Journal of Animal and Veterinary Advances 2 (3): 162-169, 2003

Molecular and Phylogenetic Analyses of Duck Hepatitis B Viruses

[1,2]Chi-Young J. Wang and [1]Joseph J. Giambrone
[1]Department of Poultry Science, Auburn University, Auburn, AL 36849-5416, USA
[2]Scott-Ritchey Research Center, College of Veterinary Medicine, Auburn, AL 36849, USA

Abstract: *Hepadnaviridae* is a DNA virus family including human hepatitis B virus (HBV), woodchuck hepatitis B virus (WHV), and duck hepatitis B virus (DHBV). Despite differences in genomic structures, DHBV is similar to HBV in replication strategy, infection outcome, and sensitivity to antiviral treatment. It is a safe animal model for HBV and has facilitated understanding of HBV. In this study, newly isolated DHBVs from California and Indiana were cloned by PCR and their nucleotide sequences (3021 bp) determined. Both isolates contained three overlapping open reading frames encoding surface (Pre-S/S), core structure (Pre-C/C), and polymerase (P) proteins. Phylogenetic analysis based on nucleotide sequences showed they were closely related to other USA isolates (≥99% identity) and were placed in "Western country" DHBV group. When the amino acid sequences of surface protein (S) were compared, a similar genetic tree with closer distances was produced. More diversity occurred between mammalian and avian hepadnaviruses. Pre-S and P proteins varied most among viral proteins. The Pre-C/C proteins were highly conserved with divergence less than 3.8%. Variations in P protein were located in its non-functional overlapping part, called spacer domain, within pre-S proteins. Data provided knowledge of the molecular and phenotypic characteristics of *avihepadnaviridae*.

Key word: Hepadnaviridae; Duck hepatitis B virus (DHBV); Polymerase chain reaction (PCR); Phylogenetic analysis

Introduction

Hepadnaviridae is a DNA virus family including two genera, *Orthohepadnavirus* and *Avihepadnavirus*. Each has restricted host specificity. *Orthohepadnaviruses* include human hepatitis B virus (HBV), woodchuck hepatitis virus (WHV), and ground squirrel hepatitis virus (GSHV) (Mandart *et al.*, 1984). *Avihepadnaviruses* include duck (duck hepatitis B virus, DHBV), grey heron (heron hepatitis B virus, HHBV), ross goose (ross goose hepatitis B virus, RGHBV), and snow goose (snow goose hepatitis B virus, SGHBV) (Chang *et al.*, 1999; Mason *et al.*, 1980; Sprengel *et al.*, 1988). Mammalian and avian hepadnaviruses show similarity in genetic organization, virus replication, and outcome in their respective hosts (Cova *et al.*, 1993; Duflot *et al.*, 1995; Schodel *et al.*, 1989). Although sequence divergence exists between those genera, DHBV has provided a safe animal model for HBV study (Orito *et al.*, 1989; Sprengel *et al.*, 1985). Studies of DHBV infection *in vitro* and *in vivo* have made significant contributions to HBV knowledge.

DHBV has a small (3 kb) partially double-stranded DNA genome with three overlapping open reading frames (ORF), Pre-C/C gene, Pre-S/S gene, and P gene. They encoded nucleocapsid/precursor proteins, presurface/surface proteins, and polymerase protein, respectively (Ishikawa *et al.*, 1995). This virus was reported from all regions and more than 13 nucleotide sequences of DHBV isolates are available in the GenBank. They were divided into two groups (Chinese and Western DHBVs) based on MEGA and splitstree analysis. They had distinct divergence with other avian viruses such as RGHBV and SGHBV (Chang et al, Triyantni *et al.*, 2001). Phylogenetic relationship between newly California and Indiana DHBV isolates and other hepadnaviruses were determined herein by primer walking and sequencing. Distinct evolutionary distance was reflected among different species lineage including *Orthohepadnavirus* and *Avihepadnavirus*. As expected, these new isolates were closest to other USA isolates, thus offering additional support to importance of geographic confinement in worldwide DHBV distribution.

Materials and Methods

Isolation and propagation of DHBV : Fertilized Pekin duck eggs were from DHBV negative flocks (Metzer Farms, CA). Eggs were incubated for 24 days at 37 °C. Embryos were killed and livers minced and digested for 30 minutes in DMEM (Gibco) containing 0.5% collagenase A (Boehringer), 0.24% $NaHCO_3$, 10 mM HEPES (pH 7.4), and 100 units per ml of penicillin. Cells were strained through gauze, pelleted at 5000 × g for

2 minutes, and washed twice with Medium 199 containing 4 × 10^3 units per ml of insulin, 1μM corticosterone, 0.034% $NaHCO_3$, 10 mM HEPES (pH 7.4), 100 units per ml of penicillin, and 2.5% FBS. Cells were seeded at 10^7 cells per well in culture plates with Medium 199. Confluent cells were infected with 20 μl of virus-containing serum, obtained from duck flocks in Indiana (Culver farm, IN) or California (Metzer farm, CA), per well adjusted to a final volume of 0.5 ml with MEM medium without FBS for 1.5 hours at 37 °C. Inoculum was removed, and L-15 medium with 2.5% FBS was added.

