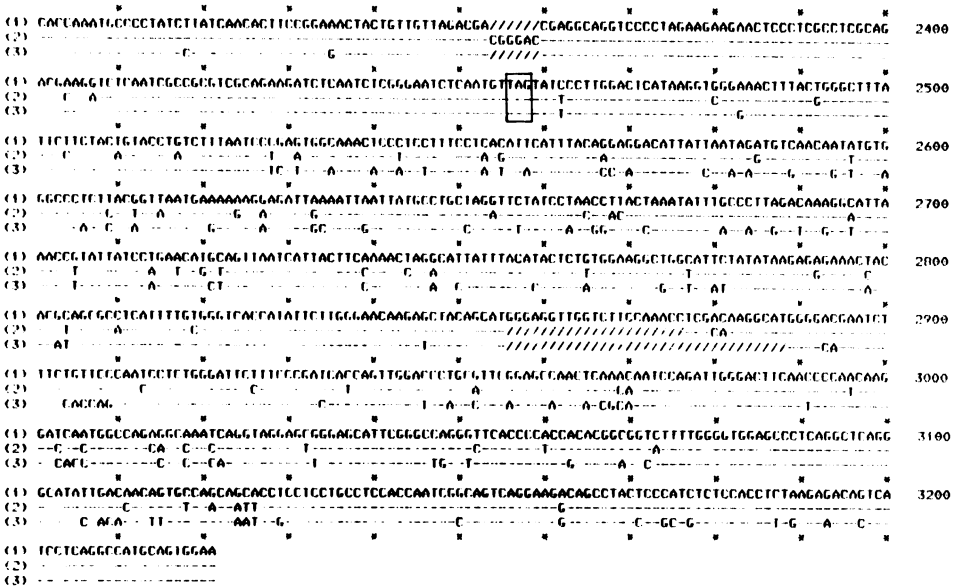


Fig. 2. Comparison of the nucleotide sequences among the different subtypes of HBV DNA. (1) adr clone, pHBr330 insert, (2) adw clone, pHBV933 insert, (3) the reported ayw HBV DNA(8). In order to compare the three sequences, the sequences are aligned at the EcoRI site. The nucleotide numbers are tentative, because the deletion and additions exist in each sequence. The bars indicate the same nucleotide as in lane (1). The oblique lines indicate deletions. The initiation codons and the stop codons of HBsAg and HBcAg are boxed.

The adr subtype differed from the reported ayw HBV DNA (3182bp) by deletion of 27bp (1791-1817) and the addition of 33bp (2856-2888), while the adw subtype differed from ayw by the addition of 18bp (2354-2359 and 2877-2888). A two by two analysis of the three nucleotide sequences (adr,adw,and ayw) demonstrated a relatively high degree of divergence. The nucleotide differences were 11.2% between adr and ayw, 10.0% between adw and ayw, and 9.8% between adr and adw. No region contains a marked divergence or a mutational hot spot, except for the deletion and addition of the sequences noted above. Our adw HBV DNA sequence differed from the adw2 sequence reported by Valenzuela *et al.*(10) by only 1.6% of the nucleotides.

Search for the viral antigen coding regions

Galibert *et al.*(8), Valenzuela *et al.*(10), and Pasek *et al.*(7) have mapped the genes of the two major viral antigens, HBsAg and HBcAg(core antigen), in the HBV genome. We searched for these antigen coding regions in the adw and adr HBV genomes. According to the investigators mentioned above, the HBsAg gene is located between 155 and 832, and consisted of 678 nucleotides coding for 226 amino acids in both HBV subtype

genomes. The amino acid sequences deduced from the nucleotide sequences of the three subtype antigens are shown in Fig. 3. The overall amino acid difference was 7.5% (17 amino acids changes/226 amino acids), but there were relatively large differences in the central hydrophilic region between amino acid residues 113 to 143; 6 changes (19.4%) between adr and adw, 8 changes (25.8%) between adw and ayw, and 7 changes (22.6%) between adr and ayw. Recently, Peterson *et al.*(23) determined the amino acid sequences of the HBsAg of the adw and ayw subtypes, and demonstrated that amino acid residues 122 to 150 of each antigen occupy an exposed region of the HBsAg lipoprotein particle and the two subtypes were found to differ at two specific position in this region. Therefore, this region might be an antigenically important area of the HBsAg.

