

A New Hepadnavirus Endemic in Arctic Ground Squirrels in Alaska

PATRICE TESTUT,¹ CLAIRE-ANGÉLIQUE RENARD,¹ OLIVIER TERRADILLOS,¹ LUDMILA VITVITSKI-TREPO,²
FREDJ TEKAIA,³ CLAUDE DEGOTT,⁴ JOHN BLAKE,⁵ BERT BOYER,⁵ AND MARIE ANNICK BUENDIA^{1*}

*Unité de Recombinaison et Expression Génétique (INSERM U163)¹ and Unité de Génétique Moléculaire des Levures,³
Institut Pasteur, Paris, Unité des Hépatites, SIDA et Rétrovirus Humains (INSERM U271), Lyon,² and Laboratoire
d'Anatomopathologie, Hôpital Beaujon, Clichy,⁴ France, and Institute of Arctic Biology,
University of Alaska Fairbanks, Fairbanks, Alaska 99775⁵*

Received 17 October 1995/Accepted 27 March 1996

We present evidence for a novel member of the hepadnavirus family that is endemic in wild arctic ground squirrels (*Spermophilus parryi kennicotti*) in Alaska. This virus, designated arctic squirrel hepatitis virus (ASHV), was initially detected in the livers of animals bearing large hepatic nodules by nucleic acid hybridization with hepadnavirus probes and in plasma by cross-reactivity with antibodies to hepadnavirus surface and core antigens. The complete nucleotide sequence of the 3,302-bp-long ASHV genome was determined and compared with those of ground squirrel hepatitis virus (GSHV) and woodchuck hepatitis virus (WHV); all sequences were organized into four open reading frames, designated pre-C/C, pre-S/S, *pol*, and X. Despite roughly equivalent variability among the three rodent hepadnaviruses (around 16% base and 19% amino acid exchanges), ASHV appeared to be more closely related to GSHV than to WHV in phylogenetic analysis. Accordingly, preliminary studies of the pathology of ASHV infection suggested that ASHV may be a less efficient oncogenic agent than WHV. About one-third of aged animals maintained in captivity, including virus-infected as well as uninfected squirrels, developed large liver nodules, consisting of hepatocellular adenomas or carcinomas or nonmalignant lesions characterized by drastic microvesicular steatosis. ASHV-infected arctic ground squirrels may serve as a new model with which to analyze the contribution of hepadnavirus- and host-specific determinants to liver pathology and tumorigenesis.

Hepatitis B virus (HBV) is a common human pathogen that causes acute and chronic liver disease, and a strong epidemiological relationship has been established between persistent HBV infection and hepatocellular carcinoma (HCC) (1). Since the discovery of HBV, several related viruses have been isolated in mammalian and avian species; this family group was designated *Hepadnaviridae* (25). Woodchuck hepatitis virus (WHV), originally identified in animals that were found to die frequently with HCC accompanied by hepatitis (63), is endemic in wild woodchucks (*Marmota monax*) in a few states on the East Coast of the United States. In a search for HBV-like viruses in California relatives of the woodchucks, ground squirrel hepatitis virus (GSHV) was discovered in Beechey ground squirrels (*Spermophilus beecheyi*) in northern California (36). Related viruses may also infect tree squirrels (*Sciurus carolinensis*) and Richardson's ground squirrels (*Spermophilus richardsonii*) in North America (14, 40, 65). So far, three avian viruses, duck hepatitis B virus (DHBV) (38), heron hepatitis B virus (62), and the recently discovered Ross's goose hepatitis virus (44), have been identified as members to the *Hepadnaviridae*.

The hepadnavirus family is defined by a unique virion ultrastructure, peculiar genome size and organization, and distinctive mechanism of replication by reverse transcription of an RNA intermediate; these viruses show high liver tropism and a narrow host range (reviewed in reference 67). Hepadnavirus genomes are partially double-stranded, noncovalently closed circular DNAs. With lengths of 3 to 3.3 kb, they rank among

the smallest genomes of known animal viruses. Hepadnaviruses have evolved a remarkably compact genomic organization: extensively overlapping reading frames encode the surface and nucleocapsid antigens, the secreted e antigen, the polymerase, and in the case of mammalian viruses, a transcriptional transactivator, the X antigen. In productive infections, viral gene expression is tightly regulated by a complex interplay of regulatory elements located within coding sequences at short distances from each other (3, 10, 41, 61).

Hepadnaviruses of lower animals have been largely used as models for elucidating different aspects of the viral life cycle and for investigating the mechanisms linking persistent viral infection and the development of liver cancer. The contribution of mammalian hepadnaviruses to liver carcinogenesis has been conclusively established (reviewed in reference 2). The high oncogenicity of WHV was demonstrated by the development of HCC in the absence of chemical cocarcinogens in all WHV carrier woodchucks experimentally infected as newborns (50), and a significant association of HCC with the GSHV carrier state has been reported (34, 37, 58). Our recent studies of woodchuck liver tumors led to the elucidation of a major step of the oncogenic process driven by WHV. In a majority of tumors, WHV DNA integration into the host genome provokes insertional activation of *myc* family oncogenes, with a predominant involvement of the woodchuck N-*myc2* gene (15, 16, 27, 75). Curiously, such a mechanism has not been found in HCCs induced by GSHV (26, 70) or in human tumors associated with HBV (2). Little is known about the viral and host factors responsible for marked differences in oncogenic strategy between mammalian hepadnaviruses.

The occurrence of frequent liver disease and large hepatic nodules in wild arctic ground squirrels (*Spermophilus parryi kennicotti*) trapped in the Brooks range in Alaska prompted us to search for an hepadnavirus in this rodent species. Here we report the characterization of a new virus closely related to

* Corresponding author. Mailing address: Unité de Recombinaison et Expression Génétique, Département des Rétrovirus, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France. Phone: (33)-1-45 68 88 26. Fax: (33)-1-45 68 89 43. Electronic mail address: mbuendia@pasteur.fr.

GSHV and WHV which provides an interesting model for comparative studies of liver pathology and cancer associated with chronic hepadnavirus infections.

MATERIALS AND METHODS

Animals. Arctic ground squirrels were trapped live at various locations on the Brooks range (Alaska) from 1987 through 1995 and housed at the laboratory facility of the Institute of Arctic Biology, University of Alaska Fairbanks. The animals were fed an ad libitum diet of rat chow and water, supplemented weekly with sunflower seeds and carrots. Gross examination of the organs was performed at necropsy; liver tissues were frozen and kept at -70°C for molecular biology studies or were fixed for histological examination. The animals were handled and euthanized according to National Institutes of Health guidelines.

Histology. Tissue specimens taken at necropsy were fixed in 10% buffered formalin and embedded in paraffin. Sections $5\ \mu\text{m}$ thick were stained with hematoxylin and eosin for histopathological examination.

DNA and RNA analysis. Frozen liver and tumor samples were pulverized in liquid nitrogen, and genomic DNA and total RNA were phenol extracted separately as previously described (41). For dot blot analysis, $4\ \mu\text{g}$ of DNA was denatured with NaOH (0.4 N)-NaCl (1 M) and spotted onto a nylon membrane (Hybond N⁺; Amersham), using a Minifold1 dot blotter (Schleicher & Schuell, Dassel, Germany). For Southern blot analysis, $20\ \mu\text{g}$ of genomic DNA was digested with *Bgl*II, fractionated by electrophoresis on a 0.8% agarose gel, and transferred onto a Hybond N⁺ membrane in NaOH (0.4 N) as described previously (75). Hybridizations were performed as described by Church and Gilbert (5), using cloned WHV and GSHV DNA probes (23, 59) radiolabeled by the random primer method. Filters were washed in phosphate (40 mM; pH 7.0)-EDTA (1 mM)-sodium dodecyl sulfate (SDS; 1%) at 63°C . Before subsequent hybridization with a different probe, the filters were stripped by heating in distilled water for 5 min at 90°C . Analysis of total liver RNA was performed by Northern (RNA) blotting as previously described (75). Briefly, $40\ \mu\text{g}$ of RNA was denatured by glyoxal and dimethyl sulfoxide, run on a 1% agarose gel in 20 mM phosphate (pH 7.2), and blotted onto Hybond N⁺ in 0.05 N NaOH. The hybridization protocol and the probes were as for Southern blots. Arctic ground squirrel sera were assayed for the presence of hepadnavirus DNA by dot blot hybridization with a labeled GSHV DNA probe as described previously (57).