Viral DNA extraction, primer design, and PCR: Infected cells were lysed in 1 ml of DNAzol (MRC, Cincinnati, Ohio) 10 days post-infection. DNA was precipitated with 0.25 ml of 100% ethanol, mixed by inverting tube, and stored at room temperature for 3 minutes. DNA was pelleted at 5000 × g for 5 minutes and washed twice with 0.25 ml of 75% ethanol. DNA was dissolved in 50 μl of TE buffer and quantified by spectrophotometer. Following combinations of oligonucleotide primers designed based on DHBV USA p2-3 strain (GenBank accession number M60677) were used to amplify the DHBV genome in overlapping fragments: P1 (5'-TGACTCCAGCCAAATTCTGG-3'; positions 567-586) and P2 (5'-TTCGCATGCTGCTTTTTAATCTT-3', positions 1074-1094); P3 (5'-AGTTGTCTATGGGAGAAGAC-3'; positions 291-310) and P4 (5'-GAAAGTAGGAAATGCTCT-3'; positions 589-608); P5 (5'-AAATCCGACTCCTCAAGAGAT-3'; positions 1028-1048) and P6 (5'-TGCTGCTACTAGCAGGATTAAGAGG-3'; positions 1532-1556); P7 (5'-TCATCTTCCTCTTAATCCTGCT-3'; positions 1525-1544) and P8 (5'-GCAGCATACATTGGTTTTAACATT-3'; positions 2014-2037); P9 (5'-CAATGTATGCTGCTATTACTAACCAAG-3'; positions 2031-2058) and P10 (5'-CCGAAGGAGAGGGGTGTAAT-3'; positions 2533-2552); P11 (5'-CGACTATCCAATTACGGCTAG-3'; positions 2997-3017) and P12 (5'-GGTGTAGGGGCTCTCCTTTC-3'; positions 325-344); P13 (5'-GAATTACACCCCTCTCCTTCG-3'; positions 2531-2551) and P14 (5'-TTGCATAAGCTTTCAACTGAGC-3'; positions 4-25). PCR reactions contained 5 μl of extracted DNA, 10 mM Tris-Hcl (pH=8.3), 2.0 mM $MgCl_2$, 0.2 mM of each dNTP, 0.1 μM of each primer, and 0.5 μl of Pfu polymerase (5 units/μl) (Perkin-Elmer, NJ) and underwent 30 cycles containing denaturation 30 sec at 95 °C, annealing 30 sec at 55 °C, extension 1 min at 72 °C), and one final extension step at 72 °C for 10 min. PCR products were analyzed in agarose and bands of correct sizes were eluted.

Table 1: Characterized hepadnavirus genomes for phylogenetic analysis

Accession	Length (nt)	Species	Origin	Clone name	Reference
		Species	Location		
M32990	3027	Duck, brown	Shanghai	DHBVS5cg	Uchida *et al.*, 1989
M32991	3027	Duck, white	Shanghai	DHBVS31cg	Uchida *et al.*, 1989
AJ006350	3027	Duck, Pekin	Australia	AusDHBV	Triyatni *et al.*, 2001
X58568	3024	Duck, domestic	China	DHBV22	Sprengel *et al.*, 1991
X58569	3024	Duck, domestic	Shanghai	DHBV26	Sprengel *et al.*, 1991
K01834	3021	Duck, Pekin	USA	DHBV16	Mandart *et al.*, 1984
M60677	3021	Duck, Pekin	USA (Maple farm)	DHBVp2-3	GenBank
X12798	3021	Duck	Germany	DHBVf1-6	Mattes *et al.*, 1990
X74623	3021	Duck	India	DHBVCG	GenBank
AF047045	3021	Duck	Canada	DHBV47045	GenBank
AF110996	3024	Snow goose	Germany	SGHBV1-13	Chang *et al.*, 1999
M95589	3018	Ross goose	USA	RGHBV	GenBank
M22056	3027	Grey heron	Germany	HHBV4	Sprengel *et al.*, 1988
J04514	3323	Woodchuck	USA	WHV8	Girones *et al.*, 1989
NC_003977	3215	Human	Japan	HBV	Okamoto *et al.*, 1986
AF505512	3021	Duck, Pekin	USA, CA	DHBV-C	GenBank
AF493986	3021	Duck, Pekin	USA, IN	DHBV-I	GenBank