HBcAg gene in the adr and adw HBV genome was located at the same region as that in the ayw genome. While the HBcAg gene in the subtype adr HBV genome consisted of 183 amino acids as the subtype ayw, two additional amino acids were observed on the adw HBV genome between residues 151 and 152 (Fig. 4). This insertional sequence was the same as that of adw2 described by Valenzuela *et al.*(10). The amino acid differences were relatively smaller than those of HBsAg and there was no region of marked divergence (about a 4% difference among the three subtypes).

One of the other viral coded proteins, DNA polymerase, was speculated to be coded in the longest open region, that overlapped the surface antigen gene(8,10). We found a long continuous nucleotide sequence devoid of a stop codon that began at ATG(2307-2309) and ended at a stop codon TGA(1621-1623) in the adr HBV genome. However, in this region of adw HBV genome a stop codon of TAA(550-552) appeared.

Heterogeneity of the cloned HBV genome

When double stranded Dane particle DNA, defined as subtype adr, was digested with EcoRI, almost all the molecules seemed to be undigestible. Nevertheless, on the DNA sequencing data, the pHB330 insert has the cutting site of EcoRI. This phenomenon might indicate that heterogeneity of the HBV genome existed in the same subtype prepared from a single donor. We, therefore, attempted to characterize other clones obtained from transformants by the chemeric plasmid pHB series. As shown in Fig. 5, some variation of the restriction patterns was observed in the three recombinant plasmids. Only pHB330 insert was digested with EcoRI. It is interesting that pHB329 insert was digested with HindIII, because all of the known HBV DNA was not digested with HindIII (5,24,25).

Although we did not determine the precise nucleotide sequences of pHB329, HindIII site might be located at 1138-1143. As the sequence of this region of pHB330 insert was AACCTT, the HindIII site (AAGCTT) could be created by one base substitution. We also observed differences of the restriction patterns in the adw clones (data not shown).

(1)	Met	Glu	Asn	Thr	Thr	Ser	Gly	Phe	Leu	Gly	Pro	Leu	Leu	Val	Leu	
(2)	-	-	-	Ile	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	Ile	-	-	-	-	-	-	-	-	-	-	-	
(1)	Gln	Ala	Gly	Phe	Phe	Leu	Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	30
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Ser	Leu	Asp	Ser	Trp	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Ala	
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Ser	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Thr	
(1)	Pro	Thr	Cys	Pro	Gly	Gln	Asn	Ser	Gln	Ser	Pro	Thr	Ser	Asn	His	60
(2)	-	Val	-	Leu	-	-	-	-	-	-	-	-	-	-	-	
(3)	Thr	Val	-	Leu	-	-	-	-	-	-	-	-	-	-	-	
(1)	Ser	Pro	Thr	Ser	Cys	Pro	Pro	Ile	Cys	Pro	Gly	Tyr	Arg	Trp	Met	
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	Thr	-	-	-	-	-	-	-	
(1)	Cys	Leu	Arg	Arg	Phe	Ile	Ile	Phe	Leu	Phe	Ile	Leu	Leu	Leu	Cys	90
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Leu	Ile	Phe	Leu	Leu	Val	Leu	Leu	Asp	Tyr	Gln	Gly	Met	Leu	Pro	
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Val	Cys	Pro	Leu	Leu	Pro	Gly	Thr	Ser	Thr	Thr	Ser	Thr	Gly	Pro	120
(2)	-	-	-	-	Ile	-	-	Ser	Thr	-	-	-	-	-	-	
(3)	-	-	-	-	Ile	-	-	Ser	-	-	-	-	-	-	-	
(1)	Cys	Lys	Thr	Cys	Thr	Ile	Pro	Ala	Gln	Gly	Thr	Ser	Met	Phe	Pro	
(2)	-	-	-	-	-	Thr	-	-	-	-	Asn	-	Lys	-	-	
(3)	-	Arg	-	-	Met	Thr	Thr	-	-	-	-	-	-	Tyr	-	
(1)	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly	Asn	Cys	Thr	Cys	Ile	150
(2)	-	-	-	-	-	-	-	Thr	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Ala	Arg	Phe	Leu	Trp	Glu	Trp	
(2)	-	-	-	-	-	-	-	-	-	Lys	Tyr	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	Gly	Lys	-	-	-	-	-	
(1)	Ala	Ser	Val	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro	Phe	Val	180
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	Ala	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val	Ile	
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	Ala	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Asn	Ile	Leu	Ser	210
(2)	-	-	-	-	-	-	-	-	-	-	-	Ser	-	Val	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	Ser	-	-	-	
(1)	Pro	Phe	Leu	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val	Tyr	
(2)	-	-	Ile	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Ile															
(2)	-															
(3)	-															