Amplification and sequencing of the ASHV genome. To amplify overlapping fragments covering the entire arctic squirrel hepatitis virus (ASHV) genome, we prepared 10 pairs of oligonucleotide primers conserved between the sequences of GSHV and various WHV isolates. With the use of these primers, 10 viral fragments were amplified from liver DNA of squirrel 90-14 by PCR as described by Saiki et al. (53). PCR products were isolated in low-melting-point agarose gels and purified by using a Jetsoorb gel extraction kit (Genomed). The primers were 5' end labeled by using T4 polynucleotide kinase and [γ - ^{35}P]ATP, and direct sequencing was performed with Vent (exo⁻) DNA polymerase, using a Circum-Vent thermal cycle dideoxy DNA sequencing kit (Biolabs) according to the manufacturer's instructions. Additional primers were then synthesized according to the determined ASHV sequences to fill up the remaining gaps. The sense oligonucleotides used were as follows: 1s, 5' GGTGGACACTGCTGCTGC 3' (nucleotides 180 to 197 on the ASHV map); 2s, 5' TGCTCATGTCAATGACAC 3' (348 to 365); 3s, 5' CTCTCCGGAACATACAG 3' (515 to 532); 4s, 5' AG CTCCAAGTGTGATC 3' (639 to 656); 5s, 5' TGTGGGAATCAGGAATCC 3' (960 to 977); 6s, 5' CTGGGAACACAGACAGC 3' (1029 to 1046); 7s, 5' TCAGCAAGAACTGGAGGC 3' (1211 to 1228); 8s, 5' AGTACCACCTCAGA CTCC 3' (1403 to 1420); 9s, 5' ATCCCGAGTGCCTGGC 3' (1800 to 1817); 10s, 5' TGATTTGGTTTGGGGGC 3' (2251 to 2268); 11s, 5' CTGTGTTT GCTGACGCAA 3' (2680 to 2696); 12s, 5' TGCAGACTTGCGAACCAT 3' (3099 to 3116); and 13s, 5' ACCACTGTTTGTATTAGG 3' (3249 to 3266). The antisense oligonucleotides used were as follows: 1as, 5' CAGGTGAAAAAGA TACAT 3' (18 to 1); 2as, 5' AGGACATGGAACACAGGC 3' (48 to 31); 3as, 5' AAAGCTTGTCTGATAGCT 3' (266 to 249); 4as, 5' GCGAGGAGAGGG AGTGCG 3' (597 to 580); 5as, 5' AGGATTCCTGATCCAC 3' (978 to 961); 6as, 5' ACTTTTATGTTGTTGCC 3' (1078 to 1061); 7as, 5' TCAGCAAGAA CTGGAGGC 3' (1228 to 1211); 8as, 5' TGTAGTAAAAGGATCTCC 3' (1595 to 1578); 9as, 5' AACGCCGAGATACATCC 3' (1905 to 1888); 10as, 5' AAA CAGAGCAAATATGGG 3' (2318 to 2301); 11as, 5' CCGGAAAGGAGCTGA CAG 3' (2742 to 2725); and 12as, 5' GTTCACGGTGGAAATCCAT 3' (3132 to 3115).

Sequence data analysis. Multiple alignments of hepadnavirus sequences were performed by using the CLUSTAL W program (66). Aligned sequences were checked by eye, and gaps introduced by the outgroup sequence (DHBV) were removed. The resulting files were submitted to different programs of the PHYLIP package version 3.52c (available from Joe Felsenstein, University of Washington). To test the reliability of the final tree topology, the bootstrap technique was used. For this purpose, the SEQBOOT program was used to generate 100 data sets that are random resampled versions of the previously aligned sequences. Two different methods, (i) maximum parsimony using the PROTPARS program and (ii) neighbor joining, were used to generate phylogenetic trees. To apply the latter method, distance matrices from each of the 100 replicated data sets were first computed by using the PROTDIST program with

the Kimura two-parameter model. The obtained matrices were then used as input to the NEIGHBOR program (54). From each of the two previously used methods, a consensus tree was constructed by using the CONSENSE program with the majority rule criteria, and the drawing tree program TreeTool (available from Mike Maciukenas, Department of Microbiology, University of Illinois) was used to plot the obtained consensus tree.

Western blotting (immunoblotting) and indirect immunofluorescence. Different dilutions of squirrel sera in $30\ \mu\text{l}$ of Laemmli buffer were loaded on 15% polyacrylamide gels. HBV-positive human sera and WHV-positive woodchuck sera were analyzed in parallel, and the Rainbow colored marker (Amersham) was added as a size marker. After electrophoresis, the gels were blotted to a polyvinylidene fluoride membrane (Porablot; Macherey-Nagel, Düren, Germany). The membranes were blocked in 3% dried milk-2% glycine-0.01% Na₂S₂O₅ in phosphate-buffered saline (PBS) for 1 h at room temperature and incubated with polyclonal anti-WHV small surface protein (WHs) antibodies (60) at a 1:1,000 dilution in blockage buffer supplemented with 10% calf serum and 0.1% Tween for 90 min at room temperature. They were then incubated with horseradish peroxidase-conjugated donkey anti-rabbit secondary antibodies (Amersham Life Sciences) at a dilution of 1:15,000. Visualization of signal was by enhanced chemiluminescence (ECL kit; Amersham). In other experiments, the membranes were incubated for 1 h at 37°C with anti-HBV small surface antigen (HBsAg), anti-HBV core antigen (HBcAg), or anti-HBV pre-S1 or pre-S2 monoclonal or polyclonal antibodies (73), washed in blockage buffer, then incubated with peroxidase-labeled secondary antibodies for 30 min at 37°C , and revealed with diaminobenzidine in PBS.

Frozen squirrel liver tissues were processed as described previously (74). Sections $4\ \mu\text{m}$ thick were fixed in anesthetic ether at -20°C for 60 s, air dried, and incubated with mouse anti-HBcAg monoclonal antibodies (MAbs) for 30 min. After three washes in PBS, a second incubation was carried out with fluorescein isothiocyanate-conjugated anti-mouse antibodies for 30 min at room temperature. After three washes in PBS, the sections were mounted in 50% glycerol in PBS and photographed on a Leitz Dialux 20 microscope under UV light.

RESULTS

Identification of hepadnavirus-related nucleic acids in arctic squirrel livers and tumors. To search for a hepadnavirus infecting arctic ground squirrels, liver DNA from eight animals bearing grossly recognizable hepatic nodules was analyzed by Southern blot hybridization, using *Bgl*II-digested genomic DNA and a ^{32}P -labeled GSHV DNA probe. As shown in Fig. 1A, a pattern typical of hepadnavirus replicative intermediates was found in liver DNA from three animals, including a discrete band at 3.3 kb corresponding to full-length double-stranded viral DNA and a long smear beneath 3.3 kb representing incomplete double-stranded and single-stranded forms. Identical patterns were obtained when samples from nodules and adjacent normal livers were analyzed in parallel (Fig. 1A, lanes with suffixes a and b). After longer exposure of the autoradiogram, an additional sample (SPK14) yielded a faint signal, and identical results were obtained with a WHV DNA probe (data not shown). The amounts of viral DNA were roughly estimated to range from 100 to 1,000 genome equivalents per cell by running in parallel serial dilutions of cloned GSHV and WHV DNAs and quantifying the hybridization signals with a PhosphorImager (data not shown). Northern blot hybridization of total liver RNA from the same animals with a GSHV DNA probe showed two abundant transcripts of 2.1 and 3.7 kb, identical in size to the major WHV and GSHV mRNAs in chronically infected livers from woodchucks and Beechey ground squirrels (Fig. 1B), confirming hepadnavirus replication in the DNA-positive samples.

We next attempted to evaluate the prevalence of hepadnavirus infection in wild arctic ground squirrels trapped in the Brooks range in Alaska. Liver tissues from 38 randomly selected animals trapped from 1988 to 1995 and sera from a different panel of 10 animals were analyzed by dot blot hybridization with GSHV DNA. Signals were observed with three liver DNAs and one serum sample (Fig. 2). Therefore, of a total of 56 animals examined, 8 (14%) were found to be positive for hepadnavirus DNA. These data indicate that a virus

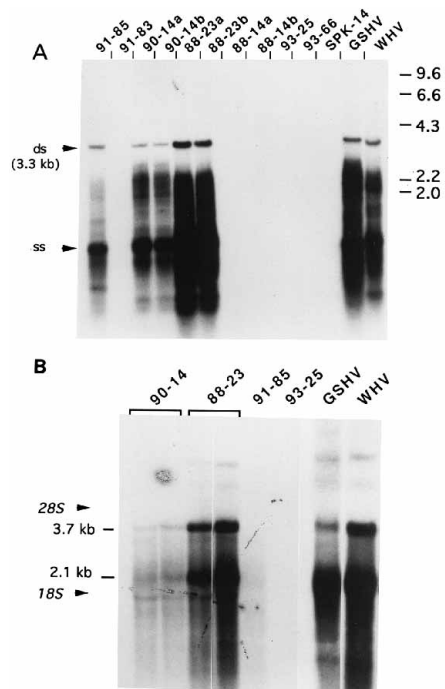


FIG. 1. Detection of hepadnavirus-related sequences in arctic squirrel livers. (A) *Bgl*II-digested genomic DNA from eight arctic squirrel livers (numbered above each lane), a GSHV-positive Beechey ground squirrel liver (lane GSHV), and a WHV-positive woodchuck liver (lane WHV) were hybridized with a GSHV probe. Suffices a and b denote normal liver and liver nodules, respectively. The positions of full-length double-stranded (ds) and single-stranded (ss) viral DNAs are indicated on the left. Sizes are in kilobase pairs. (B) Northern blot hybridization of total liver RNA from the same animals with GSHV DNA. The positions of rRNAs and the sizes of major hepadnavirus transcripts are shown on the left.

similar to GSHV and WHV is endemic in wild arctic ground squirrels in Alaska. This virus was termed ASHV.