Subcloning, DNA sequencing, and phylogenetic analysis: Eluted fragments were subcloned into pTrueBlue-rop according to the manufacturer (Genomics One, Quebec, Canada). Plasmids were examined by restriction enzyme digestion and electrophoresis. Sequences were determined in both strands by automated sequencing on the Applied Biosystems UA373 DNA Sequencing System using the PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit as recommended by the manufacturer. California and Indiana DHBV isolates were used to determine paired identity and divergence of nucleotide and deduced amino acid sequences of surface gene (S gene) with other hepadnaviruses (Table 1) using DNASTAR software packages (DNASTAR, Madison, WI, USA). Phylogenetic trees were constructed with mammalian hepadnaviruses as the outgroup using a neighbour-joining method of the Clustal X software (Thompson *et al.*, 1999). Bootstrap analyses were performed by 1,000 resamplings of data sets. A second phylogenetic analysis was produced based on deduced amino acid sequences of surface gene (S). The hydrophobicity score of S protein was determined by Kyte-Doolittle method (Kyte *et al.*, 1982). Percentages of amino acid sequence divergence of DHBV viral proteins including pre-surface protein (Pre-S), surface protein (S), pre-core protein (Pre-C), core protein (C), polymerase protein without spacer (P*), and polymerase protein (P) between California isolate and other isolates were calculated using MegAlign of DNASTAR software.

Results and Discussion

Cloning and sequencing of DHBV isolates: DHBV particles, obtained from serum, have a unique partially double-stranded DNA. This can cause interruption of PCR amplification on single strand region (Schlicht *et al.*, 1987). In this study, DHBV DNA was isolated after infection of duck embryonic hepatocytes, because the viral genome is converted into double stranded circular DNA upon entry into cells. DNA was extracted ten days after infection, which is the high yield period for DHBV replication (Tuttleman *et al.*, 1986). Complete DHBV sequences, derived from plasmids with overlapping inserts covering the entire genome, were deposited in the GenBank with the accession number AF493986, and AF505512 for DHBV Indiana, and California isolates, respectively. Both isolates contained three overlapping, open reading frames encoding polypeptides of 328/167 (Pre-S/S), 305/262 (Pre-C/C), and 786 (P) amino acids with start codons at nt 801/1284, 2518/2647, and 170 and stop codons at nt 1787, 414, and 2530, respectively. Their genome size (3021bp) was the same as other "Western country" isolates, which are shorter than Chinese isolates (3024 or 3027 bp). However, features associated with replication, including a pair of 12 nt direct repeat (DR) sequences, DR1 (nt 2535-2546) and DR2 (nt 2477-2488) for maintaining the circular genome, the polyadenylation signal (AATAAA) (nt 2772-2777) for viral transcription termination, and the tyrosine residue at position 96 within the N-terminal domain of the polymerase protein for the binding of RNA encapsidation signal, were conserved.

Nucleotide sequence and phylogenetic analysis: All DHBV isolates exhibited higher levels of sequence identity (≥88.4%) than other hepadnaviruses (Table 2). Geographic diversity was reflected in nucleotide sequences as Chinese and Western DHBV subsets with identities of 91.5%-94.9% and 88.4%-88.9%, respectively. "Western DHBV" subsets included DHBV16, DHBVp2-3, DHBVf1-6, DHBVCG and DHBV47045. California and Indiana DHBVs were closely related to other USA isolates (≥99% sequence identity) and were placed in "Western DHBV" subset. Within avian hepadnaviruses, isolates from snow geese (SGHBV) were closer to DHBV, than ross goose (RGHBV) by 12.62% and 21.56 % sequence divergence, respectively. This placed SGHBV in evolution between RGHBV and DHBV. Maximum evolutionary distance was observed on grey heron (HHBV) with 28.01% sequence divergence, which was also suggested by their distinct host range (Triyatni *et al.*, 2001). A small degree of identity in nucleotide sequence (19.8-21.0%) was uncovered between mammalian and avian hepadnavirus, because the unique genomic structure such as X gene was lacking in DHBV.

It was accepted that absence of X gene caused lack of directly oncogenic properties of DHBV in ducks (Jilbert *et al.*, 2000). Research indicated possibility of existing of X-like gene, which begins with an alternative initiation codon (TTA) in DHBV. It activated cellular and viral promoters via the Raf-mitogen-activated protein kinase signaling pathway and suggested functionally a common ancestry of *orthohepadnavirus* and *avihepadnavirus* X gene (Chang *et al.*, 2001; Triyatni *et al.*, 2001). Four major clusters were obtained from phylogenetic trees (Fig. 1). DHBVs were divided into three groups. Group I and II/III represented "Western and Chinese DHBVs" with considerable difference between them

(100% bootstrap support). Although DHBV 22 and 26 belonged to Chinese DHBVs, they were separated from other Chinese isolates as Group III with 100% bootstrap support. Their genomic length, similar to SGHBV (3024 bp), was shorter than other Chinese isolates (3027 bp) with less sequence divergence (3.5%) to "Western DHBVs". This variation in a single virus population, isolated from the same geographic location, was likely due to host factors such as different duck strains and immunological selection, or genetic mutations by the virus (Orito *et al.*, 1989). Interestingly, the Australia strain was clustered with Chinese DHBVS31cg and DHBVS5cg, which were both isolated from Shanghai (Group II). This was attributed to the importation of live ducks from China before 1949 (Triyatni *et al.*, 2001). It was an exception to the rule that DHBV phylogenetic characteristics were correlated with geographic locations.