Fig. 3. Comparison of the amino acid sequences of HBsAg deduced from the nucleotide sequences of the different subtypes of HBV. (1) subtype adr, (2) subtype adw, (3) subtype ayw(8). The bars indicate the same amino acid as in lane (1).

(1)	Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Ser	Val	Glu	Leu	
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	Thr	-	-	-	
(1)	Leu	Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Ile	Arg	Asp	Leu	30
(2)	-	-	-	-	-	-	-	-	-	-	-	Val	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	Val	-	-	-	
(1)	Leu	Asp	Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Glu	His	Cys	Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	60
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Cys	Trp	Gly	Glu	Leu	Met	Asn	Leu	Ala	Thr	Trp	Val	Gly	Ser	Asn	
(2)	-	-	-	-	-	-	Thr	-	-	-	-	-	-	Asn	-	
(3)	-	-	-	-	-	-	Thr	-	-	-	-	-	-	Val	-	
(1)	Leu	Glu	Asp	Pro	Ala	Ser	Arg	Glu	Leu	Val	Val	Ser	Tyr	Val	Asn	90
(2)	-	Gln	-	-	-	-	-	Asp	-	-	-	Asn	-	-	-	
(3)	-	-	-	-	-	-	-	Asp	-	-	-	-	-	-	-	
(1)	Val	Asn	Met	Gly	Leu	Lys	Ile	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	
(2)	Thr	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	Thr	-	-	-	-	-	Phe	-	-	-	-	-	-	-	-	
(1)	Ser	Cys	Leu	Thr	Phe	Gly	Arg	Glu	Thr	Val	Leu	Glu	Tyr	Leu	Val	120
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	Ile	-	-	-	-	
(1)	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	Pro	Pro	Ala	Tyr	Arg	Pro	Pro	
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	Glu	Thr	Thr	Val	Val	Arg	150
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Arg	/	/	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	
(2)	-	Arg	Asp	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	/	/	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	180
(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(3)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
(1)	Arg	Glu	Ser	Gln	Cys											
(2)	-	-	-	-	-											
(3)	-	-	-	-	-											

Fig. 4. Comparison of the amino acid sequences of HBcAg deduced from the nucleotide sequences of the different subtypes of HBV. (1) subtype adr, (2) subtype adw, (3) subtype ayw(8). The bars indicate the same amino acid in lane (1) and the oblique lines indicate deletions.

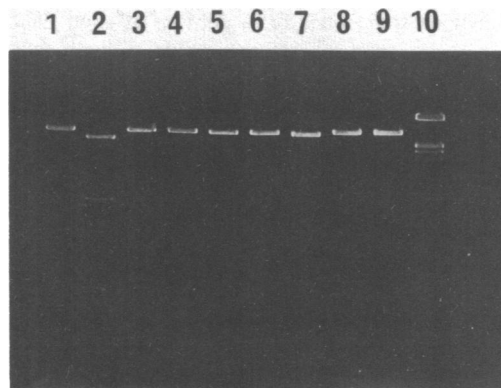


Fig. 5. Restriction fragment analysis of pHBV series . Each recombinant plasmid DNA was prepared by alkaline SDS procedures(19), digested with several restriction endonuclease without isolation of each insert DNA, and electrophoresed on 1% agarose gel. (1) pHBV329, EcoRI digestion, (2) pHBV330, EcoRI digestion, (3) pHBV347, EcoRI digestion, (4) pHBV329, XhoI digestion, (5) pHBV330, XhoI digestion, (6) pHBV347, XhoI digestion, (7) pHBV329, HindIII digestion, (8) pHBV330, HindIII digestion, (9) pHBV347, HindIII digestion, (10) size marker, lambda DNA EcoRI and HindIII digestion.