Viral proteins in arctic squirrel sera and liver tissues. Evidence that the surface and core antigens of HBV, WHV, and GSHV are antigenically cross-reactive has been previously obtained by using both MAbs and polyclonal antibodies (8, 12, 13, 56). We first analyzed the proteins of an ASHV DNA-positive arctic squirrel serum (AS4 [Fig. 2]) by SDS-polyacrylamide gel electrophoresis and immunoblotting with polyclonal antibodies directed either to the central domain (residues 96 to 166) of WHs or to the total WHV pre-S2 region (60). As shown in Fig. 3A, the anti-WHs antiserum recognized a p24-gp27 doublet corresponding to unglycosylated and glycosylated forms of the small surface protein in the ASHV-positive serum AS4. No signal was observed in a normal mouse serum or in the ASHV DNA-negative arctic squirrel serum AS5 (Fig. 3A) or when a preimmune serum was used (data not shown). In contrast, the anti-WHV pre-S2 antiserum was ineffective in recognizing the large and middle surface proteins of ASHV (Fig. 3A). In a WHV-positive woodchuck serum, these antibodies detected a p45-gp47 doublet representing different glycosylated states of the large surface protein (also called the pre-S1 protein) and a weaker 36-kDa polypeptide identified with the glycosylated form of the middle surface protein (60). We then used a panel of MAbs and polyclonal antibodies raised against HBsAg, the pre-S2 region, the pre-S1 region, and HBeAg. As shown in Fig. 3B, faint bands were revealed in the squirrel serum AS4, at 24 kDa with the anti-HBsAg antibodies (left panel, AS4 lane 2) and at 21 kDa with the anti-

HBeAg antibodies (right panel, AS4 lane+), and surprisingly, a stronger band at 45 kDa was seen with the anti-pre-S1 antibodies (left panel, AS4 lane 4), indicating cross-reactivity with the homologous ASHV surface and core proteins. Similar bands were detected by these antibodies in the serum of a WHV carrier woodchuck (lanes W+), but the anti-pre-S2 MAb revealed no signal in arctic squirrel or woodchuck sera.

Different subcellular localizations have been previously reported for the hepadnavirus core proteins in chronic carrier hosts: proteins lacking the precore sequences accumulate preferentially in the cytoplasm of WHV- and DHBV-infected hepatocytes and predominantly in the nuclei of HBV-infected cells (21). The cellular localization of ASHV core antigens was assessed in frozen liver sections of ASHV DNA-positive arctic squirrels by indirect immunofluorescence, using anti-HBeAg MAbs known to recognize the WHV core antigen (73). As observed previously for woodchuck hepatitis (18), cytoplasmic staining was observed in arctic squirrel 91-85 hepatocytes (Fig. 3C), whereas control liver from a noninfected woodchuck yielded no signal (data not shown).

Sequencing of the ASHV genome and comparison with GSHV and WHV genomes. The entire nucleotide sequence of ASHV in one of the strongly positive livers (90-14 [Fig. 1]) was determined by direct sequencing of overlapping viral fragments amplified by PCR. The oligonucleotide primers used for amplification and/or sequencing are listed in Materials and Methods, and the amplification and sequencing strategy is illustrated in Fig. 4A. Both strands of the ASHV DNA were completely sequenced. A partial sequence was also obtained by the same method from a second infected squirrel liver (88-23) to confirm some particular features of the sequence. As shown in Fig. 4B, the ASHV genome consists of 3,302 bp. Computer-based alignment of the ASHV sequence with those of other mammalian hepadnavirus genomes showed a conserved genomic organization into four long open reading frames (ORFs) in the same transcriptional orientation, referred to as pre-C/C, pre-S/S, *pol*, and X (Fig. 4A). Overall, the nucleotide sequence variation was 16% between ASHV and GSHV and ranged from 15.4 to 16% between ASHV and various WHV isolates. Compared with the HBV ayw genome, ASHV showed a nucleotide sequence variability of 37%, like that of GSHV and WHV. The 11-bp directly repeated sequences (DRs) designated DR1 and DR2, which play a critical role in viral DNA replication, are strictly identical among ASHV (positions 11 to 21 and 3091 to 3101), GSHV, and WHV. They differ from the

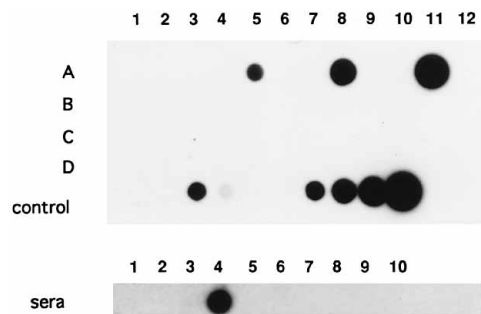


FIG. 2. Dot blot analysis of liver DNA and serum samples from randomly selected arctic squirrels. Four micrograms of total liver DNA from 38 animals (A to D) or 50 μ l of serum from a different panel of 10 squirrels (bottom) was hybridized with a labeled GSHV DNA probe. Lanes 1 and 2, mouse DNA (control); lanes 3 to 5, serial dilutions of liver DNA from a WHV carrier woodchuck; lanes 7 to 10, serial dilutions of liver DNA from a GSHV carrier Beechey ground squirrel.

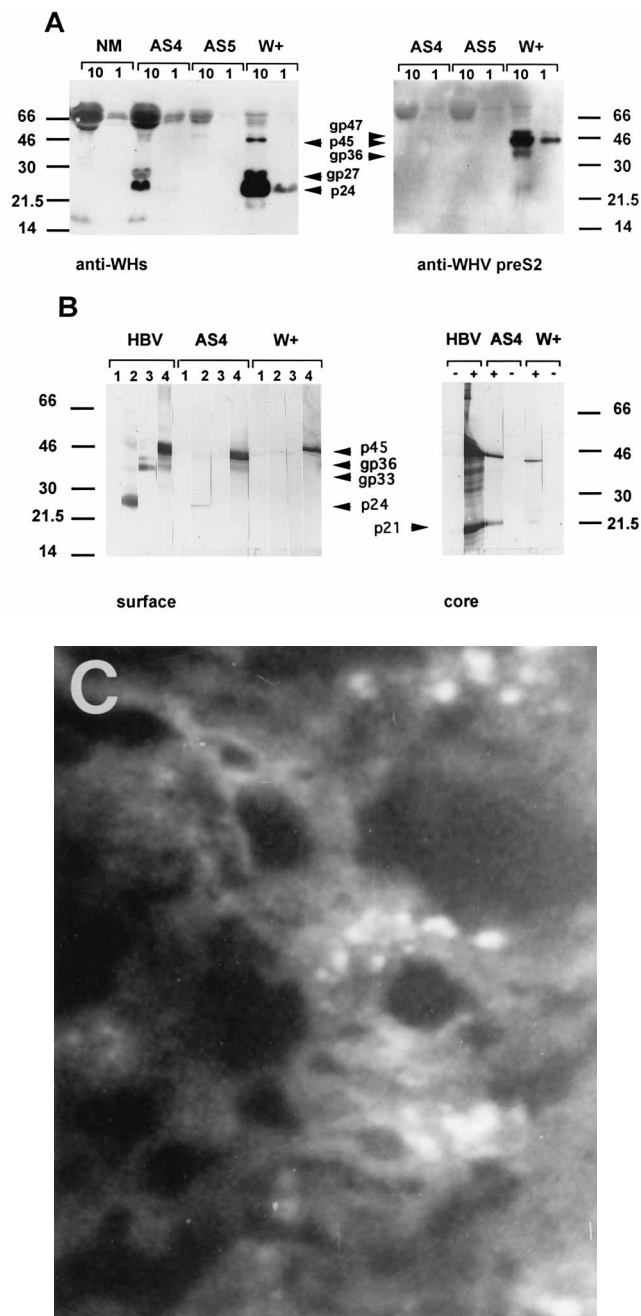


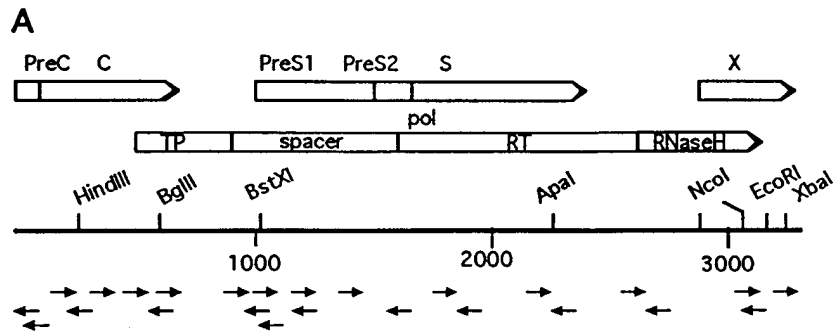
FIG. 3. Western blot analysis of viral proteins in an ASHV DNA-positive serum sample, using various MAbs and polyclonal antibodies. (A) Two dilutions of serum from a normal mouse (lanes NM), ASHV DNA-positive (lanes AS4) and -negative (lanes AS5) arctic squirrels, and a WHV carrier woodchuck (lanes W+) were analyzed by using rabbit polyclonal antisera raised against WHs (left panel) and the WHV pre-S2 region (right panel). For each sample, 3 and 0.3 μ l of serum were loaded in lanes 10 and 1 (0.3 and 0.03 μ l for lanes W+). The sizes of viral surface antigens are indicated by arrowheads. Positions (in kilodaltons) of molecular weight standards are indicated on each side. Detection was by enhanced chemiluminescence. (B) One-microliter aliquots of sera from an HBV carrier patient (lanes HBV), an ASHV-positive arctic squirrel (lanes AS4), and a WHV carrier woodchuck (lanes W+) were resolved by SDS-polyacrylamide gel electrophoresis, and Western blots were reacted with different antibodies and detected with diaminobenzidine. Lanes in the left panel: 1, no antibodies; 2, anti-HBsAg MAbs; 3, anti-HBV pre-S2 MAbs; 4, anti-HBV pre-S1 MAbs. Lanes in the right panel: +, rabbit polyclonal anti-HBsAg antiserum; -, preimmune serum. (C) Immunocytochemical localization by indirect immunofluorescence of ASHV core proteins in the cytoplasm of infected hepatocytes on a frozen liver section (animal 91-85). Magnification, $\times 400$.