Deduced amino acid sequence of surface protein (S) and phylogenetic analysis: Although little nucleotide identity existed between DHBV and HBV, high homology (51-70%) of viral proteins (Pre-S/S, Pre-C/C, and P proteins) suggested existence of a close genetic relationship. This provides evidence for common ancestry (Table 3). Additionally, similar replication strategy, virion morphology, and outcomes of infection made DHBV a useful animal model for HBV. Amino acid sequence of surface protein (S), major antigens for triggering neutralizing antibody, was compared. DHBV isolates showed higher levels of identity than nucleotide sequences (≥94%). Closer relationship between DHBV and other avian hepadnaviruses was reflected in amino acid level. The most divergent isolate was HHBV, differing by 15.2%. Percent identities between hepadnaviruses were higher for amino acid sequences than for nucleotide sequences, indicating that many nucleotide changes were silent. A similar phylogenetic tree containing four separated clusters with closer distances was attained (Fig. 2). Mammalian hepadnaviruses were placed as an independent group by

Table 2: Paired identity (the upper-right triangle) and divergence (the lower-left triangle) of the complete nucleotide sequences of hepadnaviruses

	Percent identity																
Virus isolates	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
DHBVS5cg 94.9	94.5	92.0	91.5	88.9	88.8	88.8	88.4	89.0	88.7	78.5	71.6	21.0	19.8	88.8	88.7		
DHBVS31cg	5.1		95.8	91.2	90.6	88.1	88.3	88.1	87.6	88.2	88.0	78.7	72.2	19.4	19.7	88.1	88.1
AusDHBV 5.5	4.2		91.2	90.7	88.1	88.2	88.1	87.6	88.0	87.5	78.8	72.4	20.0	18.5	87.9	88.0	
DHBV22 8.0	8.8	8.8		96.9	89.3	89.5	89.4	88.8	89.3	87.7	78.9	71.3	19.0	18.6	89.2	89.4	
DHBV26 8.5	9.4	9.3	3.1		89.3	89.5	89.6	89.0	89.3	87.4	78.6	71.4	19.5	18.9	89.1	89.1	
DHBV16 11.1	11.9	11.9	10.7	10.7		99.1	98.8	98.7	99.3	87.0	78.9	72.1	19.5	18.5	99.2	99.4	
DHBVp2-3 11.2	11.7	11.8	10.5	10.5	0.9		99.0	98.2	99.0	87.1	79.0	72.2	19.2	19.8	99.0	99.3	
DHBVf1-6 11.2	11.9	11.9	10.6	10.4	1.2	1.0		99.0	98.8	86.8	78.5	72.0	19.5	19.6	98.5	98.7	
DHBVCG 11.6	12.4	12.4	11.2	11.0	1.3	1.8	1.0		98.3	86.4	78.0	72.0	19.3	19.3	98.1	98.3	
DHBV47045	11.0	11.8	12.0	10.7	10.7	0.7	1.0	1.2	1.7		87.2	78.8	72.0	19.1	18.6	99.3	99.4
SGHBV1-13	11.3	12.0	12.5	12.3	12.6	13.0	12.9	13.2	13.6	12.8		77.0	72.3	19.7	18.7	87.0	87.0
RGHBV 21.5	21.3	21.2	22.1	21.4	21.1	21.0	21.5	22.0	21.2	23.0		73.5	18.8	19.1	78.7	78.9	
HHBV4 28.4	27.8	27.6	28.7	28.6	27.9	27.8	28.0	28.0	28.0	27.7	26.5		18.4	19.8	70.8	72.0	
WHBV4 79.0	80.6	80.0	81.0	80.5	80.5	80.8	80.5	80.7	80.9	80.3	81.2	81.6		54.2	19.7	19.4	
HBV 80.2	80.3	81.5	81.4	81.1	81.5	80.2	80.4	80.7	81.4	81.3	80.9	80.2	55.8		18.9	19.7	
DHBV-C 11.2	11.9	12.1	10.8	10.9	0.8	1.0	1.5	1.9	0.7	13.0	21.3	29.2	80.3	81.1		99.3	
DHBV-I 11.3	11.9	12.0	10.6	10.9	0.6	1.7	1.3	1.7	0.6	13.0	21.1	28.0	80.6	80.3	0.7		