DISCUSSION

In this paper, we have reported that the two HBV DNA prepared from the Dane particles with defined subtypes, adr and adw, have been cloned in *E. coli.*, and the complete nucleotide sequences have been determined. The insert of pHBV330(adr) was 3188 nucleotides long, and that of pHBV933 was 3200 nucleotides long, respectively. The length of these cloned viral genomes was different from that of the reported one previously(8). These length differences are due to deletion or additions at two positions in both HBV subtype genomes (Fig.2). The significance of these deletion and additions within the different subtype genomes is not clear, but since they are multiples of three, the reading frame is the same in all four variants (adw, adw2, adr, and ayw).

Because of inadequate information on viral proteins and messenger RNA, it is difficult to analyze the genetic organization solely from the nucleotide sequence. Nevertheless, the coding region for viral proteins could be predicted by searching for the open regions. Galibert *et al.*(8) found four relatively long open regions in the ayw HBV genome (termed by Tiollais *et al.*(32) gene S, gene C, gene P, and gene X) and predicted that gene S, gene C, and gene P coded for HBsAg, HBcAg, and DNA polymerase, respectively. These predictions were almost correct in the case of HBsAg and HBcAg genes, because these gene products expressed in yeast (26) or in *E. coli.* (27,28,31) were biologically active. However, as information on the viral DNA polymerase is limited, the hypothesis that gene P codes for DNA polymerase has not been confirmed. Though we

found the longest open region that codes for a peptide of about 90000 daltons in the adr HBV genome as in the ayw HBV genome, we could not find this long open region in the adw HBV genome, because a stop codon of TAA appears in the center of the region. This may suggest that the splicing system of viral mRNA exist if this adw HBV genome can replicate in vivo or in vitro. Analysis of viral mRNA is necessary for understanding more precisely the genetic organization.

Recently, Pourcel et al.(29) have detected the 2300 base viral specific RNA species in mouse L cell line that was co-transformed with cloned HBV DNA and the herpes simplex virus thymidine kinase gene, and they also have detected a similar RNA in a human hepatoma cell line PLC/PRF/5 that produced HBsAg. Moreover, they mapped this RNA in the viral genome, and predicted the existence of pre-HBsAg, because the initiation codon ATG was located 489bp upstream from the initiation codon of the mature HBsAg in the same reading frame, and a TATA box (30)-like sequence (TATATAA) was also found 72bp upstream from the first ATG of the pre-S region. This open region and the TATA box like sequence (2772-2778) are also conserved in our HBV DNA sequences. It is suggested, therefore, that this sequence is related to the HBsAg promoter.

We found heterogeneity of the HBV genome within the same subtype(Fig.5). Siddiqui et al. also have shown that within the same subtype restriction patterns are slightly different (4). These findings suggest that there is a significant genetic heterogeneity of the HBV genome within the same subtype, although we can not exclude the possibility that these differences reflect artifacts on the cloning procedures.

To understand the biology of HBV, it is necessary to establish a viral infection system and to express the viral specific proteins in cultured cells. We believe that our cloned HBV DNAs are suitable for this purpose.

ACKNOWLEDGEMENT

We are grateful to Dr. M. Mayumi for supplying the plasma of a donor infected persistently with adr HBV.

REFERENCES

1. Robinson, W. S., and Greenman, R. C., (1974) *J. Virol.*, 13, 1231-1236.
2. Hruska, J. F., and Robinson, W. S., (1977) *J. Med. Virol.*, 1, 119-124.
3. Lebouvier, G. L., (1973) *Ann. Intern. Med.*, 79, 894-896.
4. Siddiqui, A., Sattler, F., and Robinson, W. S., (1979) *Proc. Natl. Acad. Sci., U.S.A.*, 76, 4664-4668.
5. Sninsky, J. J., Siddiqui, A., Robinson, W. S., and Cohen, S. N., (1979) *Nature*, 279, 346-348.
6. Charnay, P., Mandart, E., Hampe, A., Fitoussi, F., Tiollais, P., and Galibert, F., (1979) *Nucleic Acids Res.*, 7, 335-346.
7. Pasek, M., Goto, T., Gilbert, W., Zink, B., Schaller, H., MacKay, P., Leadbetter, G., and Murray, K., (1979) *Nature*, 282, 575-579.

8. Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P., and Charnay, P., (1979) *Nature*, 281, 646-650.
9. Valenzuela, P., Quiraga, M., Zaldivar, J., Gray, P., and Rutter, W. J., (1980) in *Animal Virus Genetics*, Fields, B., Jaenisch, R., and Fox, C. F., Eds., pp57-70, Academic Press, New York.
10. Valenzuela, P., Gray, P., Quiraga, M., Zaldiver, J., Goodman, H. M., and Rutter, W. J., (1979) *Nature*, 280, 815-819.
11. Yamashita, Y., Kurashina, S., Miyakawa, Y., and Mayumi, M., (1975) *J. Infect. Dis.*, 131, 567-569.
12. Nishioka, K., Levin, A. G., and Simons, J., (1975) *Bull. W. H. O.*, 52, 293-300.
13. Courouse, A. M., Holland, P. V., Muller, J. Y., and Soulier, J. P., (1976) *Bibl. Haematol.*, 42, 42-59.
14. Nishioka, K., (1982) *Trans. Japan Academy*, 38, 1-19.
15. Onda, H., Matsuda, S., and Oka, Y., (1978) *Intervirology*, 10, 254-264.
16. Kaplan, P. M., Greenman, P. L., Gerin, J. L., Purcell, R. H., and Robinson, W. S., (1973) *J. Virol.*, 12, 995-1005.
17. Mason, W. S., Geeta, S., and Summers, J., (1980) *J. Virol.*, 36, 829-836.
18. Enea, V., Vovis, G. F., and Zinder, J. F., (1975) *J. Mol. Biol.*, 96, 495-509.
19. Ish-Horowicz, D. B., and Burke, J., F., (1981) *Nucleic Acids Res.*, 9, 2989-2998.
20. Sanger, F., Nicklen, S., and Coulson, A. R., (1977) *Proc. Natl. Acad. Sci., U.S.A.*, 74, 5463-5467.
21. Sanger, F., Coulson, A. R., Barrel, B. G., Smith, A. J. H., and Roe, B. A., *J. Mol. Biol.*, 143, 161-178..
22. Messing, J., Crea, R., and Seeburg, P. H., (1981) *Nucleic Acids Res.*, 9, 309-321.
23. Peterson, D. L., Nath, N., and Gauilanes, F., (1982) *J. Biol. Chem.*, 257, 10414-10420.
24. Charnay, P., Pourcel, C., Louise, A., Fritch, A., and Tiollais, P., (1979) *Proc. Natl. Acad. Sci., U.S.A.*, 76, 2222-2226.
25. Chakraborty, P. R., Ruiz-Oapazo, N., Shouval, D., and Schafritz, D. A., (1980) *Nature*, 286, 531-533.
26. Valenzuela, P., Medina, A., Rutter, W. J., Ammerer, G., and Hall, B. D., (1982) *Nature*, 298, 347-350.
27. Edman, J. C., Hallewell, R. A., Valenzuela, P., Goodman, H. M., and Rutter, W. J., (1981) *Nature*, 291, 503-506.
28. Charnay, P., Gervais, M., Louise, A., Galibert, F., and Tiollais, P. (1980) *Nature*, 286, 893-895.
29. Pourcel, C., Louise, A., Gervais, M., Chenciner, N., Dubois, M. F., and Tiollais, P., (1982) *J. Virol.*, 42, 100-105.
30. Benoist, C., and Chambon, P., (1981) *Nature*, 290, 304-310.
31. Cohen, B. J., and Richmond, J. E., (1982) *Nature*, 296, 677-678.
32. Tiollais, P., Charnay, P., and Vyas, G. N., (1981) *Science*, 213, 406-411.