HBV DRs only by one base substitution. It is notable that ASHV, like GSHV, carries an 11-bp noncoding region between the end of the X gene and the start of the pre-C region (positions 3292 to 3302).

The hepadnavirus pre-C/C ORF specifies the core and e antigens, initiated at two in-frame translation initiation codons. Two in-frame AUGs homologous to the precore and core AUGs of WHV and GSHV were found in ASHV (Fig. 4B). A difference from the WHV genome was the absence of an additional in-frame AUG located 21 bp upstream of the precore initiation codon in all established WHV isolates. The precore region (positions 1 to 90) is the sequence best conserved among hepadnaviruses (39). The ASHV pre-C region differs from that of GSHV and of the WHV isolate NY (28) only by one nucleotide substitution (T-C at position 42), resulting in one amino acid change (Ser to Pro), the latter corresponding to a residue found in most HBV subtypes. The predicted product of the pre-C/C ORF is highly homologous to its GSHV and WHV counterparts (Table 1). In fact, most amino acid positions in ASHV correspond to residues found in either GSHV or WHV or both. The ASHV C gene differs from either the GSHV or the WHV C gene at 28 nucleotide positions. The core protein of ASHV (217 amino acids) is distinguished by 10 specific residues and by the deletion of one arginine in the carboxy-terminal SPRRRR motif common to all mammalian hepadnaviruses.

Another remarkably conserved region among ASHV and other hepadnaviruses is the S gene, which encodes the small surface protein. The ASHV S gene differs from its GSHV and WHV counterparts at 36 (6%) and 46 (7%) nucleotide positions, and the predicted ASHV S protein, 222 amino acids in length like the corresponding GSHV and WHV proteins, is more closely related to the GSHV than to the WHV surface antigen (Table 1). Upstream of the S gene, two in-phase AUGs delimit the pre-S1 and pre-S2 regions that encode the amino-terminal parts of the large and middle surface proteins. This region of the ASHV genome is clearly more divergent than the C and S genes (Table 1). It consists of 615 nucleotides coding for 145 pre-S1 and 60 pre-S2 amino acid residues. Compared with WHV and GSHV, the ASHV pre-S1 domain shows a unique 12-bp deletion (position 1264). This position does not coincide with the regions deleted in GSHV compared with WHV8 or in WHV1 compared with WHV8 (19, 23, 59). The 3-bp insertion present in GSHV sequences was not found in ASHV (position 1152 in Fig. 4B). The predicted polypeptide encoded by the ASHV pre-S1 region expresses 20 (14%) specific amino acids and differs from the homologous GSHV and WHV sequences by 33 (23%) and 27 (19%) residues, whereas the translation product of the ASHV pre-S2 region differs from that of GSHV at 10 (17%) residues and from that of WHV at 15 (25%) residues. The glycosylation site common to all mammalian hepadnaviruses (Asn-Gln/Ser-Thr/Ser) is conserved in ASHV (residues 148 to 150).

In the *pol* gene product, the amino-terminal domain encoding the genome-linked protein is separated from the reverse transcriptase and RNase H domains by a spacer region of unknown function. Again, we found a strong conservation of each functional domain in ASHV *pol* compared with other rodent hepadnaviruses (Table 1), whereas the spacer region is much more variable, with only 46.5 and 58% amino acid identities with GSHV and WHV. Two specific deletions of 3 bp were found in the ASHV *pol* gene (positions 2579 and 2687), resulting in one amino acid deletion in the reverse transcriptase domain and another in the RNase H domain. The first deletion also affects the region homologous to the core of the HBV enhancer I, specifically the E element, which is perfectly con-



B

```

↗ preC -----DR1-----                               ↘ c 100
ATGTATCTTTTTCACCTGTGCCTTGTTTTTGGCCTGTGTTTCATGTCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGATGGCTTTGGGACATGGACATAG
200
ATCCCTATAAAGAATTTGGTTCATCCTACCAGTTGTGAATTTTCTCCTTTGGACTTCTTTCCTGAACCTCAATGCCTTGGTGGACACTGCTACTGCTCT
300
CTATGAAGAAGAATTAACAGGTAGGGAGCACTGCTCTCCTCATCACACAGCTATCAGACAAGCTTTAGTTTGTCTGGGAAGAATTAACAAGATTAATTGGC
400
TGGATGAGTGCTAACATTAATTCAGAAGAAGTAAGAAGAGTTATAGTTGCTCATGTCAATGACACTTGGGGACTTAAAGTTAGGCAGAAATTTATGGTTTC
500
ACTTATCCTGTCTGACTTTTGGGCAACACACAGTGCAGGAATTTTGTAGCAGCTTTGGAGTAAGGATCAGAACTCCGGCTCCTTATAGACCTCCTAATGC
600
ACCCATTCTCTCAACTCTCCGGAACATACAGTCATTAGGAGAAGAGGAAGTGAAGAGTTGTTAGTCCCCAGAAAGCGCACTCCCTCTCTCCGACAG
700
AGAAGATCTCAATCACCCGCTCGCAGGCCTCAATCTCCAGCTTCCAACCTGCTGATCTTGATTGGGTGCACTAAACTAATGCTATAACGGGCTTTTATTCT
800
ACTCAGACAGCTAAGTTAATCTGAATGGAAACAACCTGATTTTCCAAAAATTCACCTGTCTGAAGATTTATTTCTAAACTACAACAATTTTGTGGTCT
900
CTCTTACAGTTAATGAAAAAGGAAATTAATAATTAATTTTCTGCTAGGTTTTTCCCAAGGCTACTAAATATTTCCCTTCCCAAGGAAATAAAAA
1000
TAATATCTGATTTCTCTATAGAACACTTTTTGTCAGCTGCAACTATTTATGGACTTTTGGGGAATCAGGAATCTGTATTTGAGGAAAAATCAAAC
1100
ACTCTCACTTTTAAAGGTAACCATATTTCTTGGGACACAGACAGCTAGAGCAACATAATGGGCAACAACATGAAAGTCACTTCAATCCAGAGAAAGTA
1200
GCAGCATGGTGGCCAGCAGTGGGCACATATTACACAACACAGCAGCCCTCAGGACCCTCAGTTTCCCAACCAGGGATTACCAAAACAACCTTCTTGGTGA
1300
ATCCAAAAATCAGCAAGAAGTGGAGGCAGTGTAGAGAAAAGATACAAACAAATAGATTGGGATCCCTGGTAAATCAAAAATACCATTGGTCAGCAG
1400
GGTTCCTCCCAAGTCTCCTCCCAAGATCAAAGAGCTCAAACCTTTCGAAATCAAAACCCAGGCAATCATAGTTCTCGGAATCAGAGACATCCCACTGGT
1500
ATAGTACCACCTCAAACACACCCAATCGAGACAAGGGAGGAAACCTACTCCTCAGACTCCGCTTTAAGAGACTCACCCTCAATTTGAACATGAAAA
1600
ATCAGAGCCCTCATCTTCAGGGCTTTGCGGAGGGACTGAGAGTCTTAAACAACCCGGAACATCAACATTTGCTTGGAGATCCTTTACAACACTGAG
1700
CCCTGTTGGTCTACTGTCTCCACCACATGTCTCCTCCTTGAAGATTTGGGACCCCTGTACAATCAGCGGAGATGCACCATCAGTCTCCTAGGACTC
1800
CTCGCCGATTACAGGTGGTATTTCTTGTGGACAAACATCCTCACAATAGCTCAGAGTCTCGATTGGTGGTGGACTTCTCTCAGTTTCCAGGGGGGA
1900
TACCAGAGTGCATGGCCAAAATTTGCAGTTCTTAACTTGAAGCACTTGCCAACTCTTGTCCACCAACTGCAATGGCTTTCGCTGGATGTATCTGGC
2000
GCGTTTTATCATATACCTGTAGTCTGCTGCTGCTGCTCATCTTCTGTTGGTTCTCCTGGACTGAAAGGTTTACTCCAGTATGTCCATACAACCA
2100
TCCACGGAAACAACAGTAAATTCAGACAATGCACAATCTCTGCTCAAAGACACCTTTTCAACTCCTTACTGTTGTTTAAAAACCTACGGCAGGAAAT
2200
GCACCTGTGGCCCATCCCTTCATCATGGGCTTTAGGAAGCTACCTATGGGAGTGGCCCTTAGTCCGTTTCTCTTGGCTCAGTTTACTAGTCCCTTGT
2300
TCAATGGTTAGGAGGAATTTCCCTCACTGTGTGGCTTTTGCCTTATATGGATGACTTGGTTTTGGGGCCGCTACTCATGAGCATCTTACCGCCATTTATT
2400
CCCATATTTGCTCTGTTTTTCTGATTTGGGCATACATTTAAATGTAGCCAAAATAAATGGTGGGACATCAATTTACATTTTCAATGGGTTATGTTATTAC
2500
TGGAGCAGGAATTTTACCCCAAGATAAACAATGTGCAAAAAGTATCAACATATTTGAAATCCATTCACCTCAACAAAACCTTTAGATTATAAAATCTGTGAA
2600
AGGTTAACAGGCATCTGAAATTAATGTCTCCTTTTACTAAATGTGGTATGCTGCTCCTTCTCCTTTGATCAAGCTACTTCGCGTACGGCATTTGTGT
2700
TTTCTCTCTACCACAGCTGGTGTCTCCCTTTATGCTGAGTTGTGGCCTGTGCCAGGCAACGTGGCGTGGTGTCTGTCTGACGCAACCCC
2800
CACTGGTTGGGCATTTGACACCCTATCAACTCATTTCCTCCGAGGGGCGCTTTTCCCTGCGCATCGCCACCGGGAGCTCATCGCCGCTGCCTTGTCT
2900
CGCTGCTGGACAGGACTCGGCTTTGGGCACTGACAACCTCCGTGGTCTTTTCCGGCAAACTGACTTCTTATCCATGGCTGCTCGCTGTGTGCCAACT
3000
GGATTCTCGGGGACGTGTTCTGCTACGTCCCTTCGGCAGCGAATCCGGCGGACCTGCGCTCTGAGGCGCTTCTGCGGCTCTGCATCCCGTGGCGAC
3100
TCTCCGCTTCGCTCCGAGCTGAGTCGCATCTCCCTTTGGGCGCGCTCCCGCCTGTTTCCGCTCGCCGCTCGGTCGGTGTGCTGGCTGGCTTCACTGTG
3200
CAGAATTCGAGCCATGGTTCCACCGTGAACCTTTGCTACTTGGCATGCAAAAACGCCAACTGGGCATGCCAAACAAGACCTTTGGACTCCTTATGTAAG
3300
AAATCAATTATTAACCAAATGGGAGGAGGTTACTATTGATTCTAGATTACCCTGTTTGTATTAGGGGCTGTAGGCATAAATACATGTAACCTGCCGCAA
TC