Table 3: Paired identity (the upper-right triangle) and divergence (the lower-left triangle) of the deduced amino acid sequences of surface gene (S) of hepadnaviruses

	Percent identity																
Virus isolates	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
DHBVS5cg	97.0	96.4	94.0	94.0	94.6	94.6	94.6	94.6	94.6	95.2	90.4	84.4	51.0	53.5	94.6	94.6	
DHBVS31cg	3.0		99.4	92.9	92.9	92.8	92.8	92.8	92.8	92.8	94.0	90.4	84.4	51.0	53.5	92.8	92.8
AusDHBV	3.6	0.6		92.3	92.3	92.2	92.2	92.2	92.2	92.2	93.5	89.8	83.8	51.0	53.5	92.2	92.2
DHBV22	6.0	7.1	7.7		98.8	94.6	94.6	94.6	94.6	94.6	95.8	90.4	83.8	51.0	53.5	94.6	94.6
DHBV26	6.0	7.1	7.7	1.2		94.6	94.6	94.6	94.6	94.6	95.2	90.4	85.0	53.0	55.4	94.6	94.6
DHBV16	3.4	7.2	7.8	5.4	5.4		100	100	100	100	94.6	91.1	85.1	53.5	56.9	100	100
DHBVp2-3	3.4	7.2	7.8	5.4	5.4	0		100	100	100	94.6	91.1	85.1	53.5	56.9	100	100
DHBVf1-6	3.4	7.2	7.8	5.4	5.4	0	0		100	100	94.6	91.1	85.1	53.5	56.9	100	100
DHBVCG	3.4	7.2	7.8	5.4	5.4	0	0	0		100	94.6	91.1	85.1	53.5	56.9	100	100
DHBV47045	3.4	7.2	7.8	5.4	5.4	0	0	0	0		94.6	91.1	85.1	53.5	56.9	100	100
SGHBV1-13	4.8	6.0	6.5	4.2	4.8	5.4	5.4	5.4	5.4	5.4		90.4	83.8	52.0	53.5	94.6	94.6
RGHBV	9.6	9.6	10.2	9.6	9.6	8.9	8.9	8.9	8.9	8.9	9.6		86.9	54.5	57.8	91.1	91.1
HHBV4	15.6	15.6	16.2	16.2	15.0	14.9	14.9	14.9	14.9	14.9	16.2	14.1		59.4	62.7	85.1	85.1
WHBV4	49.0	49.0	49.0	49.0	47.0	46.5	46.5	46.5	46.5	46.5	48.0	55.5	40.6		68.4	53.5	53.5
HBV	46.5	46.5	46.5	46.5	45.6	43.1	43.1	43.1	43.1	43.1	46.5	42.2	37.3	32.6		56.9	56.9
DHBV-C	5.4	7.2	7.8	5.4	0.4	0	0	0	0	0	5.4	8.9	14.9	46.5	43.1		100
DHBV-I	5.4	7.2	7.8	5.4	0.4	0	0	0	0	0	5.4	8.9	14.9	46.5	43.1	0	

Table 4: Percent divergence of deduced amino acid sequences of DHBV California isolate viral proteins (Pre-S/S, Pre-C/C P and P*) with other DHBV isolates

Virus	Pre-S	S	Pre-C	C	P	P*
DHBV Indiana (USA)	1.5	0	0	0	0.6	0.6
DHBV16 (USA)	0.9	0	0	0	0.3	0.3
DHBVp2-3 (USA)	0.9	0	0	0	1	0.6
DHBV47045 (Canada)	0.9	0	0.3	0.4	0.4	0.3
DHBVf1-6 (Germany)	0.9	0	0.3	0.4	1.8	0.6
DHBVCG (India)	1.2	0	2	2.3	2.8	2.1
DHBVS5cg (Shanghai, China)	11	4.8	3	2.7	14	7.2
DHBVS31cg (Shanghai, China)	11.3	6.6	3	2.7	16.3	8.3
DHBVAus (Australia)	11.6	7.2	3	2.7	15.9	8
DHBV22 (China)	7.3	4.2	3.6	3.8	12.8	7.2
DHBV26 (Shanghai, China)	8.2	4.2	3.6	3.8	13.4	7.3

P*: P protein without spacer

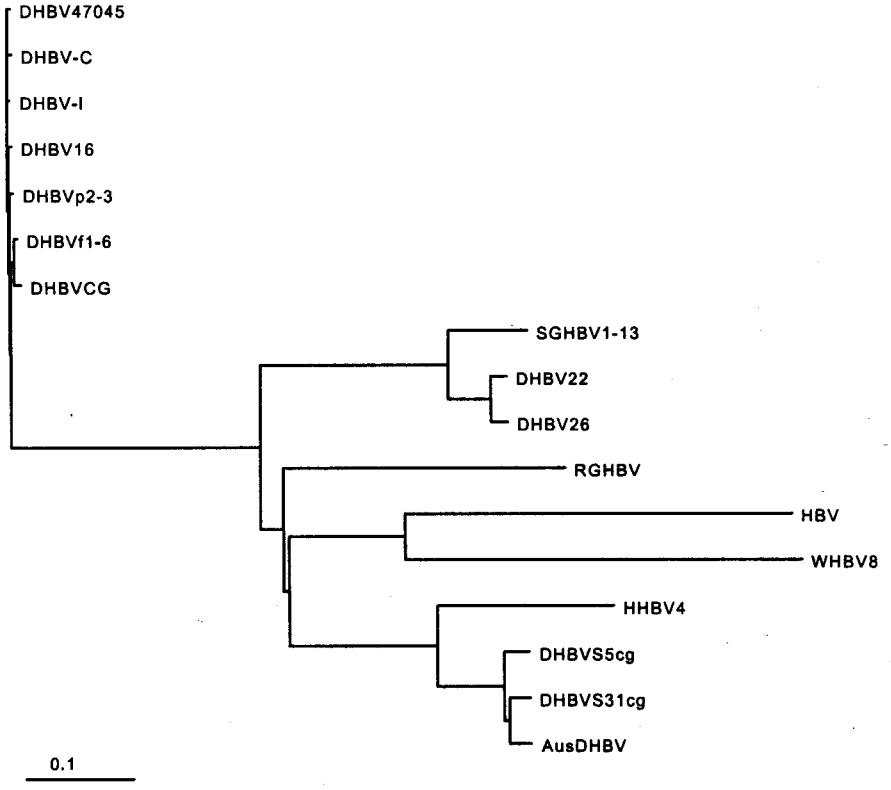

Fig. 1: Phylogenetic relationship of hepadnaviruses, inferred from the full-length nucleotide sequences by neighbour-joining analysis. Bootstrap values for the majo nodes (% support; 1, 000 iterations) are indicated

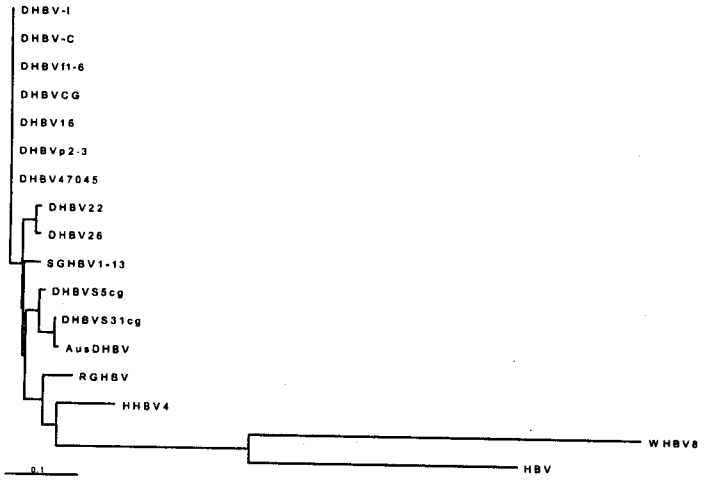

Fig. 2: Phylogenetic relationship of hepadnaviruses, inferred from the deduced amino acid sequences of surface gene (S) by neighbour-joining analysis. Bootstarp values for the major nodes (% support; 1, 000 iterations) are indicated.

```
                10   20   30   40   50   60
                |    |    |    |    |    |
DHBV-I          MSGTFGGILAGLIGLLVSFFLLIKILEILRRLDWWWISLSSPKGKMQCAFQDTGAQISPH
DHBV-C          ..............................................................I...
DHBV47045       ..............................................................I...
DHBVCG          ..............................................................I...
DHBVf1-6        ..............................................................I...
DHBVp2-3        ..............................................................I...
DHBV16          ..............................................................I...
DHBV22          ..............................................................I...
DHBV26          ..............................................................T...
DHBVS31cg       ..............................................................I...
AusDHBV         ..............................................................I...
DHBVS5cg        ..............................................................I...

                        70         80         90         100        110        120
                        |          |          |          |          |          |
DHBV-I          YVGSCPWGCPGFLWTYLRLFIIFLLILLVAAGLLYLTDNGSTILGKLQWASVSALFSSIS
DHBV-C          ...S.....................................G.T.......A..S..F....
DHBV47045       ...S.....................................G.T.......A..S..F....
DHBVCG          ...S.....................................G.T.......A..S..F....
DHBVf1-6        ...S.....................................G.T.......A..S..F....
DHBVp2-3        ...S.....................................G.T.......A..S..F....
DHBV16          ...S.....................................G.T.......A..S..F....
DHBV22          ...S.....................................G.T.......E..L..S....
DHBV26          ...S.....................................G.T.......E..L..S....
DHBVS31cg       ...F.....................................M.I.......E..S..F....
AusDHBV         ...F.....................................M.I.......E..S..F....
DHBVS5cg        ...S.....................................M.I.......E..S..F....

                        130        140        150        160
                        |          |          |          |
DHBV-I          SLLPSDPKSLVALTFGLSLIWMTSSSATQTLVTLTQLATLSALFYKS
DHBV-C          ......P......T...S.....................................
DHBV47045       ......P......T...S.....................................
DHBVCG          ......P......T...S.....................................
DHBVf1-6        ......P......T...S.....................................
DHBVp2-3        ......P......T...S.....................................
DHBV16          ......P......T...S.....................................
DHBV22          ......P......M...L.....................................
DHBV26          ......P......M...L.....................................
DHBVS31cg       ......Q......M...L.....................................
AusDHBV         ......Q......M...L.....................................
DHBVS5cg        ......Q......I...L.....................................
```