```

FIG. 4. Complete nucleotide sequence of the ASHV genome. (A) Schematic representation of the ASHV genome linearized at the pre-C initiation codon and localization of the oligonucleotide primers used for amplification and/or sequencing. (B) Nucleotide sequence, showing the 5' and 3' ends of the viral ORFs. The translation initiation codons are doubly underlined; the stop codons and polyadenylation site are underlined. DR1 and DR2 are indicated by dotted lines.

TABLE 1. Percent amino acid homologies of rodent hepadnavirus genes^a

Viruses compared	% Homology							
	Total	Pre-C/C	Pre-S	S	pol ^b			X
					Primase	RT	RNase H	
ASHV and GSHV	81.5	91.2	77.6	96.4	81.2	88.2	84.1	70.3
ASHV and WHV	79.8	89.9	76.1	90.5	81.8	86.8	82.3	71.7
GSHV and WHV	81.0	92.2	73.6	90.6	80.1	85	91.1	70.4

^a The predicted amino acid sequences of different ORFs from ASHV, GSHV (59), and WHV 8 (23) were compared by using the GAP program of the Genetics Computer Group sequence analysis software (9).

^b The hepadnavirus *pol* gene encodes a multidomain protein and can be segmented into three functional regions: primase or terminal protein (residues 1 to 181 on the ASHV map), reverse transcriptase (RT) (residues 370 to 714), and RNase H (residues 715 to 887).

served among mammalian hepadnaviruses (11). The same deletions were also found in a partial ASHV sequence derived from a second infected animal (data not shown), indicating that they represent an intrinsic feature of the ASHV genome.

The smallest ORF of mammalian hepadnavirus genomes, named X, encodes a promiscuous transcriptional activator, and it is required to establish a productive viral infection *in vivo* (4, 7, 71). The ASHV X gene is identical in size to the GSHV X gene (414 bp), from which it differs at 75 (18%) nucleotide positions. It is smaller than the WHV X gene (423 bp), but the two sequences show only 15% nucleotide exchanges. Compared with the GSHV and WHV X proteins (GSHx and WHx), the predicted X protein of ASHV (ASHx) has more amino acid variations than any other viral gene product (Table 1). As shown in Fig. 5, alignment of the mammalian hepadnavirus X proteins confirms and extends previous observations that highly conserved regions are located in two hydrophobic sequences encompassing amino-terminal residues 1 to 20, carboxy-terminal residues 124 to 136, and a moderately conserved, charged region in the central part. Internal start codons which might serve for translation of two small polypeptides (residues 81 and 98), and the regions suggested to be homologous to the Kunitz domain of Kunitz-type protease inhibitors (residues 67 to 69 and 128 to 136) (64) are also remarkably conserved in ASHx.

Phylogenetic analysis. To determine the phylogenetic relationship of ASHV to other hepadnaviruses, ASHV sequences were compared with previously published hepadnavirus sequences by using two different methods (maximum parsimony and neighbor joining) of the PHYLIP package version 3.52c. Phylogenetic trees were constructed on the entire genomes of ASHV, 1 GSHV clone (59), 6 WHV clones (6, 19, 23, 28, 30) and 11 HBV clones of various serotypes (20, 29, 31, 42, 45, 47, 51, 52, 55, 72) and were rooted by using DHBV as an outgroup

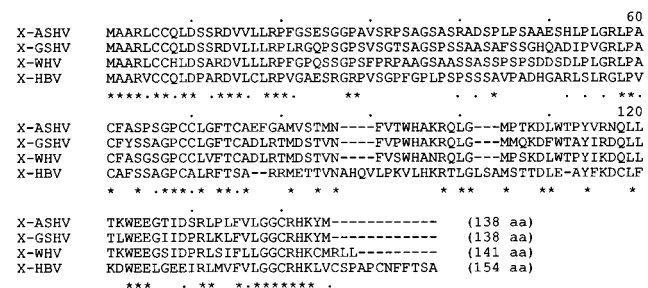


FIG. 5. Alignment of the hepadnavirus X proteins. Identical residues between the four X sequences are denoted by asterisks, and conservative changes are shown by dots under the corresponding positions.

(33). The derived topologies showed two separate virus phylogroups (human and rodent viruses) and joined ASHV with GSHV in a genomic cluster distinct from that of the WHV isolates (Fig. 6). WHV subgroups correlated with the geographical origins of the virus-infected woodchucks, and HBV genomes were clustered according to previously characterized genotypes (45, 48). However, bootstrap support for the GSHV/ASHV cluster was relatively low (67%), whereas statistical support for the other nodes of the trees was high (>90%). We therefore reconstructed phylogenetic trees on the nucleotide sequences of ASHV and GSHV isolates and only one WHV isolate (WHV8), using HBV ayw as an outgroup. In this case, the phylogenetic position of ASHV was clearly much closer to GSHV than to WHV, with a bootstrap support of 98% (data not shown). We also analyzed the protein products of different ORFs of the rodent hepadnaviruses, including the C, S, P, and X genes and the pre-S region. The topological features of the derived trees were virtually identical to those obtained with the entire genomes, except for the pre-S region and the X gene

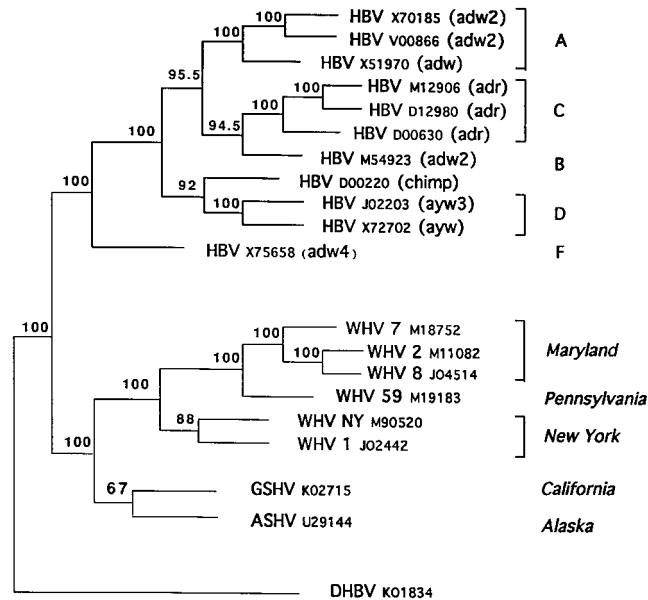


FIG. 6. Phylogenetic tree of mammalian hepadnaviruses based on complete genomic sequences of 11 HBV isolates, 6 WHV isolates, 1 GSHV isolate, and ASHV. A rooted consensus tree (DHBV being used as the outgroup) was obtained from 100 bootstrapped samples of the aligned sequences, using PROTPARS and CONSENSE programs from the PHYLIP package version 3.52c. At each node is indicated the number of its occurrences out of 100 trees obtained from the bootstrapped samples. Each sequence is designated by its accession number in databanks and, for HBV, by its serotype. HBV genotypes (A to F) and the geographical origins of the different rodent hosts are indicated on the right.