Fig. 3: Alignment and comparison of the deduced amino acid sequences of surface gene (S) of DHBV isolates. The single letter amino acid codon was used and only differences were shown. The box indicated the glycosylation signal site.

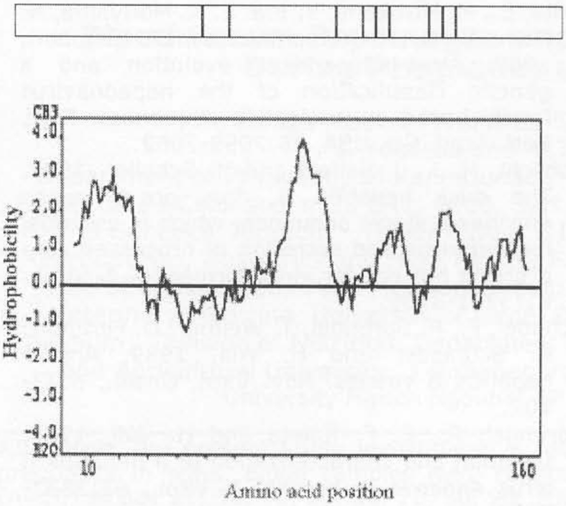

Fig. 4: Hydrophobicity in S protein of DHBV. At the top, positions of amino acid changes are indicated by vertical bars.

96% bootstrap support. This difference was attributed to absence of part of HBV/WHBV surface proteins (about 59 amino acids) in DHBV that codes for the subtype-specific antigenic determinants. Therefore, no serologic cross-reactivity between mammalian and avian hepadnaviruses was detected (Sprengel *et al.*, 1985). However, HBV was less distanced from DHBV than WHBV. Amino acid sequences of surface protein (S) were well conserved in DHBVs with similar glycosylation signals (Asn-Gly/Met-Ser) at amino acid position 99 (Fig. 3). "Chinese DHBVs" had different variations, whereas "Western DHBVs" were identical. Substitutions were mainly from uncharged polar to nonpolar amino acids, with one exception at amino acid position 110 (nonpolar to positive charged). They were located in hydrophobic regions of surface protein (Fig. 4). Possible relationships between these changes and viral immunogenicity need investigation.

Comparison of DHBV viral proteins: Distinct divergence was shown in all viral proteins between "Western and Chinese DHBV's (Table 4). Amino acids of pre-S and P proteins varied most from other proteins. Host specificity was determined by pre-S protein, due to its importance for virus binding to host cells (Uchida *et al.*, 1989). Because all viruses were isolated from ducks, no significant difference appeared (difference ≤11.6%). S protein was more conserved than pre-S protein. This small conserved surface protein facilitated maintenance of vertical transmission, because of less potential to undergo antigenic variations. This conferred an advantage for wide cross protection with a single vaccination. DNA vaccines based on PreS/S genes demonstrated effective protection against DHBV and a sustained reduction of viral loads in persistently infected ducks (Jilbert *et al.*, 2000). Pre-C/C proteins were highly conserved in all DHBVs and amino acid varied less than 3.8%. Within Chinese DHBVs, group II and group III showed differences, but no differences occurred within each group. These proteins were involved in survival such as the viral packages and signals for interaction between regulatory and replication proteins. Therefore, their divergence was prohibited (Molnar-Kimber *et al.*, 1984). P proteins exhibited highest variations (up to 16.3%). This difference was correlated with geographic distribution. Most changes in P protein were located in the overlapping area of pre-S proteins. This region was called the spacer domain (amino acid position 588 to 748) and no enzymatic function was identified by mutational analysis (Bartenschlager *et al.*, 1988; Triyatni *et al.*, 2001). Less divergence (≤ 8.3%) was observed when the rest of P protein (P* protein) was compared. Excessive changes within these enzymatic regions, in charge of reverse transcriptase and RnaseH activities, would be deleterious to the virus. It was an ideal strategy for the virus to provide no function to the spacer domain, to accommodate possible variations in pre-S regions and serve as part of the P protein, simultaneously. This provided the virus capability to tolerate emergence of nucleotide changes in pre-S protein during serial passages in the host without interfering with essential physiological functions. Moreover, it allowed the virus to adapt to the host's immune system. This study provided useful knowledge of the molecular and phylogenetic characteristics of *avihepadnaviridae*.