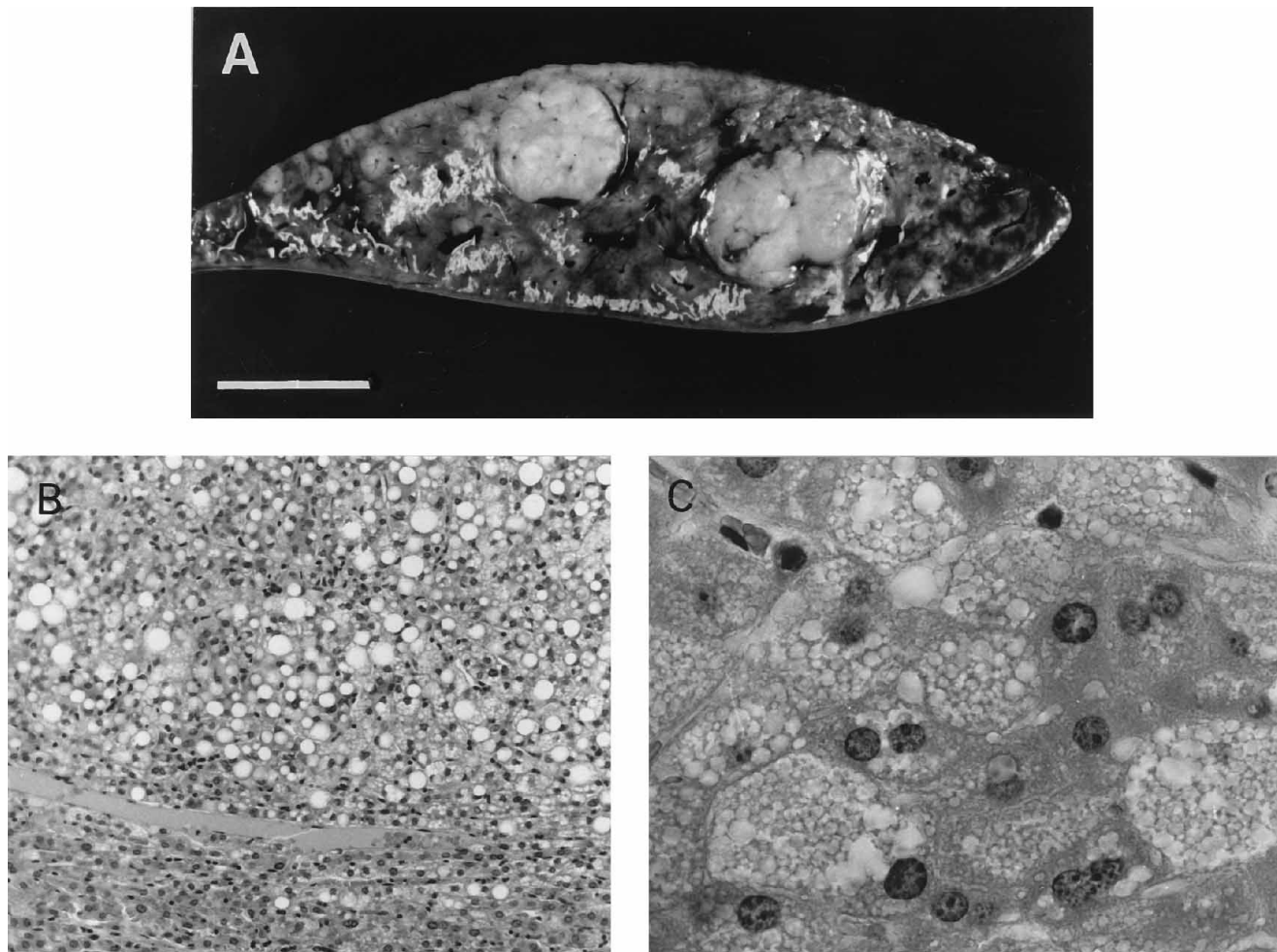


FIG. 7. Large hepatic nodules in a GSHV-infected arctic ground squirrel. (A) Gross picture of a liver section from squirrel 91-85, showing two hepatocellular adenomas. Bar = 1 cm. (B) Typical adenoma corresponding to one of the nodules shown in panel A. Magnification, $\times 84$. (C) Higher magnification ($\times 756$) showing microvesicular (spongiocytic) steatosis and normal nuclei.

(the most divergent regions within hepadnavirus genomes), for which ASHV and GSHV were placed in different clusters (data not shown).

Liver lesions in arctic ground squirrels. During postmortem examination of arctic squirrels kept at the Institute of Arctic Biology, pale yellowish liver nodules of various sizes were observed in about one-third of old animals (at 3 to 4 years of age) but rarely in young, newly trapped squirrels. Histological examinations were performed on fixed liver sections from six animals showing grossly recognizable mass lesions. In parallel blot hybridization studies (Fig. 1 and 2), three of these animals were found to be ASHV DNA positive, and the others were negative. A drastic fatty infiltration of the liver parenchyma consisting of either vacuolar, macrovesicular, or microvesicular steatosis was observed in all samples, as illustrated in Fig. 7. Fat accumulation was demonstrated by osmium tetroxide staining (data not shown). Portal or periportal lymphocytic infiltrations without necrosis or cirrhosis were occasionally seen. The livers from three animals each carried two to three nodules, about 1 cm in diameter, as shown in Fig. 7A. Two tumorous nodules developed in an ASHV DNA-positive animal (91-85 [Fig. 1]) were diagnosed as hepatocellular adenomas (Fig. 7B). The nodules were well demarcated from surrounding liver, without capsule formation, and the hepatocytes

were markedly enlarged and showed accumulation of fat droplets. Cells were arranged in cords and plates of varying thickness; sinusoids were irregular, but their differentiation was maintained; portal triads and central veins were absent. The nuclei were regular in size, but cytoplasm was enlarged to various extents, showing sometimes significant microvesicular steatosis (Fig. 7C). Silver impregnation showed marked thinning of the reticulin framework in some regions (data not shown). Adenomas and a well-differentiated HCC were observed in two ASHV DNA-negative squirrels. The HCC was trabecular and highly differentiated as usually seen in woodchucks and Beechey ground squirrels (data not shown). In liver samples from the three remaining animals examined, hepatocytes accumulated large amounts of fat, but none of the typical features of benign or malignant tumors were seen (data not shown).

DISCUSSION

A search for hepadnavirus-related DNA in liver or serum samples from arctic ground squirrels trapped during the last 7 years in Alaska revealed a 14% prevalence of hepadnavirus infection of the rodent species in this region. Hepadnavirus infections have been previously found in wild ground squirrels

and woodchucks caught at several distinct geographical locations in the United States and Canada (36, 63, 65). The present data further emphasize the importance of hepadnavirus spread in related species of the Sciuridae family throughout North America.

Analysis of the complete nucleotide sequence and deduced amino acid sequence of the ASHV genome has pinpointed its close relatedness with GSHV and WHV. These viruses are similar in genome size and virtually identical in genetic organization. The regions best conserved among the three rodent viruses, including the pre-C region, the C and S genes, and the functional domains of the P gene, are also well conserved over evolution of the different members of the hepadnavirus family, presumably as a result of high functional and structural constraints (30). We found a greater variability and distinctive deletions in the ASHV pre-S region, as previously noted when different hepadnavirus genomes were inspected (6, 45, 62). The pre-S domain of the large surface protein has been implicated in the binding of viral particles to the hepatocyte membrane, and these sequences could determine the restricted host range of hepadnaviruses (43). It has been shown that woodchucks can be infected with GSHV (58). To better define the hepatocyte attachment sites within pre-S sequences of rodent viruses, it will be interesting to test the infectivity of ASHV in woodchucks as well as that of WHV in arctic ground squirrels. Finally, the most variable amino acid sequences in ASHV were found in the X gene, despite partial overlap with the well conserved RNase H domain. Similar data were obtained when X proteins were compared among different mammalian hepadnaviruses and between WHV isolates (6, 30), suggesting that the structure of this protein might be rather flexible. Notably, ASHx and GSHx are shorter than other hepadnavirus X proteins. However, it has been shown that the short size of GSHx does not affect its transcriptional transactivator capacity in transient transfection assays (7). Whether hepadnavirus X protein variability results in altered affinities for different X-responsive elements in cellular promoters/enhancers is an interesting issue regarding a potential role of X in the oncogenicity of these viruses.

On the basis of genome sequences, ASHV differs from either WHV or GSHV to the same extent as these two viruses differ from each other (about 15% base exchanges). This is much more than the variability between different WHV isolates (<3.5% variation) (6) and even more than among different HBV genotypes (around 8 to 14% variation) (45) but less than the variability between two avian hepadnaviruses (21.6%) (62). This finding raises the question of the evolutionary relationships among the three rodent viruses. Our phylogenetic analysis indicates that ASHV and GSHV may be classified in the same phylogroup and suggests that the divergence of woodchuck and squirrel viruses predates that of GSHV and ASHV, according to the evolutionary speciation of host genera. The most parsimonious explanation for the similar genetic heterogeneity between arctic squirrel, California squirrel, and woodchuck viruses and for the relatively low statistical support at the ASHV/GSHV branch point of the tree is that GSHV and ASHV emerged very soon after the divergence of woodchuck and squirrel viruses. Nevertheless, the divergence time and the mode of spread of these viruses remain to be determined. Studies of the evolutionary relationship of reverse transcriptase-containing viruses have suggested that hepadnaviruses and retroviruses evolved from a common ancestor and that avian hepadnaviruses represent the primitive members of the *Hepadnaviridae* family, while the more recent emergence of mammalian viruses coincides with the acquisition of the X gene (39, 68, 76). It has been proposed that the hepadnavirus

family evolved independently of host species divergence. The divergence time between WHV and GSHV was estimated to be 10,000 years ago (48), whereas woodchucks and squirrels evolved more than 10 million years ago. However, markedly different estimates of the annual mutation rates of HBV and WHV genomes have been reported (24, 46). Further studies are required to clarify these issues.

It has been shown previously that GSHV and WHV display significant differences in the ability to induce hepatocarcinogenesis (37, 50) and that these differences depend on viral determinants (26, 58). Determination of the pathological and oncogenic potential of the closely related ASHV is therefore an important issue. In preliminary studies of the prevalence of ASHV infection in arctic squirrels, ASHV carriers animals were more frequently found among animals bearing gross hepatic nodules (4 of 8 cases) than in randomly selected animals (4 of 48 cases), strongly suggesting that persistent ASHV infection may increase the risk of developing mass lesions of the liver. However, as in early studies of GSHV-infected Beechey ground squirrels (22, 35), no evidence has been obtained for a strong association between persistent ASHV infection and the development of HCC in the natural host. Hepatocellular adenomas or carcinomas were seen in ASHV carrier and noncarrier arctic squirrels, and too few cases have been examined to allow estimation of the relative incidence. Further analysis of the serological status of ASHV DNA-negative animals is also needed to detect past resolved infections, considering that HCCs arise more frequently in postresolution squirrels and woodchucks than in marker-free animals (32, 34). The observation that liver tumors arise predominantly in aging arctic ground squirrels clearly indicates that this model differs markedly in tumor kinetics from the woodchuck model, in which a majority of WHV infections results in HCC development within 2 to 3 years of life.