Acknowledgements

This work was supported partially by the Alabama Agricultural Experiment Station, Auburn University, AL.

References

Bartenschlager, R., H. Schaller, 1988. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. EMBO. J., 7: 4185-4192.

Chang, S. F., H. J. Netter, M. Bruns, R. Schneider, K. Frolich and H. Will, 1999. A new avian hepadnavirus infecting snow goose (*Ansercaerulescens*) produces a significant fraction of virions containing single-stranded DNA. Virology, 262:39-54.

Chang, S. F., H. J. Netter, E. Hildt, R. Schuster, S. Schaefer, Y. C. Hsu, A. Rang and H. Will, 2001. Duck hepatitis B virus expresses a regulatory HBx-like protein from a hidden open reading frame. J. Virol., 75:161-170.

Cova, L., A. Duflot, M. Prave and C. Trepo, 1993. Duck hepatitis B virus infection, alfatoxin B1 and liver cancer in ducks. Arch. Virol., 8:81-87.

Duflot, A., R. Mehrotra, S. Yu, L. Barraud, C. Trepo and L. Cova, 1995. Spectrum of liver disease and duck hepatitis B virus infection in a large series Chinese ducks with hepatocellular carcinoma. Hepatology, 21:1483-1491.

Thompson, J. D., F. Plewniak and O. Poch, 1999. A comprehensive comparison of multiple sequence alignment programs. Nucleic Acids Res., 27:2682-2690.

Ishikawa, T., D. Ganem, 1995. The pre-S domain of the large viral envelope protein determines host range in avian hepatitis B viruses. Proc. Natl. Acad. Sci. USA, 92:6259-6263.

Jilbert, A. R., I. Kotlarski, 2000. Immune responses to duck hepatitis B virus infection. Dev. Com. Immunol., 24:285-302.

Kyte, J., R. F. Doolittle, 1982. A Simple Method for Displaying the Hydropathic Character of a Protein. J. Mol. Biol., 157:105-132.

Mandart, E., A. Kay and F. Galibert, 1984. Nucleotide sequence of a cloned duck hepatitis virus genome: comparison with woodchuck and human hepatitis B virus sequences. J. Virol., 49:782-792.

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.

Mattes, F., S. Tong, K. Teubner and H. E. Blum, 1990. Complete nucleotide sequence of a German duck hepatitis B virus. Nucleic Acids Res., 18:6140.

Molnar-Kimber, K. L., J. W. Summers and W. S. Mason, 1984. Mapping of the cohesive overlap of duck hepatitis B virus DNA and of the site of initiation of reverse transcription. J. Virol., 51:181-191.

Okamoto, H., M. Imai, M. Shimozaki, Y. Hoshi, H. Iizuka, T. Gotanda, F. Tsuda, Y. Miyakawa and M. Mayumi, 1986. Nucleotide sequence of a cloned hepatitis B virus genome, subtype ayr: comparison with genomes of the other three subtypes. J. Gen. Virol., 67:2305-2314.

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.

Schlicht, H. J., J. Salfeld and H. Schaller, 1987. The duck hepatitis B virus pre-C region encodes a single sequence, which is essential for synthesis and secretion of processed core proteins but not for virus formation. J. Virol., 61:3701-3709.

Schodel, F., R. Sprengel, T. Weimer, D. Fernholz, R. Schneider and H. Will, 1989. Animal hepatitis B viruses. Adv. Viral. Oncol., 8:73-102.

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.

Sprengel, R., C. Kuhn, H. Will and H. Schaller, 1985. Comparative sequence analysis of duck and human hepatitis B virus genome. J. Med. Virol., 15:323-333.

Sprengel, R., R. Schneider, P. L. Marion, D. Fernholz, G. Wildner and H. Will, 1991. Comparative sequence analysis of defective and infectious avian hepadnaviruses. Nucleic Acids Res., 19:4289.

Triyatni, M., P. L. Ey, T. Tran, M. L. Mire, M. Qiao, C. J. Burrell and A. R. Jilbert, 2001. Sequence comparison of an Australian duck hepatitis B virus strain with other avian hepadnaviruses. J. Gen. Virol., 82:373-378.

Tuttleman, J. S., J. C. Pugh and J. W. Summers, 1986. *In vitro* experimental infection of primary duck hepatocyte cultures with duck hepatitis B virus. J. Virol., 58:17-25.

Uchida, M., M. Esuni and T. Shikata, 1989. Molecular cloning and sequence analysis of duck hepatitis B virus genome of a new variant isolated from Shanghai ducks. Virology, 173:600-606.