An outstanding feature of arctic squirrel liver pathology is the occurrence of drastic accumulation of fat with frequent microvesicular (spongiocytic) steatosis, giving rise to grossly recognizable nodules. These lesions might represent a pathological response to seasonal development of obesity in hibernating mammals. In humans, however, microvesicular steatosis has been related to acute mitochondrial dysfunction induced by a variety of drugs (17) or associated with various viral agents. It was occasionally seen in chronic B and C hepatitis, and it is prevalent in fulminant hepatitis induced by a specific HDV strain in Central Africa and South America (49). Moreover, zonal fatty infiltration of the liver parenchyma has also been found in WHV carrier woodchucks (69), suggesting that viral agents might be implicated in this pathological process. Whether fat storage foci represent a specific preneoplastic cell population as postulated in the rat and whether ASHV infection plays a part in their development are important questions regarding the pathological properties of this virus.

Previous studies have highlighted the pivotal role of *myc* family genes in liver carcinogenesis induced by hepadnaviruses in rodent species. The strong oncogenicity of WHV has been related to the specific ability of this virus to provoke direct insertional activation of the *N-myc2* oncogene (15, 16, 75), and *c-myc* is frequently amplified in GSHV-induced tumors (26, 70). It will be important to determine whether *myc* genes are also implicated, through any mechanism, in arctic squirrel liver disease and cancer. The availability of a new rodent hepadnavirus provides an additional tool to investigate the molecular mechanisms leading to liver cancer in chronic hepadnavirus infections.

ACKNOWLEDGMENTS

We are grateful to P. Tiollais for his constant interest in this work. We thank B. Barnes for his generous gift of squirrel tissue and plasma and B. Shamoon for the gift of anti-WHs and anti-WHV pre-S2 antisera. We thank R. Lévy, C. Transy, and S. Prigent for helpful advice, J. Seeler for critical reading of the manuscript, E. Bégaud and M. Maisonnas for excellent technical assistance, and L. M. Da for secretarial assistance.

P. Testut was supported by the Centre Européen de Bioprospective. This work was supported in part by a grant from the Association pour la Recherche contre le Cancer (contract 6550).

REFERENCES

- Beasley, R. P. 1988. Hepatitis B virus—the major etiology of hepatocellular carcinoma. *Cancer* **61**:1942–1956.
- Buendia, M. A. 1992. Hepatitis B viruses and hepatocellular carcinoma. *Adv. Cancer Res.* **59**:167–226.
- Cattaneo, R., H. Will, and H. Schaller. 1984. Hepatitis B virus transcription in the infected liver. *EMBO J.* **3**:2191–2196.
- Chen, H. S., S. Kanako, R. Girones, R. W. Anderson, W. E. Hornbuckle, B. C. Tennant, P. J. Cote, J. L. Gerin, R. H. Purcell, and R. H. Miller. 1993. The woodchuck hepatitis virus X gene is important for establishment of virus infection in woodchucks. *J. Virol.* **67**:1218–1226.
- Church, G., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81**:1991–1995.
- Cohen, J. I., R. H. Miller, B. Rosenblum, K. Denniston, J. L. Gerin, and R. H. Purcell. 1988. Sequence comparison of woodchuck hepatitis virus replicative forms. *Virology* **162**:12–20.
- Colgrove, R., G. Simon, and D. Ganem. 1989. Transcriptional activation of homologous and heterologous genes by the hepatitis B virus X gene product in cells permissive for viral replication. *J. Virol.* **63**:4019–4026.
- Cote, P. J., and J. L. Gerin. 1983. Nonoverlapping antigenic sites of woodchuck hepatitis virus surface antigen and their cross-reactivity with ground squirrel hepatitis virus and hepatitis B virus surface antigens. *J. Virol.* **47**:15–23.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis program for the VAX. *Nucleic Acids Res.* **12**:387–394.
- Enders, G. H., D. Ganem, and H. Varmus. 1985. Mapping the major transcripts of ground squirrel hepatitis virus: the presumptive template for reverse transcriptase is terminally redundant. *Cell* **42**:297–308.
- Faktor, O., S. Budlovsky, R. Ben-Levy, and Y. Shaul. 1990. A single element within the hepatitis B virus enhancer binds multiple proteins and responds to multiple stimuli. *J. Virol.* **64**:1861–1863.
- Feitelson, M. A., P. L. Marion, and W. S. Robinson. 1981. Antigenic and structural relationships of the surface antigens of hepatitis B virus, ground squirrel hepatitis virus, and woodchuck hepatitis virus. *J. Virol.* **39**:447–454.
- Feitelson, M. A., P. L. Marion, and W. S. Robinson. 1983. The nature of polypeptides larger in size than the major surface antigen components of hepatitis B like viruses in ground squirrels, woodchucks and ducks. *Virology* **130**:76–90.
- Feitelson, M. A., I. Millman, T. Halbherr, H. Simmons, and B. S. Blumberg. 1986. A newly identified hepatitis B virus in tree squirrels. *Proc. Natl. Acad. Sci. USA* **83**:2233–2237.
- Fourrel, G., J. Couturier, Y. Wei, F. Apiou, P. Tiollais, and M. A. Buendia. 1994. Evidence for long-range oncogene activation by hepadnavirus insertion. *EMBO J.* **13**:2526–2534.
- Fourrel, G., C. Trépo, L. Bougueleret, B. Henglein, A. Ponzetto, P. Tiollais, and M. A. Buendia. 1990. Frequent activation of N-myc genes by hepadnavirus insertion in woodchuck liver tumours. *Nature (London)* **347**:294–298.
- Fromenty, B., and D. Pessayre. 1995. Inhibition of mitochondrial beta-oxidation as a mechanism of hepatotoxicity. *Pharmacol. Ther.* **67**:101–154.
- Frommel, D., D. Crevat, L. Vitvitsky, C. Pichoud, O. Hantz, M. Chevalier, J. A. Grimaud, J. Lindberg, and C. G. Trépo. 1984. Immunological aspects of woodchuck hepatitis. *Am. J. Pathol.* **115**:125–134.
- Galibert, F., T. N. Chen, and E. Mandart. 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence. *J. Virol.* **41**:51–65.
- Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay. 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli*. *Nature (London)* **281**:646–650.
- Ganem, D., and H. E. Varmus. 1987. The molecular biology of the hepatitis B viruses. *Annu. Rev. Biochem.* **56**:651–693.
- Ganem, D., B. Weiser, A. Barchuk, R. J. Brown, and H. E. Varmus. 1982. Biological characterization of acute infection with ground squirrel hepatitis virus. *J. Virol.* **44**:366–373.
- Girones, R., P. J. Cote, W. E. Hornbuckle, B. C. Tennant, J. L. Gerin, R. H. Purcell, and R. H. Miller. 1989. Complete nucleotide sequence of a molecular clone of woodchuck hepatitis virus that is infectious in the natural host. *Proc. Natl. Acad. Sci. USA* **86**:1846–1849.
- Girones, R., and R. H. Miller. 1989. Mutation rate of the hepadnavirus genome. *Virology* **170**:595–597.
- Gust, I. D., C. J. Burrel, A. G. Coulepis, W. S. Robinson, and A. J. Zuckerman. 1986. Taxonomic classification of human hepatitis B virus. *Intervirology* **25**:14–19.
- Hansen, L. J., B. C. Tennant, C. Seeger, and D. Ganem. 1993. Differential activation of *myc* gene family members in hepatic carcinogenesis by closely related hepatitis B viruses. *Mol. Cell. Biol.* **13**:659–667.
- Hsu, T. Y., T. Möröy, J. Etienne, A. Louise, C. Trépo, P. Tiollais, and M. A. Buendia. 1988. Activation of c-myc by woodchuck hepatitis virus insertion in hepatocellular carcinoma. *Cell* **55**:627–635.
- Kew, M. C., B. C. Tennant, R. H. Purcell, and R. H. Miller. 1993. Heterogeneity of the woodchuck hepatitis virus genome in a chronically infected woodchuck. *Virus Res.* **27**:229–237.
- Kobayashi, M., and K. Koike. 1984. Complete nucleotide sequence of hepatitis B virus DNA of subtype adr and its conserved gene organization. *Gene* **30**:227–232.
- Kodama, K., N. Ogasawara, H. Yoshikawa, and S. Murakami. 1985. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: evolutionary relationship between hepadnaviruses. *J. Virol.* **56**:978–986.
- Koechel, H. 1990. Direct submission to GenBank.
- Korba, B. E., F. V. Wells, B. Baldwin, P. J. Cote, B. C. Tennant, H. Popper, and J. L. Gerin. 1989. Hepatocellular carcinoma in woodchuck hepatitis virus-infected woodchucks: presence of viral DNA in tumor tissue from chronic carriers and animals serologically recovered from acute infections. *Hepatology* **9**:461–470.
- Mandart, E., A. Kay, and F. Galibert. 1984. Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with woodchuck and human hepatitis B virus sequences. *J. Virol.* **49**:782–792.
- Marion, P. L. 1991. Ground squirrel hepatitis virus, p. 39–51. *In* A. McLachlan (ed.), *Molecular biology of the hepatitis B virus*. CRC Press, Inc., Boca Raton, Fla.
- Marion, P. L., S. S. Knight, F. H. Salazar, H. Popper, and W. S. Robinson. 1983. Ground squirrel hepatitis virus infection. *Hepatology* **3**:519–527.
- Marion, P. L., L. S. Oshiro, D. C. Regnery, G. H. Scullard, and W. S. Robinson. 1980. A virus in beechey ground squirrels that is related to hepatitis B virus of humans. *Proc. Natl. Acad. Sci. USA* **77**:2941–2945.
- Marion, P. L., M. J. Van Davelaar, S. S. Knight, F. H. Salazar, G. Garcia, H. Popper, and W. S. Robinson. 1986. Hepatocellular carcinoma in ground squirrels persistently infected with ground squirrel hepatitis virus. *Proc. Natl. Acad. Sci. USA* **83**:4543–4546.
- Mason, W. S., G. Seal, and J. Summers. 1980. Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. *J. Virol.* **36**:829–836.
- Miller, R. H., and W. S. Robinson. 1986. Common evolutionary origin of hepatitis B virus and retroviruses. *Proc. Natl. Acad. Sci. USA* **83**:2531–2535.
- Minuk, G. Y., E. A. Shaffer, D. I. Hoar, and J. Kelly. 1986. Ground squirrel hepatitis virus (GSHV) and hepatocellular carcinoma in the Canadian Richardson ground squirrels. *Liver* **6**:350–356.
- Möröy, T., J. Etienne, C. Trépo, P. Tiollais, and M. A. Buendia. 1985. Transcription of woodchuck hepatitis virus in the chronically infected liver. *EMBO J.* **4**:1507–1514.
- Mukaide, M., T. Kumazawa, A. Hoshi, R. Kawaguchi, and K. Hikiji. 1992. The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and phylogenetic analysis. *Nucleic Acids Res.* **20**:6105.
- Neurath, A. R., S. B. H. Kent, N. Strick, and K. Parker. 1986. Identification and chemical synthesis of a host cell receptor binding site on hepatitis B virus. *Cell* **46**:429–436.
- Newbold, J. E. Personal communication.
- Norder, H., A. M. Courouce, and L. O. Magnius. 1994. Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* **198**:489–503.
- Okamoto, H., M. Imai, M. Kametani, T. Nakamura, and M. Mayumi. 1987. Genomic heterogeneity of hepatitis B virus in a 54-year-old woman who contracted the infection through materno-fetal transmission. *Jpn. J. Exp. Med.* **57**:231–236.
- Ono, Y., H. Onda, R. Sasada, K. Igarashi, Y. Sugino, and K. Nishioka. 1983. The complete nucleotide sequences of the cloned hepatitis B virus DNA: subtype adr and adw. *Nucleic Acids Res.* **11**:1747–1757.
- Orito, E., M. Mizokami, Y. Ina, E. N. Moriyama, N. Kameshima, M. Yamamoto, and T. Gojobori. 1989. Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. *Proc. Natl. Acad. Sci. USA* **86**:7059–7062.
- Parana, R., F. Gérard, J. L. Lesbordes, C. Pichoud, L. G. Costa Lyra, and C. Trépo. 1995. Serial transmission of spongicytic hepatitis to woodchucks (possible association with a specific delta strain). *J. Hepatol.* **22**:468–473.
- Popper, H., L. Roth, R. H. Purcell, B. C. Tennant, and J. L. Gerin. 1987. Hepatocarcinogenicity of the woodchuck hepatitis virus. *Proc. Natl. Acad. Sci. USA* **84**:866–870.
- Preisler-Adams, S., H. J. Schlayer, T. Peters, R. Korp, and J. Rasenack. 1993. Complete nucleotide sequence of hepatitis B virus, subtype adw2, and

- identification of three types of C open reading frame. *Nucleic Acids Res.* **21**:2258.
52. **Preisler-Adams, S., M. J. Schlayer, T. Peters, F. Hettler, W. Gerok, and J. Rasenack.** 1993. Sequence analysis of hepatitis B virus DNA in immunologically negative infection. *Arch. Virol.* **133**:385-396.
 53. **Saiki, R., D. Gelfaud, S. Stoffel, S. Scharf, R. Higuchi, G. Horn, K. Mullis, and H. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. *Science* **239**:487-491.
 54. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406-425.
 55. **Satrosowignjo, R. I., S. Omi, H. Okamoto, M. Mayumi, M. Rustam, and X. Sujudi.** 1987. The complete nucleotide sequence of HBV DNA clone of subtype adw (pMND122) from Menado in Sulawesi Island, Indonesia. *ICMR Ann.* **7**:51-60.
 56. **Schaeffer, E., R. L. Snyder, and J. J. Sninsky.** 1986. Identification and localization of pre-S-encoded polypeptides from woodchuck and ground squirrel hepatitis viruses. *J. Virol.* **57**:173-182.
 57. **Scott, J., M. Hadchouel, C. Hery, J. Yvart, P. Tiollais, and C. Bréchet.** 1983. Detection of hepatitis B virus DNA in serum by a simple spot hybridization technique: comparison with results for other viral markers. *Hepatology* **3**:279-284.
 58. **Seeger, C., B. Baldwin, W. E. Hornbuckle, A. E. Yeager, B. C. Tennant, P. Cote, L. Ferrell, D. Ganem, and H. E. Varmus.** 1991. Woodchuck hepatitis virus is a more efficient oncogenic agent than ground squirrel hepatitis virus in a common host. *J. Virol.* **65**:1673-1679.
 59. **Seeger, C., D. Ganem, and H. E. Varmus.** 1984. Nucleotide sequence of an infectious molecularly cloned genome of ground squirrel hepatitis virus. *J. Virol.* **51**:367-375.
 60. **Shamoon, B. M., A. Kay, S. Dupont de Dinechin, F. Galibert, and E. Mandart.** 1991. Production of polyclonal antibodies against the S and pre-S2 regions of woodchuck hepatitis virus: lack of detectable low glycosylated pre-S2 protein (GP33) in sera from infected animals. *J. Gen. Virol.* **72**:421-425.
 61. **Shaul, Y., W. J. Rutter, and O. Laub.** 1985. A human hepatitis B viral enhancer element. *EMBO J.* **4**:427-430.
 62. **Sprengel, R., E. F. Kaleta, and H. Will.** 1988. Isolation and characterization of a hepatitis B virus endemic in herons. *J. Virol.* **62**:3832-3839.
 63. **Summers, J., J. M. Smolec, and R. Snyder.** 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. *Proc. Natl. Acad. Sci. USA* **75**:4533-4537.
 64. **Takada, S., and K. Koike.** 1990. X protein of hepatitis B virus resembles a serine protease inhibitor. *Jpn. J. Cancer Res.* **81**:1191-1194.
 65. **Tennant, B. C., N. Mrosovsky, K. McLean, P. J. Cote, B. E. Korba, R. E. Engle, J. L. Gerin, J. Wright, G. R. Michener, E. Uhl, and J. M. King.** 1991. Hepatocellular carcinoma in Richardson's ground squirrels (*Spermophilus richardsonii*): evidence for association with hepatitis B-like virus infection. *Hepatology* **13**:1215-1222.
 66. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 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**:4673-4680.
 67. **Tiollais, P., C. Pourcel, and A. Dejean.** 1985. The hepatitis B virus. *Nature (London)* **317**:489-495.
 68. **Toh, H., H. Hyashida, and T. Miyata.** 1983. Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. *Nature (London)* **305**:827-829.
 69. **Toshkov, I., H. J. Hacker, M. Roggendorf, and P. Bannasch.** 1990. Phenotypic patterns of preneoplastic and neoplastic hepatic lesions in woodchucks infected with woodchuck hepatitis virus. *J. Cancer Res. Clin. Oncol.* **116**:581-590.
 70. **Transy, C., G. Fourel, W. S. Robinson, P. Tiollais, P. L. Marion, and M. A. Buendia.** 1992. Frequent amplification of c-myc in ground squirrel liver tumors associated with past or ongoing infection with a hepadnavirus. *Proc. Natl. Acad. Sci. USA* **89**:3874-3878.
 71. **Twu, J. S., and R. H. Schloemer.** 1987. Transcriptional *trans*-activating function of hepatitis B virus. *J. Virol.* **61**:3448-3453.
 72. **Vaudin, M., A. J. Wolstenholme, K. N. Tsiquaye, A. J. Zuckerman, and T. J. Harrison.** 1988. The complete nucleotide sequence of the genome of a hepatitis B virus isolated from a naturally infected chimpanzee. *J. Gen. Virol.* **69**:1383-1389.
 73. **Vitvitski-Trépo, L.** Unpublished results.
 74. **Vitvitski, L., M. L. Meyers, J. J. Sninsky, P. Berthillon, P. Chevalier, M. A. Sells, G. Acs, and C. Trépo.** 1988. Expression of the X gene product of hepatitis B virus and WHV in infected livers and transfected 3T3 cells: evidence for cross-reactivity and correlation with core/e gene expression, p. 341-344. *In* A. J. Zuckerman (ed.), *Viral hepatitis and liver disease*. Alan R. Liss, Inc., New York.
 75. **Wei, Y., G. Fourel, A. Ponzetto, M. Silvestro, P. Tiollais, and M. A. Buendia.** 1992. Hepadnavirus integration: mechanisms of activation of the *N-myc2* retrotransposon in woodchuck liver tumors. *J. Virol.* **66**:5265-5276.
 76. **Xiong, Y., and T. H. Eickbush.** 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**:3353-3362.